

# Linear Algebra

Aakash Jog

## Contents

<b>I</b>	<b>General Information</b>	<b>6</b>
1	Contact Information	6
2	Grades	6
<b>II</b>	<b>Fields</b>	<b>7</b>
1	Definition	7
1.1	Examples of Fields . . . . .	7
1.2	Examples of Non-fields (Rings) . . . . .	8
2	Examples	8
<b>III</b>	<b>Matrices</b>	<b>9</b>
1	Definition	9
2	Addition of Matrices	9
2.0.1	Properties . . . . .	9
3	Multiplication of a matrix by a scalar	9
4	Multiplication of matrices	10
5	Zero Divisor	10
6	Theorem ('Good properties of matrix multiplication')	11

<b>7</b>	<b>Square Matrices</b>	<b>12</b>
7.1	Diagonal Matrices . . . . .	12
7.1.1	Proof . . . . .	12
7.2	Upper-triangular Matrices . . . . .	12
7.3	Lower-triangular Matrices . . . . .	12
7.4	Theorem . . . . .	13
7.4.1	Proof . . . . .	13
7.5	Identity Matrix . . . . .	13
7.6	Theorem . . . . .	13
7.6.1	Proof . . . . .	13
7.7	Inverse of Matrix . . . . .	14
7.7.1	If $AB = I_n$ and $CA = I_n$ , then $B = C$ . . . . .	14
7.7.2	Inverse of a Matrix . . . . .	14
7.7.3	If $AB = I$ , then $BA = I$ . . . . .	15
7.7.4	If $A$ is invertible, then $A$ <u>cannot</u> be a zero divisor. . . . .	15
7.7.5	If $A$ and $B$ are invertible, then $A + B$ may or may not be invertible. . . . .	15
7.7.6	If $A$ and $B$ are invertible, then $AB$ must be invertible. . . . .	16
<b>8</b>	<b>Transpose of a Matrix</b>	<b>16</b>
8.1	Properties of $A^t$ . . . . .	16
<b>9</b>	<b>Adjoint Matrix</b>	<b>16</b>
9.0.1	Properties of Adjoint Matrices . . . . .	17
<b>10</b>	<b>Row Operations on Matrices</b>	<b>17</b>
10.1	Elementary Row Operations . . . . .	17
10.2	Theorems . . . . .	17
10.2.1	$E_I A$ = the matrix obtained from $A$ by an elementary row operation I . . . . .	17
10.2.2	$E_{II} A$ = the matrix obtained from $A$ by an elementary row operation II . . . . .	18
10.2.3	$E_{III} A$ = the matrix obtained from $A$ by an elementary row operation III . . . . .	18
10.2.4	All elementary matrices are invertible, moreover, the inverses of $E_I, E_{II}, E_{III}$ are also elementary matrices of the same type. . . . .	19
10.3	Row-equivalent of a Matrix . . . . .	21

<b>11 Row Echelon Form of a Matrix</b>	<b>22</b>
11.1 Definition . . . . .	22
11.2 Notation . . . . .	22
<b>12 Row Rank of a Matrix</b>	<b>22</b>
<b>13 Gauss Theorem</b>	<b>22</b>
13.1 Elimination Algorithm . . . . .	23
13.1.1 Example . . . . .	23
13.2 Row Spaces of Matrices . . . . .	23
13.3 Column Equivalence . . . . .	24
 <b>IV Linear Systems</b>	 <b>26</b>
<b>1 Definition</b>	<b>26</b>
<b>2 Equivalent Systems</b>	<b>26</b>
<b>3 Solution of a System of Equations</b>	<b>27</b>
<b>4 Homogeneous Systems</b>	<b>28</b>
4.1 Definition . . . . .	28
4.2 Solutions of Homogeneous Systems . . . . .	28
4.2.1 Example . . . . .	29
4.2.2 General Solution . . . . .	30
4.3 Properties . . . . .	30
4.3.1 For a homogeneous system $Ax = 0$ , if $c$ and $d$ are solutions, then $c + d$ is also a solution. . . . .	30
4.3.2 For a homogeneous system $Ax = 0$ , if $c$ is a solution and $\alpha \in \mathbb{F}$ , then, $\alpha c$ is a solution too. . . . .	30
4.4 Fundamental Solutions . . . . .	31
4.4.1 Theorem: Any solution $d$ of the system $Ax = \mathbb{O}$ can be obtained from the basic solutions $v_1, \dots, v_t$ as a linear combination of the basic solutions, $d = \alpha_1 v_1 + \dots \alpha_t v_t$ .	31
4.5 . . . . .	31
<b>5 Non-Homogeneous Systems</b>	<b>32</b>
5.1 Definition . . . . .	32
5.2 Solutions of Non-Homogeneous Systems . . . . .	32
5.2.1 Case I: $\tilde{r} = r$ . . . . .	32
5.2.2 Case II: $\tilde{r} > r$ . . . . .	32

5.3	General Solution . . . . .	33
<b>V</b>	<b>Vector Spaces</b>	<b>34</b>
<b>1</b>	<b>Definition</b>	<b>34</b>
1.1	Examples . . . . .	34
1.1.1	Geometric Vectors in Plane . . . . .	34
1.1.2	Arithmetic Vector Space . . . . .	34
1.1.3	. . . . .	35
<b>2</b>	<b>Properties</b>	<b>35</b>
2.0.4	Proof of 1 . . . . .	35
<b>3</b>	<b>Subspaces</b>	<b>35</b>
3.1	Examples . . . . .	36
3.2	Operations on Subspaces . . . . .	37
<b>4</b>	<b>Spans</b>	<b>37</b>
<b>5</b>	<b>Linear Dependence</b>	<b>40</b>
5.1	Properties of Linearly Dependent and Independent Sets . . . .	41
5.2	Changing a Basis . . . . .	45
5.3	Representation of Vectors in a Basis . . . . .	48
5.3.1	Properties of Representations . . . . .	48
<b>6</b>	<b>Determinants</b>	<b>49</b>
6.1	Definition . . . . .	49
6.2	Properties . . . . .	49
6.3	Practical Methods for Computing Determinants . . . . .	52
6.4	Expansion along a row/ column . . . . .	52
6.5	Determinant Rank . . . . .	53
<b>7</b>	<b>Linear Maps</b>	<b>53</b>
7.1	Definition . . . . .	53
7.2	Properties . . . . .	53
7.3	Matrix of a Linear Map . . . . .	53
7.4	Change of Bases . . . . .	55
7.5	Operations on Linear Maps . . . . .	56
7.6	Kernel and Image . . . . .	58
7.6.1	Dimensions of Kernel and Image . . . . .	59

<b>VI</b>	<b>Linear Operators</b>	<b>61</b>
1	Definition	61
2	Similar Matrices	61
2.1	Properties of Similar Matrices . . . . .	62
3	Diagonalization	62
4	Eigenvalues and Eigenvectors	62
5	Characteristic Polynomial	64
5.1	Properties . . . . .	65
<b>VII</b>	<b>Inner Product Spaces</b>	<b>70</b>
1	Definition	70
2	Computation of Inner Products	72
2.1	Change of Basis . . . . .	74
3	Norms	74
3.1	Definition . . . . .	74
3.2	Properties . . . . .	75
4	Orthogonality	75
4.1	Definition . . . . .	75
4.2	Properties . . . . .	75
5	Orthogonal and Orthonormal Bases	75
6	Unitary Matrices	78

## **Part I**

# **General Information**

## **1 Contact Information**

**Prof. Boris Kunyavskii**

kunyav@gmail.com

## **2 Grades**

Final Exam: 80%

Midterm Exam: 10%

Homework: 10%

Passing Criteria: 60%

# Part II

## Fields

### 1 Definition

**Definition 1** (Field). The set  $\mathbb{F}$  is a field if there are operations  $+$ ,  $\cdot$  satisfying the following properties:

$$(A1) \quad \forall a, b \in \mathbb{F}; a + b = b + a$$

$$(A2) \quad \forall a, b \in \mathbb{F}; (a + b) + c = a + (b + c)$$

$$(A3) \quad \text{There is an element } 0 \in \mathbb{F} \text{ s.t. } a + 0 = 0 + a = a$$

$$(A4) \quad \forall a \in F, \exists b \in \mathbb{F} \text{ s.t. } a + b = 0$$

$$(M1) \quad \forall a, b \in \mathbb{F}, a \cdot b = b \cdot a$$

$$(M2) \quad \forall a, b \in \mathbb{F}, (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

$$(M3) \quad \text{There is an element } 1 \in \mathbb{F} \text{ s.t. } a \cdot 1 = 1 \cdot a = a (1 \neq 0)$$

$$(M4) \quad \forall a \in \mathbb{F}, (a \neq 0), \exists b \in \mathbb{F} \text{ s.t. } a \cdot b = 1$$

$$(AM) \quad (a + b) \cdot c = (a \cdot c) + (b \cdot c)$$

If  $\mathbb{F}$  is a field, one can define subtraction and division as follows.

$$a - b \doteq a + (-b)$$

$$\frac{a}{b} \doteq a \cdot \frac{1}{b}$$

#### 1.1 Examples of Fields

1.  $\mathbb{R}$
2.  $\mathbb{C}$
3.  $\mathbb{F}_p$

## 1.2 Examples of Non-fields (Rings)

1.  $\mathbb{Z}_2$ , as M4 is not satisfied.

If we define  $\mathbb{F}_2 = 0, 1; 0 + 0 = 0; 0 + 1 = 1 + 0 = 1; 1 + 1 = 0$ , then, necessarily, 1 will have no additive inverse.

## 2 Examples

**Example 1.** Let  $p$  be a prime number.  $\mathbb{F}_p$  is defined as follows.

$$\forall m \in \mathbb{Z}, m = a \cdot p + \bar{m}$$

The operations  $+$  and  $\cdot$  are defined as

$$\bar{a} + \bar{b} = \overline{(a + b)}$$

$$\bar{a} \cdot \bar{b} = \overline{(a \cdot b)}$$

1.  $\mathbb{F}_p$  is a field.
2. If  $\mathbb{F}$  is a set of  $q$  elements, we can define on  $\mathbb{F}$  a structure of a field iff  $q = p^t$ , where  $p$  is prime,  $t \geq 1$ .

**Example 2.** For a field of 4 elements  $\{0, 1, \alpha, \beta\}$ , the addition and multiplication tables are as follows.

+	0	1	$\alpha$	$\beta$
0	0	1	$\alpha$	$\beta$
1	1	0	$\beta$	$\alpha$
$\alpha$	$\alpha$	$\beta$	0	1
$\beta$	$\beta$	$\alpha$	0	1



## Part III

# Matrices

### 1 Definition

**Definition 2** (Matrix). Let  $\mathbb{F}$  be a field,  $m, n \geq 1$ .

Then,  $A(m \times n)$  is a table consisting of  $m$  rows and  $n$  columns, filled by elements of  $\mathbb{F}$ .

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

### 2 Addition of Matrices

**Definition 3** (Addition of matrices). Let  $A, B$  be  $m \times n$  matrices over  $\mathbb{F}$ .

Then,  $C = A + B$  is defined as follows.

$$c_{ij} = a_{ij} + b_{ij}$$

#### 2.0.1 Properties

1.  $A + B = B + A, \forall A, B$  s.t. the sum is defined
2.  $(A + B) + C = A + (B + C), \forall A, B, C$  s.t. the sums are defined
3. There is a matrix  $\mathbb{O}$ , s.t.  $A + \mathbb{O} = \mathbb{O} + A = A$
4. For any  $A, \exists B$  s.t.  $B = -A$

### 3 Multiplication of a matrix by a scalar

**Definition 4** (Multiplication of a matrix by a scalar). Let  $A$  be a  $m \times n$  matrix over  $\mathbb{F}$ . Let  $\alpha \in \mathbb{F}$  be a scalar. Then,  $C = \alpha A$  is defined as follows.

$$c_{ij} = \alpha a_{ij}$$

## 4 Multiplication of matrices

**Definition 5** (Multiplication of matrices). Let  $A$  be a  $m \times n$  matrix over  $\mathbb{F}$ . Let  $B$  be a  $n \times p$  matrix over  $\mathbb{F}$ . Then,  $C = AB$  is defined as follows.

$$c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$$

**Example 3.** For matrices  $A, B$ , of same size, is  $AB = BA$ ?

*Solution.*  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$   
 $\therefore AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$   
 $\therefore AB \neq BA$

*Remark 1.*  $A \neq \mathbb{O}, B \neq \mathbb{O}$ , but  $AB = \mathbb{O}$ .

## 5 Zero Divisor

**Definition 6** (Zero divisor). We say that a square matrix  $A \neq \mathbb{O}$  is a zero divisor if either there is a square matrix  $B$  s.t.  $AB = \mathbb{O}$ , or there is a square matrix  $C$ , s.t.  $CA = \mathbb{O}$ .

*Remark 2.*  $\mathbb{O}B = C\mathbb{O} = \mathbb{O}$ .

*Remark 3.*  $AC = BC \nRightarrow A = B$ . In general, we cannot cancel matrices on either side of an equation.

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, C = \mathbb{O}$$
$$\therefore AB = CB = \mathbb{O} \& B \neq \mathbb{O}$$

But, we cannot cancel  $B$ , as  $A \neq C$ .

## 6 Theorem ('Good properties of matrix multiplication')

**Theorem 1.**

$$(AB)C = A(BC) \quad (1.1)$$

$$A(B + C) = AB + AC \quad (1.2)$$

$$(A + B)C = AC + BC \quad (1.3)$$

$$(\alpha A) = \alpha(AB) \quad (1.4)$$

*Proof.* Denote  $AB = D, BC = G, (AB)C = F, A(BC) = H$

We need to prove  $F = H$

Let the dimensions of the matrices be as follows.

$A_{m \times n}, B_{n \times p}, C_{p \times q}$

$\therefore F_{m \times q}, H_{m \times q}$

$$d_{ik} = \sum_j a_{ij}b_{jk}$$

$$\therefore g_{jl} = \sum_k b_{jk}b_{kl}$$

$$f_{il} = \sum_k d_{ik}c_{kl} = \sum_k \left( \sum_j a_{ij}b_{jk} \right) c_{kl} = \sum_k \sum_j a_{ij}b_{jk}c_{kl}$$

$$h_{il} = \sum_j a_{ij}g_{jl} = \sum_j a_{ij} \left( \sum_k b_{jk}b_{kl} \right) = \sum_k \sum_j a_{ij}b_{jk}c_{kl}$$

$$f_{il} = h_{il}$$

$$F = H$$

□

## 7 Square Matrices

Let  $A$  be a square matrix of size  $n \times n, n \geq 1$

### 7.1 Diagonal Matrices

**Definition 7** (Diagonal matrix). We say that  $A$  is a diagonal matrix if  $a_{ij} = 0$ , whenever  $i \neq j$ .

**Theorem 2.** Let  $A$  and  $B$  be diagonal  $n \times n$  matrices.

$$a_{rr} = \alpha_r, b_{rr} = \beta_r$$

Then,  $AB = BA = C, C$  is a diagonal matrix with  $c_{rr} = a_{rr}b_{rr}$ .

