Mathematical Method
MTH303

Virtual University

Virtual University of Pakistan
Knowledge beyond the boundaries

Table of Contents

Lecture 1 Introduction

Background
Linear *y=mx+c*
Quadratic *ax2+bx+c=0*
Cubic *ax3+bx2+cx+d=0*
Systems of Linear equations

ax+by+c=0
lx+my+n=0

Solution ?
Equation

Differential Operator

$$\frac{dy}{dx}=\frac{1}{x}$$

Taking anti derivative on both sides
y=ln x

From the past

- *Algebra*
- *Trigonometry*
- *Calculus*
- *Differentiation*
- *Integration*

- *Differentiation*
  - *Algebraic Functions*
  - *Trigonometric Functions*
  - *Logarithmic Functions*
  - *Exponential Functions*
  - *Inverse Trigonometric Functions*

- More Differentiation
  - Successive Differentiation
  - Higher Order
  - Leibnitz Theorem
- Applications
  - Maxima and Minima
  - Tangent and Normal
- Partial Derivatives

y=f(x)

f(x,y)=0

z=f(x,y)

Integration

- Reverse of Differentiation
- By parts
- By substitution
- By Partial Fractions
- Reduction Formula

Frequently required

- Standard Differentiation formulae
- Standard Integration Formulae

Differential Equations

- Something New
- Mostly old stuff
  - Presented differently
  - Analyzed differently
  - Applied Differently

$$\frac{dy}{dx}-5y = 1$$

$$(y-x)dx+4xdy = 0$$

$$\frac{d^2y}{dx^2}+5\left(\frac{dy}{dx}\right)^3-4y = e^x$$

$$\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x} = 0$$

$$x\frac{\partial u}{\partial x}+y\frac{\partial v}{\partial y} = u$$

$$\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial t^2}+2\frac{\partial u}{\partial t} = 0$$

Lecture 2 Fundamentals of Differential Equation

Fundamentals

- Definition of a differential equation.
- Classification of differential equations.
- Solution of a differential equation.
- Initial value problems associated to DE.
- Existence and uniqueness of solutions

Elements of the Theory

- Applicable to:
  - Chemistry
  - Physics
  - Engineering
  - Medicine
  - Biology
  - Anthropology
- Differential Equation – involves an unknown function with one or more of its derivatives
- Ordinary D.E. – a function where the unknown is dependent upon only one independent variable

Examples of DEs

$$\frac{dy}{dx}-5y=1$$

$$(y-x)dx+4xdy=0$$

$$\frac{d^2y}{dx^2}+5\left(\frac{dy}{dx}\right)^3-4y=e^x$$

$$\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}=0$$

$$x\frac{\partial u}{\partial x}+y\frac{\partial v}{\partial y}=u$$

$$\frac{\partial^2u}{\partial x^2}-\frac{\partial^2u}{\partial t^2}+2\frac{\partial u}{\partial t}=0$$

Specific Examples of ODE's

$$\frac{du}{dt} = F(t)G(u), \qquad \text{the growth equation;}$$

$$\frac{d^2\theta}{dt^2} + \frac{g}{l}\sin(\theta) = F(t), \qquad \text{the pendulum equation;}$$

$$\frac{d^2y}{dt^2} + \varepsilon(y^2+1)\frac{dy}{dt} + y = 0, \qquad \text{the van der Pol equation;}$$

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{Q}{C} = E(t), \qquad \text{the LCR oscillator equation;}$$

$$\frac{dp}{dt} = -2a(t)p + \frac{b(t)^2}{u(t)}p^2 - v(t), \qquad \text{a Riccati equation.}$$

- The order of an equation:
  - The order of the highest derivative appearing in the equation

$$\frac{d^2y}{dx^2} + 5\left(\frac{dy}{dx}\right)^3 - 4y = e^x$$

$$a^2\frac{\partial^4 y}{\partial x^4} + \frac{\partial^2 u}{\partial x^2} = 0$$

Ordinary Differential Equation

If an equation contains only ordinary derivatives of one or more dependent variables, *w.r.t* a single variable, then it is said to be an Ordinary Differential Equation **(ODE)**. For example the differential equation

$$\frac{d^2y}{dx^2} + 5\left(\frac{dy}{dx}\right)^3 - 4y = e^x$$

is an ordinary differential equation.

Partial Differential Equation

Similarly an equation that involves partial derivatives of one or more dependent variables *w.r.t* two or more independent variables is called a Partial Differential Equation **(PDE)**. For example the equation

$$a^2 \frac{\partial^4 u}{\partial x^4} + \frac{\partial^2 u}{\partial x^2} = 0$$

is a partial differential equation.

Results from ODE data

- The solution of a general differential equation:
  - $f(t, y, y', \ldots, y(n)) = 0$
  - is defined over some interval *I* having the following properties:
    - $y(t)$ and its first n derivatives exist for all t in I so that $y(t)$ and its first n - 1 derivates must be continuous in I
    - $y(t)$ satisfies the differential equation for all t in I

- General Solution – all solutions to the differential equation can be represented in this form for all constants
- Particular Solution – contains no arbitrary constants
- Initial Condition
- Boundary Condition
- Initial Value Problem (IVP)
- Boundary Value Problem (BVP)

IVP Examples

- The Logistic Equation
  - $p' = ap – bp2$
  - with initial condition $p(t0) = p0$; for $p0 = 10$ the solution is:
  - $p(t) = 10a / (10b + (a – 10b)e\text{-}a(t\text{-}t0))$
- The mass-spring system equation
  - $x'' + (a / m)\, x' + (k / m)x = g + (F(t) / m)$

BVP Examples

- Differential equations
  - $y'' + 9y = \sin(t)$
    - with initial conditions $y(0) = 1$, $y'(2p) = -1$
    - $y(t) = (1/8) \sin(t) + \cos(3t) + \sin (3t)$
  - $y'' + p2y = 0$
    - with initial conditions $y(0) = 2$, $y(1) = -2$
    - $y(t) = 2\cos(pt) + (c)\sin(pt)$

Properties of ODE's

- Linear – if the nth-order differential equation can be written:

- $a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1 y' + a_0(t)y = h(t)$

- Nonlinear – not linear

  $x^3(y''')^3 - x^2 y(y'')^2 + 3xy' + 5y = e^x$

Superposition

- Superposition – allows us to decompose a problem into smaller, simpler parts and then combine them to find a solution to the original problem.

Explicit Solution

A solution of a differential equation

$$F\left(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \cdots, \frac{d^2y}{dx^2}\right) = 0$$

that can be written as $y = f(x)$ is known as an explicit solution .

Example: The solution $y = xe^x$ is an explicit solution of the differential equation

$$\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = 0$$

Implicit Solution

A relation ***G(x,y)*** is known as an implicit solution of a differential equation, if it defines one or more explicit solution on *I*.

Example: The solution $x^2 + y^2 - 4 = 0$ is an implicit solution of the equation $y' = -x/y$ as it defines two explicit solutions $y = \pm(4-x^2)^{1/2}$

Lecture 3 Separable Equations

The differential equation of the form

$$\frac{dy}{dx} = f(x, y)$$

is called **separable** if it can be written in the form

$$\frac{dy}{dx} = h(x)g(y)$$

To solve a separable equation, we perform the following steps:

1. We solve the equation $g(y) = 0$ to find the constant solutions of the equation.

2. For non-constant solutions we write the equation in the form.

$$\frac{dy}{g(y)} = h(x)dx$$

Then integrate $$\int \frac{1}{g(y)} dy = \int h(x)dx$$

to obtain a solution of the form

$$G(y) = H(x) + C$$

3. We list the entire constant and the non-constant solutions to avoid repetition..

4. If you are given an IVP, use the initial condition to find the particular solution.

Note that:
(a) No need to use two constants of integration because $C_1 - C_2 = C$.
(b) The constants of integration may be relabeled in a convenient way.
(c) Since a particular solution may coincide with a constant solution, **step 3 is important.**

Example 1:
Find the particular solution of

$$\frac{dy}{dx} = \frac{y^2 - 1}{x}, \quad y(1) = 2$$

Solution:
1. By solving the equation

$$y^2 - 1 = 0$$

We obtain the constant solutions

$$y = \pm 1$$

2. Rewrite the equation as

$$\frac{dy}{y^2 - 1} = \frac{dx}{x}$$

Resolving into partial fractions and integrating, we obtain

$$\frac{1}{2}\int\left[\frac{1}{y-1} - \frac{1}{y+1}\right]dy = \int \frac{1}{x}dx$$

Integration of rational functions, we get

$$\frac{1}{2}\ln\frac{|y-1|}{|y+1|} = \ln|x| + C$$

3. The solutions to the given differential equation are

$$\begin{cases} \frac{1}{2}\ln\frac{|y-1|}{|y+1|} = \ln|x| + C \\ y = \pm 1 \end{cases}$$

4. Since the constant solutions do not satisfy the initial condition, we plug in the condition

$y = 2$ When $x = 1$ in the solution found in step 2 to find the value of C.

$$\frac{1}{2}\ln\left(\frac{1}{3}\right) = C$$

The above implicit solution can be rewritten in an explicit form as**:**

$$y = \frac{3 + x^2}{3 - x^2}$$

Example 2:
Solve the differential equation

$$\frac{dy}{dt} = 1 + \frac{1}{y^2}$$

Solution:

1. We find roots of the equation to find constant solutions

$$1 + \frac{1}{y^2} = 0$$

No constant solutions exist because the equation has no real roots.

2. For non-constant solutions, we separate the variables and integrate

$$\int \frac{dy}{1 + 1/y^2} = \int dt$$

Since $\displaystyle \frac{1}{1 + 1/y^2} = \frac{y^2}{y^2 + 1} = 1 - \frac{1}{y^2 + 1}$

Thus $$\int \frac{dy}{1+1/y^2} = y - \tan^{-1}(y)$$

So that $$y - \tan^{-1}(y) = t + C$$

It is **not easy** to find the **solution in an explicit form** i.e. y as a function of *t*.

3. Since $\exists$ no constant solutions, all solutions are given by the implicit equation found
in step 2.

Example 3:
Solve the initial value problem

$$\frac{dy}{dt} = 1 + t^2 + y^2 + t^2 y^2, \quad y(0) = 1$$

Solution:

1. Since $$1 + t^2 + y^2 + t^2 y^2 = (1+t^2)(1+y^2)$$
The equation is separable & has no constant solutions because $\exists$ no real roots of
$$1 + y^2 = 0.$$

2. For non-constant solutions we separate the variables and integrate.

$$\frac{dy}{1+y^2} = (1+t^2)dt$$

$$\int \frac{dy}{1+y^2} = \int (1+t^2)dt$$

$$\tan^{-1}(y) = t + \frac{t^3}{3} + C$$

Which can be written as

$$y = \tan\left(t + \frac{t^3}{3} + C\right)$$

3. Since $\exists$ no constant solutions, all solutions are given by the implicit or explicit equation.

4. The initial condition $y(0) = 1$ gives

$$C = \tan^{-1}(1) = \frac{\pi}{4}$$

The particular solution to the initial value problem is

$$\tan^{-1}(y) = t + \frac{t^3}{3} + \frac{\pi}{4}$$

or in the explicit form $$y = \tan\left(t + \frac{t^3}{3} + \frac{\pi}{4}\right)$$

Example 4:

Solve

$$(1+x)dy - ydx = 0$$

Solution:

Dividing with $(1+x)y$, we can write the given equation as

$$\frac{dy}{dx} = \frac{y}{(1+x)}$$

1. The only constant solution is $y = 0$

2. For non-constant solution we separate the variables

$$\frac{dy}{y} = \frac{dx}{1+x}$$

Integrating both sides, we have

$$\int \frac{dy}{y} = \int \frac{dx}{1+x}$$

$$\ln|y| = \ln|1+x| + c_1$$

$$y = e^{\ln|1+x|+c_1} = e^{\ln|1+x|}.e^{c_1}$$

or $$y = |1+x|e^{c_1} = \pm e^{c_1}(1+x)$$

$$y = C(1+x), \quad C = \pm e^{c_1}$$

If we use $\ln|c|$ instead of c_1 then the solution can be written as

$$\ln|y| = \ln|1+x| + \ln|c|$$

or $$\ln|y| = \ln|c(1+x)|$$

So that $$y = c(1+x).$$

3. The solutions to the given equation are

$$y = c(1+x)$$

$$y = 0$$

Example 5

Solve

$$xy^4 dx + (y^2 + 2)e^{-3x} dy = 0.$$

Solution:

The differential equation can be written as

$$\frac{dy}{dx} = \left(-xe^{3x}\right)\left(\frac{y^4}{y^2+2}\right)$$

1. Since $\frac{y^4}{y^2+2} \Rightarrow y = 0$. Therefore, the only constant solution is $y = 0$.

2. We separate the variables

$$xe^{3x}dx + \frac{y^2+2}{y^4}dy = 0 \quad \text{or} \quad xe^{3x}dx + \left(y^{-2} + 2y^{-4}\right)dy = 0$$

Integrating, with use integration by parts by parts on the first term, yields

$$\frac{1}{3}xe^{3x} - \frac{1}{9}e^{3x} - y^{-1} - \frac{2}{3}y^{-3} = c_1$$

$$e^{3x}(3x-1) = \frac{9}{y} + \frac{6}{y^3} + c \quad \text{where} \quad 9c_1 = c$$

3. All the solutions are

$$e^{3x}(3x-1) = \frac{9}{y} + \frac{6}{y^3} + c$$

$$y = 0$$

Example 6:

Solve the initial value problems

(a) $\frac{dy}{dx} = (y-1)^2, \quad y(0) = 1$ (b) $\frac{dy}{dx} = (y-1)^2, \quad y(0) = 1.01$

and compare the solutions.

Solutions**:**

1. Since $(y-1)^2 = 0 \Rightarrow y = 1$. Therefore, the only constant solution is $y = 0$.

2. We separate the variables

$$\frac{dy}{(y-1)^2} = dx \quad \text{or} \quad (y\text{-}1)^{-2}dy = dx$$

Integrating both sides we have

$$\int (y-1)^{-2}\,dy = \int dx$$

$$\frac{(y-1)^{-2+1}}{-2+1} = x + c$$

or $$-\frac{1}{y-1} = x + c$$

3. All the solutions of the equation are

$$-\frac{1}{y-1} = x + c$$
$$y = 1$$

4. We plug in the conditions to find particular solutions of both the problems

(a) $y(0)=1 \Rightarrow y=1$ when $x=0$. So we have

$$-\frac{1}{1-1} = 0 + c \Rightarrow c = -\frac{1}{0} \Rightarrow c = -\infty$$

The particular solution is

$$-\frac{1}{y-1} = -\infty \Rightarrow y-1=0$$

So that the solution is $y=1$, which is same as constant solution.

(b) $y(0)=1.01 \Rightarrow y=1.01$ when $x=0$. So we have

$$-\frac{1}{1.01-1} = 0 + c \Rightarrow c = -100$$

So that solution of the problem is

$$-\frac{1}{y-1} = x - 100 \Rightarrow y = 1 + \frac{1}{100-x}$$

5. Comparison: A radical change in the solutions of the differential equation has Occurred corresponding to a very small change in the condition**!!**

Example 7:

Solve the initial value problems

(a) $\frac{dy}{dx} = (y-1)^2 + 0.01, \quad y(0)=1$ (b) $\frac{dy}{dx} = (y-1)^2 - 0.01, \quad y(0)=1.$

Solution:

(a) First consider the problem

$$\frac{dy}{dx} = (y-1)^2 + 0.01, \quad y(0)=1$$

We separate the variables to find the non-constant solutions

$$\frac{dy}{\left(\sqrt{0.01}\right)^2+\left(y-1\right)^2}=dx$$

Integrate both sides

$$\int\frac{d\left(y-1\right)}{\left(\sqrt{0.01}\right)^2+\left(y-1\right)^2}=\int dx$$

So that $$\frac{1}{\sqrt{0.01}}\tan^{-1}\frac{y-1}{\sqrt{0.01}}=x+c$$

$$\tan^{-1}\left(\frac{y-1}{\sqrt{0.01}}\right)=\sqrt{0.01}\left(x+c\right)$$

$$\frac{y-1}{\sqrt{0.01}}=\tan\left[\sqrt{0.01}\left(x+c\right)\right]$$

or $$y=1+\sqrt{0.01}\tan\left[\sqrt{0.01}\left(x+c\right)\right]$$

Applying $y(0)=1\Rightarrow y=1$ when $x=0$, we have

$$\tan^{-1}(0)=\sqrt{0.01}\left(0+c\right)\Rightarrow 0=c$$

Thus the solution of the problem is

$$y=1+\sqrt{0.01}\tan\left(\sqrt{0.01}\,x\right)$$

(b) Now consider the problem

$$\frac{dy}{dx}=\left(y-1\right)^2-0.01,\qquad y(0)=1.$$

We separate the variables to find the non-constant solutions

$$\frac{d\ y}{\left(y-1\right)^2-\left(\sqrt{0.01}\right)^2}=dx$$

$$\int\frac{d\left(y-1\right)}{\left(y-1\right)^2-\left(\sqrt{0.01}\right)^2}=\int dx$$

$$\frac{1}{2\sqrt{0.01}}\ln\left|\frac{y-1-\sqrt{0.01}}{y-1+\sqrt{0.01}}\right|=x+c$$

Applying the condition $y(0)=1\Rightarrow y=1$ when $x=0$

$$\frac{1}{2\sqrt{0.01}}\ln\left|\frac{-\sqrt{0.01}}{\sqrt{0.01}}\right| = c \Rightarrow c = 0$$

$$\ln\left|\frac{y-1-\sqrt{0.01}}{y-1+\sqrt{0.01}}\right| = 2\sqrt{0.01}\,x$$

$$\frac{y-1-\sqrt{0.01}}{y-1+\sqrt{0.01}} = \frac{e^{2\sqrt{0.01}\,x}}{1}$$

Simplification:

By using the property $\frac{a}{b} = \frac{c}{d} \Rightarrow \frac{a+b}{a-b} = \frac{c+d}{c-d}$

$$\frac{y-1-\sqrt{0.01}+y-1+\sqrt{0.01}}{y-1-\sqrt{0.01}-y+1-\sqrt{0.01}} = \frac{e^{2\sqrt{0.01}x}+1}{e^{2\sqrt{0.01}x}-1}$$

$$\frac{2y-2}{-2\sqrt{0.01}} = \frac{e^{2\sqrt{0.01}}+1}{e^{2\sqrt{0.01}}-1}$$

$$\frac{y-1}{-\sqrt{0.01}} = \frac{e^{2\sqrt{0.01}}+1}{e^{2\sqrt{0.01}}-1}$$

$$y-1 = -\sqrt{0.01}\left(\frac{e^{2\sqrt{0.01}}+1}{e^{2\sqrt{0.01}}-1}\right)$$

$$y = 1-\sqrt{0.01}\left(\frac{e^{2\sqrt{0.01}}+1}{e^{2\sqrt{0.01}}-1}\right)$$

Comparison:

The solutions of both the problems are

(a) $y = \quad 1+\sqrt{0.01}\ \tan\left(\sqrt{0.01}\,x\right)$

(b) $y = 1-\sqrt{0.01}\left(\frac{e^{2\sqrt{0.01}}+1}{e^{2\sqrt{0.01}}-1}\right)$

Again a radical change has occurred corresponding to a very small in the differential equation**!**

Exercise:

Solve the given differential equation by separation of variables.

1. $\frac{dy}{dx} = \left(\frac{2y+3}{4x+5}\right)^2$

2. $\sec^2 x dy + \csc y dx = 0$

3. $e^y \sin 2x dx + \cos x\left(e^{2y} - y\right)dy = 0$

4. $\frac{dy}{dx} = \frac{xy+3x-y-3}{xy-2x+4y-8}$

5. $\frac{dy}{dx} = \frac{xy+2y-x-2}{xy-3y+x-3}$

6. $y\left(4-x^2\right)^{\frac{1}{2}} dy = \left(4+y^2\right)^{\frac{1}{2}} dx$

7. $\left(x+\sqrt{x}\right)\frac{dy}{dx} = y+\sqrt{y}$

Solve the given differential equation subject to the indicated initial condition.

8. $\left(e^{-y}+1\right)\sin x dx = \left(1+\cos x\right)dy$, $y(0)=0$

9. $\left(1+x^4\right)dy + x\left(1+4y^2\right)dx = 0$, $y(1)=0$

10. $ydy = 4x\left(y^2+1\right)^{\frac{1}{2}} dx$, $y(0)=1$

Lecture 4 Homogeneous Differential Equations

A differential equation of the form

$$\frac{dy}{dx} = f(x, y)$$

Is said to be *homogeneous* if the function $f(x, y)$ is homogeneous, which means

$f(tx, ty) = t^n f(x, y)$ For some real number n, for any number t.

Example 1

Determine whether the following functions are homogeneous

$$\begin{cases} f(x, y) = \dfrac{xy}{x^2 + y^2} \\ g(x, y) = \ln\left(-3x^2 y/(x^3 + 4xy^2)\right) \end{cases}$$

Solution:

The functions $f(x, y)$ is homogeneous because

$$f(tx, ty) = \frac{t^2 xy}{t^2(x^2 + y^2)} = \frac{xy}{x^2 + y^2} = f(x, y)$$

Similarly, for the function $g(x, y)$ we see that

$$g(tx, ty) = \ln\left(\frac{-3t^3 x^2 y}{t^3(x^3 + 4xy^2)}\right) = \ln\left(\frac{-3x^2 y}{x^3 + 4xy^2}\right) = g(x, y)$$

Therefore, the second function is also homogeneous.

Hence the differential equations

$$\begin{cases} \dfrac{dy}{dx} = f(x, y) \\ \dfrac{dy}{dx} = g(x, y) \end{cases}$$

Are homogeneous differential equations

Method of Solution:

To solve the homogeneous differential equation

$$\frac{dy}{dx} = f(x, y)$$

We use the substitution

$$v = \frac{y}{x}$$

If $f(x, y)$ is homogeneous of degree zero, then we have

$$f(x, y) = f(1, v) = F(v)$$

Since $y' = xv' + v$, the differential equation becomes

$$x\frac{dv}{dx} + v = f(1, v)$$

This is a separable equation. We solve and go back to old variable y through $y = xv$.

Summary:

1. Identify the equation as homogeneous by checking $f(tx, ty) = t^n f(x, y)$;

2. Write out the substitution $v = \frac{y}{x}$;

3. Through easy differentiation, find the new equation satisfied by the new function v;

$$x\frac{dv}{dx} + v = f(1, v)$$

4. Solve the new equation (which is always separable) to find v;

5. Go back to the old function y through the substitution $y = vx$;

6. If we have an IVP, we need to use the initial condition to find the constant of integration.

Caution:

- Since we have to solve a separable equation, we must be careful about the constant solutions.
- If the substitution $y = vx$ does not reduce the equation to separable form then the equation is not homogeneous or something is wrong along the way.

Illustration:

Example 2 Solve the differential equation

$$\frac{dy}{dx} = \frac{-2x + 5y}{2x + y}$$

Solution:

Step 1. It is easy to check that the function

$$f(x, y) = \frac{-2x + 5y}{2x + y}$$

is a homogeneous function.

Step 2. To solve the differential equation we substitute

$$v = \frac{y}{x}$$

Step 3. Differentiating w.r.t x, we obtain

$$xv' + v = \frac{-2x + 5xv}{2x + xv} = \frac{-2 + 5v}{2 + v}$$

which gives

$$\frac{dv}{dx} = \frac{1}{x}\left(\frac{-2 + 5v}{2 + v} - v\right)$$

This is a separable. At this stage please refer to the **Caution!**

Step 4. Solving by separation of variables all solutions are implicitly given by

$$-4\ln(|v-2|) + 3\ln|v-1| = \ln(|x|) + C$$

Step 5. Going back to the function y through the substitution $y = vx$, we get

$$-4\ln|y-2x| + 3\ln|y-x| = C$$

$$-4\ln\left|\frac{y-2x}{x}\right| + 3\ln\left|\frac{y-x}{x}\right| = \ln|x| + c$$

$$\ln\left|\frac{y-2x}{x}\right|^{-4} + \ln\left|\frac{y-x}{x}\right|^{3} = \ln x + \ln c_1,\ c = \ln c_1$$

$$\ln\left|\frac{(y-2x)^{-4}}{x^{-4}}\right| + \ln\left|\frac{(y-x)^{3}}{x^{3}}\right| = \ln c_1 x$$

$$\ln\left|\frac{(y-2x)^{-4}}{x^{-4}}.\frac{(y-x)^{3}}{x^{3}}\right| = \ln c_1 x$$

$$\frac{(y-2x)^{-4}}{x^{-4}}.\frac{(y-x)^{3}}{x^{3}} = c_1 x$$

$$x(y-2x)^{-4}(y-x)^{3} = c_1 x$$

$$(y-2x)^{-4}(y-x)^{3} = c_1$$

Note that the implicit equation can be rewritten as

$$(y-x)^3 = C_1(y-2x)^4$$

Equations reducible to homogenous form

The differential equation

$$\frac{dy}{dx} = \frac{a_1 x + b_1 y + c_1}{a_2 x + b_2 y + c_2}$$

is not homogenous. However, it can be reduced to a homogenous form as detailed below

Case 1: $\dfrac{a_1}{a_2} = \dfrac{b_1}{b_2}$

We use the substitution $z = a_1 x + b_1 y$ which reduces the equation to a separable equation in the variables x and z. Solving the resulting separable equation and replacing z with $a_1 x + b_1 y$, we obtain the solution of the given differential equation.

Case 2: $\dfrac{a_1}{a_2} \neq \dfrac{b_1}{b_2}$

In this case we substitute

$$x = X + h, \qquad y = Y + k$$

Where h and k are constants to be determined. Then the equation becomes

$$\frac{dY}{dX} = \frac{a_1 X + b_1 Y + a_1 h + b_1 k + c_1}{a_2 X + b_2 Y + a_2 h + b_2 k + c_2}$$

We choose h and k such that

$$\left.\begin{aligned} a_1 h + b_1 k + c_1 &= 0 \\ a_2 h + b_2 k + c_2 &= 0 \end{aligned}\right\}$$

This reduces the equation to

$$\frac{dY}{dX} = \frac{a_1 X + b_1 Y}{a_2 X + b_2 Y}$$

Which is homogenous differential equation in X and Y, and can be solved accordingly. After having solved the last equation we come back to the old variables x and y.

Example 3

Solve the differential equation

$$\frac{dy}{dx} = -\frac{2x+3y-1}{2x+3y+2}$$

Solution:

Since $\frac{a_1}{a_2} = 1 = \frac{b_1}{b_2}$, we substitute $z = 2x + 3y$, so that

$$\frac{dy}{dx} = \frac{1}{3}\left(\frac{dz}{dx} - 2\right)$$

Thus the equation becomes

$$\frac{1}{3}\left(\frac{dz}{dx} - 2\right) = -\frac{z-1}{z+2}$$

i.e. $$\frac{dz}{dx} = \frac{-z+7}{z+2}$$

This is a variable separable form, and can be written as

$$\left(\frac{z+2}{-z+7}\right)dz = dx$$

Integrating both sides we get

$$-z - 9\ln(z-7) = x + A$$

Simplifying and replacing z with $2x + 3y$, we obtain

$$-\ln(2x+3y-7)^9 = 3x + 3y + A$$

or $$(2x+3y-7)^{-9} = ce^{3(x+y)}, \qquad c = e^A$$

Example 4

Solve the differential equation

$$\frac{dy}{dx}=\frac{(x+2y-4)}{2x+y-5}$$

Solution:

By substitution

$$x=X+h, \qquad y=Y+k$$

The given differential equation reduces to

$$\frac{dY}{dX}=\frac{(X+2Y)+(h+2k-4)}{(2X+Y)+(2h+k-5)}$$

We choose h and k such that

$$h+2k-4=0, \quad 2h+k-5=0$$

Solving these equations we have $h=2,\ k=1$. Therefore, we have

$$\frac{dY}{dX}=\frac{X+2Y}{2X+Y}$$

This is a homogenous equation. We substitute $Y=VX$ to obtain

$$X\frac{dV}{dX}=\frac{1-V^2}{2+V} \quad \text{or} \quad \left[\frac{2+V}{1-V^2}\right]dV=\frac{dX}{X}$$

Resolving into partial fractions and integrating both sides we obtain

$$\int\left[\frac{3}{2(1-V)}+\frac{1}{2(1+V)}\right]dV=\int\frac{dX}{X}$$

or

$$-\frac{3}{2}\ln(1-V)+\frac{1}{2}\ln(1+V)=\ln X+\ln A$$

Simplifying and removing ($\ln$) from both sides, we get

$$(1-V)^3/(1+V)=CX^{-2},\ C=A^{-2}$$

$$-\frac{3}{2}\ln(1-V)+\frac{1}{2}\ln(1+V)=\ln X+\ln A$$
$$\ln(1-V)^{-3/2}+\ln(1+V)^{1/2}=\ln XA$$
$$\ln(1-V)^{-3/2}(1+V)^{1/2}=\ln XA$$
$$(1-V)^{-3/2}(1+V)^{1/2}=XA$$
taking power "−2" on both sides
$$(1-V)^{3}(1+V)^{-1}=X^{-2}A^{-2}$$
$$put\ V=\frac{Y}{X}$$
$$(1-\frac{Y}{X})^{3}\left(1+\frac{Y}{X}\right)^{-1}=X^{-2}A^{-2}$$
$$\left(\frac{X-Y}{X}\right)^{3}\left(\frac{X+Y}{X}\right)^{-1}=X^{-2}A^{-2}$$
$$\frac{(X-Y)^{3}}{X+Y}X^{-3+1}=X^{-2}A^{-2}$$
$$say, c=A^{-2}$$
$$\frac{(X-Y)^{3}}{X+Y}=c$$
$$put\ X=x-2,\ Y=y-1$$
$$(x+y-1)^{3}/x+y-3=c$$

Now substituting $V=\frac{Y}{X}, X=x-2, Y=y-1$ and simplifying, we obtain

$$(x-y-1)^{3}/(x+y-3)=C$$

This is solution of the given differential equation, an implicit one.

Exercise

Solve the following Differential Equations

1. $(x^4+y^4)dx-2x^3ydy=0$
2. $\frac{dy}{dx}=\frac{y}{x}+\frac{x^2}{y^2}+1$
3. $\left(x^2e^{\frac{-y}{x}}+y^2\right)dx=xydy$

4. $ydx + \left(y\cos\frac{x}{y} - x \right)dy = 0$

5. $\left(x^3 + y^2\sqrt{x^2 + y^2}\right)dx - xy\sqrt{x^2 + y^2}\,dy = 0$

Solve the initial value problems

6. $\left(3x^2 + 9xy + 5y^2\right)dx - \left(6x^2 + 4xy\right)dy = 0, \qquad y(2) = -6$

7. $\left(x + \sqrt{y^2 - xy}\right)\frac{dy}{dx} = y, \qquad y\left(\frac{1}{2}\right) = 1$

8. $\left(x + ye^{y/x}\right)dx - xe^{y/x}dy = 0, \qquad y(1) = 0$

9. $\frac{dy}{dx} - \frac{y}{x} = \cosh\frac{y}{x}, \qquad y(1) = 0$

Lecture 5 Exact Differential Equations

Let us first rewrite the given differential equation

$$\frac{dy}{dx} = f(x,y)$$

into the alternative form

$$M(x,y)dx + N(x,y)dy = 0 \quad \text{where} \quad f(x,y) = -\frac{M(x,y)}{N(x,y)}$$

This equation is an exact differential equation if the following condition is satisfied

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

This condition of exactness insures the existence of a function $F(x,y)$ such that

$$\frac{\partial F}{\partial x} = M(x,y), \quad \frac{\partial F}{\partial y} = N(x,y)$$

Method of Solution:

If the given equation is exact then the solution procedure consists of the following steps:

Step 1. Check that the equation is exact by verifying the condition $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$

Step 2. Write down the system $\frac{\partial F}{\partial x} = M(x,y), \ \frac{\partial F}{\partial y} = N(x,y)$

Step 3. Integrate either the 1st equation *w. r. to* ***x*** or 2nd *w. r. to* ***y***. If we choose the 1st equation then

$$F(x,y) = \int M(x,y)dx + \theta(y)$$

The function $\theta(y)$ is an arbitrary function of y, integration w.r.to x; y being constant.

Step 4. Use second equation in step 2 and the equation in step 3 to find $\theta'(y)$.

$$\frac{\partial F}{\partial y} = \frac{\partial}{\partial y}\left(\int M(x,y)dx\right) + \theta'(y) = N(x,y)$$

$$\theta'(y) = N(x,y) - \frac{\partial}{\partial y}\int M(x,y)dx$$

Step 5. Integrate to find $\theta(y)$ and write down the function $F(x,y)$;

Step 6. All the solutions are given by the implicit equation

$$F(x,y) = C$$

Step 7. If you are given an IVP, plug in the initial condition to find the constant C.

Caution: x should disappear from $\theta'(y)$. Otherwise something is **wrong!**

Example 1

Solve $(3x^2y+2)dx+(x^3+y)dy=0$

Solution: Here $M=3x^2y+2$ and $N=x^3+y$

$$\frac{\partial M}{\partial y}=3x^2, \frac{\partial N}{\partial x}=3x^2$$

i.e. $$\frac{\partial M}{\partial y}=\frac{\partial N}{\partial x}$$

Hence the equation is exact. The LHS of the equation must be an exact differential i.e. $\exists$ a function $f(x,y)$ such that

$$\frac{\partial f}{\partial x}=3x^2y+2=M$$

$$\frac{\partial f}{\partial y}=x^3+y=N$$

Integrating 1st of these equations w. r. t. *x*, have

$$f(x,y)=x^3y+2x+h(y),$$

where $h(y)$ is the constant of integration. Differentiating the above equation w. r. t. *y* and using 2nd, we obtain

$$\frac{\partial f}{\partial y}=x^3+h'(y)=x^3+y=N$$

Comparing $h'(y)=y$ is independent of *x*.

or.

Integrating, we have

$$h(y)=\frac{y^2}{2}$$

Thus $$f(x,y)=x^3y+2x+\frac{y^2}{2}$$

Hence the general solution of the given equation is given by

$$f(x,y)=c$$

i.e. $x^3 y + 2x + \frac{y^2}{2} = c$

Note that we could start with the 2nd equation

$$\frac{\partial f}{\partial y} = x^3 + y = N$$

to reach on the above solution of the given equation**!**

Example 2

Solve the initial value problem

$$\left(2y\sin x\cos x + y^2 \sin x\right)dx + \left(\sin^2 x - 2y\cos x\right)dy = 0.$$

$$y(0) = 3.$$

Solution: Here

$$M = 2y\sin x\cos x + y^2 \sin x$$

and $N = \sin^2 x - 2y\cos x$

$$\frac{\partial M}{\partial y} = 2\sin x\cos x + 2y\sin x,$$

$$\frac{\partial N}{\partial x} = 2\sin x\cos x + 2y\sin x,$$

This implies $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$

Thus given equation is exact.

Hence there exists a function $f(x, y)$ such that

$$\frac{\partial f}{\partial x} = 2y\sin x\cos x + y^2 \sin x = M$$

$$\frac{\partial f}{\partial y} = \sin^2 x - 2y\cos x = N$$

Integrating 1st of these w. r. t. x, we have

$$f(x, y) = y\sin^2 x - y^2 \cos x + h(y),$$

Differentiating this equation w. r. t. y substituting in $\frac{\partial f}{\partial y} = N$

$$\sin^2 x - 2y\cos x + h'(y) = \sin^2 x - 2y\cos x$$

$$h'(y) = 0 \quad \text{or} \quad h(y) = c_1$$

Hence the general solution of the given equation is

$$f(x,y) = c_2$$

i.e. $$y\sin^2 x - y^2 \cos x = C, \text{ where } C = c_1 - c_2$$

Applying the initial condition that when $x = 0, y = 3,$ we have

$$-9 = c$$

since $$y^2 \cos x - y\sin^2 x = 9$$

is the required solution.

Example 3:

Solve the DE $$\left(e^{2y} - y\cos xy\right)dx + \left(2xe^{2y} - x\cos x\,y + 2y\right)dy = 0$$

Solution:

The equation is neither separable nor homogenous.

Since, $$\left.\begin{aligned} M(x,y) &= e^{2y} - y\cos xy \\ N(x,y) &= 2xe^{2y} - x\cos xy + 2y \end{aligned}\right\}$$

and

$$\frac{\partial M}{\partial y} = 2e^{2y} + xy\sin xy - \cos xy = \frac{\partial N}{\partial x}$$

Hence the given equation is exact and a function $f(x,y)$ exist for which

$$M(x,y) = \frac{\partial f}{\partial x} \quad \text{and} \quad N(x,y) = \frac{\partial f}{\partial y}$$

which means that

$$\frac{\partial f}{\partial x} = e^{2y} - y\cos xy \quad \text{and} \quad \frac{\partial f}{\partial y} = 2xe^{2y} - x\cos xy + 2y$$

Let us start with the second equation i.e.

$$\frac{\partial f}{\partial y} = 2xe^{2y} - x\cos xy + 2y$$

Integrating both sides w.r.to y, we obtain

$$f(x,y) = 2x\int e^{2y}dy - x\int \cos xy\,dy + 2\int y\,dy$$

Note that while integrating w.r.to y, x is treated as constant. Therefore

$$f(x,y) = xe^{2y} - \sin xy + y^2 + h(x)$$

h is an arbitrary function of x. From this equation we obtain $\frac{\partial f}{\partial x}$ and equate it to M

$$\frac{\partial f}{\partial x} = e^{2y} - y\cos xy + h'(x) = e^{2y} - y\cos xy$$

So that
$$h'(x) = 0 \Rightarrow h(x) = C$$

Hence a one-parameter family of solution is given by

$$xe^{2y} - \sin xy + y^2 + c = 0$$

Example 4

Solve
$$2xy\,dx + \left(x^2 - 1\right)dy = 0$$

Solution:

Clearly $M(x,y) = 2xy$ and $N(x,y) = x^2 - 1$

Therefore
$$\frac{\partial M}{\partial y} = 2x = \frac{\partial N}{\partial x}$$

The equation is exact and $\exists$ a function $f(x,y)$ such that

$$\frac{\partial f}{\partial x} = 2xy \text{ and } \frac{\partial f}{\partial y} = x^2 - 1$$

We integrate first of these equations to obtain.

$$f(x,y) = x^2 y + g(y)$$

Here $g(y)$ is an arbitrary function y. We find $\frac{\partial f}{\partial y}$ and equate it to $N(x,y)$

$$\frac{\partial f}{\partial y} = x^2 + g'(y) = x^2 - 1$$

$$g'(y) = -1 \Rightarrow g(y) = -y$$

Constant of integration need not to be included as the solution is given by

$$f(x,y) = c$$

Hence a one-parameter family of solutions is given by

$$x^2 y - y = c$$

Example 5

Solve the initial value problem

$$\left(\cos x \sin x - xy^2\right)dx + y\left(1 - x^2\right)dy = 0, \quad y(0) = 2$$

Solution:

Since
$$\begin{cases} M(x,y) = \cos x.\sin x - x\,y^2 \\ N(x,y) = \quad y\left(1 - x^2\right) \end{cases}$$

and
$$\frac{\partial M}{\partial y} = -2xy = \frac{\partial N}{\partial x}$$

Therefore the equation is exact and $\exists$ a function $f(x,y)$ such that

$$\frac{\partial f}{\partial x} = \cos x\,.\,sin\,x - x\,y^2 \quad \text{and} \quad \frac{\partial f}{\partial y} = y(1 - x^2)$$

Now integrating 2nd of these equations w.r.t. 'y' keeping 'x'constant, we obtain

$$f(x,y) = \frac{y^2}{2}\left(1 - x^2\right) + h(x)$$

Differentiate w.r.t. 'x' and equate the result to $M(x,y)$

$$\frac{\partial f}{\partial x} = -xy^2 + h'(x) = \cos x \sin x - xy^2$$

The last equation implies that.

$$h'(x) = \cos x \sin x$$

Integrating w.r.to x, we obtain

$$h(x) = -\int (\cos x)(-\sin x)dx = -\frac{1}{2}\cos^2 x$$

Thus a one parameter family solutions of the given differential equation is

$$\frac{y^2}{2}\left(1 - x^2\right) - \frac{1}{2}\cos^2 x = c_1$$

or

$$y^2(1-x^2)-\cos^2 x = c$$

where $2c_1$ has been replaced by C. The initial condition $y=2$ when $x=0$ demand, that $4(1)-\cos^2(0)=c$ so that $c=3$. Thus the solution of the initial value problem is

$$y^2(1-x^2)-\cos^2 x = 3$$

Exercise

Determine whether the given equations is exact. If so, please solve.

1. $(\sin y - y\sin x)dx + (\cos x + x\cos y)dy = 0$

2. $\left(1 + \ln x + \frac{y}{x}\right)dx = (1 - \ln x)dy$

3. $(y\ln y - e^{-xy})dx + \left(\frac{1}{y} + \ln y\right)dy = 0$

4. $\left(2y - \frac{1}{x} + \cos 3x\right)\frac{dy}{dx} + \frac{y}{x^2} - 4x^3 + 3y\sin 3x = 0$

5. $\left(\frac{1}{x} + \frac{1}{x^2} - \frac{y}{x^2 + y^2}\right)dx + \left(ye^y + \frac{x}{x^2 + y^2}\right)dy = 0$

Solve the given differential equations subject to indicated initial conditions.

6. $(e^x + y)dx + (2 + x + ye^y)dy = 0, \qquad y(0) = 1$

7. $\left(\frac{3y^2 - x^2}{y^5}\right)\frac{dy}{dx} + \frac{x}{2y^4} = 0, \qquad y(1) = 1$

8. $\left(\frac{1}{1+y^2} + \cos x - 2xy\right)\frac{dy}{dx} = y(y + \sin x), \qquad y(0) = 1$

9. Find the value of ***k***, so that the given differential equation is exact.
   $\left(2xy^3 - y\sin xy + ky^4\right)dx - \left(20x^3 + x\sin xy\right)dy = 0$

10. $(6xy^3 + \cos y)dx - (kx^2y^2 - x\sin y)dy = 0$

Lecture 6 Integrating Factor Technique

If the equation

$$M(x,y)dx + N(x,y)dy = 0$$

is not exact, then we must have

$$\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$$

Therefore, we look for a function ***u (x, y)*** such that the equation

$$u(x,y)M(x,y)dx + u(x,y)N(x,y)dy = 0$$

becomes exact. The function ***u (x, y)*** (if it exists) is called the **integrating factor** **(IF)** and it satisfies the equation due to the condition of exactness.

$$\frac{\partial M}{\partial y}u + \frac{\partial u}{\partial y}M = \frac{\partial N}{\partial x}u + \frac{\partial u}{\partial x}N$$

This is a partial differential equation and is very difficult to solve. Consequently, the determination of the integrating factor is extremely difficult except for some special cases:

Example

Show that $1/(x^2+y^2)$ is an integrating factor for the equation $\left(x^2+y^2-x\right)dx - ydy = 0,$ and then solve the equation.

Solution: Since $M = x^2+y^2-x, \quad \mathrm{N} = -y$

Therefore $\frac{\partial M}{\partial y} = 2y, \quad \frac{\partial N}{\partial x} = 0$

So that $$\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$$

and the equation is not exact. However, if the equation is multiplied by $1/(x^2+y^2)$ then the equation becomes

$$\left(1-\frac{x}{x^2+y^2}\right)dx - \frac{y}{x^2+y^2}dy = 0$$

Now $M = 1-\frac{x}{x^2+y^2}$ and $N = -\frac{y}{x^2+y^2}$

Therefore $$\frac{\partial M}{\partial y} = \frac{2xy}{\left(x^2+y^2\right)^2} = \frac{\partial N}{\partial x}$$

So that this new equation is exact. The equation can be solved. However, it is simpler to observe that the given equation can also written

$$dx - \frac{xdx + ydy}{x^2+y^2} = 0 \quad \text{or} \quad dx - \frac{1}{2} d\left[\ln(x^2+y^2)\right] = 0$$

or $$d\left[x - \frac{\ln\left(x^2+y^2\right)}{2}\right] = 0$$

Hence, by integration, we have

$$x - \ln\sqrt{x^2+y^2} = k$$

Case 1:

When $\exists$ an integrating factor *u* (*x*), a function of x only. This happens if the expression

$$\frac{\dfrac{\partial M}{\partial y} - \dfrac{\partial N}{\partial x}}{N}$$

is a function of *x* only.

Then the integrating factor $u(x, y)$ is given by

$$u = \exp\left(\int \frac{\dfrac{\partial M}{\partial y} - \dfrac{\partial N}{\partial x}}{N} dx\right)$$

Case 2:

When $\exists$ an integrating factor $u(y)$, a function of *y* only. This happens if the expression

$$\frac{\dfrac{\partial N}{\partial x} - \dfrac{\partial M}{\partial y}}{M}$$

is a function of y only. Then **IF** $u(x, y)$ is given by

$$u = \exp\left(\int \frac{\dfrac{\partial N}{\partial x} - \dfrac{\partial M}{\partial y}}{M} dy\right)$$

Case 3:

If the given equation is homogeneous and

$$xM + yN \neq 0$$

Then $$u = \frac{1}{xM + yN}$$

Case 4:
If the given equation is of the form

$$yf(xy)dx + xg(xy)dy = 0$$

and $$xM - yN \neq 0$$

Then $$u = \frac{1}{xM - yN}$$

Once the **IF** is found, we multiply the old equation by *u* to get a new one, which is exact. Solve the exact equation and write the solution.

Advice: If possible, we should **check** whether or not the new equation is **exact?**

Summary:
Step 1. Write the given equation in the form

$$M(x, y)dx + N(x, y)dy = 0$$

provided the equation is not already in this form and determine M and N.
Step 2. Check for exactness of the equation by finding whether or not

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

Step 3. (a) If the equation is not exact, then evaluate

$$\frac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{N}$$

If this expression is a function of x only, then

$$u(x) = \exp\left(\int \frac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{N} dx\right)$$

Otherwise, evaluate

$$\frac{\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}}{M}$$

If this expression is a function of *y* only, then

$$u(y) = \exp\left(\int \frac{\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}}{M} dy\right)$$

In the **absence** of these **2 possibilities**, better use some other technique. However, we could also try cases 3 and 4 in step 4 and 5

Step 4. Test whether the equation is homogeneous and

$$xM + yN \neq 0$$

If yes then
$$u = \frac{1}{xM + yN}$$

Step 5. Test whether the equation is of the form

$$yf(xy)dx + xg(xy)dy = 0$$

and whether
$$xM - yN \neq 0$$

If yes then
$$u = \frac{1}{xM - yN}$$

Step 6. Multiply old equation by *u*. if possible, check whether or not the new equation is exact**?**

Step 7. Solve the new equation using steps described in the previous section.

Illustration:

Example 1

Solve the differential equation

$$\frac{dy}{dx} = -\frac{3xy + y^2}{x^2 + xy}$$

Solution:

1. The given differential equation can be written in form

$$(3xy + y^2)dx + (x^2 + xy)dy = 0$$

Therefore

$$M(x, y) = 3xy + y^2$$

$$N(x, y) = x^2 + xy$$

2. Now
$$\frac{\partial M}{\partial y} = 3x + 2y, \quad \frac{\partial N}{\partial x} = 2x + y.$$

$$\therefore \quad \frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$$

3. To find an **IF** we evaluate

$$\frac{\dfrac{\partial M}{\partial y} - \dfrac{\partial N}{\partial x}}{N} = \frac{1}{x}$$

which is a function of x only.

4.Therefore, an **IF** $u(x)$ exists and is given by

$$u(x) = e^{\int \frac{1}{x} dx} = e^{\ln(x)} = x$$

5. Multiplying the given equation with the **IF,** we obtain

$$(3x^2 y + xy^2)dx + (x^3 + x^2 y)dy = 0$$

which is exact. (Please check**!**)

6. This step consists of solving this last exact differential equation.

Solution of new exact equation:

1. Since $\dfrac{\partial M}{\partial y} = 3x^2 + 2xy = \dfrac{\partial N}{\partial x}$, the equation is exact.

2. We find *F* (*x*, *y*) by solving the system

$$\begin{cases} \dfrac{\partial F}{\partial x} = 3x^2 y + xy^2 \\ \dfrac{\partial F}{\partial y} = x^3 + x^2 y. \end{cases}$$

3. We integrate the first equation to get

$$F(x, y) = x^3 y + \frac{x^2}{2} y^2 + \theta(y)$$

4. We differentiate F w. r. t. **'y'** and use the second equation of the system in step 2 to obtain

$$\frac{\partial F}{\partial y} = x^3 + x^2 y + \theta'(y) = x^3 + x^2 y$$

$\Rightarrow \theta' = 0$, **No dependence on *x***.

5. Integrating the last equation to obtain $\theta = C$. Therefore, the function $F(x, y)$ is

$$F(x, y) = x^3 y + \frac{x^2}{2} y^2$$

We don't have to keep the constant *C*, see next step.

6. All the solutions are given by the implicit equation $F(x, y) = C$ **i.e.**

$$x^3 y + \frac{x^2 y^2}{2} = C$$

Note that it can be verified that the function

$$u(x, y) = \frac{1}{2xy(2x + y)}$$

is another integrating factor for the same equation as the new equation

$$\frac{1}{2xy(2x+y)}(3xy + y^2)dx + \frac{1}{2xy(2x+y)}(x^2 + xy)dy = 0$$

is exact. This means that we may **not have uniqueness** of the integrating factor.

Example 2. Solve

$$\left(x^2 - 2x + 2y^2\right)dx + 2xydy = 0$$

Solution:

$$M = x^2 - 2x + 2y^2$$
$$N = 2xy$$

$$\frac{\partial M}{\partial y} = 4y, \frac{\partial N}{\partial x} = 2y$$

$$\therefore \frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$$

The equation is not exact.

Here $\dfrac{M_y - N_x}{N} = \dfrac{4y - 2y}{2xy} = \dfrac{1}{x}$

Therefore, I.F. is given by

$$u = \exp\left(\int \frac{1}{x}dx\right)$$

$$u = x$$

$\therefore$ I.F is *x*.

Multiplying the equation by *x*, we have

$$\left(x^3 - 2x^2 + 2xy^2\right)dx + 2x^2 ydy = 0$$

This equation is exact. The required Solution is

$$\frac{x^4}{4} - \frac{2x^3}{3} + x^2y^2 = c_0$$

$$3x^4 - 8x^3 + 12x^2y^2 = c$$

Example 3

Solve $$dx + \left(\frac{x}{y} - \sin y\right)dy = 0$$

Solution: Here

$$M = 1, \quad N = \frac{x}{y} - \sin y$$
$$\frac{\partial M}{\partial y} = 0, \quad \frac{\partial N}{\partial x} = \frac{1}{y}$$
$$\therefore \frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$$

The equation is not exact.

Now

$$\frac{N_x - M_y}{M} = \frac{\frac{1}{y} - 0}{1} = \frac{1}{y}$$

Therefore, the IF is $$u(y) = \exp\int \frac{dy}{y} = y$$

Multiplying the equation by y, we have

$$ydx + (x - y\sin y)dy = 0$$

or $$ydx + xdy - y\sin ydy = 0$$

or $$d(xy) - y\sin ydy = 0$$

Integrating, we have

$$xy + y\cos y - \sin y = c$$

Which is the required solution.

Example 4

Solve $$\left(x^2y - 2xy^2\right)dx - \left(x^3 - 3x^2y\right)dy = 0$$

Solution: Comparing with

$$Mdx + Ndy = 0$$

we see that

$$M = x^2y - 2xy^2 \quad \text{and} \quad N = -(x^3 - 3x^2y)$$

Since both M and N are homogeneous. Therefore, the given equation is homogeneous.
Now

$$xM + yN = x^3y - 2x^2y^2 - x^3y + 3x^2y^2 = x^2y^2 \neq 0$$

Hence, the factor u is given by

$$u = \frac{1}{x^2y^2} \qquad \because u = \frac{1}{xM + yN}$$

Multiplying the given equation with the integrating factor u, we obtain.

$$\left(\frac{1}{y} - \frac{2}{x}\right)dx - \left(\frac{x}{y^2} - \frac{3}{y}\right)dy = 0$$

Now

$$M = \frac{1}{y} - \frac{2}{x} \quad \text{and} \quad N = \frac{-x}{y^2} + \frac{3}{y}$$

and therefore

$$\frac{\partial M}{\partial y} = -\frac{1}{y^2} = \frac{\partial N}{\partial x}$$

Therefore, the new equation is exact and solution of this new equation is given by

$$\frac{x}{y} - 2\ln|x| + 3\ln|y| = C$$

Example 5

Solve $$y\left(xy + 2x^2y^2\right)dx + x\left(xy - x^2y^2\right)dy = 0$$

Solution:
The given equation is of the form

$$yf(xy)dx + xg(xy)dy = 0$$

Now comparing with

$$Mdx + Ndy = 0$$

We see that

$$M = y\left(xy + 2x^2y^2\right) \quad \text{and} \quad N = x\left(xy - x^2y^2\right)$$

Further

$$xM - yN = x^2y^2 + 2x^3y^3 - x^2y^2 + x^3y^3$$
$$= \quad 3x^3y^3 \quad \neq \quad 0$$

Therefore, the integrating factor u is

$$u = \frac{1}{3x^3y^3}, \qquad \because u = \frac{1}{xM - yN}$$

Now multiplying the given equation by the integrating factor, we obtain

$$\frac{1}{3}\left(\frac{1}{x^2y} + \frac{2}{x}\right)dx + \frac{1}{3}\left(\frac{1}{xy^2} - \frac{1}{y}\right)dy = 0$$

Therefore, solutions of the given differential equation are given by

$$-\frac{1}{xy} + 2\ln|x| - \ln|y| = C$$

where $3C_0 = C$

Exercise

Solve by finding an I.F

1. $xdy - ydx = (x^2 + y^2)dx$
2. $dy + \frac{y - \sin x}{x}dx = 0$
3. $(y^4 + 2y)dx + (xy^3 + 2y^4 - 4x)dy = 0$
4. $(x^2 + y^2)dx + 2xydy = 0$
5. $(4x + 3y^2)dx + 2xydy = 0$
6. $(3x^2y^4 + 2xy)dx + (2x^3y^3)dy = 0$
7. $\frac{dy}{dx} = e^{2x} + y - 1$
8. $(3xy + y^2)dx + (x^2 + xy)dy = 0$
9. $ydx + (2xy - e^{-2y})dy = 0$
10. $(x + 2)\sin ydx + x\cos ydy = 0$

Lecture 7 First Order Linear Equations

The differential equation of the form:

$$a(x)\frac{dy}{dx}+b(x)y=c(x)$$

is a linear differential equation of first order. The equation can be rewritten in the following **famous form.**

$$\frac{dy}{dx}+p(x)y=q(x)$$

where $p(x)$ and $q(x)$ are continuous functions.

Method of solution:
The general solution of the first order linear differential equation is given by

$$y=\frac{\int u(x)q(x)dx+C}{u(x)}$$

Where $u(x)=\exp\left(\int p(x)dx\right)$

The function $u(x)$ is called the **integrating factor**. If it is an IVP then use it to find the constant *C*.

Summary:

1. Identify that the equation is 1st order linear equation. Rewrite it in the form

$$\frac{dy}{dx}+p(x)y=q(x)$$

if the equation is not already in this form.

2. Find the integrating factor

$$u(x)=e^{\int p(x)dx}$$

3. Write down the general solution

$$y=\frac{\int u(x)q(x)dx+C}{u(x)}$$

4. If you are given an IVP, use the initial condition to find the constant *C*.

5. Plug in the calculated value to write the particular solution of the problem.

Example 1:

Solve the initial value problem

$$y' + \tan(x)y = \cos^2(x), \qquad y(0) = 2$$

Solution:

1.The equation is already in the standard form

$$\frac{dy}{dx} + p(x)y = q(x)$$

with

$$\begin{cases} p(x) = \tan x \\ q(x) = \cos^2 x \end{cases}$$

2. Since

$$\int \tan x\, dx = -\ln \cos x = \ln \sec x$$

Therefore, the integrating factor is given by

$$u(x) = e^{\int \tan x\, dx} = \sec x$$

3. Further, because

$$\int \sec x \cos^2 x\, dx = \int \cos x\, dx = \sin x$$

So that the general solution is given by

$$y = \frac{\sin x + C}{\sec x} = (\sin x + C)\cos x$$

4. We use the initial condition $y(0) = 2$ to find the value of the constant C

$$y(0) = C = 2$$

5. Therefore the solution of the initial value problem is

$$y = (\sin x + 2)\cos x$$

Example 2: Solve the IVP $\dfrac{dy}{dt} - \dfrac{2t}{1+t^2} y = \dfrac{2}{1+t^2}, \qquad y(0) = 0.4$

Solution:

1.The given equation is a 1st order linear and is already in the requisite form

$$\frac{dy}{dx} + p(x)y = q(x)$$

with
$$\begin{cases} p(t) = -\dfrac{2t}{1+t^2} \\ q(t) = \dfrac{2}{1+t^2} \end{cases}$$

2. Since
$$\int \left(-\frac{2t}{1+t^2} \right) dt = -\ln|1+t^2|$$

Therefore, the integrating factor is given by

$$u(t) = e^{\int -\frac{2t}{1+t^2} dt} = (1+t^2)^{-1}$$

3. Hence, the general solution is given by

$$y = \frac{\int u(t)q(t)dt + C}{u(t)}, \quad \int u(t)q(t)dt = \int \frac{2}{(1+t^2)^2} dt$$

Now
$$\int \frac{2}{(1+t^2)^2} dt = 2\int \frac{1+t^2-t^2}{(1+t^2)^2} dt = 2\int \left(\frac{1}{1+t^2} - \frac{t^2}{(1+t^2)^2} \right) dt$$

The first integral is clearly $\tan^{-1} t$. For the 2nd we will use integration by parts with t as first function and $2t/(1+t^2)^2$ as 2nd function.

$$\int \frac{2t^2}{(1+t^2)^2} dt = t\left(-\frac{1}{1+t^2} \right) + \int \frac{1}{1+t^2} dt = -\frac{t}{1+t^2} + \tan^{-1}(t)$$

$$\int \frac{2}{(1+t^2)^2} dt = 2\tan^{-1}(t) + \frac{t}{1+t^2} - \tan^{-1}(t) = \tan^{-1}(t) + \frac{t}{1+t^2}$$

The general solution is: $y = (1+t^2)\left(\tan^{-1}(t) + \dfrac{t}{1+t^2} + C \right)$

4. The condition $y(0) = 0.4$ gives $C = 0.4$

5. Therefore, solution to the initial value problem can be written as**:**

$$y = t + (1+t^2)\tan^{-1}(t) + 0.4(1+t^2)$$

Example 3:

Find the solution to the problem

$$\cos^2 t \sin t \,.\, y' = -\cos^3 t \,.\, y + 1, \quad y\left(\frac{\pi}{4}\right) = 0$$

Solution:

1. The equation is 1st order linear and is not in the standard form

$$\frac{dy}{dx} + p(x)y = q(x)$$

Therefore we rewrite the equation as

$$y' + \frac{\cos t}{\sin t} y = \frac{1}{\cos^2 t \sin t}$$

2. Hence, the integrating factor is given by

$$u(t) = e^{\int \frac{\cos t}{\sin t} dt} = e^{\ln|\sin t|} = \sin t$$

3. Therefore, the general solution is given by

$$y = \frac{\int \sin t \frac{1}{\cos^2 t \sin t} dt + C}{\sin t}$$

Since

$$\int \sin t \frac{1}{\cos^2 t \, \sin t} dt = \int \frac{1}{\cos^2 t} dt = \tan t$$

Therefore

$$y = \frac{\tan t + C}{\sin t} = \frac{1}{\cos t} + \frac{C}{\sin t} = \sec t + C \csc t$$

(1) The initial condition $y(\pi/4) = 0$ implies

$$\sqrt{2} + C\sqrt{2} = 0$$

which gives $C = -1$.

(2) Therefore, the particular solution to the initial value problem is

$$y = \sec t - \csc t$$

Example 4

Solve $\left(x + 2y^3\right)\frac{dy}{dx} = y$

Solution:

We have

$$\frac{dy}{dx} = \frac{y}{x + 2y^3}$$

This equation is not linear in y. Let us regard x as dependent variable and y as independent variable. The equation may be written as

$$\frac{dx}{dy} = \frac{x + 2y^3}{y}$$

or

$$\frac{dx}{dy} - \frac{1}{y}x = 2y^2$$

Which is linear in x

$$IF = \exp\left[\int\left(-\frac{1}{y}\right)dy\right] = \exp\left[\ln\frac{1}{y}\right] = \frac{1}{y}$$

Multiplying with the $IF = \frac{1}{y}$, we get

$$\frac{1}{y}\frac{dx}{dy} - \frac{1}{y^2}x = 2y$$

$$\frac{d}{dy}\left(\frac{x}{y}\right) = 2y$$

Integrating, we have

$$\frac{x}{y} = y^2 + c$$

$$x = y\left(y^2 + c\right)$$

is the required solution.

Example 5

Solve

$$(x-1)^3\frac{dy}{dx}+4(x-1)^2y=x+1$$

Solution:

The equation can be rewritten as

$$\frac{dy}{dx}+\frac{4}{x-1}y=\frac{x+1}{(x-1)^3}$$

Here $P(x)=\dfrac{4}{x-1}.$

Therefore, an integrating factor of the given equation is

$$IF=\exp\left[\int\frac{4dx}{x-1}\right]=\exp\left[\ln(x-1)^4\right]=(x-1)^4$$

Multiplying the given equation by the IF, we get

$$(x-1)^4\frac{dy}{dx}+4(x-1)^3y=x^2-1$$

or $\dfrac{d}{dx}\left[y(x-1)^4\right]=x^2-1$

Integrating both sides, we obtain

$$y(x-1)^4=\frac{x^3}{3}-x+c$$

which is the required solution.

Exercise

Solve the following differential equations

1. $\dfrac{dy}{dx}+\left(\dfrac{2x+1}{x}\right)y=e^{-2x}$

2. $\dfrac{dy}{dx}+3y=3x^2e^{-3x}$

3. $x\dfrac{dy}{dx}+(1+x\cot x)y=x$

4. $(x+1)\dfrac{dy}{dx}-ny=e^x(x+1)^{n+1}$

5. $(1+x^2)\dfrac{dy}{dx}+4xy=\dfrac{1}{(1+x^2)^2}$

6. $\dfrac{dr}{d\theta}+r\sec\theta=\cos\theta$

7. $\dfrac{dy}{dx}+y=\dfrac{1-e^{-2x}}{e^x+e^{-x}}$

8. $dx=(3e^y-2x)dy$

Solve the initial value problems

9. $\dfrac{dy}{dx}=2y+x(e^{3x}-e^{2x}), \qquad y(0)=2$

10. $x(2+x)\dfrac{dy}{dx}+2(1+x)y=1+3x^2, \qquad y(-1)=1$

Lecture 8 Bernoulli Equations

A differential equation that can be written in the form

$$\frac{dy}{dx} + p(x)y = q(x)y^n$$

is called Bernoulli equation.

Method of solution:

For $n = 0,1$ the equation reduces to 1st order linear DE and can be solved accordingly.

For $n \neq 0,1$ we divide the equation with y^n to write it in the form

$$y^{-n}\frac{dy}{dx} + p(x)y^{1-n} = q(x)$$

and then put

$$v = y^{1-n}$$

Differentiating w.r.t. 'x', we obtain

$$v' = (1-n)y^{-n}y'$$

Therefore the equation becomes

$$\frac{dv}{dx} + (1-n)p(x)v = (1-n)q(x)$$

This is a linear equation satisfied by v. Once it is solved, you will obtain the function

$$y = v^{\frac{1}{(1-n)}}$$

If $n > 1$, then we add the solution $y = 0$ to the solutions found the above technique.

Summary:

1.Identify the equation

$$\frac{dy}{dx} + p(x)y = q(x)y^n$$

as Bernoulli equation.

Find ***n***. If $n \neq 0,1$ divide by y^n and substitute**;**

$$v = y^{1-n}$$

2. Through easy differentiation, find the new equation

$$\frac{dv}{dx} + (1-n)p(x)v = (1-n)q(x)$$

3. This is a linear equation. Solve the linear equation to find ***v*.**

4. Go back to the old function *y* through the substitution $y = v^{1/(1-n)}$.

6. If $n > 1$, then include *y* = 0 to in the solution**.**

7. If you have an IVP, use the initial condition to find the particular solution.

Example 1: Solve the equation $\frac{dy}{dx} = y + y^3$

Solution:

1. The given differential can be written as

$$\frac{dy}{dx} - y = y^3$$

which is a Bernoulli equation with

$$p(x) = -1, q(x) = 1, \ \mathbf{n=3.}$$

Dividing with y^3 we get

$$y^{-3}\frac{dy}{dx} - y^{-2} = 1$$

Therefore we substitute

$$v = y^{1-3} = y^{-2}$$

2. Differentiating w.r.t. 'x' we have

$$y^{-3}\frac{dy}{dx}=-\frac{1}{2}\left(\frac{dv}{dx}\right)$$

So that the equation reduces to

$$\frac{dv}{dx}+2v=-2$$

3. This is a linear equation. To solve this we find the integrating factor $u(x)$

$$u(x)=e^{\int 2dx}=e^{2x}$$

The solution of the linear equation is given by

$$v=\frac{\int u(x)q(x)dx+c}{u(x)}=\frac{\int e^{2x}(-2)dx+c}{e^{2x}}$$

Since $$\int e^{2x}(-2)dx=-e^{2x}$$

Therefore, the solution for v is given by

$$v=\frac{-e^{2x}+C}{e^{2x}}=Ce^{-2x}-1$$

4. To go back to y we substitute $v=y^{-2}$. Therefore the general solution of the given DE is

$$y=\pm\left(Ce^{-2x}-1\right)^{-\frac{1}{2}}$$

5. Since $n>1$, we include the $y=0$ in the solutions. Hence, all solutions are

$$y=0, \qquad y=\pm(Ce^{-2x}-1)^{-1/2}$$

Example 2:

Solve $$\frac{dy}{dx}+\frac{1}{x}y=xy^2$$

Solution: In the given equation we identify $P(x)=\frac{1}{x},\ q(x)=x$ and $n=2$.

Thus the substitution $w=y^{-1}$ gives

$$\frac{dw}{dx}-\frac{1}{x}w=-x.$$

The integrating factor for this linear equation is

$$e^{-\int\frac{dx}{x}}=e^{-\ln|x|}=e^{\ln|x|^{-1}}=x^{-1}$$

Hence $$\frac{d}{dx}\left[x^{-1}w\right]=-1.$$

Integrating this latter form, we get

$$x^{-1}w = -x + c \text{ or } w = -x^2 + cx.$$

Since $w = y^{-1}$, we obtain $y = \frac{1}{w}$ or

$$y = \frac{1}{-x^2 + cx}$$

For $n > 0$ the trivial solution $y = 0$ is a solution of the given equation. In this example, $y = 0$ is a singular solution of the given equation.

Example 3:

Solve:

$$\frac{dy}{dx} + \frac{xy}{1-x^2} = xy^{\frac{1}{2}} \qquad (1)$$

Solution: Dividing (1) by $y^{\frac{1}{2}}$, the given equation becomes

$$y^{\frac{-1}{2}}\frac{dy}{dx} + \frac{x}{1-x^2}y^{\frac{1}{2}} = x \qquad (2)$$

Put

$$y^{\frac{1}{2}} = v \text{ or. } \frac{1}{2}y^{-\frac{1}{2}}\frac{dy}{dx} = \frac{dv}{dx}$$

Then (2) reduces to

$$\frac{dv}{dx} + \frac{x}{2(1-x^2)}v = \frac{x}{2} \qquad (3)$$

This is linear in v.

$$\text{I.F} = \exp\left[\int \frac{x}{2(1-x^2)}dx\right] = \exp\left[\frac{-1}{4}\ln(1-x^2)\right] = (1-x^2)^{\frac{-1}{4}}$$

Multiplying (3) by $(1-x^2)^{\frac{-1}{4}}$, we get

$$(1-x^2)^{\frac{-1}{4}}\frac{dv}{dx} + \frac{x}{2(1-x^2)^{5/4}}v = \frac{x}{2(1-x^2)^{1/4}}$$

or

$$\frac{d}{dx}\left[(1-x^2)^{\frac{-1}{4}}v\right] = \frac{-1}{4}\left[-2x(1-x^2)^{\frac{-1}{4}}\right]$$

Integrating, we have

$$v\left(1-x^2\right)^{\frac{-1}{4}}=\frac{-1}{4}\frac{\left(1-x^2\right)^{\frac{3}{4}}}{3/4}+c$$

or
$$v=c\left(1-x^2\right)^{1/4}-\frac{1-x^2}{3}$$

or
$$y^{\frac{1}{2}}=c\left(1-x^2\right)^{1/4}-\frac{1-x^2}{3}$$

is the required solution.

Exercise

Solve the following differential equations

1. $x\frac{dy}{dx} + y = y^2 \ln x$

2. $\frac{dy}{dx} + y = xy^3$

3. $\frac{dy}{dx} - y = e^x y^2$

4. $\frac{dy}{dx} = y\left(xy^3 - 1\right)$

5. $x\frac{dy}{dx} - (1 + x)y = xy^2$

6. $x^2\frac{dy}{dx} + y^2 = xy$

Solve the initial-value problems

7. $x^2\frac{dy}{dx} - 2xy = 3y^4, \quad y(1) = \frac{1}{2}$

8. $y^{1/2}\frac{dy}{dx} + y^{3/2} = 1, \quad y(0) = 4$

9. $xy\left(1 + xy^2\right)\frac{dy}{dx} = 1, \quad y(1) = 0$

10. $2\frac{dy}{dx} = \frac{y}{x} - \frac{x}{y^2}, \quad y(1) = 1$

SUBSTITUTIONS

- Sometimes a differential equation can be transformed by means of a substitution into a form that could then be solved by one of the standard methods i.e. Methods used to solve separable, homogeneous, exact, linear, and Bernoulli's differential equation.

- An equation may look different from any of those that we have studied in the previous lectures, but through a sensible change of variables perhaps an apparently difficult problem may be readily solved.

- Although no firm rules can be given on the basis of which these substitution could be selected, a working axiom might be: Try something! It sometimes pays to be clever.

Example 1

The differential equation

$$y(1+2xy)dx + x(1-2xy)dy = 0$$

is not separable, not homogeneous, not exact, not linear, and not Bernoulli.
However, if we stare at the equation long enough, we might be prompted to try the substitution

$$u = 2xy \quad \text{or} \quad y = \frac{u}{2x}$$

Since

$$dy = \frac{xdu - udx}{2x^2}$$

The equation becomes, after we simplify

$$2u^2dx + (1-u)xdu = 0.$$

we obtain

$$2\ln|x| - u^{-1} - \ln|u| = c$$

$$\ln\left|\frac{x}{2y}\right| = c + \frac{1}{2xy}$$

$$\frac{x}{2y} = c_1 e^{1/2xy},$$

$$x = 2c_1 y e^{1/2xy}$$

where e^c was replaced by c_1. We can also replace $2c_1$ by c_2 if desired

Note that

The differential equation in the example possesses the trivial solution $y = 0$, but then this function is not included in the one-parameter family of solution.

Example 2

Solve

$$2xy\frac{dy}{dx} + 2y^2 = 3x - 6.$$

Solution:

The presence of the term $2y\dfrac{dy}{dx}$ prompts us to try $u = y^2$

Since

$$\frac{du}{dx} = 2y\frac{dy}{dx}$$

Therefore, the equation becomes

Now

$$x\frac{du}{dx} + 2u = 3x - 6$$

or

$$\frac{du}{dx} + \frac{2}{x}u = 3 - \frac{6}{x}$$

This equation has the form of 1st order linear differential equation

$$\frac{dy}{dx} + P(x)y = Q(x)$$

with

$$P(x) = \frac{2}{x} \text{ and } Q(x) = 3 - \frac{6}{x}$$

Therefore, the integrating factor of the equation is given by

$$\text{I.F} = e^{\int \frac{2}{x}dx} = e^{\ln x^2} = x^2$$

Multiplying with the IF gives

$$\frac{d}{dx}\left[x^2 u\right] = 3x^2 - 6x$$

Integrating both sides, we obtain

$$x^2 u = x^3 - 3x^2 + c$$

or

$$x^2 y^2 = x^3 - 3x^2 + c.$$

Example 3

Solve

$$x\frac{dy}{dx} - y = \frac{x^3}{y}e^{y/x}$$

Solution:

If we let

$$u = \frac{y}{x}$$

Then the given differential equation can be simplified to

$$ue^{-u}du = dx$$

Integrating both sides, we have

$$\int ue^{-u}du = \int dx$$

Using the integration by parts on LHS, we have

$$-ue^{-u} - e^{-u} = x + c$$

or

$$u + 1 = (c_1 - x)e^{u} \quad \text{Where } c_1 = -c$$

We then re-substitute

$$u = \frac{y}{x}$$

and simplify to obtain

$$y + x = x(c_1 - x)e^{y/x}$$

Example 4

Solve

$$\frac{d^2y}{dx^2} = 2x\left(\frac{dy}{dx}\right)^2$$

Solution:
If we let

$$u = y'$$

Then

$$du/dx = y''$$

Then, the equation reduces to

$$\frac{du}{dx} = 2xu^2$$

Which is separable form. Separating the variables, we obtain

$$\frac{du}{u^2} = 2xdx$$

Integrating both sides yields

$$\int u^{-2}du = \int 2xdx$$

or

$$-u^{-1} = x^2 + c_1^2$$

The constant is written as c_1^2 for convenience.

Since

$$u^{-1} = 1/y'$$

Therefore

$$\frac{dy}{dx} = -\frac{1}{x^2 + c_1^2}$$

or

$$dy = -\frac{dx}{x^2 + c_1^2}$$

$$\int dy = -\int \frac{dx}{x^2 + c_1^2}$$

$$y + c_2 = -\frac{1}{c_1}\tan^{-1}\frac{x}{c_1}$$

Exercise

Solve the differential equations by using an appropriate substitution.

1. $ydx + (1 + ye^{x})dy = 0$

2. $\left(2 + e^{-x/y}\right)dx + 2\left(1 - x/y\right)dy = 0$

3. $2x\csc 2y\,\dfrac{dy}{dx} = 2x - \ln(\tan y)$

4. $\dfrac{dy}{dx} + 1 = \sin x\, e^{-(x+y)}$

5. $y\dfrac{dy}{dx} + 2x\ln x = xe^{y}$

6. $x^{2}\dfrac{dy}{dx} + 2xy = x^{4}y^{2} + 1$

7. $xe^{y}\dfrac{dy}{dx} - 2e^{y} = x^{2}$

Lecture 9 Practice Examples

Example 1: $y'=\dfrac{x^2+y^2}{xy}$

Solution: $\dfrac{dy}{dx}=\dfrac{x^2+y^2}{xy}$

put y=wx then $\dfrac{dy}{dx}=w+x\dfrac{dw}{dx}$

$$w+x\frac{dw}{dx}=\frac{x^2+w^2x^2}{xxw}=\frac{1+w^2}{w}$$

$$w+x\frac{dw}{dx}=\frac{1}{w}+w$$

$$wdw=\frac{dx}{x}$$

Integrating

$$\frac{w^2}{2}=\ln x+\ln c$$

$$\frac{y^2}{2x^2}=\ln|xc|$$

$$y^2=2x^2\ln|xc|$$

Example 2: $\frac{dy}{dx}=\frac{(2\sqrt{xy}-y)}{x}$

Solution: $\frac{dy}{dx}=\frac{(2\sqrt{xy}-y)}{x}$

put y= wx

$w+x\frac{dw}{dx}=\frac{(2\sqrt{xwx}-xw)}{x}$

$w+x\frac{dw}{dx}=2\sqrt{w}-w$

$x\frac{dw}{dx}=2\sqrt{w}-2w$

$\frac{dw}{2(\sqrt{w}-w)}=\frac{dx}{x}$

$\int\frac{dw}{2(\sqrt{w}-w)}=\int\frac{dx}{x}$

$\int\frac{dw}{2\sqrt{w}(1-\sqrt{w})}=\int\frac{dx}{x}$

put $\sqrt{w}=t$

We get $\int\frac{1}{1-t}dt=\int\frac{dx}{x}$

$-\ln|1-t|=\ln|x|+\ln|c|$

$-\ln|1-t|=\ln|xc|$

$(1-t)^{-1}=xc$

$(1-\sqrt{w})^{-1}=xc$

$(1-\sqrt{y/x})^{-1}=xc$

Example 3: $(2y^2x-3)dx+(2yx^2+4)dy=0$

Solution: $(2y^2x-3)dx+(2yx^2+4)dy=0$

Here $M=(2y^2x-3)$ and $N=(2yx^2+4)$

$$\frac{\partial M}{\partial y}=4xy=\frac{\partial N}{\partial x}$$

$$\frac{\partial f}{\partial x}=(2y^2x-3) \quad \text{and} \quad \frac{\partial f}{\partial y}=(2yx^2+4)$$

Integrate w.r.t. 'x'

$f(x,y)=x^2y^2-3x+h(y)$

Differentiate w.r.t. 'y'

$$\frac{\partial f}{\partial y}=2x^2y+h'(y)=2x^2y+4=N$$

$h'(y)=4$

Integrate w.r.t. 'y'

$h(y)=4y+c$

$x^2y^2-3x+4y=C_1$

Example 4: $$\frac{dy}{dx}=\frac{2xye^{(x/y)^2}}{y^2+y^2e^{(x/y)^2}+2x^2e^{(x/y)^2}}$$

Solution: $$\frac{dx}{dy}=\frac{y^2+y^2e^{(x/y)^2}+2x^2e^{(x/y)^2}}{2xye^{(x/y)^2}}$$

put x/y=w

Aftersubsitution

$$y\frac{dw}{dy}=\frac{1+e^{w^2}}{2we^{w^2}}$$

$$\frac{dy}{y}=\frac{2we^{w^2}}{1+e^{w^2}}dw$$

Integrating

$\ln|y|=\ln|1+e^{w^2}|+\ln c$

$\ln|y|=\ln|c(1+e^{w^2})|$

$y=c(1+e^{(x/y)^2})$

Example 5: $\frac{dy}{dx}+\frac{y}{x\ln x}=\frac{3x^2}{\ln x}$

Solution: $\frac{dy}{dx}+\frac{y}{x\ln x}=\frac{3x^2}{\ln x}$

$$\frac{dy}{dx}+\frac{1}{x\ln x}y=\frac{3x^2}{\ln x}$$

$p(x)=\frac{1}{x\ln x}$ and $q(x)=\frac{3x^2}{\ln x}$

$$I.F=\exp\left(\int\frac{1}{x\ln x}dx\right)=\ln x$$

Multiply both side by lnx

$$\ln x\frac{dy}{dx}+\frac{1}{x}y=3x^2$$

$$\frac{d}{dx}(y\ln x)=3x^2$$

Integrate

$$y\ln x=\frac{3x^3}{3}+c$$

Example 6: $(y^2e^x+2xy)dx-x^2dy=0$

Solution: Here $M=y^2e^x+2xy \qquad N=-x^2$

$$\frac{\partial M}{\partial y}=2ye^x+2x, \qquad \frac{\partial N}{\partial x}=-2x$$

Clearly $\dfrac{\partial M}{\partial y}\neq\dfrac{\partial N}{\partial x}$

The given equation is not exact

divide the equation by y^2 to make it exact

$$\left[e^x+\frac{2x}{y}\right]dx+\left[-\frac{x^2}{y^2}\right]dy=0$$

Now $\dfrac{\partial M}{\partial y}=-\dfrac{2x}{y^2}=\dfrac{\partial N}{\partial x}$

Equation is exact

$$\frac{\partial f}{\partial x}=\left[e^x+\frac{2x}{y}\right] \qquad \frac{\partial f}{\partial y}=\left[-\frac{x^2}{y^2}\right]$$

Integrate w.r.t. 'x'

$$f(x,y)=e^x+\frac{x^2}{y}$$

$$e^x+\frac{x^2}{y}=c$$

Example 7:

$$x\cos x\frac{dy}{dx}+y(x\sin x+\cos x)=1$$

Solution: $x\cos x\frac{dy}{dx}+y(x\sin x+\cos x)=1$

$$\frac{dy}{dx}+y\left[\frac{x\sin x+\cos x}{x\cos x}\right]=\frac{1}{x\cos x}$$

$$\frac{dy}{dx}+y\left[\tan x+1/x\right]=\frac{1}{x\cos x}$$

$$I.F = \exp(\int(\tan x+1/x)dx)=x\sec x$$

$$x\sec x\frac{dy}{dx}+yx\sec x\left[\tan x+1/x\right]=\frac{x\sec x}{x\cos x}$$

$$x\sec x\frac{dy}{dx}+y\left[x\sec x\tan x+\sec x\right]=\sec^2 x$$

$$\frac{d}{dx}\left[xy\sec x\right]=\sec^2 x$$

$$xy\sec x=\tan x+c$$

Example 8: $xe^{2y}\frac{dy}{dx}+e^{2y}=\frac{\ln x}{x}$

Solution: $xe^{2y}\frac{dy}{dx}+e^{2y}=\frac{\ln x}{x}$

put $e^{2y}=u$

$$2e^{2y}\frac{dy}{dx}=\frac{du}{dx}$$

$$\frac{x}{2}\frac{du}{dx}+u=\frac{\ln x}{x}$$

$$\frac{du}{dx}+\frac{2}{x}u=2\frac{\ln x}{x^2}$$

Here $p(x)=2/x$ And $Q(x)=\frac{\ln x}{x^2}$

$$I.F=\exp(\int\frac{2}{x}dx)=x^2$$

$$x^2\frac{du}{dx}+2xu=2\ln x$$

$$\frac{d}{dx}(x^2u)=2\ln x$$

Integrate

$$x^2u=2[x\ln x-x]+c$$

$$x^2e^{2y}=2[x\ln x-x]+c$$

Example 9: $\frac{dy}{dx}+y\ln y=ye^{x}$

Solution: $\frac{dy}{dx}+y\ln y=ye^{x}$

$$\frac{1}{y}\frac{dy}{dx}+\ln y=e^{x}$$

put $\ln y=u$

$$\frac{du}{dx}+u=e^{x}$$

$$I.F.=e^{\int dx}=e^{x}$$

$$\frac{d}{dx}(e^{x}u)=e^{2x}$$

Integrate

$$e^{x}.u=\frac{e^{2x}}{2}+c$$

$$e^{x}\ln y=\frac{e^{2x}}{2}+c$$

Example 10: $2x\csc 2y\dfrac{dy}{dx}=2x-\ln\tan y$

Solution: $2x\csc 2y\dfrac{dy}{dx}=2x-\ln\tan y$

put $\ln\tan y=u$

$$\frac{dy}{dx}=\sin y\cos y\frac{du}{dx}$$

$$\frac{2x\sin y\cos y}{2\sin y\cos y}\frac{du}{dx}=2x-u$$

$$x\frac{du}{dx}=2x-u$$

$$\frac{du}{dx}+\frac{1}{x}u=2$$

$$\text{I.F}=\exp\left(\int 1/x\,dx\right)=x$$

$$x\frac{du}{dx}+u=2x$$

$$\frac{d}{dx}(xu)=2x$$

$$xu=x^2+c$$

$$u=x+cx^{-1}$$

$$\ln\tan y=x+cx^{-1}$$

Example 11: $\frac{dy}{dx}+x+y+1=(x+y)^2e^{3x}$

Solution: $\frac{dy}{dx}+x+y+1=(x+y)^2e^{3x}$

Put $x+y=u$

$$\frac{du}{dx}+u=u^2e^{3x}$$

$$\frac{du}{dx}+u=u^2e^{3x} \text{ (Bernouli's)}$$

$$\frac{1}{u^2}\frac{du}{dx}+\frac{1}{u}=e^{3x}$$

put $1/u=w$

$$-\frac{dw}{dx}+w=e^{3x}$$

$$\frac{dw}{dx}-w=-e^{3x}$$

$$I.F=\exp(\int -dx)=e^{-x}$$

$$e^{-x}\frac{dw}{dx}-we^{-x}=-e^{2x}$$

$$\frac{d}{dx}(e^{-x}w)=-e^{2x}$$

Integrate

$$e^{-x}w=\frac{-e^{2x}}{2}+c$$

$$\frac{1}{u}=\frac{-e^{3x}}{2}+ce^{x}$$

$$\frac{1}{x+y}=\frac{-e^{3x}}{2}+ce^{x}$$

Example 12: $\frac{dy}{dx}=(4x+y+1)^2$

Solution: $\frac{dy}{dx}=(4x+y+1)^2$

put $4x+y+1=u$

we get

$$\frac{du}{dx}-4=u^2$$

$$\frac{du}{dx}=u^2+4$$

$$\frac{1}{u^2+4}du=dx$$

Integrate

$$\frac{1}{2}\tan^{-1}\frac{u}{2}=x+c$$

$$\tan^{-1}\frac{u}{2}=2x+c_1$$

$$u=2\tan(2x+c_1)$$

$$4x+y+1=2\tan(2x+c_1)$$

Example 13: $(x+y)^2 \frac{dy}{dx} = a^2$

Solution: $(x+y)^2 \frac{dy}{dx} = a^2$

put $x+y = u$

$$u^2\left(\frac{du}{dx} - 1\right) = a^2$$

$$u^2 \frac{du}{dx} - u^2 = a^2$$

$$\frac{u^2}{u^2+a^2} du = dx$$

Integrate

$$\int \frac{u^2+a^2-a^2}{u^2+a^2} du = \int dx$$

$$\int \left(1 - \frac{a^2}{u^2+a^2}\right) du = \int dx$$

$$u - a\tan^{-1}\frac{u}{a} = x + c$$

$$(x+y) - a\tan^{-1}\frac{x+y}{a} = x + c$$

Example 14: $2y\frac{dy}{dx}+x^2+y^2+x=0$

Solution: $2y\frac{dy}{dx}+x^2+y^2+x=0$

put $x^2+y^2=u$

$$\frac{du}{dx}-2x+u+x=0$$

$$\frac{du}{dx}+u=x$$

$$\text{I.F}=\text{Exp}(\int dx)=e^x$$

$$e^x\frac{du}{dx}+ue^x=xe^x$$

$$\frac{d}{dx}(e^x u)=xe^x$$

Integrating

$$e^x u=xe^x-e^x+c$$

Example 15: $y' + 1 = e^{-(x+y)} \sin x$

Solution: $y' + 1 = e^{-(x+y)} \sin x$

put $x + y = u$

$$\frac{du}{dx} = e^{-u} \sin x$$

$$\frac{1}{e^{-u}} du = \sin x dx$$

$$e^{u} du = \sin x dx$$

Integrate

$$e^{u} = -\cos x + c$$

$$u = \ln|-\cos x + c|$$

$$x + y = \ln|-\cos x + c|$$

Example 16: $x^4y^2y'+x^3y^3=2x^3-3$

Solution: $x^4y^2y'+x^3y^3=2x^3-3$

put $x^3y^3=u$

$$3x^2y^3+3x^3y^2\frac{dy}{dx}=\frac{du}{dx}$$

$$3x^3y^2\frac{dy}{dx}=\frac{du}{dx}-3x^2y^3$$

$$x^4y^2\frac{dy}{dx}=\frac{x}{3}\frac{du}{dx}-x^3y^3$$

$$\frac{x}{3}\frac{du}{dx}=2x^3-3$$

$$\frac{du}{dx}=6x^2-9/x$$

Integrate

$u=2x^3-9\ln x+c$

$x^3y^3=2x^3-9\ln x+c$

Example 17: $\cos(x+y)dy=dx$

Solution: $\cos(x+y)dy=dx$

put $x+y=v$ or $1+\frac{dy}{dx}=\frac{dv}{dx}$, we get

$$\cos v\left[\frac{dv}{dx}-1\right]=1$$

$$dx=\frac{\cos v}{1+\cos v}dv=\left[1-\frac{1}{1+\cos v}\right]dv$$

$$dx=\left[1-\frac{1}{2}\sec^2\frac{v}{2}\right]dv$$

Integrate

$$x+c=v-\tan\frac{v}{2}$$

$$x+c=v-\tan\frac{x+y}{2}$$

Lecture 10 Applications of First Order Differential Equations

In order to translate a physical phenomenon in terms of mathematics, we strive for a set of equations that describe the system adequately. This set of equations is called a **Model** for the phenomenon. The basic steps in building such a model consist of the following steps:

Step 1: We clearly state the assumptions on which the model will be based. These assumptions should describe the relationships among the quantities to be studied.

Step 2: Completely describe the parameters and variables to be used in the model.

Step 3: Use the assumptions (from Step 1) to derive mathematical equations relating the parameters and variables (from Step 2).

The mathematical models for physical phenomenon often lead to a differential equation or a set of differential equations. The applications of the differential equations we will discuss in next two lectures include:

- Orthogonal Trajectories.
- Population dynamics.
- Radioactive decay.
- Newton's Law of cooling.
- Carbon dating.
- Chemical reactions.

etc.

Orthogonal Trajectories

- We know that that the solutions of a 1st order differential equation, e.g. separable equations, may be given by an implicit equation

$$F(x, y, C) = 0$$

with 1 parameter C, which represents a family of curves. Member curves can be obtained by fixing the parameter C. Similarly an nth order DE will yields an n-parameter family of curves/solutions.

$$F(x, y, C_1, C_1, \cdots, C_n) = 0$$

- The question arises that whether or not we can turn the problem around: Starting with an n-parameter family of curves, can we find an associated n^{th} order differential equation free of parameters and representing the family. The answer in most cases is yes.

- Let us try to see, with reference to a 1-parameter family of curves, how to proceed if the answer to the question is yes.

1. Differentiate with respect to x, and get an equation-involving x, y, $\frac{dy}{dx}$ and C.
2. Using the original equation, we may be able to eliminate the parameter C from the new equation.
3. The next step is doing some algebra to rewrite this equation in an explicit form

$$\frac{dy}{dx} = f(x, y)$$

- For illustration we consider an example:

Illustration

Example

Find the differential equation satisfied by the family

$$x^2 + y^2 = C\,x$$

Solution:

1. We differentiate the equation with respect to x, to get

$$2x + 2y\frac{dy}{dx} = C$$

2. Since we have from the original equation that

$$C = \frac{x^2 + y^2}{x}$$

then we get

$$2x + 2y\frac{dy}{dx} = \frac{x^2 + y^2}{x}$$

3. The explicit form of the above differential equation is

$$\frac{dy}{dx} = \frac{y^2 - x^2}{2xy}$$

This last equation is the desired DE free of parameters representing the given family.

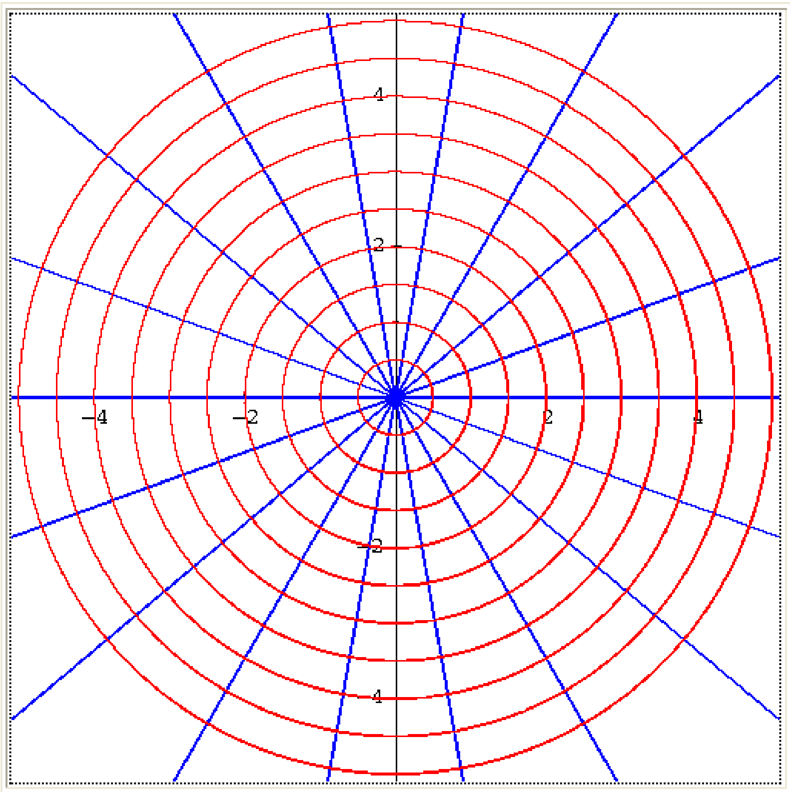
Example.

Let us consider the example of the following two families of curves

$$\begin{cases} y & = mx \\ x^2 + y^2 & = C^2 \end{cases}$$

The first family describes all the straight lines passing through the origin while the second family describes all the circles centered at the origin

If we draw the two families together on the same graph we get



Clearly whenever one line intersects one circle, the tangent line to the circle (at the point of intersection) and the line are perpendicular i.e. orthogonal to each other. We say that the two families of curves are orthogonal at the point of intersection.

Orthogonal curves:

Any two curves C_1 and C_2 are said to be orthogonal if their tangent lines T_1 and T_2 at their point of intersection are perpendicular. This means that slopes are negative reciprocals of each other, except when T_1 and T_2 are parallel to the coordinate axes.

Orthogonal Trajectories (OT):

When all curves of a family $\Im_1 : G(x, y, c_1) = 0$ orthogonally intersect all curves of another family $\Im_2 : H(x, y, c_2) = 0$ then each curve of the families is said to be orthogonal trajectory of the other.

Example:

As we can see from the previous figure that the family of straight lines $y = mx$ and the family of circles $x^2 + y^2 = C^2$ are orthogonal trajectories.

Orthogonal trajectories occur naturally in many areas of physics, fluid dynamics, in the study of electricity and magnetism etc. For example the lines of force are perpendicular to the equipotential curves i.e. curves of constant potential.

Method of finding Orthogonal Trajectory:

Consider a family of curves $\Im$. Assume that an associated DE may be found, which is given by**:**

$$\frac{dy}{dx} = f(x, y)$$

Since $\frac{dy}{dx}$ gives slope of the tangent to a curve of the family $\Im$ through (x, y).

Therefore, the slope of the line orthogonal to this tangent is $-\frac{1}{f(x, y)}$. So that the slope of the line that is tangent to the orthogonal curve through (x, y) is given by $-\frac{1}{f(x, y)}$. In other words, the family of orthogonal curves are solutions to the differential equation

$$\frac{dy}{dx} = -\frac{1}{f(x, y)}$$

The steps can be summarized as follows:

Summary:

In order to find Orthogonal Trajectories of a family of curves $\Im$ we perform the following steps:

Step 1. Consider a family of curves $\Im$ and find the associated differential equation.
Step 2. Rewrite this differential equation in the explicit form

$$\frac{dy}{dx} = f(x,y)$$

Step 3. Write down the differential equation associated to the orthogonal family

$$\frac{dy}{dx} = -\frac{1}{f(x,y)}$$

Step 4. Solve the new equation. The solutions are exactly the family of orthogonal curves.
Step 5. A specific curve from the orthogonal family may be required, something like an IVP.

Example 1
Find the orthogonal Trajectory to the family of circles

$$x^2 + y^2 = C^2$$

Solution:
The given equation represents a family of concentric circles centered at the origin.
Step 1. We differentiate w.r.t. 'x' to find the DE satisfied by the circles.

$$2y\frac{dy}{dx} + 2x = 0$$

Step 2. We rewrite this equation in the explicit form

$$\frac{dy}{dx} = -\frac{x}{y}$$

Step 3. Next we write down the DE for the orthogonal family

$$\frac{dy}{dx} = -\frac{1}{-(x/y)} = \frac{y}{x}$$

Step 4.This is a linear as well as a separable DE. Using the technique of linear equation, we find the integrating factor

$$u(x) = e^{-\int \frac{1}{x}dx} = \frac{1}{x}$$

which gives the solution

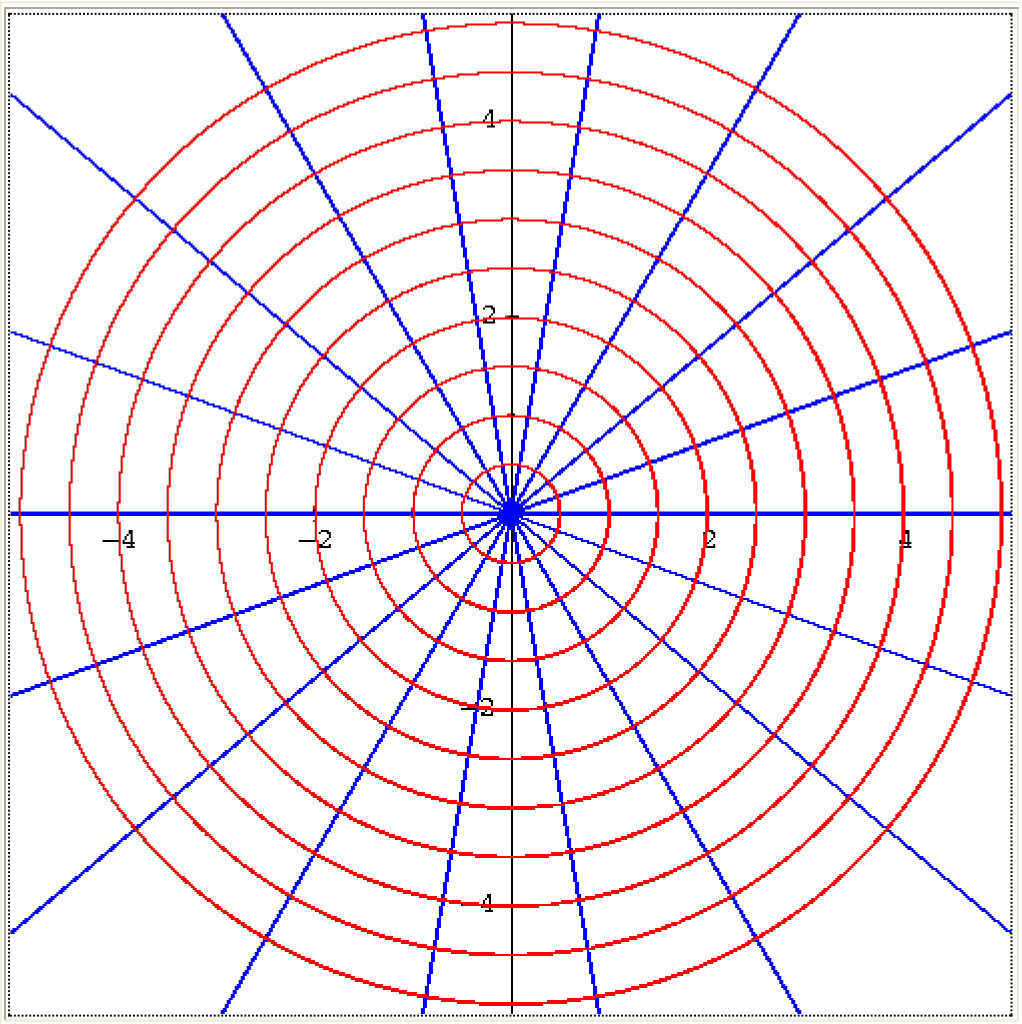
$$y\,.u(x) = m$$

or

$$y = \frac{m}{u(x)} = mx$$

Which represent a family of straight lines through origin. Hence the family of straight lines $y = mx$ and the family of circles $x^2 + y^2 = C^2$ are Orthogonal Trajectories.

Step 5. A geometrical view of these Orthogonal Trajectories is:



Example 2

Find the Orthogonal Trajectory to the family of circles

$$x^2 + y^2 = 2Cx$$

Solution:

1. We differentiate the given equation to find the DE satisfied by the circles.

$$y\frac{dy}{dx} + x = C, \qquad C = \frac{x^2 + y^2}{2x}$$

2. The explicit differential equation associated to the family of circles is

$$\frac{dy}{dx} = \frac{y^2 - x^2}{2xy}$$

3. Hence the differential equation for the orthogonal family is

$$\frac{dy}{dx} = \frac{2xy}{x^2 - y^2}$$

4. This DE is a homogeneous, to solve this equation we substitute $v = y/x$ or equivalently $y = vx$. Then we have

$$\frac{dy}{dx} = x\frac{dv}{dx} + v \text{ and } \frac{2xy}{x^2 - y^2} = \frac{2v}{1-v^2}$$

Therefore the homogeneous differential equation in step 3 becomes

$$x\frac{dv}{dx} + v = \frac{2v}{1-v^2}$$

Algebraic manipulations reduce this equation to the separable form:

$$\frac{dv}{dx} = \frac{1}{x}\left\{\frac{v+v^3}{1-v^2}\right\}$$

The constant solutions are given by

$$v + v^3 = 0 \Rightarrow v(1+v^2) = 0$$

The only constant solution is $v = 0$.

To find the non-constant solutions we separate the variables

$$\frac{1-v^2}{v+v^3}dv = \frac{1}{x}dx$$

Integrate

$$\int \frac{1-v^2}{v+v^3}\,dv = \int \frac{1}{x}\,dx$$

Resolving into partial fractions the integrand on LHS, we obtain

$$\frac{1-v^2}{v+v^3} = \frac{1-v^2}{v(1+v^2)} = \frac{1}{v} - \frac{2v}{1+v^2}$$

Hence we have

$$\int \frac{1-v^2}{v+v^3}\,dv = \int \left\{\frac{1}{v} - \frac{2v}{1+v^2}\right\} dv = \ln|v| - \ln[v^2+1]$$

Hence the solution of the separable equation becomes

$$\ln|v| - \ln[v^2+1] = \ln|x| + \ln C$$

which is equivalent to

$$\frac{v}{v^2+1} = C\,x$$

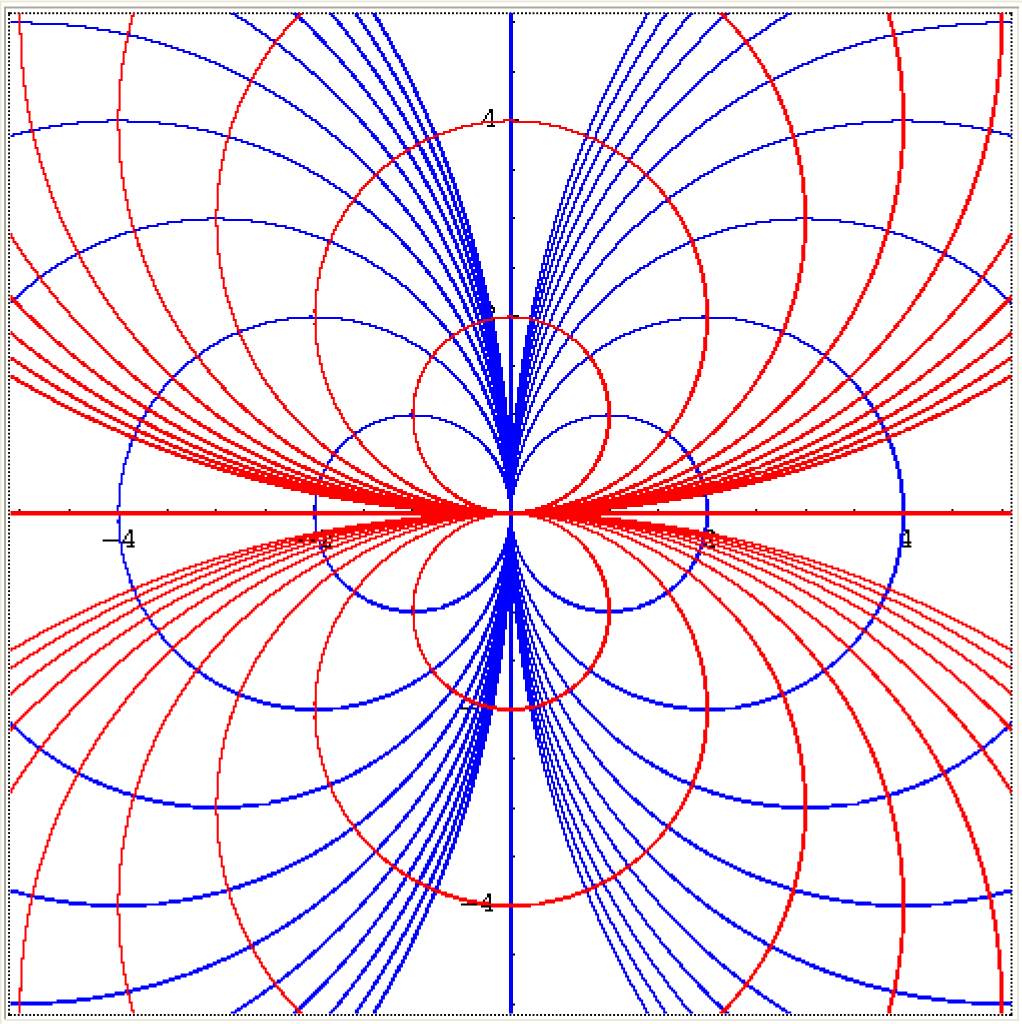
where $C \neq 0$. Hence all the solutions are

$$\begin{cases} v = 0 \\ \dfrac{v}{v^2+1} = Cx \end{cases}$$

We go back to y to get $y = 0$ and $\dfrac{y}{y^2+x^2} = C$ which is equivalent to

$$\begin{cases} y = 0 \\ x^2+y^2 = my \end{cases}$$

5. Which is x-axis and a family of circles centered on y-axis. A geometrical view of both the families is shown in the next slide.



Population Dynamics

Some natural questions related to population problems are the following:

- What will the population of a certain country after e.g. ten years?
- How are we protecting the resources from extinction?

The easiest population dynamics model is the **exponential model.** This model is based on the assumption:

The rate of change of the population is proportional to the existing population.

If $P(t)$ measures the population of a species at any time t then because of the above mentioned assumption we can write

$$\frac{dP}{dt} = kP$$

where the rate k is constant of proportionality. Clearly the above equation is linear as well as separable. To solve this equation we multiply the equation with the integrating factor e^{-kt} to obtain

$$\frac{d}{dt}\left[P\,e^{-kt}\right] = 0$$

Integrating both sides we obtain

$$P\,e^{-kt} = C \quad \text{or} \quad P = C\,e^{kt}$$

If P_0 is the initial population then $P(0) = P_0$. So that $C = P_0$ and obtain

$$P(t) = P_0\,e^{kt}$$

Clearly, we must have $k > 0$ for growth and $k < 0$ for the decay.

Illustration

Example:

The population of a certain community is known to increase at a rate proportional to the number of people present at any time. The population has doubled in 5 years, how long would it take to triple?. If it is known that the population of the community is 10,000 after 3 years. What was the initial population? What will be the population in 30 years?

Solution:

Suppose that P_0 is initial population of the community and $P(t)$ the population at any time t then the population growth is governed by the differential equation

$$\frac{dP}{dt} = kP$$

As we know solution of the differential equation is given by

$$P(t) = P_0\,e^{kt}$$

Since $P(5) = 2P_0$. Therefore, from the last equation we have

$$2P_0 = P_0\,e^{5k} \Rightarrow e^{5k} = 2$$

This means that

$5k = \ln 2 = 0.69315$ or $k = \frac{0.69315}{5} = 0.13863$

Therefore, the solution of the equation becomes

$$P(t) = P_0\, e^{0.13863t}$$

If t_1 is the time taken for the population to triple then

$$3P_0 = P_0\, e^{0.1386t_1} \Rightarrow e^{0.1386t_1} = 3$$

$$t_1 = \frac{\ln 3}{0.1386} = 7.9265 \approx 8 \text{ years}$$

Now using the information $P(3) = 10{,}000$, we obtain from the solution that

$$10{,}000 = P_0\, e^{(0.13863)(3)} \Rightarrow P_0 = \frac{10{,}000}{e^{0.41589}}$$

Therefore, the initial population of the community was

$$P_0 \approx 6598$$

Hence solution of the model is

$$P(t) = 6598e^{0.13863t}$$

So that the population in 30 years is given by

$$P(30) = 6598e^{(30)(0.13863)} = 6598e^{4.1589}$$

or $P(30) = (6598)(64.0011)$

or $P(30) \approx 422279$

Lecture 11 Radioactive Decay

In physics a radioactive substance disintegrates or transmutes into the atoms of another element. Many radioactive materials disintegrate at a rate proportional to the amount present. Therefore, if $A(t)$ is the amount of a radioactive substance present at time t, then the rate of change of $A(t)$ with respect to time t is given by

$$\frac{dA}{dt} = kA$$

where k is a constant of proportionality. Let the initial amount of the material be A_0 then $A(0) = A_0$. As discussed in the population growth model the solution of the differential equation is

$$A(t) = A_0 e^{kt}$$

The constant k can be determined using half-life of the radioactive material.

The half-life of a radioactive substance is the time it takes for one-half of the atoms in an initial amount A_0 to disintegrate or transmute into atoms of another element. The half-life measures stability of a radioactive substance. The longer the half-life of a substance, the more stable it is. If T denotes the half-life then

$$A(T) = \frac{A_0}{2}$$

Therefore, using this condition and the solution of the model we obtain

$$\frac{A_0}{2} = A_0 e^{kt}$$

So that $$kT = -\ln 2$$

Therefore, if we know T, we can get k and vice-versa. The half-life of some important radioactive materials is given in many textbooks of Physics and Chemistry. For example the half-life of $C-14$ is 5568 ± 30 years.

Example 1:
A radioactive isotope has a half-life of 16 days. We have 30 g at the end of 30 days. How much radioisotope was initially present?

Solution: Let $A(t)$ be the amount present at time t and A_0 the initial amount of the isotope. Then we have to solve the initial value problem.

$$\frac{dA}{dt} = kA, \quad A(30) = 30$$

We know that the solution of the IVP is given by

$$A(t) = A_0 e^{kt}$$

If T the half-life then the constant is given k by

$$kT = -\ln 2 \quad \text{or} \quad k = -\frac{\ln 2}{T} = -\frac{\ln 2}{16}$$

Now using the condition $A(30) = 30$, we have

$$30 = A_0 e^{30k}$$

So that the initial amount is given by

$$A_0 = 30e^{-30k} = 30e^{\frac{30\ln 2}{16}} = 110.04\text{ g}$$

Example 2:

A breeder reactor converts the relatively stable uranium 238 into the isotope plutonium 239. After 15 years it is determined that 0.043% of the initial amount A_0 of the plutonium has disintegrated. Find the half-life of this isotope if the rate of disintegration is proportional to the amount remaining.

Solution:

Let $A(t)$ denotes the amount remaining at any time t, then we need to find solution to the initial value problem

$$\frac{dA}{dt} = kA, \qquad A(0) = A_0$$

which we know is given by

$$A(t) = A_0 e^{kt}$$

If 0.043% disintegration of the atoms of A_0 means that 99.957% of the substance remains. Further 99.957% of A_0 equals $(0.99957)A_0$. So that

$$A(15) = (0.99957)A_0$$

So that

$$A_0 e^{15k} = (0.99957)A_0$$

$$15k = \ln(0.99957)$$

Or $$k = \frac{\ln(0.99957)}{15} = -0.00002867$$

Hence $$A(t) = A_0 e^{-0.00002867\,t}$$

If T denotes the half-life then $A(T) = \frac{A_0}{2}$. Thus

$$\frac{A_0}{2} = A_0 e^{-0.00002867\ T} \quad \text{or} \quad \frac{1}{2} = e^{-0.00002867\ T}$$

$$-0.00002867\ \mathrm{T} = \ln\left(\frac{1}{2}\right) = -\ln 2$$

$$T = \frac{\ln 2}{0.00002867} \approx 24{,}180 \ \text{ years}$$

Newton's Law of Cooling

From experimental observations it is known that the temperature ***T (t)*** of an object changes at a rate proportional to the difference between the temperature in the body and the temperature T_m of the surrounding environment. This is what is known as **Newton's law of cooling**.

If initial temperature of the cooling body is T_0 then we obtain the initial value problem

$$\frac{dT}{dt} = k(T - T_m), \quad T(0) = T_0$$

where k is constant of proportionality. The differential equation in the problem is linear as well as separable.

Separating the variables and integrating we obtain

$$\int \frac{dT}{T - T_m} = \int k\, dt$$

This means that

$$\ln|T - T_m| = kt + C$$

$$T - T_m = e^{kt+C}$$

$$T(t) = T_m + C_1 e^{kt} \quad \text{where} \quad C_1 = e^C$$

Now applying the initial condition $T(0) = T_0$, we see that $C_1 = T_0 - T_m$. Thus the solution of the initial value problem is given by

$$T(t) = T_m + (T_0 - T_m)e^{kt}$$

Hence, If temperatures at times t_1 and t_2 are known then we have

$$T(t_1) - T_m = (T_0 - T_m)e^{kt_1}, \; T(t_2) - T_m = (T_0 - T_m)e^{kt_2}$$

So that we can write

$$\frac{T(t_1) - T_m}{T(t_2) - T_m} = e^{k(t_1 - t_2)}$$

This equation provides the value of k if the interval of time '$t_1 - t_2$' is known and vice-versa.

