

Masters
Subscribes to Masters

Linear Algebra

MTH 501

highlighted by Masters



Virtual University of Pakistan
Knowledge beyond the boundaries

TABLE OF CONTENTS :

✓ Lesson 1: Introduction and Overview	2
✓ Lesson 2: Introduction to Matrices	10
✓ Lesson 3: Systems of Linear Equations	23
✓ Lesson 4: Row Reduction and Echelon Forms	36
✓ Lesson 5: Vector Equations	46
✓ Lesson 6: Matrix Equations	62
✓ Lesson 7: Solution Sets of Linear Systems.....	74
✓ Lesson 8: Linear Independence	84
✓ Lesson 9: Linear Transformations	94
✓ Lesson 10: The Matrix of a Linear Transformation.....	106
✓ Lesson 11: Matrix Operations.....	116
✓ Lesson 12: The Inverse of a Matrix	129
✓ Lesson 13: Characterizations of Invertible Matrices	143
✓ Lesson 14: Partitioned Matrices.....	157
✓ Lesson 15: Matrix Factorizations.....	175
✓ Lesson 16: Iterative Solutions of Linear Systems.....	192
✓ Lesson 17: Introduction to Determinant	203
✓ Lesson 18: Properties of Determinants	219
✓ Lesson 19: Cramer's Rule, Volume, and Linear Transformations	234
✓ Lesson 20: Vector Spaces and Subspaces.....	245
✓ Lesson 21: Null Spaces, Column Spaces, and Linear Transformations	269
✓ Lesson 22: Linearly Independent Sets; Bases.....	292
✓ Lesson 23: Coordinate System.....	308
✓ Lesson 24: Dimension of a Vector Space	323
✓ Lesson 25: Rank.....	341
✓ Lesson 26: Change of Basis	359
✓ Lesson 27: Applications to Difference Equations	366
✓ Lesson 28: Eigenvalues and Eigenvectors	374
✓ Lesson 29: The Characteristic Equation	384
✓ Lesson 30: Diagonalization.....	396
✓ Lesson 31: Eigenvectors and Linear Transformations.....	406
✓ Lesson 32: Complex Eigenvalues.....	413
✓ Lesson 33: Discrete Dynamical Systems	419
✓ Lesson 34: Applications to Differential Equations	429
✓ Lesson 35: Iterative Estimates for Eigenvalues	442
✓ Lesson 36: Revision.....	451
✓ Lesson 37: Revision.....	452
✓ Lesson 38: Inner Product	453
✓ Lesson 39: Orthogonal and Orthonormal set	460
✓ Lesson 40: Orthogonal Decomposition	477
✓ Lesson 41: Orthogonal basis, Gram-Schmidt Process, Orthonormal basis	485
✓ Lesson 42: Least Square Solution.....	493
✓ Lesson 43: Inner Product Space.....	502
✓ Lesson 44: Application of inner product spaces	512
✓ Lesson 45: Revision.....	523

Masters

Subscribes to Masters

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 2×2 and 3×3 systems of linear equations learned in previous classes.

It is exactly what the name suggests. Simply put, it is the algebra of systems of linear equations. While you could solve a system of, say, five linear equations involving five unknowns, it might not take a finite amount of time. With linear algebra we develop techniques to solve m linear equations and n unknowns, or show when no solution exists. We can even describe situations where an infinite number of solutions exist, and describe them geometrically.

Linear algebra is the study of linear sets of equations and their transformation properties.

Linear algebra, sometimes disguised as matrix theory, considers sets and functions, which preserve linear structure. In practice this includes a very wide portion of mathematics!

Thus linear algebra includes axiomatic treatments, computational matters, algebraic structures, and even parts of geometry; moreover, it provides tools used for analyzing differential equations, statistical processes, and even physical phenomena.

Linear Algebra consists of studying matrix calculus. It formalizes and gives geometrical interpretation of the resolution of equation systems. It creates a formal link between matrix calculus and the use of linear and quadratic transformations. It develops the idea of trying to solve and analyze systems of linear equations.

Applications of Linear algebra

Linear algebra makes it possible to work with large arrays of data. It has many applications in many diverse fields, such as

- Computer Graphics,
- Electronics,
- Chemistry,
- Biology,
- Differential Equations,
- Economics,
- Business,
- Psychology,
- Engineering,
- Analytic Geometry,
- Chaos Theory,
- Cryptography,
- Fractal Geometry,
- Game Theory,
- Graph Theory,
- Linear Programming,
- Operations Research

It is very important that the theory of linear algebra is first understood, the concepts are cleared and then computation work is started. Some of you might want to just use the

computer, and skip the theory and proofs, but if you don't understand the theory, then it can be very hard to appreciate and interpret computer results.

Why using Linear Algebra?

Linear Algebra allows for formalizing and solving many typical problems in different engineering topics. It is generally the case that (input or output) data from an experiment is given in a discrete form (discrete measurements). Linear Algebra is then useful for solving problems in such applications in topics such as Physics, Fluid Dynamics, Signal Processing and, more generally Numerical Analysis.

Linear algebra is not like algebra. It is mathematics of linear spaces and linear functions. So we have to know the term "linear" a lot. Since the concept of linearity is fundamental to any type of mathematical analysis, this subject lays the foundation for many branches of mathematics.

Objects of study in linear algebra

Linear algebra merits study at least because of its ubiquity in mathematics and its applications. The broadest range of applications is through the concept of vector spaces and their transformations. These are the central objects of study in linear algebra

1. The solutions of homogeneous systems of linear equations form paradigm examples of vector spaces. Of course they do not provide the only examples.
2. The vectors of physics, such as force, as the language suggests, also provide paradigmatic examples.
3. Binary code is another example of a vector space, a point of view that finds application in computer sciences.
4. Solutions to specific systems of differential equations also form vector spaces.
5. Statistics makes extensive use of linear algebra.
6. Signal processing makes use of linear algebra.
7. Vector spaces also appear in number theory in several places, including the study of field extensions.
8. Linear algebra is part of and motivates much abstract algebra. Vector spaces form the basis from which the important algebraic notion of module has been abstracted.

9. Vector spaces appear in the study of differential geometry through the tangent bundle of a manifold.
10. Many mathematical models, especially discrete ones, use matrices to represent critical relationships and processes. This is especially true in engineering as well as in economics and other social sciences.

There are two principal aspects of linear algebra: theoretical and computational. A major part of mastering the subject consists in learning how these two aspects are related and how to move from one to the other.

Many computations are similar to each other and therefore can be confusing without reasonable level of grasp of their theoretical context and significance. It will be very tempting to draw false conclusions.

On the other hand, while many statements are easier to express elegantly and to understand from a purely theoretical point of view, to apply them to concrete problems you will need to “get your hands dirty”. Once you have understood the theory sufficiently and appreciate the methods of computation, you will be well placed to use software effectively, where possible, to handle large or complex calculations.

Course Segments

The course is covered in 45 Lectures spanning over six major segments, which are given below;

1. Linear Equations
2. Matrix Algebra
3. Determinants
4. Vector spaces
5. Eigen values and Eigenvectors, and
6. Orthogonal sets

Course Objectives

The main purpose of the course is to introduce the concept of linear algebra, to explain the underline theory, the computational techniques and then try to apply them on real life problems. Major course objectives are as under;

- To master techniques for solving systems of linear equations
- To introduce matrix algebra as a generalization of the single-variable algebra of high school.
- To build on the background in Euclidean space and formalize it with vector space theory.
- To develop an appreciation for how linear methods are used in a variety of applications.
- To relate linear methods to other areas of mathematics such as calculus and, differential equations.

Recommended Books and Supported Material

I am indebted to several authors whose books I have freely used to prepare the lectures that follow. The lectures are based on the material taken from the books mentioned below.

1. **Linear Algebra and its Applications** (3rd Edition) by David C. Lay.
2. **Contemporary Linear Algebra** by Howard Anton and Robert C. Busby.
3. **Introductory Linear Algebra** (8th Edition) by Howard Anton and Chris Rorres.
4. **Introduction to Linear Algebra** (3rd Edition) by L. W. Johnson, R.D. Riess and J.T. Arnold.
5. **Linear Algebra** (3rd Edition) by S. H. Friedberg, A.J. Insel and L.E. Spence.
6. **Introductory Linear Algebra with Applications** (6th Edition) by B. Kolman.

I have taken the structure of the course as proposed in the book of David C. Lay. I would be following this book. I suggest that the students should purchase this book, which is easily available in the market and also does not cost much. For further study and supplement, students can consult any of the above mentioned books.

I strongly suggest that the students should also browse on the Internet; there is plenty of supporting material available. In particular, I would suggest the website of David C. Lay; www.laylinalgebra.com, where the entire material, study guide, transparencies are readily available. Another very useful website is www.wiley.com/college/anton, which contains a variety of useful material including the data sets. A number of other books are also available in the market and on the internet with free access.

I will try to keep the treatment simple and straight. The lectures will be presented in simple Urdu and easy English. These lectures are supported by the handouts in the form of lecture notes. The theory will be explained with the help of examples. There will be enough exercises to practice with. Students are advised to go through the course on daily basis and do the exercises regularly.

Schedule and Assessment

The course will be spread over 45 lectures. Lectures one and two will be introductory and the Lecture 45 will be the summary. The first two lectures will lay the foundations and would provide the overview of the course. These are important from the conceptual point of view. I suggest that these two lectures should be viewed again and again.

The course will be interesting and enjoyable, if the student will follow it regularly and completes the exercises as they come along. To follow the tradition of a semester system or of a term system, there will be a series of assignments (Max eight assignments) and a mid term exam. Finally there will be terminal examination.

The assignments have weights and therefore they have to be taken seriously.

Masters

Subscribes to Masters

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 $A, B, \dots, Y, Z$. The numbers or functions are called elements of the matrix. The elements of a matrix are denoted by small letters $a, b, \dots, y, z$.

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

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

The element, or entry, in the i th row and j th 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

$$(a) \quad 5 \cdot \begin{bmatrix} 2 & -3 \\ 4 & -1 \\ 1/5 & 6 \end{bmatrix} = \begin{bmatrix} 10 & -15 \\ 20 & -5 \\ 1 & 30 \end{bmatrix}$$

$$(b) \quad e^t \cdot \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} e^t \\ -2e^t \\ 4e^t \end{bmatrix}$$

Since we know that $kA = Ak$. Therefore, we can write

$$e^{-3t} \cdot \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2e^{-3t} \\ 5e^{-3t} \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix} e^{-3t}$$

Addition of Matrices Only matrices of the same order may be added by adding corresponding elements.

If $A = [a_{ij}]$ and $B = [b_{ij}]$ are two $m \times n$ matrices then $A + B = [a_{ij} + b_{ij}]$

Obviously order of the matrix $A + B$ is $m \times n$

Example 2 Consider the following two matrices of order 3×3

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

Since the given matrices have same orders, therefore, these matrices can be added and their sum is given by

$$A + B = \begin{pmatrix} 2+4 & -1+7 & 3+(-8) \\ 0+9 & 4+3 & 6+5 \\ -6+1 & 10+(-1) & -5+2 \end{pmatrix} = \begin{pmatrix} 6 & 6 & -5 \\ 9 & 7 & 11 \\ -5 & 9 & -3 \end{pmatrix}$$

Example 3 Write the following single column matrix as the sum of three column vectors

$$\begin{pmatrix} 3t^2 - 2e^t \\ t^2 + 7t \\ 5t \end{pmatrix}$$

Solution

$$\begin{pmatrix} 3t^2 - 2e^t \\ t^2 + 7t \\ 5t \end{pmatrix} = \begin{pmatrix} 3t^2 \\ t^2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 7t \\ 5t \end{pmatrix} + \begin{pmatrix} -2e^t \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} t^2 + \begin{pmatrix} 0 \\ 7 \\ 5 \end{pmatrix} t + \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} e^t$$

Difference of Matrices The difference of two matrices A and B of same order $m \times n$ is defined to be the matrix $A - B = A + (-B)$

The matrix $-B$ is obtained by multiplying the matrix B with -1 . So that $-B = (-1)B$

Multiplication of Matrices We can multiply two matrices if and only if, the number of columns in the first matrix equals the number of rows in the second matrix.

Otherwise, the product of two matrices is not possible.

OR

If the order of the matrix A is $m \times n$ then to make the product AB possible order of the matrix B must be $n \times p$. Then the order of the product matrix AB is $m \times p$. Thus

$$A_{m \times n} \cdot B_{n \times p} = C_{m \times p}$$

If the matrices A and B are given by

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

Then

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

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

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

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.

$$\begin{aligned} AB &= \begin{pmatrix} 5 & 8 \\ 1 & 0 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} -4 & -3 \\ 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 5 \cdot (-4) + 8 \cdot 2 & 5 \cdot (-3) + 8 \cdot 0 \\ 1 \cdot (-4) + 0 \cdot 2 & 1 \cdot (-3) + 0 \cdot 0 \\ 2 \cdot (-4) + 7 \cdot 2 & 2 \cdot (-3) + 7 \cdot 0 \end{pmatrix} = \begin{pmatrix} -4 & -15 \\ -4 & -3 \\ 6 & -6 \end{pmatrix} \\ AB &= \begin{pmatrix} 78 & 48 \\ 57 & 34 \end{pmatrix}, \quad BA = \begin{pmatrix} 30 & 53 \\ 48 & 82 \end{pmatrix} \end{aligned}$$

Clearly $AB \neq BA$.

$$AB = \begin{pmatrix} -4 & -15 \\ -4 & -3 \\ 6 & -6 \end{pmatrix}$$

However, the product BA is not possible.

Example 5

$$(a) \quad \begin{pmatrix} 2 & -1 & 3 \\ 0 & 4 & 5 \\ 1 & -7 & 9 \end{pmatrix} \begin{pmatrix} -3 \\ 6 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \cdot (-3) + (-1) \cdot 6 + 3 \cdot 4 \\ 0 \cdot (-3) + 4 \cdot 6 + 5 \cdot 6 \\ 1 \cdot (-3) + (-7) \cdot 6 + 9 \cdot 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 44 \\ -9 \end{pmatrix}$$

$$(b) \quad \begin{pmatrix} -4 & 2 \\ 3 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -4x + 2y \\ 3x + 8y \end{pmatrix}$$

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 2×2 and 3×3

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 $O = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \quad O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$

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

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

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.

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

Since order of the matrix A is $m \times n$, the order of the transpose matrix A^{tr} is $n \times m$.

Properties of the Transpose

The following properties are valid for the transpose;

- The transpose of the transpose of a matrix is the matrix itself: $(A^{tr})^{tr} = A$
- The transpose of a matrix times a scalar (k) is equal to the constant times the transpose of the matrix: $(\underline{ABC})^T = \underline{C}^T \underline{B}^T \underline{A}^T$ $(kA)^T = kA^T$
- 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: $(\underline{AB})^T = \underline{B}^T \underline{A}^T$
- The same is true for the product of multiple matrices: $(\underline{ABC})^T = \underline{C}^T \underline{B}^T \underline{A}^T$

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 $X^T = [5 \ 0 \ 3]$

Multiplicative Inverse Suppose that A is a square matrix of order $n \times n$. If there exists an $n \times n$ matrix B such that $AB = BA = I$, then B is said to be the multiplicative inverse of the matrix A and is denoted by $B = A^{-1}$.

For example: If $A = \begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix}$ then the matrix $B = \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix}$ is multiplicative inverse of A

because $AB = \begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$

Similarly we can check that $BA = I$

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

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

$$\begin{aligned} |A| &= 2(5-0) - 3(5-0) - 1(-3-2) \\ &= 10 - 15 + 5 = 0 \end{aligned}$$

which means that A is singular.

Minor of an element of a matrix

Let A be a square matrix of order $n \times n$. 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 i th row and j th 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 $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

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

Therefore, inverse of the matrix A is given by $A^{-1} = \frac{1}{\det A} \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix}$.

Example 8 Find, if possible, the multiplicative inverse for the matrix $A = \begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix}$.

Solution The matrix A is non-singular because $\det(A) = \begin{vmatrix} 1 & 4 \\ 2 & 10 \end{vmatrix} = 10 - 8 = 2$

Therefore, A^{-1} exists and is given by $A^{-1} = \frac{1}{2} \begin{pmatrix} 10 & -4 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix}$

Check $AA^{-1} = \begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix} = \begin{pmatrix} 5-4 & -2+2 \\ 10-10 & -4+5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$

$$AA^{-1} = \begin{pmatrix} 5 & -2 \\ -1 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} 5-4 & 20-20 \\ -1+1 & -4+5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Example 9 Find, if possible, the multiplicative inverse of the following matrix

$$A = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$$

Solution The matrix is singular because

$$\det(A) = \begin{vmatrix} 2 & 2 \\ 3 & 3 \end{vmatrix} = 2 \cdot 3 - 2 \cdot 3 = 0$$

Therefore, the multiplicative inverse A^{-1} of the matrix does not exist.

Example 10 Find the multiplicative inverse for the following matrix

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

Solution Since $\det(A) = \begin{vmatrix} 2 & 2 & 0 \\ -2 & 1 & 1 \\ 3 & 0 & 1 \end{vmatrix} = 2(1-0) - 2(-2-3) + 0(0-3) = 12 \neq 0$

Therefore, the given matrix is non singular. So, the multiplicative inverse A^{-1} of the matrix A exists. The cofactors corresponding to the entries in each row are

$$\begin{aligned} C_{11} &= \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1, & C_{12} &= -\begin{vmatrix} -2 & 1 \\ 3 & 1 \end{vmatrix} = 5, & C_{13} &= \begin{vmatrix} -2 & 1 \\ 3 & 0 \end{vmatrix} = -3 \\ C_{21} &= -\begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} = -2, & C_{22} &= \begin{vmatrix} 2 & 0 \\ 3 & 1 \end{vmatrix} = 2, & C_{23} &= -\begin{vmatrix} 2 & 2 \\ 3 & 0 \end{vmatrix} = 6 \\ C_{31} &= \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} = 2, & C_{32} &= -\begin{vmatrix} 2 & 0 \\ -2 & 1 \end{vmatrix} = -2, & C_{33} &= \begin{vmatrix} 2 & 2 \\ -2 & 1 \end{vmatrix} = 6 \end{aligned}$$

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

We can also verify that $A \cdot A^{-1} = A^{-1} \cdot A = I$

Derivative of a Matrix of functions

Suppose that

$$A(t) = [a_{ij}(t)]_{m \times n}$$

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 $A(t) = (a_{ij}(t))_{m \times n}$ 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

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

Exercise

Write the given sum as a single column matrix

- $3t \begin{pmatrix} 2 \\ t \\ -1 \end{pmatrix} + (t-1) \begin{pmatrix} -1 \\ -t \\ 3 \end{pmatrix} - 2 \begin{pmatrix} 3t \\ 4 \\ -5t \end{pmatrix}$
- $\begin{pmatrix} 1 & -3 & 4 \\ 2 & 5 & -1 \\ 0 & -4 & -2 \end{pmatrix} \begin{pmatrix} t \\ 2t-1 \\ -t \end{pmatrix} + \begin{pmatrix} -t \\ 1 \\ 4 \end{pmatrix} - \begin{pmatrix} 2 \\ 8 \\ -6 \end{pmatrix}$

Determine whether the given matrix is singular or non-singular. If singular, find A^{-1} .

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

Find $\frac{dX}{dt}$

- $X = \begin{pmatrix} \frac{1}{2} \sin 2t - 4 \cos 2t \\ -3 \sin 2t + 5 \cos 2t \end{pmatrix}$
- 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) $\int_0^t A(s) ds$.
- 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 $\mathbf{R}^2$ (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 $\mathbf{R}^3$ (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 $x_1, x_2, \dots, x_n$ can be expressed in the form

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

where $a_1, a_2, \dots, a_n$ 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 $a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$ (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

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

(2) The following equations are also linear

$$\begin{aligned} x + 3y &= 7 & x_1 - 2x_2 - 3x_3 + x_4 &= 0 \\ \frac{1}{2}x - y + 3z &= -1 & x_1 + x_2 + \dots + x_n &= 1 \end{aligned}$$

(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.

Consistent and inconsistent system

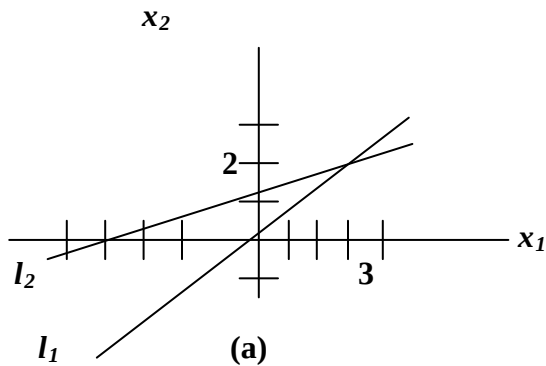
A linear system is said to be **consistent** if it has at least one solution and it is called **inconsistent** if it has no solutions.

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

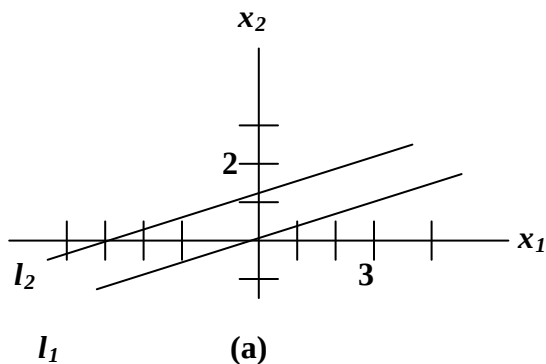


This system has exactly one solution

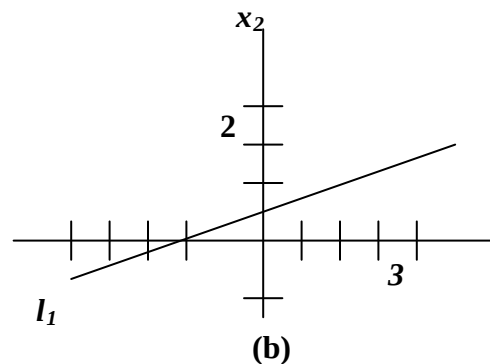
See the graphs to the following linear systems:

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

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



(a) No solution.



(b) Infinitely many solutions.

Linear System with Three Unknowns

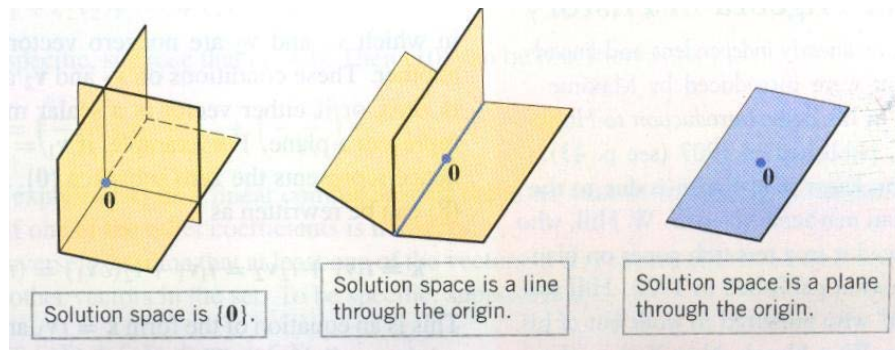
Consider a linear system of three equations in three unknowns:

$$a_1x + b_1y + c_1z = d_1$$

$$a_2x + b_2y + c_2z = d_2$$

$$a_3x + b_3y + c_3z = d_3$$

In this case, the graph of each equation is a plane, so the solutions of the system, if any, correspond to points where all three planes intersect; and again we see that there are only three possibilities – no solutions, one solution, or infinitely many solutions as shown in figure.



Theorem 1 Every system of linear equations has zero, one or infinitely many solutions; there are no other possibilities.

Example 1 Solve the linear system

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

Solution

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

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

Example 2 Solve the linear system

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

Solution

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

$$0 = 6$$

This equation is contradictory.

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

Example 3 Solve the linear system
$$\begin{aligned} 4x - 2y &= 1 \\ 16x - 8y &= 4 \end{aligned}$$

Solution

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

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

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 $4x - 2y = 1$)

$$\begin{aligned} 4x - 2t &= 1, \quad y = t \\ x &= \frac{1}{4} + \frac{1}{2}t, \quad y = t \end{aligned}$$

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

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

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**.

$$x_1 - 2x_2 + x_3 = 0$$

Given the system

$$2x_2 - 8x_3 = 8$$

$$-4x_1 + 5x_2 + 9x_3 = -9$$

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:

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

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

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

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

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

Next, multiply equation 2 by $\frac{1}{2}$ in order to obtain 1 as the coefficient for x_2

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

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

The new system has a triangular form

$$\begin{array}{rcl} x_1 - 2x_2 + x_3 & = & 0 \\ x_2 - 4x_3 & = & 4 \\ x_3 & = & 3 \end{array} \quad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

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

$$\begin{array}{rcl} x_1 - 2x_2 & = & -3 \\ x_2 & = & 16 \\ x_3 & = & 3 \end{array} \quad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

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

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

This completes the solution.

