Lecture Notes on Linear Algebra

Arbind K Lal Sukant Pati

July 10, 2018

Chapter 1

Introduction to Matrices

1.1 Definition of a Matrix

Definition 1.1.1. A rectangular array of numbers is called a **matrix**.

In this book, we shall mostly be concerned with complex numbers. The horizontal arrays of a matrix are called its **rows** and the vertical arrays are called its **columns**. Let A be a matrix having m rows and n columns. Then, A is said to have **order** $m \times n$ or is called a matrix of **size** $m \times n$ and can be represented in either of the following forms:

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

where a_{ij} is the entry at the intersection of the i^{th} row and j^{th} column. One writes $A \in \mathbb{M}_{m,n}(\mathbb{C})$ to say that A is an $m \times n$ matrix with complex entries, $A \in \mathbb{M}_{m,n}(\mathbb{R})$ to say that A is an $m \times n$ matrix with real entries and $A = [a_{ij}]$, when the order of the matrix is understood from the context. We will also use $A[i,:]$ to denote the i-th row of A, $A[:,j]$ to denote the j-th column of A and a_{ij} or $(A)_{ij}$, for the (i,j)-th entry of A.

For example, if $A = \begin{bmatrix} 1 & 3+\mathbf{i} & 7 \\ 4 & 5 & 6-5\mathbf{i} \end{bmatrix}$ then $A[1,:] = [1 \quad 3+\mathbf{i} \quad 7]$, $A[:,3] = \begin{bmatrix} 7 \\ 6-5\mathbf{i} \end{bmatrix}$ and $a_{22} = 5$. In general, in row vector commas are inserted to differentiate between entries. Thus, $A[1,:] = [1,\ 3+\mathbf{i},\ 7]$. A matrix having only one column is called a **column vector** and a matrix with only one row is called a **row vector**. All our vectors will be column vectors and will be represented by bold letters. Thus, $A[1,:]$ is a row vector and $A[:,3]$ is a column vector.

Definition 1.1.3. Two matrices $A = [a_{ij}], B = [b_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$ are said to be **equal** if $a_{ij} = b_{ij}$, for each $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

In other words, two matrices are said to be equal if they have the same order and their corresponding entries are equal.

1.1.1 Special Matrices

Definition 1.1.4. 1. A matrix in which each entry is zero is called a **zero-matrix**, denoted $\mathbf{0}$. For example,

$$\mathbf{0}_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \mathbf{0}_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

2. A matrix that has the same number of rows as the number of columns, is called a **square matrix**. A square matrix is said to have order n if it's order is $n \times n$ and is denoted either by writing $A \in \mathbb{M}_n(\mathbb{R})$ or $A \in \mathbb{M}_n(\mathbb{C})$, depending on whether the entries are real or complex numbers, respectively.

3. Let $A = [a_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$.

   (a) Then, the entries $a_{11}, a_{22}, \ldots, a_{nn}$ are called the diagonal entries and they constitute the **principal diagonal** of A.

   (b) Then, A is said to be a **diagonal** matrix if $a_{ij} = 0$ for $i \neq j$, denoted $\text{diag}(a_{11}, \ldots, a_{nn})$. For example, the zero matrix $\mathbf{0}_n$ and $\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$ are two diagonal matrices.

   (c) If $A = \text{diag}(a_{11}, \ldots, a_{nn})$ and $a_{ii} = d$ for all $i = 1, \ldots, n$ then the diagonal matrix A is called a **scalar matrix**.

   (d) Then, $A = \text{diag}(1, \ldots, 1)$ is called the **identity matrix**, denoted I_n, or in short I. For example, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

4. For $1 \leq i \leq n$, define $\mathbf{e}_i = I_n[:, i]$, a matrix of order $n \times 1$. Then, the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, where $\mathbf{e}_i \in \mathbb{M}_{n,1}(\mathbb{C})$, for $1 \leq i \leq n$, is called the **standard basis** of $\mathbb{C}^n$. Note that even though the order of the column vectors $\mathbf{e}_i$'s depend on n, we don't mention it as the size is understood from the context. For example, if $\mathbf{e}_1 \in \mathbb{C}^2$ then, $\mathbf{e}_1^T = [1, \ 0]$. If $\mathbf{e}_1 \in \mathbb{C}^3$ then, $\mathbf{e}_1^T = [1, \ 0, \ 0]$ and so on.

5. Let $A = [a_{ij}]$ be a square matrix.

   (a) Then, A is said to be an **upper triangular** matrix if $a_{ij} = 0$ for $i > j$.

   (b) Then, A is said to be a **lower triangular** matrix if $a_{ij} = 0$ for $i < j$.

   (c) Then, A is said to be **triangular** if it is an upper or a lower triangular matrix. For example, $\begin{bmatrix} 0 & 1 & 4 \\ 0 & 3 & -1 \\ 0 & 0 & -2 \end{bmatrix}$ is upper triangular, $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ is lower triangular and the matrices $\mathbf{0}, I$ are upper as well as lower triangular matrices.

6. An $m \times n$ matrix $A = [a_{ij}]$ is said to have an **upper triangular form** if $a_{ij} = 0$ for all $i > j$. For example, the matrices $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$

have upper triangular forms.

1.2 Operations on Matrices

Definition 1.2.1. Let $A = [a_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$.

1. Then, the **transpose** of A, denoted $A^T = [b_{ij}] \in \mathbb{M}_{n,m}(\mathbb{C})$ and $b_{ij} = a_{ji}$, for all i, j.
2. Then, the **conjugate transpose** of A, denoted $A^* = [c_{ij}] \in \mathbb{M}_{n,m}(\mathbb{C})$ and $c_{ij} = \overline{a_{ji}}$, for all i, j, where for $a \in \mathbb{C}$, $\overline{a}$ denotes the complex-conjugate of a.

Thus, if $\mathbf{x}$ is a column vector then $\mathbf{x}^T$ and $\mathbf{x}^*$ are row vectors and vice-versa. For example, if $A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix}$ then $A^* = A^T = \begin{bmatrix} 1 & 0 \\ 4 & 1 \\ 5 & 2 \end{bmatrix}$, whereas if $A = \begin{bmatrix} 1 & 4+\mathbf{i} \\ 0 & 1-\mathbf{i} \end{bmatrix}$ then $A^T = \begin{bmatrix} 1 & 0 \\ 4+\mathbf{i} & 1-\mathbf{i} \end{bmatrix}$ and $A^* = \begin{bmatrix} 1 & 0 \\ 4-\mathbf{i} & 1+\mathbf{i} \end{bmatrix}$. Note that $A^* \neq A^T$.

Theorem 1.2.2. *For any matrix A, $(A^*)^* = A$. Thus, $(A^T)^T = A$.*

Proof. Let $A = [a_{ij}]$, $A^* = [b_{ij}]$ and $(A^*)^* = [c_{ij}]$. Clearly, the order of A and $(A^*)^*$ is the same. Also, by definition $c_{ij} = \overline{b_{ji}} = \overline{\overline{a_{ij}}} = a_{ij}$ for all i, j and hence the result follows. ■

Definition 1.2.3. Let $A = [a_{ij}], B = [b_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, the **sum** of A and B, denoted $A + B$, is defined to be the matrix $C = [c_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$ with $c_{ij} = a_{ij} + b_{ij}$.

Definition 1.2.4. Let $A = [a_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, the **product** of $k \in \mathbb{C}$ with A, denoted kA, is defined as $kA = [ka_{ij}] = [a_{ij}k] = Ak$.

For example, if $A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix}$ then $5A = \begin{bmatrix} 5 & 20 & 25 \\ 0 & 5 & 10 \end{bmatrix}$ and $(2+\mathbf{i})A = \begin{bmatrix} 2+\mathbf{i} & 8+4\mathbf{i} & 10+5\mathbf{i} \\ 0 & 2+\mathbf{i} & 4+2\mathbf{i} \end{bmatrix}$.

Theorem 1.2.5. *Let $A, B, C \in \mathbb{M}_{m,n}(\mathbb{C})$ and let $k, \ell \in \mathbb{C}$. Then,*

1. $A + B = B + A$ *(commutativity).*
2. $(A + B) + C = A + (B + C)$ *(associativity).*
3. $k(\ell A) = (k\ell)A$.
4. $(k + \ell)A = kA + \ell A$.

Proof. Part 1.
Let $A = [a_{ij}]$ and $B = [b_{ij}]$. Then, by definition

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A$$

as complex numbers commute. The reader is required to prove the other parts as all the results follow from the properties of complex numbers. ■

Definition 1.2.6. Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$.

1. Then, the matrix $\mathbf{0}_{m\times n}$ is called the **additive identity** as $A+\mathbf{0}=\mathbf{0}+A=A$.

2. Then, there exists a matrix B with $A+B=\mathbf{0}$. This matrix B is called the **additive inverse** of A, and is denoted by $-A=(-1)A$.

1.2.1 Multiplication of Matrices

Definition 1.2.8. Let $A=[a_{ij}]\in\mathbb{M}_{m,n}(\mathbb{C})$ and $B=[b_{ij}]\in\mathbb{M}_{n,r}(\mathbb{C})$. Then, the **product** of A and B, denoted AB, is a matrix $C=[c_{ij}]\in\mathbb{M}_{m,r}(\mathbb{C})$ with

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj},\ \ 1\le i\le m,\ 1\le j\le r.$$

Thus, AB is defined if and only if **number of columns of** A = **number of rows of** B.

For example, if $A=\begin{bmatrix}a&b&c\\d&e&f\end{bmatrix}$ and $B=\begin{bmatrix}\alpha&\beta&\gamma&\delta\\x&y&z&t\\u&v&w&s\end{bmatrix}$ then

$$AB=\begin{bmatrix}a\alpha+bx+cu & a\beta+by+cv & a\gamma+bz+cw & a\delta+bt+cs\\ d\alpha+ex+fu & d\beta+ey+fv & d\gamma+ez+fw & d\delta+et+fs\end{bmatrix}. \tag{1.2.1}$$

Note that the rows of the matrix AB can be written directly as

$$\begin{aligned}(AB)[1,:] &= a\,[\alpha,\beta,\gamma,\delta] + b\,[x,y,z,t] + c\,[u,v,w,s] = aB[1,:] + bB[2,:] + cB[3,:]\\ (AB)[2,:] &= dB[1,:] + eB[2,:] + fB[3,:] \qquad (1.2.2)\end{aligned}$$

and similarly, the columns of the matrix AB can be written directly as

$$(AB)[:,1] = \begin{bmatrix} a\alpha + bx + cu \\ d\alpha + ex + fu \end{bmatrix} = \alpha\ A[:,1] + x\ A[:,2] + u\ A[:,3], \qquad (1.2.3)$$

$(AB)[:,2] = \beta\ A[:,1] + y\ A[:,2] + v\ A[:,3], \cdots, (AB)[:,4] = \delta\ A[:,1] + t\ A[:,2] + s\ A[:,3]$.

Remark 1.2.9. *Observe the following:*

1. *In this example, while AB is defined, the product BA is not defined. However, for square matrices A and B of the same order, both the product AB and BA are defined.*

2. *The product AB corresponds to operating (adding or subtracting multiples of different rows) on the* rows of the matrix *B(see Equation (1.2.2)). This is* **row method** *for calculating the matrix product.*

3. *The product AB also corresponds to operating (adding or subtracting multiples of different columns) on the* columns of the matrix *A (see Equation (1.2.3)). This is* **column method** *for calculating the matrix product.*

4. *Let A and B be two matrices such that the product AB is defined. Then, verify that*

   (a) *Then, verify that $(AB)[i,:] = A[i,:]B$. That is, the i-th row of AB is obtained by multiplying the i-th row of A with B.*

   (b) *Then, verify that $(AB)[:,j] = AB[:,j]$. That is, the j-th column of AB is obtained by multiplying A with the j-th column of B.*

   Hence,

$$AB \;=\; \begin{bmatrix} A[1,:]B \\ A[2,:]B \\ \vdots \\ A[n,:]B \end{bmatrix} = [A\ B[:,1],\ A\ B[:,2],\ \ldots,\ A\ B[:,p]]. \qquad (1.2.4)$$

Definition 1.2.12. Two square matrices A and B are said to **commute** if $AB = BA$.

Remark 1.2.13. *Note that if A is a square matrix of order n and if B is a scalar matrix of order n then $AB = BA$. In general, the matrix product is not commutative. For example, consider* $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$*. Then, verify that* $AB = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = BA$.

Theorem 1.2.14. *Suppose that the matrices A, B and C are so chosen that the matrix multiplications are defined.*

1. *Then, $(AB)C = A(BC)$. That is, the matrix multiplication is associative.*

2. *For any $k \in \mathbb{R}$, $(kA)B = k(AB) = A(kB)$.*

3. *Then, $A(B + C) = AB + AC$. That is, multiplication distributes over addition.*

4. *If $A \in \mathbb{M}_n(\mathbb{C})$ then $AI_n = I_n A = A$.*

Proof. Part 1. Let $A = [a_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$, $B = [b_{ij}] \in \mathbb{M}_{n,p}(\mathbb{C})$ and $C = [c_{ij}] \in \mathbb{M}_{p,q}(\mathbb{C})$. Then,

$$(BC)_{kj} = \sum_{\ell=1}^{p} b_{k\ell}c_{\ell j} \quad \text{and} \quad (AB)_{i\ell} = \sum_{k=1}^{n} a_{ik}b_{k\ell}.$$

Therefore,

$$\begin{aligned} \big(A(BC)\big)_{ij} &= \sum_{k=1}^{n} a_{ik}\big(BC\big)_{kj} = \sum_{k=1}^{n} a_{ik}\big(\sum_{\ell=1}^{p} b_{k\ell}c_{\ell j}\big) = \sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{ik}\big(b_{k\ell}c_{\ell j}\big) \\ &= \sum_{k=1}^{n}\sum_{\ell=1}^{p}\big(a_{ik}b_{k\ell}\big)c_{\ell j} = \sum_{\ell=1}^{p}\big(\sum_{k=1}^{n} a_{ik}b_{k\ell}\big)c_{\ell j} = \sum_{\ell=1}^{T}\big(AB\big)_{i\ell}c_{\ell j} = \big((AB)C\big)_{ij}. \end{aligned}$$

Using a similar argument, the next part follows. The other parts are left for the reader. ∎

1.2.2 Inverse of a Matrix

Definition 1.2.16. Let $A \in \mathbb{M}_n(\mathbb{C})$.

1. Then, a square matrix B is said to be a **left inverse** of A, if $BA = I_n$.
2. Then, a square matrix C is called a **right inverse** of A, if $AC = I_n$.
3. Then, A is said to be **invertible** (or is said to have an **inverse**) if there exists a matrix B such that $AB = BA = I_n$.

Lemma 1.2.17. *Let $A \in \mathbb{M}_n(\mathbb{C})$. If that there exist $B, C \in \mathbb{M}_n(\mathbb{C})$ such that $AB = I_n$ and $CA = I_n$ then $B = C$.*

Proof. Note that $C = CI_n = C(AB) = (CA)B = I_nB = B$. ∎

Remark 1.2.18. *Lemma 1.2.17 implies that whenever A is invertible, the inverse is unique. Thus, we denote the inverse of A by A^{-1}. That is, $AA^{-1} = A^{-1}A = I$.*

Theorem 1.2.20. *Let A and B be two invertible matrices. Then,*

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

Proof. Part 1. Let $B = A^{-1}$ be the inverse of A. Then, $AB = BA = I$. Thus, by definition, B is invertible and $B^{-1} = A$. Or equivalently, $(A^{-1})^{-1} = A$.

Part 2. By associativity $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = I = (B^{-1}A^{-1})(AB)$.

Part 3. As $AA^{-1} = A^{-1}A = I$, we get $(AA^{-1})^* = (A^{-1}A)^* = I^*$. Or equivalently, $(A^{-1})^*A^* = A^*(A^{-1})^* = I$. Thus, by definition $(A^*)^{-1} = (A^{-1})^*$. ■

We will again come back to the study of invertible matrices in Sections 2.2 and 2.3.1.

1.3 Some More Special Matrices

Definition 1.3.1. 1. For $1 \leq k \leq m$ and $1 \leq \ell \leq n$, define a matrix $\mathcal{E}_{k\ell} \in \mathbb{M}_{m,n}(\mathbb{C})$ by $(\mathcal{E}_{k\ell})_{ij} = \begin{cases} 1, & \text{if } (k,\ell) = (i,j) \\ 0, & \text{otherwise.} \end{cases}$ Then, the matrices $\mathcal{E}_{k\ell}$, for $1 \leq k \leq m$ and $1 \leq \ell \leq n$ are called the **standard basis** elements for $\mathbb{M}_{m,n}(\mathbb{C})$.

So, if $\mathcal{E}_{k\ell} \in \mathbb{M}_{2,3}(\mathbb{C})$ then $\mathcal{E}_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \mathcal{E}_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$ and $\mathcal{E}_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$.

2. Let $A \in \mathbb{M}_{m,n}(\mathbb{R})$.

   (a) Then, A is called **symmetric** if $A^T = A$. For example, $A = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$.

   (b) Then, A is called **skew-symmetric** if $A^T = -A$. For example, $A = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}$.

   (c) Then, A is called **orthogonal** if $AA^T = A^TA = I$. For example, $A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

(d) Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A is said to be a **permutation matrix** if A has exactly one non-zero entry, namely 1, in each row and column. For example, I_n, for each positive integer n, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ are permutation matrices. Verify that permutation matrices are Orthogonal matrices.

3. Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$.

   (a) Then, A is called **normal** if $A^*A = AA^*$. For example, $\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$ is a normal matrix.

   (b) Then, A is called **Hermitian** if $A^* = A$. For example, $A = \begin{bmatrix} 1 & 1+\mathbf{i} \\ 1-\mathbf{i} & 2 \end{bmatrix}$.

   (c) Then, A is called **skew-Hermitian** if $A^* = -A$. For example, $A = \begin{bmatrix} 0 & 1+\mathbf{i} \\ -1+\mathbf{i} & 0 \end{bmatrix}$.

   (d) Then, A is called **unitary** if $AA^* = A^*A = I$. For example, $A = \frac{1}{\sqrt{3}} \begin{bmatrix} 1+\mathbf{i} & 1 \\ -1 & 1-\mathbf{i} \end{bmatrix}$.

   Verify that Hermitian, skew-Hermitian and Unitary matrices are normal matrices.

4. Then, A is called **idempotent** if $A^2 = A$. For example, $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ is idempotent.

5. A vector $\mathbf{u} \in \mathbb{M}_{n,1}(\mathbb{C})$ such that $\mathbf{u}^*\mathbf{u} = 1$ is called a **unit vector**.

6. A matrix that is symmetric and idempotent is called a **projection** matrix. For example, let $\mathbf{u} \in \mathbb{M}_{n,1}(\mathbb{R})$ be a unit vector. Then, $A = \mathbf{u}\mathbf{u}^T$ is a symmetric and an idempotent matrix. Hence, A is a projection matrix. In particular, let $\mathbf{u} = \frac{1}{\sqrt{5}}[1, 2]^T$ and $A = \mathbf{u}\mathbf{u}^T$. Then, $\mathbf{u}^T\mathbf{u} = 1$ and for any vector $\mathbf{x} = [x_1, x_2]^T \in \mathbb{M}_{2,1}(\mathbb{R})$ note that

$$A\mathbf{x} = (\mathbf{u}\mathbf{u}^T)\mathbf{x} = \mathbf{u}(\mathbf{u}^T\mathbf{x}) = \frac{x_1 + 2x_2}{\sqrt{5}}\mathbf{u} = \left[\frac{x_1 + 2x_2}{5}, \frac{2x_1 + 4x_2}{5}\right]^T.$$

   Thus, $A\mathbf{x}$ is the foot of the perpendicular from the point $\mathbf{x}$ on the vector $[1 \;\; 2]^T$.

7. Fix a unit vector $\mathbf{a} \in \mathbb{M}_{n,1}(\mathbb{R})$ and let $A = 2\mathbf{a}\mathbf{a}^T - I_n$. Then, verify that $A \in \mathbb{M}_n(\mathbb{R})$ and $A\mathbf{y} = 2(\mathbf{a}^T\mathbf{y})\mathbf{a} - \mathbf{y}$, for all $\mathbf{y} \in \mathbb{R}^n$. This matrix is called the **reflection** matrix about the line containing the points $\mathbf{0}$ and $\mathbf{a}$.

8. Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A is said to be **nilpotent** if there exists a positive integer n such that $A^n = \mathbf{0}$. The least positive integer k for which $A^k = \mathbf{0}$ is called the **order of nilpotency**. For example, if $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{C})$ with a_{ij} equal to 1 if $i - j = 1$ and 0, otherwise then $A^n = \mathbf{0}$ and $A^\ell \neq \mathbf{0}$ for $1 \leq \ell \leq n - 1$.

1.3.1 Submatrix of a Matrix

Definition 1.3.3. For a positive integer k, let us denote $[k] = \{1, \ldots, k\}$. Also, let $A \in \mathbb{M}_{m \times n}(\mathbb{C})$.

1. Then, a matrix obtained by deleting some of the rows and/or columns of A is said to be a **submatrix** of A.

2. If $S \subseteq [m]$ and $T \subseteq [n]$ then by $\mathbf{A(S|T)}$, we denote the submatrix obtained from A by deleting the rows with indices in S and columns with indices in T. By $A[S, T]$, we mean $A(S^c|T^c)$, where S^c is the complement of S in $[m]$ and T^c is the complement of T in $[n]$. Whenever, S or T consist of a single element, then we just write the element. If $S = [m]$, then we write $A[S, T] = A[:, T]$ and if $T = [n]$ then $A[S, T] = A[S, :]$ which matches with our notation in Definition 1.1.1.

3. If $m = n$, the submatrix $A[S, S]$ is called a **principal submatrix** of A.

Theorem 1.3.5. *Let* $A = [a_{ij}] = [P\ Q]$ *and* $B = [b_{ij}] = \begin{bmatrix} H \\ K \end{bmatrix}$ *be defined as above. Then,*

$$AB = PH + QK.$$

Proof. The matrix products PH and QK are valid as the order of the matrices P, H, Q and K are respectively, $n \times r$, $r \times p$, $n \times (m - r)$ and $(m - r) \times p$. Also, the matrices PH and QK are of the same order and hence their sum is justified. Now, let $P = [P_{ij}]$, $Q = [Q_{ij}]$, $H = [H_{ij}]$, and $K = [K_{ij}]$. Then, for $1 \le i \le n$ and $1 \le j \le p$, we have

$$\begin{aligned} (AB)_{ij} &= \sum_{k=1}^{m} a_{ik}b_{kj} = \sum_{k=1}^{r} a_{ik}b_{kj} + \sum_{k=r+1}^{m} a_{ik}b_{kj} = \sum_{k=1}^{r} P_{ik}H_{kj} + \sum_{k=r+1}^{m} Q_{ik}K_{kj} \\ &= (PH)_{ij} + (QK)_{ij} = (PH + QK)_{ij}. \end{aligned}$$

Thus, the required result follows. ∎

Remark 1.3.6. *Theorem 1.3.5 is very useful due to the following reasons:*

1. *The order of the matrices* P, Q, H *and* K *are smaller than that of* A *or* B.
2. *The matrices* P, Q, H *and* K *can be further partitioned so as to form blocks that are either identity or zero or matrices that have nice forms. This partition may be quite useful during different matrix operations.*
3. *If we want to prove results using induction then after proving the initial step, one assumes the result for all* $r \times r$ *submatrices and then try to prove it for* $(r+1) \times (r+1)$ *submatrices.*

For example, if $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$ then $AB = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Suppose $A = \begin{array}{c c} & \begin{array}{cc} m_1 & m_2 \end{array} \\ \begin{array}{c} n_1 \\ n_2 \end{array} & \begin{bmatrix} P & Q \\ R & S \end{bmatrix} \end{array}$ and $B = \begin{array}{c c} & \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{array}{c} r_1 \\ r_2 \end{array} & \begin{bmatrix} E & F \\ G & H \end{bmatrix} \end{array}$. Then, the matrices P, Q, R, S and E, F, G, H, are called the blocks of the matrices A and B, respectively. Note that even if $A+B$ is defined, the orders of P and E need not be the same. But, if the block sums are defined then $A + B = \begin{bmatrix} P+E & Q+F \\ R+G & S+H \end{bmatrix}$. Similarly, if the product AB is defined, the product PE may not be defined. Again, if the block products are defined, one can verify that $AB = \begin{bmatrix} PE+QG & PF+QH \\ RE+SG & RF+SH \end{bmatrix}$. That is, once a partition of A is fixed, the partition of B has to be properly chosen for purposes of block addition or multiplication.

1.4 Summary

In this chapter, we started with the definition of a matrix and came across lots of examples. We recall these examples as they will be used in later chapters to relate different ideas:

1. The zero matrix of size $m \times n$, denoted $\mathbf{0}_{m \times n}$ or $\mathbf{0}$.
2. The identity matrix of size $n \times n$, denoted I_n or I.
3. Triangular matrices.
4. Hermitian/Symmetric matrices.
5. Skew-Hermitian/skew-symmetric matrices.
6. Unitary/Orthogonal matrices.
7. Idempotent matrices.

8. nilpotent matrices.

We also learnt product of two matrices. Even though it seemed complicated, it basically tells that multiplying by a matrix on the

1. left to a matrix A is same as operating on the rows of A.

2. right to a matrix A is same as operating on the columns of A.

Chapter 2

System of Linear Equations

Definition 2.1.3. **[Linear System]** A system of m linear equations in n variables $x_1, x_2, \ldots, x_n$ is a set of equations of the form

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array} \qquad (2.1.1)$$

where for $1 \leq i \leq m$ and $1 \leq j \leq n$; $a_{ij}, b_i \in \mathbb{R}$. Linear System (2.1.1) is called **homogeneous** if $b_1 = 0 = b_2 = \cdots = b_m$ and **non-homogeneous**, otherwise.

Definition 2.1.4. **[Coefficient and Augmented Matrices]** Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$. Then, (2.1.1) can be re-written as $A\mathbf{x} = \mathbf{b}$. In this setup, the matrix

A is called the **coefficient** matrix and the block matrix $[A\ \ \mathbf{b}]$ is called the **augmented** matrix of the linear system (2.1.1).

Remark 2.1.5. *Consider the linear system $A\mathbf{x} = \mathbf{b}$, where $A \in \mathbb{M}_{m,n}(\mathbb{C})$, $\mathbf{b} \in \mathbb{M}_{m,1}(\mathbb{C})$ and $\mathbf{x} \in \mathbb{M}_{n,1}(\mathbb{C})$. If $[A\,\mathbf{b}]$ is the augmented matrix and $\mathbf{x}^T = [x_1, \ldots, x_n]$ then,*

1. *for $j = 1, 2, \ldots, n$, the variable x_j corresponds to the column $([A\,\mathbf{b}])[:, j]$.*
2. *the vector $\mathbf{b} = ([A\,\mathbf{b}])[:, n+1]$.*
3. *for $i = 1, 2, \ldots, m$, the i^{th} equation corresponds to the row $([A\,\mathbf{b}])[i, :]$.*

Definition 2.1.6. **[Solution of a Linear System]** A **solution** of $A\mathbf{x} = \mathbf{b}$ is a vector $\mathbf{y}$ such that $A\mathbf{y}$ indeed equals $\mathbf{b}$. The set of all solutions is called the **solution set** of the system. For example, the solution set of $A\mathbf{x} = \mathbf{b}$, with $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 2 \\ 4 & 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ equals $\left\{ \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \right\}$.

Definition 2.1.7. **[Consistent, Inconsistent]** Consider a linear system $A\mathbf{x} = \mathbf{b}$. Then, this linear system is called **consistent** if it admits a solution and is called **inconsistent** if it admits no solution. For example, the homogeneous system $A\mathbf{x} = \mathbf{0}$ is always consistent as $\mathbf{0}$ is a solution whereas, verify that the system $x + y = 2, 2x + 2y = 3$ is inconsistent.

Definition 2.1.8. **[Associated Homogeneous System]** Consider a linear system $A\mathbf{x} = \mathbf{b}$. Then, the corresponding linear system $A\mathbf{x} = \mathbf{0}$ is called the **associated homogeneous system**. $\mathbf{0}$ is always a solution of the associated homogeneous system.

The readers are advised to supply the proof of the next theorem that gives information about the solution set of a homogeneous system.

Theorem 2.1.9. *Consider a homogeneous linear system $A\mathbf{x} = \mathbf{0}$.*

1. *Then, $\mathbf{x} = \mathbf{0}$, the zero vector, is always a solution, called the* **trivial** *solution.*
2. *Let $\mathbf{u} \neq \mathbf{0}$ be a solution of $A\mathbf{x} = \mathbf{0}$. Then, $\mathbf{y} = c\mathbf{u}$ is also a solution, for all $c \in \mathbb{C}$. A nonzero solution is called a* **non-trivial** *solution. Note that, in this case, the system $A\mathbf{x} = \mathbf{0}$ has an infinite number of solutions.*
3. *Let $\mathbf{u_1}, \ldots, \mathbf{u}_k$ be solutions of $A\mathbf{x} = \mathbf{0}$. Then, $\sum\limits_{i=1}^{k} a_i\mathbf{u}_i$ is also a solution of $A\mathbf{x} = \mathbf{0}$, for each choice of $a_i \in \mathbb{C}, 1 \leq i \leq k$.*

Remark 2.1.10. 1. *Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Then, $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is a non-trivial solution of $A\mathbf{x} = \mathbf{0}$.*

2. *Let $\mathbf{u} \neq \mathbf{v}$ be solutions of a non-homogeneous system $A\mathbf{x} = \mathbf{b}$. Then, $\mathbf{x}_h = \mathbf{u} - \mathbf{v}$ is a solution of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$. That is, any two distinct solutions of $A\mathbf{x} = \mathbf{b}$ differ by a solution of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$. Or equivalently, the solution set of $A\mathbf{x} = \mathbf{b}$ is of the form, $\{\mathbf{x}_0 + \mathbf{x}_h\}$, where $\mathbf{x}_0$ is a particular solution of $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x}_h$ is a solution of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$.*

2.1.1 Elementary Row Operations

Definition 2.1.13. **[Elementary Row Operations]** Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, the **elementary row operations** are

1. E_{ij}: Interchange the i-th and j-th rows, namely, interchange $A[i,:]$ and $A[j,:]$.
2. $E_k(c)$ for $c \neq 0$: Multiply the k-th row by c, namely, multiply $A[k,:]$ by c.
3. $E_{ij}(c)$ for $c \neq 0$: Replace the i-th row by i-th row plus c-times the j-th row, namely, replace $A[i,:]$ by $A[i,:] + cA[j,:]$.

Definition 2.1.14. **[Row Equivalent Matrices]** Two matrices are said to be **row equivalent** if one can be obtained from the other by a finite number of elementary row operations.

Definition 2.1.15. **[Row Equivalent Linear Systems]** The linear systems $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ are said to be **row equivalent** if their respective augmented matrices, $[A \;\; \mathbf{b}]$ and $[C \;\; \mathbf{d}]$, are row equivalent.

Thus, note that the linear systems at each step in Example 2.1.12 are row equivalent to each other. We now prove that the solution set of two row equivalent linear systems are same.

Lemma 2.1.16. *Let $C\mathbf{x} = \mathbf{d}$ be the linear system obtained from $A\mathbf{x} = \mathbf{b}$ by application of a single elementary row operation. Then, $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have the same solution set.*

Proof. We prove the result for the elementary row operation $E_{jk}(c)$ with $c \neq 0$. The reader is advised to prove the result for the other two elementary operations.

In this case, the systems $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ vary only in the j^{th} equation. So, we need to show that $\mathbf{y}$ satisfies the j^{th} equation of $A\mathbf{x} = \mathbf{b}$ if and only if $\mathbf{y}$ satisfies the j^{th} equation of $C\mathbf{x} = \mathbf{d}$. So, let $\mathbf{y}^T = [\alpha_1, \ldots, \alpha_n]$. Then, the j^{th} and k^{th} equations of $A\mathbf{x} = \mathbf{b}$ are $a_{j1}\alpha_1 + \cdots + a_{jn}\alpha_n = b_j$ and $a_{k1}\alpha_1 + \cdots + a_{kn}\alpha_n = b_k$. Therefore, we see that α_i's satisfy

$$(a_{j1} + ca_{k1})\alpha_1 + \cdots + (a_{jn} + ca_{kn})\alpha_n = b_j + cb_k. \tag{2.1.2}$$

Also, by definition the j^{th} equation of $C\mathbf{x} = \mathbf{d}$ equals

$$(a_{j1} + ca_{k1})x_1 + \cdots + (a_{jn} + ca_{kn})x_n = b_j + cb_k. \tag{2.1.3}$$

Therefore, using Equation (2.1.2), we see that $\mathbf{y}^T = [\alpha_1, \ldots, \alpha_n]$ is also a solution for Equation (2.1.3). Now, use a similar argument to show that if $\mathbf{z}^T = [\beta_1, \ldots, \beta_n]$ is a solution of $C\mathbf{x} = \mathbf{d}$ then it is also a solution of $A\mathbf{x} = \mathbf{b}$. Hence, the required result follows. ■

The readers are advised to use Lemma 2.1.16 as an induction step to prove the next result.

Theorem 2.1.17. *Let $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ be two row equivalent linear systems. Then, they have the same solution set.*

2.2 Main Ideas of Linear Systems

In the previous section, we saw that two row equivalent linear systems have the same solution set. Sometimes it helps to imagine an elementary row operation as left multiplication by a suitable matrix. In this section, we will try to understand this relationship and use them to obtain results for linear system. As special cases, we also obtain results that are very useful in the study of square matrices.

2.2.1 Elementary Matrices and the Row-Reduced Echelon Form (RREF)

Definition 2.2.1. [**Elementary Matrix**] A matrix $E \in \mathbb{M}_n(\mathbb{C})$ is called an **elementary matrix** if it is obtained by applying exactly one elementary row operation to the identity matrix I_n.

Remark 2.2.2. *The elementary matrices are of three types and they correspond to elementary row operations.*

1. *E_{ij}: Matrix obtained by applying elementary row operation E_{ij} to I_n.*
2. *$E_k(c)$ for $c \neq 0$: Matrix obtained by applying elementary row operation $E_k(c)$ to I_n.*
3. *$E_{ij}(c)$ for $c \neq 0$: Matrix obtained by applying elementary row operation $E_{ij}(c)$ to I_n.*

When an elementary matrix is multiplied on the left of a matrix A, it gives the same result as that of applying the corresponding elementary row operation on A.

Remark 2.2.5. *Observe that*

1. $(E_{ij})^{-1} = E_{ij}$ *as* $E_{ij}E_{ij} = I = E_{ij}E_{ij}$.
2. *Let* $c \neq 0$. *Then,* $(E_k(c))^{-1} = E_k(1/c)$ *as* $E_k(c)E_k(1/c) = I = E_k(1/c)E_k(c)$.
3. *Let* $c \neq 0$. *Then,* $(E_{ij}(c))^{-1} = E_{ij}(-c)$ *as* $E_{ij}(c)E_{ij}(-c) = I = E_{ij}(-c)E_{ij}(c)$.

   Thus, each elementary matrix is invertible. Also, the inverse is an elementary matrix of the same type.

Proposition 2.2.6. *Let A and B be two row equivalent matrices. Then, prove that $B = E_1 \cdots E_k A$, for some elementary matrices $E_1, \ldots, E_k$.*

Proof. By definition of row equivalence, the matrix B can be obtained from A by a finite number of elementary row operations. But by Remark 2.2.2, each elementary row operation on A corresponds to left multiplication by an elementary matrix to A. Thus, the required result follows. ∎

We now give an alternate prove of Theorem 2.1.17. To do so, we state the theorem once again.

Theorem 2.2.7. *Let $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ be two row equivalent linear systems. Then, they have the same solution set.*

Proof. Let $E_1, \ldots, E_k$ be the elementary matrices such that $E_1 \cdots E_k[A \;\; \mathbf{b}] = [C \;\; \mathbf{d}]$. Put $E = E_1 \cdots E_k$. Then, by Remark 2.2.5

$$EA = C, \quad E\mathbf{b} = \mathbf{d}, \quad A = E^{-1}C \text{ and } \mathbf{b} = E^{-1}\mathbf{d}. \tag{2.2.1}$$

Now assume that $A\mathbf{y} = \mathbf{b}$ holds. Then, by Equation (2.2.1)

$$C\mathbf{y} = EA\mathbf{y} = E\mathbf{b} = \mathbf{d}. \tag{2.2.2}$$

On the other hand if $C\mathbf{z} = \mathbf{d}$ holds then using Equation (2.2.1), we have

$$A\mathbf{z} = E^{-1}C\mathbf{z} = E^{-1}\mathbf{d} = \mathbf{b}. \tag{2.2.3}$$

Therefore, using Equations (2.2.2) and (2.2.3) the required result follows. ∎

The following result is a particular case of Theorem 2.2.7.

Corollary 2.2.8. *Let A and B be two row equivalent matrices. Then, the systems $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same solution set.*

Definition 2.2.10. [**Pivot/Leading Entry**] Let A be a nonzero matrix. Then, in each nonzero row of A, the left most nonzero entry is called a **pivot/leading entry**. The column containing the pivot is called a **pivotal column**. If a_{ij} is a pivot then we denote it by $\boxed{a_{ij}}$. For example, the entries a_{12} and a_{23} are pivots in $A = \begin{bmatrix} 0 & \boxed{3} & 4 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{2} & 1 \end{bmatrix}$. Thus, columns 2 and 3 are pivotal columns.

Definition 2.2.11. [**Row Echelon Form**] A matrix is in **row echelon form (REF)** (ladder like)

1. if the zero rows are at the bottom;
2. if the pivot of the $(i+1)$-th row, if it exists, comes to the right of the pivot of the i-th row.
3. if the entries below the pivot in a pivotal column are 0.

Definition 2.2.13. [**Row-Reduced Echelon Form (RREF)**] A matrix C is said to be in **row-reduced echelon form (RREF)**

1. if C is already in echelon form,
2. if the pivot of each nonzero row is 1,
3. if every other entry in each pivotal column is zero.

A matrix in RREF is also called a row-reduced echelon matrix.

Theorem 2.2.17. *Let A and B be two row equivalent matrices in RREF. Then $A = B$.*

As an immediate corollary, we obtain the following important result.

Corollary 2.2.18. *The RREF of a matrix A is unique.*

Proof. Suppose there exists a matrix A with two different RREFs, say B and C. As the RREFs are obtained by left multiplication of elementary matrices, there exist elementary matrices $E_1, \ldots, E_k$ and $F_1, \ldots, F_\ell$ such that $B = E_1 \cdots E_k A$ and $C = F_1 \cdots F_\ell A$. Let $E = E_1 \cdots E_k$ and $F = F_1 \cdots F_\ell$. Thus, $B = EA = EF^{-1}C$.

As inverse of an elementary matrix is an elementary matrix, F^{-1} is a product of elementary matrices and hence, B and C are row equivalent. As B and C are in RREF, using Theorem 2.2.17, $B = C$. ∎

Remark 2.2.19. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$.*

1. *Then, by Corollary 2.2.18, it's RREF is unique.*
2. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, the uniqueness of RREF implies that RREF(A) is independent of the choice of the row operations used to get the final matrix which is in RREF.*
3. *Let $B = EA$, for some elementary matrix E. Then, RREF$(A) =$ RREF(B).*

Proof. *Let* $E_1, \ldots, E_k$ *and* $F_1, \ldots, F_\ell$ *be elementary matrices such that* $RREF(A) = E_1 \cdots E_k A$ *and* $RREF(B) = F_1 \cdots F_\ell B$. *Then,*

$$RREF(B) = F_1 \cdots F_\ell B = (F_1 \cdots F_\ell)EA = (F_1 \cdots F_\ell)E(E_k^{-1} \cdots E_1^{-1})RREF(A).$$

Thus, the matrices $RREF(A)$ *and* $RREF(B)$ *are row equivalent. Since they are also in RREF by Theorem 2.2.17,* $RREF(A) = RREF(B)$. ■

4. *Then, there exists an invertible matrix* P*, a product of elementary matrices, such that* $PA = RREF(A)$.

   Proof. *By definition,* $RREF(A) = E_1 \cdots E_k A$*, for certain elementary matrices* $E_1, \ldots, E_k$. *Take* $P = E_1 \cdots E_k$. *Then,* P *is invertible (product of invertible matrices is invertible) and* $PA = RREF(A)$. ■

5. *Let* $F = RREF(A)$ *and* $B = [A[:,1], \ldots, A[:,s]]$*, for some* $s \leq n$. *Then,*

$$RREF(B) = [F[:,1], \ldots, F[:,s]].$$

Proof. *By Remark 2.2.19.4, there exist an invertible matrix* P*, such that*

$$F = PA = [PA[:,1], \ldots, PA[:,n]] = [F[:,1], \ldots, F[:,n]].$$

Thus, $PB = [PA[:,1], \ldots, PA[:,s]] = [F[:,1], \ldots, F[:,s]]$. *As* F *is in RREF, it's first* s *columns are also in RREF. Hence, by Corollary 2.2.18,* $RREF(PB) = [F[:,1], \ldots, F[:,s]]$. *Now, a repeated use of Remark 2.2.19.3 gives* $RREF(B) = [F[:,1], \ldots, F[:,s]]$. *Thus, the required result follows.* ■

Proposition 2.2.21. *Let* $A \in \mathbb{M}_n(\mathbb{C})$. *Then,* A *is invertible if and only if* $RREF(A) = I_n$. *That is, every invertible matrix is a product of elementary matrices.*

Proof. If $\text{RREF}(A) = I_n$ then $I_n = E_1 \cdots E_k A$, for some elementary matrices $E_1, \ldots, E_k$. As E_i's are invertible, $E_1^{-1} = E_2 \cdots E_k A$, $E_2^{-1} E_1^{-1} = E_3 \cdots E_k A$ and so on. Finally, one obtains $A = E_k^{-1} \cdots E_1^{-1}$. A similar calculation now gives $AE_1 \cdots E_k = I_n$. Hence, by definition of invertibility $A^{-1} = E_1 \cdots E_k$.

Now, let A be invertible with $B = \text{RREF}(A) = E_1 \cdots E_k A$, for some elementary matrices $E_1, \ldots, E_k$. As A and E_i's are invertible, the matrix B is invertible. Hence, B doesn't have any zero row. Thus, all the n rows of B have pivots. Therefore, B has n pivotal columns. As B has exactly n columns, each column is a pivotal column and hence $B = I_n$. Thus, the required result follows. ∎

As a direct application of Proposition 2.2.21 and Remark 2.2.19.3 one obtains the following.

Theorem 2.2.22. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, for any invertible matrix S, RREF$(SA) =$ RREF(A).*

Proposition 2.2.23. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be an invertible matrix. Then, for any matrix B, define $C = \begin{bmatrix} A & B \end{bmatrix}$ and $D = \begin{bmatrix} A \\ B \end{bmatrix}$. Then, RREF$(C) = \begin{bmatrix} I_n & A^{-1}B \end{bmatrix}$ and RREF$(D) = \begin{bmatrix} I_n \\ \mathbf{0} \end{bmatrix}$.*

Proof. Using matrix product,

$$A^{-1}C = \begin{bmatrix} A^{-1}A & A^{-1}B \end{bmatrix} = \begin{bmatrix} I_n & A^{-1}B \end{bmatrix}.$$

As $\begin{bmatrix} I_n & A^{-1}B \end{bmatrix}$ is in RREF, by Remark 2.2.19.1, $\text{RREF}(C) = \begin{bmatrix} I_n & A^{-1}B \end{bmatrix}$.

For the second part, note that the matrix $X = \begin{bmatrix} A^{-1} & \mathbf{0} \\ -BA^{-1} & I_n \end{bmatrix}$ is an invertible matrix. Thus, by Proposition 2.2.21, X is a product of elementary matrices. Now, verify that $XD = \begin{bmatrix} I_n \\ \mathbf{0} \end{bmatrix}$. As $\begin{bmatrix} I_n \\ \mathbf{0} \end{bmatrix}$ is in RREF, a repeated application of Remark 2.2.19.1 gives the required result. ∎

As an application of Proposition 2.2.23, we have the following observation.

Let $A \in \mathbb{M}_n(\mathbb{C})$. Suppose we start with $C = [A \;\; I_n]$ and compute RREF(C). If RREF$(C) = [G \;\; H]$ then, either $G = I_n$ or $G \neq I_n$. Thus, if $G = I_n$ then we must have $F = A^{-1}$. If $G \neq I_n$ then, A is not invertible. We explain this with an example.

2.2.2 Rank of a Matrix

Definition 2.2.26. [Rank of a Matrix] Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, the **rank** of A, denoted $\mathsf{Rank}(A)$, is the number of pivots in the RREF(A). For example, $\mathsf{Rank}(I_n) = n$ and $\mathsf{Rank}(\mathbf{0}) = 0$.

Remark 2.2.27. *Before proceeding further, for $A \in \mathbb{M}_{m,n}(\mathbb{C})$, we observe the following.*

1. *The number of pivots in the RREF(A) is same as the number of pivots in REF of A. Hence, we need not compute the RREF(A) to determine the rank of A.*
2. *Since, the number of pivots cannot be more than the number of rows or the number of columns, one has $\mathsf{Rank}(A) \le \min\{m, n\}$.*
3. *If $B = \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ then $\mathsf{Rank}(B) = \mathsf{Rank}(A)$ as $RREF(B) = \begin{bmatrix} RREF(A) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$.*
4. *If $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ then, by definition*
$$\mathsf{Rank}(A) \le \mathsf{Rank}\left(\begin{bmatrix} A_{11} & A_{12} \end{bmatrix}\right) + \mathsf{Rank}\left(\begin{bmatrix} A_{21} & A_{22} \end{bmatrix}\right).$$
   Further, using Remark 2.2.19,
   (a) $\mathsf{Rank}(A) \ge \mathsf{Rank}\left(\begin{bmatrix} A_{11} & A_{12} \end{bmatrix}\right)$.
   (b) $\mathsf{Rank}(A) \ge \mathsf{Rank}\left(\begin{bmatrix} A_{21} & A_{22} \end{bmatrix}\right)$.
   (c) $\mathsf{Rank}(A) \ge \mathsf{Rank}\left(\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}\right)$.

We now illustrate the calculation of the rank by giving a few examples.

Lemma 2.2.29. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. If S is an invertible matrix then $\mathsf{Rank}(SA) = \mathsf{Rank}(A)$.*

Proof. By Theorem 2.2.22, $\text{RREF}(A) = \text{RREF}(SA)$. Hence, $\mathsf{Rank}(SA) = \mathsf{Rank}(A)$. ∎

We now have the following result.

Corollary 2.2.30. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$ and $B \in \mathbb{M}_{n,q}(\mathbb{C})$. Then, $\mathsf{Rank}(AB) \le \mathsf{Rank}(A)$.*

In particular, if $B \in \mathbb{M}_n(\mathbb{C})$ is invertible then $\mathsf{Rank}(AB) = \mathsf{Rank}(A)$.

Proof. Let $\mathsf{Rank}(A) = r$. Then, there exists an invertible matrix P and $A_1 \in \mathbb{M}_{r,n}(\mathbb{C})$ such that $PA = \text{RREF}(A) = \begin{bmatrix} A_1 \\ \mathbf{0} \end{bmatrix}$. Then, $PAB = \begin{bmatrix} A_1 \\ \mathbf{0} \end{bmatrix} B = \begin{bmatrix} A_1 B \\ \mathbf{0} \end{bmatrix}$. So, using Lemma 2.2.29 and Remark 2.2.27.2, we get

$$\mathsf{Rank}(AB) = \mathsf{Rank}(PAB) = \mathsf{Rank}\left(\begin{bmatrix} A_1 B \\ \mathbf{0} \end{bmatrix}\right) = \mathsf{Rank}(A_1 B) \le r = \mathsf{Rank}(A). \tag{2.2.4}$$

In particular, if B is invertible then, using Equation (2.2.4), we get

$$\mathsf{Rank}(A) = \mathsf{Rank}(ABB^{-1}) \le \mathsf{Rank}(AB)$$

and hence the required result follows. ∎

Theorem 2.2.31. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. If $\mathsf{Rank}(A) = r$ then, there exist invertible matrices P and Q such that*

$$P\ A\ Q = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Proof. Let $C = \text{RREF}(A)$. Then, by Remark 2.2.19.4 there exists as invertible matrix P such that $C = PA$. Note that C has r pivots and they appear in columns, say $i_1 < i_2 < \cdots < i_r$.

Now, let $D = CE_{1i_1}E_{2i_2}\cdots E_{ri_r}$. As E_{ji_j}'s are elementary matrices that interchange the columns of C, one has $D = \begin{bmatrix} I_r & B \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, where $B \in \mathbb{M}_{r,n-r}(\mathbb{C})$.

Put $Q_1 = E_{1i_1}E_{2i_2}\cdots E_{ri_r}$. Then, Q_1 is invertible. Let $Q_2 = \begin{bmatrix} I_r & -B \\ \mathbf{0} & I_{n-r} \end{bmatrix}$. Then, verify that Q_2 is invertible and

$$CQ_1Q_2 = DQ_2 = \begin{bmatrix} I_r & B \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} I_r & -B \\ 0 & I_{n-r} \end{bmatrix} = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Thus, if we put $Q = Q_1Q_2$ then Q is invertible and $PAQ = CQ = CQ_1Q_2 = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ and hence, the required result follows. ∎

We now prove the following result.

Proposition 2.2.32. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be an invertible matrix.*

1. *If $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$ with $A_1 \in \mathbb{M}_{n,r}(\mathbb{C})$ and $A_2 \in \mathbb{M}_{n,n-r}(\mathbb{C})$ then $\mathsf{Rank}(A_1) = r$ and $\mathsf{Rank}(A_2) = n - r$.*

2. *If* $A = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ *with* $B_1 \in \mathbb{M}_{s,n}(\mathbb{C})$ *and* $B_2 \in \mathbb{M}_{n-s,n}(\mathbb{C})$ *then* $\mathsf{Rank}(B_1) = s$ *and* $\mathsf{Rank}(B_2) = n - s$.

In particular, if $B = A[S,:]$ *and* $C = A[:,T]$*, for some subsets* S, T *of* $[n]$ *then* $\mathsf{Rank}(B) = |S|$ *and* $\mathsf{Rank}(C) = |T|$.

Proof. Since A is invertible, $\text{RREF}(A) = I_n$. Hence, by Remark 2.2.19.4, there exists an invertible matrix P such that $PA = I_n$. Thus,

$$\begin{bmatrix} PA_1 & PA_2 \end{bmatrix} = P\begin{bmatrix} A_1 & A_2 \end{bmatrix} = PA = I_n = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & I_{n-r} \end{bmatrix}.$$

Thus, $PA_1 = \begin{bmatrix} I_r \\ \mathbf{0} \end{bmatrix}$ and $PA_2 = \begin{bmatrix} \mathbf{0} \\ I_{n-r} \end{bmatrix}$. So, using Corollary 2.2.30, $\mathsf{Rank}(A_1) = r$. Also, note that $\begin{bmatrix} \mathbf{0} & I_{n-r} \\ I_r & \mathbf{0} \end{bmatrix}$ is an invertible matrix and

$$\begin{bmatrix} \mathbf{0} & I_{n-r} \\ I_r & \mathbf{0} \end{bmatrix} PA_2 = \begin{bmatrix} \mathbf{0} & I_{n-r} \\ I_r & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ I_{n-r} \end{bmatrix} = \begin{bmatrix} I_{n-r} \\ \mathbf{0} \end{bmatrix}.$$

So, again by using Corollary 2.2.30, $\mathsf{Rank}(A_2) = n - r$, completing the proof of the first part.

For the second part, let us assume that $\mathsf{Rank}(B_1) = t < s$. Then, by Remark 2.2.19.4, there exists an invertible matrix Q such that

$$QB_1 = \text{RREF}(B_1) = \begin{bmatrix} C \\ \mathbf{0} \end{bmatrix}, \tag{2.2.5}$$

for some matrix C, where C is in RREF and has exactly t pivots. Since $t < s$, QB_1 has at least one zero row.

As $PA = I_n$, one has $AP = I_n$. Hence, $\begin{bmatrix} B_1P \\ B_2P \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} P = AP = I_n = \begin{bmatrix} I_s & \mathbf{0} \\ \mathbf{0} & I_{n-s} \end{bmatrix}$. Thus,

$$B_1P = \begin{bmatrix} I_s & \mathbf{0} \end{bmatrix} \text{ and } B_2P = \begin{bmatrix} \mathbf{0} & I_{n-s} \end{bmatrix}. \tag{2.2.6}$$

Further, using Equations (2.2.5) and (2.2.6), we see that

$$\begin{bmatrix} CP \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} C \\ \mathbf{0} \end{bmatrix} P = QB_1P = Q\begin{bmatrix} I_s & \mathbf{0} \end{bmatrix} = \begin{bmatrix} Q & \mathbf{0} \end{bmatrix}.$$

Thus, Q has a zero row, contradicting the assumption that Q is invertible. Hence, $\mathsf{Rank}(B_1) = s$. Similarly, $\mathsf{Rank}(B_2) = n - s$ and thus, the required result follows. ∎

As a direct corollary of Theorem 2.2.31 and Proposition 2.2.32, we have the following result which improves Corollary 2.2.30.

Corollary 2.2.33. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{C})$*. If* $\mathsf{Rank}(A) = r < n$ *then, there exists an invertible matrix* Q *and* $B \in \mathbb{M}_{m,r}(\mathbb{C})$ *such that* $AQ = \begin{bmatrix} B & \mathbf{0} \end{bmatrix}$*, where* $\mathsf{Rank}(B) = r$.

Proof. By Theorem 2.2.31, there exist invertible matrices P and Q such that $PAQ = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$. If $P^{-1} = \begin{bmatrix} B & C \end{bmatrix}$, where $B \in \mathbb{M}_{m,r}(\mathbb{C})$ and $C \in \mathbb{M}_{m,m-r}(\mathbb{C})$ then,

$$AQ = P^{-1} \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} B & C \end{bmatrix} \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} B & \mathbf{0} \end{bmatrix}.$$

Now, by Proposition 2.2.32, $\mathsf{Rank}(B) = r = \mathsf{Rank}(A)$ as the matrix $P^{-1} = \begin{bmatrix} B & C \end{bmatrix}$ is an invertible matrix. Thus, the required result follows. ∎

As an application of Corollary 2.2.33, we have the following result.

Corollary 2.2.34. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$ and $B \in \mathbb{M}_{n,p}(\mathbb{C})$. Then,* $\mathsf{Rank}(AB) \leq \mathsf{Rank}(B)$.

Proof. Let $\mathsf{Rank}(B) = r$. Then, by Corollary 2.2.33, there exists an invertible matrix Q and a matrix $C \in \mathbb{M}_{n,r}(\mathbb{C})$ such that $BQ = \begin{bmatrix} C & \mathbf{0} \end{bmatrix}$ and $\mathsf{Rank}(C) = r$. Hence, $ABQ = A\begin{bmatrix} C & \mathbf{0} \end{bmatrix} = \begin{bmatrix} AC & \mathbf{0} \end{bmatrix}$. Thus, using Corollary 2.2.30 and Remark 2.2.27.2, we get

$$\mathsf{Rank}(AB) = \mathsf{Rank}(ABQ) = \mathsf{Rank}\left(\begin{bmatrix} AC & \mathbf{0} \end{bmatrix}\right) = \mathsf{Rank}(AC) \leq r = \mathsf{Rank}(B).$$ ∎

We end this section by relating the rank of the sum of two matrices with sum of their ranks.

Proposition 2.2.35. *Let $A, B \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, prove that* $\mathsf{Rank}(A + B) \leq \mathsf{Rank}(A) + \mathsf{Rank}(B)$. *In particular, if $A = \sum\limits_{i=1}^{k} \mathbf{x}_i \mathbf{y}_i^*$, for some $\mathbf{x}_i, \mathbf{y}_i \in \mathbb{C}$, for $1 \leq i \leq k$, then* $\mathsf{Rank}(A) \leq k$.

Proof. Let $\mathsf{Rank}(A) = r$. Then, there exists an invertible matrix P and a matrix $A_1 \in \mathbb{M}_{r,n}(\mathbb{C})$ such that $PA = \mathrm{RREF}(A) = \begin{bmatrix} A_1 \\ \mathbf{0} \end{bmatrix}$. Then,

$$P(A+B) = PA + PB = \begin{bmatrix} A_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_1 + B_1 \\ B_2 \end{bmatrix}.$$

Now using Corollary 2.2.30, Remark 2.2.27.4 and the condition $\mathsf{Rank}(A) = \mathsf{Rank}(A_1) = r$, the number of rows of A_1, we have

$$\mathsf{Rank}(A+B) = \mathsf{Rank}(P(A+B)) \leq r + \mathsf{Rank}(B_2) \leq r + \mathsf{Rank}(B) = \mathsf{Rank}(A) + \mathsf{Rank}(B).$$

Thus, the required result follows. The other part follows, as $\mathsf{Rank}(\mathbf{x}_i \mathbf{y}_i^*) = 1$, for $1 \leq i \leq k$. ∎

2.2.3 Solution set of a Linear System

Definition 2.2.37. **[Basic, Free Variables]** Consider the linear system $A\mathbf{x} = \mathbf{b}$. If $\text{RREF}([A \;\; \mathbf{b}]) = [C \;\; \mathbf{d}]$. Then, the variables corresponding to the pivotal columns of C are called the **basic** variables and the variables that are not basic are called **free** variables.

Theorem 2.2.40. *Let $A\mathbf{x} = \mathbf{b}$ be a linear system in n variables with RREF$([A \;\; \mathbf{b}]) = [C \;\; \mathbf{d}]$ with* $\mathsf{Rank}(A) = r$ *and* $\mathsf{Rank}([A\,\mathbf{b}]) = r_a$.

1. *Then, the system $A\mathbf{x} = \mathbf{b}$ is inconsistent if $r < r_a$*

2. *Then, the system $A\mathbf{x} = \mathbf{b}$ is consistent if $r = r_a$.*

   (a) *Further, $A\mathbf{x} = \mathbf{b}$ has* A UNIQUE SOLUTION *if $r = n$.*

(b) *Further,* $A\mathbf{x} = \mathbf{b}$ *has* INFINITE NUMBER OF SOLUTIONS *if* $r < n$. *In this case, there exist vectors* $\mathbf{x}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-r} \in \mathbb{R}^n$ *with* $A\mathbf{x}_0 = \mathbf{b}$ *and* $A\mathbf{u}_i = \mathbf{0}$, *for* $1 \le i \le n-r$. *Furthermore, the solution set is given by*

$$\{\mathbf{x}_0 + k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_{n-r}\mathbf{u}_{n-r} \mid k_i \in \mathbb{C},\ 1 \le i \le n-r\}.$$

Proof. PART 1: As $r < r_a$, by Remark 2.2.19.5 $([C \;\; \mathbf{d}])[r+1,:] = [\mathbf{0}^T \;\; 1]$. Note that this row corresponds to the linear equation

$$0 \cdot x_1 + 0 \cdot x_2 + \cdots + 0 \cdot x_n = 1$$

which clearly has no solution. Thus, by definition and Theorem 2.1.17, $A\mathbf{x} = \mathbf{b}$ is inconsistent.

PART 2: As $r = r_a$, by Remark 2.2.19.5, $[C \;\; \mathbf{d}]$ doesn't have a row of the form $[\mathbf{0}^T \;\; 1]$. Further, the number of pivots in $[C \;\; \mathbf{d}]$ and that in C is same, namely, r pivots. Suppose the pivots appear in columns $i_1, \ldots, i_r$ with $1 \le i_1 < \cdots < i_r \le n$. Thus, the variables x_{i_j}, for $1 \le j \le r$, are basic variables and the remaining $n-r$ variables, say $x_{t_1}, \ldots, x_{t_{n-r}}$, are free variables with $t_1 < \cdots < t_{n-r}$. Since C is in RREF, in terms of the free variables and basic variables, the ℓ-th row of $[C \;\; \mathbf{d}]$, for $1 \le \ell \le r$, corresponds to the equation

$$x_{i_\ell} + \sum_{k=1}^{n-r} c_{\ell t_k} x_{t_k} = d_\ell \Leftrightarrow x_{i_\ell} = d_\ell - \sum_{k=1}^{n-r} c_{\ell t_k} x_{t_k}.$$

Thus, the system $C\mathbf{x} = \mathbf{d}$ is consistent. Hence, by Theorem 2.1.17 the system $A\mathbf{x} = \mathbf{b}$ is consistent and the solution set of the system $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ are the same. Therefore, the solution set of the system $C\mathbf{x} = \mathbf{d}$ (or equivalently $A\mathbf{x} = \mathbf{b}$) is given by

$$\begin{bmatrix} x_{i_1} \\ \vdots \\ x_{i_r} \\ x_{t_1} \\ x_{t_2} \\ \vdots \\ x_{t_{n-r}} \end{bmatrix} = \begin{bmatrix} d_1 - \sum\limits_{k=1}^{n-r} c_{1t_k} x_{t_k} \\ \vdots \\ d_r - \sum\limits_{k=1}^{n-r} c_{rt_k} x_{t_k} \\ x_{t_1} \\ x_{t_2} \\ \vdots \\ x_{t_{n-r}} \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_{t_1} \begin{bmatrix} c_{1t_1} \\ \vdots \\ c_{rt_1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_{t_2} \begin{bmatrix} c_{1t_2} \\ \vdots \\ c_{rt_2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_{t_{n-r}} \begin{bmatrix} c_{1t_{n-r}} \\ \vdots \\ c_{rt_{n-r}} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{2.2.7}$$

PART 2A: As $r = n$, there are no free variables. Hence, $x_i = d_i$, for $1 \le i \le n$, is the unique solution.

PART 2B: Define $\mathbf{x}_0 = \begin{bmatrix} d_1 \\ \vdots \\ d_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ and $\mathbf{u}_1 = \begin{bmatrix} c_{1t_1} \\ \vdots \\ c_{rt_1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{u}_{n-r} = \begin{bmatrix} c_{1t_{n-r}} \\ \vdots \\ c_{rt_{n-r}} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$. Then, it can be easily verified that $A\mathbf{x}_0 = \mathbf{b}$ and, for $1 \le i \le n-r$, $A\mathbf{u}_i = \mathbf{0}$. Also, by Equation (2.2.7) the solution set has indeed the required form, where k_i corresponds to the free variable x_{t_i}. As there is at least one free variable the system has infinite number of solutions. Thus, the proof of the theorem is complete. ∎

Corollary 2.2.42. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. If $\mathsf{Rank}(A) = r < \min\{m, n\}$ then $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions. In particular, if $m < n$, then $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions. Hence, in either case, the homogeneous system $A\mathbf{x} = \mathbf{0}$ has at least one non-trivial solution.*

Remark 2.2.43. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, Theorem 2.2.40 implies that $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathsf{Rank}(A) = \mathsf{Rank}([A\ \mathbf{b}])$. Further, the vectors associated to the free variables in Equation (2.2.7) are solutions to the associated homogeneous system $A\mathbf{x} = \mathbf{0}$.*

We end this subsection with some applications.

2.3 Square Matrices and Linear Systems

In this section the coefficient matrix of the linear system $A\mathbf{x} = \mathbf{b}$ will be a square matrix. We start with proving a few equivalent conditions that relate different ideas.

Theorem 2.3.1. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, the following statements are equivalent.*

1. *A is invertible.*
2. *$RREF(A) = I_n$.*
3. *A is a product of elementary matrices.*
4. *The homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
5. *$\mathsf{Rank}(A) = n$.*

Proof. $1 \Leftrightarrow 2$ Already done in Proposition 2.2.21.

$2 \Leftrightarrow 3$ Again, done in Proposition 2.2.21.

$3 \Longrightarrow 4$ Let $A = E_1 \cdots E_k$, for some elementary matrices $E_1, \ldots, E_k$. Then, by previous equivalence A is invertible. So, A^{-1} exists and $A^{-1}A = I_n$. Hence, if $\mathbf{x}_0$ is any solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$ then,

$$\mathbf{x}_0 = I_n \cdot \mathbf{x}_0 = (A^{-1}A)\mathbf{x}_0 = A^{-1}(A\mathbf{x}_0) = A^{-1}\,\mathbf{0} = \mathbf{0}.$$

Thus, $\mathbf{0}$ is the only solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$.

$4 \Longrightarrow 5$ Let if possible $\mathsf{Rank}(A) = r < n$. Then, by Corollary 2.2.42, the homogeneous system $A\mathbf{x} = \mathbf{0}$ has infinitely many solution. A contradiction. Thus, A has full rank.

$5 \Longrightarrow 2$ Suppose $\mathsf{Rank}(A) = n$. So, RREF(A) has n pivotal columns. But, RREF(A) has exactly n columns and hence each column is a pivotal column. Thus, RREF$(A) = I_n$. ■

We end this section by giving two more equivalent conditions for a matrix to be invertible.

Theorem 2.3.2. *The following statements are equivalent for $A \in \mathbb{M}_n(\mathbb{C})$.*

1. *A is invertible.*
2. *The system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.*
3. *The system $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$.*

Proof. $1 \Longrightarrow 2$ Note that $\mathbf{x}_0 = A^{-1}\mathbf{b}$ is the unique solution of $A\mathbf{x} = \mathbf{b}$.

$2 \Longrightarrow 3$ The system is consistent as $A\mathbf{x} = \mathbf{b}$ has a solution.

$3 \Longrightarrow 1$ For $1 \le i \le n$, define $\mathbf{e}_i^T = I_n[i, :]$. By assumption, the linear system $A\mathbf{x} = \mathbf{e}_i$ has a solution, say $\mathbf{x}_i$, for $1 \le i \le n$. Define a matrix $B = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$. Then,

$$AB = A[\mathbf{x}_1, \mathbf{x}_2 \ldots, \mathbf{x}_n] = [A\mathbf{x}_1, A\mathbf{x}_2 \ldots, A\mathbf{x}_n] = [\mathbf{e}_1, \mathbf{e}_2 \ldots, \mathbf{e}_n] = I_n.$$

Therefore, $n = \mathsf{Rank}(I_n) = \mathsf{Rank}(AB) \leq \mathsf{Rank}(A)$ and hence $\mathsf{Rank}(A) = n$. Thus, by Theorem 2.3.1, A is invertible. ∎

We now give an immediate application of Theorem 2.3.2 and Theorem 2.3.1 without proof.

Theorem 2.3.3. *The following two statements cannot hold together for $A \in \mathbb{M}_n(\mathbb{C})$.*

1. *The system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.*

2. *The system $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution.*

As an immediate consequence of Theorem 2.3.1, the readers should prove that one needs to compute either the left or the right inverse to prove invertibility of $A \in \mathbb{M}_n(\mathbb{C})$.

Corollary 2.3.4. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, the following holds.*

1. *Suppose there exists C such that $CA = I_n$. Then, A^{-1} exists.*

2. *Suppose there exists B such that $AB = I_n$. Then, A^{-1} exists.*

2.3.1 Determinant

In this section, we associate a number with each square matrix. To start with, recall the notations used in Section 1.3.1. Then, for $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \\ 2 & 4 & 7 \end{bmatrix}$, $A(1 \mid 2) = \begin{bmatrix} 1 & 2 \\ 2 & 7 \end{bmatrix}$ and $A(\{1,2\} \mid \{1,3\}) = [4]$.

With the notations as above, we are ready to give an inductive definition of the determinant of a square matrix. The advanced students can find an alternate definition of the determinant in Appendix 9.2.22, where it is proved that the definition given below corresponds to the expansion of determinant along the first row.

Definition 2.3.6. Let A be a square matrix of order n. Then, the determinant of A, denoted $\det(A)$ (or $\mid A \mid$) is defined by

$$\det(A) = \begin{cases} a, & \text{if } A = [a] \text{ (corresponds to } n = 1), \\ \sum\limits_{j=1}^{n} (-1)^{1+j} a_{1j} \det\big(A(1 \mid j)\big), & \text{otherwise.} \end{cases}$$

Definition 2.3.9. **[Singular, Non-Singular Matrices]** A matrix A is said to be a SINGULAR if $\det(A) = 0$ and is called NON-SINGULAR if $\det(A) \neq 0$.

The next result relates the determinant with row operations. For proof, see Appendix 9.3.

Theorem 2.3.10. *Let A be an $n \times n$ matrix.*

1. *If* $B = E_{ij}A$, *for* $1 \leq i \neq j \leq n$, *then* $\det(B) = -\det(A)$.

2. *If* $B = E_i(c)A$, *for* $c \neq 0, 1 \leq i \leq n$, *then* $\det(B) = c\det(A)$.

3. *If* $B = E_{ij}(c)A$, *for* $c \neq 0$ *and* $1 \leq i \neq j \leq n$, *then* $\det(B) = \det(A)$.

4. *If* $A[i,:]^T = \mathbf{0}$, *for* $1 \leq i, j \leq n$ *then* $\det(A) = 0$.

5. *If* $A[i,:] = A[j,:]$ *for* $1 \leq i \neq j \leq n$ *then* $\det(A) = 0$.

6. *If* A *is a triangular matrix with* $d_1, \ldots, d_n$ *on the diagonal then* $\det(A) = d_1 \cdots d_n$.

As $\det(I_n) = 1$, we have the following result.

Corollary 2.3.11. *Fix a positive integer* n*.*

1. *Then,* $\det(E_{ij}) = -1$.

2. *If* $c \neq 0$ *then,* $\det(E_k(c)) = c$.

3. *If* $c \neq 0$ *then,* $\det(E_{ij}(c)) = 1$.

Remark 2.3.14. *Theorem 2.3.10.1 implies that the determinant can be calculated by expanding along any row. Hence, the readers are advised to verify that*

$$\det(A) = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det(A(k \mid j)), \quad \textit{for } 1 \leq k \leq n.$$

2.3.2 Adjugate (classical Adjoint) of a Matrix

Definition 2.3.16. Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, the **cofactor** matrix, denoted $\mathsf{Cof}(A)$, is an $\mathbb{M}_n(\mathbb{C})$ matrix with $\mathsf{Cof}(A) = [C_{ij}]$, where

$$C_{ij} = (-1)^{i+j} \det\left(A(i \mid j)\right), \text{ for } 1 \leq i, j \leq n.$$

And, the **Adjugate** (classical Adjoint) of A, denoted $Adj(A)$, equals $\mathsf{Cof}^T(A)$.

The next result relates adjugate matrix with the inverse, in case $\det(A) \neq 0$.

Theorem 2.3.18. *Let* $A \in \mathbb{M}_n(\mathbb{C})$.

1. *Then,* $\sum\limits_{j=1}^{n} a_{ij}\, C_{ij} = \sum\limits_{j=1}^{n} a_{ij}(-1)^{i+j}\, \det(A(i|j)) = \det(A)$, *for* $1 \leq i \leq n$.
2. *Then,* $\sum\limits_{j=1}^{n} a_{ij}\, C_{\ell j} = \sum\limits_{j=1}^{n} a_{ij}(-1)^{i+j}\, \det(A(\ell|j)) = 0$, *for* $i \neq \ell$.
3. *Thus,* $A(Adj(A)) = \det(A) I_n$. *Hence,*

$$\text{whenever } \det(A) \neq 0 \ \text{ one has } \ A^{-1} = \frac{1}{\det(A)} Adj(A). \tag{2.3.1}$$

Proof. Part 1: It follows directly from Remark 2.3.14 and the definition of the cofactor.

Part 2: Fix positive integers i, ℓ with $1 \leq i \neq \ell \leq n$ and let $B = [b_{ij}]$ be a square matrix with $B[\ell,:] = A[i,:]$ and $B[t,:] = A[t,:]$, for $t \neq \ell$. As $\ell \neq i$, $B[\ell,:] = B[i,:]$ and thus, by Theorem 2.3.10.5, $\det(B) = 0$. As $A(\ell \mid j) = B(\ell \mid j)$, for $1 \leq j \leq n$, using Remark 2.3.14

$$\begin{aligned} 0 = \det(B) &= \sum_{j=1}^{n} (-1)^{\ell+j} b_{\ell j} \det\big(B(\ell \mid j)\big) = \sum_{j=1}^{n} (-1)^{\ell+j} a_{ij} \det\big(B(\ell \mid j)\big) \\ &= \sum_{j=1}^{n} (-1)^{\ell+j} a_{ij} \det\big(A(\ell \mid j)\big) = \sum_{j=1}^{n} a_{ij} C_{\ell j}. \end{aligned} \tag{2.3.2}$$

This completes the proof of Part 2.

Part 3: Using Equation (2.3.2) and Remark 2.3.14, observe that

$$\Big[A\big(\mathrm{Adj}(A)\big)\Big]_{ij} = \sum_{k=1}^{n} a_{ik} \big(\mathrm{Adj}(A)\big)_{kj} = \sum_{k=1}^{n} a_{ik} C_{jk} = \begin{cases} 0, & \text{if } i \neq j, \\ \det(A), & \text{if } i = j. \end{cases}$$

Thus, $A(Adj(A)) = \det(A) I_n$. Therefore, if $\det(A) \neq 0$ then $A\left(\frac{1}{\det(A)} Adj(A)\right) = I_n$. Hence, by Proposition 2.2.21, $A^{-1} = \frac{1}{\det(A)} Adj(A)$. ∎

Corollary 2.3.20. *Let A be a non-singular matrix. Then,*

$$\sum_{i=1}^{n} C_{ik}\, a_{ij} = \begin{cases} \det(A), & \text{if } j = k, \\ 0, & \text{if } j \neq k. \end{cases}$$

The next result gives another equivalent condition for a square matrix to be invertible.

Theorem 2.3.21. *A square matrix A is non-singular if and only if A is invertible.*

Proof. Let A be non-singular. Then, $\det(A) \neq 0$ and hence $A^{-1} = \frac{1}{\det(A)} Adj(A)$.

Now, let us assume that A is invertible. Then, using Theorem 2.3.1, $A = E_1 \cdots E_k$, a product of elementary matrices. Also, by Corollary 2.3.11, $\det(E_i) \neq 0$, for $1 \leq i \leq k$. Thus, a repeated application of Parts 1, 2 and 3 of Theorem 2.3.10 gives $\det(A) \neq 0$. ∎

The next result relates the determinant of a matrix with the determinant of its transpose. Thus, the determinant can be computed by expanding along any column as well.

Theorem 2.3.22. *Let A be a square matrix. Then,* $\det(A) = \det(A^T)$.

Proof. If A is a non-singular, Corollary 2.3.20 gives $\det(A) = \det(A^T)$.

If A is singular then, by Theorem 2.3.21, A is not invertible. So, A^T is also not invertible and hence by Theorem 2.3.21, $\det(A^T) = 0 = \det(A)$. ∎

The next result relates the determinant of product of two matrices with their determinants.

Theorem 2.3.23. *Let A and B be square matrices of order n. Then,*

$$\det(AB) = \det(A) \cdot \det(B) = \det(BA).$$

Proof. Case 1: Let A be non-singular. Then, by Theorem 2.3.18.3, A is invertible and by Theorem 2.3.1, $A = E_1 \cdots E_k$, a product of elementary matrices. Thus, a repeated application of Parts 1, 2 and 3 of Theorem 2.3.10 gives the desired result as

$$\begin{aligned} \det(AB) &= \det(E_1 \cdots E_k B) = \det(E_1)\det(E_2 \cdots E_k B) = \det(E_1)\det(E_2)\det(E_3 \cdots E_k B) \\ &= \cdots = \det(E_1) \cdots \det(E_k)\det(B) = \cdots = \det(E_1 E_2 \cdots E_k)\det(B) \\ &= \det(A)\det(B). \end{aligned}$$

Case 2: Let A be singular. Then, by Theorem 2.3.21 A is not invertible. So, by Proposition 2.2.21 there exists an invertible matrix P such that $PA = \begin{bmatrix} C_1 \\ \mathbf{0} \end{bmatrix}$. So, $A = P^{-1}\begin{bmatrix} C_1 \\ \mathbf{0} \end{bmatrix}$. As P is invertible, using Part 1, we have

$$\begin{aligned} \det(AB) &= \det\left(\left(P^{-1}\begin{bmatrix} C_1 \\ \mathbf{0} \end{bmatrix}\right)B\right) = \det\left(P^{-1}\begin{bmatrix} C_1 B \\ \mathbf{0} \end{bmatrix}\right) = \det(P^{-1}) \cdot \det\left(\begin{bmatrix} C_1 B \\ \mathbf{0} \end{bmatrix}\right) \\ &= \det(P) \cdot 0 = 0 = 0 \cdot \det(B) = \det(A)\det(B). \end{aligned}$$

Thus, the proof of the theorem is complete. ∎

2.3.3 Cramer's Rule

Let A be a square matrix. Then, combining Theorem 2.3.2 and Theorem 2.3.21, one has the following result.

Corollary 2.3.26. *Let A be a square matrix. Then, the following statements are equivalent:*

1. *A is invertible.*
2. *The linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.*
3. $\det(A) \neq 0$.

Thus, $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ if and only if $\det(A) \neq 0$. The next theorem gives a direct method of finding the solution of the linear system $A\mathbf{x} = \mathbf{b}$ when $\det(A) \neq 0$.

Theorem 2.3.27 (Cramer's Rule)**.** *Let A be an $n \times n$ non-singular matrix. Then, the unique solution of the linear system $A\mathbf{x} = \mathbf{b}$ with $\mathbf{x}^T = [x_1, \ldots, x_n]$ is given by*

$$x_j = \frac{\det(A_j)}{\det(A)}, \quad \textit{for } j = 1, 2, \ldots, n,$$

where A_j is the matrix obtained from A by replacing $A[:, j]$ by $\mathbf{b}$.

Proof. Since $\det(A) \neq 0$, A is invertible. Thus, there exists an invertible matrix P such that $PA = I_n$ and $P[A \mid \mathbf{b}] = [I \mid P\mathbf{b}]$. Then $A^{-1} = P$. Let $\mathbf{d} = P\mathbf{b} = A^{-1}\mathbf{b}$. Then, $A\mathbf{x} = \mathbf{b}$ has the unique solution $x_j = \mathbf{d}_j$, for $1 \leq j \leq n$. Also, $[\mathbf{e}_1, \ldots, \mathbf{e}_n] = I = PA = [PA[:,1], \ldots, PA[:,n]]$. Thus,

$$\begin{aligned} PA_j &= P[A[:,1], \ldots, A[:,j-1], \mathbf{b}, A[:,j+1], \ldots, A[:,n]] \\ &= [PA[:,1], \ldots, PA[:,j-1], P\mathbf{b}, PA[:,j+1], \ldots, PA[:,n]] \\ &= [\mathbf{e}_1, \ldots, \mathbf{e}_{j-1}, \mathbf{d}, \mathbf{e}_{j+1}, \ldots, \mathbf{e}_n]. \end{aligned}$$

Thus, $\det(PA_j) = \mathbf{d}_j$, for $1 \leq j \leq n$. Also, $\mathbf{d}_j = \dfrac{\mathbf{d}_j}{1} = \dfrac{\det(PA_j)}{\det(PA)} = \dfrac{\det(P)\det(A_j)}{\det(P)\det(A)} = \dfrac{\det(A_j)}{\det(A)}$. Hence, $x_j = \dfrac{\det(A_j)}{\det(A)}$ and the required result follows. ■

2.5 Summary

In this chapter, we started with a system of m linear equations in n variables and formally wrote it as $A\mathbf{x} = \mathbf{b}$ and in turn to the augmented matrix $[A \mid \mathbf{b}]$. Then, the basic operations on equations led to multiplication by elementary matrices on the right of $[A \mid \mathbf{b}]$. These elementary matrices are invertible and applying the GJE on a matrix A, resulted in getting the RREF of A. We used the pivots in RREF matrix to define the rank of a matrix. So, if $\mathsf{Rank}(A) = r$ and $\mathsf{Rank}([A \mid \mathbf{b}]) = r_a$

1. then, $r < r_a$ implied the linear system $A\mathbf{x} = \mathbf{b}$ is inconsistent.

2. then, $r = r_a$ implied the linear system $A\mathbf{x} = \mathbf{b}$ is consistent. Further,

   (a) if $r = n$ then the system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

   (b) if $r < n$ then the system $A\mathbf{x} = \mathbf{b}$ has an infinite number of solutions.

We have also seen that the following conditions are equivalent for $A \in \mathbb{M}_n(\mathbb{C})$.

1. A is invertible.

2. The homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

3. The row reduced echelon form of A is I.

4. A is a product of elementary matrices.

5. The system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.

6. The system $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$.

7. $\mathsf{Rank}(A) = n$.

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

So, overall we have learnt to solve the following type of problems:

1. Solving the linear system $A\mathbf{x} = \mathbf{b}$. This idea will lead to the question "is the vector $\mathbf{b}$ a linear combination of the columns of A"?