Example 3:

Suppose that a dead body was discovered at midnight in a room when its temperature was $80^o F$. The temperature of the room is kept constant at $60^o F$. Two hours later the temperature of the body dropped to $75^o F$. Find the time of death.

Solution:

Assume that the dead person was not sick, then

$$T(0) = 98.6^o F = T_0 \text{ and } T_m = 60^o F$$

Therefore, we have to solve the initial value problem

$$\frac{dT}{dt} = k(T - 60), \quad T(0) = 98.6$$

We know that the solution of the initial value problem is

$$T(t) = T_m + (T_0 - T_m)e^{kt}$$

So that

$$\frac{T(t_1) - T_m}{T(t_2) - T_m} = e^{k(t_1 - t_2)}$$

The observed temperatures of the cooling object, i.e. the dead body, are

$$T(t_1) = 80^o F \quad \text{and} \quad T(t_2) = 75^o F$$

Substituting these values we obtain

$$\frac{80 - 60}{75 - 60} = e^{2k} \quad \text{as} \quad t_1 - t_2 = 2 \quad \text{hours}$$

So

$$k = \frac{1}{2} \ln \frac{4}{3} = 0.1438$$

Now suppose that t_1 and t_2 denote the times of death and discovery of the dead body then

$$T(t_1) = T(0) = 98.6^o F \quad \text{and} \quad T(t_2) = 80^o F$$

For the time of death, we need to determine the interval $t_1 - t_2 = t_d$. Now

$$\frac{T(t_1) - T_m}{T(t_2) - T_m} = e^{k(t_1 - t_2)} \Rightarrow \frac{98.6 - 60}{80 - 60} = e^{kt_d}$$

or

$$t_d = \frac{1}{k} \ln \frac{38.6}{20} \approx 4.573$$

Hence the time of death is 7:42 PM.

Carbon Dating

- The isotope *C–14* is produced in the atmosphere by the action of cosmic radiation on nitrogen.
- The ratio of C-14 to ordinary carbon in the atmosphere appears to be constant.
- The proportionate amount of the isotope in all living organisms is same as that in the atmosphere.
- When an organism dies, the absorption of $C-14$ by breathing or eating ceases.
- Thus comparison of the proportionate amount of $C-14$ present, say, in a fossil with constant ratio found in the atmosphere provides a reasonable estimate of its age.
- The method has been used to date wooden furniture in Egyptian tombs.
- Since the method is based on the knowledge of half-life of the radio active $C-14$ (5600 years approximately), the initial value problem discussed in the radioactivity model governs this analysis.

Example:

A fossilized bone is found to contain $1/1000$ of the original amount of *C–14*. Determine the age of the fissile.

Solution:

Let $A(t)$ be the amount present at any time $\boldsymbol{t}$ and $\boldsymbol{A_0}$ the original amount of C–14. Therefore, the process is governed by the initial value problem.

$$\frac{dA}{dt} = kA, \quad A(0) = A_0$$

We know that the solution of the problem is

$$A(t) = A_0 e^{kt}$$

Since the half life of the carbon isotope is 5600 years. Therefore,

$$A(5600) = \frac{A_0}{2}$$

So that
$$\frac{A_0}{2} = A_0 e^{5600k} \quad \text{or} \quad 5600k = -\ln 2$$

$$k = -0.00012378$$

Hence

$$A(t) = A_0 e^{-(0.00012378)t}$$

If t denotes the time when fossilized bone was found then $A(t) = \frac{A_0}{1000}$

$$\frac{A_0}{1000} = A_0 e^{-(0.00012378)t} \Rightarrow -0.00012378t = -\ln 1000$$

Therefore

$$t = \frac{\ln 1000}{0.00012378} = 55{,}800 \text{ years}$$

Lecture 12 Application of Non Linear Equations

As we know that the solution of the exponential model for the population growth is

$$P(t) = P_0\, e^{kt}$$

P_0 being the initial population. From this solution we conclude that

(a) If $k > 0$ the population grows and expand to infinity i.e. $\lim_{t\to\infty} P(t) = +\infty$

(b) If $k < 0$ the population will shrink to approach 0, which means extinction.

Note that:
(1) The prediction in the first case ($k > 0$) differs substantially from what is actually observed, population growth is eventually limited by some factor**!**
(2) Detrimental effects on the environment such as pollution and excessive and competitive demands for food and fuel etc. can have inhibitive effects on the population growth.

Logistic equation:

Another model was proposed to remedy this flaw in the exponential model. This is called the **logistic model** (also called **Verhulst-Pearl model**).

Suppose that $a > 0$ is constant average rate of birth and that the death rate is proportional to the population $P(t)$ at any time t. Thus if $\frac{1}{P}\frac{dP}{dt}$ is the rate of growth per individual then

$$\frac{1}{P}\frac{dP}{dt} = a - bP \quad \text{or} \quad \frac{dP}{dt} = P(a - bP)$$

where b is constant of proportionality. The term $-bP^2,\ b > 0$ can be interpreted as inhibition term. When $b = 0$, the equation reduces to the one in exponential model.
Solution to the logistic equation is also very important in ecological, sociological and even in managerial sciences.

Solution of the Logistic equation:

The logistic equation

$$\frac{dP}{dt} = P(a - bP)$$

can be easily identified as a **nonlinear** equation that is separable. The constant solutions of the equation are given by

$$P(a - bP) = 0$$

$$\Rightarrow \ \mathrm{P} = 0 \quad \text{and} \quad P = \frac{a}{b}$$

For non-constant solutions we separate the variables

$$\frac{dP}{P(a-bP)} = dt$$

Resolving into partial fractions we have

$$\left[\frac{1/a}{P} + \frac{b/a}{a-bP}\right]dP = dt$$

Integrating $$\frac{1}{a}\ln|P| - \frac{1}{a}\ln|a-bP| = t + C$$

$$\ln\left|\frac{P}{a-bP}\right| = at + aC$$

or $$\frac{P}{a-bP} = C_1 e^{at} \quad \text{where } C_1 = e^{aC}$$

Easy algebraic manipulations give

$$P(t) = \frac{aC_1 e^{at}}{1 + bC_1 e^{at}} = \frac{aC_1}{bC_1 + e^{-at}}$$

Here C_1 is an arbitrary constant. If we are given the initial condition $P(0) = P_0$, $P_0 \neq \frac{a}{b}$ we obtain $C_1 = \frac{P_0}{a - bP_0}$. Substituting this value in the last equation and simplifying, we obtain

$$P(t) = \frac{aP_0}{bP_0 + (a - bP_0)e^{-at}}$$

Clearly $$\lim_{t\to\infty} P(t) = \frac{aP_0}{bP_0} = \frac{a}{b},$$ **limited growth**

Note that $P = \frac{a}{b}$ is a **singular solution** of the logistic equation.

Special Cases of Logistic Equation:

1. Epidemic Spread

Suppose that one person infected from a contagious disease is introduced in a fixed population of n people.

The natural assumption is that the rate $\frac{dx}{dt}$ of spread of disease is proportional to the number $x(t)$ of the infected people and number $y(t)$ of people not infected people. Then

$$\frac{dx}{dt} = kxy$$

Since $$x + y = n + 1$$

Therefore, we have the following initial value problem

$$\frac{dx}{dt} = kx(n+1-x), \quad x(0) = 1$$

The last equation is a **special case of the logistic equation** and has also been used for the **spread of information** and the **impact of advertising** in centers of population.

2. A Modification of LE:

A modification of the nonlinear logistic differential equation is the following

$$\frac{dP}{dt} = P(a - b\ln P)$$

has been used in the studies of solid tumors, in actuarial predictions, and in the growth of revenue from the sale of a commercial product in addition to growth or decline of population.

Example:

Suppose a student carrying a flu virus returns to an isolated college campus of 1000 students. If it is assumed that the rate at which the virus spreads is proportional not only to the number x of infected students but also to the number of students not infected, determine the number of infected students after 6 days if it is further observed that after 4 days $x(4) = 50$.

Solution

Assume that no one leaves the campus throughout the duration of the disease. We must solve the initial-value problem

$$\frac{dx}{dt} = kx(1000 - x), \quad x(0) = 1 \quad .$$

We identify

$$a = 1000k \quad \text{and} \quad \mathrm{b = k}$$

Since the solution of logistic equation is

$$P(t) = \frac{aP_0}{bP_0 + (a - bP_0)e^{-at}}$$

Therefore we have

$$x(t) = \frac{1000\,k}{k + 999\,ke^{-1000\,kt}} = \frac{1000}{1 + 999\,e^{-1000\,kt}}.$$

Now, using x(4)= 50, we determine *k*

$$50 = \frac{1000}{1 + 999\,e^{-4000\,k}}$$

We find

$$k = \frac{-1}{4000}\ln\frac{19}{999} = 0.0009906.$$

Thus

$$x(t) = \frac{1000}{1 + 999\,e^{-0.9906\ t}}$$

Finally

$$x(6) = \frac{1000}{1 + 999\,e^{-5.9436}} = 276\ \text{ students}\ .$$

Chemical reactions:

In a first order chemical reaction, the molecules of a substance *A* decompose into smaller molecules. This decomposition takes place at a rate proportional to the amount of the first substance that has not undergone conversion. The disintegration of a radioactive substance is an example of the first order reaction. If X is the remaining amount of the substance A at any time t then

$$\frac{dX}{dt} = k\ X$$

$k < 0$ because X is decreasing.

In a 2nd order reaction two chemicals A and B react to form another chemical C at a rate proportional to the product of the remaining concentrations of the two chemicals.

If X denotes the amount of the chemical C that has formed at time t. Then the instantaneous amounts of the first two chemicals A and B not converted to the chemical C are $\alpha - X$ and $\beta - X$, respectively. Hence the rate of formation of chemical C is given by

$$\frac{dX}{dt} = k\left(\alpha - X\right)\left(\beta - X\right)$$

where k is constant of proportionality.

Example:

A compound $\boldsymbol{C}$ is formed when two chemicals A and B are combined. The resulting reaction between the two chemicals is such that for each gram of A, 4 grams of B are used. It is observed that 30 grams of the compound C are formed in 10 minutes. Determine the amount of C at any time if the rate of reaction is proportional to the amounts of A and B remaining and if initially there are 50 grams of A and 32 grams of B. How much of the compound C is present at 15 minutes? Interpret the solution as $t \to \infty$

Solution:

If $X(t)$ denote the number of grams of chemical $\boldsymbol{C}$ present at any time t. Then

$$X(0) = 0 \text{ and } X(10) = 30$$

Suppose that there are 2 grams of the compound $\boldsymbol{C}$ and we have used a grams of A and b grams of B then

$$a + b = 2 \quad \text{and} \quad b = 4a$$

Solving the two equations we have

$$a = \frac{2}{5} = 2\,(1/5) \quad \text{and} \quad b = \frac{8}{5} = 2\,(4/5)$$

In general, if there were for X grams of C then we must have

$$a = \frac{X}{5} \quad \text{and} \quad b = \frac{4}{5}X$$

Therefore the amounts of A and B remaining at any time t are then

$$50 - \frac{X}{5} \text{ and } 32 - \frac{4}{5}X$$

respectively .

Therefore, the rate at which chemical C is formed satisfies the differential equation

$$\frac{dX}{dt} = \lambda\left(50 - \frac{X}{5}\right)\left(32 - \frac{4}{5}X\right)$$

or

$$\frac{dX}{dt} = k(250 - X)(40 - X), \quad k = 4\lambda/25$$

We now solve this differential equation.

By separation of variables and partial fraction, we can write

$$\frac{dX}{(250 - X)(40 - X)} = kdt$$

$$-\frac{1/210}{250 - X}dX + \frac{1/210}{40 - X}dX = kdt$$

$$\ln\left|\frac{250 - X}{40 - X}\right| = 210kt + c_1$$

$$\frac{250 - X}{40 - X} = c_2 e^{210kt} \quad Where \ c_2 = e^{c_1}$$

When $t = 0$, $X = 0$, so it follows at this point that $c_2 = 25/4$. Using $X = 30$ at $t = 10$, we find

$$210k = \frac{1}{10}\ln\frac{88}{25} = 0.1258$$

With this information we solve for X :

$$X(t) = 1000\left(\frac{1 - e^{-0.1258t}}{25 - 4e^{-0.1258t}}\right)$$

It is clear that as $e^{-0.1258t} \to 0$ as $t \to \infty$. Therefore $X \to 40$ as $t \to \infty$. This fact can also be verified from the following table that $X \to 40$ as $t \to \infty$.

t	10	15	20	25	30	35
X	30	34.78	37.25	38.54	39.22	39.59

This means that there are 40 grams of compound C formed, leaving

$$50 - \frac{1}{5}(40) = 42 \quad \text{grams of chemical } A$$

and $$32 - \frac{4}{5}(40) = 0 \quad \text{grams of chemical } B$$

Miscellaneous Applications

- The velocity v of a falling mass m, subjected to air resistance proportional to instantaneous velocity, is given by the differential equation

$$m\frac{dv}{dx} = mg - kv$$

Here $k > 0$ is constant of proportionality.

- The rate at which a drug disseminates into bloodstream is governed by the differential equation

$$\frac{dx}{dt} = A - Bx$$

Here A, B are positive constants and $x(t)$ describes the concentration of drug in the bloodstream at any time t.

- The rate of memorization of a subject is given by

$$\frac{dA}{dt} = k_1(M - A) - k_2 A$$

Here $k_1 > 0,\ k_2 > 0$ and $A(t)$ is the amount of material memorized in time t, M is the total amount to be memorized and $M - A$ is the amount remaining to be memorized.

Lecture 13 Higher Order Linear Differential Equations

Preliminary theory

- A differential equation of the form
$$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=g(x)$$
or $$a_n(x)y^{(n)}+a_{n-1}(x)y^{(n-1)}+\cdots+a_1(x)y'+a_0(x)y=g(x)$$
where $a_0(x), a_1(x), \ldots, a_n(x), g(x)$ are functions of x and $a_n(x)\neq 0$, is called a linear differential equation with variable coefficients.
- However, we shall first study the differential equations with constant coefficients i.e. equations of the type
$$a_n\frac{d^n y}{dx^n}+a_{n-1}\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1\frac{dy}{dx}+a_0 y=g(x)$$
where $a_0, a_1, \ldots, a_n$ are real constants. This equation is non-homogeneous differential equation and
- If $g(x)=0$ then the differential equation becomes
$$a_n\frac{d^n y}{dx^n}+a_{n-1}\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1\frac{dy}{dx}+a_0 y=0$$
which is known as the **associated homogeneous differential equation**.

Initial -Value Problem

For a linear nth-order differential equation, the problem:

Solve: $$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=g(x)$$

Subject to: $y(x_0)=y_0,\quad y^{/}(x_0)=y_0^{/},\ldots y^{n-1}(x_0)=y_0^{n-1}$

$y_0, y_0^{/}, \ldots, y_0^{n-1}$ being arbitrary constants, is called an **initial-value problem** (IVP).

The specified values $y(x_0)=y_0$, $y^{/}(x_0)=y_0^{/},\ldots,y^{n-1}(x_0)=y_0^{n-1}$ are called initial-conditions.

For $n=2$ the initial-value problem reduces to

Solve: $$a_2(x)\frac{d^2 y}{dx^2}+a_1(x)\frac{dy}{dx}+a_0(x)y=g(x)$$

Subject to: $y(x_0)=y_0, \ldots,\quad y^{/}(x_0)=y_0^{/}$

Solution of IVP

A function satisfying the differential equation on I whose graph passes through (x_0, y_0) such that the slope of the curve at the point is the number $y_0^{/}$ is called solution of the initial value problem.

Theorem: *Existence and Uniqueness of Solutions*

Let $a_n(x), a_{n-1}(x),..., a_1(x), a_0(x)$ and $g(x)$ be continuous on an interval I and let $a_n(x) \neq 0, \; \forall\, x \in I$. If $x = x_0 \in I$, then a solution $y(x)$ of the initial-value problem exist on I and is unique.

Example 1

Consider the function $y = 3e^{2x} + e^{-2x} - 3x$

This is a solution to the following initial value problem

$$y'' - 4y = 12x, \quad y(0) = 4, \; y'(0) = 1$$

Since $\dfrac{d^2y}{dx^2} = 12e^{2x} + 4e^{-2x}$

and $\dfrac{d^2y}{dx^2} - 4y = 12e^{2x} + 4e^{-2x} - 12e^{2x} - 4e^{-2x} + 12x = 12x$

Further $y(0) = 3 + 1 - 0 = 4$ and $y'(0) = 6 - 2 - 3 = 1$

Hence $y = 3e^{2x} + e^{-2x} - 3x$

is a solution of the initial value problem.

We observe that

- The equation is linear differential equation.
- The coefficients being constant are continuous.
- The function $g(x) = 12x$ being polynomial is continuous.
- The leading coefficient $a_2(x) = 1 \neq 0$ for all values of x.

Hence the function $y = 3e^{2x} + e^{-2x} - 3x$ is the unique solution.

Example 2

Consider the initial-value problem

$$3y''' + 5y'' - y' + 7y = 0,$$
$$y(1) = 0, \quad y'(1) = 0, \quad y''(1) = 0$$

Clearly the problem possesses the trivial solution $y = 0$.

Since

- The equation is homogeneous linear differential equation.
- The coefficients of the equation are constants.
- Being constant the coefficient are continuous.
- The leading coefficient $a_3 = 3 \neq 0$.

Hence $y = 0$ is the only solution of the initial value problem.

Note: If $a_n = 0$?

If $a_n(x) = 0$ in the differential equation

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = g(x)$$

for some $x \in I$ then

- Solution of initial-value problem may not be unique.
- Solution of initial-value problem may not even exist.

Example 4

Consider the function

$$y = cx^2 + x + 3$$

and the initial-value problem

$$x^2 y'' - 2xy' + 2y = 6$$

$$y(0) = 3, \qquad y'(0) = 1$$

Then $\quad y' = 2cx + 1 \quad$ and $\quad y'' = 2c$

Therefore $\quad x^2 y'' - 2xy' + 2y = x^2(2c) - 2x(2cx + 1) + 2(cx^2 + x + 3)$

$$= 2cx^2 - 4cx^2 - 2x + 2cx^2 + 2x + 6$$

$$= 6.$$

Also $\quad y(0) = 3 \Rightarrow c(0) + 0 + 3 = 3$

and $\quad y'(0) = 1 \Rightarrow 2c(0) + 1 = 1$

So that for any choice of c, the function 'y' satisfies the differential equation and the initial conditions. Hence the solution of the initial value problem is not unique.

Note that

- The equation is linear differential equation.
- The coefficients being polynomials are continuous everywhere.
- The function $g(x)$ being constant is constant everywhere.
- The leading coefficient $a_2(x) = x^2 = 0$ at $x = 0 \in (-\infty, \infty)$.

Hence $a_2(x) = 0$ brought non-uniqueness in the solution

Boundary-value problem (BVP)

For a 2nd order linear differential equation, the problem

Solve: $$a_2(x)\frac{d^2y}{dx^2}+a_1(x)\frac{dy}{dx}+a_0(x)y=g(x)$$

Subject to: $y(a)=y_0,\quad y(b)=y_1$

is called a ***boundary-value problem***. The specified values $y(a)=y_0$, and $y(b)=y_1$ are called ***boundary conditions*.**

Solution of BVP

A solution of the boundary value problem is a function satisfying the differential equation on some interval I, containing a and b, whose graph passes through two points (a, y_0) and (b, y_1).

Example 5

Consider the function

$$y=3x^2-6x+3$$

We can prove that this function is a solution of the boundary-value problem

$$x^2y''-2xy'+2y=6,$$
$$y(1)=0,\qquad y(2)=3$$

Since $\dfrac{dy}{dx}=6x-6,\ \dfrac{d^2y}{dx^2}=6$

Therefore $x^2\dfrac{d^2y}{dx^2}-2x\dfrac{dy}{dx}+2y=6x^2-12x^2+12x+6x^2-12x+6=6$

Also $y(1)=3-6+3=0,\quad y(2)=12-12+3=3$

Therefore, the function 'y' satisfies both the differential equation and the boundary conditions. Hence y is a solution of the boundary value problem.

.

Possible Boundary Conditions

For a 2nd order linear non-homogeneous differential equation

$$a_2(x)\frac{d^2y}{dx^2}+a_1(x)\frac{dy}{dx}+a_0(x)y=g(x)$$

all the possible pairs of boundary conditions are

$$y(a)=y_0,\qquad y(b)=y_1,$$
$$y'(a)=y'_0,\qquad y(b)=y_1,$$
$$y(a)=y_0,\qquad y'(b)=y'_1,$$
$$y'(a)=y'_0,\qquad y'(b)=y'_1$$

where y_0, y'_0, y_1 and y'_1 denote the arbitrary constants.

In General

All the four pairs of conditions mentioned above are just special cases of the general boundary conditions

$$\alpha_1 y(a) + \beta_1 y'(a) = \gamma_1$$
$$\alpha_2 y(b) + \beta_2 y'(b) = \gamma_2$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \{0,1\}$

Note that

A boundary value problem may have

- Several solutions.
- A unique solution, or
- No solution at all.

Example 1

Consider the function

$$y = c_1 \cos 4x + c_2 \sin 4x$$

and the boundary value problem

$$y'' + 16y = 0, \quad y(0) = 0, \quad y(\pi/2) = 0$$

Then

$$y' = -4c_1 \sin 4x + 4c_2 \cos 4x$$

$$y'' = -16(c_1 \cos 4x + c_2 \sin 4x)$$

$$y'' = -16y$$

$$y'' + 16y = 0$$

Therefore, the function

$$y = c_1 \cos 4x + c_2 \sin 4x$$

satisfies the differential equation

$$y'' + 16y = 0.$$

Now apply the boundary conditions

Applying $y(0) = 0$

We obtain

$$0 = c_1 \cos 0 + c_2 \sin 0$$
$$\Rightarrow c_1 = 0$$

So that

$$y = c_2 \sin 4x.$$

But when we apply the 2nd condition $y(\pi/2) = 0$, we have

$$0 = c_2 \sin 2\pi$$

Since $\sin 2\pi = 0$, the condition is satisfied for any choice of c_2, solution of the problem is the one-parameter family of functions

$$y = c_2 \sin 4x$$

Hence, there are an ***infinite number of solutions*** of the boundary value problem.

Example 2

Solve the boundary value problem

$$y'' + 16y = 0$$

$$y(0) = 0, \ y\left(\frac{\pi}{8}\right) = 0,$$

Solution:

As verified in the previous example that the function

$$y = c_1 \cos 4x + c_2 \sin 4x$$

satisfies the differential equation

$$y'' + 16y = 0$$

We now apply the boundary conditions

$$y(0) = 0 \Rightarrow 0 = c_1 + 0$$

and $y(\pi/8) = 0 \Rightarrow 0 = 0 + c_2$

So that $c_1 = 0 = c_2$

Hence

$$y = 0$$

is the only solution of the boundary-value problem.

Example 3

Solve the differential equation

$$y'' + 16y = 0$$

subject to the boundary conditions

$$y(0) = 0, \quad y(\pi/2) = 1$$

Solution:

As verified in an earlier example that the function

$$y = c_1 \cos 4x + c_2 \sin 4x$$

satisfies the differential equation

$$y'' + 16y = 0$$

We now apply the boundary conditions

$$y(0) = 0 \Rightarrow 0 = c_1 + 0$$

Therefore $c_1 = 0$

So that $y = c_2 \sin 4x$

However $y(\pi/2) = 1 \Rightarrow c_2 \sin 2\pi = 1$

or $1 = c_2 .0 \Rightarrow 1 = 0$

This is a clear contradiction. Therefore, the boundary value problem has ***no solution***.

Definition: **Linear Dependence**

A set of functions

$$\{f_1(x), f_2(x), \dots, f_n(x)\}$$

is said to be ***linearly dependent*** on an interval I if $\exists$ constants $c_1, c_2, \dots, c_n$ not all zero, such that

$$c_1 f_1(x) + c_2 f_2(x) + \cdots . + c_n f_n(x) = 0, \quad \forall\, x \in I$$

Definition: Linear Independence

A set of functions

$$\{f_1(x), f_2(x), \dots, f_n(x)\}$$

is said to be linearly independent on an interval I if

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0, \quad \forall\, x \in I,$$

only when

$$c_1 = c_2 = \cdots = c_n = 0.$$

Case of two functions:

If $n = 2$ then the set of functions becomes

$$\{f_1(x), f_2(x)\}$$

If we suppose that

$$c_1 f_1(x) + c_2 f_2(x) = 0$$

Also that the functions are linearly dependent on an interval I then either $c_1 \neq 0$ or $c_2 \neq 0$.

Let us assume that $c_1 \neq 0$, then

$$f_1(x) = -\frac{c_2}{c_1} f_2(x);$$

Hence $f_1(x)$ is the constant multiple of $f_2(x)$.

Conversely, if we suppose

$$f_1(x) = \mathrm{c}_2\ f_2(x)$$

Then $$(-1) f_1(x) + c_2 f_2(x) = 0, \ \forall x \in I$$

So that the functions are linearly dependent because $c_1 = -1$.

Hence, we conclude that:

- Any two functions $f_1(x)$ and $f_2(x)$ are linearly dependent on an interval *I* if and only if one is the constant multiple of the other.
- Any two functions are linearly independent when neither is a constant multiple of the other on an interval *I*.
- In general a set of n functions $\{f_1(x), f_2(x), \ldots, f_n(x)\}$ is linearly dependent if at least one of them can be expressed as a linear combination of the remaining.

Example 1

The functions

$$f_1(x) = \sin 2x, \qquad \forall x \in (-\infty, \infty)$$

$$f_2(x) = \sin x \cos x, \quad \forall x \in (-\infty, \infty)$$

If we choose $c_1 = \frac{1}{2}$ and $c_2 = -1$ then

$$c_1 \sin 2x + c_2 \sin x \cos x = \frac{1}{2}(2\sin x \cos x) - \sin x \cos x = 0$$

Hence, the two functions $f_1(x)$ and $f_2(x)$ are linearly dependent.

Example 3

Consider the functions

$$f_1(x) = \cos^2 x, \; f_2(x) = \sin^2 x, \quad \forall x \in (-\pi/2, \pi/2),$$

$$f_3(x) = \sec^2 x, \; f_4(x) = \tan^2 x, \quad \forall x \in (-\pi/2, \pi/2)$$

If we choose $c_1 = c_2 = 1, c_3 = -1, c_4 = 1$, then

$$\begin{aligned} & c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) + c_4 f_4(x) \\ &= c_1 \cos^2 x + c_2 \sin^2 x + c_3 \sec^2 x + c_4 \tan^2 x \\ &= \cos^2 x + \sin^2 x + -1 - \tan^2 x + \tan^2 x \\ &= 1 - 1 + 0 = 0 \end{aligned}$$

Therefore, the given functions are linearly dependent.

Note that

The function $f_3(x)$ can be written as a linear combination of other three functions $f_1(x), f_2(x)$ and $f_4(x)$ because $\sec^2 x = \cos^2 x + \sin^2 x + \tan^2 x$.

Example 3

Consider the functions

$$f_1(x) = 1 + x, \ \forall\, x \in (-\infty, \infty)$$
$$f_2(x) = x, \quad \forall\, x \in (-\infty, \infty)$$
$$f_3(x) = x^2, \quad \forall\, x \in (-\infty, \infty)$$

Then

$$c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) = 0$$

means that

$$c_1(1 + x) + c_2 x + c_3 x^2 \qquad = 0$$

or $\qquad c_1 + (c_1 + c_2)x + c_3 x^2 \qquad = 0$

Equating coefficients of x and x^2 constant terms we obtain

$$c_1 = 0 = c_3$$
$$c_1 + c_2 = 0$$

Therefore $\qquad c_1 = c_2 = c_3 = 0$

Hence, the three functions $f_1(x), f_2(x)$ and $f_3(x)$ are linearly independent.

Definition: Wronskian

Suppose that the function $f_1(x), f_2(x), \ldots, f_n(x)$ possesses at least $n-1$ derivatives then the determinant

$$\begin{vmatrix} f_1 & f_2 & \cdots\cdots\cdots & f_n \\ f_1^{/} & f_2^{/} & \cdots\cdots\cdots & f_n^{/} \\ \vdots & \vdots & & \vdots \\ f_1^{n-1} & f_2^{n-1} & \cdots\cdots\cdots & f_n^{n-1} \end{vmatrix}$$

is called Wronskian of the functions $f_1(x), f_2(x), \ldots, f_n(x)$ and is denoted by $W\big(f_1(x), f_1(x), \ldots, f_1(x)\big)$.

Theorem: Criterion for Linearly Independent Functions

Suppose the functions $f_1(x), f_2(x), \ldots, f_n(x)$ possess at least *n-1* derivatives on an interval I. If

$$W(f_1(x), f_2(x), \ldots, f_n(x)) \neq 0$$

for at least one point in I, then functions $f_1(x), f_2(x), \ldots, f_n(x)$ are linearly independent on the interval I.

Note that

This is only a sufficient condition for linear independence of a set of functions.

In other words

If $f_1(x), f_2(x), \ldots, f_n(x)$ possesses at least $n-1$ derivatives on an interval and are linearly dependent on I, then

$$W(f_1(x),\ f_2(x),\ldots,\ f_n(x)) = 0, \quad \forall x \in I$$

However, the converse is not true. i.e. a Vanishing Wronskian does not guarantee linear dependence of functions.

Example 1

The functions

$$f_1(x) = \sin^2 x$$
$$f_2(x) = 1 - \cos 2x$$

are linearly dependent because

$$\sin^2 x = \frac{1}{2}(1 - \cos 2x)$$

We observe that for all $x \in (-\infty, \infty)$

$$W\left(f_1(x),\ f_2(x)\right) = \begin{vmatrix} \sin^2 x & 1-\cos 2x \\ 2\sin x\cos x & 2\sin 2x \end{vmatrix}$$

$$= 2\sin^2 x \sin 2x - 2\sin x\cos x$$
$$+ 2\sin x\cos x\cos 2x$$
$$= \sin 2x[2\sin^2 x - 1 + \cos 2x]$$
$$= \sin 2x[2\sin^2 x - 1 + \cos^2 x - \sin^2 x]$$
$$= \sin 2x[\sin^2 x + \cos^2 x - 1]$$
$$= 0$$

Example 2

Consider the functions

$$f_1(x) = e^{m_1 x},\ f_2(x) = e^{m_2 x}, \quad m_1 \neq m_2$$

The functions are linearly independent because

$$c_1 f_1(x) + c_2 f_2(x) = 0$$

if and only if $\quad c_1 = 0 = c_2 \quad$ as $\quad m_1 \neq m_2$

Now for all $x \in R$

$$W\left(e^{m_1x}, e^{m_2x}\right) = \begin{vmatrix} e^{m_1x} & e^{m_2x} \\ m_1e^{m_1x} & m_2e^{m_2x} \end{vmatrix}$$
$$= (m_2 - m_1)e^{(m_1+m_2)x}$$
$$\neq 0$$

Thus f_1 and f_2 are linearly independent of any interval on x-axis.

Example 3

If α and β are real numbers, $\beta \neq 0$, then the functions

$$y_1 = e^{\alpha x}\cos\beta x \text{ and } y_2 = e^{\alpha x}\sin\beta x$$

are linearly independent on any interval of the x-axis because

$$W\left(e^{\alpha x}\cos\beta x, e^{\alpha x}\sin\beta x\right)$$

$$= \begin{vmatrix} e^{\alpha x}\cos\beta x & e^{\alpha x}\sin\beta x \\ -\beta e^{\alpha x}\sin\beta x + \alpha e^{\alpha x}\cos\beta x & \beta e^{\alpha x}\cos\beta x + \alpha e^{\alpha x}\sin\beta x \end{vmatrix}$$

$$= \beta e^{2\alpha x}\left(\cos^2\beta x + \sin^2\beta x\right) = \beta e^{2\alpha x} \neq 0.$$

Example 4

The functions

$$f_1(x) = e^x,\ f_2(x) = xe^x, \text{ and } f_3(x) = x^2e^x$$

are linearly independent on any interval of the x-axis because for all $x \in R$, we have

$$W\left(e^x, xe^x, x^2e^x\right) = \begin{vmatrix} e^x & xe^x & x^2e^x \\ e^x & xe^x + e^x & x^2e^x + 2xe^x \\ e^x & xe^x + 2e^x & x^2e^x + 4xe^x + 2e^x \end{vmatrix}$$

$$= 2e^{3x} \neq 0$$

Exercise

1. Given that

$$y = c_1 e^x + c_2 e^{-x}$$

is a two-parameter family of solutions of the differential equation

$$y'' - y = 0$$

on $(-\infty, \infty)$, find a member of the family satisfying the boundary conditions

$$y(0) = 0, \quad y'(1) = 1.$$

2. Given that

$$y = c_1 + c_2 \cos x + c_3 \sin x$$

is a three-parameter family of solutions of the differential equation

$$y''' + y' = 0$$

on the interval $(-\infty, \infty)$, find a member of the family satisfying the initial conditions $y(\pi) = 0,\ y'(\pi) = 2, y''(\pi) = -1$.

3. Given that

$$y = c_1 x + c_2 x \ln x$$

is a two-parameter family of solutions of the differential equation $x^2 y'' - xy' + y = 0$ on $(-\infty, \infty)$. Find a member of the family satisfying the initial conditions

$$y(1) = 3,\ y'(1) = -1.$$

Determine whether the functions in problems 4-7 are linearly independent or dependent on $(-\infty, \infty)$.

4. $f_1(x) = x, \quad f_2(x) = x^2, \quad f_3(x) = 4x - 3x^2$
5. $f_1(x) = 0, \quad f_2(x) = x, \ f_3(x) = e^x$
6. $f_1(x) = \cos 2x, \quad f_2(x) = 1, \quad f_3(x) = \cos^2 x$
7. $f_1(x) = e^x, \quad f_2(x) = e^{-x}, \quad f_3(x) = \sinh x$

Show by computing the Wronskian that the given functions are linearly independent on the indicated interval.

8. $\tan x,\ \cot x; \quad (-\infty, \infty)$
9. $e^x,\ e^{-x},\ e^{4x}; \quad (-\infty, \infty)$
10. $x, x\ln x, x^2 \ln x; \quad (0, \infty)$

Lecture 14 Solutions of Higher Order Linear Equations

Preliminary Theory

- In order to solve an *n*th order non-homogeneous linear differential equation
$$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=g(x)$$
we first solve the associated homogeneous differential equation
$$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=0$$

- Therefore, we first concentrate upon the preliminary theory and the methods of solving the homogeneous linear differential equation.

- We recall that a function $y=f(x)$ that satisfies the associated homogeneous equation
$$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=0$$
is called solution of the differential equation.

Superposition Principle

Suppose that $y_1, y_2, \ldots, y_n$ are solutions on an interval I of the homogeneous linear differential equation

$$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=0$$

Then

$$y=c_1y_1(x)+c_2y_2(x)+\cdots+c_ny_n(x),$$

$c_1, c_2, \ldots, c_n$ being arbitrary constants is also a solution of the differential equation.

Note that

- A constant multiple $y=c_1y_1(x)$ of a solution $y_1(x)$ of the homogeneous linear differential equation is also a solution of the equation.
- The homogeneous linear differential equations always possess the trivial solution $y=0$.
- The superposition principle is a property of linear differential equations and it does not hold in case of non-linear differential equations.

Example 1

The functions

$$y_1 = e^x, y_2 = c^{2x}, \text{ and } y_3 = e^{3x}$$

all satisfy the homogeneous differential equation

$$\frac{d^3y}{dx^3} - 6\frac{d^2y}{dx^2} + 11\frac{dy}{dx} - 6y = 0$$

on $(-\infty, \infty)$. Thus y_1, y_2 and y_3 are all solutions of the differential equation

Now suppose that

$$y = c_1 e^x + c_2 e^{2x} + c_3 e^{3x}.$$

Then

$$\frac{dy}{dx} = c_1 e^x + 2c_2 e^{2x} + 3c_3 e^{3x}.$$

$$\frac{d^2y}{dx^2} = c_1 e^x + 4c_2 e^{2x} + 9c_3 e^{3x}.$$

$$\frac{d^3y}{dx^3} = c_1 e^x + 8c_2 e^{2x} + 27c_3 e^{3x}.$$

Therefore

$$\begin{aligned}
&\frac{d^3y}{dx^3} - 6\frac{d^2y}{dx^2} + 11\frac{dy}{dx} - 6y \\
&= c_1\left(e^x - 6e^x + 11e^x - 6e^x\right) + c_2\left(8e^{2x} - 24e^{2x} + 22e^{2x} - 6e^{2x}\right) \\
&\quad + c_3\left(27e^{3x} - 54e^{3x} + 33e^{3x} - 6e^{3x}\right) \\
&= c_1(12 - 12)e^x + c_2(30 - 30)e^{2x} + c_3(60 - 60)e^{3x} \\
&= 0
\end{aligned}$$

Thus $$y = c_1 e^x + c_2 e^{2x} + c_3 e^{3x}.$$

is also a solution of the differential equation.

Example 2

The function

$$y = x^2$$

is a solution of the homogeneous linear equation

$$x^2 y'' - 3xy' + 4y = 0$$

on $(0, \infty)$.

Now consider

$$y = cx^2$$

Then $y' = 2cx$ and $y'' = 2c$

So that $x^2 y'' - 3xy' + 4y = 2cx^2 - 6cx^2 + 4cx^2 = 0$

Hence the function

$$y = cx^2$$

is also a solution of the given differential equation.

The Wronskian

Suppose that y_1, y_2 are 2 solutions, on an interval I, of the second order homogeneous linear differential equation

$$a_2 \frac{d^2 y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = 0$$

Then either $W(y_1, y_2) = 0, \quad \forall\, x \in I$

or $W(y_1, y_2) \neq 0, \quad \forall\, x \in I$

To verify this we write the equation as

$$\frac{d^2 y}{dx^2} + \frac{Pdy}{dx} + Qy = 0$$

Now $W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_1' y_2$

Differentiating *w.r.to* x, we have

$$\frac{dW}{dx} = y_1 y_2'' - y_1'' y_2$$

Since y_1 and y_2 are solutions of the differential equation

$$\frac{d^2 y}{dx^2} + \frac{Pdy}{dx} + Qy = 0$$

Therefore

$$y_1'' + Py_1' + Qy_1 = 0$$

$$y_2'' + Py_2' + Qy_2 = 0$$

Multiplying 1st equation by y_2 and 2nd by y_1 the have

$$y_1''y_2 + Py_1'y_2 + Qy_1y_2 = 0$$

$$y_1y_2'' + Py_1y_2' + Qy_1y_2 = 0$$

Subtracting the two equations we have:

$$(y_1y_2'' - y_2y_1'') + P(y_1y_2' - y_1'y_2) = 0$$

or
$$\frac{dW}{dx} + PW = 0$$

This is a linear 1st order differential equation in W, whose solution is

$$W = ce^{-\int P dx}$$

Therefore

- If $c \neq 0$ then $W(y_1, y_2) \neq 0, \quad \forall\, x \in I$
- If $c = 0$ then $W(y_1, y_2) = 0, \quad \forall\, x \in I$

Hence Wronskian of y_1 and y_2 is either identically zero or is never zero on I.

In general

If $y_1, y_2, \ldots, y_n$ are n solutions, on an interval I, of the homogeneous *n*th order linear differential equation with constants coefficients

$$a_n \frac{d^n y}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1\frac{dy}{dx} + a_0 y = 0$$

Then

Either
$$W(y_1, y_2, \ldots, y_n) = 0, \quad \forall x \in I$$

or
$$W(y_1, y_2, \ldots, y_n) \neq 0, \quad \forall x \in I$$

Linear Independence of Solutions:

Suppose that

$$y_1, y_2, \ldots, y_n$$

are n solutions, on an interval I, of the homogeneous linear nth-order differential equation

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = 0$$

Then the set of solutions is linearly independent on I if and only if

$$W(y_{1,} y_2, \ldots, y_n) \neq 0$$

In other words

The solutions

$$y_1, y_2, \ldots, y_n$$

are linearly dependent if and only if

$$W(y_{1,} y_2, \ldots, y_n) = 0, \quad \forall x \in I$$

Fundamental Set of Solutions

A set

$$\{y_1, y_2, \ldots, y_n\}$$

of n linearly independent solutions, on interval I, of the homogeneous linear nth-order differential equation

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = 0$$

is said to be a fundamental set of solutions on the interval I.

Existence of a Fundamental Set

There always exists a fundamental set of solutions for a linear nth-order homogeneous differential equation

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = 0$$

on an interval I.

General Solution-Homogeneous Equations

Suppose that

$$\{y_1, y_2, \ldots, y_n\}$$

is a fundamental set of solutions, on an interval I, of the homogeneous linear nth-order differential equation

$$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=0$$

Then the general solution of the equation on the interval I is defined to be

$$y=c_1y_1(x)+c_2y_2(x)+\cdots+c_ny_n(x)$$

Here $c_1,c_2,\ldots,c_n$ are arbitrary constants.

Example 1

The functions

$$y_1=e^{3x} \text{ and } y_2=e^{-3x}$$

are solutions of the differential equation

$$y''-9y=0$$

Since $$W\left(e^{3x},e^{-3x}\right)=\begin{vmatrix} e^{3x} & e^{-3x} \\ 3e^{3x} & -3e^{-3x}\end{vmatrix}=-6\neq 0,\quad \forall x\in I$$

Therefore y_1 and y_2 from a fundamental set of solutions on $(-\infty,\infty)$. **Hence** general solution of the differential equation on the $(-\infty,\infty)$ is

$$y=c_1e^{3x}+c_2e^{-3x}$$

Example 2

Consider the function $y=4\sinh 3x-5e^{-3x}$

Then $y'=12\cosh 3x+15e^{-3x}$, $y''=36\sinh 3x-45e^{-3x}$

$\Rightarrow$ $y''=9\left(4\sinh 3x-5e^{-3x}\right)$ or $y''=9y$,

Therefore $y''-9y=0$

Hence $$y=4\sinh 3x-5e^{-3x}$$

is a particular solution of differential equation.

$$y''-9y=0$$

The general solution of the differential equation is

$$y=c_1e^{3x}+c_2e^{-3x}$$

Choosing $c_1=2, c_2=-7$

We obtain $y = 2e^{3x} - 7e^{-3x}$

$$y = 2e^{3x} - 2e^{-3x} - 5e^{-3x}$$

$$y = 4\left(\frac{e^{3x} - e^{-3x}}{2}\right) - 5e^{-3x}$$

$$y = 4\sinh 3x - 5e^{-3x}$$

Hence, the particular solution has been obtained from the general solution.

Example 3

Consider the differential equation

$$\frac{d^3y}{dx^3} - 6\frac{d^2y}{dx^2} + 11\frac{dy}{dx} - 6y = 0$$

and suppose that $y_1 = e^x,\ y_2 = e^{2x}$ and $y_3 = e^{3x}$

Then $\dfrac{dy_1}{dx} = e^x = \dfrac{d^2y_1}{dx^2} = \dfrac{d^3y_1}{dx^3}$

Therefore $\dfrac{d^3y_1}{dx^3} - 6\dfrac{d^2y_1}{dx^2} + 11\dfrac{dy_1}{dx} - 6y_1 = e^x - 6e^x + 11e^x - 6e^x$

or $\dfrac{d^3y_1}{dx^3} - 6\dfrac{d^2y_1}{dx^2} + 11\dfrac{dy_1}{dx} - 6y_1 = 12e^x - 12e^x = 0$

Thus the function y_1 is a solution of the differential equation. Similarly, we can verify that the other two functions i.e. y_2 and y_3 also satisfy the differential equation.

Now for all $x \in R$

$$W\left(e^x, e^{2x}, e^{3x}\right) = \begin{vmatrix} e^x & e^{2x} & e^{3x} \\ e^x & 2e^{2x} & 3e^{3x} \\ e^x & 4e^{2x} & 9e^{3x} \end{vmatrix} = 2e^{6x} \neq 0 \quad \forall\, x \in I$$

Therefore $y_1, y_2,$ and y_3 form a fundamental solution of the differential equation on $(-\infty, \infty)$. We conclude that

$$y = c_1e^x + c_2e^{2x} + c_3e^{3x}$$

is the general solution of the differential equation on the interval $(-\infty, \infty)$.

Non-Homogeneous Equations

A function y_p that satisfies the non-homogeneous differential equation

$$a_n(x)\frac{d^ny}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = g(x)$$

and is free of parameters is called the particular solution of the differential equation

Example 1

Suppose that

$$y_p = 3$$

Then $$y''_p = 0$$

So that

$$y''_p + 9y_p = 0 + 9(3)$$
$$= 27$$

Therefore

$$y_p = 3$$

is a particular solution of the differential equation

$$y''_p + 9y_p = 27$$

Example 2

Suppose that

$$y_p = x^3 - x$$

Then $$y'_p = 3x^2 - 1, \quad y''_p = 6x$$

Therefore $$x^2 y''_p + 2xy'_p - 8y_p = x^2(6x) + 2x\left(3x^2 - 1\right) - 8\left(x^3 - x\right)$$

$$= 4x^3 + 6x$$

Therefore

$$y_p = x^3 - x$$

is a particular solution of the differential equation

$$x^2 y'' + 2xy' - 8y = 4x^3 + 6x$$

Complementary Function

The general solution

$$y_c = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n$$

of the homogeneous linear differential equation

$$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=0$$

is known as the complementary function for the non-homogeneous linear differential equation.

$$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=g(x)$$

General Solution of Non-Homogeneous Equations

Suppose that

- The particular solution of the non-homogeneous equation

  $$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=g(x)$$

  is y_p.

- The complementary function of the non-homogeneous differential equation

  $$a_n(x)\frac{d^n y}{dx^n}+a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1(x)\frac{dy}{dx}+a_0(x)y=0$$

  is

  $$y_c=c_1y_1+c_2y_2+\cdots+c_ny_n.$$

- Then general solution of the non-homogeneous equation on the interval I is given by

  $$y=y_c+y_p$$

  or

  $$y=c_1y_1(x)+c_2y_2(x)+\cdots+c_ny_n(x)+y_p(x)=y_c(x)+y_p(x)$$

Hence

General Solution = Complementary solution + any particular solution.

Example

Suppose that

$$y_p=-\frac{11}{12}-\frac{1}{2}x$$

Then $$y'_p=-\frac{1}{2},\ y''_p=0=y'''_p$$

$\therefore \qquad \dfrac{d^3y_p}{dx^3} - 6\dfrac{d^2y_p}{dx^2} + 11\dfrac{dy_p}{dx} - 6y_p = 0 - 0 - \dfrac{11}{2} + \dfrac{11}{2} + 3x = 3x$

Hence

$$y_p = -\frac{11}{12} - \frac{1}{2}x$$

is a particular solution of the non-homogeneous equation

$$\frac{d^3y}{dx^3} - 6\frac{d^2y}{dx^2} + 11\frac{dy}{dx} - 6y = 3x$$

Now consider

$$y_c = c_1e^x + c_2e^{2x} + c_3e^{3x}$$

Then

$$\frac{dy_c}{dx} = c_1e^x + 2c_2e^{2x} + 3c_3e^{3x}$$
$$\frac{d^2y_c}{dx^2} = c_1e^x + 4c_2e^{2x} + 9c_3e^{3x}$$
$$\frac{d^3y_c}{dx^3} = c_1e^x + 8c_2e^{2x} + 27c_3e^{3x}$$

Since,

$$\frac{d^3y_c}{dx^3} - 6\frac{d^2y_c}{dx^2} + 11\frac{dy_c}{dx} - 6y_c$$
$$= c_1e^x + 8c_2e^{2x} + 27c_3e^{3x} - 6\left(c_1e^x + 4c_2e^{2x} + 9c_3e^{3x}\right)$$
$$+11\left(c_1e^x + 2c_2e^{2x} + 3c_3e^{3x}\right) - 6\left(c_1e^x + c_2e^{2x} + c_3e^{3x}\right)$$
$$= 12c_1e^x - 12c_1e^x + 30c_2e^{2x} - 30c_2e^{2x} + 60c_3e^{3x} - 60c_3e^{3x}$$
$$= 0$$

Thus y_c is general solution of associated homogeneous differential equation

$$\frac{d^3y}{dx^3} - 6\frac{d^2y}{dx^2} + 11\frac{dy}{dx} - 6y = 0$$

Hence general solution of the non-homogeneous equation is

$$y = y_c + y_p = c_1e^x + c_2e^{2x} + c_3e^{3x} - \frac{11}{12} - \frac{1}{2}x$$

Superposition Principle for Non-homogeneous Equations

Suppose that

$$y_{p_1}, y_{p_2}, \ldots, y_{p_k}$$

denote the particular solutions of the k differential equation

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = g_i(x),$$

$i = 1, 2, \ldots k$, on an interval I. Then

$$y_p = y_{p_1}(x) + y_{p_2}(x) + \cdots + y_{p_k}(x)$$

is a particular solution of

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = g_1(x) + g_2(x) + \cdots + g_k(x)$$

Example

Consider the differential equation

$$y'' - 3y' + 4y = -16x^2 + 24x - 8 + 2e^{2x} + 2xe^x - e^x$$

Suppose that

$$y_{p_1} = -4x^2, \quad y_{p_2} = e^{2x}, \quad y_{p_3} = xe^x$$

Then $$y''_{p_1} - 3y'_{p_1} + 4y_{p_1} = -8 + 24x - 16x^2$$

Therefore $$y_{p_1} = -4x^2$$

is a particular solution of the non-homogenous differential equation

$$y'' - 3y' + 4y = -16x^2 + 24x - 8$$

Similarly, it can be verified that

$$y_{p_2} = e^{2x} \text{ and } y_{p_3} = xe^x$$

are particular solutions of the equations:

$$y'' - 3y' + 4y = 2e^{2x}$$

and $$y''\text{-}3y' + 4y = 2xe^x - e^x$$

respectively.

Hence $$y_p = y_{p_1} + y_{p_2} + y_{p_3} = -4x^2 + e^{2x} + xe^x$$

is a particular solution of the differential equation

$$y'' - 3y' + 4y = -16x^2 + 24x - 8 + 2e^{2x} + 2xe^x - e^x$$

Exercise

Verify that the given functions form a fundamental set of solutions of the differential equation on the indicated interval. Form the general solution.

11. $y'' - y' - 12y = 0; \quad e^{-3x}, e^{4x}, \quad (-\infty, \infty)$

12. $y''-2y'+5y=0;\quad e^x\cos 2x, e^x\sin 2x,\quad (-\infty,\infty)$

13. $x^2y''+xy'+y=0;\quad \cos(\ln x), \sin(\ln x),\quad (0,\infty)$

14. $4y''-4y'+y=0;\quad e^{x/2}, xe^{x/2},\quad (-\infty,\infty)$

15. $x^2y''-6xy'+12y=0;\quad x^3,\ x^4\quad (0,\infty)$

16. $y''-4y=0;\quad \cosh 2x,\ \sinh 2x,\quad (-\infty,\infty)$

Verify that the given two-parameter family of functions is the general solution of the non-homogeneous differential equation on the indicated interval.

17. $y''+y=\sec x,\quad y=c_1\cos x+c_2\sin x+x\sin x+(\cos x)\ln(\cos x),\quad (-\pi/2,\pi/2)$.

18. $y''-4y'+4y=2e^{2x}+4x-12,\quad y=c_1e^{2x}+c_2xe^{2x}+x^2e^{2x}+x-2$

19. $y''-7y'+10y=24e^x,\quad y=c_1e^{2x}+c_2e^{5x}+6e^x,\quad (-\infty,\infty)$

20. $x^2y''+5xy'+y=x^2-x,\quad y=c_1x^{-1/2}+c_2x^{-1}+\frac{1}{15}x^2-\frac{1}{6}x,\quad (0,\infty)$

Lecture 15 Construction of a Second Solution

General Case

Consider the differential equation

$$a_2(x)\frac{d^2y}{dx^2}+a_1(x)\frac{dy}{dx}+a_0(x)y=0$$

We divide by $a_2(x)$ to put the above equation in the form

$$y'' + P(x)y' + Q(x)y = 0$$

Where $P(x)$ and $Q(x)$ are continuous on some interval I .

Suppose that $y_1(x) \neq 0,\ \forall\, x \in \mathrm{I}$ is a solution of the differential equation

Then $$y_1'' + P\,y_1' + Q\,y_1 = 0$$

We define $y = u(x)y_1\ (x)$ then

$$y' = uy_1' + y_1u', \qquad y'' = uy_1'' + 2y_1'u_1' + y_1u''$$

$$y'' + Py' + Qy = u[\underbrace{y_1'' + Py_1' + Qy_1}_{zero}] + y_1u'' + (2y_1' + Py_1)u' = 0$$

This implies that we must have

$$y_1u'' + (2y_1' + Py_1)u' = 0$$

If we suppose $w = u'$, then

$$y_1w' + (2y_1' + Py_1)w = 0$$

The equation is separable. Separating variables we have from the last equation

$$.\frac{dw}{w} + (2\frac{y_1'}{y_1} + P)dx = 0$$

Integrating

$$\ln|w| + 2\ln|y_1| = -\int P\,dx + c$$

$$\ln\left|wy_1^2\right| = -\int P\,dx + c$$

$$wy_1^2 = c_1e^{-\int Pdx}$$

$$w = \frac{c_1e^{-\int Pdx}dx}{y_1^2}$$

or $$u' = \frac{c_1 e^{-\int P dx}}{{y_1}^2}$$

Integrating again, we obtain

$$u = c_1 \int \frac{e^{-\int P dx}}{{y_1}^2} dx + c_2$$

Hence $$y = u(x) y_1(x) = c_1 y_1(x) \int \frac{e^{-\int P dx}}{{y_1}^2} dx + c_2 y_1(x).$$

Choosing $c_1 = 1 \text{and} c_2 = 0$, we obtain a second solution of the differential equation

$$y_2 = y_1(x) \int \frac{e^{-\int P dx}}{{y_1}^2} dx$$

The Woolskin

$$W\left(y_1(x), y_2(x)\right) = \begin{vmatrix} y_1 & y_1 \int \frac{e^{-\int P dx}}{{y_1}^2} dx \\ y_1' & \frac{e^{-\int P dx}}{y_1} + y_1' \int \frac{e^{-\int P dx}}{{y_1}^2} dx \end{vmatrix}$$

$$= e^{-\int P dx} \neq 0, \forall x$$

Therefore $y_1(x)$ and $y_2(x)$ are linear independent set of solutions. So that they form a fundamental set of solutions of the differential equation

$$y'' + P(x) y' + Q(x) y = 0$$

Hence the general solution of the differential equation is

$$y(x) = c_1 y_1(x) + c_2 y_2(x)$$

Example 1

Given that

$$y_1 = x^2$$

is a solution of

$$x^2 y'' - 3xy' + 4y = 0$$

Find general solution of the differential equation on the interval $(0,\infty)$.

Solution:

The equation can be written as

$$y'' - \frac{3}{x} y' + \frac{4}{x^2} y = 0,$$

The 2nd solution y_2 is given by

$$y_2 = y_1(x)\int \frac{e^{-\int P dx}}{y_1^2} dx$$

or $$y_2 = x^2 \int \frac{e^{3\int dx/x}}{x^4} dx = x^2 \int \frac{e^{\ln x^3}}{x^4} dx$$

$$y_2 = x^2 \int \frac{1}{x} dx = x^2 \ln x$$

Hence the general solution of the differential equation on $(0,\infty)$ is given by

$$y = c_1 \, y_1 + c_2 y_2$$

or $$y = c_1 \, x^2 + c_2 x^2 \ln x$$

Example 2

Verify that

$$y_1 = \frac{\sin x}{\sqrt{x}}$$

is a solution of

$$x^2 y'' + xy' + (x^2 - 1/4)y = 0$$

on $(0, \pi)$. Find a second solution of the equation.

Solution:

The differential equation can be written as

$$y'' + \frac{1}{x}y' + (1 - \frac{1}{4x^2})y = 0$$

The 2nd solution is given by

$$y_2 = y_1 \int \frac{e^{-\int P dx}}{y_1^2} dx$$

Therefore

$$y_2 = \frac{\sin x}{\sqrt{x}} \int \frac{e^{-\int \frac{dx}{x}}}{(\frac{\sin x}{\sqrt{x}})^2} dx$$

$$= \frac{-\sin x}{\sqrt{x}} \int \frac{x}{x \sin^2 x} dx$$

$$= \frac{-\sin x}{\sqrt{x}} \int \csc^2 x dx$$

$$= \frac{-\sin x}{\sqrt{x}}(-\cot x) = \frac{\cos x}{\sqrt{x}}$$

Thus the second solution is

$$y_2 = \frac{\cos x}{\sqrt{x}}$$

Hence, general solution of the differential equation is

$$y = c_1\left(\frac{\sin x}{\sqrt{x}}\right) + c_2\left(\frac{\cos x}{\sqrt{x}}\right)$$

Order Reduction

Example 3

Given that

$$y_1 = x^3$$

is a solution of the differential equation

$$x^2 y'' - 6y = 0,$$

Find second solution of the equation

Solution

We write the given equation as:

$$y'' - \frac{6}{x^2} y = 0$$

So that $$P(x) = -\frac{6}{x^2}$$

Therefore

$$y_2 = y_1 \int \frac{e^{-\int P dx}}{y_1{}^2} dx$$

$$y_2 = x^3 \int \frac{e^{-\int \frac{6}{x^2}}}{x^6} dx$$

$$y_2 = x^3 \int \frac{e^{\frac{6}{x}}}{x^6} dx$$

Therefore, using the formula

$$y_2 = y_1 \int \frac{e^{-\int P dx}}{y_1{}^2} dx$$

We encounter an integral that is difficult or impossible to evaluate.

Hence, we conclude sometimes use of the formula to find a second solution is not suitable. We need to try something else.

Alternatively, we can try the reduction of order to find y_2. For this purpose, we again define

$$y(x) = u(x) y_1(x) \quad \text{or} \quad y = u(x).x^3$$

then

$$y' = 3x^2 u + x^3 u'$$
$$y'' = x^3 u'' + 6x^2 u' + 6xu$$

Substituting the values of y, y'' in the given differential equation

$$x^2 y'' - 6y = 0$$

we have

$$x^2(x^3u'' + 6x^2u' + 6xu) - 6ux^3 = 0$$

or
$$x^5u'' + 6x^4u' = 0$$

or
$$u'' + \frac{6}{x}u' = 0,$$

If we take $w = u'$ then

$$w' + \frac{6}{x}w = 0$$

This is separable as well as linear first order differential equation in w. For using the latter, we find the integrating factor

$$I.F = e^{6\int \frac{1}{x}dx} = e^{6\ln x} = x^6$$

Multiplying with the $IF = x^6$, we obtain

$$x^6w' + 6x^5w = 0$$

or
$$\frac{d}{dx}(x^6w) = 0$$

Integrating w.r.t. ' x ', we have

$$x^6w = c_1$$

or
$$u' = \frac{c_1}{x^6}$$

Integrating once again, gives

$$u = -\frac{c_1}{5x^5} + c_2$$

Therefore
$$y = ux^3 = \frac{-c_1}{5x^2} + c_2x^3$$

Choosing $c_2 = 0$ and $c_1 = -5$, we obtain

$$y_2 = \frac{1}{x^2}$$

Thus the second solution is given by

$$y_2 = \frac{1}{x^2}$$

Hence, general solution of the given differential equation is

$$y = c_1 y_1 + c_2 y_2$$

i.e.
$$y = c_1 x^3 + c_2 \left(1/x^2\right)$$

Where c_1 and c_2 are constants.

Exercise

Find the 2nd solution of each of Differential equations by reducing order or by using the formula.

1. $y'' - y' = 0; \quad y_1 = 1$
2. $y'' + 2y' + y = 0; \quad y_1 = xe^{-x}$
3. $y'' + 9y = 0; \quad y_1 = \sin x$
4. $y'' - 25y = 0; \quad y_1 = e^{5x}$
5. $6y'' + y' - y = 0; \quad y_1 = e^{x/2}$
6. $x^2 y'' + 2xy' - 6y = 0; \quad y_1 = x^2$
7. $4x^2 y'' + y = 0; \quad y_1 = x^{1/2} \ln x$
8. $(1 - x^2)y'' - 2xy' = 0; \quad y_1 = 1$
9. $x^2 y'' - 3xy' + 5y = 0; \quad y_1 = x^2 \cos(\ln x)$
10. $(1 + x)y'' + xy' - y = 0; \quad y_1 = x$

Lecture 16 Homogeneous Linear Equations with Constant Coefficients

We know that the linear first order differential equation

$$\frac{dy}{dx} + my = 0$$

m being a constant, has the exponential solution on $(-\infty, \infty)$

$$y = c_1 e^{-mx}$$

The question?

- The question is whether or not the exponential solutions of the higher-order differential equations
$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_2 y^{//} + a_1 y^{/} + a_0 y = 0,$$
exist on $(-\infty, \infty)$.
- In fact all the solutions of this equation are exponential functions or constructed out of exponential functions.

Recall

That the linear differential of order n is an equation of the form

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x) y = g(x)$$

Method of Solution

Taking $n = 2$, the nth-order differential equation becomes

$$a_2 \frac{d^2 y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = 0$$

This equation can be written as

$$a\frac{d^2 y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

We now try a solution of the exponential form

$$y = e^{mx}$$

Then

$$y' = me^{mx} \text{ and } y'' = m^2 e^{mx}$$

Substituting in the differential equation, we have

$$e^{mx}(am^2 + bm + c) = 0$$

Since $e^{mx} \neq 0, \quad \forall x \in (-\infty, \infty)$

Therefore $am^2 + bm + c = 0$

This algebraic equation is known as the Auxiliary equation (AE).The solution of the auxiliary equation determines the solutions of the differential equation.

Case 1: *Distinct Real Roots*

If the auxiliary equation has distinct real roots m_1 and m_2 then we have the following two solutions of the differential equation.

$$y_1 = e^{m_1 x} \text{ and } y_2 = e^{m_2 x}$$

These solutions are linearly independent because

$$W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = (m_2 - m_1)e^{(m_1+m_2)x}$$

Since $m_1 \neq m_2$ and $e^{(m_1+m_2)x} \neq 0$

Therefore $W(y_1, y_2) \neq 0 \ \ \forall x \in (-\infty, \infty)$

Hence

- y_1 and y_2 form a fundamental set of solutions of the differential equation.
- The general solution of the differential equation on $(-\infty, \infty)$ is

$$y = c_1 e^{m_1 x} + c_2 e^{m_2 x}$$

Case 2. *Repeated Roots*

If the auxiliary equation has real and equal roots i.e

$$m = m_1, m_2 \quad \text{with} \quad m_1 = m_2$$

Then we obtain only one exponential solution

$$y = c_1 e^{mx}$$

To construct a second solution we rewrite the equation in the form

$$y'' + \frac{b}{a} y' + \frac{c}{a} y = 0$$

Comparing with $y'' + Py' + Qy = 0$

We make the identification

$$P = \frac{b}{a}$$

Thus a second solution is given by

$$y_2 = y_1 \int \frac{e^{-\int P dx}}{{y_1}^2} dx = e^{mx} \int \frac{e^{-\frac{b}{a}x}}{e^{2mx}} dx$$

Since the auxiliary equation is a quadratic algebraic equation and has equal roots

Therefore, $Disc. = b^2 - 4ac = 0$

We know from the quadratic formula

$$m = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

we have $2m = -\frac{b}{a}$

Therefore

$$y_2 = e^{mx} \int \frac{e^{2mx}}{e^{2mx}} dx = xe^{mx}$$

Hence the general solution is

$$y = c_1 e^{mx} + c_2 x e^{mx} = (c_1 + c_2 x) e^{mx}$$

Case 3: *Complex Roots*

If the auxiliary equation has complex roots $\alpha \pm i\beta$ then, with

$$m_1 = \alpha + i\beta \text{ and } m_2 = \alpha - i\beta$$

Where α >0 and β >0 are real, the general solution of the differential equation is

$$y = c_1 e^{(\alpha + i\beta)x} + c_2 e^{(\alpha - i\beta)x}$$

First we choose the following two pairs of values of c_1 and c_2

$$c_1 = c_2 = 1$$

$$c_1 = 1, c_2 = -1$$

Then we have

$$y_1 = e^{(\alpha + i\beta)x} + e^{(\alpha - i\beta)x}$$
$$y_2 = e^{(\alpha + i\beta)x} - e^{(\alpha - i\beta)x}$$

We know by the Euler's Formula that

$$e^{i\theta} = \cos\theta + i \sin\theta, \quad \theta \in \mathrm{R}$$

Using this formula, we can simplify the solutions y_1 and y_2 as

$$y_1 = e^{\alpha x}(e^{i\beta x} + e^{-i\beta x}) = 2e^{\alpha x} \cos \beta x$$
$$y_2 = e^{\alpha x}(e^{i\beta x} - e^{-i\beta x}) = 2ie^{\alpha x} \sin \beta x$$

We can drop constant to write

$$y_1 = e^{\alpha x}\cos\beta x\,,\quad y_2 = e^{\alpha x}\sin\beta x$$

The Wronskian

$$W\left(e^{\alpha x}\cos\beta x\,,\ e^{\alpha x}\sin\beta x\right) = \beta e^{2\alpha x} \neq 0 \quad \forall\ x$$

Therefore, $e^{\alpha x}\cos(\beta\, x),\ e^{\alpha x}\sin(\beta\, x)$

form a fundamental set of solutions of the differential equation on $(-\infty,\infty)$.

Hence general solution of the differential equation is

$$y = c_1 e^{\alpha x}\cos\beta x + c_2 e^{\alpha x}\sin\beta x$$

or $$y = e^{\alpha x}(c_1\cos\beta x + c_2\sin\beta x)$$

Example:

Solve

$$2y'' - 5y' - 3y = 0$$

Solution:

The given differential equation is

$$2y'' - 5y' - 3y = 0$$

Put $y = e^{mx}$

Then $y' = me^{mx},\ \ y'' = m^2 e^{mx}$

Substituting in the give differential equation, we have

$$\left(2m^2 - 5m - 3\right)e^{mx} = 0$$

Since $e^{mx} \neq 0\ \ \forall\ x$, the auxiliary equation is

$$2m^2 - 5m - 3 = 0 \quad \text{as} \quad e^{mx} \neq 0$$

$$(2m+1)(m-3) = 0 \Rightarrow m = -\frac{1}{2},\ 3$$

Therefore, the auxiliary equation has distinct real roots

$$m_1 = -\frac{1}{2} \text{ and } m_2 = 3$$

Hence the general solution of the differential equation is

$$y = c_1 e^{(-1/2)x} + c_2 e^{3x}$$

Example 2

Solve $y'' - 10y' + 25y = 0$

Solution:

We put $y = e^{mx}$

Then $y' = me^{mx}, y'' = m^2 e^{mx}$

Substituting in the given differential equation, we have

$$(m^2 - 10m + 25)e^{mx} = 0$$

Since $e^{mx} \neq 0 \; \forall \; x$, the auxiliary equation is

$$m^2 - 10m + 25 = 0$$

$$(m-5)^2 = 0 \Rightarrow m = 5, 5$$

Thus the auxiliary equation has repeated real roots i.e

$$m_1 = 5 = m_2$$

Hence general solution of the differential equation is

$$y = c_1 e^{5x} + c_2 x e^{5x}$$

or $$y = (c_1 + c_2 x)e^{5x}$$

Example 3

Solve the initial value problem

$$y'' - 4y' + 13y = 0$$
$$y(0) = \text{-}1, \; y'(0) = 2$$

Solution:

Given that the differential equation

$$y'' - 4y' + 13y = 0$$

Put $y = e^{mx}$

Then $y' = me^{mx}, \; y'' = m^2 e^{mx}$

Substituting in the given differential equation, we have:

$$(m^2 - 4m + 13)e^{mx} = 0$$

Since $e^{mx} \neq 0 \forall x$, the auxiliary equation is

$$m^2 - 4m + 13 = 0$$

By quadratic formula, the solution of the auxiliary equation is

$$m = \frac{4 \pm \sqrt{16 - 52}}{2} = 2 \pm 3i$$

Thus the auxiliary equation has complex roots

$$m_1 = 2 + 3i, \qquad m_2 = 2 - 3i$$

Hence general solution of the differential equation is

$$y = e^{2x}(c_1 \cos 3x + c_2 \sin 3x)$$

Example 4

Solve the differential equations

(a) $y'' + k^2 y = 0$

(b) $y'' - k^2 y = 0$

Solution

First consider the differential equation

$$y'' + k^2 y = 0,$$

Put $y = e^{mx}$

Then $y' = me^{mx}$ and $y'' = m^2 e^{mx}$

Substituting in the given differential equation, we have:

$$\left(m^2 + k^2\right) e^{mx} = 0$$

Since $e^{mx} \neq 0 \forall x$, the auxiliary equation is

$$m^2 + k^2 = 0$$

or $m = \pm ki$,

Therefore, the auxiliary equation has complex roots

$$m_1 = 0 + ki, \quad m_2 = 0 - ki$$

Hence general solution of the differential equation is

$$y = c_1 \cos kx + c_2 \sin kx$$

Next consider the differential equation

$$\frac{d^2 y}{dx^2} - k^2 y = 0$$

Substituting values y and y'', we have.

$$\left(m^2 - k^2\right) e^{mx} = 0$$

Since $e^{mx} \neq 0$, the auxiliary equation is

$$m^2 - k^2 = 0$$

$$\Rightarrow \quad m = \pm k$$

Thus the auxiliary equation has distinct real roots

$$m_1 = +k, \; m_2 = -k$$

Hence the general solution is

$$y = c_1 e^{kx} + c_2 e^{-kx}.$$

Higher Order Equations

If we consider nth order homogeneous linear differential equation

$$a_n \frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \ldots + a_1 \frac{dy}{dx} + a_0 y = 0$$

Then, the auxiliary equation is an nth degree polynomial equation

$$a_n m^n + a_{n-1} m^{n-1} + \ldots + a_1 m + a_0 = 0$$

Case 1: *Real distinct roots*

If the roots $m_1, m_2, \ldots, m_n$ of the auxiliary equation are all real and distinct, then the general solution of the equation is

$$y = c_1 e^{m_1 x} + c_2 e^{m_2 x} + \ldots + c_n e^{m_n x}$$

Case 2: *Real & repeated roots*

We suppose that m_1 is a root of multiplicity n of the auxiliary equation, then it can be shown that

$$e^{m_1 x}, xe^{m_1 x}, \ldots, x^{n-1} e^{m_1 x}$$

are n linearly independent solutions of the differential equation. Hence general solution of the differential equation is

$$y = c_1 e^{m_1 x} + c_2 x e^{m_1 x} + \ldots + c_n x^{n-1} e^{m_1 x}$$

Case 3: *Complex roots*

Suppose that coefficients of the auxiliary equation are real.

- We fix n at 6, all roots of the auxiliary are complex, namely

  $$\alpha_1 \pm i\beta_1, \quad \alpha_2 \pm i\beta_2, \; \alpha_3 \pm i\beta_3$$

  - Then the general solution of the differential equation

$$y = e^{\alpha_1 x}(c_1 \cos\beta_1 x + c_2 \sin\beta_1 x) + e^{\alpha_2 x}(c_3 \cos\beta_2 x + c_4 \sin\beta_2 x)$$
$$+ e^{\alpha_3 x}(c_5 \cos\beta_3 x + c_6 \sin\beta_3 x)$$

- If $n = 6$, two roots of the auxiliary equation are real and equal and the remaining 4 are complex, namely $\alpha_1 \pm i\beta_1, \quad \alpha_2 \pm i\beta_2$

  Then the general solution is

  $$y = e^{\alpha_1 x}(c_1 \cos\beta_1 x + c_2 \sin\beta_1 x) + e^{\alpha_2 x}(c_3 \cos\beta_2 x + c_4 \sin\beta_2 x) + c_5 e^{m_1 x} + c_6 x e^{m_1 x}$$

- If $m_1 = \alpha + i\beta$ is a complex root of multiplicity k of the auxiliary equation. Then its conjugate $m_2 = \alpha - i\beta$ is also a root of multiplicity k. Thus from Case 2 **,** the differential equation has $2k$ solutions

  $$e^{(\alpha+i\beta)x}, \; xe^{(\alpha+i\beta)x}, \; x^2 e^{(\alpha+i\beta)x}, \ldots, x^{k-1} e^{(\alpha+i\beta)x}$$
  $$e^{(\alpha-i\beta)x}, \; xe^{(\alpha-i\beta)x}, \; x^2 e^{(\alpha-i\beta)x}, \ldots, x^{k-1} e^{(\alpha-i\beta)x}$$

- By using the Euler’s formula, we conclude that the general solution of the differential equation is a linear combination of the linearly independent solutions

  $$e^{\alpha x}\cos\beta x, \; xe^{\alpha x}\cos\beta x, \; x^2 e^{\alpha x}\cos\beta x, \ldots, x^{k-1} e^{\alpha x}\cos\beta x$$
  $$e^{\alpha x}\sin\beta x, \; xe^{\alpha x}\sin\beta x, \; x^2 e^{\alpha x}\sin\beta x, \ldots, x^{k-1} e^{\alpha x}\sin\beta x$$

- Thus if $k = 3$ then

  $$y = e^{\alpha x}\left[\left(c_1 + c_2 x + c_3 x^2\right)\cos\beta x + \left(d_1 + d_2 x + d_3 x^2\right)\sin\beta x\right]$$

Solving the Auxiliary Equation

Recall that the auxiliary equation of nth degree differential equation is nth degree polynomial equation

- Solving the auxiliary equation could be difficult

  $$P_n(m) = 0, \quad \text{n} > 2$$

- One way to solve this polynomial equation is to guess a root m_1. Then $m - m_1$ is a factor of the polynomial $P_n(m)$.
- Dividing with $m - m_1$ synthetically or otherwise, we find the factorization
$$P_n(m) = (m - m_1)\,Q(m)$$
- We then try to find roots of the quotient i.e. roots of the polynomial equation
$$Q(m) = 0$$
- Note that if $m_1 = \frac{p}{q}$ is a rational real root of the equation
$$P_n(m) = 0, \quad \text{n} > 2$$
then p is a factor of a_0 and q of a_n.
- By using this fact we can construct a list of all possible rational roots of the auxiliary equation and test each of them by synthetic division.

Example 1

Solve the differential equation
$$y''' + 3y'' - 4y = 0$$

Solution:

Given the differential equation
$$y''' + 3y'' - 4y = 0$$

Put $\quad y = e^{mx}$

Then $\quad \text{y}^{/} = me^{mx}, y^{//} = m^2 e^{mx}$ and $y^{///} = m^3 e^{mx}$

Substituting this in the given differential equation, we have
$$(m^3 + 3m^2 - 4)e^{mx} = 0$$

Since $\quad e^{mx} \neq 0$

Therefore $\quad m^3 + 3m^2 - 4 = 0$

So that the auxiliary equation is
$$m^3 + 3m^2 - 4 = 0$$

Solution of the AE

If we take $m = 1$ then we see that
$$m^3 + 3m^2 - 4 = 1 + 3 - 4 = 0$$
Therefore $m = 1$ satisfies the auxiliary equations so that m-1 is a factor of the polynomial
$$m^3 + 3m^2 - 4$$
By synthetic division, we can write
$$m^3 + 3m^2 - 4 = (m-1)(m^2 + 4m + 4)$$

or $m^3 + 3m^2 - 4 = (m-1)(m+2)^2$

Therefore $m^3 + 3m^2 - 4 = 0$

$\Rightarrow (m-1)(m+2)^2 = 0$

or $m = 1, -2, -2$

Hence solution of the differential equation is

$$y = c_1 e^x + c_2 e^{-2x} + c_3 x e^{-2x}$$

Example 2

Solve

$$3y''' + 5y'' + 10y' - 4y = 0$$

Solution:

Given the differential equation

$$3y''' + 5y'' + 10y' - 4y = 0$$

Put $y = e^{mx}$

Then $y' = me^{mx}, y'' = m^2 e^{mx}$ and $y''' = m^3 e^{mx}$

Therefore the auxiliary equation is

$$3m^3 + 5m^2 + 10m - 4 = 0$$

Solution of the auxiliary equation:

a) $a_\circ = -4$ and all its factors are:

$p: \quad \pm 1, \pm 2, \pm 4$

b) $a_n = 3$ and all its factors are:

$q: \quad \pm 1, \pm 3$

c) List of possible rational roots of the auxiliary equation is

$\frac{p}{q}: \quad -1, 1, -2, 2, -4, 4, \frac{-1}{3}, \frac{1}{3}, \frac{-2}{3}, \frac{2}{3}, \frac{-4}{3}, \frac{4}{3}$

d) Testing each of these successively by synthetic division we find

$$\begin{array}{c|cccc} \frac{1}{3} & 3 & 5 & 10 & -4 \\ & & 1 & 2 & 4 \\ \hline & 3 & 6 & 12 & |\,0 \end{array}$$

Consequently a root of the auxiliary equation is

$$m = 1/3$$

The coefficients of the quotient are

3 6 12

Thus we can write the auxiliary equation as:

$$(m - 1/3)(3m^2 + 6m + 12) = 0$$

$$m - \frac{1}{3} = 0 \quad \text{or} \quad 3m^2 + 6m + 12 = 0$$

Therefore $m = 1/3$ or $m = -1 \pm i\sqrt{3}$

Hence solution of the given differential equation is

$$y = c_1 e^{(1/3)x} + e^{-x}\left(c_2 \cos\sqrt{3}x + c_3 \sin\sqrt{3}x\right)$$

Example 3

Solve the differential equation

$$\frac{d^4y}{dx^4} + 2\frac{d^2y}{dx^2} + y = 0$$

Solution:

Given the differential equation

$$\frac{d^4y}{dx^4} + 2\frac{d^2y}{dx^2} + y = 0$$

Put $y = e^{mx}$

Then $y' = me^{mx},\ y'' = m^2 e^{mx}$

Substituting in the differential equation, we obtain

$$\left(m^4 + 2m^2 + 1\right) e^{mx} = 0$$

Since $e^{mx} \neq 0$, the auxiliary equation is

$$m^4 + 2m^2 + 1 = 0$$

$$(m^2 + 1)^2 = 0$$

$$\Rightarrow m = \pm i,\ \pm i$$

$$m_1 = m_3 = i \quad \text{and} \quad m_2 = m_4 = -i$$

Thus i is a root of the auxiliary equation of multiplicity 2 and so is $-i$.

Now $\alpha = 0$ and $\beta = 1$

Hence the general solution of the differential equation is

$$y = e^{0x}\left[(c_1 + c_2 x)\cos x + (d_1 + d_2 x)\sin x\right]$$

or $y = c_1 \cos x + d_1 \sin x + c_2 x \cos x + d_2 x \sin x$

Exercise

Find the general solution of the given differential equations.

1. $y'' - 8y = 0$
2. $y'' - 3y' + 2y = 0$
3. $y'' + 4y' - y = 0$
4. $2y'' - 3y' + 4y = 0$
5. $4y''' + 4y'' + y' = 0$
6. $y''' + 5y'' = 0$
7. $y''' + 3y'' - 4y' - 12y = 0$

Solve the given differential equations subject to the indicated initial conditions.

8. $y''' + 2y'' - 5y' - 6y = 0, \qquad y(0) = y'(0) = 0,\ y''(0) = 1$
9. $\dfrac{d^4y}{dx^4} = 0,\ y(0) = 2,\ y'(0) = 3,\ y''(0) = 4,\ y'''(0) = 5$
10. $\dfrac{d^4y}{dx^4} - y = 0,\ y(0) = y'(0) = y''(0) = 0,\ y'''(0) = 1$

Lecture 17 Method of Undetermined Coefficients Superposition Approach

Recall

1. That a non-homogeneous linear differential equation of order n is an equation of the form

$$a_n \frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1 \frac{dy}{dx} + a_0 y = g(x)$$

The coefficients $a_0, a_1, \ldots, a_n$ can be functions of x. However, we will discuss equations with constant coefficients.

2. That to obtain the general solution of a non-homogeneous linear differential equation we must find**:**

   - The complementary function y_c, which is general solution of the associated homogeneous differential equation.
   - Any particular solution y_p of the non-homogeneous differential equation.

3. That the general solution of the non-homogeneous linear differential equation is given by

General solution = Complementary function + Particular Integral

Finding

Complementary function has been discussed in the previous lecture. In the next three lectures we will discuss methods for finding a particular integral for the non-homogeneous equation, namely

- The method of undetermined coefficients-*superposition approach*
- The method undetermined coefficients-*annihilator operator approach*.
- The method of variation of parameters.

The Method of Undetermined Coefficient

The method of undetermined coefficients developed here is limited to non-homogeneous linear differential equations

- That have constant coefficients, and
- Where the function $g(x)$ has a specific form.

The form of $g(x)$

The input function $g(x)$ can have one of the following forms:

- A constant function k.
- A polynomial function
- An exponential function e^x
- The trigonometric functions $\sin(\beta x)$, $\cos(\beta x)$
- Finite sums and products of these functions.

Otherwise, we cannot apply the method of undetermined coefficients.

The method

Consist of performing the following steps.

Step 1 Determine the form of the input function $g(x)$.

Step 2 Assume the general form of y_p according to the form of $g(x)$

Step 3 Substitute in the given non-homogeneous differential equation.

Step 4 Simplify and equate coefficients of like terms from both sides.

Step 5 Solve the resulting equations to find the unknown coefficients.

Step 6 Substitute the calculated values of coefficients in assumed y_p

Restriction on g ?

The input function g is restricted to have one of the above stated forms because of the reason**:**

- The derivatives of sums and products of polynomials, exponentials etc are again sums and products of similar kind of functions.
- The expression $ay_p^{\prime\prime} + by_p^{\prime} + cy_p$ has to be identically equal to the input function $g(x)$.

Therefore, to make an educated guess, y_p is assured to have the same form as g.

Caution!

- In addition to the form of the input function $g(x)$, the educated guess for y_p must take into consideration the functions that make up the complementary function y_c.
- No function in the assumed y_p must be a solution of the associated homogeneous differential equation. This means that the assumed y_p should not contain terms that duplicate terms in y_c.

Taking for granted that no function in the assumed y_p is duplicated by a function in y_c, some forms of g and the corresponding forms of y_p are given in the following table.

Trial particular solutions

Number	The input function $g(x)$	The assumed particular solution y_p
1	Any constant e.g. 1	A
2	$5x+7$	$Ax+B$
3	$3x^2-2$	Ax^2+Bx+c
4	x^3-x+1	Ax^3+Bx^2+Cx+D
5	$\sin 4x$	$A\cos 4x+B\sin 4x$
6	$\cos 4x$	$A\cos 4x+B\sin 4x$
7	e^{5x}	Ae^{5x}
8	$(9x-2)e^{5x}$	$(Ax+B)e^{5x}$
9	x^2e^{5x}	$(Ax^2+Bx+C)e^{5x}$
10	$e^{3x}\sin 4x$	$Ae^{3x}\cos 4x+Be^{3x}\sin 4x$
11	$5x^2\sin 4x$	$(A_1x^2+B_1x+C_1)\cos 4x+(A_2x^2+B_2x+C_2)\sin 4x$
12	$xe^{3x}\cos 4x$	$(Ax+B)e^{3x}\cos 4x+(Cx+D)e^{3x}\sin 4x$

If $g(x)$ *equals a sum?*

Suppose that

- The input function $g(x)$ consists of a sum of m terms of the kind listed in the above table i.e.

$$g(x)=g_1(x)+g_2(x)+\cdots+g_m(x).$$

- The trial forms corresponding to $g_1(x), g_2(x), \ldots, g_m(x)$ be $y_{p_1}, y_{p_2}, \ldots, y_{p_m}$.

Then the particular solution of the given non-homogeneous differential equation is

$$y_p = y_{p_1} + y_{p_2} + \cdots + y_{p_m}$$

In other words, *the form of* y_p *is a linear combination of all the linearly independent functions generated by repeated differentiation of the input function* $g(x)$.

Example 1

Solve $y'' + 4y' - 2y = 2x^2 - 3x + 6$

Solution:

Complementary function

To find y_c, we first solve the associated homogeneous equation

$$y'' + 4y' - 2y = 0$$

We put $y = e^{mx}, \quad y' = me^{mx}, \; y'' = m^2 e^{mx}$

Then the associated homogeneous equation gives

$$(m^2 + 4m - 2)e^{mx} = 0$$

Therefore, the auxiliary equation is

$$m^2 + 4m - 2 = 0 \quad \text{as} \quad e^{mx} \neq 0, \quad \forall\, x$$

Using the quadratic formula, roots of the auxiliary equation are

$$m = -2 \pm \sqrt{6}$$

Thus we have real and distinct roots of the auxiliary equation

$$m_1 = -2 - \sqrt{6} \quad \text{and} \quad m_2 = -2 + \sqrt{6}$$

Hence the complementary function is

$$y_c = c_1 e^{-(2+\sqrt{6})x} + c_2 e^{(-2+\sqrt{6})x}$$

Next we find a particular solution of the non-homogeneous differential equation.

Particular Integral

Since the input function

$$g(x) = 2x^2 - 3x + 6$$

is a quadratic polynomial. Therefore, we assume that

$$y_p = Ax^2 + Bx + C$$

Then $y_p' = 2Ax + B$ and $y_p'' = 2A$

Therefore $y_p'' + 4y_p' - 2y_p = 2A + 8Ax + 4B - 2Ax^2 - 2Bx - 2C$

Substituting in the given equation, we have

$$2A + 8Ax + 4B - 2Ax^2 - 2Bx - 2C = 2x^2 - 3x + 6$$

or $-2Ax^2 + (8A - 2B)x + (2A + 4B - 2C) = 2x^2 - 3x + 6$

Equating the coefficients of the like powers of x, we have

$$-2A=2,\ 8A-2B=-3,\quad 2A+4B-2C=6$$

Solving this system of equations leads to the values

$$A=-1,\quad B=-5/2,\quad C=-9.$$

Thus a particular solution of the given equation is

$$y_p=-x^2-\frac{5}{2}x-9.$$

Hence, the general solution of the given non-homogeneous differential equation is given by

$$y=y_c+y_p$$

or $$y=-x^2-\frac{5}{2}x-9+c_1e^{-(2+\sqrt{6})x}+c_2e^{(-2+\sqrt{6})x}$$

Example 2

Solve the differential equation

$$y''-y'+y=2\sin 3x$$

Solution:

Complementary function

To find y_c, we solve the associated homogeneous differential equation

$$y''-y'+y=0$$

Put $$y=e^{mx},\quad y'=me^{mx},\quad y''=m^2e^{mx}$$

Substitute in the given differential equation to obtain the auxiliary equation

$$m^2-m+1=0 \qquad \text{or} \qquad m=\frac{1\pm i\sqrt{3}}{2}$$

Hence, the auxiliary equation has complex roots. Hence the complementary function is

$$y_c=e^{(1/2)x}\left(c_1\cos\frac{\sqrt{3}}{2}x+c_2\sin\frac{\sqrt{3}}{2}x\right)$$

Particular Integral

Since successive differentiation of

$$g(x)=\sin 3x$$

produce $\sin 3x$ and $\cos 3x$

Therefore, we include both of these terms in the assumed particular solution, see table

$$y_p = A\cos 3x + B\sin 3x.$$

Then $$y'_p = -3A\sin 3x + 3B\cos 3x.$$

$$y''_p = -9A\cos 3x - 9B\sin 3x.$$

Therefore $$y_p'' - y_p' + y_p = (-8A - 3B)\cos 3x + (3A - 8B)\sin 3x.$$

Substituting in the given differential equation

$$(-8A - 3B)\cos 3x + (3A - 8B)\sin 3x = 0\cos 3x + 2\sin 3x.$$

From the resulting equations

$$-8A - 3B = 0,\ 3A - 8B = 2$$

Solving these equations, we obtain

$$A = 6/73, B = -16/73$$

A particular solution of the equation is

$$y_p = \frac{6}{73}\cos 3x - \frac{16}{73}\sin 3x$$

Hence the general solution of the given non-homogeneous differential equation is

$$y = e^{(1/2)x}\left(c_1\cos\frac{\sqrt{3}}{2}x + c_2\sin\frac{\sqrt{3}}{2}x\right) + \frac{6}{73}\cos 3x - \frac{16}{73}\sin 3x$$

Example 3

Solve $$y'' - 2y' - 3y = 4x - 5 + 6xe^{2x}$$

Solution:

Complementary function

To find y_c, we solve the associated homogeneous equation

$$y'' - 2y' - 3y = 0$$

Put $$y = e^{mx},\quad y' = me^{mx},\ y'' = m^2e^{mx}$$

Substitute in the given differential equation to obtain the auxiliary equation

$$m^2 - 2m - 3 = 0$$

$$\Rightarrow (m+1)(m-3) = 0$$

$$m = -1, 3$$

Therefore, the auxiliary equation has real distinct root

$$m_1 = -1, m_2 = 3$$

Thus the complementary function is

$$y_c = c_1e^{-x} + c_2e^{3x}.$$

Particular integral

Since $$g(x) = (4x - 5) + 6xe^{2x} = g_1(x) + g_2(x)$$

Corresponding to $g_1(x)$ $$y_{p_1} = Ax + B$$

Corresponding to $g_2(x)$ $$y_{p_2} = (Cx + D)e^{2x}$$

The superposition principle suggests that we assume a particular solution

$$y_p = y_{p_1} + y_{p_2}$$

i.e. $$y_p = Ax + B + (Cx + D)\,e^{2x}$$

Then $$y'_p = A + 2(Cx + D)\,e^{2x} + Ce^{2x}$$

$$y''_p = 4(Cx + D)\,e^{2x} + 4Ce^{2x}$$

Substituting in the given

$$y_p{}'' - 2y_p{}' - 3y_p = 4Cxe^{2x} + 4De^{2x} + 4Ce^{2x} - 2A - 4Cxe^{2x}$$
$$- 4De^{2x} - 2Ce^{2x} - 3Ax - 3B - 3Cxe^{2x} - 3De^{2x}$$

Simplifying and grouping like terms

$$y_p{}'' - 2y_p{}' - 3y_p = -3Ax - 2A - 3B - 3Cxe^{2x} + (2C - 3D)e^{2x} = 4x - 5 + 6xe^{2x}.$$

Substituting in the non-homogeneous differential equation, we have

$$-3Ax - 2A - 3B - 3Cxe^{2x} + (2C - 3D)e^{2x} = 4x - 5 + 6xe^{2x} + 0e^{2x}$$

Now equating constant terms and coefficients of x, xe^{2x} and e^{2x}, we obtain

$$-2A - 3B = -5, \qquad -3A = 4$$
$$-3C = 6, \qquad 2C - 3D = 0$$

Solving these algebraic equations, we find

$$A = -4/3, \quad B = 23/9$$
$$C = -2, \quad D = -4/3$$

Thus, a particular solution of the non-homogeneous equation is

$$y_p = -(4/3)x + (23/9) - 2\,xe^{2x} - (4/3)e^{2x}$$

The general solution of the equation is

$$y = y_c + y_p = c_1e^{-x} + c_2e^{3x} - (4/3)x + (23/9) - 2\,x\,e^{2x} - (4/3)e^{2x}$$

Duplication between y_p and y_c?

- If a function in the assumed y_p is also present in y_c then this function is a solution of the associated homogeneous differential equation. In this case the obvious assumption for the form of y_p is not correct.
- In this case we suppose that the input function is made up of terms of n kinds i.e.
$$g(x) = g_1(x) + g_2(x) + \cdots + g_n(x)$$
and corresponding to this input function the assumed particular solution y_p is
$$y_p = y_{p_1} + y_{p_2} + \cdots + y_{p_n}$$
- If a y_{p_i} contain terms that duplicate terms in y_c, then that y_{p_i} must be multiplied with x^n, n being the least positive integer that eliminates the duplication.

Example 4

Find a particular solution of the following non-homogeneous differential equation

$$y'' - 5y' + 4y = 8e^x$$

Solution:

To find y_c, we solve the associated homogeneous differential equation

$$y'' - 5y' + 4y = 0$$

We put $y = e^{mx}$ in the given equation, so that the auxiliary equation is

$$m^2 - 5m + 4 = 0 \Rightarrow m = 1, 4$$

Thus $\quad y_c = c_1 e^x + c_2 e^{4x}$

Since $\quad g(x) = 8e^x$

Therefore, $\quad y_p = Ae^x$

Substituting in the given non-homogeneous differential equation, we obtain

$$Ae^x - 5Ae^x + 4Ae^x = 8e^x$$

So $\quad 0 = 8e^x$

Clearly we have made a wrong assumption for y_p, as we did not remove the duplication.

Since Ae^x is present in y_c. Therefore, it is a solution of the associated homogeneous differential equation

$$y'' - 5y' + 4y = 0$$

To avoid this we find a particular solution of the form

$$y_p = Axe^x$$

We notice that there is no duplication between y_c and this new assumption for y_p

Now $\quad y_p' = Axe^x + Ae^x, \quad y_p'' = Axe^x + 2Ae^x$

Substituting in the given differential equation, we obtain

$$Axe^x + 2Ae^x - 5Axe^x - 5Ae^x + 4Axe^x = 8e^x.$$

or $\quad -3Ae^x = 8e^x \Rightarrow A = -8/3.$

So that a particular solution of the given equation is given by

$$y_p = -(8/3)e^x$$

Hence, the general solution of the given equation is

$$y = c_1 e^x + c_2 e^{4x} - (8/3)\, x\, e^x$$

Example 5

Determine the form of the particular solution

(a) $\quad y'' - 8y' + 25y = 5x^3 e^{-x} - 7e^{-x}$

(b) $\quad y'' + 4y = x\cos x.$

Solution:

(a) To find y_c we solve the associated homogeneous differential equation

$$y'' - 8y' + 25y = 0$$

Put $y = e^{mx}$

Then the auxiliary equation is

$$m^2 - 8m + 25 = 0 \Rightarrow m = 4 \pm 3i$$

Roots of the auxiliary equation are complex

$$y_c = e^{4x}(c_1 \cos 3x + c_2 \sin 3x)$$

The input function is

$$g(x) = 5x^3 e^{-x} - 7e^{-x} = (5x^3 - 7)e^{-x}$$

Therefore, we assume a particular solution of the form

$$y_p = (Ax^3 + Bx^2 + Cx + D)e^{-x}$$

Notice that there is ***no duplication*** between the terms in y_p and the terms in y_c. Therefore, while proceeding further we can easily calculate the value A, B, C and D.

(b) Consider the associated homogeneous differential equation

$$y'' + 4y = 0$$

Since $g(x) = x\cos x$

Therefore, we assume a particular solution of the form

$$y_p = (Ax + B)\cos x + (Cx + D)\sin x$$

Again observe that there is ***no duplication*** of terms between y_c and y_p

Example 6

Determine the form of a particular solution of

$$y'' - y' + y = 3x^2 - 5\sin 2x + 7xe^{6x}$$

Solution:

To find y_c, we solve the associated homogeneous differential equation

$$y'' - y' + y = 0$$

Put $y = e^{mx}$

Then the auxiliary equation is

$$m^2 - m + 1 = 0 \Rightarrow m = \frac{1 \pm i\sqrt{3}}{2}$$

Therefore $$y_c = e^{(1/2)x}\left(c_1 \cos\frac{\sqrt{3}}{2}x + c_2 \sin\frac{\sqrt{3}}{2}x\right)$$

Since $$g(x) = 3x^2 - 5\sin 2x + 7xe^{6x} = g_1(x) + g_2(x) + g_3(x)$$

Corresponding to $g_1(x) = 3x^2$: $\quad y_{p_1} = Ax^2 + Bx + C$

Corresponding to $g_2(x) = -5\sin 2x$: $\quad y_{p_2} = D\cos 2x + E\sin 2x$

Corresponding to $g_3(x) = 7xe^{6x}$: $\quad y_{p_3} = (Fx + G)\,e^{6x}$

Hence, the assumption for the particular solution is
$$y_p = y_{p_1} + y_{p_2} + y_{p_3}$$
or $$y_p = Ax^2 + Bx + C + D\cos 2x + E\sin 2x + (Fx + G)e^{6x}$$
No term in this assumption duplicate any term in the complementary function
$$y_c = c_1 e^{2x} + c_2 e^{7x}$$

Example 7

Find a particular solution of
$$y'' - 2y' + y = e^x$$

Solution:

Consider the associated homogeneous equation
$$y'' - 2y' + y = 0$$
Put $$y = e^{mx}$$
Then the auxiliary equation is
$$m^2 - 2m + 1 = (m-1)^2 = 0$$
$$\Rightarrow m = 1,\ 1$$
Roots of the auxiliary equation are real and equal. Therefore,
$$y_c = c_1 e^x + c_2 x e^x$$

Since $$g(x) = e^x$$
Therefore, we assume that
$$y_p = Ae^x$$
This assumption fails because of duplication between y_c and y_p. We multiply with x

Therefore, we now assume
$$y_p = Axe^x$$
However, the duplication is still there. Therefore, we again multiply with x and assume
$$y_p = Ax^2 e^x$$
Since there is no duplication, this is acceptable form of the trial y_p
$$y_p = \frac{1}{2}x^2 e^x$$

Example 8

Solve the initial value problem

$$y'' + y = 4x + 10\sin x,$$

$$y(\pi) = 0,\ y'(\pi) = 2$$

Solution

Consider the associated homogeneous differential equation

$$y'' + y = 0$$

Put $y = e^{mx}$

Then the auxiliary equation is

$$m^2 + 1 = 0 \Rightarrow m = \pm i$$

The roots of the auxiliary equation are complex. Therefore, the complementary function is

$$y_c = c_1\cos x + c_2\sin x$$

Since $g(x) = 4x + 10\sin x = g_1(x) + g_2(x)$

Therefore, we assume that

$$y_{p_1} = Ax + B, \quad y_{p_2} = C\cos x + D\sin x$$

So that $y_p = Ax + B + C\cos x + D\sin x$

Clearly, there is duplication of the functions $\cos x$ and $\sin x$. To remove this duplication we multiply y_{p_2} with x. Therefore, we assume that

$$y_p = Ax + B + Cx\cos x + Dx\sin x.$$

$$y_p'' = -2C\sin x - Cx\cos x + 2D\cos x - Dx\sin x$$

So that $y_p'' + y_p = Ax + B - 2C\sin x + 2Dx\cos x$

Substituting into the given non-homogeneous differential equation, we have

$$Ax + B - 2C\sin x + 2Dx\cos x = 4x + 10\sin x$$

Equating constant terms and coefficients of x, $\sin x$, $x\cos x$, we obtain

$$B = 0,\ A = 4,\ -2C = 10,\ 2D = 0$$

So that $A = 4, B = 0, C = -5,\ D = 0$

Thus $y_p = 4x - 5x\cos x$

Hence the general solution of the differential equation is

$$y = y_c + y_p = c_1\cos x + c_2\sin x + 4x - 5x\cos x$$

We now apply the initial conditions to find c_1 and c_2.

$$y(\pi) = 0 \Rightarrow c_1\cos\pi + c_2\sin\pi + 4\pi - 5\pi\cos\pi = 0$$

Since $\sin\pi = 0, \cos\pi = -1$

Therefore $c_1 = 9\pi$

Now $y' = -9\pi\sin x + c_2\cos x + 4 + 5x\sin x - 5\cos x$

Therefore $y'(\pi) = 2 \Rightarrow -9\pi\sin\pi + c_2\cos\pi + 4 + 5\pi\sin\pi - 5\cos\pi = 2$

$\therefore$ $c_2 = 7.$

Hence the solution of the initial value problem is

$$y = 9\pi\cos x + 7\sin x + 4x - 5x\cos x.$$

Example 9

Solve the differential equation

$$y'' - 6y' + 9y = 6x^2 + 2 - 12e^{3x}$$

Solution:

The associated homogeneous differential equation is

$$y'' - 6y' + 9y = 0$$

Put $y = e^{mx}$

Then the auxiliary equation is

$$m^2 - 6m + 9 = 0 \Rightarrow m = 3, 3$$

Thus the complementary function is

$$y_c = c_1 e^{3x} + c_2 x e^{3x}$$

Since $g(x) = (x^2 + 2) - 12e^{3x} = g_1(x) + g_2(x)$

We assume that

Corresponding to $g_1(x) = x^2 + 2$**:** $y_{p_1} = Ax^2 + Bx + C$

Corresponding to $g_2(x) = -12e^{3x}$**:** $y_{p_2} = De^{3x}$

Thus the assumed form of the particular solution is

$$y_p = Ax^2 + Bx + C + De^{3x}$$

The function e^{3x} in y_{p_2} is duplicated between y_c and y_p. Multiplication with x does not remove this duplication. However, if we multiply y_{p_2} with x^2, this duplication is removed.

Thus the operative from of a particular solution is

$$y_p = Ax^2 + Bx + C + Dx^2 e^{3x}$$

Then $y'_p = 2Ax + B + 2Dxe^{3x} + 3Dx^2 e^{3x}$

and $y''_p = 2A + 2De^{3x} + 6Dxe^{3x} + 9Dx^2 e^{3x}$

Substituting into the given differential equation and collecting like term, we obtain

$$y_p'' - 6y_p' + y_p = 9Ax^2 + (-12A + 9B)x + 2A - 6B + 9C + 2De^{3x} = 6x^2 + 2 - 12e^{3x}$$

Equating constant terms and coefficients of x, x^2 and e^{3x} yields

$$2A - 6B + 9C = 2, \quad -12A + 9B = 0$$

$$9A = 6, \quad 2D = -12$$

Solving these equations, we have the values of the unknown coefficients

$$A = 2/3, B = 8/9, C = 2/3 \text{ and } D = -6$$

Thus $y_p = \frac{2}{3}x^2 + \frac{8}{9}x + \frac{2}{3} - 6x^2 e^{3x}$

Hence the general solution

$$y = y_c + y_p = c_1 e^{3x} + c_2 x e^{3x} + \frac{2}{3}x^2 + \frac{8}{9}x + \frac{2}{3} - 6x^2 e^{3x}.$$

Higher –Order Equation

The method of undetermined coefficients can also be used for higher order equations of the form

$$a_n \frac{d^n y}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + ... + a_1 \frac{dy}{dx} + a_0 y = g(x)$$

with constant coefficients. The only requirement is that $g(x)$ consists of the proper kinds of functions as discussed earlier.

Example 10

Solve $$y^{///} + y^{//} = e^x \cos x$$

Solution:

To find the complementary function we solve the associated homogeneous differential equation

$$y^{///} + y^{//} = 0$$

Put $$y = e^{mx}, y' = me^{mx}, y'' = m^2 e^{mx}$$

Then the auxiliary equation is

$$m^3 + m^2 = 0$$

or $$m^2(m+1) = 0 \Rightarrow m = 0,0,-1$$

The auxiliary equation has equal and distinct real roots. Therefore, the complementary function is

$$y_c = c_1 + c_2 x + c_3 e^{-x}$$

Since $$g(x) = e^x \cos x$$

Therefore, we assume that

$$y_p = Ae^x \cos x + Be^x \sin x$$

Clearly, there is no duplication of terms between y_c and y_p.

Substituting the derivatives of y_p in the given differential equation and grouping the like terms, we have

$$y_p^{///} + y_p^{//} = (-2A + 4B)e^x \cos x + (-4A - 2B)e^x \sin x = e^x \cos x.$$

Equating the coefficients, of $e^x \cos x$ and $e^x \sin x$, yields

$$-2A + 4B = 1, -4A - 2B = 0$$

Solving these equations, we obtain

$$A = -1/10, B = 1/5$$

So that a particular solution is

$$y_p = c_1 + c_2 x + c_3 e^{-x} - (1/10)e^x \cos x + (1/5)e^x \sin x$$

Hence the general solution of the given differential equation is

$$y_p = c_1 + c_2 x + c_3 e^{-x} - (1/10)e^x \cos x + (1/5)e^x \sin x$$

Example 12

Determine the form of a particular solution of the equation

$$y'''' + y''' = 1 - e^{-x}$$

Solution:

Consider the associated homogeneous differential equation

$$y'''' + y''' = 0$$

The auxiliary equation is

$$m^4 + m^3 = 0 \Rightarrow m = 0, 0, 0, -1$$

Therefore, the complementary function is

$$y_c = c_1 + c_2 x + c_3 x^2 + c_4 e^{-x}$$

Since $\quad g(x) = 1 - e^{-x} = g_1(x) + g_2(x)$

Corresponding to $g_1(x) = 1$: $\quad y_{p_1 = A}$

Corresponding to $g_2(x) = -e^{-x}$: $\quad y_{p_2} = Be^{-x}$

Therefore, the normal assumption for the particular solution is

$$y_p = A + Be^{-x}$$

Clearly there is duplication of

(i) The constant function between y_c and y_{p_1}.

(ii) The exponential function e^{-x} between y_c and y_{p_2}.

To remove this duplication, we multiply y_{p_1} with x^3 and y_{p_2} with x. This duplication can't be removed by multiplying with x and x^2. Hence, the correct assumption for the particular solution y_p is

$y_p = Ax^3 + Bxe^{-x}$

Exercise

Solve the following differential equations using the undetermined coefficients.

1. $\frac{1}{4}y'' + y' + y = x^2 + 2x$

2. $y'' - 8y' + 20y = 100x^2 - 26xe^x$

3. $y'' + 3y = -48x^2 e^{3x}$

4. $4y'' - 4y' - 3y = \cos 2x$

5. $y'' + 4y = (x^2 - 3)\sin 2x$

6. $y'' - 5y' = 2x^3 - 4x^2 - x + 6$

7. $y'' - 2y' + 2y = e^{2x}(\cos x - 3\sin x)$

Solve the following initial value problems.

8. $y'' + 4y' + 4y = (3 + x)e^{-2x}, \quad y(0) = 2, y'(0) = 5$

9. $\frac{d^2x}{dt^2}+\omega^2 x = F_0 \cos\gamma t, \qquad x(0)=0, x^{/}(0)=0$

10. $y^{///}+8y=2x-5+8e^{-2x}, \quad y(0)=-5, \; y^{/}(0)=3, y^{//}(0)=-4$

Lecture 18 Undetermined Coefficient: Annihilator Operator Approach

Recall

1. That a non-homogeneous linear differential equation of order n is an equation of the form

$$a_n \frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1 \frac{dy}{dx} + a_0 y = g(x)$$

The following differential equation is called the associated homogeneous equation

$$a_n \frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1 \frac{dy}{dx} + a_0 y = 0$$

The coefficients $a_0, a_1, \ldots, a_n$ can be functions of x. However, we will discuss equations with constant coefficients.

2. That to obtain the general solution of a non-homogeneous linear differential equation we must find**:**

- The complementary function y_c, which is general solution of the associated homogeneous differential equation.
- Any particular solution y_p of the non-homogeneous differential equation.

3. That the general solution of the non-homogeneous linear differential equation is given by

General Solution = Complementary Function + Particular Integral

- Finding the complementary function has been completely discussed in an earlier lecture

- In the previous lecture, we studied a method for finding particular integral of the non-homogeneous equations. This was the *method of undetermined coefficients developed from the viewpoint of superposition* principle.

- In the present lecture, we will learn to find particular integral of the non-homogeneous equations by the same method utilizing the concept of differential annihilator operators.

Differential Operators

- In calculus, the differential coefficient d/dx is often denoted by the capital letter D. So that

$$\frac{dy}{dx} = Dy$$

The symbol D is known as differential operator.

- This operator transforms a differentiable function into another function, e.g.

$$D(e^{4x}) = 4e^{4x},\ D(5x^3 - 6x^2) = 15x^2 - 12x,\ D(\cos 2x) = -2\sin 2x$$

- The differential operator D possesses the property of linearity. This means that if f, g are two differentiable functions, then

$$D\{af(x) + bg(x)\} = aDf(x) + bDg(x)$$

Where a and b are constants. Because of this property, we say that D is a linear differential operator.

- Higher order derivatives can be expressed in terms of the operator D in a natural manner:

$$\frac{d^2y}{d^2x} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = D(Dy) = D^2y$$

Similarly

$$\frac{d^3y}{dx^3} = D^3y, \ldots, \frac{d^ny}{d^nx} = D^ny$$

- The following polynomial expression of degree n involving the operator D

$$a_nD^n + a_{n-1}D^{n-1} + \cdots + a_1D + a_0$$

is also a linear differential operator.

For example, the following expressions are all linear differential operators

$$D+3,\ D^2 + 3D - 4,\ 5D^3 - 6D^2 + 4D$$

Differential Equation in Terms of D

Any linear differential equation can be expressed in terms of the notation D. Consider a 2nd order equation with constant coefficients

$$ay'' + by' + cy = g(x)$$

Since $$\frac{dy}{dx} = Dy, \frac{d^2y}{dx^2} = D^2y$$

Therefore the equation can be written as

$$aD^2y + bDy + cy = g(x)$$

or $$(aD^2 + bD + c)y = g(x)$$

Now, we define another differential operator L as

$$L = aD^2 + bD + c$$

Then the equation can be compactly written as

$$L(y) = g(x)$$

The operator L is a second-order linear differential operator with constant coefficients.

Example 1

Consider the differential equation

$$y'' + y' + 2y = 5x - 3$$

Since $$\frac{dy}{dx} = Dy, \frac{d^2y}{dx^2} = D^2y$$

Therefore, the equation can be written as

$$(D^2 + D + 2)y = 5x - 3$$

Now, we define the operator L as

$$L = D^2 + D + 2$$

Then the given differential can be compactly written as

$$L(y) = 5x - 3$$

Factorization of a differential operator

- An nth-order linear differential operator

  $$L = a_n D^n + a_{n-1} D^{n-1} + \cdots + a_1 D + a_0$$

  with constant coefficients can be factorized, whenever the characteristics polynomial equation

  $$L = a_n m^n + a_{n-1} m^{n-1} + \cdots + a_1 m + a_0$$

  can be factorized.

- The factors of a linear differential operator with constant coefficients commute.

Example 2

(a) Consider the following 2nd order linear differential operator

$$D^2 + 5D + 6$$

If we treat D as an algebraic quantity, then the operator can be factorized as

$$D^2 + 5D + 6 = (D+2)(D+3)$$

(b) To illustrate the commutative property of the factors, we consider a twice-differentiable function $y = f(x)$. Then we can write

$$(D^2 + 5D + 6)y = (D+2)(D+3)y = (D+3)(D+2)y$$

To verify this we let

$$w = (D+3)y = y' + 3y$$

Then

$$(D+2)w = Dw + 2w$$

or $$(D+2)w = (y'' + 3y') + (2y' + 6y)$$

or $$(D+2)w = y'' + 5y' + 6y$$

or $(D+2)(D+3)y = y'' + 5y' + 6y$

Similarly if we let

$$w = (D+2)y = (y' + 2y)$$

Then $(D+3)w = Dw + 3w = (y'' + 2y') + (3y' + 6y)$

or $(D+3)w = y'' + 5y' + 6y$

or $(D+3)(D+2)y = y'' + 5y' + 6y$

Therefore, we can write from the two expressions that

$$(D+3)(D+2)y = (D+2)(D+3)y$$

Hence $(D+3)(D+2)y = (D+2)(D+3)y$

Example 3

(a) The operator $D^2 - 1$ can be factorized as

$$D^2 - 1 = (D+1)(D-1) .$$

or $D^2 - 1 = (D-1)(D+1)$

(b) The operator $D^2 + D + 2$ does not factor with real numbers.

Example 4

The differential equation

$$y'' + 4y' + 4y = 0$$

can be written as

$$\left(D^2 + 4D + 4\right)y = 0$$

or $(D+2)(D+2)y = 0$

or $(D+2)^2 y = 0.$

Annihilator Operator

Suppose that

- ***L*** is a linear differential operator with constant coefficients.
- *y = f(x)* defines a sufficiently differentiable function.
- The function *f* is such that *L(y)=0*

Then the differential operator *L* is said to be an **annihilator operator** of the function *f*.

Example 5
Since

$$Dx = 0,\ \ D^2 x = 0,\ \ D^3 x^2 = 0,\ \ D^4 x^3 = 0,\ \ldots$$

Therefore, the differential operators

$$D,\ \ D^2,\ D^3,\ \ D^4,\ \ldots$$

are annihilator operators of the following functions

$$k(\text{a constant}),\ \ x,\ \ x^2,\ \ x^3,\ \ldots$$

In general, the differential operator D^n annihilates each of the functions

$$1, x, x^2, \ldots, x^{n-1}$$

Hence, we conclude that the polynomial function

$$c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$$

can be annihilated by finding an operator that annihilates the highest power of x.

Example 6

Find a differential operator that annihilates the polynomial function

$$y = 1 - 5x^2 + 8x^3.$$

Solution

Since $\quad D^4 x^3 = 0,$

Therefore $\quad D^4 y = D^4\left(1 - 5x^2 + 8x^3\right) = 0.$

Hence, D^4 is the differential operator that annihilates the function y.

