Advanced Linear Algebra: Math 363

Dr. Abdullah Al-Azemi

Mathematics Department
Kuwait University

December 21, 2019

Contents

Chapter

0 Review

Section 0.1: Fields

Definition 0.1.1

A **field** $\mathbb{F}$ is a set on which two operations $+$ and $\cdot$ (called addition and multiplication, respectively) are defined, such that for each pair of elements $x, y \in \mathbb{F}$, there are unique elements $x+y$ and $x \cdot y$ in $\mathbb{F}$ for which the following properties hold for all elements $a, b, c \in \mathbb{F}$.

F1. $a + b = b + a$ and $a \cdot b = b \cdot a$ (Commutativity).

F2. $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associativity).

F3. There are unique elements 0 and 1 in $\mathbb{F}$ such that

$$\text{(identities):} \qquad 0 + a = a \text{ and } 1 \cdot a = a.$$

F4. For each element $a \in \mathbb{F}$ and each nonzero element $b \in \mathbb{F}$, there exist unique elements c and d in $\mathbb{F}$ such that

$$\text{(inverses):} \qquad a + c = 0 \text{ and } b \cdot d = 1.$$

F5. $a \cdot (b + c) = a \cdot b + a \cdot c$ (distributivity).

Example 0.1.1

The following sets are fields with the usual definitions of addition and multiplication:

1. **real numbers** $\mathbb{R}$, and **rational numbers** $\mathbb{Q}$.
2. $\{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\} \subseteq \mathbb{R}$.

Example 0.1.2

The field $\mathbb{Z}_2 = \{0, 1\}$ with the operations of addition and multiplication defined by

$$0 + 0 = 0, \qquad 0 + 1 = 1 + 0 = 1, \qquad 1 + 1 = 0,$$
$$0 \cdot 0 = 0, \qquad 0 \cdot 1 = 1 \cdot 0 = 0, \qquad \text{and } 1 \cdot 1 = 1.$$

Remark 0.1.1

The sets $\mathbb{Z}^+$, $\mathbb{Z}^-$, and $\mathbb{Z}$ are not fields since the property **F4** does not hold for all of the three sets.

Theorem 0.1.1

For any elements a, b, and c in a field $\mathbb{F}$, the following statements hold:

1. The **Cancellation Laws** $\begin{cases} \text{If } a + c = b + c, \text{ then } a = b, \\ \text{If } ac = bc \text{ and } c \neq 0, \text{ then } a = b. \end{cases}$

2. $a \cdot 0 = 0$.

3. $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$.

4. $(-a) \cdot (-b) = a \cdot b$.

Definition 0.1.2

In a field $\mathbb{F}$, the smallest positive integer p such that the sum of p $1's$ is 0 is called the **characteristic** of $\mathbb{F}$. If no such positive integer exists, then $\mathbb{F}$ is said to have characteristic zero.

Note that $\mathbb{Z}_2$ has characteristic 2, while $\mathbb{R}$ has characteristic zero.

Section 0.2: Some Facts About Complex Numbers $\mathbb{C}$

Definition 0.2.1

A **compleax number** is an expression of the form $z = a+bi$, where a and b are real numbers called the **real part** and the **imaginary part** of z, respectively. Note that $i = \sqrt{-1}$ and hence $i^2 = -1$.

The sum and product of two complex numbers $z = a + bi$ and $w = c + di$ are defined by

$$z + w = (a + c) + (b + d)i, \text{ and } zw = (a + bi)(c + di) = (ac - db) + (ad + bc)i.$$

Definition 0.2.2

The **complex conjugate** of a complex number $z = a + bi$ is the complex number $\overline{z} = a - bi$. Moreover, the **absolute value** (or **modulus**) of z is the real number $\sqrt{a^2 + b^2}$.

Let $z = a + ib, w = c + di \in \mathbb{C}$ for some $a, b, c, d \in \mathbb{R}$, then the following statements are true:

Facts

1. $\overline{\overline{z}} = z$.
2. $\overline{z + w} = \overline{z} + \overline{w}$.
3. $\overline{zw} = \overline{z} \cdot \overline{w}$.
4. $\overline{\left(\frac{z}{w}\right)} = \frac{\overline{z}}{\overline{w}}$, if $w \neq 0$.
5. $z\,\overline{z} = |\,z\,|^2$.
6. $|zw| = |z| \cdot |w|$.
7. $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$, if $w \neq 0$.
8. $|z| - |w| \leq |z + w| \leq |z| + |w|$.
9. $z + \overline{z} = 2Re(z) = 2a$.
10. $z - \overline{z} = 2Im(z) = 2b$.

Chapter

1 Vector Spaces

Section 1.2: Vector Spaces

An object of the form $(x_1, x_2, \cdots, x_n)$, where $x_1, \cdots, x_n$ are elements of a field $\mathbb{F}$, is called an **n-tuple**. Such object is called a **vector**. Moreover, the set of all vectors with entries from $\mathbb{F}$ is denoted by $\mathbb{F}^n$. The elements $x_1. \cdots, x_n$ are called the **entries** or **components**.

Definition 1.2.1

A **vector space** (or **linear space**) $\mathbb{V}$ over a field $\mathbb{F}$ is a set of elements on which two operations (called addition and scalar multiplication) are defined so that

(α) If $x, y \in \mathbb{V}$, then $x + y \in \mathbb{V}$; that is, "$\mathbb{V}$ is closed under $+$".

VS1. $x + y = y + x$ for all $x, y \in \mathbb{V}$.

VS2. $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathbb{V}$.

VS3. There exists an element $\mathbf{0}$ in $\mathbb{V}$ such that $x + \mathbf{0} = x$ for each $x \in \mathbb{V}$.

VS4. For each $x \in \mathbb{V}$, there exists an element $y \in \mathbb{V}$ such that $x + y = \mathbf{0}$.

(β) If $x \in \mathbb{V}$ and $a \in \mathbb{F}$, then $ax \in \mathbb{V}$; that is, "$\mathbb{V}$ is closed under $\cdot$".

VS5. For each $x \in \mathbb{V}$, $1x = x$.

VS6. For each pair of elements $a, b \in \mathbb{F}$ and each element $x \in \mathbb{V}$, $(ab)x = a(bx)$.

VS7. For each $a \in \mathbb{F}$ and $x, y \in \mathbb{V}$, $a(x + y) = ax + ay$.

VS8. For each $a, b \in \mathbb{F}$ and $x \in \mathbb{V}$, $(a + b)x = ax + bx$.

Remark 1.2.1

A vector space $\mathbb{V}$ along with operation $+$ and $\cdot$ is denoted by $(\mathbb{V}, +, \cdot)$.

Theorem 1.2.1

For any positive integer n, $(\mathbb{R}^n, +, \cdot)$ is a vector space.

Example 1.2.1

Let $M_{m\times n}(\mathbb{F}) = \{\text{all } m \times n \text{ matrices over a field } \mathbb{F}\}$. Then $(\, M_{m\times n}(\mathbb{F}), +, \cdot\,)$ is a vector space where for any $A = (a_{ij}), B = (b_{ij}) \in M_{m\times n}(\mathbb{F})$ and for $c \in \mathbb{F}$, we have

$$(A + B)_{ij} = (a_{ij} + b_{ij}) \text{ and } (cA)_{ij} = c\, a_{ij},$$

for all $1 \le i \le m$ and $1 \le j \le n$.

Example 1.2.2

Let S be a nonempty set and $\mathbb{F}$ be any field, and let $\mathcal{F}(S, \mathbb{F})$ denote the set of all functions from S to $\mathbb{F}$. Two functions $f, g \in \mathcal{F}(S, \mathbb{F})$ are called equal if $f(x) = g(x)$ for each $x \in S$. The set $\mathcal{F}(S, \mathbb{F})$ is a vector space with the operations of addition and scalar multiplication defined for $f, g \in \mathcal{F}(S, \mathbb{F})$ and $c \in \mathbb{F}$ by

$$(f + g)(x) = f(x) + g(x) \qquad \text{and} \qquad (cf)(x) = c\, f(x),$$

for each $x \in S$.

Example 1.2.3

Let $S = \{(a, b) \,:\, a, b \in \mathbb{R}\}$. For any $(a, b), (x, y) \in S$ and $c \in \mathbb{R}$, define

$$(a, b) \oplus (x, y) = (a + x, b - y) \qquad \text{and} \qquad c \odot (a, b) = (ca, cb).$$

Is $(\, S, \oplus, \odot\,)$ a vector space?

Solution:

No. Since (**VS1**), (**VS2**), and (**VS8**) are not satisfied (verify!). For instace, $(1, 2) \oplus (1, 3) \ne (1, 3) \oplus (1, 2)$.

Theorem 1.2.2: Cancellation Law for Vector Addition

If x, y, and z are vectors in a vector space $\mathbb{V}$ such that $x + z = y + z$, then $x = y$.

Proof:

There is a vector $v \in \mathbb{V}$ such that $z + v = \mathbf{0}$. Then

$$\begin{aligned} x &= x + \mathbf{0} = x + (z + v) = (x + z) + v \\ &= (y + z) + v = y + (z + v) = y + \mathbf{0} = y. \end{aligned}$$

Theorem 1.2.3

Let $(\mathbb{V}, +, \cdot)$ be a vector space. Then

(a) The zero vector in $\mathbb{V}$ is unique.

(b) The addition inverse for each element in $\mathbb{V}$ is unique.

Proof:

(a): Assume that $\mathbf{0}_1$ and $\mathbf{0}_2$ are two zeros in $\mathbb{V}$, then for any $x \in \mathbb{V}$, we have $x+\mathbf{0}_1 = x = x+\mathbf{0}_2$. Thus, using the cancellation law we have

$$x + \mathbf{0}_1 = x + \mathbf{0}_2 \qquad \Rightarrow \qquad \mathbf{0}_1 = \mathbf{0}_2.$$

(b): For any $x \in \mathbb{V}$, assume that y and z are two additive inverses for x. Then, by cancellation law we have

$$x + y = \mathbf{0} = x + z \qquad \Rightarrow \qquad y = z.$$

Theorem 1.2.4

In any vector space $\mathbb{V}$, the following statements are true.

(a) $0\,x = \mathbf{0}$ for each $x \in \mathbb{V}$.

(b) $(-a)x = -(ax) = a(-x)$ for each $a \in \mathbb{F}$ and each $x \in \mathbb{V}$.

(c) $a\,\mathbf{0} = \mathbf{0}$ for each $a \in \mathbb{F}$.

Proof:

(a): Clearly $0x + \mathbf{0} = 0x = (0 + 0)x = 0x + 0x$, and by cancellation law, $0x = \mathbf{0}$.

(b): The element $-(ax)$ is the unique element in $\mathbb{V}$ such that $ax + [-(ax)] = \mathbf{0}$. But $ax + (-a)x = (a + (-a))x = 0x = \mathbf{0}$ as well. Hence, $-(ax) = (-a)x$. Moreover,

$$a(-x) = a[(-1)x] = (a(-1))x = (-a)x.$$

(c): Note that $\mathbf{0} = \mathbf{0} + \mathbf{0}$. Thus,

$$a\mathbf{0} + \mathbf{0} = a\mathbf{0} = a(\mathbf{0} + \mathbf{0}) = a\mathbf{0} + a\mathbf{0}.$$

By the cancellation law, we get $a\mathbf{0} = \mathbf{0}$.

Exercise 1.2.1

Solve the following exercises from the book at pages 12 - 16:

- 13, 17, 18.

Section 1.3: Subspaces

Definition 1.3.1

A subset $\mathbb{W}$ of a vector space $\mathbb{V}$ over a field $\mathbb{F}$ is called **subspace** of $\mathbb{V}$ if $\mathbb{W}$ is a vector space over $\mathbb{F}$ with operations of addition and scalar multiplication defined on $\mathbb{V}$.

Note that, if $\mathbb{V}$ is any vector space, then $\{\mathbf{0}\}$ and $\mathbb{V}$ are both subspaces of $\mathbb{V}$.

Theorem 1.3.1

Let $\mathbb{V}$ be a vector space over a field $\mathbb{F}$ and $\mathbb{W}$ is a subset of $\mathbb{V}$. Then, $\mathbb{W}$ is a subspace of $\mathbb{V}$ if and only if:

1. $\mathbf{0} \in \mathbb{W}$.
2. For any $x, y \in \mathbb{W}$, $x + y \in \mathbb{W}$.
3. For any $x \in \mathbb{W}$ and any $a \in \mathbb{F}$, $ax \in \mathbb{W}$.

Example 1.3.1

Show that the set $\mathbb{W}$ of all symmetric matrices (that is matrices with property $A^t = A$) is a subspace of $M_{n\times n}(\mathbb{F})$.

Solution:

We need to show the three conditions of Theorem 1.3.1.

1. Clearly, $\mathbf{0}^t_{n\times n} = \mathbf{0}_{n\times n}$ and hence $\mathbf{0}_{n\times n} \in \mathbb{W}$.
2. Let $A, B \in \mathbb{W}$. Then $A^t = A$ and $B^t = B$ and hence $(A + B)^t = A^t + B^t = A + B$. Thus, $A + B \in \mathbb{W}$.
3. Let $A \in \mathbb{W}$ and $a \in \mathbb{F}$. Then $A^t = A$ and hence $(aA)^t = aA^t = aA$. Thus, $aA \in \mathbb{W}$.

Therefore, $\mathbb{W}$ is a subspace of $M_{n\times n}(\mathbb{F})$.

Note that the set $\mathbb{W}$ of all non-singular matrices in $M_{n\times n}(\mathbb{F})$ is not a subspace of $M_{n\times n}(\mathbb{F})$. Can you guess why!?

Definition 1.3.2

The **trace** of an $n \times n$ matrix A, denoted $tr(A)$, is the sum of the diagonal entries of A. That is, for $A = (a_{ij})$,

$$tr(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}.$$

Example 1.3.2: Exercise #6 @ page 20

Show that $tr(cA + dB) = c\,tr(A) + d\,tr(B)$ for any $n \times n$ matrices A and B.

Solution:

If $A = (a_{ij})$ and $B = (b_{ij})$, then $c\,A = (c\,a_{ij})$ and $d\,B = (d\,b_{ij})$ for $1 \le i, j \le n$. Thus

$$\begin{aligned} tr(c\,A + d\,B) &= (c\,a_{11} + d\,b_{11}) + (c\,a_{22} + d\,b_{22}) + \cdots + (c\,a_{nn} + d\,b_{nn}) \\ &= c\,(a_{11} + a_{22} + \cdots + a_{nn}) + d\,(b_{11} + b_{22} + \cdots + b_{nn}) \\ &= c\,tr(A) + d\,tr(B). \end{aligned}$$

Example 1.3.3

Show that the set $\mathbb{W} = \{\, A \in M_{n\times n}(\mathbb{F}) \,:\, tr(A) = 0 \,\}$ is a subspace of $M_{n\times n}(\mathbb{F})$.

Solution:

We need to show the three conditions of Theorem 1.3.1.

1. $tr(\mathbf{0}_{n\times n}) = \sum_{i=1}^{n} 0 = 0$ and hence $\mathbf{0}_{n\times n} \in \mathbb{W}$.
2. Let $A, B \in \mathbb{W}$. Then $tr(A) = tr(B) = 0$ and hence

   $$tr(A + B) = tr(A) + tr(B) = 0 + 0 = 0.$$

   Thus $A + B \in \mathbb{W}$.
3. Let $A \in \mathbb{W}$ and $c \in \mathbb{F}$, then $tr(c\,A) = c\,tr(A) = 0$ and hence $c\,A \in \mathbb{W}$.

Therefore, $\mathbb{W}$ is a subspace of $M_{n\times n}(\mathbb{F})$.

Example 1.3.4

Let $\mathbb{W} = \{ (x, y, z) : z = x - y \}$. Show that $\mathbb{W}$ is a subspace of $\mathbb{R}^3$.

Solution:

1. Clearly $\mathbf{0} = (0, 0, 0) \in \mathbb{W}$ since $0 = 0 - 0$.
2. Let $x = (a, b, c), y = (d, e, f) \in \mathbb{W}$. Then $c = a - b$ and $f = d - e$, and hence $x + y = (a + d, b + e, c + f)$ which is in $\mathbb{W}$ since

$$c + f = (a - b) + (d - e) = (a + d) - (b + e).$$

3. Let $x = (a, b, c) \in \mathbb{W}$ and $k \in \mathbb{F}$. Then $c = a - b$ and hence $kc = ka - kb$; that is $kx = (ka, kb, kc) \in \mathbb{W}$.

Therefore, $\mathbb{W}$ is a subspace of $\mathbb{R}^3$.

Definition 1.3.3

Let $\mathbb{P}(\mathbb{F})$ denote the set of all polynomials with coefficients from a field $\mathbb{F}$. For integer $n \geq 0$, let $\mathbb{P}_n(\mathbb{F})$ be the set of all polynomials of degree less than or equal n with coefficients from $\mathbb{F}$.

For instance, $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{P}_n(\mathbb{F})$. Note that $f(x) = 0$ means that $a_n = a_{n-1} = \cdots = a_1 = a_0 = 0$ and hence f is called the **zero polynomial**. For our convenience, we define the degree of the zero polynomial as -1.

Example 1.3.5

Show that $\mathbb{P}_n(\mathbb{F})$ is a subspace of $\mathbb{P}(\mathbb{F})$.

Solution:

1. Note that the zero polynomial is of degree -1 and hence it is in $\mathbb{P}_n(\mathbb{F})$.
2. Clearly the sum of two polynomial of degrees less than or equal n is another polynomial of degree less than or equal n.
3. The product of a scalar and a polynomial of degree less than or equal n is a polynomial of degree less than or equal n.

Therefore, $\mathbb{P}_n(\mathbb{F})$ is a subspace of $\mathbb{P}(\mathbb{F})$.

Exercise 1.3.1

Solve the following exercises from the book at pages 19 - 23:

- 6, 8 : a, b, c.
- 11.

Section 1.4: Linear Combinations and Systems of Linear Equations

Definition 1.4.1

Let $S = \{ x_1, x_2, \cdots, x_n \}$ be a nonempty subset of vectors in a vector space $\mathbb{V}$ over a field $\mathbb{F}$. A vector $x \in \mathbb{V}$ is called a **linear combination** of vectors in S if there exist $c_1, c_2, \cdots, c_n \in \mathbb{F}$ such that

$$x = c_1\, x_1 + c_2\, x_2 + \cdots + c_n\, x_n. \tag{1.4.1}$$

In that case, the scalars $c_1, c_2, \cdots, c_n$ are called the **coefficients** of the linear combination.

Recall from Math-111: To solve a system of linear equations $Ax = B$, we simplify the original system $[A|B]$ to its *reduced row echelon form* (r.r.e.f for short) using the following elementary row operations:

1. Interchanging two rows.
2. Multiplying a row by a nonzero scalar.
3. Adding a multiple of a row to another.

Example 1.4.1

Is $x = (2, 1, 5)$ a linear combination of $S = \{ x_1, x_2, x_3 \} \subseteq \mathbb{R}^3$, where $x_1 = (1, 2, 1), x_2 = (1, 0, 2)$, and $x_3 = (1, 1, 0)$? Explain.

Solution:

Note that x is a linear combination of $\{ x_1, x_2, x_3 \}$ if we find scalars $c_1, c_2, c_3 \in \mathbb{R}$ such that $x = c_1 x_1 + c_2 x_2 + c_3 x_3$. Thus, we consider

$$(2, 1, 5) = c_1(1, 2, 1) + c_2(1, 0, 2) + c_3(1, 1, 0).$$

That is

$$c_1 + c_2 + c_3 = 2$$
$$2c_1 + 0 + c_3 = 1$$
$$c_1 + 2c_2 + 0 = 5$$

We then find the r.r.e.f. of that system as follows:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & 0 & 1 & 1 \\ 1 & 2 & 0 & 5 \end{array}\right] \xrightarrow{\text{r.r.e.f.}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

That is, $c_1 = 1$, $c_2 = 2$, and $c_3 = -1$ and therefore $x = x_1 + 2x_2 - x_3$.

Definition 1.4.2

Let $S = \{ x_1, x_2, \cdots, x_n \}$ be a nonempty subset of vectors in a vector space $\mathbb{V}$ over a field $\mathbb{F}$. The **span** of S, denoted **span** S, is the set of all linear combinations of the vectors in S. For convenience, we define **span** $\phi = \{0\}$, where ϕ is the empty set.

Theorem 1.4.1

Let $S = \{ x_1, x_2, \cdots, x_n \}$ be a subset of vectors in a vector space $\mathbb{V}$. The **span** S is a subspace of $\mathbb{V}$.

Proof:

Proved in Math 111. Let $\mathbb{W} = \textbf{span } S = \{ z : z = c_1x_1 + \cdots c_nx_n \} \subseteq \mathbb{V}$. Then

1. $0z = 0x_1 + 0x_2 + \cdots + 0x_n = 0 \in \mathbb{W}$.
2. Let $z_1 = c_1x_1 + \cdots + c_nx_n$, $z_2 = d_1x_1 + \cdots + d_nx_n \in \mathbb{W}$. Then,

$$z_1 + z_2 = (c_1x_1 + \cdots + c_nx_n) + (d_1x_1 + \cdots + d_nx_n) = (c_1 + d_1)x_1 + \cdots + (c_n + d_n)x_n \in \mathbb{W}.$$

3. Let $z = c_1x_1 + \cdots + c_nx_n \in \mathbb{W}$ and let a be any scalar. Then

$$az = a(c_1x_1 + \cdots + c_nx_n) = ac_1x_1 + \cdots + ac_nx_n \in \mathbb{W}.$$

Therefore, $\mathbb{W}$ is a subspace of $\mathbb{V}$.

Example 1.4.2

Let $S = \{ 1 + x, 2 - x^2, 1 + x + x^2 \}$ be a subset of $\mathbb{P}_2(\mathbb{R})$. Is x^2 a linear combination of S? Explain.

Solution:

Considering the system $x^2 = c_1(1+x) + c_2(2-x^2) + c_3(1+x+x^2)$, we get

$$x^2 = (c_1 + 2c_2 + c_3)\cdot 1 + (c_1 + c_3)\cdot x + (-c_2 + c_3)\cdot x^2.$$

Hence

$$\begin{array}{ccccccc} c_1 & + & 2c_2 & + & c_3 & = & 0 \\ c_1 & + & 0 & + & c_3 & = & 0 \\ 0 & - & c_2 & + & c_3 & = & 1 \end{array}$$

We then find the r.r.e.f. of that system as follows:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{array}\right] \xrightarrow{\text{r.r.e.f.}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Therefore, $x^2 = -1\cdot(1+x) + 0\cdot(2-x^2) + 1\cdot(1+x+x^2)$, and x^2 is a linear combination of S.

Example 1.4.3: Solving Example 1.3.4 in a different way

Show that $\mathbb{W} = \{\,(x,y,z) \,:\, z = x - y\,\}$ is a subspace of $\mathbb{R}^3$.

Solution:

Note that $\mathbb{W} = \{\,(x,y,x-y) : x,y \in \mathbb{R}\,\} = \{\,x(1,0,1) + y(0,1,-2) : x,y \in \mathbb{R}\,\}$. That is, $\mathbb{W} = \textbf{span}\,\{\,(1,0,1),(0,1,-1)\,\}$. Therefore, $\mathbb{W}$ is a subspace of $\mathbb{R}^3$.

Example 1.4.4

Show that $\mathbb{W} = \left\{ \begin{pmatrix} a & a-b \\ a+b & b \end{pmatrix} \,:\, a,b \in \mathbb{R} \right\}$ is a subspace of $M_{2\times 2}(\mathbb{R})$.

Solution:

Clearly $\mathbb{W} = \left\{ a\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \,:\, a,b \in \mathbb{R} \right\} = \textbf{span}\left\{ \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \right\}$ and therefore it is a subspace of $M_{2\times 2}(\mathbb{R})$.

Example 1.4.5

Determine whether $x = \begin{pmatrix} 1 & 2 \\ -3 & 4 \end{pmatrix}$ is in the **span** S, where $S = \left\{ \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right\}$.

Solution:

Consider the system $\begin{pmatrix} 1 & 2 \\ -3 & 4 \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + c \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. Thus,

$$1 = a + c, 2 = b + c, -3 = -a, \text{ and } 4 = b.$$

Therefore, $a = 3, b = 4, c = -2$ and hence $x \in$ **span** S since it is a linear combination of S.

Definition 1.4.3

Let $S = \{ x_1, x_2, \cdots, x_n \}$ be a subset of a vector space $\mathbb{V}$. If every vector in $\mathbb{V}$ is a linear combination of S, we say that S **spans** (or **generates**) $\mathbb{V}$ or that $\mathbb{V}$ is **spanned** (or **generated**) by S.

Example 1.4.6

Show that $S = \{ x_1, x_2, x_3 \}$ spans $\mathbb{R}^3$, where $x_1 = (1, 1, 0)$, $x_2 = (1, 0, 1)$, and $x_3 = (0, 1, 1)$.

Solution (1):

Let $x = (a, b, c) \in \mathbb{R}^3$ by an arbitrary vector. Consider the system $x = c_1x_1 + c_2x_2 + c_3x_3$ and work its matrix form to get the system in its reduced form as follows:

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 1 & 0 & 1 & b \\ 0 & 1 & 1 & c \end{array}\right] \xrightarrow{\text{r.r.e.f.}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2}(a + b - c) \\ 0 & 1 & 0 & \frac{1}{2}(a - b + c) \\ 0 & 0 & 1 & \frac{1}{2}(-a + b + c) \end{array}\right]$$

Thus, $c_1 = \frac{1}{2}(a + b - c), c_2 = \frac{1}{2}(a - b + c), c_3 = \frac{1}{2}(-a + b + c)$ and hence S generates $\mathbb{R}^3$.

Solution (2):

We can solve the problem if we know that this system has at least one solution. So, we compute the determinant of the associate matrix to the system

$$\begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -2 \neq 0$$

Therefore, the system has a unique soltion and hence S spans $\mathbb{R}^3$.

Remark 1.4.1

For any nonnegative n, $S = \{1, x, x^2, \cdots, x^n\}$ spans $\mathbb{P}_n(\mathbb{R})$.

Example 1.4.7

Does the set $S = \{1 - x, x - x^2, 1 + x^2\}$ spans $\mathbb{P}_2(\mathbb{R})$? Explain.

Solution:

Consider any polynomial $ax^2 + bx + c \in \mathbb{P}_2(\mathbb{R})$. Then

$$ax^2 + bx + c = c_1(1-x) + c_2(x - x^2) + c_3(1 + x^2) = (c_1 + c_3)\cdot 1 + (-c_1 + c_2)\cdot x + (-c_2 + c_3)\cdot x^2.$$

Thus $\left[\begin{array}{ccc|c} 1 & 0 & 1 & c \\ -1 & 1 & 0 & b \\ 0 & -1 & 1 & a \end{array}\right]$ has a unique solution since

$$\begin{vmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{vmatrix} = 2 \neq 0.$$

Thus, S spans $\mathbb{P}_2(\mathbb{R})$.

Example 1.4.8: Exercise #13 @ page 34

Show that if S_1 and S_2 are subsets of a vector space $\mathbb{V}$ such that $S_1 \subseteq S_2$, then **span** $S_1 \subseteq$ **span** S_2. If moreover, **span** $S_1 = \mathbb{V}$, then **span** $S_2 = \mathbb{V}$.

Solution:

Let $S_1 = \{x_1, x_2, \cdots, x_k\} \subseteq S_2$ and let $x \in$ **span** S_1. Then x can be written as a linear combination of vectos of S_1; that is

$$x = c_1 x_1 + c_2 x_2 + \cdots + c_k x_k,$$

for some scalars $c_1, \cdots, c_k$. But then x is also a linear combination of vectors in S_2 since all vectors $x_1, \cdots, x_k \in S_2$. Thus **span** $S_1 \subseteq$ **span** S_2.

If **span** $S_1 = \mathbb{V}$, then we know that **span** S_2 is a subspace of $\mathbb{V}$ containing **span** $S_1 = \mathbb{V}$. Therefore, **span** $S_2 = \mathbb{V}$.

Exercise 1.4.1

Solve the following exercises from the book at pages 32 - 35:

- $2 : a, b, c,\ 3 : a, b, c,\ 4 : a, b.$
- $5 : a, b, e, f, g, h.$
- $6 - 9.$
- $13.$

Section 1.5: Linear Dependence and Linear Independence

It is clear that there are many different subsets that generates a subspace $\mathbb{W}$ of a vector space $\mathbb{V}$. In this section, we will try to get these subsets as small as possible by removing unnecessary vectors from those subsets.

Remark 1.5.1

$\mathbb{R}^n$ is generated by $\{E_1, E_2, \cdots, E_n\}$ where E_i is the vector whose all entries are 0 except for entry at position i which equals 1.

Definition 1.5.1

The set of vectors $S = \{x_1, x_2, \cdots, x_n\}$ in a vector space $\mathbb{V}$ is said to be **linearly dependent** if there exist scalars $c_1, c_2, \cdots, c_n$, not all zero, such that

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = 0. \tag{1.5.1}$$

Otherwise, S is said to be **linearly independent**. That is, if whenever Equation (1.5.1) hold, we must have $c_1 = c_2 = \cdots = c_n = 0$. In that case, we say that the zero vector has only the **trivial representation** as a linear combination of the vectors of S.

Remark 1.5.2

The homogenous system $Ax = 0$ (with a square matrix A) has only trivial solution if and only if $|A| \neq 0$.

Example 1.5.1

Determine whether the set $S = \{x_1 = (1, 0, 1), x_2 = (2, 1, 2), x_3 = (1, 1, 1)\}$ is linearly dependent or independent in $\mathbb{R}^3$.

Solution (1):

We consider the homogenous system: $c_1x_1 + c_2x_2 + c_3x_3 = 0$. Solving this system, we see that

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 \end{array}\right] \xrightarrow{\text{r.r.e.f.}} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

That is, $c_1 - c_3 = 0$ and $c_2 + c_3 = 0$. If $c_3 = t \in \mathbb{R}$, the system has non-trivial solutions $c_1 = t, c_2 = -t, c_3 = t$ and hence S is linearly dependent.

Solution (2):

Note that the determinant of matrix A (the matrix whose columns are the vectors of S) is 0, and hence the set S is linearly dependent.

Example 1.5.2

Find the values, if any, of α so that the set S is linearly independent in $\mathbb{R}^3$, where

$$S = \{\, x_1 = (\,-1, 0, -1\,), x_2 = (\,2, 1, 2\,), x_3 = (\,\alpha, 1, 1\,)\,\}$$

Solution:

Simply use the determinant of a matrix whose columns are the vectors of S. Consider the homogenous system $Ax = 0$ where $A = \begin{bmatrix} -1 & 2 & \alpha \\ 0 & 1 & 1 \\ -1 & 2 & 1 \end{bmatrix}$. Thus the system has only trivial solution if and only if S is linearly independent. Therefore, the $|\,A\,| \neq 0$. That is,

$$\begin{vmatrix} -1 & 2 & \alpha \\ 0 & 1 & 1 \\ -1 & 2 & 1 \end{vmatrix} \neq 0 \;\Leftrightarrow\; -1 \begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} - \begin{vmatrix} 2 & \alpha \\ 1 & 1 \end{vmatrix} \neq 0 \;\Leftrightarrow\; 1 - (2 - \alpha) \neq 0 \;\Leftrightarrow\; \alpha \neq 1.$$

Thus, S is linearly independent only if $\alpha \neq 1$.

Theorem 1.5.1

Let S_1 and S_2 be two subsets of a vector space $\mathbb{V}$ with $S_1 \subseteq S_2$. Then

1. If S_1 is linealry dependent, then S_2 is linearly dependent.
2. If S_2 is linearly independent, then S_1 is linealry independent.

Example 1.5.3: Exercise #2(a) @ page 40

Determine whether $S = \left\{ \begin{pmatrix} 1 & -3 \\ -2 & 4 \end{pmatrix}, \begin{pmatrix} -2 & 6 \\ 4 & -8 \end{pmatrix} \right\}$ is linearly dependent or linealry independent set in $M_{2\times 2}(\mathbb{R})$?

Solution:

Consider the system $a \begin{pmatrix} 1 & -3 \\ -2 & 4 \end{pmatrix} + b \begin{pmatrix} -2 & 6 \\ 4 & -8 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Thus, we solve the following system

$$\left[\begin{array}{cc|c} 1 & -2 & 0 \\ -3 & 6 & 0 \\ -2 & 4 & 0 \\ 4 & -8 & 0 \end{array}\right] \xrightarrow{\text{r.r.e.f.}} \left[\begin{array}{cc|c} 1 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

That is $a = 2b$ and the set is linearly dependent.

