Advanced Linear Algebra: Math 403

Dr. Abdullah Al-Azemi

Mathematics Department
Kuwait University

May 7, 2022

Contents

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, $(1,2) \oplus (1,3) \neq (1,3) \oplus (1,2)$.

Theorem 1.2.2

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.

Definition 1.2.2

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 1.2.3

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 1.2.4

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 following three conditions.

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 1.2.3

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 1.2.5

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 1.2.6

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 following three conditions.

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 1.2.7

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 1.2.4

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

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.

Section 1.3: Bases and Dimensions

Definition 1.3.1

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

Theorem 1.3.1

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 $\mathbb{V}$.

Example 1.3.1

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 1.3.2

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 $\mathbb{R}^3$.

Definition 1.3.2

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.

Definition 1.3.3

Let $A \in M_{m\times n}(\mathbb{F})$. Then the **null space of** A is defined as

$$\mathcal{N}(A) = \{ x \in \mathbb{F}^n : Ax = \mathbf{0} \}.$$

The dimension of $\mathcal{N}(A)$ is called the **nullity of** A. This null space of A is a subspace of $\mathbb{F}^n$.

Definition 1.3.4

A set $\beta = \{ s_1, s_2, \cdots, s_n \}$ 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**) $\mathbb{V}$; that is, any element $x \in \mathbb{V}$ can be represented as a linear combination of elements of β: $x = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n$, and
2. β is **linearly independent** set in $\mathbb{V}$; that is, $a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = \mathbf{0}$ implies that $a_1 = a_2 = \cdots = a_n = 0$.

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

Remark 1.3.1

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. 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 $\mathbf{dim}(\mathbb{P}_n(\mathbb{F})) = n + 1$.

Remark 1.3.2

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

Theorem 1.3.2

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 1.3.3

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

Example 1.3.3

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 1.3.4

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 1.3.5

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, $\mathbb{W} = \mathbf{span}\ S$, 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 $\mathbf{dim}(\mathbb{W}) = 2$.

Exercise 1.3.1

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 $\mathbb{W}$.

Exercise 1.3.2

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

Exercise 1.3.3

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.

Exercise 1.3.4

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 $\mathbb{V}$.

Section 1.4: Coordinates of a Vector

Definition 1.4.1

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

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

Consider the vector space $\mathbb{R}^2$ with the usual vector addition and scalar multiplication. The set $\beta = \{ e_1 = (1,0), e_2 = (0,1) \}$ is clearly a basis for $\mathbb{R}^2$.

The vector $x = (1,-2) \in \mathbb{R}^2$ has its coordinates in the basis β defined as

$$[x]_\beta = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

The coordinate of x in another basis for $\mathbb{R}^2$, say γ, is in general different from the coordinate of x in β. This can be seen in the following example.

Example 1.4.1

Let $\beta = \{ e_1 = (1,0), e_2 = (0,1) \}$ be the natural (standard) basis for $\mathbb{R}^2$, and let $\gamma = \{ f_1, f_2 \}$, where $f_1 = (1,1)$ and $f_2 = (1,2)$.

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

Solution:

(1): Note that $|\,\beta\,| = |\,\gamma\,| = 2 = \mathbf{dim}(\mathbb{R}^2)$. So, we only need to show that γ is linearly independent (or γ spans $\mathbb{R}^2$). Consider $c_1f_1 + c_2f_2 = \mathbf{0}$ which is a homogenous system with $Av = 0$, where $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ and $v = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$. Clearly then $|\,A\,| = 1 \neq 0$ and hence γ is linearly independent and it is a basis for $\mathbb{R}^2$.

(2): Note that $[x]_\beta = \begin{bmatrix} a \\ b \end{bmatrix}$ since $x = ae_1 + be_2$.

Now consider

$$\begin{aligned}(a,b) &= c_1 f_1 + c_2 f_2 = c_1(1,1) + c_2(1,2)\\ &= (c_1 + c_2, c_1 + 2c_2).\end{aligned}$$

Therefore, $a = c_1 + c_2$ and $b = c_1 + 2c_2$. Hence $c_1 = 2a - b$ and $c_2 = b - a$. That is, $[x]_\gamma = \begin{bmatrix} 2a-b \\ b-a \end{bmatrix}$.

