Linear Algebra [1]

M.T.Nair
Department of Mathematics, IIT Madras

Contents

Page Number

[1]Lectures for the course MA5310, July-November 2012.

1 Vector Spaces

1.1 Definition and Examples

Definition 1.1. A nonempty set V is said to be **vector space** over a field $\mathbb{F}$ if there are two maps

$$V \times V \to V \quad \text{and} \quad \mathbb{F} \times V \to V$$

denoted by

$$(x, y) \mapsto x + y \quad \text{and} \quad (\alpha, x) \mapsto \alpha x,$$

respectively, called **vector addition** and **scalar multiplication**, respectively, which satisfy the following conditions:

1. $x + y = y + x \quad \forall\, x, y \in V$,
2. $(x + y) + z = x + (y + z) \quad \forall\, x, y, z \in V$,
3. $\exists$ an element, denoted by $0_V \in V$ such that $x + 0_V = x \quad \forall\, x \in V$.
4. $\forall\, x \in V$, $\exists$ an element, denoted by $-x \in V$ such that $x + (-x) = 0_V$.
5. $\alpha(x + y) = \alpha x + \alpha y \quad \forall\, x, y \in X$, $\alpha \in \mathbb{F}$,
6. $(\alpha + \beta)x = \alpha x + \beta x \quad \forall\, \alpha, \beta \in \mathbb{F}$, $x \in V$,
7. $(\alpha\beta)x = \alpha(\beta x) \quad \forall\, \alpha, \beta \in \mathbb{F}$, $x \in V$,
8. $1x = x \quad \forall\, x \in V$.

$\diamondsuit$

Definition 1.2.

(i) Elements of a vector space are called **vectors**.

(ii) Elements of the field $\mathbb{F}$ are called **scalars**.

(iii) The element 0_V is unique, and it is called the **zero vector** in V. (If $u, v \in V$ are such that $x + u = x = x + v$ for all $x \in V$, then $u = u + v = v + u = u$.) The zero vector is usually denoted by 0, which is distinguished from the zero in $\mathbb{F}$ by the context in which it occurs. $\diamondsuit$

- It can be verified that

$$0x = 0_V, \qquad (-1)x = -x \quad \forall\, x \in V.$$

- Condition 1 follows from Conditions 5 and 6: Let $x, v \in V$. By Conditions 6 and 8,

$$2(x+y) = (1+1)(x+y) = 1(x+y) + 1(x+y) = x+y+x+y$$

and by Condition 5 and 6,

$$2(x+y) = 2x + 2y = x + x + y + y.$$

Thus,

$$x+y+x+y = x+x+y+y$$

so that adding $-x$ and $-y$ on the left and right, respectively, we obtain $y + x = x + y$.

- Using condition 8, it can verified that for $x \in V$ and $\alpha \in \mathbb{F}$, if $\alpha x = 0$, then either $\alpha = 0$ or $x = 0$.

Example 1.3. The assertions in the following examples must be verified by the reader.

1. $\mathbb{R}^n$ with coordinate-wise addition and scalar multiplication is a vector space over $\mathbb{R}$, but not a vector space over $\mathbb{C}$.
2. $\mathbb{C}^n$ with coordinate-wise addition and scalar multiplication is a vector space over $\mathbb{C}$.
3. $\mathbb{F}^n$ with coordinate-wise addition and scalar multiplication is a vector space over $\mathbb{F}$ but not a vector space over a field $\widetilde{\mathbb{F}} \supseteq \mathbb{F}$ with $\widetilde{\mathbb{F}} \neq \mathbb{F}$.
4. $\mathcal{P}_n(\mathbb{F})$, the set of all polynomials with coefficients from $\mathbb{F}$ and of degree atmost n, is a vector space over $\mathbb{F}$.
5. $\mathcal{P}(\mathbb{F}) := \bigcup_{n=1}^{\infty} \mathcal{P}_n(\mathbb{F})$, the set of all polynomials with coefficients from $\mathbb{F}$, is a vector space over $\mathbb{F}$.
6. $\mathbb{R}^{m \times n}$, the set of all real $m \times n$ matrices is a vector space over $\mathbb{R}$ under usual matrix multiplication and scalar multiplication.
7. Let Ω be a nonempty set. Then the set $\mathcal{F}(\Omega, \mathbb{F})$, the set of all $\mathbb{F}$-valued functions defined on Ω, is a vector space over $\mathbb{F}$ with respect to the following vector space operations: For $x, y \in \mathcal{F}(\Omega, \mathbb{F})$ and $\alpha \in \mathbb{F}$, $x + y$ and αx are defined by

$$(x+y)(t) = x(t) + y(t) \quad \forall\, t \in \Omega,$$

$$(\alpha x)(t) = \alpha x(t) \quad \forall\, t \in \Omega.$$

The zero function is the zero vector and for

$$(-x)(t) = -x(t) \quad \forall\, t \in \Omega.$$

Note that if $\Omega = \mathbb{N}$, then $\mathcal{F}(\mathbb{N}, \mathbb{F})$ is the set of all scalar sequences.

$\diamondsuit$

Definition 1.4. Let V_1 and V_2 be vector spaces over the same field $\mathbb{F}$. Then V_1 and V_2 are said to be **isomorphic** if there exists a function $T : V_1 \to V_2$ which is bijective (i.e., one-one and onto) and

$$T(x+y) = T(x) + T(y), \quad T(\alpha x) = \alpha T(x)$$

for all $x, y \in V$ and $\alpha \in \mathbb{F}$, and the map T is called an isomorphism. ♢

Example 1.5. The assertions in the following examples must be verified by the reader.

1. The spaces $\mathcal{P}_n(\mathbb{F})$ and $\mathbb{F}^{n+1}$ are isomorphic, and an isomorphism is given by

$$a_0 + a_1 t + \cdots + a_n t^n \mapsto (a_0, a_1, \ldots, a_n).$$

2. The space $\underline{\mathbb{R}}^n := \mathbb{R}^{n \times 1}$, the space of all column n-vectors is isomorphic with $\mathbb{R}^n$.
3. The space $\mathbb{R}^{m \times n}$ is isomorphic with $\mathbb{R}^{mn}$.

♢

1.2 Subspaces

Definition 1.6. A subset S of a vector space V is called a **subspace** if S itself is a vector space under the vector addition and scalar multiplication for the space V. ♢

THEOREM 1.7. *A subset S of a vector space V is a subspace if and only if S is closed under vector addition and scalar multiplication, i.e.,*

$$x, y \in S,\ \alpha \in \mathbb{F} \implies x + y \in S, \quad \alpha x \in S.$$

Example 1.8. The assertions in the following examples must be verified by the reader.

1. $S = \{(\alpha_1, \alpha_2) \in \mathbb{R}^2 : \alpha_1 + \alpha_2 = 0\}$ is a subspace of $\mathbb{R}^2$.
2. $S = \{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^2 : \alpha_1 + \alpha_2 - \alpha_3 = 0\}$ is a subspace of $\mathbb{R}^3$.
3. For each $k \in \{1, \ldots, n\}$,

$$S_k = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n : \alpha_k = 0\}$$

is a subspace of $\mathbb{F}^n$.

4. For $n \in \mathbb{N}$ with $n \geq 2$ and each $k \in \{1, \ldots, n-1\}$,

$$S_k = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n : \alpha_i = 0\, \forall i > k\}$$

is a subspace of $\mathbb{F}^n$.

5. For each $n \in \mathbb{N}$, $\mathcal{P}_n$ is a subspace of $\mathcal{P}$.

6. For each $n \in \mathbb{N}$,

$$V_n := \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : x(j) = 0 \, \forall \, j \geq n\}$$

is a subspace of $\mathcal{F}(\mathbb{N}, \mathbb{F})$, and

$$c_{00} := \bigcup_{n=1}^{\infty} V_n$$

is a subspace of $\mathcal{F}(\mathbb{N}, \mathbb{F})$. Note that elements of W are sequences having only a finite number of nonzero entries.

7. For an interval $\Omega := [a, b] \subseteq \mathbb{R}$,

   (a) $C(\Omega)$, the set of all real valued continuous functions defined on Ω is a subspace of $\mathcal{F}(\Omega, \mathbb{R})$.

   (b) $\mathcal{R}(\Omega)$, the set of all Riemann integrable real valued continuous functions defined on Ω is a subspace of $\mathcal{F}(\Omega, \mathbb{R})$.

   (c) $C(\Omega)$ is a subspace of $\mathcal{R}(\Omega)$

   (d) $C^1(\Omega)$, the set of all real valued continuous functions defined on Ω and having continuous derivative in Ω is a subspace of $C(\Omega)$.

   (e) $S = \{x \in C(\Omega) : \int_a^b x(t)dt = 0\}$ is a subspace of $C(\Omega)$.

   (f) $S = \{x \in C(\Omega) : x(a) = 0\}$ is a subspace of $C(\Omega)$.

   (g) $S = \{x \in C(\Omega) : x(a) = 0 = x(b)\}$ is a subspace of $C(\Omega)$.

8. Let $A \in \mathbb{R}^{m \times n}$. Then

   (a) $\{x \in \underline{\mathbb{R}}^n : Ax = 0\}$ is a subspace of $\underline{\mathbb{R}}^n$,

   (b) $\{Ax : x \in \underline{\mathbb{R}}^n\}$ is a subspace of $\underline{\mathbb{R}}^m$,

9. $\{(\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 = 0\}$ is a subspace of $\mathbb{R}^2$.

10. $\{(\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 - \alpha_3 = 0\}$ is a subspace of $\mathbb{R}^3$.

11. For $i, j \in \mathbb{N}$, let $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$ and for $i \in \{1, \ldots, n\}$, let $e_i = (\delta_{i1}, \ldots, \delta_{in})$. Let $V = \mathbb{R}^n$. Then $\{(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \alpha_i = 0 \text{ for } i > k\}$ is a subspace of $\mathbb{R}^n$.

12. If V_1 and V_2 are subspaces of V, then $V_1 + V_2 = \text{span}(V_1 \cup V_2)$.

13. If V_1 and V_2 are subspaces of V and if $V_1 \subseteq V_2$, then $V_1 \cup V_2$ is a subspace of V.

14. If V_1 and V_2 are subspaces of V, then $V_1 \cap V_2$ is a subspace of V; but, $V_1 \cup V_2$ need not be a subspace of V.

$\diamondsuit$

1.3 Linear combination and span

Definition 1.9. Let $x_1, \ldots, x_n$ be vectors in a vector space V. A **linear combination** of $x_1, \ldots, x_n$ is a vector of the form $\alpha_1 x_1 + \ldots + \alpha_n x_n$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. ♢

Definition 1.10. Let S be a subset of a vector space V. The set of all linear combinations of elements from S is called the **span** of S, and it is denoted by **span(S)**. ♢

THEOREM 1.11. *Let V be a vector space and $S \subseteq V$.*

1. $\mathrm{span}(S)$ *is a subspace of* V.
2. *If* V_0 *is a subspace of* V *such that* $S \subseteq V_0$*, then* $\mathrm{span}(S) \subseteq V_0$.
3. $S = \mathrm{span}(S)$ *if and only if* S *is a subspace of* V.

Example 1.12. The assertions in the following examples must be verified by the reader.

1. If $V = \mathbb{R}^2$, then $\mathrm{span}(\{(1, -1)\} = \{(\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 = 0\}$.
2. If $V = \mathbb{R}^3$, then $\mathrm{span}(\{(1, -1, 0),\ (1, 0, 1)\} = \{(\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 - \alpha_3 = 0\}$.
3. For $i, j \in \mathbb{N}$, let $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$ and for $i \in \{1, \ldots, n\}$, let $e_i = (\delta_{i1}, \ldots, \delta_{in})$. Let $V = \mathbb{R}^n$. Then
   (a) $\mathrm{span}(\{e_1, \ldots, e_k\}) = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \alpha_i = 0 \text{ for } i > k\}$.
   (b) $\mathrm{span}(\{e_1, \ldots, e_n\}) = \mathbb{R}^n$.
4. If $V = \mathcal{P}$, then $\mathrm{span}(\{1, t, \ldots, t^n\}) = \mathcal{P}_n$ and $\mathrm{span}(\{1, t, t^2, \ldots\}) = \mathcal{P}$.
5. For each $i \in \mathbb{N}$, let $e_i = (\delta_{i1}, \delta_{i2}, \ldots,)$. Then $\mathrm{span}(\{e_1, e_2, \ldots\}) = c_{00}$.

♢

Exercise 1.13. Let S be a subset of a vector space V. Prove that

1. $\mathrm{span}(S)$ is the intersection of all subspaces which contain S,
2. $\mathrm{span}[\mathrm{span}(S)] = \mathrm{span}(S)$.

♢

Notation: If S_1 and S_2 are subsets of a vector space V, then the we denote

$$S_1 + S_2 := \{x + y : x \in S_1,\ y \in S_2\}.$$

- If V_1 and V_2 are subspaces of V, then $V_1 + V_2$ is a subspace of V and

$$V_1 + V_2 = \text{span}(V_1 \cup V_2).$$

Definition 1.14. If V_1 and V_2 are subspaces of V, then the subspace $\{x + y : x \in V_1, y \in V_2\}$ is called the **sum of subspaces** V_1 and V_2. ♢

1.4 Linear dependence, linear independence, basis and dimension

Definition 1.15. Let V be a vector space and $x_1, \ldots, x_n$ are in V.

1. Vectors $x_1, \ldots, x_n$ are said to be **linearly dependent** if there exist scalars $\alpha_1, \ldots, \alpha_n$ with atleast one of them is nonzero such that $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$.

2. Vectors $x_1, \ldots, x_n$ are said to be **linearly independent** if they are not linearly dependent, i.e., for scalars $\alpha_1, \ldots, \alpha_n$,

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0 \implies \alpha_1 = 0, \ldots, \alpha_n = 0.$$

♢

Definition 1.16. Let V be a vector space and $S \subseteq V$.

1. S is said to be **linearly dependent** if S contains a finite subset which is linearly dependent.

2. S is said to be **linearly independent** if every finite subset of S is linearly independent.

♢

Definition 1.17. Let V be a vector space. A subset E of V is said to be a **basis** of V if it is linearly independent and $\text{span}(E) = V$. ♢

Example 1.18. The assertions in the following examples must be verified by the reader.

1. $\{e_1, \ldots, e_n\}$ is a basis of $\mathbb{R}^n$ and $\mathbb{C}^n$.

2. $\{1, t, \ldots, t^n\}$ is a basis of $\mathcal{P}_n$.

3. $\{1, 1+t, 1+t+t^2, \ldots, 1+t+\cdots+t^n\}$ is a basis of $\mathcal{P}_n$.

4. $\{1, t, t^2, \ldots\}$ is a basis of $\mathcal{P}$.

5. For each $i \in \mathbb{N}$, let $e_i = (\delta_{i1}, \delta_{i2}, \ldots,)$. Then $\{e_1, e_2, \ldots\}$ is a basis of c_{00}.

6. If E is linearly independent in a vector space, then E is a basis for $V_0 := \text{span}(E)$.

♢

- If E is linearly independent and if $x \in V$ with $x \notin \text{span}(E)$, then $E \cup \{x\}$ is linearly independent.

THEOREM 1.19. *If V has a finite spanning set and if E_0 is linearly independent in V, then there exists a basis $E \supseteq E_0$.*

Proof. Suppose $S = \{u_1, \ldots, u_n\}$ is such that $\text{span}(S) = V$. Let

$$E_1 = \begin{cases} E_0 & \text{if } u_1 \in \text{span}(E_0), \\ E_0 \cup \{u_1\} & \text{if } u_1 \notin \text{span}(E_0). \end{cases}$$

Then E_1 is linearly independent and $\text{span}(E_1) = \text{span}(E_0 \cup \{u_1\})$. Having defined $E_1, \ldots, E_{k-1}$, define

$$E_k = \begin{cases} E_{k-1} & \text{if } u_k \in \text{span}(E_{k-1}), \\ E_{k-1} \cup \{u_k\} & \text{if } u_k \notin \text{span}(E_{k-1}). \end{cases}$$

Thus, for $k = 1, \ldots, n$, E_k is linearly independent and

$$\text{span}(E_k) = \text{span}(E_{k-1} \cup \{u_k\}) = \text{span}(E_0 \cup \{u_1, \ldots, u_k\}).$$

Hence,

$$V = \text{span}\{u_1, \ldots, u_n\} \subseteq \text{span}(E_0 \cup \{u_1, \ldots, u_n\}) = \text{span}(E_n) \subseteq V.$$

Thus, $E := E_n$ satisfies the requirements. □

Taking $E_0 = \varnothing$ in the above theorem, we obtain the following corollary.

COROLLARY 1.20. *Every vector space having a finite spanning set has a finite basis.*

THEOREM 1.21. *Suppose a basis of a vector space V contains n vectors. Then every subset containing more than n vectors is linearly dependent.*

Proof. Suppose $E = \{u_1, \ldots, u_n\}$ be a basis of V. Its enough to prove that every subset containing $n+1$ vectors is linearly dependent. Let $S = \{x_1, \ldots, x_{n+1}\} \subseteq V$. We prove S is linearly dependent. Without loss of generality assume that $\{x_1, \ldots, x_n\}$ is linearly independent. Since $\{u_1, \ldots, u_n\}$ is a basis of V, there exists scalars $\alpha_1^{(1)}, \ldots, \alpha_n^{(1)}$ such that

$$x_1 = \alpha_1^{(1)} u_1 + \ldots, \alpha_n^{(1)} u_n.$$

Since $x_1 \neq 0$, all of $\alpha_1^{(1)}, \ldots, \alpha_n^{(1)}$ cannot be zero. So, atleast one of $\alpha_1^{(1)}, \ldots, \alpha_n^{(1)}$ is nonzero. Without loss of generality assume that $\alpha_1^{(1)} \neq 0$. Then

$$u_1 \in \text{span}\{x_1, u_2, \ldots, u_n\}.$$

But, $V = \text{span}\{u_1, u_2, \ldots, u_n\}$. Hence,

$$\text{span}\{x_1, u_2, \ldots, u_n\} = V.$$

There exists scalars $\alpha_1^{(2)}, \ldots, \alpha_n^{(2)}$ such that

$$x_2 = \alpha_1^{(2)} x_1 + \alpha_2^{(2)} u_2 \ldots, \alpha_n^{(2)} u_n.$$

Since x_1, x_2 are linearly independent, all of $\alpha_2^{(2)}, \ldots, \alpha_n^{(2)}$ cannot be zero. So, atleast one of $\alpha_2^{(2)}, \ldots, \alpha_n^{(2)}$ is nonzero. Without loss of generality assume that $\alpha_2^{(2)} \neq 0$. Then

$$u_2 \in \text{span}\{x_1, x_2, u_3, \ldots, u_n\}.$$

But, $V = \text{span}\{x_1, u_2, \ldots, u_n\}$. Hence,

$$\text{span}\{x_1, x_2, \ldots, u_n\} = V.$$

Proceeding like this, we obtain at the n^{th} step,

$$\text{span}\{x_1, x_2, \ldots, x_n\} = V.$$

Thus, $x_{n+1} \in \text{span}\{x_1, x_2, \ldots, x_n\} = V$ so that $x_1, \ldots, x_{n+1}$ are linearly dependent. □

COROLLARY 1.22. *If a vector space V has a finite basis, then any two basis of V contains the same number of vectors.*

Definition 1.23. Let V be a vector space. Then

1. V is said to be a **finite dimensional space**, if V has a finite basis, and in that case the number of elements in a basis is called the **dimension** of V, and it is denoted by **dim(V)**.
2. V is said to be an **infinite dimensional space**, if V does not have a finite basis, and we write $\dim(V) = \infty$.

♢

Example 1.24. The assertions in the following examples must be verified by the reader.

1. $\mathbb{F}^n$ and $\mathcal{P}_n$ are finite dimensional spaces, and $\dim(\mathbb{F}^n) = n, \quad \dim(\mathcal{P}_n) = n + 1$.
2. $\dim(\{\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \alpha_1 + \cdots + \alpha_n = 0\} = n - 1$.
3. $\mathcal{P}$, $C[a, b]$, c_{00} are infinite dimensional spaces.
4. Every vector space containing an infinite linearly independent set is infinite dimensional.
5. If $A \in \mathbb{R}^{m \times n}$ with $n > m$, then there exists $\underline{x} \in \underline{\mathbb{R}}^n$ such that $A\underline{x} = 0$.

♢

Exercise 1.25. 1. If V_1 and V_2 are subspaces of a vector space V such that $V_1 \cap V_2 = \{0\}$, and if E_1 and E_2 are bases of V_1 and V_2, respectively, then $E_1 \cup E_2$ is a basis of $V_1 + V_2$; and in particular,

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2).$$

2. Let V_1 and V_2 be subspaces of a vector space V. Then

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2).$$

3. Let V_1, V_2, W_1, W_2 be subspaces of a vector space V such that

$$V_1 \cap V_2 = \{0\}, \quad W_1 \cap W_2 = \{0\} \quad \text{and} \quad V_1 + V_2 = W_1 + W_2.$$

   If $V_1 \subseteq W_1$ and $V_2 \subseteq W_2$, then prove that $V_1 = W_1$ and $V_2 \subseteq W_2$.

4. Let V_1 and V_2 be vector spaces and let T be an isomorphism from V_1 onto V_2. Let $E \subseteq V_1$. Then E is a basis of V_1 if and only if $\{T(u) : u \in E\}$ is a basis of V_2.

5. Let $\{u_1, \ldots, u_n\}$ be a subset of a vector space V and $T : \mathbb{F}^n \to V$ be defined by

$$T(\alpha_1, \ldots, \alpha_n) = \alpha_1 u_1 + \ldots + \alpha_n u_n, \quad (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n.$$

   Prove that $\{u_1, \ldots, u_n\}$ is linearly independent if and only if T is one-one.

