Linear Algebra

MTH 501

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. Mayor 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

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \ldots & \vdots & \ldots & \vdots \\ \ldots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

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

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

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1p} + a_{12}b_{2p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & \cdots & a_{21}b_{1p} + a_{22}b_{2p} + \cdots + a_{2n}b_{np} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1p} + a_{m2}b_{2p} + \cdots + a_{mn}b_{np} \end{bmatrix}$$

$$= \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)_{n \times p}$$

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: $(\underline{A}\underline{B}\underline{C})^T=\underline{C}^T\underline{B}^T\underline{A}^T$ $(k\underline{A})^T=k\underline{A}^T$
- 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 $B = \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix}$ is multiplicative inverse of A because AB = $\begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix}\begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ =I

Similarly we can check that BA = I

Singular and Non-Singular Matrices A square matrix A is said to be a **non-singular** matrix if $\det(A) \neq 0$, otherwise the square matrix A is said to be **singular**. Thus for a singular matrix A we must have $\det(A) = 0$

Example: $A = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 1 & 0 \\ 2 & -3 & 5 \end{bmatrix}$

$$|A| = 2(5-0) - 3(5-0) - 1(-3-2)$$
$$= 10 - 15 + 5 = 0$$

which means that A is singular.

Minor of an element of a matrix

Let A be a square matrix of order n x n. Then minor M_{ij} of the element $a_{ij} \in A$ is the determinant of $(n-1)\times(n-1)$ matrix obtained by deleting the *ith* row and *jth* column from A.

Example If $A = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 1 & 0 \\ 2 & -3 & 5 \end{bmatrix}$ is a square matrix. The Minor of $3 \in A$ is denoted by M_{12} and is defined to be $M_{12} = \begin{vmatrix} 1 & 0 \\ 2 & 5 \end{vmatrix}$ = 5-0 = 5

Cofactor of an element of a matrix

Let A be a non singular matrix of order $n \times n$ and let C_{ij} denote the cofactor (signed minor) of the corresponding entry $a_{ij} \in A$, then it is defined to be $C_{ij} = (-1)^{i+j} M_{ij}$

Example If $A = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 1 & 0 \\ 2 & -3 & 5 \end{bmatrix}$ is a square matrix. The cofactor of $3 \in A$ is denoted by C_{12} and is defined to be $C_{12} = = (-1)^{1+2} \begin{vmatrix} 1 & 0 \\ 2 & 5 \end{vmatrix} =$ - (5 - 0) = -5

Theorem If A is a square matrix of order $n \times n$ then the matrix has a multiplicative inverse A^{-1} if and only if the matrix A is non-singular.

Theorem Then inverse of the matrix A is given by $A^{-1} = \frac{1}{\det(A)}(C_{ij})^{tr}$

1. For further reference we take $n = 2$ so that A is a 2×2 non-singular matrix given by

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Therefore $C_{11} = a_{22}$, $C_{12} = -a_{21}$, $C_{21} = -a_{12}$ and $C_{22} = a_{11}$. So that

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}^{tr} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

2. For a 3×3 non-singular matrix A= $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$$C_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, C_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad \mathrm{C}_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \text{and so on.}$$

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

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

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

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 $\dfrac{dX}{dt}$

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

$$\begin{array}{ccccccc} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \tag{3}$$

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:

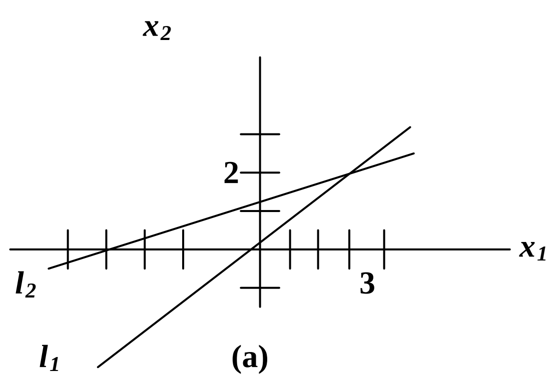
1. The lines may be parallel and distinct, in which case there is no intersection and consequently ***no solution***.
2. The lines may intersect at only one point, in which case the system has exactly ***one solution***.
3. The lines may coincide, in which case there are infinitely many points of intersection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

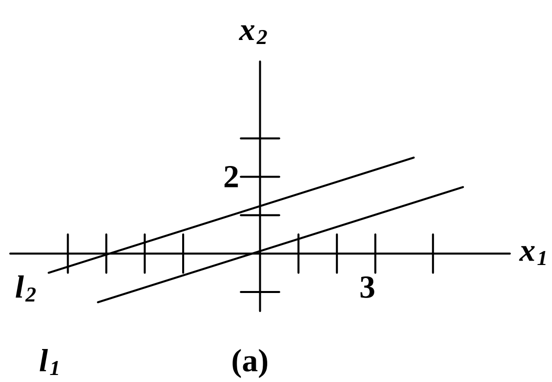
Example consider the system of linear equations in two variables $x_1 - 2x_2 = -1, \; -x_1 + 3x_2 = 3$
Solve the equation simultaneously:
Adding both equations we get $x_2 = 2$, Put $x_2 = 2$ in any one of the above equation we get $x_1 = 3$. So the solution is the single point **(3, 2)**. See the graph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

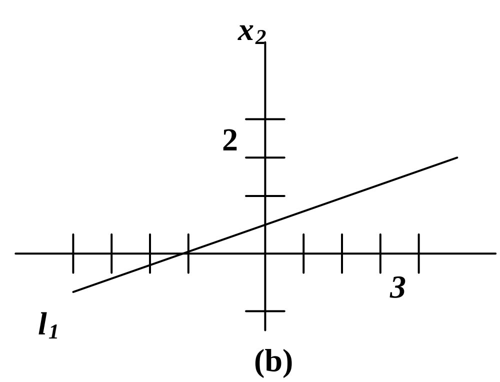


This system has exactly one solution

See the graphs to the following linear systems:

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





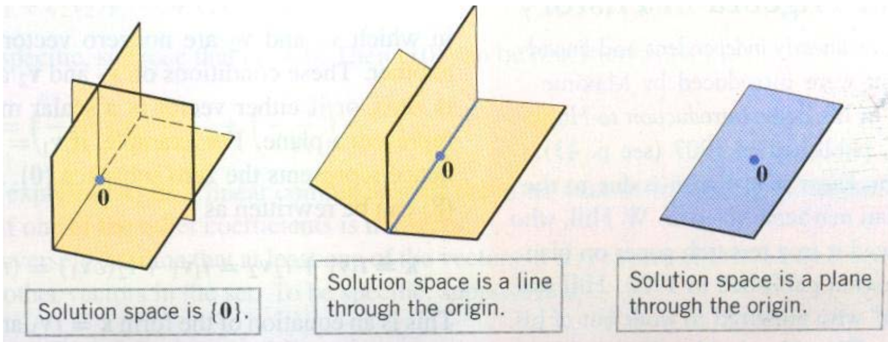
(a) No solution. (b) Infinitely many solutions.

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

$$x - y = 1$$
$$2x + y = 6$$

Solution

Adding both equations, we get $x = \frac{7}{3}$. Putting this value of x in 1st equation, we get $y = \frac{4}{3}$. Thus, the system has the ***unique solution*** $x = \frac{7}{3}, y = \frac{4}{3}$.

Geometrically, this means that the lines represented by the equations in the system intersect at a single point $\left(\frac{7}{3}, \frac{4}{3}\right)$ and thus has a unique solution.

Example 2 Solve the linear system

$$x + \quad y = 4$$
$$3x + 3y = 6$$

Solution

Multiply first equation by 3 and then subtract the second equation from this. We obtain

$$0 = 6$$

This equation is contradictory.

Geometrically, this means that the lines corresponding to the equations in the original system are parallel and distinct. So the given system has ***no solution.***

Example 3 Solve the linear system

$$\begin{aligned} 4x - 2y &= 1 \\ 16x - 8y &= 4 \end{aligned}$$

Solution

Multiply the first equation by -4 and then add in second equation.

$$\begin{array}{r} -16x + 8y = -4 \\ 16x - 8y = \ \ 4 \\ \hline 0 \qquad = 0 \end{array}$$

Thus, the solutions of the system are those values of *x* and *y* that satisfy the single equation $4x - 2y = 1$

Geometrically, this means the lines corresponding to the two equations in the original system coincide and thus the system has infinitely many solutions.

Parametric Representation

It is very convenient to describe the solution set in this case is to ***express it parametrically.*** We can do this by letting $y = t$ and solving for *x* in terms of *t*, or by letting $x = t$ and solving for *y* in terms of *t*.

The first approach yields the following parametric equations (by taking *y=t in the equation* $4x - 2y = 1$)

$$4x - 2t = 1, \quad y = t$$

$$x = \frac{1}{4} + \frac{1}{2}t, \quad y = t$$

We can now obtain some solutions of the above system by substituting some numerical values for the parameter.

Example For $t = 0$ the solution is $(\frac{1}{4}, 0)$. For $t = 1$, the solution is $(\frac{3}{4}, 1)$ and for $t = -1$ the solution is $(-\frac{1}{4}, -1)$ *etc.*

Example 4 Solve the linear system

$$\begin{aligned} x - \ \ y + 2z &= 5 \\ 2x - 2y + 4z &= 10 \\ 3x - 3y + 6z &= 15 \end{aligned}$$

Solution
Since the second and third equations are multiples of the first.

Geometrically, this means that the three planes coincide and those values of *x*, *y* and *z* that satisfy the equation $x - y + 2z = 5$ automatically satisfy all three equations.

We can express the solution set parametrically as

$$x = 5 + t_1 - 2t_2, y = t_1, \quad z = t_2$$

Some solutions can be obtained by choosing some numerical values for the parameters.

For example if we take $y = t_1 = 2$ and $z = t_2 = 3$ then

$$\begin{aligned} x &= 5 + t_1 - 2t_2 \\ &= 5 + 2 - 2(3) \\ &= 1 \end{aligned}$$

Put these values of x, y, and z in any equation of linear system to verify

$$\begin{aligned} x - y + 2z &= 5 \\ 1 - 2 + 2(3) &= 5 \\ 1 - 2 + 6 &= 5 \\ 5 &= 5 \end{aligned}$$

Hence x = 1, y = 2, z = 3 is the solution of the system. Verified.

Matrix Notation

The essential information of a linear system can be recorded compactly in a rectangular array called a **matrix**.

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

With the coefficients of each variable aligned in columns, the matrix $\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$

is called the coefficient matrix (or matrix of coefficients) of the system.

An augmented matrix of a system consists of the coefficient matrix with an added column containing the constants from the right sides of the equations. It is always denoted by A_b

$$A_b = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

Solving a Linear System

In order to solve a linear system, we use a number of methods. 1st of them is given below.

Successive elimination method In this method the x_1 term in the first equation of a system is used to eliminate the x_1 terms in the other equations. Then we use the x_2 term in the second equation to eliminate the x_2 terms in the other equations, and so on, until we finally obtain a very simple equivalent system of equations.

Example 5 Solve

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

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

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

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

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

The result of the calculation is written in place of the original third equation:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

Next, multiply equation 2 by ½ in order to obtain 1 as the coefficient for x_2

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

To eliminate the x_2 term from third equation add 3 times equation 2 to equation 3,

The new system has a triangular form

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

Now using 3rd equation eliminate the x_3 term from first and second equation i.e. multiply 3rd equation with 4 and add in second equation. Then subtract the third equation from first equation we get

$$\begin{aligned} x_1 - 2x_2 &= -3 \\ x_2 &= 16 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Adding 2 times equation 2 to equation 1, we obtain the result

$$\begin{cases} x_1 = 29 \\ x_2 = 16 \\ x_3 = 3 \end{cases} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

This completes the solution.
Our work indicates that the only solution of the original system is (29, 16, 3).

To verify that (29, 16, 3) is a solution, substitute these values into the left side of the original system for x_1, x_2 and x_3 and after computing, we get

$$(29) - 2(16) + (3) = 29 - 32 + 3 = 0$$
$$2(16) - 8(3) = 32 - 24 = 8$$
$$-4(29) + 5(16) + 9(3) = -116 + 80 + 27 = -9$$

The results agree with the right side of the original system, so (29, 16, 3) is a solution of the system.

This example illustrates how operations on equations in a linear system correspond to operations on the appropriate rows of the augmented matrix. The three basic operations listed earlier correspond to the following operations on the augmented matrix.

Elementary Row Operations

1. (Replacement) Replace one row by the sum of itself and a nonzero multiple of another row.
2. (Interchange) Interchange two rows.
3. (Scaling) Multiply all entries in a row by a nonzero constant.

Row equivalent matrices

A matrix B is said to be row equivalent to a matrix A of the same order if B can be obtained from A by performing a finite sequence of elementary row operations of A.
If A and B are row equivalent matrices, then we write this expression mathematically as $A \sim B$.

For example $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$ are row equivalent matrices because we add 4 times of 1^{st} row in 3^{rd} row in 1^{st} matrix.

Note If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

Row operations are extremely easy to perform, but they have to be learnt and practice.

Two Fundamental Questions

1. Is the system consistent; that is, does at least one solution exist?
2. If a solution exists is it the only one; that is, is the solution unique?

We try to answer these questions via row operations on the augmented matrix.

Example 6 Determine if the following system of linear equations is consistent

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

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.

<u>Example 7</u> 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.

$$\begin{bmatrix} 1\ {}^{Pivot} & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

Pivot Column

Since all entries in a column below a leading entry should be zero. For this add row 1 in row 2, and multiply row 1 by 2 and add in row 3.

Pivot

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \qquad \boxed{\begin{matrix} R_1 + R_2 \\ 2R_1 + R_3 \end{matrix}}$$

Next pivot column

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

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

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

$$\underbrace{\overbrace{\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}^{\text{Pivot}}}_{\text{Pivot column}} \qquad \textit{General} \quad \textit{form} \quad \begin{bmatrix} \circ & * & * & * & * \\ 0 & \circ & * & * & * \\ 0 & 0 & 0 & \circ & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This is in echelon form and thus columns 1, 2, and 4 of $\boldsymbol{A}$ are pivot columns.

$$\underbrace{\overbrace{\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}}^{\text{Pivot positions}}}_{\text{Pivot columns}}$$

<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.

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

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.)

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

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

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

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

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 2^{nd} Row by 2

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

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

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

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.)

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

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

$$\begin{aligned} x_1 + 6x_2 \quad + 3x_4 &= 0 \\ x_3 - 4x_4 &= 5 \\ x_5 &= 7 \end{aligned} \tag{6}$$

The pivot columns of the matrix are 1, 3 and 5, so the basic variables are x_1, x_3, and x_5. The remaining variables, x_2 and x_4, must be free.

Solving for the basic variables, we obtain the general solution:

$$\begin{cases} x_1 = -6x_2 - 3x_4 \\ x_2 \text{ is free} \\ x_3 = 5 + 4x_4 \\ x_4 \text{ is free} \\ x_5 = 7 \end{cases} \qquad (7)$$

Note that the value of x_5 is already fixed by the third equation in system (6).

Exercise

1. Find the general solution of the linear system whose augmented matrix is

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

2. Find the general solution of the system

$$\begin{aligned} x_1 - 2x_2 - x_3 + 3x_4 &= 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 &= 3 \\ 3x_1 - 6x_2 - 6x_3 + 8x_4 &= 2 \end{aligned}$$

Find the general solutions of the systems whose augmented matrices are given in Exercises 3-12

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

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

5. $\begin{bmatrix} 0 & 3 & 6 & 9 \\ -1 & 1 & -2 & -1 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 3 & -3 & 7 \\ 3 & 9 & -4 & 1 \end{bmatrix}$

7. $\begin{pmatrix} 1 & 2 & -7 \\ -1 & -1 & 1 \\ 2 & 1 & 5 \end{pmatrix}$

8. $\begin{pmatrix} 1 & 2 & 4 \\ -2 & -3 & -5 \\ 2 & 1 & -1 \end{pmatrix}$

9. $\begin{pmatrix} 2 & -4 & 3 \\ -6 & 12 & -9 \\ 4 & -8 & 6 \end{pmatrix}$

10. $\begin{pmatrix} 1 & 0 & -9 & 0 & 4 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

11. $\begin{pmatrix} 1 & -2 & 0 & 0 & 7 & -3 \\ 0 & 1 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 1 & 5 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

12. $\begin{pmatrix} 1 & 0 & -5 & 0 & -8 & 3 \\ 0 & 1 & 4 & -1 & 0 & 6 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Determine the value(s) of h such that the matrix is the augmented matrix of a consistent linear system.

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

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

Choose h and k such that the system has (a) no solution, (b) a unique solution, and (c) many solutions. Give separate answer for each part.

15. $x_1 + hx_2 = 1$
$2x_1 + 3x_2 = k$

16. $x_1 - 3x_2 = 1$
$2x_1 + hx_2 = k$

Lecture 5

Vector Equations

This lecture is devoted to connect equations involving vectors to ordinary systems of equations. The term vector appears in a variety of mathematical and physical contexts, which we will study later, while studying "Vector Spaces". Until then, we will use vector to mean a list of numbers. This simple idea enables us to get interesting and important applications as quickly as possible.

Column Vector

"*A matrix with only one column is called column vector or simply a vector*".

e.g. $u = \begin{bmatrix} 3 & -1 \end{bmatrix}^T = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, $\quad v = \begin{bmatrix} 2 & 3 & 5 \end{bmatrix}^T = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$, $\quad w = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 \end{bmatrix}^T$ are all column vectors or simply vectors.

Vectors in R^2

If $\mathbb{R}$ is the set of all real numbers then the set of all vectors with two entries is denoted by $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.

For example: the vector $u = \begin{bmatrix} 3 & -1 \end{bmatrix}^T = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \in \mathbb{R}^2$

Here real numbers are appeared as entries in the vectors, and the exponent **2** indicates that the vectors contain only two entries.

Similarly $\boldsymbol{R^3}$ and $\boldsymbol{R^4}$ contain all vectors with three and four entries respectively. The entries of the vectors are always taken from the set of real numbers ***R.*** The entries in vectors are assumed to be the elements of a set, called a ***Field***. It is denoted by F .

Algebra of Vectors

Equality of vectors in $\mathbb{R}^2$

Two vectors in $\boldsymbol{R^2}$ *are equal if and only if their corresponding entries are equal.*

If $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$ *then* $u = v$ *iff* $\boxed{u_1 = v_1} \wedge \boxed{u_2 = v_2}$

So $\begin{bmatrix} 4 \\ 6 \end{bmatrix} \neq \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ as $4 = 4$ but $6 \neq 3$

Note In fact, vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\boldsymbol{R^2}$ are nothing but ordered pairs (x, y) of real numbers both representing the position of a point with respect to origin.

Addition of Vectors

Given two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\boldsymbol{R^2}$, their sum is the *vector* $\boldsymbol{u} + \boldsymbol{v}$ *obtained by adding corresponding entries of the vectors* **u** *and* **v**, which is again a vector in $\mathbb{R}^2$

For $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$ Then $u + v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix} \in \mathbb{R}^2$

For example, $\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1+2 \\ -2+5 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

Scalar Multiplication of a vector

Given a vector $\boldsymbol{u}$ and a real number c, the scalar multiple of $\boldsymbol{u}$ by c is the *vector* $c\boldsymbol{u}$ *obtained by multiplying each entry in* **u** *by c.*

For example, if $u = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ *and* $c = 5,$ *then* $cu = 5\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ -5 \end{bmatrix}$

Notations The number c in $c\boldsymbol{u}$ is a **scalar**; it is written in lightface type to distinguish it from the boldface vector $\boldsymbol{u}$.

Example 1 Given $u = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ *and* $v = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, find 4$\boldsymbol{u}$, (-3) $\boldsymbol{v}$, and 4u + (-3) $\boldsymbol{v}$

Solution $4u == 4\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 4\times 1 \\ 4\times(-2) \end{bmatrix} = \begin{bmatrix} 4 \\ -8 \end{bmatrix}$, $(-3)v = (-3)\begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} -6 \\ 15 \end{bmatrix}$

And $4u + (-3)v = \begin{bmatrix} 4 \\ -8 \end{bmatrix} + \begin{bmatrix} -6 \\ 15 \end{bmatrix} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$

Note: Sometimes for our convenience, we write a column vector $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ in the form (3, –1). In this case, we use *parentheses and a comma to distinguish the vector* (3, –1) *from the* 1×2 *row matrix* [3 –1], written with brackets and no comma.

Thus $\begin{bmatrix} 3 \\ -1 \end{bmatrix} \neq [3 \quad -1]$ but $\begin{bmatrix} 3 \\ -1 \end{bmatrix} = (3, -1)$

Geometric Descriptions of R^2

Consider a rectangular coordinate system in the plane. Because each point in the plane is determined by an ordered pair of numbers, *we can identify a geometric point (a, b) with the column vector* $\begin{bmatrix} a \\ b \end{bmatrix}$. So we may regard $\boldsymbol{R^2}$ as the set of all points in the plane.

See Figure 1.

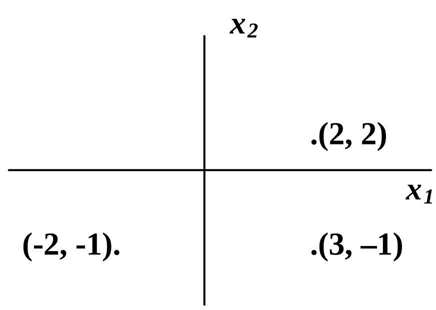


Figure 1 Vectors as points.

Vectors in R^3

Vectors in $\boldsymbol{R^3}$ are 3×1 column matrices with three entries. They are represented geometrically by points in a three-dimensional coordinate space, with arrows from the origin sometimes included for visual clarity.

Vectors in R^n

If **n** is a positive integer, $\boldsymbol{R^n}$ (read “r-n”) denotes the collection of all lists (or ordered n- tuples) of **n** real numbers, usually written as $n\times 1$ column matrices, such as

$$u = \begin{bmatrix} u_1 & u_2 & \cdots u_n \end{bmatrix}^T$$

The vector whose all entries are zero is called the **zero vector** and is denoted by **O.** (The number of entries in **O** will be clear from the context.)

Algebraic Properties of R^n

For all ***u, v, w*** in $\boldsymbol{R^n}$ and all scalars *c* and *d*:

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (Commutative)

(ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (Associative)

(iii) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$ (Additive Identity)

(iv) $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$ (Additive Inverse)

where $-\mathbf{u}$ denotes $(-1)\mathbf{u}$

(v) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ (Scalar Distribution over Vector Addition)

(vi) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ (Vector Distribution over Scalar Addition)

(vii) $c(d\mathbf{u}) = (cd)\mathbf{u}$

(viii) $1\mathbf{u}=\mathbf{u}$

Linear Combinations Given vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_p}$ in $\boldsymbol{R^n}$ and given scalars $c_1, c_2, \ldots, c_p$ the vector defined by

$$y = c_1 v_1 + c_2 v_2 + \cdots + c_p v_p$$

is called a **linear combination** of $\mathbf{v_1}, \ldots, \mathbf{v_p}$ using weights $\mathbf{c_1}, \ldots, \mathbf{c_p}$.

Property (ii) above permits us to omit parenthesis when forming such a linear combination. The weights in a linear combination can be any real numbers, including zero.

Example

For $v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, if $w = \frac{5}{2}v_1 - \frac{1}{2}v_2$ then we say that w is a linear combination of v_1 and v_2.

Example As (3, 5 , 2) = 3(1, 0 , 0) + 5(0, 1 , 0) + 2(0, 0 , 1)

(3, 5 , 2) = $3v_1 + 5v_2 + 2v_3$ where $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$ $v_3 = (0, 0, 1)$

So (3, 5 , 2) is a vector which is linear combination of v_1, v_2, v_3

Example 5 Let $a_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}, a_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}, and \quad b = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$.

Determine whether ***b*** can be generated (or written) as a linear combination of $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$. That is, determine whether weights $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ exist such that

$$x_1 \boldsymbol{a_1} + x_2 \boldsymbol{a_2} = \boldsymbol{b} \qquad (1)$$

If the vector equation (1) has a solution, find it.

Solution Use the definitions of scalar multiplication and vector addition to rewrite the vector equation

$$\underbrace{x_1\begin{bmatrix}1\\-2\\-5\end{bmatrix}}_{\boldsymbol{a_1}} + \underbrace{x_2\begin{bmatrix}2\\5\\6\end{bmatrix}}_{\boldsymbol{a_2}} = \underbrace{\begin{bmatrix}7\\4\\-3\end{bmatrix}}_{\boldsymbol{b}}$$

$$\Rightarrow \quad \begin{bmatrix}x_1\\-2x_1\\-5x_1\end{bmatrix} + \begin{bmatrix}2x_2\\5x_2\\6x_2\end{bmatrix} = \begin{bmatrix}7\\4\\-3\end{bmatrix}$$

$$\Rightarrow \quad \begin{bmatrix}x_1+2x_2\\-2x_1+5x_2\\-5x_1+6x_2\end{bmatrix} = \begin{bmatrix}7\\4\\-3\end{bmatrix} \qquad (2)$$

$$\Rightarrow \quad \begin{aligned}x_1+2x_2&=7\\-2x_1+5x_2&=4\\-5x_1+6x_2&=-3\end{aligned} \qquad (3)$$

We solve this system by row reducing the augmented matrix of the system as follows:

$$\begin{bmatrix}1&2&7\\-2&5&4\\-5&6&-3\end{bmatrix}$$

$$By\ R_2+2R_1\ ;R_3+5R_1$$

$$\sim\begin{bmatrix}1&2&7\\0&9&18\\0&16&32\end{bmatrix}$$

$$By\ \left(\frac{1}{9}\right)R_2;\left(\frac{1}{16}\right)R_3^{/}$$

$$\sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

$By\, R_3 - R_2; R_1 - 2R_2$

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

The solution of (3) is $x_1 = 3$ and $x_2 = 2$. Hence $\boldsymbol{b}$ is a linear combination of $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$**,** with weights $x_1 = 3$ and $x_2 = 2$.

Spanning Set

If $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ are in $\mathbf{R}^n$, then the set of all linear combinations of $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ is denoted by Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ and is called the **subset of $\mathbf{R}^n$ spanned** (or **generated**) by $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$. That is, Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ is the collection of all vectors that can be written in the form of $c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_p\boldsymbol{v_p}$, with $c_1, \ldots, c_p$ scalars.

If we want to check whether a vector **b** is in Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ then we will see whether the vector equation

$x_1\boldsymbol{v_1} + x_2\boldsymbol{v_2} + \ldots + x_p\boldsymbol{v_p} = \boldsymbol{b}$ has a solution, or

Equivalently, whether the linear system with augmented matrix $[\boldsymbol{v_1}, \ldots, \boldsymbol{v_p} \quad \boldsymbol{b}]$ has a solution.

Note

(1) The set Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ contains every scalar multiple of $\boldsymbol{v_1}$

because $c\boldsymbol{v_1} = c\boldsymbol{v_1} + 0\boldsymbol{v_2} + \ldots + 0\boldsymbol{v_p}$ i.e every $c\boldsymbol{v_i}$ can be written as a linear combination of $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$

(2) Zero vector $= 0 \in Span\{v_1, v_2, \cdots v_n\}$ as 0 can be written as the linear combination of $v_1, v_2, \cdots v_n$ that is $0_v = 0_F v_1 + 0_F v_2 + \cdots + 0_F v_n$ here for the convenience it is mentioned that 0_v is the vector(zero vector) while 0_F is zero scalar (weight of all $v_1, v_2, \cdots v_n$) and in particular not to make confusion that 0_v and 0_F are same!

A Geometric Description of Span {*v*} and Span {*u*, *v*}

Let ***v*** be a nonzero vector in R^3. Then Span **{*v*}** is the set of all linear combinations of ***v*** or in particular set of scalar multiples of ***v***, and we visualize it as the set of points on the line in R^3 through ***v*** and ***0***.

If ***u*** and ***v*** are nonzero vectors in R^3, *with* ***v*** *not a multiple of* ***u***, then Span **{*u*, *v*}** is the plane in R^3 that contains ***u***, ***v*** and **0**. In particular, Span **{*u*, *v*}** contains the line in R^3 through ***u*** and **0** and the line through ***v*** and **0**.

Example 6 Let $a_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, a_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}, and\ b = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$.

Then Span $\{\boldsymbol{a_1}, \boldsymbol{a_2}\}$ is a plane through the origin in R^3. Does ***b*** lie in that plane?

Solution First we see the equation $x_1\boldsymbol{a_1} + x_2\boldsymbol{a_2} = \boldsymbol{b}$ has a solution?

To answer this, row-reduce the augmented matrix $[\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \boldsymbol{b}]$:

$$\begin{bmatrix} 1 & 5 & -3 \\ -2 & -13 & 8 \\ 3 & -3 & 1 \end{bmatrix}$$

$By\ R_2 + 2R_1$

$$\sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 18 & 10 \end{bmatrix}$$

$By\ R_3 + 6R_2$

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

Last row $\Rightarrow 0x_2 = -2$ which can not be true for any value of $x_2 \in \mathbb{R}$

$\Rightarrow$ Given system has no solution

$\therefore b \notin Span\{a_1, a_2\}$ and

in geometrical meaning, vector b does not lie in the plane spanned by vectors a_1 *and* a_2

Linear Combinations in Applications

The final example shows how scalar multiples and linear combinations can arise when a quantity such as "cost" is broken down into several categories. The basic principle for the example concerns the cost of producing several units of an item when the cost per unit is known:

$$\begin{Bmatrix} \text{number} \\ \text{of units} \end{Bmatrix} \cdot \begin{Bmatrix} \cos t \\ \text{per unit} \end{Bmatrix} = \begin{Bmatrix} \text{total} \\ \cos t \end{Bmatrix}$$

Example 7 A Company manufactures two products. For one dollar's worth of product B, the company spends \$0.45 on materials, \$0.25 on labor, and \$0.15 on overhead. For one dollar's worth of product C, the company spends \$0.40 on materials, \$0.30 on labor and \$0.15 on overhead.

Let $b = \begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix}$ *and* $c = \begin{bmatrix} .40 \\ .30 \\ .15 \end{bmatrix}$, then ***b*** and ***c*** represent the "costs per dollar of income" for the two products.

a) What economic interpretation can be given to the vector 100***b***?
b) Suppose the company wishes to manufacture x_1 dollars worth of product B and x_2 dollars worth of product C. Give a vector that describes the various costs the company will have (for materials, labor and overhead).

Solution

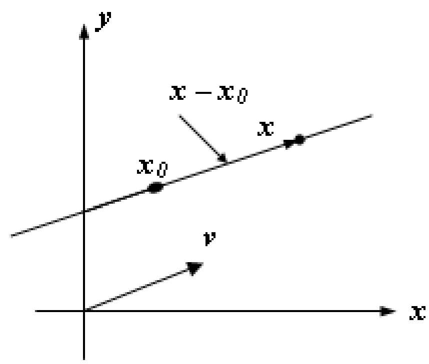
(a) We have $100b = 100\begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} = \begin{bmatrix} 45 \\ 25 \\ 15 \end{bmatrix}$

The vector 100***b*** represents a list of the various costs for producing \$100 worth of product B, namely, \$45 for materials, \$25 for labor, and \$15 for overhead.

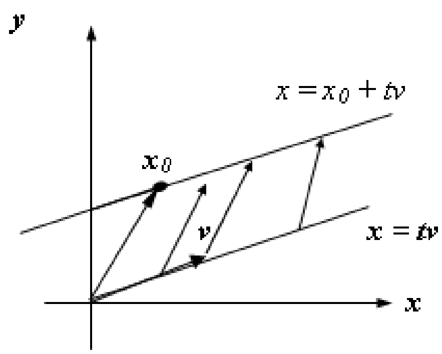
(b) The costs of manufacturing x_1 dollars worth of B are given by the vector $x_1\boldsymbol{b}$ and the costs of manufacturing x_2 dollars worth of C are given by $x_2\boldsymbol{c}$. Hence the total costs for both products are given by the vector $x_1\boldsymbol{b} + x_2\boldsymbol{c}$.

Vector Equation of a Line

Let $\mathbf{x_0}$ be a fixed point on the line and $\mathbf{v}$ be a nonzero vector that is parallel to the required line. Thus, if $\mathbf{x}$ is a variable point on the line through $\mathbf{x_0}$ that is parallel to $\mathbf{v}$, then the vector $\mathbf{x} - \mathbf{x_0}$ is a vector parallel to $\mathbf{v}$ as shown in fig below,



(b)



So by definition of parallel vectors $\boldsymbol{x} - \boldsymbol{x_0} = t\boldsymbol{v}$ *for some scalar t.*

it is also called a ***parameter*** which varies from $-\infty$ to $+\infty$. The variable point x traces out the line, so the line can be represented by the equation

$$\boldsymbol{x} - \boldsymbol{x_0} = t\boldsymbol{v} \text{ --------------(1)} \qquad (-\infty < t < +\infty)$$

This is a ***vector equation of the line*** through $\boldsymbol{x_0}$ and parallel to $\boldsymbol{v}$.

In the special case, where $\boldsymbol{x_0} = \mathbf{0}$, the line passes through the origin, it simplifies to

$$\boldsymbol{x} = t\boldsymbol{v} \qquad (-\infty < t < +\infty)$$

Parametric Equations of a Line in R^2

Let $\boldsymbol{x} = (x, y) \in R^2$ be a general point of the line through $\boldsymbol{x_0} = (x_0, y_0) \in R^2$ which is parallel to

$\boldsymbol{v} = (a, b) \in R^2$, then eq. 1 takes the form

$$(x, y) - (x_0, y_0) = t(a, b) \qquad (-\infty < t < +\infty)$$

$$\Rightarrow (x - x_0 \,, y - y_0) = (ta \,, tb) \qquad (-\infty < t < +\infty)$$

$$\Rightarrow x = x_0 + at, \quad y = y_0 + bt \quad (-\infty < t < +\infty)$$

These are called ***parametric equations*** of the line $in\ R^2$.

Parametric Equations of a Line in R^3

Similarly, if we let $\boldsymbol{x} = (x, y, z) \in R^3$ be a general point on the line through

$\boldsymbol{x_0} = (x_0, y_0, z_0) \in R^3$ that is parallel to $\boldsymbol{v} = (a, b, c) \in R^3$, then again eq. 1 takes the form

$$(x, y, z) = (x_0, y_0, z_0) + t(a, b, c) \; (-\infty < t < +\infty)$$

$$\Rightarrow \quad x = x_0 + at, \quad y = y_0 + bt, \quad z = z_0 + ct \quad (-\infty < t < +\infty)$$

These are the ***parametric equations*** of the line $in\ R^3$

Example 8

(a) Find a vector equation and parametric equations of the line in $\boldsymbol{R^2}$ that passes through the origin and is parallel to the vector $\boldsymbol{v} = (-2, 3)$.

(b) Find a vector equation and parametric equations of the line in $\boldsymbol{R^3}$ that passes through the point $P_0(1, 2, -3)$ and is parallel to the vector $\boldsymbol{v} = (4, -5, 1)$.

(c) Use the vector equation obtained in part (b) to find two points on the line that are different from P_0.

Solution

(a) We know that a vector equation of the line passing through origin is $\boldsymbol{x = tv}$.

Let $\boldsymbol{x} = (x, y)$. Then this equation can be expressed in component form as

$$(x, y) = t\,(-2, 3)$$

This is the vector equation of the line.

Equating corresponding components on the two sides of this equation yields the parametric equations $\qquad x = -\,2t, \quad y = 3t$

(b) The vector equation of the line is $\boldsymbol{x} = \boldsymbol{x_0} + t\boldsymbol{v}$.

Let $\boldsymbol{x} = (x, y, z)$, Here $\boldsymbol{x_0} = (1, 2, -3)$ and $\boldsymbol{v} = (4, -5, 1)$, then above equation can be expressed in component form as

$$(x, y, z) = (1, 2, -3) + t\,(4, -5, 1)$$

Equating corresponding components on the two sides of this equation yields the parametric equations

$$x = 1 + 4t, \quad y = 2 - 5t, \quad z = -3 + t$$

(c) Specific points on a line can be found by substituting numerical values for the parameter t.

For example, if we take $t = 0$ in part (b), we obtain the point $(x, y, z) = (1, 2, -3)$, which is the given point P_0.

$t = 1$ yields the point $(5, -3, -2)$ and

$t = -1$ yields the point $(-3, 7, -4)$.

Vector Equation of a Plane

Let x_0 be a fixed point on the required plane W and $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ be two nonzero vectors that are parallel to W and are not scalar multiples of one another. If x is any variable point in the plane W. Suppose $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ have their initial points at x_0, we can create a parallelogram with adjacent side's $t_1\boldsymbol{v_1}$ and $t_2\boldsymbol{v_2}$ in which $\boldsymbol{x} - \boldsymbol{x_0}$ is the diagonal given by the sum

$$\boldsymbol{x} - \boldsymbol{x_0} = t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2}$$

or, equivalently, $\boldsymbol{x} = \boldsymbol{x_0} + t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2}$ **---------------------(1)**

where t_1 and t_2 *are parameters* vary independently from $-\infty$ to $+\infty$,

This is a ***vector equation of the plane*** through $\boldsymbol{x_0}$ and parallel to the vectors $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$. In the special case where $x_0 = 0$, then vector equation of the plane passes through the origin takes the form

$$\boldsymbol{x} = t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2} \qquad (-\infty < t_1 < +\infty, -\infty < t_2 < +\infty)$$

Parametric Equations of a Plane

Let $\boldsymbol{x} = (x, y, z)$ be a general or variable point in the plane passes through a fixed point $\boldsymbol{x_0} = (x_0, y_0, z_0)$ and parallel to the vectors $\boldsymbol{v_1} = (a_1, b_1, c_1)$ and $\boldsymbol{v_2} = (a_2, b_2, c_2)$, then the component form of eq. 1 will be

$$(x, y, z) = (x_0, y_0, z_0) + t_1(a_1, b_1, c_1) + t_2(a_2, b_2, c_2)$$

Equating corresponding components, we get

$$x = x_0 + a_1t_1 + a_2t_2$$
$$y = y_0 + b_1t_1 + b_2t_2 \qquad (-\infty < t_1 < +\infty, -\infty < t_2 < +\infty)$$
$$z = z_0 + c_1t_1 + c_2t_2$$

These are called the parametric equations for this plane.

Example 9 (Vector and Parametric Equations of Planes)

(a) Find vector and parametric equations of the plane that passes through the origin of $\boldsymbol{R}^3$ and is parallel to the vectors $\boldsymbol{v_1} = (1, -2, 3)$ and $\boldsymbol{v_2} = (4, 0, 5)$.

(b) Find three points in the plane obtained in part (a).

Solution

(a) As vector equation of the plane passing through origin is $x = t_1v_1 + t_2v_2$.

Let $\boldsymbol{x} = (x, y, z)$ then this equation can be expressed in component form as

$$(x, y, z) = t_1(1, -2, 3) + t_2\,(4, 0, 5)$$

This is the vector equation of the plane.

Equating corresponding components, we get

$$x = t_1 + 4t_2, \quad y = -2t_1, \quad z = 3t_1 + 5t_2$$

These are the parametric equations of the plane.

(b) Points in the plane can be obtained by assigning some real values to the parameters t_1 and t_2:

$t_1 = 0$ and $t_2 = 0$ produces the point (0, 0, 0)

$t_1 = -2$ and $t_2 = 1$ produces the point (2, 4, –1)

$t_1 = ½$ and $t_2 = ½$ produces the point (5/2, –1, 4)

Vector equation of Plane through Three Points

If x_0, x_1 and x_2 are three non collinear points in the required plane, then, obviously, the vectors $\boldsymbol{v_1} = \boldsymbol{x_1} - \boldsymbol{x_0}$ and $\boldsymbol{v_2} = \boldsymbol{x_2} - \boldsymbol{x_0}$ are parallel to the plane. So, a vector equation of the plane is

$$\boldsymbol{x} = \boldsymbol{x_0} + t_1(\boldsymbol{x_1} - \boldsymbol{x_0}) + t_2(\boldsymbol{x_2} - \boldsymbol{x_0})$$

Example Find vector and parametric equations of the plane that passes through the points. P(2, – 4, 5), Q (–1, 4, –3) and R(1, 10, –7).

Solution

Let $\boldsymbol{x} = (x, y, z)$, and if we take $\boldsymbol{x_0}, \boldsymbol{x_1}$ and $\boldsymbol{x_2}$ to be the points P, Q and R respectively, then $x_1 - x_0 = \overrightarrow{PQ} = (-3, 8, -8)$ and $x_2 - x_0 = \overrightarrow{PR} = (-1, 14, -12)$

So the component form will be

$$(x, y, z) = (2, -4, 5) + t_1(-3, 8, -8) + t_2(-1, 14, -12)$$

This is the required vector equation of the plane.

By equating corresponding components, we get

$$x = 2 - 3t_1 - t_2, \quad y = -4 + 8t_1 + 14t_2, \quad z = 5 - 8t_1 - 12t_2$$

These are the parametric equations of the required plane.

Question: How can you tell that the points P, Q and R are not collinear?

Finding a Vector Equation from Parametric Equations

Example 11 Find a vector equation of the plane whose parametric equations are

$$x = 4 + 5t_1 - t_2, \quad y = 2 - t_1 + 8t_2, \quad z = t_1 + t_2$$

Solution First we rewrite the three equations as the single vector equation

$$(x, y, z) = (4 + 5t_1 - t_2, 2 - t_1 + 8t_2, t_1 + t_2)$$

$$\Rightarrow (x, y, z) = (4, 2, 0) + (5t_1, -t_1, t_1) + (-t_2, 8t_2, t_2)$$

$$\Rightarrow (x, y, z) = (4, 2, 0) + t_1(5, -1, 1) + t_2(-1, 8, 1)$$

This is a vector equation of the plane that passes through the point (4, 2, 0) and is parallel to the vectors $\boldsymbol{v_1}$ = (5, –1, 1) and $\boldsymbol{v_2}$ = (–1, 8, 1).

Finding Parametric Equations from a General Equation

Example 12 Find parametric equations of the plane $x - y + 2z = 5$.

Solution First we solve the given equation for x in terms of y and z

$$x = 5 + y - 2z$$

Now make y and z into parameters, and then express x in terms of these parameters.

Let y = t_1 and $z = t_2$

Then the parametric equations of the given plane are

$$x = 5 + t_1 - 2t_2, \quad y = t_1, \quad z = t_2$$

Exercises

1. Prove that $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ for any $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\boldsymbol{R}^n$.

2. For what value(s) of $\boldsymbol{h}$, $\boldsymbol{y}$ belongs to Span **{v_1, v_2, v_3}?** Where

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix} \quad v_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, and \quad y = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

3. Determine whether $\boldsymbol{b}$ is a linear combination of $\boldsymbol{a_1}$, $\boldsymbol{a_2}$, and $\boldsymbol{a_3}$.

i). $a_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, a_2 = \begin{bmatrix} -2 \\ 3 \\ -2 \end{bmatrix}, a_3 = \begin{bmatrix} -6 \\ 7 \\ 5 \end{bmatrix}, b = \begin{bmatrix} 11 \\ -5 \\ 9 \end{bmatrix}$

ii). $a_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, a_2 = \begin{bmatrix} -4 \\ 3 \\ 8 \end{bmatrix}, a_3 = \begin{bmatrix} 2 \\ 5 \\ -4 \end{bmatrix}, b = \begin{bmatrix} 3 \\ -7 \\ -3 \end{bmatrix}$

4. Determine if ***b*** is a linear combination of the vectors formed from the columns of the matrix ***A***.

i). $A = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 5 & 0 \\ 2 & 5 & 8 \end{bmatrix}, b = \begin{bmatrix} -5 \\ 11 \\ -7 \end{bmatrix}$

ii). $A = \begin{bmatrix} 1 & 0 & 5 \\ -2 & 1 & -6 \\ 0 & 2 & 8 \end{bmatrix}, b = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$

In exercises 7-10, list seven vectors in Span **{$\boldsymbol{v_1}$, $\boldsymbol{v_2}$}.** For each vector, show that the weights on $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ used to generate the vector and list the three entries of the vector. Give also geometric description of the Span **{$\boldsymbol{v_1}$, $\boldsymbol{v_2}$}.**

7. $v_1 = \begin{pmatrix} 5 \\ -1 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix}$

8. $v_1 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$

9. $v_1 = \begin{pmatrix} 2 \\ 6 \\ -4 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -3 \\ -9 \\ 6 \end{pmatrix}$

10. $v_1 = \begin{pmatrix} -3.7 \\ -0.4 \\ 11.2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 5.8 \\ 2.1 \\ 5.3 \end{pmatrix}$

11. Let $a_1 = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}, a_2 = \begin{bmatrix} -5 \\ -8 \\ 2 \end{bmatrix}, b = \begin{bmatrix} 3 \\ -5 \\ h \end{bmatrix}$. For what value(s) of h is ***b*** in the plane spanned by $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$?

12. Let $v_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, v_2 = \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix}, and\ y = \begin{bmatrix} h \\ -3 \\ -5 \end{bmatrix}$. For what value(s) of h is ***y*** in the plane generated by $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$?

13. Let $u = \begin{bmatrix} 2 \\ -1 \end{bmatrix} and \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Show that $\begin{bmatrix} h \\ k \end{bmatrix}$ is in Span{u, v} for all h and k.

Lecture 6

Matrix Equations

A fundamental idea in linear algebra is to view a linear combination of vectors as the product of a matrix and a vector. The following definition will permit us to rephrase some of the earlier concepts in new ways.

Definition If $\boldsymbol{A}$ is an $m \times n$ matrix, with columns $\boldsymbol{a_1}, \boldsymbol{a_2}, \dots, \boldsymbol{a_n}$ and if $\boldsymbol{x}$ is in $\boldsymbol{R^n}$, then the product of $\boldsymbol{A}$ and $\boldsymbol{x}$ denoted by $\boldsymbol{Ax}$, is the linear combination of the columns of $\boldsymbol{A}$ using the corresponding entries in $\boldsymbol{x}$ as weights, that is,

$$Ax = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 a_1 + x_2 a_2 + \dots + x_n a_n$$

Note that $\boldsymbol{Ax}$ is defined only if the number of columns of $\boldsymbol{A}$ equals the number of entries in $\boldsymbol{x}$.

Example 1

a) $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$

b) $\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$

Example 2 For $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$ in $\boldsymbol{R^m}$, write the linear combination $3\boldsymbol{v_1} - 5\boldsymbol{v_2} + 7\boldsymbol{v_3}$ as a matrix times a vector.

Solution Place $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$ into the columns of a matrix $\boldsymbol{A}$ and place the weights 3, -5, and 7 into a vector $\boldsymbol{x}$.

That is, $3v_1 - 5v_2 + 7v_3 = [v_1 \quad v_2 \quad v_3] \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} = Ax$

We know how to write a system of linear equations as a vector equation involving a linear combination of vectors. For example, we know that the system

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 4 \\ -5x_2 + 3x_3 &= 1 \end{aligned} \qquad \text{is equivalent to} \qquad x_1\begin{bmatrix}1\\0\end{bmatrix} + x_2\begin{bmatrix}2\\-5\end{bmatrix} + x_3\begin{bmatrix}-1\\3\end{bmatrix} = \begin{bmatrix}4\\1\end{bmatrix}$$

Writing the linear combination on the left side as a matrix times a vector, we get

$$\begin{bmatrix}1 & 2 & -1\\0 & -5 & 3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}4\\1\end{bmatrix}$$

Which has the form ***Ax* = *b***, and we shall call such an equation a **matrix equation**, to distinguish it from a vector equation.

<u>Theorem</u> 1 If ***A*** is an $m \times n$ matrix, with columns $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_n}$ and if ***b*** is in $\boldsymbol{R^m}$, the matrix equation ***Ax* = *b*** has the same solution set as the vector equation

$$x_1\boldsymbol{a_1} + x_2\boldsymbol{a_2} + \ldots + x_n\boldsymbol{a_n} = \boldsymbol{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is $\begin{bmatrix} a_1 & a_2 & \ldots & a_n & b \end{bmatrix}$

<u>Existence of Solutions</u> The equation ***Ax* = *b*** has a solution if and only if ***b*** is a linear combination of the columns of ***A***.

<u>Example 3</u> Let $A = \begin{bmatrix}1 & 3 & 4\\-4 & 2 & -6\\-3 & -2 & -7\end{bmatrix}$ and $b = \begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$.

Is the equation ***Ax* = *b*** consistent for all possible b_1, b_2, b_3?

<u>Solution</u> Row reduce the augmented matrix for ***Ax* = *b***:

$$\begin{bmatrix}1 & 3 & 4 & b_1\\-4 & 2 & -6 & b_2\\-3 & -2 & -7 & b_3\end{bmatrix} \overset{4R_1+R_2,\,3R_1+R_3}{\sim} \begin{bmatrix}1 & 3 & 4 & b_1\\0 & 14 & 10 & b_2+4b_1\\0 & 7 & 5 & b_3+3b_1\end{bmatrix}$$

$$R_3 - \frac{1}{2}R_2$$

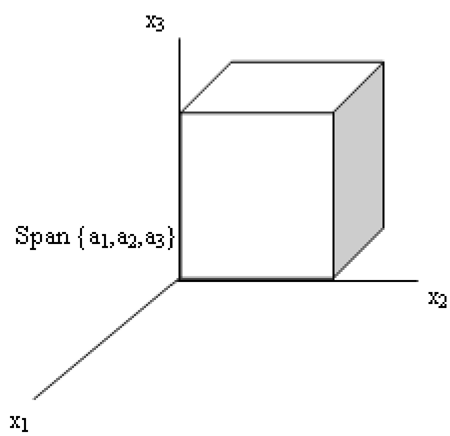
$$\sim \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{bmatrix}$$

The third entry in the augmented column is $b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1)$

The equation $\boldsymbol{Ax} = \boldsymbol{b}$ is not consistent for every $\boldsymbol{b}$ because some choices of $\boldsymbol{b}$ can make $b_1 - \frac{1}{2}b_2 + b_3$ nonzero.

The entries in $\boldsymbol{b}$ must satisfy $b_1 - \frac{1}{2}b_2 + b_3 = 0$

This is the equation of a plane through the origin in $\boldsymbol{R}^3$. The plane is the set of all linear combinations of the three columns of $\boldsymbol{A}$. See figure below.



The equation $\boldsymbol{Ax} = \boldsymbol{b}$ fails to be consistent for all $\boldsymbol{b}$ because the echelon form of $\boldsymbol{A}$ has a row of zeros. If $\boldsymbol{A}$ had a pivot in all three rows, we would not care about the calculations in the augmented column because in this case an echelon form of the augmented matrix could not have a row such as [0 0 0 1].

Example 4 Which of the following are linear combinations of

$$A = \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

(a) $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ **(c)** $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix}$

Solution

(a) $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$ = a A + b B + c C

$$= a\begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + b\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + c\begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4a+b & -b+2c \\ -2a+2b+c & -2a+3b+4c \end{bmatrix}$$

$\Rightarrow$ 4a + b = 6 (1)

-b + 2c = -8 (2)

-2a + 2b + c = -1 (3)

-2a + 3b + 4c = -8 (4)

Subtracting equation (4) from equation (3), we obtain

-b – 3c = 7 (5)

Subtracting equation (5) from equation (2):

5c = -15 $\Rightarrow$ c = -3

From (2), -b + 2(-3) = -8 $\Rightarrow$ b = 2

From (3), -2a + 2(2) – 3 = -1 $\Rightarrow$ a = 1

Now we check whether these values satisfy equation (1).

4(1) + 2 = 6

It means that $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$ is the linear combination of **A**, **B** and **C**.

Thus

$$\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix} = 1A + 2B - 3C$$

(b) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = a\,A + b\,B + c\,C$

$$= a\begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + b\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + c\begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4a+b & -b+2c \\ -2a+2b+c & -2a+3b+4c \end{bmatrix}$$

$\Rightarrow$ $4a + b = 0$ (1)

$-b + 2c = 0$ (2)

$-2a + 2b + c = 0$ (3)

$-2a + 3b + 4c = 0$ (4)

Subtracting equation (3) from equation (4) we get

$b + 3c = 0$ (5)

Adding equation (2) and equation (5), we get

$5c = 0 \quad \Rightarrow \quad c = 0$

Put $c = 0$ in equation (5), we get $b = 0$

Put $b = c = 0$ in equation (3), we get $a = 0$

$\Rightarrow$ $a = b = c = 0$

It means that $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is the linear combination of **A**, **B** and **C**.

Thus $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0A + 0B + 0C$

(c) $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix} = a\,A + b\,B + c\,C$

$$= a\begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + b\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + c\begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4a+b & -b+2c \\ -2a+2b+c & -2a+3b+4c \end{bmatrix}$$

$\Rightarrow$ 4a + b = 6 (1)

-b + 2c = 0 (2)

-2a + 2b + c = 3 (3)

-2a + 3b + 4c = 8 (4)

Subtracting (4) from (3), we obtain

-b – 3c = -5 (5)

Subtracting (5) from (2):

5c = 5 $\Rightarrow$ c = 1

From (2), -b + 2(1) = 0 $\Rightarrow$ b = 2

From (3), -2a + 2(2) + 1 = 3 $\Rightarrow$ a = 1

Now we check whether these values satisfy (1).

4(1) + 2 = 6

It means that $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix}$ is the linear combination of **A**, **B** and **C**.

Thus $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix}$ = 1**A** + 2**B** +1**C**

Theorem 2 Let $\boldsymbol{A}$ be an $m \times n$ matrix. Then the following statements are logically equivalent.

(a) For each $\boldsymbol{b}$ in $\boldsymbol{R^m}$, the equation $\boldsymbol{Ax = b}$ has a solution.

(b) The columns of $\boldsymbol{A}$ Span $\boldsymbol{R^m}$.

(c) A has a pivot position in every row.

This theorem is one of the most useful theorems. It is about a coefficient matrix, not an augmented matrix. If an augmented matrix [***A***   ***b***] *has a pivot position in every row, then the equation* ***Ax = b*** *may or may not be consistent.*

Example 4 Compute ***Ax***, where $A = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$ *and* $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Solution From the definition,

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 \\ -x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix}$$

<u>Note</u>

In above example the first entry in ***Ax*** is a sum of products (sometimes called a **dot product**), using the first row of ***A*** and the entries in ***x***.

That is $\begin{bmatrix} 2 & 3 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \end{bmatrix}$

<u>Examples</u>

In each part determine whether the given vector span R^3

(a) $v_1 = (2,\ 2,\ 2),\ v_2 = (0,\ 0,\ 3),$
$v_3 = (0,\ 1,\ 1)$

(b) $v_1 = (3,\ 1,\ 4),\ v_2 = (2, -3,\ 5),$
$v_3 = (5, -2,\ 9),\ v_4 = (1,\ 4,\ -1)$

(c) $v_1 = (1,\ 2, 6),\ v_2 = (3,\ 4,\ 1),$
$v_3 = (4, 3, 1),\ v_4 = (3,\ 3,\ 1)$

Solutions

(a) We have to determine whether arbitrary vectors $b=(b_1,\, b_2,\, b_3)$ in R^3 can be expressed as a linear combination $b=k_1v_1+k_2v_2+k_3v_3$ of the vectors v_1, v_2, v_3

Expressing this in terms of components given by

$$(b_1,\, b_2,\, b_3)=k_1(2,2,2)+k_2(0,0,3)+k_3(0,1,1)$$
$$(b_1,\, b_2,\, b_3)=(2k_1+0k_2+0k_3, 2k_1+0k_2+k_3, 2k_1+3k_2+k_3)$$
$$2k_1+0k_2+0k_3=b_1$$
$$2k_1+0k_2+k_3=b_2$$
$$2k_1+3k_2+k_3=b_3$$

$A=\begin{bmatrix} 2 & 0 & 0 \\ 2 & 0 & 1 \\ 2 & 3 & 1 \end{bmatrix}$ has a non zero determinant

Now

$\det(A)=-6\neq 0$

Therefore v_1, v_2, v_3 span R^3

(b) The set S$\{v_1, v_2, v_3, v_4\}$of vectors in R^3 spans V=R^3 if

$$c_1v_1+c_2v_2+c_3v_3+c_4v_4=d_1w_1+d_2w_2+d_3w_3 \quad ....(1)$$

with

$$w_1=(1,0,0)$$
$$w_2=(0,1,0)$$
$$w_3=(0,0,1)$$

With our vectors v_1, v_2, v_3, v_4 equation (1) becomes

$$c_1(3,1,4)+c_2(2,-3,5)+c_3(5,-2,9)+c_4(1,4,-1)=d_1(1,,0,0)+d_2(0,1,0)+d_3(0,0,1)$$

Rearranging the left hand side yields

$$3c_1+2c_2+5c_3+1c_4=1d_1+0d_2+0d_3$$

$$1c_1-3c_2-2c_3+4c_4=0d_1+1d_2+0d_3$$

$$4c_1+5c_2+9c_3-1c_4=0d_1+0d_2+1d_3$$

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

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

The reduce row echelon form

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & \frac{5}{17} & \frac{3}{17} \\ 0 & 1 & 1 & -1 & 0 & \frac{-4}{17} & \frac{1}{17} \\ 0 & 0 & 0 & 0 & 1 & \frac{-7}{17} & -\frac{11}{17} \end{bmatrix}$$ Corresponds to the system of equations

$$1c_1 + 1c_3 + 1c_4 = (\frac{5}{17})d_2 + (\frac{3}{17})d_3$$

$$1c_2 + 1c_3 + -1c_4 = (\frac{-4}{17})d_2 + (\frac{1}{17})d_3 \qquad \ldots\ldots\ldots\ldots\ldots(2)$$

$$0 = 1d_1 + (\frac{-7}{17})d_2 + (-\frac{11}{17})d_3$$

So this system is inconsistent. The set S does not span the space V.

Similarly Part C can be solved by the same way.

Exercise

1. Let $A = \begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix}, x = \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}, and \quad b = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}.$

   It can be shown that $\boldsymbol{Ax} = \boldsymbol{b}$. Use this fact to exhibit $\boldsymbol{b}$ as a specific linear combination of the columns of $\boldsymbol{A}$.

2. Let $A=\begin{bmatrix}2 & 5\\ 3 & 1\end{bmatrix}, u=\begin{bmatrix}4\\ -1\end{bmatrix}, and\ \ v=\begin{bmatrix}-3\\ 5\end{bmatrix}$. Verify ***A(u + v)* = *Au* +*Av***.

3. Solve the equation ***Ax* = *b***, with $A=\begin{bmatrix}2 & 4 & -6\\ 0 & 1 & 3\\ -3 & -5 & 7\end{bmatrix}, b=\begin{bmatrix}2\\ 5\\ -3\end{bmatrix}$.

4. Let $u=\begin{bmatrix}-5\\ -3\\ -6\end{bmatrix} and\ A=\begin{bmatrix}3 & 5\\ 1 & 1\\ -2 & -8\end{bmatrix}$. Is ***u*** belongs to the plane in R^3 spanned by the columns of A? Why or why not?

5. Let $u=\begin{bmatrix}8\\ 2\\ 3\end{bmatrix} and\ A=\begin{bmatrix}4 & 3 & 5\\ 0 & 1 & -1\\ 1 & 2 & 0\end{bmatrix}$. Is ***u*** in the subset of R^3 spanned by the columns of A? Why or why not?

6. Let $A=\begin{bmatrix}-3 & 1\\ 6 & -2\end{bmatrix} and\ b=\begin{bmatrix}b_1\\ b_2\end{bmatrix}$. Show that the equation ***Ax* = *b*** is not consistent for all possible ***b***, and describe the set of all ***b*** for which ***Ax* = *b*** is consistent.

7. How many rows of $A=\begin{bmatrix}1 & 3 & -2 & -2\\ 0 & 1 & -1 & 5\\ -1 & -2 & 1 & 7\\ 1 & 1 & 0 & -6\end{bmatrix}$ contain pivot positions?

In exercises 8 to 13 , explain how your calculations justify your answer, and mention an appropriate theorem.

8. Do the columns of the matrix $A=\begin{bmatrix}1 & 3 & -4\\ 3 & 2 & -6\\ -5 & -1 & 8\end{bmatrix}$ span R^3?

9. Do the columns of the matrix $A = \begin{bmatrix} 1 & 3 & -2 & -2 \\ 0 & 1 & -1 & 5 \\ -1 & -2 & 1 & 7 \\ 1 & 1 & 0 & -6 \end{bmatrix}$ span R^4?

10. Do the columns of the matrix $A = \begin{bmatrix} 0 & 0 & 2 \\ 0 & -5 & 1 \\ 4 & 6 & -3 \end{bmatrix}$ span R^3?

11. Do the columns of the matrix $A = \begin{bmatrix} 3 & 5 \\ 1 & 1 \\ -2 & -8 \end{bmatrix}$ span R^3?

12. Let $v_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$. Does $\{v_1, v_2, v_3\}$ span R^4?

13. Let $v_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, v_2 = \begin{bmatrix} -1 \\ 3 \\ 7 \end{bmatrix}, v_3 = \begin{bmatrix} 3 \\ -2 \\ -2 \end{bmatrix}$. Does $\{ v_1, v_2, v_3\}$ span R^3?

14. It can be shown that $\begin{bmatrix} 4 & 1 & 2 \\ -2 & 0 & 8 \\ 3 & 5 & -6 \end{bmatrix}\begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 18 \\ 5 \end{bmatrix}$. Use this fact(and no row operations) to find scalars c_1, c_2, c_3 such that $\begin{bmatrix} 4 \\ 18 \\ 5 \end{bmatrix} = c_1\begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} + c_3\begin{bmatrix} 2 \\ 8 \\ -6 \end{bmatrix}$.

15. Let $u=\begin{bmatrix}3\\8\\4\end{bmatrix}, v=\begin{bmatrix}1\\3\\1\end{bmatrix}, and\ w=\begin{bmatrix}1\\1\\3\end{bmatrix}$. It can be shown that 2u – 5v – w = 0. Use this fact(and no row operations) to solve the equation $\begin{bmatrix}3 & 1\\8 & 3\\4 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}1\\1\\3\end{bmatrix}$.

Determine if the columns of the matrix span R^4.

16. $\begin{bmatrix}7 & 2 & -5 & 8\\-5 & -3 & 4 & -9\\6 & 10 & -2 & 7\\-7 & 9 & 2 & 15\end{bmatrix}$

17. $\begin{bmatrix}12 & -7 & 11 & -9 & 5\\-9 & 4 & -8 & 7 & -3\\-6 & 11 & -7 & 3 & -9\\4 & -6 & 10 & -5 & 12\end{bmatrix}$

Lecture 7

Solution Sets of Linear Systems

Solution Set

A solution of a linear system is an assignment of values to the variables $x_1, x_2, \ldots, x_n$ such that each of the equations in the linear system is satisfied. The set of all possible solutions is called the Solution Set

Homogeneous Linear System

A system of linear equations is said to be **homogeneous** if it can be written in the form ***Ax = 0***, where ***A*** is an $m \times n$ matrix and ***0*** is the zero vector in $\boldsymbol{R^m}$.

Trivial Solution

A homogeneous system ***Ax = 0*** always has at least one solution, namely, ***x = 0*** (the zero vector in $\boldsymbol{R^n}$). This zero solution is usually called the trivial solution of the homogeneous system.

Nontrivial solution

A solution of a linear system other than trivial is called its nontrivial solution.
i.e the solution of a homogenous equation ***Ax = 0*** such that $\boldsymbol{x \neq 0}$ is called **nontrivial solution**, that is, a nonzero vector ***x*** that satisfies ***Ax = 0***.

Existence and Uniqueness Theorem

*The homogeneous equation **Ax = 0** has a nontrivial solution if and only if the equation has at least one free variable.*

Example 1 Find the solution set of the following system

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 0 \\ 3x_1 + 2x_2 - 4x_3 &= 0 \\ 6x_1 + x_2 - 8x_3 &= 0 \end{aligned}$$

Solution

$$\text{Let } A = \begin{bmatrix} 3 & 5 & -4 \\ 3 & 2 & -4 \\ 6 & 1 & -8 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The augmented matrix is

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ 3 & 2 & -4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix}$$

For solution set, row reduce to reduced echelon form

$$\sim\begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} \qquad -1R_1+R_2,-2R_1+R_3$$

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

$$\sim\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad 1/3R_1,\, 1/3R_2,\, 5/3R_2+R_1$$

$$\sim\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (-1)R_2$$

$$\begin{aligned} x_1 \quad -\frac{4}{3}x_3 &= 0 \\ x_2 \quad &= 0 \\ 0 &= 0 \end{aligned}$$

It is clear that x_3 is a free variable, so $Ax = 0$ has nontrivial solutions (one for each choice of x_3). From above equations we have,

$$x_1 = \frac{4}{3}x_3, \quad x_2 = 0, \text{ with } x_3 \text{ free.}$$

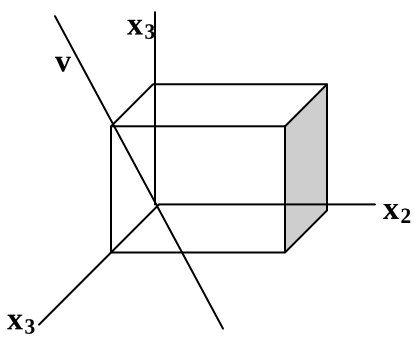
As a vector, the general solution of $Ax = 0$ is given by:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} = x_3 v, \quad where \quad v = \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

This shows that every solution of $\boldsymbol{Ax = 0}$ in this case is a scalar multiple of $\boldsymbol{v}$ (it means that v generate or spans the whole general solution).The trivial solution is obtained by choosing $\boldsymbol{x_3 = 0}$.

Geometric Interpretation

Geometrically, the solution set is a line through $\boldsymbol{0}$ in $\boldsymbol{R^3}$, as given in the **Figure below:**



Note: A nontrivial solution $\boldsymbol{x}$ can have some zero entries so long as not all of its entries are zero.

Example 2

Solve the following system

$$10x_1 - 3x_2 - 2x_3 = 0 \qquad (1)$$

Solution

Solving for the basic variable $\boldsymbol{x_1}$ in terms of the free variables,
dividing eq. 1 by 10 and solve for x

$\boldsymbol{x_1 = 0.3x_2 + 0.2x_3}$ *where* $\boldsymbol{x_2}$ and $\boldsymbol{x_3}$ free variables.

As a vector, the general solution is:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 + 0.2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.2x_3 \\ 0 \\ x_3 \end{bmatrix}$$

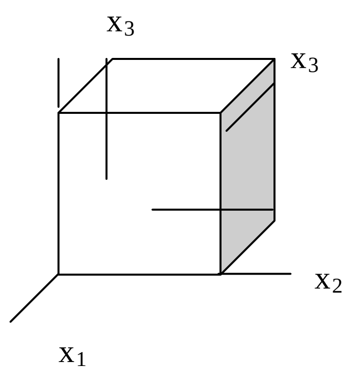
$$= x_2 \underbrace{\begin{bmatrix} 0.3 \\ 1 \\ 0 \end{bmatrix}}_{\boldsymbol{u}} + x_3 \underbrace{\begin{bmatrix} 0.2 \\ 0 \\ 1 \end{bmatrix}}_{\boldsymbol{v}} \qquad (2)$$

This calculation shows that every solution of (1) is a linear combination of the vector ***u, v*** shown in (2). That is, the solution set is Span {***u, v***}

.

Geometric Interpretation

Since neither ***u*** nor ***v*** is a scalar multiple of the other, so these are not parallel, the solution set is a plane through the origin, see the Figure below:



Note:

Above examples illustrate the fact that the solution set of a homogeneous equation $Ax = 0$ can be expressed explicitly as Span **{v_1, v_2, … , v_p}** for suitable vectors **v_1, v_2, ... , v_p**(because solution sets can be written in the form of linear combination of these vectors). If the only solution is the zero-vector then the solution set is **Span {0}**.

Example 3 (For Practice) Find the solution set of the following homogenous system:

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 0 \\ -4x_1 - 9x_2 + 2x_3 &= 0 \\ -3x_2 - 6x_3 &= 0 \end{aligned}$$

Solution:

Let $A = \begin{bmatrix} 1 & 3 & 1 \\ -4 & -9 & 2 \\ 0 & -3 & -6 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

The augmented matrix is:

$$\begin{bmatrix} 1 & 3 & 1 & 0 \\ -4 & -9 & 2 & 0 \\ 0 & -3 & -6 & 0 \end{bmatrix}$$

$$\sim\begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 3 & 6 & 0 \\ 0 & -3 & -6 & 0 \end{bmatrix} \qquad 4R_1 + R_2,$$

$$\sim\begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 3 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad R_2 + R_3$$

$$\sim\begin{bmatrix} 1 & 0 & -5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \frac{1}{2}R_2\ ,\ (-3)R_2 + R_1$$

SO

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

From above results, it is clear that $\boldsymbol{x_3}$ is a free variable, so $\boldsymbol{Ax = 0}$ has nontrivial solutions (one for each choice of $\boldsymbol{x_3}$).
From above equations we have,

$$x_1 = 5x_3, \quad x_2 = -2x_3, \text{ with } \boldsymbol{x_3} \text{ a free variable.}$$

As a vector, the general solution of $\boldsymbol{Ax = 0}$ is given by

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5x_3 \\ -2x_3 \\ x_3 \end{bmatrix} = x_3\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} = x_3 v, \qquad where \quad v = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$$

<u>**Parametric Vector Form of the solution**</u>

Whenever a solution set is described explicitly with vectors, we say that the solution is in <u>**parametric vector form**</u>

The equation

$$\boldsymbol{x} = s\boldsymbol{u} + t\boldsymbol{v} \qquad (s, t \text{ in } \boldsymbol{R})$$

is called a **parametric vector equation** of the plane. It is written in this form to emphasize that the parameters vary over all real numbers.

Similarly, the equation $\boldsymbol{x = x_3 v}$ (with $\boldsymbol{x_3}$ free), or $\boldsymbol{x = tv}$ (with $\boldsymbol{t}$ in $\boldsymbol{R}$), is a parametric vector equation of a line.

Solutions of Non-homogeneous Systems

When a non-homogeneous linear system has many solutions, the general solution can be written in parametric vector form as one vector plus an arbitrary linear combination of vectors that satisfy the corresponding homogeneous system.
To clear this concept consider the following examples,

Example: 5 Describe all solutions of ***Ax = b***, where

$$A=\begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix} \quad and \quad b=\begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}$$

Solution
Row operations on **[A b]** produce

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix}$$

$$\sim\begin{bmatrix} 3 & 5 & -4 & 7 \\ 0 & 3 & 0 & 6 \\ 0 & -9 & 0 & -18 \end{bmatrix} \qquad R_1+R_2, -2R_1+R_3$$

$$\sim\begin{bmatrix} 3 & 5 & -4 & 7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad 3R_2+R_3,\ \frac{1}{3}R_2$$

$$\sim\begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad -5R_2+R_1,\ \frac{1}{3}R_1$$

$$\begin{array}{lllll} & x_1 & & -\frac{4}{3}x_3 & =-1 \\ OR & & x_2 & & =2 \\ & & & 0 & =0 \end{array}$$

Thus $x_1=-1+\frac{4}{3}x_3$, $x_2=2$, and $\boldsymbol{x_3}$ is free.
As a vector, the general solution of ***Ax = b*** has the form

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1+\frac{4}{3}x_3 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}}_{\boldsymbol{p}} + x_3 \underbrace{\begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}}_{\boldsymbol{v}}$$

The equation $\boldsymbol{x} = \boldsymbol{p} + x_3\boldsymbol{v}$, or, writing t as a general parameter,

$$\boldsymbol{x} = \boldsymbol{p} + t\boldsymbol{v} \quad (t \text{ in } \boldsymbol{R}) \tag{3}$$

Note

We know that the solution set of this question when $A\boldsymbol{x} = \boldsymbol{0}$ (example 1) has the parametric vector equation

$$\boldsymbol{x} = t\boldsymbol{v} \ (t \text{ in } \boldsymbol{R}) \tag{4}$$

With the same $\boldsymbol{v}$ that appears in equation (3) in above example.
Thus the solutions of $A\boldsymbol{x} = \boldsymbol{b}$ are obtained by adding the vector $\boldsymbol{p}$ to the solutions of $A\boldsymbol{x} = \boldsymbol{0}$. The vector $\boldsymbol{p}$ itself is just one particular solution of $A\boldsymbol{x} = \boldsymbol{b}$ (corresponding to $t = 0$ in (3)).

The following theorem gives the precise statement.

Theorem

Suppose the equation $Ax = b$ is consistent for some given b, and let p be a solution. Then the solution set of $Ax = b$ is the set of all vectors of the form $w = p + v_h$, where v_h is any solution of the homogeneous equation Ax = 0.

Example 6: (For practice)

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 1 \\ -4x_1 - 9x_2 + 2x_3 &= -1 \\ -3x_2 - 6x_3 &= -3 \end{aligned}$$

Solution

Let $A = \begin{bmatrix} 1 & 3 & 1 \\ -4 & -9 & 2 \\ 0 & -3 & -6 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$

The augmented matrix is

$$\begin{bmatrix} 1 & 3 & 1 & 1 \\ -4 & -9 & 2 & -1 \\ 0 & -3 & -6 & -3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 3 & 6 & 3 \\ 0 & -3 & -6 & -3 \end{bmatrix} \qquad 4R_1 + R_2,$$

$$\sim \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 3 & 6 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad R_2 + R_3$$

$$\sim \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \frac{1}{3}R_2$$

$$\sim \begin{bmatrix} 1 & 0 & -5 & -2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (-3)R_2 + R_1$$

SO

$$\begin{aligned} x_1 \qquad - 5x_3 &= -2 \\ x_2 + 2x_3 &= \ \ 1 \\ 0 &= 0 \end{aligned}$$

Thus $x_1 = -2 + 5x_3, \quad x_2 = 1 - 2x_3$, and $\boldsymbol{x_3}$ is free.

As a vector, the general solution of ***Ax = b*** has the form

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 + 5x_3 \\ 1 - 2x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 5x_3 \\ -2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}}_{\boldsymbol{p}} + x_3 \underbrace{\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}}_{\boldsymbol{v}}$$

So we can write solution set in parametric vector form as

$$x = p + x_3 v$$

Steps of Writing a Solution Set (of a Consistent System) in a Parametric Vector Form

Step 1:
Row reduces the augmented matrix to reduced echelon form.

Step 2:
Express each basic variable in terms of any free variables appearing in an equation.

Step 3:
Write a typical solution ***x*** as a vector whose entries depend on the free variables if any.

Step 4:
Decompose ***x*** into a linear combination of vectors (with numeric entries) using the free variables as parameters.

Exercise

Determine if the system has a nontrivial solution. Try to use as few row operations as possible.

1. $x_1 - 5x_2 + 9x_3 = 0$
   $-x_1 + 4x_2 - 3x_3 = 0$
   $2x_1 - 8x_2 + 9x_3 = 0$

2. $3x_1 + 6x_2 - 4x_3 - x_4 = 0$
   $-5x_1 + 8x_3 + 3x_4 = 0$
   $8x_1 - x_2 + 7x_4 = 0$

3. $5x_1 - x_2 + 3x_3 = 0$
   $4x_1 - 3x_2 + 7x_3 = 0$

Write the solution set of the given homogeneous system in parametric vector form.

4. $x_1 - 3x_2 - 2x_3 = 0$
   $x_2 - x_3 = 0$
   $-2x_1 + 3x_2 + 7x_3 = 0$

5. $x_1 + 2x_2 - 7x_3 = 0$
   $-2x_1 - 3x_2 + 9x_3 = 0$
   $-2x_2 + 10x_3 = 0$

In exercises 6-8, describe all solutions of $\boldsymbol{Ax = 0}$ in parametric vector form where A is row equivalent to the matrix shown.

6. $\begin{bmatrix} 1 & -5 & 0 & 2 & 0 & -4 \\ 0 & 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

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

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

9. Describe the solution set in R^3 of $x_1 - 4x_2 + 3x_3 = 0$, compare it with the solution set of $x_1 - 4x_2 + 3x_3 = 7$.

10. Find the parametric equation of the line through ***a*** parallel to ***b***.

$$a = \begin{bmatrix} 3 \\ -8 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$$

11. Find a parametric equation of the line *M* through ***p*** and ***q***.

$$p = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, q = \begin{bmatrix} 0 \\ 7 \end{bmatrix}$$

12. Given $A = \begin{bmatrix} 5 & 10 \\ -8 & -16 \\ 7 & 14 \end{bmatrix}$, find one nontrivial solution of $\boldsymbol{Ax = 0}$ by inspection.

13. Given $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 9 \end{bmatrix}$, find one nontrivial solution of $\boldsymbol{Ax = 0}$ by inspection.

Lecture 8

Linear Independence

Definition

An indexed set of vectors **{v_1, v_2, ... , v_p}** in R^n is said to be **linearly independent** if the vector equation $x_1v_1 + x_2v_2 + \cdots + x_pv_p = 0$ has only the trivial solution. The set **{v_1, v_2, ... , v_p}** is said to be **linearly dependent** if there exist weights $c_1, \ldots, c_p$, not all zero, such that $c_1v_1 + c_2v_2 + \cdots + c_pv_p = 0$ (1)

Equation (1) is called a **linear dependence relation** among $v_1, \ldots, v_p$, when the weights are not all zero.

Example 1

Let $v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},\ v_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix},\ v_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

(a) Determine whether the set of vectors **{v_1, v_2, v_3}** is linearly independent or not.
(b) If possible, find a linear dependence relation among v_1, v_2, v_3.

Solution

(a) Row operations on the associated augmented matrix show that

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & -6 & -6 & 0 \end{bmatrix} \qquad (-2)R_1 + R_2,\ (-3)R_1 + R_3$$

$$\sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad R_2 + R_3 \qquad (2)$$

Clearly, x_1 and x_2 are basic variables and x_3 is free. Each nonzero value of x_3 determines a nontrivial solution.
Hence v_1, v_2, v_3 are linearly dependent (and not linearly independent).

(b) To find a linear dependence relation among v_1, v_2, v_3, completely row reduce the augmented matrix and write the new system:

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

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

$$\Rightarrow \begin{array}{l} x_1 \qquad -2x_3 = 0 \\ \quad x_2 \qquad +x_3 = 0 \\ \qquad\qquad 0 \quad = 0 \end{array}$$

Thus ***$x_1 = 2x_3$***, ***$x_2 = -x_3$***, and ***x_3*** is free.

Choose any nonzero value for ***x_3***, say, ***$x_3 = 5$***, then ***$x_1 = 10$***, and ***$x_2 = -5$***.
Substitute these values into $x_1v_1 + x_2v_2 + x_3v_3 = 0$

$$\Rightarrow \ \boldsymbol{10v_1 - 5v_2 + 5v_3 = 0}$$

This is one (out of infinitely many) possible linear dependence relation among ***v_1, v_2, v_3***.

<u>Example (for practice)</u>

Check whether the vectors are linearly dependent or linearly independent.

$v_1 = (3, -1) \qquad v_2 = (-2, 2)$

<u>Solution</u>

Consider two constants C_1 and C_2. Suppose:

$$c_1(3, -1) + c_2(-2, 2) = 0$$

$$(3c_1 - 2c_2, -c_1 + 2c_2) = (0, 0)$$

Now, set each of the components equal to zero to arrive at the following system of equations:

$$\begin{aligned} 3c_1 - 2c_2 &= 0 \\ -c_1 + 2c_2 &= 0 \end{aligned}$$

Solving this system gives the following solution,

$$c_1 = 0 \qquad c_2 = 0$$

The trivial solution is the only solution, so these two vectors are linearly independent.

Linear Independence of Matrix Columns

Suppose that we begin with a matrix $A = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix}$ instead of a set of vectors. The matrix equation ***Ax = 0*** can be written as $x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = 0$

Each linear dependence relation among the columns of ***A*** corresponds to a nontrivial solution of ***Ax = 0***.

Thus we have the following important fact.

The columns of a matrix A are linearly independent if and only if the equation $Ax = 0$ has only the trivial solution.

Example 2 Determine whether the columns of $A = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 2 & -1 \\ 5 & 8 & 0 \end{bmatrix}$ are linearly independent.

Solution To study ***Ax = 0,*** row reduce the augmented matrix:

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

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix} \qquad R_{12}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & -2 & 5 & 0 \end{bmatrix} \qquad (-5)R_1 + R_3$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 13 & 0 \end{bmatrix} \qquad (2)R_2 + R_3$$

At this point, it is clear that there are three basic variables and no free variables. So the equation ***Ax = 0*** has only the trivial solution, and the columns of ***A*** are linearly independent.

Sets of One or Two Vectors

A set containing only one vector (say, ***v***) is linearly independent if and only if ***v*** is not the zero vector. This is because the vector equation $\boldsymbol{x_1 v = 0}$ has only the trivial solution when $v \neq 0$. The zero vector is linearly dependent because $\boldsymbol{x_1 0 = 0}$ has many nontrivial solutions.

Example 3
Check the following sets for linearly independence and dependence.

a. $v_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$

b. $v_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$

Solution

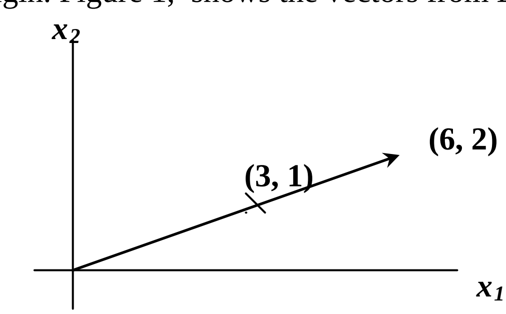
a) Notice that $\boldsymbol{v_2}$ is a multiple of $\boldsymbol{v_1}$, namely, $\boldsymbol{v_2} = \boldsymbol{2v_1}$.
Hence $\boldsymbol{-2v_1 + v_2 = 0}$, which shows that $\boldsymbol{\{v_1, v_2\}}$ is linearly dependent.

b) $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are certainly not multiples of one another. Could they be linearly dependent?

Suppose $\boldsymbol{c}$ and $\boldsymbol{d}$ satisfy $\boldsymbol{cv_1 + dv_2 = 0}$

If $c \neq 0$, then we can solve for $\boldsymbol{v_1}$ in terms of $\boldsymbol{v_2}$, namely, $\boldsymbol{v_1 = (-d/c)\, v_2}$. This result is impossible because $\boldsymbol{v_1}$ is not a multiple of $\boldsymbol{v_2}$. So, $\boldsymbol{c}$ must be zero. Similarly, $\boldsymbol{d}$ must also be zero.
Thus $\boldsymbol{\{v_1, v_2\}}$ is a linearly independent set.

Note A set of two vectors $\{v_1, v_2\}$ is linearly dependent if and only if one of the vectors is a multiple of the other.
In geometric terms, two vectors are linearly dependent if and only if they lie on the same line through the origin. Figure 1, shows the vectors from Example 3.



Linearly dependent

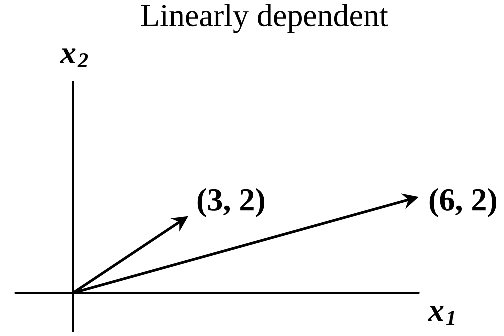


Figure 1 Linearly independent

Sets of Two or More Vectors

Theorem (Characterization of Linearly dependent Sets)

An indexed set $s = \{v_1, v_2, \cdots, v_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in ***S*** is a linear combination of the others. In fact, if **S** is linearly dependent, and $v \neq 0$, then some v_j (with $j > 1$) is a linear combination of the preceding vectors, $v_1, \cdots, v_{j-1}$.

Proof

If some v_j in ***S*** equals a linear combination of the other vectors, then v_j can be subtracted from both sides of the equation, producing a linear dependence relation with a nonzero weight (–1) on v_j.

For instance, if $\boldsymbol{v_1 = c_2v_2 + c_3v_3}$, then $\boldsymbol{0 = (-1)v_1 + c_2v_2 + c_3v_3 + 0v_4 + \ldots + 0v_p}$.
Thus ***S*** is linearly dependent.

Conversely, suppose ***S*** is linearly dependent. If $\boldsymbol{v_1}$ is zero, then it is a (trivial) linear combination of the other vectors in ***S***.

If $v \neq 0$ and there exist weights $\boldsymbol{c_1, \ldots, c_p}$, not all zero(because vectors are linearly dependent), such that

$$\boldsymbol{c_1v_1 + c_2v_2 + \ldots + c_pv_p = 0}$$

Let j be the largest subscript for which $c_j \neq 0$. If $j = 1$, then $\boldsymbol{c_1v_1 = 0}$, which is impossible because $v_j \neq 0$.

So $j > 1$, and $c_1v_1 + \cdots + c_jv_j + 0v_{j+1} + \cdots + 0v_p = 0$

$$c_jv_j = -c_1v_1 - c_2v_2 - \cdots - c_{j-1}v_{j-1}$$

$$v_j = \left(-\frac{c_1}{c_j}\right)v_1 + \left(-\frac{c_2}{c_j}\right)v_2 + \cdots + \left(\frac{c_{j-1}}{c_j}\right)v_{j-1}$$

Note: This theorem does not say that every vector in a linearly dependent set is a linear combination of the preceding vectors. A vector in a linearly dependent set may fail to be a linear combination of the other vectors.

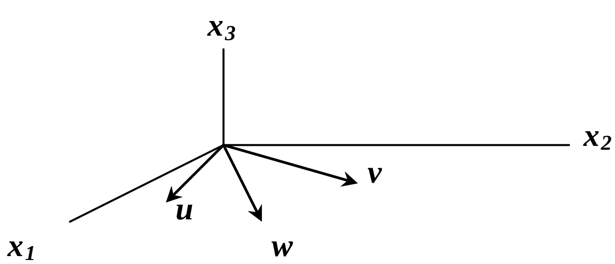
Example 4 Let $u = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Describe the set spanned by ***u*** and ***v***, and prove that a vector ***w*** is in Span **{*u, v*}** if and only if **{*u, v, w*}** is linearly dependent.

Solution

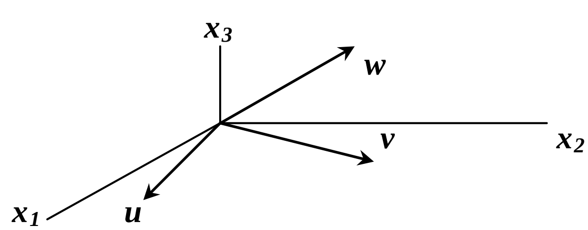
The vectors ***u*** and ***v*** are linearly independent because neither vector is a multiple of the other, nor so they span a plane in $\boldsymbol{R^3}$. In fact, Span **{*u, v*}** is the $\boldsymbol{x_1x_2}$**-plane** (with $\boldsymbol{x_3 = 0}$). If ***w*** is a linear combination of ***u*** and ***v***, then **{*u, v, w*}** is linearly dependent.