Example 1.5.4

Let $S = \{ 1 - x, x - x^2, -1 + x^2 \} \subseteq \mathbb{P}_2(\mathbb{R})$. Determine whether or not S is linearly dependent.

Solution:

Consider

$$c_1(1 - x) + c_2(x - x^2) + c_3(-1 + x^2) = 0$$
$$(c_1 - c_3) \cdot 1 + (-c_1 + c_2) \cdot x + (-c_2 + c_3) \cdot x^2 = 0$$

By equating the coefficients of x^n on both sides of the equation for $n = 0, 1, 2$, we obtain the following homogenous system:

$$c_1 - c_3 = 0$$
$$-c_1 + c_2 = 0$$
$$-c_2 + c_3 = 0$$

That is

$$\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

But

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

Which implies that the system has a non-trivial solution and hence S is linearly dependent.

Exercise 1.5.1

Solve the following exercises from the book at pages 40 - 42:

- $2 : a, b, c, d, e, f.$
- $4, 5, 6,$ 9

Exercise 1.5.2

Let x and y be two linearly independent vectors in a vector space $\mathbb{V}$. Show that the condition for the vectors $ax + by$ and $cx + dy$ to be linearly dependent is $ad - bc = 0$.

Solution:

Consider

$$r_1(ax + by) + r_2(cx + dy) = 0.$$

Then, $(r_1a + r_2c)x + (r_1b + r_2d)y = 0$ and hence $(r_1a + r_2c) = (r_1b + r_2d) = 0$ since x and y are linearly independent. Considering the second system

$$\begin{aligned} ar_1 + cr_2 &= 0 \\ br_1 + dr_2 &= 0 \end{aligned} \tag{1.5.2}$$

For $ax + by$ and $cx + dy$ to be linear dependent, we must have nontrivial solutions to the system represented in (1.5.2). That is , $\begin{vmatrix} a & c \\ b & d \end{vmatrix} = 0$. That is $ad - bc = 0$.

Section 1.6: Bases and Dimension

Let $\mathbb{V}$ be a vector space with a subspace $\mathbb{W}$. We note that if S is a generating set for $\mathbb{W}$ and no proper subset of S is a generating set for $\mathbb{W}$, then S must be a linearly independent set.

Definition 1.6.1

A set β of distinct nonzero vectors in a vector space $\mathbb{V}$ is called a **basis** for $\mathbb{V}$ if and only if

1. β spans (generates) $\mathbb{V}$, and
2. β is linearly independent set in $\mathbb{V}$.

Moreover, the **dimension** of $\mathbb{V}$ is the number of vectors in its finite basis β, denoted by $\mathbf{dim}(\mathbb{V})$. In that case, we say that $\mathbb{V}$ is a **finite-dimensional vector space**.

Remark 1.6.1

1. In $\mathbb{F}^n$, the set $\{ E_1 = (1, 0, \cdots, 0), E_2 = (0, 1, 0, \cdots, 0), \cdots, E_n = (0, \cdots, 0, 1) \}$ is a basis for $\mathbb{F}^n$. This basis is called **the standard basis** for $\mathbb{F}^n$. Therefore, $\mathbf{dim}(\mathbb{F}^n) = n$.
2. Let E^{ij} denote the matrix in $M_{m\times n}(\mathbb{F})$ whose all entries are 0 except the ij-entry is 1. The set $\{ E^{ij} : 1 \leq i \leq m, 1 \leq j \leq n \}$ is the standard basis for $M_{m\times n}(\mathbb{F})$. Therefore, $\mathbf{dim}(M_{m\times n}(\mathbb{F})) = mn$.
3. The set $\beta = \{ 1, x, x^2, \cdots, x^n \}$ is the standard basis for the vector space $\mathbb{P}_n(\mathbb{F})$, and therefore $\mathbf{dim}(\mathbb{P}_n(\mathbb{F})) = n + 1$.

Theorem 1.6.1

Let $\mathbb{V}$ be a vector space and $\beta = \{ x_1, x_2, \cdots, x_n \}$ be a nonempty subset of $\mathbb{V}$. Then β is a basis for $\mathbb{V}$ if and only if each $x \in \mathbb{V}$ can be uniquely expressed as a linear combination of vectors in β, that is, can be expressed in the form

$$x = a_1x_1 + a_2x_2 + \cdots a_nx_n, \qquad \text{for unique scalars } a_1, a_2, \cdots, a_n.$$

Proof:

Proved in Math-111. „ $\Rightarrow$": Let β be a basis for $\mathbb{V}$. If $x \in \mathbb{V}$, then $x \in \mathbf{span}\ \beta = \mathbb{V}$, and hence

$$x = a_1x_1 + a_2x_2 + \cdots + a_nx_n,$$

for some scalars $a_1, \cdots, a_n$. Assume that x has another expression as

$$x = b_1x_1 + b_2x_2 + \cdots + b_nx_n,$$

for some scalars $b_1, \cdots, b_n$. Thus

$$0 = x - x = (a_1 - b_1)x_1 + (a_2 - b_2)x_2 + \cdots + (a_n - b_n)x_n.$$

But β is linearly independent set and hence $a_i - b_i = 0$ and therefore $a_i = b_i$ for $i = 1, \cdots, n$. Thus, x has a unique expression as a linear combination of vectors in β.
„ $\Leftarrow$": Assume that every vector $x \in \mathbb{V}$ can be uniquely expressed as a linear combination of vectors in β. Then $\mathbb{V} = \textbf{span }\beta$.
Also, $0 \in \mathbb{V}$, and there is unique scalars $a_1, \cdots, a_n$ such that $0 = a_1x_1 + \cdots + a_nx_n$. Note that multiplying both sides by a constant does not change the expression by assumption. Hence, $a_1 = a_2 = \cdots = a_n = 0$. Thus β is linearly independent and hence β is a basis for $\mathbb{V}$.

Theorem 1.6.2

If a vector space $\mathbb{V}$ is generated by a finite set S, then some subset of S is a basis for $\mathbb{V}$.

Corollary 1.6.1

Every basis for a finite-dimensional vector space $\mathbb{V}$ contains the same number of vectors.

Theorem 1.6.3

Let $\mathbb{V}$ be an n-dimensional vector space and let $\beta = \{x_1, x_2, \cdots, x_n\}$ be a subset (with n vectors) of $\mathbb{V}$. Then,

1. If β spans $\mathbb{V}$, then β is a basis for $\mathbb{V}$.
2. If β is linearly independent, then β is a basis for $\mathbb{V}$.

Theorem 1.6.4

Let $\mathbb{W}$ be a subspace of a finite-dimensional vector space $\mathbb{V}$. Then $\mathbb{W}$ is finite-dimensional subspace and $\textbf{dim}(\mathbb{W}) \leq \textbf{dim}(\mathbb{V})$. Moreover, if $\textbf{dim}(\mathbb{W}) = \textbf{dim}(\mathbb{V})$, then $\mathbb{W} = \mathbb{V}$.

Example 1.6.1

Determine whether $S = \{ x_1 = (1, 0, -1), x_2 = (2, 5, 1), x_3 = (0, -4, 3) \}$ is a basis for $\mathbb{R}^3$.

Solution:

Note that S contains $3 = \mathbf{dim}(\mathbb{R}^3)$, and thus it is enough to show that S is linearly independent (or S spans $\mathbb{R}^3$). In either cases, we can simply show that the associate matrix of the system is not equal to zero. That is

$$\begin{vmatrix} 1 & 2 & 0 \\ 0 & 5 & -4 \\ -1 & 1 & 3 \end{vmatrix} = (15 + 4) - (-8) = 27 \neq 0.$$

Thus S is a basis for $\mathbb{R}^3$.

Example 1.6.2

Let $\mathbb{W} = \{ (x, y, z) : 2x + 3y - z = 0 \}$. Show that $\mathbb{W}$ is a subspace of $\mathbb{R}^3$ and find its dimension.

Solution:

Clearly, $\mathbb{W} = \{ (x, y, 2x + 3y) : x, y \in \mathbb{R} \} = \{ x(1, 0, 2) + y(0, 1, 3) \}$. Therefore, $\mathbb{W} = \mathbf{span}\ \{ (1, 0, 2), (0, 1, 3) \}$ which shows that $\mathbb{W}$ is a subspace of $\mathbb{R}^3$. Moreover, the set $\{ (1, 0, 2), (0, 1, 3) \}$ is linearly independent set and hence it is a basis for $\mathbb{W}$. Therefore, $\mathbf{dim}(\mathbb{W}) = 2$.

Example 1.6.3

Let $\mathbb{W} = \{ (x, y, z, w) : x + y + z = 0 \text{ and } w = 2x \}$.

1. Show that $\mathbb{W}$ is a subspace of $\mathbb{R}^4$.
2. Find a basis for $\mathbb{W}$.

Solution:

(1): Clearly,

$$\begin{aligned} \mathbb{W} = \{ (x, y, -x - y, 2x) : x, y \in \mathbb{R} \} &= \{ x(1, 0, -1, 2) + y(0, 1, -1, 0) : x, y \in \mathbb{R} \} \\ &= \mathbf{span}\ \{ (1, 0, -1, 2), (0, 1, -1, 0) \} \end{aligned}$$

Therefore, $\mathbb{W}$ is a subspace of $\mathbb{R}^4$.

(2): Consider the system $c_1(1,0,-1,2)+c_2(0,1,-1,0) = (0,0,0,0)$. It is clear that $c_1 = c_2 = 0$ and hence $\{(1,0,-1,2),(0,1,-1,0)\}$ is linearly independent set and is a basis for $\mathbb{W}$.

Example 1.6.4

Let $\mathbb{W} = \left\{ \begin{pmatrix} a+b & c \\ 2c & a-b \end{pmatrix} \in M_{2\times 2}(\mathbb{R}) \right\}$.

1. Show that $\mathbb{W}$ is a subspace of $M_{2\times 2}(\mathbb{R})$.
2. What is $\mathbf{dim}(\mathbb{W})$?

Solution:

(1): Note that

$$\begin{aligned} \mathbb{W} &= \left\{ a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \right\} \\ &= \mathbf{span}\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \right\}. \end{aligned}$$

So, $\mathbb{W}$ is a subspace of $M_{2\times 2}(\mathbb{R})$.

(2): Consider the homogenous system $c_1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + c_2\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + c_3\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Thus,

$$c_1 + c_2 = 0,\ c_3 = 0,\ 2c_3 = 0, \quad \text{and } c_1 - c_2 = 0.$$

Hence $c_1 = c_2 = c_3 = 0$. Therefore, $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \right\}$ is a basis for $\mathbb{W}$ and $\mathbf{dim}(\mathbb{W}) = 3$.

Example 1.6.5

Let $\mathbb{W} = \{ f(x) \in \mathbb{P}_2(\mathbb{R}) : f(1) = 0 \}$.

1. Show that $\mathbb{W}$ is a subspace of $\mathbb{P}_2(\mathbb{R})$.
2. What is $\mathbf{dim}(\mathbb{W})$?

Solution:

Note that $f(x) = a + bx + cx^2$ so that $f(1) = a + b + c = 0$. That is $c = -a - b$. Hence $f(x) = a + bx + (-a - b)x^2 = a(1 - x^2) + b(x - x^2)$. Therefore, $\mathbb{W} = \textbf{span}\ S$, where $S = \{1 - x^2, x - x^2\}$. Clearly, S is linearly independent (each element is not a composite of the other). Hence S is a basis for $\mathbb{W}$ and $\textbf{dim}(\mathbb{W}) = 2$.

Definition 1.6.2

Let $\mathbb{V}$ be a vector space with a basis $\beta = \{x_1, x_2, \cdots, x_n\}$. If $x \in \mathbb{V}$, then $x = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ is uniquely represented with scalars $c_1, c_2, \cdots, c_n$. We call thses scalars the **coordinates** of x in the basis β, denoted by

$$[x]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Example 1.6.6

Let $\beta = \{E_1, E_2, E_3\}$ be the standard basis for $\mathbb{R}^3$, and let $\gamma = \{x_1, x_2, x_3\}$, where $x_1 = (1, 1, 1), x_2 = (0, 1, 1)$, and $x_3 = (0, 0, 1)$.

1. Show that γ is another basis for $\mathbb{R}^3$.
2. Find $[x]_\beta$ and $[x]_\gamma$ for $x = (2, -1, 4)$.

Solution:

(1): Note that $|\beta| = |\gamma| = 3 = \textbf{dim}(\mathbb{R}^3)$. So, we only need to show that γ is linearly independent (or γ spans $\mathbb{R}^3$). Consider $c_1x_1 + c_2x_2 + c_3x_3 = 0$ which is a homogenous system with $Ax = 0$, where $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ and $x = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$. Clearly then $|A| = 1 \neq 0$ and hence γ is linearly independent and it is a basis for $\mathbb{R}^3$.

(2): Note that $[x]_\beta = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$ since $x = 2E_1 - E_2 + 4E_3$.

Now consider $c_1(1, 1, 1) + c_2(0, 1, 1) + c_3(0, 0, 1) = (2, -1, 4)$ to get $c_1 = 2, c_1 + c_2 = -1$, and

$c_1 + c_2 + c_3 = 4$. Therefore, $x = 2x_1 + (-3)x_2 + 5x_3$ and hence $[x]_\gamma = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$.

Exercise 1.6.1

Solve the following exercises from the book at pages 53 - 58:

- $2 : a, b,\ 3 : a, b.$
- $4, 5,\ 7.$
- $11, 12.$

Exercise 1.6.2

Let $\mathbb{W} = \{ f(x) \in \mathbb{P}_3(\mathbb{R}) : f(0) = f'(0) \text{ and } f(1) = f'(1) \}$. Find a basis for $\mathbb{W}$.

Solution:

Note that any $f(x) \in \mathbb{W}$ is of the form $f(x) = a + bx + cx^2 + dx^3$. Thus, $f(0) = f'(0)$ implies that $a = b$. Also, $f(1) = f'(1)$ implies $a+b+c+d = b+2c+3d$. These two equations implies $a = b = c + 2d$. Thus

$$f(x) = (c + 2d) + (c + 2d)x + cx^2 + dx^3 = c(1 + x + x^2) + d(2 + 2x + x^3).$$

Therefore, $\mathbb{W} = \textbf{span}\ \{ 1 + x + x^2, 2 + 2x + x^3 \}$. Clearly $S = \{ 1 + x + x^2, 2 + 2x + x^3 \}$ is a basis for $\mathbb{W}$.

Exercise 1.6.3

Let $\mathbb{W} = \{ a + bx + cx^2 \in \mathbb{P}_2(\mathbb{R}) : a = b = c \}$. Show that $\mathbb{W}$ is a subspace of $\mathbb{P}_2(\mathbb{R})$.

Solution:

Note that $\mathbb{W} = \{ a(1 + x + x^2) : a \in \mathbb{R} \}$. Thus, $\mathbb{W} = \textbf{span}\ S$, where $S = \{ 1 + x + x^2 \}$ and hence $\mathbb{W}$ is a subspace of $\mathbb{P}_2(\mathbb{R})$.

Exercise 1.6.4

Let $\mathbb{W} = \{ a + bx \in \mathbb{P}_1(\mathbb{R}) : b = a^2 \}$. Is $\mathbb{W}$ a subspace of $\mathbb{P}_1(\mathbb{R})$? Explain your answer.

Solution:

No. Clearly $f(x) = 1 + x, g(x) = 2 + 4x \in \mathbb{W}$, but $f(x) + g(x) = 3 + 5x \notin \mathbb{W}$.

Exercise 1.6.5

Exercise #11 @ page 55: Let x and y be distinct vectors of a vector space $\mathbb{V}$. Show that if $\beta = \{x, y\}$ is a basis for $\mathbb{V}$ and a and b are nonzero scalars, then both $\gamma_1 = \{x + y, ax\}$ and $\gamma_2 = \{ax, by\}$ are also bases for $\mathbb{V}$.

Solution:

Since β is a basis for $\mathbb{V}$, then $\mathbf{dim}(\mathbb{V}) = 2$. So it is enough to check if both γ_1 and γ_2 are linearly independent.

For γ_1: Assume that $s(x + y) + t(ax) = 0$. Then, $(s + ta)x + (s)y = 0$, and hence $s = 0$ and $s + ta = 0$ which implies that $t = 0$ since $a \neq 0$. Therefore, γ_1 is linearly independent and hence it is a basis for $\mathbb{V}$.

For γ_2: Assume that $s(ax) + t(by) = 0$. Then, $(sa)x + (tb)y = 0$ and hence $sa = tb = 0$ implies that $s = t = 0$ since a and b are both nonzero. Therefore, γ_2 is linearly independent and hence it is a basis for $\mathbb{V}$.

Chapter

2 Linear Transformations and Matrices

In this chapter we consider special functions defined on vector spaces that preserve the structure. These special functions are called **linear transformations**.

The preserved structure of vector space $\mathbb{V}$ over a field $\mathbb{F}$ is its *addition* and *scalar multiplication* operations, or, simply, its linear combinations.

Note that we assume that all vector spaces in this chapter are over a common field $\mathbb{F}$.

Section 2.1: Linear Transformations, Null Space, and Ranges

Definition 2.1.1

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces. A **linear transformation** $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ is a function such that:

1. $\mathbf{T}(x + y) = \mathbf{T}(x) + \mathbf{T}(y)$ for any $x, y \in \mathbb{V}$.
2. $\mathbf{T}(c\,x) = c\,\mathbf{T}(x)$ for any $c \in \mathbb{F}$ and any $x \in \mathbb{V}$.

Note that the addition operation in $x + y$ refers to that defined in $\mathbb{V}$, while the addition in $\mathbf{T}(x) + \mathbf{T}(y)$ refers to that defined in $\mathbb{W}$. Moreover, if $\mathbb{V} = \mathbb{W}$, we say that $\mathbf{T}$ is a **linear operator** on $\mathbb{V}$. We sometime simply call $\mathbf{T}$ **linear**.

Remark 2.1.1

Let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a function for vector spaces $\mathbb{V}$ and $\mathbb{W}$. Then for any scalar c, and any $x, y \in \mathbb{V}$, we have

1. If $\mathbf{T}$ is linear, then $\mathbf{T}(0_{\mathbb{V}}) = 0_{\mathbb{W}}$: For any $x \in \mathbb{V}$, $\mathbf{T}(0) = \mathbf{T}(0\,x) = 0\mathbf{T}(x) = 0$.
2. $\mathbf{T}$ is linear iff $\mathbf{T}(cx + y) = c\mathbf{T}(x) + \mathbf{T}(y)$.
3. $\mathbf{T}(x - y) = \mathbf{T}(x) - \mathbf{T}(y)$.
4. $\mathbf{T}$ is linear iff $\mathbf{T}\left(\sum_{i=1}^{n} c_i x_i\right) = \sum_{i=1}^{n} c_i \mathbf{T}(x_i)$, for scalars $c_1, \cdots, c_n$ and $x_1, \cdots, x_n \in \mathbb{V}$.

To see that a linear transformation $\mathbf{T} : \mathbb{V} \to \mathbb{V}$ preserves linear combination, assume that $v \in \mathbb{V}$

such that $v = 3s + 5t - 2u$ for some vectors $s, t, u \in \mathbb{V}$. Then, $\mathbf{T}(v) = \mathbf{T}(3s + 5t - 2u) = 3\mathbf{T}(s) + 5\mathbf{T}(t) - 2\mathbf{T}(u)$.

In what follows, we usually use property (2) above to prove that a given transformation is linear.

Definition 2.1.2

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces. We define the **trivial** linear transformation $\mathbf{T}_0 : \mathbb{V} \to \mathbb{W}$ defined by $\mathbf{T}_0(x) = 0$ for all $x \in \mathbb{V}$. Also, we define the **identity** linear transformation $\mathbf{I}_{\mathbb{V}} : \mathbb{V} \to \mathbb{V}$ defined by $\mathbf{T}(x) = x$ for all $x \in \mathbb{V}$.

Example 2.1.1

Define $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ by $\mathbf{T}(x, y) = (x, -y)$. Such linear transformation (show it) is called **reflection**.

Example 2.1.2

Define $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ by $\mathbf{T}(x, y) = (x, 0)$. Such linear transformation (show it) is called **projection**.

Example 2.1.3

Define $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\mathbf{T}(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Such linear transformation (show it) is called **rotation**.

Example 2.1.4

Define $\mathbf{T} : M_{m\times n}(\mathbb{F}) \to M_{n\times m}(\mathbb{F})$ by $\mathbf{T}(A) = A^t$. Show that $\mathbf{T}$ is linear.

Solution (1):

We show that $\mathbf{T}$ is linear by showing that $\mathbf{T}$ satisfies the conditions of the definition of linear transformation.

(1): For any $A, B \in M_{m\times n}(\mathbb{F})$, $\mathbf{T}(A + B) = (A + B)^t = A^t + B^t = \mathbf{T}(A) + \mathbf{T}(B)$.

(2): For any $c \in \mathbb{F}$ and any $A \in M_{m\times n}(\mathbb{F})$, $\mathbf{T}(cA) = (cA)^t = cA^t = c\mathbf{T}(A)$.

Therefore, $\mathbf{T}$ is linear.

Solution (2):

We use Remark 2.1.1 to show that $\mathbf{T}$ is linear. For all $A, B \in M_{m\times n}(\mathbb{F})$ and $c \in \mathbb{F}$, we have

$$\mathbf{T}(cA + B) = (cA + B)^t = (cA)^t + B^t = cA^t + B^t = c\mathbf{T}(A) + \mathbf{T}(B).$$

Therefore, $\mathbf{T}$ is linear.

Example 2.1.5

Show that $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$, defined by $\mathbf{T}(x, y) = (2x + y, x - y)$ is linear.

Solution:

We use Remark 2.1.1 to show that $\mathbf{T}$ is linear. Let $c \in \mathbb{R}$ and $(a, b), (x, y) \in \mathbb{R}^2$. Then

$$\begin{aligned}\mathbf{T}\big(c(a, b) + (x, y)\big) &= \mathbf{T}\big((ca + x, cb + y)\big) = \big(2(ca + x) + (cb + y), (ca + x) - (cb + y)\big)\\ &= \big((2ca + cb) + (2x + y), (ca - cb) + (x - y)\big)\\ &= (2ca + cb, ca - cb) + (2x + y, x - y) = c(2a + b, a - b) + (2x + y, x - y)\\ &= c\mathbf{T}(a, b) + \mathbf{T}(x, y).\end{aligned}$$

Therefore, $\mathbf{T}$ is linear.

Example 2.1.6

Define $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ by $\mathbf{T}\big(f(x)\big) = xf(x) + x^2$. Is $\mathbf{T}$ a linear transformation? Explain.

Solution:

For any $f(x), g(x) \in \mathbb{P}_2(\mathbb{R})$ and any $c \in \mathbb{R}$, we have

$$\mathbf{T}\big(cf(x) + g(x)\big) = x\big(cf(x) + g(x)\big) + x^2 = c(xf(x)) + xg(x) + x^2,$$

but

$$c\mathbf{T}\big(f(x)\big) + \mathbf{T}\big(g(x)\big) = c(xf(x) + x^2) + xg(x) + x^2 = c(xf(x)) + xg(x) + (c+1)x^2.$$

Therefore, $\mathbf{T}$ is not linear.

Example 2.1.7

Let $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}$ be a linear transformation for which $\mathbf{T}(3, -1, 2) = 5$ and $\mathbf{T}(1, 0, 1) = 2$. What is $\mathbf{T}(-1, 1, 0)$?

Solution:

We first write $(-1, 1, 0)$ as a linear combination of $(3, -1, 2)$ and $(1, 0, 1)$. Consider

$$(-1, 1, 0) = c_1(3, -1, 2) + c_2(1, 0, 1).$$

Thus, $c_1 = -1$ and $c_2 = 2$. Therefore,

$$\begin{aligned}
\mathbf{T}(-1, 1, 0) &= \mathbf{T}\Big[(-1)(3, -1, 2) + (2)(1, 0, 1)\Big] \\
&= -\mathbf{T}(3, -1, 2) + 2\mathbf{T}(1, 0, 1) \\
&= -1(5) + 2(2) = -5 + 4 = -1.
\end{aligned}$$

Example 2.1.8

Let $\mathbf{T} : \mathbb{P}_1(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ be a linear for which $\mathbf{T}(t + 1) = t^2 - 1$ and $\mathbf{T}(t - 1) = t^2 + t$. What is $\mathbf{T}(7t + 3)$?

Solution:

Consider $7t + 3 = c_1(t + 1) + c_2(t - 1)$ which implies that $c_1 + c_2 = 7$ and $c_1 - c_2 = 3$. That is, $c_1 = 5$, and $c_2 = 2$. Therefore,

$$\begin{aligned}
\mathbf{T}(7t + 3) &= \mathbf{T}\Big[5(t + 1) + 2(t - 1)\Big] \\
&= 5\mathbf{T}(t + 1) + 2\mathbf{T}(t - 1) \\
&= 5(t^2 - 1) + 2(t^2 + t) = 7t^2 + 2t - 5.
\end{aligned}$$

Definition 2.1.3

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces (over $\mathbb{F}$), and let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. The **null space** (or **kernel**) of $\mathbf{T}$, denoted $\mathcal{N}(\mathbf{T})$, is the set of all vectors $x \in \mathbb{V}$ such that

$\mathbf{T}(x) = 0$; that is

$$\mathcal{N}(\mathbf{T}) = \{\, x \in \mathbb{V} \,:\, \mathbf{T}(x) = 0 \,\} \subseteq \mathbb{V}.$$

The **range** (or **image**) of $\mathbf{T}$, denoted $\mathcal{R}(\mathbf{T})$, is the set of all images (under $\mathbf{T}$) of vectors in $\mathbb{V}$. That is

$$\mathcal{R}(\mathbf{T}) = \{\, \mathbf{T}(x) \,:\, x \in \mathbb{V} \,\} \subseteq \mathbb{W}.$$

Example 2.1.9

Find the null space and the range of: ① $\mathbf{I}_{\mathbb{V}} : \mathbb{V} \to \mathbb{V}$. ② $\mathbf{T}_0 : \mathbb{V} \to \mathbb{V}$.

Solution:

$$\text{①}: \ \mathcal{N}(\mathbf{I}_{\mathbb{V}}) = \{\, x \in \mathbb{V} \,:\, \mathbf{I}_{\mathbb{V}}(x) = 0 \,\} = \{0\}.$$

$$\text{①}: \ \mathcal{R}(\mathbf{I}_{\mathbb{V}}) = \{\, \mathbf{I}_{\mathbb{V}}(x) \,:\, x \in \mathbb{V} \,\} = \mathbb{V}.$$

$$\text{②}: \ \mathcal{N}(\mathbf{T}_0) = \{\, x \in \mathbb{V} \,:\, \mathbf{T}_0(x) = 0 \,\} = \mathbb{V}.$$

$$\text{②}: \ \mathcal{R}(\mathbf{T}_0) = \{\, \mathbf{T}_0(x) \,:\, x \in \mathbb{V} \,\} = \{0\}.$$

Theorem 2.1.1

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces and $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be linear. Then $\mathcal{N}(\mathbf{T})$ and $\mathcal{R}(\mathbf{T})$ are subspaces of $\mathbb{V}$ and $\mathbb{W}$, respectively.

Proof:

We first show that $\mathcal{N}(\mathbf{T})$ is a subspace of $\mathbb{V}$:

1. $\mathbf{T}(0_{\mathbb{V}}) = 0_{\mathbb{W}}$ and hence $0_{\mathbb{V}} \in \mathcal{N}(\mathbf{T})$.
2. Let $x, y \in \mathcal{N}(\mathbf{T})$, then $\mathbf{T}(x) = \mathbf{T}(y) = 0_{\mathbb{W}}$ and

$$\mathbf{T}(x+y) = \mathbf{T}(x) + \mathbf{T}(y) = 0_{\mathbb{W}} + 0_{\mathbb{W}} = 0_{\mathbb{W}} \qquad \Rightarrow \qquad x + y \in \mathcal{N}(\mathbf{T}).$$

3. Let $c \in \mathbb{F}$ and $x \in \mathcal{N}(\mathbf{T})$, then $\mathbf{T}(cx) = c\mathbf{T}(x) = c0_{\mathbb{W}} = 0_{\mathbb{W}}$, and hence $cx \in \mathcal{N}(\mathbf{T})$.

Therefore, $\mathcal{N}(\mathbf{T})$ is a subspace of $\mathbb{V}$.

Next we show that $\mathcal{R}(\mathbf{T})$ is a subspace of $\mathbb{W}$.

1. $\mathbf{T}(0_{\mathbb{V}}) = 0_{\mathbb{W}}$ and hence $0_{\mathbb{W}} \in \mathcal{R}(\mathbf{T})$.
2. Let $x, y \in \mathcal{R}(\mathbf{T})$, then there exist $u, v \in \mathbb{V}$ such that $\mathbf{T}(u) = x$ and $\mathbf{T}(v) = y$ and hence
$$\mathbf{T}(u+v) = \mathbf{T}(u) + \mathbf{T}(v) = x + y \qquad \Rightarrow \qquad x + y \in \mathcal{R}(\mathbf{T}).$$
3. Let $c \in \mathbb{F}$ and $x \in \mathcal{R}(\mathbf{T})$, then there exists $u \in \mathbb{V}$ such that $\mathbf{T}(u) = x$, and as $cu \in \mathbb{V}$, we have $\mathbf{T}(cu) = c\mathbf{T}(u) = cx \in \mathcal{R}(\mathbf{T})$.

Therefore, $\mathcal{R}(\mathbf{T})$ is a subspace of $\mathbb{W}$.

Remark 2.1.2

The next theorem provides a method for finding a spanning set (and therefore a basis) for the range of $\mathbf{T}$, namely for $\mathcal{R}(\mathbf{T})$.

Theorem 2.1.2

Let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. If $\beta = \{ x_1, x_2, \cdots, x_n \}$ is a basis for $\mathbb{V}$, then
$$\mathcal{R}(\mathbf{T}) = \textbf{span } \mathbf{T}(\beta) = \textbf{span } \{ \mathbf{T}(x_1), \mathbf{T}(x_2), \cdots, \mathbf{T}(x_n) \}.$$

Proof:

Since $x_i \in \mathbb{V}$, then $\mathbf{T}(x_i) \in \mathcal{R}(\mathbf{T})$ for each i. Because $\mathcal{R}(\mathbf{T})$ is a subspace of $\mathbb{W}$, $\mathcal{R}(\mathbf{T})$ contains $\textbf{span } \{ \mathbf{T}(x_1), \mathbf{T}(x_2), \cdots, \mathbf{T}(x_n) \} = \textbf{span } \mathbf{T}(\beta)$. Thus, $\textbf{span } \mathbf{T}(\beta) \subseteq \mathcal{R}(\mathbf{T})$.
Now suppose that $y \in \mathcal{R}(\mathbf{T})$. Then $y = \mathbf{T}(x)$ for some $x \in \mathbb{V}$. But because β is a basis for $\mathbb{V}$, we have $x = \sum_{i=1}^{n} c_i x_i$, for $c_1, c_2, \cdots, c_n \in \mathbb{F}$. Thus,
$$\begin{aligned} y = \mathbf{T}(x) &= \mathbf{T}(c_1x_1 + c_2x_2 + \cdots + c_nx_n) \\ &= c_1\mathbf{T}(x_1) + c_2\mathbf{T}(x_2) + \cdots + c_n\mathbf{T}(x_n). \end{aligned}$$

Thus, $y \in \textbf{span } \mathbf{T}(\beta)$. Hence $\mathcal{R}(\mathbf{T}) \subseteq \textbf{span } \mathbf{T}(\beta)$. Therefore, $\mathcal{R}(\mathbf{T}) = \textbf{span } \mathbf{T}(\beta)$.

Example 2.1.10

Let $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ be a linear transformation defined by

$$\mathbf{T}\big(f(x)\big) = \begin{pmatrix} f(1) - f(2) & 0 \\ 0 & f(0) \end{pmatrix}.$$

Find a basis for $\mathcal{R}(\mathbf{T})$.