$\diamondsuit$

1.5 Quotient space

Let V be vector space and W be a subspace of V. For $x \in V$, define

$$W_x := \{x + u : u \in W\},$$

and let

$$V_W := \{W_x : x \in V\}.$$

On V_W, define addition and scalar multiplication as follows:

$$W_x + W_y := W_{x+y}, \quad \alpha W_x := W_{\alpha x}.$$

Note hat

$$W_x = W \iff x \in W.$$

- V_W is a vector space with respect to the above operations with zero W_0 and additive inverse $-W_x := W_{-x}$.

Definition 1.26. The vector space V_W is called a quotient space of V with respect to W. This vector space is usually denoted by V/W, and its elements are also denoted by $x + W$ instead of W_x. $\diamondsuit$

Example 1.27. The reader is advised to verify the following assertions:

1. If V is $\mathbb{R}^2$ or $\mathbb{R}^3$ and W is a straight line passing through origin, then V/W is the set of all straight lines parallel to W.

2. If $V = \mathbb{R}^3$ and W is a plane passing through origin, then V/W is the set of all planes having the same normal as of W.

$\diamondsuit$

THEOREM 1.28. *Let V be a finite dimensional vector space and W be a subspace of V. Then*

$$\dim(V_W) = \dim(V) - \dim(W).$$

Proof. If $W = \{0\}$ or $W = V$, then the result can be seen easily. Hence, assume that $\{0\} not = W \neq V$. Let $\{u_1, \ldots . u_k\}$ be a basis of W and let $v_1, \ldots, v_m$ be in V such that $\{u_1, \ldots . u_k, v_1, \ldots, v_m\}$ is a basis of V. We have to show that $\dim(V/W) = m$. We show this by proving that $\{W_{v_1}, \ldots, W_{v_m}\}$ is a basis of V/W.

Let $\alpha_1, \ldots, \alpha_m$ be in $\mathbb{F}$ such that

$$\alpha_1 W_{v_1} + \cdots + \alpha_m W_{v_m} = W, \quad \text{i.e.,} \quad W_{\alpha_1 v_1 + \cdots + \alpha_m v_m} = W,$$

i.e.,

$$\alpha_1 v_1 + \cdots + \alpha_m v_m \in W.$$

Hence, there are $\beta_1, \ldots, \beta_k$ in $\mathbb{F}$ such that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = \beta_1 u_1 + \cdot + \beta_k u_k,$$

i.e.,

$$(\alpha_1 v_1 + \cdots + \alpha_m v_m) - (\beta_1 u_1 + \cdot + \beta_k u_k) = 0.$$

Since $\{u_1, \ldots . u_k, v_1, \ldots, v_m\}$ is a basis of V, we have $\alpha_i = 0$, $\beta_j = 0$ for $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, k\}$. Thus, $\{W_{v_1}, \ldots, W_{v_m}\}$ is linearly independent.

It remains to show that $\text{span}\{W_{v_1}, \ldots, W_{v_m}\} = V/W$. For this, let $x \in V$ and let Let $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_k$ in $\mathbb{F}$ such that

$$x = \alpha_1 v_1 + \cdots + \alpha_m v_m + \beta_1 u_1 + \cdot + \beta_k u_k.$$

Then

$$W_x = \alpha_1 v_1 + \cdots + \alpha_m v_m + W = \alpha_1 W_{v_1} + \cdots + \alpha_m W_{v_m}.$$

This completes the proof. $\square$

1.6 Existence of a basis

We have seen that if a vector space has a finite spanning set, then it has a finite basis.

Does every vector space have a basis?

This question cannot be answered that easily. If we assume *Zorn's lemma*, then we can answer the above question affirmatively. In order to state Zorn's lemma we have to recall some concepts

Definition 1.29. A relation $\mathcal{R}$ on set S is *partial order* on S if it is

1. Reflexive: $x\mathcal{R}x$ for every $x \in S$,
2. Antisymmetric: For $x, y \in S, \quad x\mathcal{R}x, \, \& \, y\mathcal{R}x \Longrightarrow x = y$,
3. Transitive: For $x, y, z \in S, \quad x\mathcal{R}y, \, \& \, y\mathcal{R}z \Longrightarrow x\mathcal{R}z$.

A set together with a partial order is called a partially ordered set. A partial order is usually denoted by $\preceq$. ♢

Definition 1.30. Let S be a partially ordered set with partial order $\preceq$.

1. An element $b \in S$ is called an *upper bound* for a subset T of S if $x \preceq b$ for all $x \in T$.
2. A subset T of S is said to be a *totally ordered subset* of S if any two elements of T can be compared, that is, for every $x, y \in T$, either $x \preceq y$ or $y \preceq x$.
3. An element $x_0 \in S$ is called a *maximal element* of S if for any $x \in S$,
$$x_0 \preceq x \quad \Longrightarrow \quad x = x_0.$$

♢

Example 1.31. The reader is advised to verify the following assertions:

1. The set $\mathbb{R}$ with usual order $\leq$ is a partially ordered set.
2. Any subset of $\mathbb{R}$ is a totally ordered subset of $\mathbb{R}$, and if a subset T of $\mathbb{R}$ is bounded above, then every $b \geq \sup(T)$ is an upper bound of T.
3. $\mathbb{R}$ does not have any maximal element.
4. Any subset of $\mathbb{R}$ is a partially ordered set with the partial order $\leq$. If $S \subseteq \mathbb{R}$ is bounded above and then $b := \sup(S) \in S$, then b is a maximal element of S.

♢

Example 1.32. Let X be any set and S be the *power set*, i.e., the set of all subsets of X. For A, B in S, define $A \preceq B \iff A \subseteq B$. Then $\preceq$ is a partial order on S. ♢

Example 1.33. Consider the closed unit disc in the plane, $D = \{re^{i\theta} : 0 \leq r \leq 1, \, 0 \leq \theta < 2\pi\}$. For $r_1 e^{i\theta_1}, r_2 e^{i\theta_2}$ in D, define

$$r_1 e^{i\theta_1} \preceq r_2 e^{i\theta_2} \iff \theta_1 = \theta_2 \quad \& \quad r_1 \leq r_2.$$

Then $\preceq$ is a partial order on D. For each $\theta \in [0, 2\pi)$, the set

$$D_\theta := \{re^{i\theta} : 0 \leq r \leq 1\}$$

is a totally order subset of D, and the point $e^{i\theta}$ is an upper bound for D_θ. Further, D does not have any upper bound. However, every point on the boundary of D is a maximal element of D. ♢

Zorn's lemma: *Suppose S is a nonempty partially ordered set. If every totally ordered subset of S has an upper bound, then S has a maximal element.*

THEOREM 1.34. *Every nonzero vector space has a basis. In fact, if E_0 is a linearly independent subset of vector space V, then there exists a basis E for V such that $E_0 \subseteq E$.*

Proof. Let V be a nonzero vector space and let E_0 be a linearly independent subset of V. Let $\mathcal{E}$ be the family of all linearly independent subsets of V which contains E_0. That is, $E \in \mathcal{E}$ if and only if E is a linearly independent subset of V such that $E_0 \subseteq E$. For E_1, E_2 in $\mathcal{E}$ define

$$E_1 \preceq E_2 \quad \iff \quad E_1 \subseteq E_2.$$

Then $\preceq$ is a partial order on $\mathcal{E}$. Since $E_0 \in \mathcal{E}$, $\mathcal{E}$ is nonempty. Let $\mathcal{T}$ be a totally ordered subset of $\mathcal{T}$. Let

$$T_0 = \bigcup_{T \in \mathcal{T}} T.$$

Then $T_0 \in \mathcal{E}$ and T_0 is an upper bound of $\mathcal{T}$. Hence, by Zorn's lemma, $\mathcal{E}$ has a maximal element, say E. If $\operatorname{span}(E) \neq V$, then there exists $x_0 \in V \setminus \operatorname{span}(E)$, and in that case $\widetilde{E} := \{x_0\} \cup \operatorname{span}(E) \in \mathcal{E}$, which contradicts the maximality of E. Thus, E is linearly independent such that $E_0 \subseteq E$ and $\operatorname{span}(E) = V$. In particular, E is a basis of V. □

2 Linear Transformations

Recall that if A is an $m \times n$ matrix with entries from $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, and if $\underline{x}, \underline{y} \in \underline{\mathbb{F}}^n$, and $\alpha \in \mathbb{F}$, then

$$A(\underline{x} + \underline{y}) = A\underline{x} + A\underline{y},, \quad A(\alpha \underline{x}) = \alpha A \underline{x}.$$

Generalization of the above properties of matrices we define the concept of a linear transformation between any two vector spaces.

2.1 Definition properties and examples

Definition 2.1. Let V_1 and V_2 be vector spaces over the same space $\mathbb{F}$. A function $T : V_1 \to V_2$ is called a **linear transformation** or a **linear operator** if

$$T(x + y) = T(x) + T(y) \quad \text{and} \quad T(\alpha x) = \alpha T(x)$$

for every $x, y \in V_1$ and $\alpha \in \mathbb{F}$. ♢

We observe: Let $T : V_1 \to V_2$ be a linear transformation.

- $T(0) = 0$.

- $N(T) := \{x \in V_1 : T(x) = 0\}$ is a subspace of V_1.
- $R(T) := \{T(x) : x \in V_1\}$ is a subspace of V_2.

Definition 2.2. Let $T : V_1 \to V_2$ be a linear transformation.

1. The subspaces $N(T)$ and $R(T)$ are called the **null space** and **range space** of T.
2. The $\dim[R(T)]$ is called the **rank** of T and $\dim[N(T)]$ is called the **nullity** of T.

$\diamondsuit$

Convention: If $T : V_1 \to V_2$ is a linear transformation and $x \in V_1$, then the $T(x)$ is usually denoted by Tx, i.e., $Tx := T(x)$ for all $x \in V_1$.

Example 2.3. The assertion in each of the following is to be verified by the reader. The space $C[a, b]$ and $C^1[a, b]$ are vector spaces over $\mathbb{R}$.

1. Let $A \in \mathbb{R}^{m\times n}$ and let $T : \underline{\mathbb{R}}^n \to \underline{\mathbb{R}}^m$ be defined by
$$T\underline{x} = A\underline{x}\,, \quad \underline{x} \in \underline{\mathbb{R}}^n.$$
Then T is a linear transformation.

2. For $x \in C[a, b]$, define
$$T(x) = \int_a^b x(t)dt.$$
Then $T : C[a, b] \to \mathbb{R}$ is a linear transformation.

3. For $x \in C^1[a, b]$, define
$$(Tx)(t) = x'(t), \quad t \in [a, b].$$
Then $T : C^1[a, b] \to C[a, b]$ is a linear transformation.

4. For $\tau \in [a, b]$ and $x \in C^1[a, b]$, define
$$T(x) = x'(\tau).$$
Then $T : C^1[a, b] \to \mathbb{R}$ is a linear transformation.

5. Let V_1 and V_2 be vector spaces over the same field $\mathbb{F}$.

   (a) $T : V_1 \to V_2$ defined by
   $$Tx = 0 \quad \forall\, x \in V_1$$
   is a linear transformation. This transformation is called the **zero transformation**.

(b) The map $T : V \to V$ defined by

$$Tx = x \quad \forall x \in V$$

is a linear transformation. This transformation is called the **identity transformation** on V.

(c) For each $\lambda \in \mathbb{F}$, $T_\lambda : V_1 \to V_2$ defined by

$$Tx = \lambda x \quad \forall x \in V_1$$

is a linear transformation. This transformation is called a **scalar transformation**.

6. Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$ and V be any of the spaces $c_{00}, \ell^1, \ell^\infty$. Recall that

$$c_{00} = \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : \exists\, k \in \mathbb{N} \text{ with } x(j) = 0\, \forall\, j \geq k\},$$

$$\ell^1 = \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : \sum_{j=1}^{\infty} |x(j)| \text{ converges}\},$$

$$\ell^\infty = \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : (x(n)) \text{ bounded}\}.$$

(a) $T : V \to V$ defined by

$$T(\alpha_1, \alpha_2, \ldots,) = (0, \alpha_1, \alpha_2, \ldots,)$$

is a linear transformation, called the **right shift operator**.

(b) $T : V \to V$ defined by

$$T(\alpha_1, \alpha_2, \ldots,) = (\alpha_2, \alpha_3, \ldots,)$$

is a linear transformation, called the **left shift operator**.

♢

THEOREM 2.4. *Let $T : V_1 \to V_2$ be a linear transformation. Then T is one-one if and only if $N(T) = \{0\}$.*

THEOREM 2.5. *Let $T : V_1 \to V_2$ be a linear transformation.*

1. *If $u_1, \ldots, u_n$ are in V_1 such that $Tu_1, \ldots, Tu_n$ are linearly independent in V_2, then $u_1, \ldots, u_n$ are linearly independent in V_1.*

2. *If T is one-one and $u_1, \ldots, u_n$ are linearly independent in V_1, then $Tu_1, \ldots, Tu_n$ are linearly independent in V_2.*

COROLLARY 2.6. *Let $T : V_1 \to V_2$ be a linear transformation.*

1. *If E_1 is a basis of V_1, then $R(T) = \operatorname{span}(T(E_1))$.*

2. $\dim R(T) \leq \dim(V_1)$.

3. *If T is one-one, then* $\dim R(T) = \dim(V_1)$.

4. *If V_1 and V_2 are finite dimensional such that* $\dim(V_1) = \dim(V_2)$*, then T is one-one if and only if T is onto.*

THEOREM 2.7. (Sylvester's law of nullity) *Let $T : V_1 \to V_2$ be a linear transformation. Then*

$$\mathrm{rank}(T) + \mathrm{null}(T) = \dim(V_1).$$

Proof. We know that $\mathrm{rank}(T) \leq \dim(V_1$ and $\mathrm{null}(T) \leq \dim(V_1$. Thus, if either $\mathrm{rank}(T) = \infty$ or $\mathrm{null}(T) = \infty$, then the Theorem holds. Next assume that $r = \mathrm{rank}(T) < \infty$ and $k = \mathrm{null}(T) < \infty$. Let $\{u_1, \ldots, u_k\}$ be a basis of $N(T)$ and $\{v_1, \ldots, v_r\}$ be a basis of $R(T)$. Let $w_1, \ldots, w_r$ in V_1 be such that $Tw_j = v_j$ for $j = 1, \ldots, r$. The reader may verify that

$$\{u_1, \ldots, u_k, w_1, \ldots, w_r\}$$

is a basis for V_1, which would compete the proof. □

Exercise 2.8. Let $A \in \mathbb{R}^{m \times n}$ and let $T : \underline{\mathbb{R}}^n \to \underline{\mathbb{R}}^m$ be the linear transformation defined by

$$T\underline{x} = A\underline{x}\,, \quad \underline{x} \in \underline{\mathbb{R}}^n.$$

1. T is one-one if and only if the columns of A are linearly independent.

2. $R(T)$ is the space spanned by the columns of A, and $\mathrm{rank}(T)$ is the dimension of the space spanned by the columns of A.

♢

Exercise 2.9. Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ and let $\{u_1, \ldots, u_n\}$ be a basis of V_1. Let $\{v_1, \ldots, v_n\} \subseteq V_2$. Define $T : V_1 \to V_2$ be

$$T\Big(\sum_{i=1}^{n} \alpha_i u_i\Big) = \sum_{i=1}^{n} \alpha_i v, \quad (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n.$$

1. Show that T is a linear transformation such that $T(u_j) = v_j$ for $j \in \{1, \ldots, n\}$.

2. T is one-one if and only if $\{v_1, \ldots, v_n\}$ is linearly independent.

3. T is onto if and only if $\mathrm{span}(\{v_1, \ldots, v_n\}) = V_2$.

♢

Exercise 2.10. Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ and let $E := \{u_1, \ldots, u_n\}$ be a linearly independent subset of V_1. Let $\{v_1, \ldots, v_n\} \subseteq V_2$. Show that there exists a linear transformation $T : V_1 \to V_2$ such that $T(u_j) = v_j$ for $j \in \{1, \ldots, n\}$. ♢

THEOREM 2.11. *Let $\mathcal{L}(V_1, V_2)$ be the set of all linear transformations from V_1 to V_2. For T, T_1, T_2 in $\mathcal{L}(V_1, V_2)$ and $\alpha \in \mathbb{F}$, define $T_1 + T_2$ and αT by*

$$(T_1 + T_2)(x) = T_1 x + T_2 x \quad \forall\, x \in V_1,$$

$$T(\alpha x) = \alpha T x \quad \forall\, x \in V_1.$$

Then $\mathcal{L}(V_1, V_2)$ is a vector space with respect to *the above addition and scalar multiplication its zero as the zero–transformation and $(-T)(x) := -Tx$ for all $x \in V_1$.*

Definition 2.12. The space $\mathcal{L}(V, \mathbb{F})$ is called the **dual space** of V and it is denoted by V'. Elements of V' are usually denoted by lover case letters f, g, etc. ♢

THEOREM 2.13. *Let V be a finite dimensional space and $E = \{u_1, \ldots, u_n\}$ be an order basis of V. For each $j \in \{1, \ldots, n\}$, let $f_j : V \to \mathbb{F}$ be defined by*

$$f_j(x) = \alpha_j \quad \textit{for} \quad x := \sum_{i=1}^{n} \alpha_i u_i.$$

Then

1. *$f_1, \ldots, f_n$ are in V' and they satisfy $f_i(u_j) = \delta_{ij}$ for $i, j \in \{1, \ldots, n\}$,*
2. *$\{f_1, \ldots, f_n\}$ is a basis of V'.*

COROLLARY 2.14. *Let V be a finite dimensional space. Then V and V' are linearly isomorphic.*

Definition 2.15. Let V be a finite dimensional space and $E = \{u_1, \ldots, u_n\}$ be an order basis of V. The basis $\{f_1, \ldots, f_n\}$ of V' obtained in the above theorem is called the **dual basis** of V corresponding to the ordered basis E. ♢

Note that if $E = \{u_1, \ldots, u_n\}$ is an order basis of V and $F = \{f_1, \ldots, f_n\}$ is the corresponding ordered dual basis, then for every $x \in V$ and $f \in V'$,

$$x = \sum_{i=1}^{n} f_i(x) u_i, \quad f = \sum_{i=1}^{n} f(u_i) f_i.$$

THEOREM 2.16. *Let $E = \{u_1, \ldots, u_n\}$ be an order basis of V. If $f_1, \ldots, f_n$ are in V' such that $f_i(u_j) = \delta_{ij}$, then $\{f_1, \ldots, f_n\}$ is the dual basis of V.*

THEOREM 2.17. *Let V_1, V_2, V_3 be vector spaces over the same field $\mathbb{F}$. If $T_1 \in \mathcal{L}(V_1, V_2)$ and $T_2 \in \mathcal{L}(V_2, V_3)$. Then the composition of T_2 and T_1, namely $T_2 \circ T_1$ belongs to $\mathcal{L}(V_1, V_3)$.*

Notation: The composition operator $T_2 \circ T_1$ is usually denoted by $T_2 T_1$.

THEOREM 2.18. *Let $T : V_1 \to V_2$ be a linear transformation which is one-one and onto. Then its inverse $T^{-1} : V_2 \to V_1$ is a linear transformation, and*

$$TT^{-1} = I_{V_2} \quad \textit{and} \quad T^{-1}T = I_{V_1},$$

where I_{V_1} and I_{V_1} are the identity transformations on V_1 and V_2, respectively.

Exercise 2.19. Let $T_1 \in \mathcal{L}(V_1, V_2)$ and $T_2 \in \mathcal{L}(V_2, V_3)$. Show that

1. $T_2 T_1$ one-one implies T_1 one-one.

2. $T_2 T_1$ onto implies T_2 one-one.

$\diamondsuit$

THEOREM 2.20. *Let V be a vector space and W be a subspace of V. Then the map $\eta : V \to V/W$ defined by*

$$\eta(x) = x + W, \quad x \in V,$$

is a linear transformation.

Definition 2.21. The map η in the above theorem is called the **quotient map** associated with the subspace W. $\diamondsuit$

Now, we give another proof for the Sylvester's law of nullity (Theorem 2.7) in the case of $\dim(V_1) < \infty$.

Another proof for Theorem 2.7. Let $\dim(V_1) < \infty$. Consider the operator $\tilde{T} : V_1/N(T) \to V_2$ defined by

$$\tilde{T}(x + N(T)) = Tx, \quad x \in V_1.$$

Then, it can be easily seen that $\tilde{T}$ is one-one. Hence, $V_1/N(T)$ is linearly isomorphic with $R(\tilde{T}) = R(T)$. Consequently,

$$\dim[R(T)] = \dim[V_1/N(T)] = \dim(V_1) - \dim[N(T)].$$

This completes the proof. $\square$

2.2 Matrix representation

Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ and let $E_1 := \{u_1, \ldots, u_n\}$ and $E_2 := \{v_1, \ldots, v_m\}$ be ordered bases of V_1 and V_2, respectively. Let $T : V_1 \to V_2$ be a linear transformation. For each $j \in \{1, \ldots, n\}$, let $a_{1j}, \ldots, a_{mj}$ in $\mathbb{F}$ be such that

$$Tu_j = \sum_{i=1}^{m} a_{ij} v_i.$$

Then for every $x \in V_1$, if $(\alpha_1, \ldots, \alpha_n)$ are the n-tuple of scalars such that $x = \sum_{j=1}^{n} \alpha_j u_j$, then

$$Tx = \sum_{j=1}^{n} \alpha_j T u_j = \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{m} a_{ij} v_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \alpha_j \right) v_i.$$

Definition 2.22. The matrix (a_{ij}) in the above discussion is called the **matrix representation** of T with respect to the ordered bases E_1, E_2 of V_1 and V_2, respectively. This matrix is usually denoted by $[T]_{E_1E_2}$, i.e.,

$$[T]_{E_1E_2} := (a_{ij}).$$

$\diamondsuit$

- For each j, $[Tu_j]_{E_2}$ is the j^{th} column of $[T]_{E_1E_2}$.