Conversely, suppose that **{*u, v, w*}** is linearly dependent.

Some vector in **{*u, v, w*}** is a linear combination of the preceding vectors (since $u \neq 0$). That vector must be ***w***, since ***v*** is not a multiple of ***u***. So ***w*** is in Span **{*u, v*}**



Linearly dependent ***w*** in Span **{*u, v*}**.



Linearly independent ***w*** not in **Span {*u, v*}**

Figure 2: Linear dependence in $\boldsymbol{R^3}$.

This example generalizes to any set **{*u, v, w*}** in $\boldsymbol{R^3}$ with ***u*** and ***v*** linearly independent. The set **{*u, v, w*}** will be linearly dependent if and only if ***w*** is in the plane spanned by ***u*** and ***v***.

Theorem

If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{v_1, v_2, \ldots, v_p\}$ in R^n is linearly dependent if $p > n$.

Example 5 The vectors $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ are linearly dependent, because there are three vectors in the set and there are only two entries in each vector.

Notice, however, that none of the vectors is a multiple of one of the other vectors. See Figure 4.

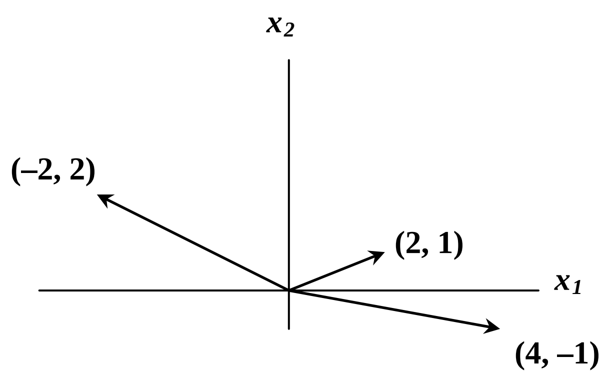


Figure 4 A linearly dependent set in $\boldsymbol{R}^2$.

Theorem

If a set $S = \{v_1, v_2, \ldots, v_p\}$ in R^n contains the zero vector, then the set is linearly dependent**.**

Proof

By renumbering the vectors, we may suppose that $\boldsymbol{v_1} = \boldsymbol{0}$.

Then $(1)\boldsymbol{v_1} + \boldsymbol{0v_2} + \ldots + \boldsymbol{0v_p} = \boldsymbol{0}$ shows that $\boldsymbol{S}$ is linearly dependent(because in this relation coefficient of v_1 is non zero).

Example 6 Determine by inspection if the given set is linearly dependent.

a. $\begin{bmatrix} 1 \\ 7 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 8 \end{bmatrix}$ **b.** $\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 8 \end{bmatrix}$ **c.** $\begin{bmatrix} -2 \\ 4 \\ 6 \\ 10 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \\ -9 \\ 15 \end{bmatrix}$

Solution

a) The set contains four vectors that each has only three entries. So,the set is linearly dependent by the Theorem above.
b) The same theorem does not apply here because the number of vectors does not exceed the number of entries in each vector. Since the zero vector is in the set, the set is linearly dependent by the next theorem.

c) As we compare corresponding entries of the two vectors, the second vector seems to be –3/2 times the first vector. This relation holds for the first three pairs of entries, but fails for the fourth pair. Thus neither of the vectors is a multiple of the other, and hence they are linearly independent.

Exercise

1. Let $u = \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix}, v = \begin{bmatrix} -6 \\ 1 \\ 7 \end{bmatrix}, w = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}, and \quad z = \begin{bmatrix} 3 \\ 7 \\ -5 \end{bmatrix}$.

(i) Are the sets **{*u, v*}, {*u, w*}, {*u, z*}, {*v, w*}, {*v, z*},** and **{*w, z*}** each linearly independent? Why or why not?
(ii) Does the answer to Problem (i) imply that **{*u, v, w, z*}** is linearly independent?
(iii) To determine if **{*u, v, w, z*}** is linearly dependent, is it wise to check if, say ***w*** is a linear combination of ***u, v*** and ***z***?
(iv) Is **{*u, v, w, z*}** linear dependent?

Decide if the vectors are linearly independent. Give a reason for each answer.

2. $\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ 0 \end{bmatrix}$

3. $\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ -5 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -6 \end{bmatrix}$

Determine if the columns of the given matrix form a linearly dependent set.

4. $\begin{bmatrix} 1 & 3 & -2 & 0 \\ 3 & 10 & -7 & 1 \\ -5 & -5 & 3 & 7 \end{bmatrix}$

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

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

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

For what values of h is v_3 in span $\{v_1, v_2\}$ and for what values of h is $\{v_1, v_2, v_3\}$ linearly dependent?

8. $v_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, v_2 = \begin{bmatrix} -2 \\ -6 \\ 4 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 2 \\ h \end{bmatrix}$

9. $v_1 = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}, v_2 = \begin{bmatrix} 3 \\ 9 \\ -1 \end{bmatrix}, v_3 = \begin{bmatrix} -2 \\ -6 \\ h \end{bmatrix}$

Find the value(s) of h for which the vectors are linearly dependent.

10. $\begin{bmatrix}1\\3\\-3\end{bmatrix},\begin{bmatrix}-2\\-4\\1\end{bmatrix},\begin{bmatrix}-1\\1\\h\end{bmatrix}$

11. $\begin{bmatrix}1\\-5\\-2\end{bmatrix},\begin{bmatrix}-3\\8\\6\end{bmatrix},\begin{bmatrix}4\\h\\-8\end{bmatrix}$

Determine by inspection whether the vectors are linearly independent. Give reasons for your answers.

12. $\begin{bmatrix}5\\5\end{bmatrix},\begin{bmatrix}6\\1\end{bmatrix},\begin{bmatrix}2\\4\end{bmatrix},\begin{bmatrix}3\\-6\end{bmatrix}$

13. $\begin{bmatrix}2\\-5\\1\end{bmatrix},\begin{bmatrix}-6\\5\\3\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix}$

14. $\begin{bmatrix}6\\2\\-8\end{bmatrix},\begin{bmatrix}3\\1\\-2\end{bmatrix}$

15. Given $A=\begin{bmatrix}2&3&5\\-5&1&-4\\-3&-1&-4\\1&0&1\end{bmatrix}$, observe that the third column is the sum of the first two columns. Find a nontrivial solution of $\boldsymbol{Ax} = \boldsymbol{0}$ without performing row operations.

Each statement in exercises 16-18 is either true(in all cases) or false(for at least one example). If false, construct a specific example to show that the statement is not always true. If true, give a justification.

16. If $v_1, \ldots, v_4$ are in R^4, and $v_3 = 2v_1 + v_2$, then $\{v_1, v_2, v_3, v_4\}$ is linearly dependent.

17. If v_1 and v_2 are in R^4, and v_1 is not a scalar multiple of v_2, then $\{v_1, v_2\}$ is linearly independent.

18. . If $v_1, \ldots, v_4$ are in R^4, and$\{v_1, v_2, v_3\}$ is linearly dependent, then $\{v_1, v_2, v_3, v_4\}$ is also linearly dependent.

19. Use as many columns of $A=\begin{bmatrix}8&-3&0&-7&2\\-9&4&5&11&-7\\6&-2&2&-4&4\\5&-1&7&0&10\end{bmatrix}$ as possible to construct a matrix B with the property that equation $\boldsymbol{Bx} = \boldsymbol{0}$ has only the trivial solution. Solve

Bx = 0 to verify your work.

Lecture 9

Linear Transformations

Outlines

- Matrix Equation
- Transformation, Examples, Matrix as Transformations
- Linear Transformation, Examples, Some Properties

Matrix Equation

An equation ***Ax = b*** is called a matrix equation in which a matrix ***A*** acts on a vector ***x*** by multiplication to produce a new vector called ***b.***

For instance, the equations

$$\underset{A}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}\underset{x}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}} = \underset{b}{\begin{bmatrix} 5 \\ 8 \end{bmatrix}}$$

and

$$\underset{A}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}\underset{u}{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}} = \underset{\underline{o}}{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

Solution of Matrix Equation

Solution of the ***Ax = b*** consists of those vectors ***x*** in ***the domain*** that are transformed into the vector ***b*** in ***range***.

Matrix equation ***Ax = b*** is an important example of transformation we would see later in the lecture.

Transformation or Function or Mapping

A transformation (or **function** or **mapping**) ***T*** from ***R*** ^{n} to ***R*** ^{m} is a rule that assigns to each vector ***x*** in ***R***^{n}, an image vector ***T(x)*** in ***R***^{m}.

$$T : R^n \to R^m$$

The set ***R*** ^{n} is called the **domain** of ***T***, and ***R*** ^{m} is called the **co-domain** of ***T***. For ***x*** in ***R***^{n} the set of all images ***T(x)*** is called the **range** of ***T***.

<u>Note</u>

To define a mapping or function, domain and co-domain are the ordinary sets. However to define a linear transformation, the domain and co-domain has to be $\mathbb{R}^m (or\ \mathbb{R}^n)$. Moreover a map $T:\mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation if for any two vectors say $u, v \in \mathbb{R}^m$ and the scalars c_1, c_2, the following equation is satisfied

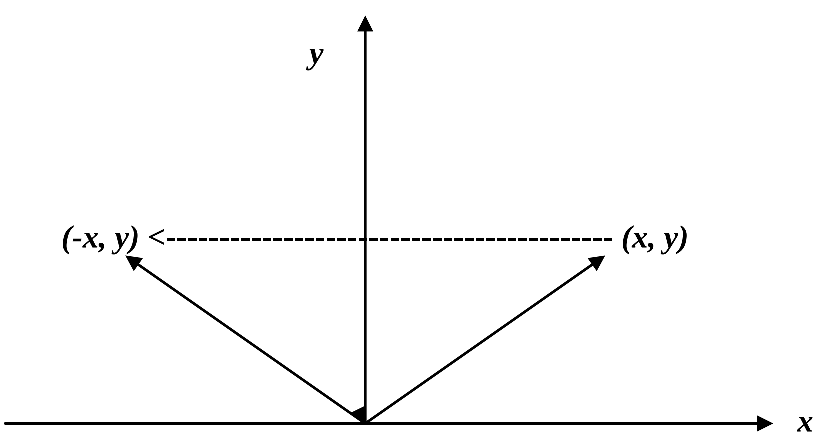
$$T\left(c_1u + c_2v\right) = c_1T(u) + c_2T(v)$$

<u>Example1</u> Consider a mapping $T : R^2 \to R^2$ defined by $T(x, y) = (-x, y)$. This transformation is a reflection about y-axis in xy plane.

Here $T(1,2) = (-1,2)$. T has transformed vector (1,2) into another vector (-1,2).

In matrix form:

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



Further the projection transformation $T : R^2 \to R^2$ defined by $T(x, y) = (x, 0)$ is given as:

$$Tv = Av = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

<u>Example 2</u> Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $u = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $c = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$, and define a transformation $T: R^2 \to R^3$ by ***T(x) = Ax***, so that

$$T(x) = Ax = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

a) Find ***T (u)***, the image of ***u*** under the transformation ***T***.
b) Find an ***x*** in ***R²***, whose image under ***T*** is ***b***.
c) Is there more than one ***x*** whose image under ***T*** is ***b***?
d) Determine if ***c*** is in the range of the transformation ***T***.

Solution (a)

$$T(u) = Au = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

Here $$T(u) = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

Here the matrix transformation has transformed $u = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ into another vector $\begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$

(b) We have to find an x such that **T (x) = b or Ax = b**

i. e $$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix} \qquad (1)$$

Now row reduced augmented matrix will be:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} -3R_1 + R_2,\ R_1 + R_3$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & -2 \end{bmatrix} \frac{1}{14}R_2,\ -4R_2 + R_3,$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} 3R_2 + R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence $x_1 = 1.5$, $x_2 = -0.5$, and $x = \begin{bmatrix} 1.5 \\ -.5 \end{bmatrix}$.

The image of this ***x*** under ***T*** is the given vector ***b***.

(c) From (2) it is clear that equation
(1) has a unique solution. So, there is exactly one ***x*** whose image is ***b***.

(d) The vector ***c*** is in the range of ***T*** if ***c*** is the image of some ***x*** in $\boldsymbol{R^2}$, that is, if $c = T(x)$ for some ***x***. This is just another way of asking if the system ***Ax = c*** is consistent. To find the answer, we will row reduce the augmented matrix:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} -3R_1 + R_2,\ R_1 + R_3$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & 8 \end{bmatrix} \frac{1}{4}R_3, R_{23}$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 14 & -7 \end{bmatrix} -14R_2 + R_3$$

$$\sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}$$

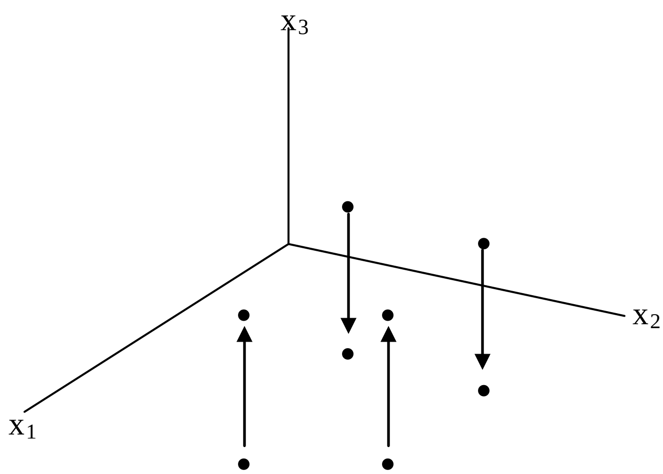
$$x_1 - 3x_2 = 3$$
$$0x_1 + x_2 = 2$$
$$0x_1 + 0x_2 = -35 \Rightarrow 0 = 35 \; but \; 0 \neq 35$$

Hence the system is inconsistent. So **c** is not in the range of **T**.

So from the above example we can view a transformation in the form of a matrix. We'll see that a transformation $T : R^n \rightarrow R^m$ can be transformed into a matrix of order $m \times n$ and every matrix of order $m \times n$ can be viewed as a linear transformation.

The next two matrix transformations can be viewed geometrically. They reinforce the dynamic view of a matrix as something that transforms vectors into other vectors.

Example 3 If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, then the transformation $x \rightarrow Ax$ projects points in $\boldsymbol{R^3}$ onto the $x_1x_2 - coordinate \;\; plane$ because $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$

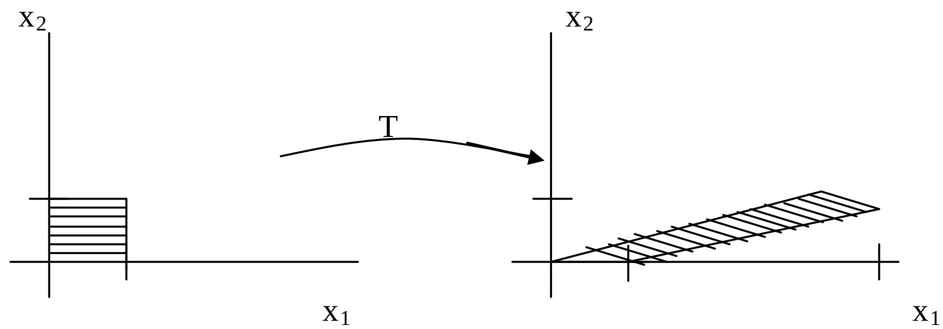


A projection transformation

Example 4 Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$., the transformation $T : R^2 \rightarrow R^2$ defined by ***T (x) = Ax*** is called a **shear** transformation.

The image of the point $u = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ *is* $T(u) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$,
and the image of $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ *is* $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$.

Here, ***T*** deforms the square as if the top of the square was pushed to the right while the base is held fixed. Shear transformations appear in physics, geology and crystallography.



A shear transformation

Linear Transformations

We know that if ***A*** is $m \times n$ matrix, then the transformation $x \to Ax$ has the properties $A(u+v) = Au + Av$ *and* $A(cu) = cAu$ for all ***u, v*** in $\boldsymbol{R^n}$ and all scalars ***c***.
These properties for a transformation identify the most important class of transformations in linear algebra.

Definition A transformation (or mapping) ***T*** is linear if:

1. ***T(u + v) = T(u) + T(v)*** for all ***u, v*** in the domain of ***T***;
2. ***T(cu) = cT(u)*** for all ***u*** and all scalars ***c***.

Example 5 Every matrix transformation is a linear transformation.

Example 6 Let $L: R^3 \to R^2$ be defined by $L(x, y, z) = (x, y)$.

we let $\boldsymbol{u} = (x_1, y_1, z_1)$ and $\boldsymbol{v} = (x_2, y_2, z_2)$.

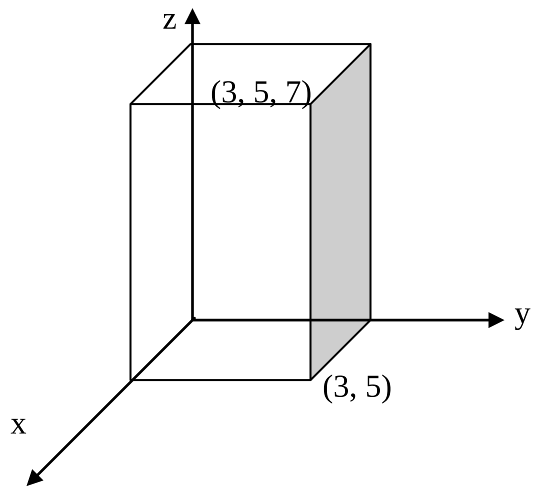
$$L(\boldsymbol{u}+\boldsymbol{v}) = L((x_1, y_1, z_1) + (x_2, y_2, z_2))$$
$$= L(x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

$$= (x_1 + x_2, y_1 + y_2)$$
$$= (x_1, y_1) + (x_2, y_2) = L(\boldsymbol{u}) + L(\boldsymbol{v})$$

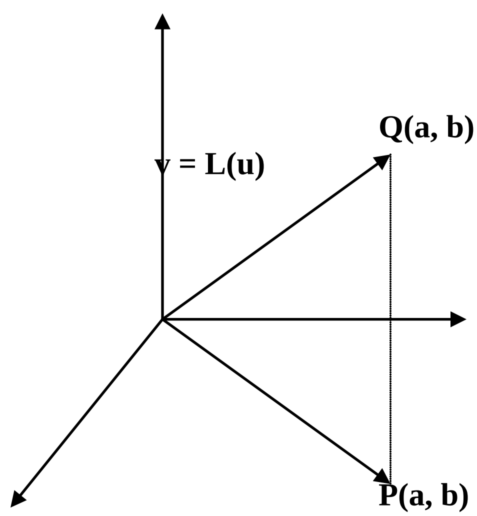
Also, if k is a real number, then

$$L(k\boldsymbol{u}) = L(kx_1, ky_1, kz_1) = (kx_1, ky_1) = kL(\boldsymbol{u})$$

Hence, ***L*** is a linear transformation, which is called a projection. The image of the vector (or point) (3, 5, 7) is the vector (or point) (3, 5) in xy-plane. See figure below:



Geometrically the image under ***L*** of a vector (a, b, c) in $\boldsymbol{R}^3$ is (a, b) in $\boldsymbol{R}^2$ can be found by drawing a line through the end point P(a, b, c) of ***u*** and perpendicular to $\boldsymbol{R}^2$, the *xy*-plane. The intersection Q(a, b)of this line with the *xy*-plane will give the image under ***L***. See the figure below:



Example 7 Let $L : R \to R$ be defined by $L(x) = x^2$

Let x and y in R and

$$L(x+y) = (x+y)^2 = x^2 + y^2 + 2xy \neq x^2 + y^2 = L(x) + L(y)$$
$$\Rightarrow L(x+y) \neq L(x) + L(y)$$

So we conclude that the function ***L*** is not a linear transformation.

Linear transformations preserve the operations of vector addition and scalar multiplication.

Properties

If ***T*** is a linear transformation, then

1. $T(0) = 0$
2. $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$
3. $T(c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \ldots + c_pT(\mathbf{v}_p)$

for all vectors ***u, v*** in the domain of ***T*** and all scalars *c*, *d*.

Proof

1. By the definition of Linear Transformation we have $T(cu) = cT(u)$ for all ***u*** and all scalars ***c***. Put $c = 0$ we'll get $T(0u) = 0T(u)$ This implies $T(0) = 0$

2. Just apply the definition of linear transformation. i. e

$$T(cu + dv) = T(cu) + T(dv) = cT(u) + dT(v)$$

Property (3) follows from (ii), because $T(0) = T(0u) = 0T(u) = 0$.
Property (4) requires both (i) and (ii):

OBSERVATION Observe that if a transformation satisfies property 2 for all ***u, v*** and *c*, *d*, it must be linear (Take *c* = *d* = *1* for preservation of addition, and take *d* = *0*)

3. Generalizing Property 2 we'll get 3

$$T(c_1\mathbf{v_1} + \ldots + c_p\mathbf{v_p}) = c_1T(\mathbf{v_1}) + \ldots + c_pT(\mathbf{v_p})$$

Applications in Engineering

In engineering and physics, property 3 is referred to as a superposition principle.
Think of $\mathbf{v_1}, \ldots, \mathbf{v_p}$ as signals that go into a system or process and $T(\mathbf{v_1}), \ldots, T(\mathbf{v_p})$ as the responses of that system to the signals. The system satisfies the superposition principle if an input is expressed as a linear combination of such signals, the system's response is the same linear combination of the responses to the individual signals.

Example 8 Given a scalar *r*, define $T: R \to R$ by

$$T(x) = x+1.$$

T is not a linear transformation (why!) because $T(0) \neq 0$ (by property 3)

Example 9 Given a scalar *r*, define $T: R^2 \to R^2$ by $T(\mathbf{x}) = r\mathbf{x}$.
T is called a **contraction** when $0 \le r < 1$
and a **dilation** when $r \ge 1$.

Let *r* = 3 and show that ***T*** is a linear transformation.

Solution Let ***u, v*** be in $\boldsymbol{R^2}$ and let *c*, *d* be scalars, then

$$\begin{aligned} T(c\mathbf{u} + d\mathbf{v}) &= 3(c\mathbf{u} + d\mathbf{v}) && \text{Definition of } \mathbf{T} \\ &= 3c\mathbf{u} + 3d\mathbf{v} \\ &= c(3\mathbf{u}) + d(3\mathbf{v}) && \text{Vector arithmetic} \\ &= c\mathbf{T}(\mathbf{u}) + d\mathbf{T}(\mathbf{v}) \end{aligned}$$

Thus ***T*** is a linear transformation because it satisfies (4).

Example 10 Define a linear transformation $T: R^2 \to R^2$ by

$$T(x) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

Find the images under ***T*** of $u = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad v = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad and \quad u+v = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$

Solution $T(u) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$, $T(v) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$

$$T(u+v) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$

In the above example, ***T*** rotates ***u, v*** and ***u + v*** counterclockwise through 90^0.
In fact, ***T*** transforms the entire parallelogram determined by ***u*** and ***v*** into the one determined by ***T (u)*** and ***T (v)***

Example 11 Let L: $R^3 \to R^2$ be a linear transformation for which:
L (1, 0, 0) = (2, -1),
L (0, 1, 0) = (3, 1), and
L (0, 0, 1) = (-1, 2).
Then find L (-3, 4, 2).

Solution Since (-3, 4, 2) = -3**i** + 4**j** + 2**k**,

$$L(-3,4,2) = L(-3\boldsymbol{i}+4\boldsymbol{j}+2\boldsymbol{k}) = -3L(\boldsymbol{i})+4L(\boldsymbol{j})+2L(\boldsymbol{k})$$
$$= -3(2,-1)+4(3,1)+2(-1,2) = (4,11)$$

Exercise

1. Suppose that $T: R^5 \to R^2$ and ***T(x) = Ax*** for some matrix ***A*** and each ***x*** in $\boldsymbol{R^5}$. How many rows and columns do ***A*** have?

2. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Give a geometric description of the transformation $x \to Ax$.

3. The line segment from **0** to a vector ***u*** is the set of points of the form *t**u***, where $0 \le t \le 1$. Show that a linear transformation ***T*** maps this segment into the segment between **0** and ***T(u)***.

4. Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, u = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}, and \begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix}$. Define T:$R^3 \to R^3$ by T (x) = Ax. Find T (u) and T (v).

In exercises 5 and 6, with T defined by T (x) = Ax, find an x whose image under T is b, and determine if x is unique.

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

6. $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -4 \\ 3 & 2 & 1 \\ -2 & -1 & -2 \end{bmatrix}, b=\begin{bmatrix} 1 \\ -5 \\ -7 \\ 3 \end{bmatrix}$

Find all x in R^4 that are mapped into the zero vector by the transformation $x \to Ax$.

7. $A=\begin{bmatrix} 1 & 2 & -7 & 5 \\ 0 & 1 & -4 & 0 \\ 1 & 0 & 1 & 6 \\ 2 & -1 & 6 & 8 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 3 & 4 & -3 \\ 0 & 1 & 3 & -2 \\ 3 & 7 & 6 & -5 \end{bmatrix}$

9. Let $b=\begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix}$ and let A be the matrix in exercise 8. Is b in the range of the linear transformation $x \to Ax$?

10. Let $b=\begin{bmatrix} 9 \\ 5 \\ 0 \\ -9 \end{bmatrix}$ and let A be the matrix in exercise 7. Is b in the range of the linear transformation $x \to Ax$?

Let T (x) = Ax for x in R^2.

(a) On a rectangular coordinate system, plot the vectors u, v, T (u) and T (v).
(b) Give a geometric description of what T does to a vector x in R^2.

11. $A=\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, u=\begin{bmatrix} 5 \\ 2 \end{bmatrix}, \text{and } v=\begin{bmatrix} 3 \\ -1 \end{bmatrix}$

12. $A=\begin{bmatrix} .5 & 0 \\ 0 & .5 \end{bmatrix}, u=\begin{bmatrix} 4 \\ 2 \end{bmatrix}, \text{and } v=\begin{bmatrix} -5 \\ -2 \end{bmatrix}$

13. Let T: $R^2 \to R^2$ be a linear transformation that maps $u=\begin{bmatrix} 1 \\ 5 \end{bmatrix} \text{into} \begin{bmatrix} 2 \\ 0 \end{bmatrix} \text{and maps } v=\begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{into} \begin{bmatrix} 1 \\ -4 \end{bmatrix}$. Use the fact that T is linear find the images under T of 2u, 3v, and 2u + 3v.

14. Let $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, y_1 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}, and\ y_2 = \begin{bmatrix} -2 \\ 7 \end{bmatrix}.$ Let T: $R^2 \to R^2$ be a linear transformation that maps e_1 into y_1 and maps e_2 into y_2. Find the images of $\begin{bmatrix} 7 \\ 6 \end{bmatrix} and \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$

15. Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, v_1 = \begin{bmatrix} -7 \\ 4 \end{bmatrix} and\ v_2 = \begin{bmatrix} 3 \\ -8 \end{bmatrix}.$ Let T: $R^2 \to R^2$ be a linear transformation that maps x into $x_1 v_1 + x_2 v_2$. Find a matrix A such that T (x) is Ax for each x.

Lecture 10

The Matrix of a Linear Transformation

<u>Outline</u>

- Matrix of a Linear Transformation.
- Examples, Geometry of Transformation, Reflection and Rotation
- Existence and Uniqueness of solution of ***T(x)=0***

In the last lecture we discussed that every linear transformation from $\boldsymbol{R^n}$ to $\boldsymbol{R^m}$ is actually a matrix transformation $x \to Ax$, where A is a matrix of order $m \times n$. First see an example

<u>Example 1</u> The columns of $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Suppose ***T*** is a linear transformation from $\boldsymbol{R^2}$ into $\boldsymbol{R^3}$ such that

$$T(e_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad and \quad T(e_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

with no additional information, find a formula for the image of an arbitrary ***x*** in $\boldsymbol{R^2}$.

<u>Solution</u> Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1 e_1 + x_2 e_2$

Since ***T*** is a linear transformation, $T(x) = x_1 T(e_1) + x_2 T(e_2)$

$$T(x) = x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix}$$

$$Hence \quad T(x) = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix}$$

<u>Theorem</u> Let $T : R^n \to R^m$ be a linear transformation. Then there exists a unique matrix ***A*** such that $T(x) = Ax$ for all ***x*** in $\boldsymbol{R^n}$

In fact, ***A*** is the $m \times n$ matrix whose jth column is the vector ***T*** **(*e_j*)**, where e_j is the jth column of the identity matrix in $\boldsymbol{R^n}$.

$$A = \begin{bmatrix} T(e_1) & \ldots & T(e_n) \end{bmatrix}$$

<u>Proof</u> Write

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{bmatrix} + \ldots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \ldots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

$$= x_1 e_1 + \ldots + x_n e_n = \begin{bmatrix} e_1 & \ldots & e_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} e_1 & \ldots & e_n \end{bmatrix} x$$

Since T is Linear, So

$$T(x) = T(x_1 e_1 + \ldots + x_n e_n) = x_1 T(e_1) + \ldots + x_n T(e_n)$$

$$= \begin{bmatrix} T(e_1) & \ldots & T(e_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax \qquad (1)$$

The matrix ***A*** in (1) is called the **standard matrix for the linear transformation *T***. We know that every linear transformation from $\boldsymbol{R^n}$ to $\boldsymbol{R^m}$ is a matrix transformation and vice versa.

The term linear transformation focuses on a property of a mapping, while matrix transformation describes how such a mapping is implemented, as the next three examples illustrate.

<u>Example 2</u> Find the standard matrix ***A*** for the dilation transformation ***T (x) = 3x***, $x \in R^2$.

<u>Solution</u> Write

$$T(e_1) = 3e_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \quad and \quad T(e_2) = 3e_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

Example 3 Let $L: R^3 \to R^3$ is the linear operator defined by $L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ y-z \\ x+z \end{bmatrix}$.

Find the standard matrix representing ***L*** and verify ***L (x) = Ax***.

Solution

The standard matrix ***A*** representing ***L*** is the 3 x 3 matrix whose columns are ***L*** ($\boldsymbol{e_1}$), ***L*** ($\boldsymbol{e_2}$), and ***L*** ($\boldsymbol{e_3}$) respectively. Thus

$$L(e_1) = L\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1+0 \\ 0-0 \\ 1+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = col_1(A)$$

$$L(e_2) = L\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0+1 \\ 1-0 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = col_2(A)$$

$$L(e_3) = L\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0+0 \\ 0-1 \\ 0+1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = col_3(A)$$

Hence $$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

Now $$\boldsymbol{Ax} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y-z \\ x+z \end{bmatrix} = L(x)$$

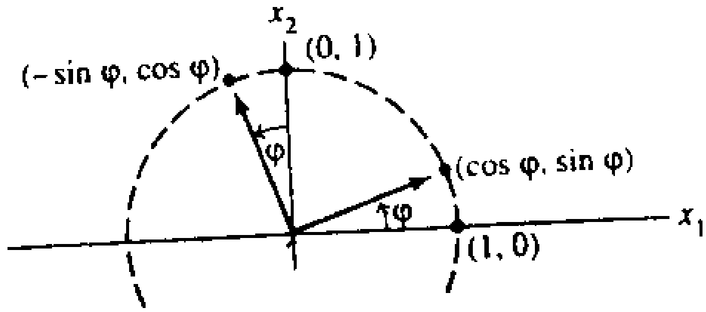
Hence verified.

Example 4 Let $T: R^2 \to R^2$ be the transformation that rotates each point in $\boldsymbol{R^2}$ through an angle φ, with counterclockwise rotation for a positive angle. We could show geometrically that such a transformation is linear. Find the standard matrix ***A*** of this transformation.

Solution $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ rotates into $\begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ rotates into $\begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}$.

See figure below.

By above theorem $A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$



A rotation transformation

Example 5 A reflection with respect to the ***x***-axis of a vector ***u*** in $\boldsymbol{R^2}$ is defined by the

linear operator $L(u) = L\left(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}\right) = \begin{bmatrix} a_1 \\ -a_2 \end{bmatrix}$.

Then $L(e_1) = L\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $L(e_2) = L\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

Hence the standard matrix representing ***L*** is $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Thus we have $L(u) = Au = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ -a_2 \end{bmatrix}$

To illustrate a reflection with respect to the x-axis in computer graphics, let the triangle ***T*** have vertices (-1, 4), (3, 1), and (2, 6).

To reflect ***T*** with respect to ***x***-axis, we let $u_1 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, u_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, u_3 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ and compute the images ***L*** $(\boldsymbol{u_1})$, ***L*** $(\boldsymbol{u_2})$, and ***L*** $(\boldsymbol{u_3})$ by forming the products

$$Au_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \end{bmatrix},$$

$$Au_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix},$$

$$Au_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ -6 \end{bmatrix}.$$

Thus the image of ***T*** has vertices (-1, -4), (3, -1), and (2, -6).

Geometric Linear Transformations of R^2

Examples 3-5 illustrate linear transformations that are described geometrically. In example 4 transformations is a rotation in the plane. It rotates each point in the plane through an angle φ. Example 5 is reflection in the plane.

Existence and Uniqueness of the solution of *T(x)=b*

The concept of a linear transformation provides a new way to understand existence and uniqueness questions asked earlier. The following two definitions give the appropriate terminology for transformations.

Definition A mapping $T : R^n \to R^m$ is said to be onto $\boldsymbol{R^m}$ if each ***b*** in $\boldsymbol{R^m}$ is the image of at least one ***x*** in $\boldsymbol{R^n}$.

OR

Equivalently, ***T*** is onto $\boldsymbol{R^m}$ if for each ***b*** in $\boldsymbol{R^m}$ there exists at least one solution of ***T (x) = b***. "Does ***T*** map $\boldsymbol{R^n}$ onto $\boldsymbol{R^m}$?" is an existence question.

The mapping ***T*** is not onto when there is some ***b*** in $\boldsymbol{R^m}$ such that the equation ***T (x) = b*** has no solution.

Definition A mapping $T : R^n \to R^m$ is said to be one-to-one (or 1:1) if each ***b*** in $\boldsymbol{R^m}$ is the image of at most one ***x*** in $\boldsymbol{R^n}$.

OR

Equivalently, **T** is one-to-one if for each ***b*** in $\boldsymbol{R^m}$ the equation ***T (x) = b*** has either a unique solution or none at all, "Is ***T*** one-to-one?" is a uniqueness question.

The mapping ***T*** is not one-to-one when some ***b*** in $\boldsymbol{R^m}$ is the image of more than one vector in $\boldsymbol{R^n}$. If there is no such ***b***, then ***T*** is one-to-one.

Example 6 Let ***T*** be the linear transformation whose standard matrix is

$$A = \begin{bmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

Does ***T*** map $\boldsymbol{R^4}$ onto $\boldsymbol{R^3}$? Is ***T*** a one-to-one mapping?

Solution Since ***A*** happens to be in echelon form, we can see at once that ***A*** has a pivot position in each row.

We know that for each ***b*** in $\boldsymbol{R^3}$, the equation ***Ax = b*** is consistent. In other words, the linear transformation ***T*** maps $\boldsymbol{R^4}$ (its domain) onto $\boldsymbol{R^3}$.

However, since the equation ***Ax = b*** has a free variable (because there are four variables and only three basic variables), each ***b*** is the image of more than one ***x***. That is, ***T*** is not one-to-one.

Theorem Let $T : R^n \to R^m$ be a linear transformation. Then ***T*** is one-to-one if and only if the equation ***T (x) = 0*** has only the trivial solution.

Proof: Since ***T*** is linear, ***T (0) = 0*** if ***T*** is one-to-one, then the equation ***T (x) = 0*** has at most one solution and hence only the trivial solution. If ***T*** is not one-to-one, then there is a ***b*** that is the image of at least two different vectors in $\boldsymbol{R^n}$ (say, ***u*** and ***v*)**.
That is, ***T (u) = b*** and ***T (v) = b***.

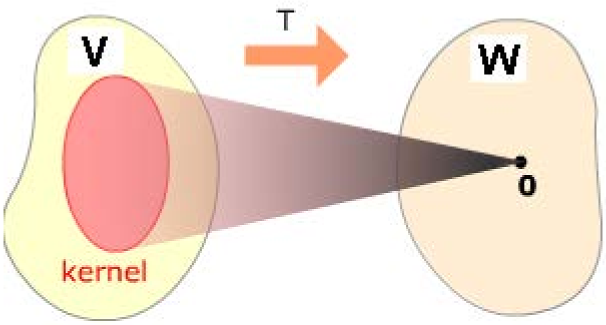
But then, since ***T*** is linear $T(u-v) = T(u) = b - b = 0$

The vector ***u – v*** is not zero, since $u \neq v$. Hence the equation ***T (x) = 0*** has more than one solution. So either the two conditions in the theorem are both true or they are both false.

Kernel of a Linear Transformation

Let $T : V \to W$ be a linear transformation. Then kernel of T (usually written as Ker*T*), is the set of those elements in V which maps onto the zero vector in W .Mathematically:

$$KerT = \{ v \in V \mid T(v) = 0 \ in \ W \}$$



Remarks

i) Ker$T$ is subspace of V

ii) T is one-one iff $KerT = 0$ in V

One-One Linear Transformation

Let $T: V \to W$ be a linear transformation. Then T is said to be one-one if for any $u, v \in V$ with $u \neq v$ implies $Tu \neq Tv$. Equivalently if $Tu = Tv$ then $u = v$.

T is said to be one-to-one or bijective if

i) T is one-one

ii)T is onto

Theorem Let $T: R^n \to R^m$ be a linear transformation and let $\boldsymbol{A}$ be the standard matrix for $\boldsymbol{T}$. Then

(a) $\boldsymbol{T}$ maps $\boldsymbol{R^n}$ onto $\boldsymbol{R^m}$ if and only if the columns of $\boldsymbol{A}$ span $\boldsymbol{R^m}$;

(b) $\boldsymbol{T}$ is one-to-one if and only if the columns of **A** are linearly independent.

Proof:

(a) The columns of $\boldsymbol{A}$ span $\boldsymbol{R^m}$ if and only if for each $\boldsymbol{b}$ the equation $\boldsymbol{Ax = b}$ is consistent – in other words, if and only if for every $\boldsymbol{b}$, the equation $\boldsymbol{T(x) = b}$ has at least one solution. This is true if and only if $\boldsymbol{T}$ maps $\boldsymbol{R^n}$ onto $\boldsymbol{R^m}$.

(b) The equations $\boldsymbol{T(x) = 0}$ and $\boldsymbol{Ax = 0}$ are the same except for notation. So $\boldsymbol{T}$ is one-to-one if and only if $\boldsymbol{Ax = 0}$ has only the trivial solution. This happens if and only if the columns of $\boldsymbol{A}$ are linearly independent.

We can also write column vectors in rows, using parentheses and commas. Also, when we apply a linear transformation $\boldsymbol{T}$ to a vector – say, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (x_1, x_2)$ we write $T(x_1, x_2)$ instead of the more formal $T\left((x_1, x_2)\right)$.

Example 7 Let $\boldsymbol{T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)}$.

Show that $\boldsymbol{T}$ is a one-to-one linear transformation.

Does $\boldsymbol{T}$ map $\boldsymbol{R^2}$ onto $\boldsymbol{R^3}$?

<u>Solution</u> When ***x*** and ***T (x)*** are written as column vectors, it is easy to see that ***T*** is described by the equation

$$\begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} \qquad (4)$$

so ***T*** is indeed a linear transformation, with its standard matrix ***A*** shown in (4). The columns of ***A*** are linearly independent because they are not multiples. Hence ***T*** is one-to-one. To decide if ***T*** is onto $\boldsymbol{R^3}$, we examine the span of the columns of ***A***. Since ***A*** is 3×2, the columns of ***A*** span $\boldsymbol{R^3}$ if and only if ***A*** has 3 pivot positions. This is impossible, since ***A*** has only 2 columns. So the columns of ***A*** do not span $\boldsymbol{R^3}$ and the associated linear transformation is not onto $\boldsymbol{R^3}$.

Exercises

1. Let $T: R^2 \to R^2$ be transformation that first performs a horizontal shear that maps e_2 into $e_2 - .5e_1$ (but leaves e_1 unchanged) and then reflects the result in the $\boldsymbol{x_2}$ – axis. Assuming that ***T*** is linear, find its standard matrix.

Assume that T is a linear transformation. Find the standard matrix of T.

2. $T: R^2 \to R^3, T(1,0) = (4,-1,2)\ and\ T(0,1) = (-5,3,-6)$

3. $T: R^3 \to R^2, T(e_1) = (1,4), T(e_2) = (-2,9), and\ T(e_3) = (3,-8),$ where e_1, e_2, and e_3 are the columns of the identity matrix.

4. $T: R^2 \to R^2$ rotates points clockwise through π radians.

5. $T: R^2 \to R^2$ is a "vertical shear" transformation that maps e_1 into $e_1 + 2e_2$ but leaves the vector e_2 unchanged.

6. $T: R^2 \to R^2$ is a "horizontal shear" transformation that maps e_2 into $e_2 - 3e_1$ but leaves the vector e_1 unchanged.

7. $T: R^3 \to R^3$ projects each point (x_1, x_2, x_3) vertically onto the x_1x_2-plane (where x_3=0).

8. $T: R^2 \to R^2$ first performs a vertical shear mapping e_1 into $e_1 - 3e_2$(leaving e_2 unchanged) and then reflects the result in the x_2-axis.

9. $T: R^2 \to R^2$ first rotates points counterclockwise through π/4 radians and then reflects the result in the x_2-axis.

Show that T is a linear transformation by finding a matrix that implements the mapping. Note that x_1, x_2, … are not vectors but are entries in vectors.

10. $T(x_1, x_2, x_3, x_4) = (x_1 + x_2, x_2 + x_3, x_3 + x_4, 0)$

11. $T(x_1, x_2, x_3) = (3x_2 - x_3, x_1 + 4x_2 + x_3)$

12. $T(x_1, x_2, x_3, x_4) = 3x_1 - 4x_2 + 8x_4$

13. Let $T: R^2 \to R^2$ be a linear transformation such that $T(x_1, x_2) = (x_1 + x_2, 4x_1 + 7x_2)$. Find ***x*** such that T (x) =(-2, -5).

13. Let $T : R^2 \to R^3$ be a linear transformation such that $T(x_1, x_2) = (x_1 + 2x_2, -x_1 - 3x_2, -3x_1 - 2x_2)$. Find ***x*** such that T (x) =(-4, 7, 0).

In exercises 14 and 15, let T be the linear transformation whose standard matrix is given.

14. Decide if T is one-to-one mapping. Justify your answer.

$$\begin{bmatrix} -5 & 10 & -5 & 4 \\ 8 & 3 & -4 & 7 \\ 4 & -9 & 5 & -3 \\ -3 & -2 & 5 & 4 \end{bmatrix}$$

15. Decide if T maps R^5 onto R^5. Justify your answer.

$$\begin{bmatrix} 4 & -7 & 3 & 7 & 5 \\ 6 & -8 & 5 & 12 & -8 \\ -7 & 10 & -8 & -9 & 14 \\ 3 & -5 & 4 & 2 & -6 \\ -5 & 6 & -6 & -7 & 3 \end{bmatrix}$$

Lecture 11

Matrix Operations

(i-j)th Element of a matrix

Let A be an $m \times n$ matrix, where m and n are number of rows and number of columns respectively, then a_{ij} represents the ***i***-th row and ***j***-th column entry of the matrix. For example a_{12} represents 1st row and 2nd column entry.

Similarly a_{32} represents 3rd row and 2nd column entry. The columns of ***A*** are vectors in $\boldsymbol{R^m}$ and are denoted by (boldface) $\boldsymbol{a_1}, \boldsymbol{a_2}, \cdots, \boldsymbol{a_n}$.

These columns are $A = [\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \ldots \quad \boldsymbol{a_n}]$

The number a_{ij} is the ***i***-*th* entry (from the top) of ***j***-*th* column vector a_j.

$$\begin{array}{cc}
 & \textit{Column } j \\
\textit{Row}\ \ i & \begin{bmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \ldots & a_{ij} & \ldots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \ldots & a_{mj} & \ldots & a_{mn} \end{bmatrix} = A \\
 & \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \boldsymbol{a_1} & \boldsymbol{a_j} & \boldsymbol{a_n} \end{array}
\end{array}$$

Figure 1 Matrix notation.

Definitions

A **diagonal matrix** is a square matrix whose non-diagonal entries are zero.

$$D = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix}$$

The diagonal entries in $A = \left[a_{ij} \right]$ are $a_{11}, a_{22}, a_{33}, \cdots$ and they form the **main diagonal** of ***A.***

For example $\begin{bmatrix} 5 & 0 \\ 0 & 7 \end{bmatrix}$ $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 11 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ are all diagonal matrices.

Null Matrix or Zero Matrix

An $m \times n$ matrix whose entries are all zero is a Null or **zero matrix** and is always written as ***O***. A null matrix may be of any order.

For example $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3 x 3 3 x 2 4 x 5

are all Zero Matrices

Equal Matrices

Two matrices are said to be **equal** if they have the same size (i.e., the same number of rows and columns) and same corresponding entries.

Example 1 Consider the matrices

$$A = \begin{bmatrix} 2 & 1 \\ 3 & x+1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 4 & 0 \end{bmatrix}$$

The matrices ***A*** and ***B*** are equal if and only if $x+1 = 5$ or $x = 4$. There is no value of x for which ***A = C***, since ***A*** and ***C*** have different sizes.

If ***A*** and B are $m \times n$ matrices, then the **sum, *A + B***, is the $m \times n$ matrix whose columns are the sums of the corresponding columns in ***A*** and ***B***. Each entry in ***A + B*** is the sum of the corresponding entries in ***A*** and ***B***. The sum ***A + B*** is defined only when ***A*** and ***B*** are of the same size.

If *r* is a scalar and ***A*** is a matrix, then the **scalar multiple** ***rA*** is the matrix whose columns are *r* times the corresponding columns in ***A***.

Note: Negative of a matrix A is defined as $-A$ to mean $(-1)A$ and the difference of A and B is written as $A-B$, which means $A + (-1)B$.

Example 2 Let $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$

Then $A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$

But ***A + C*** is not defined because ***A*** and ***C*** have different sizes.

$$2B = 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}$$

$$A - 2B = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix}$$

Theorem 1 Let ***A, B,*** and ***C*** are matrices of the same size, and let *r* and *s* are scalars.

a. ***A + B = B + A***
b. ***(A + B) + C = A + (B + C)***
c. ***A + 0 = A***
d. *r* ***(A + B)***= *r* ***A*** + *r* ***B***
e. *(r + s)* ***A*** = *r* ***A*** + *s* ***A***
f. *r* ***(sA)*** = *(rs)* ***A***

Each equality in Theorem 1 can be verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal. Size is no problem because ***A, B,*** and ***C*** are equal in size. The equality of columns follows immediately from analogous properties of vectors.

For instance, if the *jth* columns of ***A, B,*** and ***C*** are a_j, b_j and c_j, respectively, then the *jth* columns of ***(A + B) + C*** and ***A + (B + C)*** are

$$(a_j + b_j) + c_j \qquad and \qquad a_j + (b_j + c_j)$$

respectively. Since these two vector sums are equal for each *j*, property (b) is verified.

Because of the associative property of addition, we can simply write ***A + B + C*** for the sum, which can be computed either as ***(A + B) + C*** or ***A + (B + C)***. The same applies to sums of four or more matrices.

Matrix Multiplication

Multiplying an m×n matrix with an n×p matrix results in an m×p matrix. If many matrices are multiplied together, and their dimensions are written in a list in order, e.g. m×n, n×p, p×q, q×r, the size of the result is given by the first and the last numbers (m×r).

It is important to keep in mind that this definition requires the number of columns of the first factor A to be the same as the number of rows of the second factor B. When this condition is satisfied, the sizes of A and B are said to conform for the product AB. If the sizes of A and B do not conform for the product AB, then this product is undefined.

Definition If ***A*** is an $m \times n$ matrix, and if ***B*** is an $n \times p$ matrix with columns $b_1, \cdots, b_p$, then the product ***AB*** is the $m \times p$ matrix whose columns are $Ab_1, \cdots, Ab_p$.

That is $$AB = A\begin{bmatrix} b_1 & b_2 & \ldots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_p \end{bmatrix}$$

This definition makes equation (1) true for all $\boldsymbol{x}$ in $\boldsymbol{R}^{p}$. Equation (1) proves that the composite mapping ***(AB)*** is a linear transformation and that its standard matrix is ***AB***. Multiplication of matrices corresponds to composition of linear transformations.

A convenient way to determine whether ***A*** and ***B*** conform for the product ***AB*** and, if so, to find the size of the product is to write the sizes of the factors side by side as in Figure below (the size of the first factor on the left and the size of the second factor on the right).

$$\begin{array}{ccccc} \boldsymbol{A} & & \boldsymbol{B} & = & \boldsymbol{AB} \\ m \times s & & s \times n & & m \times n \end{array}$$

Inside

Outside

If the inside numbers are the same, then the product ***AB*** is defined and the outside numbers then give the size of the product.

Example 3 Compute ***AB***, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ *and* $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$

Solution: Here $\boldsymbol{B} = [\boldsymbol{b_1} \quad \boldsymbol{b_2} \quad \boldsymbol{b_3}]$, therefore

$$Ab_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad Ab_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad Ab_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Then

$$AB = A[b_1 \quad b_2 \quad b_3] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$$\boldsymbol{Ab_1} \quad \boldsymbol{Ab_2} \quad \boldsymbol{Ab_3}$$

Note from the definition of ***AB*** that its first column, $\boldsymbol{Ab_1}$, is a linear combination of the columns of ***A***, using the entries in $\boldsymbol{b_1}$ as weights. The same holds true for each column of ***AB***. Each column of ***AB*** is a linear combination of the columns of ***A*** using weights from the corresponding column of ***B***.

<u>Example 4</u> Find the product ***AB*** for

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$$

Solution It follows from definition that the product ***AB*** is formed in a column-by-column manner by multiplying the successive columns of ***B*** by ***A***. The computations are

$$\begin{array}{ccc} c_1 & c_2 & c_3 \end{array}$$

$$\left[\begin{array}{c|c|c} 1 & 2 & 4 \\ 2 & 6 & 0 \end{array}\right] \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} = 4c_1 + 0c_2 + 2c_3 = (4)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (0)\begin{bmatrix} 2 \\ 6 \end{bmatrix} + (2)\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 12 \\ 8 \end{bmatrix}$$

Similarly, $$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix} = (1)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (-1)\begin{bmatrix} 2 \\ 6 \end{bmatrix} + (7)\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 27 \\ -4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix} = (4)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (3)\begin{bmatrix} 2 \\ 6 \end{bmatrix} + (5)\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 30 \\ 26 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = (3)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (1)\begin{bmatrix} 2 \\ 6 \end{bmatrix} + (2)\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 13 \\ 12 \end{bmatrix}$$

Thus, $$AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

<u>Example 5</u> (An Undefined Product) Find the product ***BA*** for the matrices

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad \textit{and} \quad B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$$

Solution The number of columns of B is not equal to number of rows of A so BA multiplication is not possible.

The matrix ***B*** has size 3×4 and the matrix ***A*** has size 2×3. The "inside" numbers are not the same, so the product ***BA*** is undefined.

Obviously, the number of columns of ***A*** must match the number of rows in ***B*** in order for a linear combination such as $\boldsymbol{Ab_1}$ to be defined. Also, the definition of ***AB*** shows that ***AB*** has the same number of rows as ***A*** and the same number of columns as ***B***.

Example 6 If ***A*** is a 3×5 matrix and ***B*** is a 5×2 matrix, what are the sizes of ***AB*** and ***BA***, if they are defined?

Solution The product of matrices A and B of orders 3×5 and 5×2 will result in 3×2 matrix AB.
But for BA we have 5×2 and 3×5, here number of columns in1st matrix are 2 which is not equal to number of rows in 2nd matrix. So BA is not possible.

Since ***A*** has 5 columns and ***B*** has 5 rows, the product ***AB*** is defined and is a 3×2 matrix:

$$\begin{array}{ccccc} A & & B & & AB \\ \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & & \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} & = & \begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} \\ 3\times 5 & & 5\times 2 & & 3\times 2 \end{array}$$

Match

Size of AB

The product ***BA*** is not defined because the 2 columns of ***B*** do not match the 3 rows of ***A***.

The definition of ***AB*** is important for theoretical work and applications, but the following rule provides a more efficient method for calculating the individual entries in ***AB*** when working small problems by hand.

Row-Column Rule for Computing *AB*

Explanation

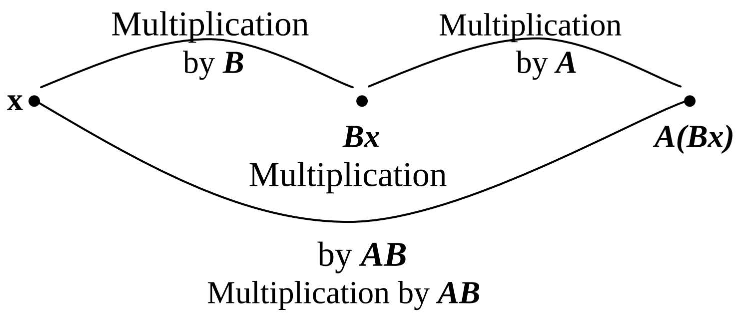
If a matrix ***B*** is multiplied with a vector ***x***, it transforms ***x*** into a vector ***Bx***. If this vector is then multiplied in turn by a matrix ***A***, the resulting vector is ***A (Bx)***.

$$\begin{tikzcd} \mathbf{x} \arrow[r, bend left, "\text{Multiplication by } \boldsymbol{B}"] & \boldsymbol{Bx} \arrow[r, bend left, "\text{Multiplication by } \boldsymbol{A}"] & \boldsymbol{A(Bx)} \end{tikzcd}$$

Multiplication by ***B*** and then ***A***

Thus ***A(Bx)*** is produced from ***x*** by a composition of mappings. Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by ***AB***, so that

$$\boldsymbol{A\,(Bx) = (AB)\,x} \text{---} (1)$$



If ***A*** is $m \times n$, ***B*** is $n \times p$, and ***x*** is in $\boldsymbol{R^p}$, denote the columns of ***B*** by $b_1, \cdots, b_p$ and the entries in ***x*** by $x_1, \cdots, x_p$, then $Bx = x_1 b_1 + x_2 b_2 + \cdots + x_p b_p$

By the linearity of multiplication by ***A***,

$$\begin{aligned} A(Bx) &= A(x_1 b_1) + A(x_2 b_2) + \cdots + A(x_p b_p) \\ &= x_1 A b_1 + x_2 A b_2 + \cdots + x_p A b_p \end{aligned}$$

The vector ***A (Bx)*** is a linear combination of the vectors $Ab_1, \cdots, Ab_p$, using the entries in ***x*** as weights. If we rewrite these vectors as the columns of a matrix, we have

$$A(Bx) = \begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_p \end{bmatrix} x$$

Thus multiplication by $\begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_p \end{bmatrix}$ transforms ***x*** into ***A(Bx)***.
We have found the matrix we sought!

Row-Column Rule for Computing *AB*

If the product ***AB*** is defined, then the entry in row ***i*** and column ***j*** of ***AB*** is the sum of the products of corresponding entries from row ***i*** of ***A*** and column ***j*** of ***B***. If $\boldsymbol{(AB)_{ij}}$ denotes the *(i, j)* – entry in ***AB***, and if ***A*** is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots + a_{in} b_{nj}$$

To verify this rule, let $B = \begin{bmatrix} b_1 & \ldots & b_p \end{bmatrix}$. Column ***j*** of ***AB*** is $\boldsymbol{Ab_j}$, and we can compute $\boldsymbol{Ab_j}$. The *ith* entry in $\boldsymbol{Ab_j}$ is the sum of the products of corresponding entries from row ***i*** of ***A*** and the vector $\boldsymbol{b_j}$, which is precisely the computation described in the rule for computing the *(i, j)* – entry of ***AB***.

Finding Specific Entries in a Matrix Product Sometimes we will be interested in finding a specific entry in a matrix product without going through the work of computing the entire column that contains the entry.

Example 7 Use the row-column rule to compute two of the entries in ***AB*** for the matrices in Example 3.

Solution: To find the entry in row 1 and column 3 of ***AB***, consider row 1 of ***A*** and column 3 of ***B***. Multiply corresponding entries and add the results, as shown below:

$$\downarrow$$

$$AB = \rightarrow\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6)+3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

For the entry in row 2 and column 2 of ***AB***, use row 2 of ***A*** and column 2 of ***B***:

$$\downarrow$$

$$\rightarrow\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3)+-5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$

Example 8 Use the dot product rule to compute the individual entries in the product of ***AB*** where

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad \textit{and} \quad B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}.$$

Solution Since ***A*** has size 2×3 and ***B*** has size 3×4, the product ***AB*** is a 2×4 matrix of the form

$$AB = \begin{bmatrix} r_1(A)\times c_1(B) & r_1(A)\times c_2(B) & r_1(A)\times c_3(B) & r_1(A)\times c_4(B) \\ r_2(A)\times c_1(B) & r_2(A)\times c_2(B) & r_2(A)\times c_3(B) & r_2(A)\times c_4(B) \end{bmatrix}$$

where $r_1(A)$ and $r_2(A)$ are the row vectors of ***A*** and $c_1(B), c_2(B), c_3(B)$ *and* $c_4(B)$ are the column vectors of ***B***. For example, the entry in row 2 and column 3 of ***AB*** can be computed as

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}\begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \boxed{26} & \square \end{bmatrix}$$

$$(2\times 4)+(6\times 3)+(0\times 5) = 26$$

and the entry in row 1 and column 4 of ***AB*** can be computed as

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}\begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \boxed{13} \\ \square & \square & \square & \square \end{bmatrix}$$

$$(1\times 3)+(2\times 1)+(4\times 2)=13$$

Here is the complete set of computations:

$$(AB)_{11} = (1\times 4)+(2\times 0)+(4\times 2)=12$$
$$(AB)_{12} = (1\times 1)+(2\times -1)+(4\times 7)=27$$
$$(AB)_{13} = (1\times 4)+(2\times 3)+(4\times 5)=30$$
$$(AB)_{14} = (1\times 3)+(2\times 1)+(4\times 2)=13$$
$$(AB)_{21} = (2\times 4)+(6\times 0)+(0\times 2)=8$$
$$(AB)_{22} = (2\times 1)+(6\times -1)+(0\times 7)=-4$$
$$(AB)_{23} = (2\times 4)+(6\times 3)+(0\times 5)=26$$
$$(AB)_{24} = (2\times 3)+(6\times 1)+(0\times 2)=12$$

Finding Specific Rows and Columns of a Matrix Product

The specific column of ***AB*** is given by the formula

$$AB = A\begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_n \end{bmatrix}$$

Similarly, the specific row of ***AB*** is given by the formula $AB = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} B = \begin{bmatrix} a_1 B \\ a_2 B \\ \vdots \\ a_m B \end{bmatrix}$

<u>Example 9</u> Find the entries in the second row of ***AB***, where

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

Solution: By the row-column rule, the entries of the second row of ***AB*** come from row 2 of ***A*** (and the columns of ***B***):

$$\begin{array}{cc} & \downarrow \quad \downarrow \\ \rightarrow \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} & \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} \end{array} = \begin{bmatrix} \square & \square \\ -4+21-12 & 6+3-8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

Example 10 **(Finding a Specific Row and Column of AB)**

Let $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$

Find the second column and the first row of ***AB***.

Solution $c_2(AB) = Ac_2(B) = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix} = \begin{bmatrix} 27 \\ -4 \end{bmatrix}$

$$r_1(AB) = r_1(A)B = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix}\begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \end{bmatrix}$$

Properties of Matrix Multiplication

These are standard properties of matrix multiplication. Remember that I_m represents the $m \times m$ identity matrix and $I_m x = x$ for all ***x*** *belong to* $\boldsymbol{R^m}$.

Theorem 2 Let ***A*** be $m \times n$, and let ***B*** and ***C*** have sizes for which the indicated sums and products are defined.

- **a.** ***A (BC) = (AB) C*** (associative law of multiplication)
- **b.** ***A (B + C) = AB + AC*** (left distributive law)
- **c.** ***(B + C)A = BA + CA*** (right distributive law)
- ***d.*** *r* ***(AB) = (****r* ***A)B = A(****r* ***B)*** (for any scalar *r*)
- **e.** $I_m A = A = A I_m$ (identity for matrix multiplication)

Proof. Properties (b) to (e) are considered exercises for you. We start property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known (or easy to check) that the composition of functions is associative.

Here is another proof of (a) that rests on the "column definition" of the product of two matrices. Let $C = \begin{bmatrix} c_1 & \ldots & c_p \end{bmatrix}$

By definition of matrix multiplication $BC = \begin{bmatrix} Bc_1 & \ldots & Bc_p \end{bmatrix}$

$$A(BC) = \begin{bmatrix} A(Bc_1) & \ldots & A(Bc_p) \end{bmatrix}$$

From above, we know that ***A(Bx) = (AB)x*** for all ***x***, so

$$A(BC) = \begin{bmatrix} (AB)c_1 & \ldots & (AB)c_p \end{bmatrix} = (AB)C$$

The associative and distributive laws say essentially that pairs of parentheses in matrix expressions can be inserted and deleted in the same way as in the algebra of real numbers. In particular, we can write ***ABC*** for the product, which can be computed as ***A(BC)*** or as ***(AB)C***. Similarly, a product ***ABCD*** of four matrices can be computed as

A(BCD) or ***(ABC)D*** or ***A(BC)D***, and so on. It does not matter how we group the matrices when computing the product, so long as the left-to-right order of the matrices is preserved.

The left-to-right order in products is critical because, in general, ***AB*** and ***BA*** are not the same. This is not surprising, because the columns of ***AB*** are linear combinations of the columns of ***A***, whereas the columns of ***BA*** are constructed from the columns of ***B***.

If ***AB = BA***, we say that ***A*** and ***B*** **commute** with one another.

Example 11 Let $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ *and* $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$

Show that these matrices don not commute, i.e. $AB \neq BA$.

Solution:

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 10+4 & 0+3 \\ 6-8 & 0-6 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10+0 & 2-0 \\ 20+9 & 4-6 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$

For emphasis, we include the remark about commutativity with the following list of important differences between matrix algebra and ordinary algebra of real numbers.

WARNINGS

1. In general, $AB \neq BA$. Clear from the Example # 11.

2. The cancellation laws do not hold for matrix multiplication. That is, if $AB = AC$, then it is not true in general that $B = C$.

For example: Consider the following three matrices

$$A = \begin{bmatrix} -3 & 2 \\ -6 & 4 \end{bmatrix} \quad B = \begin{bmatrix} -1 & 2 \\ 3 & -2 \end{bmatrix} \quad C = \begin{bmatrix} -1 & 2 \\ 3 & -2 \end{bmatrix}$$

$$AB = \begin{bmatrix} 9 & -10 \\ 18 & -20 \end{bmatrix} = \text{A}C \quad \text{But } B \neq C$$

3. If a product ***AB*** is the zero matrix, you cannot conclude in general that either ***A = 0*** *or* ***B = 0***.

For example:

$$If \;\; AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \textit{ then it can be either}$$

$$A = \begin{bmatrix} 1 & 4 \\ 6 & 2 \end{bmatrix} \ and \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad or$$

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \ and \quad B = \begin{bmatrix} 1 & 4 \\ 6 & 2 \end{bmatrix}$$

Example

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 15-1 & 0+3 \\ 6-8 & 0-6 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

Powers of a Matrix If ***A*** is an $n \times n$ matrix and if *k* is a positive integer, $\boldsymbol{A^k}$ denotes the product of *k* copies of ***A***, $A^k = \underbrace{A \ldots A}_{k}$ Also, we interpret $\boldsymbol{A^0}$ as ***I***.

Transpose of a Matrix Given an $m \times n$ matrix ***A***, the **transpose** of ***A*** is the $n \times m$ matrix, denoted by $\boldsymbol{A^t}$, whose columns are formed from the corresponding rows of ***A***.
OR, if ***A*** is an *m x n* matrix, then transpose of ***A*** is denoted by $\boldsymbol{A^t}$, is defined to be the *nxm* matrix that is obtained by making the rows of ***A*** into columns; that is, the first column of $\boldsymbol{A^t}$ is the first row of ***A***, the second column of $\boldsymbol{A^t}$ is the second row of ***A***, and so forth.

Example 12 (Transpose of a Matrix)
The following is an example of a matrix and its transpose.

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

$$A^t = \begin{bmatrix} 2 & 1 & 5 \\ 3 & 4 & 6 \end{bmatrix}$$

Example 13 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$

Then $A^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^t = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^t = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$

Theorem 3 Let ***A*** and ***B*** denote matrices whose sizes are appropriate for the following sums and products.

$a.\ \ (A^t)^t = A$

$b.\ \ (A+B)^t = A^t + B^t$

$c.\ \ For\ any\ scalar\ r, (rA)^t = rA^t$

$d.\ \ (AB)^t = B^t A^t$

The generalization of (d) to products of more than two factors can be stated in words as follows.

> "The transpose of a product of matrices equals the product of their transposes in the reverse order."

Lecture 12

The Inverse of a Matrix

In this lecture and the next, we will consider only square matrices and we will investigate the matrix analogue of the reciprocal or multiplicative inverse of a nonzero real number.

Inverse of a square Matrix

If ***A*** is an $n \times n$ matrix, A matrix C of order $n \times n$ is called multiplicative inverse of A if

$AC = CA = I$ where I is the $n \times n$ identity matrix.

Invertible Matrix

If the inverse of a square matrix exists, it is called an invertible matrix.

In this case, we say that ***A*** is **invertible** and we call ***C*** an **inverse** of ***A***.

Note: If ***B*** is another inverse of ***A***, then we would have

$$B = BI = B(AC) = (BA)C = IC = C.$$

Thus when ***A*** is **invertible**, its inverse is unique.

The inverse of ***A*** is denoted by $\boldsymbol{A^{-1}}$, so that

$$AA^{-1} = I \qquad and \qquad A^{-1}A = I$$

Note: A matrix that is not invertible is sometimes called a **singular** matrix, and an invertible matrix is called a **non-singular** matrix.

Warning

By No means $A^{-1}.A = A.A^{-1} = I \quad The\ identity\ matrix\ \ A^{-1} = \dfrac{1}{A}$, as in case of real number, we have $3^{-1} = \dfrac{1}{3}$.

A^{-1} is, in fact the $n \times n$ matrix corresponding to the $n \times n$ matrix A, which satisfies the property

$$A^{-1}.A = A.A^{-1} = I \quad The\ identity\ matrix$$

Example 1 If $A=\begin{bmatrix}2 & 5\\ -3 & -7\end{bmatrix}$ *and* $C=\begin{bmatrix}-7 & -5\\ 3 & 2\end{bmatrix}$, then

$$AC=\begin{bmatrix}2 & 5\\ -3 & -7\end{bmatrix}\begin{bmatrix}-7 & -5\\ 3 & 2\end{bmatrix}=\begin{bmatrix}-14+15 & -10+10\\ 21-21 & 15-14\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \quad \textit{and}$$

$$CA=\begin{bmatrix}-7 & -5\\ 3 & 2\end{bmatrix}\begin{bmatrix}2 & 5\\ -3 & -7\end{bmatrix}=\begin{bmatrix}-14+15 & -35+35\\ 6-6 & 15-14\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

Thus **C =** $\boldsymbol{A^{-1}}$.

Theorem Let $A=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$.

If $ad-bc\neq 0$, then ***A*** is invertible or non singular and $A^{-1}=\dfrac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$

If $ad-bc=0$, then ***A*** is not invertible or singular.

The quantity $ad-bc$ is called the **determinant of *A***, and we write

$$\det A=ad-bc$$

This implies that a 2×2 matrix ***A*** is invertible if and only if $\det A\neq 0$.

Example 2 Find the inverse of $A=\begin{bmatrix}3 & 4\\ 5 & 6\end{bmatrix}$.

Solution We have det ***A*** = 3(6) – 4(5) = $-2\neq 0$.

Hence ***A*** is invertible $A^{-1}=\dfrac{1}{-2}\begin{bmatrix}6 & -4\\ -5 & 3\end{bmatrix}=\begin{bmatrix}6/(-2) & -4/(-2)\\ -5/(-2) & 3/(-2)\end{bmatrix}=\begin{bmatrix}-3 & 2\\ 5/2 & -3/2\end{bmatrix}$

The next theorem provides three useful facts about invertible matrices.

Theorem

a. If ***A*** is an invertible matrix, then $\boldsymbol{A^{-1}}$ is invertible and $(A^{-1})^{-1}=A$

b. If ***A*** and ***B*** are $n\times n$ invertible matrices, then so is ***AB***, and the inverse of ***AB*** is the product of the inverses of ***A*** and ***B*** in the reverse order. That is $(AB)^{-1}=B^{-1}A^{-1}$

c. If ***A*** is an invertible matrix, then so is $\boldsymbol{A^T}$, and the inverse of $\boldsymbol{A^T}$ is the transpose of $\boldsymbol{A^{-1}}$. That is $(A^T)^{-1}=(A^{-1})^T$

Proof

(a) We must find a matrix ***C*** such that $A^{-1}C = I \quad and \quad CA^{-1} = I$

However, we already know that these equations are satisfied with ***A*** in place of ***C***. Hence $\boldsymbol{A^{-1}}$ is invertible and ***A*** is its inverse.

(b) We use the associative law for multiplication:

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AIA^{-1} \\ &= AA^{-1} \\ &= I\end{aligned}$$

A similar calculation shows that $(B^{-1}A^{-1})(AB) = I$.

Hence ***AB*** is invertible, and its inverse is $B^{-1}A^{-1}$ i.e $(AB)^{-1} = B^{-1}A^{-1}$

Generalization

Similarly we can prove the same results for more than two matrices i.e

$$\left((A_1)(A_2)(A_3)...(A_n)\right)^{-1} = A_n^{-1}A_{n-1}^{-1}...\ A_3^{-1}A_2^{-1}A_1^{-1}$$

The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.

<u>**Example 3**</u> **(Inverse of a Transpose).** Consider a general 2×2 invertible matrix and its transpose:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad and \quad A^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Since A is invertible, its determinant *(ad – bc)* is nonzero. But the determinant of $\boldsymbol{A^t}$ is also *(ad – bc)*, so $\boldsymbol{A^t}$ is also invertible. It follows that

$$(A^t)^{-1} = \begin{bmatrix} \dfrac{d}{ad-bc} & -\dfrac{c}{ad-bc} \\ -\dfrac{b}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix} \text{---------------(1)}$$

Now $$A^{-1} = \begin{bmatrix} \dfrac{d}{ad-bc} & -\dfrac{b}{ad-bc} \\ -\dfrac{c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix}$$

Therefore, $(A^{-1})^t = \begin{bmatrix} \dfrac{d}{ad-bc} & -\dfrac{c}{ad-bc} \\ -\dfrac{b}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix}$ -----------------(2)

From (1) and (2), we have

$$\boldsymbol{(A^t)^{-1} = (A^{-1})^t}\ .$$

<u>Example 4</u> (The Inverse of a Product). Consider the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

$$AB = \begin{bmatrix} 7 & 6 \\ 9 & 8 \end{bmatrix},$$

Here

$$(AB)^{-1} = \frac{1}{|AB|} Adj(AB) = \frac{1}{-2}\begin{bmatrix} 8 & -6 \\ -9 & 7 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|} Adj(A) = \frac{1}{1}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix},$$

$$B^{-1} = \frac{1}{|B|} Adj(B) = \frac{1}{2}\begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix},$$

$$B^{-1}A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

Thus, $(AB)^{-1} = B^{-1}A^{-1}$

Theorem: If A is invertible and n is a non-negative integer, then:

(a) $\boldsymbol{A^n}$ is invertible and $\boldsymbol{(A^n)^{-1} = A^{-n} = (A^{-1})^n}$

(b) $k\boldsymbol{A}$ is invertible for any nonzero scalar k, and $(k\boldsymbol{A})^{-1} = k^{-1}\boldsymbol{A}^{-1}$.

Example 5 (Related to the above theorem)

(a) Let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad then\ A^{-1} = \frac{1}{|A|} Adj(A) = \frac{1}{1}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$$

Now $$A^{-3} = \left(A^{-1}\right)^3 = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}$$

Also, $$A^3 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 30 \\ 15 & 41 \end{bmatrix}$$

$$(A^3)^{-1} = \frac{1}{(11)(41)-(30)(15)}\begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} = (A^{-1})^3$$

(b)

Take $A=\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$ and k=3

kA=3A=$\begin{bmatrix} 3 & 6 \\ 9 & 3 \end{bmatrix}$, $(kA)^{-1} = (3A)^{-1} = \dfrac{1}{9-54}\begin{bmatrix} 3 & -6 \\ -9 & 3 \end{bmatrix} = \begin{bmatrix} -1/15 & 2/15 \\ 1/5 & -1/15 \end{bmatrix}$------(1)

$A^{-1} = -\dfrac{1}{5}\begin{bmatrix} 1 & -2 \\ -3 & 1 \end{bmatrix}$

A= *So* $k^{-1}A^{-1}=3^{-1}A^{-1}=\dfrac{1}{3}.\dfrac{-1}{5}\begin{bmatrix} 1 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -1/15 & 2/15 \\ 1/5 & -1/15 \end{bmatrix}$-----------------------(2)

From (1) and (2), we have
$(3A)^{-1} = 3^{-1}A^{-1}$

There is an important connection between invertible matrices and row operations that leads to a method for computing inverses. As we shall see, an invertible matrix $\boldsymbol{A}$ is row equivalent to an identity matrix, and we can find $\boldsymbol{A^{-1}}$ by watching the row reduction of $\boldsymbol{A}$ to $\boldsymbol{I}$.

<u>Elementary Matrices</u>

As we have studied that there are three types of elementary row operations that can be performed on a matrix:
There are three types of elementary operations

- Interchanging of any two rows
- Multiplication to a row by a nonzero constant
- Adding a multiple of one row to another

<u>Elementary matrix</u>

An elementary matrix is a matrix that results from applying a single elementary row operation to an identity matrix.

Some examples are given below:

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

Multiply the second row of I_2 by -3.	Interchange the second and fourth rows of I_4.	Add 3 times the third row of I_3 to the first row.	Multiply the first row of I_3 by 1.

From Def it is clear that elementary matrices are always square.

Elementary matrices are important because they can be used to execute elementary row operations by matrix multiplication.

Theorem If ***A*** is an $n \times n$ identity matrix, and if the elementary matrix ***E*** results by performing a certain row operation on the identity matrix, then the product *EA* is the matrix that results when the same row operation is performed on ***A***.

In short, this theorem states that an elementary row operation can be performed on a matrix ***A*** using a left multiplication by and appropriate elementary matrix.

Example 6 **(Performing Row Operations by Matrix Multiplication)**. Consider the matrix

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

Find an elementary matrix ***E*** such that ***EA*** is the matrix that results by adding 4 times the first row of ***A*** to the third row.

Solution: The matrix ***E*** must be 3×3 to conform to the product ***EA***. Thus, we obtain ***E*** by adding 4 times the first row of I_3 to the third row. This gives us $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$

As a check, the product ***EA*** is $EA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 5 & 4 & 12 & 12 \end{bmatrix}$

So left multiplication by ***E*** does, in fact, add 4 times the first row of ***A*** to the third row.

If an elementary row operation is applied to an identity matrix ***I*** to produce an elementary matrix ***E***, then there is a second row operation that, when applied to ***E***, produces ***I*** back again.

For example, if ***E*** is obtained by multiplying the *i-th* row of ***I*** by a nonzero scalar c, then ***I*** can be recovered by multiplying the *i-th* row of ***E*** by 1/c. The following table explains how to recover the identity matrix from an elementary matrix for each of the three elementary row operations. The operations on the right side of this table are called the **inverse operations** of the corresponding operations on the left side.

Row operation on I that produces E	**Row operation on E that reproduces I**
Multiply row *i* by $c \neq 0$	Multiply row *i* by 1/c
Interchange rows *i* and *j*	Interchange rows *i* and *j*
Add c times row *i* to row *j*	Add –c times row *i* to row *j*

<u>Example 7</u> (Recovering Identity Matrices from Elementary Matrices). Here are three examples that use inverses of row operations to recover the identity matrix from

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \xrightarrow[\text{row by 7.}]{\text{Multiply the second}} \begin{bmatrix}1 & 0\\0 & 7\end{bmatrix} \xrightarrow[\text{row by 1/7.}]{\text{Multiply the second}} \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \xrightarrow[\text{second rows.}]{\text{Interchange the first and}} \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} \xrightarrow[\text{second rows.}]{\text{Interchange the first and}} \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \xrightarrow[\text{to the first.}]{\text{Add 5 times the second row}} \begin{bmatrix}1 & 5\\0 & 1\end{bmatrix} \xrightarrow[\text{to the first.}]{\text{Add -5 times the second row}} \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

The next theorem is the basic result on the invertibility of elementary matrices.

<u>Theorem</u> An elementary matrix is invertible and the inverse is also an elementary matrix.

<u>Example 8</u> Let $E_1 = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\-4 & 0 & 1\end{bmatrix}$, $E_2 = \begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix}$, $E_3 = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 5\end{bmatrix}$

$$A = \begin{bmatrix}a & b & c\\d & e & f\\g & h & i\end{bmatrix}$$

Compute $\boldsymbol{E_1A}$, $\boldsymbol{E_2A}$, $\boldsymbol{E_3A}$ and describe how these products can be obtained by elementary row operations on $\boldsymbol{A}$.

Solution We have

$$E_1A = \begin{bmatrix}a & b & c\\d & e & f\\g-4a & h-4b & i-4c\end{bmatrix}, \quad E_2A = \begin{bmatrix}d & e & f\\a & b & c\\g & h & i\end{bmatrix}, \; E_3A = \begin{bmatrix}a & b & c\\d & e & f\\5g & 5h & 5i\end{bmatrix}$$

Addition of (–4) times row 1 of $\boldsymbol{A}$ to row 3 produces $\boldsymbol{E_1 A}$. (This is a row replacement operation.) An interchange of rows 1 and 2 of $\boldsymbol{A}$ produces $\boldsymbol{E_2 A}$ and multiplication of row 3 of $\boldsymbol{A}$ by 5 produces $\boldsymbol{E_3 A}$.

Left-multiplication (that is, multiplication on the left) by $\boldsymbol{E_1}$ in Example 8 has the same effect on any $3\times n$ matrix. It adds – 4 times row 1 to row 3. In particular, since $\boldsymbol{E_1 I = E_1}$, we see that $\boldsymbol{E_1}$ itself is produced by the same row operation on the identity. Thus Example 8 illustrates the following general fact about elementary matrices.

Note: Since row operations are reversible, elementary matrices are invertible, for if $\boldsymbol{E}$ is produced by a row operation on $\boldsymbol{I}$, then there is another row operation of the same type that changes $\boldsymbol{E}$ back into $\boldsymbol{I}$. Hence there is an elementary matrix $\boldsymbol{F}$ such that $\boldsymbol{FE = I}$. Since $\boldsymbol{E}$ and $\boldsymbol{F}$ correspond to reverse operations, $\boldsymbol{EF = I}$.

Each elementary matrix $\boldsymbol{E}$ is invertible. The inverse of $\boldsymbol{E}$ is the elementary matrix of the same type that transforms $\boldsymbol{E}$ back into $\boldsymbol{I}$.

Example Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

Solution: To transform $\boldsymbol{E_1}$ into $\boldsymbol{I}$, add + 4 times row 1 to row 3.

The elementary matrix which does that is $E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$

Theorem An $n\times n$ matrix $\boldsymbol{A}$ is invertible if and only if $\boldsymbol{A}$ is row equivalent to $\boldsymbol{I_n}$, and in this case, any sequence of elementary row operations that reduces $\boldsymbol{A}$ to $\boldsymbol{I_n}$ also transforms $\boldsymbol{I_n}$ into $\boldsymbol{A^{-1}}$.

An Algorithm for Finding $\boldsymbol{A^{-1}}$ If we place $\boldsymbol{A}$ and $\boldsymbol{I}$ side-by-side to form an augmented matrix $[\boldsymbol{A} \quad \boldsymbol{I}]$, then row operations on this matrix produce identical operations on $\boldsymbol{A}$ and $\boldsymbol{I}$. Then either there are row operations that transform $\boldsymbol{A}$ to $\boldsymbol{I_n}$, and $\boldsymbol{I_n}$ to $\boldsymbol{A^{-1}}$, or else $\boldsymbol{A}$ is not invertible.

Algorithm for Finding $\boldsymbol{A^{-1}}$
Row reduce the augmented matrix $[\boldsymbol{A} \quad \boldsymbol{I}]$. If $\boldsymbol{A}$ is row equivalent to $\boldsymbol{I}$, then $[\boldsymbol{A} \quad \boldsymbol{I}]$ is row equivalent to $[\boldsymbol{I} \quad \boldsymbol{A^{-1}}]$. Otherwise, $\boldsymbol{A}$ does not have an inverse.

Example 9 Find the inverse of the matrix $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, if it exists.

Solution $[A \quad I] = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} R_{12}$

$$\begin{array}{cc} -4R_1 + R_3 & 3R_2 + R_3 \\ \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix} & \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{bmatrix} \end{array}$$

$$\begin{array}{cc} -2R_3 + R_2 & -3R_3 + R_1 \\ \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix} & \sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix} \end{array}$$

Since ***A*** ~ **I**, we conclude that ***A*** is invertible, and $A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$

It is a good idea to check the final answer:

$$AA^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It is not necessary to check that $\boldsymbol{A^{-1}A = I}$ since ***A*** is invertible.

Example 10 Find the inverse of the matrix $A = \begin{bmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 5 \\ 3 & 1 & -2 & 5 \end{bmatrix}$, if it exists.

Consider $\quad detA = \begin{vmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 5 \\ 3 & 1 & -2 & 5 \end{vmatrix} \quad = \begin{vmatrix} 1 & 2 & -3 & 1 \\ 0 & 5 & -6 & -1 \\ 0 & -4 & 7 & 3 \\ 0 & -5 & 7 & 2 \end{vmatrix}$

operating $R_2 + R_1, R_3 - 2R_1, R_4 - 3R_1$

Expand from first column $= \begin{vmatrix} 5 & -6 & -1 \\ -4 & 7 & 3 \\ -5 & 7 & 2 \end{vmatrix} = \begin{vmatrix} 5 & -6 & -1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{vmatrix} = 5(1-2) + 6(1-0) - 1(1-0) = 0$

As the given matrix is singular, so it is not invertible.

<u>Example 11</u> Find the inverse of the given matrix if possible $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix}$

Solution $\quad \det A = \begin{vmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{vmatrix} = -1$

As the given matrix is non-singular therefore, inverse is possible.

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

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

$R_2 - 2R_1, R_3 - 3R_1$

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

$R_3 - R_2$

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

$Multiply R_3\ by -1$

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

$R_1 - R_3, R_2 + R_3$

Hence the inverse of matrix A is $A^{-1} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & 1 & -1 \end{bmatrix}$

<u>Example 12</u> Find the inverse of the matrix $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 3 \\ 5 & 2 & 3 \end{bmatrix}$

Solution $\det A = \begin{vmatrix} 1 & 2 & 2 \\ 2 & 2 & 3 \\ 5 & 2 & 3 \end{vmatrix} = 6$

As the given matrix is non-singular, therefore, inverse of the matrix is possible.
We reduce it to reduce echelon form.

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

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & -2 & -1 \\ 0 & -8 & -7 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -5 & 0 & 1 \end{bmatrix}$$

$R_2 - 2R_1, R_3 - 5R_1$

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 1/2 \\ 0 & -8 & -7 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1/2 & 0 \\ -5 & 0 & 1 \end{bmatrix}$$

multiply 2nd row by -1/2

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 1/2 \\ 0 & 0 & -3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1/2 & 0 \\ 3 & -4 & 1 \end{bmatrix}$$

$R_3 + 8R_2$

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1/2 & 0 \\ -1 & 4/3 & -1/3 \end{bmatrix}$$

Multiply 3rd row by -1/3

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 3 & -\frac{8}{3} & \frac{2}{3} \\ 3/2 & -7/6 & 1/6 \\ -1 & 4/3 & -1/3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & -1/3 & 1/3 \\ 3/2 & -7/6 & 1/6 \\ -1 & 4/3 & -1/3 \end{bmatrix}$$

$R_2 - (1/2)R_3, R_1 - 2R_3$

Hence the inverse of the original matrix $A^{-1} = \begin{bmatrix} 0 & -1/3 & 1/3 \\ 3/2 & -7/6 & 1/6 \\ -1 & 4/3 & -1/3 \end{bmatrix}$

Exercises

In exercises 1 to 4, find the inverses of the matrices, if they exist. Use elementary row operations.

1. $\begin{bmatrix} 1 & 2 \\ 5 & 9 \end{bmatrix}$

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

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

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

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

6. $\begin{bmatrix} -1 & 3 & -4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{bmatrix}$

7. Let $A = \begin{bmatrix} -1 & -5 & -7 \\ 2 & 5 & 6 \\ 1 & 3 & 4 \end{bmatrix}$. Find the third column of A^{-1} without computing the other columns.

8. Let $A = \begin{bmatrix} -25 & -9 & -27 \\ 546 & 180 & 537 \\ 154 & 50 & 149 \end{bmatrix}$. Find the second and third columns of A^{-1} without computing the first column.

9. Find an elementary matrix E that satisfies the equation.
(a) EA = B (b) EB = A (c) EA = C (d) EC = A
where $A = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{bmatrix}, B = \begin{bmatrix} 8 & 1 & 5 \\ 2 & -7 & -1 \\ 3 & 4 & 1 \end{bmatrix}, C = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 2 & -7 & 3 \end{bmatrix}$.

10. Consider the matrix $A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(a) Find elementary matrices E_1 and E_2 such that $E_2E_1A=I$.
(b) Write A^{-1} as a product of two elementary matrices.

(c) Write A as a product of two elementary matrices.

In exercises 11 and 12, express A and A^{-1} as products of elementary matrices.

11. $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

13. Factor the matrix $A = \begin{bmatrix} 0 & 1 & 7 & 8 \\ 1 & 3 & 3 & 8 \\ -2 & -5 & 1 & -8 \end{bmatrix}$ as A = EFGR, where E, F, and G are elementary matrices and R is in row echelon form.