To check that this is the right coordinate representation of x in γ, simply assert that

$$x = A\,[x]_\gamma\,.$$

Getting back to the vector $x = (1,-2) \in \mathbb{R}^2$, we have its coordinates in $\gamma = \{\, f_1 = (1,1), f_2 = (1,2)\,\}$ as

$$[x]_\gamma = \begin{bmatrix} 4 \\ -3 \end{bmatrix}.$$

This can be seen as

$$x = \begin{bmatrix} 1 \\ -2 \end{bmatrix} = A\,[x]_\gamma = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ -3 \end{bmatrix}.$$

Exercise 1.4.1

Let $\beta = \{\, e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)\,\}$ be the natural (standard) basis for $\mathbb{R}^3$, and let $\gamma = \{\, f_1, f_2, f_3\,\}$, where $f_1 = (1,1,1)$, $f_2 = (0,1,1)$ and $f_3 = (0,0,1)$.

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

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 $\mathbf{T} : M_{m\times n}(\mathbb{F}) \to M_{n\times m}(\mathbb{F})$ 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 2.1.2

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 2.1.3

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

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

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

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 2.1.4

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:

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

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

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

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

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

$$\mathcal{R}(\mathbf{T}) = \mathbf{span}(\mathbf{T}(\beta)) = \mathbf{span}(\{\, \mathbf{T}(x_1), \mathbf{T}(x_2), \cdots, \mathbf{T}(x_n) \,\}).$$

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 $\text{rank}(\mathbf{T})$, 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 $\mathcal{N}(\mathbf{T}) = \{0\}$.

Theorem 2.1.5

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces **of equal finite dimension**, 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 2.1.5

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. That is, show that $\mathbf{T}$ is a bijection.

Solution:

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

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

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

Example 2.1.6

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 2.1.7

Let $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ be the linear transformation defined by $\mathbf{T}\left(f(x)\right) = f'(x) + \int_0^x f(t)\, dt$.
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,

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

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 2.1.8

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.

Definition 2.1.6

A linear transformation is called an **isomorphism** of vector spaces if it is a bijection.

Theorem 2.1.6

Let $\mathbb{V}$ and $\mathbb{W}$ be two finite dimensional vector spaces on the same field. Then there exists an isomorphism of vector spaces $\mathbf{T}: \mathbb{V} \to \mathbb{W}$ if and only if $\mathbf{dim}(\mathbb{V}) = \mathbf{dim}(\mathbb{W})$.

Exercise 2.1.1

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

Exercise 2.1.2

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

Exercise 2.1.3

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 $\mathbf{T}\big(f(x)\big) = \int_a^b f(x)\, dx$ for all $a, b \in \mathbb{R}$ with $a < b$. Show that $\mathbf{T}$ is linear.

Exercise 2.1.4

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 $rank(\mathbf{T})$.

Exercise 2.1.5

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

Exercise 2.1.6

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

1. Find $\mathcal{N}(\mathbf{T})$ and $\mathcal{R}(\mathbf{T})$.
2. Find bases for $\mathcal{N}(\mathbf{T})$ and $\mathcal{R}(\mathbf{T})$.
3. Is $\mathbf{T}$ one-to-one or onto? Explain.

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 bases.
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 $f(x) = -9x^2 + 7x + 3$ relative to β.

Solution:

Clearly $f(x) = -9x^2 + 7x + 3 = 3 + 7x - 9x^2 = 3 \cdot 1 + 7 \cdot x + (-9) \cdot x^2 + 0 \cdot x^3$, 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, $1 \le j \le n$, we have $\mathbf{T}(x_j) \in \mathbb{W}$ and there exist unique scalars $c_{ij} \in \mathbb{F}$, $1 \le i \le m$, 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

$\star$ 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

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

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

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 $\mathbb{F}$.

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 transformations 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 2.2.3

Let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$ and $\mathbf{U} : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformations respectively defined by

$$\mathbf{T}(x, y) = (x + 3y, 0, 2x - 4y) \quad \text{and} \quad \mathbf{U}(x, y) = (x - y, 2x, 3x + 2y).$$

Let β and γ be the standard ordered bases of $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Find the matrix representation of $\mathbf{T} + \mathbf{U}$; that is, $[\mathbf{T} + \mathbf{U}]_\beta^\gamma$.