2. Solving the linear system $A\mathbf{x} = \mathbf{0}$. This will lead to the question "are the columns of A linearly independent/dependent"? In particular, we will see that

   (a) if $A\mathbf{x} = \mathbf{0}$ has a unique solution then the columns of A are linear independent.

   (b) if $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution then the columns of A are linearly dependent.

Chapter 3

Vector Spaces

3.1 Vector Spaces: Definition and Examples

Let $A \in \mathbb{M}_{m,n}(\mathbb{F})$ and let $\mathbb{V}$ denote the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$. Then, by Theorem 2.1.9, $\mathbb{V}$ satisfies:

1. $\mathbf{0} \in \mathbb{V}$ as $A\mathbf{0} = \mathbf{0}$.

2. if $\mathbf{x} \in \mathbb{V}$ then $\alpha\mathbf{x} \in \mathbb{V}$, for all $\alpha \in \mathbb{F}$. In particular, for $\alpha = -1$, $-\mathbf{x} \in \mathbb{V}$.

3. if $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ then, for any $\alpha, \beta \in \mathbb{F}$, $\alpha\mathbf{x} + \beta\mathbf{y} \in \mathbb{V}$.

That is, the solution set of a homogeneous linear system satisfies some nice properties. The Euclidean plane, $\mathbb{R}^2$, and the Euclidean space, $\mathbb{R}^3$, also satisfy the above properties. In this chapter, our aim is to understand sets that satisfy such properties. We start with the following definition.

Definition 3.1.1. **[Vector Space]** A **vector space** $\mathbb{V}$ over $\mathbb{F}$, denoted $\mathbb{V}(\mathbb{F})$ or in short $\mathbb{V}$ (if the field $\mathbb{F}$ is clear from the context), is a non-empty set, satisfying the following conditions:

1. **Vector Addition:** To every pair $\mathbf{u}, \mathbf{v} \in \mathbb{V}$ there corresponds a unique element $\mathbf{u} \oplus \mathbf{v} \in \mathbb{V}$ (called the **addition of vectors**) such that
   (a) $\mathbf{u} \oplus \mathbf{v} = \mathbf{v} \oplus \mathbf{u}$ (Commutative law).
   (b) $(\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = \mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w})$ (Associative law).
   (c) $\mathbb{V}$ has a unique element, denoted $\mathbf{0}$, called **the zero vector** that satisfies $\mathbf{u} \oplus \mathbf{0} = \mathbf{u}$, for every $\mathbf{u} \in \mathbb{V}$ (called **the additive identity**).
   (d) for every $\mathbf{u} \in \mathbb{V}$ there is an element $\mathbf{w} \in \mathbb{V}$ that satisfies $\mathbf{u} \oplus \mathbf{w} = \mathbf{0}$.

2. **Scalar Multiplication:** For each $\mathbf{u} \in \mathbb{V}$ and $\alpha \in \mathbb{F}$, there corresponds a unique element $\alpha \odot \mathbf{u}$ in $\mathbb{V}$ (called the **scalar multiplication**) such that
   (a) $\alpha \cdot (\beta \odot \mathbf{u}) = (\alpha\beta) \odot \mathbf{u}$ for every $\alpha, \beta \in \mathbb{F}$ and $\mathbf{u} \in \mathbb{V}$ ($\cdot$ is multiplication in $\mathbb{F}$).
   (b) $1 \odot \mathbf{u} = \mathbf{u}$ for every $\mathbf{u} \in \mathbb{V}$, where $1 \in \mathbb{F}$.

3. **Distributive Laws: relating vector addition with scalar multiplication**
   For any $\alpha, \beta \in \mathbb{F}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{V}$, the following **distributive laws** hold:
   (a) $\alpha \odot (\mathbf{u} \oplus \mathbf{v}) = (\alpha \odot \mathbf{u}) \;\oplus\; (\alpha \odot \mathbf{v})$.
   (b) $(\alpha + \beta) \odot \mathbf{u} = (\alpha \odot \mathbf{u}) \;\oplus\; (\beta \odot \mathbf{u})$ ($+$ is addition in $\mathbb{F}$).

Remark 3.1.2. **[Real / Complex Vector Space]**

1. *The elements of* $\mathbb{F}$ *are called* **scalars**.
2. *The elements of* $\mathbb{V}$ *are called* **vectors**.
3. *We denote the zero element of* $\mathbb{F}$ *by* 0*, whereas the zero element of* $\mathbb{V}$ *will be denoted by* $\mathbf{0}$.
4. *Observe that Condition 3.1.1.1d implies that for every* $\mathbf{u} \in \mathbb{V}$*, the vector* $\mathbf{w} \in \mathbb{V}$ *such that* $\mathbf{u} + \mathbf{w} = \mathbf{0}$ *holds, is unique. For if,* $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{V}$ *with* $\mathbf{u} + \mathbf{w}_i = \mathbf{0}$*, for* $i = 1, 2$ *then by commutativity of vector addition, we see that*

$$\mathbf{w}_1 = \mathbf{w}_1 + \mathbf{0} = \mathbf{w}_1 + (\mathbf{u} + \mathbf{w}_2) = (\mathbf{w}_1 + \mathbf{u}) + \mathbf{w}_2 = \mathbf{0} + \mathbf{w}_2 = \mathbf{w}_2.$$

   Hence, we represent this unique vector by $-\mathbf{u}$ *and call it* **the additive inverse**.
5. *If* $\mathbb{V}$ *is a vector space over* $\mathbb{R}$ *then,* $\mathbb{V}$ *is called a* **real vector space**.
6. *If* $\mathbb{V}$ *is a vector space over* $\mathbb{C}$ *then* $\mathbb{V}$ *is called a* **complex vector space**.
7. *In general, a vector space over* $\mathbb{R}$ *or* $\mathbb{C}$ *is called a* **linear space**.

Some interesting consequences of Definition 3.1.1 is stated next. Intuitively, they seem obvious but for better understanding of the given conditions, it is desirable to go through the proof.

Theorem 3.1.3. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then,*

1. $\mathbf{u} \oplus \mathbf{v} = \mathbf{u}$ *implies* $\mathbf{v} = \mathbf{0}$.

2. $\alpha \odot \mathbf{u} = \mathbf{0}$ *if and only if either* $\mathbf{u} = \mathbf{0}$ *or* $\alpha = 0$.

3. $(-1) \odot \mathbf{u} = -\mathbf{u}$, *for every* $\mathbf{u} \in \mathbb{V}$.

Proof. Part 1: By Condition 3.1.1.1d, for each $\mathbf{u} \in \mathbb{V}$ there exists $-\mathbf{u} \in \mathbb{V}$ such that $-\mathbf{u} \oplus \mathbf{u} = \mathbf{0}$. Hence, $\mathbf{u} \oplus \mathbf{v} = \mathbf{u}$ is equivalent to

$$-\mathbf{u} \oplus (\mathbf{u} \oplus \mathbf{v}) = -\mathbf{u} \oplus \mathbf{u} \Longleftrightarrow (-\mathbf{u} \oplus \mathbf{u}) \oplus \mathbf{v} = \mathbf{0} \Longleftrightarrow \mathbf{0} \oplus \mathbf{v} = \mathbf{0} \Longleftrightarrow \mathbf{v} = \mathbf{0}.$$

Part 2: As $\mathbf{0} = \mathbf{0} \oplus \mathbf{0}$, using Condition 3.1.1.3, we have

$$\alpha \odot \mathbf{0} = \alpha \odot (\mathbf{0} \oplus \mathbf{0}) = (\alpha \odot \mathbf{0}) \ \oplus \ (\alpha \odot \mathbf{0}).$$

Thus, using Part 1, $\alpha \odot \mathbf{0} = \mathbf{0}$ for any $\alpha \in \mathbb{F}$. In the same way, using Condition 3.1.1.3b,

$$0 \odot \mathbf{u} = (0 + 0) \odot \mathbf{u} = (0 \odot \mathbf{u}) \ \oplus (0 \odot \mathbf{u}).$$

Hence, using Part 1, one has $0 \odot \mathbf{u} = \mathbf{0}$ for any $\mathbf{u} \in \mathbb{V}$.

Now suppose $\alpha \odot \mathbf{u} = \mathbf{0}$. If $\alpha = 0$ then the proof is over. Therefore, assume that $\alpha \neq 0, \alpha \in \mathbb{F}$. Then, $(\alpha)^{-1} \in \mathbb{F}$ and

$$\mathbf{0} = (\alpha)^{-1} \odot \mathbf{0} = (\alpha)^{-1} \odot (\alpha \odot \mathbf{u}) = ((\alpha)^{-1} \cdot \alpha) \odot \mathbf{u} = 1 \odot \mathbf{u} = \mathbf{u}$$

as $1 \odot \mathbf{u} = \mathbf{u}$ for every vector $\mathbf{u} \in \mathbb{V}$ (see Condition 2.2b). Thus, if $\alpha \neq 0$ and $\alpha \odot \mathbf{u} = \mathbf{0}$ then $\mathbf{u} = \mathbf{0}$.

Part 3: As $\mathbf{0} = 0 \cdot \mathbf{u} = (1 + (-1))\mathbf{u} = \mathbf{u} \oplus (-1) \cdot \mathbf{u}$, one has $(-1) \cdot \mathbf{u} = -\mathbf{u}$. ∎

3.1.1 Subspaces

Definition 3.1.7. [**Vector Subspace**] Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then, a non-empty subset S of $\mathbb{V}$ is called a **subspace** of $\mathbb{V}$ if S is also a vector space with vector addition and scalar multiplication inherited from $\mathbb{V}$.

Theorem 3.1.9. *Let* $\mathbb{V}(\mathbb{F})$ *be a vector space and* $\mathbb{W} \subseteq \mathbb{V}, \mathbb{W} \neq \emptyset$*. Then,* $\mathbb{W}$ *is a subspace of* $\mathbb{V}$ *if and only if* $\alpha\mathbf{u} + \beta\mathbf{v} \in \mathbb{W}$ *whenever* $\alpha, \beta \in \mathbb{F}$ *and* $\mathbf{u}, \mathbf{v} \in \mathbb{W}$*.*

Proof. Let $\mathbb{W}$ be a subspace of $\mathbb{V}$ and let $\mathbf{u}, \mathbf{v} \in \mathbb{W}$. Then, for every $\alpha, \beta \in \mathbb{F}$, $\alpha\mathbf{u}, \beta\mathbf{v} \in \mathbb{W}$ and hence $\alpha\mathbf{u} + \beta\mathbf{v} \in \mathbb{W}$.

Now, we assume that $\alpha\mathbf{u} + \beta\mathbf{v} \in \mathbb{W}$, whenever $\alpha, \beta \in \mathbb{F}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{W}$. To show, $\mathbb{W}$ is a subspace of $\mathbb{V}$:

1. Taking $\alpha = 1$ and $\beta = 1$, we see that $\mathbf{u} + \mathbf{v} \in \mathbb{W}$, for every $\mathbf{u}, \mathbf{v} \in \mathbb{W}$.

2. Taking $\alpha = 0$ and $\beta = 0$, we see that $\mathbf{0} \in \mathbb{W}$.

3. Taking $\beta = 0$, we see that $\alpha\mathbf{u} \in \mathbb{W}$, for every $\alpha \in \mathbb{F}$ and $\mathbf{u} \in \mathbb{W}$. Hence, using Theorem 3.1.3.3, $-\mathbf{u} = (-1)\mathbf{u} \in \mathbb{W}$ as well.

4. The commutative and associative laws of vector addition hold as they hold in $\mathbb{V}$.

5. The conditions related with scalar multiplication and the distributive laws also hold as they hold in $\mathbb{V}$.

Thus, one obtains the required result. ∎

3.1.2 Linear Span

Definition 3.1.11. **[Linear Combination]** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then, for any $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{V}$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$, the vector $\alpha_1\mathbf{u}_1 + \cdots + \alpha_n\mathbf{u}_n = \sum\limits_{i=1}^{n} \alpha_i\mathbf{u}_i$ is said to be a **linear combination** of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$.

Definition 3.1.14. **[Linear Span]** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and $S \subseteq \mathbb{V}$. Then, the **linear span** of S, denoted $LS(S)$, is defined as

$$LS(S) \;=\; \{\alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n \mid \alpha_i \in \mathbb{F}, \mathbf{u}_i \in S, \text{ for } 1 \le i \le n\}.$$

That is, $LS(S)$ is the set of all possible linear combinations of finitely many vectors of S. If S is an empty set, we define $LS(S) = \{\mathbf{0}\}$.

Definition 3.1.17. **[Finite Dimensional Vector Space]** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then, $\mathbb{V}$ is called **finite dimensional** if there exists $S \subseteq \mathbb{V}$, such that S has finite number of elements and $\mathbb{V} = LS(S)$. If such an S does not exist then $\mathbb{V}$ is called **infinite dimensional**.

Lemma 3.1.19 (Linear Span is a Subspace)**.** *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and $S \subseteq \mathbb{V}$. Then, $LS(S)$ is a subspace of $\mathbb{V}$.*

Proof. By definition, $\mathbf{0} \in LS(S)$. So, $LS(S)$ is non-empty. Let $\mathbf{u}, \mathbf{v} \in LS(S)$. To show, $a\mathbf{u} + b\mathbf{v} \in LS(S)$ for all $a, b \in \mathbb{F}$. As $\mathbf{u}, \mathbf{v} \in LS(S)$, there exist $n \in \mathbb{N}$, vectors $\mathbf{w}_i \in S$ and scalars $\alpha_i, \beta_i \in \mathbb{F}$ such that $\mathbf{u} = \alpha_1\mathbf{w}_1 + \cdots + \alpha_n\mathbf{w}_n$ and $\mathbf{v} = \beta_1\mathbf{w}_1 + \cdots + \beta_n\mathbf{w}_n$. Hence,

$$a\mathbf{u} + b\mathbf{v} = (a\alpha_1 + b\beta_1)\mathbf{w}_1 + \cdots + (a\alpha_n + b\beta_n)\mathbf{w}_n \in LS(S)$$

as $a\alpha_i + b\beta_i \in \mathbb{F}$ for $1 \le i \le n$. Thus, by Theorem 3.1.9, $LS(S)$ is a vector subspace. ∎

EXERCISE **3.1.20.** *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and $W \subseteq \mathbb{V}$.*

1. *Then, $LS(W) = W$ if and only if W is a subspace of $\mathbb{V}$.*
2. *If W is a subspace of $\mathbb{V}$ and $S \subseteq W$ then $LS(S)$ is a subspace of W as well.*

Theorem 3.1.21. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and $S \subseteq \mathbb{V}$. Then, $LS(S)$ is the smallest subspace of $\mathbb{V}$ containing S.*

Proof. For every $\mathbf{u} \in S$, $\mathbf{u} = 1 \cdot \mathbf{u} \in LS(S)$. Thus, $S \subseteq LS(S)$. Need to show that $LS(S)$ is the smallest subspace of $\mathbb{V}$ containing S. So, let $\mathbb{W}$ be any subspace of $\mathbb{V}$ containing S. Then, by Exercise 3.1.20, $LS(S) \subseteq \mathbb{W}$ and hence the result follows. ∎

Definition 3.1.22. **[Sum of two subsets]** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$.

1. Let S and T be two subsets of $\mathbb{V}$. Then, the **sum** of S and T, denoted $S + T$ equals $\{s + t | s \in S, t \in T\}$. For example,
   (a) if $\mathbb{V} = \mathbb{R}$, $S = \{0, 1, 2, 3, 4, 5, 6\}$ and $T = \{5, 10, 15\}$ then $S + T = \{5, 6, \ldots, 21\}$.
   (b) if $\mathbb{V} = \mathbb{R}^2$, $S = \left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$ and $T = \left\{\begin{bmatrix}-1\\1\end{bmatrix}\right\}$ then $S + T = \left\{\begin{bmatrix}0\\2\end{bmatrix}\right\}$.
   (c) if $\mathbb{V} = \mathbb{R}^2$, $S = \left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$ and $T = LS\left(\begin{bmatrix}-1\\1\end{bmatrix}\right)$ then $S + T = \left\{\begin{bmatrix}1\\1\end{bmatrix} + c\begin{bmatrix}-1\\1\end{bmatrix} | c \in \mathbb{R}\right\}$.
2. Let P and Q be two subspaces of $\mathbb{R}^2$. Then, $P + Q = \mathbb{R}^2$, if
   (a) $P = \{(x, 0)^T \mid x \in \mathbb{R}\}$ and $Q = \{(0, x)^T \mid x \in \mathbb{R}\}$ as $(x, y) = (x, 0) + (0, y)$.
   (b) $P = \{(x, 0)^T \mid x \in \mathbb{R}\}$ and $Q = \{(x, x)^T \mid x \in \mathbb{R}\}$ as $(x, y) = (x - y, 0) + (y, y)$.
   (c) $P = LS((1, 2)^T)$ and $Q = LS((2, 1)^T)$ as $(x, y) = \dfrac{2y - x}{3}(1, 2) + \dfrac{2x - y}{3}(2, 1)$.

We leave the proof of the next result for readers.

Lemma 3.1.23. *Let P and Q be two subspaces of a vector space $\mathbb{V}$ over $\mathbb{F}$. Then, $P + Q$ is a subspace of $\mathbb{V}$. Furthermore, $P + Q$ is the smallest subspace of $\mathbb{V}$ containing both P and Q.*

3.2 Fundamental Subspaces Associated with a Matrix

Definition 3.2.1. **[Fundamental Subspaces]** Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, we define the four fundamental subspaces associated with A as

1. $\text{Col}(A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{C}^n\} \subseteq \mathbb{C}^m$, called the **Column space**. Observe that $\text{Col}(A)$ is the linear span of the columns of A.
2. $\text{Row}(A) = \{\mathbf{x}^T A \mid \mathbf{x} \in \mathbb{C}^m\}$, called the **row space** of A. Observe that $\text{Row}(A)$ is the linear span of the rows of A.
3. $\text{Null}(A) = \{\mathbf{x} \in \mathbb{C}^n \mid A\mathbf{x} = \mathbf{0}\}$, called the **Null space** of A.
4. $\text{Null}(A^*) = \{\mathbf{x} \in \mathbb{C}^m \mid A^*\mathbf{x} = \mathbf{0}\}$.

Remark 3.2.2. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$.*

1. *Then, $\text{Col}(A)$ is a subspace of $\mathbb{C}^m$ and $\text{Col}(A^*)$ is a subspace of $\mathbb{C}^n$.*
2. *Then, $\text{Null}(A)$ is a subspace of $\mathbb{C}^n$ and $\text{Null}(A^*)$ is a subspace of $\mathbb{C}^m$.*

Remark 3.2.4. *Let $A \in \mathbb{M}_{m,n}(\mathbb{R})$. Then, in Example 3.2.3, observe that the direction ratios of normal vectors of $\text{Col}(A)$ matches with vector in $\text{Null}(A^T)$. Similarly, the direction ratios of normal vectors of $\text{Row}(A)$ matches with vectors in $\text{Null}(A)$. Are these true in the general setting? Do similar relations hold if $A \in \mathbb{M}_{m,n}(\mathbb{C})$? We will come back to these spaces again and again.*

3.3 Linear Independence

Definition 3.3.1. [Linear Independence and Dependence] Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be a non-empty subset of a vector space $\mathbb{V}$ over $\mathbb{F}$. Then, S is said to be **linearly independent** if the linear system

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_m\mathbf{u}_m = \mathbf{0}, \tag{3.3.1}$$

in the variables α_i's, $1 \le i \le m$, has only the trivial solution. If Equation (3.3.1) has a non-trivial solution then S is said to be **linearly dependent**.

If S has infinitely many vectors then S is said to be **linearly independent** if for every finite subset T of S, T is linearly independent.

Observe that we are solving a linear system over $\mathbb{F}$. Hence, linear independence and dependence depend on $\mathbb{F}$, the set of scalars.

3.3.1 Basic Results on Linear Independence

The reader is expected to supply the proof of parts that are not given.

Proposition 3.3.3. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$.*

1. *Then, $\mathbf{0}$, the zero-vector, cannot belong to a linearly independent set.*
2. *Then, every subset of a linearly independent set in $\mathbb{V}$ is also linearly independent.*
3. *Then, a set containing a linearly dependent set of $\mathbb{V}$ is also linearly dependent.*

Proof. Let $\mathbf{0} \in S$. Then, $1 \cdot \mathbf{0} = \mathbf{0}$. That is, a non-trivial linear combination of some vectors in S is $\mathbf{0}$. Thus, the set S is linearly dependent. ■

We now prove a couple of results which will be very useful in the next section.

Proposition 3.3.4. *Let S be a linearly independent subset of a vector space $\mathbb{V}$ over $\mathbb{F}$. If T_1, T_2 are two subsets of S such that $T_1 \cap T_2 = \emptyset$ then, $LS(T_1) \cap LS(T_2) = \{\mathbf{0}\}$. That is, if $\mathbf{v} \in LS(T_1) \cap LS(T_2)$ then $\mathbf{v} = \mathbf{0}$.*

Proof. Let $\mathbf{v} \in LS(T_1) \cap LS(T_2)$. Then, there exist vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in T_1$, $\mathbf{w}_1, \ldots, \mathbf{w}_\ell \in T_2$ and scalars α_i's and β_j's (not all zero) such that $\mathbf{v} = \sum\limits_{i=1}^{k} \alpha_i \mathbf{u}_i$ and $\mathbf{v} = \sum\limits_{j=1}^{\ell} \beta_j \mathbf{w}_j$. Thus, we see that $\sum\limits_{i=1}^{k} \alpha_i \mathbf{u}_i + \sum\limits_{j=1}^{\ell} (-\beta_j) \mathbf{w}_j = \mathbf{0}$. As the scalars α_i's and β_j's are not all zero, we see that a non-trivial linear combination of some vectors in $T_1 \cup T_2 \subseteq S$ is $\mathbf{0}$. This contradicts the assumption that S is a linearly independent subset of $\mathbb{V}$. Hence, each of α's and β_j's is zero. That is $\mathbf{v} = \mathbf{0}$. ■

We now prove another useful result.

Theorem 3.3.5. *Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be a non-empty subset of a vector space $\mathbb{V}$ over $\mathbb{F}$. If $T \subseteq LS(S)$ having more than k vectors then, T is a linearly dependent subset in $\mathbb{V}$.*

Proof. Let $T = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$. As $\mathbf{w}_i \in LS(S)$, there exist $a_{ij} \in \mathbb{F}$ such that

$$\mathbf{w}_i = a_{i1}\mathbf{u}_1 + \cdots + a_{ik}\mathbf{u}_k, \text{ for } 1 \le i \le m.$$

So,

$$\begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{bmatrix} = \begin{bmatrix} a_{11}\mathbf{u}_1 + \cdots + a_{1k}\mathbf{u}_k \\ \vdots \\ a_{m1}\mathbf{u}_1 + \cdots + a_{mk}\mathbf{u}_k \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mk} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{bmatrix}.$$

As $m > k$, using Corollary 2.2.42, the linear system $\mathbf{x}^T A = \mathbf{0}^T$ has a non-trivial solution, say $\mathbf{Y} \neq \mathbf{0}^T$. That is, $\mathbf{Y}^T A = \mathbf{0}^T$. Thus,

$$\mathbf{Y}^T \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{bmatrix} = \mathbf{Y}^T \left(A \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{bmatrix} \right) = (\mathbf{Y}^T A) \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{bmatrix} = \mathbf{0}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{bmatrix} = \mathbf{0}^T.$$

As $\mathbf{Y} \neq \mathbf{0}$, a non-trivial linear combination of vectors in T is $\mathbf{0}$. Thus, the set T is linearly dependent subset of $\mathbb{V}$. ∎

Corollary 3.3.6. *Fix $n \in \mathbb{N}$. Then, any subset S of $\mathbb{R}^n$ with $|\, S \,| \ge n+1$ is linearly dependent.*

Proof. Observe that $\mathbb{R}^n = LS(\{\mathbf{e}_1, \ldots, \mathbf{e}_n\})$, where $\mathbf{e}_i = I_n[:, i]$, is the i-th column of I_n. Hence, using Theorem 3.3.5, the required result follows. ∎

Theorem 3.3.7. *Let S be a linearly independent subset of a vector space $\mathbb{V}$ over $\mathbb{F}$. Then, for any $\mathbf{v} \in \mathbb{V}$ the set $S \cup \{\mathbf{v}\}$ is linearly dependent if and only if $\mathbf{v} \in LS(S)$.*

Proof. Let us assume that $S \cup \{\mathbf{v}\}$ is linearly dependent. Then, there exist $\mathbf{v}_i$'s in S such that the linear system

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_p\mathbf{v}_p + \alpha_{p+1}\mathbf{v} = \mathbf{0} \tag{3.3.3}$$

in the variables α_i's has a non-trivial solution, say $\alpha_i = c_i$, for $1 \le i \le p+1$. We claim that $c_{p+1} \neq 0$.

For, if $c_{p+1} = 0$ then, Equation (3.3.3) has a non-trivial solution corresponds to having a non-trivial solution of the linear system $\alpha_1\mathbf{v}_1 + \cdots + \alpha_p\mathbf{v}_p = \mathbf{0}$ in the variables $\alpha_1, \ldots, \alpha_p$. This contradicts Proposition 3.3.3.2 as $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\} \subseteq S$, a linearly independent set. Thus, $c_{p+1} \neq 0$ and we get

$$\mathbf{v} = -\frac{1}{c_{p+1}}(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) \in LS(\mathbf{v}_1, \ldots, \mathbf{v}_p)$$

as $-\frac{c_i}{c_{p+1}} \in \mathbb{F}$, for $1 \le i \le p$. That is, $\mathbf{v}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

Now, assume that $\mathbf{v} \in LS(S)$. Then, there exists $\mathbf{v}_i \in S$ and $c_i \in \mathbb{F}$, not all zero, such that $\mathbf{v} = \sum\limits_{i=1}^{p} c_i\mathbf{v}_i$. Thus, the linear system $\alpha_1\mathbf{v}_1 + \cdots + \alpha_p\mathbf{v}_p + \alpha_{p+1}\mathbf{v} = \mathbf{0}$ in the variables α_i's has a non-trivial solution $[c_1, \ldots, c_p, -1]$. Hence, $S \cup \{\mathbf{v}\}$ is linearly dependent. ∎

We now state a very important corollary of Theorem 3.3.7 without proof. This result can also be used as an alternative definition of linear independence and dependence.

Corollary 3.3.8. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and let S be a subset of $\mathbb{V}$ containing a non-zero vector $\mathbf{u}_1$.*

1. *If S is linearly dependent then, there exists k such that $LS(\mathbf{u}_1, \ldots, \mathbf{u}_k) = LS(\mathbf{u}_1, \ldots, \mathbf{u}_{k-1})$.*

2. *If S linearly independent then, $\mathbf{v} \in \mathbb{V} \setminus LS(S)$ if and only if $S \cup \{\mathbf{v}\}$ is also a linearly independent subset of $\mathbb{V}$.*

3. *If S is linearly independent then, $LS(S) = \mathbb{V}$ if and only if each proper superset of S is linearly dependent.*

3.3.2 Application to Matrices

We start with our understanding of the RREF.

Theorem 3.3.9. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, the rows of A corresponding to the pivotal rows of RREF(A) are linearly independent. Also, the columns of A corresponding to the pivotal columns of RREF(A) are linearly independent.*

Proof. Let RREF$(A) = B$. Then, the pivotal rows of B are linearly independent due to the pivotal 1's. Now, let B_1 be the submatrix of B consisting of the pivotal rows of B. Also, let A_1 be the submatrix of A whose rows corresponds to the rows of B_1. As the RREF of a matrix is unique (see Corollary 2.2.18) there exists an invertible matrix Q such that $QA_1 = B_1$. So, if there exists $\mathbf{c} \neq \mathbf{0}$ such that $\mathbf{c}^T A_1 = \mathbf{0}^T$ then

$$\mathbf{0}^T = \mathbf{c}^T A_1 = \mathbf{c}^T(Q^{-1}B_1) = (\mathbf{c}^T Q^{-1})B_1 = \mathbf{d}^T B_1,$$

with $\mathbf{d}^T = \mathbf{c}^T Q^{-1} \neq \mathbf{0}^T$ as Q is an invertible matrix (see Theorem 2.3.1). This contradicts the linear independence of the rows of B_1.

Let $B[:, i_1], \ldots, B[:, i_r]$ be the pivotal columns of B. Then, they are linearly independent due to pivotal 1's. As $B = $ RREF(A), there exists an invertible matrix P such that $B = PA$. Then, the corresponding columns of A satisfy

$$[A[:, i_1], \ldots, A[:, i_r]] = [P^{-1}B[:, i_1], \ldots, P^{-1}B[:, i_r]] = P^{-1}[B[:, i_1], \ldots, B[:, i_r]].$$

As P is invertible, the systems $[A[:, i_1], \ldots, A[:, i_r]]\begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \mathbf{0}$ and $[B[:, i_1], \ldots, B[:, i_r]]\begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \mathbf{0}$ are row-equivalent. Thus, they have the same solution set. Hence, $\{A[:, i_1], \ldots, A[:, i_r]\}$ is linearly independent if and only if $\{B[:, i_1], \ldots, B[:, i_r]\}$ is linear independent. Thus, the required result follows. ∎

The next result follows directly from Theorem 3.3.9 and hence the proof is left to readers.

Corollary 3.3.10. *The following statements are equivalent for $A \in \mathbb{M}_n(\mathbb{C})$.*

1. *A is invertible.*

2. *The columns of A are linearly independent.*

3. *The rows of A are linearly independent.*

We give an example for better understanding.

3.3.3 Linear Independence and Uniqueness of Linear Combination

We end this section with a result that states that linear combination with respect to linearly independent set is unique.

Lemma 3.3.12. *Let S be a linearly independent subset of a vector space $\mathbb{V}$ over $\mathbb{F}$. Then, each $\mathbf{v} \in LS(S)$ is a unique linear combination of vectors from S.*

Proof. Suppose there exists $\mathbf{v} \in LS(S)$ with $\mathbf{v} \in LS(T_1), LS(T_2)$ with $T_1, T_2 \subseteq S$. Let $T_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ and $T_2 = \{\mathbf{w}_1, \ldots, \mathbf{w}_\ell\}$, for some $\mathbf{v}_i$'s and $\mathbf{w}_j$'s in S. Define $T = T_1 \cup T_2$. Then, T is a subset of S. Hence, using Proposition 3.3.3, the set T is linearly independent. Let $T = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$. Then, there exist α_i's and β_j's in $\mathbb{F}$, not all zero, such that $\mathbf{v} = \alpha_1\mathbf{u}_1 + \cdots + \alpha_p\mathbf{u}_p$ as well as $\mathbf{v} = \beta_1\mathbf{u}_1 + \cdots + \beta_p\mathbf{u}_p$. Equating the two expressions for $\mathbf{v}$ gives

$$(\alpha_1 - \beta_1)\mathbf{u}_1 + \cdots + (\alpha_p - \beta_p)\mathbf{u}_p = \mathbf{0}. \tag{3.3.4}$$

As T is a linearly independent subset of $\mathbb{V}$, the system $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$, in the variables $c_1, \ldots, c_p$, has only the trivial solution. Thus, in Equation (3.3.4), $\alpha_i - \beta_i = 0$, for $1 \leq i \leq p$. Thus, for $1 \leq i \leq p$, $\alpha_i = \beta_i$ and the required result follows. ■

3.4 Basis of a Vector Space

Definition 3.4.1. **[Maximality]** Let S be a subset of a set T. Then, S is said to be a **maximal subset** of T having property P if

1. S has property P and
2. no proper superset of S in T has property P.

Definition 3.4.3. **[Maximal linearly independent set]** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then, S is called a **maximal linearly independent** subset of $\mathbb{V}$ if

1. S is linearly independent and
2. no proper superset of S in $\mathbb{V}$ is linearly independent.

Theorem 3.4.5. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and S a linearly independent set in $\mathbb{V}$. Then, S is maximal linearly independent if and only if $LS(S) = \mathbb{V}$.*

Proof. Let $\mathbf{v} \in \mathbb{V}$. As S is linearly independent, using Corollary 3.3.8.2, the set $S \cup \{\mathbf{v}\}$ is linearly independent if and only if $\mathbf{v} \in \mathbb{V} \setminus LS(S)$. Thus, the required result follows. ∎

Let $\mathbb{V} = LS(S)$ for some set S with $|S| = k$. Then, using Theorem 3.3.5, we see that if $T \subseteq \mathbb{V}$ is linearly independent then $|T| \leq k$. Hence, a maximal linearly independent subset of $\mathbb{V}$ can have at most k vectors. Thus, we arrive at the following important result.

Theorem 3.4.6. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and let S and T be two finite maximal linearly independent subsets of $\mathbb{V}$. Then, $|S| = |T|$.*

Proof. By Theorem 3.4.5, S and T are maximal linearly independent if and only if $LS(S) = \mathbb{V} = LS(T)$. Now, use the previous paragraph to get the required result. ∎

Let $\mathbb{V}$ be a finite dimensional vector space. Then, by Theorem 3.4.6, the number of vectors in any two maximal linearly independent set is the same. We use this number to define the dimension of a vector space. We do so now.

Definition 3.4.7. **[Dimension of a finite dimensional vector space]** Let $\mathbb{V}$ be a finite dimensional vector space over $\mathbb{F}$. Then, the number of vectors in any maximal linearly independent set is called the **dimension** of $\mathbb{V}$, denoted $\dim(\mathbb{V})$. By convention, $\dim(\{\mathbf{0}\}) = 0$.

Definition 3.4.9. Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then, a maximal linearly independent subset of $\mathbb{V}$ is called a **basis/Hamel basis** of $\mathbb{V}$. The vectors in a basis are called **basis** vectors. By convention, a basis of $\{\mathbf{0}\}$ is the empty set.

Existence of Hamel basis

Definition 3.4.10. **[Minimal Spanning Set]** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then, a subset S of $\mathbb{V}$ is called **minimal spanning** if $LS(S) = \mathbb{V}$ and no proper subset of S spans $\mathbb{V}$.

Remark 3.4.11 (Standard Basis)**.** *The readers should verify the statements given below.*

1. *All the maximal linearly independent set given in Example 3.4.8 form the standard basis of the respective vector space.*

2. *$\{1, x, x^2, \ldots\}$ is the standard basis of $\mathbb{R}[x]$ over $\mathbb{R}$.*

3. *Fix a positive integer n. Then, $\{1, x, x^2, \ldots, x^n\}$ is the standard basis of $\mathbb{R}[x; n]$ over $\mathbb{R}$.*

4. *Let $\mathbb{V} = \{A \in \mathbb{M}_n(\mathbb{R}) \mid A = A^T\}$. Then, $\mathbb{V}$ is a vector space over $\mathbb{R}$ with standard basis $\{\mathcal{E}_{ii}, \mathcal{E}_{ij} + \mathcal{E}_{ji} \mid 1 \leq i < j \leq n\}$.*

5. *Let* $\mathbb{V} = \{A \in \mathbb{M}_n(\mathbb{R}) \mid A^T = -A\}$. *Then,* $\mathbb{V}$ *is a vector space over* $\mathbb{R}$ *with standard basis* $\{\mathcal{E}_{ij} - \mathcal{E}_{ji} \mid 1 \leq i < j \leq n\}$.

3.4.1 Main Results associated with Bases

Theorem 3.4.13. *Let* $\mathbb{V}$ *be a non-zero vector space over* $\mathbb{F}$. *Then, the following statements are equivalent.*

1. $\mathcal{B}$ *is a basis (maximal linearly independent subset) of* $\mathbb{V}$.
2. $\mathcal{B}$ *is linearly independent and spans* $\mathbb{V}$.
3. $\mathcal{B}$ *is a minimal spanning set in* $\mathbb{V}$.

Proof. $1 \Rightarrow 2$ By definition, every basis is a maximal linearly independent subset of $\mathbb{V}$. Thus, using Corollary 3.3.8.2, we see that $\mathcal{B}$ spans $\mathbb{V}$.

$2 \Rightarrow 3$ Let S be a linearly independent set that spans $\mathbb{V}$. As S is linearly independent, for any $\mathbf{x} \in S$, $\mathbf{x} \notin LS\left(S - \{\mathbf{x}\}\right)$. Hence $LS\left(S - \{\mathbf{x}\}\right) \subsetneq LS(S) = \mathbb{V}$.

$3 \Rightarrow 1$ If $\mathcal{B}$ is linearly dependent then using Corollary 3.3.8.1, $\mathcal{B}$ is not minimal spanning. A contradiction. Hence, $\mathcal{B}$ is linearly independent.

We now need to show that $\mathcal{B}$ is a maximal linearly independent set. Since $LS(\mathcal{B}) = \mathbb{V}$, for any $\mathbf{x} \in \mathbb{V} \setminus \mathcal{B}$, using Corollary 3.3.8.2, the set $\mathcal{B} \cup \{\mathbf{x}\}$ is linearly dependent. That is, every proper superset of $\mathcal{B}$ is linearly dependent. Hence, the required result follows. ■

Now, using Lemma 3.3.12, we get the following result.

Remark 3.4.14. *Let $\mathcal{B}$ be a basis of a vector space $\mathbb{V}$ over $\mathbb{F}$. Then, for each $\mathbf{v} \in \mathbb{V}$, there exist unique $\mathbf{u}_i \in \mathcal{B}$ and unique $\alpha_i \in \mathbb{F}$, for $1 \leq i \leq n$, such that $\mathbf{v} = \sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i$.*

The next result is generally known as "every linearly independent set can be extended to form a basis of a finite dimensional vector space".

Theorem 3.4.15. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ with $\dim(\mathbb{V}) = n$. If S is a linearly independent subset of $\mathbb{V}$ then there exists a basis T of $\mathbb{V}$ such that $S \subseteq T$.*

Proof. If $LS(S) = \mathbb{V}$, done. Else, choose $\mathbf{u}_1 \in \mathbb{V} \setminus LS(S)$. Thus, by Corollary 3.3.8.2, the set $S \cup \{\mathbf{u}_1\}$ is linearly independent. We repeat this process till we get n vectors in T as $\dim(\mathbb{V}) = n$. By Theorem 3.4.13, this T is indeed a required basis. ■

3.4.2 Constructing a Basis of a Finite Dimensional Vector Space

We end this section with an algorithm which is based on the proof of the previous theorem.

Step 1: Let $\mathbf{v}_1 \in \mathbb{V}$ with $\mathbf{v}_1 \neq \mathbf{0}$. Then, $\{\mathbf{v}_1\}$ is linearly independent.

Step 2: If $\mathbb{V} = LS(\mathbf{v}_1)$, we have got a basis of $\mathbb{V}$. Else, pick $\mathbf{v}_2 \in \mathbb{V} \setminus LS(\mathbf{v}_1)$. Then, by Corollary 3.3.8.2, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent.

Step i: Either $\mathbb{V} = LS(\mathbf{v}_1, \ldots, \mathbf{v}_i)$ or $LS(\mathbf{v}_1, \ldots, \mathbf{v}_i) \subsetneq \mathbb{V}$. In the first case, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ is a basis of $\mathbb{V}$. Else, pick $\mathbf{v}_{i+1} \in \mathbb{V} \setminus LS(\mathbf{v}_1, \ldots, \mathbf{v}_i)$. Then, by Corollary 3.3.8.2, the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}\}$ is linearly independent.

This process will finally end as $\mathbb{V}$ is a finite dimensional vector space.

3.5 Application to the subspaces of $\mathbb{C}^n$

In this section, we will study results that are intrinsic to the understanding of linear algebra from the point of view of matrices, especially the fundamental subspaces (see Definition 3.2.1) associated with matrices. We start with an example.

Lemma 3.5.3. *Let $A \in M_{m\times n}(\mathbb{C})$ and let E be an elementary matrix. If*

1. *$B = EA$ then*
   (a) $\textsc{Null}(A) = \textsc{Null}(B)$, $\textsc{Row}(A) = \textsc{Row}(B)$. *Thus, the dimensions of the corresponding spaces are equal.*
   (b) $\textsc{Null}(\overline{A}) = \textsc{Null}(\overline{B})$, $\textsc{Row}(\overline{A}) = \textsc{Row}(\overline{B})$. *Thus, the dimensions of the corresponding spaces are equal.*

2. *$B = AE$ then*
   (a) $\textsc{Null}(A^*) = \textsc{Null}(B^*)$, $\textsc{Col}(\overline{A}) = \textsc{Col}(\overline{B})$. *Thus, the dimensions of the corresponding spaces are equal.*
   (b) $\textsc{Null}(A^T) = \textsc{Null}(B^T)$, $\textsc{Col}(A) = \textsc{Col}(B)$. *Thus, the dimensions of the corresponding spaces are equal.*

Proof. Part 1a: Let $\mathbf{x} \in \textsc{Null}(A)$. Then, $B\mathbf{x} = EA\mathbf{x} = E\mathbf{0} = \mathbf{0}$. So, $\textsc{Null}(A) \subseteq \textsc{Null}(B)$. Further, if $\mathbf{x} \in \textsc{Null}(B)$, then $A\mathbf{x} = (E^{-1}E)A\mathbf{x} = E^{-1}(EA)\mathbf{x} = E^{-1}B\mathbf{x} = E^{-1}\mathbf{0} = \mathbf{0}$. Hence, $\textsc{Null}(B) \subseteq \textsc{Null}(A)$. Thus, $\textsc{Null}(A) = \textsc{Null}(B)$.

Let us now prove $\textsc{Row}(A) = \textsc{Row}(B)$. So, let $\mathbf{x}^T \in \textsc{Row}(A)$. Then, there exists $\mathbf{y} \in \mathbb{C}^m$ such that $\mathbf{x}^T = \mathbf{y}^T A$. Thus, $\mathbf{x}^T = \left(\mathbf{y}^T E^{-1}\right) EA = \left(\mathbf{y}^T E^{-1}\right) B$ and hence $\mathbf{x}^G \in \textsc{Row}(B)$. That is, $\textsc{Row}(A) \subseteq \textsc{Row}(B)$. A similar argument gives $\textsc{Row}(B) \subseteq \textsc{Row}(A)$ and hence the required result follows.

Part 1b: E is invertible implies $\overline{E}$ is invertible and $\overline{B} = \overline{EA}$. Thus, an argument similar to the previous part gives us the required result.

For Part 2, note that $B^* = E^*A^*$ and E^* is invertible. Hence, an argument similar to the first part gives the required result. ■

Let $A \in M_{m\times n}(\mathbb{C})$ and let $B = \text{RREF}(A)$. Then, as an immediate application of Lemma 3.5.3, we get $\dim(\text{Row}(A)) = \text{Row rank}(A)$. We now prove that $\dim(\text{Row}(A)) = \dim(\text{Col}(A))$.

Theorem 3.5.4. *Let* $A \in M_{m\times n}(\mathbb{C})$. *Then,* $\dim(\text{Row}(A)) = \dim(\text{Col}(A))$.

Proof. Let $\dim(\text{Row}(A)) = r$. Then, there exist $i_1, \ldots, i_r$ such that $\{A[i_1,:], \ldots, A[i_r,:]\}$ forms a basis of $\text{Row}(A)$. Then, $B = \begin{bmatrix} A[i_1,:] \\ \vdots \\ A[i_r,:] \end{bmatrix}$ is an $r\times n$ matrix and it's rows are a basis of $\text{Row}(A)$. Therefore, there exist $\alpha_{ij} \in \mathbb{C}$, $1 \le i \le m$, $1 \le j \le r$ such that $A[t,:] = [\alpha_{t1}, \ldots, \alpha_{tr}]B$, for $1 \le t \le m$. So, using matrix multiplication

$$A = \begin{bmatrix} A[1,:] \\ \vdots \\ A[m,:] \end{bmatrix} = \begin{bmatrix} [\alpha_{11}, \ldots, \alpha_{1r}]B \\ \vdots \\ [\alpha_{m1}, \ldots, \alpha_{mr}]B \end{bmatrix} = CB = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1r} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mr} \end{bmatrix} B,$$

where $C = [\alpha_{ij}]$ is an $m \times r$ matrix. Thus, using matrix multiplication, we see that each column of A is a linear combination of r columns of C. Hence, $\dim(\text{Col}(A)) \le r = \dim(\text{Row}(A))$. A similar argument gives $\dim(\text{Row}(A)) \le \dim(\text{Col}(A))$. Hence, we have the required result. ■

Remark 3.5.5. *The proof also shows that for every* $A \in M_{m\times n}(\mathbb{C})$ *of rank* r *there exists matrices* $B_{r\times n}$ *and* $C_{m\times r}$, *each of rank* r, *such that* $A = CB$.

Let $\mathbb{W}_1$ and $\mathbb{W}_1$ be two subspaces of a vector space $\mathbb{V}$ over $\mathbb{F}$. Then, recall that (see Exercise 3.1.24.4d) $\mathbb{W}_1 + \mathbb{W}_2 = \{\mathbf{u} + \mathbf{v} \mid \mathbf{u} \in \mathbb{W}_1, \ \mathbf{v} \in \mathbb{W}_2\} = LS(\mathbb{W}_1 \cup \mathbb{W}_2)$ is the smallest subspace of $\mathbb{V}$ containing both $\mathbb{W}_1$ and $\mathbb{W}_2$. We now state a result similar to a result in Venn diagram that states $\mid A \mid + \mid B \mid = \mid A \cup B \mid + \mid A \cap B \mid$, whenever the sets A and B are finite (for a proof, see Appendix 9.4.1).

Theorem 3.5.6. *Let* V *be a finite dimensional vector space over* $\mathbb{F}$. *If* $\mathbb{W}_1$ *and* $\mathbb{W}_2$ *are two subspaces of* V *then*

$$\dim(\mathbb{W}_1) + \dim(\mathbb{W}_2) = \dim(\mathbb{W}_1 + \mathbb{W}_2) + \dim(\mathbb{W}_1 \cap \mathbb{W}_2). \tag{3.5.2}$$

For better understanding, we give an example for finite subsets of $\mathbb{R}^n$. The example uses Theorem 3.3.9 to obtain bases of $LS(S)$, for different choices S. The readers are advised to see Example 3.3.9 before proceeding further.

Theorem 3.5.9 (Rank-Nullity Theorem)**.** *Let $A \in M_{m \times n}(\mathbb{C})$. Then,*

$$\dim(\textsc{Col}(A)) + \dim(\textsc{Null}(A)) = n. \tag{3.5.4}$$

Proof. Let $\dim(\textsc{Null}(A)) = r \leq n$ and let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ be a basis of $\textsc{Null}(A)$. Since $\mathcal{B}$ is a linearly independent set in $\mathbb{R}^n$, extend it to get $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ as a basis of $\mathbb{R}^n$. Then,

$$\begin{aligned} \textsc{Col}(A) &= LS(\mathcal{B}) = LS(A\mathbf{u}_1, \ldots, A\mathbf{u}_n) \\ &= LS(\mathbf{0}, \ldots, \mathbf{0}, A\mathbf{u}_{r+1}, \ldots, A\mathbf{u}_n) = LS(A\mathbf{u}_{r+1}, \ldots, A\mathbf{u}_n). \end{aligned}$$

So, $\mathcal{C} = \{A\mathbf{u}_{r+1}, \ldots, A\mathbf{u}_n\}$ spans $\text{Col}(A)$. We further need to show that $\mathcal{C}$ is linearly independent. So, consider the linear system

$$\alpha_1 A\mathbf{u}_{r+1} + \cdots + \alpha_{n-r} A\mathbf{u}_n = \mathbf{0} \Leftrightarrow A(\alpha_1 \mathbf{u}_{r+1} + \cdots + \alpha_{n-r}\mathbf{u}_n) = \mathbf{0} \tag{3.5.5}$$

in the variables $\alpha_1, \ldots, \alpha_{n-r}$. Thus, $\alpha_1 \mathbf{u}_{r+1} + \cdots + \alpha_{n-r}\mathbf{u}_n \in \text{Null}(A) = LS(\mathcal{B})$. Therefore, there exist scalars β_i, $1 \le i \le r$, such that $\sum\limits_{i=1}^{n-r} \alpha_i \mathbf{u}_{r+i} = \sum\limits_{j=1}^{r} \beta_j \mathbf{u}_j$. Or equivalently,

$$\beta_1 \mathbf{u}_1 + \cdots + \beta_r \mathbf{u}_r - \alpha_1 \mathbf{u}_{r+1} - \cdots - \alpha_{n-r}\mathbf{u}_n = \mathbf{0}. \tag{3.5.6}$$

As $\mathcal{B}$ is a linearly independent set, the only solution of Equation (3.5.6) is

$$\alpha_i = 0, \text{ for } 1 \le i \le n-r \text{ and } \beta_j = 0, \text{ for } 1 \le j \le r.$$

In other words, we have shown that the only solution of Equation (3.5.5) is the trivial solution. Hence, $\{A\mathbf{u}_{r+1}, \ldots, A\mathbf{u}_n\}$ is a basis of $\text{Col}(A)$. Thus, the required result follows. ■

Theorem 3.5.9 is part of what is known as the fundamental theorem of linear algebra (see Theorem 5.2.16). The following are some of the consequences of the rank-nullity theorem. The proofs are left as an exercise for the reader.

3.6 Ordered Bases

Let $\mathbb{V}$ be a vector space over $\mathbb{C}$ with $\dim(\mathbb{V}) = n$, for some positive integer n. Also, let $\mathbb{W}$ be a subspace of $\mathbb{V}$ with $\dim(\mathbb{W}) = k$. Then, a basis of $\mathbb{W}$ may not look like a standard basis. Our problem may force us to look for some other basis. In such a case, it is always helpful to fix the vectors in a particular order and then concentrate only on the coefficients of the vectors as was done for the system of linear equations where we didn't worry about the variables. It may also happen that k is very-very small as compared to n in which case it is better to work with k vectors in place of n vectors.

Definition 3.6.1. **[Ordered Basis, Basis Matrix]** Let $\mathbb{W}$ be a vector space over $\mathbb{F}$ with a basis $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$. Then, an **ordered basis** for $\mathbb{W}$ is a basis $\mathcal{B}$ together with a one-to-one correspondence between $\mathcal{B}$ and $\{1, 2, \ldots, m\}$. Since there is an order among the elements of $\mathcal{B}$, we write $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$. The vector $B = [\mathbf{u}_1, \ldots, \mathbf{u}_m]$ is an element of $\mathbb{W}^m$ and is generally called the **basis matrix**.

Definition 3.6.3. **[Coordinate Vector]** Let $B = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$ be the basis matrix corresponding to an ordered basis $\mathcal{B}$ of $\mathbb{W}$. Since $\mathcal{B}$ is a basis of $\mathbb{W}$, for each $\mathbf{v} \in \mathbb{W}$, there exist $\beta_i, 1 \le i \le m$, such that $\mathbf{v} = \sum\limits_{i=1}^{m} \beta_i \mathbf{v}_i = B \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$. The vector $\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$, denoted $[\mathbf{v}]_{\mathcal{B}}$, is called **the coordinate vector of v with respect to** $\mathcal{B}$. Thus,

$$\mathbf{v} = B[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}_1, \ldots, \mathbf{v}_m][\mathbf{v}]_{\mathcal{B}}, \text{ or equivalently, } \mathbf{v} = [\mathbf{v}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_m \end{bmatrix}. \tag{3.6.1}$$

The last expression is generally viewed as a symbolic expression.

Remark 3.6.5. [Basis representation of v]

1. *Let $\mathcal{B}$ be an ordered basis of a vector space $\mathbb{V}$ over $\mathbb{F}$ of dimension n.*

   (a) *Then,*
   $$[\alpha\mathbf{v} + \mathbf{w}]_{\mathcal{B}} = \alpha[\mathbf{v}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}, \quad \textit{for all } \ \alpha \in \mathbb{F} \textit{ and } \mathbf{v}, \mathbf{w} \in \mathbb{V}.$$

   (b) *Further, let $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\} \subseteq \mathbb{V}$. Then, observe that S is linearly independent if and only if $\{[\mathbf{w}_1]_{\mathcal{B}}, \ldots, [\mathbf{w}_m]_{\mathcal{B}}\}$ is linearly independent in $\mathbb{F}^n$.*

2. *Suppose $V = \mathbb{F}^n$ in Definition 3.6.3. Then, $B = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ is an $n \times n$ invertible matrix (see Exercise 3.4.16.4). Thus, using Equation (3.6.1), we have*
   $$B[\mathbf{v}]_{\mathcal{B}} = \mathbf{v} = (BB^{-1})\mathbf{v} = B\left(B^{-1}\mathbf{v}\right), \ \textit{for every } \mathbf{v} \in \mathbb{V}. \tag{3.6.2}$$
   As B is invertible, $[\mathbf{v}]_{\mathcal{B}} = B^{-1}\mathbf{v}$, for every $\mathbf{v} \in \mathbb{V}$.

Definition 3.6.6. **[Change of Basis Matrix]** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ with $\dim(\mathbb{V}) = n$. Let $A = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ and $B = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ be basis matrices corresponding to the ordered bases $\mathcal{A}$ and $\mathcal{B}$, respectively, of $\mathbb{V}$. Thus, using Equation (3.6.1), we have

$$[\mathbf{v}_1, \ldots, \mathbf{v}_n] = [B[\mathbf{v}_1]_{\mathcal{B}}, \ldots, B[\mathbf{v}_n]_{\mathcal{B}}] = B\,[[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_n]_{\mathcal{B}}] = B[\mathcal{A}]_{\mathcal{B}},$$

where $[\mathcal{A}]_{\mathcal{B}} = [[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_n]_{\mathcal{B}}]$. Or equivalently, verify the symbolic equality

$$\begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} = [\mathcal{A}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix}. \tag{3.6.3}$$

The matrix $[\mathcal{A}]_{\mathcal{B}}$ is called the matrix of $\mathcal{A}$ **with respect to the ordered basis** $\mathcal{B}$ or the **change of basis matrix** from $\mathcal{A}$ to $\mathcal{B}$.

We now summarize the above discussion.

Theorem 3.6.7. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ with $\dim(\mathbb{V}) = n$. Further, let $\mathcal{A} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ be two ordered bases of $\mathbb{V}$*

1. *Then, the matrix $[\mathcal{A}]_{\mathcal{B}}$ is invertible.*
2. *Similarly, the matrix $[\mathcal{B}]_{\mathcal{A}}$ is invertible.*
3. *Moreover, $[\mathbf{x}]_{\mathcal{B}} = [\mathcal{A}]_{\mathcal{B}}[\mathbf{x}]_{\mathcal{A}}$, for all $\mathbf{x} \in \mathbb{V}$. Thus, again note that the matrix $[\mathcal{A}]_{\mathcal{B}}$ takes coordinate vector of $\mathbf{x}$ with respect to $\mathcal{A}$ to the coordinate vector of $\mathbf{x}$ with respect to $\mathcal{B}$. Hence, $[\mathcal{A}]_{\mathcal{B}}$ was called the change of basis matrix from $\mathcal{A}$ to $\mathcal{B}$.*
4. *Similarly, $[\mathbf{x}]_{\mathcal{A}} = [\mathcal{B}]_{\mathcal{A}}[\mathbf{x}]_{\mathcal{B}}$, for all $\mathbf{x} \in \mathbb{V}$.*
5. *Furthermore, $([\mathcal{A}]_{\mathcal{B}})^{-1} = [\mathcal{B}]_{\mathcal{A}}$.*

Proof. Part 1: Note that using Equation (3.6.3), we have $\begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} = [\mathcal{A}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix}$ and hence by Exercise 3.3.13.8, the matrix $[\mathcal{A}]_{\mathcal{B}}^T$ or equivalently $[\mathcal{A}]_{\mathcal{B}}$ is invertible, which proves Part 1. A similar argument gives Part 2.

Part 3: Using Equations (3.6.1) and (3.6.3). for any $\mathbf{x} \in \mathbb{V}$, we have

$$[\mathbf{x}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix} = \mathbf{x} = [\mathbf{x}]_{\mathcal{A}}^T \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} = [\mathbf{x}]_{\mathcal{A}}^T [\mathcal{A}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix}.$$

Since the basis representation of an element is unique, we get $[\mathbf{x}]_{\mathcal{B}}^T = [\mathbf{x}]_{\mathcal{A}}^T[\mathcal{A}]_{\mathcal{B}}^T$. Or equivalently, $[\mathbf{x}]_{\mathcal{B}} = [\mathcal{A}]_{\mathcal{B}}[\mathbf{x}]_{\mathcal{A}}$. This completes the proof of Part 3. We leave the proof of other parts to the reader. ∎

Remark 3.6.9. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ with $\mathcal{A} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ as an ordered basis. Then, by Theorem 3.6.7, $[\mathbf{v}]_{\mathcal{A}}$ is an element of $\mathbb{F}^n$, for each $\mathbf{v} \in \mathbb{V}$. Therefore,*

1. *if $\mathbb{F} = \mathbb{R}$ then, the elements of $\mathbb{V}$ correspond to vectors in $\mathbb{R}^n$.*
2. *if $\mathbb{F} = \mathbb{C}$ then, the elements of $\mathbb{V}$ correspond to vectors in $\mathbb{C}^n$.*

3.7 Summary

In this chapter, we defined vector spaces over $\mathbb{F}$. The set $\mathbb{F}$ was either $\mathbb{R}$ or $\mathbb{C}$. To define a vector space, we start with a non-empty set $\mathbb{V}$ of vectors and $\mathbb{F}$ the set of scalars. We also needed to do the following:

1. first define vector addition and scalar multiplication and
2. then verify the conditions in Definition 3.1.1.

If all conditions in Definition 3.1.1 are satisfied then $\mathbb{V}$ is a vector space over $\mathbb{F}$. If $\mathbb{W}$ was a non-empty subset of a vector space $\mathbb{V}$ over $\mathbb{F}$ then for $\mathbb{W}$ to be a space, we only need to check whether the vector addition and scalar multiplication inherited from that in $\mathbb{V}$ hold in $\mathbb{W}$.

We then learnt linear combination of vectors and the linear span of vectors. It was also shown that the linear span of a subset S of a vector space $\mathbb{V}$ is the smallest subspace of $\mathbb{V}$ containing S. Also, to check whether a given vector $\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_n$, we needed to solve the linear system $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{v}$ in the variables $c_1, \ldots, c_n$. Or equivalently, the system $A\mathbf{x} = \mathbf{b}$, where in some sense $A[:, i] = \mathbf{u}_i$, $1 \leq i \leq n$, $\mathbf{x}^T = [c_1, \ldots, c_n]$ and $\mathbf{b} = \mathbf{v}$. It was also shown that the geometrical representation of the linear span of $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is equivalent to finding conditions in the entries of $\mathbf{b}$ such that $A\mathbf{x} = \mathbf{b}$ was always consistent.

Then, we learnt linear independence and dependence. A set $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is linearly independent set in the vector space $\mathbb{V}$ over $\mathbb{F}$ if the homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution in $\mathbb{F}$. Else S is linearly dependent, where as before the columns of A correspond to the vectors $\mathbf{u}_i$'s.

We then talked about the maximal linearly independent set (coming from the homogeneous system) and the minimal spanning set (coming from the non-homogeneous system) and culminating in the notion of the basis of a finite dimensional vector space $\mathbb{V}$ over $\mathbb{F}$. The following important results were proved.

1. A linearly independent set can be extended to form a basis of $\mathbb{V}$.

2. Any two bases of $\mathbb{V}$ have the same number of elements.

This number was defined as the dimension of $\mathbb{V}$, denoted $\dim(\mathbb{V})$.

Now let $A \in \mathbb{M}_n(\mathbb{R})$. Then, combining a few results from the previous chapter, we have the following equivalent conditions.

1. A is invertible.

2. The homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

3. $\text{RREF}(A) = I_n$.

4. A is a product of elementary matrices.

5. The system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.

6. The system $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$.

7. $\text{Rank}(A) = n$.

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

9. $\text{Col}(A^T) = \text{Row}(A) = \mathbb{R}^n$.

10. Rows of A form a basis of $\mathbb{R}^n$.

11. $\text{Col}(A) = \mathbb{R}^n$.

12. Columns of A form a basis of $\mathbb{R}^n$.

13. $\text{Null}(A) = \{\mathbf{0}\}$.

Chapter 4

Linear Transformations

4.1 Definitions and Basic Properties

Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ with $\dim(\mathbb{V}) = n$. Also, let $\mathcal{B}$ be an ordered basis of $\mathbb{V}$. Then, in the last section of the previous chapter, it was shown that for each $\mathbf{x} \in \mathbb{V}$, the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ is a column vector of size n and has entries from $\mathbb{F}$. So, in some sense, each element of $\mathbb{V}$ looks like elements of $\mathbb{F}^n$. In this chapter, we concretize this idea. We also show that matrices give rise to functions between two finite dimensional vector spaces. To do so, we start with the definition of functions over vector spaces that commute with the operations of vector addition and scalar multiplication.

Definition 4.1.1. [**Linear Transformation, Linear Operator**] Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$. A function (map) $T : \mathbb{V} \to \mathbb{W}$ is called a **linear transformation** if for all $\alpha \in \mathbb{F}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{V}$ the function T satisfies

$$T(\alpha \cdot \mathbf{u}) = \alpha \odot T(\mathbf{u}) \ \text{ and } \ T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) \oplus T(\mathbf{v}),$$

where $+, \cdot$ are binary operations in $\mathbb{V}$ and $\oplus, \odot$ are the binary operations in $\mathbb{W}$. By $\mathcal{L}(\mathbb{V}, \mathbb{W})$, we denote the set of all linear transformations from $\mathbb{V}$ to $\mathbb{W}$. In particular, if $\mathbb{W} = \mathbb{V}$ then the linear transformation T is called a **linear operator** and the corresponding set of linear operators is denoted by $\mathcal{L}(\mathbb{V})$.

Definition 4.1.2. [**Equality of Linear Transformation**] Let $S, T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Then, S and T are said to be **equal** if $S(\mathbf{x}) = T(\mathbf{x})$, for all $\mathbf{x} \in \mathbb{V}$.

We now give examples of linear transformations.

Remark 4.1.4. *Let $A \in M_n(\mathbb{C})$ and define $T_A : \mathbb{C}^n \to \mathbb{C}^n$ by $T_A(\mathbf{x}) = A\mathbf{x}$, for every $\mathbf{x} \in \mathbb{C}^n$. Then, verify that $T_A^k(\mathbf{x}) = \underbrace{(T_A \circ T_A \circ \cdots \circ T_A)}_{k \text{ times}}(\mathbf{x}) = A^k\mathbf{x}$, for any positive integer k.*

Also, for any two linear transformations $S \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ and $T \in \mathcal{L}(\mathbb{W}, \mathbb{Z})$, we will interchangeably use $T \circ S$ and TS, for the corresponding linear transformation in $\mathcal{L}(\mathbb{V}, \mathbb{Z})$.

We now prove that any linear transformation sends the zero vector to a zero vector.

Proposition 4.1.5. *Let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Suppose that $\mathbf{0}_{\mathbb{V}}$ is the zero vector in $\mathbb{V}$ and $\mathbf{0}_{\mathbb{W}}$ is the zero vector of $\mathbb{W}$. Then $T(\mathbf{0}_{\mathbb{V}}) = \mathbf{0}_{\mathbb{W}}$.*

Proof. Since $\mathbf{0}_{\mathbb{V}} = \mathbf{0}_{\mathbb{V}} + \mathbf{0}_{\mathbb{V}}$, we get $T(\mathbf{0}_{\mathbb{V}}) = T(\mathbf{0}_{\mathbb{V}} + \mathbf{0}_{\mathbb{V}}) = T(\mathbf{0}_{\mathbb{V}}) + T(\mathbf{0}_{\mathbb{V}})$. As $T(\mathbf{0}_{\mathbb{V}}) \in \mathbb{W}$,

$$\mathbf{0}_{\mathbb{W}} + T(\mathbf{0}_{\mathbb{V}}) = T(\mathbf{0}_{\mathbb{V}}) = T(\mathbf{0}_{\mathbb{V}}) + T(\mathbf{0}_{\mathbb{V}}).$$

Hence, $T(\mathbf{0}_{\mathbb{V}}) = \mathbf{0}_{\mathbb{W}}$. ■

From now on $\mathbf{0}$ will be used as the zero vector of the domain and codomain. We now consider a few more examples.

Lemma 4.1.7. *Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Then T is known, if the image of T on basis vectors of $\mathbb{V}$ are known. In particular, if $\mathbb{V}$ is finite dimensional and $\mathcal{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is an ordered basis of $\mathbb{V}$ over $\mathbb{F}$ then, $T(\mathbf{v}) = \begin{bmatrix} T(\mathbf{v}_1) & \cdots & T(\mathbf{v}_n) \end{bmatrix} [\mathbf{v}]_{\mathcal{B}}$.*

Proof. Let $\mathcal{B}$ be a basis of $\mathbb{V}$ over $\mathbb{F}$. Then, for each $\mathbf{v} \in \mathbb{V}$, there exist vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ in $\mathcal{B}$ and scalars $c_1, \ldots, c_k \in \mathbb{F}$ such that $\mathbf{v} = \sum\limits_{i=1}^{k} c_i \mathbf{u}_i$. Thus, by definition $T(\mathbf{v}) = \sum\limits_{i=1}^{k} c_i T(\mathbf{u}_i)$. Or equivalently, whenever

$$\mathbf{v} = [\mathbf{u}_1, \ldots, \mathbf{u}_k] \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} \text{ then, } T(\mathbf{v}) = \begin{bmatrix} T(\mathbf{u}_1) & \cdots & T(\mathbf{u}_k) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}. \tag{4.1.1}$$

Thus, the image of T on $\mathbf{v}$ just depends on where the basis vectors are mapped. In particular, if $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ then, $T(\mathbf{v}) = \begin{bmatrix} T(\mathbf{u}_1) & \cdots & T(\mathbf{u}_k) \end{bmatrix} [\mathbf{v}]_{\mathcal{B}}$. Hence, the required result follows. ■

As another application of Lemma 4.1.7, we have the following result. The proof is left for the reader.

Corollary 4.1.8. *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ and let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. If $\mathcal{B}$ is a basis of $\mathbb{V}$ then,* $\text{RNG}(T) = LS(T(\mathbf{x}) | \mathbf{x} \in \mathcal{B})$.

Recall that by Example 4.1.3.6, for each $\mathbf{a} \in \mathbb{F}^n$, the map $T(\mathbf{x}) = \mathbf{a}^T\mathbf{x}$, for each $\mathbf{x} \in \mathbb{F}^n$, is a linear transformation. We now show that these are the only ones.

Corollary 4.1.9. **[Reisz Representation Theorem]** *Let $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$. Then, there exists $\mathbf{a} \in \mathbb{R}^n$ such that $T(\mathbf{x}) = \mathbf{a}^T\mathbf{x}$.*

Proof. By Lemma 4.1.7, T is known if we know the image of T on $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, the standard basis of $\mathbb{R}^n$. As T is given, for $1 \le i \le n$, $T(\mathbf{e}_i) = a_i$, for some $a_i \in \mathbb{R}$. So, consider the vector $\mathbf{a} = [a_1, \ldots, a_n]^T$. Then, for $\mathbf{x} = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$, we see that

$$T(\mathbf{x}) = T\left(\sum_{i=1}^{n} x_i \mathbf{e}_i\right) = \sum_{i=1}^{n} x_i T(\mathbf{e}_i) = \sum_{i=1}^{n} x_i a_i = \mathbf{a}^T\mathbf{x}.$$

Thus, the required result follows. ■

Before proceeding further, we define two spaces related with a linear transformation.

Definition 4.1.10. **[Range and Kernel of a Linear Transformation]** Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ and let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then,

1. the set $\{T(\mathbf{v}) | \mathbf{v} \in \mathbb{V}\}$ is called the **range space** of T, denoted $\text{RNG}(T)$.
2. the set $\{\mathbf{v} \in \mathbb{V} | T(\mathbf{v}) = \mathbf{0}\}$ is called the **kernel** of T, denoted $\text{KER}(T)$. In certain books, it is also called the **null space** of T.

4.2 Rank-Nullity Theorem

The readers are advised to see Exercise 3.3.13.9 and Theorem 3.5.9 for clarity and similarity with the results in this section. We start with the following result.

Theorem 4.2.1. *Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$.*

1. *If $S \subseteq \mathbb{V}$ is linearly dependent then $T(S) = \{T(\mathbf{v}) \mid \mathbf{v} \in \mathbb{V}\}$ is linearly dependent.*
2. *Suppose $S \subseteq \mathbb{V}$ such that $T(S)$ is linearly independent then S is linearly independent.*

Proof. As S is linearly dependent, there exist $k \in \mathbb{N}$ and $\mathbf{v}_i \in S$, for $1 \le i \le k$, such that the system $\sum\limits_{i=1}^{k} x_i\mathbf{v}_i = \mathbf{0}$, in the variable x_i's, has a non-trivial solution, say $x_i = a_i \in \mathbb{F}, 1 \le i \le k$. Thus, $\sum\limits_{i=1}^{k} a_i\mathbf{v}_i = \mathbf{0}$. Now, consider the system $\sum\limits_{i=1}^{k} y_iT(\mathbf{v}_i) = \mathbf{0}$, in the variable y_i's. Then,

$$\sum_{i=1}^{k} a_iT(\mathbf{v}_i) = \sum_{i=1}^{k} T(a_i\mathbf{v}_i) = T\left(\sum_{i=1}^{k} a_i\mathbf{v}_i\right) = T(\mathbf{0}) = \mathbf{0}.$$

Thus, a_i's give a non-trivial solution of $\sum\limits_{i=1}^{k} y_iT(\mathbf{v}_i) = \mathbf{0}$ and hence the required result follows.

The second part is left as an exercise for the reader. ∎

Definition 4.2.2. [Rank and Nullity] Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$. If $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ and $\dim(\mathbb{V})$ is finite then we define $\textsc{Rank}(T) = \dim(\textsc{Rng}(T))$ and $\textsc{Nullity}(T) = \dim(\textsc{Ker}(T))$.

We now prove the rank-nullity Theorem. The proof of this result is similar to the proof of Theorem 3.5.9. We give it again for the sake of completeness.

Theorem 4.2.3 (Rank-Nullity Theorem)**.** *Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$. If $\dim(\mathbb{V})$ is finite and $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ then,*

$$\textsc{Rank}(T) + \textsc{Nullity}(T) = \dim(\textsc{Rng}(T)) + \dim(\textsc{Ker}(T)) = \dim(\mathbb{V}).$$

Proof. By Exercise 4.1.12.2.2a, $\dim(\textsc{Ker}(T)) \le \dim(\mathbb{V})$. Let $\mathcal{B}$ be a basis of $\textsc{Ker}(T)$. We extend it to form a basis $\mathcal{C}$ of $\mathbb{V}$. As, $T(\mathbf{v}) = \mathbf{0}$, for all $\mathbf{v} \in \mathcal{B}$, using Corollary 4.1.8, we get

$$\textsc{Rng}(T) = LS(\{T(\mathbf{v})|\mathbf{v} \in \mathcal{C}\}) = LS(\{T(\mathbf{v})|\mathbf{v} \in \mathcal{C} \setminus \mathcal{B}\}).$$

We claim that $\{T(\mathbf{v})|\mathbf{v} \in \mathcal{C} \setminus \mathcal{B}\$$ is linearly independent subset of $\mathbb{W}$.

Let, if possible, the claim be false. Then, there exists $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathcal{C} \setminus \mathcal{B}$ and $\mathbf{a} = [a_1, \ldots, a_k]^T$ such that $\mathbf{a} \ne \mathbf{0}$ and $\sum\limits_{i=1}^{k} a_iT(\mathbf{v}_i) = \mathbf{0}$. Thus, we see that

$$T\left(\sum_{i=1}^{k} a_i\mathbf{v}_i\right) = \sum_{i=1}^{k} a_iT(\mathbf{v}_i) = \mathbf{0}.$$

That is, $\sum\limits_{i=1}^{k} a_i\mathbf{v}_i \in \textsc{Ker}(T)$. Hence, there exists $b_1, \ldots, b_\ell \in \mathbb{F}$ and $\mathbf{u}_1, \ldots, \mathbf{u}_\ell \in \mathcal{B}$ such that $\sum\limits_{i=1}^{k} a_i\mathbf{v}_i = \sum\limits_{j=1}^{k} b_j\mathbf{u}_j$. Or equivalently, the system $\sum\limits_{i=1}^{k} x_i\mathbf{v}_i + \sum\limits_{j=1}^{k} y_j\mathbf{u}_j = \mathbf{0}$, in the variables x_i's and y_j's, has a non-trivial solution $[a_1, \ldots, a_k, -b_1, \ldots, -b_\ell]^T$ (non-trivial as $\mathbf{a} \neq \mathbf{0}$). Hence, $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ is linearly dependent subset in $\mathbb{V}$. A contradiction to $S \subseteq \mathcal{C}$. That is,

$$\dim(\textsc{Rng}(T)) + \dim(\textsc{Ker}(T)) = |\mathcal{C} \setminus \mathcal{B}| + |\mathcal{B}| = |\mathcal{C}| = \dim(\mathbb{V}).$$

Thus, we have proved the required result. ∎

As an immediate corollary, we have the following result. The proof is left for the reader.

Corollary 4.2.4. *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. If* $\dim(\mathbb{V}) = \dim(\mathbb{W})$ *then, the following statements are equivalent.*

1. *T is one-one.*
2. $\textsc{Ker}(T) = \{\mathbf{0}\}$.
3. *T is onto.*
4. $\dim(\textsc{Rng}(T)) = \dim(\mathbb{V})$.

Corollary 4.2.5. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ with* $\dim(\mathbb{V}) = n$. *If $S, T \in \mathcal{L}(\mathbb{V})$. Then*

1. $\textsc{Nullity}(T) + \textsc{Nullity}(S) \geq \textsc{Nullity}(ST) \geq \max\{\textsc{Nullity}(T), \textsc{Nullity}(S)\}$.
2. $\min\{\mathsf{Rank}(S), \mathsf{Rank}(T)\} \geq \mathsf{Rank}(ST) \geq n - \mathsf{Rank}(S) - \mathsf{Rank}(T)$.

Proof. The prove of Part 2 is omitted as it directly follows from Part 1 and Theorem 4.2.3. Part 1: We first prove the second inequality. Suppose $\mathbf{v} \in \textsc{Ker}(T)$. Then

$$(ST)(\mathbf{v}) = S(T(\mathbf{v})) = S(\mathbf{0}) = \mathbf{0}$$

implies $\textsc{Ker}(T) \subseteq \textsc{Ker}(ST)$. Thus, $\textsc{Nullity}(T) \leq \textsc{Nullity}(ST)$.

By Theorem 4.2.3, $\textsc{Nullity}(S) \leq \textsc{Nullity}(ST)$ is equivalent to $\textsc{Rng}(ST) \subseteq \textsc{Rng}(S)$. And this holds as $\textsc{Rng}(T) \subseteq \mathbb{V}$ implies $\textsc{Rng}(ST) = S(\textsc{Rng}(T)) \subseteq S(\mathbb{V}) = \textsc{Rng}(S)$.

To prove the first inequality, let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis of $\textsc{Ker}(T)$. Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq \textsc{Ker}(ST)$. So, let us extend it to get a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ of $\textsc{Ker}(ST)$.

Claim: $\{T(\mathbf{u}_1), T(\mathbf{u}_2), \ldots, T(\mathbf{u}_\ell)\}$ is a linearly independent subset of $\textsc{Ker}(S)$.

Clearly, $\{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_\ell)\} \subseteq \textsc{Ker}(S)$. Now, consider the system $c_1 T(\mathbf{u}_1) + \cdots + c_\ell T(\mathbf{u}_\ell) = \mathbf{0}$ in the variables $c_1, \ldots, c_\ell$. As $T \in \mathcal{L}(\mathbb{V})$, we get $T\left(\sum\limits_{i=1}^{\ell} c_i \mathbf{u}_i\right) = \mathbf{0}$. Thus, $\sum\limits_{i=1}^{\ell} c_i \mathbf{u}_i \in \textsc{Ker}(T)$. Hence, $\sum\limits_{i=1}^{\ell} c_i \mathbf{u}_i$ is a unique linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$, a basis of $\textsc{Ker}(T)$. Therefore,

$$c_1 \mathbf{u}_1 + \cdots + c_\ell \mathbf{u}_\ell = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k \tag{4.2.1}$$

for some scalars $\alpha_1, \ldots, \alpha_k$. But by assumption, $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ is a basis of $\textsc{Ker}(ST)$ and hence linearly independent. Therefore, the only solution of Equation (4.2.1) is given by $c_i = 0$, for $1 \leq i \leq \ell$ and $\alpha_j = 0$, for $1 \leq j \leq k$. Thus, we have proved the claim. Hence, $\textsc{Nullity}(S) \geq \ell$ and $\textsc{Nullity}(ST) = k + \ell \leq \textsc{Nullity}(T) + \textsc{Nullity}(S)$. ∎

4.2.1 Algebra of Linear Transformations

We start with the following definition.

Definition 4.2.7. **[Sum and Scalar Multiplication of Linear Transformations]** Let $\mathbb{V}, \mathbb{W}$ be vector spaces over $\mathbb{F}$ and let $S, T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Then, we define the point-wise

1. **sum** of S and T, denoted $S+T$, by $(S+T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$, for all $\mathbf{v} \in \mathbb{V}$.
2. **scalar multiplication**, denoted cT for $c \in \mathbb{F}$, by $(cT)(\mathbf{v}) = c\,(T(\mathbf{v}))$, for all $\mathbf{v} \in \mathbb{V}$.

Theorem 4.2.8. *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$. Then $\mathcal{L}(\mathbb{V}, \mathbb{W})$ is a vector space over $\mathbb{F}$. Furthermore, if $\dim \mathbb{V} = n$ and $\dim \mathbb{W} = m$, then $\dim \mathcal{L}(\mathbb{V}, \mathbb{W}) = mn$.*

Proof. It can be easily verified that for $S, T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$, if we define $(S+\alpha T)(\mathbf{v}) = S(\mathbf{v}) + \alpha T(\mathbf{v})$ (point-wise addition and scalar multiplication) then $\mathcal{L}(\mathbb{V}, \mathbb{W})$ is indeed a vector space over $\mathbb{F}$. We now prove the other part. So, let us assume that $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ are bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. For $1 \le i \le n, 1 \le j \le m$, we define the functions $\mathbf{f}_{ij}$ on the basis vectors of $\mathbb{V}$ by

$$\mathbf{f}_{ij}(\mathbf{v}_k) = \begin{cases} \mathbf{w}_j, & \text{if } k = i \\ \mathbf{0}, & \text{if } k \neq i. \end{cases}$$

For other vectors of $\mathbb{V}$, we extend the definition by linearity. That is, if $\mathbf{v} = \sum\limits_{s=1}^{n} \alpha_s \mathbf{v}_s$ then,

$$\mathbf{f}_{ij}(\mathbf{v}) = \mathbf{f}_{ij}\left(\sum_{s=1}^{n} \alpha_s \mathbf{v}_s\right) = \sum_{s=1}^{n} \alpha_s \mathbf{f}_{ij}(\mathbf{v}_s) = \alpha_i \mathbf{f}_{ij}(\mathbf{v}_i) = \alpha_i \mathbf{w}_j. \tag{4.2.2}$$

Thus, $\mathbf{f}_{ij} \in \mathcal{L}(\mathbb{V}, \mathbb{W})$.

Claim: $\{\mathbf{f}_{ij} | 1 \le i \le n, 1 \le j \le m\}$ is a basis of $\mathcal{L}(\mathbb{V}, \mathbb{W})$.

So, consider the linear system $\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} c_{ij} \mathbf{f}_{ij} = \mathbf{0}$, in the variables c_{ij}'s, for $1 \le i \le n, 1 \le j \le m$. Using the point-wise addition and scalar multiplication, we get

$$\mathbf{0} = \mathbf{0}(\mathbf{v}_k) = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \mathbf{f}_{ij}\right)(\mathbf{v}_k) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \mathbf{f}_{ij}(\mathbf{v}_k) = \sum_{j=1}^{m} c_{kj} \mathbf{w}_j.$$

But, the set $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is linearly independent and hence the only solution equals $c_{kj} = 0$, for $1 \le j \le m$. Now, as we vary $\mathbf{v}_k$ from $\mathbf{v}_1$ to $\mathbf{v}_n$, we see that $c_{ij} = 0$, for $1 \le j \le m$ and $1 \le i \le n$. Thus, we have proved the linear independence.