Lecture 13

Characterizations of Invertible Matrices

This chapter involves a few techniques of solving the system of n linear equations in **n** unknowns and transformation associated with a matrix.

Solving Linear Systems by Matrix Inversion

Theorem

Let $\boldsymbol{A}$ be an $n \times n$ invertible matrix. For any $\boldsymbol{b} \in \boldsymbol{R}^n$, the equation $\boldsymbol{Ax = b}$ …..(1) has the unique solution i.e. $\boldsymbol{x = A^{-1}b}$.

Proof

Since A is invertible and $\boldsymbol{b} \in \boldsymbol{R}^n$ be any vector. Then, we must have a matrix $A^{-1}b$ which is a solution of eq. (1) .i.e. $\boldsymbol{Ax = A\ (A^{-1}b) = Ib = b}$.

Uniqueness

For uniqueness, we assume that there is another solution $\boldsymbol{u}$. Indeed, it is a solution of eq.(1) so it must be $\boldsymbol{u = A^{-1}b}$, it means x = $\boldsymbol{A^{-1}b}$ = $\boldsymbol{u}$. This shows that $\boldsymbol{u = x}$.

Theorem

Let A and B be the square matrices such that AB = I. Then, A and B are invertible with B = A^{-1} and A = B^{-1}

Example 1

Solve the system of linear equations

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 1 \\ 5x_1 + x_2 + 3x_3 &= 3 \\ x_1 + \qquad 4x_3 &= 6 \end{aligned}$$

Solution

Consider the linear system

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 1 \\ 5x_1 + x_2 + 3x_3 &= 3 \\ x_1 + \qquad 4x_3 &= 6 \end{aligned}$$

The Matrix form of system is $\boldsymbol{Ax = b}$, where

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 5 & 1 & 3 \\ 1 & 0 & 4 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$$

Here, det (***A***) = 2(4) – 1 (20 – 3) + 1(0 – 1) = 8 – 17 – 1 = -10 $\neq$ 0

So, ***A*** is invertible. Now, we apply the inversion algorithm:

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 5 & 1 & 3 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 5 & 1 & 3 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & -17 & 0 & 1 & -5 \\ 0 & 1 & -7 & 1 & 0 & -2 \end{array}\right], \quad -5R_1 + R_2, \quad -2R_1 + R_3$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & -17 & 0 & 1 & -5 \\ 0 & 0 & 10 & 1 & -1 & 3 \end{array}\right], \quad -R_2 + R_3,$$

$$\left[\begin{array}{cccccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & -17 & 0 & 1 & -5 \\ 0 & 0 & 1 & \frac{1}{10} & \frac{-1}{10} & \frac{3}{10} \end{array}\right], \quad \frac{R_3}{10}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-2}{5} & \frac{2}{5} & \frac{-1}{5} \\ 0 & 1 & 0 & \frac{17}{10} & \frac{-7}{10} & \frac{1}{10} \\ 0 & 0 & 1 & \frac{1}{10} & \frac{-1}{10} & \frac{3}{10} \end{array}\right], \quad -4R_3 + R_1 \quad , -17R_3 + R_2$$

Hence, $A^{-1} = \begin{bmatrix} \frac{-2}{5} & \frac{2}{5} & \frac{-1}{5} \\ \frac{17}{10} & \frac{-7}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{-1}{10} & \frac{3}{10} \end{bmatrix}$

Thus, the solution of the linear system is $x = A^{-1}b = A^{-1} = \begin{bmatrix} \frac{-2}{5} & \frac{2}{5} & \frac{-1}{5} \\ \frac{17}{10} & \frac{-7}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{-1}{10} & \frac{3}{10} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{-2}{5} \\ \frac{1}{5} \\ \frac{8}{5} \end{bmatrix}$

Example # 2 Solve the system $\begin{cases} x_1 + 2x_2 + 3x_3 = 5 \\ 2x_1 + 5x_2 + 3x_3 = 3 \\ x_1 + \quad\quad 8x_3 = 17 \end{cases}$ by inversion method.

Solution Consider the linear system $\begin{cases} x_1 + 2x_2 + 3x_3 = 5 \\ 2x_1 + 5x_2 + 3x_3 = 3 \\ x_1 + \quad\quad 8x_3 = 17 \end{cases}$

This system can be written in matrix form as ***Ax = b***, where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 5 \\ 3 \\ 17 \end{bmatrix}$$

Here, [det (***A***)] = 40 – 2 (16 – 3) + 3(0 – 5) = 40 – 26 – 15 = -1 ≠ 0
Therefore, ***A*** is invertible.

Now, we apply the inversion algorithm:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right] -2R_1 + R_2, \quad -1R_1 + R_3$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right] 2R_2 + R_3$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] -1R_3$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] 3R_3 + R_2, \quad -3R_3 + R_1$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] -2R_2 + R_1$$

Hence, $A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$

Thus, the solution of the linear system is $x = A^{-1}b = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \\ 17 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

Thus $x_1 = 1, x_2 = -1, x_3 = 2$.

Note: It is only applicable when the number of equations = number of unknown and fails if given matrix is not invertible.

Example 3

Solve the system of linear equation

$$x_1 + 6x_2 + 4x_3 = 2$$
$$2x_1 + 4x_2 - x_3 = 3$$
$$-x_1 + 2x_2 + 5x_3 = 3$$

Solution

The matrix coefficient

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

det (***A***) = 1(20 + 2) – 6(10 – 1) + 4 (4+4)

$= 22 - 54 + 32$

$= 0$

Thus, A is not invertible. Hence, the inversion method fails.

Solution for the system of a $n \times n$ Homogeneous linear equations with an Invertible Coefficient Matrix:-

Let see if the system is considered as homogeneous then what does the above theorem say?

Theorem:-

Let ***Ax = 0*** be a homogeneous linear system of **n** equations and **n** unknowns. Then, the coefficient matrix ***A*** is invertible iff this system has only a trivial solution.

Example 4

State whether the following system of linear equation has a solution or not?

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

Solution

We see

$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -3 & 1 \\ 1 & 0 & -4 \end{bmatrix}$ is an invertible matrix (det (A) $\neq 0$)

Thus, this homogeneous linear system has only the trivial solution.

Example 5

Solve $\begin{cases} x_1 + x_2 + x_3 = 8 \\ \quad 2x_2 + 3x_3 = 24 \\ 5x_1 + 5x_2 + x_3 = 8 \end{cases}$

Solution This system can be written in matrix form as ***Ax = b***, where

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 8 \\ 24 \\ 8 \end{bmatrix}$$

Here [det (***A***)] = 1(2 – 15) – 1(0 – 15) + 1(0 – 10) = -13 + 15 – 10 = -8 $\neq 0$
Therefore, ***A*** is invertible.

Now, we apply the inversion algorithm:

$$\begin{array}{c} \begin{array}{cc} A & I_3 \end{array} \\ \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 5 & 5 & 1 & 0 & 0 & 1 \end{array}\right] \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array}\right] \quad -5R_1 + R_3$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array}\right] \quad \frac{1}{2}R_2$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array}\right] \quad -1R_2 + R_1$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{5}{4} & 0 & -\frac{1}{4} \end{array}\right] \quad -\frac{1}{4}R_3$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{15}{8} & \frac{1}{2} & \frac{3}{8} \\ 0 & 0 & 1 & \frac{5}{4} & 0 & -\frac{1}{4} \end{array}\right] \quad -\frac{3}{2}R_3 + R_2$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{13}{8} & -\frac{1}{2} & -\frac{1}{8} \\ 0 & 1 & 0 & -\frac{15}{8} & \frac{1}{2} & \frac{3}{8} \\ 0 & 0 & 1 & \frac{5}{4} & 0 & -\frac{1}{4} \end{array}\right] \quad \frac{1}{2}R_3 + R_1$$

Hence, $$A^{-1} = \begin{bmatrix} \frac{13}{8} & -\frac{1}{2} & -\frac{1}{8} \\ -\frac{15}{8} & \frac{1}{2} & \frac{3}{8} \\ \frac{5}{4} & 0 & -\frac{1}{4} \end{bmatrix}$$

Thus, the solution of the linear system is

$$x = A^{-1}b = \begin{bmatrix} \frac{13}{8} & -\frac{1}{2} & -\frac{1}{8} \\ -\frac{15}{8} & \frac{1}{2} & \frac{3}{8} \\ \frac{5}{4} & 0 & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} 8 \\ 24 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}$$

Thus $x_1 = 0, x_2 = 0, x_3 = 8$.

<u>Theorem (Invertible Matrix Theorem)</u> Let ***A*** be a square $n \times n$ matrix. Then the following statements are equivalent. (Means if any one holds then all are true).

(a) ***A*** is an invertible matrix.

(b) ***A*** is row equivalent to the $n \times n$ identity matrix.

(c) ***A*** has **n** pivot positions.

(d) The equation ***Ax = 0*** has only the trivial solution.

(e) The columns of ***A*** form a linearly independent set.

(f) The linear transformation $x \to Ax$ is one-to-one.

(g) The equation ***Ax = b*** has at least one solution for each ***b*** in $\boldsymbol{R^n}$.

(h) The columns of ***A*** span $\boldsymbol{R^n}$.

(i) The linear transformation $x \to Ax$ maps $\boldsymbol{R^n}$ onto $\boldsymbol{R^n}$.

(j) There is a $n \times n$ matrix ***C*** such that ***CA = I***.

(k) There is a $n \times n$ matrix ***D*** such that ***AD = I.***

(l) $\boldsymbol{A^T}$ is an invertible matrix.

<u>Example 6</u>

Show that the matrix $A = \begin{bmatrix} 1 & 0 & -4 \\ 1 & 1 & 5 \\ 0 & 1 & 2 \end{bmatrix}$ is invertible by using Invertible Matrix Theorem

Solution

By row equivalent,

$$A=\begin{bmatrix}1 & 0 & -4\\ 1 & 1 & 5\\ 0 & 1 & 2\end{bmatrix}, -R_1+R_2$$

$$\begin{bmatrix}1 & 0 & -4\\ 0 & 1 & 9\\ 0 & 1 & 2\end{bmatrix}, -R_2+R_3$$

$$\begin{bmatrix}1 & 0 & -4\\ 0 & 1 & 9\\ 0 & 0 & -7\end{bmatrix}$$

It shows that A has three pivot positions and hence is invertible, by the Invertible Matrix Theorem (c).

Example 7

Use the Invertible Matrix Theorem to decide if ***A*** is invertible, where

$$A=\begin{bmatrix}1 & 0 & -2\\ 3 & 1 & -2\\ -5 & -1 & 9\end{bmatrix}$$

Solution

$$A=\begin{bmatrix}1 & 0 & -2\\ 0 & 1 & 4\\ 0 & -1 & -1\end{bmatrix}\sim\begin{bmatrix}1 & 0 & -2\\ 0 & 1 & 4\\ 0 & 0 & 3\end{bmatrix}$$ Here, ***A*** has three pivot positions and hence is invertible by the Invertible Matrix Theorem (c).

Example 7 Find A^t and show that A^t is an invertible matrix.

$$A=\begin{bmatrix}1 & 2 & 1 & 0\\ 2 & 4 & 1 & 1\\ 1 & 0 & 1 & 1\\ 0 & 1 & 1 & 1\end{bmatrix}$$

Solution

$$A^t=\begin{bmatrix}1 & 2 & 1 & 0\\ 2 & 4 & 0 & 1\\ 1 & 1 & 1 & 1\\ 0 & 1 & 1 & 1\end{bmatrix}$$

Now, by row equivalent of A,

$$A=\begin{bmatrix}1&2&1&0\\2&4&1&1\\1&0&1&1\\0&1&1&1\end{bmatrix}\overset{R_2-2R_1,\,R_3-R_1}{\sim}\begin{bmatrix}1&2&1&0\\0&0&-1&1\\0&-2&0&1\\0&1&1&1\end{bmatrix}\overset{R_{23}}{\sim}\begin{bmatrix}1&2&1&0\\0&-2&0&1\\0&0&-1&1\\0&1&1&1\end{bmatrix}$$

$$\overset{-\frac{1}{2}R_2}{\sim}\begin{bmatrix}1&2&1&0\\0&1&0&-\frac{1}{2}\\0&0&-1&1\\0&1&1&1\end{bmatrix}\overset{R_4-R_2}{\sim}\begin{bmatrix}1&2&1&0\\0&1&0&-\frac{1}{2}\\0&0&-1&1\\0&0&1&\frac{3}{2}\end{bmatrix}\overset{(-1)R_3}{\sim}\begin{bmatrix}1&2&1&0\\0&1&0&\frac{-1}{2}\\0&0&1&-1\\0&0&1&\frac{3}{2}\end{bmatrix}\overset{-R_3+R_4}{\sim}\begin{bmatrix}1&2&1&0\\0&1&0&\frac{-1}{2}\\0&0&1&-1\\0&0&0&\frac{5}{2}\end{bmatrix}$$

Here ***A*** has 4 pivot positions so by Invertible Matrix Theorem (c) A is invertible. Thus, by (l) A^t is invertible.

Example 8

Use the Invertible Matrix Theorem to decide if ***A*** is invertible, where

$$A=\begin{bmatrix}1&2&1&0\\2&4&1&2\\1&0&1&2\\0&1&1&2\end{bmatrix}$$

Solution

$$A=\begin{bmatrix}1&2&1&0\\2&4&1&2\\1&0&1&2\\0&1&1&2\end{bmatrix}\overset{R_2-2R_1,\,R_3-R_1}{\sim}\begin{bmatrix}1&2&1&0\\0&0&-1&2\\0&-2&0&2\\0&1&1&2\end{bmatrix}\overset{R_{23}}{\sim}\begin{bmatrix}1&2&1&0\\0&-2&0&2\\0&0&-1&2\\0&1&1&2\end{bmatrix}$$

$$\begin{array}{cccc} -\frac{1}{2}R_2 & R_4 - R_2 & (-1)R_3 & -R_3 + R_4 \end{array}$$

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

Here, ***A*** has 4 pivot positions and hence is invertible by Invertible Matrix Theorem (c).

Solving Multiple Linear Systems with a Common Coefficient Matrix

This technique is used in solving a sequence of linear systems

$$Ax_1 = b_1,\ Ax_2 = b_2,\ \cdots\ ,\ Ax_k = b_k \qquad (1)$$

where coefficient matrix ***A*** *remains* same and off course if it is invertible, then we have a sequence of solutions. i.e.

$x_1 = A^{-1}b_1,\ x_2 = A^{-1}b_2,\ \cdots\ ,\ x_k = A^{-1}b_k$.

Find a Matrix from Linear Transformation

We can find a matrix corresponding to every transformation. In this section we will learn how to find a matrix attached with a linear transformation.

Example 9

Let L be the linear transformation from R^2 to P_2 (Polynomials of order 2) defined by

T(x, y) = x y t + (x + y)t^2

Find the matrix representing T with respect to the standard bases.

Solution

Let A = {(1,0),(0,1)} be the basis of R^2, then

T(1,0) = t^2 = (0,0,1) (This triple represents the coefficients of polynomial t^2)

i.e. $t^2 = 0.1 + 0.t + 1.t^2$

Similarly, T (0, 1) = t^2 = (0,0,1). Hence, the matrix is given by

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}$$

Now, we will proceed with a more complicated example.

Example 10

Let T be the linear transformation from R^2 to R^2 such that T(x, y) = (x , y + 2x). Find a matrix A for T.

Solution

This matrix is found by finding T (1, 0) = (1, 2) and T (0, 1) = (0, 1)The matrix is $A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$.

Important Note

It should be clear that the Invertible Matrix Theorem applies only to square matrices. For example, if the columns of a 4×3 matrix are linearly independent, we cannot use the Invertible Matrix Theorem to conclude any thing about the existence or nonexistence of solutions to equations of the form $\boldsymbol{Ax = b}$.

Invertible Linear Transformations

Recall that matrix multiplication corresponds to composition of linear transformations. When a matrix $\boldsymbol{A}$ is invertible, the equation $\boldsymbol{A^{-1}Ax = x}$ can be viewed as a statement about linear transformations. See Figure 2.

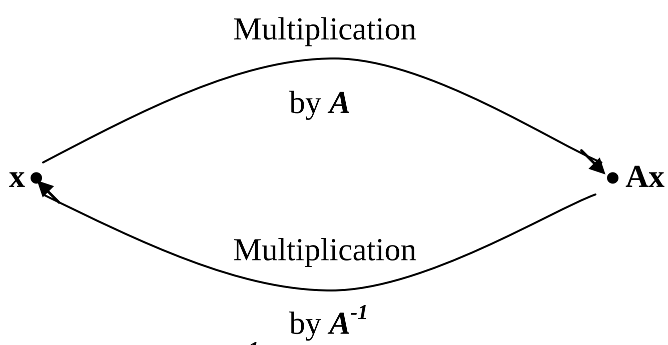


Figure 2 $\boldsymbol{A^{-1}}$ transforms $\boldsymbol{Ax}$ back to $\boldsymbol{x}$

Definition

A linear transformation $T : R^n \longrightarrow R^n$ be a linear Transformation and A be a standard matrix for T. Then, T is invertible if and only if A is an invertible matrix in that case linear transformation S given by $S(x) = A^{-1}x$ is a unique function satisfying (1) and (2)

$$S(T(x)) = x \;\forall\; x \in R^n \qquad (1)$$

$$T(S(x)) = x \;\forall\; x \in R^n \qquad (2)$$

Important Note

If the inverse of a linear transformation exits then it is unique.

Proposition

Let $T : R^n \to R^m$ be linear transformation, given as $T(x) = Ax \quad , \forall\ x \in R^n$, where A is a $m \times n$ matrix. The mapping T is invertible if the system $y = Ax$ has a unique solution.

Case 1:

If $m < n$, then the system $Ax = y$ has either no solution or infinitely many solutions, for any y in $R^{m.}$ Therefore, $y = Ax$ is non-invertible.

Case 2:

If $m = n$, then the system $Ax = y$ has a unique solution if and only if Rank (A) = n.

Case 3:

If $m > n$, then the transformation $y = Ax$ is non-invertible because we can find a vector y in R^m such that $Ax = y$ is inconsistent.

Exercises

1.Solve the system of linear equations by inverse matrix method.

$$\begin{aligned} x_1 + x_2 + 4x_3 &= 2 \\ 2x_2 + 3x_3 &= 4 \\ 5x_1 + x_2 - x_3 &= 3 \end{aligned}$$

2. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix}, b_1 = \begin{bmatrix} 5 \\ 7 \end{bmatrix}, b_2 = \begin{bmatrix} -2 \\ -2 \end{bmatrix}, b_3 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, and\ b_4 = \begin{bmatrix} 1 \\ 7 \end{bmatrix}.$

(a) Find A^{-1} and use it to solve the equations $Ax = b_1$, $Ax = b_2$, $Ax = b_3$, $Ax = b_4$.

(b) Solve the four equations in part (a) by row reducing the augmented matrix

$\begin{bmatrix} A & b_1 & b_2 & b_3 & b_4 \end{bmatrix}.$

3. (a) Solve the two systems of linear equations

$$\begin{aligned} x_1 + 2x_2 + x_3 &= -1 \\ x_1 + 3x_2 + 2x_3 &= 3 \\ x_2 + 2x_3 &= 4 \end{aligned}$$

&

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ x_1 + 3x_2 + 2x_3 &= 0 \\ x_2 + 2x_3 &= 4 \end{aligned}$$

by row reduction.

(b) Write the systems in (a) as $Ax = b_1$ and $Ax = b_2$, and then solve each of them by the method of inversion.

Determine which of the matrices in exercises 4 to 10 are invertible?

4. $\begin{bmatrix} -4 & 16 \\ 3 & -9 \end{bmatrix}$ 5. $\begin{bmatrix} 5 & 0 & 3 \\ 7 & 0 & 2 \\ 9 & 0 & 1 \end{bmatrix}$ 6. $\begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$ 7. $\begin{bmatrix} 5 & -9 & 3 \\ 0 & 3 & 4 \\ 1 & 0 & 3 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 1 & -2 & -1 \\ -2 & -6 & 3 & 2 \\ 3 & 5 & 8 & -3 \end{bmatrix}$ 9. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 10 \end{bmatrix}$ 10. $\begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$

11. $\begin{bmatrix} 5 & 4 & 3 & 6 & 3 \\ 7 & 6 & 5 & 9 & 5 \\ 8 & 6 & 4 & 10 & 4 \\ 9 & -8 & 9 & -5 & 8 \\ 10 & 8 & 7 & -9 & 7 \end{bmatrix}$

12. Suppose that ***A*** and ***B*** are $n \times n$ matrices and the equation ***ABx = 0*** has a nontrivial solution. What can you say about the matrix ***AB***?

13**.** What can we say about a one-to-one linear transformation ***T*** from $\boldsymbol{R^n}$ into $\boldsymbol{R^n}$?

14. Let $T : R^2 \to R^2$ be a linear transformation given as $T(x) = 5x$, then find a matrix A of linear transformation T.

In exercises 15 and 16, T is a linear transformation from R^2 into R^2. Show that T is invertible.

15. $T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2)$

16. $T(x_1, x_2) = (6x_1 - 8x_2, -5x_1 + 7x_2)$

17. Let $T : R^n \to R^n$ be a linear transformation and let ***A*** be the standard matrix for ***T***. Then ***T*** is invertible if and only if ***A*** is an invertible matrix.

Lecture 14

Partitioned Matrices

A **block matrix** or a **partitioned matrix** is a partition of a matrix into rectangular smaller matrices called **blocks**.Partitioned matrices appear often in modern applications of linear algebra because the notation simplifies many discussions and highlights essential structure in matrix calculations. This section provides an opportunity to review matrix algebra and use of the Invertible Matrix Theorem.

General Partitioning of a Matrix -

A matrix can be **partitioned** (subdivided) into **sub matrices** (also called **blocks**) in various ways by inserting lines between selected rows and columns.

Example 1

The matrix

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

can be partitioned into four 2×2 blocks

$$P_{11} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, P_{12} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, P_{21} = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}, P_{22} = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix},$$

The partitioned matrix can then be written as

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}$$

Example 2

$$\text{The matrix } A = \left[\begin{array}{ccc|cc|c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

can also be written as the 2×3 partitioned (or block) matrix

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

whose entries are the blocks or sub matrices.

Note:-

It is important to know that in how many ways to block up an ordinary matrix A. See the following example in which a matrix A is block up into three different ways.

Example 3

Let A be a general matrix of 5×3 order, we have

Partition (a)

$$A = \left[\begin{array}{c|cc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \hline a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

In this case we partitioned the matrix into four sub matrices. Also notice that we simplified the matrix into a more compact form and in this compact form we've mixed and matched some of our notation. The partitioned matrix can be thought of as a smaller matrix with four entries, except this time each of the entries are matrices instead of numbers and so we used capital letters to represent the entries and subscripted each one with the location in the partitioned matrix.

Be careful not to confuse the location subscripts on each of the sub matrices with the size of each sub matrix. In this case A_{11} is a 2×1 sub matrix of *A*, A_{12} is a 2×2 sub matrix of *A*, A_{21} is a 3×1 sub matrix of A and A_{22} is a 3×3 sub matrix of A.

Partition (b)

$$A = \left[\begin{array}{c|c|c} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{array}\right] = \left[c_1 \mid c_2 \mid c_3\right]$$

In this case, we partitioned *A* into three column matrices each representing one column in the original matrix. Again, note that we used the standard column matrix notation (the bold face letters) and subscripted each one with the location in the partitioned matrix. The c_i in the partitioned matrix are sometimes called the **column matrices of *A***.

Partition (c)

$$A = \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ \hline a_{21} & a_{22} & a_{23} \\ \hline a_{31} & a_{32} & a_{33} \\ \hline a_{41} & a_{42} & a_{43} \\ \hline a_{51} & a_{52} & a_{53} \end{array}\right] = \left[\begin{array}{c} r_1 \\ \hline r_2 \\ \hline r_3 \\ \hline r_4 \\ \hline r_5 \end{array}\right]$$

Just as we can partition a matrix into each of its columns as we did in the previous part we can also partition a matrix into each of its rows. The r_i in the partitioned matrix are sometimes called the **row matrices** of ***A*.**

<u>Addition of Blocked Matrices</u>

If matrices ***A*** and ***B*** are the same size and are partitioned in exactly the same way, then it is natural to make the same partition of the ordinary matrix sum ***A* + *B***. In this case, each block of ***A* + *B*** is the (matrix) sum of the corresponding blocks of ***A*** and ***B***.

<u>Theorem</u>

If A is $m\times n$ and B is $n\times p$, then

$$\text{AB} = Col_1(A) + Col_2(A) + ... + Col_n(A) \begin{bmatrix} Row_1(B) \\ Row_2(B) \\ \vdots \\ Row_n(B) \end{bmatrix}$$

$$= Col_1(A)Row_1(B) + Col_2(A)Row_2(B) + ... + Col_n(A)Row_n(B)$$

Proof

For each row index I and column index j, the (i,j) entry in col_k (A) row_k (B) is the product of a_{ik} from col_k (A) and b_{kj} from row_k (B).

Hence, the (i,j)-entry in the sum shown in (2) is

$a_{i1}b_{1j} + a_{21}b_{2j} + ... + a_{n1}b_{nj}$

$(k = 1, 2, ..., n)$

This sum is also the (i,j)- entry in AB by the row column rule

Example 4

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad and \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$, then

$$A + B = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix}$$

Similarly, subtraction of blocked matrices is defined.

Multiplication of a partitioned matrix by a scalar is also computed block by block.

Multiplication of Partitioned Matrices

If $$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \quad and \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

and if the sizes of the blocks confirm for the required operations, then

$$AB=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}=\begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} \\ A_{31}B_{11}+A_{32}B_{21} & A_{31}B_{12}+A_{32}B_{22} \end{bmatrix}$$

It is known as **block multiplication**.

Example 5

Find the block multiplication of the following partitioned matrices:

$$A=\left[\begin{array}{ccc|cc} 1 & 2 & 1 & 0 & 2 \\ -1 & -2 & 3 & 1 & 1 \\ \hline 2 & 1 & -2 & 1 & 3 \end{array}\right]=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B=\left[\begin{array}{cc} 2 & 1 \\ 3 & 0 \\ -5 & 1 \\ \hline 0 & 8 \\ 0 & 2 \end{array}\right]=\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$$

Solution

Let

$$A_{11}=\begin{bmatrix} 1 & 2 & 1 \\ -1 & -2 & 3 \end{bmatrix},\ A_{12}=\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix},\ A_{21}=\begin{bmatrix} 2 & 1 & -2 \end{bmatrix}\ and\ \ A_{22}=\begin{bmatrix} 1 & 3 \end{bmatrix}\vdots$$

$$B_{11}=\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix},\ B_{21}=\begin{pmatrix} 0 & 8 \\ 0 & 2 \end{pmatrix}\ \ So\ \ B=\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$$

Now

$$AB=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}=\begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} \\ A_{21}B_{11}+A_{22}B_{21} \end{bmatrix}$$

This is a valid formula because the sizes of the blocks are such that all of the operations can be performed:

$$A_{11}B_{11}+A_{12}B_{21}=\begin{bmatrix} 1 & 2 & 1 \\ -1 & -2 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix}+\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}\begin{bmatrix} 0 & 8 \\ 0 & 2 \end{bmatrix}=\begin{bmatrix} 3 & 4 \\ -23 & 14 \end{bmatrix}$$

$$A_{21}B_{11}+A_{22}B_{21}=\begin{bmatrix} 2 & 1 & -2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix}+\begin{bmatrix} 1 & 3 \end{bmatrix}\begin{bmatrix} 0 & 8 \\ 0 & 2 \end{bmatrix}$$

$$=\begin{bmatrix} 17 & 0 \end{bmatrix}+\begin{bmatrix} 0 & 14 \end{bmatrix}=\begin{bmatrix} 17 & 14 \end{bmatrix}$$

Thus,

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ -23 & 12 \\ 17 & 14 \end{bmatrix}$$

Note - We see result is same when we multiply A and B without partitions

$$AB = \begin{bmatrix} 1 & 2 & 1 & 0 & 2 \\ -1 & -2 & 3 & 1 & 1 \\ 2 & 1 & -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -5 & 1 \\ 0 & 8 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ -23 & 12 \\ 17 & 14 \end{bmatrix}$$

Note - Sometimes it is more useful to find the square and cube powers of a matrix.

<u>Example 6</u>

The block version of the row –column rule for the product AB of the partitioned matrices

$$A = \left[\begin{array}{ccc|cc} 3 & -4 & 1 & 0 & 2 \\ -1 & 5 & -3 & 1 & 4 \\ \hline 2 & 0 & -2 & 1 & 6 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \left[\begin{array}{cc} 2 & -1 \\ 3 & 0 \\ -5 & 1 \\ \hline 4 & -3 \\ 0 & 2 \end{array}\right] = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$$

where

$$B_{11} = \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix} \text{ and } B_{21} = \begin{pmatrix} 4 & -3 \\ 0 & 2 \end{pmatrix}$$

So

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix}$$

This is a valid formula because the sizes of the blocks are such that all of the operations can be performed:

$$A_{11}B_{11} + A_{12}B_{21} = \begin{bmatrix} 3 & -4 & 1 \\ -1 & 5 & -3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -11 & 2 \\ 32 & 3 \end{bmatrix}$$

$$A_{21}B_{11} + A_{22}B_{21} = \begin{bmatrix} 2 & 0 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 6 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 18 & 5 \end{bmatrix}$$

Thus,

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix} = \left[\begin{array}{cc} -11 & 2 \\ 32 & 3 \\ \hline 18 & 5 \end{array}\right]$$

This result can be confirmed by performing the computation.

$$AB = \begin{bmatrix} 3 & -4 & 1 & 0 & 2 \\ -1 & 5 & -3 & 1 & 4 \\ 2 & 0 & -2 & 1 & 6 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 1 \\ 4 & -3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -11 & 2 \\ 32 & 3 \\ 18 & 5 \end{bmatrix}$$

Example 7

Making block up of matrix

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}, \text{ evaluate } A^2 ?$$

Solution

We partition A as shown below

$$A = \left[\begin{array}{ccc|cc} 2 & 0 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 & 2 \\ \hline 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right] \quad \text{where } A_{32} = \begin{bmatrix} 0 & 2 \\ 0 & 2 \\ 0 & 2 \end{bmatrix},$$

Now

$$A^2 = \begin{pmatrix} 2I_3 & A_{32} \\ O_{23} & 2I_2 \end{pmatrix} \begin{pmatrix} 2I_3 & A_{32} \\ O_{23} & 2I_2 \end{pmatrix}$$

$$= \begin{pmatrix} 4I_3 & 4A_{32} \\ O_{23} & 4I_2 \end{pmatrix}$$

Hence

$$A^2 = \begin{bmatrix} 4 & 0 & 0 & 0 & 8 \\ 0 & 4 & 0 & 0 & 8 \\ 0 & 0 & 4 & 0 & 8 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

Example 8

Let $A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$ Evaluate A^2 ?

Solution

We partition A as shown below

$$A = \left[\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} I_3 & O_{32} & A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t & O_{12} & 1 \end{bmatrix} \quad \text{where } A_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Now $$A^2 = \begin{bmatrix} I_3 & O_{32} & A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t & O_{12} & 1 \end{bmatrix} \begin{bmatrix} I_3 & O_{32} & A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t & O_{12} & 1 \end{bmatrix} = \begin{bmatrix} I_3 + A_1A_1^t & O_{32} & A_1 + A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t + A_1^t & O_{12} & A_1^tA_1 + 1 \end{bmatrix}$$

$$I_3 + A_1A_1^t = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, A_1 + A_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, A_1^t + A_1^t = \begin{bmatrix} 2 & 2 & 2 \end{bmatrix}$$

$$A_1^tA_1 + 1 = \begin{bmatrix} 4 \end{bmatrix}$$

Hence $$A^2 = \begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 2 \\ 1 & 2 & 1 & 0 & 0 & 2 \\ 1 & 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 2 & 2 & 2 & 0 & 0 & 4 \end{bmatrix}$$

Toeplitz matrix

A matrix in which each descending diagonal from left to right is constant is called a **Toeplitz matrix** or **diagonal-constant matrix**

Example 9 The matrix $A = \begin{bmatrix} a & 1 & 2 & 3 \\ 4 & a & 1 & 2 \\ 5 & 4 & a & 1 \\ 6 & 5 & 4 & a \end{bmatrix}$ is a Toeplitz matrix.

Block Toeplitz matrix

A blocked matrix in which blocks (blocked matrices) are repeated down the diagonals of the matrix is called a blocked Toeplitz matrix.

A block Toeplitz matrix **B** has the form

$$B = \begin{bmatrix} B(1,1) & B(1,2) & B(1,3) & B(1,4) & B(1,5) \\ B(2,1) & B(1,1) & B(1,2) & B(1,3) & B(1,4) \\ B(3,1) & B(2,1) & B(1,1) & B(1,2) & B(1,3) \\ B(4,1) & B(3,1) & B(2,1) & B(1,1) & B(1,2) \\ B(5,1) & B(4,1) & B(3,1) & B(2,1) & B(1,1) \end{bmatrix}$$

Inverses of Partitioned Matrices

In this section, we will study about the techniques of inverse of blocked matrices.

Block Diagonal Matrices

A partitioned matrix ***A*** is said to be block diagonal if the matrices on the main diagonal are square and all other position matrices are zero, i.e.

$$A = \begin{bmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k \end{bmatrix} \qquad (1)$$

where the matrices ***D₁, D₂, …, Dₖ*** are square. It can be shown that the matrix A in (1) is invertible if and only if each matrix on the diagonal is invertible. i.e.

$$A^{-1} = \begin{bmatrix} D_1^{-1} & 0 & \dots & 0 \\ 0 & D_2^{-1} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k^{-1} \end{bmatrix}$$

<u>Example 10</u> Let A be a block diagonal matrix

$$A = \left[\begin{array}{cc|cc|c} 1 & 2 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 5 & 4 & 0 \\ \hline 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

Find A^{-1}.

<u>Solution</u>

There are three matrices on the main diagonal; two are 2x2 matrices and one is 1×1 matrix.

In order to find A^{-1}, we evaluate the inverses of three matrices lie in main diagonal of A.

Let $A_{11} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}, A_{22} = \begin{pmatrix} 1 & 1 \\ 5 & 4 \end{pmatrix}$ *and* $A_{33} = (2)$ are matrices of main diagonal of A. Then

$$A_{11}^{-1} = \frac{AdjA_{11}}{A_{11}} = \frac{\begin{pmatrix} -1 & -2 \\ -1 & 1 \end{pmatrix}}{-3} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{-1}{3} \end{pmatrix}.$$

Similarly $A_{22}^{-1} = \dfrac{\begin{pmatrix} 4 & -1 \\ -5 & 1 \end{pmatrix}}{-1}$

$$A_{22}^{-1} = \begin{pmatrix} -4 & 1 \\ 5 & -1 \end{pmatrix}$$

$$\Rightarrow \quad A^{-1} = \left[\begin{array}{cc|cc|c} \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{-1}{3} & 0 & 0 & 0 \\ \hline 0 & 0 & -4 & 1 & 0 \\ 0 & 0 & 5 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 & \frac{1}{2} \end{array}\right]$$

Example 11 Consider the block diagonal matrix

$$A = \left[\begin{array}{cc|cc|c} 8 & -7 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 5 & 2 & 0 \\ \hline 0 & 0 & 0 & 0 & 4 \end{array}\right]$$

There are three matrices on the main diagonal – two 2x2 matrices and one 1x1 matrix.

$$\Rightarrow \quad A^{-1} = \left[\begin{array}{cc|cc|c} 1 & -7 & 0 & 0 & 0 \\ 1 & -8 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & -5 & 3 & 0 \\ \hline 0 & 0 & 0 & 0 & \frac{1}{4} \end{array}\right]$$

Block Upper Triangular Matrices

A partitioned square matrix ***A*** is said to be block upper triangular if the matrices on the main diagonal are square and all matrices below the main diagonal are zero; that is, the matrix is partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ O & A_{22} & \dots & A_{2k} \\ \vdots & \vdots & & \vdots \\ O & O & \dots & A_{kk} \end{bmatrix}$$ where the matrices $\boldsymbol{A_{11}}$, $\boldsymbol{A_{22}},\dots,$ $\boldsymbol{A_{kk}}$ are square.

Note The definition of block lower triangular matrix is similar.

Here, we are going to introduce a formula for finding inverse of a block upper triangular matrix in the following example.

<u>Example 12</u>

Let A be a block upper triangular matrix of the form

$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ where the orders of $\boldsymbol{A_{11}}$ and $\boldsymbol{A_{22}}$ are $p \times p$ *and* $q \times q$ *respectively*. Find $\boldsymbol{A^{-1}}$.

<u>Solution</u>

Let B = $\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ be inverse of A .i.e. $\boldsymbol{A^{-1} = B}$, then

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & O \\ O & I_q \end{bmatrix}$$

$$\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{22}B_{21} & A_{22}B_{22} \end{bmatrix} = \begin{bmatrix} I_p & O \\ O & I_q \end{bmatrix}$$

By comparing corresponding entries, we have

$$A_{11}B_{11} + A_{12}B_{21} = I_p \qquad (1)$$

$$A_{11}B_{12} + A_{12}B_{22} = O \qquad (2)$$

$$A_{22}B_{21} = O \qquad (3)$$

$$A_{22}B_{22} = I_q \qquad (4)$$

Since A_{22} is a square matrix, so by Invertible Matrix Theorem, we have $\boldsymbol{A_{22}}$ is invertible. Thus by eq.(4), $\boldsymbol{B_{22} = A_{22}^{-1}}$. Now by eq. (3), we have $B_{21} = A_{22}^{-1}O = O$. From eq.(1)

$$A_{11}B_{11} + O = I_p$$
$$\Rightarrow A_{11}B_{11} = I_p$$
$$\Rightarrow B_{11} = A_{11}{}^{-1}$$

Finally, form (2),

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \quad and \quad B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

Thus

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{bmatrix} \qquad (5)$$

Example 13

Find A^{-1} of

$$A = \begin{bmatrix} 1 & 9 & -5 & 0 \\ 3 & 3 & 3 & -2 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 3 & 1 \end{bmatrix}$$

Solution

Let partition given matrix A in form

$$A = \left[\begin{array}{cc|cc} 1 & 9 & -5 & 0 \\ 3 & 3 & 3 & -2 \\ \hline 0 & 0 & 7 & 0 \\ 0 & 0 & 3 & 1 \end{array}\right]$$

Put $A_{11} = \begin{bmatrix} 1 & 9 \\ 3 & 3 \end{bmatrix}, A_{12} = \begin{bmatrix} -5 & 0 \\ 3 & -2 \end{bmatrix} \quad and \quad A_{22} = \begin{bmatrix} 7 & 0 \\ 3 & 1 \end{bmatrix}$

Thus $A_{11}{}^{-1} = \dfrac{AdjA}{\det(A)} = \dfrac{\begin{bmatrix} 3 & -9 \\ -3 & 1 \end{bmatrix}}{-24} = \begin{pmatrix} \frac{1}{-8} & \frac{9}{24} \\ \frac{1}{8} & \frac{1}{-24} \end{pmatrix} and\ A_{22}{}^{-1} = \begin{bmatrix} \frac{1}{7} & 0 \\ \frac{-3}{7} & 1 \end{bmatrix}$

Moreover,

$$-A_{11}^{-1}A_{12}A_{22}^{-1} = \begin{pmatrix} \frac{1}{-8} & \frac{9}{24} \\ \frac{1}{8} & \frac{1}{-24} \end{pmatrix}\begin{bmatrix} -5 & 0 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} \frac{1}{7} & 0 \\ \frac{-3}{7} & 1 \end{bmatrix} = \begin{pmatrix} \frac{-4}{7} & \frac{3}{4} \\ \frac{1}{4} & \frac{-1}{12} \end{pmatrix}$$

So by (5), we have

$$A^{-1} = \begin{bmatrix} \frac{-1}{8} & \frac{3}{8} & \frac{-4}{7} & \frac{3}{4} \\ \frac{1}{8} & \frac{-1}{24} & \frac{1}{4} & \frac{-1}{12} \\ 0 & 0 & \frac{1}{7} & 0 \\ 0 & 0 & \frac{-3}{7} & 1 \end{bmatrix}$$

Example 14

Confirm that $A = \begin{bmatrix} 4 & 7 & -5 & 3 \\ 3 & 5 & 3 & -2 \\ 0 & 0 & 7 & 2 \\ 0 & 0 & 3 & 1 \end{bmatrix}$ is an invertible block upper triangular matrix, and then find its inverse by using formula (9).

Solution The matrix is block upper triangular because it can be partitioned into form

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ O & A_{22} & \dots & A_{2k} \\ \vdots & \vdots & & \vdots \\ O & O & \dots & A_{kk} \end{bmatrix}$$

as

$$A = \left[\begin{array}{cc|cc} 4 & 7 & -5 & 3 \\ 3 & 5 & 3 & -2 \\ \hline 0 & 0 & 7 & 2 \\ 0 & 0 & 3 & 1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix}$$

where $A_{11} = \begin{bmatrix} 4 & 7 \\ 3 & 5 \end{bmatrix}, A_{12} = \begin{bmatrix} -5 & 3 \\ 3 & -2 \end{bmatrix}, A_{22} = \begin{bmatrix} 7 & 2 \\ 3 & 1 \end{bmatrix}$

Now $A_{11}^{-1} = \begin{bmatrix} -5 & 7 \\ 3 & -4 \end{bmatrix}$ *and* $A_{22}^{-1} = \begin{bmatrix} 1 & -2 \\ -3 & 7 \end{bmatrix}$

Moreover, $$-A_{11}^{-1}A_{12}A_{22}^{-1} = -\begin{bmatrix} -5 & 7 \\ 3 & -4 \end{bmatrix}\begin{bmatrix} -5 & 3 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} -133 & 295 \\ 78 & -173 \end{bmatrix}$$

So it follows from (9) that the inverse of A is $$A^{-1} = \begin{bmatrix} -5 & 7 & -133 & 295 \\ 3 & -4 & 78 & -173 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -3 & 7 \end{bmatrix}$$

NUMERICAL NOTES

1. When matrices are too large to fit in a computer's high-speed memory, partitioning permits a computer to work with only two or three sub-matrices at a time. For instance, in recent work on linear programming, a research team simplified a problem by partitioning the matrix into 837 rows and 51 columns. The problem's solution took about 4 minutes on a Cray supercomputer.

2. Some high-speed computers, particularly those with vector pipeline architecture, perform matrix calculations more efficiently when the algorithms use partitioned matrices.

3. The latest professional software for high-performance numerical linear algebra, **LAPACK**, makes intensive use of partitioned matrix calculations.

The exercises that follow give practice with matrix algebra and illustrate typical calculations found in applications.

Exercises

In exercises 1 to 3, the matrices A, B, C, X, Y, Z, and I are all n x n and satisfy the indicated equation.

1. $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ X & Y \end{bmatrix} = \begin{bmatrix} 0 & I \\ Z & 0 \end{bmatrix}$

2. $\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}\begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

3. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

4. Suppose that A_{11} is an invertible matrix. Find matrices X and Y such that the product below has the form indicated. Also compute B_{22}.

$$\begin{bmatrix} I & 0 & 0 \\ X & I & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \\ 0 & B_{32} \end{bmatrix}$$

5. The inverse of $\begin{bmatrix} I & 0 & 0 \\ C & I & 0 \\ A & B & I \end{bmatrix} is \begin{bmatrix} I & 0 & 0 \\ Z & I & 0 \\ X & Y & I \end{bmatrix}$. Find X, Y and Z.

6. Find the Inverse of matrix A.

$$A = \left[\begin{array}{cc|cc|c} 8 & 3 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 5 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 8 \end{array}\right]$$

7. Show that $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ is invertible and find its inverse.

8. Compute $\boldsymbol{X^T X}$, when $\boldsymbol{X}$ is partitioned as $[\boldsymbol{X_1} \quad \boldsymbol{X_2}]$.

In exercises 9 and 10, determine whether block multiplication can be used to compute the product using the partitions shown. If so, compute the product by block multiplication.

9. (a) $\left[\begin{array}{cc|cc} -1 & 2 & 1 & 5 \\ 0 & -3 & 4 & 2 \\ \hline 1 & 5 & 6 & 1 \end{array}\right]\left[\begin{array}{cc|c} -2 & 1 & 4 \\ -3 & 5 & 2 \\ \hline 7 & -1 & 5 \\ 0 & 3 & -3 \end{array}\right]$ (b) $\left[\begin{array}{ccc|c} -1 & 2 & 1 & 5 \\ \hline 0 & -3 & 4 & 2 \\ 1 & 5 & 6 & 1 \end{array}\right]\left[\begin{array}{cc|c} -2 & 1 & 4 \\ -3 & 5 & 2 \\ 7 & -1 & 5 \\ \hline 0 & 3 & -3 \end{array}\right]$

10. (a) $\left[\begin{array}{ccc|c} 3 & -1 & 0 & -3 \\ 2 & 1 & 4 & 5 \end{array}\right]\left[\begin{array}{ccc} 2 & -4 & 1 \\ 3 & 0 & 2 \\ 1 & -3 & 5 \\ \hline 2 & 1 & 4 \end{array}\right]$ (b) $\left[\begin{array}{cc} 2 & -5 \\ 1 & 3 \\ 0 & 5 \\ \hline 1 & 4 \end{array}\right]\left[\begin{array}{cc|cc} 2 & -1 & 3 & -4 \\ 0 & 1 & 5 & 7 \end{array}\right]$

11. Compute the product $\begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ -1 & 6 & 2 \end{bmatrix}$ using the column row-rule, and check your answer by calculating the product directly.

In exercises 11 and 12, find the inverse of the block diagonal matrix A.

12. (a) $A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 2 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}$

13. (a) $A = \begin{bmatrix} 5 & 1 & 0 & 0 \\ 4 & 1 & 0 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & -3 & 5 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 3 & 7 & 0 & 0 \\ 0 & 0 & 0 & 4 & 9 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$

14. Find the inverse of the block upper triangular matrix A.

(i) $A = \begin{bmatrix} 2 & 1 & 3 & -6 \\ 1 & 1 & 7 & 4 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 2 & 3 \end{bmatrix}$ (ii) $A = \begin{bmatrix} -1 & -1 & 2 & 5 \\ 2 & 1 & -3 & 8 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 7 & 2 \end{bmatrix}$

15. Find B_1, given that $\begin{bmatrix} A_1 & B_1 \\ 0 & C_1 \end{bmatrix}\begin{bmatrix} A_2 & B_2 \\ 0 & C_2 \end{bmatrix} = \begin{bmatrix} A_3 & B_3 \\ 0 & C_3 \end{bmatrix}$

and

$$A_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, B_2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, C_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, B_3 = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

16. Consider the partitioned linear system

$$\left[\begin{array}{cc|cc} 5 & 2 & 2 & 3 \\ 2 & 1 & -3 & 1 \\ \hline 1 & 0 & 4 & 1 \\ 0 & 1 & 0 & 2 \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \\ \hline x_3 \\ x_4 \end{array}\right] = \left[\begin{array}{c} 2 \\ 6 \\ \hline 0 \\ 0 \end{array}\right]$$

Solve this system by first expressing it as

$$\begin{bmatrix} A & B \\ I & D \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix} \text{ or equivalently, } \begin{array}{l} Au + Bv = b \\ u + Dv = 0 \end{array}$$

next solving the second equation for u in terms of v, and then substituting in the first equation. Check your answer by solving the system directly.

17. Let $A = \begin{bmatrix} -3 & 1 & 2 \\ 1 & -4 & 5 \end{bmatrix}$ *and* $B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}.$

Verify that $AB = col_1(A)row_1(B) + col_2(A)row_2(B) + col_3(A)row_3(B)$

18. Find inverse of the matrix $A = \begin{bmatrix} 4 & 7 & -5 & 3 \\ 3 & 5 & 3 & -2 \\ 0 & 0 & 7 & 2 \\ 0 & 0 & 3 & 1 \end{bmatrix}.$

Lecture 15

Matrix Factorizations

Matrix Factorization

A factorization of a matrix as a product of two or more matrices is called Matrix Factorization.

Uses of Matrix Factorization

Matrix factorizations will appear at a number of key points throughout the course. This lecture focuses on a factorization that lies at the heart of several important computer programs widely used in applications.

***LU* Factorization or LU-decomposition**

LU factorization is a matrix decomposition which writes a matrix as the product of a lower triangular matrix and an upper triangular matrix. This decomposition is used to solve systems of linear equations or calculate the determinant.

Assume A is an $m \times n$ matrix that can be row reduced to echelon form, without row interchanges. Then A can be written in the form A = LU, where L is an $m \times m$ lower triangular matrix with 1's on the diagonal and U is an $m \times n$ echelon form of A. For instance, such a factorization is called LU factorization of A. The matrix L is invertible and is called a unit lower triangular matrix.

$$A = \underset{L}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} \bullet & * & * & * & * \\ 0 & \bullet & * & * & * \\ 0 & 0 & 0 & \bullet & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}$$

***LU* factorization.**

Remarks

1) If A is the square matrix of order m, then the order of both L and U will also be m.

2) In general, not every square matrix A has an LU-decomposition, nor is an LU-decomposition unique if it exists.

<u>Theorem</u> If a square matrix *A* can be reduced to row echelon form with no row interchanges, then *A* has an *LU*-decomposition.

Note

The computational efficiency of the ***LU*** factorization depends on knowing ***L*** and ***U***. The next algorithm shows that the row reduction of ***A*** to an echelon form ***U*** amounts to an ***LU*** factorization because it produces ***L*** with essentially no extra work.

An *LU* Factorization Algorithm

Suppose ***A*** can be reduced to an echelon form ***U*** without row interchanges. Then, since row scaling is not essential, ***A*** can be reduced to ***U*** with only row replacements, adding a multiple of one row to another row below it. In this case, there exist lower triangular elementary matrices $E_1, \ldots, E_p$ such that

$$E_p \ldots E_1 A = U \qquad (1)$$

So $$A = (E_p \ldots E_1)^{-1} U = LU$$

Where $$L = (E_p \ldots E_1)^{-1} \qquad (2)$$

It can be shown that products and inverses of unit lower triangular matrices are also unit-lower triangular. Thus, L is unit-lower triangular.

Note that the row operations in (1), which reduce A to U, also reduce the L in (2) to I, because $E_p \ldots E_1 L = (Ep \ldots E_1)(E_p \ldots E_1)^{-1} = I$. This observation is the key to constructing L.

Procedure for finding an *LU*-decomposition

Step 1: Reduce matrix A to row echelon form U without using row interchanges, keeping track of the multipliers used to introduce the leading 1's and the multipliers used to introduce zeros below the leading 1's.

Step 2: In each position along the main diagonal of L, place the reciprocal of the multiplier that introduced the leading 1 in that position in U.

Step 3: In each position below the main diagonal of L, place the negative of the multiplier used to introduce the zero in that position in U.

Step 4: Form the decomposition $A = LU$.

Example 1 Find an LU-decomposition of

$$A = \begin{bmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix}$$

Solution We will reduce A to a row echelon form U and at each step we will fill in an entry of L in accordance with the four-step procedure above.

$$A = \begin{bmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \qquad \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

* denotes an unknown entry of L.

$$\begin{bmatrix} \boxed{1} & -\frac{1}{3} & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier = \frac{1}{6} \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ \boxed{0} & 2 & 1 \\ \boxed{0} & 8 & 5 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier = -9 \\ \leftarrow multiplier = -3 \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & * & 0 \\ 3 & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \dfrac{1}{2} \\ 0 & 8 & 5 \end{bmatrix} \leftarrow multiplier = \tfrac{1}{2} \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & \boxed{0} & 1 \end{bmatrix} \leftarrow multiplier = -8 \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & * \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & \boxed{1} \end{bmatrix} \leftarrow multiplier = 1 \qquad L = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix}$$

(No actual operation is performed here since there is already a leading 1 in the third row.)
So

$$A = LU = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

OR

Solution We will reduce ***A*** to a row echelon form ***U*** and at each step we will fill in an entry of ***L*** in accordance with the four-step procedure above.

$$A = \begin{bmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \qquad \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

* denotes an unknown entry of ***L***.

$$\approx \begin{bmatrix} \dfrac{6}{6} & \dfrac{-2}{6} & \dfrac{0}{6} \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \quad \frac{1}{6}R_1 \qquad \begin{bmatrix} 6 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \dfrac{-1}{3} & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix}$$

$$\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 9-9(1) & -1-9(\frac{-1}{3}) & 1-9(0) \\ 3-3(1) & 7-3(\frac{-1}{3}) & 5-3(0) \end{bmatrix} \begin{matrix} R_2 - 9R_1 \\ R_3 - 3R_1 \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & * & 0 \\ 3 & * & * \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 2 & 1 \\ 0 & 8 & 5 \end{bmatrix}$$

$$\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ \frac{0}{2} & \frac{2}{2} & \frac{1}{2} \\ 0 & 8 & 5 \end{bmatrix} \frac{1}{2}R_2 \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & * & * \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 8 & 5 \end{bmatrix}$$

$$\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0-8(0) & 8-8(1) & 5-8(\frac{1}{2}) \end{bmatrix} R_3 - 8R_2 \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & * \end{bmatrix}$$

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

$$U = \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & \boxed{1} \end{bmatrix} \leftarrow multiplier = 1 \qquad L = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix}$$

So

$$A = LU = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

Example 2 Find an ***LU*** factorization of $A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$

Solution

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ * & * & 1 & 0 & 0 \\ * & * & * & 1 & 0 \\ * & * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of L.

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \begin{matrix} \leftarrow multiplier \frac{1}{2} \\ \\ \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ * & * & 1 & 0 & 0 \\ * & * & * & 1 & 0 \\ * & * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier -6 \\ \leftarrow multiplier -2 \\ \leftarrow multiplier -4 \\ \leftarrow multiplier\, 6 \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 & 0 \\ 2 & * & 1 & 0 & 0 \\ 4 & * & * & 1 & 0 \\ -6 & * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier\,\frac{1}{3} \\ \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & * & 1 & 0 & 0 \\ 4 & * & * & 1 & 0 \\ -6 & * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier\,3 \\ \leftarrow multiplier\,-6 \\ \leftarrow multiplier\,9 \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & * & 1 & 0 \\ -6 & -9 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 10 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier\,-\frac{1}{5} \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & 0 & -5 & 0 \\ -6 & -9 & 0 & * & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \leftarrow multiplier\,-10 \qquad L = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & 0 & -5 & 0 \\ -6 & -9 & 0 & 10 & 1 \end{bmatrix}$$

Thus, we have constructed the ***LU***-decomposition

$$A = LU = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & 0 & -5 & 0 \\ -6 & -9 & 0 & 10 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Example 3 Find ***LU***–decomposition of $A = \begin{bmatrix} 6 & -2 & -4 & 4 \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix}$

Solution

$$A = \begin{bmatrix} 6 & -2 & -4 & 4 \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of ***L***.

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix} \begin{matrix} \leftarrow \textit{multiplier}\ \frac{1}{6} \\ \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & -2 & -4 & -1 \\ 0 & 4 & 13 & 0 \\ 0 & -2 & -14 & 11 \end{bmatrix} \begin{matrix} \\ \leftarrow \textit{multiplier} -3 \\ \leftarrow \textit{multiplier}\ 12 \\ \leftarrow \textit{multiplier}\ 6 \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -12 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 4 & 13 & 0 \\ 0 & -2 & -14 & 11 \end{bmatrix} \begin{matrix} \\ \leftarrow \textit{multiplier} -\frac{1}{2} \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 5 & -2 \\ 0 & 0 & -10 & 12 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier - 4 \\ \leftarrow multiplier\, 2 \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 1 & 0 \\ -6 & -2 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & -10 & 12 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier\, \frac{1}{5} \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 8 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier\, 10 \\ \\ \end{matrix} \qquad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow multiplier\, \frac{1}{8} \\ \\ \end{matrix} \qquad L = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 8 \end{bmatrix}$$

$$Thus\ \ A = LU = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 8 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Example 4 Find an ***LU*** factorization of $A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$

Solution

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of ***L***.

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \begin{matrix} \leftarrow \textit{multiplier}\ \frac{1}{2} \\ \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} \begin{matrix} \\ \leftarrow \textit{multiplier}\ 4 \\ \leftarrow \textit{multiplier}\ -2 \\ \leftarrow \textit{multiplier}\ 6 \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 1 & 0 & 0 \\ 2 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} \begin{matrix} \\ \leftarrow \textit{multiplier}\ \frac{1}{3} \\ \\ \\ \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow \textit{multiplier}\ 9 \\ \leftarrow \textit{multiplier}\ -12 \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & -9 & 1 & 0 \\ -6 & 12 & * & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & \frac{7}{4} \end{bmatrix} \leftarrow multiplier\ \frac{1}{4} \qquad L = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & -9 & 1 & 0 \\ -6 & 12 & 0 & 4 \end{bmatrix}$$

Thus $A = LU = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & -9 & 1 & 0 \\ -6 & 12 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & \frac{7}{4} \end{bmatrix}$

Matrix Inversion by *LU*-Decomposition

Many of the best algorithms for inverting matrices use ***LU***-decomposition. To understand how this can be done, let ***A*** be an invertible $n \times n$ matrix, let $A^{-1} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ be its unknown inverse partitioned into column vectors, and let $I = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix}$ be then $n \times n$ identity matrix partitioned into column vectors. The matrix equation $AA^{-1} = I$ can be expressed as

$$A\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix}$$
$$\begin{bmatrix} Ax_1 & Ax_2 & \cdots & Ax_n \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix}$$

which tells us that the unknown column vectors of A^{-1} can be obtained by solving the *n*-linear systems.

$$Ax_1 = e_1, \quad Ax_2 = e_2, \cdots, \quad Ax_n = e_n \qquad (1^*)$$

As discussed above, this can be done by finding an ***LU***-decomposition of ***A***, and then using that decomposition to solve each of the n systems in (1*).

Solving Linear System by LU-Factorization

When ***A* =*LU***, the equation ***Ax* = *b*** can be written as ***L* (*Ux*) = *b***. Writing ***y*** for ***Ux***, we can find ***x*** by solving the pair of equations; ***Ly* = *b*** and ***Ux* = *y*** (2*)

First solve ***Ly* = *b*** for ***y*** and then solve ***Ux* = *y*** for ***x***. Each equation is easy to solve because ***L*** and ***U*** are triangular.

Procedure

Step 1: Rewrite the system $\boldsymbol{A}\ \boldsymbol{x} = \boldsymbol{b}$ as $\boldsymbol{LU}\ \boldsymbol{x} = \boldsymbol{b}$ (3*)

Step 2: Define a new unknown $\boldsymbol{y}$ by letting $\boldsymbol{U}\ \boldsymbol{x} = \boldsymbol{y}$ (4*)
And rewrite (3*) as $\boldsymbol{L}\ \boldsymbol{y} = \boldsymbol{b}$

Step 3: Solve the system $\boldsymbol{L}\ \boldsymbol{y} = \boldsymbol{b}$ for the unknown $\boldsymbol{y}$.
Step 4: Substitute the known vector $\boldsymbol{y}$ into (4*) and solve for $\boldsymbol{x}$.

This procedure is called the method of ***LU***-Decomposition.

Although ***LU***-Decomposition converts the problem of solving the single system $\boldsymbol{A}\ \boldsymbol{x} = \boldsymbol{b}$ into the problem of solving the two systems, $\boldsymbol{L}\ \boldsymbol{y} = \boldsymbol{b}$ and $\boldsymbol{U}\ \boldsymbol{x} = \boldsymbol{y}$, these systems are easy to solve because their co-efficient matrices are triangular.

Example 5 Solve the given system ($\boldsymbol{Ax} = \boldsymbol{b}$) by ***LU***-Decomposition

$$\begin{aligned} 2x_1 + 6x_2 + 2x_3 &= 2 \\ -3x_1 - 8x_2 \qquad &= 2 \\ 4x_1 + 9x_2 + 2x_3 &= 3 \end{aligned} \qquad (1)$$

Solution We express the system (1) in matrix form:

$$\underset{\boldsymbol{A}}{\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix}} \underset{\boldsymbol{x}}{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} = \underset{\boldsymbol{b}}{\begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}}$$

We derive an ***LU***-decomposition of ***A***.

$$A = \begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \qquad\qquad L = \begin{bmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \leftarrow multiplier\ \frac{1}{2} \qquad\qquad \begin{bmatrix} 2 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & -3 & -2 \end{bmatrix} \begin{matrix} \\ \leftarrow multiplier\ 3 \\ \leftarrow multiplier\ -4 \end{matrix} \qquad\qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow multiplier\ 3 \qquad\qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow multiplier\,\frac{1}{7} \qquad\qquad L = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix}$$

Thus

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \qquad (2)$$

$$\boldsymbol{A} \quad = \quad \boldsymbol{L} \quad \boldsymbol{U}$$

From (2) we can rewrite this system as

$$\begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \qquad (3)$$

$$\boldsymbol{L} \quad \boldsymbol{U} \quad \boldsymbol{x} = \boldsymbol{b}$$

As specified in Step 2 above, let us define y_1, y_2 and y_3 by the equation

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \qquad (4)$$

$$\boldsymbol{U} \quad \boldsymbol{x} = \boldsymbol{y}$$

which allows us to rewrite (3) as

$$\begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \qquad (5)$$

$$\boldsymbol{L} \quad \boldsymbol{y} = \boldsymbol{b}$$

or equivalently, as

$$\begin{aligned} 2y_1 \qquad\qquad &= 2 \\ -3y_1 + y_2 \qquad &= 2 \\ 4y_1 - 3y_2 + 7y_3 &= 3 \end{aligned}$$

This system can be solved by a procedure that is similar to back substitution, except that we solve the equations from the top down instead of from the bottom up. This procedure, called **forward substitution**, yields

$$y_1 = 1, \quad y_2 = 5, \quad y_3 = 2.$$

As indicated in Step 4 above, we substitute these values into (4), which yields the linear system

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix}$$

$$x_1 + 3x_2 + x_3 = 1$$

or equivalently, $$x_2 + 3x_3 = 5$$

$$x_3 = 2$$

Solving this system by back substitution yields $x_1 = 2$, $x_2 = -1$, $x_3 = 2$

Example 6 It can be verified that

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

Use this ***LU*** factorization of ***A*** to solve ***Ax = b***, where $\boldsymbol{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$

Solution The solution of ***Ly = b*** needs only 6 multiplications and 6 additions, because the arithmetic takes place only in column 5. (The zeros below each pivot in ***L*** are created automatically by our choice of row operations.)

$$[L \quad b] = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = [I \quad y]$$

Then, for ***Ux = y***, the "backwards" phase of row reduction requires 4 divisions, 6 multiplications, and 6 additions. (For instance, creating the zeros in column 4 of **[*U* *y*]** requires 1 division in row 4 and 3 multiplication – addition pairs to add multiples of row 4 to the rows above.)

$$[U \quad y] = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \quad x = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

To find ***x*** requires 28 arithmetic operations, or "flops" (floating point operations), excluding the cost of finding ***L*** and ***U***. In contrast, row reduction of **[*A* *b*]** to **[*I* *x*]** takes 62 operations.

Numerical Notes

The following operation counts apply to an $n \times n$ dense matrix ***A*** (with most entries nonzero) for *n* moderately large, say, $n \geq 30$.

1. Computing an ***LU*** factorization of ***A*** takes about $2n^3/3$ flops (about the same as row reducing [***A*** ***b***]), whereas finding $\boldsymbol{A^{-1}}$ requires about $2n^3$ flops.
2. Solving ***Ly = b*** and ***Ux = y*** requires about $2n^2$ flops, because $n \times n$ triangular system can be solved in about n^2 flops.
3. Multiplication of ***b*** by $\boldsymbol{A^{-1}}$ also requires about $2n^2$ flops, but the result may not be as accurate as that obtained from ***L*** and ***U*** (because of round off error when computing both $\boldsymbol{A^{-1}}$ and $\boldsymbol{A^{-1}b}$).
4. If ***A*** is sparse (with mostly zero entries), then ***L*** and ***U*** may be sparse, too, whereas $\boldsymbol{A^{-1}}$ is likely to be dense. In this case, a solution of ***Ax = b*** with an ***LU*** factorization is much faster than using $\boldsymbol{A^{-1}}$.

Example 7(Gaussian Elimination Performed as an LU-Decomposition)
In Example 5, we showed how to solve the linear system

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \qquad (6)$$

using an ***LU***-decomposition of the coefficient matrix, but we did not discuss how the factorization was derived. In the course of solving the system, we obtained the intermediate vector $\boldsymbol{y} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix}$ by using forward substitution to solve system (5).

We will now use the procedure discussed above to find both the ***LU***-decomposition and the vector ***y*** by row operations on the augmented matrix for (6).

$$\left[A|b\right] = \left[\begin{array}{ccc|c} 2 & 6 & 2 & 2 \\ -3 & -8 & 0 & 2 \\ 4 & 9 & 2 & 3 \end{array}\right] \qquad \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix} = L \quad (* = unknown \quad entries)$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 1 \\ -3 & -8 & 0 & 2 \\ 4 & 9 & 2 & 3 \end{array}\right] \qquad \begin{bmatrix} 2 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & -3 & -2 & -1 \end{array}\right] \qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & * & 0 \\ 4 & * & * \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 7 & 14 \end{array}\right] \qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & * \end{bmatrix}$$

$$\begin{bmatrix} U \,|\, y \end{bmatrix} = \left[\begin{array}{ccc|c} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{array}\right] \qquad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} = L$$

These results agree with those in Example 5, so we have found an ***LU***-decomposition of the coefficient matrix and simultaneously have completed the forward substitution required to find ***y***.

All that remains to solve the given system is to solve the system ***Ux* = *y*** by back substitution. The computations were performed in Example5.

A Matrix Factorization in Electrical Engineering

Matrix factorization is intimately related to the problem of constructing an electrical network with specified properties. The following discussion gives just a glimpse of the connection between factorization and circuit design.

Suppose the box in below Figure represents some sort of electric circuit, with an input and output. Record the input voltage and current by $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$ (with voltage ***v*** in volts and current ***i*** in amps), and record the output voltage and current by $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$. Frequently, the transformation $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} \to \begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$ is linear. That is, there is a matrix A, called the transfer matrix, such that $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix} = A \begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$

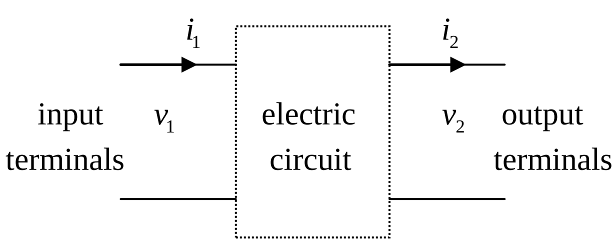


Figure A circuit with input and output terminals.

Above Figure shows a ladder network, where two circuits (there could be more) are connected in series, so that the output of one circuit becomes the input of the next circuit. The left circuit in Figure is called a series circuit, with resistance R_1 (in ohms);

The right circuit is a shunt circuit, with resistance R_2. Using Ohm's law and Kirchhoff's laws, one can show that the transfer matrices of the series and shunt circuits, respectively, are:

$$\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} \quad and \quad \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}$$

Transfer matrix of series circuit   Transfer matrix of shunt circuit

Example 8

a) Compute the transfer matrix of the ladder network in the above Figure .

b) Design a ladder network whose transfer matrix is $\begin{bmatrix} 1 & -8 \\ -0\cdot 5 & 5 \end{bmatrix}$.

Solution

a) Let $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$ be the transfer matrices of the series of the series and shunt circuits, respectively. Then an input vector $\boldsymbol{x}$ is transformed first into $\boldsymbol{A_1 x}$ and then into $\boldsymbol{A_2\ (A_1x)}$. The series connection of the circuits corresponds to composition of linear transformations; and the transfer matrix of the ladder network is (note the order)

$$A_2A_1 = \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix} \quad (6)$$

b) We seek to factor the matrix $\begin{bmatrix} 1 & -8 \\ -0\cdot 5 & 5 \end{bmatrix}$ into the product of transfer matrices, such as in (6). So we look for R_1 and R_2 to satisfy

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -0\cdot 5 & 5 \end{bmatrix}$$

From the (1, 2) – entries, R_1 = 8 ohms, and from the (2, 1) – entries, $1/R_2$ = 0.5 ohm and R_2 = 1/0.5 = 2 ohms. With these values, the network has the desired transfer matrix.

Note

A network transfer matrix summarizes the input-output behavior (“Design specifications”) of the network without reference to the interior circuits. To physically build a network with specified properties, an engineer first determines if such a network can be constructed (or realized). Then the engineer tries to factor the transfer matrix into matrices corresponding to smaller circuits that perhaps are already manufactured and ready for assembly. In the common case of alternating current, the entries in the transfer matrix are usually rational complex-valued functions. A standard problem is to find a minimal realization that uses the smallest number of electrical components.

Exercises

Find an LU factorization of the matrices in exercises 1 to 8.

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

2. $\begin{bmatrix} 3 & -1 & 2 \\ -3 & -2 & 10 \\ 9 & -5 & 6 \end{bmatrix}$

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

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

5. $\begin{bmatrix} 2 & -4 & 4 & -2 \\ 6 & -9 & 7 & -3 \\ -1 & -4 & 8 & 0 \end{bmatrix}$

6. $\begin{bmatrix} 2 & -6 & 6 \\ -4 & 5 & -7 \\ 3 & 5 & -1 \\ -6 & 4 & -8 \\ 8 & -3 & 9 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 4 & -1 & 5 \\ 3 & 7 & -2 & 9 \\ -2 & -3 & 1 & -4 \\ -1 & 6 & -1 & 7 \end{bmatrix}$

8. $\begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$

Solve the equation ***Ax = b*** by using ***LU***-factorization.

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

10. $A = \begin{bmatrix} 4 & 3 & -5 \\ -4 & -5 & 7 \\ 8 & 6 & -8 \end{bmatrix}, b = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$

11. $A = \begin{bmatrix} 2 & -1 & 2 \\ -6 & 0 & -2 \\ 8 & -1 & 5 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$

12. $A = \begin{bmatrix} 2 & -2 & 4 \\ 1 & -3 & 1 \\ 3 & 7 & 5 \end{bmatrix}, b = \begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix}$

13. $A = \begin{bmatrix} 1 & -2 & -4 & -3 \\ 2 & -7 & -7 & -6 \\ -1 & 2 & 6 & 4 \\ -4 & -1 & 9 & 8 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 7 \\ 0 \\ 3 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 3 & 4 & 0 \\ -3 & -6 & -7 & 2 \\ 3 & 3 & 0 & -4 \\ -5 & -3 & 2 & 9 \end{bmatrix}, b = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$

Lecture 16

Iterative Solutions of Linear Systems

Consistent linear systems are solved in one of two ways by direct calculation (matrix factorization) or by an iterative procedure that generates a sequence of vectors that approach the exact solution. When the coefficient matrix is large and sparse (with a high proportion of zero entries), iterative algorithms can be more rapid than direct methods and can require less computer memory. Also, an iterative process may be stopped as soon as an approximate solution is sufficiently accurate for practical work.

General Framework for an Iterative Solution of *Ax = b*

Throughout the section, ***A*** is an invertible matrix. The goal of an iterative algorithm is to produce a sequence of vectors,

$$x^{(0)}, x^{(1)}, \ldots, x^{(k)}, \ldots$$

that converges to the unique solution say x^* of ***Ax = b***, in the sense that the entries in $x^{(k)}$ are as close as desired to the corresponding entries in x^* for all *k* sufficiently large.

To describe a recursion algorithm that produces $x^{(k+1)}$ from $x^{(k)}$, we write ***A = M – N*** for suitable matrices ***M*** and ***N,*** and then we rewrite the equation ***Ax = b*** as ***Mx – Nx = b*** and

$$Mx = Nx + b$$

If a sequence { $x^{(k)}$ } satisfies

$$\boxed{Mx^{(k+1)} = Nx^{(k)} + b \qquad (k = 0, 1, \ldots)} \qquad (1)$$

and if the sequence converges to some vector x^*, then it can be shown that $Ax^* = b.$[The vector on the left in (1) approaches Mx^*, while the vector on the right in (1) approaches $Nx^* + b.$ This implies that $Mx^* = Nx^* + b$ and $Ax^* = b.$

For $x^{(k+1)}$ to be uniquely specified in (1), ***M*** must be invertible. Also, ***M*** should be chosen so that $x^{(k+1)}$ is easy to calculate. There are two iterative methods below to illustrate two simple choices for ***M***.

1) Jacobi's Method

This method assumes that the diagonal entries of ***A*** are all nonzero.
Choosing ***M*** as the diagonal matrix formed from the diagonal entries of ***A***. So next ***N = M – A***,

$$\therefore (1) \Rightarrow \; Mx^{(k+1)} = (M - A)x^{(k)} + b \qquad (k = 0, 1, \ldots)$$

For simplicity, we take the zero vector as $x^{(0)}$ as the initial approximation.

Example 1

Apply Jacobi's method to the system

$$\begin{aligned} 10x_1 + x_2 - x_3 &= 18 \\ x_1 + 15x_2 + x_3 &= -12 \\ -x_1 + x_2 + 20x_3 &= 17 \end{aligned} \qquad (2)$$

Take $\boldsymbol{x}^{(0)} = (0, 0, 0)$ as an initial approximation to the solution, and use six iterations (that is, compute $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(6)}$).

Solution

For some $\boldsymbol{k}$, let $\boldsymbol{x}^{(k)} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (x_1, x_2, x_3)$ and $\boldsymbol{x}^{(k+1)} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = (y_1, y_2, y_3)$

Firstly we will construct $\boldsymbol{M}$ and $\boldsymbol{N}$ from $\boldsymbol{A}$.

Here

$$A = \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix}$$

Its diagonal entries will give

$$M = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 20 \end{bmatrix} \text{ and}$$

$$N = M - A = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 20 \end{bmatrix} - \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

Now the recursion: $Mx^{(k+1)} = (M - A)x^{(k)} + b \qquad (here\ k = 0,1,\ldots 6)$

implies

$$\begin{bmatrix} 10 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 20 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 10y_1 \\ 15y_2 \\ 20y_3 \end{bmatrix} = \begin{bmatrix} 0x_1 - 1x_2 + 1x_3 \\ -1x_1 + 0x_2 - x_3 \\ 1x_1 - 1x_2 + 0x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 10y_1 \\ 15y_2 \\ 20y_3 \end{bmatrix} = \begin{bmatrix} 0x_1 - 1x_2 + 1x_3 + 18 \\ -1x_1 + 0x_2 - x_3 - 12 \\ 1x_1 - 1x_2 + 0x_3 + 17 \end{bmatrix}$$

Comparing the corresponding entries on both sides, we have

$$10y_1 = -x_2 + x_3 + 18$$

$$15y_2 = -x_1 \quad - x_3 - 12$$
$$20y_3 = x_1 - x_2 \quad + 17$$

And

$$y_1 = (- x_2 + x_3 + 18)/10$$
$$y_2 = (- x_1 - x_3 - 12)/15 \qquad (3)$$
$$y_3 = (x_1 - x_2 + 17)/20$$

1st Iteration

For $\boldsymbol{k} = 0$, put $\boldsymbol{x}^{(0)} = (x_1, x_2, x_3) = (0, 0, 0)$ in (3) and compute

$$\boldsymbol{x}^{(1)} = (y_1, y_2, y_3) = (18/10, - 12/15, 17/20) = (1.8, -0.8, 0.85)$$

2nd Iteration

For k = 1, put $\boldsymbol{x}^{(1)} = (1.8, -.8, .85)$

$$y_1 = [- (-0.8) + (0.85) + 18]/10 = 1.965$$
$$y_2 = [- (1.8) - (0.85) - 12]/15 = -0.9767$$
$$y_3 = [(1.8) - (-0.8) + 17]/20 = 0.98$$

Thus, $\boldsymbol{x}^{(2)} = (1.965, -.9767, .98)$.

The entries in $\boldsymbol{x}^{(2)}$ are used on the right in (3) to compute the entries in $\boldsymbol{x}^{(3)}$, and so on. Here are the results, with calculations using **MATLAB** and results reported to four decimal places:

$$\begin{array}{ccccccc} x^{(0)} & x^{(1)} & x^{(2)} & x^{(3)} & x^{(4)} & x^{(5)} & x^{(6)} \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1.8 \\ -.8 \\ .85 \end{bmatrix} & \begin{bmatrix} 1.965 \\ -.9767 \\ .98 \end{bmatrix} & \begin{bmatrix} 1.9957 \\ -.9963 \\ .9971 \end{bmatrix} & \begin{bmatrix} 1.9993 \\ -.9995 \\ .9996 \end{bmatrix} & \begin{bmatrix} 1.9999 \\ -.9999 \\ .9999 \end{bmatrix} & \begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix} \end{array}$$

If we decide to stop when the entries in $x^{(k)}$ and $x^{(k-1)}$ differ by less than .001, then we need five iterations ($k = 5$).

Alternative Approach

If we express the above system as

$$10x_1 + \quad x_2 - \quad x_3 = 18 \Rightarrow x_1 = \frac{18 - x_2 + x_3}{10}$$

$$x_1 + 15x_2 + \quad x_3 = -12 \Rightarrow x_2 = \frac{-12 - x_1 - x_3}{15}$$

$$-x_1 + \quad x_2 + 20x_3 = 17 \Rightarrow x_3 = \frac{17 + x_1 - x_2}{20}$$

$\therefore$ the equivalent system is

$$x_1 = \frac{18 - x_2 + x_3}{10}$$
$$x_2 = \frac{-12 - x_1 - x_3}{15}$$
$$x_3 = \frac{17 + x_1 - x_2}{20}$$

Now put $(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3})=(0,0,0)= \boldsymbol{x}^{(0)}$ in the RHS to have

$\boldsymbol{x_1}$=(18-0+0)/10 = 1.80

$\boldsymbol{x_2}$= (-12-0-0)/15 = -0.80

$\boldsymbol{x_3}$ = (17+0-0)/20 = 0.85

Which gives $\boldsymbol{x}^{(1)}$ = (1.80,-0.80,0.85) ----put this again on RHS of the equivalent system to get

$\boldsymbol{x_1}$=(18+0.80+0.85)/10 =1.965

$\boldsymbol{x_2}$= (-12-1.80-0.85)/15 = -0.9767

$\boldsymbol{x_3}$= (17+1.80+0.80)/20 = 0.98

So in the similar fashion, we can get the next approximate solutions: $\boldsymbol{x}^{(3)}, \boldsymbol{x}^{(4)}, \boldsymbol{x}^{(5)}$ and $\boldsymbol{x}^{(6)}$

Next example will be solved by following this approach.

Example 2

Use Jacobi iteration to approximate the solution of the system

$$\begin{aligned} &20 x_1 + x_2 - x_3 = 17 \\ &x_1 - 10x_2 + x_3 = 13 \\ &-x_1 + x_2 + 10x_3 = 18 \end{aligned}$$

Stop the process when the entries in two successive iterations are the same when rounded to four decimal places.

Solution

As required for Jacobi iteration, we begin by solving the first equation for x_1, the second for x_2, and the third for x_3. This yields

$$\begin{aligned} x_1 &= \frac{17}{20} - \frac{1}{20}x_2 + \frac{1}{20}x_3 = 0.85 - 0.05x_2 + 0.05x_3 \\ x_2 &= -\frac{13}{10} + \frac{1}{10}x_1 + \frac{1}{10}x_3 = -1.3 + 0.1x_1 + 0.1x_3 \qquad (4) \\ x_3 &= \frac{18}{10} + \frac{1}{10}x_1 - \frac{1}{10}x_2 = 1.8 + 0.1x_1 - 0.1x_2 \end{aligned}$$

which we can write in matrix form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -0.05 & 0.05 \\ 0.1 & 0 & 0.1 \\ 0.1 & -0.1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix} \qquad (5)$$

Since we have no special information about the solution, we will take the initial approximation to be $x_1 = x_2 = x_3 = 0$. To obtain the first iterate, we substitute these values into the right side of (5). This yields

$$y_1 = \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix}$$

To obtain the second iterate, we substitute the entries of y_1 into the right side of (5). This yields

$$y_2 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -0.05 & 0.05 \\ 0.1 & 0 & 0.1 \\ 0.1 & -0.1 & 0 \end{bmatrix} \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix} + \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1.005 \\ -1.035 \\ 2.015 \end{bmatrix}$$

Repeating this process until two successive iterations match to four decimal places yields the results in the following table:

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7
x_1	0	0.8500	1.0050	1.0025	1.0001	1.0000	1.0000	1.0000
x_2	0	-1.3000	-1.0350	-0.9980	-0.9994	-1.0000	-1.0000	-1.0000
x_3	0	1.8000	2.0150	2.0040	2.0000	1.9999	2.0000	2.0000

The Gauss-Seidel Method

This method uses the recursion (1) with *M* the lower triangular part of *A*. That is, *M* has the same entries as *A* on the diagonal and below, and *M* has zeros above the diagonal. See Fig. 1. As in Jacobi's method, the diagonal entries of *A* must be nonzero in order for *M* to be invertible.

$$A = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \qquad M = \begin{bmatrix} * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & * & 0 \\ * & * & * & * & * \end{bmatrix}$$

Figure 01:The Lower Triangular Part of A

Example 3

Apply the Gauss – Seidel method to the system in Example 1 with $x^{(0)} = (0,0,0)$ and six iterations.

$$\begin{aligned} 10x_1 + x_2 - x_3 &= 18 \\ x_1 + 15x_2 + x_3 &= -12 \\ -x_1 + x_2 + 20x_3 &= 17 \end{aligned} \qquad (6)$$

Solution

For some $\boldsymbol{k}$, let $\mathbf{x}^{(\mathbf{k})} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (x_1, x_2, x_3)$ and $\mathbf{x}^{(\mathbf{k+1})} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = (y_1, y_2, y_3)$

Again, firstly we construct matrices $\boldsymbol{M}$ and $\boldsymbol{N}$ from the coefficient matrix $\boldsymbol{A}$.

Here $A = \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix}$

Since matrix M is constructed by

1) taking the values along the diagonal and below the diagonal of coefficient matrix A.
2) putting the zeros above the diagonal at upper trianular position.

So

$$M = \begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix}$$

Now,

$$N = M - A = \begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix} - \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Now the recursion: $Mx^{(k+1)} = (M - A)x^{(k)} + b \qquad (here\ k = 0,1,...6)$

implies

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 10y_1 & 0 & 0 \\ y_1 & 15y_2 & 0 \\ -y_1 & y_2 & 20y_3 \end{bmatrix} = \begin{bmatrix} -x_2 + x_3 \\ -x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 10y_1 & 0 & 0 \\ y_1 & 15y_2 & 0 \\ -y_1 & y_2 & 20y_3 \end{bmatrix} = \begin{bmatrix} -x_2 + x_3 + 18 \\ -x_3 - 12 \\ 0 + 17 \end{bmatrix}$$

Comparing the corresponding entries on both sides, we have

$$10y_1 = -x_2 + x_3 + 18$$
$$y_1 + 15y_2 = -x_3 - 12$$
$$-y_1 + y_2 + 20y_3 = 17$$

This further implies as

$$\left.\begin{aligned} 10y_1 = -x_1 - x_3 + 18 &\Rightarrow y_1 = (-x_2 + x_3 + 18)/10 \\ y_1 + 15y_2 = -x_3 - 12 &\Rightarrow y_2 = (-y_1 - x_3 - 12)/15 \\ -y_1 + y_2 + 20y_3 = 17 &\Rightarrow y_3 = (y_1 - y_2 + 17)/20 \end{aligned}\right\} \text{-----------(7)}$$

Another way to view (7) is to solve each equation in (6) for x_1, x_2, x_3, respectively and regard the highlighted *x's* as the values:

$$\begin{aligned} x_1 &= (-x_2 + x_3 + 18)/10 \\ x_2 &= (-x_1 - x_3 - 12)/15 \\ x_3 &= (x_1 - x_2 + 17)/20 \end{aligned} \qquad (8)$$

Use the first equation to calculate the new x_1 [called y_1 in (7)] from x_2 and x_3. Then, use this new x_1 along with x_3 in the second equation to compute the new x_2. Finally, in the third equation, use the new values for x_1 and x_2 to compute x_3. In this way, the latest information about the variables is used to compute new values. [A computer program would use statements corresponding to the equations in (8).]

From $x^{(0)} = (0, 0, 0)$, we obtain

$$x_1 = [-(0) + (0) + 18]/10 = 1.8$$

$$x_2 = [-(1.8) - (0) - 12]/15 = -.92$$

$$x_3 = [+(1.8) - (-.92) + 17]/20 = .986$$

Thus, $x^{(1)} = (1.8, -.92, .986)$. The entries in $x^{(1)}$ are used in (8) to produce $x^{(2)}$ and so on. Here are the **MATLAB** calculations reported to four decimal places:

$$\begin{array}{ccccccc} x^{(0)} & x^{(1)} & x^{(2)} & x^{(3)} & x^{(4)} & x^{(5)} & x^{(6)} \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1.8 \\ -.92 \\ .986 \end{bmatrix} & \begin{bmatrix} 1.9906 \\ -.9984 \\ .9995 \end{bmatrix} & \begin{bmatrix} 1.9998 \\ -.9999 \\ 1.0000 \end{bmatrix} & \begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix} & \begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix} & \begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix} \end{array}$$

Observe that when k is 4, the entries in $x^{(4)}$ and $x^{(k-1)}$ differ by less than $.001$. The values in $x^{(6)}$ in this case happen to be accurate to eight decimal places.

<u>**Alternative Approach**</u>

If we express the above system as

$$10x_1 + x_2 - x_3 = 18 \Rightarrow x_1 = (-x_2 + x_3 + 18)/10 \text{ ------(a)}$$

$$x_1 + 15x_2 + x_3 = -12 \Rightarrow x_2 = (-x_1 - x_3 - 12)/15 \text{------(b)}$$

$$-x_1 + x_2 + 20x_3 = 17 \Rightarrow x_3 = (x_1 - x_2 + 17)/20 \text{ ------(c)}$$

<u>**Ist Iteration**</u>

Put $\boldsymbol{x_2 = x_3 = 0}$ in (a)

$\boldsymbol{x_1}$= 18/10 =1.80
Put $\boldsymbol{x_1}$=1.80 and $\boldsymbol{x}_3$=0 in (b)
$\boldsymbol{x}_2$ = (-1.80-0-12)/15= -0.92
Put $\boldsymbol{x}_1$=1.80, $\boldsymbol{x}_2$ = -0.92 in (c)
$\boldsymbol{x}_3$=(1.80+0.92+17)/20 = 0.9863

So, $\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 1.8 \\ -.92 \\ .986 \end{bmatrix} = x^{(1)}$

2nd iteration

Put $\boldsymbol{x}_2$ = -0.92, x3=0.9863 in (a)
$\boldsymbol{x}_1$=(0.92+0.9863+18)/10 = 1.9906
Put $\boldsymbol{x}_1$=1.9906(from 2nd iteration) and $\boldsymbol{x}_3$=0.9863(from 1st iteration) in (b)
$\boldsymbol{x}_2$=(-1.9906-0.9863-12)/15 = -0.9984
Put $\boldsymbol{x}_1$=1.9906, $\boldsymbol{x}$2=-0.9984(both from 2nd iteration) in (c)
$\boldsymbol{x}_3$=(1.9906+0.9984+17)/20 = 0.9995

So, $\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 1.9906 \\ -.9984 \\ .9995 \end{bmatrix} = x^{(2)}$

So in the similar fashion, we can get the next approximate solutions: $\boldsymbol{x}^{(3)}$,$\boldsymbol{x}^{(4)}$,$\boldsymbol{x}^{(5)}$ and $\boldsymbol{x}^{(6)}$
Next example will be solved by following this approach.

