National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Linear Algebra (MT-1004)

Lecturer

USAMA ANTULEY

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

LECTURE # 01

Course Details

- **Text Book:**

Elementary Linear Algebra 12th edition By Howard Anton & Anton Kaul

- **Reference Book (s):**

1. **Linear Algebra & its Applications By Gilbert Strang**
2. **Coding the Matrix: Linear Algebra through Applications to Computer Science By Philip N Klein**

Marking Division

S. No	Particulars	% Marks
1.	Assignment	12
2.	Quiz	08
3.	Mid Exam	30
4.	Final Exam	50
	Total	**100**

Linear Equation

"The equation of a straight line is known as "linear equation".

OR

"An equation in which variable's highest index/power is 1"

- In two dimensions a line in a rectangular xy-coordinate system can be represented by an equation of the form:

$$\boldsymbol{ax} + \boldsymbol{by} = \boldsymbol{c} \quad (a\ , b\ not\ both\ 0)$$

- In three dimensions a plane in a rectangular xyz-coordinate system can be represented by an equation of the form

$$\boldsymbol{ax} + \boldsymbol{by} + \boldsymbol{cz} = \boldsymbol{d} \quad (a\ , b, c\ not\ all\ 0)$$

Linear Equation

- The general form of linear equation is:

$$a_1x_1 + a_2x_2 + \ \ldots\ldots\ldots\ldots. + a_nx_n = b$$

System of Linear Equations

"A group (combination) of two or more linear equations having same variables is known as System of linear equation"

OR

"A finite set of linear equations is called a system of linear equations or, more briefly, linear system. The variables are called ***"unknowns"****.*

- For Example:

i) $2x + 3y = 1$

$x - y = 10$

ii) $2x + 3y + z = 1$

$x - y - z = 4$

$x + y + z = 10$

System of Linear Equations

- **Non-Homogeneous Equations:** A system of linear equations having m x n order,

$$
\begin{array}{l}
a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots\ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots\ldots\ldots\ldots + a_{2n}x_n = b_2 \\
\ldots \quad \ldots\ldots \quad \ldots\ldots \quad \ldots \\
\ldots \quad \ldots\ldots \quad \ldots\ldots \quad \ldots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots\ldots\ldots + a_{mn}x_n = b_m
\end{array}
\longrightarrow \textbf{System 1}
$$

- The **System 1** is known as Non-homogeneous system of linear equations.

System of Linear Equations

Homogeneous Equations: A system of linear equations having m x n order,

$$a_{11}x_1 + a_{12}x_2 + \dots\dots\dots\dots. + a_{1n}x_n = 0$$

$$a_{21}x_1 + a_{22}x_2 + \dots\dots\dots\dots. + a_{2n}x_n = 0$$

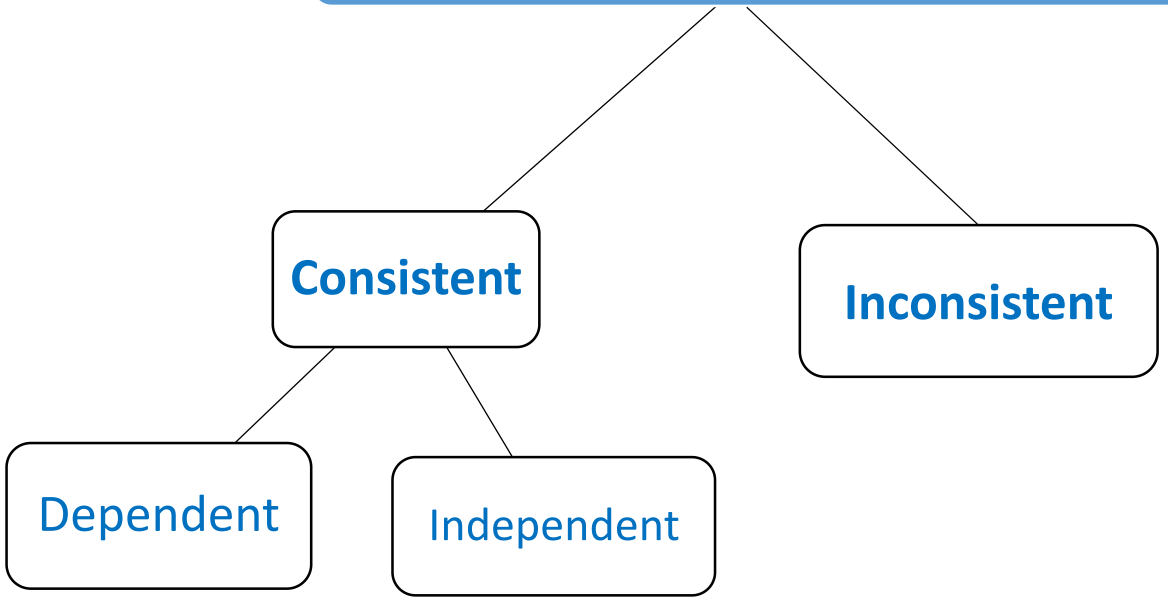
$$\dots \quad \dots.. \quad \dots\dots \quad \dots$$

$$\dots \quad \dots.. \quad \dots\dots \quad \dots$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots\dots\dots\dots. + a_{mn}x_n = 0 \longrightarrow \text{System 2}$$

The System 2 is known as Homogeneous system of linear equations.

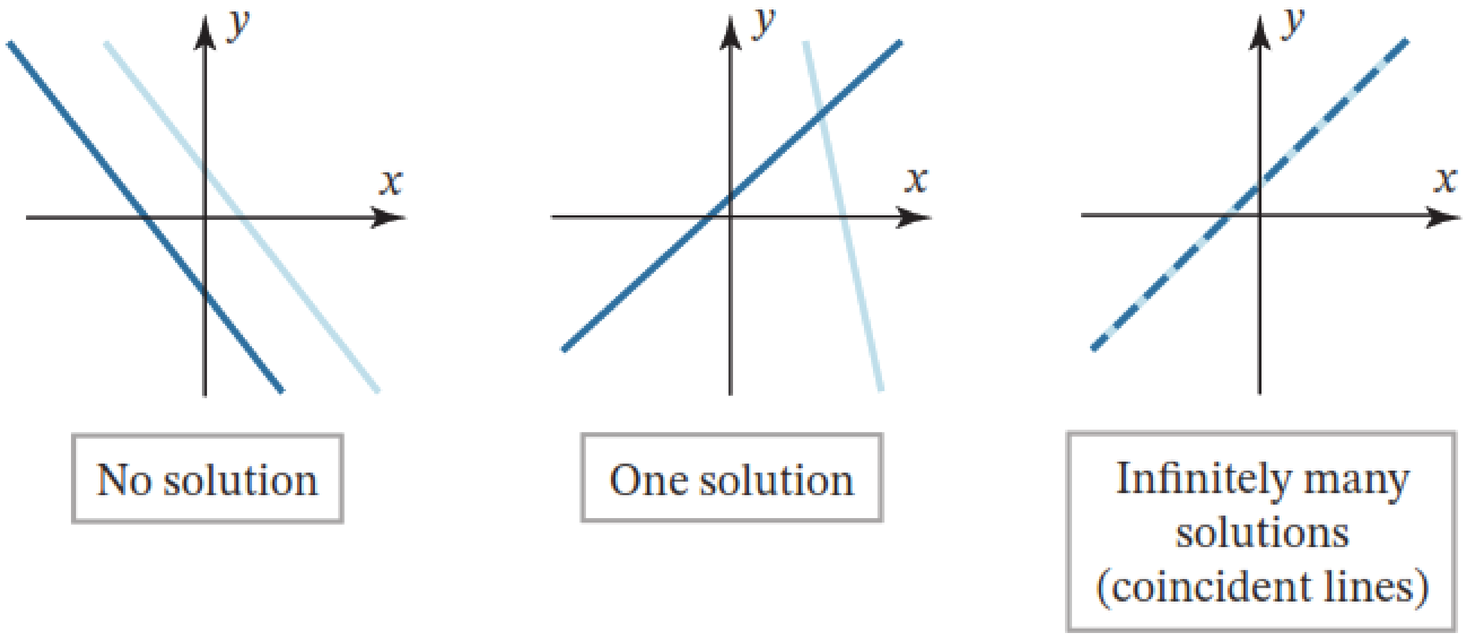
System of Linear Equation



- A system of equations that has ***no solution*** is said to be ***Inconsistent***
- A system of equations that has at least ***one solution*** is said to be ***Consistent***
- Dependent system has ***many solutions***
- Independent system has ***one solution***

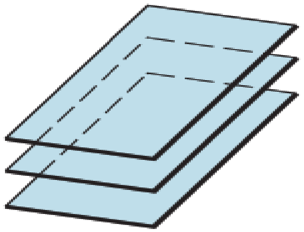
Types of Solution of System of Linear Equations

In two dimensions, For two equations having two unknowns

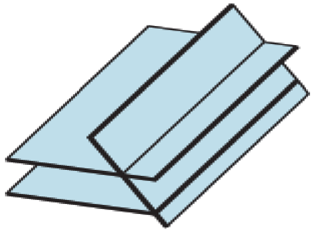


Types of Solution of System of Linear Equations

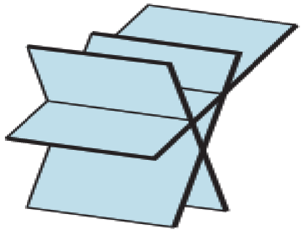
In three dimensions, For three equations having three unknowns



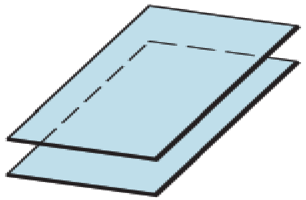
No solutions
(three parallel planes;
no common intersection)



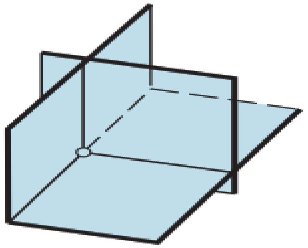
No solutions
(two parallel planes;
no common intersection)



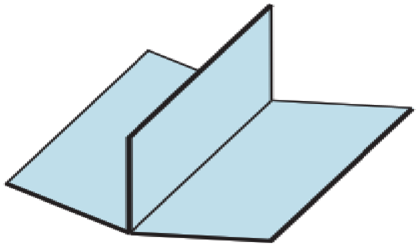
No solutions
(no common intersection)



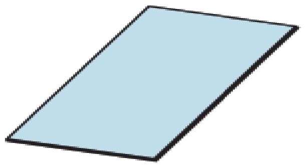
No solutions
(two coincident planes
parallel to the third;
no common intersection)



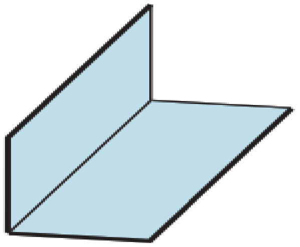
One solution
(intersection is a point)



Infinitely many solutions
(intersection is a line)



Infinitely many solutions
(planes are all coincident;
intersection is a plane)



Infinitely many solutions
(two coincident planes;
intersection is a line)

Dependent Linear System & Parametric Equations

We can describe the solution set of dependent linear system by the following steps

i. Taking the main equation and make x as subject in terms of y ***(Note: you can make y as a subject as well)***

ii. Assign an arbitrary value t (called a parameter) to y.

For example; if we have linear system of two variables x & y

$$\begin{aligned} x + y &= 6 \\ 2x + 2y &= 12 \end{aligned}$$

We can eliminate x from the second equation by adding -2 times the first equation to the second. This yields the simplified system

$$\begin{aligned} x + y &= 6 \\ 0 &= 0 \end{aligned}$$

Dependent Linear System & Parametric Equations

The second equation does not impose any restrictions on x and y and hence can be omitted. Thus, the solutions of the system are those values of x and y that satisfy the single equation

$$\boldsymbol{x + y = 6}$$

Now make x as a subject i.e. $\boldsymbol{x = 6 - y}$

and then assign an arbitrary value t (called a parameter) to y i.e. $y = t$

This allows us to express the solution by the pair of equations (called parametric equations)

$$x = 6 - t \quad \& \quad y = t$$

We can obtain specific numerical solutions from these equations by substituting numerical values for the parameter t.

i.e. when t = 0, $y = 0$ & $x = 6$ *i.e. yields solution (6,0) etc*

Augmented Matrix

The essential information of a linear system can be recorded compactly in a rectangular array called a ***"Matrix"***. Given the system,

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 \quad -5x_3 &= 10 \end{aligned}$$

with the coefficients of each variable aligned in columns, the matrix. In terms of ***Ax = b***

A= $\begin{bmatrix} \mathbf{1} & \mathbf{-2} & \mathbf{1} \\ \mathbf{0} & \mathbf{2} & \mathbf{-8} \\ \mathbf{5} & \mathbf{0} & \mathbf{-5} \end{bmatrix}$ is called the coefficient matrix (or matrix of coefficients) ; *x* = $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ *for variables ; b* = $\begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}$ *for constants (RHS)*

If we write the same system like $\begin{bmatrix} \mathbf{1} & \mathbf{-2} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{2} & \mathbf{-8} & \mathbf{8} \\ \mathbf{5} & \mathbf{0} & \mathbf{-5} & \mathbf{10} \end{bmatrix}$ It is called the ***Augmented matrix*** 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".

Elementary Row Operations (ERO)

ERO are of three types:

i. Multiply a given row by a non-zero number (i.e. kR_i which means multiply row R_i by a constant k)

ii. Interchanging any two rows of a matrix (i.e. R_{ij} which means interchange row R_i with R_j)

iii. Addition of any multiple of one row to another row (i.e. $kR_i + R_j$ which means multiply row R_i by a constant k and result so obtained is added in R_j)

Row Equivalent Matrices

The matrices ***A & B*** are called ***Row Equivalent Matrices***, written as $A \approx B$, if ***one can be obtained from the other*** by performing a finite sequence of ***ERO***

For example;

$$A = \begin{bmatrix} 1 & 5 & 2 & 3 \\ 3 & 1 & 8 & -1 \\ 2 & 5 & 1 & 6 \end{bmatrix}$$

*(Apply an ERO on R_2 i.e. $3R_2$, after multiplication, name the new matrices as **B** which is equivalent to **A**)*

$$B = \begin{bmatrix} 1 & 5 & 2 & 3 \\ 9 & 3 & 24 & -3 \\ 2 & 5 & 1 & 6 \end{bmatrix}$$

Here $\boldsymbol{B} \approx A$

Row Equivalent Matrices

$$A = \begin{bmatrix} 1 & 5 & 2 & 3 \\ 3 & 1 & 8 & -1 \\ 2 & 5 & 1 & 6 \end{bmatrix}$$

(Apply an ERO as R_{13} $i.e.interchange$ R_1 to R_3 then name the new matrices as ***C*** *which is equivalent to* ***A*** *)*

$$C = \begin{bmatrix} 2 & 5 & 1 & 6 \\ 3 & 1 & 8 & -1 \\ 1 & 5 & 2 & 3 \end{bmatrix}$$

Here $\mathbf{C} \approx A$

Row Equivalent Matrices

$$\boldsymbol{A} = \begin{bmatrix} 1 & 5 & 2 & 3 \\ 3 & 1 & 8 & -1 \\ 2 & 5 & 1 & 6 \end{bmatrix}$$

*(Apply an ERO as $R_2 + 4R_1$ then name the new matrices as **D** which is equivalent to **A**)*

$$\boldsymbol{D} = \begin{bmatrix} 1 & 5 & 2 & 3 \\ 7 & 21 & 16 & 11 \\ 2 & 5 & 1 & 6 \end{bmatrix}$$

Here $\boldsymbol{D} \approx \boldsymbol{A}$

Zero Row

"A row of any matrix is called zero if its entries are zeros"

Examples;

i. $\begin{bmatrix} \mathbf{1} & \mathbf{3} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$

ii. $\begin{bmatrix} \mathbf{1} & \mathbf{5} & \mathbf{2} & \mathbf{4} \\ \mathbf{0} & \mathbf{1} & \mathbf{6} & \mathbf{4} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$

Echelon Form OR Row Echelon Form

A matrix is in echelon form if :

- If a row does not consist entirely of zeros, then ***the first non-zero number in the row is a 1***. We call this a ***leading 1***
- If there are any rows that consist ***entirely of zeros***, then ***they are grouped together at the bottom of the matrix***
- In any two ***successive rows*** that do not consist entirely of zeros, ***the leading 1 in the lower row*** occurs ***farther to the right*** than the leading 1 in ***the higher/above row.***

Examples; $\begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$ $\begin{bmatrix} \mathbf{1} & \mathbf{3} & \mathbf{-1} \\ \mathbf{0} & \mathbf{1} & \mathbf{3} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$ $\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{3} & \mathbf{4} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{4} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$

Reduced Echelon Form

If possess the additional property that:

- ***Each column*** that contains ***a leading 1 has zeros everywhere else in that column***

Then it becomes ***Reduced Echelon Form***. *(Thus, a matrix in reduced row echelon form is of necessity in row echelon form, but not conversely)*

Examples; $\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{2} \end{bmatrix}$ $\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$ $\begin{bmatrix} \mathbf{1} & \mathbf{3} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$

Types of Solutions of Augmented Matrices on applying Reduced Echelon Form

a) Unique Solution:

Suppose that the augmented matrix for a linear system in the unknowns $x_1, x_2, x_3,$ and x_4 has been reduced by elementary row operations to

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

This matrix is in reduced row echelon form and corresponds to the equations

$$\begin{aligned} x_1 \qquad\qquad\qquad &= 3 \\ x_2 \qquad\qquad &= -1 \\ x_3 \qquad &= 0 \\ x_4 &= 5 \end{aligned}$$

Thus, the system has a unique solution, namely, $x_1 = 3, x_2 = -1, x_3 = 0, x_4 = 5$, which can also be expressed as the 4-tuple $(3, -1, 0, 5)$.

b) No Solution:

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

The equation that corresponds to the last row of the augmented matrix is

$$0x + 0y + 0z = 1$$

Since this equation is not satisfied by any values of x, y, and z, the system is inconsistent.

b) Infinitely many solutions: (Concept of leading & free variables)

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

The equation that corresponds to the last row of the augmented matrix is

$$0x + 0y + 0z = 0$$

This equation can be omitted since it imposes no restrictions on x, y, and z; hence, the linear system corresponding to the augmented matrix is

$$\begin{aligned} x \quad\quad + 3z &= -1 \\ y - 4z &= 2 \end{aligned}$$

In general, the variables in a linear system that correspond to the leading l's in its augmented matrix are called the ***leading variables***, and the remaining variables are called the ***free variables***. In this case the leading variables are x and y, and the variable z is the only free variable. Solving for the leading variables in terms of the free variables gives

$$\begin{aligned} x &= -1 - 3z \\ y &= 2 + 4z \end{aligned}$$

From these equations we see that the free variable z can be treated as a parameter and assigned an arbitrary value t, which then determines values for x and y. Thus, the solution set can be represented by the parametric equations

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

By substituting various values for t in these equations we can obtain various solutions of the system. For example, setting $t = 0$ yields the solution

$$x = -1, \quad y = 2, \quad z = 0$$

Gauss Elimination Method: (Based on Echelon Form_Forward phase)

Following are the steps for *Gauss Elimination Method:*

Step # 01: Change the linear system to form Ax=b

Step # 02: Form the augmented matrix A_b by including the elements of b as an extra column in matrix A

Steps # 03: Convert the augmented matrix in to ***Echelon Form*** by using ***ERO***

Step # 04: Convert the echelon matrix in to ***equation form*** and find unknowns by using backward substitution

Gauss Jordan Method: (Based on Reduced Echelon Form_Forward phase+Backward Phase)

Following are the steps for *Gauss Jordan Method: (Here we don't need backward substitution)*

Step # 01: Change the linear system to form Ax=b

Step # 02: Form the augmented matrix A_b by including the elements of b as an extra column in matrix A

Steps # 03: Convert the augmented matrix in to ***Reduced Echelon Form*** by using ***ERO***

Step # 04: Convert the reduced echelon matrix in to ***equation form*** and find unknowns directly

Elementary Matrix:

"A matrix (E) is called Elementary Matrix if it is obtained from an identity matrix by performing a single elementary row operation"

Examples:

Listed below are four elementary matrices and the operations that produce them.

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

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.

Elementary Matrix:

{ Theorem 1.5.1 }

Row Operations by Matrix Multiplication: If the elementary matrix E results from performing a certain row operation on I_m and if A is an m × n matrix, then the product EA is the matrix that results when this same row operation is performed on A.

Example: Consider the matrix

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

and consider the elementary matrix

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

which results from adding 3 times the first row of I_3 to the third row. The product EA is

$$EA = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 9 \end{bmatrix}$$

which is precisely the matrix that results when we add 3 times the first row of A to the third row.

Elementary Matrix:

{ Theorem 1.5.2 }

"Every elementary matrix is invertible, and the inverse is also an elementary matrix"

i.e. if E is an elementary matrix and E_0 is its inverse then It always follows:

$$\boldsymbol{E_0 E = I} \textbf{ and } \boldsymbol{E E_0 = I}$$

Method for Inverting Matrices:

Assume that the reduced row echelon form of A is I_n, so that A can be reduced to I_n by a finite sequence of elementary row operations. By Theorem 1.5.1, each of these operations can be accomplished by multiplying on the left by an appropriate elementary matrix. Thus we can find elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k \cdots E_2 E_1 A = I_n \tag{3}$$

assume for the moment, that A is an invertible $n \times n$ matrix. In Equation (3), the elementary matrices execute a sequence of row operations that reduce A to I_n. If we multiply both sides of this equation on the right by A^{-1} and simplify, we obtain

$$A^{-1} = E_k \cdots E_2 E_1 I_n$$

But this equation tells us that *the same sequence of row operations that reduces A to I_n will transform I_n to A^{-1}*. Thus, we have established the following result.

Method for Inverting Matrices:

Inversion Algorithm To find the inverse of an invertible matrix A, find a sequence of elementary row operations that reduces A to the identity and then perform that same sequence of operations on I_n to obtain A^{-1}.

Solved examples/pattern of inverse of a 3x3 & 4x4 matrices are from next slides

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Inverse of a Matrix by ERO :-

$$[A \mid I] \xrightarrow[ERO]{} [I \mid A^{-1}]$$

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

Number of Solutions of Linear System:

{ Theorem 1.6.1 }

"A system of linear equations has zero, one, or infinitely many solutions. There are no other possibilities"

Solving Linear Systems by Matrix Inversion:

{ Theorem 1.6.2 }

If A is an invertible n × n matrix, then for every n × 1 matrix b, the system of equations $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ has **exactly one solution** , namely, $\mathbf{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$

Task for Students
Refer Lecture & Example #1 (Pg # 63)
Question # 1 till 8 (Ex #1.6 ; Pg # 67) is related to above topic. Do for Practice.

Linear Systems with a Common Co-efficient Matrix (A) :

The method below is applicable for both invertible & not invertible co-efficient matrix A

$$[A \mid \boldsymbol{b_1} \mid \boldsymbol{b_2} \mid \cdots \mid \boldsymbol{b_k}]$$

in which the coefficient matrix A is “augmented” by all k of the matrices $\boldsymbol{b_1}$, $\boldsymbol{b_2}$, . . . , $\boldsymbol{b_k}$, and then reduce to reduced row echelon form by Gauss–Jordan elimination

Task for Students

Refer Lecture & Example #2 (Pg # 63)

Question # 9 till 12 (Ex #1.6 ; Pg # 67 & 68) is related to above topic. Do for Practice.

NOTE:

Theorem 1.6.3

Let A be a square matrix.

(a) If B is a square matrix satisfying $BA = I$, then $B = A^{-1}$.

(b) If B is a square matrix satisfying $AB = I$, then $B = A^{-1}$.

Theorem 1.6.4

Equivalent Statements

If A is an $n \times n$ matrix, then the following are equivalent.

(a) A is invertible.

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(c) The reduced row echelon form of A is I_n.

(d) A is expressible as a product of elementary matrices.

(e) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.

(f) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.

Theorem 1.6.5

Let A and B be square matrices of the same size. If AB is invertible, then A and B must also be invertible.

Determining Consistency by Elimination

If A is an invertible matrix, Theorem 1.6.2 completely solves this problem by asserting that for every m × 1 matrix b, the linear system Ax = b has the unique solution x $= A^{-1}$b.

If A is not square, or if A is square but not invertible, then Theorem 1.6.2 does not apply.

In these cases **b** must usually satisfy certain conditions in order for Ax = b to be consistent.

Task for Students

Refer Lecture & Example #3 & 4 (Pg # 66 & 67)

Question # 13 till 17 (Ex #1.6 ; Pg # 67 & 68) is related to above topic. Do for Practice.