Now, let us prove that $LS\left(\{\mathbf{f}_{ij} | 1 \le i \le n, 1 \le j \le m\}\right) = \mathcal{L}(\mathbb{V}, \mathbb{W})$. So, let $f \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Then, for $1 \le s \le n$, $f(\mathbf{v}_s) \in \mathbb{W}$ and hence there exists β_{st}'s such that $f(\mathbf{v}_s) = \sum\limits_{t=1}^{m} \beta_{st}\mathbf{w}_t$. So, if $\mathbf{v} = \sum\limits_{s=1}^{n} \alpha_s \mathbf{v}_s \in \mathbb{V}$ then, using Equation (4.2.2), we get

$$\begin{aligned}
f(\mathbf{v}) &= f\left(\sum_{s=1}^{n} \alpha_s \mathbf{v}_s\right) = \sum_{s=1}^{n} \alpha_s f(\mathbf{v}_s) = \sum_{s=1}^{n} \alpha_s \left(\sum_{t=1}^{m} \beta_{st}\mathbf{w}_t\right) = \sum_{s=1}^{n}\sum_{t=1}^{m} \beta_{st}(\alpha_s \mathbf{w}_t) \\
&= \sum_{s=1}^{n}\sum_{t=1}^{m} \beta_{st}\mathbf{f}_{st}(\mathbf{v}) = \left(\sum_{s=1}^{n}\sum_{t=1}^{m} \beta_{st}\mathbf{f}_{st}\right)(\mathbf{v}).
\end{aligned}$$

Since the above is true for every $\mathbf{v} \in \mathbb{V}$, $LS\left(\{\mathbf{f}_{ij} | 1 \le i \le n, 1 \le j \le m\}\right) = \mathcal{L}(\mathbb{V}, \mathbb{W})$ and thus the required result follows. ■

Before proceeding further, recall the following definition about a function.

Definition 4.2.9. **[Inverse of a Function]** Let $f : S \to T$ be any function.

1. Then, a function $g : T \to S$ is called a **left inverse** of f if $(g \circ f)(x) = x$, for all $x \in S$. That is, $g \circ f = \mathrm{Id}$, the identity function on S.
2. Then, a function $h : T \to S$ is called a **right inverse** of f if $(f \circ h)(y) = y$, for all $y \in T$. That is, $f \circ h = \mathrm{Id}$, the identity function on T.
3. Then f is said to be **invertible** if it has a right inverse and a left inverse.

Remark 4.2.10. *Let $f : S \to T$ be invertible. Then, it can be easily shown that any right inverse and any left inverse are the same. Thus, the inverse function is unique and is denoted by f^{-1}. It is well known that f is invertible if and only if f is both one-one and onto.*

Lemma 4.2.11. *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. If T is one-one and onto then, the map $T^{-1} : \mathbb{W} \to \mathbb{V}$ is also a linear transformation. The map T^{-1} is called the* **inverse linear transform** *of T and is defined by $T^{-1}(\mathbf{w}) = \mathbf{v}$, whenever $T(\mathbf{v}) = \mathbf{w}$.*

Proof. PART 1: As T is one-one and onto, by Theorem 4.2.3, $\dim(\mathbb{V}) = \dim(\mathbb{W})$. So, by Corollary 4.2.4, for each $\mathbf{w} \in \mathbb{W}$ there exists a unique $\mathbf{v} \in \mathbb{V}$ such that $T(\mathbf{v}) = \mathbf{w}$. Thus, one defines $T^{-1}(\mathbf{w}) = \mathbf{v}$.

We need to show that $T^{-1}(\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2) = \alpha_1 T^{-1}(\mathbf{w}_1) + \alpha_2 T^{-1}(\mathbf{w}_2)$, for all $\alpha_1, \alpha_2 \in \mathbb{F}$ and $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{W}$. Note that by previous paragraph, there exist unique vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}$ such that $T^{-1}(\mathbf{w}_1) = \mathbf{v}_1$ and $T^{-1}(\mathbf{w}_2) = \mathbf{v}_2$. Or equivalently, $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{w}_2$. So, $T(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2$, for all $\alpha_1, \alpha_2 \in \mathbb{F}$. Hence, for all $\alpha_1, \alpha_2 \in \mathbb{F}$, we get

$$T^{-1}(\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2) = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \alpha_1 T^{-1}(\mathbf{w}_1) + \alpha_2 T^{-1}(\mathbf{w}_2).$$

Thus, the required result follows. ■

Definition 4.2.13. **[Singular, Non-singular Transformations]** Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Then, T is said to be **singular** if $\mathbf{0} \subsetneqq \textsc{Ker}(T)$. That is, $\textsc{Ker}(T)$ contains a non-zero vector. If $\textsc{Ker}(T) = \{\mathbf{0}\}$ then, T is called **non-singular**.

Theorem 4.2.15. *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Then the following statements are equivalent.*

1. *T is one-one.*
2. *T is non-singular.*
3. *Whenever $S \subseteq \mathbb{V}$ is linearly independent then $T(S)$ is necessarily linearly independent.*

Proof. 1⇒2 Let T be singular. Then, there exists $\mathbf{v} \neq \mathbf{0}$ such that $T(\mathbf{v}) = \mathbf{0} = T(\mathbf{0})$. This implies that T is not one-one, a contradiction.

2⇒3 Let $S \subseteq \mathbb{V}$ be linearly independent. Let if possible $T(S)$ be linearly dependent. Then, there exists $\mathbf{v}_1, \ldots, \mathbf{v}_k \in S$ and $\alpha = (\alpha_1, \ldots, \alpha_k)^T \neq \mathbf{0}$ such that $\sum\limits_{i=1}^{k} \alpha_i T(\mathbf{v}_i) = \mathbf{0}$. Thus, $T\left(\sum\limits_{i=1}^{k} \alpha_i \mathbf{v}_i\right) = \mathbf{0}$. But T is nonsingular and hence we get $\sum\limits_{i=1}^{k} \alpha_i \mathbf{v}_i = \mathbf{0}$ with $\alpha \neq \mathbf{0}$, a contradiction to S being a linearly independent set.

3⇒1 Suppose that T is not one-one. Then, there exists $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ such that $\mathbf{x} \neq \mathbf{y}$ but $T(\mathbf{x}) = T(\mathbf{y})$. Thus, we have obtained $S = \{\mathbf{x} - \mathbf{y}\}$, a linearly independent subset of $\mathbb{V}$ with $T(S) = \{\mathbf{0}\}$, a linearly dependent set. A contradiction to our assumption. Thus, the required result follows. ∎

Definition 4.2.16. **[Isomorphism of Vector Spaces]** Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Then, T is said to be an **isomorphism** if T is one-one and onto. The vector spaces $\mathbb{V}$ and $\mathbb{W}$ are said to be **isomorphic**, denoted $\mathbb{V} \cong \mathbb{W}$, if there is an isomorphism from $\mathbb{V}$ to $\mathbb{W}$.

We now give a formal proof of the statement in Remark 3.6.9.

Theorem 4.2.17. *Let $\mathbb{V}$ be an n-dimensional vector space over $\mathbb{F}$. Then $\mathbb{V} \cong \mathbb{F}^n$.*

Proof. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of $\mathbb{V}$ and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, the standard basis of $\mathbb{F}^n$. Now define $T(\mathbf{v}_i) = \mathbf{e}_i$, for $1 \leq i \leq n$ and $T\left(\sum\limits_{i=1}^{n} \alpha_i \mathbf{v}_i\right) = \sum\limits_{i=1}^{n} \alpha_i \mathbf{e}_i$, for $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. Then, it is easy to observe that $T \in \mathcal{L}(\mathbb{V}, \mathbb{F}^n)$, T is one-one and onto. Hence, T is an isomorphism. ∎

As a direct application using the countability argument, one obtains the following result

Corollary 4.2.18. *The vector space $\mathbb{R}$ over $\mathbb{Q}$ is not finite dimensional. Similarly, the vector space $\mathbb{C}$ over $\mathbb{Q}$ is not finite dimensional.*

We now summarize the different definitions related with a linear operator on a finite dimensional vector space. The proof basically uses the rank-nullity theorem and they appear in some form in previous results. Hence, we leave the proof for the reader.

Theorem 4.2.19. *Let* $\mathbb{V}$ *be a vector space over* $\mathbb{F}$ *with* $\dim \mathbb{V} = n$. *Then the following statements are equivalent for* $T \in \mathcal{L}(\mathbb{V})$.

1. T *is one-one.*
2. $\textsc{Ker}(T) = \{\mathbf{0}\}$.
3. $\mathsf{Rank}(T) = n$.
4. T *is onto.*
5. T *is an isomorphism.*
6. *If* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *is a basis for* $\mathbb{V}$ *then so is* $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$.
7. T *is non-singular.*
8. T *is invertible.*

Exercise **4.2.20.** *Let* $\mathbb{V}$ *and* $\mathbb{W}$ *be two vector spaces over* $\mathbb{F}$ *and let* $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. *If* $\dim(\mathbb{V})$ *is finite then prove that*

1. T *cannot be onto if* $\dim(\mathbb{V}) < \dim(\mathbb{W})$.
2. T *cannot be one-one if* $\dim(\mathbb{V}) > \dim(\mathbb{W})$.

4.3 Matrix of a linear transformation

In Example 4.1.3.8, we saw that for each $A \in M_{m\times n}(\mathbb{C})$ there exists a linear transformation $T \in \mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$ given by $T(\mathbf{x}) = A\mathbf{x}$, for each $\mathbf{x} \in \mathbb{C}^n$. In this section, we prove that if $\mathbb{V}$ and $\mathbb{W}$ are vector spaces over $\mathbb{F}$ with dimensions n and m, respectively, then any $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ corresponds to a set of $m \times n$ matrices. Before proceeding further, the readers should recall the results on ordered basis (see Section 3.6).

So, let $\mathcal{A} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and $\mathcal{B} = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ be ordered bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. Also, let $A = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ and $B = [\mathbf{w}_1, \ldots, \mathbf{w}_m]$ be the basis matrix of $\mathcal{A}$ and $\mathcal{B}$, respectively. Then, using Equation (3.6.1), $\mathbf{v} = A[\mathbf{v}]_{\mathcal{A}}$ and $\mathbf{w} = B[\mathbf{w}]_{\mathcal{B}}$, for all $\mathbf{v} \in \mathbb{V}$ and $\mathbf{w} \in \mathbb{W}$. Thus, using the above discussion and Equation (4.1.1), we see that for $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ and $\mathbf{v} \in \mathbb{V}$,

$$\begin{aligned}
B[\mathbf{T}(\mathbf{v})]_{\mathcal{B}} &= T(\mathbf{v}) = T(A[\mathbf{v}]_{\mathcal{A}}) = T(A)\ [\mathbf{v}]_{\mathcal{A}} = \begin{bmatrix} T(\mathbf{v}_1) & \cdots & T(\mathbf{v}_n) \end{bmatrix} [\mathbf{v}]_{\mathcal{A}} \\
&= \begin{bmatrix} B[T(\mathbf{v}_1)]_{\mathcal{B}} & \cdots & B[T(\mathbf{v}_n)]_{\mathcal{B}} \end{bmatrix} [\mathbf{v}]_{\mathcal{A}} = B\begin{bmatrix} [T(\mathbf{v}_1)]_{\mathcal{B}} & \cdots & [T(\mathbf{v}_n)]_{\mathcal{B}} \end{bmatrix}[\mathbf{v}]_{\mathcal{A}}.
\end{aligned}$$

Therefore, $[\mathbf{T}(\mathbf{v})]_{\mathcal{B}} = [[T(\mathbf{v}_1)]_{\mathcal{B}}, \ldots, [T(\mathbf{v}_n)]_{\mathcal{B}}]\ [\mathbf{v}]_{\mathcal{A}}$ as a vector in $\mathbb{W}$ has a unique expansion in terms of basis elements. Note that the matrix $\begin{bmatrix} [T(\mathbf{v}_1)]_{\mathcal{B}} & \cdots & [T(\mathbf{v}_n)]_{\mathcal{B}} \end{bmatrix}$, denoted $T[\mathcal{A}, \mathcal{B}]$, is an $m \times n$ matrix and is unique with respect to the ordered basis $\mathcal{B}$ as the i-th column equals $[T(\mathbf{v}_i)]_{\mathcal{B}}$, for $1 \leq i \leq n$. So, we immediately have the following definition and result.

Definition 4.3.1. **[Matrix of a Linear Transformation]** Let $\mathcal{A} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and $\mathcal{B} = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ be ordered bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. If $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ then the matrix $T[\mathcal{A}, \mathcal{B}]$ is called the **coordinate matrix** of T or the **matrix of the linear transformation** T with respect to the basis $\mathcal{A}$ and $\mathcal{B}$, respectively. When there is no mention of bases, we take it to be the standard ordered bases and denote the corresponding matrix by $[T]$.

Note that if $\mathbf{c}$ is the coordinate vector of an element $\mathbf{v} \in \mathbb{V}$ then, $T[\mathcal{A}, \mathcal{B}]\mathbf{c}$ is the coordinate vector of $T(\mathbf{v})$. That is, the matrix $T[\mathcal{A}, \mathcal{B}]$ takes coordinate vector of the domain points to the coordinate vector of its images.

Theorem 4.3.2. *Let $\mathcal{A} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and $\mathcal{B} = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ be ordered bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. If $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ then there exists a matrix $S \in M_{m \times n}(\mathbb{F})$ with*

$$S = T[\mathcal{A}, \mathcal{B}] = \begin{bmatrix} [T(\mathbf{v}_1)]_{\mathcal{B}} & \cdots & [T(\mathbf{v}_n)]_{\mathcal{B}} \end{bmatrix} \text{ and } [T(\mathbf{x})]_{\mathcal{B}} = S\ [\mathbf{x}]_{\mathcal{A}}, \text{ for all } \mathbf{x} \in \mathbb{V}.$$

Remark 4.3.3. *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ with ordered bases $\mathcal{A}_1 = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and $\mathcal{B}_1 = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$, respectively. Also, for $\alpha \in \mathbb{F}$ with $\alpha \neq 0$, let $\mathcal{A}_2 = (\alpha\mathbf{v}_1, \ldots, \alpha\mathbf{v}_n)$ and $\mathcal{B}_1 = (\alpha\mathbf{w}_1, \ldots, \alpha\mathbf{w}_m)$ be another set of ordered bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. Then, for any $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$*

$$T[\mathcal{A}_2, \mathcal{B}_2] = \begin{bmatrix} [T(\alpha\mathbf{v}_1)]_{\mathcal{B}_2} & \cdots & [T(\alpha\mathbf{v}_1)]_{\mathcal{B}_2} \end{bmatrix} = \begin{bmatrix} [T(\mathbf{v}_n)]_{\mathcal{B}_1} & \cdots & [T(\mathbf{v}_1)]_{\mathcal{B}_1} \end{bmatrix} = T[\mathcal{A}_1, \mathcal{B}_1].$$

Thus, we see that the same matrix can be the matrix representation of T for two different pairs of bases.

We now give a few examples to understand the above discussion and Theorem 4.3.2.

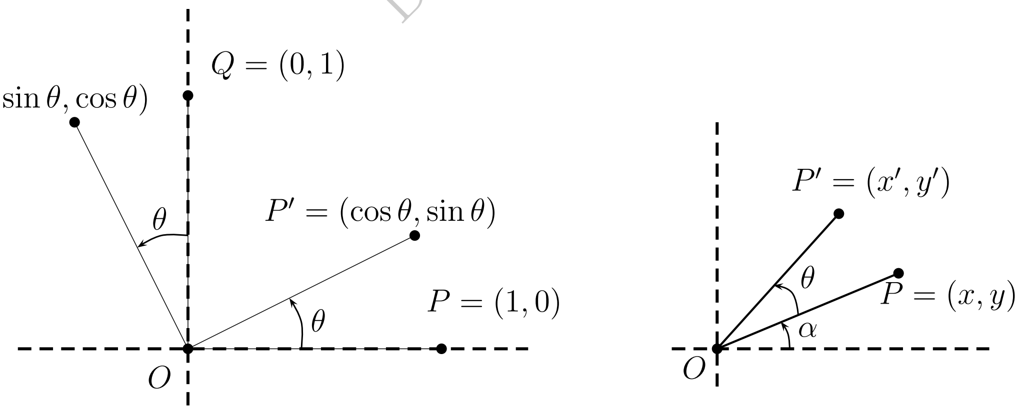


Figure 4.1: Counter-clockwise Rotation by an angle θ

4.4 Similarity of Matrices

Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ with $\dim(\mathbb{V}) = n$ and ordered basis $\mathcal{B}$. Then any $T \in \mathcal{L}(\mathbb{V})$ corresponds to a matrix in $\mathbb{M}_n(\mathbb{F})$. What happens if the ordered basis needs to change? We answer this in this subsection.

$$\begin{tikzcd}
(\mathbb{V}, \mathcal{B}, n) \arrow[r, "{T[\mathcal{B},\mathcal{C}]_{m\times n}}"] \arrow[rr, bend right, "{(ST)[\mathcal{B},\mathcal{D}]_{p\times n} = S[\mathcal{C},\mathcal{D}] \cdot T[\mathcal{B},\mathcal{C}]}"'] & (\mathbb{W}, \mathcal{C}, m) \arrow[r, "{S[\mathcal{C},\mathcal{D}]_{p\times m}}"] & (\mathbb{Z}, \mathcal{D}, p)
\end{tikzcd}$$

Figure 4.2: Composition of Linear Transformations

Theorem 4.4.1 (Composition of Linear Transformations)**.** *Let $\mathbb{V}$, $\mathbb{W}$ and $\mathbb{Z}$ be finite dimensional vector spaces over $\mathbb{F}$ with ordered bases $\mathcal{B}, \mathcal{C}$ and $\mathcal{D}$, respectively. Also, let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ and $S \in \mathcal{L}(\mathbb{W}, \mathbb{Z})$. Then $S \circ T = ST \in \mathcal{L}(\mathbb{V}, \mathbb{Z})$ (see Figure 4.2). Then*

$$(ST)\ [\mathcal{B}, \mathcal{D}] = S[\mathcal{C}, \mathcal{D}]\ \cdot\ T[\mathcal{B}, \mathcal{C}].$$

Proof. Let $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$, $\mathcal{C} = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ and $\mathcal{D} = (\mathbf{w}_1, \ldots, \mathbf{w}_p)$ be the ordered bases of $\mathbb{V}, \mathbb{W}$ and $\mathbb{Z}$, respectively. Then using Theorem 4.3.2, we have

$$\begin{aligned}
(ST)[\mathcal{B}, \mathcal{D}] &= [[ST(\mathbf{u}_1)]_{\mathcal{D}}, \ldots, [ST(\mathbf{u}_n)]_{\mathcal{D}}] = [[S(T(\mathbf{u}_1))]_{\mathcal{D}}, \ldots, [S(T(\mathbf{u}_n))]_{\mathcal{D}}] \\
&= [S[\mathcal{C}, \mathcal{D}]\ [T(\mathbf{u}_1)]_{\mathcal{C}}, \ldots, S[\mathcal{C}, \mathcal{D}]\ [T(\mathbf{u}_n)]_{\mathcal{C}}] \\
&= S[\mathcal{C}, \mathcal{D}]\ \ [[T(\mathbf{u}_1)]_{\mathcal{C}}, \ldots, [T(\mathbf{u}_n)]_{\mathcal{C}}] = S[\mathcal{C}, \mathcal{D}]\ \cdot\ T[\mathcal{B}, \mathcal{C}].
\end{aligned}$$

Hence, the proof of the theorem is complete. ■

As an immediate corollary of Theorem 4.4.1 we have the following result.

Theorem 4.4.2 (Inverse of a Linear Transformation)**.** *Let $\mathbb{V}$ is a vector space with $\dim(\mathbb{V}) = n$. If $T \in \mathcal{L}(\mathbb{V})$ is invertible then for any ordered basis $\mathcal{B}$ and $\mathcal{C}$ of the domain and co-domain, respectively, one has $(T[\mathcal{C}, \mathcal{B}])^{-1} = T^{-1}[\mathcal{B}, \mathcal{C}]$. That is, the inverse of the coordinate matrix of T is the coordinate matrix of the inverse linear transform.*

Proof. As T is invertible, $TT^{-1} = \mathrm{Id}$. Thus, Example 4.3.4.8a and Theorem 4.4.1 imply

$$I_n = \mathrm{Id}[\mathcal{B}, \mathcal{B}] = (TT^{-1})[\mathcal{B}, \mathcal{B}] = T[\mathcal{C}, \mathcal{B}]\ \cdot T^{-1}[\mathcal{B}, \mathcal{C}].$$

Hence, by definition of inverse, $T^{-1}[\mathcal{B}, \mathcal{C}] = (T[\mathcal{C}, \mathcal{B}])^{-1}$ and the required result follows. ■

Theorem 4.4.4. *Let $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ and $\mathcal{C} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be two ordered bases of $\mathbb{V}$ and Id the identity operator. Then, for any linear operator $T \in \mathcal{L}(\mathbb{V})$*

$$T[\mathcal{C},\mathcal{C}] = Id[\mathcal{B},\mathcal{C}] \cdot T[\mathcal{B},\mathcal{B}] \cdot Id[\mathcal{C},\mathcal{B}] = (Id[\mathcal{C},\mathcal{B}])^{-1} \cdot T[\mathcal{B},\mathcal{B}] \cdot Id[\mathcal{C},\mathcal{B}]. \tag{4.4.1}$$

Proof. As Id is an identity operator, $T[\mathcal{B},\mathcal{C}]$ as $(\mathrm{Id} \circ T \circ \mathrm{Id})[\mathcal{B},\mathcal{C}]$ (see Figure 4.3 for clarity). Thus, using Theorem 4.4.1, we get

$$T[\mathcal{B},\mathcal{C}] = (\mathrm{Id} \circ T \circ \mathrm{Id})[\mathcal{B},\mathcal{C}] = \mathrm{Id}[\mathcal{B},\mathcal{C}] \cdot T[\mathcal{B},\mathcal{B}] \cdot \mathrm{Id}[\mathcal{C},\mathcal{B}].$$

Hence, using Theorem 4.4.2, the required result follows. ■

Let $\mathbb{V}$ be a vector space and let $T \in \mathcal{L}(\mathbb{V})$. If $\dim(\mathbb{V}) = n$ then every ordered basis $\mathcal{B}$ of $\mathbb{V}$ gives an $n \times n$ matrix $T[\mathcal{B},\mathcal{B}]$. So, as we change the ordered basis, the coordinate matrix of T changes. Theorem 4.4.4 tells us that all these matrices are related by an invertible matrix. Thus, we are led to the following definitions.

Definition 4.4.5. [Change of Basis Matrix] Let $\mathbb{V}$ be a vector space with ordered bases $\mathcal{B}$ and $\mathcal{C}$. If $T \in \mathcal{L}(\mathbb{V})$ then, $T[\mathcal{C},\mathcal{C}] = \mathrm{Id}[\mathcal{B},\mathcal{C}] \cdot T[\mathcal{B},\mathcal{B}] \cdot \mathrm{Id}[\mathcal{C},\mathcal{B}]$. The matrix $\mathrm{Id}[\mathcal{B},\mathcal{C}]$ is called the **change of basis matrix** (also, see Theorem 3.6.7) from $\mathcal{B}$ to $\mathcal{C}$.

Definition 4.4.6. [Similar Matrices] Let $X, Y \in \mathbb{M}_n(\mathbb{C})$. Then, X and Y are said to be **similar** if there exists a non-singular matrix P such that $P^{-1}XP = Y \Leftrightarrow XP = PY$.

4.5 Dual Space*

Definition 4.5.1. **[linear Functional]** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then a map $T \in \mathcal{L}(\mathbb{V}, \mathbb{F})$ is called a **linear functional** on $\mathbb{V}$.

Definition 4.5.3. [**Dual Space**] Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then $\mathcal{L}(\mathbb{V}, \mathbb{F})$ is called the **dual space** of $\mathbb{V}$ and is denoted by $\mathbb{V}^*$. The **double dual space** of $\mathbb{V}$, denoted $\mathbb{V}^{**}$, is the dual space of $\mathbb{V}^*$.

We first give an immediate corollary of Theorem 4.2.17.

Corollary 4.5.4. *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ with $\dim \mathbb{V} = n$ and $\dim \mathbb{W} = m$.*

1. *Then $\mathcal{L}(\mathbb{V}, \mathbb{W}) \cong \mathbb{F}^{mn}$. Moreover, $\{\mathbf{f}_{ij} | 1 \le i \le n, 1 \le j \le m\}$ is a basis of $\mathcal{L}(\mathbb{V}, \mathbb{W})$.*
2. *In particular, if $\mathbb{W} = \mathbb{F}$ then $\mathcal{L}(\mathbb{V}, \mathbb{F}) = \mathbb{V}^* \cong \mathbb{F}^n$. Moreover, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of $\mathbb{V}$ then the set $\{\mathbf{f}_i | 1 \le i \le n\}$ is a basis of $\mathbb{V}^*$, where $\mathbf{f}_i(\mathbf{v}_k) = \begin{cases} 1, & \textit{if } k = i \\ \mathbf{0}, & k \neq i. \end{cases}$ The basis $\{\mathbf{f}_i | 1 \le i \le n\}$ is called the* **dual basis** *of $\mathbb{F}^n$.*

Exercise **4.5.5.** *Let $\mathbb{V}$ be a vector space. Suppose there exists $\mathbf{v} \in \mathbb{V}$ such that $\mathbf{f}(\mathbf{v}) = 0$, for all $\mathbf{f} \in \mathbb{V}^*$. Then prove that $\mathbf{v} = \mathbf{0}$.*

So, we see that $\mathbb{V}^*$ can be understood through a basis of $\mathbb{V}$. Thus, one can understand $\mathbb{V}^{**}$ again via a basis of $\mathbb{V}^*$. But, the question arises "can we understand it directly via the vector space $\mathbb{V}$ itself?" We answer this in affirmative by giving a canonical isomorphism from $\mathbb{V}$ to $\mathbb{V}^{**}$. To do so, for each $\mathbf{v} \in \mathbb{V}$, we define a map $L_{\mathbf{v}} : \mathbb{V}^* \to \mathbb{F}$ by $L_{\mathbf{v}}(\mathbf{f}) = \mathbf{f}(\mathbf{v})$, for each $\mathbf{f} \in \mathbb{V}^*$. Then $L_{\mathbf{v}}$ is a linear functional as

$$L_{\mathbf{v}}(\alpha \mathbf{f} + \mathbf{g}) = (\alpha \mathbf{f} + \mathbf{g})(\mathbf{v}) = \alpha \mathbf{f}(\mathbf{v}) + \mathbf{g}(\mathbf{v}) = \alpha L_{\mathbf{v}}(\mathbf{f}) + L_{\mathbf{v}}(\mathbf{g}).$$

So, for each $\mathbf{v} \in \mathbb{V}$, we have obtained a linear functional $L_{\mathbf{v}} \in \mathbb{V}^{**}$. Note that, if $\mathbf{v} \neq \mathbf{w}$ then, $L_{\mathbf{v}} \neq L_{\mathbf{w}}$. Indeed, if $L_{\mathbf{v}} = L_{\mathbf{w}}$ then, $L_{\mathbf{v}}(f) = L_{\mathbf{w}}(f)$, for all $f \in \mathbb{V}^*$. Thus, $f(\mathbf{v}) = f(\mathbf{w})$, for all $f \in \mathbb{V}^*$. That is, $f(\mathbf{v} - \mathbf{w}) = 0$, for each $f \in \mathbb{V}^*$. Hence, using Exercise 4.5.5, we get $\mathbf{v} - \mathbf{w} = \mathbf{0}$, or equivalently, $\mathbf{v} = \mathbf{w}$.

We use the above argument to give the required canonical isomorphism.

Theorem 4.5.6. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. If $\dim(\mathbb{V}) = n$ then the canonical map $T : \mathbb{V} \to \mathbb{V}^{**}$ defined by $T(\mathbf{v}) = L_{\mathbf{v}}$ is an isomorphism.*

Proof. Note that for each $\mathbf{f} \in \mathbb{V}^*$,

$$L_{\alpha \mathbf{v} + \mathbf{u}}(\mathbf{f}) = \mathbf{f}(\alpha \mathbf{v} + \mathbf{u}) = \alpha \mathbf{f}(\mathbf{v}) + \mathbf{f}(\mathbf{u}) = \alpha L_{\mathbf{v}}(\mathbf{f}) + L_{\mathbf{u}}(\mathbf{f}) = (\alpha L_{\mathbf{v}} + L_{\mathbf{u}})(\mathbf{f}).$$

Thus, $L_{\alpha\mathbf{v}+\mathbf{u}} = \alpha L_{\mathbf{v}} + L_{\mathbf{u}}$. Hence, $T(\alpha\mathbf{v}+\mathbf{u}) = \alpha T(\mathbf{v}) + T(\mathbf{u})$. Thus, T is a linear transformation. For verifying T is one-one, assume that $T(\mathbf{v}) = T(\mathbf{u})$, for some $\mathbf{u}, \mathbf{v} \in \mathbb{V}$. Then, $L_{\mathbf{v}} = L_{\mathbf{u}}$. Now, use the argument just before this theorem to get $\mathbf{v} = \mathbf{u}$. Therefore, T is one-one.

Thus, T gives an inclusion (one-one) map from $\mathbb{V}$ to $\mathbb{V}^{**}$. Further, applying Corollary 4.5.4.2 to $\mathbb{V}^*$, gives $\dim(\mathbb{V}^{**}) = \dim(\mathbb{V}^*) = n$. Hence, the required result follows. ∎

We now give a few immediate consequences of Theorem 4.5.6.

Corollary 4.5.7. *Let $\mathbb{V}$ be a vector space of dimension n with basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.*

1. *Then, a basis of $\mathbb{V}^{**}$, the double dual of $\mathbb{V}$, equals $\mathcal{D} = \{L_{\mathbf{v}_1}, \ldots, L_{\mathbf{v}_n}\}$. Thus, for each $T \in \mathbb{V}^{**}$ there exists $\mathbf{x} \in \mathbb{V}$ such that $T(\mathbf{f}) = \mathbf{f}(\mathbf{x})$, for all $\mathbf{f} \in \mathbb{V}^*$. Or equivalently, there exists $\mathbf{x} \in \mathbb{V}$ such that $T = T_{\mathbf{x}}$.*
2. *If $\mathcal{C} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ is the dual basis of $\mathbb{V}^*$ defined using the basis $\mathcal{B}$ (see Corollary 4.5.4.2) then $\mathcal{D}$ is indeed the dual basis of $\mathbb{V}^{**}$ obtained using the basis $\mathcal{C}$ of $\mathbb{V}^*$. Thus, each basis of $\mathbb{V}^*$ is the dual basis of some basis of $\mathbb{V}$.*

Proof. Part 1 is direct as $T : \mathbb{V} \to \mathbb{V}^{**}$ was a canonical inclusion map. For Part 2, we need to show that

$$L_{\mathbf{v}_i}(\mathbf{f}_j) = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{cases} \quad \text{or equivalently } \mathbf{f}_j(\mathbf{v}_i) = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{cases}$$

which indeed holds true using Corollary 4.5.4.2. ∎

Let $\mathbb{V}$ be a finite dimensional vector space. Then Corollary 4.5.7 implies that the spaces $\mathbb{V}$ and $\mathbb{V}^*$ are naturally dual to each other.

We are now ready to prove the main result of this subsection. To start with, let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$. Then, for each $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$, we want to define a map $\widehat{T} : \mathbb{W}^* \to \mathbb{V}^*$. So, if $g \in \mathbb{W}^*$ then, $\widehat{T}(\mathbf{g})$ a linear functional from $\mathbb{V}$ to $\mathbb{F}$. So, we need to be evaluate $\widehat{T}(\mathbf{g})$ at an element of $\mathbb{V}$. Thus, we define $\left(\widehat{T}(\mathbf{g})\right)(\mathbf{v}) = g\left(T(\mathbf{v})\right)$, for all $\mathbf{v} \in \mathbb{V}$. Now, we note that $\widehat{T} \in \mathcal{L}(\mathbb{W}^*, \mathbb{V}^*)$, as for every $g, h \in \mathbb{W}^*$,

$$\left(\widehat{T}(\alpha\mathbf{g} + \mathbf{h})\right)(\mathbf{v}) = (\alpha\mathbf{g} + \mathbf{h})\left(T(\mathbf{v})\right) = \alpha\mathbf{g}\left(T(\mathbf{v})\right) + \mathbf{h}\left(T(\mathbf{v})\right) = \left(\alpha\widehat{T}(\mathbf{g}) + \widehat{T}(\mathbf{h})\right)(\mathbf{v}),$$

for all $\mathbf{v} \in \mathbb{V}$ implies that $\widehat{T}(\alpha\mathbf{g} + \mathbf{h}) = \alpha\widehat{T}(\mathbf{g}) + \widehat{T}(\mathbf{h})$.

Theorem 4.5.8. *Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$ with ordered bases $\mathcal{A} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and $\mathcal{B} = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$, respectively. Also, let $\mathcal{A}^* = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$ and $\mathcal{B}^* = (\mathbf{g}_1, \ldots, \mathbf{g}_m)$ be the corresponding ordered bases of the dual spaces $\mathbb{V}^*$ and $\mathbb{W}^*$, respectively. Then,*

$$\widehat{T}[\mathcal{B}^*, \mathcal{A}^*] = \left(T[\mathcal{A}, \mathcal{B}]\right)^T,$$

the transpose of the coordinate matrix T.

Proof. Note that we need to compute $\widehat{T}[\mathcal{B}^*, \mathcal{A}^*] = \left[\left[\widehat{T}(\mathbf{g}_1)\right]_{\mathcal{A}^*}, \ldots, \left[\widehat{T}(\mathbf{g}_m)\right]_{\mathcal{A}^*}\right]$ and prove that it equals the transpose of the matrix $T[\mathcal{A}, \mathcal{B}]$. So, let

$$T[\mathcal{A}, \mathcal{B}] = \left[[T(\mathbf{v}_1)]_{\mathcal{B}}, \ldots, [T(\mathbf{v}_n)]_{\mathcal{B}}\right] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Thus, to prove the required result, we need to show that

$$\left[\widehat{T}(\mathbf{g}_j)\right]_{\mathcal{A}^*} = [\mathbf{f}_1, \ldots, \mathbf{f}_n] \begin{bmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jn} \end{bmatrix} = \sum_{k=1}^{n} a_{jk}\mathbf{f}_k, \text{ for } 1 \le j \le m. \tag{4.5.1}$$

Now, recall that the functionals $\mathbf{f}_i$'s and $\mathbf{g}_j$'s satisfy $\left(\sum\limits_{k=1}^{n} \alpha_k \mathbf{f}_k\right)(\mathbf{v}_t) = \sum\limits_{k=1}^{n} \alpha_k \left(\mathbf{f}_k(\mathbf{v}_t)\right) = \alpha_t$, for $1 \le t \le n$ and $[\mathbf{g}_j(\mathbf{w}_1), \ldots, \mathbf{g}_j(\mathbf{w}_m)] = \mathbf{e}_j^T$, a row vector with 1 at the j-th place and 0, elsewhere. So, let $B = [\mathbf{w}_1, \ldots, \mathbf{w}_m]$ and evaluate $\widehat{T}(\mathbf{g}_j)$ at $\mathbf{v}_t$'s, the elements of $\mathcal{A}$.

$$\begin{aligned}
\left(\widehat{T}(\mathbf{g}_j)\right)(\mathbf{v}_t) &= \mathbf{g}_j\left(T(\mathbf{v}_t)\right) = \mathbf{g}_j\left(B\,[T(\mathbf{v}_t)]_{\mathcal{B}}\right) = [\mathbf{g}_j(\mathbf{w}_1), \ldots, \mathbf{g}_j(\mathbf{w}_m)]\,[T(\mathbf{v}_t)]_{\mathcal{B}} \\
&= \mathbf{e}_j^T \begin{bmatrix} a_{1t} \\ a_{2t} \\ \vdots \\ a_{mt} \end{bmatrix} = a_{jt} = \left(\sum_{k=1}^{n} a_{jk}\mathbf{f}_k\right)(\mathbf{v}_t).
\end{aligned}$$

Thus, the linear functional $\widehat{T}(\mathbf{g}_j)$ and $\sum\limits_{k=1}^{n} a_{jk}\mathbf{f}_k$ are equal at $\mathbf{v}_t$, for $1 \le t \le n$, the basis vectors of $\mathbb{V}$. Hence $\widehat{T}(\mathbf{g}_j) = \sum\limits_{k=1}^{n} a_{jk}\mathbf{f}_k$ which gives Equation (4.5.1). ∎

Remark 4.5.9. *The proof of Theorem 4.5.8 also shows the following.*

1. *For each $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ there exists a unique map $\widehat{T} \in \mathcal{L}(\mathbb{W}^*, \mathbb{V}^*)$ such that*

$$\left(\widehat{T}(\mathbf{g})\right)(\mathbf{v}) = \mathbf{g}\left(T(\mathbf{v})\right), \textit{ for each } \mathbf{g} \in \mathbb{W}^*.$$

2. *The coordinate matrices $T[\mathcal{A}, \mathcal{B}]$ and $\widehat{T}[\mathcal{B}^*, \mathcal{A}^*]$ are transpose of each other, where the ordered bases $\mathcal{A}^*$ of $\mathbb{V}^*$ and $\mathcal{B}^*$ of $\mathbb{W}^*$ correspond, respectively, to the ordered bases $\mathcal{A}$ of $\mathbb{V}$ and $\mathcal{B}$ of $\mathbb{W}$.*

3. *Thus, the results on matrices and its transpose can be re-written in the language a vector space and its dual space.*

4.6 Summary

Chapter 5

Inner Product Spaces

5.1 Definition and Basic Properties

Recall the dot product in $\mathbb{R}^2$ and $\mathbb{R}^3$. Dot product helped us to compute the length of vectors and angle between vectors. This enabled us to rephrase geometrical problems in $\mathbb{R}^2$ and $\mathbb{R}^3$ in the language of vectors. We generalize the idea of dot product to achieve similar goal for a general vector space over $\mathbb{R}$ or $\mathbb{C}$. So, in this chapter $\mathbb{F}$ will denote either $\mathbb{R}$ or $\mathbb{C}$.

Definition 5.1.1. [**Inner Product**] Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. An **inner product** over $\mathbb{V}$, denoted by $\langle\ ,\ \rangle$, is a map from $\mathbb{V} \times \mathbb{V}$ to $\mathbb{F}$ satisfying

1. $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w}\rangle = a\langle \mathbf{u}, \mathbf{w}\rangle + b\langle \mathbf{v}, \mathbf{w}\rangle$, for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{V}$ and $a, b \in \mathbb{F}$,
2. $\langle \mathbf{u}, \mathbf{v}\rangle = \overline{\langle \mathbf{v}, \mathbf{u}\rangle}$, the complex conjugate of $\langle \mathbf{u}, \mathbf{v}\rangle$, for all $\mathbf{u}, \mathbf{v} \in \mathbb{V}$ and
3. $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$ for all $\mathbf{u} \in \mathbb{V}$. Furthermore, equality holds if and only if $\mathbf{u} = \mathbf{0}$.

Remark 5.1.2. *Using the definition of inner product, we immediately observe that*

1. $\langle \mathbf{v}, \alpha\mathbf{w}\rangle = \overline{\langle \alpha\mathbf{w}, \mathbf{v}\rangle} = \overline{\alpha}\overline{\langle \mathbf{w}, \mathbf{v}\rangle} = \overline{\alpha}\langle \mathbf{v}, \mathbf{w}\rangle$, *for all* $\alpha \in \mathbb{F}$ *and* $\mathbf{v}, \mathbf{w} \in \mathbb{V}$.
2. *If* $\langle \mathbf{u}, \mathbf{v}\rangle = 0$ *for all* $\mathbf{v} \in \mathbb{V}$ *then in particular* $\langle \mathbf{u}, \mathbf{u}\rangle = 0$. *Hence,* $\mathbf{u} = \mathbf{0}$.

Definition 5.1.3. [**Inner Product Space**] Let $\mathbb{V}$ be a vector space with an inner product $\langle\ ,\ \rangle$. Then, $(\mathbb{V}, \langle\ ,\ \rangle)$ is called an **inner product space** (in short, IPS).

Example 5.1.4. Examples 1 and 2 that appear below are called the **standard inner product** or the **dot product** on $\mathbb{R}^n$ and $\mathbb{C}^n$, respectively. Whenever an inner product is not clearly mentioned, it will be assumed to be the standard inner product.

1. For $\mathbf{u} = (u_1, \ldots, u_n)^T, \mathbf{v} = (v_1, \ldots, v_n)^T \in \mathbb{R}^n$ define $\langle \mathbf{u}, \mathbf{v}\rangle = u_1v_1 + \cdots + u_nv_n = \mathbf{v}^T\mathbf{u}$. Then, $\langle\ ,\ \rangle$ is indeed an inner product and hence $\big(\mathbb{R}^n, \langle\ ,\ \rangle\big)$ is an IPS.
2. For $\mathbf{u} = (u_1, \ldots, u_n)^*, \mathbf{v} = (v_1, \ldots, v_n)^* \in \mathbb{C}^n$ define $\langle \mathbf{u}, \mathbf{v}\rangle = u_1\overline{v_1} + \cdots + u_n\overline{v_n} = \mathbf{v}^*\mathbf{u}$. Then, $\big(\mathbb{C}^n, \langle\ ,\ \rangle\big)$ is an IPS.
3. For $\mathbf{x} = (x_1, x_2)^T, \mathbf{y} = (y_1, y_2)^T \in \mathbb{R}^2$ and $A = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$, define $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^TA\mathbf{x}$. Then, $\langle\ ,\ \rangle$ is an inner product as $\langle \mathbf{x}, \mathbf{x}\rangle = (x_1 - x_2)^2 + 3x_1^2 + x_2^2$.

4. Fix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ with $a, c > 0$ and $ac > b^2$. Then, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T A \mathbf{x}$ is an inner product on $\mathbb{R}^2$ as $\langle \mathbf{x}, \mathbf{x} \rangle = ax_1^2 + 2bx_1x_2 + cx_2^2 = a\left[x_1 + \frac{bx_2}{a}\right]^2 + \frac{1}{a}\left[ac - b^2\right]x_2^2$.

5. Verify that for $\mathbf{x} = (x_1, x_2, x_3)^T, \mathbf{y} = (y_1, y_2, y_3)^T \in \mathbb{R}^3$, $\langle \mathbf{x}, \mathbf{y} \rangle = 10x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2 + x_2y_3 + x_3y_2 + x_3y_3$ defines an inner product.

6. For $\mathbf{x} = (x_1, x_2)^T, \mathbf{y} = (y_1, y_2)^T \in \mathbb{R}^2$, we define three maps that satisfy at least one condition out of the three conditions for an inner product. Determine the condition which is not satisfied. Give reasons for your answer.

   (a) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1$.

   (b) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1^2 + y_1^2 + x_2^2 + y_2^2$.

   (c) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1^3 + x_2y_2^3$.

7. Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix. Then, for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, define $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* A \mathbf{x}$. Then, $\langle \ , \ \rangle$ satisfies $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ and $\langle \mathbf{x} + \alpha \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \alpha \langle \mathbf{z}, \mathbf{y} \rangle$, for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{C}^n$ and $\alpha \in \mathbb{C}$. Does there exist conditions on A such that $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ for all $\mathbf{x} \in \mathbb{C}$? This will be answered in affirmative in the chapter on eigenvalues and eigenvectors.

8. For $A, B \in M_n(\mathbb{R})$, define $\langle A, B \rangle = tr(B^T A)$. Then,

$$\begin{aligned} \langle A + B, C \rangle &= \mathrm{tr}\big(C^T (A + B)\big) = \mathrm{tr}(C^T A) + \mathrm{tr}(C^T B) = \langle A, C \rangle + \langle B, C \rangle \text{ and} \\ \langle A, B \rangle &= \mathrm{tr}(B^T A) = \mathrm{tr}\left((B^T A)^T \right) = \mathrm{tr}(A^T B) = \langle B, A \rangle. \end{aligned}$$

   If $A = [a_{ij}]$ then $\langle A, A \rangle = \mathrm{tr}(A^T A) = \sum\limits_{i=1}^{n} (A^T A)_{ii} = \sum\limits_{i,j=1}^{n} a_{ij} a_{ij} = \sum\limits_{i,j=1}^{n} a_{ij}^2$ and therefore, $\langle A, A \rangle > 0$ for all nonzero matrix A.

9. Consider the complex vector space $\mathcal{C}[-1, 1]$ and define $\langle f, g \rangle = \int\limits_{-1}^{1} f(x)\overline{g(x)}dx$. Then,

   (a) $\langle \mathbf{f}, \mathbf{f} \rangle = \int\limits_{-1}^{1} \mid \mathbf{f}(x) \mid^2 dx \geq 0$ as $\mid \mathbf{f}(x) \mid^2 \geq 0$ and this integral is 0 if and only if $\mathbf{f} \equiv \mathbf{0}$ as f is continuous.

   (b) $\overline{\langle \mathbf{g}, \mathbf{f} \rangle} = \overline{\int\limits_{-1}^{1} \mathbf{g}(x)\overline{\mathbf{f}(x)}dx} = \int\limits_{-1}^{1} \overline{\mathbf{g}(x)\overline{\mathbf{f}(x)}}dx = \int\limits_{-1}^{1} \mathbf{f}(x)\overline{\mathbf{g}(x)}dx = \langle \mathbf{f}, \mathbf{g} \rangle$.

   (c) $\langle \mathbf{f} + \mathbf{g}, \mathbf{h} \rangle = \int\limits_{-1}^{1} (\mathbf{f} + \mathbf{g})(x)\overline{\mathbf{h}(x)}dx = \int\limits_{-1}^{1} [\mathbf{f}(x)\overline{\mathbf{h}(x)} + \mathbf{g}(x)\overline{\mathbf{h}(x)}]dx = \langle \mathbf{f}, \mathbf{h} \rangle + \langle \mathbf{g}, \mathbf{h} \rangle$.

   (d) $\langle \alpha \mathbf{f}, \mathbf{g} \rangle = \int\limits_{-1}^{1} (\alpha \mathbf{f}(x))\overline{\mathbf{g}(x)}dx = \alpha \int\limits_{-1}^{1} \mathbf{f}(x)\overline{\mathbf{g}(x)}dx = \alpha \langle \mathbf{f}, \mathbf{g} \rangle$.

   (e) Fix an ordered basis $\mathcal{B} = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ of a complex vector space $\mathbb{V}$. Then, for any $\mathbf{u}, \mathbf{v} \in \mathbb{V}$, with $[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ and $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$, define $\langle \mathbf{u}, \mathbf{v} \rangle = \sum\limits_{i=1}^{n} a_i \overline{b_i}$. Then, $\langle \ , \ \rangle$ is indeed an inner product in $\mathbb{V}$. So, any finite dimensional vector space can be endowed with an inner product.

5.1.1 Cauchy Schwartz Inequality

As $\langle \mathbf{u}, \mathbf{u} \rangle > 0$, for all $\mathbf{u} \neq \mathbf{0}$, we use inner product to define length of a vector.

Definition 5.1.5. [Length / Norm of a Vector] Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then, for any vector $\mathbf{u} \in \mathbb{V}$, we define the **length (norm)** of $\mathbf{u}$, denoted $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$, the positive square root. A vector of norm 1 is called a **unit vector**. Thus, $\dfrac{\mathbf{u}}{\|\mathbf{u}\|}$ is called the **unit vector in the direction** of $\mathbf{u}$.

Example 5.1.6. 1. Let $\mathbb{V}$ be an IPS and $\mathbf{u} \in \mathbb{V}$. Then, for any scalar α, $\|\alpha \mathbf{u}\| = |\alpha| \cdot \|\mathbf{u}\|$.

2. Let $\mathbf{u} = (1, -1, 2, -3)^T \in \mathbb{R}^4$. Then, $\|\mathbf{u}\| = \sqrt{1+1+4+9} = \sqrt{15}$. Thus, $\frac{1}{\sqrt{15}}\mathbf{u}$ and $-\frac{1}{\sqrt{15}}\mathbf{u}$ are vectors of norm 1. Moreover $\frac{1}{\sqrt{15}}\mathbf{u}$ is a unit vector in the direction of $\mathbf{u}$.

EXERCISE **5.1.7.** *1. Let $\mathbf{u} = (-1, 1, 2, 3, 7)^T \in \mathbb{R}^5$. Find all $\alpha \in \mathbb{R}$ such that $\|\alpha \mathbf{u}\| = 1$.*

2. Let $\mathbf{u} = (-1, 1, 2, 3, 7)^T \in \mathbb{C}^5$. Find all $\alpha \in \mathbb{C}$ such that $\|\alpha \mathbf{u}\| = 1$.

3. Prove that $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\big(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\big)$, for all $\mathbf{x}^T, \mathbf{y}^T \in \mathbb{R}^n$. This equality is called the PARALLELOGRAM LAW *as in a parallelogram the sum of square of the lengths of the diagonals is equal to twice the sum of squares of the lengths of the sides.*

4. **Apollonius' Identity:** *Let the length of the sides of a triangle be $a, b, c \in \mathbb{R}$ and that of the median be $d \in \mathbb{R}$. If the median is drawn on the side with length a then prove that $b^2 + c^2 = 2\left(d^2 + \left(\frac{a}{2}\right)^2\right)$.*

5. Let $\mathbf{u} = (1, 2)^T, \mathbf{v} = (2, -1)^T \in \mathbb{R}^2$. Then, does there exist an inner product in $\mathbb{R}^2$ such that $\|\mathbf{u}\| = 1, \|\mathbf{v}\| = 1$ and $\langle \mathbf{u}, \mathbf{v} \rangle = 0$? [Hint: Let $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ and define $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T A \mathbf{x}$. Use given conditions to get a linear system of 3 equations in the variables a, b, c.]

6. Let $\mathbf{x} = (x_1, x_2)^T$, $\mathbf{y} = (y_1, y_2)^T \in \mathbb{R}^2$. Then, $\langle \mathbf{x}, \mathbf{y} \rangle = 3x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2$ defines an inner product. Use this inner product to find

(a) the angle between $\mathbf{e}_1 = (1, 0)^T$ and $\mathbf{e}_2 = (0, 1)^T$.

(b) $\mathbf{v} \in \mathbb{R}^2$ such that $\langle \mathbf{v}, \mathbf{e}_1 \rangle = 0$.

(c) $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ such that $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$ and $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

A very useful and a fundamental inequality, commonly called the Cauchy-Schwartz inequality, concerning the inner product is proved next.

Theorem 5.1.8 (Cauchy-Bunyakovskii-Schwartz inequality)**.** *Let $\mathbb{V}$ be an inner product space over $\mathbb{F}$. Then, for any $\mathbf{u}, \mathbf{v} \in \mathbb{V}$*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \le \|\mathbf{u}\| \, \|\mathbf{v}\|. \tag{5.1.1}$$

Moreover, equality holds in Inequality (5.1.1) if and only if $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent. Furthermore, if $\mathbf{u} \neq \mathbf{0}$ then $\mathbf{v} = \left\langle \mathbf{v}, \dfrac{\mathbf{u}}{\|\mathbf{u}\|} \right\rangle \dfrac{\mathbf{u}}{\|\mathbf{u}\|}$.

Proof. If $\mathbf{u} = \mathbf{0}$ then Inequality (5.1.1) holds. Hence, let $\mathbf{u} \neq \mathbf{0}$. Then, by Definition 5.1.1.3, $\langle \lambda\mathbf{u}+\mathbf{v}, \lambda\mathbf{u}+\mathbf{v}\rangle \geq 0$ for all $\lambda \in \mathbb{F}$ and $\mathbf{v} \in \mathbb{V}$. In particular, for $\lambda = -\dfrac{\langle \mathbf{v},\mathbf{u}\rangle}{\|\mathbf{u}\|^2}$,

$$\begin{aligned} 0 &\leq \langle \lambda\mathbf{u}+\mathbf{v}, \lambda\mathbf{u}+\mathbf{v}\rangle = \lambda\overline{\lambda}\|\mathbf{u}\|^2 + \lambda\langle \mathbf{u},\mathbf{v}\rangle + \overline{\lambda}\langle \mathbf{v},\mathbf{u}\rangle + \|\mathbf{v}\|^2 \\ &= \frac{\langle \mathbf{v},\mathbf{u}\rangle}{\|\mathbf{u}\|^2}\frac{\overline{\langle \mathbf{v},\mathbf{u}\rangle}}{\|\mathbf{u}\|^2}\|\mathbf{u}\|^2 - \frac{\langle \mathbf{v},\mathbf{u}\rangle}{\|\mathbf{u}\|^2}\langle \mathbf{u},\mathbf{v}\rangle - \frac{\overline{\langle \mathbf{v},\mathbf{u}\rangle}}{\|\mathbf{u}\|^2}\langle \mathbf{v},\mathbf{u}\rangle + \|\mathbf{v}\|^2 = \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v},\mathbf{u}\rangle|^2}{\|\mathbf{u}\|^2}. \end{aligned}$$

Or, in other words $|\langle \mathbf{v},\mathbf{u}\rangle|^2 \leq \|\mathbf{u}\|^2\|\mathbf{v}\|^2$ and the proof of the inequality is over.

Now, note that equality holds in Inequality (5.1.1) if and only if $\langle \lambda\mathbf{u}+\mathbf{v}, \lambda\mathbf{u}+\mathbf{v}\rangle = 0$, or equivalently, $\lambda\mathbf{u}+\mathbf{v} = \mathbf{0}$. Hence, $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent. Moreover,

$$0 = \langle \mathbf{0},\mathbf{u}\rangle = \langle \lambda\mathbf{u}+\mathbf{v},\mathbf{u}\rangle = \lambda\langle \mathbf{u},\mathbf{u}\rangle + \langle \mathbf{v},\mathbf{u}\rangle$$

implies that $\mathbf{v} = -\lambda\mathbf{u} = -\dfrac{\langle \mathbf{v},\mathbf{u}\rangle}{\|\mathbf{u}\|^2}\mathbf{u} = \left\langle \mathbf{v}, \dfrac{\mathbf{u}}{\|\mathbf{u}\|}\right\rangle \dfrac{\mathbf{u}}{\|\mathbf{u}\|}$. ■

Corollary 5.1.9. *Let* $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$. *Then,* $\left(\sum\limits_{i=1}^{n} \mathbf{x}_i\mathbf{y}_i\right)^2 \leq \left(\sum\limits_{i=1}^{n} \mathbf{x}_i^2\right)\left(\sum\limits_{i=1}^{n} \mathbf{y}_i^2\right)$.

5.1.2 Angle between two Vectors

Let $\mathbb{V}$ be a real vector space. Then, for $\mathbf{u},\mathbf{v} \in \mathbb{V}$, the Cauchy-Schwartz inequality implies that $-1 \leq \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\ \|\mathbf{v}\|} \leq 1$. We use this together with the properties of the cosine function to define the angle between two vectors in an inner product space.

Definition 5.1.10. **[Angle between Vectors]** Let $\mathbb{V}$ be a real vector space. If $\theta \in [0,\ \pi]$ is the angle between $\mathbf{u},\mathbf{v} \in \mathbb{V} \setminus \{\mathbf{0}\}$ then we define

$$\cos\theta = \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\ \|\mathbf{v}\|}.$$

Example 5.1.11.

1. Take $(1,0)^T, (1,1)^T \in \mathbb{R}^2$. Then, $\cos\theta = \frac{1}{\sqrt{2}}$. So $\theta = \pi/4$.
2. Take $(1,1,0)^T, (1,1,1)^T \in \mathbb{R}^3$. Then, angle between them, say $\beta = \cos^{-1}\frac{2}{\sqrt{6}}$.
3. Angle depends on the IP. Take $\langle \mathbf{x},\mathbf{y}\rangle = 2\mathbf{x}_1\mathbf{y}_1 + \mathbf{x}_1\mathbf{y}_2 + \mathbf{x}_2\mathbf{y}_1 + \mathbf{x}_2\mathbf{y}_2$ on $\mathbb{R}^2$. Then, angle between $(1,0)^T, (1,1)^T \in \mathbb{R}^2$ equals $\cos^{-1}\frac{3}{\sqrt{10}}$.
4. As $\langle \mathbf{x},\mathbf{y}\rangle = \langle \mathbf{y},\mathbf{x}\rangle$ for any real vector space, the angle between $\mathbf{x}$ and $\mathbf{y}$ is same as the angle between $\mathbf{y}$ and $\mathbf{x}$.
5. Let $a,b \in \mathbb{R}$ with $a,b > 0$. Then, prove that $(a+b)\left(\dfrac{1}{a}+\dfrac{1}{b}\right) \geq 4$.
6. For $1 \leq i \leq n$, let $a_i \in \mathbb{R}$ with $a_i > 0$. Then, use Corollary 5.1.9 to show that $\left(\sum\limits_{i=1}^{n} a_i\right)\left(\sum\limits_{i=1}^{n} \dfrac{1}{a_i}\right) \geq n^2$.
7. Prove that $|z_1 + \cdots + z_n| \leq \sqrt{n(|z_1|^2 + \cdots + |z_n|^2)}$, for $z_1,\ldots,z_n \in \mathbb{C}$. When does the equality hold?
8. Let $\mathbb{V}$ be an IPS. If $\mathbf{u},\mathbf{v} \in \mathbb{V}$ with $\|\mathbf{u}\| = 1, \|\mathbf{v}\| = 1$ and $\langle \mathbf{u},\mathbf{v}\rangle = 1$ then prove that $\mathbf{u} = \alpha\mathbf{v}$ for some $\alpha \in \mathbb{F}$. Is $\alpha = 1$?

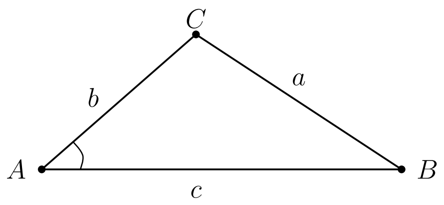


Figure 5.1: Triangle with vertices A, B and C

We will now prove that if A, B and C are the vertices of a triangle (see Figure 5.1) and a, b and c, respectively, are the lengths of the corresponding sides then $\cos(A) = \frac{b^2+c^2-a^2}{2bc}$. This in turn implies that the angle between vectors has been rightly defined.

Lemma 5.1.12. *Let A, B and C be the vertices of a triangle (see Figure 5.1) with corresponding side lengths a, b and c, respectively, in a real inner product space $\mathbb{V}$ then*

$$\cos(A) = \frac{b^2 + c^2 - a^2}{2bc}.$$

Proof. Let $\mathbf{0}$, $\mathbf{u}$ and $\mathbf{v}$ be the coordinates of the vertices A, B and C, respectively, of the triangle ABC. Then, $\vec{AB} = \mathbf{u}, \vec{AC} = \mathbf{v}$ and $\vec{BC} = \mathbf{v} - \mathbf{u}$. Thus, we need to prove that

$$\cos(A) = \frac{\|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 - \|\mathbf{v}-\mathbf{u}\|^2}{2\|\mathbf{v}\|\|\mathbf{u}\|} \Leftrightarrow \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 - \|\mathbf{v}-\mathbf{u}\|^2 = 2\,\|\mathbf{v}\|\,\|\mathbf{u}\|\,\cos(A).$$

Now, by definition $\|\mathbf{v}-\mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 - 2\langle \mathbf{v}, \mathbf{u}\rangle$ and hence $\|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 - \|\mathbf{v}-\mathbf{u}\|^2 = 2\,\langle \mathbf{u}, \mathbf{v}\rangle$. As $\langle \mathbf{v}, \mathbf{u}\rangle = \|\mathbf{v}\|\,\|\mathbf{u}\|\,\cos(A)$, the required result follows. ■

Definition 5.1.13. **[Orthogonality / Perpendicularity]** Let $\mathbb{V}$ be an inner product space over $\mathbb{R}$. Then,

1. the vectors $\mathbf{u}, \mathbf{v} \in \mathbb{V}$ are called **orthogonal/perpendicular** if $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.
2. Let $S \subseteq \mathbb{V}$. Then, the **orthogonal complement** of S in $\mathbb{V}$, denoted $S^\perp$, equals

$$S^\perp = \{\mathbf{v} \in \mathbb{V} : \langle \mathbf{v}, \mathbf{w}\rangle = 0, \text{ for all } \mathbf{w} \in S\}.$$

Example 5.1.14. 1. $\mathbf{0}$ is orthogonal to every vector as $\langle \mathbf{0}, \mathbf{x}\rangle = 0$ for all $\mathbf{x} \in \mathbb{V}$.

2. If $\mathbb{V}$ is a vector space over $\mathbb{R}$ or $\mathbb{C}$ then $\mathbf{0}$ is the only vector that is orthogonal to itself.

3. Let $\mathbb{V} = \mathbb{R}$.

   (a) $S = \{0\}$. Then, $S^\perp = \mathbb{R}$.

   (b) $S = \mathbb{R}$, Then, $S^\perp = \{0\}$.

   (c) Let S be any subset of $\mathbb{R}$ containing a nonzero real number. Then, $S^\perp = \{0\}$.

4. Let $\mathbf{u} = (1, 2)^T$. What is $\mathbf{u}^\perp$ in $\mathbb{R}^2$?
   Solution: $\{(x, y)^T \in \mathbb{R}^2 \mid x + 2y = 0\}$. Is this $\textsc{Null}(\mathbf{u})$? Note that $(2, -1)^T$ is a basis of $\mathbf{u}^\perp$ and for any vector $\mathbf{x} \in \mathbb{R}^2$,

$$\mathbf{x} = \langle \mathbf{x}, \mathbf{u}\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|^2} + \left(\mathbf{x} - \langle \mathbf{x}, \mathbf{u}\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|^2}\right) = \frac{x_1 + 2x_2}{5}(1, 2)^T + \frac{2x_1 - x_2}{5}(2, -1)^T$$

   is a decomposition of $\mathbf{x}$ into two vectors, one parallel to $\mathbf{u}$ and the other parallel to $\mathbf{u}^\perp$.

5. Fix $\mathbf{u} = (1,1,1,1)^T, \mathbf{v} = (1,1,-1,0)^T \in \mathbb{R}^4$. Determine $\mathbf{z}, \mathbf{w} \in \mathbb{R}^4$ such that $\mathbf{u} = \mathbf{z} + \mathbf{w}$ with the condition that $\mathbf{z}$ is parallel to $\mathbf{v}$ and $\mathbf{w}$ is orthogonal to $\mathbf{v}$.

   Solution: As $\mathbf{z}$ is parallel to $\mathbf{v}$, $\mathbf{z} = k\mathbf{v} = (k,k,-k,0)^T$, for some $k \in \mathbb{R}$. Since $\mathbf{w}$ is orthogonal to $\mathbf{v}$ the vector $\mathbf{w} = (a,b,c,d)^T$ satisfies $a+b-c=0$. Thus, $c = a+b$ and

   $$(1,1,1,1)^T = \mathbf{u} = \mathbf{z} + \mathbf{w} = (k,k,-k,0)^T + (a,b,a+b,d)^T.$$

   Comparing the corresponding coordinates, gives the linear system $d = 1$, $a+k = 1$, $b+k = 1$ and $a+b-k = 1$ in the variables a,b,d and k. Thus, solving for a,b,d and k gives $\mathbf{z} = \dfrac{1}{3}(1,1,-1,0)^T$ and $\mathbf{w} = \dfrac{1}{3}(2,2,4,3)^T$.

6. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then prove that

   (a) $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \Longleftrightarrow \|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ (Pythagoras Theorem).
   Solution: Use $\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle$ to get the required result follows.

   (b) $\|\mathbf{x}\| = \|\mathbf{y}\| \Longleftrightarrow \langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle = 0$ ($\mathbf{x}$ and $\mathbf{y}$ form adjacent sides of a rhombus as the diagonals $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ are orthogonal).
   Solution: Use $\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle = \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2$ to get the required result follows.

   (c) $4\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2$ (polarization identity in $\mathbb{R}^n$).
   Solution: Just expand the right hand side to get the required result follows.

   (d) $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$ (parallelogram law: the sum of squares of the diagonals of a parallelogram equals twice the sum of squares of its sides).
   Solution: Just expand the left hand side to get the required result follows.

7. Let $P = (1,1,1)^T$, $Q = (2,1,3)^T$ and $R = (-1,1,2)^T$ be three vertices of a triangle in $\mathbb{R}^3$. Compute the angle between the sides PQ and PR.

   Solution: Method 1: Note that $\vec{PQ} = (2,1,3)^T - (1,1,1)^T = (1,0,2)^T$, $\vec{PR} = (-2,0,1)^T$ and $\vec{RQ} = (-3,0,-1)^T$. As $\langle \vec{PQ}, \vec{PR} \rangle = 0$, the angle between the sides PQ and PR is $\dfrac{\pi}{2}$.

   Method 2: $\|PQ\| = \sqrt{5}, \|PR\| = \sqrt{5}$ and $\|QR\| = \sqrt{10}$. As $\|QR\|^2 = \|PQ\|^2 + \|PR\|^2$, by Pythagoras theorem, the angle between the sides PQ and PR is $\dfrac{\pi}{2}$.

Exercise **5.1.15.** *1. Let $\mathbb{V}$ be an* ips.

   (a) If $S \subseteq \mathbb{V}$ then $S^\perp$ is a subspace of $\mathbb{V}$ and $S^\perp = (LS(S))^\perp$.

   (b) Furthermore, if $\mathbb{V}$ is finite dimensional then $S^\perp$ and $LS(S)$ are complementary. That is, $\mathbb{V} = LS(S) + S^\perp$. Equivalently, $\langle \mathbf{u}, \mathbf{w} \rangle = 0$, for all $\mathbf{u} \in LS(S)$ and $\mathbf{w} \in S^\perp$.

2. *Consider $\mathbb{R}^3$ with the standard inner product. Find*

   (a) $S^\perp$ for $S = \{(1,1,1)^T, (0,1,-1)^T\}$ and $S = LS((1,1,1)^T, (0,1,-1)^T)$.

   (b) vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ such that $\mathbf{v}, \mathbf{w}, \mathbf{u} = (1,1,1)^T$ are mutually orthogonal.

   (c) the line passing through $(1,1,-1)^T$ and parallel to $(a,b,c) \neq \mathbf{0}$.

   (d) the plane containing $(1,1-1)$ with $(a,b,c) \neq \mathbf{0}$ as the normal vector.

   (e) the area of the parallelogram with three vertices $\mathbf{0}^T$, $(1,2,-2)^T$ and $(2,3,0)^T$.

   (f) the area of the parallelogram when $\|\mathbf{x}\| = 5, \|\mathbf{x}-\mathbf{y}\| = 8$ and $\|\mathbf{x}+\mathbf{y}\| = 14$.

(g) *the plane containing* $(2,-2,1)^T$ *and perpendicular to the line with parametric equation* $x = t-1, y = 3t+2, z = t+1$.

(h) *the plane containing the lines* $(1,2,-2)+t(1,1,0)$ *and* $(1,2,-2)+t(0,1,2)$.

(i) k *such that* $\cos^{-1}(\langle \mathbf{u}, \mathbf{v}\rangle) = \pi/3$, *where* $\mathbf{u} = (1,-1,1)^T$ *and* $\mathbf{v} = (1,k,1)^T$.

(j) *the plane containing* $(1,1,2)^T$ *and orthogonal to the line with parametric equation* $x = 2+t, y = 3$ *and* $z = 1-t$.

(k) *a parametric equation of a line containing* $(1,-2,1)^T$ *and orthogonal to* $x+3y+2z = 1$.

3. *Let* $P = (3,0,2)^T, Q = (1,2,-1)^T$ *and* $R = (2,-1,1)^T$ *be three points in* $\mathbb{R}^3$. *Then,*

   (a) *find the area of the triangle with vertices* P, Q *and* R.

   (b) *find the area of the parallelogram built on vectors* $\vec{PQ}$ *and* $\vec{QR}$.

   (c) *find a nonzero vector orthogonal to the plane of the above triangle.*

   (d) *find all vectors* $\mathbf{x}$ *orthogonal to* $\vec{PQ}$ *and* $\vec{QR}$ *with* $\|\mathbf{x}\| = \sqrt{2}$.

   (e) *the volume of the parallelepiped built on vectors* $\vec{PQ}$ *and* $\vec{QR}$ *and* $\mathbf{x}$, *where* $\mathbf{x}$ *is one of the vectors found in Part 3d. Do you think the volume would be different if you choose the other vector* $\mathbf{x}$*?*

4. *Let* p_1 *be a plane containing* $A = (1,2,3)^T$ *and* $(2,-1,1)^T$ *as its normal vector. Then,*

   (a) *find the equation of the plane* p_2 *that is parallel to* p_1 *and contains* $(-1,2,-3)^T$.

   (b) *calculate the distance between the planes* p_1 *and* p_2.

5. *In the parallelogram* $ABCD$, $AB\|DC$ *and* $AD\|BC$ *and* $A = (-2,1,3)^T, B = (-1,2,2)^T$ *and* $C = (-3,1,5)^T$. *Find the*

   (a) *coordinates of the point* D,

   (b) *cosine of the angle* BCD.

   (c) *area of the triangle* ABC

   (d) *volume of the parallelepiped determined by* AB, AD *and* $(0,0,-7)^T$.

6. *Let* $\mathbb{W} = \{(x,y,z,w)^T \in \mathbb{R}^4 : x+y+z-w = 0\}$. *Find a basis of* $\mathbb{W}^\perp$.

7. *Recall the* IPS $\mathbb{M}_n(\mathbb{R})$ *(see Example 5.1.4.8). If* $\mathbb{W} = \{A \in \mathbb{M}_n(\mathbb{R}) \mid A^T = A\}$ *then* $\mathbb{W}^\perp$*?*

5.1.3 Normed Linear Space

To proceed further, recall that a vector space over $\mathbb{R}$ or $\mathbb{C}$ was a linear space.

Definition 5.1.16. [Normed Linear Space] Let $\mathbb{V}$ be a linear space.

1. Then, a **norm** on $\mathbb{V}$ is a function $f(\mathbf{x}) = \|\mathbf{x}\|$ from $\mathbb{V}$ to $\mathbb{R}$ such that

   (a) $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in \mathbb{V}$ and if $\|\mathbf{x}\| = 0$ then $\mathbf{x} = \mathbf{0}$.

   (b) $\|\alpha\mathbf{x}\| = |\alpha| \; \|\mathbf{x}\|$ for all $\alpha \in \mathbb{F}$ and $\mathbf{x} \in \mathbb{V}$.

   (c) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ (triangle inequality).

2. A linear space with a norm on it is called a **normed linear space** (NLS).

Theorem 5.1.17. *Let $\mathbb{V}$ be a normed linear space and $\mathbf{x}, \mathbf{y} \in \mathbb{V}$. Then, $\big| \|\mathbf{x}\| - \|\mathbf{y}\| \big| \leq \|\mathbf{x} - \mathbf{y}\|$.*

Proof. As $\|\mathbf{x}\| = \|\mathbf{x} - \mathbf{y} + \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\|$ one has $\|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$. Similarly, one obtains $\|\mathbf{y}\| - \|\mathbf{x}\| \leq \|\mathbf{y} - \mathbf{x}\| = \|\mathbf{x} - \mathbf{y}\|$. Combining the two, the required result follows. ■

Example 5.1.18.
1. On $\mathbb{R}^3$, $\|\mathbf{x}\| = \sqrt{\mathbf{x}_1^2 + \mathbf{x}_2^2 + \mathbf{x}_3^2}$ is a norm. Also, observe that this norm corresponds to $\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$, where $\langle , \rangle$ is the standard inner product.
2. Let $\mathbb{V}$ be an IPS. Is it true that $f(\mathbf{x}) = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ is a norm?
   Solution: Yes. The readers should verify the first two conditions. For the third condition, recalling the Cauchy-Schwartz inequality, we get

$$\begin{aligned} f(\mathbf{x}+\mathbf{y})^2 &= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (f(\mathbf{x}) + f(\mathbf{y}))^2. \end{aligned}$$

   Thus, $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ is a norm, called the norm **induced** by the inner product $\langle \cdot, \cdot \rangle$.

EXERCISE **5.1.19.**
1. *Let $\mathbb{V}$ be an IPS. Then,*

$$4\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 + i\|\mathbf{x}+i\mathbf{y}\|^2 - i\|\mathbf{x}-i\mathbf{y}\|^2 \quad (\textbf{Polarization Identity}).$$

2. *Consider the complex vector space $\mathbb{C}^n$. If $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ then prove that*
   (a) *If $\mathbf{x} \neq \mathbf{0}$ then $\|\mathbf{x} + i\mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|i\mathbf{x}\|^2$, even though $\langle \mathbf{x}, i\mathbf{x} \rangle \neq 0$.*
   (b) *$\langle \mathbf{x}, \mathbf{y} \rangle = 0$ whenever $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ and $\|\mathbf{x}+i\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|i\mathbf{y}\|^2$.*
3. *Let $A \in \mathbb{M}_n(\mathbb{C})$ satisfy $\|A\mathbf{x}\| \leq \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{C}^n$. Then, prove that if $\alpha \in \mathbb{C}$ with $|\alpha| > 1$ then $A - \alpha I$ is invertible.*

The next result is stated without proof as the proof is beyond the scope of this book.

Theorem 5.1.20. *Let $\|\cdot\|$ be a norm on a NLS $\mathbb{V}$. Then, $\|\cdot\|$ is induced by some inner product if and only if $\|\cdot\|$ satisfies the PARALLELOGRAM LAW: $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$.*

Example 5.1.21.
1. For $\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$, we define $\|\mathbf{x}\|_1 = |\mathbf{x}_1| + |\mathbf{x}_2|$. Verify that $\|\mathbf{x}\|_1$ is indeed a norm. But, for $\mathbf{x} = \mathbf{e}_1$ and $\mathbf{y} = \mathbf{e}_2$, $2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) = 4$ whereas

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = \|(1,1)\|^2 + \|(1,-1)\|^2 = (|1| + |1|)^2 + (|1| + |-1|)^2 = 8.$$

   So, the parallelogram law fails. Thus, $\|\mathbf{x}\|_1$ is not induced by any inner product in $\mathbb{R}^2$.
2. Does there exist an inner product in $\mathbb{R}^2$ such that $\|\mathbf{x}\| = \max\{|x_1|, |x_2|\}$?
3. If $\|\cdot\|$ is a norm in $\mathbb{V}$ then $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, for $\mathbf{x}, \mathbf{y} \in \mathbb{V}$, defines a distance function as
   (a) $d(\mathbf{x}, \mathbf{x}) = 0$, for each $\mathbf{x} \in \mathbb{V}$.
   (b) using the triangle inequality, for any $\mathbf{z} \in \mathbb{V}$, we have

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \|(\mathbf{x}-\mathbf{z}) - (\mathbf{y}-\mathbf{z})\| \leq \|(\mathbf{x}-\mathbf{z})\| + \|(\mathbf{y}-\mathbf{z})\| = d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}).$$

5.2 Gram-Schmidt Orthonormalization Process

We start with the definition of an orthonormal set.

Definition 5.2.1. Let $\mathbb{V}$ be an IPS. Then, a non-empty set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \mathbb{V}$ is called an **orthogonal set** if $\mathbf{v}_i$ and $\mathbf{v}_j$ are **mutually orthogonal**, for $1 \le i \ne j \le n$, *i.e.*,

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0, \text{ for } 1 \le i < j \le n.$$

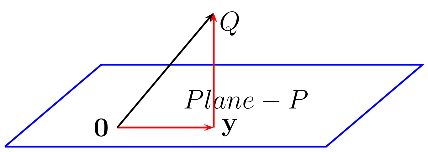
Further, if $\|\mathbf{v}_i\| = 1$, for $1 \le i \le n$, Then S is called an **orthonormal set**. If S is also a basis of $\mathbb{V}$ then S is called an **orthonormal basis** of $\mathbb{V}$.

Example 5.2.2.
1. A few orthonormal sets in $\mathbb{R}^2$ are
$$\{(1,0)^T, (0,1)^T\}, \{\frac{1}{\sqrt{2}}(1,1)^T, \frac{1}{\sqrt{2}}(1,-1)^T\} \text{ and } \{\frac{1}{\sqrt{5}}(2,1)^T, \frac{1}{\sqrt{5}}(1,-2)^T\}.$$
2. Let $S = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{R}^n$. Then, S is an orthonormal set as
   (a) $\|\mathbf{e}_i\| = 1$, for $1 \le i \le n$.
   (b) $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$, for $1 \le i \ne j \le n$.
3. The set $\left\{ \left[\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right]^T, \left[0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]^T, \left[\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right]^T \right\}$ is an orthonormal set in $\mathbb{R}^3$.
4. Recall that $\langle f(x), g(x) \rangle = \int\limits_{-\pi}^{\pi} f(x)\overline{g(x)}dx$ defines the standard inner product in $\mathcal{C}[-\pi, \pi]$. Consider $S = \{\mathbf{1}\} \cup \{\mathbf{e}_m \mid m \ge 1\} \cup \{\mathbf{f}_n \mid n \ge 1\}$, where $\mathbf{1}(x) = 1$, $\mathbf{e}_m(x) = \cos(mx)$ and $\mathbf{f}_n(x) = \sin(nx)$, for all $m, n \ge 1$ and for all $x \in [-\pi, \pi]$. Then,
   (a) S is a linearly independent set.
   (b) $\|\mathbf{1}\|^2 = 2\pi$, $\|\mathbf{e}_m\|^2 = \pi$ and $\|\mathbf{f}_n\|^2 = \pi$.
   (c) the functions in S are orthogonal.

   Hence, $\left\{\frac{\mathbf{1}}{\sqrt{2\pi}}\right\} \cup \left\{\frac{1}{\sqrt{\pi}}\mathbf{e}_m \mid m \ge 1\right\} \cup \left\{\frac{1}{\sqrt{\pi}}\mathbf{f}_n \mid n \ge 1\right\}$ is an orthonormal set in $\mathcal{C}[-\pi, \pi]$.

To proceed further, we consider a few examples for better understanding.

Example 5.2.3. Which point on the plane P is closest to the point, say Q?



Solution: Let $\mathbf{y}$ be the foot of the perpendicular from Q on P. Thus, by Pythagoras Theorem, this point is unique. So, the question arises: how do we find $\mathbf{y}$?

Note that $\overrightarrow{\mathbf{y}Q}$ gives a normal vector of the plane P. Hence, $\overrightarrow{Q} = \mathbf{y} + \overrightarrow{\mathbf{y}Q}$. So, need to decompose $\overrightarrow{Q}$ into two vectors such that one of them lies on the plane P and the other is orthogonal to the plane.

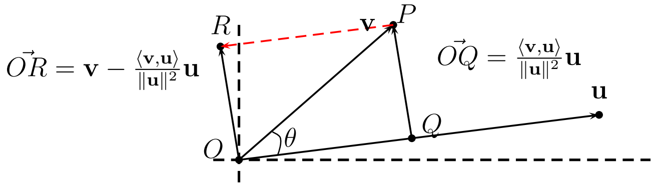


Figure 5.2: Decomposition of vector $\mathbf{v}$

Thus, we see that given $\mathbf{u}, \mathbf{v} \in \mathbb{V} \setminus \{\mathbf{0}\}$, we need to find two vectors, say $\mathbf{y}$ and $\mathbf{z}$, such that $\mathbf{y}$ is parallel to $\mathbf{u}$ and $\mathbf{z}$ is perpendicular to $\mathbf{u}$. Thus, $\mathbf{y} = \mathbf{u}\cos(\theta)$ and $\mathbf{z} = \mathbf{u}\sin(\theta)$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

We do this as follows (see Figure 5.2). Let $\hat{\mathbf{u}} = \dfrac{\mathbf{u}}{\|\mathbf{u}\|}$ be the unit vector in the direction of $\mathbf{u}$. Then, using trigonometry, $\cos(\theta) = \frac{\|\vec{OQ}\|}{\|\vec{OP}\|}$. Hence $\|\vec{OQ}\| = \|\vec{OP}\| \cos(\theta)$. Now using Definition 5.1.10, $\|\vec{OQ}\| = \|\mathbf{v}\| \left|\frac{\langle \mathbf{v},\mathbf{u}\rangle}{\|\mathbf{v}\|\,\|\mathbf{u}\|}\right| = \left|\frac{\langle \mathbf{v},\mathbf{u}\rangle}{\|\mathbf{u}\|}\right|$, where the absolute value is taken as the length/norm is a positive quantity. Thus,

$$\vec{OQ} = \|\vec{OQ}\|\,\hat{\mathbf{u}} = \left\langle \mathbf{v}, \frac{\mathbf{u}}{\|\mathbf{u}\|}\right\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Hence, $\mathbf{y} = \vec{OQ} = \left\langle \mathbf{v}, \frac{\mathbf{u}}{\|\mathbf{u}\|}\right\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|}$ and $\mathbf{z} = \mathbf{v} - \left\langle \mathbf{v}, \frac{\mathbf{u}}{\|\mathbf{u}\|}\right\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|}$. In literature, the vector $\mathbf{y} = \vec{OQ}$ is called the **orthogonal projection** of $\mathbf{v}$ on $\mathbf{u}$, denoted $\mathrm{Proj}_{\mathbf{u}}(\mathbf{v})$. Thus,

$$\mathrm{Proj}_{\mathbf{u}}(\mathbf{v}) = \left\langle \mathbf{v}, \frac{\mathbf{u}}{\|\mathbf{u}\|}\right\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|} \text{ and } \|\mathrm{Proj}_{\mathbf{u}}(\mathbf{v})\| = \|\vec{OQ}\| = \left|\frac{\langle \mathbf{v},\mathbf{u}\rangle}{\|\mathbf{u}\|}\right|. \tag{5.2.1}$$

Moreover, the distance of $\mathbf{u}$ from the point P equals $\|\vec{OR}\| = \|\vec{PQ}\| = \left\|\mathbf{v} - \langle \mathbf{v}, \frac{\mathbf{u}}{\|\mathbf{u}\|}\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|}\right\|$.