Example 4

Use Gauss-Seidel to approximate the solution of the linear system in example 2 to four decimal places.

Solution

As before, we will take $x_1 = x_2 = x_3 = 0$ as the initial approximation. First, we will substitute $x_2 = 0$ and $x_3 = 0$ into the right side of the first equation of (4) to obtain the new x_1. Then, we will substitute $x_3 = 0$ and the new x_1 into the right side of the second equation to obtain the new x_2, and finally, we will substitute the new x_1 and new x_2 into the right side of the third equation to obtain the new x_3. The computations are as follows:

$$x_1 = 0.85 - (0.05)(0) + (0.05)(0) = 0.85$$
$$x_2 = -1.3 + (0.1)(0.85) + (0.1)(0) = -1.215$$
$$x_3 = 1.8 + (0.1)(0.85) - (0.1)(-1.215) = 2.0065$$

Thus, the first Gauss-Seidel iterate is

$$y_1 = \begin{bmatrix} 0.8500 \\ -1.2150 \\ 2.0065 \end{bmatrix}$$

Similarly, the computations for second iterate are

$$x_1 = 0.85 - (0.05)(-1.215) + (0.05)(2.0065) = 1.011075$$
$$x_2 = -1.3 + (0.1)(1.011075) + (0.1)(2.0065) = -0.9982425$$
$$x_3 = 1.8 + (0.1)(1.011075) - (0.1)(-0.9982425) = 2.00093175$$

Thus, the second Gauss-Seidel iterate to four decimal places is

$$y_2 \approx \begin{bmatrix} 1.0111 \\ -0.9982 \\ 2.0009 \end{bmatrix}$$

The following table shows the first four Gauss-Seidel iterates to four decimal places. Comparing both tables, we see that the Gauss-Seidel method produced the solution to four decimal places in four iterations whereas, the Jacobi method required six.

	y_0	y_1	y_2	y_3	y_4
x_1	0	0.8500	1.0111	1.0000	1.0000
x_2	0	-1.2150	-0.9982	-0.9999	-1.0000
x_3	0	2.0065	2.0009	2.0000	2.0000

Comparison of Jacobi's and Gauss-Seidel method

There exist examples where Jacobi's method is faster than the Gauss-Seidel method, but usually a Gauss-Seidel sequence converges faster (means to say iterative solution approaches to the unique solution), as in Example 2. (If parallel processing is available, Jacobi might be faster because the entries in $x^{(k)}$ can be computed simultaneously.) There are also examples where one or both methods fail to produce a convergent sequence, and other examples where a sequence is convergent, but converges too slowly for practical use.

Condition for the Convergence of both Iterative Mthods

Fortunately, there is a simple condition that guarantees (but is not essential for) the convergence of both Jacobi and Gauss-Seidel sequences. This condition is often satisfied, for instance, in large-scale systems that can occur during numerical solutions of partial differential equations (such as Laplace's equation for steady-state heat flow).

An $n \times n$ matrix A is said to be ***strictly diagonally dominant*** *if the absolute value of each diagonal entry exceeds the sum of the absolute values of the other entries in the same row.*

In this case it can be shown that A is invertible and that both the Jacobi and Gauss-Seidel sequences converge to the unique solution of $Ax = b$, for any initial $x^{(0)}$. (The speed of the convergence depends on how much the diagonal entries dominate the corresponding row sums.)

The coefficient matrices in Examples 1 and 2 are strictly diagonally dominant, but the following matrix is not. Examine each row:

$$\begin{bmatrix} -6 & 2 & -3 \\ 1 & 4 & -2 \\ 3 & -5 & 8 \end{bmatrix} \qquad \begin{aligned} |-6| &> |2| + |-3| \\ |4| &> |1| + |-2| \\ |8| &= |3| + |-5| \end{aligned}$$

The problem lies in the third row, because $|8|$ is not *larger* than the sum of the magnitudes of the other entries.

Note

The practice problem below suggests a TRICK(rearrangement of the system of equations) that sometimes works when a system is not strictly diagonally dominant.

Example 5

Show that the Gauss-Seidel method will produce a sequence converging to the solution of the following systems, provided the equations are arranged properly:

$$\begin{aligned} x_1 - 3x_2 + x_3 &= -2 \\ -6x_1 + 4x_2 + 11x_3 &= 1 \\ 5x_1 - 2x_2 - 2x_3 &= 9 \end{aligned}$$

Solution

The system is not strictly diagonally dominant, as for the 1st row

$$|coefficient\ of\ x_1| < |coefficient\ of\ x_2| + |coefficient\ of\ x_3|$$

$$or \quad |1| < |-3| + |1|$$

so neither Jacobi nor Gauss- Seidel is guaranteed to work. In fact, both iterative methods produce sequences that fail to converge, even though the system has the unique solution $x_1 = 3, x_2 = 2, x_3 = 1$. However, the equations can be rearranged as

$$\begin{aligned} 5x_1 - 2x_2 - 2x_3 &= 9 \\ x_1 - 3x_2 + x_3 &= -2 \\ -6x_1 + 4x_2 + 11x_3 &= 1 \end{aligned}$$

So,

for 1st equation (row);

$$|coefficient\ of\ x_1| > |coefficient\ of\ x_2| + |coefficient\ of\ x_3|$$

$$or \quad |5| > |-2| + |-2|$$

for 2nd equation(row);

$$|coefficient\ of\ x_2| > |coefficient\ of\ x_1| + |coefficient\ of\ x_3|$$

$$or \quad |-3| > |1| + |1|$$

for 3rd equation(row);

$$|coefficient\ of\ x_3| > |coefficient\ of\ x_1| + |coefficient\ of\ x_2|$$

$$or \quad |11| > |-6| + |4|$$

Now the coefficient matrix is strictly diagonally dominant, so we know Gauss-Seidel works with any initial vector. In fact, if $x^{(0)} = 0$, then $x^{(8)} = (2.9987, 1.9992, .9996)$.

Exercises

Solve the system in exercise 1 to 3 using Jacobi's method, with $x^{(0)} = 0$ and three iterations. Repeat the iterations until two successive approximations agree within a tolerance of .001 in each entry.

1. $\begin{aligned} 4x_1 + x_2 &= 7 \\ -x_1 + 5x_2 &= -7 \end{aligned}$

2. $\begin{aligned} 10x_1 - x_2 &= 25 \\ x_1 + 8x_2 &= 43 \end{aligned}$

3. $\begin{aligned} 3x_1 + x_2 &= 11 \\ -x_1 - 5x_2 + 2x_3 &= 15 \\ 3x_2 + 7x_3 &= 17 \end{aligned}$

4. $\begin{aligned} 50x_1 - x_2 &= 149 \\ x_1 - 100x_2 + 2x_3 &= -101 \\ 2x_2 + 50x_3 &= -98 \end{aligned}$

In exercises 5 to 8, use the Gauss Seidel method, with $x^{(0)} = 0$ and two iterations. Compare the number of iterations needed by Gauss Seidel and Jacobi to make two successive approximations agree within a tolerance of .001.

5. The system in exercise 1

6. The system in exercise 2

7. The system in exercise 3

8. The system in exercise 4

Determine which of the matrices in exercises 9 and 10 are strictly diagonally dominant.

9. (a) $\begin{bmatrix} 5 & 4 \\ 4 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 9 & -5 & 2 \\ 5 & -8 & -1 \\ -2 & 1 & 4 \end{bmatrix}$

10. (a) $\begin{bmatrix} 3 & -2 \\ 2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 3 & 1 \\ 3 & 6 & -4 \\ 1 & -4 & 7 \end{bmatrix}$

Show that the Gauss Seidel method will produce a sequence converging to the solution of the following system, provided the equations are arranged properly:

11. $\begin{aligned} x_1 - 3x_2 + x_3 &= -2 \\ -6x_1 + 4x_2 + 11x_3 &= 1 \\ 5x_1 - 2x_2 - 2x_3 &= 9 \end{aligned}$

12. $\begin{aligned} -x_1 + 4x_2 - x_3 &= 3 \\ 4x_1 - x_2 &= 10 \\ -x_2 + 4x_3 &= 6 \end{aligned}$

Lecture 17

Introduction to Determinant

In algebra, the **determinant** is a special number associated with any square matrix. As we have studied in earlier classes, that the determinant of 2 x 2 matrix is defined as the product of the entries on the main diagonal minus the product of the entries off the main diagonal. The determinant of a matrix A is denoted by det (***A***) or |A|

For example: $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Then det (***A***) = *ad-bc*.

or |A| = ad – bc

Example Find the determinant of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

$$|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1\times 4 - 2\times 3 = 4-6 = -2$$

To extend the definition of the det(***A***) to matrices of higher order, we will use subscripted entries for ***A***.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{vmatrix} = a_{11}b_{22} - a_{12}b_{21}$$

This is called a 2x2 determinant.

The determinant of a 3x3 matrix is also called a 3x3 determinant is defined by the following formula.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \qquad (1)$$

For finding the determinant of the 3x3 matrix, we look at the following diagram:

$$\left|\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{array}$$

We write 1st and 2nd columns again beside the determinant. The first arrow goes from a_{11} to a_{33}, which gives us product: $a_{11}a_{22}a_{33}$. The second arrow goes from a_{12} to a_{31}, which gives us product: $a_{12}a_{23}a_{31}$. The third arrow goes from a_{13} to a_{32}, which gives us the product: $a_{13}a_{21}a_{32}$. These values are taken with positive signs.

The same method is used for the next three arrows that go from right to left downwards, but these products are taken as negative signs.

$$= a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Example 2 Find the determinant of the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{bmatrix}$

$$detA = \begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} \begin{matrix} 1 & 2 \\ -4 & 5 \\ 7 & -8 \end{matrix}$$

$$= 1\times5\times9 + 2\times6\times7 + 3\times(-4)\times(-8) - 3\times5\times7 - 1\times6\times(-8) - 2\times(-4)\times9$$

$$= 45+84+96-105+48+72$$

$$= 240$$

We saw earlier that a 2×2 matrix is invertible if and only if its determinant is nonzero. In simple words, a matrix has its inverse if its determinant is nonzero. To extend this useful fact to larger matrices, we need a definition for the determinant of the $n\times n$ matrix. We can discover the definition for the 3×3 case by watching what happens when an invertible 3×3 matrix $\boldsymbol{A}$ is row reduced.

<u>Gauss' algorithm for evaluation of determinants</u>

1) Firstly, we apply it for 2×2 matrix say

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

$R_2' \to R_2 - 2R_1$ (Multiplying 1st row by 2 and then subtracting from 2nd row)

$$\sim \begin{bmatrix} 2 & 3 \\ 4-2(2) & 3-2(3) \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & -3 \end{bmatrix}$$

Now the determinant of this upper triangular matrix is the product of its entries on main diagonal that is

$Det(A) = 2(-3) - 0\times 3 \;\; = -6 - 0 \;\; = -6$

2) For 3×3 matrix say

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

By $R'_{12} \to R_{21}$ (Interchanging of 1st and 2nd rows)

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

$R'_2 \to R_2 - 2R_1$ (Multiplying 1st row by '-2' and then adding in the 2nd row)
$R'_3 \to R_3 + 2R_1$ (Multiplying 1st row by '2' and then adding in the 3rd row)

$$\sim \begin{bmatrix} -1 & 1 & 3 \\ 0 & 0 & -9 \\ 0 & 2 & 5 \end{bmatrix}$$

By $R'_{23} \to R_{32}$ (Interchanging of 2nd and 3rd rows)

$$\sim \begin{bmatrix} -1 & 1 & 3 \\ 0 & 2 & 5 \\ 0 & 0 & -9 \end{bmatrix}$$

Now the determinant of this upper triangular matrix is the product of its entries on main diagonal and that is

$Det(B) = (-1)\cdot 2 \cdot (-9) = 18$

So in general,

<u>For a 1×1 matrix</u>

say, $A = [a_{ij}]$ - we define $\det A = a_{11}$.

<u>For 2×2 matrix</u>

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

By $R'_2 \to R_2 - \left(\dfrac{a_{21}}{a_{11}}\right) R_1$ provided that $a_{11} \neq 0$

$$\sim \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \dfrac{a_{21}}{a_{11}} a_{12} \end{bmatrix}$$

$\therefore\ \Delta = \det A =$ product of the diagonal entries

$$= a_{11}\left(a_{22} - \frac{a_{21}}{a_{11}}a_{12}\right) = a_{11}a_{22} - a_{12}a_{21}$$

For 3×3 matrix say

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

By $R_2' \to R_2 - \left(\frac{a_{21}}{a_{11}}\right)R_1,\ R_3' \to R_3 - \left(\frac{a_{31}}{a_{11}}\right)R_1$ provided that $a_{11} \neq 0$

$$\sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} & \frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}} \\ 0 & \frac{a_{32}a_{11} - a_{12}a_{31}}{a_{11}} & \frac{a_{11}a_{33} - a_{13}a_{31}}{a_{11}} \end{bmatrix}$$

By $R_3' \to R_3 - \left(\frac{\frac{a_{32}a_{11} - a_{12}a_{31}}{a_{11}}}{\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}}}\right)R_2$ provided that $\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} \neq 0$

$$\sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} & \frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}} \\ 0 & 0 & \frac{a_{11}a_{33} - a_{13}a_{31}}{a_{11}} - \left(\frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}}\right)\left(\frac{\frac{a_{32}a_{11} - a_{12}a_{31}}{\cancel{a_{11}}}}{\frac{a_{22}a_{11} - a_{12}a_{21}}{\cancel{a_{11}}}}\right) \end{bmatrix} \because a_{11} \neq 0$$

Which is in echelon form.Now,

$\Delta = \det A =$ product of the diagonal entries

$$= \cancel{a}_{11}\left(\frac{a_{22}a_{11}-a_{12}a_{21}}{\cancel{a}_{11}}\right)\left(\frac{a_{11}a_{33}-a_{13}a_{31}}{a_{11}} - \left(\frac{a_{23}a_{11}-a_{13}a_{21}}{a_{11}}\right)\left(\frac{a_{32}a_{11}-a_{12}a_{31}}{a_{22}a_{11}-a_{12}a_{21}}\right)\right)$$
$$= \left(a_{22}a_{11}-a_{12}a_{21}\right)\left(\frac{a_{11}a_{33}-a_{13}a_{31}}{a_{11}}\right) - \left(a_{22}a_{11}-a_{12}a_{21}\right)\left(\frac{a_{23}a_{11}-a_{13}a_{21}}{a_{11}}\right)\left(\frac{a_{32}a_{11}-a_{12}a_{31}}{a_{22}a_{11}-a_{12}a_{21}}\right)$$
$$= \frac{1}{a_{11}}\left\{\left(a_{22}a_{11}-a_{12}a_{21}\right)\left(a_{11}a_{33}-a_{13}a_{31}\right)-\left(a_{23}a_{11}-a_{13}a_{21}\right)\left(a_{32}a_{11}-a_{12}a_{31}\right)\right\}$$
$$= \frac{1}{a_{11}}\left\{a_{11}^2a_{22}a_{33}-a_{11}a_{22}a_{13}a_{31}-a_{12}a_{21}a_{11}a_{33}+a_{12}a_{21}a_{13}a_{31}-a_{23}a_{11}^2a_{32}+a_{23}a_{11}a_{12}a_{31}+a_{13}a_{21}a_{32}a_{11}-a_{12}a_{21}a_{13}a_{31}\right\}$$
$$= \frac{1}{a_{11}}\left\{a_{11}^2a_{22}a_{33}-a_{11}a_{22}a_{13}a_{31}-a_{12}a_{21}a_{11}a_{33}-a_{23}a_{11}^2a_{32}+a_{23}a_{11}a_{12}a_{31}+a_{13}a_{21}a_{32}a_{11}\right\}$$
$$= \frac{a_{11}}{a_{11}}\left\{a_{11}a_{22}a_{33}-a_{22}a_{13}a_{31}-a_{12}a_{21}a_{33}-a_{23}a_{11}a_{32}+a_{23}a_{12}a_{31}+a_{13}a_{21}a_{32}\right\}$$
$$= a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}-a_{13}a_{22}a_{31}$$

□

Since ***A*** is invertible, Δ must be nonzero. *The converse is true as well*.

To generalize the definition of the determinant to larger matrices, we will use 2×2 determinants to rewrite the 3×3 determinant Δ described above. Since the terms in Δ can be grouped as:

$$\begin{aligned}\Delta &= (a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32})-(a_{12}a_{23}a_{31}-a_{12}a_{21}a_{33})+(a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}) \\ &= a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{12}(a_{21}a_{33}-a_{23}a_{31})+a_{13}(a_{21}a_{32}-a_{22}a_{31})\end{aligned},$$

$$\Delta = a_{11}\cdot\det\begin{bmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{bmatrix} - a_{12}\cdot\det\begin{bmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{bmatrix} + a_{13}\cdot\det\begin{bmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{bmatrix}$$

$$\Delta = a_{11}\cdot\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix} - a_{12}\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix} + a_{13}\cdot\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}$$

For brevity, we write $\Delta = a_{11}\cdot\det A_{11} - a_{12}\cdot\det A_{12} + a_{13}\cdot\det A_{13}$ (3)

$$\det(A_{11}) = \begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}, \quad \det(A_{12}) = \begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix} \quad \text{and} \quad \det(A_{13}) = \begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}$$

where

A*₁₁** is obtained from ***A by deleting the first row and first column.

A*₁₂** is obtained from ***A by deleting the first row and second column.

A*₁₃** is obtained from ***A by deleting the first row and third column.

So in general, for any square matrix ***A***, let $\boldsymbol{A_{ij}}$ denote the sub-matrix formed by deleting the *ith* row and *jth* column of ***A***.

Let's understand it with the help of an example.

Example3

Find the determinant of the matrix $A=\begin{bmatrix}1&4&3\\5&2&4\\3&6&3\end{bmatrix}$

Solution Given $A=\begin{bmatrix}1&4&3\\5&2&4\\3&6&3\end{bmatrix}$

$$
\begin{aligned}
|A| &= \begin{vmatrix}1&4&3\\5&2&4\\3&6&3\end{vmatrix}\\
&=1\begin{vmatrix}2&4\\6&3\end{vmatrix}-4\begin{vmatrix}5&4\\3&3\end{vmatrix}+3\begin{vmatrix}5&2\\3&6\end{vmatrix}\\
&=1(2\times3-4\times6)-4(5\times3-4\times3)+3(5\times6-2\times3)\\
&=1(6-24)-4(15-12)+3(30-6)\\
&=1(-18)-4(3)+3(24)\\
&=-18-12+72\\
&=42
\end{aligned}
$$

For instance, if $A=\begin{bmatrix}1&-2&5&0\\2&0&4&-1\\3&1&0&7\\0&4&-2&0\end{bmatrix}$

then $\boldsymbol{A_{32}}$ is obtained by crossing out row 3 and column 2,

$$\begin{bmatrix}1&-2&5&0\\2&0&4&-1\\3&1&0&7\\0&4&-2&0\end{bmatrix} \quad \text{so that} \quad A_{32}=\begin{bmatrix}1&5&0\\2&4&-1\\0&-2&0\end{bmatrix}$$

We can now give a recursive definition of a determinant.

When <u>$n = 3$,</u> det $\boldsymbol{A}$ is defined using determinants of the 2×2 submatrices A_{1j}.
When <u>$n = 4$</u>, det $\boldsymbol{A}$ uses determinants of the 3×3 submatrices A_{1j}
In general, an nxn determinant is defined by determinants of $(n-1)\times(n-1)$ sub matrices.

Definition

For $n \geq 2,$ the **determinant** of $n \times n$ matrix $A = [a_{ij}]$ is the sum of n terms of the form $\pm a_{1j} \times (\det A_{1j})$, with plus and minus signs alternating, where the entries $a_{11}, a_{12}, \cdots, a_{1n}$ are from the first row of ***A***.

Here for a_{ij},

$i = 1, 2, 3, \ldots, n \quad (1 \leq i \leq n)$

$j = 1, 2, 3, \ldots, n \quad (1 \leq j \leq n)$

In symbols, $\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + ... + (-1)^{1+n} a_{1n} \det A_{1n} = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$

Example 4

Compute the determinant of $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$

Solution

Here A is $n \times n = 3 \times 3$ matrix such that

$i = 1, 2, 3$

$j = 1, 2, 3$

$\because \det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$ and here $j = 1, 2, 3$

$$\therefore \det(A) = \sum_{j=1}^{3} (-1)^{1+j} a_{1j} \det A_{1j} = (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} + (-1)^{1+3} a_{13} \det A_{13}$$

$$= a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}$$

$$\det A = 1.\det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5.\det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0.\det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix}$$

$$\det A = 1.\begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5.\begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0.\begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

= 1 [4(0) – (-1)(-2)] -5 [2(0) – 0(-1)] +0[2(-2) – 4(0)]

$$= 1(0 - 2) - 5(0 - 0) + 0(-4 - 0) = -2$$

Minor of an element

If ***A*** is a square matrix, then the **Minor** of entry a_{ij} (called the ijth minor of ***A***) is denoted by M_{ij} and is defined to be the determinant of the sub matrix that remains when the *ith* row and *jth* column of ***A*** are deleted.

In the above example, Minors are as follows:

$$M_{11} = \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix}, \quad M_{12} = \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix}, \quad M_{13} = \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

Cofactor of an element

The number $C_{ij}=(-1)^{i+j}M_{ij}$ is called the cofactor of entry a_{ij}(or the *ijth* cofactor of A). When the + or – sign is attached to the Minor, then Minor becomes a cofactor.

In the above example, following are the Cofactors:

$$C_{11} = (-1)^{1+1} M_{11}, \qquad C_{12} = (-1)^{1+2} M_{12}\ , \qquad C_{13} = (-1)^{1+3} M_{13}$$

$$C_{11} = (-1)^{1+1} \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix}, \quad C_{12} = (-1)^{1+2} \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix}\ , \quad C_{13} = (-1)^{1+3} \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

Example 5 Find the minor and the cofactor of the matrix $A = \begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}$

Solution Here $A = \begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}$

The minor of entry a_{11} is

$$M_{11} = \begin{vmatrix} \cancel{3} & \cancel{1} & \cancel{-4} \\ \cancel{2} & 5 & 6 \\ \cancel{1} & 4 & 8 \end{vmatrix} = \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} = 5\times 8 - 6\times 4 = 40 - 24 = 16$$

and the corresponding cofactor is

$$C_{11} = (-1)^{1+1} M_{11} = M_{11} = 16$$

The minor of entry a_{32} is

$$\mathrm{M}_{32} = \begin{vmatrix} 3 & \cancel{1} & -4 \\ 2 & \cancel{5} & 6 \\ \cancel{1} & \cancel{4} & \cancel{8} \end{vmatrix} = \begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = 26$$

and the corresponding cofactor is

$$C_{32} = (-1)^{3+2} M_{32} = -M_{32} = -\begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = -26$$

Alternate Definition

Given $A = [a_{ij}]$, the (i, j)-cofactor of ***A*** is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij} \qquad (4)$$

Then $\det A = a_{11}C_{11} + a_{12}C_{12} + ... + a_{1n}C_{1n}$

This formula is called the *cofactor expansion across the first row* of ***A***.

Example 6 Expand a 3x3 determinant using cofactor concept $A = \begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix}$

Solution Using cofactor expansion along the first column;

$$\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} = (1)(-1)^{1+1}\begin{vmatrix} 5 & 6 \\ -8 & 9 \end{vmatrix} + (-4)(-1)^{2+1}\begin{vmatrix} 2 & 3 \\ -8 & 9 \end{vmatrix} + (7)(-1)^{3+1}\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

Now if we compare it with the formula (4),

$$= 1C_{11} + (-4)C_{21} + 7C_{31}$$

$$= (1)(-1)^2\begin{vmatrix} 5 & 6 \\ -8 & 9 \end{vmatrix} + (-4)(-1)^3\begin{vmatrix} 2 & 3 \\ -8 & 9 \end{vmatrix} + (7)(-1)^4\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

$$= (1)(1)\begin{vmatrix} 5 & 6 \\ -8 & 9 \end{vmatrix} + (-4)(-1)\begin{vmatrix} 2 & 3 \\ -8 & 9 \end{vmatrix} + (7)(1)\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

$$= 1\begin{vmatrix} 5 & 6 \\ -8 & 9 \end{vmatrix} + 4\begin{vmatrix} 2 & 3 \\ -8 & 9 \end{vmatrix} + 7\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

$$= 1(45 - (-48)) + 4(18 - (-24)) + 7(12 - 15)$$

$$= 1(45 + 48) + 4(18 + 24) + 7(12 - 15)$$

$$= (1)(93) + (4)(42) + (7)(-3) = 240$$

Using cofactor expansion along the second column,

$$\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} = (2)(-1)^{1+2}\begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(-1)^{2+2}\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1)^{3+2}\begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}$$

$$= (2)(-1)^{3}\begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(-1)^{4}\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1)^{5}\begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}$$

$$= (2)(-1)\begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(1)\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1)\begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}$$

$$= -2\begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + 5\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + 8\begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix}$$

$$= -2(-36-42) + 5(9-21) + 8(6-(-12))$$

$$= (-2)(-78) + (5)(-12) + (8)(18) = 240$$

Theorem 1 The determinant of an $n \times n$ matrix $\boldsymbol{A}$ can be computed by a cofactor expansion across any row or down any column. The expansion across the ith row using the cofactors in (4) is

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

The cofactor expansion down the jth column is

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

The plus or minus sign in the (i, j)-cofactor depends on the position of a_{ij} in the matrix, regardless of the sign of a_{ij} itself. The factor $(-1)^{i+j}$ determines the following checkerboard pattern of signs:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

Example 7 Use a cofactor expansion across the third row to compute det ***A***, where

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

Solution Compute $\det A = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33}$

$$= (-1)^{3+1} a_{31} \det A_{31} + (-1)^{3+2} a_{32} \det A_{32} + (-1)^{3+3} a_{33} \det A_{33}$$

$$= 0\begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix}$$

$$= 0 + 2(-1) + 0 = -2$$

Theorem 1 is helpful for computing the determinant of a matrix that contains many zeros. For example, if a row is mostly zeros, then the cofactor expansion across that row has many terms that are zero, and the cofactors in those terms need not be calculated.
The same approach works with a column that contains many zeros.

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

Solution $\det(A) = \begin{vmatrix} 2 & 0 & 0 & 5 \\ -1 & 2 & 4 & 1 \\ 3 & 0 & 0 & 3 \\ 8 & 6 & 0 & 0 \end{vmatrix}$

Expand from third column

$$\det(A) = 0\times C_{13} + 4\times C_{23} + 0\times C_{33} + 0\times C_{43}$$

$$= 0 \;+ 4\times C_{23} \;+ 0 + 0$$

$$= \; 4\times C_{23}$$

$$= 4\times(-1)^{2+3}\begin{vmatrix} 2 & 0 & 5 \\ 3 & 0 & 3 \\ 8 & 6 & 0 \end{vmatrix}$$

Expand from second column

$$= -4\left(0 \;+\; 0 \;+\; (-6)\begin{vmatrix} 2 & 5 \\ 3 & 3 \end{vmatrix}\right)$$

$$= (-4)\;(-6)\begin{vmatrix} 2 & 5 \\ 3 & 3 \end{vmatrix}$$

$$= -216$$

Example 9 Show that the value of the determinant is independent of θ

$$A = \begin{vmatrix} \sin\theta & \cos\theta & 0 \\ -\cos\theta & \sin\theta & 0 \\ \cos\theta - \sin\theta & \sin\theta + \cos\theta & 1 \end{vmatrix}$$

Solution Consider $A = \begin{vmatrix} \sin\theta & \cos\theta & 0 \\ -\cos\theta & \sin\theta & 0 \\ \cos\theta - \sin\theta & \sin\theta + \cos\theta & 1 \end{vmatrix}$

Expand the given determinant from 3rd column we have

$$= 0 - 0 + (-1)^{3+3}[\sin^2\theta + \cos^2\theta] = 1$$

Example 10 Compute det ***A***, where $A = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$

Solution The cofactor expansion down the first column of ***A*** has all terms equal to zero except the first.

Thus $$\det A = 3\begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} - 0.C_{21} + 0.C_{31} - 0.C_{41} + 0.C_{51}$$

Henceforth, we will omit the zero terms in the cofactor expansion.

Next, expand this 4×4 determinant down the first column, in order to take advantage of the zeros there.

We have $$\det A = 3 \times 2\begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}$$

This 3×3 determinant was computed above and found to equal –2.

Hence, det ***A*** = $3 \times 2 \times (-2) = -12$.

The matrix in this example was nearly triangular. The method in that example is easily adapted to prove the following theorem.

Triangular Matrix

A triangular matrix is a special kind of m x n matrix where the entries either below or above the main diagonal are zero.

$\begin{bmatrix} 1 & 4 & 2 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix}$ is upper triangular and 25×25 $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 8 & 0 \\ 4 & 9 & 7 \end{bmatrix}$ is lower triangular matrices.

Determinants of Triangular Matrices

Determinants of the triangular matrices are also easy to evaluate regardless of size.

Theorem If $\boldsymbol{A}$ is triangular matrix, then det ($\boldsymbol{A}$) is the product of the entries on the main diagonal.

Consider a 4×4 lower triangular matrix.

$$A=\begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

Keeping in mind that an elementary product must have exactly one factor from each row and one factor from each column, the only elementary product that does not have one of the six zeros as a factor is $(a_{11}a_{22}a_{33}a_{44})$. The column indices of this elementary product are in natural order, so the associated signed elementary product takes a +.

Thus, det (A)= $a_{11}\times a_{22}\times a_{33}\times a_{44}$

Example 11

$$\begin{vmatrix} -2 & 5 & 7 \\ 0 & 3 & 8 \\ 0 & 0 & 5 \end{vmatrix} = (-2)(3)(5) = -30$$

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 4 & 9 & 0 & 0 \\ -7 & 6 & -1 & 0 \\ 3 & 8 & -5 & -2 \end{vmatrix} = (1)(9)(-1)(-2) = 18$$

$$\begin{vmatrix} 1 & 2 & 7 & -3 \\ 0 & 1 & -4 & 1 \\ 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 3 \end{vmatrix} = (1)(1)(2)(3) = 6$$

The strategy in the above Example of looking for zeros works extremely well when an entire row or column consists of zeros. In such a case, the cofactor expansion along such a row or column is a sum of zeros. So, the determinant is zero. Unfortunately, most cofactor expansions are not so quickly evaluated.

Numerical Note By today’s standards, a 25×25 matrix is small. Yet it would be impossible to calculate a 25×25 determinant by cofactor expansion. In general, a cofactor expansion requires over *n!* multiplications, and $25! \sim 1.5x10^{25}$.

If a supercomputer could make one trillion multiplications per second, it would have to run for over 500,000 years to compute a 25×25 determinant by this method. Fortunately, there are faster methods, as we'll soon discover.

Example 12 Compute $\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix}$

Solution Take advantage of the zeros. Begin with a cofactor expansion down the third column to obtain a 3×3 matrix, which may be evaluated by an expansion down its first column,

$$\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix} = (-1)^{1+3}2\begin{vmatrix} 0 & 3 & -4 \\ -5 & -8 & 3 \\ 0 & 5 & -6 \end{vmatrix}$$

$$= 2\cdot(-1)^{2+1}(-5)\begin{vmatrix} 3 & -4 \\ 5 & -6 \end{vmatrix} = 20$$

The –1 in the next-to-last calculation came from the position of the –5 in the 3×3 determinant.

Exercises

Compute the determinants in exercises 1 to 6 by cofactor expansions. At each step, choose a row or column that involves the least amount of computation.

1. $\begin{vmatrix} 6 & 0 & 0 & 5 \\ 1 & 7 & 2 & -5 \\ 2 & 0 & 0 & 0 \\ 8 & 3 & 1 & 8 \end{vmatrix}$

2. $\begin{vmatrix} 1 & -2 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 2 & -6 & -7 & 5 \\ 5 & 0 & 4 & 4 \end{vmatrix}$

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

4. $\begin{vmatrix} 4 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ 2 & 6 & 3 & 0 \\ 5 & -8 & 4 & -3 \end{vmatrix}$

5. $\begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 0 & 2 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix}$

6. $\begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$

Use the method of Example 2 to compute the determinants in exercises 7 and 8. In exercises 9 to 11, compute the determinant of elementary matrix. In exercises 12 and 13, verify that det ***EA*** = (det ***E***) . (det ***A***), where ***E*** is the elementary matrix and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

7. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$

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

9. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$

10. $\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

11. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

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

14. Let $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$. Write 5***A***. Is det 5***A*** = 5 det ***A***?

15. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and k be a scalar. Find a formula that relates det (k***A***)to k and det ***A***.

Lecture 18

Properties of Determinants

In this lecture, we will study the properties of the determinants. Some of them have already been discussed and you will be familiar with these. These properties become helpful, while computing the values of the determinants. The secret of determinants lies in how they change when row or column operations are performed.

Theorem 3 (Row Operations)**:** Let ***A*** be a square matrix.

a. If a multiple of one row of *A* is added to another row, the resulting determinant will remain same.
b. If two rows of ***A*** are interchanged to produce ***B***, then *det* ***B*** = –*det* ***A***.
c. If one row of ***A*** is multiplied by *k* to produce ***B***, then *det* ***B*** = *k* . *det* ***A***.

The following examples show how to use Theorem 3 to find determinants efficiently.

a. If a multiple of one row of *A* is added to another row, the resulting determinant will remain same.

Example

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$Multiplying\ 2nd\ row\ by\ non-zero\ scalar\ say\ 'k'\ as$$

$$ka_{21}\quad ka_{22}\quad ka_{23} ----adding\ this\ in\ 1st\ row\ then\ 'A'\ becomes$$

$$=\begin{vmatrix} a_{11}+ka_{21} & a_{12}+ka_{22} & a_{13}+ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \qquad R_1' \to R_1 + kR_2$$

If each element of any row(column) can be expressed as sum of two elements then the resulting determinant can be expressed as sum of two determinants, so in this case

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}+\begin{vmatrix} ka_{21} & ka_{22} & ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}+k\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$ By using property (c) of above theorem 3.

If any two rows or columns in a determinant are identical then value of this determinant is zero. So in this case $R_1 \equiv R_2$

$$\therefore \quad \Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + k(0)$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = A$$

b. If two rows of *A* are interchanged to produce *B*, then *det B = –det A*.

Example 1

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 1 & 1 \\ 0 & 8 & 9 \end{bmatrix}$$

$$\text{Now, } \det A = \begin{vmatrix} 1 & 2 & 3 \\ 5 & 1 & 1 \\ 0 & 8 & 9 \end{vmatrix} = 1(9-8) - 2(45-0) + 3(40-0) = 1 - 90 + 120 = 31$$

Now interchange column 1st with 2nd we get a new matrix, $B = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 5 & 1 \\ 8 & 0 & 9 \end{bmatrix}$

$$\det B = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 5 & 1 \\ 8 & 0 & 9 \end{vmatrix} = 2(45-0) - 1(9-8) + 3(0-40) = 90 - 1 - 120 = -31$$

c. If one row of *A* is multiplied by *k* to produce *B*, then *det B = k . det A.*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 0 & 1 \\ 0 & 8 & 9 \end{bmatrix}$$

$$|A| = 1(0-8) - 2(45-0) + 3(40-0)$$

$$= -8 - 90 + 120 = 22$$

$$Multipling R_1 \; by \; k, we \; get \; say$$

$$B = \begin{bmatrix} 1k & 2k & 3k \\ 5 & 0 & 1 \\ 0 & 8 & 9 \end{bmatrix}$$

$$\left|B\right| = k(40-0) - 2k(45-0) + 3k(40-0)$$
$$= 40k - 90k + 120k = 22k$$
$$= k\left|A\right|$$

Example 2

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

Solution

$$\det A = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -7 \\ 0 & -2 & -8 & -10 \\ 0 & -7 & -10 & -13 \end{vmatrix} \quad by\ R_2{}' \to R_2 + (-2)R_1, R_3{}' \to R_3 + (-3)R_1, R_4{}' \to R_4 + (-4)R_1$$

$$= \begin{vmatrix} -1 & -2 & -7 \\ -2 & -8 & -10 \\ -7 & -10 & -13 \end{vmatrix} \quad \text{expanding from Ist column}$$

$$= (-1)(-2)(-1)\begin{vmatrix} 1 & 2 & 7 \\ 1 & 4 & 5 \\ 7 & 10 & 13 \end{vmatrix} \quad taking\ (-1), (-2)\ and\ \left(-1\right) common\ from 1st, 2nd, 3rd\ rows$$

$$= (-2)\begin{vmatrix} 1 & 2 & 7 \\ 0 & 2 & -2 \\ 0 & -4 & -36 \end{vmatrix} \quad by\ R_2{}' \to R_2 + (-1)R_1, R_3{}' \to R_3 + (-7)R_1$$

$$= (-2)\begin{vmatrix} 2 & -2 \\ -4 & -36 \end{vmatrix} \quad \text{expanding by 1st column}$$

$$= (-2)(2)(-4)\begin{vmatrix} 1 & -1 \\ 1 & 9 \end{vmatrix} \quad \text{taking 2 and (-4) common from 1st and 2nd rows respectively.}$$

$$= 16\begin{vmatrix} 1 & -1 \\ 0 & 10 \end{vmatrix} \quad by\ R_2 + (-1)R_1$$

$$= 160$$

Example 3 Evaluate the determinant of the matrix $A = \begin{bmatrix} 4 & 2 & 5 & 10 \\ 1 & 1 & 6 & 3 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{bmatrix}$

Solution

$$\det A = \begin{vmatrix} 4 & 2 & 5 & 10 \\ 1 & 1 & 6 & 3 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 1 & 6 & 3 \\ 4 & 2 & 5 & 10 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{vmatrix} \quad \text{interchanging } R_1 \text{ and } R_2 (R'_{12})$$

$$= -\begin{vmatrix} 1 & 1 & 6 & 3 \\ 0 & -2 & -19 & -2 \\ 0 & -4 & -42 & -16 \\ 0 & 2 & 5 & 8 \end{vmatrix} \quad By\ R_2' \to R_2 + (-4)R_1, R_3' \to R_3 + (-7)R_1$$

$$= -\begin{vmatrix} -2 & -19 & -2 \\ -4 & -42 & -16 \\ 2 & 5 & 8 \end{vmatrix} \quad \text{expanding from 1st column}$$

$$=(-1)^3 \begin{vmatrix} 2 & 19 & 2 \\ 4 & 42 & 16 \\ 2 & 5 & 8 \end{vmatrix} \quad \text{taking (-1) as a common factor from } R_1 \text{ and } R_2$$

$$= -\begin{vmatrix} 2 & 19 & 2 \\ 4 & 42 & 16 \\ 2 & 5 & 8 \end{vmatrix}$$

$$= -2\begin{vmatrix} 1 & 19 & 2 \\ 2 & 42 & 16 \\ 1 & 5 & 8 \end{vmatrix}$$

$$= (-2)\begin{vmatrix} 1 & 19 & 2 \\ 0 & 4 & 12 \\ 0 & -14 & 6 \end{vmatrix} \quad By\ R_2' \to R_2 + (-2)R_1, R_3' \to R_3 + (-1)R_1$$

$$= (-2)\begin{vmatrix} 1 & 19 & 2 \\ 0 & 4 & 12 \\ 0 & -14 & 6 \end{vmatrix} \quad R_2 + (-2)R_1, R_3 + (-1)R_1$$

$$= -2\begin{vmatrix} 4 & 12 \\ -14 & 6 \end{vmatrix} \text{ expand from Ist column}$$

$$= -2(24+168) = -384$$

Example 4 Without expansion, show that $\begin{vmatrix} x & a+x & b+c \\ x & b+x & c+a \\ x & c+x & a+b \end{vmatrix} = 0$

Solution

$$\begin{vmatrix} x & a+x & b+c \\ x & b+x & c+a \\ x & c+x & a+b \end{vmatrix}$$

$$= \begin{vmatrix} x & a+x-x & b+c \\ x & b+x-x & c+a \\ x & c+x-x & a+b \end{vmatrix} \quad By\ C_2' \to C_2 - C_1$$

$$= \begin{vmatrix} x & a & b+c \\ x & b & c+a \\ x & c & a+b \end{vmatrix}$$

$Taking\ 'x'\ common\ from\ C_1$

$$= x\begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix}$$

$$= x\begin{vmatrix} 1 & a+b+c & b+c \\ 1 & b+c+a & c+a \\ 1 & c+a+b & a+b \end{vmatrix} \quad By\ C_2' \to C_2 + C_3$$

Now taking (a+b+c) common form C_2

$$= x(a+b+c)\begin{vmatrix} 1 & 1 & b+c \\ 1 & 1 & c+a \\ 1 & 1 & a+b \end{vmatrix}$$

$= 0$ as column Ist and 2nd are identical ($C_1 \equiv C_2$). So its value will be zero.

Example 5 Evaluate $A = \begin{vmatrix} 2 & 3 & 1 & 0 & 1 \\ 1 & 1 & 3 & 1 & 2 \\ 2 & 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 1 & 2 \\ 4 & 1 & 1 & 0 & 0 \end{vmatrix}$

Solution Interchanging R_1 and R_2, we get

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

$R_2' \to R_2 - 2R_1, R_3' \to R_3 - 2R_1, R_4' \to R_4 - 3R_1, R_5' \to R_5 - 4R_1$

$$= -\begin{vmatrix} 1 & 1 & 3 & 1 & 2 \\ 0 & 1 & -5 & -2 & -3 \\ 0 & -1 & -4 & 1 & 0 \\ 0 & -1 & -8 & -2 & -4 \\ 0 & -3 & -11 & -4 & -8 \end{vmatrix}$$

expand from C_1

$$= -\begin{vmatrix} 1 & -5 & -2 & -3 \\ -1 & -4 & 1 & 0 \\ -1 & -8 & -2 & -4 \\ -3 & -11 & -4 & -8 \end{vmatrix}$$

$R_2' \to R_2 + R_1, R_3' \to R_3 + R_1, R_4' \to R_4 + 3R_1$

$$= -\begin{vmatrix} 1 & -5 & -2 & -3 \\ 0 & -9 & -1 & -3 \\ 0 & -13 & -4 & -7 \\ 0 & -26 & -10 & -17 \end{vmatrix}$$

expand from C_1

$$= -\begin{vmatrix} -9 & -1 & -3 \\ -13 & -4 & -7 \\ -26 & -10 & -17 \end{vmatrix}$$

taking (-1) common from Ist, 2nd and 3rd row

$$=\begin{vmatrix} 9 & 1 & 3 \\ 13 & 4 & 7 \\ 26 & 10 & 17 \end{vmatrix}$$

interchange Ist and 2nd Column(C'_{12})

$$=-\begin{vmatrix} 1 & 9 & 3 \\ 4 & 13 & 7 \\ 10 & 26 & 17 \end{vmatrix}$$

$$C'_2 \to C_2 - 9C_1, C'_3 \to C_3 - 3C_1$$

$$=-\begin{vmatrix} 1 & 0 & 0 \\ 4 & -23 & -5 \\ 10 & -64 & -13 \end{vmatrix}$$

expand from Ist row

$$=-\begin{vmatrix} -23 & -5 \\ -64 & -13 \end{vmatrix} = -(299-320) = 21$$

An Algorithm to evaluate the determinant

Algorithm means a sequence of a finite number of steps to get a desired result. The word Algorithm comes from the famous Muslim mathematician AL-Khwarizmi who invented the word algebra.

The step-by-step evaluation of det(***A***) of order ***n*** is obtained as follows:

Step 1: By an interchange of rows of ***A*** (and taking the resulting sign into account) bring a non zero entry to (1,1) the position (unless all the entries in the first column are zero in which case det ***A***=0).

Step 2: By adding suitable multiples of the first row to all the other rows, reduce the (n-1) entries, except (1,1) in the first column, to 0. Expand det(***A***) by its first column. Repeat this process or continue the following steps.

Step 3: Repeat step 1 and step 2 with the last remaining rows concentrating on the second column.

Step 4: Repeat step 1,step2 and step 3 with the remaining (n-2) rows, (n-3) rows and so on, until a triangular matrix is obtained.

Step5: Multiply all the diagonal entries of the resulting triangular matrix and then multiply it by its sign to get det(A)

Example 6 Compute det ***A***, where $A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.

Solution *The strategy is to reduce **A** to echelon form and then to use the fact that the determinant of a triangular matrix is the product of the diagonal entries.* The first two row replacements in column 1 do not change the determinant:

$$\det A = \begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} \quad By\ R_2' \to R_2 + 2R_1, R_3' \to R_1 + R_3$$

An interchange of rows 2 and 3 (R'_{23}), it reverses the sign of the determinant, so

$$\det A = -\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} = -(1)(3)(-5) = 15$$

Example 7 Compute det ***A***, where

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

Solution Taking'2' common from 1st row

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix}$$

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix} By\ R_2' \to R_2 - 3R_1, R_3' \to R_3 + 3R_1, R_4' \to R_4 - R_1$$

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (By\ R_4' \to R_4 - \tfrac{1}{2}R_3)$$

$$= 2.\{(1)(3)(-6)(1)\} = -36$$

Example 8 Show that $\begin{vmatrix} x & 2 & 2 & 2 \\ 2 & x & 2 & 2 \\ 2 & 2 & x & 2 \\ 2 & 2 & 2 & x \end{vmatrix} = (x+6)(x-2)^3$

Solution

$$\begin{vmatrix} x & 2 & 2 & 2 \\ 2 & x & 2 & 2 \\ 2 & 2 & x & 2 \\ 2 & 2 & 2 & x \end{vmatrix}$$

$$= \begin{vmatrix} x+6 & 2 & 2 & 2 \\ x+6 & x & 2 & 2 \\ x+6 & 2 & x & 2 \\ x+6 & 2 & 2 & x \end{vmatrix} \quad By\ \ C_1' \to C_1 + (C_2 + C_3 + C_4)$$

Taking (x+6) common from 1st column

$$= (x+6)\begin{vmatrix} 1 & 2 & 2 & 2 \\ 1 & x & 2 & 2 \\ 1 & 2 & x & 2 \\ 1 & 2 & 2 & x \end{vmatrix}$$

$$= (x+6)\begin{vmatrix} 1 & 2 & 2 & 2 \\ 0 & x-2 & 0 & 0 \\ 0 & 0 & x-2 & 0 \\ 0 & 0 & 0 & x-2 \end{vmatrix} \quad By\ R_2' \to R_2 - R_1, R_3' \to R_3 - R_1, R_4' \to R_4 - R_1$$

And this is the triangular matrix and its determinant is the prodcut of main diagonal's entries.

$$= (x+6)(x-2)^3$$

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

Solution $A=\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$

$$\det A=\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix} R_3' \to R_3+2R_1$$

$= \mathbf{0}$ as $R_2 \equiv R_3$

Example10 Compute det $\boldsymbol{A}$, where

$$A=\begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{bmatrix}$$

Solution

$$A=\begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{bmatrix}$$

$$=\begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{bmatrix} \quad R_4' \to R_4+R_2$$

$$=(-1)\begin{bmatrix} 2 & 1 & 2 & -1 \\ 0 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{bmatrix} \qquad By\ R_{12}'$$

Expanding from 1st row and 1st column

$$=-2\begin{vmatrix} 5 & -7 & 3 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix}$$

$$=(-2)\{5(6+6)-(-7)(3-0)+3(-9-0)\}$$

$$=54$$

Remarks

Suppose that a square matrix ***A*** has been reduced to an echelon form ***U*** by row replacements and row interchanges.

If there are r interchanges, then $\det(A)=(-1)^r\det(U)$

Furthermore, all of the pivots are still visible in ***U*** (because they have not been scaled to ones). If ***A*** is invertible, then the pivots in ***U*** are on the diagonal (since ***A*** is row equivalent to the identity matrix). In this case, *det* ***U*** is the product of the pivots. If ***A*** is not invertible, then ***U*** has a row of zero and *det* ***U*** *= 0*.

$$U=\begin{bmatrix} \bullet & \circ & \circ & \circ \\ 0 & \bullet & \circ & \circ \\ 0 & 0 & \bullet & \circ \\ 0 & 0 & 0 & \bullet \end{bmatrix} \qquad U=\begin{bmatrix} \bullet & \circ & \circ & \circ \\ 0 & \bullet & \circ & \circ \\ 0 & 0 & \bullet & \circ \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\det U\neq 0 \qquad\qquad \det U=0$$

Thus we have the following formula

$$\det A=\begin{cases} (-1)^r.(Product\ of\ \ pivots\ in\ U) & When\ A\ is\ invertible \\ 0 & When\ A\ is\ not\ invertible \end{cases} \qquad (1)$$

Example

Case-01 For 2×2 invertible matrix

Reducing given 2×2 invertible matrix into Echelon form as follows;

$$A=\begin{bmatrix} 4 & 5 \\ 3 & 2 \end{bmatrix}$$

By interchanging 1st and 2nd rows(R'_{12})

$$\sim\begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix} \because \text{ one replacement of rows has occurred, } \therefore r=1$$

$$\sim\begin{bmatrix} 3 & 2 \\ 0 & \frac{7}{3} \end{bmatrix} \text{ By } R_2'\to R_2-\frac{4}{3}R_1\text{, we have desired row-echelon form } U=\begin{bmatrix} 3 & 2 \\ 0 & \frac{7}{3} \end{bmatrix}.$$

Thus using the above formula as follows;

$$\det A = (-1)^r .(Product\ of\ \ pivots\ in\ U) = (-1)^1(3\cdot\frac{7}{3}) = -7$$

Case-02 For 2×2 non-invertible matrix
In this case say;

$$A = \begin{bmatrix} 4 & 5 \\ 8 & 10 \end{bmatrix}$$

$$\sim \begin{bmatrix} 4 & 5 \\ 0 & 0 \end{bmatrix} \quad By\ R_2' \to R_2 - 2R_1, \text{desired row-echelon form is } U = \begin{bmatrix} 4 & 5 \\ 0 & 0 \end{bmatrix}$$

Here no interchange of rows has occurred. So, $r = 0$ and
$\therefore \det A = (-1)^r .(Product\ of\ \ pivots\ in\ U) = (-1)^0(4\cdot 0) = 0$

Theorem 5 If A is an $n\times n$ matrix, then *det* $\boldsymbol{A^T}$ = *det* $\boldsymbol{A}$.

Example 11 If $A = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 3 \end{bmatrix}$, find det($\boldsymbol{A}$) and det ($\boldsymbol{A^T}$)

$$\det A = \begin{vmatrix} 1 & 4 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 3 \end{vmatrix} = 1(3-2) - 4(6-6) + 1(2-3) = 1 - 0 - 1 = 0$$

Now

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

$$\det A^t = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 1 & 1 \\ 1 & 2 & 3 \end{vmatrix} = 1(3-2) - 2(12-1) + 3(8-1) = 1 - 22 + 21 = 0$$

Remark
Column operations are useful for both theoretical purposes and hand computations. However, for simplicity we'll perform only row operations in numerical calculations.

Theorem 6 (Multiplicative Property)

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are $n\times n$ matrices, then $\det(AB) = (\det A)(\det B)$.

Example 12 Verify Theorem 6 for $A = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}$ *and* $B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$

Solution $AB = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 25 & 20 \\ 14 & 13 \end{bmatrix}$

and $\det AB = 25.13 - 20.14 = 325 - 280 = 45$

Since *det* ***A*** *= 9* and *det* ***B*** *= 5*, $(\det A)(\det B) = 9.5 = 45 = \det AB$

Remark

det (***A + B***) $\neq$ det ***A*** + det ***B***, in general.

For example**,**

If $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & -3 \\ -1 & 5 \end{bmatrix}$. Then

$A + B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \det(A + B) = 0$

$\det A + \det B = \begin{vmatrix} 2 & 3 \\ 1 & -5 \end{vmatrix} + \begin{vmatrix} -2 & -3 \\ -1 & 5 \end{vmatrix} = (-10 - 3) + (-10 - 3) = -26 \neq \det(A + B)$

Exercise

Find the determinants in exercises 1 to 6 by row reduction to echelon form.

1. $\begin{vmatrix} 1 & 3 & 0 & 2 \\ -2 & -5 & 7 & 4 \\ 3 & 5 & 2 & 1 \\ 1 & -1 & 2 & -3 \end{vmatrix}$

2. $\begin{vmatrix} 1 & 3 & 3 & -4 \\ 0 & 1 & 2 & -5 \\ 2 & 5 & 4 & -3 \\ -3 & -7 & -5 & 2 \end{vmatrix}$

3. $\begin{vmatrix} 1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 2 & 8 & 5 \\ 3 & -1 & -2 & 3 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 3 & -1 & 0 & -2 \\ 0 & 2 & -4 & -1 & -6 \\ -2 & -6 & 2 & 3 & 9 \\ 3 & 7 & -3 & 8 & -7 \\ 3 & 5 & 5 & 2 & 7 \end{vmatrix}$

5. $\begin{vmatrix} 1 & -2 & 3 & 1 \\ 5 & -9 & 6 & 3 \\ -1 & 2 & -6 & -2 \\ 2 & 8 & 6 & 1 \end{vmatrix}$

6. $\begin{vmatrix} 1 & 3 & 1 & 5 & 3 \\ -2 & -7 & 0 & -4 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{vmatrix}$

Combine the methods of row reduction and cofactor expansion to compute the determinants in exercises 7 and 8.

7. $\begin{vmatrix} 2 & 5 & -3 & -1 \\ 3 & 0 & 1 & -3 \\ -6 & 0 & -4 & 9 \\ 4 & 10 & -4 & -1 \end{vmatrix}$

8. $\begin{vmatrix} 2 & 5 & 4 & 1 \\ 4 & 7 & 6 & 2 \\ 6 & -2 & -4 & 0 \\ -6 & 7 & 7 & 0 \end{vmatrix}$

9. Use determinant to find out whether the matrix is invertible $\begin{bmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{bmatrix}$

10. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be 3×3 matrices, with det $\boldsymbol{A}$ = 4 and det $\boldsymbol{B}$ = -3. Use properties of determinants to compute

(a) det $\boldsymbol{AB}$ (b) det $7\boldsymbol{A}$ (c) det $\boldsymbol{B}^{\boldsymbol{T}}$ (d) det $\boldsymbol{A}^{\boldsymbol{T}}$
(e) det $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$

11 Show that

(a) $\begin{vmatrix} a_1 & b_1 & a_1+b_1+c_1 \\ a_2 & b_2 & a_2+b_2+c_2 \\ a_3 & b_3 & a_3+b_3+c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

(b) $\begin{vmatrix} a_1+b_1 & a_1-b_1 & c_1 \\ a_2+b_2 & a_2-b_2 & c_2 \\ a_3+b_3 & a_3-b_3 & c_3 \end{vmatrix} = -2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

12 Show that

(a) $\begin{vmatrix} a_1+b_1t & a_2+b_2t & a_3+b_3t \\ a_1t+b_1 & a_2t+b_2 & a_3t+b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = (1-t^2)\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$

(b) $\begin{vmatrix} a_1 & b_1+ta_1 & c_1+rb_1+sa_1 \\ a_2 & b_2+ta_2 & c_2+rb_2+sa_2 \\ a_3 & b_3+ta_3 & c_3+rb_3+sa_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$

13. Show that $\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = (y-x)(z-x)(z-y)$

Lecture 19

Cramer's Rule, Volume, and Linear Transformations

In this lecture, we shall apply the theory discussed in the last two lectures to obtain important theoretical formulae and a geometric interpretation of the determinant.

Cramer's Rule Cramer's rule is needed in a variety of theoretical calculations. For instance, it can be used to study how the solution of ***Ax* = *b*** is affected by changes in the entries of ***b***. However, the formula is inefficient for hand calculations, except for 2×2 or perhaps 3×3 matrices.

Theorem 1 (Crammer's Rule) Let ***A*** be an invertible $n\times n$ matrix. For any ***b*** in $\boldsymbol{R^n}$, the unique solution *x* of ***Ax* = *b*** has entries given by

$$x_i = \frac{\det A_i(b)}{\det A}, \qquad i = 1, 2, \ldots, n \tag{1}$$

Example 1 Use Cramer's rule to solve the system

$$\begin{aligned} 3x_1 - 2x_2 &= 6 \\ -5x_1 + 4x_2 &= 8 \end{aligned}$$

Solution Write the system in matrix form, ***Ax* = *b***

$$\begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$$

where

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \& b = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$$

$$\det A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix} = 12 - 10 = 2$$

$$A_1(b) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, A_2(b) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

Since det ***A*** = *2*, the system has a unique solution. By Cramer's rule,

$$x_1 = \frac{\det A_1(b)}{\det A} = \frac{24+16}{2} = 20$$

$$x_2 = \frac{\det A_2(b)}{\det A} = \frac{24+30}{2} = 27$$

Example 2 Consider the following system in which *s* is an unspecified parameter. Determine the values of *s* for which the system has a unique solution and use Cramer's rule to describe the solution.

$$\begin{aligned} 3sx_1 - 2x_2 &= 4 \\ -6x_1 + sx_2 &= 1 \end{aligned}$$

Solution Here

$$A = \begin{bmatrix} 3s & -2 \\ -6 & s \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad A_1(b) = \begin{bmatrix} 4 & -2 \\ 1 & s \end{bmatrix}, \qquad A_2(b) = \begin{bmatrix} 3s & 4 \\ -6 & 1 \end{bmatrix}$$

Since $\det A = 3s^2 - 12 = 3(s+2)(s-2)$

the system has a unique solution when

$\det A \neq 0$

$\Rightarrow 3(s+2)(s-2) \neq 0$

$\Rightarrow s^2 - 4 \neq 0$

$\Rightarrow s \neq \pm 2$

For such an *s*, the solution is (x_1, x_2), where

$$x_1 = \frac{\det A_1(b)}{\det A} = \frac{4s+2}{3(s+2)(s-2)}, \qquad s \neq \pm 2$$

$$x_2 = \frac{\det A_2(b)}{\det A} = \frac{3s+24}{3(s+2)(s-2)} = \frac{s+8}{(s+2)(s-2)}, \qquad s \neq \pm 2$$

Example 3 Solve, by Cramer's Rule, the system of equations

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 1 \\ x_1 + 2x_2 - x_3 &= 2 \\ 3x_1 + 2x_2 + 2x_3 &= 3 \end{aligned}$$

Solution Here $A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 2 & -1 \\ 3 & 2 & 2 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 2 & -1 \\ 3 & 2 & 2 \end{bmatrix}$

$$A_2 = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -1 \\ 3 & 3 & 2 \end{bmatrix}, \; A_3 = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 2 \\ 3 & 2 & 3 \end{bmatrix}$$

$$D = \det A = 2\cdot 6 - 1\cdot(-5) + 3(-4) = 5$$
$$D_1 = \det A_1 b = 1\cdot 6 + 2\cdot 8 + 3(-5) = 7$$
$$D_2 = \det A_2 b = 2\cdot(7) - 1\cdot(5) + 3(-3) = 0$$
$$D_3 = \det A_3 b = 2\cdot(2) + 1\cdot(-3) + 1(-4) = -3$$
$$So\ \ x_1 = \frac{D_1}{D} = \frac{7}{5},\quad x_2 = \frac{D_2}{D} = 0,\quad x_3 = \frac{D_3}{D} = -\frac{3}{5}$$

Example 4 Use Cramer's Rule to solve.

$$\begin{aligned} x_1 + \qquad 2x_3 &= 6 \\ -3x_1 + 4x_2 + 6x_3 &= 30 \\ -\ x_1 - 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solution

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{bmatrix}, b = \begin{bmatrix} 6 \\ 30 \\ 8 \end{bmatrix}$$

$$\therefore A_1 = \begin{bmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & -8 & 3 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{bmatrix}$$

Therefore,

$$x_1 = \frac{\det(A_1 b)}{\det(A)} = \frac{-40}{44} = \frac{-10}{11}, \qquad x_2 = \frac{\det(A_2 b)}{\det(A)} = \frac{72}{44} = \frac{18}{11}$$
$$x_3 = \frac{\det(A_3 b)}{\det(A)} = \frac{152}{44} = \frac{38}{11}$$

Note For any $n\times n$ matrix ***A*** and any ***b*** in $\boldsymbol{R^n}$, *let* $\boldsymbol{A_i(b)}$ be the matrix obtained from ***A*** by replacing ***i***th column by the vector ***b***.

$$A_i(b) = \begin{bmatrix} a_1 & \ldots & \underset{\substack{\uparrow \\ i\,th\,column}}{b} & \ldots & a_n \end{bmatrix}$$

Formula for $\mathbf{A^{-1}}$

Cramer's rule leads easily to a general formula for the inverse of $n\times n$ matrix ***A***. The ***j****th* column of $\boldsymbol{A^{-1}}$ is a vector ***x*** that satisfies $\boldsymbol{Ax = e_j}$
where $\boldsymbol{e_j}$ is the ***j****th* column of the identity matrix, and the ***i****th* entry of ***x*** is the (*i*, *j*)-entry of $\boldsymbol{A^{-1}}$**.** By Cramer's rule,

$$\left\{(i, j) - entry\, of\ A^{-1}\right\} = x_{ij} = \frac{\det A_i(e_j)}{\det A} \qquad (2)$$

Recall that A_{ji} denotes the submatrix of A formed by deleting row j and column i. A cofactor expansion down column i of $A_i(e_j)$ shows that

$$\det A_i(e_j) = (-1)^{i+j} \det A_{ji} = C_{ji} \qquad (3)$$

where C_{ji} is a cofactor of A.
By (2), the (i, j)-entry of A^{-1} is the cofactor C_{ji} divided by det A.
[Note that the subscripts on C_{ji} are the reverse of (i, j).] Thus

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \qquad (4)$$

The matrix of cofactors on the right side of (4) is called the **adjugate** (or **classical adjoint**) of A, denoted by adj A. (The term adjoint also has another meaning in advance texts on linear transformations.) The next theorem simply restates (4).

Theorem 2 (An Inverse Formula)

Let A be an invertible $n \times n$ matrix, then $A^{-1} = \dfrac{1}{\det A} adj\ A$

Example

For the matrrix say

$$A = \begin{bmatrix} 2 & 3 \\ -1 & 5 \end{bmatrix} \Rightarrow \det A = 10 - (-3) = 13$$

$\Rightarrow A^{-1}$ will also be a 2×2 matrix

As

$Aji =$ submatrix of A formed by deleting row j and column i

So in this case

$A_{11} =$ submatrix of A formed by deleting row ***1*** and column ***1*** $= [5]$

$A_{12} =$ submatrix of A formed by deleting row ***1*** and column ***2*** $= [-1]$

$A_{21} =$ submatrix of A formed by deleting row ***2*** and column ***1*** $= [3]$

$A_{22} =$ submatrix of A formed by deleting row ***2*** and column ***2*** $= [2]$

and

$$\det A_i(e_j) = (-1)^{i+j} \det(A_{ji}) = C_{ji}$$

where e_j is the jth column of identity matrix $I_{n\times n}$

So in this case

$$C_{11} = \det A_1(e_1) = (-1)^{1+1}\det A_{11} = (+1)\det[5] = 5$$
$$C_{12} = \det A_2(e_1) = (-1)^{1+2}\det A_{12} = (-1)\det[-1] = (-1)(-1) = 1$$
$$C_{21} = \det A_1(e_2) = (-1)^{2+1}\det A_{21} = (-1)\det[3] = -3$$
$$C_{22} = \det A_2(e_2) = (-1)^{2+2}\det A_{22} = (+1)\det[2] = 2$$

By Cramer's rule,

$$\left\{(i,j) - entry\, of\ A^{-1}\right\} = x_{ij} = \frac{\det A_i(e_j)}{\det A} = \frac{C_{ji}}{\det A}$$

So for the current matrix;

$$\left\{(1,1) - entry\, of\ A^{-1}\right\} = x_{11} = \frac{\det A_1(e_1)}{\det A} = \frac{C_{11}}{\det A} = \frac{5}{13}$$
$$\left\{(1,2) - entry\, of\ A^{-1}\right\} = x_{12} = \frac{\det A_1(e_2)}{\det A} = \frac{C_{21}}{\det A} = \frac{-3}{13}$$
$$\left\{(2,1) - entry\, of\ A^{-1}\right\} = x_{21} = \frac{\det A_2(e_1)}{\det A} = \frac{C_{12}}{\det A} = \frac{1}{13}$$
$$\left\{(2,2) - entry\, of\ A^{-1}\right\} = x_{22} = \frac{\det A_2(e_2)}{\det A} = \frac{C_{22}}{\det A} = \frac{2}{13}$$

Hence by using equation # 4, we get

$$A^{-1} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{C_{11}}{\det A} & \dfrac{C_{21}}{\det A} \\ \dfrac{C_{12}}{\det A} & \dfrac{C_{22}}{\det A} \end{bmatrix} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{5}{13} & \dfrac{-3}{13} \\ \dfrac{1}{13} & \dfrac{2}{13} \end{bmatrix}$$

Example 5 Find the inverse of the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$.

Solution The nine cofactors are

$$C_{11} = +\begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2,\quad C_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3,\quad C_{13} = +\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14,\quad C_{22} = +\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7,\quad C_{23} = -\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = +\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4,\quad C_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1,\quad C_{33} = +\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

The adjoint matrix is the transpose of the matrix of cofactors. [For instance, C_{12} goes in the ($2, 1$) position.] Thus

$$adjA = \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

We could compute det ***A*** directly, but the following computation provides a check on the calculations above and produces det ***A***:

$$(adjA).A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

Since (adj ***A***) ***A*** = 14 ***I***, Theorem 2 shows that det ***A*** = 14 and

$$A^{-1} = \frac{1}{14}\begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} = \begin{bmatrix} -2/14 & 14/14 & 4/14 \\ 3/14 & -7/14 & 1/14 \\ 5/14 & -7/14 & -3/14 \end{bmatrix}$$

Determinants as Area or Volume

In the next application, we verify the geometric interpretation of determinants and we assume here that the usual Euclidean concepts of length, area, and volume are already understood for $\boldsymbol{R^2}$ and $\boldsymbol{R^3}$.

Theorem 3 If ***A*** is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|\det A|$. If ***A*** is a 3×3 matrix, the volume of the parallelepiped determined by the columns of ***A*** is $|\det A|$.

Example 6 Calculate the area of the parallelogram determined by the points (-2, -2), (0, 3), (4, -1) and (6, 4).

Solution

Let A(-2,-2), B(0,3), C(4,-1) and D(6,4). Fixing one point say A(-2,-2) and find the adjacent lengths of parallelogram which are given by the column vectors as follows**;**

$$AB = \begin{bmatrix} 0-(-2) \\ 3-(-2) \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$AC = \begin{bmatrix} 4-(-2) \\ -1-(-2) \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

So the area of parallelogram ABCD determined by above column vectors

$$= \left|\det\begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}\right| = |2-30| = |-28| = 28$$

Now we translate the parallelogram ABCD to one having the origin as a vertex. For which we subtract the vertex (-2, -2) from each of the four vertices. The new parallelogram has the vertices say

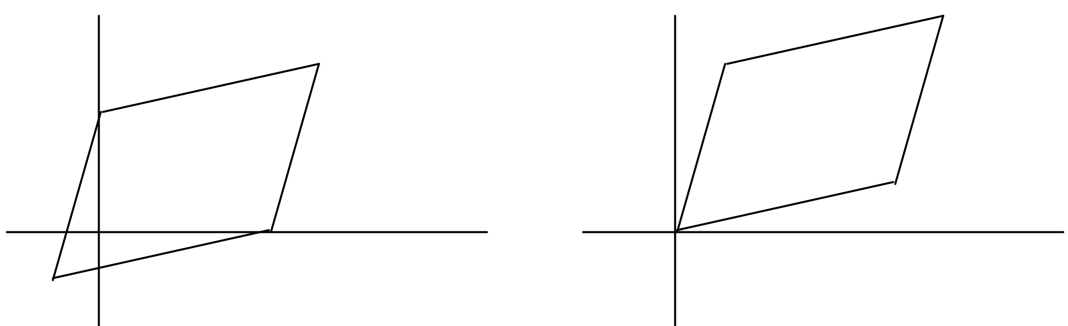
A′=(-2-(-2),-2-(-2))=(0, 0)
B′=(0-(-2),3-(-2))=(2, 5)
C′=(4-(-2),-1-(-2))=(6,1)
D′=(6-(-2),4-(-2))=(8, 6)

And fixing $A'(0,0)$ in this case, so

$$A'B' = \begin{bmatrix} 2-0 \\ 5-0 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$A'C' = \begin{bmatrix} 6-0 \\ 5-0 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \end{bmatrix}$$

See Fig below. The area of this parallelogram is also determined by the above columns vectors = $\left|\det\begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}\right| = |2-30| = |-28| = 28$



Translating a parallelogram does not change its area

Linear Transformations

Determinants can be used to describe an important geometric property of linear transformations in the plane and in $\boldsymbol{R^3}$. If ***T*** is a linear transformation and ***S*** is a set in the domain of ***T***, let ***T* (*S*)** denote the set of images of points in ***S***. We are interested in how the area (or volume) of ***T* (*S*)** compares with the area (or volume) of the original set ***S*.** For convenience, when ***S*** is a region bounded by a parallelogram, we also refer to ***S*** as a parallelogram.

Theorem 4 Let $T: R^2 \to R^2$ be the linear transformation determined by a 2×2 matrix ***A***. If ***S*** is a parallelogram in $\boldsymbol{R^2}$, then

{area of ***T (S)***} = |det***A***|. {area of ***S***}

If ***T*** is determined by a 3 x 3 matrix ***A***, and if ***S*** is a parallelepiped in $\boldsymbol{R^3}$, then

{volume of ***T (S)***} = |det***A***|. {volume of ***S***}

Example 7 Let *a* and *b* be positive numbers. Find the area of the region *E* bounded by the ellipse whose equation is $\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} = 1$.

Solution We claim that E is the image of the unit disk D under the linear transformation **A:D→E** determined by the matrix $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, given as

$A\boldsymbol{u} = \boldsymbol{x}$ where $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in D$, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in E$.

$$\Rightarrow \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Now $A\boldsymbol{u} = \boldsymbol{x} \Rightarrow \begin{bmatrix} au_1 \\ bu_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ then

$$\Rightarrow au_1 = x_1 \quad and \quad bu_2 = x_2$$

$$\Rightarrow u_1 = \frac{x_1}{a} \text{ and } u_2 = \frac{x_2}{b}$$

Since $u \in D$ (in the circular disk),it follows that the distance of $\boldsymbol{u}$ from origin will be less than unity i-e

$$\left(u_1^2 - 0\right) + \left(u_2^2 - 0\right) \le 1$$

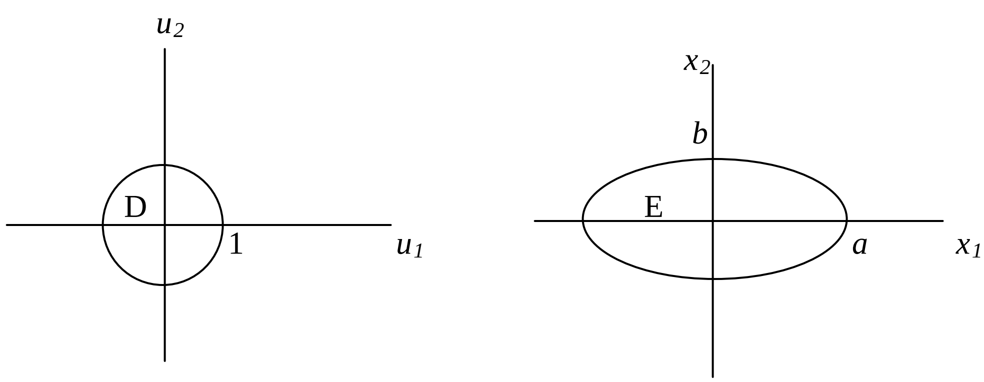
$$\Rightarrow \left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 \le 1 \qquad \because u_1 = \frac{x_1}{a}, u_2 = \frac{x_2}{b}$$

Hence by the generalization of theorem 4,

{area of ellipse} = {area of ***A(D)***} (*here* $T \equiv A$)

= |det ***A***|. {area of ***D***}

= $ab.\pi\,(1)^2 = \pi$ ab



Example 8 Let ***S*** be the parallelogram determined by the vectors $b_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, and let $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$. Compute the area of image of ***S*** under the mapping $x \to Ax$.

Solution The area of ***S*** is $\left|\det \begin{bmatrix} 1 & 5 \\ 3 & 1 \end{bmatrix}\right| = 14$, and det ***A*** = 2. By theorem 4, the area of image of ***S*** under the mapping $x \to Ax$ is |det ***A***|. {area of ***S***} = 2.14 = 28

Exercises

Use Cramer's Rule to compute the solutions of the systems in exercises 1 and 2.

1. $\begin{aligned} 2x_1 + x_2 \quad &= 7 \\ -3x_1 + \quad x_3 &= -8 \\ x_2 + 2x_3 &= -3 \end{aligned}$

2. $\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ -x_1 + \quad 2x_3 &= 2 \\ 3x_1 + x_2 + 3x_3 &= -2 \end{aligned}$

In exercises 3-6, determine the values of the parameter s for which the system has a unique solution, and describe the solution.

3. $\begin{aligned} 6sx_1 + 4x_2 &= 5 \\ 9x_1 + 2sx_2 &= -2 \end{aligned}$

4. $\begin{aligned} 3sx_1 - 5x_2 &= 3 \\ 9x_1 + 5sx_2 &= 2 \end{aligned}$

5. $\begin{aligned} sx_1 - 2sx_2 &= -1 \\ 3x_1 + 6sx_2 &= 4 \end{aligned}$

6. $\begin{aligned} 2sx_1 + \quad x_2 &= 1 \\ 3sx_1 + 6sx_2 &= 2 \end{aligned}$

In exercises 7 and 8, compute the adjoint of the given matrix, and then find the inverse of the matrix.

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

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

In exercises 9 and 10, find the area of the parallelogram whose vertices are listed.

9. *(0, 0), (5, 2), (6, 4), (11, 6)*

10. *(-1, 0), (0, 5), (1, -4), (2, 1)*

11. Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices at *(1, 0, -2), (1, 2, 4), (7, 1, 0)*.

12. Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices at *(1, 4, 0), (-2, -5, 2), (-1, 2, -1)*.

13. Let S be the parallelogram determined by the vectors $\boldsymbol{b_1} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ and $\boldsymbol{b_2} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$, and let $A = \begin{bmatrix} 6 & -2 \\ -3 & 2 \end{bmatrix}$. Compute the area of the image of S under the mapping $\boldsymbol{x} \to A\boldsymbol{x}$.

14. Let ***S*** be the parallelogram determined by the vectors $\boldsymbol{b_1} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$ and $\boldsymbol{b_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and let $\boldsymbol{A} = \begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix}$. Compute the area of the image of ***S*** under the mapping $\boldsymbol{x} \to \boldsymbol{Ax}$.

15. Let $\boldsymbol{T}: \boldsymbol{R^3} \to \boldsymbol{R^3}$ be the linear transformation determined by the matrix $\boldsymbol{A} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, where a, b, c are positive numbers. Let ***S*** be the unit ball, whose bounding surface has the equation $x_1^2 + x_2^2 + x_3^2 = 1$.

a. Show that ***T*** (***S***) is bounded by the ellipsoid with the equation $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1$.

b. Use the fact that the volume of the unit ball is $4\pi / 3$ to determine the volume of the region bounded by the ellipsoid in part (a).

Lecture 20

Vector Spaces and Subspaces

Case Example

The space shuttle's control systems are absolutely critical for flight. Because the shuttle is an unstable airframe, it requires constant computer monitoring during atmospheric flight. The flight control system sends a stream of commands to aerodynamic control surfaces.

Mathematically, the input and output signals to an engineering system are functions. It is important in applications that these functions can be added, and multiplied by scalars. These two operations on functions have algebraic properties that are completely analogous to the operation of adding vectors in $\boldsymbol{R^n}$ and multiplying a vector by a scalar, as we shall see in the lectures 20 and 27. For this reason, the set of all possible inputs (functions) is called a *vector space.* The mathematical foundation for systems engineering rests on vector spaces of functions, and we need to extend the theory of vectors in $\boldsymbol{R^n}$ to include such functions. Later on, we will see how other vector spaces arise in engineering, physics, and statistics.

Definition Let $\boldsymbol{V}$ be an arbitrary nonempty set of objects on which two operations are defined, addition and multiplication by scalars (numbers). If the following axioms are satisfied by all objects $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ in $\boldsymbol{V}$ and all scalars k and l, then we call $\boldsymbol{V}$ a vector space.

Axioms of Vector Space

1. Closure Property For any two vectors $\boldsymbol{u}$ & $\boldsymbol{v} \in \boldsymbol{V}$, implies $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{V}$

2. Commutative Property For any two vectors $\boldsymbol{u}$ & $\boldsymbol{v} \in \boldsymbol{V}$, implies $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

3. Associative Property For any three vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{V}$, $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$

4. Additive Identity For any vector $\boldsymbol{u} \in \boldsymbol{V}$, there exist a zero vector $\boldsymbol{0}$ such that

$$\boldsymbol{0} + \boldsymbol{u} = \boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$$

5. Additive Inverse For each vector $\boldsymbol{u} \in \boldsymbol{V}$, there exist a vector $-\boldsymbol{u}$ in $\boldsymbol{V}$ such that

$$-\boldsymbol{u} + \boldsymbol{u} = \boldsymbol{0} = \boldsymbol{u} + (-\boldsymbol{u})$$

6. Scalar Multiplication For any scalar k and a vector $\boldsymbol{u} \in \boldsymbol{V}$ implies k $\boldsymbol{u} \in \boldsymbol{V}$

7. Distributive Law For any scalar k if $\boldsymbol{u}$ & $\boldsymbol{v} \in \boldsymbol{V}$, then $k\,(\boldsymbol{u} + \boldsymbol{v}) = k\,\boldsymbol{u} + k\,\boldsymbol{v}$

8. For scalars m, n and for any vector $\boldsymbol{u} \in \boldsymbol{V}$, $(m + n)\,\boldsymbol{u} = m\,\boldsymbol{u} + n\,\boldsymbol{u}$

9. For scalars m, n and for any vector $\boldsymbol{u} \in \boldsymbol{V}$, $m\,(n\,\boldsymbol{u}) = (m\,n)\,\boldsymbol{u} = n\,(m\,\boldsymbol{u})$

10. For any vector $\boldsymbol{u} \in \boldsymbol{V}$, $1\boldsymbol{u} = \boldsymbol{u}$ where 1 is the multiplicative identity of real numbers.

Examples of vector spaces The following examples will specify a non empty set $\boldsymbol{V}$ and two operations: addition and scalar multiplication; then we shall verify that the ten vector space axioms are satisfied.

Example 1 Show that the set of all ordered n-tuple $\boldsymbol{R^n}$ is a vector space under the standard operations of addition and scalar multiplication.

Solution

(i) **Closure Property:**

Suppose that $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \boldsymbol{R^n}$

Then by definition, $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)$

$= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \in \boldsymbol{R^n}$ (By closure property)

Therefore, $\boldsymbol{R^n}$ is closed under addition.

(ii) **Commutative Property**

Suppose that $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \boldsymbol{R^n}$

Now $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)$

$= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$ (By closure property)

$= (v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n)$ (By commutative law of real numbers)

$= (v_1, v_2, \ldots, v_n) + (u_1, u_2, \ldots, u_n)$ (By closure property)

$= \boldsymbol{v} + \boldsymbol{u}$

Therefore, $\boldsymbol{R^n}$ is commutative under addition.

(iii) **Associative Property**

Suppose that $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ *and* $\boldsymbol{w} = (w_1, w_2, \ldots, w_n) \in \boldsymbol{R^n}$

Now $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = [(u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)] + (w_1, w_2, \ldots, w_n)$

$= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) + (w_1, w_2, \ldots, w_n)$ (By closure property)

$= ((u_1 + v_1) + w_1, (u_2 + v_2) + w_2, \ldots, (u_n + v_n) + w_n))$ (By closure property)

$= (u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), \ldots, u_n + (v_n + w_n))$ (By associative law of real numbers)

$= (u_1, u_2, \ldots, u_n) + (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$ (By closure property)

$= (u_1, u_2, \ldots, u_n) + [(v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n)]$ (By closure property)

$= \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$

Hence $\boldsymbol{R^n}$ is associative under addition.

(iv) **Additive Identity**

Suppose $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R^n}$. There exists $\boldsymbol{0} = (0, 0, \ldots, 0) \in \boldsymbol{R^n}$ such that
$\boldsymbol{0} + \boldsymbol{u} = (0, 0, \ldots, 0) + (u_1, u_2, \ldots, u_n)$

$= (0 + u_1, 0 + u_2, \ldots, 0 + u_n)$ (By closure property)

$= (u_1, u_2, \ldots, u_n) = \boldsymbol{u}$ (Existence of identity of real numbers)

Similarly, $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$

Hence $\boldsymbol{0} = (0, 0, \ldots, 0)$ is the additive identity for $\boldsymbol{R^n}$.

(v) **Additive Inverse**

Suppose $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R^n}$. There exists $-\boldsymbol{u} = (-u_1, -u_2, \ldots, -u_n) \in \boldsymbol{R^n}$

Such that $\boldsymbol{u} + (-\boldsymbol{u}) = (u_1, u_2, \ldots, u_n) + (-u_1, -u_2, \ldots, -u_n)$

$= (u_1 + (-u_1), u_2 + (-u_2), \ldots, u_n + (-u_n))$ (By closure property)

$= (0, 0, \ldots, 0) = \boldsymbol{0}$

Similarly, $(-\boldsymbol{u}) + \boldsymbol{u} = \boldsymbol{0}$

Hence the inverse of each element of $\boldsymbol{R^n}$ exists in $\boldsymbol{R^n}$.

(vi) **Scalar Multiplication**

If k is any scalar and $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R^n}$.

Then by definition, $k\,\boldsymbol{u} = k\,(u_1, u_2, \ldots, u_n) = (k\,u_1, k\,u_2, \ldots, k\,u_n) \in \boldsymbol{R^n}$
(By closure property)

(vii) **Distributive Law**

Suppose k is any scalar and $\boldsymbol{u} = (u_1, u_2, \ldots, u_n), \boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \boldsymbol{R^n}$

Now $k\,(\boldsymbol{u} + \boldsymbol{v}) = k\,[(u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)]$

$= k\,(u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$ (By closure property)

$= (k\,(u_1 + v_1), k\,(u_2 + v_2), \ldots, k\,(u_n + v_n))$ (By scalar multiplication)

$= (k\,u_1 + k\,v_1, k\,u_2 + k\,v_2, \ldots, k\,u_n + k\,v_n)$ (By Distributive Law)

$= (k\,u_1, k\,u_2, \ldots, k\,u_n) + (k\,v_1, k\,v_2, \ldots, k\,v_n)$ (By closure property)

$= k\,(u_1, u_2, \ldots, u_n) + k\,(v_1, v_2, \ldots, v_n)$ (By scalar multiplication)

$= k\,\boldsymbol{u} + k\,\boldsymbol{v}$

(viii) Suppose k and l be any scalars and $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R^n}$

Then $(k + l)\,\boldsymbol{u} = (k + l)\,(u_1, u_2, \ldots, u_n)$

$= ((k + l)u_1, (k + l)u_2, \ldots, (k + l)u_n)$ (By scalar multiplication)

$= (k\,u_1 + l\,u_1, k\,u_2 + l\,u_2, \ldots, k\,u_n + l\,u_n)$ (By Distributive Law)

$= (k\,u_1, k\,u_2, \ldots, k\,u_n) + (l\,u_1, l\,u_2, \ldots, l\,u_n)$ (By closure property)

$= k\,(u_1, u_2, \ldots, u_n) + l\,(u_1, u_2, \ldots, u_n)$ (By scalar multiplication)

$= k\,\boldsymbol{u} + l\,\boldsymbol{u}$

(ix) Suppose k and l be any scalars and $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R^n}$

Then $k\,(l\,\boldsymbol{u}) = k\,[l\,(u_1, u_2, \ldots, u_n)]$

$= k\,(l\,u_1, l\,u_2, \ldots, l\,u_n)$ (By scalar multiplication)

$= (k\,(l\,u_1), k\,(l\,u_2), \ldots, k\,(l\,u_n))$ (By scalar multiplication)

$= ((k\,l)u_1, (k\,l)u_2, \ldots, (k\,l)u_n)$ (By associative law)

$= (k\,l)\,(u_1, u_2, \ldots, u_n)$ (By scalar multiplication)

$= (k\,l)\,\boldsymbol{u}$

(x) Suppose $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{R^n}$

Then $1\,\boldsymbol{u} = 1\,(u_1, u_2, \ldots, u_n)$

$= (1u_1, 1u_2, \ldots, 1u_n)$ (By scalar multiplication)

$= (u_1, u_2, \ldots, u_n) = \boldsymbol{u}$ (Existence of identity in scalrs)

Hence, $\boldsymbol{R^n}$ is the real vector space with the standard operations of addition and scalar multiplication.

Note The three most important special cases of $\boldsymbol{R^n}$ are $\boldsymbol{R}$ (the real numbers), $\boldsymbol{R^2}$ (the vectors in the plane), and $\boldsymbol{R^3}$ (the vectors in 3-space).

Example 2 Show that the set $\boldsymbol{V}$ of all 2x2 matrices with real entries is a vector space if vector addition is defined to be matrix addition and vector scalar multiplication is defined to be matrix scalar multiplication.

Solution Suppose that $\boldsymbol{u} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \in \boldsymbol{V}$

and k and l be two any scalars.

