Part I
Linear Algebra

Chapter 1

7

1.1 Definition of a Matrix

Definition 1.1.1 (Matrix) A rectangular array of numbers is called a matrix. We shall mostly be concerned with matrices having real numbers as entries. The horizontal arrays of a matrix are called its rows and the vertical

arrays are called its columns. A matrix having m rows and n columns is said to have the order m × n. A matrix A of order m × n can be represented in the following form: $a_{11}\ a_{12} \cdots a_{1n}$

$a_{21}\ a_{22} \cdots a_{2n}$

$a_{m1}\ a_{m2} \cdots a_{mn}$ □□ □,

10 CHAPTER 1. MATRICES Example 1.1.4 The linear system of equations 2x + 3y = 5 and 3x + 2y = 5 can be identified with the matrix 2 3 : 5

3 2 : 5 ·

1.1.1 Special Matrices

Definition 1.1.5 1. A matrix in which each entry is zero is called a zero-matrix, denoted by0. For $0_{2\times 2} =$ 0 0

0 0 and $0_{2\times 3} =$ 0 0 0

0 0 0 ·

2. A matrix having the number of rows equal to the number of columns is called a square matrix. Thus, its order is m × m (for some m) and is represented by m only. 3. In a square matrix, A = [a_{ij}], of order n, the entries

$a_{11}, a_{22}, \ldots, a_{nn}$ are called the diagonal entries and form the principal diagonal of A. 4. A square matrix A = [a_{ij}] is said to be a diagonal matrix if a_{ij} = 0 for i '=j. In other words, the non-zero entries appear only on the

principal diagonal. For example, the zero matrix 0_n and 4 0 0 1 are a few diagonal matrices. A diagonal matrix D of order n with the diagonal entries $d_1, d_2, \ldots, d_n$ is denoted by D=diag($d_1, \ldots, d_n$). If d_i = d for all i = 1, 2, .

. . , n then the diagonal matrix D is called a scalar matrix. 5. A square matrix A = [a_{ij}] with a_{ij} = 1 if i = j 0 if i '= j is called the identity matrix, denoted by I_n. For example, I_2 = 1 0

0 1 , and I_3 = □□ □1 0 0

where a_{ij} is the entry at the intersection of the i th row and j th column. In a order. Remark 1.1.2 Some books also use □□□ □$a_{11}\ a_{12} \cdots a_{1n}$ example,

more concise manner, we also denote the matrix A by [a_{ij}] by suppressing its 0 1 0

Matrices

$a_{21}\ a_{22} \cdots a_{2n}$

$a_{m1}\ a_{m2} \cdots a_{mn}$ □□□ □to represent a matrix.

Let A = 1 3 7

4 5 6 . Then a_{11} = 1, a_{12} = 3, a_{13} = 7, a_{21} = 4, a_{22} = 5, and a_{23} = 6. A matrix having only one column is called a column vector; and a matrix with only one row is called a row vector. Whenever a vector is used, it should be understood from the context whether it is a row vector or a column vector. Definition 1.1.3 (Equality of two Matrices) Two matrices A = [a_{ij}] and B = [b_{ij}] having the same order

□

m × n are equal if a_{ij} = b_{ij} for each i = 1, 2, . . . , m and j = 1, 2, . . . , n. In other words, two matrices are said to be equal if they have the same order and their corresponding

□

A =

□

0 0 1 □□ □. The subscript n is suppressed in case the order is clear from the context or if no confusion arises. 6. A square matrix A = [a_{ij}] is said to be an upper triangular matrix if a_{ij} = 0 for i > j. A square matrix A = [a_{ij}] is

said to be an lower triangular matrix if a_{ij} = 0 for i < j. A square matrix A is said to be triangular if it is an upper or a lower triangular matrix. For example □□ □2 1 4 0 0 −2 □□ □ is an upper triangular matrix. An upper

triangular matrix will be represented

0 3 −1

$a_{11}\ a_{12} \cdots a_{1n}$

0 $a_{22} \cdots a_{2n}$

0 0 · · · a_{nn} □□□ □.

1.2 Operations on Matrices

Definition 1.2.1 (Transpose of a Matrix) The transpose of an m× n matrix A= [a_{ij}] is defined as the entries are equal.

.. ..

. . . . . .

9

n × m matrix B = [b_{ij}], with b_{ij} = a_{ji} for 1 ≤ i ≤ m and 1 ≤ j ≤ n. The transpose of Ais denoted byA^t.

□

□

by

□

□

.. ..

. . . .

1.2. OPERATIONS ON MATRICES 11 That is, by the transpose of an m × n matrix A, we mean a matrix of order n × m having the rows

of A as its columns and the columns of A as its rows. For example, if A = 1 4 5

0 1 2 then A^t = 1 0

4 1

5 2 . Thus, the transpose of a row vector is a column vector and vice-versa. Theorem 1.2.2 For any matrix A, we have $(A^t)^t$ = A. Proof. Let A = $[a_{ij}]$, A^t = $[b_{ij}]$ and $(A^t)^t$ = $[c_{ij}]$. Then, the definition of transpose gives c_{ij} = b_{ji} = a_{ij} for all i, j and the result follows. Definition 1.2.3 (Addition of Matrices) let A = $[a_{ij}]$ and B = $[b_{ij}]$ be are two m× n matrices. Then the sum A + B is defined to be the matrix C = $[c_{ij}]$ with c_{ij} = a_{ij} + b_{ij}. Note that, we define the sum of two matrices only when the order of the two matrices are same. Definition 1.2.4 (Multiplying a Scalar to a Matrix) Let A = $[a_{ij}]$ be an m × n matrix. Then for any element k ∈ R, we define kA = $[ka_{ij}]$. For example, if A = 1 4 5

0 1 2 and k = 5, then 5A = 5 20 25

0 5 10 . Theorem 1.2.5 Let A, B and C be matrices of order m × n, and let k, H ∈ R. Then 2. (A + B) + C = A + (B + C) (associativity).

3. k(HA) = (kH)A.
Proof. Part 1.
Let A = $[a_{ij}]$ and B = $[b_{ij}]$. Then
A + B = $[a_{ij}]$ + $[b_{ij}]$ = $[a_{ij} + b_{ij}]$ = $[b_{ij} + a_{ij}]$ = $[b_{ij}]$ + $[a_{ij}]$ = B + A as real numbers commute. The reader is required to prove the other parts as all the results follow from the properties of real numbers. Exercise 1.2.6 1.

Suppose A + B = A. Then show that B = 0. 2. Suppose A + B = 0. Then show that B = (−1)A = $[-a_{ij}]$. Definition 1.2.7 (Additive Inverse) Let A be an m × n matrix. 1. Then there exists a matrix B with A + B = 0. This matrix B is called the additive inverse of A, and is denoted by −A = (−1)A.

1. A + B = B + A (commutativity).
2. Also, for the matrix $0_{m\times n}$, A+0 = 0+ A = A. Hence, the matrix $0_{m\times n}$ is called the additive identity.
4. (k + H)A = kA + HA.

 = (AB)C_{ij}. Part 5. For all j = 1, 2, . . . , n, we have
as d_{ik} = 0 whenever i '= k. Hence, the required result follows. The reader is required to prove the other parts. Exercise 1.2.12 1. Let A and B be two matrices. If the matrix addition A + B is defined, then prove

1.2.1 Multiplication of Matrices

Definition 1.2.8 (Matrix Multiplication / Product) Let A = $[a_{ij}]$ be an m×n matrix and B= $[b_{ij}]$ bean n × r matrix. The product AB is a matrix C = $[c_{ij}]$ of order m × r, with $a_{ik}b_{kj}$ = $a_{i1}b_{1j}$ + $a_{i2}b_{2j}$ + · · · + $a_{in}b_{nj}$. Observe that the product AB is defined if and only if the number of columns of A = the number of rows of B. For example, if A = 1 2 3 2 4 1 and B = 1 2 1

0 0 3

1 0 4 then

AB = 1 + 0 + 3 2 + 0 + 0 1 + 6 + 12

2 + 0 + 1 4 + 0 + 0 2 + 12 + 4 = 4 2 19

3 4 18 . Note that in this example, while AB is defined, the product BA is not defined. However, for square matrices A and B of the same order, both the product AB and BA are defined. Definition 1.2.9 Two square matrices A and B are said to commute if AB =BA. Remark 1.2.10 1. Note that if A is a square matrix of order n then AI_n =I_nA. Also for anyd∈R, the matrix dI_n commutes with every square matrix of order n. The matrices dI_n for anyd∈Rare called scalar matrices. 2. In general, the matrix product is not commutative. For example, consider the following two matrices A = 1 1

0 0 and B = 1 0

1 0 . Then check that the matrix product

c_{ij} = ^{*n}

AB = 2 0

0 0 '= 1 1

$^{k=1}$

1 1 = BA. Theorem 1.2.11 Suppose that the matrices A, B and C are so chosen that the matrix multiplications are defined. 1. Then (AB)C = A(BC). That is, the matrix multiplication is associative. 2. For any k ∈ R, (kA)B = k(AB) = A(kB).

3. Then A(B + C) = AB + AC. That is, multiplication distributes over addition. 4. If A is an n × n matrix then AI_n = I_nA = A. 5. For any square matrix A of order n and D = diag(d_1, d_2, . . . , d_n), we have • the first row of DA is d_1 times the first row of A;

• for 1 ≤ i ≤ n, the ith row of DA is $d_i$times the ith row of A. A similar statement holds for the columns of A when A is multiplied on the right by D. Proof. Part 1. Let A = $[a_{ij}]_{m\times n}$, B = $[b_{ij}]_{n\times p}$ and C = $[c_{ij}]_{p\times q}$. Then $b_{k0}c_{0j}$ and

$(AB)_{i0}$ = ^{*n}

2. Let A = √1 √1 √13 √1 − √1 0

. Then A is an orthogonal matrix. 3. Let A = $[a_{ij}]$ be an n×n matrix with a_{ij} = 1 if i = j + 1 n − 1. The matrices A for which a positive integer k exists such that A^k =0 are called nilpotent matrices. The least positive integer k for which A^k = 0 is called the order of nilpotency. 4. Let A = 1 0

0 0 . Then A^2 = A. The matrices that satisfy the condition that A^2=Aare called

$(BC)_{kj}$ = ^{*p}

$a_{ik}b_{k0}$.

$_{0=1}$

$^{k=1}$

idempotent matrices. Exercise 1.3.3 1. Show that for any square matrix A, S =$^1_2$(A + A^t) is symmetric, T=$^1_2$(A−A^t)is skew-symmetric, and A = S + T. 2. Show that the product of two lower triangular matrices is a lower triangular matrix. A similar statementholds for upper triangular matrices. 3. Let A and B be symmetric matrices. Show that AB is symmetric if and only if AB=BA.

that $(A + B)^t$= A^t + B^t. Also, if the matrix product AB is defined then prove that $(AB)^t$ = B^tA^t.
4. Show that the diagonal entries of a skew-symmetric matrix are zero. 5. Let A, B be skew-symmetric matrices with AB = BA. Is the matrix ABsymmetric or skew-symmetric? 6. Let A be a symmetric matrix of order n with A^2 = 0. Is it necessarily true that A=0?

2. Let A = [a_1, a_2, . . . , a_n] and B = b_1
Therefore,
. Compute the matrix products AB and BA. a_{ik} ^{*p}

A(BC) $_{ij}$ = ^{*n}

a_{ik} BC $_{kj}$ = ^{*n}

$b_{k0}c_{0j}$

3. Let n be a positive integer. Compute $A^n$for the following matrices:

$^{k=1}$

$^{k=1}$

$_{0=1}$

^{*p}

^{*p}

= ^{*n}

$a_{ik}\ b_{k0}c_{0j}\ =\ ^{*n}$

0 1 , 1 1 1

0 0 1 , 1 1 1

$a_{ik}b_{k0}\ c_{0j}$

1 1

0 1 1

1 1 1

0=1

0=1

$=\ ^{*p}$

^{*n}

7. Let A be a nilpotent matrix. Show that there exists a matrix B such that B(I +A) =I = (I +A)B.

$\frac{1}{\sqrt{6}}\ \frac{1}{\sqrt{6}}\ -\ \frac{\sqrt{2}}{6}$

1.3.1 Submatrix of a Matrix

Definition 1.3.4 A matrix obtained by deleting some of the rows and/or columns of a matrix is said to be a submatrix of the given matrix.

For example, if A = 1 4 5 0 otherwise. Then A^n =0 and A^0 '=0for 1≤H≤

k=1 $a_{ik}b_{k0}\ c_{0j}\ =\ ^{*t}$ k=1

1 1 1 . Can you guess a formula for A^n and prove it by induction?

AB $_{i0}c_{0j}$

0=1

k=1

0=1

4. Find examples for the following statements. (a) Suppose that the matrix product AB is defined. Then the product BA need not be defined. $(DA)_{lj}\ =\ ^{*n}$

0 1 2 , a few submatrices of A are [1], [2], 1 0 , [1 5], 1 5

, A.

0 2 But the matrices 1 4

(b) Suppose that the matrix products AB and BA are defined. Then the matrices AB and BA can

$d_{ik}a_{kj}\ =\ d_ia_{ij}$

have different orders. (c) Suppose that the matrices A and B are square matrices of order n. Then AB and BA may or may not be equal. k=1

1 0 and 1 4

0 2 are not submatrices of A. (The reader is advised to give reasons.)

Miscellaneous Exercises

Exercise 1.3.5 1. Complete the proofs of Theorems 1.2.5 and 1.2.11.

2. Let x = x^1 x_2 , y = y_1

1.3 Some More Special Matrices

Definition 1.3.1 1. A matrix A over R is called symmetric if A^t = A and skew-symmetric if A^t = −A.

2. A matrix A is said to be orthogonal if AA^t = A^tA = I.

Example 1.3.2 1. Let A = 1 2 3

b^2..

3 −1 4 and B = 0 1 2 −1 0 −3 −2 3 0 . Then A is a symmetric matrix and

b_n

2 4 −1

B is a skew-symmetric matrix.

, A = 1 0

. Geometrically interpret y=Ax

3. Consider the two coordinate transformations

and B = cos θ − sin θ

0 −1

sin θ cos θ

y_2

and y = Bx.

$x_1 = a_{11}y_1 + a_{12}y_2$

$x_2 = a_{21}y_1 + a_{22}y_2$ and $y_1 = b_{11}z_1 + b_{12}z_2$ $y_2 = b_{21}z_1 + b_{22}z_2$.

 (a) Compose the two transformations to express x_1, x_2 in terms of z_1, z_2.

(b) If x^t = [x_1, x_2], y^t = [y_1, y_2] and z^t = [z_1, z_2] then find matrices A, B and C such that (c) Is C = AB?

4. For a square matrix A of order n, we define trace of A, denoted by tr (A) as

tr (A) = a_{11} + a_{22} + · · · a_{nn}. Then for two square matrices, A and B of the same order, show the following: (a) tr (A + B) = tr (A) + tr (B). (b) tr (AB) = tr (BA).

5. Show that, there do not exist matrices A and B such that AB − BA = cI_n for any c '= 0.

Theorem 1.3.6 is very useful due to the following reasons: 1. The order of the matrices P, Q, H and K are smaller than that of Aor B. 2. It may be possible to block the matrix in such a way that a few blocks are either identity matrices or zero matrices. In this case, it may be easy to handle the matrix product using the block form. 3. Or when we want to prove results using induction, then we may assume the result for r×rsubmatrices and then look for (r + 1) × (r + 1) submatrices, etc.

For example, if A = 1 2 0 2 5 0 and B = a b c d

e f , Then

a b 0

2 5 c d + 0 [e f] = a + 2c b + 2d

AB = 1 2

6. Let A and B be two m × n matrices and let x be an n × 1 column vector. (a) Prove that if Ax = 0 for all x, then A is the zero matrix. (b) Prove that if Ax = Bx for all x, then A = B.

0 −1 2

2a + 5c 2b + 5d ·

7. Let A be an n × n matrix such that AB = BA for all n × n matrices B. Show that A = αI for some

3 1 . Show that there exist infinitely many matrices B such that BA = I_2. Also, show 2 1

that there does not exist any matrix C such that AC = I_3.

x = Ay, y = Bz and x = Cz.

8. Let A = 1 2 0 −1 2

3 1 4

3 1 4

−2 5 −3 , or

−2 5 −3 and so on. 3 1 4

1.3.1 Block Matrices

−2 5 −3 , then A can be decomposed as follows: 3 1 4

Let A be an n × m matrix and B be an m × p matrix. Suppose r < m. Then, we can decompose the

G H . Then the matrices P, Q, R, SandE, F, G, H, are called the blocks of the matrices A and B, respectively. Even if A + B is defined, the orders of P and E may not be same and hence, we may not be able to add A and B in the block form. But, if A+ B and P + E is defined then A+B= P+E Q+F R+G S+H . Similarly, if the product AB is defined, the product P E need not be defined. Therefore, we

can talkof matrix product AB as block product of matrices, if both the products ABand P Eare defined. And in this case, we have AB = P E + QG P F + QH

P Q

E F

matrices A and B as A = [P Q] and B = H K ; where P has order n × r and H has order r × p. That is, the matrices P and Q are

submatrices of A and P consists of the first r columns of A and Q consists of the last m − r columns of A. Similarly, H and K are

submatrices of B and H consists of the first r

R S and B = s_1 s_2

rows of B and K consists of the last m − r rows of B. We now prove the following important theorem. Theorem 1.3.6 Let A

Proof. First note that the matrices P H and QK are each of order n × p. The matrix products PH

and QK are valid as the order of the matrices P, H, Q and K are respectively, n × r, r × p, n × (m − r)

and (m−r)×p. Let P = [P_{ij}], Q = [Q_{ij}], H = [H_{ij}], and K = [k_{ij}]. Then, for 1 ≤ i ≤ n and 1 ≤ j ≤ p, we have

α ∈ R.

= $(P H)_{ij}$ + $(QK)_{ij}$ = $(P H + QK)_{ij}$.

0 −1 2

0 1 1 1

A =

1 1 2 1

, or A

=

0 1 1 0

1 1 1 1

r_2

n_2

3. Let A = [a_{ij}] and B = [b_{ij}] be two matrices. Suppose a_1, a_2, . . . , a_n are the rows of A and b_1, b_2, . . . ,

b_p are the columns of B. If the product AB is defined, then show that

AB = [Ab_1, Ab_2, . . . , Ab_p] = a_1B

$(AB)_{ij}$ = ^{*m}

$a_{ik}b_{kj}$ = ^{*r}

$a_{ik}b_{kj}$ + ^{*m}

$a_{ik}b_{kj}$

. [That is, left multiplication by A, is same as multiplying each column of B by A. Similarly, right multiplication by B, is same as multiplying each row of A by B.]

k=1

k=1

k=r+1

= ^{*r}

k=1 $P_{ik}H_{kj}$ + ^{*m}

$Q_{ik}K_{kj}$

k=r+1

1.4 Matrices over Complex Numbers

Here the entries of the matrix are complex numbers. All the definitions still hold. One just needs to look at the following additional definitions. Definition 1.4.1 (Conjugate Transpose of a Matrix) 1. Let A be an m×n matrix

over C. If A = [a_{ij}] then the Conjugate of A, denoted by A, is the matrix B = [b_{ij}] with b_{ij} = a_{ij} . For example, Let A = 1 4 + 3i i 0 1 i − 2 . Then A = 1 4 − 3i −i 0 1 −i − 2 .

2. Let A be an m × n matrix over C. If A = [a_{ij}] then the Conjugate Transpose of A, denoted by A^*, is

the matrix B = [b_{ij}] with b_{ij} = a_{ji} . For example, Let A = 1 4 + 3i i 0 1 i − 2 . Then

B..

a_2 .

a_nB

1 0

3. A square matrix A over C is called Hermitian if A^* = A.

4. A square matrix A over C is called skew-Hermitian if A^* = −A. 5. A square matrix A over C is called unitary if A^*A = AA^* = I. 6. A square matrix A over C is called Normal if AA^* = A^*A. Remark 1.4.2 If A = [a_{ij}] with a_{ij} ∈

R, then A^* = A^t. Exercise 1.4.3 1. Give examples of Hermitian, skew-Hermitian and unitary matrices that have entries with non-zero imaginary parts. 2. Restate the results on transpose in terms of conjugate transpose. 3.

Show that for any square matrix A, S = $^{A+A_*}$

$_2$is Hermitian, T = $^{A-A_*}$

$_2$is skew-Hermitian, and

4. Show that if A is a complex triangular matrix and AA^* = A^*A then A is a diagonal matrix.

A^* =

.

4 − 3i 1

−i −i − 2

= [a_{ij}] = [P Q] and B = [b_{ij}] = H K be defined as above. Then AB = P H + QK.

RE + SG RF + SH . That is, once a partition of A is fixed, the partition of B has to be properly chosen forpurposes of

block addition or multiplication. Exercise 1.3.7 1. Compute the matrix product AB using the block matrix

multiplication for the matrices

If A =

0 1 0 1

1 0 0 1

and B =

1 2 2 1

−2 5 −3

−1 1 −1 1

.

0 −1 2

2. Let A = P Q

A =

R S

. If P, Q, R and S are symmetric, what can you say about A? Are P, Q, RandS symmetric, when A is

symmetric? Suppose A = m_1 m^2 n_1

r_1

A = S + T.
A =

Chapter 2
Linear System of Equations 2.1 Introduction

Let us look at some examples of linear systems. 1. Suppose a, b ∈ R. Consider the system ax = b. (a) If a '= 0 then the system has a unique solution x = $\frac{b}{a}$.

(b) If a = 0 and
i. b '= 0 then the system has no solution.
ii. b = 0 then the system has infinite number of solutions, namely all x ∈ R. 2. We now consider a system with 2 equations in 2 unknowns. Consider the equation ax + by = c. If one of the coefficients, a or b is non-zero,

then this linear equation represents a line in R^2. Thus for the system $a_1x + b_1y = c_1$ and $a_2x + b_2y = c_2$, the set of solutions is given by the points of intersection of the two lines. There are three cases to be considered.

Each case is illustrated by an example. (a) Unique Solution

x + 2y = 1 and x + 3y = 1. The unique solution is $(x, y)^t = (1, 0)^t$. Observe that in this case, $a_1b_2 - a_2b_1$ '= 0.
(b) Infinite Number of Solutions
20 CHAPTER 2. LINEAR SYSTEM OF EQUATIONS (a) Unique Solution Consider the system x+y+z = 3, x+4y+2z = 7 and 4x+ 10y−z = 13. The unique solution to this system is $(x, y, z)^t = (1, 1, 1)^t$; i.e. the three planes

intersect at a point. (b) Infinite Number of Solutions Consider the system x + y + z = 3, x + 2y + 2z = 5 and 3x + 4y + 4z = 11. The set of solutions to this system is $(x, y, z)^t = (1, 2 - z, z)^t = (1, 2, 0)^t + z(0, -1, 1)^t$, with z

arbitrary; the three planes intersect on a line. (c) No Solution The system x + y + z = 3, x + 2y + 2z = 5 and 3x + 4y + 4z = 13 has no solution. In thiscase, we get three parallel lines as intersections of the above planes

taken two at a time. The readers are advised to supply the proof. 2.2 Definition and a Solution Method Definition 2.2.1 (Linear System) A linear system of m equations in n unknowns $x_1, x_2, \ldots, x_n$ is a set

of equations of the form $a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$

$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$
$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$ where for $1 \le i \le n$, and $1 \le j \le m$; a_{ij}, $b_i \in R$. Linear System (2.2.1) is called homogeneousif $b_1 = 0 = b_2 = \cdots = b_m$ and non-homogeneous otherwise. We rewrite the above equations in

the form Ax = b, where $a_{11}\ a_{12} \cdots a_{1n}$

$a_{21}\ a_{22} \cdots a_{2n}$
$a_{m1}\ a_{m2} \cdots a_{mn}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, and $b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$
The matrix A is called the coefficient matrix and the block matrix [Ab] , is the augmentedmatrix of the linear system (2.2.1). Remark 2.2.2 Observe that the ith row of the augmented matrix [A b] represents the

i^{th} equation and the jth column of the coefficient matrix A corresponds to coefficients of the jth variable x_j. That is, for $1 \le i \le m$ and $1 \le j \le n$, the entry a_{ij} of the coefficient matrix A corresponds to the ith equation

x + 2y = 1 and 2x + 4y = 2. The set of solutions is $(x, y)^t = (1 - 2y, y)^t = (1, 0)^t + y(-2, 1)^t$ with y arbitrary. In other words, both the equations

represent the same line. Observe that in this case, $a_1b_2 - a_2b_1 = 0$, $a_1c_2 - a_2c_1 = 0$ and $b_1c_2 - b_2c_1 = 0$.

(c) No Solution

x + 2y = 1 and 2x + 4y = 3. The equations represent a pair of parallel lines and hence there is no point of intersection.

Observe that in this case, $a_1b_2 - a_2b_1 = 0$ but $a_1c_2 - a_2c_1$ '= 0. 3. As a last example, consider 3 equations in 3 unknowns. A linear

equation ax + by + cz = d represent a plane in R^3 provided (a, b, c) '= (0, 0, 0). As in the case of 2 equations in 2 unknowns, we

have to look at the points of intersection of the given three planes. Here again, we have three cases. The three cases are illustrated

by examples.

and j^{th} variable x_j. For a system of linear equations Ax = b, the system Ax = 0 is called the associated homogeneous system.

Definition 2.2.3 (Solution of a Linear System) A solution of the linear systemAx =bis a column vector

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}. \quad (2.2.1)$$

y with entries $y_1, y_2, \ldots, y_n$ such that the linear system (2.2.1) is satisfied by substituting y_i in place of x_i. That is, if $y^t = [y_1, y_2, \ldots, y_n]$ then Ay = b holds. Note: The zero n-tuple x = 0 is always a solution of the systemAx =0, and is called the trivialsolution. A non-zero n-tuple x, if it satisfies Ax = 0, is called a non-trivial solution.

A =
$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$.

2.2.1 A Solution Method

Example 2.2.4 Let us solve the linear system x + 7y + 3z = 11, x + y + z = 3, and 4x + 10y − z = 13.

1. The above linear system and the linear system
x + y + z = 3 Interchange the first two equations.
x + 7y + 3z = 11 (2.2.2)
4x + 10y − z = 13 have the same set of solutions. (why?) 2. Eliminating x from 2nd and 3rd equation, we get the linear system

19
x + y + z = 3
6y + 2z = 8 (obtained by subtracting the first
equation from the second equation.)
6y − 5z = 1 (obtained by subtracting 4 times the first equation
from the third equation.) (2.2.3)
This system and the system (2.2.2) has the same set of solution. (why?) 3. Eliminating y from the last two equations of system (2.2.3), we get the system x + y + z = 3

Solution:
6y + 2z = 8
7z = 7 obtained by subtracting the third equation
from the second equation. (2.2.4)
which has the same set of solution as the system (2.2.3). (why?)
4. The system (2.2.4) and system

2. multiply a non-zero constant throughout an equation, say “multiply the kth equation by c '= 0”; (compare the system (2.2.5) and the system (2.2.4).) 3. replace an equation by itself plus a constant multiple of another equation, say “replace the kth equation

$x + y + z = 3$
$3y + z = 4$ divide the second equation by 2
$z = 1$ divide the second equation by 2 (2.2.5)
has the same set of solution. (why?) 5. Now, $z = 1$ implies $y = \frac{4-1}{3} = 1$ and $x = 3-(1+1) = 1$. Or in terms of a vector, the set of solution is $\{(x, y, z)^t : (x, y, z) = (1, 1, 1)\}$.

2.3 Row Operations and Equivalent Systems

Definition 2.3.1 (Elementary Operations) The following operations 1, 2 and 3 are called elementary operations.

1. interchange of two equations, say "interchange the ith and j^{th} equations"; (compare the system (2.2.2) with the original system.)
2. multiply a non-zero constant throughout a row, say "multiply the kth row by $c \neq 0$", denoted $R_k(c)$; (compare the system (2.2.3) with (2.2.2) or the system (2.2.4) with (2.2.3).)
3. replace a row by itself plus a constant multiple of another row, say "replace the kth row by k^{th} row plus c times the jth row", denoted $R_{kj}(c)$.

Exercise 2.3.6 Find the inverse row operations corresponding to the elementary row operations that have been defined just above.

Definition 2.3.7 (Row Equivalent Matrices) Two matrices are said to be row-equivalent if one can be obtained from the other by a finite number of elementary row operations.

Example 2.3.8 The three matrices given below are row equivalent.

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

Solution: In this case, the augmented matrix is $\begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \\ 2 & 0 & 3 & 5 \end{bmatrix}$. The method proceeds along the following steps.

1. Interchange 1st and 2nd equation (or R_{12}).

$$\begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \\ 2 & 0 & 3 & 5 \end{bmatrix} \xrightarrow{R_{12}} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \end{bmatrix}.$$

Whereas the matrix $\begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \\ 2 & 0 & 3 & 5 \end{bmatrix}$ is not row equivalent to the matrix $\begin{bmatrix} 1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 3 \\ 2 & 0 & 3 & 5 \end{bmatrix}$.

2. Divide the 1st equation by 2 (or $R_1(1/2)$).

$$\begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \\ 2 & 0 & 3 & 5 \end{bmatrix} \xrightarrow{R_1(1/2)} \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \end{bmatrix}.$$

3. Add -1 times the 1st equation to the 3rd equation (or $R_{31}(-1)$).

$$\begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \end{bmatrix} \xrightarrow{R_{31}(-1)} \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 & 2 \end{bmatrix}.$$

4. Add -1 times the 2nd equation to the 3rd equation (or $R_{32}(-1)$).

$$\begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 & 2 \end{bmatrix} \xrightarrow{R_{32}(-1)} \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{3}{2} \end{bmatrix}.$$

$y + z = 2$
$x + y + z = 3$
$2x + 3z = 5$

Definition 2.3.9 (Forward/Gauss Elimination Method) Gaussian elimination is a method of solving a linear system $Ax = b$ (consisting of m equations in n unknowns) by bringing the augmented matrix

$$[A\ b] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

to an upper triangular form

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} & d_1 \\ 0 & c_{22} & \cdots & c_{2n} & d_2 \\ 0 & 0 & \cdots & c_{mn} & d_m \end{bmatrix}.$$

This elimination process is also called the forward elimination method. The following examples illustrate the Gauss elimination procedure.

Example 2.3.10 Solve the linear system by Gauss elimination method.

$y + z = 2$
$2x + 3z = 5$
$x + y + z = 3$

a rectangular array gives us the augmented matrix $[A\ b]$.

Definition 2.3.5 (Elementary Row Operations) The elementary row operations are defined as:

1. interchange of two rows, say "interchange the ith and j^{th} rows", denoted R_{ij};
2. multiply a non-zero constant throughout a row, say "multiply the kth row by $c \neq 0$", denoted $R_k(c)$;
3. replace a row by itself plus a constant multiple of another row, say "replace the kth row by k^{th} row plus c times the jth row", denoted $R_{kj}(c)$.

Exercise 2.3.6 Find the inverse row operations corresponding to the elementary row operations that have been defined just above.

Definition 2.3.7 (Row Equivalent Matrices) Two matrices are said to be row-equivalent if one can be obtained from the other by a finite number of elementary row operations.

