Advanced Linear Algebra

MTH 706

Virtual University of Pakistan

Knowledge beyond the boundaries

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 introduce 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 a 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, to explain the underline theory, the computational techniques and then try to apply them on real life problems. Major 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 should 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 should also browse on the Internet; there is plenty of supporting 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 go through 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 are important from the conceptual 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 will follow it regularly and completes 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

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

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

$$(A+B)C=AC+BC$$

Remarks

It is important to note that some rules arithmetic for real numbers$\mathbb{R}$ do not carry over the matrix arithmetic.

For example, $\forall a,b,c$ and $d\in\mathbb{R}$

i) if $ab=cd$ and $a\neq 0$, then $b=c$ (Law of Cancellation)

ii) if $ab=0$, then least one of the factors *a* or *b* (or both) are zero.

However the following examples shows that the corresponding results are not true in case of matrices.

Example

Let $A=\begin{bmatrix}0&1\\0&2\end{bmatrix}$, $B=\begin{bmatrix}1&1\\3&4\end{bmatrix}$, $C=\begin{bmatrix}2&5\\3&4\end{bmatrix}$ and $D=\begin{bmatrix}1&7\\0&0\end{bmatrix}$, then one can easily check that

$AB=AC=\begin{bmatrix}3&4\\6&8\end{bmatrix}$. But $B\neq C$.

Similarly neither A nor B are zero matrices but $AD=\begin{bmatrix}0&0\\0&0\end{bmatrix}$

But if D is diagonal say $D=\begin{bmatrix}1&0\\0&7\end{bmatrix}$, then $AD\neq DA$.

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 $|A|$. There is a special way to find the determinant of a given matrix.

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 $$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \text{ then } A^{tr} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

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:`
- The transpose of a matrix times a scalar (k) is equal to the constant times the transpose of the matrix: $(k\underline{A})^T = k\underline{A}^T$
- The transpose of the sum of two matrices is equivalent to the sum of their transposes: $(\underline{A}+\underline{B})^T = \underline{A}^T + \underline{B}^T$
- The transpose of the product of two matrices is equivalent to the product of their transposes in reversed order: $(\underline{A}\underline{B})^T = \underline{B}^T\underline{A}^T$
- The same is true for the product of multiple matrices: $(\underline{A}\underline{B}\underline{C})^T = \underline{C}^T\underline{B}^T\underline{A}^T$

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 $X^T = \begin{bmatrix} 5 & 0 & 3 \end{bmatrix}$

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 $\mathrm{B} = \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix}$ 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 $A^{-1} = \dfrac{1}{\det(A)} (C_{ij})^{tr}$

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 $A^{-1} = \dfrac{1}{\det A}\begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix}$.

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 $A^{-1} = \dfrac{1}{2}\begin{pmatrix} 10 & -4 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix}$

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

$$C_{11} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1, \quad C_{12} = -\begin{vmatrix} -2 & 1 \\ 3 & 1 \end{vmatrix} = 5, \quad C_{13} = \begin{vmatrix} -2 & 1 \\ 3 & 0 \end{vmatrix} = -3$$

$$C_{21} = -\begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} = -2, \quad C_{22} = \begin{vmatrix} 2 & 0 \\ 3 & 1 \end{vmatrix} = 2, \quad C_{23} = -\begin{vmatrix} 2 & 2 \\ 3 & 0 \end{vmatrix} = 6$$

$$C_{31} = \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} = 2, \quad C_{32} = -\begin{vmatrix} 2 & 0 \\ -2 & 1 \end{vmatrix} = -2, \quad C_{33} = \begin{vmatrix} 2 & 2 \\ -2 & 1 \end{vmatrix} = 6$$

Hence $$\mathrm{A}^{-1} = \frac{1}{12}\begin{pmatrix} 1 & -2 & 2 \\ 5 & 2 & -2 \\ -3 & 6 & 6 \end{pmatrix} = \begin{pmatrix} 1/12 & -1/6 & 1/6 \\ 5/12 & 1/6 & -1/6 \\ -1/4 & 1/2 & 1/2 \end{pmatrix}$$

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 in 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

$$X'(t)=\begin{pmatrix}\frac{d}{dt}(\sin 2t)\\ \frac{d}{dt}(e^{3t})\\ \frac{d}{dt}(8t-1)\end{pmatrix}=\begin{pmatrix}2\cos 2t\\ 3e^{3t}\\ 8\end{pmatrix} \quad\text{and}\quad \int_0^t X(s)ds=\begin{pmatrix}\int_0^t \sin 2s ds\\ \int_0^t e^{3s}ds\\ \int_0^t 8s-1ds\end{pmatrix}=\begin{pmatrix}-1/2\cos 2t+1/2\\ 1/3e^{3t}-1/3\\ 4t^2-t\end{pmatrix}$$

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

We know that the equation of a straight line is written as $y = mx + c$, where m is the slope of line(Tan of the angle of line with x-axis) and c is the y-intercept(the distance at which the straight line meets y-axis from origin).
Thus 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

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \text{(hyper plane in } \mathbb{R}^n\text{) --------(1)}$$

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

$$2x_1 + 3x_2 + 2 = x_3 \quad \textit{and} \quad x_2 = 2\left(\sqrt{5} + x_1\right) + 2x_3 \text{ are both linear}$$

(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,

$$\begin{aligned} 4x_1 - x_2 + 3x_3 &= -1 \\ 3x_1 + x_2 + 9x_3 &= -4 \end{aligned}$$

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

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \qquad (3)$$

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 $s_1, s_2, \cdots, s_n$ such 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 such solutions $\{s_1, s_2, \cdots, s_n\}$ 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

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

The graphs of these equations are straight lines in the *xy-plane*, so a solution (*x, y*) of this system is infact a point of intersection of these lines.

Note that there are three possibilities for a pair of straight lines in xy-plane:

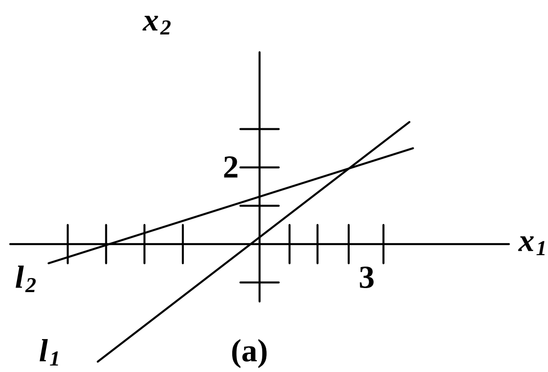
1. The lines may be parallel and distinct, in which case there is no intersection and consequently ***no solution***.
2. The lines may intersect at only one point, in which case the system has exactly ***one solution***.
3. The lines may coincide, in which case there are infinitely many points of intersection (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.

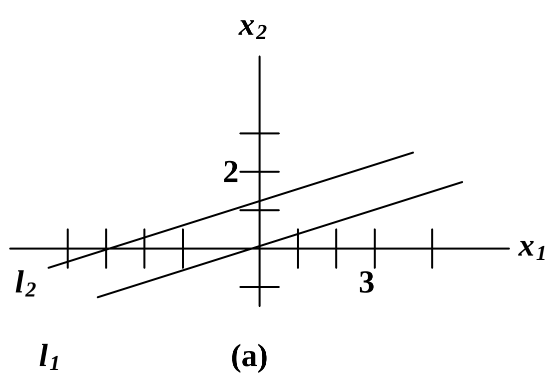
Example consider the system of linear equations in two variables $x_1 - 2x_2 = -1, \; -x_1 + 3x_2 = 3$
Solve the equation simultaneously:
Adding both equations we get $x_2 = 2$, Put $x_2 = 2$ in any one of the above equation we get $x_1 = 3$. So the solution is the single point **(3, 2)**. See the graph of this linear system

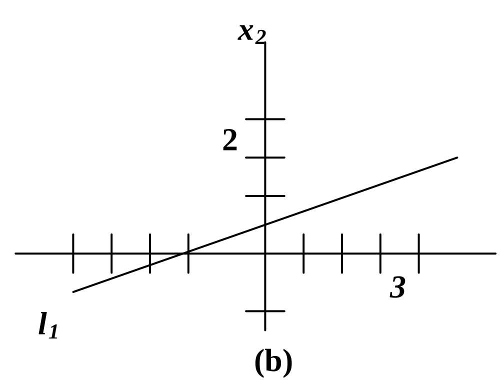


This system has exactly one solution

See the graphs to the following linear systems:

$$(a) \quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned} \qquad\qquad (b) \quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 1 \end{aligned}$$





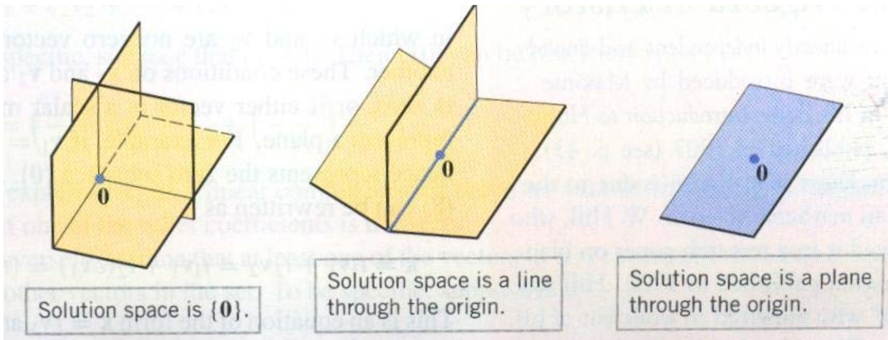
(a) No solution. (b) Infinitely many solutions.

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

$$\begin{aligned} x-y&=1\\ 2x+y&=6 \end{aligned}$$

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

$$\begin{aligned} x+\ \ y&=4\\ 3x+3y&=6 \end{aligned}$$

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
$$\begin{aligned} x_1 - 2x_2 + \quad x_3 &= 0 \\ 2x_2 - 8x_3 &= \ \ 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

Solution We perform the elimination procedure with and without matrix notation, and place the results side by side for comparison:

$$\begin{aligned} x_1 - 2x_2 + \quad x_3 &= 0 \\ 2x_2 - 8x_3 &= \ \ 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

To eliminate the x_1 term from third equation add 4 times equation 1 to equation 3,

$$\begin{array}{r} 4x_1 - 8x_2 + 4x_3 = 0 \\ \underline{-4x_1 + 5x_2 + 9x_3 = -9} \\ -3x_2 + 13x_3 = -9 \end{array}$$

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

$$\begin{aligned} x_1 - 2x_2 + \ x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

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

$$\begin{gathered} (29) - 2(16) + (3) = 29 - 32 + 3 = 0 \\ 2(16) - 8(3) = 32 - 24 = 8 \\ -4(29) + 5(16) + 9(3) = -116 + 80 + 27 = -9 \end{gathered}$$

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

$$\begin{aligned} x_1 - 2x_2 + \quad x_3 &= 0 \\ 2x_2 - 8x_3 &= \ 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

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.
$$\begin{aligned} x_1 + 4x_2 - 2x_3 + 8x_4 &= 12 \\ x_2 - 7x_3 + 2x_4 &= -4 \\ 5x_3 - x_4 &= 7 \\ x_3 + 3x_4 &= -5 \end{aligned}$$

b.
$$\begin{aligned} x_1 - 3x_2 + 5x_3 - 2x_4 &= 0 \\ x_2 + 8x_3 &= -4 \\ 2x_3 &= 7 \\ x_4 &= 1 \end{aligned}$$

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. While applying the algorithm 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 obtained 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.

$$\begin{array}{l} \text{Pivot} \\ \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \\ \text{Pivot Column} \end{array}$$

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.

$$\begin{array}{l} \text{Pivot} \\ \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \quad \boxed{\begin{array}{c} R_1+R_2 \\ 2R_1+R_3 \end{array}} \\ \text{Next pivot column} \end{array}$$

Add –5/2 times row 2 to row 3, and add 3/2 times row 2 to row 4.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \qquad \begin{array}{l} -\frac{5}{2}R_2 + R_3 \\ \frac{3}{2}R_2 + R_4 \end{array}$$

Interchange rows 3 and 4, we can produce a leading entry in column 4.

Pivot

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \textit{General} \quad \textit{form} \quad \begin{bmatrix} \circ & * & * & * & * \\ 0 & \circ & * & * & * \\ 0 & 0 & 0 & \circ & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

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

<u>Pivot element</u>

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.

<u>Example 3</u> 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.

$$\begin{array}{l} \begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \\ \text{Pivot column} \end{array}$$

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.)

$$\begin{array}{l} \text{Pivot} \\ \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \end{array}$$

STEP 3 Use row replacement operations to create zeros in all positions below the pivot

Subtract Row 1 from Row 2. i.e. $R_2 - R_1$

$$\begin{array}{l} \text{Pivot} \\ \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \end{array}$$

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 change 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 2nd Row by 2

$$\begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad \frac{1}{2}R_2 \;,\; \frac{1}{3}R_1$$

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 is 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

$$\begin{aligned} x_1 \quad -5x_3 &= 1 \\ x_2 + x_3 &= 4 \\ 0 &= 0 \quad \text{which means } x_3 \text{ is free} \end{aligned} \qquad (1)$$

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.)

$$\begin{aligned} x_1 &= 1 + 5x_3 \\ x_2 &= 4 - x_3 \\ x_3 & \ \ is \quad free \end{aligned} \qquad (2)$$

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 05

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.

<u>Null Space of a Matrix</u>:

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.

<u>NULL SPACE</u>

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

$$Nul(A) = \{x / \forall x \in \square , Ax = 0\}$$

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

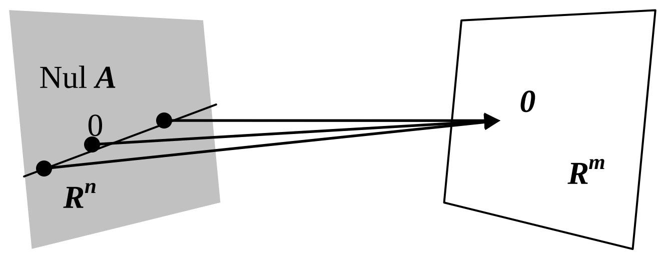


Figure 1

Example 1: Let $\boldsymbol{A} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$ and let $\boldsymbol{u} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$. Determine if $\boldsymbol{u} \in \text{Nul}\,\boldsymbol{A}$.

Solution: To test if ***u*** satisfies ***Au = 0***, simply compute

$$\boldsymbol{Au} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 5-9+4 \\ -25+27-2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$ 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:

$$\begin{pmatrix} 4 & 0 \\ -8 & 20 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 4x_1 & +\ 0x_2 \\ -8x_1 & +20x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow 4x_1 + 0x_2 = 0 \qquad \Rightarrow x_1 = 0$$

$$and \qquad \Rightarrow -8x_1 + 20x_2 = 0 \qquad \Rightarrow x_2 = 0$$

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 ***Ax=0***. Therefore, we can say that it does not change the null space of ***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.

<u>Theorem 2</u>: 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.* $Ax = 0$*).*

<u>Proof</u>: 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,
$\mathbf{A}\ (c\mathbf{u}) = c\ (\mathbf{Au}) = c\ (\mathbf{0}) = \boldsymbol{0}$
which shows that $c\boldsymbol{u}$ is in Nul $\boldsymbol{A}$. Thus Nul $\boldsymbol{A}$ is a subspace of $\boldsymbol{R^n}$.

<u>Example 2</u>: 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 $\boldsymbol{R^4}$.

<u>Solution</u>: Since $a - 2b + 5c = d$
$c - a = b$

By rearranging the equations, we get

$$\begin{aligned} a - 2b + 5c - d &= 0 \\ -a - b + c \quad\quad &= 0 \end{aligned}$$

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 $\boldsymbol{Ax} = \boldsymbol{0}$, 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 $\boldsymbol{Ax} = \boldsymbol{0}$ in terms of free variables.
After transforming the augmented matrix $[\boldsymbol{A} \quad \boldsymbol{0}]$ 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 $Ax = 0$.

<u>Example 4</u>: Find a spanning set for the null space of $\boldsymbol{A} = \begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$.

<u>Solution</u>: 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

$$\begin{aligned} 1\,x_1 + \qquad\qquad 2\,x_4 + (4/3)\,x_5 &= 0 \\ 1\,x_2 + \qquad\quad (-1/6)\,x_5 &= 0 \\ 1\,x_3 + \quad (-5/12)\,x_5 &= 0\,. \\ 0 &= 0 \\ 0 &= 0 \end{aligned}$$

No equation of this system has a form zero = nonzero; Therefore, the system is **consistent**. The system has **infinitely many solutions**:

$$x_1 = -2\,x_4 + (-4/3)\,x_5 \qquad x_2 = +(1/6)\,x_5 \quad x_3 = +(5/12)\,x_5$$
$$x_4 = \text{arbitrary} \qquad\qquad x_5 = \text{arbitrary}$$

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*** = [$\boldsymbol{a_1}$ … $\boldsymbol{a_n}$], then

Col ***A*** = *Span* {$\boldsymbol{a_1}$,… , $\boldsymbol{a_n}$ }

Since Span {$\boldsymbol{a_1}$, …, $\boldsymbol{a_n}$ } 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.

<u>Example</u>, 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***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***A*** = { a·(1, 2, 3, 4) + b·(1, 1, 1, 1) + c·(3, 4, 5, 6) }. In general, **the column space** Col***A*** **contains all the linear combinations of columns of A**.

The next theorem follows from the definition of Col ***A*** and the fact that the columns of ***A*** are in $\boldsymbol{R^m}$.

<u>Theorem 3</u>: The column space of an $m \times n$ matrix *A* is a subspace of R^m.

Note that a typical vector in Col ***A*** can be written as ***Ax*** for some ***x*** because the notation ***Ax*** stands for a linear combination of the columns of ***A***. That is,

Col ***A*** = {***b***: ***b*** = ***Ax*** for some ***x*** in $\boldsymbol{R^n}$}

The notation ***Ax*** for vectors in Col ***A*** also shows that Col ***A*** is the range of the linear transformation $\boldsymbol{x} \rightarrow \boldsymbol{Ax}$.

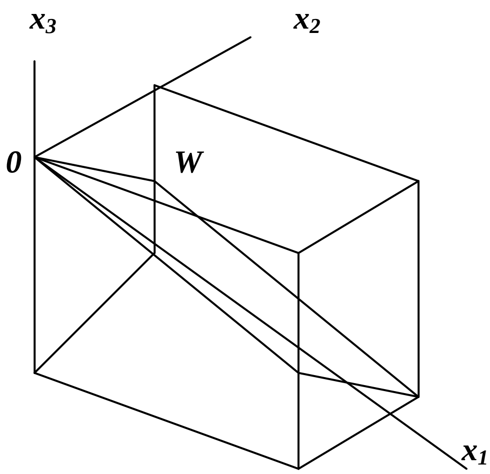
<u>Example 6</u>: Find a matrix ***A*** such that ***W*** = Col ***A***. $\boldsymbol{W} = \left\{ \begin{bmatrix} 6a - b \\ a+b \\ -7a \end{bmatrix} : a,b \text{ in } R \right\}$

<u>Solution</u>: First, write ***W*** as a set of linear combinations.

$$\boldsymbol{W} = \left\{ a\begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a,b \text{ in } R \right\} = \text{Span}\left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Second, use the vectors in the spanning set as the columns of ***A***. Let $\boldsymbol{A} = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ -7 & 0 \end{bmatrix}$.

Then ***W*** = Col ***A***, as desired.



We know that the columns of $\boldsymbol{A}$ span $\boldsymbol{R^m}$ if and only if the equation $\boldsymbol{Ax = b}$ has a solution for each $\boldsymbol{b}$. 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 $\boldsymbol{b}$ in the column space of $\boldsymbol{A}$.

Example 6: A vector $\boldsymbol{b}$ in the column space of $\boldsymbol{A}$. 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 $\boldsymbol{b}$ is in the column space of $\boldsymbol{A}$, and express $\boldsymbol{b}$ as a linear combination of the column vectors of $\boldsymbol{A}$.

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

$$\begin{aligned} x_1 \quad + x_3 &= 0 \\ x_2 + x_3 &= 0 \\ 0 &= 1 \end{aligned}$$

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. $A=\begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 2 & 0 & 1 \\ 1 & 1 & 1 & -3 \\ 0 & 2 & 2 & 1 \end{pmatrix}$; $b=\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$

<u>Theorem 5</u>: If x_0 denotes any single solution of a consistent linear system $Ax=b$ and if $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, ..., \boldsymbol{v_k}$ form the solution space of the homogeneous system $Ax=0$, then every solution of $Ax=b$ can be expressed in the form $\boldsymbol{x} = \boldsymbol{x_0} + c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + ... + c_k\boldsymbol{v_k}$ and, conversely, for all choices of scalars $c_1, c_2, c_3, ..., c_k$, 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 $\boldsymbol{Ax}=\boldsymbol{0}$.

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + \quad\quad 2x_5 \quad\quad &= 0 \\ 2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1 \\ 5x_3 + 10x_4 \quad\quad + 15x_6 &= 5 \\ 2x_1 + 6x_2 \quad\quad + 8x_4 + 4x_5 + 18x_6 &= 6 \end{aligned}$$

<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

$$\begin{aligned} 1x_1 + 3x_2 + \quad 4x_4 + 2x_5 &= 0 \\ 1x_3 + 2x_4 &= 0 \\ 1x_6 &= (1/3) \\ 0 &= 0 \end{aligned}$$

No equation of this system has a form zero = nonzero; Therefore, the system is **consistent**. The system has **infinitely many solutions**:

$$\begin{array}{lll} x_1 = -3x_2 - 4x_4 - 2x_5 & x_2 = r & x_3 = -2x_4 \\ x_4 = s & x_5 = t & x_6 = 1/3 \end{array}$$

$$\begin{array}{lll} x_1 = -3r - 4s - 2t & x_2 = r & x_3 = -2s \\ x_4 = s & x_5 = t & x_6 = \dfrac{1}{3} \end{array}$$

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 $\boldsymbol{A} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$, find a nonzero vector in Col $\boldsymbol{A}$ and a nonzero vector in Nul $\boldsymbol{A}$

Solution: It is easy to find a vector in Col $\boldsymbol{A}$. Any column of $\boldsymbol{A}$ will do, say, $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$. To find a nonzero vector in Nul $\boldsymbol{A}$, we have to do some work. We row reduce the augmented matrix $[\boldsymbol{A} \quad \boldsymbol{0}]$ to obtain $[\boldsymbol{A} \quad \boldsymbol{0}] \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$. Thus if $\boldsymbol{x}$ satisfies $\boldsymbol{Ax} = \boldsymbol{0}$, 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 $\boldsymbol{A}$, namely, $\boldsymbol{x} = (-9, 5, 1, 0)$.

Example 10: With $\boldsymbol{A} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$, let $\boldsymbol{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

(a) Determine if $\boldsymbol{u}$ is in Nul $\boldsymbol{A}$. Could $\boldsymbol{u}$ be in Col $\boldsymbol{A}$?

(b) Determine if $\boldsymbol{v}$ is in Col $\boldsymbol{A}$. Could $\boldsymbol{v}$ be in Nul $\boldsymbol{A}$?

<u>Solution</u>: (a) An explicit description of Nul $\boldsymbol{A}$ is not needed here. Simply compute the product

$$\boldsymbol{Au} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Obviously, $\boldsymbol{u}$ is not a solution of $\boldsymbol{Ax} = \boldsymbol{0}$, so $\boldsymbol{u}$ is not in Nul $\boldsymbol{A}$.

Also, with four entries, $\boldsymbol{u}$ could not possibly be in Col $\boldsymbol{A}$, since Col $\boldsymbol{A}$ is a subspace of $\boldsymbol{R^3}$.

(b) Reduce $[\boldsymbol{A} \quad \boldsymbol{v}]$ to an echelon form:

$$[\boldsymbol{A} \quad \boldsymbol{v}] = \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ -2 & -5 & 7 & 3 & -1 \\ 3 & 7 & -8 & 6 & 3 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ 0 & 1 & -5 & -4 & 2 \\ 0 & 0 & 0 & 17 & 1 \end{bmatrix}$$

At this point, it is clear that the equation $\boldsymbol{Ax} = \boldsymbol{v}$ is consistent, so $\boldsymbol{v}$ is in Col $\boldsymbol{A}$. With only three entries, $\boldsymbol{v}$ could not possibly be in Nul $\boldsymbol{A}$, since Nul $\boldsymbol{A}$ is a subspace of $\boldsymbol{R^4}$.

The following table summarizes what we have learned about Nul $\boldsymbol{A}$ and Col $\boldsymbol{A}$.

1. Nul A is a subspace of R^n.	**1. Col A is a subspace of R^m.**
2. Nul A is implicitly defined; i.e. we are given only a condition ($Ax = 0$) that vectors in Nul A must satisfy.	**2. Col A is explicitly defined; that is, we are told how to build vectors in Col A.**
3. It takes time to find vectors in Nul A. Row operations on $[A \quad 0]$ are required.	**3. It is easy to find vectors in Col A The columns of A are displayed; others are formed from them.**
4. There is no obvious relation between Nul A and the entries in A.	**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 Nul A has the property that $Av = 0$.	**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 is easy to tell if v is in Nul A. Just compute Av.	**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. Nul $A = \{0\}$ if and only if the equation $Ax = 0$ has only the trivial solution.	**7. Col A.= R^m if and only if the equation $Ax = b$ has a solution for every b in R^m.**
8. Nul $A = \{0\}$ if and only if the linear transformation $x \to Ax$ is one-to-one.	**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 $\boldsymbol{R^n}$ 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** $\boldsymbol{T}$ from a vector space $\boldsymbol{V}$ into a vector space $\boldsymbol{W}$ is a rule that assigns to each vector $\boldsymbol{x}$ in $\boldsymbol{V}$ a unique vector $\boldsymbol{T}(\boldsymbol{x})$ in $\boldsymbol{W}$, such that

(i) $\boldsymbol{T}(\boldsymbol{u}+\boldsymbol{v}) = \boldsymbol{T}(\boldsymbol{u}) + \boldsymbol{T}(\boldsymbol{v})$ for all $\boldsymbol{u}, \boldsymbol{v}$ in $\boldsymbol{V}$, and
(ii) $\boldsymbol{T}(c\boldsymbol{u}) = c\,\boldsymbol{T}(\boldsymbol{u})$ for all $\boldsymbol{u}$ in $\boldsymbol{V}$ and all scalars c.

The **kernel** (or **null space**) of such a $\boldsymbol{T}$ is the set of all $\boldsymbol{u}$ in $\boldsymbol{V}$ such that $\boldsymbol{T}(\boldsymbol{u}) = \boldsymbol{0}$ (the zero vector in $\boldsymbol{W}$). The **range** of $\boldsymbol{T}$ is the set of all vectors in $\boldsymbol{W}$ of the form $\boldsymbol{T}(\boldsymbol{x})$ for some $\boldsymbol{x}$ in $\boldsymbol{V}$. If $\boldsymbol{T}$ happens to arise as a matrix transformation, say, $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ for some matrix $\boldsymbol{A}$ – then the kernel and the range of $\boldsymbol{T}$ are just the null space and the column space of $\boldsymbol{A}$, as defined earlier. So if $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, col $\boldsymbol{A}$ = range of $\boldsymbol{T}$.

Definition: If $T: V \to W$ is a linear transformation, then the set of vectors in V that $\boldsymbol{T}$ maps into **0** is called the **kernel** of $\boldsymbol{T}$; it is denoted by ***ker(T)***. The set of all vectors in $\boldsymbol{W}$ that are images under $\boldsymbol{T}$ of at least one vector in $\boldsymbol{V}$ is called the **range** of $\boldsymbol{T}$; it is denoted by ***R(T)***.

Example: 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 $\boldsymbol{A}$ and the range of T_A is the column space of $\boldsymbol{A}$.

Remarks: The kernel of $\boldsymbol{T}$ is a subspace of $\boldsymbol{V}$ and the range of $\boldsymbol{T}$ is a subspace of $\boldsymbol{W}$.

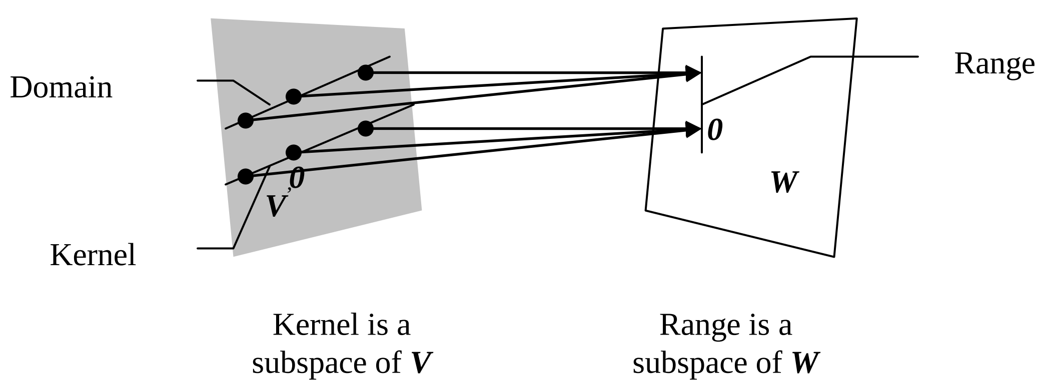


Figure 2 Subspaces associated with a linear transformation.

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.

Example 11: Let $\boldsymbol{V}$ 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 $\boldsymbol{W}$ 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 $\boldsymbol{V}$ 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, $\boldsymbol{D}$ is a linear transformation. It can be shown that the kernel of $\boldsymbol{D}$ is the set of constant functions of [*a, b*] and the range of $\boldsymbol{D}$ is the set $\boldsymbol{W}$ of all continuous functions on [*a, b*].

Example 12: 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

$$\begin{bmatrix}3b+c\\b\\c\end{bmatrix}=b\underbrace{\begin{bmatrix}3\\1\\0\end{bmatrix}}_{\uparrow\ \boldsymbol{v}_1}+c\underbrace{\begin{bmatrix}1\\0\\1\end{bmatrix}}_{\uparrow\ \boldsymbol{v}_2}$$

This calculation shows that W = 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.

<u>Example 14</u>: 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 $\boldsymbol{Ax} = \boldsymbol{v} + \boldsymbol{w}$?

<u>Solution</u>: 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 $\boldsymbol{V}$ and $\boldsymbol{W}$ be any two vector spaces. The mapping $T: V \rightarrow W$ such that $\boldsymbol{T}(\boldsymbol{v}) = \boldsymbol{0}$ for every $\boldsymbol{v}$ in $\boldsymbol{V}$ is a linear transformation called the **zero transformation**. Find the kernel and range of the zero transformation.

2. Let $\boldsymbol{V}$ be any vector space. The mapping $I: V \rightarrow V$ defined by $\boldsymbol{I(v)} = \boldsymbol{v}$ is called the **identity operator** on $\boldsymbol{V}$. 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 $\boldsymbol{B^T}=\boldsymbol{B}$. 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 $\boldsymbol{B^T}=\boldsymbol{B}$.
(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 06

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 $S=\left\{v_1, v_2, v_3, ...v_p\right\}$ 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 $\boldsymbol{R^n}$.

Definition: An ***indexed set*** of vectors $\{\boldsymbol{v_1},\ldots, \boldsymbol{v_p}\}$ in ***V*** is said to be **linearly independent** if the vector equation

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + ... + c_p\boldsymbol{v_p} = 0 \qquad (1)$$

has only the trivial solution, i.e. $c_1 = 0, \ldots , c_p = 0$.
The set $\{\boldsymbol{v_1},\ldots,\boldsymbol{v_p}\}$ is said to be **linearly dependent** if (1) has a nontrivial solution, that is, if there are some weights, $c_1,\ldots,c_p$, not all zero, such that (1) holds. In such a case, (1) is called a **linear dependence relation** among $\boldsymbol{v_1}, \ldots , \boldsymbol{v_p}$. 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 $\boldsymbol{v} \neq \boldsymbol{0}$. 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:

$$-2c_1 + 2c_2 = 0 \quad ....(1) \Rightarrow -c_1 + c_2 = 0 ....(4) \, (\textit{dividing by } 2 \textit{ on both sides of } (1))$$
$$c_1 + c_2 = 0 \quad ....(2)$$
$$c_1 - 2c_2 + c_3 = 0 \quad ......(3)$$

Solving (2) *and* (4) *implies* :

$$\begin{array}{l|l|l|} c_1 + c_2 = 0 & & \textit{Solving } (3) \textit{ implies}: \\ -c_1 + c_2 = 0 & \textit{Solving } (2) \textit{ implies}: & 0+0+c_3 = 0 \\ \hline 0 + 2c_2 = 0 & c_1 + 0 = 0 & \Rightarrow c_3 = 0 \\ \Rightarrow c_2 = 0 & \Rightarrow c_1 = 0 & \end{array}$$

$\Rightarrow c_1 = c_2 = c_3 = 0$; *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 $3p_1 - p_2 + 2\,p_3 = 0$.

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 { $\boldsymbol{v_1}$, … , $\boldsymbol{v_p}$ } 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, $\boldsymbol{v_1}$, … , $\boldsymbol{v_{j-1}}$.

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*.

Useful results:

- 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

Definition: Let $\boldsymbol{H}$ be a subspace of a vector space $\boldsymbol{V}$. An indexed set of vectors $\boldsymbol{B} = \{\boldsymbol{b_1},\ldots,\boldsymbol{b_p}\}$ in **V** is a **basis** for $\boldsymbol{H}$ if

(i) $\boldsymbol{B}$ is a linearly independent set, and
(ii) the subspace spanned by $\boldsymbol{B}$ coincides with $\boldsymbol{H}$; that is,
$\boldsymbol{H} = Span\ \{\boldsymbol{b_1},\ldots,\boldsymbol{b_p}\}$

The definition of a basis applies to the case when $\boldsymbol{H} = \boldsymbol{V}$, because any vector space is a subspace of itself. Thus a basis of $\boldsymbol{V}$ is a linearly independent set that spans $\boldsymbol{V}$. Observe that when $\boldsymbol{H} \neq \boldsymbol{V}$, condition (ii) includes the requirement that each of the vectors $\boldsymbol{b_1},\ldots,\boldsymbol{b_p}$ must belong to $\boldsymbol{H}$, because $Span\ \{\boldsymbol{b_1},\ldots,\boldsymbol{b_p}\}$ contains $\boldsymbol{b_1},\ldots,\boldsymbol{b_p}$, as we saw in lecture 21.

Example 3: Let A be an invertible $n \times n$ matrix – say, $A = [\boldsymbol{a_1} \ldots \boldsymbol{a_n}]$. Then the columns of A form a basis for $\boldsymbol{R^n}$ because they are linearly independent and they span $\boldsymbol{R^n}$, by the Invertible Matrix Theorem.

Example 4: Let $\boldsymbol{e_1},\ldots,\boldsymbol{e_n}$ be the columns of the $n \times n$ identity matrix, I_n. That is,

$$\boldsymbol{e_1} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{e_2} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots \quad \boldsymbol{e_n} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

The set $\{\boldsymbol{e_1},\ldots,\boldsymbol{e_n}\}$ is called the standard basis for $\boldsymbol{R}^n$ (Fig. 1).

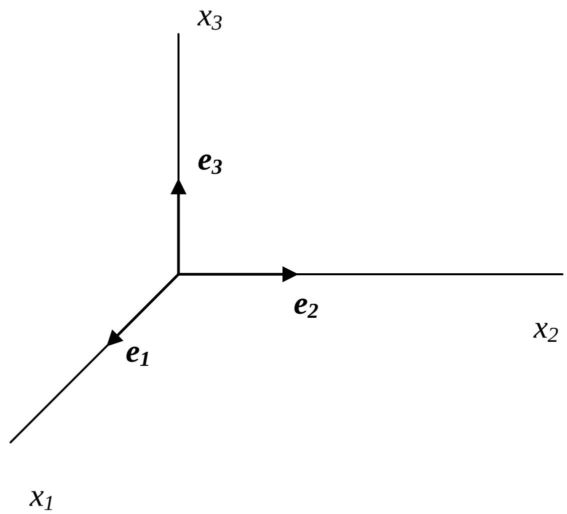


Figure 1 - The standard basis for $\mathbf{R}^3$

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

$$c_0.\mathbf{1} + c_1\mathbf{t} + c_2\mathbf{t}^2 + \ldots.. + c_n\mathbf{t}^n = \mathbf{0}\ (\mathbf{t}) \qquad (2)$$

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.

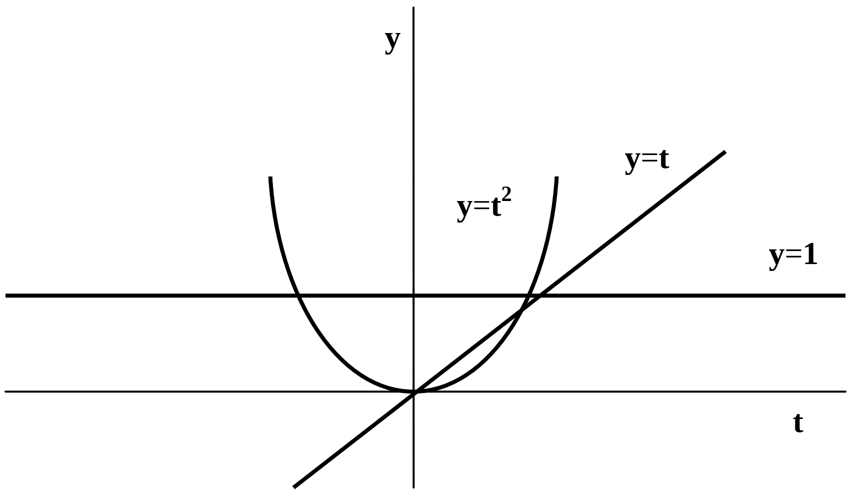


Figure 2 – The standard basis for P_2

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

$$\begin{aligned} 2c_1 + 4c_2 + 0c_3 &= 1d_1 + 0d_2 + 0d_3 \\ -3c_1 + 1c_2 - 7c_3 &= 0d_1 + 1d_2 + 0d_3 \\ 1c_1 + 1c_2 + 1c_3 &= 0d_1 + 0d_2 + 1d_3 \end{aligned} \qquad \text{(A)}$$

$$\Rightarrow \begin{bmatrix} 2 & 4 & 0 \\ -3 & 1 & -7 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

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***.