Example 2.23. Let $A \in \mathbb{R}^{m\times n}$ and let $T : \underline{\mathbb{R}}^n \to \underline{\mathbb{R}}^m$ be defined by

$$T\underline{x} = A\underline{x}\,, \quad \underline{x} \in \underline{\mathbb{R}}^n.$$

Recall that T is a linear transformation. Now, taking the standard basis E_1 and E_2 for $\underline{\mathbb{R}}^n$ and $\underline{\mathbb{R}}^m$, respectively, it can be seen that $[T]_{E_1E_2} = A$. $\diamondsuit$

Let V be an n-dimensional vector space and $E = \{u_1, \ldots, u_n\}$ be an ordered basis of V. Recall the **canonical isomorphism** $J : V \to \underline{\mathbb{R}}^n$ defined by

$$J(x) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad x := \sum_{i=1}^{n} \alpha_i u_i.$$

Let us denote

$$[x] := J(x).$$

In fact,

$$[x]_E := [f_1(x), \ldots, f_n(x)]^T,$$

where $F = \{f_1, \ldots, f_n\}$ is the dual basis of V, i.e., $F = \{f_1, \ldots, f_n\}$ is a basis of V' such that $f_i(u_j) = \delta_{ij}$.

- For each j, $[u_j]_E$ is the j^{th} standard basis vector of $\mathbb{F}^n$, i.e., $[u_j]_E = [\delta_{1j}\, \delta_{2j}\, \ldots \delta_{nj}]^T$,

THEOREM 2.24. *Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ with $\dim(V_1) = n$ and $\dim(V_2) = m$ and let E_1 and E_2 be ordered bases of V_1 and V_2, respectively. Let $T : V_1 \to V_2$ be a linear transformation. Then the following hold:*

1. *$[Tx]_{E_2} = [T]_{E_1E_2}[x]_{E_1}$ for all $x \in V_1$.*
2. *T is one-one (respectively, onto) if and only if $[T]_{E_1E_2} : \underline{\mathbb{R}}^n \to \underline{\mathbb{R}}^m$ is one-one (respectively, onto).*
3. *For $A \in \mathbb{R}^{m\times n}$,*

$$A = [T]_{E_1E_2} \iff [Tx]_{E_2} = A[x]_{E_1} \quad \forall\, x \in V_1.$$

4. $T = J_2[T]_{E_1E_2}J_1^{-1}$*, where* $J_1 : V_1 \to \underline{\mathbb{R}}^n$ *and* $J_2 : V_2 \to \underline{\mathbb{R}}^m$ *are the canonical isomorphisms,*

THEOREM 2.25. *Let* V_1, V_2, V_3 *be finite dimensional vector spaces over the same field* $\mathbb{F}$*, and let* E_1, E_2, E_3 *be ordered bases of* V_1, V_2, V_3*, respectively. If* $T_1 \in \mathcal{L}(V_1, V_2)$ *and* $T_2 \in \mathcal{L}(V_2, V_3)$*. Then the*

$$[T_2T_1]_{E_1E_3} = [T_2]_{E_2E_3}[T_1]_{E_1E_2}.$$

Proof. Note that for every $x \in V_1$,

$$[T_2T_1x]_{E_3} = [T_2]_{E_2E_3}[T_1x]_{E_2} = [T_2]_{E_2E_3}[T_1]_{E_1E_2}[x]_{E_1}.$$

Hence, by Theorem 2.24(3), $[T_2T_1]_{E_1E_3} = [T_2]_{E_2E_3}[T_1]_{E_1E_2}$. □

Exercise 2.26. For $n \in \mathbb{N}$, let $D : \mathcal{P}_n \to \mathcal{P}_{n-1}$ and $T : \mathcal{P}_n \to \mathcal{P}_{n+1}$ be defined by

$$D(a_0 + a_1t + \cdots + a_nt^n) = a_1t + 2a_2t + \cdots + na_nt^{n-1},$$

$$T(a_0 + a_1t + \cdot + a_nt^n) = a_0t + \frac{a_1}{2}t^2 + \cdots + \frac{a_n}{n+1}t^{n+1}.$$

Let $E_k = \{1, t, \ldots, t^k\}$ for $k \in \mathbb{N}$. Find

$$[D]_{E_nE_{n-1}}, \quad [T]_{E_nE_{n+1}}, \quad [TD]_{E_nE_n}, \quad [DT]_{E_nE_n}.$$

♢

2.3 Matrix representation under change of basis

Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ and let $T : V_1 \to V_2$ be a linear transformation. Let $E_1 = \{u_1, \ldots, u_n\}$ and $\tilde{E}_1 = \{\tilde{u}_1, \ldots, \tilde{u}_n\}$ be two bases of V_1 and $E_2 = \{v_1, \ldots, v_m\}$ and $\tilde{E}_2 = \{\tilde{v}_1, \ldots, \tilde{v}_m\}$ be two bases of V_2. One may want to know the relation between $[T]_{E_1E_2}$ and $[T]_{\tilde{E}_1\tilde{E}_2}$. For this purpose we consider the linear transformations $\Phi_1 : V_1 \to V_1$ and $\Phi_2 : V_2 \to V_2$ such that

$$\Phi_1(u_i) = \tilde{u}_i, \quad \Phi_2(v_j) = \tilde{v}_j$$

for $i = 1, \ldots, n$; $j = 1, \ldots, m$.

THEOREM 2.27.

$$[T]_{\tilde{E}_1\tilde{E}_2} = [\Phi_2]^{-1}_{E_2E_2}[T]_{E_1E_2}[\Phi_1]_{E_1E_1}.$$

Proof. Note that

$$[\Phi_1]_{E_1\tilde{E}_1} = (\delta_{ij}) = I_{n\times n}, \quad [\Phi_2]_{E_2\tilde{E}_2} = (d_{ij}) = I_{m\times m}.$$

Let

$$[T]_{E_1E_2} = (a_{ij}), \quad [T]_{\tilde{E}_1\tilde{E}_2} = (\tilde{a}_{ij}),$$

$$[\Phi_1]_{E_1E_1} = (s_{ij}), \quad [\Phi_2]_{E_2E_2} = (t_{ij}).$$

Then

$$T\tilde{u}_j = \sum_i \tilde{a}_{ij}\tilde{v}_i, \quad Tu_j = \sum_i a_{ij}v_i,$$

$$\Phi_1 u_j = \sum_i s_{ij}u_i, \quad \Phi_2 v_j = \sum_i t_{ij}v_i, .$$

Hence,

$$T\tilde{u}_j = T\Phi_1 u_j = \sum_i s_{ij}Tu_i = \sum_i s_{ij}\sum_k a_{ki}v_k = \sum_k \Big(\sum_i a_{ki}s_{ij}\Big)v_k,$$

$$\sum_i \tilde{a}_{ij}\tilde{v}_i = \sum_i \tilde{a}_{ij}\Phi_2 v_i = \sum_i \tilde{a}_{ij}\sum_k t_{ki}v_k = \sum_k \Big(\sum_i t_{ki}\tilde{a}_{ij}\Big)v_k.$$

Thus,

$$\sum_i a_{ki}s_{ij} = \sum_i t_{ki}\tilde{a}_{ij}$$

consequently,

$$[T]_{E_1E_2}[\Phi_1]_{E_1E_1} = [\Phi_2]_{E_2E_2}[T]_{\tilde{E}_1\tilde{E}_2},$$

i.e.,

$$[T]_{\tilde{E}_1\tilde{E}_2} = [\Phi_2]^{-1}_{E_2E_2}[T]_{E_1E_2}[\Phi_1]_{E_1E_1}.$$

□

3 Inner Product Spaces

Recall that in the Euclidian space $\mathbb{R}^3$ we have the concept of **dot product** and **absolute value**: For $x = (\alpha_1, \alpha - 2, \alpha_3)$, $y = (\beta_1, \beta_2, \beta_3)$ in $\mathbb{R}^3$,

$$x.y = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3,$$

$$|x| = \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2}.$$

We consider the generalization of these concepts to any vector space. Throughout this section we assume that $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.

Definition 3.1. Let V be a vector space over $\mathbb{F}$. An **inner product** on V is a map which associates each pair (x, y) of elements from V to a unique number in $\mathbb{F}$, denoted by $\langle x, y\rangle$ such that the following conditions are satisfied:

1. $\langle x, x\rangle \geq 0 \quad \forall\, x \in V$, and for every $x \in V$, $\langle x, x\rangle = 0 \iff x = 0$.
2. $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle \quad \forall\, x, y, z \in V$,
3. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle \quad \forall\, x \in V, \quad \forall\, \alpha \in \mathbb{F}$,
4. $\langle x, y\rangle = \overline{\langle y, x\rangle} \quad \forall\, x, y \in V$.

A vector space together with an inner product is called an **inner product space**. ♢

Definition 3.2. Let V be an inner product space and $x \in V$. The the number

$$\|x\| := \sqrt{\langle x, x\rangle} \text{ (positive square root)}$$

is called the **norm** of x. A vector x with $\|x\| = 1$ is called a **unit vector**. ♢

Exercise 3.3. Prove the following:

1. $\|x\| \geq 0 \quad \forall\, x \in V$ and $\|x\| = 0 \iff x = 0$.
2. $\|\alpha x\| = |\alpha|\, \|x\| \quad \forall\, x \in V, \forall\, \alpha \in \mathbb{F}$.

♢

Example 3.4. The assertions in the following are to be verified:

1. On the vector space c_{00}, $\langle x, y\rangle := \sum_{j=1}^{\infty} x(j)\overline{y(j)}$ defines an inner product.
2. On the vector space $C[a, b]$, $\langle, y\rangle := \int_a^b x(t)\overline{y(t)}dt$ defines an inner product.
3. Let $\tau_1, \ldots, \tau_{n+1}$ be distinct real numbers. On the vector space $\mathcal{P}_n$, $\langle p, q\rangle := \sum_{i=1}^{n+1} p(\tau_i)\overline{q(\tau_i)}$ defines an inner product.

♢

Exercise 3.5. Let V be an inner product space. Prove the following:

1. For $x \in V$, $\langle x, u\rangle = 0\, \forall\, u \in V \Longrightarrow x = 0$.
2. For $u \in V$, if $f : V \to \mathbb{F}$ is defined by $f(x) = \langle x, u\rangle$ for all $x \in V$, then $f \in V'$.
3. If $S \subseteq V$ is such that $\text{span}(S) = V$, then for every $x \in V$,

$$\langle x, u\rangle = 0 \quad \forall\, u \in S \quad \Longrightarrow \quad x = 0.$$

4. Let $u_1, u_2, \ldots, u_n$ be linearly independent vectors in V and let $x \in V$. Then

$$\langle x, u_i\rangle = 0 \quad \forall\, i \in \{1, \ldots, n\} \iff \langle x, y\rangle = 0 \quad \forall\, y \in \text{span}\{u_1, \ldots, u_n\}.$$

In particular, if $\{u_1, u_2, \ldots, u_n\}$ is a basis of V, then

$$\langle x, u_i\rangle = 0 \quad \forall\, i = 1, \ldots, n \iff x = 0.$$

♢

Exercise 3.6. Let V be an inner product space. Show that, for each $y \in V$, the map $x \mapsto \langle x, y\rangle$ is a linear functional on V. ♢

Exercise 3.7. Let $V = c_{00}$ with usual inner product, and let $f(x) = \sum_{j=1}^{\infty} x(j)$ for $x \in c_{00}$. Show that $f \in V'$, but there does not exist $y \in c_{00}$ such that $f(x) = \langle x, y \rangle$ for all $x \in c_{00}$. ♢

THEOREM 3.8. (Parallelogram law) *Let V be an inner product space and $x, y \in V$. Then*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Definition 3.9. Let V be an inner product space and $x \in V$.

1. Vectors x, y are said to be **orthogonal vectors** if $\langle x, y \rangle = 0$, and in that case we write $x \perp y$.
2. A subset S of V said to be an **orthogonal set** if $x \perp y$ for every $x, y \in V$ with $x \neq y$.
3. A subset S of V said to be an **orthonormal set** if it is an orthogonal set and $\|x\| = 1$ for $x \in S$.
4. For a subset S of V, then set

   $$S^{\perp} := \{x \in V : x \perp u \quad \forall\, u \in S\}$$

   is called the **orthogonal compliment** of S.

♢

THEOREM 3.10. *Every orthogonal set which does not contain 0 in it is a linearly independent set. In particular, every orthonormal set is linearly independent.*

THEOREM 3.11. (Pythagoras[2] theorem) *Let V be an inner product space and $x, y \in V$. If $x \perp y$, then*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

THEOREM 3.12. (Cauchy–Schwarz inequality) *Let V be an inner product space. Then for every $x, y \in V$,*

$$|\langle x, y \rangle| \leq \|x\|\,\|y\|. \qquad (*)$$

Equality holds if and only if x and y are linearly dependent.

Proof. If $y = 0$, then clearly the inequality holds. Hence, assume that $y \neq 0$, and let $u := \dfrac{\langle x, y \rangle}{\|y\|^2}$. Then we note that

$$x - u \perp u$$

so that, writing $x = u + (x - u)$ and using by Pythagoras theorem we obtain

$$\|x\|^2 = \|u\|^2 + \|x - u\|^2. \qquad (**)$$

Hence, $\|u\|^2 \leq \|x\|^2$; equivalently,

$$|\langle x, y \rangle| \leq \|x\|\,\|y\|.$$

Clearly, if x and y are linearly dependent, then equality holds in $(*)$. Conversely, from $(**)$, equality holds in $(*)$ implies $x = u$ and hence x and y are linearly dependent. □

[2]Greek Philosopher and Mathematician born around 570 BC

More generally we have the following:

THEOREM 3.13. *Suppose $\{u_1, \ldots, u_n\}$ is an orthonormal set in an inner product space V and $x \in V$. Then*

$$x - \sum_{i=1}^{n} \langle x, u_i \rangle u_i \perp \operatorname{span}\{u_1, \ldots, u_n\}$$

and

$$\sum_{i=1}^{n} |\langle x, u_i \rangle|^2 \leq \|x\|^2.$$

Further, the following are equivalent:

1. $x \in \operatorname{span}\{u_1, \ldots, u_n\}$
2. $x = \sum_{i=1}^{n} \langle x, u_i \rangle u_i$
3. $\|x\|^2 = \sum_{i=1}^{n} |\langle x, u_i \rangle|^2$.

Definition 3.14. Let $S := \{u_1, \ldots, u_n\}$ is an orthonormal set in an inner product space V.

1. The inequality in Theorem 3.13 is called the **Bessel's inequality**.
2. For $x \in \operatorname{span}(S)$, the equality in Theorem 3.13 (2) is called the **Fourier expansion** of x.
3. For $x \in \operatorname{span}(S)$, the equality in Theorem 3.13 (3) is called the **Parseval's identity** for x.

$\diamondsuit$

Exercise 3.15. For $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ and $(\beta_1, \ldots, \beta_n) \in \mathbb{F}^n$, show that

$$\sum_{j=1}^{n} |\alpha_j \beta_j| \leq \Big(\sum_{j=1}^{n} |\alpha_j|^2\Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{n} |\beta_j|^2\Big)^{\frac{1}{2}}.$$

$\diamondsuit$

Exercise 3.16. For $x, y \in \mathcal{F}(\mathbb{N})$ prove that

$$\sum_{j=1}^{\infty} |\alpha_j \beta_j| \leq \Big(\sum_{j=1}^{\infty} |\alpha_j|^2\Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{\infty} |\beta_j|^2\Big)^{\frac{1}{2}}.$$

Hint: Use Exercise 10. $\diamondsuit$

Exercise 3.17. Let

$$\ell^2 = \{x \in \mathcal{F}(\mathbb{N}) : \sum_{j=1}^{\infty} |x(j)|^2 < \infty\}.$$

Prove that

1. ℓ^2 is a subspace $\mathcal{F}(\mathbb{N})$.

2. For $x, y \in \ell^2$, $\sum_{j=1}^{\infty} |x(j)\overline{y(j)}|$ converges.

3. $\langle x, y \rangle := \sum_{j=1}^{\infty} |x(j)\overline{y(j)}|$ defines an inner product on ℓ^2.

$\diamondsuit$

Using Cauchy-Schwarz inequality we obtain the following:

THEOREM 3.18. *Let V be an inner product space and $x, y \in V$. For every $x, y \in V$,*

$$\|x + y\| \leq \|x\| + \|y\|.$$

COROLLARY 3.19. *Let V be an inner product space. Then the map $(x, y) \mapsto \|x - y\|$ is a metric on V.*

Definition 3.20. The metric defined in Corollary 3.19 is called the **metric induced by the inner product**. $\diamondsuit$

Definition 3.21. An inner product space V is called a **Hilbert space** if it is complete with respect to the metric induced by the inner product. $\diamondsuit$

Exercise 3.22. For $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ and $(\beta_1, \ldots, \beta_n) \in \mathbb{F}^n$, show that

$$\Big(\sum_{j=1}^{n} |\alpha_j + \beta_j|^2\Big)^{\frac{1}{2}} \leq \Big(\sum_{j=1}^{n} |\alpha_j|^2\Big)^{\frac{1}{2}} + \Big(\sum_{j=1}^{n} |\beta_j|^2\Big)^{\frac{1}{2}}.$$

$\diamondsuit$

Exercise 3.23. For $x, y \in \mathcal{F}(\mathbb{N})$ prove that

$$\Big(\sum_{j=1}^{\infty} |\alpha_j + \beta_j|^2\Big)^{\frac{1}{2}} \leq \Big(\sum_{j=1}^{\infty} |\alpha_j|^2\Big)^{\frac{1}{2}} + \Big(\sum_{j=1}^{\infty} |\beta_j|^2\Big)^{\frac{1}{2}}.$$

Hint: Use Exercise 12. $\diamondsuit$

Exercise 3.24. Let V be an inner product space. Show that

1. $\|x - y\| \geq \|x\| - \|y\|$ for all $x, y \in V$,

2. $x \mapsto \|x\|$ is continuous on V,

3. $S \subseteq V$ implies $S^{\perp}$ is a closed subset of V.

$\diamondsuit$

Definition 3.25. An orthonormal set E in an inner product space V is called an **orthonormal basis** if it is a maximal orthonormal set. $\diamondsuit$

THEOREM 3.26. *Let V be an inner product space. If E is a basis of V which is also an orthonormal set, then it is an orthonormal basis.*

Proof. Suppose E is a basis of V which is also an orthonormal set. If E is not an orthonormal basis, then there exists an orthonormal set $\tilde{E} \supseteq E$ such that $\tilde{E} \neq E$. In particular there exists $x \in \tilde{E} \setminus E$. Then $\langle x, u \rangle = 0$ for every $u \in E$ so that $x = 0$ which contradicts the fact that $\tilde{E}$ is linearly independent. □

Remark 3.27. An orthonormal basis need not be a basis: For example, consider the inner product space ℓ^2 and $E = \{e_1, e_2, \ldots\}$, where $e_j(i) = \delta_{ij}$. Then E is an orthonormal basis, since

$$\langle x, e_j \rangle = 0 \quad \forall j \in \nu \quad \Longrightarrow \quad x = 0.$$

But, E is not a basis of ℓ^2. For instance $(\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots)$ in ℓ^2 is not in the span of E. ♢

Now we show that for a finite dimensional inner product space, every orthonormal basis is a basis. For this, first we observe the following.

THEOREM 3.28. (Gram-Schmidt orthogonalization process) *Let V be an inner product space and $\{x_1, \ldots, x_n\}$ be an ordered linearly independent set for $n \geq 2$. Let $u_1 = x_1$ and for $j = 1, \ldots, n-1$, let*

$$u_{j+1} = x_{j+1} - \sum_{i=1}^{j} \frac{\langle x_{j+1}, u_i \rangle}{\|u_i\|^2} u_i.$$

Then $\{u_1, \ldots, u_n\}$ is an orthonormal set and

$$\text{span}\{u_1, \ldots, u_j\} = \text{span}\{x_1, \ldots, x_j\}, \quad j = 1, \ldots, n.$$

COROLLARY 3.29. *Every finite dimensional inner product space has an orthonormal basis, and every orthonormal basis of a finite dimensional inner product space is a basis.*

As a corollary to Theorem 3.13 we have the following:

THEOREM 3.30. *Let V be a finite dimensional inner product space and $\{u_1, \ldots, u_n\}$ be an orthonormal basis of V. Then the following hold.*

1. **(Fourier expansion)** *For all $x \in V$, $x = \sum_{j=1}^{n} \langle x, u_j \rangle u_j$, and*
2. **(Riesz representation theorem)** *For every $f \in V'$, there exists a unique $y \in V$ such that $f(x) = \langle x, y \rangle$ for all $x \in V$.*

Proof. Part 1 follows from Theorem 3.13. For part 2, let $f \in V'$ and $x \in V$. From part 1,

$$f(x) = \sum_{j=1}^{n} \langle x, u_j \rangle f(u_j) = \langle x, \sum_{j=1}^{n} \overline{f(u_j)} u_j \rangle.$$

Thus, $f(x) = \langle x, y \rangle$, where $y = \sum_{j=1}^{n} \overline{f(u_j)} u_j$.