(i) **Closure property** To prove axiom (i), we must show that $\boldsymbol{u} + \boldsymbol{v}$ is an object in $\boldsymbol{V}$: that is , we must show that $\boldsymbol{u} + \boldsymbol{v}$ is a 2x2 matrix. But this is clear from the definition of matrix addition, since $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix}$

(By closure property)

(ii) **Commutative property** Now it is very easy to verify the Axiom (ii)

$$\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix} \quad \text{(By closure property)}$$

$$= \begin{bmatrix} v_{11}+u_{11} & v_{12}+u_{12} \\ v_{21}+u_{21} & v_{22}+u_{22} \end{bmatrix} \quad \text{(Commutative property of real numbers)}$$

$$= \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \boldsymbol{v} + \boldsymbol{u}$$

(iii) **Associative property** $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \left(\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \right) + \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}$

$$= \begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix} + \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \quad \text{(By closure property)}$$

$$= \begin{bmatrix} (u_{11}+v_{11})+w_{11} & (u_{12}+v_{12})+w_{12} \\ (u_{21}+v_{21})+w_{21} & (u_{22}+v_{22})+w_{22} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11}+(v_{11}+w_{11}) & u_{12}+(v_{12}+w_{12}) \\ u_{21}+(v_{21}+w_{21}) & u_{22}+(v_{22}+w_{22}) \end{bmatrix} \quad \text{(By associative property of real numbers)}$$

$$= \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11}+w_{11} & v_{12}+w_{12} \\ v_{21}+w_{21} & v_{22}+w_{22} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \left(\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} + \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \right) = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$$

Therefore, ***V*** is associative under '+'.

(iv) **Additive Identity** Now to prove the axiom (iv), we must find an object ***0*** in ***V*** such that ***0*** + ***v*** = ***v*** + ***0*** = ***v*** for all ***u*** in ***V***. This can be done by defining $\boldsymbol{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

$$\boldsymbol{0} + \boldsymbol{u} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} 0+u_{11} & 0+u_{12} \\ 0+u_{21} & 0+u_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \boldsymbol{u}$$

and similarly ***u*** + ***0*** = ***u***.

(v) **Additive Inverse** Now to prove the axiom (v) we must show that each object ***u*** in ***V*** has a negative –***u*** such that ***u*** + (-***u***) = **0** = (-***u***) + **0**. Defining the negative of ***u*** to be

$$-\boldsymbol{u} = \begin{bmatrix} -u_{11} & -u_{12} \\ -u_{21} & -u_{22} \end{bmatrix}.$$

$$\boldsymbol{u} + (-\boldsymbol{u}) = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} -u_{11} & -u_{12} \\ -u_{21} & -u_{22} \end{bmatrix} = \begin{bmatrix} u_{11}+(-u_{11}) & u_{12}+(-u_{12}) \\ u_{21}+(-u_{21}) & u_{22}+(-u_{22}) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \boldsymbol{0}$$

Similarly, (-***u***) + ***u*** = ***0***

(vi) **Scalar Multiplication**

Axiom (vi) also holds because for any real number *k* we have

$$k\boldsymbol{u} = k \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix} \qquad \text{(By closure property)}$$

so that k ***u*** is a 2x2 matrix and consequently is an object in ***V***.

(vii) **Distributive Law**

$$k(\boldsymbol{u} + \boldsymbol{v}) = k\left(\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \right)$$

$$= k\begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix} = \begin{bmatrix} k(u_{11}+v_{11}) & k(u_{12}+v_{12}) \\ k(u_{21}+v_{21}) & k(u_{22}+v_{22}) \end{bmatrix}$$

$$= \begin{bmatrix} ku_{11}+kv_{11} & ku_{12}+kv_{12} \\ ku_{21}+kv_{21} & ku_{22}+kv_{22} \end{bmatrix} = \begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix} + \begin{bmatrix} kv_{11} & kv_{12} \\ kv_{21} & kv_{22} \end{bmatrix}$$

$$= k\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + k\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = k\boldsymbol{u} + k\boldsymbol{v}$$

(viii) $(k+l)\boldsymbol{u} = (k+l)\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} (k+l)u_{11} & (k+l)u_{12} \\ (k+l)u_{21} & (k+l)u_{22} \end{bmatrix}$

$$= \begin{bmatrix} ku_{11}+lu_{11} & ku_{12}+lu_{12} \\ ku_{21}+lu_{21} & ku_{22}+lu_{22} \end{bmatrix} = \begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix} + \begin{bmatrix} lu_{11} & lu_{12} \\ lu_{21} & lu_{22} \end{bmatrix}$$

$$= k\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + l\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = k\boldsymbol{u} + l\boldsymbol{u}$$

(ix) $k(l\boldsymbol{u}) = k\left(l\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}\right) = k\begin{bmatrix} lu_{11} & lu_{12} \\ lu_{21} & lu_{22} \end{bmatrix}$

$$= \begin{bmatrix} k(lu_{11}) & k(lu_{12}) \\ k(lu_{21}) & k(lu_{22}) \end{bmatrix} = \begin{bmatrix} (kl)u_{11} & (kl)u_{12} \\ (kl)u_{21} & (kl)u_{22} \end{bmatrix} = (kl)\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = (kl)\boldsymbol{u}$$

(x) Finally axiom (x) is a simple computation

$$1\boldsymbol{u} = 1\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} 1u_{11} & 1u_{12} \\ 1u_{21} & 1u_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \boldsymbol{u}$$

Hence the set of all 2x2 matrices with real entries is vector space under matrix addition and matrix scalar multiplication.

Note Example 2 is a special case of a more general class of vector spaces. The arguments in that example can be adapted to show that a set $\boldsymbol{V}$ of all $\boldsymbol{m}\times\boldsymbol{n}$ matrices with real entries, together with the operations of matrix addition and scalar multiplication, is a vector space.

Example 3 Let $\boldsymbol{V}$ be the set of all real-valued functions defined on the entire real line $(-\infty, \infty)$. If $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{V}$, then $\boldsymbol{f} + \boldsymbol{g}$ is a function defined by

$$(\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x}), \text{ for all } x \in \boldsymbol{R}.$$

The product of a scalar $a \in R$ and a function $\boldsymbol{f}$ in $\boldsymbol{V}$ is defined by

$$(a\,\boldsymbol{f})(\boldsymbol{x}) = a\,\boldsymbol{f}(\boldsymbol{x}), \text{ for all } \boldsymbol{x} \in \boldsymbol{R}.$$

Solution

(i) **Closure Property** If $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{V}$, then by definition
$(\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x}) \in \boldsymbol{V}$. Therefore, $\boldsymbol{V}$ is closed under addition.

(ii) **Commutative Property** If $\boldsymbol{f}$ and $\boldsymbol{g}$ are in $\boldsymbol{V}$, then for all $\boldsymbol{x} \in \boldsymbol{R}$

$(\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})$ (By definition)

$= \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x})$ (By commutative property)

$= (\boldsymbol{g} + \boldsymbol{f})(\boldsymbol{x})$ (By definition)

So that $\boldsymbol{f} + \boldsymbol{g} = \boldsymbol{g} + \boldsymbol{f}$

(iii) **Associative Property** If $\boldsymbol{f}$, $\boldsymbol{g}$ and $\boldsymbol{h}$ are in $\boldsymbol{V}$, then for all $\boldsymbol{x} \in \boldsymbol{R}$

$((\boldsymbol{f} + \boldsymbol{g}) + \boldsymbol{h})(\boldsymbol{x}) = (\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{x}) + \boldsymbol{h}(\boldsymbol{x})$ (By definition)

$= (\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})) + \boldsymbol{h}(\boldsymbol{x})$ (By definition)

$= \boldsymbol{f}(\boldsymbol{x}) + (\boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{h}(\boldsymbol{x}))$ (By associative property)

$= \boldsymbol{f}(\boldsymbol{x}) + (\boldsymbol{g} + \boldsymbol{h})(\boldsymbol{x})$ (By definition)

$= (\boldsymbol{f} + (\boldsymbol{g} + \boldsymbol{h}))(\boldsymbol{x})$

And so $(\boldsymbol{f} + \boldsymbol{g}) + \boldsymbol{h} = \boldsymbol{f} + (\boldsymbol{g} + \boldsymbol{h})$

(iv) **Additive Identity** The additive identity of $\boldsymbol{V}$ is the zero function defined by
$\boldsymbol{0}(\boldsymbol{x}) = \boldsymbol{0}$, for all $\boldsymbol{x} \in \boldsymbol{R}$ because $(\boldsymbol{0} + \boldsymbol{f})(\boldsymbol{x}) = \boldsymbol{0}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x})$ (By definition)

$= \boldsymbol{0} + \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$ (Existence of identity)

i.e. $\boldsymbol{0} + \boldsymbol{f} = \boldsymbol{f}$. Similarly, $\boldsymbol{f} + \boldsymbol{0} = \boldsymbol{f}$.

(v) **Additive Inverse** The additive inverse of a function $\boldsymbol{f}$ in $\boldsymbol{V}$ is $(-1)\,\boldsymbol{f} = -\boldsymbol{f} \in \boldsymbol{V}$ because

$(\boldsymbol{f} + (-\boldsymbol{f}))(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) + (-\boldsymbol{f})(\boldsymbol{x})$ (By definition)

$= \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})$ (By definition)

$= \boldsymbol{0}$ (Existence of inverse) i.e. $\boldsymbol{f} + (-\boldsymbol{f}) = \boldsymbol{0}$. Similarly, $(-\boldsymbol{f}) + \boldsymbol{f} = \boldsymbol{0}$.

(vi) **Scalar Multiplication** If $\boldsymbol{f}$ is in $\boldsymbol{V}$ and a is in R, then by definition $(a\,\boldsymbol{f})(\boldsymbol{x}) = a\,\boldsymbol{f}(\boldsymbol{x}) \in \boldsymbol{V}$.

(vii) **Distributive Law** If $\boldsymbol{f}, \boldsymbol{g}$ are in $\boldsymbol{V}$ and $a \in R$, then
$(a(\boldsymbol{f} + \boldsymbol{g}))(\boldsymbol{x}) = a(\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{x}) = a(\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})) = a\,\boldsymbol{f}(\boldsymbol{x}) + a\,\boldsymbol{g}(\boldsymbol{x})$

$= (a\,\boldsymbol{f})(\boldsymbol{x}) + (a\,\boldsymbol{g})(\boldsymbol{x}) = (a\,\boldsymbol{f} + a\,\boldsymbol{g})(\boldsymbol{x})$ and, therefore, $a(\boldsymbol{f} + \boldsymbol{g}) = a\,\boldsymbol{f} + a\,\boldsymbol{g}$

(viii) Let a, b in R and $\boldsymbol{f} \in \boldsymbol{V}$, then

$((a + b)\,\boldsymbol{f})(\boldsymbol{x}) = (a + b)\,\boldsymbol{f}(\boldsymbol{x}) = a\,\boldsymbol{f}(\boldsymbol{x}) + b\,\boldsymbol{f}(\boldsymbol{x}) = (a\,\boldsymbol{f})(\boldsymbol{x}) + (b\,\boldsymbol{f})(\boldsymbol{x}) = (a\,\boldsymbol{f} + b\,\boldsymbol{f})(\boldsymbol{x})$

Thus $(a + b)\,\boldsymbol{f} = a\,\boldsymbol{f} + b\,\boldsymbol{f}$

(ix) $a(b\,\boldsymbol{f})(\boldsymbol{x}) = a(b\,\boldsymbol{f}(\boldsymbol{x})) = (a\,b)\,\boldsymbol{f}(\boldsymbol{x})$ showing that $a(b\,\boldsymbol{f}) = (a\,b)\,\boldsymbol{f}$

(x) $(1.\boldsymbol{f})(\boldsymbol{x}) = 1\,\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$ (Existence of identity)

And so $1.\boldsymbol{f} = \boldsymbol{f}$

Hence $\boldsymbol{V}$ is a real vector space.

Example 4 If $\boldsymbol{p}(\boldsymbol{x}) = a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + \ldots + a_n\boldsymbol{x}^n$
and $\boldsymbol{q}(\boldsymbol{x}) = b_0 + b_1\boldsymbol{x} + b_2\boldsymbol{x}^2 + \ldots + b_n\boldsymbol{x}^n$
We define

$$\boldsymbol{p}(\boldsymbol{x}) + \boldsymbol{q}(\boldsymbol{x}) = (a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + \ldots + a_n\boldsymbol{x}^n) + (b_0 + b_1\boldsymbol{x} + b_2\boldsymbol{x}^2 + \ldots + b_n\boldsymbol{x}^n)$$

$= (a_0 + b_0) + (a_1 + b_1)\boldsymbol{x} + (a_2 + b_2)\boldsymbol{x}^2 + \ldots + (a_n + b_n)\boldsymbol{x}^n$ and for any scalar k,

$$k\boldsymbol{p}(\boldsymbol{x}) = k(a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + \ldots + a_n\boldsymbol{x}^n) = ka_0 + ka_1\boldsymbol{x} + ka_2\boldsymbol{x}^2 + \ldots + k\,a_n\boldsymbol{x}^n$$

Clearly the given polynomial is a vector space under the addition and scalar multiplication.

Example 5 **(The Zero Vector Space)** Let $\boldsymbol{V}$ consists of a single object, which we define by $\boldsymbol{0}$ and $\boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$ and k $\boldsymbol{0} = \boldsymbol{0}$ for all scalars k. It is easy to check that all the vector space axioms are satisfied. We call $\boldsymbol{V=\{0\}}$ as the zero vector space.

Example 6 (Every plane through the origin is a vector space)

Let $\boldsymbol{V}$ be any plane through the origin in $\boldsymbol{R}^3$. We shall show that the points in $\boldsymbol{V}$ form a vector space under a standard addition and scalar multiplication operations for vectors in $\boldsymbol{R}^3$.

From example1, we know that $\boldsymbol{R}^3$ itself is a vector space under these operations. Thus, Axioms 2, 3, 7, 8, 9 and 10 hold for all points in $\boldsymbol{R}^3$ and consequently for all points in the plane $\boldsymbol{V}$. We therefore need only show that Axioms 1, 4, 5 and 6 are satisfied.

Since the plane is passing through the origin, it has an equation of the form

$$a x + b y + c z = 0 \qquad (1)$$

Thus, if $\boldsymbol{u} = (u_1, u_2, u_3)$ and $\boldsymbol{v} = (v_1, v_2, v_3)$ are points in $\boldsymbol{V}$, then

$$a u_1 + b u_2 + c u_3 = 0 \text{ and } a v_1 + b v_2 + c v_3 = 0.$$

Adding these equations gives $a (u_1 + v_1) + b (u_2 + v_2) + c (u_3 + v_3) = 0$

This equality tell us that the coordinates of the point

$$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3)$$

satisfies (1); thus, $\boldsymbol{u} + \boldsymbol{v}$ lies in plane $\boldsymbol{V}$. This proves that the Axiom 1 is satisfied.

There exists $\boldsymbol{0} = (0, 0, 0)$ such that $a (0) + b (0) + c (0) = 0$. Therefore, Axiom 4 is satisfied.

Multiplying $a u_1 + b u_2 + c u_3 = 0$ through by k gives

$$a (ku_1) + b (ku_2) + c (ku_3) = 0$$

Thus, $(ku_1, k u_2, k u_3) = k (u_1, u_2, u_3) = k \boldsymbol{u} \in \boldsymbol{V}$. Hence, Axiom 6 is satisfied.

We shall prove the axiom 5 is satisfied. Multiplying $a u_1 + b u_2 + c u_3 = 0$ through by -1 gives a $(-1u_1) + b (-1u_2) + c (-1u_3) = 0$

Thus, $(-u_1, - u_2, - u_3) = - (u_1, u_2, u_3) = -\boldsymbol{u} \in \mathbf{V}$. This establishes Axiom 5.

Example 7 (A set that is not a vector space)

Let $\boldsymbol{V}=\boldsymbol{R}^2$ and define addition and scalar multiplication operation as follows. If $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$ then define

$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2)$ and if k is any real number then define $k\boldsymbol{u} = (ku_1, 0)$.

For any vector $\boldsymbol{u} \in \boldsymbol{V}$, $1\boldsymbol{u} = 1(u_1, u_2) = (1\, u_1, 0) = (u_1, 0) \neq \boldsymbol{u}$ where 1 is the multiplicative identity of real numbers. Therefore, the axiom 10 is not satisfied.

Hence, $\boldsymbol{V}=\boldsymbol{R}^2$ is not a vector space.

Theorem 1 Let $\boldsymbol{V}$ be a vector space, $\boldsymbol{u}$ a vector in $\boldsymbol{V}$, and k is a scalar, then

(i) $0\boldsymbol{u} = \boldsymbol{0}$ (ii) $k\boldsymbol{0} = \boldsymbol{0}$

(iii) $(-1)\,\boldsymbol{u} = -\boldsymbol{u}$ (iv) If $k\,\boldsymbol{u} = \boldsymbol{0}$ then $k = 0$ or $\boldsymbol{u} = \boldsymbol{0}$

Definition A subset $\boldsymbol{W}$ of a vector space $\boldsymbol{V}$ is called a subspace of $\boldsymbol{V}$ if $\boldsymbol{W}$ itself a vector space under the addition and scalar multiplication defined on $\boldsymbol{V}$.

Note If $\boldsymbol{W}$ is a part of a larger set $\boldsymbol{V}$ that is already known to be a vector space, then certain axioms need not be verified for $\boldsymbol{W}$ because they are "inherited" from $\boldsymbol{V}$. For example, there is no need to check that $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ (Axiom 2) for $\boldsymbol{W}$ because this holds for all vectors in $\boldsymbol{V}$ and consequently for all vectors in $\boldsymbol{W}$. Other Axioms are inherited by $\boldsymbol{W}$ from $\boldsymbol{V}$ are 3, 7, 8, 9, and 10. Thus, to show that a set $\boldsymbol{W}$ is a subspace of a vector space $\boldsymbol{V}$, we need only verify Axioms 1, 4, 5 and 6. The following theorem shows that even Axioms 4 and 5 can be omitted.

Theorem 2 If $\boldsymbol{W}$ is a set of one or more vectors from a vector space $\boldsymbol{V}$, then $\boldsymbol{W}$ is subspace of $\boldsymbol{V}$ if and only if the following conditions hold.

(a) If $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors in $\boldsymbol{W}$, then $\boldsymbol{u} + \boldsymbol{v}$ is in $\boldsymbol{W}$

(b) If k is any scalar and $\boldsymbol{u}$ is any vector in $\boldsymbol{W}$, then $k\,\boldsymbol{u}$ is in $\boldsymbol{W}$.

Proof If $\boldsymbol{W}$ is a subspace of $\boldsymbol{V}$, then all the vector space axioms are satisfied; in particular, Axioms 1 and 6 hold. But these are precisely conditions (a) and (b).

Conversely, assume conditions (a) and (b) hold. Since these conditions are vector space Axioms 1 and 6, we need only show that $\boldsymbol{W}$ satisfies the remaining 8 axioms. The vectors in $\boldsymbol{W}$ automatically satisfy axioms 2, 3, 7, 8, 9, and 10 since they are satisfied by all vectors in $\boldsymbol{V}$. Therefore, to complete the proof, we need only to verify that vectors in $\boldsymbol{W}$ satisfy axioms 4 and 5.

Let $\boldsymbol{u}$ be any vector in $\boldsymbol{W}$. By condition (b), $k\,\boldsymbol{u}$ is in $\boldsymbol{W}$ for every scalar k. Setting $k = 0$, it follows from theorem 1 that $0\,\boldsymbol{u} = \boldsymbol{0}$ is in $\boldsymbol{W}$, and setting $k = -1$, it follows that $(-1)\,\boldsymbol{u} = -\boldsymbol{u}$ is in $\boldsymbol{W}$. ☐

Remark

(1) The theorem states that $\boldsymbol{W}$ is a subspace of $\boldsymbol{V}$ if and only if $\boldsymbol{W}$ is closed under addition and closed under scalar multiplication.

(2) Every vector space has at least two subspaces, itself and the subspace $\{\boldsymbol{0}\}$ consisting only of the zero vector. Thus the subspace $\{\boldsymbol{0}\}$ is called the zero subspace.

Example 8 Let $\boldsymbol{W}$ be the subset of $\boldsymbol{R}^3$ consisting of the all the vectors of the form $(a, b, 0)$, where a and b are real numbers. To check if $\boldsymbol{W}$ is subspace of $\boldsymbol{R}^3$, we first see that axiom 1 and 6 of a vector space holds.

Let $\boldsymbol{u} = (a_1, b_1, 0)$ and $\boldsymbol{v} = (a_2, b_2, 0)$ be vectors in $\boldsymbol{W}$ then
$\boldsymbol{u} + \boldsymbol{v} = (a_1, b_1, 0) + (a_2, b_2, 0) = (a_1 + a_2, b_1 + b_2, 0)$ is in $\boldsymbol{W}$. Since the third component is zero. Also c is scalar, and then $c\boldsymbol{u} = c(a_1, b_1, 0) = (ca_1, cb_1, 0)$ is in $\boldsymbol{W}$. Therefore the I[st] and 6[th] axioms of the vector space holds. We can also verify the other axioms of vector space. Hence $\boldsymbol{W}$ is a subspace.

Example 9 Consider the set $\boldsymbol{W}$ consisting of all 2x3 matrices of the form $\begin{bmatrix} a & b & 0 \\ 0 & c & d \end{bmatrix}$, Where a, b, c and d are arbitrary real numbers. Show that the $\boldsymbol{W}$ is a subspace $\boldsymbol{M_{2x3}}$.

Solution Consider $\boldsymbol{u} = \begin{bmatrix} a_1 & b_1 & 0 \\ 0 & c_1 & d_1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} a_2 & b_2 & 0 \\ 0 & c_2 & d_2 \end{bmatrix}$ in $\boldsymbol{W}$

Then $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} a_1 & b_1 & 0 \\ 0 & c_1 & d_1 \end{bmatrix} + \begin{bmatrix} a_2 & b_2 & 0 \\ 0 & c_2 & d_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 & b_1 + b_2 & 0 \\ 0 & c_1 + c_2 & d_1 + d_2 \end{bmatrix}$ is in $\boldsymbol{W}$.

So that the (a) part of the theorem is satisfied. Also k is a scalar, and then
$k\boldsymbol{u} = \begin{bmatrix} ka_1 & kb_1 & 0 \\ 0 & kc_1 & kd_1 \end{bmatrix}$ is in $\boldsymbol{W}$. So the (b) part of the above theorem is also satisfied.
Hence $\boldsymbol{W}$ is a subspace of $\boldsymbol{M_{2x3}}$.

Note Let $\boldsymbol{V}$ is a vector space then every subset of $\boldsymbol{V}$ is not necessary a subspace of $\boldsymbol{V}$. For example, let $\boldsymbol{V} = \boldsymbol{R^2}$ then any line in $\boldsymbol{R^2}$ not passing through origin is not a subspace of $\boldsymbol{R^2}$. Similarly, a plane in $\boldsymbol{R^3}$ not passing through the origin is not a subspace of $\boldsymbol{R^3}$.

Example 10 Let $\boldsymbol{W}$ be the subset of $\boldsymbol{R^3}$ consisting of all vectors of the form *(a, b, 1)*, where *a*, *b* are any real numbers. To check whether property (a) and (b) of the above theorem holds. Let $\boldsymbol{u} = (a_1, b_1, 1)$ and $\boldsymbol{v} = (a_2, b_2, 1)$ be vectors in $\boldsymbol{W}$.
Then $\boldsymbol{u} + \boldsymbol{v} = (a_1, b_1, 1) + (a_2, b_2, 1) = (a_1 + a_2, b_1 + b_2, 1 + 1)$ which is not in $\boldsymbol{W}$ because the third component 2 is not 1. As the Ist property does not hold therefore, the given set of vectors is not a vector space.

Example 11 Which of the following are subspaces of $\boldsymbol{R^3}$
(i) All vectors of the form *(a, 0, 0)*
(ii) All vectors of the form *(a, 1, 1)*
(iii) All vectors of the form *(a, b, c)*, where $b = a + c$
(iv) All vectors of the form *(a, b, c)*, where $b = a + c + 1$

Solution Let $\boldsymbol{W}$ is the set of all vectors of the form *(a, 0, 0)*.

(i) Suppose $\boldsymbol{u} = (u_1, 0, 0)$ and $\boldsymbol{v} = (v_1, 0, 0)$ are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = (u_1, 0, 0) + (v_1, 0, 0) = (u_1 + v_1, 0, 0)$ which is of the form *(a, 0, 0)*.

Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = (u_1, 0, 0)$ is any vector in $\boldsymbol{W}$, then $k\,\boldsymbol{u} = k\,(u_1, 0, 0) = (k\,u_1, 0, 0)$
which is of the form $(a, 0, 0)$. Therefore, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{R^3}$.

(ii) Let $\boldsymbol{W}$ is the set of all vectors of the form $(a, 1, 1)$.

Suppose $\boldsymbol{u} = (u_1, 1, 1)$ and $\boldsymbol{v} = (v_1, 1, 1)$ are in $\boldsymbol{W}$. Then $\boldsymbol{u} + \boldsymbol{v} = (u_1, 1, 1) + (v_1, 1, 1) = (u_1 + v_1, 2, 2)$ which is not of the form $(a, 1, 1)$. Therefore, $\boldsymbol{u} + \boldsymbol{v} \notin \boldsymbol{W}$. Hence $\boldsymbol{W}$ is not the subspace of $\boldsymbol{R^3}$.

(iii) Suppose $\boldsymbol{W}$ is the set of all vectors of the form (a, b, c), where $b = a + c$

Suppose $\boldsymbol{u} = (u_1, u_1 + u_3, u_3)$ and $\boldsymbol{v} = (v_1, v_1 + v_3, v_3)$ are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_1 + u_3, u_3) + (v_1, v_1 + v_3, v_3)$
$= (u_1 + v_1, u_1 + u_3 + v_1 + v_3, u_3 + v_3)$
$= (u_1 + v_1, (u_1 + v_1) + (u_3 + v_3), u_3 + v_3)$, which is of the form $(a, a + c, c)$.

Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = (u_1, u_1 + u_3, u_3)$ is any vector in $\boldsymbol{W}$, then
$k\,\boldsymbol{u} = k\,(u_1, u_1 + u_3, u_3) = (k\,u_1, k\,(u_1 + u_3), k\,u_3)$ (By definition)
$= (k\,u_1, k\,u_1 + k\,u_3, k\,u_3)$ (By Distributive Law)
Which is of the form $(a, a + c, c)$. Therefore, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{R^3}$.

(iv) Let $\boldsymbol{W}$ is the set of all vectors of the form (a, b, c), where $b = a + c + 1$
Suppose $\boldsymbol{u} = (u_1, u_1 + u_3 + 1, u_3)$ and $\boldsymbol{v} = (v_1, v_1 + v_3 + 1, v_3)$ are in $\boldsymbol{W}$.
Then $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_1 + u_3 + 1, u_3) + (v_1, v_1 + v_3 + 1, v_3)$
$= (u_1 + v_1, u_1 + u_3 + 1 + v_1 + v_3 + 1, u_3 + v_3)$
$= (u_1 + v_1, (u_1 + v_1) + (u_3 + v_3) + 2, u_3 + v_3)$
Which is not of the form $(a, a + c + 1, c)$. Therefore, $\boldsymbol{u} + \boldsymbol{v} \notin \boldsymbol{W}$. Hence $\boldsymbol{W}$ is not the subspace of $\boldsymbol{R^3}$.

Example 12 Determine which of the following are subspaces of $\boldsymbol{P_3}$.
(i) All polynomials $a_0 + a_1\,\mathbf{x} + a_2\,\mathbf{x}^2 + a_3\,\mathbf{x}^3$ for which $a_0 = 0$
(ii) All polynomials $a_0 + a_1\,\mathbf{x} + a_2\,\mathbf{x}^2 + a_3\,\mathbf{x}^3$ for which $a_0 + a_1 + a_2 + a_3 = 0$
(iii) All polynomials $a_0 + a_1\,\mathbf{x} + a_2\,\mathbf{x}^2 + a_3\,\mathbf{x}^3$ for which a_0, a_1, a_2, and a_3 are integers
(iv) All polynomials of the form $a_0 + a_1\,\mathbf{x}$, where a_0 and a_1 are real numbers.

Solution (i) Let W is the set of all polynomials $a_0 + a_1 x + a_2 x^2 + a_3 x^3$ for which $a_0 = 0$.
Suppose that $u = c_0 + c_1 x + c_2 x^2 + c_3 x^3$ (where $c_0 = 0$) and $v = b_0 + b_1 x + b_2 x^2 + b_3 x^3$ (where $b_0 = 0$) are in W. Then $u + v = (c_0 + c_1x + c_2x^2 + c_3x^3) + (b_0 + b_1x + b_2x^2 + b_3x^3) = (c_0 + b_0) + (c_1 + b_1)x + (c_2 + b_2)x^2 + (c_3 + b_3) x^3$, where $c_0 + b_0 = 0$.
Therefore, $u + v \in W$

If k is any scalar and $u = c_0 + c_1 x + c_2 x^2 + c_3 x^3$ (where $c_0 = 0$) is any vector in W. Then $k u = k (c_0 + c_1x + c_2x^2 + c_3x^3) = (kc_0) + (kc_1) x + (kc_2) x^2 + (kc_3) x^3$ where $k c_0 = 0$. Therefore, $k u \in W$. Hence W is the subspace of P_3.

(ii) Let W is the set of all polynomials $a_0 + a_1x + a_2x^2 + a_3x^3$ for which $a_0+a_1+a_2+a_3 = 0$.
Suppose that $u = c_0 + c_1 x + c_2 x^2 + c_3 x^3$ (where $c_0 + c_1 + c_2 + c_3 = 0$) and $v = b_0 + b_1 x + b_2 x^2 + b_3 x^3$ (where $b_0 + b_1 + b_2 + b_3 = 0$) are in W.

Now
$u + v = (c_0 + c_1 x + c_2 x^2 + c_3 x^3) + (b_0 + b_1 x + b_2 x^2 + b_3 x^3)$

$= (c_0 + b_0) + (c_1 + b_1) x + (c_2 + b_2) x^2 + (c_3 + b_3) x^3$

Where $(c_0+b_0) + (c_1+b_1) + (c_2+b_2) + (c_3+b_3) = (c_0+c_1+c_2+c_3) + (b_0+b_1+b_2+b_3) = 0 + 0 = 0$. Therefore, $u + v \in W$

If k is any scalar and $u = c_0 + c_1 x + c_2 x^2 + c_3 x^3$ (where $c_0 + c_1 + c_2 + c_3 = 0$) is any vector in W. Then $k u = k (c_0 + c_1x + c_2x^2 + c_3x^3) = (kc_0) + (kc_1) x + (kc_2) x^2 + (kc_3) x^3$
Where $(k c_0) + (k c_1) + (k c_2) + (k c_3) = k(c_0 + c_1 + c_2 + c_3) = k.0 = 0$

Therefore, $k u \in W$. Hence W is the subspace of P_3.

(iii) Let W is the set of all polynomials $a_0 + a_1 x + a_2 x^2 + a_3 x^3$ for which $a_0, a_1, a_2,$ and a_3 are integers.
Suppose that the vectors $u = c_0 + c_1 x + c_2 x^2 + c_3 x^3$ (where $c_0, c_1, c_2,$ and c_3 are integers) and $v = b_0 + b_1 x + b_2 x^2 + b_3 x^3$ (where $b_0, b_1, b_2,$ and b_3 are integers) are in W.
Now
$u + v = (c_0 + c_1 x + c_2 x^2 + c_3 x^3) + (b_0 + b_1 x + b_2 x^2 + b_3 x^3)$
$= (c_0 + b_0) + (c_1 + b_1) x + (c_2 + b_2) x^2 + (c_3 + b_3) x^3$, where
$(c_0 + b_0), (c_1 + b_1), (c_2 + b_2),$ and $(c_3 + b_3)$ are integers (integers are closed under addition). Therefore, $u + v \in W$

If k is any scalar and $u = c_0 + c_1 x + c_2 x^2 + c_3 x^3$ (where $c_0, c_1, c_2,$ and c_3 are integers) is any vector in W. Then $k u = k (c_0 + c_1x + c_2x^2 + c_3x^3 = (kc_0) + (kc_1) x + (kc_2) x^2 + (kc_3) x^3$, where $(k c_0), (k c_1), (k c_2), and (k c_3)$ are not integers (product of real number and integer). Therefore, $k u \notin W$. Hence, W is not the subspace of P_3.

(iv) Let $\boldsymbol{W}$ is the set of all polynomials of the form $a_0 + a_1\,\boldsymbol{x}$, where a_0 and a_1 are real numbers. Suppose that $\boldsymbol{u} = c_0 + c_1\,\boldsymbol{x}$ (where c_0 and c_1 are real numbers) and
$\boldsymbol{v} = b_0 + b_1\,\boldsymbol{x}$ (where b_0 and b_1 are real numbers) are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = (c_0 + c_1\,\boldsymbol{x}) + (b_0 + b_1\,\boldsymbol{x}) = (c_0 + b_0) + (c_1 + b_1)\,\boldsymbol{x}$
Where $(c_0 + b_0)$ and $(c_1 + b_1)$ are real numbers.
Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = c_0 + c_1\,\boldsymbol{x}$ (where c_0 and c_1 are real numbers) is any vector in $\boldsymbol{W}$.
Then $k\,\boldsymbol{u} = k\,(c_0 + c_1\,\boldsymbol{x}) = (k\,c_0) + (k\,c_1)\,\boldsymbol{x}$
Where $(k\,c_0)$ and $(k\,c_1)$ are real numbers.
Therefore, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{P_3}$.

Example 13 Determine which of the following are subspaces of $\boldsymbol{M_{22}}$.

(i) All matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where $a + b + c + d = 0$

(ii) All 2 x 2 matrices $\boldsymbol{A}$ such that det $(\boldsymbol{A}) = 0$

(iii) All the matrices of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$

Solution Let $\boldsymbol{W}$ is the set of all matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where $a + b + c + d = 0$.

(i) Suppose $\boldsymbol{u} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ (where $e + f + g + h = 0$) and $\boldsymbol{v} = \begin{bmatrix} l & m \\ n & p \end{bmatrix}$
(Where $l + m + n + p = 0$) are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} l & m \\ n & p \end{bmatrix} = \begin{bmatrix} e+l & f+m \\ g+n & h+p \end{bmatrix}$ (By definition)

Where $(e + l) + (f + m) + (g + n) + (h + p)$
$= (e + f + g + h) + (l + m + n + p) = 0 + 0 = 0$
Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ (where $e + f + g + h = 0$) is any vector in $\boldsymbol{W}$.

Then $k\,\boldsymbol{u} = k\begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ke & kf \\ kg & kh \end{bmatrix}$ (by definition)

Where $k\,e + k\,f + k\,g + k\,h = k\,(e + f + g + h) = k\,0 = 0$
Hence, $k\,\boldsymbol{u} \in \boldsymbol{W}$. Therefore, $\boldsymbol{W}$ is subspace of $\boldsymbol{M_{22}}$.

(ii) Let $\boldsymbol{W}$ is the set of all 2 x 2 matrices $\boldsymbol{A}$ such that det $(\boldsymbol{A}) = 0$

Suppose $\boldsymbol{u} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ (Where det $(\boldsymbol{u}) = e\,h - f\,g = 0$) and $\boldsymbol{v} = \begin{bmatrix} l & m \\ n & p \end{bmatrix}$

(Where det $(\boldsymbol{v}) = l\,p - m\,n = 0$) are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} l & m \\ n & p \end{bmatrix} = \begin{bmatrix} e+l & f+m \\ g+n & h+p \end{bmatrix}$ (By definition)

Where det $(\boldsymbol{u} + \boldsymbol{v}) = (e + l)\,(h + p) - (f + m)\,(g + n)$

$= e\,h + e\,p + l\,h + l\,p - f\,g - f\,n - m\,g - m\,n$

$= (e\,h - f\,g) + (l\,p - m\,n) + e\,p + l\,h - f\,n - m\,g = e\,p + l\,h - f\,n - m\,g \neq 0$

Therefore, $\boldsymbol{u} + \boldsymbol{v} \notin \boldsymbol{W}$. Therefore, $\boldsymbol{W}$ is not subspace of $\boldsymbol{M_{22}}$.

(iii) Let $\boldsymbol{W}$ is the set of all matrices of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$

Suppose $\boldsymbol{u} = \begin{bmatrix} e & f \\ 0 & g \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} l & m \\ 0 & n \end{bmatrix}$ are in $\boldsymbol{W}$.

Then $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} e & f \\ 0 & g \end{bmatrix} + \begin{bmatrix} l & m \\ 0 & n \end{bmatrix} = \begin{bmatrix} e+l & f+m \\ 0 & g+n \end{bmatrix}$ (By definition)

Which is of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$. Therefore, $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{u} = \begin{bmatrix} e & f \\ 0 & g \end{bmatrix}$ is any vector in $\boldsymbol{W}$.

Then $k\,\boldsymbol{u} = k\begin{bmatrix} e & f \\ 0 & g \end{bmatrix} = \begin{bmatrix} ke & kf \\ 0 & kg \end{bmatrix}$ (By definition)

Which is of the form $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$. Hence, $k\,\boldsymbol{u} \in \boldsymbol{W}$

Therefore, $\boldsymbol{W}$ is subspace of $\boldsymbol{M_{22}}$.

Example 14 Determine which of the following are subspaces of the space $\boldsymbol{F}(-\infty,\infty)$.

(i) All $\boldsymbol{f}$ such that $\boldsymbol{f}\,(\boldsymbol{x}) \leq \boldsymbol{0}$ for all $\boldsymbol{x}$ (ii) all $\boldsymbol{f}$ such that $\boldsymbol{f}\,(0) = \boldsymbol{0}$

(iii) All $\boldsymbol{f}$ such that $\boldsymbol{f}\,(0) = \boldsymbol{2}$ (iv) all constant functions

(v) All $\boldsymbol{f}$ of the form $k_1 + k_2 \sin x$, where k_1 and k_2 are real numbers

(vi) All everywhere differentiable functions that satisfy $\boldsymbol{f}' + 2\,\boldsymbol{f} = 0$.

Solution **(i)** Let $\boldsymbol{W}$ is the set of all $\boldsymbol{f}$ such that $\boldsymbol{f}\,(\boldsymbol{x}) \leq \boldsymbol{0}$ for all $\boldsymbol{x}$.

Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$. Then $\boldsymbol{g}\,(\boldsymbol{x}) \leq \boldsymbol{0}$ for all $\boldsymbol{x}$ and $\boldsymbol{h}\,(\boldsymbol{x}) \leq \boldsymbol{0}$ for all $\boldsymbol{x}$.

Now $(\boldsymbol{g} + \boldsymbol{h})\,(\boldsymbol{x}) = \boldsymbol{g}\,(\boldsymbol{x}) + \boldsymbol{h}\,(\boldsymbol{x}) \leq \boldsymbol{0}$. Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{g}$ is any vector in $\boldsymbol{W}$. Then $\boldsymbol{g}\,(\boldsymbol{x}) \leq \boldsymbol{0}$ for all $\boldsymbol{x}$

Now $(k\,\boldsymbol{g})\,(\boldsymbol{x}) = k\,\boldsymbol{g}\,(\boldsymbol{x})$, which is greater than $\boldsymbol{0}$ for negative real values of k.

$\therefore k\boldsymbol{g} \notin \boldsymbol{W} \quad \forall\; k < 0$.

Hence $\boldsymbol{W}$ is not the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(ii) Let $\boldsymbol{W}$ is the set of all $\boldsymbol{f}$ such that $\boldsymbol{f}(0) = \boldsymbol{0}$.
Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$. Then $\boldsymbol{g}(0) = \boldsymbol{0}$ and $\boldsymbol{h}(0) = \boldsymbol{0}$
Now $(\boldsymbol{g} + \boldsymbol{h})(0) = \boldsymbol{g}(0) + \boldsymbol{h}(0) = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$. Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$
If k is any scalar and $\boldsymbol{g}$ is any vector in $\boldsymbol{W}$. Then $\boldsymbol{g}(0) = \boldsymbol{0}$
Now $(k\,\boldsymbol{g})(0) = k\,\boldsymbol{g}(0) = k\,\boldsymbol{0} = \boldsymbol{0}$. $\therefore k\boldsymbol{g} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(iii) Let $\boldsymbol{W}$ is the set of all $\boldsymbol{f}$ such that $\boldsymbol{f}(0) = \boldsymbol{2}$
Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$. Then $\boldsymbol{g}(0) = \boldsymbol{2}$ and $\boldsymbol{h}(0) = \boldsymbol{2}$
Now $(\boldsymbol{g} + \boldsymbol{h})(0) = \boldsymbol{g}(0) + \boldsymbol{h}(0) = \boldsymbol{2} + \boldsymbol{2} \neq \boldsymbol{2} \therefore k\boldsymbol{g} \notin \boldsymbol{W}$. Hence $\boldsymbol{W}$ is not the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(iv) Let $\boldsymbol{W}$ is the set of all constant functions. Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$.
Then $\boldsymbol{g}(x) = \boldsymbol{a}$ and $\boldsymbol{h}(x) = \boldsymbol{b}$, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are constants.
Now $(\boldsymbol{g} + \boldsymbol{h})(x) = \boldsymbol{g}(x) + \boldsymbol{h}(x) = \boldsymbol{a} + \boldsymbol{b}$, which is constant. Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$
If k is any scalar and $\boldsymbol{g}$ is any vector in $\boldsymbol{W}$. Then $\boldsymbol{g}(x) = \boldsymbol{a}$, where $\boldsymbol{a}$ is any constant.
Now $(k\,\boldsymbol{g})(x) = k\,\boldsymbol{g}(x) = k\,\boldsymbol{a}$, which is a constant. $\therefore k\boldsymbol{g} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(v) Let $\boldsymbol{W}$ is the set of all $\boldsymbol{f}$ of the form $k_1 + k_2 \sin x$, where k_1 and k_2 are real numbers
Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$. Then $\boldsymbol{g}(x) = m_1+m_2\sin x$ and $\boldsymbol{h}(x) = n_1+n_2 \sin x$, where m_1, m_2, $n_1$and n_2 are real numbers.

Now $(\boldsymbol{g} + \boldsymbol{h})(x) = \boldsymbol{g}(x) + \boldsymbol{h}(x) = [m_1+m_2\sin x]+[n_1+n_2\sin x] = (m_1+n_1)+(m_2+n_2)\sin x$
Which is of the form $k_1 + k_2 \sin x$. Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$

If k is any scalar and $\boldsymbol{g}$ is any vector in $\boldsymbol{W}$. Then $\boldsymbol{g}(x) = m_1 + m_2 \sin x$, where m_1 and m_2 are any real numbers.

Now $(k\,\boldsymbol{g})(x) = k\,\underline{\boldsymbol{g}}(x) = k\,[m_1 + m_2 \sin x] = (k\,m_1) + (k\,m_2) \sin x$
Which is of the form $k_1 + k_2 \sin x$. $\therefore k\boldsymbol{g} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{F}(-\infty,\infty)$.

(vi) Let $\boldsymbol{W}$ is the set of all everywhere differentiable functions that satisfy $\boldsymbol{f}' + 2\boldsymbol{f} = \boldsymbol{0}$.
Suppose $\boldsymbol{g}$ and $\boldsymbol{h}$ are the vectors in $\boldsymbol{W}$**.** Then $\boldsymbol{g}' + 2\boldsymbol{g} = \boldsymbol{0}$ and $\boldsymbol{h}' + 2\boldsymbol{h} = \boldsymbol{0}$
Now $(\boldsymbol{g} + \boldsymbol{h})' + 2(\boldsymbol{g} + \boldsymbol{h}) = \boldsymbol{g}' + \boldsymbol{h}' + 2(\boldsymbol{g} + \boldsymbol{h}) = (\boldsymbol{g}' + 2\boldsymbol{g}) + (\boldsymbol{h}' + 2\boldsymbol{h}) = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$
Therefore, $\boldsymbol{g} + \boldsymbol{h} \in \boldsymbol{W}$
If k is any scalar and $\boldsymbol{g}$ is any vector in $\boldsymbol{W}$. Then $\boldsymbol{g}' + 2\boldsymbol{g} = \boldsymbol{0}$
Now $(k\boldsymbol{g})' + 2(k\boldsymbol{g}) = k\boldsymbol{g}' + 2k\boldsymbol{g} = k(\boldsymbol{g}' + 2\boldsymbol{g}) = k.\boldsymbol{0} = \boldsymbol{0}$
$\therefore k\boldsymbol{g} \in \boldsymbol{W}$. Hence $\boldsymbol{W}$ is the subspace of $\boldsymbol{F}(-\infty,\infty)$.

Remark Let n be a non-negative integer, and let $\boldsymbol{P_n}$ be the set of real valued function of the form $\boldsymbol{p(x)} = a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + ... + a_n\boldsymbol{x}^n$ where $a_0, a_1, a_2, ..., a_n$ *are* real numbers, then $\boldsymbol{P_n}$ is a subspace $\boldsymbol{F}(-\infty,\infty)$.

Example 15 Show that the invertible n x n matrices do not form a subspace of $\boldsymbol{M}_{n\times n}$.

Solution Let $\boldsymbol{W}$ is the set of invertible matrices in $\boldsymbol{M}_{n\times n}$. This set fails to be a subspace on both counts- it is closed under neither scalar multiplication nor addition.

For example consider invertible matrices $\boldsymbol{W} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \boldsymbol{V} = \begin{bmatrix} -1 & 2 \\ -2 & 5 \end{bmatrix}$ in $\boldsymbol{M}_{n\times n}$.

The matrix $\boldsymbol{0.U}$ is a 2x2 zero matrix, hence is not invertible; and the matrix $\boldsymbol{U} + \boldsymbol{V}$ has a column of zeros, hence is not invertible.

Theorem If $\boldsymbol{Ax = 0}$ is a homogeneous linear system of m equations in n unknowns, then the set of solution vectors is a subspace of $\boldsymbol{R}^n$.

Example 16 Consider the linear systems

$$(a)\quad \begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (b)\quad \begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ -2 & 4 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$(c)\quad \begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ 4 & 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (d)\quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Each of the systems has three unknowns, so the solutions form subspaces of $\boldsymbol{R}^3$. Geometrically, this means that each solution space must be a line through origin, a plane through origin, the origin only, or all of $\boldsymbol{R}^3$.

Solution *(a)* The solutions are $x = 2s - 3t$, $y = s$, $z = t$. From which it follows that $x = 2y - 3z$ *or* $x - 2y + 3z = 0$.
This is the equation of the plane through the origin with $\boldsymbol{n} = (1, -2, 3)$ as a normal vector.

(b) The solutions are $x = -5t$, $y = -t$, $z = t$, which are parametric equations for the line through the origin parallel to the vector $\boldsymbol{v} = (-5, -1, 1)$.

(c) The solution is $x = 0$, $y = 0$, $z = 0$ so the solution space is the origin only, that is $\{\boldsymbol{0}\}$.

(d) The solutions are $x = r$, $y = s$, $z = t$. where r, s and t have arbitrary values, so the solution space is all $\boldsymbol{R}^3$.

A Subspace Spanned by a Set: The next example illustrates one of the most common ways of describing a subspace. We know that the term linear combination refers to any sum of scalar multiples of vectors, and Span $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ denotes the set of all vectors that can be written as linear combinations of $\boldsymbol{v_1}, \dots, \boldsymbol{v_p}$.

Example 17 Given $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ in a vector space $\boldsymbol{V}$, let $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. Show that $\boldsymbol{H}$ is a subspace of $\boldsymbol{V}$.

Solution The zero vector is in $\boldsymbol{H}$, since $\boldsymbol{0} = 0\boldsymbol{v_1} + 0\boldsymbol{v_2}$. To show that $\boldsymbol{H}$ is closed under vector addition, take two arbitrary vectors in $\boldsymbol{H}$, say,

$$\boldsymbol{u} = s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2} \quad \text{and} \quad \boldsymbol{w} = t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2}$$

By Axioms 2, 3 and 8 for the vector space $\boldsymbol{V}$.

$$\begin{aligned}\boldsymbol{u} + \boldsymbol{w} &= (s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2}) + (t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2}) \\ &= (s_1 + t_1)\,\boldsymbol{v_1} + (s_2 + t_2)\,\boldsymbol{v_2}\end{aligned}$$

So $\boldsymbol{u} + \boldsymbol{w}$ is in $\boldsymbol{H}$. Furthermore, if c is any scalar, then by Axioms 7 and 9,

$$c\boldsymbol{u} = c\,(s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2}) = (cs_1)\,\boldsymbol{v_1} + (cs_2)\boldsymbol{v_2}$$

Which shows that $c\boldsymbol{u}$ is in $\boldsymbol{H}$ and $\boldsymbol{H}$ is closed under scalar multiplication.

Thus $\boldsymbol{H}$ is a subspace of $\boldsymbol{V}$. □

Later on we will prove that every nonzero subspace of $\boldsymbol{R^3}$, other than $\boldsymbol{R^3}$ itself, is either Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ for some linearly independent $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ or Span $\{\boldsymbol{v}\}$ for $\boldsymbol{v} \neq \boldsymbol{0}$. In the first case the subspace is a plane through the origin and in the second case a line through the origin. (See Figure below) It is helpful to keep these geometric pictures in mind, even for an abstract vector space.

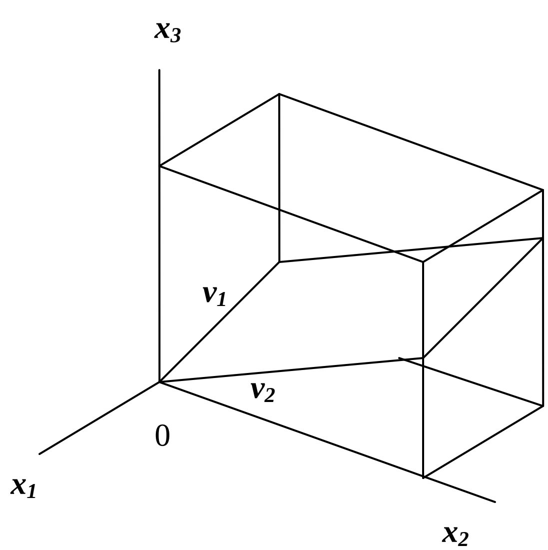


Figure 9 – An example of a subspace

The argument in Example 17 can easily be generalized to prove the following theorem.

Theorem 3 If $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ are in a vector space $\boldsymbol{V}$, then Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ is a subspace of $\boldsymbol{V}$.

We call Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ the subspace spanned (or generated) by $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$. Given any subspace $\boldsymbol{H}$ of $\boldsymbol{V}$, a spanning (or generating) set for $\boldsymbol{H}$ is a set $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ in $\boldsymbol{H}$ such that $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$.

Proof

The zero vector is in $\boldsymbol{H}$, since $\boldsymbol{0} = 0\boldsymbol{v_1} + 0\boldsymbol{v_2} + \ldots + 0\boldsymbol{v_n} = \sum_{j=0}^{j=n} 0v_j = 0\left(\sum_{j=0}^{j=n} v_j\right) = \boldsymbol{0}$

To show that $\boldsymbol{H}$ is closed under vector addition, take two arbitrary vectors in $\boldsymbol{H}$, say,

$\boldsymbol{u} = s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2} + \ldots + t_n\boldsymbol{v_n} = \sum_{i=0}^{i=n} s_i v_i$

and

$\boldsymbol{w} = t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2} \ldots + t_n\boldsymbol{v_n} = \sum_{k=0}^{k=n} t_k v_k$

By Axioms 2, 3 and 8 for the vector space $\boldsymbol{V}$.

$\boldsymbol{u} + \boldsymbol{w} = \sum_{i=0}^{i=n} s_i v_i + \sum_{k=0}^{k=n} t_k v_k = (s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2} + \ldots + s_n\boldsymbol{v_n}) + (t_1\boldsymbol{v_1} + t_2\boldsymbol{v_2} \ldots + t_n\boldsymbol{v_n})$

$= (s_1 + t_1)\,\boldsymbol{v_1} + (s_2 + t_2)\,\boldsymbol{v_2} + \ldots + (s_n + t_n)\,\boldsymbol{v_n} = \sum_{p=0}^{p=n} (s_p + t_p) v_p$

So $\boldsymbol{u} + \boldsymbol{w}$ is in $\boldsymbol{H}$. Furthermore, if c is any scalar, then by Axioms 7 and 9,

$c\boldsymbol{u} = c(s_1\boldsymbol{v_1} + s_2\boldsymbol{v_2} + \ldots + s_n\boldsymbol{v_n}) = (cs_1)\,\boldsymbol{v_1} + (cs_2)\boldsymbol{v_2} + \ldots + (cs_n)\boldsymbol{v_n} = \sum_{r=0}^{r=n} cs_r v_r$.

Which shows that $c\boldsymbol{u}$ is in $\boldsymbol{H}$ and $\boldsymbol{H}$ is closed under scalar multiplication.

Thus $\boldsymbol{H}$ is a subspace of $\boldsymbol{V}$. □

Example 18 Let $\boldsymbol{H}$ be the set of all vectors of the form $(a - 3b, b - a, a, b)$, where a and b are arbitrary scalars. That is, let $\boldsymbol{H} = \{(a - 3b, b - a, a, b)$: a and b in $\boldsymbol{R}\}$. Show that $\boldsymbol{H}$ is a subspace of $\boldsymbol{R^4}$.

Solution Write the vectors in $\boldsymbol{H}$ as column vectors. Then an arbitrary vector in $\boldsymbol{H}$ has the form

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a\underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}}_{\uparrow \; \boldsymbol{v_1}} + b\underbrace{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{\uparrow \; \boldsymbol{v_2}}$$

This calculation shows that $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$, where $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are the vectors indicated above. Thus $\boldsymbol{H}$ is a subspace of $\boldsymbol{R^4}$ by Theorem 3. □

Example18 illustrates a useful technique of expressing a subspace $\boldsymbol{H}$ as the set of linear combinations of some small collection of vectors. If $\boldsymbol{H}$ = Span $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$, we can think of the vectors $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ in the spanning set as "handles" that allow us to hold on to the subspace $\boldsymbol{H}$. *Calculations with the infinitely many vectors in* $\boldsymbol{H}$ *are often reduced to operations with the finite number of vectors in the spanning set.*

Exercises

In exercises 1-13 a set of objects is given together with operations of addition and scalar multiplication. Determine which sets are vector spaces under the given operations. For those that are not, list all axioms that fail to hold.

1. The set of all triples of real numbers *(x, y, z)* with the operations
$$(x,y,z)+(x',y',z') = (x+x',y+y',z+z') \text{ and } k(x,y,z) = (kx,y,z)$$

2. The set of all triples of real numbers *(x, y, z)* with the operations
$$(x,y,z)+(x',y',z') = (x+x',y+y',z+z') \text{ and } k(x,y,z) = (0,0,0)$$

3. The set of all pairs of real numbers *(x, y)* with the operations
$$(x,y)+(x',y') = (x+x',y+y') \text{ and } k(x,y) = (2kx,2ky)$$

4. The set of all pairs of real numbers of the form *(x, 0)* with the standard operations on $\boldsymbol{R^2}$.

5. The set of all pairs of real numbers of the form *(x, y)*, where $x \geq 0$, with the standard operations on $\boldsymbol{R^2}$.

6. The set of all n-tuples of real numbers of the form *(x, x, …, x)* with the standard operations on $\boldsymbol{R^n}$.

7. The set of all pairs of real numbers *(x, y)* with the operations.
$$(x,y)+(x',y') = (x+x'+1,y+y'+1) \text{ and } k(x,y) = (kx,ky)$$

8. The set of all 2x2 matrices of the form $\begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}$ with matrix addition and scalar multiplication.

9. The set of all 2x2 matrices of the form $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ with matrix addition and scalar multiplication.

10. The set of all pairs of real numbers of the form *(1, x)* with the operations
$$(1,y)+(1,y') = (1,y+y') \text{ and } k(1,y) = (1,ky)$$

11. The set of polynomials of the form $\boldsymbol{a} + b\boldsymbol{x}$ with the operations
$$(\boldsymbol{a_0}+a_1\boldsymbol{x})+(\boldsymbol{b_0}+b_1\boldsymbol{x}) = (\boldsymbol{a_0}+\boldsymbol{b_0})+(a_1+b_1)\boldsymbol{x} \text{ and } k(\boldsymbol{a_0}+a_1\boldsymbol{x}) = (k\boldsymbol{a_0})+(ka_1)\boldsymbol{x}$$

12. The set of all positive real numbers with operations $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{xy}$ and $k\boldsymbol{x} = \boldsymbol{x}^k$

13. The set of all real numbers *(x, y)* with operations

$$(x, y) + (x', y') = (xx', yy') \text{ and } k(x, y) = (kx, ky)$$

14. Determine which of the following are subspaces of $\boldsymbol{M_{nn}}$.
(a) all n x n matrices $\boldsymbol{A}$ such that tr $(\boldsymbol{A}) = 0$
(b) all n x n matrices $\boldsymbol{A}$ such that $\boldsymbol{A}^T = -\boldsymbol{A}$
(c) all n x n matrices $\boldsymbol{A}$ such that the linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution
(d) all n x n matrices $\boldsymbol{A}$ such that $\boldsymbol{AB} = \boldsymbol{BA}$ for a fixed n x n matrix $\boldsymbol{B}$

15. Determine whether the solution space of the system $\boldsymbol{Ax} = \boldsymbol{0}$ is a line through the origin, a plane through the origin, or the origin only. If it is a plane, find an equation for it; if it is a line, find parametric equations for it.

(a) $\boldsymbol{A} = \begin{bmatrix} -1 & 1 & 1 \\ 3 & -1 & 0 \\ 2 & -4 & -5 \end{bmatrix}$ (b) $\boldsymbol{A} = \begin{bmatrix} 1 & -2 & 3 \\ -3 & 6 & 9 \\ -2 & 4 & -6 \end{bmatrix}$

(c) $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$ (d) $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -6 \\ 1 & 4 & 4 \\ 3 & 10 & 6 \end{bmatrix}$

16. Determine if the set "all polynomial in $\boldsymbol{p_n}$ such that $\boldsymbol{p(0) = 0}$" is a subspace of $\boldsymbol{P_n}$ for an appropriate value of n. Justify your answer.

17. Let $\boldsymbol{H}$ be the set of all vectors of the form $\begin{bmatrix} s \\ 3s \\ 2s \end{bmatrix}$. Find a vector $\boldsymbol{v}$ in $\boldsymbol{R^3}$ such that $\boldsymbol{H}$ = Span $\{\boldsymbol{v}\}$. Why does this show that $\boldsymbol{H}$ is a subspace of $\boldsymbol{R^3}$?

18. Let $\boldsymbol{W}$ be the set of all vectors of the form $\begin{bmatrix} 5b + 2c \\ b \\ c \end{bmatrix}$, where b and c are arbitrary. Find vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ such that $\boldsymbol{W}$ = Span $\{\boldsymbol{u}, \boldsymbol{v}\}$. Why does this show that $\boldsymbol{W}$ is a subspace of $\boldsymbol{R^3}$?

19. Let $\boldsymbol{W}$ be the set of all vectors of the form $\begin{bmatrix} s + 3t \\ s - t \\ 2s - t \\ 4t \end{bmatrix}$. Show that $\boldsymbol{W}$ is a subspace of $\boldsymbol{R^4}$.

20. Let $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \boldsymbol{v_2} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \boldsymbol{v_3} = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}, \text{and } \boldsymbol{w} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}.$

(a) Is $\boldsymbol{w}$ in $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$? How many vectors are in $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$?
(b) How many vectors are in Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$?
(c) Is $\boldsymbol{w}$ in the subspace spanned by $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$? Why?

In exercises 21 and 22, let $\boldsymbol{W}$ be the set of all vectors of the form shown, where a, b and c represent arbitrary real numbers. In each case, either find a set $\boldsymbol{S}$ of vectors that spans $\boldsymbol{W}$ or give an example to show that $\boldsymbol{W}$ is not a vector space.

21. $\begin{bmatrix} 3a+b \\ 4 \\ a-5b \end{bmatrix}$

22. $\begin{bmatrix} a-b \\ b-c \\ c-a \\ b \end{bmatrix}$

23. Show that $\boldsymbol{w}$ is in the subspace of $\boldsymbol{R^4}$ spanned by $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$, where

$$\boldsymbol{w} = \begin{bmatrix} -9 \\ 7 \\ 4 \\ 8 \end{bmatrix}, \boldsymbol{v_1} = \begin{bmatrix} 7 \\ -4 \\ -2 \\ 9 \end{bmatrix}, \boldsymbol{v_2} = \begin{bmatrix} -4 \\ 5 \\ -1 \\ -7 \end{bmatrix}, \boldsymbol{v_3} = \begin{bmatrix} -9 \\ 4 \\ 4 \\ -7 \end{bmatrix}$$

24. Determine if $\boldsymbol{y}$ is in the subspace of $\boldsymbol{R^4}$ spanned by the columns of $\boldsymbol{A}$, where

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

Lecture 21

Null Spaces, Column Spaces, and Linear Transformations

Subspaces arise in as set of all solutions to a system of homogenous linear equations as the set of all linear combinations of certain specified vectors. In this lecture, we compare and contrast these two descriptions of subspaces, allowing us to practice using the concept of a subspace. In applications of linear algebra, subspaces of $\boldsymbol{R}^n$ usually arise in one of two ways:

- as the set of all solutions to a system of homogeneous linear equations or
- as the set of all linear combinations of certain specified vectors.

Our work here will provide us with a deeper understanding of the relationships between the solutions of a linear system of equations and properties of its coefficient matrix.

Null Space of a Matrix

Consider the following system of homogeneous equations:

$$\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -5x_1 + 9x_2 + x_3 &= 0 \end{aligned} \qquad (1)$$

In matrix form, this system is written as $\boldsymbol{Ax = 0}$, where

$$\boldsymbol{A} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \qquad (2)$$

Recall that the set of all $\boldsymbol{x}$ that satisfy (1) is called the solution set of the system (1). Often it is convenient to relate this set directly to the matrix $\boldsymbol{A}$ and the equation $\boldsymbol{Ax = 0}$. We call the set of $\boldsymbol{x}$ that satisfy $\boldsymbol{Ax = 0}$ the **null space** of the matrix $\boldsymbol{A}$. The reason for this name is that if matrix A is viewed as a linear operator that maps points of some vector space V into itself, it can be viewed as mapping all the elements of this solution space of AX = 0 into the null element "0". Thus the null space N of A is that subspace of all vectors in V which are imaged into the null element "0" by the matrix A.

NULL SPACE

Definition The **null space** of an $m \times n$ matrix $\boldsymbol{A}$, written as Nul $\boldsymbol{A}$, is the set of all solutions to the homogeneous equation $\boldsymbol{Ax = 0}$. In set notation,

$$\text{Nul } \boldsymbol{A} = \{\boldsymbol{x} : \boldsymbol{x} \text{ is in } \boldsymbol{R}^n \text{ and } \boldsymbol{Ax = 0}\}$$

OR

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

A more dynamic description of Nul $\boldsymbol{A}$ is the set of all $\boldsymbol{x}$ in $\boldsymbol{R}^n$ that are mapped into the zero vector of $\boldsymbol{R}^m$ via the linear transformation $\boldsymbol{x} \to \boldsymbol{Ax}$, where A is a matrix of transformation. See Figure1

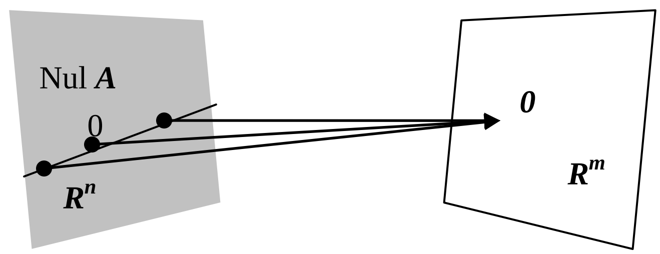


Figure 1

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

Solution To test if $\boldsymbol{u}$ satisfies $\boldsymbol{Au} = \boldsymbol{0}$, 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 $\boldsymbol{u}$ is in Nul $\boldsymbol{A}$.

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.

Theorem 2 The null space of an $m \times n$ matrix A is a subspace of R^n. Equivalently, the set of all solutions to a system $Ax = 0$ of m homogeneous linear equations in n unknowns is a subspace of R^n.

Or simply, the null space is the space of all the vectors of a Matrix A of any order those are mapped (assign) onto zero vector in the space R^n *(i.e.* $Ax = 0$*).*

Proof We know that the subspace of A consists of all the solution to the system $\boldsymbol{Ax = 0}$.

First, we should point out that the zero vector, **0**, in $\boldsymbol{R^n}$ will be a solution to this system and so we know that the null space is not empty. This is a good thing since a vector space (subspace or not) must contain at least one element.

Now we know that the null space is not empty. Consider $\boldsymbol{u}$, $\boldsymbol{v}$ be two any vectors (elements) (in) from the null space and let c be any scalar. We just need to show that the sum ($\boldsymbol{u+v}$) and scalar multiple ($c.\boldsymbol{u}$) of these are also in the null space.

Certainly Nul $\boldsymbol{A}$ is a subset of $\boldsymbol{R^n}$ because $\boldsymbol{A}$ has n columns. To show that Nul(A) is the subspace, we have to check three conditions whether they are satisfied or not. If Nul(A) satisfies the all three condition, we say Nul(A) is a subspace otherwise not.

First, zero vector "0" must be in the space and subspace. If zero vector does not in the space we can not say that is a vector space (generally, we use space for vector space).

And we know that zero vector maps on zero vector so 0 is in Nul(A). Now choose any vectors u, v from Null space and using definition of Null space (i.e. Ax=0)

$$\boldsymbol{Au = 0} \text{ and } \boldsymbol{Av = 0}$$

Now the other two conditions are vector addition and scalar multiplication. For this we proceed as follow:

Let start with vector addition:

To show that $\boldsymbol{u + v}$ is in Nul $\boldsymbol{A}$, we must show that $\boldsymbol{A (u + v) = 0}$. Using the property of matrix multiplication, we find that

$\boldsymbol{A (u + v) = Au + Av = 0 + 0 = 0}$

Thus $\boldsymbol{u + v}$ is in Nul $\boldsymbol{A}$, and Nul $\boldsymbol{A}$ is closed under vector addition.

For Matrix multiplication, consider any scalar , say c,

$\mathbf{A}\ (c\boldsymbol{u}) = c\ (\boldsymbol{Au}) = c\ (\boldsymbol{0}) = \boldsymbol{0}$

which shows that $c\boldsymbol{u}$ is in Nul $\boldsymbol{A}$. Thus Nul $\boldsymbol{A}$ is a subspace of $\boldsymbol{R^n}$.

Example 2 The set $\boldsymbol{H}$, of all vectors in $\boldsymbol{R^4}$ whose coordinates a, b, c, d satisfy the equations

$$a - 2b + 5c = d$$
$$c - a = b$$

is a subspace of $\boldsymbol{R^4}$.

Solution Since

$$a - 2b + 5c = d$$
$$c - a = b$$

By rearranging the equations, we get

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

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 ***Ax*** = ***0***, we obtain an explicit description of Nul ***A***.

Example 3 Find a spanning set for the null space of the matrix

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

Solution The first step is to find the general solution of ***Ax = 0*** in terms of free variables.
After transforming the augmented matrix [***A*** ***0***] to the reduced row echelon form and we get;

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

which corresponds to the system

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

The general solution is

$$\begin{aligned} x_1 &= 2x_2 + x_4 - 3x_5 \\ x_2 &= \text{free variable} \\ x_3 &= -2x_4 + 2x_5 \\ x_4 &= \text{free variable} \\ x_5 &= \text{free variable} \end{aligned}$$

Next, decompose the vector giving the general solution into a linear combination of vectors where the weights are the free variables. That is,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\substack{\uparrow \\ \boldsymbol{u}}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\substack{\uparrow \\ \boldsymbol{v}}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\substack{\uparrow \\ \boldsymbol{w}}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}}$$

$$= x_2\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w} \tag{3}$$

Every linear combination of ***u, v*** and ***w*** is an element of Nul ***A***. Thus $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ 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

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

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

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

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 $\boldsymbol{A}$.

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.

Definition (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

$$\text{Col } \boldsymbol{A} = Span \{\boldsymbol{a_1}, \dots, \boldsymbol{a_n}\}$$

Since Span $\{\boldsymbol{a_1}, \dots, \boldsymbol{a_n}\}$ is a subspace, by Theorem of lecture 20 i.e. if $v_1,...,v_p$ are in a vector space V, then Span $\{v_1,...,v_p\}$ is a subspace of V .

The column space of a matrix is that subspace spanned by the columns of the matrix (columns viewed as vectors). It is that space defined by all linear combinations of the column of the matrix.

Example, in the given matrix,

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 1 & 4 \\ 3 & 1 & 5 \\ 4 & 1 & 6 \end{pmatrix}$$

The column space Col***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 $\boldsymbol{R^m}$.

Theorem 3 The column space of an $m \times n$ matrix A is a subspace of R^m.

Note that a typical vector in Col ***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,

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

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

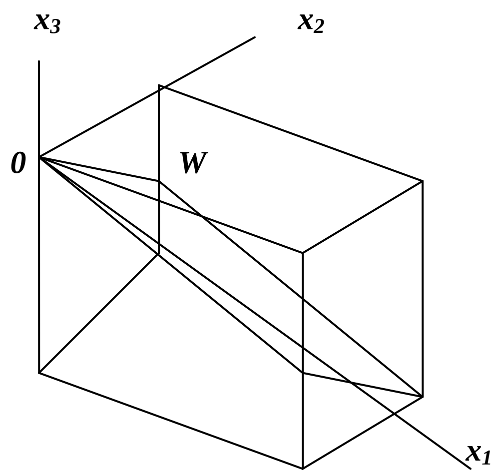
Example 6 Find a matrix ***A*** such that ***W*** **=** Col ***A***. $\boldsymbol{W} = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \text{ in } R \right\}$

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

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

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

General and Particular Solutions: The vector x_0 is called a particular solution of $Ax=b$ .The expression $x_0 + c_1 v_1 + c_2 v_2 + \ldots + c_k v_k$ is called the general solution of $Ax=b$, and the expression $c_1 v_1 + c_2 v_2 + \ldots + c_k v_k$ is called the general solution of $Ax=0$.

Example 7 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 $\boldsymbol{Ax}=\boldsymbol{0}$.

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

Solution 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 $\boldsymbol{A}$?

Solution (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 $\boldsymbol{Ax = 0}$, 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 [$\boldsymbol{A}$ $\boldsymbol{v}$] 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 $\boldsymbol{Ax = v}$ 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 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 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** ***T*** from a vector space ***V*** into a vector space ***W*** is a rule that assigns to each vector ***x*** in ***V*** a unique vector ***T*** (***x***) in ***W***, such that

(i) $T(u + v) = T(u) + T(v)$ for all ***u, v*** in ***V***, and

(ii) $T(cu) = c\,T(u)$ for all ***u*** in ***V*** and all scalars *c*.

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

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

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 ***A*** and the range of T_A is the column space of ***A***.

Remarks The kernel of ***T*** is a subspace of ***V*** and the range of ***T*** is a subspace of ***W***.

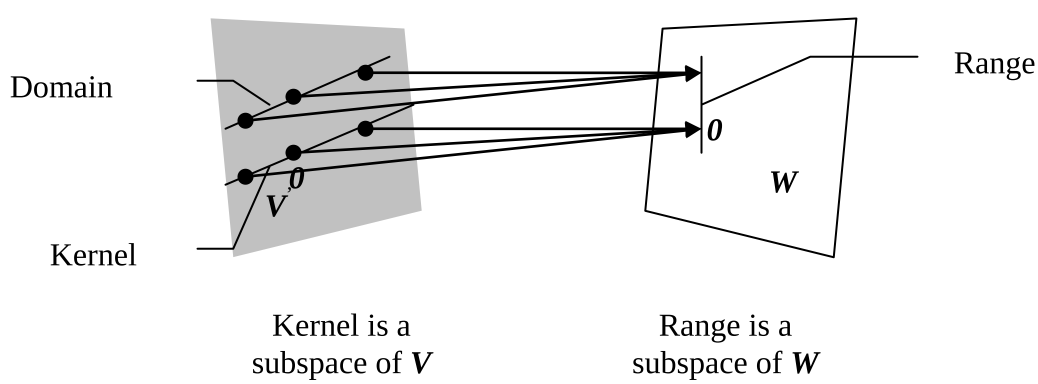


Figure 2 Subspaces associated with a linear transformation.

In applications, a subspace usually arises as either the kernel or the range of an appropriate linear transformation. For instance, the set of all solutions of a homogeneous linear differential equation turns out to be the kernel of a linear transformation. Typically, such a linear transformation is described in terms of one or more derivatives of a function. To explain this in any detail would take us too far a field at this point. So we present only two examples. The first explains why the operation of differentiation is a linear transformation.

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

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

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.

Example 14 Let $\boldsymbol{A} = \begin{bmatrix} 7 & -3 & 5 \\ -4 & 1 & -5 \\ -5 & 2 & -4 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$, *and* $\boldsymbol{W} = \begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix}$

Suppose you know that the equations $\boldsymbol{Ax = v}$ and $\boldsymbol{Ax = w}$ are both consistent. What can you say about the equation $\boldsymbol{Ax = v + w}$?

Solution Both $\boldsymbol{v}$ and $\boldsymbol{w}$ are in Col $\boldsymbol{A}$. Since Col $\boldsymbol{A}$ is a vector space, $\boldsymbol{v + w}$ must be in Col $\boldsymbol{A}$. That is, the equation $\boldsymbol{Ax = v + w}$ is consistent.

Activity

1. Let $\boldsymbol{V}$ and $\boldsymbol{W}$ be any two vector spaces. The mapping $T: V \to W$ 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 $I: V \to V$ defined by $\boldsymbol{I(v) = v}$ 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 $\boldsymbol{T}$: $\boldsymbol{P_2} \rightarrow \boldsymbol{R^2}$ 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 $\boldsymbol{T}$: $\boldsymbol{P_2} \rightarrow \boldsymbol{R^2}$ 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 $\boldsymbol{T}$: $\boldsymbol{M_{2x2}} \rightarrow \boldsymbol{M_{2x2}}$ by $\boldsymbol{T}(\boldsymbol{A}) = \boldsymbol{A} + \boldsymbol{A^T}$, 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 22

Linearly Independent Sets; Bases

First we revise some definitions and theorems from the Vector Space:

Definition Let V be an arbitrary nonempty set of objects on which two operations are defined, addition and multiplication by scalars.

If the following axioms are satisfied by all objects u, v, w in V and all scalars l and m, then we call V a vector space.

Axioms of Vector Space

For any set of vectors u, v, w in V *and scalars* l, m, n:

1. $u + v$ is in V
2. $u + v = v + u$
3. $u + (v + w) = (u + v) + w$
4. There exist a zero vector 0 such that
   $0 + u = u + 0 = u$
5. There exist a vector $-u$ in V such that
   $-u + u = 0 = u + (-u)$
6. $(l\ u)$ is in V
7. $l\ (u + v)= l\ u + l\ v$
8. $m\ (n\ u) = (m\ n)\ u = n\ (m\ u)$
9. $(l + m)\ u= I\ u+ m\ u$
10. $1u = u$ where 1 is the multiplicative identity

Definition A subset W of a vector space V is called a subspace of V if W itself is a vector space under the addition and scalar multiplication defined on V.

Theorem If W is a set of one or more vectors from a vector space V, then W is subspace of V if and only if the following conditions hold:

(a) If u and v are vectors in W, then $u + v$ is in W
(b) If k is any scalar and u is any vector in W, then $k\ u$ is in W.

Definition The null space of an $m \times n$ matrix A $(Nul\ A)$ is the set of all solutions of the hom equation $Ax = 0$

$$Nul\ A = \{x: x \text{ is in } R^n \text{ and } Ax = 0\}$$

Definition The column space of an m x n matrix A $(Col\ A)$ is the set of all linear combinations of the columns of A.
If $A = [a_1 \ \ \dots \ \ a_n]$,
then
$Col\ A = Span\ \{ a_1, \dots , a_n \}$

Since we know that a set of vectors $S = \{v_1, v_2, v_3, ...v_p\}$ spans a given vector space ***V*** if every vector in ***V*** is expressible as a linear combination of the vectors in ***S***. In general there may be more than one way to express a vector in ***V*** as linear combination of vectors in a spanning set. We shall study conditions under which each vector in ***V*** is expressible as a linear combination of the spanning vectors in exactly one way. Spanning sets with this property play a fundamental role in the study of vector spaces.
In this Lecture, we shall identify and study the subspace ***H*** as "efficiently" as possible. The key idea is that of linear independence, defined as in $\boldsymbol{R^n}$.

<u>Definition</u> An ***indexed set*** of vectors $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\}$ 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:

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

$Solving\ (2)\ and\ (4)\ implies:$

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

$\Rightarrow c_1 = c_2 = c_3 = 0\ ; scalars\, c_1, c_2, c_3 \in R\ are\ all\ zero$

$\therefore$ The system has trivial solution.

Hence the given vectors $v_1,\ v_2,\ v_3$ *are linearly independent.*

Example **(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 $3v_1 + v_2 - v_3 = 0$

Example **(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.

Theorem 1 An indexed set $\{\ \boldsymbol{v_1}, \ldots, \boldsymbol{v_p}\ \}$ 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, $\boldsymbol{v_1}, \ldots, \boldsymbol{v_{j-1}}$.

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 $\boldsymbol{A}$ in order to study the equation $\boldsymbol{Ax} = \boldsymbol{0}$. We must rely instead on the definition of linear dependence and on Theorem 1.

Example 1 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 $\boldsymbol{P}$ because $\boldsymbol{p}_3 = 4\boldsymbol{p}_1 - \boldsymbol{p}_2$.

Example 2 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,
$\boldsymbol{H} = Span\ \{\boldsymbol{b_1},...,\boldsymbol{b_p}\}$

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 $\boldsymbol{b_1},...,\boldsymbol{b_p}$ must belong to $\boldsymbol{H}$, because $Span\ \{\boldsymbol{b_1},...,\boldsymbol{b_p}\}$ 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).

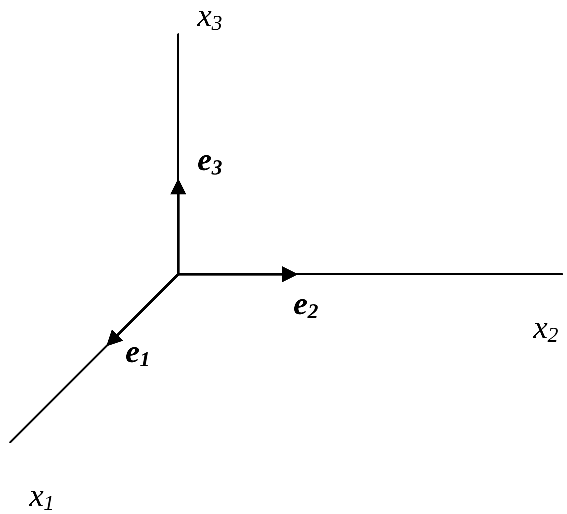


Figure 1 - The standard basis for R^3

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.

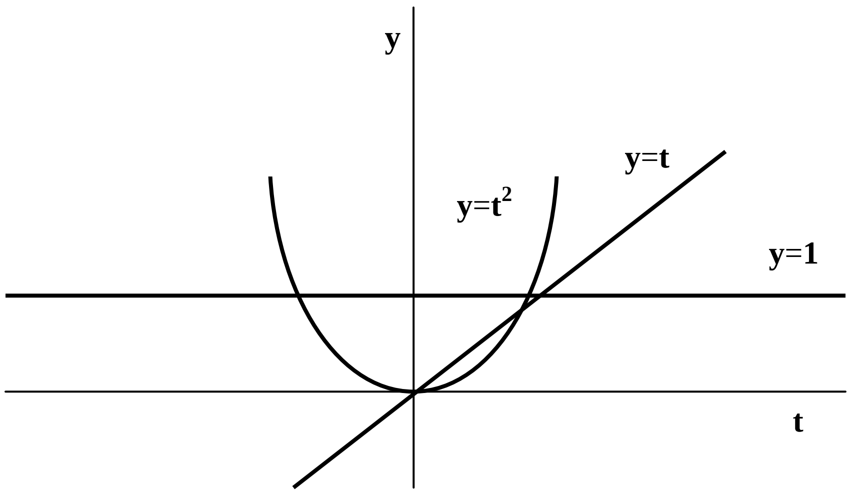


Figure 2 – The standard basis for P_2

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

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

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

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 $\boldsymbol{P_2}$?

Solution The set $\boldsymbol{S} = \{\boldsymbol{p_1(t)}, \boldsymbol{p_2(t)}, \boldsymbol{p_3(t)}\}$ of vectors in $\boldsymbol{P_2}$ spans $\boldsymbol{V} = \boldsymbol{P_2}$ if

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

with $\boldsymbol{q_1(t)} = \mathbf{1} + 0\,\mathbf{t} + 0\,\mathbf{t}^2$, $\boldsymbol{q_2(t)} = \mathbf{0} + 1\,\mathbf{t} + 0\,\mathbf{t}^2$, $\boldsymbol{q_3(t)} = \mathbf{0} + 0\,\mathbf{t} + 1\,\mathbf{t}^2$ 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 $\boldsymbol{p_1(t)}, \boldsymbol{p_2(t)}, \boldsymbol{p_3(t)}$, (*) 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

$$(-4\,c_1 + 6\,c_2 + 8\,c_3)\mathbf{1} + (1\,c_1 + 5\,c_2 + 4\,c_3)\,\mathbf{t} + (3\,c_1 + 2\,c_2 + 1\,c_3)\,\mathbf{t}^2 =$$
$$(1\,d_1 + 0\,d_2 + 0\,d_3)\mathbf{1} + (0\,d_1 + 1\,d_2 + 0\,d_3)\,\mathbf{t} + (0\,d_1 + 0\,d_2 + 1\,d_3)\,\mathbf{t}^2$$

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:

$$\begin{aligned} -4\,c_1 + 6\,c_2 + 8\,c_3 &= 1\,d_1 + 0\,d_2 + 0\,d_3 \\ 1\,c_1 + 5\,c_2 + 4\,c_3 &= 0\,d_1 + 1\,d_2 + 0\,d_3 \\ 3\,c_1 + 2\,c_2 + 1\,c_3 &= 0\,d_1 + 0\,d_2 + 1\,d_3 \end{aligned} \qquad \text{(A)}$$

$$\Rightarrow \begin{bmatrix} -4 & 6 & 8 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

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

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

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:

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

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***.

<u>Example 9</u> 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.

<u>Solution</u> 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

<u>Example 10</u> 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 ***S*** spans the space ***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 ***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 $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}$, 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 $\quad \det(A) = 48 \neq 0$

Therefore the set $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}\}$ is linearly independent. Consequently, the set ***S*** forms a basis for span ***S***.

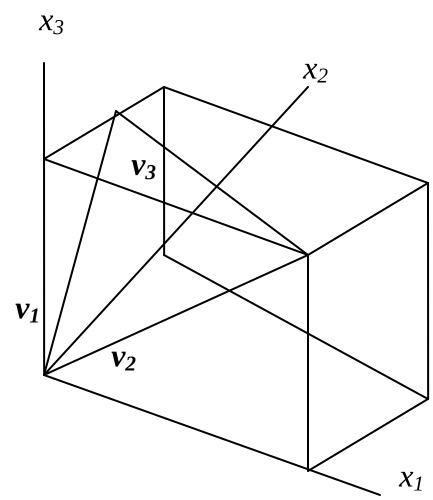
<u>Example 11</u> Let $\boldsymbol{v_1} = \begin{bmatrix}1\\-2\\-3\end{bmatrix}$, $\boldsymbol{v_2} = \begin{bmatrix}-3\\5\\7\end{bmatrix}$, $\boldsymbol{v_3} = \begin{bmatrix}-4\\5\\6\end{bmatrix}$, and $\boldsymbol{H} = Span\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$.

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 ***H***.

<u>Solution</u>

Every vector in Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ belongs to ***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.

$$v_1 = (1,\ 2,\ 3),\ v_2 = (0,\ 1,\ 1),\ v_1 = (0,\ 1,\ 3)$$

<u>The Spanning Set Theorem</u>

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.

<u>Theorem 2</u> **(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.

Remark Let $\boldsymbol{V} = \boldsymbol{R}^m$ and let $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ be a set of nonzero vectors in $\boldsymbol{V}$.

Procedure

The procedure for finding a subset of S that is a basis for W = span S is as follows:

Step 1 Write the Equation,

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_n\,\boldsymbol{v_n} = 0 \qquad (3)$$

Step 2 Construct the augmented matrix associated with the homogeneous system of Equation (1) and transforms it to reduced row echelon form.

Step 3 The vectors corresponding to the columns containing the leading 1's form a basis for $\boldsymbol{W}$ = span $\boldsymbol{S}$.

Thus if $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_6}\}$ and the leading 1's occur in columns $1, 3,$ and 4, then $\{\boldsymbol{v_1}, \boldsymbol{v_3}, \boldsymbol{v_4}\}$ is a basis for span $\boldsymbol{S}$.

Note In step 2 of the procedure above, it is sufficient to transform the augmented matrix to row echelon form.

Example 12 Let $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}, \boldsymbol{v_5}\}$ be a set of vectors in $\boldsymbol{R}^4$, where $\boldsymbol{v_1} = (1,2,-2,1)$, $\boldsymbol{v_2} = (-3,0,-4,3)$, $\boldsymbol{v_3} = (2,1,1,-1)$, $\boldsymbol{v_4} = (-3,3,-9,6)$, and $\boldsymbol{v_5} = (9,3,7,-6)$.

Find a subset of $\boldsymbol{S}$ that is a basis for $\boldsymbol{W}$ = span $\boldsymbol{S}$.

Solution Step 1 Form Equation (3),

$c_1\,(1,2,-2,1) + c_2(-3,0,-4,3) + c_3(2,1,1,-1) + c_4(-3,3,-9,6) + c_5(9,3.7,-6) = (0,0,0,0)$.

Step 2 Equating corresponding components, we obtain the homogeneous system

$$\begin{aligned}
c_1 - 3c_2 + 2c_3 - 3c_4 + 9c_5 &= 0\\
2c_1 + c_3 + 3c_4 + 3c_5 &= 0\\
-2c_1 - 4c_2 + c_3 - 9c_4 + 7c_5 &= 0\\
c_1 + 3c_2 - c_3 + 6c_4 - 6c_5 &= 0
\end{aligned}$$

The reduced row echelon form of the associated augmented matrix is

$$\begin{bmatrix}
1 & 0 & ½ & 3/2 & 3/2 & : & 0\\
0 & 1 & -1/2 & 3/2 & -5/2 & : & 0\\
0 & 0 & 0 & 0 & 0 & : & 0\\
0 & 0 & 0 & 0 & 0 & : & 0
\end{bmatrix}$$

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 $\boldsymbol{R}^4 = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}\}$. Explain why $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}\}$ is a basis for $\boldsymbol{R}^4$.