Solution:

Note that β and γ are the standard ordered bases for $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Then,

$$\mathbf{T}(1, 0) = (1, 0, 2) = 1 \cdot e_1 + 0 \cdot e_2 + 2e_3 \quad \Rightarrow [\mathbf{T}(1, 0)]_\gamma = (1, 0, 2)$$
$$\mathbf{T}(0, 1) = (3, 0, -4) = 3 \cdot e_1 + 0e_2 + (-4)e_3 \quad \Rightarrow [\mathbf{T}(0, 1)]_\gamma = (3, 0, -4).$$

That is,

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

and

$$\mathbf{U}(1,0) = (1,2,3) = 1 \cdot e_1 + 2 \cdot e_2 + 3e_3 \qquad \Rightarrow [\mathbf{U}(1,0)]_\gamma = (1,2,3)$$
$$\mathbf{U}(0,1) = (-1,0,2) = (-1) \cdot e_1 + 0e_2 + 2e_3 \quad \Rightarrow [\mathbf{U}(0,1)]_\gamma = (-1,0,2).$$

That is,

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

If we compute $\mathbf{T} + \mathbf{U}$ using their definitions, we get

$$\mathbf{T} + \mathbf{U} = (2x + 2y, 2x, 5x - 2y).$$

Thus,

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

which is simply $[\mathbf{T}]_\beta^\gamma + [\mathbf{U}]_\beta^\gamma$.

Exercise 2.2.1

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 $u = (2, 4)$.

Exercise 2.2.2

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 $f(x) = 3x^2 + 1$.

Exercise 2.2.3

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 $g(x) = 2x^2 - 3x + 1$.

Exercise 2.2.4

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.

Exercise 2.2.5

Let

$$\alpha = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}, \beta = \left\{ 1, x, x^2 \right\}, \text{ and } \gamma = \{ 1 \}.$$

1. Define $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ by $\mathbf{T}(f(x)) = \begin{pmatrix} f'(0) & 2f(1) \\ 0 & f''(3) \end{pmatrix}$. Compute $[\mathbf{T}]_\beta^\alpha$.
2. Define $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}$ by $\mathbf{T}(f(x)) = f(2)$. Compute $[\mathbf{T}]_\beta^\gamma$.

Exercise 2.2.6

Let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear defined by $\mathbf{T}(x, y) = (x + 3y, 0, 2x - 4y)$. Find $[\mathbf{T}(f(x))]_\beta^\gamma$ and $[\mathbf{T}(f(x))]_\beta^{\gamma'}$, where β, γ and γ' are $\{e_1, e_2\}$, $\{e_1, e_2, e_3\}$ and $\{e_3, e_2, e_1\}$, respectively.

Exercise 2.2.7

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!!

Exercise 2.2.8

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 $\mathcal{N}(\mathbf{T})$.

Section 2.3: Compositions and Isomorphisms of Linear Transformations

In this section, we consider *invertible* linear transformations. We furthermore consider *isomorphism* linear tranformations.

Definition 2.3.1

A linear transformation is called an **isomorphism** of vector spaces if it is a bijection.

Example 2.3.1

Let $\mathbf{T}:\mathbb{R}^3\to\mathbb{R}^3$ be the linear transformation defined by

$$\mathbf{T}(x,y,z)=(x+y+z,x+y,x).$$

Find $[\mathbf{T}]_\beta$, where β is the standard ordered basis of $\mathbb{R}^3$, and show that $\mathbf{T}$ is an isomorphism.

Solution:

Clearly

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

Therefore, $A=[\mathbf{T}]_\beta=\begin{bmatrix}1&1&1\\1&1&0\\1&0&0\end{bmatrix}$. Note that A is nonsingular with $|\,A\,|=-1\neq 0$.

Now to prove that $\mathbf{T}$ is an isomorphism, we need to show that it is a bijection.

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

That is $\mathbf{T}$ is a one-to-one and hence it is onto as well. Therefore, $\mathbf{T}$ is a bijection and thus it is an isomorphism.

We now prove that the composite of linear transformations is linear. Note that, we write $\mathbf{TU}$ instead of $\mathbf{T}\circ\mathbf{U}$.

Theorem 2.3.1

Let $\mathbb{V}$, $\mathbb{W}$ and $\mathbb{Z}$ be vector spaces over the same field $\mathbb{F}$. Let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$, $\mathbf{U} : \mathbb{W} \to \mathbb{Z}$ be linear. Then $\mathbf{UT} : \mathbb{V} \to \mathbb{Z}$ is linear.