Solution:

Consider the standard basis $\beta = \{1, x, x^2\}$ for $\mathbb{P}_2(\mathbb{R})$. Then

$$\begin{aligned} \mathcal{R}(\mathbf{T}) = \textbf{span } \mathbf{T}(\beta) &= \textbf{span } \left\{ \mathbf{T}(1), \mathbf{T}(x), \mathbf{T}(x^2) \right\} \\ &= \textbf{span } \left\{ \begin{pmatrix} 1-1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1-2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1-4 & 0 \\ 0 & 0 \end{pmatrix} \right\} \\ &= \textbf{span } \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -3 & 0 \\ 0 & 0 \end{pmatrix} \right\}. \end{aligned}$$

Considering the system $c_1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + c_2 \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + c_3 \begin{pmatrix} -3 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, we note that $c_2 = -3c_3$. Thus

$$\mathcal{R}(\mathbf{T}) = \textbf{span } \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \right\}.$$

Therefore, $\left\{ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \right\}$ is a basis for $\mathcal{R}(\mathbf{T})$ and $\textbf{dim}(\mathcal{R}(\mathbf{T})) = 2$.

Definition 2.1.4

Let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. If $\mathcal{N}(\mathbf{T})$ and $\mathcal{R}(\mathbf{T})$ are finite-dimensional, then we define the **nullity** of $\mathbf{T}$, denoted *nullity*$(\mathbf{T})$, and the **rank** of $\mathbf{T}$, denoted rank$(\mathbf{T})$, to be the dimensions of $\mathcal{N}(\mathbf{T})$ and $\mathcal{R}(\mathbf{T})$, respectively.

Theorem 2.1.3

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces, and let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. If $\mathbb{V}$ is finite-demensional, then

$$nullity(\mathbf{T}) + rank(\mathbf{T}) = \mathbf{dim}(\mathbb{V}).$$

Definition 2.1.5

Let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then $\mathbf{T}$ is said to be **one-to-one** (or simply $1-1$) if for all $x, y \in \mathbb{V}$, if $\mathbf{T}(x) = \mathbf{T}(y)$, then $x = y$.

Moreover, $\mathbf{T}$ is said to be **onto** $\mathbb{W}$ if $\mathcal{R}(\mathbf{T}) = \mathbb{W}$. That is for all $y \in \mathbb{W}$, there is $x \in \mathbb{V}$ such that $\mathbf{T}(x) = y$.

Theorem 2.1.4

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces, and let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then, $\mathbf{T}$ is ono-to-one iff $\mathcal{N}(\mathbf{T}) = \{0\}$.

Proof:

„ $\Rightarrow$": Assume that $\mathbf{T}$ is 1-1. If $x \in \mathcal{N}(\mathbf{T})$, then $\mathbf{T}(x) = 0 = \mathbf{T}(0)$, and hence $x = 0$. Therefore, $\mathcal{N}(\mathbf{T}) = \{0\}$.

„ $\Leftarrow$": Now let $\mathcal{N}(\mathbf{T}) = \{0\}$. Assume that $\mathbf{T}(x) = \mathbf{T}(y)$ for $x, y \in \mathbb{V}$. Then,

$$\mathbf{T}(x) - \mathbf{T}(y) = \mathbf{T}(x - y) = 0.$$

Hence $x - y \in \mathcal{N}(\mathbf{T}) = \{0\}$ and thus $x - y = 0$ which implies that $x = y$. Therefore, $\mathbf{T}$ is 1-1.

Theorem 2.1.5

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces <u>**of equal finite dimension**</u>, and $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then the following statements are equivalent:

1. $\mathbf{T}$ is 1-1
2. $\mathbf{T}$ is onto.
3. $\mathcal{N}(\mathbf{T}) = \{0\}$.
4. $rank(\mathbf{T}) = \mathbf{dim}(\mathbb{V})$.
5. $nullity(\mathbf{T}) = 0$.

Proof:

Note that $nullity(\mathbf{T}) + rank(\mathbf{T}) = \mathbf{dim}(\mathbb{V})$. Then,

$$\begin{aligned} \mathbf{T} \text{ is 1-1} \quad &\Leftrightarrow \quad \mathcal{N}(\mathbf{T}) = \{0\} \quad \Leftrightarrow \quad nullity(\mathbf{T}) = 0 \\ &\Leftrightarrow \quad rank(\mathbf{T}) = \mathbf{dim}(\mathbb{V}) \quad \Leftrightarrow \quad rank(\mathbf{T}) = \mathbf{dim}(\mathcal{R}(\mathbf{T})) = \mathbf{dim}(\mathbb{W}) \\ &\Leftrightarrow \quad \mathcal{R}(\mathbf{T}) = \mathbb{W} \quad \Leftrightarrow \quad \mathbf{T} \text{ is onto.} \end{aligned}$$

Example 2.1.11

Let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be linear transformation defined by

$$\mathbf{T}(x, y) = (2x - 3y, y).$$

Show that $\mathbf{T}$ is 1-1 and onto.

Solution:

We simply show that $\mathcal{N}(\mathbf{T}) = \{(0, 0)\}$.

$$\begin{aligned} \mathcal{N}(\mathbf{T}) &= \left\{ (x, y) \in \mathbb{R}^2 \,:\, \mathbf{T}(x, y) = (0, 0) \right\} \\ &= \left\{ (x, y) \in \mathbb{R}^2 \,:\, (2x - 3y, y) = (0, 0) \right\} \\ &= \left\{ (x, y) \in \mathbb{R}^2 \,:\, 2x - 3y = 0 \text{ and } y = 0 \right\} \\ &= \{ (0, 0) \}. \end{aligned}$$

Therefore, $\mathbf{T}$ is 1-1 and onto.

Example 2.1.12

Let $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by $\mathbf{T}(x, y, z) = (x, y)$. Find $\mathcal{N}(\mathbf{T}), \mathcal{R}(\mathbf{T}), nullity(\mathbf{T})$ and $rank(\mathbf{T})$.

Solution:

First,

$$\mathcal{N}(\mathbf{T}) = \left\{ (x, y, z) \in \mathbb{R}^3 : \mathbf{T}(x, y, z) = (x, y) = (0, 0) \right\} = \{ (0, 0, z) : z \in \mathbb{R} \}.$$

Thus $\{ (0, 0, 1) \}$ is a basis for $\mathcal{N}(\mathbf{T})$ and hence $nullity(\mathbf{T}) = 1$.

Next,

$$\mathcal{R}(\mathbf{T}) = \left\{ \mathbf{T}(x,y,z) = (x,y) \in \mathbb{R}^2 \right\} = \{ x(1,0) + y(0,1) : x, y \in \mathbb{R} \} = \mathbb{R}^2.$$

Thus, $rank(\mathbf{T}) = 2$.

Example 2.1.13

Let $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ be the linear transformation defined by $\mathbf{T}(f(x)) = f'(x) + \int_0^x f(t)\,dt$. (1) Is $\mathbf{T}$ one-to-one? (2) Is $\mathbf{T}$ onto? Explain.

Solution:

(1) : We show that $\mathbf{T}$ is 1-1 iff $\mathcal{N}(\mathbf{T}) = \{0\}$. Consider the basis $\beta = \{1, x, x^2\}$ for $\mathbb{P}_2(\mathbb{R})$. Then,

$$\mathcal{R}(\mathbf{T}) = \textbf{span}\left\{ \mathbf{T}(1), \mathbf{T}(x), \mathbf{T}(x^2) \right\} = \textbf{span}\left\{ x, 1 + \frac{x^2}{2}, 2x + \frac{x^3}{3} \right\}.$$

Since $\left\{ x, 1 + \frac{x^2}{2}, 2x + \frac{x^3}{3} \right\}$ is linearly independent set (It can be shown easily), it is a basis for $\mathcal{R}(\mathbf{T})$. Thus, $rank(\mathbf{T}) = \mathbf{dim}(\mathcal{R}(\mathbf{T})) = 3 = \mathbf{dim}(\mathbb{P}_2(\mathbb{R}))$. Therefore, $nullity(\mathbf{T}) = 0$ and hence $\mathcal{N}(\mathbf{T}) = \{0\}$ and then $\mathbf{T}$ is 1-1.

(2) : $rank(\mathbf{T}) = 3 < \mathbf{dim}(\mathbb{P}_3(\mathbb{R}))$ and hence $\mathcal{R}(\mathbf{T}) \neq \mathbb{P}_3(\mathbb{R})$. Therefore, $\mathbf{T}$ is not onto.

Example 2.1.14

For each of the following linear transformations, determine $\mathcal{N}(\mathbf{T})$ and $\mathcal{R}(\mathbf{T})$; find their bases; is $\mathbf{T}$ 1-1 or onto? Explain.

1. $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^2$ given by $\mathbf{T}(x,y,z) = (x - y, 2z)$.
2. $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$ given by $\mathbf{T}(x,y) = (x + y, 0, 2x - y)$.
3. $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^2$ given by $\mathbf{T}(x,y,z) = (x + y, x - y)$.
4. $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$ given by $\mathbf{T}(x,y) = (x + y, x - y, x)$.

Solution:

(1):

$$\begin{aligned}\mathcal{N}(\mathbf{T}) &= \{\,(x,y,z)\,:\,\mathbf{T}(x,y,z)=(0,0)\,\}\\ &= \{\,(x,y,z)\,:\,x-y=0 \text{ and } 2z=0\,\}\\ &= \{\,(x,x,0)\,:\,x\in\mathbb{R}\,\}=\{\,x(1,1,0)\,\}.\end{aligned}$$

Then, $nullity(\mathbf{T})=1$ since $\{\,(1,1,0\,\}$ is a basis for $\mathcal{N}(\mathbf{T})$, and $\mathbf{T}$ is not 1-1. Note that $rank(\mathbf{T}) = 3 - nullity(\mathbf{T}) = 2$. Thus, $rank(\mathbf{T}) = 2$ and hence $\mathcal{R}(\mathbf{T}) = \mathbb{R}^2$. Therefore, $\{\,(1,0),(0,1)\,\}$ is a basis for $\mathcal{R}(\mathbf{T})$ and $\mathbf{T}$ is onto. We note that we can compute $\mathcal{R}(\mathbf{T})$ by considering

$$\mathcal{R}(\mathbf{T}) = \mathbf{span}\ \{\,\mathbf{T}(1,0,0),\mathbf{T}(0,1,0),\mathbf{T}(0,0,1)\,\}.$$

Parts (2), (3), and (4) are left as exercises. (2):

$$\begin{aligned}\mathcal{N}(\mathbf{T}) &= \{\,(x,y)\,:\,\mathbf{T}(x,y)=(0,0,0)\,\}\\ &= \{\,(x,y)\,:\,x+y=0 \text{ and } 2x-y=0\,\}=\{(0,0,0)\}\end{aligned}$$

Thus, $nullity(\mathbf{T})=0$ and $\mathbf{T}$ is 1-1 and basis for $\mathcal{N}(\mathbf{T})=\phi$. The $rank(\mathbf{T}) = \mathbf{dim}(\mathcal{R}(\mathbf{T})) = 2 < \mathbf{dim}(\mathbb{R}^3)$ and hence $\mathbf{T}$ is not onto.

$$\begin{aligned}\mathcal{R}(\mathbf{T}) &= \mathbf{span}\ \{\,\mathbf{T}(1,0),\mathbf{T}(0,1)\,\}\\ &= \mathbf{span}\ \{\,(1,0,2),(1,0,-1)\,\}\end{aligned}$$

It is clear that $\{\,(1,0,2),(1,0,-1)\,\}$ is linearly independent and hence is a basis for $\mathcal{R}(\mathbf{T})$.

Theorem 2.1.6

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces, and suppose that $\{\,x_1,x_2,\cdots,x_n\,\}$ is a basis for $\mathbb{V}$. For any vectors $y_1,y_2,\cdots,y_n\in\mathbb{W}$, there exists exactly one linear transformation $\mathbf{T}:\mathbb{V}\to\mathbb{W}$ such that $\mathbf{T}(x_i)=y_i$ for $i=1,\cdots,n$.

Corollary 2.1.1

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces, and suppose that $\mathbb{V}$ has a finite basis $\{x_1,x_2,\cdots,x_n\}$. If $\mathbf{T},\mathbf{U}:\mathbb{V}\to\mathbb{W}$ are linear transformations and $\mathbf{U}(x_i)=\mathbf{T}(x_i)$ for $i=1,\cdots,n$, then $\mathbf{U}=\mathbf{T}$.

Example 2.1.15

Let $\mathbf{U}, \mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be linear transformations and let $\mathbf{T}$ be defined by $\mathbf{T}(x, y) = (2y - x, 3x)$. If $\mathbf{U}(1, 2) = (3, 3)$ and $\mathbf{U}(1, 1) = (1, 3)$, show that $\mathbf{T} = \mathbf{U}$.

Solution:

Note that $\{\,(1, 1), (1, 2)\,\}$ is a basis for $\mathbb{R}^2$ and that $\mathbf{T}(1, 1) = (1, 3) = \mathbf{U}(1, 1)$ and $\mathbf{T}(1, 2) = (3, 3) = \mathbf{U}(1, 2)$. Therefore, $\mathbf{U} = \mathbf{T}$.

Exercise 2.1.1

Solve the following exercises from the book at pages 74 - 79:

- $2, 3, 4, 5.$
- $8, 11, 12, 13.$

Exercise 2.1.2

Show that $\mathbf{T} : \mathbb{R}^4 \to \mathbb{R}^2$ defined by $\mathbf{T}(x, y, z, w) = (x, y)$ is linear.

Solution:

Let $k \in \mathbb{R}$ and $(a, b, c, d), (x, y, z, w) \in \mathbb{R}^4$. Then

$$\begin{aligned}
\mathbf{T}\Big(k(a, b, c, d) + (x, y, z, w)\Big) &= \mathbf{T}\Big((ka, kb, kc, kd) + (x, y, z, w)\Big) \\
&= \mathbf{T}(ka + x, kb + y, kc + z, kd + w) = (ka + x, kb + y) \\
&= k(a, b) + (x, y) \\
&= k\mathbf{T}(a, b, c, d) + \mathbf{T}(x, y, z, w).
\end{aligned}$$

Exercise 2.1.3

Show that $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\mathbf{T}(x, y) = (x + y, 3x)$ is linear.

Solution:

Let $k \in \mathbb{R}$ and $(a, b), (x, y) \in \mathbb{R}^2$. Then

$$\begin{aligned}
\mathbf{T}\left(k(a, b) + (x, y)\right) &= \mathbf{T}\left((ka, kb) + (x, y)\right) \\
&= \mathbf{T}\left(ka + x, kb + y\right) = \left((ka + x) + (kb + y), 3ka + 3x\right) \\
&= (ka + kb, 3ka) + (x + y, 3x) \\
&= k\mathbf{T}(a, b) + \mathbf{T}(x, y).
\end{aligned}$$

Exercise 2.1.4

Let $\mathbf{C}(\mathbb{R})$ denote the set of all real valued continuous functions on $\mathbb{R}$. Define $\mathbf{T} : \mathbf{C}(\mathbb{R}) \to \mathbb{R}$ by $\mathbf{T}\Big(f(x)\Big) = \displaystyle\int_a^b f(x)\,dx$ for all $a, b \in \mathbb{R}$ with $a < b$. Show that $\mathbf{T}$ is linear.

Solution:

For any $f(x), g(x) \in \mathbf{C}(\mathbb{R})$ and any $c \in \mathbb{R}$, we have

$$\begin{aligned}\mathbf{T}\Big(cf(x) + g(x)\Big) &= \int_a^b \Big(cf(x) + g(x)\Big)\ dx = c\int_a^b f(x)\ dx + \int_a^b g(x)\ dx \\ &= c\mathbf{T}\Big(f(x)\Big) + \mathbf{T}\Big(g(x)\Big).\end{aligned}$$

Therefore, $\mathbf{T}$ is linear.

Exercise 2.1.5

Let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation defined by $\mathbf{T}(x, y) = (2x + y, x - y)$. Find $\mathcal{N}(\mathbf{T}), \mathcal{R}(\mathbf{T}), nullity(\mathbf{T})$ and $rank(\mathbf{T})$.

Solution:

$\mathcal{N}(\mathbf{T}) = \{\, (x, y) \in \mathbb{R}^2 \,:\, \mathbf{T}(x, y) = (2x + y, x - y) = (0, 0) \,\}$. Thus, $2x + y = 0$ and $x - y = 0$ which implies that $x = y = 0$. Thus, $\mathcal{N}(\mathbf{T}) = \{(0, 0)\}$. Therefore, $nullity(\mathbf{T}) = 0$ and hence $rank(\mathbf{T}) = 2 = \mathbf{dim}(\mathbb{R}^2)$. Therefore, $\mathcal{R}(\mathbf{T}) = \mathbb{R}^2$, and we are done.

Or, we can compute the basis of $\mathcal{R}(\mathbf{T})$ as follows

$$\mathcal{R}(\mathbf{T}) = \Big\{\, \mathbf{T}(x, y) = (2x + y, x - y) \in \mathbb{R}^2 \,\Big\} = \{\, x(2, 1) + y(1, -1) \,\}.$$

Therefore, $\{\, (2, 1), (1, -1) \,\}$ is a basis for $\mathcal{R}(\mathbf{T})$ and $rank(\mathbf{T}) = 2$.

Section 2.2: The Matrix Representation of Linear Transformation

In this section, we consider the *representation* of a linear transformation by a matrix. That is, we develope a one-to-one correspondence between matrices and linear transformations that allows us to utilize properties of one to study properties of the other.

Definition 2.2.1

Let $\mathbb{V}$ be a finite-dimensional vector space. An **ordered basis** for $\mathbb{V}$ is a finite sequence of linearly independent vectors in $\mathbb{V}$ that generates $\mathbb{V}$.

Remark 2.2.1

Note that $\beta_1 = \{E_1, E_2, E_3\}$ can be considered as ordered basis for $\mathbb{R}^3$, while $\beta_2 = \{E_2, E_1, E_3\}$ is also an ordered basis for $\mathbb{R}^3$, but $\beta_1 \neq \beta_2$ as ordered basis.
In particular, $\{E_1, \cdots, E_n\}$ is the **standard ordered basis** for $\mathbb{R}^n$. Also, $\{1, x, x^2, \cdots, x^n\}$ is the **standard ordered basis** for $\mathbb{P}_n(\mathbb{R})$.

Definition 2.2.2

Let $\beta = \{x_1, \cdots, x_n\}$ be an ordered basis for a finite-dimensional vector space $\mathbb{V}$. For $x \in \mathbb{V}$, let $c_1, \cdots, c_n \in \mathbb{F}$ be the unique scalars such that $x = c_1x_1 + c_2x_2 + \cdots + c_nx_n$. We define the **coordinate vector of x relative to β**, denoted $[x]_\beta$, by

$$[x]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Example 2.2.1

Consider the vector space $\mathbb{P}_3(\mathbb{R})$ and the standard ordered basis $\beta = \{1, x, x^2, x^3\}$. Find the coordinate vector of $f(x) = 3 + 7x - 9x^2$ relative to β.

Solution:

Clearly $f(x) = 3 + 7x - 9x^2 = 3 \cdot 1 + 7 \cdot x + (-9) \cdot x^2 + 0 \cdot x^3$, and hence

$$[f(x)]_\beta = (3, 7, -9, 0) = [3 \quad 7 \quad -9 \quad 0]^t.$$

Definition 2.2.3

Let $\mathbb{V}$ and $\mathbb{W}$ be two finite-dimensional vector spaces with ordered bases $\beta = \{ x_1, x_2, \cdots, x_n \}$ and $\gamma = \{ y_1, y_2, \cdots, y_m \}$, respectively, and let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. For each j, $1 \leq j \leq n$, we have $\mathbf{T}(x_j) \in \mathbb{W}$ and there exist unique scalars $c_{ij} \in \mathbb{F}$, $1 \leq i \leq m$, such that

$$\mathbf{T}(x_j) = \sum_{i=1}^{m} c_{ij}\, y_i.$$

Then the $m \times n$ matrix $A = (c_{ij})$ is called the **matrix representation of T in the ordered bases β and γ** and is written $A = [\mathbf{T}]_\beta^\gamma$. If $\mathbb{V} = \mathbb{W}$ and $\beta = \gamma$, then we write simply $A = [\mathbf{T}]_\beta$. Note that the j^{th} column of $A = [\mathbf{T}]_\beta^\gamma$ then is simply $[\mathbf{T}(x_j)]_\gamma$. That is,

$$A = \Big[[\mathbf{T}(x_1)]_\gamma \quad [\mathbf{T}(x_2)]_\gamma \quad \cdots \quad [\mathbf{T}(x_n)]_\gamma \Big].$$

Remark 2.2.2

Following Definition 2.2.3, the following statements hold:

1. If $\mathbf{U} : \mathbb{V} \to \mathbb{W}$ is a linear transformation such that $[\mathbf{U}]_\beta^\gamma = [\mathbf{T}]_\beta^\gamma$, then $\mathbf{U} = \mathbf{T}$.
2. If $x \in \mathbb{V}$, then $[\mathbf{T}(x)]_\gamma = A\ [x]_\beta$, where $[x]_\beta$ and $[\mathbf{T}(x)]_\gamma$ are the coordinate vectors of x and $\mathbf{T}(x)$, respectively, with respect to the respective bases β and γ.
3. If $x \in \mathbb{V}$, then $\mathbf{T}\,(x) = \sum_{i=1}^{m} \left([\mathbf{T}(x)]_\gamma \right)_i\, y_i = \sum_{i=1}^{m} c_i\, y_i$.

Remark 2.2.3

⋆ Finding $[\mathbf{T}]_\beta^\gamma$:

Let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be linear transformation from n-dimensional vector space $\mathbb{V}$ into m-dimensional vector space $\mathbb{W}$, and let $\beta = \{ x_1, \cdots, x_n \}$ and $\gamma = \{ y_1, \cdots, y_m \}$ be bases for $\mathbb{V}$ and $\mathbb{W}$, respectively. Then we compute the matrix representation of $\mathbf{T}$ as follows:

1. Compute $\mathbf{T}(x_j)$ for $j = 1, 2, \cdots, n$.
2. Find the coordinate vector $[\mathbf{T}(x_j)]_\gamma$ for $\mathbf{T}(x_j)$ with respect to γ. That is, express $\mathbf{T}(x_j)$ as a linear combination of vectors in γ.
3. Form the matrix representation A of $\mathbf{T}$ with respect to β and γ by choosing $[\mathbf{T}(x_j)]_\gamma$ as the j^{th} column of A.

Example 2.2.2

Let $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear defined by $\mathbf{T}(x, y, z) = (x+y, y-z)$. Find a matrix representation A for $\mathbf{T}$. Use A to evaluate $\mathbf{T}(u)$, where $u = (1, 2, 3)$.

Solution:

We use the method described in Remark 2.2.3 and consider $\beta = \{ (1,0,0), (0,1,0), (0,0,1) \}$ and $\gamma = \{ (1,0), (0,1) \}$ as standard ordered bases for $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively. Then

$$\begin{aligned}
\mathbf{T}(1,0,0) &= (1,0) = 1 \cdot (1,0) + 0 \cdot (0,1) &\Rightarrow\quad [\mathbf{T}(1,0,0)]_\gamma = (1,0) \\
\mathbf{T}(0,1,0) &= (1,1) = 1 \cdot (1,0) + 1 \cdot (0,1) &\Rightarrow\quad [\mathbf{T}(0,1,0)]_\gamma = (1,1) \\
\mathbf{T}(0,0,1) &= (0,-1) = 0 \cdot (1,0) + (-1) \cdot (0,1) &\Rightarrow\quad [\mathbf{T}(0,0,1)]_\gamma = (0,-1).
\end{aligned}$$

Therefore, $A = [\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$.

Note that $(1,2,3) = (1,0,0) + 2(0,1,0) + 3(0,0,1)$, and that $\mathbf{T}(E_i) = column_i(A)$, for $i = 1,2,3$. Hence, we can compute $\mathbf{T}(1,2,3)$ as follows:

$$\mathbf{T}(1,2,3) = \mathbf{T}(E_1) + 2\mathbf{T}(E_2) + 3\mathbf{T}(E_3) = (3,-1).$$

On the other hand, we simply can use Remark 2.2.2 as follows:

$$[\mathbf{T}(1,2,3)]_\gamma = A \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Therefore, $\mathbf{T}(1,2,3) = (3,-1)$.

Example 2.2.3

Let $\mathbf{T} : \mathbb{P}_1(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ be a linear defined by $\mathbf{T}(f(x)) = x\,f(x)$. (1): Find the matrix representation A for $\mathbf{T}$. (2): If $f(x) = 3x - 2 \in \mathbb{P}_1(\mathbb{R})$, compute $[\mathbf{T}(f(x))]_\gamma$, where γ is the standard ordered basis in $\mathbb{P}_2(\mathbb{R})$. (3): Evaluate $\mathbf{T}(f(x))$ using A.

Solution:

①: We use the method described in Remark 2.2.3 and consider $\beta = \{1, x\}$ and $\gamma = \{1, x, x^2\}$ as standard ordered bases for $\mathbb{P}_1(\mathbb{R})$ and $\mathbb{P}_2(\mathbb{R})$, respectively. Then

$$\mathbf{T}(f(x) = 1) = x \cdot 1 = x = 0 \cdot 1 + 1 \cdot x + 0 \cdot x^2 \quad \Rightarrow \quad [\mathbf{T}(1)]_\gamma = (0, 1, 0)$$
$$\mathbf{T}(f(x) = x) = x \cdot x = x^2 = 0 \cdot 1 + 0 \cdot x + 1 \cdot x^2 \quad \Rightarrow \quad [\mathbf{T}(x)]_\gamma = (0, 0, 1).$$

Therefore, $A = [\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$.

②: We can simply compute $[\mathbf{T}(f(x))]_\gamma$ directly:

$$\mathbf{T}(f(x)) = xf(x) = x(3x-2) = 3x^2 - 2x = 0 \cdot 1 + (-2) \cdot x + 3 \cdot x^2 \ \Rightarrow \ [\mathbf{T}((f(x))]_\gamma = (0, -2, 3).$$

Or, we can use Remark 2.2.2 part(2) using A. We first write $f(x)$ as a linear combination of β:

$$f(x) = -2 \cdot 1 + 3 \cdot x \qquad \Rightarrow \qquad [f(x)]_\beta = (-2, 3).$$

Then using Remark 2.2.2 part(2), we have

$$[\mathbf{T}(f(x))]_\gamma = A\ [f(x)]_\beta = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix}.$$

Hence $[\mathbf{T}(f(x))]_\gamma = (0, -2, 3)$.

③: Use the result in part ②, to get $\mathbf{T}(f(x)) = -2x + 3x^2$.

Definition 2.2.4

Let $\mathbf{T}, \mathbf{U} : \mathbb{V} \to \mathbb{W}$ be arbitrary functions where $\mathbb{V}$ and $\mathbb{W}$ are vector spaces over $\mathbb{F}$, and let $a \in \mathbb{F}$. We define the usual addition of functions $\mathbf{T} + \mathbf{U} : \mathbb{V} \to \mathbb{W}$ by

$$(\mathbf{T} + \mathbf{U})(x) = \mathbf{T}(x) + \mathbf{U}(x) \qquad \text{for all } \ x \in \mathbb{V},$$

and $a\,\mathbf{T} : \mathbb{V} \to \mathbb{W}$ by

$$(a\,\mathbf{T})(x) = a\,\mathbf{T}(x) \qquad \text{for all } \ x \in \mathbb{V}.$$

Theorem 2.2.1

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$, and let $\mathbf{T}, \mathbf{U} : \mathbb{V} \to \mathbb{W}$ be linear transformations. Then

1. For all $a \in \mathbb{F}$, $(a\mathbf{T} + \mathbf{U})$ is linear transformation.
2. The collection of all linear transformations from $\mathbb{V}$ to $\mathbb{W}$ is a vector space over $\mathbb{F}$.

Proof:

(1) Let $x, y \in \mathbb{V}$ and $c \in \mathbb{F}$. Then

$$\begin{aligned}(a\mathbf{T} + \mathbf{U})(cx + y) &= a\mathbf{T}(cx + y) + \mathbf{U}(cx + y) = a\left[\mathbf{T}(cx + y)\right] + c\mathbf{U}(x) + \mathbf{U}(y) \\ &= a\left[c\mathbf{T}(x) + \mathbf{T}(y)\right] + c\mathbf{U}(x) + \mathbf{U}(y) \\ &= ac\mathbf{T}(x) + a\mathbf{T}(y) + c\mathbf{U}(x) + \mathbf{U}(y) \\ &= c\left(a\mathbf{T}(x) + \mathbf{U}(x)\right) + a\mathbf{T}(y) + \mathbf{U}(y) \\ &= c(a\mathbf{T} + \mathbf{U})(x) + (a\mathbf{T} + \mathbf{U})(y).\end{aligned}$$

Thus, $a\mathbf{T} + \mathbf{U}$ is linear transformation.

(2): Note that the zero transformation $\mathbf{T}_0$ is the zero vector. The other conditions of a vector space can be easily proved.

Definition 2.2.5

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$. We denote the vector space of all linear transformation from $\mathbb{V}$ into $\mathbb{W}$ by $\mathcal{L}(\mathbb{V}, \mathbb{W})$. If $\mathbb{V} = \mathbb{W}$, we simply write $\mathcal{L}(\mathbb{V})$.

Theorem 2.2.2

Let $\mathbb{V}$ and $\mathbb{W}$ be finite-dimensional vector spaces with ordered bases β and γ, respectively, and let $\mathbf{T}, \mathbf{U} : \mathbb{V} \to \mathbb{W}$ be linear transformations. Then

1. $[\mathbf{T} + \mathbf{U}]_\beta^\gamma = [\mathbf{T}]_\beta^\gamma + [\mathbf{U}]_\beta^\gamma$, and
2. $[a\mathbf{T}]_\beta^\gamma = a\,[\mathbf{T}]_\beta^\gamma$ for all scalars a.

Example 2.2.4

Define $\mathbf{T} : M_{2\times 2}(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ by $\mathbf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a+b) + 2d\,x + b\,x^2$.

Let $\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ and $\gamma = \{1, x, x^2\}$ be ordered bases for $M_{2\times 2}(\mathbb{R})$ and $\mathbb{P}_2(\mathbb{R})$, respectively. Find $[\mathbf{T}]_\beta^\gamma$. Use $[\mathbf{T}]_\beta^\gamma$ to evaluate $\mathbf{T}(D)$, where $D = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}$.

Solution:

We use the method described in Remark 2.2.3. That is,

$$\mathbf{T}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1 = 1\cdot 1 + 0\cdot x + 0\cdot x^2 \;\Rightarrow\; [\mathbf{T}]_\gamma = (1,0,0),$$

$$\mathbf{T}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1 + x^2 = 1\cdot 1 + 0\cdot x + 1\cdot x^2 \;\Rightarrow\; [\mathbf{T}]_\gamma = (1,0,1),$$

$$\mathbf{T}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = 0 = 0\cdot 1 + 0\cdot x + 0\cdot x^2 \;\Rightarrow\; [\mathbf{T}]_\gamma = (0,0,0),$$

$$\mathbf{T}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 2x = 0\cdot 1 + 2\cdot x + 0\cdot x^2 \;\Rightarrow\; [\mathbf{T}]_\gamma = (0,2,0).$$

Thus,

$$A = [\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Note that $[D]_\beta = (1,3,-1,2)$, and hence $[\mathbf{T}(D)]_\gamma = A\,[D]_\beta = (4,4,3)$. Hence, $\mathbf{T}(D) = 4 + 4x + 3x^2$.

Example 2.2.5

Let $\beta = \{x^4, x^3, x^2, x, 1\}$ be an ordered basis for $\mathbb{P}_4(\mathbb{R})$ and let γ be the standard ordered basis for $\mathbb{R}^3$. Define $\mathbf{T} : \mathbb{P}_4(\mathbb{R}) \to \mathbb{R}^3$ by $\mathbf{T}(f(x)) = (f(1) - f(0), f'(0), f''(1))$, and let $\mathbf{U} : \mathbb{P}_4(\mathbb{R}) \to \mathbb{R}^3$ be a linear transformation having the matrix representation

$$[\mathbf{U}]_\beta^\gamma = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 1 & 2 \\ 1 & -1 & 1 & 1 & 1 \end{bmatrix}.$$

1. Find $\mathbf{U}(x^4 - x^2 + 1)$.
2. Find the matrix representation of $\mathbf{T} + \mathbf{U}$; that is, $[\mathbf{T} + \mathbf{U}]_\beta^\gamma$.
3. Find the rank and the nullity of $\mathbf{U}$. Exercise!!