Example 5.2.4.

1. Determine the foot of the perpendicular from the point $(1,2,3)$ on the XY-plane.

   Solution: Verify that the required point is $(1,2,0)$?

2. Determine the foot of the perpendicular from the point $Q = (1,2,3,4)$ on the plane generated by $(1,1,0,0), (1,0,1,0)$ and $(0,1,1,1)$.

   Answer: (x,y,z,w) lies on the plane $x-y-z+2w = 0 \Leftrightarrow \langle (1,-1,-1,2),(x,y,z,w)\rangle = 0$.

   So, the required point equals

$$\begin{aligned}(1,2,3,4) - \langle (1,2,3,4), \frac{1}{\sqrt{7}}(1,-1,-1,2)\rangle \frac{1}{\sqrt{7}}(1,-1,-1,2) \\ = \quad (1,2,3,4) - \frac{4}{7}(1,-1,-1,2) = \frac{1}{7}(3,18,25,20).\end{aligned}$$

3. Determine the projection of $\mathbf{v} = (1,1,1,1)^T$ on $\mathbf{u} = (1,1,-1,0)^T$.

   Solution: By Equation (5.2.1), we have $\mathrm{Proj}_{\mathbf{v}}(\mathbf{u}) = \langle \mathbf{v},\mathbf{u}\rangle \dfrac{\mathbf{u}}{\|\mathbf{u}\|^2} = \frac{1}{3}(1,1,-1,0)^T$ and $\mathbf{w} = (1,1,1,1)^T - \mathrm{Proj}_{\mathbf{v}}(\mathbf{u}) = \frac{1}{3}(2,2,4,3)^T$ is orthogonal to $\mathbf{u}$.

4. Let $\mathbf{u} = (1,1,1,1)^T, \mathbf{v} = (1,1,-1,0)^T, \mathbf{w} = (1,1,0,-1)^T \in \mathbb{R}^4$. Write $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, where $\mathbf{v}_1$ is parallel to $\mathbf{u}$ and $\mathbf{v}_2$ is orthogonal to $\mathbf{u}$. Also, write $\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3$ such that $\mathbf{w}_1$ is parallel to $\mathbf{u}$, $\mathbf{w}_2$ is parallel to $\mathbf{v}_2$ and $\mathbf{w}_3$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}_2$.

   Solution: Note that

   (a) $\mathbf{v}_1 = \mathrm{Proj}_{\mathbf{u}}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|^2} = \frac{1}{4}\mathbf{u} = \frac{1}{4}(1,1,1,1)^T$ is parallel to $\mathbf{u}$.

   (b) $\mathbf{v}_2 = \mathbf{v} - \frac{1}{4}\mathbf{u} = \frac{1}{4}(3,3,-5,-1)^T$ is orthogonal to $\mathbf{u}$.

   Note that $\mathrm{Proj}_{\mathbf{u}}(\mathbf{w})$ is parallel to $\mathbf{u}$ and $\mathrm{Proj}_{\mathbf{v}_2}(\mathbf{w})$ is parallel to $\mathbf{v}_2$. Hence, we have

   (a) $\mathbf{w}_1 = \mathrm{Proj}_{\mathbf{u}}(\mathbf{w}) = \langle \mathbf{w}, \mathbf{u}\rangle \frac{\mathbf{u}}{\|\mathbf{u}\|^2} = \frac{1}{4}\mathbf{u} = \frac{1}{4}(1,1,1,1)^T$ is parallel to $\mathbf{u}$,

   (b) $\mathbf{w}_2 = \mathrm{Proj}_{\mathbf{v}_2}(\mathbf{w}) = \langle \mathbf{w}, \mathbf{v}_2\rangle \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|^2} = \frac{7}{44}(3,3,-5,-1)^T$ is parallel to $\mathbf{v}_2$ and

   (c) $\mathbf{w}_3 = \mathbf{w} - \mathbf{w}_1 - \mathbf{w}_2 = \frac{3}{11}(1,1,2,-4)^T$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}_2$.

We now prove the most important initial result of this section.

Theorem 5.2.5. *Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be an orthonormal subset of an* IPS $\mathbb{V}(\mathbb{F})$.

1. *Then, S is a linearly independent subset of $\mathbb{V}$.*

2. *Suppose $\mathbf{v} \in LS(S)$ with $\mathbf{v} = \sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i$, for some α_i's in $\mathbb{F}$. Then,*

   (a) $\alpha_i = \langle \mathbf{v}, \mathbf{u}_i\rangle$.

   (b) $\|\mathbf{v}\|^2 = \|\sum\limits_{i=1}^{n} \alpha_i \mathbf{u}_i\|^2 = \sum\limits_{i=1}^{n} \mid \alpha_i \mid^2$.

3. *Let $\mathbf{z} \in \mathbb{V}$ and $\mathbf{w} = \sum\limits_{i=1}^{n} \langle \mathbf{z}, \mathbf{u}_i\rangle \mathbf{u}_i$. Then, $\mathbf{z} = \mathbf{w} + (\mathbf{z} - \mathbf{w})$ with $\langle \mathbf{z} - \mathbf{w}, \mathbf{w}\rangle = 0$, i.e., $\mathbf{z} - \mathbf{w} \in LS(S)^{\perp}$. Further, $\|\mathbf{z}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{z} - \mathbf{w}\|^2 \geq \|\mathbf{w}\|^2$.*

4. *Let $\dim(\mathbb{V}) = n$. Then, $\langle \mathbf{v}, \mathbf{u}_i\rangle = 0$ for all $i = 1, 2, \ldots, n$ if and only if $\mathbf{v} = \mathbf{0}$.*

Proof. Part 1: Consider the linear system $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$ in the variables $c_1, \ldots, c_n$. As $\langle \mathbf{0}, \mathbf{u}\rangle = 0$ and $\langle \mathbf{u}_j, \mathbf{u}_i\rangle = 0$, for all $j \neq i$, we have

$$0 = \langle \mathbf{0}, \mathbf{u}_i\rangle = \langle c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n, \mathbf{u}_i\rangle = \sum_{j=1}^{n} c_j \langle \mathbf{u}_j, \mathbf{u}_i\rangle = c_i \langle \mathbf{u}_i, \mathbf{u}_i\rangle = c_i.$$

Hence, $c_i = 0$, for $1 \leq i \leq n$. Thus, the above linear system has only the trivial solution. So, the set S is linearly independent.

Part 2: Note that $\langle \mathbf{v}, \mathbf{u}_i\rangle = \langle \sum\limits_{j=1}^{n} \alpha_j \mathbf{u}_j, \mathbf{u}_i\rangle = \sum_{j=1}^{n} \alpha_j \langle \mathbf{u}_j, \mathbf{u}_i\rangle = \alpha_i \langle \mathbf{u}_i, \mathbf{u}_i\rangle = \alpha_i$. This completes the first sub-part. For the second sub-part, we have

$$\begin{aligned} \|\sum_{i=1}^{n} \alpha_i \mathbf{u}_i\|^2 &= \left\langle \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \sum_{i=1}^{n} \alpha_i \mathbf{u}_i \right\rangle = \sum_{i=1}^{n} \alpha_i \left\langle \mathbf{u}_i, \sum_{j=1}^{n} \alpha_j \mathbf{u}_j \right\rangle \\ &= \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \overline{\alpha_j} \langle \mathbf{u}_i, \mathbf{u}_j\rangle = \sum_{i=1}^{n} \alpha_i \overline{\alpha_i} \langle \mathbf{u}_i, \mathbf{u}_i\rangle = \sum_{i=1}^{n} \mid \alpha_i \mid^2. \end{aligned}$$

Part 3: Note that for $1 \le i \le n$,

$$\begin{aligned}\langle \mathbf{z}-\mathbf{w}, \mathbf{u}_i\rangle &= \langle \mathbf{z}, \mathbf{u}_i\rangle - \langle \mathbf{w}, \mathbf{u}_i\rangle = \langle \mathbf{z}, \mathbf{u}_i\rangle - \left\langle \sum_{j=1}^{n} \langle \mathbf{z}, \mathbf{u}_j\rangle \mathbf{u}_j, \mathbf{u}_i \right\rangle \\ &= \langle \mathbf{z}, \mathbf{u}_i\rangle - \sum_{j=1}^{n} \langle \mathbf{z}, \mathbf{u}_j\rangle \langle \mathbf{u}_j, \mathbf{u}_i\rangle = \langle \mathbf{z}, \mathbf{u}_i\rangle - \langle \mathbf{z}, \mathbf{u}_i\rangle = 0.\end{aligned}$$

So, $\mathbf{z}-\mathbf{w} \in \mathrm{LS}(S)^{\perp}$. Hence, $\langle \mathbf{z}-\mathbf{w}, \mathbf{w}\rangle = 0$ as $\mathbf{w} \in \mathrm{LS}(S)$. Further,

$$\|\mathbf{z}\|^2 = \|\mathbf{w} + (\mathbf{z}-\mathbf{w})\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{z}-\mathbf{w}\|^2 \ge \|\mathbf{w}\|^2.$$

Part 4: Follows directly using Part 2b as $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is a basis of $\mathbb{V}$. ■

A rephrasing of Theorem 5.2.5.2b gives a generalization of the pythagoras theorem, popularly known as the Parseval's formula. The proof is left as an exercise for the reader.

Theorem 5.2.6. *Let $\mathbb{V}$ be a finite dimensional* IPS *with an orthonormal basis $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$. Then, for each $\mathbf{x}, \mathbf{y} \in \mathbb{V}$,*

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{v}_i\rangle \overline{\langle \mathbf{y}, \mathbf{v}_i\rangle}.$$

Furthermore, if $\mathbf{x} = \mathbf{y}$ then $\|\mathbf{x}\|^2 = \sum\limits_{i=1}^{n} \mid \langle \mathbf{x}, \mathbf{v}_i\rangle \mid^2$ (generalizing the **Pythagoras Theorem***).*

As a corollary to Theorem 5.2.5, we have the following result.

Theorem 5.2.7 (Bessel's Inequality). *Let $\mathbb{V}$ be an* IPS *with $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ as an orthogonal set. Then, $\sum\limits_{k=1}^{n} \dfrac{\mid \langle \mathbf{z}, \mathbf{v}_k\rangle \mid^2}{\|\mathbf{v}_k\|^2} \le \|\mathbf{z}\|^2$, for each $\mathbf{z} \in \mathbb{V}$. Equality holds if and only if $\mathbf{z} = \sum\limits_{k=1}^{n} \dfrac{\langle \mathbf{z}, \mathbf{v}_k\rangle}{\|\mathbf{v}_k\|^2} \mathbf{v}_k$.*

Proof. For $1 \le k \le n$, define $\mathbf{u}_k = \dfrac{\mathbf{v}_k}{\|\mathbf{v}_k\|}$ and use Theorem 5.2.5.4 to get the required result. ■

Remark 5.2.8. *Using Theorem 5.2.5, we see that if $\mathcal{B} = \left[\mathbf{v}_1, \ldots, \mathbf{v}_n\right]$ is an ordered orthonormal basis of an* IPS *$\mathbb{V}$ then for each $\mathbf{u} \in \mathbb{V}$, $[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{u}, \mathbf{v}_1\rangle \\ \vdots \\ \langle \mathbf{u}, \mathbf{v}_n\rangle \end{bmatrix}$. Thus, in place of solving a linear system to get the coordinates of a vector, we just need to compute the inner product with basis vectors.*

EXERCISE **5.2.9.** 1. *Find $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ such that $\mathbf{v}, \mathbf{w}, (1,-1,-2)^T$ are mutually orthogonal.*

2. *Let $\mathcal{B} = \left[\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}\right]$ be an ordered basis of $\mathbb{R}^2$. Then, $\left[\begin{bmatrix}x\\y\end{bmatrix}\right]_{\mathcal{B}} = \begin{bmatrix}\frac{x+y}{\sqrt{2}}\\ \frac{x-y}{\sqrt{2}}\end{bmatrix}$.*

3. *For the ordered basis $\mathcal{B} = \left[\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\\0\end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\-2\end{bmatrix}\right]$ of $\mathbb{R}^3$, $[(2,3,1)^T]_{\mathcal{B}} = \begin{bmatrix}2\sqrt{3}\\ \frac{-1}{\sqrt{2}}\\ \frac{3}{\sqrt{6}}\end{bmatrix}$.*

In view of the importance of Theorem 5.2.5, we inquire into the question of extracting an orthonormal basis from a given basis. The process of extracting an orthonormal basis from a finite linearly independent set is called the **Gram-Schmidt Orthonormalization process**. We first consider a few examples. Note that Theorem 5.2.5 also gives us an algorithm for doing so, *i.e.*, from the given vector subtract all the orthogonal projections/components. If the new vector is nonzero then this vector is orthogonal to the previous ones. The proof follows directly from Theorem 5.2.5 but we give it again for the sake of completeness.

Theorem 5.2.10 (Gram-Schmidt Orthogonalization Process)**.** *Let $\mathbb{V}$ be an* IPS*. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of linearly independent vectors in $\mathbb{V}$ then there exists an orthonormal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ in $\mathbb{V}$. Furthermore, $LS(\mathbf{w}_1, \ldots, \mathbf{w}_i) = LS(\mathbf{v}_1, \ldots, \mathbf{v}_i)$, for $1 \le i \le n$.*

Proof. Note that for orthonormality, we need $\|\mathbf{w}_i\| = 1$, for $1 \le i \le n$ and $\langle \mathbf{w}_i, \mathbf{w}_j\rangle = 0$, for $1 \le i \ne j \le n$. Also, by Corollary 3.3.8.2, $\mathbf{v}_i \notin LS(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1})$, for $2 \le i \le n$, as $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a linearly independent set. We are now ready to prove the result by induction.

Step 1: Define $\mathbf{w}_1 = \dfrac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$ then $LS(\mathbf{v}_1) = LS(\mathbf{w}_1)$.

Step 2: Define $\mathbf{u}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1\rangle \mathbf{w}_1$. Then, $\mathbf{u}_2 \ne \mathbf{0}$ as $\mathbf{v}_2 \notin LS(\mathbf{v}_1)$. So, let $\mathbf{w}_2 = \dfrac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$.

Note that $\{\mathbf{w}_1, \mathbf{w}_2\}$ is orthonormal and $LS(\mathbf{w}_1, \mathbf{w}_2) = LS(\mathbf{v}_1, \mathbf{v}_2)$.

Step 3: For induction, assume that we have obtained an orthonormal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}\}$ such that $LS(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}) = LS(\mathbf{w}_1, \ldots, \mathbf{w}_{k-1})$. Now, note that $\mathbf{u}_k = \mathbf{v}_k - \sum\limits_{i=1}^{k-1} \langle \mathbf{v}_k, \mathbf{w}_i\rangle \mathbf{w}_i = \mathbf{v}_k - \sum\limits_{i=1}^{k-1} \mathrm{Proj}_{\mathbf{w}_i}(\mathbf{v}_k) \ne \mathbf{0}$ as $\mathbf{v}_k \notin LS(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$. So, let us put $\mathbf{w}_k = \dfrac{\mathbf{u}_k}{\|\mathbf{u}_k\|}$. Then, $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ is orthonormal as $\|\mathbf{w}_k\| = 1$ and

$$\begin{aligned}
\|\mathbf{u}_k\|\langle \mathbf{w}_k, \mathbf{w}_1\rangle &= \langle \mathbf{u}_k, \mathbf{w}_1\rangle = \langle \mathbf{v}_k - \sum_{i=1}^{k-1}\langle \mathbf{v}_k, \mathbf{w}_i\rangle \mathbf{w}_i, \mathbf{w}_1\rangle = \langle \mathbf{v}_k, \mathbf{w}_1\rangle - \langle \sum_{i=1}^{k-1}\langle \mathbf{v}_k, \mathbf{w}_i\rangle \mathbf{w}_i,\ \mathbf{w}_1\rangle \\
&= \langle \mathbf{v}_k, \mathbf{w}_1\rangle - \sum_{i=1}^{k-1}\langle \mathbf{v}_k, \mathbf{w}_i\rangle\langle \mathbf{w}_i,\ \mathbf{w}_1\rangle = \langle \mathbf{v}_k, \mathbf{w}_1\rangle - \langle \mathbf{v}_k, \mathbf{w}_1\rangle = \mathbf{0}.
\end{aligned}$$

Similarly, $\langle \mathbf{w}_k, \mathbf{w}_i\rangle = 0$, for $2 \le i \le k-1$. Clearly, $\mathbf{w}_k = \mathbf{u}_k/\|\mathbf{u}_k\| \in LS(\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}, \mathbf{v}_k)$. So, $\mathbf{w}_k \in LS(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.

As $\mathbf{v}_k = \|\mathbf{u}_k\|\mathbf{w}_k + \sum\limits_{i=1}^{k-1}\langle \mathbf{v}_k, \mathbf{w}_i\rangle \mathbf{w}_i$, we get $\mathbf{v}_k \in LS(\mathbf{w}_1, \ldots, \mathbf{w}_k)$. Hence, by the principle of mathematical induction $LS(\mathbf{w}_1, \ldots, \mathbf{w}_k) = LS(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ and the required result follows. ∎

We now illustrate the Gram-Schmidt process with a few examples.

Example 5.2.11.

1. Let $S = \{(1,-1,1,1), (1,0,1,0), (0,1,0,1)\} \subseteq \mathbb{R}^4$. Find an orthonormal set T such that $LS(S) = LS(T)$.

   Solution: Let $\mathbf{v}_1 = (1,0,1,0)^T, \mathbf{v}_2 = (0,1,0,1)^T$ and $\mathbf{v}_3 = (1,-1,1,1)^T$. Then, $\mathbf{w}_1 = \frac{1}{\sqrt{2}}(1,0,1,0)^T$. As $\langle \mathbf{v}_2, \mathbf{w}_1\rangle = 0$, we get $\mathbf{w}_2 = \frac{1}{\sqrt{2}}(0,1,0,1)^T$. For the third vector, let $\mathbf{u}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{w}_1\rangle \mathbf{w}_1 - \langle \mathbf{v}_3, \mathbf{w}_2\rangle \mathbf{w}_2 = (0,-1,0,1)^T$. Thus, $\mathbf{w}_3 = \frac{1}{\sqrt{2}}(0,-1,0,1)^T$.

2. Let $S = \{\mathbf{v}_1 = \begin{bmatrix}2 & 0 & 0\end{bmatrix}^T, \mathbf{v}_2 = \begin{bmatrix}\frac{3}{2} & 2 & 0\end{bmatrix}^T, \mathbf{v}_3 = \begin{bmatrix}\frac{1}{2} & \frac{3}{2} & 0\end{bmatrix}^T, \mathbf{v}_4 = \begin{bmatrix}1 & 1 & 1\end{bmatrix}^T\}$. Find an orthonormal set T such that $LS(S) = LS(T)$.

Solution: Take $\mathbf{w}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{bmatrix}1 & 0 & 0\end{bmatrix}^T = \mathbf{e}_1$. For the second vector, consider $\mathbf{u}_2 = \mathbf{v}_2 - \frac{3}{2}\mathbf{w}_1 = \begin{bmatrix}0 & 2 & 0\end{bmatrix}^T$. So, put $\mathbf{w}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \begin{bmatrix}0 & 1 & 0\end{bmatrix}^T = \mathbf{e}_2$.

For the third vector, let $\mathbf{u}_3 = \mathbf{v}_3 - \sum\limits_{i=1}^{2}\langle \mathbf{v}_3, \mathbf{w}_i\rangle \mathbf{w}_i = (0,0,0)^T$. So, $\mathbf{v}_3 \in LS((\mathbf{w}_1, \mathbf{w}_2))$. Or equivalently, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent.

So, for again computing the third vector, define $\mathbf{u}_4 = \mathbf{v}_4 - \sum\limits_{i=1}^{2}\langle \mathbf{v}_4, \mathbf{w}_i\rangle \mathbf{w}_i$. Then, $\mathbf{u}_4 = \mathbf{v}_4 - \mathbf{w}_1 - \mathbf{w}_2 = \mathbf{e}_3$. So $\mathbf{w}_4 = \mathbf{e}_3$. Hence, $T = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_4\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

3. Find an orthonormal set in $\mathbb{R}^3$ containing $(1,2,1)^T$.
   Solution: Let $(x,y,z)^T \in \mathbb{R}^3$ with $\big\langle (1,2,1), (x,y,z)\big\rangle = 0$. Thus,

$$(x,y,z) = (-2y-z, y, z) = y(-2,1,0) + z(-1,0,1).$$

   Observe that $(-2,1,0)$ and $(-1,0,1)$ are orthogonal to $(1,2,1)$ but are themselves not orthogonal.

   METHOD 1: Apply Gram-Schmidt process to $\{\frac{1}{\sqrt{6}}(1,2,1)^T, (-2,1,0)^T, (-1,0,1)^T\} \subseteq \mathbb{R}^3$.

   METHOD 2: Valid only in $\mathbb{R}^3$ using the cross product of two vectors.

   In either case, verify that $\{\frac{1}{\sqrt{6}}(1,2,1), \frac{-1}{\sqrt{5}}(2,-1,0), \frac{-1}{\sqrt{30}}(1,2,-5)\}$ is the required set.

We now state two immediate corollaries without proof.

Corollary 5.2.12. *Let $\mathbb{V} \neq \{\mathbf{0}\}$ be an IPS. If*

1. *$\mathbb{V}$ is finite dimensional then $\mathbb{V}$ has an orthonormal basis.*
2. *S is a non-empty orthonormal set and $\dim(\mathbb{V})$ is finite then S can be extended to form an orthonormal basis of $\mathbb{V}$.*

Remark 5.2.13. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \neq \{\mathbf{0}\}$ be a non-empty subset of a finite dimensional vector space $\mathbb{V}$. If we apply Gram-Schmidt process to*

1. *S then we obtain an orthonormal basis of $LS(\mathbf{v}_1, \ldots, \mathbf{v}_n)$.*
2. *a re-arrangement of elements of S then we may obtain another orthonormal basis of $LS(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. But, observe that the size of the two bases will be the same.*

EXERCISE **5.2.14.**

1. *Let $\mathbb{V}$ be an IPS with $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ as a basis. Then, prove that $\mathcal{B}$ is orthonormal if and only if for each $x \in \mathbb{V}$, $x = \sum\limits_{i=1}^{n}\langle \mathbf{x}, \mathbf{v}_i\rangle \mathbf{v}_i$. [Hint: Since $\mathcal{B}$ is a basis, each $\mathbf{x} \in \mathbb{V}$ has a unique linear combination in terms of $\mathbf{v}_i$'s.]*
2. *Let S be a subset of $\mathbb{V}$ having 101 elements. Suppose that the application of the Gram-Schmidt process yields $\mathbf{u}_5 = \mathbf{0}$. Does it imply that $LS(\mathbf{v}_1, \ldots, \mathbf{v}_5) = LS(\mathbf{v}_1, \ldots, \mathbf{v}_4)$? Give reasons for your answer.*
3. *Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal set in $\mathbb{R}^n$. For $1 \le k \le n$, define $A_k = \sum\limits_{i=1}^{k}\mathbf{v}_i\mathbf{v}_i^T$. Then, prove that $A_k^T = A_k$ and $A_k^2 = A_k$. Thus, A_k's are projection matrices.*
4. *Determine an orthonormal basis of $\mathbb{R}^4$ containing $(1,-2,1,3)^T$ and $(2,1,-3,1)^T$.*

5. *Let* $\mathbf{x} \in \mathbb{R}^n$ *with* $\|\mathbf{x}\| = 1$.

   (a) *Then, prove that* $\{\mathbf{x}\}$ *can be extended to form an orthonormal basis of* $\mathbb{R}^n$.

   (b) *Let the extended basis be* $\{\mathbf{x}, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *and* $\mathcal{B} = [\mathbf{e}_1, \ldots, \mathbf{e}_n]$ *the standard ordered basis of* $\mathbb{R}^n$. *Prove that* $A = \begin{bmatrix} [\mathbf{x}]_{\mathcal{B}}, & [\mathbf{x}_2]_{\mathcal{B}}, & \ldots, & [\mathbf{x}_n]_{\mathcal{B}} \end{bmatrix}$ *is an orthogonal matrix.*

6. *Let* $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n, n \geq 1$ *with* $\|\mathbf{u}\| = \|\mathbf{w}\| = 1$. *Prove that there exists an orthogonal matrix* A *such that* $A\mathbf{v} = \mathbf{w}$. *Prove also that* A *can be chosen such that* $\det(A) = 1$.

7. *Let* $(\mathbb{V}, \langle\ ,\ \rangle)$ *be an* n*-dimensional* IPS. *If* $\mathbf{u} \in \mathbb{V}$ *with* $\|\mathbf{u}\| = 1$ *then give reasons for the following statements.*

   (a) *Let* $S^\perp = \{\mathbf{v} \in \mathbb{V} \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0\}$. *Then,* $\dim(S^\perp) = n - 1$.

   (b) *Let* $0 \neq \beta \in \mathbb{F}$. *Then,* $S = \{\mathbf{v} \in \mathbb{V} : \langle \mathbf{v}, \mathbf{u} \rangle = \beta\}$ *is not a subspace of* $\mathbb{V}$.

   (c) *Let* $\mathbf{v} \in \mathbb{V}$. *Then,* $\mathbf{v} = \mathbf{v}_0 + \langle \mathbf{v}, \mathbf{u} \rangle \mathbf{u}$ *for a vector* $\mathbf{v}_0 \in S^\perp$. *That is,* $\mathbb{V} = LS(\mathbf{u}, S^\perp)$.

5.2.1 Application to Fundamental Spaces

We end this section by proving the fundamental theorem of linear algebra. So, the readers are advised to recall the four fundamental subspaces and also to go through Theorem 3.5.9 (the rank-nullity theorem for matrices). We start with the following result.

Lemma 5.2.15. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{R})$. *Then,* $\text{NULL}(A) = \text{NULL}(A^T A)$.

Proof. Let $\mathbf{x} \in \text{NULL}(A)$. Then, $A\mathbf{x} = \mathbf{0}$. So, $(A^T A)\mathbf{x} = A^T(A\mathbf{x}) = A^T\mathbf{0} = \mathbf{0}$. Thus, $\mathbf{x} \in \text{NULL}(A^T A)$. That is, $\text{NULL}(A) \subseteq \text{NULL}(A^T A)$.

Suppose that $\mathbf{x} \in \text{NULL}(A^T A)$. Then, $(A^T A)\mathbf{x} = \mathbf{0}$ and $0 = \mathbf{x}^T\mathbf{0} = \mathbf{x}^T(A^T A)\mathbf{x} = (A\mathbf{x})^T(A\mathbf{x}) = \|A\mathbf{x}\|^2$. Thus, $A\mathbf{x} = \mathbf{0}$ and the required result follows. ∎

Theorem 5.2.16 (Fundamental Theorem of Linear Algebra)**.** *Let* $A \in \mathbb{M}_n(\mathbb{C})$. *Then,*

1. $\dim(\text{NULL}(A)) + \dim(\text{COL}(A)) = n$.

2. $\text{NULL}(A) = \big(\text{COL}(A^*)\big)^\perp$ *and* $\text{NULL}(A^*) = \big(\text{COL}(A)\big)^\perp$.

3. $\dim(\text{COL}(A)) = \dim(\text{COL}(A^*))$.

Proof. PART 1: Proved in Theorem 3.5.9.

PART 2: We first prove that $\text{NULL}(A) \subseteq \text{COL}(A^*)^\perp$. Let $\mathbf{x} \in \text{NULL}(A)$. Then, $A\mathbf{x} = \mathbf{0}$ and

$$0 = \langle \mathbf{0}, \mathbf{u} \rangle = \langle A\mathbf{x}, \mathbf{u} \rangle = \mathbf{u}^* A\mathbf{x} = (A^*\mathbf{u})^*\mathbf{x} = \langle \mathbf{x}, A^*\mathbf{u} \rangle, \text{ for all } \mathbf{u} \in \mathbb{C}^n.$$

But $\text{COL}(A^*) = \{A^*\mathbf{u} \mid \mathbf{u} \in \mathbb{C}^n\}$. Thus, $\mathbf{x} \in \text{COL}(A^*)^\perp$ and $\text{NULL}(A) \subseteq \text{COL}(A^*)^\perp$.

We now prove that $\text{COL}(A^*)^\perp \subseteq \text{NULL}(A)$. Let $\mathbf{x} \in \text{COL}(A^*)^\perp$. Then, for every $\mathbf{y} \in \mathbb{C}^n$,

$$0 = \langle \mathbf{x}, A^*\mathbf{y} \rangle = (A^*\mathbf{y})^*\mathbf{x} = \mathbf{y}^*(A^*)^*\mathbf{x} = \mathbf{y}^* A\mathbf{x} = \langle A\mathbf{x}, \mathbf{y} \rangle.$$

In particular, for $\mathbf{y} = A\mathbf{x} \in \mathbb{C}^n$, we get $\|A\mathbf{x}\|^2 = 0$. Hence $A\mathbf{x} = \mathbf{0}$. That is, $\mathbf{x} \in \text{NULL}(A)$. Thus, the proof of the first equality in Part 2 is over. We omit the second equality as it proceeds on the same lines as above.

PART 3: Use the first two parts to get the required result.

Hence the proof of the fundamental theorem is complete. ∎

Remark 5.2.17. *Theorem 5.2.16.2 implies that* $\text{NULL}(A) = \big(\text{COL}(A^*)\big)^{\perp}$. *This statement is just stating the usual fact that if* $\mathbf{x} \in \text{NULL}(A)$ *then* $A\mathbf{x} = \mathbf{0}$ *and hence the usual dot product of every row of* A *with* $\mathbf{x}$ *equals* 0.

As an implication of Theorem 5.2.16.2 and Theorem 5.2.16.3, we show the existence of an invertible linear map $T : \text{COL}(A^*) \to \text{COL}(A)$.

Corollary 5.2.18. *Let* $A \in \mathbb{M}_n(\mathbb{C})$. *Then, the function* $T : \text{COL}(A^*) \to \text{COL}(A)$ *defined by* $T(\mathbf{x}) = A\mathbf{x}$ *is invertible.*

Proof. In view of Theorem 5.2.16.3 and the rank-nullity theorem, we just need to show that the map is one-one. So, suppose that there exist $\mathbf{x}, \mathbf{y} \in \text{COL}(A^*)$ such that $T(\mathbf{x}) = T(\mathbf{y})$. Or equivalently, $A\mathbf{x} = A\mathbf{y}$. Thus, $\mathbf{x} - \mathbf{y} \in \text{NULL}(A) = (\text{COL}(A^*))^{\perp}$ (by Theorem 5.2.16.2). Therefore, $\mathbf{x} - \mathbf{y} \in (\text{COL}(A^*))^{\perp} \cap \text{COL}(A^*) = \{\mathbf{0}\}$. Thus, $\mathbf{x} = \mathbf{y}$ and hence the map is one-one. Thus, the required result follows. ■

The readers should look at Example 3.2.3 and Remark 3.2.4. We give one more example.

Example 5.2.19. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 3 & 2 & 1 \end{bmatrix}$. Then, verify that

1. $\{(0,1,1)^T, (1,1,2)^T\}$ is a basis of $\text{COL}(A)$.
2. $\{(1,1,-1)^T\}$ is a basis of $\text{NULL}(A^T)$.
3. $\text{NULL}(A^T) = (\text{COL}(A))^{\perp}$.

EXERCISE **5.2.20.**
1. *Find distinct subspaces* $\mathbb{W}_1$ *and* $\mathbb{W}_2$
   (a) *in* $\mathbb{R}^2$ *such that* $\mathbb{W}_1$ *and* $\mathbb{W}_2$ *are orthogonal but not orthogonal complement.*
   (b) *in* $\mathbb{R}^3$ *such that* $\mathbb{W}_1 \neq \{\mathbf{0}\}$ *and* $\mathbb{W}_2 \neq \{\mathbf{0}\}$ *are orthogonal, but not orthogonal complement.*
2. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{C})$. *Then,* $\text{NULL}(A) = \text{NULL}(A^*A)$.
3. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{R})$. *Then,* $\text{COL}(A) = \text{COL}(A^TA)$.
4. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{R})$. *Then, Rank*$(A) = n$ *if and only if Rank*$(A^TA) = n$.
5. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{C})$. *Then, for every*
   (a) $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = \mathbf{u} + \mathbf{v}$, *where* $\mathbf{u} \in \text{COL}(A^T)$ *and* $\mathbf{v} \in \text{NULL}(A)$ *are unique.*
   (b) $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{y} = \mathbf{w} + \mathbf{z}$, *where* $\mathbf{w} \in \text{COL}(A)$ *and* $\mathbf{z} \in \text{NULL}(A^T)$ *are unique.*

For more information related with the fundamental theorem of linear algebra the interested readers are advised to see the article "The Fundamental Theorem of Linear Algebra, Gilbert Strang, *The American Mathematical Monthly*, Vol. 100, No. 9, Nov., 1993, pp. 848 - 855."

5.2.2 QR Decomposition*

The next result gives the proof of the QR decomposition for real matrices. The readers are advised to prove similar results for matrices with complex entries. This decomposition and its generalizations are helpful in the numerical calculations related with eigenvalue problems (see Chapter 6).

Theorem 5.2.1 (QR Decomposition)**.** *Let $A \in \mathbb{M}_n(\mathbb{R})$ be invertible. Then, there exist matrices Q and R such that Q is orthogonal and R is upper triangular with $A = QR$. Furthermore, if $\det(A) \neq 0$ then the diagonal entries of R can be chosen to be positive. Also, in this case, the decomposition is unique.*

Proof. As A is invertible, it's columns form a basis of $\mathbb{R}^n$. So, an application of the Gram-Schmidt orthonormalization process to $\{A[:,1], \ldots, A[:,n]\}$ gives an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{R}^n$ satisfying

$$LS(A[:,1], \ldots, A[:,i]) = LS(\mathbf{v}_1, \ldots, \mathbf{v}_i), \quad \text{for } 1 \leq i \leq n.$$

Since $A[:,i] \in LS(\mathbf{v}_1, \ldots, \mathbf{v}_i)$, for $1 \leq i \leq n$, there exist $\alpha_{ji} \in \mathbb{R}, 1 \leq j \leq i$, such that $A[:,i] = [\mathbf{v}_1, \ldots, \mathbf{v}_i]\begin{bmatrix}\alpha_{1i}\\ \vdots \\ \alpha_{ii}\end{bmatrix}$. Thus, if $Q = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ and $R = \begin{bmatrix}\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n}\\ 0 & \alpha_{22} & \cdots & \alpha_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \alpha_{nn}\end{bmatrix}$ then

1. Q is an orthogonal matrix (see Exercise 5.4.8.5),
2. R is an upper triangular matrix, and
3. $A = QR$.

Thus, this completes the proof of the first part. Note that

1. $\alpha_{ii} \neq 0$, for $1 \leq i \leq n$, as $A[:,1] \neq \mathbf{0}$ and $A[:,i] \notin LS(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1})$.
2. if $\alpha_{ii} < 0$, for some $i, 1 \leq i \leq n$ then we can replace $\mathbf{v}_i$ in Q by $-\mathbf{v}_i$ to get a new Q ad R in which the diagonal entries of R are positive.

Uniqueness: Suppose $Q_1R_1 = Q_2R_2$ for some orthogonal matrices Q_i's and upper triangular matrices R_i's with positive diagonal entries. As Q_i's and R_i's are invertible, we get $Q_2^{-1}Q_1 = R_2R_1^{-1}$. Now, using

1. Exercises 2.3.25.1, 1.2.15.1, the matrix $R_2R_1^{-1}$ is an upper triangular matrix.
2. Exercises 1.3.2.3, $Q_2^{-1}Q_1$ is an orthogonal matrix.

So, the matrix $R_2R_1^{-1}$ is an orthogonal upper triangular matrix and hence, by Exercise 1.2.11.4, $R_2R_1^{-1} = I_n$. So, $R_2 = R_1$ and therefore $Q_2 = Q_1$. ∎

Let A be an $n \times k$ matrix with $\text{Rank}(A) = r$. Then, by Remark 5.2.13, an application of the Gram-Schmidt orthonormalization process to columns of A yields an orthonormal set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subseteq \mathbb{R}^n$ such that

$$LS(A[:,1], \ldots, A[:,j]) = LS(\mathbf{v}_1, \ldots, \mathbf{v}_i), \quad \text{for } 1 \leq i \leq j \leq k.$$

Hence, proceeding on the lines of the above theorem, we have the following result.

Theorem 5.2.2 (Generalized QR Decomposition)**.** *Let A be an $n \times k$ matrix of rank r. Then, $A = QR$, where*

1. *$Q = [\mathbf{v}_1, \ldots, \mathbf{v}_r]$ is an $n \times r$ matrix with $Q^TQ = I_r$,*

2. *$LS(A[:,1],\ldots,A[:,j]) = LS(\mathbf{v}_1,\ldots,\mathbf{v}_i)$, for $1 \le i \le j \le k$ and*

3. *R is an $r \times k$ matrix with $Rank(R) = r$.*

Example 5.2.3. 1. Let $A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$. Find an orthogonal matrix Q and an upper triangular matrix R such that $A = QR$.

Solution: From Example 5.2.11, we know that $\mathbf{w}_1 = \frac{1}{\sqrt{2}}(1,0,1,0)^T$, $\mathbf{w}_2 = \frac{1}{\sqrt{2}}(0,1,0,1)^T$ and $\mathbf{w}_3 = \frac{1}{\sqrt{2}}(0,-1,0,1)^T$. We now compute $\mathbf{w}_4$. If $\mathbf{v}_4 = (2,1,1,1)^T$ then

$$\mathbf{u}_4 = \mathbf{v}_4 - \langle \mathbf{v}_4, \mathbf{w}_1\rangle \mathbf{w}_1 - \langle \mathbf{v}_4, \mathbf{w}_2\rangle \mathbf{w}_2 - \langle \mathbf{v}_4, \mathbf{w}_3\rangle \mathbf{w}_3 = \frac{1}{2}(1,0,-1,0)^T.$$

Thus, $\mathbf{w}_4 = \frac{1}{\sqrt{2}}(-1,0,1,0)^T$. Hence, we see that $A = QR$ with

$$Q = \begin{bmatrix}\mathbf{w}_1,\ldots,\mathbf{w}_4\end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \text{ and } R = \begin{bmatrix} \sqrt{2} & 0 & \sqrt{2} & -\frac{3}{\sqrt{2}} \\ 0 & \sqrt{2} & 0 & -\sqrt{2} \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

2. Let $A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & 0 & -2 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \end{bmatrix}$. Find a 4×3 matrix Q satisfying $Q^TQ = I_3$ and an upper triangular matrix R such that $A = QR$.

Solution: Let us apply the Gram-Schmidt orthonormalization process to the columns of A. As $\mathbf{v}_1 = (1,-1,1,1)^T$, we get $\mathbf{w}_1 = \frac{1}{2}\mathbf{v}_1$. Let $\mathbf{v}_2 = (1,0,1,0)^T$. Then,

$$\mathbf{u}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1\rangle \mathbf{w}_1 = (1,0,1,0)^T - \mathbf{w}_1 = \frac{1}{2}(1,1,1,-1)^T.$$

Hence, $\mathbf{w}_2 = \frac{1}{2}(1,1,1,-1)^T$. Let $\mathbf{v}_3 = (1,-2,1,2)^T$. Then,

$$\mathbf{u}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{w}_1\rangle \mathbf{w}_1 - \langle \mathbf{v}_3, \mathbf{w}_2\rangle \mathbf{w}_2 = \mathbf{v}_3 - 3\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{0}.$$

So, we again take $\mathbf{v}_3 = (0,1,0,1)^T$. Then,

$$\mathbf{u}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{w}_1\rangle \mathbf{w}_1 - \langle \mathbf{v}_3, \mathbf{w}_2\rangle \mathbf{w}_2 = \mathbf{v}_3 - 0\mathbf{w}_1 - 0\mathbf{w}_2 = \mathbf{v}_3.$$

So, $\mathbf{w}_3 = \frac{1}{\sqrt{2}}(0,1,0,1)^T$. Hence,

$$Q = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{-1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{-1}{2} & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } R = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix}.$$

The readers are advised to check the following:

(a) $\text{Rank}(A) = 3$,

(b) $A = QR$ with $Q^T Q = I_3$, and

(c) R is a 3×4 upper triangular matrix with $\text{Rank}(R) = 3$.

Remark 5.2.4. *Let $A \in \mathbb{M}_{m,n}(\mathbb{R})$.*

1. *If $A = QR$ with $Q = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ then* $R = \begin{bmatrix} \langle \mathbf{v}_1, A[:,1]\rangle & \langle \mathbf{v}_1, A[:,2]\rangle & \langle \mathbf{v}_1, A[:,3]\rangle & \cdots \\ 0 & \langle \mathbf{v}_2, A[:,2]\rangle & \langle \mathbf{v}_2, A[:,3]\rangle & \cdots \\ 0 & 0 & \langle \mathbf{v}_3, A[:,3]\rangle & \cdots \\ \vdots & \vdots & & \ddots \end{bmatrix}.$

   In case Rank$(A) < n$ then a slight modification gives the matrix R.

2. *Further, let Rank$(A) = n$.*

   (a) Then, $A^T A$ is invertible (see Exercise 5.2.20.4).

   (b) By Theorem 5.2.2, $A = QR$ with Q a matrix of size $m \times n$ and R an upper triangular matrix of size $n \times n$. Also, $Q^T Q = I_n$ and Rank$(R) = n$.

   (c) Thus, $A^T A = R^T Q^T Q R = R^T R$. As $A^T A$ is invertible, the matrix $R^T R$ is invertible. Since R is a square matrix, by Exercise 2.3.5.1, the matrix R itself is invertible. Hence, $(R^T R)^{-1} = R^{-1}(R^T)^{-1}$.

   (d) So, if $Q = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ then

   $$A(A^T A)^{-1} A^T = QR(R^T R)^{-1} R^T Q^T = (QR)(R^{-1}(R^T)^{-1}) R^T Q^T = QQ^T.$$

   (e) Hence, using Theorem 5.3.7, we see that the matrix

   $$P = A(A^T A)^{-1} A^T = QQ^T = [\mathbf{v}_1, \ldots, \mathbf{v}_n] \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} = \sum_{i=1}^{n} \mathbf{v}_i \mathbf{v}_i^T$$

   is the orthogonal projection matrix on $\text{Col}(A)$.

3. *Further, let Rank$(A) = r < n$. If $j_1, \ldots, j_r$ are the pivot columns of A then $\text{Col}(A) = \text{Col}(B)$, where $B = [A[:,j_1], \ldots, A[:,j_r]]$ is an $m \times r$ matrix with Rank$(B) = r$. So, using Part 2e we see that $B(B^T B)^{-1} B^T$ is the orthogonal projection matrix on $\text{Col}(A)$. So, compute RREF of A and choose columns of A corresponding to the pivot columns.*

5.3 Orthogonal Projections and Applications

Till now, our main interest was to understand the linear system $A\mathbf{x} = \mathbf{b}$, for $A \in \mathbb{M}_{m,n}(\mathbb{C}), \mathbf{x} \in \mathbb{C}^n$ and $\mathbf{b} \in \mathbb{C}^m$, from different view points. But, in most practical situations the system has no solution. So, we are interested in finding a point $\mathbf{x}_0 \in \mathbb{R}^n$ such that the $\mathbf{err} = \mathbf{b} - A\mathbf{x}_0$ is the least. Thus, we consider the problem of finding $\mathbf{x}_0 \in \mathbb{R}^n$ such that

$$\|\mathbf{err}\| = \|\mathbf{b} - A\mathbf{x}_0\| = \min\{\|\mathbf{b} - A\mathbf{x}_0\| : \mathbf{x} \in \mathbb{R}^n\},$$

i.e., we try to find the vector $\mathbf{x}_0 \in \mathbb{R}^n$ which is nearest to $\mathbf{b}$.

To begin with, recall the following result from Page 129.

Theorem 5.3.1 (Decomposition)**.** *Let $\mathbb{V}$ be an IPS having $\mathbb{W}$ as a finite dimensional subspace. Suppose $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ is an orthonormal basis of $\mathbb{W}$. Then, for each $\mathbf{b} \in \mathbb{V}$, $\mathbf{y} = \sum\limits_{i=1}^{k} \langle \mathbf{b}, \mathbf{f}_i \rangle \mathbf{f}_i$ is the closest point in $\mathbb{W}$ from $\mathbf{b}$. Furthermore, $\mathbf{b} - \mathbf{y} \in \mathbb{W}^{\perp}$.*

We now give a definition and then an implication of Theorem 5.3.1.

Definition 5.3.2. **[Orthogonal Projection]** Let $\mathbb{W}$ be a finite dimensional subspace of an IPS $\mathbb{V}$. Then, by Theorem 5.3.1, for each $\mathbf{v} \in \mathbb{V}$ there exist unique vectors $\mathbf{w} \in \mathbb{W}$ and $\mathbf{u} \in \mathbb{W}^{\perp}$ with $\mathbf{v} = \mathbf{w} + \mathbf{u}$. We thus define the **orthogonal projection** of $\mathbb{V}$ onto $\mathbb{W}$, denoted $P_{\mathbb{W}}$, by

$$P_{\mathbb{W}} : \mathbb{V} \to \mathbb{W} \;\text{ by }\; P_{\mathbb{W}}(\mathbf{v}) = \mathbf{w}.$$

The vector $\mathbf{w}$ is called the **projection** of $\mathbf{v}$ on $\mathbb{W}$.

Remark 5.3.3. *Let $A \in \mathbb{M}_{m,n}(\mathbb{R})$ and $\mathbb{W} = \mathrm{COL}(A)$. Then, to find the orthogonal projection $P_{\mathbb{W}}(\mathbf{b})$, we can use either of the following ideas:*

1. *Determine an orthonormal basis $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ of $\mathrm{COL}(A)$ and get $P_{\mathbb{W}}(\mathbf{b}) = \sum\limits_{i=1}^{k} \langle \mathbf{b}, \mathbf{f}_i \rangle \mathbf{f}_i$.*
2. *By Theorem 5.2.16.2, $\mathrm{COL}(A) = \mathrm{NULL}(A^T)^{\perp}$. Hence, for $\mathbf{b} \in \mathbb{R}^m$ there exists unique $\mathbf{u} \in \mathrm{COL}(A)$ and $\mathbf{v} \in \mathrm{NULL}(A^T)$ such that $\mathbf{b} = \mathbf{u} + \mathbf{v}$. Thus, using Definition 5.3.2 and Theorem 5.3.1, $P_{\mathbb{W}}(\mathbf{b}) = \mathbf{u}$.*

Before proceeding to projections, we give an application of Theorem 5.3.1 to a linear system.

Corollary 5.3.4. *Let $A \in \mathbb{M}_{m,n}(\mathbb{R})$ and $\mathbf{b} \in \mathbb{R}^m$. Then, every least square solution of $A\mathbf{x} = \mathbf{b}$ is a solution of the system $A^TA\mathbf{x} = A^T\mathbf{b}$. Conversely, every solution of $A^TA\mathbf{x} = A^T\mathbf{b}$ is a least square solution of $A\mathbf{x} = \mathbf{b}$.*

Proof. As $\mathbf{b} \in \mathbb{R}^m$, by Remark 5.3.3, there exists $\mathbf{y} \in \mathrm{COL}(A)$ and $\mathbf{v} \in \mathrm{NULL}(A^T)$ such that $\mathbf{b} = \mathbf{y} + \mathbf{v}$ and $\min\{\|\mathbf{b} - \mathbf{w}\| \mid \mathbf{w} \in \mathrm{COL}(A)\} = \|\mathbf{b} - \mathbf{y}\|$. As $\mathbf{y} \in \mathrm{COL}(A)$, there exists $\mathbf{x}_0 \in \mathbb{R}^n$ such that $A\mathbf{x}_0 = \mathbf{y}$, *i.e.*, $\mathbf{x}_0$ is the least square solution of $A\mathbf{x} = \mathbf{b}$. Hence,

$$(A^TA)\mathbf{x}_0 = A^T(A\mathbf{x}_0) = A^T\mathbf{y} = A^T(\mathbf{b} - \mathbf{v}) = A^T\mathbf{b} - \mathbf{0} = A^T\mathbf{b}.$$

Conversely, let $\mathbf{x}_1 \in \mathbb{R}^n$ be a solution of $A^TA\mathbf{x} = A^T\mathbf{b}$, *i.e.*, $A^T(A\mathbf{x}_1 - \mathbf{b}) = \mathbf{0}$. To show

$$\min\{\|\mathbf{b} - A\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n\} = \|\mathbf{b} - A\mathbf{x}_1\|.$$

Note that $A^T(A\mathbf{x}_1 - \mathbf{b}) = \mathbf{0}$ implies $0 = (\mathbf{x} - \mathbf{x}_1)^TA^T(A\mathbf{x}_1 - \mathbf{b}) = (A\mathbf{x} - A\mathbf{x}_1)^T(A\mathbf{x}_1 - \mathbf{b})$ and hence $\langle \mathbf{b} - A\mathbf{x}_1, A(\mathbf{x} - \mathbf{x}_1)\rangle = 0$. Thus,

$$\|\mathbf{b} - A\mathbf{x}\|^2 = \|\mathbf{b} - A\mathbf{x}_1 + A\mathbf{x}_1 - A\mathbf{x}\|^2 = \|\mathbf{b} - A\mathbf{x}_1\|^2 + \|A\mathbf{x}_1 - A\mathbf{x}\|^2 \geq \|\mathbf{b} - \mathbf{y}\|^2.$$

Hence, the required result follows. ∎

The above corollary gives the following result.

Corollary 5.3.5. *Let $A \in \mathbb{M}_{m,n}(\mathbb{R})$ and $\mathbf{b} \in \mathbb{R}^m$. If*

1. *A^TA is invertible then the least square solution of $A\mathbf{x} = \mathbf{b}$ equals $\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b}$.*

2. *A^TA is not invertible then the least square solution of $A\mathbf{x} = \mathbf{b}$ equals $\mathbf{x} = (A^TA)^{-}A^T\mathbf{b}$, where $(A^TA)^{-}$ is the pseudo-inverse of A^TA.*

Proof. Part 1 directly follows from Corollary 5.3.5. For Part 2, let $\mathbf{b} = \mathbf{y} + \mathbf{v}$, for $\mathbf{y} \in \text{COL}(A)$ and $\mathbf{v} \in \text{NULL}(A^T)$. As $\mathbf{y} \in \text{COL}(A)$, there exists $\mathbf{x}_0 \in \mathbb{R}^n$ such that $A\mathbf{x}_0 = \mathbf{y}$. Thus, by Remark 5.3.3, $A^T\mathbf{b} = A^T(\mathbf{y}+\mathbf{v}) = A^T\mathbf{y} = A^TA\mathbf{x}_0$. Now, using the definition of pseudo-inverse (see Exercise 1.3.7.14), we see that

$$(A^AA)\left((A^TA)^{-}A^T\mathbf{b}\right) = (A^TA)(A^TA)^{-}(A^TA)\mathbf{x}_0 = (A^TA)\mathbf{x}_0 = A^T\mathbf{b}.$$

Thus, we see that $(A^TA)^{-}A^T\mathbf{b}$ is a solution of the system $A^TA\mathbf{x} = A^T\mathbf{b}$. Hence, by Corollary 5.3.4, the required result follows. ■

We now give a few examples to understand projections.

Example 5.3.6. Use the fundamental theorem of linear algebra to compute the vector of the orthogonal projection.

1. Determine the projection of $(1,1,1,1,1)^T$ on $\text{NULL}\left([1,-1,1,-1,1]\right)$.
   Solution: Here $A = [1,-1,1,-1,1]$. So, a basis of $\text{COL}(A^T)$ equals $\{(1,-1,1,-1,1)^T\}$ and that of $\text{NULL}(A)$ equals $\{(1,1,0,0,0)^T, (1,0,-1,0,0)^T, (1,0,0,1,0)^T, (1,0,0,0,-1)^T\}$. Then, the solution of the linear system
   $$B\mathbf{x} = \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, \text{ where } B = \begin{bmatrix}1&1&1&1&1\\1&0&0&0&-1\\0&-1&0&0&1\\0&0&1&0&-1\\0&0&0&-1&1\end{bmatrix} \text{ equals } \mathbf{x} = \frac{1}{5}\begin{bmatrix}6\\-4\\6\\-4\\1\end{bmatrix}.$$ Thus, the projection is
   $$\frac{1}{5}\left(6(1,1,0,0,0)^T - 4(1,0,-1,0,0)^T + 6(1,0,0,1,0)^T - 4(1,0,0,0,-1)^T\right) = \frac{2}{5}(2,3,2,3,2)^T.$$
2. Determine the projection of $(1,1,1)^T$ on $\text{NULL}\left([1,1,-1]\right)$.
   Solution: Here $A = [1,1,-1]$. So, a basis of $\text{NULL}(A)$ equals $\{(1,-1,0)^T, (1,0,1)^T\}$ and that of $\text{COL}(A^T)$ equals $\{(1,1,-1)^T\}$. Then, the solution of the linear system
   $$B\mathbf{x} = \begin{bmatrix}1\\1\\1\end{bmatrix}, \text{ where } B = \begin{bmatrix}1&1&1\\-1&0&1\\0&1&-1\end{bmatrix} \text{ equals } \mathbf{x} = \frac{1}{3}\begin{bmatrix}-2\\4\\1\end{bmatrix}.$$ Thus, the projection is
   $$\frac{1}{3}\left((-2)(1,-1,0)^T + 4(1,0,1)^T\right) = \frac{2}{3}(1,1,2)^T.$$
3. Determine the projection of $(1,1,1)^T$ on $\text{COL}\left([1,2,1]^T\right)$.
   Solution: Here, $A^T = [1,2,1]$, a basis of $\text{COL}(A)$ equals $\{(1,2,1)^T\}$ and that of $\text{NULL}(A^T)$ equals $\{(1,0,-1)^T, (2,-1,0)^T\}$. Then, using the solution of the linear system
   $$B\mathbf{x} = \begin{bmatrix}1\\1\\1\end{bmatrix}, \text{ where } B = \begin{bmatrix}1&2&1\\0&-1&2\\-1&0&1\end{bmatrix} \text{ gives } \frac{2}{3}(1,2,1)^T$$ as the required vector.

To use the first idea in Remark 5.3.3, we prove the following result which helps us to get the matrix of the orthogonal projection from an orthonormal basis.

Theorem 5.3.7. *Let $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ be an orthonormal basis of a finite dimensional subspace $\mathbb{W}$ of an IPS $\mathbb{V}$. Then $P_{\mathbb{W}} = \sum\limits_{i=1}^{k} \mathbf{f}_i\mathbf{f}_i^*$.*

Proof. Let $\mathbf{v} \in \mathbb{V}$. Then,

$$P_{\mathbb{W}}\mathbf{v} = \left(\sum_{i=1}^{k} \mathbf{f}_i\mathbf{f}_i^*\right)\mathbf{v} = \sum_{i=1}^{k} \mathbf{f}_i\left(\mathbf{f}_i^*\mathbf{v}\right) = \sum_{i=1}^{k}\langle \mathbf{v}, \mathbf{f}_i\rangle \mathbf{f}_i.$$

As $P_{\mathbb{W}}\mathbf{v}$ is the only closet point (see Theorem 5.3.1), the required result follows. ■

Example 5.3.8. In each of the following, determine the matrix of the orthogonal projection. Also, verify that $P_{\mathbb{W}} + P_{\mathbb{W}^\perp} = I$. What can you say about Rank$(P_{\mathbb{W}^\perp})$ and Rank$(P_{\mathbb{W}})$? Also, verify the orthogonal projection vectors obtained in Example 5.3.6.

1. $\mathbb{W} = \{(x_1, \ldots, x_5)^T \in \mathbb{R}^5 \mid x_1 - x_2 + x_3 - x_4 + x_5 = 0\} = \textsc{Null}\left([1, -1, 1, -1, 1]\right)$.

   Solution: An orthonormal basis of $\mathbb{W}$ is $\left\{\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\\0\\0\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}0\\0\\1\\1\\0\end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix}1\\-1\\0\\0\\-2\end{bmatrix}, \frac{1}{\sqrt{30}}\begin{bmatrix}-2\\2\\3\\-3\\-2\end{bmatrix}\right\}$. Thus,

   $$P_{\mathbb{W}} = \sum_{i=1}^{4}\mathbf{f}_i\mathbf{f}_i^T = \frac{1}{5}\begin{bmatrix}4 & 1 & -1 & 1 & -1\\ 1 & 4 & 1 & -1 & 1\\ -1 & 1 & 4 & 1 & -1\\ 1 & -1 & 1 & 4 & 1\\ -1 & 1 & -1 & 1 & 4\end{bmatrix} \text{ and } P_{\mathbb{W}^\perp} = \frac{1}{5}\begin{bmatrix}1 & -1 & 1 & -1 & 1\\ -1 & 1 & -1 & 1 & -1\\ 1 & -1 & 1 & -1 & 1\\ -1 & 1 & -1 & 1 & -1\\ 1 & -1 & 1 & -1 & 1\end{bmatrix}.$$

2. $\mathbb{W} = \{(x, y, z)^T \in \mathbb{R}^3 \mid x + y - z = 0\} = \textsc{Null}\left([1, 1, -1]\right)$.

   Solution: Note $\{(1, 1, -1)\}$ is a basis of $\mathbb{W}^\perp$ and $\left\{\frac{1}{\sqrt{2}}(1, -1, 0), \frac{1}{\sqrt{6}}(1, 1, 2)\right\}$ an orthonormal basis of $\mathbb{W}$. So,

   $$P_{\mathbb{W}^\perp} = \frac{1}{3}\begin{bmatrix}1 & 1 & -1\\ 1 & 1 & -1\\ -1 & -1 & 1\end{bmatrix} \text{ and } P_{\mathbb{W}} = \frac{1}{3}\begin{bmatrix}2 & -1 & 1\\ -1 & 2 & 1\\ 1 & 1 & 2\end{bmatrix}.$$

   Verify that $P_{\mathbb{W}} + P_{\mathbb{W}^\perp} = I_3$, Rank$(P_{\mathbb{W}^\perp}) = 2$ and Rank$(P_{\mathbb{W}}) = 1$.

3. $\mathbb{W} = LS(\ (1, 2, 1)\ ) = \textsc{Col}\left([1, 2, 1]^T\right) \subseteq \mathbb{R}^3$.

   Solution: Using Example 5.2.11.3 and Equation (5.2.1)

   $$\mathbb{W}^\perp = LS(\{(-2, 1, 0), (-1, 0, 1)\}) = LS(\{(-2, 1, 0), (1, 2, -5)\}).$$

   So, $P_{\mathbb{W}} = \frac{1}{6}\begin{bmatrix}1 & 2 & 1\\ 2 & 4 & 2\\ 1 & 2 & 1\end{bmatrix}$ and $P_{\mathbb{W}^\perp} = \frac{1}{6}\begin{bmatrix}5 & -2 & -1\\ -2 & 2 & -2\\ -1 & -2 & 5\end{bmatrix}$.

We advise the readers to give a proof of the next result.

Theorem 5.3.9. *Let $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ be an orthonormal basis of a subspace $\mathbb{W}$ of $\mathbb{R}^n$. If $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ is an extended orthonormal basis of $\mathbb{R}^n$, $P_{\mathbb{W}} = \sum\limits_{i=1}^{k}\mathbf{f}_i\mathbf{f}_i^T$ and $P_{\mathbb{W}^\perp} = \sum\limits_{i=k+1}^{n}\mathbf{f}_i\mathbf{f}_i^T$ then prove that*

1. $I_n - P_{\mathbb{W}} = P_{\mathbb{W}^\perp}$.
2. $(P_{\mathbb{W}})^T = P_{\mathbb{W}}$ *and* $(P_{\mathbb{W}^\perp})^T = P_{\mathbb{W}^\perp}$. *That is,* $P_{\mathbb{W}}$ *and* $P_{\mathbb{W}^\perp}$ *are symmetric.*

3. $(P_{\mathbb{W}})^2 = P_{\mathbb{W}}$ *and* $(P_{\mathbb{W}^\perp})^2 = P_{\mathbb{W}^\perp}$. *That is,* $P_{\mathbb{W}}$ *and* $P_{\mathbb{W}^\perp}$ *are idempotent.*

4. $P_{\mathbb{W}} \circ P_{\mathbb{W}^\perp} = P_{\mathbb{W}^\perp} \circ P_{\mathbb{W}} = \mathbf{0}$.

EXERCISE **5.3.10.** *1. Let* $\mathbb{W} = \{(x, y, z, w) \in \mathbb{R}^4 : x = y, z = w\}$ *be a subspace of* $\mathbb{R}^4$. *Determine the matrix of the orthogonal projection.*

2. *Let* $P_{\mathbb{W}_1}$ *and* $P_{\mathbb{W}_2}$ *be the orthogonal projections of* $\mathbb{R}^2$ *onto* $\mathbb{W}_1 = \{(x, 0) : x \in \mathbb{R}\}$ *and* $\mathbb{W}_2 = \{(x, x) : x \in \mathbb{R}\}$, *respectively. Note that* $P_{\mathbb{W}_1} \circ P_{\mathbb{W}_2}$ *is a projection onto* $\mathbb{W}_1$. *But, it is not an orthogonal projection. Hence or otherwise, conclude that the composition of two orthogonal projections need not be an orthogonal projection?*

3. *Let* $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. *Then, A is idempotent but not symmetric. Now, define* $P : \mathbb{R}^2 \to \mathbb{R}^2$ *by* $P(\mathbf{v}) = A\mathbf{v}$, *for all* $\mathbf{v} \in \mathbb{R}^2$. *Then,*

   (a) *P is idempotent.*

   (b) $\text{NULL}(P) \cap \text{RNG}(P) = \text{NULL}(A) \cap \text{COL}(A) = \{\mathbf{0}\}$.

   (c) $\mathbb{R}^2 = \text{NULL}(P) + \text{RNG}(P)$. *But,* $(\text{RNG}(P))^\perp = (\text{COL}(A))^\perp \neq \text{NULL}(A)$.

   (d) *Since* $(\text{COL}(A))^\perp \neq \text{NULL}(A)$, *the map P is not an orthogonal projector. In this case, P is called a projection of* $\mathbb{R}^2$ *onto* $\text{RNG}(P)$ *along* $\text{NULL}(P)$.

4. *Find all* 2×2 *real matrices A such that* $A^2 = A$. *Hence, or otherwise, determine all projection operators of* $\mathbb{R}^2$.

5. *Let* $\mathbb{W}$ *be an* $(n-1)$*-dimensional subspace of* $\mathbb{R}^n$ *with ordered basis* $\mathcal{B}_{\mathbb{W}} = [\mathbf{f}_1, \ldots, \mathbf{f}_{n-1}]$. *Suppose* $\mathcal{B} = [\mathbf{f}_1, \ldots, \mathbf{f}_{n-1}, \mathbf{f}_n]$ *is an orthogonal ordered basis of* $\mathbb{R}^n$ *obtained by extending* $\mathcal{B}_{\mathbb{W}}$. *Now, define a function* $Q : \mathbb{R}^n \to \mathbb{R}^n$ *by* $Q(\mathbf{v}) = \langle \mathbf{v}, \mathbf{f}_n \rangle \mathbf{f}_n - \sum\limits_{i=1}^{n-1} \langle \mathbf{v}, \mathbf{f}_i \rangle \mathbf{f}_i$. *Then,*

   (a) *Q fixes every vector in* $\mathbb{W}^\perp$.

   (b) *Q sends every vector* $\mathbf{w} \in \mathbb{W}$ *to* $-\mathbf{w}$.

   (c) $Q \circ Q = I_n$.

   The function Q is called the **reflection operator** *with respect to* $\mathbb{W}^\perp$.

5.3.1 Orthogonal Projections as Self-Adjoint Operators*

Theorem 5.3.9 implies that the matrix of the projection operator is symmetric. We use this idea to proceed further.

Definition 5.3.11. **[Self-Adjoint Operator]** Let $\mathbb{V}$ be an IPS with inner product $\langle \, , \, \rangle$. A linear operator $P : \mathbb{V} \to \mathbb{V}$ is called **self-adjoint** if $\langle P(\mathbf{v}), \mathbf{u} \rangle = \langle \mathbf{v}, P(\mathbf{u}) \rangle$, for every $\mathbf{u}, \mathbf{v} \in \mathbb{V}$.

A careful understanding of the examples given below shows that self-adjoint operators and Hermitian matrices are related. It also shows that the vector spaces $\mathbb{C}^n$ and $\mathbb{R}^n$ can be decomposed in terms of the null space and column space of Hermitian matrices. They also follow directly from the fundamental theorem of linear algebra.

Example 5.3.12. 1. Let A be an $n \times n$ real symmetric matrix. If $P : \mathbb{R}^n \to \mathbb{R}^n$ is defined by $P(\mathbf{x}) = A\mathbf{x}$, for every $\mathbf{x} \in \mathbb{R}^n$ then

(a) P is a self adjoint operator as $A = A^T$, for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, implies

$$\langle P(\mathbf{x}), \mathbf{y}\rangle = (\mathbf{y}^T)A\mathbf{x} = (\mathbf{y}^T)A^T\mathbf{x} = (A\mathbf{y})^T\mathbf{x} = \langle \mathbf{x}, A\mathbf{y}\rangle = \langle \mathbf{x}, P(\mathbf{y})\rangle.$$

(b) $\text{NULL}(P) = (\text{RNG}(P))^{\perp}$ as $A = A^T$. Thus, $\mathbb{R}^n = \text{NULL}(P) \oplus \text{RNG}(P)$.

2. Let A be an $n \times n$ Hermitian matrix. If $P : \mathbb{C}^n \to \mathbb{C}^n$ is defined by $P(\mathbf{z}) = A\mathbf{z}$, for all $\mathbf{z} \in \mathbb{C}^n$ then using similar arguments (see Example 5.3.12.1) prove the following:

   (a) P is a self-adjoint operator.

   (b) $\text{NULL}(P) = (\text{RNG}(P))^{\perp}$ as $A = A^*$. Thus, $\mathbb{C}^n = \text{NULL}(P) \oplus \text{RNG}(P)$.

We now state and prove the main result related with orthogonal projection operators.

Theorem 5.3.13. *Let $\mathbb{V}$ be a finite dimensional* IPS. *If $\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^{\perp}$ then the orthogonal projectors $P_{\mathbb{W}} : \mathbb{V} \to \mathbb{V}$ on $\mathbb{W}$ and $P_{\mathbb{W}^{\perp}} : \mathbb{V} \to \mathbb{V}$ on $\mathbb{W}^{\perp}$ satisfy*

1. $\text{NULL}(P_{\mathbb{W}}) = \{\mathbf{v} \in V : P_{\mathbb{W}}(\mathbf{v}) = \mathbf{0}\} = \mathbb{W}^{\perp} = \text{RNG}(P_{\mathbb{W}^{\perp}})$.

2. $\text{RNG}(P_{\mathbb{W}}) = \{P_{\mathbb{W}}(\mathbf{v}) : \mathbf{v} \in \mathbb{V}\} = \mathbb{W} = \text{NULL}(P_{\mathbb{W}^{\perp}})$.

3. $P_{\mathbb{W}} \circ P_{\mathbb{W}} = P_{\mathbb{W}}$, $P_{\mathbb{W}^{\perp}} \circ P_{\mathbb{W}^{\perp}} = P_{\mathbb{W}^{\perp}}$ *(*IDEMPOTENT*).*

4. $P_{\mathbb{W}^{\perp}} \circ P_{\mathbb{W}} = \mathbf{0}_{\mathbb{V}}$ *and* $P_{\mathbb{W}} \circ P_{\mathbb{W}^{\perp}} = \mathbf{0}_{\mathbb{V}}$, *where* $\mathbf{0}_{\mathbb{V}}(\mathbf{v}) = \mathbf{0}$, *for all* $\mathbf{v} \in \mathbb{V}$

5. $P_{\mathbb{W}} + P_{\mathbb{W}^{\perp}} = I_{\mathbb{V}}$, *where* $I_{\mathbb{V}}(\mathbf{v}) = \mathbf{v}$, *for all* $\mathbf{v} \in \mathbb{V}$.

6. *The operators $P_{\mathbb{W}}$ and $P_{\mathbb{W}^{\perp}}$ are self-adjoint.*

Proof. PART 1: As $\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^{\perp}$, for each $\mathbf{u} \in \mathbb{W}^{\perp}$, one uniquely writes $\mathbf{u} = \mathbf{0} + \mathbf{u}$, where $\mathbf{0} \in \mathbb{W}$ and $\mathbf{u} \in \mathbb{W}^{\perp}$. Hence, by definition, $P_{\mathbb{W}}(\mathbf{u}) = \mathbf{0}$ and $P_{\mathbb{W}^{\perp}}(\mathbf{u}) = \mathbf{u}$. Thus, $\mathbb{W}^{\perp} \subseteq \text{NULL}(P_{\mathbb{W}})$ and $\mathbb{W}^{\perp} \subseteq \text{RNG}(P_{\mathbb{W}^{\perp}})$.

Now suppose that $\mathbf{v} \in \text{NULL}(P_{\mathbb{W}})$. So, $P_{\mathbb{W}}(\mathbf{v}) = \mathbf{0}$. As $\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^{\perp}$, $\mathbf{v} = \mathbf{w} + \mathbf{u}$, for unique $\mathbf{w} \in W$ and unique $\mathbf{u} \in \mathbb{W}^{\perp}$. So, by definition, $P_{\mathbb{W}}(\mathbf{v}) = \mathbf{w}$. Thus, $\mathbf{w} = P_{\mathbb{W}}(\mathbf{v}) = \mathbf{0}$. That is, $\mathbf{v} = \mathbf{0} + \mathbf{u} = \mathbf{u} \in \mathbb{W}^{\perp}$. Thus, $\text{NULL}(P_{\mathbb{W}}) \subseteq \mathbb{W}^{\perp}$.

A similar argument implies $\text{RNG}(P_{\mathbb{W}^{\perp}}) \subseteq W^{\perp}$ and thus completing the proof of the first part.

PART 2: Use an argument similar to the proof of Part 1.

PART 3, PART 4 AND PART 5: Let $\mathbf{v} \in \mathbb{V}$. Then, $\mathbf{v} = \mathbf{w} + \mathbf{u}$, for unique $\mathbf{w} \in \mathbb{W}$ and unique $\mathbf{u} \in \mathbb{W}^{\perp}$. Thus, by definition,

$$\begin{aligned}
(P_{\mathbb{W}} \circ P_{\mathbb{W}})(\mathbf{v}) &= P_{\mathbb{W}}\big(P_{\mathbb{W}}(\mathbf{v})\big) = P_{\mathbb{W}}(\mathbf{w}) = \mathbf{w} \text{ and } P_{\mathbb{W}}(\mathbf{v}) = \mathbf{w} \\
(P_{\mathbb{W}^{\perp}} \circ P_{\mathbb{W}})(\mathbf{v}) &= P_{\mathbb{W}^{\perp}}\big(P_{\mathbb{W}}(\mathbf{v})\big) = P_{\mathbb{W}^{\perp}}(\mathbf{w}) = \mathbf{0} \text{ and} \\
(P_{\mathbb{W}} \oplus P_{\mathbb{W}^{\perp}})(\mathbf{v}) &= P_{\mathbb{W}}(\mathbf{v}) + P_{\mathbb{W}^{\perp}}(\mathbf{v}) = \mathbf{w} + \mathbf{u} = \mathbf{v} = I_{\mathbb{V}}(\mathbf{v}).
\end{aligned}$$

Hence, $P_{\mathbb{W}} \circ P_{\mathbb{W}} = P_{\mathbb{W}}$, $P_{\mathbb{W}^{\perp}} \circ P_{\mathbb{W}} = \mathbf{0}_{\mathbb{V}}$ and $I_{\mathbb{V}} = P_{\mathbb{W}} \oplus P_{\mathbb{W}^{\perp}}$.

PART 6: Let $\mathbf{u} = \mathbf{w}_1 + \mathbf{x}_1$ and $\mathbf{v} = \mathbf{w}_2 + \mathbf{x}_2$, for unique $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{W}$ and unique $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{W}^{\perp}$. Then, by definition, $\langle \mathbf{w}_i, \mathbf{x}_j\rangle = 0$ for $1 \le i, j \le 2$. Thus,

$$\langle P_{\mathbb{W}}(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{w}_1, \mathbf{v}\rangle = \langle \mathbf{w}_1, \mathbf{w}_2\rangle = \langle \mathbf{u}, \mathbf{w}_2\rangle = \langle \mathbf{u}, P_{\mathbb{W}}(\mathbf{v})\rangle$$

and the proof of the theorem is complete. ∎

Remark 5.3.14. *Theorem 5.3.13 gives us the following:*

1. *The orthogonal projectors $P_{\mathbb{W}}$ and $P_{\mathbb{W}^\perp}$ are idempotent and self-adjoint.*

2. *Let $\mathbf{v} \in \mathbb{V}$. Then, $\mathbf{v} - P_{\mathbb{W}}(\mathbf{v}) = (I_{\mathbb{V}} - P_{\mathbb{W}})(\mathbf{v}) = P_{\mathbb{W}^\perp}(\mathbf{v}) \in \mathbb{W}^\perp$. Thus, $\langle \mathbf{v} - P_{\mathbb{W}}(\mathbf{v}), \mathbf{w}\rangle = 0$, for every $\mathbf{v} \in \mathbb{V}$ and $\mathbf{w} \in \mathbb{W}$.*

3. *As $P_{\mathbb{W}}(\mathbf{v}) - \mathbf{w} \in \mathbb{W}$, for each $\mathbf{v} \in \mathbb{V}$ and $\mathbf{w} \in \mathbb{W}$, we have*

$$\begin{aligned}\|\mathbf{v}-\mathbf{w}\|^2 &= \|\mathbf{v} - P_{\mathbb{W}}(\mathbf{v}) + P_{\mathbb{W}}(\mathbf{v}) - \mathbf{w}\|^2 \\ &= \|\mathbf{v} - P_{\mathbb{W}}(\mathbf{v})\|^2 + \|P_{\mathbb{W}}(\mathbf{v}) - \mathbf{w}\|^2 + 2\langle \mathbf{v} - P_{\mathbb{W}}(\mathbf{v}), P_{\mathbb{W}}(\mathbf{v}) - \mathbf{w}\rangle \\ &= \|\mathbf{v} - P_{\mathbb{W}}(\mathbf{v})\|^2 + \|P_{\mathbb{W}}(\mathbf{v}) - \mathbf{w}\|^2.\end{aligned}$$

Therefore, $\|\mathbf{v}-\mathbf{w}\| \geq \|\mathbf{v} - P_{\mathbb{W}}(\mathbf{v})\|$ and equality holds if and only if $\mathbf{w} = P_{\mathbb{W}}(\mathbf{v})$. Since $P_{\mathbb{W}}(\mathbf{v}) \in \mathbb{W}$, we see that

$$d(\mathbf{v}, \mathbb{W}) = \inf\,\{\|\mathbf{v}-\mathbf{w}\| \ : \mathbf{w} \in W\} = \|\mathbf{v} - P_{\mathbb{W}}(\mathbf{v})\|.$$

That is, $P_{\mathbb{W}}(\mathbf{v})$ is the vector nearest to $\mathbf{v} \in \mathbb{W}$. This can also be stated as: the vector $P_{\mathbb{W}}(\mathbf{v})$ solves the following minimization problem:

$$\inf_{\mathbf{w}\in\mathbb{W}} \|\mathbf{v}-\mathbf{w}\| = \|\mathbf{v} - P_{\mathbb{W}}(\mathbf{v})\|.$$

The next theorem is a generalization of Theorem 5.3.13. We omit the proof as the arguments are similar and uses the following:
Let $\mathbb{V}$ be a finite dimensional IPS with $\mathbb{V} = \mathbb{W}_1 \oplus \cdots \oplus \mathbb{W}_k$, for certain subspaces $\mathbb{W}_i$'s of $\mathbb{V}$. Then, for each $\mathbf{v} \in \mathbb{V}$ there exist unique vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ such that

1. $\mathbf{v}_i \in \mathbb{W}_i$, for $1 \leq i \leq k$,

2. $\langle \mathbf{v}_i, \mathbf{v}_j\rangle = 0$ for each $\mathbf{v}_i \in \mathbb{W}_i, \mathbf{v}_j \in \mathbb{W}_j, 1 \leq i \neq j \leq k$ and

3. $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_k$.

Theorem 5.3.15. *Let $\mathbb{V}$ be a finite dimensional* IPS *with subspaces $\mathbb{W}_1, \ldots, \mathbb{W}_k$ of $\mathbb{V}$ such that $\mathbb{V} = W_1 \oplus \cdots \oplus \mathbb{W}_k$. Then, for each $i, j,\ 1 \leq i \neq j \leq k$, there exist orthogonal projectors $P_{\mathbb{W}_i} : \mathbb{V} \to \mathbb{V}$ of $\mathbb{V}$ onto $\mathbb{W}_i$ satisfying the following:*

1. $\textsc{Null}(P_{\mathbb{W}_i}) = W_i^\perp = \mathbb{W}_1 \oplus \mathbb{W}_2 \oplus \cdots \oplus W_{i-1} \oplus \mathbb{W}_{i+1} \oplus \cdots \oplus W_k$.

2. $\textsc{Rng}(P_{\mathbb{W}_i}) = \mathbb{W}_i$.

3. $P_{\mathbb{W}_i} \circ P_{\mathbb{W}_i} = P_{\mathbb{W}_i}$.

4. $P_{\mathbb{W}_i} \circ P_{\mathbb{W}_j} = \mathbf{0}_{\mathbb{V}}$.

5. *$P_{\mathbb{W}_i}$ is a self-adjoint operator, and*

6. $I_{\mathbb{V}} = P_{\mathbb{W}_1} \oplus P_{\mathbb{W}_2} \oplus \cdots \oplus P_{\mathbb{W}_k}$.

5.4 Orthogonal Operator and Rigid Motion

We now give the definition and a few properties of an orthogonal operator.

Definition 5.4.1. **[Orthogonal Operator]** Let $\mathbb{V}$ be a vector space. Then, a linear operator $T : \mathbb{V} \to \mathbb{V}$ is said to be an **orthogonal operator** if $\|T(\mathbf{x})\| = \|\mathbf{x}\|$, for all $\mathbf{x} \in \mathbb{V}$.

Example 5.4.2. Each $T \in \mathcal{L}(\mathbb{V})$ given below is an orthogonal operator.

1. Fix a unit vector $\mathbf{a} \in \mathbb{V}$ and define $T(\mathbf{x}) = 2\langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a} - \mathbf{x}$, for all $\mathbf{x} \in \mathbb{V}$.
   Solution: Note that $\text{Proj}_{\mathbf{a}}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a}$. So, $\big\langle \langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a},\ \mathbf{x} - \langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a} \big\rangle = 0$. Also, by Pythagoras theorem $\|\mathbf{x} - \langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a}\|^2 = \|\mathbf{x}\|^2 - (\langle \mathbf{x}, \mathbf{a}\rangle)^2$. Thus,
   $$\|T(\mathbf{x})\|^2 = \|(\langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a}) + (\langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a} - \mathbf{x})\|^2 = \|\langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a}\|^2 + \|\mathbf{x} - \langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a}\|^2 = \|\mathbf{x}\|^2.$$
2. Let $n = 2, \mathbb{V} = \mathbb{R}^2$ and $0 \le \theta < 2\pi$. Now define $T(\mathbf{x}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$.

We now show that an operator is orthogonal if and only if it preserves the angle.

Theorem 5.4.3. *Let $T \in \mathcal{L}(\mathbb{V})$. Then, the following statements are equivalent.*

1. *T is an orthogonal operator.*
2. *$\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$. That is, T preserves inner product.*

Proof. $1 \Rightarrow 2$ Let T be an orthogonal operator. Then, $\|T(\mathbf{x} + \mathbf{y})\|^2 = \|\mathbf{x} + \mathbf{y}\|^2$. So, $\|T(\mathbf{x})\|^2 + \|T(\mathbf{y})\|^2 + 2\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \|T(\mathbf{x}) + T(\mathbf{y})\|^2 = \|T(\mathbf{x} + \mathbf{y})\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\langle \mathbf{x}, \mathbf{y}\rangle$. Thus, using definition again $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$.