Proof:

Let $x, y \in \mathbb{V}$ and $a \in \mathbb{F}$. Then

$$\begin{aligned}\mathbf{UT}(ax+y) &= \mathbf{U}(\mathbf{T}(ax+y)) = \mathbf{U}(a\mathbf{T}(x)+\mathbf{T}(y)) \\ &= a\mathbf{U}(\mathbf{T}(x)) + \mathbf{U}(\mathbf{T}(y)) = a(\mathbf{UT})(x) + \mathbf{UT}(y).\end{aligned}$$

Theorem 2.3.2

Let $\mathbb{V}$ be a vector space. Let $\mathbf{T}, \mathbf{U}_1, \mathbf{U}_2 \in \mathcal{L}(\mathbb{V})$. Then,

1. $\mathbf{T}(\mathbf{U}_1 + \mathbf{U}_2) = \mathbf{TU}_1 + \mathbf{TU}_2$ and $(\mathbf{U}_1 + \mathbf{U}_2)\mathbf{T} = \mathbf{U}_1\mathbf{T} + \mathbf{U}_2\mathbf{T}$.
2. $\mathbf{T}(\mathbf{U}_1\mathbf{U}_2) = (\mathbf{TU}_1)\mathbf{U}_2$.
3. $\mathbf{TI}_{\mathbb{V}} = \mathbf{I}_{\mathbb{V}}\mathbf{T} = \mathbf{T}$.
4. $a(\mathbf{U}_1\mathbf{U}_2) = (a\mathbf{U}_1)\mathbf{U}_2 = \mathbf{U}_1(a\mathbf{U}_2)$ for all scalars a.

Theorem 2.3.3

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

$$[\mathbf{UT}]_\alpha^\gamma = [\mathbf{U}]_\beta^\gamma \, [\mathbf{T}]_\alpha^\beta \, .$$

Moreover, for each $x \in \mathbb{V}$, we have

$$[\mathbf{T}(x)]_\beta = [\mathbf{T}]_\alpha^\beta \, [x]_\alpha \, .$$

We now illustrate Theorem 2.3.2 in the next example.

Example 2.3.2

Let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$ and $\mathbf{U} : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformations respectively defined by

$$\mathbf{T}(x, y) = (x+y, x-y, x) \quad \text{and} \quad \mathbf{U}(x, y, z) = (z, x-z).$$

Find a matrix representation for $\mathbf{UT}$, denoted by $[UT]_\beta^\gamma$, where β and γ are the standard bases of $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively.

Moreover, use Theorem 2.3.3 to find $[\mathbf{T}(2,3)]_\gamma$.

Solution:

Note that $\mathbf{UT} : \mathbb{R}^2 \to \mathbb{R}^2$ and

$$\mathbf{UT}(x,y) = \mathbf{U}(x+y, x-y, x) = (x, (x+y) - x) = (x, y).$$

That is, $[\mathbf{UT}]_\beta = [\mathbf{I}_{\mathbb{R}^2}]_\beta = I_2$.

On the other hand, we can compute

$$[\mathbf{U}]_\gamma^\beta = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad [\mathbf{T}]_\beta^\gamma = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix},$$

to get

$$[\mathbf{UT}]_\beta = [\mathbf{U}]_\gamma^\beta \, [\mathbf{T}]_\beta^\gamma = I_2.$$

Moreover,

$$[\mathbf{T}(2,3)]_\gamma = [\mathbf{T}]_\beta^\gamma \, [(2,3)]_\gamma = (5, -1, 2).$$

Definition 2.3.2

A matrix $A \in M_{n\times n}(\mathbb{R})$ is called **invertible** if there exists $B \in M_{n\times n}(\mathbb{R})$ such that $AB = BA = I_n$.

In this case, we say that B is the inverse of A and we write $B = A^{-1}$.

Definition 2.3.3

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces, and let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be linear. A function $\mathbf{U} : \mathbb{W} \to \mathbb{V}$ is said to be an inverse of $\mathbf{T}$ if $\mathbf{TU} = \mathbf{I}_{\mathbb{W}}$ and $\mathbf{UT} = \mathbf{I}_{\mathbb{V}}$.

In that case, we say that $\mathbf{T}$ is **invertible** and its unique inverse $\mathbf{U}$ is denoted by $\mathbf{T}^{-1}$.

Remark 2.3.1

For any functions $\mathbf{T}$ and $\mathbf{U}$, the following facts hold.

1. $(\mathbf{TU})^{-1} = \mathbf{U}^{-1}\mathbf{T}^{-1}$.
2. $(\mathbf{T}^{-1})^{-1} = \mathbf{T}$; in particular, $\mathbf{T}^{-1}$ is invertible.
3. A function f is invertible if and only if it is a bijection. In particular, if $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ is linear, then $\mathbf{T}$ is invertible if and only if $\mathbf{T}$ is a bijection (that is, if and only if $\mathbf{rank}(\mathbf{T}) = \mathbf{dim}(\mathbb{V})$).