Note that the functions that are annihilated by an nth-order linear differential operator L are simply those functions that can be obtained from the general solution of the homogeneous differential equation

$$L(y) = 0.$$

Example 7
Consider the homogeneous linear differential equation of order n

$$(D - \alpha)^n y = 0$$

The auxiliary equation of the differential equation is

$$(m - \alpha)^n = 0$$

$$\Rightarrow \qquad m = \alpha, \alpha, \ldots, \alpha \ \ (n \ \text{ times})$$

Therefore, the auxiliary equation has a real root α of multiplicity n. So that the differential equation has the following linearly independent solutions:

$$e^{\alpha x}, xe^{\alpha x}, x^2 e^{\alpha x}, \ldots, x^{n-1} e^{\alpha x}.$$

Therefore, the general solution of the differential equation is

$$y = c_1 e^{\alpha x} + c_2 x e^{\alpha x} + c_3 x^2 e^{\alpha x} + \cdots + c_n x^{n-1} e^{\alpha x}$$

So that the differential operator

$$(D-\alpha)^n$$

annihilates each of the functions

$$e^{\alpha x}, xe^{\alpha x}, x^2 e^{\alpha x}, \ldots, x^{n-1} e^{\alpha x}$$

Hence, as a consequence of the fact that the differentiation can be performed term by term, the differential operator

$$(D-\alpha)^n$$

annihilates the function

$$y = c_1 e^{\alpha x} + c_2 x e^{\alpha x} + c_3 x^2 e^{\alpha x} + \cdots + c_n x^{n-1} e^{\alpha x}$$

Example 8

Find an annihilator operator for the functions

(a) $\qquad f(x) = e^{5x}$

(b) $\qquad g(x) = 4e^{2x} - 6xe^{2x}$

Solution

(a) Since

$$(D-5)e^{5x} = 5e^{5x} - 5e^{5x} = 0.$$

Therefore, the annihilator operator of function f is given by

$$L = D - 5$$

We notice that in this case $\alpha = 5, \ n = 1$.

(b) Similarly

$$(D-2)^2\left(4e^{2x} - 6xe^{2x}\right) = (D^2 - 4D + 4)(4e^{2x}) - (D^2 - 4D + 4)(6xe^{2x})$$

or $\qquad (D-2)^2\left(4e^{2x} - 6xe^{2x}\right) = 32e^{2x} - 32e^{2x} + 48xe^{2x} - 48xe^{2x} + 24e^{2x} - 24e^{2x}$

or $\qquad (D-2)^2\left(4e^{2x} - 6xe^{2x}\right) = 0$

Therefore, the annihilator operator of the function g is given by

$$L = (D-2)^2$$

We notice that in this case $\alpha = 2 = n$.

Example 9

Consider the differential equation

$$\left(D^2 - 2\alpha D + \left(\alpha^2 + \beta^2\right)\right)^n y = 0$$

The auxiliary equation is

$$\left(m^2 - 2\alpha m + \left(\alpha^2 + \beta^2\right)\right)^n = 0$$

$$\Rightarrow \qquad m^2 - 2\alpha m + \left(\alpha^2 + \beta^2\right) = 0$$

Therefore, when $\alpha, \ \beta$ are real numbers, we have from the quadratic formula

$$m = \frac{2\alpha \pm \sqrt{4\alpha^2 - 4\left(\alpha^2 + \beta^2\right)}}{2} = \alpha \pm i\beta$$

Therefore, the auxiliary equation has the following two complex roots of multiplicity n.

$$m_1 = \alpha + i\beta, \quad m_2 = \alpha - i\beta$$

Thus, the general solution of the differential equation is a linear combination of the following linearly independent solutions

$$e^{\alpha x}\cos\beta x, \ xe^{\alpha x}\cos\beta x, \ x^2 e^{\alpha x}\cos\beta x, \ \cdots, \ x^{n-1}e^{\alpha x}\cos\beta x$$

$$e^{\alpha x}\sin\beta x, \ xe^{\alpha x}\sin\beta x, \ x^2 e^{\alpha x}\sin\beta x, \ \cdots, \ x^{n-1}e^{\alpha x}\sin\beta x$$

Hence, the differential operator

$$\left(D^2 - 2\alpha D + \left(\alpha^2 + \beta^2\right)\right)^n$$

is the annihilator operator of the functions

$$e^{\alpha x}\cos\beta x, \ xe^{\alpha x}\cos\beta x, \ x^2 e^{\alpha x}\cos\beta x, \ \cdots, \ x^{n-1}e^{\alpha x}\cos\beta x$$

$$e^{\alpha x}\sin\beta x, \ xe^{\alpha x}\sin\beta x, \ x^2 e^{\alpha x}\sin\beta x, \ \cdots, \ x^{n-1}e^{\alpha x}\sin\beta x$$

Example 10

If we take

$$\alpha = -1, \ \beta = 2, \ n = 1$$

Then the differential operator

$$\left(D^2 - 2\alpha D + \left(\alpha^2 + \beta^2\right)\right)^n$$

becomes $\quad D^2 + 2D + 5$.

Also, it can be verified that

$$\left(D^2 + 2D + 5\right)e^{-x}\cos 2x = 0$$

$$\left(D^2 + 2D + 5\right)e^{-x}\sin 2x = 0$$

Therefore, the linear differential operator

$$D^2 + 2D + 5$$

annihilates the functions

$$y_1(x) = e^{-x}\cos 2x$$
$$y_2(x) = e^{-x}\sin 2x$$

Now, consider the differential equation

$$\left(D^2 + 2D + 5\right)y = 0$$

The auxiliary equation is

$$m^2 + 2m + 5 = 0$$
$$\Rightarrow m = -1 \pm 2i$$

Therefore, the functions

$$y_1(x) = e^{-x}\cos 2x$$
$$y_2(x) = e^{-x}\sin 2x$$

are the two linearly independent solutions of the differential equation

$$\left(D^2 + 2D + 5\right)y = 0,$$

Therefore, the operator also annihilates a linear combination of y_1 and y_2, e.g.

$$5y_1 - 9y_2 = 5e^{-x}\cos 2x - 9e^{-x}\sin 2x.$$

Example 11

If we take

$$\alpha = 0,\ \beta = 1,\ n = 2$$

Then the differential operator

$$\left(D^2 - 2\alpha D + \left(\alpha^2 + \beta^2\right)\right)^n$$

becomes

$$(D^2 + 1)^2 = D^4 + 2D^2 + 1$$

Also, it can be verified that

$$\left(D^4 + 2D^2 + 1\right)\cos x = 0$$
$$\left(D^4 + 2D^2 + 1\right)\sin x = 0$$

and

$$\left(D^4 + 2D^2 + 1\right)x\cos x = 0$$
$$\left(D^4 + 2D^2 + 1\right)x\sin x = 0$$

Therefore, the linear differential operator

$$D^4 + 2D^2 + 1$$

annihilates the functions

$$\cos x, \quad \sin x$$
$$x\cos x, \quad x\sin x$$

Example 12

Taking $\alpha = 0$, $n = 1$, the operator

$$\left(D^2 - 2\alpha D + \left(\alpha^2 + \beta^2\right)\right)^n$$

becomes

$$D^2 + \beta^2$$

Since $\left(D^2 + \beta^2\right)\cos\beta x = -\beta^2 \cos\beta\, x + \beta^2 \cos\beta\, x = 0$

$$\left(D^2 + \beta^2\right)\sin\beta\, x = -\beta^2 \sin\beta\, x + \beta^2 \sin\beta\, x = 0$$

Therefore, the differential operator annihilates the functions

$$f(x) = \cos\beta\, x, \quad g(x) = \sin\beta\, x$$

Note that

- If a linear differential operator with constant coefficients is such that
  $$L(y_1) = 0, \qquad L(y_2) = 0$$
  i.e. the operator L annihilates the functions y_1 and y_2. Then the operator L annihilates their linear combination.
  $$L[c_1 y_1(x) + c_2 y_2(x)] = 0.$$
  This result follows from the linearity property of the differential operator L.

- Suppose that L_1 and L_2 are linear operators with constant coefficients such that
  $$L_1(y_1) = 0, \quad L_2(y_2) = 0$$
  and $\quad L_1(y_2) \neq 0, \quad L_2(y_1) \neq 0$

  then the product of these differential operators $L_1 L_2$ annihilates the linear sum
  $$y_1(x) + y_2(x)$$
  So that $\quad L_1 L_2 [y_1(x) + y_2(x)] = 0$

  To demonstrate this fact we use the linearity property for writing
  $$L_1 L_2 (y_1 + y_2) = L_1 L_2 (y_1) + L_1 L_2 (y_2)$$

  Since $\quad L_1 L_2 = L_2 L_1$

  therefore $\quad L_1 L_2 (y_1 + y_2) = L_2 L_1 (y_1) + L_1 L_2 (y_2)$

  or $\quad L_1 L_2 (y_1 + y_2) = L_2 [L_1 (y_1)] + L_1 [L_2 (y_2)]$

  But we know that $\quad L_1(y_1) = 0, \quad L_2(y_2) = 0$

  Therefore $\quad L_1 L_2 (y_1 + y_2) = L_2 [0] + L_1 [0] = 0$

Example 13

Find a differential operator that annihilates the function

$$f(x) = 7 - x + 6\sin 3x$$

Solution

Suppose that

$$y_1(x) = 7 - x, \quad y_2(x) = 6\sin 3x$$

Then

$$D^2 y_1(x) = D^2(7 - x) = 0$$

$$(D^2 + 9)y_2(x) = (D^2 + 9)\sin 3x = 0$$

Therefore, $D^2(D^2 + 9)$ annihilates the function $f(x)$.

Example 14

Find a differential operator that annihilates the function

$$f(x) = e^{-3x} + xe^x$$

Solution

Suppose that

$$y_1(x) = e^{-3x}, \quad y_2(x) = xe^x$$

Then

$$(D+3)y_1 = (D+3)e^{-3x} = 0,$$

$$(D-1)^2 y_2 = (D-1)^2 xe^x = 0.$$

Therefore, the product of two operators

$$(D+3)(D-1)^2$$

annihilates the given function $f(x) = e^{-3x} + xe^x$

Note that

- The differential operator that annihilates a function is not unique. For example,

  $$(D-5)e^{5x} = 0,$$

  $$(D-5)(D+1)e^{5x} = 0,$$

  $$(D-5)D^2 e^{5x} = 0$$

  Therefore, there are 3 annihilator operators of the functions, namely

  $$(D-5),\ (D-5)(D+1),\ (D-5)D^2$$

- When we seek a differential annihilator for a function, we want the operator of lowest possible order that does the job.

Exercises

Write the given differential equation in the form $L(y) = g(x)$, where L is a differential operator with constant coefficients.

1. $\dfrac{dy}{dx} + 5y = 9\sin x$
2. $4\dfrac{dy}{dx} + 8y = x + 3$

3. $\frac{d^3y}{dx^3} - 4\frac{d^2y}{dx^2} + 5\frac{dy}{dx} = 4x$
4. $\frac{d^3y}{dx^3} - 2\frac{d^2y}{dx^2} + 7\frac{dy}{dx} - 6y = 1 - \sin x$

Factor the given differentiable operator, if possible.

5. $9D^2 - 4$
6. $D^2 - 5$
7. $D^3 + 2D^2 - 13D + 10$
8. $D^4 - 8D^2 + 16$

Verify that the given differential operator annihilates the indicated functions

9. $2D - 1;\quad y = 4e^{x/2}$
10. $D^4 + 64;\quad y = 2\cos 8x - 5\sin 8x$

Find a differential operator that annihilates the given function.

11. $x + 3xe^{6x}$
12. $1 + \sin x$

Lecture 19 Undetermined Coefficients: Annihilator Operator Approach

The method of undetermined coefficients that utilizes the concept of annihilator operator approach is also limited to non-homogeneous linear differential equations

- That have constant coefficients, and
- Where the function $g(x)$ has a specific form.

The form of $g(x)$ **:**The input function $g(x)$ has to have one of the following forms:

- A constant function k.
- A polynomial function
- An exponential function e^{x}
- The trigonometric functions $\sin(\beta x)$, $\cos(\beta x)$
- Finite sums and products of these functions.

Otherwise, we cannot apply the method of undetermined coefficients.

The Method

Consider the following non-homogeneous linear differential equation with constant coefficients of order n

$$a_n \frac{d^n y}{dx^n} + a_{n-1}\frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1 \frac{dy}{dx} + a_0 y = g(x)$$

If L denotes the following differential operator

$$L = a_n D^n + a_{n-1} D^{n-1} + \cdots + a_1 D + a_0$$

Then the non-homogeneous linear differential equation of order n can be written as

$$L(y) = g(x)$$

The function $g(x)$ should consist of finite sums and products of the proper kind of functions as already explained.

The method of undetermined coefficients, annihilator operator approach, for finding a particular integral of the non-homogeneous equation consists of the following steps**:**

Step 1 Write the given non-homogeneous linear differential equation in the form

$$L(y) = g(x)$$

Step 2 Find the complementary solution y_c by finding the general solution of the associated homogeneous differential equation**:**

$$L(y) = 0$$

Step 3 Operate on both sides of the non-homogeneous equation with a differential operator L_1 that annihilates the function g(x).

Step 4 Find the general solution of the higher-order homogeneous differential equation

$$L_1 L(y) = 0$$

Step 5 Delete all those terms from the solution in step 4 that are duplicated in the complementary solution y_c, found in step 2.

Step 6 Form a linear combination y_p of the terms that remain. This is the form of a particular solution of the non-homogeneous differential equation

$$L(\text{y}) = g(x)$$

Step 7 Substitute y_p found in step 6 into the given non-homogeneous linear differential equation

$$L(y) = g(x)$$

Match coefficients of various functions on each side of the equality and solve the resulting system of equations for the unknown coefficients in y_p.

Step 8 With the particular integral found in step 7, form the general solution of the given differential equation as:

$$y = y_c + y_p$$

Example 1

Solve $$\frac{d^2y}{dx^2} + 3\frac{dy}{dx} + 2y = 4x^2.$$

Solution:

Step 1 Since $$\frac{dy}{dx} = Dy, \ \frac{d^2y}{dx^2} = D^2y$$

Therefore, the given differential equation can be written as

$$\left(D^2 + 3D + 2\right)y = 4x^2$$

Step 2 To find the complementary function y_c, we consider the associated homogeneous differential equation

$$\left(D^2 + 3D + 2\right)y = 0$$

The auxiliary equation is

$$m^2 + 3m + 2 = (m+1)(m+2) = 0$$

$$\Rightarrow \quad m = -1, -2$$

Therefore, the auxiliary equation has two distinct real roots.

$$m_1 = -1, \ m_2 = -2,$$

Thus, the complementary function is given by

$$y_c = c_1e^{-x} + c_2e^{-2x}$$

Step 3 In this case the input function is

$$g(x) = 4x^2$$

Further $$D^3g(x) = 4D^3x^2 = 0$$

Therefore, the differential operator D^3 annihilates the function g. Operating on both sides of the equation in step 1, we have

$$D^3(D^2 + 3D + 2)y = 4D^3x^2$$

$$D^3(D^2 + 3D + 2)y = 0$$

This is the homogeneous equation of order 5. Next we solve this higher order equation.

Step 4 The auxiliary equation of the differential equation in step 3 is

$$m^3(m^2+3m+2)=0$$

$$m^3(m+1)(m+2)=0$$

$$m=0,0,0,-1,-2$$

Thus its general solution of the differential equation must be

$$y=c_1+c_2x+c_3x^2+c_4e^{-x}+c_5e^{-2x}$$

Step 5 The following terms constitute y_c

$$c_4e^{-x}+c_5e^{-2x}$$

Therefore, we remove these terms and the remaining terms are

$$c_1+c_2x+c_3x^2$$

Step 6 This means that the basic structure of the particular solution y_p is

$$y_p=A+Bx+Cx^2,$$

Where the constants c_1, c_2 and c_3 have been replaced, with A, *B*, and *C*, respectively.

Step 7 Since

$$y_p=A+Bx+Cx^2$$

$$y'_p=B+2Cx,$$

$$y''_p=2C$$

Therefore

$$y''_p+3y'_p+2y_p=2C+3B+6Cx+2A+2Bx+2Cx^2$$

or

$$y''_p+3y'_p+2y_p=(2C)x^2+(2B+6C)x+(2A+3B+2C)$$

Substituting into the given differential equation, we have

$$(2C)x^2+(2B+6C)x+(2A+3B+2C)=4x^2+0x+0$$

Equating the coefficients of x^2, x and the constant terms, we have

$$2C = 4$$

$$2B + 6C = 0$$

$$2A+3B+2C = 0$$

Solving these equations, we obtain

$$A=7,\quad B=-6,\quad C=2$$

Hence

$$y_p=7-6x+2x^2$$

Step 8 The general solution of the given non-homogeneous differential equation is

$$y=y_c+y_p$$

$$y=c_1e^{-x}+c_2e^{-2x}+7-6x+2x^2.$$

Example 2

Solve $$\frac{d^2y}{dx^2}-3\frac{dy}{dx}=8e^{3x}+4\sin x$$

Solution:

Step 1 Since $$\frac{dy}{dx}=Dy,\ \frac{d^2y}{dx^2}=D^2y$$

Therefore, the given differential equation can be written as

$$\left(D^2-3D\right)y=8e^{3x}+4\sin x$$

Step 2 We first consider the associated homogeneous differential equation to find y_c

The auxiliary equation is

$$m(m-3)=0 \Rightarrow m=0,3$$

Thus the auxiliary equation has real and distinct roots. So that we have

$$y_c=c_1+c_2e^{3x}$$

Step 3 In this case the input function is given by

$$g(x)=8e^{3x}+4\sin x$$

Since $$(D-3)(8e^{3x})=0,\ (D^2+1)(4\sin x)=0$$

Therefore, the operators $D-3$ and D^2+1 annihilate $8e^{3x}$ and $4\sin x$, respectively. So the operator $(D-3)(D^2+1)$ annihilates the input function $g(x)$. This means that

$$(D-3)(D^2+1)g(x)=(D-3)(D^2+1)(8e^{3x}+\sin x)=0$$

We apply $(D-3)(D^2+1)$ to both sides of the differential equation in step 1 to obtain

$$(D-3)(D^2+1)(D^2-3D)y=0\,.$$

This is homogeneous differential equation of order 5.

Step 4 The auxiliary equation of the higher order equation found in step 3 is

$$(m-3)(m^2+1)(m^2-3m)=0$$

$$m(m-3)^2(m^2+1)=0$$

$$\Rightarrow m=0,\ 3,\ 3,\ \pm i$$

Thus, the general solution of the differential equation

$$y=c_1+c_2e^{3x}+c_3xe^{3x}+c_4\cos x+c_5\sin x$$

Step 5 First two terms in this solution are already present in y_c

$$c_1+c_2e^{3x}$$

Therefore, we eliminate these terms. The remaining terms are

$$c_3xe^{3x}+c_4\cos x+c_5\sin x$$

Step 6 Therefore, the basic structure of the particular solution y_p must be

$$y_p=Axe^{3x}+B\cos x+C\sin x$$

The constants c_3, c_4 and c_5 have been replaced with the constants A, B and C, respectively.

Step 7 Since $$y_p = Axe^{3x} + B\cos x + C\sin x$$

Therefore $$y''_p - 3y'_p = 3Ae^{3x} + (-B-3C)\cos x + (3B-C)\sin x$$

Substituting into the given differential equation, we have

$$3Ae^{3x} + (-B-3C)\cos x + (3B-C)\sin x = 8e^{3x} + 4\sin x.$$

Equating coefficients of e^{3x}, $\cos x$ and $\sin x$, we obtain

$$3A = 8,\ -B-3C = 0,\ 3B-C = 4$$

Solving these equations we obtain

$$A = 8/3,\quad B = 6/5,\quad C = -2/5$$

$$y_p = \frac{8}{3}xe^{3x} + \frac{6}{5}\cos x - \frac{2}{5}\sin x.$$

Step 8 The general solution of the differential equation is then

$$y = c_1 + c_2 e^{3x} + \frac{8}{3}xe^{3x} + \frac{6}{5}\cos x - \frac{2}{5}\sin x.$$

Example 3

Solve $$\frac{d^2y}{dx^2} + 8y = 5x + 2e^{-x}.$$

Solution:

Step 1 The given differential equation can be written as

$$(D^2 + 8)y = 5x + 2e^{-x}$$

Step 2 The associated homogeneous differential equation is

$$(D^2 + 8)y = 0$$

Roots of the auxiliary equation are complex

$$m = \pm 2\sqrt{2}\, i$$

Therefore, the complementary function is

$$y_c = c_1\cos 2\sqrt{2}\,x + c_2\sin 2\sqrt{2}\,x$$

Step 3 Since $$D^2 x = 0,\quad (D+1)e^{-x} = 0$$

Therefore the operators D^2 and $D+1$ annihilate the functions $5x$ and $2e^{-x}$. We apply $D^2(D+1)$ to the non-homogeneous differential equation

$$D^2(D+1)(D^2+8)y = 0.$$

This is a homogeneous differential equation of order 5.

Step 4 The auxiliary equation of this differential equation is

$$m^2(m+1)(m^2+8)=0$$
$$\Rightarrow m=0,0,-1,\pm 2\sqrt{2}\,i$$
Therefore, the general solution of this equation must be
$$y=c_1\cos 2\sqrt{2}x+c_2\sin 2\sqrt{2}x+c_3+c_4x+c_5e^{-x}$$

Step 5 Since the following terms are already present in y_c
$$c_1\cos 2\sqrt{2}x+c_2\sin 2\sqrt{2}x$$
Thus we remove these terms. The remaining ones are
$$c_3+c_4x+c_5e^{-x}$$

Step 6 The basic form of the particular solution of the equation is
$$y_p=A+Bx+Ce^{-x}$$
The constants c_3, c_4 and c_5 have been replaced with A, B and C.

Step 7 Since $\quad y_p=A+Bx+Ce^{-x}$

Therefore $\quad y''_p+8y_p=8A+8Bx+9Ce^{-x}$

Substituting in the given differential equation, we have
$$8A+8Bx+9Ce^{-x}=5x+2e^{-x}$$
Equating coefficients of x, e^{-x} and the constant terms, we have
$$A=0,\ B=5/8,\ C=2/9$$
Thus $\quad y_p=\frac{5}{8}x+\frac{2}{9}e^{-x}$

Step 8 Hence, the general solution of the given differential equation is
$$y=y_c+y_p$$
or $\quad y=c_1\cos 2\sqrt{2}x+c_2\sin 2\sqrt{2}x+\frac{5}{8}x+\frac{2}{9}e^{-x}.$

Example 4

Solve $\quad \frac{d^2y}{dx^2}+y=x\cos x-\cos x$

Solution:

Step 1 The given differential equation can be written as
$$(D^2+1)y=x\cos x-\cos x$$

Step 2 Consider the associated differential equation
$$(D^2+1)y=0$$
The auxiliary equation is

$$m^2+1=0 \Rightarrow m=\pm i$$

Therefore $y_c = c_1 \cos x + c_2 \sin x$

Step 3 Since $(D^2+1)^2(x\cos x)=0$

$$(D^2+1)^2 \cos x = 0 \ ; \quad \because x \neq 0$$

Therefore, the operator $(D^2+1)^2$ annihilates the input function

$$x\cos x - \cos x$$

Thus operating on both sides of the non-homogeneous equation with $(D^2+1)^2$, we have

$$(D^2+1)^2(D^2+1)y=0$$

or $(D^2+1)^3 y = 0$

This is a homogeneous equation of order 6.

Step 4 The auxiliary equation of this higher order differential equation is

$$(m^2+1)^3=0 \Rightarrow m=i,i,i,-i,-i,-i$$

Therefore, the auxiliary equation has complex roots i, and $-i$ both of multiplicity 3. We conclude that

$$y = c_1\cos x + c_2 \sin x + c_3 x\cos x + c_4 x\sin x + c_5 x^2\cos x + c_6 x^2 \sin x$$

Step 5 Since first two terms in the above solution are already present in y_c

$$c_1\cos x + c_2 \sin x$$

Therefore, we remove these terms.

Step 6 The basic form of the particular solution is

$$y_p = Ax\cos x + Bx\sin x + Cx^2\cos x + Ex^2 \sin x$$

Step 7 Since $y_p = Ax\cos x + Bx\sin x + Cx^2\cos x + Ex^2 \sin x$

Therefore

$$y_p'' + y_p = 4Ex\cos x - 4Cx\sin x + (2B+2C)\cos x + (-2A+2E)\sin x$$

Substituting in the given differential equation, we obtain

$$4Ex\cos x - 4Cx\sin x + (2B+2C)\cos x + (-2A+2E)\sin x = x\cos x - \cos x$$

Equating coefficients of $x\cos x, x\sin x, \cos x$ and $\sin x$, we obtain

$$4E = 1, \quad -4C = 0$$

$$2B+2C=-1, \quad -2A+2E=0$$

Solving these equations we obtain

$$A=1/4,\ B=-1/2,\ C=0,\ E=1/4$$

Thus $y_p = \frac{1}{4}x\cos x - \frac{1}{2}x\sin x + \frac{1}{4}x^2 \sin x$

Step 8 Hence the general solution of the differential equation is

$$y = c_1\cos x + c_2\sin x + \frac{1}{4}x\cos x - \frac{1}{2}x\sin x + \frac{1}{4}x^2\sin x.$$

Example 5
Determine the form of a particular solution for

$$\frac{d^2y}{dx^2}-2\frac{dy}{dx}+y=10e^{-2x}\cos x$$

Solution

Step 1 The given differential equation can be written as

$$(D^2-2D+1)y=10e^{-2x}\cos x$$

Step 2 To find the complementary function, we consider

$$y''-2y'+y=0$$

The auxiliary equation is

$$m^2-2m+1=0 \Rightarrow (m-1)^2=0 \Rightarrow m=1,1$$

The complementary function for the given equation is

$$y_c=c_1e^x+c_2xe^x$$

Step 3 Since $(D^2+4D+5)e^{-2x}\cos x=0$
Applying the operator (D^2+4D+5) to both sides of the equation, we have

$$(D^2+4D+5)(D^2-2D+1)y=0$$

This is homogeneous differential equation of order 4.

Step 4 The auxiliary equation is

$$(m^2+4m+5)(m^2-2m+1)=0$$

$$\Rightarrow m=-2\pm i,1,1$$

Therefore, general solution of the 4th order homogeneous equation is

$$y=c_1e^x+c_2xe^x+c_3e^{-2x}\cos x+c_4e^{-2x}\sin x$$

Step 5 Since the terms $c_1e^x+c_2xe^x$ are already present in y_c, therefore, we remove these and the remaining terms are $c_3e^{-2x}\cos x+c_4e^{-2x}\sin x$

Step 6 Therefore, the form of the particular solution of the non-homogeneous equation is

$$\therefore \qquad y_p=Ae^{-2x}\cos x+Be^{-2x}\sin x$$

Note that the steps 7 and 8 are not needed, as we don't have to solve the given differential equation.

Example 6
Determine the form of a particular solution for

$$\frac{d^3y}{dx^3}-4\frac{d^2y}{dx^2}+4\frac{dy}{dx}=5x^2-6x+4x^2e^{2x}+3e^{5x}.$$

Solution:

Step 1 The given differential can be rewritten as

$$\left(D^3 - 4D^2 + 4D\right)y = 5x^2 - 6x + 4x^2 e^{2x} + 3e^{5x}$$

Step 2 To find the complementary function, we consider the equation

$$\left(D^3 - 4D^2 + 4D\right)y = 0$$

The auxiliary equation is

$$m^3 - 4m^2 + 4m = 0$$
$$m(m^2 - 4m + 4) = 0$$
$$m(m-2)^2 = 0 \Rightarrow m = 0, 2, 2$$

Thus the complementary function is

$$y_c = c_1 + c_2 e^{2x} + c_3 x e^{2x}$$

Step 3 Since $\qquad g(x) = 5x^2 - 6x + 4x^2 e^{2x} + 3e^{5x}$

Further $\qquad D^3(5x^2 - 6x) = 0$

$$(D-2)^3 x^2 e^{2x} = 0$$
$$(D-5)e^{5x} = 0$$

Therefore the following operator must annihilate the input function $g(x)$. Therefore, applying the operator $D^3(D-2)^3(D-5)$ to both sides of the non-homogeneous equation, we have

$$D^3(D-2)^3(D-5)(D^3 - D^2 + 4D)y = 0$$

or $\qquad D^4(D-2)^5(D-5)y = 0$

This is homogeneous differential equation of order 10.

Step 4 The auxiliary equation for the 10th order differential equation is

$$m^4(m-2)^5(m-5) = 0$$
$$\Rightarrow m = 0, 0, 0, 0, 2, 2, 2, 2, 2, 5$$

Hence the general solution of the 10th order equation is

$$y = c_1 + c_2 x + c_3 x^2 + c_4 x^3 + c_5 e^{2x} + c_6 x e^{2x} + c_7 x^2 e^{2x} + c_8 x^3 e^{2x} + c_9 x^4 e^{2x} + c_{10} e^{5x}$$

Step 5 Since the following terms constitute the complementary function y_c, we remove these $\qquad c_1 + c_5 e^{2x} + c_6 x e^{2x}$

Thus the remaining terms are

$$c_2 x + c_3 x^2 + c_4 x^3 + c_7 x^2 e^{2x} + c_8 x^3 e^{2x} + c_9 x^4 e^{2x} + c_{10} e^{5x}$$

Hence, the form of the particular solution of the given equation is

$$y_p = Ax + Bx^2 + Cx^3 + Ex^2 e^{2x} + Fx^3 e^{2x} + Gx^4 e^{2x} + He^{5x}$$

Exercise

Solve the given differential equation by the undetermined coefficients.

1. $2y'' - 7y' + 5y = -29$
2. $y'' + 3y' = 4x - 5$
3. $y'' + 2y' + 2y = 5e^{6x}$
4. $y'' + 4y = 4\cos x + 3\sin x - 8$
5. $y'' + 2y' + y = x^2 e^{-x}$
6. $y'' + y = 4\cos x - \sin x$
7. $y''' - y'' + y' - y = xe^x - e^{-x} + 7$
8. $y'' + y = 8\cos 2x - 4\sin x$, $y(\pi/2) = -1$, $y'(\pi/2) = 0$
9. $y''' - 2y'' + y' = xe^x + 5$, y(0)=2, $y'(0) = 2$, $y''(0) = -1$
10. $y^{(4)} - y''' = x + e^x$, y(0)=0, $y'(0) = 0$, $y''(0) = 0$, $y'''(0) = 0$

Lecture 20 Variation of Parameters

Recall

- That a non-homogeneous linear differential equation with constant coefficients is an equation of the form

$$a_n \frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1 \frac{dy}{dx} + a_0 y = g(x)$$

- The general solution of such an equation is given by

General Solution = Complementary Function + Particular Integral

- Finding the complementary function has already been completely discussed.

- In the last two lectures, we learnt how to find the particular integral of the non-homogeneous equations by using the undetermined coefficients.

- That the general solution of a linear first order differential equation of the form

$$\frac{dy}{dx} + P(x)y = f(x)$$

is given by $$y = e^{-\int P dx} . \int e^{\int P dx} f(x) dx + c_1 e^{-\int P dx}$$

Note that

- In this last equation, the 2nd term

$$y_c = c_1 e^{-\int P dx}$$

is solution of the associated homogeneous equation:

$$\frac{dy}{dx} + P(x)y = 0$$

- Similarly, the 1st term

$$y_p = e^{-\int P dx} . \int e^{\int P dx} . f(x) dx$$

is a particular solution of the first order non-homogeneous linear differential equation.

- Therefore, the solution of the first order linear differential equation can be written in the form

$$y = y_c + y_p$$

In this lecture, we will use the variation of parameters to find the particular integral of the non-homogeneous equation.

The Variation of Parameters

First order equation

The particular solution y_p of the first order linear differential equation is given by

$$y_p = e^{-\int P dx} . \int e^{\int P dx} . f(x) dx$$

This formula can also be derived by another method, known as the variation of parameters. The basic procedure is same as discussed in the lecture on construction of a second solution

Since $$y_1 = e^{-\int P dx}$$

is the solution of the homogeneous differential equation

$$\frac{dy}{dx} + P(x)y = 0,$$

and the equation is linear. Therefore, the general solution of the equation is

$$y = c_1 y_1(x)$$

The variation of parameters consists of finding a function $u_1(x)$ such that

$$y_p = u_1(x)\, y_1(x)$$

is a particular solution of the non-homogeneous differential equation

$$\frac{dy}{dx} + P(x)\, y = f(x)$$

Notice that the parameter c_1 has been replaced by the variable u_1. We substitute y_p in the given equation to obtain

$$u_1\left[\frac{dy_1}{dx} + P(x)y_1\right] + y_1\frac{du_1}{dx} = f(x)$$

Since y_1 is a solution of the non-homogeneous differential equation. Therefore we must have

$$\frac{dy_1}{dx} + P(x)\, y_1 = 0$$

So that we obtain

$$\therefore \qquad y_1\frac{du_1}{dx} = f(x)$$

This is a variable separable equation. By separating the variables, we have

$$du_1 = \frac{f(x)}{y_1(x)}dx$$

Integrating the last expression *w.r.to* x, we obtain

$$u_1(x) = \int \frac{f(x)}{y_1} dx = \int e^{\int P dx} \cdot f(x) dx$$

Therefore, the particular solution y_p of the given first-order differential equation is .

$$y = u_1(x) y_1$$

or $$y_p = e^{-\int P dx} . \int e^{\int P dx} . f(x) dx$$

$$u_1 = \int \frac{f(x)}{y_1(x)} dx$$

Second Order Equation

Consider the 2nd order linear non-homogeneous differential equation

$$a_2(x) y'' + a_1(x) y' + a_0(x) y = g(x)$$

By dividing with $a_2(x)$, we can write this equation in the standard form

$$y'' + P(x) y' + Q(x) y = f(x)$$

The functions $P(x)$, $Q(x)$ and $f(x)$ are continuous on some interval I. For the complementary function we consider the associated homogeneous differential equation

$$y'' + P(x) y' + Q(x) y = 0$$

Complementary function

Suppose that y_1 and y_2 are two linearly independent solutions of the homogeneous equation. Then y_1 and y_2 form a fundamental set of solutions of the homogeneous equation on the interval I. Thus the complementary function is

$$y_c = c_1 y_1(x) + c_2 y_2(x)$$

Since y_1 and y_2 are solutions of the homogeneous equation. Therefore, we have

$$y_1'' + P(x) y_1' + Q(x) y_1 = 0$$

$$y_2'' + P(x) y_2' + Q(x) y_2 = 0$$

Particular Integral

For finding a particular solution y_p, we replace the parameters c_1 and c_2 in the complementary function with the unknown variables $u_1(x)$ and $u_2(x)$. So that the assumed particular integral is

$$y_p = u_1(x) y_1(x) + u_2(x) y_2(x)$$

Since we seek to determine two unknown functions u_1 and u_2, we need two equations involving these unknowns. One of these two equations results from substituting the assumed y_p in the given differential equation. We impose the other equation to simplify the first derivative and thereby the 2nd derivative of y_p.

$$y_p' = u_1 y_1' + y_1 u_1' + u_2 y_2' + u_2' y_2 = u_1 y_1' + u_2 y_2' + u_1' y_1 + u_2' y_2$$

To avoid 2nd derivatives of u_1 and u_2, we impose the condition

$$u_1' y_1 + u_2' y_2 = 0$$

Then $$y_p' = u_1 y_1' + u_2 y_2'$$

So that

$$y_p'' = u_1 y_1'' + u_1' y_1' + u_2 y_2'' + u_2' y_2'$$

Therefore

$$y_p'' + P y_p' + Q y_p = u_1 y_1'' + u_1' y_1' + u_2 y_2'' + u_2' y_2' + P u_1 y_1' + P u_2 y_2' + Q u_1 y_1 + Q u_2 y_2$$

Substituting in the given non-homogeneous differential equation yields

$$u_1 y_1'' + u_1' y_1' + u_2 y_2'' + u_2' y_2' + P u_1 y_1' + P u_2 y_2' + Q u_1 y_1 + Q u_2 y_2 = f(x)$$

or $$u_1[y_1'' + P y_1' + Q y_1] + u_2[y_2'' + P y_2' + Q y_2] + u_1' y_1' + u_2' y_2' = f(x)$$

Now making use of the relations

$$y_1'' + P(x) y_1' + Q(x) y_1 = 0$$

$$y_2'' + P(x) y_2' + Q(x) y_2 = 0$$

we obtain

$$u_1' y_1' + u_2' y_2' = f(x)$$

Hence u_1 and u_2 must be functions that satisfy the equations

$$u_1' y_1 + u_2' y_2 = 0$$

$$u_1' y_1' + u_2' y_2' = f(x)$$

By using the Cramer's rule, the solution of this set of equations is given by

$$u_1' = \frac{W_1}{W}, \qquad u_2' = \frac{W_2}{W}$$

Where W, W_1 and W_2 denote the following determinants

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}, \qquad W_1 = \begin{vmatrix} 0 & y_2 \\ f(x) & y_2' \end{vmatrix}, \qquad W_2 = \begin{vmatrix} y_1 & 0 \\ y_1' & f(x) \end{vmatrix}$$

The determinant W can be identified as the Wronskian of the solutions y_1 and y_2. Since the solutions y_1 and y_2 are linearly independent on I. Therefore

$$W(y_1(x), y_2(x)) \neq 0, \ \forall \ x \in I.$$

Now integrating the expressions for u_1' and u_2', we obtain the values of u_1 and u_2, hence the particular solution of the non-homogeneous linear differential equation.

Summary of the Method

To solve the 2nd order non-homogeneous linear differential equation

$$a_2 y'' + a_1 y' + a_0 y = g(x),$$

using the variation of parameters, we need to perform the following steps:

Step 1 We find the complementary function by solving the associated homogeneous differential equation

$$a_2 y'' + a_1 y' + a_0 y = 0$$

Step 2 If the complementary function of the equation is given by

$$y_c = c_1 y_1 + c_2 y_2$$

then y_1 and y_2 are two linearly independent solutions of the homogeneous differential equation. Then compute the Wronskian of these solutions.

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}$$

Step 3 By dividing with a_2, we transform the given non-homogeneous equation into the standard form

$$y'' + P(x)y' + Q(x)y = f(x)$$

and we identify the function $f(x)$.

Step 4 We now construct the determinants W_1 and W_2 given by

$$W_1 = \begin{vmatrix} 0 & y_2 \\ f(x) & y_2' \end{vmatrix}, \quad W_2 = \begin{vmatrix} y_1 & 0 \\ y_1' & f(x) \end{vmatrix}$$

Step 5 Next we determine the derivatives of the unknown variables u_1 and u_2 through the relations

$$u_1' = \frac{W_1}{W}, \quad u_2' = \frac{W_2}{W}$$

Step 6 Integrate the derivatives u_1' and u_2' to find the unknown variables u_1 and u_2. So that

$$u_1 = \int \frac{W_1}{W}\, dx, \quad u_2 = \int \frac{W_2}{W}\, dx$$

Step 7 Write a particular solution of the given non-homogeneous equation as

$$y_p = u_1 y_1 + u_2 y_2$$

Step 8 The general solution of the differential equation is then given by

$$y = y_c + y_p = c_1 y_1 + c_2 y_2 + u_1 y_1 + u_2 y_2.$$

Constants of Integration

We don't need to introduce the constants of integration, when computing the indefinite integrals in step 6 to find the unknown functions of u_1 and u_2. For, if we do introduce these constants, then

$$y_p = (u_1 + a_1)y_1 + \quad (u_2 + b_1)y_2$$

So that the general solution of the given non-homogeneous differential equation is

$$y = y_c + y_p = c_1 y_1 + c_2 y_2 + (u_1 + a_1)y_1 + (u_2 + b_1)y_2$$

or

$$y = (c_1 + a_1)y_1 + (c_2 + b_1)y_2 + u_1 y_1 + u_2 y_2$$

If we replace $c_1 + a_1$ with C_1 and $c_2 + b_1$ with C_2, we obtain

$$y = C_1 y_1 + C_2 y_2 + u_1 y_1 + u_2 y_2$$

This does not provide anything new and is similar to the general solution found in step 8, namely

$$y = c_1 y_1 + c_2 y_2 + u_1 y_1 + u_2 y_2$$

Example 1

Solve $$y'' - 4y' + 4y = (x+1)e^{2x}.$$

Solution:

Step 1 To find the complementary function

$$y'' - 4y' + 4y = 0$$

Put $$y = e^{mx}, y' = me^{mx}, y'' = m^2 e^{mx}$$

Then the auxiliary equation is

$$m^2 - 4m + 4 = 0$$

$$(m-2)^2 = 0 \Rightarrow m = 2, 2$$

Repeated real roots of the auxiliary equation

$$y_c = c_1 e^{2x} \quad + \quad c_2 x e^{2x}$$

Step 2 By the inspection of the complementary function y_c, we make the identification

$$y_1 = e^{2x} \text{ and } y_2 = xe^{2x}$$

Therefore $$W(y_1, y_2) = W\left(e^{2x}, xe^{2x}\right) = \begin{vmatrix} e^{2x} & xe^{2x} \\ 2e^{2x} & 2xe^{2x} + e^{2x} \end{vmatrix} = e^{4x} \neq 0, \forall x$$

Step 3 The given differential equation is

$$y'' - 4y' + 4y = (x+1)e^{2x}$$

Since this equation is already in the standard form

$$y'' + P(x)y' + Q(x)y = f(x)$$

Therefore, we identify the function $f(x)$ as

$$f(x) = (x+1)e^{2x}$$

Step 4 We now construct the determinants

$$W_1 = \begin{vmatrix} 0 & xe^{2x} \\ (x+1)e^{2x} & 2xe^{2x} + e^{2x} \end{vmatrix} = -(x+1)xe^{4x}$$

$$W_2 = \begin{vmatrix} e^{2x} & 0 \\ 2e^{2x} & (x+1)e^{2x} \end{vmatrix} = (x+1)e^{4x}$$

Step 5 We determine the derivatives of the functions u_1 and u_2 in this step

$$u_1' = \frac{W_1}{W} = -\frac{(x+1)xe^{4x}}{e^{4x}} = -x^2 - x$$

$$u_2' = \frac{W_2}{W} = \frac{(x+1)e^{4x}}{e^{4x}} = x+1$$

Step 6 Integrating the last two expressions, we obtain

$$u_1 = \int (-x^2 - x)dx = -\frac{x^3}{3} - \frac{x^2}{2}$$

$$u_2 = \int (x+1)\,dx \quad = \frac{x^2}{2} + x.$$

Remember! We don't have to add the constants of integration.

Step 7 Therefore, a particular solution of then given differential equation is

$$y_p = \left(-\frac{x^3}{3} - \frac{x^2}{2}\right)e^{2x} + \left(\frac{x^2}{2} + x\right)xe^{2x}$$

or $$y_p = \left(\frac{x^3}{6} + \frac{x^2}{2}\right)e^{2x}$$

Step 8 Hence**,** the general solution of the given differential equation is

$$y = y_c + y_p = c_1e^{2x} + c_2xe^{2x} + \left(\frac{x^3}{6} + \frac{x^2}{2}\right)e^{2x}$$

Example 2

Solve $$4y'' + 36y = \csc 3x.$$

Solution:

Step 1 To find the complementary function we solve the associated homogeneous differential equation

$$4y''+36y=0 \Rightarrow y''+9y=0$$

The auxiliary equation is

$$m^2+9=0 \Rightarrow m=\pm 3i$$

Roots of the auxiliary equation are complex. Therefore, the complementary function is

$$y_c=c_1\cos 3x+c_2\sin 3x$$

Step 2 From the complementary function, we identify

$$y_1=\cos 3x,\ \ y_2=\sin 3x$$

as two linearly independent solutions of the associated homogeneous equation. Therefore

$$W(\cos 3x, \sin 3x)=\begin{vmatrix}\cos 3x & \sin 3x\\ -3\sin 3x & 3\cos 3x\end{vmatrix}=3$$

Step 3 By dividing with 4, we put the given equation in the following standard form

$$y''+9y=\frac{1}{4}\csc 3x.$$

So that we identify the function $f(x)$ as

$$f(x)=\frac{1}{4}\csc 3x$$

Step 4 We now construct the determinants W_1 and W_2

$$W_1=\begin{vmatrix}0 & \sin 3x\\ \frac{1}{4}\csc 3x & 3\cos 3x\end{vmatrix}=-\frac{1}{4}\csc 3x\cdot\sin 3x=-\frac{1}{4}$$

$$W_2=\begin{vmatrix}\cos 3x & 0\\ -3\sin 3x & \frac{1}{4}\csc 3x\end{vmatrix}=\frac{1}{4}\frac{\cos 3x}{\sin 3x}$$

Step 5 Therefore, the derivatives u_1' and u_2' are given by

$$u_1'=\frac{W_1}{W}=-\frac{1}{12},\quad u_2'=\frac{W_2}{W}=\frac{1}{12}\frac{\cos 3x}{\sin 3x}$$

Step 6 Integrating the last two equations $w.r.to\ x$, we obtain

$$u_1 = -\frac{1}{12}x \quad \text{and} \quad u_2 = \frac{1}{36}\ln|\sin 3x|$$

Note that no constants of integration have been added.

Step 7 The particular solution of the non-homogeneous equation is

$$y_p = -\frac{1}{12}x\cos 3x + \frac{1}{36}(\sin 3x)\ln|\sin 3x|$$

Step 8 Hence, the general solution of the given differential equation is

$$y = y_c + y_p = c_1\cos 3x + c_2\sin 3x - \frac{1}{12}x\cos 3x + \frac{1}{36}(\sin 3x)\ln|\sin 3x|$$

Example 3

Solve $\qquad y'' - y = \frac{1}{x}.$

Solution:

Step 1 For the complementary function consider the associated homogeneous equation

$$y'' - y = 0$$

To solve this equation we put

$$y = e^{mx}, y' = m e^{mx}, y'' = m^2 e^{mx}$$

Then the auxiliary equation is:

$$m^2 - 1 = 0 \Rightarrow m = \pm 1$$

The roots of the auxiliary equation are real and distinct. Therefore, the complementary function is

$$y_c = c_1 e^x + c_2 e^{-x}$$

Step 2 From the complementary function we find

$$y_1 = e^x, \quad y_2 = e^{-x}$$

The functions y_1 and y_2 are two linearly independent solutions of the homogeneous equation. The Wronskian of these solutions is

$$W\left(e^x, e^{-x}\right) = \begin{vmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{vmatrix} = -2$$

Step 3 The given equation is already in the standard form

$$y'' + p(x)y' + Q(x)y = f(x)$$

Here $\qquad f(x) = \frac{1}{x}$

Step 4 We now form the determinants

$$W_1 = \begin{vmatrix} 0 & e^{-x} \\ 1/x & -e^{-x} \end{vmatrix} = -e^{-x}(1/x)$$

$$W_2 = \begin{vmatrix} e^{x} & 0 \\ e^{x} & 1/x \end{vmatrix} = e^{x}(1/x)$$

Step 5 Therefore, the derivatives of the unknown functions u_1 and u_2 are given by

$$u_1' = \frac{W_1}{W} = -\frac{e^{-x}(1/x)}{-2} = \frac{e^{-x}}{2x}$$

$$u_2' = \frac{W_2}{W} = \frac{e^{x}(1/x)}{-2} = -\frac{e^{x}}{2x}$$

Step 6 We integrate these two equations to find the unknown functions u_1 and u_2.

$$u_1 = \frac{1}{2}\int \frac{e^{-x}}{x}dx, \quad u_2 = -\frac{1}{2}\int \frac{e^{x}}{x}dx$$

The integrals defining u_1 and u_2 cannot be expressed in terms of the elementary functions and it is customary to write such integral as:

$$u_1 = \frac{1}{2}\int_{x_\circ}^{x} \frac{e^{-t}}{t}dt, \quad u_2 = -\frac{1}{2}\int_{x_\circ}^{x} \frac{e^{t}}{t}dt$$

Step 7 A particular solution of the non-homogeneous equations is

$$y_p = \frac{1}{2}e^{x}\int_{x_\circ}^{x} \frac{e^{-t}}{t}dt - \frac{1}{2}e^{-x}\int_{x_\circ}^{x} \frac{e^{t}}{t}dt$$

Step 8 Hence, the general solution of the given differential equation is

$$y = y_c + y_p = c_1e^{x} + c_2e^{-x} + \frac{1}{2}e^{x}\int_{x_\circ}^{x} \frac{e^{-t}}{t}dt - \frac{1}{2}e^{-x}\int_{x_\circ}^{x} \frac{e^{t}}{t}dt$$

Exercise

Solve the differential equations by variations of parameters.

1. $y'' + y = \tan x$
2. $y'' + y = \sec x \tan x$
3. $y'' + y = \sec^2 x$
4. $y'' - y = 9x/e^{3x}$
5. $y'' - 2y' + y = e^x/(1 + x^2)$
6. $4y'' - 4y' + y = e^{x/2}\sqrt{1 - x^2}$
7. $y''' + 4y' = \sec 2x$
8. $2y''' - 6y'' = x^2$

Solve the initial value problems.

9. $2y'' + y' - y = x + 1$
10. $y'' - 4y' + 4y = (12x^2 - 6x)e^{2x}$

Mth303 (Mathematical Method) Part II

Table of contents

Lecture 1

Introduction and Overview

What is Algebra?

History:

Algebra is named in honor of Mohammed Ibn-e- Musa al-Khowârizmî. Around 825, he wrote a book entitled Hisb al-jabr u'l muqubalah, ("the science of reduction and cancellation"). His book, Al-jabr, presented rules for solving equations.

Algebra is a branch of Mathematics that uses mathematical statements to describe relationships between things that vary over time. These variables include things like the relationship between supply of an object and its price. When we use a mathematical statement to describe a relationship, we often use letters to represent the quantity that varies, since it is not a fixed amount. These letters and symbols are referred to as variables.

Algebra is a part of mathematics in which unknown quantities are found with the help of relations between the unknown and known.

In algebra, letters are sometimes used in place of numbers.

The mathematical statements that describe relationships are expressed using algebraic terms, expressions, or equations (mathematical statements containing letters or symbols to represent numbers). Before we use algebra to find information about these kinds of relationships, it is important to first cover some basic terminology.

Algebraic Term:

The basic unit of an algebraic expression is a term. In general, a term is either a product of a number and with one or more variables.

For example 4x is an algebraic term in which 4 is coefficient and x is said to be variable.

Study of Algebra:

Today, algebra is the study of the properties of operations on numbers. Algebra generalizes arithmetic by using symbols, usually letters, to represent numbers or unknown quantities. Algebra is a problem-solving tool. It is like a tractor, which is a farmer's tool. Algebra is the mathematician's tool for solving problems. Algebra has applications to every human endeavor. From art to medicine to zoology, algebra can be a tool. People who say that they will never use algebra are people who do not know about algebra. Learning algebra is a bit like learning to read and write. If you truly learn algebra, you will use it. Knowledge of algebra can give you more power to solve problems and accomplish what you want in life. Algebra is a mathematicians' shorthand!

Algebraic Expressions:

An expression is a collection of numbers, variables, and +ve sign or –ve sign, of operations that must make mathematical and logical behaviour.

For example $8x^2 + 9x - 1$ is an algebraic expression.

What is Linear Algebra?

One of the most important problems in mathematics is that of solving systems of linear equations. It turns out that such problems arise frequently in applications of mathematics in the physical sciences, social sciences, and engineering. Stated in its simplest terms, the world is not linear, but the only problems that we know how to solve are the linear ones. What this often means is that only recasting them as linear systems can solve non-linear problems. A comprehensive study of linear systems leads to a rich, formal structure to analytic geometry and solutions to 2x2 and 3x3 systems of linear equations learned in previous classes.

It is exactly what the name suggests. Simply put, it is the algebra of systems of linear equations. While you could solve a system of, say, five linear equations involving five unknowns, it might not take a finite amount of time. With linear algebra we develop techniques to solve m linear equations and n unknowns, or show when no solution exists. We can even describe situations where an infinite number of solutions exist, and describe them geometrically.

Linear algebra is the study of linear sets of equations and their transformation properties.

Linear algebra, sometimes disguised as matrix theory, considers sets and functions, which preserve linear structure. In practice this includes a very wide portion of mathematics! Thus linear algebra includes axiomatic treatments, computational matters, algebraic structures, and even parts of geometry; moreover, it provides tools used for analyzing differential equations, statistical processes, and even physical phenomena.

Linear Algebra consists of studying matrix calculus. It formalizes and gives geometrical interpretation of the resolution of equation systems. It creates a formal link between matrix calculus and the use of linear and quadratic transformations. It develops the idea of trying to solve and analyze systems of linear equations.

Applications of Linear algebra:

Linear algebra makes it possible to work with large arrays of data. It has many applications in many diverse fields, such as

- Computer Graphics,
- Electronics,
- Chemistry,
- Biology,
- Differential Equations,
- Economics,
- Business,
- Psychology,
- Engineering,
- Analytic Geometry,
- Chaos Theory,
- Cryptography,
- Fractal Geometry,
- Game Theory,
- Graph Theory,
- Linear Programming,
- Operations Research

It is very important that the theory of linear algebra is first understood, the concepts are cleared and then computation work is started. Some of you might want to just use the computer, and skip the theory and proofs, but if you don't understand the theory, then it can be very hard to appreciate and interpret computer results.

Why using Linear Algebra?

Linear Algebra allows for formalizing and solving many typical problems in different engineering topics. It is generally the case that (input or output) data from an experiment is given in a discrete form (discrete measurements). Linear Algebra is then useful for solving problems in such applications in topics such as Physics, Fluid Dynamics, Signal Processing and, more generally Numerical Analysis.

Linear algebra is not like algebra. It is mathematics of linear spaces and linear functions. So we have to know the term "linear" a lot. Since the concept of linearity is fundamental to any type of mathematical analysis, this subject lays the foundation for many branches of mathematics.

Objects of study in linear algebra:

Linear algebra merits study at least because of its ubiquity in mathematics and its applications. The broadest range of applications is through the concept of vector spaces and their transformations. These are the central objects of study in linear algebra

1. The solutions of homogeneous systems of linear equations form paradigm examples of vector spaces. Of course they do not provide the only examples.
2. The vectors of physics, such as force, as the language suggests, also provide paradigmatic examples.
3. Binary code is another example of a vector space, a point of view that finds application in computer sciences.
4. Solutions to specific systems of differential equations also form vector spaces.
5. Statistics makes extensive use of linear algebra.
6. Signal processing makes use of linear algebra.
7. Vector spaces also appear in number theory in several places, including the study of field extensions.

8. Linear algebra is part of and motivates much abstract algebra. Vector spaces form the basis from which the important algebraic notion of module has been abstracted.
9. Vector spaces appear in the study of differential geometry through the tangent bundle of a manifold.
10. Many mathematical models, especially discrete ones, use matrices to represent critical relationships and processes. This is especially true in engineering as well as in economics and other social sciences.

There are two principal aspects of linear algebra: theoretical and computational. A major part of mastering the subject consists in learning how these two aspects are related and how to move from one to the other.

Many computations are similar to each other and therefore can be confusing without reasonable level of grasp of their theoretical context and significance. It will be very tempting to draw false conclusions.

On the other hand, while many statements are easier to express elegantly and to understand from a purely theoretical point of view, to apply them to concrete problems you will need to "get your hands dirty". Once you have understood the theory sufficiently and appreciate the methods of computation, you will be well placed to use software effectively, where possible, to handle large or complex calculations.

Course Segments:

The course is covered in 45 Lectures spanning over six major segments, which are given below;

1. Linear Equations
2. Matrix Algebra
3. Determinants
4. Vector spaces
5. Eigen values and Eigenvectors, and
6. Orthogonal sets

Course Objectives:

The main purpose of the course is to introduce the concept of linear algebra, explain the underline theory, explain the computational techniques and then try to apply them on real life problems. Broad course objectives are as under;

- To master techniques for solving systems of linear equations
- To introduce matrix algebra as a generalization of the single-variable algebra of high school.
- To build on the background in Euclidean space and formalize it with vector space theory.
- To develop an appreciation for how linear methods are used in a variety of applications.
- To relate linear methods to other areas of mathematics such as calculus and, differential equations.

Recommended Books and Supported Material:

I am indebted to several authors whose books I have freely used to prepare the lectures that follow. The lectures are based on the material taken from the books mentioned below.

1. **Linear Algebra and its Applications** (3rd Edition) by David C. Lay.
2. **Contemporary Linear Algebra** by Howard Anton and Robert C. Busby.
3. **Introductory Linear Algebra** (8th Edition) by Howard Anton and Chris Rorres.
4. **Introduction to Linear Algebra** (3rd Edition) by L. W. Johnson, R.D. Riess and J.T. Arnold.
5. **Linear Algebra** (3rd Edition) by S. H. Friedberg, A.J. Insel and L.E. Spence.
6. **Introductory Linear Algebra with Applications** (6th Edition) by B. Kolman.

I have taken the structure of the course as proposed in the book of David C. Lay. I would be following this book. I suggest that the students purchase this book, which is easily available in the market and also does not cost much. For further study and supplement, students can consult any of the above mentioned books.

I strongly suggest that the students also browse on the Internet; there is plenty of support material available. In particular, I would suggest the website of David C. Lay; *www.laylinalgebra.com*, where the entire material, study guide, transparencies are readily available. Another very useful website is *www.wiley.com/college/anton*, which contains a variety of useful material including the data sets. A number of other books are also available in the market and on the internet with free access.

I will try to keep the treatment simple and straight. The lectures will be presented in simple Urdu and easy English. These lectures are supported by the handouts in the form of lecture notes. The theory will be explained with the help of examples. There will be enough exercises to practice with. Students are advised to follow the course on daily basis and do the exercises regularly.

Schedule and Assessment:

The course will be spread over 45 lectures. Lectures one and two will be introductory and the Lecture 45 will be the summary. The first two lectures will lay the foundations and would provide the overview of the course. These will be important from the concept point of view. I suggest that these two lectures should be viewed again and again.

The course will be interesting and enjoyable, if the student follow it regularly and complete the exercises as they come along. To follow the tradition of a semester system or of a term system, there will be a series of assignments (Max eight assignments) and a mid term exam. Finally there will be terminal examination.

The assignments have weights and therefore they have to be taken seriously.

Lecture 2
Background

Introduction to Matrices

Matrix: A matrix is a collection of numbers or functions arranged into rows and columns.

Matrices are denoted by capital letters $A, B, \ldots, Y, Z$. The numbers or functions are called elements of the matrix. The elements of a matrix are denoted by small letters $a, b, \ldots, y, z$.

Rows and Columns: The horizontal and vertical lines in a matrix are, respectively, called the rows and columns of the matrix.

Order of a Matrix: The size (or dimension) of matrix is called as order of matrix. Order of matrix is based on the number of rows and number of columns. It can be written as $r \times c$; r means no. of row and c means no. of columns.

If a matrix has m rows and n columns then we say that the size or order of the matrix is $m \times n$. If A is a matrix having m rows and n columns then the matrix can be written as

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

The element, or entry, in the ith row and jth column of a $m \times n$ matrix A is written as a_{ij}

For example: The matrix $A = \begin{pmatrix} 2 & -1 & 3 \\ 0 & 4 & 6 \end{pmatrix}$ has two rows and three columns. So order of A will be 2×3

Square Matrix: A matrix with equal number of rows and columns is called square matrix.

For Example The matrix $A = \begin{pmatrix} 4 & 7 & -8 \\ 9 & 3 & 5 \\ 1 & -1 & 2 \end{pmatrix}$ has three rows and three columns. So it is a square matrix of order 3.

Equality of matrices: The two matrices will be equal if they must have

a) The same dimensions (i.e. same number of rows and columns)

b) Corresponding elements must be equal.

Example: The matrices $A = \begin{pmatrix} 4 & 7 & -8 \\ 9 & 3 & 5 \\ 1 & -1 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 4 & 7 & -8 \\ 9 & 3 & 5 \\ 1 & -1 & 2 \end{pmatrix}$ equal matrices

(i.e A = B) because they both have same orders and same corresponding elements.

Column Matrix: A column matrix X is any matrix having n rows and only one column. Thus the column matrix X can be written as

$$X = \begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \\ \vdots \\ b_{n1} \end{pmatrix} = [b_{i1}]_{n\times 1}$$

A column matrix is also called a column vector or simply a vector.

Multiple of matrix: A multiple of a matrix A by a nonzero constant k is defined to be

$$kA = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{bmatrix} = [ka_{ij}]_{m\times n}$$

Notice that the product kA is same as the product Ak. Therefore, we can write $kA = Ak$.

It implies that if we multiply a matrix by a constant *k*, then each element of the matrix is to be multiplied by *k*.

Example 1:

(a)
$$5\cdot\begin{bmatrix}2 & -3\\ 4 & -1\\ 1/5 & 6\end{bmatrix}=\begin{bmatrix}10 & -15\\ 20 & -5\\ 1 & 30\end{bmatrix}$$

(b)
$$e^{t}\cdot\begin{bmatrix}1\\ -2\\ 4\end{bmatrix}=\begin{bmatrix}e^{t}\\ -2e^{t}\\ 4e^{t}\end{bmatrix}$$

Since we know that $kA = Ak$. Therefore, we can write

$$e^{-3t}\cdot\begin{bmatrix}2\\ 5\end{bmatrix}=\begin{bmatrix}2e^{-3t}\\ 5e^{-3t}\end{bmatrix}=\begin{bmatrix}2\\ 5\end{bmatrix}e^{-3t}$$

Addition of Matrices: Only matrices of the same order may be added by adding corresponding elements.
If $A=[a_{ij}]$ and $B=[b_{ij}]$ are two $m\times n$ matrices then $A+B=[a_{ij}+b_{ij}]$
Obviously order of the matrix A + B is $m\times n$

Example 2: Consider the following two matrices of order 3×3

$$A=\begin{pmatrix}2 & -1 & 3\\ 0 & 4 & 6\\ -6 & 10 & -5\end{pmatrix},\quad B=\begin{pmatrix}4 & 7 & -8\\ 9 & 3 & 5\\ 1 & -1 & 2\end{pmatrix}$$

Since the given matrices have same orders, therefore, these matrices can be added and their sum is given by

$$A+B=\begin{pmatrix}2+4 & -1+7 & 3+(-8)\\ 0+9 & 4+3 & 6+5\\ -6+1 & 10+(-1) & -5+2\end{pmatrix}=\begin{pmatrix}6 & 6 & -5\\ 9 & 7 & 11\\ -5 & 9 & -3\end{pmatrix}$$

Example 3: Write the following single column matrix as the sum of three column vectors

$$\begin{pmatrix}3t^{2}-2e^{t}\\ t^{2}+7t\\ 5t\end{pmatrix}$$

Solution:

$$\begin{pmatrix} 3t^2-2e^t \\ t^2+7t \\ 5t \end{pmatrix} = \begin{pmatrix} 3t^2 \\ t^2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 7t \\ 5t \end{pmatrix} + \begin{pmatrix} -2e^t \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} t^2 + \begin{pmatrix} 0 \\ 7 \\ 5 \end{pmatrix} t + \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} e^t$$

Difference of Matrices: The difference of two matrices A and B of same order $m \times n$ is defined to be the matrix $A - B = A + (-B)$

The matrix $-B$ is obtained by multiplying the matrix B with -1. So that $-B = (-1)B$

Multiplication of Matrices: We can multiply two matrices if and only if, the number of columns in the first matrix equals the number of rows in the second matrix.
Otherwise, the product of two matrices is not possible.
OR
If the order of the matrix A is $m \times n$ then to make the product AB possible order of the matrix B must be $n \times p$. Then the order of the product matrix AB is $m \times p$. Thus

$$A_{m\times n} \cdot B_{n\times p} = C_{m\times p}$$

If the matrices A and B are given by

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \; B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

Then

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1p} + a_{12}b_{2p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & \cdots & a_{21}b_{1p} + a_{22}b_{2p} + \cdots + a_{2n}b_{np} \\ \vdots & \vdots & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1p} + a_{m2}b_{2p} + \cdots + a_{mn}b_{np} \end{bmatrix}$$

$$= \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)_{n \times p}$$

Example 4: If possible, find the products AB and BA, when

(a) $A = \begin{pmatrix} 4 & 7 \\ 3 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 9 & -2 \\ 6 & 8 \end{pmatrix}$

(b) $A = \begin{pmatrix} 5 & 8 \\ 1 & 0 \\ 2 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} -4 & -3 \\ 2 & 0 \end{pmatrix}$

Solution: (a) The matrices A and B are square matrices of order 2. Therefore, both of the products AB and BA are possible.

$$AB = \begin{pmatrix} 4 & 7 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} 9 & -2 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 4\cdot 9 + 7\cdot 6 & 4\cdot(-2) + 7\cdot 8 \\ 3\cdot 9 + 5\cdot 6 & 3\cdot(-2) + 5\cdot 8 \end{pmatrix} = \begin{pmatrix} 78 & 48 \\ 57 & 34 \end{pmatrix}$$

Similarly $BA = \begin{pmatrix} 9 & -2 \\ 6 & 8 \end{pmatrix}\begin{pmatrix} 4 & 7 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 9\cdot 4 + (-2)\cdot 3 & 9\cdot 7 + (-2)\cdot 5 \\ 6\cdot 4 + 8\cdot 3 & 6\cdot 7 + 8\cdot 5 \end{pmatrix} = \begin{pmatrix} 30 & 53 \\ 48 & 82 \end{pmatrix}$

Note: From above example it is clear that generally a matrix multiplication is not commutative i.e. $AB \neq BA$.

(b) The product AB is possible as the number of columns in the matrix A and the number of rows in B is 2. However, the product BA is not possible because the number of column in the matrix B and the number of rows in A is not same.

$$AB = \begin{pmatrix} 5 & 8 \\ 1 & 0 \\ 2 & 7 \end{pmatrix}\begin{pmatrix} -4 & -3 \\ 2 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 5\cdot(-4) + 8\cdot 2 & 5\cdot(-3) + 8\cdot 0 \\ 1\cdot(-4) + 0\cdot 2 & 1\cdot(-3) + 0\cdot 0 \\ 2\cdot(-4) + 7\cdot 2 & 2\cdot(-3) + 7\cdot 0 \end{pmatrix} = \begin{pmatrix} -4 & -15 \\ -4 & -3 \\ 6 & -6 \end{pmatrix}$$

$$AB = \begin{pmatrix} 78 & 48 \\ 57 & 34 \end{pmatrix}, \quad BA = \begin{pmatrix} 30 & 53 \\ 48 & 82 \end{pmatrix}$$

Clearly $AB \neq BA$.

$$AB = \begin{pmatrix} -4 & -15 \\ -4 & -3 \\ 6 & -6 \end{pmatrix}$$

However, the product BA is not possible.

Example 5:

(a) $$\begin{pmatrix} 2 & -1 & 3 \\ 0 & 4 & 5 \\ 1 & -7 & 9 \end{pmatrix}\begin{pmatrix} -3 \\ 6 \\ 4 \end{pmatrix} = \begin{pmatrix} 2\cdot(-3)+(-1)\cdot 6+3\cdot 4 \\ 0\cdot(-3)+4\cdot 6+5\cdot 6 \\ 1\cdot(-3)+(-7)\cdot 6+9\cdot 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 44 \\ -9 \end{pmatrix}$$

(b) $$\begin{pmatrix} -4 & 2 \\ 3 & 8 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -4x+2y \\ 3x+8y \end{pmatrix}$$

Multiplicative Identity: For a given any integer n, the $n \times n$ matrix

$$I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

is called the multiplicative identity matrix. If A is a matrix of order $n \times n$, then it can be verified that $I \cdot A = A \cdot I = A$

Example: $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ are identity matrices of orders 2 x 2 and 3 x 3 respectively and If $B = \begin{pmatrix} 9 & -2 \\ 6 & 8 \end{pmatrix}$ then we can easily prove that BI = IB = B

Zero Matrix or Null matrix: A matrix whose all entries are zero is called zero matrix or null matrix and it is denoted by O.

For example $O=\begin{pmatrix}0\\0\end{pmatrix}$; $O=\begin{pmatrix}0&0\\0&0\end{pmatrix}$; $O=\begin{pmatrix}0&0\\0&0\\0&0\end{pmatrix}$

and so on. If A and O are the matrices of same orders, then $A+O=O+A=A$

Associative Law: The matrix multiplication is associative. This means that if A, B and C are $m\times p$, $p\times r$ and $r\times n$ matrices, then $A(BC)=(AB)C$

The result is a $m\times n$ matrix. This result can be verified by taking any three matrices which are confirmable for multiplication.

Distributive Law: If B and C are matrices of order $r\times n$ and A is a matrix of order $m\times r$, then the distributive law states that

$$A(B+C)=AB+AC$$

Furthermore, if the product $(A+B)C$ is defined, then

$$(A+B)C=AC+BC$$

Determinant of a Matrix: Associated with every square matrix A of constants, there is a number called the determinant of the matrix, which is denoted by $\det(A)$ or $|A|$

Example 6: Find the determinant of the following matrix $A=\begin{pmatrix}3&6&2\\2&5&1\\-1&2&4\end{pmatrix}$

Solution: The determinant of the matrix A is given by

$$\det(A)=\begin{vmatrix}3&6&2\\2&5&1\\-1&2&4\end{vmatrix}$$

We expand the $\det(A)$ by first row, we obtain

$$\det(A)=\begin{vmatrix}3&6&2\\2&5&1\\-1&2&4\end{vmatrix}=3\begin{vmatrix}5&1\\2&4\end{vmatrix}-6\begin{vmatrix}2&1\\-1&4\end{vmatrix}+2\begin{vmatrix}2&5\\-1&2\end{vmatrix}$$

or $$\det(A)=3(20-2)-6(8+1)+2(4+5)=18$$

Transpose of a Matrix: The transpose of $m \times n$ matrix A is denoted by A^{tr} and it is obtained by interchanging rows of A into its columns. In other words, rows of A become the columns of A^{tr}. Clearly A^{tr} is $n \times m$ matrix.

If $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, then $A^{tr} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$

Since order of the matrix A is $m \times n$, the order of the transpose matrix A^{tr} is $n \times m$.

Properties of the Transpose:

The following properties are valid for the transpose;

- The transpose of the transpose of a matrix is the matrix itself: $(A^T)^T = A$
- The transpose of a matrix times a scalar (*k*) is equal to the constant times the transpose of the matrix: $(kA)^T = kA^T$
- The transpose of the sum of two matrices is equivalent to the sum of their transposes: $(A + B)^T = A^T + B^T$
- The transpose of the product of two matrices is equivalent to the product of their transposes in reversed order: $(AB)^T = B^T A^T$
- The same is true for the product of multiple matrices: $(ABC)^T = C^T B^T A^T$

Example 7: (a) The transpose of matrix $A = \begin{pmatrix} 3 & 6 & 2 \\ 2 & 5 & 1 \\ -1 & 2 & 4 \end{pmatrix}$ is $A^T = \begin{pmatrix} 3 & 2 & -1 \\ 6 & 5 & 2 \\ 2 & 1 & 4 \end{pmatrix}$

(b) If $X = \begin{pmatrix} 5 \\ 0 \\ 3 \end{pmatrix}$, then $X^T = [5 \quad 0 \quad 3]$

Multiplicative Inverse: Suppose that A is a square matrix of order $n \times n$. If there exists an $n \times n$ matrix B such that $AB = BA = I$, then B is said to be the multiplicative inverse of the matrix A and is denoted by $B = A^{-1}$.

For example: If $A=\begin{pmatrix}1 & 4\\ 2 & 10\end{pmatrix}$ then the matrix $B=\begin{pmatrix}5 & -2\\ -1 & 1/2\end{pmatrix}$ is multiplicative inverse of A because AB = $\begin{pmatrix}1 & 4\\ 2 & 10\end{pmatrix}\begin{pmatrix}5 & -2\\ -1 & 1/2\end{pmatrix}$ = $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ =I

Similarly we can check that BA = I

Singular and Non-Singular Matrices: A square matrix A is said to be a **non-singular** matrix if $\det(A)\neq 0$, otherwise the square matrix A is said to be **singular**. Thus for a singular matrix A we must have $\det(A)=0$

Example: $A = \begin{bmatrix}2 & 3 & -1\\ 1 & 1 & 0\\ 2 & -3 & 5\end{bmatrix}$

$$|A| = 2(5-0) - 3(5-0) - 1(-3-2)$$
$$= 10 - 15 + 5 = 0$$

which means that A is singular.

Minor of an element of a matrix:

Let A be a square matrix of order n x n. Then minor M_{ij} of the element $a_{ij}\in A$ is the determinant of $(n-1)\times(n-1)$ matrix obtained by deleting the *ith* row and *jth* column from A.

Example: If $A = \begin{bmatrix}2 & 3 & -1\\ 1 & 1 & 0\\ 2 & -3 & 5\end{bmatrix}$ is a square matrix. The Minor of $3\in A$ is denoted by M_{12} and is defined to be $M_{12} = \begin{vmatrix}1 & 0\\ 2 & 5\end{vmatrix}$ = 5-0 = 5

Cofactor of an element of a matrix:

Let A be a non singular matrix of order $n\times n$ and let C_{ij} denote the cofactor (signed minor) of the corresponding entry $a_{ij}\in A$, then it is defined to be $C_{ij} = (-1)^{i+j}M_{ij}$

Example: If $A = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 1 & 0 \\ 2 & -3 & 5 \end{bmatrix}$ is a square matrix. The cofactor of $3 \in A$ is denoted by C_{12} and is defined to be $C_{12} = = (-1)^{1+2}\begin{vmatrix} 1 & 0 \\ 2 & 5 \end{vmatrix} =$ - (5 - 0) = -5

Theorem: If A is a square matrix of order $n \times n$ then the matrix has a multiplicative inverse A^{-1} if and only if the matrix A is non-singular.

Theorem: Then inverse of the matrix A is given by $A^{-1} = \frac{1}{\det(A)}(C_{ij})^{tr}$

1. For further reference we take $n = 2$ so that A is a 2×2 non-singular matrix given by

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Therefore $C_{11} = a_{22}$, $C_{12} = -a_{21}$, $C_{21} = -a_{12}$ and $C_{22} = a_{11}$. So that

$$A^{-1} = \frac{1}{\det(A)}\begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}^{tr} = \frac{1}{\det(A)}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

2. For a 3×3 non-singular matrix A= $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$$C_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, C_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad C_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \text{ and so on.}$$

Therefore, inverse of the matrix A is given by $A^{-1} = \frac{1}{\det A}\begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix}$.

Example 8: Find, if possible, the multiplicative inverse for the matrix $A = \begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix}$.

Solution: The matrix A is non-singular because $\det(A) = \begin{vmatrix} 1 & 4 \\ 2 & 10 \end{vmatrix} = 10 - 8 = 2$

Therefore, A^{-1} exists and is given by $A^{-1} = \frac{1}{2}\begin{pmatrix} 10 & -4 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix}$

Check:

$$AA^{-1} = \begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix}\begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix} = \begin{pmatrix} 5-4 & -2+2 \\ 10-10 & -4+5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$$AA^{-1} = \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} 5-4 & 20-20 \\ -1+1 & -4+5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Example 9: Find, if possible, the multiplicative inverse of the following matrix

$$A = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$$

Solution: The matrix is singular because

$$\det(A) = \begin{vmatrix} 2 & 2 \\ 3 & 3 \end{vmatrix} = 2\cdot 3 - 2\cdot 3 = 0$$

Therefore, the multiplicative inverse A^{-1} of the matrix does not exist.

Example 10: Find the multiplicative inverse for the following matrix

$$A = \begin{pmatrix} 2 & 2 & 0 \\ -2 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix}.$$

Solution: Since $\det(A) = \begin{vmatrix} 2 & 2 & 0 \\ -2 & 1 & 1 \\ 3 & 0 & 1 \end{vmatrix} = 2(1-0) - 2(-2-3) + 0(0-3) = 12 \neq 0$

Therefore, the given matrix is non singular. So, the multiplicative inverse A^{-1} of the matrix A exists. The cofactors corresponding to the entries in each row are

$$C_{11} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1, \qquad C_{12} = -\begin{vmatrix} -2 & 1 \\ 3 & 1 \end{vmatrix} = 5, \quad C_{13} = \begin{vmatrix} -2 & 1 \\ 3 & 0 \end{vmatrix} = -3$$

$$C_{21} = -\begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} = -2, \quad C_{22} = \begin{vmatrix} 2 & 0 \\ 3 & 1 \end{vmatrix} = 2, \quad C_{23} = -\begin{vmatrix} 2 & 2 \\ 3 & 0 \end{vmatrix} = 6$$

$$C_{31} = \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} = 2, \quad C_{32} = -\begin{vmatrix} 2 & 0 \\ -2 & 1 \end{vmatrix} = -2, \quad C_{33} = \begin{vmatrix} 2 & 2 \\ -2 & 1 \end{vmatrix} = 6$$

Hence $$A^{-1} = \frac{1}{12}\begin{pmatrix} 1 & -2 & 2 \\ 5 & 2 & -2 \\ -3 & 6 & 6 \end{pmatrix} = \begin{pmatrix} 1/12 & -1/6 & 1/6 \\ 5/12 & 1/6 & -1/6 \\ -1/4 & 1/2 & 1/2 \end{pmatrix}$$

We can also verify that $A \cdot A^{-1} = A^{-1} \cdot A = I$

Derivative of a Matrix of functions:

Suppose that

$$A(t) = \left[a_{ij}(t) \right]_{m\times n}$$

is a matrix whose entries are functions those are differentiable on a common interval, then derivative of the matrix $A(t)$ is a matrix whose entries are derivatives of the corresponding entries of the matrix $A(t)$. Thus

$$\frac{dA}{dt} = \left[\frac{da_{ij}}{dt} \right]_{m\times n}$$

The derivative of a matrix is also denoted by $A'(t)$.

Integral of a Matrix of Functions:

Suppose that $A(t) = \left(a_{ij}(t)\right)_{m\times n}$ is a matrix whose entries are functions those are continuous on a common interval containing t, then integral of the matrix $A(t)$ is a matrix whose entries are integrals of the corresponding entries of the matrix $A(t)$. Thus

$$\int_{t_0}^{t} A(s)ds = \left(\int_{t_0}^{t} a_{ij}(s)ds \right)_{m\times n}$$

Example 11: Find the derivative and the integral of the following matrix $X(t) = \begin{pmatrix} \sin 2t \\ e^{3t} \\ 8t-1 \end{pmatrix}$

Solution: The derivative and integral of the given matrix are, respectively, given by

$$X'(t) = \begin{pmatrix} \frac{d}{dt}(\sin 2t) \\ \frac{d}{dt}(e^{3t}) \\ \frac{d}{dt}(8t-1) \end{pmatrix} = \begin{pmatrix} 2\cos 2t \\ 3e^{3t} \\ 8 \end{pmatrix} \quad \text{and} \int_0^t X(s)ds = \begin{pmatrix} \int_0^t \sin 2s ds \\ \int_0^t e^{3s} ds \\ \int_0^t 8s-1 ds \end{pmatrix} = \begin{pmatrix} -1/2\cos 2t + 1/2 \\ 1/3e^{3t} - 1/3 \\ 4t^2 - t \end{pmatrix}$$

Exercise:

Write the given sum as a single column matrix

1. $3t\begin{pmatrix} 2 \\ t \\ -1 \end{pmatrix} + (t-1)\begin{pmatrix} -1 \\ -t \\ 3 \end{pmatrix} - 2\begin{pmatrix} 3t \\ 4 \\ -5t \end{pmatrix}$

2. $\begin{pmatrix} 1 & -3 & 4 \\ 2 & 5 & -1 \\ 0 & -4 & -2 \end{pmatrix}\begin{pmatrix} t \\ 2t-1 \\ -t \end{pmatrix} + \begin{pmatrix} -t \\ 1 \\ 4 \end{pmatrix} - \begin{pmatrix} 2 \\ 8 \\ -6 \end{pmatrix}$

Determine whether the given matrix is singular or non-singular. If singular, find A^{-1}.

3. $A = \begin{pmatrix} 3 & 2 & 1 \\ 4 & 1 & 0 \\ -2 & 5 & -1 \end{pmatrix}$

4. $A = \begin{pmatrix} 4 & 1 & -1 \\ 6 & 2 & -3 \\ -2 & -1 & 2 \end{pmatrix}$

Find $\frac{dX}{dt}$

5. $X = \begin{pmatrix} \frac{1}{2}\sin 2t - 4\cos 2t \\ -3\sin 2t + 5\cos 2t \end{pmatrix}$

6. If $A(t) = \begin{pmatrix} e^{4t} & \cos \pi t \\ 2t & 3t^2 - 1 \end{pmatrix}$ then find (a) $\int_0^2 A(t)dt$, (b) $\int_0^t A(s)ds.$

7. Find the integral $\int_1^2 B(t)dt$ if $B(t) = \begin{pmatrix} 6t & 2 \\ 1/t & 4t \end{pmatrix}$

Lecture 3

Systems of Linear Equations

In this lecture we will discuss some ways in which systems of linear equations arise, how to solve them, and how their solutions can be interpreted geometrically.

Linear equations:

A line in $\boldsymbol{R}^2$ (2-dimensions) can be represented by an equation of the form $a_1x + a_2y = b$ (where a_1, a_2 not both zero). Similarly a plane in $\boldsymbol{R}^3$ (3-dimensional space) can be represented by an equation of the form $a_1x + a_2y + a_3z = b$ (where a_1, a_2, a_3 not all zero).

A linear equation in *n* variables $x_1, x_2, \cdots, x_n$ can be expressed in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \qquad (1)$$

where $a_1, a_2, \cdots, a_n$ and *b* are constants and the "*a's*" are not all zero.

Homogeneous linear equation:

In the special case if $b = 0$, Equation (1) has the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ (2)
This equation is called homogeneous linear equation.

Note: A linear equation does not involve any products or square roots of variables. All variables occur only to the first power and do not appear, as arguments of trigonometric, logarithmic, or exponential functions.

Examples of Linear Equations:

(1) The equations

$$2x_1 + 3x_2 + 2 = x_3 \quad \textit{and} \quad x_2 = 2\left(\sqrt{5} + x_1\right) + 2x_3$$ are both linear

(2) The following equations are also linear

$$x + 3y = 7 \qquad x_1 - 2x_2 - 3x_3 + x_4 = 0$$

$$\tfrac{1}{2}x - y + 3z = -1 \qquad x_1 + x_2 + \cdots + x_n = 1$$

(3) The equations $3x_1 - 2x_2 = x_1x_2$ *and* $x_2 = 4\sqrt{x_1} - 6$
are ***not linear*** because of the presence of x_1x_2 in the first equation and $\sqrt{x_1}$ in the second.

System of linear equations:

A finite set of linear equations is called a system of linear equations or ***linear system***. The variables in a linear system are called the ***unknowns***.

For example,

$$4x_1 - x_2 + 3x_3 = -1$$
$$3x_1 + x_2 + 9x_3 = -4$$

is a linear system of two equations in three unknowns x_1, x_2, and x_3.

General system of linear equations:

A general linear system of *m* equations in n-unknowns $x_1, x_2, \cdots, x_n$ can be written as

$$\begin{matrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{matrix} \qquad (3)$$

Solution of a system of linear equations:

A solution of a linear system in the unknowns $x_1, x_2, \cdots, x_n$ is a sequence of *n* numbers that when substituted for $x_1, x_2, \cdots, x_n$ respectively, makes every equation in the system a true statement. The set of all solutions of a linear system is called its ***solution set***.

Linear System with Two Unknowns:

When two lines intersect in $\boldsymbol{R}^2$, we get system of linear equations with two unknowns

For example, consider the linear system

$$a_1x + b_1y = c_1$$
$$a_2x + b_2y = c_2$$

The graphs of these equations are straight lines in the *xy-plane*, so a solution (*x*, *y*) of this system is infect a point of intersection of these lines.

Thus, there are three possibilities:

1. The lines may be parallel and distinct, in which case there is no intersection and consequently ***no solution***.
2. The lines may intersect at only one point, in which case the system has exactly ***one solution***.
3. The lines may coincide, in which case there are infinitely many points of intersection (the points on the common line) and consequently ***infinitely many solutions.***

Consistent and inconsistent system:

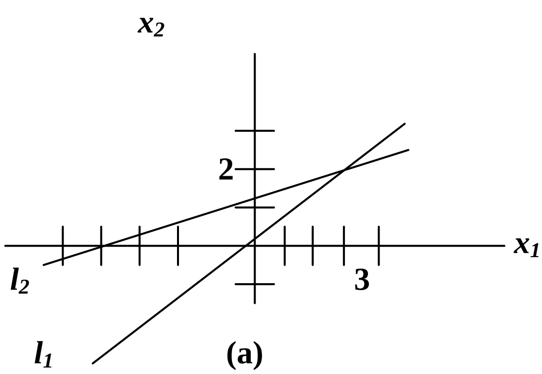
A linear system is said to be **consistent** if it has at least one solution and it is called **inconsistent** if it has no solutions.

Thus, a consistent linear system of two equations in two unknowns has either one solution or infinitely many solutions – there is no other possibility.

Example: consider the system of linear equations in two variables $x_1 - 2x_2 = -1,\ -x_1 + 3x_2 = 3$

Solve the equation simultaneously:

Adding both equations we get $x_2 = 2$, Put $x_2 = 2$ in any one of the above equation we get $x_1 = 3$. So the solution is the single point **(3, 2)**. See the graph of this linear system

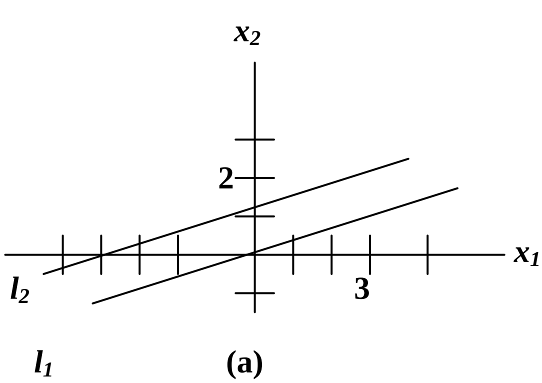


(a)

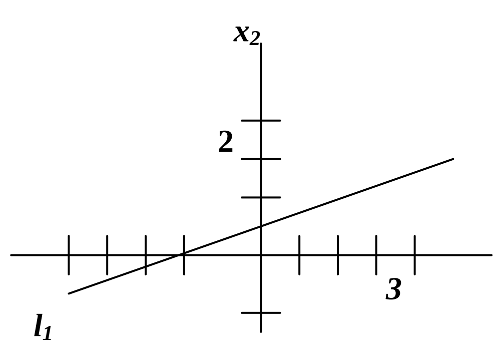
This system has exactly one solution

See the graphs to the following linear systems:

$$(a)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= \ \ 3 \end{aligned} \qquad\qquad (b)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= \ \ 1 \end{aligned}$$



(a)



(b)

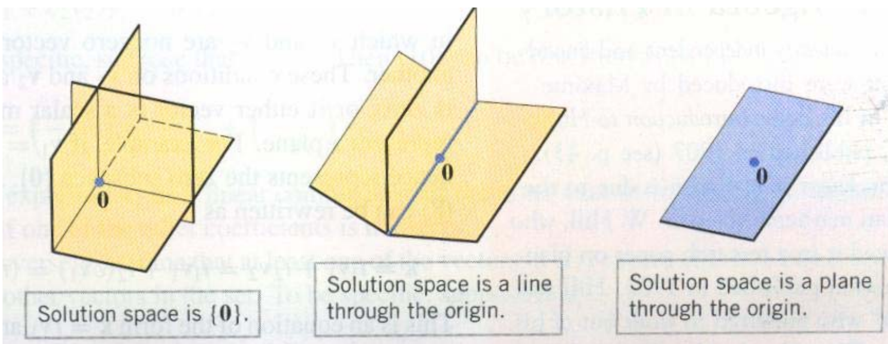
(a) No solution. (b) Infinitely many solutions.

Linear System with Three Unknowns:

Consider r a linear system of three equations in three unknowns:

$$a_1x + b_1y + c_1z = d_1$$
$$a_2x + b_2y + c_1z = d_2$$
$$a_3x + b_3y + c_3z = d_3$$

In this case, the graph of each equation is a plane, so the solutions of the system, If any correspond to points where all three planes intersect; and again we see that there are only three possibilities – no solutions, one solution, or infinitely many solutions as shown in figure.



Theorem 1: Every system of linear equations has zero, one or infinitely many solutions; there are no other possibilities.

Example 1: Solve the linear system

$$x - y = 1$$
$$2x + y = 6$$

Solution:

Adding both equations, we get $x = \frac{7}{3}$. Putting this value of x in 1st equation, we get $y = \frac{4}{3}$. Thus, the system has the ***unique solution*** $x = \frac{7}{3}, y = \frac{4}{3}$.

Geometrically, this means that the lines represented by the equations in the system intersect at a single point $\left(\frac{7}{3}, \frac{4}{3}\right)$ and thus has a unique solution.

Example 2: Solve the linear system

$$x + \quad y = 4$$
$$3x + 3y = 6$$

Solution:

Multiply first equation by 3 and then subtract the second equation from this. We obtain

$$0 = 6$$

This equation is contradictory.

Geometrically, this means that the lines corresponding to the equations in the original system are parallel and distinct. So the given system has ***no solution.***

Example 3: Solve the linear system

$$\begin{aligned} 4x - 2y &= 1 \\ 16x - 8y &= 4 \end{aligned}$$

Solution:

Multiply the first equation by -4 and then add in second equation.

$$\begin{array}{r} -16x + 8y = -4 \\ 16x - 8y = \;\;4 \\ \hline 0 \qquad = 0 \end{array}$$

Thus, the solutions of the system are those values of x and y that satisfy the single equation $4x - 2y = 1$

Geometrically, this means the lines corresponding to the two equations in the original system coincide and thus the system has infinitely many solutions.

Parametric Representation:

It is very convenient to describe the solution set in this case is to ***express it parametrically.*** We can do this by letting $y = t$ and solving for x in terms of t, or by letting $x = t$ and solving for y in terms of t.

The first approach yields the following parametric equations (by taking *y=t in the equation* $4x - 2y = 1$)

$$4x - 2t = 1, \quad y = t$$

$$x = \frac{1}{4} + \frac{1}{2}t, \quad y = t$$

We can now obtain some solutions of the above system by substituting some numerical values for the parameter.

Example: For $t = 0$ the solution is $(\frac{1}{4}, 0)$. For $t = 1$, the solution is $(\frac{3}{4}, 1)$ and for $t = -1$ the solution is $(-\frac{1}{4}, -1)$ *etc.*

.

Example 4: Solve the linear system

$$\begin{aligned} x - y + 2z &= 5 \\ 2x - 2y + 4z &= 10 \\ 3x - 3y + 6z &= 15 \end{aligned}$$

Solution:
Since the second and third equations are multiples of the first.

Geometrically, this means that the three planes coincide and those values of *x*, *y* and *z* that satisfy the equation $x - y + 2z = 5$ automatically satisfy all three equations.

We can express the solution set parametrically as

$$x = 5 + t_1 - 2t_2, y = t_1, \quad z = t_2$$

Some solutions can be obtained by choosing some numerical values for the parameters.

For example if we take $y = t_1 = 2$ and $z = t_2 = 3$ then

$$\begin{aligned} x &= 5 + t_1 - 2t_2 \\ &= 5 + 2 - 2(3) \\ &= 1 \end{aligned}$$

Put these values of x, y, and z in any equation of linear system to verify

$$\begin{aligned} x - y + 2z &= 5 \\ 1 - 2 + 2(3) &= 5 \\ 1 - 2 + 6 &= 5 \\ 5 &= 5 \end{aligned}$$

Hence x = 1, y = 2, z = 3 is the solution of the system. Verified.

Matrix Notation:

The essential information of a linear system can be recorded compactly in a rectangular array called a **matrix**.

Given the system

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

With the coefficients of each variable aligned in columns, the matrix $\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$

is called the coefficient matrix (or matrix of coefficients) of the system.

An augmented matrix of a system consists of the coefficient matrix with an added column containing the constants from the right sides of the equations. It is always denoted by A_b

$$A_b = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

Solving a Linear System:

In order to solve a linear system, we use a number of methods. 1st of them is given below.

Successive elimination method: In this method the x_1 term in the first equation of a system is used to eliminate the x_1 terms in the other equations. Then we use the x_2 term in the second equation to eliminate the x_2 terms in the other equations, and so on, until we finally obtain a very simple equivalent system of equations.

Example 5: Solve

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

Solution: We perform the elimination procedure with and without matrix notation, and place the results side by side for comparison:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

To eliminate the x_1 term from third equation add 4 times equation 1 to equation 3,

$$\begin{aligned} 4x_1 - 8x_2 + 4x_3 &= 0 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \\ \hline -3x_2 + 13x_3 &= -9 \end{aligned}$$

The result of the calculation is written in place of the original third equation:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

Next, multiply equation 2 by ½ in order to obtain 1 as the coefficient for x_2

$$\begin{aligned} x_1 - 2x_2 + \quad x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

To eliminate the x_2 term from third equation add 3 times equation 2 to equation 3,

The new system has a triangular form

$$\begin{aligned} & x_1 - 2x_2 + x_3 = 0 \\ & x_2 - 4x_3 = 4 \\ & x_3 = 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Now using 3rd equation eliminate the x_3 term from first and second equation i.e. multiply 3rd equation with 4 and add in second equation. Then subtract the third equation from first equation we get

$$\begin{aligned} x_1 - 2x_2 &= -3 \\ x_2 &= 16 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Adding 2 times equation 2 to equation 1, we obtain the result

$$\begin{cases} x_1 = 29 \\ x_2 = 16 \\ x_3 = 3 \end{cases} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

This completes the solution.
Our work indicates that the only solution of the original system is (29, 16, 3).

To verify that (29, 16, 3) is a solution, substitute these values into the left side of the original system for x_1, x_2 and x_3 and after computing, we get

$$(29) – 2(16) + (3) = 29 – 32 + 3 = 0$$
$$2(16) – 8(3) = 32 – 24 = 8$$
$$–4(29) + 5(16) + 9(3) = –116 + 80 + 27 = –9$$

The results agree with the right side of the original system, so (29, 16, 3) is a solution of the system.

This example illustrates how operations on equations in a linear system correspond to operations on the appropriate rows of the augmented matrix. The three basic operations listed earlier correspond to the following operations on the augmented matrix.

Elementary Row Operations:

1. (Replacement) Replace one row by the sum of itself and a nonzero multiple of another row.
2. (Interchange) Interchange two rows.
3. (Scaling) Multiply all entries in a row by a nonzero constant.

Row equivalent matrices:

A matrix B is said to be row equivalent to a matrix A of the same order if B can be obtained from A by performing a finite sequence of elementary row operations of A.
If A and B are row equivalent matrices, then we write this expression mathematically as $A \sim B$.

For example $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$ are row equivalent matrices because we add 4 times of 1^{st} row in 3^{rd} row in 1^{st} matrix.

Note: If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

Row operations are extremely easy to perform, but they have to be learnt and practice.

Two Fundamental Questions:

1. Is the system consistent; that is, does at least one solution exist?
2. If a solution exists is it the only one; that is, is the solution unique?

We try to answer these questions via row operations on the augmented matrix.

Example 6: Determine if the following system of linear equations is consistent

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

Solution:

First obtain the triangular matrix by removing x_1 and x_2 term from third equation and removing x_2 from second equation.

First divide the second equation by 2 we get

$$\begin{aligned} x_1 - 2x_2 + \quad x_3 &= 0 \\ x_2 - 4x_3 &= \ 4 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

Now multiply equation 1 with 4 and add in equation 3 to eliminate x_1 from third equation.

$$\begin{aligned} x_1 - 2x_2 + \quad x_3 &= 0 \\ x_2 - 4x_3 &= \ 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

Now multiply equation 2 with 3 and add in equation 3 to eliminate x_2 from third equation.

$$\begin{aligned} &x_1 - 2x_2 + x_3 = 0 \\ &x_2 - 4x_3 = 4 \\ &x_3 = 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Put value of x_3 in second equation we get

$x_2 - 4(3) = 4$

$x_2 = 16$

Now put these values of x_2 and x_3 in first equation we get

$x_1 - 2(16) + 3 = 0$

$x_1 = 29$

So a solution exists and the system is consistent and has a unique solution.

Example 7: Solve if the following system of linear equations is consistent.

$$\begin{aligned} x_2 - 4x_3 &= 8 \\ 2x_1 - 3x_2 + 2x_3 &= 1 \\ 5x_1 - 8x_2 + 7x_3 &= 1 \end{aligned}$$

Solution: The augmented matrix is

$$\begin{bmatrix} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

To obtain x_1 in the first equation, interchange rows 1 and 2:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

To eliminate the $5x_1$ term in the third equation, add –5/2 times row 1 to row 3:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & -1/2 & 2 & -3/2 \end{bmatrix}$$

Next, use the x_2 term in the second equation to eliminate the –(1/2) x_2 term from the third equation. Add ½ times row 2 to row 3:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix}$$

The augmented matrix is in triangular form.
To interpret it correctly, go back to equation notation:

$$\begin{aligned} 2x_1 - 3x_2 + 2x_3 &= 1 \\ x_2 - 4x_3 &= 8 \\ 0 &= 2.5 \end{aligned}$$

There are no values of x_1, x_2, x_3 that will satisfy because the equation 0 = 2.5 is never true.
Hence original ***system is inconsistent (i.e., has no solution)***.

Exercises:

1. State in words the next elementary "row" operation that should be performed on the system in order to solve it. (More than one answer is possible in (a).)

$$a.\quad \begin{aligned} x_1 + 4x_2 - 2x_3 + 8x_4 &= 12 \\ x_2 - 7x_3 + 2x_4 &= -4 \\ 5x_3 - x_4 &= 7 \\ x_3 + 3x_4 &= -5 \end{aligned} \qquad b.\quad \begin{aligned} x_1 - 3x_2 + 5x_3 - 2x_4 &= 0 \\ x_2 + 8x_3 &= -4 \\ 2x_3 &= 7 \\ x_4 &= 1 \end{aligned}$$

2. The augmented matrix of a linear system has been transformed by row operations into the form below. Determine if the system is consistent.

$$\begin{bmatrix} 1 & 5 & 2 & -6 \\ 0 & 4 & -7 & 2 \\ 0 & 0 & 5 & 0 \end{bmatrix}$$

3. Is (3, 4, –2) a solution of the following system?

$$\begin{aligned} 5x_1 - x_2 + 2x_3 &= 7 \\ -2x_1 + 6x_2 + 9x_3 &= 0 \\ -7x_1 + 5x_2 - 3x_3 &= -7 \end{aligned}$$

4. For what values of *h* and *k* is the following system consistent?

$$\begin{aligned} 2x_1 - x_2 &= h \\ -6x_1 + 3x_2 &= k \end{aligned}$$

Solve the systems in the exercises given below;

5. $$\begin{aligned} x_2 + 5x_3 &= -4 \\ x_1 + 4x_2 + 3x_3 &= -2 \\ 2x_1 + 7x_2 + x_3 &= -1 \end{aligned}$$

6. $$\begin{aligned} x_1 - 5x_2 + 4x_3 &= -3 \\ 2x_1 - 7x_2 + 3x_3 &= -2 \\ 2x_1 - x_2 - 7x_3 &= 1 \end{aligned}$$

7. $$\begin{aligned} x_1 + 2x_2 &= 4 \\ x_1 - 3x_2 - 3x_3 &= 2 \\ x_2 + x_3 &= 0 \end{aligned}$$

8. $$\begin{aligned} 2x_1 - 4x_3 &= -10 \\ x_2 + 3x_3 &= 2 \\ 3x_1 + 5x_2 + 8x_3 &= -6 \end{aligned}$$

Determine the value(s) of h such that the matrix is augmented matrix of a consistent linear system.

9. $\begin{bmatrix} 1 & -3 & h \\ -2 & 6 & -5 \end{bmatrix}$

10. $\begin{bmatrix} 1 & h & -2 \\ -4 & 2 & 10 \end{bmatrix}$

Find an equation involving g, h, and that makes the augmented matrix correspond to a consistent system.

11. $\begin{bmatrix} 1 & -4 & 7 & g \\ 0 & 3 & -5 & h \\ -2 & 5 & -9 & k \end{bmatrix}$

12. $\begin{bmatrix} 2 & 5 & -3 & g \\ 4 & 7 & -4 & h \\ -6 & -3 & 1 & k \end{bmatrix}$

Find the elementary row operations that transform the first matrix into the second, and then find the reverse row operation that transforms the second matrix into first.

13. $\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -4 \\ 0 & -3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & -2 \\ 0 & -3 & 4 \end{bmatrix}$

14. $\begin{bmatrix} 0 & 5 & -3 \\ 1 & 5 & -2 \\ 2 & 1 & 8 \end{bmatrix}, \begin{bmatrix} 1 & 5 & -2 \\ 0 & 5 & -3 \\ 2 & 1 & 8 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 1 & -4 & 2 \\ 0 & 2 & -5 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 1 & -4 & 2 \\ 0 & 0 & 3 & -5 \end{bmatrix}$

Lecture 4

Row Reduction and Echelon Forms

To analyze system of linear equations we shall discuss how to refine the row reduction algorithm. The algorithm applies to any matrix, we begin by introducing a non zero row or column (i.e. contains at least one nonzero entry) in a matrix,

Echelon form of a matrix:

A rectangular matrix is in *echelon form* (or row echelon form) if it has the following three properties:

1. All nonzero rows are above any rows of all zeros
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zero.

Reduced Echelon Form of a matrix:

If a matrix in echelon form satisfies the following additional conditions, then it is in *reduced echelon form* (or reduced row echelon form):

4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

Examples of Echelon Matrix form:

The following matrices are in echelon form. The leading entries ($\circ$) may have any nonzero value; the started entries (*) may have any values (including zero).

1. $\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix}$

2. $\begin{bmatrix} \circ & * & * & * \\ 0 & \circ & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 0 & \circ & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \circ & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \circ & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \circ & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \circ & * \end{bmatrix}$

4. $\begin{bmatrix} 1 & 4 & -3 & 7 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

6. $\begin{bmatrix} 0 & 1 & 2 & 6 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Examples of Reduced Echelon Form:

The following matrices are in reduced echelon form because the leading entries are 1's, and there are 0's below and above each leading 1.

1. $\begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$

4. $\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -1 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

6. $\begin{bmatrix} 0 & 1 & -2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Note: A matrix may be row reduced into more than one matrix in echelon form, using different sequences of row operations. However, the reduced echelon form one obtains from a matrix is unique.

Theorem 1 (Uniqueness of the Reduced Echelon Form): Each matrix is row equivalent to one and only one reduced echelon matrix.

Pivot Positions:

A pivot position in a matrix ***A*** is a location in ***A*** that corresponds to a leading entry in an echelon form of ***A***.

Note: When row operations on a matrix produce an echelon form, further row operations to obtain the reduced echelon form do not change the positions of the leading entries.

Pivot column:

A **pivot column** is a column of ***A*** that contains a pivot position.

Example 2: Reduce the matrix ***A*** below to echelon form, and locate the pivot columns

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

Solution: Leading entry in first column of above matrix is zero which is the pivot position. A nonzero entry, or pivot, must be placed in this position. So interchange first and last row.

$$\begin{bmatrix} 1 \text{ (Pivot)} & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

Pivot Column (first column)

Since all entries in a column below a leading entry should be zero. For this add row 1 in row 2, and multiply row 1 by 2 and add in row 3.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \quad \begin{array}{|l|} R_1 + R_2 \\ 2R_1 + R_3 \end{array}$$

Pivot (entry 2 in row 2); Next pivot column (second column)

Add –5/2 times row 2 to row 3, and add 3/2 times row 2 to row 4.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \qquad \begin{matrix} -\frac{5}{2}R_2 + R_3 \\ \frac{3}{2}R_2 + R_4 \end{matrix}$$

Interchange rows 3 and 4, we can produce a leading entry in column 4.

Pivot

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \textit{General} \quad \textit{form} \quad \begin{bmatrix} \circ & * & * & * & * \\ 0 & \circ & * & * & * \\ 0 & 0 & 0 & \circ & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivot column

This is in echelon form and thus columns 1, 2, and 4 of ***A*** are pivot columns.

Pivot positions

$$\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

Pivot columns

Pivot element:

A pivot is a nonzero number in a pivot position that is used as needed to create zeros via row operations

The Row Reduction Algorithm consists of four steps, and it produces a matrix in echelon form. A fifth step produces a matrix in reduced echelon form.

The algorithm is explained by an example.

Example 3: Apply elementary row operations to transform the following matrix first into echelon form and then into reduced echelon form.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Solution:

STEP 1: Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Pivot column (first column)

STEP 2: Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position

Interchange rows 1 and 3. (We could have interchanged rows 1 and 2 instead.)

Pivot (entry 3 in row 1, column 1)

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

STEP 3*:* Use row replacement operations to create zeros in all positions below the pivot

Subtract Row 1 from Row 2. i.e. $R_2 - R_1$

Pivot (entry 3 in row 1, column 1)

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

STEP 4: Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1 –3 to the sub-matrix, which remains. Repeat the process until there are no more nonzero rows to modify.

With row 1 covered, step 1 shows that column 2 is the next pivot column; for step 2, we'll select as a pivot the "top" entry in that column.

Pivot

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

Next pivot column

According to step 3 "*All entries in a column below a leading entry are zero*". For this subtract 3/2 time R_2 from R_3

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad R_3 - \frac{3}{2}R_2$$

When we cover the row containing the second pivot position for step 4, we are left with a new sub matrix having only one row:

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Pivot

This is the Echelon form of the matrix.
To make it in reduced echelon form we need to do one more step:

STEP 5: Make the leading entry in each nonzero row 1. Make all other entries of that column to 0.

Divide first Row by 3 and 2^{nd} Row by 2

$$\begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad \frac{1}{2}R_2 \ , \quad \frac{1}{3}R_1$$

Multiply second row by 3 and then add in first row.

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad 3R_2 + R_1$$

Subtract row 3 from row 2, and multiply row 3 by 5 and then subtract it from first row

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad \begin{matrix} R_2 - R_3 \\ R_1 - 5R_3 \end{matrix}$$

This is the matrix is in reduced echelon form.

Solutions of Linear Systems:

When this algorithm applied to the augmented matrix of the system it gives solution set of linear system.
Suppose, for example, that the augmented matrix of a linear system has been changed into the equivalent reduced echelon form

$$\begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

There are three variables because the augmented matrix has four columns. The associated system of equations is

$$\begin{aligned} x_1 \quad -5x_3 &= 1 \\ x_2 + x_3 &= 4 \qquad\qquad (1) \\ 0 &= 0 \quad \text{which means } x_3 \text{ is free} \end{aligned}$$

The variables x_1 and x_2 corresponding to pivot columns in the above matrix are called basic variables. The other variable, x_3 is called a free variable.

Whenever a system is consistent, the solution set can be described explicitly by solving the reduced system of equations for the basic variables in terms of the free variables. This operation is possible because the reduced echelon form places each basic variable in one and only one equation.

In (4), we can solve the first equation for x_1 and the second for x_2. (The third equation is ignored; it offers no restriction on the variables.)

$$\begin{aligned} &x_1 = 1 + 5x_3 \\ &x_2 = 4 - x_3 \qquad\qquad (2) \\ &x_3 \quad is \quad free \end{aligned}$$

By saying that x_3 is "free", we mean that we are free to choose any value for x_3. When $x_3 = 0$, the solution is (1, 4, 0); when $x_3 = 1$, the solution is (6, 3, 1 etc).

Note: The solution in (2) is called a **general solution** of the system because it gives an explicit description of all solutions.

Example 4: Find the general solution of the linear system whose augmented matrix has been reduced to $\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$

Solution: The matrix is in echelon form, but we want the reduced echelon form before solving for the basic variables. The symbol "~" before a matrix indicates that the matrix is row equivalent to the preceding matrix.

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

By $R_1 + 2R_3$ and $R_2 + R_3$ We get

$$\sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 2 & -8 & 0 & 10 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

By $\frac{1}{2}R_2$ we get

$$\sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

By $R_1 - 2R_2$ we get

$$\sim \begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

The matrix is now in reduced echelon form.
The associated system of linear equations now is

$$\begin{aligned} x_1 + 6x_2 \quad + 3x_4 &= 0 \\ x_3 - 4x_4 &= 5 \\ x_5 &= 7 \end{aligned} \qquad (6)$$

The pivot columns of the matrix are 1, 3 and 5, so the basic variables are x_1, x_3, and x_5. The remaining variables, x_2 and x_4, must be free.

Solving for the basic variables, we obtain the general solution:

$$\begin{cases} x_1 = -6x_2 - 3x_4 \\ x_2 \text{ is free} \\ x_3 = 5 + 4x_4 \\ x_4 \text{ is free} \\ x_5 = 7 \end{cases} \qquad (7)$$

Note that the value of x_5 is already fixed by the third equation in system (6).

Exercise:

1. Find the general solution of the linear system whose augmented matrix is

$$\begin{bmatrix} 1 & -3 & -5 & 0 \\ 0 & 1 & 1 & 3 \end{bmatrix}$$

2. Find the general solution of the system

$$\begin{aligned} x_1 - 2x_2 - \ \ x_3 + 3x_4 &= 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 &= 3 \\ 3x_1 - 6x_2 - 6x_3 + 8x_4 &= 2 \end{aligned}$$

Find the general solutions of the systems whose augmented matrices are given in Exercises 3-12

3. $\begin{bmatrix} 1 & 0 & 2 & 5 \\ 2 & 0 & 3 & 6 \end{bmatrix}$

4. $\begin{bmatrix} 1 & -3 & 0 & -5 \\ -3 & 7 & 0 & 9 \end{bmatrix}$

5. $\begin{bmatrix} 0 & 3 & 6 & 9 \\ -1 & 1 & -2 & -1 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 3 & -3 & 7 \\ 3 & 9 & -4 & 1 \end{bmatrix}$

7. $\begin{pmatrix} 1 & 2 & -7 \\ -1 & -1 & 1 \\ 2 & 1 & 5 \end{pmatrix}$

8. $\begin{pmatrix} 1 & 2 & 4 \\ -2 & -3 & -5 \\ 2 & 1 & -1 \end{pmatrix}$

9. $\begin{pmatrix} 2 & -4 & 3 \\ -6 & 12 & -9 \\ 4 & -8 & 6 \end{pmatrix}$

10. $\begin{pmatrix} 1 & 0 & -9 & 0 & 4 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

11. $\begin{pmatrix} 1 & -2 & 0 & 0 & 7 & -3 \\ 0 & 1 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 1 & 5 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

12. $\begin{pmatrix} 1 & 0 & -5 & 0 & -8 & 3 \\ 0 & 1 & 4 & -1 & 0 & 6 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Determine the value(s) of h such that the matrix is the augmented matrix of a consistent linear system.

13. $\begin{bmatrix} 1 & 4 & 2 \\ -3 & h & -1 \end{bmatrix}$

14. $\begin{bmatrix} 1 & h & 3 \\ 2 & 8 & 1 \end{bmatrix}$

Choose h and k such that the system has (a) no solution, (b) a unique solution, and (c) many solutions. Give separate answer for each part.

15. $x_1 + hx_2 = 1$
$2x_1 + 3x_2 = k$

16. $x_1 - 3x_2 = 1$
$2x_1 + hx_2 = k$

Lecture 5

Vector Equations

This lecture is devoted to connect equations involving vectors to ordinary systems of equations. The term vector appears in a variety of mathematical and physical contexts, which we will study later, while studying "Vector Spaces". Until then, we will use vector to mean a list of numbers. This simple idea enables us to get to interesting and important applications as quickly as possible.

Column Vector:

"*A matrix with only one column is called column vector or simply a vector".*

e.g. $u = \begin{bmatrix} 3 & -1 \end{bmatrix}^T = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, $v = \begin{bmatrix} 2 & 3 & 5 \end{bmatrix}^T = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$, $w = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 \end{bmatrix}^T$ are all column vectors or simply vectors.

Vectors in $\mathbf{R^2}$:

If $\mathbb{R}$ is the set of all real numbers then the set of all vectors with two entries is denoted by $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.

For example: the vector $u = \begin{bmatrix} 3 & -1 \end{bmatrix}^T = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \in \mathbb{R}^2$

Here real numbers are appeared as entries in the vectors, and the exponent **2** indicates that the vectors contain only two entries.

Similarly $\boldsymbol{R^3}$ & $\boldsymbol{R^4}$ contains all vectors with three and four entries respectively. The entries of the vectors are always taken from the set of real numbers ***R.*** The entries in vectors are assumed to be the elements of a set, called as ***Field***. It is denoted by F .

Algebra of Vectors:

Equality of vectors in $\mathbb{R}^2$:

Two vectors in $\boldsymbol{R^2}$ are equal if and only if their corresponding entries are equal.

If $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$ *then* $u = v$ *iff* $\boxed{u_1 = v_1} \wedge \boxed{u_2 = v_2}$

So $\begin{bmatrix} 4 \\ 6 \end{bmatrix} \neq \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ as $4 = 4$ but $6 \neq 3$

Note: In fact, vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\boldsymbol{R^2}$ are nothing but ordered pairs (x, y) of real numbers both representing the position of a point with respect to origin.

Addition of Vectors:

Given two vectors ***u*** and ***v*** in $\boldsymbol{R^2}$, their sum is the *vector* ***u + v*** *obtained by adding corresponding entries of the vectors* ***u*** *and* ***v***.

For $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$ Then $u + v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix} \in \mathbb{R}^2$

For example, $\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1+2 \\ -2+5 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

Scalar Multiplication of a vector:

Given a vector ***u*** and a real number *c*, the scalar multiple of ***u*** by *c* is the *vector c**u** obtained by multiplying each entry in **u** by c.*

For example, if $u = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ *and* $c = 5$, *then* $cu = 5\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ -5 \end{bmatrix}$

Notations: The number *c* in *c**u*** is a **scalar**; it is written in lightface type to distinguish it from the boldface vector ***u***.