Example 8: 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 $S = \{\mathbf{p_1}(t), \mathbf{p_2}(t), \mathbf{p_3}(t)\}$ of vectors in $\mathbf{P_2}$ spans $\mathbf{V} = \mathbf{P_2}$ if

$$c_1\mathbf{p_1}(t) + c_2\mathbf{p_2}(t) + c_3\mathbf{p_3}(t) = d_1\mathbf{q_1}(t) + d_2\mathbf{q_2}(t) + d_3\mathbf{q_3}(t) \qquad (*)$$

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 $S = \{\mathbf{p_1}(t), \mathbf{p_2}(t), \mathbf{p_3}(t)\}$ of vectors in $\mathbf{P_2}$ is linearly independent if the only solution of

$$c_1\mathbf{p_1}(t) + c_2\mathbf{p_2}(t) + c_3\mathbf{p_3}(t) = \mathbf{0} \qquad (**)$$

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 $\boldsymbol{p_1(t)}, \boldsymbol{p_2(t)}, \boldsymbol{p_3(t)}$, (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:

$$\begin{aligned} -4\,c_1 + 6\,c_2 + 8\,c_3 &= 0 \\ 1\,c_1 + 5\,c_2 + 4\,c_3 &= 0 \\ 3\,c_1 + 2\,c_2 + 1\,c_3 &= 0 \end{aligned} \quad \Rightarrow \begin{bmatrix} -4 & 6 & 8 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

As $det \begin{bmatrix} -4 & 6 & 8 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} = -26 \neq 0$. Therefore the set $\boldsymbol{S} = \{\boldsymbol{p_1(t)}, \boldsymbol{p_2(t)}, \boldsymbol{p_3(t)}\}$ 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

$$c_1 \boldsymbol{v_1} + c_2 \boldsymbol{v_2} + c_3 \boldsymbol{v_3} + c_4 \boldsymbol{v_4} = d_1 \boldsymbol{w_1} + d_2 \boldsymbol{w_2} + d_3 \boldsymbol{w_3} + d_4 \boldsymbol{w_4} \quad (*)$$

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 $\boldsymbol{S}$.

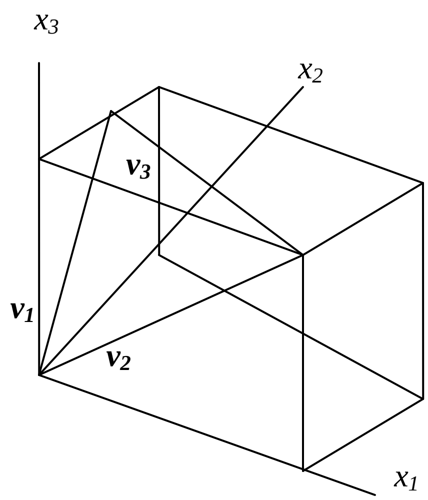
<u>Example 11</u>: Let $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ -2 \\ -3 \end{bmatrix}$, $\boldsymbol{v_2} = \begin{bmatrix} -3 \\ 5 \\ 7 \end{bmatrix}$, $\boldsymbol{v_3} = \begin{bmatrix} -4 \\ 5 \\ 6 \end{bmatrix}$, and $\boldsymbol{H} = Span\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$.

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 $\boldsymbol{S} = \{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ be a set in $\boldsymbol{V}$ and let $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$.

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.

<u>Remark</u>: 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 $\boldsymbol{V}$.

<u>Procedure</u>:

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.

<u>Example 12</u>: 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 $\boldsymbol{S}$.

<u>Solution:</u> 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

$$\begin{bmatrix}
1 & 0 & \frac{1}{2} & 3/2 & 3/2 & : & 0\\
0 & 1 & -1/2 & 3/2 & -5/2 & : & 0\\
0 & 0 & 0 & 0 & 0 & : & 0\\
0 & 0 & 0 & 0 & 0 & : & 0
\end{bmatrix}$$

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.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix},\begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

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 $\boldsymbol{R^4}$ = 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 $\boldsymbol{v_1}$ = Cos^2 x, $\boldsymbol{v_2}$ = 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 07

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 can generate the whole vector space, then such a vector space is called infinite dimensional.

Note:

(1) The dimension of the zero vector space $\{\boldsymbol{0}\}$ is defined to be zero.
(2) Every finite dimensional vector space contains a basis.

Example 1: The n dimensional set of real numbers $\mathbf{R^n}$, set of polynomials of order n $\mathbf{P_n}$, and set of matrices of order $m \times n$ $\mathbf{M_{mn}}$ are all finite- dimensional vector spaces. However, the vector spaces **F** $(-\infty, \infty)$, **C** $(-\infty, \infty)$, and $\mathbf{C^m}$ $(-\infty, \infty)$ 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 $\quad \boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

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 $\quad \boldsymbol{A_1} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$, $\boldsymbol{A_2} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, 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 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. $\{\boldsymbol{A_1}, \boldsymbol{A_2}, \boldsymbol{A_3}\}$ is said to be linearly independent if

$aA_1 + bA_2 + cA_3 = 0 \Rightarrow a=b=c=0$ *i.e.,*

$$a\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} a & 0 \\ 0 & -2a \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & b \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c & -3c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

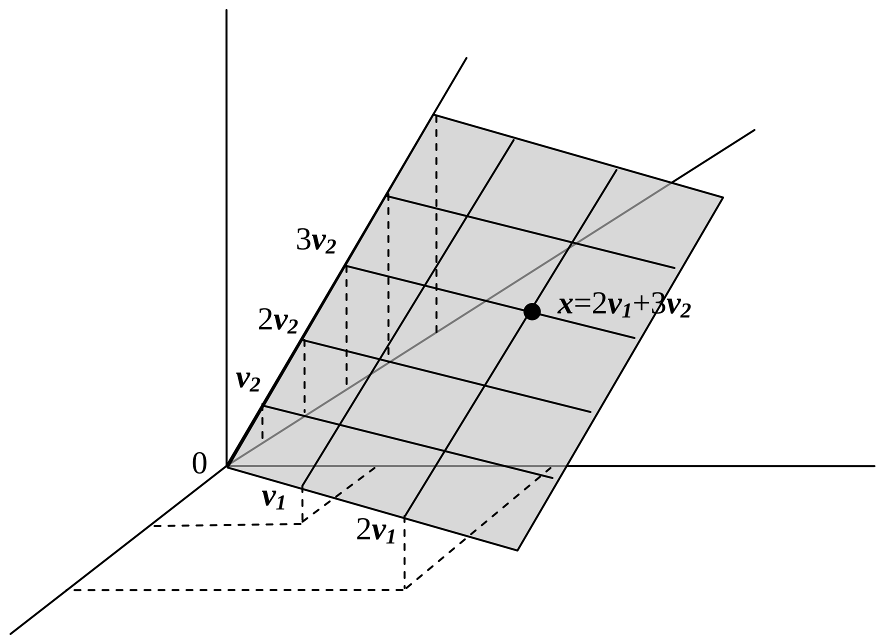
$$\begin{bmatrix} a & b \\ c & -2a+b-3c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Equating the elements, we get

$a=0, b=0, c=0$

This implies $\{\mathbf{A_1}, \mathbf{A_2}, \mathbf{A_3}\}$ is a linearly independent set that spans W. Hence, it's the basis of W with dim(W)= 3.

Example 4: 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

$$\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{v_2} = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{v_3} = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad \boldsymbol{v_4} = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Clearly, $\boldsymbol{v_1} \neq \boldsymbol{0}, \boldsymbol{v_2}$ is not a multiple of $\boldsymbol{v_1}$, but $\boldsymbol{v_3}$ is a multiple of $\boldsymbol{v_2}$. By the Spanning Set Theorem, we may discard $\boldsymbol{v_3}$ and still have a set that spans ***H***. Finally; $\boldsymbol{v_4}$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. So $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is linearly independent and hence is a basis for $\boldsymbol{H}$. Thus dim $\boldsymbol{H} = 3$.

Example 6: The subspaces of $\boldsymbol{R}^3$ can be classified by various dimensions as shown in Fig. 1.

***0*-dimensional subspaces:**

The only 0-dimensional subspace of $\boldsymbol{R}^3$ is zero space.

***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.

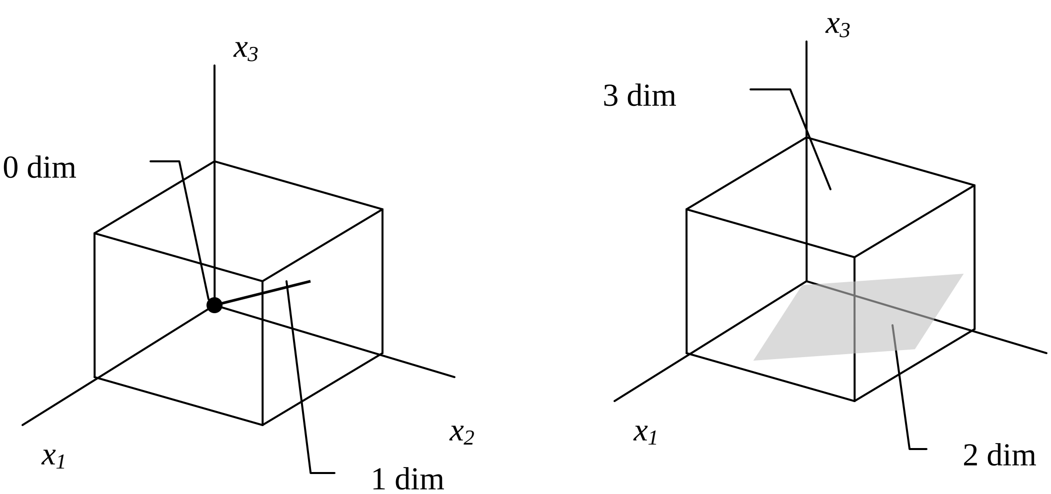


Figure 1 – Sample subspaces of $\boldsymbol{R}^3$

Bases for Nul *A* and Col *A*:

We already know how to find vectors that span the null space of a matrix $\boldsymbol{A}$. The discussion in Lecture 21 pointed out that our method always produces a linearly independent set. Thus the method produces a basis for Nul $\boldsymbol{A}$.

Example 7: Find a basis for the null space of $\boldsymbol{A} = \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$.

Solution: The null space of $\boldsymbol{A}$ is the solution space of homogeneous system

$$\begin{aligned} 2x_1 + 2x_2 - x_3 \qquad + x_5 &= 0 \\ -x_1 - x_2 + 2x_3 - 3x_4 + x_5 &= 0 \\ x_1 + x_2 - 2x_3 \qquad - x_5 &= 0 \\ x_3 + x_4 + x_5 &= 0 \end{aligned}$$

The most appropriate way to solve this system is to reduce its augmented matrix into reduced echelon form.

$$\begin{bmatrix} 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \\ 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} R_4 \leftrightarrow R_2,\ R_3 \leftrightarrow R_1$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{bmatrix} R_3 - 2R_1,\ R_3 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 0 & 3 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{bmatrix} R_3 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{bmatrix} -\frac{1}{3}R_3$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \end{bmatrix} R_4 + R_1$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \end{bmatrix} R_4 + 3R_3$$

$$\sim \begin{bmatrix} 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} R_1 + 2R_2$$

$$\sim \begin{bmatrix} 1 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} R_2 - R_3,\ R_1 - 2R_3$$

$$\sim \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

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

$$\begin{aligned} 1x_1 + 1x_2 + \quad\quad 1x_5 &= 0 \\ 1x_3 + \quad 1x_5 &= 0 \\ 1x_4 \quad\quad &= 0 \\ 0 &= 0 \end{aligned}$$

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.

$$x_1 = -1x_2 - 1x_5$$
$$x_2 = s$$
$$x_3 = -1x_5$$
$$x_4 = 0$$
$$x_5 = t$$

The general solution of the given system is

$$x_1 = -s-t \quad , \quad x_2 = s \quad , \quad x_3 = -t \quad , \quad x_4 = 0 \quad , \quad x_5 = t$$

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 $\boldsymbol{v_1} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$ span the solution space .Since they are also linearly independent,$\{\boldsymbol{v_1},\boldsymbol{v_2}\}$ is a basis for Nul $\boldsymbol{A}$.

The next two examples describe a simple algorithm for finding a basis for the column space.

Example 8: Find a basis for Col $\boldsymbol{B}$, where $\boldsymbol{B} = [\boldsymbol{b_1} \quad \boldsymbol{b_2}, \quad ..., \quad \boldsymbol{b_5}] = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Solution Each non-pivot column of $\boldsymbol{B}$ is a linear combination of the pivot columns. In fact, $\boldsymbol{b_2} = 4\boldsymbol{b_1}$ and $\boldsymbol{b_4} = 2\boldsymbol{b_1} - \boldsymbol{b_3}$. By the Spanning Set Theorem, we may discard $\boldsymbol{b_2}$ and $\boldsymbol{b_4}$ and $\{\boldsymbol{b_1}, \boldsymbol{b_3}, \boldsymbol{b_5}\}$ will still span Col $\boldsymbol{B}$. Let

$$\boldsymbol{S} = \{\boldsymbol{b_1}, \quad \boldsymbol{b_3}, \quad \boldsymbol{b_5}\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Since $\boldsymbol{b_1} \neq \boldsymbol{0}$ and no vector in $\boldsymbol{S}$ is a linear combination of the vectors that precede it, $\boldsymbol{S}$ is linearly independent. Thus $\boldsymbol{S}$ is a basis for Col $\boldsymbol{B}$.

What about a matrix $\boldsymbol{A}$ that is not in reduced echelon form? Recall that any linear dependence relationship among the columns of $\boldsymbol{A}$ can be expressed in the form $\boldsymbol{Ax}$

= ***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

$$\boldsymbol{A} = [\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \ldots \quad \boldsymbol{a_5}] = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$$

is row equivalent to the matrix ***B*** in Example 8. Find a basis for Col ***A***.

Solution: In Example 8, we have seen that $\boldsymbol{b_2} = 4\boldsymbol{b_1}$ *and* $\boldsymbol{b_4} = 2\boldsymbol{b_1} - \boldsymbol{b_3}$ so we can expect that $\boldsymbol{a_2} = 4\boldsymbol{a_1}$ *and* $\boldsymbol{a_4} = 2\boldsymbol{a_1} - \boldsymbol{a_3}$. 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***. In fact, $\{\boldsymbol{a_1}, \boldsymbol{a_3}, \boldsymbol{a_5}\}$ 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 $v_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ and $v_2 = \begin{bmatrix} -2 \\ 7 \\ -9 \end{bmatrix}$. Determine if $\{v_1, v_2\}$ is a basis for R^3. Is $\{v_1, v_2\}$ a basis for R^2?

Solution Let $A = [v_1 \quad v_2]$. 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 R^3, by Theorem 4 in Lecture 6. Hence $\{v_1, v_2\}$ is not a basis for R^3. Since v_1 and v_2 are not in R^2, they cannot possibly be a basis for R^2. However, since v_1 and v_2 are obviously linearly independent, they are a basis for a subspace of R^3, namely, Span $\{v_1, v_2\}$.

Example 11: Let $v_1 = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}, v_2 = \begin{bmatrix} 6 \\ 2 \\ -1 \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}, v_4 = \begin{bmatrix} -4 \\ -8 \\ 9 \end{bmatrix}$. Find a basis for the subspace W spanned by $\{v_1, v_2, v_3, v_4\}$.

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 $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}\}$ in $\boldsymbol{R^n}$, the following procedure produces a subset of these vectors that form a basis for span ($\boldsymbol{S}$) and expresses those vectors of $\boldsymbol{S}$ that are not in the basis as linear combinations of the basis vector.

Step1: Form the matrix $\boldsymbol{A}$ having $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}$ as its column vectors.

Step2: Reduce the matrix $\boldsymbol{A}$ to its reduced row echelon form $\boldsymbol{R}$, and let $\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}$ be the column vectors of $\boldsymbol{R}$.

Step3: Identify the columns that contain the leading entries i.e., 1's in $\boldsymbol{R}$. The corresponding column vectors of $\boldsymbol{A}$ are the basis vectors for span ($\boldsymbol{S}$).

Step4: Express each column vector of $\boldsymbol{R}$ 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 $\boldsymbol{R}$. The corresponding equations for the column vectors of $\boldsymbol{A}$ 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 $\boldsymbol{v_1} = (1, -2, 0, 3)$, $\boldsymbol{v_2} = (2, -4, 0, 6)$, $\boldsymbol{v_3} = (-1, 1, 2, 0)$ and $\boldsymbol{v_4} = (0, -1, 2, 3)$ 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 $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}$ as its column vectors

$$\begin{array}{c} \begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -4 & 1 & -1 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 0 & 3 \end{bmatrix} \\ \begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ \boldsymbol{v_1} & \boldsymbol{v_2} & \boldsymbol{v_3} & \boldsymbol{v_4} \end{array} \end{array} \qquad (A)$$

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

$$\begin{array}{cccc} \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\ w_1 \quad w_2 \quad w_3 \quad w_4 \end{array} \qquad (B)$$

The leading entries occur in column 1 and 3 so $\{\boldsymbol{w_1}, \boldsymbol{w_3}\}$ is a basis for the column space of (B) and consequently $\{\boldsymbol{v_1}, \boldsymbol{v_3}\}$ is the basis for column space of (A).

(b) We shall start by expressing $\boldsymbol{w_2}$ and $\boldsymbol{w_4}$ as linear combinations of the basis vector $\boldsymbol{w_1}$ and $\boldsymbol{w_3}$. The simplest way of doing this is to express $\boldsymbol{w_2}$ and $\boldsymbol{w_4}$ in term of basis vectors with smaller subscripts. Thus we shall express $\boldsymbol{w_2}$ as a linear combination of $\boldsymbol{w_1}$, and we shall express $\boldsymbol{w_4}$ as a linear combination of $\boldsymbol{w_1}$ and $\boldsymbol{w_3}$. By inspection of (B), these linear combinations are $\boldsymbol{w_2} = 2\boldsymbol{w_1}$ and $\boldsymbol{w_4} = \boldsymbol{w_1} + \boldsymbol{w_3}$. We call them the dependency equations. The corresponding relationship of (A) are $\boldsymbol{v_3} = 2\boldsymbol{v_1}$ and $\boldsymbol{v_5} = \boldsymbol{v_1} + \boldsymbol{v_3}$.

Example 13: Basis and Linear Combinations

(a) Find a subset of the vectors $\boldsymbol{v_1} = (1, -1, 5, 2)$, $\boldsymbol{v_2} = (-2, 3, 1, 0)$, $\boldsymbol{v_3} = (4, -5, 9, 4)$, $\boldsymbol{v_4} = (0, 4, 2, -3)$ and $\boldsymbol{v_5} = (-7, 18, 2, -8)$ 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

Solution: (a) We begin by constructing a matrix that has $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_5}$ as its column vectors

$$\begin{array}{c} \begin{bmatrix} 1 & -2 & 4 & 0 & -7 \\ -1 & 3 & -5 & 4 & 18 \\ 5 & 1 & 9 & 2 & 2 \\ 2 & 0 & 4 & -3 & -8 \end{bmatrix} \\ \begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \boldsymbol{v_1} & \boldsymbol{v_2} & \boldsymbol{v_3} & \boldsymbol{v_4} & \boldsymbol{v_5} \end{matrix} \end{array} \qquad \text{(A)}$$

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

$$\begin{array}{ccccc} \begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\ \boldsymbol{w_1} \quad \boldsymbol{w_2} \quad \boldsymbol{w_3} \quad \boldsymbol{w_4} \quad \boldsymbol{w_5} \end{array} \qquad \text{(B)}$$

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

$$\begin{array}{c} \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix} \\ \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\ \boldsymbol{v_1} \quad \boldsymbol{v_2} \quad \boldsymbol{v_3} \quad \boldsymbol{v_4} \quad \boldsymbol{v_5} \end{array} \qquad \text{(A)}$$

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

$$\begin{array}{c} \begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\ \boldsymbol{w_1} \quad \boldsymbol{w_2} \quad \boldsymbol{w_3} \quad \boldsymbol{w_4} \quad \boldsymbol{w_5} \end{array} \qquad \text{(B)}$$

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 $\boldsymbol{w_3 = 2w_1 - w_2}$ and $\boldsymbol{w_5 = w_1 + w_2 + w_4}$
The corresponding relationship of (A) are $\boldsymbol{v_3 = 2v_1 - v_2}$ ***and*** $\boldsymbol{v_5 = v_1 + v_2 + v_4}$

Subspaces of a Finite-Dimensional Space: The next theorem is a natural counterpart to the Spanning Set Theorem.

Theorem 5: Let $\boldsymbol{H}$ be a subspace of a finite-dimensional vector space $\boldsymbol{V}$. Any linearly independent set in $\boldsymbol{H}$ can be expanded, if necessary, to a basis for $\boldsymbol{H}$. Also, $\boldsymbol{H}$ 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 $\boldsymbol{V}$ be a p-dimensional vector space, $p \geq 1$. Any linearly independent set of exactly p elements in $\boldsymbol{V}$ is automatically a basis for $\boldsymbol{V}$. Any set of exactly p elements that spans $\boldsymbol{V}$ is automatically a basis for $\boldsymbol{V}$.

The Dimensions of Nul *A* and Col *A*: Since the pivot columns of a matrix $\boldsymbol{A}$ form a basis for Col $\boldsymbol{A}$, we know the dimension of Col $\boldsymbol{A}$ as soon as we know the pivot columns. The dimension of Nul $\boldsymbol{A}$ might seem to require more work, since finding a basis for Nul $\boldsymbol{A}$ usually takes more time than a basis for Col $\boldsymbol{A}$. Yet, there is a shortcut.

Let A be an $m \times n$ matrix, and suppose that the equation $\boldsymbol{Ax = 0}$ 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, $\boldsymbol{u_1}, \dots, \boldsymbol{u_k}$, one for each free variable. So $\{\boldsymbol{u_1}, \dots, \boldsymbol{u_k}\}$ 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 $\boldsymbol{Ax = 0}$, 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 08

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 $\boldsymbol{A}$ and then determining both the maximum number of linearly independent columns in $\boldsymbol{A}$ and the maximum number of linearly independent columns in $\boldsymbol{A}^T$ (rows in $\boldsymbol{A}$). 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 $\boldsymbol{A}$.

<u>The Row Space</u>: If $\boldsymbol{A}$ is an $m \times n$ matrix, each row of $\boldsymbol{A}$ has n entries and thus can be identified with a vector in $\boldsymbol{R}^n$. The set of all linear combinations of the row vectors is called the row space of $\boldsymbol{A}$ and is denoted by Row $\boldsymbol{A}$. Each row has n entries, so Row $\boldsymbol{A}$ is a subspace of $\boldsymbol{R}^n$. Since the rows of $\boldsymbol{A}$ are identified with the columns of $\boldsymbol{A}^T$, we could also write Col $\boldsymbol{A}^T$ in place of Row $\boldsymbol{A}$.

<u>Example 1</u>: Let $\boldsymbol{A} = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$ and $\boldsymbol{r}_1 = (-2,-5,8,0,-17)$, $\boldsymbol{r}_2 = (1,3,-5,1,5)$, $\boldsymbol{r}_3 = (3,11,-19,7,1)$, $\boldsymbol{r}_4 = (1,7,-13,5,-3)$

The row space of $\boldsymbol{A}$ is the subspace of $\boldsymbol{R}^5$ spanned by $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4\}$. That is, Row $\boldsymbol{A}$ = Span $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4\}$. Naturally, we write row vectors horizontally; however, they could also be written as column vectors

<u>Example</u>: Let

$\boldsymbol{A} = \begin{bmatrix} 2 & 1 & 0 \\ 3 & -1 & 4 \end{bmatrix}$ and $\boldsymbol{r}_1 = (2,1,0)$, $\boldsymbol{r}_2 = (3,-1,4)$

That is Row **A**=Span $\{\boldsymbol{r}_1, \boldsymbol{r}_2\}$.

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 $\boldsymbol{A}$ and $\boldsymbol{B}$ are row equivalent, then their row spaces are the same. If $\boldsymbol{B}$ is in echelon form, the nonzero rows of $\boldsymbol{B}$ form a basis for the row space of $\boldsymbol{A}$ as well as $\boldsymbol{B}$.

<u>Theorem 2</u>: If $\boldsymbol{A}$ and $\boldsymbol{B}$ are row equivalent matrices, then

(a) A given set of column vectors of $\boldsymbol{A}$ is linearly independent if and only if the corresponding column vectors of $\boldsymbol{B}$ are linearly independent.
(b) A given set of column vector of $\boldsymbol{A}$ forms a basis for the column space of $\boldsymbol{A}$ if and only if the corresponding column vector of $\boldsymbol{B}$ forms a basis for the column space of $\boldsymbol{B}$.

Example 2: (Bases for Row and Column Spaces)

Find the bases for the row and column spaces of $\boldsymbol{A} = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$.

Solution: We can find a basis for the row space of $\boldsymbol{A}$ by finding a basis for the row space of any row-echelon form of $\boldsymbol{A}$.

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 $\boldsymbol{A}$: $\boldsymbol{R} = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Here Theorem 1 implies that that the non zero rows are the basis vectors of the matrix. So these bases vectors are

$$\boldsymbol{r_1} = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \end{bmatrix}$$
$$\boldsymbol{r_2} = \begin{bmatrix} 0 & 0 & 1 & 3 & -2 & -6 \end{bmatrix}$$
$$\boldsymbol{r_3} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 5 \end{bmatrix}$$

$\boldsymbol{A}$ and $\boldsymbol{R}$ may have different column spaces, we cannot find a basis for the column space of $\boldsymbol{A}$ directly from the column vectors of $\boldsymbol{R}$. 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.

<u>Example 3</u>: (Basis for a Vector Space using Row Operation)
Find bases for the space spanned by the vectors

$\boldsymbol{v_1} = (1,-2,0,0,3) \qquad \boldsymbol{v_2} = (2,-5,-3,-2,6)$

$\boldsymbol{v_3} = (0,5,15,10,0) \qquad \boldsymbol{v_4} = (2,6,18,8,6)$

<u>Solution</u>: The 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 ***A*** leads to bases for the three spaces: Row ***A***, Col ***A***, and Nul ***A***.

<u>Example 5</u>: Find bases for the row space, the column space and the null space of the matrix

$$\boldsymbol{A} = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

<u>Solution</u>: To find bases for the row space and the column space, row reduce ***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 ***B*** form a basis for the row space of ***A*** (as well as the row space of ***B***). Thus Basis for Row ***A***:

$$\{(1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20)\}$$

For the column space, observe from ***B*** that the pivots are in columns *1*, *2* and *4*. Hence columns *1*, *2* and *4* of ***A*** (not ***B***) form a basis for Col ***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 ***A*** provides (in its nonzero rows) a basis for Row ***A*** and also identifies the pivot columns of ***A*** for Col ***A***. However, for Nul ***A***, we need the reduced echelon form. Further row operations on ***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 $\boldsymbol{x} \rightarrow \boldsymbol{Ax}$; the range of this mapping, Col $\boldsymbol{A}$, is shown in a separate copy of $\boldsymbol{R}^3$, along with Nul $\boldsymbol{A}^T$.

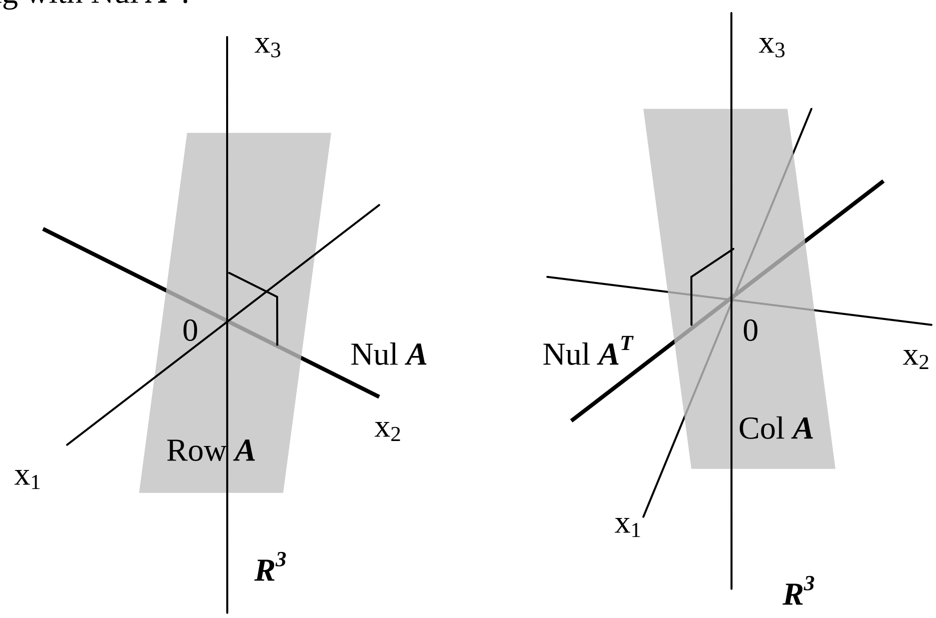


Figure 1 – Subspaces associated with a matrix *A*

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.

Example 9: Find the rank and nullity of the matrix $\boldsymbol{A}=\begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$.

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 $\boldsymbol{A}$ 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

$$\begin{aligned} x_1 &= 4r + 28s + 37t - 13u \\ x_2 &= 2r + 12s + 16t - 5u \\ x_3 &= r \\ x_4 &= s \\ x_5 &= t \\ x_6 &= u \end{aligned}$$

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 $(\boldsymbol{A}) = 7\text{-}3 = 4$
Thus, there are four parameters.

Example: 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 $(\boldsymbol{A}) = 4\text{-}0 = 4$
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 $\boldsymbol{A}$	r
Column space of $\boldsymbol{A}$	r
Null space of $\boldsymbol{A}$	n-r
Null space of $\boldsymbol{A}^T$	m-r

Example 12: If $\boldsymbol{A}$ is a 7 x 4 matrix, then the rank of $\boldsymbol{A}$ is at most 4 and, consequently, the seven row vectors must be linearly dependent. If $\boldsymbol{A}$ is a 4 x 7 matrix, then again the rank of $\boldsymbol{A}$ 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 $\boldsymbol{A}$ be an n x n matrix. Then the following statements are each equivalent to the statement that $\boldsymbol{A}$ is an invertible matrix.

m. The columns of $\boldsymbol{A}$ form a basis of $\boldsymbol{R}^n$.
n. Col $\boldsymbol{A} = \boldsymbol{R}^n$.
o. dim Col $\boldsymbol{A} = n$
p. rank $\boldsymbol{A} = n$
q. Nul $\boldsymbol{A} = \{\boldsymbol{0}\}$
r. dim Nul $\boldsymbol{A} = 0$

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 $\boldsymbol{A}$ 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 $\boldsymbol{A}$, because the row space is the column space of $\boldsymbol{A}^T$. Recall from (1) of the Invertible Matrix Theorem that $\boldsymbol{A}$ 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 $\boldsymbol{A}$ is often determined from the singular value decomposition of $\boldsymbol{A}$.

Example 13: The matrices below are row equivalent

$$\boldsymbol{A} = \begin{bmatrix} 2 & -1 & 1 & -6 & 8 \\ 1 & -2 & -4 & 3 & -2 \\ -7 & 8 & 10 & 3 & -10 \\ 4 & -5 & -7 & 0 & 4 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 1 & -2 & -4 & 3 & -2 \\ 0 & 3 & 9 & -12 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. Find rank $\boldsymbol{A}$ and dim Nul $\boldsymbol{A}$.
2. Find bases for Col $\boldsymbol{A}$ and Row $\boldsymbol{A}$.
3. What is the next step to perform if one wants to find a basis for Nul $\boldsymbol{A}$?
4. How many pivot columns are in a row echelon form of $\boldsymbol{A}^{\boldsymbol{T}}$?

Solution:

1. $\boldsymbol{A}$ has two pivot columns, so rank $\boldsymbol{A} = 2$. Since $\boldsymbol{A}$ has *5* columns altogether, dim Nul $\boldsymbol{A} = 5 - 2 = 3$.
2. The pivot columns of $\boldsymbol{A}$ are the first two columns. So a basis for Col $\boldsymbol{A}$ is

$$\{\boldsymbol{a}_1, \boldsymbol{a}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 8 \\ -5 \end{bmatrix} \right\}$$

   The nonzero rows of $\boldsymbol{B}$ form a basis for Row $\boldsymbol{A}$, namely $\{(1, -2, -4, 3, -2), (0, 3, 9, -12, 12)\}$. In this particular example, it happens that any two rows of $\boldsymbol{A}$ form a basis for the row space, because the row space is two-dimensional and none of the rows of $\boldsymbol{A}$ is a multiple of another row. In general, the nonzero rows of an echelon form of $\boldsymbol{A}$ should be used as a basis for Row $\boldsymbol{A}$, not the rows of $\boldsymbol{A}$ itself.
3. For Nul $\boldsymbol{A}$, the next step is to perform row operations on $\boldsymbol{B}$ to obtain the reduced echelon form of $\boldsymbol{A}$.
4. Rank $\boldsymbol{A}^{\boldsymbol{T}}$ = rank $\boldsymbol{A}$, by the Rank Theorem, because Col $\boldsymbol{A}^{\boldsymbol{T}}$ = Row $\boldsymbol{A}$. So $\boldsymbol{A}^{\boldsymbol{T}}$ has two pivot positions.

Exercises:

In exercises 1 to 4, assume that the matrix $\boldsymbol{A}$ is row equivalent to $\boldsymbol{B}$. Without calculations, list rank $\boldsymbol{A}$ and dim Nul $\boldsymbol{A}$. Then find bases for Col $\boldsymbol{A}$, Row $\boldsymbol{A}$, and Nul $\boldsymbol{A}$.

1. $\boldsymbol{A} = \begin{bmatrix} 1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

2. $\boldsymbol{A} = \begin{bmatrix} 1 & -3 & 4 & -1 & 9 \\ -2 & 6 & -6 & -1 & -10 \\ -3 & 9 & -6 & -6 & -3 \\ 3 & -9 & 4 & 9 & 0 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 1 & -3 & 0 & 5 & -7 \\ 0 & 0 & 2 & -3 & 8 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3. $\boldsymbol{A} = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ -2 & 3 & -3 & -3 & -4 \\ 4 & -6 & 9 & 5 & 9 \\ -2 & 3 & 3 & -4 & 1 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ 0 & 0 & 3 & -1 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

4. $\boldsymbol{A} = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 1 & 2 & -4 & 10 & 13 & -12 \\ 1 & -1 & -1 & 1 & 1 & -3 \\ 1 & -3 & 1 & -5 & -7 & 3 \\ 1 & -2 & 0 & 0 & -5 & -4 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 0 & 1 & -1 & 3 & 4 & -3 \\ 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

5. If a 3 x 8 matrix $\boldsymbol{A}$ has rank 3, find dim Nul $\boldsymbol{A}$, dim Row $\boldsymbol{A}$, and rank $\boldsymbol{A}^{\boldsymbol{T}}$.

6. If a 6 x 3 matrix $\boldsymbol{A}$ has rank 3, find dim Nul $\boldsymbol{A}$, dim Row $\boldsymbol{A}$, and rank $\boldsymbol{A}^{\boldsymbol{T}}$.

7. Suppose that a 4 x 7 matrix $\boldsymbol{A}$ has four pivot columns. Is Col $\boldsymbol{A} = \boldsymbol{R}^4$? Is Nul $\boldsymbol{A} = \boldsymbol{R}^3$? Explain your answers.

8. Suppose that a 5 x 6 matrix $\boldsymbol{A}$ has four pivot columns. What is dim Nul $\boldsymbol{A}$? Is Col $\boldsymbol{A} = \boldsymbol{R}^4$? Why or why not?

9. If the null space of a 5 x 6 matrix $\boldsymbol{A}$ is 4-dimensional, what is the dimension of the column space of $\boldsymbol{A}$?

10. If the null space of a 7 x 6 matrix $\boldsymbol{A}$ is 5-dimensional, what is the dimension of the column space of $\boldsymbol{A}$?

11. If the null space of an 8 x 5 matrix $\boldsymbol{A}$ is 2-dimensional, what is the dimension of the row space of $\boldsymbol{A}$?

12. If the null space of a 5 x 6 matrix $\boldsymbol{A}$ is 4-dimensional, what is the dimension of the row space of $\boldsymbol{A}$?

13. If $\boldsymbol{A}$ is a 7 x 5 matrix, what is the largest possible rank of $\boldsymbol{A}$? If $\boldsymbol{A}$ is a 5 x 7 matrix, what is the largest possible rank of $\boldsymbol{A}$? Explain your answers.

14. If $\boldsymbol{A}$ is a 4 x 3 matrix, what is the largest possible dimension of the row space of $\boldsymbol{A}$? If $\boldsymbol{A}$ is a 3 x 4 matrix, what is the largest possible dimension of the row space of $\boldsymbol{A}$? Explain.

15. If $\boldsymbol{A}$ is a 6 x 8 matrix, what is the smallest possible dimension of Nul $\boldsymbol{A}$?

16. If $\boldsymbol{A}$ is a 6 x 4 matrix, what is the smallest possible dimension of Nul $\boldsymbol{A}$?

Lecture 09

Solution of Linear System of Equations and Matrix Inversion

Jacobi's Method

This is an iterative method, where initial approximate solution to a given system of equations is assumed and is improved towards the exact solution in an iterative way.