8. Explain why the following sets of vectors are not bases for the indicated vector spaces. (Solve this problem by inspection).
(a) $\boldsymbol{u_1} = (1, 2)$, $\boldsymbol{u_2} = (0, 3)$, $\boldsymbol{u_3} = (2, 7)$ for $\boldsymbol{R}^2$
(b) $\boldsymbol{u_1} = (-1, 3, 2)$, $\boldsymbol{u_2} = (6, 1, 1)$ for $\boldsymbol{R}^3$
(c) $\boldsymbol{p_1} = \boldsymbol{1} + \boldsymbol{x} + \boldsymbol{x}^2$, $\boldsymbol{p_2} = \boldsymbol{x} - \boldsymbol{1}$ for $\boldsymbol{P_2}$
(d) $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 6 & 0 \\ -1 & 4 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 3 & 0 \\ 1 & 7 \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}, \boldsymbol{E} = \begin{bmatrix} 7 & 1 \\ 2 & 9 \end{bmatrix}$ for $\boldsymbol{M_{22}}$

9. Which of the following sets of vectors are bases for $\boldsymbol{R}^2$?
(a) (2, 1), (3, 0) (b) (4, 1), (-7, -8) (c) (0, 0), (1, 3) (d) (3, 9), (-4, -12)

10. Let $\boldsymbol{V}$ be the space spanned by $\boldsymbol{v_1} = \text{Cos}^2 x$, $\boldsymbol{v_2} = \text{Sin}^2 x$, $\boldsymbol{v_3} = \cos 2x$.
(a) Show that $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ is not a basis for $\boldsymbol{V}$ (b) Find a basis for $\boldsymbol{V}$

In exercises 11-13, determine a basis for the solution space of the system.

11. $$\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ -2x_1 - x_2 + 2x_3 &= 0 \\ -x_1 + x_3 &= 0 \end{aligned}$$

12. $$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ x_1 + 5x_3 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$$

13.
$$\begin{aligned} x + \ \ y + z &= 0 \\ 3x + 2y - 2z &= 0 \\ 4x + 3y - \ \ z &= 0 \\ 6x + 5y + \ \ z &= 0 \end{aligned}$$

14. Determine bases for the following subspace of $\boldsymbol{R}^3$
(a) the plane 3x – 2y + 5z = 0 (b) the plane x – y = 0
(c) the line x = 2t, y = -t, z = 4t (d) all vectors of the form (a, b, c), where b = a + c

15. Find a standard basis vector that can be added to the set $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ to produce a basis for $\boldsymbol{R}^3$.
(a) $\boldsymbol{v_1}$ = (-1, 2, 3), $\boldsymbol{v_2}$ = (1, -2, -2) (b) $\boldsymbol{v_1}$ = (1, -1, 0), $\boldsymbol{v_2}$ = (3, 1, -2)

16. Find a standard basis vector that can be added to the set $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ to produce a basis for $\boldsymbol{R}^4$.
$\boldsymbol{v_1}$ = (1, -4, 2, -3), $\boldsymbol{v_2}$ = (-3, 8, -4, 6)

Lecture No.23
Coordinate System

OBJECTIVES

The objectives of the lecture are to learn about:

- Unique representation theorem.
- Coordinate of the element of a vector space relative to the basis B.
- Some examples in which B- coordinate vector is uniquely determined using basis of a vector space.
- Graphical interpretation of coordinates.
- Coordinate Mapping

Theorem

Let $B=\{b_1,b_2,...,b_n\}$ be a basis for a vector space V. Then for each x in *V*, there exist a unique set of scalars $c_1,c_2,...,c_n$ such that

$$x=c_1b_1,c_2b_2,...,c_nb_n \text{ …………… (1)}$$

Proof

Since *B* is a basis for a vector space *V,* then by definition of basis every element of *V* can be written as a linear combination of basis vectors. That is if $x\in V$, then $x=c_1b_1,c_2b_2,...,c_nb_n$. Now, we show that this representation for x is unique.

For this, suppose that we have two representations for x .

i.e.

$$x=c_1b_1,c_2b_2,...,c_nb_n \text{ …………… (2)}$$

and

$$x=d_1b_1,d_2b_2,...,d_nb_n \text{ ………….. (3)}$$

We will show that the coefficients are actually equal. To do this, subtracting (3) from (2), we have

$$0=(c_1-d_1)b_1+(c_2-d_2)b_2+...+(c_n-d_n)b_n .$$

Since *B* is a basis, it is linearly independent set. Thus the coefficients in the last linear combination must all be zero. That is

$$c_1=d_1,...,c_n=d_n .$$

Thus the representation for x is unique.

Definition (B-Coordinate of x)

Suppose that the set $B=\{b_1,b_2,...,b_n\}$ is a basis for *V* and x is in *V*. The coordinates of x relative to basis *B* (**or the *B*-coordinate of** x) are the weights $c_1,c_2,...,c_n$ such that

$$x=c_1b_1,c_2b_2,...,c_nb_n .$$

Note

If $c_1, c_2, ..., c_n$ are the ***B***- coordinates of x, then the vector in R^n, $[x]_B = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ c_n \end{bmatrix}$ is the coordinate vector of x (relative to ***B)*** or ***B***- coordinates of x.

Example 1

Consider a basis $B = \{b_1, b_2\}$ for R^2, where $b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Suppose an x in R^2 has the coordinate vector $[x]_B = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Find x

Solution

Using above definition x is uniquely determined using coordinate vector and the basis. That is

$$
\begin{aligned}
x &= c_1 b_1, c_2 b_2 \\
&= (-2)b_1 + (3)b_2 \\
&= (-2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 6 \end{bmatrix} \\
x &= \begin{bmatrix} 1 \\ 6 \end{bmatrix}
\end{aligned}
$$

Example 2

Let ***S* = {*v*₁, *v*₂, *v*₃}** be the basis for $\boldsymbol{R^3}$, where $\boldsymbol{v_1}$ = (1, 2, 1), $\boldsymbol{v_2}$ = (2, 9, 0), and $\boldsymbol{v_3}$ = (3, 3, 4).

(a) Find the coordinates vector of ***v*** = (5, -1, 9) with respect to ***S***.

(b) Find the vector ***v*** in $\boldsymbol{R^3}$ whose coordinate vector with respect to the basis ***S*** is

$[v]_S =$ (-1, 3, 2)

Solution

Since **S** is a basis for $\mathbf{R^3}$, Thus

$$x = c_1 v_1 + c_2 v_2 + c_3 v_3.$$

Further

$(5,-1,9) = c_1(1,2,1) + c_2(2,9,0) + c_3(3,3,4)$ ………….. (A)

To find the coordinate vector of ***v***, we have to find scalars c_1, c_2, c_3.

For this equating corresponding components in (A) gives

$c_1 + 2c_2 + 3c_3 = 5 \qquad (1)$

$2c_1 + 9c_2 + 3c_3 = -1 \qquad (2)$

$c_1 + \qquad 4c_3 = 9 \qquad (3)$

Now find values of c_1, c_2 *and* c_3 *from these equations.*

From equation (3)

$c_1 = 9 - 4c_3$

Put this value of c_1 *in equations* (1) *and* (2)

$9 - 4c_3 + 2c_2 + 3c_3 = 5$

$2c_2 - c_3 = -4 \qquad (4)$

and

$2(9 - 4c_3) + 9c_2 + 3c_3 = -1$

$18 - 8c_3 + 9c_2 + 3c_3 = -1$

$9c_2 - 5c_3 = -19 \qquad (5)$

Multiply equation (4) *by 5*

$10c_2 - 5c_3 = -20$

Subtract equation (5) *from above equation*

$$\begin{array}{l} 10c_2 - 5c_3 = -20 \\ \pm 9c_2 \mp 5c_3 = \mp 19 \\ \hline c_2 = -1 \end{array}$$

Put value of c_2 *in equation* (4) *to get* c_3

$2(-1) - c_3 = -4$

$-2 - c_3 = -4$

$c_3 = 4 - 2 = 2$

Put value of c_3 *in equation* (3) *to get* c_1

$c_1 + 4(2) = 9$

$c_1 = 9 - 8 = 1$

Thus, we obtain $c_1 = 1, c_2 = -1, c_3 = 2$

Therefore, $[\boldsymbol{v}]_S = (1, -1, 2)$

Using the definition of coordinate vector, we have

$$\begin{aligned} v &= c_1v_1 + c_2v_2 + c_3v_3 \\ &= (-1)v_1 + 3v_2 + 2v_3 \\ &= (-1)(1,2,1) + 3(2,9,0) + 2(3,3,4) \\ &= (-1+6+6, -2+27+6, -1+0+8) \\ &= (11,31,7) \end{aligned}$$

Therefore

$$v = (11,31,7)$$

Example 3

Find the coordinates vector of the polynomial $\boldsymbol{p} = \boldsymbol{a_0} + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2$ relative to the basis $\boldsymbol{S} = \{\boldsymbol{1}, \boldsymbol{x}, \boldsymbol{x}^2\}$ for $\boldsymbol{p_2}$.

Solution

To find the coordinator vector of the polynomial p, we write it as a linear combination of the basis set S. That is

$$a_0 + a_1x + a_2x^2 = c_1(1) + c_2(x) + c_3(x^2)$$
$$\Rightarrow c_1 = a_0, c_2 = a_1, c_3 = a_2$$

Therefore

$$[p]_s = (a_0, a_1, a_2)$$

Example 4

Find the coordinates vector of the polynomial $\boldsymbol{p} = \boldsymbol{5} - 4\boldsymbol{x} + 3\boldsymbol{x}^2$ relative to the basis $\boldsymbol{S} = \{\boldsymbol{1}, \boldsymbol{x}, \boldsymbol{x}^2\}$ for $\boldsymbol{p_2}$.

Solution

To find the coordinator vector of the polynomial p, we write it as a linear combination of the basis set S. That is

$$5 - 4x + 3x^2 = c_1(1) + c_2(x) + c_3(x^2)$$
$$\Rightarrow c_1 = 5, c_2 = -4, c_3 = 3$$

Therefore

$$[p]_s = (5,-4,3)$$

Example 5

Find the coordinate vector of A relative to the basis S = {A1, A2, A3, A4}

$$\boldsymbol{A} = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix}; \boldsymbol{A_1} = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}; \boldsymbol{A_2} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}; \boldsymbol{A_3} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}; \boldsymbol{A_4} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Solution

To find the coordinator vector of A, we write it as a linear combination of the basis set S. That is

$\boldsymbol{A} = c_1 \boldsymbol{A_1} + c_2 \boldsymbol{A_2} + c_3 \boldsymbol{A_3} + c_4 \boldsymbol{A_4}$

$$\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} = c_1 \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -c_1 & c_1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} c_2 & c_2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c_3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c_4 \end{bmatrix}$$

$$= \begin{bmatrix} -c_1 + c_2 + 0 + 0 & c_1 + c_2 + 0 + 0 \\ 0 + 0 + c_3 + 0 & 0 + 0 + 0 + c_4 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} -c_1 + c_2 & c_1 + c_2 \\ c_3 & c_4 \end{bmatrix}$$

$$-c_1 + c_2 = 2 \qquad (1)$$

$$c_1 + c_2 = 0 \qquad (2)$$

$$c_3 = -1 \qquad (3)$$

$$c_4 = 3 \qquad (4)$$

Adding (1) and (2), gives

$2c_2 = 2 \Rightarrow c_2 = 1$

Putting the value of c_2 in (2) to get c_1, $c_1 = -1$

So $c_1 = -1, c_2 = 1, c_3 = -1, c_4 = 3$

Therefore, $[\boldsymbol{v}]_S = (-1, 1, -1, 3)$

Graphical Interpretation of Coordinates

A coordinate system on a set consists of a one-to-one mapping of the points in the set into $\boldsymbol{R^n}$. For example, ordinary graph paper provides a coordinate system for the plane when one selects perpendicular axes and a unit of measurement on each axis. Figure 1 shows the standard basis $\{\boldsymbol{e_1}, \boldsymbol{e_2}\}$, the vectors
$\boldsymbol{b_1}$ (= $\boldsymbol{e_1}$) and $\boldsymbol{b_2}$ from Example 1, that is,
$\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Vector $\boldsymbol{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$, the coordinates 1 and 6 give the location of $\boldsymbol{x}$ relative to the standard basis: 1 unit in the $\boldsymbol{e_1}$ direction and 6 units in the $\boldsymbol{e_2}$ direction.

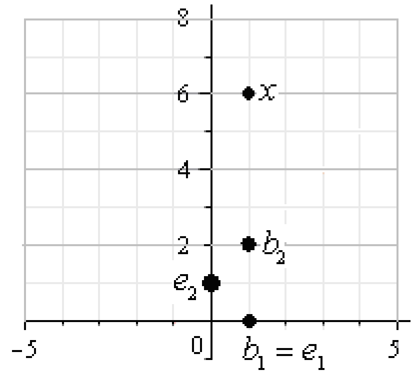


Figure 1

Figure 2 shows the vectors $\boldsymbol{b_1}$, $\boldsymbol{b_2}$, and $\boldsymbol{x}$ from Figure 1. (Geometrically, the three vectors lie on a vertical line in both figures.) However, the standard coordinate grid was erased and replaced by a grid especially adapted to the basis $\boldsymbol{B}$ in Example 1. The coordinate vector $[\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ gives the location of $\boldsymbol{x}$ on this new coordinate system: – 2 units in the $\boldsymbol{b_1}$ direction and *3* units in the $\boldsymbol{b_2}$ direction.

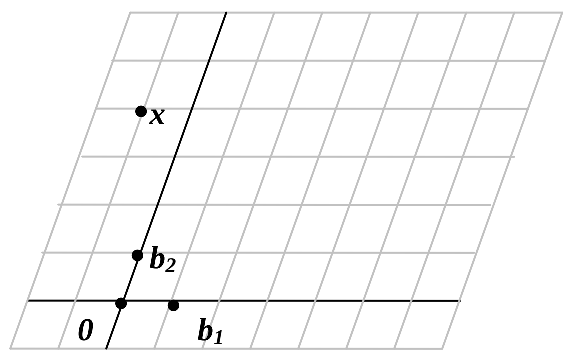
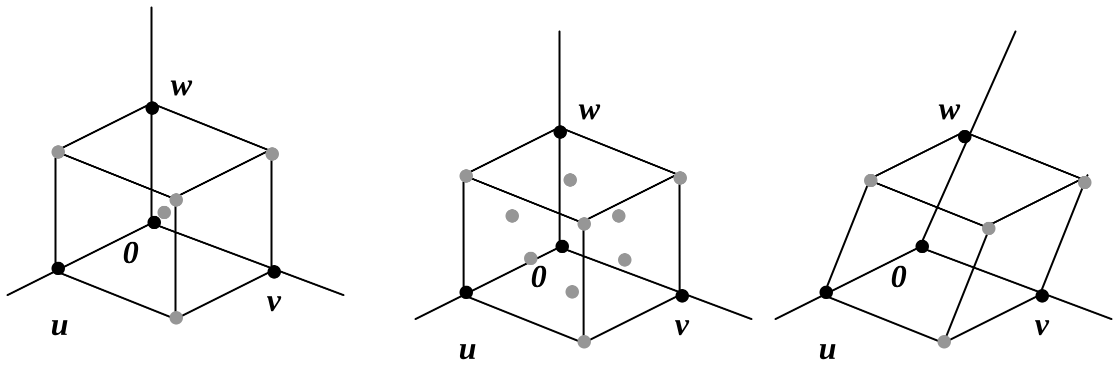


Figure 2

<u>Example 6</u>

In crystallography, the description of a crystal lattice is aided by choosing a basis $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ for $\boldsymbol{R^3}$ that corresponds to three adjacent edges of one "unit cell" of the crystal. An entire lattice is constructed by stacking together many copies of one cell. There are fourteen basic types of unit cells; three are displayed in Figure 3.



(a) Body centered cubic

(b) Face centered orthorhombic

(c) Simple monoclinic

Figure 3 – Examples of unit cells

The coordinates of atoms within the crystal are given relative to the basis for the lattice. For instance, $\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$ identifies the top face-centered atom in the cell in Figure 3(b).

Coordinates in R^n When a basis B for R^n is fixed, the B-coordinate vector of a specified x is easily found, as in the next example.

<u>**Example 7**</u> Let $b_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, x = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, and $B = \{b_1, b_2\}$.

Find the coordinate vector $[x]_B$ of x relative to B.

Solution The B – coordinates c_1, c_2 of x satisfy

$$\underbrace{c_1\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{b_1} + \underbrace{c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}}_{b_2} = \underbrace{\begin{bmatrix} 4 \\ 5 \end{bmatrix}}_{x}$$

or

$$\underbrace{\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}}_{b_1 \quad b_2}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 4 \\ 5 \end{bmatrix}}_{x} \qquad (3)$$

Now, inverse of matrix $\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{3} & \frac{2}{3} \end{bmatrix}$

From equation (3) we get

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3}(4) + \frac{1}{3}(5) \\ \frac{-1}{3}(4) + \frac{2}{3}(5) \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Thus, $c_1 = 3$, $c_2 = 2$.
(Equation (3) can also be solved by row operations on an augmented matrix. Try it yourself)

Thus $\boldsymbol{x} = 3\boldsymbol{b_1} + 2\boldsymbol{b_2}$ and $[\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

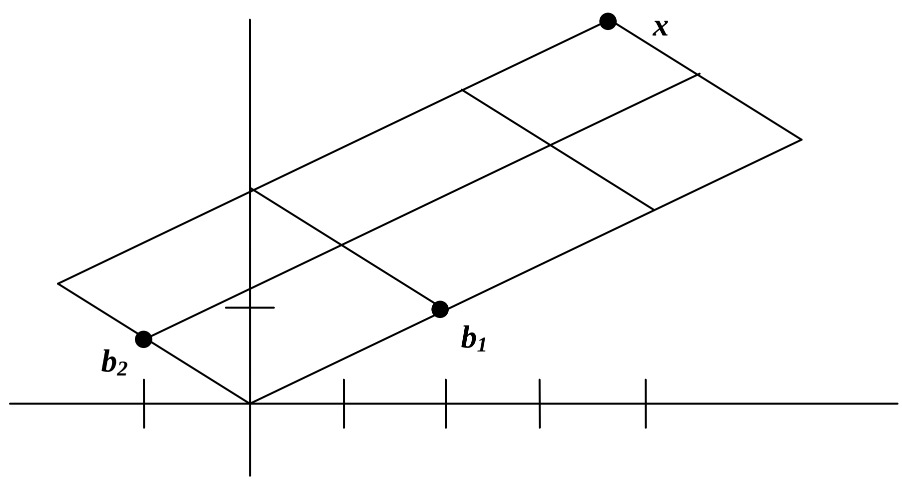


Figure 4 – The *B*-coordinate vector of *x* is (3,2)

The matrix in (3) changes the ***B***-coordinates of a vector ***x*** into the standard coordinates for ***x***. An analogous change of coordinates can be carried out in $\boldsymbol{R^n}$ for a basis
$\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \dots, \boldsymbol{b_n}\}$.

Let $\boldsymbol{P_B} = [\boldsymbol{b_1} \quad \boldsymbol{b_2} \quad \dots \quad \boldsymbol{b_n}]$

Then the vector equation $\boldsymbol{x} = c_1\boldsymbol{b_1} + c_2\boldsymbol{b_2} + \dots + c_n\boldsymbol{b_n}$

is equivalent to $\boldsymbol{x} = \boldsymbol{P_B}[\boldsymbol{x}]_{\boldsymbol{B}}$ (4)

We call $\boldsymbol{P_B}$ the **change-of-coordinates matrix** from ***B*** to the standard basis in $\boldsymbol{R^n}$.
Left-multiplication by $\boldsymbol{P_B}$ transforms the coordinate vector $[\boldsymbol{x}]_{\boldsymbol{B}}$ into ***x***. The change-of-coordinates equation (4) is important and will be needed at several points in next lectures.

Since the columns of $\boldsymbol{P_B}$ form a basis for $\boldsymbol{R^n}$, $\boldsymbol{P_B}$ is invertible (by the Invertible Matrix Theorem). Left-multiplication by $\boldsymbol{P_B^{-1}}$ converts $\boldsymbol{x}$ into its $\boldsymbol{B}$-coordinate vector:

$$\boldsymbol{P_B^{-1}x} = [\boldsymbol{x}]_{\boldsymbol{B}}$$

The correspondence $\boldsymbol{x} \to [\boldsymbol{x}]_{\boldsymbol{B}}$ produced here by $\boldsymbol{P_B^{-1}}$, is the coordinate mapping mentioned earlier. Since $\boldsymbol{P_B^{-1}}$ is an invertible matrix, the coordinate mapping is a one-to-one linear transformation from $\boldsymbol{R^n}$ onto $\boldsymbol{R^n}$, by the Invertible Matrix Theorem. (See also Theorem 3 in lecture 10) This property of the coordinate mapping is also true in a general vector space that has a basis, as we shall see.

The Coordinate Mapping Choosing a basis $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \dots, \boldsymbol{b_n}\}$ for a vector space $\boldsymbol{V}$ introduces a coordinate system in $\boldsymbol{V}$. The coordinate mapping $\boldsymbol{x} \to [\boldsymbol{x}]_{\boldsymbol{B}}$ connects the possibly unfamiliar space $\boldsymbol{V}$ to the familiar space $\boldsymbol{R^n}$. See Figure 5. Points in $\boldsymbol{V}$ can now be identified by their new "names".

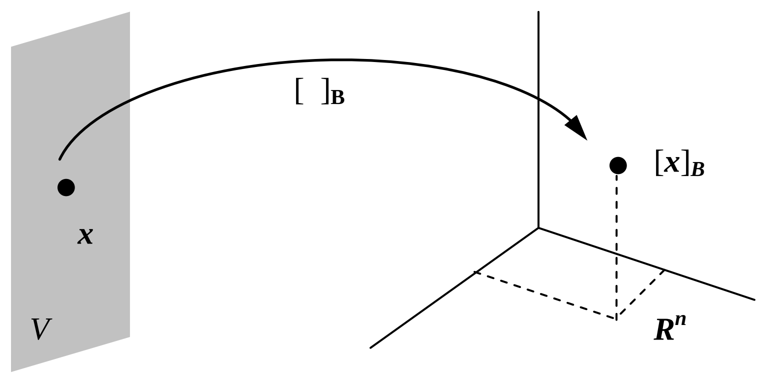


Figure 5 – The coordinate mapping from $\boldsymbol{V}$ onto $\boldsymbol{R^n}$

Theorem 2 Let $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \dots, \boldsymbol{b_n}\}$ be a basis for a vector space $\boldsymbol{V}$. Then the coordinate mapping $\boldsymbol{x} \to [\boldsymbol{x}]_{\boldsymbol{B}}$ is a one-to-one linear transformation from $\boldsymbol{V}$ onto $\boldsymbol{R^n}$.

Proof Take two typical vectors in $\boldsymbol{V}$, say

$$\boldsymbol{u} = c_1\boldsymbol{b_1} + c_2\boldsymbol{b_2} + \dots + c_n\boldsymbol{b_n}$$

$$\boldsymbol{w} = d_1\boldsymbol{b_1} + d_2\boldsymbol{b_2} + \dots + d_n\boldsymbol{b_n}$$

Then, using vector operations, $\boldsymbol{u} + \boldsymbol{w} = (c_1 + d_1)\boldsymbol{b_1} + (c_2 + d_2)\boldsymbol{b_2} + \dots + (c_n + d_n)\boldsymbol{b_n}$

It follows that $[\boldsymbol{u} + \boldsymbol{w}]_{\boldsymbol{B}} = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [\boldsymbol{u}]_{\boldsymbol{B}} + [\boldsymbol{w}]_{\boldsymbol{B}}$

Thus the coordinate mapping preserves addition. If r is any scalar, then

$$r\boldsymbol{u} = r(c_1\boldsymbol{b_1} + c_2\boldsymbol{b_2} + \cdots + c_n\boldsymbol{b_n}) = (rc_1)\boldsymbol{b_1} + (rc_2)\boldsymbol{b_2} + \cdots + (rc_n)\boldsymbol{b_n}$$

So $$[r\boldsymbol{u}]_{\boldsymbol{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[\boldsymbol{u}]_{\boldsymbol{B}}$$

Thus the coordinate mapping also preserves scalar multiplication and hence is a linear transformation. It can be verified that the coordinate mapping is one-to-one and maps $\boldsymbol{V}$ onto $\boldsymbol{R^n}$.

The linearity of the coordinate mapping extends to linear combinations, just as in lecture 9. If $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_p}$ are in $\boldsymbol{V}$ and if $c_1, c_2, \ldots, c_p$ are scalars, then

$$[c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \ldots + c_p\boldsymbol{u_p}]_{\boldsymbol{B}} = c_1[\boldsymbol{u_1}]_{\boldsymbol{B}} + c_2[\boldsymbol{u_2}]_{\boldsymbol{B}} + \ldots + c_p[\boldsymbol{u_p}]_{\boldsymbol{B}} \qquad (5)$$

In words, (5) says that the $\boldsymbol{B}$-coordinate vector of a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_p}$ is the same linear combination of their coordinate vectors.

The coordinate mapping in Theorem 2 is an important example of an isomorphism from $\boldsymbol{V}$ onto $\boldsymbol{R^n}$. In general, a one-to-one linear transformation from a vector space $\boldsymbol{V}$ onto a vector space $\boldsymbol{W}$ is called an **isomorphism** from $\boldsymbol{V}$ onto $\boldsymbol{W}$ (iso from the Greek for “the same”, and morph from the Greek for “form” or “structure”). The notation and terminology for $\boldsymbol{V}$ and $\boldsymbol{W}$ may differ, but the two spaces are indistinguishable as vector spaces. Every vector space calculation in $\boldsymbol{V}$ is accurately reproduced in $\boldsymbol{W}$, and vice versa.

Example 8 Let $\boldsymbol{B}$ be the standard basis of the space $\boldsymbol{P_3}$ of polynomials; that is, let $\boldsymbol{B} = \{\boldsymbol{1}, \boldsymbol{t}, \boldsymbol{t^2}, \boldsymbol{t^3}\}$. A typical element $\boldsymbol{p}$ of $\boldsymbol{P_3}$ has the form $\boldsymbol{p}(\boldsymbol{t}) = \boldsymbol{a_0} + a_1\boldsymbol{t} + a_2\boldsymbol{t^2} + a_3\boldsymbol{t^3}$
Since $\boldsymbol{p}$ is already displayed as a linear combination of the standard basis vectors, we conclude that $[\boldsymbol{p}]_{\boldsymbol{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$. Thus the coordinate mapping $\boldsymbol{p} \to [\boldsymbol{p}]_{\boldsymbol{B}}$ is an isomorphism from $\boldsymbol{P_3}$ onto $\boldsymbol{R^4}$. All vector space operations in $\boldsymbol{P_3}$ correspond to operations in $\boldsymbol{R^4}$.

If we think of $\boldsymbol{P_3}$ and $\boldsymbol{R^4}$ as displays on two computer screens that are connected via the coordinate mapping, then every vector space operation in $\boldsymbol{P_3}$ on one screen is exactly duplicated by a corresponding vector operation in $\boldsymbol{R^4}$ on the other screen. The vectors on the $\boldsymbol{P_3}$ screen look different from those on the $\boldsymbol{R^4}$ screen, but they “act” as vectors in exactly the same way. See Figure 6.

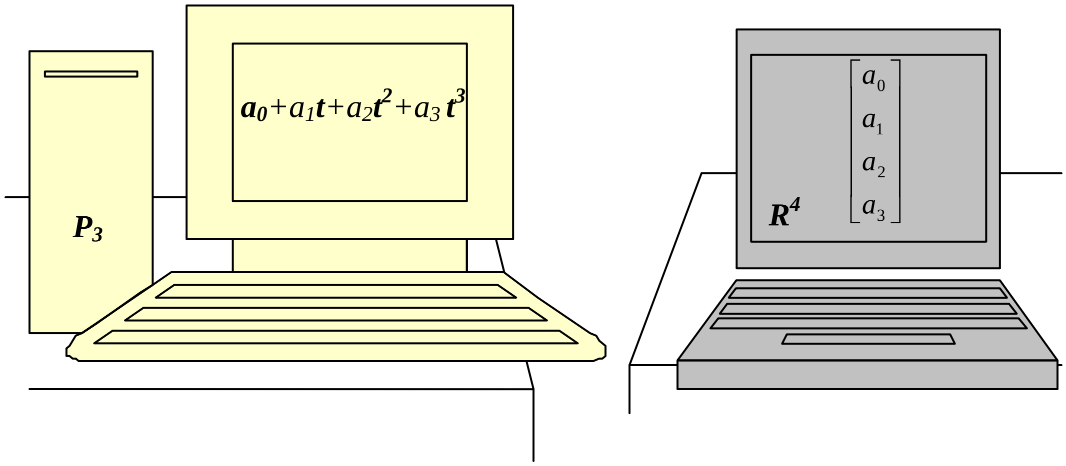


Figure 6 – The space P_3 is isomorphic to R^4

Example 9 Use coordinate vector to verify that the polynomials $\boldsymbol{1} + 2\boldsymbol{t}^2$, $\boldsymbol{4} + \boldsymbol{t} + 5\boldsymbol{t}^2$ and $\boldsymbol{3} + 2\boldsymbol{t}$ are linearly dependent in $\boldsymbol{P_2}$.

Solution The coordinate mapping from Example 8 produces the coordinate vectors $(1, 0, 2)$, $(4, 1, 5)$ and $(3, 2, 0)$, respectively. Writing these vectors as the columns of a matrix $\boldsymbol{A}$, we can determine their independence by row reducing the augmented matrix for $\boldsymbol{Ax} = \boldsymbol{0}$: $\begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 5 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

The columns of $\boldsymbol{A}$ are linearly dependent, so the corresponding polynomials are linearly dependent. In fact, it is easy to check that column 3 of $\boldsymbol{A}$ is 2 times column 2 minus 5 times column 1. The corresponding relation for the polynomials is

$$\boldsymbol{3} + 2\boldsymbol{t} = 2(\boldsymbol{4} + \boldsymbol{t} + 5\boldsymbol{t}^2) - 5(\boldsymbol{1} + 2\boldsymbol{t}^2)$$

Example 10 Let $\boldsymbol{v_1} = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $\boldsymbol{v_2} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$, and $\boldsymbol{B} = \{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. Then $\boldsymbol{B}$ is a basis for $\boldsymbol{H} = Span\ \{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. Determine if $\boldsymbol{x}$ is in $\boldsymbol{H}$ and if it is, find the coordinate vector of $\boldsymbol{x}$ relative to $\boldsymbol{B}$.

Solution If $\boldsymbol{x}$ is in $\boldsymbol{H}$, then the following vector equation is consistent.

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

The scalars, c_1 and c_2, if they exist, are the $\boldsymbol{B}$ – coordinates of $\boldsymbol{x}$.

Using row operations, we obtain $\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$.

Thus $c_1 = 2$, $c_2 = 3$ and $[\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. The coordinate system on $\boldsymbol{H}$ determined by $\boldsymbol{B}$ is shown in Figure 7.

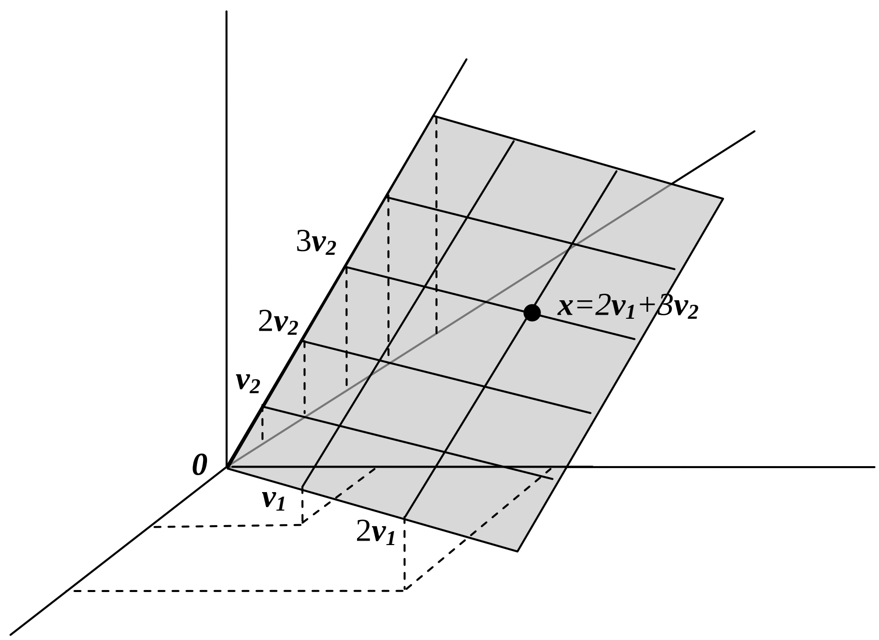


Figure 7 – A coordinate system on a plane *H* in $\boldsymbol{R}^3$

If a different basis for $\boldsymbol{H}$ were chosen, would the associated coordinate system also make $\boldsymbol{H}$ isomorphic to $\boldsymbol{R}^2$? Surely, this must be true. We shall prove it in the next lecture.

Example 11 Let $\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\boldsymbol{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}$, $\boldsymbol{b}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}$, and $\boldsymbol{x} = \begin{bmatrix} -8 \\ 2 \\ 3 \end{bmatrix}$.

a. Show that the set $\boldsymbol{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$ is a basis of $\boldsymbol{R}^3$.
b. Find the change-of-coordinates matrix from $\boldsymbol{B}$ to the standard basis.
c. Write the equation that relates $\boldsymbol{x}$ in $\boldsymbol{R}^3$ to $[\boldsymbol{x}]_{\boldsymbol{B}}$.
d. Find $[\boldsymbol{x}]_{\boldsymbol{B}}$, for the $\boldsymbol{x}$ given above.

Solution

a. It is evident that the matrix $\boldsymbol{P}_{\boldsymbol{B}} = [\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \boldsymbol{b}_3]$ is row equivalent to the identity matrix. By the Invertible Matrix Theorem, $\boldsymbol{P}_{\boldsymbol{B}}$ is invertible and its columns form a basis for $\boldsymbol{R}^3$.

b. From part (a), the change-of-coordinates matrix is $\boldsymbol{P_B} = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}$.

c. $\boldsymbol{x} = \boldsymbol{P_B}[\boldsymbol{x}]_{\boldsymbol{B}}$.

d. To solve part (c), it is probably easier to row reduce an augmented matrix instead of computing $\boldsymbol{P_B}^{-1}$. We have

$$\underbrace{\begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}}_{\boldsymbol{P_B}} \underbrace{\begin{matrix} -8 \\ 2 \\ 3 \end{matrix}}_{\boldsymbol{x}} \sim \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{I}} \underbrace{\begin{matrix} -5 \\ 2 \\ 1 \end{matrix}}_{[\boldsymbol{x}]_{\boldsymbol{B}}}$$

Hence $[\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$

Example 12 The set $\boldsymbol{B} = \{\boldsymbol{1} + \boldsymbol{t}, \boldsymbol{1} + \boldsymbol{t^2}, \boldsymbol{t} + \boldsymbol{t^2}\}$ is a basis for $\boldsymbol{P_2}$. Find the coordinate vector of $\boldsymbol{p(t)} = \boldsymbol{6} + 3\boldsymbol{t} - \boldsymbol{t^2}$ relative to $\boldsymbol{B}$.

Solution The coordinates of $\boldsymbol{p(t)} = \boldsymbol{6} + 3\boldsymbol{t} - \boldsymbol{t^2}$ with respect to $\boldsymbol{B}$ satisfy

$$c_1(\boldsymbol{1} + \boldsymbol{t}) + c_2(\boldsymbol{1} + \boldsymbol{t^2}) + c_3(\boldsymbol{t} + \boldsymbol{t^2}) = \boldsymbol{6} + 3\boldsymbol{t} - \boldsymbol{t^2}$$

$$c_1 + c_1t + c_2 + c_2t^2 + c_3t + c_3t^2 = 6 + 3t - t^2$$

$$c_1 + c_2 + c_1t + c_3t + c_2t^2 + c_3t^2 = 6 + 3t - t^2$$

$$c_1 + c_2 + (c_1 + c_3)t + (c_2 + c_3)t^2 = 6 + 3t - t^2$$

Equating coefficients of like powers of $\boldsymbol{t}$, we have

$c_1 + c_2 \quad = \quad 6$ --------------(1)

$c_1 \quad + c_3 = \quad 3$ --------------(2)

$c_2 + c_3 = -1$ --------------(3)

Subtract equation (2) *from* (1) *we get*

$c_2 - c_3 = 6 - 3 = 3$

Add this equation with equation (3)

$2c_2 = -1 + 3 = 2$

$\Rightarrow c_2 = 1$

Put value of c_2 *in equation* (3)

$1 + c_3 = -1$

$\Rightarrow c_3 = -2$

From equation (1) *we have*

$$c_1 + c_2 = 6$$
$$c_1 = 6 - 1 = 5$$

Solving, we find that $c_1 = 5$, $c_2 = 1$, $c_3 = -2$, and $[\boldsymbol{p}]_{\boldsymbol{B}} = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$.

Exercises

In exercises 1 and 2, find the vector $\boldsymbol{x}$ determined by the given coordinate vector $[\boldsymbol{x}]_{\boldsymbol{B}}$ and the given basis $\boldsymbol{B}$.

1. $\boldsymbol{B} = \left\{\begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix}\right\}, [\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

2. $\boldsymbol{B} = \left\{\begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -7 \\ 0 \end{bmatrix}\right\}, [\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$

In exercises 3-6, find the coordinate vector $[\boldsymbol{x}]_{\boldsymbol{B}}$ of $\boldsymbol{x}$ relative to the given basis $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \ldots, \boldsymbol{b_n}\}$.

3. $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

4. $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 5 \\ -6 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

5. $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} -3 \\ 4 \\ 9 \end{bmatrix}, \boldsymbol{b_3} = \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 8 \\ -9 \\ 6 \end{bmatrix}$

6. $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix}, \boldsymbol{b_3} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix}$

In exercises 7 and 8, find the change of coordinates matrix from $\boldsymbol{B}$ to standard basis in $\boldsymbol{R}^n$.

7. $\boldsymbol{B} = \left\{\begin{bmatrix} 2 \\ -9 \end{bmatrix}, \begin{bmatrix} 1 \\ 8 \end{bmatrix}\right\}$

8. $\boldsymbol{B} = \left\{\begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}, \begin{bmatrix} 8 \\ -2 \\ 7 \end{bmatrix}\right\}$

In exercises 9 and 10, use an inverse matrix to find $[\boldsymbol{x}]_{\boldsymbol{B}}$ for the given $\boldsymbol{x}$ and $\boldsymbol{B}$.

9. $\boldsymbol{B} = \left\{ \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \end{bmatrix} \right\}, \boldsymbol{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

10. $\boldsymbol{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}, \boldsymbol{x} = \begin{bmatrix} 2 \\ -6 \end{bmatrix}$

11. The set $\boldsymbol{B} = \{\mathbf{1} + \boldsymbol{t}^2, \boldsymbol{t} + \boldsymbol{t}^2, \mathbf{1} + 2\boldsymbol{t} + \boldsymbol{t}^2\}$ is a basis for $\boldsymbol{P_2}$. Find the coordinate vector of $\boldsymbol{p}(\boldsymbol{t}) = \boldsymbol{1} + 4\boldsymbol{t} + 7\boldsymbol{t}^2$ relative to $\boldsymbol{B}$.

12. The vectors $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ -8 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$ span $\boldsymbol{R}^2$ but do not form a basis. Find two different ways to express $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as a linear combination of $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$.

13. Let $\boldsymbol{B} = \left\{ \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \end{bmatrix} \right\}$. Since the coordinate mapping determined by $\boldsymbol{B}$ is a linear transformation from $\boldsymbol{R}^2$ into $\boldsymbol{R}^2$, this mapping must be implemented by some 2 x 2 matrix $\boldsymbol{A}$. Find it.

In exercises 14-16, use coordinate vectors to test the linear independence of the sets of polynomials.

14. $\mathbf{1} + \boldsymbol{t}^3, \mathbf{3} + \boldsymbol{t} - 2\boldsymbol{t}^2, -\boldsymbol{t} + 3\boldsymbol{t}^2 - \boldsymbol{t}^3$

15. $(\boldsymbol{t}-\mathbf{1})^2, \boldsymbol{t}^3 - \mathbf{2}, (\mathbf{t} - \mathbf{2})^3$

16. $\mathbf{3} + 7\boldsymbol{t}, \mathbf{5} + \boldsymbol{t} - 2\boldsymbol{t}^3, \boldsymbol{t} - 2\boldsymbol{t}^2, \mathbf{1} + 16\boldsymbol{t} - 6\boldsymbol{t}^2 + 2\boldsymbol{t}^3$

17. Let $\boldsymbol{H} = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ and $\boldsymbol{B} = \{\boldsymbol{v_1}, \boldsymbol{v_2}\}$. Show that $\boldsymbol{x}$ is in $\boldsymbol{H}$ and find the $\boldsymbol{B}$-coordinate vector of $\boldsymbol{x}$, for $\boldsymbol{v}_1 = \begin{bmatrix} 11 \\ -5 \\ 10 \\ 7 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 14 \\ -8 \\ 13 \\ 10 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 19 \\ -13 \\ 18 \\ 15 \end{bmatrix}$.

18. Let $\boldsymbol{H} = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ and $\boldsymbol{B} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$. Show that $\boldsymbol{B}$ is a basis for $\boldsymbol{H}$ and $\boldsymbol{x}$ is in $\boldsymbol{H}$, and find the $\boldsymbol{B}$-coordinate vector of $\boldsymbol{x}$, for $\boldsymbol{v}_1 = \begin{bmatrix} -6 \\ 4 \\ -9 \\ 4 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 8 \\ -3 \\ 7 \\ -3 \end{bmatrix}, \boldsymbol{v_3} = \begin{bmatrix} -9 \\ 5 \\ -8 \\ 3 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 4 \\ 7 \\ -8 \\ 3 \end{bmatrix}$.

Lecture 24

Dimension of a Vector Space

In this lecture, we will focus over the dimension of the vector spaces. The dimension of a vector space V is the cardinality or the number of vectors in the basis B of the given vector space. If the basis B has *n* (say) elements then this number *n* (called the dimension) is an intrinsic property of the space ***V.*** That is it does not depend on the particular choice of basis rather, all the bases of V will have the same cardinality. Thus, we can say that the dimension of a vector space is always unique. The discussion of dimension will give additional insight into properties of bases.
The first theorem generalizes a well-known result about the vector space $\boldsymbol{R}^n$.

Note

A vector space ***V*** with a basis ***B*** containing *n* vectors is isomorphic to $\boldsymbol{R}^n$ i.e., there exist a one-to-one linear transformation from V to $\boldsymbol{R}^n$.

Theorem 1 If a vector space ***V*** has a basis $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$, then any set in ***V*** containing more than *n* vectors must be linearly dependent.

Theorem 2 If a vector space ***V*** has a basis of *n* vectors, then every basis of ***V*** must consist of exactly *n* vectors.

Finite and infinite dimensional vector spaces

If the vector space ***V*** is spanned or generated by a finite set, then ***V*** is said to be **finite-dimensional**, and the **dimension** of ***V***, written as dim *V*, is the number of vectors in a basis for ***V***. If ***V*** is not spanned by a finite set, then ***V*** is said to be **infinite-dimensional**. That is, if we are unable to find a finite set that can generate the whole vector space, then such a vector space is called infinite dimensional.

Note

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

Example 1 The 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 $\le n$. It is called the standard basis with dim $(\boldsymbol{P}_n) = n + 1$.

(c) The set of 2 x 2 matrices with real entries (elements) $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ where

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{u}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \boldsymbol{u}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

is a linearly independent and every 2 x 2 matrix with real entries can be expressed as their linear combination. Therefore, they form a basis for the vector space $\boldsymbol{M}_{2X2}$. This basis is called the standard basis for $\boldsymbol{M}_{2X2}$ with dim $(\boldsymbol{M}_{2X2}) = 4$.

Note

(1) dim $(\boldsymbol{R}^n) = n$ { The standard basis has n vectors}.

(2) dim $(\boldsymbol{P}_n) = n + 1$ { The standard basis has $n+1$ vectors}.

(3) dim $(\boldsymbol{M}_{m\times n}) = mn$ { The standard basis has mn vectors.}

Example 3 Let $\boldsymbol{W}$ be the subspace of the set of all (2 x 2) matrices defined by

$$\boldsymbol{W} = \{\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : 2a - b + 3c + d = 0\}.$$

Determine the dimension of $\boldsymbol{W}$.
Solution The algebraic specification for $\boldsymbol{W}$ can be rewritten as $d = -2a + b - 3c$.

Now $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Substituting the value of d, it becomes

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

This can be written as

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

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

$$= a\boldsymbol{A_1} + b\boldsymbol{A_2} + c\boldsymbol{A_3}$$

where $\boldsymbol{A_1} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$, $\boldsymbol{A_2} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, and $\boldsymbol{A_3} = \begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix}$

The matrix $\boldsymbol{A}$ is in $\boldsymbol{W}$ if and only if $\boldsymbol{A} = a\boldsymbol{A_1} + b\boldsymbol{A_2} + c\boldsymbol{A_3}$, so $\{\boldsymbol{A_1}, \boldsymbol{A_2}, \boldsymbol{A_3}\}$ is a spanning set for $\boldsymbol{W}$. Now, check 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.*,

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

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

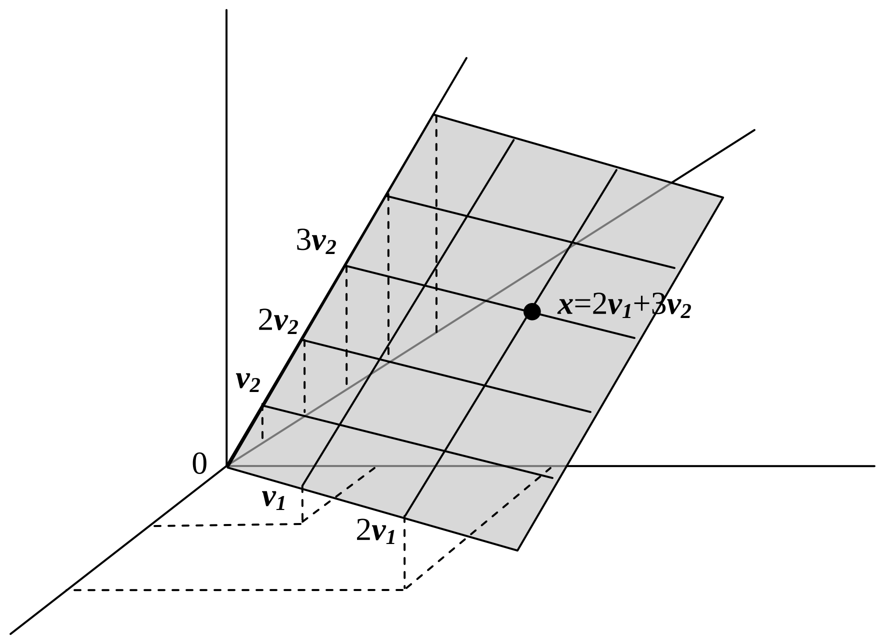
$$\begin{bmatrix} a & b \\ c & -2a+b-3c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

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 $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$. So $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_4}\}$ 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.

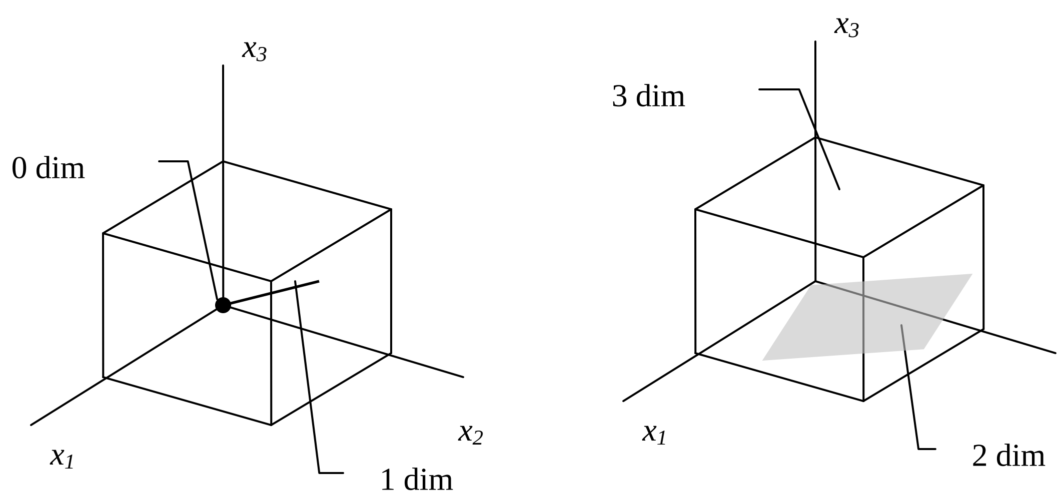


Figure 1 – Sample subspaces of $\boldsymbol{R^3}$

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

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

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

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

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

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

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

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

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

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

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

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

Thus, the reduced row echelon form of the augmented matrix is

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

which corresponds to the system

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

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

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

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 $\boldsymbol{Ax}$

$= \boldsymbol{0}$, where $\boldsymbol{x}$ is a column of weights. (If some columns are not involved in a particular dependence relation, then their weights are zero.) When $\boldsymbol{A}$ is row reduced to a matrix $\boldsymbol{B}$, the columns of $\boldsymbol{B}$ are often totally different from the columns of $\boldsymbol{A}$. However, the equations $\boldsymbol{Ax} = \boldsymbol{0}$ and $\boldsymbol{Bx} = \boldsymbol{0}$ have exactly the same set of solutions. That is, the columns of $\boldsymbol{A}$ have exactly the same linear dependence relationships as the columns of $\boldsymbol{B}$.

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 $\boldsymbol{B}$ in Example 8. Find a basis for Col $\boldsymbol{A}$.

Solution In Example 8, we have seen that $\boldsymbol{b_2} = 4\boldsymbol{b_1} \quad and \quad \boldsymbol{b_4} = 2\boldsymbol{b_1} - \boldsymbol{b_3}$ so we can expect that $\boldsymbol{a_2} = 4\boldsymbol{a_1} \quad and \quad \boldsymbol{a_4} = 2\boldsymbol{a_1} - \boldsymbol{a_3}$. This is indeed the case. Thus, we may discard $\boldsymbol{a_2}$ and $\boldsymbol{a_4}$ while selecting a minimal spanning set for Col $\boldsymbol{A}$. 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 $\boldsymbol{A}$. The columns we have used for this basis are the pivot columns of $\boldsymbol{A}$.

Examples 8 and 9 illustrate the following useful fact.

Theorem 3 The pivot columns of a matrix $\boldsymbol{A}$ form a basis for Col $\boldsymbol{A}$.

Proof The general proof uses the arguments discussed above. Let $\boldsymbol{B}$ be the reduced echelon form of $\boldsymbol{A}$. The set of pivot columns of $\boldsymbol{B}$ is linearly independent, for no vector in the set is a linear combination of the vectors that precede it. Since $\boldsymbol{A}$ is row equivalent to $\boldsymbol{B}$, the pivot columns of $\boldsymbol{A}$ are linearly independent too, because any linear dependence relation among the columns of $\boldsymbol{A}$ corresponds to a linear dependence relation among the columns of $\boldsymbol{B}$. For this same reason, every non-pivot column of $\boldsymbol{A}$ is a linear combination of the pivot columns of $\boldsymbol{A}$. Thus the non-pivot columns of $\boldsymbol{A}$ may be discarded from the spanning set for Col $\boldsymbol{A}$, by the Spanning Set Theorem. This leaves the pivot columns of $\boldsymbol{A}$ as a basis for Col $\boldsymbol{A}$.

Note Be careful to use pivot columns of $\boldsymbol{A}$ itself for the basis of Col $\boldsymbol{A}$. The columns of an echelon form $\boldsymbol{B}$ are often not in the column space of $\boldsymbol{A}$. For instance, the columns of the $\boldsymbol{B}$ in Example 8 all have zeros in their last entries, so they cannot span the column space of the $\boldsymbol{A}$ in Example 9.

Example 10 Let $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} -2 \\ 7 \\ -9 \end{bmatrix}$. Determine if $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is a basis for $\boldsymbol{R^3}$. Is $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ a basis for $\boldsymbol{R^2}$?

Solution Let $A = [\boldsymbol{v_1} \quad \boldsymbol{v_2}]$. Row operations show that $A = \begin{bmatrix} 1 & -2 \\ -2 & 7 \\ 3 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Not every row of A contains a pivot position. So the columns of A do not span $\boldsymbol{R^3}$, by Theorem 4 in Lecture 6. Hence $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is not a basis for $\boldsymbol{R^3}$. Since $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are not in $\boldsymbol{R^2}$, they cannot possibly be a basis for $\boldsymbol{R^2}$. However, since $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are obviously linearly independent, they are a basis for a subspace of $\boldsymbol{R^3}$, namely, Span $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$.

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

Solution Let $\boldsymbol{A}$ be the matrix whose column space is the space spanned by $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}\}$,

$$\boldsymbol{A} = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix}$$

Reduce the matrix $\boldsymbol{A}$ into its echelon form in order to find its pivot columns.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 20 & 4 & -20 \\ 0 & -25 & -5 & 25 \end{bmatrix} by\, R_2 + 3R_1,\ R_3 - 4R_1$$

$$\sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 5 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix} by\ \frac{1}{4}R_2,\ -\frac{1}{5}R_3,\ R_3 - R_2$$

The first two columns of $\boldsymbol{A}$ are the pivot columns and hence form a basis of Col $\boldsymbol{A} = \boldsymbol{W}$. Hence $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is a basis for $\boldsymbol{W}$.

Note that the reduced echelon form of $\boldsymbol{A}$ 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

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

Finding a basis for column space of this matrix can solve the first part of our problem.
Transforming Matrix to Reduced Row Echelon Form:

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -4 & 1 & -1 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 0 & 3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix} \begin{array}{l} \\ 2R_1 + R_2 \\ -3R_1 + R_4 \\ \\ \end{array}$$

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

$$\sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} -2R_2 + R_3 \\ -3R_2 + R_4 \end{matrix}$$

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

Labeling the column vectors of the resulting matrix as $\boldsymbol{w_1}$, $\boldsymbol{w_2}$, $\boldsymbol{w_3}$ and $\boldsymbol{w_4}$ yields

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

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

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

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

$$\begin{array}{ccccc} \left[\begin{matrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}\right] \\ \uparrow \quad\; \uparrow \quad\; \uparrow \quad\; \uparrow \quad\; \uparrow \\ \boldsymbol{w_1} \quad \boldsymbol{w_2} \quad \boldsymbol{w_3} \quad \boldsymbol{w_4} \quad \boldsymbol{w_5} \end{array} \qquad (B)$$

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

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

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

$$\begin{array}{c} \left[\begin{matrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}\right] \\ \uparrow \quad\; \uparrow \quad\; \uparrow \quad\; \uparrow \quad\; \uparrow \\ \boldsymbol{w_1} \quad \boldsymbol{w_2} \quad \boldsymbol{w_3} \quad \boldsymbol{w_4} \quad \boldsymbol{w_5} \end{array} \qquad (B)$$

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

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

Theorem 5 Let $\boldsymbol{H}$ 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 $\boldsymbol{Ax} = \boldsymbol{0}$ has k free variables. From lecture 21, we know that the standard method of finding a spanning set for Nul $\boldsymbol{A}$ will produce exactly k linearly independent vectors say, $\boldsymbol{u_1}, \dots, \boldsymbol{u_k}$, 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 $\boldsymbol{Ax} = \boldsymbol{0}$, and the dimension of Col $\boldsymbol{A}$ is the number of pivot columns in $\boldsymbol{A}$.

Example 15 Find the dimensions of the null space and column space of

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

Solution Row reduce the augmented matrix $[\boldsymbol{A} \quad \boldsymbol{0}]$ to echelon form and obtain

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

Writing it in equations form, we get

$$x_1 - 2x_2 + 2x_3 + 3x_4 - x_5 = 0$$

$$x_3 + 2x_4 - 2x_5 = 0$$

Since the number of unknowns is more than the number of equations, we will introduce free variables here (say) x_2, x_4 and x_5. Hence the dimension of Nul $\boldsymbol{A}$ is *3*. Also dim Col $\boldsymbol{A}$ *is 2* because $\boldsymbol{A}$ has two pivot columns.

Example 16 Decide whether each statement is true or false, and give a reason for each answer. Here $\boldsymbol{V}$ is a non-zero finite-dimensional vector space.

1. If dim $\boldsymbol{V} = p$ and if $\boldsymbol{S}$ is a linearly dependent subset of $\boldsymbol{V}$, then $\boldsymbol{S}$ contains more than p vectors.
2. If $\boldsymbol{S}$ spans $\boldsymbol{V}$ and if $\boldsymbol{T}$ is a subset of $\boldsymbol{V}$ that contains more vectors than $\boldsymbol{S}$, then $\boldsymbol{T}$ is linearly dependent.

Solution

1. False. Consider the set $\{\boldsymbol{0}\}$.
2. True. By the Spanning Set Theorem, $\boldsymbol{S}$ contains a basis for $\boldsymbol{V}$; call that basis $\boldsymbol{S}'$. Then $\boldsymbol{T}$ will contain more vectors than $\boldsymbol{S}'$. By Theorem 1, $\boldsymbol{T}$ is linearly dependent.

Exercises

For each subspace in exercises 1-6, (a) find a basis and (b) state the dimension.

1. $\left\{ \begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix} : s, t \text{ in } \boldsymbol{R} \right\}$

2. $\left\{ \begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix} : a, b, c \text{ in } \boldsymbol{R} \right\}$

3. $\left\{ \begin{bmatrix} a-4b-2c \\ 2a+5b-4c \\ -a+2c \\ -3a+7b+6c \end{bmatrix} : a, b, c \text{ in } \boldsymbol{R} \right\}$

4. $\left\{ \begin{bmatrix} 3a+6b-c \\ 6a-2b-2c \\ -9a+5b+3c \\ -3a+b+c \end{bmatrix} : a, b, c \text{ in } \boldsymbol{R} \right\}$

5. {(a, b, c): a – 3b + c = 0, b – 2c = 0, 2b – c = 0}

6 {(a, b, c, d): a - 3b + c = 0}

7. Find the dimension of the subspace $\boldsymbol{H}$ of $\boldsymbol{R}^2$ spanned by

$$\begin{bmatrix} 2 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 10 \end{bmatrix}, \begin{bmatrix} -3 \\ 6 \end{bmatrix}$$

8. Find the dimension of the subspace spanned by the given vectors.

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} -7 \\ -3 \\ 1 \end{bmatrix}$$

Determine the dimensions of Nul $\boldsymbol{A}$ and Col $\boldsymbol{A}$ for the matrices shown in exercises 9 to 12.

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

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

11. $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 9 & 5 \\ 0 & 0 & 1 & -4 \end{bmatrix}$

12. $\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$

13. The first four Hermite polynomials are **1**, 2***t***, **-2** + 4t^2, and -12***t*** + 8t^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 P_3.

14. Let ***B*** be the basis of P_3 consisting of the Hermite polynomials in exercise 13, and let ***p*** (***t***) = **7** – 12 ***t*** – 8 t^2 + 12 t^3. Find the coordinate vector of ***p*** relative to ***B***.

15. Extend the following vectors to a basis for R^5:

$$v_1 = \begin{bmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{bmatrix}, v_2 = \begin{bmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{bmatrix}, v_3 = \begin{bmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{bmatrix}$$

Lecture 25

Rank

With the help of vector space concepts, for a matrix several interesting and useful relationships in matrix rows and columns have been discussed.
For instance, imagine placing 2000 random numbers into a 40 x 50 matrix ***A*** and then determining both the maximum number of linearly independent columns in ***A*** and the maximum number of linearly independent columns in A^T (rows in ***A***). Remarkably, the two numbers are the same. Their common value is called the rank of the matrix. To explain why, we need to examine the subspace spanned by the subspace spanned by the rows of ***A***.

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

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

$$r_1 = (-2,-5,8,0,-17)$$
$$r_2 = (1,3,-5,1,5)$$
$$r_3 = (3,11,-19,7,1)$$
$$r_4 = (1,7,-13,5,-3)$$

The row space of ***A*** is the subspace of R^5 spanned by $\{r_1, r_2, r_3, r_4\}$. That is, Row ***A*** = Span $\{r_1, r_2, r_3, r_4\}$. Naturally, we write row vectors horizontally; however, they could also be written as column vectors

Example **Let**

$A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & -1 & 4 \end{bmatrix}$ and

$$r_1 = (2,1,0)$$
$$r_2 = (3,-1,4)$$

That is Row **A**=Span $\{r_1, r_2\}$.

We could use the Spanning Set Theorem to shrink the spanning set to a basis.
Some times row operation on a matrix will not give us the required information but row reducing certainly worthwhile, as the next theorem shows

Theorem 1 If two matrices ***A*** and ***B*** are row equivalent, then their row spaces are the same. If ***B*** is in echelon form, the nonzero rows of ***B*** form a basis for the row space of ***A*** as well as ***B***.

Theorem 2 If ***A*** and ***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***.

Example

The matrix

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

is in row-echelon form.

The vectors

$$\boldsymbol{r}_1 = \begin{bmatrix} 1 & -2 & 5 & 0 & 3 \end{bmatrix}$$

$$\boldsymbol{r}_2 = \begin{bmatrix} 0 & 1 & 3 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{r}_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

form a basis for the row space of R, and the vectors

$$c_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ c_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, c_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R.

Example 3 **(Basis for a Vector Space using Row Operation)**
Find bases for the space spanned by the vectors

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

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

Solution **T**he space spanned by these vectors is the row space of the matrix

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

Transforming Matrix to Row Echelon Form:

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 5 & 15 & 10 & 0 \\ 0 & 10 & 18 & 8 & 0 \end{bmatrix} \begin{matrix} (-2)R_1 + R_2 \\ (-2)R_1 + R_4 \\ (-1)R_2 \end{matrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -12 & -12 & 0 \end{bmatrix} \begin{matrix} (-5)R_2 + R_3 \\ (-10)R_2 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & -12 & -12 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} R_{34}$$

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3$$

Therefore, $\quad \boldsymbol{R} = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

The non-zero row vectors in this matrix are

$$\boldsymbol{w_1} = (1,-2,0,0,3), \boldsymbol{w_2} = (0,1,3,2,0), \boldsymbol{w_3} = (0,0,1,1,0)$$

These vectors form a basis for the row space and consequently form a basis for the subspace of $\boldsymbol{R}^5$ spanned by $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$.

Example 4 **(Basis for the Row Space of a Matrix)**

Find a basis for the row space of $\boldsymbol{A} = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$ consisting entirely of row vectors from $\boldsymbol{A}$.

Solution We find $\boldsymbol{A}^T$; then we will use the method of example (2) to find a basis for the column space of $\boldsymbol{A}^T$; and then we will transpose again to convert column vectors back to row vectors. Transposing $\boldsymbol{A}$ yields

$$\boldsymbol{A}^T = \begin{bmatrix} 1 & 2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 6 \end{bmatrix}$$

Transforming Matrix to Row Echelon Form:

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & -1 & 5 & 10 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} 2R_1 + R_2 \\ (-3)R_1 + R_5 \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} (-1)R_2$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} (3)R_2 + R_3 \\ (2)R_2 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} 12\,R_3 + R_4$$

Now $$\boldsymbol{R} = \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The first, second and fourth columns contain the leading 1's, so the corresponding column vectors in $\boldsymbol{A}^T$ form a basis for the column space of $\boldsymbol{A}^T$; these are

$$\boldsymbol{c_1} = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 3 \end{bmatrix},\ \boldsymbol{c_2} = \begin{bmatrix} 2 \\ -5 \\ -3 \\ -2 \\ 6 \end{bmatrix} \text{and}\ \boldsymbol{c_4} = \begin{bmatrix} 2 \\ 6 \\ 18 \\ 8 \\ 6 \end{bmatrix}$$

Transposing again and adjusting the notation appropriately yields the basis vectors

$$\boldsymbol{r_1} = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \end{bmatrix}, \boldsymbol{r_2} = \begin{bmatrix} 2 & -5 & -3 & -2 & 6 \end{bmatrix} \textit{and } \boldsymbol{r_4} = \begin{bmatrix} 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

for the row space of $\boldsymbol{A}$.

The following example shows how one sequence of row operations on $\boldsymbol{A}$ leads to bases for the three spaces: Row $\boldsymbol{A}$, Col $\boldsymbol{A}$, and Nul $\boldsymbol{A}$.

Example 5 Find bases for the row space, the column space and the null space of the matrix

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

Solution To find bases for the row space and the column space, row reduce $\boldsymbol{A}$ to an echelon form: $\boldsymbol{A} \sim \boldsymbol{B} = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

By Theorem (1), the first three rows of $\boldsymbol{B}$ form a basis for the row space of $\boldsymbol{A}$ (as well as the row space of $\boldsymbol{B}$). Thus Basis for Row $\boldsymbol{A}$:

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

For the column space, observe from $\boldsymbol{B}$ that the pivots are in columns *1, 2* and *4*. Hence columns *1, 2* and *4* of $\boldsymbol{A}$ (not $\boldsymbol{B}$) form a basis for Col $\boldsymbol{A}$:

$$\text{Basis for Col } \boldsymbol{A} : \left\{ \begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix} \right\}$$

Any echelon form of $\boldsymbol{A}$ provides (in its nonzero rows) a basis for Row $\boldsymbol{A}$ and also identifies the pivot columns of $\boldsymbol{A}$ for Col $\boldsymbol{A}$. However, for Nul $\boldsymbol{A}$, we need the reduced echelon form. Further row operations on $\boldsymbol{B}$ yield

$$\boldsymbol{A} \sim \boldsymbol{B} \sim \boldsymbol{C} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The equation $\boldsymbol{Ax} = \boldsymbol{0}$ is equivalent to $\boldsymbol{Cx} = \boldsymbol{0}$, that is,

$$\begin{aligned} x_1 + \quad x_3 \quad + x_5 &= 0 \\ x_2 - 2x_3 \quad + 3x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

So $x_1 = -x_3 - x_5$, $x_2 = 2x_3 - 3x_5$, $x_4 = 5x_5$, with x_3 and x_5 free variables. The usual calculations (discussed in lecture 21) show that

$$\text{Basis for Nul } \boldsymbol{A}: \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

Observe that, unlike the bases for Col ***A***, the bases for Row ***A*** and Nul ***A*** have no simple connection with the entries in ***A*** itself.

Note

1. Although the first three rows of ***B*** in Example (5) are linearly independent, it is wrong to conclude that the first three rows of ***A*** are linearly independent. (In fact, the third row of A is 2 times the first row plus 7 times the second row).
2. Row operations do not preserve the linear dependence relations among the rows of a matrix.

Definition The **rank** of ***A*** is the dimension of the column space of ***A***.
Since Row ***A*** is the same as Col $\boldsymbol{A}^{\boldsymbol{T}}$, the dimension of the row space of ***A*** is the rank of $\boldsymbol{A}^{\boldsymbol{T}}$. 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 $\boldsymbol{R}^{\boldsymbol{6}}$ 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 $\boldsymbol{A}^{\boldsymbol{T}}$ (= Col ***A***) and Nul $\boldsymbol{A}^{\boldsymbol{T}}$. 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 A is the x_2 – axis, Row A is the x_1x_3 – plane, Col A is the plane whose equation is $x_1 - x_2 = 0$ and Nul A^T is the set of all multiples of $(1, -1, 0)$. Figure 1 shows Nul A and Row A in the domain of the linear transformation $x \to Ax$; the range of this mapping, Col A, is shown in a separate copy of R^3, along with Nul A^T.

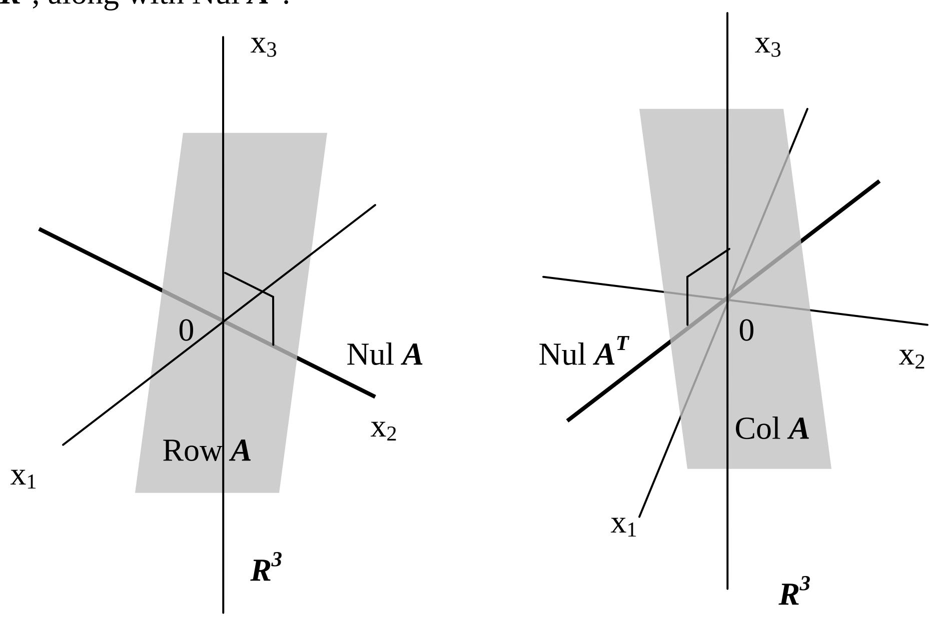


Figure 1 – Subspaces associated with a matrix A

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 A be the 40×42 coefficient matrix of the system. The given information implies that the two solutions are linearly independent and span Nul A. So dim Nul $A = 2$. By the Rank Theorem, dim Col $A = 42 - 2 = 40$. Since R^{40} is the only subspace of R^{40} whose dimension is 40, Col A must be all of R^{40}. This means that every non-homogeneous equation $Ax = b$ has a solution.

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

Verify that values obtained verify the dimension theorem.

Solution

$$\begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix} (-1)R_1$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \end{bmatrix} \begin{matrix} (-3)R_1 + R_2 \\ (-2)R_1 + R_3 \\ (-4)R_1 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & -1 & 2 & 12 & 16 & -5 \\ 0 & -1 & 2 & 12 & 16 & -5 \end{bmatrix} (-1)R_2$$

$$\begin{bmatrix} 1 & -2 & 0 & -4 & -5 & 3 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} R_2 + R_3 \\ R_2 + R_4 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} 2R_2 + R_1$$

The reduced row-echelon form of A is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (1)$$

The corresponding system of equations will be

$$x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 = 0$$
$$x_2 - 2x_3 - 12x_4 - 16x_5 + \ 5x_6 = 0$$

or, on solving for the leading variables,

$$\begin{aligned} x_1 &= 4x_3 - 28x_4 + 37x_5 - 13x_6 \\ x_2 &= 2x_3 + 12x_4 + 16x_5 - 5x_6 \end{aligned} \qquad (2)$$

it follows that the general solution of the system is

$$x_1 = 4r + 28s + 37t - 13u$$
$$x_2 = 2r + 12s + 16t - 5u$$
$$x_3 = r$$
$$x_4 = s$$
$$x_5 = t$$
$$x_6 = u$$

or equivalently, $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad (3)$$

The four vectors on the right side of (3) form a basis for the solution space, so nullity $(\boldsymbol{A}) = 4$. The matrix $\boldsymbol{A} = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$ has 6 columns,

so rank$(\boldsymbol{A})$ + nullity$(\boldsymbol{A}) = 2 + 4 = 6 = n$

Example 10 Find the rank and nullity of the matrix; then verify that the values obtained satisfy the dimension theorem $\boldsymbol{A} = \begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$

Solution Transforming Matrix to the Reduced Row Echelon Form:

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & 6 & 0 & -3 \end{bmatrix} \begin{matrix} \\ (-2)R_1 + R_3 \\ (-3)R_1 + R_4 \\ 2R_1 + R_5 \\ \\ \end{matrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & -6 & 0 & 2 \\ 0 & 3 & 6 & 0 & -3 \end{bmatrix} (1/3)R_2$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \\ (-3)\, R_2 + R_3 \\ (-3)\, R_2 + R_4 \\ (-3)R_2 + R_5 \\ \\ \end{matrix}$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & -12 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (-1/12)R_3$$

$$\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} 12R_3 + R_4$$

$$\begin{bmatrix} 1 & -3 & 0 & 2 & 11/6 \\ 0 & 1 & 0 & 0 & -1/6 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} (-2\ ) R_3 + R_2 \\ (-2\ ) R_3 + R_1 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 2 & 4/3 \\ 0 & 1 & 0 & 0 & -1/6 \\ 0 & 0 & 1 & 0 & -5/12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (3) R_2 + R_1 \qquad (1)$$

Since there are three nonzero rows (or equivalently, three leading 1's) the row space and column space are both three dimensional so rank $(\boldsymbol{A}) = 3$.
To find the nullity of $\boldsymbol{A}$, we find the dimension of the solution space of the linear system $\boldsymbol{Ax} = \boldsymbol{0}$. The system can be solved by reducing the augmented matrix to reduced row echelon form. The resulting matrix will be identical to (1), except with an additional last column of zeros, and the corresponding system of equations will be

$$x_1 + 0x_2 + 0x_3 + 2x_4 + \frac{4}{3}x_5 = 0$$

$$0x_1 + x_2 + 0x_3 + 0x_4 - \frac{1}{6}x_5 = 0$$

$$0x_1 + 0x_2 + x_3 + 0x_4 - \frac{5}{12}x_5 = 0$$

The system has infinitely many solutions:

$x_1 = -2\, x_4 + (-4/3)\, x_5$ $\qquad$ $x_2 = (1/6)\, x_5$

$x_3 = (5/12)\, x_5$ $\qquad$ $x_4 = s$

$x_5 = t$

The solution can be written in the vector form:

$\boldsymbol{c_4} = (-2, 0, 0, 1, 0)$ $\qquad$ $\boldsymbol{c_5} = (-4/3, 1/6, 5/12, 0, 1)$

Therefore the **null space** has a basis formed by the set

$$\{(-2, 0, 0, 1, 0), (-4/3, 1/6, 5/12, 0, 1)\}$$

The nullity of the matrix is 2. Now Rank $(\boldsymbol{A})$ + nullity $(\boldsymbol{A}) = 3 + 2 = 5 = n$

Theorem 4 If $\boldsymbol{A}$ is an m x n, matrix, then
(a) rank ($\boldsymbol{A}$) = the number of leading variables in the solution of $\boldsymbol{Ax} = \boldsymbol{0}$
(b) nullity ($\boldsymbol{A}$) = the number of parameters in the general solution of $\boldsymbol{Ax} = \boldsymbol{0}$

Example 11 Find the number of parameters in the solution set of $\boldsymbol{Ax} = \boldsymbol{0}$ if $\boldsymbol{A}$ is a 5×7 matrix of rank 3.
Solution nullity $(\boldsymbol{A}) = n$ – rank $(\boldsymbol{A}) = 7-3 = 4$
Thus, there are four parameters.

Example Find the number of parameters in the solution set of $\boldsymbol{Ax} = \boldsymbol{0}$ if $\boldsymbol{A}$ is a 4×4 matrix of rank 0.
Solution nullity $(\boldsymbol{A}) = n$ – rank $(\boldsymbol{A}) = 4-0 = 4$
Thus, there are four parameters.

Theorem 5 If $\boldsymbol{A}$ is any matrix, then rank $(\boldsymbol{A})$ = rank $(\boldsymbol{A}^T)$

Four fundamental matrix spaces

If we consider a matrix $\boldsymbol{A}$ and its transpose $\boldsymbol{A}^T$ together, then there are six vectors spaces of interest:

Row space of $\boldsymbol{A}$	row space of $\boldsymbol{A}^T$
Column space of $\boldsymbol{A}$	column space of $\boldsymbol{A}^T$
Null space of $\boldsymbol{A}$	null space of $\boldsymbol{A}^T$

However, transposing a matrix converts row vectors into column vectors and column vectors into row vectors, so that, except for a difference in notation, the row space of $\boldsymbol{A}^T$ is the same as the column space of $\boldsymbol{A}$ and the column space of $\boldsymbol{A}^T$ is the same as row space of of $\boldsymbol{A}$.
This leaves four vector spaces of interest:

Row space of $\boldsymbol{A}$	column space of $\boldsymbol{A}$
Null space of $\boldsymbol{A}$	null space of $\boldsymbol{A}^T$

These are known as the **fundamental matrix spaces** associated with $\boldsymbol{A}$, if $\boldsymbol{A}$ is an m x n matrix, then the row space of $\boldsymbol{A}$ and null space of $\boldsymbol{A}$ are subspaces of $\boldsymbol{R}^n$ and the column space of $\boldsymbol{A}$ and the null space of $\boldsymbol{A}^T$ are subspaces of $\boldsymbol{R}^m$.

Suppose now that $\boldsymbol{A}$ is an m x n matrix of rank r, it follows from theorem (5) that $\boldsymbol{A}^T$ is an n x m matrix of rank r . Applying theorem (3) on $\boldsymbol{A}$ and $\boldsymbol{A}^T$ yields

Nullity $(\boldsymbol{A}) = n-r$, 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 7×4 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 4×7 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 $n \times n$ 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 ***A*** is often determined from the singular value decomposition of ***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 ***A*** and dim Nul ***A***.
2. Find bases for Col ***A*** and Row ***A***.
3. What is the next step to perform if one wants to find a basis for Nul ***A***?
4. How many pivot columns are in a row echelon form of $\boldsymbol{A^T}$?

Solution

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

$$\{\boldsymbol{a_1}, \boldsymbol{a_2}\} = \left\{ \begin{bmatrix} 2 \\ 1 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 8 \\ -5 \end{bmatrix} \right\}$$

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

Exercises

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

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

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 A is a 4 x 3 matrix, what is the largest possible dimension of the row space of A? If A is a 3 x 4 matrix, what is the largest possible dimension of the row space of A? Explain.

15. If A is a 6 x 8 matrix, what is the smallest possible dimension of Nul A?

16. If A is a 6 x 4 matrix, what is the smallest possible dimension of Nul A?

Lecture 26

Change of Basis

When a basis $\boldsymbol{B}$ is chosen for an n-dimensional vector space $\boldsymbol{V}$, the associated coordinate mapping onto $\boldsymbol{R}^n$ provides a coordinate system for $\boldsymbol{V}$. Each $\boldsymbol{x}$ in $\boldsymbol{V}$ is identified uniquely by its $\boldsymbol{B}$-coordinate vector $[\boldsymbol{x}]_{\boldsymbol{B}}$.

In some applications, a problem is described initially using a basis $\boldsymbol{B}$, but the problem's solution is aided by changing $\boldsymbol{B}$ to a new basis $\boldsymbol{C}$. Each vector is assigned a new $\boldsymbol{C}$-coordinate vector. In this section, we study how $[\boldsymbol{x}]_{\boldsymbol{C}}$ and $[\boldsymbol{x}]_{\boldsymbol{B}}$ are related for each $\boldsymbol{x}$ in $\boldsymbol{V}$.

To visualize the problem, consider the two coordinate systems in Fig. 1.

In Fig. 1(a), $\boldsymbol{x} = 3\boldsymbol{b}_1 + \boldsymbol{b}_2$, while in Fig. 1 (b), the same $\boldsymbol{x}$ is shown as $\boldsymbol{x} = 6\boldsymbol{c}_1 + 4\boldsymbol{c}_2$. That is

$$[\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad [\boldsymbol{x}]_{\boldsymbol{C}} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Example 1 shows how to find the connection between the two coordinate vectors. provided we know how $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are formed from $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$.

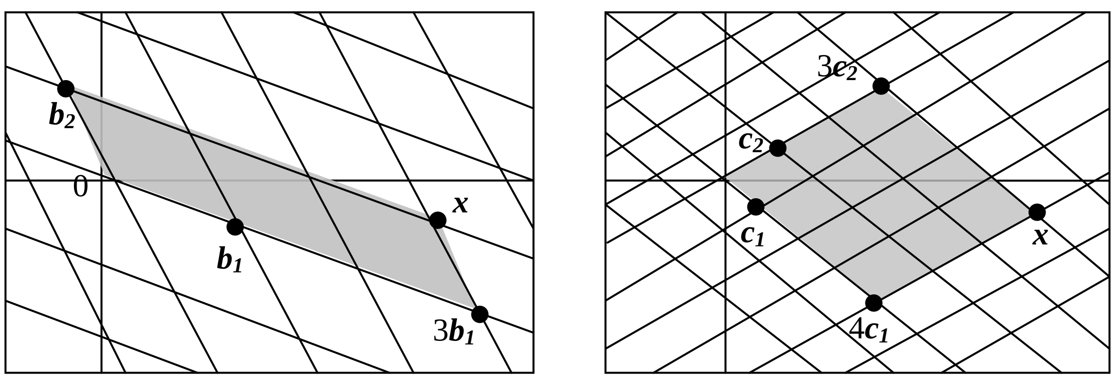


Figure 1 – Two coordinate systems for the same vector space

Example 1 Consider two bases $\boldsymbol{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ and $\boldsymbol{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2\}$ for a vector space $\boldsymbol{V}$, such that

$$\boldsymbol{b}_1 = 4\boldsymbol{c}_1 + \boldsymbol{c}_2 \text{ and } \boldsymbol{b}_2 = -6\boldsymbol{c}_1 + \boldsymbol{c}_2 \qquad (1)$$

Suppose that $\boldsymbol{x} = 3\boldsymbol{b}_1 + \boldsymbol{b}_2$ (2)

suppose that $[\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Find $[\boldsymbol{x}]_{\boldsymbol{C}}$.

Solution Since the coordinate mapping is a linear transformation, we apply (2)

$$[\boldsymbol{x}]_{\boldsymbol{C}} = [3\boldsymbol{b_1} + \boldsymbol{b_2}]_{\boldsymbol{C}}$$
$$= 3[\boldsymbol{b_1}]_C + [\boldsymbol{b_2}]_{\boldsymbol{C}}$$

as a matrix equation,

$$[\boldsymbol{x}]_{\boldsymbol{C}} = \begin{bmatrix} [\boldsymbol{b_1}]_{\boldsymbol{C}} & [\boldsymbol{b_2}]_{\boldsymbol{C}} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{3}$$

From (1),

$$[\boldsymbol{b_1}]_{\boldsymbol{C}} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \text{and} [\boldsymbol{b_2}]_{\boldsymbol{C}} = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$$

from (3) $\quad [\boldsymbol{x}]_{\boldsymbol{C}} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$

The C **–** coordinates of $\boldsymbol{x}$ match those of the $\boldsymbol{x}$ in Fig. 1.
The argument used to derive formula (3) is easily generalized to yield the following result.

Theorem Let $\boldsymbol{B} = \{\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}\}$ and $\boldsymbol{C} = \{\boldsymbol{c_1}, \ldots, \boldsymbol{c_n}\}$ be bases of a vector space $\boldsymbol{V}$. Then there exists an n x n matrix $\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}}$ such that

$$[\boldsymbol{x}]_{\boldsymbol{C}} = \underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}} [\boldsymbol{x}]_{\boldsymbol{B}} \tag{4}$$

The columns of $\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}}$ are the $\boldsymbol{C}$-coordinate vectors of the vectors in the basis $\boldsymbol{B}$. So

$$\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}} = \begin{bmatrix} [\boldsymbol{b_1}]_C & [\boldsymbol{b_2}]_C & \cdots & [\boldsymbol{b_n}]_C \end{bmatrix}$$

The matrix $\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}}$ in above theorem is called the **change-of-coordinates matrix from *B* to *C***. Multiplication by $\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}}$ converts ***B***-coordinates into ***C***-coordinates. Figure 2 illustrates the change-of-coordinates equation (4).
The columns of $\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}}$ are linearly independent because they are the coordinate vectors of the linearly independent set $\boldsymbol{B}$. Thus $\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}}$ is invertible. Left-multiplying both sides of (4) by $\left(\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}}\right)^{-1}$, we obtain

$$\left(\underset{\boldsymbol{C \leftarrow B}}{\boldsymbol{P}}\right)^{-1} [\boldsymbol{x}]_{\boldsymbol{C}} = [\boldsymbol{x}]_{\boldsymbol{B}}$$

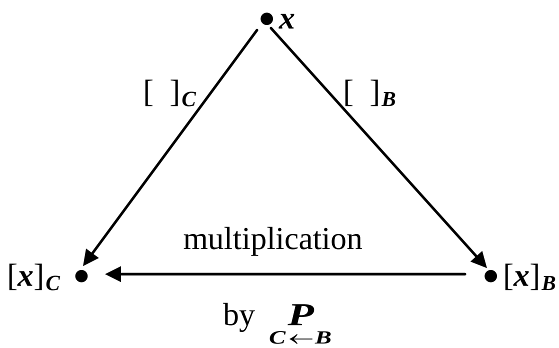


Figure 2 Two coordinate system for ***V***.

Thus $\left(\underset{C\leftarrow B}{P}\right)^{-1}$ is the matrix that converts ***C***-coordinates into ***B***-coordinates. That is,

$$\left(\underset{C\leftarrow B}{P}\right)^{-1} = \underset{B\leftarrow C}{P} \tag{5}$$

Change of Basis in R^n If $B = \{b_1, \ldots, b_n\}$ and ***E*** is the standard basis $\{e_1, \ldots, e_n\}$ in R^n, then $[b_1]_E = b_1$, and likewise for the other vectors in ***B***. In this case, $\underset{E\leftarrow B}{P}$ is the same as the change-of-coordinates matrix P_B introduced in Lecture 23, namely,

$$P_B = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}$$

To change coordinates between two nonstandard bases in R^n, we need above Theorem. The theorem shows that to solve the change-of-basis problem, we need the coordinate vector of the old basis relative to the new basis.

Example 2 Let $D = \{d_1, d_2, d_3\}$ and $F = \{f_1, f_2, f_3\}$ be basis for vector space ***V***, and suppose that $f_1 = 2d_1 - d_2 + d_3$, $f_2 = 3d_2 + d_3$, $f_3 = -3d_1 + 2d_3$.

(a) Find the change-of-coordinates matrix from ***F*** to ***D***.

(b) Find $[x]_D$ for $x = f_1 - 2f_2 + 2f_3$.