Solution:

(1): Let $f(x) = x^4 - x^2 + 1$. We compute $\left[\mathbf{U}\big(f(x)\big)\right]_\gamma = [\mathbf{U}]_\beta^\gamma \, [f(x)]_\beta$ using Remark 2.2.2:

$$f(x) = x^4 - x^2 + 1 = 1 \cdot x^4 + 0 \cdot x^3 + (-1) \cdot x^2 + 0 \cdot x + 1 \cdot 1 \;\Rightarrow\; [f(x)]_\beta = (1, 0, -1, 0, 1).$$

Thus

$$\left[\mathbf{U}\big(f(x)\big)\right]_\gamma = [\mathbf{U}]_\beta^\gamma \, [f(x)]_\beta = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 1 & 2 \\ 1 & -1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 1 \end{bmatrix} \in \mathbb{R}^3.$$

We note that, this can be computed directly as follows:

$$\begin{aligned} \left[\mathbf{U}\big(x^4 - x^2 + 1\big)\right]_\gamma &= \left[\mathbf{U}\big(x^4\big) - \mathbf{U}\big(x^2\big) + \mathbf{U}\big(1\big)\right]_\gamma = \left[\mathbf{U}\big(x^4\big)\right]_\gamma - \left[\mathbf{U}\big(x^2\big)\right]_\gamma + \left[\mathbf{U}\big(1\big)\right]_\gamma \\ &= col_1([\mathbf{U}]_\beta^\gamma) - col_3([\mathbf{U}]_\beta^\gamma) + col_5([\mathbf{U}]_\beta^\gamma) = (1, 3, 1). \end{aligned}$$

Therefore, $\mathbf{U}\big(x^4 - x^2 + 1\big) = (1, 3, 1)$.

(2): Note that $[\mathbf{T} + \mathbf{U}]_\beta^\gamma = [\mathbf{T}]_\beta^\gamma + [\mathbf{U}]_\beta^\gamma$. Thus,

$$\begin{aligned} \mathbf{T}\left(x^4\right) &= (1, 0, 12) = 1 \cdot (1, 0, 0) + 0 \cdot (0, 1, 0) + 12 \cdot (0, 0, 1) \;\Rightarrow\; \left[\mathbf{T}(x^4)\right]_\gamma = (1, 0, 12) \\ \mathbf{T}\left(x^3\right) &= (1, 0, 6) = 1 \cdot (1, 0, 0) + 0 \cdot (0, 1, 0) + 6 \cdot (0, 0, 1) \;\Rightarrow\; \left[\mathbf{T}(x^3)\right]_\gamma = (1, 0, 6) \\ \mathbf{T}\left(x^2\right) &= (1, 0, 2) = 1 \cdot (1, 0, 0) + 0 \cdot (0, 1, 0) + 2 \cdot (0, 0, 1) \;\Rightarrow\; \left[\mathbf{T}(x^2)\right]_\gamma = (1, 0, 2) \\ \mathbf{T}\left(x\right) &= (1, 1, 0) = 1 \cdot (1, 0, 0) + 1 \cdot (0, 1, 0) + 0 \cdot (0, 0, 1) \;\Rightarrow\; [\mathbf{T}(x)]_\gamma = (1, 1, 0) \\ \mathbf{T}\left(1\right) &= (0, 0, 0) = 0 \cdot (1, 0, 0) + 0 \cdot (0, 1, 0) + 0 \cdot (0, 0, 1) \;\Rightarrow\; [\mathbf{T}(1)]_\gamma = (0, 0, 0) \end{aligned}$$

Hence, $[\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 12 & 6 & 2 & 0 & 0 \end{bmatrix}$ and therefore

$$[\mathbf{T} + \mathbf{U}]_\beta^\gamma = [\mathbf{T}]_\beta^\gamma + [\mathbf{U}]_\beta^\gamma = \begin{bmatrix} 2 & 1 & 2 & 1 & 1 \\ 0 & 1 & -1 & 2 & 2 \\ 13 & 5 & 3 & 1 & 1 \end{bmatrix}.$$

(3): $\mathcal{R}(\mathbf{U}) = \textbf{span}\ \{\mathbf{U}(x^4), \mathbf{U}(x^3), \mathbf{U}(x^2), \mathbf{U}(x), \mathbf{U}(1)\}$, and hence $rank(\mathbf{U}) = rank\Big([\mathbf{U}]_\beta^\gamma\Big)$ and $nullity(\mathbf{U}) = nullity\Big([\mathbf{U}]_\beta^\gamma\Big)$

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 1 & 2 \\ 1 & -1 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{\text{r.r.e.f.}} \begin{bmatrix} 1 & 0 & 0 & 2 & 3 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -2 & -2 \end{bmatrix}$$

That is $rank(\mathbf{U}) = 3$ and hence $nullity(\mathbf{U}) = 2$.

Also note that $\{(1,0,1), (0,1,-1), (1,-1,1)\}$ is a basis for $\mathcal{R}(\mathbf{U})$. Also, $\mathbf{U}$ is not 1-1 since its nullity $\neq 0$.

Example 2.2.6

Let $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$, defined by $\mathbf{T}(f(x)) = \begin{pmatrix} f'(0) & 2f(1) \\ 0 & f''(3) \end{pmatrix}$.

1. Compute $[\mathbf{T}]_\beta^\gamma$ where β and γ are the standard ordered bases for $\mathbb{P}_2(\mathbb{R})$ and $M_{2\times 2}(\mathbb{R})$, respectively. Use $[\mathbf{T}]_\beta^\gamma$ to compute $\mathbf{T}(g(x))$, where $g(x) = x^2 + 2x$.
2. Is $\mathbf{T}$ 1-1? Explain. Exercise!!
3. Is $\mathbf{T}$ onto? Explain. Exercise!!

Solution:

(1): Let $\beta = \{1, x, x^2\}$ and $\gamma = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$. Then

$$\mathbf{T}(1) = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
$$\mathbf{T}(x) = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
$$\mathbf{T}(x^2) = \begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus

$$A = [\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Note that $[g(x)]_\beta = (0,2,1)$, and hence $[\mathbf{T}(g(x))]_\beta^\gamma = A[g(x)]_\beta = (2,6,0,2)$. Therefore,

$$\mathbf{T}(g(x)) = \begin{pmatrix} 2 & 6 \\ 0 & 2 \end{pmatrix}.$$

(2): Note that $rank(\mathbf{T}) = rank([\mathbf{T}]_\beta^\gamma)$. Thus

$$\begin{bmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \xrightarrow{\text{r.r.e.f.}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore, $rank(\mathbf{T}) = 3$ and hence $nullity(\mathbf{T}) = 0$. Thus, $\mathbf{T}$ is 1-1.

(3): $\mathbf{T}$ is not onto since $rank(\mathbf{T}) = 3 < rank(M_{2\times 2}(\mathbb{R})) = 4$.

Example 2.2.7

Let $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$ be a linear transformation satisfying:

$$\mathbf{T}(1) = (1,1,1),\ \mathbf{T}(1+x) = (1,2,1), \quad \text{and} \quad \mathbf{T}(1+x+x^2) = (1,0,1).$$

1. Find a matrix representation of $\mathbf{T}$ relative to the standard ordered bases for $\mathbb{P}_2(\mathbb{R})$ and $\mathbb{R}^3$. Evaluate $\mathbf{T}(g(x))$, where $g(x) = x^2 - 3x + 1$.
2. Find bases for $\mathcal{R}(\mathbf{T})$ and $\mathcal{N}(\mathbf{T})$. Exercise!!

Solution:

(1): Consider $\beta = \{1, x, x^2\}$ and $\gamma = \{(1,0,0),(0,1,0),(0,0,1)\}$. Then

$$\mathbf{T}(1) = (1,1,1) = 1\cdot(1,0,0) + 1\cdot(0,1,0) + 1\cdot(0,0,1) \Rightarrow [\mathbf{T}(1)]_\gamma = (1,1,1)$$
$$\mathbf{T}(x) = \mathbf{T}(1+x) - \mathbf{T}(1) = (0,1,0) \Rightarrow [\mathbf{T}(x)]_\gamma = (0,1,0)$$
$$\mathbf{T}(x^2) = \mathbf{T}(1+x+x^2) - \mathbf{T}(1+x) = (0,-2,0) \Rightarrow \left[\mathbf{T}(x^2)\right]_\gamma = (0,-2,0).$$

Thus

$$A = [\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & -2 \\ 1 & 0 & 0 \end{bmatrix}.$$

Note that, $[g(x)]_\beta = (1,-3,1)$ and hence $[\mathbf{T}(g(x))]_\gamma = A[g(x)]_\beta = (1,-4,1)$. Therefore, $\mathbf{T}(g(x)) = (1,-4,1)$.

(2): Note that

$$\begin{aligned}\mathcal{R}(\mathbf{T}) &= \mathbf{span}\,\Big\{\,\mathbf{T}(1), \mathbf{T}(x), \mathbf{T}(x^2)\,\Big\}\\ &= \mathbf{span}\,\{\,(1,1,1),(0,1,0),(0,-2,0)\,\} = \mathbf{span}\,\{\,(1,1,1),(0,1,0)\,\}.\end{aligned}$$

Therefore, $\{\,(1,1,1),(0,1,0)\,\}$ is a basis for $\mathcal{R}(\mathbf{T})$.

$$\begin{aligned}\mathcal{N}(\mathbf{T}) &= \Big\{\,f(x) = a+bx+cx^2 \in \mathbb{P}_2(\mathbb{R}) \,:\, \mathbf{T}(a+bx+cx^2) = (0,0,0)\,\Big\}\\ &= \Big\{\,a+bx+cx^2 \,:\, a\mathbf{T}(1)+b\mathbf{T}(x)+c\mathbf{T}(x^2) = (0,0,0)\,\Big\}\\ &= \Big\{\,a+bx+cx^2 \,:\, a(1,1,1)+b(0,1,0)+c(0,-2,0) = (0,0,0)\,\Big\}\\ &= \Big\{\,a+bx+cx^2 \,:\, (a, a+b-2c, a) = (0,0,0)\,\Big\}\\ &= \Big\{\,a+bx+cx^2 \,:\, a=0 \text{ and } b=2c\Big\}\\ &= \Big\{\,2cx+cx^2 \,:\, c\in\mathbb{R}\,\Big\} = \mathbf{span}\,\Big\{\,2x+x^2\,\Big\}.\end{aligned}$$

Thus, $\{\,2x+x^2\,\}$ is a basis for $\mathcal{N}(\mathbf{T})$.

Note that we could use Remark 2.2.3 to find a basis for $\mathcal{N}(\mathbf{T})$ using the following technique:

$$\Big[\mathbf{T}(a+bx+cx^2)\Big]_\beta = A\,\Big[\mathbf{T}(a+bx+cx^2)\Big]_\beta = \begin{bmatrix}1&0&0\\1&1&-2\\1&0&0\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}a\\a+b-2c\\a\end{bmatrix}.$$

Therefore,

$$\begin{bmatrix}a\\a+b-2c\\a\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

implies that $a=0$ and $b=2c$. Hence, $f(x) = 0+2cx+cx^2$ and thus $\{\,2x+x^2\,\}$ is a basis for $\mathcal{N}(\mathbf{T})$.

Exercise 2.2.1

Solve the following exercises from the book at pages 84 - 86:

- $2 : a, b, c,$ and d.
- $3, 4, 5$.
- 8.

Exercise 2.2.2

Let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear defined by $\mathbf{T}(x, y) = (2x - 3y, -x, x + 4y)$. Find a matrix representation A for $\mathbf{T}$. Use A to evaluate $\mathbf{T}(u)$, where $u = (2, 4)$.

Solution:

We use the method described in Remark 2.2.3 and consider $\beta = \{ (1, 0), (0, 1) \}$ and $\gamma = \{ (1, 0, 0), (0, 1, 0), (0, 0, 1) \}$ as standard ordered bases for $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Then

$$\begin{aligned} \mathbf{T}(1, 0) = (2, -1, 1) = 2 \cdot (1, 0, 0) + (-1) \cdot (0, 1, 0) + 1 \cdot (0, 0, 1) \quad &\Rightarrow [\mathbf{T}(1, 0)]_\gamma = (2, -1, 1) \\ \mathbf{T}(0, 1) = (-3, 0, 4) = -3 \cdot (1, 0, 0) + 0 \cdot (0, 1, 0) + 4 \cdot (0, 0, 1) \quad &\Rightarrow [\mathbf{T}(0, 1)]_\gamma = (-3, 0, 4). \end{aligned}$$

Therefore, $A = [\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 2 & -3 \\ -1 & 0 \\ 1 & 4 \end{bmatrix}$.

Simply, $[\mathbf{T}(u)]_\gamma = A\,[u]_\beta = (-8, -2, 18)$. Hence, $\mathbf{T}(u) = (-8, -2, 18)$.

Exercise 2.2.3

Let $\mathbf{T} : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ be the linear defined by $\mathbf{T}(f(x)) = f'(x)$. Let β and γ be the standard ordered bases for $\mathbb{P}_3(\mathbb{R})$ and $\mathbb{P}_2(\mathbb{R})$, respectively. Find the matrix representation A for $\mathbf{T}$ with respect to β and γ. Use A to evaluate $\mathbf{T}(f(x))$, where $f(x) = 3x^2 + 1$.

Solution:

Let $\beta = \{1, x, x^2, x^3\}$ and $\gamma = \{1, x, x^2\}$. Thus

$$\begin{aligned}
\mathbf{T}(1) &= 0 = 0\cdot 1 + 0\cdot x + 0\cdot x^2 \Rightarrow [\mathbf{T}(1)]_\gamma = (0,0,0)\\
\mathbf{T}(x) &= 1 = 1\cdot 1 + 0\cdot x + 0\cdot x^2 \Rightarrow [\mathbf{T}(x)]_\gamma = (1,0,0)\\
\mathbf{T}(x^2) &= 2x = 0\cdot 1 + 2\cdot x + 0\cdot x^2 \Rightarrow \left[\mathbf{T}(x^2)\right]_\gamma = (0,2,0)\\
\mathbf{T}(x^3) &= 3x^2 = 0\cdot 1 + 0\cdot x + 3\cdot x^2 \Rightarrow \left[\mathbf{T}(x^3)\right]_\gamma = (0,0,3).
\end{aligned}$$

Therefore, $A = [\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

Note that $[f(x)]_\beta = (1,0,3,0)$ and hence $[\mathbf{T}(f(x))]_\gamma = A\,[f(x)]_\beta = (0,6,0)$. Therefore, $\mathbf{T}(f(x)) = 6x$.

Exercise 2.2.4

Let $A = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 1 & 3 \end{bmatrix}$. Assume that $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_1(\mathbb{R})$ is the linear tranformation defined by A using the standard ordered bases β and γ for $\mathbb{P}_2(\mathbb{R})$ and $\mathbb{P}_1(\mathbb{R})$, respectively. Evaluate $\mathbf{T}(g(x))$, where $g(x) = 2x^2 - 3x + 1$.

Solution:

We solve in two methods: 1. Note that $[\mathbf{T}(2x^2 - 3x + 1)]_\gamma = 2\,[\mathbf{T}(x^2)]_\gamma - 3\,[\mathbf{T}(x)]_\gamma + [\mathbf{T}(1)]_\gamma = 2\begin{bmatrix} 2 \\ 3 \end{bmatrix} - 3\begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$. Hence, $\mathbf{T}(g(x)) = 8 + 7x$.

2. Another way: Note that $[g(x)]_\beta = (1,-3,2)$ and hence $[\mathbf{T}(g(x))]_\gamma = A\,[g(x)]_\beta = (8,7)$. Thus, $\mathbf{T}(g(x)) = 8 + 7x$.

Section 2.5: The Change of Coordinate Matrix

Definition 2.5.1

Let β and γ be ordered bases for a finite-dimensional vector space $\mathbb{V}$, and let $Q = [\mathbf{I}_{\mathbb{V}}]_{\gamma}^{\beta}$, where $\mathbf{I}_{\mathbb{V}} : \mathbb{V} \to \mathbb{V}$ is the identity linear transformation. Then Q is called the **change of coordinate matrix** (it changes γ-coordinate into β-coordinate). Moreover, Q is invertible and Q^{-1} changes β-coordinate into γ-coordinate.

Theorem 2.5.1

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$. Let β and γ be two ordered bases for $\mathbb{V}$, and let Q be the change of coordinate matrix that changes γ-coordinates into β-coordinates. Then

1. For any $x \in \mathbb{V}$, $[x]_\beta = Q\,[x]_\gamma$, and
2. $[\mathbf{T}]_\gamma = Q^{-1}\,[\mathbf{T}]_\beta\,Q$.

Example 2.5.1

Let $\beta = \{(1,0),(0,1)\}$ and $\gamma = \{(1,-1),(2,1)\}$ be two ordered bases for $\mathbb{R}^2$, and let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\mathbf{T}(a,b) = (a+b, a-2b)$. Find the change of coordinate matrix Q, that changes γ-coordinates into β-coordinates, and use it to find $[\mathbf{T}]_\gamma$. Find $[(5,1)]_\beta$ using Q.

Solution:

Note that

$$\mathbf{I}_{\mathbb{R}^2}(1,-1) = (1,-1) = 1\cdot(1,0) + (-1)\cdot(0,1) \quad \& \quad \mathbf{I}_{\mathbb{R}^2}(2,1) = (2,1) = 2\cdot(1,0) + 1\cdot(0,1).$$

Thus, the matrix that changes γ-coordinates into β-coordinates is

$$Q = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \;\Rightarrow\; Q^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}.$$

To find $[\mathbf{T}]_\gamma$, we use $[\mathbf{T}]_\gamma = Q^{-1}\,[\mathbf{T}]_\beta\,Q$ and

$$\left.\begin{array}{l} \mathbf{T}(1,0) = (1,1) = 1\cdot(1,0) + 1\cdot(0,1) \\ \mathbf{T}(0,1) = (1,-2) = 1\cdot(1,0) + (-2)\cdot(0,1) \end{array}\right\} \Rightarrow [\mathbf{T}]_\beta = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}.$$

Thus, $[\mathbf{T}]_\gamma = Q^{-1}[\mathbf{T}]_\beta Q = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}$.

⋆ Confirmation:

$$\mathbf{T}(1,-1) = (0,3) = -2\cdot(1,-1) + 1\cdot(2,1), \text{ and}$$
$$\mathbf{T}(2,1) = (3,0) = 1\cdot(1,-1) + 1\cdot(2,1).$$

Finally, note that $[(5,1)]_\beta = Q\,[(5,1)]_\gamma$, where $[(5,1)]_\gamma = (1,2)$ since $(5,1) = 1\cdot(1,-1) + 2\cdot(2,1)$. Therefore, $[(5,1)]_\beta = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$ which is true since $(5,1) = 5\cdot(1,0) + 1\cdot(0,1)$.

Example 2.5.2

Let $\beta = \{(1,1),(1,-1)\}$ and $\gamma = \{(2,4),(3,1)\}$ be bases for $\mathbb{R}^2$. (a) What is the matrix Q that changes γ-coordinates into β-coordinates, and use it to find $[(1,7)]_\beta$ and $[(1,7)]_\gamma$. (b) If $\mathbf{T}:\mathbb{R}^2\to\mathbb{R}^2$ is the linear operator on $\mathbb{R}^2$ defined by $\mathbf{T}(a,b) = (3a-b, a+3b)$, find $[\mathbf{T}]_\gamma$.

Solution:

(a): We first consider:

$$\mathbf{I}_{\mathbb{R}^2}(2,4) = (2,4) = c_1(1,1) + c_2(1,-1) = 3(1,1) + (-1)(1,-1), \quad \text{and}$$
$$\mathbf{I}_{\mathbb{R}^2}(3,1) = (3,1) = c_1(1,1) + c_2(1,-1) = 2(1,1) + 1(1,-1).$$

Thus, the matrix that changes γ-coordinates into β-coordinates is

$$Q = \begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix} \quad\Rightarrow\quad Q^{-1} = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}.$$

To compute $[(1,7)]_\beta$, consider $(1,7) = 2(2,4) + (-1)(3,1)$; hence $[(1,7)]_\gamma = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Therefore,

$$[(1,7)]_\beta = Q\,[(1,7)]_\gamma = \begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \end{bmatrix},$$

which is true since $(1,7) = 4(1,1) + (-3)(1,-1)$.

To compute $[(1,7)]_\gamma$, consider $(1,7) = 4(1,1) + (-3)(1,-1)$; hence $[(1,7)]_\beta = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$. Therefore,

$$[(1,7)]_\gamma = Q^{-1}\,[(1,7)]_\beta = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}\begin{bmatrix} 4 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

which is true since $(1,7) = 2(2,4) + (-1)(3,1)$.

ⓑ: Note that

$$\mathbf{T}(1,1) = (2,4) = 3\cdot(1,1) + (-1)\cdot(1,-1),$$
$$\mathbf{T}(1,-1) = (4,-2) = 1\cdot(1,1) + 3\cdot(1,-1).$$

Thus $[\mathbf{T}]_\beta = \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}$ and hence

$$[\mathbf{T}]_\gamma = Q^{-1}\ [\mathbf{T}]_\beta\ Q = \cdots = \begin{pmatrix} 4 & 1 \\ -2 & 2 \end{pmatrix}.$$

Which can be seen if we consider

$$\mathbf{T}(2,4) = (2,14) = \textcircled{4}\cdot(2,4) + \textcircled{-2}\cdot(3,1) \ \Rightarrow\ [\mathbf{T}(2,4)]_\gamma = (4,-2). \text{ "}1^{st}\text{ column of } [\mathbf{T}]_\gamma\text{"}$$
$$\mathbf{T}(3,1) = (8,6) = \textcircled{1}\cdot(2,4) + \textcircled{2}\cdot(3,1) \ \Rightarrow\ [\mathbf{T}(3,1)]_\gamma = (1,2). \text{ "}2^{nd}\text{ column of } [\mathbf{T}]_\gamma\text{"}$$

Example 2.5.3

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^3$ defined by

$$\mathbf{T}(a,b,c) = (2a+b, a+b+3c, -b),$$

and let $\beta = \{(1,0,0),(0,1,0),(0,0,1)\}$ and $\gamma = \{(-1,0,0),(2,1,0),(1,1,1)\}$ be bases for $\mathbb{R}^3$. Find $[\mathbf{T}]_\beta$, $[\mathbf{T}]_\gamma$, and the matrix Q that changes the γ-coordinates into β-coordinates.

Solution:

Clearly,

$$\mathbf{I}_{\mathbb{R}^3}(-1,0,0) = (-1,0,0) = -1(1,0,0) + 0(0,1,0) + 0(0,0,1)$$
$$\mathbf{I}_{\mathbb{R}^3}(2,1,0) = (2,1,0) = 2(1,0,0) + 1(0,1,0) + 0(0,0,1)$$
$$\mathbf{I}_{\mathbb{R}^3}(1,1,1) = (1,1,1) = 1(1,0,0) + 1(0,1,0) + 1(0,0,1).$$

Hence

$$Q = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow Q^{-1} = \begin{pmatrix} -1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Computing $[\mathbf{T}]_\beta$:

$$\mathbf{T}(1,0,0) = (2,1,0) = 2(1,0,0) + 1(0,1,0) + 0(0,0,1)$$
$$\mathbf{T}(0,1,0) = (1,1,-1) = 1(1,0,0) + 1(0,1,0) + (-1)(0,0,1)$$
$$\mathbf{T}(0,0,1) = (0,3,0) = 0(1,0,0) + 3(0,1,0) + 0(0,0,1).$$

Thus $[\mathbf{T}]_\beta = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & -1 & 0 \end{pmatrix}$, and hence $[\mathbf{T}]_\gamma = Q^{-1}[\mathbf{T}]_\beta Q = \begin{pmatrix} 0 & 2 & 8 \\ -1 & 4 & 6 \\ 0 & -1 & -1 \end{pmatrix}$.

Confirming:

$$\mathbf{T}(-1,0,0) = (-2,-1,0) = 0(-1,0,0) + (-1)(2,1,0) + 0(1,1,1)$$
$$\mathbf{T}(2,1,0) = (5,3,-1) = 2(-1,0,0) + 4(2,1,0) + (-1)(1,1,1)$$
$$\mathbf{T}(1,1,1) = (3,5,-1) = 8(-1,0,0) + 6(2,1,0) + (-1)(1,1,1).$$

Exercise 2.5.1

Solve the following exercises from the book at pages 116 - 117:

- $2, 3, 4, 5, 6.$

Chapter

5 Diagonalization

Section 5.1: Eigenvalues and Eigenvectors

Definition 5.1.1

Let $A \in M_{m\times n}(\mathbb{F})$. We define the mapping $\mathbf{L}_A : \mathbb{F}^n \to \mathbb{F}^m$ by $\mathbf{L}_A(x) = Ax$ for every column vector $x \in \mathbb{F}^n$. We call $\mathbf{L}_A$, the **left multiplication transformation**.

Example 5.1.1

Let $A = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 2 \end{bmatrix}$ and $\mathbf{L}_A : \mathbb{R}^3 \to \mathbb{R}^2$. Find $\mathbf{L}_A(x)$ where $x = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$.

Solution:

$$\mathbf{L}_A(x) = Ax = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 3 \end{bmatrix} \in \mathbb{R}^2.$$

Remark 5.1.1

Let $A, B \in M_{m\times n}(\mathbb{F})$ and let $c \in \mathbb{F}$. Then

1. $\mathbf{L}_A$ is a linear transformation.
2. $[\mathbf{L}_A]_\beta^\gamma = A$, where β and γ are the standard ordered bases for $\mathbb{F}^n$ and $\mathbb{F}^m$, respectively.
3. $\mathbf{L}_A = \mathbf{L}_B$ if and only if $A = B$.
4. $\mathbf{L}_{A+B} = \mathbf{L}_A + \mathbf{L}_B$ and $\mathbf{L}_{cA} = c\mathbf{L}_A$.

Proof of (2): Let $\beta = (E_1, \cdots, E_n)$ and $\gamma = (E_1, \cdots, E_m)$ be the standard ordered bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. For any column vector $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, we have

$$x = x_1E_1 + \cdots + x_nE_n,$$

and thus $[x]_\beta = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x$. Similarly, we have $[y]_\gamma = y$ for all $y \in \mathbb{R}^m$.

Now let $A \in M_{m\times n}(\mathbb{R})$, and let $x \in \mathbb{R}^n$. By definition, $\mathbf{L}_A(x) = Ax$. Also by Remark 2.2.2, we have $[\mathbf{L}_A]_\gamma = [\mathbf{L}_A]_\beta^\gamma [x]_\beta$. Note that since $[\mathbf{L}_A]_\gamma \in \mathbb{R}^m$ and $[x]_\beta \in \mathbb{R}^n$, we have

$$\mathbf{L}_A(x) = [\mathbf{L}_A]_\beta^\gamma \; x.$$

Thus, $[\mathbf{L}_A]_\beta^\gamma \; x = \mathbf{L}_A(x) = Ax$ for all $x \in \mathbb{R}^n$. Applying this to $E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, we see that the first column of $[\mathbf{L}_A]_\beta^\gamma$ and A are the same. Similarly, we apply it for all E_i for $i = 1, \cdots, n$, we get $[\mathbf{L}_A]_\beta^\gamma = A$ as desired.

Definition 5.1.2

A linear operator $\mathbf{T}$ on a finite-dimensional vector space $\mathbb{V}$ is called **diagonalizable** if there is an ordered basis β for $\mathbb{V}$ such that $[\mathbf{T}]_\beta$ is a diagonal matrix. A square matrix A is called diagonalizable if $\mathbf{L}_A$ is diagonalizable.

Definition 5.1.3

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$. A nonzero vector $x \in \mathbb{V}$ is called **eigenvector** (or **e-vector** for short) of $\mathbf{T}$ if there exists a scalar λ such that $\mathbf{T}(x) = \lambda \; x$. The scalar λ is then called **eigenvalue** (or **e-value** for short) corresponding to x.

Remark 5.1.2

Let $A \in M_{n\times n}(\mathbb{F})$.

- A nonzero vector $x \in \mathbb{F}^n$ is called e-vector of A if and only if x is an e-vector of $\mathbf{L}_A$.
- λ is an e-value of A if and only if λ is an e-value of $\mathbf{L}_A$.

Theorem 5.1.1

A linear operator $\mathbf{T}$ on a finite-dimensional vector space $\mathbb{V}$ is diagonalizable if and only if there exists an ordered basis β for $\mathbb{V}$ consisting of e-vectors of $\mathbf{T}$. Furthermore, if $\mathbf{T}$ is diagonalizable, $\beta = \{ x_1, x_2, \cdots, x_n \}$ is an ordered basis of e-vectors of $\mathbf{T}$, and $D = [\mathbf{T}]_\beta = (d_{ij})$, then D is a diagonal matrix and d_{jj} is the e-values corresponding to x_j for $1 \le j \le n$.

Note that to *diagonalize* a matrix or a linear operator is to find a basis of e-vectors and the corresponding e-values.

Example 5.1.2

Consider $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$, $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Then

$$\mathbf{L}_A(x) = Ax = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2x \qquad \text{and} \qquad \mathbf{L}_A(y) = Ay = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 3y.$$

That is 2 and 3 are e-values of $\mathbf{L}_A$ corresponding to e-vectors x and y, respectively.
Note that $\beta = \{x, y\}$ is an ordered basis for $\mathbb{R}^2$ consisting e-vectors of both A and $\mathbf{L}_A$, and therefore A and $\mathbf{L}_A$ are both diagonalizable. Moreover,

$$[\mathbf{L}_A]_\beta = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix},$$

where $[\mathbf{L}_A(x)]_\beta = (2, 0)$, and $[\mathbf{L}_A(y)]_\beta = (0, 3)$.

Theorem 5.1.2

Let $A \in M_{n\times n}(\mathbb{F})$. Then a scalar λ is an e-value of A if and only if $|\, A - \lambda I_n \,| = 0$.

Proof:

A scalar λ is an e-value of A iff there exists a nonzero vector $x \in \mathbb{F}^n$ such that

$$Ax = \lambda x \Leftrightarrow Ax - \lambda x = 0 \Leftrightarrow (A - \lambda I_n)x = 0 \Leftrightarrow A - \lambda I_n \text{ is singular} \Leftrightarrow |\, A - \lambda I_n \,| = 0.$$

Definition 5.1.4

- Let $A \in M_{n\times n}(\mathbb{F})$. The polynomial $f(t) = |\, A - tI_n \,|$ is called the **characteristic polynomial** of A.
- Let $\mathbf{T}$ be a linear operator on an n-dimensional vector space $\mathbb{V}$ with ordered basis β. We

define the characteristic polynomial $f(t)$ of $\mathbf{T}$ to be

$$f(t) = |\, A - tI_n \,|\,, \quad \text{where} \quad A = [\mathbf{T}]_\beta\,.$$

Example 5.1.3

Find the e-values of $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$.

Solution

We use the characteristic polynomial $f(\lambda) = |\, A - \lambda I_2 \,| = 0$.

$$\begin{vmatrix} 1-\lambda & 1 \\ 4 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 4 = \lambda^2 - 2\lambda - 3 = (\lambda-3)(\lambda+1) = 0.$$

Therefore, $\lambda = -1$ and 3 are the e-values of A.

Example 5.1.4

Let $\mathbf{T}$ be a linear operator on $\mathbb{P}_2(\mathbb{R})$ defined by

$$\mathbf{T}\,(f(x)) = f(x) + (x+1)f'(x).$$

Find the e-values of $\mathbf{T}$.

Solution:

Let $A = [\mathbf{T}]_\beta$ where $\beta = \{\,1, x, x^2\,\}$ is the standard ordered basis for $\mathbb{P}_2(\mathbb{R})$. Then

$$\begin{aligned} \mathbf{T}(1) &= 1 = 1\cdot 1 + 0\cdot x + 0\cdot x^2 \\ \mathbf{T}(x) &= x + (x+1) = 2x + 1 = 1\cdot 1 + 2\cdot x + 0\cdot x^2 \\ \mathbf{T}(x^2) &= x^2 + (x+1)2x = 3x^2 + 2x = 0\cdot 1 + 2\cdot x + 3\cdot x^2. \end{aligned}$$

Thus, $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$, and hence

$$f(\lambda) = |\, A - \lambda I_3 \,| = \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 3-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)(3-\lambda) = 0.$$

Therefore, λ is an e-value of A iff $\lambda = 1, 2,$ or 3.

Note that if A is an $n \times n$ matrix, then $f(t) = |\, A - tI_n \,| = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$, is of degree n.

Theorem 5.1.3

Let $A \in M_{n\times n}(\mathbb{F})$ with characteristic polynomial $f(t)$. Then

1. $f(t)$ is a polynomial of degree n with leading coefficient $(-1)^n$.
2. A has at most n distinct e-values.
3. $f(0) = a_0 = |\, A \,|$.