In general, when the coefficient matrix of the system of equations is a sparse matrix (many elements are zero), iterative methods have definite advantage over direct methods in respect of economy of computer memory

Such sparse matrices arise in computing the numerical solution of partial differential equations

Let us consider

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \;\vdots \qquad \vdots \qquad\qquad \vdots \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_1 \end{array}\right\}$$

In this method, we assume that the coefficient matrix [A] is strictly diagonally dominant, that is, in each row of [A] the modulus of the diagonal element exceeds the sum of the off-diagonal elements.

We also assume that the diagonal element do not vanish. If any diagonal element vanishes, the equations can always be rearranged to satisfy this condition.

Now the above system of equations can be written as

$$\left.\begin{array}{l} x_1 = \dfrac{b_1}{a_{11}} - \dfrac{a_{12}}{a_{11}}x_2 - \cdots - \dfrac{a_{1n}}{a_{11}}x_n \\ x_2 = \dfrac{b_2}{a_{22}} - \dfrac{a_{21}}{a_{22}}x_1 - \cdots - \dfrac{a_{2n}}{a_{22}}x_n \\ \vdots \qquad \vdots \qquad \vdots \qquad\qquad \vdots \\ x_n = \dfrac{b_n}{a_{nn}} - \dfrac{a_{n1}}{a_{nn}}x_1 - \cdots - \dfrac{a_{n(n-1)}}{a_{nn}}x_{n-1} \end{array}\right\}$$

We shall take this solution vector $(x_1, x_2, ..., x_n)^T$ as a first approximation to the exact solution of system. For convenience, let us denote the first approximation vector by $(x_1^{(1)}, x_2^{(1)}, ..., x_n^{(1)})$ got after taking as an initial starting vector.

Substituting this first approximation in the right-hand side of system, we obtain the second approximation to the given system in the form

$$\left.\begin{aligned}
x_1^{(2)} &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} x_2^{(1)} - \cdots - \frac{a_{1n}}{a_{11}} x_n^{(1)} \\
x_2^{(2)} &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1^{(1)} - \cdots - \frac{a_{2n}}{a_{22}} x_n^{(1)} \\
\vdots \quad & \quad \vdots \qquad \vdots \qquad\qquad \vdots \\
x_n^{(2)} &= \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}} x_1^{(1)} - \cdots - \frac{a_{n(n-1)}}{a_{nn}} x_{n-1}^{(1)}
\end{aligned}\right\}$$

This second approximation is substituted into the right-hand side of Equations and obtain the third approximation and so on.
This process is repeated and *(r+1)th* approximation is calculated

$$\left.\begin{aligned}
x_1^{(r+1)} &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} x_2^{(r)} - \cdots - \frac{a_{1n}}{a_{11}} x_n^{(r)} \\
x_2^{(r+1)} &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1^{(r)} - \cdots - \frac{a_{2n}}{a_{22}} x_n^{(r)} \\
\vdots \quad & \quad \vdots \qquad \vdots \qquad\qquad \vdots \\
x_n^{(r+1)} &= \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}} x_1^{(r)} - \cdots - \frac{a_{n(n-1)}}{a_{nn}} x_{n-1}^{(r)}
\end{aligned}\right\}$$

Briefly, we can rewrite these Equations as

$$x_i^{(r+1)} = \frac{b_i}{a_{ii}} - \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(r)},$$

$$r = 1, 2, \ldots, \qquad i = 1, 2, \ldots, n$$

It is also known as method of simultaneous displacements,
since no element of $x_i^{(r+1)}$ is used in this iteration until every element is computed.

A sufficient condition for convergence of the iterative solution to the exact solution is

$\left|a_{ii}\right| > \sum_{\substack{j=1 \\ j\neq 1}}^{n} \left|a_{ij}\right|, \qquad i = 1, 2, \ldots, n$ When this condition (diagonal dominance) is true, Jacobi's method converges

Example

Find the solution to the following system of equations using Jacobi's iterative method for the first five iterations:

$$83x + 11y - 4z = 95$$
$$7x + 52y + 13z = 104$$
$$3x + 8y + 29z = 71$$

Solution

$$\left.\begin{aligned} x &= \frac{95}{83} - \frac{11}{83}y + \frac{4}{83}z \\ y &= \frac{104}{52} - \frac{7}{52}x - \frac{13}{52}z \\ z &= \frac{71}{29} - \frac{3}{29}x - \frac{8}{29}y \end{aligned}\right\}$$

Taking the initial starting of solution vector as $(0,0,0)^T$, from Eq. ,we have the first approximation as

$$\begin{pmatrix} x^{(1)} \\ y^{(1)} \\ z^{(1)} \end{pmatrix} = \begin{pmatrix} 1.1446 \\ 2.0000 \\ 2.4483 \end{pmatrix}$$

Now, using Eq. ,the second approximation is computed from the equations

$$\left.\begin{aligned} x^{(2)} &= 1.1446 - 0.1325y^{(1)} + 0.0482z^{(1)} \\ y^{(2)} &= 2.0 - 0.1346x^{(1)} - 0.25z^{(1)} \\ z^{(2)} &= 2.4483 - 0.1035x^{(1)} - 0.2759y^{(1)} \end{aligned}\right\}$$

Making use of the last two equations we get the second approximation as

$$\begin{pmatrix} x^{(2)} \\ y^{(2)} \\ z^{(2)} \end{pmatrix} = \begin{pmatrix} 0.9976 \\ 1.2339 \\ 1.7424 \end{pmatrix}$$

Similar procedure yields the third, fourth and fifth approximations to the required solution and they are tabulated as below;

Variables			
Iteration number *r*	*x*	*y*	*z*
1	1.1446	2.0000	2.4483
2	0.9976	1.2339	1.7424
3	1.0651	1.4301	2.0046

4	1.0517	1.3555	1.9435
5	1.0587	1.3726	1.9655

Example

Solve the system by jacobi's iterative method

$8x-3y+2z=20$

$4x+11y-z=33$

$6x+3y+12z=35$

(Perform only four iterations)

Solution

Consider the given system as

$$8x - 3y + 2z = 20$$
$$4x + 11y - z = 33$$
$$6x + 3y + 12z = 35$$

*the system is diagonally do*min*ant*

$$x = \frac{1}{8}\left[20 + 3y - 2z\right]$$
$$y = \frac{1}{11}\left[33 - 4x + z\right]$$
$$z = \frac{1}{12}\left[35 - 6x - 3y\right]$$

we start with an initial aproximation $x_0 = y_0 = z_0 = 0$

substituting these

first iteration

$$x_1 = \frac{1}{8}\left[20 + 3(0) - 2(0)\right] = 2.5$$
$$y_1 = \frac{1}{11}\left[33 - 4(0) + 0\right] = 3$$
$$z_1 = \frac{1}{12}\left[35 - 6(0) - 3(0)\right] = 2.916667$$

Second iteration

$$x_2 = \frac{1}{8}\left[20 + 3(3) - 2(2.9166667)\right] = 2.895833$$
$$y_2 = \frac{1}{11}\left[33 - 4(2.5) + 2.9166667\right] = 2.3560606$$
$$z_2 = \frac{1}{12}\left[35 - 6(2.5) - 3(3)\right] = 0.9166666$$

third iteration

$$x_3 = \frac{1}{8}\left[20 + 3(2.3560606) - 2(0.9166666)\right] = 3.1543561$$
$$y_3 = \frac{1}{11}\left[33 - 4(2.8958333) + 0.9166666\right] = 2.030303$$
$$z_3 = \frac{1}{12}\left[35 - 6(2.8958333) - 3(2.3560606)\right] = 0.8797348$$

fourth iteration

$$x_4 = \frac{1}{8}\left[20 + 3(2.030303) - 2(0.8797348)\right] = 3.0419299$$

$$y_4 = \frac{1}{11}\left[33 - 4(3.1543561) + 0.8797348\right] = 1.9329373$$

$$z_4 = \frac{1}{12}\left[35 - 6(3.1543561) - 3(2.030303)\right] = 0.8319128$$

Example

Solve the system by jacobi's iterative method

$$3x + 4y + 15z = 54.8$$

$$x + 12y + 3z = 39.66$$

$$10x + y - 2z = 7.74$$

(Perform only four iterations)

Solution

Consider the given system as

$$3x + 4y + 15z = 54.8$$

$$x + 12y + 3z = 39.66$$

$$10x + y - 2z = 7.74$$

*the system is not diagonally do*min *ant we rearrange the system*

$$10x + y - 2z = 7.74$$

$$x + 12y + 3z = 39.66$$

$$3x + 4y + 15z = 54.8$$

$$x = \frac{1}{10}\left[7.74 - y + 2z\right]$$

$$y = \frac{1}{12}\left[39.66 - x - 3z\right]$$

$$z = \frac{1}{15}\left[54.8 - 3x - 4y\right]$$

we start with an initial aproximation $x_0 = y_0 = z_0 = 0$

substituting these

first iteration

$$x_1 = \frac{1}{10}\left[7.74 - (0) + 2(0)\right] = 0.774$$

$$y_1 = \frac{1}{12}\left[39.66 - (0) - 3(0)\right] = 1.1383333$$

$$z_1 = \frac{1}{15}\left[54.8 - 3(0) - 4(0)\right] = 3.6533333$$

Second iteration

$$x_2 = \frac{1}{10}\left[7.74 - 1.1383333 + 2(3.6533333)\right] = 1.3908333$$

$$y_2 = \frac{1}{12}\left[39.66 - 0.774 - 3(3.6533333)\right] = 2.3271667$$

$$z_2 = \frac{1}{15}\left[54.8 - 3(0.774) - 4(1.1383333)\right] = 3.1949778$$

third iteration

$$x_3 = \frac{1}{10}\left[7.74 - 2.3271667 + 2(3.1949778)\right] = 1.1802789$$

$$y_3 = \frac{1}{12}\left[39.66 - 1.3908333 - 3(3.1949778)\right] = 2.3903528$$

$$z_3 = \frac{1}{15}\left[54.8 - 3(1.3908333) - 4(2.3271667)\right] = 2.7545889$$

fourth iteration

$$x_4 = \frac{1}{10}\left[7.74 - 2.5179962 + 2(2.7798501)\right] = 1.0781704$$

$$y_4 = \frac{1}{12}\left[39.66 - 1.1802789 - 3(2.7545889)\right] = 2.51779962$$

$$z_4 = \frac{1}{15}\left[54.8 - 3(1.1802789) - 4(2.3903528)\right] = 2.7798501$$

Lecture 10

Solution of Linear System of Equations and Matrix Inversion
Gauss–Seidel Iteration Method

It is another well-known iterative method for solving a system of linear equations of the form

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}\right\}$$

In Jacobi's method, the $(r + 1)th$ approximation to the above system is given by Equations

$$\left.\begin{array}{l} x_1^{(r+1)} = \dfrac{b_1}{a_{11}} - \dfrac{a_{12}}{a_{11}}x_2^{(r)} - \cdots - \dfrac{a_{1n}}{a_{11}}x_n^{(r)} \\ x_2^{(r+1)} = \dfrac{b_2}{a_{22}} - \dfrac{a_{21}}{a_{22}}x_1^{(r)} - \cdots - \dfrac{a_{2n}}{a_{22}}x_n^{(r)} \\ \vdots \quad\quad \vdots \quad\quad \vdots \quad\quad\quad \vdots \\ x_n^{(r+1)} = \dfrac{b_n}{a_{nn}} - \dfrac{a_{n1}}{a_{nn}}x_1^{(r)} - \cdots - \dfrac{a_{n(n-1)}}{a_{nn}}x_{n-1}^{(r)} \end{array}\right\}$$

Here we can observe that no element of $x_i^{(r+1)}$ replaces $x_i^{(r)}$ entirely for the next cycle of computation.

In Gauss-Seidel method, the corresponding elements of $x_i^{(r+1)}$ replaces those of $x_i^{(r)}$ as soon as they become available.

Hence, it is called the method of successive displacements. For illustration consider

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}\right\}$$

In Gauss-Seidel iteration, the (r + 1)th approximation or iteration is computed from:

$$\left.\begin{array}{l} x_1^{(r+1)} = \dfrac{b_1}{a_{11}} - \dfrac{a_{12}}{a_{11}}x_2^{(r)} - \cdots - \dfrac{a_{1n}}{a_{11}}x_n^{(r)} \\ x_2^{(r+1)} = \dfrac{b_2}{a_{22}} - \dfrac{a_{21}}{a_{22}}x_1^{(r+1)} - \cdots - \dfrac{a_{2n}}{a_{22}}x_n^{(r)} \\ \vdots \quad\quad \vdots \quad\quad \vdots \quad\quad\quad \vdots \\ x_n^{(r+1)} = \dfrac{b_n}{a_{nn}} - \dfrac{a_{n1}}{a_{nn}}x_1^{(r+1)} - \cdots - \dfrac{a_{n(n-1)}}{a_{nn}}x_{n-1}^{(r+1)} \end{array}\right\}$$

Thus, the general procedure can be written in the following compact form

$$x_i^{(r+1)} = \frac{b_i}{a_{ii}} - \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(r+1)} - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(r)} \qquad \text{for all } i = 1, 2, ..., n \text{ and } r = 1, 2, ...$$

To describe system in the first equation, we substitute the r-th approximation into the right-hand side and denote the result by $x_1^{(r+1)}$. In the second equation, we substitute $(x_1^{(r+1)}, x_3^{(r)}, ..., x_n^{(r)})$ and denote the result by $x_2^{(r+1)}$

In the third equation, we substitute $(x_1^{(r+1)}, x_2^{(r+1)}, x_4^{(r)}, ..., x_n^{(r)})$ and denote the result by $x_3^{(r+1)}$, and so on. This process is continued till we arrive at the desired result. For illustration, we consider the following example :

Note

The difference between jacobi's method and gauss Seidel method is that in jacobi's method the approximation calculated are used in the next iteration for next approximation but in Gauss-seidel method the new approximation calculated is instantly replaced by the previous one.

Example

Find the solution of the following system of equations using Gauss-Seidel method and perform the first five iterations:

$$\begin{aligned} 4x_1 - x_2 - x_3 &= 2 \\ -x_1 + 4x_2 - x_4 &= 2 \\ -x_1 + 4x_3 - x_4 &= 1 \\ -x_2 - x_3 + 4x_4 &= 1 \end{aligned}$$

Solution

The given system of equations can be rewritten as

$$\left.\begin{aligned} x_1 &= 0.5 + 0.25x_2 + 0.25x_3 \\ x_2 &= 0.5 + 0.25x_1 + 0.25x_4 \\ x_3 &= 0.25 + 0.25x_1 + 0.25x_4 \\ x_4 &= 0.25 + 0.25x_2 + 0.25x_3 \end{aligned}\right\}$$

Taking $x_2 = x_3 = x_4 = 0$ on the right-hand side of the first equation of the system , we get $x_1^{(1)} = 0.5.$ Taking $x_3 = x_4 = 0$ and the current value of x_1, we get from the 2nd equation of the system

$$x_2^{(1)} = 0.5 + (0.25)(0.5) + 0 = 0.625$$

Further, we take *x4* = 0 and the current value of *x1* we obtain from the third equation of the system

$$\begin{aligned} x_3^{(1)} &= 0.25 + (0.25)(0.5) + 0 \\ &= 0.375 \end{aligned}$$

Now, using the current values of *x2* and *x3* the fourth equation of system gives

$$x_4^{(1)} = 0.25 + (0.25)(0.625)$$
$$+(0.25)(0.375) = 0.5$$

The Gauss-Seidel iterations for the given set of equations can be written as

$$x_1^{(r+1)} = 0.5 + 0.25x_2^{(r)} + 0.25x_3^{(r)}$$
$$x_2^{(r+1)} = 0.5 + 0.25x_1^{(r+1)} + 0.25x_4^{(r)}$$
$$x_3^{(r+1)} = 0.25 + 0.25x_1^{(r+1)} + 0.25x_4^{(r)}$$
$$x_4^{(r+1)} = 0.25 + 0.25x_2^{(r+1)} + 0.25x_3^{(r+1)}$$

Now, by Gauss-Seidel procedure, the 2nd and subsequent approximations can be obtained and the sequence of the first five approximations are tabulated as below:

	Variables			
Iteration number r	x1	x2	x3	x4
1	0.5	0.625	0.375	0.5
2	0.75	0.8125	0.5625	0.59375
3	0.84375	0.85938	0.60938	0.61719
4	0.86719	0.87110	0.62110	0.62305
5	0.87305	0.87402	0.62402	0.62451

Example

Solve the system by Gauss-Seidel iterative method

$$8x - 3y + 2z = 20$$
$$4x + 11y - z = 33$$
$$6x + 3y + 12z = 35$$

(Perform only four iterations)

Solution

Consider the given system as

$8x-3y+2z=20$

$4x+11y-z=33$

$6x+3y+12z=35$

the system is diagonally dominant

$x=\frac{1}{8}[20+3y-2z]$

$y=\frac{1}{11}[33-4x+z]$

$z=\frac{1}{12}[35-6x-3y]$

we start with an initial aproximation $x_0=y_0=z_0=0$

substituting these

first iteration

$x_1=\frac{1}{8}[20+3(0)-2(0)]=2.5$

$y_1=\frac{1}{11}[33-4(2.5)+0]=2.0909091$

$z_1=\frac{1}{12}[35-6(2.5)-3(2.0909091)]=1.1439394$

Second iteration

$x_2=\frac{1}{8}[20+3y_1-z_1]=\frac{1}{8}[20+3(2.0909091)-2(1.1439394)]=2.9981061$

$y_2=\frac{1}{11}[33-4x_2+z_1]=\frac{1}{11}[33-4(2.9981061)+1.1439394]=2.0137741$

$z_2=\frac{1}{12}[35-6x_2-3y_2]=\frac{1}{12}[35-6(2.9981061)-3(2.0137741)]=0.9141701$

third iteration

$x_3=\frac{1}{8}[20+3(2.0137741)-2(0.9141701)]=3.0266228$

$y_3=\frac{1}{11}[33-4(3.0266228)+0.9141701]=1.9825163$

$z_3=\frac{1}{12}[35-6(3.0266228)-3(1.9825163)]=0.9077262$

fourth iteration

$$x_4 = \frac{1}{8}\left[20 + 3(1.9825163) - 2(0.9077262)\right] = 3.0165121$$

$$y_4 = \frac{1}{11}\left[33 - 4(3.0165121) + 0.9077262\right] = 1.9856071$$

$$z_4 = \frac{1}{12}\left[35 - 6(3.0165121) - 3(1.9856071)\right] = 0.8319128$$

Example

Solve the system by suing Gauss-seidel iteration method

$$28x + 4y - z = 32$$
$$x + 3y + 10z = 24$$
$$2x + 17y + 4z = 35$$

Solution

$$28x + 4y - z = 32$$
$$x + 3y + 10z = 24$$
$$2x + 17y + 4z = 35$$

the given system is diagonally dominant so we will make it diagonaaly dominant by iterchanaginhg the equations

$$28x + 4y - z = 32$$
$$2x + 17y + 4z = 35$$
$$x + 3y + 10z = 24$$

hence we can apply Gauss – Seidel method
from the above equations

$$x = \frac{1}{28}[32 - 4y + z]$$

$$y = \frac{1}{17}[35 - 2x - 4z]$$

$$z = \frac{1}{10}[24 - x - 3y]$$

First approximation

putting $y = z = 0$

$$x_1 = \frac{1}{28}[32] = 1.1428571$$

puting $x = 1.1428571 \quad , z = 0$

$$y_1 = \frac{1}{17}[35 - 2(1.1428571) - 4(0)] = 1.9243697$$

putting $x = 1.1428571 \; , \; y = 1.9243697$

$$z_1 = \frac{1}{10}[24 - 1.1428571 - 3(1.9243697)] = 1.7084034$$

Second iteration

$$x_2 = \frac{1}{28}[32 - 4(1.9243697) + 1.7084034] = 0.9289615$$

$$y_2 = \frac{1}{17}[35 - 2(0.9289615) - 4(1.7084034)] = 1.5475567$$

$$z_2 = \frac{1}{10}[24 - 0.9289615 - 3(1.5475567)] = 1.8408368$$

third iteration

$$x_3 = \frac{1}{28}[32 - 4(1.5475567) + 1.8428368] = 0.9875932$$

$$y_3 = \frac{1}{17}[35 - 2(0.9875932) - 4(1.8428368)] = 1.5090274$$

$$z_3 = \frac{1}{10}[24 - 0.9875932 - 3(1.5090274)] = 1.8485325$$

fourth iteration

$$x_4 = \frac{1}{28}[32 - 4(1.5090274) + 1.8485325] = 0.9933008$$

$$y_4 = \frac{1}{17}[35 - 2(0.9933008) - 4(1.8428368)] = 1.5070158$$

$$z_4 = \frac{1}{10}[24 - 0.9933008 - 3(1.5070158)] = 1.8485652$$

Example

Using Gauss-Seidel iteration method, solve the system of the equation.

$$10x - 2y - z - w = 3$$

$$-2x + 10y - z - w = 15$$

$$-x - y + 10z - 2w = 27$$

$$-x - y - 2z + 10w = -9$$

(Perform only four iterations)

Solution

$$10x-2y-z-w=3$$
$$-2x+10y-z-w=15$$
$$-x-y+10z-2w=27$$
$$-x-y-2z+10w=-9$$

it is diagonally dominant and we may write eqaution as

$$x=\frac{1}{10}[3+2y+z+w]$$
$$y=\frac{1}{10}[15+2x+z+w]$$
$$z=\frac{1}{10}[27+x+y+2w]$$
$$w=\frac{1}{10}[-9+x+y+2z]$$

first approximation

putting $y=z=w=0$ *on RHS of* (1) *, we get*

$$x_1=0.3$$
$$y_1=\frac{1}{10}[15+2(0.3)]=1.56$$

putting $x=0.3, y=1.56$ *and* $w=0$

$$z_1=\frac{1}{10}[27+0.3+1.56]=2.886$$

putting $x=0.3, y=1.56$ *and* $z=2.886$

$$w_1=\frac{1}{10}[-9+0.3+1.56+2(2.886)]=-0.1368$$

second iteration

$$x_2=\frac{1}{10}[3+2(1.56)+2.886-0.1368]=0.88692$$
$$y_2=\frac{1}{10}[15+2(0.88692)+2.886-0.1368]=1.952304$$
$$z_2=\frac{1}{10}[27+0.88692+1.952304+2(-0.1368)]=2.9565624$$
$$w_2=\frac{1}{10}[-9+0.88692+1.952304+2(2.9565624)]=-0.0247651$$

third iteration

$$x_3=\frac{1}{10}[3+2(1.952304)+2.9565624-0.0.0247651]=0.9836405$$

$$y_3 = \frac{1}{10}[15 + 2(0.9836405) + 2.9565624 - 0.0247651] = 1.9899087$$

$$z_3 = \frac{1}{10}[27 + 0.9836405 + 1.9899087 + 2(-0.0247651)] = 2.9924019$$

$$w_3 = \frac{1}{10}[-9 + 0.983405 + 1.9899087 + 2(2.9924019)] = -0.0041647$$

fourth iteration

$$x_4 = \frac{1}{10}[3 + 2(1.9899087) + 2.9924019 - 0.0041647] = 0.9968054$$

$$y_4 = \frac{1}{10}[15 + 2(0.9968054) + 2.9924019 - 0.0041647] = 1.9981848$$

$$z_4 = \frac{1}{10}[27 + 0.9968054 + 1.9981848 + 2(-0.0041647)] = 2.9986661$$

$$w_4 = \frac{1}{10}[-9 + 0.9968054 + 1.9981848 + 2(2.9986661)] = -0.0007677$$

Note

When to stop the iterative processes ,we stop the iterative process when we get the required accuracy means if your are asked that find the accurate up to four places of decimal then we will simply perform up to that iteration after which we will get the required accuracy. If we calculate the root of the equation and its consecutive values are 1.895326125, 1.916366125, 1.919356325, 1.919326355, 1.919327145, 1.919327128

Here the accuracy up to seven places of decimal is achieved so if you are asked to acquire the accuracy up to six places of decimal then we will stop here .

But in the solved examples only some iteration are carried out and accuracy is not considered here.

Lecture 11

Solution of Linear System of Equations and Matrix Inversion

Relaxation Method
This is also an iterative method and is due to Southwell.To explain the details, consider again the system of equations

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad \vdots \qquad\qquad \vdots \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}\right\}$$

Let

$$X^{(p)} = (x_1^{(p)}, x_2^{(p)}, ..., x_n^{(p)})^T$$

be the solution vector obtained iteratively after p-th iteration. If $R_i^{(p)}$ denotes the residual of the i-th equation of system given above , that is of $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ defined by

$$R_i^{(p)} = b_i - a_{i1}x_1^{(p)} - a_{i2}x_2^{(p)} - \cdots - a_{in}x_n^{(p)}$$

we can improve the solution vector successively by reducing the largest residual to zero at that iteration. This is the basic idea of relaxation method.

To achieve the fast convergence of the procedure, we take all terms to one side and then reorder the equations so that the largest negative coefficients in the equations appear on the diagonal.

Now, if at any iteration, R_i is the largest residual in magnitude, then we give an increment to x_i; a_{ii} being the coefficient of *xi*

$$dx_i = \frac{R_i}{a_{ii}}$$

In other words, we change x_i. to $(x_i + dx_i)$ to relax R_i that is to reduce R_i to zero.

Example

Solve the system of equations

$$6x_1 - 3x_2 + x_3 = 11$$
$$2x_1 + x_2 - 8x_3 = -15$$
$$x_1 - 7x_2 + x_3 = 10$$

by the relaxation method, starting with the vector (0, 0, 0).

Solution

At first, we transfer all the terms to the right-hand side and reorder the equations, so that the largest coefficients in the equations appear on the diagonal.
Thus, we get

$$\left.\begin{aligned} 0 &= 11 - 6x_1 + 3x_2 - x_3 \\ 0 &= 10 - x_1 + 7x_2 - x_3 \\ 0 &= -15 - 2x_1 - x_2 + 8x_3 \end{aligned}\right\}$$

after interchanging the 2nd and 3rd equations.

Starting with the initial solution vector (0, 0, 0), that is taking $x_1 = x_2 = x_3 = 0,$

we find the residuals $R_1 = 11,\ R_2 = 10,\ R_3 = -15$

of which the largest residual in magnitude is R3, i.e. the 3rd equation has more error and needs immediate attention for improvement.

Thus, we introduce a change, *dx3*in *x*3 which is obtained from the formula

$$dx_3 = -\frac{R_3}{a_{33}} = \frac{15}{8} = 1.875$$

Similarly, we find the new residuals of large magnitude and relax it to zero, and so on.
We shall continue this process, until all the residuals are zero or very small.

Iteration	**Residuals**			**Maximum**	**Difference**	**Variables**		
number	**R1**	**R2**	**R3**	R_i	dx_i	***x*1**	***x*2**	***x*3**
0	**11**	**10**	**-15**	**-15**	**1.875**	**0**	**0**	**0**
1	**9.125**	**8.125**	**0**	**9.125**	**1.5288**	**0**	**0**	**1.875**
2	**0.0478**	**6.5962**	**-3.0576**	**6.5962**	**-0.9423**	**1.5288**	**0**	**1.875**

Iteration		Residuals		Maximum	Difference		Variables	
number	R1	R2	R3	R_i	dx_i	$x1$	$x2$	$x3$
0	11	10	-15	-15	15/8 =1.875	0	0	0
1	9.125	8.125	0	9.125	-9.125/(-6) =1.5288	0	0	1.875
2	0.0478	6.5962	-3.0576	6.5962	-6.5962/7 =-0.9423	1.5288	0	1.875
3	-2.8747	0.0001	-2.1153	-2.8747	2.8747/(-6) =-0.4791	1.0497	-0.9423	1.875
4	-0.0031	0.4792	-1.1571	-1.1571	1.1571/8 =0.1446	1.0497	-0.9423	1.875

Iteration		Residuals		Maximum	Difference		Variables	
number	R1	R2	R3	R_i	dx_i	$x1$	$x2$	$x3$
5	-0.1447	0.3346	0.0003	0.3346	-.3346/7 =-0.0478	1.0497	-0.9423	2.0196
6	0.2881	0.0000	0.0475	0.2881	-.2881/(-6) =0.0480	1.0497	-0.9901	2.0196
7	-0.0001	0.048	0.1435	0.1435	=-0.0179	1.0017	-0.9901	2.0196
8	0.0178	0.0659	0.0003	-	-	1.0017	-0.9901	2.0017

At this stage, we observe that all the residuals *R1*, *R2* and *R3* are small enough and therefore we may take the corresponding values of *xi* at this iteration as the solution. Hence, the numerical solution is given by

$x_1 = 1.0017, \quad x_2 = -0.9901, \quad x_3 = 2.0017,$

The exact solution is

$x_1 = 1.0, \quad x_2 = -1.0, \quad x_3 = 2.0$

Example

Solve by relaxation method, the equation

$$10x - 2y - 2z = 6$$
$$-x - 10y - 2z = 7$$
$$-x - y + 10z = 8$$

Solution

The residual r_1, r_2, r_3 are given by

$$r_1 = 6 - 10x + 2y + 2z$$
$$r_2 = 7 + x - 10y + 2z$$
$$r_3 = 8 + x + y - 10z$$

The operation table is as follows

x	y	z	r1	r2	r3	
1	0	0	-10	1	1	L1
0	1	0	2	-10	1	L2
0	0	1	2	2	-10	L3

The relaxation table is as follows

x	y	z	r1	r2	r3	
0	0	0	6	7	8	L4
0	0	1	8	9	-2	L5=L4+L3
0	1	0	10	-1	-1	L6=L5+L2
1	0	0	0	0	0	L7=L6+L1

Explanation

(1) In L4 ,the largest residual is 8.to reduce it, To reduce it ,we give an increment of

$$\frac{8}{c_3} = \frac{8}{10} = 0.8 \cong 1$$

the resulting residulas are obtained by

$L_4 + (1)L_3, i.e \; line \; L_5$

(2) In line L_5 the largest residual is 9

Increment= $\frac{9}{b_2} = \frac{9}{10} = 0.9 \cong 1$

The resulting residuals $(= L_6) = L_5 + 1.L_2$

(3) In line L_6 ,the largest residual is 10

Increment = $\frac{10}{a_1} = \frac{10}{10} \cong 1$

The resulting residuals $(= L_6) = L_5 + 1.L_2$

Exact solution is arrived and it is x=1,y=1,z=1

Example

Solve the system by relaxation method, the equations

$$9x - y + 2z = 7$$
$$x + 10y - 2z = 15$$
$$2x - 2y - 13z = -17$$

Solution

The residuals r_1, r_2, r_3 are given by

$$9x - y + 2z = 9$$
$$x + 10y - 2z = 15$$
$$2x - 2y - 13z = -17$$

here

$$r_1 = 9 - 9x + y - 2z$$
$$r_2 = 15 - x - 10y + 2z$$
$$r_3 = -17 - 2x + 2y + 13z$$

Operation table

x	y	z	r1	r2	r3
1	0	0	-9	-1	-2
0	1	0	1	-10	2
0	0	1	-2	2	13

Relaxation table is

x	y	z	r1	r2	r3
0	0	0	9	15	-17
0	0	1	7	17	-4
0	1	0	8	7	-2
0.89	0	0	-0.01	6.11	-3.78
0	0.61	0	0.6	0.01	-2.56
0	0	0.19	0.22	0.39	-0.09
0	0.039	0	0.259	0	-0.012
0.028	0	0	0.007	-0.028	-0.068
0	0	0.00523	-0.00346	-1.01754	-0.00001

Then x=0.89+0.028=0.918;y=1+0.61+0.039=1.694
And z=1+0.19+0.00523=1.19523
Now substituting the values of x,y,z in (1) ,we get
r1=9-9(0.918)+1.649-2(1.19523)=-0.00346
r2=15-0.918-10(1.649)+2(1.19523)=-0.1754
r3=-17-2(0.918) +2(1.649) +13(1.19523) =-0.00001
Which is agreement with the final residuals.

Lecture 12

Norms of Vectors and Matrices, Matrix Norms and Distances

PPT's slides are available in VULMS/downloads

Lecture 13

Error Bounds and Iterative Refinement

PPT's slides are available in VULMS/downloads

Lecture 14

Eigenvalues and Eigenvectors

In this lecture we will discuss linear equations of the form $Ax = x$ and, more generally, equations of the form $Ax = \lambda x$, where λ is a scalar. Such equations arise in a wide variety of important applications and will be a recurring theme in the rest of this course.

Fixed Points:

A fixed point of an $n \times n$ matrix $\boldsymbol{A}$ is a vector $\boldsymbol{x}$ in $\boldsymbol{R}^{\boldsymbol{n}}$ such that $\boldsymbol{Ax} = \boldsymbol{x}$. Every square matrix $\boldsymbol{A}$ has at least one fixed point, namely $\boldsymbol{x} = \boldsymbol{0}$. We call this the trivial fixed point of $\boldsymbol{A}$.

The general procedure for finding the fixed points of a matrix $\boldsymbol{A}$ is to rewrite the equation $\boldsymbol{Ax} = \boldsymbol{x}$ as $\boldsymbol{Ax} = \boldsymbol{Ix}$ or, alternatively, as

$$(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \qquad (1)$$

Since this can be viewed as a homogeneous linear system of n equations in n unknowns with coefficient matrix $\boldsymbol{I} - \boldsymbol{A}$, we see that the set of fixed points of an $n \times n$ matrix is a subspace of $\boldsymbol{R}^{\boldsymbol{n}}$ that can be obtained by solving (1).

The following theorem will be useful for ascertaining the nontrivial fixed points of a matrix.

Theorem 1:

If $\boldsymbol{A}$ is an n x n matrix, then the following statements are equivalent.

(a) $\boldsymbol{A}$ has nontrivial fixed points.

(b) $\boldsymbol{I} - \boldsymbol{A}$ is singular.

(c) $\det(\boldsymbol{I} - \boldsymbol{A}) = 0$.

Example 1:

In each part, determine whether the matrix has nontrivial fixed points; and, if so, graph the subspace of fixed points in an *xy*-coordinate system.

$$(a) \quad \boldsymbol{A} = \begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix} \qquad (b) \quad \boldsymbol{A} = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}$$

Solution:

(a) The matrix has only the trivial fixed point since.

$$(I - A) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & -6 \\ -1 & -1 \end{pmatrix}$$

$$\det(I - A) = \det\begin{pmatrix} -2 & -6 \\ -1 & -1 \end{pmatrix} = (-1)(-2) - (-1)(-6) = -4 \neq 0$$

(b) The matrix has nontrivial fixed points since

$$(I-A)=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}-\begin{pmatrix}0 & 2\\ 0 & 1\end{pmatrix}=\begin{pmatrix}1 & -2\\ 0 & 0\end{pmatrix}$$

$$\det(I-A)=\det\begin{pmatrix}1 & -2\\ 0 & 0\end{pmatrix}=0$$

The fixed points $\boldsymbol{x}=(x, y)$ are the solutions of the linear system $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$, which we can express in component form as

$$\begin{bmatrix}1 & -2\\ 0 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

A general solution of this system is

$$x = 2t,\ y = t \qquad (2)$$

which are parametric equations of the line $y=\frac{1}{2}x$. It follows from the corresponding vector form of this line that the fixed points are

$$\boldsymbol{x}=\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}2t\\ t\end{bmatrix}=t\begin{bmatrix}2\\ 1\end{bmatrix} \qquad (3)$$

As a check, $\boldsymbol{Ax}=\begin{bmatrix}0 & 2\\ 0 & 1\end{bmatrix}\begin{bmatrix}2t\\ t\end{bmatrix}=\begin{bmatrix}2t\\ t\end{bmatrix}=\boldsymbol{x}$

so every vector of form (3) is a fixed point of $\boldsymbol{A}$.

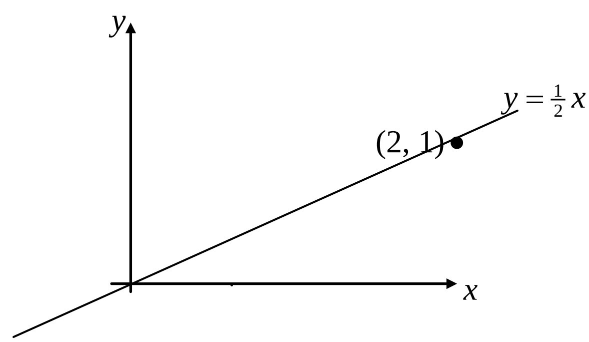


Figure 1

Eigenvalues and Eigenvectors:

In a fixed point problem one looks for nonzero vectors that satisfy the equation $\boldsymbol{Ax}=\boldsymbol{x}$. One might also consider whether there are nonzero vectors that satisfy such equations as

$$\boldsymbol{Ax} = 2\boldsymbol{x},\ \boldsymbol{Ax} = -3\boldsymbol{x},\ \boldsymbol{Ax} = \sqrt{2}\boldsymbol{x}$$

or, more generally, equations of the form $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ in which λ is a scalar.

Definition: If $\boldsymbol{A}$ is an n x n matrix, then a scalar λ is called an **eigenvalue** of $\boldsymbol{A}$ if there is a nonzero vector $\boldsymbol{x}$ such that $\boldsymbol{Ax}=\lambda\boldsymbol{x}$. If λ is an eigenvalue of $\boldsymbol{A}$, then every nonzero vector $\boldsymbol{x}$ such that $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ is called an **eigenvector** of $\boldsymbol{A}$ corresponding to λ.

Example 2:

Let $\boldsymbol{A}=\begin{bmatrix}1 & 6\\ 5 & 2\end{bmatrix}$, $\boldsymbol{u}=\begin{bmatrix}6\\ -5\end{bmatrix}$ and $\boldsymbol{v}=\begin{bmatrix}3\\ -2\end{bmatrix}$. Are $\boldsymbol{u}$ and $\boldsymbol{v}$ eigenvectors of $\boldsymbol{A}$?