Example 1: Given $u = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ *and* $v = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, find 4***u***, (-3) ***v***, and 4u + (-3) ***v***

Solution: $4u == 4\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 4\times 1 \\ 4\times(-2) \end{bmatrix} = \begin{bmatrix} 4 \\ -8 \end{bmatrix}$, $(-3)v = (-3)\begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} -6 \\ 15 \end{bmatrix}$

And $4u + (-3)v = \begin{bmatrix} 4 \\ -8 \end{bmatrix} + \begin{bmatrix} -6 \\ 15 \end{bmatrix} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$

Note: Sometimes for our convenience, we write a column vector $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ in the form (3, –1). In this case, we use *parentheses and a comma to distinguish the vector* (3, –1) *from the* 1×2 *row matrix* [3 –1], written with brackets and no comma.

Thus $\begin{bmatrix} 3 \\ -1 \end{bmatrix} \neq [3 \quad -1]$ but $\begin{bmatrix} 3 \\ -1 \end{bmatrix} = (3, -1)$

Geometric Descriptions of R^2:

Consider a rectangular coordinate system in the plane. Because each point in the plane is determined by an ordered pair of numbers, *we can identify a geometric point (a, b) with the column vector* $\begin{bmatrix} a \\ b \end{bmatrix}$. So we may regard R^2 as the set of all points in the plane.

See Figure 1.

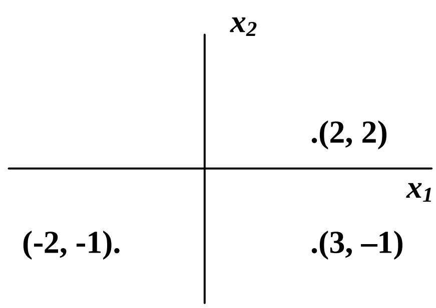


Figure 1 Vectors as points.

Vectors in R^3:

Vectors in R^3 are 3×1 column matrices with three entries. They are represented geometrically by points in a three-dimensional coordinate space, with arrows from the origin sometimes included for visual clarity.

Vectors in R^n:

If **n** is a positive integer, R^n (read "r-n") denotes the collection of all lists (or ordered n- tuples) of **n** real numbers, usually written as $n\times 1$ column matrices, such as

$$u = \begin{bmatrix} u_1 & u_2 & \cdots u_n \end{bmatrix}^T$$

The vector whose entries are all zero is called the **zero vector** and is denoted by **O.** (The number of entries in **O** will be clear from the context.)

Algebraic Properties of R^n:

For all ***u, v, w*** in $\boldsymbol{R^n}$ and all scalars *c* and *d*:

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (Commutative)

(ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (Associative)

(iii) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$ (Additive Identity)

(iv) $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$ (Additive Inverse)

where $-\mathbf{u}$ denotes $(-1)\mathbf{u}$

(v) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ (Scalar Distribution over Vector Addition)

(vi) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ (Vector Distribution over Scalar Addition)

(vii) $c(d\mathbf{u}) = (cd)\mathbf{u}$

(viii) $1\mathbf{u}=\mathbf{u}$

Linear Combinations: Given vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_p}$ in $\mathbf{R^n}$ and given scalars $c_1, c_2, \ldots, c_p$ the vector defined by

$$y = c_1v_1 + c_2v_2 + \cdots + c_pv_p$$

is called a **linear combination** of $\mathbf{v_1}, \ldots, \mathbf{v_p}$ using weights $\mathbf{c_1}, \ldots, \mathbf{c_p}$.

Property (ii) above permits us to omit parentheses when forming such a linear combination. The weights in a linear combination can be any real numbers, including zero.

Example:

For $v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, if $w = \frac{5}{2}v_1 - \frac{1}{2}v_2$ the we say that w is a linear combination of v_1 and v_2.

Example: As (3, 5 , 2) = 3(1, 0 , 0) + 5(0, 1 , 0) + 2(0, 0 , 1)

(3, 5 , 2) = $3v_1 + 5v_2 + 2v_3$ where v_1 = (1, 0 , 0) , v_2 = (0, 1 , 0) v_3 = (0, 0 , 1)

So (3, 5 , 2) is a vector which is linear combination of v_1, v_2, v_3

Example 5: Let $a_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}, a_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}, and \quad b = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$.

Determine whether ***b*** can be generated (or written) as a linear combination of $\mathbf{a_1}$ and $\mathbf{a_2}$. That is, determine whether weights x_1 and x_2 exist such that

$$x_1 \mathbf{a_1} + x_2 \mathbf{a_2} = \mathbf{b} \qquad (1)$$

If the vector equation (1) has a solution, find it.

Solution: Use the definitions of scalar multiplication and vector addition to rewrite the vector equation

$$x_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

$$\mathbf{a_1} \qquad \mathbf{a_2} \qquad \mathbf{b}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ -2x_1 \\ -5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \qquad (2)$$

$$\Rightarrow \begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= -3 \end{aligned} \qquad (3)$$

We solve this system by row reducing the augmented matrix of the system as follows:

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix}$$

$By\ R_2 + 2R_1\ ; R_3 + 5R_1$

$$\sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix}$$

$By \left(\frac{1}{9}\right) R_2; \left(\frac{1}{16}\right) R_3^{/}$

$$\sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

$By\, R_3 - R_2; R_1 - 2R_2$

$$\sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

The solution of (3) is $x_1 = 3$ and $x_2 = 2$. Hence ***b*** is a linear combination of $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$, with weights $x_1 = 3$ and $x_2 = 2$.

Spanning Set:

If $\boldsymbol{v_1, \ldots, v_p}$ *are in* $\boldsymbol{R^n}$, *then the set of all linear combinations of* $\boldsymbol{v_1, \ldots, v_p}$ *is denoted by Span* $\{\boldsymbol{v_1, \ldots, v_p}\}$ *and is called the* ***subset of*** $\boldsymbol{R^n}$ ***spanned*** *(or* ***generated****) by* $\boldsymbol{v_1, \ldots, v_p}$. *That is, Span* $\{\boldsymbol{v_1, \ldots, v_p}\}$ *is the collection of all vectors that can be written in the form* $c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_p\boldsymbol{v_p}$, *with* $c_1, \ldots, c_p$ *scalars.*

If we want to check whether a vector ***b*** is in Span $\{\boldsymbol{v_1, \ldots, v_p}\}$ then we will see whether the vector equation

$x_1\boldsymbol{v_1} + x_2\boldsymbol{v_2} + \ldots + x_p\boldsymbol{v_p} = \boldsymbol{b}$ has a solution, or

Equivalently, whether the linear system with augmented matrix $[\,\boldsymbol{v_1, \ldots, v_p} \quad \boldsymbol{b}]$ has a solution.

Note:

(1) The set Span $\{\boldsymbol{v_1, \ldots, v_p}\}$ contains every scalar multiple of $\boldsymbol{v_1}$

because $c\boldsymbol{v_1} = c\boldsymbol{v_1} + 0\boldsymbol{v_2} + \ldots + 0\boldsymbol{v_p}$ i.e every $c\boldsymbol{v_i}$ can be written as a linear combination of $\boldsymbol{v_1, \ldots, v_p}$

(2) Zero vector $= 0 \in Span\{v_1, v_2, \cdots v_n\}$ as 0 can be written as the linear combination of $v_1, v_2, \cdots v_n$ that is $0_v = 0_F v_1 + 0_F v_2 + \cdots + 0_F v_n$ here for the convenience it is mentioned that 0_v is the vector(zero vector) while 0_F is zero scalar (weight of all $v_1, v_2, \cdots v_n$) and in particular not to make confusion that 0_v and 0_F are same!

A Geometric Description of Span {*v*} and Span {*u, v*}:

Let $\boldsymbol{v}$ be a nonzero vector in $\boldsymbol{R^3}$. Then Span $\{\boldsymbol{v}\}$ is the set of all linear combinations of $\boldsymbol{v}$ or in particular set of scalar multiples of $\boldsymbol{v}$, and we visualize it as the set of points on the line in $\boldsymbol{R^3}$ through $\boldsymbol{v}$ and $\boldsymbol{0}$.

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are nonzero vectors in $\boldsymbol{R^3}$, *with* **v** *not a multiple of* $\boldsymbol{u}$, then Span $\{\boldsymbol{u}, \boldsymbol{v}\}$ is the plane in $\boldsymbol{R^3}$ that contains $\boldsymbol{u}$, $\boldsymbol{v}$ and **0**. In particular, Span $\{\boldsymbol{u}, \boldsymbol{v}\}$ contains the line in $\boldsymbol{R^3}$ through $\boldsymbol{u}$ and **0** and the line through $\boldsymbol{v}$ and **0**.

Example 6: Let $a_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, a_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}, and\ b = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}.$

Then Span $\{\boldsymbol{a_1}, \boldsymbol{a_2}\}$ is a plane through the origin in $\boldsymbol{R^3}$. Is $\boldsymbol{b}$ in that plane?

Solution: First we see the equation $x_1\boldsymbol{a_1} + x_2\boldsymbol{a_2} = \boldsymbol{b}$ has a solution?

To answer this, row-reduce the augmented matrix $[\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \boldsymbol{b}]$:

$$\begin{bmatrix} 1 & 5 & -3 \\ -2 & -13 & 8 \\ 3 & -3 & 1 \end{bmatrix}$$

$By\ R_2 + 2R_1$

$$\sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 18 & 10 \end{bmatrix}$$

$By\ R_3 + 6R_2$

$$\sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{bmatrix}$$

Last row $\Rightarrow 0x_2 = -2$ which can not be true for any value of $x_2 \in \mathbb{R}$

$\Rightarrow$ Given system has no solution

$\therefore b \notin Span\{a_1, a_2\}$ and

in geometrical meaning, vector b does not lie in the plane spanned by vectors a_1 *and* a_2

Linear Combinations in Applications:

The final example shows how scalar multiples and linear combinations can arise when a quantity such as "cost" is broken down into several categories. The basic principle for the example concerns the cost of producing several units of an item when the cost per unit is known:

$$\begin{Bmatrix}\text{number}\\ \text{of units}\end{Bmatrix}\cdot\begin{Bmatrix}\cos t\\ \text{per unit}\end{Bmatrix}=\begin{Bmatrix}\text{total}\\ \cos t\end{Bmatrix}$$

Example 7: A Company manufactures two products. For one dollar's worth of product B, the company spends \$0.45 on materials, \$0.25 on labor, and \$0.15 on overhead. For one dollar's worth of product C, the company spends \$0.40 on materials, \$0.30 on labor and \$0.15 on overhead.

Let $b=\begin{bmatrix}.45\\ .25\\ .15\end{bmatrix}$ *and* $c=\begin{bmatrix}.40\\ .30\\ .15\end{bmatrix}$, then ***b*** and ***c*** represent the "costs per dollar of income" for the two products.

a) What economic interpretation can be given to the vector 100***b***?

b) Suppose the company wishes to manufacture x_1 dollars worth of product B and x_2 dollars worth of product C. Give a vector that describes the various costs the company will have (for materials, labor and overhead).

Solution:

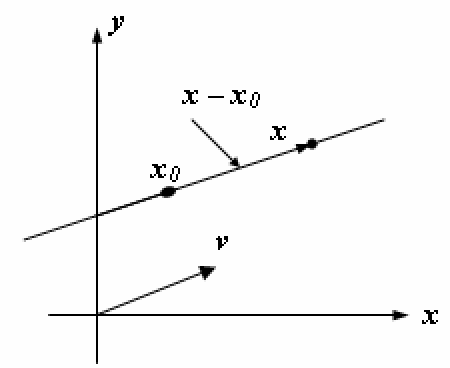
(a) We have $100b=100\begin{bmatrix}.45\\ .25\\ .15\end{bmatrix}=\begin{bmatrix}45\\ 25\\ 15\end{bmatrix}$

The vector 100***b*** represents a list of the various costs for producing \$100 worth of product B, namely, \$45 for materials, \$25 for labor, and \$15 for overhead.

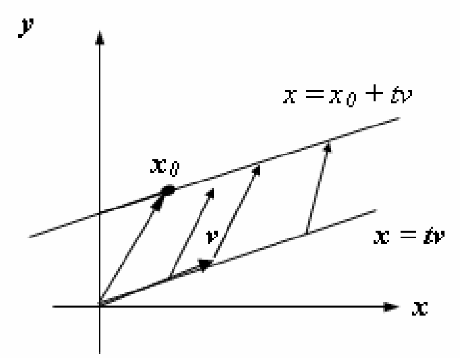
(b) The costs of manufacturing x_1 dollars worth of B are given by the vector $x_1\boldsymbol{b}$ and the costs of manufacturing x_2 dollars worth of C are given by $x_2\boldsymbol{c}$. Hence the total costs for both products are given by the vector $x_1\boldsymbol{b} + x_2\boldsymbol{c}$.

Vector Equation of a Line:

Let $\mathbf{x_0}$ be a fixed point on the line and $\mathbf{v}$ be a nonzero vector that is parallel to the required line. Thus, if $\mathbf{x}$ is a variable point on the line through $\mathbf{x_0}$ that is parallel to $\mathbf{v}$, then the vector $\mathbf{x} - \mathbf{x_0}$ is a vector parallel to $\mathbf{v}$ as shown in fig below,



(b)



So by definition of parallel vectors $\boldsymbol{x} - \boldsymbol{x_0} = t\boldsymbol{v}$ *for some scalar t.*

t is also called a ***parameter*** which varies from $-\infty$ to $+\infty$. The variable point x traces out the line, so the line can be represented by the equation

$$\boldsymbol{x} - \boldsymbol{x_0} = t\boldsymbol{v} \text{ --------------(1)} \qquad (-\infty < t < +\infty)$$

This is a ***vector equation of the line*** through $\boldsymbol{x_0}$ and parallel to $\boldsymbol{v}$.

In the special case where $\boldsymbol{x_0} = \mathbf{0}$, the line passes through the origin, it simplifies to

$$\boldsymbol{x} = t\boldsymbol{v} \qquad (-\infty < t < +\infty)$$

Parametric Equations of a Line in R^2:

Let $\boldsymbol{x} = (x, y) \in R^2$ be a general point of the line through $\boldsymbol{x_0} = (x_0, y_0) \in R^2$ which is parallel to

$\boldsymbol{v} = (a, b) \in R^2$, then eq. 1 takes the form

$$(x, y) - (x_0, y_0) = t(a, b) \qquad (-\infty < t < +\infty)$$

$$\Rightarrow (x - x_0 , y - y_0) = (ta , tb) \qquad (-\infty < t < +\infty)$$

$$\Rightarrow x = x_0 + at, \quad y = y_0 + bt \quad (-\infty < t < +\infty)$$

These are called ***parametric equations*** of the line $in \ R^2$.

<u>Parametric Equations of a Line in R^3:</u>

Similarly, if we let $\boldsymbol{x} = (x, y, z) \in R^3$ be a general point on the line through $\boldsymbol{x_0} = (x_0, y_0 , z_0) \in R^3$ that is parallel to $\boldsymbol{v} = (a, b, c) \in R^3$, then again eq. 1 takes the form

$$(x, y, z) = (x_0, y_0, z_0) + t(a, b, c) \quad (-\infty < t < +\infty)$$

$$\Rightarrow \quad x = x_0 + at, \quad y = y_0 + bt, \quad z = z_0 + ct \quad (-\infty < t < +\infty)$$

These are the ***parametric equations*** of the line $in \ R^3$

<u>Example 8:</u>

(a) Find a vector equation and parametric equations of the line in $\boldsymbol{R^2}$ that passes through the origin and is parallel to the vector $\boldsymbol{v} = (–2, 3)$.

(b) Find a vector equation and parametric equations of the line in $\boldsymbol{R^3}$ that passes through the point $P_0(1, 2, –3)$ and is parallel to the vector $\boldsymbol{v} = (4, –5, 1)$.

(c) Use the vector equation obtained in part (b) to find two points on the line that are different from P_0.

<u>Solution:</u>

(a) We know that a vector equation of the line passing through origin is $\boldsymbol{x = tv}$.

Let $\boldsymbol{x} = (x, y)$ then this equation can be expressed in component form as

$$(x, y) = t\,(–2, 3)$$

This is the vector equation of the line.

Equating corresponding components on the two sides of this equation yields the parametric equations $\quad x = – 2t, \quad y = 3t$

(b) The vector equation of the line is $\boldsymbol{x} = \boldsymbol{x_0} + t\boldsymbol{v}$.

Let $\boldsymbol{x} = (x, y, z)$, Here $\boldsymbol{x_0} = (1, 2, -3)$ and $\boldsymbol{v} = (4, -5, 1)$, then above equation can be expressed in component form as

$$(x, y, z) = (1, 2, -3) + t\,(4, -5, 1)$$

Equating corresponding components on the two sides of this equation yields the parametric equations

$$x = 1 + 4t, \quad y = 2 - 5t, \quad z = -3 + t$$

(d) Specific points on a line can be found by substituting numerical values for the parameter t.

For example, if we take $t = 0$ in part (b), we obtain the point $(x, y, z) = (1, 2, -3)$, which is the given point P_0.

$t = 1$ yields the point $(5, -3, -2)$ and

$t = -1$ yields the point $(-3, 7, -4)$.

Vector Equation of a Plane:

Let x_0 be a fixed point on the required plane W and $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ be two nonzero vectors that are parallel to W and are not scalar multiples of one another. If x is any variable point in the plane W. Suppose $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ have their initial points at x_0, we can create a parallelogram with adjacent side's $t_1\boldsymbol{v_1}$ and $t_2\boldsymbol{v_2}$ in which $\boldsymbol{x} - \boldsymbol{x_0}$ is the diagonal given by the sum

$$\boldsymbol{x} - \boldsymbol{x_0} = t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2}$$

or, equivalently, $\boldsymbol{x} = \boldsymbol{x_0} + t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2}$ **---------------------(1)**

where t_1 and t_2 *are parameters* vary independently from $-\infty$ to $+\infty$,

This is a ***vector equation of the plane*** through $\boldsymbol{x_0}$ and parallel to the vectors $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$. In the special case where $x_0 = 0$, then vector equation of the plane passes through the origin takes the form

$$\boldsymbol{x} = t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2} \qquad (-\infty < t_1 < +\infty, -\infty < t_2 < +\infty)$$

Parametric Equations of a Plane:

Let $\boldsymbol{x} = (x, y, z)$ be a general or variable point in the plane passes through a fixed point $\boldsymbol{x_0} = (x_0, y_0, z_0)$ and parallel to the vectors $\boldsymbol{v_1} = (a_1, b_1, c_1)$ and $\boldsymbol{v_2} = (a_2, b_2, c_2)$, then the component form of eq. 1 will be

$$(x, y, z) = (x_0, y_0, z_0) + t_1(a_1, b_1, c_1) + t_2(a_2, b_2, c_2)$$

Equating corresponding components, we get

$$x = x_0 + a_1t_1 + a_2t_2$$
$$y = y_0 + b_1t_1 + b_2t_2 \qquad (-\infty < t_1 < +\infty, -\infty < t_2 < +\infty)$$
$$z = z_0 + c_1t_1 + c_2t_2$$

These are called the parametric equations for this plane.

Example 9: (Vector and Parametric Equations of Planes)

(a) Find vector and parametric equations of the plane that passes through the origin of $\boldsymbol{R^3}$ and is parallel to the vectors $\boldsymbol{v_1} = (1, -2, 3)$ and $\boldsymbol{v_2} = (4, 0, 5)$.

(b) Find three points in the plane obtained in part (a).

Solution:

(a) As vector equation of the plane passing through origin is $x = t_1v_1 + t_2v_2$.

Let $\boldsymbol{x} = (x, y, z)$ then this equation can be expressed in component form as

$$(x, y, z) = t_1(1, -2, 3) + t_2 (4, 0, 5)$$

This is the vector equation of the plane.

Equating corresponding components, we get

$$x = t_1 + 4t_2, \quad y = -2t_1, \quad z = 3t_1 + 5t_2$$

These are the parametric equations of the plane.

(b) Points in the plane can be obtained by assigning some real values to the parameters t_1 and t_2:

$t_1 = 0$ and $t_2 = 0$ produces the point (0, 0, 0)

$t_1 = -2$ and $t_2 = 1$ produces the point (2, 4, –1)

$t_1 = ½$ and $t_2 = ½$ produces the point (5/2, –1, 4)

Vector equation of Plane through Three Points:

If x_0, x_1 and x_2 are three non collinear points in the required plane. Then obviously the vectors $\mathbf{v_1} = \mathbf{x_1} - \mathbf{x_0}$ and $\mathbf{v_2} = \mathbf{x_2} - \mathbf{x_0}$ are parallel to the plane. So a vector equation of the plane is

$$\mathbf{x} = \mathbf{x_0} + t_1(\mathbf{x_1} - \mathbf{x_0}) + t_2(\mathbf{x_2} - \mathbf{x_0})$$

Example: Find vector and parametric equations of the plane that passes through the points. $P(2, -4, 5)$, $Q(-1, 4, -3)$ and $R(1, 10, -7)$.

Solution:

Let $\mathbf{x} = (x, y, z)$, and if we take $\mathbf{x_0}, \mathbf{x_1}$ and $\mathbf{x_2}$ to be the points P, Q and R respectively, then

$$x_1 - x_0 = \overrightarrow{PQ} = (-3, 8, -8) \quad \text{and} \quad x_2 - x_0 = \overrightarrow{PR} = (-1, 14, -12)$$

So the component form will be

$$(x, y, z) = (2, -4, 5) + t_1(-3, 8, -8) + t_2(-1, 14, -12)$$

This is the required vector equation of the plane.

Equating corresponding components, we get

$$x = 2 - 3t_1 - t_2, \quad y = -4 + 8t_1 + 14t_2, \quad z = 5 - 8t_1 - 12t_2$$

These are the parametric equations of the required plane.

Question: How can you tell from here that the points P, Q and R are not collinear?

Finding a Vector Equation from Parametric Equations

Example 11: Find a vector equation of the plane whose parametric equations are

$$x = 4 + 5t_1 - t_2, \quad y = 2 - t_1 + 8t_2, \quad z = t_1 + t_2$$

Solution: First we rewrite the three equations as the single vector equation

$$(x, y, z) = (4 + 5t_1 - t_2, 2 - t_1 + 8t_2, t_1 + t_2)$$
$$\Rightarrow (x, y, z) = (4, 2, 0) + (5t_1, -t_1, t_1) + (-t_2, 8t_2, t_2)$$
$$\Rightarrow (x, y, z) = (4, 2, 0) + t_1(5, -1, 1) + t_2(-1, 8, 1)$$

This is a vector equation of the plane that passes through the point (4, 2, 0) and is parallel to the vectors $\mathbf{v_1}$ = (5, –1, 1) and $\mathbf{v_2}$ = (–1, 8, 1).

Finding Parametric Equations from a General Equation

Example 12: Find parametric equations of the plane $x - y + 2z = 5$.

Solution: First we solve the given equation for x in terms of y and z

$$x = 5 + y - 2z$$

Now make y and z into parameters, and then express x in terms of these parameters.

Let y = t_1 and $z = t_2$

Then the parametric equations of the given plane are

$$x = 5 + t_1 - 2t_2, \quad y = t_1, \quad z = t_2$$

Exercises:

1. Prove that $\mathbf{u + v = v + u}$ for any $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^n$.

2. For what value(s) of h will y be in Span $\{v_1, v_2, v_3\}$ if

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix} \quad v_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, and \quad y = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

Determine if ***b*** is a linear combination of $\boldsymbol{a_1}$, $\boldsymbol{a_2}$, and $\boldsymbol{a_3}$.

3. $a_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, a_2 = \begin{bmatrix} -2 \\ 3 \\ -2 \end{bmatrix}, a_3 = \begin{bmatrix} -6 \\ 7 \\ 5 \end{bmatrix}, b = \begin{bmatrix} 11 \\ -5 \\ 9 \end{bmatrix}$

4. $a_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, a_2 = \begin{bmatrix} -4 \\ 3 \\ 8 \end{bmatrix}, a_3 = \begin{bmatrix} 2 \\ 5 \\ -4 \end{bmatrix}, b = \begin{bmatrix} 3 \\ -7 \\ -3 \end{bmatrix}$

Determine if ***b*** is a linear combination of the vectors formed from the columns of the matrix ***A***.

5. $A = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 5 & 0 \\ 2 & 5 & 8 \end{bmatrix}, b = \begin{bmatrix} -5 \\ 11 \\ -7 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 0 & 5 \\ -2 & 1 & -6 \\ 0 & 2 & 8 \end{bmatrix}, b = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$

In exercises 3-6, list seven vectors in Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. For each vector, show the weights on $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ used to generate the vector and list the three entries of the vector. Give also geometric description of the Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$.

7. $v_1 = \begin{pmatrix} 5 \\ -1 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix}$

8. $v_1 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$

9. $v_1 = \begin{pmatrix} 2 \\ 6 \\ -4 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -3 \\ -9 \\ 6 \end{pmatrix}$

10. $v_1 = \begin{pmatrix} -3.7 \\ -0.4 \\ 11.2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 5.8 \\ 2.1 \\ 5.3 \end{pmatrix}$

11. Let $a_1 = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}, a_2 = \begin{bmatrix} -5 \\ -8 \\ 2 \end{bmatrix}, b = \begin{bmatrix} 3 \\ -5 \\ h \end{bmatrix}$. For what value(s) of h is ***b*** in the plane spanned by $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$?

12. Let $v_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, v_2 = \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix}, and\ y = \begin{bmatrix} h \\ -3 \\ -5 \end{bmatrix}$. For what value(s) of h is ***y*** in the plane generated by $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$?

13. Let $u = \begin{bmatrix} 2 \\ -1 \end{bmatrix} and \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Show that $\begin{bmatrix} h \\ k \end{bmatrix}$ is in Span{u, v} for all h and k.

Lecture 6

Matrix Equations

A fundamental idea in linear algebra is to view a linear combination of vectors as the product of a matrix and a vector. The following definition will permit us to rephrase some of the earlier concepts in new ways.

Definition: If ***A*** is an $m \times n$ matrix, with columns $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_n}$ and if ***x*** is in $\boldsymbol{R^n}$, then the product of ***A*** and ***x*** denoted by ***Ax***, is the linear combination of the columns of ***A*** using the corresponding entries in ***x*** as weights, that is,

$$Ax = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 a_1 + x_2 a_2 + \ldots + x_n a_n$$

Note that ***Ax*** is defined only if the number of columns of ***A*** equals the number of entries in ***x***.

Example 1

a) $$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

b) $$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$

Example 2: For $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$ in $\boldsymbol{R^m}$, write the linear combination $3\boldsymbol{v_1} - 5\boldsymbol{v_2} + 7\boldsymbol{v_3}$ as a matrix times a vector.

Solution: Place $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$ into the columns of a matrix ***A*** and place the weights 3, -5, and 7 into a vector ***x***.

That is, $$3v_1 - 5v_2 + 7v_3 = [v_1 \quad v_2 \quad v_3] \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} = Ax$$

We know how to write a system of linear equations as a vector equation involving a linear combination of vectors. For example, we know that the system

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 4 \\ -5x_2 + 3x_3 &= 1 \end{aligned} \qquad \text{is equivalent to} \qquad x_1\begin{bmatrix}1\\0\end{bmatrix} + x_2\begin{bmatrix}2\\-5\end{bmatrix} + x_3\begin{bmatrix}-1\\3\end{bmatrix} = \begin{bmatrix}4\\1\end{bmatrix}$$

Writing the linear combination on the left side as a matrix times a vector, we get

$$\begin{bmatrix}1 & 2 & -1\\0 & -5 & 3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}4\\1\end{bmatrix}$$

Which has the form ***Ax = b***, and we shall call such an equation a **matrix equation**, to distinguish it from a vector equation.

Theorem: 1 If ***A*** is an $m \times n$ matrix, with columns $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_n}$ and if ***b*** is in $\boldsymbol{R^m}$, the matrix equation ***Ax = b*** has the same solution set as the vector equation

$$x_1\boldsymbol{a_1} + x_2\boldsymbol{a_2} + \ldots + x_n\boldsymbol{a_n} = \boldsymbol{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is $\begin{bmatrix}a_1 & a_2 & \ldots & a_n & b\end{bmatrix}$

Existence of Solutions: The equation ***Ax = b*** has a solution if and only if ***b*** is a linear combination of the columns of ***A***.

Example 3: Let $A = \begin{bmatrix}1 & 3 & 4\\-4 & 2 & -6\\-3 & -2 & -7\end{bmatrix}$ and $b = \begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$.

Is the equation ***Ax = b*** consistent for all possible b_1, b_2, b_3?

Solution Row reduce the augmented matrix for ***Ax = b***:

$$\begin{bmatrix}1 & 3 & 4 & b_1\\-4 & 2 & -6 & b_2\\-3 & -2 & -7 & b_3\end{bmatrix} \sim \overset{4R_1 + R_2,\, 3R_1 + R_3}{\begin{bmatrix}1 & 3 & 4 & b_1\\0 & 14 & 10 & b_2 + 4b_1\\0 & 7 & 5 & b_3 + 3b_1\end{bmatrix}}$$

$$R_3 - \frac{1}{2}R_2$$

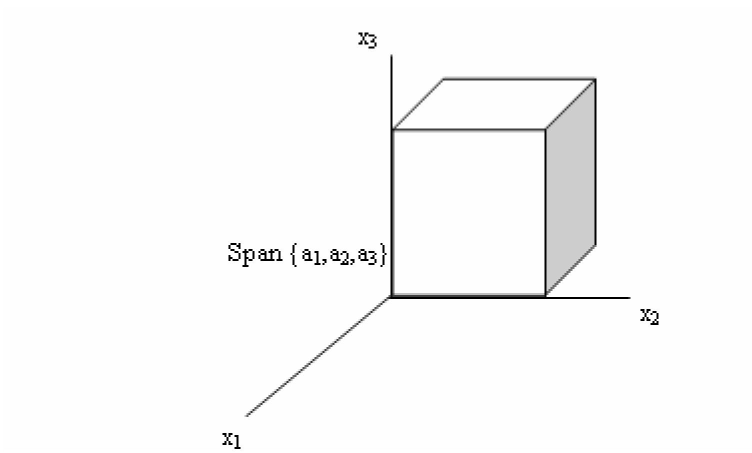
$$\sim \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{bmatrix}$$

The third entry in the augmented column is $b_1 - \frac{1}{2}b_2 + b_3$.

The equation ***Ax* = *b*** is not consistent for every ***b*** because some choices of ***b*** can make $b_1 - \frac{1}{2}b_2 + b_3$ nonzero.

The entries in ***b*** must satisfy $b_1 - \frac{1}{2}b_2 + b_3 = 0$

This is the equation of a plane through the origin in $\boldsymbol{R^3}$. The plane is the set of all linear combinations of the three columns of ***A***. See figure below.



The equation ***Ax* = *b*** fails to be consistent for all ***b*** because the echelon form of ***A*** has a row of zeros. If ***A*** had a pivot in all three rows, we would not care about the calculations in the augmented column because in this case an echelon form of the augmented matrix could not have a row such as [0 0 0 1].

Example 4: Which of the following are linear combinations of

$$A = \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

(a) $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ **(c)** $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix}$

Solution:

(a) $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$ = a A + b B + c C

$$= a\begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + b\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + c\begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4a+b & -b+2c \\ -2a+2b+c & -2a+3b+4c \end{bmatrix}$$

$\Rightarrow$ 4a + b = 6 (1)

-b + 2c = -8 (2)

-2a + 2b + c = -1 (3)

-2a + 3b + 4c = -8 (4)

Subtracting (4) from (3), we obtain

-b – 3c = 7 (5)

Subtracting (5) from (2):

5c = -15 $\Rightarrow$ c = -3

From (2), -b + 2(-3) = -8 $\Rightarrow$ b = 2

From (3), -2a + 2(2) – 3 = -1 $\Rightarrow$ a = 1

Now we check whether these values satisfy (1).

4(1) + 2 = 6

It means that $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$ is the linear combination of **A**, **B** and **C**.

Thus

$$\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix} = 1A + 2B - 3C$$

(b) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = a\,A + b\,B + c\,C$

$$= a\begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + b\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + c\begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4a + b & -b + 2c \\ -2a + 2b + c & -2a + 3b + 4c \end{bmatrix}$$

$\Rightarrow$ $4a + b = 0$ (1)

$-b + 2c = 0$ (2)

$-2a + 2b + c = 0$ (3)

$-2a + 3b + 4c = 0$ (4)

Subtracting eq. 3 from eq. 4 we get

$b + 3c = 0$ (5)

Adding eq. 2 and eq. 5, we get

$5c = 0 \Rightarrow c = 0$

Put c = 0 in eq. 5, we get b = 0

Put b = c = 0 in eq. 3, we get a = 0

$\Rightarrow$ $a = b = c = 0$

It means that $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is the linear combination of **A**, **B** and **C**.

Thus $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0A + 0B + 0C$

(c) $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix} = a\,A + b\,B + c\,C$

$$= a\begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + b\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + c\begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4a+b & -b+2c \\ -2a+2b+c & -2a+3b+4c \end{bmatrix}$$

$\Rightarrow$ 4a + b = 6 (1)

-b + 2c = 0 (2)

-2a + 2b + c = 3 (3)

-2a + 3b + 4c = 8 (4)

Subtracting (4) from (3), we obtain

-b – 3c = -5 (5)

Subtracting (5) from (2):

5c = 5 $\Rightarrow$ c = 1

From (2), -b + 2(1) = 0 $\Rightarrow$ b = 2

From (3), -2a + 2(2) + 1 = 3 $\Rightarrow$ a = 1

Now we check whether these values satisfy (1).

4(1) + 2 = 6

It means that $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix}$ is the linear combination of **A**, **B** and **C**.

Thus $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix}$ = 1**A** + 2**B** +1**C**

Theorem 2: Let ***A*** be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular ***A***, either they are all true statements or they are all false.

(a) For each ***b*** in $\boldsymbol{R^m}$, the equation ***Ax = b*** has a solution.

(b) The columns of ***A*** Span $\boldsymbol{R^m}$.

(c) A has a pivot position in every row.

This theorem is one of the most useful theorems. It is about a coefficient matrix, not an augmented matrix. If an augmented matrix [***A*** **b**] *has a pivot position in every row, then the equation* ***Ax = b*** *may or may not be consistent.*

Example 4: Compute ***Ax***, where $A = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$ *and* $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Solution From the definition,

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 \\ -x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix}$$

Note:

In above example the first entry in the product ***Ax*** is a sum of products (sometimes called a **dot product**), using the first row of ***A*** and the entries in ***x***.

That is $\begin{bmatrix} 2 & 3 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \end{bmatrix}$

Examples:

In each part determine whether the given vector span R^3

(a) $v_1 = (2,\ 2,\ 2),\ v_2 = (0,\ 0,\ 3),$
$v_3 = (0, 1, 1)$

(b) $v_1 = (3, 1, 4),\ v_2 = (2, -3, 5),$
$v_3 = (5, -2, 9),\ v_4 = (1,\ 4,\ -1)$

(c) $v_1 = (1,\ 2, 6),\ v_2 = (3,\ 4,\ 1),$
$v_3 = (4, 3, 1),\ v_4 = (3,\ 3,\ 1)$

Solutions:

(a) We have to determine whether arbitrary vectors $b=(b_1, b_2, b_3)$ in R^3 can be expressed as a linear combination $b=k_1v_1+k_2v_2+k_3v_3$ of the vectors v_1,v_2,v_3

Expressing this in terms of components given by

$$(b_1, b_2, b_3)=k_1(2,2,2)+k_2(0,0,3)+k_3(0,1,1)$$
$$(b_1, b_2, b_3)=(2k_1+0k_2+0k_3, 2k_1+0k_2+k_3, 2k_1+3k_2+k_3)$$
$$2k_1+0k_2+0k_3=b_1$$
$$2k_1+0k_2+k_3=b_2$$
$$2k_1+3k_2+k_3=b_3$$

$$A=\begin{bmatrix}2&0&0\\2&0&1\\2&3&1\end{bmatrix}$$ has a non zero determinant

Now

$$\det(A)=-6\neq 0$$

Therefore v_1,v_2,v_3 span R^3

(b) The set S{ v_1,v_2,v_3,v_4 }of vectors in R^3 spans V= R^3 if

$$c_1v_1+c_2v_2+c_3v_3+c_4v_4=d_1w_1+d_2w_2+d_3w_3 \quad ....(1)$$

with

$$w_1=(1,0,0)$$
$$w_2=(0,1,0)$$
$$w_3=(0,0,1)$$

With our vectors v_1,v_2,v_3,v_4 equation (1) becomes

$$c_1(3,1,4)+c_2(2,-3,5)+c_3(5,-2,9)+c_4(1,4,-1)=d_1(1,,0,0)+d_2(0,1,0)+d_3(0,0,1)$$

Rearranging the left hand side yields

$$3c_1+2c_2+5c_3+1c_4=1d_1+0d_2+0d_3$$

$$1c_1-3c_2-2c_3+4c_4=0d_1+1d_2+0d_3$$

$$4c_1+5c_2+9c_3-1c_4=0d_1+0d_2+1d_3$$

$$\begin{bmatrix}3&2&5&1&1&0&0\\1&-3&-2&4&0&1&0\\4&5&9&-1&0&0&1\end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 1 & -1 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & -3 & -2 \end{bmatrix}$$

The reduce row echelon form

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & \frac{5}{17} & \frac{3}{17} \\ 0 & 1 & 1 & -1 & 0 & \frac{-4}{17} & \frac{1}{17} \\ 0 & 0 & 0 & 0 & 1 & \frac{-7}{17} & -\frac{11}{17} \end{bmatrix}$$ Corresponds to the system of equations

$$1c_1 + 1c_3 + 1c_4 = (\frac{5}{17})d_2 + (\frac{3}{17})d_3$$
$$1c_2 + 1c_3 + -1c_4 = (\frac{-4}{17})d_2 + (\frac{1}{17})d_3 \qquad \text{……………(2)}$$

$$0 = 1d_1 + (\frac{-7}{17})d_2 + (-\frac{11}{17})d_3$$

So this system is inconsistent. The set S does not spans the space V.

Similarly Part C can be solved by the same way as above.

Exercise:

1. Let $A = \begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix}, x = \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}, and \quad b = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$.

   It can be shown that ***Ax* = *b***. Use this fact to exhibit ***b*** as a specific linear combination of the columns of ***A***.

2. Let $A = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}, u = \begin{bmatrix} 4 \\ -1 \end{bmatrix}, and \quad v = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. Verify ***A(u + v)* = *Au* +*Av***.

3. Solve the equation ***Ax = b***, with $A = \begin{bmatrix} 2 & 4 & -6 \\ 0 & 1 & 3 \\ -3 & -5 & 7 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 5 \\ -3 \end{bmatrix}$.

4. Let $u = \begin{bmatrix} -5 \\ -3 \\ -6 \end{bmatrix} and\ A = \begin{bmatrix} 3 & 5 \\ 1 & 1 \\ -2 & -8 \end{bmatrix}$. Is ***u*** in the plane in R^3 spanned by the columns of A? Why or why not?

5. Let $u = \begin{bmatrix} 8 \\ 2 \\ 3 \end{bmatrix} and\ A = \begin{bmatrix} 4 & 3 & 5 \\ 0 & 1 & -1 \\ 1 & 2 & 0 \end{bmatrix}$. Is ***u*** in the subset of R^3 spanned by the columns of A? Why or why not?

6. Let $A = \begin{bmatrix} -3 & 1 \\ 6 & -2 \end{bmatrix} and\ b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Show that the equation ***Ax = b*** is not consistent for all possible ***b***, and describe the set of all ***b*** for which ***Ax = b*** is consistent.

7. How many rows of $A = \begin{bmatrix} 1 & 3 & -2 & -2 \\ 0 & 1 & -1 & 5 \\ -1 & -2 & 1 & 7 \\ 1 & 1 & 0 & -6 \end{bmatrix}$ contain pivot positions?

In exercises 8 to 13 , explain how your calculations justify your answer, and mention an appropriate theorem.

8. Do the columns of the matrix $A = \begin{bmatrix} 1 & 3 & -4 \\ 3 & 2 & -6 \\ -5 & -1 & 8 \end{bmatrix}$ span R^3?

9. Do the columns of the matrix $A = \begin{bmatrix} 1 & 3 & -2 & -2 \\ 0 & 1 & -1 & 5 \\ -1 & -2 & 1 & 7 \\ 1 & 1 & 0 & -6 \end{bmatrix}$ span R^4?

10. Do the columns of the matrix $A=\begin{bmatrix}0 & 0 & 2\\ 0 & -5 & 1\\ 4 & 6 & -3\end{bmatrix}$ span R^3?

11. Do the columns of the matrix $A=\begin{bmatrix}3 & 5\\ 1 & 1\\ -2 & -8\end{bmatrix}$ span R^3?

12. Let $v_1=\begin{bmatrix}1\\0\\-1\\0\end{bmatrix}, v_2=\begin{bmatrix}0\\1\\0\\-1\end{bmatrix}, v_3=\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$. Does $\{v_1, v_2, v_3\}$ span R^4?

13. Let $v_1=\begin{bmatrix}1\\0\\-1\end{bmatrix}, v_2=\begin{bmatrix}-1\\3\\7\end{bmatrix}, v_3=\begin{bmatrix}3\\-2\\-2\end{bmatrix}$. Does $\{v_1, v_2, v_3\}$ span R^3?

14. It can be shown that $\begin{bmatrix}4 & 1 & 2\\ -2 & 0 & 8\\ 3 & 5 & -6\end{bmatrix}\begin{bmatrix}-1\\4\\2\end{bmatrix}=\begin{bmatrix}4\\18\\5\end{bmatrix}$. Use this fact(and no row operations) to find scalars c_1, c_2, c_3 such that $\begin{bmatrix}4\\18\\5\end{bmatrix}=c_1\begin{bmatrix}4\\-2\\3\end{bmatrix}+c_2\begin{bmatrix}1\\0\\5\end{bmatrix}+c_3\begin{bmatrix}2\\8\\-6\end{bmatrix}$.

15. Let $u=\begin{bmatrix}3\\8\\4\end{bmatrix}, v=\begin{bmatrix}1\\3\\1\end{bmatrix}, and\ w=\begin{bmatrix}1\\1\\3\end{bmatrix}$. It can be show that 2u – 5v – w = 0. Use this fact(and no row operations) to solve the equation $\begin{bmatrix}3 & 1\\ 8 & 3\\ 4 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}1\\1\\3\end{bmatrix}$.

Determine if the columns of the matrix span R^4.

16. $\begin{bmatrix} 7 & 2 & -5 & 8 \\ -5 & -3 & 4 & -9 \\ 6 & 10 & -2 & 7 \\ -7 & 9 & 2 & 15 \end{bmatrix}$

17. $\begin{bmatrix} 12 & -7 & 11 & -9 & 5 \\ -9 & 4 & -8 & 7 & -3 \\ -6 & 11 & -7 & 3 & -9 \\ 4 & -6 & 10 & -5 & 12 \end{bmatrix}$

Lecture 7

Solution Sets of Linear Systems

Solution Set:

A solution of a linear system is an assignment of values to the variables x_1, x_2,... , x_n such that each of the equations in the linear system is satisfied. The set of all possible solutions is called the Solution Set

Homogeneous Linear Systems

A system of linear equations is said to be **homogeneous** if it can be written in the form ***Ax = 0***, where ***A*** is an $m \times n$ matrix and ***0*** is the zero vector in $\boldsymbol{R}^m$.

Trivial solution:
A homogeneous system ***Ax = 0*** always has at least one solution, namely, ***x = 0*** (the zero vector in $\boldsymbol{R}^n$). This zero solution is usually called the trivial solution of the homogeneous system.

Nontrivial solution:
A solution of a linear system other than trivial is called its nontrivial solution.
i.e the solution of a homogenous equation ***Ax = 0*** such that $\boldsymbol{x} \neq \boldsymbol{0}$ is called **nontrivial solution**, that is, a nonzero vector ***x*** that satisfies ***Ax = 0***.

Existence and Uniqueness Theorem:

The homogeneous equation ***Ax = 0*** *has a nontrivial solution if and only if the equation has at least one free variable.*

Example 1 Find the solution set of the following system

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 0 \\ 3x_1 + 2x_2 - 4x_3 &= 0 \\ 6x_1 + x_2 - 8x_3 &= 0 \end{aligned}$$

Solution.

$$\text{Let } A = \begin{bmatrix} 3 & 5 & -4 \\ 3 & 2 & -4 \\ 6 & 1 & -8 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The augmented matrix is

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ 3 & 2 & -4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix}$$

For solution set, row reduce to reduced echelon form

$$\sim\begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} \qquad -1R_1 + R_2, -2R_1 + R_3$$

$$\sim\begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad -3R_2 + R_3$$

$$\sim\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad 1/3R_1,\ 1/3R_2,\ 5/3R_2 + R_1$$

$$\sim\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (-1)R_2$$

$$\begin{aligned} x_1 \quad - \frac{4}{3}x_3 &= 0 \\ x_2 \quad &= 0 \\ 0 &= 0 \end{aligned}$$

It is clear that ***x₃*** is a free variable, so ***Ax = 0*** has nontrivial solutions (one for each choice of ***x₃***). From above equations we have,

$$x_1 = \frac{4}{3}x_3, \quad x_2 = 0, \text{ with } \boldsymbol{x_3} \text{ free.}$$

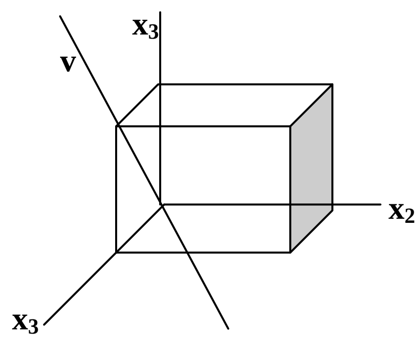
As a vector, the general solution of ***Ax = 0*** is given by

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} = x_3 v, \quad where \quad v = \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

This shows that every solution of $Ax = 0$ in this case is a scalar multiple of v (it means that v generate or spans the whole general solution).The trivial solution is obtained by choosing $x_3 = 0$.

Geometric Interpretation:

Geometrically, the solution set is a line through 0 in R^3, as given in the **Figure below**



Note: A nontrivial solution x can have some zero entries so long as not all of its entries are zero.

Example 2:

Solve the following system

$$10x_1 - 3x_2 - 2x_3 = 0 \qquad (1)$$

Solution: We solve for the basic variable x_1 in terms of the free variables.
Dividing eq. 1 by 10 and solve for x

$x_1 = 0.3x_2 + 0.2x_3$ *where* x_2 and x_3 free variables.

As a vector, the general solution is

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 + 0.2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.2x_3 \\ 0 \\ x_3 \end{bmatrix}$$

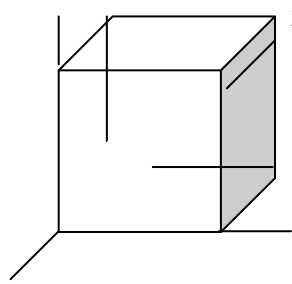
$$= x_2 \begin{bmatrix} 0.3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0.2 \\ 0 \\ 1 \end{bmatrix} \qquad (2)$$

u ***v***

This calculation shows that every solution of (1) is a linear combination of the vector ***u, v*** shown in (2). That is, the solution set is Span {***u, v***}
.

Geometric Interpretation:

Since neither ***u*** nor ***v*** is a scalar multiple of the other so these are not parallel, the solution set is a plane through the origin, see Figure below



Note:

Above examples illustrate the fact that the solution set of a homogeneous equation $Ax = 0$ can be expressed explicitly as Span **{*v*₁, *v*₂, … , *v*ₚ}** for suitable vectors $v_1, v_2, \ldots, v_p$ (because solution sets can be written in the form of linear combination of these vectors). If the only solution is the zero-vector then the solution set is **Span {0}**.

Example 3 (For Practice) Find the solution set of the following homogenous system

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 0 \\ -4x_1 - 9x_2 + 2x_3 &= 0 \\ -3x_2 - 6x_3 &= 0 \end{aligned}$$

Solution:

Let $A = \begin{bmatrix} 1 & 3 & 1 \\ -4 & -9 & 2 \\ 0 & -3 & -6 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

The augmented matrix is

$$\begin{bmatrix} 1 & 3 & 1 & 0 \\ -4 & -9 & 2 & 0 \\ 0 & -3 & -6 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 3 & 6 & 0 \\ 0 & -3 & -6 & 0 \end{bmatrix} \qquad 4R_1 + R_2,$$

$$\sim \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 3 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad R_2 + R_3$$

$$\sim \begin{bmatrix} 1 & 0 & -5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \frac{1}{2}R_2 \ , \ (-3)R_2 + R_1$$

SO

$$\begin{aligned} x_1 \qquad -5x_3 &= 0 \\ x_2 + 2x_3 &= 0 \\ 0 &= 0 \end{aligned}$$

~

From above results, it is clear that ***x*₃** is a free variable, so ***Ax* = *0*** has nontrivial solutions (one for each choice of ***x*₃**).

From above equations we have,

$$x_1 = 5x_3, \quad x_2 = -2x_3, \text{ with } \boldsymbol{x_3} \text{ a free variable.}$$

As a vector, the general solution of ***Ax* = *0*** is given by

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5x_3 \\ -2x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} = x_3 v, \quad where \quad v = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$$

Parametric Vector Form of the solution:

Whenever a solution set is described explicitly with vectors, we say that the solution is in **parametric vector form**:

The equation

$$\boldsymbol{x} = s\boldsymbol{u} + t\boldsymbol{v} \qquad (s, t \text{ in } \boldsymbol{R})$$

is called a **parametric vector equation** of the plane. It is written in this form to emphasize that the parameters vary over all real numbers.

Similarly, the equation $\boldsymbol{x} = \boldsymbol{x_3 v}$ (with $\boldsymbol{x_3}$ free), or $\boldsymbol{x} = \boldsymbol{tv}$ (with $\boldsymbol{t}$ in $\boldsymbol{R}$), is a parametric vector equation of a line.

Solutions of Non-homogeneous Systems:

When a non-homogeneous linear system has many solutions, the general solution can be written in parametric vector form as one vector plus an arbitrary linear combination of vectors that satisfy the corresponding homogeneous system.
To clear this concept consider the following examples,

Example: 5. Describe all solutions of ***Ax = b***, where

$$A=\begin{bmatrix}3 & 5 & -4\\ -3 & -2 & 4\\ 6 & 1 & -8\end{bmatrix} \quad and \quad b=\begin{bmatrix}7\\ -1\\ -4\end{bmatrix}$$

Solution
Row operations on **[A b]** produce

$$\begin{bmatrix}3 & 5 & -4 & 7\\ -3 & -2 & 4 & -1\\ 6 & 1 & -8 & -4\end{bmatrix}$$

$$\sim\begin{bmatrix}3 & 5 & -4 & 7\\ 0 & 3 & 0 & 6\\ 0 & -9 & 0 & -18\end{bmatrix} \qquad R_1+R_2, -2R_1+R_3$$

$$\sim\begin{bmatrix}3 & 5 & -4 & 7\\ 0 & 1 & 0 & 2\\ 0 & 0 & 0 & 0\end{bmatrix} \qquad 3R_2+R_3,\ \frac{1}{3}R_2$$

$$\sim\begin{bmatrix}1 & 0 & -\frac{4}{3} & -1\\ 0 & 1 & 0 & 2\\ 0 & 0 & 0 & 0\end{bmatrix} \qquad -5R_2+R_1,\ \frac{1}{3}R_1$$

$$\begin{aligned} x_1 \quad - \tfrac{4}{3}x_3 &= -1\\ OR \qquad x_2 \quad &= 2\\ 0 &= 0\end{aligned}$$

Thus $x_1 = -1 + \frac{4}{3}x_3$, $x_2 = 2$, and ***x₃*** is free.

As a vector, the general solution of ***Ax = b*** has the form

$$x=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-1+\frac{4}{3}x_3\\2\\x_3\end{bmatrix}=\begin{bmatrix}-1\\2\\0\end{bmatrix}+\begin{bmatrix}\frac{4}{3}x_3\\0\\x_3\end{bmatrix}=\underbrace{\begin{bmatrix}-1\\2\\0\end{bmatrix}}_{p}+x_3\underbrace{\begin{bmatrix}\frac{4}{3}\\0\\1\end{bmatrix}}_{v}$$

The equation $\boldsymbol{x = p + x_3 v}$, or, writing $\boldsymbol{t}$ as a general parameter,

$$\boldsymbol{x = p + tv} \quad (\boldsymbol{t} \text{ in } \boldsymbol{R}) \qquad (3)$$

Note:
We know that the solution set of this question when ***Ax = 0*** (example 1) has the parametric vector equation

$$\boldsymbol{x = tv} \ (\boldsymbol{t} \text{ in } \boldsymbol{R}) \qquad (4)$$

With the same ***v*** that appears in equation (3) in above example.
Thus the solutions of ***Ax = b*** are obtained by adding the vector ***p*** to the solutions of ***Ax = 0***. The vector ***p*** itself is just one particular solution of ***Ax = b*** (corresponding to ***t = 0*** in (3)).

The following theorem gives the precise statement.

Theorem:

Suppose the equation *Ax = b* is consistent for some given *b*, and let *p* be a solution. Then the solution set of *Ax = b* is the set of all vectors of the form $\boldsymbol{w = p + v_h}$, where $\boldsymbol{v_h}$ is any solution of the homogeneous equation Ax = 0.

Example 6: (For pratice)

$$\begin{aligned} x_1+3x_2+x_3 &= 1\\ -4x_1-9x_2+2x_3 &= -1\\ -3x_2-6x_3 &= -3\end{aligned}$$

Solution:

Let $A=\begin{bmatrix}1&3&1\\-4&-9&2\\0&-3&-6\end{bmatrix}$, $X=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$, $b=\begin{bmatrix}1\\-1\\-3\end{bmatrix}$

The augmented matrix is

$$\begin{bmatrix} 1 & 3 & 1 & 1 \\ -4 & -9 & 2 & -1 \\ 0 & -3 & -6 & -3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 3 & 6 & 3 \\ 0 & -3 & -6 & -3 \end{bmatrix} \qquad 4R_1 + R_2,$$

$$\sim \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 3 & 6 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad R_2 + R_3$$

$$\sim \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \frac{1}{3}R_2$$

$$\sim \begin{bmatrix} 1 & 0 & -5 & -2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (-3)R_2 + R_1$$

SO

$$\begin{aligned} x_1 \qquad - 5x_3 &= -2 \\ x_2 + 2x_3 &= 1 \\ 0 &= 0 \end{aligned}$$

Thus $x_1 = -2 + 5x_3$, $x_2 = 1 - 2x_3$, and $\boldsymbol{x_3}$ is free.

As a vector, the general solution of $\boldsymbol{Ax = b}$ has the form

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 + 5x_3 \\ 1 - 2x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 5x_3 \\ -2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}}_{\boldsymbol{p}} + x_3 \underbrace{\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}}_{\boldsymbol{v}}$$

So we can write solution set in parametric vector form as

$$x = p + x_3 v$$

Steps of Writing a Solution Set (of a Consistent System) in a Parametric Vector Form

Step 1:
Row reduces the augmented matrix to reduced echelon form.

Step 2:
Express each basic variable in terms of any free variables appearing in an equation.

Step 3:
Write a typical solution ***x*** as a vector whose entries depend on the free variables if any.

Step 4:
Decompose ***x*** into a linear combination of vectors (with numeric entries) using the free variables as parameters.

Exercise:

Determine if the system has a nontrivial solution. Try to use as few row operations as possible.

1. $$\begin{aligned} x_1 - 5x_2 + 9x_3 &= 0 \\ -x_1 + 4x_2 - 3x_3 &= 0 \\ 2x_1 - 8x_2 + 9x_3 &= 0 \end{aligned}$$

2. $$\begin{aligned} 3x_1 + 6x_2 - 4x_3 - x_4 &= 0 \\ -5x_1 \qquad + 8x_3 + 3x_4 &= 0 \\ 8x_1 - x_2 \qquad\quad + 7x_4 &= 0 \end{aligned}$$

3. $$\begin{aligned} 5x_1 - x_2 + 3x_3 &= 0 \\ 4x_1 - 3x_2 + 7x_3 &= 0 \end{aligned}$$

Write the solution set of the given homogeneous system in parametric vector form.

4. $$\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ x_2 - x_3 &= 0 \\ -2x_1 + 3x_2 + 7x_3 &= 0 \end{aligned}$$

5. $$\begin{aligned} x_1 + 2x_2 - 7x_3 &= 0 \\ -2x_1 - 3x_2 + 9x_3 &= 0 \\ -2x_2 + 10x_3 &= 0 \end{aligned}$$

In exercises 6-8, describe all solutions of $\boldsymbol{Ax = 0}$ in parametric vector form where A is row equivalent to the matrix shown.

6. $$\begin{bmatrix} 1 & -5 & 0 & 2 & 0 & -4 \\ 0 & 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

7. $$\begin{bmatrix} 1 & 6 & 0 & 8 & -1 & -2 \\ 0 & 0 & 1 & -3 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

8. $$\begin{bmatrix} 1 & -5 & 0 & 4 \end{bmatrix}$$

9. Describe the solution set in R^3 of $x_1 - 4x_2 + 3x_3 = 0$, compare it with the solution set of $x_1 - 4x_2 + 3x_3 = 7$.

10. Find the parametric equation of the line through ***a*** parallel to ***b***.

$$a = \begin{bmatrix} 3 \\ -8 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$$

11. Find a parametric equation of the line *M* through ***p*** and ***q***.

$$p = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, q = \begin{bmatrix} 0 \\ 7 \end{bmatrix}$$

12. Given $A = \begin{bmatrix} 5 & 10 \\ -8 & -16 \\ 7 & 14 \end{bmatrix}$, find one nontrivial solution of $\boldsymbol{Ax = 0}$ by inspection.

13. Given $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 9 \end{bmatrix}$, find one nontrivial solution of $\boldsymbol{Ax = 0}$ by inspection.

Lecture 8

Linear Independence

Definition:

An indexed set of vectors **{*v₁*, *v₂*, ... , *v_p*}** in $\boldsymbol{R}^n$ is said to be **linearly independent** if the vector equation $x_1v_1 + x_2v_2 + \cdots + x_pv_p = 0$ has only the trivial solution. The set **{$v_1, v_2, \ldots, v_p$}** is said to be **linearly dependent** if there exist weights $c_1, \ldots, c_p$, not all zero, such that $c_1v_1 + c_2v_2 + \cdots + c_pv_p = 0$ (1)

Equation (1) is called a **linear dependence relation** among $v_1, \ldots, v_p$ when the weights are not all zero.

Example 1:

Let $v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \; v_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \; v_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

(a) Determine whether the set of vectors **{v_1, v_2, v_3}** is linearly independent or not.

(b) If possible, find a linear dependence relation among v_1, v_2, v_3.

Solution:

(a) Row operations on the associated augmented matrix show that

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & -6 & -6 & 0 \end{bmatrix} \qquad (-2)R_1 + R_2, \; (-3)R_1 + R_3$$

$$\sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad R_2 + R_3 \qquad (2)$$

Clearly, x_1 and x_2 are basic variables and x_3 is free. Each nonzero value of x_3 determines a nontrivial solution.

Hence v_1, v_2, v_3 are linearly dependent (and not linearly independent).

(b) To find a linear dependence relation among v_1, v_2, v_3, completely row reduce the augmented matrix and write the new system:

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \frac{-1}{3}R_2$$

$$\sim \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad R_1 - 4R_2$$

$$\begin{aligned} x_1 \qquad & -2x_3 = 0 \\ \Rightarrow \quad x_2 \qquad & +x_3 = 0 \\ & 0 \quad = 0 \end{aligned}$$

Thus $\boldsymbol{x_1 = 2x_3}$, $\boldsymbol{x_2 = -x_3}$, and $\boldsymbol{x_3}$ is free.

Choose any nonzero value for $\boldsymbol{x_3}$, say, $\boldsymbol{x_3 = 5}$, then $\boldsymbol{x_1 = 10}$, and $\boldsymbol{x_2 = -5}$.
Substitute these values into $x_1v_1 + x_2v_2 + x_3v_3 = 0$

$$\Rightarrow \; \boldsymbol{10v_1 - 5v_2 + 5v_3 = 0}$$

This is one (out of infinitely many) possible linear dependence relation among $\boldsymbol{v_1, v_2, v_3}$.

Example (for practice):

Check whether the vectors are linearly dependent or linearly independent

$\mathbf{v}_1 = (3, -1) \qquad \mathbf{v}_2 = (-2, 2)$

Solution:

Consider two constants C_1 and C_2. Suppose

$$c_1(3,-1) + c_2(-2,2) = 0$$
$$(3c_1 - 2c_2, -c_1 + 2c_2) = (0,0)$$

Now, set each of the components equal to zero to arrive at the following system of equations.

$$\begin{aligned} 3c_1 - 2c_2 &= 0 \\ -c_1 + 2c_2 &= 0 \end{aligned}$$

Solving this system gives to following solution,

$$c_1 = 0 \qquad c_2 = 0$$

The trivial solution is the only solution and so these two vectors are linearly independent.

Linear Independence of Matrix Columns:

Suppose that we begin with a matrix $A=\begin{bmatrix} a_1 & ... & a_n \end{bmatrix}$ instead of a set of vectors. The matrix equation ***Ax = 0*** can be written as $x_1a_1 + x_2a_2 + \cdots + x_na_n = 0$

Each linear dependence relation among the columns of ***A*** corresponds to a nontrivial solution of ***Ax = 0***.

Thus we have the following important fact.

The columns of a matrix A are linearly independent if and only if the equation $Ax = 0$ has only the trivial solution.

Example 2: Determine whether the columns of $A=\begin{bmatrix} 0 & 1 & 4 \\ 1 & 2 & -1 \\ 5 & 8 & 0 \end{bmatrix}$ are linearly independent.

Solution. To study ***Ax = 0*** row reduce the augmented matrix:

$$\begin{bmatrix} 0 & 1 & 4 & 0 \\ 1 & 2 & -1 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix} \qquad R_{12}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & -2 & 5 & 0 \end{bmatrix} \qquad (-5)R_1 + R_3$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 13 & 0 \end{bmatrix} \qquad (2)R_2 + R_3$$

At this point, it is clear that there are three basic variables and no free variables. So the equation ***Ax = 0*** has only the trivial solution, and the columns of ***A*** are linearly independent.

Sets of One or Two Vectors:

A set containing only one vector (say, ***v***) is linearly independent if and only if ***v*** is not the zero vector. This is because the vector equation

$x_1v = 0$ has only the trivial solution when $v \neq 0$. The zero vector is linearly dependent because $x_10 = 0$ has many nontrivial solutions.

Example 3: Determine if the following sets of vectors are linearly independent.

a. $v_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$

b. $v_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$

Solution:

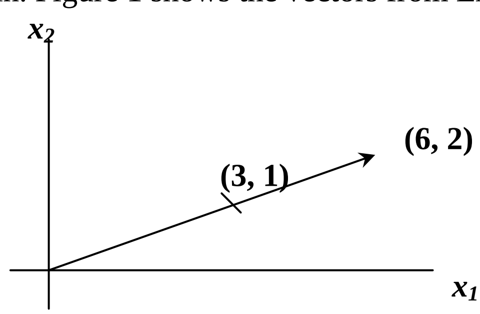
a) Notice that v_2 is a multiple of v_1, namely, $v_2 = 2v_1$.
Hence $-2v_1 + v_2 = 0$, which shows that $\{v_1, v_2\}$ is linearly dependent.

b) v_1 and v_2 are certainly not multiples of one another. Could they be linearly dependent?

Suppose c and d satisfy $cv_1 + dv_2 = 0$

If $c \neq 0$, then we can solve for v_1 in terms of v_2, namely, $v_1 = (-d/c)\, v_2$. This result is impossible because v_1 is not a multiple of v_2. So c must be zero. Similarly, d must also be zero.
Thus $\{v_1, v_2\}$ is a linearly independent set.

Note: A set of two vectors $\{v_1, v_2\}$ is linearly dependent if and only if one of the vectors is a multiple of the other.
In geometric terms, two vectors are linearly dependent if and only if they lie on the same line through the origin. Figure 1 shows the vectors from Example 3.



Linearly dependent

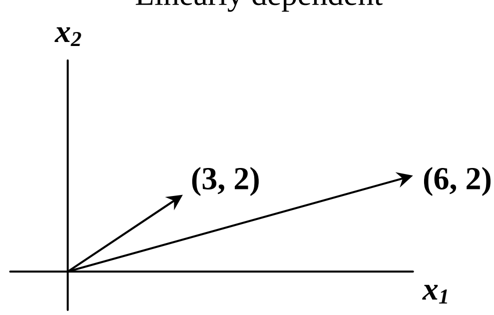


Figure 1 Linearly independent

Sets of Two or More Vectors;

Theorem (Characterization of Linearly dependent Sets):
An indexed set $s = \{v_1, v_2, \cdots, v_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in ***S*** is a linear combination of the others. In fact, if **S** is linearly dependent, and $v \neq 0$, then some v_j (with $j > 1$) is a linear combination of the preceding vectors, $v_1, \cdots, v_{j-1}$.

Proof:
If some v_j in ***S*** equals a linear combination of the other vectors, then v_j can be subtracted from both sides of the equation, producing a linear dependence relation with a nonzero weight (–1) on v_j.

For instance, if $v_1 = c_2v_2 + c_3v_3$, then $0 = (-1)v_1 + c_2v_2 + c_3v_3 + 0v_4 + \ldots + 0v_p$.
Thus ***S*** is linearly dependent.

Conversely, suppose ***S*** is linearly dependent. If v_1 is zero, then it is a (trivial) linear combination of the other vectors in ***S***.

If $v \neq 0$ and there exist weights $c_1, \ldots, c_p$, not all zero(because vectors are linearly dependent), such that

$$c_1v_1 + c_2v_2 + \ldots + c_pv_p = 0$$

Let j be the largest subscript for which $c_j \neq 0$. If $j = 1$, then $c_1v_1 = 0$, which is impossible because $v_j \neq 0$.

So $j > 1$, and $c_1v_1 + \cdots + c_jv_j + 0v_{j+1} + \cdots + 0v_p = 0$

$$c_jv_j = -c_1v_1 - c_2v_2 - \cdots - c_{j-1}v_{j-1}$$

$$v_j = \left(-\frac{c_1}{c_j}\right)v_1 + \left(-\frac{c_2}{c_j}\right)v_2 + \cdots + \left(\frac{c_{j-1}}{c_j}\right)v_{j-1}$$

Note: This theorem does not say that every vector in a linearly dependent set is a linear combination of the preceding vectors. A vector in a linearly dependent set may fail to be a linear combination of the other vectors.

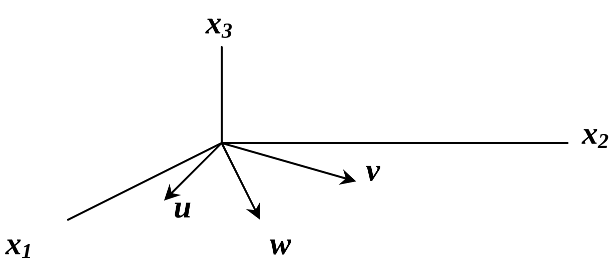
Example 4: Let $u = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Describe the set spanned by ***u*** and ***v***, and explain why a vector ***w*** is in Span **{*u, v*}** if and only if **{*u, v, w*}** is linearly dependent.

Solution:

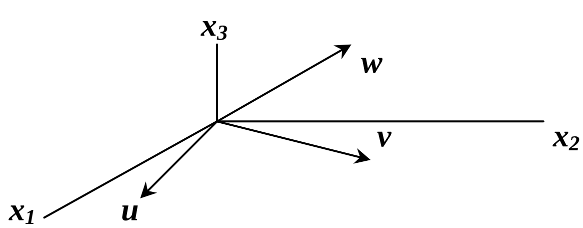
The vectors ***u*** and ***v*** are linearly independent because neither vector is a multiple of the other, nor so they span a plane in $\boldsymbol{R^3}$. In fact, Span **{*u, v*}** is the $\boldsymbol{x_1x_2}$**-plane** (with $\boldsymbol{x_3 = 0}$). If ***w*** is a linear combination of ***u*** and ***v***, then **{*u, v, w*}** is linearly dependent.

Conversely, suppose that **{*u, v, w*}** is linearly dependent.

Some vector in **{*u, v, w*}** is a linear combination of the preceding vectors (since $u \neq 0$). That vector must be ***w***, since ***v*** is not a multiple of ***u***. So ***w*** is in Span **{*u, v*}**



Linearly dependent ***w*** in Span **{*u, v*}**.



Linearly independent ***w*** not in **Span {*u, v*}**

Figure 2: Linear dependence in $\boldsymbol{R^3}$.

This example generalizes to any set **{*u, v, w*}** in $\boldsymbol{R^3}$ with ***u*** and ***v*** linearly independent. The set **{*u, v, w*}** will be linearly dependent if and only if ***w*** is in the plane spanned by ***u*** and ***v***.

Theorem:

If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{v_1, v_2, \ldots, v_p\}$ in R^n is linearly dependent if $p > n$.

Example 5: The vectors $\begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}4\\-1\end{bmatrix}, \begin{bmatrix}-2\\2\end{bmatrix}$ are linearly dependent, because there are three vectors in the set and there are only two entries in each vector.

Notice, however, that none of the vectors is a multiple of one of the other vectors. See Figure 4.

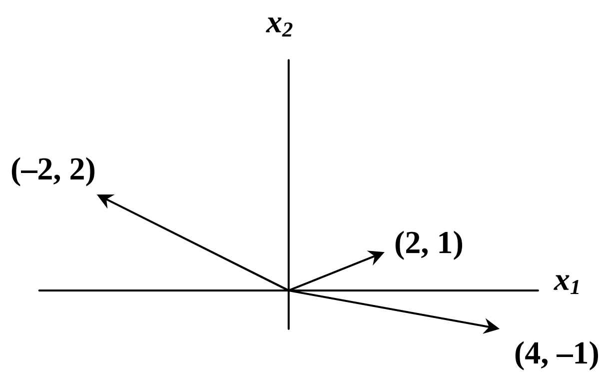


Figure 4 A linearly dependent set in $\boldsymbol{R}^2$.

Theorem:

If a set $S = \{v_1, v_2, \ldots, v_p\}$ in R^n contains the zero vector, then the set is linearly dependent**.**

Proof:

By renumbering the vectors, we may suppose that $\boldsymbol{v_1 = 0}$.

Then $(1)\boldsymbol{v_1} + \boldsymbol{0v_2} + \ldots + \boldsymbol{0v_p} = \boldsymbol{0}$ shows that $\boldsymbol{S}$ is linearly dependent(because in this relation coefficient of v_1 is non zero).

Example 6. Determine by inspection if the given set is linearly dependent.

a. $\begin{bmatrix}1\\7\\6\end{bmatrix}, \begin{bmatrix}2\\0\\9\end{bmatrix}, \begin{bmatrix}3\\1\\5\end{bmatrix}, \begin{bmatrix}4\\1\\8\end{bmatrix}$ **b.** $\begin{bmatrix}2\\3\\5\end{bmatrix}, \begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\8\end{bmatrix}$ **c.** $\begin{bmatrix}-2\\4\\6\\10\end{bmatrix}, \begin{bmatrix}3\\-6\\-9\\15\end{bmatrix}$

Solution:

a) The set contains four vectors that each has only three entries. So the set is linearly dependent by the Theorem above.
b) The same theorem does not apply here because the number of vectors does not exceed the number of entries in each vector. Since the zero vector is in the set, the set is linearly dependent by the next theorem.

c) As we compare corresponding entries of the two vectors, the second vector seems to be –3/2 times the first vector. This relation holds for the first three pairs of entries, but fails for the fourth pair. Thus neither of the vectors is a multiple of the other, and hence they are linearly independent.

Exercise:

1. Let $u=\begin{bmatrix}3\\2\\-4\end{bmatrix}, v=\begin{bmatrix}-6\\1\\7\end{bmatrix}, w=\begin{bmatrix}0\\-5\\2\end{bmatrix}, and \quad z=\begin{bmatrix}3\\7\\-5\end{bmatrix}.$

(i) Are the sets **{*u, v*}, {*u, w*}, {*u, z*}, {*v, w*}, {*v, z*},** and **{*w, z*}** each linearly independent? Why or why not?
(ii) Does the answer to Problem (i) imply that **{*u, v, w, z*}** is linearly independent?
(iii) To determine if **{*u, v, w, z*}** is linearly dependent, is it wise to check if, say, ***w*** is a linear combination of ***u, v*** and ***z***?
(iv) Is **{*u, v, w, z*}** linear dependent?

Decide if the vectors are linearly independent. Give a reason for each answer.

2. $\begin{bmatrix}3\\0\\0\end{bmatrix},\begin{bmatrix}-3\\2\\3\end{bmatrix},\begin{bmatrix}6\\4\\0\end{bmatrix}$

3. $\begin{bmatrix}1\\3\\-2\end{bmatrix},\begin{bmatrix}-3\\-5\\6\end{bmatrix},\begin{bmatrix}0\\5\\-6\end{bmatrix}$

Determine if the columns of the given matrix form a linearly dependent set.

4. $\begin{bmatrix}1&3&-2&0\\3&10&-7&1\\-5&-5&3&7\end{bmatrix}$

5. $\begin{bmatrix}3&4&3\\-1&-7&7\\1&3&-2\\0&2&-6\end{bmatrix}$

6. $\begin{bmatrix}1&1&0&4\\-1&0&3&-1\\0&-2&1&1\\1&0&-1&3\end{bmatrix}$

7. $\begin{bmatrix}1&-1&-3&0\\0&1&5&4\\-1&2&8&5\\3&-1&1&3\end{bmatrix}$

For what values of h is v_3 in span {v_1, v_2} and for what values of h is {v_1, v_2, v_3} linearly dependent?

8. $v_1=\begin{bmatrix}1\\3\\-2\end{bmatrix}, v_2=\begin{bmatrix}-2\\-6\\4\end{bmatrix}, v_3=\begin{bmatrix}1\\2\\h\end{bmatrix}$

9. $v_1=\begin{bmatrix}1\\3\\3\end{bmatrix}, v_2=\begin{bmatrix}3\\9\\-1\end{bmatrix}, v_3=\begin{bmatrix}-2\\-6\\h\end{bmatrix}$

Find the value(s) of h for which the vectors are linearly dependent.

10. $\begin{bmatrix} 1 \\ 3 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ h \end{bmatrix}$

11. $\begin{bmatrix} 1 \\ -5 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ 8 \\ 6 \end{bmatrix}, \begin{bmatrix} 4 \\ h \\ -8 \end{bmatrix}$

Determine by inspection whether the vectors are linearly independent. Give reasons for your answers.

12. $\begin{bmatrix} 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 6 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \end{bmatrix}$

13. $\begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} -6 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

14. $\begin{bmatrix} 6 \\ 2 \\ -8 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$

15. Given $A = \begin{bmatrix} 2 & 3 & 5 \\ -5 & 1 & -4 \\ -3 & -1 & -4 \\ 1 & 0 & 1 \end{bmatrix}$, observe that the third column is the sum of the first two columns. Find a nontrivial solution of $\boldsymbol{Ax = 0}$ without performing row operations.

Each statement in exercises 16-18 is either true(in all cases) or false(for at least one example). If false, construct a specific example to show that the statement is not always true. If true, give a justification.

16. If $v_1, \ldots, v_4$ are in R^4 and $v_3 = 2v_1 + v_2$, then $\{v_1, v_2, v_3, v_4\}$ is linearly dependent.

17. If v_1 and v_2 are in R^4 and v_1 is not a scalar multiple of v_2, then $\{v_1, v_2\}$ is linearly independent.

18. . If $v_1, \ldots, v_4$ are in R^4 and$\{v_1, v_2, v_3\}$ is linearly dependent, then $\{v_1, v_2, v_3, v_4\}$ is also linearly dependent.

19. Use as many columns of $A = \begin{bmatrix} 8 & -3 & 0 & -7 & 2 \\ -9 & 4 & 5 & 11 & -7 \\ 6 & -2 & 2 & -4 & 4 \\ 5 & -1 & 7 & 0 & 10 \end{bmatrix}$ as possible to construct a matrix B with the property that equation ***Bx = 0*** has only the trivial solution. Solve ***Bx = 0*** to verify your work.

Lecture 9

Linear Transformations

Outlines of the Lecture:

- Matrix Equation
- Transformation, Examples, Matrix as Transformations
- Linear Transformation, Examples, Some Properties

Matrix Equation:

An equation ***Ax = b*** is called a matrix equation in which a matrix ***A*** acts on a vector ***x*** by multiplication to produce a new vector called ***b.***

For instance, the equations

$$\underset{A}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}} \underset{x}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}} = \underset{b}{\begin{bmatrix} 5 \\ 8 \end{bmatrix}}$$

and

$$\underset{A}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}} \underset{u}{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}} = \underset{\underline{o}}{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

Solution of Matrix equation:

Solution of the ***Ax = b*** consists of those vectors ***x*** in ***the domain*** that are transformed into the vector ***b*** in ***range***.

Matrix equation ***Ax = b*** is an important example of transformation we would see later in the lecture.

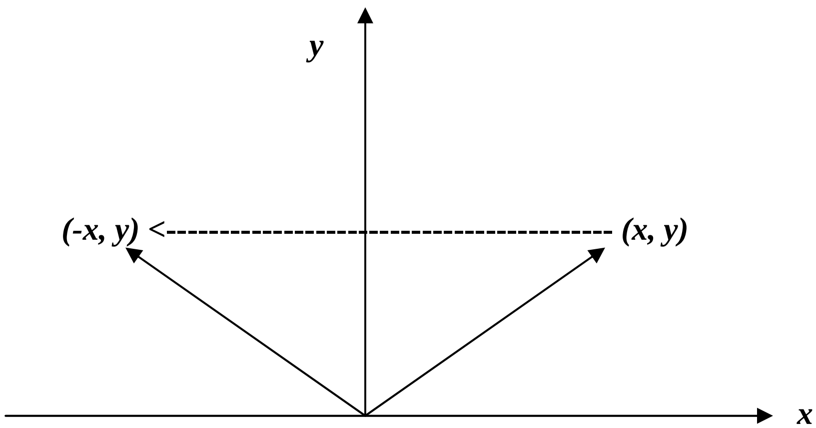
Transformation or Function or Mapping:

A transformation (or **function** or **mapping**) ***T*** from $\boldsymbol{R}^{n}$ to $\boldsymbol{R}^{m}$ is a rule that assigns to each vector ***x*** in $\boldsymbol{R}^{n}$ an image vector ***T(x)*** in $\boldsymbol{R}^{m}$.

$$T : R^n \rightarrow R^m$$

The set $\boldsymbol{R}^{n}$ is called the **domain** of ***T***, and $\boldsymbol{R}^{m}$ is called the **co-domain** of ***T***. For ***x*** in $\boldsymbol{R}^{n}$ the set of all images ***T(x)*** is called the **range** of ***T***.

Example 1: Consider a mapping $T: R^2 \rightarrow R^2$ defined by $T(x, y) = (-x, y)$. This transformation is a reflection about y-axis in xy plane.
Here $T(1,2) = (-1,2)$. T has transformed vector (1,2) into another vector (-1,2)



Example 2: Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $u = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $c = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$,

and define a transformation $T: R^2 \rightarrow R^3$ by ***T(x) = Ax***, so that

$$T(x) = Ax = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

a) Find ***T (u)***, the image of ***u*** under the transformation ***T***.
b) Find an ***x*** in $\boldsymbol{R}^{2}$ whose image under ***T*** is ***b***.
c) Is there more than one ***x*** whose image under ***T*** is ***b***?
d) Determine if ***c*** is in the range of the transformation ***T***.

Solution: (a)

$$T(u) = Au = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

Here $T(u) = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$

Here the matrix transformation has transformed $u = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ into another vector $\begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$

(b) We have to find an x such that **T (x) = b or Ax = b**

i. e $$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix} \qquad (1)$$

Now row reduced augmented matrix will be:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} -3R_1 + R_2,\ R_1 + R_3$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & -2 \end{bmatrix} \frac{1}{14}R_2,\ -4R_2 + R_3,$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} 3R_2 + R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence x_1 = 1.5, x_2 = - 0.5, and $x = \begin{bmatrix} 1.5 \\ -.5 \end{bmatrix}$.

The image of this ***x*** under ***T*** is the given vector ***b***.

(c) From (2) it is clear that equation (1) has a unique solution. So there is exactly one ***x*** whose image is ***b***.

(d) The vector ***c*** is in the range of ***T*** if ***c*** is the image of some ***x*** in $\boldsymbol{R^2}$, that is, if $c = T(x)$ for some ***x***. This is just another way of asking if the system ***Ax = c*** is consistent. To find the answer, we will row reduce the augmented matrix:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} -3R_1 + R_2,\ R_1 + R_3$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & 8 \end{bmatrix} \frac{1}{4}R_3, R_{23}$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 14 & -7 \end{bmatrix} -14R_2 + R_3$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}$$

$$x_1 - 3x_2 = 3$$
$$0x_1 + x_2 = 2$$
$$0x_1 + 0x_2 = -35 \Rightarrow 0 = 35 \; but \; 0 \neq 35$$

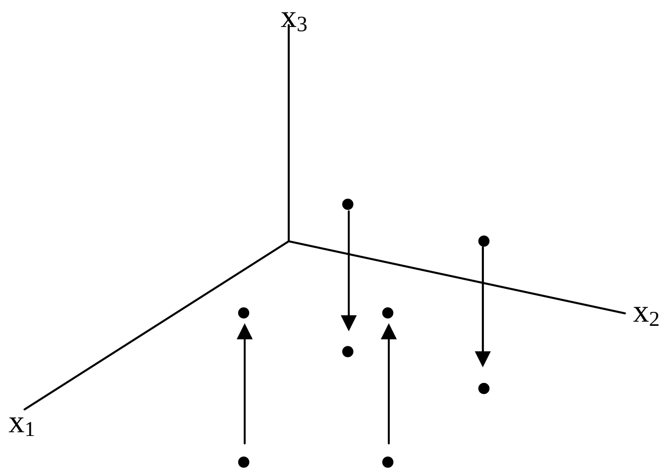
Hence the system is inconsistent. So **c** is not in the range of **T**.

So from above example we can view a transformation in the form of a matrix. We'll see that a transformation $T: R^n \to R^m$ can be transformed into a matrix of order $m \times n$ and every matrix of order $m \times n$ can be viewed as a linear transformation.

The next two matrix transformations can be viewed geometrically. They reinforce the dynamic view of a matrix as something that transforms vectors into other vectors.

<u>Example 3</u>: If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, then the transformation $x \to Ax$ projects points in $\boldsymbol{R^3}$

onto the $x_1x_2 - coordinate\ \ plane$ because $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$



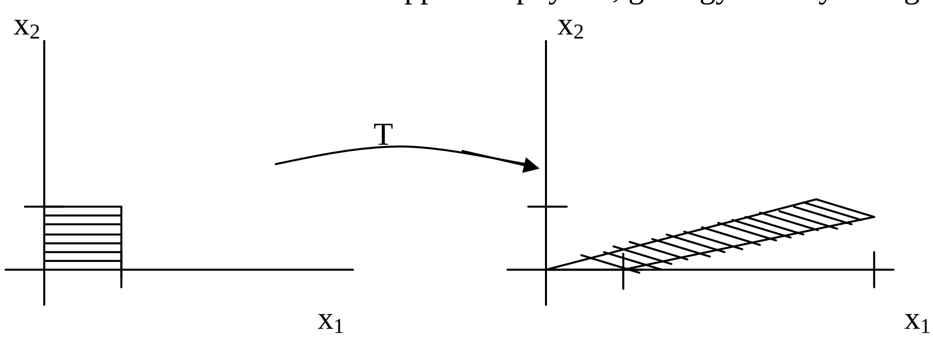
A projection transformation

Example 4: Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$., the transformation $T : R^2 \to R^2$ defined by ***T (x) = Ax*** is called a **shear** transformation.

The image of the point $u = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ is $T(u) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$,

and the image of $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ is $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$.

Here ***T*** deforms the square as if the top of the square were pushed to the right while the base is held fixed. Shear transformations appear in physics, geology and crystallography.



A shear transformation

Linear Transformations:

We know that if ***A*** is $m \times n$ matrix, then the transformation $x \to Ax$ has the properties $A(u+v) = Au + Av \ and \ A(cu) = cAu$ for all ***u, v*** in $\boldsymbol{R^n}$ and all scalars ***c***.
These properties for a transformation identify the most important class of transformations in linear algebra.

Definition: A transformation (or mapping) ***T*** is linear if:

1. ***T(u + v) = T(u) + T(v)*** for all ***u, v*** in the domain of ***T***;
2. ***T(cu) = cT(u)*** for all ***u*** and all scalars ***c***.

Example 5: Every matrix transformation is a linear transformation.

Example 6: Let $L: R^3 \to R^2$ be defined by $L(x, y, z) = (x, y)$.

we let $\boldsymbol{u} = (x_1, y_1, z_1)$ and $\boldsymbol{v} = (x_2, y_2, z_2)$.

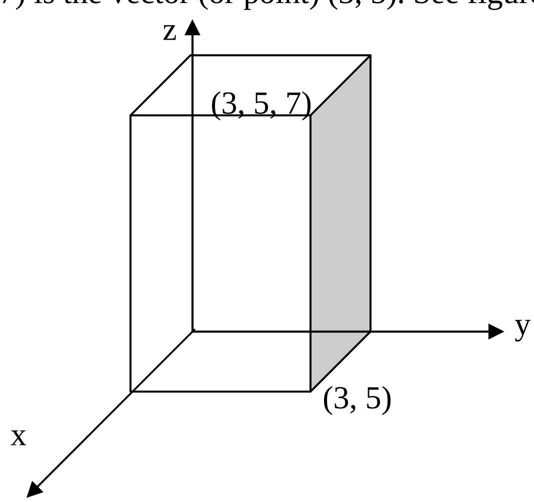
$$L(\boldsymbol{u}+\boldsymbol{v}) = L((x_1, y_1, z_1) + (x_2, y_2, z_2))$$
$$= L(x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

$$= (x_1 + x_2, y_1 + y_2)$$
$$= (x_1, y_1) + (x_2, y_2) = L(\boldsymbol{u}) + L(\boldsymbol{v})$$

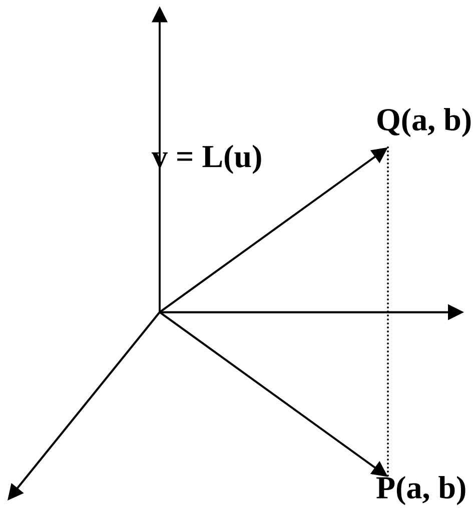
Also, if k is a real number, then

$$L(k\boldsymbol{u}) = L(kx_1, ky_1, kz_1) = (kx_1, ky_1) = kL(\boldsymbol{u})$$

Hence ***L*** is a linear transformation, which is called a projection. The image of the vector (or point) (3, 5, 7) is the vector (or point) (3, 5). See figure below.



Geometrically the image under ***L*** of a vector (a, b, c) in $\boldsymbol{R^3}$ is (a, b) in $\boldsymbol{R^2}$ can be found by drawing a line through the end point P(a, b, c) of ***u*** and perpendicular to $\boldsymbol{R^2}$, the *xy*-plane. The intersection Q(a, b)of this line with the *xy*-plane will give the image under ***L***. See figure below
.



Example 7: Let $L: R \to R$ be defined by $L(x) = x^2$

Let x and y in R and

$$L(x+y) = (x+y)^2 = x^2 + y^2 + 2xy \neq x^2 + y^2 = L(x) + L(y)$$
$$\Rightarrow L(x+y) \neq L(x) + L(y)$$

So we conclude that the function ***L*** is not a linear transformation.

Linear transformations preserve the operations of vector addition and scalar multiplication

Properties:

If ***T*** is a linear transformation, then

1. ***T(0) = 0***
2. ***T(cu +dv) = cT(u) + dT(v)***
3. $\boldsymbol{T(c_1 v_1 + \ldots + c_p v_p) = c_1 T(v_1) + \ldots + c_p T(v_p)}$

for all vectors ***u, v*** in the domain of ***T*** and all scalars *c*, *d*.

Proof:

1. By the definition of Linear Transformation we have ***T(cu) = cT(u)*** for all ***u*** and all scalars ***c***. Put $c = 0$ we'll get ***T(0u)= 0T(u)*** This implies ***T(0) = 0***

2. Just apply the definition of linear transformation. i. e

$$T(c\mathbf{u} + d\mathbf{v}) = T(c\mathbf{u}) + T(d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$$

Property (3) follows from (ii), because ***T(0) = T (0u) =0T(u) = 0***.
Property (4) requires both (i) and (ii):

OBSERVATION: Observe that if a transformation satisfies property 2 for all ***u, v*** and *c*, *d*, it must be linear (Take $c = d = 1$ for preservation of addition, and take $d = 0$)

3. Generalizing Property 2 we'll get 3

$$T(c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \ldots + c_pT(\mathbf{v}_p)$$

Applications in Engineering:

In engineering and physics, property 3 is referred to as a superposition principle.
Think of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ as signals that go into a system or process and $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)$ as the responses of that system to the signals. The system satisfies the superposition principle if whenever an input is expressed as a linear combination of such signals, the system's response is the same linear combination of the responses to the individual signals

Example 8: Given a scalar *r*, define $T : R \to R$ by

$$T(x) = x+1.$$

T is not a linear transformation (why!) because $T(0) \neq 0$ (by property 3)

Example 9: Given a scalar *r*, define $T : R^2 \to R^2$ by $T(\mathbf{x}) = r\mathbf{x}$.
T is called a **contraction** when $0 \leq r < 1$
and a **dilation** when $r \geq 1$.

Let r = 3 and show that ***T*** is a linear transformation.

Solution: Let ***u, v*** be in $\boldsymbol{R^2}$ and let *c, d* be scalars, then

$$
\begin{aligned}
T(c\mathbf{u} + d\mathbf{v}) &= 3(c\mathbf{u} + d\mathbf{v}) && \text{Definition of } \mathbf{T} \\
&= 3c\mathbf{u} + 3d\mathbf{v} && \\
&= c(3\mathbf{u}) + d(3\mathbf{v}) && \text{Vector arithmetic} \\
&= c\mathbf{T}(\mathbf{u}) + d\mathbf{T}(\mathbf{v}) &&
\end{aligned}
$$

Thus ***T*** is a linear transformation because it satisfies (4).

Example 10: Define a linear transformation $T : R^2 \to R^2$ by

$$T(x) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

Find the images under ***T*** of $u = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad v = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad and \quad u + v = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$.

Solution: $T(u) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \quad T(v) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$

$$T(u+v) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$

In above example ***T*** rotates ***u, v*** and ***u + v*** counterclockwise through 90^0.
In fact, ***T*** transforms the entire parallelogram determined by ***u*** and ***v*** into the one determined by ***T (u)*** and ***T (v)***

Example 11: Let L: $R^3 \to R^2$ be a linear transformation for which we know that
L (1, 0, 0) = (2, -1),
L (0, 1, 0) = (3, 1), and
L (0, 0, 1) = (-1, 2).
Then find L (-3, 4, 2).

Solution: Since (-3, 4, 2) = -3**i** + 4**j** + 2**k**,

$$
\begin{aligned}
L(-3,4,2) &= L(-3\boldsymbol{i} + 4\boldsymbol{j} + 2\boldsymbol{k}) = -3L(\boldsymbol{i}) + 4L(\boldsymbol{j}) + 2L(\boldsymbol{k}) \\
&= -3(2,-1) + 4(3,1) + 2(-1,2) = (4,11)
\end{aligned}
$$

Exercise

1. Suppose that $T: R^5 \to R^2$ and ***T(x) = Ax*** for some matrix ***A*** and each ***x*** in $\boldsymbol{R^5}$. How many rows and columns does ***A*** have?

2. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Give a geometric description of the transformation $x \to Ax$.

3. The line segment from **0** to a vector ***u*** is the set of points of the form ***tu***, where $0 \le t \le 1$. Show that a linear transformation ***T*** maps this segment into the segment between **0** and ***T(u)***.

4. Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, u = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}, and \begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix}$. Define T:$R^3 \to R^3$ by T (x) = Ax. Find T (u) and T (v).

In exercises 5 and 6, with T defined by T (x) = Ax, find an x whose image under T is b, and determine if x is unique.

5. $A = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 1 & -5 \\ -4 & 2 & 1 \end{bmatrix}, b = \begin{bmatrix} 0 \\ -5 \\ -6 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -4 \\ 3 & 2 & 1 \\ -2 & -1 & -2 \end{bmatrix}, b = \begin{bmatrix} 1 \\ -5 \\ -7 \\ 3 \end{bmatrix}$

Find all x in R^4 that are mapped into the zero vector by the transformation $x \to Ax$.

7. $A = \begin{bmatrix} 1 & 2 & -7 & 5 \\ 0 & 1 & -4 & 0 \\ 1 & 0 & 1 & 6 \\ 2 & -1 & 6 & 8 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 3 & 4 & -3 \\ 0 & 1 & 3 & -2 \\ 3 & 7 & 6 & -5 \end{bmatrix}$

9. Let $b = \begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix}$ and let A be the matrix in exercise 8. Is b in the range of the linear transformation $x \to Ax$?

10. Let $b = \begin{bmatrix} 9 \\ 5 \\ 0 \\ -9 \end{bmatrix}$ and let A be the matrix in exercise 7. Is b in the range of the linear transformation $x \rightarrow Ax$?

Let T (x) = Ax for x in R^2.

(a) On a rectangular coordinate system, plot the vectors u, v, T (u) and T (v).

(b) Give a geometric description of what T does to a vector x in R^2.

11. $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, u = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, and\ v = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ 12. $A = \begin{bmatrix} .5 & 0 \\ 0 & .5 \end{bmatrix}, u = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, and\ v = \begin{bmatrix} -5 \\ -2 \end{bmatrix}$

13. Let T: $R^2 \rightarrow R^2$ be a linear transformation that maps $u = \begin{bmatrix} 1 \\ 5 \end{bmatrix} into \begin{bmatrix} 2 \\ 0 \end{bmatrix} and\ maps\ v = \begin{bmatrix} 3 \\ 1 \end{bmatrix} into \begin{bmatrix} 1 \\ -4 \end{bmatrix}$. Use the fact that T is linear find the images under T of 2u, 3v, and 2u + 3v.

14. Let $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, y_1 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}, and\ y_2 = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$. Let T: $R^2 \rightarrow R^2$ be a linear transformation that maps e_1 into y_1 and maps e_2 into y_2. Find the images of $\begin{bmatrix} 7 \\ 6 \end{bmatrix} and \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

15. Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, v_1 = \begin{bmatrix} -7 \\ 4 \end{bmatrix} and\ v_2 = \begin{bmatrix} 3 \\ -8 \end{bmatrix}$. Let T: $R^2 \rightarrow R^2$ be a linear transformation that maps x into $x_1 v_1 + x_2 v_2$. Find a matrix A such that T (x) is Ax for each x.

Lecture 10

The Matrix of a Linear Transformation

Outlines of the Lecture:

- Matrix of a Linear Transformation.
- Examples, Geometry of Transformation, Reflection and Rotation
- Existence and Uniqueness of solution of ***T(x)=0***

In the last lecture we discussed that every linear transformation from $\boldsymbol{R^n}$ to $\boldsymbol{R^m}$ is actually a matrix transformation $x \to Ax$, where A is a matrix of order $m \times n$. First see an example

Example 1: The columns of $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Suppose ***T*** is a linear transformation from $\boldsymbol{R^2}$ into $\boldsymbol{R^3}$ such that

$$T(e_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad and \quad T(e_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

with no additional information, find a formula for the image of an arbitrary ***x*** in $\boldsymbol{R^2}$.

Solution: Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1 e_1 + x_2 e_2$

Since ***T*** is a linear transformation, $T(x) = x_1 T(e_1) + x_2 T(e_2)$

$$T(x) = x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix}$$

$$Hence \quad T(x) = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix}$$

Theorem: Let $T : R^n \to R^m$ be a linear transformation. Then there exists a unique matrix ***A*** such that $T(x) = Ax$ for all ***x*** in $\boldsymbol{R^n}$

In fact, ***A*** is the $m \times n$ matrix whose jth column is the vector ***T*** **(*e_j*)**, where e_j is the jth column of the identity matrix in $\boldsymbol{R^n}$.

$$A = \begin{bmatrix} T(e_1) & \dots & T(e_n) \end{bmatrix}$$

<u>Proof:</u> Write

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{bmatrix} + \dots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

$$= x_1e_1 + \dots + x_ne_n = \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} x$$

Since T is Linear, So

$$T(x) = T(x_1e_1 + \dots + x_ne_n) = x_1T(e_1) + \dots + x_nT(e_n)$$

$$= \begin{bmatrix} T(e_1) & \dots & T(e_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax \qquad (1)$$

The matrix ***A*** in (1) is called the **standard matrix for the linear transformation *T***. We know that every linear transformation from $\boldsymbol{R^n}$ to $\boldsymbol{R^m}$ is a matrix transformation and vice versa.

The term linear transformation focuses on a property of a mapping, while matrix transformation describes how such a mapping is implemented, as the next three examples illustrate.

<u>Example 2:</u> Find the standard matrix ***A*** for the dilation transformation ***T (x) = 3x***, $x \in R^2$.

<u>Solution:</u> Write

$$T(e_1) = 3e_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \quad and \quad T(e_2) = 3e_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

Example 3: Let $L: R^3 \to R^3$ is the linear operator defined by $L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ y-z \\ x+z \end{bmatrix}$.

Find the standard matrix representing ***L*** and verify ***L (x) = Ax***.

Solution:
The standard matrix ***A*** representing ***L*** is the 3 x 3 matrix whose columns are ***L*** (e_1), ***L*** (e_2), and ***L*** (e_3) respectively. Thus

$$L(e_1) = L\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1+0 \\ 0-0 \\ 1+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = col_1(A)$$

$$L(e_2) = L\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0+1 \\ 1-0 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = col_2(A)$$

$$L(e_3) = L\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0+0 \\ 0-1 \\ 0+1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = col_3(A)$$

Hence $$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

Now $$\boldsymbol{Ax} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y-z \\ x+z \end{bmatrix} = L(x)$$

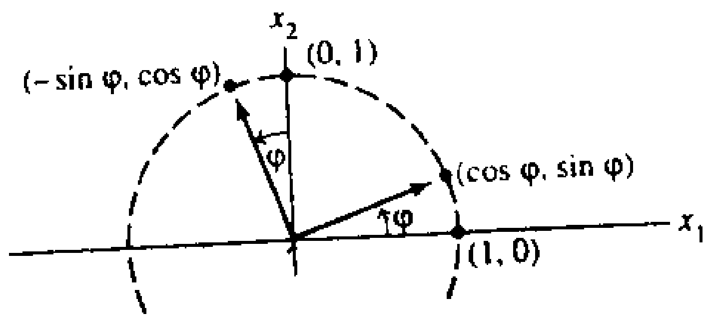
Hence verified.

Example 4: Let $T: R^2 \to R^2$ be the transformation that rotates each point in $\boldsymbol{R^2}$ through an angle φ, with counterclockwise rotation for a positive angle. We could show geometrically that such a transformation is linear. Find the standard matrix ***A*** of this transformation.

Solution $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ rotates into $\begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ rotates into $\begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}$.

See figure below.

By above theorem $A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$



A rotation transformation

Example 5: A reflection with respect to the ***x***-axis of a vector ***u*** in $\boldsymbol{R^2}$ is defined by the

linear operator $L(u) = L\left(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}\right) = \begin{bmatrix} a_1 \\ -a_2 \end{bmatrix}$.

Then $L(e_1) = L\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $L(e_2) = L\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

Hence the standard matrix representing ***L*** is $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Thus we have $L(u) = Au = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ -a_2 \end{bmatrix}$

To illustrate a reflection with respect to the x-axis in computer graphics, let the triangle ***T*** have vertices (-1, 4), (3, 1), and (2, 6).

To reflect ***T*** with respect to ***x***-axis, we let $u_1 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, u_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, u_3 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ and compute the images ***L*** (u_1), ***L*** (u_2), and ***L*** (u_3) by forming the products

$$Au_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \end{bmatrix},$$

$$Au_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix},$$

$$Au_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ -6 \end{bmatrix}.$$

Thus the image of ***T*** has vertices (-1, -4), (3, -1), and (2, -6).

Geometric Linear Transformations of R^2:

Examples 3-5 illustrate linear transformations that are described geometrically. In example 4 transformations is a rotation in the plane. It rotates each point in the plane through an angle φ . Example 5 is reflection in the plane.

Existence and Uniqueness of the solution of *T(x)=b:*

The concept of a linear transformation provides a new way to understand existence and uniqueness questions asked earlier. The following two definitions give the appropriate terminology for transformations.

Definition: A mapping $T : R^n \to R^m$ is said to be onto $\boldsymbol{R^m}$ if each ***b*** in $\boldsymbol{R^m}$ is the image of at least one ***x*** in $\boldsymbol{R^n}$.

OR

Equivalently, ***T*** is onto $\boldsymbol{R^m}$ if for each ***b*** in $\boldsymbol{R^m}$ there exists at least one solution of ***T (x) = b***. "Does ***T*** map $\boldsymbol{R^n}$ onto $\boldsymbol{R^m}$?" is an existence question.

The mapping ***T*** is not onto when there is some ***b*** in $\boldsymbol{R^m}$ such that the equation ***T (x) = b*** has no solution.

Definition: A mapping $T : R^n \to R^m$ is said to be one-to-one (or 1:1) if each ***b*** in $\boldsymbol{R^m}$ is the image of at most one ***x*** in $\boldsymbol{R^n}$.

OR

Equivalently, **T** is one-to-one if for each ***b*** in $\boldsymbol{R^m}$ the equation ***T (x) = b*** has either a unique solution or none at all, "Is ***T*** one-to-one?" is a uniqueness question.

The mapping ***T*** is not one-to-one when some ***b*** in $\boldsymbol{R^m}$ is the image of more than one vector in $\boldsymbol{R^n}$. If there is no such ***b***, then ***T*** is one-to-one.

Example 6: Let ***T*** be the linear transformation whose standard matrix is

$$A = \begin{bmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

Does ***T*** map $\boldsymbol{R^4}$ onto $\boldsymbol{R^3}$? Is ***T*** a one-to-one mapping?

Solution: Since ***A*** happens to be in echelon form, we can see at once that ***A*** has a pivot position in each row.

We know that for each ***b*** in $\boldsymbol{R^3}$, the equation ***Ax = b*** is consistent. In other words, the linear transformation ***T*** maps $\boldsymbol{R^4}$ (its domain) onto $\boldsymbol{R^3}$.

However, since the equation ***Ax = b*** has a free variable (because there are four variables and only three basic variables), each ***b*** is the image of more than one ***x***. That is, ***T*** is not one-to-one.

Theorem: Let $T : R^n \rightarrow R^m$ be a linear transformation. Then ***T*** is one-to-one if and only if the equation ***T (x) = 0*** has only the trivial solution.

Proof: Since ***T*** is linear, ***T (0) = 0*** if ***T*** is one-to-one, then the equation ***T (x) = 0*** has at most one solution and hence only the trivial solution. If ***T*** is not one-to-one, then there is a ***b*** that is the image of at least two different vectors in $\boldsymbol{R^n}$ (say, ***u*** and ***v***).
That is, ***T (u) = b*** and ***T (v) = b***.

But then, since ***T*** is linear $T(u-v) = T(u) = b - b = 0$

The vector ***u – v*** is not zero, since $u \neq v$. Hence the equation ***T (x) = 0*** has more than one solution. So either the two conditions in the theorem are both true or they are both false.

Theorem: Let $T : R^n \rightarrow R^m$ be a linear transformation and let ***A*** be the standard matrix for ***T***. Then

(a) ***T*** maps $\boldsymbol{R^n}$ onto $\boldsymbol{R^m}$ if and only if the columns of ***A*** span $\boldsymbol{R^m}$;
(b) ***T*** is one-to-one if and only if the columns of **A** are linearly independent.

Proof:

(a) The columns of ***A*** span $\boldsymbol{R^m}$ if and only if for each ***b*** the equation ***Ax = b*** is consistent – in other words, if and only if for every ***b***, the equation ***T(x) = b*** has at least one solution. This is true if and only if ***T*** maps $\boldsymbol{R^n}$ onto $\boldsymbol{R^m}$.

(b) The equations ***T (x) = 0*** and ***Ax = 0*** are the same except for notation. So ***T*** is one-to-one if and only if ***Ax = 0*** has only the trivial solution. This happens if and only if the columns of ***A*** are linearly independent.

We can also write column vectors in rows, using parentheses and commas. Also, when we apply a linear transformation ***T*** to a vector – say, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (x_1, x_2)$ we write $T(x_1, x_2)$ instead of the more formal $T\left((x_1, x_2)\right)$.

Example 7: Let $\boldsymbol{T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)}$.

Show that ***T*** is a one-to-one linear transformation.
Does ***T*** map $\boldsymbol{R^2}$ onto $\boldsymbol{R^3}$?

Solution: When ***x*** and ***T (x)*** are written as column vectors, it is easy to see that ***T*** is described by the equation

$$\begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} \quad (4)$$

so ***T*** is indeed a linear transformation, with its standard matrix ***A*** shown in (4). The columns of ***A*** are linearly independent because they are not multiples. Hence ***T*** is one-to-one. To decide if ***T*** is onto $\boldsymbol{R^3}$, we examine the span of the columns of ***A***. Since ***A*** is 3×2, the columns of ***A*** span $\boldsymbol{R^3}$ if and only if ***A*** has 3 pivot positions. This is impossible, since ***A*** has only 2 columns. So the columns of ***A*** do not span $\boldsymbol{R^3}$ and the associated linear transformation is not onto $\boldsymbol{R^3}$.

Exercises:

1. Let $T:R^2 \to R^2$ be transformation that first performs a horizontal shear that maps e_2 into $e_2 - .5e_1$ (but leaves e_1 unchanged) and then reflects the result in the $\boldsymbol{x_2}$ – axis. Assuming that ***T*** is linear, find its standard matrix.

Assume that T is a linear transformation. Find the standard matrix of T.

2. $T:R^2 \to R^3, T(1,0)=(4,-1,2)\, and\, T(0,1)=(-5,3,-6)$

3. $T:R^3 \to R^2, T(e_1)=(1,4), T(e_2)=(-2,9), and\, T(e_3)=(3,-8),$ where e_1, e_2, and e_3 are the columns of the identity matrix.

4. $T:R^2 \to R^2$ rotates points clockwise through π radians.

5. $T:R^2 \to R^2$ is a "vertical shear" transformation that maps e_1 into $e_1 + 2e_2$ but leaves the vector e_2 unchanged.

6. $T:R^2 \to R^2$ is a "horizontal shear" transformation that maps e_2 into $e_2 - 3e_1$ but leaves the vector e_1 unchanged.

7. $T:R^3 \to R^3$ projects each point (x_1, x_2, x_3) vertically onto the x_1x_2-plane (where x_3=0).

8. $T:R^2 \to R^2$ first performs a vertical shear mapping e_1 into $e_1 - 3e_2$(leaving e_2 unchanged) and then reflects the result in the x_2-axis.

9. $T:R^2 \to R^2$ first rotates points counterclockwise through π /4 radians and then reflects the result in the x_2-axis.

Show that T is a linear transformation by finding a matrix that implements the mapping. Note that x_1, x_2, … are not vectors but are entries in vectors.

10. $T(x_1,x_2,x_3,x_4)=(x_1+x_2, x_2+x_3, x_3+x_4, 0)$

11. $T(x_1,x_2,x_3)=(3x_2-x_3, x_1+4x_2+x_3)$

12. $T(x_1,x_2,x_3,x_4)=3x_1-4x_2+8x_4$

13. Let $T:R^2 \to R^2$ be a linear transformation such that $T(x_1,x_2)=(x_1+x_2, 4x_1+7x_2)$. Find ***x*** such that T (x) =(-2, -5).

13. Let $T:R^2 \to R^3$ be a linear transformation such that $T(x_1,x_2)=(x_1+2x_2, -x_1-3x_2, -3x_1-2x_2)$. Find ***x*** such that T (x) =(-4, 7, 0).

In exercises 14 and 15, let T be the linear transformation whose standard matrix is given.

14. Decide if T is one-to-one mapping. Justify your answer.

$$\begin{bmatrix} -5 & 10 & -5 & 4 \\ 8 & 3 & -4 & 7 \\ 4 & -9 & 5 & -3 \\ -3 & -2 & 5 & 4 \end{bmatrix}$$

15. Decide if T maps R^5 onto R^5. Justify your answer.

$$\begin{bmatrix} 4 & -7 & 3 & 7 & 5 \\ 6 & -8 & 5 & 12 & -8 \\ -7 & 10 & -8 & -9 & 14 \\ 3 & -5 & 4 & 2 & -6 \\ -5 & 6 & -6 & -7 & 3 \end{bmatrix}$$

Lecture 11

Matrix Operations

(i-j)th Element of a matrix

Let A be an $m \times n$ matrix, where m and n are number of rows and number of columns respectively, then a_{ij} represents the ***i***-th row and ***j***-th column entry of the matrix. For example a_{12} represents 1st row and 2nd column entry.

Similarly a_{32} represents 3rd row and 2nd column entry. The columns of ***A*** are vectors in $\boldsymbol{R^m}$ and are denoted by (boldface) $\boldsymbol{a_1, a_2, \cdots, a_n}$.

These columns are $A = [\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \ldots \quad \boldsymbol{a_n}]$

The number a_{ij} is the ***i****-th* entry (from the top) of ***j****-th* column vector a_j.

$$\begin{array}{c} \\ Row \quad i \\ \\ \end{array} \begin{array}{c} \text{Column} \\ j \\ \begin{bmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \ldots & a_{ij} & \ldots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \ldots & a_{mj} & \ldots & a_{mn} \end{bmatrix} = A \\ \uparrow \qquad\quad \uparrow \qquad\quad \uparrow \\ \boldsymbol{a_1} \qquad\quad \boldsymbol{a_j} \qquad\quad \boldsymbol{a_n} \end{array}$$

Figure 1 Matrix notation.

Definitions

A **diagonal matrix** is a square matrix whose non-diagonal entries are zero.

$$D = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix}$$

The diagonal entries in $A = \left[a_{ij} \right]$ are $a_{11}, a_{22}, a_{33}, \cdots$ and they form the **main diagonal** of ***A.***

For example $\begin{bmatrix} 5 & 0 \\ 0 & 7 \end{bmatrix}$ $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 11 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ are all diagonal matrices.

Null Matrix or Zero Matrix

An $m \times n$ matrix whose entries are all zero is a Null or **zero matrix** and is always written as ***O***. A null matrix may be of any order.

For example $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3 x 3 3 x 2 4 x 5

are all Zero Matrices

Equal Matrices

Two matrices are said to be **equal** if they have the same size (i.e., the same number of rows and columns) and same corresponding entries.

Example 1 Consider the matrices

$$A = \begin{bmatrix} 2 & 1 \\ 3 & x+1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 4 & 0 \end{bmatrix}$$

The matrices ***A*** and ***B*** are equal if and only if *x*+1 = 5 or *x* = 4. There is no value of *x* for which ***A = C***, since ***A*** and ***C*** have different sizes.

If ***A*** and B are $m \times n$ matrices, then the **sum, *A + B***, is the $m \times n$ matrix whose columns are the sums of the corresponding columns in ***A*** and ***B***. Each entry in ***A + B*** is the sum of the corresponding entries in ***A*** and ***B***. The sum ***A + B*** is defined only when ***A*** and ***B*** are of the same size.

If *r* is a scalar and ***A*** is a matrix, then the **scalar multiple** ***rA*** is the matrix whose columns are *r* times the corresponding columns in ***A***.

Note: Negative of a matrix *A* is defined as – *A* to mean (–1)*A* and the difference of *A* and *B* is written as *A*–*B*, which means *A* + (–1) *B*.

Example 2 Let $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$

Then $A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$

But ***A + C*** is not defined because ***A*** and ***C*** have different sizes.

$$2B = 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}$$

$$A - 2B = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix}$$

Theorem 1: Let ***A, B,*** and ***C*** are matrices of the same size, and let *r* and *s* are scalars.

a. ***A + B = B + A***

b. ***(A + B) + C = A + (B + C)***

c. ***A + 0 = A***

d. ***r (A + B)= r A + r B***

e. ***(r + s) A = r A + s A***

f. ***r (sA) = (rs) A***

Each equality in Theorem 1 can be verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal. Size is no problem because ***A, B,*** and ***C*** are equal in size. The equality of columns follows immediately from analogous properties of vectors.