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

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

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

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 1st row in 3rd row in 1st matrix.

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

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

Two Fundamental Questions

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

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

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

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

Solution

First obtain the triangular matrix by removing x_1 and x_2 term from third equation and removing x_2 from second equation.

First divide the second equation by 2 we get

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

Now multiply equation 1 with 4 and add in equation 3 to eliminate x_1 from third equation.

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

Now multiply equation 2 with 3 and add in equation 3 to eliminate x_2 from third equation.

$$\begin{array}{rcl} x_1 - 2x_2 + x_3 & = & 0 \\ x_2 - 4x_3 & = & 4 \\ x_3 & = & 3 \end{array} \quad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Put value of x_3 in second equation we get

$$x_2 - 4(3) = 4$$

$$x_2 = 16$$

Now put these values of x_2 and x_3 in first equation we get

$$x_1 - 2(16) + 3 = 0$$

$$x_1 = 29$$

So a solution exists and the system is consistent and has a unique solution.

Example 7 Solve if the following system of linear equations is consistent.

$$\begin{array}{rcl} x_2 - 4x_3 & = & 8 \\ 2x_1 - 3x_2 + 2x_3 & = & 1 \\ 5x_1 - 8x_2 + 7x_3 & = & 1 \end{array}$$

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 $1/2$ 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:

$$2x_1 - 3x_2 + 2x_3 = 1$$

$$x_2 - 4x_3 = 8$$

$$0 = 2.5$$

There are no values of x_1, x_2, x_3 that will satisfy because the equation $0 = 2.5$ is never true.

Hence original **system is inconsistent (i.e., has no solution)**.

Exercises

1. State in words the next elementary “row” operation that should be performed on the system in order to solve it. (More than one answer is possible in (a).)

$$a. \quad x_1 + 4x_2 - 2x_3 + 8x_4 = 12$$

$$x_2 - 7x_3 + 2x_4 = -4$$

$$5x_3 - x_4 = 7$$

$$x_3 + 3x_4 = -5$$

$$b. \quad x_1 - 3x_2 + 5x_3 - 2x_4 = 0$$

$$x_2 + 8x_3 = -4$$

$$2x_3 = 7$$

$$x_4 = 1$$

2. The augmented matrix of a linear system has been transformed by row operations into the form below. Determine if the system is consistent.

$$\left[\begin{array}{cccc} 1 & 5 & 2 & -6 \\ 0 & 4 & -7 & 2 \\ 0 & 0 & 5 & 0 \end{array} \right]$$

3. Is $(3, 4, -2)$ a solution of the following system?

$$5x_1 - x_2 + 2x_3 = 7$$

$$-2x_1 + 6x_2 + 9x_3 = 0$$

$$-7x_1 + 5x_2 - 3x_3 = -7$$

4. For what values of h and k is the following system consistent?

$$2x_1 - x_2 = h$$

$$-6x_1 + 3x_2 = k$$

Solve the systems in the exercises given below;

$$x_2 + 5x_3 = -4$$

$$5. \quad x_1 + 4x_2 + 3x_3 = -2$$

$$2x_1 + 7x_2 + x_3 = -1$$

6.

$$x_1 - 5x_2 + 4x_3 = -3$$

$$2x_1 - 7x_2 + 3x_3 = -2$$

$$2x_1 - x_2 - 7x_3 = 1$$

$$7. \quad x_1 + 2x_2 = 4$$

$$x_1 - 3x_2 - 3x_3 = 2$$

$$x_2 + x_3 = 0$$

8.

$$2x_1 - 4x_3 = -10$$

$$x_2 + 3x_3 = 2$$

$$3x_1 + 5x_2 + 8x_3 = -6$$

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 k 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).

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

$$2. \begin{bmatrix} \circ & * & * & * \\ 0 & \circ & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

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

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

$$3. \begin{bmatrix} 0 & \circ & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \circ & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \circ & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \circ & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \circ & * \end{bmatrix}$$

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

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.

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

↓
Pivot Column

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

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

Pivot
Next pivot column

$R_1 + R_2$
 $2R_1 + R_3$

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

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

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

$$\begin{array}{ccccc} & & & \text{Pivot} & \\ \left[\begin{array}{ccccc} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] & \text{General form} & \left[\begin{array}{ccccc} \circ & * & * & * & * \\ 0 & \circ & * & * & * \\ 0 & 0 & 0 & \circ & * \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \\ \text{Pivot column} & & \end{array}$$

This is in echelon form and thus columns 1, 2, and 4 of **A** are pivot columns.

$$\begin{array}{ccccc} & & & \text{Pivot positions} & \\ \left[\begin{array}{ccccc} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{array} \right] & & \text{Pivot columns} & & \end{array}$$

Pivot element

A pivot is a nonzero number in a pivot position that is used as needed to create zeros via row operations

The Row Reduction Algorithm consists of four steps, and it produces a matrix in echelon form. A fifth step produces a matrix in reduced echelon form.

The algorithm is explained by an example.

Example 3 Apply elementary row operations to transform the following matrix first into echelon form and then into reduced echelon form.

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

Solution

STEP 1 Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

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

└ Pivot column

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.)

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

└ Pivot

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

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

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

└ Pivot

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.

$$\left[\begin{array}{cc|cc|cc} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array} \right]$$

Pivot
Next pivot column

According to step 3 “All entries in a column below a leading entry are zero”. For this subtract $\frac{3}{2}$ time R_2 from R_3

$$\left[\begin{array}{cc|cc|cc} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array} \right] R_3 - \frac{3}{2}R_2$$

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:

$$\left[\begin{array}{cc|cc|cc} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array} \right]$$

Pivot

This is the Echelon form of the matrix.

To change it in reduced echelon form we need to do one more step:

STEP 5 Make the leading entry in each nonzero row 1. Make all other entries of that column to 0.

Divide first Row by 3 and 2nd Row by 2

$$\left[\begin{array}{cc|cc|cc} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array} \right] \frac{1}{2}R_2, \frac{1}{3}R_1$$

Multiply second row by 3 and then add in first row.

$$\left[\begin{array}{cc|cc|cc} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array} \right] 3R_2 + R_1$$

Subtract row 3 from row 2, and multiply row 3 by 5 and then subtract it from first row

$$\left[\begin{array}{cccccc} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array} \right] \begin{array}{l} R_2 - R_3 \\ R_1 - 5R_3 \end{array}$$

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

$$\left[\begin{array}{cccc} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

There are three variables because the augmented matrix has four columns. The associated system of equations is

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

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.)

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

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 “ $\sim$ ” before a matrix indicates that the matrix is row equivalent to the preceding matrix.

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

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

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

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:

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

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

Exercise

- Find the general solution of the linear system whose augmented matrix is

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

- 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

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

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

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

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

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

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

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

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

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

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

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.

$$15. \begin{aligned} x_1 + hx_2 &= 1 \\ 2x_1 + 3x_2 &= k \end{aligned}$$

$$16. \begin{aligned} x_1 - 3x_2 &= 1 \\ 2x_1 + hx_2 &= k \end{aligned}$$

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”.

$$\text{e.g. } u = [3 \ -1]^T = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \quad v = [2 \ 3 \ 5]^T = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, \quad w = [w_1 \ w_2 \ w_3 \ w_4]^T \text{ are all}$$

column vectors or simply vectors.

Vectors in $\mathbb{R}^2$

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

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

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

Similarly $\mathbb{R}^3$ and $\mathbb{R}^4$ contain all vectors with three and four entries respectively. The entries of the vectors are always taken from the set of real numbers $\mathbb{R}$. 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 $\mathbb{R}^2$ are equal if and only if their corresponding entries are equal.

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

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

Note In fact, vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$ 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 $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, their sum is the vector $\mathbf{u} + \mathbf{v}$ obtained by adding corresponding entries of the vectors $\mathbf{u}$ and $\mathbf{v}$, which is again a vector in $\mathbb{R}^2$

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

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

Scalar Multiplication of a vector

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

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

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

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

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

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

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

Geometric Descriptions of R^2

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

See Figure 1.

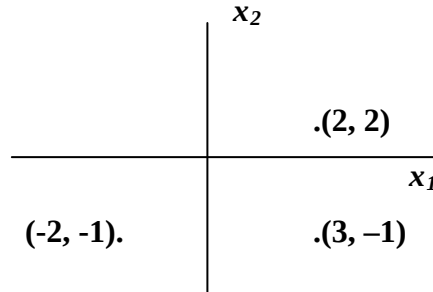


Figure 1 **Vectors as points.**

Vectors in R^3

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

Vectors in R^n

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

$$u = [u_1 \ u_2 \ \cdots u_n]^T$$

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 R^n and all scalars c and d :

- (i) $u + v = v + u$ (Commutative)
- (ii) $(u + v) + w = u + (v + w)$ (Associative)
- (iii) $u + 0 = 0 + u = u$ (Additive Identity)
- (iv) $u + (-u) = (-u) + u = 0$ (Additive Inverse)
where $-u$ denotes $(-1)u$

- (v) $c(u + v) = cu + cv$ (Scalar Distribution over Vector Addition)
- (vi) $(c + d)u = cu + du$ (Vector Distribution over Scalar Addition)
- (vii) $c(du) = (cd)u$
- (viii) $1u = u$

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

$$y = c_1v_1 + c_2v_2 + \dots + c_pv_p$$

is called a **linear combination** of $v_1, \dots, v_p$ using weights $c_1, \dots, c_p$.

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 $(3, 5, 2) = 3(1, 0, 0) + 5(0, 1, 0) + 2(0, 0, 1)$

$$(3, 5, 2) = 3v_1 + 5v_2 + 2v_3 \text{ where } v_1 = (1, 0, 0), v_2 = (0, 1, 0), v_3 = (0, 0, 1)$$

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 $b = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$.

Determine whether $\mathbf{b}$ can be generated (or written) as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$.
That is, determine whether weights x_1 and x_2 exist such that

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 = \mathbf{b} \quad (1)$$

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

$$\begin{aligned}
 & x_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \\
 & \quad \quad \quad \mathbf{a}_1 \quad \quad \mathbf{a}_2 \quad \quad \mathbf{b} \\
 \Rightarrow & \begin{bmatrix} x_1 \\ -2x_1 \\ -5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \\
 \Rightarrow & \begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \quad (2) \\
 \Rightarrow & \begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= -3 \end{aligned} \quad (3)
 \end{aligned}$$

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

$$\begin{aligned}
 & \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \\
 & \text{By } R_2 + 2R_1; R_3 + 5R_1 \\
 & \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \\
 & \text{By } \left(\frac{1}{9}\right)R_2; \left(\frac{1}{16}\right)R_3
 \end{aligned}$$

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

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 $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$, with weights $x_1 = 3$ and $x_2 = 2$.

Spanning Set

If $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in $\mathbf{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_p$ is denoted by $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and is called the **subset of $\mathbf{R}^n$ spanned** (or **generated**) by $\mathbf{v}_1, \dots, \mathbf{v}_p$. That is, $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form of $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_p\mathbf{v}_p$, with $c_1, \dots, c_p$ scalars.

If we want to check whether a vector $\mathbf{b}$ is in $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ then we will see whether the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_p\mathbf{v}_p = \mathbf{b} \text{ has a solution, or}$$

Equivalently, whether the linear system with augmented matrix $[\mathbf{v}_1, \dots, \mathbf{v}_p \quad \mathbf{b}]$ has a solution.

Note

(1) The set $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ contains every scalar multiple of $\mathbf{v}_1$

because $c\mathbf{v}_1 = c\mathbf{v}_1 + 0\mathbf{v}_2 + \dots + 0\mathbf{v}_p$ i.e every $c\mathbf{v}_i$ can be written as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_p$

(2) Zero vector $= \mathbf{0} \in \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ as $\mathbf{0}$ can be written as the linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ that is $\mathbf{0}_v = 0_F\mathbf{v}_1 + 0_F\mathbf{v}_2 + \dots + 0_F\mathbf{v}_n$ here for the convenience it is mentioned that $\mathbf{0}_v$ is the vector (zero vector) while 0_F is zero scalar (weight of all $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$) and in particular not to make confusion that $\mathbf{0}_v$ and 0_F are same!

A Geometric Description of Span $\{v\}$ and Span $\{u, v\}$

Let v be a nonzero vector in $\mathbf{R}^3$. 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 $\mathbf{R}^3$ through v and 0 .

If u and v are nonzero vectors in $\mathbf{R}^3$, with v not a multiple of u , then Span $\{u, v\}$ is the plane in $\mathbf{R}^3$ that contains u , v and 0 . In particular, Span $\{u, v\}$ contains the line in $\mathbf{R}^3$ 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 $\{a_1, a_2\}$ is a plane through the origin in $\mathbf{R}^3$. Does b lie in that plane?

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

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

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

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

$\therefore b \notin \text{Span}\{a_1, a_2\}$ 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:

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

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.

- What economic interpretation can be given to the vector $100b$?
- 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

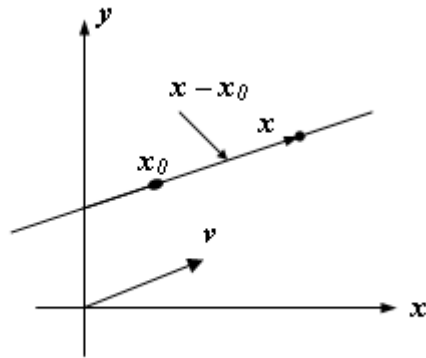
(a) We have $100b = 100 \begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} = \begin{bmatrix} 45 \\ 25 \\ 15 \end{bmatrix}$

The vector $100b$ 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.

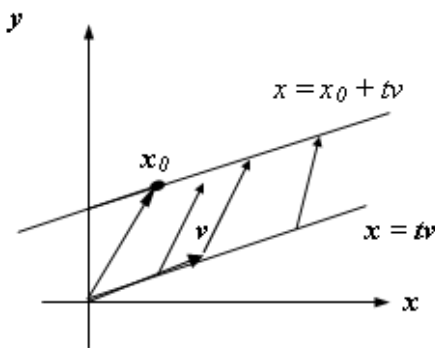
- The costs of manufacturing x_1 dollars worth of B are given by the vector x_1b and the costs of manufacturing x_2 dollars worth of C are given by x_2c . Hence the total costs for both products are given by the vector $x_1b + x_2c$.

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 $\mathbf{x} - \mathbf{x}_0 = t\mathbf{v}$ for some scalar t .

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

$$\mathbf{x} - \mathbf{x}_0 = t\mathbf{v} \text{ -----(1)} \quad (-\infty < t < +\infty)$$

This is a **vector equation of the line** through $\mathbf{x}_0$ and parallel to $\mathbf{v}$.

In the special case, where $\mathbf{x}_0 = \mathbf{0}$, the line passes through the origin, it simplifies to

$$\mathbf{x} = t\mathbf{v} \quad (-\infty < t < +\infty)$$

Parametric Equations of a Line in $\mathbb{R}^2$

Let $\mathbf{x} = (x, y) \in \mathbb{R}^2$ be a general point of the line through $\mathbf{x}_0 = (x_0, y_0) \in \mathbb{R}^2$ which is parallel to

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

$$\begin{aligned}(x, y) - (x_0, y_0) &= t(a, b) \quad (-\infty < t < +\infty) \\ \Rightarrow (x - x_0, y - y_0) &= (ta, tb) \quad (-\infty < t < +\infty) \\ \Rightarrow x &= x_0 + at, \quad y = y_0 + bt \quad (-\infty < t < +\infty)\end{aligned}$$

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

Parametric Equations of a Line in R^3

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

$\mathbf{x}_0 = (x_0, y_0, z_0) \in R^3$ that is parallel to $\mathbf{v} = (a, b, c) \in R^3$, then again eq. 1 takes the form

$$\begin{aligned}(x, y, z) &= (x_0, y_0, z_0) + t(a, b, c) \quad (-\infty < t < +\infty) \\ \Rightarrow x &= x_0 + at, \quad y = y_0 + bt, \quad z = z_0 + ct \quad (-\infty < t < +\infty)\end{aligned}$$

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

Example 8

- Find a vector equation and parametric equations of the line in R^2 that passes through the origin and is parallel to the vector $\mathbf{v} = (-2, 3)$.
- Find a vector equation and parametric equations of the line in R^3 that passes through the point $P_0(1, 2, -3)$ and is parallel to the vector $\mathbf{v} = (4, -5, 1)$.
- Use the vector equation obtained in part (b) to find two points on the line that are different from P_0 .

Solution

- We know that a vector equation of the line passing through origin is $\mathbf{x} = t\mathbf{v}$.

Let $\mathbf{x} = (x, y)$. 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

$$x = -2t, \quad y = 3t$$

(b) The vector equation of the line is $\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}$.

Let $\mathbf{x} = (x, y, z)$, Here $\mathbf{x}_0 = (1, 2, -3)$ and $\mathbf{v} = (4, -5, 1)$, 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 $\mathbf{x}_0$ be a fixed point on the required plane W and $\mathbf{v}_1$ and $\mathbf{v}_2$ be two nonzero vectors that are parallel to W and are not scalar multiples of one another. If $\mathbf{x}$ is any variable point in the plane W . Suppose $\mathbf{v}_1$ and $\mathbf{v}_2$ have their initial points at $\mathbf{x}_0$, we can create a parallelogram with adjacent side's $t_1\mathbf{v}_1$ and $t_2\mathbf{v}_2$ in which $\mathbf{x} - \mathbf{x}_0$ is the diagonal given by the sum

$$\mathbf{x} - \mathbf{x}_0 = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

or, equivalently, $\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{v}_1 + t_2\mathbf{v}_2$ -----(1)

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

This is a **vector equation of the plane** through $\mathbf{x}_0$ and parallel to the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. In the special case where $\mathbf{x}_0 = \mathbf{0}$, then vector equation of the plane passes through the origin takes the form

$$\mathbf{x} = t_1 \mathbf{v}_1 + t_2 \mathbf{v}_2 \quad (-\infty < t_1 < +\infty, -\infty < t_2 < +\infty)$$

Parametric Equations of a Plane

Let $\mathbf{x} = (x, y, z)$ be a general or variable point in the plane passes through a fixed point $\mathbf{x}_0 = (x_0, y_0, z_0)$ and parallel to the vectors $\mathbf{v}_1 = (a_1, b_1, c_1)$ and $\mathbf{v}_2 = (a_2, b_2, c_2)$, 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

$$x = x_0 + a_1 t_1 + a_2 t_2$$

$$y = y_0 + b_1 t_1 + b_2 t_2 \quad (-\infty < t_1 < +\infty, -\infty < t_2 < +\infty)$$

$$z = z_0 + c_1 t_1 + c_2 t_2$$

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 $\mathbf{R}^3$ and is parallel to the vectors $\mathbf{v}_1 = (1, -2, 3)$ and $\mathbf{v}_2 = (4, 0, 5)$.
- (b) Find three points in the plane obtained in part (a).

Solution

- (a) As vector equation of the plane passing through origin is $\mathbf{x} = t_1 \mathbf{v}_1 + t_2 \mathbf{v}_2$.

Let $\mathbf{x} = (x, y, z)$ 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 :

$$t_1 = 0 \text{ and } t_2 = 0 \quad \text{produces the point } (0, 0, 0)$$

$$t_1 = -2 \text{ and } t_2 = 1 \quad \text{produces the point } (2, 4, -1)$$

$$t_1 = \frac{1}{2} \text{ and } t_2 = \frac{1}{2} \quad \text{produces the point } (5/2, -1, 4)$$

Vector equation of Plane through Three Points

If x_0, x_1 and x_2 are three non collinear points in the required plane, then, obviously, the vectors $\mathbf{v}_1 = \mathbf{x}_1 - \mathbf{x}_0$ and $\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{x}_0$ are parallel to the plane. So, a vector equation of the plane is

$$\mathbf{x} = \mathbf{x}_0 + t_1(\mathbf{x}_1 - \mathbf{x}_0) + t_2(\mathbf{x}_2 - \mathbf{x}_0)$$

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

Solution

Let $\mathbf{x} = (x, y, z)$, and if we take $\mathbf{x}_0, \mathbf{x}_1$ and $\mathbf{x}_2$ to be the points P, Q and R respectively, then $\mathbf{x}_1 - \mathbf{x}_0 = \overrightarrow{PQ} = (-3, 8, -8)$ and $\mathbf{x}_2 - \mathbf{x}_0 = \overrightarrow{PR} = (-1, 14, -12)$

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

$$\begin{aligned}(x, y, z) &= (4 + 5t_1 - t_2, 2 - t_1 + 8t_2, t_1 + t_2) \\ \Rightarrow (x, y, z) &= (4, 2, 0) + (5t_1, -t_1, t_1) + (-t_2, 8t_2, t_2) \\ \Rightarrow (x, y, z) &= (4, 2, 0) + t_1(5, -1, 1) + t_2(-1, 8, 1)\end{aligned}$$

This is a vector equation of the plane that passes through the point $(4, 2, 0)$ and is parallel to the vectors $\mathbf{v}_1 = (5, -1, 1)$ and $\mathbf{v}_2 = (-1, 8, 1)$.

Finding Parametric Equations from a General Equation

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

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

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

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

Let $y = t_1$ and $z = t_2$

Then the parametric equations of the given plane are

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

Exercises

1. Prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for any $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^n$.
2. For what value(s) of h , $\mathbf{y}$ belongs to $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? Where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \text{ and } \mathbf{y} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

3. Determine whether $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$.

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

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

4. Determine if $\mathbf{b}$ is a linear combination of the vectors formed from the columns of the matrix $\mathbf{A}$.

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

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

In exercises 7-10, list seven vectors in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. For each vector, show that the weights on $\mathbf{v}_1$ and $\mathbf{v}_2$ used to generate the vector and list the three entries of the vector. Give also geometric description of the $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

$$7. \quad \mathbf{v}_1 = \begin{pmatrix} 5 \\ -1 \\ 3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix}$$

$$8. \quad \mathbf{v}_1 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

$$9. \quad \mathbf{v}_1 = \begin{pmatrix} 2 \\ 6 \\ -4 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -3 \\ -9 \\ 6 \end{pmatrix}$$

$$10. \quad \mathbf{v}_1 = \begin{pmatrix} -3.7 \\ -0.4 \\ 11.2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 5.8 \\ 2.1 \\ 5.3 \end{pmatrix}$$

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 a_1 and 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 v_1 and 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 $\text{Span}\{u, v\}$ for all h and k .

Lecture 6

Matrix Equations

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

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

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

Note that Ax is defined only if the number of columns of A equals the number of entries in x .

Example 1

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

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

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

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

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

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

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

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 $A\mathbf{x} = \mathbf{b}$, and we shall call such an equation a **matrix equation**, to distinguish it from a vector equation.

Theorem 1 If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ and if $\mathbf{b}$ is in $\mathbf{R}^m$, the matrix equation $A\mathbf{x} = \mathbf{b}$ has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is $[a_1 \ a_2 \ \dots \ a_n \ b]$

Existence of Solutions The equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of A .

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

Is the equation $A\mathbf{x} = \mathbf{b}$ consistent for all possible b_1, b_2, b_3 ?