The following theorem describes a procedure for computing the e-vectors corresponding to a given e-value.

Theorem 5.1.4

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$, and let λ be an e-value of $\mathbf{T}$. A vector $x \in \mathbb{V}$ is an e-vector of $\mathbf{T}$ corresponding to λ if and only if $x \neq 0$ and $x \in \mathcal{N}(\mathbf{T} - \lambda I)$.

Example 5.1.5

Let $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$. Find all e-vectors of A.

Solution:

We start finding the e-values using $f(\lambda) = |\, A - \lambda I_2 \,| = 0$. Thus

$$|\, A - \lambda I_2 \,| = \begin{vmatrix} 1-\lambda & 1 \\ 4 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 4 = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1) = 0.$$

Thus $\lambda_1 = -1$ and $\lambda_2 = 3$.

For $\underline{\lambda_1 = -1}$: Let $B_1 = A - \lambda_1 I_2 = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$. Then $x_1 = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ is an e-vector corresponding to $\lambda_1 = -1$ iff $x_1 \neq 0$ and $x_1 \in \mathcal{N}(\mathbf{L}_{B_1})$. That is

$$\mathbf{L}_{B_1}(x_1) = B_1 x_1 = 0 \;\Rightarrow\; \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left[\begin{array}{cc|c} 2 & 1 & 0 \\ 4 & 2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right] \;\Rightarrow\; a + \frac{1}{2}b = 0 \;\Rightarrow\; b = -2a.$$

That is, $x_1 \in \mathcal{N}(\mathbf{L}_{B_1}) = \left\{ t \begin{pmatrix} 1 \\ -2 \end{pmatrix} : 0 \neq t \in \mathbb{R} \right\}$. Thus x_1 is an e-vector of A corresponding to $\lambda_1 = -1$ iff $x_1 = t \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ for some nonzero $t \in \mathbb{R}$.

For $\underline{\lambda_2 = 3}$: Let $B_2 = A - \lambda_2 I_2 = \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix}$. Then $x_2 = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ is an e-vector corresponding to $\lambda_2 = 3$ iff $x_2 \neq 0$ and $x_2 \in \mathcal{N}(\mathbf{L}_{B_2})$. That is

$$\mathbf{L}_{B_2}(x_2) = B_2 x_2 = 0 \;\Rightarrow\; \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left[\begin{array}{cc|c} -2 & 1 & 0 \\ 4 & -2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right] \;\Rightarrow\; a - \frac{1}{2}b = 0 \;\Rightarrow\; b = 2a.$$

That is, $x_2 \in \mathcal{N}(\mathbf{L}_{B_2}) = \left\{ t \begin{pmatrix} 1 \\ 2 \end{pmatrix} : 0 \neq t \in \mathbb{R} \right\}$. Thus x_2 is an e-vector of A corresponding to $\lambda_2 = 3$ iff $x_2 = t \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ for some nonzero $t \in \mathbb{R}$.

Remark:

Note that $\gamma = \left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$ is an ordered basis for $\mathbb{R}^2$ containing e-vectors of A. Thus $\mathbf{L}_A$, and hence A, is diagonalizable and if $Q = \begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix}$, then $Q^{-1}AQ = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$.

Remark 5.1.3

Note that to find the e-vectors of a linear operator $\mathbf{T}$ on an n-dimensional vector space $\mathbb{V}$:

1. Select an ordered basis for $\mathbb{V}$, say β.
2. Let $A = [\mathbf{T}]_\beta$. Then $x \in \mathbb{V}$ is an e-vector of $\mathbf{T}$ corresponding to λ if and only if $[x]_\beta$, the coordinate vector of x relative to β, is an e-vector of A corresponding to λ.

Example 5.1.6

Let $\mathbf{T}$ be the linear operator defined on $\mathbb{P}_2(\mathbb{R})$ by

$$\mathbf{T}\,(f(x)) = f(x) + (x+1)f'(x).$$

Find the e-vectors of $\mathbf{T}$ and an ordered basis γ for $\mathbb{P}_2(\mathbb{R})$ so that $[\mathbf{T}]_\gamma$ is diagonalizable.

Solution:

Let $\beta = \{\,1, x, x^2\,\}$ be an ordered basis for $\mathbb{P}_2(\mathbb{R})$. Then

$$\begin{aligned}\mathbf{T}(1) &= 1 = 1\cdot 1 + 0\cdot x + 0\cdot x^2\\ \mathbf{T}(x) &= x + (x+1) = 2x+1 = 1\cdot 1 + 2\cdot x + 0\cdot x^2\\ \mathbf{T}(x^2) &= x^2 + (x+1)2x = 3x^2 + 2x = 0\cdot 1 + 2\cdot x + 3\cdot x^2.\end{aligned}$$

Thus, $A = [\mathbf{T}]_\beta = \begin{bmatrix} 1 & 1 & 0\\ 0 & 2 & 2\\ 0 & 0 & 3\end{bmatrix}$, and hence

$$f(\lambda) = |\,A - \lambda I_3\,| = \begin{vmatrix} 1-\lambda & 1 & 0\\ 0 & 2-\lambda & 2\\ 0 & 0 & 3-\lambda\end{vmatrix} = (1-\lambda)(2-\lambda)(3-\lambda) = 0.$$

Therefore, $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$.

For $\underline{\lambda_1 = 1}$: Let $B_1 = A - \lambda_1 I_3 = \begin{pmatrix} 0 & 1 & 0\\ 0 & 1 & 2\\ 0 & 0 & 2\end{pmatrix}$. Then $x_1 = \begin{pmatrix} a\\ b\\ c\end{pmatrix}$ is an e-vector corresponding to $\lambda_1 = 1$ iff $x_1 \neq 0$ and $x_1 \in \mathcal{N}(\mathbf{L}_{B_1})$. That is

$$\mathbf{L}_{B_1}(x_1) = B_1 x_1 = 0 \;\Rightarrow\; \begin{pmatrix} 0 & 1 & 0\\ 0 & 1 & 2\\ 0 & 0 & 2\end{pmatrix}\begin{pmatrix} a\\ b\\ c\end{pmatrix} = \begin{pmatrix} 0\\ 0\\ 0\end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left[\begin{array}{ccc|c} 0 & 1 & 0 & 0\\ 0 & 1 & 2 & 0\\ 0 & 0 & 2 & 0\end{array}\right] \sim \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\end{array}\right] \;\Rightarrow\; b = c = 0 \text{ and } a = t \in \mathbb{R}\backslash\{0\}.$$

That is, $x_1 \in \mathcal{N}(\mathbf{L}_{B_1}) = \left\{ t\begin{pmatrix} 1\\ 0\\ 0\end{pmatrix} : 0 \neq t \in \mathbb{R}\right\}$. Thus x_1 is an e-vector of A corresponding to $\lambda_1 = 1$ iff $x_1 = t\begin{pmatrix} 1\\ 0\\ 0\end{pmatrix}$ for some nonzero $t \in \mathbb{R}$. Consequently, (using the ordered basis β) the e-vectors of $\mathbf{T}$ corresponding to $\lambda_1 = 1$ are of the form

$$\left\{ t(1 \cdot 1 + 0 \cdot x + 0 \cdot x^2) : t \in \mathbb{R} \backslash \{0\} \right\} = \{ t \; : \; t \in \mathbb{R} \backslash \{0\} \}.$$

For $\underline{\lambda_2 = 2}$: Let $B_2 = A - \lambda_2 I_3 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Then $x_2 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is an e-vector corresponding to $\lambda_2 = 2$ iff $x_2 \neq 0$ and $x_2 \in \mathcal{N}(\mathbf{L}_{B_2})$. That is

$$\mathbf{L}_{B_2}(x_2) = B_2 x_2 = 0 \;\Rightarrow\; \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left[\begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \;\Rightarrow\; a - b = 0 \text{ and } c = 0 \Rightarrow a = b = t \in \mathbb{R} \backslash \{0\}.$$

That is, $x_2 \in \mathcal{N}(\mathbf{L}_{B_2}) = \left\{ t \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} : 0 \neq t \in \mathbb{R} \right\}$. Thus x_2 is an e-vector of A corresponding to $\lambda_2 = 2$ iff $x_2 = t \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ for some nonzero $t \in \mathbb{R}$. Consequently, (using the ordered basis β) the e-vectors of $\mathbf{T}$ corresponding to $\lambda_2 = 2$ are of the form

$$\left\{ t(1 \cdot 1 + 1 \cdot x + 0 \cdot x^2) : t \in \mathbb{R} \backslash \{0\} \right\} = \{ t(1 + x) \; : \; t \in \mathbb{R} \backslash \{0\} \}.$$

For $\underline{\lambda_3 = 3}$: Let $B_3 = A - \lambda_3 I_3 = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. Then $x_3 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is an e-vector corresponding to $\lambda_3 = 3$ iff $x_3 \neq 0$ and $x_3 \in \mathcal{N}(\mathbf{L}_{B_3})$. That is

$$\mathbf{L}_{B_3}(x_3) = B_3 x_3 = 0 \;\Rightarrow\; \begin{pmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left[\begin{array}{ccc|c} -2 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \;\Rightarrow\; a = c \text{ and } b = 2c \Rightarrow c = t \in \mathbb{R} \backslash \{0\}.$$

That is, $x_3 \in \mathcal{N}(\mathbf{L}_{B_3}) = \left\{ t \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} : 0 \neq t \in \mathbb{R} \right\}$. Thus x_3 is an e-vector of A corresponding to $\lambda_3 = 3$ iff $x_3 = t \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ for some nonzero $t \in \mathbb{R}$. Consequently, (using the ordered basis β) the e-vectors of $\mathbf{T}$ corresponding to $\lambda_3 = 3$ are of the form

$$\left\{ t(1 \cdot 1 + 2 \cdot x + 1 \cdot x^2) : t \in \mathbb{R}\backslash\{0\} \right\} = \left\{ t(1 + 2x + x^2) \, : \, t \in \mathbb{R}\backslash\{0\} \right\}.$$

Therefore, setting $t = 1$, we get $\gamma = \{ 1, 1 + x, 1 + 2x + x^2 \}$ which is an ordered basis for $\mathbb{P}_2(\mathbb{R})$ containing e-vectors of $\mathbf{T}$ and hence $\mathbf{T}$ is diagonalizable and

$$[\mathbf{T}]_\gamma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = Q^{-1} A \, Q, \quad \text{where } Q = [\ \mathbf{I}\ ]_\gamma^\beta = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that the columns of Q are the vectors $[u_i]_\beta$ for $i = 1, 2, 3$ and $u_i \in \gamma$. That is $Q = [[u_1]_\beta \quad [u_2]_\beta \quad [u_3]_\beta]$ where u_i is the i^{th} vector of γ.

Example 5.1.7

Let $\mathbf{T}$ be a linear operator defined on $M_{2\times 2}(\mathbb{R})$ by $\mathbf{T}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} d & b \\ c & a \end{pmatrix}$. Find the e-vectors of $\mathbf{T}$ and an ordered basis γ for $M_{2\times 2}(\mathbb{R})$ such that $[\mathbf{T}]_\gamma$ is a diagonal matrix.

Solution:

Let $\beta = \left\{ E^{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E^{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E^{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E^{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$. Then,

$$\begin{aligned}
\mathbf{T}\left(E^{11}\right) &= E^{22} = 0 \cdot E^{11} + 0 \cdot E^{12} + 0 \cdot E^{21} + 1 \cdot E^{22} \\
\mathbf{T}\left(E^{12}\right) &= E^{12} = 0 \cdot E^{11} + 1 \cdot E^{12} + 0 \cdot E^{21} + 0 \cdot E^{22} \\
\mathbf{T}\left(E^{21}\right) &= E^{21} = 0 \cdot E^{11} + 0 \cdot E^{12} + 1 \cdot E^{21} + 0 \cdot E^{22} \\
\mathbf{T}\left(E^{22}\right) &= E^{11} = 1 \cdot E^{11} + 0 \cdot E^{12} + 0 \cdot E^{21} + 0 \cdot E^{22}
\end{aligned}$$

Thus, $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ and hence the e-values of A are

$$f(\lambda) = |\,A - \lambda I_4\,| = \begin{vmatrix} -\lambda & 0 & 0 & 1 \\ 0 & 1-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 1 & 0 & 0 & -\lambda \end{vmatrix} = (1-\lambda)^2(\lambda^2 - 1) = 0$$

Thus, $\lambda_{1,2,3} = 1$ and $\lambda_4 = -1$.

For $\underline{\lambda = \lambda_{1,2,3} = 1}$: Let $B = A - \lambda I_4 = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$. Then $x = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$ is an e-vector corresponding to λ iff $x \neq 0$ and $x \in \mathcal{N}(\mathbf{L}_B)$. That is

$$\mathbf{L}_B(x) = Bx = 0 \;\Rightarrow\; \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left(\begin{array}{cccc|c} -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \;\Rightarrow\; a - d = 0 \Rightarrow a = d = s; b = t; c = r,$$

where $s, t, r \in \mathbb{R}\backslash\{0\}$. That is, x are of the form

$$\left\{ \begin{pmatrix} s \\ t \\ r \\ s \end{pmatrix} : s, t, r \in \mathbb{R} \right\} = \left\{ s \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + r \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} : s, t, r \in \mathbb{R} \right\}.$$

Note that $s, t,$ and r are in $\mathbb{R}$ not all zeros. Consequently, (using the ordered basis β) the e-vectors of $\mathbf{T}$ corresponding to λ are of the form

$$s \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, t \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, r \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

For $\underline{\lambda = \lambda_4 = -1}$: Let $B = A - \lambda I_4 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$. Then $x = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$ is an e-vector corresponding to λ iff $x \neq 0$ and $x \in \mathcal{N}(\mathbf{L}_B)$. That is

$$\mathbf{L}_B(x) = Bx = 0 \;\Rightarrow\; \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \;\Rightarrow\; a + d = 0; b = c = 0 \Rightarrow d = t; a = -t,$$

where $t \in \mathbb{R} \backslash \{0\}$. That is, x are of the form

$$\left\{ t \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \; : \; t \in \mathbb{R} \backslash \{0\} \right\}.$$

Consequently, (using the ordered basis β) the e-vectors of $\mathbf{T}$ corresponding to λ are of the form

$$t \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ for some } t \in \mathbb{R} \backslash \{0\}.$$

Thus, $\gamma = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ is an ordered basis for $M_{2\times 2}(\mathbb{R})$ consisting of e-vectors of $\mathbf{T}$. Therefore $\mathbf{T}$ is diagonalizable and

$$[\mathbf{T}]_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = Q^{-1} A \, Q.$$

Where Q is the matrix whose columns are $[u_i]_\beta$ for $i = 1, 2, 3, 4$ and $u_i \in \gamma$.

Example 5.1.8

Let $\mathbf{T}$ be the linear operator defined on $\mathbb{R}^2$ by $\mathbf{T}(a,b) = (-2a+3b, -10a+9b)$. Find the e-values of $\mathbf{T}$ and an ordered basis γ for $\mathbb{R}^2$ such that $[\mathbf{T}]_\gamma$ is a diagonal matrix.

Solution:

Let $\beta = \{\,(1,0),(0,1)\,\}$. Then

$$\mathbf{T}(1,0) = (-2,-10) = -2\cdot(1,0) + (-10)\cdot(0,1)$$
$$\mathbf{T}(0,1) = (3,9) = 3\cdot(1,0) + 9\cdot(0,1)$$

Thus $A = \begin{pmatrix} -2 & 3 \\ -10 & 9 \end{pmatrix}$ and the e-values of A are

$$f(\lambda) = |\,A - \lambda I_2\,| = \begin{vmatrix} -2-\lambda & 3 \\ -10 & 9-\lambda \end{vmatrix} = \cdots = (\lambda-3)(\lambda-4) = 0.$$

Therefore, $\lambda_1 = 3$ and $\lambda_2 = 4$.

For $\underline{\lambda_1 = 3}$: Let $B_1 = A - \lambda_1 I_2 = \begin{pmatrix} -5 & 3 \\ -10 & 6 \end{pmatrix}$. Then $x_1 = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ is an e-vector corresponding to $\lambda_1 = 3$ iff $x_1 \neq 0$ and $x_1 \in \mathcal{N}(\mathbf{L}_{B_1})$. That is

$$\mathbf{L}_{B_1}(x_1) = B_1 x_1 = 0 \;\Rightarrow\; \begin{pmatrix} -5 & 3 \\ -10 & 6 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left[\begin{array}{cc|c} -5 & 3 & 0 \\ -10 & 6 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -\frac{3}{5} & 0 \\ 0 & 0 & 0 \end{array}\right] \;\Rightarrow\; a - \frac{3}{5}b = 0 \;\Rightarrow\; a = \frac{3}{5}b.$$

That is, $x_1 \in \mathcal{N}(\mathbf{L}_{B_1}) = \left\{ t\begin{pmatrix} 3 \\ 5 \end{pmatrix} \;:\; 0 \neq t \in \mathbb{R} \right\}$. Thus x_1 is an e-vector of A corresponding to $\lambda_1 = 3$ iff $x_1 = t\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ for some nonzero $t \in \mathbb{R}$. Consequently, (using the ordered basis β) the e-vectors of $\mathbf{T}$ corresponding to $\lambda = 3$ are of the form

$$t\begin{pmatrix} 3 \\ 5 \end{pmatrix}, \text{ for some } t \in \mathbb{R}\backslash\{0\}.$$

For $\underline{\lambda_2 = 4}$: Let $B_2 = A - \lambda_2 I_2 = \begin{pmatrix} -6 & 3 \\ -10 & 5 \end{pmatrix}$. Then $x_2 = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ is an e-vector corresponding to $\lambda_2 = 4$ iff $x_2 \neq 0$ and $x_2 \in \mathcal{N}(\mathbf{L}_{B_2})$. That is

$$\mathbf{L}_{B_2}(x_2) = B_2 x_2 = 0 \;\Rightarrow\; \begin{pmatrix} -6 & 3 \\ -10 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is a homogenous system which can be solved using r.r.e.f. as follows:

$$\left[\begin{array}{cc|c} -6 & 3 & 0 \\ -10 & 5 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right] \;\Rightarrow\; a - \frac{1}{2}b = 0 \;\Rightarrow\; a = \frac{1}{2}b.$$

That is, $x_2 \in \mathcal{N}(\mathbf{L}_{B_2}) = \left\{ t \begin{pmatrix} 1 \\ 2 \end{pmatrix} \;:\; 0 \neq t \in \mathbb{R} \right\}$. Thus x_2 is an e-vector of A corresponding to $\lambda_2 = 4$ iff $x_2 = t \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ for some nonzero $t \in \mathbb{R}$. Consequently, (using the ordered basis β) the e-vectors of $\mathbf{T}$ corresponding to $\lambda = 4$ are of the form

$$t \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \text{ for some } t \in \mathbb{R}\backslash\{0\}.$$

Thus, $\gamma = \{\,(3,5),(1,2)\,\}$ is an ordered basis for $\mathbb{R}^2$ consisting of e-vectors of $\mathbf{T}$. Therefore $\mathbf{T}$ is diagonalizable and

$$[\mathbf{T}]_\gamma = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} = Q^{-1} A\, Q.$$

Where Q is the matrix whose columns are the vectors of γ. That is, $Q = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$.

Exercise 5.1.1

Solve the following exercises from the book at pages 256 - 260:

- 2.
- $3 : a, b$, and d.
- $4, 5$.
- $11 : a$, and c.
- $12 : a$.
- $14, 15$.

Section 5.2: Diagonalizability

In this section, we introduce a simple test to determine whether an operator or a matrix can be diagonalized. Also, we present a method for finding an ordered basis of e-vectors.

Theorem 5.2.1

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be **distinct** e-values of $\mathbf{T}$. If $x_1, x_2, \cdots, x_k$ are e-vectors of $\mathbf{T}$ such that λ_i correspond to x_i $(1 \le i \le k)$, then $\{x_1, x_2, \cdots, x_k\}$ is linearly independent set in $\mathbb{V}$.

Theorem 5.2.2

Let $\mathbf{T}$ be a linear operator on an n-dimensional vector space $\mathbb{V}$. If $\mathbf{T}$ has n distinct e-values, then $\mathbf{T}$ is diagonalizable.

Example 5.2.1

Is $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ diagonalizable? Explain.

Solution:

We first start to find the e-values of A (and hence of $\mathbf{L}_A$) using its characteristic polynomial:

$$f(\lambda) = |\, A - \lambda I_2 \,| = \begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 1 = \lambda^2 - 2\lambda = \lambda(\lambda - 2) = 0.$$

Therefore, $\lambda_1 = 0$ and $\lambda_2 = 2$. Since $\mathbf{L}_A$ is a linear operator on $\mathbb{R}^2$ and has two distinct e-values (0 and 2), then $\mathbf{L}_A$ (and hence A) is diagonalizable.

Remark 5.2.1

The converse of Theorem 5.2.1 is not true in general. That is if $\mathbf{T}$ is diagonalizable, then $\mathbf{T}$ not necessary has distinct e-values.

Definition 5.2.1

We say that a polynomial $f(t) \in \mathbb{P}(\mathbb{F})$ **splits over** $\mathbb{F}$ if there are scalars $c, a_1, a_2, \cdots, a_n$ (not necessary distinct) in $\mathbb{F}$ such that

$$f(t) = c\,(t - a_1)(t - a_2)\cdots(t - a_n).$$

Example 5.2.2

Note that $f(t) = t^2 - 1$ splits over $\mathbb{R}$, but $g(t) = t^2 + 1$ does not.

Theorem 5.2.3

The characteristic polynomial of any diagonalizable linear operator splits.

Proof:

Let $\mathbf{T}$ be a diagonalizable linear operator on the n-dimensional vector space $\mathbb{V}$ with an ordered basis β such that $[\mathbf{T}]_\beta = D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ is a diagonal matrix. The characteristic polynomial of $\mathbf{T}$ is

$$f(t) = \left| [\mathbf{T}]_\beta - tI_n \right| = |\, D - tI \,| = \begin{vmatrix} \lambda_1 - t & & 0 \\ & \ddots & \\ 0 & & \lambda_n - t \end{vmatrix}$$

$$= (\lambda_1 - t)(\lambda_2 - t)\cdots(\lambda_n - t) = (-1)^n (t - \lambda_1)\cdots(t - \lambda_n).$$

Definition 5.2.2

Let λ be an e-value of a linear operator or a matrix with characteristic polynomial $f(t)$. The (**algebraic**) **multiplicity** of λ is the largest positive integer k for which $(t - \lambda)^k$ is a factor of $f(t)$. We write $m(\lambda)$ to denote λ's multiplicity.

Example 5.2.3

Consider the characteristic polynomial $f(t) = (t - 2)^4(t - 3)^2(t - 1)$. Hence $\lambda = 2, 3, 1$ are the e-values with multiplicities: $m(\lambda = 2) = 4$, $m(\lambda = 3) = 2$, and $m(\lambda = 1) = 1$.

Definition 5.2.3

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$, and let λ be an e-value of $\mathbf{T}$. Define

$$E_\lambda = \{\, x \in \mathbb{V} \,:\, \mathbf{T}\,(x) = \lambda\, x \,\} = \mathcal{N}(\mathbf{T} - \lambda \mathbf{I}_V).$$

The set E_λ is called the **eigenspace** (or **e-space** for short) of $\mathbf{T}$ corresponding to λ. We also define the eigen space of a square matrix A to be the eigen space of $\mathbf{L}_A$.

Remark 5.2.2

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$, and let λ be an e-value of $\mathbf{T}$. Then

1. E_λ is a subspace of $\mathbb{V}$.
2. E_λ consists of the zero vector and the e-vector of $\mathbf{T}$ corresponding to λ.
3. $\mathbf{dim}(E_\lambda)$ is the maximum number of linearly independent e-vectors corresponding to λ.

Theorem 5.2.4

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and let λ be an e-value of $\mathbf{T}$ having multiplicity m. Then $1 \le \mathbf{dim}(E_\lambda) \le m$.

Theorem 5.2.5: Diagonalization Test

Let $\mathbf{T}$ be a linear operator on an n-dimensional vector space $\mathbb{V}$. Then, $\mathbf{T}$ is diagonalizable if and only if both of the following conditions hold.

1. The characteristic polynomial of $\mathbf{T}$ splits, and
2. For each e-value λ of $\mathbf{T}$, $m(\lambda) = \mathbf{dim}(E_\lambda) = n - rank(\mathbf{T} - \lambda \mathbf{I}_V)$.

Moreover, if $\mathbf{T}$ is diagonalizable and β_i is an ordered basis for E_{λ_i} for $i = 1, \cdots, k$, then $\beta = \beta_1 \cup \cdots \cup \beta_k$ (in corresponding order of e-values) is an ordered basis for $\mathbb{V}$ consisting of e-vectors of $\mathbf{T}$.

Example 5.2.4

Let $\mathbf{T}$ be a linear operator on $\mathbb{P}_2(\mathbb{R})$ defined by $\mathbf{T}(f(x)) = f'(x)$. Is $\mathbf{T}$ diagonalizable? Explain.

Solution:

Choose the standard ordered basis $\beta = \{1, x, x^2\}$ for $\mathbb{P}_2(\mathbb{R})$. Then,

$$\left.\begin{aligned} \mathbf{T}(1) &= 0 = 0\cdot 1 + 0\cdot x + 0\cdot x^2 \\ \mathbf{T}(x) &= 1 = 1\cdot 1 + 0\cdot x + 0\cdot x^2 \\ \mathbf{T}(x^2) &= 2x = 0\cdot 1 + 2\cdot x + 0\cdot x^2 \end{aligned}\right\} \Rightarrow A = [\mathbf{T}]_\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

The characteristic polynomial of $\mathbf{T}$ is

$$f(\lambda) = |A - \lambda I_3| = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda^3 = 0.$$

Therefore, $\mathbf{T}$ has one e-value $\lambda = 0$ with multiplicity $m(0) = 3$. The e-space E_λ corresponding to $\lambda = 0$ is $E_\lambda = \mathcal{N}(\mathbf{T} - \lambda I_3) = \mathcal{N}(\mathbf{T})$. That is,

$$E_\lambda = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \left\{ t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Hence E_λ is the subspace of $\mathbb{P}_2(\mathbb{R})$ consisting of the constant polynomials. So, $\{1\}$ is a basis for E_λ and hence $\mathbf{dim}(E_\lambda) = 1 \neq m(0) = 3$.

Therefore, there is no ordered basis for $\mathbb{P}_2(\mathbb{R})$ consisting of e-vectors of $\mathbf{T}$. Therefore, $\mathbf{T}$ is not diagonalizable.

Example 5.2.5

Let $\mathbf{T}$ be a linear operator on $\mathbb{R}^3$ defined by $\mathbf{T}(a, b, c) = (4a + c, 2a + 3b + 2c, a + 4c)$. Determine the e-space corresponding to each e-value of $\mathbf{T}$.

Solution:

Choose $\beta = \{ E_1, E_2, E_3 \}$ the standard ordered basis for $\mathbb{R}^3$. Then,

$$\left.\begin{aligned} \mathbf{T}(E_1) &= (4,2,1) = 4\cdot E_1 + 2\cdot E_2 + 1\cdot E_3 \\ \mathbf{T}(E_2) &= (0,3,0) = 0\cdot E_1 + 3\cdot E_2 + 0\cdot E_3 \\ \mathbf{T}(E_3) &= (1,2,4) = 1\cdot E_1 + 2\cdot E_2 + 4\cdot E_3 \end{aligned}\right\} \Rightarrow A = [\mathbf{T}]_\beta = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}.$$

The characteristic polynomial of $\mathbf{T}$ is

$$f(\lambda) = |\, A - \lambda I_3 \,| = \begin{vmatrix} 4-\lambda & 0 & 1 \\ 2 & 3-\lambda & 2 \\ 1 & 0 & 4-\lambda \end{vmatrix} = \cdots = (3-\lambda)(\lambda-3)(\lambda-5) = 0$$

Thus, $\mathbf{T}$ has e-values: $\lambda_1 = 3$ with $m(3) = 2$ and $\lambda_2 = 5$ with $m(5) = 1$.

For $\underline{E_{\lambda_1}}$: The e-space E_{λ_1} corresponding to $\lambda_1 = 3$ is $E_{\lambda_1} = \mathcal{N}(\mathbf{T} - 3I_3)$. Therefore

$$E_{\lambda_1} = \left\{ (a,b,c) \in \mathbb{R}^3 \; : \; \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow a = -c; c = r, b = t \in \mathbb{R}.$$

Setting $r, t \in \mathbb{R}$, we get

$$E_{\lambda_1} = \left\{ r \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \; : \; t, r \in \mathbb{R} \right\}.$$

Therefore, $\gamma_1 = \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ is a basis for E_{λ_1}. Thus, $\mathbf{dim(}E_{\lambda_1}\mathbf{)} = 2 = m(\lambda_1)$.

For $\underline{E_{\lambda_2}}$: The e-space E_{λ_2} corresponding to $\lambda_2 = 5$ is $E_{\lambda_2} = \mathcal{N}(\mathbf{T} - 5I_3)$. Therefore

$$E_{\lambda_2} = \left\{ (a,b,c) \in \mathbb{R}^3 \; : \; \begin{pmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 2 & -2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow a = c, b = 2c; c = t \in \mathbb{R}.$$

Setting $r, t \in \mathbb{R}$, we get

$$E_{\lambda_2} = \left\{ t \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\}.$$

Therefore, $\gamma_2 = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$ is a basis for E_{λ_2}. Thus, $\mathbf{dim}(E_{\lambda_2}) = 1 = m(\lambda_2)$.

Afterall, $\gamma = \gamma_1 \cup \gamma_2 = \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$ is a basis for $\mathbb{R}^3$ consisting e-vectors of $\mathbf{T}$.

Therefore, $\mathbf{T}$ is diagonalizable and

$$[\mathbf{T}]_\gamma = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Example 5.2.6

Let $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. Is A diagonalizable? Explain.

Solution:

The characteristic polynomial of A is

$$f(t) = |\,A - tI_3\,| = \begin{vmatrix} 3-t & 1 & 0 \\ 0 & 3-t & 0 \\ 0 & 0 & 4-t \end{vmatrix} = (3-t)^2(4-t) = 0.$$

Thus, $\lambda_1 = 3$ with $m(3) = 2$ and $\lambda_2 = 4$ with $m(4) = 1$. But we note that

$$A - \lambda_1 I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

has rank 2 and hence $\mathbf{dim}(E_{\lambda_1}) = 3 - 2 = 1$ which is different from the multiplicity of λ_1.

Therefore, A is not diagonalizable.

Example 5.2.7

Let $\mathbf{T}$ be the linear operator on $\mathbb{P}_2(\mathbb{R})$ defined by

$$\mathbf{T}\,(f(x)) = f(1) + f'(0)\cdot x + (f'(0) + f''(0))\cdot x^2.$$

Is $\mathbf{T}$ diagonalizable? Explain.

Solution:

Let $\beta = \{\,1, x, x^2\,\}$ be the standard ordered basis for $\mathbb{P}_2(\mathbb{R})$. Then

$$\left.\begin{array}{l} \mathbf{T}(1) = 1 = 1\cdot 1 + 0\cdot x + 0\cdot x^2 \\ \mathbf{T}(x) = 1 + x + (1+0)x^2 = 1\cdot 1 + 1\cdot x + 1\cdot x^2 \\ \mathbf{T}(x^2) = 1 + 2x^2 = 1\cdot 1 + 0\cdot x + 2\cdot x^2 \end{array}\right\} \Rightarrow A = [\mathbf{T}]_\beta = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}.$$

The characteristic polynomial of $\mathbf{T}$ is

$$f(t) = |\,A - tI_3\,| = \begin{vmatrix} 1-t & 1 & 1 \\ 0 & 1-t & 0 \\ 0 & 1 & 2-t \end{vmatrix} = (1-t)^2(2-t) = 0.$$

Thus, $\lambda_1 = 1$ with $m(1) = 2$ and $\lambda_2 = 2$ with $m(2) = 1$.

For $\underline{E_{\lambda_1}}$: The e-space E_{λ_1} corresponding to $\lambda_1 = 1$ is $E_{\lambda_1} = \mathcal{N}(\mathbf{T} - 1I_3)$. Therefore

$$E_{\lambda_1} = \left\{ (a,b,c) \in \mathbb{R}^3 \;:\; \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow b = -c.$$

Setting $a = t$ and $c = r$ both in $\mathbb{R}$, we get

$$E_{\lambda_1} = \left\{ t\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + r\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \;:\; t, r \in \mathbb{R} \right\}.$$

Therefore, $\gamma_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$ is a basis for E_{λ_1}.