Example 2.3.8 The three matrices given below are row equivalent.

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

 5. Multiply the 3rd equation by $-\frac{2}{3}$(or $R_3(-\frac{2}{3})$).

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

The last equation gives z = 1, the second equation now gives y = 1. Finally the first equation gives x = 1. Hence the set of solutions is $(x, y, z)^t = (1, 1, 1)^t$, a unique solution. Example 2.3.11 Solve the linear system by Gauss elimination method.

$R_1(1/2)$

$$\begin{aligned} x + y + z &= 3 \\ x + 2y + 2z &= 5 \\ 3x + 4y + 4z &= 11 \end{aligned}$$

Solution: In this case, the augmented matrix is $\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & 2 & 2 & 5 \\ 3 & 4 & 4 & 11 \end{bmatrix}$ and the method proceeds as follows:

1. Add −1 times the first equation to the second equation.

$$\begin{aligned} 3x + 4y + 4z &= 11 \\ x + y + z &= 3 \\ y + z &= 2 \end{aligned} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 3 & 4 & 4 & 11 \end{bmatrix}.$$

2. Add −3 times the first equation to the third equation.

$$\begin{aligned} x + y + z &= 3 \\ y + z &= 2 \\ x + \frac{3}{2}z &= \frac{5}{2} \end{aligned} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix}.$$

3. Add −1 times the second equation to the third equation

$$\begin{aligned} y + z &= 2 \\ x + y + z &= 3 \\ 0 &= 0 \end{aligned} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus, the set of solutions is $(x, y, z)^t = (1, 2 - z, z)^t = (1, 2, 0)^t + z(0, -1, 1)^t$, with z arbitrary. In other words, the system has infinite number of solutions. Example 2.3.12 Solve the linear system by Gauss elimination method.

$$\begin{aligned} x + y + z &= 3 \\ x + 2y + 2z &= 5 \\ 3x + 4y + 4z &= 12 \end{aligned}$$

Solution: In this case, the augmented matrix is $\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & 2 & 2 & 5 \\ 3 & 4 & 4 & 12 \end{bmatrix}$ and the method proceeds as follows:

1. Add −1 times the first equation to the second equation.

$$\begin{aligned} 3x + 4y + 4z &= 12 \\ x + y + z &= 3 \\ y + z &= 2 \end{aligned} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 3 & 4 & 4 & 12 \end{bmatrix}.$$

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

$$\begin{aligned} x + \frac{3}{2}z &= \frac{5}{2} \end{aligned}$$

2. Add −3 times the first equation to the third equation.

$$\begin{aligned} x + y + z &= 3 \\ y + z &= 3 \\ y + z &= 2 \end{aligned} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 \end{bmatrix}.$$

$$\begin{aligned} x + \frac{3}{2}z &= \frac{5}{2} \end{aligned}$$

3. Add −1 times the second equation to the third equation

$$\begin{aligned} y - \frac{1}{2}z &= \frac{1}{2} \\ x + y + z &= 3 \\ 0 &= 1 \end{aligned} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

$$\begin{aligned} x + \frac{3}{2}z &= \frac{5}{2} \end{aligned}$$

The third equation in the last step is

$$-\frac{3}{2}z = -\frac{3}{2}$$

0x + 0y + 0z = 1. This can never hold for any value of x, y, z. Hence, the system has no solution. Remark 2.3.13 Note that to solve a linear system, Ax = b, one needs to apply only the elementary row operations to the augmented matrix [A b].

2.4 Row Reduced Echelon Form of a Matrix

Definition 2.4.1 (Row Reduced Form of a Matrix) A matrix C is said to be in the row reduced form if 1. the first non-zero entry in each row of C is 1; 2. the column containing this 1 has all its other entries zero. A matrix in the row reduced form is also called a row reduced matrix. Example 2.4.2 1. One of the most important examples of a row reduced matrix is the n×nidentitymatrix, I_n. Recall that the (i, j)th entry of the identity matrix is

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

δ_{ij} is usually referred to as the Kronecker delta function.

2. The matrices $\begin{bmatrix} 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ are also in row reduced form.

3. The matrix $\begin{bmatrix} 1 & 0 & 0 & 5 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$ is not in the row reduced form. (why?)

Definition 2.4.3 (Leading Term, Leading Column) For a row-reduced matrix, the first non-zero entry of any row is called a leading term. The columns containing the leading terms are called the leading

columns.

 Definition 2.4.4 (Basic, Free Variables) Consider the linear system Ax = b in n variables and m equations. Let [C d] be the row-reduced matrix obtained by applying the Gauss elimination method to the augmented matrix [A b]. Then the variables corresponding to the leading columns in the first n columns of [C d] are called the basic variables. The variables which are not basic are called free variables. The free variables are called so as they can be assigned arbitrary values and the value of the basic variables can then be written in terms of the free variables. Observation: In Example 2.3.11, the solution set was given by

Definition 2.4.8 (Row Reduced Echelon Matrix) A matrix which is in the row reduced echelon form isalso called a row reduced echelon matrix. Definition 2.4.9 (Back Substitution/Gauss-Jordan Method) The procedure to get to Step II of Example2.3.10 from Step 5 of Example 2.3.10 is called the back substitution. The elimination process applied to obtain the row reduced echelon form of the augmented matrix is calledthe Gauss-Jordan elimination. That is, the Gauss-Jordan elimination method consists of both the forward elimination and the backward substitution. Method to get the row-reduced echelon form of a given matrix A

Let A be an m × n matrix. Then the following method is used to obtain the row-reduced echelon formthe matrix A. Step 1: Consider the first column of the matrix A. If all the entries in the first column are zero, move to the second column. Else, find a row, say i th row, which contains a non-zero entry in the first column. Now, interchange the first row with the i th row. Suppose the non-zero entry in the (1, 1)-position is α'= 0. Dividethe whole row by α so that the (1, 1)-entry of the new matrix is 1. Now, use the 1 to make all theentries below this 1 equal to 0. Step 2: If all entries in the first column after the first step are zero, consider the right m×(n−1)submatrix of the matrix obtained in step 1 and proceed as in step 1. Else, forget the first row and first column. Start with the lower (m−1) ×(n −1) submatrix of thematrix obtained in the first step and proceed as in step 1. Step 3: Keep repeating this process till we reach a stage where all the entries below a particular row, of the matrix then there will be r leading terms. That is, there will be r leading columns. Therefore, if there are r leading terms and n variables, then there will be r basic variables and n − r free variables.

2.4.1 Gauss-Jordan Elimination

We now start with Step 5 of Example 2.3.10 and apply the elementary operations once again. But this time, we start with the 3rd row.

I. Add −1 times the third equation to the second equation (or $R_{23}(-1)$).

z = 1 1 0 $\frac{3}{2}$ $\frac{5}{2}$

y = 2

0 1 0 1

0 0 1 1 .

II. Add $\frac{-3}{2}$ times the third equation to the first equation (or $R_{13}(-\frac{3}{2})$).

x = 1

z = 1 1 0 0 1

y = 1

0 1 0 1

0 0 1 1 .

III. From the above matrix, we directly have the set of solution as $(x, y, z)^t = (1, 1, 1)^t$. Definition 2.4.6 (Row Reduced Echelon Form of a Matrix) A matrix C is said to be in the row reduced echelon form if

1. C is already in the row reduced form;
2. The rows consisting of all zeros comes below all non-zero rows; and
3. the leading terms appear from left to right in successive rows. That is, for 1 ≤ H ≤ k, let i_0 be the leading column of the H th row. Then $i_1 < i_2 < \cdots < i_k$. Example 2.4.7 Suppose A = 0 1 0 2

0 0 1 1 and B = 0 0 0 1 0

0 0 0 1 are in row reduced form. Then the

0 0 0 0

1 1 0 0 0

$x + \frac{3}{2}z = \frac{5}{2}$

corresponding matrices in the row reduced echelon form are respectively, 0 1 0 2 0 0 0 0 and 1 1 0 0 0

0 0 0 1 .

0 0 1 1

0 0 0 1 0

say r, are zero. Suppose at this stage we have obtained a matrix C. Then Chas the following 1. the first non-zero entry in each row of C is 1. These 1's are the leading terms of C and the columns containing these leading terms are the leading columns. 2. the entries of C below the leading term are all zero. Step 4: Now use the leading term in the r th row to make all entries in the r th leading column equal to zero. Step 5: Next, use the leading term in the (r − 1) th row to make all entries in the (r −1) thleadingcolumn equal to zero and continue till we come to the first leading term or column. The final matrix is the row-reduced echelon form of the matrix A. Remark 2.4.10 Note that the row reduction involves only row operations and proceeds fromleft toright. Hence, if A is a matrix consisting of first s columns of a matrix C, then the row reduced form of A will be the first s columns of the row reduced form of C. The proof of the following theorem is beyond the scope of this book and is omitted. Theorem 2.4.11 The row reduced echelon form of a matrix is unique. Exercise 2.4.12 1. Solve the following linear system.

form:

 (a) x + y + z + w = 0, x − y + z + w = 0 and −x + y + 3z + 3w = 0. (b) x + 2y + 3z = 1 and x + 3y + 2z = 1.

(c) x + y + z = 3, x + y − z = 1 and x + y + 7z = 6.
(d) x + y + z = 3, x + y − z = 1 and x + y + 4z = 6.
(e) x + y + z = 3, x + y − z = 1, x + y + 4z = 6 and x + y − 4z = −1. 2. Find the row-reduced echelon form of the following matrices.

−1 1 3 5

−3 −1 13 , 2. 10 8 6 4

1 3 5 7

2 0 −2 −4

9 11 13 15

2.4.2 Elementary Matrices

Definition 2.4.13 A square matrix E of order n is called an elementary matrix if it is obtained by applying exactly one elementary row operation to the identity matrix, I_n. Remark 2.4.14 There are three types of elementary matrices. 1. E_{ij}, which is obtained by the application of the elementary row operation R_{ij} to the identity matrix, I_n. Thus, the (k, H) th entry of E_{ij} is $(E_{ij})_{(k,0)} =$ 1 if k = H and H '= i, j 1 if (k, H) = (i, j) or (k, H) = (j, i) 0 otherwise. 2. $E_k(c)$, which is obtained by the application of the elementary row operation $R_k(c)$ to the identity matrix, I_n. The (i, j) th entry of $E_k(c)$ is $(E_k(c))_{(i,j)} =$ 1 if i = j and i '= k 3. $E_{ij}(c)$, which is obtained by the application of the elementary row operation $R_{ij}(c)$ to the identity

2. Consider the augmented matrix [A b] = 0 1 1 2

1 1 1 3 . Then the result of the steps given below is same as the matrix product 2 0 3 5

0 1 1 2

$$\begin{bmatrix} 1 & 1 & 1 & 3 \end{bmatrix} \xrightarrow{R_{13}} \begin{bmatrix} 1 & 1 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 1 & 2 \end{bmatrix} \xrightarrow{R_{21}(-2)} \begin{bmatrix} 1 & 1 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 1 & 2 \end{bmatrix} \xrightarrow{R_{23}} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 2 & 0 & 3 & 5 \\ 2 & 0 & 3 & 5 \\ 0 & -2 & 1 & -1 \end{bmatrix}$$

$$\xrightarrow{R_{32}(2)} \begin{bmatrix} 1 & 1 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 3 & 3 \end{bmatrix} \xrightarrow{R_3(1/3)} \begin{bmatrix} 1 & 1 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{R_{12}(-1)} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\xrightarrow{R_{23}(-1)} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Now, consider an $m \times n$ matrix A and an elementary matrix E of order n. Then multiplying byEon the right to A corresponds to applying column transformation on the matrix A. Therefore, for eachelementary matrix, there is a corresponding column transformation. We summarize:

Definition 2.4.16 The column transformations obtained by right multiplication of elementary matrices are called elementary column operations.

Example 2.4.17 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix}$ and consider the elementary column operation f which interchanges the second and the third column of A. Then $f(A) = \begin{bmatrix} 1 & 3 & 2 & 2 & 0 & 3 \\ 3 & 5 & 4 \end{bmatrix} = A \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

$E_{23}(-1)E_{12}(-1)E_3(1/3)E_{32}(2)E_{23}E_{21}(-2)E_{13}[Ab]$.

Exercise 2.4.18 1. Let e be an elementary row operation and let $E = e(I)$ be the corresponding ele- mentary matrix. That is, E is the matrix obtained from I by applying the elementary row operatione.

Show that $e(A) = EA$.

2. Show that the Gauss elimination method is same as multiplying by a series of elementary matrices onthe left to the augmented matrix. Does the Gauss-Jordan method also corresponds to multiplying by elementary matrices on the left?Give reasons. 3. Let A and B be two $m \times n$ matrices. Then prove that the two matrices A, Bare row-equivalent if and only if $B = P A$, where P is product of elementary matrices. When is this P unique?

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_1(c) = \begin{bmatrix} c & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } E_{23}(c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}.$$

Example 2.4.15 1. Let $A = \begin{bmatrix} 1 & 2 & 3 & 0 & 2 & 0 & 3 & 4 \\ 3 & 4 & 5 & 6 \end{bmatrix}$. Then

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 3 & 4 & 5 & 6 \end{bmatrix} \xrightarrow{R_{23}} \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 0 & 3 & 4 \\ 3 & 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} A = E_{23} A.$$

That is, interchanging the two rows of the matrix A is same as multiplying on the left by the corre- sponding elementary matrix. In other words, we see that the left multiplication of elementary matrices to a matrix results in elementary row operations.

$$E_{23} = \begin{cases} c & \text{if } i = j = k \\ 0 & \text{otherwise} \end{cases}$$

2.5 Rank of a Matrix

In previous sections, we solved linear systems using Gauss elimination method or the Gauss-Jordan method. In the examples considered, we have encountered three possibilities, namely 1. existence of a unique solution, 2. existence of an infinite number of solutions, and 3. no solution. Based on the above possibilities, we have the following definition.

Definition 2.5.1 (Consistent, Inconsistent) A linear system is called consistent if it admits a solution and is called inconsistent if it admits no solution.

The question arises, as to whether there are conditions under which the linear system $Ax = b$ is consistent. The answer to this question is in the affirmative. To proceed further, we need a few definitions and remarks.

From the last matrix in Step 2b, we deduce row-rank(A) = 2. Remark 2.5.4 Let $Ax = b$ be a linear system with m equations and n unknowns. Then the row-reduced echelon form of A agrees with the first n columns of [A b], and hence row-rank(A) ≤ row-rank([A b]). The reader is advised to supply a proof. Remark 2.5.5 Consider a matrix A. After application of a finite number of elementary column oper ations (see Definition 2.4.16) to the matrix A, we can have a matrix, say B, which has the following rows is a unique number. Also, note that the number of non-zero rows in either the row reduced form or the row reduced echelon form of a matrix are same. Definition 2.5.2 (Row rank of a Matrix) The number of non-zero rows in the row reduced form of a matrix is called the row-rank of the matrix. By the very definition, it is clear that row-equivalent matrices have the same row-rank. For a matrix A, we write 'row-rank (A)' to denote the row-rank of A.

Recall that the row reduced echelon form of a matrix is unique and therefore, the number of non-zero

Example 2.5.3 1. Determine the row-rank of $A = \begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

Solution: To determine the row-rank of A, we proceed as follows.

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix} \xrightarrow{R_{21}(-2),\ R_{31}(-1)} \begin{bmatrix} 1 & 2 & 1 \\ \end{bmatrix}$$

1 1 2 ------------→
0 −1 1 .
1 2 1
0 −1 1 -----------→ $R_2(-1)$, $R_{32}(1)$ 1 2 1
0 1 1
0 0 2 .
1 2 1
0 0 2 ------------→ $R_3(1/2)$, $R_{12}(-2)$ 1 0 −1
0 1 1
0 1 1
0 0 1 .
1 0 −1
$R_{23}(-1)$, $R_{13}(1)$ 1 0 0
0 1 1
0 1 0
0 0 1 The last matrix in Step 1d is the row reduced form of A which has 3 non-zero rows. Thus, row-rank(A) = 3. This result can also be easily deduced from the last matrix in Step 1b. 2. Determine the row-rank of A
= 1 2 1
0 0 1 -----------→
2 3 1
1 1 0 .
Solution: Here we have
1 2 1
1 1 0 ------------→ $R_{21}(-2)$, $R_{31}(-1)$ 1 2 1
2 3 1

properties: 1. The first nonzero entry in each column is 1.
2. A column containing only 0's comes after all columns with at least one non-zero entry. 3. The first non-zero entry (the leading term) in each non-zero column moves down in successive columns. Therefore, we can define column-rank of A as the number of non-zero columns in B. It will be proved later that row-rank(A) = column-rank(A). Thus we are led to the following definition. Definition 2.5.6 The number of non-zero rows in the row reduced form of a matrix A is called the rank of A, denoted rank (A). Theorem 2.5.7 Let A be a matrix of rank r. Then there exist elementary matrices $E_1, E_2, \ldots, E_s$ and $F_1, F_2, \ldots, F_0$ such that

$E_1E_2 \ldots E_s A F_1F_2 \ldots F_0 = I_r$ 0
0 0 .

Proof. Let C be the row reduced echelon matrix obtained by applying elementary row operations to the given matrix A. As rank(A) = r, the matrix C will have the first r rows as the non-zero rows. So by Remark 2.4.5, C will have r leading columns, say $i_1, i_2, \ldots, i_r$. Note that, for $1 \le s \le r$, the i_sth column will have 1 in the sth row and zero elsewhere. We now apply column operations to the matrix C. Let D be the matrix obtained from C by successively interchanging the sth and i_sth column of C for $1 \le s \le r$. Then the matrix D can be written in the form $\begin{bmatrix} I_r & B \\ 0 & 0 \end{bmatrix}$, where B is a matrix of appropriate size. As the (1, 1) block of D is an identity matrix, the block (1, 2) can be made the zero matrix by application of column operations to D. This gives the required result. Exercise 2.5.8 1. Determine the ranks of the coefficient and the augmented matrices that appear in Part 1 and Part 2 of Exercise 2.4.12. 2. For any matrix A, prove that $\text{rank}(A) = \text{rank}(A^t)$.

3. Let A be an $n \times n$ matrix with rank(A) = n. Then prove that A is row-equivalent to I_n.

(a) 0 −1 −1
1 2 1
$R_2(-1)$, $R_{32}(1)$ 1 2 1 0 1 1
(b)
0 0 0 .
0 −1 −1
(c)
(d)

2.6 Existence of Solution of Ax = b

We try to understand the properties of the set of solutions of a linear system through an example, using the Gauss-Jordan method. Based on this observation, we arrive at the existence and uniqueness results for the linear system Ax = b. This example is more or less a motivation.

2.6.1 Example

Consider a linear system Ax = b which after the application of the Gauss-Jordan method reduces to a matrix [C d] with

(a)
(b) 0 −1 −1
0 −1 −1

$$[C\ d] = \begin{bmatrix} 1 & 0 & 2 & -1 & 0 & 0 & 2 & 8 \\ 0 & 1 & 1 & 3 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

For this particular matrix [C d], we want to see the set of solutions. We start with some observations. 1. The number of non-zero rows in C is 4. This number is also equal to the number of non-zero rows in [C d].

2.6.2 Main Theorem

Theorem 2.6.1 [Existence and Non-existence] Consider a linear system Ax = b, where A is a m×n matrix, and x, b are vectors with orders n×1, and m×1, respectively. Suppose rank (A) = r and rank([A b]) = r_a. Then exactly one of the following statement holds: 1. if $r_a = r < n$, the set of solutions of the linear system is an infinite set and has the form

$\{u_0 + k_1u_1 + k_2u_2 + \cdots + k_{n-r}u_{n-r} : k_i \in \mathbb{R}, 1 \le i \le n-r\}$, where $u_0, u_1, \ldots, u_{n-r}$ are $n \times 1$ vectors satisfying $Au_0 = b$ and $Au_i = 0$ for $1 \le i \le n-r$. 2. if $r_a = r = n$, the solution set of the linear system has a unique $n \times 1$ vector x_0 satisfying $Ax_0 = b$.

3. If $r < r_a$, the linear system has no solution. Remark 2.6.2 Let A be an m × n matrix and consider the linear system Ax = b. Then by Theorem 2.6.1, we see that the linear system Ax = b is consistent if and only if

where k_1, k_2 and k_3 are arbitrary.

0 −1 −1 0 −1 −1

. Then it can easily be verified that $Cu_0 = d$, and for $1 \le i \le 3$, $Cu_i = 0$. A similar idea is used in the proof of the next theorem and is omitted. The interested readers can read the proof in Appendix 14.1.

8
−2
1
−2

2. The first non-zero entry in the non-zero rows appear in columns 1, 2, 5 and 6.

3. Thus, the respective variables x_1, x_2, x_5 and x_6 are the basic variables.

Let u_0 = , u_1 = , u_2 = and u_3 =

4. The remaining variables, x_3, x_4 and x_7 are free variables.

5. We assign arbitrary constants k_1, k_2 and k_3 to the free variables x_3, x_4 and x_7, respectively. Hence, we have the set of solutions as

$2 + k_3$

Observations:

$8 - 2k_1 + k_2 - 2k_3$

x_1

$1 - k_1 - 3k_2 - 5k_3$

x_2

x_3

k_1

=

x_4

k_2

x_5

$4 - k_3$

x_6

system Ax = 0. Corollary 2.6.3 Let A be an m×n matrix. Then the homogeneous systemAx =0 has a non-trivial solution if and only if rank(A) < n. Proof. Suppose the system Ax = 0 has a non-trivial solution, x_0. That is, Ax_0 =0 and x_0 '=0. Under this assumption, we need to show that rank(A) < n. On the contrary, assume that rank(A) =n. So, n = rank(A) = rank [A 0] = r_a.

Also A0 = 0 implies that 0 is a solution of the linear system Ax = 0. Hence, by the uniqueness of the solution under the condition r = r_a = n (see Theorem 2.6.1), we get x_0 =0. A contradiction to the factthat x_0 was a given non-trivial solution. Now, let us assume that rank(A) < n. Then

r_a = rank [A 0] = rank(A) < n.

So, by Theorem 2.6.1, the solution set of the linear system Ax = 0 has infinite number of vectorsxsatisfying Ax = 0. From this infinite set, we can choose any vector x_0 that is different from0. Thus, we have a solution x_0 '= 0. That is, we have obtained a non-trivial solution x_0. We now state another important result whose proof is immediate from Theorem 2.6.1 and Corollary

rank (A) = rank([A b]). The following corollary of Theorem 2.6.1 is a very important result about the homogeneous linear

8 −2 −2 1

1 −1 −3 −5

0 1 0 0

=

+ k_1 + k_2 + k_3 ,

0 0 1 0

2 0 0 1

4 0 0

−1

0 0

0 0

1 2.6.3.

 Proposition 2.6.4 Consider the linear system Ax = b. Then the two statements given below cannot hold together. 1. The system Ax = b has a unique solution for every b.

2. The system Ax = 0 has a non-trivial solution. Remark 2.6.5 1. Suppose x_1, x_2 are two solutions of Ax = 0. Then $k_1x_1 + k_2x_2$ is also a solution of Ax = 0 for any $k_1, k_2 \in R$.

 Proof. Note that

$C = CI_n = C(AB) = (CA)B = I_nB = B$. Remark 2.7.3 1. From the above lemma, we observe that if a matrix Ais invertible, then the inverse is unique. 2. As the inverse of a matrix A is unique, we denote it by A^{-1}. That is, $AA^{-1} = A^{-1}A = I$. Theorem 2.7.4 Let A and B be two matrices with inverses $A^{-1}$and B^{-1}, respectively. Then

2. If u, v are two solutions of Ax = b then u − v is a solution of the system Ax = 0. That is, $u - v = x_h$ for some solution x_h of Ax = 0. That is, any two solutions of Ax = b differ by a solution of the associated homogeneous system Ax = 0. In conclusion, for b '= 0, the set of solutions of the system Ax = b is of the form, $\{x_0 + x_h\}$; where x_0 is a particular solution of Ax = b and x_h is a solution Ax = 0.

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

2.6.3 Exercises

Exercise 2.6.6 1. For what values of c and k-the following systems have i) no solution, ii) a unique solution and iii) infinite number of solutions.

(a) x + y + z = 3, x + 2y + cz = 4, 2x + 3y + 2cz = k. (b) x + y + z = 3, x + y + 2cz = 7, x + 2y + 3cz = k. (c) x + y + 2z = 3, x + 2y + cz = 5, x + 2y + 4z = k. (d) kx + y + z = 1, x + ky + z = 1, x + y + kz = 1.

(e) x + 2y − z = 1, 2x + 3y + kz = 3, x + ky + 3z = 2.
(f) x − 2y = 1, x − y + kz = 1, ky + 4z = 6.
2. Find the condition on a, b, c so that the linear system x + 2y − 3z = a, 2x + 6y − 11z = b, x − 2y + 7z = c is consistent.

Proof. Proof of Part 1. By definition $AA^{-1} = A^{-1}A = I$. Hence, if we denote $A^{-1}$by B, then we get AB=BA=I. This again by definition, implies $B^{-1} = A$, or equivalently $(A^{-1})^{-1} = A$. Proof of Part 2. Verify that $(AB)(B^{-1}A^{-1}) = I = (B^{-1}A^{-1})(AB)$. Hence, the result follows by definition. Proof of Part 3. We know $AA^{-1} = A^{-1}A = I$. Taking transpose, we get $(AA^{-1})^t = (A^{-1}A)^t = I^t \Longleftrightarrow (A^{-1})^tA^t = A^t(A^{-1})^t = I$. Hence, by definition $(A^t)^{-1} = (A^{-1})^t$. Exercise 2.7.5 1. If A is a symmetric matrix, is the matrix A^{-1}symmetric? 2. Show that every elementary matrix is invertible. Is the inverse of an elementary matrix, also an elementary matrix? 3. Let $A_1, A_2, \ldots, A_r$ be invertible matrices. Prove that the product $A_1A_2 \cdots A_r$ is also an invertible

3. Let A be an n × n matrix. If the system $A^2x = 0$ has a non trivial solution then show that Ax = 0 also has a non trivial solution.

2.7 Invertible Matrices

2.7.1 Inverse of a Matrix

Definition 2.7.1 (Inverse of a Matrix) Let A be a square matrix of order n. 1. A square matrix B is said to be a left inverse of A if $BA = I_n$. 2. A square matrix C is called a right inverse of A, if $AC = I_n$.

1. $(A^{-1})^{-1} = A$.
2. $(AB)^{-1} = B^{-1}A^{-1}$.

4. If P and Q are invertible matrices and P AQ is defined then show that rank (P AQ) =rank(A). 5. Find matrices P and Q which are product of elementary matrices such that B=P AQwhereA=2 4 8 1 3 2 and B = 1 0 0 0 1 0 . 6. Let A and B be two matrices. Show that (a) if A + B is defined, then rank(A + B) ≤ rank(A) + rank(B), (b) if AB is defined, then rank(AB) ≤ rank(A) and rank(AB) ≤ rank(B).

3. A matrix A is said to be invertible (or is said to have an inverse) if there exists a matrix B such that $AB = BA = I_n$. Lemma 2.7.2 Let A be an n × n matrix. Suppose that there exist n × n matrices B and C such that 2.7.

$B_1A = R_1R^2$ 0 0 0 , $AC_1 = S_1$ 0 S_3 0 0 , $B_2AC_2 = A_1$ 0 0

8. Let A be an m × n matrix of rank r. Then A can be written as A = BC, where both B and C have rank r and B is a matrix of size m × r and C is a matrix of size r × n. 9. Let A and B be two matrices such that AB is defined and rank (A) = rank (AB). Then show that

7. Let A be any matrix of rank r. Then show that there exists invertible matrices B_i, $C_i$such that 0 0 0 0 , and $B_3AC_3 = I_r$ 0 0 0 . Also, prove that the matrix A_1 is an r × r invertible matrix.

Suppose A is of full rank. This implies, the row reduced echelon form of Ahas all non-zero rows.

But A has as many columns as rows and therefore, the last row of the row reduced echelon form of Awill be (0, 0, . . . , 0, 1).

Hence, the row reduced echelon form of A is the identity matrix. 3 ⇒ 4

A = ABX for some matrix X. Similarly, if BA is defined and rank (A) = rank (BA), then A = Y BA

for some matrix Y. [Hint: Choose non-singular matrices P, Q and R such that P AQ = A_1 0 0 0 and

0 0 . Define X = R $C^{-1}A_1$ 0

0 0 Q^{-1}.]

10. Let A = [a_{ij}] be an invertible matrix and let B = [$p^{i-j}a_{ij}$] for some nonzero real number p. Find the inverse of B.

AB = I_n and CA = I_n, then B = C.

11. If matrices B and C are invertible and the involved partitioned products are defined, then show that A B

12. Suppose A is the inverse of a matrix B. Partition A and B as follows: A = A_{11} A 12 A_{21} A_{22} , B = B_{11} B 12 B_{21} B_{22} .

If A_{11} is invertible and P = A_{22} − A_{21}(A^{-1}

$_{11}$ A_{12}), then show that

$_{11}$ + (A^{-1}

$_{11}$), B_{21} = −P^{-1}($A_{21}A^{-1}$

$_{11}$), B_{12} = −(A^{-1}

2.7.2 Equivalent conditions for Invertibility

Definition 2.7.6 A square matrix A or order n is said to be of full rank if rank (A) = n.

Theorem 2.7.7 For a square matrix A of order n, the following statements are equivalent.

P(AB)R = C 0

1. A is invertible.
2. A is of full rank.
3. A is row-equivalent to the identity matrix.

$^{-1}$

Since A is row-equivalent to the identity matrix there exist elementary matrices E_1, E_2, . . . , $E_k$suchthat A = E_1E_2 · · ·E_kI_n. That is, A is product of elementary matrices. 4 ==⇒ 1

Suppose A = E_1E_2 · · ·E_k; where the E_i's are elementary matrices. We know that elementary matricesare invertible and product of invertible matrices is also invertible, we get the required result. The ideas of Theorem 2.7.7 will be used in the next subsection to find the inverse of an invertiblematrix. The idea used in the proof of the first part also gives the following important Theorem. Werepeat the proof for the sake of clarity. Theorem 2.7.8 Let A be a square matrix of order n. 1. Suppose there exists a matrix B such that AB = I_n. Then A^{-1}exists. 2. Suppose there exists a matrix C such that CA = I_n. Then A^{-1}exists. Proof. Suppose that AB = I_n. We will prove that the matrix A is of full rank. That is, rank (A) =n. Let if possible, rank(A) = r < n. Then there exists an invertible matrix P (a product of elementary matrices) such that P A = C_1 C 2 0 0 . Let B = B 1 B_2 , where B_1 is an r ×n matrix. ThenP = P I_n = P(AB) = (P A)B = C_1 C 2 0 0 B 1 B_2 = C_1B_1 +C_2B^2 0 . (2.7.2)Thus the matrix P has n − r rows as zero rows. So, P cannot be invertible. A contradiction toPbeing a product of invertible matrices. Thus, rank (A) = n. That is, A is of full rank. Hence, using Theorem2.7.7, A is an invertible matrix. That is, BA = I_n as well. Using the first part, it is clear that the matrix C in the second part, is invertible. Hence Thus, A is invertible as well. Remark 2.7.9 This theorem implies the following: "if we want to show that a square matrixAof order

0 C^{-1}

=

4. A is a product of elementary matrices.

.