#### 7.1.1 Proof

$$a_{ij} = \begin{cases} 0, i \neq j \\ \alpha_i, i = j \end{cases}$$

$$b_{ij} = \begin{cases} 0, i \neq j \\ \beta_i, i = j \end{cases}$$

$$c_{ik} = \sum_{j=1}^n a_{ij}b_{jk} = a_{ii}b_{ik} = \alpha_i b_{ik} = \begin{cases} 0, i \neq k \\ \alpha_i \beta_i, i = k \end{cases}$$

Similarly for  $BA$ .

### 7.2 Upper-triangular Matrices

We say that  $A$  is an upper-triangular matrix if  $a_{ij} = 0$ , whenever  $i > j$ .

### 7.3 Lower-triangular Matrices

We say that  $A$  is a lower-triangular matrix if  $a_{ij} = 0$ , whenever  $i < j$ .

#### Remark

Diagonal matrices are upper-triangular and lower-triangular. Conversely, if a matrix is both upper-triangular and lower-triangular, it is a diagonal matrix.

## 7.4 Theorem

If  $A$  and  $B$  are both upper-triangular, then  $AB$  and  $BA$  are upper-triangular too.

### 7.4.1 Proof

Denote  $C = AB$ .

$$\therefore c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$$

Suppose  $i > k$ , then, either  $i > j$  or  $j > k$ . So, in each case, atleast one of  $a_{ij}$  or  $b_{jk}$  is 0.

## 7.5 Identity Matrix

Let  $n \geq 1$ . We call  $I_n$  the  $n \times n$  identity matrix.

$$I_n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

## 7.6 Theorem

Let  $I_n$  be the identity  $n \times n$  matrix. Then, for any  $n \times n$  matrix  $B$ , we have

$$I_n B = B I_n = B$$

### 7.6.1 Proof

$$I_n = (e_{ij}); e_{ij} = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$$

Denote  $C = I_n B$ . We have

$$c_{ik} = \sum_{j=1}^n e_{ij}b_{jk} = e_{ii}b_{ik} = 1 \cdot b_{ik} = b_{ik}$$

$$\therefore C = B \Rightarrow I_n B = B$$

Similarly for  $B I_n = B$ .

## 7.7 Inverse of Matrix

Let  $A$  be an  $n \times n$  matrix. We say that  $A$  is invertible if there exist  $B, C$ , s.t.  $AB = I_n$  and  $CA = I_n$

### Remark

$A = \mathbb{O}$  is not invertible because  $\mathbb{O}B = C\mathbb{O} = \mathbb{O} \neq I_n$

### Remark

There are non-zero matrices which are not invertible.

$$\text{Let } A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

If possible, let there be  $C$  s.t.  $CA = I_2$ .

$$\text{Let } B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

We have  $CA = I$ .

$$\therefore (CA)B = IB$$

$$\therefore C(AB) = B$$

$$\therefore C\mathbb{O} = B$$

$$\therefore \mathbb{O} = B$$

But,  $B \neq 0$ . Therefore,  $C$  does not exist.

### 7.7.1 If $AB = I_n$ and $CA = I_n$ , then $B = C$

$$\begin{aligned} C &= CI \\ &= C(AB) \\ &= (CA)B \\ &= IB \\ &= B \end{aligned}$$

### 7.7.2 Inverse of a Matrix

If  $A$  is invertible, i.e. if there exists  $B$ , s.t.  $AB = BA = I$ , then,  $B$  is called the inverse of  $A$ , and is denoted by  $A^{-1}$ .

**7.7.3** If  $AB = I$ , then  $BA = I$ .

**7.7.4** If  $A$  is invertible, then  $A$  cannot be a zero divisor.

If possible, let  $A$  be a zero divisor.

Therefore, either  $AB = \mathbb{O}$ , for some  $B \neq \mathbb{O}$ ; or  $CA = \mathbb{O}$ , for some  $C \neq \mathbb{O}$

**Case I:**  $AB = \mathbb{O}$

$$\begin{aligned} AB &= \mathbb{O} \\ \therefore A^{-1}(AB) &= A^{-1}\mathbb{O} \\ \therefore (A^{-1}A)B &= \mathbb{O} \\ \therefore IB &= \mathbb{O} \\ \therefore B &= \mathbb{O} \end{aligned}$$

This contradicts the assumption  $B \neq \mathbb{O}$

**Case II:**  $CA = \mathbb{O}$

$$\begin{aligned} CA &= \mathbb{O} \\ \therefore (CA)A^{-1} &= \mathbb{O}A^{-1} \\ \therefore C(A^{-1}A) &= \mathbb{O} \\ \therefore CI &= \mathbb{O} \\ \therefore C &= \mathbb{O} \end{aligned}$$

This contradicts the assumption  $C \neq \mathbb{O}$

**7.7.5** If  $A$  and  $B$  are invertible, then  $A + B$  may or may not be invertible.

If  $A = B$ , then  $A + B = 2A$  is invertible.

If  $A = -B$ , then  $A + B = \mathbb{O}$  is not invertible.

**7.7.6** If  $A$  and  $B$  are invertible, then  $AB$  must be invertible.

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AIA^{-1} \\ &= AA^{-1} \\ &= I\end{aligned}$$

$$\begin{aligned}\text{Similarly, } (B^{-1}A^{-1})(AB) &= I \\ \therefore (AB)^{-1} &= B^{-1}A^{-1}\end{aligned}$$

## 8 Transpose of a Matrix

Let  $A$  be a  $m \times n$  matrix,  $A = (a_{ij})_{1 \leq i \leq m; 1 \leq j \leq n}$

$B = A^t$  is defined as follows.

$$b_{ji} = a_{ij}$$

### 8.1 Properties of $A^t$

1.  $(A + B)^t = A^t + B^t$
2.  $(\alpha A)^t = \alpha A^t$
3.  $(AB)^t = B^t A^t$
4. If  $A$  is invertible, then,  $A^t$  must be invertible, and  $(A^t)^{-1} = (A^{-1})^t$

## 9 Adjoint Matrix

$$A^* \doteq \overline{A}^t$$

For example,

$$A = \begin{pmatrix} 1 & 1+i & 2-i \\ i & -5i & 3 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & -i \\ 1-i & 5i \\ 2+i & 3 \end{pmatrix}$$



### 9.0.1 Properties of Adjoint Matrices

1.  $(A + B)^* = A^* + B^*$
2.  $(\alpha A)^* = \bar{\alpha} A^*$
3.  $(AB)^* = B^* A^*$
4. If  $A$  is invertible, then  $A^*$  is invertible, and  $(A^*)^{-1} = (A^{-1})^*$

## 10 Row Operations on Matrices

### 10.1 Elementary Row Operations

Let  $A$  be a  $m \times n$  matrix with rows  $a_1, \dots, a_m$ . We define 3 types of elementary row operations.

- I  $a_i \leftrightarrow a_j$  (Switch of the  $i^{\text{th}}$  and  $j^{\text{th}}$  rows.)
- II  $a_i \rightarrow \alpha a_i (\alpha \neq 0)$  (Multiplication of a row by a non-zero scalar.)
- III  $a_i \rightarrow a_i + \alpha a_j (j \neq i)$  (Addition of a row multiplied by a scalar, and another row.)

$E_I, E_{II}, E_{III}$  are matrices obtained from the identity matrix by applying elementary row operations I, II, III, respectively. These matrices are called elementary matrices.

### 10.2 Theorems

Let  $e_i = \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{pmatrix}$  be a  $1 \times m$  matrix.

Let  $A$  be any  $m \times n$  matrix.

Then,  $e_i A =$  the  $i^{\text{th}}$  row of  $A$ .

#### 10.2.1 $E_I A =$ the matrix obtained from $A$ by an elementary row operation I

##### Proof

Let  $A$  be any  $m \times n$  matrix.

$$\therefore E_I A = \begin{pmatrix} e_1 A \\ \vdots \\ e_j A \\ \vdots \\ e_i A \\ \vdots \\ e_m A \end{pmatrix}$$

**10.2.2**  $E_{II} A$  = the matrix obtained from  $A$  by an elementary row operation II

**Proof**

Let  $A$  be any  $m \times n$  matrix.

$$\therefore E_{II} A = \begin{pmatrix} e_1 A \\ \vdots \\ \alpha e_i A \\ \vdots \\ e_m A \end{pmatrix}$$

**10.2.3**  $E_{III} A$  = the matrix obtained from  $A$  by an elementary row operation III

**Proof**

Let  $A$  be any  $m \times n$  matrix.

$$\begin{aligned}
\therefore E_I A &= \begin{pmatrix} e_1 A \\ \vdots \\ a_{i1} + \alpha a_{j1} \cdots + a_{in} + \alpha a_{jn} \\ \vdots \\ e_j A \\ \vdots \\ e_m A \end{pmatrix} \\
&= \begin{pmatrix} 1^{\text{st}} \text{ row of } A \\ \vdots \\ i^{\text{th}} \text{ row of } A + \alpha(j^{\text{th}}) \text{ row of } A \\ \vdots \\ j^{\text{th}} \text{ row of } A \\ \vdots \\ m^{\text{th}} \text{ row of } A \end{pmatrix}
\end{aligned}$$

**10.2.4** All elementary matrices are invertible, moreover, the inverses of  $E_I, E_{II}, E_{III}$  are also elementary matrices of the same type.

$$\begin{aligned}
E_I^{-1} &= E_I \\
\Leftrightarrow E_I^2 &= I_m
\end{aligned}$$

$$\begin{aligned}
E_1^2 &= E_1 E_1 \\
&= \begin{pmatrix} e_1 E_1 \\ \vdots \\ e_j E_1 \\ \vdots \\ e_i E_1 \\ \vdots \\ e_m E_1 \end{pmatrix} \\
&= \begin{pmatrix} 1^{\text{st}} \text{ row of } A \\ \vdots \\ j^{\text{th}} \text{ row of } A \\ \vdots \\ i^{\text{th}} \text{ row of } A \\ \vdots \\ m^{\text{th}} \text{ row of } A \end{pmatrix} \\
&= \begin{pmatrix} e_1 \\ \vdots \\ e_j \\ \vdots \\ e_i \\ \vdots \\ e_m \end{pmatrix} = I_m
\end{aligned}$$

Similarly for  $E_{\text{II}}$ , to get the inverse,  $\alpha$  is replaced by  $\frac{1}{\alpha}$

$$E_{\text{II}} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

$$\therefore E_{\text{II}}^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{\alpha} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Similarly for  $E_{\text{III}}$ , to get the inverse,  $\alpha$  is replaced by  $-\alpha$

$$E_{\text{III}} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & \alpha & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

$$\therefore E_{\text{III}}^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & -\alpha & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

### 10.3 Row-equivalent of a Matrix

A matrix  $A'$  is a row-equivalent of  $A$ , if  $A'$  is obtained for  $A$ , by a finite sequence of elementary row operations.

## 11 Row Echelon Form of a Matrix

### 11.1 Definition

Let  $A$  be an  $m \times n$  matrix.

Denote the  $i^{\text{th}}$  row of  $A$  by  $a_i$ .

The leading entry of a non-zero row  $a_i$  is its first non-zero entry.

Denote the column where the leading entry occurs by  $l_i$ .

$$\begin{aligned} a_{ij} &= 0 \text{ if } j < l(i) \\ a_{ij} &\neq 0 \text{ if } j = l(i) \end{aligned}$$

We say that  $A$  is in row echelon form (REF) if the following conditions hold.

1. The non-zero rows are at the top of  $A$ . ( $r$  = the number of non-zero rows)
2. The leading entries go right as we go down, i.e.  $l(1) < l_2 < \dots < l(r)$
3. All leading entries equal 1, i.e. if  $j = l(i)$ , then,  $a_{ij} = 1$
4. Any column which contains a leading entry must have all other entries equal to 0, i.e. if  $j = l(i)$ , then,  $a_{kj} = 0; \forall k \neq i$

### 11.2 Notation

The REF of  $A$  will be denoted by  $A_R$ .

## 12 Row Rank of a Matrix

The number of non-zero rows in  $A_R$  is called the row rank of  $A$ . It is denoted by  $r$ .

$$r \leq n$$

## 13 Gauss Theorem

Any  $m \times n$  matrix  $A$  can be brought to REF by a sequence of elementary row operations.

### 13.1 Elimination Algorithm

Step 1 Find the first non-zero column  $C_p$  of  $A$ .

Step 2 Denote by  $a_{ip}$  the first non-zero entry of  $C_p$ .

Step 3 Switch the 1<sup>st</sup> and  $i^{\text{th}}$  rows.

Step 4 Multiply the 1<sup>st</sup> row by  $\frac{1}{a_{ip}}$ .

Step 5 Using row operations of type III, make all other entries of the  $p^{\text{th}}$  column zeros.

Step 6 Ignoring the top row and  $C_p$ , repeat steps Step 1 to Step 5.

#### 13.1.1 Example

$$\begin{aligned}
 & \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & -1 & 4 & 7 \\ 0 & -1 & 7 & 6 \end{pmatrix} \xrightarrow{R_1 \rightarrow R_2} \begin{pmatrix} 0 & -1 & 4 & 7 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & 7 & 6 \end{pmatrix} \xrightarrow{R_1 \rightarrow -R_1} \begin{pmatrix} 0 & 1 & -4 & -7 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & 7 & 6 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + R_1} \\
 & \begin{pmatrix} 0 & 1 & -4 & -7 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 3 & -1 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 0 & 1 & -4 & -7 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \xrightarrow{R_2 \rightarrow \frac{R_2}{3}} \begin{pmatrix} 0 & 1 & -4 & -7 \\ 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & -1 \end{pmatrix} \xrightarrow{R_1 \rightarrow R_1 + 4R_2} \\
 & \begin{pmatrix} 0 & 1 & 0 & -\frac{25}{3} \\ 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & -1 \end{pmatrix} \xrightarrow{R_3 \rightarrow -R_3} \begin{pmatrix} 0 & 1 & 0 & -\frac{25}{3} \\ 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_1 \rightarrow R_1 + \frac{25}{3}R_3} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 + \frac{1}{3}R_3} \\
 & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
 \end{aligned}$$

### 13.2 Row Spaces of Matrices

**Definition 8** (Row space of a matrix). Let  $A$  be a  $m \times n$  matrix over  $\mathbb{F}$ .  $R(A)$  is defined as

$$R(A) = \text{span } v_1, \dots, v_m$$

where  $v_1, \dots, v_m$  are rows of  $A$ .

$R(A)$  a subspace of the vector space of all rows of length  $n$ , is called the row space of  $A$ .

**Definition 9** (Row rank of a matrix).  $\dim R(A)$  is called the row-rank of  $A$ , and is denoted by  $\text{rr}(A)$ .

**Theorem 3.** *Let  $P$  be a  $l \times m$  matrix. Then*

1.  $R(PA) \subseteq R(A)$
2. *If  $P$  is an invertible  $m \times m$  matrix, then  $R(PA) = R(A)$*

**Corollary 3.1.**

$$A' \stackrel{R}{\sim} A \implies R(A') = R(A)$$

**Theorem 4.** *If  $A$  is in REF, and if  $r$  is the number of non-zero rows in  $A$ , then*

$$\text{rr}(A) = r$$

**Corollary 4.1.** *The following are equivalent*

1.  $A \stackrel{R}{\sim} A'$
2. *There is an invertible matrix  $P$ , s.t.  $A' = PA$*
3.  $R(A) = R(A')$
4.  *$A$  and  $A'$  have the same REF*

### 13.3 Column Equivalence

**Definition 10** (Elementary column operations, column equivalence, column echelon form, column space and column rank). If  $A$  is a  $m \times n$  matrix, we can define elementary column operations, column equivalence ( $A \stackrel{C}{\sim}$ ) and column echelon form (CEF), the column space of  $A$  ( $C(A)$ ), and the column rank of  $A$  ( $\text{cr}(A)$ ).