Theorem 2.3.4

Let β and γ be two ordered bases for the two vector spaces $\mathbb{V}$ and $\mathbb{W}$ of the same dimension, respectively. Let $\mathbf{T} : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then

1. $\mathbf{T}$ is invertible (a bijection) if and only if $[\mathbf{T}]_\beta^\gamma$ is invertible.
2. If $\mathbf{T}$ is invertible, then $\mathbf{T}^{-1} : \mathbb{W} \to \mathbb{V}$ is linear.
3. If $\mathbf{T}$ is invertible, then $[\mathbf{T}^{-1}]_\gamma^\beta = \left([\mathbf{T}]_\beta^\gamma\right)^{-1}$.

Example 2.3.3

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

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

Show that $\mathbf{T}$ is invertible, find $[\mathbf{T}^{-1}]_\beta$, where β is the standard ordered basis of $\mathbb{R}^2$ and use it to find $\mathbf{T}^{-1}(x, y)$ for any $(x, y) \in \mathbb{R}^2$.

Moreover, find $[\mathbf{T}(x, y)]_\beta$ and $[\mathbf{T}^{-1}(x, y)]_\beta$.

Solution:

Clearly,

$$\mathbf{T}(e_1) = \mathbf{T}(1, 0) = (1, 0) = e_1 \quad \text{and} \quad \mathbf{T}(e_2) = \mathbf{T}(0, 1) = (2, 1) = 2e_1 + e_2.$$

That is,

$$[\mathbf{T}]_\beta = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

which is invertible as $\left|[\mathbf{T}]_\beta\right| = 1 \neq 0$. Moreover,

$$\left[\mathbf{T}^{-1}\right]_\beta = \left([\mathbf{T}]_\beta\right)^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix},$$

and hence

$$\mathbf{T}^{-1}(e_1) = \mathbf{T}^{-1}(1,0) = (1,0) = e_1 \quad \text{and} \quad \mathbf{T}^{-1}(e_2) = \mathbf{T}^{-1}(0,1) = (-2,1) = -2e_1 + e_2.$$

Thus, for any $(x, y) \in \mathbb{R}^2$, we have

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

Finally,

$$[\mathbf{T}(x,y)]_\beta = [\mathbf{T}]_\beta^\beta \, [(x,y)]_\beta = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ y \end{bmatrix},$$

and

$$\left[\mathbf{T}^{-1}(x,y)\right]_\beta = \left[\mathbf{T}^{-1}\right]_\beta^\beta [(x,y)]_\beta = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - 2y \\ y \end{bmatrix}.$$

Example 2.3.4

Let $\mathbf{T} : \mathbb{P}_1(\mathbb{R}) \to \mathbb{R}^2$ be the linear transformation defined by

$$\mathbf{T}(a + bx) = (a, a + b).$$

Show that $\mathbf{T}$ is invertible and determine $\mathbf{T}^{-1}$.

Solution:

Consider β and γ as the standard ordered bases for $\mathbb{R}^2$ and $\mathbb{P}_1(\mathbb{R})$, respectively. Clearly, $\mathbf{T}(1) = (1,1)$ and $\mathbf{T}(x) = (0,1)$ which implies

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

Note that $[\mathbf{T}]_\gamma^\beta$ is invertible since $\left|[\mathbf{T}]_\gamma^\beta\right| = 1 \neq 0$. Thus,

$$\left[\mathbf{T}^{-1}\right]_\beta^\gamma = \left([\mathbf{T}]_\gamma^\beta\right)^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

That is

$$\left[\mathbf{T}^{-1}(a,b)\right]_{\gamma} = \left[\mathbf{T}^{-1}\right]_{\beta}^{\gamma}[(a,b)]_{\beta} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ -a+b \end{bmatrix}.$$

Therefore,

$$\mathbf{T}^{-1}(a,b) = a\cdot 1 + (-a+b)\cdot x = a + (b-a)x.$$

Definition 2.3.4

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

Example 2.3.5

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 2.3.2

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 $\mathbf{L}_{cA} = c\mathbf{L}_A$.
5. If $C \in M_{n\times p}(\mathbb{F})$, then $\mathbf{L}_{AC} = \mathbf{L}_A\mathbf{L}_C$.
6. If $A \in M_{n\times n}(\mathbb{F})$, then A is invertible if and only if $\mathbf{L}_A$ is invertible. Furthermore, $(\,\mathbf{L}_A\,)^{-1} = \mathbf{L}_{A^{-1}}$.