$2 \Rightarrow 1$ If $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ then T is an orthogonal operator as $\|T(\mathbf{x})\|^2 = \langle T(\mathbf{x}), T(\mathbf{x})\rangle = \langle \mathbf{x}, \mathbf{x}\rangle = \|\mathbf{x}\|^2$. ∎

As an immediate corollary, we obtain the following result.

Corollary 5.4.4. *Let $T \in \mathcal{L}(\mathbb{V})$. Then, T is an orthogonal operator if and only if "for every orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ of $\mathbb{V}$, $\{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)\}$ is an orthonormal basis of $\mathbb{V}$". Thus, if $\mathcal{B}$ is an orthonormal ordered basis of $\mathbb{V}$ then $T[\mathcal{B}, \mathcal{B}]$ is an orthogonal matrix.*

Definition 5.4.5. **[Isometry, Rigid Motion]** Let $\mathbb{V}$ be a vector space. Then, a map $T : \mathbb{V} \to \mathbb{V}$ is said to be an **isometry or a rigid motion** if $\|T(\mathbf{x}) - T(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$. That is, an isometry is distance preserving.

Observe that if T and S are two rigid motions then ST is also a rigid motion. Furthermore, it is clear from the definition that every rigid motion is invertible.

Example 5.4.6. The maps given below are rigid motions/isometry.

1. Let $\mathbb{V}$ be a linear space with norm $\|\cdot\|$. If $\mathbf{a} \in \mathbb{V}$ then the translation map $T_{\mathbf{a}} : \mathbb{V} \to \mathbb{V}$ (see Exercise 7), defined by $T_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} + \mathbf{a}$ for all $\mathbf{x} \in \mathbb{V}$, is an isometry/rigid motion as
   $$\|T_{\mathbf{a}}(\mathbf{x}) - T_{\mathbf{a}}(\mathbf{y})\| = \|\,(\mathbf{x} + \mathbf{a}) - (\mathbf{y} + \mathbf{a})\,\| = \|\mathbf{x} - \mathbf{y}\|.$$
2. Let $\mathbb{V}$ be an ips. Then, using Theorem 5.4.3, we see that every orthogonal operator is an isometry.

We now prove that every rigid motion that fixes origin is an orthogonal operator.

Theorem 5.4.7. *Let $\mathbb{V}$ be a real ips. Then, the following statements are equivalent for any map $T : \mathbb{V} \to \mathbb{V}$.*

1. *T is a rigid motion that fixes origin.*
2. *T is linear and $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ (preserves inner product).*
3. *T is an orthogonal operator.*

Proof. We have already seen the equivalence of Part 2 and Part 3 in Theorem 5.4.3. Let us now prove the equivalence of Part 1 and Part 2/Part 3.

If T is an orthogonal operator then $T(\mathbf{0}) = \mathbf{0}$ and $\|T(\mathbf{x}) - T(\mathbf{y})\| = \|T(\mathbf{x} - \mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$. This proves Part 3 implies Part 1.

We now prove Part 1 implies Part 2. So, let T be a rigid motion that fixes $\mathbf{0}$. Thus, $T(\mathbf{0}) = \mathbf{0}$ and $\|T(\mathbf{x}) - T(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$. Hence, in particular for $\mathbf{y} = \mathbf{0}$, we have $\|T(\mathbf{x})\| = \|\mathbf{x}\|$, for all $\mathbf{x} \in \mathbb{V}$. So,

$$\begin{aligned}
\|T(\mathbf{x})\|^2 + \|T(\mathbf{y})\|^2 - 2\langle T(\mathbf{x}), T(\mathbf{y})\rangle &= \langle T(\mathbf{x}) - T(\mathbf{y}), T(\mathbf{x}) - T(\mathbf{y})\rangle = \|T(\mathbf{x}) - T(\mathbf{y})\|^2 \\
&= \|\mathbf{x} - \mathbf{y}\|^2 = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y}\rangle \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle.
\end{aligned}$$

Thus, using $\|T(\mathbf{x})\| = \|\mathbf{x}\|$, for all $\mathbf{x} \in \mathbb{V}$, we get $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$. Now, to prove T is linear, we use $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ in 3-rd and 4-th line to get

$$\begin{aligned}
\|T(\mathbf{x}+\mathbf{y}) - (T(\mathbf{x}) + T(\mathbf{y}))\|^2 &= \langle T(\mathbf{x}+\mathbf{y}) - (T(\mathbf{x}) + T(\mathbf{y})), T(\mathbf{x}+\mathbf{y}) - (T(\mathbf{x}) + T(\mathbf{y}))\rangle \\
&= \langle T(\mathbf{x}+\mathbf{y}), T(\mathbf{x}+\mathbf{y})\rangle - 2\langle T(\mathbf{x}+\mathbf{y}), T(\mathbf{x})\rangle \\
&\quad -2\langle T(\mathbf{x}+\mathbf{y}), T(\mathbf{y})\rangle + \langle T(\mathbf{x}) + T(\mathbf{y}), T(\mathbf{x}) + T(\mathbf{y})\rangle \\
&= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle - 2\langle \mathbf{x}+\mathbf{y}, \mathbf{x}\rangle - 2\langle \mathbf{x}+\mathbf{y}, \mathbf{y}\rangle \\
&\quad +\langle T(\mathbf{x}), T(\mathbf{x})\rangle + 2\langle T(\mathbf{x}), T(\mathbf{y})\rangle + \langle T(\mathbf{y}), T(\mathbf{y})\rangle \\
&= -\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{x}\rangle + 2\langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle = 0.
\end{aligned}$$

Thus, $T(\mathbf{x}+\mathbf{y}) - (T(\mathbf{x}) + T(\mathbf{y})) = \mathbf{0}$ and hence $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$. A similar calculation gives $T(\alpha\mathbf{x}) = \alpha T(\mathbf{x})$ and hence T is linear. ∎

EXERCISE **5.4.8.** 1. *Let $A, B \in \mathbb{M}_n(\mathbb{C})$. Then, A and B are said to be*

   (a) **Orthogonally Congruent** *if $B = S^T AS$, for some invertible matrix S.*

   (b) **Unitarily Congruent** *if $B = S^* AS$, for some invertible matrix S.*

   Prove that Orthogonal and Unitary congruences are equivalence relations on $\mathbb{M}_n(\mathbb{R})$ and $\mathbb{M}_n(\mathbb{C})$, respectively.

2. *Let $\mathbf{x} \in \mathbb{C}^2$. Identify it with the complex number $\mathbf{x} = \mathbf{x}_1 + i\mathbf{x}_2$. If we rotate $\mathbf{x}$ by a counterclockwise rotation $\theta, 0 \le \theta < 2\pi$ then, we have*

$$\mathbf{x}e^{i\theta} = (\mathbf{x}_1 + i\mathbf{x}_2)(\cos\theta + i\sin\theta) = \mathbf{x}_1\cos\theta - \mathbf{x}_2\sin\theta + i[\mathbf{x}_1\sin\theta + \mathbf{x}_2\cos\theta].$$

Thus, the corresponding vector in $\mathbb{R}^2$ *is*

$$\begin{bmatrix} x_1\cos\theta - x_2\sin\theta \\ x_1\sin\theta + x_2\cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Is the matrix, $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$*, the matrix of the corresponding rotation? Justify.*

3. *Let* $A \in M_2(\mathbb{R})$ *and* $T(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$*, for* $\theta \in \mathbb{R}$*. Then, A is an orthogonal matrix if and only if* $A = T(\theta)$ *or* $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} T(\theta)$*, for some* $\theta \in \mathbb{R}$*.*

4. *Let* $A \in \mathbb{M}_n(\mathbb{C})$*. Then, the following statements are equivalent.*

   (a) *A is an orthogonal matrix.*

   (b) $A^{-1} = A^T$.

   (c) A^T *is orthogonal.*

   (d) *the columns of A form an orthonormal basis of the real vector space* $\mathbb{R}^n$*.*

   (e) *the rows of A form an orthonormal basis of the real vector space* $\mathbb{R}^n$*.*

   (f) *for any two vectors* $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$*,* $\langle A\mathbf{x}, A\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ Orthogonal matrices preserve angle.

   (g) *for any vector* $\mathbf{x} \in \mathbb{C}^n$*,* $\|A\mathbf{x}\| = \|\mathbf{x}\|$ Orthogonal matrices preserve length.

5. *Let U be an* $n \times n$ *matrix. Then, prove that the following statements are equivalent.*

   (a) *U is a unitary matrix.*

   (b) $U^{-1} = U^*$.

   (c) U^* *is unitary.*

   (d) *the columns of U form an orthonormal basis of the complex vector space* $\mathbb{C}^n$*.*

   (e) *the rows of U form an orthonormal basis of the complex vector space* $\mathbb{C}^n$*.*

   (f) *for any two vectors* $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$*,* $\langle U\mathbf{x}, U\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ Unitary matrices preserve angle.

   (g) *for any vector* $\mathbf{x} \in \mathbb{C}^n$*,* $\|U\mathbf{x}\| = \|\mathbf{x}\|$ Unitary matrices preserve length.

6. *Let A be an* $n \times n$ *orthogonal matrix. Then, prove that* $\det(A) = \pm 1$*.*

7. *Let A be an* $n \times n$ *upper triangular matrix. If A is also an orthogonal matrix then A is a diagonal matrix with diagonal entries* ± 1*.*

8. *Prove that in* $M_5(\mathbb{R})$*, there are infinitely many orthogonal matrices of which only finitely many are diagonal (in fact, there number is just* 32*).*

9. *Prove that permutation matrices are real orthogonal.*

10. *Let* $A, B \in \mathbb{M}_n(\mathbb{C})$ *be two unitary matrices. Then, prove that* AB *and* BA *are unitary matrices.*

11. *If $A = [a_{ij}]$ and $B = [b_{ij}]$ are unitarily equivalent then prove that* $\sum\limits_{ij} |a_{ij}|^2 = \sum\limits_{ij} |b_{ij}|^2$.

12. *Let U be a unitary matrix and for every $\mathbf{x} \in \mathbb{C}^n$, define*

$$\|\mathbf{x}\|_1 = \max\{|\mathbf{x}_i| : \mathbf{x}^T = [\mathbf{x}_1, \ldots, \mathbf{x}_n]\}.$$

Then, is it necessary that $\|U\mathbf{x}\|_1 = \|\mathbf{x}\|_1$?

5.5 Summary

In the previous chapter, we learnt that if $\mathbb{V}$ is vector space over $\mathbb{F}$ with $\dim(\mathbb{V}) = n$ then $\mathbb{V}$ basically looks like $\mathbb{F}^n$. Also, any subspace of $\mathbb{F}^n$ is either $\textsc{Col}(A)$ or $\textsc{Null}(A)$ or both, for some matrix A with entries from $\mathbb{F}$.

So, we started this chapter with inner product, a generalization of the dot product in $\mathbb{R}^3$ or $\mathbb{R}^n$. We used the inner product to define the length/norm of a vector. The norm has the property that "the norm of a vector is zero if and only if the vector itself is the zero vector". We then proved the Cauchy-Bunyakovskii-Schwartz Inequality which helped us in defining the angle between two vector. Thus, one can talk of geometrical problems in $\mathbb{R}^n$ and proved some geometrical results.

We then independently defined the notion of a norm in $\mathbb{R}^n$ and showed that a norm is induced by an inner product if and only if the norm satisfies the parallelogram law (sum of squares of the diagonal equals twice the sum of square of the two non-parallel sides).

The next subsection dealt with the fundamental theorem of linear algebra where we showed that if $A \in \mathbb{M}_{m,n}(\mathbb{C})$ then

1. $\dim(\textsc{Null}(A)) + \dim(\textsc{Col}(A)) = n$.
2. $\textsc{Null}(A) = \big(\textsc{Col}(A^*)\big)^{\perp}$ and $\textsc{Null}(A^*) = \big(\textsc{Col}(A)\big)^{\perp}$.
3. $\dim(\textsc{Col}(A)) = \dim(\textsc{Col}(A^*))$.

We then saw that having an orthonormal basis is an asset as determining the

1. coordinates of a vector boils down to computing the inner product.
2. projection of a vector on a subspace boils down to finding an orthonormal basis of the subspace and then summing the corresponding rank 1 matrices.

So, the question arises, how do we compute an orthonormal basis? This is where we came across the Gram-Schmidt Orthonormalization process. This algorithm helps us to determine an orthonormal basis of $LS(S)$ for any finite subset S of a vector space. This also lead to the QR-decomposition of a matrix.

Thus, we observe the following about the linear system $A\mathbf{x} = \mathbf{b}$. If

1. $\mathbf{b} \in \textsc{Col}(A)$ then we can use the Gauss-Jordan method to get a solution.
2. $\mathbf{b} \notin \textsc{Col}(A)$ then in most cases we need a vector $\mathbf{x}$ such that the least square error between $\mathbf{b}$ and $A\mathbf{x}$ is minimum. We saw that this minimum is attained by the projection of $\mathbf{b}$ on $\textsc{Col}(A)$. Also, this vector can be obtained either using the fundamental theorem of linear algebra or by computing the matrix $B(B^TB)^{-1}B^T$, where the columns of B are either the pivot columns of A or a basis of $\textsc{Col}(A)$.

Chapter 6

Eigenvalues, Eigenvectors and Diagonalizability

6.1 Introduction and Definitions

In this chapter, every matrix is an element of $\mathbb{M}_n(\mathbb{C})$ and $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$, for some $n \in \mathbb{N}$. We start with a few examples to motivate this chapter.

Example 6.1.1. 1. Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$.

(a) Then A magnifies the nonzero vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ three times as $A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and behaves by changing the direction of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ as $A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = -1\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Further, the vectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are orthogonal.

(b) B magnifies both the vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ as $B\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $B\begin{bmatrix} 2 \\ -1 \end{bmatrix} = 10\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Here again, the vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ are orthogonal.

(c) $\mathbf{x}^T A\mathbf{x} = 3\dfrac{(x+y)^2}{2} - \dfrac{(x-y)^2}{2}$. Here, the displacements occur along perpendicular lines $x+y=0$ and $x-y=0$, where $x+y = (x,y)\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $x-y = (x,y)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Whereas, $\mathbf{x}^T B\mathbf{x} = 5\dfrac{(x+2y)^2}{5} + 10\dfrac{(2x-y)^2}{5}$. Here also the maximum/minimum displacements occur along the orthogonal lines $x+2y=0$ and $2x-y=0$, where $x+2y = (x,y)\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $2x-y = (x,y)\begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

(d) the curve $\mathbf{x}^T A\mathbf{x} = 10$ represents a hyperbola, where as the curve $\mathbf{x}^T B\mathbf{x} = 10$ represents an ellipse (see Figure 6.1 drawn using the package "Sagemath").

2. Let $C = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$, a non-symmetric matrix. Then, does there exist a nonzero $\mathbf{x} \in \mathbb{C}^2$ which

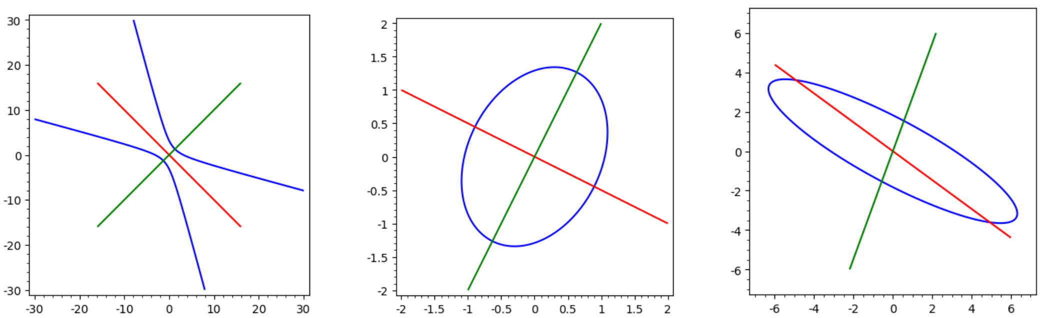


Figure 6.1: A Hyperbola and two Ellipses (first one has orthogonal axes)

.

gets magnified by C?

So, we need $\mathbf{x} \neq \mathbf{0}$ and $\alpha \in \mathbb{C}$ such that $C\mathbf{x} = \alpha\mathbf{x} \Leftrightarrow [\alpha I_2 - C]\mathbf{x} = \mathbf{0}$. As $\mathbf{x} \neq 0$, $[\alpha I_2 - C]\mathbf{x} = \mathbf{0}$ has a solution if and only if $\det[\alpha I - A] = 0$. But,

$$\det[\alpha I - A] = \det\left(\begin{bmatrix} \alpha - 1 & -2 \\ -1 & \alpha - 3 \end{bmatrix}\right) = \alpha^2 - 4\alpha + 1.$$

So, $\alpha = 2 \pm \sqrt{3}$. For $\alpha = 2 + \sqrt{3}$, verify that the $\mathbf{x} \neq \mathbf{0}$ that satisfies $\begin{bmatrix} 1+\sqrt{3} & -2 \\ -1 & \sqrt{3}-1 \end{bmatrix}\mathbf{x} = \mathbf{0}$ equals $\mathbf{x} = \begin{bmatrix} \sqrt{3}-1 \\ 1 \end{bmatrix}$. Similarly, for $\alpha = 2 - \sqrt{3}$, the vector $\mathbf{x} = \begin{bmatrix} \sqrt{3}+1 \\ -1 \end{bmatrix}$ satisfies $\begin{bmatrix} 1-\sqrt{3} & -2 \\ -1 & -\sqrt{3}-1 \end{bmatrix}\mathbf{x} = \mathbf{0}$. In this example,

(a) we still have magnifications in the directions $\begin{bmatrix} \sqrt{3}-1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} \sqrt{3}+1 \\ -1 \end{bmatrix}$.

(b) the maximum/minimum displacements do not occur along the lines $(\sqrt{3}-1)x + y = 0$ and $(\sqrt{3}+1)x - y = 0$ (see the third curve in Figure 6.1).

(c) the lines $(\sqrt{3}-1)x + y = 0$ and $(\sqrt{3}+1)x - y = 0$ are not orthogonal.

3. Let A be a real symmetric matrix. Consider the following problem:

$$\text{Maximize (Minimize)} \;\; \mathbf{x}^T A\mathbf{x} \text{ such that } \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{x}^T\mathbf{x} = 1.$$

To solve this, consider the Lagrangian

$$L(\mathbf{x}, \lambda) = \mathbf{x}^T A\mathbf{x} - \lambda(\mathbf{x}^T\mathbf{x} - 1) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j - \lambda\left(\sum_{i=1}^{n} x_i^2 \; - 1\right).$$

Partially differentiating $L(\mathbf{x}, \lambda)$ with respect to x_i for $1 \le i \le n$, we get

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= 2a_{11}x_1 + 2a_{12}x_2 + \cdots + 2a_{1n}x_n - 2\lambda x_1, \\ \vdots &= \vdots \\ \frac{\partial L}{\partial x_n} &= 2a_{n1}x_1 + 2a_{n2}x_2 + \cdots + 2a_{nn}x_n - 2\lambda x_n. \end{aligned}$$

Therefore, to get the points of extremum, we solve for

$$\mathbf{0}^T = \left(\frac{\partial L}{\partial x_1}, \frac{\partial L}{\partial x_2}, \ldots, \frac{\partial L}{\partial x_n}\right)^T = \frac{\partial L}{\partial \mathbf{x}} = 2(A\mathbf{x} - \lambda\mathbf{x}).$$

Thus, to solve the extremal problem, we need $\lambda \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} \neq \mathbf{0}$ and $A\mathbf{x} = \lambda\mathbf{x}$.

We observe the following about the matrices A, B and C that appear in Example 6.1.1.

1. $\det(A) = -3 = 3 \times -1$, $\det(B) = 50 = 5 \times 10$ and $\det(C) = 1 = (2+\sqrt{3}) \times (2-\sqrt{3})$.
2. $\mathsf{tr(A)} = 2 = 3 - 1$, $\mathsf{tr(B)} = 15 = 5 + 10$ and $\det(C) = 4 = (2+\sqrt{3}) + (2-\sqrt{3})$.
3. The sets $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right\}$, $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\-1\end{bmatrix}\right\}$ and $\left\{\begin{bmatrix}\sqrt{3}-1\\1\end{bmatrix}, \begin{bmatrix}\sqrt{3}+1\\-1\end{bmatrix}\right\}$ are linearly independent.
4. If $\mathbf{v}_1 = \begin{bmatrix}1\\1\end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix}1\\-1\end{bmatrix}$ and $S = [\mathbf{v}_1, \mathbf{v}_2]$ then
   (a) $AS = [A\mathbf{v}_1, A\mathbf{v}_2] = [3\mathbf{v}_1, -\mathbf{v}_2] = S\begin{bmatrix}3 & 0\\0 & -1\end{bmatrix} \Leftrightarrow S^{-1}AS = \begin{bmatrix}3 & 0\\0 & -1\end{bmatrix} = \operatorname{diag}(3,-1)$.
   (b) Let $\mathbf{u}_1 = \frac{1}{\sqrt{2}}\mathbf{v}_1$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\mathbf{v}_2$. Then, $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthonormal unit vectors, *i.e.*, if $U = [\mathbf{u}_1, \mathbf{u}_2]$ then $I = UU^* = \mathbf{u}_1\mathbf{u}_1^* + \mathbf{u}_2\mathbf{u}_2^*$ and $A = 3\mathbf{u}_1\mathbf{u}_1^* - \mathbf{u}_2\mathbf{u}_2^*$.
5. If $\mathbf{v}_1 = \begin{bmatrix}1\\2\end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix}2\\-1\end{bmatrix}$ and $S = [\mathbf{v}_1, \mathbf{v}_2]$ then
   (a) $AS = [A\mathbf{v}_1, A\mathbf{v}_2] = [5\mathbf{v}_1, 10\mathbf{v}_2] = S\begin{bmatrix}5 & 0\\0 & 10\end{bmatrix} \Leftrightarrow S^{-1}AS = \begin{bmatrix}5 & 0\\0 & 10\end{bmatrix} = \operatorname{diag}(3,-1)$.
   (b) Let $\mathbf{u}_1 = \frac{1}{\sqrt{5}}\mathbf{v}_1$ and $\mathbf{u}_2 = \frac{1}{\sqrt{5}}\mathbf{v}_2$. Then, $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthonormal unit vectors, *i.e.*, if $U = [\mathbf{u}_1, \mathbf{u}_2]$ then $I = UU^* = \mathbf{u}_1\mathbf{u}_1^* + \mathbf{u}_2\mathbf{u}_2^*$ and $A = 5\mathbf{u}_1\mathbf{u}_1^* + 10\mathbf{u}_2\mathbf{u}_2^*$.
6. If $\mathbf{v}_1 = \begin{bmatrix}\sqrt{3}-1\\1\end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix}\sqrt{3}+1\\-1\end{bmatrix}$ and $S = [\mathbf{v}_1, \mathbf{v}_2]$ then

$$S^{-1}CS = \begin{bmatrix}2+\sqrt{3} & 0\\0 & 2-\sqrt{3}\end{bmatrix} = \operatorname{diag}(2+\sqrt{3}, 2-\sqrt{3}).$$

Thus, we see that given $A \in \mathbb{M}_n(\mathbb{C})$, the number $\lambda \in \mathbb{C}$ and $\mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq \mathbf{0}$ satisfying $A\mathbf{x} = \lambda\mathbf{x}$ have certain nice properties. For example, there exists a basis of $\mathbb{C}^2$ in which the matrices A, B and C behave like diagonal matrices. To understand the ideas better, we start with the following definitions.

Definition 6.1.2. **[Eigenvalues, Eigenvectors and Eigenspace]** Let $A \in \mathbb{M}_n(\mathbb{C})$. Then,

1. the equation
$$A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow (A - \lambda I_n)\mathbf{x} = \mathbf{0} \tag{6.1.1}$$
   is called the **eigen-condition**.

2. an $\alpha \in \mathbb{C}$ is called a **characteristic value/root** or **eigenvalue** or **latent root** of A if there exists $\mathbf{x} \neq \mathbf{0}$ satisfying $A\mathbf{x} = \alpha\mathbf{x}$.
3. an $\mathbf{x} \neq \mathbf{0}$ satisfying Equation (6.1.1) is called a **characteristic vector** or **eigenvector** or **invariant/latent vector** of A corresponding to λ.
4. the tuple $(\alpha, \mathbf{x})$ with $\mathbf{x} \neq \mathbf{0}$ and $A\mathbf{x} = \alpha\mathbf{x}$ is called an **eigen-pair** or **characteristic-pair**.
5. for an eigenvalue $\alpha \in \mathbb{C}$, $\text{N}\textsc{ull}(A - \alpha I) = \{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = \alpha\mathbf{x}\}$ is called the **eigenspace** or **characteristic vector space** of A corresponding to α.

Theorem 6.1.3. *Let $A \in \mathbb{M}_n(\mathbb{C})$ and $\alpha \in \mathbb{C}$. Then, the following statements are equivalent.*

1. *α is an eigenvalue of A.*
2. $\det(A - \alpha I_n) = 0$.

Proof. We know that α is an eigenvalue of A if any only if the system $(A - \alpha I_n)\mathbf{x} = \mathbf{0}$ has a non-trivial solution. By Theorem 2.2.40 this holds if and only if $\det(A - \alpha I) = 0$. ∎

Definition 6.1.4. **[Characteristic Polynomial / Equation, Spectrum and Spectral Radius]** Let $A \in \mathbb{M}_n(\mathbb{C})$. Then,

1. $\det(A - \lambda I)$ is a polynomial of degree n in λ and is called the **characteristic polynomial** of A, denoted $P_A(\lambda)$, or in short $P(\lambda)$.
2. the equation $P_A(\lambda) = 0$ is called the **characteristic equation** of A.
3. The multi-set (collection with multiplicities) $\{\alpha \in \mathbb{C} : P_A(\alpha) = 0\}$ is called the **spectrum** of A, denoted $\sigma(A)$. Hence, $\sigma(A)$ contains all the eigenvalues of A.
4. The **Spectral Radius**, denoted $\rho(A)$ of $A \in \mathbb{M}_n(\mathbb{C})$, equals $\max\{|\alpha| : \alpha \in \sigma(A)\}$.

We thus observe the following.

Remark 6.1.5. *Let $A \in \mathbb{M}_n(\mathbb{C})$.*

1. *Then, A is singular if and only if $0 \in \sigma(A)$.*
2. *Further, if $\alpha \in \sigma(A)$ then the following statements hold.*
   (a) *$\{\mathbf{0}\} \subsetneqq \text{N}\textsc{ull}(A - \alpha I)$. Therefore, if $\text{R}\textsc{ank}(A - \alpha I) = r$ then $r < n$. Hence, by Theorem 2.2.40, the system $(A - \alpha I)\mathbf{x} = \mathbf{0}$ has $n - r$ linearly independent solutions.*
   (b) *$\mathbf{x} \in \text{N}\textsc{ull}(A - \alpha I)$ if and only if $c\mathbf{x} \in \text{N}\textsc{ull}(A - \alpha I)$, for $c \neq 0$.*
   (c) *If $\mathbf{x}_1, \ldots, \mathbf{x}_r \in \text{N}\textsc{ull}(A - \alpha I)$ are linearly independent then $\sum\limits_{i=1}^{r} c_i\mathbf{x}_i \in \text{N}\textsc{ull}(A - \alpha I)$, for all $c_i \in \mathbb{C}$. Hence, if S is a collection of eigenvectors then, we necessarily want the set S to be* LINEARLY INDEPENDENT.
   (d) *Thus, an eigenvector $\mathbf{v}$ of A is in some sense a line $\ell = Span(\{\mathbf{v}\})$ that passes through $\mathbf{0}$ and $\mathbf{v}$ and has the property that the image of ℓ is either ℓ itself or $\mathbf{0}$.*
3. *Since the eigenvalues of A are roots of the characteristic equation, A has exactly n eigenvalues, including multiplicities.*
4. *If the entries of A are real and $\alpha \in \sigma(A)$ is also real then the corresponding eigenvector has real entries.*

5. *Further, if $(\alpha, \mathbf{x})$ is an eigenpair for A and $f(A) = b_0 I + b_1 A + \cdots + b_k A^k$ is a polynomial in A then $(f(\alpha), \mathbf{x})$ is an eigenpair for $f(A)$.*

Almost all books in mathematics differentiate between characteristic value and eigenvalue as the ideas change when one moves from complex numbers to any other scalar field. We give the following example for clarity.

Remark 6.1.6. *Let $A \in \mathbb{M}_2(\mathbb{F})$. Then, A induces a map $T \in \mathcal{L}(\mathbb{F}^2)$ defined by $T(\mathbf{x}) = A\mathbf{x}$, for all $\mathbf{x} \in \mathbb{F}^2$. We use this idea to understand the difference.*

1. *Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then, $p_A(\lambda) = \lambda^2 + 1$. So, $\pm i$ are the roots of $P(\lambda) = 0$ in $\mathbb{C}$. Hence,*
   (a) *A has $(i, (1, i)^T)$ and $(-i, (i, 1)^T)$ as eigen-pairs or characteristic-pairs.*
   (b) *A has no characteristic value over $\mathbb{R}$.*
2. *Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$. Then, $2 \pm \sqrt{3}$ are the roots of the characteristic equation. Hence,*
   (a) *A has characteristic values or eigenvalues over $\mathbb{R}$.*
   (b) *A has no characteristic value over $\mathbb{Q}$.*

Let us look at some more examples.

Example 6.1.7. 1. Let $A = \operatorname{diag}(d_1, \ldots, d_n)$ with $d_i \in \mathbb{C}, 1 \le i \le n$. Then, $p(\lambda) = \prod\limits_{i=1}^{n} (\lambda - d_i)$ and thus verify that $(d_1, \mathbf{e}_1), \ldots, (d_n, \mathbf{e}_n)$ are the eigen-pairs.

2. Let $A = (a_{ij})$ be an $n \times n$ triangular matrix. Then, $p(\lambda) = \prod\limits_{i=1}^{n} (\lambda - a_{ii})$ and thus verify that $\sigma(A) = \{a_{11}, a_{22}, \ldots, a_{nn}\}$. What can you say about the eigen-vectors of an upper triangular matrix if the diagonal entries are all distinct?

3. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Then, $p(\lambda) = (1 - \lambda)^2$. Hence, $\sigma(A) = \{1, 1\}$. But the complete solution of the system $(A - I_2)\mathbf{x} = \mathbf{0}$ equals $\mathbf{x} = c\mathbf{e}_1$, for $c \in \mathbb{C}$. Hence using Remark 6.1.5.2, $\mathbf{e}_1$ is an eigenvector. Therefore, 1 IS A REPEATED EIGENVALUE WHEREAS THERE IS ONLY ONE EIGENVECTOR.

4. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then, 1 is a repeated eigenvalue of A. In this case, $(A - I_2)\mathbf{x} = \mathbf{0}$ has a solution for every $\mathbf{x} \in \mathbb{C}^2$. Hence, any two LINEARLY INDEPENDENT vectors $\mathbf{x}^T, \mathbf{y}^T \in \mathbb{C}^2$ gives $(1, \mathbf{x})$ and $(1, \mathbf{y})$ as the two eigen-pairs for A. In general, if $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is a basis of $\mathbb{C}^n$ then $(1, \mathbf{x}_1), \ldots, (1, \mathbf{x}_n)$ are eigen-pairs of I_n, the identity matrix.

5. Let $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Then, $\left(1 + i, \begin{bmatrix} i \\ 1 \end{bmatrix}\right)$ and $\left(1 - i, \begin{bmatrix} 1 \\ i \end{bmatrix}\right)$ are the eigen-pairs of A.

6. Let $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. Then, $\sigma(A) = \{0, 0, 0\}$ with $\mathbf{e}_1$ as the only eigenvector.

7. Let $A = \left[\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$. Then, $\sigma(A) = \{0,0,0,0,0\}$. Note that $A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \mathbf{0}$ implies $x_2 = 0 = x_3 = x_5$. Thus, $\mathbf{e}_1$ and $\mathbf{e}_4$ are the only eigenvectors. Note that the diagonal blocks of A are nilpotent matrices.

Exercise **6.1.8.** *1. Let $A \in \mathbb{M}_n(\mathbb{R})$. Then, prove that*

(a) if $\alpha \in \sigma(A)$ then $\alpha^k \in \sigma(A^k)$, for all $k \in \mathbb{N}$.

(b) if A is invertible and $\alpha \in \sigma(A)$ then $\alpha^k \in \sigma(A^k)$, for all $k \in \mathbb{Z}$.

2. Find eigen-pairs over $\mathbb{C}$, for each of the following matrices:
$\begin{bmatrix} 1 & 1+i \\ 1-i & 1 \end{bmatrix}$, $\begin{bmatrix} i & 1+i \\ -1+i & i \end{bmatrix}$, $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ *and* $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$.

3. Let $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{C})$ with $\sum\limits_{j=1}^{n} a_{ij} = a$, for all $1 \le i \le n$. Then, prove that a is an eigenvalue of A with corresponding eigenvector $\mathbf{1} = [1,1,\ldots,1]^T$.

4. Prove that the matrices A and A^T have the same set of eigenvalues. Construct a 2×2 matrix A such that the eigenvectors of A and A^T are different.

5. Prove that $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if $\overline{\lambda} \in \mathbb{C}$ is an eigenvalue of A^*.

6. Let A be an idempotent matrix. Then, prove that its eigenvalues are either 0 or 1 or both.

7. Let A be a nilpotent matrix. Then, prove that its eigenvalues are all 0.

8. Let $J = \mathbf{1}\mathbf{1}^T \in \mathbb{M}_n(\mathbb{C})$. Then, J is a matrix with each entry 1. Show that

(a) $(n, \mathbf{1})$ is an eigenpair for J.

(b) $0 \in \sigma(J)$ with multiplicity $n-1$. Find a set of $n-1$ linearly independent eigenvectors for $0 \in \sigma(J)$.

9. Let $B \in \mathbb{M}_n(\mathbb{C})$ and $C \in \mathbb{M}_m(\mathbb{C})$. Now, define the **Direct Sum** *$B \oplus C = \begin{bmatrix} B & \mathbf{0} \\ \mathbf{0} & C \end{bmatrix}$. Then, prove that*

(a) if $(\alpha, \mathbf{x})$ is an eigen-pair for B then $\left(\alpha, \begin{bmatrix} \mathbf{x} \\ \mathbf{0} \end{bmatrix}\right)$ is an eigen-pair for $B \oplus C$.

(b) if $(\beta, \mathbf{y})$ is an eigen-pair for C then $\left(\beta, \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix}\right)$ is an eigen-pair for $B \oplus C$.

Definition 6.1.9. Let $A \in \mathcal{L}(\mathbb{C}^n)$. Then, a vector $\mathbf{y} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ satisfying $\mathbf{y}^* A = \lambda \mathbf{y}^*$ is called a **left eigenvector** of A for λ.

Example 6.1.10. 1. Let $A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$. Then, $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a left eigenvector of A corresponding to the eigenvalue 0 and $\left(0, \mathbf{y} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$ is a (right) eigenpair of A.

2. Let $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$. Then, $\left(0, \mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$ and $\left(3, \mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$ are (right) eigen-pairs of A. Also, $\left(3, \mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$ and $\left(0, \mathbf{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right)$ are left eigen-pairs of A. Note that $\mathbf{x}$ is orthogonal to $\mathbf{u}$ and $\mathbf{y}$ is orthogonal to $\mathbf{v}$. This is true in general and is proved next.

3. Let S be a nonsingular matrix such that its columns are left eigenvectors of A. Then, prove that the columns of $(S^*)^{-1}$ are right eigenvectors of A.

Theorem 6.1.11. **[Principle of bi-orthogonality]** *Let $(\lambda, \mathbf{x})$ be a (right) eigenpair and $(\mu, \mathbf{y})$ be a left eigenpair of A, where $\lambda \neq \mu$. Then, $\mathbf{y}$ is orthogonal to $\mathbf{x}$.*

Proof. Verify that $\mu \mathbf{y}^* \mathbf{x} = (\mathbf{y}^* A)\mathbf{x} = \mathbf{y}^*(\lambda \mathbf{x}) = \lambda \mathbf{y}^* \mathbf{x}$. Thus, $\mathbf{y}^* \mathbf{x} = 0$. ∎

EXERCISE **6.1.12.** *Let $A\mathbf{x} = \lambda \mathbf{x}$ and $\mathbf{x}^* A = \mu \mathbf{x}^*$. Then $\mu = \lambda$.*

Definition 6.1.13. **[Eigenvalues of a linear Operator]** Let $T \in \mathcal{L}(\mathbb{C}^n)$. Then, $\alpha \in \mathcal{C}$ is called an **eigenvalue** of T if there exists $\mathbf{v} \in \mathbb{C}^n$ with $\mathbf{v} \neq \mathbf{0}$ such that $T(\mathbf{v}) = \alpha \mathbf{v}$.

Proposition 6.1.14. *Let $T \in L(\mathbb{C}^n)$ and let $\mathcal{B}$ be an ordered basis in $\mathbb{C}^n$. Then, $(\alpha, \mathbf{v})$ is an eigenpair for T if and only if $(\alpha, [\mathbf{v}]_{\mathcal{B}})$ is an eigenpair of $A = T[\mathcal{B}, \mathcal{B}]$.*

Proof. Note that, by definition, $T(\mathbf{v}) = \alpha \mathbf{v}$ if and only if $[Tv]_{\mathcal{B}} = [\alpha \mathbf{v}]_{\mathcal{B}}$. Or equivalently, $\alpha \in \sigma(T)$ if and only if $A[\mathbf{v}]_{\mathcal{B}} = \alpha [\mathbf{v}]_{\mathcal{B}}$. Thus, the required result follows. ∎

Remark 6.1.15. **[A linear operator on an infinite dimensional space may not have any eigenvalue]** *Let $\mathbb{V}$ be the space of all real sequences (see Example 3.1.4.8a). Now, define a linear operator $T \in \mathcal{L}(\mathbb{V})$ by*

$$T\,(a_0, a_1, \ldots) = (0, a_1, a_2, \ldots).$$

We now show that T doesn't have any eigenvalue.

Solution: *Let if possible α be an eigenvalue of T with corresponding eigenvector $\mathbf{x} = (x_1, x_2, \ldots)$. Then, the eigen-condition $T(\mathbf{x}) = \alpha \mathbf{x}$ implies that*

$$(0, x_1, x_2, \ldots) = \alpha(x_1, x_2, \ldots) = (\alpha x_1, \alpha x_2, \ldots).$$

So, if $\alpha \neq 0$ then $x_1 = 0$ and this in turn implies that $\mathbf{x} = \mathbf{0}$, a contradiction. If $\alpha = 0$ then $(0, x_1, x_2, \ldots) = (0, 0, \ldots)$ and we again get $\mathbf{x} = \mathbf{0}$, a contradiction. Hence, the required result follows.

Theorem 6.1.16. *Let $\lambda_1, \ldots, \lambda_n$, not necessarily distinct, be the $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{C})$. Then, $\det(A) = \prod\limits_{i=1}^{n} \lambda_i$ and $\mathsf{tr}(\mathsf{A}) = \sum\limits_{\mathsf{i}=1}^{\mathsf{n}} \mathsf{a}_{\mathsf{ii}} = \sum\limits_{\mathsf{i}=1}^{\mathsf{n}} \lambda_{\mathsf{i}}$.*

Proof. Since $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, by definition,

$$\det(A - xI_n) = (-1)^n \prod_{i=1}^{n} (x - \lambda_i) \tag{6.1.2}$$

is an identity in x as polynomials. Therefore, by substituting $x = 0$ in Equation (6.1.2), we get $\det(A) = (-1)^n(-1)^n \prod_{i=1}^n \lambda_i = \prod_{i=1}^n \lambda_i$. Also,

$$\det(A - xI_n) = \begin{bmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - x \end{bmatrix} \tag{6.1.3}$$

$$= a_0 - xa_1 + \cdots + (-1)^{n-1}x^{n-1}a_{n-1} + (-1)^n x^n \tag{6.1.4}$$

for some $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$. Then, a_{n-1}, the coefficient of $(-1)^{n-1}x^{n-1}$, comes from the term

$$(a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x).$$

So, $a_{n-1} = \sum\limits_{i=1}^{n} a_{ii} = \mathsf{tr}(\mathsf{A})$, the trace of A. Also, from Equation (6.1.2) and (6.1.4), we have

$$a_0 - xa_1 + \cdots + (-1)^{n-1}x^{n-1}a_{n-1} + (-1)^n x^n = (-1)^n \prod_{i=1}^{n}(x - \lambda_i).$$

Therefore, comparing the coefficient of $(-1)^{n-1}x^{n-1}$, we have

$$\mathsf{tr}(\mathsf{A}) = \mathsf{a}_{\mathsf{n}-1} = (-1)\left\{(-1)\sum_{\mathsf{i}=1}^{\mathsf{n}} \lambda_{\mathsf{i}}\right\} = \sum_{\mathsf{i}=1}^{\mathsf{n}} \lambda_{\mathsf{i}}.$$

Hence, we get the required result. ■

Exercise **6.1.17.** 1. *Let A be a 3×3 orthogonal matrix ($AA^T = I$). If $\det(A) = 1$, then prove that there exists $\mathbf{v} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ such that $A\mathbf{v} = \mathbf{v}$.*

2. *Let $A \in \mathbb{M}_{2n+1}(\mathbb{R})$ with $A^T = -A$. Then, prove that 0 is an eigenvalue of A.*

3. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A is invertible if and only if 0 is not an eigenvalue of A.*

4. *Let $A \in \mathbb{M}_n(\mathbb{C})$ satisfy $\|A\mathbf{x}\| \le \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{C}^n$. Then, prove that if $\alpha \in \mathbb{C}$ with $|\alpha| > 1$ then $A - \alpha I$ is invertible.*

6.1.1 Spectrum of a Matrix

Definition 6.1.18. **[Algebraic, Geometric Multiplicity]** Let $A \in \mathbb{M}_n(\mathbb{C})$. Then,

1. the multiplicity of $\alpha \in \sigma(A)$ is called the **algebraic multiplicity** of A, denoted $\text{Alg.Mul}_\alpha(A)$.
2. for $\alpha \in \sigma(A)$, $\dim(\text{Null}(A - \alpha I))$ is called the **geometric multiplicity** of A, $\text{Geo.Mul}_\alpha(A)$.

We now state the following observations.

Remark 6.1.19. *Let $A \in \mathbb{M}_n(\mathbb{C})$.*

1. *Then, for each $\alpha \in \sigma(A)$, using Theorem 2.2.40 $\dim(\text{Null}(A - \alpha I)) \ge 1$. So, we have at least one eigenvector.*
2. *If the algebraic multiplicity of $\alpha \in \sigma(A)$ is $r \ge 2$ then the Example 6.1.7.7 implies that we need not have r linearly independent eigenvectors.*

Theorem 6.1.20. *Let A and B be two similar matrices. Then,*

1. *$\alpha \in \sigma(A)$ if and only if $\alpha \in \sigma(B)$.*
2. *for each $\alpha \in \sigma(A)$, $\text{ALG.MUL}_\alpha(A) = \text{ALG.MUL}_\alpha(B)$ and $\text{GEO.MUL}_\alpha(A) = \text{GEO.MUL}_\alpha(B)$.*

Proof. Since A and B are similar, there exists an invertible matrix S such that $A = SBS^{-1}$. So, $\alpha \in \sigma(A)$ if and only if $\alpha \in \sigma(B)$ as

$$\begin{aligned}\det(A - xI) &= \det(SBS^{-1} - xI) = \det\left(S(B - xI)S^{-1}\right) \\ &= \det(S)\det(B - xI)\det(A^{-1}) = \det(B - xI).\end{aligned} \tag{6.1.5}$$

Note that Equation (6.1.5) also implies that $\text{ALG.MUL}_\alpha(A) = \text{ALG.MUL}_\alpha(B)$. We will now show that $\text{GEO.MUL}_\alpha(A) = \text{GEO.MUL}_\alpha(B)$.

So, let $Q_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis of $\text{NULL}(A - \alpha I)$. Then, $B = SAS^{-1}$ implies that $Q_2 = \{S\mathbf{v}_1, \ldots, S\mathbf{v}_k\} \subseteq \text{NULL}(B - \alpha I)$. Since Q_1 is linearly independent and S is invertible, we get Q_2 is linearly independent. So, $\text{GEO.MUL}_\alpha(A) \leq \text{GEO.MUL}_\alpha(B)$. Now, we can start with eigenvectors of B and use similar arguments to get $\text{GEO.MUL}_\alpha(B) \leq \text{GEO.MUL}_\alpha(A)$ and hence the required result follows. ∎

Remark 6.1.21.

1. *Let $A = S^{-1}BS$. Then, from the proof of Theorem 6.1.20, we see that $\mathbf{x}$ is an eigenvector of A for λ if and only if $S\mathbf{x}$ is an eigenvector of B for λ.*
2. *Let A and B be two similar matrices then $\sigma(A) = \sigma(B)$. But, the converse is not true. For example, take $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.*
3. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, for any invertible matrix B, the matrices AB and $BA = B(AB)B^{-1}$ are similar. Hence, in this case the matrices AB and BA have*
   (a) *the same set of eigenvalues.*
   (b) *$\text{ALG.MUL}_\alpha(AB) = \text{ALG.MUL}_\alpha(BA)$, for each $\alpha \in \sigma(A)$.*
   (c) *$\text{GEO.MUL}_\alpha(AB) = \text{GEO.MUL}_\alpha(BA)$, for each $\alpha \in \sigma(A)$.*

We will now give a relation between the geometric multiplicity and the algebraic multiplicity.

Theorem 6.1.22. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, for $\alpha \in \sigma(A)$, $\text{GEO.MUL}_\alpha(A) \leq \text{ALG.MUL}_\alpha(A)$.*

Proof. Let $\text{GEO.MUL}_\alpha(A) = k$. Suppose $Q_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthonormal basis of $\text{NULL}(A - \alpha I)$. Extend Q_1 to get $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ as an orthonormal basis of $\mathbb{C}^n$. Put $P = [\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n]$. Then, $P^* = P^{-1}$ and

$$\begin{aligned} P^*AP &= P^*\left[A\mathbf{v}_1, \ldots, A\mathbf{v}_k, A\mathbf{v}_{k+1}, \ldots, A\mathbf{v}_n\right] \\ &= \begin{bmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_k^* \\ \mathbf{v}_{k+1}^* \\ \vdots \\ \mathbf{v}_n^* \end{bmatrix} [\alpha\mathbf{v}_1, \ldots, \alpha\mathbf{v}_k, *, \ldots, *] = \left[\begin{array}{ccc|ccc} \alpha & \cdots & 0 & * & \cdots & * \\ 0 & \ddots & 0 & * & \cdots & * \\ 0 & \cdots & \alpha & * & \cdots & * \\ \hline 0 & \cdots & 0 & * & \cdots & * \\ \vdots & & & & & \\ 0 & \cdots & 0 & * & \cdots & * \end{array}\right]. \end{aligned}$$

Now, if we denote the lower diagonal submatrix as D then

$$P_A(x) = \det(A - xI) = \det(P^*AP - xI) = (\alpha - x)^k \det(D - xI). \tag{6.1.6}$$

So, $\text{ALG.MUL}_\alpha(A) = \text{ALG.MUL}_\alpha(P^*AP) \geq k = \text{GEO.MUL}_\alpha(A)$. ∎

Remark 6.1.23. *Note that in the proof of Theorem 6.1.22, the remaining eigenvalues of A are the eigenvalues of D (see Equation (6.1.6)). This technique is called* **deflation**.

EXERCISE **6.1.24.**

1. *Let* $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix}$. *Notice that* $\mathbf{x}_1 = \frac{1}{\sqrt{3}}\mathbf{1}$ *is an eigenvector for A. Find an ordered basis* $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ *of* $\mathbb{C}^3$. *Put* $X = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}$. *Compute* $X^{-1}AX$ *to get a block-triangular matrix. Can you now find the remaining eigenvalues of A?*

2. *Let* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ *and* $B \in \mathbb{M}_{n \times m}(\mathbb{R})$.
   (a) *If* $\alpha \in \sigma(AB)$ *and* $\alpha \neq 0$ *then*
      i. $\alpha \in \sigma(BA)$.
      ii. $\text{ALG.MUL}_\alpha(AB) = \text{ALG.MUL}_\alpha(BA)$.
      iii. $\text{GEO.MUL}_\alpha(AB) = \text{GEO.MUL}_\alpha(BA)$.
   (b) *If* $0 \in \sigma(AB)$ *and* $n = m$ *then* $\text{ALG.MUL}_0(AB) = \text{ALG.MUL}_0(BA)$ *as there are n eigenvalues, counted with multiplicity.*
   (c) *Give an example to show that* $\text{GEO.MUL}_0(AB)$ *need not equal* $\text{GEO.MUL}_0(BA)$ *even when* $n = m$.

3. *Let* $A \in \mathbb{M}_n(\mathbb{R})$ *be an invertible matrix and let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *with* $\mathbf{x} \neq \mathbf{0}$ *and* $\mathbf{y}^T A^{-1}\mathbf{x} \neq 0$. *Define* $B = \mathbf{x}\mathbf{y}^T A^{-1}$. *Then, prove that*
   (a) $\lambda_0 = \mathbf{y}^T A^{-1}\mathbf{x}$ *is an eigenvalue of B of multiplicity* 1.
   (b) 0 *is an eigenvalue of B of multiplicity* $n - 1$ *[Hint: Use Exercise 6.1.24.2a].*
   (c) $1 + \alpha\lambda_0$ *is an eigenvalue of* $I + \alpha B$ *of multiplicity* 1, *for any* $\alpha \in \mathbb{R}$.
   (d) 1 *is an eigenvalue of* $I + \alpha B$ *of multiplicity* $n - 1$, *for any* $\alpha \in \mathbb{R}$.
   (e) $\det(A + \alpha\mathbf{x}\mathbf{y}^T)$ *equals* $(1 + \alpha\lambda_0)\det(A)$, *for any* $\alpha \in \mathbb{R}$. *This result is known as the Shermon-Morrison formula for determinant.*

4. *Let* $A, B \in \mathbb{M}_2(\mathbb{R})$ *such that* $\det(A) = \det(B)$ *and* $\mathsf{tr(A)} = \mathsf{tr(B)}$.
   (a) *Do A and B have the same set of eigenvalues?*
   (b) *Give examples to show that the matrices A and B need not be similar.*

5. *Let* $A, B \in \mathbb{M}_n(\mathbb{R})$. *Also, let* $(\lambda_1, \mathbf{u})$ *and* $(\lambda_2, \mathbf{v})$ *are eigen-pairs of A and B, respectively.*
   (a) *If* $\mathbf{u} = \alpha\mathbf{v}$ *for some* $\alpha \in \mathbb{R}$ *then* $(\lambda_1 + \lambda_2, \mathbf{u})$ *is an eigen-pair for* $A + B$.
   (b) *Give an example to show that if* $\mathbf{u}$ *and* $\mathbf{v}$ *are linearly independent then* $\lambda_1 + \lambda_2$ *need not be an eigenvalue of* $A + B$.

6. *Let $A \in \mathbb{M}_n(\mathbb{R})$ be an invertible matrix with eigen-pairs $(\lambda_1, \mathbf{u}_1), \ldots, (\lambda_n, \mathbf{u}_n)$. Then, prove that $\mathcal{B} = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ forms a basis of $\mathbb{R}^n$. If $[\mathbf{b}]_{\mathcal{B}} = (c_1, \ldots, c_n)^T$ then the system $A\mathbf{x} = \mathbf{b}$ has the unique solution*
$$\mathbf{x} = \frac{c_1}{\lambda_1}\mathbf{u}_1 + \frac{c_2}{\lambda_2}\mathbf{u}_2 + \cdots + \frac{c_n}{\lambda_n}\mathbf{u}_n.$$

6.2 Diagonalization

Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $T \in \mathcal{L}(\mathbb{C}^n)$ be defined by $T(\mathbf{x}) = A\mathbf{x}$, for all $\mathbf{x} \in \mathbb{C}^n$. In this section, we first find conditions under which one can obtain a basis $\mathcal{B}$ of $\mathbb{C}^n$ such that $T[\mathcal{B}, \mathcal{B}]$ (see Theorem 4.4.4) is a diagonal matrix. And, then it is shown that normal matrices satisfy the above conditions. To start with, we have the following definition.

Definition 6.2.1. [**Matrix Diagonalizability**] A matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix. Or equivalently, $P^{-1}AP = D \Leftrightarrow AP = PD$, for some diagonal matrix D and invertible matrix P.

Example 6.2.2. 1. Let A be an $n \times n$ diagonalizable matrix. Then, by definition, A is similar to a diagonal matrix, say $D = \text{diag}(d_1, \ldots, d_n)$. Thus, by Remark 6.1.21, $\sigma(A) = \sigma(D) = \{d_1, \ldots, d_n\}$.

2. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then, A cannot be diagonalized.

   Solution: Suppose A is diagonalizable. Then, A is similar to $D = \text{diag}(d_1, d_2)$. Thus, by Theorem 6.1.20, $\{d_1, d_2\} = \sigma(D) = \sigma(A) = \{0, 0\}$. Hence, $D = \mathbf{0}$ and therefore, $A = SDS^{-1} = \mathbf{0}$, a contradiction.

3. Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. Then, A cannot be diagonalized.

   Solution: Suppose A is diagonalizable. Then, A is similar to $D = \text{diag}(d_1, d_2, d_3)$. Thus, by Theorem 6.1.20, $\{d_1, d_2, d_3\} = \sigma(D) = \sigma(A) = \{2, 2, 2\}$. Hence, $D = 2I_3$ and therefore, $A = SDS^{-1} = 2I_3$, a contradiction.

4. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then, $\left(i, \begin{bmatrix} i \\ 1 \end{bmatrix}\right)$ and $\left(-i, \begin{bmatrix} -i \\ 1 \end{bmatrix}\right)$ are two eigen-pairs of A. Define $U = \frac{1}{\sqrt{2}}\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$. Then, $U^*U = I_2 = UU^*$ and $U^*AU = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}$.

Theorem 6.2.3. *Let $A \in \mathbb{M}_n(\mathbb{R})$.*

1. *Let S be an invertible matrix such that $S^{-1}AS = \text{diag}(d_1, \ldots, d_n)$. Then, for $1 \le i \le n$, the i-th column of S is an eigenvector of A corresponding to d_i.*

2. *Then, A is diagonalizable if and only if A has n linearly independent eigenvectors.*

Proof. Let $S = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$. Then, $AS = SD$ gives
$$[A\mathbf{u}_1, \ldots, A\mathbf{u}_n] = A[\mathbf{u}_1, \ldots, \mathbf{u}_n] = AS = SD = S\ \text{diag}(d_1, \ldots, d_n) = [d_1\mathbf{u}_1, \ldots, d_n\mathbf{u}_n].$$
Or equivalently, $A\mathbf{u}_i = d_i\mathbf{u}_i$, for $1 \le i \le n$. As S is invertible, $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ are linearly independent. Hence, $(d_i, \mathbf{u}_i)$, for $1 \le i \le n$, are eigen-pairs of A. This proves Part 1 and "only if" part of Part 2.

Conversely, let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be n linearly independent eigenvectors of A corresponding to eigenvalues $\alpha_1, \ldots, \alpha_n$. Then, by Corollary 3.3.10, $S = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ is non-singular and

$$AS \;=\; [A\mathbf{u}_1, \ldots, A\mathbf{u}_n] = [\alpha_1\mathbf{u}_1, \ldots, \lambda_n\mathbf{u}_n] = [\mathbf{u}_1, \ldots, \mathbf{u}_n]\begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \alpha_n \end{bmatrix} = SD,$$

where $D = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$. Therefore, $S^{-1}AS = D$ and hence A is diagonalizable. ∎

Definition 6.2.4. 1. A matrix $A \in \mathbb{M}_n(\mathbb{C})$ is called **defective** if for some $\alpha \in \sigma(A)$, $\text{GEO.MUL}_\alpha(A) < \text{ALG.MUL}_\alpha(A)$.

2. A matrix $A \in \mathbb{M}_n(\mathbb{C})$ is called **non-derogatory** if $\text{GEO.MUL}_\alpha(A) = 1$, for each $\alpha \in \sigma(A)$.

As a direct consequence of Theorem 6.2.3, we obtain the following result.

Corollary 6.2.5. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then,*

1. *A is non-defective if and only if A is diagonalizable.*

2. *A has distinct eigenvalues if and only if A is non-derogatory and non-defective.*

Theorem 6.2.6. *Let $(\alpha_1, \mathbf{v}_1), \ldots, (\alpha_k, \mathbf{v}_k)$ be k eigen-pairs of $A \in \mathbb{M}_n(\mathbb{C})$ with α_i's distinct. Then, $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent.*

Proof. Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly dependent. Then, there exists a smallest $\ell \in \{1, \ldots, k-1\}$ and $\beta \neq \mathbf{0}$ such that $\mathbf{v}_{\ell+1} = \beta_1\mathbf{v}_1 + \cdots + \beta_\ell\mathbf{v}_\ell$. So,

$$\alpha_{\ell+1}\mathbf{v}_{\ell+1} = \alpha_{\ell+1}\beta_1\mathbf{v}_1 + \cdots + \alpha_{\ell+1}\beta_\ell\mathbf{v}_\ell. \tag{6.2.1}$$

and

$$\alpha_{\ell+1}\mathbf{v}_{\ell+1} \;=\; A\mathbf{v}_{\ell+1} = A\left(\beta_1\mathbf{v}_1 + \cdots + \beta_\ell\mathbf{v}_\ell\right) = \alpha_1\beta_1\mathbf{v}_1 + \cdots + \alpha_\ell\beta_\ell\mathbf{v}_\ell. \tag{6.2.2}$$

Now, subtracting Equation (6.2.2) from Equation (6.2.1), we get

$$\mathbf{0} = (\alpha_{\ell+1} - \alpha_1)\,\beta_1\mathbf{v}_1 + \cdots + (\alpha_{\ell+1} - \alpha_\ell)\,\beta_\ell\mathbf{v}_\ell.$$

So, $\mathbf{v}_\ell \in LS(\mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1})$, a contradiction to the choice of ℓ. Thus, the required result follows. ∎

An immediate corollary of Theorem 6.2.3 and Theorem 6.2.6 is stated next without proof.

Corollary 6.2.7. *Let $A \in \mathbb{M}_n(\mathbb{C})$ have n distinct eigenvalues. Then, A is diagonalizable.*

The converse of Theorem 6.2.6 is not true as I_n has n linearly independent eigenvectors corresponding to the eigenvalue 1, repeated n times.

Corollary 6.2.8. *Let $\alpha_1, \ldots, \alpha_k$ be k distinct eigenvalues $A \in \mathbb{M}_n(\mathbb{C})$. Also, for $1 \leq i \leq k$, let $\dim(\text{NULL}(A - \alpha_i I_n)) = n_i$. Then, A has $\sum\limits_{i=1}^{k} n_i$ linearly independent eigenvectors.*

Proof. For $1 \leq i \leq k$, let $S_i = \{\mathbf{u}_{i1}, \ldots, \mathbf{u}_{in_i}\}$ be a basis of $\text{Null}(A - \alpha_i I_n)$. Then, we need to prove that $\bigcup\limits_{i=1}^{k} S_i$ is linearly independent. To do so, denote $p_j(A) = \left(\prod\limits_{i=1}^{k} (A - \alpha_i I_n)\right) / (A - \alpha_j I_n)$, for $1 \leq j \leq k$. Then, note that $p_j(A)$ is a polynomial in A of degree $k-1$ and

$$p_j(A)\mathbf{u} = \begin{cases} \mathbf{0}, & \text{if } \mathbf{u} \in \text{Null}(A - \alpha_i I_n), \text{ for some } i \neq j \\ \prod\limits_{i \neq j} (\alpha_j - \alpha_i)\mathbf{u} & \text{if } \mathbf{u} \in \text{Null}(A - \alpha_j I_n) \end{cases} \tag{6.2.3}$$

So, to prove that $\bigcup\limits_{i=1}^{k} S_i$ is linearly independent, consider the linear system

$$c_{11}\mathbf{u}_{11} + \cdots + c_{1n_1}\mathbf{u}_{1n_1} + \cdots + c_{k1}\mathbf{u}_{k1} + \cdots + c_{kn_k}\mathbf{u}_{kn_k} = \mathbf{0}$$

in the variables c_{ij}'s. Now, applying the matrix $p_j(A)$ and using Equation (6.2.3), we get

$$\prod_{i \neq j} (\alpha_j - \alpha_i) \left(c_{j1}\mathbf{u}_{j1} + \cdots + c_{jn_j}\mathbf{u}_{jn_j}\right) = \mathbf{0}.$$

But $\prod\limits_{i \neq j} (\alpha_j - \alpha_i) \neq 0$ as α_i's are distinct. Hence, $c_{j1}\mathbf{u}_{j1} + \cdots + c_{jn_j}\mathbf{u}_{jn_j} = \mathbf{0}$. As S_j is a basis of $\text{Null}(A - \alpha_j I_n)$, we get $c_{jt} = 0$, for $1 \leq t \leq n_j$. Thus, the required result follows. ∎

Corollary 6.2.9. *Let $A \in \mathbb{M}_n(\mathbb{C})$ with distinct eigenvalues $\alpha_1, \ldots, \alpha_k$. Then, A is diagonalizable if and only if $\text{Geo.Mul}_{\alpha_i}(A) = \text{Alg.Mul}_{\alpha_i}(A)$, for each $1 \leq i \leq k$.*

Proof. Let $\text{Alg.Mul}_{\alpha_i}(A) = m_i$. Then, $\sum\limits_{i=1}^{k} m_i = n$. Let $\text{Geo.Mul}_{\alpha_i}(A) = n_i$, for $1 \leq i \leq k$. Then, by Corollary 6.2.8 A has $\sum\limits_{i=1}^{k} n_i$ linearly independent eigenvectors. Also, by Theorem 6.1.22, $n_i \leq m_i$, for $1 \leq i \leq m_i$.

Now, let A be diagonalizable. Then, by Theorem 6.2.3, A has n linearly independent eigenvectors. So, $n = \sum\limits_{i=1}^{k} n_i$. As $n_i \leq m_i$ and $\sum\limits_{i=1}^{k} m_i = n$, we get $n_i = m_i$.

Now, assume that $\text{Geo.Mul}_{\alpha_i}(A) = \text{Alg.Mul}_{\alpha_i}(A)$, for $1 \leq i \leq k$. Then, for each $i, 1 \leq i \leq n$, A has $n_i = m_i$ linearly independent eigenvectors. Thus, A has $\sum\limits_{i=1}^{k} n_i = \sum\limits_{i=1}^{k} m_i = n$ linearly independent eigenvectors. Hence by Theorem 6.2.3, A is diagonalizable. ∎

Example 6.2.10. Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & -1 & 1 \end{bmatrix}$. Then, $\left(1, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right)$ and $\left(2, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right)$ are the only eigen-pairs. Hence, by Theorem 6.2.3, A is not diagonalizable.

Exercise **6.2.11.**

1. *Let A be diagonalizable. Then, prove that $A + \alpha I$ is diagonalizable for every $\alpha \in \mathbb{C}$.*
2. *Let A be an strictly upper triangular matrix. Then, prove that A is not diagonalizable.*
3. *Let A be an $n \times n$ matrix with $\lambda \in \sigma(A)$ with $\text{alg.mul}_\lambda(A) = m$. If $\text{Rank}[A - \lambda I] \neq n - m$ then prove that A is not diagonalizable.*
4. *If $\sigma(A) = \sigma(B)$ and both A and B are diagonalizable then prove that A is similar to B. That is, they are two basis representation of the same linear transformation.*

5. *Let A and B be two similar matrices such that A is diagonalizable. Prove that B is diagonalizable.*

6. *Let $A \in \mathbb{M}_n(\mathbb{R})$ and $B \in \mathbb{M}_m(\mathbb{R})$. Suppose $C = \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & B \end{bmatrix}$. Then, prove that C is diagonalizable if and only if both A and B are diagonalizable.*

7. *Is the matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ diagonalizable?*

8. *Let J_n be an $n \times n$ matrix with all entries 1. Then, $\text{Geo.Mul}_1(J_n) = \text{Alg.Mul}_1(J_n) = 1$ and $\text{Geo.Mul}_0(J_n) = \text{Alg.Mul}_0(J_n) = n - 1$.*

9. *Let $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{R})$, where $a_{ij} = a$, if $i = j$ and b, otherwise. Then, verify that $A = (a - b)I_n + bJ_n$. Hence, or otherwise determine the eigenvalues and eigenvectors of J_n. Is A diagonalizable?*

10. *Let $T : \mathbb{R}^5 \longrightarrow \mathbb{R}^5$ be a linear operator with $\text{Rank}(T - I) = 3$ and*

$$\text{Null}(T) = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \mid x_1 + x_4 + x_5 = 0,\ x_2 + x_3 = 0\}.$$

   (a) Determine the eigenvalues of T?

   (b) For each distinct eigenvalue α of T, determine $\text{Geo.Mul}_\alpha(T)$.

   (c) Is T diagonalizable? Justify your answer.

11. *Let $A \in \mathbb{M}_n(\mathbb{R})$ with $A \neq \mathbf{0}$ but $A^2 = \mathbf{0}$. Prove that A cannot be diagonalized.*

12. *Are the following matrices diagonalizable?*
   i) $\begin{bmatrix} 1 & 3 & 2 & 1 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}$, *ii)* $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$, *iii)* $\begin{bmatrix} 1 & -3 & 3 \\ 0 & -5 & 6 \\ 0 & -3 & 4 \end{bmatrix}$ *and iv)* $\begin{bmatrix} 2 & i \\ i & 0 \end{bmatrix}$.

13. *Let $A \in \mathbb{M}_n(\mathbb{C})$.*

   (a) Then, prove that $\mathsf{Rank}(A) = 1$ if and only if $A = \mathbf{x}\mathbf{y}^*$, for some non-zero vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$.

   (b) If $\mathsf{Rank}(A) = 1$ then

      i. A has at most one nonzero eigenvalue of algebraic multiplicity 1.

      ii. find this eigenvalue and its geometric multiplicity.

      iii. when is A diagonalizable?

14. *Let $A \in \mathbb{M}_n(\mathbb{C})$. If $\mathsf{Rank}(A) = k$ then there exists $\mathbf{x}_i, \mathbf{y}_i \in \mathbb{C}^n$ such that $A = \sum\limits_{i=1}^{k} \mathbf{x}_i\mathbf{y}_i^*$. Is the converse true?*

6.2.1 Schur's Unitary Triangularization

We now prove one of the most important results in diagonalization, called the Schur's Lemma or the Schur's unitary triangularization.

Lemma 6.2.12 (Schur's unitary triangularization (SUT))**.** *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, there exists a unitary matrix U such that A is similar to an upper triangular matrix. Further, if $A \in \mathbb{M}_n(\mathbb{R})$ and $\sigma(A)$ have real entries then U is a real orthogonal matrix.*

Proof. We prove the result by induction on n. The result is clearly true for $n = 1$. So, let $n > 1$ and assume the result to be true for $k < n$ and prove it for n.

Let $(\lambda_1, \mathbf{x}_1)$ be an eigen-pair of A with $\|\mathbf{x}_1\| = 1$. Now, extend it to form an orthonormal basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{u}_n\}$ of $\mathbb{C}^n$ and define $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{u}_n]$. Then, X is a unitary matrix and

$$X^*AX = X^*[A\mathbf{x}_1, A\mathbf{x}_2, \ldots, A\mathbf{x}_n] = \begin{bmatrix} \mathbf{x}_1^* \\ \mathbf{x}_2^* \\ \vdots \\ \mathbf{x}_n^* \end{bmatrix} [\lambda_1\mathbf{x}_1, A\mathbf{x}_2, \ldots, A\mathbf{x}_n] = \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{bmatrix}, \tag{6.2.4}$$

where $B \in \mathbb{M}_{n-1}(\mathbb{C})$. Now, by induction hypothesis there exists a unitary matrix $U \in \mathbb{M}_{n-1}(\mathbb{C})$ such that $U^*BU = T$ is an upper triangular matrix. Define $\widehat{U} = X \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & U \end{bmatrix}$. Then, using Exercise 5.4.8.10, the matrix $\widehat{U}$ is unitary and

$$\begin{aligned} \left(\widehat{U}\right)^* A\widehat{U} &= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & U^* \end{bmatrix} X^*AX \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & U \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & U^* \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & U \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & U^*B \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & U \end{bmatrix} = \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & U^*BU \end{bmatrix} = \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & T \end{bmatrix}. \end{aligned}$$

Since T is upper triangular, $\begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & T \end{bmatrix}$ is upper triangular.

Further, if $A \in \mathbb{M}_n(\mathbb{R})$ and $\sigma(A)$ has real entries then $\mathbf{x}_1 \in \mathbb{R}^n$ with $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$. Now, one uses induction once again to get the required result. ∎

Remark 6.2.13. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, by Schur's Lemma there exists a unitary matrix U such that $U^*AU = T = [t_{ij}]$, a triangular matrix. Thus,*

$$\{\alpha_1, \ldots, \alpha_n\} = \sigma(A) = \sigma(U^*AU) = \{t_{11}, \ldots, t_{nn}\}. \tag{6.2.5}$$

Furthermore, we can get the α_i's in the diagonal of T in any prescribed order.

Definition 6.2.14. **[Unitary Equivalence]** Let $A, B \in \mathbb{M}_n(\mathbb{C})$. Then, A and B are said to be **unitarily equivalent/similar** if there exists a unitary matrix U such that $A = U^*BU$.

Remark 6.2.15. *We know that if two matrices are unitarily equivalent then they are necessarily similar as $U^* = U^{-1}$, for every unitary matrix U. But, similarity doesn't imply unitary equivalence (see Exercise 6.2.17.6). In numerical calculations, unitary transformations are preferred as compared to similarity transformations due to the following main reasons:*

1. *Exercise 5.4.8.5g implies that* $\|A\mathbf{x}\| = \|\mathbf{x}\|$, *whenever* A *is a normal matrix. This need not be true under a similarity change of basis.*

2. *As* $U^{-1} = U^*$, *for a unitary matrix, unitary equivalence is computationally simpler.*

3. *Also, computation of "conjugate transpose" doesn't create round-off error in calculation.*

Example 6.2.16. Consider the two matrices $A = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. Then, we show that they are similar but not unitarily similar.

Solution: Note that $\sigma(A) = \sigma(B) = \{1, 2\}$. As the eigenvalues are distinct, by Theorem 6.2.7, the matrices A and B are diagonalizable and hence there exists invertible matrices S and T such that $A = S\Lambda S^{-1}$, $B = T\Lambda T^{-1}$, where $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. Thus, $A = ST^{-1}B(ST^{-1})^{-1}$. That is, A and B are similar. But, $\sum |a_{ij}|^2 \neq \sum |b_{ij}|^2$ and hence by Exercise 5.4.8.11, they cannot be unitarily similar.

Exercise **6.2.17.**

1. *If* A *is unitarily similar to an upper triangular matrix* $T = [t_{ij}]$ *then prove that* $\sum\limits_{i<j} |t_{ij}|^2 = \mathsf{tr}(\mathsf{A}^*\mathsf{A}) - \sum |\lambda_\mathsf{i}|^2$.

2. *Use the exercises given below to conclude that the upper triangular matrix obtained in the "Schur's Lemma" need not be unique.*

   (a) *Prove that* $B = \begin{bmatrix} 2 & -1 & 3\sqrt{2} \\ 0 & 1 & \sqrt{2} \\ 0 & 0 & 3 \end{bmatrix}$ *and* $C = \begin{bmatrix} 2 & 1 & 3\sqrt{2} \\ 0 & 1 & -\sqrt{2} \\ 0 & 0 & 3 \end{bmatrix}$ *are unitarily equivalent.*

   (b) *Prove that* $D = \begin{bmatrix} 2 & 0 & 3\sqrt{2} \\ 1 & 1 & \sqrt{2} \\ 0 & 0 & 1 \end{bmatrix}$ *and* $E = \begin{bmatrix} 2 & 0 & 3\sqrt{2} \\ -1 & 1 & -\sqrt{2} \\ 0 & 0 & 1 \end{bmatrix}$ *are unitarily equivalent.*

   (c) *Let* $A_1 = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ *and* $A_2 = \begin{bmatrix} 1 & 1 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$. *Then, prove that*

      i. A_1 *and* D *are unitarily equivalent.*

      ii. A_2 *and* B *are unitarily equivalent.*

      iii. *Do the above results contradict Exercise 5.4.8.5c? Give reasons for your answer.*

3. *Prove that* $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ *and* $B = \begin{bmatrix} 2 & -1 & \sqrt{2} \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ *are unitarily equivalent.*

4. *Let* A *be a normal matrix. If all the eigenvalues of* A *are* 0 *then prove that* $A = \mathbf{0}$. *What happens if all the eigenvalues of* A *are* 1?

5. *Let* $A \in \mathbb{M}_n(\mathbb{C})$. *Then, Prove that if* $\mathbf{x}^* A\mathbf{x} = 0$, *for all* $\mathbf{x} \in \mathbb{C}^n$, *then* $A = \mathbf{0}$. *Do these results hold for arbitrary matrices?*

6. *Show that the matrices* $A = \begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix}$ *and* $B = \begin{bmatrix} 10 & 9 \\ -4 & -2 \end{bmatrix}$ *are similar. Is it possible to find a unitary matrix* U *such that* $A = U^*BU$?

We now use Lemma 6.2.12 to give another proof of Theorem 6.1.16.

Corollary 6.2.18. *Let $A \in \mathbb{M}_n(\mathbb{C})$. If $\sigma(A) = \{\alpha_1, \ldots, \alpha_n\}$ then* $\det(A) = \prod\limits_{i=1}^{n} \alpha_i$ *and* $\mathsf{tr}(\mathsf{A}) = \sum\limits_{i=1}^{n} \alpha_i$.

Proof. By Schur's Lemma there exists a unitary matrix U such that $U^*AU = T = [t_{ij}]$, a triangular matrix. By Remark 6.2.13, $\sigma(A) = \sigma(T)$. Hence, $\det(A) = \det(T) = \prod\limits_{i=1}^{n} t_{ii} = \prod\limits_{i=1}^{n} \alpha_i$ and $\mathsf{tr}(\mathsf{A}) = \mathsf{tr}(\mathsf{A}(\mathsf{U}\mathsf{U}^*)) = \mathsf{tr}(\mathsf{U}^*(\mathsf{A}\mathsf{U})) = \mathsf{tr}(\mathsf{T}) = \sum\limits_{i=1}^{n} t_{ii} = \sum\limits_{i=1}^{n} \alpha_i$. ∎

6.2.2 Diagonalizability of some Special Matrices

We now use Schur's unitary triangularization Lemma to state the main theorem of this subsection. Also, recall that A is said to be a normal matrix if $AA^* = A^*A$.

Theorem 6.2.19 (Spectral Theorem for Normal Matrices)**.** *Let $A \in \mathbb{M}_n(\mathbb{C})$. If A is a normal matrix then there exists a unitary matrix U such that $U^*AU = \text{diag}(\alpha_1, \ldots, \alpha_n)$.*

Proof. By Schur's Lemma there exists a unitary matrix U such that $U^*AU = T = [t_{ij}]$, a triangular matrix. Since A is a normal

$$T^*T = (U^*AU)^*(U^*AU) = U^*A^*AU = U^*AA^*U = (U^*AU)(U^*AU)^* = TT^*.$$

Thus, we see that T is an upper triangular matrix with $T^*T = TT^*$. Thus, by Exercise 1.2.11.4, T is a diagonal matrix and this completes the proof. ∎

Exercise **6.2.20.** *Let $A \in \mathbb{M}_n(\mathbb{C})$. If A is either a Hermitian, skew-Hermitian or Unitary matrix then A is a normal matrix.*

We re-write Theorem 6.2.19 in another form to indicate that A can be decomposed into linear combination of orthogonal projectors onto eigen-spaces. Thus, it is independent of the choice of eigenvectors.

Remark 6.2.21. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a normal matrix with eigenvalues $\alpha_1, \ldots, \alpha_n$.*

1. *Then, there exists a unitary matrix $U = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ such that*
   - (a) $I_n = \mathbf{u}_1\mathbf{u}_1^* + \cdots + \mathbf{u}_n\mathbf{u}_n^*$.
   - (b) *the columns of U form a set of orthonormal eigenvectors for A (use Theorem 6.2.3).*
   - (c) $A = A \cdot I_n = A\left(\mathbf{u}_1\mathbf{u}_1^* + \cdots + \mathbf{u}_n\mathbf{u}_n^*\right) = \alpha_1\mathbf{u}_1\mathbf{u}_1^* + \cdots + \alpha_n\mathbf{u}_n\mathbf{u}_n^*$.
2. *Let $\alpha_1, \ldots, \alpha_k$ be the distinct eigenvalues of A. Also, let $W_i = \text{Null}(A - \alpha_i I_n)$, for $1 \leq i \leq k$, be the corresponding eigen-spaces.*
   - (a) *Then, we can group the $\mathbf{u}_i$'s such that they form an orthonormal basis of W_i, for $1 \leq i \leq k$. Hence, $\mathbb{C}^n = W_1 \oplus \cdots \oplus W_k$.*
   - (b) *If P_{α_i} is the orthogonal projector onto W_i, for $1 \leq i \leq k$ then $A = \alpha_1 P_1 + \cdots + \alpha_k P_k$. Thus, A depends only on eigen-spaces and not on the computed eigenvectors.*

We now give the spectral theorem for Hermitian matrices.

Theorem 6.2.22. **[Spectral Theorem for Hermitian Matrices]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix. Then,*

1. *the eigenvalues α_i, for $1 \leq i \leq n$, of A are real.*
2. *there exists a unitary matrix U, say $U = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ such that*
   (a) *$I_n = \mathbf{u}_1\mathbf{u}_1^* + \cdots + \mathbf{u}_n\mathbf{u}_n^*$.*
   (b) *$\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ forms a set of orthonormal eigenvectors for A.*
   (c) *$A = \alpha_1\mathbf{u}_1\mathbf{u}_1^* + \cdots + \alpha_n\mathbf{u}_n\mathbf{u}_n^*$, or equivalently, $U^*AU = D$, where $D =$ diag$(\alpha_1, \ldots, \alpha_n)$.*

Proof. The second part is immediate from Theorem 6.2.19 as Hermitian matrices are also normal matrices. For Part 1, let $(\alpha, \mathbf{x})$ be an eigen-pair. Then, $A\mathbf{x} = \alpha\mathbf{x}$. As A is Hermitian $A^* = A$. Thus, $\mathbf{x}^*A = \mathbf{x}^*A^* = (A\mathbf{x})^* = (\alpha\mathbf{x})^* = \overline{\alpha}\mathbf{x}^*$. Hence, using $\mathbf{x}^*A = \overline{\alpha}\mathbf{x}^*$, we get

$$\alpha\mathbf{x}^*\mathbf{x} = \mathbf{x}^*(\alpha\mathbf{x}) = \mathbf{x}^*(A\mathbf{x}) = (\mathbf{x}^*A)\mathbf{x} = (\overline{\alpha}\mathbf{x}^*)\mathbf{x} = \overline{\alpha}\mathbf{x}^*\mathbf{x}.$$

As $\mathbf{x}$ is an eigenvector, $\mathbf{x} \neq \mathbf{0}$. Hence, $\|\mathbf{x}\|^2 = \mathbf{x}^*\mathbf{x} \neq 0$. Thus $\alpha = \overline{\alpha}$, *i.e.*, $\alpha \in \mathbb{R}$. ∎

As an immediate corollary of Theorem 6.2.22 and the second part of Lemma 6.2.12, we give the following result without proof.