**Theorem 5.**

$$\text{cr}(A) = \text{rr}(A) = r$$



*Proof.* Let  $r = \text{rr}(A) = \dim \text{R}(A)$ .

Choose  $r$  rows of  $A$  which form a basis of  $\text{R}(A)$ , WLG, say  $v_1, \dots, v_r$ .

Let

$$X_{r \times n} = \begin{pmatrix} v_1 \\ \vdots \\ v_r \end{pmatrix}$$

$$\text{span}(X) = \text{R}(A)$$

Hence, any row of  $A$  can be expressed as a linear combination of  $v_1, \dots, v_r$

$$v_i = \sum_{j=1}^r y_{ij} v_j$$

Let

$$Y_{m \times r} = (y_{ij})$$

Therefore,

$$A = YX$$

Considering each column of  $A$  as a linear combination of columns of  $Y$ ,

$$\begin{aligned} \text{C}(A) &\subseteq \text{C}(Y) \\ \therefore \text{cr}(A) &\leq \text{cr}(Y) \leq r = \text{rr}(A) \\ \therefore \text{cr}(A) &\leq \text{rr}(A) \end{aligned}$$

Similarly,

$$\text{rr}(A) \leq \text{cr}(A) \therefore \text{cr}(A) = \text{rr}(A)$$

□

**Corollary 5.1.** *The following are equivalent*

1.  $A \stackrel{\mathcal{C}}{\sim} A'$
2. There is an invertible matrix  $Q$ , s.t.  $A' = QA$
3.  $\text{C}(A) = \text{C}(A')$
4.  $A$  and  $A'$  have the same CEF

# Part IV

## Linear Systems

### 1 Definition

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

Here, all  $x_i$  are taken to be unknowns, and all  $a_{ij}, b_i$  are given.

A solution to such a system is a collection  $d_1, \dots, d_n$ , s.t. after replacing  $x_i$  by  $d_i$ , we get equalities.

We assume that all  $a_{ij}, b_i$  belong to  $\mathbb{F}$ , and we are looking for solutions  $d_i \in \mathbb{F}$ .

Given such a system, we define  $A_{m \times n} = (a_{ij}), b_{m \times 1} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, x_{n \times 1} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

Then, we can write the system as

$$Ax = b$$

A solution to this system is  $d_n = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$ , s.t.  $Ad = b$

Let  $D$  be the set of all  $d = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$

$D$  may be empty, infinite, or a singleton set.

### 2 Equivalent Systems

Two systems  $Ax = b$  and  $A'x = b'$  are called equivalent, if every solution of the first system is also a solution of the second system, and vice versa.

### 3 Solution of a System of Equations

We want to bring a given system

$$Ax = b$$

to the form

$$A_R x = b_R$$

using elementary row operations.

We denote the augmented or extended matrix of the system as follows.

$$\overline{A}_{m \times (n+1)} = (A_{m \times n} | b_{m \times 1})$$

Then apply Gaussian elimination method to  $\overline{A}$ , in order to get the matrix

$$(A_R | b_R)$$

As  $A_R$  is obtained from  $A$  using elementary row operations,

$$A_R = E_n \dots E_2 E_1 A$$

where every  $E_i$  is an elementary matrix.

Let  $P = E_n \dots E_2 E_1$ .  $P$  is invertible, as it is a product of elementary matrices.

$$\begin{aligned} A_R &= PA \\ \therefore A_R d &= PAd \\ &= Pb \\ &= b_R \end{aligned}$$

Conversely, let  $d$  be a solution to

$$\begin{aligned} A_R d &= b_R \\ \therefore PAd &= b_R \\ \therefore P^{-1}(PAd) &= P^{-1}b_R \\ \therefore Ad &= b \end{aligned}$$

If we have a system  $Ax = b$ , we may and will assume that  $A$  is in REF, i.e.  $A = A_R, b = b_R$ .

Let  $l(1), \dots, l(r)$  denote the numbers of the columns containing leading entries.

$$\text{Let } b = \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ b_{r+1} \\ \vdots \\ b_m \end{pmatrix}$$

Therefore,

$$\begin{aligned} 1 \cdot x_{l(1)} + \dots &= b_1 \\ 1 \cdot x_{l(2)} + \dots &= b_2 \\ &\vdots \\ 1 \cdot x_{l(r)} &= b_r \\ 0 &= b_{r+1} \\ &\vdots \\ 0 &= b_m \end{aligned}$$

## 4 Homogeneous Systems

### 4.1 Definition

A system of the form

$$Ax = \mathbb{O}$$

is called a homogeneous system.

#### Remark

Any homogeneous system is consistent and has a trivial solution  $x = \mathbb{O}$

### 4.2 Solutions of Homogeneous Systems

If  $r$  = number of non-zero rows, let  $t = n - r$  = number of free variables.  
If  $t > 0$ , denote the numbers of the columns that do not contain leading entries by  $z(1), \dots, z(t)$

#### 4.2.1 Example

$$A = \begin{pmatrix} 0 & 1 & 2 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore,

$$m = 4$$

$$n = 6$$

$$r = 3$$

$$t = 3$$

$$l(1) = 2$$

$$l(2) = 4$$

$$l(3) = 5$$

$$z(1) = 1$$

$$z(2) = 3$$

$$z(3) = 6$$

Therefore,

$$x_2 + 2x_3 - 3x_6 = 0$$

$$x_4 - x_6 = 0$$

$$x_5 + 7x_6 = 0$$

Therefore,

$$x_2 = -2x_3 + 3x_6$$

$$x_4 = x_6$$

$$x_5 = -7x_6$$

$$\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix} = C_{3 \times 3} \begin{pmatrix} x_1 \\ x_3 \\ x_6 \end{pmatrix}$$

$$\text{where } C_{3 \times 3} = \begin{pmatrix} 0 & -2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -7 \end{pmatrix}$$

The free variables  $x_1, x_3, x_6$  can be considered as parameters,  $x_1 = \gamma_1, x_2 =$

$$\gamma_2, x_3 = \gamma_3.$$

Therefore,

$$x_2 = -2\gamma_3 + 3\gamma_6$$

$$x_4 = \gamma_6$$

$$x_5 = -7\gamma_6$$

#### 4.2.2 General Solution

##### 4.2.2.1 Case I: $t = 0$

If  $t = 0$ , there are no free variables, and the system has a unique trivial solution.

##### 4.2.2.2 Case II: $t > 0$

$$\begin{pmatrix} x_{l(1)} \\ x_{l(2)} \\ \vdots \\ x_{l(r)} \end{pmatrix} = C_{r \times t} \begin{pmatrix} x_{z(1)} \\ x_{z(2)} \\ \vdots \\ x_{z(t)} \end{pmatrix}$$

$C$  is filled by coefficients of the equations obtained after shifting the terms containing all  $z_i$  to the RHS.

### 4.3 Properties

**4.3.1 For a homogeneous system  $Ax = 0$ , if  $c$  and  $d$  are solutions, then  $c + d$  is also a solution.**

$$\begin{aligned} Ac &= \mathbb{O} \\ Ad &= \mathbb{O} \\ \therefore A(c + d) &= Ac + Ad \\ &= \mathbb{O} + \mathbb{O} \\ &= \mathbb{O} \end{aligned}$$

**4.3.2 For a homogeneous system  $Ax = 0$ , if  $c$  is a solution and  $\alpha \in \mathbb{F}$ , then,  $\alpha c$  is a solution too.**

$$\begin{aligned} Ac &= \mathbb{O} \\ \therefore A(\alpha c) &= \alpha(Ac) \\ &= \alpha \mathbb{O} \\ &= \mathbb{O} \end{aligned}$$

## 4.4 Fundamental Solutions

We define  $t$  fundamental solutions or basic solutions,  $v_1, \dots, v_t$ .

We define  $t$  columns, each of length  $n$  as follows.

For the  $i^{\text{th}}$  column  $v_i$ , we set

$$\begin{aligned} x_{z(1)} &= 0 \\ x_{z(i)} &= 1 \\ &\vdots \\ x_{z(t)} &= 0 \end{aligned}$$

and for  $x_{l(1)}, \dots, x_{l(r)}$ ,

$$\begin{pmatrix} x_{l(1)} \\ \vdots \\ x_{l(r)} \end{pmatrix} = C \begin{pmatrix} x_{z(1)} \\ \vdots \\ x_{z(t)} \end{pmatrix} = i^{\text{th}} \text{column of } C$$

**4.4.1 Theorem:** Any solution  $d$  of the system  $Ax = \mathbb{O}$  can be obtained from the basic solutions  $v_1, \dots, v_t$  as a linear combination of the basic solutions,  $d = \alpha_1 v_1 + \dots \alpha_t v_t$

One can choose another collection  $v'_1, \dots, v'_t$  s.t. any solution of  $Ax = \mathbb{O}$  can be obtained as a linear combination of  $v'_1, \dots, v'_t$ . In such a case, we get another form of the general solution.

## 4.5

$$r \leq \min m, n$$

If  $r = n$ , i.e.  $t = 0$ , the system has a unique solution.

If  $r < n$ , i.e.  $t > 0$ , the system has more than one solutions. Its general solution can be expressed as in terms of  $t$  parameters, where each free variable serves as a parameter, whose value can be any element of  $\mathbb{F}$ .

If  $m < n$ , then  $r < n$ . Therefore, the system has more than one solution.

## 5 Non-Homogeneous Systems

### 5.1 Definition

Consider a system  $Ax = b; b \neq \mathbb{O}$ . The extended matrix is defined as

$$\tilde{A} = (A|b) = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

### 5.2 Solutions of Non-Homogeneous Systems

Let  $\tilde{r}$  be the number of non-zero rows in the REF of  $\tilde{A}$ , i.e.  $\tilde{A}_R$ .

#### 5.2.1 Case I: $\tilde{r} = r$

$$b'_{r+1} = \dots = b'_m = 0$$

##### 5.2.1.1 Case a: $r = n$ , i.e. $t = 0$

Therefore,

$$\begin{aligned} x_1 &= b'_1 \\ &\dots \\ x_r &= b'_r \end{aligned}$$

Hence, the system has a unique solution.

##### 5.2.1.2 Case b: $r < n$ , i.e. $t > 0$

Therefore,

$$\begin{aligned} x_l(1) &= b'_1 + c_{11}x_{z(1)} + \dots + c_{1t}x_{z(t)} \\ &\vdots \\ x_l(r) &= b'_1 + c_{r1}x_{z(1)} + \dots + c_{rt}x_{z(t)} \end{aligned}$$

#### 5.2.2 Case II: $\tilde{r} > r$

In this case, the  $(r + 1)^{\text{th}}$  row represents an equation of the form  $0 = 1$ . Therefore, the system is inconsistent.



### 5.3 General Solution

The general solution of  $Ax = b$  can be expressed by adding the general solution of  $Ax = b$  and any particular solution of  $Ax = b$ .

If  $c$  is a solution of  $Ax = \mathbb{O}$ , and  $d$  is a solution of  $Ax = b$ , then  $c + d$  is a solution of  $Ax = b$ .

Conversely, if  $d$  and  $d'$  are solutions of  $Ax = b$ , then,  $c = d' - d$  is a solution of  $Ax = \mathbb{O}$ .

# Part V

## Vector Spaces

### 1 Definition

Let  $\mathbb{F}$  be a field. A vector space  $V$ , over  $\mathbb{F}$ , is a set on which there are two operations, denoted by  $+$  and  $\cdot$ , where

$+$  is the addition of elements of  $V$

$\cdot$  is the multiplication of an element of  $V$  by an element of  $\mathbb{F}$

s.t. the sum of elements of  $V$  lies in  $V$ , and the product of an element of  $V$  by an element of  $\mathbb{F}$  lies in  $V$ , and the following properties hold.

$$(A1) \quad x + y = y + x; \forall x, y \in V$$

$$(A2) \quad (x + y) + z = x + (y + z); \forall x, y, z \in V$$

$$(A3) \quad \exists \mathbb{O} \in V, \text{ s.t. } \mathbb{O} + x = x + \mathbb{O} = x; \forall x \in V$$

$$(A4) \quad \forall x \in V, \exists y \in V, \text{ s.t. } x + y = \mathbb{O}. \text{ (} y \text{ is denoted as } -x \text{.)}$$

$$(M1) \quad \alpha(x + y) = \alpha x + \alpha y; \forall \alpha \in \mathbb{F}, \forall x, y \in V$$

$$(M2) \quad (\alpha + \beta)x = \alpha x + \beta x; \forall \alpha, \beta \in \mathbb{F}, \forall x \in V$$

$$(M3) \quad (\alpha\beta)x = \alpha(\beta x) = \beta(\alpha x); \forall \alpha, \beta \in \mathbb{F}, \forall x \in V$$

$$(M4) \quad 1 \cdot x = x; \forall x \in V$$

Elements of  $V$  are called vectors, and elements of  $\mathbb{F}$  are called scalars.

### 1.1 Examples

#### 1.1.1 Geometric Vectors in Plane

#### 1.1.2 Arithmetic Vector Space

Let  $\mathbb{F}$  be a field, and  $n \geq 1 \in \mathbb{Z}$ .

Let  $V = \mathbb{F}^n$  be a set of ordered n-tuples.

We define

$$\begin{aligned} (\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) &= (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n) \\ \alpha(\alpha_1, \dots, \alpha_n) &= (\alpha\alpha_1, \dots, \alpha\alpha_n) \end{aligned}$$

### 1.1.3

Let  $\mathbb{F}$  be a field, and  $m, n \geq 1 \in \mathbb{Z}$ .

Let  $V = \mathbb{F}^{mn}$  be the set of all  $(m \times n)$  matrices over  $\mathbb{F}$ , i.e. a set of ordered  $mn$ -tuples. For  $X, Y \in V$ , we use the usual definitions of  $X + Y$  and  $\alpha X$  from algebra of matrices.

## 2 Properties

1.  $\alpha \mathbb{O} = \mathbb{O}; \forall \alpha \in F$
2.  $\alpha(-x) = -(\alpha x)$
3.  $x - y \doteq x + (-y)$
4.  $0x = \mathbb{O}; \forall x \in V$
5.  $(-1)x = -x; \forall x \in V$
6.  $(\alpha - \beta)x = \alpha x - \beta x; \forall \alpha, \beta \in F, \forall x \in V$

### 2.0.4 Proof of 1

$$\begin{aligned}\alpha \mathbb{O} &= \alpha(\mathbb{O} + \mathbb{O}) \\ &= \alpha \mathbb{O} + \alpha \mathbb{O}\end{aligned}$$

For  $\alpha \mathbb{O} \exists y$  s.t.  $\alpha \mathbb{O} + y = \mathbb{O}$ .