Exercise 2.3.1

Let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$ and $\mathbf{U} : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformations respectively defined by

$$\mathbf{T}(x, y) = (x + y, x - y, x) \quad \text{and} \quad \mathbf{U}(x, y, z) = (x + y - z, x - z).$$

Find a matrix representation for $\mathbf{UT}$.

Exercise 2.3.2

Let $g(x) = 3 + x$. Let $\mathbf{T} : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ and $\mathbf{U} : \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}^3$ be the linear transformations respectively defined by

$$\mathbf{T}(f(x)) = f'(x)\, g(x) + 2f(x) \quad \text{and} \quad \mathbf{U}(a + bx + cx^2) = (a + b, c, a - b).$$

Let $\beta = \{ 1, x, x^2 \}$ and $\gamma = \{ (1, 0, 0), (0, 1, 0), (0, 0, 1) \}$ be the standard bases for $\mathbb{P}_2(\mathbb{R})$ and $\mathbb{R}^3$, respectively.

1. Compute $[\mathbf{U}]_\beta^\gamma$, $[\mathbf{T}]_\beta$, and $[\mathbf{UT}]_\beta^\gamma$ directly. Then use Theorem 2.3.3 to verify your result.
2. Let $h(x) = 3 - 2x + x^2$. Compute $[h(x)]_\beta$ and $[\mathbf{U}(h(x))]_\gamma$. Then use $[\mathbf{U}]_\beta^\gamma$ from part (1) and Theorem 2.3.3 to verify your result.

Exercise 2.3.3

Let $\mathbf{T}$ be the linear transformation defined in the corresponding part of Exercise 2.2.5 of Section 2.2. Use Theorem 2.3.3 to compute the following vectors:

1. $[\mathbf{T}(f(x))]_\alpha$, where $f(x) = 4 - 6x + 3x^2$.
2. $[\mathbf{T}(f(x))]_\gamma$, where $f(x) = 6 - x + 2x^2$.

Exercise 2.3.4

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

$$\mathbf{T}(x, y) = (2x + y, 5x + 3y).$$

Show that $\mathbf{T}$ is invertible and determine $\mathbf{T}^{-1}$.

Exercise 2.3.5

Let $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$\mathbf{T}(a, b, c) = (3a - 2c, b, 3a + 4b).$$

Show that $\mathbf{T}$ is invertible and determine $\mathbf{T}^{-1}$.

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

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

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

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

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 $(1,7) = 2(2,4) + (-1)(3,1)$.

b: 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

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

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:

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

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

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

Definition 2.5.2

Let A and B be matrices in $M_{n\times n}(\mathbb{F})$. We say that B is **similar** to A if there exits a non-singular matrix P such that $B = P^{-1}AP$. In that case, we write $B \equiv A$.

Remark 2.5.1

Note that if $\mathbf{T}$ is a linear operator on a finite-dimensional vector space $\mathbb{V}$, and if β and γ are any ordered bases for $\mathbb{V}$, then $[\mathbf{T}]_\gamma$ is similar to $[\mathbf{T}]_\beta$.

Theorem 2.5.2

Let $A, B, C \in M_{n\times n}(\mathbb{F})$. Then

1. $A \equiv A$.
2. If $B \equiv A$, then $A \equiv B$.
3. If $A \equiv B$ and $B \equiv C$, then $A \equiv C$.
4. If $A \equiv B$, then $|\,A\,| = |\,B\,|$.

Exercise 2.5.1

Let $\beta = \{ (-4, 3), (2, -1) \}$ and $\gamma = \{ (2, 1), (-4, 1) \}$ be two ordered bases for $\mathbb{R}^2$. Find the change of coordinate matrix Q, that changes γ-coordinates into β-coordinates.

Exercise 2.5.2

Let $\beta = \{ (1, 0), (0, 1) \}$ and $\gamma = \{ (1, 1), (1, 2) \}$ be two ordered bases for $\mathbb{R}^2$, and let $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\mathbf{T}(a, b) = (2a + b, a - 3b)$. Find the change of coordinate matrix Q, that changes γ-coordinates into β-coordinates, and use it to find $[\mathbf{T}]_\gamma$. Find $[(5, 1)]_\gamma$ using Q.