Some Special Types of Matrices

Diagonal Matrices:

A square matrix in which all the entries off the main diagonal are zero is called a diagonal matrix. Here are some examples:

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

A general $n \times n$ diagonal matrix D can be written as

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

A diagonal matrix is **invertible** if and only if all of its **diagonal entries are nonzero**; in this case the inverse of D as mentioned above is:

$$D^{-1} = \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix}$$

Powers of diagonal matrices:

$$D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix}$$

Matrix products that involve diagonal matrix:

$$\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} d_1a_{11} & d_1a_{12} & d_1a_{13} & d_1a_{14} \\ d_2a_{21} & d_2a_{22} & d_2a_{23} & d_2a_{24} \\ d_3a_{31} & d_3a_{32} & d_3a_{33} & d_3a_{34} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} = \begin{bmatrix} d_1a_{11} & d_2a_{12} & d_3a_{13} \\ d_1a_{21} & d_2a_{22} & d_3a_{23} \\ d_1a_{31} & d_2a_{32} & d_3a_{33} \\ d_1a_{41} & d_2a_{42} & d_3a_{43} \end{bmatrix}$$

In words, *to multiply a matrix A on the left by a diagonal matrix D, multiply successive rows of A by the successive diagonal entries of D, and to multiply A on the right by D, multiply successive columns of A by the successive diagonal entries of D.*

Task for Students

Refer Lecture & Example #1 (Pg # 69)

Question # 3 till 10 (Ex #1.7 ; Pg # 74) is related to above topic. Do for Practice.

Triangular Matrices:

A square matrix in which all the entries **above** the main diagonal are zero is called lower triangular, and a square matrix in which all the entries below the main diagonal are zero is called upper triangular. A matrix that is either upper triangular or lower triangular is called triangular

Note:

- Diagonal matrices are both upper triangular and lower triangular since they have zeros below and above the main diagonal.
- Observe also that a square matrix in row echelon form is upper triangular since it has zeros below the main diagonal.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

A general 4×4 upper triangular matrix

A general 4×4 lower triangular matrix

Properties of Triangular Matrices:

Theorem 1.7.1

(*a*) The transpose of a lower triangular matrix is upper triangular, and the transpose of an upper triangular matrix is lower triangular.

(*b*) The product of lower triangular matrices is lower triangular, and the product of upper triangular matrices is upper triangular.

(*c*) A triangular matrix is invertible if and only if its diagonal entries are all nonzero.

(*d*) The inverse of an invertible lower triangular matrix is lower triangular, and the inverse of an invertible upper triangular matrix is upper triangular.

Task for Students

Observe Example #3 (Pg #71)

Symmetric Matrices:

Definition 1

A square matrix A is said to be ***symmetric*** if $A = A^T$.

EXAMPLE 4 | Symmetric Matrices

The following matrices are symmetric since each is equal to its own transpose (verify).

$$\begin{bmatrix} 7 & -3 \\ -3 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 4 & 5 \\ 4 & -3 & 0 \\ 5 & 0 & 7 \end{bmatrix}, \quad \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}$$

Important Points/Properties related to Symmetric Matrices:

Theorem 1.7.2

If A and B are symmetric matrices with the same size, and if k is any scalar, then:

(a) A^T is symmetric.

(b) $A+B$ and $A-B$ are symmetric.

(c) kA is symmetric.

Theorem 1.7.3

The product of two symmetric matrices is symmetric if and only if the matrices commute.

Task for Students

Observe Example #5 (Pg #73)

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Invertibility of Symmetric Matrices

Theorem 1.7.4

If A is an invertible symmetric matrix, then A^{-1} is symmetric.

Theorem 1.7.5

If A is an invertible matrix, then AA^T and A^TA are also invertible.

Task for Students

Observe Example # 6 (Pg #74)

As per Course outline : Do (Q.1 till 10 & 19 till 28 from Ex # 1.7)

LECTURE # 03 & 04

Fundamental Points before moving on Linear Transformation

We defined an "ordered n-tuple" to be a sequence of n real numbers, and we observed that a solution of a linear system in **n unknowns**, say $x_1 = s_1, \quad x_2 = s_2, \ldots, \quad x_n = s_n$

can be expressed as the ordered n-tuple The set of all $(s_1, s_2, \ldots, s_n)$

The set of all ordered **n-tuples** of real numbers is denoted by the symbol $\boldsymbol{R^n}$. The elements of $\boldsymbol{R^n}$ are called **vectors** and are denoted in **boldface** type, such **as a, b, v, w,** and **x**. When convenient, ordered n-tuples can be denoted in matrix notation as column vectors. For example, the matrix

$$\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

For each $i = 1, 2, \ldots, n$, let $\mathbf{e}_i$ denote the vector in R^n with a 1 in the ith position and zeros elsewhere. In column form these vectors are

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

We call the vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ the ***standard basis vectors*** for R^n. For example, the vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

are the standard basis vectors for R^3.

Transformation or Function or Mapping:

Recall that a ***function*** is a rule that associates with each element of a set A one and only one element in a set B. If f associates the element b with the element a, then we write

$$b = f(a)$$

and we say that b is the ***image*** of a under f or that $f(a)$ is the ***value*** of f at a. The set A is called the ***domain*** of f and the set B the ***codomain*** of f (Figure 1.8.1). The subset of the codomain that consists of all images of elements in the domain is called the ***range*** of f.

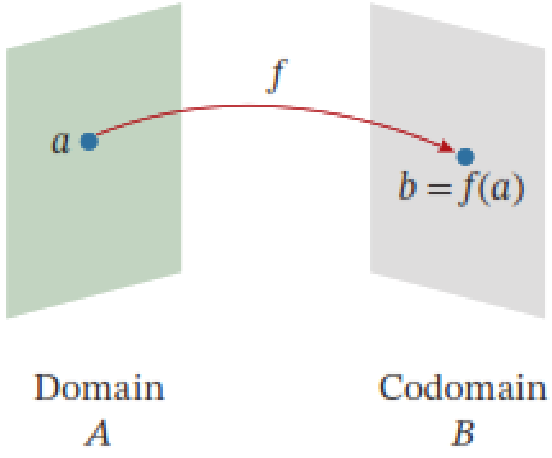


FIGURE 1.8.1

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 $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$. The set $\mathbb{R}^n$ is called the **domain** of T, and $\mathbb{R}^m$ is called the **codomain** of T. The notation $T : \mathbb{R}^n \to \mathbb{R}^m$ indicates that the domain of T is $\mathbb{R}^n$ and the codomain is $\mathbb{R}^m$. For $\mathbf{x}$ in $\mathbb{R}^n$, the vector $T(\mathbf{x})$ in $\mathbb{R}^m$ is called the **image** of $\mathbf{x}$ (under the action of T). The set of all images $T(\mathbf{x})$ is called the **range** of T. See Figure 2.

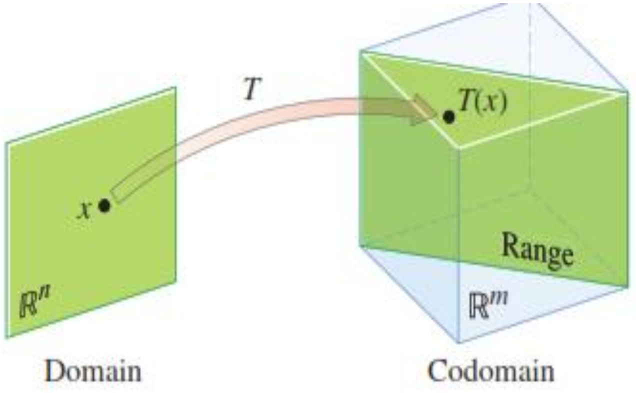


FIGURE 2 Domain, codomain, and range of $T : \mathbb{R}^n \to \mathbb{R}^m$.

In the special case where $m = n$, a transformation is sometimes called an ***operator*** on R^n.

Basic Idea _ Matrix Transformation

The difference between a matrix equation $\boldsymbol{Ax} = \boldsymbol{b}$ and the associated vector equation $\boldsymbol{x_1 a_1} + \boldsymbol{x_2 a_2} + \cdots + \boldsymbol{x_n a_n} = \boldsymbol{b}$ is merely a matter of notation. However, a matrix equation $\boldsymbol{Ax} = \boldsymbol{b}$ can arise in linear algebra (and in applications such as computer graphics and signal processing) in a way that is not directly connected with linear combinations of vectors. This happens when we think of the matrix $\boldsymbol{A}$ as an object that "acts" on a ***vector x*** by multiplication to produce a new vector called $\boldsymbol{Ax}$.

For instance, the equations

$$\underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 5 \\ 8 \end{bmatrix}}_{\mathbf{b}} \quad \text{and} \quad \underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}}_{\mathbf{u}} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\mathbf{0}}$$

say that multiplication by A transforms $\mathbf{x}$ into $\mathbf{b}$ and transforms $\mathbf{u}$ into the zero vector. See Figure 1.

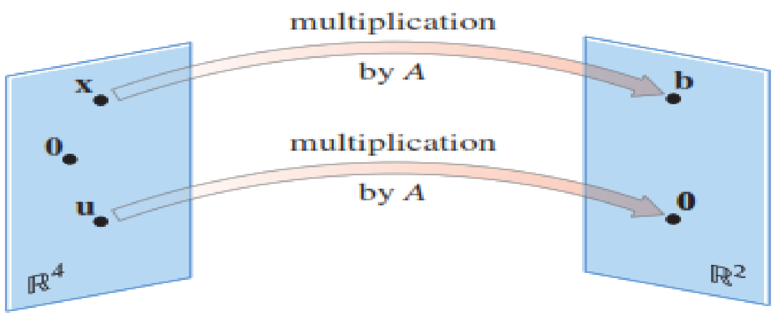


FIGURE 1 Transforming vectors via matrix multiplication.

Matrix Transformation:

The rest of this section focuses on mappings associated with matrix multiplication. For each $\mathbf{x}$ in $\mathbb{R}^n$, $T(\mathbf{x})$ is computed as $A\mathbf{x}$, where A is an $m \times n$ matrix. For simplicity, we sometimes denote such a *matrix transformation* by $\mathbf{x} \mapsto A\mathbf{x}$. Observe that the domain of

T is $\mathbb{R}^n$ when A has n columns and the codomain of T is $\mathbb{R}^m$ when each column of A has m entries. The range of T is the set of all linear combinations of the columns of A, because each image $T(\mathbf{x})$ is of the form $A\mathbf{x}$.

suppose that we have the system of linear equations

$$\begin{array}{ccccccccc} w_1 & = & a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ w_2 & = & a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n \\ \vdots & & \vdots & & \vdots & & & & \vdots \\ w_m & = & a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n \end{array}$$

which we can write in matrix notation as

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

or more briefly as

$$\mathbf{w} = A\mathbf{x}$$

$T_A(x) = Ax$

Finding the Standard Matrix for a Matrix Transformation

Step 1. Find the images of the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ for R^n.

Step 2. Construct the matrix that has the images obtained in Step 1 as its successive columns. This matrix is the standard matrix for the transformation.

EXAMPLE 4 | Finding a Standard Matrix

Find the standard matrix A for the linear transformation $T: R^2 \to R^3$ defined by the formula

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - 3x_2 \\ -x_1 + x_2 \end{bmatrix} \tag{16}$$

Solution We leave it for you to verify that

$$T(\mathbf{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$$

Thus, it follows from Formulas (15) and (16) that the standard matrix is

$$A = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2)] = \begin{bmatrix} 2 & 1 \\ 1 & -3 \\ -1 & 1 \end{bmatrix}$$

As a check, observe that

$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - 3x_2 \\ -x_1 + x_2 \end{bmatrix}$$

which shows that multiplication by A produces the same result as the transformation T (see Equation (16)).

EXAMPLE 5 | Computing with Standard Matrices

For the linear transformation in Example 4, use the standard matrix A obtained in that example to find

$$T\left(\begin{bmatrix} 1 \\ 4 \end{bmatrix}\right)$$

Solution The transformation is multiplication by A, so

$$T\left(\begin{bmatrix} 1 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 2 & 1 \\ 1 & -3 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ -11 \\ 3 \end{bmatrix}$$

EXAMPLE 6 | Finding a Standard Matrix

Rewrite the transformation $T(x_1, x_2) = (3x_1 + x_2, 2x_1 - 4x_2)$ in column-vector form and find its standard matrix.

Solution

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 + x_2 \\ 2x_1 - 4x_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Thus, the standard matrix is

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

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Matrix Transformation:

EXAMPLE 1 Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

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

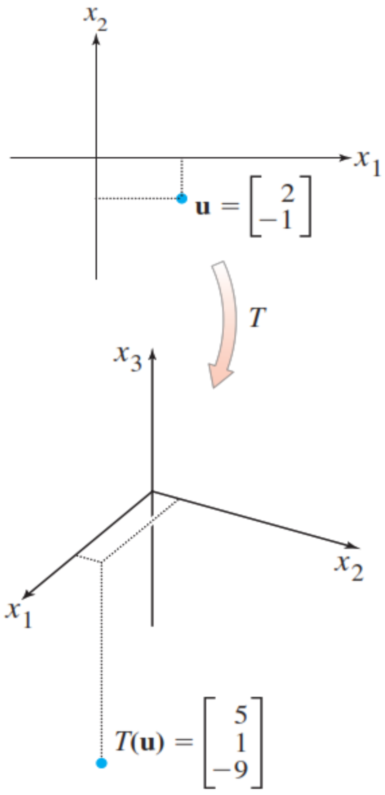
$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

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

Find $T(\mathbf{u})$, the image of $\mathbf{u}$ under the transformation T.

Compute

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



Geometrical Representation of Transformation for this case i.e.
$T_A(u)$: $R^2 \longrightarrow R^3$

NOTE: We take ***Matrix A*** *as our Transformation Parameter and Apply on vector* $\mathbf{u}$ $(\boldsymbol{R^2})$ *by means of multiplication to get* T**(u) *i.e.*** $(\boldsymbol{R^3})$ ***Here,*** $T_A(u)$: $R^2 \longrightarrow R^3$

Zero & Identity Transformations:

i.e. A as Zero Matrix & A=I

EXAMPLE 2 | Zero Transformations

If 0 is the $m \times n$ zero matrix, then

$$T_0(\mathbf{x}) = 0\mathbf{x} = \mathbf{0}$$

so multiplication by zero maps every vector in R^n into the zero vector in R^m. We call T_0 the ***zero transformation*** from R^n to R^m.

EXAMPLE 3 | Identity Operators

If I is the $n \times n$ identity matrix, then

$$T_I(\mathbf{x}) = I\mathbf{x} = \mathbf{x}$$

so multiplication by I maps every vector in R^n to itself. We call T_I the ***identity operator*** on R^n.

Properties of Matrix Transformations

Theorem 1.8.1

For every matrix A the matrix transformation $T_A : R^n \rightarrow R^m$ has the following properties for all vectors $\mathbf{u}$ and $\mathbf{v}$ and for every scalar k:

(a) $T_A(\mathbf{0}) = \mathbf{0}$

(b) $T_A(k\mathbf{u}) = kT_A(\mathbf{u})$ **[Homogeneity property]**

(c) $T_A(\mathbf{u} + \mathbf{v}) = T_A(\mathbf{u}) + T_A(\mathbf{v})$ **[Additivity property]**

(d) $T_A(\mathbf{u} - \mathbf{v}) = T_A(\mathbf{u}) - T_A(\mathbf{v})$

Theorem 1.8.2

$T : R^n \rightarrow R^m$ is a matrix transformation if and only if the following relationships hold for all vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n and for every scalar k:

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ **[Additivity property]**

(ii) $T(k\mathbf{u}) = kT(\mathbf{u})$ **[Homogeneity property]**

Reflection Operators:

Some of the most basic matrix operators on R^2 and R^3 are those that map each point into its symmetric image about a fixed line or a fixed plane that contains the origin; these are called reflection operators

TABLE 1

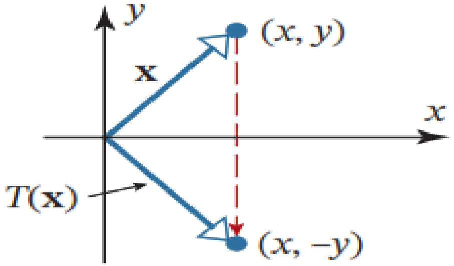
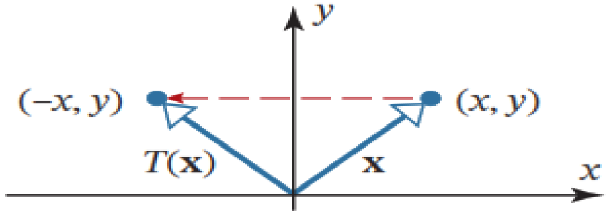
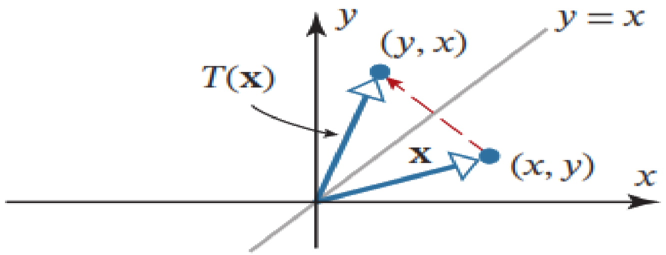
Operator	Illustration	Images of $\mathbf{e}_1$ and $\mathbf{e}_2$	Standard Matrix
Reflection about the x-axis $T(x, y) = (x, -y)$		$T(\mathbf{e}_1) = T(1, 0) = (1, 0)$ $T(\mathbf{e}_2) = T(0, 1) = (0, -1)$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflection about the y-axis $T(x, y) = (-x, y)$		$T(\mathbf{e}_1) = T(1, 0) = (-1, 0)$ $T(\mathbf{e}_2) = T(0, 1) = (0, 1)$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflection about the line $y = x$ $T(x, y) = (y, x)$		$T(\mathbf{e}_1) = T(1, 0) = (0, 1)$ $T(\mathbf{e}_2) = T(0, 1) = (1, 0)$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

TABLE 2

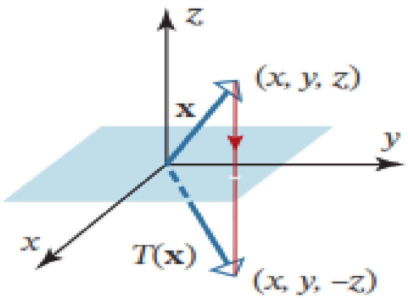
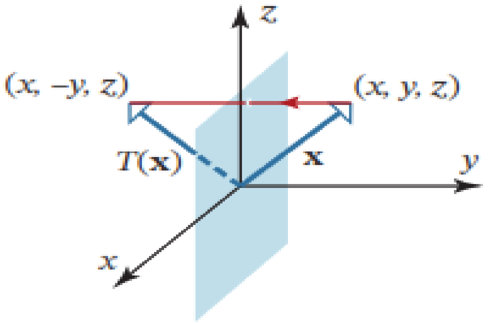
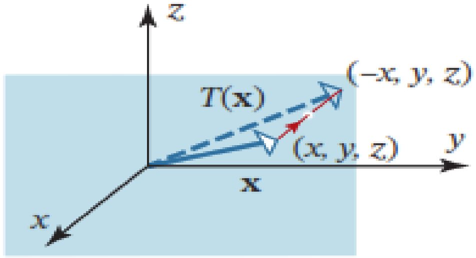
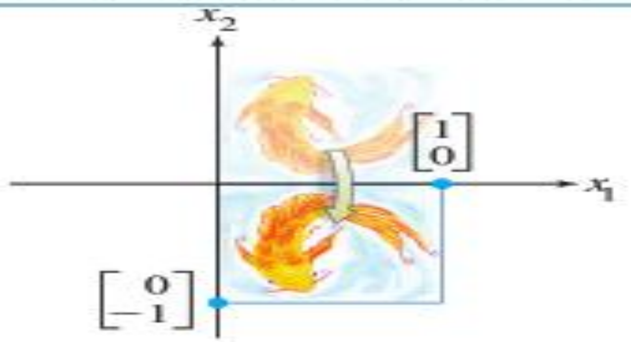
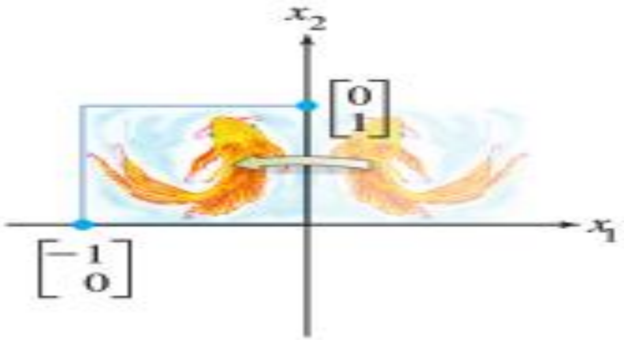
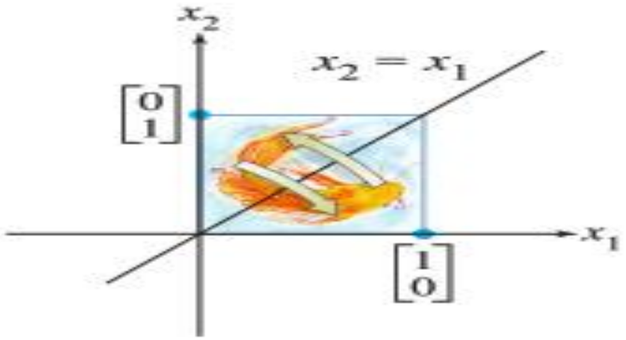
Operator	Illustration	Images of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$	Standard Matrix
Reflection about the xy-plane $T(x, y, z) = (x, y, -z)$		$T(\mathbf{e}_1) = T(1,0,0) = (1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,-1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Reflection about the xz-plane $T(x, y, z) = (x, -y, z)$		$T(\mathbf{e}_1) = T(1,0,0) = (1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,-1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Reflection about the yz-plane $T(x, y, z) = (-x, y, z)$		$T(\mathbf{e}_1) = T(1,0,0) = (-1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,1)$	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

TABLE 1 Reflections

Transformation	Image of the Unit Square	Standard Matrix
Reflection through the x_1-axis		$\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$
Reflection through the x_2-axis		$\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}$
Reflection through the line $x_2 = x_1$		$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$

Projection Operators:

Matrix operators on R^2 and R^3 that map each point into its orthogonal projection onto a fixed line or plane through the origin are called projection operators (or more precisely, orthogonal projection operators)

TABLE 3

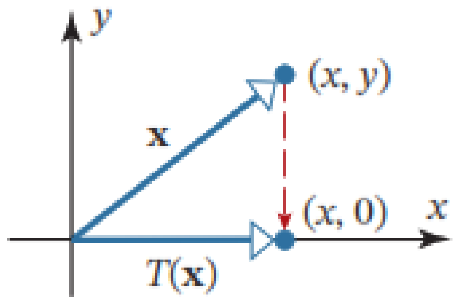
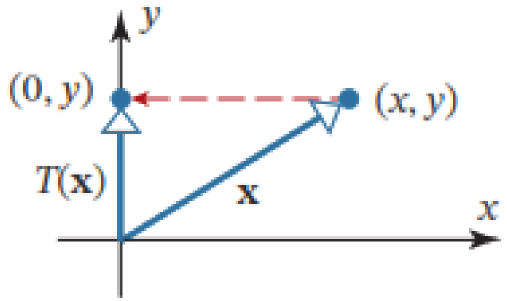
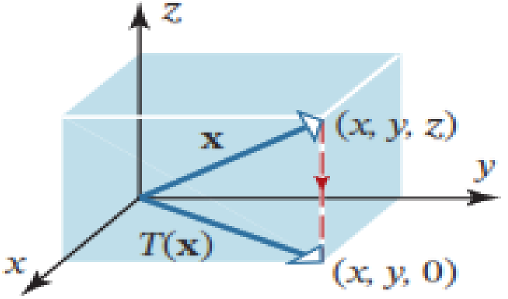
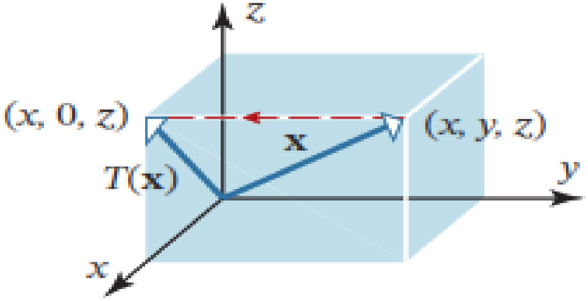
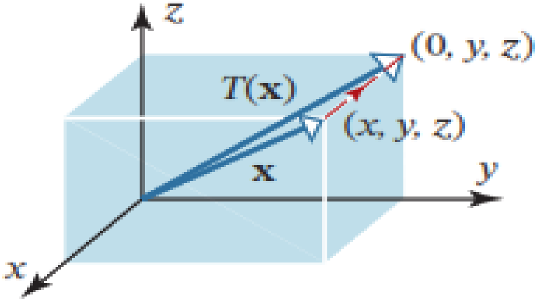
Operator	Illustration	Images of $\mathbf{e}_1$ and $\mathbf{e}_2$	Standard Matrix
Orthogonal projection onto the x-axis $T(x, y) = (x, 0)$		$T(\mathbf{e}_1) = T(1,0) = (1,0)$ $T(\mathbf{e}_2) = T(0,1) = (0,0)$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Orthogonal projection onto the y-axis $T(x, y) = (0, y)$		$T(\mathbf{e}_1) = T(1,0) = (0,0)$ $T(\mathbf{e}_2) = T(0,1) = (0,1)$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