B^{-1} −$B^{-1}$$AC^{-1}$

C 0

Proof. 1 ==⇒ 2 Let if possible rank(A) = r < n. Then there exists an invertible matrix P (a product of elementary matrices) such that P A = B_1 B^2 0 0 , where B_1 is an r×r matrix. Since A is invertible, let A^{-1} = C 1 C_2 , where

C_1 is an r × n matrix. Then

n is invertible, it is enough to show the existence of 1. either a matrix B such that AB = I_n 2. or a matrix C such that CA = I_n. Theorem 2.7.10 The following statements are equivalent for a square matrix Aof order n.

P = P I_n = P(AA^{-1}) = (P A)A^{-1} = B_1 B^2 0 0 C 1 C_2 = B_1C_1 + B_2C^2 0 . (2.7.1) Thus the matrix P has n − r rows as zero rows. Hence, P cannot be invertible. A contradiction to P

1. A is invertible.

B_{11} = A^{-1}

$_{11}$ A_{12})P^{-1}($A_{21}A^{-1}$

$_{11}$ A_{12})P^{-1},

and B_{22} = P^{-1}.

being a product of invertible matrices. Thus, A is of full rank. 2 ==⇒ 3

2. Ax = 0 has only the trivial solution x = 0. 3. Ax = b has a solution x for every b.

AC = I_n = CA.

2.7. INVERTIBLE MATRICES 39 Proof. 1 ==⇒ 2

40 CHAPTER 2. LINEAR SYSTEM OF EQUATIONS

2 1 1 1 0 0

1 1 2 0 0 1 ------→ R_1(1/2) 1 $^1_2$ $^1_2$ 0 0

Since A is invertible, by Theorem 2.7.7 A is of full rank. That is, for the linear system Ax = 0, the number of unknowns is equal to the rank of the matrix A. Hence, by Theorem 2.6.1 the system Ax = 0

1 2 1 0 1 0

1 2 1 0 1 0

has a unique solution x = 0. 2 ==⇒ 1 Let if possible A be non-invertible. Then by Theorem 2.7.7, the matrix A is not of full rank. Thus by Corollary 2.6.3, the linear system Ax = 0 has infinite number of solutions. This

contradicts the assumption that Ax = 0 has only the trivial solution x = 0. 1 ==⇒ 3

Since A is invertible, for every b, the system Ax = b has a unique solution x = A^{-1}b. 3 ==⇒ 1 For 1 ≤ i ≤ n, define e_i = (0, . . . , 0, 1

1 1 2 0 0 1

1 $^1_2$ $^1_2$ 0 0

1 $^1_2$ $^1_2$ 0 0

1 2 1 0 1 0

0 $^3_2$ $^1_2$ −$^1_2$ 1 0

1 1 2 0 0 1 ------→ R_{21}(−1) ------→ 0 $^1_2$ $^3_2$ −$^1_2$ 0 1

1 $^1_2$ $^1_2$ 0 0

0 $^1_2$ $^3_2$ −$^1_2$ 0 1 ------→ R_2(2/3) 1 $^1_2$ $^1_2$ 0 0 0 $^3_2$ $^1_2$ −$^1_2$ 1 0

0 1 $^1_3$ −$^1_3$ $^2_3$ 0

0 $^1_2$ $^3_2$ −$^1_2$ 0 1

1 $^1_2$ $^1_2$ 0 0

0 $^1_2$ $^3_2$ −$^1_2$ 0 1 --------→ R_{32}(−1/2) 1 $^1_2$ $^1_2$ 0 0

i^{th} position, 0, . . . , 0)^t, and consider the linear system Ax = e_i. By assumption, this system has a solution $x_i$for each i, 1 ≤ i ≤ n. Define a matrix B = [x_1, x_2, . . . , x_n]. That is, the ith column of B is the solution of the system Ax = e_i. Then

0 1 $^1_3$ −$^1_3$ $^2_3$ 0

0 1 $^1_3$ −$^1_3$ $^2_3$ 0

0 0 4/3 −1/3 −1/3 1

1 1/2 1/2 1/2 0 0

0 0 4/3 −1/3 −1/3 1 ------→ $R_3(3/4)$ 1 1/2 1/2 1/2 0 0

0 1 1/3 −1/3 2/3 0

AB = A[x1, x2 . . . , xn] = [Ax1, Ax2 . . . , Axn] = [e1, e2 . . . , en] = In. Therefore, by Theorem 2.7.8, the matrix A is invertible.

Exercise 2.7.11 1. Show that a triangular matrix A is invertible if and only if each diagonal entry of A

0 1 1/3 −1/3 2/3 0 0 0 1 −1/4 −1/4 3/4

is non-zero.

1 1/2 1/2 1/2 0 0

0 1 1/3 −1/3 2/3 0

0 1 0 −1/4 3/4 −1/4

0 0 1 −1/4 −1/4 3/4 -------→ 0 0 1 −1/4 −1/4 3/4

0 1 0

invertible? Give reasons.

1.

3. Let A be an n × m matrix and B be an m × n matrix. Prove that the matrix I − BA is invertible if and only if the matrix I − AB is invertible.

We first give a consequence of Theorem 2.7.3 Inverse and Gauss-Jordan Method

2.7.7 and then use it to find the inverse of an invertible matrix.

Corollary 2.7.12 Let A be an invertible n×n matrix. Suppose that a sequence of elementary row-operations reduces A to the identity matrix. Then the same sequence of elementary row-operations when applied to the identity matrix yields A^{-1}.

40,7

0 0 1 −1/4 −1/4 3/4 -------→ 0 0 1 −1/4 −1/4 3/4.

Proof. Let A be a square matrix of order n. Also, let $E_1, E_2, \ldots, E_k$ be a sequence of elementary row operations such that $E_1E_2 \cdots E_kA = I_n$. Then $E_1E_2 \cdots E_kI_n = A^{-1}$. This implies $A^{-1} = E_1E_2 \cdots E_k$.

Notation: For an n × n matrix A, by A(α|β), we mean the submatrix Bof A, which is obtained by deleting the αth row and β^{th} column.

Example 2.8.1 Consider a matrix A = 1 2 3

5.

Summary: Let A be an n × n matrix. Apply the Gauss-Jordan method to the matrix $[A\ I_n]$. Suppose the row reduced echelon form A is not invertible.

Example 2.7.13 Find the inverse of the matrix 2 1 1

1 1 2 using Gauss-Jordan method. 1 2 1

A(1, 2|1, 3) = [4]. Definition 2.8.2 (Determinant of a Square Matrix) Let A be a square matrix of order n. WithA, we associate inductively (on n) a number, called the determinant of A, written det(A) (or |A|) by

Solution: Consider the matrix 2 1 1 1 0 0

1 2 1 0 1 0

7.

1 1 2 0 0 1 . A sequence of steps in the Gauss-Jordan method

Definition 2.8.3 (Minor, Cofactor of a Matrix) The number det (A(i|j)) is called the $(i, j)^{th}$ minor of A. We write A_{ij} = det (A(i|j)). The $(i, j)^{th}$ cofactor of A, denoted C_{ij}, is the number $(-1)^{i+j}A_{ij}$.

Example 2.8.4 1. Let A = a11 a12

det(A) = a if A = [a] (n = 1),

(i)

$R_{23}(-1/3)$ -------→ $R_{13}(-1/2)$ 1 1/2 0 5/8 1/8 −3/8

$R_{12}(-1/2)$

1 1/2 0 5/8 1/8 −3/8

turns out that theway we have defined determinant is usually called the expansion of the determinant alongthe first row. 2. Part 1 of Lemma 2.8.7 implies that "one can also calculate the determinant by expanding along any row." Hence, for an n × n matrix A, for every k, 1 ≤ k ≤ n, one also has

. Then, det(A) = |A| = a11A11 − a12A12 = a11a22 − a12a21. For example, for A = 1 2 2 1 det(A) = |A| = 1 − 2 · 2 = −3.

det(A) = α^n

Remark 2.8.8 1. Many authors define the determinant using "Permutations." It

2. Let A = a11 a12 a13

are:

det(A) = |A| = a11A11 − a12A12 + a13A13

= a11(a22a33 − a23a32) − a12(a21a33 − a31a23) + a13(a21a32 − a31a22) = a11a22a33 − a11a23a32 − a12a21a33 + a12a23a31 + a13a21a32 − a13a22a31 (2.8.1)

For example, if A = 1 2 3 2 3 1

1 2 2 then

det(A) = |A| = 1 · 3 1

2 2 − 2 · 2 1

1 2 + 3 · 2 3

= 4 − 2(3) + 3(1) = 1.

1 2

Exercise 2.8.5 1. Find the determinant of the following matrices.

0 0 0 5, ii) 3 5 2 1

1 2 7 8

a21 a22

2 0 3 0, iii) 1 a a2

0 4 3 2

0 2 0 5

1 b b2

0 0 2 3

6 −7 1 0

1 c c2.

2. Show that the determinant of a triangular matrix is the product of its diagonal entries.

Definition 2.8.6 A matrix A is said to be a singular matrix if det(A) = 0. It is called non-singular if

. Then,

a21 a22 a23

a31 a32 a33

det(A) '= 0.

The proof of the next theorem is omitted. The interested reader is advised to go through Appendix

= a11 a22 a23

− a12 a21 a23

2. Let A be a 1 × 2 matrix and B be a 2 × 1 matrix having positive entries. Which of BA or AB is

1 0 0 3/4 −1/4 −1/4

0 1 0 −1/4 3/4 −1/4

8. Thus, the inverse of the given matrix is 3/4 −1/4 −1/4

2.

−1/4 −1/4 3/4. Exercise 2.7.14 Find the inverse of the following matrices using Gauss-Jordan method.

$R_{31}(-1)$

1 2 3

2 4 7, (ii) 1 3 3

3.

1 3 2

2 3 2

2 4 7, (iii) 2 −1 3

−1 3 −2

2 4 1.

2.8 Determinant

4.

of the matrix $[A\ I_n]$ is [B C]. If B = In, then A^{-1} = C or else

2 4 7. Then A(1|2) = 1 2

2 7 , A(1|3) = 1 3

2 4 , and

1 3 2

6.

A(1, 2|1, 3) = [4].

+ a_{13} " a_{21} a_{22}

a_{32} a_{33}

a_{31} a_{33}

a_{31} a_{32}

14.3. Theorem 2.8.7 Let A be an n × n matrix. Then

1. if B is obtained from A by interchanging two rows, then det(B) = − det(A), 2. if B is obtained from A by multiplying a row by c then det(B) = c det(A), 3. if all the elements of one row or column are 0 then det(A) = 0,

$(-1)^{k+j}a_{kj}$ det A(k|j) . Remark 2.8.9 1. Let $u^t = (u_1, u_2)$ and $v^t = (v_1, v_2)$ be two vectors in R^2. Then consider the par- allelogram, PQRS, formed by the vertices {P = $(0, 0)^t$, Q = u, S =v, R=u +v}. We

$(-1)^{1+j}a_{1j}$ det A(1|j) , otherwise.

j=1

Claim: Area (P QRS) = " det u_1 v 1 u_2 v_2 " $=|u_1v_2 - u_2v_1|$. Recall that the dot product, u • v = $u_1v_1 + u_2v_2$, and √ u • u = $(u^2_1 + u^2_2)$, is the length of thevector u. We denote the length by H(u). With the above notation, if θ is

the angle between the cos(θ) = u • v

H(u)H(v) . Which tells us,

Area(PQRS) = H(u)H(v) sin(θ) = H(u)H(v) (1 − / u • v = $|u_1v_2 - u_2v_1|$. Hence, the claim holds. That is, in R^2, the determinant is ± times the area of the parallelogram. 2. Let u = (u_1, u_2, u_3), v = (v_1, v_2, v_3) and w = (w_1, w_2, w_3)

be three elements of R^3. Recall that the cross product of two vectors in $R^3$is,

j=1

u × v = $(u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)$. Note here that if A = $[u^t, v^t, w^t]$, then det(A) = " u_1 v_1 w 1 u_2 v_2 w_2

= u • (v × w) = v • (w × u) =w• (u×v).

Let P be the parallelopiped formed with (0, 0, 0) as a vertex and the vectors u, v, was adjacent vertices. Then observe that u × v is a vector perpendicular to the plane that contains the paral lelogram formed by the

vectors u and v. So, to compute the volume of the parallelopipedP, weneed to look at cos(θ), where θ is the angle between the vector wand the normal vector to the parallelogram formed by u and v. So,

vectors u and v, then

volume (P) = |w • (u × v)|.

3_2

H(u)H(v)

4. if B is obtained from A by replacing the jth row by itself plus k times the ith row, where i '= j then

= $H(u)^2 + H(v)^2 - (u \cdot v)^2$ = $(u_1v_2 - u_2v_1)^2$

det(B) = det(A),

□

5. if A is a square matrix having two rows equal then det(A) = 0.

□

i)

□

3. Let $u_1, u_2, \ldots, u_n \in R^{n\times 1}$and let A = $[u_1, u_2, \ldots, u_n]$ be an n ×n matrix. Then the following properties of det(A) also hold for the volume of an n-dimensional parallelopiped formed with0 ∈ $R^{n\times 1}$as one vertex and the

vectors $u_1, u_2, \ldots, u_n$ as adjacent vertices:

"

u_3 v_3 w_3

Hence, | det(A)| = volume (P).

2.8. DETERMINANT 43 (a) If $u_1 = (1, 0, \ldots, 0)^t$, $u_2 = (0, 1, 0, \ldots, 0)^t$, . . . , and $u_n = (0, \ldots, 0, 1)^t$, then det(A) = 1. Also, volume of a unit n-dimensional cube is 1. (b) If we replace the vector u_i by αu_i, for some α ∈ R,

then the determinant of the new matrix is α · det(A). This is also true for the volume, as the original volume gets multiplied by α.

44 CHAPTER 2. LINEAR SYSTEM OF EQUATIONS= 0 if i '= j det(A) if i = j

(c) If $u_1 = u_i$for some i, 2 ≤ i ≤ n, then the vectors $u_1, u_2, \ldots, u_n$ will give rise to an (n − 1)- dimensional parallelopiped. So, this parallelopiped lies on an (n−1)-dimensional hyperplane. Thus, its n-dimensional volume will be zero. Also, | det(A)| = |0| = 0. In general, for any n × n matrix A, it can be proved that | det(A)| is indeed equal to the volume of the n-dimensional parallelepiped. The actual proof is beyond the scope of this book.

Thus, A(Adj(A)) = det(A)I_n. Since, det(A) '= 0, A 1 det(A)Adj(A) =I_n. Therefore, Ahas a right inverse. Hence, by Theorem 2.7.8 A has an inverse and

2.8.1 Adjoint of a Matrix Recall that for a square matrix A, the notations A_{ij} and $C_{ij} = (-1)^{i+j}A_{ij}$ were respectively used to denote the (i, j) th minor and the (i, j) th cofactor of A. Definition 2.8.10 (Adjoint of a Matrix) Let A

be an n × n matrix. The matrix B = $[b_{ij}]$ with $b_{ij} = C_{ji}$,

A^{-1} = 1 det(A) Adj(A). Example 2.8.13 Let A = □□ 1 −1 0

0 1 1

for 1 ≤ i, j ≤ n is called the Adjoint of A, denoted Adj(A). Example 2.8.11 Let A = □□ □1 Adj(A) = □□ □−1 1 −1

2 3 1 2 2 □□ □. Then Adj(A) = □□ □ 4 2 −7 1 1 −1

−3 −1 5 −1 −3 1 □□ □

2 3 1

Now,

1 2 1 □□ □ . Then

1 0 −1 □□ □ ;

as $C_{11} = (-1)^{1+1}A_{11} = 4$, $C_{12} = (-1)^{1+2}A_{12} = -3$, $C_{13} = (-1)^{1+3}A_{13}$ = 1, and so on. Theorem 2.8.12 Let A be an n × n matrix. Then

A Adj(A) 3

ij = *n

and det(A) = −2. By Theorem 2.8.12.3, A^{-1} = □□ □ 1/2 −1/2 1/2 −1/2 −1/2 1/2

a_{ik} Adj(A) kj = *n

/

$a_{ik}C_{jk}$

det(B) = 0. By construction again, det A(H|j) = det B(H|j) for 1 ≤ j ≤ n. Thus, by Remark 2.8.8, we have 0 = det(B) = *n

1. for 1 ≤ i ≤ n, !n

a_{ij} $(-1)^{i+j}$ A_{ij} = det(A),

2. for i '= H, !n

a_{ij} $(-1)^{0+j}$ A_{0j} = 0, and

3. A(Adj(A)) = det(A)I_n. Thus,

det(A) '= 0 ⇒ A^{-1} = 1 det(A)Adj(A). (2.8.2)

Proof. Let B = $[b_{ij}]$ be a square matrix with

• the H th row of B as the i th row of A,

• the other rows of B are the same as that of A.

By the construction of B, two rows (i th and H th) are equal. By Part 5 of Lemma 2.8.7,

j=1 a_{ij} C_{ij} = !n

j=1

j=1 a_{ij} C_{0j} = !n

j=1

2.8. DETERMINANT 45 Step 2. Suppose det(A) = 0. Then A is not invertible. Hence, there exists an invertible matrix P such that P A =

C, where C = $\begin{pmatrix} 1 & \\ 0 & \end{pmatrix}$. So, A = P^{-1}C, and therefore

det(AB) = det((P $^{-1}$C)B) = det(P $^{-1}$(CB)) = det $\begin{pmatrix} P^{-1} & C_1B \end{pmatrix}$

$(-1)^{0+j}b_{0j}$ det B(H|j) = $\sum_{j=1}^{n}$

= det(P $^{-1}$) · det $\begin{pmatrix} C_1B \end{pmatrix}$

as P $^{-1}$is non-singular

$(-1)^{0+j}a_{1j}$ det B(H|j)

= det(P) · 0 = 0 = 0 · det(B) = det(A) det(B). Thus, the proof of the theorem is complete. Corollary 2.8.16 Let A be a square matrix.

Then A is non-singular if and only if A has an inverse.

= $\sum_{j=1}^{n}$

$(-1)^{0+j}a_{1j}$ det A(H|j) = $\sum_{j=1}^{n}$

$a_{1j}C_{0j}$.

1/2 3/2 −1/2 $\begin{pmatrix} \square & \square & \square \end{pmatrix}$. The next corollary is an easy consequence of Theorem 2.8.12 (recall Theorem 2.7.8).

Corollary 2.8.14 If A is a non-singular matrix, then

Adj(A) A = det(A)I_n and I^n

0 if j '= k. Theorem 2.8.15 Let A and B be square matrices of order n. Then det(AB) = det(A) det(B). Proof. Step 1. Let

det(A) '= 0. This means, A is invertible. Therefore, either A is an elementary matrix or is a product of

elementary matrices (see Theorem 2.7.7). So, let $E_1, E_2, \ldots, E_k$ be elementary matrices such that A=$E_1E_2 \cdots E_k$. Then,

det(A) det(B) = det(AB) = det(I) = 1. This implies that det(A) '= 0. Thus, A is non-singular.

Proof. If A is a non-singular Corollary 2.8.14 gives det(A) = det(A^t).

If A is singular, then det(A) = 0. Hence, by Corollary 2.8.16, A doesn't have an inverse. Therefore, A^t also doesn't have an inverse (for if A^t has an inverse then A^{-1} = $(A^t)^{-1\,t}$). Thus again by Corollary 2.8.16, det(A^t) = 0.

Therefore, we again have det(A) = 0 = det(A^t). Hence, we have det(A) = det(A^t).

0

0

2.8.2 Cramer's Rule

Recall the following:

• The linear system Ax = b has a unique solution for every b if and only if A^{-1}exists. • A has an inverse if and only if det(A) '= 0.

for j = 2, 3, . . . , n.

Example 2.8.19 Suppose that A = $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 2 \\ 2 & 3 & 1 \end{pmatrix}$ and b = $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Use Cramer's rule to find a vector xsuch

Solution: Check that det(A) = 1. Therefore x_1 = $\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 2 \\ 1 & 3 & 1 \end{vmatrix}$ = −1,

$\begin{vmatrix} 1 & 1 & 3 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{vmatrix}$ = 1, and x_3 = $\begin{vmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 2 & 3 & 1 \end{vmatrix}$ = 0. That is, x^t = (−1, 1, 0).

x_1 = $\frac{1}{\det(A)}$

2.9 Miscellaneous Exercises

Exercise 2.9.1 1. Let A be an orthogonal matrix. Show that det A = ±1.

2. If A and B are two n × n non-singular matrices, are the matrices A +Band A−Bnon-singular?Justify your answer. 3. For an n × n matrix A, prove that the following conditions are equivalent: x_j = $\frac{1}{\det(A)}$

Thus, Ax = b has a unique solution for every b if and only if det(A) '= 0. The following theorem gives a direct method of system with n equations in n unknowns. If

finding the solution of the linear system Ax = b when det(A) '= 0. Theorem 2.8.18 (Cramer's Rule) Let Ax = b be a linear

det(A) '= 0, then the unique solution to this system is

det(A), for j = 1, 2, . . . , n,

where A_j is the matrix obtained from A by replacing the jth column of A by the column vector b. x_j =det(A_j)

(a) A is singular (A^{-1}doesn't exist).

(b) rank(A) '= n.

(c) det(A) = 0.

(d) A is not row-equivalent to I_n, the identity matrix of order n. (e) Ax = 0 has a non-trivial solution for x.

that Ax = b.

(f) Ax = b doesn't have a unique solution, i.e., it has no solutions or it has infinitely many solutions.x_2 =

2 0 6 0 4

5 3 2 2 7

4. Let A = $\begin{pmatrix} \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \end{pmatrix}$. We know that the numbers 20604, 53227, 25755, 20927 and 78421 are 2 5 7 5 5

2 0 9 2 7

all divisible by 17. Does this imply 17 divides det(A)?

5. Let A = $[a_{ij}]_{n\times n}$ where a_{ij} = x^{i-1}

^{i}. Show that det(A) = $\prod_{i>j}$

$(x_j - x_i)$. [The matrix A is usually

called the Van-dermonde matrix.] 6. Let A = $[a_{ij}]$ with a_{ij} = max{i, j} be an n × n matrix. Compute det A.

7. Let A = $[a_{ij}]$ with a_{ij} = 1/(i + j) be an n × n matrix. Show that A is invertible. 8. Solve the following system of equations by Cramer's rule.

i) x + y + z − w = 1, x + y − z + w = 2, 2x + y + z − w = 7, x + y + z + w = 3. ii) x − y + z − w = 1, x + y − z + w = 2, 2x + y − z − w = 7, x − y − z + w = 3. 9. Suppose A = $[a_{ij}]$ and B = $[b_{ij}]$ are two n × n matrices such that b_{ij}

= $p^{i-j}a_{ij}$ for 1 ≤ i, j ≤ n for some non-zero real number p. Then compute det(B) in terms of det(A). 10. The position of an element a_{ij} of a determinant is called even or odd according as i + j is even or odd. Show that (a) If

by using Parts 1, 2 and 4 of Lemma 2.8.7 repeatedly, we get det(AB) = det($E_1E_2 \cdots E_kB$) = det(E_1) det($E_2 \cdots E_kB$)

= det(E_1) det(E_2) det($E_3 \cdots E_kB$)

= det(E_1E_2) det($E_3 \cdots E_kB$)

= det($E_1E_2 \cdots E_k$) det(B)

= det(A) det(B). Thus, we get the required result in case A is non-singular.

$\sum_{i=1}^{n} a_{ij}C_{ik}$ = det(A) if j = k

Proof. Since det(A) '= 0, A^{-1} = $\frac{1}{\det(A)}$Adj(A). Thus, the linear systemAx

=bhas the solution x = $\frac{1}{\det(A)}$Adj(A)b. Hence, x_j , the jth coordinate of x is given by x_j =$b_1C_{1j} + b_2C_{2j} + \cdots + b_nC_{nj}$ det(A)

=det(A_j)

= $\begin{vmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{vmatrix}$.

det(A). The theorem implies that

$\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}$, and in general

Proof. Suppose A is non-singular. Then det(A) '= 0 and therefore, A^{-1} = $\frac{1}{\det(A)}$Adj(A). Thus, A has an inverse.

Suppose A has an inverse. Then there exists a matrix B such that AB = I = BA. Taking determinant of both sides, we get

$\begin{vmatrix} a_{11} & \cdots & a_{1j-1} & b_1 & a_{1j+1} & \cdots & a_{1n} \\ a_{12} & \cdots & a_{2j-1} & b_2 & a_{2j+1} & \cdots & a_{2n} \\ a_{1n} & \cdots & a_{nj-1} & b_n & a_{nj+1} & \cdots & a_{nn} \end{vmatrix}$

Theorem 2.8.17 Let A be a square matrix. Then det(A) = det(A^t).

all the entries in odd positions are multiplied with −1 then the value of the determinant doesn't change. (b) If all entries in even positions are multiplied with −1 then the determinant

i. does not change if the matrix is of even order.

ii. is multiplied by −1 if the matrix is of odd order. 11. Let A be an n×n Hermitian matrix, that is, $A^* = A$. Show that det A is a real number. [A is a matrix with complex entries and $A^* = A^t$.]

12. Let A be an n × n matrix. Then show that
A is invertible ⇐⇒ Adj(A) is invertible.

13. Let A and B be invertible matrices. Prove that Adj(AB) = Adj(B)Adj(A). 14. Let P = $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ be a rectangular matrix with A a square matrix of order n and |A| '= 0. Then show that rank (P) = n if and only if $D = CA^{-1}B$.

Chapter 3

(called the additive⊕ is called vector addition.

2. Scalar Multiplication: For each u ∈ V and α ∈ F, there corresponds a unique element α$uinV such that

(a) α · (β $ u) = (αβ) $ u for every α, β ∈ F and u ∈ V. (b) 1 $ u = u for every u ∈ V, where 1 ∈ R.

3. Distributive Laws: relating vector addition with scalar multiplication. For any α, β ∈ F and u, v ∈ V, the following distributive laws hold:

(d) For every u ∈ V there is a unique element −u ∈ V such that u⊕(−u) =0

Finite Dimensional Vector Spaces

Consider the problem of finding the set of points of intersection of the two planes 2x + 3y + z + u = 0 and 3x + y + 2z + u = 0.
Let V be the set of points of intersection of the two planes. Then V has the following properties:

1. The point (0, 0, 0, 0) is an element of V.
2. For the points (−1, 0, 1, 1) and (−5, 1, 7, 0) which belong to V ; the point (−6, 1, 8, 1) = (−1, 0, 1, 1)+ (−5, 1, 7, 0) ∈ V.
3. Let α ∈ R. Then the point α(−1, 0, 1, 1) = (−α, 0, α, α) also belongs to V.

Similarly, for an m × n real matrix A, consider the set V, of solutions of the homogeneous linear system Ax = 0. This set satisfies the following properties:

1. If Ax = 0 and Ay = 0, then x, y ∈ V. Then x + y ∈ V as A(x + y) = Ax + Ay = 0 + 0 = 0. Also,

(a) α $ (u ⊕ v) = (α $ u) ⊕ (α $ v).
(b) (α + β) $ u = (α $ u) ⊕ (β $ u).

Note: the number 0 is the element of F whereas 0 is the zero vector. Remark 3.1.2 The elements of F are called scalars, and that of V are called vectors. If F=R, the vector space is called a real vector space. If F = C, the vector space is called a complex vector space. We may sometimes write V for a vector space if F is understood from the context. Some interesting consequences of Definition 3.1.1 is the following useful result. Intuitively, these results seem to be obvious but for better understanding of the axioms it is desirable to go through the inverse).

2. It is clear that if x, y, z ∈ V then (x + y) + z = x + (y + z). 3. The vector 0 ∈ V as A0 = 0.

4. If Ax = 0 then A(−x) = −Ax = 0. Hence, −x ∈ V.
5. Let α ∈ R and x ∈ V. Then αx ∈ V as A(αx) = αAx = 0.

Thus we are lead to the following.

3.1 Vector Spaces

3.1.1 Definition

Definition 3.1.1 (Vector Space) A vector space over F, denoted V (F), is a non-empty set, satisfying the following axioms:

(−5, 1, 7, 0) ∈ V.
1. Vector Addition: To every pair u, v ∈ V there corresponds a unique element u⊕ v in V such that (a) u ⊕ v = v ⊕ u (Commutative law).

(b) (u ⊕ v) ⊕ w = u ⊕ (v ⊕ w) (Associative law).
(c) There is a unique element 0 in V (the zero vector) such that u ⊕ 0 = u, for every u ∈ V (called the additive identity).

x + y = y + x.

proof. Theorem 3.1.3 Let V be a vector space over F. Then
2. α $ u = 0 if and only if either u is the zero vector or α = 0. 3. (−1) $ u = −u for every u ∈ V.

Proof. Proof of Part 1.
For u ∈ V, by Axiom 1d there exists −u ∈ V such that −u ⊕ u = 0. Hence, u ⊕ v = u is equivalent to

−u ⊕ (u ⊕ v) = −u ⊕ u ⇐⇒ (−u ⊕ u) ⊕ v = 0 ⇐⇒0 ⊕ v =0 ⇐⇒v =0. Proof of Part 2. As 0 = 0 ⊕ 0, using the distributive law, we have

α $ 0 = α $ (0 ⊕ 0) = (α $ 0) ⊕ (α $ 0). Thus, for any α ∈ F, the first part implies α $ 0 = 0. In the same way, 0 $ u = (0 + 0) $ u = (0 $ u) ⊕ (0 $ u). Hence, using the first part, one has 0 $ u = 0 for any u ∈ V. Now suppose

α $ u = 0. If α = 0 then the proof is over. Therefore, let us assume α'= 0 (note that α is a real or complex number, hence $\frac{1}{\alpha}$ exists and

0 = $\frac{1}{\alpha}$ $ 0 = $\frac{1}{\alpha}$ $ (α $ u) = ($\frac{1}{\alpha}$α) $ u = 1 $ u =u1. u ⊕ v = u implies v = 0.

as 1 $ u = u for every vector u ∈ V. Thus we have shown that if α '= 0 and α $ u = 0 then u = 0. Proof of Part 3. We have 0 = 0u = (1 + (−1))u = u + (−1)u and hence (−1)u = −u.

7. Consider the set $C^n = \{(z_1, z_2, \ldots, z_n) : z_i \in C \text{ for } 1 \le i \le n\}$. For $(z_1, \ldots, z_n),(w_1, \ldots, w_n) \in C^n$ $(z_1, \ldots, z_n) \oplus (w_1, \ldots, w_n) = (z_1 + w_1, \ldots, z_n + w_n)$ and α $ $(z_1, \ldots, z_n) = (\alpha z_1, \ldots, \alpha z_n)$.

(a) If the set F is the set C of complex numbers, then C^n is a complex vector space havingn-tupleof complex numbers as its vectors. (b) If the set F is the set R of real numbers, then $C^n$is a real vector space having n-tuple of complexnumbers as its vectors. Remark 3.1.5 In Example 7a, the scalars are Complex numbers and hence i(1, 0) = (i, 0). Whereas, in Example 7b, the scalars are Real Numbers and hence we cannot writei(1, 0) =

3.1.2 Examples

Example 3.1.4 1. The set R of real numbers, with the usual addition and multiplication (i.e., $\oplus \equiv +$ and $\$ \equiv \cdot$) forms a vector space over R.