For $\underline{E_{\lambda_2}}$: The e-space E_{λ_2} corresponding to $\lambda_2 = 2$ is $E_{\lambda_2} = \mathcal{N}(\mathbf{T} - 2I_3)$. Therefore

$$E_{\lambda_2} = \left\{ (a,b,c) \in \mathbb{R}^3 : \begin{pmatrix} -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{ccc|c} -1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow b = 0; a = c.$$

Setting $c = t \in \mathbb{R}$, we get

$$E_{\lambda_1} = \left\{ t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\}.$$

Therefore, $\gamma_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$ is a basis for E_{λ_2}.

Thus, $\gamma = \gamma_1 \cup \gamma_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$ is a basis for $\mathbb{R}^3$ consisting of e-vectors of A.

Therefore, the vectors in γ are the coordinate vectors relative to β of the vectors in the set $\alpha = \{ 1, -x + x^2, 1 + x^2 \}$ which is an ordered basis for $\mathbb{P}_2(\mathbb{R})$ consisting e-vectors of $\mathbf{T}$. Thus,

$$[\mathbf{T}]_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Example 5.2.8

Let $A = \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}$. Is A diagonalizable? Explain your answer and compute A^n for positive integer n.

Solution:

The characteristic polynomial of A is

$$f(t) = |\, A - tI_2 \,| = \begin{vmatrix} -t & -2 \\ 1 & 3 - t \end{vmatrix} = t^2 - 3t + 2 = (t-1)(t-2) = 0.$$

Thus, $\lambda_1 = 1$ with $m(1) = 1$ and $\lambda_2 = 2$ with $m(2) = 1$. Then the operator $\mathbf{L}_A$ has two distinct e-values and hence A is diagonalizable.

For $\underline{E_{\lambda_1}}$: The e-space E_{λ_1} corresponding to $\lambda_1 = 1$ is $E_{\lambda_1} = \mathcal{N}(A - 1I_2)$. Therefore

$$E_{\lambda_1} = \left\{ (a,b) \in \mathbb{R}^2 \, : \, \begin{pmatrix} -1 & -2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{cc|c} -1 & -2 & 0 \\ 1 & 2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right] \;\Rightarrow\; a = -2b.$$

Setting $b = t \in \mathbb{R}$, we get

$$E_{\lambda_1} = \left\{ t \begin{pmatrix} -2 \\ 1 \end{pmatrix} \, : \, t \in \mathbb{R} \right\}.$$

Therefore, $\gamma_1 = \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$ is a basis for E_{λ_1}.

For $\underline{E_{\lambda_2}}$: The e-space E_{λ_2} corresponding to $\lambda_2 = 2$ is $E_{\lambda_2} = \mathcal{N}(A - 2I_2)$. Therefore

$$E_{\lambda_2} = \left\{ (a,b) \in \mathbb{R}^2 \, : \, \begin{pmatrix} -2 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{cc|c} -2 & -2 & 0 \\ 1 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \;\Rightarrow\; a = -b.$$

Setting $b = t \in \mathbb{R}$, we get

$$E_{\lambda_2} = \left\{ t \begin{pmatrix} -1 \\ 1 \end{pmatrix} \, : \, t \in \mathbb{R} \right\}.$$

Therefore, $\gamma_2 = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$ is a basis for E_{λ_2}.

Thus, $\gamma = \gamma_1 \cup \gamma_2 = \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$ is a basis for $\mathbb{R}^2$ consisting of e-vectors of A.

Note that $D := [\mathbf{L}_A]_\gamma = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = Q^{-1} A\, Q$ where $Q = \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix}$ and $Q^{-1} = \begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix}$.

Therefore, $A = Q\, D\, Q^{-1}$ and hence $A^n = Q\, D^n\, Q^{-1}$; that is

$$A^n = \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1^n & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix} = \cdots = \begin{pmatrix} 2 - 2^n & 2 - 2^{n+1} \\ -1 + 2^n & -1 + 2^{n+1} \end{pmatrix}.$$

Exercise 5.2.1

Solve the following exercises from the book at pages 279 - 283:

- $2, 3$.
- $7, 8$.

Section 5.4: Invariant Subspaces and The Cayley-Hamilton Theorem

Definition 5.4.1

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$. A subspace $\mathbb{W}$ of $\mathbb{V}$ is called **T-invariant subspace** of $\mathbb{V}$ if $\mathbf{T}(\mathbb{W}) \subseteq \mathbb{W}$; that is if $\mathbf{T}(x) \in \mathbb{W}$ for all $x \in \mathbb{W}$.

Remark 5.4.1

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$. Then the following subspaces of $\mathbb{V}$ are $\mathbf{T}$-invariant:

1. $\{0\}$.
2. $\mathbb{V}$.
3. $\mathcal{R}(\mathbf{T})$.
4. $\mathcal{N}(\mathbf{T})$.
5. E_λ for any e-value λ of $\mathbf{T}$.

Example 5.4.1

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^3$ defined by $\mathbf{T}(a,b,c) = (a+b, b+c, 0)$. Show that the subspaces of $\mathbb{R}^3$, $\mathbb{W}_1$ and $\mathbb{W}_2$, are $\mathbf{T}$-invariant, where

$$\textcircled{1} : \mathbb{W}_1 = \{(a,b,0) : a,b \in \mathbb{R}\}, \quad \text{and} \quad \textcircled{2} : \mathbb{W}_2 = \{(a,0,0) : a \in \mathbb{R}\}.$$

Solution:

①: Clearly, $\mathbf{T}(a,b,0) = (a+b,b,0) \in \mathbb{W}_1$ for all $(a,b,0) \in \mathbb{W}_1$. Thus, $\mathbb{W}_1$ is a $\mathbf{T}$-invariant subspace of $\mathbb{R}^3$.

②: Clearly, $\mathbf{T}(a,0,0) = (a,0,0) \in \mathbb{W}_2$ for all $(a,0,0) \in \mathbb{W}_1$. Thus, $\mathbb{W}_2$ is a $\mathbf{T}$-invariant subspace of $\mathbb{R}^3$.

Definition 5.4.2

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$, and let x be a nonzero vector in $\mathbb{V}$. The subspace

$$\mathbb{W} = \textbf{span}\left\{x, \mathbf{T}(x), \mathbf{T}^2(x), \cdots\right\},$$

where $\mathbf{T}^2(x) = \mathbf{T}(\mathbf{T}(x))$, $\mathbf{T}^3(x) = \mathbf{T}(\mathbf{T}(\mathbf{T}(x)))$, and so on, is called a **T-cyclic subspace** of $\mathbb{V}$ generated by x.

Example 5.4.2

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^3$ defined by $\mathbf{T}(a, b, c) = (-b + c, a + c, 3c)$. Determine the $\mathbf{T}$-cyclic subspace of $\mathbb{R}^3$ generated by $E_1 = (1, 0, 0)$.

Solution:

We simply compute the set containing E_1 and $\mathbf{T}^i(E_1)$ for $i = 1, 2, \cdots$.

$$\mathbf{T}(E_1) = \mathbf{T}(1, 0, 0) = (0, 1, 0) = E_2,$$
$$\mathbf{T}^2(E_1) = \mathbf{T}\left(\mathbf{T}\left(E_1\right)\right) = \mathbf{T}\left(E_2\right) = (-1, 0, 0) = -E_1.$$

Therefore, $\mathbb{W} = \mathbf{span}\,\{\, E_1, \mathbf{T}\left(E_1\right), \mathbf{T}^2(E_1), \cdots \} = \mathbf{span}\,\{\, E_1, E_2 \,\} = \{\, (s, t, 0) \,:\, s, t \in \mathbb{R} \,\}$ is the $\mathbf{T}$-cyclic subspace of $\mathbb{R}^3$ generated by E_1.

Remark 5.4.2

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$, and let x be a nonzero vector in $\mathbb{V}$. The subspace $\mathbb{W}$ generated by x is the smallest $\mathbf{T}$-invariant subsapce which contains x. That is, any $\mathbf{T}$-invariant subspace of $\mathbb{V}$ containing x must contain $\mathbb{W}$.

Example 5.4.3

Let $\mathbf{T}$ be the linear operator on $\mathbb{P}_2(\mathbb{R})$ defined by $\mathbf{T}\left(f(x)\right) = f'(x)$. Determine the $\mathbf{T}$-cyclic subspace of $\mathbb{P}_2(\mathbb{R})$ generated by x^2.

Solution:

Note that $\mathbf{T}(x^2) = 2x$, $\mathbf{T}^2(x^2) = \mathbf{T}(2x) = 2$, and $\mathbf{T}^3(x^2) = \mathbf{T}(2) = 0$. Therefore, $\mathbb{W} = \mathbf{span}\,\{\, x^2, 2x, 2 \,\} = \mathbb{P}_2(\mathbb{R})$ is the $\mathbf{T}$-cyclic subspace of $\mathbb{P}_2(\mathbb{R})$ generated by x^2.

Example 5.4.4

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^4$ defined by $\mathbf{T}(a,b,c,d) = (a+b+2c-d, b+d, 2c-d, c+d)$, and let $\mathbb{W} = \{ (t,s,0,0) : t,s \in \mathbb{R} \}$. Show that $\mathbb{W}$ is a $\mathbf{T}$-invariant subspace of $\mathbb{R}^4$.

Solution:

Choose arbitrary $x = (t,s,0,0) \in \mathbb{W}$. Then

$$\mathbf{T}(x) = (t+s, s, 0, 0) \in \mathbb{W}.$$

Thus, $\mathbf{T}(\mathbb{W}) \subseteq \mathbb{W}$ and hence $\mathbb{W}$ is a $\mathbf{T}$-invariant subsapce of $\mathbb{R}^4$.

Theorem 5.4.1

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and let $\mathbb{W}$ be a $\mathbf{T}$-cyclic subspace of $\mathbb{V}$ generated by $x \in \mathbb{V}$. Let $\mathbf{dim}(\mathbb{W}) = k$. Then $\left\{ x, \mathbf{T}(x), \cdots, \mathbf{T}^{k-1}(x) \right\}$ is a basis for $\mathbb{W}$.

Example 5.4.5

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^3$ defined by $\mathbf{T}(a,b,c) = (-b+c, a+c, 3c)$, and let $\mathbb{W}$ be the $\mathbf{T}$-cyclic subspace of $\mathbb{R}^3$ generated by E_1.

Solution:

Clearly, $E_1 = (1,0,0)$, $\mathbf{T}(E_1) = (0,1,0) = E_2$, and $\mathbf{T}^2(E_1) = \mathbf{T}(E_2) = (-1,0,0) = -E_1$. Therefore, $\mathbb{W} = \mathbf{span}\,\{ E_1, E_2 \}$ and hence $\mathbf{dim}(\mathbb{W}) = 2$. Thus, by Theorem 5.4.1, $\gamma = \{ E_1, E_2 \}$ is an ordered basis for $\mathbb{W}$.

Theorem 5.4.2: The Cayley-Hamilton Theorem

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and let $f(t)$ be the characteristic polynomial of $\mathbf{T}$. Then $f(\mathbf{T}) = \mathbf{T}_0$, the zero transformation. That is, $\mathbf{T}$ "satisfies" its characteristic equation.

Theorem 5.4.3: The Cayley-Hamilton Theorem for Matrices

Let A be an $n \times n$ matrix, and let $f(t)$ be the characteristic polynomial of A. Then $f(A) = 0$, the $n \times n$ zero matrix.

Example 5.4.6

Verify the Cayley-Hamilton theorem for the linear operator $\mathbf{T}$ defined on $\mathbb{R}^2$ by $\mathbf{T}(a, b) = (a + 2b, -2a + b)$.

Solution:

Let $\beta = \{ E_1, E_2 \}$ be an ordered basis for $\mathbb{R}^2$. Then

$$\mathbf{T}(E_1) = (1, -2) = E_1 + (-2)E_2$$
$$\mathbf{T}(E_2) = (2, 1) = 2E_1 + E_2.$$

Thus, $A = [\mathbf{T}]_\beta = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$. The characteristic polynomial of $\mathbf{T}$ is therefore

$$f(t) = |\, A - tI_2 \,| = \begin{vmatrix} 1-t & 2 \\ -2 & 1-t \end{vmatrix} = (1-t)^2 + 4 = t^2 - 2t + 5 = 0.$$

That is,

$$\begin{aligned} f(\mathbf{T}) &= (\mathbf{T}^2 - 2\mathbf{T} + 5\mathbf{I}_T) \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \mathbf{T}^2 \begin{pmatrix} a \\ b \end{pmatrix} - 2\mathbf{T} \begin{pmatrix} a \\ b \end{pmatrix} + 5\mathbf{I}_T \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \mathbf{T} \begin{pmatrix} a+2b \\ -2a+b \end{pmatrix} - 2 \begin{pmatrix} a+2b \\ -2a+b \end{pmatrix} + 5 \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} (a+2b) + 2(-2a+b) \\ -2(a+2b) + (-2a+b) \end{pmatrix} + \begin{pmatrix} -2a-4b \\ 4a-2b \end{pmatrix} + \begin{pmatrix} 5a \\ 5b \end{pmatrix} \\ &= \begin{pmatrix} -3a+4b \\ -4a-3b \end{pmatrix} + \begin{pmatrix} -2a-4b \\ 4a-2b \end{pmatrix} + \begin{pmatrix} 5a \\ 5b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{T}_0 \begin{pmatrix} a \\ b \end{pmatrix}. \end{aligned}$$

Note that

$$f(A) = A^2 - 2A + 5I = \begin{pmatrix} -3 & 4 \\ -4 & -3 \end{pmatrix} + \begin{pmatrix} -2 & 4 \\ 4 & -2 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

Example 5.4.7

Let $\mathbf{T}$ be the linear operator defined on $\mathbb{P}_1(\mathbb{R})$ by $\mathbf{T}(f(x)) = f(x) + f'(x)$. Verify the Cayley-Hamilton Theorem for $\mathbf{T}$.

Solution:

Let $\beta = \{1, x\}$. Then,

$$\mathbf{T}(1) = 1 + 0 = 1 \cdot 1 + 0 \cdot x$$
$$\mathbf{T}(x) = x + 1 = 1 \cdot 1 + 1 \cdot x$$

Thus, $[\mathbf{T}]_\beta = A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, and the characteristic polynomial of $\mathbf{T}$ is therefore,

$$f(t) \, | \, A - tI_2 \, | = \begin{vmatrix} 1-t & 1 \\ 0 & 1-t \end{vmatrix} = (1-t)^2 = t^2 - 2t + 1.$$

Therefore,

$$\begin{aligned} f(\mathbf{T}) &= (\mathbf{T}^2 - 2\mathbf{T} + \mathbf{I}_T)\begin{pmatrix} a \\ bx \end{pmatrix} = \mathbf{T}^2\begin{pmatrix} a \\ bx \end{pmatrix} - 2\mathbf{T}\begin{pmatrix} a \\ bx \end{pmatrix} + \begin{pmatrix} a \\ bx \end{pmatrix} \\ &= \mathbf{T}\begin{pmatrix} a \\ b+bx \end{pmatrix} - 2\begin{pmatrix} a \\ b+bx \end{pmatrix} + \begin{pmatrix} a \\ bx \end{pmatrix} = \begin{pmatrix} a \\ (b+bx)+b \end{pmatrix} + \begin{pmatrix} -2a \\ -2b-2bx \end{pmatrix} + \begin{pmatrix} a \\ bx \end{pmatrix} \\ &= \begin{pmatrix} a \\ 2b+bx \end{pmatrix} + \begin{pmatrix} -2a \\ -2b-2bx \end{pmatrix} + \begin{pmatrix} a \\ bx \end{pmatrix} = \begin{pmatrix} 2a-2a \\ (2b-2b)+(-2bx+2bx) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{T}_0\begin{pmatrix} a \\ bx \end{pmatrix}. \end{aligned}$$

Note that,

$$\begin{aligned} f(A) = (A^2 - 2A + I_2) &= \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} - 2\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -2 & -2 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 \end{aligned}$$

Example 5.4.8

Use Cayley-Hamilton Theorem to find A^{-1} if $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix}$.

Solution:

Note that $|A| = -2 \neq 0$ and hence A^{-1} exists. The characteristic polynomial of A is

$$\begin{aligned} f(t) = |A - tI_3| &= \begin{vmatrix} 1-t & 2 & 1 \\ 0 & 2-t & 3 \\ 0 & 0 & -1-t \end{vmatrix} \\ &= (1-t)(2-t)(-1-t) = -(2 - 3t + t^2)(1+t) \\ &= -((2+2t) - 3t - 3t^2 + t^2 + t^3) = -t^3 + 2t^2 + t - 2. \end{aligned}$$

Thus,

$$\begin{aligned} f(A) = -A^3 + 2A^2 + A - 2I_3 &= 0 \\ \Rightarrow 2I_3 &= -A^3 + A^2 + A \\ \Rightarrow I_3 &= -\frac{1}{2}A^3 + A^2 + \frac{1}{2}A \\ \Rightarrow I_3 &= \left(-\frac{1}{2}A^2 + A + \frac{1}{2}I_3\right) A. \end{aligned}$$

Hence $A^{-1} = -\frac{1}{2}A^2 + A + \frac{1}{2}I_3$.

Exercise 5.4.1

Solve the following exercises from the book at pages 321 - 327:

- $2, 3$, and 6.

Chapter

6 Inner Product Spaces

Section 6.1: Inner Product and Norms

Remark 6.1.1

Let $z = a + ib \in \mathbb{C}$ for some $a, b \in \mathbb{R}$, then

1. $|\,z\,| = \sqrt{a^2 + b^2}$ is called the absolute value for modulus of z.
2. $z\,\overline{z} = |\,z\,|^2$.
3. $z + \overline{z} = 2Re(z) = 2a$.
4. $z - \overline{z} = 2Im(z) = 2b$.
5. $Re(z) \leq |\,z\,|$.
6. $\overline{\overline{z}} = z$.
7. $\overline{z + w} = \overline{z} + \overline{w}$ and $\overline{zw} = \overline{z}\,\overline{w}$.

Definition 6.1.1

Let $\mathbb{V}$ be a vector space over a field $\mathbb{F}$. An **inner product** on $\mathbb{V}$ is a function that assigns, to every pair of vectors $x, y \in \mathbb{V}$, a scalar in $\mathbb{F}$, denoted by $\langle\, x, y\,\rangle$, such that for all $x, y, z \in \mathbb{V}$ and all $c \in \mathbb{F}$, the following conditions hold:

1. $\langle\, x + y, z\,\rangle = \langle\, x, z\,\rangle + \langle\, y, z\,\rangle$.
2. $\langle\, cx, y\,\rangle = c\langle\, x, y\,\rangle$.
3. $\overline{\langle\, x, y\,\rangle} = \langle\, y, x\,\rangle$, where the bar denotes the complex conjugation.
4. $\langle\, x, x\,\rangle > 0$ if $x \neq 0$.

Note that, Condition (3) reduces to $\langle\, x, y\,\rangle = \langle\, y, x\,\rangle$ if $\mathbb{F} = \mathbb{R}$.

Example 6.1.1

Let $\mathbb{V} = C([0, 1])$, the vector space of real valued continuous function on $[0, 1]$. Define

$$\langle\, f, g\,\rangle = \int_0^1 f(t)g(t)\;dt.$$

Show that $\langle f, g \rangle$ is an inner product on $\mathbb{V}$.

Solution:

For every $f, g, h \in \mathbb{V}$ and every $c \in \mathbb{R}$, we have

1.
$$\begin{aligned}\langle f+g, h \rangle &= \int_0^1 (f(t)+g(t))h(t)\ dt = \int_0^1 (f(t)h(t)+g(t)h(t))\ dt \\ &= \int_0^1 f(t)h(t)\ dt + \int_0^1 g(t)h(t)\ dt = \langle f, h \rangle + \langle g, h \rangle.\end{aligned}$$

2. $\langle cf, g \rangle = \int_0^1 c\, f(t)g(t)\ dt = c \int_0^1 f(t)g(t)\ dt = c\langle f, g \rangle$.
3. $\langle f, g \rangle = \int_0^1 f(t)g(t)\ dt = \int_0^1 g(t)f(t)\ dt = \langle g, f \rangle$.
4. If $f \neq 0$, $\langle f, f \rangle = \int_0^1 f(t)f(t)\ dt = \int_0^1 f^2(t)\ dt > 0$.

Thus, $\langle f, g \rangle$ is an inner product on $C([0, 1)]$.

Example 6.1.2

For $x = (a_1, a_2, \cdots, a_n), y = (b_1, b_2, \cdots, b_n) \in \mathbb{F}^n$, define $\langle x, y \rangle = \sum_{i=1}^n a_i \overline{b_i}$. Show that $\langle x, y \rangle$ is an inner product on $\mathbb{F}^n$.

Solution:

For any $x = (a_1, \cdots, a_n), y = (b_1, \cdots, b_n), z = (c_1, \cdots, c_n) \in \mathbb{F}^n$ and $k \in \mathbb{F}$, we have

1. $\langle x+y, z \rangle = \sum_{i=1}^n (a_i + b_i)\overline{c_i} = \sum_{i=1}^n (a_i\overline{c_i} + b_i\overline{c_i}) = \sum_{i=1}^n a_i\overline{c_i} + \sum_{i=1}^n b_i\overline{c_i} = \langle x, z \rangle + \langle y, z \rangle$.
2. $\langle kx, y \rangle = \sum_{i=1}^n k a_i \overline{b_i} = k \sum_{i=1}^n a_i \overline{b_i} = k\langle x, y \rangle$.
3. $\overline{\langle x, y \rangle} = \overline{\sum_{i=1}^n a_i\overline{b_i}} = \sum_{i=1}^n \overline{a_i}\overline{\overline{b_i}} = \sum_{i=1}^n \overline{a_i} b_i = \sum_{i=1}^n b_i \overline{a_i} = \langle y, x \rangle$.
4. If $x \neq 0$, $\langle x, x \rangle = \sum_{i=1}^n a_i\overline{a_i} = \sum_{i=1}^n |\, a_i \,|^2 > 0$.

Remark 6.1.2

Note that, the inner product defined in Example 6.1.2, is called the **standard inner product** on $\mathbb{F}^n$. In case of $\mathbb{F} = \mathbb{R}$, we have $\langle x, y \rangle = \sum_{i=1}^n a_i b_i = x \cdot y$ which is the usual **dot (or scalar) product** of x and y in $\mathbb{R}^n$.

Definition 6.1.2

Let $A = (a_{ij}) \in M_{m\times n}(\mathbb{F})$. We define the **conjugate transpose** or **adjoint** of A to be the $n \times m$ matrix A^* such that $a^*_{ij} = \overline{a_{ji}}$ for all $1 \le i \le m$ and $1 \le j \le n$. Note that if $\mathbb{F} = \mathbb{R}$, then we simply write A^t instead of A^*.

Example 6.1.3

If $A = \begin{pmatrix} i & 1+2i \\ 2 & 3+4i \end{pmatrix}$, then $A^* = \begin{pmatrix} -i & 2 \\ 1-2i & 3-4i \end{pmatrix}$.

Example 6.1.4

Let $\mathbb{V} = M_{n\times n}(\mathbb{F})$, and define $\langle A, B \rangle = tr(B^*A)$ for $A, B \in \mathbb{V}$. Show that $\langle A, B \rangle$ is an inner product on $\mathbb{V}$.

Solution:

For any $A, B, C \in \mathbb{V}$ and $c \in \mathbb{F}$, we have

1. $\langle A + B, C \rangle = tr(C^*(A + B)) = tr(C^*A + C^*B) = tr(C^*A) + tr(C^*B)$
$= \langle A, C \rangle + \langle B, C \rangle$.
2. $\langle cA, B \rangle = tr(B^*(cA)) = c\,tr(B^*A) = c\langle A, B \rangle$.
3. $\overline{\langle A, B \rangle} = \overline{tr(B^*A)} = tr(\overline{B^*A}) = tr(A^*B) = \langle B, A \rangle$.
4. $\langle A, A \rangle = tr(A^*A) = \sum_{i=1}^{n} b_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} a^*_{ik}a_{ki} = \sum_{i=1}^{n}\sum_{k=1}^{n} \overline{a_{ki}}a_{ki} = \sum_{i=1}^{n}\sum_{k=1}^{n} |\,a_{ki}\,|^2$. Note that if $A \ne 0$, then $a_{ki} \ne 0$ for some k and i. So, $\langle A, A \rangle > 0$.

Here is a detailed proof of $tr(\overline{B^*A}) = tr(A^*B)$: Assuming the $C = (c_{ij}) = B^*A$, we have

$$\begin{aligned} tr(\overline{B^*A}) &= \sum_{i}^{n} \overline{c_{ii}} = \sum_{i}^{n}\sum_{j}^{n} \overline{b^*_{ij}a_{ji}} = \sum_{i}^{n}\sum_{j}^{n} \overline{\overline{b_{ji}}a_{ji}} \\ &= \sum_{i}^{n}\sum_{j}^{n} b_{ji}\overline{a_{ji}} = \sum_{i}^{n}\sum_{j}^{n} a^*_{ij}b_{ji} = tr(A^*B). \end{aligned}$$

Note that, a vector space $\mathbb{V}$ over a field $\mathbb{F}$ together with specific inner product on $\mathbb{V}$ is called an **inner product space**. If $\mathbb{F} = \mathbb{C}$, we call $\mathbb{V}$ a complex inner product space, and if $\mathbb{F} = \mathbb{R}$, we call $\mathbb{V}$ a real inner product space.

Theorem 6.1.1

Let $\mathbb{V}$ be an inner product space. Then for $x, y, z \in \mathbb{V}$ and $c \in \mathbb{F}$, the following statements are true:

1. $\langle\, x, y+z \,\rangle = \langle\, x, y \,\rangle + \langle\, x, z \,\rangle$.
2. $\langle\, x, cy \,\rangle = \overline{c}\, \langle\, x, y \,\rangle$.
3. $\langle\, x, 0 \,\rangle = \langle\, 0, x \,\rangle = 0$.
4. $\langle\, x, x \,\rangle = 0$ iff $x = 0$.
5. If $\langle\, x, y \,\rangle = \langle\, x, z \,\rangle$ for all $x \in \mathbb{V}$, then $y = z$. $\langle\, y-z, y-z \,\rangle = 0 \Rightarrow y - z = 0 \Rightarrow y = z$.

Definition 6.1.3

Let $\mathbb{V}$ be an inner product space. For $x \in \mathbb{V}$, we define the **norm** or **length** of x by

$$\|x\| = \sqrt{\langle\, x, x \,\rangle}.$$

Note that if $\mathbb{V} = \mathbb{R}$, then $\|x\| = |\, x \,|$ and if $\mathbb{V} = \mathbb{R}^n$, then $\|x\| = \sqrt{\langle\, x, x \,\rangle} = \sqrt{x \cdot x}$.

Theorem 6.1.2

Let $\mathbb{V}$ be an inner product space over a field $\mathbb{F}$. Then for all $x, y \in \mathbb{V}$ and $c \in \mathbb{F}$, the following statements are true:

1. $\|c\,x\| = |\,c\,|\; \|x\|$.
2. $\|x\| \geq 0$; and $\|x\| = 0$ iff $x = 0$.
3. (Cauchy-Schwarz Inequality) $|\,\langle\, x, y \,\rangle\,| \leq \|x\| \cdot \|y\|$.
4. (Triangle Inequality) $\|x + y\| \leq \|x\| + \|y\|$.

Proof:

1. $\|c\,x\| = \sqrt{\langle\, cx, cx \,\rangle} = \sqrt{c\bar{c}\langle\, x, x \,\rangle} = \sqrt{|\,c\,|^2\, \langle\, x, x \,\rangle} = |\,c\,| \cdot \|x\|$.
2. $\|x\| = \sqrt{\langle\, x, x \,\rangle}$. If $x = 0$, then $\langle\, x, x \,\rangle = \langle\, 0, 0 \,\rangle = 0$. Otherwise, $\langle\, x, x \,\rangle > 0$ and hence $\|x\| \geq 0$.
3. If $y = 0$, then the Cauchy-Schwarz Inequality clearly hold. Assume now that $y \neq 0$.

For any $c \in \mathbb{F}$, we have

$$\begin{aligned} 0 \le \|x - cy\|^2 = \langle x - cy, x - cy \rangle &= \langle x, x - cy \rangle - c\langle y, x - cy \rangle \\ &= \langle x, x \rangle - \bar{c}\langle x, y \rangle - c\langle y, x \rangle + c\bar{c}\langle y, y \rangle. \end{aligned}$$

Let $c = \dfrac{\langle x, y \rangle}{\langle y, y \rangle}$, then

$$\begin{aligned} 0 &\le \langle x, x \rangle - \frac{\langle y, x \rangle}{\langle y, y \rangle}\langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle}\langle y, x \rangle + \frac{\langle x, y \rangle}{\langle y, y \rangle} \cdot \frac{\langle y, x \rangle}{\cancel{\langle y, y \rangle}}\cancel{\langle y, y \rangle} \\ 0 &\le \langle x, x \rangle - \frac{\langle y, x \rangle}{\langle y, y \rangle}\langle x, y \rangle - \cancel{\frac{\langle x, y \rangle}{\langle y, y \rangle}\langle y, x \rangle} + \cancel{\frac{\langle x, y \rangle}{\langle y, y \rangle} \cdot \langle y, x \rangle} \\ 0 &\le \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}, \quad \text{where } \langle x, y \rangle\langle y, x \rangle = \langle x, y \rangle\overline{\langle x, y \rangle} = |\langle x, y \rangle|^2. \end{aligned}$$

Therefore, $|\langle x, y \rangle|^2 \le \|x\|^2 \|y\|^2$ and hence $|\langle x, y \rangle| \le \|x\| \|y\|$.

4. Consider $\|x + y\|^2 = \langle x + y, x + y \rangle$. Then

$$\begin{aligned} \|x + y\|^2 &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2 \\ &= \|x\|^2 + 2Re\langle x, y \rangle + \|y\|^2, \quad \text{where } Re\langle x, y \rangle \le |\langle x, y \rangle| \\ &\le \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Therefore, $\|x + y\| \le \|x\| + \|y\|$.

Definition 6.1.4

If $x \neq 0$ is any vector in an inner product space $\mathbb{V}$, then $u = \frac{1}{\|x\|}x$ is a **unit vector**; that is a vector with length 1. This procedure is called **normalizing**.

Definition 6.1.5

Two vectors x and y in $\mathbb{V}$ are called **orthogonal** (or **perpendicular**) if $\langle x, y \rangle = 0$. Moreover, x and y are called **orthonormal** if they are orthogonal and $\|x\| = \|y\| = 1$.

Example 6.1.5

Note that the set $S = \{\, (1,1,0), (1,-1,1), (-1,1,2) \,\}$ in $\mathbb{F}^3$ is an orthogonal set of nonzero vectors, but it is not orthonormal. However, normalizing S, we get

$$B = \left\{ \frac{1}{\sqrt{2}}(1,1,0), \frac{1}{\sqrt{3}}(1,-1,1), \frac{1}{\sqrt{6}}(-1,1,2) \right\},$$

which is orthonormal in $\mathbb{F}^3$.

Example 6.1.6

Let H be the vector space of complex valued functions defined on the interval $[0, 2\pi]$, with the inner product on H defined by

$$\langle\, f, g \,\rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)\overline{g(t)}\; dt.$$

Show that $S = \left\{ f_n(t) = e^{int} \;:\; n \in \mathbb{Z} \text{ and } t \in [0, 2\pi] \right\}$ is an orthonormal subset of H. Recall that $e^{ix} = \cos x + i \sin x$, $\overline{e^{ix}} = e^{-ix}$ for all $x \in \mathbb{R}$, and $\int e^{ax} dx = \frac{1}{a} e^{ax}$.

Solution:

For any $m \neq n$ in $\mathbb{Z}$, we have

$$\begin{aligned} \langle\, f_m(t), f_n(t) \,\rangle &= \frac{1}{2\pi} \int_0^{2\pi} f_m(t)\overline{f_n(t)}\; dt = \frac{1}{2\pi} \int_0^{2\pi} e^{imt}\; \overline{e^{int}}\; dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)t}\; dt = \frac{1}{2\pi i}\, \frac{1}{(m-n)} e^{i(m-n)t}\Big|_0^{2\pi} \\ &= \frac{1}{2\pi i(m-n)} \left[e^{i(m-n)2\pi} - e^0 \right] = \frac{1}{2\pi i(m-n)} [1-1] = 0. \end{aligned}$$

Also, $\langle\, f_n, f_n \,\rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{i(n-n)t}\; dt = \frac{1}{2\pi} \int_0^{2\pi} 1\; dt = \frac{1}{2\pi}(2\pi - 0) = \frac{2\pi}{2\pi} = 1.$ Therefore, S is orthonormal subset of H.