Solution Row reduce the augmented matrix for $A\mathbf{x} = \mathbf{b}$:

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

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

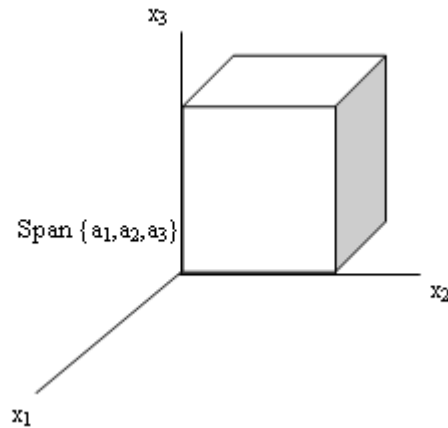
$$\sim \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{bmatrix}$$

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

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

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

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



The equation $\mathbf{Ax} = \mathbf{b}$ fails to be consistent for all $\mathbf{b}$ because the echelon form of $\mathbf{A}$ has a row of zeros. If $\mathbf{A}$ 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

$$\begin{aligned} \text{(a)} \quad \begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix} &= aA + bB + cC \\ &= a \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + b \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + c \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 4a + b & -b + 2c \\ -2a + 2b + c & -2a + 3b + 4c \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad 4a + b &= 6 & (1) \\ -b + 2c &= -8 & (2) \\ -2a + 2b + c &= -1 & (3) \\ -2a + 3b + 4c &= -8 & (4) \end{aligned}$$

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

$$-b - 3c = 7 \quad (5)$$

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

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

$$\begin{aligned} \text{(b)} \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} &= aA + bB + cC \\ &= a \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + b \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + c \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 4a + b & -b + 2c \\ -2a + 2b + c & -2a + 3b + 4c \end{bmatrix} \end{aligned}$$

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

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

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

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

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

$$b + 3c = 0 \quad (5)$$

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

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

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

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

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

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

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

$$\text{(c)} \quad \begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix} = aA + bB + cC$$

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

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

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

$$-2a + 2b + c = 3 \quad (3)$$

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

Subtracting (4) from (3), we obtain

$$-b - 3c = -5 \quad (5)$$

Subtracting (5) from (2):

$$5c = 5 \Rightarrow c = 1$$

$$\text{From (2),} \quad -b + 2(1) = 0 \Rightarrow b = 2$$

$$\text{From (3),} \quad -2a + 2(2) + 1 = 3 \Rightarrow a = 1$$

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**.

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

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

- (a) For each **b** in $\mathbf{R}^m$, the equation $\mathbf{Ax} = \mathbf{b}$ has a solution.
- (b) The columns of **A** Span $\mathbf{R}^m$.
- (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,

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

Note

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

Examples

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

- (a) $v_1 = (2, 2, 2)$, $v_2 = (0, 0, 3)$,
 $v_3 = (0, 1, 1)$
- (b) $v_1 = (3, 1, 4)$, $v_2 = (2, -3, 5)$,
 $v_3 = (5, -2, 9)$, $v_4 = (1, 4, -1)$
- (c) $v_1 = (1, 2, 6)$, $v_2 = (3, 4, 1)$,
 $v_3 = (4, 3, 1)$, $v_4 = (3, 3, 1)$

Solutions

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

Expressing this in terms of components given by

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

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

$$2k_1 + 0k_2 + 0k_3 = b_1$$

$$2k_1 + 0k_2 + k_3 = b_2$$

$$2k_1 + 3k_2 + k_3 = b_3$$

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

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

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

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

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

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

$$1c_2 + 1c_3 + -1c_4 = \left(\frac{-4}{17}\right)d_2 + \left(\frac{1}{17}\right)d_3 \quad \dots\dots\dots(2)$$

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

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

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

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

2. Let $A = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}$, $u = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$, and $v = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. 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 $\mathbb{R}^3$ 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 $\mathbb{R}^3$ 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 $\mathbb{R}^3$?

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 $\mathbb{R}^4$?

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

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

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 $\mathbb{R}^4$?

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 $\mathbb{R}^3$?

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 $\mathbb{R}^4$.

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 $x_1, x_2, \dots, x_n$ 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 R^m .

Trivial Solution

A homogeneous system $Ax = 0$ always has at least one solution, namely, $x = 0$ (the zero vector in 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 $x \neq 0$ 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

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

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

$$6x_1 + x_2 - 8x_3 = 0$$

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

$$\begin{aligned} & \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} & -1R_1 + R_2, -2R_1 + R_3 \\ & \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & -3R_2 + R_3 \\ & \sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & 1/3R_1, 1/3R_2, 5/3R_2 + R_1 \\ & \sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & (-1)R_2 \\ & \begin{array}{lcl} x_1 & -\frac{4}{3}x_3 & = 0 \\ x_2 & & = 0 \\ 0 & & = 0 \end{array} \end{aligned}$$

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

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

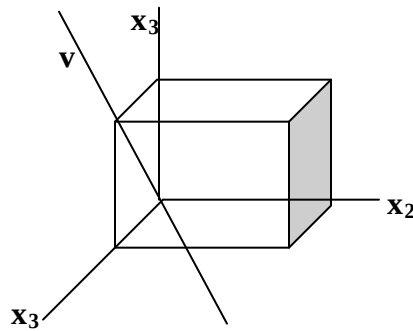
As a vector, the general solution of $A\mathbf{x} = \mathbf{0}$ is given by:

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

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

Geometric Interpretation

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



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

Example 2

Solve the following system

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

Solution

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

$x_1 = 0.3x_2 + 0.2x_3$ where x_2 and x_3 free variables.

As a vector, the general solution is:

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

$$= x_2 \begin{bmatrix} 0.3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0.2 \\ 0 \\ 1 \end{bmatrix} \quad (2)$$

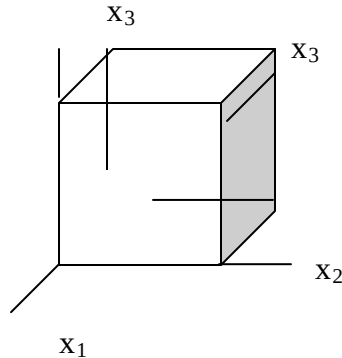
$\downarrow$
 $\mathbf{u}$

$\downarrow$
 $\mathbf{v}$

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

Geometric Interpretation

Since neither $\mathbf{u}$ nor $\mathbf{v}$ 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 $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ for suitable vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ (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 **$\text{Span}\{\mathbf{0}\}$** .

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:

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

The augmented matrix is:

$$\begin{aligned}
 & \begin{bmatrix} 1 & 3 & 1 & 0 \\ -4 & -9 & 2 & 0 \\ 0 & -3 & -6 & 0 \end{bmatrix} \\
 & \sim \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 3 & 6 & 0 \\ 0 & -3 & -6 & 0 \end{bmatrix} & 4R_1 + R_2, \\
 & \sim \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 3 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & R_2 + R_3 \\
 & \sim \begin{bmatrix} 1 & 0 & -5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \frac{1}{2}R_2, (-3)R_2 + R_1 \\
 & \text{SO}
 \end{aligned}$$

$$x_1 - 5x_3 = 0$$

$$x_2 + 2x_3 = 0$$

$$0 = 0$$

From above results, it is clear that x_3 is a free variable, so $A\mathbf{x} = \mathbf{0}$ has nontrivial solutions (one for each choice of x_3).

From above equations we have,

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

As a vector, the general solution of $A\mathbf{x} = \mathbf{0}$ is given by

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

Parametric Vector Form of the solution

Whenever a solution set is described explicitly with vectors, we say that the solution is in parametric vector form

The equation

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

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 $\mathbf{x} = x_3\mathbf{v}$ (with x_3 free), or $\mathbf{x} = t\mathbf{v}$ (with t in $\mathbf{R}$), 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 $A\mathbf{x} = \mathbf{b}$, where

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

Solution

Row operations on $[A \ \mathbf{b}]$ produce

$$\begin{aligned} & \begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix} \\ & \sim \begin{bmatrix} 3 & 5 & -4 & 7 \\ 0 & 3 & 0 & 6 \\ 0 & -9 & 0 & -18 \end{bmatrix} & R_1 + R_2, -2R_1 + R_3 \\ & \sim \begin{bmatrix} 3 & 5 & -4 & 7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} & 3R_2 + R_3, \frac{1}{3}R_2 \\ & \sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} & -5R_2 + R_1, \frac{1}{3}R_1 \\ & \begin{matrix} x_1 & -\frac{4}{3}x_3 = -1 \\ \text{OR} & x_2 & = 2 \\ & 0 & = 0 \end{matrix} \end{aligned}$$

Thus $x_1 = -1 + \frac{4}{3}x_3$, $x_2 = 2$, and x_3 is free.

As a vector, the general solution of $A\mathbf{x} = \mathbf{b}$ has the form

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

$\underbrace{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}}_{\mathbf{p}} + x_3 \underbrace{\begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}}$

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

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

Note

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

$$x = t\mathbf{v} \quad (t \text{ in } R) \quad (4)$$

With the same $\mathbf{v}$ that appears in equation (3) in above example.

Thus the solutions of $Ax = \mathbf{b}$ are obtained by adding the vector $\mathbf{p}$ to the solutions of $Ax = \mathbf{0}$. The vector $\mathbf{p}$ itself is just one particular solution of $Ax = \mathbf{b}$ (corresponding to $t = 0$ in (3)).

The following theorem gives the precise statement.

Theorem

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

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

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

The augmented matrix is

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

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

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

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

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

SO

$$x_1 - 5x_3 = -2$$

$$x_2 + 2x_3 = 1$$

$$0 = 0$$

Thus $x_1 = -2 + 5x_3$, $x_2 = 1 - 2x_3$, and x_3 is free.

As a vector, the general solution of $\mathbf{Ax} = \mathbf{b}$ has the form

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 + 5x_3 \\ 1 - 2x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 5x_3 \\ -2x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$$

$\downarrow$
 $\mathbf{p}$

$\downarrow$
 $\mathbf{v}$

So we can write solution set in parametric vector form as

$$x = \mathbf{p} + x_3 \mathbf{v}$$

**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 $\mathbf{x}$ as a vector whose entries depend on the free variables if any.
- Step 4:
Decompose $\mathbf{x}$ 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 $A\mathbf{x} = \mathbf{0}$ 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. $[1 \quad -5 \quad 0 \quad 4]$

9. Describe the solution set in $\mathbb{R}^3$ 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 $\mathbf{a}$ parallel to $\mathbf{b}$.

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

11. Find a parametric equation of the line M through $\mathbf{p}$ and $\mathbf{q}$.

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

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

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

Lecture 8

Linear Independence

Definition

An indexed set of vectors $\{v_1, v_2, \dots, v_p\}$ in R^n is said to be **linearly independent** if the vector equation $x_1v_1 + x_2v_2 + \dots + x_pv_p = 0$ has only the trivial solution.

The set $\{v_1, v_2, \dots, v_p\}$ is said to be **linearly dependent** if there exist weights $c_1, \dots, c_p$, not all zero, such that $c_1v_1 + c_2v_2 + \dots + c_pv_p = 0$ (1)

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

Example 1

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

(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

$$\begin{aligned} & \begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix} \\ & \sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & -6 & -6 & 0 \end{bmatrix} \quad (-2)R_1 + R_2, (-3)R_1 + R_3 \\ & \sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad R_2 + R_3 \end{aligned} \quad (2)$$

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:

$$\begin{aligned} & \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{-\frac{1}{3}R_2} \\ & \sim \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 - 4R_2} \\ & \Rightarrow \begin{array}{lcl} x_1 & & -2x_3 = 0 \\ & x_2 & +x_3 = 0 \\ & & 0 = 0 \end{array} \end{aligned}$$

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 $x_1v_1 + x_2v_2 + x_3v_3 = 0$

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

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

Example (for practice)

Check whether the vectors are linearly dependent or linearly independent.

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

Solution

Consider two constants C_1 and C_2 . Suppose:

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

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

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

$$3c_1 - 2c_2 = 0$$

$$-c_1 + 2c_2 = 0$$

Solving this system gives the following solution,

$$c_1 = 0 \quad c_2 = 0$$

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 $A = [a_1 \ \dots \ a_n]$ instead of a set of vectors. The matrix equation $Ax = 0$ can be written as $x_1a_1 + x_2a_2 + \dots + x_na_n = 0$

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

Thus we have the following important fact.

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

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

independent.

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

$$\begin{array}{l} \begin{bmatrix} 0 & 1 & 4 & 0 \\ 1 & 2 & -1 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix} \\ \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix} \quad R_{12} \\ \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & -2 & 5 & 0 \end{bmatrix} \quad (-5)R_1 + R_3 \\ \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 13 & 0 \end{bmatrix} \quad (2)R_2 + R_1 \end{array}$$

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, $\mathbf{v}$) is linearly independent if and only if $\mathbf{v}$ is not the zero vector. This is because the vector equation $x_1\mathbf{v} = 0$ has only the trivial solution when $\mathbf{v} \neq 0$. The zero vector is linearly dependent because $x_1\mathbf{0} = 0$ has many nontrivial solutions.

Example 3

Check the following sets for linearly independence and dependence.

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

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

Solution

a) Notice that v_2 is a multiple of v_1 , namely, $v_2 = 2v_1$.

Hence $-2v_1 + v_2 = 0$, which shows that $\{v_1, v_2\}$ is linearly dependent.

b) v_1 and v_2 are certainly not multiples of one another. Could they be linearly dependent?

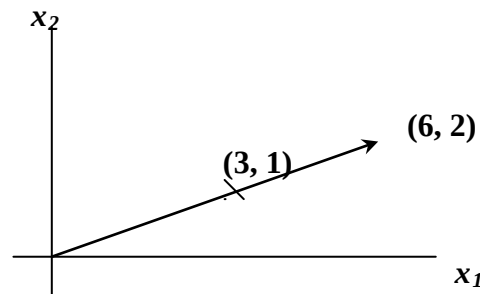
Suppose c and d satisfy $cv_1 + dv_2 = 0$

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

Thus $\{v_1, v_2\}$ is a linearly independent set.

Note A set of two vectors $\{v_1, v_2\}$ is linearly dependent if and only if one of the vectors is a multiple of the other.

In geometric terms, two vectors are linearly dependent if and only if they lie on the same line through the origin. Figure 1, shows the vectors from Example 3.



Linearly dependent

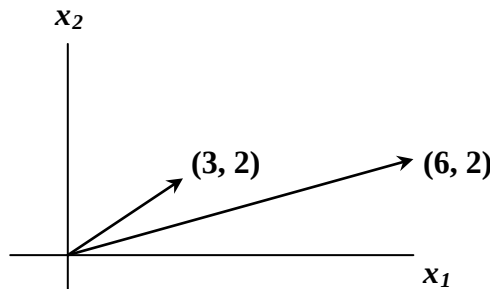


Figure 1 Linearly independent**Sets of Two or More Vectors****Theorem (Characterization of Linearly dependent Sets)**

An indexed set $s = \{v_1, v_2, \dots, v_p\}$ 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, $v_1, \dots, v_{j-1}$.

Proof

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

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

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

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

$$c_1v_1 + c_2v_2 + \dots + c_pv_p = 0$$

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

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

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

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

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.

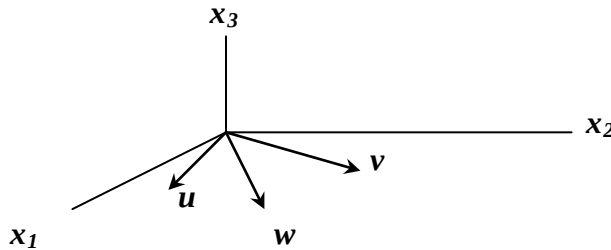
Example 4 Let $u = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Describe the set spanned by u and v , and prove that a vector w is in $\text{Span}\{u, v\}$ if and only if $\{u, v, w\}$ is linearly dependent.

Solution

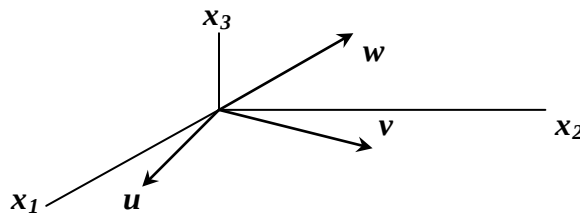
The vectors u and v are linearly independent because neither vector is a multiple of the other, nor so they span a plane in $\mathbb{R}^3$. In fact, $\text{Span}\{u, v\}$ is the x_1x_2 -plane (with $x_3 = 0$). If w is a linear combination of u and v , then $\{u, v, w\}$ is linearly dependent.

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 $\text{Span}\{u, v\}$.



Linearly dependent w in $\text{Span}\{u, v\}$.



Linearly independent w not in $\text{Span}\{u, v\}$

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

This example generalizes to any set $\{u, v, w\}$ in $\mathbb{R}^3$ 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 $\{v_1, v_2, \dots, v_p\}$ in $\mathbb{R}^n$ 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.

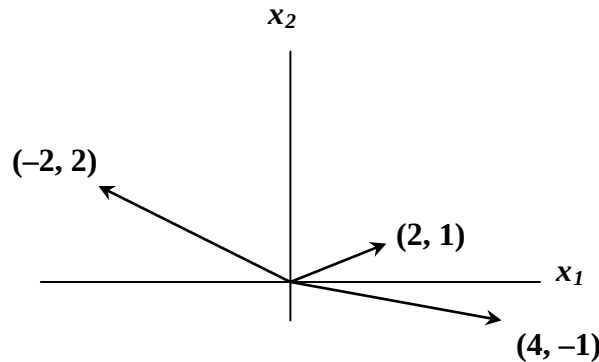


Figure 4 A linearly dependent set in $\mathbb{R}^2$.

Theorem

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

Proof

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

Then $(1)v_1 + 0v_2 + \dots + 0v_p = \mathbf{0}$ shows that 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

- The set contains four vectors that each has only three entries. So, the set is linearly dependent by the Theorem above.
- 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 $z = \begin{bmatrix} 3 \\ 7 \\ -5 \end{bmatrix}$.

- Are the sets $\{u, v\}$, $\{u, w\}$, $\{u, z\}$, $\{v, w\}$, $\{v, z\}$, and $\{w, z\}$ each linearly independent? Why or why not?
- Does the answer to Problem (i) imply that $\{u, v, w, z\}$ is linearly independent?
- 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 ?
- 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 $\text{span}\{v_1, v_2\}$ 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.

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

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

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

$$15. \text{ Given } A = \begin{bmatrix} 2 & 3 & 5 \\ -5 & 1 & -4 \\ -3 & -1 & -4 \\ 1 & 0 & 1 \end{bmatrix}, \text{ observe that the third column is the sum of the first two}$$

columns. Find a nontrivial solution of $A\mathbf{x} = \mathbf{0}$ 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 $\mathbf{v}_1, \dots, \mathbf{v}_4$ are in $\mathbb{R}^4$, and $\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly dependent.

17. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are in $\mathbb{R}^4$, and $\mathbf{v}_1$ is not a scalar multiple of $\mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent.

18. . If $\mathbf{v}_1, \dots, \mathbf{v}_4$ are in $\mathbb{R}^4$, and $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is also linearly dependent.

$$19. \text{ Use as many columns of } A = \begin{bmatrix} 8 & -3 & 0 & -7 & 2 \\ -9 & 4 & 5 & 11 & -7 \\ 6 & -2 & 2 & -4 & 4 \\ 5 & -1 & 7 & 0 & 10 \end{bmatrix} \text{ as possible to construct a}$$

matrix B with the property that equation $B\mathbf{x} = \mathbf{0}$ 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

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

and

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

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 $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$, an image vector $T(\mathbf{x})$ in $\mathbb{R}^m$.

$$T : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

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

Note

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 $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 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

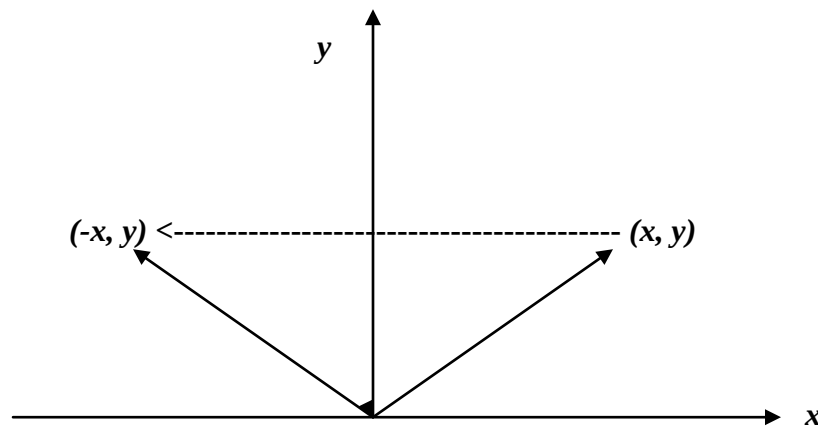
$$T(c_1u + c_2v) = c_1T(u) + c_2T(v)$$

Example 1 Consider a mapping $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(x, y) = (x, 0)$ is given as:

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

Example 2 Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $u = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $c = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$,

and define a transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$, so that

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

- Find $T(\mathbf{u})$, the image of $\mathbf{u}$ under the transformation T .
- Find an $\mathbf{x}$ in $\mathbb{R}^2$, whose image under T is $\mathbf{b}$.
- Is there more than one $\mathbf{x}$ whose image under T is $\mathbf{b}$?
- Determine if $\mathbf{c}$ is in the range of the transformation T .

Solution (a)

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

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

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

(b) We have to find an $\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{b}$ or $A\mathbf{x} = \mathbf{b}$

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

Now row reduced augmented matrix will be:

$$\begin{aligned}
& \begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} \begin{matrix} -3R_1 + R_2, \\ R_1 + R_3 \end{matrix} \\
& \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & -2 \end{bmatrix} \begin{matrix} \frac{1}{14}R_2, \\ -4R_2 + R_3, \end{matrix} \\
& \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & -0.5 \\ 0 & 0 & 0 \end{bmatrix} 3R_2 + R_1 \\
& \sim \begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -0.5 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned}$$

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

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 $\mathbf{R}^2$, 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:

$$\begin{aligned}
& \begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} \begin{matrix} -3R_1 + R_2, \\ R_1 + R_3 \end{matrix} \\
& \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & 8 \end{bmatrix} \begin{matrix} \frac{1}{4}R_3, \\ R_{23} \end{matrix} \\
& \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 14 & -7 \end{bmatrix} -14R_2 + R_3 \\
& \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}
\end{aligned}$$

$$x_1 - 3x_2 = 3$$

$$0x_1 + x_2 = 2$$

$$0x_1 + 0x_2 = -35 \Rightarrow 0 = 35 \text{ but } 0 \neq 35$$

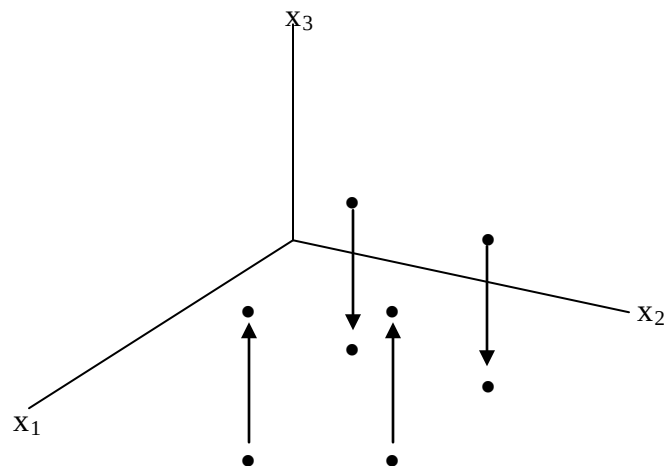
Hence the system is inconsistent. So $\mathbf{c}$ is not in the range of $\mathbf{T}$.

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 R^3

onto the x_1x_2 -coordinate plane because $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$

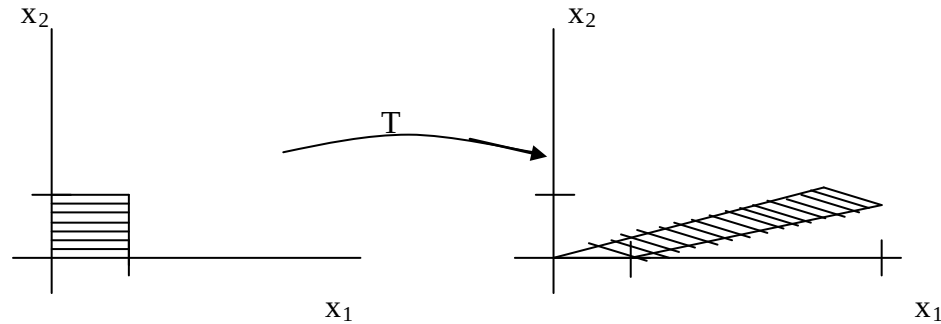


A projection transformation

Example 4 Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$, the transformation $T : R^2 \rightarrow R^2$ defined by $T(\mathbf{x}) = A\mathbf{x}$ 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 $x \rightarrow Ax$ has the properties $A(u+v) = Au + Av$ and $A(cu) = cAu$ for all u, v in $\mathbb{R}^n$ 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 $L: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be defined by $L(x, y, z) = (x, y)$.