Therefore,

$$\begin{aligned}\alpha \mathbb{O} + y &= (\alpha \mathbb{O} + \alpha \mathbb{O}) + y \\ \therefore \mathbb{O} &= \alpha \mathbb{O} + (\mathbb{O} + y) \\ &= \alpha \mathbb{O} + \mathbb{O} \\ &= \alpha \mathbb{O}\end{aligned}$$

## 3 Subspaces

Let  $V$  be a vector space over  $\mathbb{F}$ . Let  $U \subseteq V$ .  $U$  is called a subspace of  $V$  if the following properties hold.

Axiom 1  $\mathbb{O} \in U$

Axiom 2 If  $x, y \in U$ , then,  $(x + y) \in U$

Axiom 3 If  $x \in U, \alpha \in \mathbb{F}$ , then,  $\alpha x \in U$

### 3.1 Examples

**Example 4.** Let  $V$  be the set of all geometric vectors in plane.

If  $U_1$  is the set of all vectors along the  $x$ -axis,  $U_2$  is the singleton set of a specific vector along the  $x$ -axis, and  $U_3$  is the set of all vectors along the  $x$ -axis and a specific vector not along the  $x$ -axis. Which of  $U_1, U_2, U_3$  are subspaces of  $V$ ?

*Solution.*  $U_1$  is a subspace of  $V$  as it satisfies all three axioms.

$U_2$  is not a subspace of  $V$  as it does not satisfy any of the three axioms.

$U_3$  is not a subspace of  $V$  as it does not satisfy Axiom 3

**Example 5.**

$$\mathbb{F} = \mathbb{R}$$

$$V = \mathbb{C} = \{\alpha + \beta i; \alpha, \beta \in \mathbb{R}\}$$

where  $+$  is addition in  $\mathbb{C}$  and  $\cdot$  is multiplication by real scalars.

$$U_1 = \{\alpha + 0i\}$$

$$U_2 = \{0 + \beta i\}$$

Which of  $U_1, U_2, U_3$  are subspaces of  $V$ ?

*Solution.* Both  $U_1$  and  $U_2$  are subspaces of  $V$ , as they satisfy all three axioms.

**Example 6.** Let  $V = \mathbb{F}$ , where  $+$  is addition in  $\mathbb{F}$ , and  $\cdot$  is multiplication in  $\mathbb{F}$ .

$$U_1 = \{\alpha + 0i\}$$

$$U_2 = \{0 + \beta i\}$$

Which of  $U_1, U_2$  are subspaces of  $V$ ?

*Solution.* Neither  $U_1$  nor  $U_2$  are subspaces of  $V$ .

**Example 7.** Let  $V = \{f : [0, 1] \rightarrow \mathbb{R}\}$ , where  $+$  and  $\cdot$  is defined as follows.

$$(f + g)(x) = f(x) + g(x)$$

$$(\alpha f)(x) = \alpha f(x)$$

$\mathbb{O}$  is the function with graph  $x = 0$ .

$$U = \{\text{all continuous functions } [0, 1] \rightarrow \mathbb{R}\}$$

Is  $U$  is subspace?

*Solution.*  $\mathbb{O} \in \mathbb{R}$ . Therefore, Axiom 1 is satisfied. Similarly, Axiom 2 and Axiom 3 are also satisfied.

### 3.2 Operations on Subspaces

Let  $V/F$  be a vector space, and  $U_1, U_2$  be subspaces of  $V$ .

$$U_1 \cap U_2 = \{x \in V : x \in U_1 \text{ and } x \in U_2\}$$

$$U_1 \cup U_2 = \{x \in V : x \in U_1 \text{ or } x \in U_2\}$$

$$U_1 + U_2 = \{x \in V : x = x_1 + x_2, x_1 \in U_1, x_2 \in U_2\}$$

**Example 8.** Let  $V$  be a set of geometric vectors in 3D space.

Let  $U_1$  be the  $xy$ -plane, and  $U_2$  be the  $yz$ -plane. Is  $U_1 \cap U_2$  a subspace of  $V$ ?

*Solution.*

$$\mathbb{O} \in U_1, \mathbb{O} \in U_2 \Rightarrow \mathbb{O} \in U_1 \cap U_2$$

$$x, y \in U_1 \cap U_2 \Rightarrow x, y \in U_1, x, y \in U_2$$

$$\Rightarrow x + y \in U_1, x + y \in U_2$$

$$= x + y \in U_1 \cap U_2$$

Similarly, if  $x \in U_1 \cap U_2, \alpha \in \mathbb{F}$ , then,  $\alpha x \in U_1 \cap U_2$ . Therefore,  $U_1 \cap U_2$  is a subspace of  $V$ .

## 4 Spans

**Definition 11** (Span). Let  $V/\mathbb{F}$  be a vector space. Let  $S \subset V$  be non-empty.

$$\text{span}(S) = \{x \in V : x = \alpha_1 v_1 + \cdots + \alpha_m v_m, \alpha_1, \dots, \alpha_m \in \mathbb{F}, v_1, \dots, v_m \in S\}$$

$\text{span}(S)$  is the collection of all linear combinations of finite number of vectors of  $S$  with coefficients from  $\mathbb{F}$

**Theorem 1.**  $\text{span}(S)$  is a subspace of  $V$

*Proof.*

$$\mathbb{O} = 0v \Rightarrow \mathbb{O} \in \text{span}(S)$$

$$\begin{aligned} x, y \in \text{span}(S) &\Rightarrow x = \alpha_1 v_1 + \cdots + \alpha_m v_m, \beta_1 w_1 + \cdots + \beta_m w_m \\ &\Rightarrow x + y = \alpha_1 v_1 + \cdots + \alpha_m v_m + \beta_1 w_1 + \cdots + \beta_m w_m \in \text{span}(S) \end{aligned}$$

$$\begin{aligned} x \in \text{span}(S), \alpha \in \mathbb{F} &\Rightarrow \alpha_1 v_1 + \cdots + \alpha_m v_m \\ &\Rightarrow \alpha x = \alpha(\alpha_1 v_1 + \cdots + \alpha_m v_m) \\ &\Rightarrow \alpha x = \alpha \alpha_1 v_1 + \cdots + \alpha \alpha_m v_m \in \text{span}(S) \end{aligned}$$

□

**Definition 12** (Spanning sets and dimensionality). Let  $V/\mathbb{F}$  be a vector space. A set  $S \subseteq V$  is said to be a spanning set, if  $\text{span}(S) = V$ . If  $V$  has at least one finite spanning set,  $V$  is said to be finite-dimensional. Otherwise,  $V$  is said to be infinite-dimensional.

*Remark 4.*  $V$  may have many finite spanning sets, of different sizes

**Definition 13** (Basis of a vector space). Let  $V/\mathbb{F}$  be a vector space. We say that  $B = \{v_1, \dots, v_n\} \subset V$  is a basis of  $V$  if every vector  $v \in V$  can be expressed in a unique way

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n \quad ; \alpha_1, \dots, \alpha_n \in \mathbb{F}$$

that is, as a linear combination of elements of  $B$ .

**Definition 14** (Isomorphic spaces). Let  $V/\mathbb{F}$  and  $W/\mathbb{F}$  be vector spaces. We say that  $V$  is isomorphic to  $W$  if there is a map  $\varphi : V \rightarrow W$ , s.t.

1.  $\varphi$  is one-to-one and onto
2.  $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2); \forall v_1, v_2 \in V$
3.  $\varphi(\alpha v) = \alpha \varphi(v); \forall v \in V, \forall \alpha \in \mathbb{F}$

**Theorem 2.** If a vector space  $V/\mathbb{F}$  has a basis  $B = \{v_1, \dots, v_n\}$  consisting of  $n$  elements, then it is isomorphic to the space

$$W = \mathbb{F}^n = \left\{ \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \right\}$$

*Proof.* Let  $B' = \{e_1, \dots, e_n\}$ , where

$$e_1 = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$$

$B'$  is a basis of  $Q$ , as any  $w = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in W$  can be expressed in a unique way

$$w = \alpha_1 e_1 + \dots + \alpha_n e_n$$

Let  $\varphi : V \rightarrow W$ ,

$$\begin{aligned} \varphi(v_1) &= e_1 \\ &\vdots \\ \varphi(v_n) &= e_n \end{aligned}$$

For any  $v = \alpha_1 v_1 + \dots + \alpha_n v_n \in V$ ,

$$\varphi(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Therefore,

$$\begin{aligned} \varphi(\alpha_1 v_1 + \dots + \alpha_n v_n) &= \alpha_1 e_1 + \dots + \alpha_n e_n \\ &= \alpha_1 \varphi(v_1) + \dots + \alpha_n \varphi(v_n) \end{aligned}$$

If  $v \neq v'$ ,

$$\begin{aligned} v &= \alpha_1 v_1 + \dots + \alpha_n v_n \\ v' &= \alpha'_1 v_1 + \dots + \alpha'_n v_n \end{aligned}$$

Hence  $\varphi$  is one-to-one.

For any  $w = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in W$ .

Let  $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ .

Therefore,

$$\varphi(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = w$$

Therefore,  $\varphi$  is onto. □

## 5 Linear Dependence

**Definition 15** (Linearly dependent subsets). Let  $V/\mathbb{F}$  be a vector space. Let  $S \subseteq V$  be a finite subset.  $S$  is said to be linearly dependent if there exist scalars  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ , not all equal to zero, s.t.

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = \mathbb{O}$$

Otherwise,  $S$  is said to be linearly independent if all  $\alpha_1 = \cdots = \alpha_n = 0$ .

**Example 9.** Is  $S = \{v_1, \dots, v_l, v, \alpha v\}$  linearly dependent?

*Solution.*

$$(0)v_1 + \cdots + (0)v_l + (-\alpha)v + (1)\alpha v = \mathbb{O}$$

Therefore, as not all coefficients are zero,  $S$  is linearly dependent.

**Example 10.** Is  $S = \{v_1, \dots, v_l, \mathbb{O}\}$  linearly dependent?

*Solution.*

$$(0)v_1 + \cdots + (0)v_l + (1)\mathbb{O} = \mathbb{O}$$

Therefore, as not all coefficients are zero,  $S$  is linearly dependent.

**Theorem 3.** Any basis  $B = \{v_1, \dots, v_n\}$  of a vector space  $V$  is linearly independent.

*Proof.* Let

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = \mathbb{O}$$

Also,

$$(0)v_1 + \cdots + (0)v_n = \mathbb{O} \tag{3.1}$$



Therefore, there are two representations of  $v = \mathbb{O}$  as linear combinations of elements of  $B$ . By the definition of basis, they must coincide.

Therefore,

$$\begin{aligned}\alpha_1 &= 0 \\ \vdots \\ \alpha_n &= 0\end{aligned}$$

Hence,  $B$  is linearly independent.  $\square$

## 5.1 Properties of Linearly Dependent and Independent Sets

**Theorem 4.** *If  $S \subseteq S'$  and  $S$  is linearly dependent, then  $S'$  is also linearly dependent.*

**Theorem 5.** *If  $S \subseteq S'$  and  $S'$  is linearly independent, then  $S$  is also linearly independent.*

**Theorem 6.** *Let  $S = \{v_1, \dots, v_n\}$ .  $S$  is linearly dependent iff one of the  $v_i$ s is a linear combination of the others.*

*Proof of statement.* Suppose

$$\begin{aligned}v_n &= \alpha_1 v_1 + \dots + \alpha_{n-1} v_{n-1} \\ \therefore \alpha_1 v_1 + \dots + \alpha_{n-1} v_{n-1} + (-1)v_n &= \mathbb{O}\end{aligned}$$

Therefore,  $S$  is linearly dependent.  $\square$

*Proof of converse.* Suppose

$$\alpha_1 v_1 + \dots + \alpha_{n-1} v_{n-1} + \alpha_n v_n = \mathbb{O}$$

not all of  $\alpha_i$ s are 0. WLG, let  $\alpha_n \neq 0$

$$\therefore v_n = -\frac{\alpha_1}{\alpha_n} v_1 - \dots - \frac{\alpha_{n-1}}{\alpha_n} v_{n-1}$$

$\square$

**Theorem 7.** *Let  $S = \{v_1, \dots, v_m\}$ . Let  $w \in V$ . Suppose  $w$  is a linear combination of  $v_i$ s*

$$w = \alpha_1 v_1 + \dots + \alpha_n v_n$$

*Then, such an expression is unique iff  $S$  is linearly dependent.*

*Proof of statement.* Let

$$w = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

be unique.

If possible, let

$$\beta_1 v_1 + \cdots + \beta_n v_n = \mathbb{O}$$

not all  $\beta_i$ s are zero.

Then,

$$(\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_n + \beta_n)v_n = w$$

This is another expression for  $w$ , and contradicts the assumption.  $\square$

*Proof of converse.* If possible, let  $S$  be linearly independent. Assume

$$w = \alpha'_1 v_1 + \cdots + \alpha'_n v_n$$

Therefore,

$$(\alpha_1 - \alpha'_1)v_1 + \cdots + (\alpha_n - \alpha'_n)v_n = \mathbb{O}$$

Therefore,  $S$  is linearly dependent, which contradicts the assumption.  $\square$

**Theorem 8** (Main Lemma on Linear Independence). *Suppose  $V$  is spanned by  $n$  vectors.*

*Let  $S = \{v_1, \dots, v_m\} \subset V$ . Suppose  $m > n$ .*

*Then,  $S$  is linearly dependent.*

*Proof.* Let  $E = \{w_1, \dots, w_n\}$  be a spanning set for  $V$ ,  $V = \text{span}(E)$ .

Therefore, all elements of  $S$  can be represented as linear combinations of elements of  $E$ .

$$\begin{aligned} v_1 &= \beta_{11}w_1 + \cdots + \beta_{1n}w_n \\ &\vdots \\ v_m &= \beta_{m1}w_1 + \cdots + \beta_{mn}w_n \end{aligned}$$

Let

$$\begin{aligned} &\alpha_1 v_1 + \cdots + \alpha_m v_m = \mathbb{O} \\ \therefore \alpha_1(\beta_{11}w_1 + \cdots + \beta_{1n}w_n) + \cdots + \alpha_m(\beta_{m1}w_1 + \cdots + \beta_{mn}w_n) &= \mathbb{O} \\ \therefore (\alpha_1\beta_{11} + \cdots + \alpha_m\beta_{m1})w_1 + \cdots + (\alpha_1\beta_{1n} + \cdots + \alpha_m\beta_{mn}) &= \mathbb{O} \end{aligned}$$

Therefore

$$\begin{aligned}\alpha_1\beta_{11} + \cdots + \alpha_m\beta_{m1} &= 0 \\ \vdots \\ \alpha_1\beta_{1n} + \cdots + \alpha_m\beta_{mn} &= 0\end{aligned}$$

These equations form a homogeneous linear system with respect to  $\alpha_1, \dots, \alpha_m$ . As  $m > n$ , the system has a non-zero solution. Therefore not all  $\alpha_i$ s are zero. Hence  $S$  is linearly dependent.  $\square$

**Definition 16** (Alternative definition of a basis).  $B = \{v_1, \dots, v_n\}$  is said to be a basis of  $V$  if  $B$  is a spanning set and  $B$  is linearly independent.