Solution (a) $\underset{D\leftarrow F}{P} = \begin{bmatrix} [f_1]_D & [f_2]_D & [f_3]_D \end{bmatrix}$

But $[f_1]_D = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, [f_2]_D = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}, [f_3]_D = \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}$

$\therefore \quad \underset{D\leftarrow F}{P} = \begin{bmatrix} 2 & 0 & -3 \\ -1 & 3 & 0 \\ 1 & 1 & 2 \end{bmatrix}$

(b) $[x]_D = [f_1 - 2f_2 + 2f_3]_D = [f_1]_D - 2[f_2]_D + 2[f_3]_D$

$$= \left[[\boldsymbol{f}_1]_{\boldsymbol{D}} \quad [\boldsymbol{f_2}]_{\boldsymbol{D}} \quad [\boldsymbol{f_3}]_{\boldsymbol{D}}\right] \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & -3 \\ -1 & 3 & 0 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ -7 \\ 3 \end{bmatrix}$$

Example 3 Let $\boldsymbol{b_1} = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$, $\boldsymbol{b_2} = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, $\boldsymbol{c_1} = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\boldsymbol{c_2} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$ and consider the bases for $\boldsymbol{R^2}$ given by $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}\}$ and $\boldsymbol{C} = \{\boldsymbol{c_1}, \boldsymbol{c_2}\}$. Find the change of coordinates matrix from $\boldsymbol{B}$ to $\boldsymbol{C}$.

Solution The matrix $\underset{C \leftarrow B}{\boldsymbol{P}}$ involves the $\boldsymbol{C}$ – coordinate vectors of $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$.

Let $[\boldsymbol{b_1}]_{\boldsymbol{C}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $[\boldsymbol{b_2}]_{\boldsymbol{C}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Then, by definition,

$$[\boldsymbol{c_1} \quad \boldsymbol{c_2}] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \boldsymbol{b_1} \qquad \text{and} \qquad [\boldsymbol{c_1} \quad \boldsymbol{c_2}] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \boldsymbol{b_2}$$

To solve both systems simultaneously augment the coefficient matrix with $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$ and row reduce:

$$[\boldsymbol{c_1} \quad \boldsymbol{c_2} \quad \boldsymbol{b_1} \quad \boldsymbol{b_2}] = \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right]$$

$$\sim \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 7 & -35 & -21 \end{array}\right] R_2 + 4R_1$$

$$\sim \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 1 & -5 & -3 \end{array}\right] \frac{1}{7} R_2$$

$$\sim \left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right] R_1 - 3R_2 \qquad (6)$$

$$[\boldsymbol{b_1}]_{\boldsymbol{C}} = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \quad \text{and} \quad [\boldsymbol{b_2}]_{\boldsymbol{C}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

therefore

$$\underset{C \leftarrow B}{\boldsymbol{P}} = \left[[\boldsymbol{b_1}]_C \quad [\boldsymbol{b_2}]_C\right] = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$$

Example 4 Let $\boldsymbol{b_1} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, $\boldsymbol{b_2} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\boldsymbol{c_1} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $\boldsymbol{c_2} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and consider the bases for $\boldsymbol{R^2}$ given by $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}\}$ and $\boldsymbol{C} = \{\boldsymbol{c_1}, \boldsymbol{c_2}\}$.

(a) Find the change – of – coordinates matrix from $\boldsymbol{C}$ to $\boldsymbol{B}$.

(b) Find the change – of – coordinates matrix from $\boldsymbol{B}$ to $\boldsymbol{C}$.

Solution

(a) Notice that $\underset{B\leftarrow C}{\boldsymbol{P}}$ is needed rather than $\underset{C\leftarrow B}{\boldsymbol{P}}$ and compute

$$[\boldsymbol{b_1} \quad \boldsymbol{b_2} \mid \boldsymbol{c_1} \quad \boldsymbol{c_2}] = \left[\begin{array}{cc|cc} 7 & 2 & 4 & 5 \\ -2 & -1 & 1 & 2 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & \frac{2}{7} & \frac{4}{7} & \frac{5}{7} \\ -2 & -1 & 1 & 2 \end{array}\right] \frac{1}{7}R_1$$

$$\begin{bmatrix} 1 & \frac{2}{7} & \frac{4}{7} & \frac{5}{7} \\ 0 & -\frac{3}{7} & \frac{15}{7} & \frac{24}{7} \end{bmatrix} 2R_1 + R_2$$

$$\begin{bmatrix} 1 & \frac{2}{7} & \frac{4}{7} & \frac{5}{7} \\ 0 & 1 & -5 & -8 \end{bmatrix} -\frac{7}{3}R_2$$

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -5 & -8 \end{bmatrix} -\frac{2}{7}R_2 + R_1$$

so $$\underset{B\leftarrow C}{\boldsymbol{P}} = \begin{bmatrix} 2 & 3 \\ -5 & -8 \end{bmatrix}$$

(b) By part (a) and property (5) above (with $\boldsymbol{B}$ and $\boldsymbol{C}$ interchanged),

$$\underset{C\leftarrow B}{\boldsymbol{P}} = (\underset{B\leftarrow C}{\boldsymbol{P}})^{-1} = \frac{1}{-1}\begin{bmatrix} -8 & -3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 3 \\ -5 & -2 \end{bmatrix}.$$

Example 5 Let $\boldsymbol{b_1} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\boldsymbol{b_2} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $\boldsymbol{c_1} = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $\boldsymbol{c_2} = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$ and consider the bases for $\boldsymbol{R^2}$ given by $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}\}$ and $\boldsymbol{C} = \{\boldsymbol{c_1}, \boldsymbol{c_2}\}$.

(a) Find the change – of – coordinates matrix from $\boldsymbol{C}$ to $\boldsymbol{B}$.

(b) Find the change – of – coordinates matrix from $\boldsymbol{B}$ to $\boldsymbol{C}$.

Solution

(a) Notice that $\underset{B\leftarrow C}{\boldsymbol{P}}$ is needed rather than $\underset{C\leftarrow B}{\boldsymbol{P}}$ and compute

$$[\boldsymbol{b_1} \quad \boldsymbol{b_2} \mid \boldsymbol{c_1} \quad \boldsymbol{c_2}] = \left[\begin{array}{cc|cc} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array}\right]$$

so $$\underset{B\leftarrow C}{P} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}$$

(b) By part (a) and property (5) above (with ***B*** and ***C*** interchanged),

$$\underset{C\leftarrow B}{P} = (\underset{B\leftarrow C}{P})^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix}.$$

Example 6

1. Let $\boldsymbol{F} = \{\boldsymbol{f_1}, \boldsymbol{f_2}\}$ and $\boldsymbol{G} = \{\boldsymbol{g_1}, \boldsymbol{g_2}\}$ be bases for a vector space ***V***, and let ***P*** be a matrix whose columns are $[\boldsymbol{f_1}]_{\boldsymbol{G}}$. Which of the following equations is satisfied by ***P*** for all ***v*** in ***V***?

(i) $[\boldsymbol{v}]_{\boldsymbol{F}} = \boldsymbol{P}[\boldsymbol{v}]_{\boldsymbol{G}}$ (ii) $[\boldsymbol{v}]_{\boldsymbol{G}} = \boldsymbol{P}[\boldsymbol{v}]_{\boldsymbol{F}}$

2. Let ***B*** and ***C*** be as in Example 1. Use the results of that example to find the change-of-coordinates matrix from ***C*** to ***B***.

Solution

1. Since the columns of ***P*** are ***G***-coordinate vectors, a vector of the form ***Px*** must be a ***G***-coordinate vector. Thus ***P*** satisfies equation (ii).
2. The coordinate vectors found in Example 1 show that

$$\underset{C\leftarrow B}{P} = \begin{bmatrix} [\boldsymbol{b_1}]_C & [\boldsymbol{b_2}]_C \end{bmatrix} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}$$

Hence $$\underset{B\leftarrow C}{P} = \left(\underset{C\leftarrow B}{P}\right)^{-1} = \frac{1}{10}\begin{bmatrix} 1 & 6 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.6 \\ -0.1 & 0.4 \end{bmatrix}$$

Exercises

1. Let $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}\}$ and $\boldsymbol{C} = \{\boldsymbol{c_1}, \boldsymbol{c_2}\}$ be bases for a vector space ***V***, and suppose that $\boldsymbol{b_1} = 6\,\boldsymbol{c_1} - 2\,\boldsymbol{c_2}$ and $\boldsymbol{b_2} = 9\,\boldsymbol{c_1} - 4\,\boldsymbol{c_2}$.
(a) Find the change-of-coordinates matrix from ***B*** to ***C***.
(b) Find $[\boldsymbol{x}]_{\boldsymbol{C}}$ for $\boldsymbol{x} = -3\,\boldsymbol{b_1} + 2\,\boldsymbol{b_2}$. Use part (a).

2. Let $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}\}$ and $\boldsymbol{C} = \{\boldsymbol{c_1}, \boldsymbol{c_2}\}$ be bases for a vector space ***V***, and suppose that $\boldsymbol{b_1} = -\boldsymbol{c_1} + 4\,\boldsymbol{c_2}$ and $\boldsymbol{b_2} = 5\,\boldsymbol{c_1} - 3\,\boldsymbol{c_2}$.
(a) Find the change-of-coordinates matrix from ***B*** to ***C***.
(b) Find $[\boldsymbol{x}]_{\boldsymbol{C}}$ for $\boldsymbol{x} = 5\,\boldsymbol{b_1} + 3\,\boldsymbol{b_2}$.

3. Let $\boldsymbol{U} = \{\boldsymbol{u_1}, \boldsymbol{u_2}\}$ and $\boldsymbol{W} = \{\boldsymbol{w_1}, \boldsymbol{w_2}\}$ be bases for ***V***, and let ***P*** be a matrix whose columns are $[\boldsymbol{u_1}]_{\boldsymbol{W}}$ and $[\boldsymbol{u_2}]_{\boldsymbol{W}}$. Which of the following equations is satisfied by ***P*** for all ***x*** in ***V***?
(i) $[\boldsymbol{x}]_{\boldsymbol{U}} = \boldsymbol{P}[\boldsymbol{x}]_{\boldsymbol{W}}$ (ii) $[\boldsymbol{x}]_{\boldsymbol{W}} = \boldsymbol{P}[\boldsymbol{x}]_{\boldsymbol{U}}$

4. Let $\boldsymbol{A} = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$and $\boldsymbol{D} = \{\boldsymbol{d_1}, \boldsymbol{d_2}, \boldsymbol{d_3}\}$ be bases for ***V***, and let $\boldsymbol{P} = \begin{bmatrix} [\boldsymbol{d_1}]_{\boldsymbol{A}} & [\boldsymbol{d_2}]_{\boldsymbol{A}} & [\boldsymbol{d_3}]_{\boldsymbol{A}} \end{bmatrix}$. Which of the following equations is satisfied by ***P*** for all ***x*** in ***V***?
(i) $[\boldsymbol{x}]_{\boldsymbol{A}} = \boldsymbol{P}[\boldsymbol{x}]_{\boldsymbol{D}}$ (ii) $[\boldsymbol{x}]_{\boldsymbol{D}} = \boldsymbol{P}[\boldsymbol{x}]_{\boldsymbol{A}}$

5. Let $\boldsymbol{A} = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}\}$ and $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \boldsymbol{b_3}\}$ be bases for a vector space $\boldsymbol{V}$, and suppose that $\boldsymbol{a_1} = 4\boldsymbol{b_1} - \boldsymbol{b_2}$, $\boldsymbol{a_2} = -\boldsymbol{b_1} + \boldsymbol{b_2} + \boldsymbol{b_3}$, and $\boldsymbol{a_3} = \boldsymbol{b_2} - 2\,\boldsymbol{b_3}$.
(a) Find the change-of-coordinates matrix from $\boldsymbol{A}$ to $\boldsymbol{B}$.
(b) Find $[\boldsymbol{x}]_{\boldsymbol{B}}$ for $\boldsymbol{x} = 3\,\boldsymbol{a_1} + 4\,\boldsymbol{a_2} + \boldsymbol{a_3}$.

In exercises 6 to 9, let $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}\}$ and $\boldsymbol{C} = \{\boldsymbol{c_1}, \boldsymbol{c_2}\}$ be bases for $\boldsymbol{R}^2$. In each exercise, find the change-of-coordinates matrix from $\boldsymbol{B}$ to $\boldsymbol{C}$, and change-of-coordinates matrix from $\boldsymbol{C}$ to $\boldsymbol{B}$.

6. $\boldsymbol{b_1} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \boldsymbol{c_1} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \boldsymbol{c_2} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

7 $\boldsymbol{b_1} = \begin{bmatrix} -1 \\ 8 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \boldsymbol{c_1} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \boldsymbol{c_2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

8. $\boldsymbol{b_1} = \begin{bmatrix} -6 \\ -1 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \boldsymbol{c_1} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \boldsymbol{c_2} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$

9. $\boldsymbol{b_1} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}, \boldsymbol{b_2} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \boldsymbol{c_1} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \boldsymbol{c_2} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$

10. In $\boldsymbol{P_2}$, find the change-of-coordinate matrix from the basis $\boldsymbol{B} = \{\mathbf{1} - 2\,\boldsymbol{t} + \boldsymbol{t}^2, \mathbf{3} - 5\,\boldsymbol{t} + 4\,\boldsymbol{t}^2, 2\,\boldsymbol{t} + 3\,\boldsymbol{t}^2\}$ to the standard basis $\boldsymbol{C} = \{\mathbf{1}, \boldsymbol{t}, \boldsymbol{t}^2\}$. Then find the $\boldsymbol{B}$-coordinate vector for $-\mathbf{1} + 2\,\boldsymbol{t}$.

11. In $\boldsymbol{P_2}$, find the change-of-coordinates matrix from the basis $\boldsymbol{B} = \{\mathbf{1} - 3\,\boldsymbol{t}^2, \mathbf{2} + \boldsymbol{t} - 5\,\boldsymbol{t}^2, \mathbf{1} + 2\,\boldsymbol{t}\}$ to the standard basis. Then write $\boldsymbol{t}^2$ as a linear combination of the polynomials in $\boldsymbol{B}$.

12. Let $\boldsymbol{P} = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -5 & 0 \\ 4 & 6 & 1 \end{bmatrix}, \boldsymbol{v_1} = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix}, \boldsymbol{v_2} = \begin{bmatrix} -8 \\ 5 \\ 2 \end{bmatrix}, \boldsymbol{v_3} = \begin{bmatrix} -7 \\ 2 \\ 6 \end{bmatrix}$
(a). Find a basis $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ for $\boldsymbol{R}^3$ such that $\boldsymbol{P}$ is the change-of-coordinates matrix from $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ to the basis $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$.
(b). Find a basis $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_3}\}$ for $\boldsymbol{R}^3$ such that $\boldsymbol{P}$ is the change-of-coordinates matrix from $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ to $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_3}\}$.

Lecture 27

Applications to Difference Equations

Recently Discrete or Digital data has been used preferably rather then continuous data in scientific and engineering problems. Difference equation is considered more reliable tool to analyze such type of data even if we are using a differential equation to analyze continuous process, a numerical solution is often produced from a related difference equation.
In this lecture we will study some fundamental properties of linear difference equation that are considered best tool in Linear Algebra.

Discrete-Time Signals Let $\boldsymbol{S}$ is the space of discrete-time signals. A signal in $\boldsymbol{S}$ is a function defined only on the integers and is visualized as a sequence of numbers, say, $\{\boldsymbol{y_k}\}$.
Digital signals obviously arise in electrical and control systems engineering, but discrete-data sequences are also generated in biology, physics, economics, demography and many other areas, wherever a process is measured, or sampled, at discrete time intervals.
When a process beings at a specific time, it is sometimes convenient to write a signal as a sequence of the form $(y_0, y_1, y_2, ...)$. The term y_k for k<0 either are assumed to be zero or are simply omitted.
Discrete time signals are functions defined on integers that are sequences or it takes only a discrete set of values.
For Example a Radio Station broadcast's weather report once in a day so the sampling period is discrete, but sampling may be uniform or constant accordingly.

Example 1 The crystal clear sounds from a compact disc player are produced from music that has been sampled at the rate of 44,100 times per second. At each measurement, the amplitude of the music signal is recorded as a number, say, $\boldsymbol{y_k}$. The original music is composed of many different sounds of varying frequencies, yet the sequence $\{\boldsymbol{y_k}\}$ contains enough information to reproduce all the frequencies in the sound up to about 20,000 cycles per second, higher than the human ear can sense.

Linear Independence in the Space *S* of Signals To simplify notation, we consider a set of only three signals in $\boldsymbol{S}$, say, $\{\boldsymbol{u_k}\}$,$\{\boldsymbol{v_k}\}$and $\{\boldsymbol{w_k}\}$. They are linearly independent precisely when the equation

$$c_1\boldsymbol{u_k} + c_2\boldsymbol{v_k} + c_3\boldsymbol{w_k} = 0 \quad \text{for all } k \qquad (1)$$

implies that $c_1 = c_2 = c_3 = 0$. The phrase "for all k" means for all integers positive, negative and zero. One could also consider signals that start with $k = 0$, for example, in a case "for all k" would mean for all integers $k \geq 0$.
Suppose c_1, c_2, c_3 satisfy (1). Then the equation (1) holds for any three consecutive values of k, say, k, $k + 1$ and $k + 2$. Thus (1) implies that

$$c_1\boldsymbol{u_{k+1}} + c_2\boldsymbol{v_{k+1}} + c_3\boldsymbol{w_{k+1}} = 0 \text{ for all } k$$

and
$$c_1\boldsymbol{u_{k+2}} + c_2\boldsymbol{v_{k+2}} + c_3\boldsymbol{w_{k+2}} = 0 \quad \text{for all } k$$

Hence c_1, c_2, c_3 satisfy

$$\begin{bmatrix} \boldsymbol{u_k} & \boldsymbol{v_k} & \boldsymbol{w_k} \\ \boldsymbol{u_{k+1}} & \boldsymbol{v_{k+1}} & \boldsymbol{w_{k+1}} \\ \boldsymbol{u_{k+2}} & \boldsymbol{v_{k+2}} & \boldsymbol{w_{k+2}} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{for all } k \qquad (2)$$

The coefficient matrix in this system is called the **Casorati matrix** of the signals, and the determinant of the matrix is called the **Casoratian** of $\{\boldsymbol{u_k}\}$,$\{\boldsymbol{v_k}\}$, and $\{\boldsymbol{w_k}\}$. If for at least one value of k the Casorati matrix is invertible, then (2) will imply that $c_1 = c_2 = c_3 = 0$, which will prove that the three signals are linearly independent.

Example 2 Verify that $\boldsymbol{1^k}$, $(\boldsymbol{-2})^k$ and $\boldsymbol{3^k}$ are linearly independent signals.

Solution The Casorati matrix is $\begin{bmatrix} \mathbf{1}^k & (\mathbf{-2})^k & \mathbf{3}^k \\ \mathbf{1}^{k+1} & (\mathbf{-2})^{k+1} & \mathbf{3}^{k+1} \\ \mathbf{1}^{k+2} & (\mathbf{-2})^{k+2} & \mathbf{3}^{k+2} \end{bmatrix}$

Row operations can show fairly easily that this matrix is always invertible. However; it is faster to substitute a value for k – say, $k = 0$ – and row reduce the numerical matrix:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 3 \\ 1 & 4 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 1 & 4 & 9 \end{bmatrix} R_2 - R_1$$

$$\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 3 & 8 \end{bmatrix} R_3 - R_1 \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 0 & 10 \end{bmatrix} R_3 + R_2$$

The Casorati matrix is invertible for $k = 0$. So $\boldsymbol{1^k}$, $(\boldsymbol{-2})^k$ and $\boldsymbol{3^k}$ are linearly independent.

If a Casorati matrix is not invertible, the associated signals being tested may or may not be linearly dependent. However, it can be shown that if the signals are all solutions of the same homogeneous difference equation (described below), then either the Casorati matrix is invertible for all k and the signals are linearly independent, or else for all k the Casorati matrix is not invertible and the signals are linearly dependent.

Activity

Verify that 2^k, $(-1)^k$ and 0^k are linearly independent signals.

Linear Difference Equations Given scalars $\boldsymbol{a_0}, \dots, \boldsymbol{a_n}$, with $\boldsymbol{a_0}$ and $\boldsymbol{a_n}$ nonzero, and given a signal $\{\boldsymbol{z_k}\}$, the equation

$$a_0 \boldsymbol{y_{k+n}} + a_1 \boldsymbol{y_{k+n-1}} + \dots + a_{n-1} \boldsymbol{y_{k+1}} + a_n \boldsymbol{y_k} = \boldsymbol{z_k} \quad \text{for all } k \qquad (3)$$

is called a **linear difference equation** (or **linear recurrence relation**) of order n. For simplicity, $\boldsymbol{a_0}$ is often taken equal to 1. If $\{\boldsymbol{z_k}\}$ is the zero sequence, the equation is **homogeneous;** otherwise, the equation is **non-homogeneous**.

In simple words, an equation which expresses a value of a sequence as a function of the other terms in the sequence is called a difference equation.
In particular an equation which expresses the value a_n of a sequence $\{a_n\}$ as a function of the term a_{n-1} is called a first order difference equation.

Example 3 In digital signal processing, a difference equation such as (3) above describes a linear filter and $a_0, \ldots, a_n$ are called the **filter coefficients**. If $\{y_k\}$ is treated as the input and $\{z_k\}$ the output, then the solutions of the associated homogeneous equation are the signals that are filtered out and transformed into the zero signal. Let us feed two different signals into the filter

$$0.35y_{k+2} + 0.5y_{k+1} + 0.35y_k = z_k$$

Here 0.35 is an abbreviation for $\sqrt{2}/4$. The first signal is created by sampling the continuous signal $y = \cos(\pi t/4)$ at integer values of t, as in Fig. 3 (a). The discrete signals is $\{y_k\} = \{\ldots \cos(0),\ \cos(\pi/4), \cos(2\pi/4), \cos(3\pi/4) \ldots\}$

For simplicity, write ± 0.7 in place of $\pm\sqrt{2}/2$, so that

$$\{y_k\} = \{\ \ldots,\ \underset{\substack{\uparrow \\ k=0}}{1}, 0.7, 0, -0.7, -1, -0.7, 0, 0.7, 1, 0.7, 0, \ldots\}$$

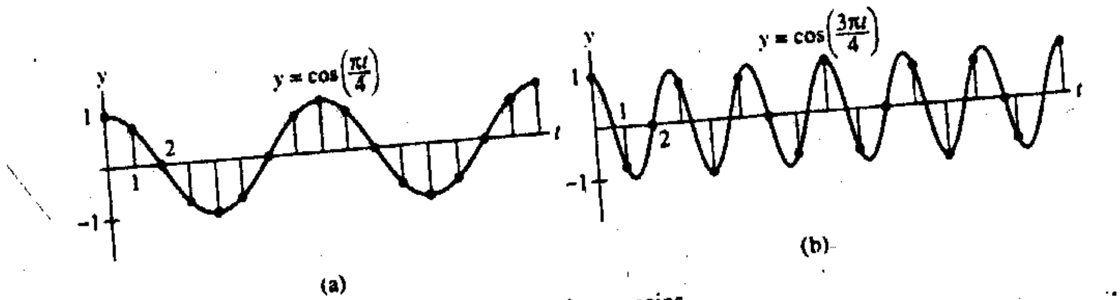


Figure Discrete signals with different frequencies

The following table shows a calculation of the output sequence $\{z_k\}$, where $0.35(0.7)$ is an abbreviation for $(\sqrt{2}/4)(\sqrt{2}/2) = 0.25$. The output is $\{y_k\}$, shifted by one term.

Table Computing the output of a filter

K	y_k	y_{k+1}	y_{k+2}	$0.35y_k+0.5y_{k+1}+0.35y_{k+2} = z_k$
0	**1**	**0.7**	**0**	$0.35(\mathbf{1})+0.5(\mathbf{0.7})+0.35(\mathbf{0})=0.7$
1	**0.7**	**0**	**-0.7**	**0**
2	**0**	**-0.7**	**-1**	**-0.7**
3	**-0.7**	**-1**	**-0.7**	**-1**
4	**-1**	**-0.7**	**0**	**-0.7**
5	**-0.7**	**0**	**0.7**	**0**
⋮	⋮	⋮	⋮	⋮

A different input signal is produced from the higher frequency signal $y = \cos(3\pi t/4)$.
Sampling at the same rate as before produces a new input sequence:

$$\{w_k\} = \{\ldots, 1, -0.7, 0, 0.7, -1, 0.7, 0, -0.7, 1, -0.7, 0, \ldots\}$$
$$\uparrow\ k = 0$$

When $\{w_k\}$ is fed into the filter, the output is the zero sequence. The filter, called a **low pass filter,** lets $\{y_k\}$ pass through, but stops the higher frequency $\{w_k\}$.
In many applications, a sequence $\{z_k\}$ is specified for the right side of a difference equation (3) and a $\{y_k\}$ that satisfies (3) is called a solution of the equation. The next example shows how to find solution for a homogeneous equation.

Example 4 Solutions of a homogeneous difference equation often have the form $y_k = r^k$ for some r. Find some solutions of the equation

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \text{ for all } k \qquad (4)$$

Solution Substitute r^k for y_k in the equation and factor the left side:

$$r^{k+3} - 2r^{k+2} - 5r^{k+1} + 6r^k = 0 \qquad (5)$$

By synthetic division

$$r^k(r^3 - 2r^2 - 5r + 6) = 0$$

	1	-2	-5	6
1	0	1	-1	-6
	1	-1	-6	0

So we get

$$r^k[(x-1)(x^2 - x - 6)] = 0$$
$$r^k[(x-1)(x^2 - 3x + 2x - 6)] = 0$$
$$r^k[(x-1)\{x(x-3) + 2(x-3)\}] = 0$$
$$r^k(r-1)(r+2)(r-3) = 0 \qquad (6)$$

since (5) is equivalent to (6), r^k satisfies the difference equation (4) if and only if r satisfies (6). Thus 1^k, $(-2)^k$ and 3^k are all solutions of (4). For instance, to verify that 3^k is a solution of (4), compute

$$3^{k+3} - 2.3^{k+2} - 5.3^{k+1} + 6.3^k = 3^k(27 - 18 - 15 + 6) = 0 \text{ for all k}$$

In general, a nonzero signal r^k satisfies the homogeneous difference equation

$$y_{k+n} + a_1 y_{k+n-1} + \ldots + a_{n-1} y_{k+1} + a_n y_k = 0 \text{ for all } k$$

if and only if r is a root of the **auxiliary equation**

$$r^n + a_1 r^{n-1} + \ldots + a_{n-1} r + a_n .1 = 0$$

We will not consider the case when r is a repeated root of the auxiliary equation. When the auxiliary equation has a complex root, the difference equation has solutions of the form $s^k \cos kw$ and $s^k \sin kw$, for constants s and w. This happened in Example 3.

Solution Sets of Linear Difference Equations

Given $\boldsymbol{a_1}, \dots, \boldsymbol{a_n}$, consider the mapping $\boldsymbol{T}: \boldsymbol{S} \to \boldsymbol{S}$ that transforms a signal $\{\boldsymbol{y_k}\}$into a signal $\{\boldsymbol{w_k}\}$ given by

$$\boldsymbol{w_k} = \boldsymbol{y_{k+n}} + a_1\boldsymbol{y_{k+n-1}} + \dots + a_{n-1}\boldsymbol{y_{k+1}} + a_n\boldsymbol{y_k}$$

It is readily checked that $\boldsymbol{T}$ is a linear transformation. This implies that the solution set of the homogeneous equation

$$\boldsymbol{y_{k+n}} + a_1\boldsymbol{y_{k+n-1}} + \dots + a_{n-1}\boldsymbol{y_{k+1}} + a_n\boldsymbol{y_k} = \boldsymbol{0} \text{ for all } k$$

is the kernel of $\boldsymbol{T}$ (kernel is the set of signals that $\boldsymbol{T}$ maps into the zero signal) and hence the solution set is a subspace of $\boldsymbol{S}$. Any linear combination of solutions is again a solution.

Theorem If $a_n \neq 0$ and if $\{\boldsymbol{z_k}\}$ is given, the equation

$$\boldsymbol{y_{k+n}} + a_1\boldsymbol{y_{k+n-1}} + \dots + a_{n-1}\boldsymbol{y_{k+1}} + a_n\boldsymbol{y_k} = \boldsymbol{z_k} \text{ for all } k \tag{7}$$

has a unique solution whenever $\boldsymbol{y_0}, \dots, \boldsymbol{y_{n-1}}$ are specified.

Proof If $\boldsymbol{y_0}, \dots, \boldsymbol{y_{n-1}}$ are specified, use (7) to define

$$\boldsymbol{y_n} = \boldsymbol{z_0} - [a_1\boldsymbol{y_{n-1}} + \dots + a_n\boldsymbol{y_0}]$$

And now that $\boldsymbol{y_1}, \dots, \boldsymbol{y_n}$ are specified, use (7) to define $\mathbf{y_{n+1}}$. In general, use the recursion relation

$$\boldsymbol{y_{n+k}} = \boldsymbol{z_k} - [a_1\boldsymbol{y_{k+n-1}} + \dots + a_n\boldsymbol{y_k}] \tag{8}$$

to define $\boldsymbol{y_{n+k}}$ for $k \geq 0$. To define $\boldsymbol{y_k}$ for $k<0$, use the recursion relation

$$\boldsymbol{y_k} = \frac{1}{a_n}\boldsymbol{z_k} - \frac{1}{a_n}[\boldsymbol{y_{k+n}} + a_1\boldsymbol{y_{k+n-1}} + \dots + a_{n-1}\boldsymbol{y_{k+1}}] \tag{9}$$

This produces a signal that satisfies (7). Conversely, any signal that satisfies (7) for all k certainly satisfies (8) and (9) so the solution of (7) is unique.

Theorem The set $\boldsymbol{H}$ of all solutions of the nth-order homogeneous linear difference equation

$$\boldsymbol{y_{k+n}} + a_1\boldsymbol{y_{k+n-1}} + \dots + a_{n-1}\boldsymbol{y_{k+1}} + a_n\boldsymbol{y_k} = \boldsymbol{0} \text{ for all } k \tag{10}$$

is an n-dimensional vector space.

Proof We explained earlier why $\boldsymbol{H}$ is a subspace of $\boldsymbol{S}$. For $\{\boldsymbol{y_k}\}$in $\boldsymbol{H}$, let $F\{\boldsymbol{y_k}\}$be the vector in $\boldsymbol{R^n}$ given by $(\boldsymbol{y_0}, \boldsymbol{y_1}, \dots, \boldsymbol{y_{n-1}})$. It is readily verified that $\boldsymbol{F}:\boldsymbol{H} \to \boldsymbol{R^n}$ is a linear transformation. Given any vector $(\boldsymbol{y_0}, \boldsymbol{y_1}, \dots, \boldsymbol{y_{n-1}})$ in $\boldsymbol{R^n}$, the previous theorem says that there is a unique signal $\{\boldsymbol{y_k}\}$ in $\boldsymbol{H}$ such that $\boldsymbol{F}\{\boldsymbol{y_k}\}=(\boldsymbol{y_0}, \boldsymbol{y_1}, \dots, \boldsymbol{y_{n-1}})$. It means that $\boldsymbol{F}$ is a one-to-one linear transformation of $\boldsymbol{H}$ onto $\boldsymbol{R^n}$; that is, F is an isomorphism. Thus dim $\boldsymbol{H}$ = dim $\boldsymbol{R^n} = n$.

Example 5 Find a basis for the set of all solutions to the difference equation

$$\boldsymbol{y_{k+3}} - 2\boldsymbol{y_{k+2}} - 5\boldsymbol{y_{k+1}} + 6\boldsymbol{y_k} = \boldsymbol{0} \text{ for all } k$$

Solution

Generally it is difficult to identify directly that a set of signals spans the solution space, but this problem is being resolved by two above key theorems, from the last theorem the solution space is exactly three – dimensional and the Basis Theorem describes that a linearly independent set of n vectors in an n – dimensional space is automatically a basis. So 1^k, $(-2)^k$ and 3^k form a basis for the solution space.

The standard way to describe the "general solution" of (10) is to exhibit a basis for the subspace of all solutions. Such a basis is usually called a **fundamental set of solutions** of (10). In practice, if we can find n linearly independent signals that satisfy (10), they will automatically span the n-dimensional solution space, as we saw in above example.

Non-homogeneous Equations The general solution of the non-homogeneous difference equation

$$\boldsymbol{y_{k+n}} + a_1\boldsymbol{y_{k+n-1}} + \ldots + a_{n-1}\boldsymbol{y_{k+1}} + a_n\boldsymbol{y_k} = \boldsymbol{z_k} \quad \text{for all } k \tag{11}$$

can be written as one particular solution of (11) plus an arbitrary linear combination of a fundamental set of solutions of the corresponding homogeneous equation (10). This fact is analogous to the result in lecture 7 about how the solution sets of $\boldsymbol{Ax} = \boldsymbol{b}$ and $\boldsymbol{Ax} = \boldsymbol{0}$ are parallel. Both results have the same explanation: The mapping $\boldsymbol{x} \rightarrow \boldsymbol{Ax}$ is linear, and the mapping that transforms the signal $\{\boldsymbol{y_k}\}$ into the signal $\{\boldsymbol{z_k}\}$ in (11) is linear.

Example 6 Verify that the signal $\boldsymbol{y_k} = k^2$ satisfies the difference equation

$$\boldsymbol{y_{k+2}} - 4\boldsymbol{y_{k+1}} + 3\boldsymbol{y_k} = -4k \quad \text{for all } k \tag{12}$$

Then find a description of all solutions of this equation.

Solution Substitute k^2 for $\boldsymbol{y_k}$ in the left side of (12):

$$(k+2)^2 - 4(k+1)^2 + 3k^2 = (k^2 + 4k + 4) - 4(k^2 + 2k + 1) + 3k^2 = -4k$$

So k^2 is indeed a solution of (12). The next step is to solve the homogeneous equation

$$\boldsymbol{y_{k+2}} - 4\boldsymbol{y_{k+1}} + 3\boldsymbol{y_k} = 0 \tag{13}$$

The auxiliary equation is

$$r^2 - 4r + 3 = r^2-3r-r+3 = r(r-3)-1(r-3) = (r-1)(r-3) = 0$$

The roots are $r = 1, 3$. So two solutions of the homogeneous difference equation are $\boldsymbol{1^k}$ and $\boldsymbol{3^k}$. They are obviously not multiples of each other, so they are linearly independent signals. (The Casorati test could have been used, too.) By the last Theorem, the solution space is two – dimensional, so $\boldsymbol{3^k}$ and $\boldsymbol{1^k}$ form a basis for the set of solutions of (13). Translating that set by a particular solution of the non-homogeneous equation (12), we obtain the general solution of (12):

$$\boldsymbol{k^2} + c_1\boldsymbol{1^k} + c_2\boldsymbol{3^k}, \quad \text{or} \quad \boldsymbol{k^2} + \boldsymbol{c_1} + c_2\boldsymbol{3^k}$$

Reduction to Systems of First-Order Equations A modern way to study a homogeneous nth-order linear difference equation is to replace it by an equivalent system of first order difference equations, written in the form $\boldsymbol{x_{k+1}} = \boldsymbol{Ax_k}$ for k = 0, 1, 2, … Where the vectors $\boldsymbol{x_k}$ are in $\boldsymbol{R^n}$ and $\boldsymbol{A}$ is an n x n matrix.

Example 7 Write the following difference equation as a first order system:

$$\boldsymbol{y_{k+3}} - 2\boldsymbol{y_{k+2}} - 5\boldsymbol{y_{k+1}} + 6\boldsymbol{y_k} = 0 \text{ for all } k$$

Solution For each k, set $\boldsymbol{x_k} = \begin{bmatrix} \boldsymbol{y_k} \\ \boldsymbol{y_{k+1}} \\ \boldsymbol{y_{k+2}} \end{bmatrix}$

The difference equation says that $y_{k+3} = 2y_{k+2} + 5y_{k+1} - 6y_k$, so

$$x_{k+1} = \begin{bmatrix} y_{k+1} \\ y_{k+2} \\ y_{k+3} \end{bmatrix} = \begin{bmatrix} 0 + y_{k+1} + 0 \\ 0 + 0 + y_{k+2} \\ -6y_k + 5y_{k+1} + 2y_{k+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix} = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

That is, $x_{k+1} = Ax_k$ for all k, where $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix}$

In general, the equation $y_{k+n} + a_1 y_{k+n-1} + \dots + a_{n-1} y_{k+1} + a_n y_k = 0$ for all k can be written as $x_{k+1} = Ax_k$ for all k, where

$$x_k = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+n-1} \end{bmatrix}, A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

Example 8 It can be shown that the signals 2^k, $3^k \, Sin\frac{k\pi}{2}$, and $3^k \, Cos\frac{k\pi}{2}$ are solutions of $y_{k+3} - 2y_{k+2} + 9y_{k+1} - 18y_k = 0$. Show that these signals form a basis for the set of all solutions of the difference equation.

Solution Examine the Casorati matrix:

$$C(k) = \begin{bmatrix} 2^k & 3^k \, Sin\frac{k\pi}{2} & 3^k \, Cos\frac{k\pi}{2} \\ 2^{k+1} & 3^{k+1} Sin\frac{(k+1)\pi}{2} & 3^{k+1} Cos\frac{(k+1)\pi}{2} \\ 2^{k+2} & 3^{k+2} Sin\frac{(k+2)\pi}{2} & 3^{k+2} Cos\frac{(k+2)\pi}{2} \end{bmatrix}$$

Set $k = 0$ and row reduce the matrix to verify that it has three pivot positions and hence is invertible:

$$C(0) = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 0 \\ 4 & 0 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -2 \\ 4 & 0 & -9 \end{bmatrix} R_2 - 2R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -2 \\ 0 & 0 & -13 \end{bmatrix} R_3 - 4R_1$$

The Casorati matrix is invertible at $k = 0$, so signals are linearly independent. Since there are three signals, and the solution space $\boldsymbol{H}$ of the difference equation has three-dimensions (Theorem 2), the signals form a basis for $\boldsymbol{H}$, by the Basis Theorem.

Exercises

Verify that the signals in exercises 1 and 2 are solution of the accompanying difference equation.

1. $\mathbf{2}^{\mathbf{k}}, (\mathbf{-4})^{\mathbf{k}}; \; y_{k+2} + 2\,y_{k+1} - 8\,y_k = 0$

2. $\mathbf{3}^{\mathbf{k}}, (\mathbf{-3})^{\mathbf{k}}; \; y_{k+2} - 9\,y_k = 0$

Show that the signals in exercises 3 to 6 form a basis for the solution set of accompanying difference equation.

3. $\mathbf{2}^{\mathbf{k}}, (\mathbf{-4})^{\mathbf{k}}; \; y_{k+2} + 2\,y_{k+1} - 8\,y_k = 0$

4. $\mathbf{3}^{\mathbf{k}}, (\mathbf{-3})^{\mathbf{k}}; \; y_{k+2} - 9\,y_k = 0$

5. $(\mathbf{-3})^{\mathbf{k}}, \boldsymbol{k}(\mathbf{-3})^{\mathbf{k}}; \; y_{k+2} + 6\,y_{k+1} + 9\,y_k = 0$

6. $\mathbf{5}^{\mathbf{k}}\text{Cos}\dfrac{k\pi}{2}, \mathbf{5}^{\mathbf{k}}\text{Sin}\dfrac{k\pi}{2}; \; y_{k+2} + 25\,y_k = 0$

In exercises 7 to 10, assume that the signals listed are solutions of the given difference equation. Determine if the signals form a basis for the solution space of the equation.

7. $\mathbf{1}^{\mathbf{k}}, \mathbf{3}^{\mathbf{k}}\text{Cos}\dfrac{k\pi}{2}, \mathbf{3}^{\mathbf{k}}\text{Sin}\dfrac{k\pi}{2}; \; y_{k+3} - y_{k+2} + 9\,y_{k+1} - 9\,y_k = 0$

8. $(\mathbf{-1})^{\mathbf{k}}, \boldsymbol{k}(\mathbf{-1})^{\mathbf{k}}, \mathbf{5}^{\mathbf{k}}; \; y_{k+3} - 3\,y_{k+2} - 9\,y_{k+1} - 5\,y_k = 0$

9. $(\mathbf{-1})^{\mathbf{k}}, \mathbf{3}^{\mathbf{k}}; \; y_{k+3} + y_{k+2} - 9\,y_{k+1} - 9\,y_k = 0$

10. $\mathbf{1}^{\mathbf{k}}, (\mathbf{-1})^{\mathbf{k}}; \; y_{k+4} - 2\,y_{k+2} + y_k = 0$

In exercises 11 and 12, find a basis for the solution space of the difference equation.

11. $y_{k+2} - 7\,y_{k+1} + 12\,y_k = 0$

12. $16y_{k+2} + 8\,y_{k+1} - 3\,y_k = 0$

In exercises 13 and 14, show that the given signal is a solution of the difference equation. Then find the general solution of that difference equation.

13. $y_k = \mathrm{k}^2; \; y_{k+2} + 3\,y_{k+1} - 4\,y_k = 10\,\mathrm{k} + 7$

14. $y_k = 2 - 2\,\mathrm{k}; \; y_{k+2} - (9/2)\,y_{k+1} + 2\,y_k = 3\,\mathrm{k} + 2$

Write the difference equations in exercises 15 and 16 as first order systems, $\boldsymbol{x}_{k+1} = \boldsymbol{A}\,\boldsymbol{x}_k$, for all k.

15. $y_{k+4} - 6\,y_{k+3} + 8\,y_{k+2} + 6\,y_{k+1} - 9\,y_k = 0$

16. $y_{k+3} - (3/4)y_{k+2} + (1/16)y_k = 0$

Lecture 28

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 $\boldsymbol{R^n}$ 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 $\boldsymbol{R^n}$ 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 $\boldsymbol{A}$.

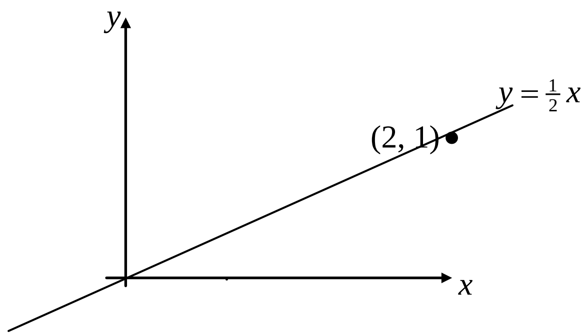


Figure 1

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.

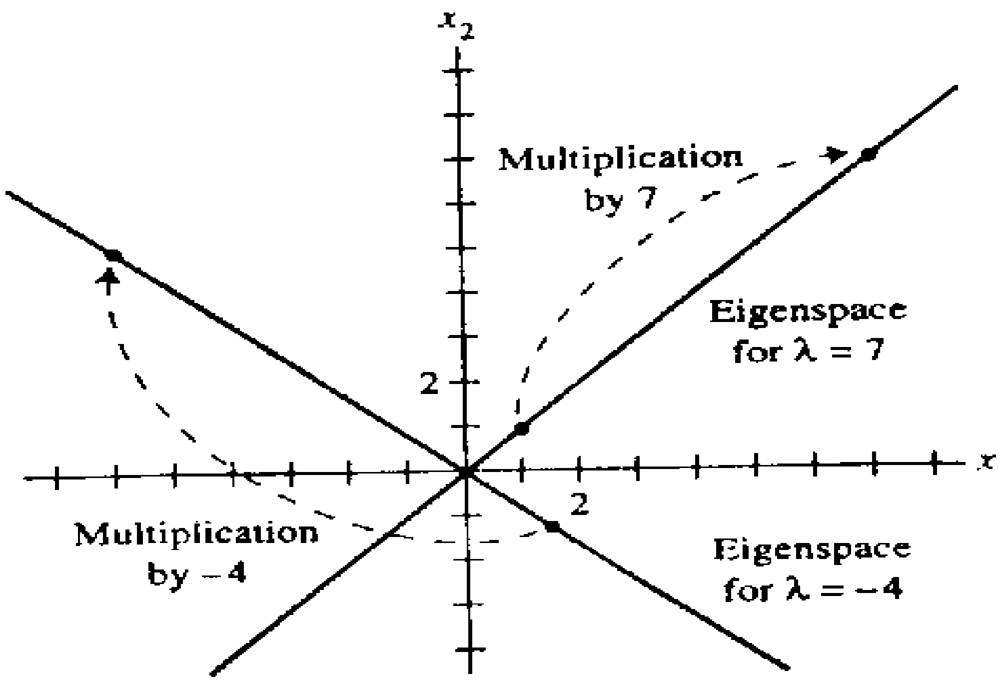


FIGURE 2 Eigenspaces for λ = −4 and λ = 7.

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 $A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$ and row reduce the augmented matrix for (A – 2I) x = 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 A because the equation (A – 2I) x = 0 has free variables. The general solution is

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

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

$$2x_1 = t - 6s$$

$$x_1 = \left(\tfrac{1}{2}\right)t - 3s$$

$$then$$

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

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 R^3. A basis is

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

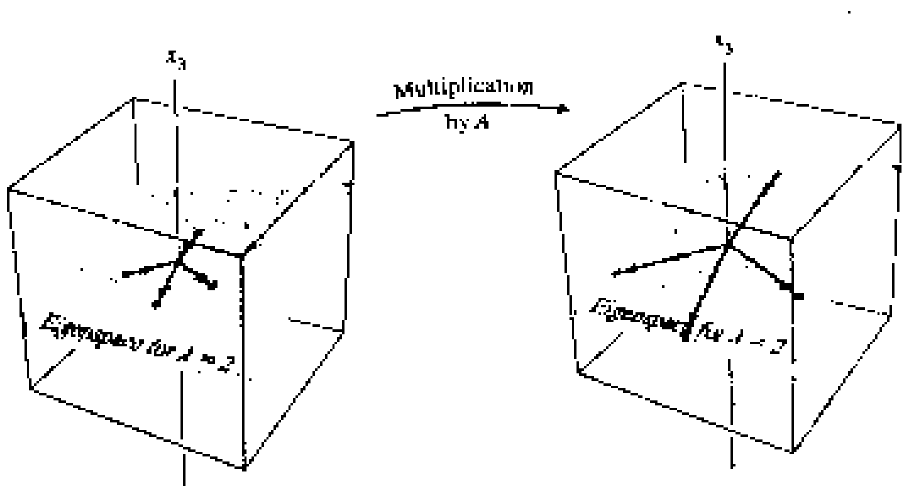


FIGURE 3 A acts as a dilation on the eigenspace.

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 29

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

$$(\lambda + 2)(\lambda - 5) = 0. \qquad (1)$$

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

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

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,

$$x = -t, y = t \ . \qquad (3)$$

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

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

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,

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

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).

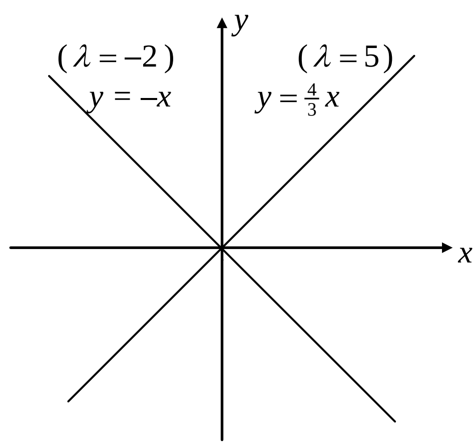


Figure 1(a)

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.

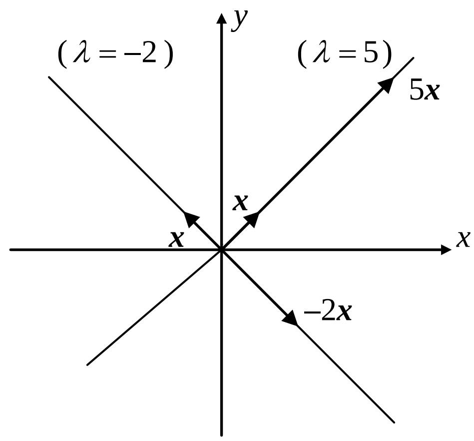


Figure 1(b)

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

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

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

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

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 $\pm 1, \pm 2,$ and ± 4. Substituting these values successively into Eq. (8) yields that λ =4 is an integer solution. This implies that λ – 4 is a factor of Eq.(7), Thus, dividing the polynomial by λ – 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

$\lambda = 4, \qquad \lambda = 2+\sqrt{3}, \qquad \lambda = 2-\sqrt{3}$

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

$$\det(B - \lambda I) = \det\left[P^{-1}(A - \lambda I)P\right]$$

$$= \det(P^{-1}).\det(A - \lambda I).\det(P) \qquad \text{(A)}$$

Since

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

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

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

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

$$= \lambda^2 - 1.92\lambda + .92$$

By the quadratic formula

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

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

$$x_1 = t \text{ and } x_2 = -t$$

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

$$\boldsymbol{x_1} = \boldsymbol{Ax_0} = \boldsymbol{c_1Av_1} + \boldsymbol{c_2Av_2} \quad \text{(Using linearity of } x \to Ax\text{)}$$

$$= c_1\boldsymbol{v_1} + c_2\,(.92)\boldsymbol{v_2} \quad (\boldsymbol{v_1} \text{ and } \boldsymbol{v_2} \text{ are eigenvectors.})$$

$$\boldsymbol{x_2} = \boldsymbol{Ax_1} = c_1\boldsymbol{Av_1} + c_2\,(.92)\boldsymbol{Av_2} = c_1\boldsymbol{v_1} + c_2\,(.92)^2\,\boldsymbol{v_2}.$$

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

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

which is a quadratic equation whose roots are given as

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

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 30

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 ***A*** = ***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, ***A*** = ***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 $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$ and if ***D*** is any diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$ then

$$\boldsymbol{AP} = \boldsymbol{A}\begin{bmatrix} \boldsymbol{v_1} & \boldsymbol{v_2} & \ldots & \boldsymbol{v_n} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Av_1} & \boldsymbol{Av_2} & \ldots & \boldsymbol{Av_n} \end{bmatrix}, \tag{1}$$

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

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

$$\begin{bmatrix} \boldsymbol{Av_1} & \boldsymbol{Av_2} & \cdots & \boldsymbol{Av_n} \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{v_1} & \lambda_2\boldsymbol{v_2} & \cdots & \lambda_n\boldsymbol{v_n} \end{bmatrix} \tag{3}$$

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 $\boldsymbol{v_1}, \ldots, v_n$ must be linearly independent. Also, since these columns are nonzero, Eq.(4) shows that $\lambda_1, \ldots, \lambda_n$ are eigenvalues and $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$ 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 $\boldsymbol{P}$ 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 $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}$ and if $\boldsymbol{P} = [\boldsymbol{v_1} \ldots \boldsymbol{v_n}]$, 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 $\lambda_1, \ldots, \lambda_p$.

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.

$$\text{Basis for } \lambda = 5 : \boldsymbol{v}_1 = \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} \text{ and } \boldsymbol{v}_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{Basis for } \lambda = -3 : \boldsymbol{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \boldsymbol{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The set $\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_4\}$is linearly independent, by Theorem 3. So the matrix $P = [\boldsymbol{v}_1 \ldots \boldsymbol{v}_4]$ 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 $A = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}, \boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, and $\boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suppose you are told that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ 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 31

Eigenvectors and Linear Transformations

The goal of this lecture is to investigate the relationship between eigenvectors and linear transformations. Previously, we have learned how to find the eigenvalues and eigenvectors. We shall see that the matrix factorization $\boldsymbol{A = PDP^{-1}}$ is also a type of linear transformations for some invertible matrix $\boldsymbol{P}$ and some diagonal matrix $\boldsymbol{D}$.

The Matrix of a Linear Transformation

Let $\boldsymbol{V}$ be an n-dimensional vector space, $\boldsymbol{W}$ an m-dimensional vector space, and $\boldsymbol{T}$ any linear transformation from $\boldsymbol{V}$ to $\boldsymbol{W}$. To associate a matrix with $\boldsymbol{T}$, we choose (ordered) bases $\boldsymbol{B}$ and $\boldsymbol{C}$ for $\boldsymbol{V}$ and $\boldsymbol{W}$, respectively.
Given any $\boldsymbol{x}$ in $\boldsymbol{V}$, the coordinate vector $[\boldsymbol{x}]_{\mathbf{B}}$ is in $\boldsymbol{R^n}$ and the coordinate vector of its image, $[\boldsymbol{T}(\boldsymbol{x})]_{\mathbf{C}}$, is in $\boldsymbol{R^m}$, as shown in Fig. 1.

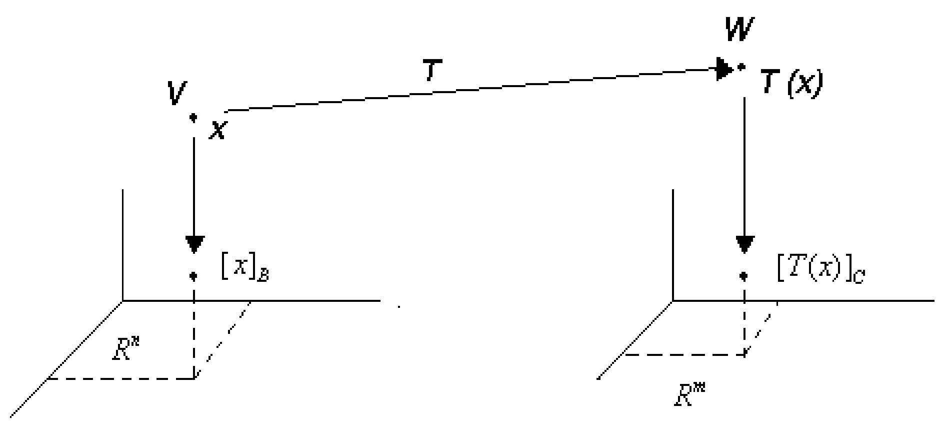


Figure 1 A linear transformation from $\boldsymbol{V}$ to $\boldsymbol{W}$

Let $\boldsymbol{B}=\{\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}\}$ be the basis for $\boldsymbol{V}$. If $\boldsymbol{x} = r_1\boldsymbol{b_1} + \cdots + r_n\boldsymbol{b_n}$,

then $[\boldsymbol{x}]_{\boldsymbol{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$

and T is a linear transformation

$$\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{T}(r_1\boldsymbol{b_1} + \cdots + r_n\boldsymbol{b_n}) = r_1\boldsymbol{T}(\boldsymbol{b_1}) + \cdots + r_n\boldsymbol{T}(\boldsymbol{b_n}) \quad (1)$$

Using the basis $\boldsymbol{C}$ in $\boldsymbol{W}$, we can rewrite (1) in terms of $\boldsymbol{C}$-coordinate vectors:

$$[\boldsymbol{T}(\boldsymbol{x})]_{\boldsymbol{C}} = r_1[\boldsymbol{T}(\boldsymbol{b_1})]_{\boldsymbol{C}} + \cdots + r_n[\boldsymbol{T}(\boldsymbol{b_n})]_{\boldsymbol{C}} \qquad (2)$$

Since $\boldsymbol{C}$-coordinate vectors are in $\boldsymbol{R^m}$, the vector equation (2) can be written as a matrix equation, namely

$$[\boldsymbol{T}(\boldsymbol{x})]_{\boldsymbol{C}} = \boldsymbol{M}[\boldsymbol{x}]_{\boldsymbol{B}} \qquad (3)$$

where $$\boldsymbol{M} = \left[[\boldsymbol{T}(\boldsymbol{b_1})]_{\boldsymbol{C}} \quad [\boldsymbol{T}(\boldsymbol{b_2})]_{\boldsymbol{C}} \quad \cdots \quad [\boldsymbol{T}(\boldsymbol{b_n})]_{\boldsymbol{C}}\right] \qquad (4)$$

The matrix $\boldsymbol{M}$ is a matrix representation of $\boldsymbol{T}$, called the **matrix for $\boldsymbol{T}$ relative to the bases $\boldsymbol{B}$ and $\boldsymbol{C}$**.

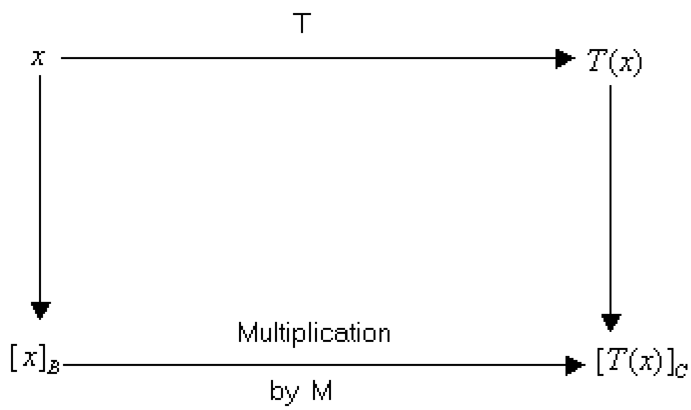


Figure 2

Example 1

Suppose that $\boldsymbol{B} = \{\boldsymbol{b_1}, \boldsymbol{b_2}\}$ is a basis for $\boldsymbol{V}$ and $\boldsymbol{C} = \{\boldsymbol{c_1}, \boldsymbol{c_2}, \boldsymbol{c_3}\}$ is a basis for $\boldsymbol{W}$. Let $\boldsymbol{T}:\boldsymbol{V} \to \boldsymbol{W}$ be a linear transformation with the property that

$$\boldsymbol{T}(\boldsymbol{b_1}) = 3\boldsymbol{c_1} - 2\boldsymbol{c_2} + 5\boldsymbol{c_3} \text{ and } \boldsymbol{T}(\boldsymbol{b_2}) = 4\boldsymbol{c_1} + 7\boldsymbol{c_2} - \boldsymbol{c_3}$$

Find the matrix $\boldsymbol{M}$ for $\boldsymbol{T}$ relative to $\boldsymbol{B}$ and $\boldsymbol{C}$.

Solution

Since $\boldsymbol{M} = \left[[\boldsymbol{T}(\boldsymbol{b_1})]_{\boldsymbol{C}} \quad [\boldsymbol{T}(\boldsymbol{b_2})]_{\boldsymbol{C}} \quad \cdots \quad [\boldsymbol{T}(\boldsymbol{b_n})]_{\boldsymbol{C}}\right]$ and here

$$[\boldsymbol{T}(\boldsymbol{b_1})]_{\boldsymbol{C}} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \text{ and } [\boldsymbol{T}(\boldsymbol{b_2})]_{\boldsymbol{C}} = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$$

Hence, $$\boldsymbol{M} = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$$

Linear Transformations from V into V

In general when ***W*** is the same as ***V*** and the basis ***C*** is the same as ***B***, the matrix ***M*** in (4) is called the **matrix for *T* relative to *B***, or simply the ***B*-matrix** for ***T***, is denoted by $[T]_B$. See Fig. 3.

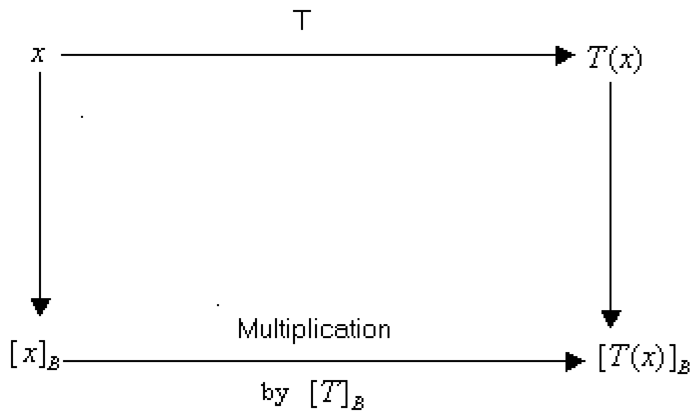


Figure 3

The ***B***-matrix of $T: V \to V$ satisfies

$$[T(x)]_B = [T]_B[x]_B, \quad \text{for all } x \text{ in } V \tag{5}$$

Example 2

The mapping $T: P_2 \to P_2$ defined by $T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$ is a linear transformation.

(a) Find the ***B***-matrix for ***T***, when ***B*** is the basis $\{1, t, t^2\}$.
(b) Verify that $[T(p)]_B = [T]_B[p]_B$ for each ***p*** in P_2.

Solution

(a) We have to find the ***B***-matrix of $T: V \to V$ satisfies

$$[T(x)]_B = [T]_B[x]_B, \quad \text{for all } x \text{ in } V$$

Since $T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$ therefore

$$T(1) = 0, \; T(t) = 1, \; T(t^2) = 2t$$

$$[T(1)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, [T(t)]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, [T(t^2)]_B = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$[T]_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

(b) For a general $p(t) = a_0 + a_1 t + a_2 t^2$, we have

$$[T(p)]_B = [a_1 + 2a_2 t]_B = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_B [p]_B$$

See figure 4.

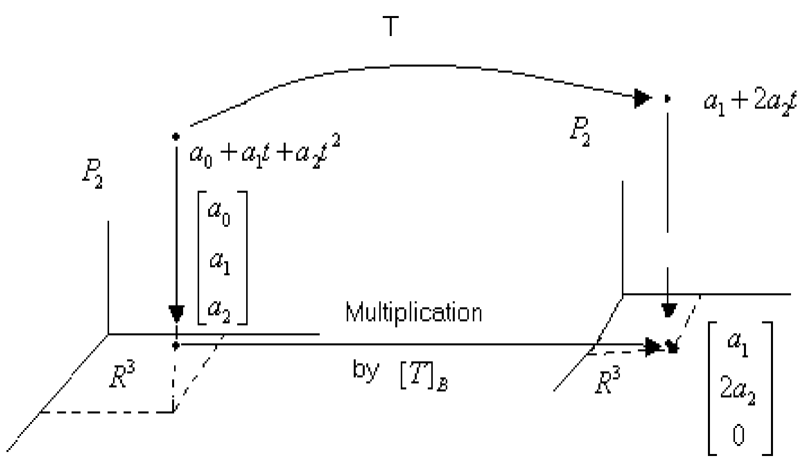


Figure 4 Matrix representation of a linear transformation

Linear Transformations on R^n

In an applied problem involving R^n, a linear transformation T usually appears first as a matrix transformation, $x \to Ax$. If A is diagonalizable then there is a basis B for R^n consisting of eigenvectors of A. Theorem below shows that in this case, the B-matrix of T is diagonal. Diagonalizing A amounts to finding a diagonal matrix representation of $x \to Ax$.

Theorem: Diagonal Matrix Representation

Suppose $\boldsymbol{A = PDP^{-1}}$, where $\boldsymbol{D}$ is diagonal n x n matrix. If $\boldsymbol{B}$ is the basis for $\boldsymbol{R^n}$ formed from the columns of $\boldsymbol{P}$, then $\boldsymbol{D}$ is the $\boldsymbol{B}$-matrix of the transformation $\boldsymbol{x} \to \boldsymbol{Ax}$.

Proof Denote the columns of $\boldsymbol{P}$ by $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$, so that $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ and $\boldsymbol{P} = [\boldsymbol{b}_1 \ \ \ldots \ \ \boldsymbol{b}_n]$. In this case, $\boldsymbol{P}$ is the change-of-coordinates matrix $\boldsymbol{P}_B$ discussed in lecture 23, where $\boldsymbol{P}[\boldsymbol{x}]_B = \boldsymbol{x}$ and $[\boldsymbol{x}]_B = \boldsymbol{P}^{-1}\boldsymbol{x}$

If $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{Ax}$ for $\boldsymbol{x}$ in $\boldsymbol{R^n}$, then

$$
\begin{aligned}
[\boldsymbol{T}]_B &= \left[[\boldsymbol{T}(\boldsymbol{b}_1)]_B \quad \cdots \quad [\boldsymbol{T}(\boldsymbol{b}_n)]_B\right] && \text{Definition of } [\boldsymbol{T}]_B \\
&= \left[[\boldsymbol{Ab}_1]_B \quad \cdots \quad [\boldsymbol{Ab}_n]_B\right] && \text{Since } \boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{Ax} \\
&= \left[\boldsymbol{P}^{-1}\boldsymbol{Ab}_1 \quad \cdots \quad \boldsymbol{P}^{-1}\boldsymbol{Ab}_n\right] && \text{Change of Coordinates} \\
&= \boldsymbol{P}^{-1}\boldsymbol{A}\left[\boldsymbol{b}_1 \quad \cdots \quad \boldsymbol{b}_n\right] && \text{Matrix multiplication} \\
&= \boldsymbol{P}^{-1}\boldsymbol{AP} && (6)
\end{aligned}
$$

Since $\boldsymbol{A} = \boldsymbol{PDP^{-1}}$, we have $[\boldsymbol{T}]_B = \boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}$.

Example 3

Define $\boldsymbol{T}: \boldsymbol{R}^2 \to \boldsymbol{R}^2$ by $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{Ax}$, where $\boldsymbol{A} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a basis $\boldsymbol{B}$ for $\boldsymbol{R^2}$ with the property that the $\boldsymbol{B}$-matrix of $\boldsymbol{T}$ is a diagonal matrix.

Solution

Since the eigenvalues of matrix A are 5 and 3 and the eigenvectors corresponding to eigenvalue 5 is $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and the eigenvectors corresponding to eigenvalue 3 is $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$

Therefore $A = \boldsymbol{PDP^{-1}}$, where $\boldsymbol{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ and $\boldsymbol{D} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$

By the above theorem $\boldsymbol{D}$ matrix is the $\boldsymbol{B}$ matrix of T when $\boldsymbol{B} = \left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\}$.

Similarity of Matrix Representations

We know that A is similar to C if there is an invertible matrix P such that $A = PCP^{-1}$, therefore if A is similar to C , then C is the $\boldsymbol{B}$-matrix of the transformation $\boldsymbol{x} \to \boldsymbol{Ax}$ **when the basis** B **is formed from the columns of** P by the theorem above.(Since in the proof , the information that $\boldsymbol{D}$ is a diagonal matrix was not used).

The factorization $\boldsymbol{A} = \boldsymbol{PCP^{-1}}$ is shown in Fig. 5.

$$\begin{tikzcd}[column sep=huge, row sep=huge]
x \arrow[r, "\text{Multiplication}", "\text{by } A"'] \arrow[d, "\text{Multiplication by } P^{-1}"'] & Ax \arrow[d, "\text{Multiplication by P}"] \\
{[x]_B} \arrow[r, "\text{Multiplication}", "\text{by } C"'] & {[Ax]_B}
\end{tikzcd}$$

Figure 5 Similarity of two matrix representations: $\boldsymbol{A} = \boldsymbol{PCP}^{-1}$

Conversely, if $\boldsymbol{T} : \boldsymbol{R}^n \to \boldsymbol{R}^n$ is defined by $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{Ax}$, and if $\boldsymbol{B}$ is any basis for $\boldsymbol{R}^n$, then the $\boldsymbol{B}$-matrix of $\boldsymbol{T}$ is similar to $\boldsymbol{A}$. In fact, the calculations in (6) show that if $\boldsymbol{P}$ is the matrix whose columns come from the vectors in $\boldsymbol{B}$, then $[\boldsymbol{T}]_{\boldsymbol{B}} = \boldsymbol{P}^{-1}\boldsymbol{AP}$. Thus, the set of all matrices similar to a matrix $\boldsymbol{A}$ coincides with the set of all matrix representations of the transformation $\boldsymbol{x} \to \boldsymbol{Ax}$.

An efficient way to compute a $\boldsymbol{B}$-matrix $\boldsymbol{P}^{-1}\boldsymbol{AP}$ is to compute $\boldsymbol{AP}$ and then to row reduce the augmented matrix $[\boldsymbol{P} \quad \boldsymbol{AP}]$ to $[\boldsymbol{I} \quad \boldsymbol{P}^{-1}\boldsymbol{AP}]$. A separate computation of $\boldsymbol{P}^{-1}$ is unnecessary.

Example 4

Find $\boldsymbol{T}(\boldsymbol{a_0} + a_1\boldsymbol{t} + a_2\,\boldsymbol{t}^2)$, if $\boldsymbol{T}$ is the linear transformation from $\boldsymbol{P_2}$ to $\boldsymbol{P_2}$ whose matrix relative to $\boldsymbol{B} = \{\boldsymbol{1}, \boldsymbol{t}, \boldsymbol{t}^2\}$ is $[\boldsymbol{T}]_{\boldsymbol{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$

Solution

Let $\boldsymbol{p}(\mathrm{t}) = \boldsymbol{a_0} + a_1\boldsymbol{t} + a_2\,\boldsymbol{t}^2$ and compute

$$[\boldsymbol{T}(\boldsymbol{p})]_{\boldsymbol{B}} = [\boldsymbol{T}]_{\boldsymbol{B}}[\boldsymbol{p}]_{\boldsymbol{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}\begin{bmatrix} \boldsymbol{a}_0 \\ \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{bmatrix} = \begin{bmatrix} 3\boldsymbol{a}_0 + 4\boldsymbol{a}_1 \\ 5\boldsymbol{a}_1 - \boldsymbol{a}_2 \\ \boldsymbol{a}_0 - 2\boldsymbol{a}_1 + 7\boldsymbol{a}_2 \end{bmatrix}$$

So $\boldsymbol{T}(\boldsymbol{p}) = (3a_0+4a_1)\boldsymbol{1} + (5a_1-a_2)\boldsymbol{t} + (a_0-2a_1+7a_2)\boldsymbol{t}^2$.

Example 5

Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ be n x n matrices. Verify that
(a) $\boldsymbol{A}$ is similar to $\boldsymbol{A}$.
(b) If $\boldsymbol{A}$ is similar to $\boldsymbol{B}$ and $\boldsymbol{B}$ is similar to $\boldsymbol{C}$, then $\boldsymbol{A}$ is similar to $\boldsymbol{C}$.

Solution

(a) $\boldsymbol{A} = (\boldsymbol{I})^{-1}\boldsymbol{A}\boldsymbol{I}$, so $\boldsymbol{A}$ is similar to $\boldsymbol{A}$.
(b) By hypothesis, there exist invertible matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ with the property that $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ and $\boldsymbol{C} = \boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}$. Substituting, and using facts about the inverse of a product we have

$$\boldsymbol{C} = \boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q} = \boldsymbol{Q}^{-1}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})\boldsymbol{Q} = (\boldsymbol{P}\boldsymbol{Q})^{-1}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{Q})$$

This equation has the proper form to show that A is similar to C.

Lecture 32

Eigenvalues and Eigenvectors

Definition

A complex scalar λ satisfies

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

if and only if there is a nonzero vector x in C^n such that

$$Ax = \lambda x.$$

We call λ a (complex) eigenvalue and x a (complex) eigenvector corresponding to λ.

Example 1

If $\boldsymbol{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the linear transformation $\boldsymbol{x} \to \boldsymbol{Ax}$ on $\boldsymbol{R^2}$ rotates the plane counterclockwise through a quarter-turn.

Solution

The action of ***A*** is periodic since after four quarter-turns, a vector is back where it started. Obviously, no nonzero vector is mapped into a multiple of itself, so ***A*** has no eigenvectors in $\boldsymbol{R^2}$ and hence no real eigenvalues. In fact, the characteristic equation of ***A*** is $\lambda^2 + 1 = 0 \Rightarrow \lambda^2 = -1 \Rightarrow \lambda = \pm i$

The only roots are complex: $\lambda = i$ and $\lambda = -i$. However, if we permit ***A*** to act on $\boldsymbol{C^2}$ then

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} 1 \\ -i \end{bmatrix}$$
$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i\begin{bmatrix} 1 \\ i \end{bmatrix}$$

Thus *i* and *–i* are the eigenvalues, with $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$ as corresponding eigenvectors.

Example 2

Let $\boldsymbol{A} = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}$.Find the eigenvalues of A, and find a basis for each eigen-space.

Solution

The characteristic equation of ***A*** is

$$0 = \det\begin{bmatrix} 0.5-\lambda & -0.6 \\ 0.75 & 1.1-\lambda \end{bmatrix} = (0.5-\lambda)(1.1-\lambda) - (-0.6)(0.75)$$
$$= \lambda^2 - 1.6\lambda + 1$$

From the quadratic formula, $\lambda = \frac{1}{2}[1.6 \pm \sqrt{(-1.6)^2 - 4}] = 0.8 \pm 0.6i.$ For the eigenvalue $\lambda = 0.8 - 0.6i,$ we study

$$A - (0.8 - 0.6i) = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix} - \begin{bmatrix} 0.8 - 0.6i & 0 \\ 0 & 0.8 - 0.6i \end{bmatrix} = \begin{bmatrix} -0.3 + 0.6i & -0.6 \\ 0.75 & 0.3 + 0.6i \end{bmatrix} \quad (1)$$

Since 0.8 – 0.6i is an eigenvalue, we know that the system

$$\begin{aligned} (-0.3 + 0.6i)x_1 - \quad 0.6x_2 &= 0 \\ 0.75x_1 + (0.3 + 0.6i)x_2 &= 0 \end{aligned} \quad (2)$$

has a nontrivial solution (with x_1 and x_2 possibly complex numbers).

$$0.75x_1 = (-0.3 - 0.6i)x_2$$

$$\Rightarrow \quad x_1 = (-0.4 - 0.8i)x_2$$

Taking x_2 = 5 to eliminate the decimals, we have x_1 = –2 – 4i. A basis for the eigenspace corresponding to λ = 0.8 – 0.6i is $\boldsymbol{v_1} = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$

Analogous calculations for λ =0 .8 + 0.6i produce the eigenvector $\boldsymbol{v_2} = \begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix}$

Check
Compute

$$\boldsymbol{Av_2} = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix} = \begin{bmatrix} -4 + 2i \\ 4 + 3i \end{bmatrix} = (0.8 + 0.6i)\boldsymbol{v_2}$$

Example 3

One way to see how multiplication by the $\boldsymbol{A}$ in Example 2 affects points is to plot an arbitrary initial point-say, $\mathbf{x_0}$ = (2, 0) and then to plot successive images of this point under repeated multiplications by $\boldsymbol{A}$. That is, plot

$$\boldsymbol{x_1} = \boldsymbol{Ax_0} = \begin{bmatrix} 0.5 & -0.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix}$$

$$\boldsymbol{x_2} = \boldsymbol{Ax_1} = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -0.4 \\ 2.4 \end{bmatrix}$$

$$\boldsymbol{x_3} = \boldsymbol{Ax_2}, \ldots$$

Figure 1 shows $\mathbf{x_0}, \ldots, \mathbf{x_8}$ as heavy dots. The smaller dots are the locations of $\boldsymbol{x_9}, \ldots, \boldsymbol{x_{100}}$. The sequence lies along an elliptical orbit.

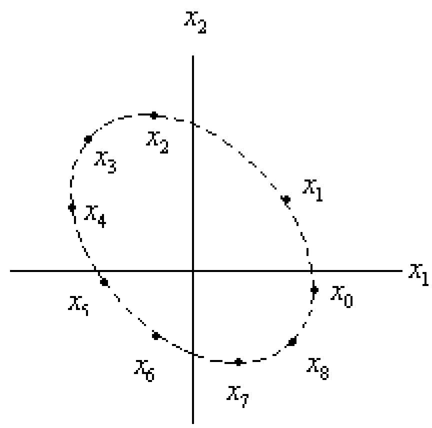


Figure 1 Iterates of a point x_0 under the action of a matrix with a complex eigenvalue

Of course Figure 1 does not explain why the rotation occurs. The secret to the rotation is hidden in the real and imaginary parts of a complex eigenvector.

Real and Imaginary Parts of Vectors

The complex conjugate of a complex vector x in C^n is the vector $\overline{\boldsymbol{x}}$ in C^n whose entries are the complex conjugates of the entries in x. The real and imaginary parts of a complex vector x are the vectors **Re** x and **Im** x formed from the real and imaginary parts of the entries of x.

Example 4

$$\text{If } \boldsymbol{x} = \begin{bmatrix} 3-i \\ i \\ 2+5i \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}, \text{ then}$$

$$\operatorname{Re} \boldsymbol{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \operatorname{Im} \boldsymbol{x} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}, \text{and } \quad \overline{\boldsymbol{x}} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} - i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+i \\ -i \\ 2-5i \end{bmatrix}$$

If $\boldsymbol{B}$ is an $m \times n$ matrix with possibly complex entries, then $\overline{\boldsymbol{B}}$ denotes the matrix whose entries are the complex conjugates of the entries in $\boldsymbol{B}$. Properties of conjugates for complex numbers carry over to complex matrix algebra:

$$\overline{r\boldsymbol{x}} = \overline{r}\,\overline{\boldsymbol{x}}, \quad \overline{B\boldsymbol{x}} = \overline{B}\,\overline{\boldsymbol{x}}, \text{and } \overline{rB} = \overline{r}\,\overline{B}$$

Eigenvalues and Eigenvectors of a Real Matrix that Acts on C^n

Let A be $n \times n$ matrix whose entries are real. Then $\overline{A\boldsymbol{x}} = \overline{A}\,\overline{\boldsymbol{x}} = A\overline{\boldsymbol{x}}$. If λ is an eigenvalue of A with x a corresponding eigenvector in C^n, then $A\overline{\boldsymbol{x}} = \overline{A\boldsymbol{x}} = \overline{\lambda \boldsymbol{x}} = \overline{\lambda}\,\overline{\boldsymbol{x}}$
Hence $\overline{\lambda}$ is also an eigenvalue of $\boldsymbol{A}$, with $\overline{\boldsymbol{x}}$ a corresponding eigenvector. This shows that when $\boldsymbol{A}$ is real, its complex eigenvalues occur in conjugate pairs. (Here and elsewhere, we use the term complex eigenvalue to refer to an eigenvalue $\lambda = a + bi$ with $b \neq 0$.)

Example 5

The eigenvalues of the real matrix in example 2 are complex conjugates namely, $0.8 - 0.6i$ and $0.8 + 0.6i$. The corresponding eigenvectors found in example 2 are also conjugates: $\boldsymbol{v_1} = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \overline{\boldsymbol{v_1}}$

Example 6

If $\boldsymbol{C} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where a and b are real and not both zero, then the eigenvalues of $\boldsymbol{C}$ are $\lambda = a \pm bi$. Also, if $r = |\lambda| = \sqrt{a^2 + b^2}$, then

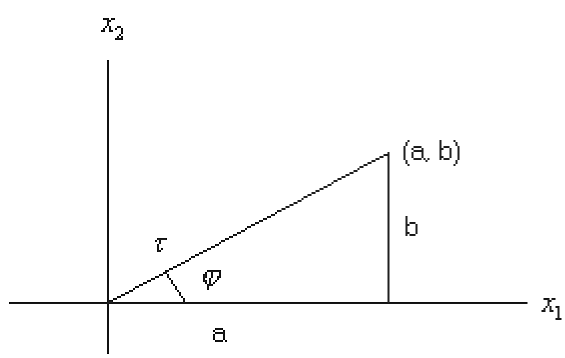


Figure 2

$$\boldsymbol{C} = r\begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

where φ is the angle between the positive x-axis and the ray from (0, 0) through (a, b). See Fig. 2. The angle φ is called the argument of $\lambda = a + bi$.

Thus, the transformation $\boldsymbol{x} \to \boldsymbol{Cx}$ may be viewed as the composition of a rotation through the angle φ and a scaling by $|\lambda|$.

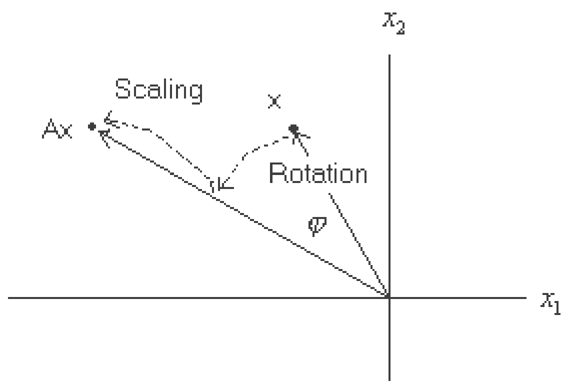


Figure 3 A rotation followed by a scaling

Finally, we are ready to uncover the rotation that is hidden within a real matrix having a complex eigenvalue.

Example 7

Let $\boldsymbol{A} = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}, \lambda = 0.8 - 0.6i,$ and $\boldsymbol{v_1} = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$, as in example 2. Also, let $\boldsymbol{P}$ be the 2×2 real matrix $\boldsymbol{P} = \begin{bmatrix} \text{Re}\,\boldsymbol{v_1} & \text{Im}\,\boldsymbol{v_1} \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$

and let $$\boldsymbol{C} = \boldsymbol{P}^{-1}\boldsymbol{AP} = \frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}\begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$$

By Example 6, $\boldsymbol{C}$ is a pure rotation because $|\lambda|^2 = (.8)^2 + (.6)^2 = 1.$

From $\boldsymbol{C} = \boldsymbol{P}^{-1}\boldsymbol{AP}$, we obtain

$$\boldsymbol{A} = \boldsymbol{PCP}^{-1} = \boldsymbol{P}\begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}\boldsymbol{P}^{-1}$$

Here is the rotation “inside” A! The matrix $\boldsymbol{P}$ provides a change of variable, say, $\boldsymbol{x} = \boldsymbol{Pu}$. The action of $\boldsymbol{A}$ amounts to a change of variable from $\boldsymbol{x}$ to $\boldsymbol{u}$, followed by a rotation, and then return to the original variable. See Figure 4.

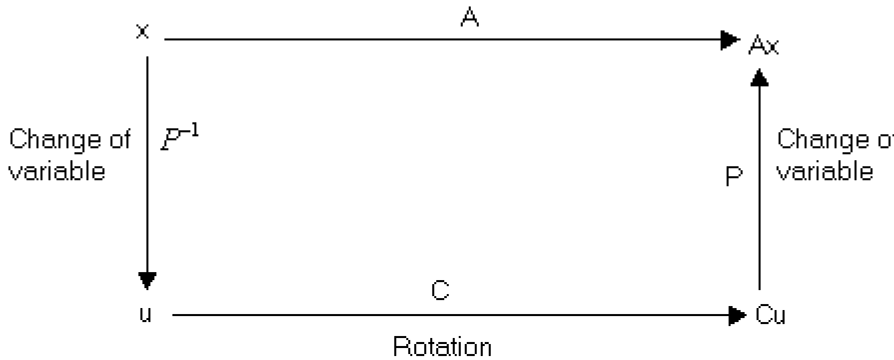


Figure 4 Rotation due to a complex eigenvalue

Theorem

Let A be a real 2 x 2 matrix with complex Eigen values $\lambda = a - b_i \quad (b \neq 0)$ and associated eigenvectors v in $\mathbb{C}^2$, then

$$A = PCP^{-1}, \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

$$P = \begin{bmatrix} \operatorname{Re} v & \operatorname{Im} v \end{bmatrix}$$

Lecture 33

Discrete Dynamical Systems

For a dynamical system described by the difference equation $x_{k+1} = Ax_k$ **,**eigenvalues and eigenvectors provide the key to understand the long term behavior and evolution of the system. Here the vector x_k gives information about the system as time (which is denoted by k) passes. Discrete dynamical system has a great application in many scientific fields. For example, modern state space design method of standard undergraduate courses in control system and steady state response of a control system relies heavily on dynamical system.

In this section, we will suppose that a matrix *A* is diagonalizable, with *n* linearly independent eigenvectors, $v_1, \ldots, v_n$, and corresponding eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$ **.**

For convenience, assume that the eigenvectors are arranged so that

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$$

Since { $v_1, \ldots, v_n$ } is a basis for R^n, any initial vector x_0 can be written uniquely as

$$x_0 = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$$

Since the *v*i are eigenvectors,

$$\begin{aligned} x_1 &= Ax_0 \\ &= c_1 Av_1 + c_2 Av_2 + \ldots + c_n Av_n \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \ldots + c_n \lambda_n v_n \end{aligned}$$

In general,

$$x_k = c_1 (\lambda_1)^k v_1 + \ldots + c_n (\lambda_n)^k v_n \qquad (k = 0,1,2\ldots)$$

We will discuss in the following examples what can happen to x_k if $k \to \infty$.

Example 1

Denote the owl and wood rat populations at time *k* by $\boldsymbol{x_k} = \begin{bmatrix} O_k \\ R_k \end{bmatrix}$, where *k* is the time in months, O_k is the number of owls in the region studied, and R_k is the number of rats (measured in thousands). Suppose that

$$\begin{aligned} O_{k+1} &= (.5)O_k + (.4)R_k \\ R_{k+1} &= -p \cdot O_k + (1.1)R_k \end{aligned}$$

where *p* is a positive parameter to be specified. The $(.5)O_k$ in the first equation says that with no wood rats for food only half of the owls will survive each month, while the $(1.1)R_k$ in the second equation says that with no owls as predators, the rat population will grow by 10% per month. If rats are plentiful, then $(.4)R_k$ will tend to make the owl population rise, while the negative term $-p.O_k$ measures the deaths of rats due to predation by owls. (In fact, 1000*p* is the average number of rats eaten by one owl in one month.) Determine the evolution of this system when the predation parameter *p* is 0.104.

Solution

The coefficient matrix ***A*** for the given equations is

$$A = \begin{bmatrix} 0.5 & 0.4 \\ -p & 1.1 \end{bmatrix}$$

When $p = .104$,

The matrix becomes

$$A = \begin{bmatrix} 0.5 & 0.4 \\ -0.104 & 1.1 \end{bmatrix}$$

For eigenvalues of the coefficient matrix put

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

The eigenvectors corresponding to the eigenvalues $\lambda_1 = 1.02$ and $\lambda_2 = .58$ are $\boldsymbol{v_1} = \begin{bmatrix} 10 \\ 13 \end{bmatrix}, \boldsymbol{v_2} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

An initial $\boldsymbol{x_0}$ can be written as $\boldsymbol{x_0} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2}$ then, for $k \geq 0$,

$$\boldsymbol{x_k} = c_1(1.02)^k \boldsymbol{v_1} + c_2(.58)^k \boldsymbol{v_2} = c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} + c_2(.58)^k \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

As $k \to \infty, (.58)^k$ rapidly approaches zero. Assume $c_1 \neq 0$. Then, for all sufficiently large *k*, $\boldsymbol{x_k}$ is approximately the same as $c_1(1.02)^k\boldsymbol{v_1}$, and we write

$$\boldsymbol{x_k} \approx c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \qquad \text{..................(1)}$$

The approximation in (1) improves as *k* increases, and so for large *k*,

$$\boldsymbol{x_{k+1}} \approx c_1(1.02)^{k+1} \begin{bmatrix} 10 \\ 13 \end{bmatrix} = (1.02)c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \approx 1.02\boldsymbol{x_k}$$

This approximation says that eventually both x_k entries i.e owls and rats grow by a factor of almost 1.02 each month. Also equation (1) shows that x_k is approximately a multiple of (10,13).So we can say that the entries in x_k are almost in the same ratio as 10 to 13.It means that for every 10 owls there are about 13 thousand rats.