TABLE 4

Operator	Illustration	Images of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$	Standard Matrix
Orthogonal projection onto the xy-plane $T(x, y, z) = (x, y, 0)$		$T(\mathbf{e}_1) = T(1,0,0) = (1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,0)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Orthogonal projection onto the xz-plane $T(x, y, z) = (x, 0, z)$		$T(\mathbf{e}_1) = T(1,0,0) = (1,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,0,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Orthogonal projection onto the yz-plane $T(x, y, z) = (0, y, z)$		$T(\mathbf{e}_1) = T(1,0,0) = (0,0,0)$ $T(\mathbf{e}_2) = T(0,1,0) = (0,1,0)$ $T(\mathbf{e}_3) = T(0,0,1) = (0,0,1)$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Rotation Operators:

Matrix operators on R^2 that move points along arcs of circles centered at the origin are called rotation operators

$$A = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2)] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

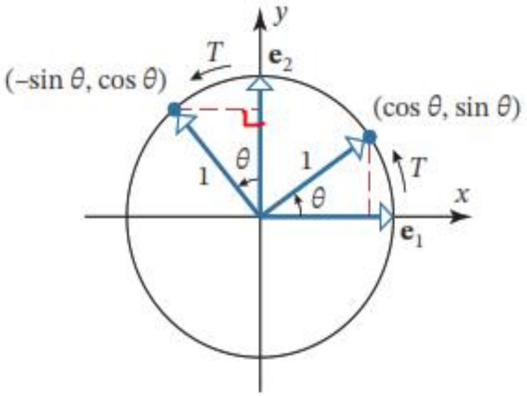
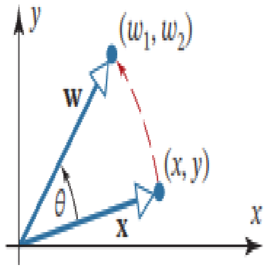


TABLE 5

Operator	Illustration	Images of $\mathbf{e}_1$ and $\mathbf{e}_2$	Standard Matrix
Counterclockwise rotation about the origin through an angle θ		$T(\mathbf{e}_1) = T(1,0) = (\cos\theta, \sin\theta)$ $T(\mathbf{e}_2) = T(0,1) = (-\sin\theta, \cos\theta)$	$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

In the plane, counterclockwise angles are positive and clockwise angles are negative. The rotation matrix for a *clockwise* rotation of $-\theta$ radians can be obtained by replacing θ by $-\theta$ in (19). After simplification this yields

$$R_{-\theta} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Some Extra Example of Transformation (not included in Ex # 1.8)

TABLE 2 Contractions and Expansions

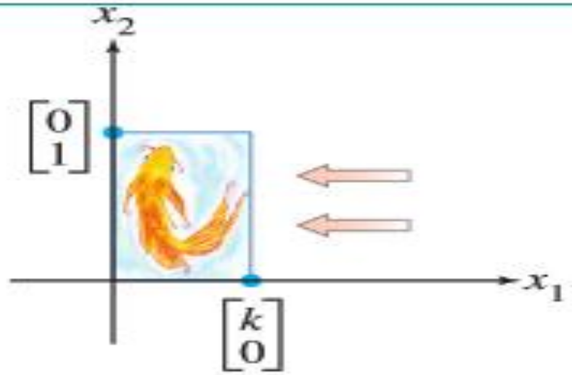
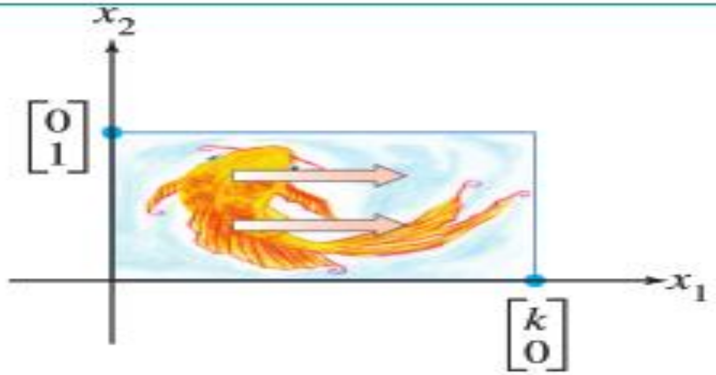
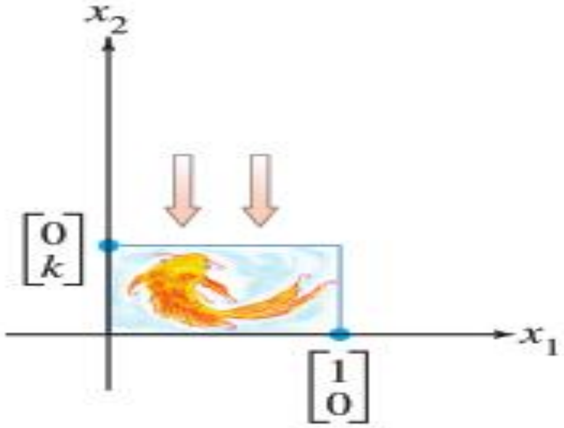
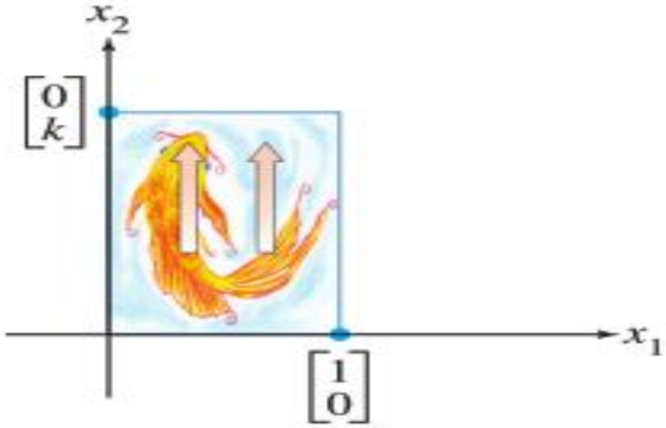
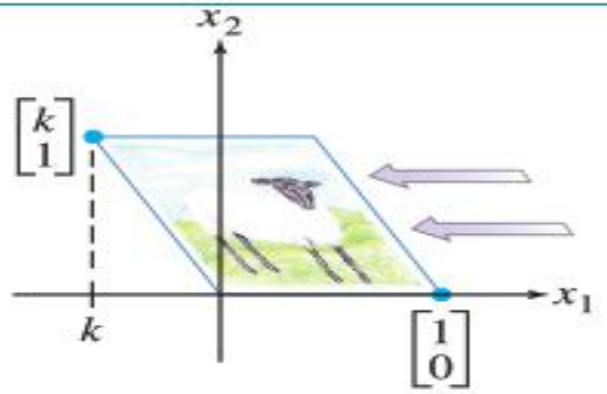
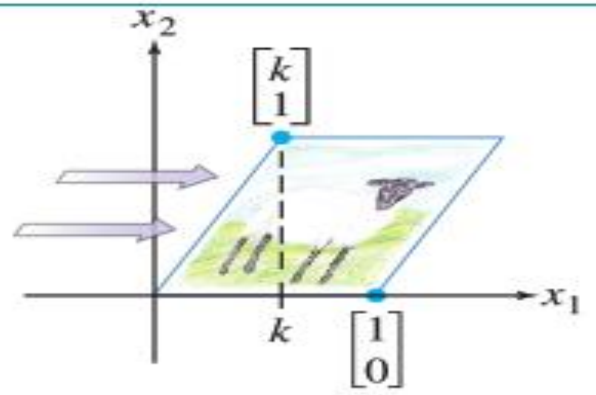
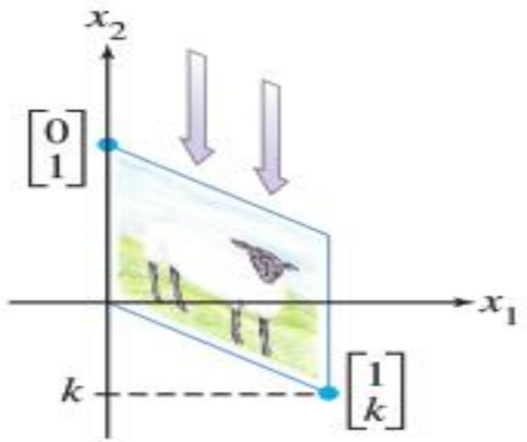
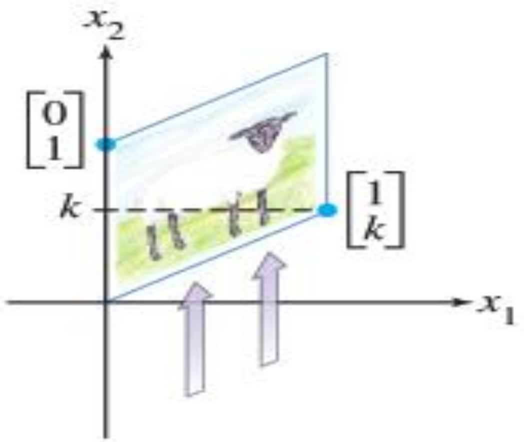
Transformation	Image of the Unit Square		Standard Matrix
Horizontal contraction and expansion	 $0 < k < 1$	 $k > 1$	$\begin{bmatrix}k & 0\\0 & 1\end{bmatrix}$
Vertical contraction and expansion	 $0 < k < 1$	 $k > 1$	$\begin{bmatrix}1 & 0\\0 & k\end{bmatrix}$

TABLE 3 Shears

Transformation	Image of the Unit Square		Standard Matrix
Horizontal shear	 $k < 0$	 $k > 0$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
Vertical shear	 $k < 0$	 $k > 0$	$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

In Exercises 1–2, find the domain and codomain of the transformation $T_A(\mathbf{x}) = A\mathbf{x}$.

1. **a.** A has size 3×2. **b.** A has size 2×3.

c. A has size 3×3. **d.** A has size 1×6.

2. **a.** A has size 4×5. **b.** A has size 5×4.

c. A has size 4×4. **d.** A has size 3×1.

In Exercises 3–4, find the domain and codomain of the transformation defined by the equations.

3. **a.**
$$\begin{aligned} w_1 &= 4x_1 + 5x_2 \\ w_2 &= x_1 - 8x_2 \end{aligned}$$

b.
$$\begin{aligned} w_1 &= 5x_1 - 7x_2 \\ w_2 &= 6x_1 + x_2 \\ w_3 &= 2x_1 + 3x_2 \end{aligned}$$

4. **a.**
$$\begin{aligned} w_1 &= x_1 - 4x_2 + 8x_3 \\ w_2 &= -x_1 + 4x_2 + 2x_3 \\ w_3 &= -3x_1 + 2x_2 - 5x_3 \end{aligned}$$

b.
$$\begin{aligned} w_1 &= 2x_1 + 7x_2 - 4x_3 \\ w_2 &= 4x_1 - 3x_2 + 2x_3 \end{aligned}$$

In Exercises 5–6, find the domain and codomain of the transformation defined by the matrix product.

5. **a.** $\begin{bmatrix} 3 & 1 & 2 \\ 6 & 7 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ **b.** $\begin{bmatrix} 2 & -1 \\ 4 & 3 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

6. **a.** $\begin{bmatrix} 6 & 3 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ **b.** $\begin{bmatrix} 2 & 1 & -6 \\ 3 & 7 & -4 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Solutions

1. **(a)** $T_A(\mathbf{x}) = A\mathbf{x}$ maps any vector $\mathbf{x}$ in R^2 into a vector $\mathbf{w} = A\mathbf{x}$ in R^3. The domain of T_A is R^2; the codomain is R^3.

(b) $T_A(\mathbf{x}) = A\mathbf{x}$ maps any vector $\mathbf{x}$ in R^3 into a vector $\mathbf{w} = A\mathbf{x}$ in R^2. The domain of T_A is R^3; the codomain is R^2.

(c) $T_A(\mathbf{x}) = A\mathbf{x}$ maps any vector $\mathbf{x}$ in R^3 into a vector $\mathbf{w} = A\mathbf{x}$ in R^3. The domain of T_A is R^3; the codomain is R^3.

(d) $T_A(\mathbf{x}) = A\mathbf{x}$ maps any vector $\mathbf{x}$ in R^6 into a vector $\mathbf{w} = A\mathbf{x}$ in $R^1 = R$. The domain of T_A is R^6; the codomain is R.

2. **(a)** $T_A(\mathbf{x}) = A\mathbf{x}$ maps any vector $\mathbf{x}$ in R^5 into a vector $\mathbf{w} = A\mathbf{x}$ in R^4. The domain of T_A is R^5; the codomain is R^4.

(b) $T_A(\mathbf{x}) = A\mathbf{x}$ maps any vector $\mathbf{x}$ in R^4 into a vector $\mathbf{w} = A\mathbf{x}$ in R^5. The domain of T_A is R^4; the codomain is R^5.

(c) $T_A(\mathbf{x}) = A\mathbf{x}$ maps any vector $\mathbf{x}$ in R^4 into a vector $\mathbf{w} = A\mathbf{x}$ in R^4. The domain of T_A is R^4; the codomain is R^4.

(d) $T_A(\mathbf{x}) = A\mathbf{x}$ maps any vector $\mathbf{x}$ in $R^1 = R$ into a vector $\mathbf{w} = A\mathbf{x}$ in R^3. The domain of T_A is R; the codomain is R^3.

3. **(a)** The transformation maps any vector $\mathbf{x}$ in R^2 into a vector $\mathbf{w}$ in R^2. Its domain is R^2; the codomain is R^2.

(b) The transformation maps any vector $\mathbf{x}$ in R^2 into a vector $\mathbf{w}$ in R^3. Its domain is R^2; the codomain is R^3.

4. **(a)** The transformation maps any vector $\mathbf{x}$ in R^3 into a vector $\mathbf{w}$ in R^3. Its domain is R^3; the codomain is R^3.

14. Find the standard matrix for the operator T defined by the formula.

a. $T(x_1, x_2) = (2x_1 - x_2, x_1 + x_2)$

b. $T(x_1, x_2) = (x_1, x_2)$

c. $T(x_1, x_2, x_3) = (x_1 + 2x_2 + x_3, x_1 + 5x_2, x_3)$

d. $T(x_1, x_2, x_3) = (4x_1, 7x_2, -8x_3)$

15. Find the standard matrix for the operator $T : R^3 \to R^3$ defined by

$$\begin{aligned} w_1 &= 3x_1 + 5x_2 - x_3 \\ w_2 &= 4x_1 - x_2 + x_3 \\ w_3 &= 3x_1 + 2x_2 - x_3 \end{aligned}$$

and then compute $T(-1, 2, 4)$ by directly substituting in the equations and then by matrix multiplication.

16. Find the standard matrix for the transformation $T : R^4 \to R^2$ defined by

$$\begin{aligned} w_1 &= 2x_1 + 3x_2 - 5x_3 - x_4 \\ w_2 &= x_1 - 5x_2 + 2x_3 - 3x_4 \end{aligned}$$

and then compute $T(1, -1, 2, 4)$ by directly substituting in the equations and then by matrix multiplication.

14. **(a)** $T(x_1,x_2)=\begin{bmatrix}2x_1-x_2\\x_1+x_2\end{bmatrix}=\begin{bmatrix}2&-1\\1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$; the standard matrix is $\begin{bmatrix}2&-1\\1&1\end{bmatrix}$

(b) $T(x_1,x_2)=\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$; the standard matrix is $\begin{bmatrix}1&0\\0&1\end{bmatrix}$

(c) $T(x_1,x_2,x_3)=\begin{bmatrix}x_1+2x_2+x_3\\x_1+5x_2\\x_3\end{bmatrix}=\begin{bmatrix}1&2&1\\1&5&0\\0&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$; the standard matrix is $\begin{bmatrix}1&2&1\\1&5&0\\0&0&1\end{bmatrix}$

(d) $T(x_1,x_2,x_3)=\begin{bmatrix}4x_1\\7x_2\\-8x_3\end{bmatrix}=\begin{bmatrix}4&0&0\\0&7&0\\0&0&-8\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$; the standard matrix is $\begin{bmatrix}4&0&0\\0&7&0\\0&0&-8\end{bmatrix}$

15. The given equations can be expressed in matrix form as $\begin{bmatrix}w_1\\w_2\\w_3\end{bmatrix}=\begin{bmatrix}3&5&-1\\4&-1&1\\3&2&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ therefore the standard matrix for this operator is $\begin{bmatrix}3&5&-1\\4&-1&1\\3&2&-1\end{bmatrix}$.

By directly substituting $(-1,2,4)$ for (x_1,x_2,x_3) into the given equation we obtain

$$w_1=-(3)(1)+(5)(2)-(1)(4)=3$$

$$w_2=-(4)(1)-(1)(2)+(1)(4)=-2$$

$$w_3=-(3)(1)+(2)(2)-(1)(4)=-3$$

By matrix multiplication, $\begin{bmatrix}w_1\\w_2\\w_3\end{bmatrix}=\begin{bmatrix}3&5&-1\\4&-1&1\\3&2&-1\end{bmatrix}\begin{bmatrix}-1\\2\\4\end{bmatrix}=\begin{bmatrix}-(3)(1)+(5)(2)-(1)(4)\\-(4)(1)-(1)(2)+(1)(4)\\-(3)(1)+(2)(2)-(1)(4)\end{bmatrix}=\begin{bmatrix}3\\-2\\-3\end{bmatrix}$.

Task for Students

As per Course outline : Do (Q.1 till 20 from Ex # 1.8)

Network Analysis

The concept of a *network* appears in a variety of applications. Loosely stated, a ***network*** is a set of ***branches*** through which something "flows." For example, the branches might be electrical wires through which electricity flows, pipes through which water or oil flows, traffic lanes through which vehicular traffic flows, or economic linkages through which money flows, to name a few possibilities.

In most networks, the branches meet at points, called ***nodes*** or ***junctions***, where the flow divides. For example, in an electrical network, nodes occur where three or more wires join, in a traffic network they occur at street intersections, and in a financial network they occur at banking centers where incoming money is distributed to individuals or other institutions.

In the study of networks, there is generally some numerical measure of the rate at which the medium flows through a branch. For example, the flow rate of electricity is often measured in amperes, the flow rate of water or oil in gallons per minute, the flow rate of traffic in vehicles per hour, and the flow rate of European currency in millions of Euros per day. We will restrict our attention to networks in which there is ***flow conservation*** at each node, by which we mean that *the rate of flow into any node is equal to the rate of flow out of that node.* This ensures that the flow medium does not build up at the nodes and block the free movement of the medium through the network.

EXAMPLE 1 | Network Analysis Using Linear Systems

Figure 1.10.1 shows a network with four nodes in which the flow rate and direction of flow in certain branches are known. Find the flow rates and directions of flow in the remaining branches.

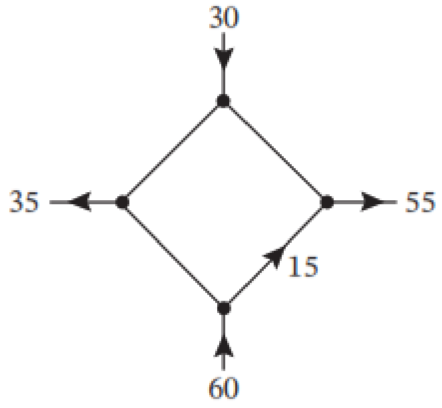


FIGURE 1.10.1

Solution As illustrated in Figure 1.10.2, we have assigned arbitrary directions to the unknown flow rates x_1, x_2, and x_3. We need not be concerned if some of the directions are incorrect, since an incorrect direction will be signaled by a negative value for the flow rate when we solve for the unknowns.

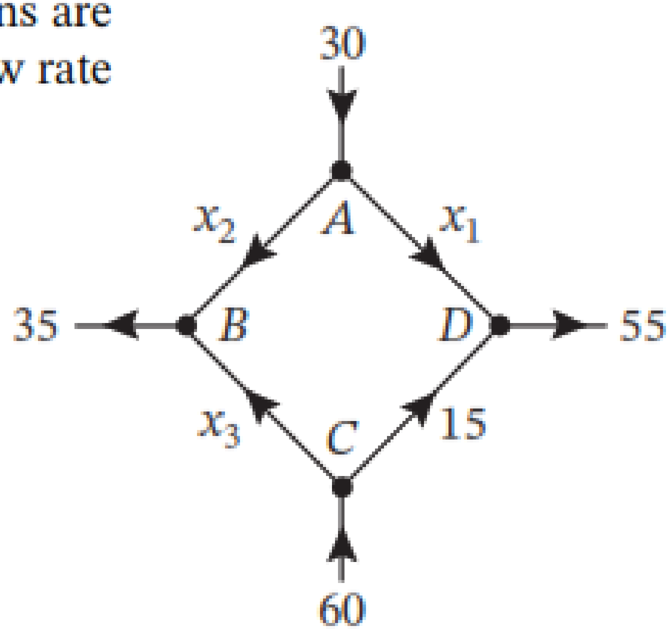


FIGURE 1.10.2

It follows from the conservation of flow at node A that

$$x_1 + x_2 = 30$$

Similarly, at the other nodes we have

$$\begin{aligned} x_2 + x_3 &= 35 \quad &(\text{node } B) \\ x_3 + 15 &= 60 \quad &(\text{node } C) \\ x_1 + 15 &= 55 \quad &(\text{node } D) \end{aligned}$$

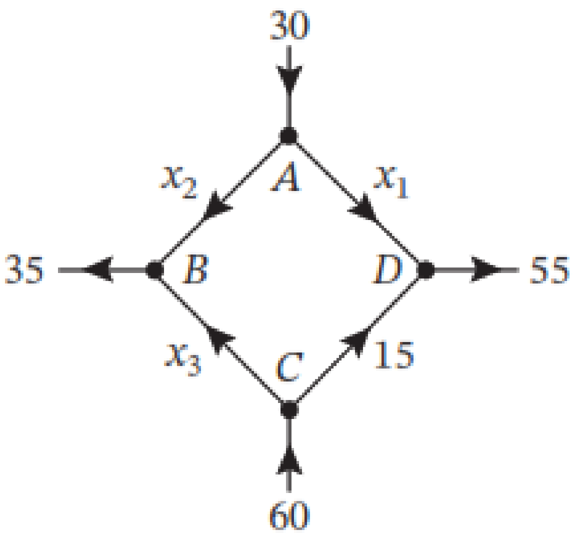


FIGURE 1.10.2

These four conditions produce the linear system

$$\begin{aligned} x_1 + x_2 \qquad\quad &= 30 \\ x_2 + x_3 &= 35 \\ x_3 &= 45 \\ x_1 \qquad\qquad\quad &= 40 \end{aligned}$$

which we can now try to solve for the unknown flow rates. In this particular case the system is sufficiently simple that it can be solved by inspection (work from the bottom up). We leave it for you to confirm that the solution is

$$x_1 = 40, \quad x_2 = -10, \quad x_3 = 45$$

The fact that x_2 is negative tells us that the direction assigned to that flow in Figure 1.10.2 is incorrect; that is, the flow in that branch is *into* node A.

EXAMPLE 2 | Design of Traffic Patterns

The network in Figure 1.10.3*a* shows a proposed plan for the traffic flow around a new park that will house the Liberty Bell in Philadelphia, Pennsylvania. The plan calls for a computerized traffic light at the north exit on Fifth Street, and the diagram indicates the average number of vehicles per hour that are expected to flow in and out of the streets that border the complex. All streets are one-way.

(*a*) How many vehicles per hour should the traffic light let through to ensure that the average number of vehicles per hour flowing into the complex is the same as the average number of vehicles flowing out?

(*b*) Assuming that the traffic light has been set to balance the total flow in and out of the complex, what can you say about the average number of vehicles per hour that will flow along the streets that border the complex?