Uniqueness follows easily (Write details!). □

COROLLARY 3.31. *Let V_1 and V_2 be a inner product spaces and $A : V_1 \to V_2$ be a linear transformation. If V_1 is finite dimensional, then there exists a unique linear transformation $B : V_2 \to V_1$ such that*

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall\,(x, y) \in V_1 \times V_2.$$

Proof. Let $\{u_1, \ldots, u_n\}$ be an orthonormal basis of V_1 and $x \in V_1$. Since $x = \sum_{j=1}^{n} \langle x, u_j\rangle u_j$, for every $y \in V_2$, we have

$$\langle Ax, y\rangle = \left\langle \sum_{j=1}^{n} \langle x, u_j\rangle Au_j, y \right\rangle = \sum_{j=1}^{n} \langle x, u_j\rangle \langle Au_j, y\rangle = \sum_{j=1}^{n} \langle x, \overline{\langle Au_j, y\rangle} u_j\rangle = \left\langle x, \sum_{j=1}^{n} \overline{\langle Au_j, y\rangle} u_j \right\rangle.$$

Thus, $\langle Ax, y\rangle = \langle x, By\rangle$ for all $(x, y) \in V_1 \times V_2$, where

$$By := \sum_{j=1}^{n} \overline{\langle Au_j, y\rangle} u_j.$$

It can be easily seen (Write details!) that $B : V_2 \to V_1$ is a linear transformation and it is the unique linear transformation satisfying $\langle Ax, y\rangle = \langle x, By\rangle$ for all $(x, y) \in V_1 \times V_2$. □

Definition 3.32. The transformation B in Corollary 3.31 is called the **adjoint** of A, and it is usually denoted by A^*. ♢

Definition 3.33. Let V be a finite dimensional inner product space and $A : V \to V$ be a linear transformation. Then A is called a

1. **self-adjoint operator** if $A^* = A$,
2. **normal operator** if $A^*A = AA^*$,
3. **unitary operator** if $A^*A = I = AA^*$.

♢

Observe the following:

- If A is self adjoint, the $\langle Ax, x\rangle \in \mathbb{R}$ for every $x \in X$.
- If A is normal, then $\|Ax\| = \|A^*x\|$ for every $x \in X$.
- If A is unitary, then $\langle Ax, Ay\rangle = \langle x, y\rangle$ for every $x, y \in X$. In particular, images of orthogonal vectors are orthogonal.

Exercise 3.34. 1. Let $V = \mathbb{F}^3$ with standard inner product. In the following, given vectors $x, y, z \in \mathbb{F}^3$ construct orthonormal vectors u, v, w in $\mathbb{F}^3$ such that $\text{span}\{u, v\} = \text{span}\{x, y\}$ and $\text{span}\{u, v, w\} = \text{span}\{x, y, z\}$.

(a) $x = (1, 0, 0)$, $y = (1, 1, 0)$, $z = (1, 1, 1)$;

(b) $x = (1,1,0)$, $y = (0,1,1)$, $z = (1,0,1)$.

2. Let $\dim(V) = n$ and let $E = \{u_1, \ldots, u_n\}$ be an ordered orthonormal set which is a basis of V. Let $A : V \to V$ be a linear transformation. Show that $[A]_{E,E} = (\langle Au_j, u_i\rangle)$. [Hint: Use Fourier expansion.]

3. Let $\dim(V) = n$ and let $E = \{u_1, \ldots, u_n\}$ be an ordered orthonormal set which is a basis of V. Let $A, B : V \to V$ be a linear transformations satisfying $\langle Ax, y\rangle = \langle x, By\rangle$ for all $x, y \in V$. Show that $[B]_{E,E} = \overline{[A]}^T_{E,E}$, conjugate transpose of $[A]_{E,E}$.

4. Let $E_1 = \{u_1, \ldots, u_n\}$ and $E_2 = \{v_1, \ldots, v_m\}$ be an ordered orthonormal bases of inner product spaces V_1 and V_2, respectively. If $A : V_1 \to V_2$ is a linear transformation, then prove that $[A^*]_{E_2,E_1} = (b_{ij})$, where $b_{ij} = \overline{\langle Au_i, u_j\rangle}$.

$\diamondsuit$

THEOREM 3.35. (Projection theorem) *Let V be an inner product space and V_0 be a finite dimensional subspace of V. Then*

$$V = V_0 + V_0^{\perp}.$$

In particular, for every $x \in V$, there exists a unique pair $(y, z) \in V_0 \times V_0^{\perp}$ such that $x = y + z$.

Proof. Let $\{u_1, \ldots, u_n\}$ be an orthonormal basis for V_0. For $x \in V$, let $y = \sum_{j=1}^{n} \langle x, u_j\rangle u_j$. Then we see that $x = y + (x - y)$ with $y \in V_0$ and $x - y \in V_0^{\perp}$. Uniqueness follows easily (Write details!). $\square$

COROLLARY 3.36. (Best approximation) *Let V be an inner product space and V_0 be a finite dimensional subspace of V. Then for every $x \in V$, there exists a unique pair $y \in V_0$ such that*

$$\|x - y\| = \inf_{u \in V_0} \|x - u\|.$$

Proof. Let $x \in V$ and let $(y, z) \in V_0 \times V_0^{\perp}$ be as in Theorem 3.35. Then, $x - y = z \in V_0^{\perp}$ and for every $u \in V_0$, $y - u \in V_0$ so that by Pythagoras theorem,

$$\|x - u\|^2 = \|(x - y) + (y - u)\|^2 = \|x - y\|^2 + \|y - u\|^2.$$

Thus, $\|x - y\| \le \|x - u\|$ for all $u \in V_0$ so that

$$\|x - y\| = \inf_{u \in V_0} \|x - u\|.$$

If there is also $y_1 \in V_0$ such that $\|x - y_1\| = \inf_{u \in V_0} \|x - u\|$, then we have

$$\|x - y_1\| = \|x - y\|$$

and hence, again by Pythagoras theorem,

$$\|x - y_1\|^2 = \|(x - y) + (y - y_1)\|^2 = \|x - y\|^2 + \|y - y_1\|^2.$$

Thus, we obtain $\|y - y_1\| = 0$. $\square$

Exercise 3.37. Let V be an inner product space and V_0 be a finite dimensional subspace of V. For $x \in V$, let (y, z) be the unique element in $V_0 \times V_0^\perp$ such that $x = y + z$. Let $P, Q : V \to V$ be defined by $P(x) = y$ and $Q(x) = z$. Prove that P and Q are liner transformations satisfying the following:

$$R(P) = V_0, \quad R(Q) = V_0^\perp, \quad P^2 = P, \quad Q^2 = Q, \quad P + Q = I,$$

$$\langle Pu, v\rangle = \langle u, Pv\rangle \quad \forall\, u, v \in V, \quad \|x - Px\| \le \|x - u\| \quad \forall\, u \in V_0.$$

♢

Exercise 3.38. Let V be an inner product space and V_0 be a subspace of V and let $x \in V$ and $y \in V_0$. Prove the following:

1. If $\langle x - y, u\rangle = 0 \quad \forall\, u \in V_0 \implies \|x - y\| = \inf_{u \in V_0} \|x - u\|$.
2. If $\mathrm{span}(S) = V_0$ and $\langle x - y, u\rangle = 0 \quad \forall\, u \in S \implies \|x - y\| = \inf_{u \in V_0} \|x - u\|$.

♢

Exercise 3.39. Let V be an inner product space, V_0 be a finite dimensional subspace of V and $x \in V$. Let $\{u_1, \ldots, u_k\}$ be a basis of V_0. Prove that for $y = \sum_{j=1}^{k} \alpha_j u_j$,

$$\langle x - y, u\rangle = 0 \quad \forall\, u \in V_0 \iff \sum_{j=1}^{k} \langle u_j, u_i\rangle \alpha_j = \langle x, u_i\rangle, \quad i = 1, \ldots, k.$$

Further, prove that there exists a unique $(\alpha_1, \ldots, \alpha_k) \in \mathbb{F}^k$ such that

$$\sum_{j=1}^{k} \langle u_j, u_i\rangle \alpha_j = \langle x, u_i\rangle, \quad i = 1, \ldots, k,$$

and in that case $\|x - y\| = \inf_{u \in V_0} \|x - u\|$. ♢

Exercise 3.40.

1. Let $V = C[0, 1]$ with inner product: $\langle f, g\rangle := \int_0^1 f(t)\overline{g(t)}dt$. Let $x(t) = t^5$. Find best approximation for x from the space V_0, where

$$(i)\, V_0 = \mathcal{P}_1, \qquad (ii)\, V_0 = \mathcal{P}_2, \qquad (iii)\, V_0 = \mathcal{P}_3, \qquad (iv)\, V_0 = \mathcal{P}_4, \qquad (v)\, V_0 = \mathcal{P}_5.$$

2. Let $V = C[0, 2\pi]$ with inner product: $\langle f, g\rangle := \int_0^{2\pi} f(t)\overline{g(t)}dt$. Let $x(t) = t^2$. Find best approximation for x from the space V_0, where

$$V_0 = \mathrm{span}\{1, \sin t, \cos t, \sin 2t, \cos 2t\}.$$

♢

Exercise 3.41. Let $V = C[0, 1]$ with inner product: $\langle f, g\rangle := \int_0^1 f(t)\overline{g(t)}dt$ for $f, g \in C[0, 1]$. Let $x(t) = \sin t$. Find the best approximation for x from the subspace $V_0 := \mathrm{span}\{u_1, u_2, u_3\}\}$, where $u_1(t) = 1$, $u_2(t) = t$, $u_3(t) = t^2$. ♢

4 Eigenvalues and Eigenvectors

4.1 Definition and Examples

Definition 4.1. Let V be a vector space (over a field $\mathbb{F}$) and $T : V \to V$ be a linear operator. A scalar λ is called an **eigenvalue** of T if there exists a non-zero $x \in V$ such that

$$Tx = \lambda x,$$

and in that case x is called an **eigenvector** of T corresponding to the eigenvalue λ.

The set of all eigenvalues of T is called the **eigen-spectrum** or **point spectrum** of T, and we denote it by $\sigma_{\text{eig}}(T)$. $\diamondsuit$

Let $T : V \to V$ be a linear operator and $\lambda \in \mathbb{F}$. Observe:

- $\lambda \in \sigma_{\text{eig}}(T) \iff T - \lambda I$ is not one-one.
- A non-zero $x \in V$ is an eigenvector of T corresponding to $\lambda \in \sigma_{\text{eig}}(T) \iff x \in N(A-\lambda I)\setminus\{0\}$.
- The set of all eigenvectors T corresponding to $\lambda \in \sigma_{\text{eig}}(T)$ is the set $N(A - \lambda I) \setminus \{0\}$.

Definition 4.2. Let $T : V \to V$ be a linear operator and λ be an eigenvalue of T.

1. The subspace $N(T - \lambda I)$ of V is called the **eigenspace** of T corresponding to the eigenvalue λ.
2. $\dim[N(T - \lambda I)]$ is called the **geometric multiplicity** of λ.

$\diamondsuit$

Remark 4.3. If V is the zero space, then zero operator is the only operator on V, and it does not have any eigenvalue as there is no non-zero vector in V. $\diamondsuit$

Example 4.4. Let $A \in \mathbb{R}^{n\times n}$, and consider it as a linear operator from $\mathbb{R}^n$ to itself. We know that

- A is not one-one if and only if
- columns of A are linearly dependent if and only if
- $\det(A) = 0$.

Thus, $\lambda \in \sigma_{\text{eig}}(A) \iff \det(A - \lambda I) = 0$. $\diamondsuit$

4.2 Existence of eigenvalues

Note that for a given $A \in \mathbb{R}^{n\times n}$, there need not exist $\lambda \in \mathbb{R}$ such that $\det(A - \lambda I) = 0$. For example, consider $n = 2$ and

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

This matrix has no eigenvalues!

However, if $A \in \mathbb{C}^{n\times n}$, then, by the *fundamental theorem of algebra*, there exists $\lambda \in \mathbb{C}$ such that $\det(A - \lambda I) = 0$. Thus, in this case

$$\sigma_{\text{eig}}(A) \neq \varnothing.$$

Now, recall that if V is a finite dimensional vector space, say of dimension n, and $\{u_1, \ldots, u_n\}$ is a basis of V and if $T : V \to V$ is a linear transformation, then

- T is one-one $\iff$ columns of $[T]_{EE}$ are linearly independent,

and hence, in this case,

- $\lambda \in \sigma_{\text{eig}}(T) \iff \det([T]_{EE} - \lambda I) = 0.$

Note that the above equivalence is true for any basis E of V. Hence, eigenvalues of a linear operator T can be found by finding the zeros of the polynomial $\det([T]_{EE} - \lambda I)$ in $\mathbb{F}$. This also shows that:

THEOREM 4.5. *If V is a finite dimensional over an algebraically closed field $\mathbb{F}$, then every linear operator on V has atleast one eigenvalue*

Recall from *algebra* that $\mathbb{C}$ is an algebraically closed field, whereas $\mathbb{R}$ and $\mathbb{Q}$ are not algebraically closed.

We shall give a proof for the above theorem without relying on the concept of determinant. Before that let us observe that the conclusion in the above theorem need not hold if the space is infinite dimensional.

Example 4.6. (i) Let $V = \mathcal{P}$, the space of all polynomials over $\mathbb{F}$, which is either $\mathbb{R}$ or $\mathbb{C}$. Let

$$Tp(t) = tp(t), \quad p(t) \in \mathcal{P}.$$

Note that for $\lambda \in \mathbb{F}$ and $p(t) \in \mathcal{P}$,

$$Tp(t) = tp(t) \iff p(t) = 0.$$

Hence, $\sigma_{\text{eig}}(T) = \varnothing$.

(ii) Let $V = c_{00}$ and T be the right shift operator on V, i.e.,

$$T(\alpha_1, \alpha_2, \ldots) = (0, \alpha_1, \alpha_2, \ldots).$$

Then we see that $\sigma_{\text{eig}}(T) = \varnothing$. ♢

Proof of Theorem 4.5 independent of determinant. Let V be an n dimensional vector space over an algebraically closed field $\mathbb{F}$. Let x be a non-zero vector in V. If $Tx = 0$, then 0 is an eigenvalue. Assume that $Tx \neq 0$. Then we known that $\{x, Tx, \ldots, T^n x\}$ is linearly dependent, so that there exist $\alpha_0, \alpha_1, \ldots, \alpha_k$ in $\mathbb{F}$ with $k\{1, \ldots, n\}$ such that $a_k \neq 0$ and

$$\alpha_0 x + \alpha_1 Tx + \cdots + \alpha_k T^k x = 0,$$

i.e.,

$$(\alpha_0 I + \alpha_1 T + \cdots + \alpha_k T^k)x = 0.$$

Thus,

$$p(T)x = 0,$$

where $p(t) := \alpha_0 + \alpha_1 t + \cdots + \alpha_k t^k$. By fundamental theorem of algebra, there exist $\lambda_1, \ldots, \lambda_k$ in $\mathbb{F}$ such that

$$p(t) = \alpha_k (t - \lambda_1) \cdots (t - \lambda_k).$$

Since $p(T)x = 0$, we have

$$\alpha_k (T - \lambda_1 I) \cdots (T - \lambda_k I)x.$$

This shows that atleast one of $T - \lambda_1 I, \ldots, T - \lambda_k I$ is not one-one. Thus, at least one of $\lambda_1, \ldots, \lambda_k$ is an eigenvalue of T, and hence, $\sigma_{\text{eig}}(T) \neq \varnothing$. □

Can we show existence of an eigenvalue by imposing more conditions on the space V and the operator? Here is an answer in this respect.

THEOREM 4.7. *Let V be a non-zero finite dimensional inner product space over $\mathbb{F}$ which is either $\mathbb{R}$ or $\mathbb{C}$, and T be a self adjoint operator on V. Then $\sigma_{\text{eig}}(T) \neq \varnothing$, $\sigma_{\text{eig}}(T) \subseteq \mathbb{R}$.*

Proof. Let x be a non-zero vector in V such that $Tx = 0$. As in the proof of Theorem 4.5, let $p(t) := \alpha_0 + \alpha_1 t + \cdots + \alpha_k t^k$ be such that $\alpha_k \neq 0$ and

$$p(T)x = 0.$$

Let $\lambda_1, \ldots, \lambda_k$ in $\mathbb{C}$ be such that

$$p(t) = \alpha_k (t - \lambda_1) \cdots (t - \lambda_k).$$

If $\lambda_j \notin \mathbb{R}$ for some j, then we know that $\overline{\lambda}_j$ is also a zero of $p(t)$. So, there is ℓ such that $\lambda_\ell = \overline{\lambda}_j$. Writing $\lambda_j = \alpha_j + i\beta_j$ with $\alpha_j, \beta_j \in \mathbb{R}$ and $\beta_j \neq 0$, we have

$$(t - \lambda_j)(t - \lambda_\ell) = [t - (\alpha_j + i\beta_j)][t - (\alpha_j - i\beta_j)] = (t - \alpha_j)^2 + \beta_j^2.$$

Since $p(T)x = 0$, it follows that either there exists some m such that $\lambda_m \in \mathbb{R}$ and $T - \lambda_m I$ is not one-one or there exists some j such that $\lambda_j \notin \mathbb{R}$ and $(T - \alpha_j I)^2 + \beta_j^2 I$ is not one-one. In the first case, $\lambda_m \in \mathbb{R}$ is an eigenvalue. In the latter case, there exists $u \neq 0$ in V such that

$$[(T - \alpha_j I)^2 + \beta_j^2 I]u = 0.$$

Now, using the self adjointness of T,

$$\begin{aligned}\langle[(T-\alpha_j I)^2+\beta_j^2 I]u,u\rangle &= \langle(T-\alpha_j I)^2 u,u\rangle+\beta_j^2 I\langle u,u\rangle \\ &= \langle(T-\alpha_j I)u,(T-\alpha_j I)u\rangle+\beta_j^2 I\langle u,u\rangle.\end{aligned}$$

Since $u \neq 0$, it follows that $\beta_j = 0$ and $(T-\alpha_j I)u = 0$. Thus, T has a real eigenvalue.

Next, suppose that $\lambda \in \sigma_{\text{eig}}(T)$. If x is an eigenvector corresponding to λ, then we have

$$\lambda\langle x,x\rangle = \langle \lambda x,x\rangle = \langle Tx,x\rangle = \langle x,Tx\rangle = \langle x,\lambda x\rangle = \overline{\lambda}\langle x,x\rangle.$$

Hence, $\lambda \in \mathbb{R}$. □

THEOREM 4.8. *Eigenvectors corresponding to distinct eigenvalues of a linear operator are linearly independent.*

Proof. Let $\lambda_1,\ldots\lambda_n$ be eigenvalues of a linear operator $T : V \to V$ and let $u_1,\ldots,u_n$ be eigenvectors corresponding to $\lambda_1,\ldots,\lambda_n$, respectively. We prove the result by induction:

Let $n = 2$, and let α_1, α_2 such that $\alpha_1 u_1 + \alpha_2 u_2 = 0$. Then

$$T(\alpha_1 u_1 + \alpha_2 u_2) = 0, \quad \lambda_2(\alpha_1 u_1 + \lambda_2 \alpha_2 u_2) = 0$$

so that

$$\alpha_1 \lambda_1 u_1 + \alpha_2 \lambda_2 u_2 = 0 \quad (i), \quad \alpha_1 \lambda_2 u_1 + \alpha_2 \lambda_2 u_2 = 0. \quad (ii)$$

Hence, $(ii) - (i)$ implies

$$\alpha_1(\lambda_2 - \lambda_1)u_1 = 0.$$

Since $\lambda_2 \neq \lambda_1$ we have $\alpha_1 = 0$. Hence, from the equation $\alpha_1 u_1 + \alpha_2 u_2 = 0$, we obtain $\alpha_2 = 0$. Next, assume that the result is true for $n = k$ for some $k \in N$, $2 \leq k < n$. Let $\alpha_1,\ldots,\alpha_{k+1}$ be such that

$$\alpha_1 u_1 + \ldots + \alpha_{k+1} u_{k+1} = 0. \qquad (iii)$$

Since

$$T(\alpha_1 u_1 + \ldots + \alpha_{k+1} u_{k+1}) = 0, \quad \lambda_n(\alpha_1 u_1 + \ldots + \alpha_{k+1} u_{k+1}) = 0,$$

we have

$$\alpha_1 \lambda_1 u_1 + \ldots + \alpha_{k+1}\lambda_{k+1} u_{k+1}) = 0 \quad (iv), \quad \alpha_1 \lambda_n u_1 + \ldots + \alpha_{k+1}\lambda_{k+1} u_{k+1}) = 0. \quad (v)$$

Hence, $(v) - (iv)$ implies

$$\alpha_1(\lambda_1 - \lambda_{k+1})u_1 + \ldots + \alpha_k(\lambda_k - \lambda_{k+1})u_k = 0.$$

By induction assumption, $u_1,\ldots,u_k$ are linearly independent. Since $\lambda_1,\ldots,\lambda_k,\lambda_{k+1}$ are distinct, it follows that $\alpha_1 = 0, \alpha - 2 =, \ldots, \alpha_k = 0$. Hence, from (iii), $\alpha_{k+1} = 0$ as well. This completes the proof. □

LEMMA 4.9. *Let V be a non-zero finite dimensional inner product space over $\mathbb{F}$ which is either $\mathbb{R}$ or $\mathbb{C}$, and T be a normal operator on V. Let $\lambda \in \mathbb{F}$ and $x \in V$. Then*

$$Tx = \lambda x \iff T^*x = \bar{\lambda}x.$$

Proof. Since T is normal, i.e., $T^*T = TT^*$, it can be seen that $T - \lambda I$ is also a normal operator. Indeed,

$$(T - \lambda I)(T^* - \bar{\lambda}I) = TT^* - \bar{\lambda}T - \lambda T^* + |\lambda|^2 I = T^*T - \bar{\lambda}T - \lambda T^* + |\lambda|^2 I = (T^* - \bar{\lambda}I)(T - \lambda I).$$

Thus,

$$\begin{aligned}
\|(T^* - \bar{\lambda}I)x\|^2 &= \langle (T^* - \bar{\lambda}I)x, (T^* - \bar{\lambda}I)x\rangle \\
&= \langle (T - \lambda I)(T^* - \bar{\lambda}I)x, x\rangle \\
&= \langle (T^* - \bar{\lambda}I)(T - \bar{\lambda}I)x, x\rangle \\
&= \langle (T - \bar{\lambda}I)x, (T - \lambda I)x\rangle \\
&= \|(T - \lambda I)x\|^2.
\end{aligned}$$

Hence, $Tx = \lambda x \iff T^*x = \bar{\lambda}x$. □

THEOREM 4.10. *Let V be a non-zero finite dimensional inner product space over $\mathbb{F}$ which is either $\mathbb{R}$ or $\mathbb{C}$, and T be a normal operator on V. Then eigenvectors associated with distinct eigenvalues are orthogonal. In particular,*

$$\lambda \neq \mu \implies N(T - \lambda I) \perp N(T - \mu I).$$

Proof. Let T be a normal operator and let λ and μ be distinct eigenvalues of T with corresponding eigenvectors x and y, respectively. Then

$$\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle Tx, y\rangle. = \langle x, T^*y\rangle = \langle x, \bar{\mu}y\rangle = \mu\langle x, y\rangle$$

so that

$$(\lambda - \mu)\langle x, y\rangle = 0.$$

Since $\lambda \neq \mu$, we have $\langle x, y\rangle = 0$. □

4.3 Diagonalizability

We observe:

> If V is a finite dimensional vector space and T be a linear operator on V such that there is a basis E for V consisting of eigenvectors of T, then $[T]_{EE}$ is a diagonal matrix.