For instance, if the *jth* columns of ***A, B,*** and ***C*** are a_j, b_j and c_j, respectively, then the *jth* columns of ***(A + B) + C*** and ***A + (B + C)*** are

$$(a_j + b_j) + c_j \qquad and \qquad a_j + (b_j + c_j)$$

respectively. Since these two vector sums are equal for each ***j***, property (b) is verified.

Because of the associative property of addition, we can simply write ***A + B + C*** for the sum, which can be computed either as ***(A + B) + C*** or ***A + (B + C)***. The same applies to sums of four or more matrices.

Matrix Multiplication:

Multiplying an m×n matrix with an n×p matrix results in an m×p matrix. If many matrices are multiplied together, and their dimensions are written in a list in order, e.g. m×n, n×p, p×q, q×r, the size of the result is given by the first and the last numbers (m×r).

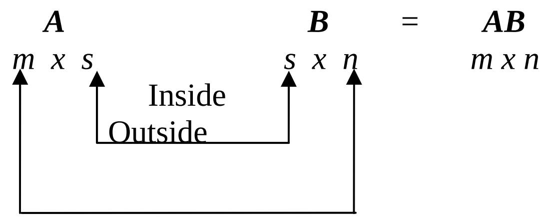
Matrix Multiplication It is important to keep in mind that this definition requires the number of columns of the first factor A to be the same as the number of rows of the second factor B. When this condition is satisfied, the sizes of A and B are said to conform for the product AB. If the sizes of A and B do not conform for the product AB, then this product is undefined.

Definition: If ***A*** is an $m \times n$ matrix, and if ***B*** is an $n \times p$ matrix with columns $b_1, \cdots, b_p$, then the product ***AB*** is the $m \times p$ matrix whose columns are $Ab_1, \cdots, Ab_p$.

That is $$AB = A\begin{bmatrix} b_1 & b_2 & \dots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_p \end{bmatrix}$$

This definition makes equation (1) true for all ***x*** in $\boldsymbol{R^p}$. Equation (1) proves that the composite mapping ***(AB)*** is a linear transformation and that its standard matrix is ***AB***. Multiplication of matrices corresponds to composition of linear transformations.

A convenient way to determine whether ***A*** and ***B*** conform for the product ***AB*** and, if so, to find the size of the product is to write the sizes of the factors side by side as in Figure below (the size of the first factor on the left and the size of the second factor on the right).



If the inside numbers are the same, then the product ***AB*** is defined and the outside numbers then give the size of the product.

<u>Example 3</u>: Compute ***AB***, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} and \quad B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$

Solution: Here $\boldsymbol{B = [b_1 \quad b_2 \quad b_3]}$, therefore

$$Ab_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad Ab_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad Ab_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Then

$$AB = A[b_1 \quad b_2 \quad b_3] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$\boldsymbol{Ab_1} \quad \boldsymbol{Ab_2} \quad \boldsymbol{Ab_3}$

Note from the definition of ***AB*** that its first column, $\boldsymbol{Ab_1}$, is a linear combination of the columns of ***A***, using the entries in $\boldsymbol{b_1}$ as weights. The same holds true for each column of ***AB***. Each column of ***AB*** is a linear combination of the columns of ***A*** using weights from the corresponding column of ***B***.

<u>Example 4</u>: Find the product ***AB*** for

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$$

Solution It follows from definition that the product ***AB*** is formed in a column-by-column manner by multiplying the successive columns of ***B*** by ***A***. The computations are

$$\begin{matrix} c_1 & c_2 & c_3 \end{matrix}$$

$$\left[\begin{array}{c:c:c} 1 & 2 & 4 \\ 2 & 6 & 0 \end{array}\right] \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} = 4c_1 + 0c_2 + 2c_3 = (4)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (0)\begin{bmatrix} 2 \\ 6 \end{bmatrix} + (2)\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 12 \\ 8 \end{bmatrix}$$

Similarly,

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix} = (1)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (-1)\begin{bmatrix} 2 \\ 6 \end{bmatrix} + (7)\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 27 \\ -4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix} = (4)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (3)\begin{bmatrix} 2 \\ 6 \end{bmatrix} + (5)\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 30 \\ 26 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = (3)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (1)\begin{bmatrix} 2 \\ 6 \end{bmatrix} + (2)\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 13 \\ 12 \end{bmatrix}$$

Thus,

$$AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

<u>Example 5</u>: (An Undefined Product) Find the product ***BA*** for the matrices

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad \textit{and} \quad B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$$

Solution The number of columns of B is not equal to number of rows of A so BA multiplication is not possible.

The matrix ***B*** has size 3×4 and the matrix ***A*** has size 2×3. The "inside" numbers are not the same, so the product ***BA*** is undefined.

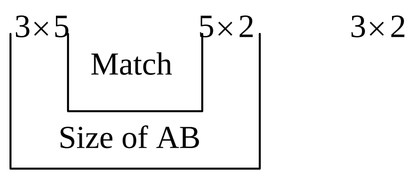
Obviously, the number of columns of ***A*** must match the number of row in ***B*** in order for a linear combination such as $\boldsymbol{Ab_1}$ to be defined. Also, the definition of ***AB*** shows that ***AB*** has the same number of rows as ***A*** and the same number of columns as ***B***.

<u>Example 6</u>: If ***A*** is a 3×5 matrix and ***B*** is a 5×2 matrix, what are the sizes of ***AB*** and ***BA***, if they are defined?

Solution: The product of matrices A and B of orders 3×5 and 5×2 will result in 3×2 matrix AB.
But for BA we have 5×2 and 3×5, here number of columns in1st matrix are 2 which is not equal to number of rows in 2nd matrix. So BA is not possible.

Since ***A*** has 5 columns and ***B*** has 5 rows, the product ***AB*** is defined and is a 3×2 matrix:

$$\underset{A}{\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}} \underset{B}{\begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix}} = \underset{AB}{\begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix}}$$

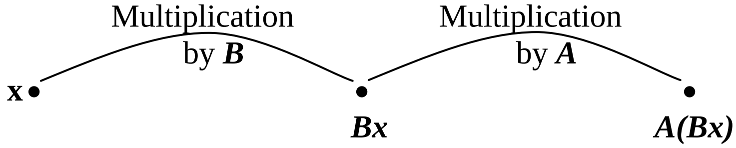


The product ***BA*** is not defined because the 2 columns of ***B*** do not match the 3 rows of ***A***.

The definition of ***AB*** is important for theoretical work and applications, but the following rule provides a more efficient method for calculating the individual entries in ***AB*** when working small problems by hand.

<u>Row-Column Rule for Computing *AB*</u>

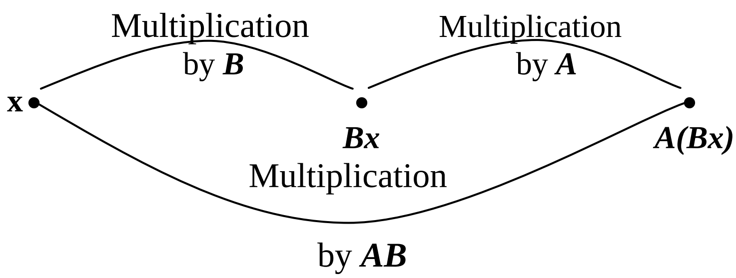
Explanation
If a matrix ***B*** is multiplied with a vector ***x***, it transforms ***x*** into a vector ***Bx***. If this vector is then multiplied in turn by a matrix ***A***, the resulting vector is ***A (Bx)***.



Multiplication by ***B*** and then ***A***

Thus ***A(Bx)*** is produced from ***x*** by a composition of mappings. Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by ***AB***, so that

$$\boldsymbol{A\ (Bx) = (AB)\ x}\text{---}(1)$$



Multiplication by ***AB***

If ***A*** is $m \times n$, ***B*** is $n \times p$, and ***x*** is in $\boldsymbol{R}^p$, denote the columns of ***B*** by $b_1, \cdots, b_p$ and the entries in ***x*** by $x_1, \cdots, x_p$, then $Bx = x_1 b_1 + x_2 b_2 + \cdots + x_p b_p$

By the linearity of multiplication by ***A***,

$$\begin{aligned} A(Bx) &= A(x_1 b_1) + A(x_2 b_2) + \cdots + A(x_p b_p) \\ &= x_1 A b_1 + x_2 A b_2 + \cdots + x_p A b_p \end{aligned}$$

The vector ***A (Bx)*** is a linear combination of the vectors $Ab_1, \cdots, Ab_p$, using the entries in ***x*** as weights. If we rewrite these vectors as the columns of a matrix, we have

$$A(Bx) = \begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_p \end{bmatrix} x$$

Thus multiplication by $\begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_p \end{bmatrix}$ transforms ***x*** into ***A(Bx)***.
We have found the matrix we sought!

<u>Row-Column Rule for Computing *AB*</u>

If the product ***AB*** is defined, then the entry in row ***i*** and column ***j*** of ***AB*** is the sum of the products of corresponding entries from row ***i*** of ***A*** and column ***j*** of ***B***. If $(AB)_{ij}$ denotes the *(i, j)* – entry in ***AB***, and if ***A*** is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots + a_{in} b_{nj}$$

To verify this rule, let $B = \begin{bmatrix} b_1 & \ldots & b_p \end{bmatrix}$. Column ***j*** of ***AB*** is Ab_j, and we can compute Ab_j. The *ith* entry in Ab_j is the sum of the products of corresponding entries from row ***i*** of ***A*** and the vector b_j, which is precisely the computation described in the rule for computing the *(i, j)* – entry of ***AB***.

Finding Specific Entries in a Matrix Product Sometimes we will be interested in finding a specific entry in a matrix product without going through the work of computing the entire column that contains the entry.

Example 7: Use the row-column rule to compute two of the entries in ***AB*** for the matrices in Example 3.

Solution: To find the entry in row 1 and column 3 of ***AB***, consider row 1 of ***A*** and column 3 of ***B***. Multiply corresponding entries and add the results, as shown below:

$$AB = \begin{matrix}\rightarrow\\ \\ \end{matrix}\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6)+3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

(the arrow $\downarrow$ points to column 3 of B)

For the entry in row 2 and column 2 of ***AB***, use row 2 of ***A*** and column 2 of ***B***:

$$\begin{matrix} \\ \rightarrow \end{matrix}\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3)+-5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$

(the arrow $\downarrow$ points to column 2 of B)

Example 8 Use the dot product rule to compute the individual entries in the product of ***AB*** where $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$.

Solution Since ***A*** has size 2×3 and ***B*** has size 3×4, the product ***AB*** is a 2×4 matrix of the form

$$AB = \begin{bmatrix} r_1(A)\times c_1(B) & r_1(A)\times c_2(B) & r_1(A)\times c_3(B) & r_1(A)\times c_4(B) \\ r_2(A)\times c_1(B) & r_2(A)\times c_2(B) & r_2(A)\times c_3(B) & r_2(A)\times c_4(B) \end{bmatrix}$$

where $r_1(A)$ and $r_2(A)$ are the row vectors of ***A*** and $c_1(B), c_2(B), c_3(B)$ *and* $c_4(B)$ are the column vectors of ***B***. For example, the entry in row 2 and column 3 of ***AB*** can be computed as

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}\begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \boxed{26} & \square \end{bmatrix}$$

$$(2\times 4)+(6\times 3)+(0\times 5) = 26$$

and the entry in row 1 and column 4 of ***AB*** can be computed as

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \boxed{13} \\ \square & \square & \square & \square \end{bmatrix}$$

$$(1\times 3)+(2\times 1)+(4\times 2)=13$$

Here is the complete set of computations:

$$(AB)_{11} = (1\times 4)+(2\times 0)+(4\times 2)=12$$
$$(AB)_{12} = (1\times 1)+(2\times -1)+(4\times 7) = 27$$
$$(AB)_{13} = (1\times 4)+(2\times 3)+(4\times 5) = 30$$
$$(AB)_{14} = (1\times 3)+(2\times 1)+(4\times 2)=13$$
$$(AB)_{21} = (2\times 4)+(6\times 0)+(0\times 2) = 8$$
$$(AB)_{22} = (2\times 1)+(6\times -1)+(0\times 7) = -4$$
$$(AB)_{23} = (2\times 4)+(6\times 3)+(0\times 5) = 26$$
$$(AB)_{24} = (2\times 3)+(6\times 1)+(0\times 2)=12$$

Finding Specific Rows and Columns of a Matrix Product

The specific column of ***AB*** is given by the formula

$$AB = A\begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_n \end{bmatrix}$$

Similarly, the specific row of ***AB*** is given by the formula $AB = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} B = \begin{bmatrix} a_1 B \\ a_2 B \\ \vdots \\ a_m B \end{bmatrix}$

<u>Example 9</u> Find the entries in the second row of ***AB***, where

$$A = \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

Solution: By the row-column rule, the entries of the second row of ***AB*** come from row 2 of ***A*** (and the columns of ***B***):

$$\begin{matrix} & & \downarrow & \downarrow \end{matrix}$$
$$\rightarrow\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}\begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square \\ -4+21-12 & 6+3-8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

<u>Example 10</u> (Finding a Specific Row and Column of AB)

Let $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$

Find the second column and the first row of ***AB***.

Solution $c_2(AB) = Ac_2(B) = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix} = \begin{bmatrix} 27 \\ -4 \end{bmatrix}$

$$r_1(AB) = r_1(A)B = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix}\begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \end{bmatrix}$$

<u>Properties of Matrix Multiplication</u>

These are standard properties of matrix multiplication. Remember that I_{m} represents the $m \times m$ identity matrix and $I_{\mathrm{m}}x = x$ for all ***x*** *belong to* $\boldsymbol{R^m}$.

<u>Theorem 2</u> Let ***A*** be $m \times n$, and let ***B*** and ***C*** have sizes for which the indicated sums and products are defined.

- **a.** ***A (BC) = (AB) C*** (associative law of multiplication)
- **b.** ***A (B + C) = AB + AC*** (left distributive law)
- **c.** ***(B + C)A = BA + CA*** (right distributive law)
- ***d.*** *r* ***(AB) = (r A)B = A(r B)*** (for any scalar *r*)
- **e.** $I_{\mathrm{m}} A = A = A\,\mathrm{I}_m$ (identity for matrix multiplication)

Proof. Properties (b) to (e) are considered exercises for you. We start property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known (or easy to check) that the composition of functions is associative.

Here is another proof of (a) that rests on the "column definition" of the product of two matrices. Let $C = \begin{bmatrix} c_1 & \ldots & c_p \end{bmatrix}$

By definition of matrix multiplication $BC = \begin{bmatrix} Bc_1 & \ldots & Bc_p \end{bmatrix}$

$$A(BC) = \begin{bmatrix} A(Bc_1) & \ldots & A(Bc_p) \end{bmatrix}$$

From above, we know that ***A(Bx) = (AB)x*** for all ***x***, so

$$A(BC) = \left[(AB)c_1 \quad \ldots \quad (AB)c_p \right] = (AB)C$$

The associative and distributive laws say essentially that pairs of parentheses in matrix expressions can be inserted and deleted in the same way as in the algebra of real numbers. In particular, we can write ***ABC*** for the product, which can be computed as ***A(BC)*** or as ***(AB)C***. Similarly, a product ***ABCD*** of four matrices can be computed as ***A(BCD)*** or ***(ABC)D*** or ***A(BC)D***, and so on. It does not matter how we group the matrices when computing the product, so long as the left-to-right order of the matrices is preserved.

The left-to-right order in products is critical because, in general, ***AB*** and ***BA*** are not the same. This is not surprising, because the columns of ***AB*** are linear combinations of the columns of ***A***, whereas the columns of ***BA*** are constructed from the columns of ***B***.

If ***AB = BA***, we say that ***A*** and ***B*** **commute** with one another.

<u>Example 11</u> Let $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} and \quad B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$

Show that these matrices don not commute, i.e. $AB \neq BA$.

Solution:

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 10+4 & 0+3 \\ 6-8 & 0-6 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10+0 & 2-0 \\ 20+9 & 4-6 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$

For emphasis, we include the remark about commutativity with the following list of important differences between matrix algebra and ordinary algebra of real numbers.

WARNINGS

1. In general, $AB \neq BA$.
2. The cancellation laws do not hold for matrix multiplication. That is, if $AB = AC$, then it is not true in general that $B = C$.
3. If a product ***AB*** is the zero matrix, you cannot conclude in general that either ***A = 0*** *or* ***B = 0***.

<u>Powers of a Matrix</u>: If ***A*** is an $n \times n$ matrix and if *k* is a positive integer, $\boldsymbol{A^k}$ denotes the product of *k* copies of ***A***, $A^k = \underbrace{A \ldots A}_{k}$ Also, we interpret $\boldsymbol{A^0}$ as ***I***.

<u>Transpose of a Matrix</u>: Given an $m \times n$ matrix ***A***, the **transpose** of ***A*** is the $n \times m$ matrix, denoted by $\boldsymbol{A^t}$, whose columns are formed from the corresponding rows of ***A***.
OR, if ***A*** is an *m x n* matrix, then transpose of ***A*** is denoted by $\boldsymbol{A^t}$, is defined to be the *nxm* matrix that is obtained by making the rows of ***A*** into columns; that is, the first column of $\boldsymbol{A^t}$ is the first row of ***A***, the second column of $\boldsymbol{A^t}$ is the second row of ***A***, and so forth.

<u>Example 12</u> (Transpose of a Matrix)

The following is an example of a matrix and its transpose.

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 5 & 6 \end{bmatrix}$$

$$A^t = \begin{bmatrix} 2 & 1 & 5 \\ 3 & 4 & 6 \end{bmatrix}$$

Example 13 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$

Then $A^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^t = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^t = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$

Theorem 3: Let ***A*** and ***B*** denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^t)^t = A$

b. $(A+B)^t = A^t + B^t$

c. For any scalar $r, (rA)^t = rA^t$

d. $(AB)^t = B^t A^t$

The generalization of (d) to products of more than two factors can be stated in words as follows.

> "The transpose of a product of matrices equals the product of their transposes in the reverse order."

Lecture 12

The Inverse of a Matrix

In this lecture and the coming next, we consider only square matrices and we investigate the matrix analogue of the reciprocal or multiplicative inverse of a nonzero real number.

Inverse of a square Matrix

If ***A*** is an $n \times n$ matrix, A matrix C of order $n \times n$ is called multiplicative inverse of A if

$AC = CA = I$ where I is the $n \times n$ identity matrix.

Invertible Matrix

If the inverse of a square matrix exist. It is called an invertible matrix.

In this case, we say that ***A*** is **invertible** and we call ***C*** an **inverse** of ***A***.

Note: If ***B*** is another inverse of ***A***, then we would have

$$B = BI = B(AC) = (BA)C = IC = C.$$

Thus when ***A*** is **invertible**, its inverse is unique.

The inverse of ***A*** is denoted by $\boldsymbol{A^{-1}}$, so that

$$AA^{-1} = I \qquad and \qquad A^{-1}A = I$$

Note: A matrix that is not invertible is sometime called a **singular** matrix, and an invertible matrix is called a **non-singular** matrix.

Example 1: If $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} and \quad C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, then

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -14+15 & -10+10 \\ 21-21 & 15-14 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad and$$

$$CA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} -14+15 & -35+35 \\ 6-6 & 15-14 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus **C =** $\boldsymbol{A^{-1}}$.

Theorem Let $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

If $ad-bc\neq 0$, then ***A*** is invertible or non singular and $A^{-1}=\dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

If $ad-bc=0$, then ***A*** is not invertible or singular.

The quantity $ad-bc$ is called the **determinant of *A***, and we write

$$\det A = ad-bc$$

This implies that a 2×2 matrix ***A*** is invertible if and only if $\det A\neq 0$.

Example 2 Find the inverse of $A=\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$.

Solution We have det ***A*** = 3(6) – 4(5) = $-2\neq 0$.

Hence ***A*** is invertible $A^{-1}=\dfrac{1}{-2}\begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix}=\begin{bmatrix} 6/(-2) & -4/(-2) \\ -5/(-2) & 3/(-2) \end{bmatrix}=\begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}$

The next theorem provides three useful facts about invertible matrices.

Theorem

a. If ***A*** is an invertible matrix, then $\boldsymbol{A^{-1}}$ is invertible and $(A^{-1})^{-1}=A$

b. If ***A*** and ***B*** are $n\times n$ invertible matrices, then so is ***AB***, and the inverse of ***AB*** is the product of the inverses of ***A*** and ***B*** in the reverse order. That is $(AB)^{-1}=B^{-1}A^{-1}$

c. If ***A*** is an invertible matrix, then so is $\boldsymbol{A^T}$, and the inverse of $\boldsymbol{A^T}$ is the transpose of $\boldsymbol{A^{-1}}$. That is $(A^T)^{-1}=(A^{-1})^T$

Proof:

(a) We must find a matrix ***C*** such that $A^{-1}C=I \quad and \quad CA^{-1}=I$

However, we already know that these equations are satisfied with ***A*** in place of ***C***. Hence $\boldsymbol{A^{-1}}$ is invertible and ***A*** is its inverse.

(b) We use the associative law for multiplication:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1}$$
$$= AIA^{-1}$$
$$= AA^{-1}$$
$$= I$$

A similar calculation shows that $(B^{-1}A^{-1})(AB) = I$.

Hence ***AB*** is invertible, and its inverse is $B^{-1}A^{-1}$ i.e $(AB)^{-1} = B^{-1}A^{-1}$

Generalization

Similarly we can prove the same results for more than two matrices i.e

$$\left((A_1)(A_2)(A_3)...(A_n)\right)^{-1} = A_n^{-1}A_{n-1}^{-1}... \; A_3^{-1}A_2^{-1}A_1^{-1}$$

The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.

Example 3: (Inverse of a Transpose). Consider a general 2×2 invertible matrix and its transpose:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad and \quad A^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Since A is invertible, its determinant *(ad – bc)* is nonzero. But the determinant of $\boldsymbol{A^t}$ is also *(ad – bc)*, so $\boldsymbol{A^t}$ is also invertible. It follows that

$$(A^t)^{-1} = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{c}{ad-bc} \\ -\frac{b}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \text{---------------(1)}$$

Now $$A^{-1} = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$$

Therefore, $$(A^{-1})^t = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{c}{ad-bc} \\ -\frac{b}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \text{-----------------(2)}$$

From (1) and (2), we have

$$\boldsymbol{(A^t)^{-1} = (A^{-1})^t} .$$

Example 4: (The Inverse of a Product). Consider the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

Here

$$AB = \begin{bmatrix} 7 & 6 \\ 9 & 8 \end{bmatrix},$$

$$(AB)^{-1} = \frac{1}{|AB|} Adj(AB) = \frac{1}{-2}\begin{bmatrix} 8 & -6 \\ -9 & 7 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|} Adj(A) = \frac{1}{1}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix},$$

$$B^{-1} = \frac{1}{|B|} Adj(B) = \frac{1}{2}\begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix},$$

$$B^{-1}A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

Thus, $(AB)^{-1} = B^{-1}A^{-1}$

Theorem: If A is invertible and n is a nonnegative integer, then:
(a) $\boldsymbol{A^n}$ is invertible and $\boldsymbol{(A^n)^{-1} = A^{-n} = (A^{-1})^n}$
(b) $k\boldsymbol{A}$ is invertible for any nonzero scalar k, and $(k\boldsymbol{A})^{-1} = k^{-1}\boldsymbol{A}^{-1}$.

Example 5 (Related to above theorem)

(a) Let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad then\ A^{-1} = \frac{1}{|A|} Adj(A) = \frac{1}{1}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$$

Now

$$A^{-3} = \left(A^{-1}\right)^3 = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}$$

Also,

$$A^3 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 30 \\ 15 & 41 \end{bmatrix}$$

$$(A^3)^{-1} = \frac{1}{(11)(41)-(30)(15)}\begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} = (A^{-1})^3$$

(b)

Take $A=\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$ and k=3

$kA=3A=\begin{bmatrix} 3 & 6 \\ 9 & 3 \end{bmatrix}$, $\quad (kA)^{-1} = (3A)^{-1} = \frac{1}{9-54}\begin{bmatrix} 3 & -6 \\ -9 & 3 \end{bmatrix} = \begin{bmatrix} -1/15 & 2/15 \\ 1/5 & -1/15 \end{bmatrix}$ ------(1)

$$A^{-1} = -\frac{1}{5}\begin{bmatrix} 1 & -2 \\ -3 & 1 \end{bmatrix}$$

A= So $k^{-1}A^{-1}=3^{-1}A^{-1}=\frac{1}{3}.\frac{-1}{5}\begin{bmatrix} 1 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -1/15 & 2/15 \\ 1/5 & -1/15 \end{bmatrix}$ -----------------------(2)

From (1) and (2), we have
$(3A)^{-1} = 3^{-1}A^{-1}$

There is an important connection between invertible matrices and row operations that leads to a method for computing inverses. As we shall see, an invertible matrix ***A*** is row equivalent to an identity matrix, and we can find $\boldsymbol{A^{-1}}$ by watching the row reduction of ***A*** to ***I***.

Elementary Matrices
As we have studied that there are three types of elementary row operations that can be performed on a matrix:
There are three types of elementary operations

- Interchanging of any two rows
- Multiplication to a row by a nonzero constant
- Adding a multiple of one row to another

Elementary matrix

An elementary matrix is a matrix that results from applying a single elementary row operation to an identity matrix.

Some examples are given below:

$$\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Multiply the second row of I_2 by -3.	Interchange the second and fourth rows of I_4.	Add 3 times the third row of I_3 to the first row.	Multiply the first row of I_3 by 1.

From Def it is clear that elementary matrices are always square.

Elementary matrices are important because they can be used to execute elementary row operations by matrix multiplication.

Theorem: If ***A*** is an $n \times n$ identity matrix, and if the elementary matrix ***E*** results by performing a certain row operation on the identity matrix, then the product *EA* is the matrix that results when the same row operation is performed on ***A***.

In short, this theorem states that an elementary row operation can be performed on a matrix ***A*** using a left multiplication by and appropriate elementary matrix.

Example 6: (Performing Row Operations by Matrix Multiplication). Consider the matrix

$$A=\begin{bmatrix}1 & 0 & 2 & 3\\ 2 & -1 & 3 & 6\\ 1 & 4 & 4 & 0\end{bmatrix}$$

Find an elementary matrix ***E*** such that ***EA*** is the matrix that results by adding 4 times the first row of ***A*** to the third row.

Solution: The matrix ***E*** must be 3×3 to conform for the product ***EA***. Thus, we obtain ***E*** by adding 4 times the first row of I_3 to the third row. This gives us $E=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 4 & 0 & 1\end{bmatrix}$

As a check, the product ***EA*** is $EA=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 4 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 2 & 3\\ 2 & -1 & 3 & 6\\ 1 & 4 & 4 & 0\end{bmatrix}=\begin{bmatrix}1 & 0 & 2 & 3\\ 2 & -1 & 3 & 6\\ 5 & 4 & 12 & 12\end{bmatrix}$

So left multiplication by ***E*** does, in fact, add 4 times the first row of ***A*** to the third row.

If an elementary row operation is applied to an identity matrix ***I*** to produce an elementary matrix ***E***, then there is a second row operation that, when applied to ***E***, produces ***I*** back again.

For example, if ***E*** is obtained by multiplying the *i-th* row of ***I*** by a nonzero scalar c, then ***I*** can be recovered by multiplying the *i-th* row of ***E*** by 1/c. The following table explains how to recover the identity matrix from an elementary matrix for each of the three elementary row operations. The operations on the right side of this table are called the **inverse operations** of the corresponding operations on the left side.

Row operation on I that produces E	**Row operation on E that reproduces I**
Multiply row *i* by $c \neq 0$	Multiply row *i* by 1/c
Interchange rows *i* and *j*	Interchange rows *i* and *j*
Add c times row *i* to row *j*	Add –c times row *i* to row *j*

Example 7: **(Recovering Identity Matrices from Elementary Matrices)**. Here are three examples that use inverses of row operations to recover the identity matrix from

$$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\rightarrow\begin{bmatrix}1 & 0\\ 0 & 7\end{bmatrix}\rightarrow\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

Multiply the second row by 7.

Multiply the second row by 1/7.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow[\text{second rows.}]{\text{Interchange the first and}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \xrightarrow[\text{second rows.}]{\text{Interchange the first and}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow[\text{to the first.}]{\text{Add 5 times the second row}} \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \xrightarrow[\text{to the first.}]{\text{Add -5 times the second row}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The next theorem is the basic result on the invertibility of elementary matrices.

<u>Theorem</u>: An elementary matrix is invertible and the inverse is also an elementary matrix.

<u>Example 8:</u> Let $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Compute ***E₁A, E₂A, E₃A*** and describe how these products can be obtained by elementary row operations on ***A***.

Solution We have

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, \quad E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix}, E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}$$

Addition of (–4) times row 1 of ***A*** to row 3 produces E_1A. (This is a row replacement operation.) An interchange of rows 1 and 2 of ***A*** produces E_2A and multiplication of row 3 of ***A*** by 5 produces E_3A.

<u>Left-multiplication</u> (that is, multiplication on the left) by E_1 in Example 8 has the same effect on any $3 \times n$ matrix. It adds – 4 times row 1 to row 3. In particular, since $E_1 I = E_1$, we see that E_1 itself is produced by the same row operation on the identity. Thus Example 8 illustrates the following general fact about elementary matrices.

Note: Since row operations are reversible, elementary matrices are invertible, for if ***E*** is produced by a row operation on ***I***, then there is another row operation of the same type that changes ***E*** back into ***I***. Hence there is an elementary matrix ***F*** such that ***FE = I***. Since ***E*** and ***F*** correspond to reverse operations, ***EF = I***.

Each elementary matrix ***E*** is invertible. The inverse of ***E*** is the elementary matrix of the same type that transforms ***E*** back into ***I***.

Example Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

Solution: To transform E_1 into ***I***, add + 4 times row 1 to row 3.

The elementary matrix that does that is $E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$

Theorem An $n \times n$ matrix ***A*** is invertible if and only if ***A*** is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces ***A*** to I_n also transforms I_n into A^{-1}.

An Algorithm for Finding A^{-1}: If we place ***A*** and ***I*** side-by-side to form an augmented matrix **[*A* *I*]**, then row operations on this matrix produce identical operations on ***A*** and ***I***. Then either there are row operations that transform ***A*** to I_n, and I_n to A^{-1}, or else ***A*** is not invertible.

Algorithm for Finding A^{-1}

Row reduce the augmented matrix **[*A* *I*]**. If ***A*** is row equivalent to ***I***, then **[*A* *I*]** is row equivalent to **[*I* A^{-1}]**. Otherwise, ***A*** does not have an inverse.

Example 9 Find the inverse of the matrix $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, if it exists.

Solution $[A \quad I] = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} R_{12}$

$$\overset{-4R_1+R_3}{\sim}\begin{bmatrix}1&0&3&0&1&0\\0&1&2&1&0&0\\0&-3&-4&0&-4&1\end{bmatrix}\overset{3R_2+R_3}{\sim}\begin{bmatrix}1&0&3&0&1&0\\0&1&2&1&0&0\\0&0&2&3&-4&1\end{bmatrix}$$

$$\overset{-2R_3+R_2}{\sim}\begin{bmatrix}1&0&3&0&1&0\\0&1&2&1&0&0\\0&0&1&3/2&-2&1/2\end{bmatrix}\overset{-3R_3+R_1}{\sim}\begin{bmatrix}1&0&0&-9/2&7&-3/2\\0&1&0&-2&4&-1\\0&0&1&3/2&-2&1/2\end{bmatrix}$$

Since $\boldsymbol{A} \sim \mathbf{I}$, we conclude that $\boldsymbol{A}$ is invertible, and $A^{-1}=\begin{bmatrix}-9/2&7&-3/2\\-2&4&-1\\3/2&-2&1/2\end{bmatrix}$

It is a good idea to check the final answer:

$$AA^{-1}=\begin{bmatrix}0&1&2\\1&0&3\\4&-3&8\end{bmatrix}\begin{bmatrix}-9/2&7&-3/2\\-2&4&-1\\3/2&-2&1/2\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

It is not necessary to check that $\boldsymbol{A^{-1}A = I}$ since $\boldsymbol{A}$ is invertible.

Example 10 Find the inverse of the matrix $A=\begin{bmatrix}1&2&-3&1\\-1&3&-3&-2\\2&0&1&5\\3&1&-2&5\end{bmatrix}$, if it exists.

Consider $detA=\begin{vmatrix}1&2&-3&1\\-1&3&-3&-2\\2&0&1&5\\3&1&-2&5\end{vmatrix}=\begin{vmatrix}1&2&-3&1\\0&5&-6&-1\\0&-4&7&3\\0&-5&7&2\end{vmatrix}$

operating $R_2+R_1, R_3-2R_1, R_4-3R_1$

Expand from first column $=\begin{vmatrix}5&-6&-1\\-4&7&3\\-5&7&2\end{vmatrix}=\begin{vmatrix}5&-6&-1\\1&1&2\\0&1&1\end{vmatrix}=5(1-2)+6(1-0)-1(1-0)=0$

As the given matrix is singular, so it is not invertible.

Example 11 Find the inverse of the given matrix if possible $A=\begin{bmatrix}1 & 0 & 1\\ 2 & 1 & 1\\ 3 & 1 & 1\end{bmatrix}$

Solution: $\det A=\begin{vmatrix}1 & 0 & 1\\ 2 & 1 & 1\\ 3 & 1 & 1\end{vmatrix}=-1$

As the given matrix is non singular therefore, inverse is possible.

$$\begin{bmatrix}1 & 0 & 1\\ 2 & 1 & 1\\ 3 & 1 & 1\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & -1\\ 0 & 1 & -2\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ -3 & 0 & 1\end{bmatrix}$$

R_2-2R_1, R_3-3R_1

$$\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & -1\\ 0 & 0 & -1\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ -1 & -1 & 1\end{bmatrix}$$

R_3-R_2

$$\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & -1\\ 0 & 0 & 1\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ 1 & 1 & -1\end{bmatrix}$$

$Multiply R_3\ by -1$

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \qquad \begin{bmatrix}0 & -1 & 1\\ -1 & 2 & -1\\ 1 & 1 & -1\end{bmatrix}$$

R_1-R_3, R_2+R_3

Hence the inverse of matrix A is $A^{-1}=\begin{bmatrix}0 & -1 & 1\\ -1 & 2 & -1\\ 1 & 1 & -1\end{bmatrix}$

<u>Example 12</u> Find the inverse of the matrix $A=\begin{bmatrix}1 & 2 & 2\\ 2 & 2 & 3\\ 5 & 2 & 3\end{bmatrix}$

Solution $\det A=\begin{vmatrix}1 & 2 & 2\\ 2 & 2 & 3\\ 5 & 2 & 3\end{vmatrix}=6$

As the given matrix is non singular, therefore, inverse of the matrix is possible.
We reduce it to reduce echelon form.

$$\begin{bmatrix}1 & 2 & 2\\ 2 & 2 & 3\\ 5 & 2 & 3\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 2 & 2\\ 0 & -2 & -1\\ 0 & -8 & -7\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ -5 & 0 & 1\end{bmatrix}$$

R_2-2R_1, R_3-5R_1

$$\begin{bmatrix}1 & 2 & 2\\ 0 & 1 & 1/2\\ 0 & -8 & -7\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ 1 & -1/2 & 0\\ -5 & 0 & 1\end{bmatrix}$$

multiply 2nd row by -1/2

$$\begin{bmatrix}1 & 2 & 2\\ 0 & 1 & 1/2\\ 0 & 0 & -3\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ 1 & -1/2 & 0\\ 3 & -4 & 1\end{bmatrix}$$

R_3+8R_2

$$\begin{bmatrix}1 & 2 & 2\\ 0 & 1 & 1/2\\ 0 & 0 & 1\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & 0\\ 1 & -1/2 & 0\\ -1 & 4/3 & -1/3\end{bmatrix}$$

Multiply 3rd row by -1/3

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 3 & -\frac{8}{3} & \frac{2}{3} \\ 3/2 & -7/6 & 1/6 \\ -1 & 4/3 & -1/3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & -1/3 & 1/3 \\ 3/2 & -7/6 & 1/6 \\ -1 & 4/3 & -1/3 \end{bmatrix}$$

$R_2 - (1/2)R_3, R_1 - 2R_3$

Hence the inverse of the original matrix $A^{-1} = \begin{bmatrix} 0 & -1/3 & 1/3 \\ 3/2 & -7/6 & 1/6 \\ -1 & 4/3 & -1/3 \end{bmatrix}$

Exercises

In exercises 1 to 4, find the inverses of the matrices, if they exist. Use elementary row operations.

1. $\begin{bmatrix} 1 & 2 \\ 5 & 9 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 0 & 5 \\ 1 & 1 & 0 \\ 3 & 2 & 6 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 4 & -3 \\ -2 & -7 & 6 \\ 1 & 7 & -2 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & 2 \\ -1 & 0 & 8 \end{bmatrix}$

5. $\begin{bmatrix} 3 & 4 & -1 \\ 1 & 0 & 3 \\ 2 & 5 & -4 \end{bmatrix}$

6. $\begin{bmatrix} -1 & 3 & -4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{bmatrix}$

7. Let $A = \begin{bmatrix} -1 & -5 & -7 \\ 2 & 5 & 6 \\ 1 & 3 & 4 \end{bmatrix}$. Find the third column of A^{-1} without computing the other columns.

8. Let $A = \begin{bmatrix} -25 & -9 & -27 \\ 546 & 180 & 537 \\ 154 & 50 & 149 \end{bmatrix}$. Find the second and third columns of A^{-1} without computing the first column.

9. Find an elementary matrix E that satisfies the equation.
(a) EA = B (b) EB = A (c) EA = C (d) EC = A
where $A = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{bmatrix}, B = \begin{bmatrix} 8 & 1 & 5 \\ 2 & -7 & -1 \\ 3 & 4 & 1 \end{bmatrix}, C = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 2 & -7 & 3 \end{bmatrix}$.

10. Consider the matrix $A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(a) Find elementary matrices E_1 and E_2 such that $E_2E_1A=I$.
(b) Write A^{-1} as a product of two elementary matrices.

(c) Write A as a product of two elementary matrices.

In exercises 11 and 12, express A and A^{-1} as products of elementary matrices.

11. $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

13. Factor the matrix $A = \begin{bmatrix} 0 & 1 & 7 & 8 \\ 1 & 3 & 3 & 8 \\ -2 & -5 & 1 & -8 \end{bmatrix}$ as A = EFGR, where E, F, and G are elementary matrices and R is in row echelon form.

Lecture 13

Characterizations of Invertible Matrices

This chapter involves a few techniques of solving the system of n linear equations in **n** unknowns.

Solving Linear Systems by Matrix Inversion

Theorem:-

If $\boldsymbol{A}$ is an invertible $n \times n$ matrix, then for each $\boldsymbol{b}$ in $\boldsymbol{R}^n$, the equation $\boldsymbol{Ax} = \boldsymbol{b}$ has the unique solution $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

Proof:-

Let $\boldsymbol{b}$ be any vector in $\boldsymbol{R}^n$. A solution must exists because when $\boldsymbol{A}^{-1}\boldsymbol{b}$ is substituted for $\boldsymbol{x}$ we have $\boldsymbol{Ax} = \boldsymbol{A}(\boldsymbol{A}^{-1}\boldsymbol{b}) = \boldsymbol{Ib} = \boldsymbol{b}$. So $\boldsymbol{A}^{-1}\boldsymbol{b}$ is solution.

To prove that the solution is unique, we show that if $\boldsymbol{u}$ is any other solution, then $\boldsymbol{u}$ must be $\boldsymbol{A}^{-1}\boldsymbol{b}$ .i.e. $\boldsymbol{u} = \boldsymbol{A}^{-1}\boldsymbol{b}$. Indeed, if $\boldsymbol{Au} = \boldsymbol{b}$, we can multiply both sides by $\boldsymbol{A}^{-1}$ and obtain

$$A^{-1}Au = A^{-1}b, \qquad Iu = A^{-1}b, \quad and \quad u = A^{-1}b$$

Example 1:-

Solve the system of linear equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 5 \\ 2x_1 + 5x_2 + 3x_3 &= 3 \\ x_1 + \qquad 8x_3 &= 17 \end{aligned}$$

by inverse matrix method..

Solution:-

Consider the linear system

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 5 \\ 2x_1 + 5x_2 + 3x_3 &= 3 \\ x_1 + \qquad 8x_3 &= 17 \end{aligned}$$

This system can be written in matrix form as $\boldsymbol{Ax} = \boldsymbol{b}$, where

$$A=\begin{bmatrix}1&2&3\\2&5&3\\1&0&8\end{bmatrix},\quad x=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},\quad b=\begin{bmatrix}5\\3\\17\end{bmatrix}$$

Here, det (***A***) = 40 – 2 (16 – 3) + 3(0 – 5) = 40 – 26 – 15 = -1 $\neq$ 0

Therefore, ***A*** is invertible. Now we apply the inversion algorithm:

$$\left[\begin{array}{ccc|ccc}1&2&3&1&0&0\\2&5&3&0&1&0\\1&0&8&0&0&1\end{array}\right]$$

$$\left[\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&1&-3&-2&1&0\\0&-2&5&-1&0&1\end{array}\right]-2R_1+R_2,\quad -1R_1+R_3$$

$$\left[\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&1&-3&-2&1&0\\0&0&-1&-5&2&1\end{array}\right]2R_2+R_3$$

$$\left[\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&1&-3&-2&1&0\\0&0&1&5&-2&-1\end{array}\right]-1R_3$$

$$\left[\begin{array}{ccc|ccc}1&2&0&-14&6&3\\0&1&0&13&-5&-3\\0&0&1&5&-2&-1\end{array}\right]3R_3+R_2,\quad -3R_3+R_1$$

$$\left[\begin{array}{ccc|ccc}1&0&0&-40&16&9\\0&1&0&13&-5&-3\\0&0&1&5&-2&-1\end{array}\right]-2R_2+R_1$$

Hence, $A^{-1}=\begin{bmatrix}-40&16&9\\13&-5&-3\\5&-2&-1\end{bmatrix}$

Thus, the solution of the linear system is $x=A^{-1}b=\begin{bmatrix}-40&16&9\\13&-5&-3\\5&-2&-1\end{bmatrix}\begin{bmatrix}5\\3\\17\end{bmatrix}=\begin{bmatrix}1\\-1\\2\end{bmatrix}$

Or, equivalently $x_1=1, x_2=-1, x_3=2$.

Note: This method applies only when the number of equations = number of unknown and fails if given matrix is not invertible.

Example 2:-

Solve the system of linear equation

$$x_1 + 6x_2 + 4x_3 = 2$$
$$2x_1 + 4x_2 - x_3 = 3$$
$$-x_1 + 2x_2 + 5x_3 = 3$$

by inversion method.

Solution:-

This system can be written in matrix form ***Ax = b***, where

$$A = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{bmatrix}$$

Here det (***A***) = 1(20 + 2) – 6(10 – 1) + 4 (4+4) = 22 – 54 + 32 = 0

Therefore, ***A*** is not invertible. Hence, the inversion method can not be used.

Theorem:-

If ***Ax = 0*** is a homogeneous linear system of **n** equations in **n** unknowns, then the system has only the trivial solution if and only if the coefficient matrix ***A*** is invertible.

Example 3:-

State whether the following system of linear equation has a solution or not?

$$x_1 + 2x_2 + 3x_3 = 0$$
$$2x_1 + 5x_2 + 3x_3 = 0$$
$$x_1 + \qquad 8x_3 = 0$$

Solution:-

By Example 1, we get

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$$ is an invertible matrix.

Then, by Theorem (above) says that the homogeneous linear system

$$x_1 + 2x_2 + 3x_3 = 0$$
$$2x_1 + 5x_2 + 3x_3 = 0$$
$$x_1 + \qquad 8x_3 = 0$$

has only the trivial solution.

Solving Multiple Linear Systems with a Common Coefficient Matrix

In many applications one is concerned with solving a sequence of linear systems

$$Ax_1 = b_1,\ \ Ax_2 = b_2,\ \cdots\ ,\ Ax_k = b_k \qquad (1)$$

each of which has the same coefficient matrix ***A***. If the coefficient matrix ***A*** in (1) is invertible, $x_1 = A^{-1}b_1,\ x_2 = A^{-1}b_2,\ \cdots\ ,\ x_k = A^{-1}b_k$. However, this procedure cannot be used unless ***A*** is invertible.

Theorem (Invertible Matrix Theorem): Let ***A*** be a square $n \times n$ matrix. Then the following statements are equivalent. (Means if any one holds then all are true).

(a) ***A*** is an invertible matrix.

(b) ***A*** is row equivalent to the $n \times n$ identity matrix.

(c) ***A*** has **n** pivot positions.

(d) The equation ***Ax = 0*** has only the trivial solution.

(e) The columns of ***A*** form a linearly independent set.

(f) The linear transformation $x \to Ax$ is one-to-one.

(g) The equation ***Ax = b*** has at least one solution for each ***b*** in $\boldsymbol{R^n}$.

(h) The columns of ***A*** span $\boldsymbol{R^n}$.

(i) The linear transformation $x \to Ax$ maps $\boldsymbol{R^n}$ onto $\boldsymbol{R^n}$.

(j) There is a $n \times n$ matrix ***C*** such that ***CA = I***.

(k) There is a $n \times n$ matrix ***D*** such that ***AD = I.***

(l) $\boldsymbol{A^T}$ is an invertible matrix.

Example 4:-

Show that the matrix

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix} \text{ is invertible.}$$

Solution:-

By row equivalent,

$$A=\begin{bmatrix}1 & 0 & -2\\ 0 & 1 & 4\\ 0 & -1 & -1\end{bmatrix}\sim\begin{bmatrix}1 & 0 & -2\\ 0 & 1 & 4\\ 0 & 0 & 3\end{bmatrix}$$

It shows that A has three pivot positions and hence is invertible, by the Invertible Matrix Theorem (c).

Example 7 Find A^t and show that A^t is invertible matrix.

$$A=\begin{bmatrix}1 & 2 & 1 & 0\\ 2 & 4 & 1 & 2\\ 1 & 0 & 1 & 2\\ 0 & 1 & 1 & 2\end{bmatrix}$$

Solution:-

$$A^t=\begin{bmatrix}1 & 2 & 1 & 0\\ 2 & 4 & 0 & 1\\ 1 & 1 & 1 & 1\\ 0 & 2 & 2 & 2\end{bmatrix}$$

Now by row equivalent of A,

$$A=\begin{bmatrix}1 & 2 & 1 & 0\\ 2 & 4 & 1 & 2\\ 1 & 0 & 1 & 2\\ 0 & 1 & 1 & 2\end{bmatrix}\underset{R_2-2R_1,\,R_3-R_1}{\sim}\begin{bmatrix}1 & 2 & 1 & 0\\ 0 & 0 & -1 & 2\\ 0 & -2 & 0 & 2\\ 0 & 1 & 1 & 2\end{bmatrix}\underset{R_{23}}{\sim}\begin{bmatrix}1 & 2 & 1 & 0\\ 0 & -2 & 0 & 2\\ 0 & 0 & -1 & 2\\ 0 & 1 & 1 & 2\end{bmatrix}$$

$$\underset{-\frac{1}{2}R_2}{\sim}\begin{bmatrix}1 & 2 & 1 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & -1 & 2\\ 0 & 1 & 1 & 2\end{bmatrix}\underset{R_4-R_2}{\sim}\begin{bmatrix}1 & 2 & 1 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & -1 & 2\\ 0 & 0 & 1 & 3\end{bmatrix}\underset{(-1)R_3}{\sim}\begin{bmatrix}1 & 2 & 1 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & 1 & -2\\ 0 & 0 & 1 & 3\end{bmatrix}\underset{-R_3+R_4}{\sim}\begin{bmatrix}1 & 2 & 1 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & 1 & -2\\ 0 & 0 & 0 & 5\end{bmatrix}$$

Here ***A*** has 4 pivot positions so by Invertible Matrix Theorem (c) A is invertible. Thus by (l) A^t is invertible.

Find a Matrix from Linear Transformation:-

We can find a matrix corresponding to every transformation. In this section we will learn how to find a matrix attached with a linear transformation.

Example:-

Let L be the linear transformation from R^2 to P_2 (Polynomials of order 2) defined by

$T(x, y) = x\,y\,t + (x + y)t^2$

Find the matrix representing T with respect to the standard bases.

Solution:-

Let A = {(1,0),(0,1)} be the basis of R^2, then

$T(1,0) = t^2 = (0,0,1)$ (This triple represents the coefficients of polynomial t^2)

Similarly, $T(0, 1) = t^2 = (0,0,1)$.Hence the matrix is given by

$$A=\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}$$

Now we will proceed with a more complicated example.

Example:-

Let T be the linear transformation from R^2 to R^2 such that $T(x, y) = (x , y + 2x)$. Find a matrix A for T.

Solution:-

This matrix is found by finding $T(1, 0) = (1, 2)$ and $T(0, 1) = (0, 1)$ the matrix is $A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$.

Important Note:-

It should be clear that the Invertible Matrix Theorem applies only to square matrices. For example, if the columns of a 4×3 matrix are linearly independent, we cannot use the Invertible Matrix Theorem to conclude any thing about the existence or nonexistence of solutions to equations of the form $\boldsymbol{Ax = b}$.

Definition (Inverse of a Linear Transformation) A linear transformation $T : R^n \to R^n$ is said to be invertible (left as well as right) if there exists a function $T' : R^n \to R^n$ such that

$T'(T(x)) = x \ \forall \quad x \in R^n$
$T(T'(x)) = x \ \forall \quad x \in R^n$

Here S is called inverse of linear transformation T.

Important Note:-

If the inverse of a linear transformation exits then it is unique.

Proposition:-

Let $T : R^n \to R^m$ be linear transformation, given as $T(x) = Ax \quad , \forall \ x \in R^n$, where A is a $m \times n$ matrix. The mapping T is invertible if the system $y = Ax$ has a unique solution.

Case 1:

If $m < n$, then the system $Ax = y$ has either no solution or infinitely many solution, for any y in $R^{m.}$ Therefore $y = Ax$ is non-invertible.

Case 2:

If $m = n$, then the system $Ax = y$ has a unique solution if and only if Rank (A) = n.

Case 3:

If $m > n$, then the transformation $y = Ax$ is non-invertible because we can find a vector y in R^m such that $Ax = y$ is inconsistent.

Exercises

1. Solve the system of linear equations

$$x_1 + x_2 + x_3 = 8$$
$$2x_2 + 3x_3 = 24$$
$$5x_1 + 5x_2 + x_3 = 8$$

by inverse matrix method.

2. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix}, b_1 = \begin{bmatrix} 5 \\ 7 \end{bmatrix}, b_2 = \begin{bmatrix} -2 \\ -2 \end{bmatrix}, b_3 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, and\ b_4 = \begin{bmatrix} 1 \\ 7 \end{bmatrix}.$

(a) Find A^{-1} and use it to solve the equations $Ax = b_1$, $Ax = b_2$, $Ax = b_3$, $Ax = b_4$.

(b) Solve the four equations in part (a) by row reducing the augmented matrix $\begin{bmatrix} A & b_1 & b_2 & b_3 & b_4 \end{bmatrix}$.

3. (a) Solve the two systems of linear equations

$$\begin{aligned} x_1 + 2x_2 + x_3 &= -1 \\ x_1 + 3x_2 + 2x_3 &= 3 \\ x_2 + 2x_3 &= 4 \end{aligned} \quad \text{and} \quad \begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ x_1 + 3x_2 + 2x_3 &= 0 \\ x_2 + 2x_3 &= 4 \end{aligned}$$

by row reduction.

(b) Write the systems in (a) as Ax = b_1 and Ax = b_2, and then solve each of them by the method of inversion.

Determine which of the matrices in exercises 4 to 10 are invertible.

4. $\begin{bmatrix} -4 & 16 \\ 3 & -9 \end{bmatrix}$ 5. $\begin{bmatrix} 5 & 0 & 3 \\ 7 & 0 & 2 \\ 9 & 0 & 1 \end{bmatrix}$ 6. $\begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$ 7. $\begin{bmatrix} 5 & -9 & 3 \\ 0 & 3 & 4 \\ 1 & 0 & 3 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 1 & -2 & -1 \\ -2 & -6 & 3 & 2 \\ 3 & 5 & 8 & -3 \end{bmatrix}$ 9. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 10 \end{bmatrix}$ 10. $\begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$

11. $\begin{bmatrix} 5 & 4 & 3 & 6 & 3 \\ 7 & 6 & 5 & 9 & 5 \\ 8 & 6 & 4 & 10 & 4 \\ 9 & -8 & 9 & -5 & 8 \\ 10 & 8 & 7 & -9 & 7 \end{bmatrix}$

12. Suppose that ***A*** and ***B*** are $n \times n$ matrices and the equation ***ABx = 0*** has a nontrivial solution. What can you say about the matrix ***AB***?

13**.** What can we say about a one-to-one linear transformation ***T*** from $\boldsymbol{R^n}$ into $\boldsymbol{R^n}$?

14. Let $T : R^2 \to R^2$ be a linear transformation given as $T(x) = 5x$, then find a matrix A of linear transformation T.

In exercises 15 and 16, T is a linear transformation from R^2 into R^2. Show that T is invertible.

15. $T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2)$

16. $T(x_1, x_2) = (6x_1 - 8x_2, -5x_1 + 7x_2)$

Lecture 14

Partitioned Matrices

A **block matrix** or a **partitioned matrix** is a partition of a matrix into rectangular smaller matrices called **blocks**.Partitioned matrices appear often in modern applications of linear algebra because the notation simplifies many discussions and highlights essential structure in matrix calculations. This section provides an opportunity to review matrix algebra and use of the Invertible Matrix Theorem.

General Partitioning

A matrix can be **partitioned** (subdivided) into **sub matrices** (also called **blocks**) in various ways by inserting lines between selected rows and columns.

Example 1:-

The matrix

$$P = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \\ 3 & 3 & 4 & 4 \end{bmatrix}$$

can be partitioned into four 2×2 blocks

$$P_{11} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, P_{12} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, P_{21} = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}, P_{22} = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix},$$

The partitioned matrix can then be written as

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}$$

Note:-

It is important to know that in how many ways to block up a ordinary matrix A? See the following example in which a matrix A is block up into three different ways.

Example 3:-

Let A be a general matrix of 5×3 order, we have

Partition (a)

$$A = \left[\begin{array}{c|cc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \hline a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

In this case we partitioned the matrix into four sub matrices. Also notice that we simplified the matrix into a more compact form and in this compact form we've mixed and matched some of our notation. The partitioned matrix can be thought of as a smaller matrix with four entries, except this time each of the entries are matrices instead of numbers and so we used capital letters to represent the entries and subscripted each on with the location in the partitioned matrix.

Be careful not to confuse the location subscripts on each of the sub matrices with the size of each sub matrix. In this case A_{11} is a 2×1 sub matrix of *A*, A_{12} is a 2×2 sub matrix of *A*, A_{21} is a 3×1 sub matrix of A and A_{22} is a 3×3 sub matrix of A.

Partition (b)

$$A = \left[\begin{array}{c|c|c} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{array}\right] = [\mathbf{c}_1 \mid \mathbf{c}_2 \mid \mathbf{c}_3]$$

In this case we partitioned *A* into three column matrices each representing one column in the original matrix. Again, note that we used the standard column matrix notation (the bold face letters) and subscripted each one with the location in the partitioned matrix.

The c_i in the partitioned matrix are sometimes called the **column matrices of *A***.

Partition (c)

$$A = \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ \hline a_{21} & a_{22} & a_{23} \\ \hline a_{31} & a_{32} & a_{33} \\ \hline a_{41} & a_{42} & a_{43} \\ \hline a_{51} & a_{52} & a_{53} \end{array}\right] = \left[\begin{array}{c} \mathbf{r}_1 \\ \hline \mathbf{r}_2 \\ \hline \mathbf{r}_3 \\ \hline \mathbf{r}_4 \\ \hline \mathbf{r}_5 \end{array}\right]$$

Just as we can partition a matrix into each of its columns as we did in the previous part we can also partition a matrix into each of its rows. The r_i in the partitioned matrix are sometimes called the **row matrices** of ***A.***

Addition of Blocked Matrices:-

If matrices ***A*** and ***B*** are the same size and are partitioned in exactly the same way, then it is natural to make the same partition of the ordinary matrix sum ***A + B***. In this case, each block of ***A + B*** is the (matrix) sum of the corresponding blocks of ***A*** and ***B***.

Multiplication of a partitioned matrix by a scalar is also computed block by block.

Multiplication of Partitioned Matrices:-

If $$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \quad and \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

and if the sizes of the blocks confirm for the required operations, then

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \\ A_{31}B_{11} + A_{32}B_{21} & A_{31}B_{12} + A_{32}B_{22} \end{bmatrix}$$

It is known as **block multiplication**.

Example 3:-

Find the block multiplication of the following partitioned matrices:

$$A = \left[\begin{array}{ccc:cc} 3 & -4 & 1 & 0 & 2 \\ -1 & 5 & -3 & 1 & 4 \\ \hdashline 2 & 0 & -2 & 1 & 6 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \left[\begin{array}{cc} 2 & -1 \\ 3 & 0 \\ -5 & 1 \\ \hdashline 4 & -3 \\ 0 & 2 \end{array}\right] = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$$

Solution:-

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

This is a valid formula because the sizes of the blocks are such that all of the operations can be performed:

$$A_{11}B_{11} + A_{12}B_{21} = \begin{bmatrix} 3 & -4 & 1 \\ -1 & 5 & -3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -11 & 2 \\ 32 & 3 \end{bmatrix}$$

$$A_{21}B_{11} + A_{22}B_{21} = \begin{bmatrix} 2 & 0 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 6 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 18 & 5 \end{bmatrix}$$

Thus,

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix} = \begin{bmatrix} -11 & 2 \\ 32 & 3 \\ 18 & 5 \end{bmatrix} \quad AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix}$$

<u>Note</u>:- We see result is same when we multiply A and B without partitions

$$AB = \begin{bmatrix} 3 & -4 & 1 & 0 & 2 \\ -1 & 5 & -3 & 1 & 4 \\ 2 & 0 & -2 & 1 & 6 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 1 \\ 4 & -3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -11 & 2 \\ 32 & 3 \\ 18 & 5 \end{bmatrix}$$

Note:- Some time it more useful to find the square and cube powers of a matrix.

<u>Example 4</u>:-

Making block up of matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}, \text{ evaluate } A^2 ?$$

Solution:-

We partition A as shown below

$$A = \left[\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} I_3 & O_{32} & A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t & O_{12} & 1 \end{bmatrix} \quad \text{where } A_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Now
$$A^2 = \begin{bmatrix} I_3 & O_{32} & A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t & O_{12} & 1 \end{bmatrix}\begin{bmatrix} I_3 & O_{32} & A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t & O_{12} & 1 \end{bmatrix} = \begin{bmatrix} I_3 + A_1A_1^t & O_{32} & A_1 + A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t + A_1^t & O_{12} & A_1^tA_1 + 1 \end{bmatrix}$$

$$I_3 + A_1A_1^t = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, A_1 + A_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, A_1^t + A_1^t = \begin{bmatrix} 2 & 2 & 2 \end{bmatrix}$$

$$A_1^tA_1 + 1 = \begin{bmatrix} 4 \end{bmatrix}$$

Hence
$$A^2 = \begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 2 \\ 1 & 2 & 1 & 0 & 0 & 2 \\ 1 & 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 2 & 2 & 2 & 0 & 0 & 4 \end{bmatrix}$$

Example 6: Let $A = \begin{bmatrix} -3 & 1 & 2 \\ 1 & -4 & 5 \end{bmatrix}$ *and* $B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$.

Verify that $AB = col_1(A)row_1(B) + col_2(A)row_2(B) + col_3(A)row_3(B)$

Solution Each term above is an outer product.

By the ordinary row – column rule,

$$col_1(A)row_1(B) = \begin{bmatrix} -3 \\ 1 \end{bmatrix}[a \quad b] = \begin{bmatrix} -3a & -3b \\ a & b \end{bmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$col_2(A)row_2(B) = \begin{bmatrix} 1 \\ -4 \end{bmatrix}[c \quad d] = \begin{bmatrix} c & d \\ -4c & -4d \end{bmatrix}$$

$$col_3(A)row_3(B) = \begin{bmatrix} 2 \\ 5 \end{bmatrix}[e \quad f] = \begin{bmatrix} 2e & 2f \\ 5e & 5f \end{bmatrix}$$

Thus $$\sum_{k=1}^{3} col_k(A)row_k(B) = \begin{bmatrix} -3a+c+2e & -3b+d+2f \\ a-4c+5e & b-4d+5f \end{bmatrix}$$

This matrix is obviously ***AB***.

Toeplitz matrix:-

A matrix in which each descending diagonal from left to right is constant is called a **Toeplitz matrix** or **diagonal-constant matrix**

Example:-

The matrix

$$A = \begin{bmatrix} a & 1 & 2 & 3 \\ 4 & a & 1 & 2 \\ 5 & 4 & a & 1 \\ 6 & 5 & 4 & a \end{bmatrix}$$ is a Toeplitz matrix.

Block Toeplitz matrix:-

A blocked matrix in which blocks (blocked matrices) are repeated down the diagonals of the matrix is called a blocked Toeplitz matrix.

A block Toeplitz matrix **B** has the form

$$\text{B} = \begin{bmatrix} B(1,1) & B(1,2) & B(1,3) & B(1,4) & B(1,5) \\ B(2,1) & B(1,1) & B(1,2) & B(1,3) & B(1,4) \\ B(3,1) & B(2,1) & B(1,1) & B(1,2) & B(1,3) \\ B(4,1) & B(3,1) & B(2,1) & B(1,1) & B(1,2) \\ B(5,1) & B(4,1) & B(3,1) & B(2,1) & B(1,1) \end{bmatrix}$$

Inverses of Partitioned Matrices:-

In this section we will study about the techniques of inverse of blocked matrices.

Block Diagonal Matrices:

A partitioned matrix ***A*** is said to be block diagonal if the matrices on the main diagonal are square and all other position matrices are zero, i.e.

$$A = \begin{bmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k \end{bmatrix} \qquad (1)$$

where the matrices ***D₁, D₂, …, Dₖ*** are square. It can be shown that the matrix A in (1) is invertible if and only if each matrix on the diagonal is invertible. i.e.

$$A^{-1} = \begin{bmatrix} D_1^{-1} & 0 & \dots & 0 \\ 0 & D_2^{-1} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k^{-1} \end{bmatrix}$$

Example 7: Let A be a block diagonal matrix

$$A = \left[\begin{array}{cc|cc|c} 8 & -7 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 5 & 2 & 0 \\ \hline 0 & 0 & 0 & 0 & 4 \end{array}\right]$$

Find A^{-1}.

Solution:-

There are three matrices on the main diagonal; two are 2x2 matrices and one is 1×1 matrix.

In order to find A^{-1}, we evaluate the inverses of three matrices lie in main diagonal of A.

Let $A_{11} = \begin{pmatrix} 8 & -7 \\ 1 & -1 \end{pmatrix}, A_{22} = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$ *and* $A_{33} = (4)$ are matrices of main diagonal of A. Then

$A_{11}^{-1} = \dfrac{AdjA_{11}}{A_{11}} = \dfrac{\begin{pmatrix} -1 & 7 \\ -1 & 8 \end{pmatrix}}{-1} = \begin{pmatrix} 1 & -7 \\ 1 & -8 \end{pmatrix}$. Similarly we can find inverses A_{22} and A_{33}. Thus

$$\Rightarrow \quad A^{-1} = \left[\begin{array}{cc:cc:c} 1 & -7 & 0 & 0 & 0 \\ 1 & -8 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & -5 & 3 & 0 \\ \hdashline 0 & 0 & 0 & 0 & \frac{1}{4} \end{array}\right]$$

Block Upper Triangular Matrices:

A partitioned square matrix ***A*** is said to be block upper triangular if the matrices on the main diagonal are square and all matrices below the main diagonal are zero; that is, the matrix is partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ O & A_{22} & \dots & A_{2k} \\ \vdots & \vdots & & \vdots \\ O & O & \dots & A_{kk} \end{bmatrix}$$ where the matrices $A_{11}, A_{22}, \dots, A_{kk}$ are square.

Note: The definition of block lower triangular matrix is similar.

Here we are going to introduce a formula for finding inverse of a block upper triangular matrix in the following example.

Example 8:-

Let A be a block upper triangular matrix of the form

$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ where the orders of A_{11} and A_{22} are $p \times p$ *and* $q \times q$ *respectively*. Find A^{-1}.

Solution:-

Let B = $\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ be inverse of A .i.e. $A^{-1} = B$, then

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & O \\ O & I_q \end{bmatrix}$$

$$\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{22}B_{21} & A_{22}B_{22} \end{bmatrix} = \begin{bmatrix} I_p & O \\ O & I_q \end{bmatrix}$$

By comparing corresponding entries, we have

$$A_{11}B_{11} + A_{12}B_{21} = I_p \qquad (1)$$

$$A_{11}B_{12} + A_{12}B_{22} = O \qquad (2)$$

$$A_{22}B_{21} = O \qquad (3)$$

$$A_{22}B_{22} = I_q \qquad (4)$$

Since A_{22} is a square matrix, so by Invertible Matrix Theorem, we have $\boldsymbol{A_{22}}$ is invertible. Thus by eq.(4), $\boldsymbol{B_{22} = A_{22}^{-1}}$. Now by eq. (3), we have $B_{21} = A_{22}^{-1}O = O$. From eq.(1)

$$A_{11}B_{11} + O = I_p$$

$$\Rightarrow A_{11}B_{11} = I_p$$

$$\Rightarrow B_{11} = A_{11}^{-1}$$

Finally, form (2),

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \quad and \quad B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

Thus

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{bmatrix} \qquad (5)$$

<u>Example 9</u>:-

Find A^{-1} of

$$A = \begin{bmatrix} 4 & 7 & -5 & 3 \\ 3 & 5 & 3 & -2 \\ 0 & 0 & 7 & 2 \\ 0 & 0 & 3 & 1 \end{bmatrix}$$

Solution:-

Let partition given matrix A in form

$$A = \left[\begin{array}{cc|cc} 4 & 7 & -5 & 3 \\ 3 & 5 & 3 & -2 \\ \hline 0 & 0 & 7 & 2 \\ 0 & 0 & 3 & 1 \end{array}\right]$$

Put $A_{11} = \begin{bmatrix} 4 & 7 \\ 3 & 5 \end{bmatrix}, A_{12} = \begin{bmatrix} -5 & 3 \\ 3 & -2 \end{bmatrix} \quad and \quad A_{22} = \begin{bmatrix} 7 & 2 \\ 3 & 1 \end{bmatrix}$

Thus $A_{11}^{-1} = \begin{bmatrix} -5 & 7 \\ 3 & -4 \end{bmatrix} and\, A_{22}^{-1} = \begin{bmatrix} 1 & -2 \\ -3 & 7 \end{bmatrix}$

Moreover,

$$-A_{11}^{-1}A_{12}A_{22}^{-1} = \begin{bmatrix} -5 & 7 \\ 3 & -4 \end{bmatrix}\begin{bmatrix} -5 & 3 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} -133 & 295 \\ 78 & -173 \end{bmatrix}$$

So by (5), we have

$$A^{-1} = \begin{bmatrix} -5 & 7 & -133 & 295 \\ 3 & -4 & 78 & -173 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -3 & 7 \end{bmatrix}$$

Exercises

In exercises 1 to 3, the matrices A, B, C, X, Y, Z, and I are all n x n and satisfy the indicated equation

1. $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ X & Y \end{bmatrix} = \begin{bmatrix} 0 & I \\ Z & 0 \end{bmatrix}$

2. $\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}\begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

3. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

4. Suppose that A_{11} is an invertible matrix. Find matrices X and Y such that the product below has the form indicated. Also compute B_{22}.

$$\begin{bmatrix} I & 0 & 0 \\ X & I & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \\ 0 & B_{32} \end{bmatrix}$$

5. The inverse of $\begin{bmatrix} I & 0 & 0 \\ C & I & 0 \\ A & B & I \end{bmatrix}$ *is* $\begin{bmatrix} I & 0 & 0 \\ Z & I & 0 \\ X & Y & I \end{bmatrix}$. Find X, Y and Z.

6. Show that $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ is invertible and find its inverse.

7. Compute $\boldsymbol{X^T X}$, when $\boldsymbol{X}$ is partitioned as $[\boldsymbol{X_1} \quad \boldsymbol{X_2}]$.

In exercises 8 and 9, determine whether block multiplication can be used to compute the product using the partitions shown. If so, compute the product by block multiplication.

8. (a) $\left[\begin{array}{cc:cc} -1 & 2 & 1 & 5 \\ 0 & -3 & 4 & 2 \\ \hdashline 1 & 5 & 6 & 1 \end{array}\right]\left[\begin{array}{cc:c} -2 & 1 & 4 \\ -3 & 5 & 2 \\ \hdashline 7 & -1 & 5 \\ 0 & 3 & -3 \end{array}\right]$ (b) $\left[\begin{array}{ccc:c} -1 & 2 & 1 & 5 \\ \hdashline 0 & -3 & 4 & 2 \\ 1 & 5 & 6 & 1 \end{array}\right]\left[\begin{array}{cc:c} -2 & 1 & 4 \\ -3 & 5 & 2 \\ 7 & -1 & 5 \\ \hdashline 0 & 3 & -3 \end{array}\right]$

9 (a) $\left[\begin{array}{ccc:c} 3 & -1 & 0 & -3 \\ 2 & 1 & 4 & 5 \end{array}\right]\left[\begin{array}{ccc} 2 & -4 & 1 \\ 3 & 0 & 2 \\ 1 & -3 & 5 \\ \hdashline 2 & 1 & 4 \end{array}\right]$ (b) $\left[\begin{array}{cc} 2 & -5 \\ 1 & 3 \\ 0 & 5 \\ \hdashline 1 & 4 \end{array}\right]\left[\begin{array}{cc:cc} 2 & -1 & 3 & -4 \\ 0 & 1 & 5 & 7 \end{array}\right]$

10. Compute the product $\begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ -1 & 6 & 2 \end{bmatrix}$ using the column row-rule, and check

your answer by calculating the product directly.

In exercises 11 and 12, find the inverse of the block diagonal matrix A.

11. (a) $A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 2 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}$

12. (a) $A = \begin{bmatrix} 5 & 1 & 0 & 0 \\ 4 & 1 & 0 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & -3 & 5 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 3 & 7 & 0 & 0 \\ 0 & 0 & 0 & 4 & 9 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$

In exercises 13 and 14, find the inverse of the block upper triangular matrix A.

13. $A=\begin{bmatrix} 2 & 1 & 3 & -6 \\ 1 & 1 & 7 & 4 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 2 & 3 \end{bmatrix}$

14. $A=\begin{bmatrix} -1 & -1 & 2 & 5 \\ 2 & 1 & -3 & 8 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 7 & 2 \end{bmatrix}$

15. Find B_1, given that $\begin{bmatrix} A_1 & B_1 \\ 0 & C_1 \end{bmatrix}\begin{bmatrix} A_2 & B_2 \\ 0 & C_2 \end{bmatrix}=\begin{bmatrix} A_3 & B_3 \\ 0 & C_3 \end{bmatrix}$

and

$$A_1=\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, B_2=\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, C_1=\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$A_2=\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, C_2=\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, B_3=\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

16. Consider the partitioned linear system

$$\left[\begin{array}{cc:cc} 5 & 2 & 2 & 3 \\ 2 & 1 & -3 & 1 \\ \hdashline 1 & 0 & 4 & 1 \\ 0 & 1 & 0 & 2 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \\ \hdashline x_3 \\ x_4 \end{array}\right]=\left[\begin{array}{c} 2 \\ 6 \\ \hdashline 0 \\ 0 \end{array}\right]$$

Solve this system by first expressing it as

$$\begin{bmatrix} A & B \\ I & D \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}=\begin{bmatrix} b \\ 0 \end{bmatrix} \text{ or equivalently, } \begin{array}{l} Au+Bv=b \\ u+Dv=0 \end{array}$$

next solving the second equation for u in terms of v, and then substituting in the first equation. Check your answer by solving the system directly.

Lecture 15

Matrix Factorizations

Matrix Factorization

A factorization of a matrix as a product of two or more matrices is called Matrix factorization.

Uses of Matrix Factorization

Matrix factorizations will appear at a number of key points throughout the course. This lecture focuses on a factorization that lies at the heart of several important computer programs widely used in applications.

***LU* Factorization or LU-decomposition**

LU factorization is a matrix decomposition which writes a matrix as the product of a lower triangular matrix and an upper triangular matrix. This decomposition is used to solve systems of linear equations or calculate the determinant.

Assume A is an $m \times n$ matrix that can be row reduced to echelon form, without row interchanges. Then A can be written in the form A = LU, where L is an $m \times m$ lower triangular matrix with 1's on the diagonal and U is an $m \times n$ echelon form of A. For instance, such a factorization is called LU factorization of A. The matrix L is invertible and is called a unit lower triangular matrix.

$$A = \underset{L}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} \bullet & * & * & * & * \\ 0 & \bullet & * & * & * \\ 0 & 0 & 0 & \bullet & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}$$

***LU* factorization.**

Remarks:

1) If A is the square matrix of order m, then the order of both L and U will also be m.

2) In general, not every square matrix A has an LU-decomposition, nor is an LU-decomposition unique if it exists.

Theorem: If a square matrix *A* can be reduced to row echelon form with no row interchanges, then *A* has an *LU*-decomposition.

Note:

The computational efficiency of the ***LU*** factorization depends on knowing ***L*** and ***U***. The next algorithm shows that the row reduction of ***A*** to an echelon form ***U*** amounts to an ***LU*** factorization because it produces ***L*** with essentially no extra work.

An *LU* Factorization Algorithm

Suppose ***A*** can be reduced to an echelon form ***U*** without row interchanges. Then, since row scaling is not essential, ***A*** can be reduced to ***U*** with only row replacements, adding a multiple of one row to another row below it. In this case, there exist lower triangular elementary matrices $E_1, \ldots, E_p$ such that

$$E_p \ldots E_1 A = U \qquad (1)$$

So $$A = (E_p \ldots E_1)^{-1} U = LU$$

Where $$L = (E_p \ldots E_1)^{-1} \qquad (2)$$

It can be shown that products and inverses of unit lower triangular matrices are also unit-lower triangular. Thus L is unit-lower triangular.

Note that the row operations in (1), which reduce A to U, also reduce the L in (2) to I, because $E_p \ldots E_1 L = (E_p \ldots E_1)(E_p \ldots E_1)^{-1} = I$. This observation is the key to constructing L.

Procedure for finding an *LU*-decomposition

Step 1 Reduce matrix A to row echelon form U without using row interchanges, keeping track of the multipliers used to introduce the leading 1's and the multipliers used to introduce zeros below the leading 1's.

Step 2 In each position along the main diagonal of L, place the reciprocal of the multiplier that introduced the leading 1 in that position in U.

Step 3 In each position below the main diagonal of L, place the negative of the multiplier used to introduce the zero in that position in U.

Step 4 Form the decomposition $A = LU$.

Example 1: Find an LU-decomposition of

$$A = \begin{bmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix}$$

Solution: We will reduce A to a row echelon form U and at each step we will fill in an entry of L in accordance with the four-step procedure above.

$$A = \begin{bmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \qquad \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

* denotes an unknown entry of L.

$$\begin{bmatrix} \boxed{1} & -\frac{1}{3} & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier = \frac{1}{6} \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ \boxed{0} & 2 & 1 \\ \boxed{0} & 8 & 5 \end{bmatrix} \begin{matrix} \leftarrow multiplier = -9 \\ \leftarrow multiplier = -3 \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & * & 0 \\ 3 & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \dfrac{1}{2} \\ 0 & 8 & 5 \end{bmatrix} \leftarrow multiplier = \tfrac{1}{2} \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & \boxed{0} & 1 \end{bmatrix} \leftarrow multiplier = -8 \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & * \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & \boxed{1} \end{bmatrix} \leftarrow multiplier = 1 \qquad L = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix}$$

(No actual operation is performed here since there is already a leading 1 in the third row.)
So

$$A = LU = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

OR

Solution: We will reduce ***A*** to a row echelon form ***U*** and at each step we will fill in an entry of ***L*** in accordance with the four-step procedure above.

$$A = \begin{bmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \qquad \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

* denotes an unknown entry of ***L***.

$$\approx \begin{bmatrix} \dfrac{6}{6} & \dfrac{-2}{6} & \dfrac{0}{6} \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \quad \frac{1}{6}R_1 \qquad \begin{bmatrix} 6 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \dfrac{-1}{3} & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix}$$

$$\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 9-9(1) & -1-9(\frac{-1}{3}) & 1-9(0) \\ 3-3(1) & 7-3(\frac{-1}{3}) & 5-3(0) \end{bmatrix} \begin{matrix} R_2-9R_1 \\ R_3-3R_1 \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & * & 0 \\ 3 & * & * \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 2 & 1 \\ 0 & 8 & 5 \end{bmatrix}$$

$$\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ \frac{0}{2} & \frac{2}{2} & \frac{1}{2} \\ 0 & 8 & 5 \end{bmatrix} \frac{1}{2}R_2 \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & * & * \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 8 & 5 \end{bmatrix}$$

$$\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0-8(0) & 8-8(1) & 5-8(\frac{1}{2}) \end{bmatrix} R_3-8R_2 \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & * \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & \boxed{1} \end{bmatrix} \leftarrow multiplier = 1 \qquad L = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix}$$

So

$$A = LU = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

Example 2 Find an ***LU*** factorization of $A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$

Solution

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ * & * & 1 & 0 & 0 \\ * & * & * & 1 & 0 \\ * & * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of L.

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \begin{matrix} \leftarrow multiplier \frac{1}{2} \\ \\ \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ * & * & 1 & 0 & 0 \\ * & * & * & 1 & 0 \\ * & * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier -6 \\ \leftarrow multiplier -2 \\ \leftarrow multiplier -4 \\ \leftarrow multiplier\, 6 \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 & 0 \\ 2 & * & 1 & 0 & 0 \\ 4 & * & * & 1 & 0 \\ -6 & * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix} \leftarrow multiplier \frac{1}{3} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & * & 1 & 0 & 0 \\ 4 & * & * & 1 & 0 \\ -6 & * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix} \begin{matrix} \leftarrow multiplier\, 3 \\ \leftarrow multiplier -6 \\ \leftarrow multiplier\, 9 \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & * & 1 & 0 \\ -6 & -9 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 10 \end{bmatrix} \leftarrow multiplier -\frac{1}{5} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & 0 & -5 & 0 \\ -6 & -9 & 0 & * & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \leftarrow multiplier -10 \qquad L = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & 0 & -5 & 0 \\ -6 & -9 & 0 & 10 & 1 \end{bmatrix}$$

Thus, we have constructed the ***LU***-decomposition

$$A = LU = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & 0 & -5 & 0 \\ -6 & -9 & 0 & 10 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Example 3 Find ***LU***–decomposition of $A = \begin{bmatrix} 6 & -2 & -4 & 4 \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix}$

Solution

$$A = \begin{bmatrix} 6 & -2 & -4 & 4 \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of ***L***.

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix} \begin{matrix} \leftarrow multiplier\ \frac{1}{6} \\ \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & -2 & -4 & -1 \\ 0 & 4 & 13 & 0 \\ 0 & -2 & -14 & 11 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier -3 \\ \leftarrow multiplier\ 12 \\ \leftarrow multiplier\ 6 \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -12 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 4 & 13 & 0 \\ 0 & -2 & -14 & 11 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier -\frac{1}{2} \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 5 & -2 \\ 0 & 0 & -10 & 12 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier -4 \\ \leftarrow multiplier\, 2 \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 1 & 0 \\ -6 & -2 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & -10 & 12 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier\, \frac{1}{5} \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 8 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier\, 10 \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier\, \frac{1}{8} \\ \\ \end{matrix} \qquad L = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 8 \end{bmatrix}$$

$$Thus\;\; A = LU = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 8 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Example 4 Find an ***LU*** factorization of $A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$

Solution

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of ***L***.

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \begin{matrix} \leftarrow multiplier\ \frac{1}{2} \\ \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier\ 4 \\ \leftarrow multiplier\ -2 \\ \leftarrow multiplier\ 6 \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 1 & 0 & 0 \\ 2 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier\ \frac{1}{3} \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier\ 9 \\ \leftarrow multiplier\ -12 \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & -9 & 1 & 0 \\ -6 & 12 & * & 1 \end{bmatrix}$$

$$U=\begin{bmatrix}1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1\\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1\\ 0 & 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 1 & \frac{7}{4}\end{bmatrix}\leftarrow multiplier\ \frac{1}{4} \qquad L=\begin{bmatrix}2 & 0 & 0 & 0\\ -4 & 3 & 0 & 0\\ 2 & -9 & 1 & 0\\ -6 & 12 & 0 & 4\end{bmatrix}$$

Thus $A=LU=\begin{bmatrix}2 & 0 & 0 & 0\\ -4 & 3 & 0 & 0\\ 2 & -9 & 1 & 0\\ -6 & 12 & 0 & 4\end{bmatrix}\begin{bmatrix}1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1\\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1\\ 0 & 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 1 & \frac{7}{4}\end{bmatrix}$

Matrix Inversion by *LU*-Decomposition

Many of the best algorithms for inverting matrices use ***LU***-decomposition. To understand how this can be done, let ***A*** be an invertible $n\times n$ matrix, let $A^{-1}=\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}$ be its unknown inverse partitioned into column vectors, and let $I=\begin{bmatrix}e_1 & e_2 & \cdots & e_n\end{bmatrix}$ be then $n\times n$ identity matrix partitioned into column vectors. The matrix equation $AA^{-1}=I$ can be expressed as

$$A\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}=\begin{bmatrix}e_1 & e_2 & \cdots & e_n\end{bmatrix}$$
$$\begin{bmatrix}Ax_1 & Ax_2 & \cdots & Ax_n\end{bmatrix}=\begin{bmatrix}e_1 & e_2 & \cdots & e_n\end{bmatrix}$$

which tells us that the unknown column vectors of A^{-1} can be obtained by solving the ***n***-linear systems

$$Ax_1=e_1,\ \ Ax_2=e_2,\cdots,\ \ Ax_n=e_n \qquad (1^*)$$

As discussed above, this can be done by finding an ***LU***-decomposition of ***A***, and then using that decomposition to solve each of the n systems in (1*).

Solving Linear System by LU-Factorization

When ***A* =*LU***, the equation ***Ax* = *b*** can be written as ***L* (*Ux*) = *b***. Writing ***y*** for ***Ux***, we can find ***x*** by solving the pair of equations; ***Ly* = *b*** and ***Ux* = *y*** (2*)

First solve ***Ly* = *b*** for ***y*** and then solve ***Ux* = *y*** for ***x***. Each equation is easy to solve because ***L*** and ***U*** are triangular.

Procedure

Step 1 Rewrite the system ***A x = b*** as ***LU x = b*** (3*)

Step 2 Define a new unknown ***y*** by letting ***U x = y*** (4*)
And rewrite (3*) as ***L y = b***

Step 3 Solve the system ***L y = b*** for the unknown ***y***.
Step 4 Substitute the known vector ***y*** into (4*) and solve for ***x***.

This procedure is called the method of ***LU***-Decomposition.

Although ***LU***-Decomposition converts the problem of solving the single system ***A x = b*** into the problem of solving the two systems, ***L y = b*** and ***U x = y***, these systems are easy to solve because their co-efficient matrices are triangular.

Example 5 Solve the given system (***Ax =b***) by ***LU***-Decomposition

$$\begin{aligned} 2x_1 + 6x_2 + 2x_3 &= 2 \\ -3x_1 - 8x_2 \qquad &= 2 \\ 4x_1 + 9x_2 + 2x_3 &= 3 \end{aligned} \qquad (1)$$

Solution We express the system (1) in matrix form:

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$$

$$\boldsymbol{A} \qquad \boldsymbol{x} = \boldsymbol{b}$$

We derive an ***LU***-decomposition of ***A***.

$$A = \begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \leftarrow multiplier\,\frac{1}{2} \qquad \begin{bmatrix} 2 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & -3 & -2 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier\,3 \\ \leftarrow multiplier - 4 \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow multiplier\,3 \qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow multiplier\,\frac{1}{7} \qquad\qquad L = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix}$$

Thus

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \qquad (2)$$

A* = *L* *U

From (2) we can rewrite this system as

$$\begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \qquad (3)$$

L* *U* *x* = *b

As specified in Step 2 above, let us define y_1, y_2 and y_3 by the equation

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \qquad (4)$$

U* *x* = *y

which allows us to rewrite (3) as

$$\begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \qquad (5)$$

L* *y* = *b

or equivalently, as

$$\begin{aligned} 2y_1 \qquad\qquad &= 2 \\ -3y_1 + y_2 \qquad &= 2 \\ 4y_1 - 3y_2 + 7y_3 &= 3 \end{aligned}$$

This system can be solved by a procedure that is similar to back substitution, except that we solve the equations from the top down instead of from the bottom up. This procedure, called **forward substitution**, yields

$y_1 = 1$, $y_2 = 5$, $y_3 = 2$.

As indicated in Step 4 above, we substitute these values into (4), which yields the linear system

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix}$$

$$x_1 + 3x_2 + x_3 = 1$$

or equivalently, $$x_2 + 3x_3 = 5$$

$$x_3 = 2$$

Solving this system by back substitution yields $x_1 = 2$, $x_2 = -1$, $x_3 = 2$

Example 6 It can be verified that

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

Use this ***LU*** factorization of ***A*** to solve ***Ax = b***, where $b = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$

Solution The solution of ***Ly = b*** needs only 6 multiplications and 6 additions, because the arithmetic takes place only in column 5. (The zeros below each pivot in ***L*** are created automatically by our choice of row operations.)

$$[L \quad b] = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = [I \quad y]$$

Then, for ***Ux = y***, the "backwards" phase of row reduction requires 4 divisions, 6 multiplications, and 6 additions. (For instance, creating the zeros in column 4 of **[*U* *y*]** requires 1 division in row 4 and 3 multiplication – addition pairs to add multiples of row 4 to the rows above.)

$$[U \quad y] = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \quad x = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

To find ***x*** requires 28 arithmetic operations, or "flops" (floating point operations), excluding the cost of finding ***L*** and ***U***. In contrast, row reduction of **[*A* *b*]** to **[*I* *x*]** takes 62 operations.

Numerical Notes

The following operation counts apply to an $n \times n$ dense matrix ***A*** (with most entries nonzero) for *n* moderately large, say, $n \geq 30$.

1. Computing an ***LU*** factorization of ***A*** takes about $2n^3/3$ flops (about the same as row reducing [***A*** ***b***]), whereas finding $\boldsymbol{A^{-1}}$ requires about $2n^3$ flops.
2. Solving ***Ly = b*** and ***Ux = y*** requires about $2n^2$ flops, because $n \times n$ triangular system can be solved in about n^2 flops.
3. Multiplication of ***b*** by $\boldsymbol{A^{-1}}$ also requires about $2n^2$ flops, but the result may not be as accurate as that obtained from ***L*** and ***U*** (because of round off error when computing both $\boldsymbol{A^{-1}}$ and $\boldsymbol{A^{-1}b}$).
4. If ***A*** is sparse (with mostly zero entries), then ***L*** and ***U*** may be sparse, too, whereas $\boldsymbol{A^{-1}}$ is likely to be dense. In this case, a solution of ***Ax = b*** with an ***LU*** factorization is much faster than using $\boldsymbol{A^{-1}}$.

Example 7(Gaussian Elimination Performed as an LU-Decomposition)
In Example 5 we showed how to solve the linear system

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \qquad (6)$$

using an ***LU***-decomposition of the coefficient matrix, but we did not discuss how the factorization was derived. In the course of solving the system we obtained the intermediate vector $\boldsymbol{y} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix}$ by using forward substitution to solve system (5).

We will now use the procedure discussed above to find both the ***LU***-decomposition and the vector ***y*** by row operations on the augmented matrix for (6).

$$\left[A|b\right] = \left[\begin{array}{ccc:c} 2 & 6 & 2 & 2 \\ -3 & -8 & 0 & 2 \\ 4 & 9 & 2 & 3 \end{array}\right] \qquad \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix} = L \quad (* = unknown \quad entries)$$

$$\left[\begin{array}{ccc:c} 1 & 3 & 1 & 1 \\ -3 & -8 & 0 & 2 \\ 4 & 9 & 2 & 3 \end{array}\right] \qquad \begin{bmatrix} 2 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

$$\left[\begin{array}{ccc:c} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & -3 & -2 & -1 \end{array}\right] \qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & * & 0 \\ 4 & * & * \end{bmatrix}$$

$$\left[\begin{array}{ccc:c} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 7 & 14 \end{array}\right] \qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & * \end{bmatrix}$$

$$\left[U\,\middle|\,y\right]=\left[\begin{array}{ccc|c}1&3&1&1\\0&1&3&5\\0&0&1&2\end{array}\right]\qquad\begin{bmatrix}2&0&0\\-3&1&0\\4&-3&7\end{bmatrix}=L$$

These results agree with those in Example 5, so we have found an ***LU***-decomposition of the coefficient matrix and simultaneously have completed the forward substitution required to find ***y***.

All that remains to solve the given system is to solve the system ***Ux = y*** by back substitution. The computations were performed in Example5.

A Matrix Factorization in Electrical Engineering

Matrix factorization is intimately related to the problem of constructing an electrical network with specified properties. The following discussion gives just a glimpse of the connection between factorization and circuit design.

Suppose the box in below Figure represents some sort of electric circuit, with an input and output. Record the input voltage and current by $\begin{bmatrix}v_1\\i_1\end{bmatrix}$ (with voltage ***v*** in volts and current ***i*** in amps), and record the output voltage and current by $\begin{bmatrix}v_2\\i_2\end{bmatrix}$. Frequently, the transformation $\begin{bmatrix}v_1\\i_1\end{bmatrix}\to\begin{bmatrix}v_2\\i_2\end{bmatrix}$ is linear. That is, there is a matrix A, called the transfer matrix, such that $\begin{bmatrix}v_2\\i_2\end{bmatrix}=A\begin{bmatrix}v_1\\i_1\end{bmatrix}$

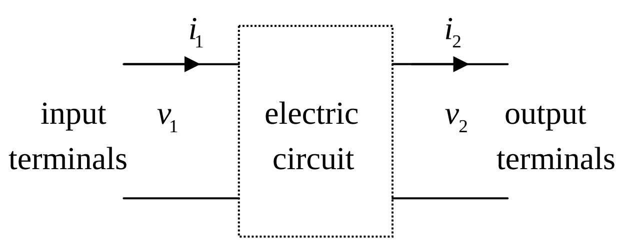


Figure A circuit with input and output terminals.

Above Figure shows a ladder network, where two circuits (there could be more) are connected in series, so that the output of one circuit becomes the input of the next circuit. The left circuit in Figure is called a series circuit, with resistance R_1 (in ohms);

The right circuit is a shunt circuit, with resistance R_2. Using Ohm's law and Kirchhoff's laws, one can show that the transfer matrices of the series and shunt circuits, respectively, are

$$\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} \quad and \quad \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}$$

Transfer matrix of series circuit — Transfer matrix of shunt circuit

Example 8

a) Compute the transfer matrix of the ladder network in above Figure .

b) Design a ladder network whose transfer matrix is $\begin{bmatrix} 1 & -8 \\ -0\cdot 5 & 5 \end{bmatrix}$.

Solution

a) Let $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$ be the transfer matrices of the series of the series and shunt circuits, respectively. Then an input vector $\boldsymbol{x}$ is transformed first into $\boldsymbol{A_1x}$ and then into $\boldsymbol{A_2 (A_1x)}$. The series connection of the circuits corresponds to composition of linear transformations; and the transfer matrix of the ladder network in (note the order)

$$A_2A_1 = \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix} \quad (6)$$

b) We seek to factor the matrix $\begin{bmatrix} 1 & -8 \\ -0\cdot 5 & 5 \end{bmatrix}$ into the product of transfer matrices, such as in (6). So we look for R_1 and R_2 to satisfy

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -0\cdot 5 & 5 \end{bmatrix}$$

From the (1, 2) – entries, R_1 = 8 ohms, and from the (2, 1) – entries, $1/R_2$ = 0.5 ohm and R_2 = 1/0.5 = 2 ohms. With these values, the network has the desired transfer matrix.

Note:

A network transfer matrix summarizes the input-output behavior (“Design specifications”) of the network without reference to the interior circuits. To physically build a network with specified properties, an engineer first determines if such a network can be constructed (or realized). Then the engineer tries to factor the transfer matrix into matrices corresponding to smaller circuits that perhaps are already manufactured and ready for assembly. In the common case of alternating current, the entries in the transfer matrix are usually rational complex-valued functions. A standard problem is to find a minimal realization that uses the smallest number of electrical components.

Exercises

Find an LU factorization of the matrices in exercises 1 to 8.

1. $\begin{bmatrix} 2 & 5 \\ -3 & -4 \end{bmatrix}$

2. $\begin{bmatrix} 3 & -1 & 2 \\ -3 & -2 & 10 \\ 9 & -5 & 6 \end{bmatrix}$

3. $\begin{bmatrix} 3 & -6 & 3 \\ 6 & -7 & 2 \\ -1 & 7 & 0 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 3 & -5 & -3 \\ -1 & -5 & 8 & 4 \\ 4 & 2 & -5 & -7 \\ -2 & -4 & 7 & 5 \end{bmatrix}$

5. $\begin{bmatrix} 2 & -4 & 4 & -2 \\ 6 & -9 & 7 & -3 \\ -1 & -4 & 8 & 0 \end{bmatrix}$

6. $\begin{bmatrix} 2 & -6 & 6 \\ -4 & 5 & -7 \\ 3 & 5 & -1 \\ -6 & 4 & -8 \\ 8 & -3 & 9 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 4 & -1 & 5 \\ 3 & 7 & -2 & 9 \\ -2 & -3 & 1 & -4 \\ -1 & 6 & -1 & 7 \end{bmatrix}$

8. $\begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$

Solve the equation ***Ax = b*** by using ***LU***-factorization.

9. $A = \begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix}, b = \begin{bmatrix} -7 \\ 5 \\ 2 \end{bmatrix}$

10. $A = \begin{bmatrix} 4 & 3 & -5 \\ -4 & -5 & 7 \\ 8 & 6 & -8 \end{bmatrix}, b = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$

11. $A = \begin{bmatrix} 2 & -1 & 2 \\ -6 & 0 & -2 \\ 8 & -1 & 5 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$

12. $A = \begin{bmatrix} 2 & -2 & 4 \\ 1 & -3 & 1 \\ 3 & 7 & 5 \end{bmatrix}, b = \begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix}$

13. $A = \begin{bmatrix} 1 & -2 & -4 & -3 \\ 2 & -7 & -7 & -6 \\ -1 & 2 & 6 & 4 \\ -4 & -1 & 9 & 8 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 7 \\ 0 \\ 3 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 3 & 4 & 0 \\ -3 & -6 & -7 & 2 \\ 3 & 3 & 0 & -4 \\ -5 & -3 & 2 & 9 \end{bmatrix}, b = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$

Lecture 16

Iterative Solutions of Linear Systems

Consistent linear systems are solved in one of two ways by direct calculation (matrix factorization) or by an iterative procedure that generates a sequence of vectors that approach the exact solution. When the coefficient matrix is large and sparse (with a high proportion of zero entries), iterative algorithms can be more rapid than direct methods and can require less computer memory. Also, an iterative process may be stopped as soon as an approximate solution is sufficiently accurate for practical work.

General Framework for an Iterative Solution of *Ax = b:*

Throughout the section, ***A*** is an invertible matrix. The goal of an iterative algorithm is to produce a sequence of vectors,

$$x^{(0)}, x^{(1)}, \ldots, x^{(k)}, \ldots$$

that converges to the unique solution say x^* of ***Ax = b***, in the sense that the entries in $x^{(k)}$ are as close as desired to the corresponding entries in x^* for all *k* sufficiently large.

To describe a recursion algorithm that produces $x^{(k+1)}$ from $x^{(k)}$, we write ***A = M – N*** for suitable matrices ***M*** and ***N,*** and then we rewrite the equation ***Ax = b*** as ***Mx – Nx = b*** and

$$Mx = Nx + b$$

If a sequence { $x^{(k)}$ } satisfies

$$\boxed{Mx^{(k+1)} = Nx^{(k)} + b \qquad (k = 0,1,\ldots)} \qquad (1)$$

and if the sequence converges to some vector x^*, then it can be shown that $Ax^* = b.$[The vector on the left in (1) approaches Mx^*, while the vector on the right in (1) approaches $Nx^* + b.$ This implies that $Mx^* = Nx^* + b$ and $Ax^* = b.$

For $x^{(k+1)}$ to be uniquely specified in (1), ***M*** must be invertible. Also, ***M*** should be chosen so that $x^{(k+1)}$ is easy to calculate. There are two iterative methods below to illustrate two simple choices for ***M***.

1) Jacobi's Method:

This method assumes that the diagonal entries of ***A*** are all nonzero.
Choosing ***M*** as the diagonal matrix formed from the diagonal entries of ***A***. So next ***N = M – A***,

$$\therefore (1) \Rightarrow Mx^{(k+1)} = (M - A)x^{(k)} + b \qquad (k = 0,1,\ldots)$$

For simplicity, we take the zero vector as $x^{(0)}$ as the initial approximation.

Example 1:

Apply Jacobi's method to the system

$$\begin{aligned} 10x_1 + x_2 - x_3 &= 18 \\ x_1 + 15x_2 + x_3 &= -12 \\ -x_1 + x_2 + 20x_3 &= 17 \end{aligned} \qquad (2)$$

Take $\mathbf{x}^{(0)} = (0, 0, 0)$ as an initial approximation to the solution, and use six iterations (that is, compute $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(6)}$).

Solution:

For some $\mathbf{k}$, let $\mathbf{x}^{(k)} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (x_1, x_2, x_3)$ and $\mathbf{x}^{(k+1)} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = (y_1, y_2, y_3)$

Firstly we construct $\boldsymbol{M}$ and $\boldsymbol{N}$ from $\boldsymbol{A}$.

Here

$$A = \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix}$$

Its diagonal entries will give

$$M = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 20 \end{bmatrix} \text{ and}$$

$$N = M - A = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 20 \end{bmatrix} - \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

Now the recursion: $Mx^{(k+1)} = (M - A)x^{(k)} + b \qquad (here\ k = 0, 1, \ldots 6)$

implies

$$\begin{bmatrix} 10 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 20 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 10y_1 \\ 15y_2 \\ 20y_3 \end{bmatrix} = \begin{bmatrix} 0x_1 - 1x_2 + 1x_3 \\ -1x_1 + 0x_2 - x_3 \\ 1x_1 - 1x_2 + 0x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 10y_1 \\ 15y_2 \\ 20y_3 \end{bmatrix} = \begin{bmatrix} 0x_1 - 1x_2 + 1x_3 + 18 \\ -1x_1 + 0x_2 - x_3 - 12 \\ 1x_1 - 1x_2 + 0x_3 + 17 \end{bmatrix}$$

Comparing the corresponding entries on both sides, we have

$$10y_1 = -x_2 + x_3 + 18$$

$15y_2 = -x_1 \quad - x_3 - 12$
$20y_3 = x_1 - x_2 \quad + 17$

And

$$y_1 = (-x_2 + x_3 + 18)/10$$
$$y_2 = (-x_1 - x_3 - 12)/15 \qquad (3)$$
$$y_3 = (x_1 - x_2 + 17)/20$$

1st Iteration:

For $\boldsymbol{k} = 0$, put $\boldsymbol{x}^{(0)} = (x_1, x_2, x_3) = (0, 0, 0)$ in (3) and compute

$$\boldsymbol{x}^{(1)} = (y_1, y_2, y_3) = (18/10, -12/15, 17/20) = (1.8, -0.8, 0.85)$$

2nd Iteration:

For k = 1, put $\boldsymbol{x}^{(1)} = (1.8, -.8, .85)$

$$y_1 = [-(-0.8) + (0.85) + 18]/10 = 1.965$$
$$y_2 = [-(1.8) - (0.85) - 12]/15 = -0.9767$$
$$y_3 = [(1.8) - (-0.8) + 17]/20 = 0.98$$

Thus $\boldsymbol{x}^{(2)} = (1.965, -.9767, .98)$.

The entries in $\boldsymbol{x}^{(2)}$ are used on the right in (3) to compute the entries in $\boldsymbol{x}^{(3)}$, and so on. Here are the results, with calculations using **MATLAB** and results reported to four decimal places:

$$\begin{matrix} x^{(0)} & x^{(1)} & x^{(2)} & x^{(3)} & x^{(4)} & x^{(5)} & x^{(6)} \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1.8 \\ -.8 \\ .85 \end{bmatrix} & \begin{bmatrix} 1.965 \\ -.9767 \\ .98 \end{bmatrix} & \begin{bmatrix} 1.9957 \\ -.9963 \\ .9971 \end{bmatrix} & \begin{bmatrix} 1.9993 \\ -.9995 \\ .9996 \end{bmatrix} & \begin{bmatrix} 1.9999 \\ -.9999 \\ .9999 \end{bmatrix} & \begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix} \end{matrix}$$

If we decide to stop when the entries in $x^{(k)}$ and $x^{(k-1)}$ differ by less than .001, then we need five iterations ($k = 5$).

Alternative Approach:

If we express the above system as

$$10x_1 + x_2 - x_3 = 18 \Rightarrow x_1 = \frac{18 - x_2 + x_3}{10}$$
$$x_1 + 15x_2 + x_3 = -12 \Rightarrow x_2 = \frac{-12 - x_1 - x_3}{15}$$
$$-x_1 + x_2 + 20x_3 = 17 \Rightarrow x_3 = \frac{17 + x_1 - x_2}{20}$$

$\therefore$ the equivalent system is

$$x_1 = \frac{18 - x_2 + x_3}{10}$$
$$x_2 = \frac{-12 - x_1 - x_3}{15}$$
$$x_3 = \frac{17 + x_1 - x_2}{20}$$

Now put $(\boldsymbol{x_1},\boldsymbol{x_2},\boldsymbol{x_3})$=(0,0,0)= $\boldsymbol{x}^{(0)}$ in the RHS to have

$\boldsymbol{x_1}$=(18-0+0)/10 = 1.80
$\boldsymbol{x_2}$= (-12-0-0)/15 = -0.80
$\boldsymbol{x_3}$ = (17+0-0)/20 = 0.85

Which gives $\boldsymbol{x}^{(1)}$ = (1.80,-0.80,0.85) ----put this again on RHS of the equivalent system to get

$\boldsymbol{x_1}$=(18+0.80+0.85)/10 =1.965
$\boldsymbol{x_2}$= (-12-1.80-0.85)/15 = -0.9767
$\boldsymbol{x_3}$= (17+1.80+0.80)/20 = 0.98

So in the similar fashion, we can get the next approximate solutions: $\boldsymbol{x}^{(3)}$,$\boldsymbol{x}^{(4)}$,$\boldsymbol{x}^{(5)}$ and $\boldsymbol{x}^{(6)}$
Next example will be solved by following this approach.

Example 2:

Use Jacobi iteration to approximate the solution of the system

$$20\,x_1 + x_2 - x_3 = 17$$
$$x_1 - 10x_2 + x_3 = 13$$
$$-x_1 + x_2 + 10x_3 = 18$$

Stop the process when the entries in two successive iterations are the same when rounded to four decimal places.

Solution:

As required for Jacobi iteration, we begin by solving the first equation for x_1, the second for x_2, and the third for x_3. This yields

$$x_1 = \frac{17}{20} - \frac{1}{20}x_2 + \frac{1}{20}x_3 = 0.85 - 0.05x_2 + 0.05x_3$$
$$x_2 = -\frac{13}{10} + \frac{1}{10}x_1 + \frac{1}{10}x_3 = -1.3 + 0.1x_1 + 0.1x_3 \qquad (4)$$
$$x_3 = \frac{18}{10} + \frac{1}{10}x_1 - \frac{1}{10}x_2 = 1.8 + 0.1x_1 - 0.1x_2$$

which we can write in matrix form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -0.05 & 0.05 \\ 0.1 & 0 & 0.1 \\ 0.1 & -0.1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix} \qquad (5)$$

Since we have no special information about the solution, we will take the initial approximation to be $x_1 = x_2 = x_3 = 0$. To obtain the first iterate, we substitute these values into the right side of (5). This yields

$$y_1 = \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix}$$

To obtain the second iterate, we substitute the entries of y_1 into the right side of (5). This yields

$$y_2 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -0.05 & 0.05 \\ 0.1 & 0 & 0.1 \\ 0.1 & -0.1 & 0 \end{bmatrix} \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix} + \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1.005 \\ -1.035 \\ 2.015 \end{bmatrix}$$

Repeating this process until two successive iterations match to four decimal places yields the results in the following table.

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7
x_1	0	0.8500	1.0050	1.0025	1.0001	1.0000	1.0000	1.0000
x_2	0	-1.3000	-1.0350	-0.9980	-0.9994	-1.0000	-1.0000	-1.0000
x_3	0	1.8000	2.0150	2.0040	2.0000	1.9999	2.0000	2.0000

The Gauss-Seidel Method:

This method uses the recursion (1) with *M* the lower triangular part of *A*. That is, *M* has the same entries as *A* on the diagonal and below, and *M* has zeros above the diagonal. See Fig. 1. As in Jacobi's method, the diagonal entries of *A* must be nonzero in order for *M* to be invertible.

$$A = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \qquad M = \begin{bmatrix} * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & * & 0 \\ * & * & * & * & * \end{bmatrix}$$

Figure 01:The Lower Triangular Part of A

Example 3:

Apply the Gauss – Seidel method to the system in Example 1 with $x^{(0)} = (0,0,0)$ and six iterations.

$$\begin{aligned} 10x_1 + x_2 - x_3 &= 18 \\ x_1 + 15x_2 + x_3 &= -12 \\ -x_1 + x_2 + 20x_3 &= 17 \end{aligned} \qquad (6)$$

Solution:

For some ***k***, let $\mathbf{x}^{(\mathbf{k})} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (x_1, x_2, x_3)$ and $\mathbf{x}^{(\mathbf{k+1})} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = (y_1, y_2, y_3)$

Again, firstly we construct matrices ***M*** and ***N*** from the coefficient matrix ***A.***

Here $A = \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix}$

Since matrix M is constructed by

1) taking the values along the diagonal and below the diagonal of coefficient matrix A.
2) putting the zeros above the diagonal at upper trianular position.

So

$$M = \begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix}$$

Now,

$$N = M - A = \begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix} - \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Now the recursion: $Mx^{(k+1)} = (M - A)x^{(k)} + b \qquad (here\ k = 0,1,...6)$

implies

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 10y_1 & 0 & 0 \\ y_1 & 15y_2 & 0 \\ -y_1 & y_2 & 20y_3 \end{bmatrix} = \begin{bmatrix} -x_2 + x_3 \\ -x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 10y_1 & 0 & 0 \\ y_1 & 15y_2 & 0 \\ -y_1 & y_2 & 20y_3 \end{bmatrix} = \begin{bmatrix} -x_2 + x_3 + 18 \\ -x_3 - 12 \\ 0 + 17 \end{bmatrix}$$

Comparing the corresponding entries on both sides, we have

$$10y_1 = -x_2 + x_3 + 18$$
$$y_1 + 15y_2 = -x_3 - 12$$
$$-y_1 + y_2 + 20y_3 = 17$$

This further implies as

$$\left.\begin{aligned} 10y_1 = -x_1 - x_3 + 18 &\Rightarrow y_1 = (-x_2 + x_3 + 18)/10 \\ y_1 + 15y_2 = -x_3 - 12 &\Rightarrow y_2 = (-y_1 - x_3 - 12)/15 \\ -y_1 + y_2 + 20y_3 = 17 &\Rightarrow y_3 = (y_1 - y_2 + 17)/20 \end{aligned}\right\} \text{-----------(7)}$$

Another way to view (7) is to solve each equation in (6) for x_1, x_2, x_3, respectively and regard the highlighted *x's* as the values:

$$\begin{aligned} x_1 &= (-x_2 + x_3 + 18)/10 \\ x_2 &= (-x_1 - x_3 - 12)/15 \\ x_3 &= (x_1 - x_2 + 17)/20 \end{aligned} \qquad (8)$$

Use the first equation to calculate the new x_1 [called y_1 in (7)] from x_2 and x_3. Then use this new x_1 along with x_3 in the second equation to compute the new x_2. Finally, in the third equation, use the new values for x_1 and x_2 to compute x_3. In this way, the latest information about the variables is used to compute new values. [A computer program would use statements corresponding to the equations in (8).]

From $x^{(0)} = (0, 0, 0)$, we obtain

$$x_1 = [-(0) + (0) + 18]/10 = 1.8$$

$$x_2 = [-(1.8) - (0) - 12]/15 = -.92$$

$$x_3 = [+(1.8) - (-.92) + 17]/20 = .986$$

Thus $x^{(1)} = (1.8, -.92, .986)$. The entries in $x^{(1)}$ are used in (8) to produce $x^{(2)}$ and so on. Here are the **MATLAB** calculations reported to four decimal places:

$$\begin{matrix} x^{(0)} & x^{(1)} & x^{(2)} & x^{(3)} & x^{(4)} & x^{(5)} & x^{(6)} \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1.8 \\ -.92 \\ .986 \end{bmatrix} & \begin{bmatrix} 1.9906 \\ -.9984 \\ .9995 \end{bmatrix} & \begin{bmatrix} 1.9998 \\ -.9999 \\ 1.0000 \end{bmatrix} & \begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix} & \begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix} & \begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix} \end{matrix}$$

Observe that when *k* is *4*, the entries in $x^{(4)}$ and $x^{(k-1)}$ differ by less than *.001*. The values in $x^{(6)}$ in this case happen to be accurate to eight decimal places.

Alternative Approach:

If we express the above system as

$10x_1 + x_2 - x_3 = 18 \Rightarrow x_1 = (-x_2 + x_3 + 18)/10$ ------(a)

$x_1 + 15x_2 + x_3 = -12 \Rightarrow x_2 = (-x_1 - x_3 - 12)/15$------(b)

$-x_1 + x_2 + 20x_3 = 17 \Rightarrow x_3 = (x_1 - x_2 + 17)/20$ ------(c)

Ist Iteration:

Put $x_2 = x_3 = 0$ in (a)

x_1= 18/10 =1.80

Put x_1=1.80 and x_3=0 in (b)

x_2 = (-1.80-0-12)/15= -0.92

Put x_1=1.80, x_2 = -0.92 in (c)

x_3=(1.80+0.92+17)/20 = 0.9863

So, $\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 1.8 \\ -.92 \\ .986 \end{bmatrix} = x^{(1)}$

2nd iteration:

Put x_2 = -0.92, x3=0.9863 in (a)

x_1=(0.92+0.9863+18)/10 = 1.9906

Put x_1=1.9906(from 2nd iteration) and x_3=0.9863(from 1st iteration) in (b)

x_2=(-1.9906-0.9863-12)/15 = -0.9984

Put x_1=1.9906, $x2$=-0.9984(both from 2nd iteration) in (c)

x_3=(1.9906+0.9984+17)/20 = 0.9995

So, $\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 1.9906 \\ -.9984 \\ .9995 \end{bmatrix} = x^{(2)}$

So in the similar fashion, we can get the next approximate solutions: $x^{(3)}, x^{(4)}, x^{(5)}$ and $x^{(6)}$
Next example will be solved by following this approach.

Example 4:

Use Gauss-Seidel to approximate the solution of the linear system in example 2 to four decimal places.

Solution:

As before, we will take $x_1 = x_2 = x_3 = 0$ as the initial approximation. First we will substitute $x_2 = 0$ and $x_3 = 0$ into the right side of the first equation of (4) to obtain the new x_1, then we will substitute $x_3 = 0$ and the new x_1 into the right side of the second equation to obtain the new x_2, and finally we will substitute the new x_1 and new x_2 into the right side of the third equation to obtain the new x_3. The computations are as follows:

$$x_1 = 0.85 - (0.05)(0) + (0.05)(0) = 0.85$$
$$x_2 = -1.3 + (0.1)(0.85) + (0.1)(0) = -1.215$$
$$x_3 = 1.8 + (0.1)(0.85) - (0.1)(-1.215) = 2.0065$$

Thus, the first Gauss-Seidel iterate is

$$y_1 = \begin{bmatrix} 0.8500 \\ -1.2150 \\ 2.0065 \end{bmatrix}$$

Similarly, the computations for second iterate are

$$x_1 = 0.85 - (0.05)(-1.215) + (0.05)(2.0065) = 1.011075$$
$$x_2 = -1.3 + (0.1)(1.011075) + (0.1)(2.0065) = -0.9982425$$
$$x_3 = 1.8 + (0.1)(1.011075) - (0.1)(-0.9982425) = 2.00093175$$

Thus, the second Gauss-Seidel iterate to four decimal places is

$$y_2 \approx \begin{bmatrix} 1.0111 \\ -0.9982 \\ 2.0009 \end{bmatrix}$$

The following table shows the first four Gauss-Seidel iterates to four decimal places. Comparing both tables, we see that the Gauss-Seidel method produced the solution to four decimal places in four iterations, whereas the Jacobi method required six.

	y_0	y_1	y_2	y_3	y_4
x_1	0	0.8500	1.0111	1.0000	1.0000
x_2	0	-1.2150	-0.9982	-0.9999	-1.0000
x_3	0	2.0065	2.0009	2.0000	2.0000

Comparison of Jacobi's and Gauss-Seidel method:

There exist examples where Jacobi's method is faster than the Gauss-Seidel method, but usually a Gauss-Seidel sequence converges faster (means to say iterative solution approaches to the unique solution), as in Example 2. (If parallel processing is available, Jacobi might be faster because the entries in $x^{(k)}$ can be computed simultaneously.) There are also examples where one or both methods fail to produce a convergent sequence, and other examples where a sequence is convergent, but converges too slowly for practical use.