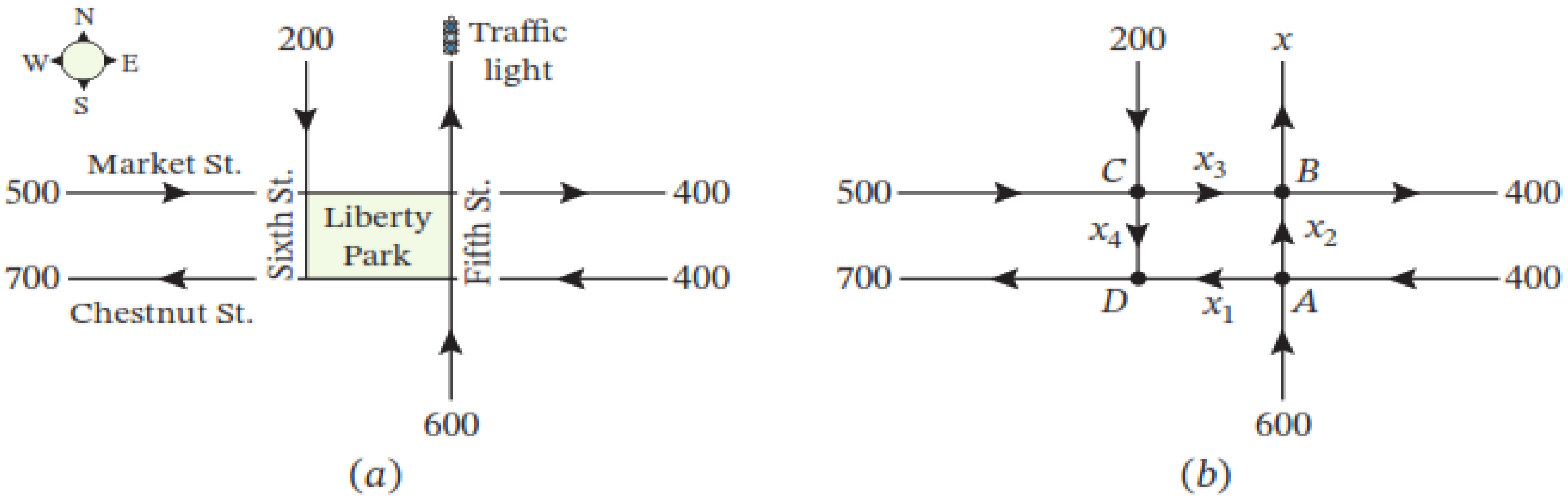
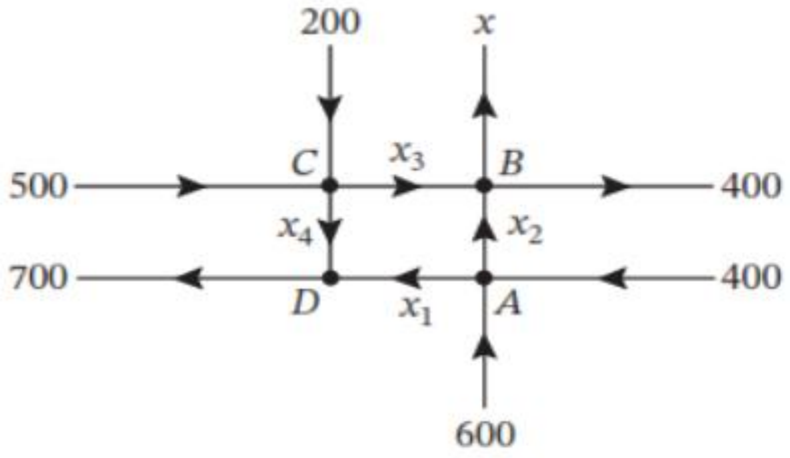


FIGURE 1.10.3



(b)

Solution (a) If, as indicated in **Figure 1.10.3b**, we let x denote the number of vehicles per hour that the traffic light must let through, then the total number of vehicles per hour that flow in and out of the complex will be

Flowing in: $500 + 400 + 600 + 200 = 1700$

Flowing out: $x + 700 + 400$

Equating the flows in and out shows that the traffic light should let $x = 600$ vehicles per hour pass through.

Solution (*b*) To avoid traffic congestion, the flow in must equal the flow out at each intersection. For this to happen, the following conditions must be satisfied:

Intersection	Flow In		Flow Out
A	$400 + 600$	$=$	$x_1 + x_2$
B	$x_2 + x_3$	$=$	$400 + x$
C	$500 + 200$	$=$	$x_3 + x_4$
D	$x_1 + x_4$	$=$	700

Thus, with $x = 600$, as computed in part (a), we obtain the following linear system:

$$\begin{aligned} x_1 + x_2 \qquad\qquad &= 1000 \\ x_2 + x_3 \qquad &= 1000 \\ x_3 + x_4 &= 700 \\ x_1 \qquad\qquad + x_4 &= 700 \end{aligned}$$

We leave it for you to show that the system has infinitely many solutions and that these are given by the parametric equations

$$x_1 = 700 - t, \quad x_2 = 300 + t, \quad x_3 = 700 - t, \quad x_4 = t \tag{1}$$

However, the parameter t is not completely arbitrary here, since there are physical constraints to be considered. For example, the average flow rates must be nonnegative since we have assumed the streets to be one-way, and a negative flow rate would indicate a flow in the wrong direction. This being the case, we see from (1) that t can be any real number that satisfies $0 \leq t \leq 700$, which implies that the average flow rates along the streets will fall in the ranges

$$0 \leq x_1 \leq 700, \quad 300 \leq x_2 \leq 1000, \quad 0 \leq x_3 \leq 700, \quad 0 \leq x_4 \leq 700$$

3. The accompanying figure shows a network of one-way streets with traffic flowing in the directions indicated. The flow rates along the streets are measured as the average number of vehicles per hour.

a. Set up a linear system whose solution provides the unknown flow rates.

b. Solve the system for the unknown flow rates.

c. If the flow along the road from A to B must be reduced for construction, what is the minimum flow that is required to keep traffic flowing on all roads?

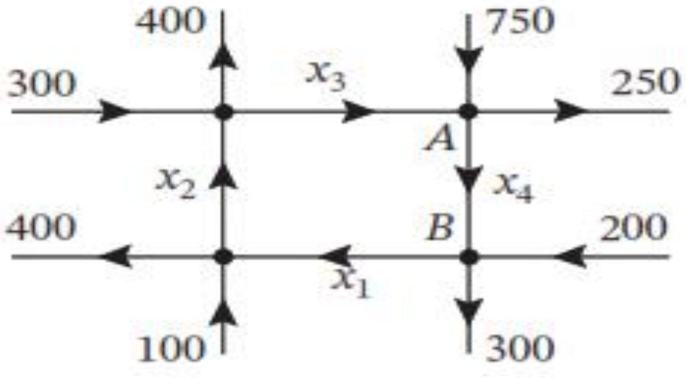


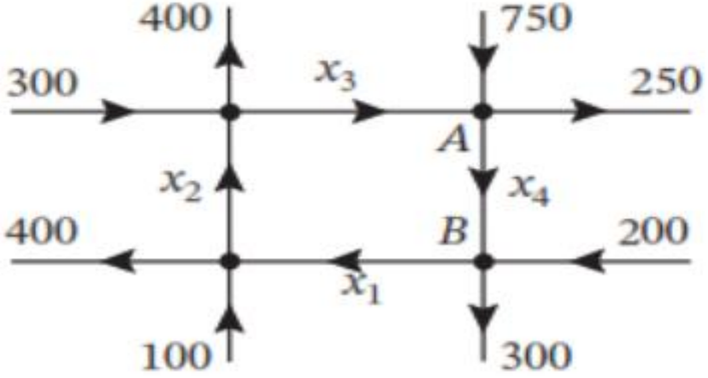
FIGURE Ex-3

(a) There are four nodes – each of them corresponds to an equation.

Network node	Flow In		Flow Out
top left	$x_2 + 300$	$=$	$x_3 + 400$
top right (A)	$x_3 + 750$	$=$	$x_4 + 250$
bottom left	$x_1 + 100$	$=$	$x_2 + 400$
bottom right (B)	$x_4 + 200$	$=$	$x_1 + 300$

This system can be rearranged as follows

$$\begin{aligned} x_2 - x_3 &= 100 \\ x_3 - x_4 &= -500 \\ x_1 - x_2 &= 300 \\ -x_1 + x_4 &= 100 \end{aligned}$$



(b) The augmented matrix of the linear system obtained in part (a) $\begin{bmatrix} 0 & 1 & -1 & 0 & 100 \\ 0 & 0 & 1 & -1 & -500 \\ 1 & -1 & 0 & 0 & 300 \\ -1 & 0 & 0 & 1 & 100 \end{bmatrix}$ has the reduced row echelon form $\begin{bmatrix} 1 & 0 & 0 & -1 & -100 \\ 0 & 1 & 0 & -1 & -400 \\ 0 & 0 & 1 & -1 & -500 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. If we assign x_4 the arbitrary value s, the general solution is given by the formulas

$$x_1 = -100 + s,\ x_2 = -400 + s,\ x_3 = -500 + s,\ x_4 = s$$

(c) In order for all x_i values to remain positive, we must have $s > 500$. Therefore, to keep the traffic flowing on all roads, the flow from A to B must exceed 500 vehicles per hour.

Task for Students

As per Course outline : Do (Q.1 till 4 from Ex # 1.10)

2.1 (1-32)

2.2 (1-23)

2.3(1-29,31,32)

Chapter # 04

Vectors in R^n

- An ordered n-tuple :

  a sequence of n real numbers $(x_1, x_2, \ldots, x_n)$

- R^n-space :

  the set of all ordered n-tuples

$n = 1$ R^1-space = set of all real numbers
(R^1-space can be represented geometrically by the x-axis)

$n = 2$ R^2-space = set of all ordered pair of real numbers (x_1, x_2)
(R^2-space can be represented geometrically by the xy-plane)

$n = 3$ R^3-space = set of all ordered triple of real numbers (x_1, x_2, x_3)
(R^3-space can be represented geometrically by the xyz-space)

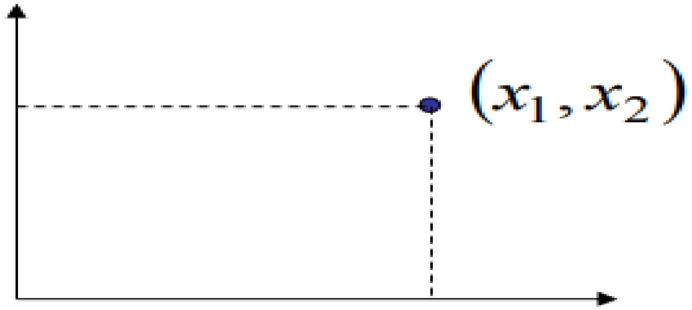
$n = 4$ R^4-space = set of all ordered quadruple of real numbers (x_1, x_2, x_3, x_4)

- Notes:

  (1) An n-tuple $(x_1, x_2, \cdots, x_n)$ can be viewed as a point in R^n with the x_i's as its coordinates

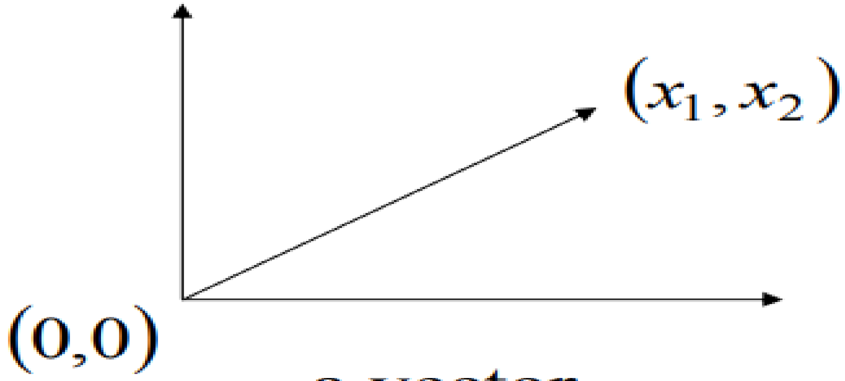
  (2) An n-tuple $(x_1, x_2, \cdots, x_n)$ also can be viewed as a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ in R^n with the x_i's as its components

- Ex:1



a point

or



a vector

※ A vector on the plane is expressed geometrically by a directed line segment whose initial point is the origin and whose terminal point is the point (x_1, x_2)

$$\mathbf{u} = (u_1, u_2, \cdots, u_n), \quad \mathbf{v} = (v_1, v_2, \cdots, v_n) \qquad \text{(two vectors in } R^n\text{)}$$

- Equality:

  $\mathbf{u} = \mathbf{v}$ if and only if $u_1 = v_1$, $u_2 = v_2, \ldots\ldots, u_n = v_n$

- Vector addition (the sum of **u** and **v**):

  $$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, \cdots, u_n + v_n)$$

- Scalar multiplication (the scalar multiple of **u** by c):

  $$c\mathbf{u} = (cu_1, cu_2, \cdots, cu_n)$$

- Notes:

  The sum of two vectors and the scalar multiple of a vector in R^n are called the standard operations in R^n

- Difference between **u** and **v**:

$$\mathbf{u}-\mathbf{v} \equiv \mathbf{u}+(-1)\mathbf{v} = (u_1 - v_1,\ u_2 - v_2,\ u_3 - v_3, \ldots,\ u_n - v_n)$$

- Zero vector :

$$\mathbf{0} = (0, 0, \ldots, 0)$$

Notations for Vectors

Up to now we have been writing vectors in R^n using the notation

$$\mathbf{v} = (v_1, v_2, \ldots, v_n) \tag{15}$$

We call this the ***comma-delimited*** form. However, since a vector in R^n is just a list of its n components in a specific order, any notation that displays those components in the correct order is a valid way of representing the vector. For example, the vector in (15) can be written as

$$\mathbf{v} = [v_1 \quad v_2 \quad \cdots \quad v_n] \tag{16}$$

which is called ***row-vector*** form, or as

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \tag{17}$$

which is called ***column-vector*** form. The choice of notation is often a matter of taste or convenience, but sometimes the nature of a problem will suggest a preferred notation. Notations (15), (16), and (17) will all be used at various places in this text.

The following theorem summarizes the most important properties of vector operations.

Theorem 3.1.1

If $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ are vectors in R^n, and if k and m are scalars, then:

(a) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

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

(c) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

(d) $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

(e) $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$

(f) $(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$

(g) $k(m\mathbf{u}) = (km)\mathbf{u}$

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

Linear Combinations

Addition, subtraction, and scalar multiplication are frequently used in combination to form new vectors. For example, if $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are vectors in R^n, then the vectors

$$\mathbf{u} = 2\mathbf{v}_1 + 3\mathbf{v}_2 + \mathbf{v}_3 \quad \text{and} \quad \mathbf{w} = 7\mathbf{v}_1 - 6\mathbf{v}_2 + 8\mathbf{v}_3$$

are formed in this way. In general, we make the following definition.

Definition 4

If $\mathbf{w}$ is a vector in R^n, then $\mathbf{w}$ is said to be a ***linear combination*** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ in R^n if it can be expressed in the form

$$\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r \tag{14}$$

where $k_1, k_2, \ldots, k_r$ are scalars. These scalars are called the ***coefficients*** of the linear combination. In the case where $r = 1$, Formula (14) becomes $\mathbf{w} = k_1\mathbf{v}_1$, so that a linear combination of a single vector is just a scalar multiple of that vector.

Note that this definition of a linear combination is consistent with that given in the context of matrices (see Definition 6 in Section 1.3).

JUST FOR UNDERSTANDING

Application of Linear Combinations to Color Models

Colors on computer monitors are commonly based on what is called the **RGB *color model***. Colors in this system are created by adding together percentages of the primary colors red (R), green (G), and blue (B). One way to do this is to identify the primary colors with the vectors

$$\mathbf{r} = (1, 0, 0) \quad \text{(pure red)},$$
$$\mathbf{g} = (0, 1, 0) \quad \text{(pure green)},$$
$$\mathbf{b} = (0, 0, 1) \quad \text{(pure blue)}$$

in R^3 and to create all other colors by forming linear combinations of $\mathbf{r}$, $\mathbf{g}$, and $\mathbf{b}$ using coefficients between 0 and 1, inclusive; these coefficients represent the percentage of each pure color in the mix. The set of all such color vectors is called **RGB *space*** or the **RGB *color cube*** (Figure 3.1.14). Thus, each color vector $\mathbf{c}$ in this cube is expressible as a linear combination of the form

$$\begin{aligned}\mathbf{c} &= k_1\mathbf{r} + k_2\mathbf{g} + k_3\mathbf{b} \\ &= k_1(1, 0, 0) + k_2(0, 1, 0) + k_3(0, 0, 1) \\ &= (k_1, k_2, k_3)\end{aligned}$$

where $0 \le k_i \le 1$. As indicated in the figure, the corners of the cube represent the pure primary colors together with the colors black, white, magenta, cyan, and yellow. The vectors along the diagonal running from black to white correspond to shades of gray.

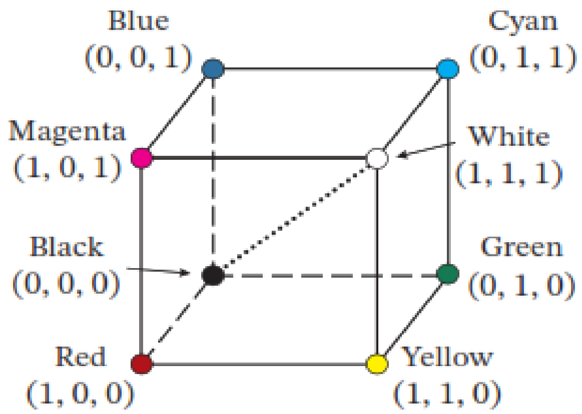


FIGURE 3.1.14

Vector Spaces

- Vector spaces:

  Let V be a set on which two operations (addition and scalar multiplication) are defined. **If the following ten axioms are satisfied** for every element **u**, **v**, and **w** in V and every scalar (real number) c and d, then V **is called a vector space**, and the **elements** in V are called **vectors**

  Addition:

  (1) $\mathbf{u}+\mathbf{v}$ is in V

  (2) $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$

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

  (4) V has a zero vector $\mathbf{0}$ such that for every $\mathbf{u}$ in V, $\mathbf{u}+\mathbf{0} = \mathbf{u}$

  (5) For every $\mathbf{u}$ in V, there is a vector in V denoted by $-\mathbf{u}$ such that $\mathbf{u}+(-\mathbf{u}) = \mathbf{0}$

Scalar multiplication:

(6) $c\mathbf{u}$ is in V

(7) $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$

(8) $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$

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

(10) $1(\mathbf{u})=\mathbf{u}$

※ This type of definition is called an **abstraction** because you abstract a collection of properties from R^n to form the axioms for defining a more general space V

※ Thus, we can conclude that R^n is of course a vector space

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

OR
WE CAN DEFINE VECTOR SPACE AS;

Definition 1

Let V be an arbitrary nonempty set of objects for which two operations are defined: addition and multiplication by numbers called ***scalars***. By ***addition*** we mean a rule for associating with each pair of objects $\mathbf{u}$ and $\mathbf{v}$ in V an object $\mathbf{u}+\mathbf{v}$, called the ***sum*** of $\mathbf{u}$ and $\mathbf{v}$; by ***scalar multiplication*** we mean a rule for associating with each scalar k and each object $\mathbf{u}$ in V an object $k\mathbf{u}$, called the ***scalar multiple*** of $\mathbf{u}$ by k. If the following axioms are satisfied by all objects $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V and all scalars k and m, then we call V a ***vector space*** and we call the objects in V ***vectors***.

1. If $\mathbf{u}$ and $\mathbf{v}$ are objects in V, then $\mathbf{u}+\mathbf{v}$ is in V.
2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$
3. $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$
4. There exists an object in V, called the ***zero vector***, that is denoted by $\mathbf{0}$ and has the property that $\mathbf{0}+\mathbf{u}=\mathbf{u}+\mathbf{0}=\mathbf{u}$ for all $\mathbf{u}$ in V.
5. For each $\mathbf{u}$ in V, there is an object $-\mathbf{u}$ in V, called a ***negative*** of $\mathbf{u}$, such that $\mathbf{u}+(-\mathbf{u})=(-\mathbf{u})+\mathbf{u}=\mathbf{0}$.
6. If k is any scalar and $\mathbf{u}$ is any object in V, then $k\mathbf{u}$ is in V.
7. $k(\mathbf{u}+\mathbf{v})=k\mathbf{u}+k\mathbf{v}$
8. $(k+m)\mathbf{u}=k\mathbf{u}+m\mathbf{u}$
9. $k(m\mathbf{u})=(km)(\mathbf{u})$
10. $1\mathbf{u}=\mathbf{u}$

In this text scalars will be either real numbers or complex numbers. Vector spaces with real scalars will be called ***real vector spaces*** and those with complex scalars will be called ***complex vector spaces***. For now we will consider only real vector spaces.

Steps to Show That a Set with Two Operations Is a Vector Space

Step 1. Identify the set V of objects that will become vectors.

Step 2. Identify the addition and scalar multiplication operations on V.

Step 3. Verify Axioms 1 and 6; that is, adding two vectors in V produces a vector in V, and multiplying a vector in V by a scalar also produces a vector in V.

Axiom 1 is called ***closure under addition***, and Axiom 6 is called ***closure under scalar multiplication***.

Step 4. Confirm that Axioms 2, 3, 4, 5, 7, 8, 9, and 10 hold.

PROPERTIES OF VECTORS:

Theorem 4.1.1

Let V be a vector space, $\mathbf{u}$ a vector in V, and k a scalar; then:

(a) $0\mathbf{u} = \mathbf{0}$

(b) $k\mathbf{0} = \mathbf{0}$

(c) $(-1)\mathbf{u} = -\mathbf{u}$

(d) If $k\mathbf{u} = \mathbf{0}$, then $k = 0$ or $\mathbf{u} = \mathbf{0}$.

SUBSPACES:

Definition 1

A subset W of a vector space V is called a ***subspace*** of V if W is itself a vector space under the addition and scalar multiplication defined on V.

- Subspace:

$(V,+,\cdot)$: a vector space

$$\left.\begin{array}{l} W \neq \Phi \\ W \subseteq V \end{array}\right\} : \text{a nonempty subset of } V$$

$(W,+,\cdot)$: The nonempty subset W is called a subspace **if W is a vector space** under the operations of addition and scalar multiplication defined on V

- Trivial subspace:

Every vector space V has at least two subspaces

(1) Zero vector space $\{\mathbf{0}\}$ is a subspace of V (It satisfies the ten axioms)

(2) V is a subspace of V

* Any subspaces other than these two are called proper (or nontrivial) subspaces

Axioms that is not inherited for Subspace (Rest are supposed to be inherited as ***W*** *lies in Vector Space* ***V)***

Axiom 1—Closure of W under addition

Axiom 4—Existence of a zero vector in W

Axiom 5—Existence of a negative in W for every vector in W

Axiom 6—Closure of W under scalar multiplication

NOTE: It is necessary to verify that $\boldsymbol{W}$ is closed under addition and scalar multiplication since it is possible that adding two vectors in $\boldsymbol{W}$ or multiplying a vector in W by a scalar produces a vector in V that is outside of W. If such condition happens, then its not a subspace (Example is in Fig. 4.2.1)

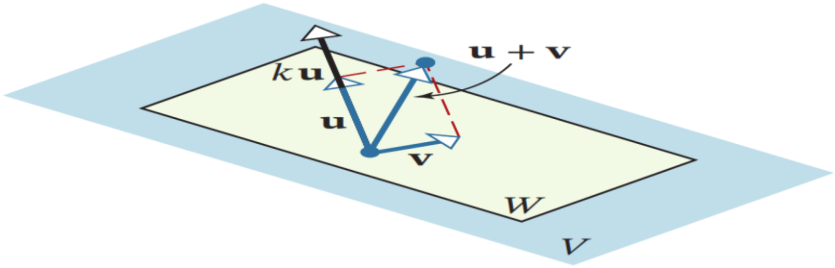


FIGURE 4.2.1 The vectors **u** and **v** are in W, but the vectors **u** + **v** and k**u** are not.

Subspace Test

Theorem 4.2.1

Subspace Test

If W is a nonempty set of vectors in a vector space V, then W is a subspace of V if and only if the following conditions are satisfied.

(*a*) If $\mathbf{u}$ and $\mathbf{v}$ are vectors in W, then $\mathbf{u} + \mathbf{v}$ is in W.

(*b*) If k is a scalar and $\mathbf{u}$ is a vector in W, then $k\mathbf{u}$ is in W.

Again NOTE:

> Note that every vector space has at least two subspaces, itself and its zero subspace.