In view of the above observation we have the following definition.

Definition 4.11. Let V be a finite dimensional vector space and T be a linear operator on V. Then T is said to be diagonalizable if there is a basis E for V consisting of eigenvectors of T such that $[T]_{EE}$ is a diagonal matrix. ♢

THEOREM 4.12. *Let V be a finite dimensional vector space and T be a linear operator on V. Then T is diagonalizable if and only if there are distinct $\lambda_1, \ldots, \lambda_k$ in $\mathbb{F}$ such that*

$$V = N(T - \lambda_1 I) + \cdots + N(T - \lambda_k I).$$

Look at the following example.

Example 4.13. Consider the matrix

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

We observe that A as a linear operator on $\mathbb{R}^2$ has only one eigenvalue which is 0 and its geometric multiplicity is 1. Hence there is no basis for $\mathbb{R}^2$ consisting of eigenvectors of A. Hence, the above operator is not diagonalizable. ♢

Remark 4.14. Let V be an n- dimensional vector space and T be a linear operator on V. Suppose T is diagonalizable. Let $\{u_1, \ldots, u_n\}$ be a basis of V consisting of eigenvectors of T, and let $\lambda_j \in \mathbb{F}$ be such that $u_j = \lambda_j u_j$ for $j = 1, \ldots, n$. Let use the notation $U := [u_1, \ldots, u_n]$ for a map from $\mathbb{F}^n$ to V defined by

$$[u_1, \ldots, u_n] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha_1 u_1 + \cdots \alpha_n u_n.$$

Then we have

$$TU = T[u_1, \ldots, u_n] = [Tu_1, \ldots, Tu_n] = [\lambda_1 u_1, \ldots, \lambda_n u_n].$$

Thus, using the standard basis $\{e_1, \ldots, e_n\}$ of $\mathbb{F}^n$, we have

$$TUe_j = \lambda_j e_j, \quad j = 1, \ldots, n.$$

Thus,

$$TU = U\Lambda,$$

equivalently,

$$U^{-1}TU = \Lambda,$$

where $\Lambda := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. If T itself is an $n \times n$-matrix, then the above relation shows that T is similar to a diagonal matrix. ♢

Under what condition on the space V and operator T can we say that T is diagonalizable?

THEOREM 4.15. *Let V be a finite dimensional vector space, say $\dim(V) = n$, and T be a linear operator on V.*

(i) If T has n distinct eigenvalues, then T is diagonalizable.

(ii) If T has an eigenvalue λ such that $N(T-\lambda I)$ is a proper subspace of $N(T-\lambda I)^2$, then T is not diagonalizable.

Proof. (i) Follows from Theorem 4.8.

(ii) Assume for a moment that T is diagonalizable. Then by Theorem 4.12, there are distinct $\lambda_1, \ldots, \lambda_k$ in $\mathbb{F}$ such that

$$V = N(T-\lambda_1 I) + \cdots + N(T-\lambda_k I).$$

Let $x \in N(T-\lambda_1)^2$, and let $x_j \in N(T-\lambda_j)$ be such that

$$x = x_1 + \cdots + x_k.$$

Then

$$(T-\lambda_1 I)x = (T-\lambda_1 I)x_1 + \cdots + (T-\lambda_1 I)x_k.$$

We observe that $(T-\lambda_1 I)x \in N(T-\lambda_1 I)$ and $(T-\lambda_1 I)x_j \in N(T-\lambda_j I)$ for $j = 1, \ldots, k$. Hence, $(T-\lambda_1 I)(x-x_1) = 0$. Consequently, $x \in N(T-\lambda_1 I)$. Since $N(T-\lambda_1) \subseteq N(T-\lambda_1)^2$, we obtain that $N(T-\lambda_1 I)^2 = N(T-\lambda_1 I)$. Similarly, we have $N(T-\lambda_j I)^2 = N(T-\lambda_j I)$ for $j = 1, \ldots, k$. □

In view of the above theorem, we introduce the following definition.

Definition 4.16. An eigenvalue λ of a linear operator $T : V \to V$ is said to be **defective** if $N(T-\lambda I)$ is a proper subspace of $N(T-\lambda I)^2$. ♢

THEOREM 4.17. *Let T be a self-adjoint operator on an inner product space V. Then every eigenvalue of T is non-defective.*

Proof. Since T is self-adjoint, for $x \in V$,

$$\langle (T-\lambda I)^2 x, x\rangle = \langle (T-\lambda I)x, (T-\lambda I)x\rangle.$$

Hence, $N(T-\lambda I)^2 = N(T-\lambda I)$. □

Still it is not clear from whatever we have proved whether a self-adjoint operator on a finite dimensional space is diagonalizable or not. We shall take up this issue in the next section. Before that let us observe some facts:

- For any linear operator $T : V \to V$,

$$\{0\} \subseteq N(T) \subseteq N(T^2) \subseteq N(T) \cdots \subseteq N(T^n) \subseteq N(T) \cdots.$$

- If there exists $k \in \mathbb{N}$ such that
$$N(T^k) = N(T^{k+1})$$
then
$$N(T^k) = N(T^{k+j}) \quad \forall\, j \in \mathbb{N}.$$

- If V is finite dimensional and $N(T) \neq \{0\}$, then there exists $k \in \mathbb{N}$ such that
$$N(T^{k-1}) \neq N(T^k) = N(T^{k+j}) \quad \forall\, j \in \mathbb{N}.$$

Definition 4.18. Let V be finite dimensional space and λ be an eigenvalue of T. Then the number
$$\ell := \min\{k : N(T - \lambda I)^{k-1} \neq N(T - \lambda I)^k = N(T - \lambda I)^{k+1}\}$$
is called the **ascent** or **index** of λ. $\diamondsuit$

Note that:

- If ℓ is the ascent of an eigenvalue λ, then
$$N(T - \lambda I)^\ell = \bigcup_{k=1}^{\infty} N(T - \lambda I)^k.$$

Definition 4.19. Let V be finite dimensional space and λ be an eigenvalue of T with ascent ℓ. Then the space $N(T - \lambda I)^\ell$ is called the **generalized eigen-space** of T corresponding to the eigenvalue λ. Members of a generalized eigen-space are called **generalized eigenvectors**. $\diamondsuit$

4.4 Spectral representation of self adjoint operators

A natural question is whether every self-adjoint operator on a finite dimensional inner product space is diagonalizable. The answer is in affirmative. In order to prove this, we shall make use of a definition and a preparatory lemma.

Definition 4.20. Let V be a vector space and T be a linear operator on V. A subspace V_0 of V is said to be **invariant** under T if $T(V_0) \subseteq V_0$, that is, for every $x \in V$, $x \in V_0 \Longrightarrow Tx \in V_0$, and in that case, we say that V_0 is an **invariant subspace** of T. $\diamondsuit$

LEMMA 4.21. *Let T be a self-adjoint operator on an inner product space V. Let V_0 of V be an invariant subspace of T. Then*

(i) $V_0^\perp$ is invariant under T,

(ii) $T_0 := T_{|_{V_0}} : V_0 \to V_0$, the restriction of T to V_0, in self-adjoint.

Proof. (i) Suppose V_0 is invariant under T. Then for every $x \in V_0^\perp$ and $u \in V_0$, we have $Tu \in V_0$, and hence,

$$\langle Tx, u\rangle = \langle x, Tu\rangle = 0$$

so that $Tx \in V_0^\perp$.

(ii) For every $x, y \in V_0$, we have

$$\langle T_0x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \langle x, T_0y\rangle.$$

This completes the proof. □

THEOREM 4.22. (Spectral representation) *Let T be a self-adjoint operator on a finite dimensional inner product space V, say of dimension n. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. Then*

$$V = N(T - \lambda_1 I) + \cdots + N(T - \lambda_k I).$$

Further, there exists a linear operator $U : \mathbb{F}^n \to V$ such that $U^*U = I_n$, $UU^* = I_V$ and $[T]_{EE} = U^*TU$ is a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_k$ such that λ_j repeated $n_j := \dim(T - \lambda_j I)$ times for $j = 1, \ldots, k$.

Proof. Let $V_0 = N(T - \lambda_1 I) + \cdots + N(T - \lambda_k I)$. By Projection Theorem,

$$V = V_0 + V_0^\perp.$$

Its enough to show that $V_0^\perp = \{0\}$. Suppose $V_0^\perp \neq \{0\}$. By Lemma 4.21, $V_0^\perp$ is invariant under T and the operator $T_1 := T_{|_{V_0^\perp}} : V_0^\perp \to V_0^\perp$, the restriction of T to $V_0^\perp$, is self-adjoint. By theorem 4.7, T has an eigenvalue $\lambda \in \mathbb{R}$. Let $x \in V_0^\perp$ be a corresponding eigenvector. Now, since $\lambda_1 x = T_1 x = Tx$, $\lambda \in \{\lambda_1, \ldots, \lambda_k\}$. Without loss of generality, assume that $\lambda = \lambda_1$. Then $x \in N(T - \lambda_1 I) \subseteq V_0$. Thus, $x \in V_0^\perp \cap V_0 = \{0\}$, a contradiction. Hence, $V_0^\perp = \{0\}$, and

$$V = N(T - \lambda_1 I) + \cdots + N(T - \lambda_k I).$$

To see the remaining part, for each $j \in \{1, \ldots, k\}$, let $\{u_{j1}, \ldots, u_{jn_j}\}$ be an ordered orthonormal basis of $N(T - \lambda_j I)$. Then we see that

$$E = \{u_{11}, \ldots, u_{1n_1}, u_{21}, \ldots, u_{2n_2}, \ldots, u_{k1}, \ldots, u_{kn_k}\}$$

is an ordered orthonormal basis for V. To simplify the notation, let us write the above ordered E as $\{u_1, \ldots, u_n\}$ and μ_i, $i = 1, \ldots, n$ such that $\mu_{n_{j-1}+i} = \lambda_i$ for $i = 1, \ldots, n_j$ with $n_0 = 0$ and $j = 1, \ldots, k$. Let $J : V \to \mathbb{F}^n$ be the canonical isomorphism defined by

$$J(x) = [x]_E, \quad x \in V.$$

Then, we have $J^* = J^{-1}$ and $U := J^*$ satisfies

$$U^*U = JJ^{-1} = I_V, \quad UU^* = J^{-1}J = I_n, \quad U^*TU = JTJ^{-1} = A := [T]_{EE}.$$

Further,

$$A\underline{e}_j = JTJ^{-1}\underline{e}_j = JTu_j = J(\mu_j u_j) = \mu_j Ju_j = \mu_j \underline{e}_j.$$

Thus, $A := [T]_{EE}$ is a diagonal matrix with diagonal entries $\mu_1, \ldots, \mu_n$. $\square$

Remark 4.23. Recall that the U introduced in the proof of Theorem 4.22 is same as the operator introduced in Remark 4.14, namely,

$$U = [u_{11}, \ldots, u_{1n_1}, u_{21}, \ldots, u_{2n_2}, \ldots, u_{k1}, \ldots, u_{kn_k}].$$

$\diamondsuit$

COROLLARY 4.24. (Spectral representation) *Let T be a self-adjoint operator on a finite dimensional inner product space V, say of dimension n. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. For each i, let $\{u_{i1}, \ldots, u_{in_i}\}$ be an ordered orthonormal basis of $N(T - \lambda_i I)$. Then*

$$Tx = \sum_{i=1}^{k}\sum_{i=1}^{n_i} \lambda_i \langle x, u_{ij}\rangle u_{ij}, \quad x \in V.$$

COROLLARY 4.25. (Spectral representation) *Let T be a self-adjoint operator on a finite dimensional inner product space V, say of dimension n. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. For each $i \in \{1, \ldots, k\}$, let P_j be the orthogonal projection onto $N(T - \lambda_i I)$. Then*

$$T = \sum_{i=1}^{k} \lambda_i P_i.$$

COROLLARY 4.26. (Diagonal representation) *Let $A \in \mathbb{F}^{n\times n}$ be a self adjoint matrix (i.e., hermitian if $\mathbb{F} = \mathbb{C}$ and symmetric if $\mathbb{F} = \mathbb{R}$). Then there exists a unitary matrix $U \in \mathbb{F}^{n\times n}$ such that U^*TU is a diagonal matrix.*

4.5 Singular value representation

Let T be a linear operator on a finite dimensional inner product space V. The we know that T^*T is a self adjoint operator. By spectral theorem, we know that V has an orthonormal basis $E; \{u_1, \ldots, u_n\}$ consisting of eigenvectors of T^*T, and if $T^*Tu_j = \lambda_j u_j$ for $j = 1, \ldots, n$ (where λ_j's need not be distinct), then

$$T^*Tx = \sum_{j=1}^{n} \lambda_j \langle x, u_j\rangle u_j, \quad x \in V.$$

Note that

$$\lambda_j = \lambda_j \langle u_j, u_j\rangle = \langle \lambda_j u_j, u_j\rangle = \langle T^*Tu_j, u_j\rangle = \langle Tu_j, Tu_j\rangle = \|Tu_j\|^2 \geq 0.$$

Let $\lambda_1, \ldots, \lambda_k$ be the nonzero (positive) numbers among $\lambda_1, \ldots, \lambda_n$. For $j \in \{1, \ldots, k\}$, let us write $\lambda_j = s_j^2$, where s_j is the positive square-root of λ_j. Thus, writing $v_j = \frac{Tu_j}{s_j}$, we obtain

$$Tu_j = s_j v_j, \quad T^*v_j = s_j u_j.$$

Further, since $x = \sum_{j=1}^{k} \langle x, u_j \rangle u_j$, we have

$$Tx = \sum_{j=1}^{k} \langle x, u_j \rangle T u_j = \sum_{j=1}^{k} s_j \langle x, u_j \rangle v_j. \tag{1}$$

Also,

$$\langle x, T^* y \rangle = \langle Tx, y \rangle = \left\langle \sum_{j=1}^{k} s_j \langle x, u_j \rangle v_j, y \right\rangle = \sum_{j=1}^{k} s_j \langle x, u_j \rangle \langle v_j, y \rangle = \left\langle x, \sum_{j=1}^{k} s_j \langle y, v_j \rangle u_j \right\rangle.$$

Hence,

$$T^* y = \sum_{j=1}^{k} s_j \langle y, v_j \rangle u_j. \tag{2}$$

Observe that

$$s_j \langle v_i, v_j \rangle = \langle v_i, s_j v_j \rangle = \langle v_i, T u_j \rangle = \langle T^* v_i, u_j \rangle = \langle s_i u_i, u_j \rangle = s_i \langle u_i, u_j \rangle.$$

Therefore, $\{v_j : j = 1, \ldots, k\}$ is an orthonormal set. From the representations (1) and (2), it can be seen that

- $\{u_1, \ldots, u_k\}$ is an orthonormal basis of $N(T)^\perp$, and
- $\{v_1, \ldots, v_k\}$ is an orthonormal basis of $R(T)$.

Definition 4.27. The numbers $s_1, \ldots, s_n$ are called the **singular values** of T and the set $\{(s_j, u_j, v_j) : j = 1, \ldots, n\}$ is called the singular system for T.

The representations (1) and (2) above are called the **singular value representations** of T and T^*, respectively. ♢

If we write

$$U_0 = [u_1, \ldots, u_k], \quad V_0 = [v_1, \ldots, v_k]$$

as the operators on $\mathbb{F}^k$ defied as in Remark 4.14, then, in view of the relations $T u_j = s_j v_j$ and $T^* v_j = s_i u_j$, we have

$$TU_0 = V_0 S_0, \quad T^* V_0 = U_0 S,$$

where $S_0 = \operatorname{diag}(s_1, \ldots, s_k)$. Suppose $n > k$. If we extend the orthonormal sets $\{u_1, \ldots, u_k\}$ and $\{v_1, \ldots, v_k\}$ to orthonormal bases $\{u_1, \ldots, u_n\}$ and $\{v_1, \ldots, v_n\}$, then for $j = k + 1, \ldots, n$, $u_j \in n(T)$ and $v_j \in R(T)^\perp$ so that,since $R(T)^\perp = N(T^*)$, we obtain

$$TU = VS, \quad T^* V = US,$$

where

$$U = [u_1, \ldots, u_n], \quad V_0 = [v_1, \ldots, v_n], \quad S = \operatorname{diag}(s_1, \ldots, s_n),$$

with $s_j = 0$ for $j > k$. Thus, we have

$$V^* TU = S, \quad U^* T^* V = S.$$

4.6 Spectral decomposition

Throughout this section we assume that V is a finite dimensional space over $\mathbb{C}$ and $T : V \to V$ is a linear operator.

In the following, if V_1 and V_2 are subspaces of V, then by $V_1 \oplus V_2$ we mean $V_1 + V_2$ whenever $V_1 \cap V_2 = \{0\}$.

The main theorem, in this section, is the following.

THEOREM 4.28. *Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T with ascents $\ell_1, \ldots, \ell_k$ be the ascents of $\lambda_1, \ldots, \lambda_k$, respectively. Then*

$$V = N(T - \lambda_1 I)^{\ell_1} \oplus \cdots \oplus N(T - \lambda_k I)^{\ell_k}.$$

where each $N(T - \lambda_j I)^{\ell_j}$ is invariant under T. In particular, T is diagonalizable if and only if ascent of each eigenvalue of T is 1.

- Since ascent of each eigenvalue of a self adjoint operator on an inner product space, an immediate corollary of the above theorem is Theorem 4.22.

For proving Theorem 4.28, we shall make use of the following lemma.

LEMMA 4.29. *Let V be a finite dimensional vector space and $T : V \to V$ be a linear operator. Let λ be an eigenvalue of T with ascent ℓ. Then the following hold.*

1. *For every $j \in \mathbb{N}$, $N(T - \lambda I)^j$ and $R(T - \lambda I)^j$ are invariant under T.*
2. *$V = N(T - \lambda I)^\ell \oplus R(T - \lambda I)^\ell$.*
3. *λ is an eigenvalue of $T_0 := T_{|_{N(T-\lambda I)^\ell}}$, and λ is the only eigenvalue of T_0.*
4. *If $\mu \neq \lambda$, then for each $j \in \mathbb{N}$, $N(T - \mu I)^j \cap N(T - \lambda I)^\ell = \{0\}$.*

Proof. 1. Let $j \in \mathbb{N}$ and $x \in N(T - \lambda I)^j$. Then

$$(T - \lambda I)^j T x = T(T - \lambda I)^j x = 0 \Longrightarrow Tx \in N(T - \lambda I)^j.$$

Hence, $Tx \in N(T - \lambda I)^j$. Let $y \in R(T - \lambda I)^j$. Then $\exists\, x \in V$ such that $(T - \lambda I)^j x = y$. Hence,

$$Ty = T(T - \lambda I)^j x = (T - \lambda I)^j T x \in R(T - \lambda I)^j.$$

Hence, $Ty \in R(T - \lambda I)^j$.

2. Since $\dim(V) < \infty$ and since $\dim[N(T - \lambda I)^\ell] + \dim[R(T - \lambda I)^\ell] = \dim(V)$, it is enough to show that $N(T - \lambda I)^\ell \cap R(T - \lambda I)^\ell = \{0\}$.

Suppose $x \in N(T-\lambda I)^\ell \cap R(T-\lambda I)^\ell = \{0\}$. Then, $(T-\lambda I)^\ell x = 0$ and there exists $u \in V$ such that $x = (T-\lambda I)^\ell u$. Then $(T-\lambda I)^\ell x = (T-\lambda I)^{2\ell} u = 0$ so that $u \in N(T-\lambda I)^{2\ell} = N(T-\lambda I)^\ell$. Thus $x = (T-\lambda I)^\ell u = 0$.

3. Note that, if $0 \neq x \in N(T-\lambda I)$, then $x \in N(T-\lambda I)^\ell$ and hence $\lambda x = Tx = T_0 x$ so that λ is an eigenvalue of T_0. Next suppose that $\mu \in \mathbb{C}$ such that $\mu \neq \lambda$ and μ is an eigenvalue of T_0 with a corresponding eigenvector $y \in N(T-\lambda I)^\ell$. Then we have

$$0 = (T-\lambda I)^\ell y = (\lambda - \mu)^\ell y$$

which is a contradiction, since $\lambda \neq \mu$ and $y \neq 0$. Thus, λ is the only eigenvalue of T_0.