**Theorem 9.** *If  $B$  and  $B'$  are bases of  $V$ , then they contain the same number of elements.*

*Proof.* If possible, let  $B$  contain  $n$  elements  $\{v_1, \dots, v_n\}$ , and  $B'$  contain  $m$  elements  $\{w_1, \dots, w_m\}$ ,  $m > n$ .

Therefore,  $B$  is a spanning set and  $B'$  contains more elements than  $n$ , hence by Main Lemma on Linear Independence,  $B'$  is linearly dependent. Also,  $B'$  is a basis, so it is linearly independent.

This is a contradiction.  $\square$

**Definition 17** (Dimension of a vector space). Let  $V/\mathbb{F}$  be a finite-dimensional vector space. The number of elements in any basis  $B$  of  $V$  is called the dimension of  $V$ .

$$n = \dim V$$

*Remark 5.* If  $V$  and  $W$  are vector spaces over  $\mathbb{F}$ , s.t.

$$\dim V = \dim W$$

then,  $V$  is isomorphic to  $W$

**Theorem 10.** *If  $S = \{v_1, \dots, v_m\}$  is a spanning set of  $V$ , and if  $S$  is not a basis of  $V$ , a basis  $B$  of  $V$  can be obtained by removing some elements from  $S$ .*

*Proof.* If  $S$  is linearly independent, then it is a basis.

Otherwise, if  $S$  is linearly dependent, it has an element, WLG, say  $v_m$ , which is a linear combination of the others.

$$v_m = \alpha_1 v_1 + \cdots + \alpha_{m-1} v_{m-1}$$

Let

$$S' = S - \{v_m\}$$

$S'$  is a spanning set.

Therefore,  $\forall v \in V$

$$\begin{aligned} v &= \beta_1 v_1 + \cdots + \beta_{m-1} v_{m-1} + \beta_m v_m \\ &= \beta_1 v_1 + \cdots + \beta_{m-1} + \beta_m (\alpha_1 v_1 + \cdots + \alpha_{m-1} v_{m-1}) \\ &= \gamma_1 v_1 + \cdots + \gamma_{m-1} v_{m-1} \end{aligned}$$

If  $S'$  is linearly independent, then it is a basis, else the same process above can be repeated till we get a basis.

Therefore, a basis is a smallest spanning set.  $\square$

**Theorem 11.** *If  $B_0 = \{v_1, \dots, v_n\}$  is a linearly independent set, and if  $B_0$  is a basis of  $V$ , a basis of  $V$  can be obtained by adding elements to  $B_0$ .*

**Theorem 12.** *Let  $V$  be a vector space, s.t.  $\dim V = n$ .*

*If  $B$  satisfies 2 out of the 3 following conditions, then it is a basis.*

1.  $B$  has  $n$  elements.
2.  $B$  is a spanning set.
3.  $B$  is linearly dependent.

**Theorem 13** (Dimension Theorem).

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

**Theorem 14.**

$$U + W = \text{span}(U \cup W)$$

*If*

$$U = \text{span}(B)$$

$$W = \text{span}(B')$$

*then,*

$$U + W = \text{span}(B \cup B')$$

*Proof.* Let  $v \in U + W$ .

Then,

$$\begin{aligned} v &= u + w \quad ; u \in U, w \in W \\ u &\in U \cup W \\ w &\in U \cup W \\ \therefore v &\in \text{span}(U \cup W) \end{aligned}$$

Let

$$v \in \text{span}(U \cup W) \therefore v = \alpha_1 v_1 + \cdots + \alpha_k v_k \quad ; v_i \in U \cup W$$

Let

$$\begin{aligned} v_1, \dots, v_l &\in U \\ v_{l+1}, \dots, v_k &\in W \end{aligned}$$

Therefore,

$$\begin{aligned} v &= (\alpha_1 v_1 + \cdots + \alpha_l v_l) + (\alpha_{l+1} v_{l+1} + \cdots + \alpha_k v_k) \\ \therefore v &\in U + W \end{aligned}$$

□

## 5.2 Changing a Basis

Let  $B = \{v_1, \dots, v_n\}$  be a basis of  $V$ , s.t.  $\dim V = n$ . Let  $B' = \{v'_1, \dots, v'_n\}$ . As  $B$  is a spanning set, all of  $v'_1, \dots, v'_n$  can be expressed as a linear combination of  $v_1, \dots, v_n$ .

$$\begin{aligned} v'_1 &= \gamma_{11} v_1 + \cdots + \gamma_{n1} v_n \\ &\vdots \\ v'_n &= \gamma_{1n} v_1 + \cdots + \gamma_{nn} v_n \end{aligned}$$

**Definition 18** (Transition matrix). The matrix

$$C = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ \vdots & & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} \end{pmatrix}$$

is called the transition matrix from  $B$  to  $B'$ .

If  $B$  and  $B'$  are considered as row vectors of length  $n$  filled by vectors,

$$\begin{aligned} v'_1 &= \gamma_{11}v_1 + \cdots + \gamma_{n1}v_n \\ &\vdots \\ v'_n &= \gamma_{1n}v_1 + \cdots + \gamma_{nn}v_n \end{aligned}$$

can be written as

$$B'_{1 \times n} = B_{1 \times n} C_{n \times n}$$

**Theorem 15.**  $B'$  is a basis of  $V$  iff  $C$  is invertible.

*Proof of statement.* Let  $B' = BC$  be a basis.

$B'$  is a basis, and hence is a spanning set. Therefore, any vector from  $B$  can be expressed as a linear combination of elements of  $B'$ .

Therefore,

$$\begin{aligned} B &= B'Q \\ &= BCQ \end{aligned}$$

Also,

$$B = BI$$

Therefore,

$$I = CQ$$

Similarly,

$$\begin{aligned} B' &= BC \\ &= B'QC \end{aligned}$$

Also,

$$B' = B'I$$

Therefore,

$$I = QC$$

Therefore,

$$CQ = QC = I$$

Hence  $C$  is invertible. □

*Proof of converse.* Let  $B' = BC$  and  $C$  be invertible. Therefore,  $B'$  is a basis iff  $B$  is a spanning set.

Let  $z \in V$ . As  $B$  is a spanning set,

$$z = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

Therefore,

$$z = Bg$$

where

$$\begin{aligned} g &= \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \\ \therefore z &= Bg \\ &= B(Ig) \\ &= B(CC^{-1})g \\ &= (BC)(C^{-1}g) \end{aligned}$$

Let  $C^{-1}g = f$

$$\therefore z = B'f$$

Therefore,  $z$  can be expressed as a linear combination of vectors from  $B'$ .  $\square$

*Remark 6.* Let  $B$  be a basis of  $V$ . If

$$BP = BQ$$

where  $P$  and  $Q$  are  $n \times n$  matrices, then

$$P = Q$$

**Example 11.** Let  $B = \{e_1, e_2\}$  and  $B' = \{e'_1, e'_2\}$ , where

$$\begin{aligned} e'_1 &= e_1 + e_2 \\ e'_2 &= -e_1 + e_2 \end{aligned}$$

*Solution.*

$$\begin{aligned} e'_1 &= e_1 + e_2 \\ e'_2 &= -e_1 + e_2 \\ \therefore C &= \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
e_1 &= \frac{1}{2}e'_1 - \frac{1}{2}e'_2 \\
e_2 &= \frac{1}{2}e'_1 + \frac{1}{2}e'_2 \\
\therefore C^{-1} &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}
\end{aligned}$$

### 5.3 Representation of Vectors in a Basis

Let  $V$  be a vector space of dimension  $n$ . Let  $B = \{v_1, \dots, v_n\}$  be a basis of  $V$ .

Let  $z \in V$ .

$z$  can be written as a unique linear combination of elements of  $B$ .

$$z = \alpha_1 v_1 + \dots + \alpha_n v_n$$

The representation of  $z$  w.r.t  $B$  can be represented as

$$[z]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

#### 5.3.1 Properties of Representations

1.  $[z_1 + z_2]_B = [z_1]_B + [z_2]_B$
2.  $[\alpha z]_B = \alpha [z]_B$
3.  $[z_1]_B = [z_2]_B \iff z_1 = z_2$
4.  $\forall \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{F}^n, \exists z \in V, \text{ s.t. } [z]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$



## 6 Determinants

### 6.1 Definition

**Definition 19** (Determinants). Given an  $n \times n$  matrix  $A$ ,  $n \geq 1$ ,  $\det(A)$  is defined as follows.

$$\begin{aligned} n = 1 & \quad \det(a) = a \\ n = 2 & \quad \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} \\ & \quad \vdots \\ n = n & \end{aligned}$$

The determinant of a  $n \times n$  matrix is the summation of  $n!$  summands. Each summand is the product of  $n$  elements, each from a different row and column.

Summand	Permutation	Number of Elementary Permutations <sup>1</sup>	Parity
$a_{11}a_{22}a_{33}$	$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$	0	even
$a_{12}a_{23}a_{31}$	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$	2 $((1, 2, 3) \rightarrow (2, 1, 3) \rightarrow (2, 3, 1))$	even
$a_{13}a_{21}a_{32}$	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$	2 $((1, 2, 3) \rightarrow (1, 3, 2) \rightarrow (3, 1, 2))$	even
$a_{13}a_{22}a_{31}$	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$	1 $((1, 2, 3) \rightarrow (3, 2, 1))$	odd
$a_{12}a_{21}a_{33}$	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$	1 $((1, 2, 3) \rightarrow (2, 1, 3))$	odd
$a_{11}a_{23}a_{32}$	$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$	1 $((1, 2, 3) \rightarrow (1, 3, 2))$	odd

### 6.2 Properties

**Theorem 16.** If  $A$ ,  $A'$  are matrices s.t. all rows except the  $i^{\text{th}}$  row are identical, and  $A''$  is obtained by addition of  $i^{\text{th}}$  row of  $A$  and  $i^{\text{th}}$  row of  $A'$ ,

---

<sup>1</sup>Any permutation can be represented as a result of a series of elementary permutations, i.e. permutations of 2 elements only. The parity of a particular permutation depends of the parity of the number of elementary functions required for it.

then

$$\det(A'') = \det(A) + \det(A')$$

**Theorem 17.** *If  $A'$  is obtained from  $A$  by switching two rows, then*

$$\det(A') = -\det(A)$$

**Theorem 18.** *If  $A'$  is obtained from  $A$  by multiplication of a row by a scalar  $\alpha$ , then*

$$\det(A') = \alpha \det(A)$$

**Theorem 19.** *If  $A'$  is obtained from  $A$  by adding to the  $i^{\text{th}}$  row the  $j^{\text{th}}$  row multiplied by a scalar  $\alpha$ , then*

$$\det(A') = \det(A)$$

**Corollary 19.1** (Corollary of Property 2). *If  $A$  has two identical rows, then  $\det(A) = 0$ .*

**Theorem 20.** *The determinant of upper triangular and lower triangular matrices is the product of the elements on the principal diagonal.*

**Theorem 21.**

$$\det(A^t) = \det(A)$$

**Corollary 21.1.** *In all above theorems, the properties which are applicable to rows, are also applicable to columns.*

**Theorem 22.** *If  $A$ ,  $B$ ,  $C$  are some matrices, and  $\mathbb{O}$  is the zero matrix,*

$$\begin{pmatrix} A_{m \times m} & B \\ \mathbb{O} & C_{n \times n} \end{pmatrix} = \det(A) \cdot \det(C)$$

**Theorem 23.**

$$\det(AB) = \det(A) \det(B)$$

**Corollary 23.1.** *If  $A$  is invertible, then*

$$\det(A) \neq 0$$

*Proof.*  $A$  is invertible.

Therefore,  $\exists P$ , s.t.

$$PA = I$$

$$\therefore \det(PA) = \det(I)$$

$$\therefore \det(P) \det(A) = 1$$

$$\therefore \det(A) \neq 0$$

□

**Theorem 24.** *If*

$$\det(A) \neq 0$$

*then  $A$  is invertible.*

*Proof.* If possible let  $A$  be non invertible.

Let the REF of  $A$  be  $A_R$ .

As  $A$  is non invertible,  $A_R$  has a zero row. Therefore,

$$\det(A_R) = 0$$

But

$$\det(A) = 0$$

This is not possible as elementary row operations cannot change a non-zero determinant to zero.

Therefore,  $A$  is invertible.

□

**Theorem 25.**

$$\det(A) \neq 0$$

*iff the rows of  $A$  are linearly independent iff the columns of  $A$  are linearly independent.*

*Proof.* If possible, let the rows of  $A$  be linearly dependent.

Therefore, either all of them are zeros, or one row is the linear combination of the others.

*Case 1* (All rows are zeros).

$$\therefore \det(A) = 0$$

Case 2 (One row is a linear combination of the others). Let

$$v_n = \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1}$$

$$\therefore A = \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix}$$

$$v_n \rightarrow v_n - \alpha_1 v_1 - \cdots - \alpha_{n-1} v_{n-1}$$

$$\therefore A' = \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ \mathbb{O} \end{pmatrix}$$

$$\therefore \det(A') = 0$$

$$\therefore \det(A) = 0$$

This contradicts  $\det(A) \neq 0$ . Therefore, the rows of  $A$  must be linearly independent.

If  $v_1, \dots, v_n$  are linearly independent,

$$\dim R(A) = n$$

$$\therefore r = n$$

Therefore, there are no zero rows in REF of  $A$ . Hence  $A$  is invertible.

$$\therefore \det(A) \neq 0$$

□

## 6.3 Practical Methods for Computing Determinants

### 6.4 Expansion along a row/ column

Let  $A$  be a  $m \times n$  matrix, and let  $A_{ij}$  be the matrix obtained by removing the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column from  $A$ .

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

## 6.5 Determinant Rank

**Definition 20** (Determinant rank). Let  $A$  be any  $m \times n$  matrix. Consider all square sub-matrices of  $A$  and compute their determinants. If there is an  $r \times r$  sub-matrix of  $A$  s.t. its determinant is non-zero, but the determinants of all  $(r+1) \times (r+1)$  sub-matrices of  $A$  are zero, then,  $r$  is called the determinant rank of  $A$ .

**Theorem 26.** *The determinant rank of  $A$  is equal to the rank of  $A$ .*

## 7 Linear Maps

### 7.1 Definition

**Definition 21** (Linear map). Let  $V$  and  $W$  be vector spaces over the same field  $\mathbb{F}$ .

$$\varphi : V \rightarrow W$$

is said to be a linear map if

1.  $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2); \forall v_1, v_2 \in V$
2.  $\varphi(\alpha v) = \alpha \varphi(v); \forall v \in V, \forall \alpha \in \mathbb{F}$

### 7.2 Properties

1.  $\varphi(\mathbb{O}) = \mathbb{O}$
2.  $\varphi(-v) = -\varphi(v)$

### 7.3 Matrix of a Linear Map

**Definition 22** (Matrix of a linear map). Let  $\varphi : V \rightarrow W$  be a linear map. Let

$$n = \dim V$$

$$m = \dim W$$

Let

$$\begin{aligned} B &= \{v_1, \dots, v_n\} \\ B' &= \{w_1, \dots, w_m\} \end{aligned}$$

be bases of  $V$  and  $W$  respectively.