Condtion for the Convergence of both Iterative Mthods:

Fortunately, there is a simple condition that guarantees (but is not essential for) the convergence of both Jacobi and Gauss-Seidel sequences. This condition is often satisfied, for instance, in large-scale systems that can occur during numerical solutions of partial differential equations (such as Laplace's equation for steady-state heat flow).

An $n \times n$ matrix A is said to be ***strictly diagonally dominant*** *if the absolute value of each diagonal entry exceeds the sum of the absolute values of the other entries in the same row.*

In this case it can be shown that *A* is invertible and that both the Jacobi and Gauss-Seidel sequences converge to the unique solution of *Ax = b*, for any initial $x^{(0)}$. (The speed of the convergence depends on how much the diagonal entries dominate the corresponding row sums.)

The coefficient matrices in Examples 1 and 2 are strictly diagonally dominant, but the following matrix is not. Examine each row:

$$\begin{bmatrix} -6 & 2 & -3 \\ 1 & 4 & -2 \\ 3 & -5 & 8 \end{bmatrix} \qquad \begin{aligned} |-6| &> |2| + |-3| \\ |4| &> |1| + |-2| \\ |8| &= |3| + |-5| \end{aligned}$$

The problem lies in the third row, because $|8|$ is not *larger* than the sum of the magnitudes of the other entries.

Note:

The practice problem below suggests a TRICK(rearrangement of the system of equations) that sometimes works when a system is not strictly diagonally dominant.

Example 5:

Show that the Gauss-Seidel method will produce a sequence converging to the solution of the following systems, provided the equations are arranged properly:

$$x_1 - 3x_2 + x_3 = -2$$
$$-6x_1 + 4x_2 + 11x_3 = 1$$
$$5x_1 - 2x_2 - 2x_3 = 9$$

Solution:

The system is not strictly diagonally dominant, as for the 1st row

$$|coefficient\ of\ x_1| < |coefficient\ of\ x_2| + |coefficient\ of\ x_3|$$

$$or \quad |1| < |-3| + |1|$$

so neither Jacobi nor Gauss- Seidel is guaranteed to work. In fact, both iterative methods produce sequences that fail to converge, even though the system has the unique solution $x_1 = 3, x_2 = 2, x_3 = 1$. However, the equations can be rearranged as

$$5x_1 - 2x_2 - 2x_3 = 9$$
$$x_1 - 3x_2 + x_3 = -2$$
$$-6x_1 + 4x_2 + 11x_3 = 1$$

So,

for 1st equation (row);

$$|coefficient\ of\ x_1| > |coefficient\ of\ x_2| + |coefficient\ of\ x_3|$$

$$or \quad |5| > |-2| + |-2|$$

For 2nd equation(row);

$$|coefficient\ of\ x_2| > |coefficient\ of\ x_1| + |coefficient\ of\ x_3|$$

$$or \quad |-3| > |1| + |1|$$

For 3rd equation(row);

$$|coefficient\ of\ x_3| > |coefficient\ of\ x_1| + |coefficient\ of\ x_2|$$

$$or \quad |11| > |-6| + |4|$$

Now the coefficient matrix is strictly diagonally dominant, so we know Gauss-Seidel works with any initial vector. In fact, if $x^{(0)} = 0$, then $x^{(8)} = (2.9987, 1.9992, .9996)$.

Exercises:

Solve the system in exercise 1 to 3 using Jacobi's method, with $x^{(0)} = 0$ and three iterations. Repeat the iterations until two successive approximations agree within a tolerance of .001 in each entry.

1. $\begin{aligned} 4x_1 + x_2 &= 7 \\ -x_1 + 5x_2 &= -7 \end{aligned}$

2. $\begin{aligned} 10x_1 - x_2 &= 25 \\ x_1 + 8x_2 &= 43 \end{aligned}$

3. $\begin{aligned} 3x_1 + x_2 &= 11 \\ -x_1 - 5x_2 + 2x_3 &= 15 \\ 3x_2 + 7x_3 &= 17 \end{aligned}$

4. $\begin{aligned} 50x_1 - x_2 &= 149 \\ x_1 - 100x_2 + 2x_3 &= -101 \\ 2x_2 + 50x_3 &= -98 \end{aligned}$

In exercises 5 to 8, use the Gauss Seidel method, with $x^{(0)} = 0$ and two iterations. Compare the number of iterations needed by Gauss Seidel and Jacobi to make two successive approximations agree within a tolerance of .001.

5. The system in exercise 1

6. The system in exercise 2

7. The system in exercise 3

8. The system in exercise 4

Determine which of the matrices in exercises 9 and 10 are strictly diagonally dominant.

9. (a) $\begin{bmatrix} 5 & 4 \\ 4 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 9 & -5 & 2 \\ 5 & -8 & -1 \\ -2 & 1 & 4 \end{bmatrix}$

10. (a) $\begin{bmatrix} 3 & -2 \\ 2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 3 & 1 \\ 3 & 6 & -4 \\ 1 & -4 & 7 \end{bmatrix}$

Show that the Gauss Seidel method will produce a sequence converging to the solution of the following system, provided the equations are arranged properly:

11. $\begin{aligned} x_1 - 3x_2 + x_3 &= -2 \\ -6x_1 + 4x_2 + 11x_3 &= 1 \\ 5x_1 - 2x_2 - 2x_3 &= 9 \end{aligned}$

12. $\begin{aligned} -x_1 + 4x_2 - x_3 &= 3 \\ 4x_1 - x_2 &= 10 \\ -x_2 + 4x_3 &= 6 \end{aligned}$

Lecture 17

Introduction to Determinant

In algebra, the **determinant** is a special number associated with any square matrix. As we have studied in earlier classes, that the determinant of 2 x 2 matrix is defined as the product of the entries on the main diagonal minus the product of the entries off the main diagonal.. The determinant of a matrix A is denoted by det (***A***) or |A|

For example: $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Then det (***A***) = *ad-bc*.

or |A| = ad – bc

Example: Find the determinant of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

$$|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1\times 4 - 2\times 3 = 4-6 = -2$$

To extend the definition of the det(***A***) to matrices of higher order, we will use subscripted entries for ***A***.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

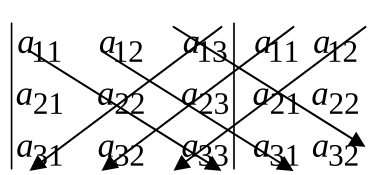
$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{vmatrix} = a_{11}b_{22} - a_{12}b_{21}$$

This is called a 2x2 determinant.

The determinant of a 3x3 matrix is also called a 3x3 determinant is defined by the following formula.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \qquad (1)$$

For finding the determinant of the 3x3 matrix, we look at the following diagram.



We write 1st and 2nd columns again beside the determinant. The first arrow goes from a_{11} to a_{33}, which gives us product: $a_{11}a_{22}a_{33}$. The second arrow goes from a_{12} to a_{31}, which gives us product: $a_{12}a_{23}a_{31}$. The third arrow goes from a_{13} to a_{32}, which gives us the product: $a_{13}a_{21}a_{32}$. These values are taken with positive signs.

The same method is used for the next three arrows that go from right to left downwards, but these product are taken as negative signs.

$$= a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Example 2: Find the determinant of the matrix $A=\begin{bmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{bmatrix}$

$$detA = \begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} \begin{matrix} 1 & 2 \\ -4 & 5 \\ 7 & -8 \end{matrix}$$

$$= 1\times5\times9 + 2\times6\times7 + 3\times(-4)\times(-8) - 3\times5\times7 - 1\times6\times(-8) - 2\times(-4)\times9$$

$$= 45+84+96-105+48+72$$

$$= 240$$

We saw earlier that a $2x2$ matrix is invertible if and only if its determinant is nonzero. In simple words, a matrix has its inverse if its determinant is nonzero. To extend this useful fact to larger matrices, we need a definition for the determinant of the $n\times n$ matrix. We can discover the definition for the 3×3 case by watching what happens when an invertible 3×3 matrix $\boldsymbol{A}$ is row reduced.

<u>Gauss' algorithm for evaluation of determinants:</u>

1) Firstly we apply it for 2×2 matrix say

$$A=\begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix}$$

$R_2' \to R_2 - 2R_1$ (Multiplying 1st row by 2 and then subtracting from 2nd row)

$$\sim \begin{bmatrix} 2 & 3 \\ 4-2(2) & 3-2(3) \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & -3 \end{bmatrix}$$

Now the determinant of this upper triangular matrix is the product of its entries on main diagonal that is

$Det(A) = 2(-3) - 0 \times 3 \ = -6 - 0 \ = -6$

2) For 3×3 matrix say

$$B = \begin{bmatrix} -2 & 2 & -3 \\ -1 & 1 & 3 \\ 2 & 0 & -1 \end{bmatrix}$$

By $R'_{12} \to R_{21}$ (Interchanging of 1st and 2nd rows)

$$\sim \begin{bmatrix} -1 & 1 & 3 \\ -2 & 2 & -3 \\ 2 & 0 & -1 \end{bmatrix}$$

$R'_2 \to R_2 - 2R_1$ (Multiplying 1st row by '-2' and then adding in the 2nd row)

$R'_3 \to R_3 + 2R_1$ (Multiplying 1st row by '2' and then adding in the 3rd row)

$$\sim \begin{bmatrix} -1 & 1 & 3 \\ 0 & 0 & -9 \\ 0 & 2 & 5 \end{bmatrix}$$

By $R'_{23} \to R_{32}$ (Interchanging of 2nd and 3rd rows)

$$\sim \begin{bmatrix} -1 & 1 & 3 \\ 0 & 2 & 5 \\ 0 & 0 & -9 \end{bmatrix}$$

Now the determinant of this upper triangular matrix is the product of its entries on main diagonal that is

$Det(B) = (-1) \cdot 2 \cdot (-9) = 18$

So in general,

For a 1×1 matrix:

say, $A = [a_{ij}]$ - we define $\det A = a_{11}$.

For 2×2 matrix:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

By $R'_2 \to R_2 - \left(\dfrac{a_{21}}{a_{11}}\right) R_1$ provided that $a_{11} \neq 0$

$$\sim \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \dfrac{a_{21}}{a_{11}} a_{12} \end{bmatrix}$$

$\therefore \ \Delta = \det A =$ product of the diagonal entries

$$= a_{11}\left(a_{22} - \frac{a_{21}}{a_{11}} a_{12}\right) = a_{11}a_{22} - a_{12}a_{21}$$

For 3×3 matrix say:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

By $R_2' \to R_2 - \left(\frac{a_{21}}{a_{11}}\right)R_1,\ R_3' \to R_3 - \left(\frac{a_{31}}{a_{11}}\right)R_1$ provided that $a_{11} \neq 0$

$$\sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} & \frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}} \\ 0 & \frac{a_{32}a_{11} - a_{12}a_{31}}{a_{11}} & \frac{a_{11}a_{33} - a_{13}a_{31}}{a_{11}} \end{bmatrix}$$

By $R_3' \to R_3 - \left(\frac{\frac{a_{32}a_{11} - a_{12}a_{31}}{a_{11}}}{\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}}}\right)R_2$ provided that $\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} \neq 0$

$$\sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} & \frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}} \\ 0 & 0 & \frac{a_{11}a_{33} - a_{13}a_{31}}{a_{11}} - \left(\frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}}\right)\left(\frac{\frac{a_{32}a_{11} - a_{12}a_{31}}{\not{a}_{11}}}{\frac{a_{22}a_{11} - a_{12}a_{21}}{\not{a}_{11}}}\right) \end{bmatrix} \because a_{11} \neq 0$$

Which is in echelon form.Now,

$\Delta = \det A =$ product of the diagonal entries

$$= \not{a}_{11}\left(\frac{a_{22}a_{11}-a_{12}a_{21}}{\not{a}_{11}}\right)\left(\frac{a_{11}a_{33}-a_{13}a_{31}}{a_{11}}-\left(\frac{a_{23}a_{11}-a_{13}a_{21}}{a_{11}}\right)\left(\frac{a_{32}a_{11}-a_{12}a_{31}}{a_{22}a_{11}-a_{12}a_{21}}\right)\right)$$
$$=\left(a_{22}a_{11}-a_{12}a_{21}\right)\left(\frac{a_{11}a_{33}-a_{13}a_{31}}{a_{11}}\right)-\left(a_{22}a_{11}-a_{12}a_{21}\right)\left(\frac{a_{23}a_{11}-a_{13}a_{21}}{a_{11}}\right)\left(\frac{a_{32}a_{11}-a_{12}a_{31}}{a_{22}a_{11}-a_{12}a_{21}}\right)$$
$$=\frac{1}{a_{11}}\left\{\left(a_{22}a_{11}-a_{12}a_{21}\right)\left(a_{11}a_{33}-a_{13}a_{31}\right)-\left(a_{23}a_{11}-a_{13}a_{21}\right)\left(a_{32}a_{11}-a_{12}a_{31}\right)\right\}$$
$$=\frac{1}{a_{11}}\left\{a_{11}^2a_{22}a_{33}-a_{11}a_{22}a_{13}a_{31}-a_{12}a_{21}a_{11}a_{33}+a_{12}a_{21}a_{13}a_{31}-a_{23}a_{11}^2a_{32}+a_{23}a_{11}a_{12}a_{31}+a_{13}a_{21}a_{32}a_{11}-a_{12}a_{21}a_{13}a_{31}\right\}$$
$$=\frac{1}{a_{11}}\left\{a_{11}^2a_{22}a_{33}-a_{11}a_{22}a_{13}a_{31}-a_{12}a_{21}a_{11}a_{33}-a_{23}a_{11}^2a_{32}+a_{23}a_{11}a_{12}a_{31}+a_{13}a_{21}a_{32}a_{11}\right\}$$
$$=\frac{a_{11}}{a_{11}}\left\{a_{11}a_{22}a_{33}-a_{22}a_{13}a_{31}-a_{12}a_{21}a_{33}-a_{23}a_{11}a_{32}+a_{23}a_{12}a_{31}+a_{13}a_{21}a_{32}\right\}$$
$$=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}-a_{13}a_{22}a_{31}$$

□

Since ***A*** is invertible, Δ must be nonzero. *The converse is true as well.*

To generalize the definition of the determinant to larger matrices, we will use 2×2 determinants to rewrite the 3×3 determinant Δ described above. Since the terms in Δ can be grouped as:

$$\Delta=(a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32})-(a_{12}a_{23}a_{31}-a_{12}a_{21}a_{33})+(a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31})$$
$$=a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{12}(a_{21}a_{33}-a_{23}a_{31})+a_{13}(a_{21}a_{32}-a_{22}a_{31}),$$

$$\Delta=a_{11}\cdot\det\begin{bmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{bmatrix}-a_{12}\cdot\det\begin{bmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{bmatrix}+a_{13}\cdot\det\begin{bmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{bmatrix}$$

$$\Delta=a_{11}\cdot\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix}-a_{12}\begin{vmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{vmatrix}+a_{13}\cdot\begin{vmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{vmatrix}$$

For brevity, we write $\Delta=a_{11}\cdot\det A_{11}-a_{12}\cdot\det A_{12}+a_{13}\cdot\det A_{13}$ (3)

$$\det(A_{11})=\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix},\quad \det(A_{12})=\begin{vmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{vmatrix}\quad\text{and}\quad \det(A_{13})=\begin{vmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{vmatrix}$$

where

$\boldsymbol{A_{11}}$ is obtained from ***A*** by deleting the first row and first column.
$\boldsymbol{A_{12}}$ is obtained from ***A*** by deleting the first row and second column.
$\boldsymbol{A_{13}}$ is obtained from ***A*** by deleting the first row and third column.
So in general, for any square matrix ***A***, let $\boldsymbol{A_{ij}}$ denote the sub-matrix formed by deleting the *ith* row and *jth* column of ***A***.

Let's understand it with the help of an example.

Example3:

Find the determinant of the matrix $A=\begin{bmatrix}1&4&3\\5&2&4\\3&6&3\end{bmatrix}$

Solution: Given $A=\begin{bmatrix}1&4&3\\5&2&4\\3&6&3\end{bmatrix}$

$$|A|=\begin{vmatrix}1&4&3\\5&2&4\\3&6&3\end{vmatrix}$$
$$=1\begin{vmatrix}2&4\\6&3\end{vmatrix}-4\begin{vmatrix}5&4\\3&3\end{vmatrix}+3\begin{vmatrix}5&2\\3&6\end{vmatrix}$$
$$=1(2\times3-4\times6)-4(5\times3-4\times3)+3(5\times6-2\times3)$$
$$=1(6-24)-4(15-12)+3(30-6)$$
$$=1(-18)-4(3)+3(24)$$
$$=-18-12+72$$
$$=42$$

For instance, if $A=\begin{bmatrix}1&-2&5&0\\2&0&4&-1\\3&1&0&7\\0&4&-2&0\end{bmatrix}$

then $\boldsymbol{A_{32}}$ is obtained by crossing out row 3 and column 2,

$$\begin{bmatrix}1&-2&5&0\\2&0&4&-1\\3&1&0&7\\0&4&-2&0\end{bmatrix}\quad \text{so that}\quad A_{32}=\begin{bmatrix}1&5&0\\2&4&-1\\0&-2&0\end{bmatrix}$$

We can now give a recursive definition of a determinant.

When <u>$n = 3$,</u> det $\boldsymbol{A}$ is defined using determinants of the 2×2 submatrices A_{1j}.
When <u>$n = 4$</u>, det $\boldsymbol{A}$ uses determinants of the 3×3 submatrices A_{1j}
In general, an nxn determinant is defined by determinants of $(n-1)\times(n-1)$ sub matrices.

Definition:

For $n \geq 2$, the **determinant** of $n \times n$ matrix $A = [a_{ij}]$ is the sum of n terms of the form $\pm a_{1j} \times (\det A_{1j})$, with plus and minus signs alternating, where the entries $a_{11}, a_{12}, \cdots, a_{1n}$ are from the first row of ***A***.

Here for a_{ij},

$i = 1, 2, 3, \ldots, n \quad (1 \leq i \leq n)$

$j = 1, 2, 3, \ldots, n \quad (1 \leq j \leq n)$

In symbols, $\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + ... + (-1)^{1+n} a_{1n} \det A_{1n} = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$

Example 4:

Compute the determinant of $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$

Solution:

Here A is $n \times n = 3 \times 3$ matrix such that

$i = 1, 2, 3$

$j = 1, 2, 3$

$\because \det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$ and here $j = 1, 2, 3$

$\therefore \det(A) = \sum_{j=1}^{3} (-1)^{1+j} a_{1j} \det A_{1j} = (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} + (-1)^{1+3} a_{13} \det A_{13}$

$= a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}$

$$\det A = 1.\det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5.\det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0.\det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix}$$

$$\det A = 1.\begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5.\begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0.\begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

= 1 [4(0) – (-1)(-2)] -5 [2(0) – 0(-1)] +0[2(-2) – 4(0)]

$= 1(0-2) - 5(0-0) + 0(-4-0) = -2$

Minor of an element:

If ***A*** is a square matrix, then the **Minor** of entry a_{ij} (called the ijth minor of ***A***) is denoted by M_{ij} and is defined to be the determinant of the sub matrix that remains when the *ith* row and *jth* column of ***A*** are deleted.

In the above example, Minors are the followings:

$$M_{11} = \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix}, \quad M_{12} = \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix}, \quad M_{13} = \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

Cofactor of an element:

The number $C_{ij}=(-1)^{i+j}M_{ij}$ is called the cofactor of entry a_{ij}(or the *ijth* cofactor of A).
When the + or – sign is attached to the Minor, then Minor becomes a cofactor.

In the above example, Cofactors are the followings:

$$C_{11} = (-1)^{1+1} M_{11}, \quad C_{12} = (-1)^{1+2} M_{12}, \quad C_{13} = (-1)^{1+3} M_{13}$$

$$C_{11} = (-1)^{1+1} \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix}, \quad C_{12} = (-1)^{1+2} \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix}, \quad C_{13} = (-1)^{1+3} \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

Example 5: Find the minor and cofactor of the matrix $A = \begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}$

Solution: Here $A = \begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}$

The minor of entry a_{11} is

$$M_{11} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} = 5\times 8 - 6\times 4 = 40 - 24 = 16$$

and the corresponding cofactor is

$$C_{11} = (-1)^{1+1} M_{11} = M_{11} = 16$$

The minor of entry a_{32} is

$$\mathrm{M}_{32} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = 26$$

and the corresponding cofactor is

$$C_{32} = (-1)^{3+2} M_{32} = -M_{32} = -\begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = -26$$

Alternate Definition:

Given $A=[a_{ij}]$, the (i, j)-cofactor of ***A*** is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij} \qquad (4)$$

Then $\det A = a_{11}C_{11} + a_{12}C_{12} + ... + a_{1n}C_{1n}$

This formula is called the *cofactor expansion across the first row* of ***A***.

Example 6: Expand a 3x3 determinant using cofactor concept $A=\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix}$

Solution: Using cofactor expansion along the first column;

$$\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} = (1)(-1)^{1+1}\begin{vmatrix} 5 & 6 \\ -8 & 9 \end{vmatrix} + (-4)(-1)^{2+1}\begin{vmatrix} 2 & 3 \\ -8 & 9 \end{vmatrix} + (7)(-1)^{3+1}\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

Now if we compare it with the formula (4),

$$= 1\,C_{11} + (-4)\,C_{21} + 7\,C_{31}$$

$$= (1)(-1)^2\begin{vmatrix} 5 & 6 \\ -8 & 9 \end{vmatrix} + (-4)(-1)^3\begin{vmatrix} 2 & 3 \\ -8 & 9 \end{vmatrix} + (7)(-1)^4\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

$$= (1)(1)\begin{vmatrix} 5 & 6 \\ -8 & 9 \end{vmatrix} + (-4)(-1)\begin{vmatrix} 2 & 3 \\ -8 & 9 \end{vmatrix} + (7)(1)\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

$$= 1\begin{vmatrix} 5 & 6 \\ -8 & 9 \end{vmatrix} + 4\begin{vmatrix} 2 & 3 \\ -8 & 9 \end{vmatrix} + 7\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

$$= 1(45-(-48)) + 4(18-(-24)) + 7(12-15)$$

$$= 1(45+48) + 4(18+24) + 7(12-15)$$

$$= (1)(93) + (4)(42) + (7)(-3) = 240$$

Using cofactor expansion along the second column,

$$\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} = (2)(-1)^{1+2}\begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(-1)^{2+2}\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1)^{3+2}\begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}$$

$$= (2)(-1)^{3}\begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(-1)^{4}\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1)^{5}\begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}$$

$$= (2)(-1)\begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(1)\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1)\begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}$$

$$= -2\begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + 5\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + 8\begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}$$

$$= -2(-36 - 42) + 5(9 - 21) + 8(6 - (-12))$$

$$= (-2)(-78) + (5)(-12) + (8)(18) = 240$$

Theorem 1: The determinant of an $n \times n$ matrix $\boldsymbol{A}$ can be computed by a cofactor expansion across any row or down any column. The expansion across the ith row using the cofactors in (4) is

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

The cofactor expansion down the jth column is

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

The plus or minus sign in the (i, j)-cofactor depends on the position of a_{ij} in the matrix, regardless of the sign of a_{ij} itself. The factor $(-1)^{i+j}$ determines the following checkerboard pattern of signs:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

Example 7: Use a cofactor expansion across the third row to compute det $\boldsymbol{A}$, where

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Solution: Compute $\det A = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33}$

$$= (-1)^{3+1} a_{31} \det A_{31} + (-1)^{3+2} a_{32} \det A_{32} + (-1)^{3+3} a_{33} \det A_{33}$$

$$= 0\begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix}$$

$$= 0 + 2(-1) + 0 = -2$$

Theorem 1 is helpful for computing the determinant of a matrix that contains many zeros. For example, if a row is mostly zeros, then the cofactor expansion across that row has many terms that are zero, and the cofactors in those terms need not be calculated.
The same approach works with a column that contains many zeros.

Example 8: Evaluate the determinant of $A = \begin{bmatrix} 2 & 0 & 0 & 5 \\ -1 & 2 & 4 & 1 \\ 3 & 0 & 0 & 3 \\ 8 & 6 & 0 & 0 \end{bmatrix}$

Solution: $\det(A) = \begin{vmatrix} 2 & 0 & 0 & 5 \\ -1 & 2 & 4 & 1 \\ 3 & 0 & 0 & 3 \\ 8 & 6 & 0 & 0 \end{vmatrix}$

Expand from third column

$$\det(A) = 0\times C_{13} + 4\times C_{23} + 0\times C_{33} + 0\times C_{43}$$

$$= 0 \ + 4\times C_{23} \ + 0 \ + 0$$

$$= \ 4\times C_{23}$$

$$= 4\times(-1)^{2+3}\begin{vmatrix} 2 & 0 & 5 \\ 3 & 0 & 3 \\ 8 & 6 & 0 \end{vmatrix}$$

Expand from second column

$$= -4\left(0 \ + \ 0 \ + \ (-6)\begin{vmatrix} 2 & 5 \\ 3 & 3 \end{vmatrix}\right)$$

$$= (-4) \ (-6)\begin{vmatrix} 2 & 5 \\ 3 & 3 \end{vmatrix}$$

$$= -216$$

Example 9: Show that the value of the determinant is independent of θ

$$A = \begin{vmatrix} \sin\theta & \cos\theta & 0 \\ -\cos\theta & \sin\theta & 0 \\ \cos\theta - \sin\theta & \sin\theta + \cos\theta & 1 \end{vmatrix}$$

Solution: Consider $A = \begin{vmatrix} \sin\theta & \cos\theta & 0 \\ -\cos\theta & \sin\theta & 0 \\ \cos\theta - \sin\theta & \sin\theta + \cos\theta & 1 \end{vmatrix}$

Expand the given determinant from 3rd column we have

$$= 0 - 0 + (-1)^{3+3}[\sin^2\theta + \cos^2\theta] = 1$$

Example 10: Compute det ***A***, where $A=\begin{bmatrix}3 & -7 & 8 & 9 & -6\\ 0 & 2 & -5 & 7 & 3\\ 0 & 0 & 1 & 5 & 0\\ 0 & 0 & 2 & 4 & -1\\ 0 & 0 & 0 & -2 & 0\end{bmatrix}$

Solution: The cofactor expansion down the first column of ***A*** has all terms equal to zero except the first.

Thus $$\det A = 3\begin{vmatrix}2 & -5 & 7 & 3\\ 0 & 1 & 5 & 0\\ 0 & 2 & 4 & -1\\ 0 & 0 & -2 & 0\end{vmatrix} - 0.C_{21} + 0.C_{31} - 0.C_{41} + 0.C_{51}$$

Henceforth we will omit the zero terms in the cofactor expansion.

Next, expand this 4×4 determinant down the first column, in order to take advantage of the zeros there.

We have $$\det A = 3\times 2\begin{vmatrix}1 & 5 & 0\\ 2 & 4 & -1\\ 0 & -2 & 0\end{vmatrix}$$

This 3×3 determinant was computed above and found to equal –2.

Hence det ***A*** = 3x2x(-2) = - 12.

The matrix in this example was nearly triangular. The method in that example is easily adapted to prove the following theorem.

Triangular Matrix:

A triangular matrix is a special kind of m x n matrix where the entries either below or above the main diagonal are zero.

$\begin{bmatrix}1 & 4 & 2\\ 0 & 3 & 4\\ 0 & 0 & 1\end{bmatrix}$ is upper triangular and $\begin{bmatrix}1 & 0 & 0\\ 2 & 8 & 0\\ 4 & 9 & 7\end{bmatrix}$ is lower triangular matrices.

Determinants of Triangular Matrices:

Determinants of the triangular matrices are also easy to evaluate regardless of size.

Theorem: If ***A*** is triangular matrix, then det (***A***) is the product of the entries on the main diagonal.

Consider a 4x4 lower triangular matrix.

$$A = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

Keeping in mind that an elementary product must have exactly one factor from each row and one factor from each column, the only elementary product that does not have one of the six zeros as a factor is $(a_{11}a_{22}a_{33}a_{44})$. The column indices of this elementary product are in natural order, so the associated signed elementary product takes a +.

Thus, det (A)= $a_{11} \times a_{22} \times a_{33} \times a_{44}$

Example 11:

$$\begin{vmatrix} -2 & 5 & 7 \\ 0 & 3 & 8 \\ 0 & 0 & 5 \end{vmatrix} = (-2)(3)(5) = -30$$

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 4 & 9 & 0 & 0 \\ -7 & 6 & -1 & 0 \\ 3 & 8 & -5 & -2 \end{vmatrix} = (1)(9)(-1)(-2) = 18$$

$$\begin{vmatrix} 1 & 2 & 7 & -3 \\ 0 & 1 & -4 & 1 \\ 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 3 \end{vmatrix} = (1)(1)(2)(3) = 6$$

The strategy in the above Example of looking for zeros works extremely well when an entire row or column consists of zeros. In such a case, the cofactor expansion along such a row or column is a sum of zeros! So the determinant is zero. Unfortunately, most cofactor expansions are not so quickly evaluated.

Numerical Note: By today's standards, a $25x25$ matrix is small. Yet it would be impossible to calculate a 25×25 determinant by cofactor expansion. In general, a cofactor expansion requires over *n!* multiplications, and $25! \sim 1.5x10^{25}$.

If a supercomputer could make one trillion multiplications per second, it would have to run for over 500,000 years to compute a 25×25 determinant by this method. Fortunately, there are faster methods, as we'll soon discover.

Example 12: Compute $\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix}$

Solution: Take advantage of the zeros. Begin with a cofactor expansion down the third column to obtain a 3×3 matrix, which may be evaluated by an expansion down its first column,

$$\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix} = (-1)^{1+3}2\begin{vmatrix} 0 & 3 & -4 \\ -5 & -8 & 3 \\ 0 & 5 & -6 \end{vmatrix}$$

$$= 2\cdot(-1)^{2+1}(-5)\begin{vmatrix} 3 & -4 \\ 5 & -6 \end{vmatrix} = 20$$

The –1 in the next-to-last calculation came from the position of the –5 in the 3×3 determinant.

Exercises:

Compute the determinants in exercises 1 to 6 by cofactor expansions. At each step, choose a row or column that involves the least amount of computation.

1. $\begin{vmatrix} 6 & 0 & 0 & 5 \\ 1 & 7 & 2 & -5 \\ 2 & 0 & 0 & 0 \\ 8 & 3 & 1 & 8 \end{vmatrix}$

2. $\begin{vmatrix} 1 & -2 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 2 & -6 & -7 & 5 \\ 5 & 0 & 4 & 4 \end{vmatrix}$

3. $\begin{vmatrix} 3 & 5 & -8 & 4 \\ 0 & -2 & 3 & -7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$

4. $\begin{vmatrix} 4 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ 2 & 6 & 3 & 0 \\ 5 & -8 & 4 & -3 \end{vmatrix}$

5. $\begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 0 & 2 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix}$

6. $\begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$

Use the method of Example 2 to compute the determinants in exercises 7 and 8. In exercises 9 to 11, compute the determinant of elementary matrix. In exercises 12 and 13, verify that det ***EA*** = (det ***E***) . (det ***A***), where ***E*** is the elementary matrix and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

7. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$

8. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$

9. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$

10. $\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

11. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

14. Let $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$. Write 5***A***. Is det 5***A*** = 5 det ***A***?

15. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and *k* be a scalar. Find a formula that relates det (*k**A***)to *k* and det ***A***.

Lecture 18

Properties of Determinants

In this lecture, we will study the properties of the determinants. Some of them have already been discussed and you will be familiar with these. These properties become helpful, while computing the values of the determinants. The secret of determinants lies in how they change when row or column operations are performed.

Theorem 3:(Row Operations)**:** Let ***A*** be a square matrix.

a. If a multiple of one row of *A* is added to another row, the resulting determinant will remain same.
b. If two rows of ***A*** are interchanged to produce ***B***, then *det* ***B*** = –*det* ***A***.
c. If one row of ***A*** is multiplied by *k* to produce ***B***, then *det* ***B*** = *k* . *det* ***A***.

The following examples show how to use Theorem 3 to find determinants efficiently.

a. If a multiple of one row of *A* is added to another row, the resulting determinant will remain same.

Example:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$Multiplying\ 2nd\ row\ by\ non-zero\ scalar\ say\ 'k'\ as$$

$$ka_{21}\quad ka_{22}\quad ka_{23} ----adding\ this\ in\ 1st\ row\ then\ 'A'\ becomes$$

$$= \begin{vmatrix} a_{11}+ka_{21} & a_{12}+ka_{22} & a_{13}+ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \qquad R_1' \to R_1 + kR_2$$

If each element of any row(column) can be expressed as sum of two elements then the resulting determinant can be expressed as sum of two determinants, so in this case

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} ka_{21} & ka_{22} & ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + k\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$ By using property (c) of above theorem 3.

If any two rows or columns in a determinant are identical then value of this determinant is zero. So in this case $R_1 \equiv R_2$

$$\therefore \quad \Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + k(0)$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = A$$

b. If two rows of *A* are interchanged to produce *B*, then *det B = –det A.*

Example 1:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 1 & 1 \\ 0 & 8 & 9 \end{bmatrix}$$

$$\text{Now, } \det A = \begin{vmatrix} 1 & 2 & 3 \\ 5 & 1 & 1 \\ 0 & 8 & 9 \end{vmatrix} = 1(9-8) - 2(45-0) + 3(40-0) = 1 - 90 + 120 = 31$$

Now interchange column 1st with 2nd we get a new matrix, $B = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 5 & 1 \\ 8 & 0 & 9 \end{bmatrix}$

$$\det B = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 5 & 1 \\ 8 & 0 & 9 \end{vmatrix} = 2(45-0) - 1(9-8) + 3(0-40) = 90 - 1 - 120 = -31$$

c. If one row of *A* is multiplied by *k* to produce *B*, then *det B = k . det A.*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 0 & 1 \\ 0 & 8 & 9 \end{bmatrix}$$

$$|A| = 1(0-8) - 2(45-0) + 3(40-0)$$

$$= -8 - 90 + 120 = 22$$

$Multipling R_1 \; by \; k, we \; get \;\; say$

$$B = \begin{bmatrix} 1k & 2k & 3k \\ 5 & 0 & 1 \\ 0 & 8 & 9 \end{bmatrix}$$

$$|B| = k(40-0) - 2k(45-0) + 3k(40-0)$$
$$= 40k - 90k + 120k = 22k$$
$$= k|A|$$

Example 2:

$$\text{Evaluate } A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix}$$

Solution:

$$\det A = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -7 \\ 0 & -2 & -8 & -10 \\ 0 & -7 & -10 & -13 \end{vmatrix} \quad by\ R_2{}' \to R_2 + (-2)R_1, R_3{}' \to R_3 + (-3)R_1, R_4{}' \to R_4 + (-4)R_1$$

$$= \begin{vmatrix} -1 & -2 & -7 \\ -2 & -8 & -10 \\ -7 & -10 & -13 \end{vmatrix} \quad \text{expanding from Ist column}$$

$$= (-1)(-2)(-1)\begin{vmatrix} 1 & 2 & 7 \\ 1 & 4 & 5 \\ 7 & 10 & 13 \end{vmatrix} \quad taking\ (-1), (-2)\ and\ \left(-1\right) common\ from\, 1st, 2nd, 3rd\ rows$$

$$= (-2)\begin{vmatrix} 1 & 2 & 7 \\ 0 & 2 & -2 \\ 0 & -4 & -36 \end{vmatrix} \quad by\ R_2{}' \to R_2 + (-1)R_1, R_3{}' \to R_3 + (-7)R_1$$

$$= (-2)\begin{vmatrix} 2 & -2 \\ -4 & -36 \end{vmatrix} \quad \text{expanding by 1st column}$$

$$= (-2)(2)(-4)\begin{vmatrix} 1 & -1 \\ 1 & 9 \end{vmatrix} \quad \text{taking 2 and (-4) common from 1st and 2nd rows respectively.}$$

$$= 16\begin{vmatrix} 1 & -1 \\ 0 & 10 \end{vmatrix} \quad by\ R_2 + (-1)R_1$$

$$= 160$$

Example 3: Evaluate the determinant of the matrix $A=\begin{bmatrix}4&2&5&10\\1&1&6&3\\7&3&0&5\\0&2&5&8\end{bmatrix}$

Solution:

$$\det A=\begin{vmatrix}4&2&5&10\\1&1&6&3\\7&3&0&5\\0&2&5&8\end{vmatrix}$$

$$=-\begin{vmatrix}1&1&6&3\\4&2&5&10\\7&3&0&5\\0&2&5&8\end{vmatrix} \quad \text{interchanging } R_1 \text{ and } R_2 (R'_{12})$$

$$=-\begin{vmatrix}1&1&6&3\\0&-2&-19&-2\\0&-4&-42&-16\\0&2&5&8\end{vmatrix} \quad By\ R_2' \to R_2+(-4)R_1, R_3' \to R_3+(-7)R_1$$

$$=-\begin{vmatrix}-2&-19&-2\\-4&-42&-16\\2&5&8\end{vmatrix} \quad \text{expanding from 1st column}$$

$$=(-1)^3\begin{vmatrix}2&19&2\\4&42&16\\2&5&8\end{vmatrix} \quad \text{taking (-1) as a common factor from } R_1 \text{ and } R_2$$

$$=-\begin{vmatrix}2&19&2\\4&42&16\\2&5&8\end{vmatrix}$$

$$=-2\begin{vmatrix}1&19&2\\2&42&16\\1&5&8\end{vmatrix}$$

$$=(-2)\begin{vmatrix}1&19&2\\0&4&12\\0&-14&6\end{vmatrix} \quad By\ R_2' \to R_2+(-2)R_1, R_3' \to R_3+(-1)R_1$$

$$= (-2)\begin{vmatrix} 1 & 19 & 2 \\ 0 & 4 & 12 \\ 0 & -14 & 6 \end{vmatrix} \quad R_2 + (-2)R_1, R_3 + (-1)R_1$$

$$= -2\begin{vmatrix} 4 & 12 \\ -14 & 6 \end{vmatrix} \text{ expand from Ist column}$$

$$= -2(24+168) = -384$$

Example 4: Without expansion, show that $\begin{vmatrix} x & a+x & b+c \\ x & b+x & c+a \\ x & c+x & a+b \end{vmatrix} = 0$

Solution:

$$\begin{vmatrix} x & a+x & b+c \\ x & b+x & c+a \\ x & c+x & a+b \end{vmatrix}$$

$$= \begin{vmatrix} x & a+x-x & b+c \\ x & b+x-x & c+a \\ x & c+x-x & a+b \end{vmatrix} \quad By\ C_2' \to C_2 - C_1$$

$$= \begin{vmatrix} x & a & b+c \\ x & b & c+a \\ x & c & a+b \end{vmatrix}$$

$Taking\ 'x'\ common\ from\ C_1$

$$= x\begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix}$$

$$= x\begin{vmatrix} 1 & a+b+c & b+c \\ 1 & b+c+a & c+a \\ 1 & c+a+b & a+b \end{vmatrix} \quad By\ C_2' \to C_2 + C_3$$

Now taking (a+b+c) common form C_2

$$= x(a+b+c)\begin{vmatrix} 1 & 1 & b+c \\ 1 & 1 & c+a \\ 1 & 1 & a+b \end{vmatrix}$$

$= 0$ as column Ist and 2nd are identical ($C_1 \equiv C_2$). So its value will be zero.

Example 5: Evaluate $A=\begin{vmatrix}2&3&1&0&1\\1&1&3&1&2\\2&1&2&3&4\\3&2&1&1&2\\4&1&1&0&0\end{vmatrix}$

Solution: Interchanging R_1 and R_2, we get

$$A=-\begin{vmatrix}1&1&3&1&2\\2&3&1&0&1\\2&1&2&3&4\\3&2&1&1&2\\4&1&1&0&0\end{vmatrix}$$

$R_2' \to R_2-2R_1, R_3' \to R_3-2R_1, R_4' \to R_4-3R_1, R_5' \to R_5-4R_1$

$$=-\begin{vmatrix}1&1&3&1&2\\0&1&-5&-2&-3\\0&-1&-4&1&0\\0&-1&-8&-2&-4\\0&-3&-11&-4&-8\end{vmatrix}$$

expand from C_1

$$=-\begin{vmatrix}1&-5&-2&-3\\-1&-4&1&0\\-1&-8&-2&-4\\-3&-11&-4&-8\end{vmatrix}$$

$R_2' \to R_2+R_1, R_3' \to R_3+R_1, R_4' \to R_4+3R_1$

$$=-\begin{vmatrix}1&-5&-2&-3\\0&-9&-1&-3\\0&-13&-4&-7\\0&-26&-10&-17\end{vmatrix}$$

expand from C_1

$$=-\begin{vmatrix}-9&-1&-3\\-13&-4&-7\\-26&-10&-17\end{vmatrix}$$

taking (-1) common from Ist, 2nd and 3rd row

$$=\begin{vmatrix} 9 & 1 & 3 \\ 13 & 4 & 7 \\ 26 & 10 & 17 \end{vmatrix}$$

interchange Ist and 2nd Column(C'_{12})

$$=-\begin{vmatrix} 1 & 9 & 3 \\ 4 & 13 & 7 \\ 10 & 26 & 17 \end{vmatrix}$$

$C'_2 \to C_2 - 9C_1, C'_3 \to C_3 - 3C_1$

$$=-\begin{vmatrix} 1 & 0 & 0 \\ 4 & -23 & -5 \\ 10 & -64 & -13 \end{vmatrix}$$

expand from Ist row

$$=-\begin{vmatrix} -23 & -5 \\ -64 & -13 \end{vmatrix} = -(299-320) = 21$$

<u>An Algorithm to evaluate the determinant:</u>

Algorithm means a sequence of a finite number of steps to get a desired result. The word Algorithm comes form the famous Muslim mathematician AL-Khwarizmi who invented the word algebra.

The step-by-step evaluation of det(***A***) of order ***n*** is obtained as follows:

<u>Step 1:</u> By an interchange of rows of ***A*** (and taking the resulting sign into account) bring a non zero entry to (1,1) the position (unless all the entries in the first column are zero in which case det ***A***=0).

<u>Step 2:</u> By adding suitable multiples of the first row to all the other rows, reduce the (n-1) entries, except (1,1) in the first column, to 0. Expand det(***A***) by its first column. Repeat this process.
Or continue the following steps.

<u>Step 3:</u> Repeat step 1 and step 2 with the last remaining rows concentrating on the second column.

<u>Step 4:</u> Repeat step 1,step2 and step 3 with the remaining (n-2) rows, (n-3) rows and so on, until a triangular matrix is obtained.

<u>Step5:</u> Multiply all the diagonal entries of the resulting triangular matrix and then multiply it by its sign to get det(A)

Example 6: Compute det ***A***, where $A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.

Solution: *The strategy is to reduce* ***A*** *to echelon form and then to use the fact that the determinant of a triangular matrix is the product of the diagonal entries.* The first two row replacements in column 1 do not change the determinant:

$$\det A = \begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} \quad By\ R_2' \to R_2 + 2R_1, R_3' \to R_1 + R_3$$

An interchange of rows 2 and 3 (R'_{23}), it reverses the sign of the determinant, so

$$\det A = -\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} = -(1)(3)(-5) = 15$$

Example 7: Compute det ***A***, where

$$A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}.$$

Solution: Taking'2' common from 1st row

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix}$$

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix} \quad By\ R_2' \to R_2 - 3R_1, R_3' \to R_3 + 3R_1, R_4' \to R_4 - R_1$$

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (By\ R_4' \to R_4 - \tfrac{1}{2}R_3)$$

$$= 2.\{(1)(3)(-6)(1)\} = -36$$

Example 8: Show that $\begin{vmatrix} x & 2 & 2 & 2 \\ 2 & x & 2 & 2 \\ 2 & 2 & x & 2 \\ 2 & 2 & 2 & x \end{vmatrix} = (x+6)(x-2)^3$

Solution:

$$\begin{vmatrix} x & 2 & 2 & 2 \\ 2 & x & 2 & 2 \\ 2 & 2 & x & 2 \\ 2 & 2 & 2 & x \end{vmatrix}$$

$$= \begin{vmatrix} x+6 & 2 & 2 & 2 \\ x+6 & x & 2 & 2 \\ x+6 & 2 & x & 2 \\ x+6 & 2 & 2 & x \end{vmatrix} \quad By\ \ C_1' \to C_1 + (C_2 + C_3 + C_4)$$

Taking (x+6) common from 1st column

$$= (x+6)\begin{vmatrix} 1 & 2 & 2 & 2 \\ 1 & x & 2 & 2 \\ 1 & 2 & x & 2 \\ 1 & 2 & 2 & x \end{vmatrix}$$

$$= (x+6)\begin{vmatrix} 1 & 2 & 2 & 2 \\ 0 & x-2 & 0 & 0 \\ 0 & 0 & x-2 & 0 \\ 0 & 0 & 0 & x-2 \end{vmatrix} \quad By\ R_2' \to R_2 - R_1, R_3' \to R_3 - R_1, R_4' \to R_4 - R_1$$

which is the triangular matrix and its determinant is the prodcut of main diagonal's entries.

$$= (x+6)(x-2)^3$$

Example 9: Compute det ***A***, where $A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$.

Solution: $A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$

$$\det A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix} R_3' \to R_3 + 2R_1$$

$= \mathbf{0}$ as $R_2 \equiv R_3$

Example10: Compute det ***A***, where

$$A = \begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{bmatrix}$$

Solution:

$$A = \begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{bmatrix} \quad R_4' \to R_4 + R_2$$

$$= (-1)\begin{bmatrix} 2 & 1 & 2 & -1 \\ 0 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{bmatrix} \qquad By\ R_{12}'$$

Expanding from 1st row and 1st column

$$= -2\begin{vmatrix} 5 & -7 & 3 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix}$$

$$= (-2)\{5(6+6)-(-7)(3-0)+3(-9-0)\}$$

$$= 54$$

Remarks:

Suppose that a square matrix ***A*** has been reduced to an echelon form ***U*** by row replacements and row interchanges.

If there are r interchanges, then $\det(A) = (-1)^r \det(U)$

Furthermore, all of the pivots are still visible in ***U*** (because they have not been scaled to ones). If ***A*** is invertible, then the pivots in ***U*** are on the diagonal (since ***A*** is row equivalent to the identity matrix). In this case, *det* ***U*** is the product of the pivots. If ***A*** is not invertible, then ***U*** has a row of zero and *det* ***U*** *= 0*.

$$U = \begin{bmatrix} \bullet & \circ & \circ & \circ \\ 0 & \bullet & \circ & \circ \\ 0 & 0 & \bullet & \circ \\ 0 & 0 & 0 & \bullet \end{bmatrix} \qquad U = \begin{bmatrix} \bullet & \circ & \circ & \circ \\ 0 & \bullet & \circ & \circ \\ 0 & 0 & \bullet & \circ \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\det U \neq 0 \qquad\qquad \det U = 0$$

Thus we have the following formula:

$$\det A = \begin{cases} (-1)^r .(Product\ of\ \ pivots\ in\ U) & When\ A\ is\ invertible \\ 0 & When\ A\ is\ not\ invertible \end{cases} \qquad (1)$$

Example:

Case-01: For 2×2 invertible matrix

Reducing given 2×2 invertible matrix into Echelon form as follows;

$$A = \begin{bmatrix} 4 & 5 \\ 3 & 2 \end{bmatrix}$$

By interchanging 1st and 2nd rows(R'_{12})

$$\sim \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix} \because \text{ one replacement of rows has occurred, } \therefore r = 1$$

$$\sim \begin{bmatrix} 3 & 2 \\ 0 & \frac{7}{3} \end{bmatrix} \text{ By } R_2' \to R_2 - \frac{4}{3}R_1 \text{, we have desired row-echelon form:} U = \begin{bmatrix} 3 & 2 \\ 0 & \frac{7}{3} \end{bmatrix}.$$

Thus using the above formula as follows;

$$\det A = (-1)^r .(Product\,of\;\; pivots\,in U) = (-1)^1(3\cdot\frac{7}{3}) = -7$$

Case-02: For 2×2 non-invertible matrix:
In this case say;

$$A = \begin{bmatrix} 4 & 5 \\ 8 & 10 \end{bmatrix}$$

$$\sim \begin{bmatrix} 4 & 5 \\ 0 & 0 \end{bmatrix} \quad By\;\; R_2{}' \to R_2 - 2R_1 \text{,desired row-echelon form is } U = \begin{bmatrix} 4 & 5 \\ 0 & 0 \end{bmatrix}$$

Here no interchange of rows has occurred. So, $r = 0$ and
$\therefore \det A = (-1)^r .(Product\,of\;\; pivots\,in U) = (-1)^0(4\cdot 0) = 0$

Theorem 5: If A is an $n\times n$ matrix, then *det* $\boldsymbol{A^T}$ = *det* $\boldsymbol{A}$.

Example 11: If $A = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 3 \end{bmatrix}$, find det($\boldsymbol{A}$) and det ($\boldsymbol{A^T}$)

$$\det A = \begin{vmatrix} 1 & 4 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 3 \end{vmatrix} = 1(3-2) - 4(6-6) + 1(2-3) = 1 - 0 - 1 = 0$$

Now

$$A^t = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$$

$$\det A^t = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 1 & 1 \\ 1 & 2 & 3 \end{vmatrix} = 1(3-2) - 2(12-1) + 3(8-1) = 1 - 22 + 21 = 0$$

Remark:
Column operations are useful for both theoretical purposes and hand computations. However, for simplicity we’ll perform only row operations in numerical calculations.

Theorem 6 (Multiplicative Property):

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are $n\times n$ matrices, then $\det(AB) = (\det A)(\det B)$.

Example 12: Verify Theorem 6 for $A = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix} \quad and \quad B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$

Solution: $AB=\begin{bmatrix}6 & 1\\ 3 & 2\end{bmatrix}\begin{bmatrix}4 & 3\\ 1 & 2\end{bmatrix}=\begin{bmatrix}25 & 20\\ 14 & 13\end{bmatrix}$

and $\det AB = 25.13 - 20.14 = 325 - 280 = 45$

Since *det* ***A*** *= 9* and *det* ***B*** *= 5*, $(\det A)(\det B) = 9.5 = 45 = \det AB$

Remark:

det (***A + B***) $\neq$ det ***A*** + det ***B***, in general.

For example**,**

If $A=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix}$ and $B=\begin{bmatrix}-2 & -3\\ -1 & 5\end{bmatrix}$. Then

$A+B=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\Rightarrow \det(A+B)=0$

$\det A+\det B=\begin{vmatrix}2 & 3\\ 1 & -5\end{vmatrix}+\begin{vmatrix}-2 & -3\\ -1 & 5\end{vmatrix}=(-10-3)+(-10-3)=-26\neq\det(A+B)$

Exercise:

Find the determinants in exercises 1 to 6 by row reduction to echelon form.

1. $\begin{vmatrix} 1 & 3 & 0 & 2 \\ -2 & -5 & 7 & 4 \\ 3 & 5 & 2 & 1 \\ 1 & -1 & 2 & -3 \end{vmatrix}$

2. $\begin{vmatrix} 1 & 3 & 3 & -4 \\ 0 & 1 & 2 & -5 \\ 2 & 5 & 4 & -3 \\ -3 & -7 & -5 & 2 \end{vmatrix}$

3. $\begin{vmatrix} 1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 2 & 8 & 5 \\ 3 & -1 & -2 & 3 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 3 & -1 & 0 & -2 \\ 0 & 2 & -4 & -1 & -6 \\ -2 & -6 & 2 & 3 & 9 \\ 3 & 7 & -3 & 8 & -7 \\ 3 & 5 & 5 & 2 & 7 \end{vmatrix}$

5. $\begin{vmatrix} 1 & -2 & 3 & 1 \\ 5 & -9 & 6 & 3 \\ -1 & 2 & -6 & -2 \\ 2 & 8 & 6 & 1 \end{vmatrix}$

6. $\begin{vmatrix} 1 & 3 & 1 & 5 & 3 \\ -2 & -7 & 0 & -4 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{vmatrix}$

Combine the methods of row reduction and cofactor expansion to compute the determinants in exercises 7 and 8.

7. $\begin{vmatrix} 2 & 5 & -3 & -1 \\ 3 & 0 & 1 & -3 \\ -6 & 0 & -4 & 9 \\ 4 & 10 & -4 & -1 \end{vmatrix}$

8. $\begin{vmatrix} 2 & 5 & 4 & 1 \\ 4 & 7 & 6 & 2 \\ 6 & -2 & -4 & 0 \\ -6 & 7 & 7 & 0 \end{vmatrix}$

9. Use determinant to find out whether the matrix is invertible $\begin{bmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{bmatrix}$

10. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be *3 x 3* matrices, with det $\boldsymbol{A}$ = 4 and det $\boldsymbol{B}$ = -3. Use properties of determinants to compute:

(a) det $\boldsymbol{AB}$ (b) det $7\boldsymbol{A}$ (c) det $\boldsymbol{B}^{\boldsymbol{T}}$ (d) det $\boldsymbol{A}^{\boldsymbol{T}}$
(e) det $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$

11 Show that

(a) $$\begin{vmatrix} a_1 & b_1 & a_1+b_1+c_1 \\ a_2 & b_2 & a_2+b_2+c_2 \\ a_3 & b_3 & a_3+b_3+c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

(b) $$\begin{vmatrix} a_1+b_1 & a_1-b_1 & c_1 \\ a_2+b_2 & a_2-b_2 & c_2 \\ a_3+b_3 & a_3-b_3 & c_3 \end{vmatrix} = -2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

12 Show that

(a) $$\begin{vmatrix} a_1+b_1t & a_2+b_2t & a_3+b_3t \\ a_1t+b_1 & a_2t+b_2 & a_3t+b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = (1-t^2)\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

(b) $$\begin{vmatrix} a_1 & b_1+ta_1 & c_1+rb_1+sa_1 \\ a_2 & b_2+ta_2 & c_2+rb_2+sa_2 \\ a_3 & b_3+ta_3 & c_3+rb_3+sa_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

13. Show that $$\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = (y-x)(z-x)(z-y)$$

Lecture 19

Cramer's Rule, Volume, and Linear Transformations

In this lecture, we shall apply the theory discussed in the last two lectures to obtain important theoretical formulas and a geometric interpretation of the determinant.

Cramer's Rule: Cramer's rule is needed in a variety of theoretical calculations. For instance, it can be used to study how the solution of ***Ax = b*** is affected by changes in the entries of ***b***. However, the formula is inefficient for hand calculations, except for 2×2 or perhaps 3×3 matrices.

Theorem 1 (Crammer's Rule): Let ***A*** be an invertible $n\times n$ matrix. For any ***b*** in $\boldsymbol{R}^n$, the unique solution *x* of ***Ax = b*** has entries given by

$$x_i = \frac{\det A_i(b)}{\det A}, \qquad i = 1, 2, \ldots, n \qquad (1)$$

Example 1: Use Cramer's rule to solve the system

$$3x_1 - 2x_2 = 6$$
$$-5x_1 + 4x_2 = 8$$

Solution: Write the system in matrix form, ***Ax = b***

$$\begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$$

where

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \& b = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$$

$$\det A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix} = 12 - 10 = 2$$

$$A_1(b) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, A_2(b) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

Since det ***A*** = *2*, the system has a unique solution. By Cramer's rule,

$$x_1 = \frac{\det A_1(b)}{\det A} = \frac{24+16}{2} = 20$$

$$x_2 = \frac{\det A_2(b)}{\det A} = \frac{24+30}{2} = 27$$

Example 2: Consider the following system in which s is an unspecified parameter. Determine the values of s for which the system has a unique solution and use Cramer's rule to describe the solution.

$$\begin{aligned} 3sx_1 - 2x_2 &= 4 \\ -6x_1 + sx_2 &= 1 \end{aligned}$$

Solution: Here

$$A = \begin{bmatrix} 3s & -2 \\ -6 & s \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad A_1(b) = \begin{bmatrix} 4 & -2 \\ 1 & s \end{bmatrix}, \quad A_2(b) = \begin{bmatrix} 3s & 4 \\ -6 & 1 \end{bmatrix}$$

Since $\det A = 3s^2 - 12 = 3(s+2)(s-2)$

the system has a unique solution when

$\det A \neq 0$

$\Rightarrow 3(s+2)(s-2) \neq 0$

$\Rightarrow s^2 - 4 \neq 0$

$\Rightarrow s \neq \pm 2$

For such an s, the solution is (x_1, x_2), where

$$x_1 = \frac{\det A_1(b)}{\det A} = \frac{4s+2}{3(s+2)(s-2)}, \qquad s \neq \pm 2$$

$$x_2 = \frac{\det A_2(b)}{\det A} = \frac{3s+24}{3(s+2)(s-2)} = \frac{s+8}{(s+2)(s-2)}, \qquad s \neq \pm 2$$

Example 3: Solve, by Cramer's Rule, the system of equations:

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 1 \\ x_1 + 2x_2 - x_3 &= 2 \\ 3x_1 + 2x_2 + 2x_3 &= 3 \end{aligned}$$

Solution: Here $A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 2 & -1 \\ 3 & 2 & 2 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $A_1 = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 2 & -1 \\ 3 & 2 & 2 \end{bmatrix}$

$$A_2 = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -1 \\ 3 & 3 & 2 \end{bmatrix}, \; A_3 = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 2 \\ 3 & 2 & 3 \end{bmatrix}$$

$$D = \det A = 2\cdot 6 - 1\cdot(-5) + 3(-4) = 5$$
$$D_1 = \det A_1 b = 1\cdot 6 + 2\cdot 8 + 3(-5) = 7$$
$$D_2 = \det A_2 b = 2\cdot(7) - 1\cdot(5) + 3(-3) = 0$$
$$D_3 = \det A_3 b = 2\cdot(2) + 1\cdot(-3) + 1(-4) = -3$$
$$So\ x_1 = \frac{D_1}{D} = \frac{7}{5},\quad x_2 = \frac{D_2}{D} = 0,\quad x_3 = \frac{D_3}{D} = -\frac{3}{5}$$

Example 4: Use Cramer's Rule to solve.

$$\begin{aligned} x_1 + \qquad 2x_3 &= 6 \\ -3x_1 + 4x_2 + 6x_3 &= 30 \\ -\ x_1 - 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solution:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{bmatrix}, b = \begin{bmatrix} 6 \\ 30 \\ 8 \end{bmatrix}$$

$$\therefore A_1 = \begin{bmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & -8 & 3 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{bmatrix}$$

Therefore,

$$x_1 = \frac{\det(A_1 b)}{\det(A)} = \frac{-40}{44} = \frac{-10}{11}, \qquad x_2 = \frac{\det(A_2 b)}{\det(A)} = \frac{72}{44} = \frac{18}{11}$$
$$x_3 = \frac{\det(A_3 b)}{\det(A)} = \frac{152}{44} = \frac{38}{11}$$

Note: For any $n\times n$ matrix ***A*** and any ***b*** in $\boldsymbol{R^n}$, *let* $\boldsymbol{A_i(b)}$ be the matrix obtained from ***A*** by replacing ***i***th column by the vector ***b***.

$$A_i(b) = \begin{bmatrix} a_1 & \dots & b & \dots & a_n \end{bmatrix}$$
$$\uparrow$$
$$i\,th\,column$$

Formula for A^{-1}:

Cramer's rule leads easily to a general formula for the inverse of $n\times n$ matrix ***A***. The ***j****th* column of $\boldsymbol{A^{-1}}$ is a vector ***x*** that satisfies $\boldsymbol{Ax = e_j}$ where $\boldsymbol{e_j}$ is the ***j****th* column of the identity matrix, and the ***i****th* entry of ***x*** is the (*i, j*)-entry of $\boldsymbol{A^{-1}}$. By Cramer's rule,

$$\left\{(i,j) - entry\,of\ A^{-1}\right\} = x_{ij} = \frac{\det A_i(e_j)}{\det A} \qquad (2)$$

Recall that A_{ji} denotes the submatrix of A formed by deleting row j and column i. A cofactor expansion down column i of $A_i(e_j)$ shows that

$$\det A_i(e_j) = (-1)^{i+j} \det A_{ji} = C_{ji} \quad (3)$$

where C_{ji} is a cofactor of A.
By (2), the (i, j)-entry of A^{-1} is the cofactor C_{ji} divided by det A.
[Note that the subscripts on C_{ji} are the reverse of (i, j).] Thus

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix} \quad (4)$$

The matrix of cofactors on the right side of (4) is called the **adjugate** (or **classical adjoint**) of A, denoted by adj A. (The term adjoint also has another meaning in advance texts on linear transformations.) The next theorem simply restates (4).

Theorem 2 (An Inverse Formula):

Let A be an invertible $n\times n$ matrix, then $A^{-1} = \dfrac{1}{\det A} adj\ A$

Example:

For the matrrix say

$$A = \begin{bmatrix} 2 & 3 \\ -1 & 5 \end{bmatrix} \Rightarrow \det A = 10-(-3) = 13$$

$\Rightarrow A^{-1}$ will also be a 2×2 matrix
As
$Aji =$ submatrix of A formed by deleting row j and column i
So in this case
$A_{11} =$ submatrix of A formed by deleting row ***1*** and column ***1*** $= [5]$
$A_{12} =$ submatrix of A formed by deleting row ***1*** and column ***2*** $= [-1]$
$A_{21} =$ submatrix of A formed by deleting row ***2*** and column ***1*** $= [3]$
$A_{22} =$ submatrix of A formed by deleting row ***2*** and column ***2*** $= [2]$
and
$\det A_i(e_j) = (-1)^{i+j} \det(A_{ji}) = C_{ji}$
where e_j is the jth column of identity matrix $I_{n\times n}$
So in this case

$C_{11} = \det A_1(e_1) = (-1)^{1+1} \det A_{11} = (+1)\det[5] = 5$
$C_{12} = \det A_2(e_1) = (-1)^{1+2} \det A_{12} = (-1)\det[-1] = (-1)(-1) = 1$
$C_{21} = \det A_1(e_2) = (-1)^{2+1} \det A_{21} = (-1)\det[3] = -3$
$C_{22} = \det A_2(e_2) = (-1)^{2+2} \det A_{22} = (+1)\det[2] = 2$

By Cramer’s rule,

$$\left\{(i,j) - entry\ of\ A^{-1}\right\} = x_{ij} = \frac{\det A_i(e_j)}{\det A} = \frac{C_{ji}}{\det A}$$

So for the current matrix;

$$\left\{(1,1) - entry\ of\ A^{-1}\right\} = x_{11} = \frac{\det A_1(e_1)}{\det A} = \frac{C_{11}}{\det A} = \frac{5}{13}$$
$$\left\{(1,2) - entry\ of\ A^{-1}\right\} = x_{12} = \frac{\det A_1(e_2)}{\det A} = \frac{C_{21}}{\det A} = \frac{-3}{13}$$
$$\left\{(2,1) - entry\ of\ A^{-1}\right\} = x_{21} = \frac{\det A_2(e_1)}{\det A} = \frac{C_{12}}{\det A} = \frac{1}{13}$$
$$\left\{(2,2) - entry\ of\ A^{-1}\right\} = x_{22} = \frac{\det A_2(e_2)}{\det A} = \frac{C_{22}}{\det A} = \frac{2}{13}$$

Hence by using equation # 4, we get

$$A^{-1} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} \frac{C_{11}}{\det A} & \frac{C_{21}}{\det A} \\ \frac{C_{12}}{\det A} & \frac{C_{22}}{\det A} \end{bmatrix} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix} = \begin{bmatrix} \frac{5}{13} & \frac{-3}{13} \\ \frac{1}{13} & \frac{2}{13} \end{bmatrix}$$

Example 5: Find the inverse of the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$.

Solution: The nine cofactors are

$$C_{11} = +\begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2, \quad C_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3, \quad C_{13} = +\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14, \quad C_{22} = +\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7, \quad C_{23} = -\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = +\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4, \quad C_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1, \quad C_{33} = +\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

The adjoint matrix is the transpose of the matrix of cofactors. [For instance, C_{12} goes in the $(2, 1)$ position.] Thus

$$adjA = \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

We could compute det ***A*** directly, but the following computation provides a check on the calculations above and produces det ***A***:

$$(adjA).A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

Since (adj ***A***) ***A*** = 14 ***I***, Theorem 2 shows that det ***A*** = 14 and

$$A^{-1} = \frac{1}{14}\begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} = \begin{bmatrix} -2/14 & 14/14 & 4/14 \\ 3/14 & -7/14 & 1/14 \\ 5/14 & -7/14 & -3/14 \end{bmatrix}$$

Determinants as Area or Volume:

In the next application, we verify the geometric interpretation of determinants and we assume here that the usual Euclidean concepts of length, area, and volume are already understood for $\boldsymbol{R}^2$ and $\boldsymbol{R}^3$.

Theorem 3: If ***A*** is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|\det A|$. If ***A*** is a 3×3 matrix, the volume of the parallelepiped determined by the columns of ***A*** is $|\det A|$.

Example 6: Calculate the area of the parallelogram determined by the points (-2, -2), (0, 3), (4, -1) and (6, 4).

Solution:

Let A(-2,-2), B(0,3), C(4,-1) and D(6,4). Fixing one point say A(-2,-2) and find the adjacent lengths of parallelogram which are given by the column vectors as follows**;**

$$AB = \begin{bmatrix} 0-(-2) \\ 3-(-2) \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$AC = \begin{bmatrix} 4-(-2) \\ -1-(-2) \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

So the area of parallelogram ABCD determined by above column vectors

$$= \left|\det\begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}\right| = |2-30| = |-28| = 28$$

Now we translate the parallelogram ABCD to one having the origin as a vertex. For which we subtract the vertex (-2, -2) from each of the four vertices. The new parallelogram has the vertices say

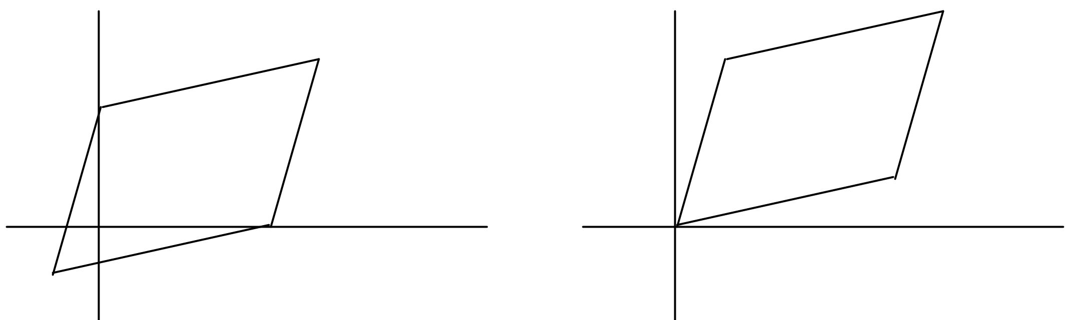
A′=(-2-(-2),-2-(-2))=(0, 0)
B′=(0-(-2),3-(-2))=(2, 5)
C′=(4-(-2),-1-(-2))=(6,1)
D′=(6-(-2),4-(-2))=(8, 6)

And fixing $A'(0,0)$ in this case, so

$$A'B' = \begin{bmatrix} 2-0 \\ 5-0 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$A'C' = \begin{bmatrix} 6-0 \\ 5-0 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \end{bmatrix}$$

See Fig below. The area of this parallelogram is also determined by the above columns vectors = $\left|\det\begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}\right| = |2-30| = |-28| = 28$



Translating a parallelogram does not change its area

Linear Transformations:

Determinants can be used to describe an important geometric property of linear transformations in the plane and in $\boldsymbol{R^3}$. If ***T*** is a linear transformation and ***S*** is a set in the domain of ***T***, let ***T*** **(*S*)** denote the set of images of points in ***S***. We are interested in how the area (or volume) of ***T*** **(*S*)** compares with the area (or volume) of the original set ***S*.** For convenience, when ***S*** is a region bounded by a parallelogram, we also refer to ***S*** as a parallelogram.

Theorem 4: Let $T: R^2 \to R^2$ be the linear transformation determined by a 2×2 matrix ***A***. If ***S*** is a parallelogram in $\boldsymbol{R^2}$, then

{area of ***T (S)***} = |det***A***|. {area of ***S***}

If ***T*** is determined by a 3 x 3 matrix ***A***, and if ***S*** is a parallelepiped in $\boldsymbol{R^3}$, then

{volume of ***T (S)***} = |det***A***|. {volume of ***S***}

Example 7: Let *a* and *b* be positive numbers. Find the area of the region *E* bounded by the ellipse whose equation is $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$.

Solution: We claim that E is the image of the unit disk D under the linear transformation **A:D→E** determined by the matrix $A=\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, given as

$A\boldsymbol{u} = \boldsymbol{x}$ where $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in D$, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in E$.

$$\Rightarrow \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Now $A\boldsymbol{u} = \boldsymbol{x} \Rightarrow \begin{bmatrix} au_1 \\ bu_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ then

$$\Rightarrow au_1 = x_1 \quad and \quad bu_2 = x_2$$

$$\Rightarrow u_1 = \frac{x_1}{a} \text{ and } u_2 = \frac{x_2}{b}$$

Since $u \in D$ (in the circular disk),it follows that the distance of $\boldsymbol{u}$ from origin will be less than unity i-e

$$\left(u_1^2 - 0\right) + \left(u_2^2 - 0\right) \le 1$$

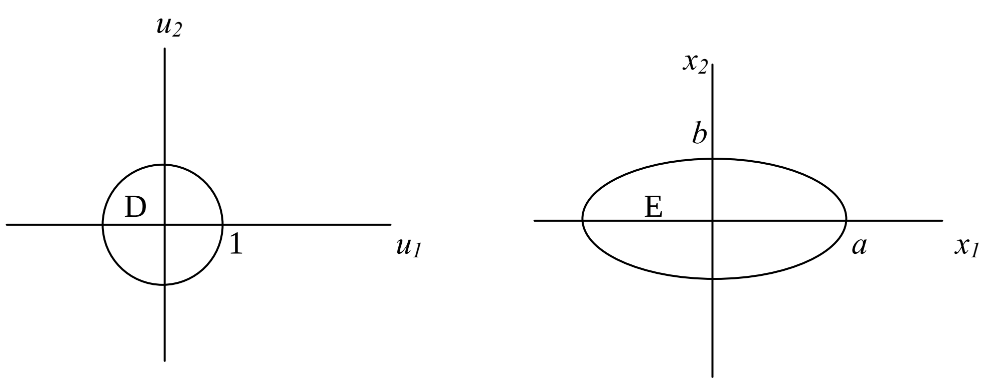
$$\Rightarrow \left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 \le 1 \qquad \because u_1 = \frac{x_1}{a}, u_2 = \frac{x_2}{b}$$

Hence by the generalization of theorem 4,

{area of ellipse} = {area of ***A(D)***} $\quad$ (*here* $T \equiv A$)

= |det ***A***|. {area of ***D***}

= $ab.\pi\,(1)^2 = \pi$ ab



Example 8: Let ***S*** be the parallelogram determined by the vectors $b_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, and let $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$. Compute the area of image of ***S*** under the mapping $x \to Ax$.

Solution: The area of ***S*** is $\left| \det \begin{bmatrix} 1 & 5 \\ 3 & 1 \end{bmatrix} \right| = 14$, and det ***A*** = 2. By theorem 4, the area of image of ***S*** under the mapping $x \to Ax$ is |det ***A***|. {area of ***S***} = 2.14 = 28

Exercises:

Use Cramer's Rule to compute the solutions of the systems in exercises 1 and 2.

1. $\begin{aligned} 2x_1 + x_2 \quad &= 7 \\ -3x_1 + \quad x_3 &= -8 \\ x_2 + 2x_3 &= -3 \end{aligned}$

2. $\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ -x_1 + \quad 2x_3 &= 2 \\ 3x_1 + x_2 + 3x_3 &= -2 \end{aligned}$

In exercises 3-6, determine the values of the parameter s for which the system has a unique solution, and describe the solution.

3. $\begin{aligned} 6sx_1 + 4x_2 &= 5 \\ 9x_1 + 2sx_2 &= -2 \end{aligned}$

4. $\begin{aligned} 3sx_1 - 5x_2 &= 3 \\ 9x_1 + 5sx_2 &= 2 \end{aligned}$

5. $\begin{aligned} sx_1 - 2sx_2 &= -1 \\ 3x_1 + 6sx_2 &= 4 \end{aligned}$

6. $\begin{aligned} 2sx_1 + \quad x_2 &= 1 \\ 3sx_1 + 6sx_2 &= 2 \end{aligned}$

In exercises 7 and 8, compute the adjoint of the given matrix, and then find the inverse of the matrix.

7. $\begin{bmatrix} 3 & 5 & 4 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix}$

8. $\begin{bmatrix} 3 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 3 & 2 \end{bmatrix}$

In exercises 9 and 10, find the area of the parallelogram whose vertices are listed.

9. *(0, 0), (5, 2), (6, 4), (11, 6)*

10. *(-1, 0), (0, 5), (1, -4), (2, 1)*

11. Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices at *(1, 0, -2), (1, 2, 4), (7, 1, 0)*.

12. Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices at *(1, 4, 0), (-2, -5, 2), (-1, 2, -1)*.

13. Let $\boldsymbol{S}$ be the parallelogram determined by the vectors $\boldsymbol{b_1} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ and $\boldsymbol{b_2} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$, and let $\boldsymbol{A} = \begin{bmatrix} 6 & -2 \\ -3 & 2 \end{bmatrix}$. Compute the area of the image of $\boldsymbol{S}$ under the mapping $\boldsymbol{x} \to \boldsymbol{Ax}$.

14. Let ***S*** be the parallelogram determined by the vectors $\boldsymbol{b_1} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$ and $\boldsymbol{b_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and let $\boldsymbol{A} = \begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix}$. Compute the area of the image of ***S*** under the mapping $\boldsymbol{x} \to \boldsymbol{Ax}$.

15. Let $\boldsymbol{T}: \boldsymbol{R^3} \to \boldsymbol{R^3}$ be the linear transformation determined by the matrix $\boldsymbol{A} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, where a, b, c are positive numbers. Let ***S*** be the unit ball, whose bounding surface has the equation $x_1^2 + x_2^2 + x_3^2 = 1$.

a. Show that ***T*** (***S***) is bounded by the ellipsoid with the equation $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1$.

b. Use the fact that the volume of the unit ball is $4\pi / 3$ to determine the volume of the region bounded by the ellipsoid in part (a).

Lecture 20

Vector Spaces and Subspaces

Case Example:

The space shuttle's control systems are absolutely critical for flight. Because the shuttle is an unstable airframe, it requires constant computer monitoring during atmospheric flight. The flight control system sends a stream of commands to aerodynamic control surfaces.

Mathematically, the input and output signals to an engineering system are functions. It is important in applications that these functions can be added, and multiplied by scalars. These two operations on functions have algebraic properties that are completely analogous to the operation of adding vectors in $\boldsymbol{R}^n$ and multiplying a vector by a scalar, as we shall see in the lectures 20 and 27. For this reason, the set of all possible inputs (functions) is called a *vector space.* The mathematical foundation for systems engineering rests on vector spaces of functions, and we need to extend the theory of vectors in $\boldsymbol{R}^n$ to include such functions. Later on, we will see how other vector spaces arise in engineering, physics, and statistics.

Definition: Let $\boldsymbol{V}$ is an arbitrary nonempty set of objects on which two operations are defined, addition and multiplication by scalars (numbers). If the following axioms are satisfied by all objects $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ in $\boldsymbol{V}$ and all scalars k and l, then we call $\boldsymbol{V}$ a vector space.

Axioms of Vector Space:

1. Closure Property For any two vectors $\boldsymbol{u}$ & $\boldsymbol{v} \in \boldsymbol{V}$, implies $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{V}$

2. Commutative Property For any two vectors $\boldsymbol{u}$ & $\boldsymbol{v} \in \boldsymbol{V}$, implies $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

3. Associative Property For any three vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{V}$, $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$

4. Additive Identity For any vector $\boldsymbol{u} \in \boldsymbol{V}$, there exist a zero vector $\boldsymbol{0}$ such that

$$\boldsymbol{0} + \boldsymbol{u} = \boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$$

5. Additive Inverse For each vector $\boldsymbol{u} \in \boldsymbol{V}$, there exist a vector $-\boldsymbol{u}$ in $\boldsymbol{V}$ such that

$$-\boldsymbol{u} + \boldsymbol{u} = \boldsymbol{0} = \boldsymbol{u} + (-\boldsymbol{u})$$

6. Scalar Multiplication For any scalar k and a vector $\boldsymbol{u} \in \boldsymbol{V}$ implies k $\boldsymbol{u} \in \boldsymbol{V}$

7. Distributive Law For any scalar k if $\boldsymbol{u}$ & $\boldsymbol{v} \in \boldsymbol{V}$, then $k\,(\boldsymbol{u} + \boldsymbol{v}) = k\,\boldsymbol{u} + k\,\boldsymbol{v}$

8. For scalars m, n and for any vector $\boldsymbol{u} \in \boldsymbol{V}$, $(m + n)\,\boldsymbol{u} = m\,\boldsymbol{u} + n\,\boldsymbol{u}$

9. For scalars m, n and for any vector $\boldsymbol{u} \in \boldsymbol{V}$, $m\,(n\,\boldsymbol{u}) = (m\,n)\,\boldsymbol{u} = n\,(m\,\boldsymbol{u})$

10. For any vector $\boldsymbol{u} \in \boldsymbol{V}$, $1\boldsymbol{u} = \boldsymbol{u}$ where 1 is the multiplicative identity of real numbers.

Examples of vector spaces: The following examples will specify a non empty set $\boldsymbol{V}$ and two operations: addition and scalar multiplication; then we shall verify that the ten vector space axioms are satisfied.

Example 1: Show that the set of all ordered n-tuple $\boldsymbol{R^n}$ is a vector space under the standard operations of addition and scalar multiplication.

Solution

(i) **Closure Property:**

Suppose that $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \boldsymbol{R^n}$

Then by definition, $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)$

$= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \in \boldsymbol{R^n}$ (By closure property)

Therefore, $\boldsymbol{R^n}$ is closed under addition.

(ii) **Commutative Property**

Suppose that $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \boldsymbol{R^n}$

Now $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)$

$= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$ (By closure property)

$= (v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n)$ (By commutative law of real numbers)

$= (v_1, v_2, \ldots, v_n) + (u_1, u_2, \ldots, u_n)$ (By closure property)

$= \boldsymbol{v} + \boldsymbol{u}$

Therefore, $\boldsymbol{R^n}$ is commutative under addition.

(iii) **Associative Property**

Suppose that $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ *and* $\boldsymbol{w} = (w_1, w_2, \ldots, w_n) \in \boldsymbol{R^n}$

Now $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = [(u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)] + (w_1, w_2, \ldots, w_n)$

$= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) + (w_1, w_2, \ldots, w_n)$ (By closure property)

$= ((u_1 + v_1) + w_1, (u_2 + v_2) + w_2, \ldots, (u_n + v_n) + w_n))$ (By closure property)

$= (u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), \ldots, u_n + (v_n + w_n))$ (By associative law of real numbers)

$= (u_1, u_2, \ldots, u_n) + (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$ (By closure property)

$= (u_1, u_2, \ldots, u_n) + [(v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n)]$ (By closure property)

$= \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$

Hence $\boldsymbol{R}^n$ is associative under addition.

(iv) **Additive Identity**

Suppose $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R}^n$. There exists $\boldsymbol{0} = (0, 0, \ldots, 0) \in \boldsymbol{R}^n$ such that
$\boldsymbol{0} + \boldsymbol{u} = (0, 0, \ldots, 0) + (u_1, u_2, \ldots, u_n)$

$= (0 + u_1, 0 + u_2, \ldots, 0 + u_n)$ (By closure property)

$= (u_1, u_2, \ldots, u_n) = \boldsymbol{u}$ (Existence of identity of real numbers)

Similarly, $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$

Hence $\boldsymbol{0} = (0, 0, \ldots, 0)$ is the additive identity for $\boldsymbol{R}^n$.

(v) **Additive Inverse**

Suppose $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R}^n$. There exists $-\boldsymbol{u} = (-u_1, -u_2, \ldots, -u_n) \in \boldsymbol{R}^n$

Such that $\boldsymbol{u} + (-\boldsymbol{u}) = (u_1, u_2, \ldots, u_n) + (-u_1, -u_2, \ldots, -u_n)$

$= (u_1 + (-u_1), u_2 + (-u_2), \ldots, u_n + (-u_n))$ (By closure property)

$= (0, 0, \ldots, 0) = \boldsymbol{0}$

Similarly, $(-\boldsymbol{u}) + \boldsymbol{u} = \boldsymbol{0}$

Hence the inverse of each element of $\boldsymbol{R}^n$ exists in $\boldsymbol{R}^n$.

(vi) **Scalar Multiplication**

If k is any scalar and $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R}^n$.

Then by definition, $k\,\boldsymbol{u} = k\,(u_1, u_2, \ldots, u_n) = (k\,u_1, k\,u_2, \ldots, k\,u_n) \in \boldsymbol{R}^n$
(By closure property)

(vii) **Distributive Law**

Suppose k is any scalar and $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \boldsymbol{R}^n$

Now $k\,(\boldsymbol{u} + \boldsymbol{v}) = k\,[(u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)]$

$= k\,(u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$ (By closure property)

$= (k\,(u_1 + v_1), k\,(u_2 + v_2), \ldots, k\,(u_n + v_n))$ (By scalar multiplication)

$= (k\,u_1 + k\,v_1, k\,u_2 + k\,v_2, \ldots, k\,u_n + k\,v_n)$ (By Distributive Law)

$= (k\,u_1, k\,u_2, \ldots, k\,u_n) + (k\,v_1, k\,v_2, \ldots, k\,v_n)$ (By closure property)

$= k\,(u_1, u_2, \ldots, u_n) + k\,(v_1, v_2, \ldots, v_n)$ (By scalar multiplication)

$= k\,\boldsymbol{u} + k\,\boldsymbol{v}$

(viii) Suppose k and l be any scalars and $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R}^n$

Then $(k + l)\,\boldsymbol{u} = (k + l)\,(u_1, u_2, \ldots, u_n)$

$= ((k + l)u_1, (k + l)u_2, \ldots, (k + l)u_n)$ (By scalar multiplication)

$= (k\,u_1 + l\,u_1, k\,u_2 + l\,u_2, \ldots, k\,u_n + l\,u_n)$ (By Distributive Law)

$= (k\,u_1, k\,u_2, \ldots, k\,u_n) + (l\,u_1, l\,u_2, \ldots, l\,u_n)$ (By closure property)

$= k\,(u_1, u_2, \ldots, u_n) + l\,(u_1, u_2, \ldots, u_n)$ (By scalar multiplication)

$= k\,\boldsymbol{u} + l\,\boldsymbol{u}$

(ix) Suppose k and l be any scalars and $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R}^n$

Then $k\,(l\,\boldsymbol{u}) = k\,[l\,(u_1, u_2, \ldots, u_n)]$

$= k\,(l\,u_1, l\,u_2, \ldots, l\,u_n)$ (By scalar multiplication)

$= (k\,(l\,u_1), k\,(l\,u_2), \ldots, k\,(l\,u_n))$ (By scalar multiplication)

$= ((k\,l)u_1, (k\,l)u_2, \ldots, (k\,l)u_n)$ (By associative law)

$= (k\,l)\,(u_1, u_2, \ldots, u_n)$ (By scalar multiplication)

$= (k\,l)\,\boldsymbol{u}$

(x) Suppose $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R}^n$

Then $1\,\boldsymbol{u} = 1\,(u_1, u_2, \ldots, u_n)$

$= (1u_1, 1u_2, \ldots, 1u_n)$ (By scalar multiplication)

$= (u_1, u_2, \ldots, u_n) = \boldsymbol{u}$ (Existence of identity in scalrs)

Hence, $\boldsymbol{R^n}$ is the real vector space with the standard operations of addition and scalar multiplication.

Note: The three most important special cases of $\boldsymbol{R^n}$ are $\boldsymbol{R}$ (the real numbers), $\boldsymbol{R^2}$ (the vectors in the plane), and $\boldsymbol{R^3}$ (the vectors in 3-space).

Example 2: Show that the set $\boldsymbol{V}$ of all 2x2 matrices with real entries is a vector space if vector addition is defined to be matrix addition and vector scalar multiplication is defined to be matrix scalar multiplication.

Solution: Suppose that $\boldsymbol{u} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \in \boldsymbol{V}$ and k and l be two any scalars.

(i) **Closure property** To prove axiom (i), we must show that $\boldsymbol{u} + \boldsymbol{v}$ is an object in $\boldsymbol{V}$: that is , we must show that $\boldsymbol{u} + \boldsymbol{v}$ is a 2x2 matrix. But this follows from the definition of matrix addition, since $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix}$

(By closure property)

(ii) **Commutative property** Now it is very easy to verify the Axiom (ii)

$$\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix} \quad \text{(By closure property)}$$

$$= \begin{bmatrix} v_{11}+u_{11} & v_{12}+u_{12} \\ v_{21}+u_{21} & v_{22}+u_{22} \end{bmatrix} \quad \text{(Commutative property of real numbers)}$$

$$= \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \boldsymbol{v} + \boldsymbol{u}$$

(iii) **Associative property** $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \left(\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \right) + \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}$

$$= \begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix} + \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \quad \text{(By closure property)}$$

$$= \begin{bmatrix} (u_{11}+v_{11})+w_{11} & (u_{12}+v_{12})+w_{12} \\ (u_{21}+v_{21})+w_{21} & (u_{22}+v_{22})+w_{22} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11}+(v_{11}+w_{11}) & u_{12}+(v_{12}+w_{12}) \\ u_{21}+(v_{21}+w_{21}) & u_{22}+(v_{22}+w_{22}) \end{bmatrix} \quad \text{(By associative property of real numbers)}$$

$$= \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11}+w_{11} & v_{12}+w_{12} \\ v_{21}+w_{21} & v_{22}+w_{22} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \left(\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} + \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \right) = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$$

Therefore, ***V*** is associative under '+'.

(iv) **Additive Identity** Now to prove the axiom (iv), we must find an object ***0*** in ***V*** such that ***0*** + ***v*** = ***v*** + ***0*** = ***v*** for all ***u*** in ***V***. This can be done by defining $\boldsymbol{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

$$\boldsymbol{0} + \boldsymbol{u} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} 0+u_{11} & 0+u_{12} \\ 0+u_{21} & 0+u_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \boldsymbol{u}$$

and similarly ***u*** + ***0*** = ***u***.

(v) **Additive Inverse** Now to prove the axiom (v) we must show that each object ***u*** in ***V*** has a negative **–*u*** such that ***u*** + (-***u***) = **0** = (-***u***) + **0**. Defining the negative of ***u*** to be

$$-\boldsymbol{u} = \begin{bmatrix} -u_{11} & -u_{12} \\ -u_{21} & -u_{22} \end{bmatrix}.$$

$$\boldsymbol{u} + (-\boldsymbol{u}) = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} -u_{11} & -u_{12} \\ -u_{21} & -u_{22} \end{bmatrix} = \begin{bmatrix} u_{11}+(-u_{11}) & u_{12}+(-u_{12}) \\ u_{21}+(-u_{21}) & u_{22}+(-u_{22}) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \boldsymbol{0}$$

Similarly, (-***u***) + ***u*** = ***0***

(vi) **Scalar Multiplication**

Axiom (vi) also holds because for any real number *k* we have

$$k\boldsymbol{u} = k \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix} \qquad \text{(By closure property)}$$

so that k ***u*** is a 2x2 matrix and consequently is an object in ***V***.

(vii) **Distributive Law:**

$$k(\boldsymbol{u}+\boldsymbol{v}) = k\left(\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \right)$$

$$= k\begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix} = \begin{bmatrix} k(u_{11}+v_{11}) & k(u_{12}+v_{12}) \\ k(u_{21}+v_{21}) & k(u_{22}+v_{22}) \end{bmatrix}$$

$$= \begin{bmatrix} ku_{11}+kv_{11} & ku_{12}+kv_{12} \\ ku_{21}+kv_{21} & ku_{22}+kv_{22} \end{bmatrix} = \begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix} + \begin{bmatrix} kv_{11} & kv_{12} \\ kv_{21} & kv_{22} \end{bmatrix}$$

$$= k\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + k\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = k\boldsymbol{u} + k\boldsymbol{v}$$

(viii) $(k + l)\boldsymbol{u} = (k + l)\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} (k+l)u_{11} & (k+l)u_{12} \\ (k+l)u_{21} & (k+l)u_{22} \end{bmatrix}$

$$= \begin{bmatrix} ku_{11}+lu_{11} & ku_{12}+lu_{12} \\ ku_{21}+lu_{21} & ku_{22}+lu_{22} \end{bmatrix} = \begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix} + \begin{bmatrix} lu_{11} & lu_{12} \\ lu_{21} & lu_{22} \end{bmatrix}$$

$$= k\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + l\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = k\boldsymbol{u} + l\boldsymbol{u}$$

(ix) $k(l\boldsymbol{u}) = k\left(l\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}\right) = k\begin{bmatrix} lu_{11} & lu_{12} \\ lu_{21} & lu_{22} \end{bmatrix}$

$$= \begin{bmatrix} k(lu_{11}) & k(lu_{12}) \\ k(lu_{21}) & k(lu_{22}) \end{bmatrix} = \begin{bmatrix} (kl)u_{11} & (kl)u_{12} \\ (kl)u_{21} & (kl)u_{22} \end{bmatrix} = (kl)\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = (kl)\boldsymbol{u}$$

(x) Finally axiom (x) is a simple computation

$$1\boldsymbol{u} = 1\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} 1u_{11} & 1u_{12} \\ 1u_{21} & 1u_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \boldsymbol{u}$$

Hence the set of all 2x2 matrices with real entries is vector space under matrix addition and matrix scalar multiplication.

Note: Example 2 is a special case of a more general class of vector spaces. The arguments in that example can be adapted to show that a set $\boldsymbol{V}$ of all $\boldsymbol{m}\times\boldsymbol{n}$ matrices with real entries, together with the operations of matrix addition and scalar multiplication, is a vector space.

Example 3: Let $\boldsymbol{V}$ be the set of all real-valued functions defined on the entire real line $(-\infty, \infty)$. If $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{V}$, then $\boldsymbol{f} + \boldsymbol{g}$ is a function defined by

$$(\boldsymbol{f} + \boldsymbol{g})(x) = \boldsymbol{f}(x) + \boldsymbol{g}(x), \text{ for all } x \in \boldsymbol{R}.$$

The product of a scalar $a \in R$ and a function $\boldsymbol{f}$ in $\boldsymbol{V}$ is defined by

$$(a\,\boldsymbol{f})(x) = a\,\boldsymbol{f}(x), \text{ for all } x \in \boldsymbol{R}.$$

Solution:

(i) **Closure Property** If $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{V}$, then by definition

$(\boldsymbol{f} + \boldsymbol{g})(x) = \boldsymbol{f}(x) + \boldsymbol{g}(x) \in \boldsymbol{V}$. Therefore, $\boldsymbol{V}$ is closed under addition.

(ii) **Commutative Property** If $\boldsymbol{f}$ and $\boldsymbol{g}$ are in $\boldsymbol{V}$, then for all $x \in \boldsymbol{R}$

$(\boldsymbol{f} + \boldsymbol{g})(x) = \boldsymbol{f}(x) + \boldsymbol{g}(x)$ (By definition)

$= \boldsymbol{g}(x) + \boldsymbol{f}(x)$ (By commutative property)

$= (\boldsymbol{g} + \boldsymbol{f})(x)$ (By definition)

So that $\boldsymbol{f} + \boldsymbol{g} = \boldsymbol{g} + \boldsymbol{f}$

(iii) **Associative Property** If $\boldsymbol{f}$, $\boldsymbol{g}$ and $\boldsymbol{h}$ are in $\boldsymbol{V}$, then for all $x \in \boldsymbol{R}$

$((\boldsymbol{f} + \boldsymbol{g}) + \boldsymbol{h})(x) = (\boldsymbol{f} + \boldsymbol{g})(x) + \boldsymbol{h}(x)$ (By definition)

$= (\boldsymbol{f}(x) + \boldsymbol{g}(x)) + \boldsymbol{h}(x)$ (By definition)

$= \boldsymbol{f}(x) + (\boldsymbol{g}(x) + \boldsymbol{h}(x))$ (By associative property)

$= \boldsymbol{f}(x) + (\boldsymbol{g} + \boldsymbol{h})(x)$ (By definition)

$= (\boldsymbol{f} + (\boldsymbol{g} + \boldsymbol{h}))(x)$

And so $(\boldsymbol{f} + \boldsymbol{g}) + \boldsymbol{h} = \boldsymbol{f} + (\boldsymbol{g} + \boldsymbol{h})$

(iv) **Additive Identity** The additive identity of $\boldsymbol{V}$ is the zero function defined by $\boldsymbol{0}(x) = \boldsymbol{0}$, for all $x \in \boldsymbol{R}$ because $(\boldsymbol{0} + \boldsymbol{f})(x) = \boldsymbol{0}(x) + \boldsymbol{f}(x)$ (By definition)

$= \boldsymbol{0} + \boldsymbol{f}(x) = \boldsymbol{f}(x)$ (Existence of identity)

i.e. $\boldsymbol{0} + \boldsymbol{f} = \boldsymbol{f}$. Similarly, $\boldsymbol{f} + \boldsymbol{0} = \boldsymbol{f}$.

(v) **Additive Inverse** The additive inverse of a function $\boldsymbol{f}$ in $\boldsymbol{V}$ is $(-1)\,\boldsymbol{f} = -\boldsymbol{f} \in \boldsymbol{V}$ because

$(\boldsymbol{f} + (-\boldsymbol{f}))(x) = \boldsymbol{f}(x) + (-\boldsymbol{f})(x)$ (By definition)

$= \boldsymbol{f}(x) - \boldsymbol{f}(x)$ (By definition)

$= \boldsymbol{0}$ (Existence of inverse) i.e. $\boldsymbol{f} + (-\boldsymbol{f}) = \boldsymbol{0}$. Similarly, $(-\boldsymbol{f}) + \boldsymbol{f} = \boldsymbol{0}$.

(vi) **Scalar Multiplication** If $\boldsymbol{f}$ is in $\boldsymbol{V}$ and a is in R, then by definition $(a\,\boldsymbol{f})(\boldsymbol{x}) = a\,\boldsymbol{f}(\boldsymbol{x}) \in \boldsymbol{V}$.

(vii) **Distributive Law** If $\boldsymbol{f}, \boldsymbol{g}$ are in $\boldsymbol{V}$ and $a \in R$, then
$(a(\boldsymbol{f} + \boldsymbol{g}))(\boldsymbol{x}) = a(\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{x}) = a(\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})) = a\,\boldsymbol{f}(\boldsymbol{x}) + a\,\boldsymbol{g}(\boldsymbol{x})$

$= (a\,\boldsymbol{f})(\boldsymbol{x}) + (a\,\boldsymbol{g})(\boldsymbol{x}) = (a\,\boldsymbol{f} + a\,\boldsymbol{g})(\boldsymbol{x})$ and, therefore, $a(\boldsymbol{f} + \boldsymbol{g}) = a\,\boldsymbol{f} + a\,\boldsymbol{g}$

(viii) Let a, b in R and $\boldsymbol{f} \in \boldsymbol{V}$, then

$((a + b)\boldsymbol{f})(\boldsymbol{x}) = (a + b)\,\boldsymbol{f}(\boldsymbol{x}) = a\,\boldsymbol{f}(\boldsymbol{x}) + b\,\boldsymbol{f}(\boldsymbol{x}) = (a\,\boldsymbol{f})(\boldsymbol{x}) + (b\,\boldsymbol{f})(\boldsymbol{x}) = (a\,\boldsymbol{f} + b\,\boldsymbol{f})(\boldsymbol{x})$

Thus $(a + b)\,\boldsymbol{f} = a\,\boldsymbol{f} + b\,\boldsymbol{f}$

(ix) $a(b\,\boldsymbol{f})(\boldsymbol{x}) = a(b\,\boldsymbol{f}(\boldsymbol{x})) = (a\,b)\,\boldsymbol{f}(\boldsymbol{x})$ showing that $a(b\,\boldsymbol{f}) = (a\,b)\,\boldsymbol{f}$

(x) $(1.\boldsymbol{f})(\boldsymbol{x}) = 1\,\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$ (Existence of identity)

And so $1.\boldsymbol{f} = \boldsymbol{f}$

Hence $\boldsymbol{V}$ is a real vector space.

Example 4: If $\boldsymbol{p}(\boldsymbol{x}) = a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + \ldots + a_n\boldsymbol{x}^n$
and $\boldsymbol{q}(\boldsymbol{x}) = b_0 + b_1\boldsymbol{x} + b_2\boldsymbol{x}^2 + \ldots + b_n\boldsymbol{x}^n$
We define

$$\boldsymbol{p}(\boldsymbol{x}) + \boldsymbol{q}(\boldsymbol{x}) = (a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + \ldots + a_n\boldsymbol{x}^n) + (b_0 + b_1\boldsymbol{x} + b_2\boldsymbol{x}^2 + \ldots + b_n\boldsymbol{x}^n)$$

$= (a_0 + b_0) + (a_1 + b_1)\boldsymbol{x} + (a_2 + b_2)\boldsymbol{x}^2 + \ldots + (a_n + b_n)\boldsymbol{x}^n$ and for any scalar k,

$$k\boldsymbol{p}(\boldsymbol{x}) = k(a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + \ldots + a_n\boldsymbol{x}^n) = ka_0 + ka_1\boldsymbol{x} + ka_2\boldsymbol{x}^2 + \ldots + k\,a_n\boldsymbol{x}^n$$

Clearly the given polynomial is a vector space under the addition and scalar multiplication.

Example 5: **(The Zero Vector Space)** Let $\boldsymbol{V}$ consists of a single object, which we define by $\boldsymbol{0}$ and $\boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$ and k $\boldsymbol{0} = \boldsymbol{0}$ for all scalars k. It is easy to check that all the vector space axioms are satisfied. We call $\boldsymbol{V} = \{\boldsymbol{0}\}$ as the zero vector space.

Example 6: (Every plane through the origin is a vector space)

Let ***V*** be any plane through the origin in $\boldsymbol{R^3}$. We shall show that the points in ***V*** form a vector space under a standard addition and scalar multiplication operations for vectors in $\boldsymbol{R^3}$.

From example1, we know that $\boldsymbol{R^3}$ itself is a vector space under these operations. Thus, Axioms 2, 3, 7, 8, 9 and 10 hold for all points in $\boldsymbol{R^3}$ and consequently for all points in the plane ***V***. We therefore need only show that Axioms 1, 4, 5 and 6 are satisfied.

Since the plane is passing through the origin, it has an equation of the form

$$a x + b y + c z = 0 \qquad (1)$$

Thus, if $\boldsymbol{u} = (u_1, u_2, u_3)$ and $\boldsymbol{v} = (v_1, v_2, v_3)$ are points in ***V***, then

$$a u_1 + b u_2 + c u_3 = 0 \text{ and } a v_1 + b v_2 + c v_3 = 0.$$

Adding these equations gives $a (u_1 + v_1) + b (u_2 + v_2) + c (u_3 + v_3) = 0$

This equality tell us that the coordinates of the point

$$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3)$$

satisfies (1); thus, $\boldsymbol{u} + \boldsymbol{v}$ lies in plane ***V***. This proves that the Axiom 1 is satisfied.

There exists $\boldsymbol{0} = (0, 0, 0)$ such that $a (0) +b (0) + c (0) = 0$. Therefore, Axiom 4 is satisfied.

Multiplying $a u_1 + b u_2 + c u_3 = 0$ through by k gives

$$a (ku_1) + b (ku_2) + c (ku_3) = 0$$

Thus, $(ku_1, k u_2, k u_3) = k (u_1, u_2, u_3) = k \boldsymbol{u} \in \boldsymbol{V}$. Hence, Axiom 6 is satisfied.

We shall prove the axiom 5 is satisfied. Multiplying $a u_1 + b u_2 + c u_3 = 0$ through by -1 gives a $(-1u_1) + b (-1u_2) + c (-1u_3) = 0$

Thus, $(-u_1, - u_2, - u_3) = - (u_1, u_2, u_3) = -\boldsymbol{u} \in \mathbf{V}$. This establishes Axiom 5.

<u>Example 7:</u> (A set that is not a vector space)

Let $\boldsymbol{V}=\boldsymbol{R^2}$ and define addition and scalar multiplication operation as follows. If $\boldsymbol{u} = (u_1, u_2) \textit{ and } \boldsymbol{v} = (v_1, v_2)$ then define

$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2)$ and if k is any real number then define $k\boldsymbol{u} = (ku_1, 0)$.

For any vector $\boldsymbol{u} \in \boldsymbol{V}$, $1\boldsymbol{u} = 1(u_1, u_2) = (1\, u_1, 0) = (u_1, 0) \neq \boldsymbol{u}$ where 1 is the multiplicative identity of real numbers. Therefore, the axiom 10 is not satisfied.

Hence, $\boldsymbol{V}=\boldsymbol{R^2}$ is not a vector space.

<u>Theorem 1:</u> Let ***V*** be a vector space, ***u*** a vector in ***V***, and *k* is a scalar, then

(i) $0\boldsymbol{u} = \boldsymbol{0}$ (ii) $k\boldsymbol{0} = \boldsymbol{0}$

(iii) $(-1)\,\boldsymbol{u} = -\boldsymbol{u}$ (iv) If $k\,\boldsymbol{u} = \boldsymbol{0}$ then $k = 0$ or $\boldsymbol{u} = \boldsymbol{0}$

Definition: A subset ***W*** of a vector space ***V*** is called a subspace of ***V*** if ***W*** itself a vector space under the addition and scalar multiplication defined on ***V***.

Note: If ***W*** is a part of a larger set ***V*** that is already known to be a vector space, then certain axioms need not be verified for ***W*** because they are "inherited" from ***V***. For example, there is no need to check that $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ (Axiom 2) for ***W*** because this holds for all vectors in ***V*** and consequently for all vectors in ***W***. Other Axioms are inherited by ***W*** from ***V*** are 3, 7, 8, 9, and 10. Thus, to show that a set ***W*** is a subspace of a vector space ***V***, we need only verify Axioms 1, 4, 5 and 6. The following theorem shows that even Axioms 4 and 5 can be omitted.

Theorem 2: If ***W*** is a set of one or more vectors from a vector space ***V***, then ***W*** is subspace of ***V*** if and only if the following conditions hold.

(a) If ***u*** and ***v*** are vectors in ***W***, then $\boldsymbol{u} + \boldsymbol{v}$ is in ***W***
(b) If k is any scalar and ***u*** is any vector in ***W***, then $k\,\boldsymbol{u}$ is in ***W***.

Proof: If ***W*** is a subspace of ***V***, then all the vector space axioms are satisfied; in particular, Axioms 1 and 6 hold. But these are precisely conditions (a) and (b).

Conversely, assume conditions (a) and (b) hold. Since these conditions are vector space Axioms 1 and 6, we need only show that ***W*** satisfies the remaining 8 axioms. The vectors in ***W*** automatically satisfy axioms 2, 3, 7, 8, 9, and 10 since they are satisfied by all vectors in ***V***. Therefore, to complete the proof, we need only verify that vectors in ***W*** satisfy axioms 4 and 5.

Let ***u*** be any vector in ***W***. By condition (b), $k\,\boldsymbol{u}$ is in ***W*** for every scalar k. Setting $k = 0$, it follows from theorem 1 that $0\,\boldsymbol{u} = \boldsymbol{0}$ is in ***W***, and setting $k = -1$, it follows that $(-1)\,\boldsymbol{u} = -\boldsymbol{u}$ is in ***W***. □

Remark:

(1) The theorem states that ***W*** is a subspace of ***V*** if and only if ***W*** is closed under addition and closed under scalar multiplication.

(2) Every vector space has at least two subspaces, itself and the subspace $\{\boldsymbol{0}\}$ consisting only of the zero vector. Thus the subspace $\{\boldsymbol{0}\}$ is called the zero subspace.

Example 8: Let ***W*** be the subset of $\boldsymbol{R}^3$ consisting of the all the vectors of the form $(a, b, 0)$, where a and b are real numbers. To check if ***W*** is subspace of $\boldsymbol{R}^3$, we first see that axiom 1 and 6 of a vector space holds.

Let $\boldsymbol{u} = (a_1, b_1, 0)$ and $\boldsymbol{v} = (a_2, b_2, 0)$ be vectors in ***W*** then $\boldsymbol{u} + \boldsymbol{v} = (a_1, b_1, 0) + (a_2, b_2, 0) = (a_1 + a_2, b_1 + b_2, 0)$ is in ***W***. Since the third component is zero. Also c is scalar, and then $c\boldsymbol{u} = c(a_1, b_1, 0) = (ca_1, cb_1, 0)$ is in ***W***. Therefore the I[st] and

6[th] axioms of the vector space holds. We can also verify the other axioms of vector space. Hence ***W*** is a subspace.

Example 9: Consider the set ***W*** consisting of all 2x3 matrices of the form $\begin{bmatrix} a & b & 0 \\ 0 & c & d \end{bmatrix}$, Where *a*, *b*, *c* and *d* are arbitrary real numbers. Show that the ***W*** is a subspace M_{2x3}.

Solution: Consider $\mathbf{u} = \begin{bmatrix} a_1 & b_1 & 0 \\ 0 & c_1 & d_1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} a_2 & b_2 & 0 \\ 0 & c_2 & d_2 \end{bmatrix}$ in ***W***

Then $\mathbf{u} + \mathbf{v} = \begin{bmatrix} a_1 & b_1 & 0 \\ 0 & c_1 & d_1 \end{bmatrix} + \begin{bmatrix} a_2 & b_2 & 0 \\ 0 & c_2 & d_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 & b_1 + b_2 & 0 \\ 0 & c_1 + c_2 & d_1 + d_2 \end{bmatrix}$ is in ***W***.

So that the (a) part of the theorem is satisfied. Also *k* is a scalar, and then

$k\mathbf{u} = \begin{bmatrix} ka_1 & kb_1 & 0 \\ 0 & kc_1 & kd_1 \end{bmatrix}$ is in ***W***. So the (b) part of the above theorem is also satisfied.

Hence ***W*** is a subspace of M_{2x3}.

Note: Let ***V*** is a vector space then every subset of ***V*** is not necessary a subspace of ***V***. For example, let $V = R^2$ then any line in R^2 not passing through origin is not a subspace of R^2. Similarly, a plane in R^3 not passing through the origin is not a subspace of R^3.

Example 10: Let ***W*** be the subset of R^3 consisting of all vectors of the form *(a, b, 1)*, where *a*, *b* are any real numbers. To check whether property (a) and (b) of the above theorem holds. Let $\mathbf{u} = (a_1, b_1, 1)$ and $\mathbf{v} = (a_2, b_2, 1)$ be vectors in ***W***.

Then $\mathbf{u} + \mathbf{v} = (a_1, b_1, 1) + (a_2, b_2, 1) = (a_1 + a_2, b_1 + b_2, 1 + 1)$ which is not in ***W*** because the third component 2 is not 1. As the Ist property does not hold therefore, the given set of vectors is not a vector space.

Example 11: Which of the following are subspaces of R^3

(i) All vectors of the form *(a, 0, 0)*

(ii) All vectors of the form *(a, 1, 1)*

(iii) All vectors of the form *(a, b, c)*, where $b = a + c$

(iv) All vectors of the form *(a, b, c)*, where $b = a + c + 1$

Solution: Let ***W*** is the set of all vectors of the form *(a, 0, 0)*.

(i) Suppose $\mathbf{u} = (u_1, 0, 0)$ and $\mathbf{v} = (v_1, 0, 0)$ are in ***W***.

Then $\mathbf{u} + \mathbf{v} = (u_1, 0, 0) + (v_1, 0, 0) = (u_1 + v_1, 0, 0)$ which is of the form *(a, 0, 0)*.

Therefore, $\mathbf{u} + \mathbf{v} \in W$

If *k* is any scalar and $\mathbf{u} = (u_1, 0, 0)$ is any vector in ***W***, then $k\,\mathbf{u} = k\,(u_1, 0, 0) = (k\,u_1, 0, 0)$

which is of the form $(a, 0, 0)$. Therefore, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{R^3}$.

(ii) Let $\boldsymbol{W}$ is the set of all vectors of the form $(a, 1, 1)$.

Suppose $\boldsymbol{u} = (u_1, 1, 1)$ and $\boldsymbol{v} = (v_1, 1, 1)$ are in $\boldsymbol{W}$. Then $\boldsymbol{u} + \boldsymbol{v} = (u_1, 1, 1) + (v_1, 1, 1) = (u_1 + v_1, 2, 2)$ which is not of the form $(a, 1, 1)$. Therefore, $\boldsymbol{u} + \boldsymbol{v} \notin \boldsymbol{W}$. Hence $\boldsymbol{W}$ is not the subspace of $\boldsymbol{R^3}$.

(iii) Suppose $\boldsymbol{W}$ is the set of all vectors of the form (a, b, c), where $b = a + c$

Suppose $\boldsymbol{u} = (u_1, u_1 + u_3, u_3)$ and $\boldsymbol{v} = (v_1, v_1 + v_3, v_3)$ are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_1 + u_3, u_3) + (v_1, v_1 + v_3, v_3)$
$= (u_1 + v_1, u_1 + u_3 + v_1 + v_3, u_3 + v_3)$
$= (u_1 + v_1, (u_1 + v_1) + (u_3 + v_3), u_3 + v_3)$, which is of the form $(a, a + c, c)$.

Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = (u_1, u_1 + u_3, u_3)$ is any vector in $\boldsymbol{W}$, then
$k\,\boldsymbol{u} = k\,(u_1, u_1 + u_3, u_3) = (k\,u_1, k\,(u_1 + u_3), k\,u_3)$ (By definition)
$= (k\,u_1, k\,u_1 + k\,u_3, k\,u_3)$ (By Distributive Law)
Which is of the form $(a, a + c, c)$. Therefore, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{R^3}$.

(iv) Let $\boldsymbol{W}$ is the set of all vectors of the form (a, b, c), where $b = a + c + 1$
Suppose $\boldsymbol{u} = (u_1, u_1 + u_3 + 1, u_3)$ and $\boldsymbol{v} = (v_1, v_1 + v_3 + 1, v_3)$ are in $\boldsymbol{W}$.
Then $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_1 + u_3 + 1, u_3) + (v_1, v_1 + v_3 + 1, v_3)$
$= (u_1 + v_1, u_1 + u_3 + 1 + v_1 + v_3 + 1, u_3 + v_3)$
$= (u_1 + v_1, (u_1 + v_1) + (u_3 + v_3) + 2, u_3 + v_3)$
Which is not of the form $(a, a + c + 1, c)$. Therefore, $\boldsymbol{u} + \boldsymbol{v} \notin \boldsymbol{W}$. Hence $\boldsymbol{W}$ is not the subspace of $\boldsymbol{R^3}$.

Example 12: Determine which of the following are subspaces of $\boldsymbol{P_3}$.
(i) All polynomials $a_0 + a_1\,\boldsymbol{x} + a_2\,\boldsymbol{x^2} + a_3\,\boldsymbol{x^3}$ for which $a_0 = 0$
(ii) All polynomials $a_0 + a_1\,\boldsymbol{x} + a_2\,\boldsymbol{x^2} + a_3\,\boldsymbol{x^3}$ for which $a_0 + a_1 + a_2 + a_3 = 0$
(iii) All polynomials $a_0 + a_1\,\boldsymbol{x} + a_2\,\boldsymbol{x^2} + a_3\,\boldsymbol{x^3}$ for which $a_0, a_1, a_2,$ and a_3 are integers
(iv) All polynomials of the form $a_0 + a_1\,\boldsymbol{x}$, where a_0 and a_1 are real numbers.