Trajectory of a dynamical system

In a 2×2 matrix, geometric description of a system's evolution can enhance the algebraic calculations. We can view that what happens to an initial point x_0 in R^2, when it is transformed repeatedly by the mapping $x \to Ax$. The graph of x_0, x_1…is called a trajectory of the dynamical system.

Example 2

Plot several trajectories of the dynamical system $\boldsymbol{x}_{\mathbf{k+1}} = \boldsymbol{A}\boldsymbol{x}_{\boldsymbol{k}}$, when

$$\boldsymbol{A} = \begin{bmatrix} .80 & 0 \\ 0 & .64 \end{bmatrix}$$

Solution

In a diagonal matrix the eigen values are the diagonal entries and as A is a diagonal matrix with diagonal enteries 0.8 and 0.64,Therefore the eigenvalues of $\boldsymbol{A}$ are .8 and .64, with eigenvectors $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

If $\boldsymbol{x_0} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2}$, then $\boldsymbol{x_k} = c_1(.8)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.64)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Of course, $\boldsymbol{x}_{\mathbf{k}}$ tends to 0 because $(.8)^k$ and $(.64)^k$ both approach to 0 as $k \to \infty$. But the way $\boldsymbol{x}_{\mathbf{k}}$ goes towards 0 is interesting. Figure 1 shows the first few terms of several trajectories that begin at points on the boundary of the box with corners at $(\pm 3, \pm 3)$. The points on a trajectory are connected by a thin curve, to make the trajectory easier to see.

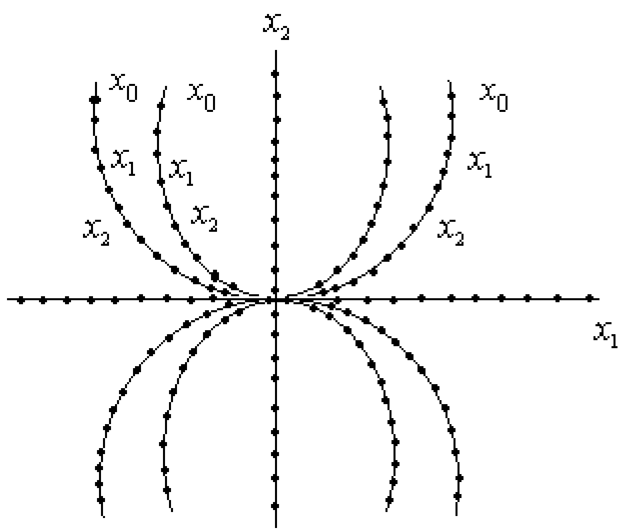


Figure 1 The origin as an attractor

In this example, the origin is called an **attractor** of the dynamical system because all trajectories tend towards **O**. This occurs whenever both eigenvalues are less than 1 in magnitude. The direction of greatest attraction is along the line through **O** and the eigenvector $\boldsymbol{v_2}$ as it has smaller eigenvalue.

Example 3

Plot several typical solution of the equation $x_{k+1} = Ax_k$, when

$$A = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.2 \end{bmatrix}$$

Solution

In a diagonal matrix the eigen values are the diagonal entries and as A is a diagonal matrix with diagonal entries 1.44 and 1.2,Therefore the eigenvalues of ***A*** are 1.44 and 1.2, with eigenvectors $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

As

$$\begin{aligned} x_0 &= c_1 v_1 + c_2 v_2 \\ &= c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} c_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c_2 \end{bmatrix} \\ &= \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \ \textit{then} \end{aligned}$$

$$x_k = c_1 (1.44)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 (1.2)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Both terms grow in size as k$\rightarrow\infty$,but the first term grows faster because it has larger eigenvalue. Figure 2 shows several trajectories that begin at points quite close to 0.

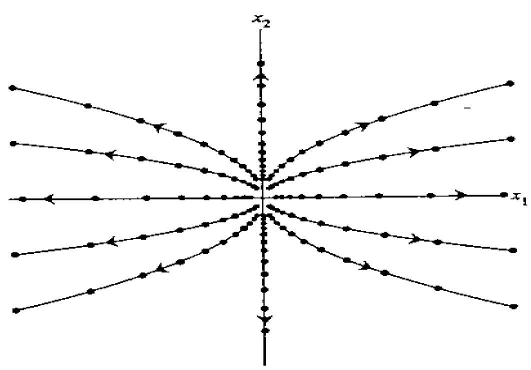


Figure 2 The origin as a repellor

In this example, the origin is called a **repellor** of the dynamical system because all trajectories tend away from **O**. This occurs whenever both eigenvalues are greater than 1 in magnitude. The direction of the greatest repulsion is the line through 0 and that eigen vector which has a larger eigenvalue.

Remark

From the above two examples it may be noticed that when eigenvalues are less than 1 in magnitude, the origin behaves as an attractor and when the eigenvalues are greater than 1 in magnitude, the origin behaves as a repellor.

Example 4

Plot several typical solutions of the equation

$y_{k+1} = Dy_k$ **,** where

$$D = \begin{bmatrix} 2.0 & 0 \\ 0 & 0.5 \end{bmatrix}$$

Show that a solution $\{y_k\}$ is unbounded if its initial point is not on the x_2-axis.

Solution

Mistakes

The eigenvalues of D are 2.0 and 0.5, with eigenvectors $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

If

$$y_0 = c_1\lambda_1 v_1 + c_2\lambda_2 v_2$$

$$y_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, then$$

$$y_k = c_1(2.0)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(0.5)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

If y_0 is on the x_2-axis,then c_1 =0 and $y_k \to 0$ as k$\to \infty$.But if y_0 is not on the x_2-axis, then the first term in the sum for y_k becomes arbitrarily large, and so $\{y_k\}$ is unbounded. Figure 3 shows ten trajectories that begin near or on the axis.

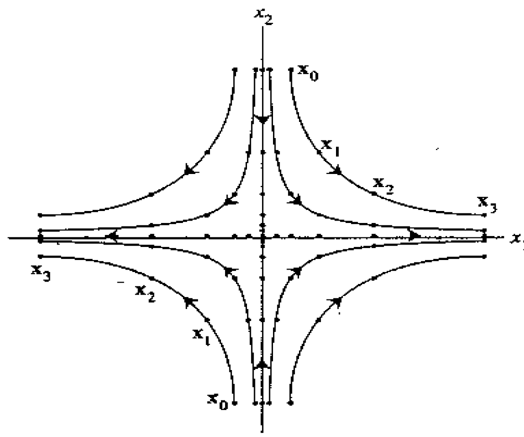


FIGURE 3 The origin as a saddle point.

In this example 'O' is called the saddle point because one eigenvalue is greater than '1' in magnitude and one is less than '1' in magnitude. The origin attracts solution from some directions and repels them in other directions.

Change of Variable

The preceding three examples involved diagonal matrices. To handle the nondiagonal case, we return for a moment to the $n \times n$ case in which eigenvectors of $\boldsymbol{A}$ form a basis $\{\boldsymbol{v_1},\ldots,\boldsymbol{v_n}\}$ for $\boldsymbol{R^n}$. Let $\boldsymbol{P} = [\boldsymbol{v_1} \ \cdots \ \boldsymbol{v_n}]$, and let $\boldsymbol{D}$ be the diagonal matrix with the corresponding eigenvalues on the diagonal. Given a sequence $\{\boldsymbol{x_k}\}$ satisfying $\boldsymbol{x_{k+1}} = \boldsymbol{Ax_k}$, define a new sequence $\{\boldsymbol{y_k}\}$ by $\boldsymbol{y_k} = \boldsymbol{P}^{-1}\boldsymbol{x_k}$, or equivalently, $\boldsymbol{x_k} = \boldsymbol{Py_k}$.

Substituting these relations into the equation $\boldsymbol{x_{k+1}} = \boldsymbol{Ax_k}$, and using the fact that $\boldsymbol{A} = \boldsymbol{PDP}^{-1}$, we find that $\boldsymbol{Py_{k+1}} = \boldsymbol{APy_k} = (\boldsymbol{PDP}^{-1})\boldsymbol{Py_k} = \boldsymbol{PDy_k}$

Left-multiplying both sides by $\boldsymbol{P}^{-1}$, we obtain $\boldsymbol{y_{k+1}} = \boldsymbol{Dy_k}$

If we write $\boldsymbol{y_k}$ as $\boldsymbol{y}(k)$ and denote the entries in $\boldsymbol{y}(k)$ by $\boldsymbol{y_1}(k), \ldots, \boldsymbol{y_n}(k)$, then

$$\begin{bmatrix} \boldsymbol{y_1}(k+1) \\ \boldsymbol{y_2}(k+1) \\ \vdots \\ \boldsymbol{y_n}(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{y_1}(k) \\ \boldsymbol{y_2}(k) \\ \vdots \\ \boldsymbol{y_n}(k) \end{bmatrix}$$

The change of variable from $\boldsymbol{x_k}$ to $\boldsymbol{y_k}$ has decoupled the system of difference equations. The evolution of $\boldsymbol{y_1}(k)$, for example, is unaffected by what happens to $\boldsymbol{y_2}(k), \ldots, \boldsymbol{y_n}(k)$, because $\boldsymbol{y_1}(k+1) = \lambda_1 . \boldsymbol{y_1}(k)$ for each k.

Example 5

Show that the origin is a saddle point for solutions of $\boldsymbol{x}_{\mathbf{k+1}} = \boldsymbol{A}\boldsymbol{x}_{\mathbf{k}}$, where

$$\boldsymbol{A} = \begin{bmatrix} 1.25 & -.75 \\ -.75 & 1.25 \end{bmatrix}$$

find the directions of greatest attraction and greatest repulsion.

Solution

For eigenvalues of the given matrix put

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

we find that $\boldsymbol{A}$ has eigenvalues 2 and .5, with corresponding eigenvectors $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively. Since $|2| > 1$ and $|.5| < 1$, the origin is a saddle point of the dynamical system. If $\boldsymbol{x_0} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2}$, then

$$\boldsymbol{x_k} = c_1(2)^k \boldsymbol{v_1} + c_2(.5)^k \boldsymbol{v_2}$$

In figure 4, the direction of the greatest repulsion is the line through v_1, the eigenvector whose eigenvalue is greater than 1.

The direction of the greatest attraction is determined by the eigenvector v_2 whose eigenvalue is less than 1.

A number of trajectories are shown in figure 4.When this graph is viewed in terms of eigenvector axes, the diagram looks same as figure 4.

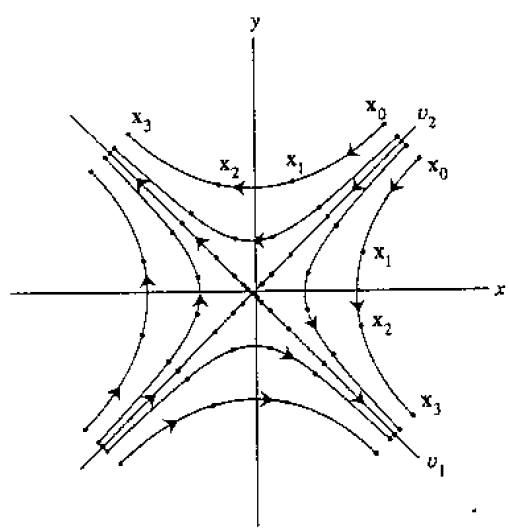


FIGURE 4 The origin as a saddle point.

Complex Eigenvalues

Since the characteristic equation of n× n matrix involves a polynomial of degree n, the equation always has exactly n roots, in which complex roots are possibly included. If we have complex eigenvalues then we get complex eigen vectors which may be divided into two parts: which is real part of the vector and imaginary part of the vector.

Note

If a matrix has two complex eigenvalues in which absolute value of one eigenvalue is greater than 1,then origin behaves as a repellor and iterates of x_0 will spiral outwards around the origin .If the absolute value of an eigenvalue is less than 1,then the origin behaves as an attractor and iterates of x_0 will spiral inward towards the origin.

Example 6

Let

$$A = \begin{bmatrix} 0.8 & 0.5 \\ -0.1 & 1.0 \end{bmatrix} \text{ has eigenvalue } 0.9 \pm 0.2i \text{ with eigen vectors} \begin{bmatrix} 1 \pm 2i \\ 1 \end{bmatrix}.$$

$$\textit{Find the trajectories of the system } x_{k+1} = Ax_k \textit{ with initial vectors} \begin{bmatrix} 0 \\ 2.5 \end{bmatrix},$$

$$\begin{bmatrix} 3 \\ 0 \end{bmatrix} \textit{ and } \begin{bmatrix} 0 \\ -2.5 \end{bmatrix}.$$

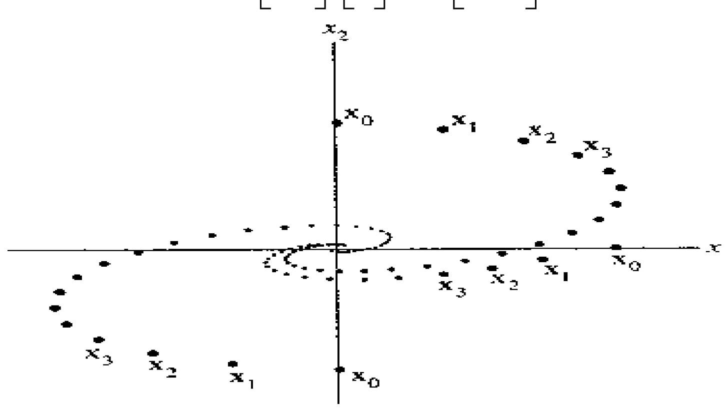
Solution

Given a matrix

$$A = \begin{bmatrix} 0.8 & 0.5 \\ -0.1 & 1.0 \end{bmatrix}$$

The eigenvalues of the matrix are $0.9 \pm 0.2i$ and corresponding eigenvectors are $\begin{bmatrix} 1 \pm 2i \\ 1 \end{bmatrix}$

also the initial vectors are given as $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix} \textit{ and } \begin{bmatrix} 0 \\ -2.5 \end{bmatrix}.$



Example 7

Suppose the search survival rate of a young bird is 50%,and the stage matrix A is

$$A = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.3 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{bmatrix}$$

What does the stage matrix model predict about this bird?

Solution

Given a matrix A= $\begin{bmatrix} 0 & 0 & 0.33 \\ 0.3 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{bmatrix}$

for eigenvalues of A we will put det(A- λ I)=0

$$\det(A-\lambda I) = \begin{vmatrix} -\lambda & 0 & 0.33 \\ 0.3 & -\lambda & 0 \\ 0 & 0.71 & 0.94-\lambda \end{vmatrix}$$

$$= -\lambda(-0.94+\lambda^2)+0.33(0.3\times 0.71)$$

$$= 0.94\lambda-\lambda^3+0.07029$$

$$\textit{for eigenvalues put } \det(A-\lambda I)=0$$

$$0.94\lambda-\lambda^3+0.07029=0$$

$$\lambda^3-0.94\lambda-0.07029=0$$

it gives 3 values of λ which are approximately 1.01,-0.03+0.26i,-0.03-0.26i
eigen vector for first value of λ will be $v_1 = (10,3,31)$ and for the next two values of λ we will get complex eigen vectors denoted by v_2 and v_3.

Now

$$x_k = c_1(1.01)^k v_1 + c_2(-0.03+0.26i)^k v_2 + c_3(-0.03-0.26i)^k v_3$$

As k$\rightarrow\infty$,the second two vectors tends to 0.So x_k becomes more and more like the (real) vector $c_1(1.01)^k v_1$ .Thus the long term growth rate of the owl population will be 1.01,and the population will grow slowly. The eigenvector describes the eventual distribution of the owls by life stages: For every 31 adults, there will be about 10 juveniles and three subadults.

Exercise

Suppose that a 3×3 matrix A has eigenvalues 2,2/5,3/5 and corresponding eigenvectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} v_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} v_3 = \begin{bmatrix} -3 \\ 3 \\ 7 \end{bmatrix}$$

$$also \;\; x_0 = \begin{bmatrix} -2 \\ -5 \\ -3 \end{bmatrix}$$

Find a general solution of the equation $x_{k+1} = Ax_k$ also find what happens to x_k when $k \to \infty$.

Lecture 34

Applications to Differential Equations

Differential equation

A differential equation is any equation which contains derivatives, either ordinary derivatives or partial derivatives. Differential equations play an extremely important and useful role in applied mathematics, engineering, and physics. Moreover, a lot of mathematical and numerical machinery has been developed for the solution of differential equations.

System of linear differential equations

A system of linear differential equations can be expressed as:

$$\begin{aligned}
x_1' &= a_{11}x_1 + \ldots + a_{1n}x_n \\
x_2' &= a_{21}x_1 + \ldots + a_{2n}x_n \\
&. \\
&. \\
&. \\
x_n' &= a_{n1}x_1 + \ldots + a_{nn}x_n
\end{aligned}$$

where $x_i(t)$ is a function of time, i = 1, . . . n, and the matrix of constant coefficient is A = $\left[a_{ij}\right]$

In many applied problems, several quantities are continuously varying in time, and they are related by a system of differential equations:

$$\begin{aligned}
x_1' &= a_{11}x_1 + \ldots + a_{1n}x_n \\
x_2' &= a_{21}x_1 + \ldots + a_{2n}x_n \\
&. \\
&. \\
&. \\
x_n' &= a_{n1}x_1 + \ldots + a_{nn}x_n
\end{aligned}$$

Here $x_1, x_2, ..., x_n$ are differentiable functions of t ,with derivatives $x_1', x_2', ..., x_n'$,and the a_{ij} are constants. Write the system as a matrix differential equation

$x^{'} = Ax$

Where

$$x(t)=\begin{bmatrix} x_1(t) \\ . \\ . \\ . \\ x_n(t) \end{bmatrix}, x^{'}(t)=\begin{bmatrix} x_1^{'}(t) \\ . \\ . \\ . \\ x_n^{'}(t) \end{bmatrix}, and\ A=\begin{bmatrix} a_{11} \ . \ . \ . \ a_{1n} \\ . \qquad . \\ . \qquad . \\ . \qquad . \\ a_{n1} \qquad a_{nn} \end{bmatrix}$$

Superposition of solutions

If u and v are two solutions of $x^{'} = Ax$, then $cu + dv$ is also a solution.

Proof:

If u and v are two solutions of $x^{'} = Ax$ then $u' = Au$ and $v' = Av$ .Therefore,

$$\begin{aligned}(cu+dv)^{'} &= cu^{'} + dv^{'} \\ &= cAu + dAv \\ &= A(cu+dv)\end{aligned}$$

Linear combination of any set of solutions is also a solution. This property is called **superposition** of solutions.

Fundamental set of Solutions

If A is an n× n matrix, then there are n linearly independent functions in a fundamental set, and each solution of $x^{'} = Ax$ is a unique linear combination of these ‘n’ functions. So we can say that fundamental set of solutions is a basis for the set of all solutions of $x^{'} = Ax$,and the solution set is n-dimensional vector space of functions.

Initial value problem

If a vector x_0 is specified, then the initial value problem is to construct the unique function x such that $x^{'} = Ax \ and \ x(0) = x_0$

Example 1

For the system $x' = Ax$ .What will be the solution when A is a diagonal matrix and is given by

$$A = \begin{pmatrix} 3 & 0 \\ 0 & -5 \end{pmatrix}$$

Solution

The solution of the given system can be obtained by elementary calculus. Consider

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

from here we have

$x_1'(t) = 3x_1(t)$ *and* $x_2'(t) = -5x_2(t)$

first of all we will find solution of $x_1'(t) = 3x_1(t)$ *by using calculus.*

$$x_1'(t) = 3x_1(t)$$

$$\frac{dx_1}{dt} = 3x_1$$

$$\frac{dx_1}{x_1} = 3dt$$

Integrating both sides

$$\int \frac{dx_1}{x_1} = 3\int dt$$

$$\ln x_1 = 3t + \ln c_1$$

$$\ln x_1 - \ln c_1 = 3t$$

$$\ln\left(\frac{x_1}{c_1}\right) = 3t$$

taking antilog on both sides

$$\frac{x_1}{c_1} = e^{3t}$$

$$x_1(t) = c_1 e^{3t}$$

Similarly Solution of

$$x_2'(t) = -5x_2(t)$$

will be

$$x_2(t) = c_2 e^{-5t}$$

And it can be written in the form

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c_1 e^{3t} \\ c_2 e^{-5t} \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-5t}$$

The given system is said to be decoupled because each derivative of the function depends only on the function itself, not on some combination or coupling of $x_1(t)$ *and* $x_2(t)$.

Observation

From the above example we observe that a solution of the general equation $x' = Ax$ might be a linear combination of functions of the form

$$x(t) = ve^{\lambda t}$$

Where λ is some scalar and v is some fixed non-zero vector.

$$x(t) = ve^{\lambda t} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

differentiating w.r.t 't'

$$x'(t) = \lambda ve^{\lambda t}$$

Multiplying both sides of (1) *by A*

$$Ax(t) = Ave^{\lambda t}$$

Since $e^{\lambda t}$ *is never zero,* $x'(t)$ *will equal to* $Ax(t)$ *iff* $\lambda v = Av.$

Eigenfunctions

Each pair of eigenvalue and its corresponding eigenvector provides a solution of the equation $x' = Ax$ which is called **eigenfunctions** of the differential equation.

Example 2

The circuit in figure 1 can be described by

$$\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \begin{bmatrix} -(R_1 + R_2)/C_1 & R_2/C_1 \\ R_2/C_2 & -R_2/C_2 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

Where $v_1(t)$ and $v_2(t)$ are the voltages across the two capacitors at time *t*.

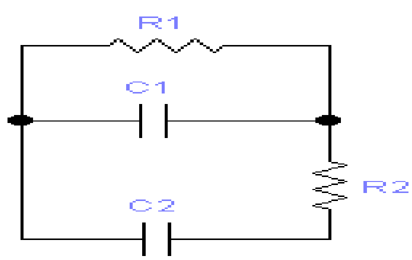


Figure 1

Suppose that resistor R_1 is 2 ohms, R_2 is 1 ohms, capacitor C_1 is 2 farads, and capacitor C_2 is 1 farad and suppose that there is an initial charge of 5 volts on capacitor C_1 and 4 volts on capacitor C_2. Find formulas for $v_1(t)$ and $v_2(t)$ that describe how the voltages change over time.

Solution

For the data given, we can form A as

$$A = \begin{bmatrix} -(R_1+R_2)/C_1 & R_2/C_1 \\ R_2/C_2 & -R_2/C_2 \end{bmatrix}$$

$R_1 = 2$

$R_2 = 1$

$C_1 = 2$

$C_2 = 1$ *so*

$-(R_1+R_2)/C_1 = {}^{-(2+1)}\!\big/_2 \Rightarrow {}^{-3}\!\big/_2 \Rightarrow -1.5$

$R_2/C_1 = {}^{1}\!\big/_2 \Rightarrow 0.5$

$R_2/C_2 = {}^{1}\!\big/_1 \Rightarrow 1$

$-R_2/C_2 = -{}^{1}\!\big/_1 \Rightarrow -1$

$\boldsymbol{A} = \begin{bmatrix} -1.5 & .5 \\ 1 & -1 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$ and $\boldsymbol{x_0} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$. The vector $\boldsymbol{x_0}$ lists the initial values of $\boldsymbol{x}$. From $\boldsymbol{A}$, we obtain eigenvalues $\lambda_1 = -.5$ and $\lambda_2 = -2$, by using det (A-λI)=0 with corresponding eigenvectors $\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\boldsymbol{v_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

The eigenfunctions $\boldsymbol{x_1}(t) = \boldsymbol{v_1} e^{\lambda_1 t}$ and $\boldsymbol{x_2}(t) = \boldsymbol{v_2} e^{\lambda_2 t}$ both satisfy $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$, and so does any linear combination of $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$. Set

$$\begin{aligned} \boldsymbol{x}(t) &= c_1 \boldsymbol{v_1} e^{\lambda_1 t} + c_2 \boldsymbol{v_2} e^{\lambda_2 t} \\ &= c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t} \text{............(1)} \end{aligned}$$

and note that $\boldsymbol{x}(0) = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2}$. Since $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are obviously linearly independent and hence span $\boldsymbol{R^2}$, c_1 and c_2 can be found to make $\boldsymbol{x}(0)$ equal to $\boldsymbol{x_0}$.

We can find c_1 *and* c_2 as

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} v_1 & v_2 & x_0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix} \textit{row reduction of this matrix gives} \\ &= \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \end{bmatrix} \end{aligned}$$

$$\textit{hence } c_1 = 3 \textit{ and } c_2 = -2$$

Put the value of c_1 *and* c_2 in (1)

$$\boldsymbol{x}(t) = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} - 2\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

Thus the desired solution of the differential equation $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$ is

$$x(t)=\begin{bmatrix}3e^{-0.5t}\\6e^{-0.5t}\end{bmatrix}-\begin{bmatrix}-2e^{-2t}\\2e^{-2t}\end{bmatrix}$$

or $$\begin{bmatrix}v_1(t)\\v_2(t)\end{bmatrix}=\begin{bmatrix}3e^{-0.5t}+2e^{-2t}\\6e^{-0.5t}-2e^{-2t}\end{bmatrix}$$

As both eigenvalues are negative so origin in this case behaves as an attractor and the direction of greatest attraction will be along the more negative eigenvalue $\lambda=-2$.

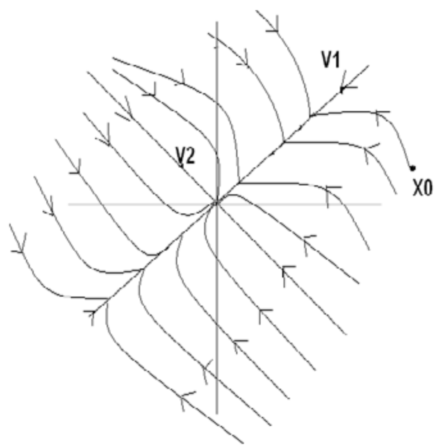


Figure 2 The Origin as an attractor

Example 3

Suppose a particle is moving in a planar force field and its position vector x satisfies $x' = Ax$ and $x(0) = x_0$, where

$$\boldsymbol{A}=\begin{bmatrix}4 & -5\\-2 & 1\end{bmatrix},\ \boldsymbol{x_0}=\begin{bmatrix}2.9\\2.6\end{bmatrix}$$

Solve this initial value problem for $t \geq 0$, and sketch the trajectory of the particle.

Solution

The eigenvalues of the given matrix can be found by using det(A-λI)=0 which are turned out to be λ_1 = 6 and λ_2 =-1 with corresponding eigen vectors $v_1=\begin{bmatrix}-5\\2\end{bmatrix}$ *and* $v_2=\begin{bmatrix}1\\1\end{bmatrix}$.

For any constants c_1 *and* c_2 ,the function

$$x(t)=c_1v_1e^{\lambda_1 t}+c_2v_2e^{\lambda_2 t}$$
$$=c_1\begin{bmatrix}-5\\2\end{bmatrix}e^{6t}+c_2\begin{bmatrix}1\\1\end{bmatrix}e^{-t}$$

We have to find c_1 *and* c_2 to satisfy x(0)= x_0 ,so

$$x_0 = c_1\begin{bmatrix}-5\\2\end{bmatrix} + c_2\begin{bmatrix}1\\1\end{bmatrix}$$
$$\begin{bmatrix}2.9\\2.6\end{bmatrix} = \begin{bmatrix}-5c_1\\2c_1\end{bmatrix} + \begin{bmatrix}c_2\\c_2\end{bmatrix}$$
$$= \begin{bmatrix}-5c_1 + c_2\\2c_1 + c_2\end{bmatrix}$$
$$2.9 = -5c_1 + c_2 ....(1)$$
$$2.6 = 2c_1 + c_2 ....(2)$$
Subtracting 1 *from* 2 *we get*

$$c_1 = -\frac{3}{70} \text{ and substituting this value of } c_1 \text{ in}(1) \text{ we get } c_2 = \frac{188}{70}$$
$$x(t) = -\frac{3}{70}\begin{bmatrix}-5\\2\end{bmatrix}e^{6t} + \frac{188}{70}\begin{bmatrix}1\\1\end{bmatrix}e^{-t}$$

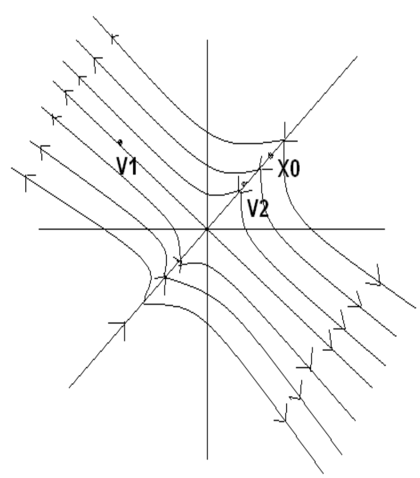


Figure 3 The Origin as a saddle point

As the matrix A has both positive and negative eigenvalues so in this case origin behaves as a saddle point. The direction of greatest repulsion is the line through v_1 *and* 0 corresponding to the positive eigenvalue. The direction of greatest attraction is the line through v_2 *and* 0 , corresponding to the negative eigenvalue.

Decoupling a Dynamical System

For any dynamical system described by $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$ when $\boldsymbol{A}$ is diagonalizable, the fundamental set of solutions can be obtained by the methods that we have discussed in examples 2 and 3.

Suppose that the eigenfunctions for $\boldsymbol{A}$ are

$$\boldsymbol{v_1}e^{\lambda_1 t}, \cdots, \boldsymbol{v_n}e^{\lambda_n t}$$

with $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ linearly independent eigenvectors. Let $\boldsymbol{P} = [\boldsymbol{v}_1 \ \ldots \ \boldsymbol{v}_n]$ and let $\boldsymbol{D}$ be the diagonal matrix with entries $\lambda_1, \ldots, \lambda_n$, so that $\boldsymbol{A} = \boldsymbol{PDP}^{-1}$. Now make a change of variable, defining a new function $\boldsymbol{y}$ by

$$\boldsymbol{y}(t) = \boldsymbol{P}^{-1}\boldsymbol{x}(t), \text{ or equivalently, } \boldsymbol{x}(t) = \boldsymbol{Py}(t)$$

Substituting these relations into $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$ gives

$$\begin{aligned}
&Py'(t) = (PDP^{-1})Py(t) \\
&Py'(t) = PDy(t) \\
&\text{pre multiplying both side by } P^{-1} \\
&P^{-1}Py'(t) = P^{-1}PDy(t) \\
&y'(t) = Dy(t) \ \ or \\
&y' = Dy
\end{aligned}$$

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

The change of variable from $\boldsymbol{x}$ to $\boldsymbol{y}$ has decoupled the system of differential equations because the derivative of each scalar function $\boldsymbol{y_k}$ depends only on $\boldsymbol{y}_k$. Since $\boldsymbol{y_1}' = \lambda_1 \boldsymbol{y_1}$, we have $\boldsymbol{y_1}(t) = c_1 e^{\lambda_1 t}$, with similar formulas for $\boldsymbol{y}_2, \ldots, \boldsymbol{y}_n$. Thus

$$\boldsymbol{y}(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}, \quad \text{where} \quad \begin{bmatrix} c_1 \\ \vdots \\ c_2 \end{bmatrix} = \boldsymbol{y}(0) = \boldsymbol{P}^{-1}\boldsymbol{x}(0) = \boldsymbol{P}^{-1}\boldsymbol{x_0}$$

To obtain the general solution $\boldsymbol{x}$ of the original system, compute

$$\boldsymbol{x}(t) = \boldsymbol{Py}(t) = [\boldsymbol{v_1} \quad \cdots \quad \boldsymbol{v_n}]\,\boldsymbol{y}(t) = c_1\boldsymbol{v_1}e^{\lambda_1 t} + \cdots + c_n\boldsymbol{v_n}e^{\lambda_n t}$$

This is the eigenfunction expansion as constructed in example 2.

Complex Eigenvalues

As we know that for a real matrix A , complex eigenvalues and associated eigen vectors come in conjugate pairs i.e if λ is an eigenvalue of A then $\overline{\lambda}$ will be the 2nd eigenvalue. Similarly if one eigen vector is v then the other will be $\overline{v}$. So there will be two solutions of $x' = Ax$ and they are

$$x_1(t) = ve^{\lambda t}, x_2(t) = \overline{v}e^{\overline{\lambda} t}$$

Here $x_2(t) = \overline{x_1(t)}$

$$\text{Re}(ve^{\lambda t}) = \frac{1}{2}\left[x_1(t) + \overline{x_1(t)}\right]$$

$$\text{Im}(ve^{\lambda t}) = \frac{1}{2i}\left[x_1(t) - \overline{x_1(t)}\right]$$

From Calculus

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots$$

$$e^{\lambda t} = 1 + \lambda t + \frac{(\lambda t)^2}{2!} + \ldots + \frac{(\lambda t)^n}{n!} + \ldots$$

If we write $\lambda = a + ib$

$$e^{(a+ib)t} = e^{at}.e^{ibt} = e^{at}(\cos bt + i\sin bt)$$

So

$$ve^{\lambda t} = (\text{Re}\, v + i\,\text{Im}\, v).e^{at}(\cos bt + i\sin bt)$$

$$ve^{\lambda t} = \text{Re}\, v.e^{at}\cos bt + \text{Re}\, v.e^{at} i\sin bt + i\,\text{Im}\, v.e^{at}\cos bt - \text{Im}\, v.e^{at}\sin bt$$

$$ve^{\lambda t} = e^{at}(\text{Re}\, v\cos bt - \text{Im}\, v\sin bt) + ie^{at}(\text{Re}\, v\sin bt + \text{Im}\, v\cos bt)$$

So two real solutions of $x' = Ax$ *are*

$$y_1(t) = \text{Re}\, x_1(t) = \left[(\text{Re}\, v)\cos bt - (\text{Im}\, v)\sin bt\right]e^{at}$$

$$y_2(t) = \text{Im}\, x_1(t) = \left[(\text{Re}\, v)\sin bt + (\text{Im}\, v)\cos bt\right]e^{at}$$

Example 4

The circuit in figure 4 can be described by the equation

$$\begin{bmatrix} \boldsymbol{i}_L' \\ \boldsymbol{v}_C' \end{bmatrix} = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1C) \end{bmatrix}\begin{bmatrix} \boldsymbol{i}_L \\ \boldsymbol{v}_C \end{bmatrix}$$

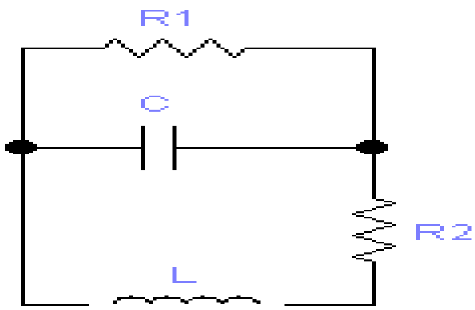


Figure 4

Where i_L is the current passing through the inductor L and v_c is the voltage drop across the capacitor C. Suppose R_1 is 5 ohms, R_2 is .8 ohm, C is .1 farad, and L is .4 henry. Find formulas for i_L and v_c, if the initial current through the inductor is 3 amperes and the initial voltage across the capacitor is 3 volts.

Solution

From the data given, We can form A as

$$A = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1C) \end{bmatrix}$$

$$R_1 = 5$$
$$R_2 = 0.8$$
$$C = 0.1$$
$$L = 0.4$$

$$-R_2/L = -0.8/0.4 \Rightarrow -8/10 \times 10/4 \Rightarrow -2$$
$$-1/L = -1/0.4 \Rightarrow -10/4 \Rightarrow -2.5$$
$$1/C = 1/0.1 \Rightarrow 10$$
$$-1/(R_1C) = -1/5\times 0.1 \Rightarrow -1/0.5 \Rightarrow -2$$
$$\boldsymbol{A} = \begin{bmatrix} -2 & -2.5 \\ 10 & -2 \end{bmatrix}$$

$$\det(A-\lambda I)=\begin{bmatrix}-2-\lambda & -2.5\\ 10 & -2-\lambda\end{bmatrix}$$
$$=(-2-\lambda)^2+25$$
$$=4+\lambda^2+4\lambda+25$$
$$=\lambda^2+4\lambda+29$$

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

$\lambda^2+4\lambda+29=0$

the eigenvalues of A is $-2\pm 5i$ with corresponding eigenvectors $\boldsymbol{v_1}=\begin{bmatrix}i\\2\end{bmatrix}$ *and* $\begin{bmatrix}-i\\2\end{bmatrix}$ The complex solutions of
$\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x}$ are complex linear combinations of

$$\boldsymbol{x_1}(t)=\begin{bmatrix}i\\2\end{bmatrix}e^{(-2+5i)t} \quad \text{and} \quad \boldsymbol{x_2}(t)=\begin{bmatrix}-i\\2\end{bmatrix}e^{(-2-5i)t}$$
$$=\begin{bmatrix}i\\2\end{bmatrix}e^{-2t}.e^{5it}$$
$$\boldsymbol{x_1}(t)=\begin{bmatrix}i\\2\end{bmatrix}e^{-2t}(\cos 5t+i\sin 5t)$$

By taking real and imaginary parts of $\boldsymbol{x_1}$, we obtain real solutions:

$$\boldsymbol{y_1}(t)=\begin{bmatrix}-\sin 5t\\2\cos 5t\end{bmatrix}e^{-2t},\ \boldsymbol{y_2}(t)=\begin{bmatrix}\cos 5t\\2\sin 5t\end{bmatrix}e^{-2t}$$

Since $\boldsymbol{y_1}$ and $\boldsymbol{y_2}$ are linearly independent functions, they form a basis for the two-dimensional real vector space of solutions $\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x}$. Thus, the general solution is

$$\boldsymbol{x}(t)=c_1\begin{bmatrix}-\sin 5t\\2\cos 5t\end{bmatrix}e^{-2t}+c_2\begin{bmatrix}\cos 5t\\2\sin 5t\end{bmatrix}e^{-2t}$$

To satisfy $\boldsymbol{x}(0)=x_0=\begin{bmatrix}3\\3\end{bmatrix}$ we need sc_1 and c_2

$$c_1\begin{bmatrix}0\\2\end{bmatrix}+c_2\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}3\\3\end{bmatrix},$$
$$\begin{bmatrix}0\\2c_1\end{bmatrix}+\begin{bmatrix}c_2\\0\end{bmatrix}=\begin{bmatrix}3\\3\end{bmatrix}$$
$$0+c_2=3$$

Thus $\quad 2c_1+0=3$

$$c_2=3$$
$$c_1=\frac{3}{2}=1.5$$

$$\boldsymbol{x}(t) = c_1\begin{bmatrix} -\sin 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + c_2\begin{bmatrix} \cos 5t \\ 2\sin 5t \end{bmatrix} e^{-2t}$$

$$= 1.5\begin{bmatrix} -\sin 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + 3\begin{bmatrix} \cos 5t \\ 2\sin 5t \end{bmatrix} e^{-2t}$$

$$= \begin{bmatrix} -1.5\sin 5te^{-2t} \\ 3\cos 5te^{-2t} \end{bmatrix} + \begin{bmatrix} 3\cos 5te^{-2t} \\ 6\sin 5te^{-2t} \end{bmatrix}$$

$$= \begin{bmatrix} -1.5\sin 5te^{-2t} + 3\cos 5te^{-2t} \\ 3\cos 5te^{-2t} + 6\sin 5te^{-2t} \end{bmatrix}$$

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -1.5\sin 5t + 3\cos 5t \\ 3\cos 5t + 6\sin 5t \end{bmatrix} e^{-2t}$$

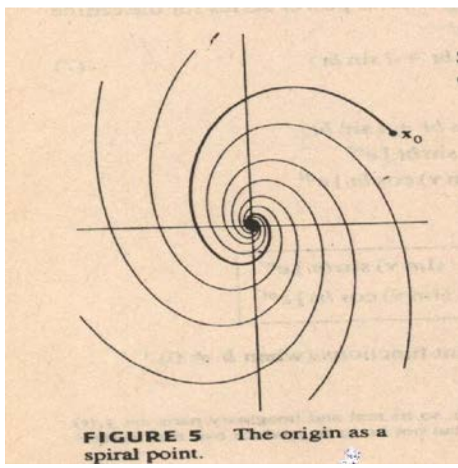


Figure 5 The origin as a spiral point

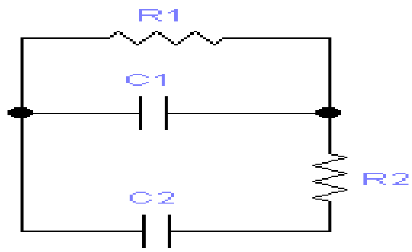
In this figure, due to the complex eigenvalues a rotation is caused by the sine and cosine function and hence the origin is called a spiral point. The trajectories spiral inwards as in this example the real part of the eigenvalue is negative.
Similarly, it is important to note that when A has a complex eigenvalue with positive real part, the trajectories spiral outwards .And if the real part of the eigenvalue is zero, then the trajectories form ellipses around the origin.

Exercise 1

The circuit in given figure can be described by

$$\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \begin{bmatrix} -\left(\frac{1}{R_1} + \frac{1}{R_2}\right)/C_1 & \frac{1}{R_2C_1} \\ \frac{1}{R_2C_2} & \frac{-1}{R_2C_2} \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

Where $v_1(t)$ and $v_2(t)$ are the voltages across the two capacitors at time t.



Suppose that resistor R_1 is 1 ohms, R_2 is 2 ohms, capacitor C_1 is 1 farad, and capacitor C_2 is 0.5 farad, and suppose that there is an initial charge of 7 volts on capacitor C_1 and 6 volts on capacitor C_2. Find formulas for $v_1(t)$ and $v_2(t)$ that describe how the voltages change over time.

Exercise 2

Find formulas for the current i_L and the voltage V_C for the circuit in Example 4, assuming that $R_1 = 10\,ohms, R_2 = 0.2\,ohm, C = 0.4\ farad, L = 0.2\,henry$ the initial current is 0 amp, and the initial voltage is 16 volts.

Lecture 35
Iterative Estimates for Eigenvalues

Eigenvalues and eigenvectors

Let A be an $n \times n$ square matrix, λ is any scalar value. If there exists a non-zero vector X, such that it satisfies the equation $AX = \lambda X$ then λ is called the eigenvalue and X is called the corresponding eigenvector.

i.e. if

$$AX = \lambda X$$
$$AX - \lambda X = 0$$
$$AX - \lambda IX = 0$$
$$(A - \lambda I)X = 0$$

To find the eigenvalues we have to solve the equation

$$|A - \lambda I| = 0$$

Example 1

Find the eigenvalues of $A = \begin{bmatrix} 1 & 0 \\ 1 & -2 \end{bmatrix}$.

Solution

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

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

$$|A - \lambda I| = (1-\lambda)(-2-\lambda)$$

Solving $|A - \lambda I| = 0$, we get

$$(1-\lambda)(-2-\lambda) = 0$$

So, either $(1-\lambda) = 0$ or $(-2-\lambda) = 0$

$so, \lambda = 1 \quad \lambda = -2$ are the eigenvalues of A.

Power Method

Let a matrix A is diagonalizable, with n linearly independent eigenvectors, $v_1, v_2, \ldots v_n$ and corresponding eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$ such that

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$$

Since $\{v_1, v_2, \ldots v_n\}$ is a basis for R^n, any vector x can be written as

$$x = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$$

Multiplying both sides by A, we get

$$\begin{aligned}Ax &= A(c_1v_1 + c_2v_2 + ... + c_nv_n)\\ &= A(c_1v_1) + A(c_2v_2) + ...A(c_nv_n)\\ &= c_1(Av_1) + c_2(Av_2) + ...c_n(Av_n)\\ &= c_1(\lambda_1 v_1) + c_2(\lambda_2 v_2) + ...c_n(\lambda_n v_n)\end{aligned}$$

Again multiplying by A and simplifying as above, we get

$$A^2x = c_1(\lambda_1{}^2 v_1) + c_2(\lambda_2{}^2 v_2) + ...c_n(\lambda_n{}^2 v_n)$$

Continuing this process we get

$$A^kx = c_1(\lambda_1{}^k v_1) + c_2(\lambda_2{}^k v_2) + ...c_n(\lambda_n{}^k v_n)$$

$$A^kx = c_1(\lambda_1)^k v_1 + c_2(\lambda_2)^k v_2) + ...c_n(\lambda_n)^k v_n \quad (k = 1, 2, ...)$$

$$(\frac{1}{\lambda_1})^k A^k x = c_1v_1 + c_2(\frac{\lambda_2}{\lambda_1})^k v_2) + ...c_n(\frac{\lambda_n}{\lambda_1})^k v_n \quad (k = 1, 2, ...)$$

$$(\lambda_1)^{-k} A^k x \to c_1v_1$$

$$as\, k \to \infty \qquad \text{Since } |\lambda_1| > |\lambda_2| \geq ... \geq |\lambda_n|$$

Thus for large k, a scalar multiple of $A^k x$ determines almost the same direction as the eigenvector c_1v_1. Since, positive scalar multiples do not change the direction of the vector, $A^k x$ itself points almost in the same direction as v_1 or $-v_1$, provided $c_1 \neq 0$.

Procedure for finding the Eigenvalues and Eigenvectors
Power Method

To compute the largest eigenvalue 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 eigenvalue and $v^{(k)}$ is the corresponding eigenvector.

Example 2

Let $A = \begin{pmatrix} 1.8 & .8 \\ 0.2 & 1.2 \end{pmatrix}$, $v_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$ and $X = \begin{pmatrix} -0.5 \\ 1 \end{pmatrix}$. Then A has Eigenvalues 2 and 1, and the eigenspace for $\lambda_1 = 2$ is the line through 0 and v_1. For k=0, …, 8, compute $A^k x$ and construct the line through 0 and $A^k x$. What happens as k increases?

Solution

The first three calculations are

$$\boldsymbol{Ax} = \begin{bmatrix} 1.8 & 0.8 \\ 0.2 & 1.2 \end{bmatrix}\begin{bmatrix} -0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.1 \\ 1.1 \end{bmatrix}$$

$$\boldsymbol{A^2 x} = \boldsymbol{A}(\boldsymbol{Ax}) = \begin{bmatrix} 1.8 & 0.8 \\ 0.2 & 1.2 \end{bmatrix}\begin{bmatrix} -0.1 \\ 1.1 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 1.3 \end{bmatrix}$$

$$\boldsymbol{A^3 x} = \boldsymbol{A}(\boldsymbol{A^2 x}) = \begin{bmatrix} 1.8 & 0.8 \\ 0.2 & 1.2 \end{bmatrix}\begin{bmatrix} 0.7 \\ 1.3 \end{bmatrix} = \begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$$

Analogous calculations complete Table 1.

Table 1 Iterates of a Vector

k	0	1	2	3	4	5	6	7	8
$\boldsymbol{A^k x}$	$\begin{bmatrix} -.5 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.1 \\ 1.1 \end{bmatrix}$	$\begin{bmatrix} 0.7 \\ 1.3 \end{bmatrix}$	$\begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$	$\begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$	$\begin{bmatrix} 11.9 \\ 4.1 \end{bmatrix}$	$\begin{bmatrix} 24.7 \\ 7.3 \end{bmatrix}$	$\begin{bmatrix} 50.3 \\ 13.7 \end{bmatrix}$	$\begin{bmatrix} 101.5 \\ 26.5 \end{bmatrix}$

The vectors $x, Ax, A^2x, ..., A^k x$ are shown in Fig. 1. The other vectors are growing too long to display. However, line segments are drawn showing the directions of those vectors. In fact, the directions of the vectors are what we really want to see, not the vectors themselves. The lines seem to be approaching the line representing the eigenspaces spanned by $\boldsymbol{v_1}$. More precisely, the angle between the line (subspace) determined by $A^k x$ and the line (eigenspaces) determined by v_1 goes to zero as $k \to \infty$.

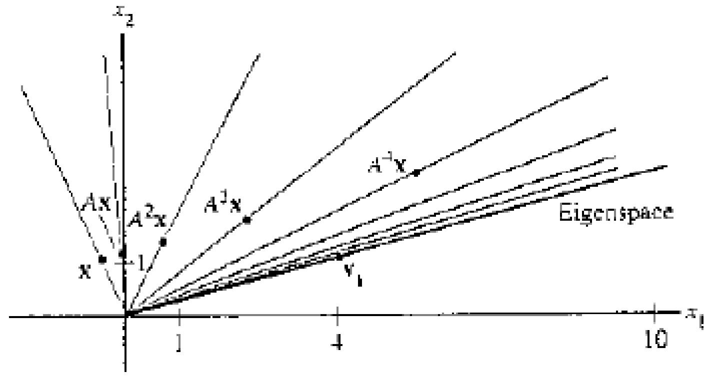


Figure 1 Directions determined by $\boldsymbol{x}$, $\boldsymbol{Ax}$, $\boldsymbol{A^2x}$, …, $\boldsymbol{A^7x}$

The vectors $(\lambda_1)^{-k}\boldsymbol{A^k x}$ in (3) are scaled to make them converge to c_1v_1, provided $c_1 \neq 0$. We cannot scale A^kx in this way because we do not know λ_1 but we can scale each A^kx to make its largest entry 1. It turns out that the resulting sequence $\{x_k\}$ will converge to a multiple of v_1 whose largest entry is 1. Figure 2 shows the scaled sequence for Example 1. The eigenvalue λ_1 can be estimated from the sequence $\{x_k\}$ too. When x_k is close to an eigenvector for λ_1 the vector Ax_k is close to $\lambda_1\boldsymbol{x_k}$ with each entry in Ax_k approximately λ_1 times the corresponding entry in x_k. Because the largest entry in x_k is 1 the largest entry in Ax_k is close to λ_1.

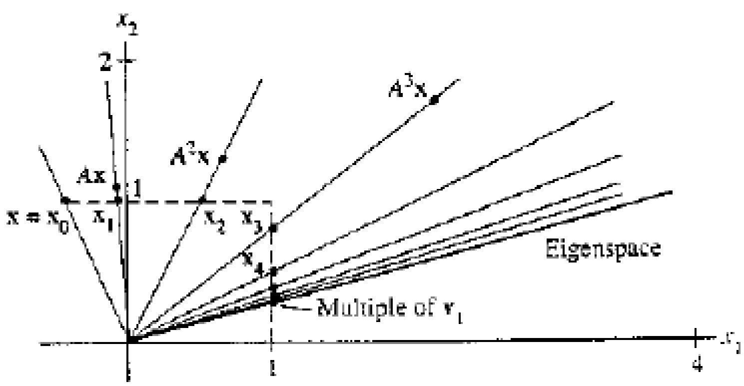


Figure 2 Scaled multiples of $x, Ax, A^2x, \ldots, A^7x$.

Steps for finding the eigenvalue and the eigenvector

1. Select an initial vector x_0 whose largest entry is 1.
2. For k=0,1,….
   a. Compute Ax_k
   b. Let μ_k be an entry in Ax_k whose absolute value is as large as possible.
   c. Compute $x_{k+1} = (\frac{1}{\mu_k})Ax_k$

3. For almost all choices of x_0, the sequence $\{\mu_k\}$ approaches the dominant eigenvalue, and the sequence $\{x_k\}$ approaches a dominant eigenvector.

Example 3

Apply the power method to $A = \begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}$ with $x_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Stop when $k = 5$, and estimate the dominant eigenvalue and a corresponding eigenvector of A.

Solution

To begin, we calculate Ax_0 and identify the largest entry μ_0 in Ax_0.

$$\boldsymbol{Ax_0} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix},\ \mu_0 = 5$$

Scale Ax_0 by $1/\mu_0$ to get x_1, compute Ax_1 and identify the largest entry in Ax_1

$$\boldsymbol{x_1} = \frac{1}{\mu_0}\boldsymbol{Ax_0} = \frac{1}{5}\begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ .4 \end{bmatrix}$$

$$\boldsymbol{Ax_1} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1.8 \end{bmatrix},\ \mu_1 = 8$$

Scale Ax_1 by $1/\mu_1$ to get $\boldsymbol{x_2}$, compute Ax_2, and identify the largest entry in Ax_2.

$$\boldsymbol{x_2} = \frac{1}{\mu_1}\boldsymbol{Ax_1} = \frac{1}{8}\begin{bmatrix} 8 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1 \\ .225 \end{bmatrix}$$

$$\boldsymbol{Ax_2} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .225 \end{bmatrix} = \begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix},\ \mu_2 = 7.125$$

Scale Ax_2 by $1/\mu_2$ to get x_3, and so on. The results of MATLAB calculations for the first five iterations are arranged in the Table 2.

The evidence from the Table 2 strongly suggests that $\{x_k\}$ approaches (1, .2) and $\{\mu_k\}$ approaches 7. If so, then (1, .2) is an eigenvector and 7 is the dominant eigenvalue. This is easily verified by computing $\boldsymbol{A} = \begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 7 \\ 1.4 \end{bmatrix} = 7\begin{bmatrix} 1 \\ .2 \end{bmatrix}$

TABLE 2 The Power Method for Example 2

k	0	1	2	3	4	5
x_k	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .4 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .225 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2035 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2005 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .20007 \end{bmatrix}$
Ax_k	$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 8 \\ 1.8 \end{bmatrix}$	$\begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}$	$\begin{bmatrix} 7.0175 \\ 1.4070 \end{bmatrix}$	$\begin{bmatrix} 7.0025 \\ 1.4010 \end{bmatrix}$	$\begin{bmatrix} 7.00036 \\ 1.40014 \end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

Example 4

Find the first three iterations of the power method applied on the following matrices

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix} \quad use\ x_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Solution

1st Iteration

$$u_1 = \mathrm{A}x_0 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0+0+0 \\ 0+0+2 \\ 0+0+2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$

Now, we normalize the resultant vector to get

$$u_1 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = q_1 x_1$$

2nd Iteration

$$u_2 = \mathrm{A}x_1 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0+1+0 \\ 0+4+2 \\ 0+1+2 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix}$$

Now, we normalize the resultant vector to get

$$u_2 = \begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix} = 6\begin{bmatrix} \frac{1}{6} \\ 1 \\ \frac{1}{2} \end{bmatrix} = q_2 x_2$$

3rd Iteration

$$u_3 = Ax_2 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{6} \\ 1 \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{6}+1+0 \\ \frac{1}{3}+4+1 \\ 0+1+1 \end{bmatrix} = \begin{bmatrix} \frac{7}{6} \\ \frac{16}{3} \\ 2 \end{bmatrix}$$

Now, we normalize the resultant vector to get

$$u_3 = \begin{bmatrix} \frac{7}{6} \\ \frac{16}{3} \\ 2 \end{bmatrix} = \frac{16}{3} \begin{bmatrix} \frac{7}{32} \\ 1 \\ \frac{3}{8} \end{bmatrix} = q_3 x_3$$

Hence the largest eigenvalue after 3 iterations is $\frac{16}{3}$.

The corresponding eigenvector is $\begin{bmatrix} \frac{7}{32} \\ 1 \\ \frac{3}{8} \end{bmatrix}$

<u>The inverse power method</u>

1. Select an initial estimate sufficiently close to λ.
2. Select an initial vector x_0 whose large entry is 1.
3. For k=0, 1, 2,…

   Solve $(A-\alpha I)y_k = x_k$.

   Let μ_k be an entry in y_k whose absolute value is as large as possible.

   Compute $v_k = \alpha + (\frac{1}{\mu_k})$.

   Compute $x_{k+1} = (\frac{1}{\mu_k})y_k$

4. For almost all the choice of x_0, the sequence $\{v_k\}$ approaches the eigenvalue λ of A, and the sequence $\{x_k\}$ approaches a corresponding eigenvector.

Example 5

It is not uncommon in some applications to need to know the smallest eigenvalue of a matrix A and to have at hand rough estimates of the eigenvalues. Suppose 21, 3.3, and 1.9 are estimates for the eigenvalues of the matrix A below. Find the smallest eigenvalue, accurate to six decimal places.

$$\boldsymbol{A} = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 4 \\ -4 & 5 & 4 \end{bmatrix}$$

Solution

The two smallest eigenvalues seem close together, so we use the inverse power method for $A-1.9I$. Results of a MATLAB calculation are shown in Table 3. Here x_0 was chosen arbitrarily, $y_k = (A-1.9I)^{-1}x_k$, μ_k is the largest entry in y_k, $\boldsymbol{v_k} = 1.9 + 1/\mu_k$, and $\boldsymbol{x_{k+1}} = (1/\mu_k)\boldsymbol{y_k}$.

Table 3: The Inverse Power Method

K	0	1	2	3	4
x_0	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.5736 \\ 0.0646 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.5054 \\ 0.0045 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.5004 \\ 0.0003 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.50003 \\ 0.00002 \\ 1 \end{bmatrix}$
y_k	$\begin{bmatrix} 4.45 \\ 0.50 \\ 7.76 \end{bmatrix}$	$\begin{bmatrix} 5.0131 \\ 0.0442 \\ 9.9197 \end{bmatrix}$	$\begin{bmatrix} 5.0012 \\ 0.0031 \\ 9.9949 \end{bmatrix}$	$\begin{bmatrix} 5.0001 \\ 0.0002 \\ 9.9996 \end{bmatrix}$	$\begin{bmatrix} 5.000006 \\ 0.000015 \\ 9.999975 \end{bmatrix}$
μ_k	7.76	9.9197	9.9949	9.9996	9.999975
v_k	2.03	2.0008	2.00005	2.000004	2.0000002

Therefore, we can say that eigenvalue is 2 from the matrix

$$\boldsymbol{A} = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 4 \\ -4 & 5 & 4 \end{bmatrix}$$

Exercise

How can you tell that if a given vector X is a good approximation to an eigenvector of a matrix A: if it is, how would you estimate the corresponding eigenvalue? Experiment with

$$A = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \text{ And } X = \begin{bmatrix} 1.0 \\ -4.3 \\ 8.1 \end{bmatrix}$$

Lecture 36

Revision

Revision of Previous Lectures

Lecture 37

Revision

Revision of Previous Lectures

Lecture 38

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

$u = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} and\ v = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$

$u^t = \begin{bmatrix} 2 & -5 & -1 \end{bmatrix}$

$$u.v = u^t v = \begin{bmatrix} 2 & -5 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = 2(3) + (-5)(2) + (-1)(-3)$$

$$= 6 - 10 + 3 = -1$$

$v^t = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix}$

$$v.u = v^t u = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = 3(2) + (2)(-5) + (-3)(-1)$$

$$= 6 - 10 + 3 = -1$$

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

$$(c_1u_1 + c_2u_2 + ...c_pu_p).w$$
$$= c_1(u_1.w) + c_2(u_2.w) + ...c_p(u_p.w)$$

Length or Norm

The length or Norm of v is the nonnegative scalar $\|v\|$ defined by

$$\|v\| = \sqrt{v.v} = \sqrt{v_1^{\ 2} + v_1^{\ 2} + ... + v_n^{\ 2}}$$
$$\|v\|^2 = v.v$$

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

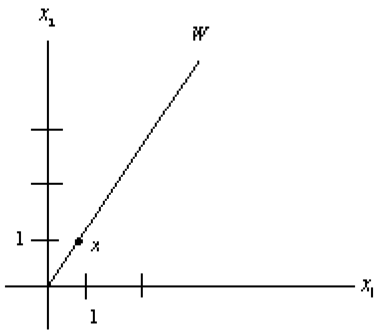
W consists of all multiples of ***x***, as in Fig. 2(a). Any nonzero vector in ***W*** is a basis for ***W***. To simplify the calculation, ***x*** is scaled to eliminate fractions. That is, multiply ***x*** by 3 to get

$$\boldsymbol{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

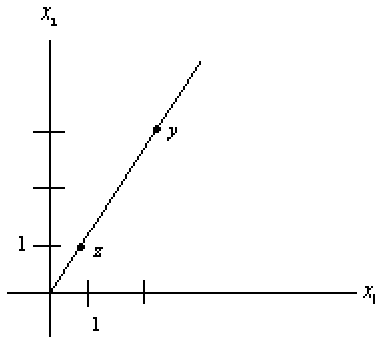
Now compute $\|\boldsymbol{y}\|^2 = 2^2 + 3^2 = 13, \|\boldsymbol{y}\| = \sqrt{13},$ and normalize ***y*** to get

$$\boldsymbol{z} = \frac{1}{\sqrt{13}}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}$$

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)

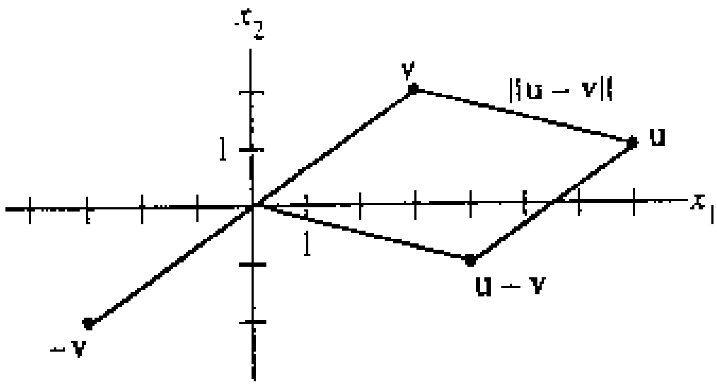
Solution

$$u-v = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 7-3 \\ 1-2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

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

$$\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos\vartheta \qquad (2)$$

To verify this formula for vectors in R^2, consider the triangle shown in Fig. 7, with sides of length $\|\boldsymbol{u}\|, \|\boldsymbol{v}\|$, and $\|\boldsymbol{u}-\boldsymbol{v}\|$. By the law of cosines, $\|\boldsymbol{u}-\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\vartheta$
which can be rearranged to produce

$$\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\vartheta = \frac{1}{2}\left[\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2\right]$$

$$= \frac{1}{2}\left[u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2\right] = u_1 v_1 + u_2 v_2 = \boldsymbol{u} \cdot \boldsymbol{v}$$

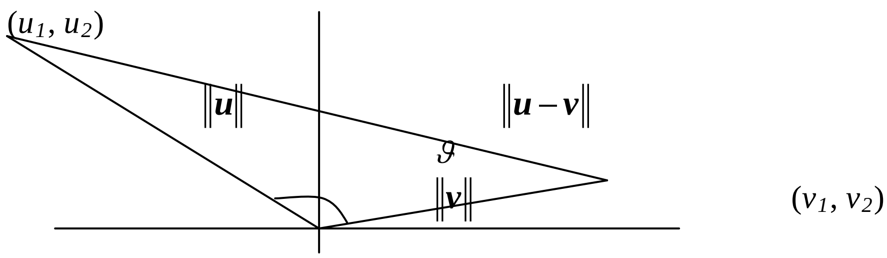


Figure 7 The angle between two vectors.

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

$\cos\vartheta=\dfrac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}$

Putting the values, we get

$\cos\theta=\dfrac{1}{\sqrt{\boldsymbol{6}}\sqrt{5}}=\dfrac{1}{\sqrt{30}}$

$\cos\theta=\dfrac{1}{\sqrt{30}}$

$\theta=\cos^{-1}\dfrac{1}{\sqrt{30}}=79.48^\circ$

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 No.39
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

$$0=c_1u_1+c_2u_2+...+c_pu_p.$$

Where $c_1,c_2...,c_p$ are scalars.

$$u_1.0=u_1.(c_1u_1+c_2u_2+...+c_pu_p)$$

$$\begin{aligned}0 &= u_1.(c_1u_1)+u_1.(c_2u_2)+...+u_1.(c_pu_p)\\&= c_1(u_1.u_1)+c_2(u_1.u_2)+...+c_p(u_1.u_p)\\&= c_1(u_1.u_1)\end{aligned}$$

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

$$\begin{aligned} y.u_1 &= (c_1u_1 + c_2u_2 + \dots + c_pu_p).u_1 \\ &= (c_1u_1).u_1 + (c_2u_2).u_1 + \dots + (c_pu_p).u_1 \\ &= c_1(u_1.u_1) + c_2(u_1.u_2) + \dots + c_p(u_1.u_p) \\ &= c_1(u_1.u_1). \end{aligned}$$

Since S is orthogonal set, so, $u_1.u_2 + \dots + u_1.u_p = 0$ but $u_1.u_1 > 0.$
Hence

$c_1 = \dfrac{y.u_1}{u_1.u_1}$ and similarly $c_2 = \dfrac{y.u_2}{u_2.u_2}, \dots\ c_p = \dfrac{y.u_p}{u_p.u_p}.$

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

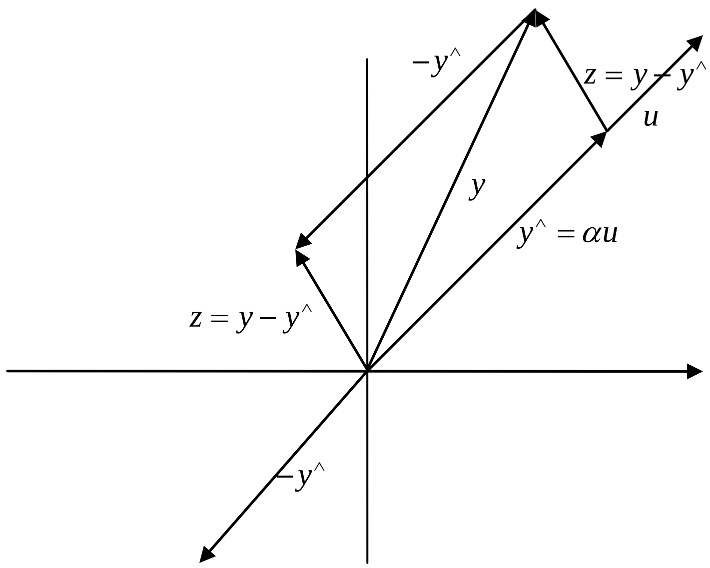
$$y = \begin{bmatrix}2\\1\\3\end{bmatrix}, u_1 = \begin{bmatrix}2\\-1\\0\end{bmatrix}, u_2 = \begin{bmatrix}\frac{1}{5}\\\frac{2}{5}\\-1\end{bmatrix} \text{ and } u_3 = \begin{bmatrix}\frac{8}{3}\\\frac{16}{3}\\\frac{8}{3}\end{bmatrix}$$

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

$$z = y - y^{\wedge}$$

$$= y - \frac{y.u}{u.u}u \quad \text{is a component of } y$$

<u>Example</u>

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**.

<u>Solution</u>

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

$$\hat{\boldsymbol{y}} \cdot (\boldsymbol{y} - \hat{\boldsymbol{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$

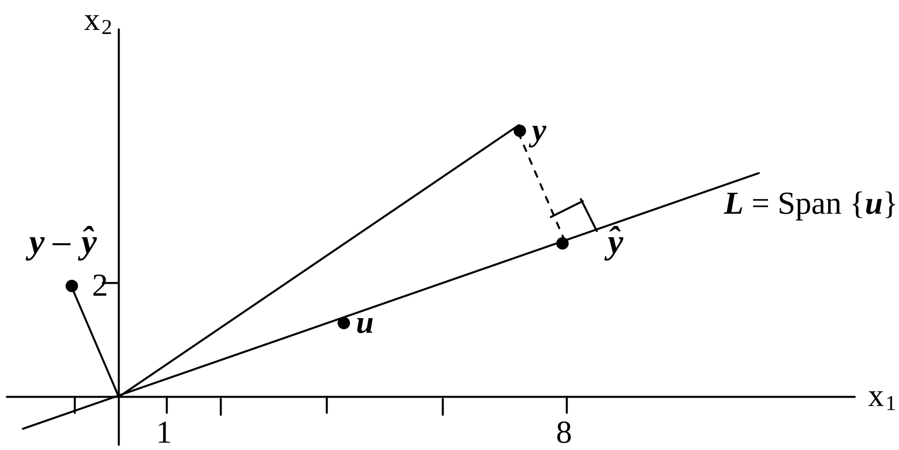
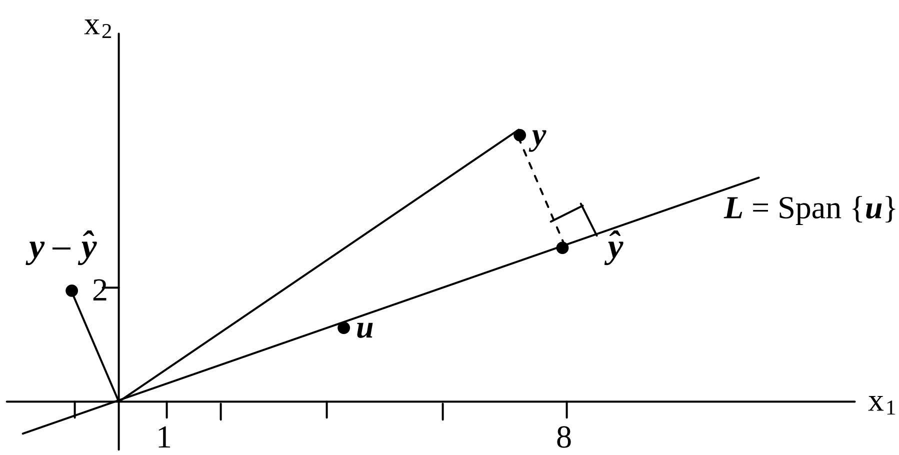


Figure 3 The orthogonal projection of ***y*** on to a line through the origin.

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,

$$y^{\wedge}=\begin{bmatrix}-\frac{3}{2}\\\frac{1}{2}\end{bmatrix}\ \ and\ \ z=\begin{bmatrix}-\frac{3}{2}\\-\frac{9}{2}\end{bmatrix}$$

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 $U^tU=I$

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

$$U=[u_1\ \ u_2\ \ u_3\ \dots\ u_m]$$

Taking transpose, it becomes

$$U^t = \begin{bmatrix} u_1^t \\ u_2^t \\ u_3^t \\ . \\ . \\ . \\ u_m^t \end{bmatrix}$$

$$U^t U = \begin{bmatrix} u_1^t \\ u_2^t \\ u_3^t \\ . \\ . \\ . \\ u_m^t \end{bmatrix} \begin{bmatrix} u_1 & u_2 & u_3 . & . & . & u_m \end{bmatrix} = \begin{bmatrix} u_1^t u_1 & u_1^t u_2 & u_1^t u_3 & \ldots & u_1^t u_m \\ u_2^t u_1 & u_2^t u_2 & u_2^t u_3 & \ldots & u_2^t u_m \\ u_3^t u_1 & u_3^t u_2 & u_3^t u_3 & \ldots & u_3^t u_m \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ u_m^t u_1 & u_m^t u_2 & u_m^t u_3 & \ldots & u_m^t u_m \end{bmatrix}$$

$As \quad u.v = v^t u$

$Therefore$

$$U^t U = \begin{bmatrix} u_1.u_1 & u_1.u_2 & u_1.u_3 & \ldots & u_1.u_m \\ u_2.u_1 & u_2.u_2 & u_2.u_3 & \ldots & u_2.u_m \\ u_3.u_1 & u_3.u_2 & u_3.u_3 & \ldots & u_3.u_m \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ u_m.u_1 & u_m.u_2 & u_m.u_3 & \ldots & u_m.u_m \end{bmatrix}$$

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

$$U^t U = \begin{bmatrix} u_1.u_1 & u_1.u_2 & u_1.u_3 & \cdots & u_1.u_m \\ u_2.u_1 & u_2.u_2 & u_2.u_3 & \cdots & u_2.u_m \\ u_3.u_1 & u_3.u_2 & u_3.u_3 & \cdots & u_3.u_m \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ u_m.u_1 & u_m.u_2 & u_m.u_3 & \cdots & u_m.u_m \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Clearly, it can be seen that

$u_i . u_j = 0 \; for \; i \neq j \quad i, j = 1, 2, \ldots m$

and

$u_i . u_j = 1 \; for \; i = j \quad i, j = 1, 2, \ldots m$

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

$u_i . u_j = 0 \; for \; i \neq j \quad i, j = 1, 2, \ldots m$

and

$u_i . u_j = 1 \; for \; i = j \quad i, j = 1, 2, \ldots m$

Hence, $U^t U = \begin{bmatrix} u_1.u_1 & u_1.u_2 & u_1.u_3 & \cdots & u_1.u_m \\ u_2.u_1 & u_2.u_2 & u_2.u_3 & \cdots & u_2.u_m \\ u_3.u_1 & u_3.u_2 & u_3.u_3 & \cdots & u_3.u_m \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ u_m.u_1 & u_m.u_2 & u_m.u_3 & \cdots & u_m.u_m \end{bmatrix}$

$$= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

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 No.40
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 $y^{\wedge} \in W$ *and* $z \in W^{\perp}$

Furthermore, if $\{u_1, u_2, ..., u_p\}$ is any orthogonal basis for W, then

$$y^{\wedge} = c_1u_1 + c_2u_2 + ... + c_nu_n, \, where\, c_j = \frac{y.u_j}{u_j.u_j}$$

Proof

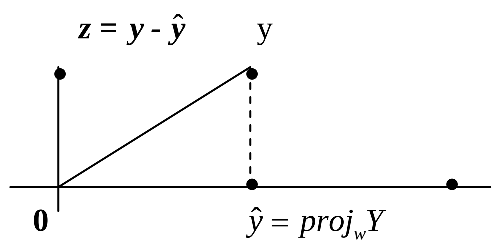


Fig: Orthogonal projection of y on to W.

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 $c_j = \dfrac{y.u_j}{u_j.u_j}$

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.

$$
\begin{aligned}
z \,.\, u_1 &= (y - y^\wedge).u_1 \\
&= y.u_1 - y^\wedge.u_1 \\
&= y.u_1 - (c_1u_1 + c_2u_2 + ... + c_pu_p).u_1 \\
&= y.u_1 - c_1(u_1.u_1) - c_2(u_2.u_1) - ...c_p(u_p.u_1) \\
&= y.u_1 - c_1(u_1.u_1) \quad \textit{where } u_j.u_1 = 0,\ j = 2, 3, ...p \\
&= y.u_1 - \frac{y.u_1}{u_1.u_1}(u_1.u_1) \\
&= y.u_1 - y.u_1 \\
&= 0
\end{aligned}
$$

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:

$$y^\wedge = c_1 u_1 + c_2 u_2$$

$$= \frac{y.u_1}{u_1.u_1} u_1 + \frac{y.u_2}{u_2.u_2} u_2$$

$$= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -2 \\ 10 \\ 1 \end{bmatrix}$$

$$y - y^\wedge = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

$$= 7/5\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

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 $\boldsymbol{y}$ to $\boldsymbol{W}$ =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 $\boldsymbol{y}$ to $\boldsymbol{W}$ is $\|y - \hat{y}\|$, where $\hat{\boldsymbol{y}}$ = proj$_{\boldsymbol{w}}\boldsymbol{y}$. Since, $\{\boldsymbol{u_1}, \boldsymbol{u_2}\}$ is an orthogonal basis for $\boldsymbol{W}$, 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 $\boldsymbol{y}$ to $\boldsymbol{W}$ is $\sqrt{45} = 3\sqrt{5}.$

Theorem

If $\{u_1, u_2, ..., u_p\}$ is an orthonormal basis for a subspace W of R^n , then

$$\Pr oj_w y = (y.u_1)u_1 + (y.u_2)u_2 + ...(y.u_p)u_p$$

$$If\ \ U = [u_1\ u_2 \ ....u_p]$$

$$then\ \Pr oj_w^y = UU^T y \quad \forall y\, in\, R^n$$

Example

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 $\boldsymbol{W}$ =span$\{u_1, u_2\}$. Use the fact that u_1 and u_2 are orthogonal to compute $\Pr oj_w y$.

Solution

$$\mathrm{Proj}_w y = \frac{y.u_1}{u_1.u_1}u_1 + \frac{y.u_2}{u_2.u_2}u_2$$
$$= \frac{88}{66}u_1 + \frac{-2}{6}u_2$$

$$= \frac{4}{3}\begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix} = y$$

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 # 41

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 $\boldsymbol{x_2} - \boldsymbol{P}$, 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

$$v_1 = x_1 \quad v_2 = x_2 - \frac{x_2.v_1}{v_1.v_1}v_1$$

$$v_3 = x_3 - \frac{x_3.v_1}{v_1.v_1}v_1 - \frac{x_3.v_2}{v_2.v_2}v_2$$

$$\vdots$$

$$v_p = x_p - \frac{x_p.v_1}{v_1.v_1}v_1 - \frac{x_p.v_2}{v_2.v_2}v_2 - \ldots - \frac{x_p.v_{p-1}}{v_{p-1}.v_{p-1}}v_{p-1}$$

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.

$$x_1 = \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} and \ x_2 = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$$

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 = ***x₁***

$$v_2 = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \frac{15}{30} \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ -5/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3/2 \\ 3/2 \end{bmatrix}$$

$$Now \left\| v_2 \right\| = \frac{\sqrt{54}}{2} = \frac{3\sqrt{6}}{2} \text{ and Since } \left\| v_1 \right\| = \sqrt{30}$$

Thus the orthonormal basis for W is

$$\left\{ \frac{\boldsymbol{v_1}}{\left\|\boldsymbol{v_1}\right\|}, \frac{\boldsymbol{v_2}}{\left\|\boldsymbol{v_2}\right\|} \right\} = \left\{ \begin{bmatrix} 2/\sqrt{30} \\ -5/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}, \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \right\}$$

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

First 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 $U = [u_1 \quad u_2 \quad ... \quad u_p]$,

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 <u>the best approximation to</u> *y* <u>by elements of</u> *W*.

Exercise 1

Let W = Span {$x1$, $x2$}, where

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \; and \; x_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$$

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 # 42
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

$$A\hat{x} = \hat{b} \tag{1}$$

Since $\hat{b}$ is the closest point in Col ***A*** to ***b***, a vector $\hat{x}$ is a least-squares solution of ***Ax*** = ***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 $\boldsymbol{A^T}$,

$$A^T (b - A\hat{x}) = 0 \tag{2}$$

$$A^T b - A^T A\hat{x} = 0$$

$$A^T A\hat{x} = A^T b \tag{3}$$

The matrix equation (3) represents a system of linear equations commonly referred to as the **normal equations for** $\hat{\boldsymbol{x}}$**.**

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 <u>least-squares solution</u> 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 $(A^T A)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$

Therefore, $\hat{x} = (A^T A)^{-1} A^T b = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

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 $\left\|\boldsymbol{b} - \boldsymbol{A\hat{x}}\right\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$

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

$$A\tilde{x} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & 0 \\ 7 & 1 & 4 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} -.676 \\ -.776 \\ .834 \end{bmatrix}$$

$$A\tilde{x} = \begin{bmatrix} 0.548 \\ 1.652 \\ -2.172 \\ 3.494 \end{bmatrix}$$

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

<u>Theorem</u>

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

$\hat{x} = R^{-1}Q^T b$

<u>Example 1</u>

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,

$$Q = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{bmatrix} \text{ And}$$

$$R = Q^T A = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}$$

Then $$Q^T b = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

The least-squares solution $\hat{\boldsymbol{x}}$ satisfies $R\hat{x} = Q^T b;$ 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 $R\hat{x} = Q^T b$ 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

$$Q = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} \; and \; Q^T = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$Now \; R = Q^T A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 15 & 5 \\ 0 & 2 \end{bmatrix} \qquad And \; Q^T b = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

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 # 43
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

$$\langle u,u\rangle = u^t Au > 0 \qquad for\, all\;\; u \neq 0$$

$$So \quad \langle u,u\rangle = u^t Au = 0 \qquad iff \;\; u = 0$$

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 ***p*** must vanish at n + 1 points: t_0, …, t_n. This is possible only if ***p*** is the zero polynomial, because the degree of ***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

$$P(-1)=3\ , P(0)=4\ ,\ P(1)=5$$
$$q(-1)=1\ ,\ q(0)=5\ ,\ q(1)=1$$

$$q.p=28 \qquad p.p=50$$

$$\hat{q}=\frac{q.p}{p.p}p=\frac{28}{50}(4+t)$$
$$=\frac{56}{25}+\frac{14}{25}t$$

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 $p(t)=12t^2$ *and* $q(t)=2t-1$

Compute $\langle p,q\rangle$ *and* $\langle q,q\rangle$

Solution

$$\langle p,q\rangle = p(0)q(0)+p\left(\tfrac{1}{2}\right)+p(1)q(1)$$
$$=(0)(-1)+(3)(0)+(12)(1)=12$$

$$\langle q,q\rangle = [q(0)]^2+\left[q\left(\frac{1}{2}\right)\right]^2+[q(1)]^2$$
$$=(-1)^2+(0)^2+(1)^2=2$$

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

$$\|p\|^2=\langle p,p\rangle=[p(0)]^2+\left[p\left(\frac{1}{2}\right)\right]^2+[p(1)]^2$$
$$=(0)^2+(3)^2+(12)^2=153$$
$$\|p\|=\sqrt{153}$$

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 $\|x\|$, $\|y\|\, and\, |\langle x,y\rangle|^2$ 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 $\langle x,y\rangle=4x_1y_1+5x_2y_2$

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{4(1)(1) + 5(1)(1)}$$
$$= \sqrt{4+5} = \sqrt{9} = 3$$
$$\|y\| = \sqrt{\langle y, y\rangle} = \sqrt{4(5)(5) + 5(-1)(-1)}$$
$$= \sqrt{100+5} = \sqrt{105}$$

$$|\langle x, y\rangle|^2 = \langle x, y\rangle\langle x, y\rangle$$
$$= [4(1)(5) + 5(1)(-1)]^2$$
$$= [20-5]^2$$
$$= [15]^2 = 225$$

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.

<u>Example 9</u>

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 $1, t \,\&\, t^2$.

<u>Solution</u>

Given polynomials **1** ***t*** $\boldsymbol{t^2}$ at -2,-1, 0, 1 and 2

Polynomial: **1** ***t*** $\boldsymbol{t^2}$

$$\text{Vector of values: } \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix}$$

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 $\frac{10}{5}p_0 + 0\,p_1$. 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:

Polynomial: $\boldsymbol{p_0}$ $\boldsymbol{p_1}$ $\boldsymbol{p_2}$

$$\text{Vector of values: } \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}, \begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}$$

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: $\boldsymbol{p_0}$ $\boldsymbol{p_1}$ $\boldsymbol{p_2}$

$$\text{Vector of values: } \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, \quad \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}, \quad \begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}$$

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.

<u>Cauchy – Schwarz Inquality</u>

$$For\, all \quad u, v \ in\ V$$

$$\left|\langle u, v\rangle\right| \le \|u\|\, \|v\|$$

<u>Triangle Inequality</u>

$$For\ all\ \ u, v\ in\ V$$

$$\|u+v\| \ \le\ \|u\| + \|v\|$$

<u>Proof</u>

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle \\
&\le \|u\|^2 + 2\left|\langle u,v\rangle\right| + \|v\|^2 \\
&\le \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\
\|u+v\|^2 &= \left(\|u\| + \|v\|\right)^2
\end{aligned}$$

$$\Rightarrow \quad \|u+v\| = \|u\| + \|v\|$$

<u>Inner product for $C[a,b]$</u>

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
$p_o(t)=1$, $p_1(t)=t,$ *and* $p_2(t)=t^2$
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
$p_o(t)=1$, $p_1(t)=t,$ *and* $p_2(t)=t^2$
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 44
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 $(x_1, y_1), ..., (x_n, y_n)$ 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,

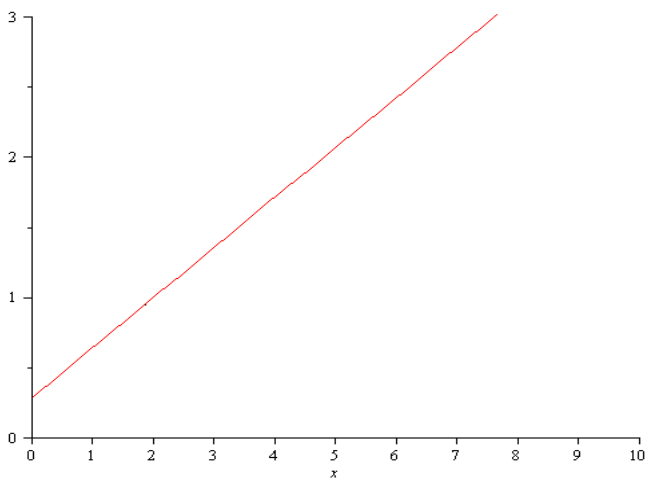
$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

$$= \frac{1}{84}\begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

$$= \frac{1}{84}\begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

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, $y_1, y_2, ..., y_n$ 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

$$y = 1.0 - 0.1x$$

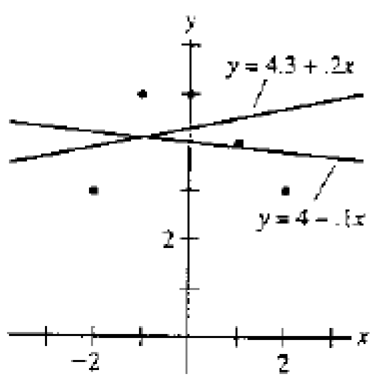


FIGURE 1 Weighted and ordinary least-squares lines.

What Does Trend Analysis Mean?

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 $t_0, t_1, ..., t_n$ 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

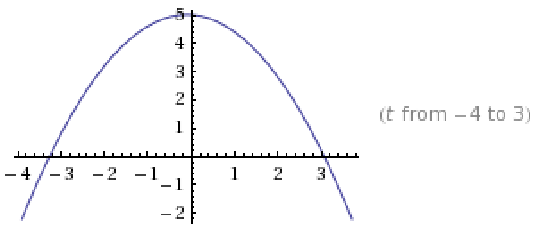
$$\begin{array}{lcccc} \text{Polynomial:} & \text{p}_0 & \text{p}_1 & \text{p}_2 & \text{data: g} \\ \textit{Vector of values:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix} \end{array}$$

$$\hat{p} = \frac{< g, p_0 >}{< p_0, p_0 >} p_0 + \frac{< g, p_1 >}{< p_1, p_1 >} p_1 + \frac{< g, p_2 >}{< p_2, p_2 >} p_2$$

$$= \frac{20}{5} p_0 - \frac{1}{10} p_1 - \frac{7}{14} p_2$$

$$\textit{and} \qquad \hat{p}(t) = 4 - 0.1t - 0.5(t^2 - 2)$$

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 $m \neq n$

$$< \cos mt, \cos nt >= \int_0^{2\pi} \cos mt \cos nt\,dt$$

$$= \frac{1}{2}\int_0^{2\pi} [\cos(mt + nt) + \cos(mt - nt)\,dt$$

$$= \frac{1}{2}\left[\frac{\sin(mt+nt)}{m+n} + \frac{\sin(mt-nt)}{m-n}\right]\Bigg|_0^{2\pi}$$
$$= 0$$

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

$f(t) = 3 - 2\sin t + 5\sin 2t - 6\cos 2t$

Revision of Previous Lectures