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

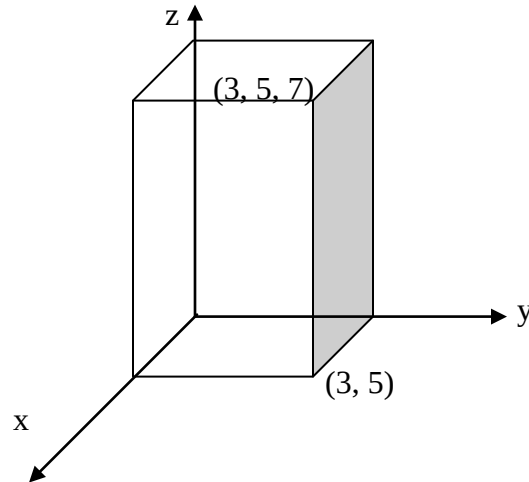
$$\begin{aligned} L(u+v) &= L((x_1, y_1, z_1) + (x_2, y_2, z_2)) \\ &= L(x_1 + x_2, y_1 + y_2, z_1 + z_2) \end{aligned}$$

$$\begin{aligned}
 &= (x_1 + x_2, y_1 + y_2) \\
 &= (x_1, y_1) + (x_2, y_2) = L(\mathbf{u}) + L(\mathbf{v})
 \end{aligned}$$

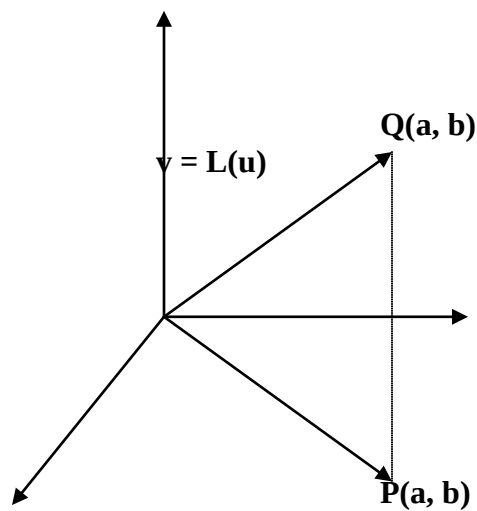
Also, if k is a real number, then

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

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 $\mathbf{R}^3$ is (a, b) in $\mathbf{R}^2$ can be found by drawing a line through the end point $P(a, b, c)$ of $\mathbf{u}$ and perpendicular to $\mathbf{R}^2$, 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 $L : R \rightarrow R$ be defined by $L(x) = x^2$

Let x and y in R and

$$\begin{aligned} L(x+y) &= (x+y)^2 = x^2 + y^2 + 2xy \neq x^2 + y^2 = L(x) + L(y) \\ \Rightarrow L(x+y) &\neq L(x) + L(y) \end{aligned}$$

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

for all vectors $\mathbf{u}, \mathbf{v}$ in the domain of T and all scalars c, d .

Proof

1. By the definition of Linear Transformation we have $T(c\mathbf{u}) = cT(\mathbf{u})$ for all $\mathbf{u}$ and all scalars c . Put $c = 0$ we'll get $T(\mathbf{0u}) = \mathbf{0}T(\mathbf{u})$ This implies $T(\mathbf{0}) = \mathbf{0}$
2. Just apply the definition of linear transformation. i. e

$$T(c\mathbf{u} + d\mathbf{v}) = T(c\mathbf{u}) + T(d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$$

Property (3) follows from (ii), because $T(\mathbf{0}) = T(\mathbf{0u}) = \mathbf{0}T(\mathbf{u}) = \mathbf{0}$.

Property (4) requires both (i) and (ii):

OBSERVATION Observe that if a transformation satisfies property 2 for all $\mathbf{u}, \mathbf{v}$ 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

$$T(c_1 \mathbf{v}_1 + \dots + c_p \mathbf{v}_p) = c_1 T(\mathbf{v}_1) + \dots + c_p T(\mathbf{v}_p)$$

Applications in Engineering

In engineering and physics, property 3 is referred to as a superposition principle.

Think of $\mathbf{v}_1, \dots, \mathbf{v}_p$ as signals that go into a system or process and $T(\mathbf{v}_1), \dots, T(\mathbf{v}_p)$ 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 $T : R \rightarrow R$ 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 $T : R^2 \rightarrow R^2$ by $T(x) = rx$.

T is called a **contraction** when $0 \leq r < 1$

and a **dilation** when $r \geq 1$.

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

Solution Let u, v be in R^2 and let c, d be scalars, then

$$\begin{aligned} T(cu + dv) &= 3(cu + dv) \\ &= 3cu + 3dv \\ &= c(3u) + d(3v) \\ &= cT(u) + dT(v) \end{aligned} \quad \left. \begin{array}{l} \text{Definition of } T \\ \text{Vector arithmetic} \end{array} \right\}$$

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

Example 10 Define a linear transformation $T : R^2 \rightarrow R^2$ 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 $u = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, and $u + v = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$.

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

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

In the above example, T rotates u , v and $u+v$ counterclockwise through 90° .

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 $L: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be a linear transformation for which:

$$L(1, 0, 0) = (2, -1),$$

$$L(0, 1, 0) = (3, 1), \text{ and}$$

$$L(0, 0, 1) = (-1, 2).$$

Then find $L(-3, 4, 2)$.

Solution Since $(-3, 4, 2) = -3\mathbf{i} + 4\mathbf{j} + 2\mathbf{k}$,

$$\begin{aligned} L(-3, 4, 2) &= L(-3\mathbf{i} + 4\mathbf{j} + 2\mathbf{k}) = -3L(\mathbf{i}) + 4L(\mathbf{j}) + 2L(\mathbf{k}) \\ &= -3(2, -1) + 4(3, 1) + 2(-1, 2) = (4, 11) \end{aligned}$$

Exercise

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

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

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

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 $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ by $T(\mathbf{x}) = \mathbf{Ax}$. Find $T(u)$ and $T(v)$.

In exercises 5 and 6, with T defined by $T(\mathbf{x}) = \mathbf{Ax}$, find an $\mathbf{x}$ whose image under T is $\mathbf{b}$, and determine if $\mathbf{x}$ 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 $\mathbb{R}^4$ that are mapped into the zero vector by the transformation $x \rightarrow Ax$.

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

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

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 $x \rightarrow Ax$?

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

Let $T(x) = Ax$ for x in $\mathbb{R}^2$.

- (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 $\mathbb{R}^2$.

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

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

13. Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear transformation that maps

$u = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$ into $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and maps $v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ into $\begin{bmatrix} 1 \\ -4 \end{bmatrix}$. Use the fact that T is linear find the images under T of $2u$, $3v$, and $2u + 3v$.

14. Let $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $y_1 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, and $y_2 = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$. Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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

Outline

- 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 $\mathbf{R}^n$ to $\mathbf{R}^m$ is actually a matrix transformation $x \rightarrow Ax$, where A is a matrix of order $m \times n$. First see an example

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

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

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

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

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

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

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

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

Theorem Let $T : \mathbf{R}^n \rightarrow \mathbf{R}^m$ be a linear transformation. Then there exists a unique matrix A such that $T(x) = Ax$ for all x in $\mathbf{R}^n$

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

$$A = [T(e_1) \quad \dots \quad T(e_n)]$$

Proof Write

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

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

Since T is Linear, So

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

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

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

Example 2 Find the standard matrix A for the dilation transformation $T(x) = 3x$, $x \in \mathbf{R}^2$.

Solution Write

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

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

Example 3 Let $L : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is the linear operator defined by $L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ y-z \\ x+z \end{pmatrix}$.

Find the standard matrix representing L and verify $L(\mathbf{x}) = A\mathbf{x}$.

Solution

The standard matrix A representing L is the 3×3 matrix whose columns are $L(\mathbf{e}_1)$, $L(\mathbf{e}_2)$, and $L(\mathbf{e}_3)$ respectively. Thus

$$L(\mathbf{e}_1) = L \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1+0 \\ 0-0 \\ 1+0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \text{col}_1(A)$$

$$L(\mathbf{e}_2) = L \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0+1 \\ 1-0 \\ 0+0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \text{col}_2(A)$$

$$L(\mathbf{e}_3) = L \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0+0 \\ 0-1 \\ 0+1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \text{col}_3(A)$$

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

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

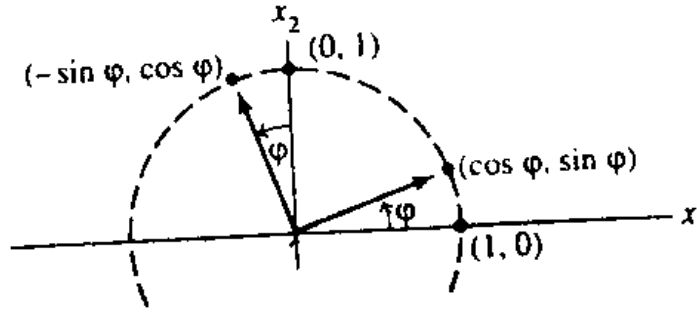
Hence verified.

Example 4 Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the transformation that rotates each point in $\mathbb{R}^2$ 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 $\mathbf{R}^2$ is defined by the

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

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

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

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

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

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

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

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

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

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

Geometric Linear Transformations of $\mathbb{R}^2$

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

OR

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

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

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

OR

Equivalently, T is one-to-one if for each $\mathbf{b}$ in $\mathbf{R}^m$ the equation $T(\mathbf{x}) = \mathbf{b}$ 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 $\mathbf{b}$ in $\mathbf{R}^m$ is the image of more than one vector in $\mathbf{R}^n$. If there is no such $\mathbf{b}$, 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 $\mathbf{R}^4$ onto $\mathbf{R}^3$? 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 $\mathbf{b}$ in $\mathbf{R}^3$, the equation $A\mathbf{x} = \mathbf{b}$ is consistent. In other words, the linear transformation T maps $\mathbf{R}^4$ (its domain) onto $\mathbf{R}^3$.

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

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

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

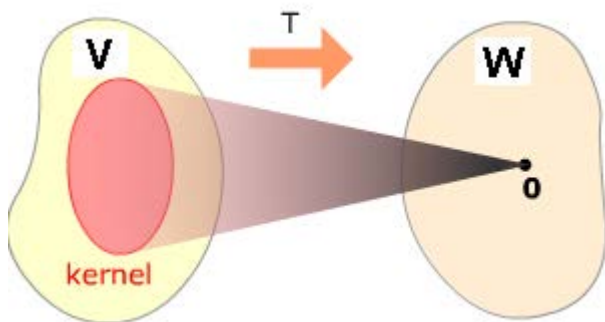
But then, since T is linear $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$

The vector $\mathbf{u} - \mathbf{v}$ is not zero, since $\mathbf{u} \neq \mathbf{v}$. Hence the equation $T(\mathbf{x}) = \mathbf{0}$ 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 $T : V \rightarrow W$ be a linear transformation. Then kernel of T (usually written as $\text{Ker}T$), is the set of those elements in V which maps onto the zero vector in W . Mathematically:

$$\text{Ker}T = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0} \text{ in } W\}$$

**Remarks**

- i) $\text{Ker } T$ is subspace of V
- ii) T is one-one iff $\text{Ker } T = \{0\}$ in V

One-One Linear Transformation

Let $T : V \rightarrow W$ 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 $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation and let A be the standard matrix for T . Then

- (a) T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$;
- (b) T is one-to-one if and only if the columns of A are linearly independent.

Proof:

- (a) The columns of A span $\mathbb{R}^m$ if and only if for each $\mathbf{b}$ the equation $A\mathbf{x} = \mathbf{b}$ is consistent – in other words, if and only if for every $\mathbf{b}$, the equation $T(\mathbf{x}) = \mathbf{b}$ has at least one solution. This is true if and only if T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$.
- (b) The equations $T(\mathbf{x}) = \mathbf{0}$ and $A\mathbf{x} = \mathbf{0}$ are the same except for notation. So T is one-to-one if and only if $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. This happens if and only if the columns of A are linearly independent.

We can also write column vectors in rows, using parentheses and commas. Also, when we apply a linear transformation T to a vector – say, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (x_1, x_2)$ we write $T(x_1, x_2)$ instead of the more formal $T((x_1, x_2))$.

Example 7 Let $T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$.

Show that T is a one-to-one linear transformation.
Does T map $\mathbb{R}^2$ onto $\mathbb{R}^3$?

Solution When $\mathbf{x}$ and $\mathbf{T}(\mathbf{x})$ are written as column vectors, it is easy to see that $\mathbf{T}$ is described by the equation

$$\begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} \quad (4)$$

so $\mathbf{T}$ is indeed a linear transformation, with its standard matrix $\mathbf{A}$ shown in (4). The columns of $\mathbf{A}$ are linearly independent because they are not multiples. Hence $\mathbf{T}$ is one-to-one. To decide if $\mathbf{T}$ is onto $\mathbf{R}^3$, we examine the span of the columns of $\mathbf{A}$. Since $\mathbf{A}$ is 3×2 , the columns of $\mathbf{A}$ span $\mathbf{R}^3$ if and only if $\mathbf{A}$ has 3 pivot positions. This is impossible, since $\mathbf{A}$ has only 2 columns. So the columns of $\mathbf{A}$ do not span $\mathbf{R}^3$ and the associated linear transformation is not onto $\mathbf{R}^3$.

Exercises

1. Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 $x_2 -$ axis. Assuming that T is linear, find its standard matrix.

Assume that T is a linear transformation. Find the standard matrix of T .

2. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3, T(1, 0) = (4, -1, 2)$ and $T(0, 1) = (-5, 3, -6)$

3. $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2, T(e_1) = (1, 4), T(e_2) = (-2, 9),$ and $T(e_3) = (3, -8),$ where $e_1, e_2,$ and e_3 are the columns of the identity matrix.

4. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ rotates points clockwise through π radians.

5. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a “vertical shear” transformation that maps e_1 into $e_1 + 2e_2$ but leaves the vector e_2 unchanged.

6. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a “horizontal shear” transformation that maps e_2 into $e_2 - 3e_1$ but leaves the vector e_1 unchanged.

7. $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ projects each point (x_1, x_2, x_3) vertically onto the x_1x_2 -plane (where $x_3=0$).

8. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ first rotates points counterclockwise through $\pi/4$ 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 $x_1, x_2, \dots$ 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 $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 $T : R^2 \rightarrow R^3$ be a linear transformation such that $T(x_1, x_2) = (x_1 + 2x_2, -x_1 - 3x_2, -3x_1 - 2x_2)$. Find $\mathbf{x}$ such that $T(\mathbf{x}) = (-4, 7, 0)$.

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 $\mathbf{R}^m$ and are denoted by (boldface) $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$.

These columns are $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]$

The number a_{ij} is the i -th entry (from the top) of j -th column vector $\mathbf{a}_j$.

$$\begin{array}{c}
 \text{Column} \\
 j \\
 \begin{array}{c}
 \left[\begin{array}{ccccc}
 a_{11} & \dots & a_{1j} & \dots & a_{1n} \\
 \vdots & & \vdots & & \vdots \\
 a_{i1} & \dots & a_{ij} & \dots & a_{in} \\
 \vdots & & \vdots & & \vdots \\
 a_{m1} & \dots & a_{mj} & \dots & a_{mn}
 \end{array} \right] = A \\
 \begin{array}{ccccc}
 \uparrow & & \uparrow & & \uparrow \\
 \mathbf{a}_1 & & \mathbf{a}_j & & \mathbf{a}_n
 \end{array}
 \end{array}
 \end{array}$$

Figure 1 Matrix notation.

Definitions

A **diagonal matrix** is a square matrix whose non-diagonal entries are zero.

$$D = \begin{bmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_{nn} \end{bmatrix}$$

The diagonal entries in $A = [a_{ij}]$ are $a_{11}, a_{22}, a_{33}, \dots$ 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 $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3×3 3×2 4×5

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** $r\mathbf{A}$ is the matrix whose columns are r times the corresponding columns in **A**.

Note: Negative of a matrix **A** is defined as $-\mathbf{A}$ to mean $(-1)\mathbf{A}$ and the difference of **A** and **B** is written as $\mathbf{A}-\mathbf{B}$, which means $\mathbf{A} + (-1)\mathbf{B}$.

Example 2 Let $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}$, $C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$

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. $A + B = B + A$	d. $r(A + B) = rA + rB$
b. $(A + B) + C = A + (B + C)$	e. $(r + s)A = rA + sA$
c. $A + 0 = A$	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 j th columns of A , B , and C are a_j, b_j and c_j , respectively, then the j th columns of $(A + B) + C$ and $A + (B + C)$ are

$$(a_j + b_j) + c_j \quad \text{and} \quad a_j + (b_j + c_j)$$

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 $m \times n$ matrix with an $n \times p$ matrix results in an $m \times p$ matrix. If many matrices are multiplied together, and their dimensions are written in a list in order, e.g. $m \times n, n \times p, p \times q, q \times r$, the size of the result is given by the first and the last numbers ($m \times r$).

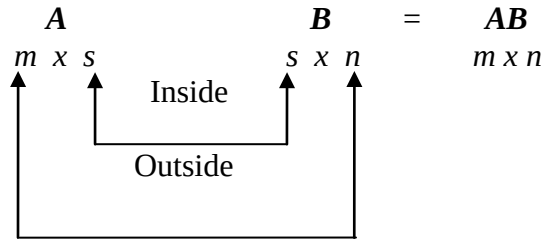
It is important to keep in mind that this definition requires the number of columns of the first factor A to be the same as the number of rows of the second factor B . When this condition is satisfied, the sizes of A and B are said to conform for the product AB . If the sizes of A and B do not conform 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 $b_1, \dots, b_p$, then the product AB is the $m \times p$ matrix whose columns are $Ab_1, \dots, Ab_p$.

That is
$$AB = A \begin{bmatrix} b_1 & b_2 & \dots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_p \end{bmatrix}$$

This definition makes equation (1) true for all $\mathbf{x}$ in $\mathbf{R}^p$. Equation (1) proves that the composite mapping $(\mathbf{AB})$ is a linear transformation and that its standard matrix is $\mathbf{AB}$. Multiplication of matrices corresponds to composition of linear transformations.

A convenient way to determine whether $\mathbf{A}$ and $\mathbf{B}$ conform for the product $\mathbf{AB}$ and, if so, to find the size of the product is to write the sizes of the factors side by side as in Figure below (the size of the first factor on the left and the size of the second factor on the right).



If the inside numbers are the same, then the product $\mathbf{AB}$ is defined and the outside numbers then give the size of the product.

Example 3 Compute $\mathbf{AB}$, where $\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$

Solution: Here $\mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3]$, therefore

$$\mathbf{Ab}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad \mathbf{Ab}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad \mathbf{Ab}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Then

$$\begin{array}{c}
 \mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix} \\
 \uparrow \quad \uparrow \quad \uparrow \\
 \mathbf{Ab}_1 \ \mathbf{Ab}_2 \ \mathbf{Ab}_3
 \end{array}$$

Note from the definition of $\mathbf{AB}$ that its first column, $\mathbf{Ab}_1$, is a linear combination of the columns of $\mathbf{A}$, using the entries in $\mathbf{b}_1$ as weights. The same holds true for each column of $\mathbf{AB}$. Each column of $\mathbf{AB}$ is a linear combination of the columns of $\mathbf{A}$ using weights from the corresponding column of $\mathbf{B}$.

Example 4 Find the product $\mathbf{AB}$ for

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$$

Solution It follows from definition that the product $\mathbf{AB}$ is formed in a column-by-column manner by multiplying the successive columns of $\mathbf{B}$ by $\mathbf{A}$. The computations are

$$c_1 \quad c_2 \quad c_3$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} = 4c_1 + 0c_2 + 2c_3 = (4) \begin{bmatrix} 1 \\ 2 \end{bmatrix} + (0) \begin{bmatrix} 2 \\ 6 \end{bmatrix} + (2) \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 12 \\ 8 \end{bmatrix}$$

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,
$$\mathbf{AB} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

Example 5 (An Undefined Product) Find the product $\mathbf{BA}$ for the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$$

Solution The number of columns of $\mathbf{B}$ is not equal to number of rows of $\mathbf{A}$ so $\mathbf{BA}$ multiplication is not possible.

The matrix $\mathbf{B}$ has size 3×4 and the matrix $\mathbf{A}$ has size 2×3 . The “inside” numbers are not the same, so the product $\mathbf{BA}$ is undefined.

Obviously, the number of columns of $\mathbf{A}$ must match the number of rows in $\mathbf{B}$ in order for a linear combination such as $\mathbf{Ab}_1$ to be defined. Also, the definition of $\mathbf{AB}$ shows that $\mathbf{AB}$ has the same number of rows as $\mathbf{A}$ and the same number of columns as $\mathbf{B}$.

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 in 1st 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:

$$\begin{array}{ccc}
 A & B & AB \\
 \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} & = \begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} \\
 3 \times 5 & 5 \times 2 & 3 \times 2 \\
 \begin{array}{c} \text{Match} \\ \text{Size of } AB \end{array} & &
 \end{array}$$

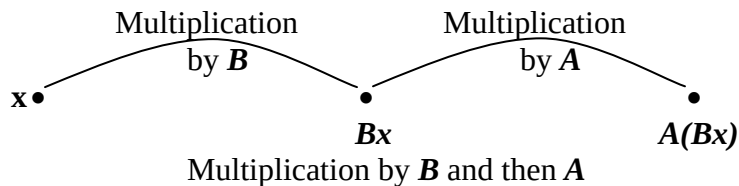
The product BA is not defined because the 2 columns of B do not match the 3 rows of A .

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

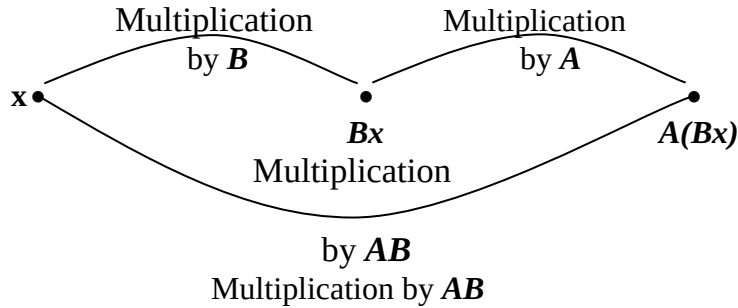
Explanation

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 $A(Bx)$.



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

$$A(Bx) = (AB)x \text{-----(1)}$$



If A is $m \times n$, B is $n \times p$, and x is in $\mathbf{R}^p$, denote the columns of B by $b_1, \dots, b_p$ and the entries in x by $x_1, \dots, x_p$, then $Bx = x_1 b_1 + x_2 b_2 + \dots + x_p b_p$

By the linearity of multiplication by A ,

$$\begin{aligned} A(Bx) &= A(x_1 b_1) + A(x_2 b_2) + \dots + A(x_p b_p) \\ &= x_1 Ab_1 + x_2 Ab_2 + \dots + x_p Ab_p \end{aligned}$$

The vector $A(Bx)$ is a linear combination of the vectors $Ab_1, \dots, Ab_p$, using the entries in x as weights. If we rewrite these vectors as the columns of a matrix, we have

$$A(Bx) = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_p \end{bmatrix} x$$

Thus multiplication by $\begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_p \end{bmatrix}$ 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 $(AB)_{ij}$ denotes the (i, j) – entry in AB , and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$$

To verify this rule, let $B = \begin{bmatrix} b_1 & \dots & b_p \end{bmatrix}$. Column j of AB is Ab_j , and we can compute Ab_j . The i th entry in Ab_j is the sum of the products of corresponding entries from row i of A and the vector b_j , which is precisely the computation described in the rule for computing the (i, j) – entry of AB .

Finding Specific Entries in a Matrix Product Sometimes we will be interested in finding a specific entry in a matrix product without going through the work of computing the entire column that contains the entry.

Example 7 Use the row-column rule to compute two of the entries in $\mathbf{AB}$ for the matrices in Example 3.