2. Consider the set $R^2 = \{(x_1, x_2) : x_1, x_2 \in R\}$. For $x_1, x_2, y_1, y_2 \in R$ and $\alpha \in R$, define, $(x_1, x_2) \oplus (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ and $\alpha \$ (x_1, x_2) = (\alpha x_1, \alpha x_2)$.

Then $R^2$is a real vector space.

3. Let $R^n = \{(a_1, a_2, \ldots, a_n) : a_i \in R, 1 \le i \le n\}$, be the set of n-tuples of real numbers. For $u = (a_1, \ldots, a_n)$, $v = (b_1, \ldots, b_n)$ in V and $\alpha \in R$, we define

$u \oplus v = (a_1 + b_1, \ldots, a_n + b_n)$ and $\alpha \$ u = (\alpha a_1, \ldots, \alpha a_n)$

(called component wise or coordinate wise operations). Then V is a real vector space with addition and scalar multiplication defined as above. This vector space is denoted by R^n, called the real vector space of n-tuples.

4. Let $V = R^+$ (the set of positive real numbers). This is not a vector space under usual operations of addition and scalar multiplication (why?). We now define a new vector addition and scalar multiplication for all $v_1, v_2, v \in R^+$ and $\alpha \in R$. Then R^+ is a real vector space with 1 as the additive identity.

8. Fix a positive integer n and let $M_n(R)$ denote the set of all n × n matrices with real entries. Then$M_n(R)$ is a real vector space with vector addition and scalar multiplication defined by $A \oplus B = [a_{ij}] \oplus [b_{ij}] = [a_{ij} + b_{ij}]$, $\alpha \$ A = \alpha \$[a_{ij}] = [\alpha a_{ij}]$. 9. Fix a positive integer n. Consider the set, $P_n(R)$, of all polynomials of degree ≤nwith coefficientsfrom R in the indeterminate x. Algebraically,

$P_n(R) = \{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n : a_i \in R, 0 \le i \le n\}$.
and $\alpha \in F$, define,

Let $f(x), g(x) \in P_n(R)$. Then $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$and $g(x) = b_0+b_1x+b_2x^2+\cdots + b_nx^n$ for some $a_i, b_i \in R$, $0 \le i \le n$. It can be verified that $P_n(R)$ is a real vector space with the addition and scalar multiplication defined by: $f(x) \oplus g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$, and

5. Let $V = R^2$. Define $(x_1, x_2) \oplus (y_1, y_2) = (x_1 + y_1 + 1, x_2 + y_2 - 3)$, $\alpha \$ (x_1, x_2) = (\alpha x_1 + \alpha - 1, \alpha x_2 - 3\alpha + 3)$ for $(x_1, x_2),(y_1, y_2) \in R^2$and $\alpha \in R$. Then it can be easily verified that the vector (−1, 3) is the additive identity and V is indeed a real vector space.

Recall $\sqrt{-1}$ is denoted i.

6. Consider the set $C = \{x + iy : x, y \in R\}$ of complex numbers. (a) For $x_1 + iy_1, x_2 + iy_2 \in C$ and $\alpha \in R$, define,

$$(x_1 + iy_1) \oplus (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2) \text{ and}$$
$$\alpha \$ (x_1 + iy_1) = (\alpha x_1) + i(\alpha y_1).$$

Then C is a real vector space.

(b) For $x_1 + iy_1, x_2 + iy_2 \in C$ and $\alpha + i\beta \in C$, define,

$$(x_1 + iy_1) \oplus (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2) \text{ and}$$
$$(\alpha + i\beta) \$ (x_1 + iy_1) = (\alpha x_1 - \beta y_1) + i(\alpha y_1 + \beta x_1).$$

as

Then C forms a complex vector space.

$$v_1 \oplus v_2 = v_1 \cdot v_2 \text{ and } \alpha \$ v = v^\alpha$$

real vector space. The readers are advised to justify the statements made in the above examples. From now on, we will use 'u + v' in place of 'u ⊕ v' and 'α · u or αu' in place of 'α $ u'.

10. Consider the set P(R), of all polynomials with real coefficients. Let f(x), g(x) ∈ P(R). Observe that a polynomial of the form $a_0 + a_1x + \cdots + a_mx^m$can be written as $a_0 + a_1x + \cdots + a_mx^m + 0 \cdot x^{m+1} + \cdots + 0 \cdot x^p$for any p > m. Hence, we can assume $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_px^pandg(x) = b_0 + b_1x + b_2x^2 + \cdots + b_px^p$for some $a_i, b_i \in R$, $0 \le i \le p$, for some large positive integer p. We now define the vector addition and scalar multiplication as

(i, 0).

$f(x) \oplus g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_p + b_p)x^p$, and $\alpha \$ f(x) = \alpha a_0 + \alpha a_1x + \cdots + \alpha a_px^p$for $\alpha \in R$. Then P(R) forms a real vector space. 11. Let C([−1, 1]) be the set of all real valued continuous functions on the interval [−1, 1].

For f, g∈$(f \oplus g)(x) = f(x) + g(x)$, and

$(\alpha \$ f)(x) = \alpha f(x)$, for all $x \in [-1, 1]$.

Then C([−1, 1]) forms a real vector space. The operations defined above are called point wise addition and scalar multiplication.

12. Let V and W be real vector spaces with binary operations (+, •) and (⊕, $), respectively. Consider the following operations on the set V × W : for $(x_1, y_1),(x_2, y_2) \in V \times W$ and $\alpha \in R$, define $(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2, y_1 \oplus y_2)$, and $\alpha \circ (x_1, y_1) = (\alpha \bullet x_1, \alpha \$ y_1)$. On the right hand side, we write $x_1 + x_2$ to mean the addition in V, while $y_1 \oplus y_2$ is the addition in W. Similarly, $\alpha \bullet x_1$ and $\alpha \$ y_1$ come from scalar multiplication in V and W, respectively. With the above definitions, V × W also forms a

(e) $\{(x_1, x_2, \ldots, x_n) :$ either x_1 or x_2 or both is0$\}$.
(f) $\{(x_1, x_2, \ldots, x_n) : |x_1| \le 1\}$.

3. Which of the following are subspaces of i)$C^n(R)$ ii)$C^n(C)$?
(a) $\{(z_1, z_2, \ldots, z_n) : z_1$is real $\}$.
(b) $\{(z_1, z_2, \ldots, z_n) : z_1 + z_2 = z_3\}$.
C([−1, 1]) and $\alpha \in R$, define
(c) $\{(z_1, z_2, \ldots, z_n) : |z_1| = |z_2|\}$.

3.1.3 Subspaces

Definition 3.1.6 (Vector Subspace) Let S be a non-empty subset of V. S(F) is said to be a subspace of V (F) if $\alpha u+\beta v \in S$ whenever $\alpha, \beta \in F$ and $u, v \in S$; where the vector addition and scalar multiplication are the same as that of V (F). Remark 3.1.7 Any subspace is a vector space in its own right with respect to the vector addition and scalar multiplication that is defined for V (F). Example 3.1.8 1. Let V (F) be a vector space.

Then

(a) S = {0}, the set consisting of the zero vector 0,
are vector subspaces of V. These are called trivial subspaces.
2. Let $S = \{(x, y, z) \in R^3 : x + y - z = 0\}$. Then S is a subspace of R^3. (S is a plane in R^3passing through the origin.)

3. Let $S = \{(x, y, z) \in R^3 : x + y + z = 3\}$. Then S is not a subspace of R^3. (S is again a plane in R^3 but it doesn't pass through the origin.)

4. Let $S = \{(x, y, z) \in R^3 : z = x\}$. Then S is a subspace of R^3.
5. The vector space $P_n(R)$ is a subspace of the vector space P(R). Exercise 3.1.9 1. Which of the following are correct statements?
(a) Let $S = \{(x, y, z) \in R^3 : z = x^2\}$. Then S is a subspace of R^3.
(b) Let V (F) be a vector space. Let $x \in V$. Then the set $\{\alpha x : \alpha \in F\}$ forms a vector subspace of V. (c) Let $W = \{f \in C([-1, 1]) : f(1/2) = 0\}$. Then W is a subspace of the real vector space, 2. Which of the following are subspaces of $R^n(R)$?

(a) $\{(x_1, x_2, \ldots, x_n) : x_1 \ge 0\}$.
(b) $\{(x_1, x_2, \ldots, x_n) : x_1 + 2x_2 = 4x_3\}$.
(b) S = V
(c) $\{(x_1, x_2, \ldots, x_n) : x_1$is rational $\}$.
(d) $\{(x_1, x_2, \ldots, x_n) : x_1 = x^2_3\}$.
C([−1, 1]).

3.1.4 Linear Combinations

Definition 3.1.10 (Linear Span) Let V (F) be a vector space and let $S = \{u_1, u_2, \ldots, u_n\}$ be a non-empty subset of V. The linear span of S is the set defined by $L(S) = \{\alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_nu_n : \alpha_i \in F, 1 \le i \le n\}$

If S is an empty set we define L(S) = {0}. Example 3.1.11 1. Note that (4, 5, 5) is a linear combination of (1, 0, 0), (1, 1, 0), and (1, 1, 1) as (4, 5, 5) =5(1, 1, 1) − 1(1, 0, 0) + 0(1, 1, 0).

For each vector, the linear combination in terms of the vectors (1, 0, 0),(1, 1, 0), and(1, 1, 1) is unique.

2. Is (4, 5, 5) a linear combination of (1, 2, 3),(−1, 1, 4) and (3, 3, 2)? Solution: We want to find $\alpha_1, \alpha_2, \alpha_3 \in R$ such that

$\alpha_1(1, 2, 3) + \alpha_2(-1, 1, 4) + \alpha_3(3, 3, 2) = (4, 5, 5)$. (3.1.1)Check that 3(1, 2, 3)+(−1)(−1, 1, 4)+0(3, 3, 2) = (4, 5, 5). Also, in this case, the vector (4, 5, 5) does not have a unique expression as linear combination of vectors (1, 2, 3),(−1, 1, 4) and3. Verify that (4, 5, 5) is not a linear combination of the vectors (1, 2, 1) and (1, 1, 0)? 4. The linear span of $S = \{(1, 1, 1),(2, 1, 3)\}$ over R is $L(S) = \{\alpha(1, 1, 1) + \beta(2, 1, 3) : \alpha, \beta \in R\}$

$= \{(\alpha + 2\beta, \alpha + \beta, \alpha + 3\beta) : \alpha, \beta \in R\}$
$= \{(x, y, z) \in R^3 : 2x - y = z\}$.
as $2(\alpha + 2\beta) - (\alpha + \beta) = \alpha + 3\beta$, and if $z = 2x - y$, take $\alpha = 2y - x$ and $\beta = x - y$. Lemma 3.1.12 (Linear Span is a subspace) Let V (F) be a vector space and let S be a non-empty subsetof V. Then L(S) is a subspace of V (F).

Proof. By definition, $S \subset L(S)$ and hence L(S) is non-empty subset of V. Let $u, v \in L(S)$. Then, for $1 \le i \le n$ there exist vectors $w_i \in S$, and scalars $\alpha_i, \beta_i \in F$ such that $u = \alpha_1w_1 + \alpha_2w_2 + \cdots + \alpha_nw_n$and $v = \beta_1w_1 + \beta_2w_2 + \cdots + \beta_nw_n$. Hence,

$u + v = (\alpha_1 + \beta)w_1 + \cdots + (\alpha_n + \beta_n)w_n \in L(S)$.
Thus, L(S) is a vector subspace of V (F).
(3, 3, 2).

Remark 3.1.13 Let V (F) be a vector space and W ⊂ V be a subspace. If S ⊂ W, then L(S) ⊂ W is a

subspace of W as W is a vector space in its own right. Theorem 3.1.14 Let S be a non-empty subset of a vector space V. Then L(S) is the smallest subspace of Proof. For every u ∈ S, u = 1.u ∈ L(S) and therefore, S ⊆

L(S). To show L(S) is the smallest subspace of V containing S, consider any subspace W of V containing S. Then by Proposition 3.1.13.

Proof. Part 1) can be easily proved. Let A be an m ×n matrix. For part 2), let Dbe the row-reduced form of A with non-zero rows $d^t_1, d^t_2, \ldots, d^t_r$. Then $B = E_kE_{k-1} \cdots E_2E_1A$ for some elementary matrices $E_1, E_2, \ldots, E_k$.

Then, a repeated application of Lemma 3.1.16 implies Row Space(A) = Row Space(B). That is, if the rows of the matrix A are $a^t_1, a^t_2, \ldots, a^t_m$, then $L(a_1, a_2, \ldots, a_m) = L(b_1, b_2, \ldots, b_r)$. Hence the required result follows.

Exercise 3.1.18 1. Show that any two row-equivalent matrices have the same row space. Give examples to show that the column space of two row-equivalent matrices need not be same. 2. Find all the vector subspaces

of R^2.

L(S) ⊆ W and hence the result follows. Definition 3.1.15 Let A be an m × n matrix with real entries. Then using the rows $a^t_1, a^t_2, \ldots, a^t_m \in R^n$

and columns $b_1, b_2, \ldots, b_n \in R^m$, we define

1. RowSpace(A) = $L(a_1, a_2, \ldots, a_m)$,
2. ColumnSpace(A) = $L(b_1, b_2, \ldots, b_n)$,
4. Range(A), denoted Im (A) = $\{y : Ax = y \text{ for some } x^t \in R^n\}$. Note that the "column space" of a matrix A consists of all b such that Ax = b has a solution. Hence, ColumnSpace(A) = Range(A). Lemma 3.1.16 Let A be a

real m × n matrix. Suppose B = EA for some elementary matrix E. Then Row Space(A) = Row Space(B). V containing S.

3. Let P and Q be two subspaces of a vector space V. Show that P ∩ Qis a subspace of V. Also show that P ∪ Q need not be a subspace of V. When is P ∪ Q a subspace of V ? 4. Let P and Q be two subspaces of a

vector space V. Define P + Q ={u +v : u∈P, v∈Q}. Showthat P + Q is a subspace of V. Also show that L(P ∪ Q) = P + Q. 5. Let S = $\{x_1, x_2, x_3, x_4\}$ where $x_1 = (1, 0, 0, 0)$, $x_2 = (1, 1, 0, 0)$, $x_3 = (1, 2, 0, 0)$, $x_4 = (1, 1, 1, 0)$.

Determine all x_i such that L(S) = L(S \ $\{x_i\}$). 6. Let C([−1, 1]) be the set of all continuous functions on the interval [−1, 1] (cf. Example 3.1.4.11). Let

$W_1 = \{f \in C([-1, 1]) : f(0.2) = 0\}$, and

$W_2 = \{f \in C([-1, 1]) : f(\frac{1}{4}) \text{ exists }\}$.

Are W_1, W_2 subspaces of C([−1, 1])?

Proof. We prove the result for the elementary matrix $E_{ij}(c)$, where c '= 0 and i < j. Let $a^t_1, a^t_2, \ldots, a^t_m$

be the rows of the matrix A. Then $B = E_{ij}(c)A$ gives us Row Space(B) = $L(a_1, \ldots, a_{i-1}, a_i + ca_j, \ldots, a_m) = \{\alpha_1a_1 + \cdots + \alpha_{i-1}a_{i-1} + \alpha_i(a_i$

$+ ca_j) + \cdots$

$= L(a_1, \ldots, a_{i-1}, a_i, \ldots, a_m)$

= Row Space(A) Theorem 3.1.17 Let A be an m × n matrix with real entries. Then

3. NullSpace(A), denoted N (A) as $\{x^t \in R^n : Ax = 0\}$. 1. N (A) is a subspace of R^n;

2. the non-zero row vectors of a matrix in row-reduced form, forms a basis for the row-space. Hence dim(Row Space(A)) = row rank of (A).

$+\alpha_m a_m : \alpha_0 \in R, 1 \le H \le m\}$

$_{0=1}\ \alpha_0 a_0 + \alpha_i \cdot ca_j : \alpha_0 \in R, 1 \le H \le m$

$= {}^{*m}$

$\beta_0 a_0 : \beta_0 \in R, 1 \le H \le m$

$= {}^{*m}$

3.2 Linear Independence

$_{0=1}$

Definition 3.2.1 (Linear Independence and Dependence) Let S = $\{u_1, u_2, \ldots, u_m\}$ be any non-empty subset of V. If there exist some

non-zero α_i's $1 \le i \le m$, such that $\alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_mu_m = 0$,

7. Let V = {(x, y) : x, y ∈ R} over R. Define $(x, y) \oplus (x_1, y_1) = (x + x_1, 0)$ and $\alpha\$(x, y) = (\alpha x, 0)$. Show that V is not a vector

then the set S is called a linearly dependent set. Otherwise, the set S is called linearly independent. Example 3.2.2 1. Let S = {(1, 2, 1),(2, 1, 4),(3, 3, 5)}.Then check that 1(1, 2, 1)+1(2, 1, 4)+(−1)(3, 3, 5) =

(0, 0, 0). Since $\alpha_1 = 1$, $\alpha_2 = 1$ and $\alpha_3 = -1$ is a solution of (3.2.1), so the set S is a linearly dependent subset of R^3. Corollary 3.2.6 Let $\{v_1, v_2, \ldots, v_p\}$ be a linearly independent subset of a vector space V. Suppose there

$L(u_1, u_2, \ldots, u_k) = L(u_1, u_2, \ldots, u_{k-1})$. The next corollary follows immediately from Theorem 3.2.4 and Corollary 3.2.5.

2. Let S = {(1, 1, 1),(1, 1, 0),(1, 0, 1)}. Suppose there exists α, β, γ ∈ R such that α(1, 1, 1)+β(1, 1, 0)+ γ(1, 0, 1) = (0, 0, 0). Then check that in this case we necessarily have α = β = γ = 0 which shows that the set S = {(1, 1, 1),(1, 1, 0),(1, 0, 1)} is a linearly independent subset of R^3. In other words, if S = $\{u_1, u_2, \ldots, u_m\}$ is a non-empty subset of a vector space V, then to check whether the set S is linearly dependent or independent, one needs to consider the equation $\alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_mu_m = 0$. (3.2.1)

In case $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ is the only solution of (3.2.1), the set S becomes a linearly independent subset of V. Otherwise, the set S becomes a linearly dependent subset of V. Proposition 3.2.3 Let V be a vector space. 1. Then the zero-vector cannot belong to a

linearly independent set. 2. If S is a linearly independent subset of V, then every subset of S is also linearly independent. 3. If S is a

linearly dependent subset of V then every set containing S is also linearly dependent. Proof. We give the proof of the first part. The reader is required to supply the proof of other parts.

exists a vector v ∈ V, such that v '∈ $L(v_1, v_2, \ldots, v_p)$. Then the set $\{v_1, v_2, \ldots, v_p, v\}$ is also linearly independent subset of

Let S = $\{0 = u_1, u_2, \ldots, u_n\}$ be a set consisting of the zero vector. Then for any γ '= o, $\gamma u_1 + ou_2 +$

$\cdots + 0u_n = 0$. Hence, for the system $\alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_mu_m = 0$, we have a non-zero solution $\alpha_1 = \gamma$

and $o = \alpha_2 = \cdots = \alpha_n$. Therefore, the set S is linearly dependent. Theorem 3.2.4 Let $\{v_1, v_2, \ldots, v_p\}$ be a linearly independent subset of a vector space V. Suppose there exists a vector $v_{p+1} \in V$, such that the set $\{v_1, v_2, \ldots, v_p, v_{p+1}\}$ is linearly dependent, then v_{p+1} is a linear combination of $v_1, v_2, \ldots, v_p$. Proof. Since the set $\{v_1, v_2, \ldots, v_p, v_{p+1}\}$ is linearly dependent, there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_{p+1}$' not all zero such that $\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_pv_p + \alpha_{p+1}v_{p+1} = 0$. (3.2.2)

Claim: α_{p+1} '= 0. Let if possible $\alpha_{p+1} = 0$. Then equation (3.2.2) gives $\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_pv_p = 0$ with not all α_i, $1 \le i \le p$ zero. Hence, by the definition of linear independence, the set $\{v_1, v_2, \ldots, v_p\}$ is linearly dependent

which is contradictory to our hypothesis. Thus, α_{p+1} '= 0 and we get $v_{p+1} = -\frac{1}{\alpha_{p+1}}(\alpha_1v_1 + \cdots + \alpha_pv_p)$.

8. Show that any set of k vectors in R^3 is linearly dependent if k ≥ 4. 9. Is the set of vectors (1, 0),(i, 0) linearly independent subset of $C^2(R)$? 10. Under what conditions on α are the vectors (1 + α, 1 − α) and (α −1, 1 +α) in $C^2(R)$ linearly 11. Let u, v ∈ V and M be a subspace of V. Further, let K be the subspace spanned by Manduand H be the subspace spanned by M and v. Show that if v ∈ K and v '∈ Mthen u ∈H.

3.3 Bases

Definition 3.3.1 (Basis of a Vector Space) 1. A non-empty subset B of a vector space Vis called a basis of V if

L(S)?

(a) B is a linearly independent set, and

(b) L(B) = V, i.e., every vector in V can be expressed as a linear combination of the elements of B.independent?

space over R. 8. Recall that $M_n(R)$ is the real vector space of all n × n real matrices. Prove that the following subsets are

subspaces of $M_n(R)$. (a) $sl_n = \{A \in M_n(R) : trace(A) = 0\}$

(b) $Sym_n = \{A \in M_n(R) : A = A^t\}$

(c) $Skew_n = \{A \in M_n(R) : A + A^t = 0\}$

9. Let V = R. Define x⊕ y = x− y and α$ x = −αx. Which vector space axioms are not satisfied here?In this section, we saw that a vector space has infinite number of vectors. Hence, one can start withany finite collection of vectors and obtain their span. It means that any vector space contains infinitenumber of other vector subspaces. Therefore, the following questions arise: 1. What are the conditions under which, the linear span of two distinct sets the same? 2. Is it possible to find/choose vectors so that the linear span of the chosen vectors is the whole vectorspace itself? 3. Suppose we are able to choose certain vectors whose linear span is the whole space. Can we findthe minimum number of such vectors? We try to answer these questions in the subsequent sections.

Note that $\alpha_i \in F$ for every i, $1 \le i \le p + 1$ and hence $-\frac{\alpha_i}{\alpha_{p+1}} \in F$ for $1 \le i \le p$. Hence the result follows. We now state two important corollaries of the above theorem. We don't give their proofs as they are easy consequence of the above theorem. Corollary 3.2.5 Let $\{v_1, v_2, \ldots, v_n\}$ be a linearly dependent subset of a vector space V. Then there exists a smallest k, $2 \le k \le n$ such that

V. Exercise 3.2.7 1. Consider the vector space R^2. Let $u_1 = (1, 0)$. Find all choices for the vector u_2 suchthat the set $\{u_1, u_2\}$ is linear independent subset of R^2. Does there exist choices for vectors u_2 and u_3 such that the set $\{u_1, u_2, u_3\}$ is linearly independent subset of R^2? 2. If none of the elements appearing along the principal diagonal of a lower triangular matrix is zero, show that the row vectors are linearly independent in R^n. The same is true for column vectors. 3. Let S = {(1, 1, 1, 1),(1, −1, 1, 2),(1, 1, −1, 1)} ⊂ R^4. Determine whether or not the vector (1, 1, 2, 1)∈4. Show that S = {(1, 2, 3),(−2, 1, 1),(8, 6, 10)} is linearly dependent in R^3. 5. Show that S = {(1, 0, 0),(1, 1, 0),(1, 1, 1)} is a linearly independent set in R^3. In general if $\{f_1, f_2, f_3\}$ is a linearly independent set then $\{f_1, f_1 + f_2, f_1 + f_2 + f_3\}$ is also a linearly independent set. 6. In R^3, give an example of 3 vectors u, v and w such that {u, v, w} is linearly dependent but any setof 2 vectors from u, v, w is linearly independent.

7. What is the maximum number of linearly independent vectors in R^3?

2. A vector in B is called a basis vector. Remark 3.3.2 Let $\{v_1, v_2, \ldots, v_p\}$ be a basis of a vector space V (F). Then any $v \in V$ is a unique

Remark 3.3.5 We can use the above results to obtain a basis of any finite dimensional vector space V as follows: Step 1: Choose a non-zero vector, say, $v_1 \in V$. Then the set $\{v_1\}$ is linearly independent. Step 2: If $V = L(v_1)$, we have got a basis of V. Else there exists a vector, say, $v_2 \in V$ such that $v_2 \notin L(v_1)$. Then by Corollary 3.2.6, the set $\{v_1, v_2\}$ is linearly independent. Step 3: If $V = L(v_1, v_2)$, then $\{v_1, v_2\}$ is a basis of V. Else there exists a vector, say, $v_3 \in V$ such that $v_3 \notin L(v_1, v_2)$. So, by Corollary 3.2.6, the set $\{v_1, v_2, v_3\}$ is linearly independent. At the ith step, either $V = L(v_1, v_2, \ldots, v_i)$, or $L(v_1, v_2, \ldots, v_i) \neq V$. In the first case, we have $\{v_1, v_2, \ldots, v_i\}$ as a basis of V. In the second case, $L(v_1, v_2, \ldots, v_i) \subset V$. So, we choose a vector, say, $v_{i+1} \in V$ such that $v_{i+1} \notin L(v_1, v_2, \ldots, v_i)$. Therefore, by Corollary 3.2.6, the set $\{v_1, v_2, \ldots, v_{i+1}\}$ is linearly independent. This process will finally end as V is a finite dimensional vector space. Exercise 3.3.6 1. Let $S = \{v_1, v_2, \ldots, v_p\}$ be a subset of a vector space V (F). Suppose $L(S) = V$ but S is not a linearly independent set. Then prove that each vector in V can be expressed in more than one way as a linear combination of vectors from S. 2. Show that the set $\{(1, 0, 1), (1, i, 0), (1, 1, 1 - i)\}$ is a basis of $C^3(C)$.

3. Let $V = \{(x, y, z) : x+y-z = 0, x, y, z \in R\}$ be a vector subspace of R^3. Then $S = \{(1, 1, 2), (2, 1, 3), (1, 2, 3)\} \subset V$. It can be easily verified that the vector $(3, 2, 5) \in V$ and $(3, 2, 5) = (1, 1, 2) + (2, 1, 3) = 4(1, 1, 2) - (1, 2, 3)$. Then by Remark 3.3.2, S cannot be a basis of V. A basis of V can be obtained by the following method: The condition $x + y - z = 0$ is equivalent to $z = x + y$. we replace the value of z with $x + y$ to get $(x, y, z) = (x, y, x + y) = (x, 0, x) + (0, y, y) = x(1, 0, 1) + y(0, 1, 1)$. Hence, $\{(1, 0, 1), (0, 1, 1)\}$ forms a basis of V.

4. Let $V = \{a + ib : a, b \in R\}$ and $F = C$. That is, V is a complex vector space. Note that any element $a + ib \in V$ can be written as $a + ib = (a + ib)1$. Hence, a basis of V is $\{1\}$.

3.3.1 Important Results

Theorem 3.3.7 Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of a given vector space V. If $\{w_1, w_2, \ldots, w_m\}$ is a set of vectors from V with $m > n$ then this set is linearly dependent. Proof. Since we want to find whether the set $\{w_1, w_2, \ldots, w_m\}$ is linearly independent or not, we consider the linear system $\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_m w_m = 0$ (3.3.1) with $\alpha_1, \alpha_2, \ldots, \alpha_m$ as the m unknowns. If the solution set of this linear system of equations has more than one solution, then this set will be linearly dependent. As $\{v_1, v_2, \ldots, v_n\}$ is a basis of V and $w_i \in V$, for each i, $1 \le i \le m$, there exist scalars a_{ij}, $1 \le i \le n$, $1 \le j \le m$, such that $w_m = a_{1m}v_1 + a_{2m}v_2 + \cdots + a_{nm}v_n$. The set of equations (3.3.1) can be rewritten as

5. Let $V = \{a + ib : a, b \in R\}$ and $F = R$. That is, V is a real vector space. Any element $a + ib \in V$ is expressible as $a \cdot 1 + b \cdot i$. Hence a basis of V is $\{1, i\}$. Observe that i is a vector in C. Also, $i \notin R$ and hence $i \cdot (1 + 0 \cdot i)$ is not defined.

6. Recall the vector space P(R), the vector space of all polynomials with real coefficients. A basis of this vector space is the set $\{1, x, x^2, \ldots, x^n, \ldots\}$. This basis has infinite number of vectors as the degree of the polynomial can be any positive integer. Definition 3.3.4 (Finite Dimensional Vector Space) A vector space V is said to be finite dimensional if there exists a basis consisting of finite number of elements. Otherwise, the vector space V is called infinite dimensional. In Example 3.3.3, the vector space of all polynomials is an example of an infinite dimensional vector space. All the other vector spaces are finite dimensional.

Since the set $\{v_1, v_2, \ldots, v_n\}$ is linearly independent, we have $\sum_{i=1} \alpha_i a_{ni} = 0$. Therefore, finding α_i's satisfying equation (3.3.1) reduces to solving the system of homogeneous equations $\sum_{i=1} \alpha_i a_{2i} = \cdots = 0$. Since n < m, i.e., the number of equations is strictly less than the number of unknowns, Corollary 2.6.3 implies that the solution

$A\alpha = 0$ where $\alpha^t = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ and $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$.

$$w_1 = a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n$$
$$w_2 = a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n$$
$$\vdots$$

2. Consider the real vector space $C^2(R)$. In this case, any vector

$(a + ib, c + id) = a(1, 0) + b(i, 0) + c(0, 1) + d(0, i)$. Hence, the set $\{(1, 0), (i, 0), (0, 1), (0, i)\}$ is a basis and $\dim(V) = 4$. Remark 3.3.13 It is important to note that the dimension of a vector space may change if the under lying field (the set of scalars) is changed. Example 3.3.14 Let V be the set of all functions $f : R^n \to R$ with the property that $f(x+y) = f(x) + f(y)$

$$\sum_{j=1} a_{ij}v_j = 0, \quad \sum_{j=1} a_{2j}v_j = 0, \quad \ldots, \quad \sum_{j=1} a_{mj}v_j = 0$$

i.e.,

set consists of infinite number of elements. Therefore, the equation (3.3.1) has a solution with not all α_i, $1 \le i \le m$, zero. Hence, the set $\{w_1, w_2, \ldots, w_m\}$ is a linearly dependent set. Remark 3.3.8 Let V be a vector subspace of R^n with spanning set S. We give a method of finding a basis of V from S.

1. Construct a matrix A whose rows are the vectors in S. 2. Use only the elementary row operations $R_i(c)$ and $R_{ij}(c)$ to get the row-reduced form B of A (in fact we just need to make as many zero-rows as possible). 3. Let B be the set of vectors in S corresponding to the non-zero rows of B. Then the set B is a basis of L(S) = V. Example 3.3.9 Let $S = \{(1, 1, 1, 1), (1, 1, -1, 1), (1, 1, 0, 1), (1, -1, 1, 1)\}$ be a subset of R^4. Find a basis of

$(f \oplus g)(x) = f(x) + g(x)$ and $(t \$ f)(x) = f(tx)$.

and $f(\alpha x) = \alpha f(x)$. For f, g ∈ V, and t ∈ R, define

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix} \xrightarrow{R_{12}(-1),\ R_{13}(-1),\ R_{14}(-1)} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & 0 \\ 1 & 1 & -1 & 1 \\ 0 & 0 & -2 & 0 \end{bmatrix} \xrightarrow{R_{32}(-2)} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$

Then V is a real vector space.