Solution: (i) Let $\boldsymbol{W}$ is the set of all polynomials $a_0 + a_1\,\boldsymbol{x} + a_2\,\boldsymbol{x^2} + a_3\,\boldsymbol{x^3}$ for which $a_0 = 0$.
Suppose that $\boldsymbol{u} = c_0 + c_1\,\boldsymbol{x} + c_2\,\boldsymbol{x^2} + c_3\,\boldsymbol{x^3}$ (where $c_0 = 0$) and $\boldsymbol{v} = b_0 + b_1\,\boldsymbol{x} + b_2\,\boldsymbol{x^2} + b_3\,\boldsymbol{x^3}$ (where $b_0 = 0$) are in $\boldsymbol{W}$. Then $\boldsymbol{u} + \boldsymbol{v} = (c_0 + c_1\boldsymbol{x} + c_2\boldsymbol{x^2} + c_3\boldsymbol{x^3}) + (b_0 + b_1\boldsymbol{x} + b_2\boldsymbol{x^2} + b_3\boldsymbol{x^3}) = (c_0 + b_0) + (c_1 + b_1)\boldsymbol{x} + (c_2 + b_2)\boldsymbol{x^2} + (c_3 + b_3)\,\boldsymbol{x^3}$, where $c_0 + b_0 = 0$.
Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = c_0 + c_1\,\boldsymbol{x} + c_2\,\boldsymbol{x^2} + c_3\,\boldsymbol{x^3}$ (where $c_0 = 0$) is any vector in $\boldsymbol{W}$.

Then $k\,\boldsymbol{u} = k\,(c_0 + c_1\mathbf{x} + c_2\mathbf{x}^2 + c_3\mathbf{x}^3) = (kc_0) + (kc_1)\,\mathbf{x} + (kc_2)\,\mathbf{x}^2 + (kc_3)\,\mathbf{x}^3$ where $k\,c_0 = 0$. Therefore, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{P_3}$.

(ii) Let $\boldsymbol{W}$ is the set of all polynomials $a_0 + a_1\mathbf{x} + a_2\mathbf{x}^2 + a_3\mathbf{x}^3$ for which $a_0+a_1+a_2+a_3 = 0$.
Suppose that $\boldsymbol{u} = c_0 + c_1\,\mathbf{x} + c_2\,\mathbf{x}^2 + c_3\,\mathbf{x}^3$ (where $c_0 + c_1 + c_2 + c_3 = 0$) and $\boldsymbol{v} = b_0 + b_1\,\mathbf{x} + b_2\,\mathbf{x}^2 + b_3\,\mathbf{x}^3$ (where $b_0 + b_1 + b_2 + b_3 = 0$) are in $\boldsymbol{W}$.

Now
$\boldsymbol{u} + \boldsymbol{v} = (c_0 + c_1\,\mathbf{x} + c_2\,\mathbf{x}^2 + c_3\,\mathbf{x}^3) + (b_0 + b_1\,\mathbf{x} + b_2\,\mathbf{x}^2 + b_3\,\mathbf{x}^3)$

$= (c_0 + b_0) + (c_1 + b_1)\,\mathbf{x} + (c_2 + b_2)\,\mathbf{x}^2 + (c_3 + b_3)\,\mathbf{x}^3$

Where $(c_0+b_0) + (c_1+b_1) + (c_2+b_2) + (c_3+b_3) = (c_0+c_1+c_2+c_3) + (b_0+b_1+b_2+b_3) = 0 + 0 = 0$. Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = c_0 + c_1\,\mathbf{x} + c_2\,\mathbf{x}^2 + c_3\,\mathbf{x}^3$ (where $c_0 + c_1 + c_2 + c_3 = 0$) is any vector in $\boldsymbol{W}$. Then $k\,\boldsymbol{u} = k\,(c_0 + c_1\mathbf{x} + c_2\mathbf{x}^2 + c_3\mathbf{x}^3) = (kc_0) + (kc_1)\,\mathbf{x} + (kc_2)\,\mathbf{x}^2 + (kc_3)\,\mathbf{x}^3$
Where $(k\,c_0) + (k\,c_1) + (k\,c_2) + (k\,c_3) = k(c_0 + c_1 + c_2 + c_3) = k.0 = 0$

Therefore, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{P_3}$.

(iii) Let $\boldsymbol{W}$ is the set of all polynomials $a_0 + a_1\,\mathbf{x} + a_2\,\mathbf{x}^2 + a_3\,\mathbf{x}^3$ for which $a_0, a_1, a_2,$ and a_3 are integers.
Suppose that the vectors $\boldsymbol{u} = c_0 + c_1\,\mathbf{x} + c_2\,\mathbf{x}^2 + c_3\,\mathbf{x}^3$ (where $c_0, c_1, c_2,$ and c_3 are integers) and $\boldsymbol{v} = b_0 + b_1\,\mathbf{x} + b_2\,\mathbf{x}^2 + b_3\,\mathbf{x}^3$ (where $b_0, b_1, b_2,$ and b_3 are integers) are in $\boldsymbol{W}$.
Now
$\boldsymbol{u} + \boldsymbol{v} = (c_0 + c_1\,\mathbf{x} + c_2\,\mathbf{x}^2 + c_3\,\mathbf{x}^3) + (b_0 + b_1\,\mathbf{x} + b_2\,\mathbf{x}^2 + b_3\,\mathbf{x}^3)$
$= (c_0 + b_0) + (c_1 + b_1)\,\mathbf{x} + (c_2 + b_2)\,\mathbf{x}^2 + (c_3 + b_3)\,\mathbf{x}^3$, where $(c_0 + b_0), (c_1 + b_1), (c_2 + b_2),$ and $(c_3 + b_3)$ are integers (integers are closed under addition). Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = c_0 + c_1\,\mathbf{x} + c_2\,\mathbf{x}^2 + c_3\,\mathbf{x}^3$ (where $c_0, c_1, c_2,$ and c_3 are integers) is any vector in $\boldsymbol{W}$. Then $k\,\boldsymbol{u} = k\,(c_0 + c_1\mathbf{x} + c_2\mathbf{x}^2 + c_3\mathbf{x}^3 = (kc_0) + (kc_1)\,\mathbf{x} + (kc_2)\,\mathbf{x}^2 + (kc_3)\,\mathbf{x}^3$, where $(k\,c_0), (k\,c_1), (k\,c_2),$ *and* $(k\,c_3)$ are not integers (product of real number and integer). Therefore, $k\,\boldsymbol{u} \notin \boldsymbol{W}$. Hence, $\boldsymbol{W}$ is not the subspace of $\boldsymbol{P_3}$.

(iv) Let $\boldsymbol{W}$ is the set of all polynomials of the form $a_0 + a_1\,\mathbf{x}$, where a_0 and a_1 are real numbers. Suppose that $\boldsymbol{u} = c_0 + c_1\,\mathbf{x}$ (where c_0 and c_1 are real numbers) and $\boldsymbol{v} = b_0 + b_1\,\mathbf{x}$ (where b_0 and b_1 are real numbers) are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = (c_0 + c_1\,\mathbf{x}) + (b_0 + b_1\,\mathbf{x}) = (c_0 + b_0) + (c_1 + b_1)\,\mathbf{x}$
Where $(c_0 + b_0)$ and $(c_1 + b_1)$ are real numbers.
Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = c_0 + c_1 \boldsymbol{x}$ (where c_0 and c_1 are real numbers) is any vector in $\boldsymbol{W}$.
Then $k\,\boldsymbol{u} = k\,(c_0 + c_1\,\boldsymbol{x}) = (k\,c_0) + (k\,c_1)\,\boldsymbol{x}$
Where $(k\,c_0)$ and $(k\,c_1)$ are real numbers.
Therefore, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{P_3}$.

Example 13: Determine which of the following are subspaces of $\boldsymbol{M_{22}}$.

(i) All matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where $a + b + c + d = 0$

(ii) All 2 x 2 matrices $\boldsymbol{A}$ such that det $(\boldsymbol{A}) = 0$

(iii) All the matrices of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$

Solution: Let $\boldsymbol{W}$ is the set of all matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where $a + b + c + d = 0$.

(i) Suppose $\boldsymbol{u} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ (where $e + f + g + h = 0$) and $\boldsymbol{v} = \begin{bmatrix} l & m \\ n & p \end{bmatrix}$
(Where $l + m + n + p = 0$) are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} l & m \\ n & p \end{bmatrix} = \begin{bmatrix} e+l & f+m \\ g+n & h+p \end{bmatrix}$ (By definition)

Where $(e + l) + (f + m) + (g + n) + (h + p)$
$= (e + f + g + h) + (l + m + n + p) = 0 + 0 = 0$
Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ (where $e + f + g + h = 0$) is any vector in $\boldsymbol{W}$.

Then $k\,\boldsymbol{u} = k\begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ke & kf \\ kg & kh \end{bmatrix}$ (by definition)

Where $k\,e + k\,f + k\,g + k\,h = k\,(e + f + g + h) = k\,0 = 0$
Hence, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Therefore, $\boldsymbol{W}$ is subspace of $\boldsymbol{M_{22}}$.

(ii) Let $\boldsymbol{W}$ is the set of all 2 x 2 matrices $\boldsymbol{A}$ such that det $(\boldsymbol{A}) = 0$

Suppose $\boldsymbol{u} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ (Where det $(\boldsymbol{u}) = e\,h - f\,g = 0$) and $\boldsymbol{v} = \begin{bmatrix} l & m \\ n & p \end{bmatrix}$
(Where det $(\boldsymbol{v}) = l\,p - m\,n = 0$) are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} l & m \\ n & p \end{bmatrix} = \begin{bmatrix} e+l & f+m \\ g+n & h+p \end{bmatrix}$ (By definition)

Where det $(\boldsymbol{u} + \boldsymbol{v}) = (e + l)\,(h + p) - (f + m)\,(g + n)$
$= e\,h + e\,p + l\,h + l\,p - f\,g - f\,n - m\,g - m\,n$
$= (e\,h - f\,g) + (l\,p - m\,n) + e\,p + l\,h - f\,n - m\,g = e\,p + l\,h - f\,n - m\,g \neq 0$
Therefore, $\boldsymbol{u} + \boldsymbol{v} \notin \boldsymbol{W}$. Therefore, $\boldsymbol{W}$ is not subspace of $\boldsymbol{M_{22}}$.

(iii) Let ***W*** is the set of all matrices of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$

Suppose $\boldsymbol{u} = \begin{bmatrix} e & f \\ 0 & g \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} l & m \\ 0 & n \end{bmatrix}$ are in ***W***.

Then $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} e & f \\ 0 & g \end{bmatrix} + \begin{bmatrix} l & m \\ 0 & n \end{bmatrix} = \begin{bmatrix} e+l & f+m \\ 0 & g+n \end{bmatrix}$ (By definition)

Which is of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$. Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If *k* is any scalar and $\boldsymbol{u} = \begin{bmatrix} e & f \\ 0 & g \end{bmatrix}$ is any vector in ***W***.

Then $k\,\boldsymbol{u} = k\begin{bmatrix} e & f \\ 0 & g \end{bmatrix} = \begin{bmatrix} ke & kf \\ 0 & kg \end{bmatrix}$ (By definition)

Which is of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$. Hence, $k\,\boldsymbol{u} \in \boldsymbol{W}$

Therefore, ***W*** is subspace of $\boldsymbol{M_{22}}$.

Example 14: Determine which of the following are subspaces of the space $\boldsymbol{F}(-\infty,\infty)$.

(i) All ***f*** such that $\boldsymbol{f}(x) \le \boldsymbol{0}$ for all $\boldsymbol{x}$
(ii) all ***f*** such that $\boldsymbol{f}(0) = \boldsymbol{0}$
(iii) All ***f*** such that $\boldsymbol{f}(0) = \boldsymbol{2}$
(iv) all constant functions
(v) All ***f*** of the form $k_1 + k_2 \sin x$, where k_1 and k_2 are real numbers
(vi) All everywhere differentiable functions that satisfy $\boldsymbol{f}' + 2\boldsymbol{f} = 0$.

Solution: **(i)** Let ***W*** is the set of all ***f*** such that $\boldsymbol{f}(x) \le \boldsymbol{0}$ for all $\boldsymbol{x}$.
Suppose ***g*** and ***h*** are the vectors in ***W***. Then $\boldsymbol{g}(x) \le \boldsymbol{0}$ for all $\boldsymbol{x}$ and $\boldsymbol{h}(x) \le \boldsymbol{0}$ for all $\boldsymbol{x}$.
Now $(\boldsymbol{g} + \boldsymbol{h})(x) = \boldsymbol{g}(x) + \boldsymbol{h}(x) \le \boldsymbol{0}$. Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$
If *k* is any scalar and ***g*** is any vector in ***W***. Then $\boldsymbol{g}(x) \le \boldsymbol{0}$ for all $\boldsymbol{x}$
Now $(k\,\boldsymbol{g})(x) = k\,\boldsymbol{g}(x)$, which is greater than $\boldsymbol{0}$ for negative real values of *k*.
$\therefore k\boldsymbol{g} \notin \boldsymbol{W} \quad \forall\ k < 0$.
Hence ***W*** is not the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(ii) Let ***W*** is the set of all ***f*** such that $\boldsymbol{f}(0) = \boldsymbol{0}$.
Suppose ***g*** and ***h*** are the vectors in ***W***. Then $\boldsymbol{g}(0) = \boldsymbol{0}$ and $\boldsymbol{h}(0) = \boldsymbol{0}$
Now $(\boldsymbol{g} + \boldsymbol{h})(0) = \boldsymbol{g}(0) + \boldsymbol{h}(0) = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$. Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$
If *k* is any scalar and ***g*** is any vector in ***W***. Then $\boldsymbol{g}(0) = \boldsymbol{0}$
Now $(k\,\boldsymbol{g})(0) = k\,\boldsymbol{g}(0) = k\,\boldsymbol{0} = \boldsymbol{0}$. $\therefore k\boldsymbol{g} \in \boldsymbol{W}$. Hence ***W*** is the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(iii) Let ***W*** is the set of all ***f*** such that $\boldsymbol{f}(0) = \boldsymbol{2}$
Suppose ***g*** and ***h*** are the vectors in ***W***. Then $\boldsymbol{g}(0) = \boldsymbol{2}$ and $\boldsymbol{h}(0) = \boldsymbol{2}$

Now $(\boldsymbol{g} + \boldsymbol{h})(0) = \boldsymbol{g}(0) + \boldsymbol{h}(0) = 2 + 2 \neq 2 \therefore k\boldsymbol{g} \notin \boldsymbol{W}$. Hence $\boldsymbol{W}$ is not the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(iv) Let W is the set of all constant functions. Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$.
Then $\boldsymbol{g}(x) = \boldsymbol{a}$ and $\boldsymbol{h}(x) = \boldsymbol{b}$, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are constants.
Now $(\boldsymbol{g} + \boldsymbol{h})(x) = \boldsymbol{g}(x) + \boldsymbol{h}(x) = \boldsymbol{a} + \boldsymbol{b}$, which is constant. Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$
If k is any scalar and $\boldsymbol{g}$ is any vector in $\boldsymbol{W}$. Then $\boldsymbol{g}(x) = \boldsymbol{a}$, where $\boldsymbol{a}$ is any constant.
Now $(k\,\boldsymbol{g})(x) = k\,\boldsymbol{g}(x) = k\,\boldsymbol{a}$, which is a constant. $\therefore k\boldsymbol{g} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(v) Let W is the set of all $\boldsymbol{f}$ of the form $k_1 + k_2 \sin x$, where k_1 and k_2 are real numbers
Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$. Then $\boldsymbol{g}(x) = m_1+m_2 \sin x$ and $\boldsymbol{h}(x) = n_1+n_2 \sin x$, where m_1, m_2, n_1 and n_2 are real numbers.

Now $(\boldsymbol{g} + \boldsymbol{h})(x) = \boldsymbol{g}(x) + \boldsymbol{h}(x) = [m_1+m_2 \sin x]+[n_1+n_2 \sin x] = (m_1+n_1)+(m_2+n_2) \sin x$
Which is of the form $k_1 + k_2 \sin x$. Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{g}$ is any vector in $\boldsymbol{W}$. Then $\boldsymbol{g}(x) = m_1 + m_2 \sin x$, where m_1 and m_2 are any real numbers.

Now $(k\,\boldsymbol{g})(x) = k\,\boldsymbol{g}(x) = k\,[m_1 + m_2 \sin x] = (k\,m_1) + (k\,m_2) \sin x$
Which is of the form $k_1 + k_2 \sin x$. $\therefore k\boldsymbol{g} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(vi) Let W is the set of all everywhere differentiable functions that satisfy $\boldsymbol{f}' + 2\boldsymbol{f} = \boldsymbol{0}$.
Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$. Then $\boldsymbol{g}' + 2\boldsymbol{g} = \boldsymbol{0}$ and $\boldsymbol{h}' + 2\boldsymbol{h} = \boldsymbol{0}$
Now $(\boldsymbol{g} + \boldsymbol{h})' + 2(\boldsymbol{g} + \boldsymbol{h}) = \boldsymbol{g}' + \boldsymbol{h}' + 2(\boldsymbol{g} + \boldsymbol{h}) = (\boldsymbol{g}' + 2\boldsymbol{g}) + (\boldsymbol{h}' + 2\boldsymbol{h}) = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$
Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$
If k is any scalar and $\boldsymbol{g}$ is any vector in $\boldsymbol{W}$. Then $\boldsymbol{g}' + 2\boldsymbol{g} = \boldsymbol{0}$
Now $(k\boldsymbol{g})' + 2(k\boldsymbol{g}) = k\boldsymbol{g}' + 2k\boldsymbol{g} = k(\boldsymbol{g}' + 2\boldsymbol{g}) = k.\boldsymbol{0} = \boldsymbol{0}$
$\therefore k\boldsymbol{g} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{F}(-\infty,\infty)$.

Remark: Let n be a nonnegative integer, and let $\boldsymbol{P}_n$ be the set of real valued function of the form $\boldsymbol{p}(\boldsymbol{x}) = a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + \ldots + a_n\boldsymbol{x}^n$ where $a_0, a_1, a_2, \ldots, a_n$ *are* real numbers, then $\boldsymbol{P}_n$ is a subspace $\boldsymbol{F}(-\infty,\infty)$.

Example 15: Show that the invertible n x n matrices do not form a subspace of $\boldsymbol{M}_{n\times n}$.

Solution: Let $\boldsymbol{W}$ is the set of invertible matrices in $\boldsymbol{M}_{n\times n}$. This set fails to be a subspace on both counts- it is closed under neither scalar multiplication nor addition.

For example consider invertible matrices $\boldsymbol{W} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \boldsymbol{V} = \begin{bmatrix} -1 & 2 \\ -2 & 5 \end{bmatrix}$ in $\boldsymbol{M}_{n\times n}$.

The matrix $\boldsymbol{0.U}$ is a 2x2 zero matrix, hence is not invertible; and the matrix $\boldsymbol{U} + \boldsymbol{V}$ has a column of zeros, hence is not invertible.

Theorem: If $\boldsymbol{Ax = 0}$ is a homogeneous linear system of m equations in n unknowns, then the set of solution vectors is a subspace of $\boldsymbol{R^n}$.

Example 16: Consider the linear systems

$$(a)\quad \begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (b)\begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ -2 & 4 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$(c)\quad \begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ 4 & 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (d)\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Each of the systems has three unknowns, so the solutions form subspaces of $\boldsymbol{R^3}$. Geometrically, this means that each solution space must be a line through origin, a plane through origin, the origin only, or all of $\boldsymbol{R^3}$.

Solution :*(a)* The solutions are $x = 2s - 3t, y = s, z = t$. From which it follows that $x = 2y - 3z$ *or* $x - 2y + 3z = 0$.
This is the equation of the plane through the origin with $\boldsymbol{n} = (1, -2, 3)$ as a normal vector.

(b) The solutions are $x = -5t, y = -t, z = t$, which are parametric equations for the line through the origin parallel to the vector $\boldsymbol{v} = (-5, -1, 1)$.

(c) The solution is $x = 0, y = 0, z = 0$ so the solution space is the origin only, that is $\{\boldsymbol{0}\}$.

(d) The solutions are $x = r, y = s, z = t$. where r, s and t have arbitrary values, so the solution space is all $\boldsymbol{R^3}$.

A Subspace Spanned by a Set: The next example illustrates one of the most common ways of describing a subspace. We know that the term linear combination refers to any sum of scalar multiples of vectors, and Span $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ denotes the set of all vectors that can be written as linear combinations of $\boldsymbol{v_1}, \dots, \boldsymbol{v_p}$.

Example 17: Given $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ in a vector space $\boldsymbol{V}$, let $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. Show that $\boldsymbol{H}$ is a subspace of $\boldsymbol{V}$.

Solution: The zero vector is in $\boldsymbol{H}$, since $\boldsymbol{0} = 0\boldsymbol{v_1} + 0\boldsymbol{v_2}$. To show that $\boldsymbol{H}$ is closed under vector addition, take two arbitrary vectors in $\boldsymbol{H}$, say,

$$\boldsymbol{u} = s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2} \quad \text{and} \quad \boldsymbol{w} = t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2}$$

By Axioms 2, 3 and 8 for the vector space $\boldsymbol{V}$.

$$\boldsymbol{u} + \boldsymbol{w} = (s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2}) + (t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2})$$

$= (s_1 + t_1)\,\boldsymbol{v_1} + (s_2 + t_2)\,\boldsymbol{v_2}$

So ***u + w*** is in ***H***. Furthermore, if *c* is any scalar, then by Axioms 7 and 9,

$$c\boldsymbol{u} = c\,(s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2}) = (cs_1)\,\boldsymbol{v_1} + (cs_2)\boldsymbol{v_2}$$

Which shows that $c\boldsymbol{u}$ is in ***H*** and ***H*** is closed under scalar multiplication.

Thus ***H*** is a subspace of ***V***. □

Later on we will prove that every nonzero subspace of $\boldsymbol{R}^3$, other than $\boldsymbol{R}^3$ itself, is either Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ for some linearly independent $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ or Span $\{\boldsymbol{v}\}$ for $\boldsymbol{v} \neq \boldsymbol{0}$. In the first case the subspace is a plane through the origin and in the second case a line through the origin. (See Figure below) It is helpful to keep these geometric pictures in mind, even for an abstract vector space.

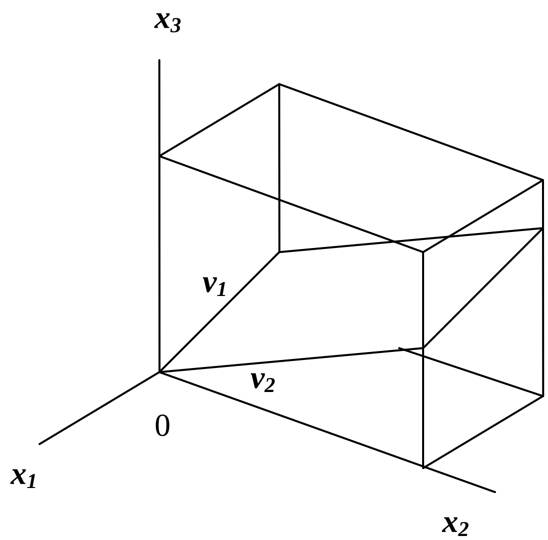


Figure 9 – An example of a subspace

The argument in Example 17 can easily be generalized to prove the following theorem.

<u>Theorem 3:</u> If $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ are in a vector space ***V***, then Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ is a subspace of ***V***.

We call Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ the subspace spanned (or generated) by $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$. Given any subspace ***H*** of ***V***, a spanning (or generating) set for ***H*** is a set $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ in ***H*** such that $\boldsymbol{H} = \text{Span}\,\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$.

<u>Proof:</u>

The zero vector is in ***H***, since $\boldsymbol{0} = 0\boldsymbol{v_1} + 0\boldsymbol{v_2} + \ldots + 0\boldsymbol{v_n} = \sum_{j=0}^{j=n} 0V_j = 0\left(\sum_{j=0}^{j=n} V_j\right) = \boldsymbol{0}$

To show that ***H*** is closed under vector addition, take two arbitrary vectors in ***H*****,** say,

$\boldsymbol{u} = s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2} + \ldots + t_n\boldsymbol{v_n} = \sum_{i=0}^{i=n} s_i V_i$

and

$\boldsymbol{w} = t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2} \ldots + t_n\boldsymbol{v_n} = \sum_{k=0}^{k=n} t_k v_k$

By Axioms 2, 3 and 8 for the vector space ***V***.

$$\boldsymbol{u} + \boldsymbol{w} = \sum_{i=0}^{i=n} s_i v_i + \sum_{k=0}^{k=n} t_k v_k = (s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2} + \ldots + s_n\boldsymbol{v_n}) + (t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2} \ldots + t_n\boldsymbol{v_n})$$

$$= (s_1 + t_1)\,\boldsymbol{v_1} + (s_2 + t_2)\,\boldsymbol{v_2} + \ldots + (s_n + t_n)\,\boldsymbol{v_n} = \sum_{p=0}^{p=n} (s_p + t_p) v_p$$

So $\boldsymbol{u} + \boldsymbol{w}$ is in ***H***. Furthermore, if *c* is any scalar, then by Axioms 7 and 9,

$$c\boldsymbol{u} = c(s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2} + \ldots + s_n\boldsymbol{v_n}) = (cs_1)\,\boldsymbol{v_1} + (cs_2)\boldsymbol{v_2} + \ldots + (cs_n)\boldsymbol{v_n} = \sum_{r=0}^{r=n} cs_r v_r\,.$$

Which shows that $c\boldsymbol{u}$ is in ***H*** and ***H*** is closed under scalar multiplication.

Thus ***H*** is a subspace of ***V***. □

Example 18: Let ***H*** be the set of all vectors of the form *(a – 3b, b – a, a, b)*, where *a* and *b* are arbitrary scalars. That is, let $\boldsymbol{H} = \{(a - 3b, b - a, a, b): a \text{ and } b \text{ in } \boldsymbol{R}\}$. Show that ***H*** is a subspace of $\boldsymbol{R}^4$.

Solution: Write the vectors in ***H*** as column vectors. Then an arbitrary vector in ***H*** has the form

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\qquad\qquad\uparrow\qquad\quad\uparrow$$

$$\qquad\qquad\boldsymbol{v_1}\qquad\quad\boldsymbol{v_2}$$

This calculation shows that ***H*** = Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$, where $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are the vectors indicated above. Thus ***H*** is a subspace of $\boldsymbol{R}^4$ by Theorem 3. □

Example18 illustrates a useful technique of expressing a subspace ***H*** as the set of linear combinations of some small collection of vectors. If ***H*** = Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$, we can think of the vectors $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ in the spanning set as "handles" that allow us to hold on to the subspace ***H***. *Calculations with the infinitely many vectors in* ***H*** *are often reduced to operations with the finite number of vectors in the spanning set.*

Exercises:

In exercises 1-13 a set of objects is given together with operations of addition and scalar multiplication. Determine which sets are vector spaces under the given operations. For those that are not, list all axioms that fail to hold.

1. The set of all triples of real numbers *(x, y, z)* with the operations
$$(x,y,z)+(x',y',z')=(x+x',y+y',z+z') \text{ and } k(x,y,z)=(kx,y,z)$$

2. The set of all triples of real numbers *(x, y, z)* with the operations
$$(x,y,z)+(x',y',z')=(x+x',y+y',z+z') \text{ and } k(x,y,z)=(0,0,0)$$

3. The set of all pairs of real numbers *(x, y)* with the operations
$$(x,y)+(x',y')=(x+x',y+y') \text{ and } k(x,y)=(2kx,2ky)$$

4. The set of all pairs of real numbers of the form *(x, 0)* with the standard operations on $\boldsymbol{R}^2$.

5. The set of all pairs of real numbers of the form *(x, y)*, where $x \geq 0$, with the standard operations on $\boldsymbol{R}^2$.

6. The set of all n-tuples of real numbers of the form *(x, x, …, x)* with the standard operations on $\boldsymbol{R}^n$.

7. The set of all pairs of real numbers *(x, y)* with the operations.
$$(x,y)+(x',y')=(x+x'+1,y+y'+1) \text{ and } k(x,y)=(kx,ky)$$

8. The set of all 2x2 matrices of the form $\begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}$ with matrix addition and scalar multiplication.

9. The set of all 2x2 matrices of the form $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ with matrix addition and scalar multiplication.

10. The set of all pairs of real numbers of the form *(1, x)* with the operations
$$(1,y)+(1,y')=(1,y+y') \text{ and } k(1,y)=(1,ky)$$

11. The set of polynomials of the form $\boldsymbol{a} + b\boldsymbol{x}$ with the operations
$$(\boldsymbol{a_0}+a_1\boldsymbol{x})+(\boldsymbol{b_0}+b_1\boldsymbol{x})=(\boldsymbol{a_0}+\boldsymbol{b_0})+(a_1+b_1)\boldsymbol{x} \text{ and } k(\boldsymbol{a_0}+a_1\boldsymbol{x})=(k\boldsymbol{a_0})+(ka_1)\boldsymbol{x}$$

12. The set of all positive real numbers with operations $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{xy}$ and $k\boldsymbol{x}=\boldsymbol{x}^k$

13. The set of all real numbers *(x, y)* with operations

$$(x, y) + (x', y') = (xx', yy') \text{ and } k(x, y) = (kx, ky)$$

14. Determine which of the following are subspaces of $\boldsymbol{M_{nn}}$.
(a) all n x n matrices $\boldsymbol{A}$ such that tr $(\boldsymbol{A}) = 0$
(b) all n x n matrices $\boldsymbol{A}$ such that $\boldsymbol{A}^T = -\boldsymbol{A}$
(c) all n x n matrices $\boldsymbol{A}$ such that the linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution
(d) all n x n matrices $\boldsymbol{A}$ such that $\boldsymbol{AB} = \boldsymbol{BA}$ for a fixed n x n matrix $\boldsymbol{B}$

15. Determine whether the solution space of the system $\boldsymbol{Ax} = \boldsymbol{0}$ is a line through the origin, a plane through the origin, or the origin only. If it is a plane, find an equation for it; if it is a line, find parametric equations for it.

(a) $\boldsymbol{A} = \begin{bmatrix} -1 & 1 & 1 \\ 3 & -1 & 0 \\ 2 & -4 & -5 \end{bmatrix}$ (b) $\boldsymbol{A} = \begin{bmatrix} 1 & -2 & 3 \\ -3 & 6 & 9 \\ -2 & 4 & -6 \end{bmatrix}$

(c) $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$ (d) $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -6 \\ 1 & 4 & 4 \\ 3 & 10 & 6 \end{bmatrix}$

16. Determine if the set "all polynomial in $\boldsymbol{p_n}$ such that $\boldsymbol{p(0) = 0}$" is a subspace of $\boldsymbol{P_n}$ for an appropriate value of n. Justify your answer.

17. Let $\boldsymbol{H}$ be the set of all vectors of the form $\begin{bmatrix} s \\ 3s \\ 2s \end{bmatrix}$. Find a vector $\boldsymbol{v}$ in $\boldsymbol{R^3}$ such that $\boldsymbol{H}$ = Span $\{\boldsymbol{v}\}$. Why does this show that $\boldsymbol{H}$ is a subspace of $\boldsymbol{R^3}$?

18. Let $\boldsymbol{W}$ be the set of all vectors of the form $\begin{bmatrix} 5b + 2c \\ b \\ c \end{bmatrix}$, where b and c are arbitrary. Find vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ such that $\boldsymbol{W}$ = Span $\{\boldsymbol{u}, \boldsymbol{v}\}$. Why does this show that $\boldsymbol{W}$ is a subspace of $\boldsymbol{R^3}$?

19. Let $\boldsymbol{W}$ be the set of all vectors of the form $\begin{bmatrix} s + 3t \\ s - t \\ 2s - t \\ 4t \end{bmatrix}$. Show that $\boldsymbol{W}$ is a subspace of $\boldsymbol{R^4}$.

20. Let $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \boldsymbol{v_2} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \boldsymbol{v_3} = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}, \text{and } \boldsymbol{w} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

(a) Is $\boldsymbol{w}$ in $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$? How many vectors are in $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$?
(b) How many vectors are in Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$?
(c) Is $\boldsymbol{w}$ in the subspace spanned by $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$? Why?

In exercises 21 and 22, let $\boldsymbol{W}$ be the set of all vectors of the form shown, where a, b and c represent arbitrary real numbers. In each case, either find a set $\boldsymbol{S}$ of vectors that spans $\boldsymbol{W}$ or give an example to show that $\boldsymbol{W}$ is not a vector space.

21. $\begin{bmatrix} 3a+b \\ 4 \\ a-5b \end{bmatrix}$

22. $\begin{bmatrix} a-b \\ b-c \\ c-a \\ b \end{bmatrix}$

23. Show that $\boldsymbol{w}$ is in the subspace of $\boldsymbol{R^4}$ spanned by $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$, where

$$\boldsymbol{w} = \begin{bmatrix} -9 \\ 7 \\ 4 \\ 8 \end{bmatrix}, \boldsymbol{v_1} = \begin{bmatrix} 7 \\ -4 \\ -2 \\ 9 \end{bmatrix}, \boldsymbol{v_2} = \begin{bmatrix} -4 \\ 5 \\ -1 \\ -7 \end{bmatrix}, \boldsymbol{v_3} = \begin{bmatrix} -9 \\ 4 \\ 4 \\ -7 \end{bmatrix}$$

24. Determine if $\boldsymbol{y}$ is in the subspace of $\boldsymbol{R^4}$ spanned by the columns of $\boldsymbol{A}$, where

$$\boldsymbol{y} = \begin{bmatrix} 6 \\ 7 \\ 1 \\ -4 \end{bmatrix}, \boldsymbol{A} = \begin{bmatrix} 5 & -5 & -9 \\ 8 & 8 & -6 \\ -5 & -9 & 3 \\ 3 & -2 & -7 \end{bmatrix}$$

Lecture 21

Null Spaces, Column Spaces, and Linear Transformations

Subspaces arise in as set of all solutions to a system of homogenous linear equations as the set of all linear combinations of certain specified vectors. In this lecture, we compare and contrast these two descriptions of subspaces, allowing us to practice using the concept of a subspace. In applications of linear algebra, subspaces of $\boldsymbol{R}^n$ usually arise in one of two ways:

- as the set of all solutions to a system of homogeneous linear equations or
- as the set of all linear combinations of certain specified vectors.

Our work here will provide us with a deeper understanding of the relationships between the solutions of a linear system of equations and properties of its coefficient matrix.

Null Space of a Matrix:

Consider the following system of homogeneous equations:

$$\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -5x_1 + 9x_2 + x_3 &= 0 \end{aligned} \qquad (1)$$

In matrix form, this system is written as $\boldsymbol{Ax = 0}$, where

$$\boldsymbol{A} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \qquad (2)$$

Recall that the set of all $\boldsymbol{x}$ that satisfy (1) is called the solution set of the system (1). Often it is convenient to relate this set directly to the matrix $\boldsymbol{A}$ and the equation $\boldsymbol{Ax = 0}$. We call the set of $\boldsymbol{x}$ that satisfy $\boldsymbol{Ax = 0}$ the **null space** of the matrix $\boldsymbol{A}$. The reason for this name is that if matrix A is viewed as a linear operator that maps points of some vector space V into itself, it can be viewed as mapping all the elements of this solution space of AX = 0 into the null element "0". Thus the null space N of A is that subspace of all vectors in V which are imaged into the null element "0" by the matrix A.

NULL SPACE

Definition The **null space** of an $m \times n$ matrix $\boldsymbol{A}$, written as Nul $\boldsymbol{A}$, is the set of all solutions to the homogeneous equation $\boldsymbol{Ax = 0}$. In set notation,

Nul $\boldsymbol{A} = \{\boldsymbol{x}: \boldsymbol{x}$ is in $\boldsymbol{R}^n$ and $\boldsymbol{Ax = 0}\}$

OR

$$Nul(A) = \{x / \forall x \in \mathbb{R}, Ax = 0\}$$

A more dynamic description of Nul $\boldsymbol{A}$ is the set of all $\boldsymbol{x}$ in $\boldsymbol{R}^n$ that are mapped into the zero vector of $\boldsymbol{R}^m$ via the linear transformation $\boldsymbol{x} \to \boldsymbol{Ax}$, Where A is a matrix of transformation. See Figure1

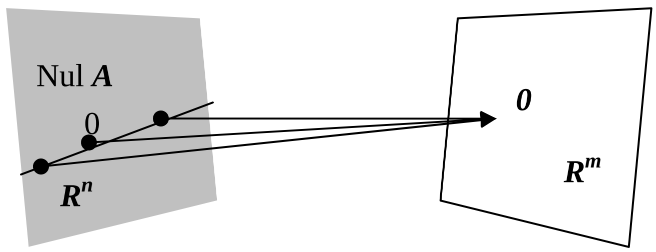


Figure 1

Example 1: Let $A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$ and let $u = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$. Determine if $u \in \text{Nul } A$.

Solution: To test if ***u*** satisfies ***Au = 0***, simply compute

$$Au = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 - 9 + 4 \\ -25 + 27 - 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$ Thus ***u*** is in Nul ***A***.

Example: Determine the null space of the following matrix:

$$A = \begin{pmatrix} 4 & 0 \\ -8 & 20 \end{pmatrix}$$

Solution: To find the null space of *A* we need to solve the following system of equations:

$$\begin{pmatrix} 4 & 0 \\ -8 & 20 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 4x_1 & + \ 0x_2 \\ -8x_1 & +20x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow 4x_1 + 0x_2 = 0 \qquad \Rightarrow x_1 = 0$$

$$and \quad \Rightarrow -8x_1 + 20x_2 = 0 \qquad \Rightarrow x_2 = 0$$

We can find Null space of a matrix with two ways i.e. with matrices or with system of linear equations. We have given this in both matrix form and (here first we convert the matrix into system of equations) equation form. In equation form it is easy to see that by solving these equations together the only solution is $x_1 = x_2 = 0$. In terms of vectors from $\mathbb{R}^2$ the solution consists of the single vector $\{0\}$ and hence the null space of *A* is $\{0\}$.

Activity: Determine the null space of the following matrices:

1. $0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

2. $M = \begin{pmatrix} 1 & -5 \\ -5 & 25 \end{pmatrix}$

In earlier (previous) lectures, we developed the technique of elementary row operations to solve a linear system. We know that performing elementary row operations on an augmented matrix does not change the solution set of the corresponding linear system $\boldsymbol{Ax=0}$. Therefore, we can say that it does not change the null space of $\boldsymbol{A}$. We state this result as a theorem:

Theorem 1: Elementary row operations do not change the null space of a matrix.

Or

Null space N(A) of a matrix A can not be changed (always same) by changing the matrix with elementary row operations.

Example: Determine the null space of the following matrix using the elementary row operations: (Taking the matrix from the above Example)

$$A = \begin{pmatrix} 4 & 0 \\ -8 & 20 \end{pmatrix}$$

Solution: First we transform the matrix to the reduced row echelon form:

$$\begin{pmatrix} 4 & 0 \\ -8 & 20 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ -8 & 20 \end{pmatrix} \quad \frac{1}{4}R_1$$

$$\sim \begin{pmatrix} 1 & 0 \\ 0 & 20 \end{pmatrix} \quad R_2 + 8R_1$$

$$\sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \frac{1}{20}R_2$$

which corresponds to the system

$$\begin{aligned} x_1 \quad &= 0 \\ x_2 &= 0 \end{aligned}$$

Since every column in the coefficient part of the matrix has a leading entry that means our system has the **trivial solution only**:

$x_1 = 0$

$x_2 = 0$

This means **the null space consists only of the zero vector**.

We can observe and compare both the above examples which show the same result.

Theorem 2: The null space of an $m \times n$ matrix A is a subspace of R^n. Equivalently, the set of all solutions to a system $Ax = 0$ of m homogeneous linear equations in n unknowns is a subspace of R^n.

Or simply, the null space is the space of all the vectors of a Matrix A of any order those are mapped (assign) onto zero vector in the space R^n *(i.e. $Ax = 0$).*

Proof: We know that the subspace of A consists of all the solution to the system $\boldsymbol{Ax = 0}$. First, we should point out that the zero vector, **0**, in $\boldsymbol{R^n}$ will be a solution to this system and so we know that the null space is not empty. This is a good thing since a vector space (subspace or not) must contain at least one element.

Now we know that the null space is not empty. Consider $\boldsymbol{u}$, $\boldsymbol{v}$ be two any vectors (elements) (in) from the null space and let c be any scalar. We just need to show that the sum ($\boldsymbol{u+v}$) and scalar multiple ($c.\boldsymbol{u}$) of these are also in the null space.

Certainly Nul $\boldsymbol{A}$ is a subset of $\boldsymbol{R^n}$ because $\boldsymbol{A}$ has n columns. To show that Nul(A) is the subspace, we have to check three conditions whether they are satisfied or not. If Nul(A) satisfies the all three condition, we say Nul(A) is a subspace otherwise not.

First, zero vector "0" must be in the space and subspace. If zero vector does not in the space we can not say that is a vector space (generally, we use space for vector space).

And we know that zero vector maps on zero vector so 0 is in Nul(A). Now choose any vectors u, v from Null space and using definition of Null space (i.e. Ax=0)

$$\boldsymbol{Au = 0} \text{ and } \boldsymbol{Av = 0}$$

Now the other two conditions are vector addition and scalar multiplication. For this we proceed as follow:

Let start with vector addition:

To show that $\boldsymbol{u + v}$ is in Nul $\boldsymbol{A}$, we must show that $\boldsymbol{A (u + v) = 0}$. Using the property of matrix multiplication, we find that

$\boldsymbol{A (u + v) = Au + Av = 0 + 0 = 0}$

Thus $\boldsymbol{u + v}$ is in Nul $\boldsymbol{A}$, and Nul $\boldsymbol{A}$ is closed under vector addition.

For Matrix multiplication, consider any scalar , say c,

$\mathbf{A}\,(c\boldsymbol{u}) = c\,(\boldsymbol{Au}) = c\,(\boldsymbol{0}) = \boldsymbol{0}$

which shows that $c\boldsymbol{u}$ is in Nul $\boldsymbol{A}$. Thus Nul $\boldsymbol{A}$ is a subspace of $\boldsymbol{R^n}$.

Example 2: The set $\boldsymbol{H}$, of all vectors in $\boldsymbol{R^4}$ whose coordinates a, b, c, d satisfy the equations

$$a - 2b + 5c = d$$
$$c - a = b$$

is a subspace of $\boldsymbol{R^4}$.

Solution: Since

$$a - 2b + 5c = d$$
$$c - a = b$$

By rearranging the equations, we get

$$a - 2b + 5c - d = 0$$
$$-a - b + c = 0$$

We see that ***H*** is the set of all solutions of the above system of homogeneous linear equations.
Therefore from the Theorem 2, ***H*** is a subspace of $\boldsymbol{R^4}$.

It is important that the linear equations defining the set ***H*** are homogeneous. Otherwise, the set of solutions will definitely not be a subspace (because the zero-vector (origin) is not a solution of a non- homogeneous system), geometrically means that a line that not passes through origin can not be a subspace, because subspace must hold the zero vector (origin). Also, in some cases, the set of solutions could be empty. In this case, we can not find any solution of a system of linear equations, geometrically says that lines are parallel or not intersecting.
If the null space having more than one vector, geometrically means that the lines intersect more than one point and must passes through origin (zero vector) .

An Explicit Description of Nul *A*:

There is no obvious relation between vectors in Nul ***A*** and the entries in ***A***. We say that Nul ***A*** is defined implicitly, because it is defined by a condition that must be checked. No explicit list or description of the elements in Nul ***A*** is given. However, when we solve the equation ***Ax*** = ***0***, we obtain an explicit description of Nul ***A***.

Example 3: Find a spanning set for the null space of the matrix

$$\boldsymbol{A} = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

Solution: The first step is to find the general solution of ***Ax = 0*** in terms of free variables.
After transforming the augmented matrix [***A*** ***0***] to the reduced row echelon form and we get;

$$\begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which corresponds to the system

$$\begin{aligned} x_1 - 2x_2 \quad - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

The general solution is

$$\begin{aligned} x_1 &= 2x_2 + x_4 - 3x_5 \\ x_2 &= \text{ free variable} \\ x_3 &= -2x_4 + 2x_5 \\ x_4 &= \text{ free variable} \\ x_5 &= \text{ free variable} \end{aligned}$$

Next, decompose the vector giving the general solution into a linear combination of vectors where the weights are the free variables. That is,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\substack{\uparrow \\ \boldsymbol{u}}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\substack{\uparrow \\ \boldsymbol{v}}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\substack{\uparrow \\ \boldsymbol{w}}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}}$$

$$= x_2\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w} \qquad (3)$$

Every linear combination of ***u, v*** and ***w*** is an element of Nul ***A***. Thus *{**u, v, w**}* is a spanning set for Nul ***A***.

Two points should be made about the solution in Example 3 that apply to all problems of this type. We will use these facts later.

1. The spanning set produced by the method in Example 3 is automatically linearly independent because the free variables are the weights on the spanning vectors. For instance, look at the 2nd, 4th and 5th entries in the solution vector in (3) and note that $x_2\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w}$ can be *0* only if the weights x_2, x_4 and x_5 are all zero.
2. When Nul *A* contains nonzero vector, the number of vectors in the spanning set for Nul *A* equals the number of free variables in the equation $Ax = 0$.

Example 4: Find a spanning set for the null space of $\boldsymbol{A} = \begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$.

Solution: The null space of ***A*** is the solution space of the homogeneous system

$$\begin{aligned} x_1 - 3x_2 + 2x_3 + 2x_4 + x_5 &= 0 \\ 0x_1 + 3x_2 + 6x_3 + 0x_4 - 3x_5 &= 0 \\ 2x_1 - 3x_2 - 2x_3 + 4x_4 + 4x_5 &= 0 \\ 3x_1 - 6x_2 + 0x_3 + 6x_4 + 5x_5 &= 0 \\ -2x_1 + 9x_2 + 2x_3 - 4x_4 - 5x_5 &= 0 \end{aligned}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 & 0 \\ 0 & 3 & 6 & 0 & -3 & 0 \\ 2 & -3 & -2 & 4 & 4 & 0 \\ 3 & -6 & 0 & 6 & 5 & 0 \\ -2 & 9 & 2 & -4 & -5 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 & 0 \\ 0 & 3 & 6 & 0 & -3 & 0 \\ 0 & 3 & -6 & 0 & 2 & 0 \\ 0 & 3 & -6 & 0 & 2 & 0 \\ 0 & 3 & 6 & 0 & -3 & 0 \end{bmatrix} \begin{matrix} -2R_1 + R_3 \\ -3R_1 + R_4 \\ 2R_1 + R_5 \end{matrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & -1 & 0 \\ 0 & 3 & -6 & 0 & 2 & 0 \\ 0 & 3 & -6 & 0 & 2 & 0 \\ 0 & 3 & 6 & 0 & -3 & 0 \end{bmatrix} (1/3)R_2$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & -12 & 0 & 5 & 0 \\ 0 & 0 & -12 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} -3R_2 + R_3 \\ -3R_2 + R_4 \\ -3R_2 + R_5 \end{matrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -5/12 & 0 \\ 0 & 0 & -12 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -5/12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} 12R_3 + R_4$$

$$\begin{bmatrix} 1 & -3 & 0 & 2 & 11/6 & 0 \\ 0 & 1 & 0 & 0 & -1/6 & 0 \\ 0 & 0 & 1 & 0 & -5/12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} -2R_3 + R_2 \\ -2R_3 + R_1 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 2 & 4/3 & 0 \\ 0 & 1 & 0 & 0 & -1/6 & 0 \\ 0 & 0 & 1 & 0 & -5/12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} 3R_2 + R_1$$

The reduced row echelon form of the augmented matrix corresponds to the system

$$\begin{aligned} 1\,x_1 + \quad\quad 2\,x_4 + (4/3)\,x_5 &= 0 \\ 1\,x_2 + \quad\quad (-1/6)\,x_5 &= 0 \\ 1\,x_3 + \quad (-5/12)\,x_5 &= 0\;. \\ 0 &= 0 \\ 0 &= 0 \end{aligned}$$

No equation of this system has a form zero = nonzero; Therefore, the system is **consistent**. The system has **infinitely many solutions**:

$$x_1 = -2\,x_4 + (-4/3)\,x_5 \quad x_2 = +(1/6)\,x_5 \quad x_3 = +(5/12)\,x_5$$
$$x_4 = \text{arbitrary} \quad x_5 = \text{arbitrary}$$

The solution can be written in the vector form:

$$\boldsymbol{c_4} = (-2,0,0,1,0) \qquad \boldsymbol{c_5} = (-4/3,1/6,5/12,0,1)$$

Therefore $\{(-2,0,0,1,0),\ (-4/3,1/6,5/12,0,1)\}$ is a spanning set for Null space of ***A***.

Activity: Find an explicit description of Nul A where:

1. $A = \begin{pmatrix} 3 & 5 & 5 & 3 & 9 \\ 5 & 1 & 1 & 0 & 3 \end{pmatrix}$

2. $A = \begin{pmatrix} 4 & 1 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$

<u>The Column Space of a Matrix</u>: Another important subspace associated with a matrix is its column space. Unlike the null space, the column space is defined explicitly via linear combinations.

<u>Definition</u>: (Column Space): The column space of an $m \times n$ matrix ***A***, written as Col ***A***, is the set of all linear combinations of the columns of ***A***. If ***A*** = [$\boldsymbol{a_1} \;\ldots\; \boldsymbol{a_n}$], then

Col $\boldsymbol{A}$ = $Span\ \{\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}\}$

Since Span $\{\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}\}$ is a subspace, by Theorem of lecture 20 i.e. if $v_1,...,v_p$ are in a vector space V, then Span $\{v_1,...,v_p\}$ is a subspace of V .

The column space of a matrix is that subspace spanned by the columns of the matrix (columns viewed as vectors). It is that space defined by all linear combinations of the column of the matrix.

Example, in the given matrix,

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 1 & 4 \\ 3 & 1 & 5 \\ 4 & 1 & 6 \end{pmatrix}$$

The column space Col$\boldsymbol{A}$ is all the linear combination of the first (1, 2, 3, 4), the second (1, 1, 1, 1) and the third column (3, 4, 5, 6). That is, Col$\boldsymbol{A}$ = { a·(1, 2, 3, 4) + b·(1, 1, 1, 1) + c·(3, 4, 5, 6) }. In general, **the column space** Col$\boldsymbol{A}$ **contains all the linear combinations of columns of A**.

The next theorem follows from the definition of Col $\boldsymbol{A}$ and the fact that the columns of $\boldsymbol{A}$ are in $\boldsymbol{R^m}$.

Theorem 3: The column space of an $m \times n$ matrix A is a subspace of R^m.

Note that a typical vector in Col $\boldsymbol{A}$ can be written as $\boldsymbol{Ax}$ for some $\boldsymbol{x}$ because the notation $\boldsymbol{Ax}$ stands for a linear combination of the columns of $\boldsymbol{A}$. That is,

Col $\boldsymbol{A}$ = $\{\boldsymbol{b}: \boldsymbol{b} = \boldsymbol{Ax}$ for some $\boldsymbol{x}$ in $\boldsymbol{R^n}\}$

The notation $\boldsymbol{Ax}$ for vectors in Col $\boldsymbol{A}$ also shows that Col $\boldsymbol{A}$ is the range of the linear transformation $\boldsymbol{x} \rightarrow \boldsymbol{Ax}$.

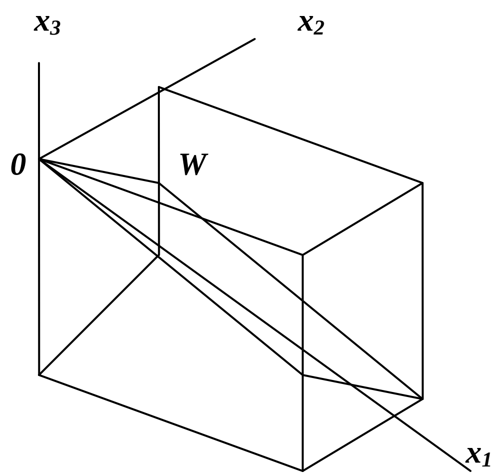
Example 6: Find a matrix $\boldsymbol{A}$ such that $\boldsymbol{W}$ = Col $\boldsymbol{A}$. $\boldsymbol{W} = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a,b \, in \, R \right\}$

Solution: First, write $\boldsymbol{W}$ as a set of linear combinations.

$$\boldsymbol{W} = \left\{ a\begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a,b \, in \, R \right\} = \text{Span}\left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Second, use the vectors in the spanning set as the columns of $\boldsymbol{A}$. Let $\boldsymbol{A} = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ -7 & 0 \end{bmatrix}$.

Then $\boldsymbol{W}$ = Col $\boldsymbol{A}$, as desired.



We know that the columns of ***A*** span $\boldsymbol{R^m}$ if and only if the equation $\boldsymbol{Ax = b}$ has a solution for each ***b***. We can restate this fact as follows:
The column space of an $m \times n$ matrix A is all of R^m if and only if the equation $Ax = b$ has a solution for each b in R^m.

Theorem 4: A system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent if and only if ***b*** in the column space of ***A***.

Example 6: A vector ***b*** in the column space of ***A***. Let $\boldsymbol{Ax} = \boldsymbol{b}$ is the linear system $\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$. Show that ***b*** is in the column space of ***A***, and express ***b*** as a linear combination of the column vectors of ***A***.

Solution: Augmented Matrix is given by

$$\begin{bmatrix} -1 & 3 & 2 & 1 \\ 1 & 2 & -3 & -9 \\ 2 & 1 & -2 & -3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -3 & -2 & -1 \\ 0 & 5 & -1 & -8 \\ 0 & 7 & 2 & -1 \end{bmatrix} \begin{matrix} -1R_1 \\ -1R_1 + R_2 \\ -2R_1 + R_3 \end{matrix}$$

$$\begin{bmatrix} 1 & -3 & -2 & -1 \\ 0 & 1 & -1/5 & -8/5 \\ 0 & 0 & 17/5 & 51/5 \end{bmatrix} \begin{matrix} 1/5R_2 \\ -7R_2 + R_3 \end{matrix}$$

$$\begin{bmatrix} 1 & -3 & 0 & 5 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix} \begin{matrix} (5/17)R_3 \\ (1/5)R_3 + R_2 \\ 2R_3 + R_1 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix} 3R_2 + R_1$$

$\Rightarrow x_1 = 2, x_2 = -1, x_3 = 3$. Since the system is consistent, ***b*** is in the column space of ***A***.

Moreover, $$2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

<u>Example</u>: Determine whether **b** is in the column space of **A** and if so, express **b** as a linear combination of the column vectors of **A**:

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{pmatrix} : \; b = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$$

<u>Solution</u>:

The coefficient matrix $Ax = b$ is:

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$$

The augmented matrix for the linear system that corresponds to the matrix equation $Ax = b$ is:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 2 \end{array}\right)$$

We reduce this matrix to the Reduced Row Echelon Form:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 2 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & -1 & -1 & 1 \\ 2 & 1 & 3 & 2 \end{array}\right) \quad R_2 + (-1)R_1$$

$$\sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & -1 & -1 & 1 \\ 0 & -1 & -1 & 4 \end{array}\right) \quad R_3 + (-2)R_1$$

$$\sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & -1 & -1 & 4 \end{array}\right) \quad (-1)R_2$$

$$\sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 3 \end{array}\right) \quad R_3 + R_2$$

$$\sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{array}\right) \quad \frac{1}{3}R_3$$

$$\sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \quad R_2 + R_3$$

$$\sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \quad R_1 + R_3$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \quad R_1 + (-1)R_2$$

The new system for the equation $Ax = b$ is

$$x_1 \quad + x_3 = 0$$
$$x_2 + x_3 = 0$$
$$0 = 1$$

Equation $0 = 1$ cannot be solved, therefore, the system has **no solution** (i.e. the system is **inconsistent**).
Since the equation $A\mathbf{x} = \mathbf{b}$ has no solution, therefore **b** is not in the column space of ***A***.

Activity: Determine whether **b** is in the column space of A and if so, express **b** as a linear combination of the column vectors of **A**:

1.

$$A=\begin{pmatrix}1 & -1 & 2\\ 9 & 3 & 1\\ 1 & 1 & 1\end{pmatrix};\quad b=\begin{pmatrix}5\\ 1\\ 0\end{pmatrix}$$

2. $$A=\begin{pmatrix}1 & -1 & 1\\ 1 & 1 & -1\\ -1 & -1 & -1\end{pmatrix};\quad b=\begin{pmatrix}1\\ 2\\ 3\end{pmatrix}$$

3. $$A=\begin{pmatrix}1 & 1 & -2 & 1\\ 0 & 2 & 0 & 1\\ 1 & 1 & 1 & -3\\ 0 & 2 & 2 & 1\end{pmatrix};\quad b=\begin{pmatrix}1\\ 2\\ 3\\ 4\end{pmatrix}$$

<u>Theorem 5</u>: If x_0 denotes any single solution of a consistent linear system $Ax=b$ and if $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \ldots, \boldsymbol{v_k}$ form the solution space of the homogeneous system $Ax=0$, then every solution of $Ax=b$ can be expressed in the form $\boldsymbol{x} = \boldsymbol{x_0} + c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_k\boldsymbol{v_k}$ and, conversely, for all choices of scalars $c_1, c_2, c_3, \ldots, c_k$, the vector x is a solution of $Ax=b$.

General and Particular Solutions: The vector x_0 is called a particular solution of $Ax=b$ .The expression $x_0 + c_1 v_1 + c_2 v_2 + \ldots + c_k v_k$ is called the general solution of $Ax=b$, and the expression $c_1 v_1 + c_2 v_2 + \ldots + c_k v_k$ is called the general solution of $Ax=0$.

<u>Example 7</u>: Find the vector form of the general solution of the given linear system $\boldsymbol{Ax} = \boldsymbol{b}$; then use that result to find the vector form of the general solution of $\boldsymbol{Ax}=\boldsymbol{0}$.

$$\begin{aligned}
x_1 + 3x_2 - 2x_3 + \; 2x_5 &= 0\\
2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1\\
5x_3 + 10x_4 + 15x_6 &= 5\\
2x_1 + 6x_2 + 8x_4 + 4x_5 + 18x_6 &= 6
\end{aligned}$$

<u>Solution</u>: We solve the non-homogeneous linear system. The augmented matrix of this system is given by

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 2 & 6 & 0 & 8 & 4 & 18 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix} \begin{matrix} \\ -2R_1 + R_2 \\ -2R_1 + R_4 \\ \\ \end{matrix}$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix} -1R_2$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \end{bmatrix} \begin{matrix} \\ -5R_2 + R_3 \\ -4R_2 + R_4 \\ \\ \end{matrix}$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} R_{34}$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} (1/6)R_3$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} -3R_3 + R_2$$

$$\begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} 2R_2 + R_1$$

The reduced row echelon form of the augmented matrix corresponds to the system

$$\begin{aligned} 1\,x_1 + 3\,x_2 + \quad 4\,x_4 + 2\,x_5 &= 0 \\ 1\,x_3 + 2\,x_4 &= 0 \\ 1\,x_6 &= (1/3) \\ 0 &= 0 \end{aligned}$$

No equation of this system has a form zero = nonzero; Therefore, the system is **consistent**. The system has **infinitely many solutions**:

$x_1 = -3\,x_2 - 4\,x_4 - 2\,x_5 \quad x_2 = r \quad x_3 = -2\,x_4$

$x_4 = s \quad x_5 = t \quad x_6 = 1/3$

$x_1 = -3r - 4s - 2t \quad x_2 = r \quad x_3 = -2s$

$x_4 = s \quad x_5 = t \quad x_6 = \dfrac{1}{3}$

This result can be written in vector form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} -3r - 4s - 2t \\ r \\ -2s \\ s \\ t \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix} + r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \qquad \text{(A)}$$

which is the general solution of the given system. The vector $\boldsymbol{x_0}$ in (A) is a particular solution of the given system; the linear combination $r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ in (A) is the general solution of the homogeneous system.

Activity:

1. Suppose that $x_1 = -1,\ x_2 = 2,\ x_3 = 4,\ x_4 = -3$ is a solution of a non-homogenous linear system $Ax = b$ and that the solution set of the homogenous system $Ax = 0$ is given by this formula:

$x_1 = -3r + 4s$,

$x_2 = r - s$,

$x_3 = r$,

$x_4 = s$

(a) Find the vector form of the general solution of $Ax = 0$.
(b) Find the vector form of the general solution of $Ax = 0$.

Find the vector form of the general solution of the following linear system $\boldsymbol{Ax} = \boldsymbol{b}$; then use that result to find the vector form of the general solution of $\boldsymbol{Ax}=\boldsymbol{0}$:

2.
$$\begin{aligned} x_1 - 2x_2 &= 1 \\ 3x_1 - 9x_2 &= 2 \end{aligned}$$

3.
$$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 3 \\ -3x_1 - x_2 + 3x_3 + x_4 &= -1 \\ -x_1 + 3x_2 - x_3 + 2x_4 &= 2 \\ 4x_1 - 5x_2 \qquad - 3x_4 &= -5 \end{aligned}$$

The Contrast between Nul *A* and Col *A*:

It is natural to wonder how the null space and column space of a matrix are related. In fact, the two spaces are quite dissimilar. Nevertheless, a surprising connection between the null space and column space will emerge later.

Example 8: Let $\boldsymbol{A} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$

(a) If the column space of $\boldsymbol{A}$ is a subspace of $\boldsymbol{R}^k$, what is k?
(b) If the null space of $\boldsymbol{A}$ is a subspace of $\boldsymbol{R}^k$, what is k?

Solution:

(a) The columns of $\boldsymbol{A}$ each have three entries, so Col $\boldsymbol{A}$ is a subspace of $\boldsymbol{R}^k$, where $k = 3$.
(b) A vector $\boldsymbol{x}$ such that $\boldsymbol{Ax}$ is defined must have four entries, so Nul $\boldsymbol{A}$ is a subspace of $\boldsymbol{R}^k$, where $k = 4$.

When a matrix is not square, as in Example 8, the vectors in Nul $\boldsymbol{A}$ and Col $\boldsymbol{A}$ live in entirely different "universes". For example, we have discussed no algebraic operations that connect vectors in $\boldsymbol{R}^3$ with vectors in $\boldsymbol{R}^4$. Thus we are not likely to find any relation between individual vectors in Nul $\boldsymbol{A}$ and Col $\boldsymbol{A}$.

Example 9: If $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$, find a nonzero vector in Col ***A*** and a nonzero vector in Nul ***A***

Solution: It is easy to find a vector in Col ***A***. Any column of ***A*** will do, say, $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$. To find a nonzero vector in Nul ***A***, we have to do some work. We row reduce the augmented matrix $[A \quad 0]$ to obtain $[A \quad 0] \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$. Thus if ***x*** satisfies $Ax = 0$, then $x_1 = -9x_3, x_2 = 5x_3, x_4 = 0$, and x_3 is free. Assigning a nonzero value to x_3 (say), $x_3 = 1$, we obtain a vector in Nul ***A***, namely, $x = (-9, 5, 1, 0)$.

Example 10: With $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$, let $u = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ and $v = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

(a) Determine if ***u*** is in Nul ***A***. Could ***u*** be in Col ***A***?
(b) Determine if ***v*** is in Col ***A***. Could ***v*** be in Nul ***A***?

Solution: (a) An explicit description of Nul ***A*** is not needed here. Simply compute the product

$$Au = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Obviously, ***u*** is not a solution of $Ax = 0$, so ***u*** is not in Nul ***A***.
Also, with four entries, ***u*** could not possibly be in Col ***A***, since Col ***A*** is a subspace of R^3.
(b) Reduce $[A \quad v]$ to an echelon form:

$$[A \quad v] = \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ -2 & -5 & 7 & 3 & -1 \\ 3 & 7 & -8 & 6 & 3 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ 0 & 1 & -5 & -4 & 2 \\ 0 & 0 & 0 & 17 & 1 \end{bmatrix}$$

At this point, it is clear that the equation $Ax = v$ is consistent, so ***v*** is in Col ***A***. With only three entries, ***v*** could not possibly be in Nul ***A***, since Nul ***A*** is a subspace of R^4.

The following table summarizes what we have learned about Nul ***A*** and Col ***A***.

1. Nul *A* is a subspace of R^n. **2. Nul *A* is implicitly defined; i.e. we are given only a condition ($Ax = 0$) that vectors in Nul *A* must satisfy.** **3. It takes time to find vectors in Nul *A*. Row operations on [$A \quad 0$] are required.** **4. There is no obvious relation between Nul *A* and the entries in *A*.** **5. A typical vector v in Nul *A* has the property that $Av = 0$.** **6. Given a specific vector *v*, it is easy to tell if *v* is in Nul *A*. Just compute *Av*.** **7. Nul $A = \{0\}$ if and only if the equation $Ax = 0$ has only the trivial solution.** **8. Nul $A = \{0\}$ if and only if the linear transformation $x \to Ax$ is one-to-one.**	**1. Col *A* is a subspace of R^m.** **2. Col *A* is explicitly defined; that is, we are told how to build vectors in Col *A*.** **3. It is easy to find vectors in Col *A* The columns of *A* are displayed; others are formed from them.** **4. There is an obvious relation between Col *A* and the entries in *A*, since each column of *A* is in Col *A*.** **5. A typical vector *v* in Col *A* has the property that the equation $Ax = v$ is consistent.** **6. Given a specific vector *v*, it may take time to tell if *v* is in Col *A*. Row operations on [$A \quad v$] are required.** **7. Col $A.= R^m$ if and only if the equation $Ax = b$ has a solution for every *b* in R^m.** **8. Col $A = R^m$ if and only if the linear transformation $x \to Ax$ maps R^n onto R^m.**

Kernel and Range of A Linear Transformation:

Subspaces of vector spaces other than R^n are often described in terms of a linear transformation instead of a matrix. To make this precise, we generalize the definition given earlier in Segment I.

Definition: A **linear transformation *T*** from a vector space ***V*** into a vector space ***W*** is a rule that assigns to each vector ***x*** in ***V*** a unique vector ***T*** (***x***) in ***W***, such that

(i) $T(u + v) = T(u) + T(v)$ for all ***u, v*** in ***V***, and

(ii) $T(cu) = c\,T(u)$ for all ***u*** in ***V*** and all scalars *c*.

The **kernel** (or **null space**) of such a ***T*** is the set of all ***u*** in ***V*** such that $T(u) = 0$ (the zero vector in ***W***). The **range** of ***T*** is the set of all vectors in ***W*** of the form ***T*** (***x***) for some ***x*** in ***V***. If ***T*** happens to arise as a matrix transformation, say, $T(x) = Ax$ for some matrix ***A*** – then the kernel and the range of ***T*** are just the null space and the column space of ***A***, as defined earlier. So if $T(x) = Ax$, col ***A*** = range of ***T***.

Definition: If $T : V \to W$ is a linear transformation, then the set of vectors in *V* that ***T*** maps into **0** is called the **kernel** of ***T***; it is denoted by ***ker(T)***. The set of all vectors in ***W***

that are images under ***T*** of at least one vector in ***V*** is called the **range** of ***T***; it is denoted by ***R(T)***.

<u>Example</u>: If $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is multiplication by the $m \times n$ matrix A, then from the above definition; the kernel of T_A is the null space of ***A*** and the range of T_A is the column space of ***A***.

<u>Remarks</u>: The kernel of ***T*** is a subspace of ***V*** and the range of ***T*** is a subspace of ***W***.

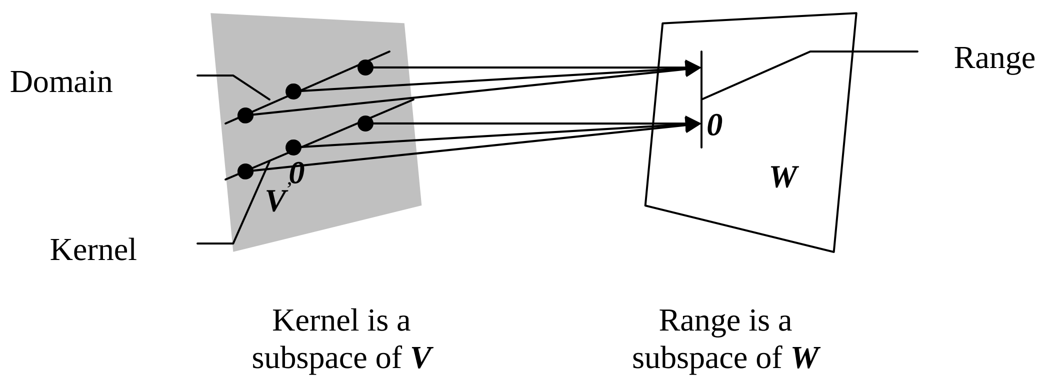


Figure 2 Subspaces associated with a linear transformation.

In applications, a subspace usually arises as either the kernel or the range of an appropriate linear transformation. For instance, the set of all solutions of a homogeneous linear differential equation turns out to be the kernel of a linear transformation. Typically, such a linear transformation is described in terms of one or more derivatives of a function. To explain this in any detail would take us too far a field at this point. So we present only two examples. The first explains why the operation of differentiation is a linear transformation.

<u>Example 11</u>: Let ***V*** be the vector space of all real-valued functions $\boldsymbol{f}$ defined on an interval [*a*, *b*] with the property that they are differentiable and their derivatives are continuous functions on [*a*, *b*]. Let ***W*** be the vector space of all continuous functions on [*a*, *b*] and let $\boldsymbol{D}: \boldsymbol{V} \to \boldsymbol{W}$ be the transformation that changes $\boldsymbol{f}$ in ***V*** into its derivative $\boldsymbol{f}'$. In calculus, two simple differentiation rules are

$$\boldsymbol{D}(\boldsymbol{f}+\boldsymbol{g}) = \boldsymbol{D}(\boldsymbol{f}) + \boldsymbol{D}(\boldsymbol{g}) \text{ and } \boldsymbol{D}(c\boldsymbol{f}) = c\boldsymbol{D}(\boldsymbol{f})$$

That is, ***D*** is a linear transformation. It can be shown that the kernel of ***D*** is the set of constant functions of [*a*, *b*] and the range of ***D*** is the set ***W*** of all continuous functions on [*a*, *b*].

<u>Example 12</u>: The differential equation $y'' + wy = 0$ (4)

where *w* is a constant, is used to describe a variety of physical systems, such as the vibration of a weighted spring, the movement of a pendulum and the voltage in an inductance – capacitance electrical circuit. The set of solutions of (4) is precisely the kernel of the linear transformation that maps a function $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{t})$ into the function $\boldsymbol{f}''(\boldsymbol{t}) + w\boldsymbol{f}(\boldsymbol{t})$. Finding an explicit description of this vector space is a problem in differential equations.

Example 13: Let $\boldsymbol{W} = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a - 3b - c = 0 \right\}$. Show that W is a subspace of R^3 in different ways.

Solution: First method: W is a subspace of R^3 by Theorem 2 because W is the set of all solutions to a system of homogeneous linear equations (where the system has only one equation). Equivalently, W is the null space of the $1x3$ matrix $\boldsymbol{A} = [1 \;\; -3 \;\; -1]$.

Second method: Solve the equation $a - 3b - c = 0$ for the leading variable a in terms of the free variables b and c.

Any solution has the form $\begin{bmatrix} 3b + c \\ b \\ c \end{bmatrix}$, where b and c are arbitrary, and

$$\begin{bmatrix} 3b + c \\ b \\ c \end{bmatrix} = b \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\uparrow \quad \uparrow$$

$$\boldsymbol{v_1} \quad \boldsymbol{v_2}$$

This calculation shows that $W = \text{Span}\{v_1, v_2\}$. Thus W is a subspace of R^3 by Theorem i.e. if $v_1, ..., v_p$ are in a vector space V, then Span $\{v_1, ..., v_p\}$ is a subspace of V. We could also solve the equation $a - 3b - c = 0$ for b or c and get alternative descriptions of W as a set of linear combinations of two vectors.

Example 14: Let $\boldsymbol{A} = \begin{bmatrix} 7 & -3 & 5 \\ -4 & 1 & -5 \\ -5 & 2 & -4 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$, and $\boldsymbol{W} = \begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix}$

Suppose you know that the equations $\boldsymbol{Ax} = \boldsymbol{v}$ and $\boldsymbol{Ax} = \boldsymbol{w}$ are both consistent. What can you say about the equation $\boldsymbol{Ax} = \boldsymbol{v} + \boldsymbol{w}$?

Solution: Both $\boldsymbol{v}$ and $\boldsymbol{w}$ are in Col $\boldsymbol{A}$. Since Col $\boldsymbol{A}$ is a vector space, $\boldsymbol{v} + \boldsymbol{w}$ must be in Col $\boldsymbol{A}$. That is, the equation $\boldsymbol{Ax} = \boldsymbol{v} + \boldsymbol{w}$ is consistent.

Activity:

1. Let ***V*** and ***W*** be any two vector spaces. The mapping $T: V \to W$ such that ***T*** (***v***) **=** ***0*** for every ***v*** in ***V*** is a linear transformation called the **zero transformation**. Find the kernel and range of the zero transformation.

2. Let ***V*** be any vector space. The mapping $I: V \to V$ defined by ***I(v)* = *v*** is called the **identity operator** on ***V***. Find the kernel and range of the identity operator.

Exercises:

1. Determine if $\boldsymbol{w} = \begin{bmatrix} 5 \\ -3 \\ 2 \end{bmatrix}$ is in Nul $\boldsymbol{A}$, where $\boldsymbol{A} = \begin{bmatrix} 5 & 21 & 19 \\ 13 & 23 & 2 \\ 8 & 14 & 1 \end{bmatrix}$.

In exercises 2 and 3, find an explicit description of Nul $\boldsymbol{A}$, by listing vectors that span the null space.

2. $\begin{bmatrix} 1 & 3 & 5 & 0 \\ 0 & 1 & 4 & -2 \end{bmatrix}$

3. $\begin{bmatrix} 1 & -2 & 0 & 4 & 0 \\ 0 & 0 & 1 & -9 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

In exercises 4-7, either use an appropriate theorem to show that the given set, $\boldsymbol{W}$ is a vector space, or find a specific example to the contrary.

4. $\left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a + b + c = 2 \right\}$

5. $\left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} : \begin{matrix} a - 2b = 4c \\ 2a = c + 3d \end{matrix} \right\}$

6. $\begin{bmatrix} b - 2d \\ 5 + d \\ b + 3d \\ d \end{bmatrix} : b, d \text{ real}$

7. $\begin{bmatrix} -a + 2b \\ a - 2b \\ 3a - 6b \end{bmatrix} : a, b \text{ real}$

In exercises 8 and 9, find $\boldsymbol{A}$ such that the given set is Col $\boldsymbol{A}$.

8. $\left\{ \begin{bmatrix} 2s + 3t \\ r + s - 2t \\ 4r + s \\ 3r - s - t \end{bmatrix} : r, s, t \text{ real} \right\}$

9. $\left\{ \begin{bmatrix} b - c \\ 2b + c + d \\ 5c - 4d \\ d \end{bmatrix} : b, c, d \text{ real} \right\}$

For the matrices in exercises 10-13, (a) find k such that Nul $\boldsymbol{A}$ is a subspace of $\boldsymbol{R}^k$, and (b) find k such that Col $\boldsymbol{A}$ is a subspace of $\boldsymbol{R}^k$.

10. $\boldsymbol{A} = \begin{bmatrix} 2 & -6 \\ -1 & 3 \\ -4 & 12 \\ 3 & -9 \end{bmatrix}$

11. $\boldsymbol{A} = \begin{bmatrix} 7 & -2 & 0 \\ -2 & 0 & -5 \\ 0 & -5 & 7 \\ -5 & 7 & -2 \end{bmatrix}$

12. $\boldsymbol{A} = \begin{bmatrix} 4 & 5 & -2 & 6 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$

13. $\boldsymbol{A} = \begin{bmatrix} 1 & -3 & 9 & 0 & -5 \end{bmatrix}$

14. Let $\boldsymbol{A} = \begin{bmatrix} -6 & 12 \\ -3 & 6 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Determine if $\boldsymbol{w}$ is in Col $\boldsymbol{A}$. Is $\boldsymbol{w}$ in Nul $\boldsymbol{A}$?

15. Let $\boldsymbol{A} = \begin{bmatrix} -8 & -2 & -9 \\ 6 & 4 & 8 \\ 4 & 0 & 4 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$. Determine if $\boldsymbol{w}$ is in Col $\boldsymbol{A}$. Is $\boldsymbol{w}$ in Nul $\boldsymbol{A}$?

16. Define $\boldsymbol{T}: \boldsymbol{P_2} \to \boldsymbol{R^2}$ by $\boldsymbol{T}(\boldsymbol{p}) = \begin{bmatrix} \boldsymbol{p}(0) \\ \boldsymbol{p}(1) \end{bmatrix}$. For instance, if $\boldsymbol{p}(t) = \boldsymbol{3} + 5\boldsymbol{t} + 7\boldsymbol{t}^2$, then

$\boldsymbol{T}(\boldsymbol{p}) = \begin{bmatrix} 3 \\ 15 \end{bmatrix}$.

a. Show that $\boldsymbol{T}$ is a linear transformation.
b. Find a polynomial $\boldsymbol{p}$ in $\boldsymbol{P_2}$ that spans the kernel of $\boldsymbol{T}$, and describe the range of $\boldsymbol{T}$.

17. Define a linear transformation $\boldsymbol{T}: \boldsymbol{P_2} \to \boldsymbol{R^2}$ by $\boldsymbol{T}(\boldsymbol{p}) = \begin{bmatrix} \boldsymbol{p}(0) \\ \boldsymbol{p}(1) \end{bmatrix}$. Find polynomials $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ in $\boldsymbol{P_2}$ that span the kernel of $\boldsymbol{T}$, and describe the range of $\boldsymbol{T}$.

18. Let $\boldsymbol{M_{2x2}}$ be the vector space of all 2x2 matrices, and define $\boldsymbol{T}: \boldsymbol{M_{2x2}} \to \boldsymbol{M_{2x2}}$ by $\boldsymbol{T}(\boldsymbol{A}) = \boldsymbol{A} + \boldsymbol{A}^{\boldsymbol{T}}$, where $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

(a) Show that $\boldsymbol{T}$ is a linear transformation.
(b) Let $\boldsymbol{B}$ be any element of $\boldsymbol{M_{2x2}}$ such that $\boldsymbol{B}^{\boldsymbol{T}}=\boldsymbol{B}$. Find an $\boldsymbol{A}$ in $\boldsymbol{M_{2x2}}$ such that $\boldsymbol{T}(\boldsymbol{A}) = \boldsymbol{B}$.
(c) Show that the range of $\boldsymbol{T}$ is the set of $\boldsymbol{B}$ in $\boldsymbol{M_{2x2}}$ with the property that $\boldsymbol{B}^{\boldsymbol{T}}=\boldsymbol{B}$.
(d) Describe the kernel of $\boldsymbol{T}$.

19. Determine whether $\boldsymbol{w}$ is in the column space of $\boldsymbol{A}$, the null space of $\boldsymbol{A}$, or both, where

(a) $\boldsymbol{w}=\begin{bmatrix}1\\1\\-1\\-3\end{bmatrix}, \boldsymbol{A}=\begin{bmatrix}7 & 6 & -4 & 1\\-5 & -1 & 0 & -2\\9 & -11 & 7 & -3\\19 & -9 & 7 & 1\end{bmatrix}$ (b) $\boldsymbol{w}=\begin{bmatrix}1\\2\\1\\0\end{bmatrix}, \boldsymbol{A}=\begin{bmatrix}-8 & 5 & -2 & 0\\-5 & 2 & 1 & -2\\10 & -8 & 6 & -3\\3 & -2 & 1 & 0\end{bmatrix}$

20. Let $\boldsymbol{a_1}, \ldots, \boldsymbol{a_5}$ denote the columns of the matrix $\boldsymbol{A}$, where

$$\boldsymbol{A}=\begin{bmatrix}5 & 1 & 2 & 2 & 0\\3 & 3 & 2 & -1 & -12\\8 & 4 & 4 & -5 & 12\\2 & 1 & 1 & 0 & -2\end{bmatrix}, \ \boldsymbol{B}=\begin{bmatrix}\boldsymbol{a_1} & \boldsymbol{a_2} & \boldsymbol{a_4}\end{bmatrix}$$

(a) Explain why $\boldsymbol{a_3}$ and $\boldsymbol{a_5}$ are in the column space of $\boldsymbol{B}$
(b) Find a set of vectors that spans Nul $\boldsymbol{A}$
(c) Let $\boldsymbol{T}$: $\boldsymbol{R^5} \rightarrow \boldsymbol{R^4}$ be defined by $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{Ax}$. Explain why $\boldsymbol{T}$ is neither one-to-one nor onto.

Lecture 22

Linearly Independent Sets; Bases

First we revise some definitions and theorems from the Vector Space:

Definition: Let V be an arbitrary nonempty set of objects on which two operations are defined, addition and multiplication by scalars.

If the following axioms are satisfied by all objects u, v, w in V and all scalars l and m, then we call V a vector space.

Axioms of Vector Space:

For any set of vectors u, v, w in V *and scalars* l, m, n:

1. $u + v$ is in V
2. $u + v = v + u$
3. $u + (v + w) = (u + v) + w$
4. There exist a zero vector 0 such that
   $0 + u = u + 0 = u$
5. There exist a vector $-u$ in V such that
   $-u + u = 0 = u + (-u)$
6. $(l\,u)$ is in V
7. $l\,(u + v) = l\,u + l\,v$
8. $m\,(n\,u) = (m\,n)\,u = n\,(m\,u)$
9. $(l + m)\,u = I\,u + m\,u$
10. $1u = u$ where 1 is the multiplicative identity

Definition: A subset W of a vector space V is called a subspace of V if W itself is a vector space under the addition and scalar multiplication defined on V.

Theorem: If W is a set of one or more vectors from a vector space V, then W is subspace of V if and only if the following conditions hold:

(a) If u and v are vectors in W, then $u + v$ is in W
(b) If k is any scalar and u is any vector in W, then $k\,u$ is in W.

Definition; The null space of an $m \times n$ matrix A $(Nul\ A)$ is the set of all solutions of the hom equation $Ax = 0$

$$Nul\ A = \{x: x \text{ is in } R^n \text{ and } Ax = 0\}$$

Definition: The column space of an m x n matrix A $(Col\ A)$ is the set of all linear combinations of the columns of A.
If $A = [a_1 \ \dots \ a_n]$,
then
$Col\ A = Span\ \{ a_1, \dots , a_n \}$

Since we know that a set of vectors $S = \{v_1, v_2, v_3, ...v_p\}$ spans a given vector space ***V*** if every vector in ***V*** is expressible as a linear combination of the vectors in ***S***. In general there may be more than one way to express a vector in ***V*** as linear combination of vectors in a spanning set. We shall study conditions under which each vector in ***V*** is expressible as a linear combination of the spanning vectors in exactly one way. Spanning sets with this property play a fundamental role in the study of vector spaces.
In this Lecture, we shall identify and study the subspace ***H*** as "efficiently" as possible. The key idea is that of linear independence, defined as in $\boldsymbol{R^n}$.

<u>Definition</u>: An ***indexed set*** of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in ***V*** is said to be **linearly independent** if the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_p\mathbf{v}_p = 0 \qquad (1)$$

has only the trivial solution, i.e. $c_1 = 0, \dots , c_p = 0$.
The set $\{\mathbf{v}_1,\dots,\mathbf{v}_p\}$ is said to be **linearly dependent** if (1) has a nontrivial solution, that is, if there are some weights, $c_1,\dots,c_p$, not all zero, such that (1) holds. In such a case, (1) is called a **linear dependence relation** among $\mathbf{v}_1, \dots , \mathbf{v}_p$. Alternatively, to say that the ***v***'s are linearly dependent is to say that the zero vector **0** can be expressed as a *nontrivial linear combination* of the ***v***'s.

If the trivial solution is the only solution to this equation then the vectors in the set are called **linearly independent** and the set is called a **linearly independent set**. If there is another solution then the vectors in the set are called **linearly dependent** and the set is called a **linearly dependent set**.

Just as in $\boldsymbol{R^n}$, a set containing a single vector ***v*** is linearly independent if and only if $\mathbf{v} \neq \mathbf{0}$. Also, a set of two vectors is linearly dependent if and only if one of the vectors is a multiple of the other. And any set containing the zero-vector is linearly dependent.

Determining whether a set of vectors $a_1, a_2, a_3, ...a_n$ is linearly independent is easy when one of the vectors is **0**: if, say, $a_1 = 0$, then we have a simple solution to $x_1a_1 + x_2a_2 + x_3\, a_3 + ... + x_na_n = 0$ given by choosing x_1 to be any nonzero value and putting all the other *x*'s equal to 0. Consequently, *if a set of vectors contains the zero vector, it must always be linearly dependent*. Equivalently, *any set of linearly independent vectors cannot contain the zero vector*.

Another situation in which it is easy to determine linear independence is when there are more vectors in the set than entries in the vectors. If $n > m$, then the *n* vectors $a_1, a_2, a_3, ...a_n$ in $\boldsymbol{R^m}$ are columns of an $m \times n$ matrix *A*. The vector equation $x_1a_1 + x_2a_2 + x_3\, a_3 + ... + x_na_n = 0$ is equivalent to the matrix equation $A\mathbf{x} = \mathbf{0}$ whose corresponding linear system has more variables than equations. Thus there must be at least one free variable in the solution, meaning that there are nontrivial solutions

to $x_1a_1 + x_2a_2 + x_3\,a_3 + ... + x_na_n = 0$: If $n > m$, then the set $\{a_1, a_2, a_3, ...a_n\}$ of vectors in $\boldsymbol{R}^m$ *must* be linearly dependent.
When n is small we have a clear geometric picture of the relation amongst linearly independent vectors. For instance, the case $n = 1$ produces the equation $x_1a_1 = 0$, and as long as $a_1 \neq 0$, we only have the trivial solution $x_1 = 0$. *A single nonzero vector always forms a linearly independent set*.
When $n = 2$, the equation takes the form $x_1a_1 + x_2a_2 = 0$. If this were a linear dependence relation, then one of the x's, say x_1, would have to be nonzero. Then we could solve the equation for a_1 and obtain a relation indicating that a_1 is a scalar multiple of a_2. Conversely, if one of the vectors is a scalar multiple of the other, we can express this in the form $x_1a_1 + x_2a_2 = 0$. Thus, *a set of two nonzero vectors is linearly dependent if and only if they are scalar multiples of each other*.

Example: (linearly independent set)
Show that the following vectors are linearly independent:

$$v_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Solution: Let there exist scalars c_1, c_2, c_3 in $\boldsymbol{R}$ such that

$$c_1v_1 + c_3v_2 + c_3v_3 = 0$$

Therefore,

$$\Rightarrow \quad c_1\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} + c_3\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

$$\Rightarrow \quad \begin{bmatrix} -2c_1 \\ c_1 \\ c_1 \end{bmatrix} + \begin{bmatrix} 2c_2 \\ c_2 \\ -2c_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ c_3 \end{bmatrix} = 0$$

$$\Rightarrow \quad \begin{bmatrix} -2c_1 + 2c_2 \\ c_1 + c_2 \\ c_1 - 2c_2 + c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The above can be written as:

$-2c_1 + 2c_2 = 0 \quad ....(1) \Rightarrow -c_1 + c_2 = 0 ....(4)\,(dividing\ by\ 2\ on\ both\ sides\ of\,(1))$

$c_1 + c_2 = 0 \quad ....(2)$

$c_1 - 2c_2 + c_3 = 0 \quad ......(3)$

$Solving\ (2)\ and\ (4)\ implies:$

$c_1 + c_2 = 0$ $-c_1 + c_2 = 0$ $\overline{}$ $0 + 2c_2 = 0$ $\Rightarrow c_2 = 0$	$Solving\ (2)\ implies:$ $c_1 + 0 = 0$ $\Rightarrow c_1 = 0$	$Solving\ (3)\ implies:$ $0 + 0 + c_3 = 0$ $\Rightarrow c_3 = 0$

$\Rightarrow c_1 = c_2 = c_3 = 0\ ; scalars\ c_1, c_2, c_3 \in R\ are\ all\ zero$

$\therefore$ The system has trivial solution.

Hence the given vectors v_1, v_2, v_3 *are linearly independent.*

<u>Example</u>: (linearly dependent set)

If $v_1 = \{2, -1, 0, 3\}, v_2 = \{1, 2, 5, -1\}$ *and* $v_3 = \{7, -1, 5, 8\}$, then the set of vectors $S = \{v_1, v_2, v_3\}$ is linearly dependent, since $3v_1 + v_2 - v_3 = 0$

<u>Example</u>; (linearly dependent set)

The polynomials $p_1 = -x + 1$, $p_2 = -2x^2 + 3x + 5$, *and* $p_3 = -x^2 + 3x + 1$ form a linearly dependent set in p_2 since $3p_1 - p_2 + 2p_3 = 0$.

Note: The linearly independent or linearly dependent sets can also be determined using the Echelon Form or the Reduced Row Echelon Form methods.

<u>Theorem 1</u>: An indexed set { $\boldsymbol{v_1}, \dots, \boldsymbol{v_p}$ } of two or more vectors, with $\boldsymbol{v_1} \neq \boldsymbol{0}$, is linearly dependent if and only if some $\boldsymbol{v_j}$ (with $j > 1$) is a linear combination of the preceding vectors, $\boldsymbol{v_1}, \dots, \boldsymbol{v_{j-1}}$.

The main difference between linear dependence in $\boldsymbol{R^n}$ and in a general vector space is that when the vectors are not n – tuples, the homogeneous equation (1) usually cannot be written as a system of *n* linear equations. That is, the vectors cannot be made into the columns of a matrix ***A*** in order to study the equation $\boldsymbol{Ax} = \boldsymbol{0}$. We must rely instead on the definition of linear dependence and on Theorem 1.

<u>Example 1</u>: Let $\boldsymbol{p_1}(t) = 1$, $\boldsymbol{p_2}(t) = t$ and $\boldsymbol{p_3}(t) = 4 - t$. Then { $\boldsymbol{p_1}, \boldsymbol{p_2}, \boldsymbol{p_3}$ } is linearly dependent in ***P*** because $\boldsymbol{p_3} = 4\boldsymbol{p_1} - \boldsymbol{p_2}$.

<u>Example 2</u>: The set {***Sin*** *t*, ***Cos*** *t*} is linearly independent in ***C*** *[0, 1]* because ***Sin*** *t* and ***Cos*** *t* are not multiples of one another as vectors in ***C*** *[0, 1]*. That is, there is no scalar *c*

such that ***Cos*** *t = c.* ***Sin t*** for all *t* in *[0, 1]*. (Look at the graphs of ***Sin*** *t* and ***Cos*** *t*.) However, *{**Sin** t **Cos** t, **Sin** 2t}* is linearly dependent because of he identity:
Sin *2t = 2* ***Sin*** *t* ***Cos*** *t*, for all *t*.

<u>Useful results</u>:

- A set containing the zero vector is linearly dependent.
- A set of two vectors is linearly dependent if and only if one is a multiple of the other.
- A set containing one nonzeoro vector is linearly independent. i.e. consider the set containing one nonzeoro vector $\{v_1\}$ so $\{v_1\}$ is linearly independent when $v_1 \neq 0$.
- A set of two vectors is linearly independent if and only if neither of the vectors is a multiple of the other.

Activity: Determine whether the following sets of vectors are linearly independent or linearly dependent:

1. $i=(1,0,0,0),\ j=(0,1,0,0), k=(0,0,0,1)$ in $\mathbb{R}^4$.

2. $v_1=(2,0,-1),\ v_2=(-3,-2,-5),\ v_3=(-6,1,-1),\ v_4=(-7,0,2)$ in $\mathbb{R}^3$.

3. $i=(1,0,0,...,0),\ j=(0,1,0,...,0), k=(0,0,0,...,1)$ in $\mathbb{R}^m$.

4. $3x^2+3x+1,\ 4x^2+x,\ 3x^2+6x+5,\ -x^2+2x+7$ in p_2

<u>Definition</u>: Let ***H*** be a subspace of a vector space ***V***. An indexed set of vectors ***B*** = *{**b**$_1$,…, **b**$_p$}* in **V** is a **basis** for ***H*** if

(i) ***B*** is a linearly independent set, and
(ii) the subspace spanned by ***B*** coincides with ***H***; that is,
H = *Span {**b**$_1$,...,**b**$_p$ }*

The definition of a basis applies to the case when ***H*** = ***V***, because any vector space is a subspace of itself. Thus a basis of ***V*** is a linearly independent set that spans ***V***. Observe that when $H \neq V$, condition (ii) includes the requirement that each of the vectors ***b**$_1$,...,**b**$_p$* must belong to ***H***, because *Span { **b**$_1$,...,**b**$_p$ }* contains ***b**$_1$,…,**b**$_p$*, as we saw in lecture 21.

<u>Example 3</u>: Let *A* be an invertible $n \times n$ matrix – say, *A = [**a**$_1$ … **a**$_n$]*. Then the columns of *A* form a basis for ***R^n*** because they are linearly independent and they span ***R^n***, by the Invertible Matrix Theorem.

<u>Example 4</u>: Let ***e**$_1$,…, **e**$_n$* be the columns of the $n \times n$ identity matrix, I_n. That is,

$$e_1 = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\quad e_2 = \begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},\ \dots\ e_n = \begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$$

The set $\{e_1, \ldots, e_n\}$ is called the standard basis for R^n (Fig. 1).

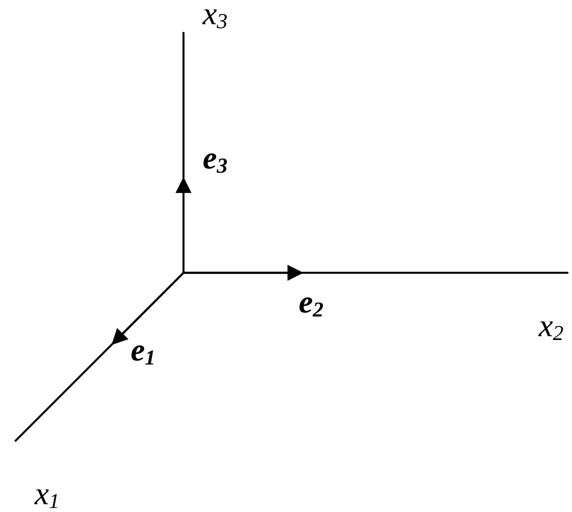


Figure 1 - The standard basis for R^3

Example 5: Let $v_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, v_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}$, and $v_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$. Determine if $\{v_1, v_2, v_3\}$ is a basis for R^3.

Solution: Since there are exactly three vectors here in R^3, we can use one of any methods to determine whether they are basis for $\mathbb{R}^3$ or not. For this, let solve with help of matrices. First form a matrix of vectors i.e. matrix $A = [v_1 \quad v_2 \quad v_3]$. If this matrix is invertible (i.e. $|A| \neq 0$ determinant should be non zero).
For instance, a simple computation shows that $det\ A = 6 \neq 0$. Thus A is invertible. As in example 3, the columns of A form a basis for R^3.

Example 6: Let $S = \{1, t, t^2, \ldots, t^n\}$. Verify that S is a basis for P_n. This basis is called the **standard basis** for P_n.

Solution: Certainly S spans P_n. To show that S is linearly independent, suppose that $c_0, \ldots, c_n$ satisfy

$$c_0.\mathbf{1} + c_1 t + c_2 t^2 + \ldots + c_n t^n = 0(t) \qquad (2)$$

This equality means that the polynomial on the left has the same values as the zero polynomial on the right. A fundamental theorem in algebra says that the only polynomial in P_n with more than n zeros is the zero polynomial. That is, (2) holds for all **t** only if $c_0 = \ldots = c_n = 0$. This proves that S is linearly independent and hence is a basis for P_n. See Figure 2.

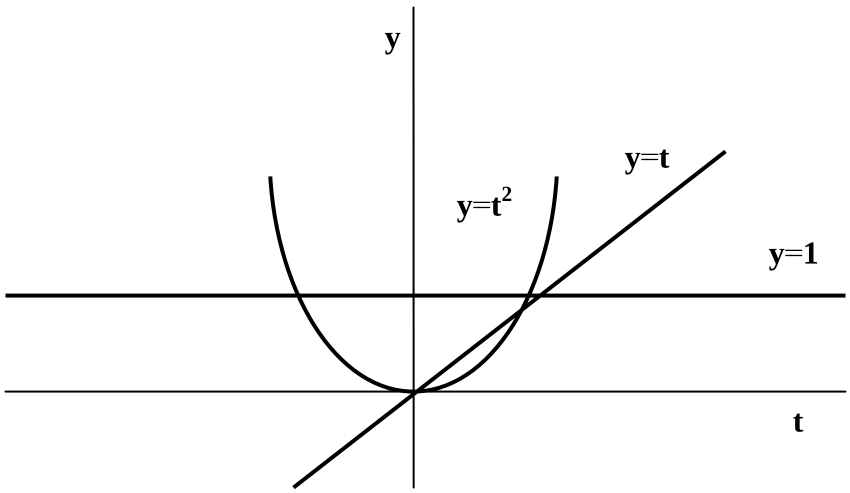


Figure 2 – The standard basis for P_2

Problems involving linear independence and spanning in $\boldsymbol{P_n}$ are handled best by a technique to be discussed later.

Example 7: Check whether the set of vectors $\{(2, -3, 1), (4, 1, 1), (0, -7, 1)\}$ is basis for $\boldsymbol{R^3}$?

Solution: The set $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ of vectors in $\boldsymbol{R^3}$ spans $\boldsymbol{V} = \boldsymbol{R^3}$ if

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + c_3\boldsymbol{v_3} = d_1\boldsymbol{w_1} + d_2\boldsymbol{w_2} + d_3\boldsymbol{w_3} \qquad (*)$$

with $\boldsymbol{w_1} = (1,0,0)$, $\boldsymbol{w_2} = (0,1,0)$, $\boldsymbol{w_3} = (0,0,1)$ has at least one solution for every set of values of the coefficients d_1, d_2, d_3. Otherwise (i.e., if no solution exists for at least some values of d_1, d_2, d_3), $\boldsymbol{S}$ does not span $\boldsymbol{V}$. With our vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$, (*) becomes

$$c_1(2,-3,1) + c_2(4,1,1) + c_3(0,-7,1) = d_1(1,0,0) + d_2(0,1,0) + d_3(0,0,1)$$

Rearranging the left hand side yields

$$\begin{aligned} 2c_1 + 4c_2 + 0c_3 &= 1d_1 + 0d_2 + 0d_3 \\ -3c_1 + 1c_2 - 7c_3 &= 0d_1 + 1d_2 + 0d_3 \\ 1c_1 + 1c_2 + 1c_3 &= 0d_1 + 0d_2 + 1d_3 \end{aligned} \qquad (A)$$

$$\Rightarrow \begin{bmatrix} 2 & 4 & 0 \\ -3 & 1 & -7 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

We now find the determinant of coefficient matrix $\begin{bmatrix} 2 & 4 & 0 \\ -3 & 1 & -7 \\ 1 & 1 & 1 \end{bmatrix}$ to determine whether the system is consistent (so that $\boldsymbol{S}$ spans $\boldsymbol{V}$), or inconsistent ($\boldsymbol{S}$ does not span $\boldsymbol{V}$).

Now $\det\begin{bmatrix} 2 & 4 & 0 \\ -3 & 1 & -7 \\ 1 & 1 & 1 \end{bmatrix} = 2(8) - 4(4) + 0 = 0$

Therefore, the system (A) is inconsistent, and, consequently, the set ***S*** does not span the space ***V***.

<u>Example 8</u>: Check whether the set of vectors
$\{-\mathbf{4} + 1\,\mathbf{t} + 3\,\mathbf{t}^2,\ \mathbf{6} + 5\,\mathbf{t} + 2\,\mathbf{t}^2,\ \mathbf{8} + 4\,\mathbf{t} + 1\,\mathbf{t}^2\}$ is a basis for $\mathbf{P_2}$?

Solution The set $\mathbf{S} = \{\mathbf{p_1(t)}, \mathbf{p_2(t)}, \mathbf{p_3(t)}\}$ of vectors in $\mathbf{P_2}$ spans $\mathbf{V} = \mathbf{P_2}$ if

$$c_1\mathbf{p_1}(\mathbf{t}) + c_2\mathbf{p_2}(\mathbf{t}) + c_3\mathbf{p_3}(\mathbf{t}) = d_1\mathbf{q_1}(\mathbf{t}) + d_2\mathbf{q_2}(\mathbf{t}) + d_3\mathbf{q_3}(\mathbf{t}) \qquad (*)$$

with $\mathbf{q_1(t)} = \mathbf{1} + 0\,\mathbf{t} + 0\,\mathbf{t}^2$, $\mathbf{q_2(t)} = \mathbf{0} + 1\,\mathbf{t} + 0\,\mathbf{t}^2$, $\mathbf{q_3(t)} = \mathbf{0} + 0\,\mathbf{t} + 1\,\mathbf{t}^2$ has at least one solution for every set of values of the coefficients d_1, d_2, d_3. Otherwise (i.e., if no solution exists for at least some values of d_1, d_2, d_3), ***S*** does not span ***V***. With our vectors $\mathbf{p_1(t)}, \mathbf{p_2(t)}, \mathbf{p_3(t)}$, (*) becomes:

$$c_1(-\mathbf{4} + 1\,\mathbf{t} + 3\,\mathbf{t}^2) + c_2(\mathbf{6} + 5\,\mathbf{t} + 2\,\mathbf{t}^2) + c_3(\mathbf{8} + 4\,\mathbf{t} + 1\,\mathbf{t}^2) =$$
$$d_1(\mathbf{1} + 0\,\mathbf{t} + 0\,\mathbf{t}^2) + d_2(\mathbf{0} + 1\,\mathbf{t} + 0\,\mathbf{t}^2) + d_3(\mathbf{0} + 0\,\mathbf{t} + 1\,\mathbf{t}^2)$$

Rearranging the left hand side yields

$$(-4c_1 + 6c_2 + 8c_3)\mathbf{1} + (1c_1 + 5c_2 + 4c_3)\,\mathbf{t} + (3c_1 + 2c_2 + 1c_3)\,\mathbf{t}^2 =$$
$$(1d_1 + 0d_2 + 0d_3)\mathbf{1} + (0d_1 + 1d_2 + 0d_3)\,\mathbf{t} + (0d_1 + 0d_2 + 1d_3)\,\mathbf{t}^2$$

In order for the equality above to hold for all values of t, the coefficients corresponding to the same power of **t** on both sides of the equation must be equal. This yields the following system of equations:

$$\begin{aligned} -4c_1 + 6c_2 + 8c_3 &= 1d_1 + 0d_2 + 0d_3 \\ 1c_1 + 5c_2 + 4c_3 &= 0d_1 + 1d_2 + 0d_3 \\ 3c_1 + 2c_2 + 1c_3 &= 0d_1 + 0d_2 + 1d_3 \end{aligned} \qquad (A)$$

$$\Rightarrow \begin{bmatrix} -4 & 6 & 8 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

We now find the determinant of coefficient matrix $\begin{bmatrix} -4 & 6 & 8 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$ to determine whether the system is consistent (so that ***S*** spans ***V***), or inconsistent (***S*** does not span ***V***).

Now $\det\begin{bmatrix} -4 & 6 & 8 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} = -26 \neq 0$. Therefore, the system (A) is consistent, and, consequently, the set ***S*** spans the space ***V***.

The set $\mathbf{S} = \{\mathbf{p_1(t)}, \mathbf{p_2(t)}, \mathbf{p_3(t)}\}$ of vectors in $\mathbf{P_2}$ is linearly independent if the only solution of

$$c_1\mathbf{p_1}(\mathbf{t}) + c_2\mathbf{p_2}(\mathbf{t}) + c_3\mathbf{p_3}(\mathbf{t}) = \mathbf{0} \qquad (**)$$

is $c_1, c_2, c_3 = 0$. In this case, the set ***S*** forms a basis for span ***S***. Otherwise (i.e., if a solution with at least some nonzero values exists), ***S*** is linearly dependent. With our vectors $\boldsymbol{p_1(t), p_2(t), p_3(t)}$, (2) becomes: $c_1\,(\mathbf{-4} + 1\,\mathbf{t} + 3\,\mathbf{t^2}) + c_2\,(\mathbf{6} + 5\,\mathbf{t} + 2\,\mathbf{t^2}) + c_3\,(\mathbf{8} + 4\,\mathbf{t} + 1\,\mathbf{t^2}) = \mathbf{0}$ Rearranging the left hand side yields

$$(-4\,c_1 + 6\,c_2 + 8\,c_3)\mathbf{1} + (1\,c_1 + 5\,c_2 + 4\,c_3)\,\mathbf{t} + (3\,c_1 + 2\,c_2 + 1\,c_3)\,\mathbf{t^2} = \mathbf{0}$$

This yields the following homogeneous system of equations:

$$\begin{aligned} -4\,c_1 + 6\,c_2 + 8\,c_3 &= 0 \\ 1\,c_1 + 5\,c_2 + 4\,c_3 &= 0 \\ 3\,c_1 + 2\,c_2 + 1\,c_3 &= 0 \end{aligned} \Rightarrow \begin{bmatrix} -4 & 6 & 8 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

As $det \begin{bmatrix} -4 & 6 & 8 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} = -26 \neq 0$. Therefore the set $\boldsymbol{S} = \{\boldsymbol{p_1(t), p_2(t), p_3(t)}\}$ is linearly independent. Consequently, the set ***S*** forms a basis for span ***S***.

Example 9: The set $\boldsymbol{S} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is a basis for the vector space ***V*** of all 2 x 2 matrices.

Solution: To verify that ***S*** is linearly independent, we form a linear combination of the vectors in ***S*** and set it equal to zero:

$$c_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

This gives $\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, which implies that $c_1 = c_2 = c_3 = c_4 = 0$. Hence ***S*** is linearly independent.

To verify that ***S*** spans ***V*** we take any vector $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in ***V*** and we must find scalars c_1, c_2, c_3, and c_4 such that

$$c_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

We find that $c_1 = a$, $c_2 = b$, $c_3 = c$, and $c_4 = d$ so that ***S*** spans ***V***.

The basis ***S*** in this example is called the standard basis for $\boldsymbol{M_{22}}$. More generally, the standard basis for $\boldsymbol{M_{mn}}$ consists of *mn* different matrices with a single *1* and zeros for the remaining entries

Example 10: Show that the set of vectors

$$\left\{ \begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \right\}$$

is a basis for the vector space *V* of all 2 x 2 matrices (i.e. $\boldsymbol{M_{22}}$).

Solution: The set $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}\}$ of vectors in $\boldsymbol{M_{22}}$ spans $\boldsymbol{V} = \boldsymbol{M_{22}}$ if

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + c_3\boldsymbol{v_3} + c_4\boldsymbol{v_4} = d_1\boldsymbol{w_1} + d_2\boldsymbol{w_2} + d_3\boldsymbol{w_3} + d_4\boldsymbol{w_4} \qquad (*)$$

with $\boldsymbol{w_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\boldsymbol{w_2} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\boldsymbol{w_3} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\boldsymbol{w_4} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

has at least one solution for every set of values of the coefficients d_1, d_2, d_3, d_4. Otherwise (i.e., if no solution exists for at least some values of d_1, d_2, d_3, d_4), $\boldsymbol{S}$ does not span $\boldsymbol{V}$.
With our vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}$, (*) becomes:

$$c_1\begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix} + c_2\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix} + c_4\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$$

$$= d_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + d_2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + d_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Rearranging the left hand side yields

$$\begin{bmatrix} 3c_1+0c_2+0c_3+1c_4 & 6c_1-1c_2-8c_3+0c_4 \\ 3c_1-1c_2-12c_3-1c_4 & -6c_1+0c_2-4c_3+2c_4 \end{bmatrix} =$$

$$\begin{bmatrix} 1d_1+0d_2+0d_3+0d_4 & 0d_1+1d_2+0d_3+0d_4 \\ 0d_1+0d_2+1d_3+0d_4 & 0d_1+0d_2+0d_3+1d_4 \end{bmatrix}$$

The matrix equation above is equivalent to the following system of equations

$$\begin{aligned} 3c_1 + 0c_2 + 0c_3 + 1c_4 &= 1d_1 + 0d_2 + 0d_3 + 0d_4 \\ 6c_1 - 1c_2 - 8c_3 + 0c_4 &= 0d_1 + 1d_2 + 0d_3 + 0d_4 \\ 3c_1 - 1c_2 - 12c_3 - 1c_4 &= 0d_1 + 0d_2 + 1d_3 + 0d_4 \\ -6c_1 + 0c_2 - 4c_3 + 2c_4 &= 0d_1 + 0d_2 + 0d_3 + 1d_4 \end{aligned}$$

$$\Rightarrow \begin{bmatrix} 3 & 0 & 0 & 1 \\ 6 & -1 & -8 & 0 \\ 3 & -1 & -12 & -1 \\ -6 & 0 & -4 & 2 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}$$

We now find the determinant of coefficient matrix $A = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 6 & -1 & -8 & 0 \\ 3 & -1 & -12 & -1 \\ -6 & 0 & -4 & 2 \end{bmatrix}$ to determine whether the system is consistent (so that $\boldsymbol{S}$ spans $\boldsymbol{V}$), or inconsistent ($\boldsymbol{S}$ does not span $\boldsymbol{V}$).

Now $det(A) = 48 \neq 0$. Therefore, the system (A) is consistent, and, consequently, the set $\boldsymbol{S}$ spans the space $\boldsymbol{V}$.
Now, the set $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}\}$ of vectors in $\boldsymbol{M_{22}}$ is linearly independent if the only solution of $c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + c_3\boldsymbol{v_3} + c_4\boldsymbol{v_4} = 0$ is $c_1, c_2, c_3, c_4 = 0$. In this case the set $\boldsymbol{S}$ forms a basis for span $\boldsymbol{S}$. Otherwise (i.e., if a solution with at least some nonzero values exists), $\boldsymbol{S}$ is linearly dependent. With our vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}$, we have

$$c_1\begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix} + c_2\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix} + c_4\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Rearranging the left hand side yields

$$\begin{bmatrix} 3c_1 + 0c_2 + 0c_3 + 1c_4 & 6c_1 - 1c_2 - 8c_3 + 0c_4 \\ 3c_1 - 1c_2 - 12c_3 - 1c_4 & -6c_1 + 0c_2 - 4c_3 + 2c_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The matrix equation above is equivalent to the following homogeneous equation.

$$\begin{bmatrix} 3 & 0 & 0 & 1 \\ 6 & -1 & -8 & 0 \\ 3 & -1 & -12 & -1 \\ -6 & 0 & -4 & 2 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

As $\det(A) = 48 \neq 0$

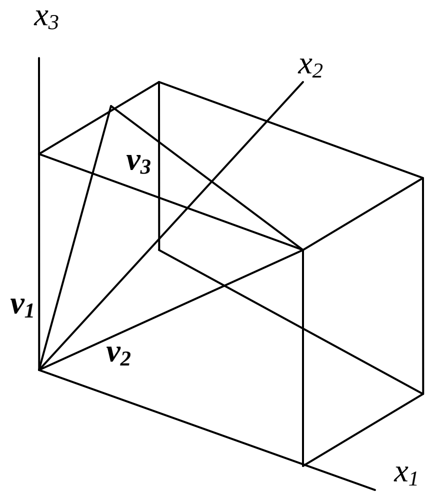
Therefore the set $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}\}$ is linearly independent. Consequently, the set $\boldsymbol{S}$ forms a basis for span $\boldsymbol{S}$.

<u>Example 11</u>: Let $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ -2 \\ -3 \end{bmatrix}$, $\boldsymbol{v_2} = \begin{bmatrix} -3 \\ 5 \\ 7 \end{bmatrix}$, $\boldsymbol{v_3} = \begin{bmatrix} -4 \\ 5 \\ 6 \end{bmatrix}$, and $\boldsymbol{H} = Span\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$.

Note that $\boldsymbol{v_3} = 5\boldsymbol{v_1} + 3\boldsymbol{v_2}$ and show that Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ = Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. Then find a basis for the subspace $\boldsymbol{H}$.

<u>Solution</u>: Every vector in Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ belongs to $\boldsymbol{H}$ because

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + 0\,\boldsymbol{v_3}$$



Now let $\boldsymbol{x}$ be any vector in $\boldsymbol{H}$ – say, $\boldsymbol{x} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + c_3\boldsymbol{v_3}$. Since $\boldsymbol{v_3} = 5\boldsymbol{v_1} + 3\boldsymbol{v_2}$, we may substitute

$$\boldsymbol{x} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + c_3(5\boldsymbol{v_1} + 3\boldsymbol{v_2})$$
$$= (c_1 + 5c_3)\boldsymbol{v_1} + (c_2 + 3c_3)\boldsymbol{v_2}$$

Thus $\boldsymbol{x}$ is in Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$, so every vector in $\boldsymbol{H}$ already belongs to Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. We conclude that $\boldsymbol{H}$ and *Span* $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ are actually the same set of vectors. It follows that $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is a basis of $\boldsymbol{H}$ since $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is obviously linearly independent.