Solution: To find the entry in row 1 and column 3 of $\mathbf{AB}$, consider row 1 of $\mathbf{A}$ and column 3 of $\mathbf{B}$. Multiply corresponding entries and add the results, as shown below:

$$\begin{array}{c} \downarrow \\ \mathbf{AB} = \rightarrow \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6)+3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix} \end{array}$$

For the entry in row 2 and column 2 of $\mathbf{AB}$, use row 2 of $\mathbf{A}$ and column 2 of $\mathbf{B}$:

$$\begin{array}{c} \downarrow \\ \rightarrow \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3)+-5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix} \end{array}$$

Example 8 Use the dot product rule to compute the individual entries in the product of

$$\mathbf{AB} \text{ where } \mathbf{A} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}.$$

Solution Since $\mathbf{A}$ has size 2×3 and $\mathbf{B}$ has size 3×4 , the product $\mathbf{AB}$ is a 2×4 matrix of the form

$$\mathbf{AB} = \begin{bmatrix} r_1(\mathbf{A}) \times c_1(\mathbf{B}) & r_1(\mathbf{A}) \times c_2(\mathbf{B}) & r_1(\mathbf{A}) \times c_3(\mathbf{B}) & r_1(\mathbf{A}) \times c_4(\mathbf{B}) \\ r_2(\mathbf{A}) \times c_1(\mathbf{B}) & r_2(\mathbf{A}) \times c_2(\mathbf{B}) & r_2(\mathbf{A}) \times c_3(\mathbf{B}) & r_2(\mathbf{A}) \times c_4(\mathbf{B}) \end{bmatrix}$$

where $r_1(\mathbf{A})$ and $r_2(\mathbf{A})$ are the row vectors of $\mathbf{A}$ and $c_1(\mathbf{B}), c_2(\mathbf{B}), c_3(\mathbf{B})$ and $c_4(\mathbf{B})$ are the column vectors of $\mathbf{B}$. For example, the entry in row 2 and column 3 of $\mathbf{AB}$ 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 $\mathbf{AB}$ 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 $\mathbf{AB}$ is given by the formula

$$\mathbf{AB} = \mathbf{A} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix} = \begin{bmatrix} \mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_n \end{bmatrix}$$

Similarly, the specific row of $\mathbf{AB}$ is given by the formula $\mathbf{AB} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \mathbf{B} = \begin{bmatrix} a_1 \mathbf{B} \\ a_2 \mathbf{B} \\ \vdots \\ a_m \mathbf{B} \end{bmatrix}$

Example 9 Find the entries in the second row of $\mathbf{AB}$, where

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

Solution: By the row-column rule, the entries of the second row of $\mathbf{AB}$ come from row 2 of $\mathbf{A}$ (and the columns of $\mathbf{B}$):

$$\rightarrow \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} \begin{matrix} \downarrow & \downarrow \\ \begin{bmatrix} \square & \square \\ -4+21-12 & 6+3-8 \\ \square & \square \\ \square & \square \end{bmatrix} \end{matrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

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 $\mathbf{R}^m$.

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. $r(AB) = (rA)B = A(rB)$ (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 $C = \begin{bmatrix} c_1 & \dots & c_p \end{bmatrix}$

By definition of matrix multiplication $BC = \begin{bmatrix} Bc_1 & \dots & Bc_p \end{bmatrix}$

$$A(BC) = \begin{bmatrix} A(Bc_1) & \dots & A(Bc_p) \end{bmatrix}$$

From above, we know that $A(Bx) = (AB)x$ for all x , so

$$A(BC) = \begin{bmatrix} (AB)c_1 & \dots & (AB)c_p \end{bmatrix} = (AB)C$$

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 do 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

$$A = \begin{bmatrix} -3 & 2 \\ -6 & 4 \end{bmatrix} \quad B = \begin{bmatrix} -1 & 2 \\ 3 & -2 \end{bmatrix} \quad C = \begin{bmatrix} -1 & 2 \\ 3 & -2 \end{bmatrix}$$

$$AB = \begin{bmatrix} 9 & -10 \\ 18 & -20 \end{bmatrix} = AC \quad \text{But } B \neq C$$

3. If a product AB is the zero matrix, you cannot conclude in general that either $A = \mathbf{0}$ or $B = \mathbf{0}$.

For example:

$$\text{If } AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ then it can be either}$$

$$A = \begin{bmatrix} 1 & 4 \\ 6 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ or}$$

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 4 \\ 6 & 2 \end{bmatrix}$$

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, A^k denotes the product of k copies of A , $A^k = \underbrace{A \dots A}_k$. Also, we interpret 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 A^t , whose columns are formed from the corresponding rows of A . OR, if A is an $m \times n$ matrix, then transpose of A is denoted by A^t , is defined to be the $n \times m$ matrix that is obtained by making the rows of A into columns; that is, the first column of A^t is the first row of A , the second column of 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 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$

Then $A^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$, $B^t = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}$, $C^t = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$

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

$$AC = CA = I \text{ where } I \text{ is the } n \times n \text{ identity matrix.}$$

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 A^{-1} , so that

$$AA^{-1} = I \quad \text{and} \quad A^{-1}A = I$$

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 $A^{-1}.A = A.A^{-1} = I$ The identity matrix $A^{-1} = \frac{1}{A}$, as in case of real number, we have $3^{-1} = \frac{1}{3}$.

A^{-1} is, in fact the $n \times n$ matrix corresponding to the $n \times n$ matrix A , which satisfies the property

$$A^{-1}.A = A.A^{-1} = I \text{ The identity matrix}$$

Example 1 If $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, then

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -14+15 & -10+10 \\ 21-21 & 15-14 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and}$$

$$CA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} -14+15 & -35+35 \\ 6-6 & 15-14 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus $C = A^{-1}$.

Theorem Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

If $ad - bc \neq 0$, then A is invertible or non singular and $A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

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 $\det A = 3(6) - 4(5) = -2 \neq 0$.

$$\text{Hence } A \text{ is invertible } A^{-1} = \frac{1}{-2} \begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 6/(-2) & -4/(-2) \\ -5/(-2) & 3/(-2) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}$$

The next theorem provides three useful facts about invertible matrices.

Theorem

- If A is an invertible matrix, then A^{-1} is invertible and $(A^{-1})^{-1} = A$
- 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 $(AB)^{-1} = B^{-1}A^{-1}$
- If A is an invertible matrix, then so is A^T , and the inverse of A^T is the transpose of A^{-1} . That is $(A^T)^{-1} = (A^{-1})^T$

Proof

(a) We must find a matrix C such that $A^{-1}C = I$ and $CA^{-1} = I$

However, we already know that these equations are satisfied with A in place of C . Hence 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

$$((A_1)(A_2)(A_3)\dots(A_n))^{-1} = A_n^{-1}A_{n-1}^{-1}\dots A_3^{-1}A_2^{-1}A_1^{-1}$$

The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.

Example 3 (Inverse of a Transpose). Consider a general 2×2 invertible matrix and its transpose:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad A^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Since A is invertible, its determinant $(ad - bc)$ is nonzero. But the determinant of A^t is also $(ad - bc)$, so A^t is also invertible. It follows that

$$(A^t)^{-1} = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{c}{ad-bc} \\ -\frac{b}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \text{-----(1)}$$

Now

$$A^{-1} = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$$

Therefore, $(A^{-1})^t = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{c}{ad-bc} \\ -\frac{b}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$ -----(2)

From (1) and (2), we have

$$(A^t)^{-1} = (A^{-1})^t.$$

Example 4 (The Inverse of a Product). Consider the matrices

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

$$AB = \begin{bmatrix} 7 & 6 \\ 9 & 8 \end{bmatrix},$$

Here

$$(AB)^{-1} = \frac{1}{|AB|} \text{Adj}(AB) = \frac{1}{-2} \begin{bmatrix} 8 & -6 \\ -9 & 7 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|} \text{Adj}(A) = \frac{1}{1} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix},$$

$$B^{-1} = \frac{1}{|B|} \text{Adj}(B) = \frac{1}{2} \begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix},$$

$$B^{-1}A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

Thus, $(AB)^{-1} = B^{-1}A^{-1}$

Theorem: If A is invertible and n is a non-negative integer, then:

(a) A^n is invertible and $(A^n)^{-1} = A^{-n} = (A^{-1})^n$

(b) kA is invertible for any nonzero scalar k, and $(kA)^{-1} = k^{-1}A^{-1}$.

Example 5 (Related to the above theorem)

(a) Let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \quad \text{then } A^{-1} = \frac{1}{|A|} \text{Adj}(A) = \frac{1}{1} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$$

Now

$$A^{-3} = (A^{-1})^3 = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}$$

Also,

$$A^3 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 30 \\ 15 & 41 \end{bmatrix}$$

$$(A^3)^{-1} = \frac{1}{(11)(41) - (30)(15)} \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} = (A^{-1})^3$$

(b)

Take $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$ and $k=3$

$$kA = 3A = \begin{bmatrix} 3 & 6 \\ 9 & 3 \end{bmatrix}, \quad (kA)^{-1} = (3A)^{-1} = \frac{1}{9-54} \begin{bmatrix} 3 & -6 \\ -9 & 3 \end{bmatrix} = \begin{bmatrix} -1/15 & 2/15 \\ 1/5 & -1/15 \end{bmatrix} \text{-----(1)}$$

$$A^{-1} = -\frac{1}{5} \begin{bmatrix} 1 & -2 \\ -3 & 1 \end{bmatrix}$$

$$A = \text{So } k^{-1}A^{-1} = 3^{-1}A^{-1} = \frac{1}{3} \cdot \frac{-1}{5} \begin{bmatrix} 1 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -1/15 & 2/15 \\ 1/5 & -1/15 \end{bmatrix} \text{-----(2)}$$

From (1) and (2), we have

$$(3A)^{-1} = 3^{-1}A^{-1}$$

There is an important connection between invertible matrices and row operations that leads to a method for computing inverses. As we shall see, an invertible matrix A is row equivalent to an identity matrix, and we can find A^{-1} by watching the row reduction of A to I .

Elementary Matrices

As we have studied that there are three types of elementary row operations that can be performed on a matrix:

There are three types of elementary operations

- Interchanging of any two rows
- Multiplication to a row by a nonzero constant
- Adding a multiple of one row to another

Elementary matrix

An elementary matrix is a matrix that results from applying a single elementary row operation to an identity matrix.

Some examples are given below:

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

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 an 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

Example 7 (Recovering Identity Matrices from Elementary Matrices). Here are three examples that use inverses of row operations to recover the identity matrix from

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Multiply the second row by 7. Multiply the second row by 1/7.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{\text{Interchange the first and second rows.}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \xrightarrow{\text{Interchange the first and second rows.}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Add 5 times the second row to the first. Add -5 times the second row to the first.

The next theorem is the basic result on the invertibility of elementary matrices.

Theorem An elementary matrix is invertible and the inverse is also an elementary matrix.

Example 8 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 E_1A , E_2A , E_3A and describe how these products can be obtained by elementary row operations on A .

Solution We have

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, \quad E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix}, \quad E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}$$

Addition of (-4) times row 1 of $\mathbf{A}$ to row 3 produces $\mathbf{E}_1\mathbf{A}$. (This is a row replacement operation.) An interchange of rows 1 and 2 of $\mathbf{A}$ produces $\mathbf{E}_2\mathbf{A}$ and multiplication of row 3 of $\mathbf{A}$ by 5 produces $\mathbf{E}_3\mathbf{A}$.

Left-multiplication (that is, multiplication on the left) by $\mathbf{E}_1$ 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 $\mathbf{E}_1 \mathbf{I} = \mathbf{E}_1$, we see that $\mathbf{E}_1$ 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 $\mathbf{E}$ is produced by a row operation on $\mathbf{I}$, then there is another row operation of the same type that changes $\mathbf{E}$ back into $\mathbf{I}$. Hence there is an elementary matrix $\mathbf{F}$ such that $\mathbf{FE} = \mathbf{I}$. Since $\mathbf{E}$ and $\mathbf{F}$ correspond to reverse operations, $\mathbf{EF} = \mathbf{I}$.

Each elementary matrix $\mathbf{E}$ is invertible. The inverse of $\mathbf{E}$ is the elementary matrix of the same type that transforms $\mathbf{E}$ back into $\mathbf{I}$.

Example Find the inverse of $\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

Solution: To transform $\mathbf{E}_1$ into $\mathbf{I}$, add $+4$ times row 1 to row 3.

The elementary matrix which does that is $\mathbf{E}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$

Theorem An $n \times n$ matrix $\mathbf{A}$ is invertible if and only if $\mathbf{A}$ is row equivalent to $\mathbf{I}_n$, and in this case, any sequence of elementary row operations that reduces $\mathbf{A}$ to $\mathbf{I}_n$ also transforms $\mathbf{I}_n$ into $\mathbf{A}^{-1}$.

An Algorithm for Finding $\mathbf{A}^{-1}$ If we place $\mathbf{A}$ and $\mathbf{I}$ side-by-side to form an augmented matrix $[\mathbf{A} \ \mathbf{I}]$, then row operations on this matrix produce identical operations on $\mathbf{A}$ and $\mathbf{I}$. Then either there are row operations that transform $\mathbf{A}$ to $\mathbf{I}_n$, and $\mathbf{I}_n$ to $\mathbf{A}^{-1}$, or else $\mathbf{A}$ is not invertible.

Algorithm for Finding $\mathbf{A}^{-1}$

Row reduce the augmented matrix $[\mathbf{A} \ \mathbf{I}]$. If $\mathbf{A}$ is row equivalent to $\mathbf{I}$, then $[\mathbf{A} \ \mathbf{I}]$ is row equivalent to $[\mathbf{I} \ \mathbf{A}^{-1}]$. Otherwise, $\mathbf{A}$ 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 $[A \ I] = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} R_{12}$

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

$$\begin{array}{c} -2R_3 + R_2 \qquad \qquad -3R_3 + R_1 \\ \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix} \end{array}$$

Since $A \sim I$, 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 $A^{-1}A = I$ 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.

$$\text{Consider } \det A = \begin{vmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 5 \\ 3 & 1 & -2 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -3 & 1 \\ 0 & 5 & -6 & -1 \\ 0 & -4 & 7 & 3 \\ 0 & -5 & 7 & 2 \end{vmatrix}$$

operating $R_2 + R_1, R_3 - 2R_1, R_4 - 3R_1$

$$\text{Expand from first column} = \begin{vmatrix} 5 & -6 & -1 \\ -4 & 7 & 3 \\ -5 & 7 & 2 \end{vmatrix} = \begin{vmatrix} 5 & -6 & -1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{vmatrix} = 5(1-2) + 6(1-0) - 1(1-0) = 0$$

As the given matrix is singular, so it is not invertible.

Example 11 Find the inverse of the given matrix if possible $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix}$

Solution

$$\det A = \begin{vmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{vmatrix} = -1$$

As the given matrix is non-singular therefore, inverse is possible.

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

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

$$R_2 - 2R_1, R_3 - 3R_1$$

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

$$R_3 - R_2$$

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

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

$$\text{Multiply } R_3 \text{ by } -1$$

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

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

$$R_1 - R_3, R_2 + R_3$$

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

Hence the inverse of matrix A is $A^{-1} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & 1 & -1 \end{bmatrix}$

Example 12 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.

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

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

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & -2 & -1 \\ 0 & -8 & -7 \end{bmatrix}$$

$$R_2 - 2R_1, R_3 - 5R_1$$

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

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 1/2 \\ 0 & -8 & -7 \end{bmatrix}$$

$$\text{multiply 2nd row by } -1/2$$

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

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

$$R_3 + 8R_2$$

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

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{Multiply 3rd row by } -1/3$$

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

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

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

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

$$R_2 - (1/2)R_3, R_1 - 2R_3$$

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

$$\text{Hence the inverse of the original matrix } A^{-1} = \begin{bmatrix} 0 & -1/3 & 1/3 \\ 3/2 & -7/6 & 1/6 \\ -1 & 4/3 & -1/3 \end{bmatrix}$$

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 $E_2 E_1 A = I$.
 (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 A be an $n \times n$ invertible matrix. For any $b \in \mathbf{R}^n$, the equation $Ax = b$ (1) has the unique solution i.e. $x = A^{-1}b$.

Proof

Since A is invertible and $b \in \mathbf{R}^n$ be any vector. Then, we must have a matrix $A^{-1}b$ which is a solution of eq. (1) i.e. $Ax = A(A^{-1}b) = Ib = b$.

Uniqueness

For uniqueness, we assume that there is another solution u . Indeed, it is a solution of eq.(1) so it must be $u = A^{-1}b$, it means $x = A^{-1}b = u$. This shows that $u = x$.

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

$$2x_1 + x_2 + x_3 = 1$$

$$5x_1 + x_2 + 3x_3 = 3$$

$$x_1 + \quad \quad 4x_3 = 6$$

Solution

Consider the linear system

$$2x_1 + x_2 + x_3 = 1$$

$$5x_1 + x_2 + 3x_3 = 3$$

$$x_1 + \quad \quad 4x_3 = 6$$

The Matrix form of system is $Ax = b$, 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:

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 5 & 1 & 3 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 5 & 1 & 3 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 & 0 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & -17 & 0 & 1 & -5 \\ 0 & 1 & -7 & 1 & 0 & -2 \end{array} \right], \quad -5R_1 + R_2, \quad -2R_1 + R_3$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & -17 & 0 & 1 & -5 \\ 0 & 0 & 10 & 1 & -1 & 3 \end{array} \right], \quad -R_2 + R_3,$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & -17 & 0 & 1 & -5 \\ 0 & 0 & 1 & \frac{1}{10} & \frac{-1}{10} & \frac{3}{10} \end{array} \right], \quad \frac{R_3}{10}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-2}{5} & \frac{2}{5} & \frac{-1}{5} \\ 0 & 1 & 0 & \frac{17}{10} & \frac{-7}{10} & \frac{1}{10} \\ 0 & 0 & 1 & \frac{1}{10} & \frac{-1}{10} & \frac{3}{10} \end{array} \right], \quad -4R_3 + R_1, \quad -17R_3 + R_2$$

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 $x = A^{-1}b = A^{-1} \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{-2}{5} \\ \frac{1}{5} \\ \frac{8}{5} \end{bmatrix}$

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, $[\det(A)] = 40 - 2(16 - 3) + 3(0 - 5) = 40 - 26 - 15 = -1 \neq 0$

Therefore, A is invertible.

Now, we apply the inversion algorithm:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array} \right] \begin{array}{l} \\ -2R_1 + R_2, \quad -1R_1 + R_3 \\ \end{array}$$

$$\begin{aligned}
 & \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array} \right] 2R_2 + R_3 \\
 & \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array} \right] -1R_3 \\
 & \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array} \right] 3R_3 + R_2, \quad -3R_3 + R_1 \\
 & \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array} \right] -2R_2 + R_1
 \end{aligned}$$

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?

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

Solution

We see

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -3 & 1 \\ 1 & 0 & -4 \end{bmatrix} \text{ is an invertible matrix (det (A) } \neq 0 \text{)}$$

Thus, this homogeneous linear system has **only the trivial solution**.

Example 5

$$\text{Solve } \begin{cases} x_1 + x_2 + x_3 = 8 \\ 2x_2 + 3x_3 = 24 \\ 5x_1 + 5x_2 + x_3 = 8 \end{cases}$$

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 $[\det(\mathbf{A})] = 1(2 - 15) - 1(0 - 15) + 1(0 - 10) = -13 + 15 - 10 = -8 \neq 0$
 Therefore, $\mathbf{A}$ is invertible.

Now, we apply the inversion algorithm:

$$\begin{array}{c}
 \begin{array}{cc} \mathbf{A} & \mathbf{I}_3 \end{array} \\
 \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 5 & 5 & 1 & 0 & 0 & 1 \end{array} \right] \\
 \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array} \right] \quad -5R_1 + R_3 \\
 \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array} \right] \quad \frac{1}{2}R_2 \\
 \left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array} \right] \quad -1R_2 + R_1 \\
 \left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{5}{4} & 0 & -\frac{1}{4} \end{array} \right] \quad -\frac{1}{4}R_3 \\
 \left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{15}{8} & \frac{1}{2} & \frac{3}{8} \\ 0 & 0 & 1 & \frac{5}{4} & 0 & -\frac{1}{4} \end{array} \right] \quad -\frac{3}{2}R_3 + R_2 \\
 \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{13}{8} & -\frac{1}{2} & -\frac{1}{8} \\ 0 & 1 & 0 & -\frac{15}{8} & \frac{1}{2} & \frac{3}{8} \\ 0 & 0 & 1 & \frac{5}{4} & 0 & -\frac{1}{4} \end{array} \right] \quad \frac{1}{2}R_3 + R_1
 \end{array}$$

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 $x_1 = 0, x_2 = 0, x_3 = 8$.

Theorem (Invertible Matrix Theorem) Let A be a square $n \times n$ matrix. Then the following statements are equivalent. (Means if any one holds then all are true).

- (a) A is an invertible matrix.
- (b) A is row equivalent to the $n \times n$ identity matrix.
- (c) A has n pivot positions.
- (d) The equation $Ax = 0$ has only the trivial solution.
- (e) The columns of A form a linearly independent set.
- (f) The linear transformation $x \rightarrow Ax$ is one-to-one.
- (g) The equation $Ax = b$ has at least one solution for each b in R^n .
- (h) The columns of A span R^n .
- (i) The linear transformation $x \rightarrow Ax$ maps R^n onto 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) A^T is an invertible matrix.

Example 6

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

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{Here, } A \text{ has three pivot positions and hence is}$$

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,

$$\begin{aligned}
 A &= \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{R_2 - 2R_1, R_3 - R_1} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & -2 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{R_{23}} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \\
 &\xrightarrow{-\frac{1}{2}R_2} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{R_4 - R_2} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & \frac{3}{2} \end{bmatrix} \xrightarrow{(-1)R_3} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & \frac{3}{2} \end{bmatrix} \xrightarrow{-R_3 + R_4} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & \frac{5}{2} \end{bmatrix}
 \end{aligned}$$

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

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix} \xrightarrow{R_2 - 2R_1, R_3 - R_1} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & -2 & 0 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix} \xrightarrow{R_{23}} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & 0 & 2 \\ 0 & 0 & -1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

$$\begin{array}{cccc}
-\frac{1}{2}R_2 & R_4 - R_2 & (-1)R_3 & -R_3 + R_4 \\
\sim \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix} & \sim \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix} & \sim \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 1 & 3 \end{bmatrix} & \sim \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}
\end{array}$$

Here, $\mathbf{A}$ 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

$$Ax_1 = b_1, Ax_2 = b_2, \dots, Ax_k = b_k \quad (1)$$

where coefficient matrix $\mathbf{A}$ remains same and off course if it is invertible, then we have a sequence of solutions. i.e.

$$x_1 = A^{-1}b_1, x_2 = A^{-1}b_2, \dots, x_k = A^{-1}b_k.$$

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

$$T(x, y) = xyt + (x + y)t^2$$

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

$$T(1,0) = t^2 = (0,0,1) \text{ (This triple represents the coefficients of polynomial } t^2)$$

$$\text{i.e. } t^2 = 0.1 + 0.t + 1.t^2$$

Similarly, $T(0,1) = t = (0,1,0)$. 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 $\mathbb{R}^2$ to $\mathbb{R}^2$ 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

$$\text{is } A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

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 anything about the existence or nonexistence of solutions to equations of the form $Ax = b$.

Invertible Linear Transformations

Recall that matrix multiplication corresponds to composition of linear transformations. When a matrix A is invertible, the equation $A^{-1}Ax = x$ can be viewed as a statement about linear transformations. See Figure 2.

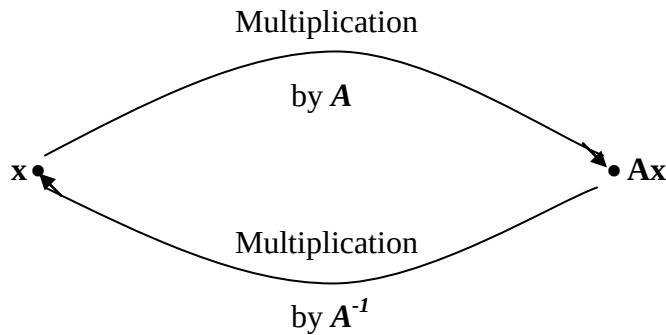


Figure 2 A^{-1} transforms Ax back to x

Definition

A linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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)

$$S(T(x)) = x \quad \forall \quad x \in \mathbb{R}^n \quad (1)$$

$$T(S(x)) = x \quad \forall \quad x \in \mathbb{R}^n \quad (2)$$

Important Note

If the inverse of a linear transformation exists then it is unique.

Proposition

Let $T : R^n \rightarrow R^m$ be linear transformation, given as $T(x) = Ax$, $\forall x \in R^n$, 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 $\text{Rank}(A) = n$.

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.

$$x_1 + x_2 + 4x_3 = 2$$

$$2x_2 + 3x_3 = 4$$

$$5x_1 + x_2 - x_3 = 3$$

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

$$[A \quad b_1 \quad b_2 \quad b_3 \quad b_4].$$

3. (a) Solve the two systems of linear equations

$$x_1 + 2x_2 + x_3 = -1$$

$$x_1 + 3x_2 + 2x_3 = 3$$

$$x_2 + 2x_3 = 4$$

&

$$x_1 + 2x_2 + x_3 = 0$$

$$x_1 + 3x_2 + 2x_3 = 0$$

$$x_2 + 2x_3 = 4$$

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

$$8. \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 1 & -2 & -1 \\ -2 & -6 & 3 & 2 \\ 3 & 5 & 8 & -3 \end{bmatrix} \quad 9. \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 10 \end{bmatrix} \quad 10. \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$$

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

12. Suppose that $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices and the equation $\mathbf{ABx} = \mathbf{0}$ has a nontrivial solution. What can you say about the matrix $\mathbf{AB}$?