For $1 \le i \le n$, consider the functions

$e_i(x) = e_i(x_1, x_2, \ldots, x_n) = x_i$. Then it can be easily verified that the set $\{e_1, e_2, \ldots, e_n\}$ is a basis of V and hence dim(V) =n. The next theorem follows directly from Corollary 3.2.6 and Theorem 3.3.7. Hence, the proof is omitted. Theorem 3.3.15 Let S be a linearly independent subset of a finite dimensional vector space V. Then the set S can be extended to form a basis of V. Theorem 3.3.15 is equivalent to the following statement:

Let V be a vector space of dimension n. Suppose, we have found a linearly independent set S=

$$\begin{matrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix}$$

$$\begin{matrix} 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -2 & 0 & 0 \end{matrix}$$

. Observe that the rows 1, 3 and 4 are non-zero. Hence, a basis of L(S) consists of the first, third and fourth vectors of the set S. Thus, B = {(1, 1, 1, 1), (1, 1, 0, 1), (1, −1, 1, 1)} is a basis of L(S). Observe that at the last step, in place of the elementary row operation $R_{32}(-2)$, we can apply $R_{23}(-{}^{1}{}_{2})$ to make the third row as the zero-row. In this case, we get {(1, 1, 1, 1), (1, 1, −1, 1), (1, −1, 1, 1)} as a basis of L(S). Corollary 3.3.10 Let V be a finite dimensional vector space. Then any two bases of V have the same number of vectors. Proof. Let $\{u_1, u_2, \ldots, u_n\}$ and $\{v_1, v_2, \ldots, v_m\}$ be two bases of V with m > n. Then by the above $\{v_1, v_2, \ldots, v_r\} \subset V$. Then there exist vectors $v_{r+1}, \ldots, v_n$ in V such that $\{v_1, v_2, \ldots, v_n\}$ is a basis of V. Corollary 3.3.16 Let V be a vector space of dimension n. Then any set of n linearly independent vectorsforms a basis of V. Also, every set of m vectors, m > n, is linearly dependent. Example 3.3.17 Let V = {(v, w, x, y, z) ∈ R^5: v + x − 3y + z = 0} and W={(v, w, x, y, z) ∈R^5: w − x − z = 0, v = y} be two subspaces of R^5. Find bases of V and Wcontaining a basis of V∩W. Solution: Let us find a basis of V ∩ W. The solution set of the linear equations

v + x − 3y + z = 0, w − x − z = 0 and v =y

is given by

theorem the set $\{v_1, v_2, \ldots, v_m\}$ is linearly dependent if we take $\{u_1, u_2, \ldots, u_n\}$ as the basis of V. This contradicts the assumption that $\{v_1, v_2, \ldots, v_m\}$ is also a basis of V. Hence, we get m = n. Definition 3.3.11 (Dimension of a Vector Space) The dimension of a finite dimensional vector space V L(S).

is the number of vectors in a basis of V, denoted dim(V).

Solution: Here A =

$(v, w, x, y, z)^t = (y, 2y, x, y, 2y - x)^t = y(1, 2, 0, 1, 2)^t + x(0, 0, 1, 0, -1)^t$. Thus, a basis of V ∩ W is {(1, 2, 0, 1, 2),(0, 0, 1, 0, −1)}. To find a basis of W containing a basis of V ∩ W, we can proceed as follows:

3.3. BASES 63 1. Find a basis of W.

2. Take the basis of V ∩W found above as the first two vectors and that of W as the next set of vectors. Now use Remark 3.3.8 to get the required basis. Heuristically, we can also find the basis in the following way: A vector of W has the form (y, x + z, x, y, z) for x, y, z ∈ R. Substituting y = 1, x = 1, and z = 0 in

64 CHAPTER 3. FINITE DIMENSIONAL VECTOR SPACES7. If M and N are 4-dimensional subspaces of a vector space V of dimension 7 then show that Mand N have at least one vector in common other than the zero vector. 8. Let P = L{(1, 0, 0),(1, 1, 0)} and Q = L{(1, 1, 1)} be vector subspaces of R^3. Show that P+Q=$R^3$and P ∩ Q = {0}. If u ∈ R^3, determine u_P , u_Q such that u = u_P + $u_Q$where u_P∈Pandu_Q∈Q. Is it necessary that u_P and u_Q are unique?

9. Let W_1 be a k-dimensional subspace of an n-dimensional vector space V (F) where k ≥1. Prove that there exists an (n − k)-dimensional subspace W_2 of V such that $W_1 \cap W_2$ ={0} and W_1+W_2=V.

(y, x + z, x, y, z) gives us the vector (1, 1, 1, 1, 0) ∈ W. It can be easily verified that a basis of W is

10. Let P and Q be subspaces of R^n such that P + Q = R^n and P ∩ Q={0}. Then show that each

{(1, 2, 0, 1, 2),(0, 0, 1, 0, −1),(1, 1, 1, 1, 0)}. Similarly, a vector of V has the form (v, w, x, y, 3y−v−x) for v, w, x, y ∈ R. Substituting v = 0, w = 1, x = 0 and y = 0, gives a vector (0, 1, 0, 0, 0) ∈ V. Also, substituting v = 0, w = 1, x = 1 and y = 1, gives another vector (0, 1, 1, 1, 2) ∈ V. So, a basis of V can be taken as

u ∈ R^ncan be uniquely expressed as u = u_P + u_Q where u_P ∈ P and u_Q∈ Q. 11. Let P = L{(1, −1, 0),(1, 1, 0)} and Q = L{(1, 1, 1),(1, 2, 1)} be vector subspaces of R^3. Show thatP + Q = $R^3$and P ∩ Q '= {0}. Show that there exists a vector u ∈ R^3such that ucannot be written uniquely in the form u = u_P + u_Q where u_P ∈ P and u_Q ∈ Q. 12. Recall the vector space P_4(R). Is the set,

{(1, 2, 0, 1, 2),(0, 0, 1, 0, −1),(0, 1, 0, 0, 0),(0, 1, 1, 1, 2)}. Recall that for two vector subspaces M and N of a vector space V (F), the vector subspace M + N is defined by

W = {p(x) ∈ P_4(R) : p(−1) = p(1) = 0}

M + N = {u + v : u ∈ M, v ∈ N}. With this definition, we have the following very important theorem (for a proof, see Appendix 14.4.1). Theorem 3.3.18 Let V (F) be a finite dimensional vector space and let M and N be two subspaces of V. dim(M) + dim(N) = dim(M + N) + dim(M ∩ N). (3.3.2) Exercise 3.3.19 1. Find a basis of the vector space P_n(R). Also, find dim(P_n(R)). What can you say about the dimension of P(R)?

a subspace of P_4(R)? If yes, find its dimension. 13. Let V be the set of all 2 × 2 matrices with complex entries and a_{11} +a_{22} = 0. Show that Vis a real vector space. Find its basis. Also let W = {A ∈ V : a_{21} = −a_{12}}. Show Wis a vector subspace of V, and find its dimension. 14. Let A =

$$\begin{matrix} 1 & 2 & 1 & 3 & 2 \\ 4 & 2 & 5 & 6 & 10 \end{matrix}, \text{ and } B = \begin{matrix} 2 & 4 & 0 & 6 & -1 \\ 0 & -2 & 5 & -3 & -5 \\ 1 & -4 & -1 & -1 & 2 \end{matrix}$$

be two matrices. For

AandBfindthe following: (a) their row-reduced echelon forms.

$$\begin{matrix} 0 & 2 & 2 & 2 & 4 \\ 2 & -2 & 4 & 0 & 8 \end{matrix}$$

(b) the matrices P_1 and P_2 such that P_1A and P_2B are in row-reduced form. (c) a basis each for the row spaces of A and B. (d) a basis each for the range spaces of A and B. (e) bases of the null spaces of A and B.

(f) the dimensions of all the vector subspaces so obtained. 15. Let M(n, R) denote the space of all n × n real matrices. For the sets given below, check that they are subspaces of M(n, R) and also find their dimension. (a) sl(n, R) = {A ∈ M(n, R) : tr(A) = 0}, where recall that tr(A) stands for trace of A. (b) S(n, R) = {A ∈ M(n, R) : A = A^t}.

4. Let W = {(x, y, z, w) ∈ R^4: x + y − z + w = 0} be a subspace of R^4. Find its basis and dimension. 5. Let V = {(x, y, z, w) ∈ R^4: x + y − z + w = 0, x + y + z + w = 0} and W = {(x, y, z, w) ∈ R^4: x − y − z + w = 0, x + 2y − w = 0} be two subspaces of R^4. Find bases and dimensions of V, W, 6. Let V be the set of all real symmetric n × n matrices. Find its basis and dimension. What if V is the complex vector space of all n × n Hermitian matrices?

(c) A(n, R) = {A ∈ M(n, R) : A + A^t = 0}. Before going to the next section, we prove that for any matrix A of order m×n Row rank(A) = Column rank(A).

2. Consider the real vector space, C([0, 2π]), of all real valued continuous functions. For each n consider the vector e_n defined by $e_n(x)$ = sin(nx). Prove that the collection of vectors $\{e_n : 1 \le n < \infty\}$ is a linearly independent set.

[Hint: On the contrary, assume that the set is linearly dependent. Then we have a finite set of vectors, say $\{e_{k_1}, e_{k_2}, \ldots, e_{k}\}$ that are linearly dependent. That is, there exist scalars α_i ∈ R for 1 ≤ i ≤ # not all zero such that $\alpha_1 \sin(k_1x) + \alpha_2 \sin(k_2x) + \cdots + \alpha \sin(k x) = 0$ for all x ∈ [0, 2π]. Now for different values of m integrate the function $\alpha_1 \sin(k_1x) + \alpha_2 \sin(k_2x) + \cdots + \alpha \sin(k x))$ dx to get the required result.]

3. Show that the set {(1, 0, 0),(1, 1, 0),(1, 1, 1)} is a basis of C^3(C). Is it a basis of C^3(R) also?

3.3. BASES 65 Proposition 3.3.20 Let A be an m × n real matrix. Then Row rank(A) = Column rank(A).

Proof. Let $R_1, R_2, \ldots, R_m$ be the rows of A and $C_1, C_2, \ldots, C_n$ be the columns of A. Note that Row rank(A) = r, means that dim L($R_1, R_2, \ldots, R_m$) = r. Hence, there exists vectors $u_1 = (u_{11}, \ldots, u_{1n})$, $u_2 = (u_{21}, \ldots, u_{2n})$, . . . ,$u_r = (u_{r1}, \ldots, u_m)$ ∈ R^n

V ∩ W and V + W.

R_i ∈ L($u_1, u_2, \ldots, u_r$) ∈ R^n, for all i, 1 ≤ i ≤ m. Therefore, there exist real numbers α_{ij} , 1 ≤ i ≤ m, 1 ≤ j ≤ r such that $R_1 = \alpha_{11}u_1 + \alpha_{12}u_2 + \cdots + \alpha_{1r}u_r$ = (

$R_2 = \alpha_{21}u_1 + \alpha_{22}u_2 + \cdots + \alpha_{2r}u_r$ = (

$R_m = \alpha_{m1}u_1 + \cdots + \alpha_{mr}u_r$ = (, and so on, till

$_{i=1}\ \alpha_{mi}u_{i2}, \ldots, {}^{*r}$

$+ \cdots + u_{r1}$ $\alpha^{1r}\ \alpha^{2r}\ldots$

. In general, for $1 \le j \le n$, we have
with

$+ \cdots + u_{rj}$ $\alpha^{1r}\ \alpha^{2r}\ldots$

$_{i=1}\ \alpha_{1i}u_{i1}, {}^{*r}$
$_{i=1}\ \alpha_{1i}u_{in}),$
$_{i=1}\ \alpha_{2i}u_{i1}, {}^{*r}$

. Therefore, we observe that the columns $C_1, C_2, \ldots, C_n$ are linear combination of the r vectors $(\alpha_{11}, \alpha_{21}, \ldots, \alpha_{m1})^t, (\alpha_{12}, \alpha_{22}, \ldots, \alpha_{m2})^t, \ldots, (\alpha_{1r}, \alpha_{2r}, \ldots, \alpha_{mr})^t$.

$_{i=1}\ \alpha_{mi}u_{i1}, {}^{*r}$
$_{i=1}\ \alpha_{mi}u_{in}).$
Therefore, Column rank(A) = dim $L(C_1, C_2, \ldots, C_n)$ $=\le r =$ Row rank(A). A similar argument gives Row rank(A) $\le$ Column rank(A). Thus, we have the required result.

3.4 Ordered Bases

Let $B = \{u_1, u_2, \ldots, u_n\}$ be a basis of a vector space V (F). As B is a set, there is no ordering of its elements. In this section, we want to associate an order among the vectors in any basis of V. Definition 3.4.1 (Ordered Basis) An ordered basis for a vector space V (F) of dimension n, is a ba- sis $\{u_1, u_2, \ldots, u_n\}$ together with a one-to-one correspondence between the sets $\{u_1, u_2, \ldots, u_n\}$ and $\{1, 2, 3, \ldots, n\}$. If the ordered basis has u_1 as the first vector, u_2 as the second vector and so on, then we denote this ordered basis by $(u_1, u_2, \ldots, u_n)$. Example 3.4.2 Consider $P_2(R)$, the vector space of all polynomials of degree less than or equal to 2 with coefficients from R. The set $\{1 - x, 1 + x, x^2\}$ is a basis of $P_2(R)$. For any element $a_0 + a_1x + a_2x^2 \in P_2(R)$, we have $2(1 - x) + a_0 + a_1$

$2(1 + x) + a_2x^2$.
2 is the second component, and a_2 is the third component of the vector $a_0 + a_1x + a_2x^2$. If $(1-x, 1+x, x^2)$ is an ordered basis, then $a_0 - a_1$ 2 is the first component, $a_0 + a_1$

2 is the first component, $a_0 - a_1$ 2 is the second component, and a_2 is the third component of the vector $a_0 + a_1x + a_2x^2$. That is, as ordered bases $(u_1, u_2, \ldots, u_n)$, $(u_2, u_3, \ldots, u_n, u_1)$, and $(u_n, u_1, u_2, \ldots, u_{n-1})$ are differenteven though they have the same set of vectors as elements. Definition 3.4.3 (Coordinates of a Vector) Let $B = (v_1, v_2, \ldots, v_n)$ be an ordered basis of a vector space

If we take $(1 + x, 1 - x, x^2)$ as an ordered basis, then $a_0 + a_1$ V (F) and let $v \in V$. If

$v = \beta_1v_1 + \beta_2v_2 + \cdots + \beta_nv_n$
then the tuple $(\beta_1, \beta_2, \ldots, \beta_n)$ is called the coordinate of the vector v with respect to the ordered basis B. Mathematically, we denote it by $[v]_B = (\beta_1, \ldots, \beta_n)^t$, a column vector. Suppose $B_1 = (u_1, u_2, \ldots, u_n)$ and $B_2 = (u_n, u_1, u_2, \ldots, u_{n-1})$ are two ordered bases of V. Then for any $x \in V$ there exists unique scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$x = \alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_nu_n = \alpha_nu_n + \alpha_1u_1 + \cdots + \alpha_{n-1}u_{n-1}$. $[x]_{B_1} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^t$ and $[x]_{B_2} = (\alpha_n, \alpha_1, \alpha_2, \ldots, \alpha_{n-1})^t$. Note that x is uniquely written as ^{n}

$a_0 + a_1x + a_2x^2 = a_0 - a_1$
α_iu_i and hence the coordinates with respect to an ordered basis are unique. Suppose that the ordered basis B_1 is changed to the ordered basis $B_3 = (u_2, u_1, u_3, \ldots, u_n)$. Then $[x]_{B_3} = (\alpha_2, \alpha_1, \alpha_3, \ldots, \alpha_n)^t$. So, the coordinates of a vector depend on the ordered basis chosen. Example 3.4.4 Let $V = R^3$. Consider the ordered bases V. Then, with respect to the above bases we have

So,

^{r}
$_{i=1}\ \alpha_{1i}u^{i1}$ ^{r}

$(1, -1, 1) = 1 \cdot (1, 0, 0) + (-1) \cdot (0, 1, 0) + 1 \cdot (0, 0, 1). = 2 \cdot (1, 0, 0) + (-2) \cdot (1, 1, 0) + 1 \cdot (1, 1, 1).$

α_{11}
α_{12}

$= 1 \cdot (1, 1, 1) + (-2) \cdot (1, 1, 0) + 2 \cdot (1, 0, 0).$
$_{i=1}\ \alpha_{2i}u_{i1}$
α_{22}
$\alpha^{21}\ldots$
$C_1 =$
$= u_{11}$
$+ u_{21}$
$\ldots$
$\ldots$
^{r}
α_{m1} α_{m2} α_{mr}
$_{i=1}\ \alpha_{mi}u_{i1}$
^{r}
$_{i=1}\ \alpha_{1i}u^{ij}$ ^{r}
α_{11}
α_{12}

□

$_{i=1}\ \alpha_{2j}u_{ij}$ □ □ Therefore,

□

□

□

□

□

α^{21}.. . C_j =

α^{22}.. □

= u_{1j} □ + u_{2j} □

.. . □

!^{r}

□ □

$\alpha_{mi}u_{ij}$

$_{i=1}$ $_{i=1}$ and B_3 = (1, 1, 1),(1, 1, 0),(1, 0, 0) of

α_{m1}

α_{m2}

α_{mr} B_1 = (1, 0, 0),(0, 1, 0),(0, 0, 1) , B_2 = (1, 0, 0),(1, 1, 0),(1, 1, 1)

3.4. ORDERED BASES 67 Therefore, if we write u = (1, −1, 1), then $[u]_{B_1}$ = (1, −1, 1)^t, $[u]_{B_2}$ = (2, −2, 1)^t, $[u]_{B_3}$ = (1, −2, 2)^t. In general, let V be an n-dimensional vector space with ordered bases B_1 = (u_1, u_2, . . . , u_n) and

B_2 = (v_1, v_2, . . . , v_n). Since, B_1 is a basis of V, there exists unique scalars a_{ij} , 1 ≤ i, j ≤ n such that $a_{ij}u_i$ for 1 ≤ i ≤ n. That is, for each i, 1 ≤ i ≤ n, $[v_i]_{B_1}$ = (a_{1i}, a_{2i}, . . . , a_{ni})^t.

Let v ∈ V with $[v]_{B_2}$ = (α_1, α_2, . . . , α_n)^t. As B_2 as ordered basis (v_1, v_2, . . . , v_n), we have Since B_1 is a basis this representation of v in terms of u_i's is unique. So,

$a_{2i}\alpha_i$, . . . , $_{*n}$

a_{11} · · · a_{1n}

a_{21} · · · a_{2n}

a_{n1} · · · a_{nn} □□ □□ □ □ □ α_1 α_2 ... α_n □□ □

= A$[v]_{B_2}$. Note that the ith column of the matrix A is equal to $[v_i]_{B_1}$, i.e., the ith column of A is the coordinate of the ith vector v_i of B_2 with respect to the ordered basis B_1. Hence, we have proved the following

v_i = $_{*n}$

theorem. Theorem 3.4.5 Let V be an n-dimensional vector space with ordered bases B_1 = (u_1, u_2, . . . , u_n) and

$_{i=1}$

B_2 = (v_1, v_2, . . . , v_n). Let

A = [$[v_1]_{B_1}$, $[v_2]_{B_1}$, . . . , $[v_n]_{B_1}$] . Then for any v ∈ V,

□

□

v = $_{*n}$

$_{i=1}$ $\alpha_i v_i$ = $_{*n}$

□ $_{*n}$

□ = $_{*n}$

u_j .

$_{i=1}$ α_i

68 CHAPTER 3. FINITE DIMENSIONAL VECTOR SPACES2. Let A = [a_{ij}] = □□ □0 2 0

1 1 0 □□ □. The columns of the matrix A are obtained by the following rule: 0 −2 1

$[(1, 1, 1)]_{B_1}$ = 0 ·(1, 0, 0) + 0 · (1, 1, 0) + 1 · (1, 1, 1) = (0, 0, 1)^t,

$[(1, −1, 1)]_{B_1}$ = 2 · (1, 0, 0) + (−2) · (1, 1, 0) + 1 · (1, 1, 1) = (2, −2, 1)^t $[(1, 1, 0)]_{B_1}$ = 0 ·(1, 0, 0) + 1 · (1, 1, 0) + 0 · (1, 1, 1) = (0, 1, 0)^t. That is, the elements of B_2 = (1, 1, 1),(1, −1, 1),(1, 1, 0) are expressed in terms of the

ordered basis3. Note that for any (x, y, z) ∈ R^3,

□

□0 2 0

$[(x, y, z)]_{B_1}$ = □□ □x − y

□ =A$[(x, y, z)]_{B_2}$.

0 −2 1

1 1 0 □□ □□ □ □y−x

4. The matrix A is invertible and hence $[(x, y, z)]_{B_2}$ = $A^{-1}[(x, y, z)]_{B_1}$. In the next chapter, we try to understand Theorem 3.4.5 again using the ideas of 'linear transforma- tions / functions'. Exercise 3.4.7 1. Determine the

coordinates of the vectors (1, 2, 1) and (4, −2, 2) with respect to the2. Consider the vector space P_3(R).

(a) Show that B_1 = (1 − x, 1 + x^2, 1 − x^3, 3 + x^2 − x^3) and B_2 = (1, 1 −x, 1 +x^2, 1 −x^3) are bases(b) Find the coordinates of the vector u = 1 + x + x^2 + x^3 with respect to the ordered basis $B_1$and

and

(c) Find the matrix A such that $[u]_{B_2}$ = A$[u]_{B_1}$. (d) Let v = a_0 + a_1x + a_2x^2 + a_3x^3. Then verify the following:

B_1.

−a_0 − a_1 + 2a_2 − a_3

□

□

$a_{ji}u_j$ $a_{ji}\alpha_i$ □

$_{j=1}$

$_{j=1}$

$_{i=1}$

$_2$ + z $_{x−y}$

□ =

$[v]_{B_1}$ = A$[v]_{B_2}$. Example 3.4.6 Consider two bases B_1 = (1, 0, 0),(1, 1, 0),(1, 1, 1) and B_2 = (1, 1, 1),(1, −1, 1),(1, 1, 0)

y − z

$_t$

$_2$

$[v]_{B_1}$ = $_{*n}$

z

$a_{1i}\alpha_i$, $_{*n}$

x − z

1. Then

0 1 0 0

$a_{ni}\alpha_i$

−1 0 1 0

$_{i=1}$

$_{i=1}$

$[(x, y, z)]_{B_1}$ = (x − y) · (1, 0, 0) + (y − z) · (1, 1, 0) + z · (1, 1, 1)

□

= (x − y, y − z, z)^t

□

=

□

.. .. ..

. . .

$[(x, y, z)]_{B_2}$ = (y − x

2 + z) · (1, 1, 1) + x − y 2 · (1, −1, 1)

= (y − x

2 + z, x − y 2, x − z)^t.

Chapter 4

and

Linear Transformations

+(x − z) · (1, 1, 0)

−1 0 0 1

1 0 0 0 □□ □ □□ □ a_0 + a_1 −a_2 +a^3 −a_1

= $[v]_{B_2}$.

of R^3.

basis B = (2, 1, 0),(2, 1, 1),(2, 2, 1) of R^3.

of $P_3(R)$.

B_2.

$-a_1$

70 CHAPTER 4. LINEAR TRANSFORMATIONS Proposition 4.1.3 Let T : V −→W be a linear transformation. Suppose that 0_V is the zero vector inVand

$[v]_{B_1}$ =

$-a_0 - a_1 + a_2 - 2a_3$

$a_0 + a_1 - a_2 + a_3$

0_W is the zero vector of W. Then $T(0_V) = 0_W$.

Proof. Since $0_V = 0_V + 0_V$, we have

called a linear transformation if

T (αu + βv) = αT(u) + βT (v), for all α, β ∈ F, and u, v ∈ V. We now give a few examples of linear transformations. Example 4.1.2 1. Define T : R−→$R^2$by T(x) = (x, 3x) for all x ∈ R. Then T is a linear transformation T(x + y) = (x + y, 3(x + y)) = (x, 3x) + (y, 3y) = T (x) + T (y). 2. Verify that the maps given below from R^n to R are linear transformations. Let x = $(x_1, x_2, \ldots, x_n)$. (a) Define T (x) = $\int^n$

(b) For any i, 1 ≤ i ≤ n, define $T_i(x) = x_i$.

(c) For a fixed vector a = $(a_1, a_2, \ldots, a_n) \in R^n$, define T (x) = $\int^n a_i x_i$. Note that examples (a)

and (b) can be obtained by assigning particular values for the vector a.

3. Define T : R^2−→$R^3$by T ((x, y)) = (x + y, 2x − y, x + 3y). Then T is a linear transformation with T ((1, 0)) = (1, 2, 1) and T((0, 1)) = (1, −1, 3).

4. Let A be an m × n real matrix. Define a map $T_A : R^n \to R^m$ by $T_A(x) = Ax$ for every $x^t = (x_1, x_2, \ldots, x_n) \in R^n$.

Then T_A is a linear transformation. That is, every m × n real matrix defines a linear transformation 5. Recall that $P_n(R)$ is the set of all polynomials of degree less than or equal to n with real coefficients. T $((a_1, a_2, \ldots, a_{n+1})) = a_1 + a_2x + \cdots + a_{n+1}x^n$

as

for $(a_1, a_2, \ldots, a_{n+1}) \in R^{n+1}$. Then T is a linear transformation.

$_{i=1}$ x_i.

$_{i=1}$

from $R^n$to R^m.

Define T : R^{n+1}−→$P_n(R)$ by

69

Then T is a linear transformation. Such a linear transformation is called the Identity transformation and ~~is~~denoted by I. We now prove a result that relates a linear transformation T with its value on a basis of the domain~~n~~ space. Theorem 4.1.6 Let T : V −→W be a linear transformation and B = $(u_1, u_2, \ldots, u_n)$ be an ordered basisof V. Then the linear transformation T is a linear combination of the vectors $T(u_1), T(u_2), \ldots, T(u_n)$. In other words, T is determined by $T(u_1), T(u_2), \ldots, T(u_n)$.

Proof. Since B is a basis of V, for any x ∈ V, there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$such that $x=\alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_nu_n$. So, by the definition of a linear transformation T (x) = $T(\alpha_1u_1 + \cdots + \alpha_nu_n) = \alpha_1T(u_1) + \cdots + \alpha_nT(u_n)$.

Observe that, given x ∈ V, we know the scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$. Therefore, to knowT (x), we just need to know the vectors $T(u_1), T(u_2), \ldots, T(u_n)$ in W. That is, for every x ∈ V, T (x) is determined by the coordinates $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of xwith respect to the ordered basis B and the vectors $T(u_1), T(u_2), \ldots, T(u_n) \in W$. Exercise 4.1.7 1. Which of the following are linear transformations T : V −→W? Justify your answers. (a) Let V = $R^2$and W = R^3 with T (x, y) = (x + y + 1, 2x − y, x + 3y)

by

(b) Let V = W = R^2 with T (x, y) = $(x - y, x^2 - y^2)$ (c) Let V = W = R^2 with T (x, y) = (x − y, |x|) (d) Let V = $R^2$and W = −→R^4 with T (x, y) = (x + y, x − y, 2x +y, 3x −4y) (e) Let V = W = R^4 with T (x, y, z, w) = (z, x, w, y)

2. Recall that $M_2(R)$ is the space of all 2 × 2 matrices with real entries. Then, which of the following are linear transformations T : $M_2(R) \to M_2(R)$?

4.1. DEFINITIONS AND BASIC PROPERTIES 71 (a) T (A) = A^t(b) T (A) = I + A (c) T (A) = A^2

72 CHAPTER 4. LINEAR TRANSFORMATIONSProof. Since T is onto, for each w ∈ W there exists a vector v ∈ V such that T (v) =w. So, the set$T^{-1}(w)$ is non-empty. Suppose there exist vectors $v_1, v_2 \in V$ such that T $(v_1) = T(v_2)$. But by assumption, Tis one-one and therefore $v_1 = v_2$. This completes the proof of Part 1. We now show that T^{-1}as defined above is a linear transformation. Let $w_1, w_2 \in W$. Then by Part 1, there exist unique vectors $v_1, v_2 \in V$ such that $T^{-1}(w_1) = v_1$ and $T^{-1}(w_2) = v_2$. Or equivalently,

(d) T (A) = BAB^{-1}, where B is some fixed 2 × 2 matrix. 3. Let T : R −→ R be a map. Then T is a linear transformation if and only if there exists a unique c ∈ R such that T (x) = cx for every x ∈ R.

4. Let A be an n × n real matrix. Consider the linear transformation

$T_A(x) = Ax$ for every $x \in R^n$. Then prove that $T^2(x) := T(T(x)) = A^2x$. In general, for k ∈ N, prove that $T^k(x) = A^kx$. 5. Use the ideas of matrices to give examples of linear transformations T, S : R^3−→R^3that satisfy: (a) T '= 0, T^2= 0, T^3 = 0.

=

a_2

T (0_V) = T $(0_V + 0_V)$ = $T(0_V)$ + T (0_V).

$-a_3$

So, $T(0_V) = 0_W$ as T $(0_V) \in W$. From now on, we write 0 for both the zero vector of the domain space and the zero vector of the range space. Definition 4.1.4 (Zero Transformation) Let V be a vector space and let T : V −→Wbe the map defined T(v) = 0 for every v ∈ V. Then T is a linear transformation. Such a linear transformation is called the zero transformation and ~~is~~denoted by 0. Definition 4.1.5 (Identity Transformation) Let V be a vector space and let T : V −→Vbe the map defined by

4.1 Definitions and Basic Properties

Throughout this chapter, the scalar field F is either always the set R or always the set C. Definition 4.1.1 (Linear Transformation) Let V and W be vector spaces over F. A map T : V −→W is

T(v) = v for every v ∈ V.

(b) $T \neq 0$, $S \neq 0$, $S \circ T \neq 0$, $T \circ S = 0$; where $T \circ S(x) = T\,S(x)$. (c) $S^2 = T^2$, $S \neq T$.

(d) $T^2 = I$, $T \neq I$.

6. Let $T : R^n \longrightarrow R^n$ be a linear transformation such that $T \neq 0$ and $T^2 = 0$. Let $x \in R^n$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$. So, for any $\alpha_1, \alpha_2 \in F$, we have $T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 w_1 + \alpha_2 w_2$. Thus for any $\alpha_1, \alpha_2 \in F$, $T^{-1}(\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 v_1 + \alpha_2 v_2 = \alpha_1 T^{-1}(w_1) + \alpha_2 T^{-1}(w_2)$. Hence $T^{-1} : W \longrightarrow V$, defined as above, is a linear transformation.

Definition 4.1.9 (Inverse Linear Transformation) Let $T : V \longrightarrow W$ be a linear transformation. If the map T is one-one and onto, then the map $T^{-1} : W \longrightarrow V$ defined by

$$T^{-1}(w) = v \text{ whenever } T(v) = w$$

is called the inverse of the linear transformation T.