- Ex : A subspace of $M_{2\times 2}$

  Let W be the set of all 2×2 symmetric matrices. Show that W is a subspace of the vector space $M_{2\times 2}$, with the standard operations of matrix addition and scalar multiplication

Sol:

First, we knon that W, the set of all 2×2 symmetric matrices, is an nonempty subset of the vector space $M_{2\times 2}$

Second,

$$A_1 \in W, A_2 \in W \Rightarrow (A_1 + A_2)^T = A_1^T + A_2^T = A_1 + A_2 \quad (A_1 + A_2 \in W)$$

$$c \in R,\ A \in W \Rightarrow (cA)^T = cA^T = cA \quad (cA \in W)$$

The definition of a symmetric matrix A is that $A^T = A$

Thus, Th. 2.4 is applied to obtain that W is a subspace of M_{2x2}

- Ex : The set of singular matrices is not a subspace of $M_{2\times 2}$

Let W be the set of singular (noninvertible) matrices of order 2. Show that W is not a subspace of $M_{2\times 2}$ with the standard matrix operations

Sol:

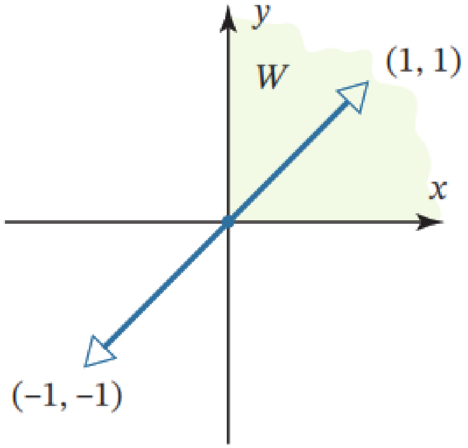
$$A=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\in W,\ B=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\in W$$

$$\because A+B=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I\notin W$$ (W is not closed under vector addition)

$\therefore W$ is not a subspace of $M_{2\times 2}$

- Ex : The set of first-quadrant vectors is not a subspace of R^2

  Show that $W=\{(x_1,x_2):x_1\geq 0 \text{ and } x_2\geq 0\}$, with the standard operations, is not a subspace of R^2



Sol:

Let $\mathbf{u}=(1,\ 1)\in W$

$\because (-1)\mathbf{u}=(-1)(1,1)=(-1,-1)\notin W$

(W is not closed under scalar multiplication)

$\therefore W$ is not a subspace of R^2

Building Subspaces

The following theorem provides a useful way of creating a new subspace from known subspaces.

Theorem 4.2.2

If $W_1, W_2, \ldots, W_r$ are subspaces of a vector space V, then the intersection of these subspaces is also a subspace of V.

Solution Spaces of Homogeneous Systems

The solutions of a homogeneous linear system $A\mathbf{x} = \mathbf{0}$ of m equations in n unknowns can be viewed as vectors in R^n. The following theorem provides an important insight into the geometric structure of the solution set.

Theorem 4.2.3

The solution set of a homogeneous system $A\mathbf{x} = \mathbf{0}$ of m equations in n unknowns is a subspace of R^n.

EXAMPLE 13 | Solution Spaces of Homogeneous Systems

In each part the solution of the linear system is provided. Give a geometric description of the solution set.

(a) $\begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ -2 & 4 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ 4 & 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Solution (*a*) The solutions are

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

from which it follows that

$$x = 2y - 3z \quad \text{or} \quad x - 2y + 3z = 0$$

This is the equation of a plane through the origin that has $\mathbf{n} = (1, -2, 3)$ as a normal.

Solution (*b*) The solutions are

$$x = -5t, \quad y = -t, \quad z = t$$

which are parametric equations for the line through the origin that is parallel to the vector $\mathbf{v} = (-5, -1, 1)$.

Solution (*c*) The only solution is $x = 0$, $y = 0$, $z = 0$, so the solution space consists of the single point $\{\mathbf{0}\}$.

Solution (*d*) This linear system is satisfied by all real values of x, y, and z, so the solution space is all of R^3.

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

SEE EXAMPLES 1 till 12 FROM TEXT BOOK

DO Questions from Ex # 4.1 (1-7,11)
Examples: 1-5,7 & Ex # 4.2 (1-3)

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Linear Algebra (MT-1004)

Lecture# 5

Linear Combination in a Vector Space

- Linear combination:

  A vector $\mathbf{u}$ in a vector space V is called a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ in V if $\mathbf{u}$ can be written in the form

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k,$$

where $c_1, c_2, \ldots, c_k$ are real-number scalars

Spanning Set

- This section introduces the spanning set, this notion is associated with the representation of any vector in a vector space as a **linear combination** of a selected set of vectors in that vector space.

- Spanning Set:
  If $S=\{\mathbf{v}_1,\mathbf{v}_2,\ldots.., \mathbf{v}_k\}$ is a set of vectors in a vector space W of V consisting of all linear combinations of the vectors in S is called space spanned by $\mathbf{v}_1$, $\mathbf{v}_2,\ldots\ldots,\mathbf{v}_k$ and we say that the vectors $\mathbf{v}_1$, $\mathbf{v}_2,\ldots\ldots..,\mathbf{v}_k$ span W. It is denoted by

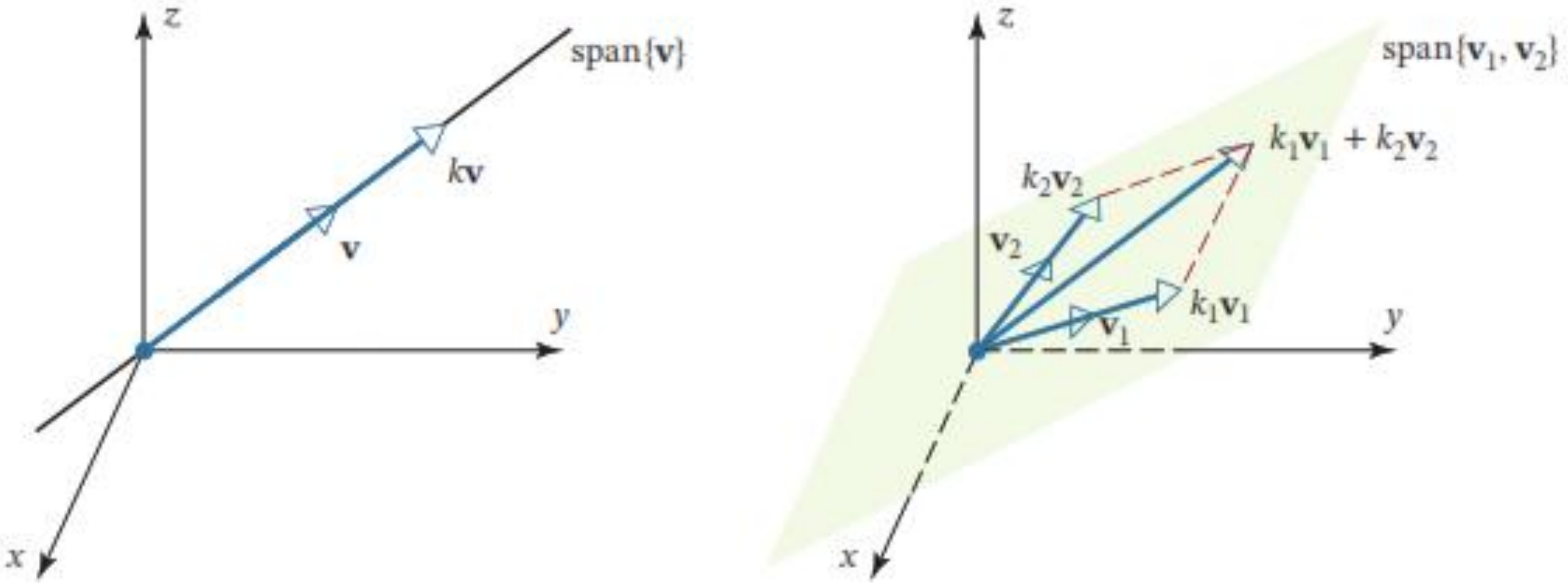
$$W = \text{Span}(S) \text{ or } W = \text{Span}\{\mathbf{v}_1,\mathbf{v}_2,\ldots\ldots..,\mathbf{v}_k\}$$

Theorem 4.3.1

If $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is a nonempty set of vectors in a vector space V, then:

(*a*) The set W of all possible linear combinations of the vectors in S is a subspace of V.

(*b*) The set W in part (a) is the "smallest" subspace of V that contains all of the vectors in S in the sense that any other subspace that contains those vectors contains W.



(*a*) span{**v**} is the line through the origin determined by **v**

(*b*) span{$\mathbf{v}_1$, $\mathbf{v}_2$} is the plane through the origin determined by $\mathbf{v}_1$ and $\mathbf{v}_2$

FIGURE 4.3.1

A Procedure for Identifying Spanning Sets

Step 1. Let $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ be a given set of vectors in V, and let $\mathbf{x}$ be an arbitrary vector in V.

Step 2. Set up the augmented matrix for the linear system that results by equating corresponding components on the two sides of the vector equation

$$k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + \cdots + k_r\mathbf{w}_r = \mathbf{x} \tag{2}$$

Step 3. Use the techniques developed in Chapters 1 and 2 to investigate the consistency or inconsistency of that system. If it is consistent for *all* choices of $\mathbf{x}$, the vectors in S span V, and if it is inconsistent for *some* vector $\mathbf{x}$, they do not.

Problem Let $\mathbf{v}_1 = (2,5)$ and $\mathbf{v}_2 = (1,3)$. Show that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a spanning set for $\mathbb{R}^2$.

Take any vector $\mathbf{w} = (a, b) \in \mathbb{R}^2$. We have to check that there exist $r_1, r_2 \in \mathbb{R}$ such that

$$\mathbf{w} = r_1\mathbf{v}_1 + r_2\mathbf{v}_2 \iff \begin{cases} 2r_1 + r_2 = a \\ 5r_1 + 3r_2 = b \end{cases}$$

Coefficient matrix: $C = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$. $\det C = 1 \neq 0$.

Since the matrix C is invertible, the system has a unique solution for any a and b.
Thus $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2) = \mathbb{R}^2$.

EXAMPLE 4 | Linear Combinations

Consider the vectors $\mathbf{u} = (1, 2, -1)$ and $\mathbf{v} = (6, 4, 2)$ in R^3. Show that $\mathbf{w} = (9, 2, 7)$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ and that $\mathbf{w}' = (4, -1, 8)$ is *not* a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Solution In order for $\mathbf{w}$ to be a linear combination of $\mathbf{u}$ and $\mathbf{v}$, there must be scalars k_1 and k_2 such that $\mathbf{w} = k_1\mathbf{u} + k_2\mathbf{v}$; that is,

$$(9, 2, 7) = k_1(1, 2, -1) + k_2(6, 4, 2) = (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2)$$

Equating corresponding components gives

$$\begin{aligned} k_1 + 6k_2 &= 9 \\ 2k_1 + 4k_2 &= 2 \\ -k_1 + 2k_2 &= 7 \end{aligned}$$

Solving this system using Gaussian elimination yields $k_1 = -3$, $k_2 = 2$, so

$$\mathbf{w} = -3\mathbf{u} + 2\mathbf{v}$$

Similarly, for $\mathbf{w}'$ to be a linear combination of $\mathbf{u}$ and $\mathbf{v}$, there must be scalars k_1 and k_2 such that $\mathbf{w}' = k_1\mathbf{u} + k_2\mathbf{v}$; that is,

$$(4, -1, 8) = k_1(1, 2, -1) + k_2(6, 4, 2) = (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2)$$

Equating corresponding components gives

$$\begin{aligned} k_1 + 6k_2 &= 4 \\ 2k_1 + 4k_2 &= -1 \\ -k_1 + 2k_2 &= 8 \end{aligned}$$

This system of equations is inconsistent (verify), so no such scalars k_1 and k_2 exist. Consequently, $\mathbf{w}'$ is not a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

EXAMPLE 5 | Testing for Spanning

Determine whether the vectors $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (1, 0, 1)$, and $\mathbf{v}_3 = (2, 1, 3)$ span the vector space R^3.

Solution We must determine whether an arbitrary vector $\mathbf{b} = (b_1, b_2, b_3)$ in R^3 can be expressed as a linear combination

$$\mathbf{b} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3$$

of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. Expressing this equation in terms of components gives

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

or

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

or

$$\begin{aligned} k_1 + k_2 + 2k_3 &= b_1 \\ k_1 + k_3 &= b_2 \\ 2k_1 + k_2 + 3k_3 &= b_3 \end{aligned}$$

Thus, our problem reduces to ascertaining whether this system is consistent for all values of b_1, b_2, and b_3. One way of doing this is to use parts (*e*) and (*g*) of Theorem 2.3.8, which state that the system is consistent if and only if its coefficient matrix

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

has a nonzero determinant. But this is *not* the case here since $\det(A) = 0$ (verify), so $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ do not span R^3.

Definition 1

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a set of two or more vectors in a vector space V, then S is said to be a ***linearly independent set*** if no vector in S can be expressed as a linear combination of the others. A set that is not linearly independent is said to be ***linearly dependent***. If S has only one vector, we will agree that it is linearly independent if and only if that vector is nonzero.

OR

Theorem Based Definition

Linear Independence and Linear Dependence

- Definitions :

  If $S=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ is a nonempty set of vectors, then the vector equation

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+\ldots+c_k\mathbf{v}_k=\mathbf{0}$$

has at least one solution, namely $c_1=0,\ c_2=0,\ldots,c_k=0$.

If this is the only solution, then S is called a linearly independent set. If there are other solutions, then S is called a linearly dependent set.

- Ex $\boxed{\text{Testing for linear independence}}$

Determine whether the following set of vectors in R^3 is L.I. or L.D.

$$S=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}=\{(1,\ 2,\ 3),(0,\ 1,\ 2),(-2,\ 0,\ 1)\}$$

Sol:

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}\ \Rightarrow\ \begin{aligned} c_1 \qquad\quad -2c_3 &= 0\\ 2c_1+\ c_2+\qquad\ &= 0\\ 3c_1+2c_2+\ c_3 &= 0\end{aligned}$$

$$\Rightarrow\left[\begin{array}{ccc|c} 1 & 0 & -2 & 0\\ 2 & 1 & 0 & 0\\ 3 & 2 & 1 & 0\end{array}\right]\xrightarrow{\text{G.-J. E.}}\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{array}\right]$$

$$\Rightarrow c_1=c_2=c_3=0\quad(\text{only the trivial solution})$$

$$(\text{or } \det(A)=-1\neq 0,\ \text{so there is only the trivial solution})$$

$$\Rightarrow S \text{ is (or } \mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3 \text{ are) linearly independent}$$

- Ex : Testing for linear independence **for polynomials**

  Determine whether the following set of vectors in P_2 is L.I. or L.D.

$$S=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}=\{1+x-2x^2, 2+5x-x^2, x+x^2\}$$

Sol:

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}$$

i.e., $c_1(1+x-2x^2)+c_2(2+5x-x^2)+c_3(x+x^2)=0+0x+0x^2$

$$\Rightarrow \begin{array}{rl} c_1+2c_2 & =0 \\ c_1+5c_2+c_3 & =0 \\ -2c_1-c_2+c_3 & =0 \end{array} \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 1 & 5 & 1 & 0 \\ -2 & -1 & 1 & 0 \end{array}\right] \xrightarrow{\text{G. E.}} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 1 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\Rightarrow$ This system has infinitely many solutions
(i.e., this system has nontrivial solutions, e.g., $c_1=2$, $c_2=-1$, $c_3=3$)

$\Rightarrow$ S is (or $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ are) linearly dependent

Theorem 4.4.2

(*a*) A set with finitely many vectors that contains **0** is linearly dependent.

(*b*) A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.

A Geometric Interpretation of Linear Independence

Linear independence has the following useful geometric interpretations in R^2 and R^3:

- Two vectors in R^2 or R^3 are linearly independent if and only if they do not lie on the same line when they have their initial points at the origin. Otherwise one would be a scalar multiple of the other (**Figure 4.4.3**).

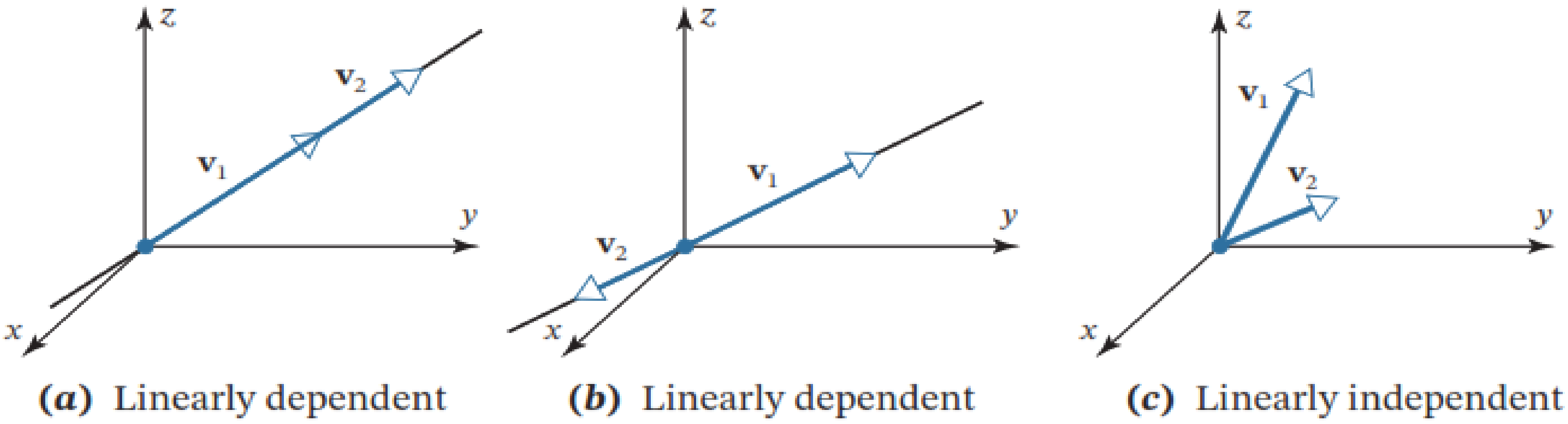
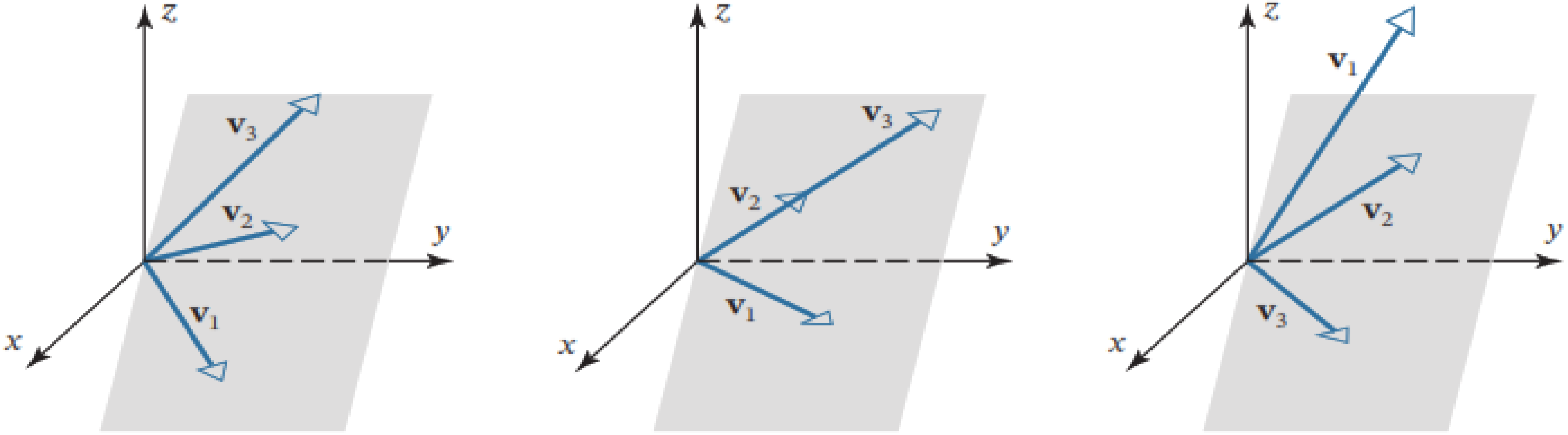


FIGURE 4.4.3

- Three vectors in R^3 are linearly independent if and only if they do not lie in the same plane when they have their initial points at the origin. Otherwise at least one would be a linear combination of the other two (**Figure 4.4.4**).



(a) Linearly dependent **(b)** Linearly dependent **(c)** Linearly independent

FIGURE 4.4.4

Theorem 4.4.3

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ be a set of vectors in R^n. If $r > n$, then S is linearly dependent.

example

Explain why the following form linearly dependent sets of vectors. (Solve this problem by inspection.)

b. $\mathbf{u}_1 = (3, -1)$, $\mathbf{u}_2 = (4, 5)$, $\mathbf{u}_3 = (-4, 7)$ in R^2

Ans: A set of 3 vectors in R^2 must be linearly dependent by Theorem 4.4.3.

Basis

- Basis :

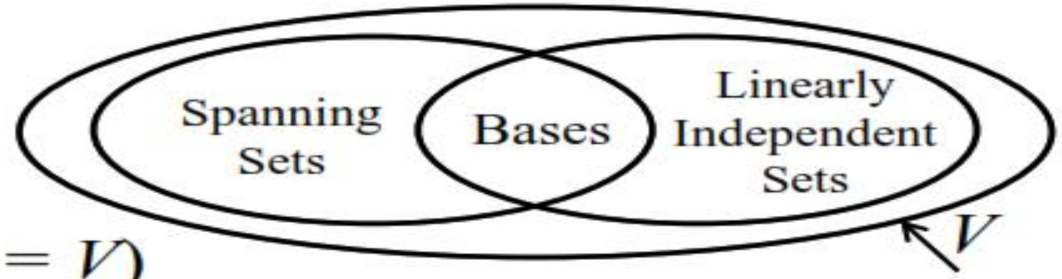
  V: a vector space

  $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subseteq V$

  1) S spans V (i.e., span(S) = V)

  2) S is linearly independent

  $\Rightarrow$ S is called a basis for V



- Notes:

  A basis S must have enough vectors to span V, but not so many vectors that one of them could be written as a linear combination of the other vectors in S

BASIS COORDINATES

If $S=\{v_1, v_2, \ldots, v_n\}$ is a basis for a vector space V, and

$$v=c_1v_1+ c_2v_2+ \ldots\ldots + c_nv_n$$

Is the expression for a vector v in terms of the basis S, then the scalars c1, c2, ….., cn are called the coordinates of v relative to the basis S. The coordinate vector of v relative to S is denoted by

$$(v)_s=(c_1, c_2, \ldots, c_n)$$

Theorem 4.5.1

Uniqueness of Basis Representation

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, then every vector $\mathbf{v}$ in V can be expressed in the form $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ in exactly one way.

Dimension

- Dimension:

  The dimension of a vector space V is defined to be the number of vectors in a basis for V

  V: a vector space $\qquad$ S: a basis for V

  $$\Rightarrow \dim(V) = \#(S) \quad \text{(the number of vectors in a basis } S\text{)}$$

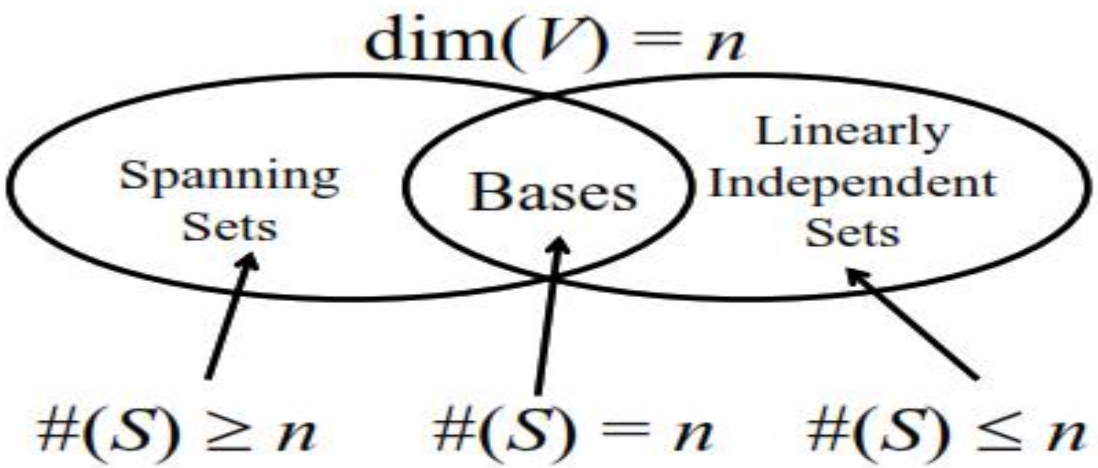
- Finite dimensional:

  A vector space V is finite dimensional if it has a basis consisting of a finite number of elements

- Infinite dimensional:

  If a vector space V is not finite dimensional, then it is called infinite dimensional

- **Notes:**



(1) $\dim(\{\mathbf{0}\}) = 0$

(If V consists of the zero vector alone, the dimension of V is defined as zero)

(2) Given $\dim(V) = n$, for $S \subseteq V$

S: a spanning set $\Rightarrow \#(S) \ge n$

S: a L.I. set $\Rightarrow \#(S) \le n$ (from Theorem 4.2)

S: a basis $\Rightarrow \#(S) = n$ (Since a basis is defined to be a set of L.I. vectors that can spans V, $\#(S) = \dim(V) = n$) (see the above figure)

(3) Given $\dim(V) = n$, if W is a subspace of $V \Rightarrow \dim(W) \le n$

※ For example, if $V = R^3$, you can infer the $\dim(V)$ is 3, which is the number of vectors in the standard basis

Theorem 4.6.2

Let V be a finite-dimensional vector space, and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any basis for V.

(*a*) If a set in V has more than n vectors, then it is linearly dependent.

(*b*) If a set in V has fewer than n vectors, then it does not span V.