13. What can we say about a one-to-one linear transformation T from $\mathbf{R}^n$ into $\mathbf{R}^n$?

14. Let $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ be a linear transformation given as $T(x) = 5x$, then find a matrix $\mathbf{A}$ of linear transformation T .

In exercises 15 and 16, T is a linear transformation from $\mathbf{R}^2$ into $\mathbf{R}^2$. 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 $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$ be a linear transformation and let $\mathbf{A}$ be the standard matrix for T . Then T is invertible if and only if $\mathbf{A}$ 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)

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

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)

$$A = \left[\begin{array}{c|c|c} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{array} \right] = [c_1 \mid c_2 \mid c_3]$$

In this case, we partitioned A into three column matrices each representing one column in the original matrix. Again, note that we used the standard column matrix notation (the bold face letters) and subscripted each one with the location in the partitioned matrix.

The c_i in the partitioned matrix are sometimes called the **column matrices of A** .

Partition (c)

$$A = \left[\begin{array}{c} a_{11} \quad a_{12} \quad a_{13} \\ a_{21} \quad a_{22} \quad a_{23} \\ a_{31} \quad a_{32} \quad a_{33} \\ a_{41} \quad a_{42} \quad a_{43} \\ a_{51} \quad a_{52} \quad a_{53} \end{array} \right] = \left[\begin{array}{c} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{array} \right]$$

Just as we can partition a matrix into each of its columns as we did in the previous part we can also partition a matrix into each of its rows. The r_i in the partitioned matrix are sometimes called the **row matrices of A** .

Addition of Blocked Matrices

If matrices A and B are the same size and are partitioned in exactly the same way, then it is natural to make the same partition of the ordinary matrix sum $A + B$. In this case, each block of $A + B$ is the (matrix) sum of the corresponding blocks of A and B .

Theorem

If A is $m \times n$ and B is $n \times p$, then

$$\begin{aligned}
 AB &= Col_1(A) + Col_2(A) + \dots + Col_n(A) \begin{bmatrix} Row_1(B) \\ Row_2(B) \\ \vdots \\ Row_n(B) \end{bmatrix} \\
 &= Col_1(A)Row_1(B) + Col_2(A)Row_2(B) + \dots + Col_n(A)Row_n(B)
 \end{aligned}$$

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

$$\begin{aligned}
 &a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} \\
 &(k = 1, 2, \dots, n)
 \end{aligned}$$

This sum is also the (i,j) -entry in AB by the row column rule

Example 4

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \text{ and } B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \text{ then}$$

$$\begin{aligned}
 A + B &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \\
 &= \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix}
 \end{aligned}$$

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

$$\text{If } A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \text{ and } B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

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

$$A_{11} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & -2 & 3 \end{bmatrix}, A_{12} = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}, A_{21} = [2 \ 1 \ -2] \text{ and } A_{22} = [1 \ 3]:$$

$$B_{11} = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix}, B_{21} = \begin{pmatrix} 0 & 8 \\ 0 & 2 \end{pmatrix} \text{ So } B = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$$

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:

$$A_{11}B_{11} + A_{12}B_{21} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & -2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix} + \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} \begin{bmatrix} 0 & 8 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ -23 & 14 \end{bmatrix}$$

$$\begin{aligned} A_{21}B_{11} + A_{22}B_{21} &= [2 \ 1 \ -2] \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix} + [1 \ 3] \begin{bmatrix} 0 & 8 \\ 0 & 2 \end{bmatrix} \\ &= [17 \ 0] + [0 \ 14] = [17 \ 14] \end{aligned}$$

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.

Example 6

The block version of the row –column rule for the product AB of the partitioned matrices

$$A = \left[\begin{array}{ccc|cc} 3 & -4 & 1 & 0 & 2 \\ -1 & 5 & -3 & 1 & 4 \\ \hline 2 & 0 & -2 & 1 & 6 \end{array} \right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \left[\begin{array}{c|c} 2 & -1 \\ 3 & 0 \\ \hline -5 & 1 \\ 4 & -3 \\ 0 & 2 \end{array} \right] = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$$

where

$$B_{11} = \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 1 \end{bmatrix} \text{ and } B_{21} = \begin{bmatrix} 4 & -3 \\ 0 & 2 \end{bmatrix}$$

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,

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix} = \begin{bmatrix} -11 & 2 \\ 32 & 3 \\ 18 & 5 \end{bmatrix}$$

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

$$\begin{aligned} A^2 &= \begin{pmatrix} 2I_3 & A_{32} \\ O_{23} & 2I_2 \end{pmatrix} \begin{pmatrix} 2I_3 & A_{32} \\ O_{23} & 2I_2 \end{pmatrix} \\ &= \begin{pmatrix} 4I_3 & 4A_{32} \\ O_{23} & 4I_2 \end{pmatrix} \end{aligned}$$

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 $A^2 = \begin{bmatrix} I_3 & O_{32} & A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t & O_{12} & 1 \end{bmatrix} \begin{bmatrix} I_3 & O_{32} & A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t & O_{12} & 1 \end{bmatrix} = \begin{bmatrix} I_3 + A_1 A_1^t & O_{32} & A_1 + A_1 \\ O_{23} & I_2 & O_{21} \\ A_1^t + A_1^t & O_{12} & A_1^t A_1 + 1 \end{bmatrix}$

$$I_3 + A_1 A_1^t = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, A_1 + A_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, A_1^t + A_1^t = \begin{bmatrix} 2 & 2 & 2 \end{bmatrix}$$

$$A_1^t A_1 + 1 = \begin{bmatrix} 4 \end{bmatrix}$$

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.

$$A = \begin{bmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k \end{bmatrix} \quad (1)$$

where the matrices $D_1, D_2, \dots, D_k$ are square. It can be shown that the matrix A in (1) is invertible if and only if each matrix on the diagonal is invertible. i.e.

$$A^{-1} = \begin{bmatrix} D_1^{-1} & 0 & \dots & 0 \\ 0 & D_2^{-1} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k^{-1} \end{bmatrix}$$

Example 10 Let A be a block diagonal matrix

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

Find A^{-1} .

Solution

There are three matrices on the main diagonal; two are 2×2 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

$$A_{11}^{-1} = \frac{\text{Adj}A_{11}}{A_{11}} = \frac{\begin{pmatrix} -1 & -2 \\ -1 & 1 \end{pmatrix}}{-3} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}.$$

$$\text{Similarly } A_{22}^{-1} = \frac{\begin{pmatrix} 4 & -1 \\ -5 & 1 \end{pmatrix}}{-1}$$

$$A_{22}^{-1} = \begin{pmatrix} -4 & 1 \\ 5 & -1 \end{pmatrix}$$

$$\Rightarrow A^{-1} = \left[\begin{array}{cc|cc|c} \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{-1}{3} & 0 & 0 & 0 \\ \hline 0 & 0 & -4 & 1 & 0 \\ 0 & 0 & 5 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 & \frac{1}{2} \end{array} \right]$$

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.

$$\Rightarrow A^{-1} = \left[\begin{array}{cc|cc|c} 1 & -7 & 0 & 0 & 0 \\ 1 & -8 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & -5 & 3 & 0 \\ \hline 0 & 0 & 0 & 0 & \frac{1}{4} \end{array} \right]$$

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

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ O & A_{22} & \dots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & A_{kk} \end{bmatrix} \text{ where the matrices } A_{11}, A_{22}, \dots, A_{kk} \text{ are}$$

square.

Note The definition of block lower triangular matrix is similar.

Here, we are going to introduce a formula for finding inverse of a block upper triangular matrix in the following example.

Example 12

Let A be a block upper triangular matrix of the form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \text{ where the orders of } A_{11} \text{ and } A_{22} \text{ are } p \times p \text{ and } q \times q \text{ respectively. Find } A^{-1}.$$

Solution

$$\text{Let } B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \text{ be inverse of } A \text{ i.e. } A^{-1} = B, \text{ then}$$

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & O \\ O & I_q \end{bmatrix}$$

$$\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{22}B_{21} & A_{22}B_{22} \end{bmatrix} = \begin{bmatrix} I_p & O \\ O & I_q \end{bmatrix}$$

By comparing corresponding entries, we have

$$A_{11}B_{11} + A_{12}B_{21} = I_p \quad (1)$$

$$A_{11}B_{12} + A_{12}B_{22} = O \quad (2)$$

$$A_{22}B_{21} = O \quad (3)$$

$$A_{22}B_{22} = I_q \quad (4)$$

Since A_{22} is a square matrix, so by Invertible Matrix Theorem, we have A_{22} is invertible.

Thus by eq.(4), $B_{22} = A_{22}^{-1}$. 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),

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \quad \text{and} \quad B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

Thus

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{bmatrix} \quad (5)$$

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

$$A = \left[\begin{array}{cc|cc} 1 & 9 & -5 & 0 \\ 3 & 3 & 3 & -2 \\ \hline 0 & 0 & 7 & 0 \\ 0 & 0 & 3 & 1 \end{array} \right]$$

$$\text{Put } A_{11} = \begin{bmatrix} 1 & 9 \\ 3 & 3 \end{bmatrix}, A_{12} = \begin{bmatrix} -5 & 0 \\ 3 & -2 \end{bmatrix} \quad \text{and} \quad A_{22} = \begin{bmatrix} 7 & 0 \\ 3 & 1 \end{bmatrix}$$

$$\text{Thus } A_{11}^{-1} = \frac{\text{Adj}A}{\det(A)} = \frac{\begin{bmatrix} 3 & -9 \\ -3 & 1 \end{bmatrix}}{-24} = \begin{pmatrix} \frac{1}{-8} & \frac{9}{24} \\ \frac{1}{8} & \frac{1}{-24} \end{pmatrix} \quad \text{and} \quad A_{22}^{-1} = \begin{bmatrix} \frac{1}{7} & 0 \\ \frac{-3}{7} & 1 \end{bmatrix}$$

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 $n \times n$ 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} .

$$\begin{bmatrix} I & 0 & 0 \\ X & I & 0 \\ Y & 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \\ 0 & B_{32} \end{bmatrix}$$

$$5. \text{ The inverse of } \begin{bmatrix} I & 0 & 0 \\ C & I & 0 \\ A & B & I \end{bmatrix} \text{ is } \begin{bmatrix} I & 0 & 0 \\ Z & I & 0 \\ X & Y & I \end{bmatrix}. \text{ Find X, Y and Z.}$$

6. Find the Inverse of matrix A.

$$A = \left[\begin{array}{cc|cc|c} 8 & 3 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 5 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 8 \end{array} \right]$$

$$7. \text{ Show that } \begin{bmatrix} I & 0 \\ A & I \end{bmatrix} \text{ is invertible and find its inverse.}$$

8. Compute $X^T X$, when X is partitioned as $[X_1 \ X_2]$.

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.

$$9. (a) \left[\begin{array}{cc|cc} -1 & 2 & 1 & 5 \\ 0 & -3 & 4 & 2 \\ \hline 1 & 5 & 6 & 1 \end{array} \right] \left[\begin{array}{cc|c} -2 & 1 & 4 \\ -3 & 5 & 2 \\ \hline 7 & -1 & 5 \\ 0 & 3 & -3 \end{array} \right]$$

$$(b) \left[\begin{array}{ccc|c} -1 & 2 & 1 & 5 \\ 0 & -3 & 4 & 2 \\ \hline 1 & 5 & 6 & 1 \end{array} \right] \left[\begin{array}{cc|c} -2 & 1 & 4 \\ -3 & 5 & 2 \\ \hline 7 & -1 & 5 \\ 0 & 3 & -3 \end{array} \right]$$

$$10. (a) \left[\begin{array}{ccc|c} 3 & -1 & 0 & -3 \\ 2 & 1 & 4 & 5 \\ \hline 2 & 1 & 4 & \end{array} \right] \left[\begin{array}{ccc} 2 & -4 & 1 \\ 3 & 0 & 2 \\ \hline 1 & -3 & 5 \\ 2 & 1 & 4 \end{array} \right]$$

$$(b) \left[\begin{array}{cc} 2 & -5 \\ 1 & 3 \\ 0 & 5 \\ \hline 1 & 4 \end{array} \right] \left[\begin{array}{cc|cc} 2 & -1 & 3 & -4 \\ 0 & 1 & 5 & 7 \end{array} \right]$$

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

$$\left[\begin{array}{cc|cc} 5 & 2 & 2 & 3 \\ 2 & 1 & -3 & 1 \\ \hline 1 & 0 & 4 & 1 \\ 0 & 1 & 0 & 2 \end{array} \right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 0 \\ 0 \end{bmatrix}$$

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 $AB = \text{col}_1(A)\text{row}_1(B) + \text{col}_2(A)\text{row}_2(B) + \text{col}_3(A)\text{row}_3(B)$

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.

$$A = \begin{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix} \\ L & \end{matrix} \begin{matrix} \begin{bmatrix} \bullet & * & * & * & * \\ 0 & \bullet & * & * & * \\ 0 & 0 & 0 & \bullet & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ U & \end{matrix}$$

LU factorization.

Remarks

- 1) If A is the square matrix of order m , then the order of both L and U will also be m .
- 2) In general, not every square matrix A has an LU-decomposition, nor is an LU-decomposition unique if it exists.

Theorem If a square matrix A can be reduced to row echelon form with no row interchanges, then A has an LU-decomposition.

Note

The computational efficiency of the LU factorization depends on knowing L and U . The next algorithm shows that the row reduction of A to an echelon form U amounts to an LU factorization because it produces L with essentially no extra work.

An LU Factorization Algorithm

Suppose A can be reduced to an echelon form U without row interchanges. Then, since row scaling is not essential, A can be reduced to U with only row replacements, adding a

multiple of one row to another row below it. In this case, there exist lower triangular elementary matrices $E_1, \dots, E_p$ such that

$$E_p \dots E_1 A = U \quad (1)$$

So $A = (E_p \dots E_1)^{-1} U = LU$

Where $L = (E_p \dots E_1)^{-1}$ (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 $E_p \dots E_1 L = (E_p \dots E_1)(E_p \dots E_1)^{-1} = I$. 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 .

$$\begin{bmatrix} \boxed{1} & -\frac{1}{3} & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \leftarrow \text{multiplier} = \frac{1}{6}$$

$$\begin{bmatrix} 6 & -2 & 0 \\ \boxed{0} & 2 & 1 \\ \boxed{0} & 8 & 5 \end{bmatrix} \leftarrow \begin{matrix} \text{multiplier} = -9 \\ \text{multiplier} = -3 \end{matrix}$$

$$\begin{bmatrix} 6 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix} \qquad \begin{bmatrix} 6 & 0 & 0 \\ 9 & * & 0 \\ 3 & * & * \end{bmatrix}$$

$$\begin{aligned}
 & \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 8 & 5 \end{bmatrix} \leftarrow \text{multiplier} = \frac{1}{2} & \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & * & * \end{bmatrix} \\
 & \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & \boxed{0} & 1 \end{bmatrix} \leftarrow \text{multiplier} = -8 & \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & * \end{bmatrix} \\
 U = & \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & \boxed{1} \end{bmatrix} \leftarrow \text{multiplier} = 1 & L = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix}
 \end{aligned}$$

(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 .

$$\begin{aligned}
 & \begin{bmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \approx \begin{bmatrix} \frac{6}{6} & \frac{-2}{6} & \frac{0}{6} \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \quad \frac{1}{6}R_1 & \begin{bmatrix} 6 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix} \\
 & = \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
 &\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 9-9(1) & -1-9(\frac{-1}{3}) & 1-9(0) \\ 3-3(1) & 7-3(\frac{-1}{3}) & 5-3(0) \end{bmatrix} \begin{array}{l} R_2 - 9R_1 \\ R_3 - 3R_1 \end{array} \quad \begin{bmatrix} 6 & 0 & 0 \\ 9 & * & 0 \\ 3 & * & * \end{bmatrix} \\
 &= \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 2 & 1 \\ 0 & 8 & 5 \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
 &\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 2 & 1 \\ 0 & 8 & 5 \end{bmatrix} \frac{1}{2}R_2 \quad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & * & * \end{bmatrix} \\
 &= \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 8 & 5 \end{bmatrix} \\
 &\approx \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0-8(0) & 8-8(1) & 5-8(\frac{1}{2}) \end{bmatrix} R_3 - 8R_2 \quad \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & * \end{bmatrix} \\
 &= \begin{bmatrix} 1 & \frac{-1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}
 \end{aligned}$$

$$U = \begin{bmatrix} 1 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & \boxed{1} \end{bmatrix} \leftarrow \text{multiplier} = 1 \quad L = \begin{bmatrix} 6 & 0 & 0 \\ 9 & 2 & 0 \\ 3 & 8 & 1 \end{bmatrix}$$

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

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ * & * & 1 & 0 & 0 \\ * & * & * & 1 & 0 \\ * & * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of L.

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{2} \quad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ * & * & 1 & 0 & 0 \\ * & * & * & 1 & 0 \\ * & * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix} \begin{matrix} \leftarrow \text{multiplier} -6 \\ \leftarrow \text{multiplier} -2 \\ \leftarrow \text{multiplier} -4 \\ \leftarrow \text{multiplier } 6 \end{matrix} \quad \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 & 0 \\ 2 & * & 1 & 0 & 0 \\ 4 & * & * & 1 & 0 \\ -6 & * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{3}$$

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

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix} \begin{array}{l} \leftarrow \text{multiplier } 3 \\ \leftarrow \text{multiplier } -6 \\ \leftarrow \text{multiplier } 9 \end{array}$$

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

$$\begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 10 \end{bmatrix} \leftarrow \text{multiplier } -\frac{1}{5}$$

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

$$U = \begin{bmatrix} 1 & -2 & -1 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \leftarrow \text{multiplier } -10 \quad L = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ 4 & 6 & 0 & -5 & 0 \\ -6 & -9 & 0 & 10 & 1 \end{bmatrix}$$

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

$$A = \begin{bmatrix} 6 & -2 & -4 & 4 \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of L .

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{6}$$

$$\begin{bmatrix} 6 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & -2 & -4 & -1 \\ 0 & 4 & 13 & 0 \\ 0 & -2 & -14 & 11 \end{bmatrix} \begin{array}{l} \leftarrow \text{multiplier } -3 \\ \leftarrow \text{multiplier } 12 \\ \leftarrow \text{multiplier } 6 \end{array}$$

$$\begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -12 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 4 & 13 & 0 \\ 0 & -2 & -14 & 11 \end{bmatrix} \leftarrow \text{multiplier } -\frac{1}{2}$$

$$\begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & * & 1 & 0 \\ -6 & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 5 & -2 \\ 0 & 0 & -10 & 12 \end{bmatrix} \begin{array}{l} \leftarrow \text{multiplier } -4 \\ \leftarrow \text{multiplier } 2 \end{array} \quad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 1 & 0 \\ -6 & -2 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & -10 & 12 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{5} \quad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 8 \end{bmatrix} \leftarrow \text{multiplier } 10 \quad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 1 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{8} \quad L = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 8 \end{bmatrix}$$

$$\text{Thus } A = LU = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ -12 & 4 & 5 & 0 \\ -6 & -2 & -10 & 8 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

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

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

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

* denotes an unknown entry of L .

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{2}$$

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} \begin{array}{l} \leftarrow \text{multiplier } 4 \\ \leftarrow \text{multiplier } -2 \\ \leftarrow \text{multiplier } 6 \end{array}$$

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

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{3}$$

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

$$\begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \begin{array}{l} \leftarrow \text{multiplier } 9 \\ \leftarrow \text{multiplier } -12 \end{array}$$

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & -9 & 1 & 0 \\ -6 & 12 & * & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & \frac{7}{4} \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{4}$$

$$L = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & -9 & 1 & 0 \\ -6 & 12 & 0 & 4 \end{bmatrix}$$

$$\text{Thus } A = LU = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & -9 & 1 & 0 \\ -6 & 12 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & -\frac{1}{2} & \frac{5}{2} & -1 \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & \frac{7}{4} \end{bmatrix}$$

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 $A^{-1} = [x_1 \ x_2 \ \cdots \ x_n]$ be its unknown inverse partitioned into column vectors, and let $I = [e_1 \ e_2 \ \cdots \ e_n]$ be then $n \times n$ identity matrix partitioned into column vectors. The matrix equation $AA^{-1} = I$ can be expressed as

$$A[x_1 \ x_2 \ \cdots \ x_n] = [e_1 \ e_2 \ \cdots \ e_n]$$

$$[Ax_1 \ Ax_2 \ \cdots \ Ax_n] = [e_1 \ e_2 \ \cdots \ e_n]$$

which tells us that the unknown column vectors of A^{-1} can be obtained by solving the n -linear systems.

$$Ax_1 = e_1, \ Ax_2 = e_2, \ \cdots, \ Ax_n = e_n \quad (1^*)$$

As discussed above, this can be done by finding an **LU**-decomposition of **A**, and then using that decomposition to solve each of the n systems in (1*).

Solving Linear System by LU-Factorization

When **A** = **LU**, the equation **Ax** = **b** can be written as **L(Ux)** = **b**. Writing **y** for **Ux**, we can find **x** by solving the pair of equations; **Ly** = **b** and **Ux** = **y** (2*)

First solve **Ly** = **b** for **y** and then solve **Ux** = **y** for **x**. Each equation is easy to solve because **L** and **U** are triangular.

Procedure

Step 1: Rewrite the system $A \mathbf{x} = \mathbf{b}$ as $LU \mathbf{x} = \mathbf{b}$ (3*)

Step 2: Define a new unknown $\mathbf{y}$ by letting $U \mathbf{x} = \mathbf{y}$ (4*)
And rewrite (3*) as $L \mathbf{y} = \mathbf{b}$

Step 3: Solve the system $L \mathbf{y} = \mathbf{b}$ for the unknown $\mathbf{y}$.

Step 4: Substitute the known vector $\mathbf{y}$ into (4*) and solve for $\mathbf{x}$.

This procedure is called the method of **LU-Decomposition**.

Although **LU-Decomposition** converts the problem of solving the single system $A \mathbf{x} = \mathbf{b}$ into the problem of solving the two systems, $L \mathbf{y} = \mathbf{b}$ and $U \mathbf{x} = \mathbf{y}$, these systems are easy to solve because their co-efficient matrices are triangular.

Example 5 Solve the given system ($A\mathbf{x} = \mathbf{b}$) by **LU-Decomposition**

$$\begin{aligned} 2x_1 + 6x_2 + 2x_3 &= 2 \\ -3x_1 - 8x_2 &= 2 \\ 4x_1 + 9x_2 + 2x_3 &= 3 \end{aligned} \quad (1)$$

Solution We express the system (1) in matrix form:

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$$

$$\mathbf{A} \quad \mathbf{x} = \mathbf{b}$$

We derive an **LU-decomposition** of A .

$$A = \begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{2} \quad \begin{bmatrix} 2 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & -3 & -2 \end{bmatrix} \begin{array}{l} \leftarrow \text{multiplier } 3 \\ \leftarrow \text{multiplier } -4 \end{array} \quad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & * & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow \text{multiplier } 3 \quad \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \leftarrow \text{multiplier } \frac{1}{7} \quad L = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix}$$

Thus

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$\mathbf{A} = \mathbf{L} \mathbf{U}$$

From (2) we can rewrite this system as

$$\begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \quad (3)$$

$$\mathbf{L} \mathbf{U} \mathbf{x} = \mathbf{b}$$

As specified in Step 2 above, let us define y_1, y_2 and y_3 by the equation

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \quad (4)$$

$$\mathbf{U} \mathbf{x} = \mathbf{y}$$

which allows us to rewrite (3) as

$$\begin{bmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \quad (5)$$

$$\mathbf{L} \mathbf{y} = \mathbf{b}$$

$$2y_1 = 2$$

or equivalently, as $-3y_1 + y_2 = 2$

$$4y_1 - 3y_2 + 7y_3 = 3$$

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

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

Use this LU factorization of A to solve $Ax = b$, where $b = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$

Solution The solution of $Ly = b$ needs only 6 multiplications and 6 additions, because the arithmetic takes place only in column 5. (The zeros below each pivot in L are created automatically by our choice of row operations.)

$$[L \ b] = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = [I \ y]$$

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.)

$$[U \ y] = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \quad x = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

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 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 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 A^{-1} and $A^{-1}b$).
4. If A is sparse (with mostly zero entries), then L and U may be sparse, too, whereas A^{-1} is likely to be dense. In this case, a solution of $Ax = b$ with an LU factorization is much faster than using A^{-1} .