Example 4.1.10 1. Define $T : R^2 \longrightarrow R^2$ by $T((x, y)) = (x + y, x - y)$. Then $T^{-1} : R^2 \longrightarrow R^2$ is defined by

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

Note that $T \circ T^{-1}((x, y)) = T(T^{-1}((x, y))) = T\left(\left(\frac{x+y}{2}, \frac{x-y}{2}\right)\right)$

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

$$= (x, y).$$

Hence, $T \circ T^{-1} = I$, the identity transformation. Verify that $T^{-1} \circ T = I$. Thus, the map T^{-1} is indeed the inverse of the linear transformation T.

2. Recall the vector space $P_n(R)$ and the linear transformation $T : R^{n+1} \longrightarrow P_n(R)$ defined by $T((a_1, a_2, \ldots, a_{n+1})) = a_1 + a_2 x + \cdots + a_{n+1} x^n$ for $(a_1, a_2, \ldots, a_{n+1}) \in R^{n+1}$. Then $T^{-1} : P_n(R) \longrightarrow R^{n+1}$ is defined as $T^{-1}(a_1 + a_2 x + \cdots + a_{n+1} x^n) = (a_1, a_2, \ldots, a_{n+1})$ for $a_1 + a_2 x + \cdots + a_{n+1} x^n \in P_n(R)$. Verify that $T \circ T^{-1} = T^{-1} \circ T = I$. Hence, conclude that the map T^{-1} is indeed the inverse of the linear transformation T.

7. Let $T : R^n \longrightarrow R^m$ be a linear transformation, and let $x_0 \in R^n$ with $T(x_0) = y$. Consider the sets $S = \{x \in R^n : T(x) = y\}$ and $N = \{x \in R^n : T(x) = 0\}$. Show that for every $x \in S$ there exists $z \in N$ such that $x = x_0 + z$. 8. Define a map $T : C \longrightarrow C$ by $T(z) = \bar{z}$, the complex conjugate of z. Is T linear on 9. Find all functions $f : R^2 \longrightarrow R^2$ that satisfy the conditions

(a) $f((x, x)) = (x, x)$ and

(b) $f((x, y)) = (y, x)$ for all $(x, y) \in R^2$. That is, f fixes the line $y = x$ and sends the point (x_1, y_1) for $x_1 \neq y_1$ to its mirror image along the line $y = x$.

Is this function a linear transformation? Justify your answer.

Theorem 4.1.8 Let $T : V \longrightarrow W$ be a linear transformation. For $w \in W$, define the set $T^{-1}(w) = \{v \in V : T(v) = w\}$. Suppose that the map T is one-one and onto. 1. Then for each $w \in W$, the set $T^{-1}(w)$ is a set consisting of a single element. 2. The map $T^{-1} : W \longrightarrow V$ defined by

$$T^{-1}(w) = v \text{ whenever } T(v) = w.$$

is a linear transformation.

(a) $\mathbb{C}$ over $\mathbb{R}$ (b) $\mathbb{C}$ over $\mathbb{C}$.

4.2 Matrix of a linear transformation

In this section, we relate linear transformation over finite dimensional vector spaces with matrices. For this, we ask the reader to recall the results on ordered basis, studied in Section 3.4. Let V and W be finite dimensional vector spaces over the set F with respective dimensions m and n. Also, let $T : V \longrightarrow W$ be a linear transformation. Suppose $B_1 = (v_1, v_2, \ldots, v_n)$ is an ordered basis of V. In the last section, we saw that a linear transformation is determined by its image on a basis of the domain space. We therefore look at the images of the vectors $v_j \in B_1$ for $1 \le j \le n$. Now for each j, $1 \le j \le n$, the vectors $T(v_j) \in W$. We now express these vectors in terms of an ordered basis $B_2 = (w_1, w_2, \ldots, w_m)$ of W. So, for each j, $1 \le j \le n$, there exist unique scalars $a_{1j}, a_{2j}, \ldots, a_{mj} \in F$ such that

Remark 4.2.2 Let $B_1 = (v_1, v_2, \ldots, v_n)$ be an ordered basis of V and $B_2 = (w_1, w_2, \ldots, w_m)$ be an ordered basis of W. Let $T : V \longrightarrow W$ be a linear transformation with $A = T[B_1, B_2]$. Then the first column of A is the coordinate of the vector $T(v_1)$ in the basis B_2. In general, the ith column of A is the coordinate of the vector $T(v_i)$ in the basis B_2. We now give a few examples to understand the above discussion and the theorem.

Example 4.2.3 1. Let $T : R^2 \longrightarrow R^2$ be a linear transformation, given by

$$T(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m$$
$$T(v_2) = a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m$$
$$T(v_n) = a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m.$$

Or in short, $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$ for $1 \le j \le n$. In other words, for each j, $1 \le j \le n$, the coordinates of $T(v_j)$ with respect to the ordered basis B_2 is the column vector $[a_{1j}, a_{2j}, \ldots, a_{mj}]^t$. Equivalently,

$$[T(v_j)]_{B_2} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Let $[x]_{B_1} = [x_1, x_2, \ldots, x_n]^t$ be the coordinates of a vector $x \in V$. Then

$$T(x) = T\left(\sum_{j=1}^{n} x_j v_j\right) = \sum_{j=1}^{n} x_j T(v_j) = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij} w_i\right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) w_i.$$

Define a matrix A by

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

Then the coordinates of the vector $T(x)$ with respect to the ordered basis B_2 is

$$A[x]_{B_1} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

The matrix A is called the matrix of the linear transformation T with respect to the ordered bases B_1 and B_2, and is denoted by $T[B_1, B_2]$. We thus have the following theorem.

Theorem 4.2.1 Let V and W be finite dimensional vector spaces with dimensions n and m, respectively. Let $T : V \longrightarrow W$ be a linear transformation. If B_1 is an ordered basis of V and B_2 is an ordered basis of W, then there exists an $m \times n$ matrix $A = T[B_1, B_2]$ such that

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \text{for } 1 \le j \le n.$$

$[T(x)]_{B_2} =$

$= \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}$

$\cdots \; !_n$

$\cdots \cdots \cdots$

$j=1 \; a_{mj}x_j$

$[T(x)]_{B_2} = A\,[x]_{B_1}$.

$T((x, y)) = (x + y, x - y)$. We obtain $T[B_1, B_2]$, the matrix of the linear transformation T with respect to the ordered bases $B_1 = (1, 0),(0, 1)$ and $B_2 = (1, 1),(1, -1)$ of R^2. For any vector $(x, y) \in R^2$, $[(x, y)]_{B_1} = \begin{matrix} x \\ y \end{matrix}$ as $(x, y) = x(1, 0) + y(0, 1)$. Also, by definition of the linear transformation T, we have $T((1, 0)) = (1, 1) = 1 \cdot (1, 1) + 0 \cdot (1, -1)$. So, $[T((1, 0))]_{B_2} = (1, 0)^t$. $T((0, 1)) = (1, -1) = 0 \cdot (1, 1) + 1 \cdot (1, -1)$. That is, $[T((0, 1))]_{B_2} = (0, 1)^t$. So the $T[B_1, B_2] = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$. Observe that in this case,

$[T((x, y))]_{B_2} = [(x + y, x - y)]_{B_2} = x(1, 1) + y(1, -1) = \begin{matrix} x \\ y \end{matrix}$, and $\begin{matrix} 0 & 1 \end{matrix} \begin{matrix} x \\ y \end{matrix} = \begin{matrix} x \\ y \end{matrix} = [T((x, y))]_{B_2}$.

$T[B_1, B_2]\,[(x, y)]_{B_1} = \begin{matrix} 1 & 0 \end{matrix}$

2. Let $B_1 = (1, 0, 0),(0, 1, 0),(0, 0, 1)$, $B_2 = (1, 0, 0),(1, 1, 0),(1, 1, 1)$ be two ordered bases of R^3. $T : R^3 \longrightarrow R^3$ by $T(x) = x$.

$T((1, 0, 0)) = 1 \cdot (1, 0, 0) + 0 \cdot (1, 1, 0) + 0 \cdot (1, 1, 1)$,

$T((0, 1, 0)) = -1 \cdot (1, 0, 0) + 1 \cdot (1, 1, 0) + 0 \cdot (1, 1, 1)$, and $T((0, 0, 1)) = 0 \cdot (1, 0, 0) + (-1) \cdot (1, 1, 0) + 1 \cdot (1, 1, 1)$. Thus, we have $T[B_1, B_2] = [[T((1, 0, 0))]_{B_2}, [T((0, 1, 0))]_{B_2}, [T((0, 0, 1))]_{B_2}]$

and

$= [(1, 0, 0)^t, (-1, 1, 0)^t, (0, -1, 1)^t]$

$\begin{matrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{matrix}$. Similarly check that $T[B_1, B_1] = \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$.

Define

Then

$=$

4.3. RANK-NULLITY THEOREM 75 3. Let $T : R^3 \longrightarrow R^2$ be define by $T((x, y, z)) = (x + y - z, x + z)$. Let $B_1 = (1, 0, 0),(0, 1, 0),(0, 0, 1)$ and $B_2 = (1, 0),(0, 1)$ be the ordered bases of the domain and range space, respectively. Then $T[B_1, B_2] = \begin{matrix} 1 & 1 & -1 \\ 1 & 0 & 1 \end{matrix}$. Check that that $[T(x, y, z)]_{B_2} = T[B_1, B_2]\,[(x, y, z)]_{B_1}$. Exercise 4.2.4 Recall the space $P_n(R)$ (the vector space of all polynomials of degree less than or equal to n). We define a linear transformation $D : P_n(R) \longrightarrow P_n(R)$ by

$D(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$. Find the matrix of the linear transformation D. However, note that the image of the linear transformation is contained in $P_{n-1}(R)$. Remark 4.2.5 1. Observe that

$T[B_1, B_2] = [[T(v_1)]_{B_2}, [T(v_2)]_{B_2}, \ldots, [T(v_n)]_{B_2}]$.

2. It is important to note that

$[T(x)]_{B_2} = T[B_1, B_2]\,[x]_{B_1}$. That is, we multiply the matrix of the linear transformation with the coordinates $[x]_{B_1}$, of the vector $x \in V$ to obtain the coordinates of the vector $T(x) \in W$. 3. If A is an $m \times n$ matrix, then A induces a linear transformation $T_A : R^n \longrightarrow R^m$, defined by

76 CHAPTER 4. LINEAR TRANSFORMATIONS3. T is one-one $\Leftarrow\Rightarrow$ $N(T) = \{0\}$ is the zero subspace of V $\Leftarrow\Rightarrow$ $\{T(u_i) : 1 \le i \le n\}$ is a basis of 4. $\dim(R(T)) = \dim(V)$ if and only if $N(T) = \{0\}$. Proof. The results about R(T) and N (T) can be easily proved. We thus leave the proof for the readers. We now assume that T is one-one. We need to show that $N(T) = \{0\}$. Let $u \in N(T)$. Then by definition, $T(u) = 0$. Also for any linear transformation (see Proposition 4.1.3), $T(0) = 0$. Thus $T(u) = T(0)$. So, T is one-one implies $u = 0$. That is, $N(T) = \{0\}$. Let $N(T) = \{0\}$. We need to show that T is one-one. So, let us assume that for some $u, v \in V$, $T(u) = T(v)$. Then, by linearity of T, $T(u - v) = 0$. This implies, $u - v \in N(T) = \{0\}$. This in turn implies $u = v$. Hence, T is one-one. The other parts can be similarly proved. Remark 4.3.3 1. The space R(T) is called the range space of T and N(T) is called thenull2. We write $\rho(T) = \dim(R(T))$ and $\nu(T) = \dim(N(T))$. 3. $\rho(T)$ is called the rank of the linear transformation T and $\nu(T)$ is called the nullity of T. Example 4.3.4 Determine the range and null space of the linear transformation

R(T).

$T : R^3 \longrightarrow R^4$ with $T(x, y, z) = (x - y + z, y - z, x, 2x - 5y + 5z)$.

$T_A(x) = Ax$. We sometimes write A for T_A. Suppose that the standard bases for R^n and R^m are the ordered bases B_1 and B_2, respectively. Then observe that

$$T[B_1, B_2] = A.$$

4.3 Rank-Nullity Theorem

Definition 4.3.1 (Range and Null Space) Let V, W be finite dimensional vector spaces over the same set of scalars and $T : V \longrightarrow W$ be a linear transformation. We define 1. $R(T) = \{T(x) : x \in V\}$, and

$$R(T) = L(1, 0, 1, 2),(-1, 1, 0, -5),(1, -1, 0, 5)$$

1. (a) R(T) is a subspace of W.

(b) $R(T) = L(T(v_1), T(v_2), \ldots, T(v_n))$.

2. (a) N (T) is a subspace of V.

4.3. RANK-NULLITY THEOREM 77 2. Let $T : R^2 \longrightarrow R^3$ be defined by

$T(1, 0) = (1, 0, 0)$, $T(0, 1) = (1, 0, 0)$. Then the vectors (1, 0) and (0, 1) are linearly independent whereas $T(1, 0)$ and $T(0, 1)$ are linearly 3. Is there a linear transformation

$T : R^3 \longrightarrow R^2$ such that $T(1, -1, 1) = (1, 2)$, and $T(-1, 1, 2) = (1, 0)$?

(c) $\dim(R(T)) \le \dim(W)$.

4. Recall the vector space $P_n(R)$. Define a linear transformation

Solution: By Definition $R(T) = L(T(1, 0, 0), T(0, 1, 0), T(0, 0, 1))$. We therefore have

$= \{\alpha(1, 0, 1, 2) + \beta(1, -1, 0, 5) : \alpha, \beta \in R\}$

$= \{(\alpha + \beta, -\beta, \alpha, 2\alpha + 5\beta) : \alpha, \beta \in R\}$

$= \{(x, y, z, w) \in R^4 : x + y - z = 0, 5y - 2z + w = 0\}$. Also, by definition $N(T) = \{(x, y, z) \in R^3 : T(x, y, z) = 0\}$

2. $N(T) = \{x \in V : T(x) = 0\}$. Proposition 4.3.2 Let V and W be finite dimensional vector spaces and let $T : V \longrightarrow W$ be a linear transformation. Suppose that $(v_1, v_2, \ldots, v_n)$ is an ordered basis of V. Then

$= \{(x, y, z) \in R^3 : (x - y + z, y - z, x, 2x - 5y + 5z) = 0\}$

(b) dim(N (T)) ≤ dim(V).

$D(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$. Describe the null space and range space of D. Note that the range space is contained in the space 5. Let $T : R^3 \longrightarrow R^3$ be defined by

T (1, 0, 0) = (0, 0, 1), T(1, 1, 0) = (1, 1, 1) and T (1, 1, 1) = (1, 1, 0).
(a) Find T (x, y, z) for x, y, z ∈ R,
(b) Find R(T) and N (T). Also calculate ρ(T) and v(T).
(c) Show that T^3 = T and find the matrix of the linear transformation with respect to the standard
dependent.
basis. 6. Let $T : R^2 \longrightarrow R^2$ be a linear transformation with
T((3, 4)) = (0, 1), T((−1, 1)) = (2, 3).

Find the matrix representation T[B, B] of T with respect to the ordered basis B = (1, 0),(1, 1) of 7. Determine a linear transformation $T : R^3 \longrightarrow R^3$ whose range space is L{(1, 2, 0),(0, 1, 1),(1, 3, 1)}.

$D : P_n(R) \longrightarrow P_n(R)$
8. Suppose the following chain of matrices is given.
by
space of T.
= {(x, y, z) ∈ R^3: x − y + z = 0, y − z = 0,
x = 0, 2x − 5y + 5z = 0}
= {(x, y, z) ∈ R^3: y − z = 0, x = 0}
= {(x, y, z) ∈ R^3: y = z, x = 0}
= {(0, y, y) ∈ R^3: y arbitrary}
= L((0, 1, 1)) Exercise 4.3.5 1. Let T : V −→W be a linear transformation and let {T (v₁), T(v₂), . . . , T(vₙ)}be linearly independent in R(T). Prove that {v₁, v₂, . . . , vₙ} ⊂ V is linearly independent.

= L (1, 0, 1, 2),(1, −1, 0, 5)

78 CHAPTER 4. LINEAR TRANSFORMATIONSProof. Let dim(V) = n and dim(N (T)) = r. Suppose $\{u_1, u_2, \ldots, u_r\}$ is a basis of N(T). Since $\{u_1, u_2, \ldots, u_r\}$ is a linearly independent set in V, we can extend it to form a

basis of V(see Corollary 3.3.15). So, there exist vectors $\{u_{r+1}, u_{r+2}, \ldots, u_n\}$ such that $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ is a basis of V. Therefore, by Proposition 4.3.2 R(T) = L(T(u₁), T (u₂), . . . , T(uₙ))

= L(0, . . . , 0, T (u_{r+1}), T (u_{r+2}), . . . , T(uₙ))
= L(T(u_{r+1}), T (u_{r+2}), . . . , T(u_n)). We now prove that the set $\{T(u_{r+1}), T(u_{r+2}), \ldots, T(u_n)\}$ is linearly independent. Suppose the set is not linearly independent. Then, there exists scalars $\alpha_{r+1}, \alpha_{r+2}, \ldots, \alpha_n$, not all zero such

that $\alpha_{r+1}T(u_{r+1}) + \alpha_{r+2}T(u_{r+2}) + \cdots + \alpha_nT(u_n) = 0$. That is, $T(\alpha_{r+1}u_{r+1} + \alpha_{r+2}u_{r+2} + \cdots + \alpha_nu_n) = 0$. So, by definition of N (T), $\alpha_{r+1}u_{r+1} + \alpha_{r+2}u_{r+2} + \cdots + \alpha_nu_n \in N(T) = L(u_1, \ldots, u_r)$. Hence, there exists scalars α_i, 1 ≤ i ≤ r

such that $\alpha_{r+1}u_{r+1} + \alpha_{r+2}u_{r+2} + \cdots + \alpha_nu_n = \alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_ru_r$. That is, $\alpha_1u_1 + + \cdots + \alpha_ru_r - \alpha_{r+1}u_{r+1} - \cdots - \alpha_nu_n = 0$. But the set $\{u_1, u_2, \ldots, u_n\}$ is a basis of V and so linearly independent. Thus by definition of linear α_i

= 0 for all i, 1 ≤ i ≤ n. In other words, we have shown that $\{T(u_{r+1}), T(u_{r+2}), \ldots, T(u_n)\}$ is a basis of R(T). Hence, dim(R(T)) + dim(N (T)) = (n − r) + r = n = dim(V). Using the Rank-nullity theorem, we give a short proof of

the following result. Corollary 4.3.7 Let T : V −→V be a linear transformation on a finite dimensional vector space V. Then

$A \longrightarrow B_1 \longrightarrow B_2 \cdots \longrightarrow B_{k-1} \longrightarrow B_k \longrightarrow B$. T is one-one ⇐⇒ T is onto ⇐⇒T is invertible.
If row space of B is in the row space of B_k and the row space of $B_l$is in the row space of B_{l-1} for
$P_{n-1}(R)$.
2 ≤ l ≤ k then show that the row space of B is in the row space of A. We now state and prove the rank-nullity Theorem. This result also follows from Proposition 4.3.2. Theorem 4.3.6 (Rank Nullity Theorem) Let T : V

−→W be a linear transformation and V be a finite dimensional vector space. Then dim(R(T)) + dim(N(T)) = dim(V), or equivalently ρ(T) + v(T) = dim(V). R^2.

Proof. By Proposition 4.3.2, T is one-one if and only if N (T) = {0}. By the rank-nullity Theorem 4.3.6 N (T) = {0} is equivalent to the condition dim(R(T)) = dim(V). Or equivalently Tis onto. By definition, T is invertible if T

is one-one and onto. But we have shown that T is one-one if and only if T is onto. Thus, we have the last equivalent condition. Remark 4.3.8 Let V be a finite dimensional vector space and let T : V −→V be a linear

transformation. If either T is one-one or T is onto, then T is invertible. The following are some of the consequences of the rank-nullity theorem. The proof is left as an exercise for the reader.independence

4.3. RANK-NULLITY THEOREM 79 Corollary 4.3.9 The following are equivalent for an m × n real matrix A. 1. Rank (A) = k.

2. There exist exactly k rows of A that are linearly independent.
3. There exist exactly k columns of A that are linearly independent.
4. There is a k × k submatrix of A with non-zero determinant and every (k + 1) × (k + 1) submatrix of A has zero determinant. 5. The dimension of the range space of A is k.

6. There is a subset of R^m consisting of exactly k linearly independent vectors $b_1, b_2, \ldots, b_k$ such that the system Ax = $b_i$for 1 ≤ i ≤ k is consistent.

7. The dimension of the null space of A = n − k. Exercise 4.3.10 1. Let T : V −→W be a linear transformation. (a) If V is finite dimensional then show that the null space and the range space of T are also finite

dimensional. (b) If V and W are both finite dimensional then show that

i. if dim(V) < dim(W) then T is onto.
ii. if dim(V) > dim(W) then T is not one-one.
2. Let A be an m × n real matrix. Then
80 CHAPTER 4. LINEAR TRANSFORMATIONS(a) Show that R(T + S) ⊂ R(T) + R(S). Deduce that ρ(T + S) ≤ρ(T) +ρ(S). Hint: For two subspaces M, N of a vector space V, recall the definition of the vector subspace(b)

Use the above and the rank-nullity Theorem 4.3.6 to prove v(T +S) ≥v(T) +v(S) −n. 6. Let V be the complex vector space of all complex polynomials of degree at most n. Givenkdistinct complex numbers $z_1, z_2, \ldots, z_k$,

we define a linear transformation $T : V \longrightarrow C^k$ by T P(z) = P(z_1), P(z_2), . . . , P(z_k). For each k ≥ 1, determine the dimension of the range space of T. 7. Let A be an n × n real matrix with A^2 = A. Consider the linear

transformation T_A: $R^n \longrightarrow R^n$ defined by $T_A(v)$ = Av for all v ∈ R^n. Prove that

(a) $T_A \circ T_A = T_A$ (use the condition A^2 = A). (b) N (T_A) ∩ R(T_A) = {0}. Hint: Let x ∈ N(T_A) ∩ R(T_A). This implies $T_A(x) = 0$ and x =$T_A(y)$ for some y∈R^n. So, x = $T_A(y) = (T_A \circ T_A)(y) = T_A\, T_A(y) = T_A(x) = 0$.

(c) R^n = N (T_A) + R(T_A). Hint: Let {v₁, . . . , vₖ} be a basis of N (T_A). Extend it to get a basis {v₁, . . . ,vₖ, vₖ₊₁, . . . ,vₙ}of R^n. Then by Rank-nullity Theorem 4.3.6, {T_A(vₖ₊₁), . . . , T_A(vₙ)} is a basis of R(T_A). M + N.

(a) if n > m, then the system Ax = 0 has infinitely many solutions, (b) if n < m, then there exists a non-zero vector b = (b₁, b₂, . . . , bₘ)ᵗsuch that the system Ax = b
does not have any solution.

4.4 Similarity of Matrices

In the last few sections, the following has been discussed in detail: Given a finite dimensional vector space V of dimension n, we fixed an ordered basis B. For anyv∈V, we calculated the column vector $[v]_B$, to obtain the coordinates of v with respect to the ordered basis

3. Let A be an m × n matrix. Prove that Row Rank (A) = Column Rank (A). [Hint: Define $T_A : R^n \longrightarrow R^m$ by $T_A(v)$ = Av for all v ∈ R^n. Let Row Rank (A) = r. Use Theorem
2.6.1 to show, Ax = 0 has n − r linearly independent solutions. This implies,
$v(T_A)$ = dim({v ∈ R^n: $T_A(v) = 0$}) = dim({v ∈ R^n: Av = 0}) = n − r. Now observe that R(T_A) is the linear span of columns of A and use
the rank-nullity Theorem 4.3.6 to get the required result.]
4. Prove Theorem 2.6.1. [Hint: Consider the linear system of equation Ax = b with the orders of A,x and b, respectively
B. Also, for any linear transformation T : V −→V, we got an n × n matrix T[B, B], the matrix of Twithrespect to the ordered basis
B. That is, once an ordered basis of V is fixed, every linear transformationis represented by a matrix with entries from the scalars. In this section, we understand the matrix representation of T in terms of different bases $B_1$and$B_2$of V. That is, we relate the two n × n matrices T [B_1, B_1] and T [B_2, B_2]. We start with the followingimportant theorem. This theorem also enables us to understand why the matrix product is definedsomewhat differently. Theorem 4.4.1 (Composition of Linear Transformations) Let V, Wand Z be finite dimensional vec- tor spaces with ordered bases B_1, B_2, B_3, respectively. Also, let T : V −→Wand S : W−→Zbe linear transformations. Then the composition map S ◦ T : V −→Z is a linear transformation and

as m × n, n × 1 and m × 1. Define a linear transformation $T : R^n \longrightarrow R^m$ by T (v) = Av. First observe that if the solution exists then b is a linear combination of the columns of A and the linear span of the columns of A give us R(T). Note that ρ(A) = column rank(A) = dim(R(T)) = H(say). Then for part i) one can proceed as follows. i) Let $C_{i_1}, C_{i_2}, \ldots, C_i$ be the linearly independent columns of A. Then rank(A) < rank([A b])

$(S \circ T)\,[B_1, B_3] = S[B_2, B_3]\,T\,[B_1, B_2]$.

implies that $\{C_{i_1}, C_{i_2}, \ldots, C_i, b\}$ is linearly independent. Hence $b' \in L(C_{i_1}, C_{i_2}, \ldots, C_i)$. Hence, the system doesn't have any solution. On similar lines prove the other two parts.] 5. Let T, S : V −→V be linear transformations with dim(V) = n.

Now for $1 \le t \le n$, $(S \circ T)(u_t) = S(T(u_t)) = S$ $[(S \circ T)(u_t)]_{B_3} = ((S[B_2, B_3]\,T\,[B_1, B_2])_{1t}, \ldots, (S[B_2, B_3]\,T[B_1, B_2])_{pt})^t$.

$(S \circ T)\,[B_1, B_3] = [(S \circ T)(u_1)]_{B_3}, \ldots, [(S \circ T)(u_n)]_{B_3}) = S[B_2, B_3]\,T\,[B_1, B_2]$. This completes the proof. Proposition 4.4.2 Let V be a finite dimensional vector space and let T, S : V −→V be a linear transformations. Then $v(T) + v(S) \ge v(T \circ S) \ge \max\{v(T), v(S)\}$.

Proof. We first prove the second inequality. Suppose that $v \in N(S)$. Then $T \circ S(v) = T(S(v)) = T(0) = 0$. So, $N(S) \subset N(T \circ S)$. Therefore,

Proof. Let $B_1 = (u_1, u_2, \ldots, u_n)$, $B_2 = (v_1, v_2, \ldots, v_m)$ and $B_3 = (w_1, w_2, \ldots, w_p)$ be ordered basesof V, W

$$(T[B_1, B_2])_{jt} v_j = \sum_{j=1}^{m} (T[B_1, B_2])_{jt} S(v_j)$$

$v(S) \le v(T \circ S)$. Suppose dim(V) = n. Then using the rank-nullity theorem, observe that

$$(T[B_1, B_2])_{jt} = \sum_{j=1}^{m} (S[B_2, B_3])_{kj} w_k$$

$$v(T \circ S) \ge v(T) \Longleftrightarrow n - v(T \circ S) \le n - v(T) \Longleftrightarrow \rho(T \circ S) \le \rho(T).$$ So, to complete the proof of the second inequality, we need to show that $R(T \circ S) \subset R(T)$. This is true

$$= \sum_{j=1}^{p} (S[B_2, B_3])_{kj} (T[B_1, B_2])_{jt} w_k$$

$$= \sum_{k=1}^{p} (S[B_2, B_3]\,T[B_1, B_2])_{kt} w_k.$$

as $R(S) \subset V$. We now prove the first inequality.

Let k = v(S) and let $\{v_1, v_2, \ldots, v_k\}$ be a basis of N (S). Clearly, $\{v_1, v_2, \ldots, v_k\} \subset N(T \circ S)$ as

So,

T(0) = 0. We extend it to get a basis $\{v_1, v_2, \ldots, v_k, u_1, u_2, \ldots, u_0\}$ of $N(T \circ S)$. Claim: The set $\{S(u_1), S(u_2), \ldots, S(u_0)\}$ is linearly independent subset of N (T).

Hence,

As $u_1, u_2, \ldots, u_0 \in N(T \circ S)$, the set $\{S(u_1), S(u_2), \ldots, S(u_0)\}$ is a subset of N(T). Let if possible the given set be linearly dependent. Then there exist non-zero scalars $c_1, c_2, \ldots, c_0$ such that $c_1S(u_1) + c_2S(u_2) + \cdots + c_0S(u_0) = 0$.

Hence,

Let V be a vector space with dim(V) = n. Let $B_1 = (u_1, u_2, \ldots, u_n)$ and $B_2 = (v_1, v_2, \ldots, v_n)$ betwo ordered bases of V. Recall from Definition 4.1.5 that I : V −→V is the identity linear transformation defined by I(x) = x for every $x \in V$. Suppose $x \in V$ with $[x]_{B_1} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^t$ and $[x]_{B_2} = (\beta_1, \beta_2, \ldots, \beta_n)^t$. We now express each vector in B_2 as a linear combination of the vectors from B_1. Since $v_i \in V$, for $1 \le i \le n$, and B_1 is a basis of V, we can find scalars a_{ij}, $1 \le i, j \le n$ such that $a_{ji}u_j$ for all i, $1 \le i \le n$. Hence, $[I(v_i)]_{B_1} = [v_i]_{B_1} = (a_{1i}, a_{2i}, \cdots, a_{ni})^t$ and $v_i = I(v_i) = \sum_{j=1}^{n}$

$I[B_2, B_1] = [[I(v_1)]_{B_1}, [I(v_2)]_{B_1}, \ldots, [I(v_n)]_{B_1}]$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$c_iu_i \in N(S)$ and is a linear combination of the basis vectors $v_1, v_2, \ldots, v_k$ of N (S). Therefore, there exist scalars $\alpha_1, \alpha_2, \alpha_k$ such that

So, the vector $\sum_{i=1}^{k} \alpha_i v_i$. Or equivalently

$$\sum_{i=1}^{k} c_i u_i = \sum_{i=1}^{n} (-\alpha_i) v_i = 0.$$

$$\sum_{i=1}^{k} c_i u_i + \sum_{i=1}^{n} (-\alpha_i) v_i = 0.$$

Thus, we have proved the following result. Theorem 4.4.5 (Change of Basis Theorem) Let V be a finite dimensional vector space with ordered basesB1 = (u1, u2, . . . ,un} and B2 = (v1, v2, . . . ,vn}.

Suppose $x \in V$ with $[x]_{B_1} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^t$ and $[x]_{B_2} = (\beta_1, \beta_2, \ldots, \beta_n)^t$. Then

$$[x]_{B_1} = I[B_2, B_1]\,[x]_{B_2}.$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \end{pmatrix}$$

and therefore,

Note: Observe that the identity linear transformation I : V −→V defined by I(x) = x for every $x \in V$ is invertible and

$I[B_2, B_1]^{-1} = I^{-1}[B_1, B_2] = I[B_1, B_2]$. Therefore, we also have $[x]_{B_2} = I[B_1, B_2]\,[x]_{B_1}$.