Activity: Show that the following set of vectors is basis for $\mathbb{R}^3$:

1.
$$v_1 = (1, 0, 0),\ v_2 = (0, 2, 1),\ v_1 = (3, 0, 1)$$

2.
$$v_1 = (1, 2, 3),\ v_2 = (0, 1, 1),\ v_1 = (0, 1, 3)$$

The Spanning Set Theorem:

As we will see, a basis is an "efficient" spanning set that contains no unnecessary vectors. In fact, a basis can be constructed from a spanning set by discarding unneeded vectors.

Theorem 2: (The Spanning Set Theorem) Let $\boldsymbol{S} = \{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ be a set in $\boldsymbol{V}$ and let $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$.

- **a.** If one of the vectors in $\boldsymbol{S}$ – say, $\boldsymbol{v_k}$ – is a linear combination of the remaining vectors in $\boldsymbol{S}$, then the set formed from $\boldsymbol{S}$ by removing $\boldsymbol{v_k}$ still spans $\boldsymbol{H}$.
- **b.** If $\boldsymbol{H} \neq \{\boldsymbol{0}\}$, some subset of $\boldsymbol{S}$ is a basis for $\boldsymbol{H}$.

Since we know that span is the set of all linear combinations of some set of vectors and basis is a set of linearly independent vectors whose span is the entire vector space. The spanning set is a set of vectors whose span is the entire vector space. "The Spanning set theorem" is that a spanning set of vectors always contains a subset that is a basis.

Remark: Let $\boldsymbol{V} = \boldsymbol{R}^m$ and let $\boldsymbol{S} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a set of nonzero vectors in $\boldsymbol{V}$.

Procedure:

The procedure for finding a subset of S that is a basis for W = span S is as follows:

Step 1 Write the Equation,

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n = 0 \qquad (3)$$

Step 2 Construct the augmented matrix associated with the homogeneous system of Equation (1) and transforms it to reduced row echelon form.

Step 3 The vectors corresponding to the columns containing the leading 1's form a basis for $\boldsymbol{W}$ = span $\boldsymbol{S}$.

Thus if $\boldsymbol{S} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_6\}$ and the leading 1's occur in columns 1, 3, and 4, then $\{\boldsymbol{v}_1, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ is a basis for span $\boldsymbol{S}$.

Note In step 2 of the procedure above, it is sufficient to transform the augmented matrix to row echelon form.

Example 12: Let $\boldsymbol{S} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4, \boldsymbol{v}_5\}$ be a set of vectors in $\boldsymbol{R}^4$, where $\boldsymbol{v}_1 = (1,2,-2,1)$, $\boldsymbol{v}_2 = (-3,0,-4,3)$, $\boldsymbol{v}_3 = (2,1,1,-1)$, $\boldsymbol{v}_4 = (-3,3,-9,6)$, and $\boldsymbol{v}_5 = (9,3,7,-6)$.
Find a subset of $\boldsymbol{S}$ that is a basis for $\boldsymbol{W}$ = span $\boldsymbol{S}$.

Solution: Step 1 Form Equation (3),

$c_1\,(1,2,-2,1) + c_2(-3,0,-4,3) + c_3(2,1,1,-1) + c_4(-3,3,-9,6) + c_5(9,3.7,-6) = (0,0,0,0)$.

Step 2 Equating corresponding components, we obtain the homogeneous system

$$\begin{aligned}
c_1 - 3c_2 + 2c_3 - 3c_4 + 9c_5 &= 0\\
2c_1 + c_3 + 3c_4 + 3c_5 &= 0\\
-2c_1 - 4c_2 + c_3 - 9c_4 + 7c_5 &= 0\\
c_1 + 3c_2 - c_3 + 6c_4 - 6c_5 &= 0
\end{aligned}$$

The reduced row echelon form of the associated augmented matrix is

$$\begin{bmatrix} 1 & 0 & \tfrac{1}{2} & 3/2 & 3/2 & : & 0\\ 0 & 1 & -1/2 & 3/2 & -5/2 & : & 0\\ 0 & 0 & 0 & 0 & 0 & : & 0\\ 0 & 0 & 0 & 0 & 0 & : & 0 \end{bmatrix}$$

Step 3 The leading 1's appear in columns 1 and 2, so $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is a basis for $\boldsymbol{W}$ = span $\boldsymbol{S}$.

Two Views of a Basis When the Spanning Set Theorem is used, the deletion of vectors from a spanning set must stop when the set becomes linearly independent. If an additional vector is deleted, it will not be a linear combination of the remaining vectors and hence the smaller set will no longer span ***V***. Thus a basis is a spanning set that is as small as possible.
A basis is also a linearly independent set that is as large as possible. If ***S*** is a basis for ***V***, and if ***S*** is enlarged by one vector – say, ***w*** – from ***V***, then the new set cannot be linearly independent, because *S* spans *V*, and *w* is therefore a linear combination of the elements in ***S***.

Example 13: The following three sets in $\boldsymbol{R^3}$ show how a linearly independent set can be enlarged to a basis and how further enlargement destroys the linear independence of the set. Also, a spanning set can be shrunk to a basis, but further shrinking destroys the spanning property.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix}\right\} \quad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix}\right\} \quad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix},\begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

Linearly independent but does not span $\boldsymbol{R^3}$	A basis for $\boldsymbol{R^3}$	Spans $\boldsymbol{R^3}$ but is linearly dependent

Example 14: Let $\boldsymbol{v_1}=\begin{bmatrix}1\\0\\0\end{bmatrix}, \boldsymbol{v_2}=\begin{bmatrix}0\\1\\0\end{bmatrix}$, and $\boldsymbol{H}=\left\{\begin{bmatrix}s\\s\\0\end{bmatrix}: s \text{ in } R\right\}$. then every vector in ***H*** is a linear combination of $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ because $\begin{bmatrix}s\\s\\0\end{bmatrix}=s\begin{bmatrix}1\\0\\0\end{bmatrix}+s\begin{bmatrix}0\\1\\0\end{bmatrix}$. Is $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ a basis for ***H***?

Solution: Neither $\boldsymbol{v_1}$ nor $\boldsymbol{v_2}$ is in ***H***, so $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ cannot a basis for ***H***. In fact, $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is a basis for the plane of all vectors of the form $(c_1, c_2, 0)$, but ***H*** is only a line.

Activity: Find a Basis for the subspace W in $\mathbb{R}^3$ spanned by the following sets of vectors:

1. $v_1=(1,0,2),\ v_2=(3,2,1),\ v_3=(1,0,6),\ v_4=(3,2,1)$
2. $v_1=(1,2,2),\ v_2=(3,2,1),\ v_3=(1,1,7),\ v_4=(7,6,4)$

Exercises:

Determine which set in exercises 1-4 are bases for $\boldsymbol{R}^2$ or $\boldsymbol{R}^3$. Of the sets that are not bases, determine which one are linearly independent and which ones span $\boldsymbol{R}^2$ or $\boldsymbol{R}^3$. Justify your answers.

1. $\begin{bmatrix}1\\0\\-2\end{bmatrix},\begin{bmatrix}3\\2\\-4\end{bmatrix},\begin{bmatrix}-3\\-5\\1\end{bmatrix}$

2. $\begin{bmatrix}1\\-3\\0\end{bmatrix},\begin{bmatrix}-2\\9\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\-3\\5\end{bmatrix}$

3. $\begin{bmatrix}1\\2\\-3\end{bmatrix},\begin{bmatrix}-4\\-5\\6\end{bmatrix}$

4. $\begin{bmatrix}1\\-4\\3\end{bmatrix},\begin{bmatrix}0\\3\\-1\end{bmatrix},\begin{bmatrix}3\\-5\\4\end{bmatrix},\begin{bmatrix}0\\2\\-2\end{bmatrix}$

5. Find a basis for the set of vectors in $\boldsymbol{R}^3$ in the plane x + 2y + z = 0.

6. Find a basis for the set of vectors in $\boldsymbol{R}^2$ on the line y = 5x.

7. Suppose $\boldsymbol{R}^4 = \text{Span}\ \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$. Explain why $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ is a basis for $\boldsymbol{R}^4$.

8. Explain why the following sets of vectors are not bases for the indicated vector spaces. (Solve this problem by inspection).
(a) $\boldsymbol{u}_1 = (1, 2)$, $\boldsymbol{u}_2 = (0, 3)$, $\boldsymbol{u}_3 = (2, 7)$ for $\boldsymbol{R}^2$
(b) $\boldsymbol{u}_1 = (-1, 3, 2)$, $\boldsymbol{u}_2 = (6, 1, 1)$ for $\boldsymbol{R}^3$
(c) $\boldsymbol{p}_1 = \boldsymbol{1} + \boldsymbol{x} + \boldsymbol{x}^2$, $\boldsymbol{p}_2 = \boldsymbol{x} - \boldsymbol{1}$ for $\boldsymbol{P}_2$
(d) $\boldsymbol{A} = \begin{bmatrix}1 & 1\\2 & 3\end{bmatrix}, \boldsymbol{B} = \begin{bmatrix}6 & 0\\-1 & 4\end{bmatrix}, \boldsymbol{C} = \begin{bmatrix}3 & 0\\1 & 7\end{bmatrix}, \boldsymbol{D} = \begin{bmatrix}5 & 1\\4 & 2\end{bmatrix}, \boldsymbol{E} = \begin{bmatrix}7 & 1\\2 & 9\end{bmatrix}$ for $\boldsymbol{M}_{22}$

9. Which of the following sets of vectors are bases for $\boldsymbol{R}^2$?
(a) (2, 1), (3, 0) (b) (4, 1), (-7, -8) (c) (0, 0), (1, 3) (d) (3, 9), (-4, -12)

10. Let $\boldsymbol{V}$ be the space spanned by $\boldsymbol{v}_1 = \text{Cos}^2\ x$, $\boldsymbol{v}_2 = \text{Sin}^2 x$, $\boldsymbol{v}_3 = \cos 2x$.
(a) Show that $\boldsymbol{S} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is not a basis for $\boldsymbol{V}$ (b) Find a basis for $\boldsymbol{V}$

In exercises 11-13, determine a basis for the solution space of the system.

11. $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ -2x_1 - x_2 + 2x_3 &= 0 \\ -x_1 + x_3 &= 0 \end{aligned}$

12. $\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ x_1 + 5x_3 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$

13.
$$\begin{aligned} x + \ \ y + z &= 0 \\ 3x + 2y - 2z &= 0 \\ 4x + 3y - \ \ z &= 0 \\ 6x + 5y + \ \ z &= 0 \end{aligned}$$

14. Determine bases for the following subspace of $\boldsymbol{R}^3$

(a) the plane $3x - 2y + 5z = 0$ (b) the plane $x - y = 0$

(c) the line $x = 2t$, $y = -t$, $z = 4t$ (d) all vectors of the form (a, b, c), where b = a + c

15. Find a standard basis vector that can be added to the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ to produce a basis for $\boldsymbol{R}^3$.

(a) $\boldsymbol{v}_1 = (-1, 2, 3)$, $\boldsymbol{v}_2 = (1, -2, -2)$ (b) $\boldsymbol{v}_1 = (1, -1, 0)$, $\boldsymbol{v}_2 = (3, 1, -2)$

16. Find a standard basis vector that can be added to the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ to produce a basis for $\boldsymbol{R}^4$.

$\boldsymbol{v}_1 = (1, -4, 2, -3)$, $\boldsymbol{v}_2 = (-3, 8, -4, 6)$

Lecture No.23
Coordinate System

OBJECTIVES:

The objectives of the lecture are to learn about:

- Unique representation theorem.
- Coordinate of the element of a vector space relative to the basis B.
- Some examples in which B- coordinate vector is uniquely determined using basis of a vector space.
- Graphical interpretation of coordinates.
- Coordinate Mapping

Theorem:

Let $B = \{b_1, b_2, ..., b_n\}$ be a basis for a vector space V. Then for each x in *V*, there exist a unique set of scalars $c_1, c_2, ..., c_n$ such that

$$x = c_1b_1, c_2b_2, ..., c_nb_n \text{ ………….. (1)}$$

Proof:

Since *B* is a basis for a vector space *V*, then by definition of basis every element of *V* can be written as a linear combination of basis vectors. That is if $x \in V$, then $x = c_1b_1, c_2b_2, ..., c_nb_n$. Now, we show that this representation for x is unique.

For this, suppose that we have two representations for x.

i.e.

$$x = c_1b_1, c_2b_2, ..., c_nb_n \text{ ………….. (2)}$$

and

$$x = d_1b_1, d_2b_2, ..., d_nb_n \text{ ………….. (3)}$$

We will show that the coefficients are actually equal. To do this, subtracting (3) from (2), we have

$$0 = (c_1 - d_1)b_1 + (c_2 - d_2)b_2 + ... + (c_n - d_n)b_n .$$

Since *B* is a basis, it is linearly independent set. Thus the coefficients in the last linear combination must all be zero. That is

$$c_1 = d_1, ..., c_n = d_n .$$

Thus the representation for x is unique.

Definition (B-Coordinate of x):

Suppose that the set $B = \{b_1, b_2, ..., b_n\}$ is a basis for *V* and x is in *V*. The coordinates of x relative to basis *B* (**or the *B*-coordinate of** x) are the weights $c_1, c_2, ..., c_n$ such that

$$x = c_1b_1, c_2b_2, ..., c_nb_n .$$

Note:

If $c_1, c_2, ..., c_n$ are the ***B***- coordinates of x, then the vector in R^n, $[x]_B = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ c_n \end{bmatrix}$ is the coordinate vector of x (relative to ***B)*** or ***B***- coordinates of x.

Example 1:

Consider a basis $B = \{b_1, b_2\}$ for R^2 , where $b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.
Suppose an x in R^2 has the coordinate vector $[x]_B = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Find x

Solution:

Using above definition x is uniquely determined using coordinate vector and the basis. That is

$$\begin{aligned} x &= c_1 b_1, c_2 b_2 \\ &= (-2)b_1 + (3)b_2 \\ &= (-2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 6 \end{bmatrix} \\ x &= \begin{bmatrix} 1 \\ 6 \end{bmatrix} \end{aligned}$$

Example 2:

Let ***S*** **=** {$\boldsymbol{v_1}$, $\boldsymbol{v_2}$, $\boldsymbol{v_3}$} be the basis for $\boldsymbol{R^3}$, where $\boldsymbol{v_1}$ **=** (1, 2, 1), $\boldsymbol{v_2}$ = (2, 9, 0), and $\boldsymbol{v_3}$ = (3, 3, 4).

(a) Find the coordinates vector of ***v*** = (5, -1, 9) with respect to ***S***.

(b) Find the vector ***v*** in $\boldsymbol{R^3}$ whose coordinate vector with respect to the basis ***S*** is

$[\boldsymbol{v}]_{\mathbf{S}} =$ (-1, 3, 2)

Solution:

Since **S** is a basis for $\mathbf{R^3}$, Thus

$$x = c_1 v_1 + c_2 v_2 + c_3 v_3.$$

Further

$(5,-1,9) = c_1(1,2,1) + c_2(2,9,0) + c_3(3,3,4)$ ………….. (A)

To find the coordinate vector of ***v***, we have to find scalars c_1, c_2, c_3.

For this equating corresponding components in (A) gives

$c_1 + 2c_2 + 3c_3 = 5$ (1)

$2c_1 + 9c_2 + 3c_3 = -1$ (2)

$c_1 + \quad 4c_3 = 9$ (3)

Now find values of c_1, c_2 *and* c_3 *from these equations.*

From equation (3)

$c_1 = 9 - 4c_3$

Put this value of c_1 *in equations* (1) *and* (2)

$9 - 4c_3 + 2c_2 + 3c_3 = 5$

$2c_2 - c_3 = -4$ (4)

and

$2(9 - 4c_3) + 9c_2 + 3c_3 = -1$

$18 - 8c_3 + 9c_2 + 3c_3 = -1$

$9c_2 - 5c_3 = -19$ (5)

Multiply equation (4) *by 5*

$10c_2 - 5c_3 = -20$

Subtract equation (5) *from above equation*

$10c_2 - 5c_3 = -20$

$\pm 9c_2 \mp 5c_3 = \mp 19$

$c_2 = -1$

Put value of c_2 *in equation* (4) *to get* c_3

$2(-1) - c_3 = -4$

$-2 - c_3 = -4$

$c_3 = 4 - 2 = 2$

Put value of c_3 *in equation* (3) *to get* c_1

$c_1 + 4(2) = 9$

$c_1 = 9 - 8 = 1$

Thus, we obtain $c_1 = 1, c_2 = -1, c_3 = 2$

Therefore, $[\boldsymbol{v}]_{\mathbf{s}} = (1, -1, 2)$

Figure 2 Using the definition of coordinate vector, we have

$$
\begin{aligned}
v &= c_1v_1 + c_2v_2 + c_3v_3 \\
&= (-1)v_1 + 3v_2 + 2v_3 \\
&= (-1)(1,2,1) + 3(2,9,0) + 2(3,3,4) \\
&= (-1+6+6, -2+27+6, -1+0+8) \\
&= (11,31,7)
\end{aligned}
$$

Therefore

$$v = (11,31,7)$$

Example 3:

Find the coordinates vector of the polynomial $\boldsymbol{p} = \boldsymbol{a_0} + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2$ relative to the basis $\boldsymbol{S} = \{\boldsymbol{1}, \boldsymbol{x}, \boldsymbol{x}^2\}$ for $\boldsymbol{p_2}$.

Solution:

To find the coordinator vector of the polynomial p, we write it as a linear combination of the basis set S. That is

$$a_0 + a_1x + a_2x^2 = c_1(1) + c_2(x) + c_3(x^2)$$
$$\Rightarrow c_1 = a_0, c_2 = a_1, c_3 = a_2$$

Therefore

$$[p]_s = (a_0, a_1, a_2)$$

Example 4:

Find the coordinates vector of the polynomial $\boldsymbol{p} = \boldsymbol{5} - 4\boldsymbol{x} + 3\boldsymbol{x}^2$ relative to the basis $\boldsymbol{S} = \{\boldsymbol{1}, \boldsymbol{x}, \boldsymbol{x}^2\}$ for $\boldsymbol{p_2}$.

Solution:

To find the coordinator vector of the polynomial p, we write it as a linear combination of the basis set S. That is

$$5 - 4x + 3x^2 = c_1(1) + c_2(x) + c_3(x^2)$$
$$\Rightarrow c_1 = 5, c_2 = -4, c_3 = 3$$

Therefore

$$[p]_s = (5,-4,3)$$

Example 5:

Find the coordinate vector of A relative to the basis S = {A1, A2, A3, A4}

$$\boldsymbol{A} = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix}; \boldsymbol{A_1} = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}; \boldsymbol{A_2} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}; \boldsymbol{A_3} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}; \boldsymbol{A_4} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Solution:

To find the coordinator vector of A, we write it as a linear combination of the basis set S. That is

$$\boldsymbol{A} = c_1 \boldsymbol{A_1} + c_2 \boldsymbol{A_2} + c_3 \boldsymbol{A_3} + c_4 \boldsymbol{A_4}$$

$$\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} = c_1 \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -c_1 & c_1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} c_2 & c_2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c_3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c_4 \end{bmatrix}$$

$$= \begin{bmatrix} -c_1 + c_2 + 0 + 0 & c_1 + c_2 + 0 + 0 \\ 0 + 0 + c_3 + 0 & 0 + 0 + 0 + c_4 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} -c_1 + c_2 & c_1 + c_2 \\ c_3 & c_4 \end{bmatrix}$$

$$-c_1 + c_2 = 2 \qquad (1)$$

$$c_1 + c_2 = 0 \qquad (2)$$

$$c_3 = -1 \qquad (3)$$

$$c_4 = 3 \qquad (4)$$

Adding (1) and (2), gives

$2c_2 = 2 \Rightarrow c_2 = 1$

Putting the value of c_2 in (2) to get c_1, $c_1 = -1$

So $c_1 = -1, c_2 = 1, c_3 = -1, c_4 = 3$

Therefore, $[\boldsymbol{v}]_s = (-1, 1, -1, 3)$

Graphical Interpretation of Coordinates

A coordinate system on a set consists of a one-to-one mapping of the points in the set into $\boldsymbol{R}^n$. For example, ordinary graph paper provides a coordinate system for the plane when one selects perpendicular axes and a unit of measurement on each axis. Figure 1 shows the standard basis $\{\boldsymbol{e_1}, \boldsymbol{e_2}\}$, the vectors

$\boldsymbol{b_1}$ (= $\boldsymbol{e_1}$) and $\boldsymbol{b_2}$ from Example 1, that is,

$\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Vector $x = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$, the coordinates 1 and 6 give the location of x relative to the standard basis: 1 unit in the e_1 direction and 6 units in the e_2 direction.

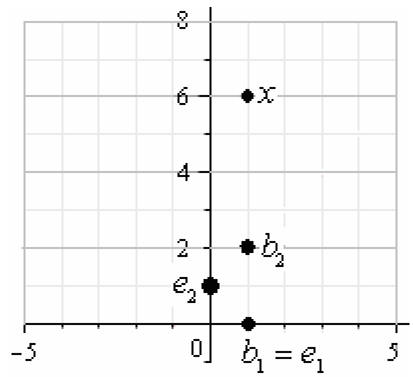


Figure 1

Figure 2 shows the vectors b_1, b_2, and x from Figure 1. (Geometrically, the three vectors lie on a vertical line in both figures.) However, the standard coordinate grid was erased and replaced by a grid especially adapted to the basis B in Example 1. The coordinate vector $[x]_B = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ gives the location of x on this new coordinate system: – 2 units in the b_1 direction and *3* units in the b_2 direction.

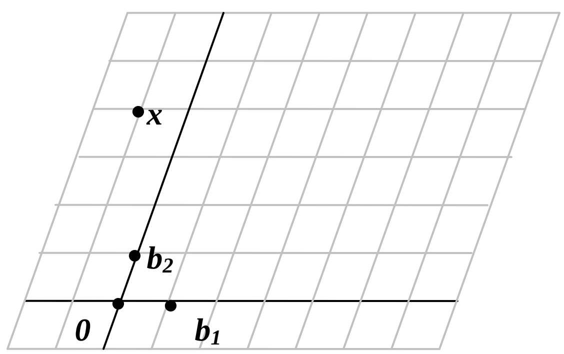
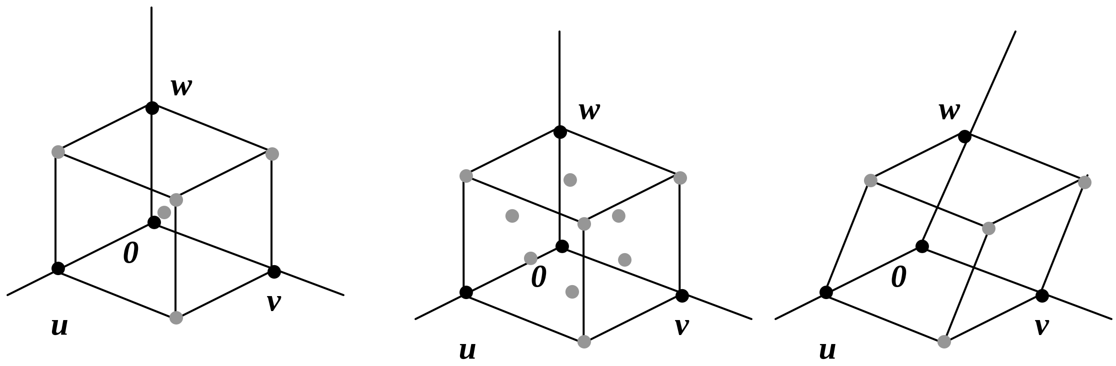


Figure 2

<u>Example 6</u>:

In crystallography, the description of a crystal lattice is aided by choosing a basis $\{u, v, w\}$ for R^3 that corresponds to three adjacent edges of one "unit cell" of the crystal. An entire lattice is constructed by stacking together many copies of one cell. There are fourteen basic types of unit cells; three are displayed in Figure 3.



(a) Body centered cubic

(b) Face centered orthorhombic

(c) Simple monoclinic

Figure 3 – Examples of unit cells

The coordinates of atoms within the crystal are given relative to the basis for the lattice. For instance, $\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$ identifies the top face-centered atom in the cell in Figure 3(b).

Coordinates in R^n When a basis B for R^n is fixed, the B-coordinate vector of a specified x is easily found, as in the next example.

Example 7: Let $b_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, x = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, and $B = \{b_1, b_2\}$.

Find the coordinate vector $[x]_B$ of x relative to B.

Solution The B – coordinates c_1, c_2 of x satisfy

$$\underset{b_1}{c_1\begin{bmatrix} 2 \\ 1 \end{bmatrix}} + \underset{b_2}{c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}} = \underset{x}{\begin{bmatrix} 4 \\ 5 \end{bmatrix}}$$

or

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \qquad (3)$$

$$b_1 \quad b_2 \qquad\qquad x$$

Now, inverse of matrix $\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{3} & \frac{2}{3} \end{bmatrix}$

From equation (3) we get

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3}(4) + \frac{1}{3}(5) \\ \frac{-1}{3}(4) + \frac{2}{3}(5) \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Thus, $c_1 = 3, c_2 = 2$.
(Equation (3) can also be solved by row operations on an augmented matrix. Try it yourself)

Thus $\boldsymbol{x} = 3\boldsymbol{b_1} + 2\boldsymbol{b_2}$ and $[\ \boldsymbol{x}\ ]_B = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

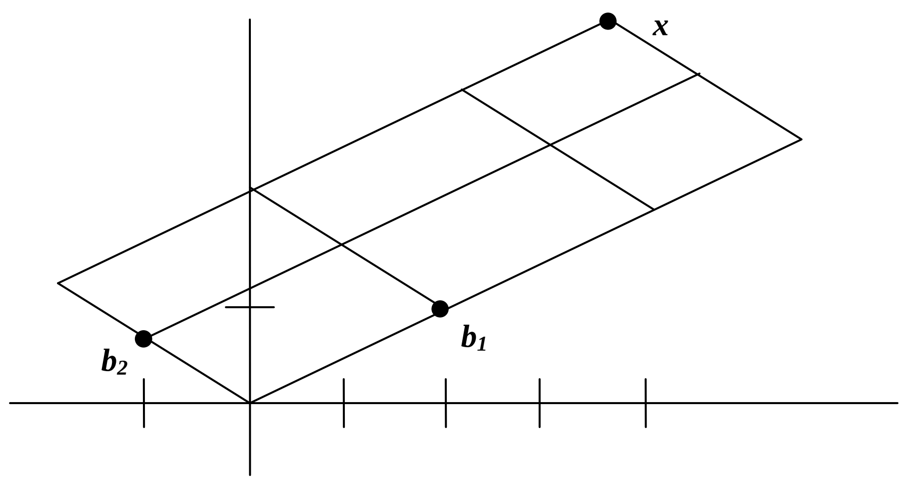


Figure 4 – The *B*-coordinate vector of *x* is (3,2)

The matrix in (3) changes the ***B***-coordinates of a vector ***x*** into the standard coordinates for ***x***. An analogous change of coordinates can be carried out in $\boldsymbol{R^n}$ for a basis
$\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \ldots, \boldsymbol{b_n}\}$.

Let $\boldsymbol{P_B} = [\boldsymbol{b_1} \quad \boldsymbol{b_2} \quad \ldots \quad \boldsymbol{b_n}]$

Then the vector equation $\boldsymbol{x} = c_1\boldsymbol{b_1} + c_2\boldsymbol{b_2} + \ldots + c_n\boldsymbol{b_n}$

is equivalent to $\boldsymbol{x} = \boldsymbol{P_B}[\boldsymbol{x}]_{\boldsymbol{B}}$ (4)

We call $\boldsymbol{P_B}$ the **change-of-coordinates matrix** from ***B*** to the standard basis in $\boldsymbol{R^n}$.
Left-multiplication by $\boldsymbol{P_B}$ transforms the coordinate vector $[\boldsymbol{x}]_{\boldsymbol{B}}$ into ***x***. The change-of-coordinates equation (4) is important and will be needed at several points in next lectures.

Since the columns of P_B form a basis for R^n, P_B is invertible (by the Invertible Matrix Theorem). Left-multiplication by P_B^{-1} converts x into its B-coordinate vector:

$$P_B^{-1}x = [x]_B$$

The correspondence $x \to [x]_B$ produced here by P_B^{-1}, is the coordinate mapping mentioned earlier. Since P_B^{-1} is an invertible matrix, the coordinate mapping is a one-to-one linear transformation from R^n onto R^n, by the Invertible Matrix Theorem. (See also Theorem 3 in lecture 10) This property of the coordinate mapping is also true in a general vector space that has a basis, as we shall see.

The Coordinate Mapping Choosing a basis $B = \{b_1, b_2, \ldots, b_n\}$ for a vector space V introduces a coordinate system in V. The coordinate mapping $x \to [x]_B$ connects the possibly unfamiliar space V to the familiar space R^n. See Figure 5. Points in V can now be identified by their new "names".

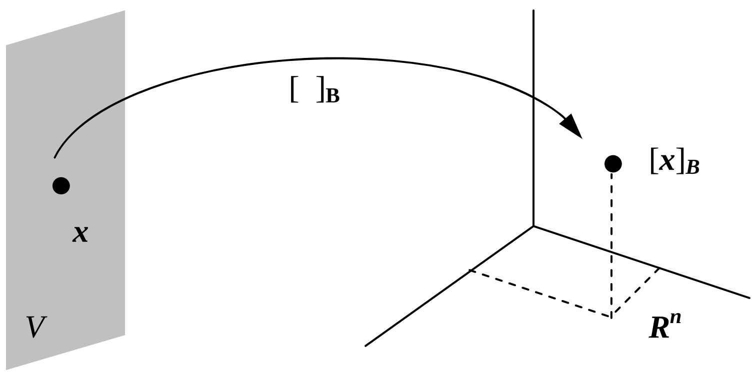


Figure 5 – The coordinate mapping from *V* onto R^n

Theorem 2: Let $B = \{b_1, b_2, \ldots, b_n\}$ be a basis for a vector space V. Then the coordinate mapping $x \to [x]_B$ is a one-to-one linear transformation from V onto R^n.

Proof Take two typical vectors in V, say

$$u = c_1b_1 + c_2b_2 + \ldots + c_nb_n$$

$$w = d_1b_1 + d_2b_2 + \ldots + d_nb_n$$

Then, using vector operations, $u + w = (c_1 + d_1)b_1 + (c_2 + d_2)b_2 + \ldots + (c_n + d_n)b_n$

It follows that $[u+w]_B = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [u]_B + [w]_B$

Thus the coordinate mapping preserves addition. If r is any scalar, then

$$ru = r(c_1b_1 + c_2b_2 + \cdots + c_nb_n) = (rc_1)b_1 + (rc_2)b_2 + \cdots + (rc_n)b_n$$

So $$[r\boldsymbol{u}]_{\boldsymbol{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[\boldsymbol{u}]_{\boldsymbol{B}}$$

Thus the coordinate mapping also preserves scalar multiplication and hence is a linear transformation. It can be verified that the coordinate mapping is one-to-one and maps ***V*** onto $\boldsymbol{R^n}$.

The linearity of the coordinate mapping extends to linear combinations, just as in lecture 9. If $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_p}$ are in ***V*** and if $c_1, c_2, \dots, c_p$ are scalars, then

$$[c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \dots + c_p\boldsymbol{u_p}]_{\boldsymbol{B}} = c_1[\boldsymbol{u_1}]_{\boldsymbol{B}} + c_2[\boldsymbol{u_2}]_{\boldsymbol{B}} + \dots + c_p[\boldsymbol{u_p}]_{\boldsymbol{B}} \qquad (5)$$

In words, (5) says that the ***B***-coordinate vector of a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_p}$ is the same linear combination of their coordinate vectors.

The coordinate mapping in Theorem 2 is an important example of an isomorphism from ***V*** onto $\boldsymbol{R^n}$. In general, a one-to-one linear transformation from a vector space ***V*** onto a vector space ***W*** is called an **isomorphism** from ***V*** onto ***W*** (iso from the Greek for "the same", and morph from the Greek for "form" or "structure"). The notation and terminology for ***V*** and ***W*** may differ, but the two spaces are indistinguishable as vector spaces. Every vector space calculation in ***V*** is accurately reproduced in ***W***, and vice versa.

Example 8: Let ***B*** be the standard basis of the space $\boldsymbol{P_3}$ of polynomials; that is, let $\boldsymbol{B} = \{\boldsymbol{1}, \boldsymbol{t}, \boldsymbol{t^2}, \boldsymbol{t^3}\}$. A typical element ***p*** of $\boldsymbol{P_3}$ has the form $\boldsymbol{p}(\boldsymbol{t}) = \boldsymbol{a_0} + a_1\boldsymbol{t} + a_2\boldsymbol{t^2} + a_3\boldsymbol{t^3}$
Since ***p*** is already displayed as a linear combination of the standard basis vectors, we conclude that $[\boldsymbol{p}]_{\boldsymbol{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$. Thus the coordinate mapping $\boldsymbol{p} \to [\boldsymbol{p}]_{\boldsymbol{B}}$ is an isomorphism from $\boldsymbol{P_3}$ onto $\boldsymbol{R^4}$. All vector space operations in $\boldsymbol{P_3}$ correspond to operations in $\boldsymbol{R^4}$.

If we think of $\boldsymbol{P_3}$ and $\boldsymbol{R^4}$ as displays on two computer screens that are connected via the coordinate mapping, then every vector space operation in $\boldsymbol{P_3}$ on one screen is exactly duplicated by a corresponding vector operation in $\boldsymbol{R^4}$ on the other screen. The vectors on the $\boldsymbol{P_3}$ screen look different from those on the $\boldsymbol{R^4}$ screen, but they "act" as vectors in exactly the same way. See Figure 6.

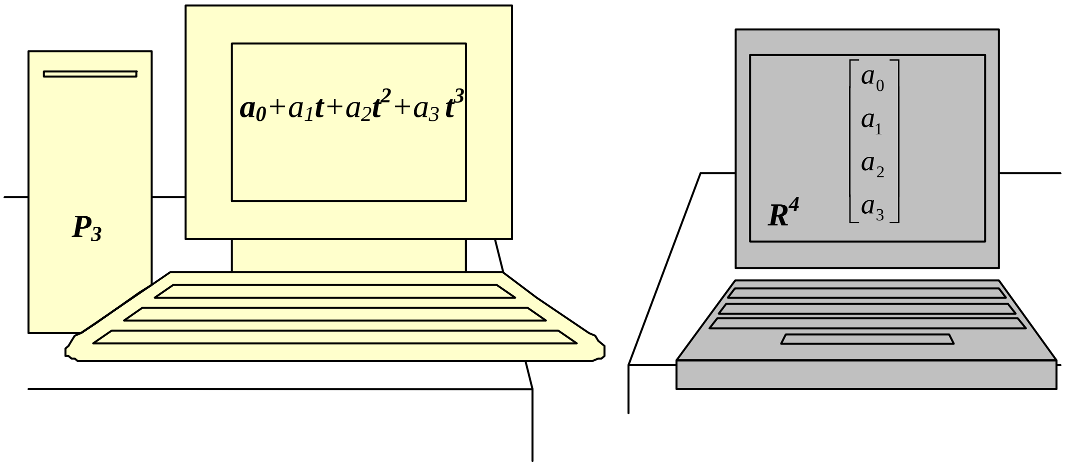


Figure 6 – The space P_3 is isomorphic to R^4

<u>Example 9</u>: Use coordinate vector to verify that the polynomials $1 + 2t^2$, $4 + t + 5t^2$ and $3 + 2t$ are linearly dependent in P_2.

<u>Solution:</u> The coordinate mapping from Example 8 produces the coordinate vectors (1, 0, 2), (4, 1, 5) and (3, 2, 0), respectively. Writing these vectors as the columns of a matrix A, we can determine their independence by row reducing the augmented matrix for $Ax = 0$: $\begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 5 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

The columns of A are linearly dependent, so the corresponding polynomials are linearly dependent. In fact, it is easy to check that column 3 of A is 2 times column 2 minus 5 times column 1. The corresponding relation for the polynomials is

$$3 + 2t = 2(4 + t + 5t^2) - 5(1 + 2t^2)$$

<u>Example 10</u>: Let $v_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $x = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$, and $B = \{v_1, v_2\}$. Then B is a basis for $H = Span\{v_1, v_2\}$. Determine if x is in H and if it is, find the coordinate vector of x relative to B.

<u>Solution:</u> If x is in H, then the following vector equation is consistent.

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

The scalars, c_1 and c_2, if they exist, are the B – coordinates of x.

Using row operations, we obtain $\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$.

Thus $c_1 = 2$, $c_2 = 3$ and $[\mathbf{x}]_B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. The coordinate system on ***H*** determined by ***B*** is shown in Figure 7.

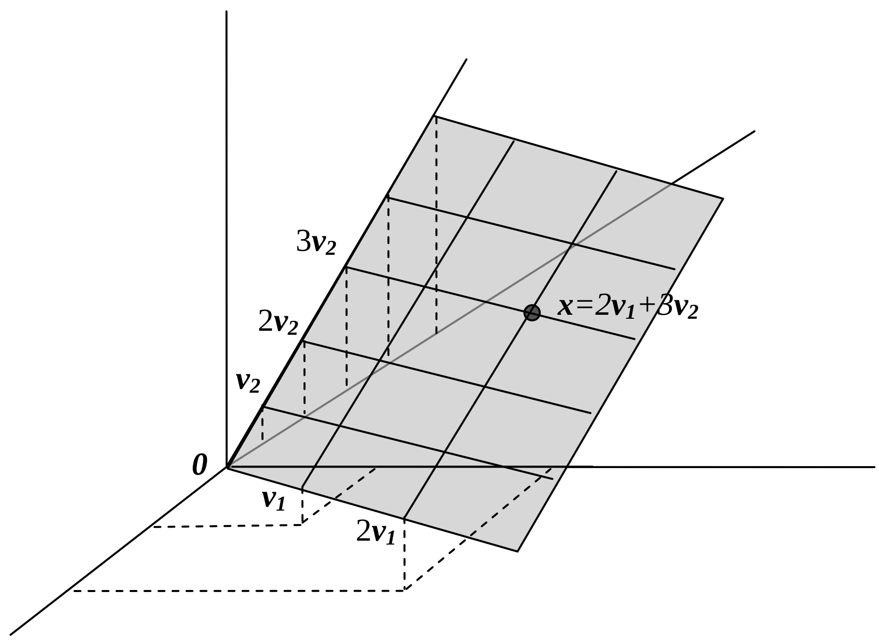


Figure 7 – A coordinate system on a plane *H* in R^3

If a different basis for ***H*** were chosen, would the associated coordinate system also make ***H*** isomorphic to R^2? Surely, this must be true. We shall prove it in the next lecture.

<u>Example 11</u>: Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} -8 \\ 2 \\ 3 \end{bmatrix}$.

a. Show that the set $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is a basis of R^3.
b. Find the change-of-coordinates matrix from ***B*** to the standard basis.
c. Write the equation that relates ***x*** in R^3 to $[\mathbf{x}]_B$.
d. Find $[\mathbf{x}]_B$, for the ***x*** given above.

<u>Solution:</u>

a. It is evident that the matrix $P_B = [\mathbf{b}_1 \ \ \mathbf{b}_2 \ \ \mathbf{b}_3]$ is row equivalent to the identity matrix. By the Invertible Matrix Theorem, P_B is invertible and its columns form a basis for R^3.

b. From part (a), the change-of-coordinates matrix is $P_B = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}$.

c. $x = P_B[x]_B$.

d. To solve part (c), it is probably easier to row reduce an augmented matrix instead of computing P_B^{-1}. We have

$$\left[\begin{array}{ccc|c} 1 & -3 & 3 & -8 \\ 0 & 4 & -6 & 2 \\ 0 & 0 & 3 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$P_B \qquad x \qquad I \qquad [x]_B$$

Hence $[x]_B = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$

Example 12: The set $B = \{1 + t,\ 1 + t^2,\ t + t^2\}$ is a basis for P_2. Find the coordinate vector of $p(t) = 6 + 3t - t^2$ relative to B.

Solution: The coordinates of $p(t) = 6 + 3t - t^2$ with respect to B satisfy

$$c_1(1 + t) + c_2(1 + t^2) + c_3(t + t^2) = 6 + 3t - t^2$$

$$c_1 + c_1 t + c_2 + c_2 t^2 + c_3 t + c_3 t^2 = 6 + 3t - t^2$$

$$c_1 + c_2 + c_1 t + c_3 t + c_2 t^2 + c_3 t^2 = 6 + 3t - t^2$$

$$c_1 + c_2 + (c_1 + c_3)t + (c_2 + c_3)t^2 = 6 + 3t - t^2$$

Equating coefficients of like powers of t, we have

$c_1 + c_2 \quad = 6$ --------------(1)

$c_1 \quad + c_3 = 3$ --------------(2)

$\quad c_2 + c_3 = -1$ --------------(3)

Subtract equation (2) *from* (1) *we get*

$c_2 - c_3 = 6 - 3 = 3$

Add this equation with equation (3)

$2c_2 = -1 + 3 = 2$

$\Rightarrow c_2 = 1$

Put value of c_2 *in equation* (3)

$1 + c_3 = -1$

$\Rightarrow c_3 = -2$

$$\textit{From equation (1) we have}$$
$$c_1 + c_2 = 6$$
$$c_1 = 6 - 1 = 5$$

Solving, we find that $c_1 = 5$, $c_2 = 1$, $c_3 = -2$, and $[\boldsymbol{p}]_B = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$.

Exercises:

In exercises 1 and 2, find the vector $\boldsymbol{x}$ determined by the given coordinate vector $[\boldsymbol{x}]_B$ and the given basis $\boldsymbol{B}$.

1. $\boldsymbol{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}, [\boldsymbol{x}]_B = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

2. $\boldsymbol{B} = \left\{ \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -7 \\ 0 \end{bmatrix} \right\}, [\boldsymbol{x}]_B = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$

In exercises 3-6, find the coordinate vector $[\boldsymbol{x}]_B$ of $\boldsymbol{x}$ relative to the given basis $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \ldots, \boldsymbol{b_n}\}$.

3. $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

4. $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 5 \\ -6 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

5. $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} -3 \\ 4 \\ 9 \end{bmatrix}, \boldsymbol{b_3} = \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 8 \\ -9 \\ 6 \end{bmatrix}$

6. $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix}, \boldsymbol{b_3} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix}$

In exercises 7 and 8, find the change of coordinates matrix from $\boldsymbol{B}$ to standard basis in $\boldsymbol{R}^n$.

7. $\boldsymbol{B} = \left\{ \begin{bmatrix} 2 \\ -9 \end{bmatrix}, \begin{bmatrix} 1 \\ 8 \end{bmatrix} \right\}$

8. $\boldsymbol{B} = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}, \begin{bmatrix} 8 \\ -2 \\ 7 \end{bmatrix} \right\}$

In exercises 9 and 10, use an inverse matrix to find $[\boldsymbol{x}]_{\boldsymbol{B}}$ for the given $\boldsymbol{x}$ and $\boldsymbol{B}$.

9. $\boldsymbol{B}=\left\{\begin{bmatrix}4\\5\end{bmatrix},\begin{bmatrix}6\\7\end{bmatrix}\right\}, \boldsymbol{x}=\begin{bmatrix}2\\0\end{bmatrix}$

10. $\boldsymbol{B}=\left\{\begin{bmatrix}3\\-5\end{bmatrix},\begin{bmatrix}-4\\6\end{bmatrix}\right\}, \boldsymbol{x}=\begin{bmatrix}2\\-6\end{bmatrix}$

11. The set $\boldsymbol{B} = \{\mathbf{1} + \boldsymbol{t}^2, \boldsymbol{t} + \boldsymbol{t}^2, \mathbf{1} + 2\boldsymbol{t} + \boldsymbol{t}^2\}$ is a basis for $\boldsymbol{P_2}$. Find the coordinate vector of $\boldsymbol{p}(\boldsymbol{t}) = \boldsymbol{1} + 4\boldsymbol{t} + 7\boldsymbol{t}^2$ relative to $\boldsymbol{B}$.

12. The vectors $\boldsymbol{v}_1=\begin{bmatrix}1\\-3\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}2\\-8\end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix}-3\\7\end{bmatrix}$ span $\boldsymbol{R}^2$ but do not form a basis. Find two different ways to express $\begin{bmatrix}1\\1\end{bmatrix}$ as a linear combination of $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$.

13. Let $\boldsymbol{B}=\left\{\begin{bmatrix}1\\-4\end{bmatrix},\begin{bmatrix}-2\\9\end{bmatrix}\right\}$. Since the coordinate mapping determined by $\boldsymbol{B}$ is a linear transformation from $\boldsymbol{R}^2$ into $\boldsymbol{R}^2$, this mapping must be implemented by some 2 x 2 matrix $\boldsymbol{A}$. Find it.

In exercises 14-16, use coordinate vectors to test the linear independence of the sets of polynomials.

14. $\mathbf{1} + \boldsymbol{t}^3, \mathbf{3} + \boldsymbol{t} - 2\boldsymbol{t}^2, -\boldsymbol{t} + 3\boldsymbol{t}^2 - \boldsymbol{t}^3$

15. $(\boldsymbol{t}-\mathbf{1})^2, \boldsymbol{t}^3 - \mathbf{2}, (\mathbf{t} - \mathbf{2})^3$

16. $\mathbf{3} + 7\boldsymbol{t}, \mathbf{5} + \boldsymbol{t} - 2\boldsymbol{t}^3, \boldsymbol{t} - 2\boldsymbol{t}^2, \mathbf{1} + 16\boldsymbol{t} - 6\boldsymbol{t}^2 + 2\boldsymbol{t}^3$

17. Let $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ and $\boldsymbol{B} = \{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. Show that $\boldsymbol{x}$ is in $\boldsymbol{H}$ and find the $\boldsymbol{B}$-coordinate vector of $\boldsymbol{x}$, for $\boldsymbol{v}_1=\begin{bmatrix}11\\-5\\10\\7\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}14\\-8\\13\\10\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}19\\-13\\18\\15\end{bmatrix}$.

18. Let $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ and $\boldsymbol{B} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$. Show that $\boldsymbol{B}$ is a basis for $\boldsymbol{H}$ and $\boldsymbol{x}$ is in $\boldsymbol{H}$, and find the $\boldsymbol{B}$-coordinate vector of $\boldsymbol{x}$, for $\boldsymbol{v}_1=\begin{bmatrix}-6\\4\\-9\\4\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}8\\-3\\7\\-3\end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix}-9\\5\\-8\\3\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}4\\7\\-8\\3\end{bmatrix}$.

Lecture 24

Dimension of a Vector Space

In this lecture, we will focus over the dimension of the vector spaces. The dimension of a vector space V is the cardinality or the number of vectors in the basis B of the given vector space. If the basis B has *n* (say) elements then this number *n* (called the dimension) is an intrinsic property of the space ***V***. That is it does not depend on the particular choice of basis rather, all the bases of V will have the same cardinality. Thus, we can say that the dimension of a vector space is always unique. The discussion of dimension will give additional insight into properties of bases.
The first theorem generalizes a well-known result about the vector space $\boldsymbol{R}^n$.

Note:
A vector space ***V*** with a basis ***B*** containing *n* vectors is isomorphic to $\boldsymbol{R}^n$ i.e., there exist a one-to-one linear transformation from V to $\boldsymbol{R}^n$.

Theorem 1: If a vector space ***V*** has a basis $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$, then any set in ***V*** containing more than *n* vectors must be linearly dependent.

Theorem 2: If a vector space ***V*** has a basis of *n* vectors, then every basis of ***V*** must consist of exactly *n* vectors.

Finite and infinite dimensional vector spaces:

If the vector space ***V*** is spanned or generated by a finite set, then ***V*** is said to be **finite-dimensional**, and the **dimension** of ***V***, written as dim *V*, is the number of vectors in a basis for ***V***.. If ***V*** is not spanned by a finite set, then ***V*** is said to be **infinite-dimensional**. That is, if we are unable to find a finite set that is able to generate the whole vector space, then such a vector space is called infinite dimensional.

Note:
(1) The dimension of the zero vector space $\{0\}$ is defined to be zero.
(2) Every finite dimensional vector space contains a basis.

Example 1: The set of real numbers of n dimension $\boldsymbol{R}^n$, *set of polynomials of order n* $\boldsymbol{P}_n$, *and set of matrices of order* $m \times n$ $\boldsymbol{M}_{mn}$ are all finite- dimensional vector spaces..
However, the vector spaces
$F(-\infty, \infty)$, $C(-\infty, \infty)$, and $C^m(-\infty, \infty)$ are infinite- dimensional.

Example 2:
(a) Any pair of non-parallel vectors ***a***, ***b*** in the xy-plane, which are necessarily linearly independent, can be regarded as a basis of the subspace $\boldsymbol{R}^2$. In particular the set of unit vectors $\{\boldsymbol{i}, \boldsymbol{j}\}$ forms a basis for $\boldsymbol{R}^2$. Therefore, dim $(\boldsymbol{R}^2) = 2$.

Any set of three non coplanar vectors $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ in ordinary (physical) space, which will be necessarily linearly independent, spans the space $\boldsymbol{R}^3$. Therefore any set of such vectors forms a basis for $\boldsymbol{R}^3$. In particular the set of unit vectors $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ forms a basis of $\boldsymbol{R}^3$. This basis is called standard basis for $\boldsymbol{R}^3$. Therefore dim $(\boldsymbol{R}^3) = 3$.

The set of vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ where

$$\begin{aligned} \boldsymbol{e}_1 &= (1, 0, 0, 0, \ldots, 0),\\ \boldsymbol{e}_2 &= (0, 1, 0, 0, \ldots, 0),\\ \boldsymbol{e}_3 &= (0, 0, 1, 0, \ldots, 0),\\ &\ldots\\ &\ldots\\ &\ldots\\ \boldsymbol{e}_n &= (0, 0, 0, 0, \ldots, 1) \end{aligned}$$

is linearly independent.

Moreover, any vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ in $\boldsymbol{R}^n$ can be expressed as a linear combination of these vectors as

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3 + \ldots + x_n\boldsymbol{e}_n.$$

Hence, the set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ forms a basis for $\boldsymbol{R}^n$. It is called the standard basis of $\boldsymbol{R}^n$, therefore dim $(\boldsymbol{R}^n) = n$. Any other set of n linearly independent vectors in $\boldsymbol{R}^n$ will form a non-standard basis.

(b) The set $\boldsymbol{B} = \{\boldsymbol{1}, \boldsymbol{x}, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n\}$ forms a basis for the vector space $\boldsymbol{P}_n$ of polynomials of degree $\leq n$. It is called the standard basis with dim $(\boldsymbol{P}_n) = n + 1$.

(c) The set of 2 x 2 matrices with real entries (elements) $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ where

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{u}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \boldsymbol{u}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

is a linearly independent and every 2 x 2 matrix with real entries can be expressed as their linear combination. Therefore, they form a basis for the vector space $\boldsymbol{M}_{2X2}$. This basis is called the standard basis for $\boldsymbol{M}_{2X2}$ with dim $(\boldsymbol{M}_{2X2}) = 4$.

Note:

(1) dim $(\boldsymbol{R}^n) = n$ { The standard basis has n vectors}.

(2) dim $(\boldsymbol{P}_n) = n + 1$ { The standard basis has $n+1$ vectors}.

(3) dim $(\boldsymbol{M}_{m\times n}) = mn$ { The standard basis has mn vectors.}

Example 3: Let $\boldsymbol{W}$ be the subspace of the set of all (2 x 2) matrices defined by

$$\boldsymbol{W} = \{\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : 2a - b + 3c + d = 0\}.$$

Determine the dimension of $\boldsymbol{W}$.

Solution: The algebraic specification for $\boldsymbol{W}$ can be rewritten as $d = -2a + b - 3c$.

Now $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Substituting the value of d, it becomes

$$A = \begin{bmatrix} a & b \\ c & -2a+b-3c \end{bmatrix}$$

This can be written as

$$A = \begin{bmatrix} a & 0 \\ 0 & -2a \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & b \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c & -3c \end{bmatrix}$$

$$= a\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix}$$

$$= a\boldsymbol{A_1} + b\boldsymbol{A_2} + c\boldsymbol{A_3}$$

where $\boldsymbol{A_1} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$, $\boldsymbol{A_2} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, and $\boldsymbol{A_3} = \begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix}$

The matrix $\boldsymbol{A}$ is in $\boldsymbol{W}$ if and only if $\boldsymbol{A} = a\boldsymbol{A_1} + b\boldsymbol{A_2} + c\boldsymbol{A_3}$, so $\{\boldsymbol{A_1}, \boldsymbol{A_2}, \boldsymbol{A_3}\}$ is a spanning set for $\boldsymbol{W}$. Now, check whether if this set is a basis for W or not, we will see whether $\{\boldsymbol{A_1}, \boldsymbol{A_2}, \boldsymbol{A_3}\}$ is linearly independent or not. For this purpose, we will see that $\{\boldsymbol{A_1}, \boldsymbol{A_2}, \boldsymbol{A_3}\}$ is linearly independent if

$aA_1 + bA_2 + cA_3 = 0 \Rightarrow a=b=c=0$ *i.e.,*

$$a\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} a & 0 \\ 0 & -2a \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & b \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c & -3c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

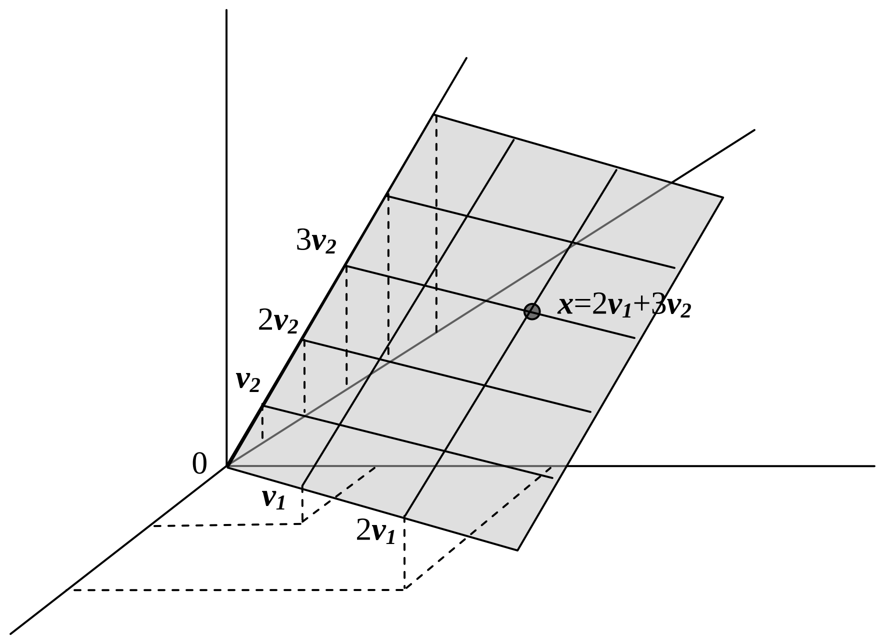
$$\begin{bmatrix} a & b \\ c & -2a+b-3c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Equating the elements, we get

$a = 0, b = 0, c = 0$

This implies $\{\boldsymbol{A_1}, \boldsymbol{A_2}, \boldsymbol{A_3}\}$ *is a linearly independent set that spans W. Hence, it's the basis of W with dim(W)= 3.*

<u>Example 4</u>: Let $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$, where $\boldsymbol{v_1} = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Then $\boldsymbol{H}$ is the plane studied in Example 10 of lecture 23**.** A basis for $\boldsymbol{H}$ is $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$, since $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are not multiples and hence are linearly independent. Thus, dim $\boldsymbol{H} = 2$.



A coordinate system on a plane H in R^3

Example 5: Find the dimension of the subspace

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} : a,b,c,d \in R \right\}$$

Solution: The representative vector of H can be written as

$$\begin{bmatrix} a-3b+6c \\ 5a+4d \\ b-2c-d \\ 5d \end{bmatrix} = a\begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix} + d\begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Now, it is easy to see that H is the set of all linear combinations of the vectors

$$v_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Clearly, $v_1 \neq 0, v_2$ is not a multiple of v_1, but v_3 is a multiple of v_2. By the Spanning Set Theorem, we may discard v_3 and still have a set that spans H. Finally; v_4 is not a linear combination of v_1 and v_2. So $\{v_1, v_2, v_4\}$ is linearly independent and hence is a basis for H. Thus dim $H = 3$.

Example 6: The subspaces of $\boldsymbol{R^3}$ can be classified by its/various dimensions as shown in Fig. 1.

***0*-dimensional subspaces**:

The only 0-dimensional subspace of $\boldsymbol{R^3}$ is zero subspace.

***1*-dimensional subspaces:**

1-dimensional subspaces include any subspace spanned by a single non-zero vector. Such subspaces are lines through the origin.

2-dimensional subspaces:

Any subspace spanned by two linearly independent vectors. Such subspaces are planes through the origin.

3-dimensional subspaces:

The only 3-dimensional subspace is $\boldsymbol{R^3}$ itself. Any three linearly independent vectors in $\boldsymbol{R^3}$ span all of $\boldsymbol{R^3}$, by the Invertible Matrix Theorem.

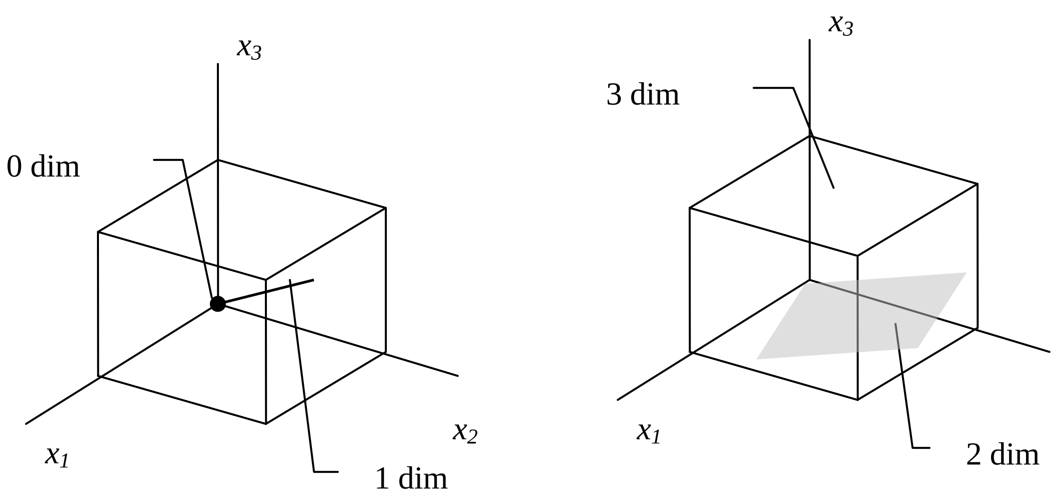


Figure 1 – Sample subspaces of $\boldsymbol{R^3}$

Bases for Nul *A* and Col *A*: We already know how to find vectors that span the null space of a matrix ***A***. The discussion in Lecture 21 pointed out that our method always produces a linearly independent set. Thus the method produces a basis for Nul ***A***.

Example 7: Find a basis for the null space of $A = \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$.

<u>Solution</u>: The null space of ***A*** is the solution space of homogeneous system

$$\begin{aligned} 2x_1 + 2x_2 - x_3 \quad\quad + x_5 &= 0 \\ -x_1 - x_2 + 2x_3 - 3x_4 + x_5 &= 0 \\ x_1 + x_2 - 2x_3 \quad\quad - x_5 &= 0 \\ x_3 + x_4 + x_5 &= 0 \end{aligned}$$

The most appropriate way to solve this system is to reduce its augmented matrix into reduced echelon form.

$$\begin{bmatrix} 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \\ 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} R_4 \sim R_2,\ R_3 \sim R_1$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{bmatrix} R_3 - 2R_1,\ R_3 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 0 & 3 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{bmatrix} R_3 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{bmatrix} -\frac{1}{3}R_3$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{bmatrix} R_4 + R_1$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \end{bmatrix} R_4 + 3R_3$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} R_1 + 2R_2$$

$$\sim\begin{bmatrix}1 & 1 & 0 & 2 & 1 & 0\\0 & 0 & 1 & 1 & 1 & 0\\0 & 0 & 0 & 1 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0\end{bmatrix} R_2-R_3,\ R_1-2R_3$$

$$\sim\begin{bmatrix}1 & 1 & 0 & 0 & 1 & 0\\0 & 0 & 1 & 0 & 1 & 0\\0 & 0 & 0 & 1 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

Thus, the reduced row echelon form of the augmented matrix is

$$\begin{bmatrix}1 & 1 & 0 & 0 & 1 & 0\\0 & 0 & 1 & 0 & 1 & 0\\0 & 0 & 0 & 1 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

which corresponds to the system

$$\begin{aligned}1x_1 + 1\,x_2 + \quad\quad 1\,x_5 &= 0\\ 1\,x_3 + \quad 1\,x_5 &= 0\\ 1\,x_4 \quad &= 0\\ 0 &= 0\end{aligned}$$

No equation of this system has a form zero = nonzero. Therefore, the system is consistent. Since the number of unknowns is more than the number of equations, we will assign some arbitrary value to some variables. This will lead to infinite many solutions of the system.

$$\begin{aligned}x_1 &= -1x_2 - 1x_5\\ x_2 &= s\\ x_3 &= -1x_5\\ x_4 &= 0\\ x_5 &= t\end{aligned}$$

The general solution of the given system is

$$x_1 = -s - t \ , \quad x_2 = s \ , \quad x_3 = -t \ , \quad x_4 = 0 \ , \quad x_5 = t$$

Therefore, the solution vector can be written as

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix} = \begin{bmatrix}-s-t\\s\\-t\\0\\t\end{bmatrix} = \begin{bmatrix}-s\\s\\0\\0\\0\end{bmatrix} + \begin{bmatrix}-t\\0\\-t\\0\\t\end{bmatrix} = s\begin{bmatrix}-1\\1\\0\\0\\0\end{bmatrix} + t\begin{bmatrix}-1\\0\\-1\\0\\1\end{bmatrix}$$

which shows that the vectors $v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$ span the solution space .Since they are also linearly independent, $\{v_1, v_2\}$ is a basis for Nul A.

The next two examples describe a simple algorithm for finding a basis for the column space.

Example 8: Find a basis for Col B, where $B = [b_1 \quad b_2, \quad ..., \quad b_5] = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Solution: Each non-pivot column of B is a linear combination of the pivot columns. Infact, $b_2 = 4b_1$ and $b_4 = 2b_1 - b_3$. By the Spanning Set Theorem, we may discard b_2 and b_4 and $\{b_1, b_3, b_5\}$ will still span Col B. Let

$$S = \{b_1, \quad b_3, \quad b_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Since $b_1 \neq 0$ and no vector in S is a linear combination of the vectors that precede it, S is linearly independent. Thus S is a basis for Col B.

What about a matrix A that is not in reduced echelon form? Recall that any linear dependence relationship among the columns of A can be expressed in the form $Ax = 0$, where x is a column of weights. (If some columns are not involved in a particular dependence relation, then their weights are zero.) When A is row reduced to a matrix B, the columns of B are often totally different from the columns of A. However, the equations $Ax = 0$ and $Bx = 0$ have exactly the same set of solutions. That is, the columns of A have exactly the same linear dependence relationships as the columns of B.

Elementary row operations on a matrix do not affect the linear dependence relations among the columns of the matrix.

Example 9: It can be shown that the matrix

$$A = [a_1 \quad a_2 \quad ... \quad a_5] = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$$

is row equivalent to the matrix B in Example 8. Find a basis for Col A.

Solution: In Example 8, we saw that $\boldsymbol{b}_2 = 4\boldsymbol{b}_1 \quad and \quad \boldsymbol{b}_4 = 2\boldsymbol{b}_1 - \boldsymbol{b}_3$ so we can expect that $\boldsymbol{a}_2 = 4\boldsymbol{a}_1 \quad and \quad \boldsymbol{a}_4 = 2\boldsymbol{a}_1 - \boldsymbol{a}_3$. This is indeed the case. Thus, we may discard $\boldsymbol{a}_2$ and $\boldsymbol{a}_4$ while selecting a minimal spanning set for Col ***A***. Infact, $\{\boldsymbol{a}_1, \boldsymbol{a}_3, \boldsymbol{a}_5\}$ must be linearly independent because any linear dependence relationship among $\boldsymbol{a}_1, \boldsymbol{a}_3, \boldsymbol{a}_5$ would imply a linear dependence relationship among $\boldsymbol{b}_1, \boldsymbol{b}_3, \boldsymbol{b}_5$. But we know that $\{\boldsymbol{b}_1, \boldsymbol{b}_3, \boldsymbol{b}_5\}$ is a linearly independent set. Thus $\{\boldsymbol{a}_1, \boldsymbol{a}_3, \boldsymbol{a}_5\}$ is a basis for Col ***A***. The columns we have used for this basis are the pivot columns of ***A***.

Examples 8 and 9 illustrate the following useful fact.

Theorem 3: The pivot columns of a matrix ***A*** form a basis for Col ***A***.

Proof: The general proof uses the arguments discussed above. Let ***B*** be the reduced echelon form of ***A***. The set of pivot columns of ***B*** is linearly independent, for no vector in the set is a linear combination of the vectors that precede it. Since ***A*** is row equivalent to ***B***, the pivot columns of ***A*** are linearly independent too, because any linear dependence relation among the columns of ***A*** corresponds to a linear dependence relation among the columns of ***B***. For this same reason, every non-pivot column of ***A*** is a linear combination of the pivot columns of ***A***. Thus the non-pivot columns of ***A*** may be discarded from the spanning set for Col ***A***, by the Spanning Set Theorem. This leaves the pivot columns of ***A*** as a basis for Col ***A***.

Note: Be careful to use pivot columns of ***A*** itself for the basis of Col ***A***. The columns of an echelon form ***B*** are often not in the column space of ***A***. For instance, the columns of the ***B*** in Example 8 all have zeros in their last entries, so they cannot span the column space of the ***A*** in Example 9.

Example 10: Let $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ and $\boldsymbol{v}_2 = \begin{bmatrix} -2 \\ 7 \\ -9 \end{bmatrix}$. Determine if $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is a basis for $\boldsymbol{R}^3$. Is $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ a basis for $\boldsymbol{R}^2$?

Solution: Let $A = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2]$. Row operations show that $A = \begin{bmatrix} 1 & -2 \\ -2 & 7 \\ 3 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Not every row of A contains a pivot position. So the columns of A do not span $\boldsymbol{R}^3$, by Theorem 4 in Lecture 6. Hence $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is not a basis for $\boldsymbol{R}^3$. Since $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are not in $\boldsymbol{R}^2$, they cannot possibly be a basis for $\boldsymbol{R}^2$. However, since $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are obviously linearly independent, they are a basis for a subspace of $\boldsymbol{R}^3$, namely, Span $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$.

Example 11: Let $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 6 \\ 2 \\ -1 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}, \boldsymbol{v}_4 = \begin{bmatrix} -4 \\ -8 \\ 9 \end{bmatrix}$. Find a basis for the subspace ***W*** spanned by $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$.

Solution: Let ***A*** be the matrix whose column space is the space spanned by $\{v_1, v_2, v_3, v_4\}$,

$$A = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix}$$

Reduce the matrix ***A*** into its echelon form in order to find its pivot columns.

$$A = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 20 & 4 & -20 \\ 0 & -25 & -5 & 25 \end{bmatrix} by\ R_2 + 3R_1,\ R_3 - 4R_1$$

$$\sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 5 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix} by\ \frac{1}{4}R_2,\ -\frac{1}{5}R_3,\ R_3 - R_2$$

The first two columns of ***A*** are the pivot columns and hence form a basis of Col ***A = W***. Hence $\{v_1, v_2\}$ is a basis for ***W***.
Note that the reduced echelon form of ***A*** is not needed in order to locate the pivot columns.

Procedure:

Basis and Linear Combinations

Given a set of vectors $S = \{v_1, v_2, \ldots, v_k\}$ in R^n, the following procedure produces a subset of these vectors that form a basis for span (***S***) and expresses those vectors of ***S*** that are not in the basis as linear combinations of the basis vector.

Step1: Form the matrix ***A*** having $v_1, v_2, \ldots, v_k$ as its column vectors.

Step2: Reduce the matrix ***A*** to its reduced row echelon form ***R***, and let $w_1, w_2, \ldots, w_k$ be the column vectors of ***R***.

Step3: Identify the columns that contain the leading entries i.e., 1's in ***R***. The corresponding column vectors of ***A*** are the basis vectors for span (***S***).

Step4: Express each column vector of ***R*** that does not contain a leading entry as a linear combination of preceding column vector that do contain leading entries (we will be able to do this by inspection). This yields a set of dependency equations involving the column vectors of ***R***. The corresponding equations for the column vectors of ***A*** express the vectors which are not in the basis as linear combinations of basis vectors.

Example 12: Basis and Linear Combinations

(a) Find a subset of the vectors $v_1 = (1, -2, 0, 3)$, $v_2 = (2, -4, 0, 6)$, $v_3 = (-1, 1, 2, 0)$ and $v_4 = (0, -1, 2, 3)$ that form a basis for the space spanned by these vectors.

(b) Express each vector not in the basis as a linear combination of the basis vectors.

Solution:

(a) We begin by constructing a matrix that has v_1, v_2, v_3, v_4 as its column vectors

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -4 & 1 & -1 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 0 & 3 \end{bmatrix} \qquad \text{(A)}$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow \\ \boldsymbol{v_1} & \boldsymbol{v_2} & \boldsymbol{v_3} & \boldsymbol{v_4} \end{matrix}$$

Finding a basis for column space of this matrix can solve the first part of our problem.
Transforming Matrix to Reduced Row Echelon Form:

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -4 & 1 & -1 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 0 & 3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix} \begin{matrix} 2R_1 + R_2 \\ -3R_1 + R_4 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix} -1R_2$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} -2R_2 + R_3 \\ -3R_2 + R_4 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} R_2 + R_1$$

Labeling the column vectors of the resulting matrix as $\boldsymbol{w_1}$, $\boldsymbol{w_2}$, $\boldsymbol{w_3}$ and $\boldsymbol{w_4}$ yields

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(B)}$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow \\ \boldsymbol{w_1} & \boldsymbol{w_2} & \boldsymbol{w_3} & \boldsymbol{w_4} \end{matrix}$$

The leading entries occur in column 1 and 3 so $\{w_1, w_3\}$ is a basis for the column space of (B) and consequently $\{v_1, v_3\}$ is the basis for column space of (A).
(b) We shall start by expressing w_2 and w_4 as linear combinations of the basis vector w_1 and w_3. The simplest way of doing this is to express w_2 and w_4 in term of basis vectors with smaller subscripts. Thus we shall express w_2 as a linear combination of w_1, and we shall express w_4 as a linear combination of w_1 and w_3. By inspection of (B), these linear combinations are $w_2 = 2w_1$ and $w_4 = w_1 + w_3$. We call them the dependency equations. The corresponding relationship of (A) are $v_3 = 2v_1$ and $v_5 = v_1 + v_3$.

<u>Example 13</u>: Basis and Linear Combinations
(a) Find a subset of the vectors $v_1 = (1, -1, 5, 2)$, $v_2 = (-2, 3, 1, 0)$, $v_3 = (4, -5, 9, 4)$, $v_4 = (0, 4, 2, -3)$ and $v_5 = (-7, 18, 2, -8)$ that form a basis for the space spanned by these vectors.
(b) Express each vector not in the basis as a linear combination of the basis vectors
<u>Solution</u>:
(a) We begin by constructing a matrix that has $v_1, v_2, \ldots, v_5$ as its column vectors

$$\begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ -1 & 3 & -5 & 4 & 18 \\ 5 & 1 & 9 & 2 & 2 \\ 2 & 0 & 4 & -3 & -8 \end{bmatrix} \qquad (A)$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ v_1 & v_2 & v_3 & v_4 & v_5 \end{matrix}$$

Finding a basis for column space of this matrix can solve the first part of our problem.
Transforming Matrix to Reduced Row Echelon Form:

$$\begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ -1 & 3 & -5 & 4 & 18 \\ 5 & 1 & 9 & 2 & 2 \\ 2 & 0 & 4 & -3 & -8 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 4 & 11 \\ 0 & 11 & -11 & 2 & 37 \\ 0 & 4 & -4 & -3 & 6 \end{bmatrix} \begin{matrix} R_1 + R_2 \\ -5R_1 + R_3 \\ -2R_1 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 4 & 11 \\ 0 & 0 & 0 & -42 & -84 \\ 0 & 0 & 0 & -19 & -38 \end{bmatrix} \begin{matrix} -11R_2 + R_3 \\ -4R_2 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 4 & 11 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -19 & -38 \end{bmatrix} (-1/42)R_3$$

$$\begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 4 & 11 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} 19R_3 + R_4$$

$$\begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (-4)R_3 + R_2$$

$$\begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} 2R_2 + R_1$$

Denoting the column vectors of the resulting matrix by $\boldsymbol{w_1}$, $\boldsymbol{w_2}$, $\boldsymbol{w_3}$, $\boldsymbol{w_4}$, and $\boldsymbol{w_5}$ yields

$$\begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (B)$$

$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ w_1 & w_2 & w_3 & w_4 & w_5 \end{matrix}$$

The leading entries occur in columns 1,2 and 4 so that $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_4}\}$ is a basis for the column space of (B) and consequently $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_4}\}$ is the basis for column space of (A).
(b) We shall start by expressing $\boldsymbol{w_3}$ and $\boldsymbol{w_5}$ as linear combinations of the basis vector $\boldsymbol{w_1}$, $\boldsymbol{w_2}$, $\boldsymbol{w_4}$. The simplest way of doing this is to express $\boldsymbol{w_3}$ and $\boldsymbol{w_5}$ in term of basis vectors with smaller subscripts. Thus we shall express $\boldsymbol{w_3}$ as a linear combination of $\boldsymbol{w_1}$ and $\boldsymbol{w_2}$, and we shall express $\boldsymbol{w_5}$ as a linear combination of $\boldsymbol{w_1}$, $\boldsymbol{w_2}$, and $\boldsymbol{w_4}$. By inspection of (B), these linear combination are $\boldsymbol{w_3} = 2\boldsymbol{w_1} - \boldsymbol{w_2}$ and $\boldsymbol{w_5} = -\boldsymbol{w_1} + 3\boldsymbol{w_2} + 2\boldsymbol{w_4}$.
The corresponding relationship of (A) are $\boldsymbol{v_3} = 2\boldsymbol{v_1} - \boldsymbol{v_2}$ and $\boldsymbol{v_5} = -\boldsymbol{v_1} + 3\boldsymbol{v_2} + 2\boldsymbol{v_4}$.

Example 14: Basis and Linear Combinations

(a) Find a subset of the vectors $\boldsymbol{v_1} = (1, -2, 0, 3)$, $\boldsymbol{v_2} = (2, -5, -3, 6)$, $\boldsymbol{v_3} = (0, 1, 3, 0)$, $\boldsymbol{v_4} = (2, -1, 4, -7)$ and $\boldsymbol{v_5} = (5 , -8, 1, 2)$ that form a basis for the space spanned by these vectors.
(b) Express each vector not in the basis as a linear combination of the basis vectors.

Solution:

(a) We begin by constructing a matrix that has $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots , \boldsymbol{v_5}$ as its column vectors

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix}$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \boldsymbol{v_1} & \boldsymbol{v_2} & \boldsymbol{v_3} & \boldsymbol{v_4} & \boldsymbol{v_5} \end{matrix} \qquad \text{(A)}$$

Finding a basis for column space of this matrix can solve the first part of our problem. Reducing the matrix to reduced-row echelon form and denoting the column vectors of the resulting matrix by $\boldsymbol{w_1}$, $\boldsymbol{w_2}$, $\boldsymbol{w_3}$, $\boldsymbol{w_4}$, and $\boldsymbol{w_5}$ yields

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(B)}$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \boldsymbol{w_1} & \boldsymbol{w_2} & \boldsymbol{w_3} & \boldsymbol{w_4} & \boldsymbol{w_5} \end{matrix}$$

The leading entries occur in columns 1, 2 and 4 so $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_4}\}$ is a basis for the column space of (B) and consequently $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_4}\}$ is the basis for column space of (A).
(b) Dependency equations are $\boldsymbol{w_3} = 2\boldsymbol{w_1} - \boldsymbol{w_2}$ and $\boldsymbol{w_5} = \boldsymbol{w_1} + \boldsymbol{w_2} + \boldsymbol{w_4}$
The corresponding relationship of (A) are $\boldsymbol{v_3} = 2\boldsymbol{v_1} - \boldsymbol{v_2}$ ***and*** $\boldsymbol{v_5} = \boldsymbol{v_1} + \boldsymbol{v_2} + \boldsymbol{v_4}$

Subspaces of a Finite-Dimensional Space: The next theorem is a natural counterpart to the Spanning Set Theorem.

Theorem 5: Let ***H*** be a subspace of a finite-dimensional vector space ***V***. Any linearly independent set in ***H*** can be expanded, if necessary, to a basis for ***H***. Also, ***H*** is finite-dimensional and $\dim \boldsymbol{H} \le \dim \boldsymbol{V}$.

When the dimension of a vector space or subspace is known, the search for a basis is simplified by the next theorem. It says that if a set has the right number of elements, then one has only to show either that the set is linearly independent or that it spans the space. The theorem is of critical importance in numerous applied problems (involving differential equations or difference equations, for example) where linear independence is much easier to verify than spanning.

Theorem 5: (The Basis Theorem) Let ***V*** be a *p*-dimensional vector space, $p \geq 1$. Any linearly independent set of exactly *p* elements in ***V*** is automatically a basis for ***V***. Any set of exactly *p* elements that spans ***V*** is automatically a basis for ***V***.

The Dimensions of Nul *A* and Col *A*: Since the pivot columns of a matrix ***A*** form a basis for Col ***A***, we know the dimension of Col ***A*** as soon as we know the pivot columns. The dimension of Nul ***A*** might seem to require more work, since finding a basis for Nul ***A*** usually takes more time than a basis for Col ***A***. Yet, there is a shortcut.

Let A be an $m \times n$ matrix, and suppose that the equation $\boldsymbol{Ax} = \boldsymbol{0}$ has k free variables. From lecture 21, we know that the standard method of finding a spanning set for Nul $\boldsymbol{A}$ will produce exactly k linearly independent vectors say, $\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}$, one for each free variable. So $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ is a basis for Nul $\boldsymbol{A}$, and the number of free variables determines the size of the basis. Let us summarize these facts for future reference.

The dimension of Nul $\boldsymbol{A}$ is the number of free variables in the equation $\boldsymbol{Ax} = \boldsymbol{0}$, and the dimension of Col $\boldsymbol{A}$ is the number of pivot columns in $\boldsymbol{A}$.

Example 15: Find the dimensions of the null space and column space of

$$\boldsymbol{A} = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

Solution: Row reduce the augmented matrix $[\boldsymbol{A} \quad \boldsymbol{0}]$ to echelon form and obtain

$$\begin{bmatrix} 1 & -2 & 2 & 3 & -1 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Writing it in equations form, we get

$$x_1 - 2x_2 + 2x_3 + 3x_4 - x_5 = 0$$

$$x_3 + 2x_4 - 2x_5 = 0$$

Since the number of unknowns is more than the number of equations, we will introduce free variables here (say) x_2, x_4 and x_5. Hence the dimension of Nul $\boldsymbol{A}$ is *3*. Also dim Col $\boldsymbol{A}$ *is 2* because $\boldsymbol{A}$ has two pivot columns.

Example 16: Decide whether each statement is true or false, and give a reason for each answer. Here $\boldsymbol{V}$ is a non-zero finite-dimensional vector space.

1. If dim $\boldsymbol{V} = p$ and if $\boldsymbol{S}$ is a linearly dependent subset of $\boldsymbol{V}$, then $\boldsymbol{S}$ contains more than p vectors.
2. If $\boldsymbol{S}$ spans $\boldsymbol{V}$ and if $\boldsymbol{T}$ is a subset of $\boldsymbol{V}$ that contains more vectors than $\boldsymbol{S}$, then $\boldsymbol{T}$ is linearly dependent.

Solution:

1. False. Consider the set $\{\boldsymbol{0}\}$.
2. True. By the Spanning Set Theorem, $\boldsymbol{S}$ contains a basis for $\boldsymbol{V}$; call that basis $\boldsymbol{S'}$. Then $\boldsymbol{T}$ will contain more vectors than $\boldsymbol{S'}$. By Theorem 1, $\boldsymbol{T}$ is linearly dependent.

Exercises:

For each subspace in exercises 1-6, (a) find a basis and (b) state the dimension.

1. $\left\{\begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix} : s, t \text{ in } \boldsymbol{R}\right\}$

2. $\left\{\begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix} : a, b, c \text{ in } \boldsymbol{R}\right\}$

3. $\left\{\begin{bmatrix} a-4b-2c \\ 2a+5b-4c \\ -a+2c \\ -3a+7b+6c \end{bmatrix} : a, b, c \text{ in } \boldsymbol{R}\right\}$

4. $\left\{\begin{bmatrix} 3a+6b-c \\ 6a-2b-2c \\ -9a+5b+3c \\ -3a+b+c \end{bmatrix} : a, b, c \text{ in } \boldsymbol{R}\right\}$

5. {(a, b, c): a – 3b + c = 0, b – 2c = 0, 2b – c = 0}

6 {(a, b, c, d): a - 3b + c = 0}

7. Find the dimension of the subspace $\boldsymbol{H}$ of $\boldsymbol{R}^2$ spanned by

$$\begin{bmatrix} 2 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 10 \end{bmatrix}, \begin{bmatrix} -3 \\ 6 \end{bmatrix}$$

8. Find the dimension of the subspace spanned by the given vectors.

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} -7 \\ -3 \\ 1 \end{bmatrix}$$

Determine the dimensions of Nul $\boldsymbol{A}$ and Col $\boldsymbol{A}$ for the matrices shown in exercises 9 to 12.

9. $\boldsymbol{A} = \begin{bmatrix} 1 & -6 & 9 & 0 & -2 \\ 0 & 1 & 2 & -4 & 5 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

10. $\boldsymbol{A} = \begin{bmatrix} 1 & 3 & -4 & 2 & -1 & 6 \\ 0 & 0 & 1 & -3 & 7 & 0 \\ 0 & 0 & 0 & 1 & 4 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

11. $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 9 & 5 \\ 0 & 0 & 1 & -4 \end{bmatrix}$

12. $\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$

13. The first four Hermite polynomials are **1**, 2$\boldsymbol{t}$, **-2** + 4$\boldsymbol{t^2}$, and -12$\boldsymbol{t}$ + 8$\boldsymbol{t^3}$. These polynomials arise naturally in the study of certain important differential equations in mathematical physics. Show that the first four Hermite polynomials form a basis of $\boldsymbol{P_3}$.

14. Let $\boldsymbol{B}$ be the basis of $\boldsymbol{P_3}$ consisting of the Hermite polynomials in exercise 13, and let $\boldsymbol{p}(\boldsymbol{t}) = \mathbf{7} - 12\,\boldsymbol{t} - 8\,\boldsymbol{t}^2 + 12\,\boldsymbol{t}^3$. Find the coordinate vector of $\boldsymbol{p}$ relative to $\boldsymbol{B}$.

15. Extend the following vectors to a basis for $\boldsymbol{R^5}$:

$$\boldsymbol{v_1} = \begin{bmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{bmatrix}, \boldsymbol{v_2} = \begin{bmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{bmatrix}, \boldsymbol{v_3} = \begin{bmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{bmatrix}$$

Lecture 25

Rank

With the help of vector space concepts, for a matrix several interesting and useful relationships in matrix rows and columns have been discussed.
For instance, imagine placing 2000 random numbers into a 40 x 50 matrix ***A*** and then determining both the maximum number of linearly independent columns in ***A*** and the maximum number of linearly independent columns in A^T (rows in ***A***). Remarkably, the two numbers are the same. Their common value is called the rank of the matrix. To explain why, we need to examine the subspace spanned by the subspace spanned by the rows of ***A***.

<u>The Row Space</u>: If ***A*** is an $m \times n$ matrix, each row of ***A*** has *n* entries and thus can be identified with a vector in R^n. The set of all linear combinations of the row vectors is called the row space of ***A*** and is denoted by Row ***A***. Each row has *n* entries, so Row ***A*** is a subspace of R^n. Since the rows of ***A*** are identified with the columns of A^T, we could also write Col A^T in place of Row ***A***.

<u>Example 1</u>: Let $A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$ and

$r_1 = (-2,-5,8,0,-17)$
$r_2 = (1,3,-5,1,5)$
$r_3 = (3,11,-19,7,1)$
$r_4 = (1,7,-13,5,-3)$

The row space of ***A*** is the subspace of R^5 spanned by $\{r_1, r_2, r_3, r_4\}$. That is, Row ***A*** = Span $\{r_1, r_2, r_3, r_4\}$. Naturally, we write row vectors horizontally; however, they could also be written as column vectors

<u>Example</u>: Let

$A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & -1 & 4 \end{bmatrix}$ and

$r_1 = (2,1,0)$
$r_2 = (3,-1,4)$

That is Row **A=**Span $\{r_1, r_2\}$.

We could use the Spanning Set Theorem to shrink the spanning set to a basis.
Some times row operation on a matrix will not give us the required information but row reducing certainly worthwhile, as the next theorem shows

<u>Theorem 1</u>: If two matrices ***A*** and ***B*** are row equivalent, then their row spaces are the same. If ***B*** is in echelon form, the nonzero rows of ***B*** form a basis for the row space of ***A*** as well as ***B***.

<u>Theorem 2</u>: If ***A*** and ***B*** are row equivalent matrices, then

(a) A given set of column vectors of ***A*** is linearly independent if and only if the corresponding column vectors of ***B*** are linearly independent.
(b) A given set of column vector of ***A*** forms a basis for the column space of ***A*** if and only if the corresponding column vector of ***B*** forms a basis for the column space of ***B***.

Example 2: (Bases for Row and Column Spaces)

Find the bases for the row and column spaces of $A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$.

Solution: We can find a basis for the row space of ***A*** by finding a basis for the row space of any row-echelon form of ***A***.

Now
$$\begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$$
$$\begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} -2R_1 + R_2 \\ -2R_1 + R_3 \\ R_1 + R_4 \end{matrix}$$
$$\begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} -1R_2 + R_3$$

Row-echelon form of ***A***: $R = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Here Theorem 1 implies that that the non zero rows are the basis vectors of the matrix. So these bases vectors are

$$r_1 = [1 \quad -3 \quad 4 \quad -2 \quad 5 \quad 4]$$
$$r_2 = [0 \quad 0 \quad 1 \quad 3 \quad -2 \quad -6]$$
$$r_3 = [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 5]$$

A and ***R*** may have different column spaces, we cannot find a basis for the column space of ***A*** directly from the column vectors of ***R***. however, it follows from the theorem (2b) if we can find a set of column vectors of ***R*** that forms a basis for the column space of ***R***, then the corresponding column vectors of ***A*** will form a basis for the column space of ***A***.

The first, third, and fifth columns of ***R*** contains the leading 1's of the row vectors, so

$$\boldsymbol{c}_1' = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{c}_3' = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{c}_5' = \begin{bmatrix} 5 \\ -2 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of ***R***, thus the corresponding column vectors of ***A***

namely, $$\boldsymbol{c}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ -1 \end{bmatrix} \qquad \boldsymbol{c}_3 = \begin{bmatrix} 4 \\ 9 \\ 9 \\ -4 \end{bmatrix} \qquad \boldsymbol{c}_5 = \begin{bmatrix} 5 \\ 8 \\ 9 \\ -5 \end{bmatrix}$$

form a basis for the column space of ***A***.

<u>Example</u>:
The matrix

$$R = \begin{bmatrix} 1 & -2 & 5 & 0 & 3 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is in row-echelon form.
The vectors

$$\boldsymbol{r}_1 = \begin{bmatrix} 1 & -2 & 5 & 0 & 3 \end{bmatrix}$$
$$\boldsymbol{r}_2 = \begin{bmatrix} 0 & 1 & 3 & 0 & 0 \end{bmatrix}$$
$$\boldsymbol{r}_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

form a basis for the row space of R, and the vectors

$$c_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; c_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, c_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R.

Example 3: **(Basis for a Vector Space using Row Operation)**
Find bases for the space spanned by the vectors

$\mathbf{v}_1 = (1,-2,0,0,3) \qquad \mathbf{v}_2 = (2,-5,-3,-2,6)$

$\mathbf{v}_3 = (0,5,15,10,0) \qquad \mathbf{v}_4 = (2,6,18,8,6)$

Solution: **T**he space spanned by these vectors is the row space of the matrix

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

Transforming Matrix to Row Echelon Form:

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 5 & 15 & 10 & 0 \\ 0 & 10 & 18 & 8 & 0 \end{bmatrix} \begin{matrix} (-2)R_1 + R_2 \\ (-2)R_1 + R_4 \\ (-1)R_2 \end{matrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -12 & -12 & 0 \end{bmatrix} \begin{matrix} (-5)R_2 + R_3 \\ (-10)R_2 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & -12 & -12 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} R_{34}$$

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3$$

Therefore, $\boldsymbol{R} = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

The non-zero row vectors in this matrix are

$$\boldsymbol{w_1} = (1,-2,0,0,3), \boldsymbol{w_2} = (0,1,3,2,0), \boldsymbol{w_3} = (0,0,1,1,0)$$

These vectors form a basis for the row space and consequently form a basis for the subspace of $\boldsymbol{R}^5$ spanned by $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$.

Example 4: **(Basis for the Row Space of a Matrix)**

Find a basis for the row space of $\boldsymbol{A} = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$ consisting entirely of row vectors from $\boldsymbol{A}$.

Solution: We find $\boldsymbol{A}^T$; then we will use the method of example (2) to find a basis for the column space of $\boldsymbol{A}^T$; and then we will transpose again to convert column vectors back to row vectors. Transposing $\boldsymbol{A}$ yields

$$\boldsymbol{A}^T = \begin{bmatrix} 1 & 2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 6 \end{bmatrix}$$

Transforming Matrix to Row Echelon Form:

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & -1 & 5 & 10 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} 2R_1 + R_2 \\ (-3)R_1 + R_5 \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} (-1)R_2$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} (3)R_2 + R_3 \\ (2)R_2 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} 12\,R_3 + R_4$$

Now $$\boldsymbol{R} = \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The first, second and fourth columns contain the leading 1's, so the corresponding column vectors in $\boldsymbol{A^T}$ form a basis for the column space of $\boldsymbol{A^T}$; these are

$$\boldsymbol{c_1} = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 3 \end{bmatrix},\ \boldsymbol{c_2} = \begin{bmatrix} 2 \\ -5 \\ -3 \\ -2 \\ 6 \end{bmatrix} \text{and}\ \boldsymbol{c_4} = \begin{bmatrix} 2 \\ 6 \\ 18 \\ 8 \\ 6 \end{bmatrix}$$

Transposing again and adjusting the notation appropriately yields the basis vectors

$$\boldsymbol{r_1} = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \end{bmatrix}, \boldsymbol{r_2} = \begin{bmatrix} 2 & -5 & -3 & -2 & 6 \end{bmatrix} \textit{and}\ \boldsymbol{r_4} = \begin{bmatrix} 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

for the row space of $\boldsymbol{A}$.

The following example shows how one sequence of row operations on $\boldsymbol{A}$ leads to bases for the three spaces: Row $\boldsymbol{A}$, Col $\boldsymbol{A}$, and Nul $\boldsymbol{A}$.

Example 5: Find bases for the row space, the column space and the null space of the matrix

$$\boldsymbol{A} = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

Solution: To find bases for the row space and the column space, row reduce $\boldsymbol{A}$ to an echelon form: $\boldsymbol{A} \sim \boldsymbol{B} = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

By Theorem (1), the first three rows of $\boldsymbol{B}$ form a basis for the row space of $\boldsymbol{A}$ (as well as the row space of $\boldsymbol{B}$). Thus Basis for Row $\boldsymbol{A}$:

$$\{(1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20)\}$$

For the column space, observe from $\boldsymbol{B}$ that the pivots are in columns *1, 2* and *4*. Hence columns *1, 2* and *4* of $\boldsymbol{A}$ (not $\boldsymbol{B}$) form a basis for Col $\boldsymbol{A}$:

$$\text{Basis for Col } \boldsymbol{A} : \left\{ \begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix} \right\}$$

Any echelon form of $\boldsymbol{A}$ provides (in its nonzero rows) a basis for Row $\boldsymbol{A}$ and also identifies the pivot columns of $\boldsymbol{A}$ for Col $\boldsymbol{A}$. However, for Nul $\boldsymbol{A}$, we need the reduced echelon form. Further row operations on $\boldsymbol{B}$ yield

$$\boldsymbol{A} \sim \boldsymbol{B} \sim \boldsymbol{C} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The equation $\boldsymbol{Ax} = \boldsymbol{0}$ is equivalent to $\boldsymbol{Cx} = \boldsymbol{0}$, that is,

$$\begin{aligned} x_1 + \quad x_3 \quad + x_5 &= 0 \\ x_2 - 2x_3 \quad + 3x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

So $x_1 = -x_3 - x_5$, $x_2 = 2x_3 - 3x_5$, $x_4 = 5x_5$, with x_3 and x_5 free variables. The usual calculations (discussed in lecture 21) show that

$$\text{Basis for Nul } \boldsymbol{A} : \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

Observe that, unlike the bases for Col ***A***, the bases for Row ***A*** and Nul ***A*** have no simple connection with the entries in ***A*** itself.

Note:

1. Although the first three rows of ***B*** in Example (5) are linearly independent, it is wrong to conclude that the first three rows of ***A*** are linearly independent. (In fact, the third row of A is 2 times the first row plus 7 times the second row).
2. Row operations do not preserve the linear dependence relations among the rows of a matrix.

Definition: The **rank** of ***A*** is the dimension of the column space of ***A***.
Since Row ***A*** is the same as Col $\boldsymbol{A}^T$, the dimension of the row space of ***A*** is the rank of $\boldsymbol{A}^T$. The dimension of the null space is sometimes called the **nullity** of ***A***.

Theorem 3: (The Rank Theorem) The dimensions of the column space and the row space of an $m \times n$ matrix ***A*** are equal. This common dimension, the rank of ***A***, also equals the number of pivot positions in ***A*** and satisfies the equation

$$\text{rank } \boldsymbol{A} + \dim \text{Nul } \boldsymbol{A} = n$$

Example 6:

(a) If ***A*** is a 7×9 matrix with a two – dimensional null space, what is the rank of ***A***?
(b). Could a 6×9 matrix have a two – dimensional null space?

Solution:

(a) Since ***A*** has *9* columns, (rank ***A***) $+ 2 = 9$ and hence rank ***A*** $= 7$.
(b) No, If a 6×9 matrix, call it ***B***, had a two – dimensional null space, it would have to have rank *7*, by the Rank Theorem. But the columns of ***B*** are vectors in $\boldsymbol{R}^6$ and so the dimension of Col ***B*** cannot exceed *6*; that is, rank ***B*** cannot exceed *6*.

The next example provides a nice way to visualize the subspaces we have been studying. Later on, we will learn that Row ***A*** and Nul ***A*** have only the zero vector in common and are actually "perpendicular" to each other. The same fact will apply to Row $\boldsymbol{A}^T$ (= Col ***A***) and Nul $\boldsymbol{A}^T$. So the figure in Example (7) creates a good mental image for the general case.

Example 7: Let $A = \begin{bmatrix} 3 & 0 & -1 \\ 3 & 0 & -1 \\ 4 & 0 & 5 \end{bmatrix}$. It is readily checked that Nul $\boldsymbol{A}$ is the x_2 – axis, Row $\boldsymbol{A}$ is the x_1x_3 – plane, Col $\boldsymbol{A}$ is the plane whose equation is $x_1 - x_2 = 0$ and Nul $\boldsymbol{A^T}$ is the set of all multiples of *(1, -1, 0)*. Figure 1 shows Nul $\boldsymbol{A}$ and Row $\boldsymbol{A}$ in the domain of the linear transformation $\boldsymbol{x} \to \boldsymbol{Ax}$; the range of this mapping, Col $\boldsymbol{A}$, is shown in a separate copy of $\boldsymbol{R^3}$, along with Nul $\boldsymbol{A^T}$.

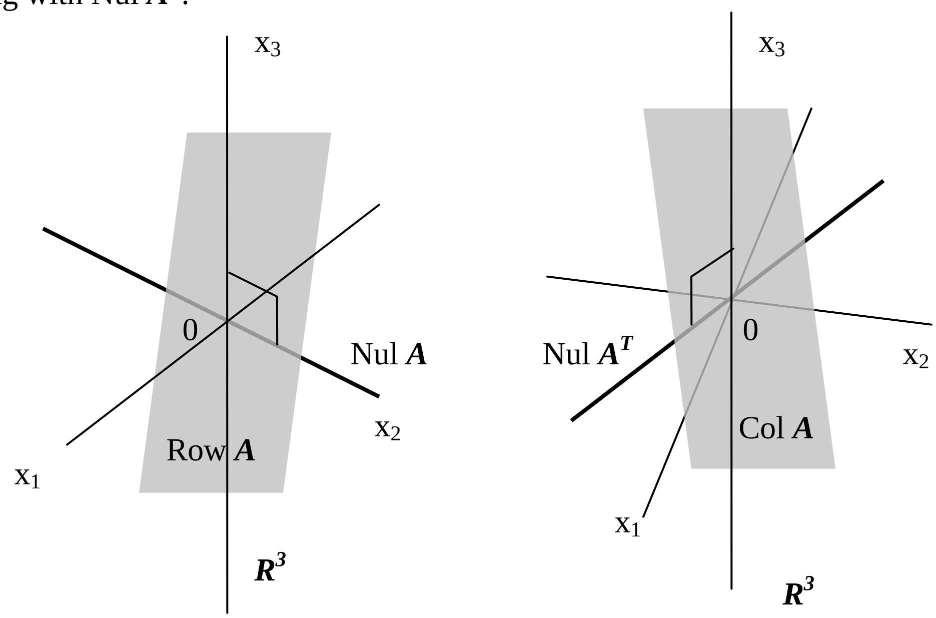


Figure 1 – Subspaces associated with a matrix *A*

Applications to Systems of Equations:

The Rank Theorem is a powerful tool for processing information about systems of linear equations. The next example simulates the way a real-life problem using linear equations might be stated, without explicit mention of linear algebra terms such as matrix, subspace and dimension.

Example 8: A scientist has found two solutions to a homogeneous system of *40* equations in *42* variables. The two solutions are not multiples and all other solutions can be constructed by adding together appropriate multiples of these two solutions. Can the scientist be certain that an associated non-homogeneous system (with the same coefficients) has a solution?

Solution: Yes. Let $\boldsymbol{A}$ be the 40×42 coefficient matrix of the system. The given information implies that the two solutions are linearly independent and span Nul $\boldsymbol{A}$. So dim Nul $\boldsymbol{A} = 2$. By the Rank Theorem, dim Col $\boldsymbol{A} = 42 - 2 = 40$. Since $\boldsymbol{R^{40}}$ is the only subspace of $\boldsymbol{R^{40}}$ whose dimension is *40*, Col $\boldsymbol{A}$ must be all of $\boldsymbol{R^{40}}$. This means that every non-homogeneous equation $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution.

<u>Example 9</u>: Find the rank and nullity of the matrix $A = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$.

Verify that values obtained verify the dimension theorem.

Solution

$$\begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix} (-1)R_1$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \end{bmatrix} \begin{matrix} (-3)R_1 + R_2 \\ (-2)R_1 + R_3 \\ (-4)R_1 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \end{bmatrix} (-1)R_2$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} R_2 + R_3 \\ R_2 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} 2R_2 + R_1$$

The reduced row-echelon form of ***A*** is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (1)$$

The corresponding system of equations will be

$$x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 = 0$$
$$x_2 - 2x_3 - 12x_4 - 16x_5 + 5x_6 = 0$$

or, on solving for the leading variables,

$$\begin{aligned} x_1 &= 4x_3 - 28x_4 + 37x_5 - 13x_6 \\ x_2 &= 2x_3 + 12x_4 + 16x_5 - 5x_6 \end{aligned} \qquad (2)$$

it follows that the general solution of the system is

$$x_1 = 4r + 28s + 37t - 13u$$
$$x_2 = 2r + 12s + 16t - 5u$$
$$x_3 = r$$
$$x_4 = s$$
$$x_5 = t$$
$$x_6 = u$$

or equivalently, $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad (3)$$

The four vectors on the right side of (3) form a basis for the solution space, so nullity $(\boldsymbol{A}) = 4$. The matrix $\boldsymbol{A} = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$ has 6 columns,

so rank$(\boldsymbol{A})$ + nullity$(\boldsymbol{A}) = 2 + 4 = 6 = n$

Example 10: Find the rank and nullity of the matrix; then verify that the values obtained satisfy the dimension theorem $\boldsymbol{A} = \begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$

Solution: Transforming Matrix to the Reduced Row Echelon Form:

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & 6 & 0 & -3 \end{bmatrix} \begin{matrix} (-2)R_1 + R_3 \\ (-3)R_1 + R_4 \\ 2R_1 + R_5 \end{matrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & 6 & 0 & -3 \end{bmatrix} (1/3)R_2$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} (-3)R_2 + R_3 \\ (-3)R_2 + R_4 \\ (-3)R_2 + R_5 \end{matrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} 12R_3 + R_4$$

$$\begin{bmatrix} 1 & -3 & 0 & 2 & 11/6 \\ 0 & 1 & 0 & 0 & -1/6 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} (-2\ ) R_3 + R_2 \\ (-2\ ) R_3 + R_1 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 2 & 4/3 \\ 0 & 1 & 0 & 0 & -1/6 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (3)\, R_2 + R_1 \qquad (1)$$

Since there are three nonzero rows (or equivalently, three leading 1's) the row space and column space are both three dimensional so rank $(\boldsymbol{A}) = 3$.

To find the nullity of $\boldsymbol{A}$, we find the dimension of the solution space of the linear system $\boldsymbol{Ax} = \boldsymbol{0}$. The system can be solved by reducing the augmented matrix to reduced row echelon form. The resulting matrix will be identical to (1), except with an additional last column of zeros, and the corresponding system of equations will be

$$x_1 + 0x_2 + 0x_3 + 2x_4 + \frac{4}{3}x_5 = 0$$

$$0x_1 + x_2 + 0x_3 + 0x_4 - \frac{1}{6}x_5 = 0$$

$$0x_1 + 0x_2 + x_3 + 0x_4 - \frac{5}{12}x_5 = 0$$

The system has infinitely many solutions:

$x_1 = -2\,x_4 + (-4/3)\,x_5$ $\qquad$ $x_2 = (1/6)\,x_5$

$x_3 = (5/12)\,x_5$ $\qquad$ $x_4 = s$

$x_5 = t$

The solution can be written in the vector form:

$\boldsymbol{c_4} = (-2, 0, 0, 1, 0)$ $\qquad$ $\boldsymbol{c_5} = (-4/3, 1/6, 5/12, 0, 1)$

Therefore the **null space** has a basis formed by the set

$$\{(-2, 0, 0, 1, 0), (-4/3, 1/6, 5/12, 0, 1)\}$$

The nullity of the matrix is 2. Now Rank $(\boldsymbol{A})$ + nullity $(\boldsymbol{A}) = 3 + 2 = 5 = n$

Theorem 4: If $\boldsymbol{A}$ is an m x n, matrix, then
(a) rank ($\boldsymbol{A}$) = the number of leading variables in the solution of $\boldsymbol{Ax} = \boldsymbol{0}$
(b) nullity ($\boldsymbol{A}$) = the number of parameters in the general solution of $\boldsymbol{Ax} = \boldsymbol{0}$

Example 11: Find the number of parameters in the solution set of $\boldsymbol{Ax} = \boldsymbol{0}$ if $\boldsymbol{A}$ is a 5×7 matrix of rank 3.
Solution: nullity $(\boldsymbol{A}) = n -$ rank $(\boldsymbol{A}) = 7-3 = 4$
Thus, there are four parameters.

Example (not in handouts) Find the number of parameters in the solution set of $\boldsymbol{Ax} = \boldsymbol{0}$ if $\boldsymbol{A}$ is a 4×4 matrix of rank 0.
Solution nullity $(\boldsymbol{A}) = n -$ rank $(\boldsymbol{A}) = 4-0 = 4$
Thus, there are four parameters.

Theorem 5: If $\boldsymbol{A}$ is any matrix, then rank $(\boldsymbol{A})$ = rank $(\boldsymbol{A}^T)$

Four fundamental matrix spaces:

If we consider a matrix $\boldsymbol{A}$ and its transpose $\boldsymbol{A}^T$ together, then there are six vectors spaces of interest:

Row space of $\boldsymbol{A}$	row space of $\boldsymbol{A}^T$
Column space of $\boldsymbol{A}$	column space of $\boldsymbol{A}^T$
Null space of $\boldsymbol{A}$	null space of $\boldsymbol{A}^T$

However, transposing a matrix converts row vectors into column vectors and column vectors into row vectors, so that, except for a difference in notation, the row space of $\boldsymbol{A}^T$ is the same as the column space of $\boldsymbol{A}$ and the column space of $\boldsymbol{A}^T$ is the same as row space of of $\boldsymbol{A}$.
This leaves four vector spaces of interest:

Row space of $\boldsymbol{A}$	column space of $\boldsymbol{A}$
Null space of $\boldsymbol{A}$	null space of $\boldsymbol{A}^T$

These are known as the **fundamental matrix spaces** associated with $\boldsymbol{A}$, if $\boldsymbol{A}$ is an m x n matrix, then the row space of $\boldsymbol{A}$ and null space of $\boldsymbol{A}$ are subspaces of $\boldsymbol{R}^n$ and the column space of $\boldsymbol{A}$ and the null space of $\boldsymbol{A}^T$ are subspaces of $\boldsymbol{R}^m$.

Suppose now that $\boldsymbol{A}$ is an m x n matrix of rank r, it follows from theorem (5) that $\boldsymbol{A}^T$ is an n x m matrix of rank r . Applying theorem (3) on $\boldsymbol{A}$ and $\boldsymbol{A}^T$ yields

$$\text{Nullity } (\boldsymbol{A}) = n-r, \text{ nullity } (\boldsymbol{A}^T) = m-r$$

From which we deduce the following table relating the dimensions of the four fundamental spaces of an m x n matrix $\boldsymbol{A}$ of rank r.

Fundamental space	**Dimension**
Row space of ***A***	*r*
Column space of ***A***	*r*
Null space of ***A***	*n-r*
Null space of $\boldsymbol{A}^T$	*m-r*

Example 12: If ***A*** is a *7* x *4* matrix, then the rank of ***A*** is at most *4* and, consequently, the seven row vectors must be linearly dependent. If ***A*** is a *4* x *7* matrix, then again the rank of ***A*** is at most *4* and, consequently, the seven column vectors must be linearly dependent.

Rank and the Invertible Matrix Theorem: The various vector space concepts associated with a matrix provide several more statements for the Invertible Matrix Theorem. We list only the new statements here, but we reference them so they follow the statements in the original Invertible Matrix Theorem in lecture 13.

Theorem 6: The Invertible Matrix Theorem (Continued)
Let ***A*** be an *n* x *n* matrix. Then the following statements are each equivalent to the statement that ***A*** is an invertible matrix.

m. The columns of ***A*** form a basis of $\boldsymbol{R}^n$.
n. Col ***A*** = $\boldsymbol{R}^n$.
o. dim Col ***A*** = *n*
p. rank ***A*** = *n*
q. Nul ***A*** = $\{\boldsymbol{0}\}$
r. dim Nul ***A*** = *0*

Proof: Statement (m) is logically equivalent to statements (e) and (h) regarding linear independence and spanning. The other statements above are linked into the theorem by the following chain of almost trivial implications:

$$(g) \Rightarrow (n) \Rightarrow (o) \Rightarrow (p) \Rightarrow (r) \Rightarrow (q) \Rightarrow (d)$$

Only the implication (p) $\Rightarrow$ (r) bears comment. It follows from the Rank Theorem because ***A*** is $n \times n$. Statements (d) and (g) are already known to be equivalent, so the chain is a circle of implications.

We have refrained from adding to the Invertible Matrix Theorem obvious statements about the row space of ***A***, because the row space is the column space of $\boldsymbol{A}^T$. Recall from (1) of the Invertible Matrix Theorem that ***A*** is invertible if and only if $\boldsymbol{A}^T$ is invertible. Hence every statement in the Invertible Matrix Theorem can also be stated for $\boldsymbol{A}^T$.

Numerical Note:

Many algorithms discussed in these lectures are useful for understanding concepts and making simple computations by hand. However, the algorithms are often unsuitable for large-scale problems in real life.

Rank determination is a good example. It would seem easy to reduce a matrix to echelon form and count the pivots. But unless exact arithmetic is performed on a matrix whose entries are specified exactly, row operations can change the apparent rank of a matrix.

For instance, if the value of x in the matrix $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ is not stored exactly as *7* in a computer, then the rank may be *1* or *2*, depending on whether the computer treats $x - 7$ as zero.
In practical applications, the effective rank of a matrix ***A*** is often determined from the singular value decomposition of ***A***.

Example 13: The matrices below are row equivalent

$$A = \begin{bmatrix} 2 & -1 & 1 & -6 & 8 \\ 1 & -2 & -4 & 3 & -2 \\ -7 & 8 & 10 & 3 & -10 \\ 4 & -5 & -7 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -2 & -4 & 3 & -2 \\ 0 & 3 & 9 & -12 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. Find rank ***A*** and dim Nul ***A***.
2. Find bases for Col ***A*** and Row ***A***.
3. What is the next step to perform if one wants to find a basis for Nul ***A***?
4. How many pivot columns are in a row echelon form of A^T?

Solution:

1. ***A*** has two pivot columns, so rank ***A*** = *2*. Since ***A*** has *5* columns altogether, dim Nul ***A*** = *5 – 2 = 3*.
2. The pivot columns of ***A*** are the first two columns. So a basis for Col ***A*** is

$$\{a_1, a_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 8 \\ -5 \end{bmatrix} \right\}$$

   The nonzero rows of ***B*** form a basis for Row ***A***, namely *{(1, –2, –4, 3, –2), (0, 3, 9, –12, 12)}*. In this particular example, it happens that any two rows of ***A*** form a basis for the row space, because the row space is two-dimensional and none of the rows of ***A*** is a multiple of another row. In general, the nonzero rows of an echelon form of ***A*** should be used as a basis for Row ***A***, not the rows of ***A*** itself.
3. For Nul ***A***, the next step is to perform row operations on ***B*** to obtain the reduced echelon form of ***A***.
4. Rank A^T = rank ***A***, by the Rank Theorem, because Col A^T = Row ***A***. So A^T has two pivot positions.

Exercises:

In exercises 1 to 4, assume that the matrix ***A*** is row equivalent to ***B***. Without calculations, list rank ***A*** and dim Nul ***A***. Then find bases for Col ***A***, Row ***A***, and Nul ***A***.

1. $A = \begin{bmatrix} 1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & -3 & 4 & -1 & 9 \\ -2 & 6 & -6 & -1 & -10 \\ -3 & 9 & -6 & -6 & -3 \\ 3 & -9 & 4 & 9 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & -3 & 0 & 5 & -7 \\ 0 & 0 & 2 & -3 & 8 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ -2 & 3 & -3 & -3 & -4 \\ 4 & -6 & 9 & 5 & 9 \\ -2 & 3 & 3 & -4 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ 0 & 0 & 3 & -1 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 1 & 2 & -4 & 10 & 13 & -12 \\ 1 & -1 & -1 & 1 & 1 & -3 \\ 1 & -3 & 1 & -5 & -7 & 3 \\ 1 & -2 & 0 & 0 & -5 & -4 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 0 & 1 & -1 & 3 & 4 & -3 \\ 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

5. If a 3 x 8 matrix A has rank 3, find dim Nul A, dim Row A, and rank A^T.

6. If a 6 x 3 matrix A has rank 3, find dim Nul A, dim Row A, and rank A^T.

7. Suppose that a 4 x 7 matrix A has four pivot columns. Is Col A = R^4? Is Nul A = R^3? Explain your answers.

8. Suppose that a 5 x 6 matrix A has four pivot columns. What is dim Nul A? Is Col A = R^4? Why or why not?

9. If the null space of a 5 x 6 matrix A is 4-dimensional, what is the dimension of the column space of A?

10. If the null space of a 7 x 6 matrix A is 5-dimensional, what is the dimension of the column space of A?

11. If the null space of an 8 x 5 matrix A is 2-dimensional, what is the dimension of the row space of A?

12. If the null space of a 5 x 6 matrix A is 4-dimensional, what is the dimension of the row space of A?

13. If A is a 7 x 5 matrix, what is the largest possible rank of A? If A is a 5 x 7 matrix, what is the largest possible rank of A? Explain your answers.

14. If ***A*** is a 4 x 3 matrix, what is the largest possible dimension of the row space of ***A***? If ***A*** is a 3 x 4 matrix, what is the largest possible dimension of the row space of ***A***? Explain.

15. If ***A*** is a 6 x 8 matrix, what is the smallest possible dimension of Nul ***A***?

16. If ***A*** is a 6 x 4 matrix, what is the smallest possible dimension of Nul ***A***?