4. By (2), it is enough to show that $N(T-\mu I)^j \subseteq R(T-\lambda I)^\ell$. We shall prove this by induction.

Let $j = 1$ and $x \in N(T-\mu I)$. By (2), there exists $u \in N(T-\lambda I)^\ell$ and $v \in R(T-\lambda I)^\ell$ such that $x = u + v$. Then

$$0 = (T-\mu I)x = (T-\mu I)u + (T-\mu I)v.$$

Since $(T-\mu I)u \in N(T-\lambda I)^\ell$ and $(T-\mu I)v \in R(T-\lambda I)^\ell$, by (2) we have $(T-\mu I)u = 0$. Now, if $u \neq 0$, then it follows that, μ is also an eigenvalue of T_0, which is a contradiction, due to (3). Thus, $u = 0$ and $x = v \in R(T-\lambda I)^\ell$.

Next assume that $N(T-\mu I)^j \subseteq R(T-\lambda I)^\ell$ for some $j \geq 1$. We have to show that $N(T-\mu I)^{j+1} \subseteq R(T-\lambda I)^\ell$. So let $x \in N(T-\mu I)^{j+1}$. By (2), there exists $u \in N(T-\lambda I)^\ell$ and $v \in R(T-\lambda I)^\ell$ such that $x = u + v$. Then

$$0 = (T-\mu I)^{j+1}x = (T-\mu I)^{j+1}u + (T-\mu I)^{j+1}v.$$

Since $(T-\mu I)^{j+1}u \in N(T-\lambda I)^\ell$ and $(T-\mu I)^{j+1}v \in R(T-\lambda I)^\ell$, by (2) we have $(T-\mu I)^{j+1}u = 0$, i.e., $(T-\mu I)u \in N(T-\mu I)^j \subseteq N(T-\lambda I)^\ell$. But, by induction hypothesis, $N(T-\mu I)^j \subseteq R(T-\lambda I)^\ell$. Thus, $(T-\mu I)u \in N(T-\lambda I)^\ell \cap R(T-\lambda I)^\ell = \{0\}$. Thus, if $u \neq 0$, then μ is also an eigenvalue of T_0. which is a contradiction, due to (3). Thus, $u = 0$ and $x = v \in R(T-\lambda I)^\ell$. □

Proof of Theorem 4.28. In view of Lemma 4.29, it is enough to prove that V is spanned by generalized eigenvectors of T. We shall prove this by induction on dimension of V. The case of $\dim(V) = 1$ is obvious, for in this case, V is spanned by the eigenspace of T, as there is only one eigenvalue and the generalized eigenspace corresponding to that is the eigenspace which is the whole space. Next assume that the result is true for all vector spaces of dimension less than n, and let $\dim(V) = n$. Let λ be an eigenvalue of T with ascent ℓ. Then, by Lemma 4.29, $V = N(T-\lambda I)^\ell + R(T-\lambda I)^\ell$ where $\dim[R(T-\lambda I)^\ell] < n$. Let $\tilde{T} := T_{|_{R(T-\lambda I)^\ell}}$. By induction assumption, $R(T-\lambda I)^\ell$ is spanned by the generalized eigenvectors of $\tilde{T}$. But, generalized eigenvectors of $\tilde{T}$ are generalized eigenvectors of T as well. Thus both $N(T-\lambda I)^\ell$ and $R(T-\lambda I)^\ell$ are spanned by the generalized eigenvectors of T. This completes the proof. □

THEOREM 4.30. *Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T with ascents of $\lambda_1, \ldots, \lambda_k$, respectively. Let*

$$p(t) = (t-\lambda_1)^{\ell_1} \cdots (t-\lambda_k)^{\ell^k}.$$

Then,

$$p(T) = 0.$$

Further, if $q(t)$ is a polynomial satisfying $q(T) = 0$, then $p(t)$ divides $q(t)$.

Proof. Since $(T - \lambda_r I)^{\ell_r}$ and $(T - \lambda_s I)^{\ell_s}$ commute, it follows that

$$p(T)u = 0$$

for every $u \in N(T - \lambda_i I)^{\ell_i}$, $i = 1, \ldots, k$. Hence, by Theorem 4.28, $p(T)x = 0$ for every $x \in V$. Consequently, $p(T) = 0$.

Next, let $q(t)$ be a polynomial such that $q(T) = 0$. Let $\mu_1, \ldots, \mu_r$ be the distinct zeros of $q(t)$ so that

$$q(t) = a(t - \mu_1)^{n_1} \cdots (t - \mu_r)^{n_r}$$

for some $0 \neq a \in \mathbb{C}$. Since $q(T) = 0$, for each $j \in \{1, \ldots, k\}$, we have

$$a(T - \mu_1 I)^{n_1} \cdots (T - \mu_r I)^{n_r} u = 0 \quad \forall\, u \in N(T - \lambda_j I)^{\ell_j}. \tag{$*$}$$

Now, if $\mu_i \neq \lambda_j$, then we know that $(T - \mu_i I)^{n_i}$ is one-one on $N(T - \lambda_j I)^{\ell_j}$. Hence, it follows that there exists i such that $\mu_i = \lambda_j$ such that

$$(T - \lambda_j I)^{n_i} u = 0 \quad \forall\, u \in N(T - \lambda_j I)^{\ell_j}.$$

Taking $u \in N(T - \lambda_j I)^{\ell_j} \setminus N(T - \lambda_j I)^{\ell_j - 1}$, it follows that $n_i \geq \ell_j$. Thus, $\{\lambda_1, \ldots, \lambda_k\} \subseteq \{\mu_1, \ldots, \mu_r\}$. Without loss of generality, we can assume that
$m_j = \lambda_j$ so that $n_j \geq \ell_j$ for $j = 1, \ldots, k$. Thus, $p(t)$ divides $q(t)$. □

Definition 4.31. A monic polynomial $p(t)$ is called a **minimal polynomial** for T if $p(T) = 0$ and for any polynomial $q(t)$ with $q(T) = 0$, $p(t)$ divides $q(t)$. ♢

- Theorem 4.30 shows that if $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T with ascents of $\lambda_1, \ldots, \lambda_k$, respectively, then

  $$p(t) := (t - \lambda_1)^{\ell_1} \cdots (t - \lambda_k)^{\ell^k}$$

  is the minimal polynomial of T.

For the next definition we recall the concept of matrix representation:

Let V be a finite dimensional vector space, and let $E_1 := \{u_1, \ldots, u_n\}$ and $E_2 := \{v_1, \ldots, v_n\}$ be bases of V. Let $T : V \to V$ be a linear operator. Let Then

$$[T]_{E_1 E_1} = [J^{-1}]_{E_2 E_1} [T]_{E_2 E_2} [J]_{E_1 E_2} = [J]_{E_2 E_1}^{-1} [T]_{E_2 E_2} [J]_{E_1 E_2},$$

where $J : V \to V$ is the isomorphism defined by

$$J(\alpha_1 u_1 + \ldots + \alpha_n u_n) = \alpha_1 v_1 + \ldots + \alpha_n v_n.$$

Hence, we have

$$\det[T]_{E_1E_1} = \det[T]_{E_2E_2}.$$

Thus, determinant of the matrix representation of an operator is independent of the basis with respect to which it is represented.

Definition 4.32. Let E be a basis of V. The monic polynomial

$$q_T(t) := \det[tI - T]_{EE}$$

is called the **characteristic polynomial** of T, where E is any basis of V. $\diamondsuit$

4.7 Cayley-Hamilton theorem

We know that eigenvalues of T are the zeros of the characteristic polynomial $q_T(t)$. Thus, $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T if and only if

$$q_T(t) = (t - \lambda_1)^{n_1} \cdots (t - \lambda_k)^{n_k}$$

with $n_1, \ldots, n_k$ in $\mathbb{N}$ such that $n_1 + \cdots + n_k = n := \dim(V)$.

THEOREM 4.33. (Cayley–Hamilton theorem)

$$q_T(T) = 0.$$

Proof. Recall that for operators $T, T_1, T_2 : V \to V$ and $\alpha \in \mathbb{C}$,

$$[T_1 + T_2]_{EE} = [T_1]_{EE} + [T_2]_{EE}, \quad [\alpha T]_{EE} = \alpha [T]_{EE}.$$

Hence, if $q_T(t) = t^n + a_1 t^{n-1} + \ldots + a_{n-1} t + a_n$, then

$$\begin{aligned} [q_T(T)]_{EE} &= [T]^n_{EE} + a_1 [T]^{n-1}_{EE} + \ldots + a_{n-1}[T]_{EE} + a_n [I]_{EE} \\ &= q_T([T]_{EE}). \end{aligned}$$

Recall that, by the Cayley–Hamilton theorem for matrices, we have $q_T([T]_{EE}) = 0$. Therefore, $[q_T(T)]_{EE} = 0$ so that $q_T(T) = 0$. $\square$

Definition 4.34. Let λ be an eigenvalue of T and λ be an eigenvalue of T. Then the order of λ as a zero of the characteristic polynomial $q_T(t)$ is called the **algebraic multiplicity** of λ. $\diamondsuit$

THEOREM 4.35. *Let λ be an eigenvalue of T with ascent ℓ. Then $m := \dim[N(T - \lambda I)^\ell]$ is the algebraic multiplicity of λ.*

In order to prove the above theorem we make use of the following observation.

PROPOSITION 4.36. *Suppose V_1 and V_2 are invariant subspaces of a linear operator $T : V \to V$ such that $V = V_1 \oplus V_2$. Let $T_1 = T_{|V_1}$ and $T_2 = T_{|V_2}$. Then*

$$\det(T) = \det(T_1)\det(T_2).$$

Proof. Writing $x \in V$ as

$$x = x_1 + x_2 \quad \text{with} \quad x_1 \in V_1,\, x_2 \in V_2,$$

we have

$$Tx = T_1 x_1 + T_2 x_2.$$

Define $\tilde{T}_1, \tilde{T}_2 : V \to V$ by

$$\tilde{T}_1 x = T_1 x_1 + x_2, \quad \tilde{T}_2 x = x_1 + T_2 x_2.$$

Then we have

$$\tilde{T}_1 \tilde{T}_2 x = \tilde{T}_1(x_1 + T_2 x_2) = T_1 x_1 + T_2 x_2 = Tx.$$

Thus, with respect to any basis E of V, we have

$$[T]_{EE} = [\tilde{T}_1]_{EE} [\tilde{T}_2]_{EE}$$

and hence

$$\det(T) = \det(\tilde{T}_1) \det(\tilde{T}_2).$$

Next we show that

$$\det(\tilde{T}_1) = \det(T_1), \quad \det(\tilde{T}_2) = \det(T_2).$$

For this, let $E_1 = \{u_1, \ldots, u_r\}$ and $E_2 = \{u_{r+1}, \ldots, u_n\}$ be bases of V_1 and V_2 respectively. Consider the basis $E = E_1 \cup E_2$ for V. Then, we have

$$\tilde{T}_1 u_j = \begin{cases} T_1 u_j, & j = 1, \ldots, r, \\ u_j, & j = r+1, \ldots, s. \end{cases} \quad \text{and} \quad \tilde{T}_2 u_j = \begin{cases} u_j, & j = 1, \ldots, r, \\ T_2 u_j, & j = r+1, \ldots, s. \end{cases}$$

Hence, we obtain,

$$\det(\tilde{T}_1) = \det(T_1), \quad \det(\tilde{T}_2) = \det(T_2).$$

This completes the proof. □

Proof of Theorem 4.35. Let $K = N(T - \lambda I)^\ell$ and $R = R(T - \lambda I)^\ell$. We know that K and R are invariant under T and $V = K \oplus R$. Let $T_1 := T_{|_K}$ and $T_2 := T_{|_R}$. We know that λ is the only eigenvalue of T_1. Also, observe that λ is not an eigenvalue of T_2. Indeed, if $x \in R$ such that $T_2 x = \lambda x$, then $x \in N(T - \lambda I) \subseteq K$ so that $x = 0$. By Proposition 4.36,

$$det(tI - T) = \det(tI_1 - T_1) \det(tI_2 - T_2),$$

where I_1 and I_2 are identity operators on K and R respectively. Since $\det(\lambda I_2 - T_2) \neq 0$, it is clear that the algebraic multiplicity of λ as an eigenvalue of T is same as the algebraic multiplicity of λ as an eigenvalue of T_1. Since λ is the only eigenvalue of T_1, we obtain that $m := \dim(K)$ is the algebraic multiplicity of λ. □

Remark 4.37. Recall that if T is a self-adjoint operator on a finite dimensional inner product space, then we have

$$T = \sum_{i=}^{k} \lambda_i P_i$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T and $P_1, \ldots, P_k$ are the orthogonal projections onto the eigenspaces $N(T - \lambda_I), \ldots, N(T - \lambda_k I)$, respectively.

Next suppose that V is a finite dimensional vector space and T is a diagonalizable operator. Again let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. We know that

$$V = N(T - \lambda_I) \oplus \cdots N(T - \lambda_k I).$$

Hence, every $x \in V$ can be written uniquely as

$$x = x_1 + \cdots + x_k \quad \text{with} \quad x_i \in N(T - \lambda_i I).$$

For $i = 1, \ldots, k$, let $P_i : V \to V$ be defined by

$$P_i x = x_i, \quad x \in V.$$

The, it can be easily seen that $P_i^2 = P_i$ so that P_i is a projection onto $N(T - \lambda_i I)$. Hence,

$$I = P_1 + \cdots + P_k$$

and

$$T = TP_1 + \cdots + TP_k = \sum_{i=1}^{k} \lambda_i P_i.$$

Next, consider any linear operator finite dimensional vector space over $\mathbb{C}$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T with ascents $\ell_, \ldots, \ell_k$, respectively. Then, by spectral decomposition theorem, we have

$$V = N(T - \lambda_I) \oplus \cdots N(T - \lambda_k I).$$

Hence, every $x \in V$ can be written uniquely as

$$x = x_1 + \cdots + x_k \quad \text{with} \quad x_i \in N(T - \lambda_i I)^{\ell_i}.$$

Again, for $i = 1, \ldots, k$, let $P_i : V \to V$ be defined by

$$P_i x = x_i, \quad x \in V.$$

Then we have

$$I = P_1 + \cdots + P_k$$

and

$$T = TP_1 + \cdots + TP_k = \sum_{i=1}^{k} \lambda_i P_i + \sum_{i=1}^{k} (T - \lambda_i I) P_i.$$

Let $D_i = (T - \lambda_i I) P_i$. Then we see that

$$D_i^{\ell_i} = 0 \quad \text{and} \quad D^{\ell_i - 1} \neq 0.$$

Thus, D_i is a *nilpotent operator* of index ℓ_i for $i = 1, \ldots, k$. $\diamondsuit$

4.8 Triangulization and Jordan representation

As in last section, we assume that V is a finite dimensional space over $\mathbb{C}$ and $T : V \to V$ is a linear operator.

THEOREM 4.38. (Triangulization) *There exists a basis E for V such that $[T]_{EE}$ is a triangular matrix.*

Proof. First let us assume that T has only one eigenvalue λ with ascent ℓ. Then $V = N(T - \lambda I)^\ell$. If $\ell = 1$, then the result is obvious. In fact, in this case T is diagonalizable. So, assume that $\ell > 1$. Let

$$K_j = N(T - \lambda I)^j \quad \text{and} \quad g_j = \dim(K_j), \quad j = 1, \ldots, \ell.$$

Then, we have $K_\ell = V$ and K_j is a proper subspace of K_{j+1} for $j = 1, \ldots, \ell - 1$. Let $E = \{u_1, \ldots, u_n\}$ be a basis of V such that $\{u_1, \ldots, u_{g_j}\}$ is a basis of K_j for $j = 1, \ldots, \ell$. Then, $\{u_1, \ldots, u_{g_1}\}$ is a basis of $K_1 := N(T - \lambda I)$ and

$$\{u_{g_j+1}, \ldots, u_{g_j}\} \subseteq K_{j+1} \setminus K_j, \quad j \in \{1, \ldots, \ell - 1\}.$$

Further,

$$\mathrm{span}(\{u_{g_j+1}, \ldots, u_{g_{j+1}}\}) \cap K_j = \{0\}.$$

Note that for each $k \in \{1, \ldots, n\}$,

$$Tu_k = \lambda u_k + (T - \lambda I)u_k.$$

Clearly, $Tu_k = \lambda u_k$ for $k = 1, \ldots, g_1$. If $k \in \{g_1 + 1, \ldots, n\}$, then there exists $j \in \{1, \ldots, \ell - 1\}$ such that $k \in \{g_j + 1, \ldots, g_{j+1}\}$, i.e., k is such that $u_k \in \{u_{g_j+1}, \ldots, u_{g_{j+1}}\}$. Then we have $(T - \lambda I)u_k \in K_j$ so that Tu_k takes the form

$$Tu_k = \lambda u_k + \sum_{i=1}^{g_j} \alpha_i^{(k)} u_i.$$

Thus, $[T]_{EE}$ is a triangular matrix with every diagonal entry λ.

Next assume that the distinct eigenvalues of T are $\lambda_1, \ldots, \lambda_r$ with $\ell_1, \ldots, \ell_r$, respectively. Let

$$V_j := N(T - \lambda I)^{\ell_j}, \quad j = 1, \ldots, r.$$

Let $T_j : V_j \to V_j$ be the restriction of T to V_j. Then we know that λ_j is the only eigenvalue of T_j. Let E_j be a basis for V_j such that $A_j := [T_j]_{E_j E_j}$ is a triangular diagonal matrix with diagonal entries λ_j. Now, taking $E = \cup_{j=1}^r E_j$, we it follows that E is a basis of V and $[T]_{EE}$ has block diagonal form with blocks $A_1, \ldots, A_r$. $\square$

THEOREM 4.39. (Jordan form) *There exists a basis E such that $[T]_E = (a_{ij})$, where*

$$a_{ii} \in \{\lambda_1, \ldots, \lambda_k\}, \quad a_{ij} = \begin{cases} 0 & \textit{if } j < i \textit{ and } j > i + 1, \\ 0 \textit{ or } 1 & \textit{if } j = i + 1 \end{cases}$$

Proof. In view of the fact that each $N(T-\lambda_j I)^{\ell_j}$ is invariant under T and the spectral decomposition theorem (Theorem 4.28), it is enough to consider the case of T having only one eigenvalue.

So, let λ be the only eigenvalue of T with ascent ℓ. Then $V = N(T-\lambda I)^\ell$. If $\ell = 1$, then we are done. In fact, in this case T is diagonalizable. So, assume that $\ell > 1$, and let

$$K_j = N(T-\lambda I)^j \quad \text{and} \quad g_j := \dim(K_j) \quad \text{for} \quad j \in \{1, \ldots, \ell\}.$$

Then for $j \in \{1, \ldots, \ell-1\}$, K_j is a proper subspace of K_{j+1}. Let $K_{j+1} = K_j \oplus Y_{j+1}$, where Y_{j+1} is spanned by $K_{j+1} \setminus K_j$. Let $h_1 = g_1$ and for $j = 1, \ldots, \ell-1$, let $h_{j+1} = g_{j+1} - g_j$. Thus, $h_{j+1} = \dim(Y_{j+1})$, $j = 1, \ldots, \ell-1$, and $h_1 + \cdots + h_\ell = g_\ell = \dim(V)$.

The idea is to identify linearly independent vectors $u_j^{(i)}$, $j = 1, \ldots, h_i$, in $K_i \setminus K_{i-1}$ for each $i = 1, \ldots, \ell$ so that their union is the basis of V with respect to which T has a the required form.

Now, let $u_1^{(\ell)}, \ldots, u_{h_\ell}^{(\ell)}$ be a basis of Y_ℓ. Let us observe that following:

1. $(T-\lambda I)u_1^{(\ell)}, \ldots, (T-\lambda I)u_{h_\ell}^{(\ell)}$ are linearly independent, and
2. $(T-\lambda I)u_j^{(\ell)} \in K_{\ell-1} \setminus K_{\ell-2}$ for $j = 1, \ldots, h_\ell$, whenever $\ell > 2$.

Let $\alpha_1, \ldots, \alpha_{h_\ell} \in \mathbb{C}$ be such that $\sum_{i=1}^{h_\ell} \alpha_i (T-\lambda I)u_i^{(\ell)} = 0$. Then

$$\sum_{i=1}^{h_\ell} \alpha_i u_i^{(\ell)} \in N(T-\lambda I) \subseteq K_{\ell-1}.$$

Hence, $\sum_{i=1}^{h_\ell} \alpha_i u_i^{(\ell)} \in K_{\ell-1} \cap Y_\ell = \{0\}$ so that $\alpha_i = 0$ for $i = 1, \ldots, h_\ell$. Thus, (1) is proved. To see (2), first we observe that $(T-\lambda I)u_j^{(\ell)} \in K_{\ell-1}$. Suppose $(T-\lambda I)u_j^{(\ell)} \in K_{\ell-2}$ for some j. Then $u_j^{(\ell)} \in K_{\ell-1} \cap Y_\ell = \{0\}$, which is not possible. This proves (2).