- Ex4.2: Find the dimension of a vector space according to the standard basis

※ The simplest way to find the dimension of a vector space is to count the number of vectors in the standard basis for that vector space

(1) Vector space R^n $\Rightarrow$ standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$

$$\Rightarrow \dim(R^n) = n$$

(2) Vector space $M_{m\times n}$ $\Rightarrow$ standard basis $\{E_{ij} \mid 1\le i\le m,\ 1\le j\le n\}$

$$\text{and in } E_{ij} \begin{cases} a_{ij} = 1 \\ \text{other entries are zero} \end{cases}$$

$$\Rightarrow \dim(M_{m\times n}) = mn$$

(3) Vector space $P_n(x)$ $\Rightarrow$ standard basis $\{1, x, x^2, \ldots, x^n\}$

$$\Rightarrow \dim(P_n(x)) = n+1$$

(4) Vector space $P(x)$ $\Rightarrow$ standard basis $\{1, x, x^2, \ldots\}$

- Ex 4.3: Determining the dimension of a subspace of R^3

  (a) $W = \{(d, c-d, c): c \text{ and } d \text{ are real numbers}\}$

  (b) $W = \{(2b, b, 0): b \text{ is a real number}\}$

Sol: (Hint: find a set of L.I. vectors that spans the subspace, i.e., find a basis for subspace.)

(a) $(d, c-d, c) = c(0, 1, 1) + d(1, -1, 0)$

$\Rightarrow S = \{(0, 1, 1), (1, -1, 0)\}$ (S is L.I. and S spans W)

$\Rightarrow S$ is a basis for W

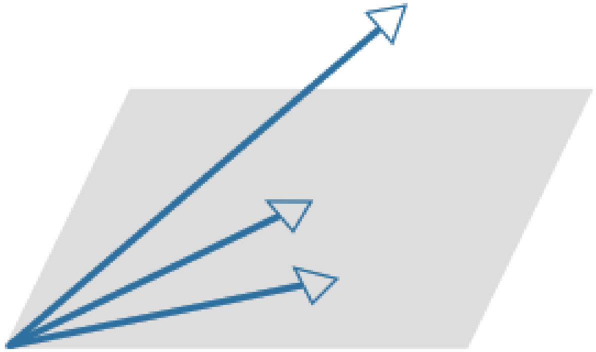
$\Rightarrow \dim(W) = \#(S) = 2$

(b) $\because (2b, b, 0) = b(2, 1, 0)$

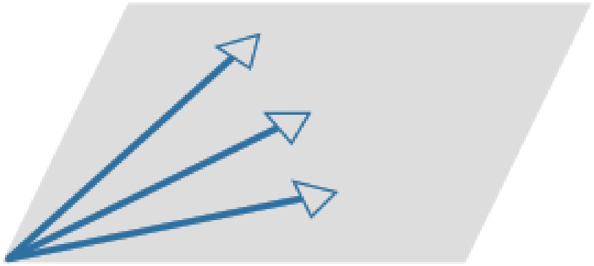
$\Rightarrow S = \{(2, 1, 0)\}$ spans W and S is L.I.

$\Rightarrow S$ is a basis for W

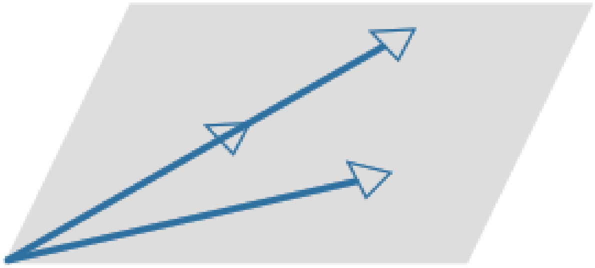
$\Rightarrow \dim(W) = \#(S) = 1$



The vector outside the plane can be adjoined to the other two without affecting their linear independence.



Any of the vectors can be removed, and the remaining two will still span the plane.



Either of the collinear vectors can be removed, and the remaining two will still span the plane.

Theorem 4.6.3

Plus/Minus Theorem

Let S be a nonempty set of vectors in a vector space V.

(*a*) If S is a linearly independent set, and if $\mathbf{v}$ is a vector in V that is outside of span(S), then the set $S \cup \{\mathbf{v}\}$ that results by inserting $\mathbf{v}$ into S is still linearly independent.

(*b*) If $\mathbf{v}$ is a vector in S that is expressible as a linear combination of other vectors in S, and if $S - \{\mathbf{v}\}$ denotes the set obtained by removing $\mathbf{v}$ from S, then S and $S - \{\mathbf{v}\}$ span the same space; that is,

$$\text{span}(S) = \text{span}(S - \{\mathbf{v}\})$$

Theorem 4.6.4

Let V be an n-dimensional vector space, and let S be a set in V with exactly n vectors. Then S is a basis for V if and only if S spans V or S is linearly independent.

1. Every spanning set for a subspace is either a basis for that subspace or has a basis as a subset.
2. Every linearly independent set in a subspace is either a basis for that subspace or can be extended to a basis for it.

Theorem 4.6.5

Let S be a finite set of vectors in a finite-dimensional vector space V.

(a) If S spans V but is not a basis for V, then S can be reduced to a basis for V by removing appropriate vectors from S.

(b) If S is a linearly independent set that is not already a basis for V, then S can be enlarged to a basis for V by inserting appropriate vectors into S.

Do the following from Book:

Spanning Sets Linear Independence		**4.3** (1-13, 17, 18) **4.4** (1-14, 17-20)
Coordinates and Bases Dimensions		**4.5** (1-9, 11-16) **4.6** (1-6)

Linear Algebra (MT-1004)

Lecture # 6

Definition 1

For an $m \times n$ matrix

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

the vectors

$$\begin{aligned} \mathbf{r}_1 &= [a_{11} \quad a_{12} \quad \cdots \quad a_{1n}] \\ \mathbf{r}_2 &= [a_{21} \quad a_{22} \quad \cdots \quad a_{2n}] \\ \vdots & \qquad\qquad\quad \vdots \\ \mathbf{r}_m &= [a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}] \end{aligned}$$

in R^n formed from the rows of A are called the ***row vectors*** of A, and the vectors

$$\mathbf{c}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \quad \mathbf{c}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

in R^m formed from the columns of A are called the ***column vectors*** of A.

Definition 2

If A is an $m \times n$ matrix, then the subspace of R^n spanned by the row vectors of A is denoted by row(A) and is called the ***row space*** of A, and the subspace of R^m spanned by the column vectors of A is denoted by col(A) and is called the ***column space*** of A. The solution space of the homogeneous system of equations $A\mathbf{x} = \mathbf{0}$, which is a subspace of R^n, is denoted by null(A) and is called the ***null space*** of A.

Theorem 4.8.1

A system of linear equations $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A.

EXAMPLE 2 | A Vector **b** in the Column Space of A

Let $A\mathbf{x} = \mathbf{b}$ be the linear system

$$\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

Show that **b** is in the column space of A by expressing it as a linear combination of the column vectors of A.

Solution Solving the system by Gaussian elimination yields (verify)

$$x_1 = 2, \quad x_2 = -1, \quad x_3 = 3$$

It follows from this and Formula (2) that

$$2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

Theorem 4.8.2

If $\mathbf{x}_0$ is any solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$, and if $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a basis for the null space of A, then every solution of $A\mathbf{x} = \mathbf{b}$ can be expressed in the form

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \tag{6}$$

Conversely, for all choices of scalars $c_1, c_2, \ldots, c_k$, the vector $\mathbf{x}$ in this formula is a solution of $A\mathbf{x} = \mathbf{b}$.

Theorem 4.8.3

(*a*) *Row equivalent matrices have the same row space.*

(*b*) *Row equivalent matrices have the same null space.*

Theorem 4.8.4

If a matrix R is in row echelon form, then the row vectors with the leading 1's (the nonzero row vectors) form a basis for the row space of R, and the column vectors with the leading 1's of the row vectors form a basis for the column space of R.

EXAMPLE 3 | Bases for the Row and Column Spaces of a Matrix in Row Echelon Form

Find bases for the row and column spaces of the matrix

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

Solution Since the matrix R is in row echelon form, it follows from Theorem 4.8.4 that the vectors

$$\begin{aligned} \mathbf{r}_1 &= [1 \quad -2 \quad 5 \quad 0 \quad 3] \\ \mathbf{r}_2 &= [0 \quad 1 \quad 3 \quad 0 \quad 0] \\ \mathbf{r}_3 &= [0 \quad 0 \quad 0 \quad 1 \quad 0] \end{aligned}$$

form a basis for the row space of R, and the vectors

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R.

EXAMPLE 4 | Basis for a Row Space by Row Reduction

Find a basis for the row space of the matrix

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

Solution Since elementary row operations do not change the row space of a matrix, we can find a basis for the row space of A by finding a basis for the row space of any row echelon form of A. Reducing A to row echelon form, we obtain (verify)

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

By Theorem 4.8.4, the nonzero row vectors of R form a basis for the row space of R and hence form a basis for the row space of A. These basis vectors are

$$\begin{aligned} \mathbf{r}_1 &= [1 \quad -3 \quad 4 \quad -2 \quad 5 \quad 4] \\ \mathbf{r}_2 &= [0 \quad 0 \quad 1 \quad 3 \quad -2 \quad -6] \\ \mathbf{r}_3 &= [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 5] \end{aligned}$$

Theorem 4.8.5

If A and B are row equivalent matrices, then:

(*a*) A given set of column vectors of A is linearly independent if and only if the corresponding column vectors of B are linearly independent.

(*b*) A given set of column vectors of A forms a basis for the column space of A if and only if the corresponding column vectors of B form a basis for the column space of B.

Theorem 4.9.1

The row space and the column space of a matrix A have the same dimension.

Definition 1

The common dimension of the row space and column space of a matrix A is called the ***rank*** of A and is denoted by rank(A); the dimension of the null space of A is called the ***nullity*** of A and is denoted by nullity(A).

Find the rank and nullity of the matrix

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

Solution The reduced row echelon form of A is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(verify). Since this matrix has two leading 1's, its row and column spaces are two-dimensional and rank$(A) = 2$. To find the nullity of A, we must find the dimension of the solution space of the linear system $A\mathbf{x} = \mathbf{0}$. This system can be solved by reducing its augmented matrix to reduced row echelon form. The resulting matrix will be identical to (1), except that it will have an additional last column of zeros, and hence the corresponding system of equations will be

$$\begin{aligned} x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 &= 0 \\ x_2 - 2x_3 - 12x_4 - 16x_5 + 5x_6 &= 0 \end{aligned}$$

from which we obtain the general solution

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

or in column vector form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{3}$$

Because the four vectors on the right side of Formula (3) form a basis for the solution space it follows that nullity$(A) = 4$.

EXAMPLE 2 | Maximum Value for Rank

What is the maximum possible rank of an $m \times n$ matrix A that is not square?

Solution Since the row vectors of A lie in R^n and the column vectors in R^m, the row space of A is at most n-dimensional and the column space is at most m-dimensional. Since the rank of A is the common dimension of its row and column space, it follows that the rank is at most the smaller of m and n. We denote this by writing

$$\text{rank}(A) \leq \min(m, n)$$

in which $\min(m, n)$ is the minimum of m and n.

Theorem 4.9.2

Dimension Theorem for Matrices

If A is a matrix with n columns, then

$$\text{rank}(A) + \text{nullity}(A) = n \tag{4}$$

EXAMPLE 3 | The Sum of Rank and Nullity

The matrix

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

has 6 columns, so

$$\text{rank}(A) + \text{nullity}(A) = 6$$

This is consistent with Example 1, where we showed that

$$\text{rank}(A) = 2 \quad \text{and} \quad \text{nullity}(A) = 4$$

Theorem 4.9.3

If A is an $m \times n$ matrix, then

(a) $\operatorname{rank}(A)$ = the number of leading variables in the general solution of $A\mathbf{x} = \mathbf{0}$.

(b) $\operatorname{nullity}(A)$ = the number of parameters in the general solution of $A\mathbf{x} = \mathbf{0}$.

EXAMPLE 4 | Rank, Nullity, and Linear Systems

(a) Find the number of parameters in the general solution of $A\mathbf{x} = \mathbf{0}$ if A is a 5×7 matrix of rank 3.

(b) Find the rank of a 5×7 matrix A for which $A\mathbf{x} = \mathbf{0}$ has a two-dimensional solution space.

Solution (a) From (4),

$$\operatorname{nullity}(A) = n - \operatorname{rank}(A) = 7 - 3 = 4$$

Thus, there are four parameters.

Solution (b) The matrix A has nullity 2, so

$$\operatorname{rank}(A) = n - \operatorname{nullity}(A) = 7 - 2 = 5$$

Theorem 4.9.4

If $A\mathbf{x} = \mathbf{b}$ is a consistent linear system of m equations in n unknowns, and if A has rank r, then the general solution of the system contains $n - r$ parameters.

Theorem 4.9.5

If A is any matrix, then $\text{rank}(A) = \text{rank}(A^T)$.

EXAMPLE 5 | Bases for the Fundamental Spaces

In Example 1 we found a basis for the null space of the 4×6 matrix

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

so in this example we will focus on finding bases for the remaining three fundamental spaces starting with the matrix

$$\underbrace{\left[\begin{array}{rrrrrr|rrrr} -1 & 2 & 0 & 4 & 5 & -3 & 1 & 0 & 0 & 0 \\ 3 & -7 & 2 & 0 & 1 & 4 & 0 & 1 & 0 & 0 \\ 2 & -5 & 2 & 4 & 6 & 1 & 0 & 0 & 1 & 0 \\ 4 & -9 & 2 & -4 & -4 & 7 & 0 & 0 & 0 & 1 \end{array}\right]}_{A \qquad\qquad\qquad I}$$

in which a 4×4 identity matrix has been adjoined to A. Using Gaussian elimination to reduce the left side to reduced row echelon form R yields (verify)

$$\underbrace{\left[\begin{array}{rrrrrr|rrrr} 1 & 0 & -4 & -28 & -37 & 13 & 0 & 0 & -\frac{9}{2} & \frac{5}{2} \\ 0 & 1 & -2 & -12 & -16 & 5 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \end{array}\right]}_{R \qquad\qquad\qquad E}$$

From R we see that A has rank $r = 2$ (two nonzero rows), has nullity $n - r = 6 - 2 = 4$, and from (7) has left nullity $m - r = 2$. The two pivot rows of R (rows 1 and 2) form a basis for the row space of A, the two pivot columns of A (columns 1 and 2) form a basis for the column space of A, and the bottom two rows of E form a basis for the left null space of A. Expressing these bases in column form we have:

$$\text{row space basis: } \left\{ \begin{bmatrix} -1 \\ 0 \\ -4 \\ -28 \\ -37 \\ 13 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \\ -12 \\ -16 \\ 5 \end{bmatrix} \right\}, \quad \text{column space basis: } \left\{ \begin{bmatrix} -1 \\ 3 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -7 \\ -5 \\ -9 \end{bmatrix} \right\}$$

$$\text{left null space basis: } \left\{ \begin{bmatrix} 1 \\ 0 \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \right\}$$

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Linear Algebra (MT-1004)

Lecture # 8

Eigen Values & Eigen Vectors:

Some Applications of Eigen Values & Eigen Vectors:

- **Eigen Values are used to reduce noise in data.**
- **Eigen Values & Eigen Vectors lives in the heart of data science field**
- **It is must-know topic for anyone who wants to understand machine learning in depth**

Historical Note

Methods of linear algebra are used in the emerging field of computerized face recognition. Researchers are working with the idea that every human face in a racial group is a combination of a few dozen primary shapes. For example, by analyzing three-dimensional scans of many faces, researchers at Rockefeller University have produced both an average head shape in the Caucasian group—dubbed the ***meanhead*** (top row left in the figure to the left)—and a set of standardized variations from that shape, called ***eigenheads*** (15 of which are shown in the picture). These are so named because they are eigenvectors of a certain matrix that stores digitized facial information. Face shapes are represented mathematically as linear combinations of the eigenheads.

[Image: © Dr. Joseph J. Atick, adapted from *Scientific American*]

Definition 1

If A is an $n \times n$ matrix, then a nonzero vector $\mathbf{x}$ in R^n is called an ***eigenvector*** of A (or of the matrix operator T_A) if $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$; that is,

$$A\mathbf{x} = \lambda\mathbf{x}$$

for some scalar λ. The scalar λ is called an ***eigenvalue*** of A (or of T_A), and $\mathbf{x}$ is said to be an ***eigenvector corresponding to λ***.

The requirement that an eigenvector be nonzero is imposed to avoid the unimportant case $A\mathbf{0} = \lambda\mathbf{0}$, which holds for every A and λ.

Theorem 5.1.1

If A is an $n \times n$ matrix, then λ is an eigenvalue of A if and only if it satisfies the equation

$$\det(\lambda I - A) = 0 \tag{1}$$

This is called the ***characteristic equation*** of A.

When the determinant $\det(\lambda I - A)$ in (1) is expanded, the characteristic equation of A takes the form

$$\boxed{\lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0} \tag{3}$$

where the left side of this equation is a polynomial of degree n in which the coefficient of λ^n is 1 (Exercise 37). The polynomial

$$\boxed{p(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_n} \tag{4}$$

is called the ***characteristic polynomial*** of A. For example, it follows from (2) that the characteristic polynomial of the 2×2 matrix in Example 2 is

$$p(\lambda) = (\lambda - 3)(\lambda + 1) = \lambda^2 - 2\lambda - 3$$

which is a polynomial of degree 2.

Theorem 5.1.2

If A is an $n \times n$ triangular matrix (upper triangular, lower triangular, or diagonal), then the eigenvalues of A are the entries on the main diagonal of A.

By inspection, the eigenvalues of the lower triangular matrix

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

By inspection, the eigenvalues of the lower triangular matrix

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

are $\lambda = \frac{1}{2}$, $\lambda = \frac{2}{3}$, and $\lambda = -\frac{1}{4}$.

Theorem 5.1.3

If A is an $n \times n$ matrix, the following statements are equivalent.

(a) λ is an eigenvalue of A.

(b) λ is a solution of the characteristic equation $\det(\lambda I - A) = 0$.

(c) The system of equations $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has nontrivial solutions.

(d) There is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$.

Finding Eigenvectors and Bases for Eigenspaces

Now that we know how to find the eigenvalues of a matrix, we will consider the problem of finding the corresponding eigenvectors. By definition, the eigenvectors of A corresponding to an eigenvalue λ are the nonzero vectors that satisfy

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

Thus, we can find the eigenvectors of A corresponding to λ by finding the nonzero vectors in the solution space of this linear system. This solution space, which is called the ***eigenspace*** of A corresponding to λ, can also be viewed as:

1. the null space of the matrix $\lambda I - A$
2. the kernel of the matrix operator $T_{\lambda I - A}: R^n \to R^n$
3. the set of vectors for which $A\mathbf{x} = \lambda\mathbf{x}$

Eigenvalues and Invertibility

The next theorem establishes a relationship between the eigenvalues and the invertibility of a matrix.

Theorem 5.1.4

A square matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.

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

Find Eigen Values.

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

Solution It follows from Formula (1) that the eigenvalues of A are the solutions of the equation $\det(\lambda I - A) = 0$, which we can write as

$$\begin{vmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{vmatrix} = 0$$

from which we obtain

$$(\lambda - 3)(\lambda + 1) = 0 \tag{2}$$

This shows that the eigenvalues of A are $\lambda = 3$ and $\lambda = -1$. Thus, in addition to the eigenvalue $\lambda = 3$ noted in Example 1, we have discovered a second eigenvalue $\lambda = -1$.

Find the eigenvalues of

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$$

$$\lambda = 4, \quad \lambda = 2 + \sqrt{3}, \quad \text{and} \quad \lambda = 2 - \sqrt{3}$$

Find bases for the eigenspaces of

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

Find bases for the eigenspaces of

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

Solution The characteristic equation of A is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, or in factored form, $(\lambda - 1)(\lambda - 2)^2 = 0$ (verify). Thus, the distinct eigenvalues of A are $\lambda = 1$ and $\lambda = 2$, so there are two eigenspaces of A.

By definition,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

is an eigenvector of A corresponding to λ if and only if $\mathbf{x}$ is a nontrivial solution of $(\lambda I - A)\mathbf{x} = \mathbf{0}$, or in matrix form,

$$\begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{6}$$

In the case where $\lambda = 2$, Formula (6) becomes

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

Solving this system using Gaussian elimination yields (verify)

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

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

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

Since

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

If $\lambda = 1$, then (6) becomes

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

Solving this system yields (verify)

$$x_1 = -2s, \quad x_2 = s, \quad x_3 = s$$

Thus, the eigenvectors corresponding to $\lambda = 1$ are the nonzero vectors of the form

$$\begin{bmatrix} -2s \\ s \\ s \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \quad \text{so that} \quad \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

is a basis for the eigenspace corresponding to $\lambda = 1$.

More on the Equivalence Theorem

As our final result in this section, we will use Theorem 5.1.4 to add one additional part to Theorem 4.9.8.

Theorem 5.1.5

Equivalent Statements

If A is an $n \times n$ matrix in which there are no duplicate rows and no duplicate columns, then the following statements are equivalent.

(a) A is invertible.
(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
(c) The reduced row echelon form of A is I_n.
(d) A is expressible as a product of elementary matrices.
(e) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.
(f) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.
(g) $\det(A) \neq 0$.
(h) The column vectors of A are linearly independent.
(i) The row vectors of A are linearly independent.
(j) The column vectors of A span R^n.
(k) The row vectors of A span R^n.
(l) The column vectors of A form a basis for R^n.
(m) The row vectors of A form a basis for R^n.
(n) A has rank n.
(o) A has nullity 0.
(p) The orthogonal complement of the null space of A is R^n.
(q) The orthogonal complement of the row space of A is $\{\mathbf{0}\}$.
(r) $\lambda = 0$ is not an eigenvalue of A.

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Do Question # 1-16 from Ex # 5.1

Diagonalization:

Definition 2

A square matrix A is said to be ***diagonalizable*** if it is similar to some diagonal matrix; that is, if there exists an invertible matrix P such that $P^{-1}AP$ is diagonal. In this case the matrix P is said to ***diagonalize*** A.

TABLE 1 Similarity Invariants

Property	Description
Determinant	A and $P^{-1}AP$ have the same determinant.
Invertibility	A is invertible if and only if $P^{-1}AP$ is invertible.
Rank	A and $P^{-1}AP$ have the same rank.
Nullity	A and $P^{-1}AP$ have the same nullity.
Trace	A and $P^{-1}AP$ have the same trace.
Characteristic polynomial	A and $P^{-1}AP$ have the same characteristic polynomial.
Eigenvalues	A and $P^{-1}AP$ have the same eigenvalues.
Eigenspace dimension	If λ is an eigenvalue of A (and hence of $P^{-1}AP$) then the eigenspace of A corresponding to λ and the eigenspace of $P^{-1}AP$ corresponding to λ have the same dimension.

Related Theorems:

Theorem 5.2.1

If A is an $n \times n$ matrix, the following statements are equivalent.

(*a*) A is diagonalizable.

(*b*) A has n linearly independent eigenvectors.

Theorem 5.2.2

(*a*) If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of a matrix A, and if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are corresponding eigenvectors, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

(*b*) An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Theorem 5.2.3

If k is a positive integer, λ is an eigenvalue of a matrix A, and $\mathbf{x}$ is a corresponding eigenvector, then λ^k is an eigenvalue of A^k and $\mathbf{x}$ is a corresponding eigenvector.

Theorem 5.2.4

Geometric and Algebraic Multiplicity

If A is a square matrix, then:

(*a*) For every eigenvalue of A, the geometric multiplicity is less than or equal to the algebraic multiplicity.

(*b*) A is diagonalizable if and only if its characteristic polynomial can be expressed as a product of linear factors, and the geometric multiplicity of every eigenvalue is equal to the algebraic multiplicity.

A Procedure for Diagonalizing an $n \times n$ Matrix

Step 1. Determine first whether the matrix is actually diagonalizable by searching for n linearly independent eigenvectors. One way to do this is to find a basis for each eigenspace and count the total number of vectors obtained. If there is a total of n vectors, then the matrix is diagonalizable, and if the total is less than n, then it is not.

Step 2. If you ascertained that the matrix is diagonalizable, then form the matrix $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]$ whose column vectors are the n basis vectors you obtained in Step 1.

Step 3. $P^{-1}AP$ will be a diagonal matrix whose successive diagonal entries are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ that correspond to the successive columns of P.

Ex 7.10: (Diagonalizing a matrix)

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}$$

Find a matrix P such that $P^{-1}AP$ is diagonal.

- Ex 7.10: (Diagonalizing a matrix)

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}$$

Find a matrix P such that $P^{-1}AP$ is diagonal.