Corollary 6.2.23. *Let $A \in \mathbb{M}_n(\mathbb{R})$ be symmetric. Then, $A = U$ diag$(\alpha_1, \ldots, \alpha_n)$ U^*, where*

1. *the α_i's are all real,*
2. *the columns of U can be chosen to have real entries,*
3. *the eigenvectors that correspond to the columns of U form an orthonormal basis of $\mathbb{R}^n$.*

Exercise **6.2.24.**
1. *Let A be a skew-symmetric matrix. Then, the eigenvalues of A are either zero or purely imaginary and A is unitarily diagonalizable.*
2. *Let A be a skew-Hermitian matrix. Then, A is unitarily diagonalizable.*
3. *Characterize all normal matrices in $\mathbb{M}_2(\mathbb{R})$.*
4. *Let $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$. Then, prove that the following statements are equivalent.*
   (a) *A is normal.*
   (b) *A is unitarily diagonalizable.*
   (c) *$\sum\limits_{i,j} |a_{ij}|^2 = \sum\limits_{i} |\lambda_i|^2$.*
   (d) *A has n orthonormal eigenvectors.*
5. *Let A be a normal matrix with $(\lambda, \mathbf{x})$ as an eigen-pair. Then,*
   (a) *$(A^*)^k\mathbf{x}$ for $k \in \mathbb{Z}^+$ is also an eigenvector corresponding to λ.*
   (b) *$(\overline{\lambda}, \mathbf{x})$ is an eigen-pair for A^*. [Hint: Verify $\|A^*\mathbf{x} - \overline{\lambda}\mathbf{x}\|^2 = \|A\mathbf{x} - \lambda\mathbf{x}\|^2$.]*
6. *Let A be an $n \times n$ unitary matrix. Then,*
   (a) *$|\lambda| = 1$ for any eigenvalue λ of A.*

(b) *the eigenvectors* $\mathbf{x}, \mathbf{y}$ *corresponding to distinct eigenvalues are orthogonal.*

7. *Let* A *be a* 2×2 *orthogonal matrix. Then, prove the following:*

   (a) *if* $\det(A) = 1$ *then* $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, *for some* $\theta, 0 \le \theta < 2\pi$. *That is,* A *counterclockwise rotates every point in* $\mathbb{R}^2$ *by an angle* θ.

   (b) *if* $\det A = -1$ *then* $A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$, *for some* $\theta, 0 \le \theta < 2\pi$. *That is,* A *reflects every point in* $\mathbb{R}^2$ *about a line passing through origin. Determine this line. Or equivalently, there exists a non-singular matrix* P *such that* $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

8. *Let* A *be a* 3×3 *orthogonal matrix. Then, prove the following:*

   (a) *if* $\det(A) = 1$ *then* A *is a rotation about a fixed axis, in the sense that* A *has an eigen-pair* $(1, \mathbf{x})$ *such that the restriction of* A *to the plane* $\mathbf{x}^{\perp}$ *is a two dimensional rotation in* $\mathbf{x}^{\perp}$.

   (b) *if* $\det A = -1$ *then* A *corresponds to a reflection through a plane* P, *followed by a rotation about the line through origin that is orthogonal to* P.

9. *Let* A *be a normal matrix. Then, prove that* $\text{Rank}(A)$ *equals the number of nonzero eigenvalues of* A.

10. **[Equivalent characterizations of Hermitian matrices]** *Let* $A \in \mathbb{M}_n(\mathbb{C})$. *Then, the following statements are equivalent.*

    (a) *The matrix* A *is Hermitian.*

    (b) *The number* x^*Ax *is real for each* $x \in \mathbb{C}^n$.

    (c) *The matrix* A *is normal and has real eigenvalues.*

    (d) *The matrix* S^*AS *is Hermitian for each* $S \in \mathbb{M}_n(\mathbb{C})$.

6.2.3 Cayley Hamilton Theorem

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, in Theorem 6.1.16, we saw that

$$P_A(x) = \det(A - xI) = (-1)^n \left(x^n - a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + (-1)^{n-1}a_1 x + (-1)^n a_0\right) \tag{6.2.6}$$

for certain $a_i \in \mathbb{C}$, $0 \le i \le n-1$. Also, if α is an eigenvalue of A then $P_A(\alpha) = 0$. So, $x^n - a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + (-1)^{n-1}a_1 x + (-1)^n a_0 = 0$ is satisfied by n complex numbers. It turns out that the expression

$$A^n - a_{n-1}A^{n-1} + a_{n-2}A^{n-2} + \cdots + (-1)^{n-1}a_1 A + (-1)^n a_0 I = \mathbf{0}$$

holds true as a matrix identity. This is a celebrated theorem called the **Cayley Hamilton Theorem**. We give a proof using Schur's unitary triangularization. To do so, we look at multiplication of certain upper triangular matrices.

Lemma 6.2.25. *Let $A_1, \ldots, A_n \in \mathbb{M}_n(\mathbb{C})$ be upper triangular matrices such that the (i,i)-th entry of A_i equals 0, for $1 \le i \le n$. Then, $A_1 A_2 \cdots A_n = \mathbf{0}$.*

Proof. We use induction to prove that the first k columns of $A_1 A_2 \cdots A_k$ is $\mathbf{0}$, for $1 \le k \le n$. The result is clearly true for $k = 1$ as the first column of A_1 is $\mathbf{0}$. For clarity, we show that the first two columns of $A_1 A_2$ is $\mathbf{0}$. Let $B = A_1 A_2$. Then, by matrix multiplication

$$B[:, i] = A_1[:, 1](A_2)_{1i} + A_1[:, 2](A_2)_{2i} + \cdots + A_1[:, n](A_2)_{ni} = \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$$

as $A_1[:, 1] = \mathbf{0}$ and $(A_2)_{ji} = 0$, for $i = 1, 2$ and $j \ge 2$. So, assume that the first $n-1$ columns of $C = A_1 \cdots A_{n-1}$ is $\mathbf{0}$ and let $B = CA_n$. Then, for $1 \le i \le n$, we see that

$$B[:, i] = C[:, 1](A_n)_{1i} + C[:, 2](A_n)_{2i} + \cdots + C[:, n](A_n)_{ni} = \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$$

as $C[:, j] = \mathbf{0}$, for $1 \le j \le n-1$ and $(A_n)_{ni} = 0$, for $i = n-1, n$. Thus, by induction hypothesis the required result follows. ∎

EXERCISE **6.2.26.** *Let $A, B \in \mathbb{M}_n(\mathbb{C})$ be upper triangular matrices with the top leading principal submatrix of A of size k being $\mathbf{0}$. If $B[k+1, k+1] = 0$ then prove that the leading principal submatrix of size $k+1$ of AB is $\mathbf{0}$.*

We now prove the Cayley Hamilton Theorem using Schur's unitary triangularization.

Theorem 6.2.27 (Cayley Hamilton Theorem)**.** *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A satisfies its characteristic equation. That is, if $P_A(x) = \det(A - xI_n) = a_0 - xa_1 + \cdots + (-1)^{n-1} a_{n-1} x^{n-1} + (-1)^n x^n$ then*

$$A^n - a_{n-1} A^{n-1} + a_{n-2} A^{n-2} + \cdots + (-1)^{n-1} a_1 A + (-1)^n a_0 I = \mathbf{0}$$

holds true as a matrix identity.

Proof. Let $\sigma(A) = \{\alpha_1, \ldots, \alpha_n\}$ then $P_A(x) = \prod\limits_{i=1}^{n} (x - \alpha_i)$. And, by Schur's unitary triangularization there exists a unitary matrix U such that $U^* A U = T$, an upper triangular matrix with $t_{ii} = \alpha_i$, for $1 \le i \le n$. Now, observe that if $A_i = T - \alpha_i I$ then the A_i's satisfy the conditions of Lemma 6.2.25. Hence,

$$(T - \alpha_1 I) \cdots (T - \alpha_n I) = \mathbf{0}.$$

Therefore,

$$P_A(A) = \prod_{i=1}^{n} (A - \alpha_i I) = \prod_{i=1}^{n} (UTU^* - \alpha_i UIU^*) = U\Big[(T - \alpha_1 I) \cdots (T - \alpha_n I)\Big]U^* = U\mathbf{0}U^* = \mathbf{0}.$$

Thus, the required result follows. ∎

We now give some examples and then implications of the Cayley Hamilton Theorem.

Remark 6.2.28. *1. Let $A = \begin{bmatrix} 1 & 2 \\ 1 & -3 \end{bmatrix}$. Then, $P_A(x) = x^2 + 2x - 5$. Hence, verify that*

$$A^2 + 2A - 5I_2 = \begin{bmatrix} 3 & -4 \\ -2 & 11 \end{bmatrix} + 2\begin{bmatrix} 1 & 2 \\ 1 & -3 \end{bmatrix} - 5\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{0}.$$

Further, verify that $A^{-1} = \frac{1}{5}(A+2I_2) = \frac{1}{5}\begin{bmatrix}3 & 2\\ 1 & -1\end{bmatrix}$. *Furthermore,* $A^2 = -2A+5I$ *implies that*

$$A^3 = A(A^2) = A(-2A+5I) = -2A^2+5I = -2(-2A+5I)+5I = 4A-10I+5I = 4A-5I.$$

We can keep using the above technique to get A^m *as a linear combination of* A *and* I*, for all* $m \geq 1$.

2. *Let* $A = \begin{bmatrix}3 & 1\\ 2 & 0\end{bmatrix}$. *Then,* $P_A(t) = t(t-3)-2 = t^2-3t-2$. *So, using* $P_A(A) = \mathbf{0}$*, we have* $A^{-1} = \frac{A-3I}{2}$. *Further,* $A^2 = 3A+2I$ *implies that* $A^3 = 3A^2+2A = 3(3A+2I)+2A = 11A+6I$. *So, as above,* A^m *is a combination of* A *and* I*, for all* $m \geq 1$.

3. *Let* $A = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$. *Then,* $P_A(x) = x^2$. *So, even though* $A \neq \mathbf{0}$, $A^2 = \mathbf{0}$.

4. *For* $A = \begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$, $P_A(x) = x^3$. *Thus, by the Cayley Hamilton Theorem* $A^3 = \mathbf{0}$. *But, it turns out that* $A^2 = \mathbf{0}$.

5. *For* $A = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$, *note that* $P_A(t) = (t-1)^3$. *So* $P_A(A) = \mathbf{0}$. *But, observe that if* $q(t) = (t-1)^2$ *then* $q(A)$ *is also* $\mathbf{0}$.

6. *Let* $A \in \mathbb{M}_n(\mathbb{C})$ *with* $P_A(x) = a_0 - xa_1 + \cdots + (-1)^{n-1}a_{n-1}x^{n-1} + (-1)^n x^n$.

   (a) *Then, for any* $\ell \in \mathbb{N}$*, the division algorithm gives* $\alpha_0, \alpha_1, \ldots, \alpha_{n-1} \in \mathbb{C}$ *and a polynomial* $f(x)$ *with coefficients from* $\mathbb{C}$ *such that*

   $$x^\ell = f(x)P_A(x) + \alpha_0 + x\alpha_1 + \cdots + x^{n-1}\alpha_{n-1}.$$

   Hence, by the Cayley Hamilton Theorem, $A^\ell = \alpha_0 I + \alpha_1 A + \cdots + \alpha_{n-1}A^{n-1}$.

      i. *Thus, to compute any power of* A*, one needs to apply the division algorithm to get* α_i*'s and know* A^i*, for* $1 \leq i \leq n-1$. *This is quite helpful in numerical computation as computing powers takes much more time than division.*

      ii. *Note that* $LS\left\{I, A, A^2, \ldots\right\}$ *is a subspace of* $\mathbb{M}_n(\mathbb{C})$. *Also,* $\dim\left(\mathbb{M}_n(\mathbb{C})\right) = n^2$. *But, the above argument implies that* $\dim\left(LS\left\{I, A, A^2, \ldots\right\}\right) \leq n$.

      iii. *In the language of graph theory, it says the following: "Let* G *be a graph on* n *vertices and* A *its adjacency matrix. Suppose there is no path of length* $n-1$ *or less from a vertex* v *to a vertex* u *in* G*. Then,* G *doesn't have a path from* v *to* u *of any length. That is, the graph* G *is disconnected and* v *and* u *are in different components of* G*."*

   (b) *Suppose* A *is non-singular. Then, by definition* $a_0 = \det(A) \neq 0$. *Hence,*

   $$A^{-1} = \frac{1}{a_0}\left[a_1 I - a_2 A + \cdots + (-1)^{n-2}a_{n-1}A^{n-2} + (-1)^{n-1}A^{n-1}\right].$$

   This matrix identity can be used to calculate the inverse.

(c) *The above also implies that if A is invertible then $A^{-1} \in LS\left\{I, A, A^2, \ldots\right\}$. That is, A^{-1} is a linear combination of the vectors $I, A, \ldots, A^{n-1}$.*

EXERCISE **6.2.29.** *Find the inverse of* $\begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 1 & 1 & 2 \end{bmatrix}$, $\begin{bmatrix} -1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ *and* $\begin{bmatrix} 1 & -2 & -1 \\ -2 & 1 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ *by the Cayley Hamilton Theorem.*

EXERCISE **6.2.30.** *Miscellaneous Exercises:*

1. *Let $A, B \in \mathbb{M}_2(\mathbb{C})$ such that $A = AB - BA$. Then, prove that $A^2 = \mathbf{0}$.*
2. *Let B be an $m \times n$ matrix and $A = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix}$. Then, prove that $\left(\lambda, \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}\right)$ is an eigen-pair if and only if $\left(-\lambda, \begin{bmatrix} \mathbf{x} \\ -\mathbf{y} \end{bmatrix}\right)$ is an eigen-pair.*
3. *Let $B, C \in \mathbb{M}_n(\mathbb{R})$. Define $A = \begin{bmatrix} B & C \\ -C & B \end{bmatrix}$. Then, prove the following:*
   (a) *if s is a real eigenvalue of A with corresponding eigenvector $\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}$ then s is also an eigenvalue corresponding to the eigenvector $\begin{bmatrix} -\mathbf{y} \\ \mathbf{x} \end{bmatrix}$.*
   (b) *if $s + it$ is a complex eigenvalue of A with corresponding eigenvector $\begin{bmatrix} \mathbf{x} + i\mathbf{y} \\ -\mathbf{y} + i\mathbf{x} \end{bmatrix}$ then $s - it$ is also an eigenvalue of A with corresponding eigenvector $\begin{bmatrix} \mathbf{x} - i\mathbf{y} \\ -\mathbf{y} - i\mathbf{x} \end{bmatrix}$.*
   (c) *$(s + it, \mathbf{x} + i\mathbf{y})$ is an eigen-pair of $B+iC$ if and only if $(s - it, \mathbf{x} - i\mathbf{y})$ is an eigen-pair of $B - iC$.*
   (d) *$\left(s + it, \begin{bmatrix} \mathbf{x} + i\mathbf{y} \\ -\mathbf{y} + i\mathbf{x} \end{bmatrix}\right)$ is an eigen-pair of A if and only if $(s + it, \mathbf{x} + i\mathbf{y})$ is an eigen-pair of $B + iC$.*
   (e) $\det(A) = |\det(B + iC)|^2$.

The next section deals with quadratic forms which helps us in better understanding of conic sections in analytic geometry.

6.3 Quadratic Forms

Definition 6.3.1. [Positive, Semi-positive and Negative definite matrices] Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A is said to be

1. **positive semi-definite** (psd) if $\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ and $\mathbf{x}^*A\mathbf{x} \geq 0$, for all $\mathbf{x} \in \mathbb{C}^n$.
2. **positive definite** (pd) if $\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ and $\mathbf{x}^*A\mathbf{x} > 0$, for all $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$.
3. **negative semi-definite** (nsd) if $\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ and $\mathbf{x}^*A\mathbf{x} \leq 0$, for all $\mathbf{x} \in \mathbb{C}^n$.
4. **negative definite** (nd) if $\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ and $\mathbf{x}^*A\mathbf{x} < 0$, for all $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$.

5. **indefinite** if $\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ and there exist $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ such that $\mathbf{x}^*A\mathbf{x} < 0 < \mathbf{y}^*A\mathbf{y}$.

Lemma 6.3.2. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then A is Hermitian if and only if at least one of the following statements hold:*

1. *S^*AS is Hermitian for all $S \in \mathbb{M}_n$.*
2. *A is normal and has real eigenvalues.*
3. *$\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ for all $\mathbf{x} \in \mathbb{C}^n$.*

Proof. Let $S \in \mathbb{M}_n$, $(S^*AS)^* = S^*A^*S = S^*AS$. Thus S^*AS is Hermitian.

Suppose $A = A^*$. Then, A is clearly normal as $AA^* = A^2 = A^*A$. Further, if $(\lambda, \mathbf{x})$ is an eigenpair then $\lambda\mathbf{x}^*\mathbf{x} = \mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ implies $\lambda \in \mathbb{R}$.

For the last part, note that $\mathbf{x}^*A\mathbf{x} \in \mathbb{C}$. Thus $\overline{\mathbf{x}^*A\mathbf{x}} = (\mathbf{x}^*A\mathbf{x})^* = \mathbf{x}^*A^*\mathbf{x} = \mathbf{x}^*A\mathbf{x}$, we get $\operatorname{Im}(\mathbf{x}^*A\mathbf{x}) = 0$. Thus, $\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$.

If S^*AS is Hermitian for all $S \in \mathbb{M}_n$ then taking $S = I_n$ gives A is Hermitian.

If A is normal then $A = U^* \operatorname{diag}(\lambda_1, \ldots, \lambda_n)U$ for some unitary matrix U. Since $\lambda_i \in \mathbb{R}$, $A^* = (U^* \operatorname{diag}(\lambda_1, \ldots, \lambda_n)U)^* = U^* \operatorname{diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n})U = U^* \operatorname{diag}(\lambda_1, \ldots, \lambda_n)U = A$. So, A is Hermitian.

If $\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ for all $\mathbf{x} \in \mathbb{C}^n$ then $a_{ii} = \mathbf{e}_i^*A\mathbf{e}_i \in \mathbb{R}$. Also, $a_{ii} + a_{jj} + a_{ij} + a_{ji} = (\mathbf{e}_i + \mathbf{e}_j)^*A(\mathbf{e}_i + \mathbf{e}_j) \in \mathbb{R}$. So, $\operatorname{Im}(a_{ij}) = -\operatorname{Im}(a_{ji})$. Similarly, $a_{ii} + a_{jj} + ia_{ij} - ia_{ji} = (\mathbf{e}_i + i\mathbf{e}_j)^*A(\mathbf{e}_i + i\mathbf{e}_j) \in \mathbb{R}$ implies that $\operatorname{Re}(a_{ij}) = \operatorname{Re}(a_{ji})$. Thus, $A = A^*$. ■

Remark 6.3.3. *Let $A \in \mathbb{M}_n(\mathbb{R})$. Then the condition $\mathbf{x}^*A\mathbf{x} \in \mathbb{R}$ in Definition 6.3.9 is always true and hence doesn't put any restriction on the matrix A. So, in Definition 6.3.9, we assume that $A^T = A$, i.e., A is a symmetric matrix.*

Example 6.3.4.
1. Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ or $A = \begin{bmatrix} 3 & 1+i \\ 1-i & 4 \end{bmatrix}$. Then, A is positive definite.
2. Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ or $A = \begin{bmatrix} \sqrt{2} & 1+i \\ 1-i & \sqrt{2} \end{bmatrix}$. Then, A is positive semi-definite but not positive definite.
3. Let $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ or $A = \begin{bmatrix} -2 & 1-i \\ 1+i & -2 \end{bmatrix}$. Then, A is negative definite.
4. Let $A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ or $A = \begin{bmatrix} -2 & 1-i \\ 1+i & -1 \end{bmatrix}$. Then, A is negative semi-definite.
5. Let $A = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$ or $A = \begin{bmatrix} 1 & 1+i \\ 1-i & 1 \end{bmatrix}$. Then, A is indefinite.

Theorem 6.3.5. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, the following statements are equivalent.*

1. *A is positive semi-definite.*
2. *$A^* = A$ and each eigenvalue of A is non-negative.*
3. *$A = B^*B$ for some $B \in \mathbb{M}_n(\mathbb{C})$.*

Proof. $1 \Rightarrow 2$: Let A be positive semi-definite. Then, by Lemma 6.3.2 A is Hermitian. If $(\alpha, \mathbf{v})$ is an eigen-pair of A then $\alpha\|\mathbf{v}\|^2 = \mathbf{v}^*A\mathbf{v} \geq 0$. So, $\alpha \geq 0$.

$2 \Rightarrow 3$: Let $\sigma(A) = \{\alpha_1, \ldots, \alpha_n\}$. Then, by spectral theorem, there exists a unitary matrix U such that $U^*AU = D$ with $D = \text{diag}(\alpha_1, \ldots, \alpha_n)$. As $\alpha_i \geq 0$, for $1 \leq i \leq n$, define $D^{\frac{1}{2}} = \text{diag}(\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_n})$. Then, $A = UD^{\frac{1}{2}}[D^{\frac{1}{2}}U^*] = B^*B$.

$3 \Rightarrow 1$: Let $A = B^*B$. Then, for $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{x}^*A\mathbf{x} = \mathbf{x}^*B^*B\mathbf{x} = \|B\mathbf{x}\|^2 \geq 0$. Thus, the required result follows. ∎

A similar argument gives the next result and hence the proof is omitted.

Theorem 6.3.6. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, the following statements are equivalent.*

1. *A is positive definite.*
2. *$A^* = A$ and each eigenvalue of A is positive.*
3. *$A = B^*B$ for a non-singular matrix $B \in \mathbb{M}_n(\mathbb{C})$.*

Remark 6.3.7. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then, there exists a unitary matrix $U = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$ and a diagonal matrix $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ such that $A = UDU^*$. Now, for $1 \leq i \leq n$, define $\alpha_i = \max\{\lambda_i, 0\}$ and $\beta_i = \min\{\lambda_i, 0\}$. Then*

1. *for $D_1 = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$, the matrix $A_1 = UD_1U^*$ is positive semi-definite.*
2. *for $D_2 = \text{diag}(\beta_1, \beta_2, \ldots, \beta_n)$, the matrix $A_2 = UD_2U^*$ is positive semi-definite.*
3. *$A = A_1 - A_2$. The matrix A_1 is generally called the positive semi-definite part of A.*

Definition 6.3.8. [**Multilinear Function**] Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then,

1. for a fixed $m \in \mathbb{N}$, a function $f : \mathbb{V}^m \to \mathbb{F}$ is called an m-**multilinear** function if f is linear in each component. That is,

$$\begin{aligned} f(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, (\mathbf{v}_i + \alpha\mathbf{u}), \mathbf{v}_{i+1} \ldots, \mathbf{v}_m) &= f(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1} \ldots, \mathbf{v}_m) \\ &\quad + \alpha f(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{u}, \mathbf{v}_{i+1} \ldots, \mathbf{v}_m) \end{aligned}$$

   for $\alpha \in \mathbb{F}$, $\mathbf{u} \in \mathbb{V}$ and $\mathbf{v}_i \in \mathbb{V}$, for $1 \leq i \leq m$.
2. An m-multilinear form is also called an m-**form**.
3. A 2-form is called a **bilinear form**.

Definition 6.3.9. [**Sesquilinear, Hermitian and Quadratic Forms**] Let $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix and let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. Then, a **sesquilinear form** in $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ is defined as $H(\mathbf{x}, \mathbf{y}) = \mathbf{y}^*A\mathbf{x}$. In particular, $H(\mathbf{x}, \mathbf{x})$, denoted $H(\mathbf{x})$, is called a **Hermitian form**. In case $A \in \mathbb{M}_n(\mathbb{R})$, $H(\mathbf{x})$ is called a **quadratic** form.

Remark 6.3.10. *Observe that*

1. *if $A = I_n$ then the bilinear/sesquilinear form reduces to the standard inner product.*
2. *$H(\mathbf{x}, \mathbf{y})$ is 'linear' in the first component and 'conjugate linear' in the second component.*
3. *the quadratic form $H(\mathbf{x})$ is a real number. Hence, for $\alpha \in \mathbb{R}$, the equation $H(\mathbf{x}) = \alpha$, represents a conic in $\mathbb{R}^n$.*

Example 6.3.11. 1. Let $\mathbf{v}_i \in \mathbb{C}^n$, for $1 \leq i \leq n$. Then, $f(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \det([\mathbf{v}_1, \ldots, \mathbf{v}_n])$ is an n-form on $\mathbb{C}^n$.

2. Let $A \in \mathbb{M}_n(\mathbb{R})$. Then, $f(\mathbf{x}, \mathbf{y}) = \mathbf{y}^T A\mathbf{x}$, for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, is a bilinear form on $\mathbb{R}^n$.
3. Let $A = \begin{bmatrix} 1 & 2-i \\ 2+i & 2 \end{bmatrix}$. Then, $A^* = A$ and for $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, verify that
$$H(\mathbf{x}) = \mathbf{x}^* A\mathbf{x} = |x|^2 + 2|y|^2 + 2\mathrm{Re}\left((2-i)\overline{x}y\right)$$
where 'Re' denotes the real part of a complex number, is a sesquilinear form.

6.3.1 Sylvester's law of inertia

The main idea of this section is to express $H(\mathbf{x})$ as sum or difference of squares. Since $H(\mathbf{x})$ is a quadratic in $\mathbf{x}$, replacing $\mathbf{x}$ by $c\mathbf{x}$, for $c \in \mathbb{C}$, just gives a multiplication factor by $|c|^2$. Hence, one needs to study only the normalized vectors. Let us consider Example 6.1.1 again. There we see that

$$\mathbf{x}^T A\mathbf{x} = 3\frac{(x+y)^2}{2} - \frac{(x-y)^2}{2} = (x+2y)^2 - 3y^2, \text{ and} \qquad (6.3.1)$$
$$\mathbf{x}^T B\mathbf{x} = 5\frac{(x+2y)^2}{5} + 10\frac{(2x-y)^2}{5} = (3x - \frac{2y}{3})^2 + \frac{50y^2}{9}. \qquad (6.3.2)$$

Note that both the expressions in Equation (6.3.1) is the difference of two non-negative terms. Whereas, both the expressions in Equation (6.3.2) consists of sum of two non-negative terms. Is this just a coincidence?

In general, let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix. Then, by Theorem 6.2.22, $\sigma(A) = \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{R}$ and there exists a unitary matrix U such that $U^*AU = D = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$. Let $\mathbf{x} = U\mathbf{z}$. Then, $\|\mathbf{x}\| = 1$ and U is unitary implies that $\|\mathbf{z}\| = 1$. If $\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_n)^*$ then

$$H(\mathbf{x}) = \mathbf{z}^*U^*AU\mathbf{z} = \mathbf{z}^*D\mathbf{z} = \sum_{i=1}^{n} \alpha_i |\mathbf{z}_i|^2 = \sum_{i=1}^{p} |\sqrt{\alpha_i}\, \mathbf{z}_i|^2 - \sum_{i=p+1}^{r} \left|\sqrt{|\alpha_i|}\, \mathbf{z}_i\right|^2, \qquad (6.3.3)$$

where $\alpha_1, \ldots, \alpha_p > 0$, $\alpha_{p+1}, \ldots, \alpha_r < 0$ and $\alpha_{r+1}, \ldots, \alpha_n = 0$. Thus, we see that the possible values of $H(\mathbf{x})$ seem to depend only on the eigenvalues of A. Since U is an invertible matrix, the components $\mathbf{z}_i$'s of $\mathbf{z} = U^{-1}\mathbf{x} = U^*\mathbf{x}$ are commonly known as the **linearly independent linear forms**. Note that each $\mathbf{z}_i$ is a linear expression in the components of $\mathbf{x}$. Also, note that in Equation (6.3.3), p corresponds to the number of positive eigenvalues and $r - p$ to the number of negative eigenvalues. For a better understanding, we define the following numbers.

Definition 6.3.12. [Inertia and Signature of a Matrix] Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix. The **inertia** of A, denoted $i(A)$, is the triplet $(i_+(A), i_-(A), i_0(A))$, where $i_+(A)$ is the number of positive eigenvalues of A, $i_-(A)$ is the number of negative eigenvalues of A and $i_0(A)$ is the nullity of A. The difference $i_+(A) - i_-(A)$ is called the **signature** of A.

EXERCISE **6.3.13.** *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix. If the signature and the rank of A is known then prove that one can find out the inertia of A.*

As a next result, we show that in any expression of $H(\mathbf{x})$ as a sum or difference of n *absolute squares of linearly independent linear forms*, the number p (respectively, $r-p$) gives the number of positive (respectively, negative) eigenvalues of A. This is popularly known as the 'Sylvester's law of inertia'.

Lemma 6.3.14. **[Sylvester's Law of Inertia]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix and let $\mathbf{x} \in \mathbb{C}^n$. Then, every Hermitian form $H(\mathbf{x}) = \mathbf{x}^*A\mathbf{x}$, in n variables can be written as*

$$H(\mathbf{x}) = |\mathbf{y}_1|^2 + \cdots + |\mathbf{y}_p|^2 - |\mathbf{y}_{p+1}|^2 - \cdots - |\mathbf{y}_r|^2$$

where $\mathbf{y}_1, \ldots, \mathbf{y}_r$ are linearly independent linear forms in the components of $\mathbf{x}$ and the integers p and r satisfying $0 \leq p \leq r \leq n$, depend only on A.

Proof. Equation (6.3.3) implies that $H(\mathbf{x})$ has the required form. We only need to show that p and r are uniquely determined by A. Hence, let us assume on the contrary that there exist $p, q, r, s \in \mathbb{N}$ with $p > q$ such that

$$\begin{aligned} H(\mathbf{x}) &= |\mathbf{y}_1|^2 + \cdots + |\mathbf{y}_p|^2 - |\mathbf{y}_{p+1}|^2 - \cdots - |\mathbf{y}_r|^2 & (6.3.4) \\ &= |\mathbf{z}_1|^2 + \cdots + |\mathbf{z}_q|^2 - |\mathbf{z}_{q+1}|^2 - \cdots - |\mathbf{z}_s|^2, & (6.3.5) \end{aligned}$$

where $\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = M\mathbf{x}$, $\mathbf{z} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = N\mathbf{x}$ with $Y_1 = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_p \end{bmatrix}$ and $Z_1 = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_q \end{bmatrix}$ for some invertible matrices M and N. Now the invertibility of M and N implies $\mathbf{z} = B\mathbf{y}$, for some invertible matrix B. Decompose $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$, where B_1 is a $q \times p$ matrix. Then $\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$. As $p > q$, the homogeneous linear system $B_1 Y_1 = \mathbf{0}$ has a nontrivial solution, say $\widetilde{Y_1} = \begin{bmatrix} \tilde{y_1} \\ \vdots \\ \tilde{y_p} \end{bmatrix}$ and consider $\widetilde{\mathbf{y}} = \begin{bmatrix} \widetilde{Y_1} \\ \mathbf{0} \end{bmatrix}$. Then for this choice of $\widetilde{\mathbf{y}}$, $Z_1 = \mathbf{0}$ and thus, using Equations (6.3.4) and (6.3.5), we have

$$H(\tilde{\mathbf{y}}) = |\tilde{y_1}|^2 + |\tilde{y_2}|^2 + \cdots + |\tilde{y_p}|^2 - 0 = 0 - (|z_{q+1}|^2 + \cdots + |z_s|^2).$$

Now, this can hold only if $\widetilde{Y_1} = \mathbf{0}$, a contradiction to $\widetilde{Y_1}$ being a non-trivial solution. Hence $p = q$. Similarly, the case $r > s$ can be resolved. This completes the proof of the lemma. □

Remark 6.3.15. *Since A is Hermitian,* $\text{Rank}(A)$ *equals the number of nonzero eigenvalues. Hence,* $\text{Rank}(A) = r$. *The number r is called the* **rank** *and the number $r - 2p$ is called the* **inertial degree** *of the Hermitian form $H(\mathbf{x})$.*

We now look at another form of the Sylvester's law of inertia. We start with the following definition.

Definition 6.3.16. **[Star Congruence]** Let $A, B \in \mathbb{M}_n(\mathbb{C})$. Then, A is said to be **$*$-congruent** (read star-congruent) to B if there exists an invertible matrix S such that $A = S^*BS$.

Theorem 6.3.17. **[Second Version: Sylvester's Law of Inertia]** *Let $A, B \in \mathbb{M}_n(\mathbb{C})$ be Hermitian. Then, A is $*$-congruent to B if and only if $i(A) = i(B)$.*

Proof. By spectral theorem $U^*AU = \Lambda_A$ and $V^*BV = \Lambda_B$, for some unitary matrices U, V and diagonal matrices Λ_A, Λ_B of the form $\operatorname{diag}(+, \cdots, +, -, \cdots, -, 0, \cdots, 0)$. Thus, there exist invertible matrices S, T such that $S^*AS = D_A$ and $T^*BT = D_B$, where D_A, D_B are diagonal matrices of the form $\operatorname{diag}(1, \cdots, 1, -1, \cdots, -1, 0, \cdots, 0)$.

If $i(A) = i(B)$, then it follows that $D_A = D_B$, *i.e.*, $S^*AS = T^*BT$ and hence $A = (TS^{-1})^*B(TS^{-1})$.

Conversely, suppose that $A = P^*BP$, for some invertible matrix P, and $i(B) = (k, l, m)$. As $T^*BT = D_B$, we have, $A = P^*(T^*)^{-1}D_BT^{-1}P = (T^{-1}P)^*D_B(T^{-1}P)$. Now, let $X = (T^{-1}P)^{-1}$. Then, $A = (X^{-1})^*D_BX^{-1}$ and we have the following observations.

1. As rank and nullity do not change under similarity transformation, $i_0(A) = i_0(D_B) = m$ as $i(B) = (k, l, m)$.
2. Using $i(B) = (k, l, m)$, we also have
$$X[:, k+1]^*AX[:, k+1] = X[:, k+1]^*\big((X^{-1})^*D_B(X^{-1})\big)X[:, k+1] = \mathbf{e}_{k+1}^*D_B\mathbf{e}_{k+1} = -1.$$
Similarly, $X[:, k+2]^*AX[:, k+2] = \cdots = X[:, k+l]^*AX[:, k+l] = -1$. As the vectors $X[:, k+1], \ldots, X[:, k+l]$ are linearly independent, using 9.7.10, we see that A has at least l negative eigenvalues.
3. Similarly, $X[:, 1]^*AX[:, 1] = \cdots = X[:, k]^*AX[:, k] = 1$. As $X[:, 1], \ldots, X[:, k]$ are linearly independent, using 9.7.10 again, we see that A has at least k positive eigenvalues.

Thus, it now follows that $i(A) = (k, l, m)$. ∎

6.3.2 Applications in Eculidean Plane and Space

We now obtain conditions on the eigenvalues of A, corresponding to the associated quadratic form, to characterize conic sections in $\mathbb{R}^2$, with respect to the standard inner product.

Definition 6.3.18. [**Associated Quadratic Form**] Let $f(x, y) = ax^2 + 2hxy + by^2 + 2fx + 2gy + c$ be a general quadratic in x and y, with coefficients from $\mathbb{R}$. Then,
$$H(\mathbf{x}) = \mathbf{x}^TA\mathbf{x} = \begin{bmatrix} x, & y \end{bmatrix} \begin{bmatrix} a & h \\ h & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2hxy + by^2$$
is called the **associated quadratic form** of the conic $f(x, y) = 0$.

Proposition 6.3.19. *Consider the general quadratic $f(x, y)$, for $a, b, c, g, f, h \in \mathbb{R}$. Then, $f(x, y) = 0$ represents*

1. *an ellipse or a circle if $ab - h^2 > 0$,*
2. *a parabola or a pair of parallel lines if $ab - h^2 = 0$,*
3. *a hyperbola or a pair of intersecting lines if $ab - h^2 < 0$.*

Proof. As A is symmetric, by Corollary 6.2.23, $A = U \operatorname{diag}(\alpha_1, \alpha_2)U^T$, where $U = [\mathbf{u}_1, \mathbf{u}_2]$ is an orthogonal matrix, with $(\alpha_1, \mathbf{u}_1)$ and $(\alpha_2, \mathbf{u}_2)$ as eigen-pairs of A. Let $[u, \ v] = \mathbf{x}^TU$. As $\mathbf{u}_1$ and

$\mathbf{u}_2$ are orthogonal, u and v represent orthogonal lines passing through origin in the (x, y)-plane. In most cases, these lines form the principal axes of the conic.

We also have $\mathbf{x}^T A\mathbf{x} = \alpha_1 u^2 + \alpha_2 v^2$ and hence $f(x, y) = 0$ reduces to

$$\alpha_1 u^2 + \alpha_2 v^2 + 2g_1 u + 2f_1 v + c = 0. \tag{6.3.6}$$

for some $g_1, f_1 \in \mathbb{R}$. Now, we consider different cases depending of the values of α_1, α_2:

1. If $\alpha_1 = 0 = \alpha_2$ then $A = \mathbf{0}$ and Equation (6.3.6) gives the straight line $2gx + 2fy + c = 0$.

2. if $\alpha_1 = 0$ and $\alpha_2 \neq 0$ then $ab - h^2 = \det(A) = \alpha_1\alpha_2 = 0$. So, after dividing by α_2, Equation (6.3.6) reduces to $(v + d_1)^2 = d_2 u + d_3$, for some $d_1, d_2, d_3 \in \mathbb{R}$. Hence, let us look at the possible subcases:
   (a) Let $d_2 = d_3 = 0$. Then, $v + d_1 = 0$ is a pair of coincident lines.
   (b) Let $d_2 = 0$, $d_3 \neq 0$.
      i. If $d_3 > 0$, then we get a pair of parallel lines given by $v = -d_1 \pm \sqrt{\frac{d_3}{\alpha_2}}$.
      ii. If $d_3 < 0$, the solution set of the corresponding conic is an empty set.
   (c) If $d_2 \neq 0$. Then, the given equation is of the form $Y^2 = 4aX$ for some translates $X = x + \alpha$ and $Y = y + \beta$ and thus represents a parabola.

      Let $H(\mathbf{x}) = x^2 + 4y^2 + 4xy$ be the associated quadratic form for a class of curves. Then, $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, $\alpha_1 = 0, \alpha_2 = 5$ and $v = x + 2y$. Now, let $d_1 = -3$ and vary d_2 and d_3 to get different curves (see Figure 6.2 drawn using the package "MATHEMATICA").

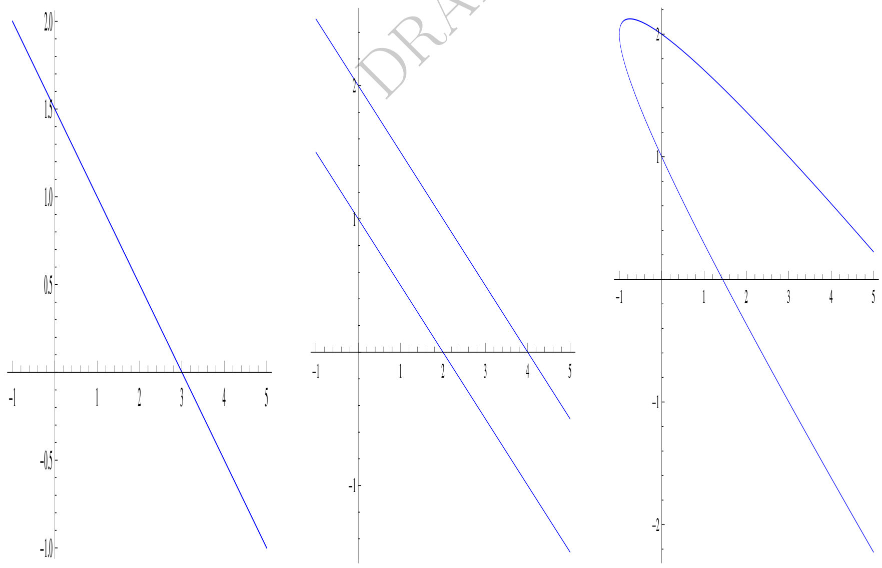


Figure 6.2: Curves for $d_2 = 0 = d_3$, $d_2 = 0, d_3 = 1$ and $d_2 = 1, d_3 = 1$

.

3. $\alpha_1 > 0$ and $\alpha_2 < 0$. Then, $ab - h^2 = \det(A) = \lambda_1\lambda_2 < 0$. If $\alpha_2 = -\beta_2$, for $\beta_2 > 0$, then Equation (6.3.6) reduces to

$$\alpha_1(u + d_1)^2 - \beta_2(v + d_2)^2 = d_3, \quad \text{for some} \quad d_1, d_2, d_3 \in \mathbb{R} \tag{6.3.7}$$

   whose understanding requires the following subcases:

(a) If $d_3 = 0$ then Equation (6.3.7) equals

$$\left(\sqrt{\alpha_1}(u+d_1) + \sqrt{\beta_2}(v+d_2)\right) \cdot \left(\sqrt{\alpha_1}(u+d_1) - \sqrt{\beta_2}(v+d_2)\right) = 0$$

or equivalently, a pair of intersecting straight lines $u + d_1 = 0$ and $v + d_2 = 0$ in the (u, v)-plane.

(b) Without loss of generality, let $d_3 > 0$. Then, Equation (6.3.7) equals

$$\frac{\lambda_1(u+d_1)^2}{d_3} - \frac{\alpha_2(v+d_2)^2}{d_3} = 1$$

or equivalently, a hyperbola with orthogonal principal axes $u+d_1 = 0$ and $v+d_2 = 0$. Let $H(\mathbf{x}) = 10x^2 - 5y^2 + 20xy$ be the associated quadratic form for a class of curves. Then, $A = \begin{bmatrix} 10 & 10 \\ 10 & -5 \end{bmatrix}$, $\alpha_1 = 15, \alpha_2 = -10$ and $\sqrt{5}u = 2x + y, \sqrt{5}v = x - 2y$. Now, let $d_1 = \sqrt{5}, d_2 = -\sqrt{5}$ to get $3(2x + y + 1)^2 - 2(x - 2y - 1)^2 = d_3$. Now vary d_3 to get different curves (see Figure 6.3 drawn using the package "MATHEMATICA").

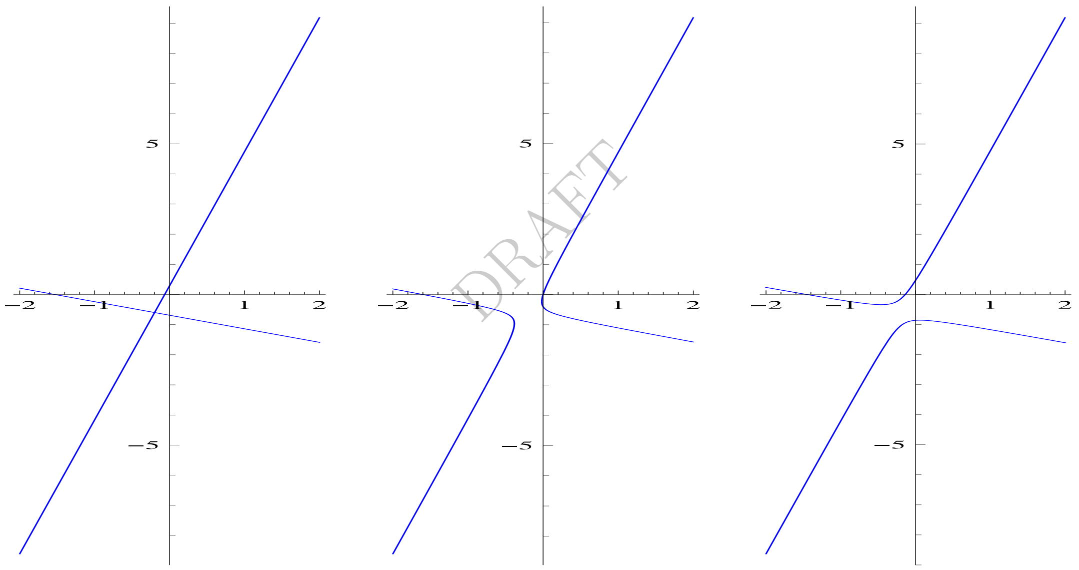


Figure 6.3: Curves for $d_3 = 0$, $d_3 = 1$ and $d_3 = -1$

.

4. $\alpha_1, \alpha_2 > 0$. Then, $ab - h^2 = \det(A) = \alpha_1\alpha_2 > 0$ and Equation (6.3.6) reduces to

$$\lambda_1(u+d_1)^2 + \lambda_2(v+d_2)^2 = d_3, \quad \text{for some} \quad d_1, d_2, d_3 \in \mathbb{R}. \tag{6.3.8}$$

We consider the following three subcases to understand this.

(a) If $d_3 = 0$ then we get a pair of orthogonal lines $u + d_1 = 0$ and $v + d_2 = 0$.

(b) If $d_3 < 0$ then the solution set of Equation (6.3.8) is an empty set.

(c) If $d_3 > 0$ then Equation (6.3.8) reduces to $\dfrac{\alpha_1(u+d_1)^2}{d_3} + \dfrac{\alpha_2(v+d_2)^2}{d_3} = 1$, an ellipse or circle with $u + d_1 = 0$ and $v + d_2 = 0$ as the orthogonal principal axes.

Let $H(\mathbf{x}) = 6x^2 + 9y^2 + 4xy$ be the associated quadratic form for a class of curves. Then, $A = \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix}$, $\alpha_1 = 10, \alpha_2 = 5$ and $\sqrt{5}u = x + 2y, \sqrt{5}v = 2x - y$. Now, let $d_1 = \sqrt{5}, d_2 = -\sqrt{5}$ to get $2(x+2y+1)^2 + (2x-y-1)^2 = d_3$. Now vary d_3 to get different curves (see Figure 6.4 drawn using the package "MATHEMATICA").

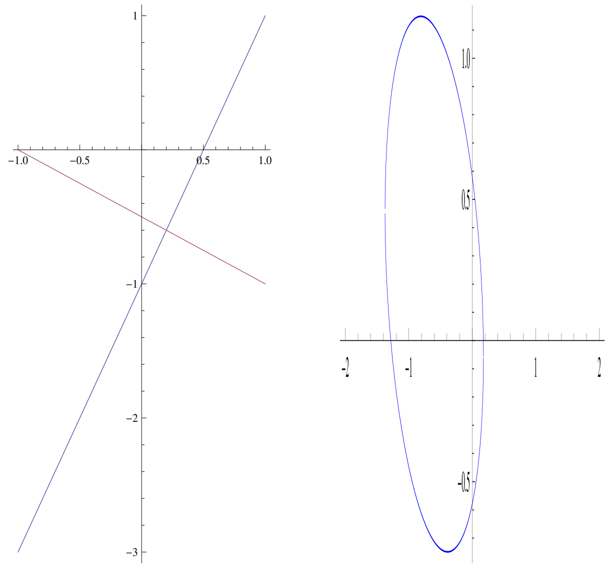


Figure 6.4: Curves for $d_3 = 0$ and $d_3 = 5$

Thus, we have considered all the possible cases and the required result follows. ■

Remark 6.3.20. *Observe that the condition* $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$ *implies that the principal axes of the conic are functions of the eigenvectors* $\mathbf{u}_1$ *and* $\mathbf{u}_2$*.*

EXERCISE **6.3.21.** *Sketch the graph of the following surfaces:*

1. $x^2 + 2xy + y^2 - 6x - 10y = 3$.

2. $2x^2 + 6xy + 3y^2 - 12x - 6y = 5$.

3. $4x^2 - 4xy + 2y^2 + 12x - 8y = 10$.

4. $2x^2 - 6xy + 5y^2 - 10x + 4y = 7$.

As a last application, we consider a quadratic in 3 variables, namely x, y and z. To do so, let $A = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} l \\ m \\ n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ with

$$\begin{aligned} f(x, y, z) &= \mathbf{x}^T A\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + q && (6.3.9) \\ &= ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + 2lx + 2my + 2nz + q && (6.3.10) \end{aligned}$$

Then, we observe the following:

1. As A is symmetric, $P^TAP = \text{diag}(\alpha_1, \alpha_2, \alpha_3)$, where $P = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$ is an orthogonal matrix and $(\alpha_i, \mathbf{u}_i)$, for $i = 1, 2, 3$ are eigen-pairs of A.

2. Let $\mathbf{y} = P^T\mathbf{x}$. Then, $f(x, y, z)$ reduces to

$$g(y_1, y_2, y_3) = \alpha_1 y_1^2 + \alpha_2 y_2^2 + \alpha_3 y_3^2 + 2l_1y_1 + 2l_2y_2 + 2l_3y_3 + q. \tag{6.3.11}$$

3. Depending on the values of α_i's, rewrite $g(y_1, y_2, y_3)$ to determine the center and the planes of symmetry of $f(x, y, z) = 0$.

Example 6.3.22. Determine the following quadrics $f(x, y, z) = 0$, where

1. $f(x, y, z) = 2x^2 + 2y^2 + 2z^2 + 2xy + 2xz + 2yz + 4x + 2y + 4z + 2$.

2. $f(x, y, z) = 3x^2 - y^2 + z^2 + 10$.

3. $f(x, y, z) = 3x^2 - y^2 + z^2 - 10$.

4. $f(x, y, z) = 3x^2 - y^2 + z - 10$.

Solution: Part 1 Here, $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ and $q = 2$. So, the orthogonal matrices $P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix}$ and $P^TAP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Hence, $f(x, y, z) = 0$ reduces to

$$4(y_1 + \frac{5}{4\sqrt{3}})^2 + (y_2 + \frac{1}{\sqrt{2}})^2 + (y_3 - \frac{1}{\sqrt{6}})^2 = \frac{9}{12}.$$

So, the standard form of the quadric is $4z_1^2 + z_2^2 + z_3^2 = \frac{9}{12}$, where $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} \frac{-5}{4\sqrt{3}} \\ \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} \frac{-3}{4} \\ \frac{1}{4} \\ \frac{-3}{4} \end{bmatrix}$ is the center and $x + y + z = 0, x - y = 0$ and $x + y - 2z = 0$ as the principal axes.

Part 2 Here $f(x, y, z) = 0$ reduces to $\frac{y^2}{10} - \frac{3x^2}{10} - \frac{z^2}{10} = 1$ which is the equation of a hyperboloid consisting of two sheets with center $\mathbf{0}$ and the axes x, y and z as the principal axes.

Part 3 Here $f(x, y, z) = 0$ reduces to $\frac{3x^2}{10} - \frac{y^2}{10} + \frac{z^2}{10} = 1$ which is the equation of a hyperboloid consisting of one sheet with center $\mathbf{0}$ and the axes x, y and z as the principal axes.

Part 4 Here $f(x, y, z) = 0$ reduces to $z = y^2 - 3x^2 + 10$ which is the equation of a hyperbolic paraboloid.

The different curves are given in Figure 6.5. These curves have been drawn using the package "MATHEMATICA".

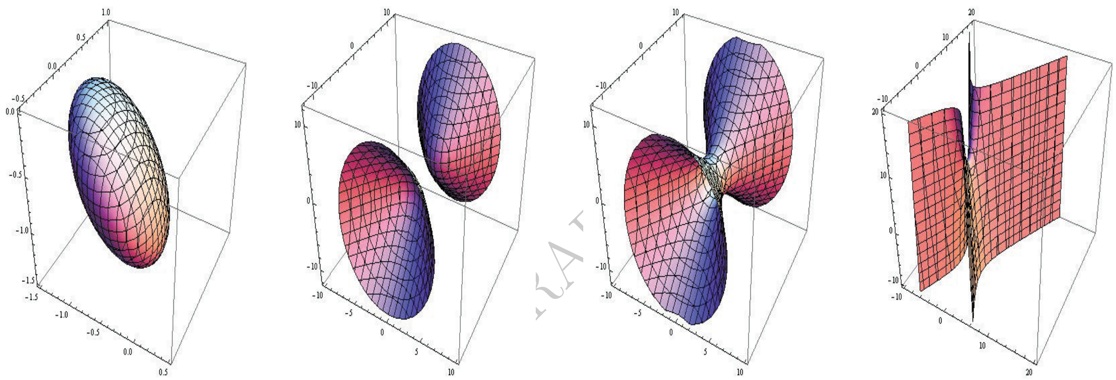


Figure 6.5: Ellipsoid, Hyperboloid of two sheets and one sheet, Hyperbolic Paraboloid

.

Chapter 7

Jordan Canonical form

7.1 Jordan Canonical form theorem

We start this chapter with the following theorem which generalizes the Schur Upper triangularization theorem.

Theorem 7.1.1. **[Generalized Schur's Theorem]** *Let $A \in \mathbb{M}_n(\mathbb{C})$. Suppose $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A with multiplicities $m_1, \ldots, m_k$, respectively. Then, there exists a non-singular matrix W such that*

$$W^{-1}AW = \bigoplus_{i=1}^{k} T_i, \ \textit{where, } T_i \in \mathbb{M}_{m_i}(\mathbb{C}), \ \textit{for } 1 \le i \le k$$

and T_i's are upper triangular matrices with constant diagonal λ_i. If A has real entries with real eigenvalues then W can be chosen to have real entries.

Proof. By Schur Upper Triangularization (see Lemma 6.2.12), there exists a unitary matrix U such that $U^*AU = T$, an upper triangular matrix with $\operatorname{diag}(T) = (\lambda_1, \ldots, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_k)$.

Now, for any upper triangular matrix B, a real number α and $i < j$, consider the matrix $F(B, i, j, \alpha) = E_{ij}(-\alpha) B E_{ij}(\alpha)$, where the matrix $E_{ij}(\alpha)$ is defined in Definition 2.1.13. Then, for $1 \le k, \ell \le n$,

$$(F(B,i,j,\alpha))_{k\ell} = \begin{cases} B_{ij} - \alpha B_{jj} + \alpha B_{ii}, & \text{whenever } k = i, \ell = j \\ B_{i\ell} - \alpha B_{j\ell}, & \text{whenever } \ell \neq j \\ B_{kj} + \alpha B_{ki}, & \text{whenever } k \neq i \\ B_{k\ell}, & \text{otherwise.} \end{cases} \tag{7.1.1}$$

Now, using Equation (7.1.1), the diagonal entries of $F(T, i, j, \alpha)$ and T are equal and

$$(F(T,i,j,\alpha))_{ij} = \begin{cases} T_{ij}, & \text{whenever } T_{jj} = T_{ii} \\ 0, & \text{whenever } T_{jj} \neq T_{ii} \text{ and } \alpha = \frac{T_{ij}}{T_{jj} - T_{ii}}. \end{cases}$$

Thus, if we denote the matrix $F(T, i, j, \alpha)$ by T_1 then $(F(T_1, i-1, j, \alpha))_{i-1,j} = 0$, for some choice of α, whenever $(T_1)_{i-1,i-1} \neq T_{jj}$. Moreover, this operation also preserves the 0 created by $F(T, i, j, \alpha)$ at (i, j)-th place. Similarly, $F(T_1, i, j+1, \alpha)$ preserves the 0 created by $F(T, i, j, \alpha)$ at (i, j)-th place. So, we can successively apply the following sequence of operations to get

$$T \to F(T, m_1, m_1+1, \alpha) = T_1 \to F(T_1, m_1-1, m_1+1, \beta) \to \cdots \to F(T_{m_1-1}, 1, m_1+1, \gamma) = T_{m_1},$$

where $\alpha, \beta, \ldots, \gamma$ are appropriately chosen and $T_{m_1}[:, m_1 + 1] = \lambda_2 \mathbf{e}_{m_1+1}$. Thus, observe that the above operation can be applied for different choices of i and j with $i < j$ to get the required result. ■

EXERCISE **7.1.2.** *Apply Theorem 7.1.1 to the matrix given below for better understanding.*

$$\left[\begin{array}{ccc|ccc|ccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
0 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline
0 & 0 & 0 & 2 & 3 & 4 & 5 & 6 & 7 \\
0 & 0 & 0 & 0 & 2 & 3 & 4 & 5 & 6 \\
0 & 0 & 0 & 0 & 0 & 2 & 3 & 4 & 5 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 3 & 4 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3
\end{array}\right].$$

Definition 7.1.3. **[Jordan Block and Jordan Matrix]**

1. Let $\lambda \in \mathbb{C}$ and k be a positive integer. Then, by the **Jordan block** $J_k(\lambda) \in \mathbb{M}_k(\mathbb{C})$, we understand the matrix
$$\begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}.$$

2. A **Jordan matrix** is a direct sum of Jordan blocks. That is, if A is a Jordan matrix having r blocks then there exist positive integers k_i's and complex numbers λ_i's (not necessarily distinct), for $1 \le i \le r$ such that
$$A = J_{k_1}(\lambda_1) \oplus \cdots \oplus J_{k_r}(\lambda_r).$$

We now give some examples of Jordan matrices with diagonal entries 0.

Example 7.1.4. 1. $J_1(0) = \begin{bmatrix} 0 \end{bmatrix}$ is the only Jordan matrix of size 1.

2. $J_1(0) \oplus J_1(0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $J_2(0) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ are Jordan matrices of size 2.
3. Even though , $J_1(0) \oplus J_2(0)$ and $J_2(0) \oplus J_1(0)$ are two Jordan matrices of size 3, we do not differentiate between them as they are similar (use permutations).
4. $J_1(0) \oplus J_1(0) \oplus J_1(0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $J_2(0) \oplus J_1(0) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $J_3(0) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ are Jordan matrices of size 3.
5. Observe that the number of Jordan matrices of size 4 with 0 on the diagonal are 5.

We now give some properties of the Jordan blocks. The proofs are immediate and hence left for the reader. They will be used in the proof of subsequent results.

Remark 7.1.5. **[Jordan blocks]** *Fix a positive integer k. Then,*

1. $J_k(\lambda)$ *is an upper triangular matrix with λ as an eigenvalue.*
2. $J_k(\lambda) = \lambda I_k + J_k(0)$.
3. $\text{ALG.MUL}_\lambda(J_k(\lambda)) = k$.
4. *The matrix $J_k(0)$ satisfies the following properties.*
   (a) $\mathsf{Rank}((J_k(0)^i) = k - i$, *for* $1 \le i \le k$.
   (b) $J_k^T(0)J_k(0) = \begin{bmatrix} 0 & 0 \\ 0 & I_{k-1} \end{bmatrix}$.
   (c) $J_k(0)^p = 0$ *whenever* $p \ge k$.
   (d) $J_k(0)\mathbf{e}_i = \mathbf{e}_{i-1}$ *for* $i = 2, \ldots, k$.
   (e) $\Big(I - J_k^T(0)J_k(0)\Big)\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{0} \end{bmatrix} = \langle \mathbf{x}, \mathbf{e}_1 \rangle \mathbf{e}_1$.
5. *Thus, using Remark 7.1.5.4d* $\text{GEO.MUL}_\lambda(J_k(\lambda)) = 1$.

EXERCISE **7.1.6.** 1. *Fix a positive integer k and a complex number λ. Then, prove that*
   (a) $\mathsf{Rank}(J_k(\lambda) - \lambda I_k) = k - 1$.
   (b) $\mathsf{Rank}(J_k(\lambda) - \alpha I_k) = k$, *whenever $\alpha \ne \lambda$. Or equivalently, for all $\alpha \ne \lambda$ the matrix $J_k(\lambda) - \alpha I_k$ is invertible.*
   (c) *for* $1 \le i \le k$, $\mathsf{Rank}((J_k(\lambda) - \lambda I_k)^i) = k - i$.
   (d) *for* $\alpha \ne \lambda$, $\mathsf{Rank}((J_k(\lambda) - \alpha I_k)^i) = k$, *for all* i.

2. *Let J be a Jordan matrix that contains ℓ Jordan blocks for λ. Then, prove that*
   (a) $\mathsf{Rank}(J - \lambda I) = n - \ell$.
   (b) *J has ℓ linearly independent eigenvectors for λ.*
   (c) $\mathsf{Rank}(J - \lambda I) \ge \mathsf{Rank}((J - \lambda I)^2) \ge \mathsf{Rank}((J - \lambda I)^3) \ge \cdots$.

3. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, prove that $AJ_n(\lambda) = J_n(\lambda)A$ if and only if $AJ_n(0) = J_n(0)A$.*

Definition 7.1.7. **[Index of an Eigenvalue]** Let J be a Jordan matrix containing $J_t(\lambda)$, for some positive integer t and some complex number λ. Then, the smallest value of k for which $\mathsf{Rank}((J - \lambda I)^k)$ stops decreasing is the order of the largest Jordan block $J_k(\lambda)$ in J. This number k is called the **index of the eigenvalue** λ.

Lemma 7.1.8. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be strictly upper triangular. Then, A is similar to a direct sum of Jordan blocks. That is, there exists a non-singular matrix S and integers $n_1 \ge \ldots \ge n_m \ge 1$ such that*

$$A = S^{-1}\Big(J_{n_1}(0) \oplus \cdots \oplus J_{n_m}(0)\Big)S.$$

If $A \in \mathbb{M}_n(\mathbb{R})$ then S can be chosen to have real entries.

Proof. We will prove the result by induction on n. For $n = 1$, the statement is trivial. So, let the result be true for matrices of size $\le n - 1$ and let $A \in \mathbb{M}_n(\mathbb{C})$ be strictly upper triangular. Then, $A = \begin{bmatrix} 0 & \mathbf{a}^T \\ 0 & A_1 \end{bmatrix}$. By induction hypothesis there exists an invertible matrix S_1 such that

$$A_1 = S_1^{-1}\Big(J_{n_1}(0) \oplus \cdots \oplus J_{n_m}(0)\Big)S_1 \text{ with } \sum_{i=1}^{m} n_i = n - 1.$$

Thus,

$$\begin{bmatrix}1 & 0\\ 0 & S_1^{-1}\end{bmatrix} A \begin{bmatrix}1 & 0\\ 0 & S_1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & S_1^{-1}\end{bmatrix}\begin{bmatrix}0 & \mathbf{a}^T\\ 0 & A_1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & S_1\end{bmatrix} = \begin{bmatrix}0 & \mathbf{a}^T S_1\\ 0 & S^{-1}A_1S_1\end{bmatrix} = \begin{bmatrix}0 & \mathbf{a}_1^T & \mathbf{a}_2^T\\ 0 & J_{n_1}(0) & 0\\ 0 & 0 & J\end{bmatrix},$$

where $S_1^{-1}\Big(J_{n_1}(0)\oplus\cdots\oplus J_{n_m}(0)\Big)S_1 = J_{n_1}(0)\oplus J$ and $\mathbf{a}^T S_1 = \begin{bmatrix}\mathbf{a}_1^T & \mathbf{a}_2^T\end{bmatrix}$. Now, writing J_{n_1} to mean $J_{n_1}(0)$ and using Remark 7.1.5.4e, we have

$$\begin{bmatrix}1 & -\mathbf{a}_1^T J_{n_1}^T & 0\\ 0 & I_{n_1} & 0\\ 0 & 0 & I\end{bmatrix}\begin{bmatrix}0 & \mathbf{a}_1^T & \mathbf{a}_2^T\\ 0 & J_{n_1} & 0\\ 0 & 0 & J\end{bmatrix}\begin{bmatrix}1 & \mathbf{a}_1^T J_{n_1}^T & 0\\ 0 & I_{n_1} & 0\\ 0 & 0 & I\end{bmatrix} = \begin{bmatrix}0 & \langle \mathbf{a}_1, \mathbf{e}_1\rangle \mathbf{e}_1^T & \mathbf{a}_2^T\\ 0 & J_{n_1} & 0\\ 0 & 0 & J\end{bmatrix}.$$

So, we now need to consider two cases depending on whether $\langle \mathbf{a}_1, \mathbf{e}_1\rangle = 0$ or $\langle \mathbf{a}_1, \mathbf{e}_1\rangle \neq 0$. In the first case, A is similar to $\begin{bmatrix}0 & 0 & \mathbf{a}_2^T\\ 0 & J_{n_1} & 0\\ 0 & 0 & J\end{bmatrix}$. This in turn is similar to $\begin{bmatrix}J_{n_1} & 0 & 0\\ 0 & 0 & \mathbf{a}_2^T\\ 0 & 0 & J\end{bmatrix}$ by permuting the first row and column. At this stage, one can apply induction and if necessary do a block permutation, in order to keep the block sizes in decreasing order.

So, let us now assume that $\langle \mathbf{a}_1, \mathbf{e}_1\rangle \neq 0$. Then, writing $\alpha = \langle \mathbf{a}_1, \mathbf{e}_1\rangle$, we have

$$\begin{bmatrix}\frac{1}{\alpha} & 0 & 0\\ 0 & I & 0\\ 0 & 0 & \frac{1}{\alpha}I\end{bmatrix}\begin{bmatrix}0 & \alpha\mathbf{e}_1^T & \mathbf{a}_2^T\\ 0 & J_{n_1} & 0\\ 0 & 0 & J\end{bmatrix}\begin{bmatrix}\alpha & 0 & 0\\ 0 & I & 0\\ 0 & 0 & \alpha I\end{bmatrix} = \begin{bmatrix}0 & \mathbf{e}_1^T & \mathbf{a}_2^T\\ 0 & J_{n_1} & 0\\ 0 & 0 & J\end{bmatrix} \equiv \begin{bmatrix}J_{n_1+1} & \mathbf{e}_1\mathbf{a}_2^T\\ 0 & J\end{bmatrix}.$$

Now, using Remark 7.1.5.4c, verify that

$$\begin{bmatrix}I & \mathbf{e}_{i+1}\mathbf{a}_2^T J^{i-1}\\ 0 & I\end{bmatrix}\begin{bmatrix}J_{n_1+1} & \mathbf{e}_i\mathbf{a}_2^T J^{i-1}\\ 0 & J\end{bmatrix}\begin{bmatrix}I & -\mathbf{e}_{i+1}\mathbf{a}_2^T J^{i-1}\\ 0 & I\end{bmatrix} = \begin{bmatrix}J_{n_1+1} & \mathbf{e}_{i+1}\mathbf{a}_2^T J^{i}\\ 0 & J\end{bmatrix}, \text{ for } i \geq 1.$$

Hence, for $p = n - n_1 - 1$, we have

$$\begin{bmatrix}I & \mathbf{e}_{p+1}\mathbf{a}_2^T J^{p-1}\\ 0 & I\end{bmatrix}\cdots\begin{bmatrix}I & \mathbf{e}_2\mathbf{a}^T\\ 0 & I\end{bmatrix}\begin{bmatrix}J_{n_1+1} & \mathbf{e}_1\mathbf{a}^T\\ 0 & J\end{bmatrix}\begin{bmatrix}I & -\mathbf{e}_2\mathbf{a}^T\\ 0 & I\end{bmatrix}\cdots\begin{bmatrix}I & -\mathbf{e}_{p+1}\mathbf{a}^T J^{p-1}\\ 0 & I\end{bmatrix} = \begin{bmatrix}J_{n_1+1} & 0\\ 0 & J\end{bmatrix}.$$

If necessary, we need to do a block permutation, in order to keep the block sizes in decreasing order. Hence, the required result follows. ■

Exercise **7.1.9.** *Convert* $\begin{bmatrix}0 & 1 & 1\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$ *to* $J_3(0)$ *and* $\begin{bmatrix}0 & 1 & 2\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$ *to* $J_2(0)\oplus J_1(0)$.

Corollary 7.1.10. $A \in \mathbb{M}_n(\mathbb{C})$. *Then, A is similar to J, a Jordan matrix.*

Proof. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of A with algebraic multiplicities $m_1, \ldots, m_k$. By Theorem 7.1.1, there exists a non-singular matrix S such that $S^{-1}AS = \bigoplus\limits_{i=1}^{k} T_i$, where T_i is an upper triangular with diagonal $(\lambda_i, \ldots, \lambda_i)$. Thus $T_i - \lambda_i I_{m_i}$ is a strictly upper triangular matrix. Thus, by Theorem 7.1.8, there exist a non-singular matrix S_i such that

$$S_i^{-1}\Big(T_i - \lambda_i I_{m_i}\Big)S_i = J(0),$$

a Jordan matrix with 0 on the diagonal and the size of the Jordan blocks decreases as we move down the diagonal. So, $S_i^{-1}T_iS_i = J(\lambda_i)$ is a Jordan matrix with λ_i on the diagonal and the size of the Jordan blocks decreases as we move down the diagonal.

Now, take $W = S\left(\bigoplus\limits_{i=1}^{k} S_i\right)$. Then, verify that $W^{-1}AW$ is a Jordan matrix. ∎

Let $A \in \mathbb{M}_n(\mathbb{C})$. Suppose $\lambda \in \sigma(A)$ and J is a Jordan matrix that is similar to A. Then, for each fixed $i, 1 \le i \le n$, by $\ell_i(\lambda)$, we denote the number of Jordan blocks $J_k(\lambda)$ in J for which $k \ge i$. Then, the next result uses Exercise 7.1.6 to determine the number $\ell_i(\lambda)$.

Remark 7.1.11. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Suppose $\lambda \in \sigma(A)$ and J is a Jordan matrix that is similar to A. Then, for $1 \le k \le n$,*

$$\ell_k(\lambda) = \mathsf{Rank}(A-\lambda I)^{k-1} - \mathsf{Rank}(A-\lambda I)^{k}.$$

Proof. In view of Exercise 7.1.6, we need to consider only the Jordan blocks $J_k(\lambda)$, for different values of k. Hence, without loss of generality, let us assume that $J = \bigoplus\limits_{i=1}^{n} a_iJ_i(\lambda)$, where a_i's are non-negative integers and J contains exactly a_i copies of the Jordan block $J_i(\lambda)$, for $1 \le i \le n$. Then, by definition and Exercise 7.1.6, we observe the following:

1. $n = \sum\limits_{i\ge 1} ia_i$.
2. $\mathsf{Rank}(J-\lambda I) = \sum\limits_{i\ge 2}(i-1)a_i$.
3. $\mathsf{Rank}((J-\lambda I)^2) = \sum\limits_{i\ge 3}(i-2)a_i$.
4. In general, for $1 \le k \le n$, $\mathsf{Rank}((J-\lambda I)^k) = \sum\limits_{i\ge k+1}(i-k)a_i$.

Thus, writing ℓ_i in place of $\ell_i(\lambda)$, we get

$$\begin{aligned}
\ell_1 &= \sum_{i\ge 1} a_i = \sum_{i\ge 1} ia_i - \sum_{i\ge 2}(i-1)a_i = n - \mathsf{Rank}(J-\lambda I),\\
\ell_2 &= \sum_{i\ge 2} a_i = \sum_{i\ge 2}(i-1)a_i - \sum_{i\ge 3}(i-2)a_i = \mathsf{Rank}(J-\lambda I) - \mathsf{Rank}((J-\lambda I)^2),\\
&\vdots\\
\ell_k &= \sum_{i\ge k} a_i = \sum_{i\ge k}(i-(k-1))a_i - \sum_{i\ge k+1}(i-k)a_i = \mathsf{Rank}((J-\lambda I)^{k-1}) - \mathsf{Rank}((J-\lambda I)^k).
\end{aligned}$$

Now, the required result follows as rank is invariant under similarity operation and the matrices J and A are similar. ∎

Lemma 7.1.12. [Similar Jordan Matrices] *Let J and J' be two similar Jordan matrices of size n. Then, J is a block permutation of J'.*

Proof. For $1 \le i \le n$, let ℓ_i and ℓ_i' be, respectively, the number of Jordan blocks of J and J' of size at least i corresponding to λ. Since J and J' are similar, the matrices $(J-\lambda I)^i$ and $(J'-\lambda I)^i$ are similar for all $i, 1 \le i \le n$. Therefore, their ranks are equal for all $i \ge 1$ and hence, $\ell_i = \ell_i'$ for all $i \ge 1$. Thus the required result follows. ∎

We now state the main result of this section which directly follows from Lemma 6.2.12, Theorem 7.1.1 and Corollary 7.1.10 and hence the proof is omitted.

Theorem 7.1.13. **[Jordan Canonical Form Theorem]** *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A is similar to a Jordan matrix J, which is unique up to permutation of Jordan blocks. If $A \in \mathbb{M}_n(\mathbb{R})$ and has real eigenvalues then the similarity transformation matrix S may be chosen to have real entries. This matrix J is called the the Jordan canonical form of A, denoted* Jordan CF(A).

We now start with a few examples and observations.

Example 7.1.14. Let us use the idea from Lemma 7.1.11 to find the Jordan Canonical Form of the following matrices.

1. Let $A = J_4(0)^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

   Solution: Note that $\ell_1 = 4 - \mathsf{Rank}(A - 0I) = 2$. So, there are two Jordan blocks.

   Also, $\ell_2 = \mathsf{Rank}(A - 0I) - \mathsf{Rank}((A - 0I)^2) = 2$. So, there are at least 2 Jordan blocks of size 2. As there are exactly two Jordan blocks, both the blocks must have size 2. Hence, Jordan CF$(A) = J_2(0) \oplus J_2(0)$.

2. Let $A_1 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

   Solution: Let $B = A_1 - I$. Then, $\ell_1 = 4 - \mathsf{Rank}(B) = 1$. So, B has exactly one Jordan block and hence A_1 is similar to $J_4(1)$.

3. $A_2 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

   Solution: Let $C = A_2 - I$. Then, $\ell_1 = 4 - \mathsf{Rank}(C) = 2$. So, C has exactly two Jordan blocks. Also, $\ell_2 = \mathsf{Rank}(C) - \mathsf{Rank}(C^2) = 1$ and $\ell_3 = \mathsf{Rank}(C^2) - \mathsf{Rank}(C^3) = 1$. So, there is at least 1 Jordan blocks of size 3.

   Thus, we see that there are two Jordan blocks and one of them is of size 3. Also, the size of the matrix is 4. Thus, A_2 is similar to $J_3(1) \oplus J_1(1)$.

4. Let $A = J_4(1)^2 \oplus A_1 \oplus A_2$, where A_1 and A_2 are given in the previous exercises.

   Solution: One can directly get the answer from the previous exercises as the matrix A is already in the block diagonal form. But, we compute it again for better understanding.

   Let $B = A - I$. Then, $\ell_1 = 16 - \mathsf{Rank}(B) = 5$, $\ell_2 = \mathsf{Rank}(B) - \mathsf{Rank}(B^2) = 11 - 7 = 4$, $\ell_3 = \mathsf{Rank}(B^2) - \mathsf{Rank}(B^3) = 7 - 3 = 4$ and $\ell_4 = \mathsf{Rank}(B^3) - \mathsf{Rank}(B^4) = 3 - 0 = 3$.

   Hence, $J_4(1)$ appears thrice (as $\ell_4 = 3$ and $\ell_5 = 0$), $J_3(1)$ also appears once (as $\ell_3 - \ell_4 = 1$), $J_2(1)$ does not appear as (as $\ell_2 - \ell_3 = 0$) and $J_1(1)$ appears once (as $\ell_1 - \ell_2 = 1$). Thus, the required result follows.

Remark 7.1.15. **[Observations about** Jordan CF(A)**]**

1. *What are the steps to find* JORDAN CFA*?*

   Ans: *Let $\lambda_1,\ldots,\lambda_k$ be the distinct eigenvalues of A. Now, apply the Schur Upper Triangularization Lemma (see Lemma 6.2.12) to get an upper triangular matrix, say T such that the diagonal entries of T are $\lambda_1,\ldots,\lambda_1,\lambda_2,\ldots,\lambda_2,\ldots,\lambda_k,\ldots,\lambda_k$. Now, apply similarity transformations (see Theorem 7.1.1) to get $T=\bigoplus\limits_{i=1}^{k} T_i$, where each diagonal entry of T_i is λ_i. Then, for each $i, 1\le i\le k$, use Theorem 7.1.8 to get an invertible matrix S_i such that $S_i^{-1}(T_i-\lambda_i I)S_i=\widetilde{J}_i$, a Jordan matrix. Thus, we obtain a Jordan matrix $J_i=\widetilde{J}_i+\lambda_i I=S_i^{-1}T_iS_i$, where each diagonal entry of J_i is λ_i. Hence, $S=\bigoplus\limits_{i=1}^{k} S_i$ converts $T=\bigoplus\limits_{i=1}^{k} T_i$ into the required Jordan matrix.*

2. *Let $A\in\mathbb{M}_n(\mathbb{C})$ be a diagonalizable matrix. Then, by definition, A is similar to $\bigoplus\limits_{i=1}^{n}\lambda_i$, where $\lambda_i\in\sigma(A)$, for $1\le i\le n$. Thus,* JORDAN CF$(A)=\bigoplus\limits_{i=1}^{n}\lambda_i$*, up to a permutation of λ_i's.*

3. *In general, the computation of* JORDAN CF(A) *is not numerically stable. To understand this, let $A_\epsilon=\begin{bmatrix}\epsilon&0\\1&0\end{bmatrix}$. Then, A_ϵ is diagonalizable as A has distinct eigenvalues. So,* JORDAN CF$(A_\epsilon)=\begin{bmatrix}\epsilon&0\\0&0\end{bmatrix}$.

   Whereas, for $A=\begin{bmatrix}0&0\\1&0\end{bmatrix}$, we know that JORDAN CF$(A)=\begin{bmatrix}0&1\\0&0\end{bmatrix}\neq\lim\limits_{\epsilon\to0}$ JORDAN CF(A_ϵ). *Thus, a small change in the entries of A may change* JORDAN CF(A) *significantly.*

4. *Let $A\in\mathbb{M}_n(\mathbb{C})$ and $\epsilon>0$ be given. Then, there exists an invertible matrix S such that $S^{-1}AS=\bigoplus\limits_{i=1}^{k} J_{n_i}(\lambda_i,\epsilon)$, where $J_{n_i}(\lambda_i,\epsilon)$ is obtained from $J_{n_i}(\lambda_i)$ by replacing each off diagonal entry 1 by an ϵ. To get this, define $Di(\epsilon)=\ diag(1,\epsilon,\epsilon^2,\ldots,\epsilon^{n_i-1})$, for $1\le i\le k$. Now compute $\bigoplus\limits_{i=1}^{k}\left((Di(\epsilon))^{-1}J_{n_i}(\lambda_i)Di(\epsilon)\right)$.*

5. *Let* JORDAN CF(A) *contain ℓ Jordan blocks for λ. Then, A has ℓ linearly independent eigenvectors for λ.*

   For if, A has at least $\ell+1$ linearly independent eigenvectors for λ, then $\dim(\text{NULL}(A-\lambda I))>\ell$. *So,* $\mathsf{Rank}(A-\lambda I)<n-\ell$. *But, the number of Jordan blocks for λ in A is ℓ. Thus, we must have* $\mathsf{Rank}(J-\lambda I)=n-\ell$*, a contradiction.*

6. *Let $\lambda\in\sigma(A)$. Then, by Remark 7.1.5.5,* GEO.MUL$_\lambda(A)$ = *the number of Jordan blocks $J_k(\lambda)$ in* JORDAN CF(A).

7. *Let $\lambda\in\sigma(A)$. Then, by Remark 7.1.5.3,* ALG.MUL$_\lambda(A)$ = *the sum of the sizes of all Jordan blocks $J_k(\lambda)$ in* JORDAN CF(A).

8. *Let $\lambda\in\sigma(A)$. Then,* JORDAN CF(A) *<u>does not</u> get determined by* ALG.MUL$_\lambda(A)$ *and* GEO.MUL$_\lambda(A)$. *For example,* $\begin{bmatrix}0\end{bmatrix}\oplus\begin{bmatrix}0&1\\0&0\end{bmatrix}\oplus\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}$ *and* $\begin{bmatrix}0&1\\0&0\end{bmatrix}\oplus\begin{bmatrix}0&1\\0&0\end{bmatrix}\oplus\begin{bmatrix}0&1\\0&0\end{bmatrix}$

are different JORDAN CF*s but they have the same algebraic and geometric multiplicities.*

9. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Suppose that, for each $\lambda \in \sigma(A)$, the values of* $\mathsf{Rank}(A - \lambda I)^k$, *for $k = 1, \ldots, n$ are known. Then, using Remark 7.1.11,* JORDAN CF(A) *can be computed. But, note here that finding rank is* <u>*numerically unstable*</u> *as $\begin{bmatrix} \epsilon \end{bmatrix}$ has rank 1 but it converges to $\begin{bmatrix} 0 \end{bmatrix}$ which has a different rank.*

Theorem 7.1.16. [**A is similar to A^T**] *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A is similar to A^T.*

Proof. Let $K_n = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}$. Then, observe that $K^{-1} = K$ and $KJ_n(a)K = J_n(a)^T$, as the (i,j)-th entry of A goes to $(n-i+1, n-j+1)$-th position in KAK. Hence,

$$\left[\bigoplus K_{n_i}\right] \left[\bigoplus J_{n_i}(\lambda_i)\right] \left[\bigoplus K_{n_i}\right] = \left[\bigoplus J_{n_i}(\lambda_i)\right]^T.$$

Thus, J is similar to J^T. But, A is similar to J and hence A^T is similar to J^T and finally we get A is similar to A^T. Therefore, the required result follows. ■

7.2 Minimal polynomial

We start this section with the following definition. Recall that a polynomial $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ with $a_n = 1$ is called a monic polynomial.

Definition 7.2.1. [**Companion Matrix**] Let $P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0$ be a monic polynomial in t of degree n. Then, the $n \times n$ matrix $A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 0 & 0 & & 1 & -a_{n-1} \end{bmatrix}$, denoted $A(n : a_0, \ldots, a_{n-1})$ or COMPANION(P), is called the **companion matrix** of $P(t)$.

Definition 7.2.2. [**Annihilating Polynomial**] Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, the polynomial $P(t)$ is said to **annihilate** (destroy) A if $P(A) = \mathbf{0}$.

Let $P(x)$ be the characteristic polynomial of A. Then, by the Cayley-Hamilton Theorem, $P(A) = \mathbf{0}$. So, if $f(x) = P(x)g(x)$, for any multiple of $g(x)$, then $f(A) = P(A)g(A) = \mathbf{0}g(A) = \mathbf{0}$. Thus, there are infinitely many polynomials which annihilate A. In this section, we will concentrate on a monic polynomial of least positive degree that annihilates A.