Let

$$\begin{aligned} \varphi(v_1) &= \alpha_{11}w_1 + \dots + \alpha_{m1}w_m \\ &\vdots \\ \varphi(v_n) &= \alpha_{1n}w_1 + \dots + \alpha_{mn}w_m \end{aligned}$$

The matrix

$$A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$$

is called the matrix of  $\varphi$  with respect to the bases  $B$  and  $B'$ .

It is denoted as

$$A = [\varphi]_{B,B'}$$

**Theorem 27.** *Let*

$$\varphi : V \rightarrow W$$

*be a linear map.*

*Let  $B$  and  $B'$  be bases of  $V$  and  $W$  respectively, and let*

$$A = [\varphi]_{B,B'}$$

*be the matrix of  $\varphi$  with respect to  $B$  and  $B'$ . Then,  $\forall x \in V$ ,*

$$[\varphi(z)]_{B'} = A[z]_B$$

*Proof.* Let

$$\begin{aligned} B &= \{v_1, \dots, v_n\} \\ B' &= \{w_1, \dots, w_m\} \end{aligned}$$

Case 3 ( $z \in B$ ). WLG, let  $z = v_i$ . Then,

$$[z]_B = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

i.e. all rows except the  $i^{\text{th}}$  row are 0.

Let this vector be  $e_i$ .

Therefore,

$$A[z]_B = Ae_i$$

is the  $i^{\text{th}}$  column of  $A$ .

$$[\varphi(z)]_{B'} = [\varphi(v_i)]_{B'}$$

is the  $i^{\text{th}}$  row in the formulae of  $\varphi(v_1), \dots, \varphi(v_n)$ .

Therefore, it is the  $i^{\text{th}}$  column of  $A$ .

Case 4 ( $z \in V$  is an arbitrary vector). Let

$$z = \alpha_1 v_1 + \dots + \alpha_n v_n$$

Therefore,

$$\begin{aligned} [\varphi(z)]_{B'} &= [\varphi(\alpha_1 v_1 + \dots + \alpha_n v_n)]_{B'} \\ &= [\alpha_1 \varphi(v_1) + \dots + \alpha_n \varphi(v_n)]_{B'} \\ &= \alpha_1 [\varphi(v_1)]_{B'} + \dots + \alpha_n [\varphi(v_n)]_{B'} \\ &= \alpha_1 \cdot (1^{\text{st}} \text{ column of } A) + \dots + \alpha_n \cdot (n^{\text{th}} \text{ column of } A) \\ &= A[z]_B \end{aligned}$$

□

## 7.4 Change of Bases

**Theorem 28.** Let  $V, W$  be vector spaces over  $\mathbb{F}$ ,  $\dim(V) = n$ ,  $\dim(W) = m$ . Let  $\varphi : V \rightarrow W$  be a linear map. Let  $B, \tilde{B}$  be bases of  $V$  and let  $B'$  and  $\tilde{B}'$  be bases of  $W$ . Let  $A = [\varphi]_{B, B'}$  and  $\tilde{A} = [\varphi]_{\tilde{B}, \tilde{B}'}$  be the matrices of  $\varphi$  w.r.t.

the pairs  $B, B'$  and  $\widetilde{B}, \widetilde{B}'$ . Let  $P$  denote the transition matrix from  $B$  to  $\widetilde{B}$ , and let  $Q$  denote the transition matrix from  $B'$  to  $\widetilde{B}'$ . Then,

$$\widetilde{A}_{m \times n} = Q_{m \times m}^{-1} A_{m \times n} P_{n \times n}$$

*Proof.*  $\forall z \in V$ ,

$$[\varphi(z)]_{B'} = A[z]_B \quad (28.1)$$

$$[\varphi(z)]_{\widetilde{B}'} = A[z]_{\widetilde{B}} \quad (28.2)$$

We have

$$[z]_B = P[z]_{\widetilde{B}} \quad (28.3)$$

$$[\varphi(z)]_{B'} = Q[\varphi(z)]_{\widetilde{B}'} \quad (28.4)$$

Therefore,

$$(28.1) \text{ in } (28.4) \implies$$

$$A[z]_B = Q[\varphi(z)]_{\widetilde{B}'} \quad (28.5)$$

$$(28.3) \text{ in } (28.5) \implies$$

$$AP[z]_{\widetilde{B}} = Q[\varphi(z)]_{\widetilde{B}'} \quad (28.6)$$

Multiplying on the left by  $Q^{-1}$ ,

$$\begin{aligned} Q^{-1}AP[z]_{\widetilde{B}} &= [\varphi(z)]_{\widetilde{B}'} \\ \therefore [\varphi(z)]_{\widetilde{B}'} &= Q^{-1}AP[z]_{\widetilde{B}} \end{aligned}$$

Comparing with (28.2),

$$\widetilde{A} = Q^{-1}AP$$

□

## 7.5 Operations on Linear Maps

**Definition 23** (Operations on linear maps). Let

$$\begin{aligned} \varphi &: V \rightarrow W \\ \varphi' &: V \rightarrow W \end{aligned}$$



be linear maps.

$$\varphi + \varphi' : V \rightarrow W$$

is defined as

$$(\varphi + \varphi')(v) = \varphi(v) + \varphi'(v)$$

and

$$\alpha\varphi : V \rightarrow W$$

is defined as

$$(\alpha\varphi)(v) = \alpha\varphi(v)$$

**Definition 24** (Composed map). Let

$$\varphi : V \rightarrow W$$

$$\varphi' : W \rightarrow U$$

be linear maps.

$$(\varphi' \circ \varphi) : V \rightarrow U$$

is defined as

$$(\varphi' \circ \varphi)(v) = \varphi'(\varphi(v))$$

**Theorem 29** (Matrix of composed map). *Let  $\varphi : V \rightarrow W$ ,  $\varphi' : W \rightarrow U$  be linear maps. Let  $(\varphi' \circ \varphi) : V \rightarrow U$  be the composed map. Let  $\dim V = n$ ,  $\dim W = m$ ,  $\dim U = l$ . Let  $B, B', B''$  be bases of  $V, W, U$  respectively. Let  $A = [\varphi]_{B,B'}$ ,  $A' = [\varphi']_{B',B''}$  be the matrices of  $\varphi, \varphi'$ . Let  $A'' = [\varphi' \circ \varphi]_{B,B''}$  be the matrix of the composed map. Then,*

$$A'' = A'A$$

*Proof.* Let  $z \in V$ .

$$\begin{aligned} [(\varphi' \circ \varphi)(z)]_{B''} &= [\varphi'(\varphi(z))]_{B''} \\ &= A'[\varphi(z)]_{B'} \\ &= A'A[z]_B \end{aligned}$$

By definition,

$$[(\varphi' \circ \varphi)(z)]_{B''} = A''[z]_B$$

Therefore,

$$A'' = A'A$$

□

## 7.6 Kernel and Image

**Definition 25** (Kernel and image). Let  $\varphi : V \rightarrow W$  be a linear map.

$$\ker \varphi \doteq \{v \in V : \varphi(v) = \mathbb{O}\}$$

$$\operatorname{im} \varphi \doteq \{\phi(v) : v \in V\}$$

**Theorem 30.**  $\ker \varphi$  is a subspace of  $V$  and  $\operatorname{im} \varphi$  is a subspace of  $W$ .

*Proof.*

$$\varphi(\mathbb{O}) = \mathbb{O}$$

$$\therefore \mathbb{O} \in \ker \varphi$$

If  $v_1, v_2 \in \ker \varphi$ , then

$$\begin{aligned}\varphi(v_1 + v_2) &= \varphi(v_1) + \varphi(v_2) \\ &= \mathbb{O} + \mathbb{O} \\ &= \mathbb{O}\end{aligned}$$

$$\therefore v_1 + v_2 \in \ker V$$

If  $v \in \ker \varphi$ ,  $\alpha \in \mathbb{F}$ , then

$$\begin{aligned}\varphi(\alpha v) &= \alpha \varphi(v) \\ &= \alpha \mathbb{O} \\ &= \mathbb{O} \therefore \alpha v \in \ker \varphi\end{aligned}$$

Therefore,  $\ker \varphi$  is a subspace of  $W$ .

$$\varphi(\mathbb{O}) = \mathbb{O}$$

$$\therefore \mathbb{O} \in \operatorname{im} \varphi$$

If  $w_1, w_2 \in \operatorname{im} \varphi$ , then

$$\begin{aligned}w_1 &= \varphi(v_1) \\ w_2 &= \varphi(v_2) \\ \therefore w_1 + w_2 &= \varphi(v_1) + \varphi(v_2) \\ &= \varphi(v_1 + v_2) \\ \therefore w_1 + w_2 &\in \operatorname{im} \varphi\end{aligned}$$

If  $w \in W$ ,  $\alpha \in \mathbb{F}$ , then

$$\begin{aligned}\alpha w &= \alpha \phi(v) \\ &= \varphi(\alpha v) \\ \therefore \alpha w &\in \operatorname{im} \varphi\end{aligned}$$

Therefore,  $\operatorname{im} \varphi$  is a subspace of  $W$ . □

### 7.6.1 Dimensions of Kernel and Image

**Theorem 31.** *Let  $\varphi : V \rightarrow W$  be a linear map. Then*

$$\dim(\ker(\varphi)) + \dim(\operatorname{im}(\varphi))$$

*Proof.* Let  $\ker \varphi = U$ ,  $U \subseteq V$ .

Let  $B_0 = \{v_1, \dots, v_k\}$  be a basis of  $U$ .

Completing  $B_0$  to a basis  $B$  of  $V$ ,

$$B = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$$

Let

$$w_{k+1} = \varphi(v_{k+1})$$

$$\vdots$$

$$w_n = \varphi(v_n)$$

Therefore, we need to prove that  $B'$  is a basis of  $W' = \operatorname{im}(\varphi)$ , by proving that  $B'$  is a spanning set and that  $B'$  is linearly independent.

Take  $w \in \operatorname{im}(\varphi)$ , so that there is  $v \in V$  s.t.  $\varphi(v) = w$ .

Representing  $v$  as a linear combination of elements of  $B$ ,

$$\begin{aligned} v &= \alpha_1 v_1 + \dots + \alpha_k v_k + \alpha_{k+1} v_{k+1} + \dots + \alpha_n v_n \\ \therefore w &= \varphi(v) \\ &= \varphi(\alpha_1 v_1 + \dots + \alpha_k v_k + \alpha_{k+1} v_{k+1} + \dots + \alpha_n v_n) \\ &= \alpha_1 \varphi(v_1) + \dots + \alpha_k \varphi(v_k) + \alpha_{k+1} \varphi(v_{k+1}) + \dots + \alpha_n \varphi(v_n) \\ &= \alpha_{k+1} \varphi(v_{k+1}) + \dots + \alpha_n \varphi(v_n) \\ &= \alpha_{k+1} w_{k+1} + \dots + \alpha_n w_n \\ &\in \operatorname{span}(B') \end{aligned}$$

Therefore,  $B'$  is a spanning set for  $W'$ .

Let

$$\beta_{k+1} w_{k+1} + \dots + \beta_n w_n = \mathbb{O}$$

Therefore,  $B'$  is linearly independent iff

$$\beta_{k+1} = \dots = \beta_n = 0$$

As  $\varphi$  is a linear map,

$$\begin{aligned} \varphi(\beta_{k+1} v_{k+1} + \dots + \beta_n v_n) &= \mathbb{O} \\ \therefore \beta_{k+1} v_{k+1} + \dots + \beta_n v_n &\in \ker \varphi \end{aligned}$$

Therefore, it can be expressed as a linear combination of vectors of  $B_0$ , which is a basis of  $\ker \varphi$ .

Let

$$\begin{aligned}\beta_{k+1}v_{k+1} + \cdots + \beta_nv_n &= \alpha_{k+1}v_{k+1} + \cdots + \alpha_nv_n \\ \therefore \alpha_{k+1}v_{k+1} + \cdots + \alpha_nv_n - \beta_{k+1}v_{k+1} - \cdots - \beta_nv_n &= \mathbb{O}\end{aligned}$$

As  $\{v_1, \dots, v_n\}$  is a basis of  $V$ , all coefficients must be 0

Therefore,

$$\beta_{k+1}v_{k+1} = \cdots = \beta_nv_n = 0$$

Hence, as  $B'$  is a spanning set of  $\text{im } \varphi$  and also linearly independent,  $B'$  is a basis of  $\text{im } \varphi$ .

Therefore,

$$\begin{aligned}\dim(\text{im } \varphi) &= \text{size of } B' \\ &= n - k \\ &= n - \dim(\ker \varphi) \\ \therefore \dim(\text{im } \varphi) + \dim(\ker \varphi) &= \dim V\end{aligned}$$

□

**Corollary 31.1.**

$$\dim(\text{im } \varphi) = r$$

where  $r$  is the rank of  $A$

**Corollary 31.2.** Let  $A_{m \times n}$  be a matrix of rank  $r$ . Let  $C(A)$  be the column space of  $A$ , and let  $\dim C(A)$  be the column rank of  $A$ . Then

$$\dim C(A) = r$$

*Proof.* Define

$$\varphi : \mathbb{F}^n \rightarrow \mathbb{F}^m$$

s.t.  $A = [\varphi]_{B, B'}$ , where  $B$  is the standard basis of  $\mathbb{F}$ .

$$\begin{aligned}B &= \left\{ \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} \right\} \\ &= \{e_1, \dots, e_n\}\end{aligned}$$

$\forall v \in \mathbb{F}^n$ , we have

$$[\varphi(v)]_{B'} = A[v]_B$$

If  $v = e_i$ ,

$$[\varphi(e_i)] = Ae_i$$

which is the  $i^{\text{th}}$  column of  $A$ . So, the space spanned by  $\{\varphi(e_1), \dots, \varphi(e_n)\}$  is equal to  $C(A)$ . But it is also in  $\text{im } \varphi$ .

Therefore,

$$\text{im } \varphi = C(A)$$

and

$$\begin{aligned} \dim(\text{im } \varphi) &= \dim(C(A)) \\ \therefore r &= \dim(C(A)) \end{aligned}$$

□

*Remark 7.* Let  $\varphi : V \rightarrow W$  be a linear map. Let  $w \in \text{im } (\varphi)$ , so that there is  $v \in V$  s.t.  $\varphi(v) = w$ . Then any  $v'$  s.t.  $\varphi(v') = w$  can be written down as  $v' = v + v_0$  where  $v_0 \in \ker \varphi$ .

## Part VI

# Linear Operators

### 1 Definition

**Definition 26** (Linear operator). A linear operator or transformation

$$T : V \rightarrow V$$

is a linear map from a vector space  $V$  to itself.

### 2 Similar Matrices

Let  $B$  and  $\tilde{B}$  be bases of  $V$ . Let  $A$  and  $\tilde{A}$  be the representing matrices

$$\begin{aligned} A &= [T]_B \\ \tilde{A} &= [T]_{\tilde{B}} \end{aligned}$$

Both these are  $n \times n$  matrices, where  $n = \dim V$ . Let  $P$  denote the transition matrix from  $B$  to  $\tilde{B}$ . Then,

$$\tilde{A} = P^{-1}AP$$

**Definition 27** (Similarity of matrices). Let  $A, \tilde{A}$  be  $n \times n$  matrices.  $A$  is said to be similar to  $\tilde{A}$ , denoted as  $A \sim \tilde{A}$ , if there exists an invertible  $n \times n$  matrix  $P$ , s.t.  $\tilde{A} = P^{-1}AP$ .