Sol: Characteristic equation:

$$|\lambda \mathrm{I} - A| = \begin{vmatrix} \lambda - 1 & 1 & 1 \\ -1 & \lambda - 3 & -1 \\ 3 & -1 & \lambda + 1 \end{vmatrix} = (\lambda - 2)(\lambda + 2)(\lambda - 3) = 0$$

The eigenvalues: $\lambda_1 = 2, \lambda_2 = -2, \lambda_3 = 3$

$$\lambda_1 = 2$$

$$\Rightarrow \lambda_1 \mathrm{I} - A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 3 & -1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -t \\ 0 \\ t \end{bmatrix} \Rightarrow \text{eigenvector } p_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda_2 = -2$$

$$\Rightarrow \lambda_2 \mathrm{I} - A = \begin{bmatrix} -3 & 1 & 1 \\ -1 & -5 & -1 \\ 3 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{4}t \\ -\frac{1}{4}t \\ t \end{bmatrix} \Rightarrow \text{eigenvector } p_2 = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$$

$\lambda_3 = 3$

$$\Rightarrow \lambda_3 \mathrm{I} - A = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 3 & -1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -t \\ t \\ t \end{bmatrix} \Rightarrow \text{eigenvector } p_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

$$P = [p_1 \quad p_2 \quad p_3] = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 1 \\ 1 & 4 & 1 \end{bmatrix},$$

$$\text{s.t.} P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Computing Powers of a Matrix

If A is an nxn matrix and P is an invertible matrix, then

$$(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}AIAP = P^{-1}A^2P$$

More generally, for any positive integer *k*, $(P^{-1}AP)^k = P^{-1}A^kP$

It follows from this equation that if *A is diagonalizable, and $P^{-1}AP$ =D is a diagonal matrix, then* $P^{-1}A^kP = (P^{-1}AP)^k = D^k$

Solving this equation for A^k yields $A^k = PD^kP^{-1}$

This last equation expresses the kth power of A in terms of the kth power of the diagonal matrix D. But D^k is easy to compute, for if

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}, \quad \text{then} \quad D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix}$$

Do Question # 1-20 from Ex # 5.2

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Linear Algebra (MT-1004)

Lecture # 09 & 10

INNER PRODUCT SPACE

Definition 1

An ***inner product*** on a real vector space V is a function that associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors in V in such a way that the following axioms are satisfied for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and all scalars k.

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ **[Symmetry axiom]**
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ **[Additivity axiom]**
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$ **[Homogeneity axiom]**
4. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$ **[Positivity axiom]**

A real vector space with an inner product is called a ***real inner product space***.

- Note:

$$\mathbf{u} \cdot \mathbf{v} = \text{dot product (Euclidean inner product for } R^n\text{)}$$

$$< \mathbf{u}\ , \mathbf{v} >= \text{general inner product for vector space } V$$

Vector space:	$(V,\ +,\ \bullet)$
Inner product space:	$(V,\ +,\ \bullet,\ <,>)$

Weighted Euclidean Inner Product

The Euclidean inner product is the most important inner product on R^n. However, there are various applications in which it is desirable to modify the Euclidean inner product by *weighting its terms differently. More precisely, if*

$$w_1, w_2, \ldots, w_n$$

are *positive real numbers, which we shall call* ***weights, and if*** $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in R^n, then it can be shown that the formula

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots w_n u_n v_n$$

defines an inner product on R^n; it is called the ***weighted Euclidean inner product with weights*** $w_1, w_2, \ldots, w_n$

Definition 2

If V is a real inner product space, then the ***norm*** (or ***length***) of a vector $\mathbf{v}$ in V is denoted by $\|\mathbf{v}\|$ and is defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

and the ***distance*** between two vectors is denoted by $d(\mathbf{u}, \mathbf{v})$ and is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$$

A vector of norm 1 is called a ***unit vector***.

Theorem 6.1.1

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real inner product space V, and if k is a scalar, then:

(*a*) $\|\mathbf{v}\| \geq 0$ with equality if and only if $\mathbf{v} = \mathbf{0}$.

(*b*) $\|k\mathbf{v}\| = |k|\|\mathbf{v}\|$.

(*c*) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.

(*d*) $d(\mathbf{u}, \mathbf{v}) \geq 0$ with equality if and only if $\mathbf{u} = \mathbf{v}$.

EXAMPLE 1 | Weighted Euclidean Inner Product

Let $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ be vectors in R^2. Verify that the weighted Euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2 \tag{3}$$

satisfies the four inner product axioms.

Solution

Axiom 1: Interchanging $\mathbf{u}$ and $\mathbf{v}$ in Formula (3) does not change the sum on the right side, so $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.

Axiom 2: If $\mathbf{w} = (w_1, w_2)$, then

$$\begin{aligned} \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle &= 3(u_1 + v_1)w_1 + 2(u_2 + v_2)w_2 \\ &= 3(u_1w_1 + v_1w_1) + 2(u_2w_2 + v_2w_2) \\ &= (3u_1w_1 + 2u_2w_2) + (3v_1w_1 + 2v_2w_2) \\ &= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \end{aligned}$$

Axiom 3:
$$\begin{aligned} \langle k\mathbf{u}, \mathbf{v} \rangle &= 3(ku_1)v_1 + 2(ku_2)v_2 \\ &= k(3u_1v_1 + 2u_2v_2) \\ &= k\langle \mathbf{u}, \mathbf{v} \rangle \end{aligned}$$

Axiom 4: Observe that $\langle \mathbf{v}, \mathbf{v} \rangle = 3(v_1v_1) + 2(v_2v_2) = 3v_1^2 + 2v_2^2 \geq 0$ with equality if and only if $v_1 = v_2 = 0$, that is, if and only if $\mathbf{v} = \mathbf{0}$.

- Ex 6.3: (A function that is not an inner product)

Show that the following function is not an inner product on R^3.

$$\langle \mathbf{u} \cdot \mathbf{v} \rangle = u_1 v_1 - 2u_2 v_2 + u_3 v_3$$

Sol:

Let $\mathbf{v} = (1, 2, 1)$

Then $< \mathbf{v}, \mathbf{v} >= (1)(1) - 2(2)(2) + (1)(1) = -6 < 0$

Axiom 4 is not satisfied.

Thus this function is not an inner product on R^3.

EXAMPLE 2 | Calculating with a Weighted Euclidean Inner Product

It is important to keep in mind that norm and distance depend on the inner product being used. If the inner product is changed, then the norms and distances between vectors also change. For example, for the vectors $\mathbf{u} = (1, 0)$ and $\mathbf{v} = (0, 1)$ in R^2 with the Euclidean inner product we have

$$\|\mathbf{u}\| = \sqrt{1^2 + 0^2} = 1$$

and

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \|(1, -1)\| = \sqrt{1^2 + (-1)^2} = \sqrt{2}$$

but if we change to the weighted Euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2$$

we have

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = [3(1)(1) + 2(0)(0)]^{1/2} = \sqrt{3}$$

and

$$\begin{aligned} d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| &= \langle (1, -1), (1, -1) \rangle^{1/2} \\ &= [3(1)(1) + 2(-1)(-1)]^{1/2} = \sqrt{5} \end{aligned}$$

Unit Circles and Spheres in Inner Product Spaces

Definition 3

If V is an inner product space, then the set of points in V that satisfy

$$\|\mathbf{u}\| = 1$$

is called the ***unit sphere*** in V (or the ***unit circle*** in the case where $V = R^2$).

Theorem 6.1.2

If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in a real inner product space V, and if k is a scalar, then:

(a) $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$

(b) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$

(c) $\langle \mathbf{u}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{w} \rangle$

(d) $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle$

(e) $k\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, k\mathbf{v} \rangle$

POLYNOMIALS

If

$$\mathbf{p} = a_0 + a_1x + \cdots + a_nx^n \quad \text{and} \quad \mathbf{q} = b_0 + b_1x + \cdots + b_nx^n$$

are polynomials in P_n, then the following formula defines an inner product on P_n (verify) that we will call the ***standard inner product*** on this space:

$$\langle \mathbf{p}, \mathbf{q} \rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n \tag{9}$$

The norm of a polynomial $\mathbf{p}$ relative to this inner product is

$$\|\mathbf{p}\| = \sqrt{\langle \mathbf{p}, \mathbf{p} \rangle} = \sqrt{a_0^2 + a_1^2 + \cdots + a_n^2}$$

- Ex 6.4: (Finding inner product)

$\langle p, q\rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n$ is an inner product

Let $p(x) = 1 - 2x^2$, $q(x) = 4 - 2x + x^2$ be polynomials in $P_2(x)$

$(a)\ \langle p, q\rangle = ? \quad (b)\ \|q\| = ? \quad (c)\ d(p, q) = ?$

Sol:

$$(a)\quad \langle p, q\rangle = (1)(4) + (0)(-2) + (-2)(1) = 2$$

$$(b)\quad \|q\| = \sqrt{\langle q, q\rangle} = \sqrt{4^2 + (-2)^2 + 1^2} = \sqrt{21}$$

$$(c)\quad \because p - q = -3 + 2x - 3x^2$$

$$\therefore d(p, q) = \|p - q\| = \sqrt{\langle p - q, p - q\rangle}$$

$$= \sqrt{(-3)^2 + 2^2 + (-3)^2} = \sqrt{22}$$

EXAMPLE 8 | The Evaluation Inner Product on P_n

If

$$\mathbf{p} = p(x) = a_0 + a_1x + \cdots + a_nx^n \quad \text{and} \quad \mathbf{q} = q(x) = b_0 + b_1x + \cdots + b_nx^n$$

are polynomials in P_n, and if $x_0, x_1, \ldots, x_n$ are distinct real numbers (called ***sample points***), then the formula

$$\langle \mathbf{p}, \mathbf{q} \rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n) \tag{10}$$

defines an inner product on P_n called the ***evaluation inner product*** at $x_0, x_1, \ldots, x_n$. Algebraically, this can be viewed as the dot product in R^n of the n-tuples

$$(p(x_0), p(x_1), \ldots, p(x_n)) \quad \text{and} \quad (q(x_0), q(x_1), \ldots, q(x_n))$$

and hence the first three inner product axioms follow from properties of the dot product. The fourth inner product axiom follows from the fact that

$$\langle \mathbf{p}, \mathbf{p} \rangle = [p(x_0)]^2 + [p(x_1)]^2 + \cdots + [p(x_n)]^2 \geq 0$$

with equality holding if and only if

$$p(x_0) = p(x_1) = \cdots = p(x_n) = 0$$

But a nonzero polynomial of degree n or less can have at most n distinct roots, so it must be that $\mathbf{p} = \mathbf{0}$, which proves that the fourth inner product axiom holds.

The norm of a polynomial $\mathbf{p}$ relative to the evaluation inner product is

$$\|\mathbf{p}\| = \sqrt{\langle \mathbf{p}, \mathbf{p} \rangle} = \sqrt{[p(x_0)]^2 + [p(x_1)]^2 + \cdots + [p(x_n)]^2} \tag{11}$$

EXAMPLE 9 | Working with the Evaluation Inner Product

Let P_2 have the evaluation inner product at the points

$$x_0 = -2, \quad x_1 = 0, \quad \text{and} \quad x_2 = 2$$

Compute $\langle \mathbf{p}, \mathbf{q} \rangle$ and $\|\mathbf{p}\|$ for the polynomials $\mathbf{p} = p(x) = x^2$ and $\mathbf{q} = q(x) = 1 + x$.

Solution It follows from (10) and (11) that

$$\langle \mathbf{p}, \mathbf{q} \rangle = p(-2)q(-2) + p(0)q(0) + p(2)q(2) = (4)(-1) + (0)(1) + (4)(3) = 8$$

$$\|\mathbf{p}\| = \sqrt{[p(x_0)]^2 + [p(x_1)]^2 + [p(x_2)]^2} = \sqrt{[p(-2)]^2 + [p(0)]^2 + [p(2)]^2}$$

$$= \sqrt{4^2 + 0^2 + 4^2} = \sqrt{32} = 4\sqrt{2}$$

Inner Products Generated by Matrices

The Euclidean inner product and the weighted Euclidean inner products are special cases of a general class of inner products on R^n called ***matrix inner products***. To define this class of inner products, let $\mathbf{u}$ and $\mathbf{v}$ be vectors in R^n that are expressed in *column form*, and let A be an *invertible* $n \times n$ matrix. It can be shown (Exercise 47) that if $\mathbf{u} \cdot \mathbf{v}$ is the Euclidean inner product on R^n, then the formula

$$\boxed{\langle \mathbf{u}, \mathbf{v} \rangle = A\mathbf{u} \cdot A\mathbf{v}} \tag{5}$$

EXAMPLE 4 | Matrices Generating Weighted Euclidean Inner Products

The standard Euclidean and weighted Euclidean inner products are special cases of matrix inner products. The standard Euclidean inner product on R^n is generated by the $n \times n$ identity matrix, since setting $A = I$ in Formula (5) yields

$$\langle \mathbf{u}, \mathbf{v} \rangle = I\mathbf{u} \cdot I\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$$

and the weighted Euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots + w_n u_n v_n \tag{7}$$

is generated by the matrix

$$A = \begin{bmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{w_n} \end{bmatrix}$$

This can be seen by observing that $A^T A$ is the $n \times n$ diagonal matrix whose diagonal entries are the weights $w_1, w_2, \ldots, w_n$.

EXAMPLE 6 | The Standard Inner Product on M_{nn}

If $\mathbf{u} = U$ and $\mathbf{v} = V$ are matrices in the vector space M_{nn}, then the formula

$$\boxed{\langle \mathbf{u}, \mathbf{v} \rangle = \operatorname{tr}(U^T V)} \tag{8}$$

defines an inner product on M_{nn} called the ***standard inner product*** on that space (see Definition 8 of Section 1.3 for a definition of trace). This can be proved by confirming that the four inner product space axioms are satisfied, but we can illustrate the idea by computing (8) for the 2×2 matrices

$$U = \begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix}$$

This yields

$$\langle \mathbf{u}, \mathbf{v} \rangle = \operatorname{tr}(U^T V) = u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4$$

which is just the dot product of the corresponding entries in the two matrices. And it follows from this that

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\operatorname{tr}\langle U^T U \rangle} = \sqrt{u_1^2 + u_2^2 + u_3^2 + u_4^2}$$

For example, if

$$\mathbf{u} = U = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = V = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$$

then

$$\langle \mathbf{u}, \mathbf{v} \rangle = \operatorname{tr}(U^T V) = 1(-1) + 2(0) + 3(3) + 4(2) = 16$$

and

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\operatorname{tr}(U^T U)} = \sqrt{1^2 + 2^2 + 3^2 + 4^2} = \sqrt{30}$$

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\operatorname{tr}(V^T V)} = \sqrt{(-1)^2 + 0^2 + 3^2 + 2^2} = \sqrt{14}$$

Definition 1

A set of two or more vectors in a real inner product space is said to be ***orthogonal*** if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is said to be ***orthonormal***.

EXAMPLE 1 | An Orthogonal Set in R^3

Let

$$\mathbf{v}_1 = (0, 1, 0), \quad \mathbf{v}_2 = (1, 0, 1), \quad \mathbf{v}_3 = (1, 0, -1)$$

and assume that R^3 has the Euclidean inner product. It follows that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set since $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$.

The above set is orthonormal??

NORMALIZING:

The Euclidean norms of the vectors in Example 1 are

$$\|\mathbf{v}_1\| = 1, \quad \|\mathbf{v}_2\| = \sqrt{2}, \quad \|\mathbf{v}_3\| = \sqrt{2}$$

Consequently, normalizing $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ yields

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = (0, 1, 0), \quad \mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right),$$

$$\mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

We leave it for you to verify that the set $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthonormal by showing that

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{u}_3 \rangle = \langle \mathbf{u}_2, \mathbf{u}_3 \rangle = 0 \quad \text{and} \quad \|\mathbf{u}_1\| = \|\mathbf{u}_2\| = \|\mathbf{u}_3\| = 1$$

Theorem 6.3.1

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space, then S is linearly independent.

Since an orthonormal set is orthogonal, and since its vectors are nonzero (norm 1), it follows from Theorem 6.3.1 that every *orthonormal* set is linearly independent.

In an inner product space, a basis consisting of orthonormal vectors is called an ***orthonormal basis***, and a basis of orthogonal vectors is called an ***orthogonal basis***. A familiar example of an orthonormal basis is the standard basis for R^n with the Euclidean inner product:

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

EXAMPLE 4 | An Orthonormal Basis

In Example 2 we showed that the vectors

$$\mathbf{u}_1 = (0, 1, 0), \quad \mathbf{u}_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \quad \text{and} \quad \mathbf{u}_3 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

form an orthonormal set with respect to the Euclidean inner product on R^3. By Theorem 6.3.1, these vectors form a linearly independent set, and since R^3 is three-dimensional, it follows from Theorem 4.6.4 that $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for R^3.

Coordinates Relative to Orthonormal Bases

One way to express a vector **u** as a linear combination of basis vectors

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

is to convert the vector equation

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

to a linear system and solve for the coefficients $c_1, c_2, \dots, c_n$. However, if the basis happens to be orthogonal or orthonormal, then the following theorem shows that the coefficients can be obtained more simply by computing appropriate inner products.

Theorem 6.3.2

(*a*) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthogonal basis for an inner product space V, and if **u** is any vector in V, then

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n \tag{3}$$

(*b*) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthonormal basis for an inner product space V, and if **u** is any vector in V, then

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n \tag{4}$$

Using the terminology and notation from Definition 2 of Section 4.5, it follows from Theorem 6.3.2 that the coordinate vector of a vector **u** in V relative to an orthogonal basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is

$$\boxed{(\mathbf{u})_S = \left(\frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}, \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}, \dots, \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\right)} \tag{6}$$

and relative to an orthonormal basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is

$$\boxed{(\mathbf{u})_S = \left(\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \dots, \langle \mathbf{u}, \mathbf{v}_n \rangle\right)} \tag{7}$$

EXAMPLE 5 | A Coordinate Vector Relative to an Orthonormal Basis

Let

$$\mathbf{v}_1 = (0, 1, 0), \quad \mathbf{v}_2 = \left(-\tfrac{4}{5}, 0, \tfrac{3}{5}\right), \quad \mathbf{v}_3 = \left(\tfrac{3}{5}, 0, \tfrac{4}{5}\right)$$

It is easy to check that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis for R^3 with the Euclidean inner product. Express the vector $\mathbf{u} = (1, 1, 1)$ as a linear combination of the vectors in S, and find the coordinate vector $(\mathbf{u})_S$.

Solution We leave it for you to verify that

$$\langle \mathbf{u}, \mathbf{v}_1 \rangle = 1, \quad \langle \mathbf{u}, \mathbf{v}_2 \rangle = -\tfrac{1}{5}, \quad \text{and} \quad \langle \mathbf{u}, \mathbf{v}_3 \rangle = \tfrac{7}{5}$$

Therefore, by Theorem 6.3.2 we have

$$\mathbf{u} = \mathbf{v}_1 - \tfrac{1}{5}\mathbf{v}_2 + \tfrac{7}{5}\mathbf{v}_3$$

that is,

$$(1, 1, 1) = (0, 1, 0) - \tfrac{1}{5}\left(-\tfrac{4}{5}, 0, \tfrac{3}{5}\right) + \tfrac{7}{5}\left(\tfrac{3}{5}, 0, \tfrac{4}{5}\right)$$

Thus, the coordinate vector of $\mathbf{u}$ relative to S is

$$(\mathbf{u})_S = (\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \langle \mathbf{u}, \mathbf{v}_3 \rangle) = \left(1, -\tfrac{1}{5}, \tfrac{7}{5}\right)$$

EXAMPLE 6 | An Orthonormal Basis from an Orthogonal Basis

(*a*) Show that the vectors

$$\mathbf{w}_1 = (0, 2, 0), \quad \mathbf{w}_2 = (3, 0, 3), \quad \mathbf{w}_3 = (-4, 0, 4)$$

form an orthogonal basis for R^3 with the Euclidean inner product, and use that basis to find an orthonormal basis by normalizing each vector.

Solution (*a*) The given vectors form an orthogonal set since

$$\langle \mathbf{w}_1, \mathbf{w}_2 \rangle = 0, \quad \langle \mathbf{w}_1, \mathbf{w}_3 \rangle = 0, \quad \langle \mathbf{w}_2, \mathbf{w}_3 \rangle = 0$$

It follows from Theorem 6.3.1 that these vectors are linearly independent and hence form a basis for R^3 by Theorem 4.6.4. We leave it for you to calculate the norms of $\mathbf{w}_1, \mathbf{w}_2$, and $\mathbf{w}_3$ and then obtain the orthonormal basis

$$\mathbf{v}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = (0, 1, 0), \quad \mathbf{v}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right),$$

$$\mathbf{v}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

(*b*) Express the vector $\mathbf{u} = (1, 2, 4)$ as a linear combination of the orthonormal basis vectors obtained in part (*a*).

Solution (*b*) It follows from Formula (4) that

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \langle \mathbf{u}, \mathbf{v}_3 \rangle \mathbf{v}_3$$

We leave it for you to confirm that

$$\langle \mathbf{u}, \mathbf{v}_1 \rangle = (1, 2, 4) \cdot (0, 1, 0) = 2$$

$$\langle \mathbf{u}, \mathbf{v}_2 \rangle = (1, 2, 4) \cdot \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) = \frac{5}{\sqrt{2}}$$

$$\langle \mathbf{u}, \mathbf{v}_3 \rangle = (1, 2, 4) \cdot \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) = \frac{3}{\sqrt{2}}$$

and hence that

$$(1, 2, 4) = 2(0, 1, 0) + \frac{5}{\sqrt{2}} \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) + \frac{3}{\sqrt{2}} \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)$$

Theorem 6.3.3

Projection Theorem

If W is a finite-dimensional subspace of an inner product space V, then every vector $\mathbf{u}$ in V can be expressed in exactly one way as

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \tag{8}$$

where $\mathbf{w}_1$ is in W and $\mathbf{w}_2$ is in $W^\perp$.

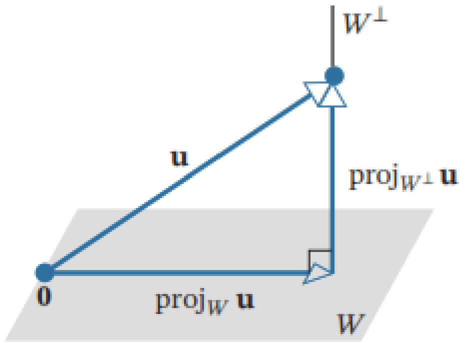
The vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in Formula (8) are commonly denoted by

$$\mathbf{w}_1 = \text{proj}_W \mathbf{u} \quad \text{and} \quad \mathbf{w}_2 = \text{proj}_{W^\perp} \mathbf{u} \tag{9}$$

These are called the ***orthogonal projection of* u *on W*** and the ***orthogonal projection of* u *on*** $W^\perp$, respectively. The vector $\mathbf{w}_2$ is also called the ***component of* u *orthogonal to W***. Using the notation in (9), Formula (8) can be expressed as

$$\boxed{\mathbf{u} = \text{proj}_W \mathbf{u} + \text{proj}_{W^\perp} \mathbf{u}} \tag{10}$$

$$\boxed{\mathbf{u} = \text{proj}_W \mathbf{u} + (\mathbf{u} - \text{proj}_W \mathbf{u})}$$



Theorem 6.3.4

Let W be a finite-dimensional subspace of an inner product space V.

(*a*) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthogonal basis for W, and $\mathbf{u}$ is any vector in V, then

$$\text{proj}_W \mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_r \rangle}{\|\mathbf{v}_r\|^2}\mathbf{v}_r \tag{12}$$

(*b*) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthonormal basis for W, and $\mathbf{u}$ is any vector in V, then

$$\text{proj}_W \mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_r \rangle \mathbf{v}_r \tag{13}$$

Although Formulas (12) and (13) are expressed in terms of orthogonal and orthonormal basis vectors, the resulting vector $\text{proj}_W \mathbf{u}$ does not depend on the basis vectors that are used.

EXAMPLE 7 | Calculating Projections

Let R^3 have the Euclidean inner product, and let W be the subspace spanned by the orthonormal vectors $\mathbf{v}_1 = (0, 1, 0)$ and $\mathbf{v}_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)$. From Formula (13) the orthogonal projection of $\mathbf{u} = (1, 1, 1)$ on W is

$$\begin{aligned}\operatorname{proj}_W \mathbf{u} &= \langle \mathbf{u}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle \mathbf{v}_2 \\ &= (1)(0, 1, 0) + \left(-\frac{1}{5}\right)\left(-\frac{4}{5}, 0, \frac{3}{5}\right) \\ &= \left(\frac{4}{25}, 1, -\frac{3}{25}\right)\end{aligned}$$

The component of $\mathbf{u}$ orthogonal to W is

$$\operatorname{proj}_{W^\perp} \mathbf{u} = \mathbf{u} - \operatorname{proj}_W \mathbf{u} = (1, 1, 1) - \left(\frac{4}{25}, 1, -\frac{3}{25}\right) = \left(\frac{21}{25}, 0, \frac{28}{25}\right)$$

Observe that $\operatorname{proj}_{W^\perp} \mathbf{u}$ is orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$, so this vector is orthogonal to each vector in the space W spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$, as it should be.