Solution:

$$\boldsymbol{Au}=\begin{bmatrix}1 & 6\\ 5 & 2\end{bmatrix}\begin{bmatrix}6\\ -5\end{bmatrix}=\begin{bmatrix}-24\\ 20\end{bmatrix}=-4\begin{bmatrix}6\\ -5\end{bmatrix}=-4\boldsymbol{u}$$

$$\boldsymbol{Av}=\begin{bmatrix}1 & 6\\ 5 & 2\end{bmatrix}\begin{bmatrix}3\\ -2\end{bmatrix}=\begin{bmatrix}-9\\ 11\end{bmatrix}\neq\lambda\begin{bmatrix}3\\ -2\end{bmatrix}$$

Thus $\boldsymbol{u}$ is an eigenvector corresponding to an eigenvalue – 4, but $\boldsymbol{v}$ is not an eigenvector of $\boldsymbol{A}$, because $\boldsymbol{Av}$ is not a multiple of $\boldsymbol{v}$.

Example 3:

Show that 7 is an eigenvalue of $\boldsymbol{A}=\begin{bmatrix}1 & 6\\ 5 & 2\end{bmatrix}$, find the corresponding eigenvectors.

Solution:

The scalar 7 is an eigenvalue of $\boldsymbol{A}$ if and only if the equation

$$\boldsymbol{Ax} = 7\boldsymbol{x} \qquad \text{(A)}$$

has a nontrivial solution. But (A) is equivalent to $\boldsymbol{Ax} - 7\boldsymbol{x} = \boldsymbol{0}$, or

$$(\boldsymbol{A} - 7\boldsymbol{I})\,\boldsymbol{x} = \boldsymbol{0} \qquad \text{(B)}$$

To solve this homogeneous equation, form the matrix

$$\boldsymbol{A}-7\boldsymbol{I}=\begin{bmatrix}1 & 6\\ 5 & 2\end{bmatrix}-\begin{bmatrix}7 & 0\\ 0 & 7\end{bmatrix}=\begin{bmatrix}-6 & 6\\ 5 & -5\end{bmatrix}$$

The columns of $\boldsymbol{A} - 7\boldsymbol{I}$ are obviously linearly dependent, so (B) has nontrivial solutions. Thus 7 is an eigenvalue of $\boldsymbol{A}$. To find the corresponding eigenvectors, use row operations:

$$\begin{bmatrix}-6 & 6 & 0\\ 5 & -5 & 0\end{bmatrix}$$

$$\sim\begin{bmatrix}1 & -1 & 0\\ 5 & -5 & 0\end{bmatrix}(-1R_1-R_2)$$

$$\sim\begin{bmatrix}1 & -1 & 0\\ 0 & 0 & 0\end{bmatrix}(R_2-5R_1)$$

The general solution has the form $x_2\begin{bmatrix}1\\ 1\end{bmatrix}$. Each vector of this form with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$.

The equivalence of equations (A) and (B) obviously holds for any λ in place of $\lambda = 7$. Thus λ is an eigenvalue of $\boldsymbol{A}$ if and only if the equation

$$(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0} \qquad \text{(C)}$$

has a nontrivial solution.

Eigen space:

The set of all solutions of $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ is just the null space of the matrix $\boldsymbol{A} - \lambda \boldsymbol{I}$. So this set is a subspace of $\boldsymbol{R}^n$ and is called the eigenspace of $\boldsymbol{A}$ corresponding to λ. The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ.

Example 3 shows that for matrix $\boldsymbol{A}$ in Example 2, the eigenspace corresponding to $\lambda = 7$ consists of all multiples of (1, 1), which is the line through (1, 1) and the origin. From Example 2, one can check that the eigenspace corresponding to $\lambda = -4$ is the line through (6, -5). These eigenspaces are shown in Fig. 1, along with eigenvectors (1, 1) and (3/2, -5/4) and the geometric action of the transformation $\boldsymbol{x} \rightarrow \boldsymbol{Ax}$ on each eigenspace.

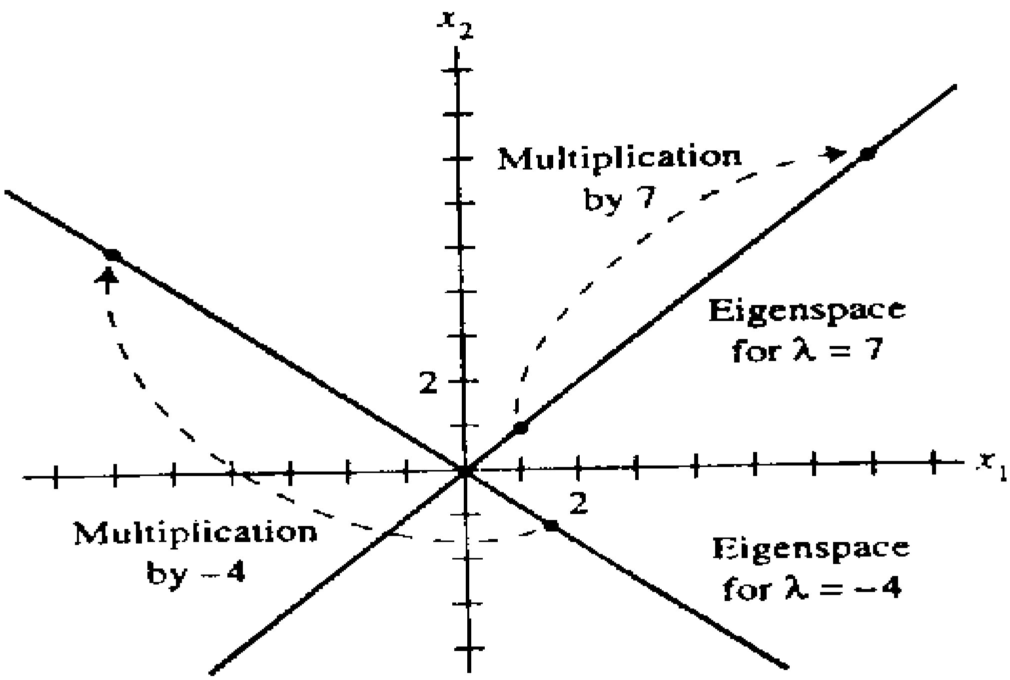


FIGURE 2 Eigenspaces for $\lambda = -4$ and $\lambda = 7$.

Example 4: Let $\boldsymbol{A} = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$.

Find a basis for the corresponding eigenspace where eigen value of matrix is 2.

Solution: Form $\boldsymbol{A}-2\boldsymbol{I}=\begin{bmatrix}4 & -1 & 6\\ 2 & 1 & 6\\ 2 & -1 & 8\end{bmatrix}-\begin{bmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{bmatrix}=\begin{bmatrix}2 & -1 & 6\\ 2 & -1 & 6\\ 2 & -1 & 6\end{bmatrix}$ and row reduce the augmented matrix for $(\boldsymbol{A}-2\boldsymbol{I})\,\boldsymbol{x}=\boldsymbol{0}$:

$$\begin{bmatrix}2 & -1 & 6 & 0\\ 2 & -1 & 6 & 0\\ 2 & -1 & 6 & 0\end{bmatrix}$$

$$\sim\begin{bmatrix}2 & -1 & 6 & 0\\ 0 & 0 & 0 & 0\\ 2 & -1 & 6 & 0\end{bmatrix}R_2-R_1$$

$$\sim\begin{bmatrix}2 & -1 & 6 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}R_3-R_1$$

At this point we are confident that 2 is indeed an eigenvalue of $\boldsymbol{A}$ because the equation $(\boldsymbol{A}-2\boldsymbol{I})\,\boldsymbol{x}=\boldsymbol{0}$ has free variables. The general solution is

$$2x_1-x_2+6x_3=0........(a)$$

$$Let\ x_2=t,\ x_3=s\ then$$

$$2x_1=t-6s$$

$$x_1=\left(1/2\right)t-3s$$

$$then$$

$$\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}t/2-3s\\ t\\ s\end{bmatrix}=\begin{bmatrix}t/2\\ t\\ 0\end{bmatrix}+\begin{bmatrix}-3s\\ 0\\ s\end{bmatrix}=t\begin{bmatrix}1/2\\ 1\\ 0\end{bmatrix}+s\begin{bmatrix}-3\\ 0\\ 1\end{bmatrix}$$

By back substitution the general solution is

$$\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=x_2\begin{bmatrix}1/2\\ 1\\ 0\end{bmatrix}+x_3\begin{bmatrix}-3\\ 0\\ 1\end{bmatrix},\ x_2 \text{ and } x_3 \text{ free}$$

The eigenspace, shown in Fig. 2, is a two – dimensional subspace of $\boldsymbol{R}^3$. A basis is

$$\left\{\begin{bmatrix}1\\ 2\\ 0\end{bmatrix},\begin{bmatrix}-3\\ 0\\ 1\end{bmatrix}\right\} \text{ is a basis.}$$

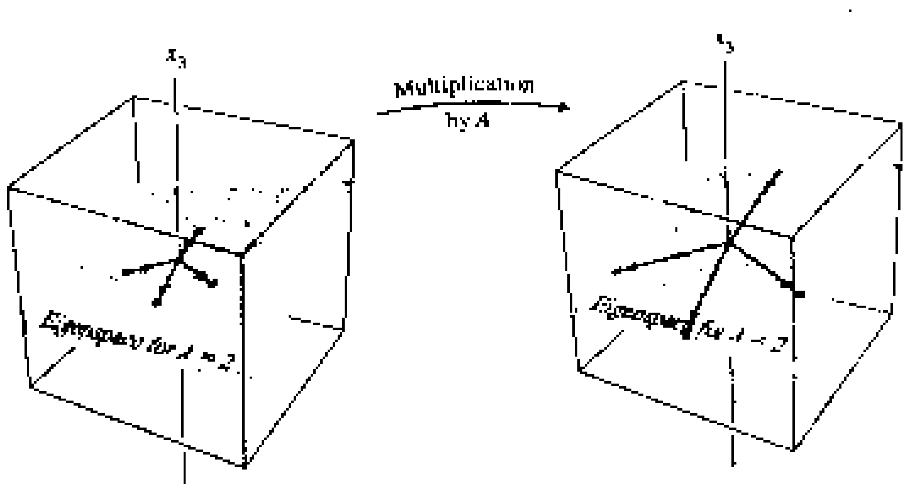


FIGURE 3 A acts as a dilation on the eigenspace.

The most direct way of finding the eigenvalues of an $n \times n$ matrix $\boldsymbol{A}$ is to rewrite the equation $Ax = \lambda x$ as $Ax = \lambda Ix$, or equivalently, as

$$(\lambda I - A)x = 0 \qquad (4)$$

and then try to determine those values of λ, if any, for which this system has nontrivial solutions. Since (4) have nontrivial solutions if and only if the coefficient matrix $\lambda \boldsymbol{I} - \boldsymbol{A}$ is singular, we see that the eigenvalues of $\boldsymbol{A}$ are the solutions of the equation

$$\det(\lambda I - A) = 0 \qquad (5)$$

Equation (5) is known as characteristic equation. Also, if λ is an eigenvalue of *A*, then equation (4) has a nonzero solution space, which we call the eigenspace of *A* corresponding to λ. It is the nonzero vectors in the eigenspace of A corresponding to λ that are the eigenvectors of A corresponding to λ.
The above discussion is summarized by the following theorem.

Theorem: If $\boldsymbol{A}$ is an $n \times n$ matrix and λ is a scalar, then the following statements are equivalent.
(i) λ is an eigenvalue of $\boldsymbol{A}$.
(ii) λ is a solution of the equation $\det(\lambda I - A) = 0$.
(iii) The linear system $(\lambda I - A)x = 0$ has nontrivial solutions.

Eigenvalues of Triangular Matrices: If $\boldsymbol{A}$ is an $n \times n$ triangular matrix with diagonal entries a_{11}, a_{22}, …, a_{nn}, then $\lambda I - A$ is a triangular matrix with diagonal entries $\lambda - a_{11}, \lambda - a_{22}, \ldots, \lambda - a_{nn}$. Thus, the characteristic polynomial of $\boldsymbol{A}$ is

$$\det(\lambda I - A) = (\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

which implies that the eigenvalues of $\boldsymbol{A}$ are

$$\lambda_1 = a_{11},\ \lambda_2 = a_{22},\ \ldots,\ \lambda_n = a_{nn}$$

Thus, we have the following theorem.

Theorem: If A is a triangular matrix (upper triangular, lower triangular, or diagonal) then the eigenvalues of A are the entries on the main diagonal of A.

Example 5: (Eigenvalues of Triangular Matrices)

By inspection, the characteristic polynomial of the matrix $\boldsymbol{A} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ -1 & -\frac{2}{3} & 0 & 0 \\ 7 & \frac{5}{8} & 6 & 0 \\ \frac{4}{9} & -4 & 3 & 6 \end{bmatrix}$ is $p(\lambda) = (\lambda - \frac{1}{2})(\lambda + \frac{2}{3})(\lambda - 6)^2$. So the distinct eigenvalues of $\boldsymbol{A}$ are $\lambda = \frac{1}{2}$, $\lambda = -\frac{2}{3}$, and $\lambda = 6$.

Eigenvalues of Powers of a Matrix: Once the eigenvalues and eigenvectors of a matrix $\boldsymbol{A}$ are found, it is a simple matter to find the eigenvalues and eigenvectors of any positive integer power of $\boldsymbol{A}$. For example, if λ is an eigenvalue of $\boldsymbol{A}$ and $\boldsymbol{x}$ is a corresponding eigenvector, then $A^2x = A(Ax) = A(\lambda x) = \lambda(Ax) = \lambda(\lambda x) = \lambda^2 x$, which shows that λ^2 is an eigenvalue of $\boldsymbol{A^2}$ and $\boldsymbol{x}$ is a corresponding eigenvector. In general we have the following result.

Theorem: If λ is an eigenvalue of a matrix A and x is a corresponding eigenvector, and if k is any positive integer, then λ^k is an eigenvalue of A^k and x is a corresponding eigenvector.
Some problems that use this theorem are given in the exercises.

A Unifying Theorem: Since λ is an eigenvalue of a square matrix $\boldsymbol{A}$ if and only if there is a nonzero vector $\boldsymbol{x}$ such that $\boldsymbol{Ax} = \lambda \boldsymbol{x}$, it follows that $\lambda = 0$ is an eigenvalue of $\boldsymbol{A}$ if and only if there is a nonzero vector $\boldsymbol{x}$ such that $\boldsymbol{Ax} = \boldsymbol{0}$. However, this is true if and only if $\det(\boldsymbol{A}) = 0$, so we list the following

Theorem: If $\boldsymbol{A}$ is an $n \times n$ matrix, then the following statements are equivalent.
(a) The reduced row echelon form of $\boldsymbol{A}$ is $\boldsymbol{I_n}$.
(b) $\boldsymbol{A}$ is expressible as a product of elementary matrices.
(c) $\boldsymbol{A}$ is invertible.
(d) $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution.
(e) $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for every vector $\boldsymbol{b}$ in $\boldsymbol{R^n}$.
(f) $\boldsymbol{Ax} = \boldsymbol{b}$ has exactly one solution for every vector $\boldsymbol{b}$ in $\boldsymbol{R}^n$
(g) The column vectors of $\boldsymbol{A}$ are linearly independent.
(h) The row vectors of $\boldsymbol{A}$ are linearly independent.
(i) $\det(\boldsymbol{A}) \neq 0$.
(j) $\lambda = 0$ is not an eigenvalue of $\boldsymbol{A}$.

Example 6:

(1) Is 5 an eigenvalue of $\boldsymbol{A} = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}$?

(2) If $\boldsymbol{x}$ is an eigenvector for $\boldsymbol{A}$ corresponding to λ, what is $\boldsymbol{A^3x}$?

Solution:

(1) The number 5 is an eigenvalue of $\boldsymbol{A}$ if and only if the equation $(\boldsymbol{A}-\lambda \boldsymbol{I})\,\boldsymbol{x} = \boldsymbol{0}$ has a nontrivial solution. Form

$$\boldsymbol{A} - 5\boldsymbol{I} = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix} - \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 1 \\ 3 & -5 & 5 \\ 2 & 2 & 1 \end{bmatrix}$$

and row reduce the augmented matrix:

$$\begin{bmatrix} 1 & -3 & 1 & 0 \\ 3 & -5 & 5 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix} R_2 - 3R_1$$

$$\sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 8 & -1 & 0 \end{bmatrix} R_3 - 2R_1$$

$$\sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix} R_3 - 2R_2$$

At this point it is clear that the homogeneous system has no free variables. Thus $\boldsymbol{A} - 5\boldsymbol{I}$ is an invertible matrix, which means that 5 is not an eigenvalue of $\boldsymbol{A}$.

(2). If $\boldsymbol{x}$ is an eigenvector for $\boldsymbol{A}$ corresponding to λ, then $\boldsymbol{Ax} = \lambda\boldsymbol{x}$ and so

$$\boldsymbol{A^2 x} = \boldsymbol{A}(\lambda \boldsymbol{x}) = \lambda \boldsymbol{Ax} = \lambda^2 \boldsymbol{x}$$

Again $\boldsymbol{A^3 x} = \boldsymbol{A}(\boldsymbol{A^2 x}) = \boldsymbol{A}(\lambda^2 \boldsymbol{x}) = \lambda^2 \boldsymbol{Ax} = \lambda^3 \boldsymbol{x}$. The general pattern, $\boldsymbol{A^k x} = \lambda^k \boldsymbol{x}$, is proved by induction.

Exercises:

1. Is $\lambda = 2$ an eigenvalue of $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$?

2. Is $\begin{bmatrix} -1+\sqrt{2} \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$? If so, find the eigenvalue.

3. Is $\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 3 & 7 & 9 \\ -4 & -5 & 1 \\ 2 & 4 & 4 \end{bmatrix}$? If so, find the eigenvalue.

4. Is $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 3 & 6 & 7 \\ 3 & 3 & 7 \\ 5 & 6 & 5 \end{bmatrix}$? If so, find the eigenvalue.

5. Is $\lambda = 4$ an eigenvalue of $\begin{bmatrix} 3 & 0 & -1 \\ 2 & 3 & 1 \\ -3 & 4 & 5 \end{bmatrix}$? If so, find one corresponding eigenvector.

6. Is $\lambda = 3$ an eigenvalue of $\begin{bmatrix} 1 & 2 & 2 \\ 3 & -2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$? If so, find one corresponding eigenvector.

In exercises 7 to 12, find a basis for the eigenspace corresponding to each listed eigenvalue.

7. $\boldsymbol{A} = \begin{bmatrix} 4 & -2 \\ -3 & 9 \end{bmatrix}, \lambda = 10$

8. $\boldsymbol{A} = \begin{bmatrix} 7 & 4 \\ -3 & -1 \end{bmatrix}, \lambda = 1.5$

9. $\boldsymbol{A} = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \lambda = 1, 2, 3$

10. $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & -3 & 0 \\ 4 & -13 & 1 \end{bmatrix}, \lambda = -2$

11. $\boldsymbol{A} = \begin{bmatrix} 4 & 2 & 3 \\ -1 & 1 & -3 \\ 2 & 4 & 9 \end{bmatrix}, \lambda = 3$

12. $\boldsymbol{A} = \begin{bmatrix} 3 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \lambda = 4$

Find the eigenvalues of the matrices in Exercises 13 and 14.

13. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 5 \\ 0 & 0 & -1 \end{bmatrix}$

14. $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -3 \end{bmatrix}$

15. For $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, find one eigenvalue, with no calculation. Justify your answer.

16. Without calculation, find one eigenvalue and two linearly independent vectors of $\boldsymbol{A} = \begin{bmatrix} 5 & 5 & 5 \\ 5 & 5 & 5 \\ 5 & 5 & 5 \end{bmatrix}$. Justify your answer.

Lecture 15

The Characteristic Equation

The Characteristic equation contains useful information about the eigenvalues of a square matrix **A.** It is defined as

$$\det(A-\lambda I)=0,$$

Where λ is the eigenvalue and I is the identity matrix. We will solve the Characteristic equation (also called the characteristic polynomial) to work out the eigenvalues of the given square matrix **A**.

Example 1: Find the eigenvalues of $\boldsymbol{A}=\begin{bmatrix}2 & 3\\ 3 & -6\end{bmatrix}$.

Solution: In order to find the eigenvalues of the given matrix, we must solve the matrix equation

$$(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=0$$

for the scalar λ such that it has a nontrivial solution (since the matrix is non singular). By the Invertible Matrix Theorem, this problem is equivalent to finding all λ such that the matrix $\boldsymbol{A}-\lambda\boldsymbol{I}$ is not invertible, where

$$\boldsymbol{A}-\lambda\boldsymbol{I}=\begin{bmatrix}2 & 3\\ 3 & -6\end{bmatrix}-\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}=\begin{bmatrix}2-\lambda & 3\\ 3 & -6-\lambda\end{bmatrix}.$$

By definition, this matrix $\boldsymbol{A}-\lambda\boldsymbol{I}$ fails to be invertible precisely when its determinant is zero. Thus, the eigenvalues of $\boldsymbol{A}$ are the solutions of the equation

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=\det\begin{bmatrix}2-\lambda & 3\\ 3 & -6-\lambda\end{bmatrix}=0.$$

Recall that $\det\begin{bmatrix}a & b\\ c & d\end{bmatrix}=ad-bc$

So $\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(2-\lambda)(-6-\lambda)-(3)(3)$

$$=-12+6\lambda-2\lambda+\lambda^2-9$$
$$=\lambda^2+4\lambda-21$$
$$\lambda^2+4\lambda-21=0,$$
$$(\lambda-3)(\lambda+7)=0,$$

so the eigenvalues of $\boldsymbol{A}$ are 3 and –7.

Example 2: Compute det $\boldsymbol{A}$ for $A=\begin{bmatrix}1 & 5 & 0\\ 2 & 4 & -1\\ 0 & -2 & 0\end{bmatrix}$

Solution:

Firstly, we will reduce the given matrix in echelon form by applying elementary row operations

$$by\, R_2 - 2R_1$$

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix},$$

$$by\, R_2 \leftrightarrow R_3$$

$$\sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix},$$

$$by\, R_3 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

which is an upper triangular matrix. Therefore,

$$\det A = (1)(-2)(-1)$$
$$= 2.$$

Theorem 1: Properties of Determinants

Let A and B be two matrices of order n then

(a) A is invertible if and only if $\det A \neq 0$.
(b) $\det AB = (\det A)(\det B)$.
(c) $\det A^T = \det A$.
(d) If A is triangular, then det A is the product of the entries on the main diagonal of A.
(e) A row replacement operation on A does not change the determinant.
(f) A row interchange changes the sign of the determinant.
(g) A row scaling also scales the determinant by the same scalar factor.

Note: These Properties will be helpful in using the characteristic equation to find eigenvalues of a matrix **A**.

Example 3: (a) Find the eigenvalues and corresponding eigenvectors of the matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$$

(b) Graph the eigenspaces of $\boldsymbol{A}$ in an xy-coordinate system.

Solution: (a) The eigenvalues will be worked out by solving the characteristic equation of $\boldsymbol{A}$. Since

$$\lambda \boldsymbol{I} - \boldsymbol{A} = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{bmatrix}.$$

The characteristic equation $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ becomes

$$\begin{vmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{vmatrix} = 0.$$

Expanding and simplifying the determinant, it yields

$$\lambda^2 - 3\lambda - 10 = 0,$$

or

$$(\lambda + 2)(\lambda - 5) = 0. \tag{1}$$

Thus, the eigenvalues of $\boldsymbol{A}$ are $\lambda = -2$ and $\lambda = 5$.

Now, to work out the eigenspaces corresponding to these eigenvalues, we will solve the system

$$\begin{bmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2}$$

for $\lambda = -2$ and $\lambda = 5$. Here are the computations for the two cases.

(i) Case $\lambda = -2$

In this case Eq. (2) becomes

$$\begin{bmatrix} -3 & -3 \\ -4 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which can be written as

$$-3x - 3y = 0,$$
$$-4x - 4y = 0 \Rightarrow x = -y.$$

In parametric form,

$$x = -t, y = t \ . \tag{3}$$

Thus, the eigenvectors corresponding to $\lambda = -2$ are the nonzero vectors of the form

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -t \\ t \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \tag{4}$$

It can be verified as

$$\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} -t \\ t \end{bmatrix} = \begin{bmatrix} 2t \\ -2t \end{bmatrix} = -2 \begin{bmatrix} -t \\ t \end{bmatrix} = -2\boldsymbol{x}$$

Thus,

$$Ax = \lambda x$$

(ii) Case $\lambda = 5$

In this case Eq. (2) becomes

$$\begin{bmatrix} 4 & -3 \\ -4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which can be written as

$$4x - 3y = 0$$
$$-4x + 3y = 0 \Rightarrow x = \frac{3}{4}y.$$

In parametric form,

$$x = \frac{3}{4}t, \;\; y = t. \tag{5}$$

Thus, the eigenvectors corresponding to $\lambda = 5$ are the nonzero vectors of the form

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{3}{4}t \\ t \end{bmatrix} = t\begin{bmatrix} \frac{3}{4} \\ 1 \end{bmatrix}. \tag{6}$$

It can be verified as

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} \frac{3}{4}t \\ t \end{bmatrix} = \begin{bmatrix} \frac{15}{4}t \\ 5t \end{bmatrix} = 5\begin{bmatrix} \frac{3}{4}t \\ t \end{bmatrix} = 5\boldsymbol{x}.$$

(b) The eigenspaces corresponding to $\lambda = -2$ and $\lambda = 5$ can be sketched from the parametric equations (3) and (5) as shown in figure 1(a).

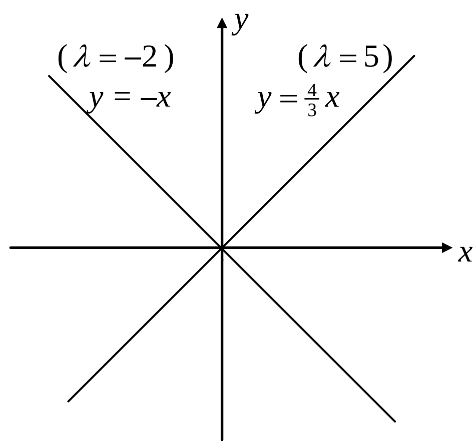


Figure 1(a)

It can also be drawn using the vector equations (4) and (6) as shown in Figure 1(b). When an eigenvector $\boldsymbol{x}$ in the eigenspace for $\lambda = 5$ is multiplied by $\boldsymbol{A}$, the resulting vector has the same direction as $\boldsymbol{x}$ but the length is increased by a factor of 5 and when an eigenvector $\boldsymbol{x}$ in the eigenspace for $\lambda = -2$ is multiplied by $\boldsymbol{A}$, the resulting vector is oppositely directed to $\boldsymbol{x}$ and the length is increased by a factor of 2. In both cases, multiplying an eigenvector by $\boldsymbol{A}$ produces a vector in the same eigenspace.

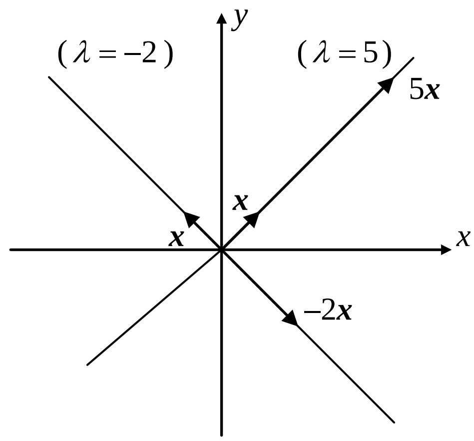


Figure 1(b)

Eigenvalues of an $n \times n$ matrix:

Eigen values of an $n\times n$ matrix can be found in the similar fashion. However, for the higher values of n, it is more convenient to work them out using various available mathematical software. Here is an example for a 3×3 matrix.

Example 4: Find the eigen values of the matrix $\boldsymbol{A}=\begin{bmatrix}0 & -1 & 0\\ 0 & 0 & 1\\ -4 & -17 & 8\end{bmatrix}$

Solution:

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \det\begin{bmatrix}\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda\end{bmatrix} - \begin{bmatrix}0 & -1 & 0\\ 0 & 0 & 1\\ -4 & -17 & 8\end{bmatrix}$$

$$= \begin{vmatrix}\lambda & 1 & 0\\ 0 & \lambda & -1\\ 4 & 17 & \lambda-8\end{vmatrix} \qquad (7)$$

$$= \lambda^3 - 8\lambda^2 + 17\lambda - 4,$$

which yields the characteristic equation

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0 \qquad (8)$$

To solve this equation, firstly, we will look for integer solutions. This can be done by using the fact that if a polynomial equation has integer coefficients, then its integer solutions, if any, must be divisors of the constant term of the given polynomial. Thus, the only possible integer solutions of Eq.(8) are the divisors of –4, namely $\pm1, \pm2,$ and ± 4. Substituting these values successively into Eq. (8) yields that λ =4 is an integer solution. This implies that $\lambda - 4$ is a factor of Eq.(7), Thus, dividing the polynomial by $\lambda - 4$ and rewriting Eq.(8), we get

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1) = 0.$$

Now, the remaining solutions of the characteristic equation satisfy the quadratic equation

$$\lambda^2 - 4\lambda + 1 = 0.$$

Solving the above equation by the quadratic formula, we get the eigenvalues of ***A*** as
$\lambda = 4, \quad \lambda = 2+\sqrt{3}, \quad \lambda = 2-\sqrt{3}$

Example 5 Find the characteristic equation of $A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Solution: Clearly, the given matrix is an upper triangular matrix. Forming $\boldsymbol{A} - \lambda \boldsymbol{I}$, we get

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \det \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix}$$

Now using the fact that determinant of a triangular matrix is equal to product of its diagonal elements, the characteristic equation becomes

$$(5-\lambda)^2(3-\lambda)(1-\lambda) = 0.$$

Expanding the product, we can also write it as

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0.$$

Here, the eigenvalue 5 is said to have multiplicity 2 because (λ - 5) occurs two times as a factor of the characteristic polynomial. In general, the (**algebraic**) multiplicity of an eigenvalue λ is its multiplicity as a root of the characteristic equation.

Note:
From the above mentioned examples, it can be easily observed that if A is an $n \times n$ matrix, then det (A – λ I) is a polynomial of degree n called the characteristic polynomial of A.

Example 6 The characteristic polynomial of a 6×6 matrix is $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Find the eigenvalues and their multiplicities.

Solution:

In order to find the eigenvalues, we will factorize the polynomial as

$$\begin{aligned} &\lambda^6 - 4\lambda^5 - 12\lambda^4 \\ &= \lambda^4(\lambda^2 - 4\lambda - 12) \\ &= \lambda^4(\lambda - 6)(\lambda + 2) \end{aligned}$$

The eigenvalues are 0 (multiplicity 4), 6 (multiplicity 1) and – 2 (multiplicity 1).We could also list the eigenvalues in Example 6 as 0, 0, 0, 0, 6 and –2, so that the eigenvalues are repeated according to their multiplicities

Activity:

Work out the eigenvalues and eigenvectors for the following square matrix.

$$A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}.$$

Similarity:

Let A and B be two n x n matrices, A is said to be similar to B if there exist an invertible matrix P such that

$$P^{-1}AP = B,$$

or equivalently,

$$A = PBP^{-1}.$$

Replacing Q by P^{-1}, we have

$$Q^{-1}BQ = A.$$

So B is also similar to A. Thus, we can say that A and B are similar.

Similarity transformation:

The act of changing A into $P^{-1}AP$ is called a similarity transformation.

The following theorem illustrates use of the characteristic polynomial and it provides the foundation for several iterative methods that approximate eigenvalues.

Theorem 2:

If n x n matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

Proof: If $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{AP}$, then

$$B - \lambda I = P^{-1}AP - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

Using the multiplicative property (b) of Theorem 1, we compute

$$\begin{aligned} \det(B - \lambda I) &= \det\left[P^{-1}(A - \lambda I)P\right] \\ &= \det(P^{-1}).\det(A - \lambda I).\det(P) \qquad \text{(A)} \end{aligned}$$

Since

$$\begin{aligned} \det(\boldsymbol{P}^{-1}).\det(\boldsymbol{P}) &= \det(\boldsymbol{P}^{-1}\boldsymbol{P}) \\ &= \det \boldsymbol{I} \\ &= 1, \end{aligned}$$

Eq. (A) implies that

$$\det(B - \lambda I) = \det(A - \lambda I).$$

Hence, both the matrices have the same characteristic polynomials and therefore, same eigenvalues.

Note: It must be clear that Similarity and row equivalence are two different concepts. (If $\boldsymbol{A}$ is row equivalent to $\boldsymbol{B}$, then $\boldsymbol{B} = \boldsymbol{EA}$ for some invertible matrix $\boldsymbol{E}$.) Row operations on a matrix usually change its eigenvalues.

Application to Dynamical Systems:

Dynamical system is the one which evolves with the passage of time. Eigenvalues and eigenvectors play a vital role in the evaluation of a dynamical system. Let's consider an example of a dynamical system.

Example 7: Let $\boldsymbol{A} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. Analyze the long term behavior of the dynamical system defined by $\boldsymbol{x_{k+1}} = \boldsymbol{Ax_k}$ $(k = 0, 1, 2, \ldots)$, with $\boldsymbol{x_0} = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$.

Solution: The first step is to find the eigenvalues of $\boldsymbol{A}$ and a basis for each eigenspace. The characteristic equation for $\boldsymbol{A}$ is

$$0 = \det(A - \lambda I)$$

$$0 = \det\begin{bmatrix} 0.95\text{-}\lambda & 0.03 \\ 0.05 & 0.97\text{-}\lambda \end{bmatrix} = (.95\text{-}\lambda)(.97\text{-}\lambda)\text{-}(.03)(.05)$$

$$= \lambda^2 \text{-} 1.92\lambda + .92$$

By the quadratic formula

$$\lambda = \frac{1.92 \pm \sqrt{(1.92)^2 \text{-} 4(.92)}}{2} = \frac{1.92 \pm \sqrt{.0064}}{2} = \frac{1.92 \pm .08}{2} = 1 \text{ or } .92$$

Firstly, the eigenvectors will be found as given below.

$$Ax = \lambda x,$$

$$(Ax - \lambda x) = 0,$$

$$(A - \lambda I)x = 0.$$

For $\lambda = 1$

$$\left[\begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0,$$

$$\begin{pmatrix} -0.05 & 0.03 \\ 0.05 & -0.03 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0,$$

which can be written as

$$-0.05x_1 + 0.03x_2 = 0$$

$$0.05x_1 - 0.03x_2 = 0 \Rightarrow x_1 = \frac{0.03}{0.05}x_2 \;\; or \;\; x_1 = \frac{3}{5}x_2.$$

In parametric form, it becomes

$$x_1 = \frac{3}{5}t \;\; and \;\; x_2 = t.$$

For $\lambda = 0.92$

$$\left[\begin{pmatrix}0.95 & 0.03\\0.05 & 0.97\end{pmatrix}-\begin{pmatrix}0.92 & 0\\0 & 0.92\end{pmatrix}\right]\begin{pmatrix}x_1\\x_2\end{pmatrix}=0,$$

$$\begin{pmatrix}0.03 & 0.03\\0.05 & 0.05\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=0.$$

It can be written as

$$0.03x_1+0.03x_2=0$$

$$0.05x_1+0.05x_2=0 \Rightarrow x_1=-x_2$$

In parametric form, it becomes

$$x_1=t \ and \ x_2=-t$$

Thus, the eigenvectors corresponding to $\lambda = 1$ and $\lambda = .92$ are multiples of

$\boldsymbol{v_1}=\begin{bmatrix}3\\5\end{bmatrix}$ and $\boldsymbol{v_2}=\begin{bmatrix}1\\-1\end{bmatrix}$ respectively.

The next step is to write the given $\boldsymbol{x_0}$ in terms of $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$. This can be done because $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is obviously a basis for $\boldsymbol{R^2}$. So there exists weights c_1 and c_2 such that

$$\boldsymbol{x_0}=c_1\boldsymbol{v_1}+c_2\boldsymbol{v_2}=[\boldsymbol{v_1} \quad \boldsymbol{v_2}]\begin{bmatrix}c_1\\c_2\end{bmatrix} \tag{1}$$

In fact, $$\begin{bmatrix}c_1\\c_2\end{bmatrix}=[\boldsymbol{v_1} \quad \boldsymbol{v_2}]^{-1}x_0=\begin{bmatrix}3 & 1\\5 & -1\end{bmatrix}^{-1}\begin{bmatrix}.60\\.40\end{bmatrix}$$

Here,

$$\begin{bmatrix}3 & 1\\5 & -1\end{bmatrix}^{-1}=\frac{1}{\begin{vmatrix}3 & 1\\5 & -1\end{vmatrix}}Adj\begin{bmatrix}3 & 1\\5 & -1\end{bmatrix}=-\frac{1}{8}\begin{bmatrix}-1 & -1\\-5 & 3\end{bmatrix}$$

Therefore,

$$\begin{bmatrix}c_1\\c_2\end{bmatrix}=\frac{1}{-8}\begin{bmatrix}-1 & -1\\-5 & 3\end{bmatrix}\begin{bmatrix}.60\\.40\end{bmatrix}=\begin{bmatrix}.125\\.225\end{bmatrix} \tag{2}$$

Because $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ in Eq.(1) are eigenvectors of $\boldsymbol{A}$, with $\boldsymbol{Av_1} = \boldsymbol{v_1}$ and $\boldsymbol{Av_2} = (.92)\,\boldsymbol{v_2}$, $\boldsymbol{x_k}$ can be computed as

$$\boldsymbol{x_1} = \boldsymbol{Ax_0} = \boldsymbol{c_1Av_1} + \boldsymbol{c_2Av_2} \quad \text{(Using linearity of } \boldsymbol{x} \to \boldsymbol{Ax}\text{)}$$
$$= c_1\boldsymbol{v_1} + c_2\,(.92)\boldsymbol{v_2} \quad (\boldsymbol{v_1} \text{ and } \boldsymbol{v_2} \text{ are eigenvectors.})$$
$$\boldsymbol{x_2} = \boldsymbol{Ax_1} = c_1\boldsymbol{Av_1} + c_2\,(.92)\boldsymbol{Av_2} = c_1\boldsymbol{v_1} + c_2\,(.92)^2\,\boldsymbol{v_2}.$$

Continuing in the same way, we get the general equation as

$$\boldsymbol{x_k} = c_1\boldsymbol{v_1} + c_2(.92)^k\boldsymbol{v_2} \qquad (k = 0, 1, 2, \ldots).$$

Using c_1 and c_2 from Eq.(2),

$$\boldsymbol{x_k}=.125\begin{bmatrix}3\\5\end{bmatrix}+.225(.92)^k\begin{bmatrix}1\\-1\end{bmatrix} \qquad (k=0,1,2,\ldots) \tag{3}$$

This explicit formula for $\boldsymbol{x_k}$ gives the solution of the difference equation $\boldsymbol{x_{k+1}} = \boldsymbol{Ax_k}$.