Definition 7.2.3. [**Minimal polynomial**] Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, the **minimal polynomial** of A, denoted $m_A(x)$, is a monic polynomial of least positive degree satisfying $m_A(A) = \mathbf{0}$.

Theorem 7.2.4. *Let A be the companion matrix of the monic polynomial $P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0$. Then, $P(t)$ is both the characteristic and the minimal polynomial of A.*

Proof. Expanding $\det(tI_n - \textsc{Companion}(P))$ along the first row, we have

$$\begin{aligned}\det(tI_n - A(n:a_0,\ldots,a_{n-1})) &= t\det(tI_{n-1} - A(n-1:a_1,\ldots,a_{n-1})) + (-1)^{n+1}a_0(-1)^{n-1}\\ &= t^2\det(tI_{n-2} - A(n-2:a_2,\ldots,a_{n-1})) + a_0 + a_1 t\\ &\ \vdots\\ &= P(t).\end{aligned}$$

Thus, $P(t)$ is the characteristic polynomial of A and hence $P(A) = \mathbf{0}$.

We will now show that $P(t)$ is the minimal polynomial of A. To do so, we first observe that $A\mathbf{e}_1 = \mathbf{e}_2, \ldots, A\mathbf{e}_{n-1} = \mathbf{e}_n$. That is,

$$A^k\mathbf{e}_1 = \mathbf{e}_{k+1}, \text{ for } 1 \le k \le n-1. \tag{7.2.1}$$

Now, Suppose we have a monic polynomial $Q(t) = t^m + b_{m-1}t^{m-1} + \cdots + b_0$, with $m < n$, such that $Q(A) = \mathbf{0}$. Then, using Equation (7.2.1), we get

$$\mathbf{0} = Q(A)\mathbf{e}_1 = A^m\mathbf{e}_1 + b_{m-1}A^{m-1}\mathbf{e}_1 + \cdots + b_0 I\mathbf{e}_1 = \mathbf{e}_{m+1} + b_{m-1}\mathbf{e}_m + \cdots + b_0\mathbf{e}_1,$$

a contradiction to the linear independence of $\{\mathbf{e}_1, \ldots, \mathbf{e}_{m+1}\} \subseteq \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. ■

The next result gives us the existence of such a polynomial for every matrix A. To do so, recall that the well-ordering principle implies that if S is a subset of natural numbers then it contains a least element.

Lemma 7.2.5. **[Existence of the Minimal Polynomial]** *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, there exists a unique monic polynomial $m(x)$ of minimum (positive) degree such that $m(A) = \mathbf{0}$. Further, if $f(x)$ is any polynomial with $f(A) = \mathbf{0}$ then $m(x)$ divides $f(x)$.*

Proof. Let $P(x)$ be the characteristic polynomial of A. Then, $\deg(P(x)) = n$ and by the Cayley-Hamilton Theorem, $P(A) = \mathbf{0}$. So, consider the set

$$S = \{\deg(f(x)) : f(x) \text{ is a nonzero polynomial, } f(A) = \mathbf{0}\}.$$

Then, S is a non-empty subset of $\mathbb{N}$ as $n \in S$. Thus, by well-ordering principle there exists a smallest positive integer, say M, and a corresponding polynomial, say $m(x)$, such that $\deg(m(x)) = M$, $m(A) = \mathbf{0}$.

Also, without loss of generality, we can assume that $m(x)$ is monic and unique (non-uniqueness will lead to a polynomial of smaller degree in S).

Now, suppose there is a polynomial $f(x)$ such that $f(A) = \mathbf{0}$. Then, by division algorithm, there exist polynomials $q(x)$ and $r(x)$ such that $f(x) = m(x)q(x) + r(x)$, where either $r(x)$ is identically the zero polynomial of $\deg(r(x)) < M = \deg(m(x))$. As

$$\mathbf{0} = f(A) = m(A)q(A) + r(A) = \mathbf{0}q(A) + r(A) = r(A),$$

we get $r(A) = \mathbf{0}$. But, $m(x)$ was the least degree polynomial with $m(A) = \mathbf{0}$ and hence $r(x)$ is the zero polynomial. That is, $m(x)$ divides $f(x)$. ■

As an immediate corollary, we have the following result.

Corollary 7.2.6. [Minimal polynomial divides the Characteristic Polynomial] *Let $m_A(x)$ and $P_A(x)$ be, respectively, the minimal and the characteristic polynomials of $A \in \mathbb{M}_n(\mathbb{C})$.*

1. *Then, $m_A(x)$ divides $P_A(x)$.*
2. *Further, if λ is an eigenvalue of A then $m_A(\lambda) = 0$.*

Proof. The first part following directly from Lemma 7.2.5. For the second part, let $(\lambda, \mathbf{x})$ be an eigen-pair. Then, $f(A)\mathbf{x} = f(\lambda)\mathbf{x}$, for any polynomial of f, implies that

$$m_A(\lambda)\mathbf{x} = m_A(A)\mathbf{x} = \mathbf{0}\mathbf{x} = \mathbf{0}.$$

But, $\mathbf{x} \neq \mathbf{0}$ and hence $m_A(\lambda) = 0$. Thus, the required result follows. ∎

we also have the following result.

Lemma 7.2.7. *Let A and B be two similar matrices. Then, they have the same minimal polynomial.*

Proof. Since A and B are similar, there exists an invertible matrix S such that $A = S^{-1}BS$. Hence, $f(A) = F(S^{-1}BS) = S^{-1}f(B)S$, for any polynomial f. Hence, $m_A(A) = \mathbf{0}$ if and only if $m_A(B) = \mathbf{0}$ and thus the required result follows. ∎

Theorem 7.2.8. *Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of A. If n_i is the size of the largest Jordan block for λ_i in $J =$ Jordan CFA then*

$$m_A(x) = \prod_{i=1}^{k} (x - \lambda_i)^{n_i}.$$

Proof. Using 7.2.6, we see that $m_A(x) = \prod\limits_{i=1}^{k} (x - \lambda_i)^{\alpha_i}$, for some α_i's with $1 \leq \alpha_i \leq \text{Alg.Mul}_{\lambda_i}(A)$. As $m_A(A) = \mathbf{0}$, using Lemma 7.2.7 we have $m_A(J) = \prod\limits_{i=1}^{k} (J - \lambda_i I)^{\alpha_i} = \mathbf{0}$. But, observe that for the Jordan block $J_{n_i}(\lambda_i)$, one has

1. $(J_{n_i}(\lambda_i) - \lambda_i I)^{\alpha_i} = \mathbf{0}$ if and only if $\alpha_i \geq n_i$, and
2. $(J_{n_m}(\lambda_m) - \lambda_i I)^{\alpha_i}$ is invertible, for all $m \neq i$.

Thus $\prod\limits_{i=1}^{k} (J - \lambda_i I)^{n_i} = \mathbf{0}$ and $\prod\limits_{i=1}^{k} (x - \lambda_i)^{n_i}$ divides $\prod\limits_{i=1}^{k} (x - \lambda_i)^{\alpha_i} = m_A(x)$ and $\prod\limits_{i=1}^{k} (x - \lambda_i)^{n_i}$ is a monic polynomial, the result follows. ∎

As an immediate consequence, we also have the following result which corresponds to the converse of the above theorem.

Theorem 7.2.9. *Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of A. If the minimal polynomial of A equals $\prod\limits_{i=1}^{k} (x - \lambda_i)^{n_i}$ then n_i is the size of the largest Jordan block for λ_i in $J =$ Jordan CFA.*

Proof. It directly follows from Theorem 7.2.8. ∎

We now give equivalent conditions for a square matrix to be diagonalizable.

Theorem 7.2.10. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, the following statements are equivalent.*

1. *A is diagonalizable.*
2. *Every zero of $m_A(x)$ has multiplicity* 1.
3. *Whenever $m_A(\alpha)=0$, for some α, then* $\frac{d}{dx}m_A(x)\big|_{x=\alpha}\neq 0$.

Proof. Part 1 $\Rightarrow$ Part 2. If A is diagonalizable, then each Jordan block in $J = \text{JORDAN CF}A$ has size 1. Hence, by Theorem 7.2.8, $m_A(x)=\prod_{i=1}^{k}(x-\lambda_i)$, where λ_i's are the distinct eigenvalues of A.

Part 2 $\Rightarrow$ Part 3. Let $m_A(x)=\prod_{i=1}^{k}(x-\lambda_i)$, where λ_i's are the distinct eigenvalues of A. Then, $m_A(x)=0$ if and only if $x=\lambda_i$, for some $i, 1\le i\le k$. In that case, it is easy to verify that $\frac{d}{dx}m_A(x)\neq 0$, for each λ_i.

Part 3 $\Rightarrow$ Part 1. Suppose that for each α satisfying $m_A(\alpha)=0$, one has $\frac{d}{dx}m_A(\alpha)\neq 0$. Then, it follows that each zero of $m_A(x)$ has multiplicity 1. Also, using Corollary 7.2.6, each zero of $m_A(x)$ is an eigenvalue of A and hence by Theorem 7.2.8, the size of each Jordan block is 1. Thus, A is diagonalizable. ∎

We now have the following remarks and observations.

Remark 7.2.11. 1. *Let $f(x)$ be a monic polynomial and $A=\text{COMPANION}(f)$ be the companion matrix of f. Then, by Theorem 7.2.4) $f(A)=\mathbf{0}$ and no monic polynomial of smaller degree annihilates A. Thus $P_A(x)=m_A(x)=f(x)$, where $P_A(x)$ is the characteristic polynomial and $m_A(x)$, the minimal polynomial of A.*

2. *Let $A\in\mathbb{M}_n(\mathbb{C})$. Then, A is similar to* COMPANION(f)*, for some monic polynomial f if and only if $m_A(x)=f(x)$.*

   Proof. Let $B=$ COMPANION (f)*. Then, using Lemma 7.2.7, we see that $m_A(x)=m_B(x)$. But, by Remark 7.2.11.1, we get $m_B(x)=f(x)$ and hence the required result follows.*

   Conversely, assume that $m_A(x)=f(x)$. But, by Remark 7.2.11.1, $m_B(x)=f(x)=P_B(x)$, the characteristic polynomial of B. Since $m_A(x)=m_B(x)$, the matrices A and B have the same largest Jordan blocks for each eigenvalue λ. As $P_B=m_B$, we know that for each λ, there is only one Jordan block in JORDAN CFB*. Thus,* JORDAN CFA = JORDAN CFB *and hence A is similar to* COMPANION (f). ∎

EXERCISE **7.2.12.** *The following are some facts and questions.*

1. *Let $A\in\mathbb{M}_n(\mathbb{C})$. If $P_A(x)$ is the minimal polynomial of A then A is similar to* COMPANION (P_A) *if and only if A is nonderogatory. T/F?*
2. *Let $A,B\in\mathbb{M}_3(\mathbb{C})$ with eigenvalues $1,2,3$. Is it necessary that A is similar to B?*
3. *Let $A,B\in\mathbb{M}_3(\mathbb{C})$ with eigenvalues $1,1,3$. Is it necessary that A is similar to B?*
4. *Let $A,B\in\mathbb{M}_4(\mathbb{C})$ with the same minimal polynomial. Is it necessary that A is similar to B?*
5. *Let $A,B\in\mathbb{M}_3(\mathbb{C})$ with the same minimal polynomial. Is it necessary that A is similar to B?*
6. *Let $A\in\mathbb{M}_n(\mathbb{C})$ be idempotent and let $J=$* JORDAN CFA*. Thus, $J^2=J$ and hence conclude that J must be a diagonal matrix. Hence, every idempotent matrix is diagonalizable.*

7. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Suppose that $m_A(x) | x(x-1)(x-2)(x-3)$. Must A be diagonalizable?*
8. *Let $A \in \mathbb{M}_9(\mathbb{C})$ be a nilpotent matrix such that $A^5 \neq \mathbf{0}$ but $A^6 = \mathbf{0}$. Determine $P_A(x)$ and $m_A(x)$.*
9. *Recall that for $A, B \in \mathbb{M}_n(\mathbb{C})$, the characteristic polynomial of AB and BA are the same. That is, $P_{AB}(x) = P_{BA}(x)$. However, they need not have the same minimal polynomial. Take $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ to verify that $m_{AB}(x) \neq m_{BA}(x)$.*

We end this section with a method to compute the minimal polynomial of a given matrix.

Example 7.2.13. **[Computing the Minimal Polynomial]** Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of $A \in \mathbb{M}_n(\mathbb{C})$.

7.3 Applications of Jordan Canonical Form

In the last section, we say that the matrices if A is a square matrix then A and A^T are similar. In this section, we look at some more applications of the Jordan Canonical Form.

7.3.1 Coupled system of linear differential equations

Consider the first order Initial Value Problem (IVP) $\mathbf{x}'(t) = \begin{bmatrix} \mathbf{x}_1'(t) \\ \vdots \\ \mathbf{x}_n'(t) \end{bmatrix} = A \begin{bmatrix} \mathbf{x}_1(t) \\ \vdots \\ \mathbf{x}_n(t) \end{bmatrix} = A\mathbf{x}(t)$, with $\mathbf{x}(0) = \mathbf{0}$. If A is not a diagonal matrix then the system is called COUPLED and is hard to solve. Note that if A can be transformed to a nearly diagonal matrix, then the amount of coupling among $\mathbf{x}_i$'s can be reduced. So, let us look at $J = \text{Jordan CF}(A) = S^{-1}AS$. Then, using $S^{-1}A = JS^{-1}$. verify that the initial problem $\mathbf{x}'(t) = A\mathbf{x}(t)$ is equivalent to the equation $S^{-1}\mathbf{x}'(t) = S^{-1}A\mathbf{x}(t)$ which in turn is equivalent to $\mathbf{y}'(t) = J\mathbf{y}(t)$, where $S^{-1}\mathbf{x}(t) = \mathbf{y}(t)$ with $\mathbf{y}(0) = S^{-1}\mathbf{x}(0) = \mathbf{0}$. Therefore, if $\mathbf{y}$ is a solution to the second equation then $\mathbf{x}(t) = S\mathbf{y}$ is a solution to the initial problem.

When J is diagonalizable then solving the second is as easy as solving $\mathbf{y}_i'(t) = \lambda_i \mathbf{y}_i(t)$ for which the required solution is given by $\mathbf{y}_i(t) = \mathbf{y}_i(0)e^{\lambda_i t}$.

If J is not diagonal, then for each Jordan block, the system reduces to

$$\mathbf{y}_1'(t) = \lambda\mathbf{y}_1(t) + \mathbf{y}_2(t), \cdots, \mathbf{y}_{k-1}'(t) = \lambda\mathbf{y}_{k-1}(t) + \mathbf{y}_k(t), \;\; \mathbf{y}_k'(t) = \lambda\mathbf{y}_k(t).$$

This problem can also be solved as in this case the solution is given by $\mathbf{y}_k = c_0 e^{\lambda t}$; $\mathbf{y}_{k-1} = (c_0 t + c_1)e^{\lambda t}$ and so on.

7.3.2 Commuting matrices

Let $P(x)$ be a polynomial and $A \in \mathbb{M}_n(\mathbb{C})$. Then, $P(A)A = AP(A)$. What about the converse? That is, suppose we are given that $AB = BA$ for some $B \in \mathbb{M}_n(\mathbb{C})$. Does it necessarily imply that $B = P(A)$, for some nonzero polynomial $P(x)$? The answer is **No** as I commutes with A for every A. We start with a set of remarks.

Theorem 7.3.1. *Let $A \in \mathbb{M}_n(\mathbb{C})$ and $B \in \mathbb{M}_m(\mathbb{C})$. Then, the linear system $AX - XB = \mathbf{0}$, in the variable matrix X of size $n \times m$, has a unique solution, namely $X = \mathbf{0}$ (the trivial solution), if and only if $\sigma(A)$ and $\sigma(B)$ are disjoint.*

Proof. Let us assume that $\sigma(A)$ and $\sigma(B)$ are disjoint.

Since $\sigma(A)$ and $\sigma(B)$ are disjoint, the matrix $P_B(A) = \left(\prod\limits_{\lambda \in \sigma(B)} [\lambda I - A]\right)$, obtained by evaluating A at the characteristic polynomial, $P_B(t)$, of B, is invertible. So, let us look at the implication of the condition $AX = XB$. This condition implies that $A^2X = AXB = XBB = XB^2$ and hence, $P(A)X = XP(B)$, for any polynomial $P(t)$. In particular, $P_B(A)X = XP_B(B) = X\mathbf{0} = \mathbf{0}$. As $P_B(A)$ is invertible, we get $X = \mathbf{0}$.

Now, conversely assume that $AX - XB = \mathbf{0}$ has only the trivial solution $X = \mathbf{0}$. Suppose on the contrary λ is a common eigenvalue of both A and B. So, choose nonzero vectors $\mathbf{x} \in \mathbb{C}^n$ and $\mathbf{y} \in \mathbb{C}^m$ such that $(\lambda, \mathbf{x})$ is an eigen-pair of A and $(\lambda, \mathbf{y})$ is a left eigen-pair of B. Now, define $X_0 = \mathbf{x}\mathbf{y}^T$. Then, X_0 is an $n \times m$ nonzero matrix and

$$AX - XB = A\mathbf{x}\mathbf{y}^T - \mathbf{x}\mathbf{y}^T B = \lambda\mathbf{x}\mathbf{y}^T - \lambda\mathbf{x}\mathbf{y}^T = \mathbf{0}.$$

Thus, we see that if λ is a common eigenvalue of A and B then the system $AX - XB = \mathbf{0}$ has a nonzero solution X_0, a contradiction. Hence, the required result follows. ∎

Corollary 7.3.2. *Let $A \in \mathbb{M}_n(\mathbb{C}), B \in \mathbb{M}_m(\mathbb{C})$ and C be an $n \times m$ matrix. Also, assume that $\sigma(A)$ and $\sigma(B)$ are disjoint. Then, it can be easily verified that the system $AX - XB = C$, in the variable matrix X of size $n \times m$, has a unique solution, for any given C.*

Proof. Consider the linear transformation $T : \mathbb{M}_{n,m}(\mathbb{C}) \to \mathbb{M}_{n,m}(\mathbb{C})$, defined by $T(X) = AX - XB$. Then, by Theorem 7.3.1, $\text{NULL}(T) = \{\mathbf{0}\}$. Hence, by the rank-nullity theorem, T is a bijection and the required result follows. ∎

Definition 7.3.3. [**Toeplitz Matrix**] A square matrix A is said to be of **Toeplitz type** if each (super/sub)-diagonal of A consists of the same element. For example, $A = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ a_1 & b_1 & b_2 & b_3 \\ a_2 & a_1 & b_1 & b_2 \\ a_3 & a_2 & a_1 & b_1 \end{bmatrix}$ is a 4×4 Toeplitz type matrix. and the matrix $B = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ 0 & b_1 & b_2 & b_3 \\ 0 & 0 & b_1 & b_2 \\ 0 & 0 & 0 & b_1 \end{bmatrix}$ is an upper triangular Toeplitz type matrix.

EXERCISE **7.3.4.** *Let $J_n(0) \in \mathbb{M}_n(\mathbb{C})$ be the Jordan block with 0 on the diagonal.*

1. *Further, if $A \in \mathbb{M}_n(\mathbb{C})$ such that $AJ_n(0) = J_n(0)A$ then prove that A is an upper Toeplitz type matrix.*
2. *Further, if $A, B \in \mathbb{M}_n(\mathbb{C})$ are two upper Toeplitz type matrices then prove that*
   (a) *there exists $a_i \in \mathbb{C}, 1 \le i \le n$, such that $A = a_0 I + a_1 J_n(0) + \cdots + a_n J_n(0)^{n-1}$.*
   (b) *$P(A)$ is a Toeplitz matrix for any polynomial $P(t)$.*

(c) AB is a Toeplitz matrix.

(d) if A is invertible then A^{-1} is also an upper Toeplitz type matrix.

To proceed further, recall that a matrix $A \in \mathbb{M}_n(\mathbb{C})$ is called **non-derogatory** if $\text{Geo.Mul}_\alpha(A) = 1$, for each $\alpha \in \sigma(A)$ (see Definition 6.2.4).

Theorem 7.3.5. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a non-derogatory matrix. Then, the matrices A and B commute if and only if $B = P(A)$, for some polynomial $P(t)$ of degree at most $n-1$.*

Proof. If $B = P(A)$, for some polynomial $P(t)$, then A and B commute. Conversely, suppose that $AB = BA$, $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$ and let $J = \text{Jordan CF}A = S^{-1}AS$ be the Jordan matrix of A. Then, $J = \begin{bmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_k}(\lambda_k) \end{bmatrix}$. Now, write $\overline{B} = S^{-1}BS = \begin{bmatrix} \overline{B}_{11} & \cdots & \overline{B}_{1k} \\ \vdots & \ddots & \vdots \\ \overline{B}_{k1} & \cdots & \overline{B}_{kk} \end{bmatrix}$, where $\overline{B}$ is partitioned conformally with J. Note that $AB = BA$ gives $J\overline{B} = \overline{B}J$. Thus, verify that

$$J_{n_1}(\lambda_1)\overline{B}_{12} = [JB]_{12} = [BJ]_{12} = \overline{B}_{12}J_{n_2}(\lambda_2),$$

and hence $\overline{B}_{12} = \mathbf{0}$. A similar argument gives $\overline{B}_{ij} = 0$, for all $i \neq j$. Hence, $JB = BJ$ implies $J_{n_i}(\lambda_i)\overline{B}_{ii} = \overline{B}_{ii}J_{n_i}(\lambda_i)$, for $1 \le i \le k$. Or equivalently, $J_{n_i}(0)B_{ii} = B_{ii}J_{n_i}(0)$, for $1 \le i \le k$ (using Exercise 7.1.6.3). Now, using Exercise 7.3.4.1, we see that $\overline{B}_{ii}$ is an upper triangular Toeplitz type matrix.

To proceed further, for $1 \le i \le k$, define $F_i(t) = \prod\limits_{j \neq i}(t-\lambda_j)^{n_j}$. Then, $F_i(t)$ is a polynomial with $\deg(F_i(t)) = n - n_i$ and $F_i(J_{n_j}(\lambda_j)) = \mathbf{0}$ if $j \neq i$. Also, note that $F_i(J_{n_i}(\lambda_i))$ is a nonsingular upper triangular Toeplitz type matrix. Hence, its inverse has the same form and using Exercise 7.3.4.1, the matrix $F_i(J_{n_i}(\lambda_i))^{-1}\overline{B}_{ii}$ is also a Toeplitz type upper triangular matrix. Hence,

$$F_i(J_{n_i}(\lambda_i))^{-1}\overline{B}_{ii} = c_1 I + c_2 J_{n_i}(0) + \cdots + c_{n_i}J_{n_i}(0)^{n_i-1} = R_i(J_{n_i}(\lambda_i)), \text{ (say)}.$$

Thus, $\overline{B}_{ii} = F_i(J_{n_i}(\lambda_i))R_i(J_{n_i}(\lambda_i))$. Putting $P_i(t) = F_i(t)R_i(t)$, for $1 \le i \le k$, we see that $P_i(t)$ is a polynomial of degree at most $n-1$ with $P_i((J_{n_j}(\lambda_j)) = \mathbf{0}$, for $j \neq i$ and $P_i((J_{n_j}(\lambda_i)) = \overline{B}_{ii}$. Taking, $P = P_1 + \cdots + P_k$, we have

$$\begin{aligned} P(J) &= P_1\left(\begin{bmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_k}(\lambda_k) \end{bmatrix}\right) + \cdots + P_k\left(\begin{bmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_k}(\lambda_k) \end{bmatrix}\right) \\ &= \begin{bmatrix} \overline{B}_{11} & & \\ & \ddots & \\ & & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & & \\ & \ddots & \\ & & \overline{B}_{kk} \end{bmatrix} = \overline{B}. \end{aligned}$$

Hence, $B = S\overline{B}S^{-1} = SP(J)S^{-1} = P(SJS^{-1}) = P(A)$ and the required result follows. ∎

Chapter 8

Advanced Topics on Diagonalizability and Triangularization*

8.1 More on the Spectrum of a Matrix

We start this subsection with a few definitions and examples. So, it will be nice to recall the notations used in Section 1.3.1 and a few results from Appendix 9.2.

Definition 8.1.1. [Principal Minor] Let $A \in \mathbb{M}_n(\mathbb{C})$.

1. Also, let $S \subseteq [n]$. Then, $\det(A[S,S])$ is called the **Principal minor** of A corresponding to S.
2. By $\mathbf{EM}_k(A)$, we denote the sum of all $k \times k$ principal minors of A.

Definition 8.1.2. [Elementary Symmetric Functions] Let k be a positive integer. Then, the kth **elementary symmetric function** of the numbers $r_1, \ldots, r_n$ is $S_k(r_1, \ldots, r_n)$ and is defined as

$$S_k(r_1, \ldots, r_n) = \sum_{i_1 < \cdots < i_k} r_{i_1} \cdots r_{i_k}.$$

Example 8.1.3. Let $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 8 & 7 & 6 \\ 5 & 4 & 3 & 2 \end{bmatrix}$. Then, note that

1. $\mathrm{EM}_1(A) = 1 + 6 + 7 + 2 = 16$ and $\mathrm{EM}_2(A) = \det A(\{1,2\},\{1,2\}) + \det A(\{1,3\},\{1,3\}) + \det A(\{1,4\},\{1,4\}) + \det A(\{2,3\},\{2,3\}) + \det A(\{2,4\},\{2,4\}) + \det A(\{3,4\},\{3,4\}) = -80$.
2. $S_1(1,2,3,4) = 10$ and $S_2(1,2,3,4) = 1 \cdot (2+3+4) + 2 \cdot (3+4) + 3 \cdot 4 = 9 + 14 + 12 = 35$.

Theorem 8.1.4. *Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$. Then,*

1. *the coefficient of t^{n-k} in $P_A(t) = \prod\limits_{i=1}^{n}(t - \lambda_i)$, the characteristic polynomial of A, is*

$$(-1)^k \sum_{i_1 < \cdots < i_k} \lambda_{i_1} \cdots \lambda_{i_k} = (-1)^k S_k(\lambda_1, \ldots, \lambda_n). \tag{8.1.1}$$

2. $EM_k(A) = S_k(\lambda_1, \ldots, \lambda_n)$.

Proof. Note that by definition,

$$\begin{aligned} P_A(t) &= \prod_{i=1}^{n}(t-\lambda_i) = t^n - S_1(\lambda_1,\ldots,\lambda_n)t^{n-1} \\ &\qquad + S_2(\lambda_1,\ldots,\lambda_n)t^{n-2} - \cdots + (-1)^n S_n(\lambda_1,\ldots,\lambda_n) \quad (8.1.2) \\ &= t^n - EM_1(A)t^{n-1} + EM_2(A)t^{n-2} - \cdots + (-1)^n EM_n(A). \quad (8.1.3) \end{aligned}$$

As the second part is just a re-writing of the first, we will just prove the first part. To do so, let $B = tI - A = \begin{bmatrix} t-a_{11} & \cdots & -a_{1n} \\ & \ddots & \\ -a_{n1} & \cdots & t-a_{nn} \end{bmatrix}$. Then, using Definition 9.2.2 in Appendix, note that $\det B = \sum\limits_{\sigma} \text{sgn}\sigma \prod\limits_{i=1}^{n} b_{i\sigma(i)}$ and hence each $S \subseteq [n]$ with $|S| = n-k$ has a contribution to the coefficient of t^{n-k} in the following way:

For all $i \in S$, consider all permutations σ such that $\sigma(i) = i$. Our idea is to select a 't' from these $b_{i\sigma(i)}$. Since we do not want any more 't', we set $t = 0$ for any other diagonal position. So the contribution from S to the coefficient of t^{n-k} is $\det[-A(S|S)] = (-1)^k \det A(S|S)$. Hence the coefficient of t^{n-k} in $P_A(t)$ is

$$(-1)^k \sum_{S\subseteq[n],\,|S|=n-k} \det A(S|S) = (-1)^k \sum_{T\subseteq[n],\,|T|=k} \det A[T,T] = (-1)^k E_k(A).$$

The proof is complete in view of Equation (8.1.2). ∎

As a direct application, we obtain Theorem 6.1.16 which we state again.

Corollary 8.1.5. *Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$. Then $\text{Tr}(A) = \sum_1^n \lambda_i$ and $\det A = \prod_1^n \lambda_i$.*

Let A and B be similar matrices. Then, by Theorem 6.1.20, we know that $\sigma(A) = \sigma(B)$. Thus, as a direct consequence of Part 2 of Theorem 8.1.4 gives the following result.

Corollary 8.1.6. *Let A and B be two similar matrices of order n. Then, $EM_k(A) = EM_k(B)$ for $1 \le k \le n$.*

So, the sum of principal minors of similar matrices are equal. Or in other words, the sum of principal minors are invariant under similarity.

Corollary 8.1.7. [Derivative of Characteristic Polynomial] *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then*

$$\frac{d}{dt}P_A(t) = P_A'(t) = \sum_{i=1}^{n} P_{A(i|i)}(t).$$

Proof. For $1 \le i \le n$, let us denote $A(i|i)$ by A_i. Then, using Equation (8.1.3), we have

$$\begin{aligned} \sum_{i=1}^{n} P_{A_i}(t) &= \sum_i t^{n-1} - \sum_i EM_1(A_i)t^{n-2} + \cdots + (-1)^{n-1}\sum_i EM_{n-1}(A_i) \\ &= nt^{n-1} - (n-1)EM_1(A)t^{n-2} + (n-2)EM_2(A)t^{n-3} - \cdots + (-1)^{n-1}EM_{n-1}(A) \\ &= P_A'(t). \end{aligned}$$

Which gives us the desired result. ∎

Corollary 8.1.8. *Let $A \in \mathbb{M}_n(\mathbb{C})$. If* $\text{ALG.MUL}_\alpha(A) = 1$ *then* $\text{Rank}[A - \lambda I] = n - 1$.

Proof. As $\text{ALG.MUL}_\alpha(A) = 1$, $P_A(t) = (t-\lambda)q(t)$, where $q(t)$ is a polynomial with $q(\lambda) \neq 0$. Thus $P'_A(t) = q(t) + (t-\lambda)q'(t)$. Hence, $P'_A(\lambda) = q(\lambda) \neq 0$. Thus, by Corollary 8.1.7, $\sum_i P_{A(i|i)}(\lambda) = P'_A(\lambda) \neq 0$. Hence, there exists $i, 1 \le i \le n$ such that $P_{A(i|i)}(\lambda) \neq 0$. That is, $\det[A(i|i) - \lambda I] \neq 0$ or $\text{Rank}[A - \lambda I] = n - 1$. ■

Remark 8.1.9. *Converse of Corollary 8.1.8 is false. Note that for the matrix* $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\text{Rank}[A - 0I] = 1 = 2 - 1 = n - 1$, *but* 0 *has multiplicity* 2 *as a root of* $P_A(t) = 0$.

As an application of Corollary 8.1.7, we have the following result.

We now relate the multiplicity of an eigenvalue with the spectrum of a principal sub-matrix.

Theorem 8.1.10. **[Multiplicity and Spectrum of a Principal Sub-Matrix]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ and k be a positive integer. Then 1 ⇒ 2 ⇒ 3, where*

1. $\text{GEO.MUL}_\lambda(A) \ge k$.
2. *If B is a principal sub-matrix of A of size $m > n - k$ then $\lambda \in \sigma(B)$.*
3. $\text{ALG.MUL}_\lambda(A) \ge k$.

Proof. Part 1⇒ Part 2. Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ be linearly independent eigenvectors for λ and let B be a principal sub-matrix of A of size $m > n - k$. Without loss, we may write $A = \begin{bmatrix} B & * \\ * & * \end{bmatrix}$. Let us partition the $\mathbf{x}_i$'s , say $\mathbf{x}_i = \begin{bmatrix} \mathbf{x}_{i1} \\ \mathbf{x}_{i2} \end{bmatrix}$, such that

$$\left[\begin{array}{c|c} B & * \\ * & * \end{array}\right] \begin{bmatrix} \mathbf{x}_{i1} \\ \mathbf{x}_{i2} \end{bmatrix} = \lambda \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix}, \text{ for } 1 \le i \le k.$$

As $m > n - k$, the size of $\mathbf{x}_{i2}$ is less than k. Thus, the set $\{\mathbf{x}_{12}, \ldots, \mathbf{x}_{k2}\}$ is linearly dependent (see Corollary 3.3.6). So, there is a nonzero linear combination $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$ of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ such that $\mathbf{y}_2 = \mathbf{0}$. Notice that $\mathbf{y}_1 \neq \mathbf{0}$ and $B\mathbf{y}_1 = \lambda \mathbf{y}_1$.

Part 2⇒ Part 3. By Corollary 8.1.7, we know that $P'_A(t) = \sum\limits_{i=1}^{n} P_{A(i|i)}(t)$. As $A(i|i)$ is of size $n-1$, we get $P_{A(i|i)}(\lambda) = 0$, for all $i = 1, 2, \ldots, n$. Thus, $P'_A(\lambda) = 0$. A similar argument now applied to each of the $A(i|i)$'s, gives $P_A^{(2)}(\lambda) = 0$, where $P_A^{(2)}(t) = \dfrac{d}{dt} P'_A(t)$. Proceeding on above lines, we finally get $P_A^{(i)}(\lambda) = 0$, for $i = 0, 1, \ldots, k-1$. This implies that $\text{ALG.MUL}_\lambda(A) \ge k$. ■

Definition 8.1.11. **[Moments]** Fix a positive integer n and let $\alpha_1, \ldots, \alpha_n$ be n complex numbers. Then, for a positive integer k, the sum $\sum\limits_{i=1}^{n} \alpha_i^k$ is called the k**-th moment** of the numbers $\alpha_1, \ldots, \alpha_n$.

Theorem 8.1.12. **[Newton's identities]** *Let $P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0$ have zeros $\lambda_1, \ldots, \lambda_n$, counted with multiplicities. Put $\mu_k = \sum\limits_{i=1}^{n} \lambda_i^k$. Then, for $1 \le k \le n$,*

$$k\, a_{n-k} + \mu_1 a_{n-k+1} + \cdots + \mu_{k-1} a_{n-1} + \mu_k = 0. \tag{8.1.4}$$

That is, the first n moments of the zeros determine the coefficients of $P(t)$.

Proof. For simplicity of expression, let $a_n = 1$. Then, using Equation (8.1.4), we see that $k = 1$ gives us $a_{n-1} = -\mu_1$. To compute a_{n-2}, put $k = 2$ in Equation (8.1.4) to verify that $a_{n-2} = \frac{-\mu_2+\mu_1^2}{2}$. This process can be continued to get all the coefficients of $P(t)$. Now, let us prove the n given equations.

Define $f(t) = \sum\limits_i \frac{1}{t-\lambda_i} = \frac{P'(t)}{P(t)}$ and take $|t| > \max\limits_i |\lambda_i|$. Then, the left hand side can be re-written as

$$f(t) = \sum_{i=1}^{n} \frac{1}{t-\lambda_i} = \sum_{i=1}^{n} \frac{1}{t\left(1-\dfrac{\lambda_i}{t}\right)} = \sum_{i=1}^{n}\left[\frac{1}{t} + \frac{\lambda_i}{t^2} + \cdots\right] = \frac{n}{t} + \frac{\mu_1}{t^2} + \cdots. \tag{8.1.5}$$

Thus, using $P'(t) = f(t)P(t)$, we get

$$na_nt^{n-1} + (n-1)a_{n-1}t^{n-2} + \cdots + a_1 = P'(t) = \left[\frac{n}{t} + \frac{\mu_1}{t^2} + \cdots\right]\left[a_nt^n + \cdots + a_0\right].$$

Now, equating the coefficient of t^{n-k-1} on both sides, we get

$$(n-k)a_{n-k} = na_{n-k} + \mu_1 a_{n-k+1} + \cdots + \mu_k a_n, \text{ for } 0 \le k \le n-1$$

which is the required Newton's identity. ∎

Remark 8.1.13. *Let $P(t) = a_nt^n + \cdots + a_1t + a_0$ with $a_n = 1$. Thus, we see that we need not find the zeros of $P(t)$ to find the k-th moments of the zeros of $P(t)$. It can directly be computed recursively using the Newton's identities.*

EXERCISE **8.1.14.** *Let $A, B \in \mathbb{M}_n(\mathbb{C})$. Then, prove that A and B have the same eigenvalues if and only if $\mathsf{tr}(\mathsf{A}^\mathsf{k}) = \mathsf{tr}(\mathsf{B}^\mathsf{k})$, for $k = 1, \ldots, n$. (Use Exercise 6.1.8 1a).*

8.2 Methods for Tridiagonalization and Diagonalization

Let $\mathcal{G} = \{A \in \mathbb{M}_n(\mathbb{C}) : A^*A = I\}$. Then, using Exercise 5.4.8, we see that

1. for every $A, B \in \mathcal{G}$, $AB \in \mathcal{G}$ (see Exercise 5.4.8.10).
2. for every $A, B, C \in \mathcal{G}$, $(AB)C = A(BC)$.
3. I_n is the identity element of $\mathcal{G}$. That is, for any $A \in \mathcal{G}, AI_n = A = I_nA$.
4. for every $A \in \mathcal{G}, A^{-1} \in \mathcal{G}$.

Thus, the set $\mathcal{G}$ forma a group with respect to multiplication. We now define this group.

Definition 8.2.1. **[Unitary Group]** Let $\mathcal{G} = \{A \in \mathbb{M}_n(\mathbb{C}) : A^*A = I\}$. Then, $\mathcal{G}$ forms a multiplicative group. This group is called the **unitary group**.

Proposition 8.2.2. **[Selection Principle of Unitary Matrices]** *Let $\{U_k : k \ge 1\}$ be a sequence of unitary matrices. Viewing them as elements of $\mathbb{C}^{n^2}$, let us assume that "for any $\epsilon > 0$, there exists a positive integer N such that $\|U_k - U\| < \epsilon$, for all $k \ge N$". That is, the matrices U_k's converge to U as elements in $\mathbb{C}^{n^2}$. Then, U is also a unitary matrix.*

Proof. Let $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{C})$ be an unitary matrix. Then $\sum\limits_{i,j=1}^{n} |a_{ij}|^2 = \mathsf{tr(A^*A) = n}$. Thus, the set of unitary matrices is a compact subset of $\mathbb{C}^{n^2}$. Hence, any sequence of unitary matrices has a convergent subsequence (Bolzano-Weierstrass Theorem), whose limit is again unitary. Thus, the required result follows. ∎

For a unitary matrix U, we know that $U^{-1} = U^*$. Our next result gives a necessary and sufficient condition on an invertible matrix A so that the matrix A^{-1} is similar to A^*.

Theorem 8.2.3. **[Generalizing a Unitary Matrix]** *Let A be an invertible matrix. Then A^{-1} is similar to A^* if and only if there exists an invertible matrix B such that $A = B^{-1}B^*$.*

Proof. Suppose $A = B^{-1}B^*$, for some invertible matrix B. Then

$$A^* = B(B^{-1})^* = B(B^{-1})^*BB^{-1} = B(B^{-1}B^*)^{-1}B^{-1} = BA^{-1}B^{-1}.$$

Conversely, let $A^* = SA^{-1}S^{-1}$, for some invertible matrix S. Need to show, $A = S^{-1}S^*$.

We first show that there exists a nonsingular Hermitian H_θ such that $A^{-1} = H_\theta^{-1}A^*H_\theta$, for some $\theta \in \mathbb{R}$.

Note that for any $\theta \in \mathbb{R}$, if we put $S_\theta = e^{i\theta}S$ then

$$S_\theta A^{-1}S_\theta^{-1} = A^* \text{ and } S_\theta = A^*S_\theta A.$$

Now, define $H_\theta = S_\theta + S_\theta^*$. Then, H_θ is a Hermitian matrix and $H_\theta = A^*H_\theta A$. Furthermore, there are infinitely many choices of $\theta \in \mathbb{R}$ such that $\det H_\theta = 0$. To see this, let us choose a $\theta \in \mathbb{R}$ such that H_θ is singular. Hence, there exists $\mathbf{x} \neq \mathbf{0}$ such that $H_\theta\mathbf{x} = \mathbf{0}$. So,

$$S_\theta\mathbf{x} = -S_\theta^*\mathbf{x} = -e^{-i\theta}S^*\mathbf{x}. \text{ Or equivalently, } -e^{2i\theta}\mathbf{x} = S^{-1}S^*\mathbf{x}.$$

That is, $-e^{2i\theta} \in \sigma(S^{-1}S^*)$. Thus, if we choose $\theta_0 \in \mathbb{R}$ such that $-e^{2i(\theta_0)} \notin \sigma(S^{-1}S^*)$ then $H_(\theta_0)$ is nonsingular.

To get our result, we finally choose $B = \beta(\alpha I - A^*)H_(\theta_0)$ such that $\beta \neq 0$ and $\alpha = e^{i\gamma} \notin \sigma(A^*)$.

Note that with α and β chosen as above, B is invertible. Furthermore,

$$BA = \alpha\beta H_(\theta_0)A - \beta A^*H_(\theta_0)A = \alpha\beta H_(\theta_0)A - \beta H_(\theta_0) = \beta H_(\theta_0)(\alpha A - I).$$

As we need, $BA = B^*$, we get $\beta H_(\theta_0)(\alpha A - I) = \overline{\beta}H_(\theta_0)(\overline{\alpha}I - A)$ and thus, we need $\overline{\beta} = -\beta\alpha$, which holds true if $\beta = e^{i(\pi-\gamma)/2}$. Thus, the required result follows. ∎

Exercise **8.2.4.** *Suppose that A is similar to a unitary matrix. Then, prove that A^{-1} is similar to A^*.*

8.2.1 Plane Rotations

Definition 8.2.5. **[Plane Rotations]** For a fixed positive integer n, consider the vector space $\mathbb{R}^n$ with standard basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. Also, for $1 \leq i, j \leq n$, let $E_{i,j} = \mathbf{e}_i\mathbf{e}_j^T$. Then, for $\theta \in \mathbb{R}$ and $1 \leq i, j \leq n$, a **plane rotation**, denoted $U(\theta; i, j)$, is defined as

$$U(\theta; i, j) = I - E_{i,i} - E_{j,j} + [E_{i,i} + E_{j,j}]\cos\theta - E_{i,j}\sin\theta + E_{j,i}\sin\theta.$$

That is, $U(\theta; i, j) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos\theta & & -\sin\theta & & \\ & & & \ddots & & & \\ & & \sin\theta & & \cos\theta & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$, where the unmentioned diagonal entries are 1 and the unmentioned off-diagonal entries are 0.

Remark 8.2.6. *Note the following about the matrix $U(\theta; i, j)$, where $\theta \in \mathbb{R}$ and $1 \le i, j \le n$.*

1. *$U(\theta; i, j)$ are orthogonal.*
2. *Geometrically $U(\theta; i, j)\mathbf{x}$ rotates $\mathbf{x}$ by the angle θ in the ij-plane.*
3. *Geometrically $(U(\theta; i, j))^T \mathbf{x}$ rotates $\mathbf{x}$ by the angle $-\theta$ in the ij-plane.*
4. *If $\mathbf{y} = U(\theta; i, j)\mathbf{x}$ then the coordinates of $\mathbf{y}$ are given by*
   (a) $\mathbf{y}_i = \mathbf{x}_i \cos\theta - \mathbf{x}_j \sin\theta$,
   (b) $\mathbf{y}_j = \mathbf{x}_i \sin\theta + \mathbf{x}_j \cos\theta$, *and*
   (c) *for $l \neq i, j$, $\mathbf{y}_l = \mathbf{x}_l$.*
5. *Thus, for $\mathbf{x} \in \mathbb{R}^n$, the choice of θ for which $\mathbf{y}_j = 0$, where $\mathbf{y} = U(\theta; i, j)\mathbf{x}$ equals*
   (a) $\theta = 0$, *whenever $\mathbf{x}_j = 0$. That is, $U(0; i, j) = I$.*
   (b) $\theta = \cot^{-1}\left(-\frac{\mathbf{x}_i}{\mathbf{x}_j}\right)$, *whenever $\mathbf{x}_j \neq 0$.*

6. [**Geometry**] *Imagine standing at $\mathbf{1} = (1, 1, 1)^T \in \mathbb{R}^3$. We want to apply a plane rotation U, so that $\mathbf{v} = U^T \mathbf{1}$ with $\mathbf{v}_2 = 0$. That is, the final point is on the xz-plane.*

   Then, we can either apply a plane rotation along the xy-plane or the yz-plane. For the xy-plane, we need the plane $z = 1$ (xy plane lifted by 1). This plane contains the vector $\mathbf{1}$. Imagine moving the tip of $\vec{\mathbf{1}}$ on this plane. Then this locus corresponds to a circle that lies on the plane $z = 1$, has radius $\sqrt{2}$ and is centred at $(0, 0, 1)$. That is, we draw the circle $x^2 + y^2 = 1$ on the xy-plane and then lifted it up by so that it lies on the plane $z = 1$. Thus, note that the xz-plane cuts this circle at two points. These two points of intersections give us the two choices for the vector $\mathbf{v}$ (see Figure 8.1). A similar calculation can be done for the yz-plane.
7. *In general, in $\mathbb{R}^n$, suppose that we want to apply plane rotation to $\mathbf{a}$ along the x_1x_2-plane so that the resulting vector has 0 in the 2-nd coordinate. In that case, our circle on x_1x_2-plane has radius $r = \sqrt{\mathbf{a}_1^2 + \mathbf{a}_2^2}$ and it gets translated by $\begin{bmatrix} 0, & 0, & \mathbf{a}_3, & \cdots & \mathbf{a}_n \end{bmatrix}^T$. So, there are two points $\mathbf{x}$ on this circle with $\mathbf{x}_2 = 0$ and they are $\begin{bmatrix} \pm r, & 0, & a_3, & \cdots & a_n \end{bmatrix}^T$.*
8. *Consider three mutually orthogonal unit vectors, say $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Then, $\mathbf{x}$ can be brought to $\mathbf{e}_1$ by two plane rotations, namely by an appropriate $U(\theta_1; 1, 3)$ and $U(\theta_2; 1, 2)$. Thus,*

$$U(\theta_2; 1, 2)U(\theta_1; 1, 3)\mathbf{x} = \mathbf{e}_1.$$

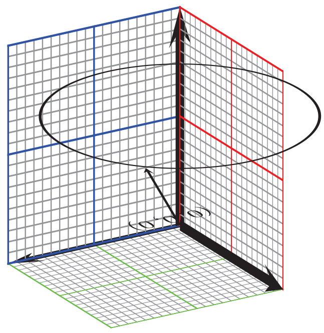


Figure 8.1: Geometry of plane rotations in $\mathbb{R}^3$

.

In this process, the unit vectors $\mathbf{y}$ *and* $\mathbf{z}$*, get shifted to say,*

$$\hat{\mathbf{y}} = U(\theta_2; 1, 2)U(\theta_1; 1, 3)\mathbf{y} \text{ and } \hat{\mathbf{z}} = U(\theta_2; 1, 2)U(\theta_1; 1, 3)\mathbf{z}.$$

As unitary transformations preserve angles, note that $\hat{\mathbf{y}}(1) = \hat{\mathbf{z}}(1) = 0$*. Now, we can apply an appropriate plane rotation* $U(\theta_3; 2, 3)$ *so that* $U(\theta_3; 2, 3)\hat{\mathbf{y}} = \mathbf{e}_2$*. Since* $\mathbf{e}_3$ *is the only unit vector in* $\mathbb{R}^3$ *orthogonal to both* $\mathbf{e}_1$ *and* $\mathbf{e}_2$*, it follows that* $U(\theta_3; 2, 3)\hat{\mathbf{z}} = \mathbf{e}_3$*. Thus,*

$$I = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix} = U(\theta_3; 2, 3)U(\theta_2; 1, 2)U(\theta_1; 1, 3)\begin{bmatrix} \mathbf{x} & \mathbf{y} & \mathbf{z} \end{bmatrix}.$$

Hence, any real orthogonal matrix $A \in \mathbb{M}_3(\mathbb{R})$ *is a product of three plane rotations.*

We are now ready to give another method to get the QR-decomposition of a square matrix (see Theorem 5.2.1 that uses the Gram-Schmidt Orthonormalization Process).

Proposition 8.2.7. **[QR Factorization Revisited: Square Matrix]** *Let* $A \in \mathbb{M}_n(\mathbb{R})$*. Then there exists a real orthogonal matrix* Q *and an upper triangular matrix* R *such that* $A = QR$*.*

Proof. We start by applying the plane rotations to A so that the positions $(2, 1), (3, 1), \ldots, (n, 1)$ of A become zero. This means, if $a_{21} = 0$, we multiply by I. Otherwise, we use the plane rotation $U(\theta; 1, 2)$, where $\theta = \cot^{-1}(-a_{11}/a_{21})$. Then, we apply a similar technique to A so that the $(3, 1)$ entry of A becomes 0. Note that this plane rotation doesn't change the $(2, 1)$ entry of A. We continue this process till all the entry in the first column of A, except possibly the $(1, 1)$ entry, is zero.

We then apply the plane rotations to make positions $(3, 2), (4, 2), \ldots, (n, 2)$ zero. Observe that this does not disturb the zeros in the first column. Thus, continuing the above process a finite number of times give us the required result. ∎

Lemma 8.2.8. **[QR Factorization Revisited: Rectangular Matrix]** *Let* $A \in \mathbb{M}_{m,n}(\mathbb{R})$*. Then there exists a real orthogonal matrix* Q *and a matrix* $R \in \mathbb{M}_{m,n}(\mathbb{R})$ *in upper triangular form such that* $A = QR$*.*

Proof. If $\text{Rank}A < m$, add some columns to A to get a matrix, say $\tilde{A}$ such that $\text{Rank}\tilde{A} = m$. Now suppose that $\tilde{A}$ has k columns. For $1 \le i \le k$, let $\mathbf{v}_i = \tilde{A}[:, i]$. Now, apply the Gram-Schmidt Orthonormalization Process to $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. For example, suppose the result is a sequence of k vectors $\mathbf{w}_1, 0, \mathbf{w}_2, 0, 0, \ldots, 0, \mathbf{w}_m, 0, \ldots, 0$, where $Q = \begin{bmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_m \end{bmatrix}$ is real orthogonal. Then $\tilde{A}[:, 1]$ is a linear combination of $\mathbf{w}_1$, $\tilde{A}[:, 2]$ is also a linear combination of $\mathbf{w}_1$, $\tilde{A}[:, 3]$ is a linear combination of $\mathbf{w}_1, \mathbf{w}_2$ and so on. In general, for $1 \le s \le k$, the column $\tilde{A}[:, s]$ is a linear combination of $\mathbf{w}_i$-s in the list that appear up to the s-th position. Thus, $\tilde{A}[:, s] = \sum\limits_{i=1}^{m} \mathbf{w}_i r_{is}$, where $r_{is} = 0$ for all $i > s$. That is, $\tilde{A} = QR$, where $R = [r_{ij}]$. Now, remove the extra columns of $\tilde{A}$ and the corresponding columns in R to get the required result. ∎

Note that Proposition 8.2.7 is also valid for any complex matrix. In this case the matrix Q will be unitary. This can also be seen from Theorem 5.2.1 as we need to apply the Gram-Schmidt Orthonormalization Process to vectors in $\mathbb{C}^n$.

To proceed further recall that a matrix $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{C})$ is called a **tri-diagonal** matrix if $a_{ij} = 0$, whenever $|i - j| > 1, 1 \le i, j \le n$.

Proposition 8.2.9. [Tridiagonalization of a Real Symmetric Matrix: Given's Method] *Let A be a real symmetric. Then, there exists a real orthogonal matrix Q such that QAQ^T is a tri-diagonal matrix.*

Proof. If $a_{31} \neq 0$, then put $U_1 = U(\theta_1; 2, 3)$, where $\theta_1 = \cot^{-1}(-a_{21}/a_{31})$. Notice that $U_1^T[:, 1] = \mathbf{e}_1$ and so

$$(U_1 A U_1^T)[:, 1] = (U_1 A)[:, 1].$$

We already know that $U_1 A[3, 1] = 0$. Hence, $U_1 A U_1^T$ is a real symmetric matrix with $(3, 1)$-th entry 0. Now, proceed to make the $(4, 1)$-th entry of $U_1 A$ equal to 0. To do so, take $U_2 = U(\theta_2; 2, 4)$. Notice that $U_2^T(:, 1) = \mathbf{e}_1$ and so

$$(U_2(U_1 A U_1^T) U_2^T)[:, 1) = (U_2 U_1 A U_1^T)[:, 1].$$

But by our choice of the plane rotation U_2, we have $U_2(U_1 A U_1^T)(4, 1) = 0$. Furthermore, as $U_2[3, :] = \mathbf{e}_3^T$, we have

$$(U_2 U_1 A U_1^T)[3, 1] = U_2[3, :](U_1 A U_1^T)[:, 1] = (U_1 A U_1^T)[3, 1] = 0.$$

That is, the previous zeros are preserved.

Continuing this way, we can find a real orthogonal matrix Q such that QAQ^T is tri-diagonal. ∎

Proposition 8.2.10. [Almost Diagonalization of a Real Symmetric Matrix: Jacobi method] *Let $A \in M_n(\mathbb{R})$ be real symmetric. Then there exists a real orthogonal matrix S, a product of plane rotations, such that SAS^T is almost a diagonal matrix.*

Proof. The idea is to reduce the off-diagonal entries of A to 0 as much as possible. So, we start with choosing $i \neq j)$ such that $i < j$ and $|a_{ij}|$ is maximum. Now, put

$$\theta = \frac{1}{2} \cot^{-1} \frac{a_{ii} - a_{jj}}{2a_{ij}}, \quad U = U(\theta; i, j), \quad \text{and} \quad B = U^T A U.$$

Then, for all $l, k \neq i, j$, we see that

$$\begin{aligned}
b_{lk} &= U^T[l,:]AU[:,k] = \mathbf{e}_l^T A\mathbf{e}_k = a_{lk} \\
b_{ik} &= U^T[i,:]AU[:,k] = (\cos\theta\mathbf{e}_i^T + \sin\theta\mathbf{e}_j^T)A\mathbf{e}_k = a_{ik}\cos\theta + a_{jk}\sin\theta \\
b_{lj} &= U^T[l,:]AU[:,j] = \mathbf{e}_l^T A(-\sin\theta\mathbf{e}_i + \cos\theta\mathbf{e}_j) = -a_{li}\sin\theta + a_{lj}\cos\theta \\
b_{ij} &= U^T[i,:]AU[:,j] = (\cos\theta\mathbf{e}_i^T + \sin\theta\mathbf{e}_j^T)A(-\sin\theta\mathbf{e}_i + \cos\theta\mathbf{e}_j) \\
&= \sin(2\theta)\tfrac{a_{jj}-a_{ii}}{2} + a_{ij}\cos(2\theta) = 0
\end{aligned}$$

Thus, using the above, we see that whenever $l, k \neq i, j$, $a_{lk}^2 = b_{lk}^2$ and for $l \neq i, j$, we have

$$b_{il}^2 + b_{lj}^2 = a_{il}^2 + a_{lj}^2.$$

As U is unitary and $B = U^T AU$, we get $\sum |a_{ij}|^2 = \sum |b_{ij}|^2$. Further, $b_{ij} = 0$ implies that

$$a_{ii}^2 + 2a_{ij}^2 + a_{jj}^2 = b_{ii}^2 + 2b_{ij}^2 + b_{jj}^2 = b_{ii}^2 + b_{jj}^2.$$

As the rest of the diagonal entries have not changed, we observe that the sum of the squares of the off-diagonal entries have reduced by $2a_{ij}^2$. Thus, a repeated application of the above process makes the matrix "close to diagonal". ■

8.2.2 Householder Matrices

We will now look at another class of unitary matrices, commonly called the Householder matrices (see Exercise 1.3.7.11).

Definition 8.2.11. **[Householder Matrix]** Let $\mathbf{w} \in \mathbb{C}^n$ be a unit vector. Then, the matrix $U_{\mathbf{w}} = I - 2\mathbf{w}\mathbf{w}^*$ is called a **Householder matrix**.

Remark 8.2.12. *We observe the following about the Householder matrix $U_{\mathbf{w}}$.*

1. *$U_{\mathbf{w}} = I - 2\mathbf{w}\mathbf{w}^*$ is the sum of two Hermitian matrices and hence is also Hermitian.*
2. *$U_{\mathbf{w}}U_{\mathbf{w}}^* = (I - 2\mathbf{w}\mathbf{w}^*)(I - 2\mathbf{w}\mathbf{w}^*)^T = I - 2\mathbf{w}\mathbf{w}^* - 2\mathbf{w}\mathbf{w}^* + 4\mathbf{w}\mathbf{w}^* = I$. Or equivalently, verify that $\|U_{\mathbf{w}}\mathbf{x}\| = \|\mathbf{x}\|$, for all $\mathbf{x} \in \mathbb{C}^n$. So $U_{\mathbf{w}}$ is unitary.*
3. *If $\mathbf{x} \in \mathbf{w}^\perp$ then $U_{\mathbf{w}}\mathbf{x} = \mathbf{x}$.*
4. *If $\mathbf{x} = c\mathbf{w}$, for some $cin\mathbb{C}$, then $U_{\mathbf{w}}\mathbf{x} = -\mathbf{x}$.*
5. *Thus, if $\mathbf{v} \in \mathbb{C}^n$ then we know that $\mathbf{v} = \mathbf{x} + \mathbf{y}$, where $\mathbf{x} \in \mathbf{w}^\perp$ and $\mathbf{y} = c\mathbf{w}$, for some $c \in \mathbb{C}$. In this case, $U_{\mathbf{w}}\mathbf{v} = U_{\mathbf{w}}(\mathbf{x} + \mathbf{y}) = \mathbf{x} - \mathbf{y}$.*
6. *Geometrically, $U_{\mathbf{w}}\mathbf{v}$ reflects the vector $\mathbf{v}$ along the vector $\mathbf{w}^\perp$. Thus, $U_{\mathbf{w}}$ is a reflection matrix along $\mathbf{w}^\perp$ (see Exercise 1.3.7.??).*

Example 8.2.13. In $\mathbb{R}^2$, let $\mathbf{w} = \mathbf{e}_2$. Then $\mathbf{w}^\perp$ is the x-axis. The vector $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{e}_1 + 2\mathbf{e}_2$, where $\mathbf{e}_1 \in \mathbf{w}^\perp$ and $2\mathbf{e}_2 \in LS(\mathbf{w})$. So

$$U_{\mathbf{w}}(\mathbf{e}_1 + 2\mathbf{e}_2) = U_{\mathbf{w}}\mathbf{v} = U_{\mathbf{w}}(\mathbf{x} + \mathbf{y}) = \mathbf{x} - \mathbf{y} = \mathbf{e}_1 - 2\mathbf{e}_2.$$

That is, the reflection of $\mathbf{v}$ along the x-axis ($\mathbf{w}^\perp$).

Recall that if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with $\mathbf{x} \neq \mathbf{y}$ and $\|\mathbf{x}\| = \|\mathbf{y}\|$ then, $(\mathbf{x}+\mathbf{y}) \perp (\mathbf{x}-\mathbf{y})$. This is not true in $\mathbb{C}^n$ as can be seen from the following example. Take $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} i \\ -1 \end{bmatrix}$. Then $\langle \begin{bmatrix} 1+i \\ 0 \end{bmatrix}, \begin{bmatrix} 1-i \\ 2 \end{bmatrix} \rangle = (1+i)^2 \neq 0$. Thus, to pick the right choice for the matrix $U_{\mathbf{w}}$, we need to be observant of the choice of the inner product space.

Example 8.2.14. Let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ with $\mathbf{x} \neq \mathbf{y}$ and $\|\mathbf{x}\| = \|\mathbf{y}\|$. Then, which $U_{\mathbf{w}}$ should be used to reflect $\mathbf{y}$ to $\mathbf{x}$?

1. **Solution in case of $\mathbb{R}^n$:** Imagine the line segment joining $\mathbf{x}$ and $\mathbf{y}$. Now, place a mirror at the midpoint and perpendicular to the line segment. Then, the reflection of $\mathbf{y}$ on that mirror is $\mathbf{x}$. So, take $\mathbf{w} = \frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|} \in \mathbb{R}^n$. Then,

$$\begin{aligned} U_{\mathbf{w}}\mathbf{y} &= (I - 2\mathbf{w}\mathbf{w}^T)\mathbf{y} = \mathbf{y} - 2\mathbf{w}\mathbf{w}^T\mathbf{y} = \mathbf{y} - 2\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|^2}(\mathbf{x}-\mathbf{y})^T\mathbf{y} \\ &= \mathbf{y} - 2\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|^2}\frac{-\|\mathbf{x}-\mathbf{y}\|^2}{2} = \mathbf{x}. \end{aligned}$$

2. **Solution in case of $\mathbb{C}^n$:** Suppose there is a unit vector $\mathbf{w} \in \mathbb{C}^n$ such that $(I - 2\mathbf{w}\mathbf{w}^*)\mathbf{y} = \mathbf{x}$. Then $\mathbf{y} - \mathbf{x} = 2\mathbf{w}\mathbf{w}^*\mathbf{y}$ and hence $\mathbf{w}^*(\mathbf{y}-\mathbf{x}) = 2\mathbf{w}^*\mathbf{w}\mathbf{w}^*\mathbf{y} = 2\mathbf{w}^*\mathbf{y}$. Thus,

$$\mathbf{w}^*(\mathbf{y}+\mathbf{x}) = 0, \text{ that is, } \mathbf{w} \perp (\mathbf{y}+\mathbf{x}). \tag{8.2.1}$$

Furthermore, again using $\mathbf{w}^*(\mathbf{y}+\mathbf{x}) = 0$, we get $\mathbf{y} - \mathbf{x} = 2\mathbf{w}\mathbf{w}^*\mathbf{y} = -2\mathbf{w}\mathbf{w}^*\mathbf{x}$. So,

$$2(\mathbf{y}-\mathbf{x}) = 2\mathbf{w}\mathbf{w}^*(\mathbf{y}-\mathbf{x}) \text{ or } \mathbf{y}-\mathbf{x} = \mathbf{w}\mathbf{w}^*(\mathbf{y}-\mathbf{x}).$$

On the other hand, using Equation (8.2.1), we get $\mathbf{w}\mathbf{w}^*(\mathbf{y}+\mathbf{x}) = 0$. So,

$$0 = [(\mathbf{y}+\mathbf{x})^*\mathbf{w}\mathbf{w}^*](\mathbf{y}-\mathbf{x}) = (\mathbf{y}+\mathbf{x})^*[\mathbf{w}\mathbf{w}^*(\mathbf{y}-\mathbf{x})] = (\mathbf{y}+\mathbf{x})^*(\mathbf{y}-\mathbf{x}).$$

Therefore, if such a $\mathbf{w}$ exists, then $(\mathbf{y}-\mathbf{x}) \perp (\mathbf{y}+\mathbf{x})$.

But, in that case, $\mathbf{w} = \frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|}$ will work as using above $\|\mathbf{x}-\mathbf{y}\|^2 = 2(\mathbf{y}-\mathbf{x})^*\mathbf{y}$ and

$$\begin{aligned} U_{\mathbf{w}}\mathbf{y} &= (I - 2\mathbf{w}\mathbf{w}^*)\mathbf{y} = \mathbf{y} - 2\mathbf{w}\mathbf{w}^*\mathbf{y} = \mathbf{y} - 2\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|^2}(\mathbf{x}-\mathbf{y})^*\mathbf{y} \\ &= \mathbf{y} - 2\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|^2}\frac{-\|\mathbf{x}-\mathbf{y}\|^2}{2} = \mathbf{x}. \end{aligned}$$

Thus, in this case, if $\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle \neq 0$ then we will not find a $\mathbf{w}$ such that $U_{\mathbf{w}}\mathbf{y} = \mathbf{x}$.

For example, taking $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} i \\ -1 \end{bmatrix}$, we have $\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle \neq 0$.

As an application, we now prove that any real symmetric matrix can be transformed into a tri-diagonal matrix.

Proposition 8.2.15. [**Householder's Tri-Diagonalization**] *Let* $\mathbf{v} \in \mathbb{R}^{n-1}$ *and* $A = \begin{bmatrix} a & \mathbf{v}^T \\ \mathbf{v} & B \end{bmatrix} \in \mathbb{M}_n(\mathbb{R})$ *be a real symmetric matrix. Then, there exists a real orthogonal matrix* Q*, a product of Householder matrices, such that* Q^TAQ *is tri-diagonal.*

Proof. If $\mathbf{v} = \mathbf{e}_1$ then we proceed to apply our technique to the matrix B, a matrix of lower order. So, without loss of generality, we assume that $\mathbf{v} \neq \mathbf{e}_1$.

As we want $Q^T AQ$ to be tri-diagonal, we need to find a vector $\mathbf{w} \in \mathbb{R}^{n-1}$ such that $U_{\mathbf{w}}\mathbf{v} = r\mathbf{e}_1 \in \mathbb{R}^{n-1}$, where $r = \|\mathbf{v}\| = \|U_{\mathbf{w}}\mathbf{v}\|$. Thus, using Example 8.2.14, choose the required vector $\mathbf{w} \in \mathbb{R}^{n-1}$. Then,

$$\begin{bmatrix} 1 & 0 \\ 0 & U_{\mathbf{w}} \end{bmatrix} \begin{bmatrix} a & \mathbf{v}^T \\ \mathbf{v} & B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & U_{\mathbf{w}}^T \end{bmatrix} = \begin{bmatrix} a & \mathbf{v}^T U_{\mathbf{w}}^T \\ U_{\mathbf{w}}\mathbf{v} & U_{\mathbf{w}} B U_{\mathbf{w}}^T \end{bmatrix} = \begin{bmatrix} a & r & 0 \\ r & * & * \\ 0 & * & * \end{bmatrix} = \begin{bmatrix} a & r\mathbf{e}_1^T \\ r\mathbf{e}_1 & S \end{bmatrix},$$

where $S \in \mathbb{M}_{n-1}(\mathbb{R})$ is a symmetric matrix. Now, use induction on the matrix S to get the required result. ∎

8.2.3 Schur's Upper Triangularization Revisited

Definition 8.2.16. Let s and t be two symbols. Then, an expression of the form

$$W(s,t) = s^{m_1}t^{n_1} \dots s^{m_k}t^{n_k} \text{ where } m_i, n_i \text{ are positive integers}$$

is called a **word in symbols** s **and** t of degree $\sum\limits_{i=1}^{k}(m_i + n_i)$.

Remark 8.2.17. [More on Unitary Equivalence] *Let s and t be two symbols and $W(s,t)$ be a word in symbols s and t.*

1. *Suppose U is a unitary matrix such that $B = U^*AU$. Then, $W(A, A^*) = U^*W(B, B^*)U$. Thus,* $\mathsf{tr}[\mathsf{W}(\mathsf{A}, \mathsf{A}^*)] = \mathsf{tr}[\mathsf{W}(\mathsf{B}, \mathsf{B}^*)]$.
2. *Let A and B be two matrices such that* $\mathsf{tr}[\mathsf{W}(\mathsf{A}, \mathsf{A}^*)] = \mathsf{tr}[\mathsf{W}(\mathsf{B}, \mathsf{B}^*)]$, *for each word W. Then, does it imply that A and B are unitarily equivalent? The answer is* **'yes'** *as provided by the following result. The proof is outside the scope of this book.*

   Theorem 8.2.18. [Specht-Pearcy] *Let $A, B \in \mathbb{M}_n(\mathbb{C})$ and suppose that* $\mathsf{tr}[\mathsf{W}(\mathsf{A}, \mathsf{A}^*)] = \mathsf{tr}[\mathsf{W}(\mathsf{B}, \mathsf{B}^*)]$ *holds for all words of degree less than or equal to $2n^2$. Then $B = U^*AU$, for some unitary matrix U.*

EXERCISE **8.2.19. [Triangularization via Complex Orthogonal Matrix need not be Possible]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ and $A = QTQ^T$, where Q is complex orthogonal matrix and T is upper triangular. Then, prove that*

1. *A has an eigenvector $\mathbf{x}$ such that $\mathbf{x}^T\mathbf{x} \neq 0$.*
2. *there is no orthogonal matrix Q such that $Q^T \begin{bmatrix} 1 & \mathbf{i} \\ \mathbf{i} & -1 \end{bmatrix} Q$ is upper triangular.*

Proposition 8.2.20. [Matrices with Distinct Eigenvalues are Dense in $\mathbb{M}_n(\mathbb{C})$] *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, for each $\epsilon > 0$, there exists a matrix $A(\epsilon) \in \mathbb{M}_n(\mathbb{C})$ such that $A(\epsilon) = [a(\epsilon)_{ij}]$ has distinct eigenvalues and $\sum |a_{ij} - a(\epsilon)_{ij}|^2 < \epsilon$.*

Proof. By Schur Upper Triangularization (see Lemma 6.2.12), there exists a unitary matrix U such that $U^*AU = T$, an upper triangular matrix. Now, choose α_i's such that $t_{ii} + \alpha_i$ are distinct and $\sum |\alpha_i|^2 < \epsilon$. Now, consider the matrix $A(\epsilon) = U\left(T + \text{ diag}(\alpha_1, \ldots, \alpha_n)\right)U^*$. Then, $B = A(\epsilon) - A = U \text{ diag}(\alpha_1, \ldots, \alpha_n)]U^*$ with

$$\sum_{i,j} |b_{ij}|^2 = \mathsf{tr}(\mathsf{B}^*\mathsf{B}) = \mathsf{trU} \text{ diag}(|\alpha_1|^2, \ldots, |\alpha_\mathsf{n}|^2)\mathsf{U}^* = \sum_{\mathsf{i}} |\alpha_\mathsf{i}|^2 < \epsilon.$$

Thus, the required result follows. ∎

Before proceeding with our next result on almost diagonalizability, we look at the following example.

Example 8.2.21. Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ and $\epsilon > 0$ be given. Then, determine a diagonal matrix D such that the non-diagonal entry of $D^{-1}AD$ is less than ϵ.

Solution: Choose $\alpha < \dfrac{\epsilon}{2}$ and define $D = \text{ diag}(1, \alpha)$. Then,

$$D^{-1}AD = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\alpha} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} = \begin{bmatrix} 1 & 2\alpha \\ 0 & 3 \end{bmatrix}.$$

As $\alpha < \dfrac{\epsilon}{2}$, the required result follows.

Proposition 8.2.22. [**A matrix is Almost Diagonalizable**] *Let $A \in \mathbb{M}_n(\mathbb{C})$ and $\epsilon > 0$ be given. Then, there exists an invertible matrix S_ϵ such that $S_\epsilon^{-1}AS_\epsilon = T$, an upper triangular matrix with $|t_{ij}| < \epsilon$, for all $i \neq j$.*

Proof. By Schur Upper Triangularization (see Lemma 6.2.12), there exists a unitary matrix U such that $U^*AU = T$, an upper triangular matrix. Now, take $t = 2 + \max\limits_{i<j} |t_{ij}|$ and choose α such that $0 < \alpha < \epsilon/t$. Then, if we take $D_\alpha = \text{ diag}(1, \alpha, \alpha^2, \ldots, \alpha^{n-1})$ and $S = UD_\alpha$, we have $S^{-1}AS = D_{\frac{1}{\alpha}}TD_\alpha = F$ (say), an upper triangular. Furthermore, note that for $i < j$, we have $|f_{ij}| = |t_{ij}|\alpha^{j-i} \leq \epsilon$. Thus, the required result follows. ∎

8.3 Commuting Matrices and Simultaneous Diagonalization

Definition 8.3.1. [**Simultaneously Diagonalizable**] Let $A, B \in \mathbb{M}_n(\mathbb{C})$. Then, they are said to be **simultaneously diagonalizable** if there exists an invertible matrix S such that $S^{-1}AS$ and $S^{-1}BS$ are both diagonal matrices.

Since diagonal matrices commute, we have our next result.

Proposition 8.3.2. *Let $A, B \in \mathbb{M}_n(\mathbb{C})$. If A and B are simultaneously diagonalizable then $AB = BA$.*

Proof. By definition, there exists an invertible matrix S such that $S^{-1}AS = \Lambda_1$ and $S^{-1}BS = \Lambda_2$. Hence,

$$AB = (S\Lambda_1S^{-1}) \cdot (S\Lambda_2S^{-1}) = S\Lambda_1\Lambda_2S^{-1} = S\Lambda_2\Lambda_1S^{-1} = S\Lambda_2S^{-1}S\Lambda_1S^{-1} = BA.$$

Thus, we have proved the required result. ∎

Theorem 8.3.3. *Let $A, B \in \mathbb{M}_n(\mathbb{C})$ be diagonalizable matrices. Then they are simultaneously diagonalizable if and only if they commute.*

Proof. One part of this theorem has already been proved in Proposition 8.3.2. For the other part, let us assume that $AB = BA$. Since A is diagonalizable, there exists an invertible matrix S such that

$$S^{-1}AS = \Lambda = \lambda_1 I \oplus \cdots \oplus \lambda_k I, \tag{8.3.1}$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A. We now use the sub-matrix structure of $S^{-1}AS$ to decompose $C = S^{-1}BS$ as $C = \begin{bmatrix} C_{11} & \cdots & C_{1k} \\ & \ddots & \\ C_{k1} & \cdots & C_{kk} \end{bmatrix}$. Since $AB = BA$ and S is invertible, we have $\Lambda C = C\Lambda$. Thus,

$$\begin{bmatrix} \lambda_1 C_{11} & \cdots & \lambda_1 C_{1k} \\ & \ddots & \\ \lambda_k C_{k1} & \cdots & \lambda_k C_{kk} \end{bmatrix} = \begin{bmatrix} \lambda_1 C_{11} & \cdots & \lambda_k C_{1k} \\ & \ddots & \\ \lambda_1 C_{k1} & \cdots & \lambda_k C_{kk} \end{bmatrix}.$$

Since $\lambda_i \neq \lambda_j$ for $1 \leq i \neq j \leq k$, we have $C_{ij} = 0$, whenever $i \neq j$. Thus, the matrix $C = C_{11} \oplus \cdots \oplus C_{kk}$.

Since B is diagonalizable, the matrix C is also diagonalizable and hence the matrices C_{ii}, for $1 \leq i]lek$, are diagonalizable. So, for $1 \leq i \leq k$, there exists invertible matrices T_i's such that $T_i^{-1} C_{ii} T_i = \Lambda_i$. Put $T = T_1 \oplus \cdots \oplus T_k$. Then,

$$T^{-1}S^{-1}AST = \begin{bmatrix} T_1^{-1} & & \\ & \ddots & \\ & & T_k^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 I & & \\ & \ddots & \\ & & \lambda_k I \end{bmatrix} \begin{bmatrix} T_1 & & \\ & \ddots & \\ & & T_k \end{bmatrix} = \begin{bmatrix} \lambda_1 I & & \\ & \ddots & \\ & & \lambda_k I \end{bmatrix}$$

and

$$T^{-1}S^{-1}BST = \begin{bmatrix} T_1^{-1} & & \\ & \ddots & \\ & & T_k^{-1} \end{bmatrix} \begin{bmatrix} C_{11} & & \\ & \ddots & \\ & & C_{kk} \end{bmatrix} \begin{bmatrix} T_1 & & \\ & \ddots & \\ & & T_k \end{bmatrix} = \begin{bmatrix} \Lambda_1 & & \\ & \ddots & \\ & & \Lambda_k \end{bmatrix}.$$

Thus A and B are simultaneously diagonalizable and the required result follows. ∎

Definition 8.3.4. **[Commuting Family of Matrices]**

1. Let $\mathcal{F} \subseteq \mathbb{M}_n(\mathbb{C})$. Then $\mathcal{F}$ is said to be a **commuting family** if each pair of matrices in $\mathcal{F}$ commutes.

2. Let $B \in \mathbb{M}_n(\mathbb{C})$ and W be a subspace of $\mathbb{C}^n$. Then, W is said to be a B-**invariant** subspace if $B\mathbf{w} \in W$, for all $\mathbf{w} \in W$ (or equivalently, $BW \subseteq W$).

3. A subspace W of $\mathbb{C}^n$ is said to be $\mathcal{F}$-**invariant** if W is B-invariant for each $B \in \mathcal{F}$.

Example 8.3.5. Let $A \in \mathbb{M}_n(\mathbb{C})$ with $(\lambda, \mathbf{x})$ as an eigenpair. Then, $W = \{c\mathbf{x} : c \in \mathbb{C}\}$ is an A-invariant subspace. Furthermore, if W is an A-invariant subspace with $\dim(W) = 1$ then verify that any non-zero vector in W is an eigenvector of A.

Theorem 8.3.6. **[An A-invariant Subspace Contains an Eigenvector of A]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ and $W \subseteq \mathbb{C}^n$ be an A-invariant subspace of dimension at least 1. Then W contains an eigenvector of A.*

Proof. Let $\mathcal{B} = \{\mathbf{f}_1, \ldots, \mathbf{f}_k\} \subseteq \mathbb{C}^n$ be an ordered basis for W. Define $T : W \to W$ as $T\mathbf{v} = A\mathbf{v}$. Then $T[\mathcal{B}, \mathcal{B}] = \begin{bmatrix} [T\mathbf{f}_1]_{\mathcal{B}} & \cdots & [T\mathbf{f}_k]_{\mathcal{B}} \end{bmatrix}$ is a $k \times k$ matrix which satisfies $[T\mathbf{w}]_{\mathcal{B}} = T[\mathcal{B}, \mathcal{B}]\ [\mathbf{w}]_{\mathcal{B}}$, for all $\mathbf{w} \in W$. As $T[\mathcal{B}, \mathcal{B}] \in \mathbb{M}_k(\mathbb{C})$, it has an eigenpair, say $(\lambda, \hat{\mathbf{x}})$ with $\hat{\mathbf{x}} \in \mathbb{C}^k$. That is,

$$T[\mathcal{B}, \mathcal{B}]\hat{\mathbf{x}} = \lambda\hat{\mathbf{x}}. \tag{8.3.2}$$

Now, put $\mathbf{x} = \sum\limits_{i=1}^{k} (\hat{\mathbf{x}})_i\, \mathbf{f}_i \in \mathbb{C}^n$. Then, verify that $\mathbf{x} \in W$ and $[\mathbf{x}]_{\mathcal{B}} = \hat{\mathbf{x}}$. Thus, $T\mathbf{x} \in W$ and now using Equation (8.3.2), we get

$$T\mathbf{x} = \sum_{i=1}^{k} \left([T\mathbf{x}]_{\mathcal{B}}\right)_i \mathbf{f}_i = \sum_{i=1}^{k} \left(T[\mathcal{B}, \mathcal{B}][\mathbf{x}]_{\mathcal{B}}\right)_i \mathbf{f}_i = \sum_{i=1}^{k} \left(T[\mathcal{B}, \mathcal{B}]\hat{\mathbf{x}}\right)_i \mathbf{f}_i = \sum_{i=1}^{k} (\lambda\hat{\mathbf{x}})_i \mathbf{f}_i = \lambda \sum_{i=1}^{k} (\hat{\mathbf{x}})_i \mathbf{f}_i = \lambda\mathbf{x}.$$

So, A has an eigenvector $\mathbf{x} \in W$ corresponding to the eigenvalue λ. ∎

Theorem 8.3.7. *Let $\mathcal{F} \subseteq \mathbb{M}_n(\mathbb{C})$ be a commuting family of matrices. Then, all the matrices in $\mathcal{F}$ have a common eigenvector.*

Proof. Note that $\mathbb{C}^n$ is $\mathcal{F}$-invariant. Let $W \subseteq \mathbb{C}^n$ be $\mathcal{F}$-invariant with minimum positive dimension. Let $\mathbf{y} \in W$ such that $\mathbf{y} \neq 0$. We claim that $\mathbf{y}$ is an eigenvector, for each $A \in \mathcal{F}$.

So, on the contrary assume $\mathbf{y}$ is not an eigenvector for some $A \in \mathcal{F}$. Then, by Theorem 8.3.6, W contains an eigenvector $\mathbf{x}$ of A for some eigenvalue, say λ. Define $W_0 = \{\mathbf{z} \in W : A\mathbf{z} = \lambda\mathbf{z}\}$. So W_0 is a proper subspace of W as $\mathbf{y} \in W \setminus W_0$. Also, for $\mathbf{z} \in W_0$ and $C \in \mathcal{F}$, we note that $A(C\mathbf{z}) = CA\mathbf{z} = \lambda(C\mathbf{z})$, so that $C\mathbf{z} \in W_0$. So W_0 is $\mathcal{F}$-invariant and $1 \leq \dim W_0 < \dim W$, a contradiction. ∎

Theorem 8.3.8. *Let $\mathcal{F} \subseteq \mathbb{M}_n(\mathbb{C})$ be a family of diagonalizable matrices. Then $\mathcal{F}$ is commuting if and only if $\mathcal{F}$ is simultaneously diagonalizable.*

Proof. We prove the result by induction on n. The result is clearly true for $n = 1$. So, let us assume the result to be valid for all $n < m$. Now, let us assume that $\mathcal{F} \subseteq \mathbb{M}_m(\mathbb{C})$ is a family of diagonalizable matrices.