Example 7(Gaussian Elimination Performed as an LU-Decomposition)

In Example 5, we showed how to solve the linear system

$$\begin{bmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \quad (6)$$

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 $y = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix}$ 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).

$$[A|b] = \left[\begin{array}{ccc|c} 2 & 6 & 2 & 2 \\ -3 & -8 & 0 & 2 \\ 4 & 9 & 2 & 3 \end{array} \right] \quad \left[\begin{array}{ccc} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{array} \right] = L \quad (* = \text{unknown entries})$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 1 \\ -3 & -8 & 0 & 2 \\ 4 & 9 & 2 & 3 \end{array} \right] \quad \left[\begin{array}{ccc} 2 & 0 & 0 \\ * & * & 0 \\ * & * & * \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & -3 & -2 & -1 \end{array} \right] \quad \left[\begin{array}{ccc} 2 & 0 & 0 \\ -3 & * & 0 \\ 4 & * & * \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 7 & 14 \end{array} \right] \quad \left[\begin{array}{ccc} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & * \end{array} \right]$$

$$[U|y] = \left[\begin{array}{ccc|c} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{array} \right] \quad \left[\begin{array}{ccc} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -3 & 7 \end{array} \right] = L$$

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 Example 5.

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 $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} \rightarrow \begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$ is linear. That is, there is a matrix **A**, called the transfer matrix, such that $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix} = A \begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$

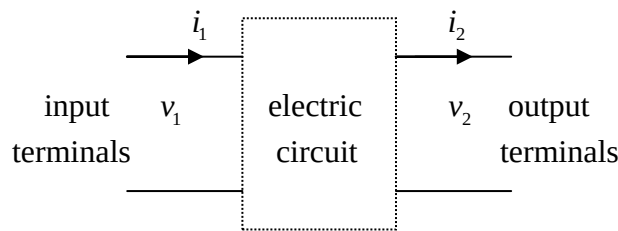


Figure A circuit with input and output terminals.

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:

$$\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}$$

Transfer matrix Transfer matrix
of series circuit 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 $\begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$.

Solution

- a) Let A_1 and A_2 be the transfer matrices of the series of the series and shunt circuits, respectively. Then an input vector $\mathbf{x}$ is transformed first into $A_1\mathbf{x}$ and then into $A_2(A_1\mathbf{x})$. 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 $\begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$ into the product of transfer matrices, such as in (6). So we look for R_1 and R_2 to satisfy

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$$

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,

$$x^{(0)}, x^{(1)}, \dots, x^{(k)}, \dots$$

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

$$\boxed{Mx^{(k+1)} = Nx^{(k)} + b \quad (k = 0, 1, \dots)} \quad (1)$$

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 $N = M - A$,

$$\therefore (1) \Rightarrow Mx^{(k+1)} = (M - A)x^{(k)} + b \quad (k = 0, 1, \dots)$$

For simplicity, we take the zero vector as $x^{(0)}$ as the initial approximation.

Example 1

Apply Jacobi's method to the system

$$\begin{aligned}
10x_1 + x_2 - x_3 &= 18 \\
x_1 + 15x_2 + x_3 &= -12 \\
-x_1 + x_2 + 20x_3 &= 17
\end{aligned} \tag{2}$$

Take $\mathbf{x}^{(0)} = (0, 0, 0)$ as an initial approximation to the solution, and use six iterations (that is, compute $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(6)}$).

Solution

For some k , let $\mathbf{x}^{(k)} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (x_1, x_2, x_3)$ and $\mathbf{x}^{(k+1)} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = (y_1, y_2, y_3)$

Firstly we will construct M and N from 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: $M\mathbf{x}^{(k+1)} = (M - A)\mathbf{x}^{(k)} + b$ (here $k = 0, 1, \dots, 6$)

implies

$$\begin{aligned}
\begin{bmatrix} 10 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 20 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} &= \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix} \\
\Rightarrow \begin{bmatrix} 10y_1 \\ 15y_2 \\ 20y_3 \end{bmatrix} &= \begin{bmatrix} 0x_1 - 1x_2 + 1x_3 \\ -1x_1 + 0x_2 - x_3 \\ 1x_1 - 1x_2 + 0x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix} \\
\Rightarrow \begin{bmatrix} 10y_1 \\ 15y_2 \\ 20y_3 \end{bmatrix} &= \begin{bmatrix} 0x_1 - 1x_2 + 1x_3 + 18 \\ -1x_1 + 0x_2 - x_3 - 12 \\ 1x_1 - 1x_2 + 0x_3 + 17 \end{bmatrix}
\end{aligned}$$

Comparing the corresponding entries on both sides, we have

$$10y_1 = -x_2 + x_3 + 18$$

$$\begin{aligned} 15y_2 &= -x_1 - x_3 - 12 \\ 20y_3 &= x_1 - x_2 + 17 \end{aligned}$$

And

$$\begin{aligned} y_1 &= (-x_2 + x_3 + 18)/10 \\ y_2 &= (-x_1 - x_3 - 12)/15 \\ y_3 &= (x_1 - x_2 + 17)/20 \end{aligned} \quad (3)$$

1st Iteration

For $k = 0$, put $\mathbf{x}^{(0)} = (x_1, x_2, x_3) = (0, 0, 0)$ in (3) and compute

$$\mathbf{x}^{(1)} = (y_1, y_2, y_3) = (18/10, -12/15, 17/20) = (1.8, -0.8, 0.85)$$

2nd Iteration

For $k = 1$, put $\mathbf{x}^{(1)} = (1.8, -0.8, 0.85)$

$$\begin{aligned} y_1 &= [-(-0.8) + (0.85) + 18]/10 = 1.965 \\ y_2 &= [-(1.8) - (0.85) - 12]/15 = -0.9767 \\ y_3 &= [(1.8) - (-0.8) + 17]/20 = 0.98 \end{aligned}$$

Thus, $\mathbf{x}^{(2)} = (1.965, -0.9767, 0.98)$.

The entries in $\mathbf{x}^{(2)}$ are used on the right in (3) to compute the entries in $\mathbf{x}^{(3)}$, and so on. Here are the results, with calculations using **MATLAB** and results reported to four decimal places:

$\mathbf{x}^{(0)}$	$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	$\mathbf{x}^{(3)}$	$\mathbf{x}^{(4)}$	$\mathbf{x}^{(5)}$	$\mathbf{x}^{(6)}$
$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1.8 \\ -0.8 \\ 0.85 \end{bmatrix}$	$\begin{bmatrix} 1.965 \\ -0.9767 \\ 0.98 \end{bmatrix}$	$\begin{bmatrix} 1.9957 \\ -0.9963 \\ 0.9971 \end{bmatrix}$	$\begin{bmatrix} 1.9993 \\ -0.9995 \\ 0.9996 \end{bmatrix}$	$\begin{bmatrix} 1.9999 \\ -0.9999 \\ 0.9999 \end{bmatrix}$	$\begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix}$

If we decide to stop when the entries in $\mathbf{x}^{(k)}$ and $\mathbf{x}^{(k-1)}$ differ by less than .001, then we need five iterations ($k = 5$).

Alternative Approach

If we express the above system as

$$10x_1 + x_2 - x_3 = 18 \Rightarrow x_1 = \frac{18 - x_2 + x_3}{10}$$

$$x_1 + 15x_2 + x_3 = -12 \Rightarrow x_2 = \frac{-12 - x_1 - x_3}{15}$$

$$-x_1 + x_2 + 20x_3 = 17 \Rightarrow x_3 = \frac{17 + x_1 - x_2}{20}$$

$\therefore$ the equivalent system is

$$x_1 = \frac{18 - x_2 + x_3}{10}$$

$$x_2 = \frac{-12 - x_1 - x_3}{15}$$

$$x_3 = \frac{17 + x_1 - x_2}{20}$$

Now put $(x_1, x_2, x_3) = (0, 0, 0) = \mathbf{x}^{(0)}$ in the RHS to have

$$x_1 = (18 - 0 + 0)/10 = 1.80$$

$$x_2 = (-12 - 0 - 0)/15 = -0.80$$

$$x_3 = (17 + 0 - 0)/20 = 0.85$$

Which gives $\mathbf{x}^{(1)} = (1.80, -0.80, 0.85)$ ----put this again on RHS of the equivalent system to get

$$x_1 = (18 + 0.80 + 0.85)/10 = 1.965$$

$$x_2 = (-12 - 1.80 - 0.85)/15 = -0.9767$$

$$x_3 = (17 + 1.80 + 0.80)/20 = 0.98$$

So in the similar fashion, we can get the next approximate solutions: $\mathbf{x}^{(3)}, \mathbf{x}^{(4)}, \mathbf{x}^{(5)}$ and $\mathbf{x}^{(6)}$

Next example will be solved by following this approach.

Example 2

Use Jacobi iteration to approximate the solution of the system

$$20x_1 + x_2 - x_3 = 17$$

$$x_1 - 10x_2 + x_3 = 13$$

$$-x_1 + x_2 + 10x_3 = 18$$

Stop the process when the entries in two successive iterations are the same when rounded to four decimal places.

Solution

As required for Jacobi iteration, we begin by solving the first equation for x_1 , the second for x_2 , and the third for x_3 . This yields

$$x_1 = \frac{17}{20} - \frac{1}{20}x_2 + \frac{1}{20}x_3 = 0.85 - 0.05x_2 + 0.05x_3$$

$$x_2 = -\frac{13}{10} + \frac{1}{10}x_1 + \frac{1}{10}x_3 = -1.3 + 0.1x_1 + 0.1x_3 \quad (4)$$

$$x_3 = \frac{18}{10} + \frac{1}{10}x_1 - \frac{1}{10}x_2 = 1.8 + 0.1x_1 - 0.1x_2$$

which we can write in matrix form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -0.05 & 0.05 \\ 0.1 & 0 & 0.1 \\ 0.1 & -0.1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0.85 \\ -1.3 \\ 1.8 \end{bmatrix} \quad (5)$$

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.

$$A = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \quad M = \begin{bmatrix} * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & * & 0 \\ * & * & * & * & * \end{bmatrix}$$

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.

$$\begin{aligned} 10x_1 + x_2 - x_3 &= 18 \\ x_1 + 15x_2 + x_3 &= -12 \\ -x_1 + x_2 + 20x_3 &= 17 \end{aligned} \quad (6)$$

Solution

For some k , let $\mathbf{x}^{(k)} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (x_1, x_2, x_3)$ and $\mathbf{x}^{(k+1)} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = (y_1, y_2, y_3)$

Again, firstly we construct matrices $\mathbf{M}$ and $\mathbf{N}$ from the coefficient matrix $\mathbf{A}$.

Here $\mathbf{A} = \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix}$

Since matrix $\mathbf{M}$ is constructed by

- 1) taking the values along the diagonal and below the diagonal of coefficient matrix $\mathbf{A}$.
- 2) putting the zeros above the diagonal at upper triangular position.

So

$$\mathbf{M} = \begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix}$$

Now,

$$\mathbf{N} = \mathbf{M} - \mathbf{A} = \begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix} - \begin{bmatrix} 10 & 1 & -1 \\ 1 & 15 & 1 \\ -1 & 1 & 20 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Now the recursion: $\mathbf{M}\mathbf{x}^{(k+1)} = (\mathbf{M} - \mathbf{A})\mathbf{x}^{(k)} + \mathbf{b}$ (here $k = 0, 1, \dots, 6$)

implies

$$\begin{aligned} \begin{bmatrix} 10 & 0 & 0 \\ 1 & 15 & 0 \\ -1 & 1 & 20 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} &= \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix} \\ \Rightarrow \begin{bmatrix} 10y_1 & 0 & 0 \\ y_1 & 15y_2 & 0 \\ -y_1 & y_2 & 20y_3 \end{bmatrix} &= \begin{bmatrix} -x_2 + x_3 \\ -x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} 18 \\ -12 \\ 17 \end{bmatrix} \\ \Rightarrow \begin{bmatrix} 10y_1 & 0 & 0 \\ y_1 & 15y_2 & 0 \\ -y_1 & y_2 & 20y_3 \end{bmatrix} &= \begin{bmatrix} -x_2 + x_3 + 18 \\ -x_3 - 12 \\ 0 + 17 \end{bmatrix} \end{aligned}$$

Comparing the corresponding entries on both sides, we have

$$\begin{aligned} 10y_1 &= -x_2 + x_3 + 18 \\ y_1 + 15y_2 &= -x_3 - 12 \\ -y_1 + y_2 + 20y_3 &= 17 \end{aligned}$$

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:

$$\begin{aligned} x_1 &= (-x_2 + x_3 + 18)/10 \\ x_2 &= (-x_1 - x_3 - 12)/15 \\ x_3 &= (x_1 - x_2 + 17)/20 \end{aligned} \quad (8)$$

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

$$\begin{aligned} x_1 &= [- (0) + (0) + 18]/10 = 1.8 \\ x_2 &= [- (1.8) - (0) - 12]/15 = -.92 \\ x_3 &= [+(1.8) - (-.92) + 17]/20 = .986 \end{aligned}$$

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:

$x^{(0)}$	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$	$x^{(6)}$
$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1.8 \\ -.92 \\ .986 \end{bmatrix}$	$\begin{bmatrix} 1.9906 \\ -.9984 \\ .9995 \end{bmatrix}$	$\begin{bmatrix} 1.9998 \\ -.9999 \\ 1.0000 \end{bmatrix}$	$\begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix}$	$\begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix}$	$\begin{bmatrix} 2.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix}$

Observe that when k is 4, the entries in $x^{(4)}$ and $x^{(k-1)}$ differ by less than .001. The values in $x^{(6)}$ in this case happen to be accurate to eight decimal places.

Alternative Approach

If we express the above system as

$$10x_1 + x_2 - x_3 = 18 \Rightarrow x_1 = (-x_2 + x_3 + 18)/10 \text{ -----}(a)$$

$$x_1 + 15x_2 + x_3 = -12 \Rightarrow x_2 = (-x_1 - x_3 - 12)/15 \text{ -----}(b)$$

$$-x_1 + x_2 + 20x_3 = 17 \Rightarrow x_3 = (x_1 - x_2 + 17)/20 \text{ -----}(c)$$

Ist Iteration

Put $x_2=x_3=0$ in (a)

$$x_1 = 18/10 = 1.80$$

Put $x_1 = 1.80$ and $x_3 = 0$ in (b)

$$x_2 = (-1.80 - 0 - 12)/15 = -0.92$$

Put $x_1 = 1.80$, $x_2 = -0.92$ in (c)

$$x_3 = (1.80 + 0.92 + 17)/20 = 0.9863$$

$$\text{So, } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1.8 \\ -0.92 \\ .986 \end{bmatrix} = x^{(1)}$$

2nd iteration

Put $x_2 = -0.92$, $x_3 = 0.9863$ in (a)

$$x_1 = (0.92 + 0.9863 + 18)/10 = 1.9906$$

Put $x_1 = 1.9906$ (from 2nd iteration) and $x_3 = 0.9863$ (from 1st iteration) in (b)

$$x_2 = (-1.9906 - 0.9863 - 12)/15 = -0.9984$$

Put $x_1 = 1.9906$, $x_2 = -0.9984$ (both from 2nd iteration) in (c)

$$x_3 = (1.9906 + 0.9984 + 17)/20 = 0.9995$$

$$\text{So, } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1.9906 \\ -0.9984 \\ .9995 \end{bmatrix} = x^{(2)}$$

So in the similar fashion, we can get the next approximate solutions: $x^{(3)}$, $x^{(4)}$, $x^{(5)}$ and $x^{(6)}$
Next example will be solved by following this approach.

Example 4

Use Gauss-Seidel to approximate the solution of the linear system in example 2 to four decimal places.

Solution

As before, we will take $x_1 = x_2 = x_3 = 0$ as the initial approximation. First, we will substitute $x_2 = 0$ and $x_3 = 0$ into the right side of the first equation of (4) to obtain the new x_1 . Then, we will substitute $x_3 = 0$ and the new x_1 into the right side of the second equation to obtain the new x_2 , and finally, we will substitute the new x_1 and new x_2 into the right side of the third equation to obtain the new x_3 . The computations are as follows:

$$x_1 = 0.85 - (0.05)(0) + (0.05)(0) = 0.85$$

$$x_2 = -1.3 + (0.1)(0.85) + (0.1)(0) = -1.215$$

$$x_3 = 1.8 + (0.1)(0.85) - (0.1)(-1.215) = 2.0065$$

Thus, the first Gauss-Seidel iterate is

$$y_1 = \begin{bmatrix} 0.8500 \\ -1.2150 \\ 2.0065 \end{bmatrix}$$

Similarly, the computations for second iterate are

$$x_1 = 0.85 - (0.05)(-1.215) + (0.05)(2.0065) = 1.011075$$

$$x_2 = -1.3 + (0.1)(1.011075) + (0.1)(2.0065) = -0.9982425$$

$$x_3 = 1.8 + (0.1)(1.011075) - (0.1)(-0.9982425) = 2.00093175$$

Thus, the second Gauss-Seidel iterate to four decimal places is

$$y_2 \approx \begin{bmatrix} 1.0111 \\ -0.9982 \\ 2.0009 \end{bmatrix}$$

The following table shows the first four Gauss-Seidel iterates to four decimal places. Comparing both tables, we see that the Gauss-Seidel method produced the solution to four decimal places in four iterations whereas, the Jacobi method required six.

	y_0	y_1	y_2	y_3	y_4
x_1	0	0.8500	1.0111	1.0000	1.0000
x_2	0	-1.2150	-0.9982	-0.9999	-1.0000
x_3	0	2.0065	2.0009	2.0000	2.0000

Comparison of Jacobi's and Gauss-Seidel method

There exist examples where **Jacobi's method is faster than the Gauss-Seidel method**, but usually a **Gauss-Seidel sequence converges faster** (means to say iterative solution approaches to the unique solution), as in Example 2. (If parallel processing is available, **Jacobi might be faster because the entries in $x^{(k)}$ can be computed simultaneously.**) There are also examples where one or both methods fail to produce a convergent sequence, and other examples where a sequence is convergent, but **converges too slowly for practical use.**

Condition for the Convergence of both Iterative Methods

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:

$$\begin{bmatrix} -6 & 2 & -3 \\ 1 & 4 & -2 \\ 3 & -5 & 8 \end{bmatrix} \quad \begin{array}{l} |-6| > |2| + |-3| \\ |4| > |1| + |-2| \\ |8| = |3| + |-5| \end{array}$$

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:

$$\begin{array}{rcl} x_1 - 3x_2 + x_3 & = & -2 \\ -6x_1 + 4x_2 + 11x_3 & = & 1 \\ 5x_1 - 2x_2 - 2x_3 & = & 9 \end{array}$$

Solution

The system is not strictly diagonally dominant, as for the 1st row

$$|\text{coefficient of } x_1| < |\text{coefficient of } x_2| + |\text{coefficient of } x_3|$$

$$\text{or } |1| < |-3| + |1|$$

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

$$\begin{array}{rcl} 5x_1 - 2x_2 - 2x_3 & = & 9 \\ x_1 - 3x_2 + x_3 & = & -2 \\ -6x_1 + 4x_2 + 11x_3 & = & 1 \end{array}$$

So,

for 1st equation (row);

$$|\text{coefficient of } x_1| > |\text{coefficient of } x_2| + |\text{coefficient of } x_3|$$

$$\text{or } |5| > |-2| + |-2|$$

for 2nd equation(row);

$$|\text{coefficient of } x_2| > |\text{coefficient of } x_1| + |\text{coefficient of } x_3|$$

$$\text{or } |-3| > |1| + |1|$$

for 3rd equation(row);

$$|\text{coefficient of } x_3| > |\text{coefficient of } x_1| + |\text{coefficient of } x_2|$$

$$\text{or } |11| > |-6| + |4|$$

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.

$$1. \begin{cases} 4x_1 + x_2 = 7 \\ -x_1 + 5x_2 = -7 \end{cases}$$

$$2. \begin{cases} 10x_1 - x_2 = 25 \\ x_1 + 8x_2 = 43 \end{cases}$$

$$3. \begin{cases} 3x_1 + x_2 = 11 \\ -x_1 - 5x_2 + 2x_3 = 15 \\ 3x_2 + 7x_3 = 17 \end{cases}$$

$$4. \begin{cases} 50x_1 - x_2 = 149 \\ x_1 - 100x_2 + 2x_3 = -101 \\ 2x_2 + 50x_3 = -98 \end{cases}$$

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:

$$11. \begin{cases} x_1 - 3x_2 + x_3 = -2 \\ -6x_1 + 4x_2 + 11x_3 = 1 \\ 5x_1 - 2x_2 - 2x_3 = 9 \end{cases}$$

$$12. \begin{cases} -x_1 + 4x_2 - x_3 = 3 \\ 4x_1 - x_2 = 10 \\ -x_2 + 4x_3 = 6 \end{cases}$$

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.

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

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{vmatrix} = a_{11}b_{22} - a_{12}b_{21}$$

This is called a 2x2 determinant.

The determinant of a 3x3 matrix is also called a 3x3 determinant is defined by the following formula.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad (1)$$

For finding the determinant of the 3x3 matrix, we look at the following diagram:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

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 $A = \begin{bmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{bmatrix}$

$$\det A = \begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} \begin{vmatrix} 1 & 2 \\ -4 & 5 \\ 7 & -8 \end{vmatrix}$$

$$= 1 \times 5 \times 9 + 2 \times 6 \times 7 + 3 \times (-4) \times (-8) - 3 \times 5 \times 7 - 1 \times 6 \times (-8) - 2 \times (-4) \times 9$$

$$= 45 + 84 + 96 - 105 + 48 + 72$$

$$= 240$$

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 A is row reduced.

Gauss' algorithm for evaluation of determinants

1) Firstly, we apply it for 2×2 matrix say

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

reducing to echelon form and then multiply only the values in diagonal.

$R'_2 \rightarrow R_2 - 2R_1$ (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

$$\text{Det}(A) = 2(-3) - 0 \times 3 = -6 - 0 = -6$$

2) For 3×3 matrix say

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

By $R'_{12} \rightarrow R_{21}$ (Interchanging of 1st and 2nd rows)

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

$R'_2 \rightarrow R_2 - 2R_1$ (Multiplying 1st row by '-2' and then adding in the 2nd row)

$R'_3 \rightarrow R_3 + 2R_1$ (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 $R'_{23} \rightarrow R_{32}$ (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

$$\text{Det}(B) = (-1) \cdot 2 \cdot (-9) = 18$$

So in general,

For a 1×1 matrix

say, $A = [a_{ij}]$ - we define $\det A = a_{11}$.