Chapter

3 Elementary Matrices and Determinants

Section 3.1: Elementary Matrices

Definition 3.1.1

Let A be an $m \times n$ matrix. Any one of the following three operations on the rows (columns) of A is called an **elementary row (column) operation**:

1. interchanging any two rows (columns) of A,
2. multiplying any row (column) of A by a nonzero scalar,
3. adding any scalar multiple of a row (column) of A to another row (column).

Any of these three operations is called **elementary operation**. These operations are called of **type 1**, **type 2**, or **type 3** depending on whether they are obtained by 1, 2, or 3.

Definition 3.1.2 Elementary Matrix

An $n \times n$ **elementary matrix** is a matrix obtained by performing an elementary operation on I_n. The elementary matrix is said to be of **type 1**, **2**, or **3** according to whether the elementary operation performed on I_n is a type 1, 2, or 3 oeration, respectively.

Now we give an important result that uses elementary matrices. Elementary matrices of the three types are explained afterward.

Theorem 3.1.1 Elementary Matrices

Every invertible matrix is a product of elementary matrices.

• Elementary Matrices (Type 1)

Consider the following product of a 3×3 matrix A with an elementary matrix (of type 1) $T_{1,2}$ obtained by changing the first and second rows (columns) of I_3:

$$T_{1,2}\,A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Note that $r_1 \leftrightarrow r_2$. Moreover the following product produces $c_1 \leftrightarrow c_2$

$$A\,T_{1,2} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{bmatrix}.$$

That is, if

$$T_{i,j} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix},$$

then

- $T_{i,j}\,A$ is the matrix obtained from A by interchanging rows i and j.
- $A\,T_{i,j}$ is the matrix obtained from A by interchanging columns i and j.

Note that $T_{i,j}^2 = T_{i,j}\,T_{i,j} = I$ which implies that the $T_{i,j}$ is invertible and its inverse is $T_{i,j}$. That is, $T_{i,j}^{-1} = T_{i,j}$.

Moreover, $\det(\,T_{i,j}\,) = -1 = -\det(\,I_n\,)$. Hence

$$\det(\,T_{i,j}\,A\,) = \det(\,A\,T_{i,j}\,) = -\det(\,A\,).$$

• Elementary Matrices (Type 2)

Now we consider elementary matrices of type 2, denoted by $D_i(m)$, obtained by multiplying row i (column i) of I_n by a nonzero scalar m. That is,

$$D_i(m) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & m & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}.$$

It follows that

- $D_i(m)\,A$ is the matrix obtained from A by multiplying row i by m.
- $A\,D_i(m)$ is the matrix obtained from A by multiplying column i by m.

Note that $\det(\,D_i(m)\,) = m \neq 0$ and hence $D_i(m)$ is invertible. Moreover,

$$D_i^{-1}(m) = D_i\left(\frac{1}{m}\right) \text{ since } D_i(m)\,D_i\left(\frac{1}{m}\right) = I_n.$$

Also,

$$\det(\,D_i(m)\,A\,) = \det(\,A\,D_i(m)\,) = m\det(\,A\,).$$

• Elementary Matrices (Type 3)

Here we consider elementary matrices of type 3, denoted by $L_{i,j}(m)$, obtained by adding row j (column i) multiplied by a scalar m to row i (column j) of I_n. That is,

$$L_{i,j}(m) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & m & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}.$$

It follows that

- $L_{i,j}(m)\,A$ is the matrix obtained from A by adding m times row j to row i.
- $A\,L_{i,j}(m)$ is the matrix obtained from A by adding m times column i to column j.

Note that $\det(\,L_{i,j}(m)\,) = 1 \neq 0$ and hence $L_{i,j}(m)$ is invertible. Moreover,

$$L_{i,j}^{-1}(m) = L_{i,j}(\,-m\,) \text{ since } L_{i,j}(m)\,L_{i,j}(\,-m\,) = I_n.$$

Also,

$$\det(\,L_{i,j}(m)\,A\,) = \det(\,A\,L_{i,j}(m)\,) = \det(\,A\,).$$

Exercise 3.1.1

Let
$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ and } E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

For $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$, compute $E_i A$ for $i = 1, 2, 3$, and conclude the effect of E_i on A in the product for $i = 1, 2, 3$.

Section 3.2: Determinants

We first start with the following facts about determinants of elementary matrices.