If $\mathcal{F}$ is simultaneously diagonalizable, then by Proposition 8.3.2, the family $\mathcal{F}$ is commuting. Conversely, let $\mathcal{F}$ be a commuting family. If each $A \in \mathcal{F}$ is a scalar matrix then they are simultaneously diagonalizable via I. So, let $A \in \mathcal{F}$ be a non-scalar matrix. As A is diagonalizable, there exists an invertible matrix S such that

$$S^{-1}AS = \lambda_1 I \oplus \cdots \oplus \lambda_k I,\ k \geq 2,$$

where λ_i's are distinct. Now, consider the family $\mathcal{G} = \{\hat{X} = S^{-1}XS \mid X \in \mathcal{F}\}$. As $\mathcal{F}$ is a commuting family, the set $\mathcal{G}$ is also a commuting family. So, each $\hat{X} \in \mathcal{G}$ has the form $\hat{X} = X_1 \oplus \cdots \oplus X_k$. Note that $\mathcal{H}_i = \{X_i \mid \hat{X} \in \mathcal{G}\}$ is a commuting family of diagonalizable matrices of size $< m$. Thus, by induction hypothesis, $\mathcal{H}_i$'s are simultaneously diagonalizable, say by the invertible matrices T_i's. That is, $T_i^{-1}X_iT_i = \Lambda_i$, a diagonal matrix, for $1 \le i \le k$. Thus, if $T = T_1 \oplus \cdots \oplus T_k$ then

$$T^{-1}S^{-1}\hat{X}ST = T^{-1}(X_1 \oplus \cdots \oplus X_k)T = T_1^{-1}X_1T_1 \oplus \cdots \oplus T_k^{-1}X_kT_k = \Lambda_1 \oplus \cdots \oplus \Lambda_k,$$

a diagonal matrix, for all $X \in \mathcal{F}$. Thus the result holds by induction. ■

We now give prove of some parts of Exercise 6.1.24.exe:eigen:1.

Remark 8.3.9. [$\sigma(AB)$ **and** $\sigma(BA)$] *Let $m \le n$, $A \in \mathbb{M}_{m\times n}(\mathbb{C})$, and $B \in \mathbb{M}_{n\times m}(\mathbb{C})$. Then $\sigma(BA) = \sigma(AB)$ with $n-m$ extra 0's. In particular, if $A, B \in \mathbb{M}_n(\mathbb{C})$ then, $P_{AB}(t) = P_{BA}(t)$.*

Proof. Note that

$$\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}\begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} AB & ABA \\ B & BA \end{bmatrix} = \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix}\begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}.$$

Thus, the matrices $\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$ are similar. Hence, AB and BA have precisely the same non-zero eigenvalues. Therefore, if they have the same size, they must have the same characteristic polynomial. ■

EXERCISE **8.3.10. [Miscellaneous Exercises]**

1. *Let A be nonsingular. Then, verify that $A^{-1}(AB)A = BA$. Hence, AB and BA are similar. Thus, $P_{AB}(t) = P_{BA}(t)$.*
2. *Fix a positive integer $k, 0 \le k \le n$. Now, define the function $f_k : \mathbb{M}_n(\mathbb{C}) \to \mathbb{C}$ by $f(A) =$ coefficient of t^k in $P_A(t)$. Prove that f_k is a continuous function.*
3. *For any matrix A, prove that there exists an $\epsilon > 0$ such that $A_\alpha = A + \alpha I$ is invertible, for all $\alpha \in (0, \epsilon)$. Thus, use the first part to conclude that for any given B, we have $P_{A_\alpha B}(t) = P_{BA_\alpha}(t)$, for all $\alpha \in (0, \epsilon)$.*
4. *Now, use continuity to argue that $P_{AB}(t) = \lim\limits_{\alpha\to 0+} P_{A_\alpha B}(t) = \lim\limits_{\alpha\to 0+} P_{BA_\alpha}(t) = P_{BA}(t)$.*
5. *Let $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$, $\sigma(B) = \{\mu_1, \ldots, \mu_n\}$ and suppose that $AB = BA$. Then,*
   (a) *prove that there is a permutation π such that $\sigma(A+B) = \{\lambda_1 + \mu_{\pi(1)}, \ldots, \lambda_n + \mu_{\pi(n)}\}$. In particular, $\sigma(A+B) \subseteq \sigma(A) + \sigma(B)$.*
   (b) *if we further assume that $\sigma(A) \cap \sigma(-B) = \emptyset$ then the matrix $A+B$ is nonsingular.*
6. *Let A and B be two non-commuting matrices. Then, give an example to show that it is difficult to relate $\sigma(A+B)$ with $\sigma(A)$ and $\sigma(B)$.*
7. *Are the matrices $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ simultaneously triangularizable?*
8. *Let $\mathcal{F} \subseteq \mathbb{M}_n(\mathbb{C})$ be a family of commuting normal matrices. Then, prove that each element of $\mathcal{F}$ is simultaneously unitarily diagonalizable.*
9. *Let $A \in \mathbb{M}_n(\mathbb{C})$ with $A^* = A$ and $\mathbf{x}^*A\mathbf{x} \ge 0$, for all $x \in \mathbb{C}^n$. Then prove that $\sigma(A) \subseteq \mathbb{R}_+$ and if $\mathsf{tr(A)} = 0$, then $A = \mathbf{0}$.*

8.3.1 Diagonalization and Real Orthogonal Matrix

Proposition 8.3.11. **[Triangularization: Real Matrix]** *Let $A \in \mathbb{M}_n(\mathbb{R})$. Then, there exists a real orthogonal matrix Q such that Q^TAQ is block upper triangular, where each diagonal block is of size either* 1 *or* 2.

Proof. If all the eigenvalues of A are real then the corresponding eigenvectors have real entries and hence, one can use induction to get the result in this case (see Lemma 6.2.12).

So, now let us assume that A has a complex eigenvalue, say $\lambda = \alpha + i\beta$ with $\beta \neq 0$ and $\mathbf{x} = \mathbf{u} + i\mathbf{v}$ as an eigenvector for λ. Thus, $A\mathbf{x} = \lambda\mathbf{x}$ and hence $A\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$. But, $\lambda \neq \overline{\lambda}$ as $\beta \neq 0$. Thus, the eigenvectors $\mathbf{x}, \overline{\mathbf{x}}$ are linearly independent and therefore, $\{\mathbf{u}, \mathbf{v}\}$ is a linearly independent set. By Gram-Schmidt Orthonormalization process, we get an ordered basis, say $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ of $\mathbb{R}^n$, where $LS(\mathbf{w}_1, \mathbf{w}_2) = LS(\mathbf{u}, \mathbf{v})$. Also, using the eigen-condition $A\mathbf{x} = \lambda\mathbf{x}$ gives

$$A\mathbf{w}_1 = a\mathbf{w}_1 + b\beta\mathbf{w}_2, \; A\mathbf{w}_2 = c\mathbf{w}_1 + d\mathbf{w}_2,$$

for some real numbers a, b, c and d.

Now, form a matrix $X = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n]$. Then, X is a real orthogonal matrix and

$$\begin{aligned} X^*AX &= X^*[A\mathbf{w}_1, A\mathbf{w}_2, \ldots, A\mathbf{w}_n] = \begin{bmatrix} \mathbf{w}_1^* \\ \mathbf{w}_2^* \\ \vdots \\ \mathbf{w}_n^* \end{bmatrix} [a\mathbf{w}_1 + b\mathbf{w}_2, c\mathbf{w}_1 + d\mathbf{w}_2, \ldots, A\mathbf{w}_n] \\ &= \begin{bmatrix} \begin{matrix} a & b \\ c & d \end{matrix} & * \\ \mathbf{0} & B \end{bmatrix} \end{aligned} \tag{8.3.3}$$

where $B \in \mathbb{M}_{n-2}(\mathbb{R})$. Now, by induction hypothesis the required result follows. ■

The next result is a direct application of Proposition 8.3.11 and hence the proof is omitted.

Corollary 8.3.12. **[Simultaneous Triangularization: Real Matrices]** *Let $\mathcal{F} \subseteq \mathbb{M}_n(\mathbb{R})$ be a commuting family. Then, there exists a real orthogonal matrix Q such that Q^TAQ is a block upper triangular matrix, where each diagonal block is of size either* 1 *or* 2*, for all $A \in \mathcal{F}$.*

Proposition 8.3.13. *Let $A \in \mathbb{M}_n(\mathbb{R})$. Then the following statements are equivalent.*

1. *A is normal.*
2. *There exists a real orthogonal matrix Q such that $Q^TAQ = \bigoplus_i A_i$, where A_i's are real normal matrices of size either* 1 *or* 2.

Proof. $2 \Rightarrow 1$ is trivial. To prove $1 \Rightarrow 2$, recall that Proposition 8.3.11 gives the existence of a real orthogonal matrix Q such that Q^TAQ is upper triangular with diagonal blocks of size either 1 or 2. So, we can write

$$Q^TAQ = \left[\begin{array}{ccc|ccc} \lambda_1 & * & * & * & * & * \\ 0 & \ddots & * & * & * & * \\ 0 & \cdots & \lambda_p & * & * & * \\ \hline 0 & \cdots & 0 & A_{11} & \cdots & A_{1k} \\ 0 & \cdots & 0 & 0 & \ddots & * \\ 0 & \cdots & 0 & 0 & \cdots & A_{kk} \end{array}\right] = \begin{bmatrix} R & C \\ 0 & B \end{bmatrix} \text{ (say).}$$

As A is normal, $\begin{bmatrix} R & C \\ 0 & B \end{bmatrix}\begin{bmatrix} R^T & 0 \\ C^T & B^T \end{bmatrix} = \begin{bmatrix} R^T & 0 \\ C^T & B^T \end{bmatrix}\begin{bmatrix} R & C \\ 0 & B \end{bmatrix}$. Thus, $\mathsf{tr}(\mathsf{CC^T}) = \mathsf{tr}(\mathsf{RR^T} - \mathsf{R^TR}) = \mathsf{0}$. Now, using Exercise 8.3.10.9, we get $C = \mathbf{0}$. Hence, $RR^T = R^TR$ and therefore, R is a diagonal matrix.

As $B^TB = BB^T$, we have $\sum A_{1i}A_{1i}^T = A_{11}A_{11}^T$. So $\mathsf{tr}\left(\sum\limits_2^k \mathsf{A_{1i}A_{1i}^T}\right) = \mathsf{0}$. Now, using Exercise 8.3.10.9 again, we have $\sum\limits_2^k A_{1i}A_{1i}^T = 0$ and so $A_{1i}A_{1i}^T = \mathbf{0}$, for all $i = 2, \ldots, k$. Thus, $A_{1i} = \mathbf{0}$, for all $i = 2, \ldots, k$. Hence, the required result follows. ∎

EXERCISE **8.3.14.** *Let $A \in \mathbb{M}_n(\mathbb{R})$. Then the following are true.*

1. *$A = -A^T$ if and only if A is real orthogonally similar to $[\bigoplus_j 0] \oplus [\bigoplus_i A_i]$, where $A_i = \begin{bmatrix} 0 & a_i \\ -a_i & 0 \end{bmatrix}$, for some real numbers a_i's.*
2. *$AA^T = I$ if and only if A is real orthogonally similar to $[\bigoplus_i \lambda_i] \oplus [\bigoplus_j A_j]$, where $\lambda_i = \pm 1$ and $A_j = \begin{bmatrix} \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j \end{bmatrix}$, for some real numbers θ_i's.*

8.3.2 Convergent and nilpotent matrices

Definition 8.3.15. **[Convergent matrices]** A matrix A is called a **convergent matrix** if $A^m \to \mathbf{0}$ as $m \to \infty$.

Remark 8.3.16. *1. Let A be a diagonalizable matrix with $\rho(A) < 1$. Then, A is a convergent matrix.*

Proof. *Let $A = U^*\ diag(\lambda_1, \ldots, \lambda_n)U$. As $\rho(A) < 1$, for each $i, 1 \le i \le n$, $\lambda_i^m \to 0$ as $m \to \infty$. Thus, $A^m = U^*\ diag(\lambda_1^m, \ldots, \lambda_n^m)U \to \mathbf{0}$.* ∎

2. Even if the matrix A is not diagonalizable, the above result holds. That is, whenever $\rho(A) < 1$, the matrix A is convergent. The converse is also true.

Proof. *Let $J_k(\lambda) = \lambda I_k + N_k$ be a Jordan block of $J =$ JORDAN CFA. Then as $N_k^k = 0$, for each fixed k, we have*

$$J_k(\lambda)^m = \lambda^m + C(m,1)\lambda^{m-1}N_k + \cdots + C(m,k-1)\lambda^{m-k+1}N_k^{k-1} \to 0, \text{ as } m \to \infty.$$

As $\lambda^m \to 0$ as $m \to \infty$, the matrix $J_k(\lambda)^m \to \mathbf{0}$ and hence J is convergent. Thus, A is a convergent matrix.

Conversely, if A is convergent, then J must be convergent. Thus each Jordan block $J_k(\lambda)$ must be convergent. Hence $|\lambda| < 1$. ∎

Theorem 8.3.17. **[Decomposition into Diagonalizable and Nilpotent Parts]** Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $A = B + C$, where B is diagonalizable matrix and C is nilpotent such that $BC = CB$.

Proof. Let $J =$ JORDAN CFA. Then, $J = D + N$, where $D =$ diag(J) and N is clearly a nilpotent matrix.

Now, note that $DN = ND$ as for each Jordan block $J_k(\lambda) = D_k + N_k$, we have $D_k = \lambda I$ and $N_k = J_k(0)$ so that $D_kN_k = N_kD_k$. As $J =$ JORDAN CFA, there exists an invertible matrix S, such that $S^{-1}AS = J$. Hence, $A = SJS^{-1} = SDS^{-1} + SNS^{-1} = B + C$, which satisfy the required conditions. ∎

Chapter 9

Appendix

9.1 Uniqueness of RREF

Definition 9.1.1. Fix $n \in \mathbb{N}$. Then, for each $f \in \mathcal{S}_n$, we associate an $n \times n$ matrix, denoted $P^f = [p_{ij}]$, such that $p_{ij} = 1$, whenver $f(j) = i$ and 0, otherwise. The matrix P^f is called the **Permutation matrix** corresponding to the permutation f. For example, I_2, corresponding to Id_2, and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = E_{12}$, corresponding to the permutation $(1,2)$, are the two permutation matrices of order 2×2.

Remark 9.1.2. *Recall that in Remark 9.2.16.1, it was observed that each permutation is a product of n transpositions, $(1,2), \ldots, (1,n)$.*

1. *Verify that the elementary matrix E_{ij} is the permutation matrix corresponding to the transposition (i,j) .*
2. *Thus, every permutation matrix is a product of elementary matrices E_{1j}, $1 \leq j \leq n$.*
3. *For $n = 3$, the permutation matrices are I_3, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = E_{23} = E_{12}E_{13}E_{12}$, $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E_{12}$, $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = E_{12}E_{13}$, $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = E_{13}E_{12}$ and $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = E_{13}$.*
4. *Let $f \in \mathcal{S}_n$ and $P^f = [p_{ij}]$ be the corresponding permutation matrix. Since $p_{ij} = \delta_{i,j}$ and $\{f(1), \ldots, f(n)\} = [n]$, each entry of P^f is either 0 or 1. Furthermore, every row and column of P^f has exactly one nonzero entry. This nonzero entry is a 1 and appears at the position $p_{i,f(i)}$.*
5. *By the previous paragraph, we see that when a permutation matrix is multiplied to A*
   1. *(a) from left then it permutes the rows of A.*
   2. *(b) from right then it permutes the columns of A.*
6. *P is a permutation matrix if and only if P has exactly one 1 in each row and column.*
   Solution: *If P has exactly one 1 in each row and column, then P is a square matrix, say $n \times n$. Now, apply GJE to P. The occurrence of exactly one 1 in each row and column*

implies that these 1*'s are the pivots in each column. We just need to interchange rows to get it in RREF. So, we need to multiply by* E_{ij}*. Thus, GJE of* P *is* I_n *and* P *is indeed a product of* E_{ij}*'s. The other part has already been explained earlier.*

We are now ready to prove Theorem 2.2.17.

Theorem 9.1.3. *Let* A *and* B *be two matrices in RREF. If they are row equivalent then* $A = B$.

Proof. Note that the matrix $A = \mathbf{0}$ if and only if $B = \mathbf{0}$. So, let us assume that the matrices $A, B \neq \mathbf{0}$. Also, the row-equivalence of A and B implies that there exists an invertible matrix C such that $A = CB$, where C is product of elementary matrices.

Since B is in RREF, either $B[:,1] = \mathbf{0}^T$ or $B[:,1] = (1,0,\ldots,0)^T$. If $B[:,1] = \mathbf{0}^T$ then $A[:,1] = CB[:,1] = C\mathbf{0} = \mathbf{0}$. If $B[:,1] = (1,0,\ldots,0)^T$ then $A[:,1] = CB[:,1] = C[:,1]$. As C is invertible, the first column of C cannot be the zero vector. So, $A[:,1]$ cannot be the zero vector. Further, A is in RREF implies that $A[:,1] = (1,0,\ldots,0)^T$. So, we have shown that if A and B are row-equivalent then their first columns must be the same.

Now, let us assume that the first $k-1$ columns of A and B are equal and it contains r pivotal columns. We will now show that the k-th column is also the same.

Define $A_k = [A[:,1],\ldots,A[:,k]]$ and $B_k = [B[:,1],\ldots,B[:,k]]$. Then, our assumption implies that $A[:,i] = B[:,i]$, for $1 \le i \le k-1$. Since, the first $k-1$ columns contain r pivotal columns, there exists a permutation matrix P such that

$$A_kP = \left[\begin{array}{cc|c} I_r & W & A[:,k] \\ \mathbf{0} & \mathbf{0} & \end{array}\right] \text{ and } B_kP = \left[\begin{array}{cc|c} I_r & W & B[:,k] \\ \mathbf{0} & \mathbf{0} & \end{array}\right].$$

If the k-th columns of A and B are pivotal columns then by definition of RREF, $A[:,k] = \begin{bmatrix} \mathbf{0} \\ \mathbf{e}_1 \end{bmatrix} = B[:,k]$, where $\mathbf{0}$ is a vector of size r and $\mathbf{e}_1 = (1,0,\ldots,0)^T$. So, we need to consider two cases depending on whether both are non-pivotal or one is pivotal and the other is not.

As $A = CB$, we get $A_k = CB_k$ and

$$\left[\begin{array}{cc|c} I_r & W & A[:,k] \\ \mathbf{0} & \mathbf{0} & \end{array}\right] = A_kP = CB_kP = \left[\begin{array}{c|c} C_1 & C_2 \\ C_3 & C_4 \end{array}\right]\left[\begin{array}{cc|c} I_r & W & B[:,k] \\ \mathbf{0} & \mathbf{0} & \end{array}\right] = \left[\begin{array}{cc|c} C_1 & C_1W & CB[:,k] \\ C_3 & C_3W & \end{array}\right].$$

So, we see that $C_1 = I_r$, $C_3 = \mathbf{0}$ and $A[:,k] = \left[\begin{array}{c|c} I_r & C_2 \\ \mathbf{0} & C_4 \end{array}\right] B[:,k]$.

Case 1: Neither $A[:,k]$ nor $B[:,k]$ are pivotal. Then

$$\begin{bmatrix} X \\ \mathbf{0} \end{bmatrix} = A[:,k] = \left[\begin{array}{c|c} I_r & C_2 \\ \mathbf{0} & C_4 \end{array}\right] B[:,k] = \left[\begin{array}{c|c} I_r & C_2 \\ \mathbf{0} & C_4 \end{array}\right]\begin{bmatrix} Y \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} Y \\ \mathbf{0} \end{bmatrix}.$$

Thus, $X = Y$ and in this case the k-th columns are equal.

Case 2: $A[:,k]$ is pivotal but $B[:,k]$ in non-pivotal. Then

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{e}_1 \end{bmatrix} = A[:,k] = \left[\begin{array}{c|c} I_r & C_2 \\ \mathbf{0} & C_4 \end{array}\right] B[:,k] = \left[\begin{array}{c|c} I_r & C_2 \\ \mathbf{0} & C_4 \end{array}\right]\begin{bmatrix} Y \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} Y \\ \mathbf{0} \end{bmatrix},$$

a contradiction as $\mathbf{e}_1 \neq \mathbf{0}$. Thus, this case cannot arise.

Therefore, combining both the cases, we get the required result. ∎

9.2 Permutation/Symmetric Groups

Definition 9.2.1. For a positive integer n, denote $[n] = \{1, 2, \ldots, n\}$. A function $f : A \to B$ is called

1. **one-one/injective** if $f(x) = f(y)$ for some $x, y \in A$ necessarily implies that $x = y$.
2. **onto/surjective** if for each $b \in B$ there exists $a \in A$ such that $f(a) = b$.
3. a **bijection** if f is both one-one and onto.

Example 9.2.2. Let $A = \{1, 2, 3\}$, $B = \{a, b, c, d\}$ and $C = \{\alpha, \beta, \gamma\}$. Then, the function

1. $j : A \to B$ defined by $j(1) = a, j(2) = c$ and $j(3) = c$ is neither one-one nor onto.
2. $f : A \to B$ defined by $f(1) = a, f(2) = c$ and $f(3) = d$ is one-one but not onto.
3. $g : B \to C$ defined by $g(a) = \alpha, g(b) = \beta, g(c) = \alpha$ and $g(d) = \gamma$ is onto but not one-one.
4. $h : B \to A$ defined by $h(a) = 2, h(b) = 2, h(c) = 3$ and $h(d) = 1$ is onto.
5. $h \circ f : A \to A$ is a bijection.
6. $g \circ f : A \to C$ is neither one-one not onto.

Remark 9.2.3. *Let $f : A \to B$ and $g : B \to C$ be functions. Then, the* **composition** *of functions, denoted $g \circ f$, is a function from A to C defined by $(g \circ f)(a) = g(f(a))$. Also, if*

1. *f and g are one-one then $g \circ f$ is one-one.*
2. *f and g are onto then $g \circ f$ is onto.*

Thus, if f and g are bijections then so is $g \circ f$.

Definition 9.2.4. A function $f : [n] \to [n]$ is called a **permutation** on n elements if f is a bijection. For example, $f, g : [2] \to [2]$ defined by $f(1) = 1, f(2) = 2$ and $g(1) = 2, g(2) = 1$ are permutations.

Exercise **9.2.5.** *Let $\mathcal{S}_3$ be the set consisting of all permutation on 3 elements. Then, prove that $\mathcal{S}_3$ has 6 elements. Moreover, they are one of the 6 functions given below.*

1. $f_1(1) = 1, f_1(2) = 2$ *and* $f_1(3) = 3$.
2. $f_2(1) = 1, f_2(2) = 3$ *and* $f_2(3) = 2$.
3. $f_3(1) = 2, f_3(2) = 1$ *and* $f_3(3) = 3$.
4. $f_4(1) = 2, f_4(2) = 3$ *and* $f_4(3) = 1$.
5. $f_5(1) = 3, f_5(2) = 1$ *and* $f_5(3) = 2$.
6. $f_6(1) = 3, f_6(2) = 2$ *and* $f_6(3) = 1$.

Remark 9.2.6. *Let $f : [n] \to [n]$ be a bijection. Then, the* **inverse** *of f, denote f^{-1}, is defined by $f^{-1}(m) = \ell$ whenever $f(\ell) = m$ for $m \in [n]$ is well defined and f^{-1} is a bijection. For example, in Exercise 9.2.5, note that $f_i^{-1} = f_i$, for $i = 1, 2, 3, 6$ and $f_4^{-1} = f_5$.*

Remark 9.2.7. *Let $\mathcal{S}_n = \{f : [n] \to [n] \ : \ \sigma \text{ is a permutation}\}$. Then, $\mathcal{S}_n$ has $n!$ elements and forms a* **group** *with respect to composition of functions, called* **product**, *due to the following.*

1. *Let $f \in \mathbf{S}_n$. Then,*
   (a) *f can be written as* $f = \begin{pmatrix} 1 & 2 & \cdots & n \\ f(1) & f(2) & \cdots & f(n) \end{pmatrix}$, *called a* **two row notation**.
   (b) *f is one-one. Hence, $\{f(1), f(2), \ldots, f(n)\} = [n]$ and thus, $f(1) \in [n], f(2) \in [n] \setminus \{f(1)\}, \ldots$ and finally $f(n) = [n] \setminus \{f(1), \ldots, f(n-1)\}$. Therefore, there are n choices for $f(1)$, $n-1$ choices for $f(2)$ and so on. Hence, the number of elements in $\mathcal{S}_n$ equals $n(n-1)\cdots 2 \cdot 1 = n!$.*
2. *By Remark 9.2.3, $f \circ g \in \mathcal{S}_n$, for any $f, g \in \mathbf{S}_n$.*
3. *Also associativity holds as $f \circ (g \circ h) = (f \circ g) \circ h$ for all functions f, g and h.*
4. *$\mathcal{S}_n$ has a special permutation called the* **identity** *permutation, denoted Id_n, such that $Id_n(i) = i$, for $1 \le i \le n$.*
5. *For each $f \in \mathcal{S}_n$, $f^{-1} \in \mathcal{S}_n$ and is called the* **inverse** *of f as $f \circ f^{-1} = f^{-1} \circ f = Id_n$.*

Lemma 9.2.8. *Fix a positive integer n. Then, the group $\mathcal{S}_n$ satisfies the following:*

1. *Fix an element $f \in \mathcal{S}_n$. Then, $\mathcal{S}_n = \{f \circ g : g \in \mathcal{S}_n\} = \{g \circ f : g \in \mathcal{S}_n\}$.*
2. *$\mathcal{S}_n = \{g^{-1} : g \in \mathcal{S}_n\}$.*

Proof. Part 1: Note that for each $\alpha \in \mathcal{S}_n$ the functions $f^{-1} \circ \alpha, \alpha \circ f^{-1} \in \mathcal{S}_n$ and $\alpha = f \circ (f^{-1} \circ \alpha)$ as well as $\alpha = (\alpha \circ f^{-1}) \circ f$.

Part 2: Note that for each $f \in \mathcal{S}_n$, by definition, $(f^{-1})^{-1} = f$. Hence the result holds. □

Definition 9.2.9. Let $f \in \mathcal{S}_n$. Then, the number of inversions of f, denoted $n(f)$, equals

$$\begin{aligned} n(f) &= \mid \{(i,j) : \ i < j, \ f(i) > f(j) \ \} \mid \\ &= \mid \{j : i + 1 \le j \le n, f(j) < f(i)\} \mid \ \text{using two row notation.} \end{aligned} \tag{9.2.1}$$

Example 9.2.10. 1. For $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$, $n(f) = \ \mid \{(1,2), (1,3), (2,3)\} \mid \ = 3$.

2. In Exercise 9.2.5, $n(f_5) = 2 + 0 = 2$.

3. Let $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 2 & 3 & 5 & 1 & 9 & 8 & 7 & 6 \end{pmatrix}$. Then, $n(f) = 3+1+1+1+0+3+2+1 = 12$.

Definition 9.2.11. [Cycle Notation] Let $f \in \mathcal{S}_n$. Suppose there exist $r, 2 \le r \le n$ and $i_1, \ldots, i_r \in [n]$ such that $f(i_1) = i_2, f(i_2) = i_3, \ldots, f(i_r) = i_1$ and $f(j) = j$ for all $j \ne i_1, \ldots, i_r$. Then, we represent such a permutation by $f = (i_1, i_2, \ldots, i_r)$ and call it an **r-cycle**. For example, $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 5 & 1 \end{pmatrix} = (1,4,5)$ and $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 2 & 4 & 5 \end{pmatrix} = (2,3)$.

Remark 9.2.12. 1. *One also write the r-cycle $(i_1, i_2, \ldots, i_r)$ as $(i_2, i_3, \ldots, i_r, i_1)$ and so on. For example, $(1,4,5) = (4,5,1) = (5,1,4)$.*

2. *The permutation* $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 5 & 1 \end{pmatrix}$ *is not a cycle.*

3. *Let* $f = (1,3,5,4)$ *and* $g = (2,4,1)$ *be two cycles. Then, their product, denoted* $f \circ g$ *or* $(1,3,5,4)(2,4,1)$ *equals* $(1,2)(3,5,4)$. *The calculation proceeds as (the arrows indicate the images):*
   $1 \to 2$. *Note* $(f \circ g)(1) = f(g(1)) = f(2) = 2$.
   $2 \to 4 \to 1$ *as* $(f \circ g)(2) = f(g(2)) = f(4) = 1$. *So,* $(1,2)$ *forms a cycle.*
   $3 \to 5$ *as* $(f \circ g)(3) = f(g(3)) = f(3) = 5$.
   $5 \to 4$ *as* $(f \circ g)(5) = f(g(5)) = f(5) = 4$.
   $4 \to 1 \to 3$ *as* $(f \circ g)(4) = f(g(4)) = f(1) = 3$. *So, the other cycle is* $(3,5,4)$.
4. *Let* $f = (1,4,5)$ *and* $g = (2,4,1)$ *be two permutations. Then,* $(1,4,5)(2,4,1) = (1,2,5)(4) = (1,2,5)$ *as* $1 \to 2, 2 \to 4 \to 5, 5 \to 1, 4 \to 1 \to 4$ *and*
   $(2,4,1)(1,4,5) = (1)(2,4,5) = (2,4,5)$ *as* $1 \to 4 \to 1, 2 \to 4, 4 \to 5, 5 \to 1 \to 2$.
5. *Even though* $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 5 & 1 \end{pmatrix}$ *is not a cycle, verify that it is a product of the cycles* $(1,4,5)$ *and* $(2,3)$.

Definition 9.2.13. A permutation $f \in \mathcal{S}_n$ is called a **transposition** if there exist $m, r \in [n]$ such that $f = (m, r)$.

Remark 9.2.14. *Verify that*

1. $(2,4,5) = (2,5)(2,4) = (4,2)(4,5) = (5,4)(5,2) = (1,2)(1,5)(1,4)(1,2)$.
2. *in general, the* r*-cycle* $(i_1, \ldots, i_r) = (1, i_1)(1, i_r)(1, i_{r-1}) \cdots (1, i_2)(1, i_1)$.
3. *So, every* r*-cycle can be written as product of transpositions. Furthermore, they can be written using the* n *transpositions* $(1,2), (1,3), \ldots, (1,n)$.

With the above definitions, we state and prove two important results.

Theorem 9.2.15. *Let* $f \in \mathcal{S}_n$. *Then,* f *can be written as product of transpositions.*

Proof. Note that using use Remark 9.2.14, we just need to show that f can be written as product of disjoint cycles.

Consider the set $S = \{1, f(1), f^{(2)}(1) = (f \circ f)(1), f^{(3)}(1) = (f \circ (f \circ f))(1), \ldots\}$. As S is an infinite set and each $f^{(i)}(1) \in [n]$, there exist i, j with $0 \le i < j \le n$ such that $f^{(i)}(1) = f^{(j)}(1)$. Now, let j_1 be the least positive integer such that $f^{(i)}(1) = f^{(j_1)}(1)$, for some i, $0 \le i < j_1$. Then, we claim that $i = 0$.

For if, $i - 1 \ge 0$ then $j_1 - 1 \ge 1$ and the condition that f is one-one gives

$$f^{(i-1)}(1) = (f^{-1} \circ f^{(i)})(1) = f^{-1}\left(f^{(i)}(1)\right) = f^{-1}\left(f^{(j_1)}(1)\right) = (f^{-1} \circ f^{(j_1)})(1) = f^{(j_1-1)}(1).$$

Thus, we see that the repetition has occurred at the $(j_1 - 1)$-th instant, contradicting the assumption that j_1 was the least such positive integer. Hence, we conclude that $i = 0$. Thus, $(1, f(1), f^{(2)}(1), \ldots, f^{(j_1-1)}(1))$ is one of the cycles in f.

Now, choose $i_1 \in [n] \setminus \{1, f(1), f^{(2)}(1), \ldots, f^{(j_1-1)}(1)\}$ and proceed as above to get another cycle. Let the new cycle by $(i_1, f(i_1), \ldots, f^{(j_2-1)}(i_1))$. Then, using f is one-one follows that

$$\left\{1, f(1), f^{(2)}(1), \ldots, f^{(j_1-1)}(1)\right\} \cap \left\{i_1, f(i_1), \ldots, f^{(j_2-1)}(i_1)\right\} = \emptyset.$$

So, the above process needs to be repeated at most n times to get all the disjoint cycles. Thus, the required result follows. $\square$

Remark 9.2.16. *Note that when one writes a permutation as product of disjoint cycles, cycles of length* 1 *are suppressed so as to match Definition 9.2.11. For example, the algorithm in the proof of Theorem 9.2.15 implies*

1. *Using Remark 9.2.14.3, we see that every permutation can be written as product of the* n *transpositions* $(1,2),(1,3),\ldots,(1,n)$.
2. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 3 & 5 & 2 \end{pmatrix} = (1)(2,4,5)(3) = (2,4,5)$.
3. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 2 & 3 & 5 & 1 & 9 & 8 & 7 & 6 \end{pmatrix} = (1,4,5)(2)(3)(6,9)(7,8) = (1,4,5)(6,9)(7,8)$.

Note that $Id_3 = (1,2)(1,2) = (1,2)(2,3)(1,2)(1,3)$, as well. The question arises, is it possible to write Id_n as a product of odd number of transpositions? The next lemma answers this question in negative.

Lemma 9.2.17. *Suppose there exist transpositions* f_i, $1 \le i \le t$, *such that*

$$Id_n = f_1 \circ f_2 \circ \cdots \circ f_t,$$

then t *is even.*

Proof. We will prove the result by mathematical induction. Observe that $t \neq 1$ as Id_n is not a transposition. Hence, $t \geq 2$. If $t = 2$, we are done. So, let us assume that the result holds for all expressions in which the number of transpositions $t \le k$. Now, let $t = k+1$.

Suppose $f_1 = (m,r)$ and let $\ell, s \in [n] \setminus \{m,r\}$. Then, the possible choices for the composition $f_1 \circ f_2$ are $(m,r)(m,r) = Id_n$, $(m,r)(m,\ell) = (r,\ell)(r,m)$, $(m,r)(r,\ell) = (\ell,r)(\ell,m)$ and $(m,r)(\ell,s) = (\ell,s)(m,r)$. In the first case, f_1 and f_2 can be removed to obtain $Id_n = f_3 \circ f_4 \circ \cdots \circ f_t$, where the number of transpositions is $t-2 = k-1 < k$. So, by mathematical induction, $t-2$ is even and hence t is also even.

In the remaining cases, the expression for $f_1 \circ f_2$ is replaced by their counterparts to obtain another expression for Id_n. But in the new expression for Id_n, m doesn't appear in the first transposition, but appears in the second transposition. The shifting of m to the right can continue till the number of transpositions reduces by 2 (which in turn gives the result by mathematical induction). For if, the shifting of m to the right doesn't reduce the number of transpositions then m will get shifted to the right and will appear only in the right most transposition. Then, this expression for Id_n does not fix m whereas $Id_n(m) = m$. So, the later case leads us to a contradiction. Hence, the shifting of m to the right will surely lead to an expression in which the number of transpositions at some stage is $t-2 = k-1$. At this stage, one applies mathematical induction to get the required result. □

Theorem 9.2.18. *Let* $f \in \mathcal{S}_n$. *If there exist transpositions* $g_1, \ldots, g_k$ *and* $h_1, \ldots, h_\ell$ *with*

$$f = g_1 \circ g_2 \circ \cdots \circ g_k = h_1 \circ h_2 \circ \cdots \circ h_\ell$$

then, either k *and* ℓ *are both even or both odd.*

Proof. As $g_1 \circ \cdots \circ g_k = h_1 \circ \cdots \circ h_\ell$ and $h^{-1} = h$ for any transposition $h \in \mathcal{S}_n$, we have

$$Id_n = g_1 \circ g_2 \circ \cdots \circ g_k \circ h_\ell \circ h_{\ell-1} \circ \cdots \circ h_1.$$

Hence by Lemma 9.2.17, $k + \ell$ is even. Thus, either k and ℓ are both even or both odd. □

Definition 9.2.19. [Even and Odd Permutation] A permutation $f \in \mathcal{S}_n$ is called an

1. **even permutation** if f can be written as product of even number of transpositions.
2. **odd permutation** if f can be written as a product of odd number of transpositions.

Definition 9.2.20. Observe that if f and g are both even or both odd permutations, then $f \circ g$ and $g \circ f$ are both even. Whereas, if one of them is odd and the other even then $f \circ g$ and $g \circ f$ are both odd. We use this to define a function sgn : $\mathcal{S}_n \to \{1, -1\}$, called the **signature** of a permutation, by

$$\operatorname{sgn}(f) = \begin{cases} 1 & \text{if } f \text{ is an even permutation} \\ -1 & \text{if } f \text{ is an odd permutation} \end{cases}.$$

Example 9.2.21. Consider the set $\mathcal{S}_n$. Then,

1. by Lemma 9.2.17, Id_n is an even permutation and $\operatorname{sgn}(Id_n) = 1$.
2. a transposition, say f, is an odd permutation and hence $\operatorname{sgn}(f) = -1$
3. using Remark 9.2.20, $\operatorname{sgn}(f \circ g) = \operatorname{sgn}(f) \cdot \operatorname{sgn}(g)$ for any two permutations $f, g \in \mathcal{S}_n$.

We are now ready to define determinant of a square matrix A.

Definition 9.2.22. Let $A = [a_{ij}]$ be an $n \times n$ matrix with complex entries. Then, the **determinant** of A, denoted $\det(A)$, is defined as

$$\det(A) = \sum_{g \in \mathcal{S}_n} \operatorname{sgn}(g) a_{1g(1)} a_{2g(2)} \ldots a_{ng(n)} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(g) \prod_{i=1}^{n} a_{ig(i)}. \tag{9.2.2}$$

For example, if $\mathcal{S}_2 = \{Id, f = (1,2)\}$ then for $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, $\det(A) = \operatorname{sgn}(Id) \cdot a_{1Id(1)} a_{2Id(2)} + \operatorname{sgn}(f) \cdot a_{1f(1)} a_{2f(2)} = 1 \cdot a_{11} a_{22} + (-1) a_{12} a_{21} = 1 - 4 = -3.$

Observe that $\det(A)$ is a scalar quantity. Even though the expression for $\det(A)$ seems complicated at first glance, it is very helpful in proving the results related with "properties of determinant". We will do so in the next section. As another examples, we verify that this definition also matches for 3×3 matrices. So, let $A = [a_{ij}]$ be a 3×3 matrix. Then, using Equation (9.2.2),

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{3} a_{i\sigma(i)} \\
&= \operatorname{sgn}(f_1) \prod_{i=1}^{3} a_{if_1(i)} + \operatorname{sgn}(f_2) \prod_{i=1}^{3} a_{if_2(i)} + \operatorname{sgn}(f_3) \prod_{i=1}^{3} a_{if_3(i)} + \\
&\qquad \operatorname{sgn}(f_4) \prod_{i=1}^{3} a_{if_4(i)} + \operatorname{sgn}(f_5) \prod_{i=1}^{3} a_{if_5(i)} + \operatorname{sgn}(f_6) \prod_{i=1}^{3} a_{if_6(i)} \\
&= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.
\end{aligned}$$

9.3 Properties of Determinant

Theorem 9.3.1 (Properties of Determinant)**.** *Let $A = [a_{ij}]$ be an $n \times n$ matrix.*

1. *If $A[i,:] = \mathbf{0}^T$ for some i then* $\det(A) = 0$.

2. *If $B = E_i(c)A$, for some $c \neq 0$ and $i \in [n]$ then* $\det(B) = c\det(A)$.

3. *If $B = E_{ij}A$, for some $i \neq j$ then* $\det(B) = -\det(A)$.

4. *If $A[i,:] = A[j,:]$ for some $i \neq j$ then* $\det(A) = 0$.

5. *Let B and C be two $n \times n$ matrices. If there exists $m \in [n]$ such that $B[i,:] = C[i,:] = A[i,:]$ for all $i \neq m$ and $C[m,:] = A[m,:] + B[m,:]$ then* $\det(C) = \det(A) + \det(B)$.

6. *If $B = E_{ij}(c)$, for $c \neq 0$ then* $\det(B) = \det(A)$.

7. *If A is a triangular matrix then* $\det(A) = a_{11} \cdots a_{nn}$, *the product of the diagonal entries.*

8. *If E is an $n \times n$ elementary matrix then* $\det(EA) = \det(E)\det(A)$.

9. *A is invertible if and only if* $\det(A) \neq 0$.

10. *If B is an $n \times n$ matrix then* $\det(AB) = \det(A)\det(B)$.

11. *If A^T denotes the transpose of the matrix A then* $\det(A) = \det(A^T)$.

Proof. Part 1: Note that each sum in $\det(A)$ contains one entry from each row. So, each sum has an entry from $A[i,:] = \mathbf{0}^T$. Hence, each sum in itself is zero. Thus, $\det(A) = 0$.
Part 2: By assumption, $B[k,:] = A[k,:]$ for $k \neq i$ and $B[i,:] = cA[i,:]$. So,

$$\begin{aligned} \det(B) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{k \neq i} b_{k\sigma(k)} \right) b_{i\sigma(i)} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{k \neq i} a_{k\sigma(k)} \right) c a_{i\sigma(i)} \\ &= c \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} a_{k\sigma(k)} = c\det(A). \end{aligned}$$

Part 3: Let $\tau = (i, j)$. Then, $\operatorname{sgn}(\tau) = -1$, by Lemma 9.2.8, $\mathcal{S}_n = \{\sigma \circ \tau : \sigma \in \mathcal{S}_n\}$ and

$$\begin{aligned} \det(B) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{i\sigma(i)} = \sum_{\sigma\circ\tau \in \mathcal{S}_n} \operatorname{sgn}(\sigma \circ \tau) \prod_{i=1}^{n} b_{i,(\sigma\circ\tau)(i)} \\ &= \sum_{\sigma\circ\tau \in \mathcal{S}_n} \operatorname{sgn}(\tau) \cdot \operatorname{sgn}(\sigma) \left(\prod_{k \neq i,j} b_{k\sigma(k)} \right) b_{i(\sigma\circ\tau)(i)} b_{j(\sigma\circ\tau)(j)} \\ &= \operatorname{sgn}(\tau) \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{k \neq i,j} b_{k\sigma(k)} \right) b_{i\sigma(j)} b_{j\sigma(i)} = -\sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} a_{k\sigma(k)} \\ &= -\det(A). \end{aligned}$$

Part 4: As $A[i,:] = A[j,:]$, $A = E_{ij}A$. Hence, by Part 3, $\det(A) = -\det(A)$. Thus, $\det(A) = 0$.

Part 5: By assumption, $C[i,:] = B[i,:] = A[i,:]$ for $i \neq m$ and $C[m,:] = B[m,:] + A[m,:]$. So,

$$\begin{aligned}\det(C) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} c_{i\sigma(i)} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{i \neq m} c_{i\sigma(i)} \right) c_{m\sigma(m)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{i \neq m} c_{i\sigma(i)} \right) (a_{m\sigma(m)} + b_{m\sigma(m)}) \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} + \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{i\sigma(i)} = \det(A) + \det(B).\end{aligned}$$

Part 6: By assumption, $B[k,:] = A[k,:]$ for $k \neq i$ and $B[i,:] = A[i,:] + cA[j,:]$. So,

$$\begin{aligned}\det(B) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} b_{k\sigma(k)} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{k \neq i} b_{k\sigma(k)} \right) b_{i\sigma(i)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{k \neq i} a_{k\sigma(k)} \right) (a_{i\sigma(i)} + c a_{j\sigma(j)}) \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{k \neq i} a_{k\sigma(k)} \right) a_{i\sigma(i)} + c \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \left(\prod_{k \neq i} a_{k\sigma(k)} \right) a_{j\sigma(j)}) \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} a_{k\sigma(k)} + c \cdot 0 = \det(A). \qquad UsePart\ 4\end{aligned}$$

Part 7: Observe that if $\sigma \in \mathcal{S}_n$ and $\sigma \neq Id_n$ then $n(\sigma) \geq 1$. Thus, for every $\sigma \neq Id_n$, there exists $m \in [n]$ (depending on σ) such that $m > \sigma(m)$ or $m < \sigma(m)$. So, if A is triangular, $a_{m\sigma(m)} = 0$. So, for each $\sigma \neq Id_n$, $\prod_{i=1}^{n} a_{i\sigma(i)} = 0$. Hence, $\det(A) = \prod_{i=1}^{n} a_{ii}$. the result follows.
Part 8: Using Part 7, $\det(I_n) = 1$. By definition $E_{ij} = E_{ij} I_n$ and $E_i(c) = E_i(c) I_n$ and $E_{ij}(c) = E_{ij}(c) I_n$, for $c \neq 0$. Thus, using Parts 2, 3 and 6, we get $\det(E_i(c)) = c, \det(E_{ij}) = -1$ and $\det(E_{ij}(k)) = 1$. Also, again using Parts 2, 3 and 6, we get $\det(EA) = \det(E)\det(A)$.
Part 9: Suppose A is invertible. Then, by Theorem 2.3.1, $A = E_1 \cdots E_k$, for some elementary matrices $E_1, \ldots, E_k$. So, a repeated application of Part 8 implies $\det(A) = \det(E_1) \cdots \det(E_k) \neq 0$ as $\det(E_i) \neq 0$ for $1 \leq i \leq k$.

Now, suppose that $\det(A) \neq 0$. We need to show that A is invertible. On the contrary, assume that A is not invertible. Then, by Theorem 2.3.1, $\text{Rank}(A) < n$. So, by Proposition 2.2.21, there exist elementary matrices $E_1, \ldots, E_k$ such that $E_1 \cdots E_k A = \begin{bmatrix} B \\ \mathbf{0} \end{bmatrix}$. Therefore, by Part 1 and a repeated application of Part 8 gives

$$\det(E_1) \cdots \det(E_k) \det(A) = \det(E_1 \cdots E_k A) = \det \left(\begin{bmatrix} B \\ \mathbf{0} \end{bmatrix} \right) = 0.$$

As $\det(E_i) \neq 0$, for $1 \leq i \leq k$, we have $\det(A) = 0$, a contradiction. Thus, A is invertible.
Part 10: Let A be invertible. Then, by Theorem 2.3.1, $A = E_1 \cdots E_k$, for some elementary matrices $E_1, \ldots, E_k$. So, applying Part 8 repeatedly gives $\det(A) = \det(E_1) \cdots \det(E_k)$ and

$$\det(AB) = \det(E_1 \cdots E_k B) = \det(E_1) \cdots \det(E_k) \det(B) = \det(A) \det(B).$$

In case A is not invertible, by Part 9, $\det(A) = 0$. Also, AB is not invertible (AB is invertible will imply A is invertible using the rank argument). So, again by Part 9, $\det(AB) = 0$. Thus, $\det(AB) = \det(A)\det(B)$.

Part 11: Let $B = [b_{ij}] = A^T$. Then, $b_{ij} = a_{ji}$, for $1 \le i, j \le n$. By Lemma 9.2.8, we know that $\mathcal{S}_n = \{\sigma^{-1} : \ \sigma \in \mathcal{S}_n\}$. As $\sigma \circ \sigma^{-1} = Id_n$, $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$. Hence,

$$\begin{aligned}\det(B) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{i\sigma(i)} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i} = \sum_{\sigma^{-1} \in \mathcal{S}_n} \operatorname{sgn}(\sigma^{-1}) \prod_{i=1}^{n} a_{i\sigma^{-1}(i)} \\ &= \det(A).\end{aligned}$$

□

Remark 9.3.2. 1. *As* $\det(A) = \det(A^T)$, *we observe that in Theorem 9.3.1, the condition on "row" can be replaced by the condition on "column".*

2. *Let* $A = [a_{ij}]$ *be a matrix satisfying* $a_{1j} = 0$, *for* $2 \le j \le n$. *Let* $B = A(1 \mid 1)$, *the submatrix of A obtained by removing the first row and the first column. Then* $\det(A) = a_{11}\det(B)$.
   Proof: *Let* $\sigma \in \mathcal{S}_n$ *with* $\sigma(1) = 1$. *Then,* σ *has a cycle* (1). *So, a disjoint cycle representation of* σ *only has numbers* $\{2, 3, \ldots, n\}$. *That is, we can think of* σ *as an element of* $\mathcal{S}_{n-1}$. *Hence,*

$$\begin{aligned}\det(A) &= \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma \in \mathcal{S}_n, \sigma(1)=1} sgn(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \\ &= a_{11} \sum_{\sigma \in \mathcal{S}_n, \sigma(1)=1} sgn(\sigma) \prod_{i=2}^{n} a_{i\sigma(i)} = a_{11} \sum_{\sigma \in \mathcal{S}_{n-1}} sgn(\sigma) \prod_{i=1}^{n-1} b_{i\sigma(i)} = a_{11}\det(B).\end{aligned}$$

We now relate this definition of determinant with the one given in Definition 2.3.6.

Theorem 9.3.3. *Let A be an $n \times n$ matrix. Then,* $\det(A) = \sum\limits_{j=1}^{n} (-1)^{1+j} a_{1j} \det\big(A(1 \mid j)\big)$, *where recall that* $A(1 \mid j)$ *is the submatrix of A obtained by removing the* 1^{st} *row and the* j^{th} *column.*

Proof. For $1 \le j \le n$, define an $n \times n$ matrix $B_j = \begin{bmatrix} 0 & 0 & \cdots & a_{1j} & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}$. Also, for each matrix B_j, we define the $n \times n$ matrix C_j by

1. $C_j[:, 1] = B_j[:, j]$,
2. $C_j[:, i] = B_j[:, i-1]$, for $2 \le i \le j$ and
3. $C_j[:, k] = B_j[:, k]$ for $k \ge j + 1$.

Also, observe that B_j's have been defined to satisfy $B_1[1, :] + \cdots + B_n[1, :] = A[1, :]$ and $B_j[i, :] = A[i, :]$ for all $i \ge 2$ and $1 \le j \le n$. Thus, by Theorem 9.3.1.5,

$$\det(A) = \sum_{j=1}^{n} \det(B_j). \tag{9.3.1}$$

Let us now compute $\det(B_j)$, for $1 \le j \le n$. Note that $C_j = E_{12}E_{23}\cdots E_{j-1,j}B_j$, for $1 \le j \le n$. Then, by Theorem 9.3.1.3, we get $\det(B_j) = (-1)^{j-1}\det(C_j)$. So, using Remark 9.3.2.2 and Theorem 9.3.1.2 and Equation (9.3.1), we have

$$\det(A) = \sum_{j=1}^{n}(-1)^{j-1}\det(C_j) = \sum_{j=1}^{n}(-1)^{j+1}a_{1j}\det\big(A(1 \mid j)\big).$$

Thus, we have shown that the determinant defined in Definition 2.3.6 is valid. □

9.4 Dimension of $\mathbb{W}_1 + \mathbb{W}_2$

Theorem 9.4.1. *Let $\mathbb{V}$ be a finite dimensional vector space over $\mathbb{F}$ and let $\mathbb{W}_1$ and $\mathbb{W}_2$ be two subspaces of $\mathbb{V}$. Then,*

$$\dim(\mathbb{W}_1) + \dim(\mathbb{W}_2) = \dim(\mathbb{W}_1 + \mathbb{W}_2) + \dim(\mathbb{W}_1 \cap \mathbb{W}_2). \tag{9.4.1}$$

Proof. Since $\mathbb{W}_1 \cap \mathbb{W}_2$ is a vector subspace of V, let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ be a basis of $\mathbb{W}_1 \cap \mathbb{W}_2$. As, $\mathbb{W}_1 \cap \mathbb{W}_2$ is a subspace of both $\mathbb{W}_1$ and $\mathbb{W}_2$, let us extend the basis $\mathcal{B}$ to form a basis $\mathcal{B}_1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{v}_1, \ldots, \mathbf{v}_s\}$ of $\mathbb{W}_1$ and a basis $\mathcal{B}_2 = \{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{w}_1, \ldots, \mathbf{w}_t\}$ of $\mathbb{W}_2$.

We now prove that $\mathcal{D} = \{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{w}_1, \ldots, \mathbf{w}_s, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t\}$ is a basis of $\mathbb{W}_1 + \mathbb{W}_2$. To do this, we show that

1. $\mathcal{D}$ is linearly independent subset of $\mathbb{V}$ and
2. $LS(\mathcal{D}) = \mathbb{W}_1 + \mathbb{W}_2$.

The second part can be easily verified. For the first part, consider the linear system

$$\alpha_1\mathbf{u}_1 + \cdots + \alpha_r\mathbf{u}_r + \beta_1\mathbf{w}_1 + \cdots + \beta_s\mathbf{w}_s + \gamma_1\mathbf{v}_1 + \cdots + \gamma_t\mathbf{v}_t = \mathbf{0} \tag{9.4.2}$$

in the variables α_i's, β_j's and γ_k's. We re-write the system as

$$\alpha_1\mathbf{u}_1 + \cdots + \alpha_r\mathbf{u}_r + \beta_1\mathbf{w}_1 + \cdots + \beta_s\mathbf{w}_s = -(\gamma_1\mathbf{v}_1 + \cdots + \gamma_t\mathbf{v}_t).$$

Then, $\mathbf{v} = -\sum\limits_{i=1}^{s}\gamma_i\mathbf{v}_i \in LS(\mathcal{B}_1) = \mathbb{W}_1$. Also, $\mathbf{v} = \sum\limits_{j=1}^{r}\alpha_r\mathbf{u}_r + \sum\limits_{k=1}^{T}\beta_k\mathbf{w}_k$. So, $\mathbf{v} \in LS(\mathcal{B}_2) = \mathbb{W}_2$. Hence, $\mathbf{v} \in \mathbb{W}_1 \cap \mathbb{W}_2$ and therefore, there exists scalars $\delta_1, \ldots, \delta_k$ such that $\mathbf{v} = \sum\limits_{j=1}^{r}\delta_j\mathbf{u}_j$.

Substituting this representation of $\mathbf{v}$ in Equation (9.4.2), we get

$$(\alpha_1 - \delta_1)\mathbf{u}_1 + \cdots + (\alpha_r - \delta_r)\mathbf{u}_r + \beta_1\mathbf{w}_1 + \cdots + \beta_t\mathbf{w}_t = \mathbf{0}.$$

So, using Exercise 3.4.16.1, $\alpha_i - \delta_i = 0$, for $1 \le i \le r$ and $\beta_j = 0$, for $1 \le j \le t$. Thus, the system (9.4.2) reduces to

$$\alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k + \gamma_1\mathbf{v}_1 + \cdots + \gamma_r\mathbf{v}_r = \mathbf{0}$$

which has $\alpha_i = 0$ for $1 \le i \le r$ and $\gamma_j = 0$ for $1 \le j \le s$ as the only solution. Hence, we see that the linear system of Equations (9.4.2) has no nonzero solution. Therefore, the set $\mathcal{D}$ is linearly independent and the set $\mathcal{D}$ is indeed a basis of $\mathbb{W}_1 + \mathbb{W}_2$. We now count the vectors in the sets $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2$ and $\mathcal{D}$ to get the required result. □

9.5 When does Norm imply Inner Product

In this section, we prove the following result. A generalization of this result to complex vector space is left as an exercise for the reader as it requires similar ideas.

Theorem 9.5.1. *Let $\mathbb{V}$ be a real vector space. A norm $\|\cdot\|$ is induced by an inner product if and only if, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$, the norm satisfies*

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 \qquad \text{(PARALLELOGRAM LAW)}. \tag{9.5.1}$$

Proof. Suppose that $\|\cdot\|$ is indeed induced by an inner product. Then, by Exercise 5.1.7.3 the result follows.

So, let us assume that $\|\cdot\|$ satisfies the parallelogram law. So, we need to define an inner product. We claim that the function $f : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ defined by

$$f(\mathbf{x}, \mathbf{y}) = \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right), \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{V}$$

satisfies the required conditions for an inner product. So, let us proceed to do so.

STEP 1: Clearly, for each $\mathbf{x} \in \mathbb{V}$, $f(\mathbf{x}, \mathbf{0}) = 0$ and $f(\mathbf{x}, \mathbf{x}) = \frac{1}{4}\|\mathbf{x}+\mathbf{x}\|^2 = \|\mathbf{x}\|^2$. Thus, $f(\mathbf{x}, \mathbf{x}) \geq 0$. Further, $f(\mathbf{x}, \mathbf{x}) = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

STEP 2: By definition $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$.

STEP 3: Now note that $\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 = 2\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right)$. Or equivalently,

$$2f(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2, \text{ for } \mathbf{x}, \mathbf{y} \in \mathbb{V}. \tag{9.5.2}$$

Thus, for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{V}$, we have

$$\begin{aligned} 4\left(f(\mathbf{x}, \mathbf{y}) + f(\mathbf{z}, \mathbf{y})\right) &= \|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 + \|\mathbf{z}+\mathbf{y}\|^2 - \|\mathbf{z}-\mathbf{y}\|^2 \\ &= 2\left(\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{z}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{z}\|^2 - 2\|\mathbf{y}\|^2\right) \\ &= \|\mathbf{x}+\mathbf{z}+2\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{z}\|^2 - \left(\|\mathbf{x}+\mathbf{z}\|^2 + \|\mathbf{x}-\mathbf{z}\|^2\right) - 4\|\mathbf{y}\|^2 \\ &= \|\mathbf{x}+\mathbf{z}+2\mathbf{y}\|^2 - \|\mathbf{x}+\mathbf{z}\|^2 - \|2\mathbf{y}\|^2 \\ &= 2f(\mathbf{x}+\mathbf{z}, 2\mathbf{y}) \text{ using Equation (9.5.2).} \end{aligned} \tag{9.5.3}$$

Now, substituting $\mathbf{z} = \mathbf{0}$ in Equation (9.5.3) and using Equation (9.5.2), we get $2f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}, 2\mathbf{y})$ and hence $4f(\mathbf{x}+\mathbf{z}, \mathbf{y}) = 2f(\mathbf{x}+\mathbf{z}, 2\mathbf{y}) = 4\left(f(\mathbf{x}, \mathbf{y}) + f(\mathbf{z}, \mathbf{y})\right)$. Thus,

$$f(\mathbf{x}+\mathbf{z}, \mathbf{y}) = f(\mathbf{x}, \mathbf{y}) + f(\mathbf{z}, \mathbf{y}), \text{ for all } \mathbf{x}, \mathbf{y} \in \mathbb{V}. \tag{9.5.4}$$

STEP 4: Using Equation (9.5.4), $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$ and the principle of mathematical induction, it follows that $nf(\mathbf{x}, \mathbf{y}) = f(n\mathbf{x}, \mathbf{y})$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ and $n \in \mathbb{N}$. Another application of Equation (9.5.4) with $f(\mathbf{0}, \mathbf{y}) = 0$ implies that $nf(\mathbf{x}, \mathbf{y}) = f(n\mathbf{x}, \mathbf{y})$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ and $n \in \mathbb{Z}$. Also, for $m \neq 0$,

$$mf\left(\frac{n}{m}\,\mathbf{x}, \mathbf{y}\right) = f(m\frac{n}{m}\,\mathbf{x}, \mathbf{y}) = f(n\mathbf{x}, \mathbf{y}) = nf(\mathbf{x}, \mathbf{y}).$$

Hence, we see that for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ and $a \in \mathbb{Q}$, $f\left(a\mathbf{x}, \mathbf{y}\right) = af(\mathbf{x}, \mathbf{y})$.

STEP 5: Fix $\mathbf{u}, \mathbf{v} \in \mathbb{V}$ and define a function $g : \mathbb{R} \to \mathbb{R}$ by

$$\begin{aligned} g(x) &= f(x\mathbf{u}, \mathbf{v}) - xf(\mathbf{u}, \mathbf{v}) \\ &= \frac{1}{2}\left(\|x\mathbf{u}+\mathbf{v}\|^2 - \|x\mathbf{u}\|^2 - \|\mathbf{v}\|^2\right) - \frac{x}{2}\left(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2\right). \end{aligned}$$

Then, by previous step $g(x) = 0$, for all $x \in \mathbb{Q}$. So, if g is a continuous function then continuity implies $g(x) = 0$, for all $x \in \mathbb{R}$. Hence, $f(x\mathbf{u}, \mathbf{v}) = xf(\mathbf{u}, \mathbf{v})$, for all $x \in \mathbb{R}$.

Note that the second term of $g(x)$ is a constant multiple of x and hence continuous. Using a similar reason, it is enough to show that $g_1(x) = \|x\mathbf{u} + \mathbf{v}\|$, for certain fixed vectors $\mathbf{u}, \mathbf{v} \in \mathbb{V}$, is continuous. To do so, note that

$$\|x_1\mathbf{u}+\mathbf{v}\| = \|(x_1 - x_2)\mathbf{u} + x_2\mathbf{u} + \mathbf{v}\| \le \|(x_1 - x_2)\mathbf{u}\| + \|x_2\mathbf{u}+\mathbf{v}\|.$$

Thus, $\Big|\|x_1\mathbf{u}+\mathbf{v}\| - \|x_2\mathbf{u}+\mathbf{v}\|\Big| \le \|(x_1-x_2)\mathbf{u}\|$. Hence, taking the limit as $x_1 \to x_2$, we get $\lim\limits_{x_1 \to x_2} \|x_1\mathbf{u}+\mathbf{v}\| = \|x_2\mathbf{u}+\mathbf{v}\|$.

Thus, we have proved the continuity of g and hence the prove of the required result. ∎

9.6 Roots of a Polynomials

The main aim of this subsection is to prove the continuous dependence of the zeros of a polynomial on its coefficients and to recall Descartes's rule of signs.

Definition 9.6.1. [Jordan Curves] A **curve** in $\mathbb{C}$ is a continuous function $f : [a, b] \to \mathbb{C}$, where $[a, b] \subseteq \mathbb{R}$.

1. If the function f is one-one on $[a, b)$ and also on $(a, b]$, then it is called a **simple curve**.
2. If $f(b) = f(a)$, then it is called a **closed curve**.
3. A closed simple curve is called a **Jordan curve**.
4. The derivative (integral) of a curve $f = u+iv$ is defined component wise. If f' is continuous on $[a, b]$, we say f is a $\mathcal{C}^1$**-curve** (at end points we consider one sided derivatives and continuity).
5. A $\mathcal{C}^1$-curve on $[a, b]$ is called a **smooth curve**, if f' is never zero on (a, b).
6. A piecewise smooth curve is called a **contour**.
7. A positively oriented simple closed curve is called a **simple closed curve** such that while traveling on it the interior of the curve always stays to the left. (Camille Jordan has proved that such a curve always divides the plane into two connected regions, one of which is called the **bounded** region and the other is called the **unbounded** region. The one which is bounded is considered as the interior of the curve.)

We state the famous Rouche Theorem of complex analysis without proof.

Theorem 9.6.2. [Rouche's Theorem] *Let C be a positively oriented simple closed contour. Also, let f and g be two analytic functions on R_C, the union of the interior of C and the curve C itself. Assume also that $|f(x)| > |g(x)|$, for all $x \in C$. Then, f and $f + g$ have the same number of zeros in the interior of C.*

Corollary 9.6.3. **[Alen Alexanderian, The University of Texas at Austin, USA.]** *Let $P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0$ have distinct roots $\lambda_1, \ldots, \lambda_m$ with multiplicities $\alpha_1, \ldots, \alpha_m$, respectively. Take any $\epsilon > 0$ for which the balls $\overline{B_\epsilon(\lambda_i)}$ are disjoint. Then, there exists a $\delta > 0$ such that the polynomial $q(t) = t^n + a'_{n-1}t^{n-1} + \cdots + a'_0$ has exactly α_i roots (counting with multiplicities) in $B_\epsilon(\lambda_i)$, whenever $|a_j - a'_j| < \delta$.*

Proof. For an $\epsilon > 0$ and $1 \le i \le m$, let $C_i = \{z \in \mathbb{C} : |z - \lambda_i| = \epsilon\}$. Now, for each $i, 1 \le i \le m$, take $\nu_i = \min\limits_{z \in C_i} |p(z)|$, $\rho_i = \max\limits_{z \in C_i} [1 + |z| + \cdots + |z|^{n-1}]$ and choose $\delta > 0$ such that $\rho_i \delta < \nu_i$. Then, for a fixed j and $z \in C_j$, we have

$$|q(z) - P(z)| = |(a'_{n-1} - a_{n-1})z^{n-1} + \cdots + (a'_0 - a_0)| \le \delta\rho_j < \nu_j \le |p(z)|.$$

Hence, by Rouche's theorem, $p(z)$ and $q(z)$ have the same number of zeros inside C_j, for each $j = 1, \ldots, m$. That is, the zeros of $q(t)$ are within the ϵ-neighborhood of the zeros of $P(t)$. ∎

As a direct application, we obtain the following corollary.

Corollary 9.6.4. *Eigenvalues of a matrix are continuous functions of its entries.*

Proof. Follows from Corollary 9.6.3. ∎

Remark 9.6.5.

1. **[Sign changes in a polynomial]** *Let $P(x) = \sum_0^n a_i x^{n-i}$ be a real polynomial, with $a_0 \neq 0$. Read the coefficient from the left $a_0, a_1, \ldots$. We say the* SIGN CHANGES OF a_i OCCUR AT $m_1 < m_2 < \cdots < m_k$ *to mean that a_{m_1} is the first after a_0 with sign opposite to a_0; a_{m_2} is the first after a_{m_1} with sign opposite to a_{m_1}; and so on.*

2. **[Descartes' Rule of Signs]** *Let $P(x) = \sum_0^n a_i x^{n-i}$ be a real polynomial. Then, the maximum number of positive roots of $P(x) = 0$ is the number of changes in sign of the coefficients and that the maximum number of negative roots is the number of sign changes in $P(-x) = 0$.*

   Proof. Assume that $a_0, a_1, \cdots, a_n$ has $k > 0$ sign changes. Let $b > 0$. Then, the coefficients of $(x - b)P(x)$ are

   $$a_0, a_1 - ba_0, a_2 - ba_1, \cdots, a_n - ba_{n-1}, -ba_n.$$

   This list has at least $k + 1$ changes of signs. To see this, assume that $a_0 > 0$ and $a_n \neq 0$. Let the sign changes of a_i occur at $m_1 < m_2 < \cdots < m_k$. Then, setting

   $$c_0 = a_0, c_1 = a_{m_1} - ba_{m_1 - 1}, c_2 = a_{m_2} - ba_{m_2 - 1}, \cdots, c_k = a_{m_k} - ba_{m_k - 1}, c_{k+1} = -ba_n,$$

   we see that $c_i > 0$ when i is even and $c_i < 0$, when i is odd. That proves the claim.

   Now, assume that $P(x) = 0$ has k positive roots $b_1, b_2, \cdots, b_k$. Then,

   $$P(x) = (x - b_1)(x - b_2) \cdots (x - b_k)Q(x),$$

   where $Q(x)$ is a real polynomial. By the previous observation, the coefficients of $(x - b_k)Q(x)$ has at least one change of signs, coefficients of $(x - b_{k-1})(x - b_k)Q(x)$ has at least two, and so on. Thus coefficients of $P(x)$ has at least k change of signs. The rest of the proof is similar. ∎

9.7 Variational characterizations of Hermitian Matrices

Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix. Then, by Theorem 6.2.22, we know that all the eigenvalues of A are real. So, we write $\lambda_i(A)$ to mean the i-th smallest eigenvalue of A. That is, the i-th from the left in the list $\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_n(A)$.

Lemma 9.7.1. **[Rayleigh-Ritz Ratio]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix. Then,*

1. $\lambda_1(A)\mathbf{x}^*\mathbf{x} \leq \mathbf{x}^*A\mathbf{x} \leq \lambda_n(A)\mathbf{x}^*\mathbf{x}$*, for each* $\mathbf{x} \in \mathbb{C}^n$.
2. $\lambda_1(A) = \min\limits_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^*A\mathbf{x}}{\mathbf{x}^*\mathbf{x}} = \min\limits_{\|\mathbf{x}\|=1} \mathbf{x}^*A\mathbf{x}$.
3. $\lambda_n(A) = \max\limits_{\mathbf{x}\neq 0} \frac{\mathbf{x}^*A\mathbf{x}}{\mathbf{x}^*\mathbf{x}} = \max\limits_{\|\mathbf{x}\|=1} \mathbf{x}^*A\mathbf{x}$.

Proof. Proof of Part 1: By spectral theorem (see Theorem 6.2.22, there exists a unitary matrix U such that $A = UDU^*$, where $D =$ diag$(\lambda_1(A), \ldots, \lambda_n(A))$ is a real diagonal matrix. Thus, the set $\{U[:,1], \ldots, U[:,n]\}$ is a basis of $\mathbb{C}^n$. Hence, for each $\mathbf{x} \in \mathbb{C}^n$, there exists **Ans** $:_i$'s (scalar) such that $\mathbf{x} = \sum \alpha_i U[:,i]$. So, note that $\mathbf{x}^*\mathbf{x} = |\alpha_i|^2$ and

$$\lambda_1(A)\mathbf{x}^*\mathbf{x} = \lambda_1(A) \sum |\alpha_i|^2 \leq \sum |\alpha_i|^2 \lambda_i(A) = \mathbf{x}^*A\mathbf{x} \leq \lambda_n \sum |\alpha_i|^2 = \lambda_n \mathbf{x}^*\mathbf{x}.$$

For Part 2 and Part 3, take $\mathbf{x} = U[:,1]$ and $\mathbf{x} = U(:,n)$, respectively. ∎

As an immediate corollary, we state the following result.

Corollary 9.7.2. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix and $\alpha = \dfrac{\mathbf{x}^*A\mathbf{x}}{\mathbf{x}^*\mathbf{x}}$. Then, A has an eigenvalue in the interval $(-\infty, \alpha]$ and has an eigenvalue in the interval $[\alpha, \infty)$.*

We now generalize the second and third parts of Theorem 9.7.2.

Proposition 9.7.3. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix with $A = UDU^*$, where U is a unitary matrix and D is a diagonal matrix consisting of the eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Then, for any positive integer $k, 1 \leq k \leq n$,*

$$\lambda_k = \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp U[:,1],\ldots,U[:,k-1]}} \mathbf{x}^*A\mathbf{x} = \max_{\substack{\|\mathbf{x}\|=1 \\ x\perp U[:,n],\ldots,U[:,k+1]}} \mathbf{x}^*A\mathbf{x}.$$

Proof. Let $\mathbf{x} \in \mathbb{C}^n$ such that $\mathbf{x}$ is orthogonal to $U[,1], \ldots, U[:,k-1]$. Then, we can write $\mathbf{x} = \sum\limits_{i=k}^{n} \alpha_i U[:,i]$, for some scalars α_i's. In that case,

$$\lambda_k \mathbf{x}^*\mathbf{x} = \lambda_k \sum_{i=k}^{n} |\alpha_i|^2 \leq \sum_{i=k}^{n} |\alpha_i|^2 \lambda_i = \mathbf{x}^*A\mathbf{x}$$

and the equality occurs for $\mathbf{x} = U[:,k]$. Thus, the required result follows. ∎

Theorem 9.7.4. **[Courant-Fischer]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Then,*

$$\lambda_k = \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}}} \mathbf{x}^*A\mathbf{x} = \min_{\mathbf{w}_n,\ldots,\mathbf{w}_{k+1}} \max_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_n,\ldots,\mathbf{w}_{k+1}}} \mathbf{x}^*A\mathbf{x}.$$

Proof. Let $A = UDU^*$, where U is a unitary matrix and $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Now, choose a set of $k-1$ linearly independent vectors from $\mathbb{C}^n$, say $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}\}$. Then, some of the eigenvectors $\{U[:,1], \ldots, U[:,k-1]\}$ may be an element of $S^\perp$. Thus, using Proposition 9.7.3, we see that

$$\lambda_k = \min_{\substack{\|\mathbf{x}\|=1, \\ \mathbf{x}\perp U[:,1],\ldots,U[:,k-1]}} \mathbf{x}^* A\mathbf{x} \geq \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\in S^\perp}} \mathbf{x}^* A\mathbf{x}.$$

Hence, $\lambda_k \geq \max\limits_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min\limits_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}}} \mathbf{x}^* A\mathbf{x}$, for each choice of $k-1$ linearly independent vectors. But, by Proposition 9.7.3, the equality holds for the linearly independent set $\{U[:,1], \ldots, U[:,k-1]\}$ which proves the first equality. A similar argument gives the second equality and hence the proof is omitted. ∎

Theorem 9.7.5. [Weyl Interlacing Theorem] *Let $A, B \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrices. Then, $\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_n(B)$. In particular, if $B = P^*P$, for some matrix P, then $\lambda_k(A+B) \geq \lambda_k(A)$. In particular, for $\mathbf{z} \in \mathbb{C}^n$, $\lambda_k(A + \mathbf{z}\mathbf{z}^*) \leq \lambda_{k+1}(A)$.*

Proof. As A and B are Hermitian matrices, the matrix $A+B$ is also Hermitian. Hence, by Courant-Fischer theorem and Lemma 9.7.1.1,

$$\begin{aligned}
\lambda_k(A+B) &= \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}}} \mathbf{x}^*(A+B)\mathbf{x} \\
&\leq \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}}} [\mathbf{x}^*A\mathbf{x} + \lambda_n(B)] = \lambda_k(A) + \lambda_n(B)
\end{aligned}$$

and

$$\begin{aligned}
\lambda_k(A+B) &= \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}}} \mathbf{x}^*(A+B)\mathbf{x} \\
&\geq \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}}} [\mathbf{x}^*A\mathbf{x} + \lambda_1(B)] = \lambda_k(A) + \lambda_1(B).
\end{aligned}$$

If $B = P^*P$, then $\lambda_1(B) = \min\limits_{\|\mathbf{x}\|=1} \mathbf{x}^*(P^*P)\mathbf{x} = \min\limits_{\|\mathbf{x}\|=1} \|P\mathbf{x}\|^2 \geq 0$. Thus,

$$\lambda_k(A+B) \geq \lambda_k(A) + \lambda_1(B) \geq \lambda_k(A).$$

In particular, for $\mathbf{z} \in \mathbb{C}^n$, we have

$$\begin{aligned}
\lambda_k(A+\mathbf{z}\mathbf{z}^*) &= \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}}} [\mathbf{x}^*A\mathbf{x} + |\mathbf{x}^*\mathbf{z}|^2] \\
&\leq \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1},\mathbf{z}}} [\mathbf{x}^*A\mathbf{x} + |\mathbf{x}^*\mathbf{z}|^2] \\
&= \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \min_{\substack{\|x\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1},\mathbf{z}}} \mathbf{x}^*A\mathbf{x} \\
&\leq \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1},\mathbf{w}_k} \min_{\substack{\|x\|=1 \\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{k-1},\mathbf{w}_k}} \mathbf{x}^*A\mathbf{x} = \lambda_{k+1}(A).
\end{aligned}$$

∎

Theorem 9.7.6. [Cauchy Interlacing Theorem] *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix. Define $\hat{A} = \begin{bmatrix} A & \mathbf{y} \\ \mathbf{y}^* & a \end{bmatrix}$, for some $a \in \mathbb{R}$ and $\mathbf{y} \in \mathbb{C}^n$. Then,*

$$\lambda_k(\hat{A}) \leq \lambda_k(A) \leq \lambda_{k+1}(\hat{A}).$$

Proof. Note that

$$\begin{aligned}\lambda_{k+1}(\hat{A}) &= \max_{\mathbf{w}_1,\ldots,\mathbf{w}_k\in\mathbb{C}^{n+1}} \min_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_k}} \mathbf{x}^*\hat{A}\mathbf{x} \le \max_{\mathbf{w}_1,\ldots,\mathbf{w}_k\in\mathbb{C}^{n+1}} \min_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_k\\ \mathbf{x}_{n+1}=0}} \mathbf{x}^*\hat{A}\mathbf{x}\\ &= \max_{\mathbf{w}_1,\ldots,\mathbf{w}_k\in\mathbb{C}^{n}} \min_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_k}} \mathbf{x}^*A\mathbf{x} = \lambda_{k+1}(A)\end{aligned}$$

and

$$\begin{aligned}\lambda_{k+1}(\hat{A}) &= \min_{\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}\in\mathbb{C}^{n+1}} \max_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}}} \mathbf{x}^*\hat{A}\mathbf{x} \ge \min_{\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}\in\mathbb{C}^{n+1}} \max_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}\\ \mathbf{x}_{n+1}=0}} \mathbf{x}^*\hat{A}\mathbf{x}\\ &= \min_{\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}\in\mathbb{C}^{n}} \max_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}}} \mathbf{x}^*A\mathbf{x} = \lambda_{k}(A)\end{aligned}$$

■

As an immediate corollary, one has the following result.

Corollary 9.7.7. **[Inclusion principle]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix and r be a positive integer with $1 \le r \le n$. If $B_{r\times r}$ is a principal submatrix of A then, $\lambda_k(A) \le \lambda_k(B) \le \lambda_{k+n-r}(A)$.*

Theorem 9.7.8. **[Poincare Separation Theorem]** *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix and $\{\mathbf{u}_1,\ldots,\mathbf{u}_r\} \subseteq \mathbb{C}^n$ be an orthonormal set for some positive integer $r, 1 \le r \le n$. If further $B = [b_{ij}]$ is an $r\times r$ matrix with $b_{ij} = \mathbf{u}_i^*A\mathbf{u}_j$, $1 \le i,j \le r$ then, $\lambda_k(A) \le \lambda_k(B) \le \lambda_{k+n-r}(A)$.*

Proof. Let us extend the orthonormal set $\{\mathbf{u}_1,\ldots,\mathbf{u}_r\}$ to an orthonormal basis, say $\{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$ of $\mathbb{C}^n$ and write $U = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n\end{bmatrix}$. Then, B is a $r\times r$ principal submatrix of U^*AU. Thus, by inclusion principle, $\lambda_k(U^*AU) \le \lambda_k(B) \le \lambda_{k+n-r}(U^*AU)$. But, we know that $\sigma(U^*AU) = \sigma(A)$ and hence the required result follows. ■

The proof of the next result is left for the reader.

Corollary 9.7.9. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix and r be a positive integer with $1 \le r \le n$. Then,*

$$\lambda_1(A) + \cdots + \lambda_r(A) = \min_{U^*U=I_r} \mathsf{tr} U^*AU \quad \text{and} \quad \lambda_{n-r+1}(A) + \cdots + \lambda_n(A) = \max_{U^*U=I_r} \mathsf{tr} U^*AU.$$

Corollary 9.7.10. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian matrix and W be a k-dimensional subspace of $\mathbb{C}^n$. Suppose, there exists a real number c such that $\mathbf{x}^*A\mathbf{x} \ge c\mathbf{x}^*\mathbf{x}$, for each $\mathbf{x} \in W$. Then, $\lambda_{n-k+1}(A) \ge c$. In particular, if $\mathbf{x}^*A\mathbf{x} > 0$, for each nonzero $x \in W$, then $\lambda_{n-k+1} > 0$. $\Bigg($Note that, a k-dimensional subspace need not contain an eigenvector of A. For example, the line $y = 2x$ does not contain an eigenvector of $\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}$.$\Bigg)$*

Proof. Let $\{\mathbf{x}_1,\ldots,\mathbf{x}_{n-k}\}$ be a basis of $W^\perp$. Then,

$$\lambda_{n-k+1}(A) = \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}} \min_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}}} \mathbf{x}^*A\mathbf{x} \ge \min_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{x}_1,\ldots,\mathbf{x}_{n-k}}} \mathbf{x}^*A\mathbf{x} \ge c.$$

Now assume that $\mathbf{x}^*A\mathbf{x} > 0$ holds for each nonzero $\mathbf{x} \in W$ and that $\lambda_{n-k+1} = 0$. Then, it follows that $\displaystyle\min_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{x}_1,\ldots,\mathbf{x}_{n-k}}} x^*Ax = 0$. Now, define $f : \mathbb{C}^n \to \mathbb{C}$ by $f(\mathbf{x}) = \mathbf{x}^*A\mathbf{x}$.

Then, f is a continuous function and $\displaystyle\min_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\in W}} f(\mathbf{x}) = 0$. Thus, f must attain its bound on the unit sphere. That is, there exists $\mathbf{y} \in W$ with $\|\mathbf{y}\| = 1$ such that $\mathbf{y}^*A\mathbf{y} = 0$, a contradiction. Thus, the required result follows. ■

Index