Let V be a finite dimensional vector space and let B_1 and B_2 be two ordered bases of V. Let T : V −→V and Z, respectively. Then

$$(S \circ T)\,[B_1, B_3] = [[S \circ T(u_1)]_{B_3}, [S \circ T(u_2)]_{B_3}, \ldots, [S \circ T(u_n)]_{B_3}].$$

That is, the 0 vector is a non-trivial linear combination of the basis vectors $v_1, v_2, \ldots, v_k, u_1, u_2, \ldots, u_0$ of N (T ◦ S). A contradiction. Thus, the set $\{S(u_1), S(u_2), \ldots, S(u_0)\}$ is a linearly independent subset of N(T) and sov(T) ≥H. $v(T \circ S) = k + H \le v(S) + v(T)$. Recall from Theorem 4.1.8 that if T is an invertible linear Transformation, then T^{-1}: V−→Vis a linear transformation defined by $T^{-1}(u) = v$ whenever T (v) = u. We now state an important result about inverse of a linear transformation. The reader is required to supply the proof (use Theorem 4.4.1). Theorem 4.4.3 (Inverse of a Linear Transformation) Let V be a finite dimensional vector space with ordered bases B_1 and B_2. Also let T : V −→V be an invertible linear transformation. Then the matrix of Tand T^{-1} are related by

$$T\,[B_1, B_2]^{-1} = T^{-1}[B_2, B_1].$$

Exercise 4.4.4 For the linear transformations given below, find the matrix T[B, B]. 1. Let B = (1, 1, 1),(1, −1, 1),(1, 1, −1) be an ordered basis of R^3. Define $T : R^3 \longrightarrow R^3$ by T(1, 1, 1) = (1, −1, 1), T (1, −1, 1) = (1, 1, −1), and T (1, 1, −1) = (1, 1, 1). Is T an invertible linear transforma- tion? Give reasons. 2. Let $B = (1, x, x^2, x^3)$ be an ordered basis of $P_3(R)$. Define $T : P_3(R) \longrightarrow P_3(R)$ by T(1) = 1, T(x) = 1 + x, $T(x^2) = (1 + x)^2$, and $T(x^3) = (1 + x)^3$. Prove that T is an invertible linear transformation. Also, find $T^{-1}[B, B]$.

Hence $T[B_2, B_1] = BA$. Also, for each j, $1 \le j \le n$,

$T(v_j) = \sum_{k=1}^{n} c_{kj} I(v_k) = \sum_{k=1}^{n} c_{kj} (\sum_{0=1}^{n} a_{0k} u_0)$ and so

be a linear transformation. We are now in a position to relate the two matrices $T[B_1, B_1]$ and $T[B_2, B_2]$. Theorem 4.4.6 Let V be a finite dimensional vector space and let $B_1 = (u_1, u_2, \ldots, u_n)$ and $B_2 = (v_1, v_2, \ldots, v_n)$ be two ordered bases of V. Let $T : V \longrightarrow V$ be a linear transformation with $B = T[B_1, B_1]$ and $C = T[B_2, B_2]$ as matrix representations of T in bases B_1 and B_2. Also, let $A = [a_{ij}] = I[B_2, B_1]$, be the matrix of the identity linear transformation with respect to the

Proof. For any $x \in V$, we represent $[T(x)]_{B_2}$ in two ways. Using Theorem 4.2.1, the first expression is

Equivalently,

$$[T(x)]_{B_2} = T[B_2, B_2][x]_{B_2}. \quad (4.4.1)$$

Using Theorem 4.4.5, the other expression is

$$[T(x)]_{B_2} = I[B_1, B_2][T(x)]_{B_1}$$

$$= I[B_1, B_2]\, T[B_1, B_1][x]_{B_1}$$

Let V be a vector space with $\dim(V) = n$, and let $T : V \longrightarrow V$ be a linear transformation. Then foreach ordered basis B of V, we get an $n \times n$ matrix $T[B, B]$. Also, we know that for any vector space wehave infinite number of choices for an ordered basis. So, as we change an ordered basis, the matrix of the linear transformation changes. Theorem 4.4.6 tells

$$= I[B_1, B_2]\, T[B_1, B_1]\, I[B_2, B_1][x]_{B_2}. \quad (4.4.2)$$

Hence, using (4.4.1) and (4.4.2), we see that for every $x \in V$, $I[B_1, B_2]\, T[B_1, B_1]\, I[B_2, B_1][x]_{B_2} = T[B_2, B_2][x]_{B_2}$. Since the result is true for all $x \in V$, we get $I[B_1, B_2]\, T[B_1, B_1]\, I[B_2, B_1] = T[B_2, B_2]$. (4.4.3) That is, $A^{-1}BA = C$ or equivalently $ACA^{-1} = B$. Another

Proof: Let $B = [b_{ij}]$ and $C = [c_{ij}]$. Then for $1 \le i \le n$, $T(u_i) = \sum_{k=1}^{n} b_{ki} u_k$ and $T(v_j) = \sum_{k=1}^{n} c_{kj} v_k = \sum_{k=1}^{n} c_{kj} (\sum_{0=1}^{n} a_{0k} u_0)$

bases B_1 and B_2. Then $BA = AC$. Equivalently $B = ACA^{-1}$. $c_{ij} v_j$. So, for each j, $1 \le j \le n$,

its columns are linearly independent and hence we can take its columns as an ordered basis B_1. Then note that $B = T_A[B_1, B_1]$. The above observations lead to the following remark and the definition. Remark 4.4.7 The identity (4.4.3) shows how the matrix representation of a linear transformation T changes if the ordered basis used to compute the matrix representation is changed. Hence, the matrix

$I[B_1, B_2]$ is called the $B_1 : B_2$ change of basis matrix. Definition 4.4.8 (Similar Matrices) Two square matrices B and C of the same order are said to be similar if there exists a non-singular matrix P such that $B = PCP^{-1}$ or equivalently $BP = PC$. Remark 4.4.9 Observe that if $A = T[B, B]$ then

$$[T(v_j)]_{B_1} = \sum_{i=1}^{n} a_{ij} T(u_i) = I[B_2, B_1] = [[I(1 + x - x^2)]_{B_1}, [I(1 + 2x + x^2)]_{B_1}, [I(2 + x + x^2)]_{B_1}]$$

$$= [[1 + x - x^2]_{B_1}, [1 + 2x + x^2]_{B_1}, [2 + x + x^2]_{B_1}]$$

$$[2 + x + x^2]_{B_1} = 1 \cdot 1 + 0 \cdot (1 + x) + 1 \cdot (1 + x + x^2) = (1, 0, 1)^t.$$

Example 4.4.10 1. Consider $P_2(R)$, with ordered bases $B_1 = 1, 1 + x, 1 + x + x^2$ and $B_2 = 1 + x - x^2$, $1 + 2x + x^2, 2 + x + x^2$. Then $[1 + x - x^2]_{B_1} = 0 \cdot 1 + 2 \cdot (1 + x) + (-1) \cdot (1 + x + x^2) = (0, 2, -1)^t$, $[1 + 2x + x^2]_{B_1} = (-1) \cdot 1 + 1 \cdot (1 + x) + 1 \cdot (1 + x + x^2) = (-1, 1, 1)^t$, and

k=1 =1 =1 k=1

2 1 0

−1 1 1 □□ □.

Find the matrices T [B_1, B_1] and T [B_2, B_2]. Also verify that

T [B_2,B_2] = I[B_1, B_2] T [B_1, B_1] I[B_2, B_1]

2. Consider two bases B_1 = (1, 0, 0),(1, 1, 0),(1, 1, 1) and B_2 = (1, 1, −1),(1, 2, 1),(2, 1, 1) of R^3. Suppose T : R^3−→$R^3$is a linear transformation defined by

T((x, y, z)) = (x + y, x + y + 2z, y − z).

T[B_1, B_1] = □□ □0 0 −2

1 1 4 0 1 0 □□ □, and T[B_2, B_2] = □□ □ −4/5 1 8/5 −2/5 2 9/5

k=1

{S^{-1}AS : S is n × n invertible matrix }

is the set of all matrices that are similar to the given matrix A. Therefore, similar matrices are just different matrix representations of a single linear transformation. 86 CHAPTER 4. LINEAR TRANSFORMATIONS(c) Let A be an n × n matrix with the property that A^{n-1}= 0 but A^n =0. Then prove that Aissimilar to the matrix given above. 2. Let T : R^3−→R^3be a linear

Find I[B_1, B_2] and verify,

8/5 0 −1/5 □□ □ .

I[B_1, B_2] T[B_1, B_1] I[B_2, B_1] = T [B_2, B_2]. Check that,

Therefore,

T[B_1, B_1] I[B_2, B_1] = I[B_2, B_1] T [B_2, B_2] = □□ □ 2 −2 −2 −2 4 5

2 1 0 □□ □ . Exercise 4.4.11 1. Let V be an n-dimensional vector space and let T : V −→V be a linear transformation.

□

=

Suppose T has the property that T^{n-1}= 0 but T^n = 0. (a) Then prove that there exists a vector u ∈ V such that the set

{u, T (u), . . . , T^{n-1}(u)}

= I^{-1}[B_2, B_1] T [B_1, B_1] I[B_2, B_1].

is a basis of V.

(b) Let B = (u, T(u), . . . , T^{n-1}(u)). Then prove that

0 0 0 · · · 0

1 0 0 · · · 0

0 1 0 · · · 0

T [B, B] = □□ □

Then

0 0 · · · 1 0 □□□ □.

□

.. ..

. . . .

R^3.

transformation given by T ((x, y, z)) = (x + y + 2z, x − y − 3z, 2x + 3y +z).

Let B be the standard basis and B_1 = (1, 1, 1),(1, −1, 1),(1, 1, 2) be another ordered basis. (a) Find the matrices T [B, B] and T [B_1, B_1].

(b) Find the matrix P such that P^{-1}T [B, B] P = T[B_1, B_1]. 3. Let T : R^3−→R^3be a linear transformation given by

T((x, y, z)) = (x, x + y, x + y + z).

Let B be the standard basis and B_1 = (1, 0, 0),(1, 1, 0),(1, 1, 1) be another ordered basis. (a) Find the matrices T [B, B] and T [B_1, B_1].

(b) Find the matrix P such that P^{-1}T [B, B] P = T[B_1, B_1]. 4. Let B_1 = (1, 2, 0),(1, 3, 2),(0, 1, 3) and B_2 = (1, 2, 1),(0, 1, 2),(1, 4, 6) be two ordered bases of (a) Find the change of basis matrix P from B_1 to B_2. (b) Find the change of basis matrix Q from B_2 to B_1. (c) Verify that PQ = I = QP.

(d) Find the change of basis matrix from the standard basis of $R^3$to B_1. What do you notice?

Chapter 5

Inner Product Spaces

We had learned that given vectors J i and J j (which are at an angle of 90˚) in a plane, any vector in the plane is a linear combination of the vectors J i and J j. In this section, we investigate a method by which any basis of a finite dimensional vector can be transferred to another basis in such a way that the vectors in the new basis are at an angle of 90˚to each other. To do this, we start by defining a notion of inner product (dot product) in a vector space. This helps us in finding out whether two vectors are at 90˚ or not.

5.1 Definition and Basic Properties

In R^2, given two vectors x = (x_1, x_2), y = (y_1, y_2), we know the inner product x·y = x_1y_1 +x_2y_2. Note that for any x, y, z ∈ $R^2$and α ∈ R, this inner product satisfies the conditions

x · (y + αz) = x · y + αx · z, x · y = y · x, and x · x ≥ 0 and x · x = 0 if and only if x = 0. Thus, we are motivated to define an inner product on an arbitrary vector space. Definition 5.1.1 (Inner Product) Let V (F) be a vector space over F. An inner product over V (F), denoted by (, %, is a map,

such that for u, v, w ∈ V and a, b ∈ F

2. (u, v% = (v, u%, the complex conjugate of (u, v%, and
3. (u, u% ≥ 0 for all u ∈ V and equality holds if and only if u = 0. Definition 5.1.2 (Inner Product Space) Let V be a vector space with an inner product (, %. Then (V,(, %) is called an inner product space, in short denoted by ips. Example 5.1.3 The first two examples given below are called the standard inner product or the dot 1. Let V = R^nbe the real vector space of dimension n. Given two vectors u = (u_1, u_2, . . . , u_n) and v = (v_1, v_2, . . . , v_n) of V, we define

(u, v% = u_1v_1 + u_2v_2 + · · · + u_nv_n = uv^t.

Verify (, % is an inner product.

(, % : V × V −→ F

1. (au + bv, w% = a(u, w% + b(v, w%,

Or, in other words and the proof of the inequality is over. Observe that if u '= 0 then the equality holds if and only of λu + v = 0 for λ = −(v,u%

by &u& = (u, u%, the positive square root. A very useful and a fundamental inequality concerning the inner product is due to Cauchy andSchwartz. The next theorem gives the statement and a proof of this inequality. Theorem 5.1.7 (Cauchy-Schwartz inequality) Let V (F) be an inner product space. Then for anyu, v∈ |(u, v%| ≤ &u& &v&. The equality holds if and only if the vectors u and v are linearly dependent. Further, if u'=0, then Proof. If u = 0, then the inequality holds. Let u '= 0. Note that (λu +v, λu +v% ≥0 for all λ∈F. In particular, for λ = −(v, u% &u&^{2}, we get

0 ≤ (λu + v, λu + v%

V

90 CHAPTER 5. INNER PRODUCT SPACES4. Let V be a complex vector space with dim(V) = n. Fix an ordered basis B = $v = (v, u\&\% u$

$(u_1, u_2, \ldots, u_n)$. Define a map

$\&u\&\cdot$

$\&u\&^2$. That is, u

and v are linearly dependent. We leave it for the reader to prove

product on $R^n$and C^n, respectively..

$\&u\&^2 \&u\&^2 \_(v, u\%$

$(v, u\%$

$\&u\&^2(v, u\% + \&v\&^2$

$\&u\&^2(u,v\% -$

$v = (v, u\&\% u\&\cdot$ Definition 5.1.8 (Angle between two vectors) Let V be a real vector space. Then for every u, v ∈ V, by the

Cauchy-Schwartz inequality, we have

$\&u\&$

is an inner product in R^3(R). With respect to this inner product, find the angle between the vectors

87

$= \&v\&^2 \_|(v, u\%|^2$

$\&u\&\ \&v\& \le 1$. We know that cos : [0, π] −→ [−1, 1] is an one-one and onto function. Therefore, for every real number $\&u\&\ \&v\&\cdot$

$\&u\&^2$.

there exists a unique θ, 0 ≤ θ ≤ π, such that

6. Consider the set $M_{n\times n}$(R) of all real square matrices of order n. For A, B∈$M_{n\times n}$(R) we define(A, B% = tr(AB^t). Then

$\&u\&\ \&v\&\cdot$

$(A + B, C\% = tr\,(A + B)C^t = tr(AC^t) + tr(BC^t) = (A, C\% + (B, C\%. (A, B\% = tr(AB^t) = tr(\,(AB^t)^t) =$

1. The real number θ with 0 ≤ θ ≤ π and satisfying cos θ =$(u, v\%$

$tr(BA^t) = (B, A\%.$

$\&u\&\ \&v\&$is called the angle between the two

Let A = (a_{ij}). Then

$|(v, u\%|^2 \le \&u\&^2\&v\&^2$

2. The vectors u and v in V are said to be orthogonal if (u, v% = 0.

$(A, A\% = tr(AA^t) = \sum_{i=1}^{n} a^{2i}$ and therefore, (A, A% > 0 for all non-zero matrices A. So, it is clear that (A, B% is an inner product on

whenever $[u]_B = (a_1, a_2, \ldots, a_n)^t$ and $[v]_B = (b_1, b_2, \ldots, b_n)^t$. Show that the above defined map isindeed an inner product. 5.

$(\ ,\ \% : V \times V \longrightarrow C$ by $(u, v\% = \sum_{i=1}^{n} a_i b_i$

Let x = (x_1, x_2, x_3), y = $(y_1, y_2, y_3) \in R^3$. Show that

7. Let V be the real vector space of all continuous functions with domain [−2π, 2π]. That is, V= C[−2π, 2π]. Then show that V is an inner product space with inner product $\int_{-1}^{1} f(x)g(x)dx$. For different values of m and n, find the angle between the functions cos(mx) and sin(nx).

$= \lambda\lambda\&u\&^2 + \lambda(u, v\% + \lambda(v, u\% + \&v\&^2$

$(x, y\% = 10x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2 + x_2y_3 + x_3y_2 + x_3y_3$

$=(v, u\%$

$(v, u\%$

3. A set of vectors $\{u_1, u_2, \ldots, u_n\}$ is called mutually orthogonal if $(u_i, u_j\%$ = 0 for all 1 ≤ i '= j ≤ n. Exercise 5.1.9 1. Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis of R^n. Then prove that with respect to the

standard inner product on R^n, the vectors $e_i$satisfy the following:

(a) $\&e_i\&$ = 1 for 1 ≤ i ≤ n.

(b) $(e_i, e_j\%$ = 0 for 1 ≤ i '= j ≤ n.

2. Recall the following inner product on R^2: for x = $(x_1, x_2)^t$ and y = $(y_1, y_2)^t$,

$-1 \le (u, v\%$

$(x, y\% = 4x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2.$

(a) Find the angle between the vectors $e_1 = (1, 0)^t$ and $e_2 = (0, 1)^t$.

$(u, v\%$

(b) Let u = $(1, 0)^t$. Find v ∈ R^2such that (v, u% = 0.

(c) Find two vectors x, y ∈ R^2, such that &x& = &y& = 1 and (x, y% = 0.

$\cos\theta = (u, v\%$

3. Find an inner product in R^2such that the following conditions hold:

$\&(1, 2)\& = \&(2, -1)\& = 1$, and $((1, 2), (2, -1)\% = 0$.

[Hint: Consider a symmetric matrix A = $\begin{matrix} a & b \\ b & c \end{matrix}$. Define $(x, y\% = y^tAx$ and solve a system of 3

equations for the unknowns a, b, c.]

8. Let V be an inner product space. Prove that

$\&u + v\& \le \&u\& + \&v\&$ for every u, v ∈ V. This inequality is called the triangle inequality.

(1, 1, 1) and (2, −5, 2).

9. Let $z_1, z_2, \ldots, z_n \in C$. Use the Cauchy-Schwartz inequality to prove that $|z_1 + z_2 + \cdots + z_n| \le n(|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2)$. When does the equality hold? 10. Let x, y ∈ R^n. Observe that (x, y% = (y, x%. Hence or

otherwise prove the following: (a) (x, y% = 0 ⇐⇒ $\&x - y\&^2 = \&x\&^2 + \&y\&^2$, (This is called Pythagoras Theorem). (b) &x& = &y& ⇐⇒ (x + y, x − y% = 0, (x and y form adjacent sides of a rhombus as the diagonals

$(AA^t)_{ii} = \sum_{i=1}^{n} a_{ij}a_{ij} = \sum_{j=1}^{n} a_{ij}a_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}a_{ij}$

x + y and x − y are orthogonal).

(c) $\&x + y\&^2 + \&x - y\&^2 = 2\&x\&^2 + 2\&y\&^2$, (This is called the Parallelogram Law). (d) $4(x, y\% = \&x + y\&^2 - \&x - y\&^2$(This is called the polarisation identity). Remark 5.1.10 i. Suppose the norm of a vector is given. Then,

the polarisation identity can be used to define an inner product.$M_{n\times n}$(R).

5.1. DEFINITION AND BASIC PROPERTIES 91 ii. Observe that if (x,y% = 0 then the parallelogram spanned by the vectors x and y is a rectangle. The above equality tells us that the lengths of the two diagonals are

equal. Are these results true if x, y ∈ C^n(C)? 11. Let x, y ∈ C^n(C). Prove that

(b) If x '= 0 then $\&x + ix\&^2 = \&x\&^2 + \&ix\&^2$, even though (x, ix% '= 0.

(c) If $\&x + y\&^2 = \&x\&^2 + \&y\&^2$and $\&x + iy\&^2 = \&x\&^2 + \&iy\&^2$then show that (x, y% = 0. 12. Let V be an n-dimensional inner product space, with an inner product (, %. Let u ∈ V be a fixed vector with &u& = 1. Then give

reasons for the following statements. (a) Let $S^\perp = \{v \in V : (v, u\% = 0\}$. Then S is a subspace of V of dimension n − 1. (b) Let 0 '= α ∈ F and let S = {v ∈ V : (v, u% = α}. Then S is not a

subspace of V.

(c) For any v ∈ S, there exists a vector $v_0 \in S^\perp$, such that $v = v_0 + \alpha u$. Theorem 5.1.11 Let V be an inner product space. Let $\{u_1, u_2, \ldots, u_n\}$ be a set of non-zero, mutually orthogonal vectors of V. 1. Then the set $\{u_1, u_2, \ldots, u_n\}$ is linearly independent.

3. Let dim(V) = n and also let $\&u_i\&$ = 1 for i = 1, 2, . . . , n. Then for any v ∈ V, In particular, $(v, u_i\%$ = 0 for all i = 1, 2, . . . , n if and only if v = 0.

(a) $4(x, y\% = \&x + y\&^2 - \&x - y\&^2 + i\&x + iy\&^2 - i\&x - iy\&^2$. Proof. Consider the set of non-zero, mutually orthogonal vectors $\{u_1, u_2, \ldots, u_n\}$. Suppose there exist

scalars $c_1, c_2, \ldots, c_n$ not all zero, such that

$c_1u_1 + c_2u_2 + \cdots + c_nu_n = 0$. Then for 1 ≤ i ≤ n, we have $0 = (0, u_i\% = (c_1u_1 + c_2u_2 + \cdots + c_nu_n, u_i\% = \sum_{j=1}^{n}$ as $(u_j, u_i\% = 0$ for all j '= i and $(u_i, u_i\% = 1$. This gives a contradiction to our assumption that some of the c_i's are

non-zero. This establishes the linear independence of a set of non-zero, mutually orthogonal $||\cdot||$

Figure 5.1: Gram-~~Schmidt~~ Process

Theorem 5.2.1 (~~Gram-Schmidt Orthogonalization Process)~~

Let V ~~be an inner product~~ space. Suppose

vectors. For the second part, using $(u_i, u_j\% = \begin{cases} 0 & \text{if } i \neq j \\ \&u_i\&^2 & \text{if } i = j \end{cases}$ for 1 ≤ i, j ≤ n, we have

2. $\&\ \int_{i=1}^{n} \alpha_i u_i\&^2 = \int_{i=1}^{n} |\alpha_i|^2\&u_i\&^2;$

$_{i=1}$

$v = {}^{*n}$
$(v, u_i\%u_i.$
$_{i=1}$

5.2. GRAM-SCHMIDT ORTHOGONALISATION PROCESS 93
$c_j (u_j , u_i\% = c_i$
$_{j=1}$

$\&^{*n}$
$_{i=1}\ \alpha_i u_i\&^2 = (^{*n}$
$_{i=1}\ \alpha_i u_i,^{*n}$
$_{i=1}\ \alpha_i u_i\% = {}^{*n}$
$_{i=1}\ \alpha_i (u_i,^{*n}$
$_{j=1}\ \alpha_j u_j\%$
$= {}^{*n}$
$_{i=1}\ \alpha_i^{*n}$
$_{j=1}\ \alpha_j (u_i, u_j\% = {}^{*n}$
$_{i=1}\ \alpha_i\alpha_i(u_i, u_i\%$
$= {}^{*n}$
$|\alpha_i|^2\&u_i\&^2.$
$_{i=1}$

$\{u_1, u_2, . . . ,u_n\}$ is ~~a set of line~~arly ~~independent vectors of~~ V. ~~Then there exists a set~~ $\{v_1, v_2, . . . , v_n\}$ of vectors of V satisfy~~ing the follo~~wing: 1. $\&v_i\& = 1$ for $1 \le i \le n$,

2. $(v_i, v_j\% = 0$ for $1 \le i, j \le n, i \ne j$ and
3. $L(v_1, v_2, . . . ,v_i) = L(u_1, u_2, . . . ,u_i)$ for $1 \le i \le n$.
Proof. We successively define the ~~vectors~~ $v_1, v_2, . . . ,v_n$ ~~as follows.~~

v

<v,u>

v
uu

Calculate $w_2 = u_2 - (u_2, v_1\%v_1$, ~~and let~~ $v_2 = w_2$
92 CHAPTER 5. INNER PRODUCT SPACESFor the third part, observe from the first part, the linear independence of the non-zero mutually orthogonal vectors $u_1, u_2, . . . , u_n$. Since dim(V) = n, they form a basis of V.

Thus, for every vector $v \in V$, there exist scalars α_i, $1 \le i \le n$, such that $v = !^n_{i=1}\ \alpha_i u_n$. Hence, $_{i=1}\ \alpha_i(u_i, u_j\% = \alpha_j$. Therefore, we have obtained the required result. Definition 5.1.12 (Orthonormal Set) Let V be an inner product space. A set of non-zero, mutually orthogonal vectors $\{v_1, v_2, . . . , v_n\}$ in V is called an orthonormal set if $\&v_i\& = 1$ for $i = 1, 2, . . . , n$. If the set $\{v_1, v_2, . . . ,v_n\}$ is also a basis of V, then the set of vectors $\{v_1, v_2, . . . , v_n\}$ is called an orthonormal basis of V. Example 5.1.13 1. Consider the vector space R^2 with the standard inner product. Then the standard ordered basis B = (1, 0),(0, 1) is an orthonormal set. Also, the basis $B_1 = \frac{1}{\sqrt{2}}(1, 1), \frac{1}{\sqrt{2}}(1, -1)$ is an orthonormal set. 2. Let R^n be endowed with the standard inner product. Then by Exercise 5.1.9.1, the standard ordered basis $(e_1, e_2, . . . , e_n)$ is an orthonormal set. In view of Theorem 5.1.11, we inquire into the question of extracting an orthonormal basis from a given basis. In the section, we describe a process (called the Gram-Schmidt Orthogonalisation process) that generates an orthonormal set from a given set containing finitely many vectors. Remark 5.1.14 The last part of the above theorem can be rephrased as "suppose $\{v_1, v_2, . . . ,v_n\}$ is an orthonormal basis of an inner product space V. Then for each $u \in V$ the numbers $(u, v_i\%$ for $1 \le i \le n$ are the coordinates of u with respect to the above basis". That is, let $B = (v_1, v_2, . . . , v_n)$ be an ordered basis. Then for any $u \in V$,

$(v, u_j\% = (^{*n}$
$_{i=1}\ \alpha_i u_i, u_j\% = {}^{*n}$

$[u]_B = ((u, v_1\%,(u, v_2\%, . . . ,(u, v_n\%)^t.$

5.2 Gram-Schmidt Orthogonalisation Process

Let V be a finite dimensional inner product space. Suppose $u_1,u_2, . . . ,u_n$ is a linearly independent subset of V. Then the Gram-Schmidt orthogonalisation process uses the vectors $u_1, u_2, . . . ,u_n$ to construct new vectors $v_1, v_2, . . . ,v_n$ such that $(v_i, v_j\% = 0$ for $i \ne j$, $\&v_i\& = 1$ and Span $\{u_1, u_2, . . . ,u_i\}$=Span $\{v_1, v_2, . . . , v_i\}$ for $i = 1, 2, . . . , n$. This process proceeds with the following idea. Suppose we are given two vectors u and v in a plane. If we want to get vectors z and y such that z is a unit vector in the direction of u and y is a unit vector perpendicular to z, then they can be obtained in the following way: Take the first vector $z = \frac{u}{\&u\&}$. Let θ be the angle between the vectors u and v. Then $\cos(\theta) = \frac{(u, v\%}{\&u\&\ \&v\&}$. Defined $\alpha = \&v\& \cos(\theta) = \frac{(u, v\%}{\&u\&}$ = (z, v%. Then $w = v - \alpha z$ is a vector perpendicular to the unit vector z, as we have removed the component of z from v. So, the vectors that we are interested in are $\frac{w}{\&w\&}$. This idea is used to give the Gram-Schmidt Orthogonalization process which we now describe. 94 CHAPTER 5. INNER PRODUCT SPACES3. $L(v_1, v_2, . . . ,v_i) = L(u_1, u_2, . . . ,u_i)$ for $1 \le i \le k$. Now, let us assume that we are given a set of n linearly independent vectors $\{u_1, u_2, . . . , u_n\}$ of V. Then by the inductive assumption, we already have vectors $v_1, v_2, . . . , v_{n-1}$ satisfying 1. $\&v_i\& = 1$ for $1 \le i \le n - 1$,

2. $(v_i, v_j\% = 0$ for $1 \le i \ne j \le n - 1$, and
3. $L(v_1, v_2, . . . ,v_i) = L(u_1, u_2, . . . ,u_i)$ for $1 \le i \le n - 1$. Using (5.2.1), we define

z and y = w

$w_n = u_n - (u_n, v_1\%v_1 - (u_n, v_2\%v_2 - \cdots - (u_n, v_{n-1}\%v_{n-1}$. (5.2.2) We first show that $w_n \ne 0$ and hence $v_n = \frac{w_n}{\&w_n\&}$ is well defined. On the contrary, assume that $w_n \in L(v_1, v_2, . . . ,v_{n-1})$. Then there exist scalars $\alpha_1, \alpha_2, . . . , \alpha_{n-1}$ such that $w_n = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{n-1}v_{n-1}$. So, by (5.2.2) $u_n = \alpha_1 + (u_n, v_1\% v_1 + \alpha_2 + (u_n, v_2\% v_2 + \cdots + (\alpha_{n-1} + (u_n, v_{n-1}\% v_{n-1}$. Thus, by the third induction assumption, $u_n \in L(v_1, v_2, . . . ,v_{n-1}) = L(u_1, u_2, . . . ,u_{n-1})$. This gives a contradiction to the given assumption that the set of vectors $\{u_1, u_2, . . . , u_n\}$ is linear

u

Obtain $w_3 = u_3 - (u_3, v_1\%v_1 - (u_3,v_2\%v_2$, and let $v_3 = \frac{w_3}{\&w_3\&}$. In general, if $v_1, v_2, v_3, v_4, . . . , v_{i-1}$ are already obtained, we compute $w_i = u_i - (u_i, v_1\%v_1 - (u_i, v_2\%v_2 - \cdots - (u_i, v_{i-1}\%v_{i-1}$, (5.2.1) and define $v_i = \frac{w_i}{\&w_i\&}$. We prove the theorem by induction on n, the number of linearly independent vectors. For n = 1, we have $v_1 = u_1$

&w_n&. Then &v_n& = 1. Also, it can be easily verified that (v_n, v_i% = 0 for1 ≤ i ≤ n − 1. Hence, by the principle of mathematical induction, the proof of the theorem is complete. We illustrate the Gram-Schmidt process by the following example. Example 5.2.2 Let {(1, −1, 1, 1),(1, 0, 1, 0),(0, 1, 0, 1)} be a linearly independent set in R^4(R). Find an orthonormal set {v_1, v_2, v_3} such that L((1, −1, 1, 1),(1, 0, 1, 0),(0, 1, 0, 1)) =L(v_1, v_2, v_3).