As $k \to \infty$, $(.92)^k$ tends to zero and $\boldsymbol{x_k}$ tends to $\begin{bmatrix} .375 \\ .625 \end{bmatrix} = .125\boldsymbol{v_1}$.

Example 8: Find the characteristic equation and eigenvalues of $\boldsymbol{A} = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$.

Solution: The characteristic equation is

$$0 = \det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \det\begin{bmatrix} 1-\lambda & -4 \\ 4 & 2-\lambda \end{bmatrix}$$
$$= (1-\lambda)(2-\lambda) - (-4)(4),$$
$$= \lambda^2 - 3\lambda + 18,$$

which is a quadratic equation whose roots are given as

$$\lambda = \frac{3 \pm \sqrt{(-3)^2 - 4(18)}}{2}$$
$$= \frac{3 \pm \sqrt{-63}}{2}$$

Thus, we see that the characteristic equation has no real roots, so $\boldsymbol{A}$ has no real eigenvalues. $\boldsymbol{A}$ is acting on the real vector space $\boldsymbol{R^2}$ and there is no non-zero vector $\boldsymbol{v}$ in $\boldsymbol{R^2}$ such that $\boldsymbol{Av} = \lambda \boldsymbol{v}$ for some scalar λ.

Exercises:

Find the characteristic polynomial and the eigenvalues of matrices in exercises 1 to 12.

1. $\begin{bmatrix} 3 & -2 \\ 1 & -1 \end{bmatrix}$

2. $\begin{bmatrix} 5 & -3 \\ -4 & 3 \end{bmatrix}$

3. $\begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$

4. $\begin{bmatrix} 3 & -4 \\ 4 & 8 \end{bmatrix}$

5. $\begin{bmatrix} 5 & 3 \\ -4 & 4 \end{bmatrix}$

6. $\begin{bmatrix} 7 & -2 \\ 2 & 3 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & -1 \\ 0 & 6 & 0 \end{bmatrix}$

8. $\begin{bmatrix} 0 & 3 & 1 \\ 3 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}$

9. $\begin{bmatrix} 4 & 0 & 0 \\ 5 & 3 & 2 \\ -2 & 0 & 2 \end{bmatrix}$

10. $\begin{bmatrix} -1 & 0 & 1 \\ -3 & 4 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

11. $\begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$

12. $\begin{bmatrix} 5 & -2 & 3 \\ 0 & 1 & 0 \\ 6 & 7 & -2 \end{bmatrix}$

For the matrices in exercises 13 to 15, list the eigenvalues, repeated according to their multiplicities.

13. $\begin{bmatrix} 4 & -7 & 0 & 2 \\ 0 & 3 & -4 & 6 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 5 & 0 & 0 & 0 \\ 8 & -4 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 1 & -5 & 2 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 & 0 \\ 3 & 8 & 0 & 0 & 0 \\ 0 & -7 & 2 & 1 & 0 \\ -4 & 1 & 9 & -2 & 3 \end{bmatrix}$

16. It can be shown that the algebraic multiplicity of an eigenvalue λ is always greater than or equal to the dimension of the eigenspace corresponding to λ. Find h in the matrix $\boldsymbol{A}$ below such that the eigenspace for $\lambda=5$ is two-dimensional:

$$\boldsymbol{A} = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & h & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Lecture 16

Diagonalization

Diagonalization is a process of transforming a vector A to the form $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$ for some invertible matrix **P** and a diagonal matrix **D**. In this lecture, the factorization enables us to compute $\boldsymbol{A^k}$ quickly for large values of k which is a fundamental idea in several applications of linear algebra. Later, the factorization will be used to analyze (and decouple) dynamical systems.
The "D" in the factorization stands for diagonal. Powers of such a D are trivial to compute.

Example 1: If $\boldsymbol{D} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, then $\boldsymbol{D^2} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$ and

$$\boldsymbol{D^3} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

In general, $\boldsymbol{D^k} = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}$ for $k \geq 1$

The next example shows that if $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$ for some invertible $\boldsymbol{P}$ and diagonal $\boldsymbol{D}$, then it is quite easy to compute $\boldsymbol{A^k}$.

Example 2: Let $\boldsymbol{A} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a formula for $\boldsymbol{A^k}$, given that $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$, where

$$\boldsymbol{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \text{and } \boldsymbol{D} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

Solution: The standard formula for the inverse of a 2×2 matrix yields

$$\boldsymbol{P^{-1}} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

By associative property of matrix multiplication,

$$\boldsymbol{A^2} = (\boldsymbol{PDP^{-1}})(\boldsymbol{PDP^{-1}}) = \boldsymbol{PD}\underbrace{(\boldsymbol{P^{-1}P})}_{1}\boldsymbol{DP^{-1}} = PDIDP^{-1} = \boldsymbol{PDDP^{-1}}$$

where I is the identity matrix.

$$= \boldsymbol{PD^2P^{-1}} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Again, $\boldsymbol{A^3} = (\boldsymbol{PDP^{-1}})\boldsymbol{A^2} = (\boldsymbol{PD}\underbrace{\boldsymbol{P^{-1}})\boldsymbol{P}}_{1}\boldsymbol{D^2P^{-1}} = \boldsymbol{PDD^2P^{-1}} = \boldsymbol{PD^3P^{-1}}$

Thus, in general, for $k \geq 1$, $\boldsymbol{A^k} = \boldsymbol{PD^kP^{-1}} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$,

$$= \begin{bmatrix} 5^k & 3^k \\ -5^k & -2.3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix},$$

$$= \begin{bmatrix} 2.5^k - 3^k & 5^k - 3^k \\ 2.3^k - 2.5^k & 2.3^k - 5^k \end{bmatrix}.$$

Activity:

Work out C^4, given that $C = PDP^{-1}$ where

$$P = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix},\ D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

Remarks:

A square matrix ***A*** is said to be **diagonalizable** if ***A*** is similar to a diagonal matrix, that is, if $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$ for some invertible matrix ***P*** and some diagonal matrix ***D***. The next theorem gives a characterization of diagonalizable matrices and tells how to construct a suitable factorization.

Theorem 1: The Diagonalization Theorem

An n x n matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$, with ***D*** a diagonal matrix, if and only if the columns of ***P*** are n linearly independent eigenvectors of ***A***. In this case, the diagonal entries of ***D*** are eigenvalues of ***A*** that correspond, respectively, to the eigenvectors in ***P***.

In other words, ***A*** is diagonalizable if and only if there are enough eigenvectors to form a basis of $\boldsymbol{R^n}$. We call such a basis an eigenvector basis.

Proof: First, observe that if ***P*** is any $n \times n$ matrix with columns $\boldsymbol{v_1}, \dots, \boldsymbol{v_n}$ and if ***D*** is any diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$ then

$$\boldsymbol{AP} = \boldsymbol{A}\left[\boldsymbol{v_1} \quad \boldsymbol{v_2} \quad \dots \quad \boldsymbol{v_n}\right] = \left[\boldsymbol{Av_1} \quad \boldsymbol{Av_2} \quad \dots \quad \boldsymbol{Av_n}\right], \tag{1}$$

while
$$\boldsymbol{PD} = \boldsymbol{P}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \left[\lambda_1\boldsymbol{v_1} \quad \lambda_2\boldsymbol{v_2} \quad \cdots \quad \lambda_n\boldsymbol{v_n}\right] \tag{2}$$

Suppose now that ***A*** is diagonalizable and $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$. Then right-multiplying this relation by ***P***, we have ***AP*** = ***PD***. In this case, (1) and (2) imply that

$$\left[\boldsymbol{Av_1} \quad \boldsymbol{Av_2} \quad \cdots \quad \boldsymbol{Av_n}\right] = \left[\lambda_1\boldsymbol{v_1} \quad \lambda_2\boldsymbol{v_2} \quad \cdots \quad \lambda_n\boldsymbol{v_n}\right] \tag{3}$$

Equating columns, we find that

$$\boldsymbol{Av_1} = \lambda_1\boldsymbol{v_1}, \boldsymbol{Av_2} = \lambda_2\boldsymbol{v_2}, \cdots, \boldsymbol{Av_n} = \lambda_n\boldsymbol{v_n} \tag{4}$$

Since ***P*** is invertible, its columns $\boldsymbol{v_1}, \dots, v_n$ must be linearly independent. Also, since these columns are nonzero, Eq.(4) shows that $\lambda_1, \dots, \lambda_n$ are eigenvalues and $\boldsymbol{v_1}, \dots, \boldsymbol{v_n}$ are corresponding eigenvectors. This argument proves the "only if" parts of the first and second statements along with the third statement, of the theorem.

Finally, given any n eigenvectors $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$ use them to construct the columns of $\boldsymbol{P}$ and use corresponding eigenvalues $\lambda_1, \ldots\ldots, \lambda_n$ to construct $\boldsymbol{D}$. By Eqs. (1) – (3). $\boldsymbol{AP} = \boldsymbol{PD}$. This is true without any condition on the eigenvectors. If, in fact, the eigenvectors are linearly independent, then $\boldsymbol{P}$ is invertible (by the Invertible Matrix Theorem), and $\boldsymbol{AP} = \boldsymbol{PD}$ implies that $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$.

Diagonalizing Matrices

Example 3: Diagonalize the following matrix, if possible $\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$

Solution: To diagonalize the given matrix, we need to find an invertible matrix **P** and a diagonal matrix **D** such that $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$ which can be done in following four steps.

Step 1: Find the eigenvalues of $\boldsymbol{A}$.

The characteristic equation becomes

$$0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4$$
$$= -(\lambda - 1)(\lambda + 2)^2$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$ (multiplicity 2)

Step 2: Find three linearly independent eigenvectors of $\boldsymbol{A}$. Since $\boldsymbol{A}$ is a 3×3 matrix and we have obtained three eigen values, we need three eigen vectors. This is the critical step. If it fails, then above Theorem says that $\boldsymbol{A}$ cannot be diagonalized. Now we will produce basis for these eigen values.

Basis vector for $\lambda = 1$:

$(A - \lambda I)x = 0$

$$\begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

After applying few row operations on the matrix $(A - \lambda I)$, we get

$$\begin{bmatrix} 0 & 1 & 1 \\ 3 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which can be written as

$x_2 + x_3 = 0$

$3x_1 + 3x_2 = 0$

In parametric form, it becomes

$x_1 = t,\ x_2 = -t,\ x_3 = t$

Thus, the basis vector for $\lambda = 1$ is $v_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

Basis vector for $\lambda =$ -2

$$(A - \lambda I)x = 0$$

$$\begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which can be written as

$$3x_1 + 3x_2 + 3x_3 = 0$$
$$-3x_1 - 3x_2 - 3x_3 = 0$$
$$3x_1 + 3x_2 + 3x_3 = 0$$

In parametric form, it becomes

$x_1 = -s - t, \;\; x_2 = s, \;\; x_3 = t$

Now,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -s-t \\ s \\ t \end{bmatrix} = \begin{bmatrix} -s \\ s \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ 0 \\ t \end{bmatrix},$$

$$= s\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

$$= x_2\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Thus, the basis for $\lambda = -2$ is $v_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $v_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

We can check that $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ is a linearly independent set.

Step 3: Check that $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ is a linearly independent set.

Construct ***P*** from the vectors in step 2. The order of the vectors is not important. Using the order chosen in step 2, form $\boldsymbol{P} = [\boldsymbol{v_1} \quad \boldsymbol{v_2} \quad \boldsymbol{v_3}] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

Step 4: Form ***D*** from the corresponding eigen values. For this purpose, the order of the eigen values must match the order chosen for the columns of ***P***. Use the eigen value λ = -2 twice, once for each of the eigenvectors corresponding to λ = -2:

$$\boldsymbol{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Now, we need to check do ***P*** and ***D*** really work. To avoid computing $\boldsymbol{P}^{-1}$, simply verify that ***AP*** = ***PD***. This is equivalent to ***A*** = $\boldsymbol{PDP}^{-1}$ when ***P*** is invertible.We compute

$$\boldsymbol{AP} = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

$$\boldsymbol{PD} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

Example 4: Diagonalize the following matrix, if possible.

$$\boldsymbol{A} = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Solution: The characteristic equation of ***A*** turns out to be exactly the same as that in example 3 i.e.,

$$\begin{aligned} 0 &= \det(\boldsymbol{A} - \lambda \boldsymbol{I}) \\ &= \begin{vmatrix} 2-\lambda & 4 & 3 \\ -4 & -6-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix} \\ &= -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

The eigen values are λ = 1 and λ = -2 (multiplicity 2). However, when we look for eigen vectors, we find that each eigen space is only one – dimensional.

Basis for $\lambda = 1$: $\quad \boldsymbol{v_1} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

Basis for $\lambda = -2$: $\quad \boldsymbol{v_2} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

There are no other eigen values and every eigen vector of $\boldsymbol{A}$ is a multiple of either $\boldsymbol{v_1}$ or $\boldsymbol{v_2}$. Hence it is impossible to form a basis of $\boldsymbol{R^3}$ using eigenvectors of $\boldsymbol{A}$. By above Theorem, $\boldsymbol{A}$ is not diagonalizable.

Theorem 2: An n x n matrix with n distinct eigenvalues is diagonalizable.

The condition in Theorem 2 is sufficient but not necessary i.e., it is not necessary for an n x n matrix to have n distinct eigen values in order to be diagonalizable. Example 3 serves as a counter example of this case where the 3 x 3 matrix is diagonalizable even though it has only two distinct eigen values.

Example 5: Determine if the following matrix is diagonalizable.

$$\boldsymbol{A} = \begin{bmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{bmatrix}$$

Solution : In the light of Theorem 2, the answer is quite obvious. Since the matrix is triangular, its eigen values are obviously 5, 0, and –2.Since A is a 3 x 3 matrix with three distinct eigen values, $\boldsymbol{A}$ is diagonalizable.

Matrices Whose Eigenvalues Are Not Distinct:

If an n x n matrix $\boldsymbol{A}$ has n distinct eigen values, with corresponding eigen vectors $\boldsymbol{v_1}$,..., $\boldsymbol{v_n}$ and if $\boldsymbol{P} = [\boldsymbol{v_1} \dots \boldsymbol{v_n}]$, then $\boldsymbol{P}$ is automatically invertible because its columns are linearly independent , by Theorem 2 of lecture 28. When $\boldsymbol{A}$ is diagonalizable but has fewer than n distinct eigen values, it is still possible to build $\boldsymbol{P}$ in a way that makes $\boldsymbol{P}$ automatically invertible, as shown in the next theorem.

Theorem 3: Let $\boldsymbol{A}$ be an n x n matrix whose distinct eigen values are $\lambda_1, \dots, \lambda_p$.

a. For $1 \le k \le p$, the dimension of the eigen space for λ_k is less than or equal to the multiplicity of the eigen value λ_k
b. The matrix $\boldsymbol{A}$ is diagonalizable if and only if the sum of the dimensions of the distinct eigen spaces is equal to n, and this happens if and only if the dimension of the eigen space for each of λ_k equals the multiplicity of λ_k .
c. If $\boldsymbol{A}$ is diagonalizable and $\boldsymbol{B_k}$ is basis for the eigen space corresponding to λ_k for each k, then the total collection of vectors in the sets $\boldsymbol{B_1}$, ..., $\boldsymbol{B_p}$ form an eigenvector basis for $\boldsymbol{R^n}$.

Example 6: Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix}$$

Solution: Since A is triangular matrix, the eigenvalues are 5 and –3, each with multiplicity 2. Using the method of lecture 28, we find a basis for each eigen space.

$$\text{Basis for } \lambda = 5 : v_1 = \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} \text{ and } v_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{Basis for } \lambda = -3 : v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } v_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The set $\{v_1,\ldots,v_4\}$ is linearly independent, by Theorem 3. So the matrix $P = [v_1 \ldots v_4]$ is invertible, and $A = PDP^{-1}$, where

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

Example 7:

(1) Compute A^8 where $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$

(2) Let $A = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}, v_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, and $v_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suppose you are told that v_1 and v_2 are eigenvectors of A. Use this information to diagonalize A.

(3) Let A be a 4 x 4 matrix with eigenvalues 5, 3, and -2, and suppose that you know the eigenspace for $\lambda=3$ is two-dimensional. Do you have enough information to determine if A is diagonalizable?

Solution:

Here, det $(A-\lambda I)=\lambda^2 -3\lambda+2=(\lambda-2)(\lambda-1)$.

The eigen values are 2 and 1, and corresponding eigenvectors are
$\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, and $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Next, form

$$\boldsymbol{P} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } \boldsymbol{P}^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

Since $\boldsymbol{A} = \boldsymbol{PDP}^{-1}$, $\boldsymbol{A}^8 = \boldsymbol{PD}^8\boldsymbol{P} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$

$$= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix}$$

(2)Here, $\boldsymbol{Av}_1 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1.\boldsymbol{v}_1$, and

$$\boldsymbol{Av}_2 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3.\boldsymbol{v}_2$$

Clearly, $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are eigenvectors for the eigenvalues 1 and 3 , respectively. Thus
$\boldsymbol{A} = \boldsymbol{PDP}^{-1}$, where $\boldsymbol{P} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$ and $\boldsymbol{D} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

(3) Yes $\boldsymbol{A}$ is diagonalizable. There is a basis $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ for the eigen space corresponding to λ=3. Moreover, there will be at least one eigenvector for λ=5 and one for λ=-2 say $\boldsymbol{v_3}$ and $\boldsymbol{v_4}$. Then $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_4}\}$ is linearly independent and $\boldsymbol{A}$ is diagonalizable , by Theorem 3. There can be no additional eigen vectors that are linearly independent from $\boldsymbol{v_1}$ to $\boldsymbol{v_4}$ because the vectors are all in $\boldsymbol{R}^4$ .Hence the eigenspaces for λ=5 and λ=-2 are both one–dimensional.

Exercise:

In exercises 1 and 2, let $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$ and compute $\boldsymbol{A^4}$.

1. $\boldsymbol{P} = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

2. $\boldsymbol{P} = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$

In exercises 3 and 4, use the factorization $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$ to compute $\boldsymbol{A^k}$, where k represents an arbitrary positive integer.

3. $\begin{bmatrix} a & 0 \\ 3(a-b) & b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$

4. $\begin{bmatrix} -2 & 12 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 4 \\ 1 & -3 \end{bmatrix}$

In exercises 5 and 6, the matrix $\boldsymbol{A}$ is factored in the form $\boldsymbol{PDP^{-1}}$. Use the Diagonalization Theorem to find the eigenvalues of $\boldsymbol{A}$ and a basis for each eigenspace.

5. $\begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}\begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & -3/4 \\ 1/4 & -1/2 & 1/4 \end{bmatrix}$

6. $\begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} -2 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 4 \\ -1 & 0 & -2 \end{bmatrix}$

Diagonalize the matrices in exercises 7 to 18, if possible.

7. $\begin{bmatrix} 3 & -1 \\ 1 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$

9. $\begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}$

10. $\begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$

11. $\begin{bmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & -2 & 2 \end{bmatrix}$

12. $\begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix}$

13. $\begin{bmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{bmatrix}$

14. $\begin{bmatrix} 0 & -4 & -6 \\ -1 & 0 & -3 \\ 1 & 2 & 5 \end{bmatrix}$

15. $\begin{bmatrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

16. $\begin{bmatrix} -7 & -16 & 4 \\ 6 & 13 & -2 \\ 12 & 16 & 1 \end{bmatrix}$

17. $\begin{bmatrix} 5 & -3 & 0 & 9 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

18. $\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$

Lecture 17

Inner Product

If u and v are vectors in R^n , then we regard u and v as $n\times 1$ matrices. The transpose u^t is a $1\times n$ matrix, and the matrix product $u^t v$ is a 1×1 matrix which we write as a single real number (a scalar) without brackets.
The number $u^t v$ is called the inner product of u and v. And often is written as $u.v$
This inner product is also referred to as a dot product.

If $u=\begin{bmatrix} u_1\\ u_2\\ .\\ .\\ .\\ u_n \end{bmatrix}$ and $v=\begin{bmatrix} v_1\\ v_2\\ .\\ .\\ .\\ v_n \end{bmatrix}$

Then the inner product of u and v is

$$\begin{bmatrix} u_1 & u_2 & . & . & . & u_n \end{bmatrix}\begin{bmatrix} v_1\\ v_2\\ .\\ .\\ .\\ v_n \end{bmatrix} = u_1v_1+u_2v_2+...+u_nv_n$$

Example 1

Compute $u.v$ and $v.u$ when $u=\begin{bmatrix} 2\\ -5\\ -1 \end{bmatrix}$ $and\ v=\begin{bmatrix} 3\\ 2\\ -3 \end{bmatrix}$.

Solution

$$u=\begin{bmatrix} 2\\ -5\\ -1 \end{bmatrix} and\ v=\begin{bmatrix} 3\\ 2\\ -3 \end{bmatrix}$$

$$u^t=\begin{bmatrix} 2 & -5 & -1 \end{bmatrix}$$

$$u.v=u^tv=\begin{bmatrix} 2 & -5 & -1 \end{bmatrix}\begin{bmatrix} 3\\ 2\\ -3 \end{bmatrix}=2(3)+(-5)(2)+(-1)(-3)$$

$$=6-10+3=-1$$

$v^t = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix}$

$$v.u = v^t u = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = 3(2) + (2)(-5) + (-3)(-1)$$

$$= 6 - 10 + 3 = -1$$

Theorem

Let u, v and w be vectors in R^n, and let c be a scalar. Then

a. $u.v = v.u$
b. $(u+v).w = u.w + v.w$
c. $(cu).v = c(u.v) = u.(cv)$
d. $u.u \geq 0$ and $u.u = 0$ if and only if $u = 0$

Observation

$$(c_1u_1 + c_2u_2 + \ldots.c_pu_p).w$$
$$= c_1(u_1.w) + c_2(u_2.w) + \ldots c_p(u_p.w)$$

Length or Norm

The length or Norm of v is the nonnegative scalar $\|v\|$ defined by

$$\|v\| = \sqrt{v.v} = \sqrt{v_1^{\,2} + v_1^{\,2} + \ldots + v_n^{\,2}}$$

$$\|v\|^2 = v.v$$

Note: For any scalar c, $\|cv\| = |c|\,\|v\|$

Unit vector

A vector whose length is 1 is called a unit vector. If we divide a non-zero vector v by its length $\|v\|$, we obtain a unit vector u as

$$u = \frac{v}{\|v\|}$$

The length of u is $\|u\| = \dfrac{1}{\|v\|}\|v\| = 1$

Definition

The process of creating the unit vector u from v is sometimes called normalizing v, and we say that u is in the same direction as v. In this case "u" is called the normalized vector.

Example 2

Let $v = (1, 2, 2, 0)$ in R^4. Find a unit vector u in the same direction as v.

Solution

The length of v is given by

$\|v\| = \sqrt{v.v} = \sqrt{{v_1}^2 + {v_2}^2 + {v_3}^2 + {v_4}^2}$

So,

$\|v\| = \sqrt{1^2 + 2^2 + 2^2 + 0^2} = \sqrt{1+4+4+0} = \sqrt{9} = 3$

The unit vector u in the direction of v is given as

$$u = \frac{1}{\|v\|} v = \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \\ 0 \end{bmatrix}$$

To check that $\|u\| = 1$

$$\|u\| = \sqrt{u.u} = \sqrt{(\frac{1}{3})^2 + (\frac{-2}{3})^2 + (\frac{2}{3})^2 + (0)^2} = \sqrt{\frac{1}{9} + \frac{4}{9} + \frac{4}{9} + 0} = 1$$

Example 3

Let W be the subspace of R^2 spanned by $X = (\frac{2}{3}, 1)$. Find a unit vector Z that is a basis for W.

Solution

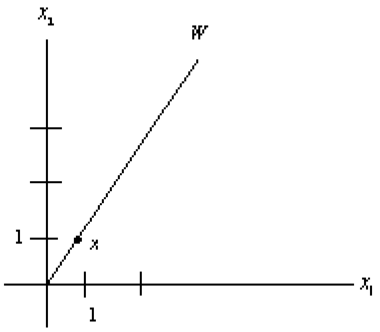
W consists of all multiples of ***x***, as in Fig. 2(a). Any nonzero vector in ***W*** is a basis for ***W***. To simplify the calculation, ***x*** is scaled to eliminate fractions. That is, multiply ***x*** by 3 to get

$$\boldsymbol{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

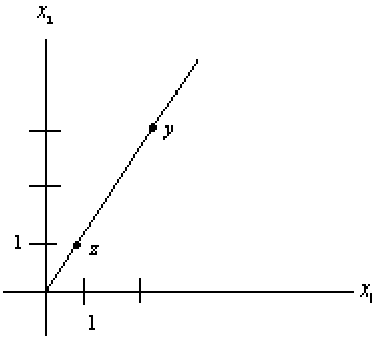
Now compute $\|\boldsymbol{y}\|^2 = 2^2 + 3^2 = 13, \|\boldsymbol{y}\| = \sqrt{13},$ and normalize ***y*** to get

$$\boldsymbol{z} = \frac{1}{\sqrt{13}}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}$$

See Fig. 2(b). Another unit vector is $(-2/\sqrt{13}, -3/\sqrt{13})$.



(a)



(b)

Figure 2 Normalizing a vector to produce a unit vector.

Definition

For u and v vectors in R^n, the distance between u and v, written as dist (u , v), is the length of the vector $u-v$. That is

$$dist(u,v)=\|u-v\|$$

Example 4

Compute the distance between the vectors u = (7, 1) and v = (3, 2)

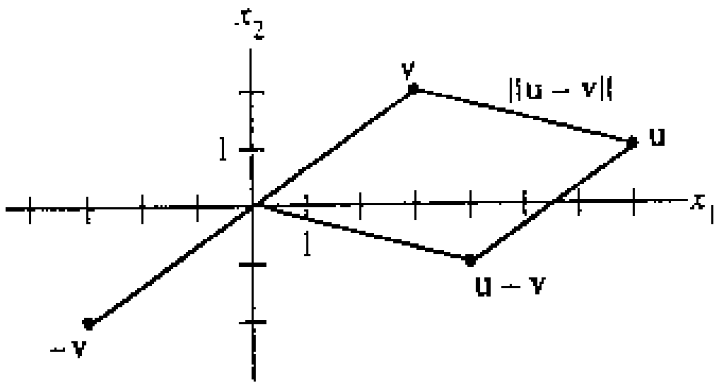
Solution

$$u-v=\begin{bmatrix}7\\1\end{bmatrix}-\begin{bmatrix}3\\2\end{bmatrix}=\begin{bmatrix}7-3\\1-2\end{bmatrix}=\begin{bmatrix}4\\-1\end{bmatrix}$$

dist (u , v) = $\|u-v\|=\sqrt{(4)^2+(-1)^2}=\sqrt{16+1}=\sqrt{17}$

Law of Parallelogram of vectors

The vectors, u , v and $u-v$ are shown in the fig. below. When the vector $u-v$ is added to v, the result is u. Notice that the parallelogram in the fig. below shows that the distance from u to v is the same as the distance of $u-v$ to o.



Example 5

If $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$, then

$$dist(u,v) = \|u - v\| = \sqrt{(u-v).(u-v)}$$

$$= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}$$

Definition

Two vectors in u and v in R^n are orthogonal (to each other) if $u.v = 0$

Note

The zero vector is orthogonal to every vector in R^n because $0^t.v = 0$ for all v in R^n.

The Pythagorean Theorem

Two vectors u and v are orthogonal if and only if $\|u + v\|^2 = \|u\|^2 + \|v\|^2$

Orthogonal Complements

The set of all vectors z that are orthogonal to w in W is called the orthogonal complement of W and is denoted by $W^{\perp}$

Example 6

Let ***W*** be a plane through the origin in $\boldsymbol{R^3}$, and let ***L*** be the line through the origin and perpendicular to ***W***. If ***z*** and ***w*** are nonzero, ***z*** is on ***L***, and ***w*** is in ***W***, then the line segment from ***0*** to ***z*** is perpendicular to the line segment from ***0*** to ***w***; that is, ***z*** . ***w*** = 0. So each vector on ***L*** is orthogonal to every ***w*** in ***W***. In fact, ***L*** consists of all vectors that are orthogonal to the ***w***'s in ***W***, and ***W*** consists of all vectors orthogonal to the ***z***'s in ***L***. That is,

$$\boldsymbol{L} = \boldsymbol{W}^{\perp} \text{ and } \boldsymbol{W} = \boldsymbol{L}^{\perp}$$

Remarks

The following two facts about $W^{\perp}$, with W a subspace of R^n, are needed later in the segment.
(1) A vector x is in $W^{\perp}$ if and only if x is orthogonal to every vector in a set that spans W.
(2) $W^{\perp}$ is a subspace of R^n.

Theorem 3

Let A be m x n matrix. Then the orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T: $(\text{Row } A)^{\perp} = \text{Nul } A,\ (\text{Col } A)^{\perp} = \text{Nul } A^T$

Proof

The row-column rule for computing Ax shows that if x is in Nul A, then x is orthogonal to each row of A (with the rows treated as vectors in R^n). Since the rows of A span the row space, x is orthogonal to Row A. Conversely, if x is orthogonal to Row A, then x is certainly orthogonal to each row of A, and hence $Ax = 0$. This proves the first statement. The second statement follows from the first by replacing A with A^T and using the fact that Col A = Row A^T.

Angles in R^2 and R^3

If u and v are nonzero vectors in either R^2 or R^3, then there is a nice connection between their inner product and the angle ϑ between the two line segments from the origin to the points identified with u and v. The formula is

$$u \cdot v = \|u\| \|v\| \cos \vartheta \qquad (2)$$

To verify this formula for vectors in R^2, consider the triangle shown in Fig. 7, with sides of length $\|u\|, \|v\|$, and $\|u - v\|$. By the law of cosines, $\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos\vartheta$
which can be rearranged to produce

$$\|u\|\|v\|\cos\vartheta = \frac{1}{2}\left[\|u\|^2 + \|v\|^2 - \|u-v\|^2\right]$$

$$= \frac{1}{2}\left[u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2\right] = u_1 v_1 + u_2 v_2 = u \cdot v$$

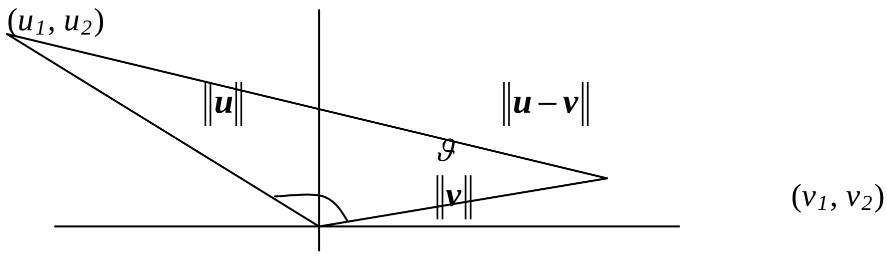


Figure 7 The angle between two vectors.

Example 7

Find the angle between the vectors $u = (1,-1,2),\quad v = (2,1,0)$

Solution

$u.v = (1)(2) + (-1)(1) + (2)(0) = 2 - 1 + 0 = 1$

And

$\|u\| = \sqrt{(1)^2 + (-1)^2 + (2)^2} = \sqrt{1+1+4} = \sqrt{6}$

$\|v\| = \sqrt{(2)^2 + (1)^2 + (0)^2} = \sqrt{4+1} = \sqrt{5}$

Angle between the two vectors is given by

$$\cos\vartheta = \frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}$$

Putting the values, we get

$$\cos\theta = \frac{1}{\sqrt{\boldsymbol{6}}\sqrt{5}} = \frac{1}{\sqrt{30}}$$

$$\cos\theta = \frac{1}{\sqrt{30}}$$

$$\theta = \cos^{-1}\frac{1}{\sqrt{30}} = 79.48^{\circ}$$

Exercises

Q.1

Compute $u.v$ and $v.u$ when $u = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$ $and\ v = \begin{bmatrix} -3 \\ 1 \\ 5 \end{bmatrix}$

Q.2

Let $v = (2,1,0,3)$ in R^4. Find a unit vector u in the direction opposite to that of v.

Q.3

Let W be the subspace of R^3 spanned by $X = (\frac{1}{2}, \frac{3}{2}, \frac{5}{2})$. Find a unit vector Z that is a basis for W.

Q.4

Compute the distance between the vectors u = (1, 5, 7) and v = (2, 3, 5).

Q.5

Find the angle between the vectors $u = (2,1,3),\quad v = (1,0,2)$.

Lecture 18
Orthogonal and Orthonormal sets

Objectives

The objectives of the lecture are to learn about:

- Orthogonal Set.
- Orthogonal Basis.
- Unique representation of a vector as a linear combination of Basis vectors.
- Orthogonal Projection.
- Decomposition of a vector into sum of two vectors.
- Orthonormal Set.
- Orthonormal Basis.
- Some examples to verify the definitions and the statements of the theorems.

Orthogonal Set

Let $S=\{u_1,u_2,...,u_p\}$ be the set of non-zero vectors in R^n, is said to be an orthogonal set if all vectors in S are mutually orthogonal. That is
$O \notin S$ and $u_i.u_j = o \; \forall \, i \neq j, i, j = 1,2,..., p.$

Example

Show that $S=\{u_1,u_2,u_3\}$ is an orthogonal set. Where

$$u_1=\begin{bmatrix}3\\1\\1\end{bmatrix}, u_2=\begin{bmatrix}-1\\2\\1\end{bmatrix} \text{ and } u_3=\begin{bmatrix}\frac{-1}{2}\\-2\\\frac{7}{2}\end{bmatrix}.$$

Solution

To show that S is orthogonal, we show that each vector in S is orthogonal to other. That is
$u_i.u_j = o \; \forall \, i \neq j, i, j = 1,2,3.$
For $i=1, j=2$

$$u_1.u_2=\begin{bmatrix}3\\1\\1\end{bmatrix}.\begin{bmatrix}-1\\2\\1\end{bmatrix}$$
$$=-3+2+1=0$$

Which implies u_1 is orthogonal to u_2.

For $i=1, j=3$

$$u_1 . u_3 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} . \begin{bmatrix} \frac{-1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix}$$

$$= \frac{-3}{2} - 2 + \frac{7}{2}$$

$$= 0.$$

Which implies u_1 is orthogonal to u_3.

For $i=2, j=3$

$$u_2 . u_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} . \begin{bmatrix} \frac{-1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix}$$

$$= \frac{1}{2} - 4 + \frac{7}{2}$$

$$= 0.$$

Which implies u_2 is orthogonal to u_3.

Thus $S = \{u_1, u_2, u_3\}$ is an orthogonal set.

Theorem

Suppose that $S = \{u_1, u_2, ..., u_p\}$ is an orthogonal set of non-zero vectors in R^n and $W = Span\{u_1, u_2, ..., u_p\}$. Then S is linearly independent set and a basis for W.

Proof

Suppose

$0 = c_1 u_1 + c_2 u_2 + ... + c_p u_p.$

Where $c_1, c_2 ..., c_p$ are scalars.

$$u_1 . 0 = u_1 . (c_1 u_1 + c_2 u_2 + ... + c_p u_p)$$

$$0 \quad = u_1 . (c_1 u_1) + u_1 . (c_2 u_2) + ... + u_1 . (c_p u_p)$$

$$= c_1 (u_1 . u_1) + c_2 (u_1 . u_2) + ... + c_p (u_1 . u_p)$$

$$= c_1 (u_1 . u_1)$$

Since S is orthogonal set, so, $u_1.u_2+...+u_1.u_p=0$ but $u_1.u_1>0.$
Therefore $c_1=0$. Similarly, it can be shown that $c_2=c_3=...=c_p=0$
Therefore by definition $S=\{u_1,u_2,...,u_p\}$ is linearly independent set and by definition of basis is a basis for subspace W.

Example

If $S=\{u_1,u_2\}$ is an orthogonal set of non-zero vector in R^2. Show that S is linearly independent set. Where
$u_1=\begin{bmatrix}3\\1\end{bmatrix}$ and $u_2=\begin{bmatrix}-1\\3\end{bmatrix}$.

Solution

To show that $S=\{u_1,u_2\}$ is linearly independent set, we show that the following vector equation
$c_1u_1+c_2u_2=0.$
has only the trivial solution. i.e. $c_1=c_2=0.$

$$c_1u_1+c_2u_2=0$$

$$c_1\begin{bmatrix}3\\1\end{bmatrix}+c_2\begin{bmatrix}-1\\3\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

$$\begin{bmatrix}3c_1\\c_1\end{bmatrix}+\begin{bmatrix}-c_2\\3c_2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

$$3c_1-c_2=0$$

$$c_1+3c_2=0$$

Solve them simultaneously, gives
$c_1=c_2=0.$
Therefore if S is an orthogonal set then it is linearly independent.

Orthogonal basis

Let $S=\{u_1,u_2,...,u_p\}$ be a basis for a subspace W of R^n, is also an orthogonal basis if S is an orthogonal set.