For 2×2 matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

By $R'_2 \rightarrow R_2 - \left(\frac{a_{21}}{a_{11}}\right)R_1$ provided that $a_{11} \neq 0$

$$\sim \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}}a_{12} \end{bmatrix}$$

$\therefore \Delta = \det A = \text{product of the diagonal entries}$

$$= a_{11} \left(a_{22} - \frac{a_{21}}{a_{11}} a_{12} \right) = a_{11} a_{22} - a_{12} a_{21}$$

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 $R_2' \rightarrow R_2 - \left(\frac{a_{21}}{a_{11}} \right) R_1$, $R_3' \rightarrow R_3 - \left(\frac{a_{31}}{a_{11}} \right) R_1$ 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 $R_3' \rightarrow R_3 - \left(\frac{\frac{a_{32}a_{11} - a_{12}a_{31}}{a_{11}}}{\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}}} \right) R_2$ provided that $\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} \neq 0$

$$\sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} & \frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}} \\ 0 & 0 & \frac{a_{11}a_{33} - a_{13}a_{31}}{a_{11}} - \left(\frac{\frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}}}{\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}}} \right) \left(\frac{\frac{a_{32}a_{11} - a_{12}a_{31}}{a_{11}}}{\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}}} \right) \end{bmatrix} \because a_{11} \neq 0$$

Which is in echelon form. Now,

$\Delta = \det A =$ product of the diagonal entries

$$\begin{aligned}
&= a_{11} \left(\frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} \right) \left(\frac{a_{11}a_{33} - a_{13}a_{31}}{a_{11}} - \left(\frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}} \right) \left(\frac{a_{32}a_{11} - a_{12}a_{31}}{a_{22}a_{11} - a_{12}a_{21}} \right) \right) \\
&= (a_{22}a_{11} - a_{12}a_{21}) \left(\frac{a_{11}a_{33} - a_{13}a_{31}}{a_{11}} \right) - (a_{22}a_{11} - a_{12}a_{21}) \left(\frac{a_{23}a_{11} - a_{13}a_{21}}{a_{11}} \right) \left(\frac{a_{32}a_{11} - a_{12}a_{31}}{a_{22}a_{11} - a_{12}a_{21}} \right) \\
&= \frac{1}{a_{11}} \{ (a_{22}a_{11} - a_{12}a_{21})(a_{11}a_{33} - a_{13}a_{31}) - (a_{23}a_{11} - a_{13}a_{21})(a_{32}a_{11} - a_{12}a_{31}) \} \\
&= \frac{1}{a_{11}} \{ a_{11}^2 a_{22} a_{33} - a_{11} a_{22} a_{13} a_{31} - a_{12} a_{21} a_{11} a_{33} + a_{12} a_{21} a_{13} a_{31} - a_{23} a_{11}^2 a_{32} + a_{23} a_{11} a_{12} a_{31} + a_{13} a_{21} a_{32} a_{11} - a_{12} a_{21} a_{13} a_{31} \} \\
&= \frac{1}{a_{11}} \{ a_{11}^2 a_{22} a_{33} - a_{11} a_{22} a_{13} a_{31} - a_{12} a_{21} a_{11} a_{33} - a_{23} a_{11}^2 a_{32} + a_{23} a_{11} a_{12} a_{31} + a_{13} a_{21} a_{32} a_{11} \} \\
&= \frac{a_{11}}{a_{11}} \{ a_{11} a_{22} a_{33} - a_{22} a_{13} a_{31} - a_{12} a_{21} a_{33} - a_{23} a_{11} a_{32} + a_{23} a_{12} a_{31} + a_{13} a_{21} a_{32} \} \\
&= a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{12} a_{21} a_{33} - a_{11} a_{23} a_{32} - a_{13} a_{22} a_{31}
\end{aligned}$$

□

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:

$$\begin{aligned}
\Delta &= (a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32}) - (a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33}) + (a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}) \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
\Delta &= a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\
\Delta &= a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}
\end{aligned}$$

$$\text{For brevity, we write } \Delta = a_{11} \cdot \det A_{11} - a_{12} \cdot \det A_{12} + a_{13} \cdot \det A_{13} \quad (3)$$

$$\det(A_{11}) = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad \det(A_{12}) = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \quad \text{and} \quad \det(A_{13}) = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

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 A_{ij} denote the sub-matrix formed by deleting the i th row and j th 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 $A = \begin{bmatrix} 1 & 4 & 3 \\ 5 & 2 & 4 \\ 3 & 6 & 3 \end{bmatrix}$

$$\begin{aligned}
 |A| &= \begin{vmatrix} 1 & 4 & 3 \\ 5 & 2 & 4 \\ 3 & 6 & 3 \end{vmatrix} \\
 &= 1 \begin{vmatrix} 2 & 4 \\ 6 & 3 \end{vmatrix} - 4 \begin{vmatrix} 5 & 4 \\ 3 & 3 \end{vmatrix} + 3 \begin{vmatrix} 5 & 2 \\ 3 & 6 \end{vmatrix} \\
 &= 1(2 \times 3 - 4 \times 6) - 4(5 \times 3 - 4 \times 3) + 3(5 \times 6 - 2 \times 3) \\
 &= 1(6 - 24) - 4(15 - 12) + 3(30 - 6) \\
 &= 1(-18) - 4(3) + 3(24) \\
 &= -18 - 12 + 72 \\
 &= 42
 \end{aligned}$$

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 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 $n = 3$, $\det A$ is defined using determinants of the 2×2 submatrices A_{ij} .

When $n = 4$, $\det A$ uses determinants of the 3×3 submatrices A_{ij} .

In general, an $n \times n$ 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 $a_{11}, a_{12}, \dots, a_{1n}$ are from the first row of A .

Here for a_{ij} ,

$$i = 1, 2, 3, \dots, n \quad (1 \leq i \leq n)$$

$$j = 1, 2, 3, \dots, n \quad (1 \leq j \leq n)$$

$$\text{In symbols, } \det A = a_{11} \det A_{11} - a_{12} \det A_{12} + \dots + (-1)^{1+n} a_{1n} \det A_{1n} = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det A_{1j}$$

Example 4

$$\text{Compute the determinant of } A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Solution

Here A is $n \times n = 3 \times 3$ matrix such that

$$i = 1, 2, 3$$

$$j = 1, 2, 3$$

$$\therefore \det(A) = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det A_{1j} \text{ and here } j = 1, 2, 3$$

$$\begin{aligned} \therefore \det(A) &= \sum_{j=1}^3 (-1)^{1+j} a_{1j} \det A_{1j} = (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} + (-1)^{1+3} a_{13} \det A_{13} \\ &= a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} \end{aligned}$$

$$\det A = 1 \cdot \det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5 \cdot \det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0 \cdot \det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix}$$

$$\det A = 1 \cdot \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5 \cdot \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0 \cdot \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

$$= 1 [4(0) - (-1)(-2)] - 5 [2(0) - 0(-1)] + 0[2(-2) - 4(0)]$$

$$= 1(0 - 2) - 5(0 - 0) + 0(-4 - 0) = -2$$

Minor of an element

If A is a square matrix, then the **Minor** of entry a_{ij} (called the ij th minor of A) is denoted by M_{ij} and is defined to be the determinant of the sub matrix that remains when the i th row and j th 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 ij th 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:

$$C_{11} = (-1)^{1+1} M_{11}, \quad C_{12} = (-1)^{1+2} M_{12}, \quad C_{13} = (-1)^{1+3} M_{13}$$

$$C_{11} = (-1)^{1+1} \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix}, \quad C_{12} = (-1)^{1+2} \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix}, \quad C_{13} = (-1)^{1+3} \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix}$$

Example 5 Find the minor and 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

$$M_{11} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} = 5 \times 8 - 6 \times 4 = 40 - 24 = 16$$

and the corresponding cofactor is

$$C_{11} = (-1)^{1+1} M_{11} = M_{11} = 16$$

The minor of entry a_{32} is

$$M_{32} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = 26$$

and the corresponding cofactor is

$$C_{32} = (-1)^{3+2} M_{32} = -M_{32} = -\begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = -26$$

Alternate Definition

Given $A = [a_{ij}]$, the (i, j) -cofactor of A is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij} \quad (4)$$

Then $\det A = a_{11}C_{11} + a_{12}C_{12} + \dots + a_{1n}C_{1n}$

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,

$$\begin{aligned}
\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} &= (2)(-1)^{1+2} \begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(-1)^{2+2} \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1)^{3+2} \begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix} \\
&= (2)(-1)^3 \begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(-1)^4 \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1)^5 \begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix} \\
&= (2)(-1) \begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + (5)(1) \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-8)(-1) \begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix} \\
&= -2 \begin{vmatrix} -4 & 6 \\ 7 & 9 \end{vmatrix} + 5 \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + 8 \begin{vmatrix} 1 & 3 \\ -4 & 6 \end{vmatrix} \\
&= -2(-36 - 42) + 5(9 - 21) + 8(6 - (-12)) \\
&= (-2)(-78) + (5)(-12) + (8)(18) = 240
\end{aligned}$$

Theorem 1 The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any row or down any column. The expansion across the i th 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 j th 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:

+	-	+	
-	+	-	
+	-	+	
⋮			⋱

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 $\det \mathbf{A}$, where $\mathbf{A} = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$

Solution The cofactor expansion down the first column of $\mathbf{A}$ has all terms equal to zero except the first.

Thus $\det \mathbf{A} = 3 \begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} - 0.C_{21} + 0.C_{31} - 0.C_{41} + 0.C_{51}$

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 $\det \mathbf{A} = 3 \times 2 \begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}$

This 3×3 determinant was computed above and found to equal -2 .

Hence, $\det \mathbf{A} = 3 \times 2 \times (-2) = -12$.

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 $m \times n$ 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 $25 \times 25 \begin{bmatrix} 1 & 0 & 0 \\ 2 & 8 & 0 \\ 4 & 9 & 7 \end{bmatrix}$ is lower triangular matrices.

Determinants of Triangular Matrices

Determinants of the triangular matrices are also easy to evaluate regardless of size.

Theorem If A is triangular matrix, then $\det(A)$ 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 $25! \sim 1.5 \times 10^{25}$.

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 $\det EA = (\det E) \cdot (\det A)$, where E is the elementary matrix and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

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

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

$$9. \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$$

$$10. \begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$11. \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$12. \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

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

$$14. \text{ Let } A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}. \text{ Write } 5A. \text{ Is } \det 5A = 5 \det A?$$

$$15. \text{ Let } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } k \text{ be a scalar. Find a formula that relates } \det (kA) \text{ to } k \text{ and } \det A.$$

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.

- If a multiple of one row of A is added to another row, the resulting determinant will remain same.
- If two rows of A are interchanged to produce B , then $\det B = -\det A$.
- If one row of A is multiplied by k to produce B , then $\det B = k \cdot \det A$.

The following examples show how to use Theorem 3 to find determinants efficiently.

- 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

$$= \begin{vmatrix} a_{11} + ka_{21} & a_{12} + ka_{22} & a_{13} + ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad R_1' \rightarrow R_1 + kR_2$$

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

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} ka_{21} & ka_{22} & ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + k \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad \text{By using property (c) of above theorem 3.}$$

If any two rows or columns in a determinant are identical then value of this determinant is zero. So in this case $R_1 \equiv R_2$

$$\begin{aligned}\therefore \Delta &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + k(0) \\ &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = A\end{aligned}$$

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 $\det B = k \cdot \det A$.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 0 & 1 \\ 0 & 8 & 9 \end{bmatrix}$$

$$\begin{aligned}|A| &= 1(0 - 8) - 2(45 - 0) + 3(40 - 0) \\ &= -8 - 90 + 120 = 22\end{aligned}$$

Multiplying R_1 by k , we get say

$$B = \begin{bmatrix} 1k & 2k & 3k \\ 5 & 0 & 1 \\ 0 & 8 & 9 \end{bmatrix}$$

$$\begin{aligned}
 |B| &= k(40 - 0) - 2k(45 - 0) + 3k(40 - 0) \\
 &= 40k - 90k + 120k = 22k \\
 &= k|A|
 \end{aligned}$$

Example 2

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

Solution

$$\begin{aligned}
 \det A &= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} \\
 &= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -7 \\ 0 & -2 & -8 & -10 \\ 0 & -7 & -10 & -13 \end{vmatrix} \quad \text{by } R_2' \rightarrow R_2 + (-2)R_1, R_3' \rightarrow R_3 + (-3)R_1, R_4' \rightarrow R_4 + (-4)R_1 \\
 &= \begin{vmatrix} -1 & -2 & -7 \\ -2 & -8 & -10 \\ -7 & -10 & -13 \end{vmatrix} \quad \text{expanding from 1st column} \\
 &= (-1)(-2)(-1) \begin{vmatrix} 1 & 2 & 7 \\ 1 & 4 & 5 \\ 7 & 10 & 13 \end{vmatrix} \quad \text{taking } (-1), (-2) \text{ and } (-1) \text{ common from 1st, 2nd, 3rd rows} \\
 &= (-2) \begin{vmatrix} 1 & 2 & 7 \\ 0 & 2 & -2 \\ 0 & -4 & -36 \end{vmatrix} \quad \text{by } R_2' \rightarrow R_2 + (-1)R_1, R_3' \rightarrow R_3 + (-7)R_1 \\
 &= (-2) \begin{vmatrix} 2 & -2 \\ -4 & -36 \end{vmatrix} \quad \text{expanding by 1st column} \\
 &= (-2)(2)(-4) \begin{vmatrix} 1 & -1 \\ 1 & 9 \end{vmatrix} \quad \text{taking 2 and } (-4) \text{ common from 1st and 2nd rows respectively.} \\
 &= 16 \begin{vmatrix} 1 & -1 \\ 0 & 10 \end{vmatrix} \quad \text{by } R_2 \rightarrow R_2 + (-1)R_1 \\
 &= 160
 \end{aligned}$$

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

$$\det A = \begin{vmatrix} 4 & 2 & 5 & 10 \\ 1 & 1 & 6 & 3 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{vmatrix}$$

$$= - \begin{vmatrix} 1 & 1 & 6 & 3 \\ 4 & 2 & 5 & 10 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{vmatrix} \quad \text{interchanging } R_1 \text{ and } R_2 (R'_{12})$$

$$= - \begin{vmatrix} 1 & 1 & 6 & 3 \\ 0 & -2 & -19 & -2 \\ 0 & -4 & -42 & -16 \\ 0 & 2 & 5 & 8 \end{vmatrix} \quad \text{By } R'_2 \rightarrow R_2 + (-4)R_1, R'_3 \rightarrow R_3 + (-7)R_1$$

$$= - \begin{vmatrix} -2 & -19 & -2 \\ -4 & -42 & -16 \\ 2 & 5 & 8 \end{vmatrix} \quad \text{expanding from 1st column}$$

$$= (-1)^3 \begin{vmatrix} 2 & 19 & 2 \\ 4 & 42 & 16 \\ 2 & 5 & 8 \end{vmatrix} \quad \text{taking } (-1) \text{ as a common factor from } R_1 \text{ and } R_2$$

$$= - \begin{vmatrix} 2 & 19 & 2 \\ 4 & 42 & 16 \\ 2 & 5 & 8 \end{vmatrix}$$

$$= -2 \begin{vmatrix} 1 & 19 & 2 \\ 2 & 42 & 16 \\ 1 & 5 & 8 \end{vmatrix}$$

$$= (-2) \begin{vmatrix} 1 & 19 & 2 \\ 0 & 4 & 12 \\ 0 & -14 & 6 \end{vmatrix} \quad \text{By } R'_2 \rightarrow R_2 + (-2)R_1, R'_3 \rightarrow R_3 + (-1)R_1$$

$$\begin{aligned}
 &= (-2) \begin{vmatrix} 1 & 19 & 2 \\ 0 & 4 & 12 \\ 0 & -14 & 6 \end{vmatrix} \quad R_2 + (-2)R_1, R_3 + (-1)R_1 \\
 &= -2 \begin{vmatrix} 4 & 12 \\ -14 & 6 \end{vmatrix} \text{ expand from Ist column} \\
 &= -2(24+168) = -384
 \end{aligned}$$

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

$$\begin{aligned}
 &\begin{vmatrix} x & a+x & b+c \\ x & b+x & c+a \\ x & c+x & a+b \end{vmatrix} \\
 &= \begin{vmatrix} x & a+x-x & b+c \\ x & b+x-x & c+a \\ x & c+x-x & a+b \end{vmatrix} \quad \text{By } C_2' \rightarrow C_2 - C_1 \\
 &= \begin{vmatrix} x & a & b+c \\ x & b & c+a \\ x & c & a+b \end{vmatrix}
 \end{aligned}$$

Taking 'x' common from C_1

$$\begin{aligned}
 &= x \begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix} \\
 &= x \begin{vmatrix} 1 & a+b+c & b+c \\ 1 & b+c+a & c+a \\ 1 & c+a+b & a+b \end{vmatrix} \quad \text{By } C_2' \rightarrow C_2 + C_3
 \end{aligned}$$

Now taking $(a+b+c)$ common from 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 1st 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

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

$$R'_2 \rightarrow R_2 - 2R_1, R'_3 \rightarrow R_3 - 2R_1, R'_4 \rightarrow R_4 - 3R_1, R'_5 \rightarrow R_5 - 4R_1$$

$$= - \begin{vmatrix} 1 & 1 & 3 & 1 & 2 \\ 0 & 1 & -5 & -2 & -3 \\ 0 & -1 & -4 & 1 & 0 \\ 0 & -1 & -8 & -2 & -4 \\ 0 & -3 & -11 & -4 & -8 \end{vmatrix}$$

expand from C_1

$$= - \begin{vmatrix} 1 & -5 & -2 & -3 \\ -1 & -4 & 1 & 0 \\ -1 & -8 & -2 & -4 \\ -3 & -11 & -4 & -8 \end{vmatrix}$$

$$R'_2 \rightarrow R_2 + R_1, R'_3 \rightarrow R_3 + R_1, R'_4 \rightarrow R_4 + 3R_1$$

$$= - \begin{vmatrix} 1 & -5 & -2 & -3 \\ 0 & -9 & -1 & -3 \\ 0 & -13 & -4 & -7 \\ 0 & -26 & -10 & -17 \end{vmatrix}$$

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})

$$= - \begin{vmatrix} 1 & 9 & 3 \\ 4 & 13 & 7 \\ 10 & 26 & 17 \end{vmatrix}$$

$C'_2 \rightarrow C_2 - 9C_1, C'_3 \rightarrow C_3 - 3C_1$

$$= - \begin{vmatrix} 1 & 0 & 0 \\ 4 & -23 & -5 \\ 10 & -64 & -13 \end{vmatrix}$$

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 $\det(\mathbf{A})$ of order n is obtained as follows:

Step 1: By an interchange of rows of $\mathbf{A}$ (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 $\det \mathbf{A}=0$).

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 $\det(\mathbf{A})$ 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, step 2 and step 3 with the remaining (n-2) rows, (n-3) rows and so on, until a triangular matrix is obtained.

Step 5: Multiply all the diagonal entries of the resulting triangular matrix and then multiply it by its sign to get $\det(\mathbf{A})$

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:

$$\begin{aligned} \det A &= \begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix} \\ &= \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} \quad \text{By } R'_2 \rightarrow R_2 + 2R_1, R'_3 \rightarrow R_1 + R_3 \end{aligned}$$

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

$$\begin{aligned} \det A &= 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} \\ &= 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix} \quad \text{By } R'_2 \rightarrow R_2 - 3R_1, R'_3 \rightarrow R_3 + 3R_1, R'_4 \rightarrow R_4 - R_1 \end{aligned}$$

$$\det A = 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (\text{By } R'_4 \rightarrow R_4 - \frac{1}{2}R_3)$$

$$= 2 \cdot \{(1)(3)(-6)(1)\} = -36$$

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

$$\begin{vmatrix} x & 2 & 2 & 2 \\ 2 & x & 2 & 2 \\ 2 & 2 & x & 2 \\ 2 & 2 & 2 & x \end{vmatrix}$$

$$= \begin{vmatrix} x+6 & 2 & 2 & 2 \\ x+6 & x & 2 & 2 \\ x+6 & 2 & x & 2 \\ x+6 & 2 & 2 & x \end{vmatrix} \quad \text{By } C'_1 \rightarrow C_1 + (C_2 + C_3 + C_4)$$

Taking (x+6) common from 1st column

$$= (x+6) \begin{vmatrix} 1 & 2 & 2 & 2 \\ 1 & x & 2 & 2 \\ 1 & 2 & x & 2 \\ 1 & 2 & 2 & x \end{vmatrix}$$

$$= (x+6) \begin{vmatrix} 1 & 2 & 2 & 2 \\ 0 & x-2 & 0 & 0 \\ 0 & 0 & x-2 & 0 \\ 0 & 0 & 0 & x-2 \end{vmatrix} \quad \text{By } R'_2 \rightarrow R_2 - R_1, R'_3 \rightarrow R_3 - R_1, R'_4 \rightarrow R_4 - R_1$$

And this is the triangular matrix and its determinant is the product of main diagonal's entries.

$$= (x+6)(x-2)^3$$

Example 9 Compute $\det A$, where $A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$.

Solution $A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$

$$\det A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix} \begin{matrix} R'_3 \rightarrow R_3 + 2R_1 \\ \\ \end{matrix}$$

$$= 0 \quad \text{as } R_2 \equiv R_3$$

Example 10 Compute $\det A$, where

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

Solution

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

$$= \begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{bmatrix} \begin{matrix} R'_4 \rightarrow R_4 + R_2 \\ \\ \end{matrix}$$

$$= (-1) \begin{bmatrix} 2 & 1 & 2 & -1 \\ 0 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{bmatrix} \quad \text{By } R'_{12}$$

Expanding from 1st row and 1st column

$$\begin{aligned}
&= -2 \begin{vmatrix} 5 & -7 & 3 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix} \\
&= (-2) \{5(6+6) - (-7)(3-0) + 3(-9-0)\} \\
&= 54
\end{aligned}$$

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 $\det U = 0$.

$$\begin{array}{cc}
U = \begin{bmatrix} \bullet & \circ & \circ & \circ \\ 0 & \bullet & \circ & \circ \\ 0 & 0 & \bullet & \circ \\ 0 & 0 & 0 & \bullet \end{bmatrix} & U = \begin{bmatrix} \bullet & \circ & \circ & \circ \\ 0 & \bullet & \circ & \circ \\ 0 & 0 & \bullet & \circ \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
\det U \neq 0 & \det U = 0
\end{array}$$

Thus we have the following formula

$$\det A = \begin{cases} (-1)^r \cdot (\text{Product of pivots in } U) & \text{When } A \text{ is invertible} \\ 0 & \text{When } A \text{ is not invertible} \end{cases} \quad (1)$$

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'_2)

$$\sim \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix} \because \text{one replacement of rows has occurred, } \therefore r = 1$$

$$\sim \begin{bmatrix} 3 & 2 \\ 0 & \frac{7}{3} \end{bmatrix} \text{ By } R'_2 \rightarrow R'_2 - \frac{4}{3}R'_1, \text{ we have desired row-echelon form } U = \begin{bmatrix} 3 & 2 \\ 0 & \frac{7}{3} \end{bmatrix}.$$

Thus using the above formula as follows;

$$\det A = (-1)^r \cdot (\text{Product of pivots in } U) = (-1)^1 \left(3 \cdot \frac{7}{3}\right) = -7$$

Case-02 For 2×2 non-invertible matrix

In this case say;

$$A = \begin{bmatrix} 4 & 5 \\ 8 & 10 \end{bmatrix}$$

$$\sim \begin{bmatrix} 4 & 5 \\ 0 & 0 \end{bmatrix} \quad \text{By } R_2' \rightarrow R_2 - 2R_1, \text{ desired row-echelon form is } U = \begin{bmatrix} 4 & 5 \\ 0 & 0 \end{bmatrix}$$

Here no interchange of rows has occurred. So, $r = 0$ and

$$\therefore \det A = (-1)^r \cdot (\text{Product of pivots in } U) = (-1)^0 (4 \cdot 0) = 0$$

Theorem 5 If A is an $n \times n$ matrix, then $\det A^T = \det A$.

Example 11 If $A = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 3 \end{bmatrix}$, find $\det(A)$ and $\det(A^T)$

$$\det A = \begin{vmatrix} 1 & 4 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 3 \end{vmatrix} = 1(3-2) - 4(6-6) + 1(2-3) = 1 - 0 - 1 = 0$$

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 A and 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 $\det AB = 25 \cdot 13 - 20 \cdot 14 = 325 - 280 = 45$

Since $\det A = 9$ and $\det B = 5$, $(\det A)(\det B) = 9 \cdot 5 = 45 = \det AB$

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

$A + B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \det(A + B) = 0$

$\det A + \det B = \begin{vmatrix} 2 & 3 \\ 1 & -5 \end{vmatrix} + \begin{vmatrix} -2 & -3 \\ -1 & 5 \end{vmatrix} = (-10 - 3) + (-10 - 3) = -26 \neq \det(A + B)$

Masters

Subscribes to Masters

highlighted by Masters

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.

$$7. \begin{vmatrix} 2 & 5 & -3 & -1 \\ 3 & 0 & 1 & -3 \\ -6 & 0 & -4 & 9 \\ 4 & 10 & -4 & -1 \end{vmatrix}$$

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

$$9. \text{ Use determinant to find out whether the matrix is invertible } \begin{bmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{bmatrix}$$

10. Let $\mathbf{A}$ and $\mathbf{B}$ be 3×3 matrices, with $\det \mathbf{A} = 4$ and $\det \mathbf{B} = -3$. Use properties of determinants to compute

- (a) $\det \mathbf{AB}$ (b) $\det 7\mathbf{A}$ (c) $\det \mathbf{B}^T$ (d) $\det \mathbf{A}^T$
 (e) $\det \mathbf{A}^T \mathbf{A}$

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

$$(a) \begin{vmatrix} a_1 + b_1 t & a_2 + b_2 t & a_3 + b_3 t \\ a_1 t + b_1 & a_2 t + b_2 & a_3 t + b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = (1 - t^2) \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$(b) \begin{vmatrix} a_1 & b_1 + ta_1 & c_1 + rb_1 + sa_1 \\ a_2 & b_2 + ta_2 & c_2 + rb_2 + sa_2 \\ a_3 & b_3 + ta_3 & c_3 + rb_3 + sa_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

13. Show that $\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = (y - x)(z - x)(z - y)$