So, w_n '= 0. Define v_n = w_n

Solution: Let u_1 = (1, 0, 1, 0). Define v_1 =(1, 0, 1, 0)

√2. Let u_2 = (0, 1, 0, 1). Then

w_2 = (0, 1, 0, 1) − ((0, 1, 0, 1), (1, 0, 1, 0)

&u_1&^{2} = 1. Hence, the result holds for n = 1. Let the result hold for all k ≤ n − 1. That is, suppose we are given any set of k, 1 ≤ k ≤ n − 1 linearly independent vectors {u_1, u_2, . . . ,u_k} of V. Then by the inductive

assumption, there exists a set {v_1, v_2, . . . ,v_k} of vectors satisfying the following:

v_1 =u_1

&u_1&.

1. &v_i& = 1 for 1 ≤ i ≤ k,
2. (v_i, v_j% = 0 for 1 ≤ i '= j ≤ k, and

√2%v_1 = (0, 1, 0, 1).

√2. Let u_3 = (1, −1, 1, 1). Then

w_3 = (1, −1, 1, 1) − (1,−1, 1, 1), (1, 0, 1, 0) √2 v_1 − (1, −1, 1, 1), (0, 1, 0, 1) √2 v_2= (0, −1, 0, 1)

&w_2&.

Hence, v_2 =(0, 1, 0, 1)

&u_1&. Since u_1 '= 0, v_1 '= 0 and &u_1&% =(u_1, u_1%

and v_3 =(0, −1, 0, 1) √2.

&v_1&^{2} = (v_1, v_1% = (u_1 &u_1&, u_1

independent.

5.2. GRAM-SCHMIDT ORTHOGONALISATION PROCESS 95 Remark 5.2.3 1. Let {u_1, u_2, . . . , u_k} be any basis of a k-dimensional subspace W of R^n. Then by Gram-Schmidt orthogonalisation process, we get an orthonormal set {v_1, v_2, . . . ,v_k} ⊂ R^n with

96 CHAPTER 5. INNER PRODUCT SPACES Note that,

v_1 ^{2} v_1, v_2 · · · v_1, v_k

W = L(v_1, v_2, . . . ,v_k), and for 1 ≤ i ≤ k,

L(v_1, v_2, . . . ,v_i) = L(u_1, u_2, . . . , u_i).

2. Suppose we are given a set of n vectors, {u_1, u_2, . . . ,u_n} of V that are linearly dependent. Then by Corollary 3.2.5, there exists a smallest k, 2 ≤ k ≤ n such that L(u_1, u_2, . . . , u_k) = L(u_1, u_2, . . . ,$u_{k−1}$). We claim that in

this case, w_k = 0. Since, we have chosen the smallest k satisfying

L(u_1, u_2, . . . ,u_i) = L(u_1, u_2, . . . ,$u_{i−1}$),

for 2 ≤ i ≤ n, the set {u_1, u_2, . . . , $u_{k−1}$} is linearly independent (use Corollary 3.2.5). So, by Theorem 5.2.1, there exists an orthonormal set {v_1, v_2, . . . ,$v_{k−1}$} such that L(u_1,u_2, . . . ,$u_{k−1}$) = L(v_1, v_2, . . . , $v_{k−1}$).

As u_k ∈ L(v_1, v_2, . . . , $v_{k−1}$), by Remark 5.1.14

u_k = (u_k, v_1%v_1 + (u_k, v_2%v_2 + · · · + (u_k, $v_{k−1}$%$v_{n−1}$. So, by definition of w_k, w_k = 0. Therefore, in this case, we can continue with the Gram-Schmidt process by replacing u_k by u_{k+1}.

[v_1, v_2, . . . ,v_k]

v_2, v_1 v_2 ^{2}. · · v_2, v_k

v_k, v_1 v_k, v_2 · · · v_k ^{2}

1 0 · · · 0

0 1 · · · 0

0 0 · · · 1 = I_k. Or using (5.2.3), in the language of matrices, we get α_{1k} α_{2k} · · · α_{nk} α_{11} α_{12} · · · α_{1k} α_{21} α_{22} · · · α_{2k}

α_{11} α_{21} · · · α_{n1}

α_{12} α_{22} · · · α_{n2}

α_{n1} α_{n2} · · · α_{nk} =I_k. Perhaps the readers must have noticed that the inverse of A is its transpose. Such matrices are calledorthogonal matrices and they have a special role to play. Definition 5.2.4 (Orthogonal Matrix)

A n × n real matrix A is said to be an orthogonal matrix if A A^t= It is worthwhile to solve the following exercises. Exercise 5.2.5 1. Let A and B be two n × n orthogonal matrices. Then prove that ABandBAareboth

orthogonal matrices. 2. Let A be an n × n orthogonal matrix. Then prove that (a) the rows of A form an orthonormal basis of R^n.

v^t_1

v^t_2

A^tA =

=

..

.

v^t_k

3. Let S be a countably infinite set of linearly independent vectors. Then one can apply the Gram- Schmidt process to get a countably infinite orthonormal set. 4. Let {v_1, v_2, . . . ,v_k} be an orthonormal subset of R^n. Let B = (e_1, e_2, . . . , e_n) be the standard ordered basis of

R^n. Then there exist real numbers α_{ij} , 1 ≤ i ≤ k, 1 ≤ j ≤ n such that $[v_i]_B$ = (α_{1i}, α_{2i}, . . . , α_{ni})^t.

Let $A = [v_1, v_2, \ldots, v_k]$. Then in the ordered basis B, we have

$$\begin{matrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1k} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2k} \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nk} \end{matrix}$$

is an $n \times k$ matrix. Also, observe that the conditions $\|v_i\| = 1$ and $(v_i, v_j) = 0$ for $1 \le i \neq j \le n$, implies that $1 = \|v_i\| = \|v_i\|^2 = (v_i, v_i) = 1^n$ and $0 = (v_i, v_j) = 1^n$

$$= \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix}$$

(b) the columns of A form an orthonormal basis of R^n.
(c) for any two vectors $x, y \in R^{n\times 1}$, $(Ax, Ay) = (x, y)$. (d) for any vector $x \in R^{n\times 1}$, $\|Ax\| = \|x\|$.

Theorem 5.2.6 (QR Decomposition) Let A be a square matrix of order n. Then there exist matrices Q and R such that Q is orthogonal and R is upper triangular with A = QR.

In case, A is non-singular, the diagonal entries of R can be chosen to be positive. Also, in this case, the decomposition is unique.

$$A = \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix}$$

Proof. We prove the theorem when A is non-singular. The proof for the singular case is left as an exercise. Let the columns of A be $x_1, x_2, \ldots, x_n$. The Gram-Schmidt orthogonalisation process applied to the vectors $x_1, x_2, \ldots, x_n$ gives the vectors $u_1, u_2, \ldots, u_n$ satisfying

$$L(u_1, u_2, \ldots, u_i) = L(x_1, x_2, \ldots, x_i),$$

$$\|u_i\| = 1, \ (u_i, u_j) = 0, \quad \text{for } 1 \le i \neq j \le n. \tag{5.2.4}$$

Now, consider the ordered basis $B = (u_1, u_2, \ldots, u_n)$. From (5.2.4), for $1 \le i \le n$, we have $L(u_1, u_2, \ldots, u_i) = L(x_1, x_2, \ldots, x_i)$. So, we can find scalars α_{ji}, $1 \le j \le i$ such that

$$x_i = \alpha_{1i}u_1 + \alpha_{2i}u_2 + \cdots + \alpha_{ii}u_i = (\alpha_{1i}, \ldots, \alpha_{ii}, 0 \ldots, 0)^t_B. \tag{5.2.5}$$

3. R is an $r \times k$ matrix with rank (R) = r.

Example 5.2.8 1. Let $A = \begin{matrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 1 \end{matrix}$. Find an orthogonal matrix Qand an upper triangular matrix R such that A = QR. Solution: From Example 5.2.2, we know that

$$v_1 = \frac{1}{\sqrt{2}}(1, 0, 1, 0), \ v_2 = \frac{1}{\sqrt{2}}(0, 1, 0, 1), \ v_3 = \frac{1}{\sqrt{2}}(0, -1, 0, 1). \tag{5.2.6}$$

We now compute w_4. If we denote $u_4 = (2, 1, 1, 1)^t$then by the Gram-Schmidt process,

$$= \frac{1}{2}(1, 0, -1, 0)^t. \tag{5.2.7}$$

Thus, using (5.2.6) and (5.2.7), we get

$$= [x_1, x_2, \ldots, x_n] = A.$$

Thus, we see that A = QR, where Q is an orthogonal matrix (see Remark 5.2.3.4) and R is an upper triangular matrix. The proof doesn't guarantee that for $1 \le i \le n$, α_{ii} is positive. But this can be achieved by replacing the vector u_i by $-u_i$ whenever α_{ii} is negative.

1. Observe the following properties of upper triangular matrices.
1. The inverse of an upper triangular matrix is also an upper triangular matrix, and
2. product of upper triangular matrices is also upper triangular. Thus the matrix $R_2R_1^{-1}$ is an upper triangular matrix. Also, by Exercise 5.2.5.1, the matrix $Q_2Q_1^{-1}$ is an orthogonal matrix. Hence, by Exercise 5.2.5.4, $R_2R_1^{-1} = I_n$. So, $R_2 = R_1$ and therefore $Q_2 = Q_1$.

Suppose we have matrix $A = [x_1, x_2, \ldots, x_k]$ of dimension $n\times k$ with rank (A) = r. Then by Remark 5.2.3.2, the application of the Gram-Schmidt orthogonalisation process yields a set $\{u_1, u_2, \ldots, u_r\}$ of orthonormal vectors of R^n. In this case, for each i, $1 \le i \le r$, we have $L(u_1, u_2, \ldots, u_i) = L(x_1, x_2, \ldots, x_j)$, for some j, $i \le j \le k$. Hence, proceeding on the lines of the above theorem, we have the following result.

Theorem 5.2.7 (Generalised QR Decomposition) Let A be an $n \times k$ matrix of rank r. ThenA=QR,

$$= \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix}$$

where

1. Q is an $n \times r$ matrix with $Q^tQ = I_r$. That is, the columns of Q form an orthonormal set,
2. If $Q = [u_1, u_2, \ldots, u_r]$, then

$$L(u_1, u_2, \ldots, u_r) = L(x_1, x_2, \ldots, x_k), \text{ and } u_i = \sum_{j=1}^{n} \alpha_{ji}e_j, \text{ for } 1 \le i \le n.$$

Prove that $A^tA = I_n$. Hence deduce that A is an orthogonal matrix.

3. Let $\{u_1, u_2, \ldots, u_n\}$ be an orthonormal basis of R^n. Let $B = (e_1, e_2, \ldots, e_n)$ be the standard basis of R^n. Construct an $n \times n$ matrix A by

$$A = [u_1, u_2, \ldots, u_n] = \begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{matrix}$$

$$A^tA = \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix}$$

$A^tA = I_n$.

4. Let A be an $n \times n$ upper triangular matrix. If A is also an orthogonal matrix, then prove that $A=I_n$.

Q = % v1, v2, v3, v4) =
0 √ 1 2 − √ 1 2 0
0 √ 1 2 √ 1 2 0
and
R =

of R^4.

5.2. GRAM-SCHMIDT ORTHOGONALISATION PROCESS 99 Let $u_1 = (1, -1, 1, 1)$. Define $v_1 = u_1$. Let $u_2 = (1, 0, 1, 0)$. Then

$w_2 = (1, 0, 1, 0) - (u_2, v_1\%v_1 = (1, 0, 1, 0) - v_1 = \frac{1}{2}(1, 1, 1, -1)$.
Hence, $v_2 = (1, 1, 1, -1)$. Let $u_3 = (1, -2, 1, 2)$. Then
$w_3 = u_3 - (u_3, v_1\%v_1 - (u_3, v_2\%v_2 = u_3 - 3v_1 + v_2 = 0$. So, we again take $u_3 = (0, 1, 0, 1)$. Then 2 1 3 0

Q = [v1, v2, v3] = 1 2 1 2 0 −1 2 1 2 1 2 √
0 1 −1 0
0 0 0 √ 2 . The readers are advised to check the following: (a) rank (A) = 3, (b) A = QR with $Q^tQ = I_3$, and

(c) R a 3 × 4 upper triangular matrix with rank (R) = 3. Exercise 5.2.9 1. Determine an orthonormal basis of R^4 containing the vectors (1, −2, 1, 3) and (2, 1, −3, 1). 2. Prove that the polynomials 1, x, $\frac{3}{2}x^2 - \frac{1}{2}$, $\frac{5}{2}x^3 - \frac{3}{2}x$

form an orthogonal set of functions in the inner product space C[−1, 1] with the inner product (f, g% = # 1 −1 f(t)g(t)dt. Find the corresponding functions, f(x) with &f (x)& = 1.

3. Consider the vector space C[−π, π] with the standard inner product defined in the above exercise. Find an orthonormal basis for the subspace spanned by x, sin x and sin(x + 1).

100 CHAPTER 5. INNER PRODUCT SPACES

5.3 Orthogonal Projections and Applications

Recall that given a k-dimensional vector subspace of a vector space V of dimension n, one can always find an (n − k)-dimensional vector subspace W_0 of V (see Exercise 3.3.19.9) satisfying $W + W_0 = V$ and $W \cap W_0 = \{0\}$. The subspace W_0 is called the complementary subspace of Win V. We now define an important class of linear transformations on an inner product space, called orthogonal projections. Definition 5.3.1 (Projection Operator) Let V be an n-dimensional vector space and let Wbe a k-dimensional subspace of V. Let W_0 be a complement of W in V. Then we define a map P_W: V−→V $P_W(v) = w$, whenever $v = w + w_0$, $w \in W$, $w_0 \in W_0$. The map P_W is called the projection of V onto W along W_0. Remark 5.3.2 The map P is well defined due to the following reasons: 1. $W + W_0 = V$ implies that for every $v \in V$, we can find $w \in W$ and $w_0 \in W_0$ such that $v = w + w_0$. 2. $W \cap W_0 = \{0\}$ implies that the expression $v = w + w_0$ is unique for every $v \in V$. The next proposition states that the map defined above is a linear transformation from V to V. We omit the proof, as it follows directly from the above remarks. Proposition 5.3.3 The map P_W: V −→ V defined above is a linear transformation. Example 5.3.4 Let $V = R^3$ and $W = \{(x, y, z) \in R^3: x + y - z = 0\}$. 1. Let $W_0 = L((1, 2, 2))$. Then $W \cap W_0 = \{0\}$ and $W + W_0 = R^3$.

Also, for any vector $(x, y, z) \in R^3$, note that $(x, y, z) = w + w_0$, where $w = (z - y, 2z - 2x - y, 3z - 2x - 2y)$, and $w_0 = (x + y - z)(1, 2, 2)$. So, by definition,

4. Let M be a subspace of R^n and dim M = m. A vector $x \in R^n$ is said to be orthogonal to M if $(x, y\% = 0$ for every $y \in M$.

(a) How many linearly independent vectors can be orthogonal to M?

(b) If $M = \{(x_1, x_2, x_3) \in R^3: x_1 + x_2 + x_3 = 0\}$, determine a maximal set of linearly independent vectors orthogonal to M in R^3.

5. Determine an orthogonal basis of vector subspace spanned by

6. Let $S = \{(1, 1, 1, 1), (1, 2, 0, 1), (2, 2, 4, 0)\}$. Find an orthonormal basis of L(S) in R^4.

7. Let R^n be endowed with the standard inner product. Suppose we have a vector $x^t = (x_1, x_2, \ldots, x_n) \in R^n$, with &x& = 1. Then prove the following: (a) the set {x} can always be extended to form an orthonormal basis of R^n.

(b) Let this basis be $\{x, x_2, \ldots, x_n\}$. Suppose $B = (e_1, e_2, \ldots, e_n)$ is the standard basis of R^n. Let $A = [[x]_B, [x_2]_B, \ldots, [x_n]_B]$. Then prove that A is an orthogonal matrix. 8. Let $v, w \in R^n$, $n \geq 1$ with &u& = &w& = 1. Prove that there exists an orthogonal matrix A such that Av = w. Prove also that A can be chosen such that det(A) = 1.

5.3. ORTHOGONAL PROJECTIONS AND APPLICATIONS 101

102 CHAPTER 5. INNER PRODUCT SPACES

Theorem 5.3.10 Let S be a subset of a finite dimensional inner product space V, with inner product (, %.

1. $S^\perp$ is a subspace of V.
2. Let S be equal to a subspace W. Then the subspaces W and $W^\perp$ are complementary. Moreover, if $w \in W$ and $u \in W^\perp$, then (u, w% = 0 and $V = W + W^\perp$.

Let $\{w_1, w_2, \ldots, w_k\}$ be a basis of W. By Gram-Schmidt orthogonalisation process, we get an orthonormal basis, say, $\{v_1, v_2, \ldots, v_k\}$ of W. Then, for any $v \in V$,

W_0 is unique if we put an additional condition that $W_0 = \{v \in V : (v, w\% = 0$ for all $w \in W\}$. We now prove some basic properties about projection maps. Theorem 5.3.6 Let W and W_0 be complementary subspaces of a vector space V. Let P_W: V −→ V be a projection operator of V onto W along W_0. Then

1. the null space of P_W, $N(P_W) = \{v \in V : P_W(v) = 0\} = W_0$. 2. the range space of P_W, $R(P_W) = \{P_W(v) : v \in V\} = W$.

Proof. We leave the prove of the first part for the reader. The prove of the second part is as follows: Let dim(V) = n and dim(W) = k.

3. $P^2_W = P_W$. The condition $P^2_W = P_W$ is equivalent to $P_W(I - P_W) = 0 = (I - P_W)P_W$.

Proof. We only prove the first part of the theorem.
subspace W_0.

Let $w_0 \in W_0$. Then $w_0 = 0 + w_0$ for $0 \in W$. So, by definition, $P(w_0) = 0$. Hence, $W_0 \subset N(P_W)$. Also, for any $v \in V$, let $P_W(v) = 0$ with $v = w + w_0$ for some $w_0 \in W_0$ and $w \in W$. Then by definition $0 = P_W(v) = w$. That is, $w = 0$ and $v = w_0$. Thus, $v \in W_0$. Hence $N(P_W) = W_0$. Exercise 5.3.7 1. Let A be an n × n real matrix with $A^2 = A$. Consider the linear transformation

So, $V \subset W + W^\perp$. Also, for any $v \in W \cap W^\perp$, by definition of $W^\perp$, 0 =(v, v% =&v&^{2}. So, v=0. That

$T_A : R^n \longrightarrow R^n$, defined by $T_A(v) = Av$ for all $v \in R^n$. Prove that

(a) $T_A \circ T_A = T_A$ (use the condition $A^2 = A$).

(b) $N(T_A) \cap R(T_A) = \{0\}$. Hint: Let $x \in N(T_A) \cap R(T_A)$. This implies $T_A(x) = 0$ and $x = T_A(y)$ for some $y \in R^n$. So, $x = T_A(y) = (T_A \circ T_A)(y) = T_A\, T_A(y) = T_A(x) = 0$. (c) $R^n = N(T_A) + R(T_A)$. Hint: Let $\{v_1, \ldots, v_k\}$ be a basis of $N(T_A)$. Extend it to get a basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ of R^n. Then by Rank-nullity Theorem 4.3.6, $\{T_A(v_{k+1}), \ldots, T_A(v_n)\}$ is a basis of $R(T_A)$.

$W^\perp$ is called the orthogonal complement of W in V. Exercise 5.3.12 1. Let $W = \{(x, y, z) \in R^3 : x + y + z = 0\}$. Find $W^\perp$ with respect to the standard inner product. 2. Let W be a subspace of a finite dimensional inner product space V . Prove that $(W^\perp)^\perp = W$.

W_0. Recall that the first three parts of this exercise was also given in Exercise 4.3.10.7. 2. Find all 2×2 real matrices A such that $A^2 = A$. Hence or otherwise, determine all projection operators of R^2. The next result uses the Gram-Schmidt orthogonalisation process to get the complementary subspace in such a way that the vectors in different subspaces are orthogonal. Definition 5.3.8 (Orthogonal Subspace of a Set) Let V be an inner product space. Let S be a non-empty subset of V . We define

$S^\perp = \{v \in V : (v, s) = 0 \text{ for all } s \in S\}$. Example 5.3.9 Let V = R. 1. $S = \{0\}$. Then $S^\perp = R$.

2. S = R, Then $S^\perp = \{0\}$.
3. Let S be any subset of R containing a non-zero real number. Then $S^\perp = \{0\}$.

Then P_W is called the orthogonal projection of V onto W along $W^\perp$. Definition 5.3.14 (Self-Adjoint Transformation/Operator) Let V be an inner product space with inner

$$v - \sum_{i=1}^{k} (v, v_i) v_i \in W^\perp.$$

product (,). A linear transformation $T : V \longrightarrow V$ is called a self-adjoint operator if $(T(v), u) = (v, T(u))$ for every $u, v \in V$. Example 5.3.15 1. Let A be an n×n real symmetric matrix. That is, $A^t = A$. Then show that the linear transformation $T_A : R^n \longrightarrow R^n$ defined by $T_A(x) = Ax$ for every $x^t \in R^n$ is a self-adjoint operator. Solution: By definition, for every $x^t, y^t \in R^n$,

$(T_A(x), y) = (y)^t Ax = (y)^t A^t x = (Ay)^t x = (x, T_A(y))$. Hence, the result follows. 2. Let A be an n × n Hermitian matrix, that is, $A^* = A$. Then the linear transformation $T_A : C^n \longrightarrow C^n$ defined by $T_A(z) = Az$ for every $z^t \in C^n$ is a self-adjoint operator. Remark 5.3.16 1. By Proposition 5.3.3, the map P_W defined above is a linear transformation. by

5.3. ORTHOGONAL PROJECTIONS AND APPLICATIONS 103 3. Let $u, v \in V$ with $u = u_1 + u_2$ and $v = v_1 + v_2$ for some $u_1, v_1 \in W$ and $u_2, v_2 \in W^\perp$. Then we know that $(u_i, v_j) = 0$ whenever $1 \le i \ne j \le 2$. Therefore, for every $u, v \in V$, $(u, P_W(v))$. Thus, the orthogonal projection operator is a self-adjoint operator. 4. Let $v \in V$ and $w \in W$. Then $P_W(w) = w$ for all $w \in W$. Therefore, using Remarks 5.3.16.2 and 5.3.16.3, we get

$$(v - P_W(v), w) = (I - P_W(v), P_W(w)) = (P_W(I - P_W(v)), w) = (0(v), w) = 0, \quad w = 0$$

for every $w \in W$.

104 CHAPTER 5. INNER PRODUCT SPACES

ordered basis $B_2 = (e_1, e_2, \ldots, e_n)$ of R^n. Therefore, if $v_i = \sum_{j=1}^{n} a_{ji} e_j$, for $1 \le i \le k$, then

$$[v]_{B_2} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}_{n \times k}.$$

Then as observed in Remark 5.2.3.4, $A^t A = I_k$. That is, for $1 \le i, j \le k$,

$$a_{si} a_{sj} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \ne j. \end{cases} \quad (5.3.1)$$

Thus, using the associativity of matrix product and (5.3.1), we get

5. In particular, $(v - P_W(v), P_W(v) - w) = 0$ as $P_W(v) \in W$. Thus, $(v - P_W(v), P_W(v) - w) = 0$, for every $w \in W$. Hence, for any $v \in V$ and $w \in W$, we have

$$\|v - w\|^2 = \|v - P_W(v) + P_W(v) - w\|^2 = \|v - P_W(v)\|^2 + \|P_W(v) - w\|^2 + 2(v - P_W(v), P_W(v) - w) = \|v - P_W(v)\|^2 + \|P_W(v) - w\|^2.$$

2. $P_W^2 = P_W$, $(I - P_W)P_W = 0 = P_W(I - P_W)$.

$(P_W(u), v) = (u_1, v) = (u_1, v_1 + v_2) = (u_1, v_1) = (u_1 + u_2, v_1)$ and the equality holds if and only if $w = P_W(v)$. Since $P_W(v) \in W$, we see that

$$d(v, W) = \inf\{\|v - w\| : w \in W\} = \|v - P_W(v)\|.$$

That is, $P_W(v)$ is the vector nearest to $v \in W$. This can also be stated as: the vector $P_W(v)$ solves

and

the following minimisation problem:

$$(AA^t)(v) = A \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1i} & a_{2i} & \cdots & a_{ni} \end{bmatrix} \begin{bmatrix} (v, v_1) \\ (v, v_2) \\ \vdots \\ (v, v_i) \end{bmatrix} = \sum_{i=1}^{k} a_{1i} (v, v_i) \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1k} \end{bmatrix} + \sum_{i=1}^{k} a_{2i} (v, v_i) \begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2k} \end{bmatrix} + \cdots + \sum_{i=1}^{k} a_{ni} (v, v_i) \begin{bmatrix} a_{n1} \\ a_{n2} \\ \vdots \\ a_{nk} \end{bmatrix}$$

$$[P_W(v)]_{B_2} = \begin{bmatrix} \sum_{i=1}^{k} a_{1i} (v, v_i) \\ \sum_{i=1}^{k} a_{2i} (v, v_i) \\ \vdots \\ \sum_{i=1}^{k} a_{ni} (v, v_i) \end{bmatrix}_{n \times 1}$$

ordered basis of W. Define an n×k matrix $A = [v_1, v_2, \ldots, v_k]$. Then the matrix of the linear transformation P_W in the standard orthogonal ordered basis $(e_1, e_2, \ldots, e_n)$ is AA^t.

5.3.1 Matrix of the Orthogonal Projection

The minimization problem stated above arises in lot of applications. So, it will be very helpful if the matrix of the orthogonal projection can be obtained under a given basis. To this end, let W be a k-dimensional subspace of R^n with $W^\perp$ as its orthogonal complement. Let $B_1 = (v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n)$ of R^n. Then by Theorem 5.1.11, for any $v \in R^n$, $v = \sum_{i=1}^{n} (v, v_i) v_i$. Therefore,

$$\|v\|^2 = \sum_{i=1}^{k} (v, v_i)^2 + \sum_{i=1}^{n-k} (v, v_{k+i})^2.$$

Thus, by definition, $P_W(v) = \sum_{i=1}^{k} (v, v_i) v_i$. Let $A = [v_1, v_2, \ldots, v_k]$. Consider the standard orthogonal

$P_W : R^n \longrightarrow R^n$ be the orthogonal projection of R^n onto W. Suppose, we are given an orthonormal basis $B = (v_1, v_2, \ldots, v_k)$ of W. Under the assumption that B is known, we explicitly give the matrix of

P_W with respect to an extended ordered basis of R^n. Let us extend the given ordered orthonormal basis B of W to get an

orthonormal ordered basis

$$= [P_W(v)]_{B_2}.$$

Thus $P_W[B_2, B_2] = AA^t$. Thus, we have proved the following theorem. Theorem 5.3.17 Let W be a k-dimensional subspace of R^n and let $P_W : R^n \longrightarrow R^n$ be the orthogonal projection of R^n onto W along $W^\perp$. Suppose, $B = (v_1, v_2, \ldots, v_k)$ is an orthonormal

n

5.3. ORTHOGONAL PROJECTIONS AND APPLICATIONS 105 Example 5.3.18 Let W = {(x, y, z, w) ∈ R^4: x = y, z = w} be a

subspace of W. Then an orthonormal ordered basis of W is $\frac{1}{\sqrt{2}}(1, 1, 0, 0), \frac{1}{\sqrt{2}}(0, 0, 1, 1)$,

$\inf_{w \in W} \&v - w\& = \&v - P_W(v)\&$.

106 CHAPTER 5. INNER PRODUCT SPACES

and that of $W^{\perp}$ is $\frac{1}{\sqrt{2}}(1, -1, 0, 0), \frac{1}{\sqrt{2}}(0, 0, 1, -1)$. Therefore, if $P_W : R^4 \longrightarrow R^4$ is an orthogonal projection of R^4 onto W along $W^{\perp}$, then

the corresponding matrix A is given by

. Hence, the matrix of the orthogonal projection P_W in the ordered basis $P_W[B, B] = AA^t = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

It is easy to see that

1. the matrix $P_W[B, B]$ is symmetric,

3. $I_4 - P_W[B, B]\, P_W[B, B] = 0 = P_W[B, B]\, I_4 - P_W[B, B]$. Also, for any (x, y, z, w) ∈ R^4, we have

$$[(x, y, z, w)]_B = \left(x + y\sqrt{\tfrac{1}{2}}, z + w\sqrt{\tfrac{1}{2}}, x - y\sqrt{\tfrac{1}{2}}, z - w\sqrt{\tfrac{1}{2}}\right)^t.$$

$$A = \begin{pmatrix} \sqrt{\tfrac{1}{2}} & 0 \\ 0 & \sqrt{\tfrac{1}{2}} \end{pmatrix}$$

Thus, $P_W(x, y, z, w) = \frac{x + y}{2}(1, 1, 0, 0) + \frac{z + w}{2}(0, 0, 1, 1)$ is the closest vector to the subspace W for any

vector (x, y, z, w) ∈ R^4. Exercise 5.3.19 1. Show that for any non-zero vector $v^t \in R^n$, the rank of the matrix vv^t is 1.

$B = \frac{1}{\sqrt{2}}(1, 1, 0, 0), \frac{1}{\sqrt{2}}(0, 0, 1, 1), \frac{1}{\sqrt{2}}(1, -1, 0, 0), \frac{1}{\sqrt{2}}(0, 0, 1, -1)$

2. Let W be a subspace of a vector space V and let P : V −→ V be the orthogonal projection of V onto W along $W^{\perp}$. Let B be an orthonormal ordered basis of V. Then prove that corresponding matrix

is

3. Let A be an n × n matrix with $A^2 = A$ and $A^t = A$. Consider the associated linear transformation $T_A : R^n \longrightarrow R^n$ defined by $T_A(v) = Av$ for all $v^t \in R^n$. Then prove that there exists a subspace W of R^n such that T_A

is the orthogonal projection of R^n onto W along $W^{\perp}$. 4. Let W_1 and W_2 be two distinct subspaces of a finite dimensional vector space V. Let P_{W_1} and P_{W_2} be the corresponding orthogonal projection operators of V

along $W^{\perp}_1$ and $W^{\perp}_2$, respectively. Then by constructing an example in R^2, show that the map $P_{W_1} \circ P_{W_2}$ is a projection but not an orthogonal

5. Let W be an (n−1)-dimensional vector subspace of R^n and let $W^{\perp}$ be its orthogonal complement. Let

$B = (v_1, v_2, \ldots, v_{n-1}, v_n)$ be an orthogonal ordered basis of R^n with $(v_1, v_2, \ldots, v_{n-1})$ an ordered basis of W. Define a map

$T : R^n \longrightarrow R^n$ by $T(v) = w_0 - w$

whenever v = w + w_0 for some w ∈ W and $w_0 \in W^{\perp}$. Then (a) prove that T is a linear transformation, (b)

find the matrix, T[B, B], and

(c) prove that T[B, B] is an orthogonal matrix.

T is called the reflection along $W^{\perp}$.

2. $P_W[B, B]^2 = P_W[B, B]$, and

satisfies $P[B, B]^t = P[B, B]$.

projection.