Example 6.1.7

Let $\mathbb{V} = \mathbb{C}^3$ with the standard inner product. Let $x = (2, 1+i, i)$ and $y = (2-i, 2, 1+2i)$.

1. Compute $\langle\, x, y \,\rangle$, $\langle\, y, x \,\rangle$, $\|x\|$, $\|y\|$, and $\|x+y\|$.
2. Verify both Cauchy-Schwarz Inequality and triangle inequality.

Solution:

1.

$$\begin{aligned}\langle x, y \rangle &= \sum_{i=1}^{3} x_i \overline{y_i} = 2(\overline{2-i}) + (1+i)(\overline{2}) + i(\overline{1+2i}) \\ &= 2(2+i) + 2 + 2i + i(1-2i) = 4 + 2i + 2 + 2i + i + 2 \\ &= 8 + 5i.\end{aligned}$$

Thus, $\langle y, x \rangle = \overline{\langle x, y \rangle} = \overline{8+5i} = 8 - 5i$. Also

$$\begin{aligned}\|x\| = \sqrt{\langle x, x \rangle} &= \sqrt{2(\overline{2}) + (1+i)(\overline{1+i}) + i(\overline{i})} \\ &= \sqrt{4 + (1+i)(1-i) + i(-i)} = \sqrt{4 + 1 - i + i + 1 + 1} = \sqrt{7}.\end{aligned}$$

$$\begin{aligned}\|y\| = \sqrt{\langle y, y \rangle} &= \sqrt{(2-i)(\overline{2-i}) + 2(\overline{2}) + (1+2i)(\overline{1+2i})} \\ &= \sqrt{(2-i)(2+i) + 4 + (1+2i)(1-2i)} = \sqrt{4+1+4+1+4} = \sqrt{14}.\end{aligned}$$

$$\begin{aligned}\|x+y\| &= \|(4-i, 3+i, 1+3i\| \\ &= \sqrt{(4-i)(4+i) + (3+i)(3-i) + (1+3i)(1-3i)} \\ &= \sqrt{16+1+9+1+1+9} = \sqrt{37}.\end{aligned}$$

2. Clearly, Cauchy-Schwarz Inequality is satisfied as

$$|\langle x, y \rangle| = \sqrt{64+25} = \sqrt{89} \leq \sqrt{7}\sqrt{14} = \sqrt{98}.$$

For triangle inequality, note that

$$\|x+y\| = \sqrt{37} \leq \|x\| + \|y\| = \sqrt{7} + \sqrt{14}.$$

Since

$$\begin{aligned}(\|x\| + \|y\|)^2 &= (\sqrt{7} + \sqrt{14})^2 = 7 + 2\sqrt{98} + 14 \\ &= 21 + 2\sqrt{98} \geq 21 + 2\sqrt{81} = 21 + 2 \cdot 9 = 39 \\ &\geq 37 = \|x+y\|^2.\end{aligned}$$

Exercise 6.1.1

Solve the following exercises from the book at pages 336 - 341:

- $2, 3$.
- 8 : a and c.
- 9.

Section 6.2: The Gram-Schmidt Orthogonalization Process

Definition 6.2.1

Let $\mathbb{V}$ be an inner product space. A subset of $\mathbb{V}$ is called an **orthonormal basis** for $\mathbb{V}$ if it is an ordered basis that is orthonormal.

Example 6.2.1

- The standard ordered basis for $\mathbb{F}^n$ is orthonormal basis for $\mathbb{F}^n$.
- $S = \left\{ \frac{1}{\sqrt{5}}(1,2), \frac{1}{\sqrt{5}}(2,-1) \right\}$ is an orthonormal basis for $\mathbb{R}^2$.

Theorem 6.2.1

Let $\mathbb{V}$ be an inner product space and let $S = \{ x_1, x_2, \cdots, x_k \}$ be an orthogonal subset of $\mathbb{V}$ consisting of nonzero vectors. If $y \in$ **span** S, then

$$y = \sum_{i=1}^{k} \frac{\langle y, x_i \rangle}{\|x_i\|^2} x_i.$$

Proof:

Write $y = \sum_{i=1}^{k} a_i x_i$, where $a_1, \cdots, a_k \in \mathbb{F}$. Then, for $1 \le j \le k$

$$\begin{aligned} \langle y, x_j \rangle &= \langle \sum_{i=1}^{k} a_i x_i, x_j \rangle = \sum_{i=1}^{k} a_i \langle x_i, x_j \rangle, \quad \text{where } \langle x_i, x_j \rangle = 0 \text{ if } i \neq j \\ &= a_j \langle x_j, x_j \rangle = a_j \|x_j\|^2. \end{aligned}$$

So, $a_j = \frac{\langle y, x_j \rangle}{\|x_j\|^2}$. Therefore,

$$y = \sum_{i=1}^{k} a_i\, x_i = \sum_{i=1}^{k} \frac{\langle y, x_i \rangle}{\|x_i\|^2}\, x_i.$$

Corollary 6.2.1

Let $\mathbb{V}$ be an inner product space and let $S = \{ x_1, x_2, \cdots, x_k \}$ be an orthonormal subset of $\mathbb{V}$. If $y \in$ **span** S, then $y = \sum_{i=1}^{k} \langle y, x_i \rangle\, x_i$.

Corollary 6.2.2

Let $\mathbb{V}$ be an inner product space and let $S = \{x_1, x_2, \cdots, x_k\}$ be an orthogonal subset of $\mathbb{V}$ consisting of nonzero vectors. Then, S is linearly independent.

Proof:

Suppose that $a_1x_1 + \cdots a_kx_k = \sum_{i=1}^{k} a_ix_i = 0$. Then for all $1 \leq j \leq k$, we have

$$\langle 0, x_j \rangle = \langle \sum_{i=1}^{k} a_ix_i, x_j \rangle = \sum_{i=1}^{k} a_i \langle x_i, x_j \rangle = a_j \langle x_j, x_j \rangle = a_j \|x_j\|^2 .$$

Thus, $a_j = \dfrac{\langle 0, x_j \rangle}{\|x_j\|^2} = 0$ for all j. So, S is linearly independent.

Theorem 6.2.2: The Gram-Schmidt Process

Let $\mathbb{V}$ be an inner product space and $S = \{y_1, y_2, \cdots, y_n\}$ be linearly independent subset of $\mathbb{V}$. Define $S' = \{x_1, x_2, \cdots, x_n\}$, where $x_1 = y_1$ and

$$x_k = y_k - \sum_{j=1}^{k-1} \frac{\langle y_k, x_j \rangle}{\|x_j\|^2} x_j, \quad \text{for} \ \ 2 \leq j \leq n.$$

Then, S' is an orthogonal set of nonzero vectors such that **span** $S' =$ **span** S.

Theorem 6.2.3

Let $\mathbb{V}$ be a nonzero finite-dimensional inner product space. Then $\mathbb{V}$ has an orthonormal basis β. Furthermore, if $\beta = \{x_1, x_2, \cdots, x_n\}$ and $y \in \mathbb{V}$, then

$$y = \sum_{i=1}^{n} \langle y, x_i \rangle x_i.$$

That is $[y]_\beta = (\langle y, x_1 \rangle, \langle y, x_2 \rangle, \cdots, \langle y, x_n \rangle)$. These scalars are called **Fourier coefficients**.

Corollary 6.2.3

Let $\mathbf{T}$ be a linear operator on a finite-dimensional inner product space $\mathbb{V}$ with an orthonorml basis $\beta = \{x_1, x_2, \cdots, x_n\}$, and let $A = [\mathbf{T}]_\beta = (a_{ij})$. Then, for any i and j, $a_{ij} = \langle \mathbf{T}(x_j), x_i \rangle$.

Example 6.2.2

Let $S = \left\{ \frac{1}{\sqrt{2}}(1,1,0), \frac{1}{\sqrt{3}}(1,-1,1), \frac{1}{\sqrt{6}}(-1,1,2) \right\}$ be an orthonormal basis for $\mathbb{R}^3$. Express $x = (2,1,3) \in \mathbb{R}^3$ as a linear combination of vectors of S.

Solution:

Consider $x = (2,1,3) = c_1 \frac{1}{\sqrt{2}}(1,1,0) + c_2 \frac{1}{\sqrt{3}}(1,-1,1) + c_3 \frac{1}{\sqrt{6}}(-1,1,2)$. Then,

$$\begin{aligned} c_1 &= \langle (2,1,3), \frac{1}{\sqrt{2}}(1,1,0) \rangle = \frac{1}{\sqrt{2}}(2+1+0) = \frac{3}{\sqrt{2}}. \\ c_2 &= \langle (2,1,3), \frac{1}{\sqrt{3}}(1,-1,1) \rangle = \frac{1}{\sqrt{3}}(2-1+3) = \frac{4}{\sqrt{3}}. \\ c_3 &= \langle (2,1,3), \frac{1}{\sqrt{6}}(-1,1,2) \rangle = \frac{1}{\sqrt{6}}(-2+1+6) = \frac{5}{\sqrt{6}}. \end{aligned}$$

Thus $x = (2,1,3) = \frac{3}{2}(1,1,0) + \frac{4}{3}(1,-1,1) + \frac{5}{6}(-1,1,2)$.

Example 6.2.3

Use the Gram-Schmidt process to find an orthonormal basis for **span** S, where

$$S = \{ y_1 = (1,0,1,0), y_2 = (1,1,1,1), y_3 = (0,1,2,1) \}$$

is a subset of $\mathbb{R}^4$.

Solution:

We first compute S' containing orthogonal vectors x_1, x_2, x_3 and then we normalize these vectors to obtain an orthonormal set S''.

- $x_1 = y_1 = (1,0,1,0)$.
- $x_2 = y_2 - \frac{\langle y_2, x_1 \rangle}{\|x_1\|^2} x_1$, where $\|x_1\|^2 = \left(\sqrt{2}\right)^2 = 2$, and $\langle y_2, x_1 \rangle = 1+0+1+0 = 2$. Then $x_2 = y_2 - \frac{2}{2}x_1 = (0,1,0,1)$.
- $x_3 = y_3 - \left(\frac{\langle y, x_1 \rangle}{\|x_1\|^2} x_1 + \frac{\langle y, x_2 \rangle}{\|x_2\|^2} x_2 \right)$, where $\|x_2\|^2 = \left(\sqrt{2}\right)^2 = 2 = \|x_1\|^2$. Moreover, $\langle y_3, x_1 \rangle = 0+0+2+0 = 2$ and $\langle y_3, x_2 \rangle = 0+1+0+1 = 2$. Therefore, $x_3 = (0,1,2,1) - \frac{2}{2}(1,0,1,0) - \frac{2}{2}(0,1,0,1) = (-1,0,1,0)$.

Thus, by Theorem 6.2.2, $S' = \{ (1,0,1,0), (0,1,0,1), (-1,0,1,0) \}$ is orthogonal set in $\mathbb{R}^4$

such that **span** $S' =$ **span** S. Therefore,

$$S'' = \left\{ \frac{1}{\sqrt{2}}(1,0,1,0), \frac{1}{\sqrt{2}}(0,1,0,1), \frac{1}{\sqrt{2}}(-1,0,1,0) \right\}$$

is orthonormal set in $\mathbb{R}^4$.

Example 6.2.4

Let $\mathbb{V} = \mathbb{P}(\mathbb{R})$ with an inner product defined by $\langle f(x), g(x) \rangle = \int_{-1}^{1} f(t)g(t)\, dt$. Use the Gram-Schmidt process to replace the standard ordered basis $S = \{1, t, t^2\}$ by an orthonormal basis for $\mathbb{P}_2(\mathbb{R})$. Represent $h(x) = 1+2x+3x^2$ as a linear combination of the vectors of the obtained orthonormal basis for $\mathbb{P}_2(\mathbb{R})$.

Solution:

Let $S = \{y_1 = 1, y_2 = t, y_3 = t^2\}$. Then $S' = \{x_1, x_2, x_3\}$, where

- $x_1 = y_1 = 1$.
- $x_2 = y_2 - \frac{\langle y_2, x_1 \rangle}{\|x_1\|^2} x_1 = t - \frac{\langle t, 1 \rangle}{\|1\|^2} 1 = t - \langle t, 1 \rangle$. Note that

  $$\|1\|^2 = \langle 1, 1 \rangle = \int_{-1}^{1} 1\, dt = t\Big|_{-1}^{1} = 2,$$

  and

  $$\langle t, 1 \rangle = \int_{-1}^{1} t \cdot 1\, dt = \frac{t^2}{2}\Big|_{-1}^{1} = \frac{1}{2} - \frac{1}{2} = 0.$$

  Therefore, $x_2 = t - \frac{0}{2}\, 1 = t$.
- $x_3 = y_3 - \left(\frac{\langle y_3, x_1 \rangle}{\|x_1\|^2} x_1 + \frac{\langle y_3, x_2 \rangle}{\|x_2\|^2} x_2 \right) = t^2 - \frac{\langle t^2, 1 \rangle}{\|1\|^2} 1 - \frac{\langle t^2, t \rangle}{\|t\|^2} t.$
  Note that $\|1\|^2 = 2$ and $\|t\|^2 = \int_{-1}^{1} t^2\, dt = \frac{t^3}{3}\Big|_{-1}^{1} = \frac{2}{3}$. Moreover, $\langle t^2, 1 \rangle = \int_{-1}^{1} t^2\, dt = \frac{t^3}{3}\Big|_{-1}^{1} = \frac{2}{3}$, and $\langle t^2, t \rangle = \int_{-1}^{1} t^3\, dt = \frac{t^4}{4}\Big|_{-1}^{1} = 0$. Therefore,

  $$x_3 = t^2 - \frac{2/3}{2}\, 1 - \frac{0}{2/3}\, t = t^2 - \frac{1}{3}.$$

We now normalize S' to obtain $S'' = \left\{ \frac{1}{\|x_1\|}x_1, \frac{1}{\|x_2\|}x_2, \frac{1}{\|x_3\|}x_3 \right\}$ as follows:

$$\|x_1\|^2 = \|1\|^2 = 2 \;\Rightarrow \|x_1\| = \sqrt{2}.$$

$$\|x_2\|^2 = \|t\|^2 = \langle\, t, t\,\rangle = \frac{2}{3} \;\Rightarrow\; \|x_2\| = \sqrt{\frac{2}{3}}.$$

$$\|x_3\|^2 = \left\|t^2 - \frac{1}{3}\right\|^2 = \langle\, t^2 - \frac{1}{3}, t^2 - \frac{1}{3}\,\rangle = \int_{-1}^{1}\left(t^2 - \frac{1}{3}\right)^2 dt$$

$$= \int_{-1}^{1} t^4 - \frac{2}{3}t^2 + \frac{1}{9}\; dt = \left[\frac{t^5}{5} - \frac{2}{3}\frac{t^3}{3} + \frac{1}{9}t\right]_{-1}^{1} = \cdots = \frac{8}{45}$$

$$\Rightarrow \|x_3\| = \sqrt{\frac{8}{45}} = \frac{2\sqrt{2}}{3\sqrt{5}}.$$

Thus, $S'' = \left\{ z_1 = \frac{1}{\sqrt{2}}\,1, z_2 = \sqrt{\frac{3}{2}}\,t, z_3 = \frac{3\sqrt{5}}{2\sqrt{2}}\left(t^2 - \frac{1}{3}\right) \right\}$ is orthonormal basis for $\mathbb{P}_2(\mathbb{R})$.
We now use Theorem 6.2.3 to represent $h(x)$ as a linear combination of the vectors of S''.
Note that

$$\langle\, h(x), z_1\,\rangle = \int_{-1}^{1} \frac{1}{\sqrt{2}}(1 + 2t + 3t^2)dt = 2\sqrt{2}$$

$$\langle\, h(x), z_2\,\rangle = \int_{-1}^{1} \sqrt{\frac{3}{2}}t(1 + 2t + 3t^2)dt = \frac{2\sqrt{6}}{3}$$

$$\langle\, h(x), z_3\,\rangle = \int_{-1}^{1} \sqrt{\frac{5}{8}}(3t^2 - 1)(1 + 2t + 3t^2)dt = \frac{2\sqrt{10}}{5}$$

Therefore, $h(x) = 2\sqrt{2}z_1 + \frac{2\sqrt{6}}{3}z_2 + \frac{2\sqrt{10}}{5}z_3$.

Example 6.2.5

Let $\mathbb{W} = \textbf{span}\,\{\,(1,1,1),(1,0,2)\,\}$ be a subspace of $\mathbb{R}^3$. Find an orthonormal basis for $\mathbb{W}$.

Solution:

Consider $x_1 = (1, 1, 1)$ and

$$x_2 = (1, 0, 2) - \frac{\langle\,(1, 0, 2), (1, 1, 1)\,\rangle}{\|(1, 1, 1)\|^2}(1, 1, 1) = (1, 0, 2) - \frac{3}{3}(1, 1, 1) = (0, -1, 1).$$

Thus, $S' = \left\{ \frac{1}{\sqrt{3}}(1, 1, 1), \frac{1}{\sqrt{2}}(0, -1, 1) \right\}$ is an orthonormal basis for $\mathbb{W}$.

Example 6.2.6

Let $\mathbb{W} = \{\,(x, y, z)\, :\, x + 3y - 2z = 0\,\}$ be a subspace of the inner product space $\mathbb{R}^3$. Find an orthonormal basis for $\mathbb{W}$.

Solution:

Note that

$$\mathbb{W} = \{\,(2z - 3y, y, z)\,\} = \textbf{span}\ \{\,(2s - 3r, r, s) : r, s \in \mathbb{R}\,\} = \textbf{span}\ \{\,(2, 0, 1), (-3, 1, 0)\,\},$$

where $S = \{\,(2, 0, 1), (-3, 1, 0)\,\}$ is an ordered basis for $\mathbb{W}$. We now construct orthogonal basis for $\mathbb{W}$ and normalize it to an orthonormal basis. Let $x_1 = (2, 0, 1)$ and

$$\begin{aligned} x_2 &= (-3, 1, 0) - \frac{\langle\,(-3, 1, 0), (2, 0, 1)\,\rangle}{\|(2, 0, 1)\|^2}(2, 0, 1) \\ &= (-3, 1, 0) - \frac{-6}{5}(2, 0, 1) = (-\frac{3}{5}, 1, \frac{6}{5}). \end{aligned}$$

Thus, $\|x_2\| = \sqrt{\frac{9}{25} + \frac{25}{25} + \frac{36}{25}} = \frac{\sqrt{70}}{5}$. Thus, $S' = \left\{\,\frac{1}{\sqrt{5}}(2, 0, 1), \frac{1}{\sqrt{70}}(-3, 5, 6)\,\right\}$ is an orthonormal basis for $\mathbb{W}$.

Exercise 6.2.1

Solve the following exercises from the book at pages 352 - 357:

- 2 : a, b, c, g, and h.

Section 6.3: The Adjoint of a Linear Operator

Recall that A^* is the conjugate transpose of a matrix. In this section, for a linear operator $\mathbf{T}$ on an inner product space $\mathbb{V}$, we define a related linear operator on $\mathbb{V}$ called the **adjoint** of $\mathbf{T}$, denoted $\mathbf{T}^*$, whose matrix representation with respect to any orthonormal basis β for $\mathbb{V}$ is $[\mathbf{T}]_\beta^*$.

Definition 6.3.1

Let $\mathbb{V}$ be a finite-dimensional inner product space , and let $\mathbf{T}$ be a linear operator on $\mathbb{V}$. The **adjoint** (sometimes called **hermitian conjugate**) of $\mathbf{T}$ is the unique linear operator $\mathbf{T}^*$ on $\mathbb{V}$ such that

$$\langle \mathbf{T}(x), y \rangle = \langle x, \mathbf{T}^*(y) \rangle, \qquad \text{for all } x, y \in \mathbb{V}.$$

Remark 6.3.1

Note that

$$\langle x, \mathbf{T}(y) \rangle = \overline{\langle \mathbf{T}(y), x \rangle} = \overline{\langle y, \mathbf{T}^*(x) \rangle} = \langle \mathbf{T}^*(x), y \rangle.$$

Theorem 6.3.1

Let $\mathbb{V}$ be a finite-dimensional inner product space , let β be an orthonormal basis for $\mathbb{V}$, and let $\mathbf{T}$ and $\mathbf{U}$ be linear operators on $\mathbb{V}$. Then:

1. $\mathbf{T}^*$ is unique linear operator on $\mathbb{V}$.
2. $[\mathbf{T}^*]_\beta = [\mathbf{T}]_\beta^*$.
3. $(\mathbf{T} + \mathbf{U})^* = \mathbf{T}^* + \mathbf{U}^*$, and $(\mathbf{T}\mathbf{U})^* = \mathbf{U}^*\mathbf{T}^*$.
4. $(c\mathbf{T})^* = \overline{c}\mathbf{T}^*$.
5. $(\mathbf{T}^*)^* = \mathbf{T}$.
6. $\mathbf{I}_V{}^* = \mathbf{I}_V$.

Example 6.3.1

Let $\mathbf{T}$ be the linear operator on $\mathbb{C}^2$ defined by $\mathbf{T}(a, b) = (2ai + 3b, a - b)$. Evaluate $\mathbf{T}^*$.

Solution:

We can find $\mathbf{T}^*$ directly by the definition:

$$\begin{aligned}\langle (a,b), \mathbf{T}^*(c,d) \rangle &= \langle \mathbf{T}(a,b), (c,d) \rangle = \langle (2ai+3b, a-b), (c,d) \rangle \\ &= (2ai+3b)\overline{c} + (a-b)\overline{d} = 2a\overline{c}i + 3b\overline{c} + a\overline{d} - b\overline{d} \\ &= a\left(2i\overline{c} + \overline{d}\right) + b\left(3\overline{c} - \overline{d}\right) = \langle (a,b), (-2ci+d, 3c-d) \rangle.\end{aligned}$$

Therefore, $\mathbf{T}^*(c,d) = (-2ci+d, 3c-d)$.

On the other hand, we can also find $\mathbf{T}^*$ using the Theorem 6.3.1. Choose β as the standard orthonormal basis for $\mathbb{C}^2$. Clearly, $[\mathbf{T}]_\beta = \begin{pmatrix} 2i & 3 \\ 1 & -1 \end{pmatrix}$. Then, $[\mathbf{T}^*]_\beta = [\mathbf{T}]_\beta^* = \begin{pmatrix} -2i & 1 \\ 3 & -1 \end{pmatrix}$. Hence, $\mathbf{T}^*(a,b) = (-2ai+b, 3a-b)$.

Example 6.3.2

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^2$ defined by $\mathbf{T}(a,b) = (2a+b, a-3b)$. Evaluate $\mathbf{T}^*$ at $x = (3,5)$.

Solution:

We can find $\mathbf{T}^*(3,5)$ directly by the definition:

$$\begin{aligned}\langle (a,b), \mathbf{T}^*(3,5) \rangle &= \langle \mathbf{T}(a,b), (3,5) \rangle = \langle (2a+b, a-3b), (3,5) \rangle \\ &= (6a+3b) + 5a - 15b = 11a - 12b \\ &= \langle (a,b), (11,-12) \rangle.\end{aligned}$$

Therefore, $\mathbf{T}^*(3,5) = (11,-12)$.

On the other hand, we can also find $\mathbf{T}^*(3,5)$ using the Remark 2.2.2. Choose β as an orthonormal basis for $\mathbb{R}^2$. Clearly, $[\mathbf{T}]_\beta = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}$. Then, $[\mathbf{T}^*]_\beta = [\mathbf{T}]_\beta^* = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}$, and $[(3,5)]_\beta = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$. Hence,

$$[\mathbf{T}^*(3,5)]_\beta = [\mathbf{T}]_\beta^*[(3,5)]_\beta = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 11 \\ -12 \end{pmatrix}.$$

Therefore, $\mathbf{T}^*(3,5) = (11,-12)$.

Example 6.3.3

Let $\mathbf{T}$ be the linear operator on $\mathbb{P}_1(\mathbb{R})$ defined by $\mathbf{T}(f) = f' + 3f$ with $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)dt$. Evaluate $\mathbf{T}^*$ at $f(x) = 4 - 2x$. Get 1 bonus point when you evaluate $\mathbf{T}^*(h(x))$, where $h(x) = a + bx \in \mathbb{P}_1(\mathbb{R})$. Hand it over to me at my office.

Solution (1):

Using the definition: Let $g(x) = a + bx$ for $a, b \in \mathbb{R}$. Then, $\mathbf{T}(g) = b + 3a + abx$.

$$\begin{aligned}\langle g, \mathbf{T}^*(f) \rangle &= \langle \mathbf{T}(g), f \rangle = \langle (3a + b + 3bx), (4 - 2x) \rangle \\ &= \int_{-1}^{1} (3a + b + 3bx)(4 - 2x) = \cdots = 24a + 4b.\end{aligned}$$

Assuming that $\mathbf{T}^*(f) = c + dx$, we get:

$$\begin{aligned}\langle g, \mathbf{T}^*(f) \rangle &= \langle (a + bx), (c + dx) \rangle \\ &= \int_{-1}^{1} (a + bx)(c + dx) = \cdots = 2ac + \frac{2}{3}bd.\end{aligned}$$

By equating the two results, we get $c = 12$ and $d = 6$ and hence $\mathbf{T}^*(f = 4 - 2x) = 12 + 6x$.

Solution (2):

We can find $\mathbf{T}^*(f)$ using the Remark 2.2.2. Choose $\beta = \left\{ v_1 = \frac{1}{\sqrt{2}}, v_2 = \sqrt{\frac{3}{2}}x \right\}$ as an orthonormal basis for $\mathbb{P}_1(\mathbb{R})$ (Use Gram-Schmidt process to find such basis). Then,

$$\begin{aligned}\mathbf{T}(v_1) &= 3\frac{1}{\sqrt{2}} = 3v_1 + 0v_2 \quad \Rightarrow \quad [\mathbf{T}(v_1)]_\beta = (3, 0). \\ \mathbf{T}(v_2) &= \sqrt{\frac{3}{2}} + 3\sqrt{\frac{3}{2}}x = \sqrt{3}v_1 + 3v_2 \quad \Rightarrow \quad [\mathbf{T}(v_2)]_\beta = (\sqrt{3}, 3).\end{aligned}$$

Hence $[\mathbf{T}]_\beta = \begin{pmatrix} 3 & \sqrt{3} \\ 0 & 3 \end{pmatrix}$ and thus $[\mathbf{T}]_\beta^* = \begin{pmatrix} 3 & 0 \\ \sqrt{3} & 3 \end{pmatrix}$. Furthermore, observe that $[f(x)]_\beta = \left(\langle 4 - 2x, \frac{1}{\sqrt{2}} \rangle, \langle 4 - 2x, \sqrt{\frac{3}{2}}x \rangle \right) = \left(4\sqrt{2}, -2\sqrt{\frac{2}{3}} \right)$. Therefore,

$$[\mathbf{T}(f(x))]_\beta^* = \begin{pmatrix} 3 & 0 \\ \sqrt{3} & 3 \end{pmatrix} \begin{pmatrix} 4\sqrt{2} \\ -2\sqrt{\frac{2}{3}} \end{pmatrix} = \begin{pmatrix} 12\sqrt{2} \\ 2\sqrt{6} \end{pmatrix}.$$

That is, $\mathbf{T}^*(4 - 2x) = 12\sqrt{2}v_1 + 2\sqrt{6}v_2 = 12 + 6x$. In the general case when $h(x) = a + bx$, we use the matrix multiplication since using the definition is rather difficult. Observe that $[h(x)]_\beta = \left(a\sqrt{2}, b\sqrt{\frac{2}{3}} \right)$. Hence

$$[\mathbf{T}(h(x))]_\beta^* = [\mathbf{T}]_\beta^* [h]_\beta = \begin{pmatrix} 3 & 0 \\ \sqrt{3} & 3 \end{pmatrix} \begin{pmatrix} a\sqrt{2} \\ b\sqrt{\frac{2}{3}} \end{pmatrix} = \begin{pmatrix} 3a\sqrt{2} \\ a\sqrt{6} + b\sqrt{6} \end{pmatrix}.$$

That is, $\mathbf{T}^*(a+bx) = \left(3a\sqrt{2}\right)v_1 + \sqrt{6}(\,a+b\,)v_2 = 3a + 3(\,a+b\,)x$.

Example 6.3.4

Let $\mathbb{V}$ be an inner product space , and let $y, z \in \mathbb{V}$. Define $\mathbf{T} : \mathbb{V} \to \mathbb{V}$ by $\mathbf{T}(x) = \langle\, x, y\,\rangle z$ for all $x \in \mathbb{V}$. Show that $\mathbf{T}$ is linear, and evaluate $\mathbf{T}^*(x)$.

Solution:

We first show that $\mathbf{T}$ is linear. For any $x_1, x_2 \in \mathbb{V}$ and any $c \in \mathbb{F}$.

$$\begin{aligned}\mathbf{T}(cx_1 + x_2) &= \langle\, cx_1 + x_2, y\,\rangle z = \langle\, cx_1, y\,\rangle z + \langle\, x_2, y\,\rangle z \\ &= c\langle\, x_1, y\,\rangle z + \langle\, x_2, y\,\rangle z = c\mathbf{T}(x_1) + \mathbf{T}(x_2)\,.\end{aligned}$$

Hence, $\mathbf{T}$ is linear. Furthermore,

$$\begin{aligned}\langle\, u, \mathbf{T}^*(x)\,\rangle &= \langle\, \mathbf{T}(u)\,, x\,\rangle = \langle\, \langle\, u, y\,\rangle z, x\,\rangle \\ &= \langle\, u, y\,\rangle\langle\, z, x\,\rangle = \langle\, u, \overline{\langle\, z, x\,\rangle} y\,\rangle = \langle\, u, \langle\, x, z\,\rangle y\,\rangle.\end{aligned}$$

Therefore, $\mathbf{T}^*(x) = \langle\, x, z\,\rangle y$.

Exercise 6.3.1

Solve the following exercises from the book at pages 352 - 357:

- $2 : a, b, c, g$, and h.

Section 6.4: Self-Adjoint, Normal, and Unitary Operators

In this section, we present more properties of special linear operators. Furthermore, we consider the diagonalization problem for these operators.

Definition 6.4.1

Let $\mathbb{V}$ be an inner product space , and let $\mathbf{T}$ be a linear operator on $\mathbb{V}$. Then:

1. $\mathbf{T}$ is called **self-adjoint** (**Hermitian**) if $\mathbf{T} = \mathbf{T}^*$.
2. An $n \times n$-real or complex matrix A is called **self-adjoint** (**Hermitian**) matrix if $A = A^*$.
3. $\mathbf{T}$ is called **normal** if $\mathbf{T}\mathbf{T}^* = \mathbf{T}^*\mathbf{T}$.
4. An $n \times n$-real or complex matrix A is called **normal** matrix if $AA^* = A^*A$.

Remark 6.4.1

If $\mathbf{T}$ is a linear operator on an inner product space $\mathbb{V}$ and β is an orthonormal basis for $\mathbb{V}$, then:

1. $\mathbf{T}$ is self-adjoint if and only if $[\mathbf{T}]_\beta$ is self-adjoint.
2. $\mathbf{T}$ is normal if and only if $[\mathbf{T}]_\beta$ is normal.
3. If $\mathbf{T}$ is self-adjoint, then $\mathbf{T}$ is normal.

Theorem 6.4.1

Let $\mathbb{V}$ be an inner product space , and let $\mathbf{T}$ be a **normal operator** on $\mathbb{V}$. Then:

1. $\|\mathbf{T}(x)\| = \|\mathbf{T}^*(x)\|$ for all $x \in \mathbb{V}$.
2. If x is an eigenvector of $\mathbf{T}$, then x is an eigenvector of $\mathbf{T}^*$. In fact, if $\mathbf{T}(x) = \lambda x$, then $\mathbf{T}^*(x) = \overline{\lambda} x$.
3. If λ_1 and λ_2 are distinct eigenvalues of $\mathbf{T}$ with corresponding eigenvectors x_1 and x_2, respectively, then x_1 and x_2 are orthogonal.