Linear Algebra (MT-1004)

Lecture # 36

The Gram–Schmidt Process

We have seen that orthonormal bases exhibit a variety of useful properties. Our next theorem, which is the main result in this section, shows that every nonzero finite-dimensional vector space has an orthonormal basis. The proof of this result is extremely important since it provides an algorithm, or method, for converting an arbitrary basis into an orthonormal basis.

Theorem 6.3.5

Every nonzero finite-dimensional inner product space has an orthonormal basis.

The Gram–Schmidt Process

To convert a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$, perform the following computations:

Step 1. $\mathbf{v}_1 = \mathbf{u}_1$

Step 2. $\mathbf{v}_2 = \mathbf{u}_2 - \dfrac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

Step 4. $\mathbf{v}_4 = \mathbf{u}_4 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3$

$\vdots$

(continue for r steps)

Optional Step. To convert the orthogonal basis into an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_r\}$, normalize the orthogonal basis vectors.

EXAMPLE 8 | Using the Gram–Schmidt Process

Assume that the vector space R^3 has the Euclidean inner product. Apply the Gram–Schmidt process to transform the basis vectors

$$\mathbf{u}_1 = (1, 1, 1), \quad \mathbf{u}_2 = (0, 1, 1), \quad \mathbf{u}_3 = (0, 0, 1)$$

into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, and then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$.

Solution

Step 1. $\mathbf{v}_1 = \mathbf{u}_1 = (1, 1, 1)$

Step 2. $\mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{W_1}\mathbf{u}_2 = \mathbf{u}_2 - \dfrac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

$$= (0, 1, 1) - \frac{2}{3}(1, 1, 1) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \text{proj}_{W_2}\mathbf{u}_3 = \mathbf{u}_3 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

$$= (0, 0, 1) - \frac{1}{3}(1, 1, 1) - \frac{1/3}{2/3}\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$$= \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

Thus,

$$\mathbf{v}_1 = (1, 1, 1), \quad \mathbf{v}_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right), \quad \mathbf{v}_3 = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

form an orthogonal basis for R^3. The norms of these vectors are

$$\|\mathbf{v}_1\| = \sqrt{3}, \quad \|\mathbf{v}_2\| = \frac{\sqrt{6}}{3}, \quad \|\mathbf{v}_3\| = \frac{1}{\sqrt{2}}$$

so an orthonormal basis for R^3 is

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \quad \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right),$$

$$\mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Theorem 6.3.7

QR-Decomposition

If A is an $m \times n$ matrix with linearly independent column vectors, then A can be factored as

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal column vectors, and R is an $n \times n$ invertible upper triangular matrix.

$$R = \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{q}_1 \rangle & \langle \mathbf{u}_2, \mathbf{q}_1 \rangle & \cdots & \langle \mathbf{u}_n, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{u}_2, \mathbf{q}_2 \rangle & \cdots & \langle \mathbf{u}_n, \mathbf{q}_2 \rangle \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \langle \mathbf{u}_n, \mathbf{q}_n \rangle \end{bmatrix}$$

EXAMPLE 10 | QR-Decomposition of a 3×3 Matrix

Find a QR-decomposition of

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

Solution The column vectors of A are

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Applying the Gram–Schmidt process with normalization to these column vectors yields the orthonormal vectors (see Example 8)

$$\mathbf{q}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Thus, it follows from Formula (16) that R is

$$R = \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{q}_1 \rangle & \langle \mathbf{u}_2, \mathbf{q}_1 \rangle & \langle \mathbf{u}_3, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{u}_2, \mathbf{q}_2 \rangle & \langle \mathbf{u}_3, \mathbf{q}_2 \rangle \\ 0 & 0 & \langle \mathbf{u}_3, \mathbf{q}_3 \rangle \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

from which it follows that a QR-decomposition of A is

$$\underset{A}{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}} = \underset{Q}{\begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}} \underset{R}{\begin{bmatrix} \frac{3}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}}$$

49. Find a QR-decomposition of the matrix

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

In partitioned form, $A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3]$. By inspection, $\mathbf{u}_3 = \mathbf{u}_1 + 2\mathbf{u}_2$, so the column vectors of A are not linearly independent and A does not have a QR-decomposition.

Linear Algebra (MT-1004)

Lecture # 37

Orthogonal Matrices

Definition 1

A square matrix A is said to be ***orthogonal*** if its transpose is the same as its inverse, that is, if

$$A^{-1} = A^T$$

or, equivalently, if

$$AA^T = A^TA = I \tag{1}$$

A matrix transformation $T_A : R^n \to R^n$ is said to be an ***orthogonal transformation*** or an ***orthogonal operator*** if A is an orthogonal matrix.

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

EXAMPLE 1 | A 3 × 3 Orthogonal Matrix

The matrix

$$A = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}$$

is orthogonal since

$$A^T A = \begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

EXAMPLE 2 | Rotation and Reflection Matrices Are Orthogonal

Recall from Table 5 of Section 1.8 that the standard matrix for the counterclockwise rotation about the origin of R^2 through an angle θ is

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

This matrix is orthogonal for all choices of θ since

$$A^T A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We leave it for you to verify that the reflection matrices in Tables 1 and 2 of Section 1.8 are all orthogonal.

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology FAST

Theorem 7.1.1

The following are equivalent for an $n \times n$ matrix A.

(a) A is orthogonal.

(b) The row vectors of A form an orthonormal set in R^n with the Euclidean inner product.

(c) The column vectors of A form an orthonormal set in R^n with the Euclidean inner product.

Warning Note that an orthogonal matrix has *orthonormal* rows and columns—not simply orthogonal rows and columns.

$$AA^T = \begin{bmatrix} \mathbf{r}_1\mathbf{c}_1^T & \mathbf{r}_1\mathbf{c}_2^T & \cdots & \mathbf{r}_1\mathbf{c}_n^T \\ \mathbf{r}_2\mathbf{c}_1^T & \mathbf{r}_2\mathbf{c}_2^T & \cdots & \mathbf{r}_2\mathbf{c}_n^T \\ \vdots & \vdots & & \vdots \\ \mathbf{r}_n\mathbf{c}_1^T & \mathbf{r}_n\mathbf{c}_2^T & \cdots & \mathbf{r}_n\mathbf{c}_n^T \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1\cdot\mathbf{r}_1 & \mathbf{r}_1\cdot\mathbf{r}_2 & \cdots & \mathbf{r}_1\cdot\mathbf{r}_n \\ \mathbf{r}_2\cdot\mathbf{r}_1 & \mathbf{r}_2\cdot\mathbf{r}_2 & \cdots & \mathbf{r}_2\cdot\mathbf{r}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{r}_n\cdot\mathbf{r}_1 & \mathbf{r}_n\cdot\mathbf{r}_2 & \cdots & \mathbf{r}_n\cdot r_n \end{bmatrix}$$

It is evident from this formula that $AA^T = I$ if and only if

$$\mathbf{r}_1\cdot\mathbf{r}_1 = \mathbf{r}_2\cdot\mathbf{r}_2 = \cdots = \mathbf{r}_n\cdot\mathbf{r}_n = 1$$

and

$$\mathbf{r}_i\cdot\mathbf{r}_j = 0 \quad \text{when } i \neq j$$

which are true if and only if $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ is an orthonormal set in R^n. ■

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Theorem 7.1.2

(*a*) The transpose of an orthogonal matrix is orthogonal.

(*b*) The inverse of an orthogonal matrix is orthogonal.

(*c*) A product of orthogonal matrices is orthogonal.

(*d*) If A is orthogonal, then $\det(A) = 1$ or $\det(A) = -1$.

Do Q.1 till 6 from Ex # 7.1

Orthogonal Diagonalization:

Definition 1

If A and B are square matrices, then we say that B is ***orthogonally similar*** to A if there is an orthogonal matrix P such that $B = P^T AP$.

If A is orthogonally similar to some diagonal matrix, say

$$P^T AP = D$$

then we say A is ***orthogonally diagonalizable*** and P ***orthogonally diagonalizes*** A.

Theorem 7.2.1

If A is an $n \times n$ matrix with real entries, then the following are equivalent.

(a) A is orthogonally diagonalizable.

(b) A has an orthonormal set of n eigenvectors.

(c) A is symmetric.

Properties of Symmetric Matrices

Our next goal is to devise a procedure for orthogonally diagonalizing a symmetric matrix, but before we can do so, we need the following critical theorem about eigenvalues and eigenvectors of symmetric matrices.

Theorem 7.2.2

If A is a symmetric matrix with real entries, then:

(a) The eigenvalues of A are all real numbers.

(b) Eigenvectors from different eigenspaces are orthogonal.

Orthogonally Diagonalizing an $n \times n$ Symmetric Matrix

Step 1. Find a basis for each eigenspace of A.

Step 2. Apply the Gram–Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.

Step 3. Form the matrix P whose columns are the vectors constructed in Step 2. This matrix will orthogonally diagonalize A, and the eigenvalues on the diagonal of $D = P^TAP$ will be in the same order as their corresponding eigenvectors in P.

EXAMPLE 1 | Orthogonally Diagonalizing a Symmetric Matrix

Find an orthogonal matrix P that diagonalizes

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

Solution We leave it for you to verify that the characteristic equation of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 4 & -2 & -2 \\ -2 & \lambda - 4 & -2 \\ -2 & -2 & \lambda - 4 \end{bmatrix} = (\lambda - 2)^2(\lambda - 8) = 0$$

Thus, the distinct eigenvalues of A are $\lambda = 2$ and $\lambda = 8$. By the method used in Example 7 of Section 5.1, it can be shown that

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \tag{5}$$

form a basis for the eigenspace corresponding to $\lambda = 2$. Applying the Gram–Schmidt process to $\{\mathbf{u}_1, \mathbf{u}_2\}$ yields the following orthonormal eigenvectors (verify):

$$\mathbf{v}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix} \tag{6}$$

The eigenspace corresponding to $\lambda = 8$ has

$$\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

as a basis. Applying the Gram–Schmidt process to $\{\mathbf{u}_3\}$ (i.e., normalizing $\mathbf{u}_3$) yields

$$\mathbf{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

Finally, using $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ as column vectors, we obtain

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

which orthogonally diagonalizes A. As a check, we leave it for you to confirm that

$$P^TAP = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Linear Algebra (MT-1004)

Lecture # 38

Spectral Decomposition

If A is a symmetric matrix with real entries that is orthogonally diagonalized by

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n]$$

$$A = PDP^T = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

$$= [\lambda_1\mathbf{u}_1 \quad \lambda_2\mathbf{u}_2 \quad \cdots \quad \lambda_n\mathbf{u}_n] \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

Multiplying out, we obtain the formula

$$\boxed{A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T} \tag{7}$$

which is called a ***spectral decomposition of A***.*

NOTE:

The terminology spectral decomposition is derived from the fact that the set of all eigenvalues of a matrix A is sometimes called the spectrum of A

EXAMPLE 2 | A Geometric Interpretation of a Spectral Decomposition

The matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$$

has eigenvalues $\lambda_1 = -3$ and $\lambda_2 = 2$ with corresponding eigenvectors

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(verify). Normalizing these basis vectors yields

$$\mathbf{u}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

so a spectral decomposition of A is

$$\begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T = (-3)\begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix} + (2)\begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

$$= (-3)\begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{4}{5} \end{bmatrix} + (2)\begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix} \tag{8}$$

where, as noted above, the 2×2 matrices on the right side of (8) are the standard matrices for the orthogonal projections onto the eigenspaces corresponding to the eigenvalues $\lambda_1 = -3$ and $\lambda_2 = 2$, respectively.

Now let us see what this spectral decomposition tells us about the image of the vector $\mathbf{x} = (1, 1)$ under multiplication by A. Writing $\mathbf{x}$ in column form, it follows that

$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \tag{9}$$

and from (8) that

$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = (-3)\begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{4}{5} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (2)\begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= (-3)\begin{bmatrix} -\frac{1}{5} \\ \frac{2}{5} \end{bmatrix} + (2)\begin{bmatrix} \frac{6}{5} \\ \frac{3}{5} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3}{5} \\ -\frac{6}{5} \end{bmatrix} + \begin{bmatrix} \frac{12}{5} \\ \frac{6}{5} \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \tag{10}$$

Formulas (9) and (10) provide two different ways of viewing the image of the vector (1, 1) under multiplication by A: Formula (9) tells us directly that the image of this vector is (3, 0), whereas Formula (10) tells us that this image can also be obtained by projecting (1, 1) onto the eigenspaces corresponding to $\lambda_1 = -3$ and $\lambda_2 = 2$ to obtain the vectors $\left(-\frac{1}{5}, \frac{2}{5}\right)$ and $\left(\frac{6}{5}, \frac{3}{5}\right)$, then scaling by the eigenvalues to obtain $\left(\frac{3}{5}, -\frac{6}{5}\right)$ and $\left(\frac{12}{5}, \frac{6}{5}\right)$, and then adding these vectors (see Figure 7.2.1).

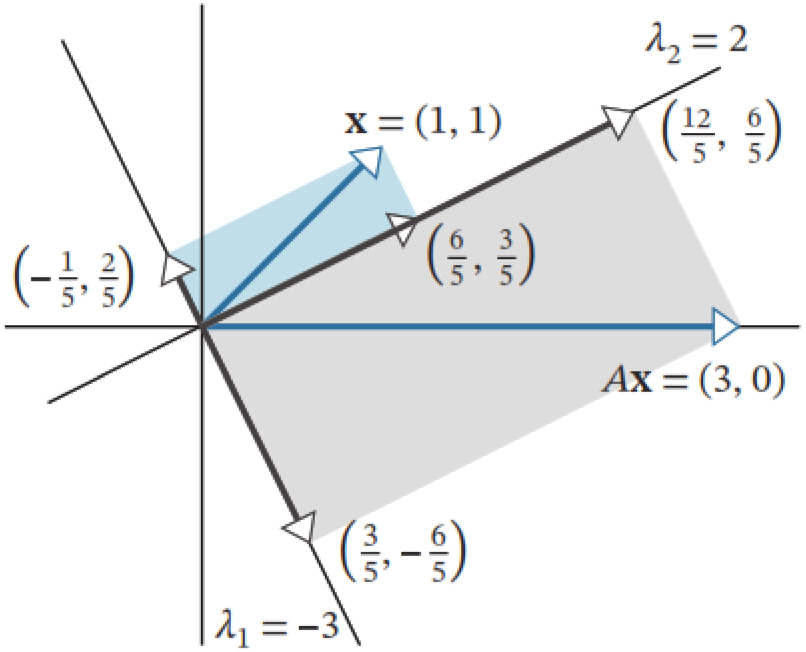

$\lambda_2 = 2$
$\mathbf{x} = (1, 1)$
$\left(\frac{12}{5}, \frac{6}{5}\right)$
$\left(\frac{6}{5}, \frac{3}{5}\right)$
$\left(-\frac{1}{5}, \frac{2}{5}\right)$
$A\mathbf{x} = (3, 0)$
$\left(\frac{3}{5}, -\frac{6}{5}\right)$
$\lambda_1 = -3$

FIGURE 7.2.1

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Linear Algebra (MT-1004)

Lecture # 39

National University of computer and emerging sciences

Foundation for Advancement of Science and Technology

Definition of a Quadratic Form

Expressions of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

occurred in our study of linear equations and linear systems. If $a_1, a_2, \ldots, a_n$ are treated as constants, then this expression is a real-valued function of the ***variables*** $x_1, x_2, \ldots, x_n$ and is called a ***linear form*** on R^n. All variables in a linear form occur to the first power and there are no products of variables. Here we will be concerned with ***quadratic forms*** on R^n, which are functions of the form

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2 + (\text{all possible terms } a_kx_ix_j \text{ in which } i \neq j)$$

The terms of the form $a_kx_ix_j$ in which i is $\neq j$ are called ***cross product terms***. It is common to combine the cross product terms involving x_ix_j with those involving x_jx_i to avoid duplication. Thus, a general quadratic form on R^2 would typically be expressed as

$$a_1x_1^2 + a_2x_2^2 + 2a_3x_1x_2 \tag{1}$$

and a general quadratic form on R^3 as

$$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + 2a_4x_1x_2 + 2a_5x_1x_3 + 2a_6x_2x_3 \tag{2}$$

EXAMPLE 1 | Expressing Quadratic Forms in Matrix Notation

In each part, express the quadratic form in the matrix notation $\mathbf{x}^T A\mathbf{x}$, where A is symmetric.

(a) $2x^2 + 6xy - 5y^2$ (b) $x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3$

Solution The diagonal entries of A are the coefficients of the squared terms, and the off-diagonal entries are half the coefficients of the cross product terms, so

$$2x^2 + 6xy - 5y^2 = [x \quad y]\begin{bmatrix} 2 & 3 \\ 3 & -5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3 = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 1 & 2 & -1 \\ 2 & 7 & 4 \\ -1 & 4 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Change of Variable in a Quadratic Form

There are three important kinds of problems that occur in applications of quadratic forms:

Problem 1	If $\mathbf{x}^T A\mathbf{x}$ is a quadratic form on R^2 or R^3, what kind of curve or surface is represented by the equation $\mathbf{x}^T A\mathbf{x} = k$?
Problem 2	If $\mathbf{x}^T A\mathbf{x}$ is a quadratic form on R^n, what conditions must A satisfy for $\mathbf{x}^T A\mathbf{x}$ to have positive values for $\mathbf{x} \neq \mathbf{0}$?
Problem 3	If $\mathbf{x}^T A\mathbf{x}$ is a quadratic form on R^n, what are its maximum and minimum values if $\mathbf{x}$ is constrained to satisfy $\|\mathbf{x}\| = 1$?

Theorem 7.3.1

The Principal Axes Theorem

If A is a symmetric $n \times n$ matrix, then there is an orthogonal change of variable that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into a quadratic form $\mathbf{y}^T D\mathbf{y}$ with no cross product terms. Specifically, if P orthogonally diagonalizes A, then making the change of variable $\mathbf{x} = P\mathbf{y}$ in the quadratic form $\mathbf{x}^T A\mathbf{x}$ yields the quadratic form

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

in which $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A corresponding to the eigenvectors that form the successive columns of P.

Proof If we make the change of variable $\mathbf{x} = P\mathbf{y}$ in the quadratic form $\mathbf{x}^T A\mathbf{x}$, then we obtain

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T A P\mathbf{y} = \mathbf{y}^T(P^T A P)\mathbf{y} \tag{6}$$

Find an orthogonal change of variable that eliminates the cross product terms in the quadratic form $Q = x_1^2 - x_3^2 - 4x_1x_2 + 4x_2x_3$, and express Q in terms of the new variables.

Solution The quadratic form can be expressed in matrix notation as

$$Q = \mathbf{x}^T A\mathbf{x} = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

The characteristic equation of the matrix A is

$$\begin{vmatrix} \lambda - 1 & 2 & 0 \\ 2 & \lambda & -2 \\ 0 & -2 & \lambda + 1 \end{vmatrix} = \lambda^3 - 9\lambda = \lambda(\lambda + 3)(\lambda - 3) = 0$$

so the eigenvalues are $\lambda = 0, -3, 3$. We leave it for you to show that orthonormal bases for the three eigenspaces are

$$\lambda = 0: \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \quad \lambda = -3: \begin{bmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \quad \lambda = 3: \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

Thus, a substitution $\mathbf{x} = P\mathbf{y}$ that eliminates the cross product terms is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

This produces the new quadratic form

$$Q = \mathbf{y}^T(P^TAP)\mathbf{y} = [y_1 \quad y_2 \quad y_3]\begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = -3y_2^2 + 3y_3^2$$

in which there are no cross product terms.

Positive Definite Quadratic Forms

We will now consider the second of the two problems posed earlier, determining conditions under which $\mathbf{x}^T A\mathbf{x} > 0$ for all nonzero values of $\mathbf{x}$. We will explain why this is important shortly, but first let us introduce some terminology.

Definition 1

A quadratic form $\mathbf{x}^T A\mathbf{x}$ is said to be

positive definite if $\mathbf{x}^T A\mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$;

negative definite if $\mathbf{x}^T A\mathbf{x} < 0$ for $\mathbf{x} \neq \mathbf{0}$;

indefinite if $\mathbf{x}^T A\mathbf{x}$ has both positive and negative values.

Theorem 7.3.2

If A is a symmetric matrix, then:

(*a*) $\mathbf{x}^T A\mathbf{x}$ is positive definite if and only if all eigenvalues of A are positive.

(*b*) $\mathbf{x}^T A\mathbf{x}$ is negative definite if and only if all eigenvalues of A are negative.

(*c*) $\mathbf{x}^T A\mathbf{x}$ is indefinite if and only if A has at least one positive eigenvalue and at least one negative eigenvalue.

Remark The three classifications in Definition 1 do not exhaust all possibilities. Specifically:

- $\mathbf{x}^T A\mathbf{x}$ is ***positive semidefinite*** if $\mathbf{x}^T A\mathbf{x} \geq 0$ if $\mathbf{x} \neq \mathbf{0}$
- $\mathbf{x}^T A\mathbf{x}$ is ***negative semidefinite*** if $\mathbf{x}^T A\mathbf{x} \leq 0$ if $\mathbf{x} \neq \mathbf{0}$

EXAMPLE 4 | Positive Definite Quadratic Forms

It is not usually possible to tell from the signs of the entries in a symmetric matrix A whether that matrix is positive definite, negative definite, or indefinite. For example, the entries of the matrix

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

are nonnegative, but the matrix is indefinite since its eigenvalues are $\lambda = 1, 4, -2$ (verify). To see this another way, write out the quadratic form as

$$\mathbf{x}^T A\mathbf{x} = [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 3x_1^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$$

We can now see, for example, that

$$\mathbf{x}^T A\mathbf{x} = 4 \quad \text{for} \quad x_1 = 0, \quad x_2 = 1, \quad x_3 = 1$$

and

$$\mathbf{x}^T A\mathbf{x} = -4 \quad \text{for} \quad x_1 = 0, \quad x_2 = 1, \quad x_3 = -1$$

Identifying Positive Definite Matrices

As positive definite matrices arise in many applications, it will be useful to learn a little more about them. We already know that a symmetric matrix is positive definite if and only if its eigenvalues are all positive; now we will give a criterion that can be used to determine whether a symmetric matrix is positive definite without the need for finding the eigenvalues. For this purpose we define the ***k*th principal submatrix** of an $n \times n$ matrix A to be the $k \times k$ submatrix consisting of the first k rows and columns of A. For example, here are the principal submatrices of a general 4×4 matrix:

$$\left[\begin{array}{c|ccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right] \quad \left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right] \quad \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right] \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

First principal submatrix | Second principal submatrix | Third principal submatrix | Fourth principal submatrix $= A$

Theorem 7.3.4

If A is a symmetric matrix, then:

(*a*) A is positive definite if and only if the determinant of every principal submatrix is positive.

(*b*) A is negative definite if and only if the determinants of the principal submatrices alternate between negative and positive values starting with a negative value for the determinant of the first principal submatrix.

(*c*) A is indefinite if and only if it is neither positive definite nor negative definite and at least one principal submatrix has a positive determinant and at least one has a negative determinant.

EXAMPLE 5 | Working with Principal Submatrices

The matrix

$$A = \begin{bmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{bmatrix}$$

is positive definite since the determinants

$$|2| = 2, \quad \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3, \quad \begin{vmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{vmatrix} = 1$$

are all positive. Thus, we are guaranteed that all eigenvalues of A are positive and $\mathbf{x}^T A \mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$.

National University of Computer & Emerging Sciences
MT-1004 Linear Algebra
Fall 2023

Course Content for Final Exam

Contents/Topics	Exercises
System of Linear Equations and Matrices: Invertible Matrices, Introduction to linear Transformations	**1.6** (1-20) **1.8** (1-24, 27-41)
Application of linear systems • Network Analysis • Polynomial interpolation	**1.10** (1-8,13-16)
General Vector Spaces: Real Vector Space, Spanning Sets, Linear Independence,	**4.1** (1-14) **4.3** (1-13,17,18) **4.4**(1-14)
Coordinates and Bases, Dimensions	**4.5**(1-22) **4.6** (Q1-8,12,13,18-20)
Bases for row, column, and null spaces, Rank and Nullity	**4.8** (Q1-15,18,19, 25-28) **4.9** (Q1-14,19-23)
Eigenvalues and Eigenvectors: Eigen values and Eigenvectors, Diagonalization	**5.1** (1-16) **5.2** (1-20)
MID-II	
Inner Product Space: Inner products , Angle and Orthogonality in inner product spaces	**6.1** (1-26) **6.2** (1-12, 17-19)
Gram-Schmidt Process, QR-Decomposition.	**6.3** (1-14, 27-31), (44-49)
Diagonalization and Quadratic Forms: Orthogonal Matrices Orthogonal Diagonalization. Spectral Decomposition	**7.1** (1-6) **7.2** (1-18)
Quadratic Forms	**7.3** (1-8 ,17-28)