Now, let us denote

$$u_j^{(\ell-1)} = (T-\lambda I)u_j^{(\ell)}, \quad j = 1, \ldots, h_\ell.$$

Find $u_j^{(\ell-1)} \in K_{\ell-1} \setminus K_{\ell-2}$ for $j = h_\ell + 1, \ldots, h_{\ell-1}$ so that $u_j^{(\ell-1)}$, $j = 1, \ldots, h_{\ell-1}$ are linearly independent. Continuing this procedure to the next level downwards, we obtain a basis for V as

$$E = E_\ell \cup E_{\ell-1} \cup \cdots \cup E_1, \quad E_i := \{u_j^{(i)} : j = 1, \ldots, h_i\}. \tag{$*$}$$

Note that

$$h_1 + h_2 + \cdots + h_\ell = g_1 + (g_2 - g_1) + \cdots + (g_\ell - g_{\ell-1}) = g_\ell.$$

Also, $Tu_j^{(1)} = \lambda u_j^{(1)}$, $j = 1, \ldots, h_1 = g_1$ and for $i > 1$,

$$Tu_j^{(i)} = \lambda u_j^{(i)} + (T-\lambda I)u_j^{(i)} = \lambda u_j^{(i)} + u_j^{(i-1)}, \quad j = 1, \ldots, h_i.$$

Reordering the basis vectors in E appropriately, we obtain the required form of the matrix representation of T. Note that at the upper off-diagonal of $[T]_E$ there are $g_1 - 1$ number of 0's and $g_\ell - g_1$ number of 1's. □

5 Problems

5.1 (On Section 1: Vector spaces)

In the following, V denotes a vector space over a filed $\mathbb{F}$.

For $i, j \in \mathbb{N}$, we denote $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$

1. Prove that $0x = 0_V$ and $(-1)x = -x$ for all $x \in V$.

2. Prove that for $x \in V$ and $\alpha \in \mathbb{F}$, if $\alpha x = 0$, then either $\alpha = 0$ or $x = 0$.

3. Verify (prove) the following:

   (a) $\mathbb{R}^n$ with coordinate-wise addition and scalar multiplication is a vector space over $\mathbb{R}$ and over $\mathbb{Q}$, but not a vector space over $\mathbb{C}$.

   (b) $\mathbb{F}^n$ with coordinate-wise addition and scalar multiplication is a vector space over $\mathbb{F}$ but not a vector space over a field $\widetilde{\mathbb{F}} \supseteq \mathbb{F}$ with $\widetilde{\mathbb{F}} \neq \mathbb{F}$.

   (c) $\mathbb{R}^{m\times n}$, the set of all real $m \times n$ matrices is a vector space over $\mathbb{R}$ under usual matrix multiplication and scalar multiplication.

   (d) Let Ω be a nonempty set. Then the set $\mathcal{F}(\Omega, \mathbb{F})$, the set of all $\mathbb{F}$-valued functions defined on Ω, is a vector space over $\mathbb{F}$ with respect to the pointwise addition and pointwise scalar multiplication.

   Is the set of all scalar sequences a special case of the above?.

4. Which of the following subset of $\mathbb{C}^3$ a subspace of $\mathbb{C}^3$?

   (a) $\{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{C}^3 : \alpha_1 \in \mathbb{R}\}$.

   (b) $\{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{C}^3 : \text{ either } \alpha_1 = 0 \text{ or } \alpha_2 = 0\}$.

   (c) $\{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{C}^3 : \alpha_1 + \alpha_2 = 1 \in \mathbb{R}\}$.

5. Which of the following subset of $\mathcal{P}$ a subspace of $\mathcal{P}$?

   (a) $\{x \in \mathcal{P} : \text{ degree of } x \text{ is } 3\}$.

   (b) $\{x \in \mathcal{P} : 2x(0) = x(1)\}$.

   (c) $\{x \in \mathcal{P} : x(t) \geq 0 \text{ for } t \in [0, 1]\}$.

   (d) $\{x \in \mathcal{P} : x(t) = x(1 - t)\, \forall\, t\}$.

6. Prove the following:

   (a) The spaces $\mathcal{P}_n(\mathbb{F})$ and $\mathbb{F}^{n+1}$ are isomorphic, and find an isomorphism.

   (b) The space $\underline{\mathbb{R}}^n := \mathbb{R}^{n\times 1}$, the space of all column n-vectors is isomorphic with $\mathbb{R}^n$, and find an isomorphism.

(c) The space $\mathbb{R}^{m\times n}$ is isomorphic with $\mathbb{R}^{mn}$, and find an isomorphism.

7. Prove the assertions in the following:

   (a) $S = \{(\alpha_1, \alpha_2) \in \mathbb{R}^2 : \alpha_1 + \alpha_2 = 0\}$ is a subspace of $\mathbb{R}^2$.

   (b) $S = \{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^2 : \alpha_1 + \alpha_2 - \alpha_3 = 0\}$ is a subspace of $\mathbb{R}^3$.

   (c) For each $k \in \{1, \ldots, n\}$, $S_k = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n : \alpha_k = 0\}$ is a subspace of $\mathbb{F}^n$.

   (d) For $n \in \mathbb{N}$ with $n \geq 2$ and each $k \in \{1, \ldots, n-1\}$,
   $S_k = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n : \alpha_i = 0\, \forall i > k\}$ is a subspace of $\mathbb{F}^n$.

   (e) For each $n \in \mathbb{N}$, $\mathcal{P}_n$ is a subspace of $\mathcal{P}$.

   (f) For each $n \in \mathbb{N}$, $V_n := \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : x(j) = 0\, \forall j \geq n\}$ is a subspace of $\mathcal{F}(\mathbb{N}, \mathbb{F})$, and $c_{00} := \bigcup_{n=1}^{\infty} V_n$ is a subspace of $\mathcal{F}(\mathbb{N}, \mathbb{F})$.
   (Note that elements of W are sequences having only a finite number of nonzero entries.)

   (g) For an interval $\Omega := [a, b] \subseteq \mathbb{R}$,

      i. $\mathcal{R}(\Omega)$, the set of all Riemann integrable real valued continuous functions defined on Ω is a subspace of $\mathcal{F}(\Omega, \mathbb{R})$.

      ii. $C(\Omega)$ is a subspace of $\mathcal{R}(\Omega)$

      iii. $C^1(\Omega)$, the set of all real valued continuous functions defined on Ω and having continuous derivative in Ω is a subspace of $C(\Omega)$.

      iv. $S = \{x \in C(\Omega) : \int_a^b x(t)dt = 0\}$ is a subspace of $C(\Omega)$.

      v. $S = \{x \in C(\Omega) : x(a) = 0\}$ is a subspace of $C(\Omega)$.

      vi. $S = \{x \in C(\Omega) : x(a) = 0 = x(b)\}$ is a subspace of $C(\Omega)$.

   (h) Let $A \in \mathbb{R}^{m\times n}$. Then

      i. $\{x \in \underline{\mathbb{R}}^n : Ax = 0\}$ is a subspace of $\underline{\mathbb{R}}^n$,

      ii. $\{Ax : x \in \underline{\mathbb{R}}^n\}$ is a subspace of $\underline{\mathbb{R}}^m$,

   (i) $\{(\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 = 0\}$ is a subspace of $\mathbb{R}^2$.

   (j) $\{(\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 - \alpha_3 = 0\}$ is a subspace of $\mathbb{R}^3$.

   (k) For $i \in \{1, \ldots, n\}$, let $e_i = (\delta_{i1}, \ldots, \delta_{in})$. Let $V = \mathbb{R}^n$. Then $\{(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \alpha_i = 0 \text{ for } i > k\}$ is a subspace of $\mathbb{R}^n$.

   (l) If V_1 and V_2 are subspaces of V, then $V_1 + V_2 = \text{span}(V_1 \cup V_2)$.

   (m) If V_1 and V_2 are subspaces of V and if $V_1 \subseteq V_2$, then $V_1 \cup V_2$ is a subspace of V.

   (n) If V_1 and V_2 are subspaces of V, then $V_1 \cap V_2$ is a subspace of V; but, $V_1 \cup V_2$ need not be a subspace of V.

8. Let V be a vector space and $S \subseteq V$. Prove the following:

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

   (b) If V_0 is a subspace of V such that $S \subseteq V_0$, then $\text{span}(S) \subseteq V_0$.

(c) $S = \text{span}(S)$ if and only if S is a subspace of V.

9. Prove the assertions in the following:

   (a) If $V = \mathbb{R}^2$, then $\text{span}(\{(1,-1)\} = \{(\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 = 0\}$.

   (b) If $V = \mathbb{R}^3$, then $\text{span}(\{(1,-1,0),\ (1,0,1)\} = \{(\alpha_1, \alpha_2) : \alpha_1 + \alpha_2 - \alpha_3 = 0\}$.

   (c) For $i \in \{1, \ldots, n\}$, let $e_i = (\delta_{i1}, \ldots, \delta_{in})$. Let $V = \mathbb{R}^n$. Then

      i. $\text{span}(\{e_1, \ldots, e_k\}) = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \alpha_i = 0 \text{ for } i > k\}$.

      ii. $\text{span}(\{e_1, \ldots, e_n\}) = \mathbb{R}^n$.

   (d) If $V = \mathcal{P}$, then $\text{span}(\{1, t, \ldots, t^n\}) = \mathcal{P}_n$ and $\text{span}(\{1, t, t^2, \ldots\}) = \mathcal{P}$.

   (e) For each $i \in \mathbb{N}$, let $e_i = (\delta_{i1}, \delta_{i2}, \ldots,)$. Then $\text{span}(\{e_1, e_2, \ldots\}) = c_{00}$.

10. Prove that a set of vectors $x_1, \ldots, x_n$ in a vector space V are linearly dependent if and only if there exists $k \in \{2, \ldots, n\}$ such that x_k is a linear combination of $x_1, \ldots, x_{k-1}$.

11. Prove that any three of the polynomials $1, t, t^2, 1 + t + t^2$ are linearly independent

12. Give vectors x_1, x_2, x_2, x_4 in $\mathbb{C}^3$ such that any three of them are linearly independent.

13. Find conditions on α such that the vectors

   (a) $(1 + \alpha, 1 - \alpha)$, $(1 - \alpha,\ 1 + \alpha)$ are linearly dependent $\mathbb{C}^2$,

   (b) $(\alpha, 1, 0)$, $(1, \alpha, 1)$, $(0, 1, \alpha)$ are linearly dependent in $\mathbb{R}^3$.

14. Suppose x, y, z are linearly independent. Is it true that $x + y, y + z, z + x$ are also linearly independent?

15. Prove the assertions in the following:

   (a) $\{e_1, \ldots, e_n\}$ is a basis of $\mathbb{R}^n$ and $\mathbb{C}^n$.

   (b) $\{1, t, \ldots, t^n\}$ is a basis of $\mathcal{P}_n$.

   (c) $\{1, 1 + t, 1 + t + t^2, \ldots, 1 + t + \cdots + t^n\}$ is a basis of $\mathcal{P}_n$.

   (d) $\{1, t, t^2, \ldots\}$ is a basis of $\mathcal{P}$.

   (e) For each $i \in \mathbb{N}$, let $e_i = (\delta_{i1}, \delta_{i2}, \ldots,)$. Then $\{e_1, e_2, \ldots\}$ is a basis of c_{00}.

   (f) If E is linearly independent in a vector space, then E is a basis for $V_0 := \text{span}(E)$.

16. Prove:

   (a) If E is linearly independent and if $x \in V$ with $x \notin \text{span}(E)$, then $E \cup \{x\}$ is linearly independent.

   (b) Every vector space having a finite spanning set has a finite basis.

   (c) If a vector space V has a finite basis, then any two basis of V contains the same number of vectors.

17. Find bases E_1, E_2 for $\mathbb{C}^4$ such that $E_1 \cap E_2 = \varnothing$ and $\{(1,0,0,0), (1,1,0,0)\} \subseteq E_1$ and $\{(1,1,1,0), (1,1,1,1)\} \subseteq E_2$.

18. Prove the assertions in the following:

    (a) $\mathbb{F}^n$ and $\mathcal{P}_n$ are finite dimensional spaces, and $\dim(\mathbb{F}^n) = n$, $\dim(\mathcal{P}_n) = n + 1$.

    (b) $\dim(\{\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \alpha_1 + \cdots + \alpha_n = 0\} = n - 1$.

    (c) $\mathcal{P}$, $C[a, b]$, c_{00} are infinite dimensional spaces.

    (d) Every vector space containing an infinite linearly independent set is infinite dimensional.

    (e) If $A \in \mathbb{R}^{m\times n}$ with $n > m$, then there exists $\underline{x} \in \underline{\mathbb{R}}^n$ such that $A\underline{x} = 0$.

19. Prove:

    (a) If V_1 and V_2 are subspaces of a vector space V such that $V_1 \cap V_2 = \{0\}$, and if E_1 and E_2 are bases of V_1 and V_2, respectively, then $E_1 \cup E_2$ is a basis of $V_1 + V_2$; and in particular,

    $$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2).$$

    (b) If V_1 and V_2 are subspaces of a vector space V, then

    $$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2).$$

    (c) Let V_1 and V_2 be vector spaces and let T be an isomorphism from V_1 onto V_2. Let $E \subseteq V_1$. Then E is a basis of V_1 if and only if $\{T(u) : u \in E\}$ is a basis of V_2.

20. Suppose V_1 and V_2 are subspaces of a vector space V. Prove:

    (a) If V_1 and V_2 are finite dimensional such that $\dim(V_1) = \dim(V_2)$ and $V_1 \subseteq V_2$, then $V_1 = V_2$.

    (b) If $V = V_1 \cup V_2$, then either $V_1 = V$ or $V_2 = V$.

21. Prove that, if V_0 is a subspace of a vector space V, then there exists a subspace V_1 of V such that

    $$V = V_0 + V_1 \quad \text{and} \quad V_0 \cap V_1 = \{0\}.$$

22. If V_1 is the set of all *odd* polynomials (i.e., $x(-t) = -x(t)$ for all t), and if V_2 is the set of all *even* polynomials (i.e., $x(-t) = x(t)$ for all t), prove that V_1 and V_2 are subspaces of $\mathcal{P}$ such that $V = V_1 + V_2$ and $V_1 \cap V_2 = \{0\}$.

23. Let V_1 and V_2 be vector spaces over the same field $\mathbb{F}$. For $x := (x_1, x_2)$, $y := (y_1, y_2)$ in $V_1 \times V_2$, and $\alpha \in \mathbb{F}$, define

    $$x + y = (x_1 + y_1, x_2 + y_2), \qquad \alpha x = (\alpha x_1, \alpha x_2).$$

    Prove:

    (a) $V_1 \times V_2$ is a vector space over $\mathbb{F}$ with respect to the above operations with its zero as $(0, 0)$ and $-x := (-x_1, -x_2)$.

(b) If V_1 and V_2 are finite dimensional, then

$$\dim(V_1 \times V_2) = \dim(V_1) + \dim(V_2).$$

(c) If $\widetilde{V}_1 := \{(x_1, x_2) \in V_1 \times V_2 : x_2 = 0\}$ and $\widetilde{V}_2 := \{(x_1, x_2) \in V_1 \times V_2 : x_1 = 0\}$, then $\widetilde{V}_1$ and $\widetilde{V}_2$ are subspaces of $V_1 \times V_2$ and

$$V_1 \times V_2 = \widetilde{V}_1 + \widetilde{V}_2, \qquad \widetilde{V}_1 \cap \widetilde{V}_2 = \{(0, 0)\}.$$

In view of the above, the space $V_1 \times V_2$ is called the *direct sum* of V_1 and V_2.

24. Let V_1 and V_2 be subspaces of a finite dimensional vector space V such that $V = V_1 + V_2$ and $V_1 \cap V_2 = \{0\}$. Prove that V is isomorphic with $V_1 \times V_2$.

25. Let V_0 be a subspaces of a finite dimensional vector space V. Prove that V is isomorphic with $(V/V_0) \times V_0$.

5.2 (On Section 2: Linear Transformations)

In the following, V_1 and V_2 are vector spaces over a filed $\mathbb{F}$.

For $i, j \in \mathbb{N}$, we denote $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$

1. Let $T : V_1 \to V_2$ be a linear transformation. Prove that

   (a) $T(0) = 0$.

   (b) T is one-one iff $N(T) = \{0\}$.

2. Verify the assertion in each of the following:

   (a) Let $A \in \mathbb{R}^{m \times n}$ and let $T : \underline{\mathbb{R}}^n \to \underline{\mathbb{R}}^m$ be defined by

   $$T\underline{x} = A\underline{x}\,, \quad \underline{x} \in \underline{\mathbb{R}}^n.$$

   Then T is a linear transformation.

   (b) For $x \in C[a, b]$, define

   $$T(x) = \int_a^b x(t)dt.$$

   Then $T : C[a, b] \to \mathbb{R}$ is a linear transformation.

   (c) For $x \in C^1[a, b]$, define

   $$(Tx)(t) = x'(t), \quad t \in [a, b].$$

   Then $T : C^1[a, b] \to C[a, b]$ is a linear transformation.

(d) For $\tau \in [a,b]$ and $x \in C^1[a,b]$, define

$$T(x) = x'(\tau).$$

Then $T : C^1[a,b] \to \mathbb{R}$ is a linear transformation.

(e) Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$ and V be any of the spaces $c_{00}, \ell^1, \ell^\infty$. Recall that

$$c_{00} = \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : \exists\, k \in \mathbb{N} \text{ with } x(j) = 0\, \forall\, j \geq k\},$$

$$\ell^1 = \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : \sum_{j=1}^{\infty} |x(j)| \text{ converges}\},$$

$$\ell^\infty = \{x \in \mathcal{F}(\mathbb{N}, \mathbb{F}) : (x(n)) \text{ bounded}\}.$$

i. $T : V \to V$ defined by

$$T(\alpha_1, \alpha_2, \ldots,) = (0, \alpha_1, \alpha_2, \ldots,)$$

is a linear transformation, called the **right shift operator**.

ii. $T : V \to V$ defined by

$$T(\alpha_1, \alpha_2, \ldots,) = (\alpha_2, \alpha_3, \ldots,)$$

is a linear transformation, called the **left shift operator**.

3. Let $T : V_1 \to V_2$ be a linear transformation. Prove:

(a) If $u_1, \ldots, u_n$ are in V_1 such that $Tu_1, \ldots, Tu_n$ are linearly independent in V_2, then $u_1, \ldots, u_n$ are linearly independent in V_1.

(b) If T is one-one and $u_1, \ldots, u_n$ are linearly independent in V_1, then $Tu_1, \ldots, Tu_n$ are linearly independent in V_2.

Let $T : V_1 \to V_2$ be a linear transformation. Prove:

(a) If E_1 is a basis of V_1, then $R(T) = \text{span}(T(E_1))$.

(b) $\dim R(T) \leq \dim(V_1)$.

(c) If T is one-one, then $\dim R(T) = \dim(V_1)$.

(d) If V_1 and V_2 are finite dimensional such that $\dim(V_1) = \dim(V_2)$, then T is one-one if and only if T is onto.

4. Find the following subspaces of the space $\mathcal{P}_n$:

(a) $V_1 = \{p(t) \in \mathcal{P}_n :\, p(1) = 0\}$,

(b) $V_1 = \{p(t) \in \mathcal{P}_n :\, p(0) = 0,\, p(1)\}$,

(c) $V_1 = \{p(t) \in \mathcal{P}_n :\, \int_0^1 p(t)dt = 0\}$.

5. Let $A \in \mathbb{R}^{m\times n}$ and let $T : \underline{\mathbb{R}}^n \to \underline{\mathbb{R}}^m$ be the linear transformation defined by

$$T\underline{x} = A\underline{x}\,, \quad \underline{x} \in \underline{\mathbb{R}}^n.$$

Prove:

(a) T is one-one if and only if the columns of A are linearly independent.

(b) $R(T)$ is the space spanned by the columns of A, and $\text{rank}(T)$ is the dimension of the space spanned by the columns of A.

6. Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ and let $\{u_1, \ldots, u_n\}$ be a basis of V_1. Let $\{v_1, \ldots, v_n\} \subseteq V_2$. Define $T : V_1 \to V_2$ be

$$T\Big(\sum_{i=1}^{n} \alpha_i u_i\Big) = \sum_{i=1}^{n} \alpha_i v, \quad (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n.$$

(a) Show that T is a linear transformation such that $T(u_j) = v_j$ for $j \in \{1, \ldots, n\}$.

(b) T is one-one if and only if $\{v_1, \ldots, v_n\}$ is linearly independent.

(c) T is onto if and only if $\text{span}(\{v_1, \ldots, v_n\}) = V_2$.

7. Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ and let $E := \{u_1, \ldots, u_n\}$ be a linearly independent subset of V_1. Let $\{v_1, \ldots, v_n\} \subseteq V_2$. Show that there exists a linear transformation $T : V_1 \to V_2$ such that $T(u_j) = v_j$ for $j \in \{1, \ldots, n\}$.

Let V be a finite dimensional space and $E = \{u_1, \ldots, u_n\}$ be an order basis of V. For each $j \in \{1, \ldots, n\}$, let $f_j : V \to \mathbb{F}$ be defined by

$$f_j(x) = \alpha_j \quad \text{for} \quad x := \sum_{i=1}^{n} \alpha_i u_i.$$

Prove:

(a) $f_1, \ldots, f_n$ are in V' and they satisfy $f_i(u_j) = \delta_{ij}$ for $i, j \in \{1, \ldots, n\}$,

(b) $\{f_1, \ldots, f_n\}$ is a basis of V'.

8. Prove: Let V be a finite dimensional space. Then V and V' are linearly isomorphic.

9. Let $E = \{u_1, \ldots, u_n\}$ be an order basis of V. If $f_1, \ldots, f_n$ are in V' such that $f_i(u_j) = \delta_{ij}$. Prove $\{f_1, \ldots, f_n\}$ is the dual basis of V.

10. Let $T_1 \in \mathcal{L}(V_1, V_2)$ and $T_2 \in \mathcal{L}(V_2, V_3)$. Show that

(a) T_2T_1 one-one implies T_1 one-one.

(b) T_2T_1 onto implies T_2 onto.

11. Prove: Let V be a vector space and W be a subspace of V. Then the map $\eta : V \to V/W$ defined by

$$\eta(x) = x + W, \quad x \in V,$$

is a linear transformation.

12. Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ and let $E_1 := \{u_1, \ldots, u_n\}$ and $E_2 := \{v_1, \ldots, v_m\}$ be ordered bases of V_1 and V_2, respectively. Let $T : V_1 \to V_2$ be a linear transformation. Prove that for each j, $[Tu_j]_{E_2}$ is the j^{th} column of $[T]_{E_1E_2}$.

13. Let $A \in \mathbb{R}^{m\times n}$ and let $T : \underline{\mathbb{R}}^n \to \underline{\mathbb{R}}^m$ be defined by $T\underline{x} = A\underline{x}\,, \quad \underline{x} \in \mathbb{R}^n$. If E_1 and E_2 are the standard bases for $\underline{\mathbb{R}}^n$ and $\underline{\mathbb{R}}^m$, respectively, then prove that $[T]_{E_1E_2} = A$.

14. Prove: Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ with $\dim(V_1) = n$ and $\dim(V_2) = m$ and let E_1 and E_2 be ordered bases of V_1 and V_2, respectively. Let $T : V_1 \to V_2$ be a linear transformation. Then the following hold:

    (a) $[Tx]_{E_2} = [T]_{E_1E_2}[x]_{E_1}$ for all $x \in V_1$.

    (b) T is one-one (respectively, onto) if and only if $[T]_{E_1E_2} : \underline{\mathbb{R}}^n \to \underline{\mathbb{R}}^m$ is one-one (respectively, onto).

    (c) For $A \in \mathbb{R}^{m\times n}$,
    $$A = [T]_{E_1E_2} \iff [Tx]_{E_2} = A[x]_{E_1} \quad \forall\, x \in V_1.$$

    (d) $T = J_2^{-1}[T]_{E_1E_2}J_1$, where $J_1 : V_1 \to \underline{\mathbb{R}}^n$ and $J_2 : V_2 \to \underline{\mathbb{R}}^m$ are the canonical isomorphisms,

15. Let V_1, V_2, V_3 be finite dimensional vector spaces over the same field $\mathbb{F}$, and let E_1, E_2, E_3 be ordered bases of V_1, V_2, V_3, respectively. If $T_1 \in \mathcal{L}(V_1, V_2)$ and $T_2 \in \mathcal{L}(V_2, V_3)$. Then the
$$[T_2T_1]_{E_1E_3} = [T_2]_{E_2E_3}[T_1]_{E_1E_2}.$$

16. For $n \in \mathbb{N}$, let $D : \mathcal{P}_n \to \mathcal{P}_{n-1}$ and $T : \mathcal{P}_n \to \mathcal{P}_{n+1}$ be defined by
$$D(a_0 + a_1t + \cdots + a_nt^n) = a_1t + 2a_2t + \cdots + na_nt^{n-1},$$
$$T(a_0 + a_1t + \cdot + a_nt^n) = a_0t + \frac{a_1}{2}t^2 + \cdots + \frac{a_n}{n+1}t^{n+1}.$$
Let $E_k = \{1, t, \ldots, t^k\}$ for $k \in \mathbb{N}$. Find
$$[D]_{E_nE_{n-1}}, \quad [T]_{E_nE_{n+1}}, \quad [TD]_{E_nE_n}, \quad [DT]_{E_nE_n}.$$

17. Let V_1 and V_2 be finite dimensional vector spaces over the same field $\mathbb{F}$ and let $T : V_1 \to V_2$ be a linear transformation. Let $E_1 = \{u_1, \ldots, u_n\}$ and $\tilde{E}_1 = \{\tilde{u}_1, \ldots, \tilde{u}_n\}$ be two bases of V_1 and $E_2 = \{v_1, \ldots, v_m\}$ and $\tilde{E}_2 = \{\tilde{v}_1, \ldots, \tilde{v}_m\}$ be two bases of V_2. Let $\Phi_1 : V_1 \to V_1$ and $\Phi_2 : V_2 \to V_2$ be the linear transformations such that
$$\Phi_1(u_i) = \tilde{u}_i, \quad \Phi_2(v_j) = \tilde{v}_j$$
for $i = 1, \ldots, n$; $j = 1, \ldots, m$. Prove that
$$[T]_{\tilde{E}_1\tilde{E}_2} = [\Phi_2]_{E_2E_2}^{-1}[T]_{E_1E_2}[\Phi_1]_{E_1E_1}.$$

18. Let $\mathcal{P}_2$ be the vector space (over $\mathbb{R}$) of all polynomials of degree at most 2 with real coefficients. Let $T : \mathcal{P}_2 \to \mathbb{R}^{2\times 2}$ be the linear transformation defined by
$$T(p(t)) = \begin{bmatrix} p(1) & p(0) \\ p(0) - p(1) & p(0) \end{bmatrix}.$$

    (a) Find a basis for $N(T)$ and a basis for $R(T)$.

    (b) If W is the space of all symmetric matrices in $\mathbb{R}^{2\times 2}$, then find a basis for $W \cap R(T)$.

5.3 (On Section 3: Inner product spaces)

In the following, V is an inner product over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

For $i, j \in \mathbb{N}$, we denote $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$

1. Verify:

   (a) On the vector space c_{00}, $\langle x, y\rangle := \sum_{j=1}^{\infty} x(j)\overline{y(j)}$ defines an inner product.

   (b) On the vector space $C[a, b]$, $\langle, y\rangle := \int_a^b x(t)\overline{y(t)}dt$ defines an inner product.

   (c) Let $\tau_1, \ldots, \tau_{n+1}$ be distinct real numbers. On the vector space $\mathcal{P}_n$, $\langle p, q\rangle := \sum_{i=1}^{n+1} p(\tau_i)\overline{q(\tau_i)}$ defines an inner product.

2. Prove the following:

   (a) For $x \in V$, $\langle x, u\rangle = 0 \,\forall\, u \in V \Longrightarrow x = 0$.

   (b) For $u \in V$, if $f : V \to \mathbb{F}$ is defined by $f(x) = \langle x, u\rangle$ for all $x \in V$, then $f \in V'$.

   (c) Let $u_1, u_2, \ldots, u_n$ be linearly independent vectors in V and let $x \in V$. Then

   $$\langle x, u_i\rangle = 0 \quad \forall\, i \in \{1, \ldots, n\} \iff \langle x, y\rangle = 0 \quad \forall\, y \in \text{span}\{u_1, \ldots, u_n\}.$$

   In particular, if $\{u_1, u_2, \ldots, u_n\}$ is a basis of V, and if $\langle x, u_i\rangle = 0$ for all $i \in \{1, \ldots, n\}$, then $x = 0$.

   (d) For $S \subseteq V$, $[\text{span}(S)]^{\perp} = S^{\perp}$.

3. Let V_1 and V_2 be subspaces of an inner product space V. Prove that $(V_1 + V_2)^{\perp} = V_1^{\perp} \cap V_2^{\perp}$.

4. Recall that $d : V \times V \to \mathbb{R}$ defined by $d(x, y) = \|x - y\|$ is a metric on V, called the metric induced by the inner product. Then, with respect to the above metric, prove the following:

   (a) The map $x \mapsto \|x\|$ is continuous on V.

   (b) For each $u \in V$, the linear functional $f : V \to \mathbb{F}$ defined by $f(x) = \langle x, u\rangle$, $x \in V$, is continuous.

   (c) For every $S \subseteq V$, the set $S^{\perp}$ is closed in V.

5. Consider the standard inner product on $\mathbb{F}^n$. For each $j \in \{1, \ldots, n\}$, let $e_j = (\delta_{1j}, \delta_{2j}, \ldots, \delta_{nj})$. Show that $(e_i + e_j) \perp (e_i - e_j)$ for every $i, j \in \{1, \ldots, n\}$.

6. Using Gram-Schmidt orthogonalization process, orthonormalise the sets S in the following:

   (a) $S = \{1, t, t^2, t^3\}$ with respect to the usual inner product on $\mathcal{P}_3$.

   (b) $S = \{(1, 1, 1, 1), (1, 1, 1, 0), (1, 1, 0, 0)\}$ with respect to the usual inner product on $\mathbb{R}^4$.

7. Consider the vector space $C[0, 2\pi]$ with inner product defined by $\langle f, g\rangle := \int_0^{2\pi} f(t)\overline{g(t)}\, dt$ for $f, g \in C[0, 2\pi]$. For $n \in \mathbb{N}$, let

$$u_n(t) := \sin(nt), \quad v_n(t) = \cos(nt), \qquad 0 \le t \le 2\pi.$$

Let $w_{2n-2} = v_n$ and $w_{2n-1} = u_n$ for $n \in \mathbb{N}$.Show that the sets

$$\{u_n : n \in \mathbb{N}\}, \quad \{v_n : n \in \mathbb{N}\}, \quad \{w_n : n \in \mathbb{N}\}$$

are orthogonal sets.

8. Suppose $\{u_1, \ldots, u_n\}$ is an orthonormal set in an inner product space V and $x \in V$. Then

$$x - \sum_{i=1}^{n} \langle x, u_i\rangle u_i \perp \operatorname{span}\{u_1, \ldots, u_n\}$$

and

$$\sum_{i=1}^{n} |\langle x, u_i\rangle|^2 \le \|x\|^2.$$

Further, the following are equivalent:

(a) $x \in \operatorname{span}\{u_1, \ldots, u_n\}$

(b) $x = \sum_{i=1}^{n} \langle x, u_i\rangle u_i$

(c) $\|x\|^2 = \sum_{i=1}^{n} |\langle x, u_i\rangle|^2$.

9. Let $V = \mathbb{F}^3$ with standard inner product. Form the given vectors $x, y, z \in \mathbb{F}^3$ in the following Construct orthonormal vectors u, v, w in $\mathbb{F}^3$ such that $\operatorname{span}\{u, v\} = \operatorname{span}\{x, y\}$ and $\operatorname{span}\{u, v, w\} = \operatorname{span}\{x, y, z\}$.

(a) $x = (1, 0, 0)$, $y = (1, 1, 0)$, $z = (1, 1, 1)$;

(b) $x = (1, 1, 0)$, $y = (0, 1, 1)$, $z = (1, 0, 1)$.

10. For $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ and $(\beta_1, \ldots, \beta_n) \in \mathbb{F}^n$, show that

$$\sum_{j=1}^{n} |\alpha_j \beta_j| \le \Big(\sum_{j=1}^{n} |\alpha_j|^2\Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{n} |\beta_j|^2\Big)^{\frac{1}{2}}.$$

11. For $x, y \in \mathcal{F}(\mathbb{N})$ prove that

$$\sum_{j=1}^{\infty} |\alpha_j \beta_j| \le \Big(\sum_{j=1}^{\infty} |\alpha_j|^2\Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{\infty} |\beta_j|^2\Big)^{\frac{1}{2}}.$$

Hint: Use Exercise 10.

Let

$$\ell^2 = \{x \in \mathcal{F}(\mathbb{N}) : \sum_{j=1}^{\infty} |x(j)|^2 < \infty\}.$$

Prove that

(a) ℓ^2 is a subspace $\mathcal{F}(\mathbb{N})$.

(b) For $x, y \in \ell^2$, $\sum_{j=1}^{\infty} |x(j)\overline{y(j)}|$ converges.

(c) $\langle x, y\rangle := \sum_{j=1}^{\infty} |x(j)\overline{y(j)}|$ defines an inner product on ℓ^2.

12. For $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ and $(\beta_1, \ldots, \beta_n) \in \mathbb{F}^n$, show that

$$\Big(\sum_{j=1}^{n} |\alpha_j + \beta_j|^2\Big)^{\frac{1}{2}} \leq \Big(\sum_{j=1}^{n} |\alpha_j|^2\Big)^{\frac{1}{2}} + \Big(\sum_{j=1}^{n} |\beta_j|^2\Big)^{\frac{1}{2}}.$$

13. For $x, y \in \mathcal{F}(\mathbb{N})$ prove that

$$\Big(\sum_{j=1}^{\infty} |\alpha_j + \beta_j|^2\Big)^{\frac{1}{2}} \leq \Big(\sum_{j=1}^{\infty} |\alpha_j|^2\Big)^{\frac{1}{2}} + \Big(\sum_{j=1}^{\infty} |\beta_j|^2\Big)^{\frac{1}{2}}.$$

Hint: Use Exercise 12.

14. Let $\dim(V) = n$ and let $E = \{u_1, \ldots, u_n\}$ be an ordered orthonormal set which is a basis of V. Let $A : V \to V$ be a linear transformation.

(a) Show that $[A]_{E,E} = (\langle Au_j, u_i\rangle)$. [Hint: Use Fourier expansion.]

(b) Define $B : V \to V$ such that $\langle Ax, y\rangle = \langle x, By\rangle$ for all $x, y \in V$.

15. Let $\dim(V) = n$ and let $E = \{u_1, \ldots, u_n\}$ be an ordered orthonormal set which is a basis of V. Let $A, B : V \to V$ be a linear transformations satisfying $\langle Ax, y\rangle = \langle x, By\rangle$ for all $x, y \in V$. Show that $[B]_{E,E} = \overline{[A]}^T_{E,E}$, conjugate transpose of $[A]_{E,E}$.

16. Let V be finite dimensional and V_0 is a subspace of V. Prove that every $x \in V$ can be written uniquely as $x = y + z$ with $y \in V_0$ and $z \in V_0^{\perp}$. [Hint: Obtain a basis of V_0, extend it to a basis of V, and consider the orthonormalization of that basis.]

17. Let V be an inner product space and V_0 be a finite dimensional subspace of V. Then for every $x \in V$, there exists a unique pair $y \in V_0$ such that

$$\|x - y\| = \inf_{u \in V_0} \|x - u\|.$$

18. Let V be an inner product space and V_0 be a subspace of V and let $x \in V$ and $y \in V_0$. Prove the following:

(a) If $\langle x - y, u\rangle = 0 \quad \forall\, u \in V_0 \quad \Longrightarrow \quad \|x - y\| = \inf_{u \in V_0} \|x - u\|$.

(b) If $\mathrm{span}(S) = V_0$ and $\langle x - y, u\rangle = 0 \quad \forall\, u \in S \quad \Longrightarrow \quad \|x - y\| = \inf_{u \in V_0} \|x - u\|$.

19. Let V be an inner product space, V_0 be a finite dimensional subspace of V and $x \in V$. Let $\{u_1, \ldots, u_k\}$ be a basis of V_0. Prove that for $y = \sum_{j=1}^{k} \alpha_j u_j$,

$$\langle x - y, u \rangle = 0 \quad \forall\, u \in V_0 \iff \sum_{j=1}^{k} \langle u_j, u_i \rangle \alpha_j = \langle x, u_i \rangle, \quad i = 1, \ldots, k.$$

Further, prove that there exists a unique $(\alpha_1, \ldots, \alpha_k) \in \mathbb{F}^k$ such that

$$\sum_{j=1}^{k} \langle u_j, u_i \rangle \alpha_j = \langle x, u_i \rangle, \quad i = 1, \ldots, k,$$

and in that case $\|x - y\| = \inf_{u \in V_0} \|x - u\|$.

20. Let $V = C[0,1]$ with inner product: $\langle f, g \rangle := \int_0^1 f(t)\overline{g(t)}dt$. Let $x(t) = t^5$. Find best approximation for x from the space V_0, where

$$(i)\, V_0 = \mathcal{P}_1, \qquad (ii)\, V_0 = \mathcal{P}_2, \qquad (iii)\, V_0 = \mathcal{P}_3, \qquad (iv)\, V_0 = \mathcal{P}_4, \qquad (v)\, V_0 = \mathcal{P}_5.$$

21. Let $V = C[0, 2\pi]$ with inner product: $\langle f, g \rangle := \int_0^{2\pi} f(t)\overline{g(t)}dt$. Let $x(t) = t^2$. Find best approximation for x from the space V_0, where

$$V_0 = \text{span}\{1, \sin t, \cos t, \sin 2t, \cos 2t\}.$$

22. Let V be finite dimensional and V_0 is a subspace of V. For $x \in V$, let y, z be as in the last problem. Define $P, Q : V \to V$ by $P(x) = y$ and $Q(x) = z$. Prove that P and Q are liner transformations satisfying the following:

$$R(P) = V_0, \quad R(Q) = V_0^{\perp}, \quad P^2 = P, \quad Q^2 = Q, \quad P + Q = I,$$

$$\langle Pu, v \rangle = \langle u, Pv \rangle \quad \forall\, u, v \in V, \quad \|x - Px\| \le \|x - u\| \quad \forall\, u \in V_0.$$

23. Prove the following:

    (a) If A is self adjoint, the $\langle Ax, x \rangle \in \mathbb{R}$ for every $x \in X$.

    (b) If A is normal, then $\|Ax\| = \|A^*x\|$ for every $x \in X$.

    (c) If A is unitary, then $\langle Ax, Ay \rangle = \langle x, y \rangle$ for every $x, y \in X$. In particular, images of orthogonal vectors are orthogonal.

5.4 (On Section 4: Eigenvalues and eigenvectors)

In the following V is a vector space over $\mathbb{F}$ which is either $\mathbb{R}$ or $\mathbb{C}$, and $T : V \to T$ is a linear operator.

1. Let $A \in \underline{\mathbb{R}}^{n \times n}$, and consider it as a linear operator from $\underline{\mathbb{R}}^n$ to itself. Prove that $\lambda \in \sigma_{\text{eig}}(A) \iff \det(A - \lambda I) = 0$.

2. Show that $\sigma_{\text{eig}}(T) = \varnothing$ in the following cases:

   (a) Let $V = \mathcal{P}$, the space of all polynomials over $\mathbb{F}$ and let $Tp(t) = tp(t)$, $p(t) \in \mathcal{P}$.

   (b) Let $V = c_{00}$ and T be the right shift operator on V.

3. Find the eigenvalues and some corresponding eigenvectors for the following cases:

   (a) $V = \mathcal{P}$ and $Tf = f''$.

   (b) $V = C(\mathbb{R}))$ and $Tf = f''$.

4. Let $V = \mathcal{P}_2$. Using a matrix representation of T, find eigenvalues of $T_1 f = f'$ and $T_2 f = f''$.

5. Find eigenspectrum of T if $T^2 = T$.

6. Prove that eigenvectors corresponding to distinct eigenvalues of T are linearly independent.

7. Prove that, for every polynomial $p(t)$ and $\lambda \in \mathbb{F}$ and $x \in V$, $Tx = \lambda x \Longrightarrow p(T)x = p(\lambda)x$.

8. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$T(\alpha, \beta, \gamma) = (\alpha,\, 2\alpha + 3\beta,\, 3\alpha + 4\gamma), \quad (\alpha, \beta, \gamma) \in \mathbb{R}^3.$$

   Find a basis for $\mathbb{R}^3$ consisting of eigenvectors of T.

9. Suppose V is an inner product space and T is a normal operator, i.e., $T^*T = TT^*$. Prove that vector x is an eigenvector of T corresponding to an eigenvalue λ if and only if x is an eigenvector of T corresponding to the eigenvalue $\bar{\lambda}$.

10. Prove that every symmetric matrix with real entries has a (real) eigenvalue.

11. Let $A = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$. Find an orthogonal matrix U such that $U^T AU$ is a diagonal matrix.

12. Prove that, if V is a finite dimensional inner product space and T is a self adjoint operator, then $\sigma_{\text{eig}}(T) \neq \varnothing$.

13. Let V be a finite dimensional vector space.

   (a) Prove that T is diagonalizable if and only if there are distinct $\lambda_1, \ldots, \lambda_k$ in $\mathbb{F}$ such that $V = N(T - \lambda_1 I) + \cdots + N(T - \lambda_k I)$.

   (b) Prove that, if T has an eigenvalue λ such that $N(T-\lambda I)$ is a proper subspace of $N(T-\lambda I)^2$, then T is not diagonalizable. Is the converse true?

   (c) Give an example of a non-diagonalizable operator on a finite dimensional vector space.

14. Let V be a finite dimensional vector space and T be diagonalizable. If $p(t)$ is a polynomial which vanishes at the eigenvalues of T, then prove that $p(T) = 0$.

15. Let V be a finite dimensional vector space.

(a) Let $\lambda \neq \mu$. Prove that $N(T-\lambda I)^i \cap N(T-\mu I)^j = \{0\}$ for every $i, j \in \mathbb{N}$.

(b) Prove that generalized eigenvectors associated with distinct eigenvalues are linearly independent.

(c) Prove Cayley-Hamilton theorem for operators.

16. Let V be finite dimensional over $\mathbb{C}$ and λ be an eigenvalue of T with ascent ℓ. Prove that $m := \dim[N(T-\lambda I)^\ell]$ is the algebraic multiplicity of λ.

17. Let V finite dimensional, $k \in \mathbb{N}$ be such that $\{0\} \neq N(T^k) \neq N(T^{k+1})$, and let Y_k be a subspace of $N(T^{k+1})$ such that $N(T^{k+1}) = N(T^k) \oplus Y_k$. Prove that $\dim(Y_k) \leq \dim[N(T^k)]$.

18. Let V be a finite dimensional vector space and T be diagonalizable. Let $u_1, \ldots, u_n$ be eigenvectors of T which for a basis of T, and let $\lambda_1, \ldots, \lambda_n$ be such that $Tu_j = \lambda_j u_j$, $j = 1, \ldots, n$. Let f be an $\mathbb{F}$-valued function defined on an opens set $\Omega \subseteq \mathbb{F}$ such that $\Omega \supset \sigma_{\rm eig}(T)$. For $x = \sum_{j=1}^n \alpha_j u_j \in V$, define
$$f(T)x = \sum_{j=1}^n \alpha_j f(\lambda_j) u_j.$$
Prove that there is a polynomial $p(t)$ such that $f(T) = p(T)$ [Hint: Lagrange interpolation].