## 2.1 Properties of Similar Matrices

1.  $A \sim A$
2. If  $A \sim \tilde{A}$ , then  $\tilde{A} \sim A$
3. If  $A \sim \tilde{A}$  and  $\tilde{A} \sim \tilde{\tilde{A}}$ , then  $A \sim \tilde{\tilde{A}}$
4. If  $A \sim \tilde{A}$ , then  $\det(A) = \det(\tilde{A})$
5. If  $A \sim I$ , then  $A = I$

## 3 Diagonalization

Given a square matrix  $A_{n \times n}$ , decide whether or not  $A$  is similar to some diagonal matrix  $D$ . If it is, find  $D$ , and  $P$  s.t.  $P^{-1}AP = D$ .

Alternatively,

Given an operator  $T : V \rightarrow V$ , decide whether or not there exists a basis  $B$  of  $V$ , s.t.  $[T]_B$  is a diagonal matrix  $D$ . If it exists, find  $D$ , and  $B$ , s.t.  $[T]_B = D$ .

**Definition 28** (Diagonalizability). If  $A$  is similar to a diagonal matrix,  $A$  is said to be diagonalizable.  $P$ , s.t.  $P^{-1}AP = D$  is called a diagonalizing matrix for  $A$ .  $D$  is called a diagonal form of  $A$ .

## 4 Eigenvalues and Eigenvectors

**Definition 29** (Eigenvalue and eigenvector). Let  $A$  be a  $n \times n$  matrix over  $\mathbb{F}$ .  $\lambda \in \mathbb{F}$  is said to be an eigenvalue of  $A$ , if  $\exists v \in \mathbb{F}, v \neq 0$ , such that

$$Av = \lambda v$$

$v$  is called an eigenvector corresponding to  $\lambda$ .

**Definition 30** (Alternate definition of eigenvalue and eigenvector). Let  $T : V \rightarrow V$  be a linear operator, where  $V$  is a vector space over  $\mathbb{F}$ .  $\lambda \in \mathbb{F}$  is said to be an eigenvalue of  $A$ , if  $\exists v \in V, v \neq 0$ , such that

$$T(v) = \lambda v$$

$v$  is called an eigenvector corresponding to  $\lambda$ .

**Definition 31** (Spectrum). The collection of all eigenvalues of a matrix, or a linear operator is called the spectrum.

**Theorem 1.** Let  $A$  be a  $n \times n$  matrix.  $\lambda \in \mathbb{F}$  is an eigenvalue of  $A$  iff

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

*Proof.*  $\lambda$  is an eigenvalue of  $A$

$$\iff \exists v \in \mathbb{F}^n, v \neq 0, \text{ s.t. } Av = \lambda v$$

$$\iff \exists v \in \mathbb{F}^n, v \neq 0, \text{ s.t. } (\lambda I - A)v = \mathbf{0}$$

$$\iff v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$\iff (\lambda I - A) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{0} \text{ has a non-zero solution}$$

$$\iff \text{there are free variables}$$

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

□

**Theorem 2** (General criterion for diagonalization). Let  $A$  be a  $n \times n$  matrix.  $A$  is diagonalizable if and only if there exists a basis  $B = \{v_1, \dots, v_n\}$  of  $\mathbb{F}^n$  consisting of eigenvectors of  $A$ . In such a case, the diagonal entries of  $D$  are eigenvalues of  $A$ , and  $B$  can be chosen as consisting of the columns of  $P$ , where  $P^{-1}AP = D$ .

**Corollary 2.1.** If  $A$  has no eigenvalues, then it is not diagonalizable.

**Theorem 3.** Let  $\lambda_1, \dots, \lambda_s$  be pairwise distinct eigenvalues of an  $n \times n$  matrix  $A$ , i.e.  $\forall i \neq j, \lambda_i \neq \lambda_j$ . Let  $v_1, \dots, v_s$  be eigenvectors of  $A$  corresponding to  $\lambda_1, \dots, \lambda_s$ . Then the set  $S = \{v_1, \dots, v_s\}$  is linearly independent.

*Proof.* If possible, let  $S$  be linearly dependent. Let  $S'$  denote a linearly dependent subset of  $S$  of smallest possible size, say  $l$ . WLG, let  $S' = \{v_1, \dots, v_l\}$ . Hence,  $\exists \alpha_1, \dots, \alpha_l \in \mathbb{F}$ , s.t.

$$\alpha_1 v_1 + \dots + \alpha_l v_l = \mathbb{O} \quad (3.1)$$

Multiplying (3.1) on both sides by  $A$ ,

$$\alpha_1 A v_1 + \dots + \alpha_l A v_l = \mathbb{O} \quad (3.2)$$

$$\alpha_1 \lambda_1 v_1 + \dots + \alpha_l \lambda_l v_l = \mathbb{O} \quad (3.3)$$

Multiplying (3.1) on both sides by  $\lambda_l$

$$\alpha_1 \lambda_1 v_1 + \dots + \alpha_l A v_l = \mathbb{O} \quad (3.4)$$

Subtracting (3.4) from (3.3)

$$\alpha_1 (\lambda_1 - \lambda_l) v_1 + \dots + \alpha_{l-1} (\lambda_{l-1} - \lambda_l) v_{l-1} = \mathbb{O} \quad (3.5)$$

Solving,

$$\alpha_1 = \alpha_l = 0$$

This is a contradiction.  $\square$

**Corollary 3.1.** *Let  $A_{n \times n}$  have  $n$  distinct eigenvalues. Then,  $A$  is diagonalizable.*

*Proof.* Let  $v_1, \dots, v_n$  be eigenvectors of  $A$ , corresponding to  $\lambda_1, \dots, \lambda_n$ . As they are distinct, by the above theorem, they are linearly independent. The number of elements in the set  $\{v_1, \dots, v_n\}$  is  $n$ . Therefore, the set is a basis. Hence, according to General criterion for diagonalization,  $A$  is diagonalizable.  $\square$

## 5 Characteristic Polynomial

**Definition 32** (Characteristic Polynomial). Let  $A$  be any  $n \times n$  matrix.

$$p_A(x) = \det(xI_n - A)$$

is called the characteristic polynomial.



## 5.1 Properties

1. The roots of  $p_A(x)$  are the eigenvalues of  $A$ .
2.  $\deg p_A(x) = n$
3. The coefficient of  $x^n$  is 1.
4. The constant term is  $\alpha_0 = (-1)^n \det(A)$ .
5. The coefficient of  $x^{n-1}$  is  $\alpha_{n-1} = -(a_{11} + \cdots + a_{nn})$ .

**Theorem 4.** *If  $A \sim A'$ , then  $p_A(x) = p_{A'}(x)$ .*

*Proof.*

$$\begin{aligned}
 A' &= P^{-1}AP \\
 \therefore p_{A'}(x) &= \det(xI - A') \\
 &= \det(xI - P^{-1}AP) \\
 &= \det(P^{-1}(xI)P - P^{-1}AP) \\
 &= \det(P^{-1}(xI - A)P) \\
 &= \cancel{\det(P^{-1})} \det(xI - A) \cancel{\det(P)} \\
 &= \det(xI - A) \\
 &= p_A(x)
 \end{aligned}$$

□

**Definition 33** (Alternative definition of characteristic polynomial). Let  $T : V \rightarrow V$  be a linear operator. The characteristic polynomial of  $T$  is defined as the characteristic polynomial of any representing matrix of  $T$ .

**Theorem 5.** *Let  $f(x)$ ,  $g(x)$  be polynomials. Then  $\exists q(x), r(x)$ , s.t.*

$$f(x) = g(x)q(x) + r(x)$$

*and  $\deg r(x) < \deg g(x)$ .*

**Definition 34** (Remainder). If

$$f(x) = g(x)q(x) + r(x)$$

$r(x)$  is called the remainder after division of  $f(x)$  by  $g(x)$ . If  $r(x) = \mathbb{O}$ ,  $f(x)$  is said to be divisible by  $g(x)$ .

**Corollary 5.1.** *Let  $f(x)$  be a polynomial and let  $\alpha$  be a root of  $f$ . Then  $f(x)$  is divisible by  $(x - \alpha)$ .*

**Definition 35** (Algebraic multiplicity of eigenvalue). Let  $A$  be a  $n \times n$  matrix, and let  $p_A(x)$  be the characteristic polynomial of  $A$ , and let  $\lambda$  be an eigenvalue of  $A$ . The algebraic multiplicity of  $\lambda$  is defined as the largest possible integer value of  $k$  such that  $p_A(x)$  is divisible by  $(x - \lambda)^k$ .

**Definition 36** (Eigenspace). Let  $A$  be a  $n \times n$  matrix, and let  $\lambda$  be an eigenvalue of  $A$ . The eigenspace of  $A$  corresponding to  $\lambda$  is defined as

$$V_\lambda = \{v \in \mathbb{F}^n; Av = \lambda v\}$$

**Theorem 6.** *An eigenspace of a matrix is a subspace of the field over which the matrix is defined.*

**Definition 37** (Geometric multiplicity of eigenvalue).  $m = \dim V_\lambda$  is called the geometric multiplicity of  $\lambda$ .

**Theorem 7.** *Let  $\lambda$  be an eigenvalue of  $A_{n \times n}$ . Let  $k$  be the algebraic multiplicity of  $\lambda$  and let  $m$  be the geometric multiplicity of  $\lambda$ . Then*

$$m \leq k$$

*Proof.*

$$m = \dim V_\lambda$$

Therefore, let  $B_0 = \{v_1, \dots, v_m\}$  be a basis of  $V_\lambda$ .

Completing  $B_0$  to  $B = \{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ , a basis of  $\mathbb{F}^n$ .

Let  $P_{n \times n}$  be a matrix with columns  $v_1, \dots, v_n$ .

$$P = \begin{pmatrix} v_1 & \dots & v_m & v_{m+1} & \dots & v_n \end{pmatrix}$$

$P$  is invertible as  $v_1, \dots, v_n$  are linearly independent.

Consider  $A' = P^{-1}AP$ .

$$\begin{aligned} \therefore P^{-1}AP &= P^{-1}A \begin{pmatrix} v_1 & \dots & v_m & v_{m+1} & \dots & v_n \end{pmatrix} \\ &= P^{-1} \begin{pmatrix} Av_1 & \dots & Av_m & Av_{m+1} & \dots & Av_n \end{pmatrix} \\ &= P^{-1} \begin{pmatrix} \lambda v_1 & \dots & \lambda v_m & \star & \dots & \star \end{pmatrix} \\ &= \begin{pmatrix} P^{-1}(\lambda v_1) & \dots & P^{-1}(\lambda v_m) & \star & \dots & \star \end{pmatrix} \\ &= \begin{pmatrix} \lambda e_1 & \dots & \lambda e_m & \star & \dots & \star \end{pmatrix} \\ &= \begin{pmatrix} \lambda I_m & \star \\ 0 & \tilde{A} \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
p_{A'}(x) &= \det(xI_n + A') \\
&= \det \left( \begin{pmatrix} xI_m & 0 \\ 0 & xI_{n-m} \end{pmatrix} - \begin{pmatrix} \lambda I_m & \star \\ 0 & \tilde{A} \end{pmatrix} \right) \\
&= \det \begin{pmatrix} (x - \lambda)I_m & \star \\ 0 & xI_{n-m} - \tilde{A} \end{pmatrix} \\
&= \det((x - \lambda)I_m) \cdot \det(xI_{n-m} - \tilde{A}) \\
&= (x - \lambda)^m \cdot p_{\tilde{A}}(x)
\end{aligned}$$

As  $A \sim \tilde{A}$ ,

$$p_A(x) = p_{A'} = (x - \lambda)^m p_{\tilde{A}}(x)$$

By the definition of Algebraic multiplicity of eigenvalue,  $k \geq m$ .  $\square$

**Theorem 8.** *If a matrix  $A_{n \times n}$  is diagonalizable, then its characteristic polynomial  $p_A(x)$  can be represented as a product of linear factors.*

$$p_A(x) = (x - \lambda_1)^{k_1} \dots (x - \lambda_s)^{k_s}$$

where  $k_i$  is the algebraic multiplicity of  $\lambda_i$  and  $\lambda_1, \dots, \lambda_s$  are pairwise distinct.

*Proof.* As  $A$  is diagonalizable, let  $A \sim D$ ,

$$\begin{pmatrix}
\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \ddots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \ddots & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \lambda_s & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \ddots & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \lambda_s
\end{pmatrix}$$

Then,

$$\begin{aligned}
p_A(x) &= p_D(x) \\
&= \det \begin{pmatrix} x - \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & x - \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x - \lambda_s & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x - \lambda_s \end{pmatrix} \\
&= (x - \lambda_1)^{k_1} \dots (x - \lambda_s)^{k_s}
\end{aligned}$$

□

**Theorem 9** (Explicit criterion for diagonalization). *Let  $A$  be an  $n \times n$  matrix, s.t.  $p_A(x)$  splits completely. Then  $A$  is diagonalizable if and only if  $\forall \lambda_i$  of  $A$ , the algebraic multiplicity coincides with the geometric multiplicity.*

*Proof of statement.* If  $p_A$  splits completely, then  $k_1 + \dots + k_n = n$ .

If  $A$  is diagonalizable, then by the General criterion for diagonalization, there is  $B = \{v_1, \dots, v_n\}$ , a basis of  $\mathbb{F}^n$ , s.t. each  $v_i$  is an eigenvector of  $A$ .

Dividing  $v_1, \dots, v_n$  into  $s$  groups corresponding to  $\lambda_1, \dots, \lambda_s$ , to each  $\lambda_i$ , there correspond at most  $m_i = \dim V_{\lambda_i}$  eigenvectors, as they are a part of a basis and hence linearly independent.

Therefore,

$$n \leq m_1 + \dots + m_s$$

As  $p_A$  splits completely,

$$n = k_1 + \dots + k_s$$

Also,  $k_i \geq m_i$

$$\therefore k_1 + \dots + k_s = m_1 + \dots + m_s$$

moreover,  $\forall i$ , s.t.  $1 \leq i \leq s$ ,

$$k_i = m_i$$

□

*Proof of converse.*

$\forall i, \text{ s.t. } 1 \leq i \leq s$

$$\therefore k_i = m_i$$

As  $k_1 + \dots + k_s = n$ ,

$$m_1 + \dots + m_s = n$$

Let the bases of the eigenspaces  $V_{\lambda_1}, \dots, V_{\lambda_s}$  be  $B_1, \dots, B_s$ .

$$|B_1| = m_1$$

$$\vdots$$

$$|B_s| = m_s$$

Let  $B = B_1 \cup \dots \cup B_s$ .  $|B| = n$ .

It is enough to prove that  $B$  is linearly independent.

Let

$$B = \{v_1, v_2, \dots, w_1, w_2, \dots, u_1, u_2, \dots\}$$

Suppose

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \beta_1 w_1 + \beta_2 w_2 + \dots + \gamma_1 u_1 + \gamma_2 u_2 + \dots = \mathbb{O}$$

If possible, let at least one coefficient be non-zero. WLG, let  $\alpha_1 \neq 0$ .