Proof:

1. For any vector $x \in \mathbb{V}$, we have:

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

Therefore, $\|\mathbf{T}(x)\| = \|\mathbf{T}^*(x)\|$.

2. Observe that for any $c \in \mathbb{F}$, $(\mathbf{T} - cI)^* = \mathbf{T}^* - \bar{c}I$ and that $(\mathbf{T} - cI)$ is normal as $\mathbf{T}$ normal (prove it!). Now assume that for some $x \in \mathbb{V}$, $\mathbf{T}(x) = \lambda x$. Then $(\mathbf{T} - \lambda I)(x) = 0$, where $\mathbf{T} - \lambda I$ is normal.

   Then, by (1), we have

$$\begin{aligned}0 &= \|(\mathbf{T} - \lambda I)(x)\| = \|(\mathbf{T} - \lambda I)^*(x)\| \\ &= \left\|\left(\mathbf{T}^* - \overline{\lambda} I\right)(x)\right\| = \left\|\mathbf{T}^*(x) - \overline{\lambda} x\right\| = \left\|\left(\mathbf{T}^* - \overline{\lambda} I\right)(x)\right\| .\end{aligned}$$

   Hence, $\mathbf{T}^*(x) = \overline{\lambda} x$.

3. Let $\lambda_1 \neq \lambda_2$ be two eigenvalues of $\mathbf{T}$ with corresponding eigenvectors x_1 and x_2. Then, by part (2), we have:

$$\begin{aligned}(\lambda_1 - \lambda_2)\langle x_1, x_2 \rangle &= \langle \lambda_1 x_1, x_2 \rangle - \langle x_1, \overline{\lambda_2} x_2 \rangle \\ &= \langle \mathbf{T}(x_1), x_2 \rangle - \langle x_1, \mathbf{T}^*(x_2) \rangle = 0.\end{aligned}$$

   But since $\lambda_1 - \lambda_2 \neq 0$, then $\langle x_1, x_2 \rangle = 0$.

Theorem 6.4.2

Let $\mathbf{T}$ be a linear operator on a finite-dimensional inner product space $\mathbb{V}$ over $\mathbb{C}$. Then $\mathbf{T}$ is normal if and only if there exists an orthonormal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$.

Theorem 6.4.3

Let $\mathbf{T}$ be a self-adjoint linear operator on a finite-dimensional inner product space $\mathbb{V}$. Then every eigenvalues of $\mathbf{T}$ is real.

Proof:

Assume that $\mathbf{T}(x) = \lambda x$ for $x \neq 0$. Then

$$\lambda x = \mathbf{T}(x) = \mathbf{T}^*(x) = \overline{\lambda} x.$$

Therefore, $\lambda = \overline{\lambda}$ and hence λ is real.

Theorem 6.4.4

Let $\mathbf{T}$ be a linear operator on a finite-dimensional inner product space over $\mathbb{R}$. Then, $\mathbf{T}$ is self-adjoint if and only if there exists an orthonormal basis β for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$.

Lemma 6.4.1

Let $\mathbf{T}$ be a self-adjoint operator on a finite-dimensional inner product space $\mathbb{V}$. If $\langle x, \mathbf{T}(x) \rangle = 0$, for all $x \in \mathbb{V}$, then $\mathbf{T} = \mathbf{T}_0$.

Proof:

Choose an orthonormal basis β for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$. If $x \in \beta$, then $\mathbf{T}(x) = \lambda x$ for some λ. Then

$$0 = \langle x, \mathbf{T}(x) \rangle = \langle x, \lambda x \rangle = \overline{\lambda} \langle x, x \rangle.$$

So, $\overline{\lambda} = 0$. Hence $\mathbf{T}(x) = 0$ for all $x \in \beta$, and thus $\mathbf{T} = \mathbf{T}_0$.

Definition 6.4.2

Let $\mathbf{T}$ be a linear operator on a finite-dimensional inner product space $\mathbb{V}$ over $\mathbb{F}$. If $\|\mathbf{T}(x)\| = \|x\|$ for all $x \in \mathbb{V}$, we call $\mathbf{T}$ a unitary operator if $\mathbb{F} = \mathbb{C}$ and an **orthogonal** operator if $\mathbb{F} = \mathbb{R}$. Moreover, a square matrix A is called an **orthogonal** matrix if $AA^T = A^T A = I$ and **unitary** matrix if $AA^* = A^* A = I$.

Remark 6.4.2

Note that, the condition $AA^* = I$ is equivalent to the statement that the rows of A form an orthonormal basis for $\mathbb{F}^n$. The same statement can be made on the columns of A and the condition $A^* A = I$.

Remark 6.4.3

A linear operator $\mathbf{T}$ on a inner product space $\mathbb{V}$ is unitary (orthogonal) if and only if $[\mathbf{T}]_\beta$ is unitary (orthogonal, respectively), for some orthonormal basis β for $\mathbb{V}$.

Theorem 6.4.5

Let $\mathbf{T}$ be a linear operator on a finite-dimensional inner product space $\mathbb{V}$. Then the following statements are equivalent:

1. $\mathbf{T}\mathbf{T}^* = \mathbf{T}^*\mathbf{T} = \mathbf{I}_V$.
2. $\langle \mathbf{T}(x), \mathbf{T}(y) \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{V}$.
3. If β is an orthonormal basis for $\mathbb{V}$, then $\mathbf{T}(\beta)$ is an orthonormal basis for $\mathbb{V}$.
4. $\|\mathbf{T}(x)\| = \|x\|$ for all $x \in \mathbb{V}$.

Proof:

We proof that each statement implies the following one as follows:

1. $1 \to 2$: Let $x, y \in \mathbb{V}$, then $\langle x, y \rangle = \langle \mathbf{T}^*\mathbf{T}(x), y \rangle = \langle \mathbf{T}(x), \mathbf{T}(y) \rangle$.
2. $2 \to 3$: Let $\beta = \{x_1, x_2, \cdots, x_n\}$ be an orthonormal basis for $\mathbb{V}$. So $\mathbf{T}(\beta) = \{\mathbf{T}(x_1), \mathbf{T}(x_2), \cdots, \mathbf{T}(x_n)\}$. It follows that $\langle \mathbf{T}(x_i), \mathbf{T}(x_j) \rangle = \langle x_i, x_j \rangle = \delta_{ij}$ where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. Hence, $\mathbf{T}(\beta)$ is an orthonormal basis for $\mathbb{V}$.
3. $3 \to 4$: Let $x \in \mathbb{V}$ and let $\beta = \{x_1, x_2, \cdots, x_n\}$ be an orthonormal basis for $\mathbb{V}$. Then $x = \sum_{i=1}^n a_i x_i$ for some scalars a_i and hence

$$\|x\|^2 = \langle x, x \rangle = \langle \sum_{i=1}^n a_i x_i, \sum_{j=1}^n a_j x_j \rangle = \sum_{i=1}^n \sum_{j=1}^n a_i \overline{a_j} \langle x_i, x_j \rangle$$
$$= \sum_{i=1}^n \sum_{j=1}^n a_i \overline{a_j} \delta_{ij} = \sum_{i=1}^n a_i \overline{a_i} = \sum_{i=1}^n |a_i|^2.$$

   In a similar way, $\mathbf{T}(x) = \sum_{i=1}^n a_i \mathbf{T}(x_i)$, and using the fact that $\mathbf{T}(\beta)$ is also orthonormal, we obtain $\|\mathbf{T}(x)\|^2 = \sum_{i=1}^n |a_i|^2$. Therefore, $\|\mathbf{T}(x)\| = \|x\|$.
4. $4 \to 1$: For any $x \in \mathbb{V}$,

$$\langle x, x \rangle = \|x\|^2 = \|\mathbf{T}(x)\|^2 = \langle \mathbf{T}(x), \mathbf{T}(x) \rangle = \langle x, \mathbf{T}^*\mathbf{T}(x) \rangle.$$

   Thus, $\mathbf{T}^*\mathbf{T}(x) = x$ and hence $(\mathbf{T}^*\mathbf{T} - \mathbf{I}_V)(x) = 0$. Thus, $\langle x, (\mathbf{T}^*\mathbf{T} - \mathbf{I}_V)(x) \rangle = 0$ for all $x \in \mathbb{V}$. Also, $(\mathbf{T}^*\mathbf{T} - \mathbf{I}_V)$ is clearly self-adjoint. By Lemma 6.4.1, we get $(\mathbf{T}^*\mathbf{T} - \mathbf{I}_V) = \mathbf{T}_0$ and therefore, $\mathbf{T}^*\mathbf{T} = \mathbf{I}_V$.

Definition 6.4.3

Two square matrices A and B are said to be **unitarily equivalent** (**orthogonally equivalent**) if and only if there exists a unitary (orthogonal, respectively) matrix P such that $A = P^*BP$.

Theorem 6.4.6

Let A be $n \times n$ matrix. Then:

1. If A is complex. Then, A is normal if and only if A is unitarily equivalent to a diagonal matrix.
2. If A is real. Then, A is symmetric if and only if A is orthogonally equivalent to a real diagonal matrix.

Example 6.4.1

Let $\mathbf{T}$ be a linear operator on an inner product space $\mathbb{V}$. Let $\mathbf{U}_1 = \mathbf{T} + \mathbf{T}^*$ and $\mathbf{U}_2 = \mathbf{T}\mathbf{T}^*$. Show that $\mathbf{U}_1$ and $\mathbf{U}_2$ are both self-adjoint.

Solution:

Clearly

$$\mathbf{U}_1^* = (\mathbf{T} + \mathbf{T}^*)^* = \mathbf{T}^* + (\mathbf{T}^*)^* = \mathbf{T}^* + \mathbf{T} = \mathbf{T} + \mathbf{T}^* = \mathbf{U}_1.$$

$$\mathbf{U}_2^* = (\mathbf{T}\mathbf{T}^*)^* = (\mathbf{T}^*)^*\mathbf{T}^* = \mathbf{T}\mathbf{T}^* = \mathbf{U}_2.$$

Example 6.4.2

Let $\mathbf{T}$ be a linear operator on $\mathbb{V} = \mathbb{R}^2$ defined by $\mathbf{T}(a, b) = (2a - 2b, -2a + 5b)$. Determine whether $\mathbf{T}$ is normal, self-adjoint, or neither. If possible, produce an orthonormal basis of eigenvectors of $\mathbf{T}$ for $\mathbb{V}$, and list the corresponding eigenvalues.

Solution:

Choose an orthonormal basis $\beta = \{(1, 0), (0, 1)\}$. Then, $A = [\mathbf{T}]_\beta = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$. Therefore, A is self-adjoint and hence it is normal. That is, $\mathbf{T}$ is self-adjoint and normal operator. We now produce an orthonormal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$. Consider the characteristic polynomial:

$$f(\lambda) = |\,A - \lambda I_2\,| = \begin{vmatrix} 2-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = \cdots = (\,\lambda - 1\,)(\,\lambda - 6\,) = 0.$$

Therefore, $\lambda_1 = 1$ and $\lambda_2 = 6$. For $\underline{E_{\lambda_1}}$: The eigenspace E_{λ_1} corresponding to $\lambda_1 = 1$ is $E_{\lambda_1} = \mathcal{N}(\mathbf{T} - I_2)$. Therefore

$$E_{\lambda_1} = \left\{ (a,b) \in \mathbb{R}^2 \,:\, \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Which implies that $a = 2b$. That is,

$$E_{\lambda_1} = \{\, t(\,2, 1\,) \,:\, t \in \mathbb{R}\,\}.$$

Therefore, $\gamma_1 = \{\,(\,2, 1\,)\,\}$ is a basis for E_{λ_1}.

For $\underline{E_{\lambda_2}}$: The eigenspace E_{λ_2} corresponding to $\lambda_2 = 6$ is $E_{\lambda_2} = \mathcal{N}(\mathbf{T} - 6I_2)$. Therefore

$$E_{\lambda_2} = \left\{ (a,b) \in \mathbb{R}^2 \,:\, \begin{pmatrix} -4 & -2 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Which implies that $a = -\frac{1}{2}b$. That is,

$$E_{\lambda_2} = \{\, t(\,1, -2\,) \,:\, t \in \mathbb{R}\,\}.$$

Therefore, $\gamma_2 = \{\,(\,1, -2\,)\,\}$ is a basis for E_{λ_2}.

Thus, $\gamma = \gamma_1 \cup \gamma_2 = \{\,(\,2, 1\,), (\,1, -2\,)\,\}$ is orthogonal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$. Normalizing the vectors of γ, we obtain γ^* which is orthonormal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$, where

$$\gamma^* = \left\{ \frac{1}{\sqrt{5}}(\,2, 1\,), \frac{1}{\sqrt{5}}(\,1, -2\,) \right\}.$$

We note that, we can confirm our solution by confirming that $Q^{-1}AQ = diag(1, 6)$, where $Q = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$.

Example 6.4.3

Let $\mathbf{T}$ be a linear operator $\mathbb{V} = \mathbb{R}^3$ defined by $\mathbf{T}\,(a, b, c) = (\,-a + b, 5b, 4a - 2b + 5c\,)$. Determine whether $\mathbf{T}$ is normal, self-adjoint, or neither. If possible, produce an orthonormal basis of eigenvectors of $\mathbf{T}$ for $\mathbb{V}$, and list the corresponding eigenvalues.

Solution:

Choose an orthonormal basis $\beta = \{(1,0,0),(0,1,0),(0,0,1)\}$. Then, $A = [\mathbf{T}]_\beta = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 5 & 0 \\ 4 & -2 & 5 \end{pmatrix}$. Then $A^* = \begin{pmatrix} -1 & 0 & 4 \\ 1 & 5 & -2 \\ 0 & 0 & 5 \end{pmatrix}$. Therefore, A is not self-adjoint as $A^* \neq A$. Furthermore, $(AA^*)_{11} = 2$ while $(A^*A)_{11} = 17$. Hence A is not normal. Therefore it has no orthonormal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$.

Example 6.4.4

Let $\mathbf{T}$ be a linear operator on $\mathbb{V} = \mathbb{C}^2$ defined by $\mathbf{T}(a,b) = (2a+bi, a+2b)$. Determine whether $\mathbf{T}$ is normal, self-adjoint, or neither. If possible, produce an orthonormal basis of eigenvectors of $\mathbf{T}$ for $\mathbb{V}$, and list the corresponding eigenvalues.

Solution:

Choose an orthonormal basis $\beta = \{(1,0),(0,1)\}$. Then, $A = [\mathbf{T}]_\beta = \begin{pmatrix} 2 & i \\ 1 & 2 \end{pmatrix}$. Then $A^* = \begin{pmatrix} 2 & 1 \\ -i & 2 \end{pmatrix}$. Therefore, A is not self-adjoint. However, $AA^* = A^*A = \begin{pmatrix} 5 & 2+2i \\ 2-2i & 5 \end{pmatrix}$. That is, $\mathbf{T}$ is normal operator. We now produce an orthonormal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$. Consider the characteristic polynomial:

$$f(\lambda) = |A - \lambda I_2| = \begin{vmatrix} 2-\lambda & i \\ 1 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + (4-i) = \left(\lambda - (2+\sqrt{i})\right)\left(\lambda - (2-\sqrt{i})\right) = 0.$$

Therefore, $\lambda_1 = 2+\sqrt{i}$ and $\lambda_2 = 2-\sqrt{i}$.

For $\underline{E_{\lambda_1}}$: The eigenspace E_{λ_1} corresponding to $\lambda_1 = 2+\sqrt{i}$ is $E_{\lambda_1} = \mathcal{N}(\mathbf{T} - (2+\sqrt{i})I_2)$. Therefore

$$E_{\lambda_1} = \left\{ (a,b) \in \mathbb{C}^2 : \begin{pmatrix} -\sqrt{i} & i \\ 1 & -\sqrt{i} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Which implies that $a = b\sqrt{i}$. That is,

$$E_{\lambda_1} = \left\{ t\left(\sqrt{i}, 1\right) : t \in \mathbb{R} \right\}.$$

Therefore, $\gamma_1 = \left\{ \left(\sqrt{i}, 1\right) \right\}$ is a basis for E_{λ_1}.

For $\underline{E_{\lambda_2}}$: The eigenspace E_{λ_2} corresponding to $\lambda_2 = 2 - \sqrt{i}$ is $E_{\lambda_2} = \mathcal{N}(\mathbf{T} - (2 - \sqrt{i})I_2)$. Therefore

$$E_{\lambda_2} = \left\{ (a, b) \in \mathbb{C}^2 : \begin{pmatrix} \sqrt{i} & i \\ 1 & \sqrt{i} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Which implies that $a = -b\sqrt{i}$. That is,

$$E_{\lambda_2} = \left\{ t\left(\sqrt{i}, -1 \right) : t \in \mathbb{R} \right\}.$$

Therefore, $\gamma_2 = \left\{ \left(\sqrt{i}, -1 \right) \right\}$ is a basis for E_{λ_2}.
Thus, $\gamma = \gamma_1 \cup \gamma_2 = \left\{ \left(\sqrt{i}, 1 \right), \left(\sqrt{i}, -1 \right) \right\}$ is orthogonal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$. Normalizing the vectors of γ, we obtain γ^* which is orthonormal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$, where

$$\gamma^* = \left\{ \frac{1}{\sqrt{2}}\left(\sqrt{i}, 1 \right), \frac{1}{\sqrt{2}}\left(\sqrt{i}, -1 \right) \right\}.$$

We note that, we can confirm our solution by confirming that $Q^{-1}AQ = diag(2+\sqrt{i}, 2-\sqrt{i})$, where $Q = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{i} & \sqrt{i} \\ 1 & -1 \end{pmatrix}$.

Example 6.4.5

Let $\mathbf{T}$ be a linear operator on $\mathbb{V} = \mathbb{P}_2(\mathbb{R})$ defined by $\mathbf{T}(f) = f'$, where $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$. Determine whether $\mathbf{T}$ is normal, self-adjoint, or neither. If possible, produce an orthonormal basis of eigenvectors of $\mathbf{T}$ for $\mathbb{V}$, and list the corresponding eigenvalues.

Solution:

We first consider the standard ordered basis for $\mathbb{P}_2(\mathbb{R})$ which $\beta = \{1, x, x^2\}$. Note that β is not orthonormal and hence we use the Gram-Schmidt process to construct orthogonal basis and then normalize it to obtain an orthonormal basis. Let $\beta = \{u_1 = 1, u_2 = x, u_3 = x^2\}$. Then, $\beta' = \{v_1, v_2, v_3\}$ is orthogonal basis for $\mathbb{P}_2(\mathbb{R})$, where $v_1 = u_1 = 1$. And,

$$v_2 = u_2 - \left(\frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1 \right) = x - \left(\frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 \right) = x - \frac{1}{2},$$

where $\langle x, 1 \rangle = \frac{1}{2}$ and $\langle 1, 1 \rangle = 1$. And,

$$v_3 = u_3 - \left(\frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2 \right) = x^2 - \left(\frac{\langle x^2, 1 \rangle}{\|1\|^2} \cdot 1 + \frac{\langle x^2, (x - \frac{1}{2}) \rangle}{\left\|x - \frac{1}{2}\right\|^2} (x - \frac{1}{2}) \right)$$
$$= x^2 - \left(\frac{1}{3}\right) - \left(\frac{1}{4} - \frac{1}{6}\right)\left(x - \frac{1}{2}\right) = x^2 - x + \frac{1}{6},$$

where $\left\|x - \frac{1}{2}\right\|^2 = \frac{1}{12}$, $\langle x^2, 1 \rangle = \frac{1}{3}$, and $\langle x^2, x - \frac{1}{2} \rangle = \frac{1}{12}$.
Thus, $\beta' = \left\{ 1, x - \frac{1}{2}, x^2 - x + \frac{1}{6} \right\}$ is orthogonal basis for $\mathbb{P}_2(\mathbb{R})$. We may observe that

$$\|1\| = 1, \quad \left\|x - \frac{1}{2}\right\| = \frac{1}{2\sqrt{3}}, \quad \text{and } \left\|x^2 - x + \frac{1}{6}\right\| = \frac{1}{6\sqrt{5}}.$$

Therefore, $\gamma = \left\{ 1, 2\sqrt{3}\left(x - \frac{1}{2}\right), 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right) \right\}$ is orthonormal basis for $\mathbb{P}_2(\mathbb{R})$. We now compute the representation of $\mathbf{T}$ relative to γ. Note that, we can compute the Fourier coefficients as in Theorem 6.2.3:

$$\begin{aligned}
&\mathbf{T}(1) = 0 && \Rightarrow \quad [\mathbf{T}(1)]_\gamma = (0, 0, 0). \\
&\mathbf{T}\left(2\sqrt{3}\left(x - \tfrac{1}{2}\right)\right) = 2\sqrt{3} && \Rightarrow \quad \left[\mathbf{T}\left(2\sqrt{3}\left(x - \tfrac{1}{2}\right)\right)\right]_\gamma = (\tfrac{2}{\sqrt{3}}, 0, 0). \\
&\mathbf{T}\left(6\sqrt{5}\left(x^2 - x + \tfrac{1}{6}\right)\right) = 12\sqrt{5}x - 6\sqrt{5} && \Rightarrow \quad \left[\mathbf{T}\left(6\sqrt{5}\left(x^2 - x + \tfrac{1}{6}\right)\right)\right]_\gamma = (0, 2\sqrt{15}, 0).
\end{aligned}$$

That is $[\mathbf{T}]_\gamma = \begin{pmatrix} 0 & 2\sqrt{3} & 0 \\ 0 & 0 & 2\sqrt{15} \\ 0 & 0 & 0 \end{pmatrix}$. Hence, $\mathbf{T}$ is not self-adjoint and not normal, for instance $([\mathbf{T}]_\gamma [\mathbf{T}]_\gamma^*)_{11} = 12$ while $([\mathbf{T}]_\gamma^* [\mathbf{T}]_\gamma)_{11} = 0$. Therefore, there is no orthonormal basis for $\mathbb{P}_2(\mathbb{R})$ consisting of eigenvectors of $\mathbf{T}$.

Example 6.4.6

Let $\mathbf{T}$ be a linear operator on $\mathbb{V} = M_{2\times 2}(\mathbb{R})$ defined by $\mathbf{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$. Determine whether $\mathbf{T}$ is normal, self-adjoint, or neither. If possible, produce an orthonormal basis of eigenvectors of $\mathbf{T}$ for $\mathbb{V}$, and list the corresponding eigenvalues.

Solution:

Choose the standard orthonormal basis $\beta = \{ E^{11}, E^{12}, E^{21}, E^{22} \}$. Then,

$$\mathbf{T}\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix} \qquad \Rightarrow \left[\mathbf{T}\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\right]_\beta = (0,0,1,0)$$

$$\mathbf{T}\begin{pmatrix}0 & 1\\0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} \qquad \Rightarrow \left[\mathbf{T}\begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}\right]_\beta = (0,0,0,1)$$

$$\mathbf{T}\begin{pmatrix}0 & 0\\1 & 0\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} \qquad \Rightarrow \left[\mathbf{T}\begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}\right]_\beta = (1,0,0,0)$$

$$\mathbf{T}\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix} \qquad \Rightarrow \left[\mathbf{T}\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right]_\beta = (0,1,0,0).$$

Therefore, $A = [\mathbf{T}]_\beta = \begin{pmatrix}0 & 0 & 1 & 0\\0 & 0 & 0 & 1\\1 & 0 & 0 & 0\\0 & 1 & 0 & 0\end{pmatrix}$. Then A and hence $\mathbf{T}$ is self-adjoint and normal.

We now produce an orthonormal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$. Consider the characteristic polynomial:

$$f(\lambda) = |\, A - \lambda I_4 \,| = \begin{vmatrix}-\lambda & 0 & 1 & 0\\0 & -\lambda & 0 & 1\\1 & 0 & -\lambda & 0\\0 & 1 & 0 & -\lambda\end{vmatrix} = \left(\lambda^2 - 1\right)^2 = 0.$$

Therefore, $\lambda_1 = -1$ and $\lambda_2 = 1$.

For $\underline{E_{\lambda_1}}$: The eigenspace E_{λ_1} corresponding to $\lambda_1 = -1$ is $E_{\lambda_1} = \mathcal{N}(\mathbf{T} + I_4)$. Therefore

$$E_{\lambda_1} = \left\{ (a,b,c,d) \in \mathbb{R}^4 \;:\; \begin{pmatrix}1 & 0 & 1 & 0\\0 & 1 & 0 & 1\\1 & 0 & 1 & 0\\0 & 1 & 0 & 1\end{pmatrix}\begin{pmatrix}a\\b\\c\\d\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix} \right\}.$$

Which implies that $a = -c$ and $b = -d$. That is,

$$E_{\lambda_1} = \{\, t(\,1,0,-1,0\,), r(\,0,1,0,-1\,) \;:\; t,r \in \mathbb{R}\,\}.$$

Therefore, $\gamma_1 = \{\,(\,1,0,-1,0\,),(\,0,1,0,-1\,)\,\}$ is a basis for E_{λ_1}.

For $\underline{E_{\lambda_2}}$: The eigenspace E_{λ_2} corresponding to $\lambda_2 = 1$ is $E_{\lambda_2} = \mathcal{N}(\mathbf{T} - I_4)$. Therefore

$$E_{\lambda_2} = \left\{ (a,b,c,d) \in \mathbb{R}^4 \;:\; \begin{pmatrix}-1 & 0 & 1 & 0\\0 & -1 & 0 & 1\\1 & 0 & -1 & 0\\0 & 1 & 0 & -1\end{pmatrix}\begin{pmatrix}a\\b\\c\\d\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix} \right\}.$$

Which implies that $a = c$ and $b = d$. That is,

$$E_{\lambda_2} = \{\, t(1,0,1,0), r(0,1,0,1) \,:\, t, r \in \mathbb{R} \,\}.$$

Therefore, $\gamma_2 = \{\, (1,0,1,0), (0,1,0,1) \,\}$ is a basis for E_{λ_2}.

Thus, $\gamma = \gamma_1 \cup \gamma_2 = \{\, (1,0,-1,0), (0,1,0,-1), (1,0,1,0), (0,1,0,1) \,\}$ is orthogonal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$. Normalizing the vectors of γ, we obtain γ^* which is orthonormal basis for $\mathbb{V}$ consisting of eigenvectors of $\mathbf{T}$, where

$$\gamma^* = \left\{ \frac{1}{\sqrt{2}}(1,0,-1,0), \frac{1}{\sqrt{2}}(1,0,-1,0), \frac{1}{\sqrt{2}}(1,0,1,0), \frac{1}{\sqrt{2}}(1,0,1,0) \right\}.$$

Example 6.4.7

Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Show that A is orthogonally equivalent to a diagonal matrix, and find an orthogonal or unitary matrix P and a diagonal matrix D such that $P^*AP = D$.

Solution:

Clearly, A is symmetric and hence it is orthogonally equivalent to a diagonal matrix. We then construct a unitary matrix P whose columns are the eigenvectors of A (chosen from orthonormal basis) so that $P^*AP = D = diag(\lambda_1, \lambda_2)$.

$$f(\lambda) = |\, A - \lambda I_2 \,| = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 4 = 0.$$

Thus, $\lambda_1 = -1$ and $\lambda_2 = 3$.

For $\underline{E_{\lambda_1}}$: The eigenspace E_{λ_1} corresponding to $\lambda_1 = -1$ is $E_{\lambda_1} = \mathcal{N}(\mathbf{T} + I_2)$. Therefore

$$E_{\lambda_1} = \left\{ (a,b) \in \mathbb{R}^2 \,:\, \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Which implies that $a = -b$. That is,

$$E_{\lambda_1} = \{\, t(1,-1) \,:\, t \in \mathbb{R} \,\}.$$

Therefore, $\gamma_1 = \{\, (1,-1) \,\}$ is a basis for E_{λ_1}.

For $\underline{E_{\lambda_2}}$: The eigenspace E_{λ_2} corresponding to $\lambda_2 = 3$ is $E_{\lambda_2} = \mathcal{N}(\mathbf{T} - 3I_2)$. Therefore

$$E_{\lambda_2} = \left\{ (a,b) \in \mathbb{R}^2 \,:\, \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Which implies that $a = b$. That is,

$$E_{\lambda_2} = \{\, t(\,1, 1\,) \;:\; t \in \mathbb{R}\,\}.$$

Therefore, $\gamma_2 = \{\,(\,1, 1\,)\,\}$ is a basis for E_{λ_2}.

Thus, $\gamma = \gamma_1 \cup \gamma_2 = \{\,(\,1, -1\,), (\,1, 1\,)\,\}$ is orthogonal basis consisting of eigenvectors of A. Normalizing the vectors of γ, we obtain γ^* which is orthonormal basis consisting of eigenvectors of A, where

$$\gamma^* = \left\{ \frac{1}{\sqrt{2}}(\,1, -1\,), \frac{1}{\sqrt{2}}(\,1, 1\,) \right\}.$$

Finally, $P^*AP = D$, where $P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ and $D = diag(-1, 3)$.

Example 6.4.8

Let $A = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$. Show that A is orthogonally equivalent to a diagonal matrix, and find an orthogonal or unitary matrix P and a diagonal matrix D such that $P^*AP = D$.

Solution:

Clearly, A is symmetric and hence it is orthogonally equivalent to a diagonal matrix. We then construct a unitary matrix P whose columns are the eigenvectors of A (chosen from orthonormal basis) so that $P^*AP = D = diag(\lambda_1, \lambda_2.\lambda_3)$.

$$f(\lambda) = |\,A - \lambda I_3\,| = \begin{vmatrix} -\lambda & 2 & 2 \\ 2 & -\lambda & 2 \\ 2 & 2 & -\lambda \end{vmatrix} = \cdots = (\lambda + 2)^2(4 - \lambda) = 0.$$

Thus, $\lambda_1 = -2$ and $\lambda_2 = 4$.

For $\underline{E_{\lambda_1}}$: The eigenspace E_{λ_1} corresponding to $\lambda_1 = -2$ is $E_{\lambda_1} = \mathcal{N}(\mathbf{T} + 2I_2)$. Therefore

$$E_{\lambda_1} = \left\{ (a, b, c) \in \mathbb{R}^3 \;:\; \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Which implies that $a = -b - c$. That is,

$$E_{\lambda_1} = \{\, t(\,1, -1, 0\,), r(\,1, 0, -1\,) \;:\; t, r \in \mathbb{R}\,\}.$$

Therefore, $\gamma_1 = \{ u_1 = (1,-1,0), u_2 = (1,0,-1) \}$ is a basis for E_{λ_1}. We note that γ_1 is not orthogonal set, and hence we use Gram-Schmidt process to orthogonalize it. Let $v_1 = u_1 = (1,-1,0)$, and

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1 = (1,0,-1) - \frac{\langle (1,0,-1),(1,-1,0) \rangle}{\|(1,-1,0)\|^2}(1,-1,0) = \left(\frac{1}{2}, \frac{1}{2}, -1\right).$$

Hence, $\gamma_1^* = \left\{ (1,-1,0), \left(\frac{1}{2}, \frac{1}{2}, -1\right) \right\}$ is orthogonal basis for E_{λ_1}.

For $\underline{E_{\lambda_2}}$: The eigenspace E_{λ_2} corresponding to $\lambda_2 = 4$ is $E_{\lambda_2} = \mathcal{N}(\mathbf{T} - 4I_2)$. Therefore

$$E_{\lambda_2} = \left\{ (a,b) \in \mathbb{R}^2 \, : \, \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Which implies that $a = b = c$. That is,

$$E_{\lambda_2} = \{ t(1,1,1) \, : \, t \in \mathbb{R} \}.$$

Therefore, $\gamma_2 = \{ (1,1,1) \}$ is a basis for E_{λ_2}.

Thus, $\gamma = \gamma_1^* \cup \gamma_2 = \left\{ (1,-1,0), \left(\frac{1}{2}, \frac{1}{2}, -1\right), (1,1,1) \right\}$ is orthogonal basis consisting of eigenvectors of A. Normalizing the vectors of γ, we obtain γ^* which is orthonormal basis consisting of eigenvectors of A, where

$$\gamma^* = \left\{ \frac{1}{\sqrt{2}}(1,-1,0), \sqrt{\frac{2}{3}}\left(\frac{1}{2}, \frac{1}{2}, -1\right), \frac{1}{\sqrt{3}}(1,1,1) \right\}.$$

Finally, $P^*AP = D$, where

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \text{ and } \qquad D = diag(-2,-2,4).$$

Exercise 6.4.1

Solve the following exercises from the book at pages 352 - 357:

- 2 : a, b, c, g, and h.

The Index