Theorem

If $S=\{u_1,u_2,...,u_p\}$ is an orthogonal basis for a subspace W of R^n. Then each y in W can be uniquely expressed as a linear combination of $u_1,u_2,...,u_p$. That is
$y=c_1u_1+c_2u_2+...+c_pu_p.$

Where

$$c_j = \frac{y.u_j}{u_j.u_j}$$

Proof

$$\begin{aligned} y.u_1 &= (c_1u_1 + c_2u_2 + ... + c_pu_p).u_1 \\ &= (c_1u_1).u_1 + (c_2u_2).\ u_1 + ... + (c_pu_p).u_1 \\ &= c_1(u_1.u_1) + c_2(u_1.u_2) + ... + c_p(u_1.u_p) \\ &= c_1(u_1.u_1). \end{aligned}$$

Since S is orthogonal set, so, $u_1.u_2 + ... + u_1.u_p = 0$ but $u_1.u_1 > 0.$
Hence

$c_1 = \dfrac{y.u_1}{u_1.u_1}$ and similarly $c_2 = \dfrac{y.u_2}{u_2.u_2}, ...\ c_p = \dfrac{y.u_p}{u_p.u_p}.$

Example

The set $S = \{u_1, u_2, u_3\}$ as in first example is an orthogonal basis for R^3. Express y as a linear combination of the vectors in S. Where
$y = [6 \quad 1 \quad -8]^T$

Solution

We want to write
$y = c_1u_1 + c_2u_2 + c_3u_3$
Where c_1, c_2 and c_3 are to be determined.
By the above theorem

$$c_1 = \frac{y.u_1}{u_1.u_1}$$

$$= \frac{\begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} . \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} . \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}} = \frac{11}{11} = 1$$

$$c_2 = \frac{y.u_2}{u_2.u_2}$$

$$= \frac{\begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} . \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}}{\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} . \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}} = \frac{-12}{6} = -2$$

And

$$c_3 = \frac{y.u_3}{u_3.u_3}$$

$$= \frac{\begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} . \begin{bmatrix} \frac{-1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix}}{\begin{bmatrix} \frac{-1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix} . \begin{bmatrix} \frac{-1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix}} = \frac{-33}{33/2} = -2$$

Hence

$y = u_1 - 2u_2 - 2u_3$.

Example

The set $S = \{u_1, u_2, u_3\}$ is an orthogonal basis for R^3. Write y as a linear combination of the vectors in S. Where

$$y = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}, u_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Solution

We want to write

$y = c_1 u_1 + c_2 u_2 + c_3 u_3$

Where c_1, c_2 and c_3 are to be determined.
By the above theorem

$$c_1 = \frac{y.u_1}{u_1.u_1} = \frac{\begin{bmatrix}3\\7\\4\end{bmatrix}.\begin{bmatrix}1\\-1\\0\end{bmatrix}}{\begin{bmatrix}1\\-1\\0\end{bmatrix}.\begin{bmatrix}1\\-1\\0\end{bmatrix}} = \frac{3-7+0}{1+1+0} = -2$$

$$c_2 = \frac{y.u_2}{u_2.u_2} = \frac{\begin{bmatrix}3\\7\\4\end{bmatrix}.\begin{bmatrix}1\\1\\0\end{bmatrix}}{\begin{bmatrix}1\\1\\0\end{bmatrix}.\begin{bmatrix}1\\1\\0\end{bmatrix}} = \frac{3+7+0}{1+1} = 5$$

And

$$c_3 = \frac{y.u_3}{u_3.u_3} = \frac{\begin{bmatrix}3\\7\\4\end{bmatrix}.\begin{bmatrix}0\\0\\1\end{bmatrix}}{\begin{bmatrix}0\\0\\1\end{bmatrix}.\begin{bmatrix}0\\0\\1\end{bmatrix}} = \frac{4}{1} = 4$$

Hence
$y = -2u_1 + 5u_2 + 4u_3$.

Exercise

The set $S = \{u_1, u_2, u_3\}$ is an orthogonal basis for R^3. Write y as a linear combination of the vectors in S. where

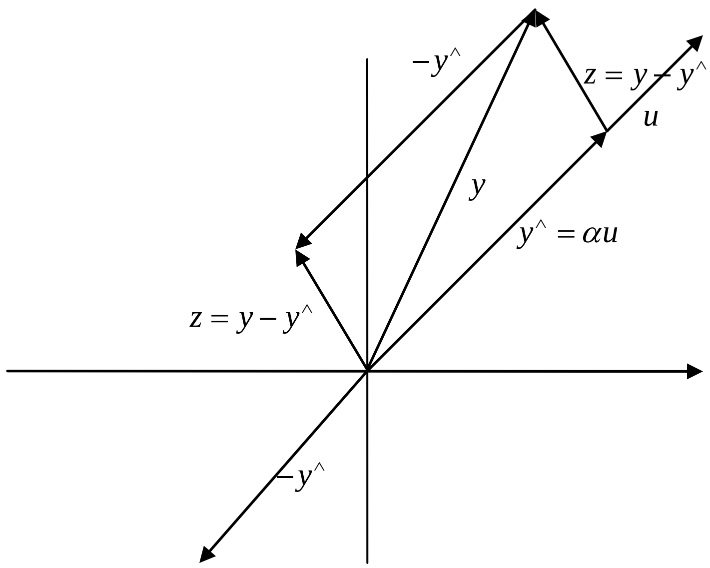
$$y = \begin{bmatrix}2\\1\\3\end{bmatrix}, u_1 = \begin{bmatrix}2\\-1\\0\end{bmatrix}, u_2 = \begin{bmatrix}\frac{1}{5}\\\frac{2}{5}\\-1\end{bmatrix} \text{ and } u_3 = \begin{bmatrix}\frac{8}{3}\\\frac{16}{3}\\\frac{8}{3}\end{bmatrix}$$

An Orthogonal Projection (Decomposition of a vector into the sum of two vectors

Decomposition of a non- zero vector $y \in R^n$ into the sum of two vectors in such a way, one is multiple of $u \in R^n$ and the other orthogonal to u. That is

$$y = y^{\wedge} + z$$

Where $y^{\wedge} = \alpha u$ for some scalar α and z is orthogonal to u.



In the above figure a vector y is decomposed into two vectors $z = y - y^{\wedge}$ and $y^{\wedge} = \alpha u$. Clearly it can be seen that $z = y - y^{\wedge}$ is orthogonal to u and $y^{\wedge} = \alpha u$ is a multiple of u. Since $z = y - y^{\wedge}$ is orthogonal to u.

Therefore

$$z.u = 0$$

$$(y - y^{\wedge}).u = 0$$

$$(y - \alpha u).u = 0$$

$$y.u - \alpha(u.u) = 0$$

$$\Rightarrow \alpha = \frac{y.u}{u.u}$$

And

$$z = y - y^{\wedge}$$

$$= y - \frac{y.u}{u.u}u$$

Hence

$y^{\wedge} = \frac{y.u}{u.u}u$, which is an orthogonal projection of y onto u .

And

$$z = y - y^{\wedge}$$

$$= y - \frac{y.u}{u.u}u \quad \text{is a component of } y$$

Example

Let $\boldsymbol{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\boldsymbol{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.

Find the orthogonal projection of **y** onto **u**. Then write **y** as a sum of two orthogonal vectors, one in span {u} and one orthogonal to **u**.

Solution

Compute

$$\boldsymbol{y} \cdot \boldsymbol{u} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$$

$$\boldsymbol{u} \cdot \boldsymbol{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

The orthogonal projection of $\boldsymbol{y}$ onto $\boldsymbol{u}$ is $\hat{y} = \frac{y \cdot u}{u \cdot u}u = \frac{40}{20}u = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$ and the component of $\boldsymbol{y}$ orthogonal to $\boldsymbol{u}$ is $y - \hat{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

The sum of these two vectors is $\boldsymbol{y}$. That is, $\underset{y}{\begin{bmatrix} 7 \\ 6 \end{bmatrix}} = \underset{\hat{y}}{\begin{bmatrix} 8 \\ 4 \end{bmatrix}} + \underset{(y-\hat{y})}{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}$

This decomposition of $\boldsymbol{y}$ is illustrated in Fig. 3. **Note**: If the calculations above are correct, then $\{\hat{\boldsymbol{y}}, \boldsymbol{y} - \hat{\boldsymbol{y}}\}$ will be an orthogonal set. As a check, compute

$$\hat{\boldsymbol{y}} \cdot (\boldsymbol{y} - \hat{\boldsymbol{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$

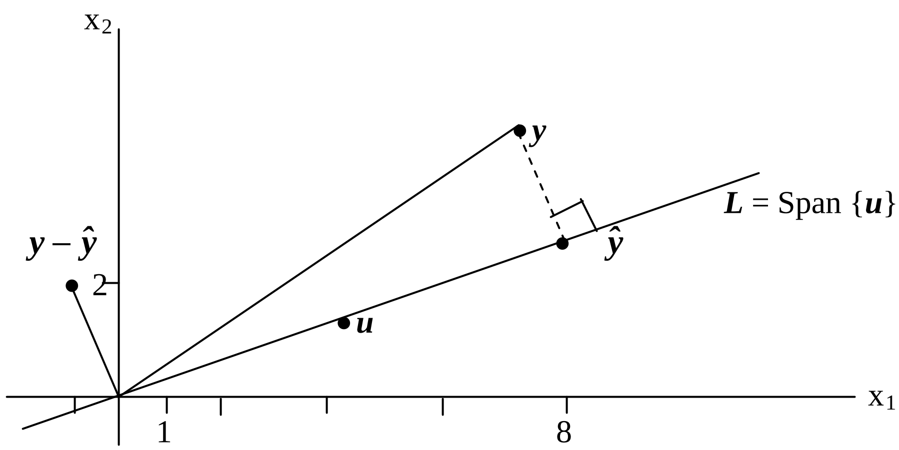
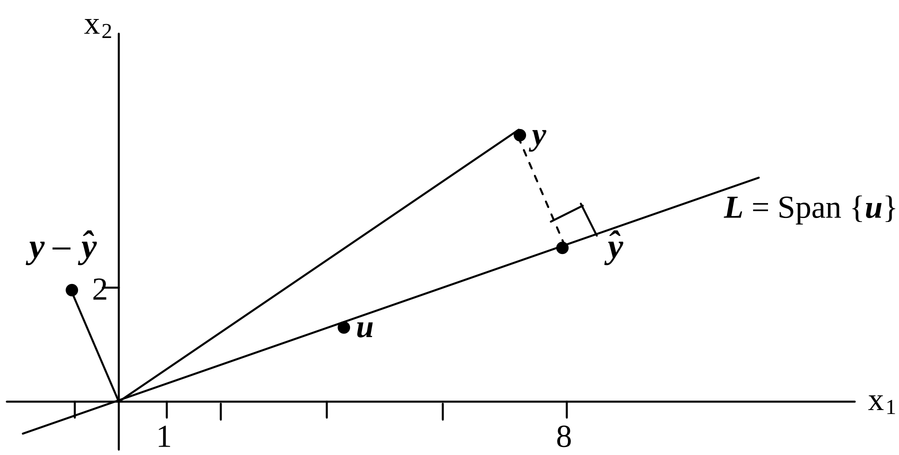


Figure 3 The orthogonal projection of ***y*** on to a line through the origin.

Example

Find the distance in figure below from *y* to *L*.



Solution

The distance from y to L is the length of the perpendicular line segment from y to the orthogonal projection $\hat{y}$.

The length equals the length of $y-\hat{y}$.

This distance is

$\|y-\hat{y}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$

Example

Decompose ***y* = (-3,-4)** into two vectors $\hat{y}$ and ***z***, where $\hat{y}$ is a multiple of ***u* = (-3, 1)** and ***z*** is orthogonal to ***u*.** Also prove that $y^{\wedge}.z=0$

Solution

It is very much clear that $y^{\wedge}$ is an orthogonal projection of ***y*** onto ***u*** and it is calculated by applying the following formula

$$y^{\wedge}=\frac{y.u}{u.u}u$$

$$=\frac{\begin{bmatrix}-3\\-4\end{bmatrix}.\begin{bmatrix}-3\\1\end{bmatrix}}{\begin{bmatrix}-3\\1\end{bmatrix}.\begin{bmatrix}-3\\1\end{bmatrix}}\begin{bmatrix}-3\\1\end{bmatrix} \quad =\frac{9-4}{9+1}\begin{bmatrix}-3\\1\end{bmatrix} \quad =\frac{1}{2}\begin{bmatrix}-3\\1\end{bmatrix} \quad =\begin{bmatrix}-\frac{3}{2}\\ \frac{1}{2}\end{bmatrix}$$

$$z=y-y^{\wedge}=\begin{bmatrix}-3\\-4\end{bmatrix}-\begin{bmatrix}\frac{-3}{2}\\ \frac{1}{2}\end{bmatrix}=\begin{bmatrix}-3+3/2\\-4-1/2\end{bmatrix}=\begin{bmatrix}-\frac{3}{2}\\-\frac{9}{2}\end{bmatrix}$$

So,

$$y^{\wedge}=\begin{bmatrix}-\frac{3}{2}\\ \frac{1}{2}\end{bmatrix} \quad and \quad z=\begin{bmatrix}-\frac{3}{2}\\-\frac{9}{2}\end{bmatrix}$$

Now

$$y^{\wedge}.z = \begin{bmatrix} -\frac{3}{2} \\ \frac{1}{2} \end{bmatrix} . \begin{bmatrix} -\frac{3}{2} \\ -\frac{9}{2} \end{bmatrix}$$
$$= \frac{9}{4} - \frac{9}{4}$$
$$= 0$$

Therefore, $y^{\wedge}$ is orthogonal to ***z***

Exercise

Find the orthogonal projection of a vector ***y* = (-3, 2)** onto ***u* = (2, 1).** Also prove that $y = y^{\wedge} + z$, where $y^{\wedge}$ a multiple of ***u*** and **z** is is an orthogonal to ***u*.**

Orthonormal Set

Let $S = \{u_1, u_2, ..., u_p\}$ be the set of non-zero vectors in R^n, is said to be an orthonormal set if S is an orthogonal set of unit vectors.

Example

Show that $S = \{u_1, u_2, u_3\}$ is an orthonormal set. Where

$$u_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{-1}{\sqrt{5}} \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix}.$$

Solution

To show that S is an orthonormal set, we show that it is an orthogonal set of unit vectors.

It can be easily prove that S is an orthogonal set because

$u_i . u_j = 0 \; \forall \, i \neq j, i, j = 1, 2, 3.$

Furthermore

$$u_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{-1}{\sqrt{5}} \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} u_3 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix}.$$

$$u_1 . u_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{-1}{\sqrt{5}} \end{bmatrix} . \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{-1}{\sqrt{5}} \end{bmatrix}$$

$$= \frac{4}{5} + 0 + \frac{1}{5}$$

$$= 1$$

$$u_2 . u_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} . \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

$$= 0 + 1 + 0$$

$$= 1$$

And

$$u_3 . u_3 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix} . \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix}$$

$$= \frac{1}{5} + \frac{4}{5}$$

$$= 1$$

Hence

$S = \{u_1, u_2, u_3\}$ is an orthonormal set.

Orthonormal basis

Let $S = \{u_1, u_2, ..., u_p\}$ be a basis for a subspace W of R^n, is also an orthonormal basis if S is an orthonormal set.

Example

Show that $S = \{u_1, u_2, u_3\}$ is an orthonormal basis of R^3, where

$$u_1 = \begin{bmatrix} \frac{3}{\sqrt{11}} \\ \frac{1}{\sqrt{11}} \\ \frac{1}{\sqrt{11}} \end{bmatrix}, u_2 = \begin{bmatrix} \frac{-1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} \frac{-1}{\sqrt{66}} \\ \frac{-4}{\sqrt{66}} \\ \frac{7}{\sqrt{66}} \end{bmatrix}.$$

Solution

To show that $S = \{u_1, u_2, u_3\}$ is an orthonormal basis, it is sufficient to show that it is an orthogonal set of unit vectors. That is

$u_i . u_j = 0 \, \forall \, i \neq j, i, j = 1, 2, 3.$

And

$u_i . u_j = 1 \, \forall \, i = j, i, j = 1, 2, 3.$

Clearly it can be seen that

$u_1 . u_2 = 0,$

$u_1 . u_3 = 0$

And

$u_2 . u_3 = 0.$

Furthermore

$u_1 . u_1 = 1,$

$u_2 . u_2 = 1$

And

$u_3 . u_3 = 1.$

Hence S is an orthonormal basis of R^3.

Theorem

A $m \times n$ matrix U has orthonormal columns if and only if $U^t U = I$

Proof

Keep in mind that in an if and only if statement, one part depends on the other, so, each part is proved separately. That is, we consider one part and then prove the other part with the help of that assumed part.

Before proving both sides of the statements, we have to do some extra work which is necessary for the better understanding.

Let $u_1, u_2, ..., u_m$ be the columns of ***U***. Then ***U*** can be written in matrix form as

$$U = [u_1 \;\; u_2 \;\; u_3 \; ... \; u_m]$$

Taking transpose, it becomes

$$U^t = \begin{bmatrix} u_1^t \\ u_2^t \\ u_3^t \\ . \\ . \\ . \\ u_m^t \end{bmatrix}$$

$$U^t U = \begin{bmatrix} u_1^t \\ u_2^t \\ u_3^t \\ . \\ . \\ . \\ u_m^t \end{bmatrix} \begin{bmatrix} u_1 & u_2 & u_3 . & . & . & u_m \end{bmatrix} = \begin{bmatrix} u_1^t u_1 & u_1^t u_2 & u_1^t u_3 & \dots & u_1^t u_m \\ u_2^t u_1 & u_2^t u_2 & u_2^t u_3 & \dots & u_2^t u_m \\ u_3^t u_1 & u_3^t u_2 & u_3^t u_3 & \dots & u_3^t u_m \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ u_m^t u_1 & u_m^t u_2 & u_m^t u_3 & \dots & u_m^t u_m \end{bmatrix}$$

$$As \quad u.v = v^t u$$

$$Therefore$$

$$U^t U = \begin{bmatrix} u_1.u_1 & u_1.u_2 & u_1.u_3 & \dots & u_1.u_m \\ u_2.u_1 & u_2.u_2 & u_2.u_3 & \dots & u_2.u_m \\ u_3.u_1 & u_3.u_2 & u_3.u_3 & \dots & u_3.u_m \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ u_m.u_1 & u_m.u_2 & u_m.u_3 & \dots & u_m.u_m \end{bmatrix}$$

Now, we come to prove the theorem.
First suppose that $U^t U = I$, and we prove that columns of U are orthonormal.
Since, we assume that

$$U^t U = \begin{bmatrix} u_1.u_1 & u_1.u_2 & u_1.u_3 & \dots & u_1.u_m \\ u_2.u_1 & u_2.u_2 & u_2.u_3 & \dots & u_2.u_m \\ u_3.u_1 & u_3.u_2 & u_3.u_3 & \dots & u_3.u_m \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ u_m.u_1 & u_m.u_2 & u_m.u_3 & \dots & u_m.u_m \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Clearly, it can be seen that

$u_i\,.u_j = 0 \;\; for\, i \neq j \quad i, j = 1, 2, \dots m$

and

$u_i\,.u_j = 1 \;\; for\, i = j \quad i, j = 1, 2, \dots m$

Therefore, columns of U are orthonormal.

Next suppose that the columns of U are orthonormal and we will show that $U^t U = I$.

Since we assume that columns of U are orthonormal, so, we can write

$u_i\,.u_j = 0 \;\; for\, i \neq j \quad i, j = 1, 2, \dots m$

and

$u_i\,.u_j = 1 \;\; for\, i = j \quad i, j = 1, 2, \dots m$

Hence, $U^t U = \begin{bmatrix} u_1.u_1 & u_1.u_2 & u_1.u_3 & \dots & u_1.u_m \\ u_2.u_1 & u_2.u_2 & u_2.u_3 & \dots & u_2.u_m \\ u_3.u_1 & u_3.u_2 & u_3.u_3 & \dots & u_3.u_m \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ u_m.u_1 & u_m.u_2 & u_m.u_3 & \dots & u_m.u_m \end{bmatrix}$

$$= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

That is
$U^t U = I$.
Which is our required result.

Exercise
Prove that the following matrices have orthonormal columns using above theorem.

(1) $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

(2) $\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}$

(3) $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

Solution (1)
Let

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$U^t = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Then

$$U^t U = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$U^t U = I$
Therefore, by the above theorem, U has orthonormal columns.

(2) And (3) are left for reader.

Theorem

Let U be an $m \times n$ matrix with orthonormal columns, and let x and y be in R^n. Then

a) $\|Ux\| = \|x\|$
b) $(Ux).(Uy) = x.y$
c) $(Ux).(Uy) = 0 \;\; iff \;\; x.y = 0$

Example

$$Let\ U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \;\; and \;\; X = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$$

Verify that $\|Ux\| = \|x\|$

Solution

Notice that U has orthonormal columns and

$$U^T U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Ux = \frac{1}{\sqrt{2}}\begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}\begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|Ux\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|x\| = \sqrt{2+9} = \sqrt{11}$$

Lecture 19
Orthogonal Decomposition

Objectives

The objectives of the lecture are to learn about:

- Orthogonal Decomposition Theorem.
- Best Approximation Theorem.

Orthogonal Projection

The orthogonal projection of a point in R^2 onto a line through the origin has an important analogue in R^n .
That is given a vector Y and a subspace W in R^n, there is a vector $\hat{y}$ in W such that

1) $\hat{y}$ is the unique vector in W for which $y-\hat{y}$ is orthogonal to W, and
2) $\hat{y}$ is the unique vector in W closest to *y*.

We observe that whenever a vector y is written as a linear combination of vectors $u_1, u_2, ..., u_n$ in a basis of R^n, the terms in the sum for y can be grouped into two parts so that y can be written as $y = z_1 + z_2$, where z_1 is a linear combination of some of the u_i's, and z_2 is a linear combination of the rest of the u_i's . This idea is particularly useful when $\{u_1, u_2, ..., u_n\}$ is an orthogonal basis.

Example 1

Let $\{u_1, u_2, ..., u_5\}$ be an orthogonal basis for R^5 and let $y = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + ... + c_5\boldsymbol{u}_5$.
Consider the subspace W= Span{u_1, u_2} and write y as the sum of a vector z_1 in W and a vector z_2 in $\boldsymbol{W}^{\perp}$.

Solution

Write

$$y = \underbrace{c_1u_1 + c_2u_2}_{z_1} + \underbrace{c_3u_3 + c_4u_4 + c_5u_5}_{z_2}$$

where $z_1 = c_1u_1 + c_2u_2$ is in Span of {$\boldsymbol{u_1}, \boldsymbol{u_2}$}, and $z_2 = c_3u_3 + c_4u_4 + c_5u_5$ is in Span of {$\boldsymbol{u_3}, \boldsymbol{u_4}, \boldsymbol{u_5}$}.

To show that $\boldsymbol{z_2}$ is in $\boldsymbol{W}^\perp$, it suffices to show that $\boldsymbol{z_2}$ is orthogonal to the vectors in the basis {$\boldsymbol{u_1}, \boldsymbol{u_2}$} for $\boldsymbol{W}$. Using properties of the inner product, compute

$$z_2 \cdot u_1 = (c_3u_3 + c_4u_4 + c_5u_5) \cdot u_1 = c_3u_3 \cdot u_1 + c_4u_4 \cdot u_1 + c_5u_5 \cdot u_1 = 0$$

since $\boldsymbol{u_1}$ is orthogonal to $\boldsymbol{u_3}$, $\boldsymbol{u_4}$, and $\boldsymbol{u_5}$, a similar calculation shows that $\boldsymbol{z_2} \cdot \boldsymbol{u_2} = 0$

Thus, $\boldsymbol{z_2}$ is in $\boldsymbol{W}^\perp$.

Orthogonal decomposition theorem

Let W be a subspace of R^n, then each y in R^n can be written uniquely in the form

$$y = \hat{y} + z$$

Where $y^\wedge \in W \text{ and } z \in W^\perp$

Furthermore, if $\{u_1, u_2, ..., u_p\}$ is any orthogonal basis for W, then

$$y^\wedge = c_1u_1 + c_2u_2 + ... + c_nu_n, \text{where } c_j = \frac{y.u_j}{u_j.u_j}$$

Proof

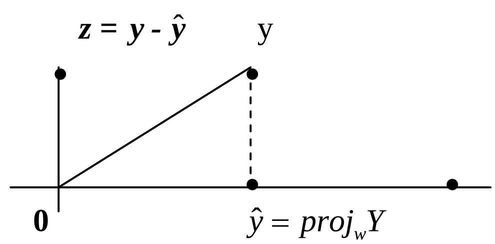


Fig: Orthogonal projection of y on to W.

Firstly, we show that $y^\wedge \in W$, $z \in W^\perp$. Then we will show that $y = y^\wedge + z$ can be represented in a unique way.

Suppose W is a subspace of R^n and let $\{u_1, u_2, ..., u_p\}$ be an orthogonal basis for W.

As $y^\wedge = c_1u_1 + c_2u_2 ... c_pu_p$ where $c_j = \frac{y.u_j}{u_j.u_j}$

Since $\{u_1, u_2, ..., u_p\}$ is the basis for W and $y^\wedge$ is written as a linear combination of these basis vectors. Therefore, by definition of basis $y^\wedge \in W$.

Now, we will show that $z = y - y^\wedge \in W^\perp$. For this it is sufficient to show that $z \perp u_j$ for each $j = 1, 2, ... p.$

Let $u_1 \in W$ be an arbitrary vector.

$$
\begin{aligned}
z \,.\, u_1 &= (y - y^\wedge).u_1 \\
&= y.u_1 - y^\wedge .u_1 \\
&= y.u_1 - (c_1 u_1 + c_2 u_2 + ... + c_p u_p).u_1 \\
&= y.u_1 - c_1(u_1.u_1) - c_2(u_2.u_1) - ... c_p(u_p.u_1) \\
&= y.u_1 - c_1(u_1.u_1) \quad \textit{where } u_j.u_1 = 0,\ j = 2, 3, ... p \\
&= y.u_1 - \frac{y.u_1}{u_1.u_1}(u_1.u_1) \\
&= y.u_1 - y.u_1 \\
&= 0
\end{aligned}
$$

Therefore, $z \perp u_1.$

Since u_1 is an arbitrary vector, therefore $z \perp u_j$ for $j = 1, 2, ... p.$

Hence by definition of $W^\perp, z \in W^\perp$

Now, we must show that $y = y^\wedge + z$ is unique by contradiction.

Let $y = y^\wedge + z$ and $y = y_1^\wedge + z_1$, where $y^\wedge, y_1^\wedge \in W$ and $z, z_1 \in W^\perp$,

also $z \neq z_1$ *and* $y^\wedge \neq y_1^\wedge$. Since above representations for y are equal, that is

$$y^\wedge + z = y_1^\wedge + z_1$$

$$\Rightarrow \quad y^\wedge - y_1^\wedge = z_1 - z$$

Let

$$s = y^\wedge - y_1^\wedge$$

Then

$$s = z_1 - z$$

Since W is a subspace, therefore, by closure property

$$s = y^\wedge - y_1^\wedge \in W$$

Furthermore, $W^\perp$ is also a subspace, therefore by closure property

$$s = z_1 - z \in w^\perp$$

Since

$s \in W$ and $s \in W^\perp$. Therefore by definition $s \perp s$

That is **s. s**=0

Therefore

$$s = y^\wedge - y_1^\wedge = 0$$

$$\Rightarrow y^\wedge = y_1^\wedge$$

Also

$z_1 = z$

This shows that representation is unique.

Example

Let $u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Observe that $\{u_1, u_2\}$ is an orthogonal basis for W=span$\{u_1, u_2\}$, write y as the sum of a vector in W and a vector orthogonal to W.

Solution

Since $y^\wedge \in W$, therefore $y^\wedge$ can be written as following:

$$y^\wedge = c_1u_1 + c_2u_2$$

$$= \frac{y.u_1}{u_1.u_1}u_1 + \frac{y.u_2}{u_2.u_2}u_2$$

$$= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -2 \\ 10 \\ 1 \end{bmatrix}$$

$$y - y^\wedge = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

$$= 7/5\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

Above theorem ensures that $y - y^\wedge$ is in $W^\perp$.

You can also verify by $(y - y^\wedge).u_1 = 0$ and $(y - y^\wedge).u_2 = 0$.

The desired decomposition of y is

$$y=\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}+\begin{bmatrix}7/5\\0\\14/5\end{bmatrix}$$

Example

Let W=span$\{u_1,u_2\}$, where $u_1=\begin{bmatrix}1\\-3\\2\end{bmatrix}$ and $u_2=\begin{bmatrix}4\\2\\1\end{bmatrix}$

Decompose $y=\begin{bmatrix}2\\-2\\5\end{bmatrix}$ into two vectors; one in W and one in $W^{\perp}$. Also verify that these two vectors are orthogonal.

Solution

Let $y^{\wedge}\in W$ and $z=y-y^{\wedge}\in W^{\perp}$.

Since $y^{\wedge}\in W$, therefore $y^{\wedge}$ can be written as following:

$$\begin{aligned}y^{\wedge}&=c_1u_1+c_2u_2\\&=\frac{y.u_1}{u_1.u_1}u_1+\frac{y.u_2}{u_2.u_2}u_2\\&=\frac{9}{7}\begin{bmatrix}1\\-3\\2\end{bmatrix}+\frac{3}{7}\begin{bmatrix}4\\2\\1\end{bmatrix}\\y^{\wedge}&=\begin{bmatrix}3\\-3\\3\end{bmatrix}\end{aligned}$$

Now

$$z=y-y^{\wedge}=\begin{bmatrix}2\\-2\\5\end{bmatrix}-\begin{bmatrix}3\\-3\\3\end{bmatrix}$$

$$z=\begin{bmatrix}-1\\1\\2\end{bmatrix}$$

Now we show that $z\perp y^{\wedge}$, i.e. $z.y^{\wedge}=0$

$$z.y^{\wedge} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} . \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix}$$
$$= 0$$

Therefore $z \perp y^{\wedge}$.

Exercise

Let W=span$\{u_1, u_2\}$, where $u_1 = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$

Write $y = \begin{bmatrix} 6 \\ -8 \\ 12 \end{bmatrix}$ as a sum of two vectors; one in W and one in $W^{\perp}$. Also verify that these two vectors are orthogonal.

Best Approximation Theorem

Let W is a finite dimensional subspace of an inner product space V and y is any vector in V. The best approximation to y from W is then $\mathrm{Pr}\,oj_w^y$, i.e for every w (that is not $\mathrm{Pr}\,oj_w^y$) in W, we have

$$\left\| y - \mathrm{Pr}\,oj_w^y \right\| < \left\| y - w \right\|.$$

Example

Let W=span$\{u_1, u_2\}$, where $u_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$, $u_2 = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$ and $y = \begin{bmatrix} 2 \\ -2 \\ 5 \end{bmatrix}$. Then using above theorem, find the distance from y to W.

Solution

Using above theorem the distance from y to W is calculated using the following formula

$$\left\| y - \mathrm{Pr}\,oj_w^y \right\| = \left\| y - y^{\wedge} \right\|$$

Since, we have already calculated

$$y - y^{\wedge} = \begin{bmatrix} 2 \\ -2 \\ 5 \end{bmatrix} - \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

So $\left\| y - y^{\wedge} \right\| = \sqrt{6}$

Example

The distance from a point y in R^n to a subspace W is defined as the distance from y to the nearest point in W.

Find the distance from ***y*** to ***W*** =span$\{u_1, u_2\}$, where

$$y = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, u_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

By the Best Approximation Theorem, the distance from ***y*** to ***W*** is $\|y-\hat{y}\|$, where $\hat{\boldsymbol{y}}$ = proj$_{\boldsymbol{w}}$***y***. Since, $\{\boldsymbol{u_1}, \boldsymbol{u_2}\}$ is an orthogonal basis for ***W***, we have

$$\hat{y} = \frac{15}{30}u_1 + \frac{-21}{6}u_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$\boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|y-\hat{y}\|^2 = 3^2 + 6^2 = 45$$

The distance from ***y*** to ***W*** is $\sqrt{45} = 3\sqrt{5}$.

<u>Theorem</u>

If $\{u_1, u_2, ..., u_p\}$ is an orthonormal basis for a subspace W of R^n, then

$$\Pr oj_w y = (y.u_1)u_1 + (y.u_2)u_2 + ...(y.u_p)u_p$$

$$If \ \ U = [u_1 \ u_2 \ ....u_p]$$

$$then \ \Pr oj_w^y = UU^T y \quad \forall y \, in \, R^n$$

<u>Example</u>

Let $u_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}, \ y = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$

and ***W*** =span$\{u_1, u_2\}$. Use the fact that u_1 and u_2 are orthogonal to compute $\Pr oj_w y$.

<u>Solution</u>

$$\operatorname{Proj}_w y = \frac{y.u_1}{u_1.u_1}u_1 + \frac{y.u_2}{u_2.u_2}u_2$$
$$= \frac{88}{66}u_1 + \frac{-2}{6}u_2$$

$$= \frac{4}{3}\begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix} = y$$

In this case,y happens to be a linear combination of u_1 and u_2 . So y is in W. The closest point in W to y is y itself.

Lecture 20

Orthogonal basis, Gram-Schmidt Process, Orthonormal basis

Example 1

Let W = Span $\{x_1, x_2\}$, where

$$x_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} \; and \; x_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Find the orthogonal basis $\{v_1, v_2\}$ for W.

Solution

Let P be a projection of $\boldsymbol{x_2}$ on to $\boldsymbol{x_1}$. The component of $\boldsymbol{x_2}$ orthogonal to $\boldsymbol{x_1}$ is $\boldsymbol{x_2 - P}$, which is in $\boldsymbol{W}$ as it is formed from $\boldsymbol{x_2}$ and a multiple of $\boldsymbol{x_1}$.

Let $\boldsymbol{v_1} = \boldsymbol{x_1}$ and compute

$$v_2 = x_2 - P = x_2 - \frac{x_2.v_1}{v_1.v_1} v_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Thus, $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is an orthogonal set of nonzero vectors in $\boldsymbol{W}$, dim $\boldsymbol{W}$ = 2 and $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is a basis of $\boldsymbol{W}$.

Example 2

For the given basis of a subspace W = Span $\{x_1, x_2\}$,

$$x_1 = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} \; and \; x_2 = \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix}$$

Find the orthogonal basis $\{v_1, v_2\}$ for W.

Solution

Set $\boldsymbol{v_1} = \boldsymbol{x_1}$ and compute

$$v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1}v_1 = \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix} - \frac{10}{20}\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ -8 \end{bmatrix}$$

Thus, an orthogonal basis for W is $\left\{ \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \\ -8 \end{bmatrix} \right\}$

Theorem

Given a basis $\{x_1, \ldots, x_p\}$ for a subspace W of R^n. Define

$$v_1 = x_1 \quad v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1}v_1$$

$$v_3 = x_3 - \frac{x_3.v_1}{v_1.v_1}v_1 - \frac{x_3.v_2}{v_2.v_2}v_2$$

$$\vdots$$

$$v_p = x_p - \frac{x_p.v_1}{v_1.v_1}v_1 - \frac{x_p.v_2}{v_2.v_2}v_2 - \ldots - \frac{x_p.v_{p\text{-}1}}{v_{p\text{-}1}.v_{p\text{-}1}}v_{p\text{-}1}$$

Then $\{v_1, \ldots, v_p\}$ is an orthogonal basis for W.

In addition

Span $\{v_1, \ldots, v_k\}$= Span $\{x_1, \ldots, x_k\}$ for $1 \le k \le p$

Example 3

The following vectors $\{x_1, x_2, x_3\}$ are linearly independent

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Construct an orthogonal basis for W by Gram-Schmidt Process.

Solution

To construct orthogonal basis we have to perform the following steps.

Step 1 Let $v_1 = x_1$

Step 2

$$\text{Let } v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1} v_1$$

Since $$\boldsymbol{v_2} = \boldsymbol{x_1} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Step 3

$$v_3 = x_3 - \left[\frac{x_3.v_1}{v_1.v_1} v_1 + \frac{x_3.v_2}{v_2\ .v_2} v_2 \right] = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \left[\frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} \right]$$

$$v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Thus, $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ is an orthogonal set of nonzero vectors in $\boldsymbol{W}$.

Example 4

Find an orthogonal basis for the column space of the following matrix by Gram-Schmidt Process.

$$\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$$

Solution

Name the columns of the above matrix as x_1, x_2, x_3 and perform the Gram-Schmidt Process on these vectors.

$$x_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \; x_2 = \begin{bmatrix} 6 \\ -8 \\ -2 \\ -4 \end{bmatrix}, \; x_3 = \begin{bmatrix} 6 \\ 3 \\ 6 \\ -3 \end{bmatrix}$$

Set $\boldsymbol{v_1} = \boldsymbol{x_1}$

$$v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1} v_1$$

$$= \begin{bmatrix} 6 \\ -8 \\ -2 \\ -4 \end{bmatrix} - (-3) \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}$$

$$v_3 = x_3 - \frac{x_3.v_1}{v_1.v_1} v_1 - \frac{x_3.v_2}{v_2 \; .v_2} v_2$$

$$= \begin{bmatrix} 6 \\ 3 \\ 6 \\ -3 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} - \frac{5}{2} \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix}$$

Thus, orthogonal basis is $\left\{ \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix} \right\}$

Example 5

Find the orthonormal basis of the subspace spanned by the following vectors.

$$x_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Solution

Since from example # 1, we have

$$v_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \; v_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Orthonormal Basis

$$u_1 = \frac{1}{\|v_1\|} v_1 = \frac{1}{\sqrt{45}} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}, \; u_2 = \frac{1}{\|v_2\|} v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Example 6

Find the orthonormal basis of the subspace spanned by the following vectors.

$$x_1 = \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$$

Solution

Firstly we find v_1 and v_2 by Gram-Schmidt Process as

$$v_2 = x_2 \; = \; x_2 - \frac{x_2 . v_1}{v_1 . v_1} v_1$$

Set v1 = $\boldsymbol{x_1}$

$$v_2 = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \frac{15}{30} \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ -5/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 3 \end{bmatrix}$$

$$Now\ \|v_2\| = \frac{1}{\sqrt{54}} = \frac{1}{3\sqrt{6}}\ and\ since\ \|v_1\| = \sqrt{30}$$

$$Thus,\ the\ orthonormal\ basis\ for\ W\ is$$

$$\left\{ \frac{\boldsymbol{v_1}}{\|\boldsymbol{v_1}\|}, \frac{\boldsymbol{v_2}}{\|\boldsymbol{v_2}\|} \right\} = \left\{ \begin{bmatrix} 2/\sqrt{30} \\ -5/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}, \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \right\}$$

Theorem

If A is an m x n matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an m x n matrix whose columns form an orthonormal basis for Col A and R is an n x n upper triangular invertible matrix with positive entries on its diagonal.