Hence, as  $v_1, v_2, \dots$  form  $B_1$  which is a basis of  $V_{\lambda_1}$ ,

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots \neq \mathbb{O}$$

Let

$$w = \beta_1 w_1 + \beta_2 w_2 + \dots$$

$$\dots$$

$$u = \gamma_1 u_1 + \gamma_2 u_2 + \dots$$

Therefore,

$$v + w + \dots + u = \mathbb{O}$$

where  $v \neq \mathbb{O}$  and  $v \in V_{\lambda_1}, w \in V_{\lambda_2}, \dots, u \in V_{\lambda_s}$ .

But as  $\lambda_1, \dots, \lambda_s$  are pairwise distinct,  $v, w, \dots, u$  are linearly independent.

This is a contradiction. Therefore,  $B$  is a basis. Hence, as  $B$  consists of eigenvectors of  $A$ , by the General criterion for diagonalization,  $A$  is diagonalizable.  $\square$

**Theorem 10** (Criterion for triangularization). *An operator  $T : V \rightarrow V$  is triangularizable, i.e. there is a basis  $B$  of  $V$  such that  $[T]_B$  is upper triangular, if and only if  $p_T(x)$  splits completely.*

**Theorem 11** (Jordan Theorem). *Let  $T : V \rightarrow V$  be a linear operator such that  $p_T(x)$  splits completely. Then there exists a basis  $B$  of  $V$  such that  $[T]_B$  is of the form*

$$[T]_B = \begin{pmatrix} J_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_l \end{pmatrix}$$

where each  $J_i$  is of the form

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

where  $\lambda$  is some eigenvalue of  $T$ .

## Part VII

# Inner Product Spaces

## 1 Definition

**Definition 38** (Inner product). Let  $\mathbb{F}$  be  $\mathbb{R}$  or  $\mathbb{C}$ . Let  $V$  be a vector space over  $\mathbb{F}$ . An inner product on  $V$  is a function in two vector arguments with scalar values which associates to two given vectors  $v, w \in V$  their product  $\langle v, w \rangle \in \mathbb{F}$  so that the following properties are satisfied.

1.  $\langle \alpha_1 v_1 + \alpha_2 v_2, w \rangle = \alpha_1 \langle v_1, w \rangle + \alpha_2 \langle v_2, w \rangle, \forall v_1, v_2, w \in V, \forall \alpha_1, \alpha_2 \in \mathbb{F}$
2.  $\langle v, w \rangle = \overline{\langle w, v \rangle}, \forall v, w \in V$
3.  $\langle v, v \rangle$  is a real non-negative number,  $\forall v \in V$

**Example 12.** The dot product of two vectors is defined as follows. Is it an inner product?

$$V = \mathbb{F}^n$$

$$\left\langle \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \right\rangle = \alpha_1 \overline{\beta_1} + \cdots + \alpha_n \overline{\beta_n}$$

*Solution.* All three axioms are satisfied by this product. Hence, it is an inner product.

**Theorem 1** (Sesquilinearity).

$$\langle v, \beta_1 w_1 + \beta_2 w_2 \rangle = \overline{\beta_1} \langle v, w_1 \rangle + \overline{\beta_2} \langle v, w_2 \rangle$$

$$\forall v, w_1, w_2 \in V, \beta_1, \beta_2 \in \mathbb{F}$$

**Definition 39** (Length). The length of a vector

$$v = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

is defined to be

$$\|v\| = \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}$$

**Example 13.** Let  $V$  be the vector space consisting of all continuous functions  $f : [a, b] \rightarrow \mathbb{R}$ .

$$\langle f, g \rangle = \int_a^b f(x)g(x) \, dx$$

*Solution.* All three axioms are satisfied by this product. Hence, it is an inner product.

## 2 Computation of Inner Products

**Definition 40** (Gram matrix). Let  $V$  be an inner product space. Let

$$B = \{v_1, \dots, v_n\}$$

be a basis of  $V$ .

$$G_B = \begin{pmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_n \rangle \\ \vdots & & \vdots \\ \langle v_n, v_1 \rangle & \dots & \langle v_n, v_n \rangle \end{pmatrix}$$

is called the Gram matrix of the inner product with respect to  $B$ .

**Example 14.** Find the Gram matrix of  $V = \mathbb{F}^n$  with standard dot product with respect to

$$B = \left\{ \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

*Solution.*

$$\begin{aligned} G_B &= \begin{pmatrix} \langle e_1, e_1 \rangle & \dots & \langle e_1, e_n \rangle \\ \vdots & & \vdots \\ \langle e_n, e_1 \rangle & \dots & \langle e_n, e_n \rangle \end{pmatrix} \\ &= \begin{pmatrix} 1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 1 \end{pmatrix} \end{aligned}$$

**Example 15.** Find the Gram matrix of  $V = \mathbb{F}^n$  with standard dot product with respect to

$$B = \left\{ \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 6 \\ 7 \end{pmatrix} \right\}$$

*Solution.*

$$\begin{aligned} G_B &= \begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle \end{pmatrix} \\ &= \begin{pmatrix} 25 & 46 \\ 46 & 85 \end{pmatrix} \end{aligned}$$



**Theorem 2.**

$$\langle v, w \rangle = [v]_B^t G_B \overline{[w]}_B$$

*Proof.* Let

$$B = \{v_1, \dots, v_n\}$$

be a basis of  $V$ .

The Gram matrix is

$$G_B = \left( \langle v_i, v_j \rangle \right) = \left( g_{ij} \right)$$

To compute  $\langle v, w \rangle$ , find

$$[v]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

$$[w]_B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

$$\begin{aligned} \langle v, w \rangle &= \langle \alpha_1 v_1 + \dots + \alpha_n v_n, \beta_1 v_1 + \dots + \beta_n v_n \rangle \\ &= \alpha_1 \overline{\beta_1} \langle v_1, v_1 \rangle + \dots + \alpha_1 \overline{\beta_n} \langle v_1, v_n \rangle \\ &\quad + \alpha_2 \overline{\beta_1} \langle v_2, v_1 \rangle + \dots + \alpha_2 \overline{\beta_n} \langle v_2, v_n \rangle \\ &\quad + \dots \\ &\quad + \alpha_n \overline{\beta_1} \langle v_n, v_1 \rangle + \dots + \alpha_n \overline{\beta_n} \langle v_n, v_n \rangle \\ &= \alpha_1 g_{11} \overline{\beta_1} + \dots + \alpha_1 g_{1n} \overline{\beta_n} \\ &\quad + \alpha_2 g_{21} \overline{\beta_1} + \dots + \alpha_2 g_{2n} \overline{\beta_n} \\ &\quad + \dots \\ &\quad + \alpha_n g_{n1} \overline{\beta_1} + \dots + \alpha_n g_{nn} \overline{\beta_n} \\ &= [v]_B^t G_B \overline{[w]}_B \end{aligned}$$

□

## 2.1 Change of Basis

**Theorem 3.** Let  $B, \tilde{B}$  be bases of  $V$ . Let  $P$  be the transition matrix from  $B$  to  $\tilde{B}$ . Then

$$G_{\tilde{B}} = P^t G_B \overline{P}$$

where  $\overline{P}$  is the matrix obtained by replacing all elements of  $P$  by their complex conjugates.

*Proof.*

$$[v]_B = P[v]_{\tilde{B}}$$

$$\begin{aligned} \langle v, w \rangle &= [v]_B^t G_B \overline{[w]_B} \\ &= (P[v]_{\tilde{B}})^t G_B \overline{(P[w]_{\tilde{B}})} \\ &= [v]_{\tilde{B}}^t (P^t G_B \overline{P}) \overline{[w]_{\tilde{B}}} \end{aligned}$$

Also,

$$\langle v, w \rangle = [v]_{\tilde{B}}^t G_{\tilde{B}} \overline{[w]_{\tilde{B}}}$$

Therefore,

$$G_{\tilde{B}} = P^t G_B \overline{P}$$

□

## 3 Norms

### 3.1 Definition

**Definition 41** (Norm). Let  $V$  be a vector space over  $\mathbb{F}$  with inner product.  $\forall v \in V$ ,

$$\|v\| \doteq \sqrt{\langle v, v \rangle}$$

$\|v\|$  is called the norm of  $v$ .

### 3.2 Properties

1. Positivity  
 $\|v\| \geq 0, \forall v \in V$   
 $\|v\| = 0 \iff v = \mathbb{O}$
2. Homogeneity  
 $\|\alpha v\| = |\alpha| \|v\|, \forall v \in V, \forall \alpha \in \mathbb{F}$
3. Triangle Inequality  
 $\|u + v\| \leq \|u\| + \|v\|, \forall u, v \in V$

## 4 Orthogonality

### 4.1 Definition

**Definition 42** (Orthogonality). A vector  $u \in V$  is said to be orthogonal to  $v \in V$  if

$$\langle u, v \rangle = 0$$

It is denoted as  $u \perp v$ .

### 4.2 Properties

1. If  $u \perp v$ , then  $v \perp u$ .
2. If  $u \perp v, \alpha, \beta \in \mathbb{F}$ , then  $\alpha u \perp \beta v$ .
3.  $\mathbb{O} \perp v, \forall v \in V$ .

## 5 Orthogonal and Orthonormal Bases

Let  $V$  be a vector space over  $\mathbb{F}$  with an inner product. Let  $S \subset V$ .

**Definition 43** (Orthogonal set).  $S$  is said to be orthogonal if any two distinct vectors from  $S$  are orthogonal.

**Definition 44** (Orthonormal set).  $S$  is said to be orthonormal if it is orthogonal and the norm of every vector is 1.

**Definition 45** (Orthogonal basis).  $S$  is said to be an orthogonal basis of  $V$  if it is orthogonal and a basis of  $V$ .

**Definition 46** (Orthonormal basis).  $S$  is said to be an orthonormal basis of  $V$  if it is orthonormal and a basis of  $V$ .

**Theorem 4.** *Let  $S$  be an orthogonal set such that  $\mathbb{O} \notin S$ . Then  $S$  is linearly independent.*

*Proof.* Let

$$\begin{aligned}\alpha_1, \dots, \alpha_m &\in \mathbb{F} \\ v_1, \dots, v_m &\in S\end{aligned}$$

Let

$$\alpha_1 v_1 + \dots + \alpha_m v_m = \mathbb{O}$$

$S$  is linearly independent if and only if

$$\alpha_1 = \dots = \alpha_m = 0$$

$$\alpha_1 v_1 + \dots + \alpha_m v_m = \mathbb{O}$$

Multiplying both sides by  $v_1$ ,

$$\begin{aligned}\langle \alpha_1 v_1 + \dots + \alpha_m v_m, v_1 \rangle &= \langle \mathbb{O}, v_1 \rangle \\ \therefore \alpha_1 \langle v_1, v_1 \rangle + \dots + \alpha_m \langle v_m, v_1 \rangle &= 0\end{aligned}$$

As  $v_1, \dots, v_m$  are orthogonal,

$$\langle v_2, v_1 \rangle = \dots = \langle v_m, v_1 \rangle$$

$$\therefore \alpha_1 \langle v_1, v_1 \rangle = 0$$

As  $v_1 \neq \mathbb{O}$

$$\begin{aligned}\langle v_1, v_1 \rangle &\neq 0 \\ \therefore \alpha_1 &= 0\end{aligned}$$

Similarly,

$$\alpha_2 = \dots = \alpha_m = 0$$

□

**Corollary 4.1.** *Any orthonormal set is linearly independent.*

**Corollary 4.2.** *Any orthonormal set consisting of  $n = \dim V$  vectors is an orthonormal basis of  $V$ .*

**Example 16.** Is the set

$$S = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

orthonormal?

*Solution.* The norm of the elements of  $S$  is not 1. Hence  $S$  is not orthonormal.

**Theorem 5.** *Let  $B = \{v_1, \dots, v_n\}$  be an orthonormal basis of  $V$ . Let  $v \in V$ .*

*Let  $[v]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ . Then,*

$$\alpha_1 = \langle v, v_1 \rangle$$

$$\vdots$$

$$\alpha_n = \langle v, v_n \rangle$$

*Proof.*

$$\begin{aligned} v &= \alpha_1 v_1 + \dots + \alpha_n v_n \\ \therefore \langle v, v_1 \rangle &= \langle \alpha_1 v_1 + \dots + \alpha_n v_n, v_1 \rangle \\ &= \alpha_1 \langle v_1, v_1 \rangle + \dots + \alpha_n \langle v_n, v_1 \rangle \\ &= \alpha_1 \end{aligned}$$

Similarly, in general,  $\forall 1 \leq i \leq n$ ,

$$\langle v, v_i \rangle = \alpha_i$$

□

**Theorem 6** (Pythagoras Theorem). *Let  $B = \{v_1, \dots, v_n\}$  be an orthonormal*

*basis of  $V$ . Let  $v \in V$ . Let  $[v]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ . Then,*

$$\|v\|^2 = |\alpha_1|^2 + \dots + |\alpha_n|^2$$

*Proof.*

$$\begin{aligned}
\|v\|^2 &= \langle v, v \rangle \\
&= \langle \alpha_1 v_1 + \cdots + \alpha_n v_n, \alpha_1 v_1 + \cdots + \alpha_n v_n \rangle \\
&= \alpha_1 \overline{\alpha_1} + \cdots + \alpha_n \overline{\alpha_n} \\
&= |\alpha_1|^2 + \cdots + |\alpha_n|^2
\end{aligned}$$

□

## 6 Unitary Matrices

**Definition 47.** Let  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$ . Let  $A$  be an  $n \times n$  matrix.  $A$  is said to be a unitary matrix if

$$A^* = \overline{A}^t = A^{-1}$$

If  $\mathbb{F} = \mathbb{R}$ , unitary matrices are called orthogonal matrices.

1.  $I$  is a unitary matrix.
2. If  $A_1$  and  $A_2$  are unitary matrices, then  $(A_1 A_2)^* = A_2^* A_1^*$ .
3. If  $A$  is unitary,  $A^{-1}$  is also unitary.

**Theorem 7.** Let  $A$  be an  $n \times n$  matrix. Let  $v_1, \dots, v_n$  be the columns of  $A$ . Let  $r_1, \dots, r_n$  be the columns of  $A$ . Then the following are equivalent.

1.  $A$  is unitary.
2.  $\{v_1, \dots, v_n\}$  is an orthonormal basis of  $\mathbb{F}^n$ , with respect to standard dot product.
3.  $\{r_1, \dots, r_n\}$  is an orthonormal basis of  $\mathbb{F}^n$ , with respect to standard dot product.

*Proof.* As  $A$  is unitary,  $A^t$  is also unitary.

$$\begin{aligned}
(A^t)^* &= (\overline{A^t})^t \\
&= (A^*)^t \\
&= (A^{-1})^t \\
&= (A^t)^{-1}
\end{aligned}$$

$$\begin{aligned}
& A \text{ is unitary} \\
& \iff A^* = A^{-1} \\
& \iff AA^* = I \\
& \iff A\overline{A}^t = I \\
& \iff (A\overline{A}^t)_{ik} = I_{ik} \\
& \quad = \sum_{j=1}^n a_{ij}\overline{a_{jk}} \\
& \quad = r_i \cdot \overline{r_k} \iff \{r_1, \dots, r_n\} \text{ is an orthonormal basis}
\end{aligned}$$

□