Example 7

Find a QR factorization of matrix $A = \begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$

Solution

Firstly find the orthonormal basis by applying Gram Schmidt process on the columns of A. We get the following matrix Q.

$$Q = \begin{bmatrix} 1/\sqrt{5} & 1/2 & 1/2 \\ -1/\sqrt{5} & 0 & 0 \\ -1/\sqrt{5} & 1/2 & 1/2 \\ 1/\sqrt{5} & -1/2 & 1/2 \\ 1/\sqrt{5} & 1/2 & -1/2 \end{bmatrix}$$

$$Now\ R = Q^T A = \begin{bmatrix} \sqrt{5} & -\sqrt{5} & 4\sqrt{5} \\ 0 & 6 & -2 \\ 0 & 0 & 4 \end{bmatrix}$$

Verify that A=QR.

Theorem

$If\ \{u_1,\ ...,\ u_p\}\ is\ an\ orthonormal\ basis\ for\ a\ subspace\ W\ of\ R^n,\ then$

$proj_w y = (y \cdot u_1)u_1 + (y \cdot u_2)u_2 + \cdots + (y \cdot u_p)u_p$

$If\ U = [u_1\ \ u_2\ \ ...\ \ u_p],$

$then\ \ proj_w y = UU^T y\ \ \forall\ y\ in\ R^n$

The Orthogonal Decomposition Theorem

Let *W* be a subspace of R^n Then each *y* in R^n can be written uniquely in the form

$y = \hat{y} + z$

where $W^{\perp}$ is in *W* and *z* is in

In fact, if $\{u_1, \ldots, u_p\}$ is any orthogonal basis of *W*, then

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \frac{y \cdot u_p}{u_p \cdot u_p} u_p$$

and $z = y - \hat{y}$. The vector $\hat{y}$ is called the orthogonal projection of *y* onto *W* and is often written as $\text{proj}_w\, y$.

Best Approximation Theorem

Let *W* be a subspace of R^n, *y* is any vector in R^n and $\hat{\boldsymbol{y}}$ the orthogonal projection of *y* onto *W*. Then $\hat{\boldsymbol{y}}$ is the closest point in *W* to *y*, in the sense that

for all *v* in *W* distinct from $\hat{\boldsymbol{y}}$.

$$\|y - \hat{y}\| < \|y - v\|$$

The vector $y^{\wedge}$ in this theorem is called the best approximation to *y* by elements of *W*.

Exercise 1

Let W = Span {$x1$, $x2$}, where

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$$

Construct an orthonormal basis for W.

Exercise 2

Find an orthogonal basis for the column space of the following matrix by Gram-Schmidt Process.

$$\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & -5 & -2 \\ 3 & -7 & 8 \end{bmatrix}$$

Exercise 3

Find a QR factorization of

$$A = \begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{bmatrix}$$

Lecture 21
Least Square Solution

Best Approximation Theorem

Let *W* be a subspace of R^n, *y be* any vector in R^n and $\hat{y}$ the orthogonal projection of *y* onto *W*. Then $\hat{y}$ is the closest point in *W* to *y*, in the sense that $\|y-\hat{y}\| < \|y-v\|$ for all *v* in *W* distinct from $\hat{y}$

The vector $\hat{y}$ in this theorem is called the best approximation to *y* by elements of *W*.

Least-squares solution

The most important aspect of the least-squares problem is that no matter what "*x*" we select, the vector *Ax* will necessarily be in the column space Col *A*. So we seek an *x* that makes *Ax* the closest point in Col *A* to *b*. Of course, if *b* happens to be in Col *A*, then *b* is *Ax* for some *x* and such an *x* is a "least-squares solution."

Solution of the General Least-Squares Problem

Given ***A*** and ***b*** as above, apply the Best Approximation Theorem stated above to the subspace Col ***A***. Let $\hat{b} = proj_{Col\,A} b$

Since $\hat{b}$ is in the column space of ***A***, the equation $Ax = \hat{b}$ is consistent, and there is an $\hat{x}$ in $\boldsymbol{R^n}$ such that

$$A\hat{x} = \hat{b} \qquad (1)$$

Since $\hat{b}$ is the closest point in Col ***A*** to ***b***, a vector $\hat{x}$ is a least-squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$ if and only if $\hat{\boldsymbol{x}}$ satisfies $A\hat{x} = \hat{b}$. Such an $\hat{x}$ in $\boldsymbol{R^n}$ is a list of weights that will build $\hat{b}$ out of the columns of ***A***.

Normal equations for $\hat{x}$

Suppose that $\hat{\boldsymbol{x}}$ satisfies $A\hat{x} = \hat{b}$. By the Orthogonal Decomposition Theorem the projection $\hat{b}$ has the property that $b-\hat{b}$ is orthogonal to Col ***A***, so $b - A\hat{x}$ is orthogonal to each column of ***A***. If $\boldsymbol{a_j}$ is any column of ***A***, then $a_j \cdot (b - A\hat{x}) = 0$, and $a_j^T(b - A\hat{x}) = 0$. Since each $\boldsymbol{a_j^T}$ is a row of $\boldsymbol{A^T}$,

$$A^T(b - A\hat{x}) = 0 \qquad (2)$$

$$A^T b - A^T A\hat{x} = 0$$

$$A^T A\hat{x} = A^T b \qquad (3)$$

The matrix equation (3) represents a system of linear equations commonly referred to as the **normal equations for $\hat{x}$.**

Since set of least-squares solutions is nonempty and any such $\hat{\boldsymbol{x}}$ satisfies the normal equations. Conversely, suppose that $\hat{\boldsymbol{x}}$ satisfies $A^T A\hat{x} = A^T b.$ Then it satisfy that $\boldsymbol{b} - \boldsymbol{A}\hat{\boldsymbol{x}}$ is orthogonal to the rows of A^T and hence is orthogonal to the columns of *A*. Since the columns of *A* span Col *A*, the vector $\boldsymbol{b} - \boldsymbol{A}\hat{\boldsymbol{x}}$ is orthogonal to all of Col *A*. Hence the equation $b = A\hat{x} + (b - A\hat{x})$ is a decomposition of *b* into the sum of a vector in Col *A* and a vector orthogonal to Col *A*. By the uniqueness of the orthogonal decomposition, $\boldsymbol{A}\hat{\boldsymbol{x}}$ must be the orthogonal projection of *b* onto Col *A*. That is, $\boldsymbol{A}\hat{\boldsymbol{x}} = \hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{x}}$ is a least-squares solution.

Definition

If *A* is *m* x *n* and *b* is in R^n , a least-squares solution of *Ax* = *b* is an I $\hat{x}$ in R^n such that

$$\|b - A\hat{x}\| \le \|b - Ax\| \quad \forall \ x \in R^n$$

Theorem

The set of least-squares solutions of *Ax* = *b* coincides with the nonempty set of solutions of the normal equations

$$A^T A\hat{x} = A^T b$$

Example 1

Find the least squares solution and its error from the following matrices,

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solution

Firstly we find

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \quad and$$

$$A^T b = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^T A\hat{x} = A^T b$ becomes $\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$

Row operations can be used to solve this system, but since A^TA is invertible and 2×2, it is probably faster to compute $(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$

Therefore, $\hat{x} = (A^TA)^{-1}A^T b = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Now again as A= $\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$

Then $\boldsymbol{A\hat{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$

Hence $\boldsymbol{b} - \boldsymbol{A\hat{x}} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$

So $\|\boldsymbol{b} - \boldsymbol{A\hat{x}}\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$

The least-squares error is $\sqrt{84}$. For any x in R^2, the distance between b and the vector Ax is at least $\sqrt{84}$.

Example 2

Find the general least-squares solution of $Ax = b$ in the form of a free variable with

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, b = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

Solution

Firstly we find, $\boldsymbol{A^T A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$ and

$$\boldsymbol{A^T b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

Then augmented matrix for $\boldsymbol{A^T A\hat{x}} = \boldsymbol{A^T b}$ is

$$\begin{bmatrix} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The general solution is $x_1 = 3 - x_4, x_2 = -5 + x_4, x_3 = -2 + x_4,$ and x_4 is free.
So the general least-squares solution of $Ax = b$ has the form

$$\boldsymbol{\hat{x}} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Theorem

The matrix $A^T A$ is invertible iff the columns of A are linearly independent. In this case, the equation $Ax = b$ has only one least-squares solution $\hat{x}$, and it is given by

$$\hat{x} = (A^T A)^{-1} A^T b$$

Example 3

Find the least squares solution to the following system of equations.

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 4 \end{bmatrix}$$

Solution

As

$$A^T = \begin{bmatrix} 2 & 1 & 7 & 1 \\ 4 & -3 & 1 & 0 \\ 6 & 0 & 4 & 5 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 2 & 1 & 7 & 1 \\ 4 & -3 & 1 & 0 \\ 6 & 0 & 4 & 5 \end{bmatrix} \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 55 & 12 & 45 \\ 12 & 26 & 28 \\ 45 & 28 & 77 \end{bmatrix}$$

Now

$$A^T b = \begin{bmatrix} 2 & 1 & 7 & 1 \\ 4 & -3 & 1 & 0 \\ 6 & 0 & 4 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} -9 \\ -5 \\ 12 \end{bmatrix}$$

As $\quad A^T A\hat{x} = A^T b$

$$\begin{bmatrix} 55 & 12 & 45 \\ 12 & 26 & 28 \\ 45 & 28 & 77 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} -9 \\ -5 \\ 12 \end{bmatrix}$$

$$\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} -.676 \\ -.776 \\ .834 \end{bmatrix}$$

Example 4

Compute the least square error for the solution of the following equation

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 4 \end{bmatrix}$$

.

Solution

$$A\hat{x} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} -.676 \\ -.776 \\ .834 \end{bmatrix}$$

$$A\hat{x} = \begin{bmatrix} 0.548 \\ 1.652 \\ -2.172 \\ 3.494 \end{bmatrix}$$

As least square error

$$\|\in\| = \|b - A\overline{x}\|$$

is as small as possible, or in other words is smaller than all other possible choices of x.

As
$$b - A\hat{x} = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 4 \end{bmatrix} - \begin{bmatrix} 0.548 \\ 1.652 \\ -2.172 \\ 3.494 \end{bmatrix} = \begin{bmatrix} -0.548 \\ 0.652 \\ 0.172 \\ 1.494 \end{bmatrix}$$

$$\|\in\|^2 = \in_1^{\,2} + \in_2^{\,2} + \in_3^{\,2} + \in_4^{\,2}$$

Thus, least square error is

$$\|\in\| = \|b - A\overline{x}\| = \sqrt{(-0.548)^2 + (0.652)^2 + (0.172)^2 + (1.494)^2}$$

$$=0.3003+0.4251+.02958+2.23=2.987$$

Theorem

Given an m x n matrix A with linearly independent columns.Let $A = QR$ be a QR factorization of A ,then for each b in Rm, the equation $Ax = b$ has a unique least-squares solution, given by

$$\hat{x} = R^{-1}Q^{T}b$$

Example 1

Find the least square solution for $A=\begin{bmatrix}1 & 3 & 5\\ 1 & 1 & 0\\ 1 & 1 & 2\\ 1 & 3 & 3\end{bmatrix}, b=\begin{bmatrix}3\\ 5\\ 7\\ -3\end{bmatrix}$

Solution

First of all we find QR factorization of the given matrix A. For this we have to find out orthonormal basis for the column space of A by applying Gram-Schmidt Process, we get the matrix of orthonormal basis Q,

$$Q=\begin{bmatrix}1/2 & 1/2 & 1/2\\ 1/2 & -1/2 & -1/2\\ 1/2 & -1/2 & 1/2\\ 1/2 & 1/2 & -1/2\end{bmatrix} \text{ And}$$

$$R=Q^TA=\begin{bmatrix}1/2 & 1/2 & 1/2 & 1/2\\ 1/2 & -1/2 & -1/2 & 1/2\\ 1/2 & -1/2 & 1/2 & -1/2\end{bmatrix}\begin{bmatrix}1 & 3 & 5\\ 1 & 1 & 0\\ 1 & 1 & 2\\ 1 & 3 & 3\end{bmatrix}=\begin{bmatrix}2 & 4 & 5\\ 0 & 2 & 3\\ 0 & 0 & 2\end{bmatrix}$$

Then $$Q^Tb=\begin{bmatrix}1/2 & 1/2 & 1/2 & 1/2\\ 1/2 & -1/2 & -1/2 & 1/2\\ 1/2 & -1/2 & 1/2 & -1/2\end{bmatrix}\begin{bmatrix}3\\ 5\\ 7\\ -3\end{bmatrix}=\begin{bmatrix}6\\ -6\\ 4\end{bmatrix}$$

The least-squares solution $\hat{\boldsymbol{x}}$ satisfies $R\hat{x}=Q^Tb;$ that is, $\begin{bmatrix}2 & 4 & 5\\ 0 & 2 & 3\\ 0 & 0 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}6\\ -6\\ 4\end{bmatrix}$

This equation is solved easily and yields $\hat{\boldsymbol{x}}=\begin{bmatrix}10\\ -6\\ 2\end{bmatrix}$.

Example 2

Find the least squares solution $R\hat{x}=Q^Tb$ to the given matrices,

$$A = \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Solution

First of all we find QR factorization of the given matrix A. Thus, we have to make
Orthonormal basis by applying Gram Schmidt process on the columns of A,

Let $\boldsymbol{v_1} = \boldsymbol{x_1}$

$$v_2 = x_2 - P \; = \; x_2 - \frac{x_2.v_1}{v_1.v_1} v_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Thus, the orthonormal basis are

$$\left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|} \right\} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Thus

$$Q = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} and \; Q^T = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$Now \; R = Q^T A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 15 & 5 \\ 0 & 2 \end{bmatrix} \quad And \; Q^T b = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

Thus, least squares solution of $R\hat{x} = Q^T b$ is

$$\begin{bmatrix} 15 & 5 \\ 0 & 2 \end{bmatrix} \hat{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$\hat{x} = \begin{bmatrix} 1 \\ 1.8 \end{bmatrix}$$

Exercise 1

Find a least-squares solution of $Ax = b$ for

$$\boldsymbol{A} = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Exercise 2

Find the least-squares solution and its error of $Ax = b$ for

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

Exercise 3

Find the least squares solution $R\hat{x} = Q^T b$ to the given matrices,

$$A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix}, b = \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix}$$

Lecture 22

Eigen Value Problems

Let [A] be an n x n square matrix. Suppose, there exists a scalar and a vector

$X = (x_1 \quad x_2 \quad \dots \quad x_n)^T$
such that
$[A](X) = \lambda(X)$

$\frac{d}{dx}(e^{ax}) = a(e^{ax})$

$\frac{d^2}{dx^2}(\sin ax) = -a^2(\sin ax)$

Then λ is the eigen value and X is the corresponding eigenvector of the matrix [A].
We can also write it as $[A - \lambda I](X) = (O)$
This represents a set of n homogeneous equations possessing non-trivial solution, provided
$|A - \lambda I| = 0$
This determinant, on expansion, gives an n-th degree polynomial which is called characteristic polynomial of [A], which has n roots. Corresponding to each root, we can solve these equations in principle, and determine a vector called eigenvector.
Finding the roots of the characteristic equation is laborious. Hence, we look for better methods suitable from the point of view of computation. Depending upon the type of matrix [A] and on what one is looking for, various numerical methods are available.

Power Method and Jacobi's Method

Note!
We shall consider only real and real-symmetric matrices and discuss power and Jacobi's methods

Power Method

To compute the largest eigen value and the corresponding eigenvector of the system
$[A](X) = \lambda(X)$
where [A] is a real, symmetric or un-symmetric matrix, the power method is widely used in practice.

Procedure
Step 1: Choose the initial vector such that the largest element is unity.

Step 2: The normalized vector $v^{(0)}$ is pre-multiplied by the matrix [A].

Step 3: The resultant vector is again normalized.

Step 4: This process of iteration is continued and the new normalized vector is repeatedly pre-multiplied by the matrix [A] until the required accuracy is obtained.
At this point, the result looks like

$$u^{(k)} = [A]v^{(k-1)} = q_k v^{(k)}$$

Here, q_k is the desired largest eigen value and $v^{(k)}$ is the corresponding eigenvector.

Example
Find the eigen value of largest modulus, and the associated eigenvector of the matrix by power method

$$[A] = \begin{bmatrix} 2 & 3 & 2 \\ 4 & 3 & 5 \\ 3 & 2 & 9 \end{bmatrix}$$

Solution
We choose an initial vector $\upsilon^{(0)}$
as $(1,1,1)^T$.
Then, compute first iteration

$$u^{(1)} = [A]v^{(0)} \begin{bmatrix} 2 & 3 & 2 \\ 4 & 3 & 5 \\ 3 & 2 & 9 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 12 \\ 14 \end{pmatrix}$$

Now we normalize the resultant vector to get

$$u^{(1)} = 14 \begin{pmatrix} \frac{1}{2} \\ \frac{6}{7} \\ 1 \end{pmatrix} = q_1 v^{(1)}$$

The second iteration gives,

$$u^{(2)} = [A]v^{(1)} \begin{bmatrix} 2 & 3 & 2 \\ 4 & 3 & 5 \\ 3 & 2 & 9 \end{bmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{6}{7} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{39}{7} \\ \frac{67}{7} \\ \frac{171}{14} \end{pmatrix}$$

$$= 12.2143 \begin{pmatrix} 0.456140 \\ 0.783626 \\ 1.0 \end{pmatrix} = q_2 v^{(2)}$$

Continuing this procedure, the third and subsequent iterations are given in the following slides

$$u^{(3)} = [A]v^{(2)} = \begin{bmatrix} 2 & 3 & 2 \\ 4 & 3 & 5 \\ 3 & 2 & 9 \end{bmatrix} \begin{pmatrix} 0.456140 \\ 0.783626 \\ 1.0 \end{pmatrix}$$

$$= \begin{pmatrix} 5.263158 \\ 9.175438 \\ 11.935672 \end{pmatrix} = 11.935672 \begin{pmatrix} 0.44096 \\ 0.776874 \\ 1.0 \end{pmatrix} = q_3 v^{(3)}$$

$$u^{(4)} = [A]v^{(3)} = \begin{pmatrix} 5.18814 \\ 9.07006 \\ 11.86036 \end{pmatrix}$$

$$= 11.8636 \begin{pmatrix} 0.437435 \\ 0.764737 \\ 1.0 \end{pmatrix} = q_4 v^{(4)}$$

$$u^{(5)} = [A]v^{(4)} = \begin{pmatrix} 5.16908 \\ 9.04395 \\ 11.84178 \end{pmatrix}$$

$$= 11.84178 \begin{pmatrix} 0.436512 \\ 0.763732 \\ 1.0 \end{pmatrix} = q_5 v^{(5)}$$

After rounding-off, the largest eigen value and the corresponding eigenvector as accurate to two decimals are

$$\lambda = 11.84 \quad (X) = \begin{pmatrix} 0.44 \\ 0.76 \\ 1.00 \end{pmatrix}$$

Example

Find the first three iterations of the power method of the given matrix

$$\begin{bmatrix} 7 & 6 & -3 \\ -12 & -20 & 24 \\ -6 & -12 & 16 \end{bmatrix}$$

Solution

$$\begin{bmatrix} 7 & 6 & -3 \\ -12 & -20 & 24 \\ -6 & -12 & 16 \end{bmatrix}$$

$we\ choose\ \ finitial\ vector\ as\ v^{(0)} = (1,1,1)^t$

$first\ iteration$

$$u^{(1)} = [A]v^{(0)} = \begin{bmatrix} 7 & 6 & -3 \\ -12 & -20 & 24 \\ -6 & -12 & 16 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7+6-3 \\ -12-20+24 \\ -6-12+16 \end{bmatrix} = \begin{bmatrix} 10 \\ -8 \\ -2 \end{bmatrix}$$

$$by\ diagonali\sin g \begin{bmatrix} 10 \\ -8 \\ -2 \end{bmatrix} \quad u^{(1)} = 10\begin{bmatrix} 1 \\ -0.8 \\ -0.2 \end{bmatrix} = q_1 v^{(1)}$$

$\sec ond\ iteration$

$$u^{(2)} = [A]v^{(1)} = \begin{bmatrix} 7 & 6 & -3 \\ -12 & -20 & 24 \\ -6 & -12 & 16 \end{bmatrix}\begin{bmatrix} 1 \\ -0.8 \\ -0.2 \end{bmatrix} = \begin{bmatrix} 7-4.8+0.6 \\ -12+16-4.8 \\ -6+9.6-3.2 \end{bmatrix} = \begin{bmatrix} 2.8 \\ -0.8 \\ 0.4 \end{bmatrix}$$

$$by\ diagonali\sin g \begin{bmatrix} 2.8 \\ -0.8 \\ 0.4 \end{bmatrix} \quad u^{(1)} = 2.8\begin{bmatrix} 1 \\ -0.2857 \\ 0.1428 \end{bmatrix} = q_2 v^{(2)}$$

$third\ iteration$

$$u^{(3)} = [A]v^{(2)} = \begin{bmatrix} 7 & 6 & -3 \\ -12 & -20 & 24 \\ -6 & -12 & 16 \end{bmatrix}\begin{bmatrix} 1 \\ -0.2857 \\ 0.1428 \end{bmatrix} = \begin{bmatrix} 7-1.7142-0.4284 \\ -12+5.714+3.4272 \\ -6+3.4284+2.2848 \end{bmatrix} = \begin{bmatrix} 4.8574 \\ -2.8588 \\ -0.2868 \end{bmatrix}$$

$now\ daigonali\sin g$

$$\begin{bmatrix} 4.8574 \\ -2.8588 \\ -0.2868 \end{bmatrix} \quad now\ normali\sin g\ 4.8574\begin{bmatrix} 1 \\ -0.5885 \\ -0.0590 \end{bmatrix}$$

Example

Find the first three iteration of the power method applied on the following matrices

$$\begin{bmatrix} 1 & -1 & 0 \\ -2 & 4 & -2 \\ 0 & -1 & 2 \end{bmatrix} \quad use\ x^0 = (-1,2,1)^t$$

Solution

$$\begin{bmatrix} 1 & -1 & 0 \\ -2 & 4 & -2 \\ 0 & -1 & 2 \end{bmatrix} \quad USE\ x^{(0)} = (-1,2,1)^t$$

$1st\ iterations$

$$u^{(1)} = [\mathrm{A}]x^{(0)} = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 4 & -2 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1-2+0 \\ 2+8-2 \\ 0-2+2 \end{bmatrix} = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

now we normalize the resul tan *t vector to get*

$$u^{(1)} = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = 8\begin{bmatrix} \frac{-3}{8} \\ 1 \\ 0 \end{bmatrix} = q_1 x^{(1)}$$

$$u^{(2)} = [\mathrm{A}]x^{(1)} = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 4 & -2 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} \frac{-3}{8} \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{-3}{8}-1+0 \\ \frac{6}{8}+4+0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1.375 \\ 4.75 \\ -1 \end{bmatrix}$$

$$u^{(2)} = \begin{bmatrix} -1.375 \\ 4.75 \\ -1 \end{bmatrix} = 4.75\begin{bmatrix} -0.28947 \\ 1 \\ -0.2152 \end{bmatrix}$$

$$u^{(3)} = [\mathrm{A}]x^{(2)} = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 4 & -2 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} -0.28947 \\ 1 \\ -0.2152 \end{bmatrix} = \begin{bmatrix} -1.28947 \\ 4.99998 \\ -1.42104 \end{bmatrix} = 4.99998\begin{bmatrix} -0.25789 \\ 1 \\ -0.28420 \end{bmatrix}$$

Exercise

Find the largest eigen value and the corresponding eigen vector by power method after fourth iteration starting with the initial vector $\upsilon^{(0)} = (0,0,1)^T$

$$[A] = \begin{bmatrix} 1 & -3 & 2 \\ 4 & 4 & -1 \\ 6 & 3 & 5 \end{bmatrix}$$

Let

$\lambda_1, \lambda_2, \ldots, \lambda_n$

be the distinct eigen values of an n x n matrix [A], such that $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$ and suppose $v_1, v_2, \ldots, v_n$ are the corresponding eigen vectors

Power method is applicable if the above eigen values are real and distinct, and hence, the corresponding eigenvectors are linearly independent.

Then, any eigenvector v in the space spanned by the eigenvectors $v_1, v_2, \ldots, v_n$ can be written as their linear combination $v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$

Pre-multiplying by A and substituting

$$Av_1 = \lambda_1 v_1, \qquad Av_2 = \lambda_2 v_2, \ \ldots \qquad Av_n = \lambda_n v_n$$

We get

$$Av = \lambda_1 \left(c_1 v_1 + c_2 \frac{\lambda_2}{\lambda_1} v_2 + \cdots + c_n \frac{\lambda_n}{\lambda_1} v_n \right)$$

Again, pre-multiplying by A and simplifying, we obtain

$$A^2 v = \lambda_1^2 \left[c_1 v_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^2 v_2 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^2 v_n \right]$$

Similarly, we have

$$A^r v = \lambda_1^r \left[c_1 v_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^r v_2 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^r v_n \right]$$

and

$$A^{r+1} v = (\lambda_1)^{r+1} \left[c_1 v_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^{r+1} v_2 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^{r+1} v_n \right]$$

Now, the eigen value λ_1

can be computed as the limit of the ratio of the corresponding components of $A^r v$ and $A^{r+1} v$.

That is,

$$\lambda_1 = \frac{\lambda_1^{r+1}}{\lambda_1^r} = \underset{r\to\infty}{Lt} \frac{(A^{r+1}v)_p}{(A^r v)_p}, \qquad p = 1, 2, \ldots, n$$

Here, the index *p* stands for the p-th component in the corresponding vector
Sometimes, we may be interested in finding the least eigen value and the corresponding eigenvector.
In that case, we proceed as follows.
We note that $[A](X) = \lambda(X)$.
Pre-multiplying by $[A^{-1}]$, we get
$[A^{-1}][A](X) = [A^{-1}]\lambda(X) = \lambda[A^{-1}](X)$
Which can be rewritten as
$[A^{-1}](X) = \dfrac{1}{\lambda}(X)$
which shows that the inverse matrix has a set of eigen values which are the reciprocals of the eigen values of [*A*].
Thus, for finding the eigen value of the least magnitude of the matrix [*A*], we have to apply power method to the inverse of [*A*].

Lecture 23

Jacobi's Method

Definition
An *n x n* matrix [*A*] is said to be orthogonal if
$[A]^T[A]=[I],$
$\text{i.e.}[A]^T=[A]^{-1}$
In order to compute all the eigen values and the corresponding eigenvectors of a real symmetric matrix, Jacobi's method is highly recommended. It is based on an important property from matrix theory, which states
that, if [*A*] is an *n x n* real symmetric matrix, its eigen values are real, and there exists an orthogonal matrix [*S*] such that the diagonal matrix D is
$[S^{-1}][A][S]$

This digitalization can be carried out by applying a series of orthogonal transformations
$S_1, S_2, ..., S_n,$
Let *A* be an *n x n* real symmetric matrix. Suppose $\left|a_{ij}\right|$ be numerically the largest element amongst the off-diagonal elements of *A*. We construct an orthogonal matrix *S*1 defined as

$$s_{ij}=-\sin\theta,\quad s_{ji}=\sin\theta,$$
$$s_{ii}=\cos\theta,\quad s_{jj}=\cos\theta$$
While each of the remaining off-diagonal elements are zero, the remaining diagonal elements are assumed to be unity. Thus, we construct *S*1 as under

$$\begin{array}{cc}
 & \begin{array}{cccccccc} & & & \text{i-th column} & & j\text{-th column} & & \\ & & & \downarrow & & \downarrow & & \end{array} \\
S_1= & \begin{bmatrix}
1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0\\
0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0\\
\vdots & \vdots & & \vdots & & \vdots & & \vdots\\
0 & 0 & \cdots & \cos\theta & \cdots & -\sin\theta & \cdots & 0\\
\vdots & \vdots & & \vdots & & \vdots & & \vdots\\
0 & 0 & \cdots & \sin\theta & \cdots & \cos\theta & \cdots & 0\\
\vdots & \vdots & & \vdots & & \vdots & & \vdots\\
0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1
\end{bmatrix}
\begin{array}{l} \\ \\ \\ \leftarrow \text{i-th row} \\ \\ \leftarrow j\text{-th row} \\ \\ \\ \end{array}
\end{array}$$

Where $\cos\theta, -\sin\theta, \sin\theta$ *and* $\cos\vartheta$ are inserted in $(i,i),(i,j),(j,i),(j,j)-\text{th}$ positions respectively, and elsewhere it is identical with a unit matrix.
Now, we compute
$D_1=S_1^{-1}AS_1=S_1^TAS_1$
Since S1 is an orthogonal matrix, such that .After the transformation, the elements at the position *(i , j), (j , i)* get annihilated, that is *dij and dji* reduce to zero, which is seen as follows:

$$\begin{bmatrix} d_{ii} & d_{ij} \\ d_{ji} & d_{jj} \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} a$$

$$\begin{bmatrix} a_{ii}\cos^2\theta \sin\theta\cos\theta + a_{jj}\sin^2\theta & (a_{jj}-a_{ii})\sin\theta\cos\theta + a_{ij}\cos 2\theta \\ (a_{ji}-a_{ii})\sin\theta\cos\theta + a_{ij}\cos 2\theta & a_{ii}\sin^2\theta + a_{jj}\cos^2\theta - 2a_{ij}\sin\theta\cos\theta \end{bmatrix}$$

Therefore, $d_{ij} = 0$ only if,

$$a_{ij}\cos 2\theta + \frac{a_{jj}-a_{ii}}{2}\sin 2\theta = 0$$

That is if

$$\tan 2\theta = \frac{2a_{ij}}{a_{ii}-a_{jj}}$$

Thus, we choose θ such that the above equation is satisfied, thereby, the pair of off-diagonal elements *dij* and *dji* reduces to zero.However, though it creates a new pair of zeros, it also introduces non-zero contributions at formerly zero positions.
Also, the above equation gives four values of , but to get the least possible rotation, we choose

Lecture 24

Example
Find all the eigen values and the corresponding eigen vectors of the matrix by Jacobi's method

$$A = \begin{bmatrix} 1 & \sqrt{2} & 2 \\ \sqrt{2} & 3 & \sqrt{2} \\ 2 & \sqrt{2} & 1 \end{bmatrix}$$

Solution
The given matrix is real and symmetric. The largest off-diagonal element is found to be $a_{13} = a_{31} = 2.$
Now, we compute

$$\tan 2\theta = \frac{2a_{ij}}{a_{ii} - a_{jj}} = \frac{2a_{13}}{a_{11} - a_{33}} = \frac{4}{0} = \infty$$

This gives, $\theta = \pi/4$
Thus, we construct an orthogonal matrix *Si* as

$$S_1 = \begin{bmatrix} \cos\frac{\pi}{4} & 0 & -\sin\frac{\pi}{4} \\ 0 & 1 & 0 \\ \sin\frac{\pi}{4} & 0 & \cos\frac{\pi}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

The first rotation gives,

$$D_1 = S_1^{-1} A S_1$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & \sqrt{2} & 2 \\ \sqrt{2} & 3 & \sqrt{2} \\ 2 & \sqrt{2} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

We observe that the elements d13 and *d31* got annihilated. To make sure that calculations are correct up to this step, we see that the sum of the diagonal elements of *D1* is same as the sum of the diagonal elements of the original matrix *A*.
As a second step, we choose the largest off-diagonal element of *D1* and is found to be $d_{12} = d_{21} = 2,$ and compute

$$\tan 2\theta = \frac{2d_{12}}{d_{11} - d_{22}} = \frac{4}{0} = \infty$$

This again gives $\theta = \pi/4$
Thus, we construct the second rotation matrix as

$$S_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

At the end of the second rotation, we get

$$D_2 = S_2^{-1} D_1 S_2$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

This turned out to be a diagonal matrix, so we stop the computation. From here, we notice that the eigen values of the given matrix are 5,1 and –1. The eigenvectors are the column vectors of $S = S_1 S_2$

Therefore

$$S = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Example

Find all the eigen values of the matrix by Jacobi's method.

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

Solution

Here all the off-diagonal elements are of the same order of magnitude. Therefore, we can choose any one of them. Suppose, we choose *a12* as the largest element and compute

$$\tan 2\theta = \frac{-1}{0} = \infty$$

Which gives, $\theta = \pi/4$.

Then $\cos\theta = \sin\theta = 1/\sqrt{2}$

and we construct an orthogonal matrix *S1* such that

$$S = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The first rotation gives

$$D_1 = S_1^{-1} A S_1$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 3 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 \end{bmatrix}$$

Now, we choose $d_{13} = -1/\sqrt{2}$

As the largest element of D1 and compute

$$\tan 2\theta = \frac{2d_{13}}{d_{11} - d_{33}} = \frac{-\sqrt{2}}{1-2}$$

$\theta = 27^o 22' 41''$.

Now we construct another orthogonal matrix S2, such that

$$S_2 = \begin{bmatrix} 0.888 & 0 & -0.459 \\ 0 & 1 & 0 \\ 0.459 & 0 & 0.888 \end{bmatrix}$$

At the end of second rotation, we obtain

$$D_2 = S_2^{-1} D_1 S_2 = \begin{bmatrix} 0.634 & -0.325 & 0 \\ 0.325 & 3 & -0.628 \\ 0 & -0.628 & 2.365 \end{bmatrix}$$

Now, the numerically largest off-diagonal element of $D2$ is found to be $d_{23} = -0.628$ and compute.

$$\tan 2\theta = \frac{-2 \times 0.628}{3 - 2.365}$$

$\theta = -31^o 35' 24''$.

Thus, the orthogonal matrix is

$$S_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.852 & 0.524 \\ 0 & -0.524 & 0.852 \end{bmatrix}$$

At the end of third rotation, we get

$$D_3 = S_3^{-1} D_2 S_3 = \begin{bmatrix} 0.634 & -0.277 & 0 \\ 0.277 & 3.386 & 0 \\ 0 & 0 & 1.979 \end{bmatrix}$$

To reduce *D3* to a diagonal form, some more rotations are required. However, we may take 0.634, 3.386 and 1.979 as eigen values of the given matrix.

Example

Using Jacobi's method, find the eigenvalues and eigenvectors of the following matrix,

$$\begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}$$

Solution:

The given matrix is real and symmetric. The largest off-diagonal element is found to be

$$a_{12} = a_{21} = \frac{1}{2}$$

Now we comute

$$\tan 2\theta = \frac{2a_{ij}}{a_{ii} - a_{jj}} = \frac{2a_{12}}{a_{11} - a_{22}} = \frac{2\left(\frac{1}{2}\right)}{1 - \frac{1}{3}} = \frac{3}{2}$$

$$\theta = \frac{\tan^{-1}\left(\frac{3}{2}\right)}{2} = 28.155$$

Thus we construct an orthogonal matrix S_1 *as*

$$S_1 = \begin{bmatrix} \cos 28.155 & -\sin 28.155 & 0 \\ \sin 28.155 & \cos 28.155 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.882 & -0.472 & 0 \\ 0.472 & 0.882 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The first rotation gives, $D_1 = S_1^{-1} A S_1$

$$= \begin{bmatrix} 0.882 & 0.472 & 0 \\ -0.472 & 0.882 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \begin{bmatrix} 0.882 & -0.472 & 0 \\ 0.472 & 0.882 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1.268 & 0.000 & 0.412 \\ 0.000 & 0.066 & 0.063 \\ 0.412 & 0.063 & 0.200 \end{bmatrix}$$

We see that sum of the diagonal elements of D_1 =1.53

And the sum of the diagonal elements of A = 1.53
This means that our question is going right.

As a second step we choose the largest of off-diagonal element of D_1, which is $d_{13} = d_{31} = 0.412$, and comput

$$\tan 2\theta = \frac{2d_{13}}{d_{11} - d_{33}} = \frac{2(0.412)}{1.268 - 0.200} = 0.772$$

$$\theta = \frac{\tan^{-1}(0.772)}{2} = 18.834$$

Thus we construct an orthogonal matrix S_2 as

$$S_2 = \begin{bmatrix} \cos 18.834 & 0 & -\sin 18.834 \\ 0 & 1 & 0 \\ \sin 18.834 & 0 & \cos 18.834 \end{bmatrix} = \begin{bmatrix} 0.946 & 0 & -0.323 \\ 0 & 1 & 0 \\ 0.323 & 0 & 0.946 \end{bmatrix}$$

Thus the rotation gives,

$$D_2 = S_2^{-1} D_1 S_2$$

$$= \begin{bmatrix} 0.946 & 0 & 0.323 \\ 0 & 1 & 0 \\ -0.323 & 0 & 0.946 \end{bmatrix} \begin{bmatrix} 1.268 & 0.000 & 0.412 \\ 0.000 & 0.066 & 0.063 \\ 0.412 & 0.063 & 0.200 \end{bmatrix} \begin{bmatrix} 0.946 & 0 & -0.323 \\ 0 & 1 & 0 \\ 0.323 & 0 & 0.946 \end{bmatrix}$$

$$= \begin{bmatrix} 1.408 & 0.020 & -0.001 \\ 0.020 & 0.066 & 0.060 \\ -0.001 & 0.060 & 0.059 \end{bmatrix}$$

We again see that sum of the diagonal elements of D_2 =1.53
Also the sum of the diagonal elements of A = 1.53
This means that our question is going right.
Hence the eigenvalues are 1.408 , .066 and .059 and the corresponding eigenvectors are the columns of S.Where

$$S = S_1 S_2$$

$$= \begin{bmatrix} 0.882 & -0.472 & 0 \\ 0.472 & 0.882 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.946 & 0 & -0.323 \\ 0 & 1 & 0 \\ 0.323 & 0 & 0.946 \end{bmatrix}$$

$$= \begin{bmatrix} .8343 & -.472 & -.2848 \\ .446 & .88 & -.1524 \\ .323 & 0 & .946 \end{bmatrix}$$

Lecture 25
Inner Product Space

Inner Product Space

In mathematics, an inner product space is a vector space with the additional structure called an inner product. This additional structure associates each pair of vectors in the space with a scalar quantity known as the inner product of the vectors. Inner products allow the rigorous introduction of intuitive geometrical notions such as the length of a vector or the angle between two vectors. They also provide the means of defining orthogonality between vectors (zero inner product). Inner product spaces generalize Euclidean spaces (in which the inner product is the dot product, also known as the scalar product) to vector spaces of any (possibly infinite) dimension, and are studied in functional analysis.

Definition

An inner product on a vector space V is a function that to each pair of vectors u and v associates a real number $\langle u,v\rangle$ and satisfies the following axioms,

For all u, v, w in V and all scalars C:

1) $\langle u,v\rangle = \langle v,u\rangle$
2) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$
3) $\langle cu,v\rangle = c\langle u,v\rangle$
4) $\langle u,u\rangle \geq 0 \quad \textit{and} \quad \langle u,u\rangle = 0 \quad \textit{iff} \quad u=0$

A vector space with an inner product is called inner product space.

Example 1

Fix any two positive numbers say 4 & 5 and for vectors $u=\langle u_1,u_2\rangle$ and $v=\langle v_1,v_2\rangle$ in R^2 set

$$\langle u,v\rangle = 4u_1v_1 + 5u_2v_2$$

Show that it defines an inner product.

Solution

Certainly Axiom 1 is satisfied, because

$$\langle u,v\rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2 = \langle \mathbf{v},u\rangle.$$

If $\boldsymbol{w} = (w_1, w_2)$, then

$$\langle u+v,w\rangle = 4(u_1+v_1)w_1 + 5(u_2+v_2)w_2$$
$$= 4u_1w_1 + 5u_2w_2 + 4v_1w_1 + 5v_2w_2 = \langle u,w\rangle + \langle v,w\rangle$$

This verifies Axiom 2.

For Axiom 3, we have $\langle c\boldsymbol{u},\boldsymbol{v}\rangle = 4(cu_1)v_1 + 5(cu_2)v_2 = c(4u_1v_1+5u_2v_2) = c\langle \boldsymbol{u},\boldsymbol{v}\rangle$

For Axiom 4, note that $\langle u,u\rangle = 4u_1^2 + 5u_2^2 \geq 0$, and $4u_1^2 + 5u_2^2 = 0$ only if $u_1 = u_2 = 0$, that is, if $\boldsymbol{u} = \mathbf{0}$. Also, $\langle 0,0\rangle = 0$.

So $\langle u,v\rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2$ defines an inner product on $\boldsymbol{R^2}$.

Example 2

Let A be symmetric, positive definite $n \times n$ matrix and let u and v be vectors in $\Re^n$.
Show that $\langle u,v\rangle = u^t Av$ defines and inner product.

Solution

We check that

$$\langle u,v\rangle = u^t Av = u.Av = Av.u$$

$$= A^t v.u = \left(v^t A\right)^t u = v^t Au = \langle v,u\rangle$$

Also

$$\langle u,v+w\rangle = u^t A\left(v+w\right) = u^t Av + u^t Aw$$

$$= \langle u,v\rangle + \langle u,w\rangle$$

And

$$\langle cu,v\rangle = \left(cu\right)^t Av = c\left(u^t Av\right) = c\langle u,v\rangle$$

Finally since A is positive definite

$$\langle u,u\rangle = u^t Au > 0 \qquad \textit{for all } u \neq 0$$

$$So \quad \langle u,u\rangle = u^t Au = 0 \qquad \textit{iff} \quad u = 0$$

So $\langle u,v\rangle = u^t Av$ is an inner product space.

Example 3

Let $t_0, \ldots, t_n$ be distinct real numbers. For p and q in $\boldsymbol{P_n}$, define

$$\langle p,q\rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n)$$

Show that it defines inner product.

Solution

Certainly Axiom 1 is satisfied, because

$$\langle p,q\rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n)$$

$$= q(t_0)p(t_0) + q(t_1)p(t_1) + \cdots + q(t_n)p(t_n) = \langle q,p\rangle$$

If $\boldsymbol{r} = r(t_0) + r(t_1) + \cdots + r(t_n)$, then

$$\langle p+q,r\rangle = \left[p(t_0)+q(t_0)\right]r(t_0) + \left[p(t_1)+q(t_1)\right]r(t_1) + \cdots + \left[p(t_n)+q(t_n)\right]r(t_n)$$

$$= \left[p(t_0)r(t_0) + p(t_1)r(t_1) + \ldots + p(t_n)r(t_n)\right] + \left[q(t_0)r(t_0) + q(t_1)r(t_1) + \ldots + q(t_n)r(t_n)\right]$$

$$= \langle p,r\rangle + \langle q,r\rangle$$

This verifies Axiom 2.
For Axiom 3, we have

$$\langle cp,q\rangle = \left[cp(t_0)\right]q(t_0) + \left[cp(t_1)\right]q(t_1) + \cdots + \left[cp(t_n)\right]q(t_n)$$

$$= c\left[p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n)\right] = c\langle p,q\rangle$$

For Axiom 4, note that

$$\langle p,p\rangle = [p(t_0)]^2 + [p(t_1)]^2 + \cdots + [p(t_n)]^2 \geq 0$$

Also, $\langle \mathbf{0}, \mathbf{0} \rangle$=0. (We still use a boldface zero for the zero polynomial, the zero vector in $\boldsymbol{P_n}$.) If $\langle p, p \rangle$=0, then $\boldsymbol{p}$ must vanish at n + 1 points: t_0, …, t_n. This is possible only if $\boldsymbol{p}$ is the zero polynomial, because the degree of $\boldsymbol{p}$ is less than n + 1. Thus $\langle p, q \rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n)$ defines an inner product on $\boldsymbol{P_n}$.

Example 4

Compute $\langle p, q \rangle$ where p(t)= 4+t q(t) = 5-4t^2

Refer to P_2 with the inner product given by evaluation at -1, 0 and 1 in example 2.

Solution

$$P(-1) = 3\ , P(0) = 4\ ,\ P(1) = 5$$
$$q(-1) = 1\ ,\ q(0) = 5\ ,\ q(1) = 1$$

$$\begin{aligned}\langle p, q \rangle &= P(-1)q(-1) + P(0)q(0) + P(1)q(1)\\ &= (3)(1) + (4)(5) + (5)(1)\\ &= 3 + 20 + 5\end{aligned}$$

Example 5

Compute the orthogonal projection of q onto the subspace spanned by p, for p and q in the above example.

Solution

The orthogonal projection of q onto the subspace spanned by p

$$P(-1) = 3\ , P(0) = 4\ ,\ P(1) = 5$$
$$q(-1) = 1\ ,\ q(0) = 5\ ,\ q(1) = 1$$

$$q.p = 28 \qquad p.p = 50$$

$$\begin{aligned}\hat{q} = \frac{q.p}{p.p}\,p &= \frac{28}{50}(4+t)\\ &= \frac{56}{25} + \frac{14}{25}t\end{aligned}$$

Example 6

Let V be P_2 , with the inner product from example 2 where

$$t_0 = 0\ ,\ t_1 = \frac{1}{2}\ \ and\ \ t_2 = 1$$

Let $\quad p(t) = 12t^2 \quad and \quad q(t) = 2t - 1$

Compute $\langle p, q \rangle \ \ and \ \ \langle q, q \rangle$

Solution

$$\langle p,q\rangle = p(0)q(0) + p\left(\tfrac{1}{2}\right) + p(1)q(1)$$
$$=(0)(-1)+(3)(0)+(12)(1) = 12$$

$$\langle q,q\rangle = \left[q(0)\right]^2 + \left[q\left(\frac{1}{2}\right)\right]^2 + \left[q(1)\right]^2$$
$$=(-1)^2+(0)^2+(1)^2 = 2$$

Norm of a Vector

Let V be and inner product space with the inner product denoted by $\langle u,v\rangle$ just as in R^n , we define the length or norm of a vector V to be the scalar

$$\|v\| = \sqrt{\langle u,v\rangle} \quad or \quad \|v\|^2 = \langle u,v\rangle$$

1) A unit vector is one whose length is 1.
2) The distance between $u \,\&\, v$ is $\|u-v\|$ vectors $u \,\&\, v$ are orthogonal if $\langle u,v\rangle = 0$

Example 7

Compute the length of the vectors in example 3.

Solution

$$\|p\|^2 = \langle p,p\rangle = \left[p(0)\right]^2 + \left[p\left(\frac{1}{2}\right)\right]^2 + \left[p(1)\right]^2$$
$$=(0)^2+(3)^2+(12)^2 = 153$$
$$\|p\| = \sqrt{153}$$

In example 3 we found that

$$\langle q,q\rangle = 2$$
$$Hence \quad \|q\| = \sqrt{2}$$

Example 8

Let $\Re^2$ have the inner product of example 1 and let x=(1,1) and y=(5,-1)

a) Find $\|x\|$, $\|y\| and \left|\langle x,y\rangle\right|^2$ b) Describe all vectors (z_1,z_2) that are orthogonal to y.

Solution

a) We have x=(1,1) and y=(5,-1)

And $\langle x,y\rangle = 4x_1y_1 + 5x_2y_2$

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{4(1)(1)+5(1)(1)}$$
$$= \sqrt{4+5} = \sqrt{9} = 3$$
$$\|y\| = \sqrt{\langle y, y\rangle} = \sqrt{4(5)(5)+5(-1)(-1)}$$
$$= \sqrt{100+5} = \sqrt{105}$$

$$\left|\langle x, y\rangle\right|^2 = \langle x, y\rangle\langle x, y\rangle$$
$$= \left[4(1)(5)+5(1)(-1)\right]^2$$
$$= \left[20-5\right]^2$$
$$= \left[15\right]^2 = 225$$

b) All vectors $z = (z_1, z_2)$ orthogonal to y=(5,-1)

$$< y, z >= 0$$
$$4(5)(z_1) + 5(-1)(z_2) = 0$$
$$20z_1 - 5z_2 = 0$$
$$4z_1 - z_2 = 0$$
$$\begin{bmatrix} 4 & -1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 0$$

So all multiples of $\begin{bmatrix} 4 \\ -1 \end{bmatrix}$ are orthogonal to y.

Example 9

Le V be P_4 with the inner product in example 2 involving evaluation of polynomials at -2,-1,0,1,2 and view P_2 as a subspace of V. Produce an orthogonal basis for P_2 by applying the Gram Schmidt process to the polynomials $1, t \,\&\, t^2$.

Solution

Given polynomials **1** ***t*** ***t*²** at -2,-1, 0, 1 and 2

Polynomial: **1** ***t*** ***t*²**

Vector of values: $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix}$

The inner product of two polynomials in ***V*** equals the (standard) inner product of their corresponding vectors in $\boldsymbol{R}^5$. Observe that ***t*** is orthogonal to the constant function **1**. So take $\boldsymbol{p_0}(t) = 1$ and $\boldsymbol{p_1}(t) = t$. For p_2, use the vectors in $\boldsymbol{R^5}$ to compute the projection of $\boldsymbol{t^2}$ onto Span $\{\boldsymbol{p_0}, \boldsymbol{p_1}\}$:

$$\langle t^2, p_0\rangle = \langle t^2, 1\rangle = 4+1+0+1+4 = 10$$

$$\langle p_0, p_0\rangle = 5$$

$$\langle t^2, p_1\rangle = \langle t^2, t\rangle = -8+(-1)+0+1+8 = 0$$

The orthogonal projection of $\boldsymbol{t^2}$ onto Span $\{\mathbf{1}, \boldsymbol{t}\}$ is $\frac{10}{5}p_0 + 0p_1$. Thus

$$p_2(t) = t^2 - 2p_0(t) = t^2 - 2$$

An orthogonal basis for the subspace $\boldsymbol{P_2}$ of $\boldsymbol{V}$ is:

$$\begin{array}{lccc} \text{Polynomial:} & \boldsymbol{p_0} & \boldsymbol{p_1} & \boldsymbol{p_2} \\ \text{Vector of values:} & \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, & \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}, & \begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix} \end{array}$$

Best Approximation in Inner Produce Spaces

A common problem in applied mathematics involves a vector space V whose elements are functions. The problem is to approximate a function f in V by a function g from a specified subspace W of V. The "closeness" of the approximation of f depends on the way $\|f-g\|$ is defined. We will consider only the case in which the distance between f and g is determined by an inner product. In this case the best approximation to f by functions in W is the orthogonal projection of f onto the subspace W.

Example 10

Let V be P_4 with the inner product in example 5 and let $P_0, P_1 \,\&\, P_2$ be the orthogonal basis for the subspace P_2, find the best approximation to $p(t) = 5 - \frac{1}{2}t^4$ by polynomials in P_2.

Solution:

The values of $\boldsymbol{p_0}$, $\boldsymbol{p_1}$, and $\boldsymbol{p_2}$ at the numbers – 2, –1, 0, 1, and 2 are listed in $\boldsymbol{R^5}$ vectors in

Polynomial: $\boldsymbol{p_0}$ $\boldsymbol{p_1}$ $\boldsymbol{p_2}$

$$\text{Vector of values: } \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}, \begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}$$

The corresponding values for ***p*** are: $-3, 9/2, 5, 9/2,$ and –3.

We compute

$$\langle p, p_0\rangle = 8 \qquad \langle p, p_1\rangle = 0 \qquad \langle p, p_2\rangle = -31$$
$$\langle p_0, p_0\rangle = 5, \qquad \langle p_2, p_2\rangle = 14$$

Then the best approximation in ***V*** to ***p*** by polynomials in $\boldsymbol{P_2}$ is

$$\hat{p} = proj_{p_2} p = \frac{\langle p, p_0\rangle}{\langle p_0, p_0\rangle} p_0 + \frac{\langle p, p_1\rangle}{\langle p_1, p_1\rangle} p_1 + \frac{\langle p, p_2\rangle}{\langle p_2, p_2\rangle} p_2$$
$$= \tfrac{8}{5} p_0 + \tfrac{-31}{14} p_2 = \tfrac{8}{5} - \tfrac{31}{14}(t^2 - 2).$$

This polynomial is the closest to ***P*** of all polynomials in $\boldsymbol{P_2}$, when the distance between polynomials is measured only at –2, –1, 0, 1, and 2.

Cauchy – Schwarz Inquality

$$For\ all \quad u, v\ in\ V$$
$$|\langle u, v\rangle| \le \|u\|\,\|v\|$$

Triangle Inequality

$$For\ all\ \ u, v\ in\ V$$
$$\|u+v\| \le \|u\| + \|v\|$$

Proof

$$\|u+v\|^2 = \langle u+v, u+v\rangle$$
$$= \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle$$
$$\le \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2$$
$$\le \|u\|^2 + 2\|u\|\|v\| + \|v\|^2$$
$$\|u+v\|^2 = (\|u\| + \|v\|)^2$$

$$\Rightarrow \quad \|u+v\| = \|u\| + \|v\|$$

Inner product for $C[a,b]$

Probably the most widely used inner product space for applications is the vector space $C[a,b]$ of all continuous functions on an interval $a \le t \le b$, with an inner product that will describe.

Example 11

For f, g in $C[a,b]$, set

$$\langle f,g\rangle = \int_b^a f(t)\, g(t)\, dt$$

Show that it defines an inner product on $C[a,b]$.

Solution

Inner product Axioms 1 to 3 follow from elementary properties of definite integrals

1. $\langle f,g\rangle = \langle g,f\rangle$
2. $\langle f+h,g\rangle = \langle f,g\rangle + \langle h,g\rangle$
3. $\langle cf,g\rangle = c\langle f,g\rangle$

For Axiom 4, observe that

$$\langle f,f\rangle = \int_a^b [f(t)]^2\, dt \ge 0$$

The function $[\boldsymbol{f}(t)]^2$ is continuous and nonnegative on $[a, b]$. If the definite integral of $[\boldsymbol{f}(t)]^2$ is zero, then $[\boldsymbol{f}(t)]^2$ must be identically zero on $[a, b]$, by a theorem in advanced calculus, in which case $\boldsymbol{f}$ is the zero function. Thus $\langle \boldsymbol{f}, \boldsymbol{f}\rangle = 0$ implies that $\boldsymbol{f}$ is the zero function of $[a, b]$.

So $\langle \boldsymbol{f},\boldsymbol{g}\rangle = \int_a^b \boldsymbol{f}(t)\boldsymbol{g}(t)\,dt$ defines an inner product on $\boldsymbol{C}[a, b]$.

Example 12

Compute $\langle f,g\rangle$ where $f(t)=1-3t^2$ *and* $g(t)=t-t^3$ *on* $v=C[0,1]$.

Solution

Let V be the space $C[a,b]$ with the inner product $\langle f,g\rangle = \int_b^a f(t)\, g(t)\, dt$

$$f(t)=1-3t^2 \quad , \quad g(t)=t-t^3$$

$$\begin{aligned}\langle f,g\rangle &= \int_0^1 (1-3t^2)\left(t-t^3\right)dt \\ &= \int_0^1 \left(3t^5-4t^3+t\right)dt \\ &= \left|\frac{1}{2}t^6 - t^4 + \frac{1}{2}t^2\right|_0^1 \\ &= 0\end{aligned}$$

Example 13

Let V be the space $C[a,b]$ with the inner product $\langle f,g\rangle=\int_b^a f(t)\,g(t)\,dt$

Let W be the subspace spanned by the polynomials $P_1(t)=1$, $P_2(t)=2t-1$ & $P_3(t)=12t^2$

Use the Gram – Schmidt process to find an orthogonal basis for W.

Solution

Let $\boldsymbol{q_1} = \boldsymbol{p_1}$, and compute

$$\langle \boldsymbol{p_2},\boldsymbol{q_1}\rangle = \int_0^1 (2t-1)(1)dt = (t^2-t)\Big|_0^1 = 0$$

So $\boldsymbol{p_2}$ is already orthogonal to $\boldsymbol{q_1}$, and we can take $\boldsymbol{q_2} = \boldsymbol{p_2}$. For the projection of $\boldsymbol{p_3}$ onto $\boldsymbol{W_2}$ = Span $\{\boldsymbol{q_1}, \boldsymbol{q_2}\}$, we compute

$$\langle \boldsymbol{p_3},\boldsymbol{q_1}\rangle = \int_0^1 12t^2\cdot 1dt = 4t^3\Big|_0^1 = 4$$

$$\langle \boldsymbol{q_1},\boldsymbol{q_1}\rangle = \int_0^1 1\cdot 1dt = t\Big|_0^1 = 1$$

$$\langle \boldsymbol{p_3},\boldsymbol{q_2}\rangle = \int_0^1 12t^2(2t-1)dt = \int_0^1 (24t^3-12t^2)dt = 2$$

$$\langle \boldsymbol{q_2},\boldsymbol{q_2}\rangle = \int_0^1 (2t-1)^2 dt = \frac{1}{6}(2t-1)^3\Big|_0^1 = \frac{1}{3}$$

Then $$\text{proj}_{w_2}\boldsymbol{p_3} = \frac{\langle \boldsymbol{p_3},\boldsymbol{q_1}\rangle}{\langle \boldsymbol{q_1},\boldsymbol{q_1}\rangle}\boldsymbol{q_1} + \frac{\langle \boldsymbol{p_3},\boldsymbol{q_2}\rangle}{\langle \boldsymbol{q_2},\boldsymbol{q_2}\rangle}\boldsymbol{q_2} = \frac{4}{1}\boldsymbol{q_1} + \frac{2}{1/3}\boldsymbol{q_2} = 4\boldsymbol{q_1} + 6\boldsymbol{q_2}$$

And $$\boldsymbol{q_3} = \boldsymbol{p_3} - \text{proj}_{w_2}\boldsymbol{p_3} = \boldsymbol{p_3} - 4\boldsymbol{q_1} - 6\boldsymbol{q_2}$$

As a function, $\boldsymbol{q_3}(t) = 12\boldsymbol{t}^2 - \mathbf{4} - 6(2\boldsymbol{t} - \mathbf{1}) = 12\boldsymbol{t}^2 - 12\boldsymbol{t} + \mathbf{2}$. The orthogonal basis for the subspace $\boldsymbol{W}$ is $\{\boldsymbol{q_1}, \boldsymbol{q_2}, \boldsymbol{q_3}\}$

Exercises

Let $\Re^2$ have the inner product of example 1 and let x=(1,1) and y=(5,-1)

a) Find $\|x\|$, $\|y\|$ and $|\langle x,y\rangle|^2$ b) Describe all vectors (z_1,z_2) that are orthogonal to y.

2) Let $\Re^2$ have the inner product of Example 1. Show that the Cauchy-Shwarz inequality holds for x=(3,-2) and y=(-2,1)

Exercise 3-8 refer to P_2 with the inner product given by evaluation at -1,0 and 1 in example 2.

3) Compute $\langle p,q\rangle$ where p(t)= 4+t q(t) = 5-4t^2

4) Compute $\langle p,q\rangle$ where p(t)= 3t - t^2 q(t) = 3 + t^2

5) Compute $\|P\|$ *and* $\|q\|$ for p and q in exercise 3.

6) Compute $\|P\|$ *and* $\|q\|$ for p and q in exercise 4.

7) Compute the orthogonal projection of q onto the subspace spanned by p, for p and q in Exercise 3.

8) Compute the orthogonal projection of q onto the subspace spanned by p, for p and q in Exercise 4.

9) Let P^3 have the inner product given by evaluation at -3,-1,1, and 3. Let $p_o(t)=1$, $p_1(t)=t, \quad and \quad p_2(t)=t^2$

a)Computer the orthogonal projection of P_2 on to the subspace spanned by P_0 and P_1. b) Find a polynomial q that is orthogonal to P_0 and P_1 such tha $\{p_0, p_1, q\}$ is an orthogonal basis for span $\{p_0, p_1, q\}$. Scale the polynomial q so that its vector of values at (-3,-1,1,3) is (1,-1,-1,1)

10) Let P^3 have the inner product given by evaluation at -3,-1,1, and 3. Let $p_o(t)=1$, $p_1(t)=t, \quad and \quad p_2(t)=t^2$

Find the best approximation to $p(t)=t^3$ by polynomials in Span $\{p_0, p_1, q\}$.

11) Let p_0, p_1, p_2 be the orthogonal polynomials described in example 5, where the inner product on P_4 is given by evaluation at -2, -1, 0, 1, and 2. Find the orthogonal projection of t^3 onto Span $\{p_0, p_1, p_2\}$

12) Compute $\langle f, g\rangle$ where $f(t)=1-3t^2 \; and \; g(t)=t-t^3 \; on \; v=C[0,1]$.

13) Compute $\langle f, g\rangle$ where

$f(t)=5t-3 \; and \; g(t)=t^3-t^2 \; on \; v=C[0,1]$.

14) Compute $\|f\|$ for f in exercise 12.

15) Compute $\|g\|$ for g in exercise 13.

16) Let V be the space C[-2,2] with the inner product of Example 7. Find an orthogonal basis for the subspace spanned by the polynomials $1, t, t^2$.

17)

Let $u=\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $v=\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be two vectors in R^2. Show that $\langle u, v\rangle=2u_1v_1+3u_2v_2$ defines an inner product.

Lecture 26
Application of inner product spaces

Definition

An inner product on a vector space *V* is a function that associates to each pair of vectors *u* and *v* in *V*, a real number $\langle u, v\rangle$ and satisfies the following axioms, for all *u, v, w* in *V* and all scalars *c*:

1. $\langle u, v\rangle = \langle v, u\rangle$
2. $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
3. $\langle cu, v\rangle = c\langle u, v\rangle$
4. $\langle u, u\rangle \geq 0$ *and* $\langle u, u\rangle = 0$ *iff* $u = 0$.

A vector space with an inner product is called an inner product space.

Least Squares Lines

The simplest relation between two variables x and y is the linear equation $y = \beta_0 + \beta_1 x$. Often experimental data produces points $(x_1, y_1), \ldots, (x_n, y_n)$ that when graphed, seem to lie close to a line. Actually we want to determine the parameters β_0 and β_1 that make the line as "close" to the points as possible. There are several ways to measure how close the line is to the data. The usual choice is to add the squares of the residuals. The least squares line is the line $y = \beta_0 + \beta_1 x$ that minimizes the sum of the squares of the residuals.

If the data points are on the line, the parameters β_0 and β_1 would satisfy the equations

predicted value	*Observed value*
$\beta_0 + \beta_1 x_1$	$= y_1$
$\beta_0 + \beta_1 x_2$	$= y_2$
.	.
.	.
.	.
$\beta_0 + \beta_1 x_n$	$= y_n$

We can write this system as

$$X\beta = y$$

$$Where \;\; X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix} \;\;,\;\; \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \;\;,\;\; y = \begin{bmatrix} y_0 \\ y_1 \\ . \\ . \\ . \\ y_n \end{bmatrix}$$

Computing the least-squares solution of $X\beta = y$ is equivalent to finding the β that determines the least-squares line.

Example 1

Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points (2, 1), (5, 2), (7, 3), (8, 3).

Solution

$$X\beta = y$$

$$Here \;\; X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \;\;,\;\; \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \;\;,\;\; y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

For the least-squares solution of $x\beta = y$, obtain the normal equations(with the new notation) :

$$X^T X \hat{\beta} = X^T y$$

i.e, compute

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

The normal equations are

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Hence,

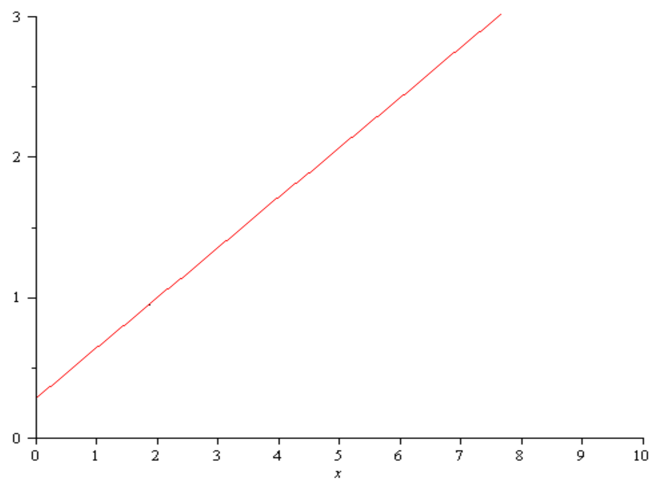
$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1}\begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

$$= \frac{1}{84}\begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix}\begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

$$= \frac{1}{84}\begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

Thus, the least -squares line has the equation

$$y = \frac{2}{7} + \frac{5}{14}x$$



Weighted Least-Squares

Let y be a vector of n observations, $y_1, y_2, \ldots, y_n$ and suppose we wish to approximate y by a vector $\hat{y}$ that belongs to some specified subspace of $\mathbb{R}^n$ (as discussed previously that $\hat{y}$ is written as Ax so that $\hat{y}$ was in the column space of A).Now suppose the approximating vector $\hat{y}$ is to be constructed from the columns of matrix A. Then we find an $\hat{x}$ that makes $A\hat{x} = \hat{y}$ as close to y as possible. So that measure of closeness is the weighted error

$$\|Wy - W\hat{y}\|^2 = \|Wy - WA\hat{y}\|^2$$

Where W is the diagonal matrix with (positive) $w_1,...,w_n$ on its diagonal, that is

$$W = \begin{bmatrix} w_1 & 0 & \cdots & & 0 \\ 0 & w_2 & & & \\ . & & . & & . \\ . & & & . & . \\ . & & & . & . \\ 0 & & \cdots & & w_n \end{bmatrix}$$

Thus, $\hat{x}$ is the ordinary least-squares solution of the equation

$$WAx = Wy$$

The normal equation for the weighted least-squares solution is

$$(WA)^T WAX = (WA)^T Wy$$

Example 2

Find the least squares line $y = \beta_0 + \beta_1 x$ that best fits the data (–2, 3), (–1, 5), (0, 5), (1, 4), (2, 3). Suppose that the errors in measuring the *y*-values of the last two data points are greater than for the other points. Weight this data half as much as the rest of the data.

Solution

Write X, β *and* y

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \ y = \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix}$$

For a weighting matrix, choose W with diagonal entries 2 , 2 , 2 , 1 and 1.
Left-multiplication by W scales the rows of X and y:

$$WX = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$WX = \begin{bmatrix} 2 & -4 \\ 2 & -2 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \ Wy = \begin{bmatrix} 6 \\ 10 \\ 10 \\ 4 \\ 3 \end{bmatrix}$$

For normal equation, compute

$$(WX)^T WX = \begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix}, \ and \ (WX)^T Wy = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

And solve

$$\begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix}^{-1} \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{269} \begin{bmatrix} 25 & 9 \\ 9 & 14 \end{bmatrix} \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{269} \begin{bmatrix} 25 & 9 \\ 9 & 14 \end{bmatrix} \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{269} \begin{bmatrix} 1169 \\ 55 \end{bmatrix} = \begin{bmatrix} 4.3 \\ 0.2 \end{bmatrix}$$

Therefore, the solution to two significant digits is $\beta_0 = 4.3 \, and \ \beta_1 = 0.20$.

Hence the required line is $y = 4.3 + 0.2x$

In contrast, the ordinary least-squares line for this data can be found as**:**

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 20 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 20 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 20 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{50} \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 20 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{50} \begin{bmatrix} 200 \\ -5 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -0.1 \end{bmatrix}$$

Hence the equation of least-squares line is

$$y = 1.0 - 0.1x$$

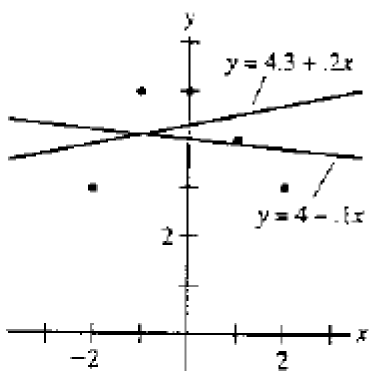


FIGURE 1 Weighted and ordinary least-squares lines.

<u>What Does Trend Analysis Mean?</u>

An aspect of technical analysis that tries to predict the future movement of a stock based on past data. Trend analysis is based on the idea that what has happened in the past gives traders an idea of what will happen in the future.

Linear Trend

A first step in analyzing a time series, to determine whether a linear relationship provides a good approximation to the long-term movement of the series computed by the method of semi averages or by the method of least squares.

Note

The simplest and most common use of trend analysis occurs when the points $t_0, t_1, \ldots, t_n$ can be adjusted so that they are evenly spaced and sum to zero.

Example

Fit a quadratic trend function to the data (-2,3), (-1,5), (0,5), (1,4) and (2,3)

Solution

The t-coordinates are suitably scaled to use the orthogonal polynomials found in Example 5 of the last lecture. We have

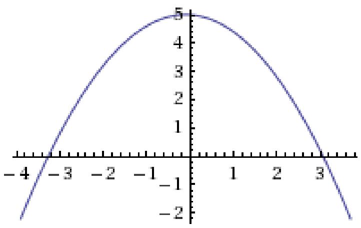
$$\begin{array}{lcccc} \text{Polynomial:} & \text{p}_0 & \text{p}_1 & \text{p}_2 & \text{data: g} \\ \textit{Vector of values:} & \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, & \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}, & \begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}, & \begin{bmatrix}3\\5\\5\\4\\3\end{bmatrix} \end{array}$$

$$\hat{p} = \frac{<g, p_0>}{<p_0, p_0>} p_0 + \frac{<g, p_1>}{<p_1, p_1>} p_1 + \frac{<g, p_2>}{<p_2, p_2>} p_2$$

$$= \frac{20}{5} p_0 - \frac{1}{10} p_1 - \frac{7}{14} p_2$$

$$and \qquad \hat{p}(t) = 4 - 0.1t - 0.5(t^2 - 2)$$

Since, the coefficient of p_2 is not extremely small, it would be reasonable to conclude that the trend is at least quadratic.



Above figure shows that approximation by a quadratic trend function

Fourier series

If f is a 2π -periodic function then

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos mt + b_m \sin mt)$$

is called Fourier series of f where

$$a_m = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos mt \, dt \qquad \text{and}$$

$$b_m = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin mt \, dt$$

Example

Let $C[0, 2\pi]$ has the inner product

$$< f, g >= \int_0^{2\pi} f(t) g(t) \, dt$$

and let m and n be unequal positive integers. Show that $\cos mt$ and $\cos nt$ are orthogonal.

Solution

When $m \neq n$

$$< \cos mt, \cos nt >= \int_0^{2\pi} \cos mt \cos nt \, dt$$

$$= \frac{1}{2} \int_0^{2\pi} [\cos(mt + nt) + \cos(mt - nt) \, dt$$

$$= \frac{1}{2}\left[\frac{\sin(mt+nt)}{m+n} + \frac{\sin(mt-nt)}{m-n}\right]_0^{2\pi}$$
$$= 0$$

Example

Find the nth-order Fourier approximation to the function

$f(t) = t$ on the interval $[0, 2\pi]$.

Solution

We compute

$$\frac{a_0}{2} = \frac{1}{2}.\frac{1}{\pi}\int_0^{2\pi} t\,dt = \frac{1}{2\pi}\left[\frac{1}{2}t^2\Big|_0^{2\pi}\right] = \pi$$

and for k>0, using integration by parts,

$$a_k = \frac{1}{\pi}\int_0^{2\pi} t\cos kt\,dt = \frac{1}{\pi}\left[\frac{1}{k^2}\cos kt + \frac{t}{k}\sin kt\right]_0^{2\pi} = 0$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} t\sin kt\,dt = \frac{1}{\pi}\left[\frac{1}{k^2}\sin kt - \frac{t}{k}\cos kt\right]_0^{2\pi} = -\frac{2}{k}$$

Thus, the nth-order Fourier approximation of $f(t) = t$ is

$$\pi - 2\sin t - \sin 2t - \frac{2}{3}\sin 3t - \cdots - \frac{2}{n}\sin nt$$

The norm of the difference between f and a Fourier approximation is called the mean square error in the approximation.
It is common to write

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty}(a_m \cos mt + b_m \sin mt)$$

This expression for f (t) is called the Fourier series for f on $[0, 2\pi]$. The term $a_m \cos mt$, for example, is the projection of f onto the one-dimensional subspace spanned by $\cos mt$.

Example

Let $q_1(t) = 1,\ q_2(t) = t, and\ q_3(t) = 3t^2 - 4$ .Verify that $\{q_1, q_2, q_3\}$ is an orthogonal set in C{-2,2] with the inner product

$$< f, g >= \int_a^b f(t)\, g(t)\, dt$$

Solution:

$$< q_1, q_2 >= \int_{-2}^{2} 1.t\, dt = \frac{1}{2} t^2\, dt \Big|_{-2}^{2} = 0$$

$$< q_1, q_3 >= \int_{-2}^{2} 1.(3t^2 - 4)\, dt = (t^2 - 4t) \Big|_{-2}^{2} = 0$$

$$< q_2, q_3 >= \int_{-2}^{2} t.(3t^2 - 4)\, dt = (\frac{3}{4} t^4 - 2t^2) \Big|_{-2}^{2} = 0$$

Exercise

1. Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points (0, 1), (1, 1), (2, 2), (3, 2).
2. Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points (-1, 0), (0, 1), (1, 2,),(2, 4).
3. Find the least-squares line $y = \beta_0 + \beta_1 x$ that best fits the data (-2, 0), (-1, 0), (0, 2,),(1, 4),(2, 4), assuming that the first and last data points are less reliable. Weight them half as much as the three interior points.

4: To make a trend analysis of six evenly spaced data points, one can use orthogonal polynomials with respect to evaluation at the points t=-5, -3, -1, 1, 3and 5

(a). Show that the first three orthogonal polynomials are

$$p_0(t) = 1, \qquad p_1(t) = t, \qquad and \qquad p_2(t) = \frac{3}{8} t^2 - \frac{35}{8}$$

(b) Fit a quadratic trend function to the data (-5, 1), (-3, 1), (-1, 4), (1, 4), (3, 6), (5, 8)

5: For the space $C[0, 2\pi]$ with the inner product defined by

$$< f, g >= \int_0^{2\pi} f(t)\, g(t)\, dt$$

(a) Show that $\sin mt$ and $\sin nt$ are orthogonal when $m \neq n$

(b) Find the third–order Fourier approximation to $f(t) = 2\pi - t$

(c) Find the third order Fourier approximation to $\cos^3 t$, without performing any integration calculations.

6: Find the first-order and third order Fourier approximations to

$f(t) = 3 - 2\sin t + 5\sin 2t - 6\cos 2t$

Lecture 27

Householder's Method and QR Algorithum

PPT's slides are available in VULMS/downloads

Lecture 28

Singular Value Decomposition

PPT's slides are available in VULMS/downloads

Lecture 29

Fixed Points for Functions of Several Variables

PPT's slides are available in VULMS/downloads

Lecture 30

Newton's Method

PPT's slides are available in VULMS/downloads

Lecture 31

Quasi-Newton Method

PPT’s slides are available in VULMS/downloads