1. If E is an elementary matrix of type 1, then $\det(E) = -1$.
2. If E is an elementary matrix of type 2 (by multiplying a row (a column) by a nonzero scalar m), then $\det(E) = m$.
3. If E is an elementary matrix of type 3, then $\det(E) = 1$.

The rest of this section recalls some important facts about properties of determinants.

Theorem 3.2.1

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

1. $\det(AB) = \det(A) \cdot \det(B)$.
2. A is invertible if and only if $\det(A) \neq 0$. Furthermore, if A is invertible, then $\det\left(A^{-1}\right) = \dfrac{1}{\det(A)}$.
3. $\det(A^t) = \det(A)$.
4. If A has rank less than n, then $\det(A) = 0$.
5. If A has a row (or a column) consisting entirely of zeros, then $\det(A) = 0$.
6. If A has two identical rows (or columns), then $\det(A) = 0$.
7. $\det(c\,A) = c^n \det(A)$.

Theorem 3.2.2 Cofactors

Let $A \in M_{n\times n}(\mathbb{F})$, where $n > 2$. Then

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \cdot \det\left(\tilde{A}_{ij}\right),$$

(if the determinant is evaluated by the entries of row i of A) or

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \cdot \det\left(\tilde{A}_{ij}\right),$$

(if the determinant is evaluated by the entries of column j of A), where A_{ij} is the cofactor of the row i and column j; and $\tilde{A}_{ij}$ is the $(n-1)\times(n-1)$ matrix obtained from A by deleting row i and column j.

A matrix $A \in M_{n\times n}(\mathbb{F})$ of the form

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{bmatrix},$$

is called **Vandermonde matrix** and has its determinant

$$\det(A) = \prod_{1\le i<j\le n}^{n} (x_j - x_i)$$

Example 3.2.1

Compute $\det(A)$, where $A = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 \\ 4 & 5 & 2 & 3 \\ 5 & 2 & 3 & 4 \end{bmatrix}$.

Solution:

$$|A| = \begin{vmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 \\ 4 & 5 & 2 & 3 \\ 5 & 2 & 3 & 4 \end{vmatrix} \xrightarrow[c_1+c_4\to c_1]{c_1+c_2\to c_1; c_1+c_3\to c_1} = \begin{vmatrix} 14 & 3 & 4 & 5 \\ 14 & 4 & 5 & 2 \\ 14 & 5 & 2 & 3 \\ 14 & 2 & 3 & 4 \end{vmatrix} = (14) \begin{vmatrix} 1 & 3 & 4 & 5 \\ 1 & 4 & 5 & 2 \\ 1 & 5 & 2 & 3 \\ 1 & 2 & 3 & 4 \end{vmatrix}$$

$$\xrightarrow[-r_1+r_4\to r_4]{-r_1+r_2\to r_2; -r_1+r_3\to r_3} 14 \begin{vmatrix} 1 & 3 & 4 & 5 \\ 0 & 1 & 1 & -3 \\ 0 & 2 & -2 & -1 \\ 0 & -1 & -1 & -1 \end{vmatrix} = 14 \begin{vmatrix} 1 & 1 & -3 \\ 2 & -2 & -1 \\ -1 & -1 & -1 \end{vmatrix}$$

$$\xrightarrow{r_1+r_3\to r_3} \begin{vmatrix} 1 & 1 & -3 \\ 2 & -2 & -1 \\ 0 & 0 & -4 \end{vmatrix} = (14)(-4) \begin{vmatrix} 1 & 1 \\ 2 & -2 \end{vmatrix} = (14)(-4)(-4) = 224.$$

Exercise 3.2.1

Verify that

$$\begin{vmatrix} 1 & -1 & 2 & -1 \\ -3 & 4 & 1 & -1 \\ 2 & -5 & -3 & 8 \\ -2 & 6 & -4 & 1 \end{vmatrix} = 154.$$

Exercise 3.2.2

Verify that

$$\begin{vmatrix} 0 & 2 & 1 & 3 \\ 1 & 0 & -2 & 2 \\ 3 & -1 & 0 & 1 \\ -1 & 1 & 2 & 0 \end{vmatrix} = -3.$$

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 $\mathbf{L}_A(x) = A\,x$ 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 $\mathbf{L}_{cA} = c\mathbf{L}_A$.

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

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

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.

Exercise 5.1.1

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.

Exercise 5.1.2

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.

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 $(1 \leq i \leq k)$, 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 $1 \leq \mathbf{dim}(E_\lambda) \leq m$.

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 $\mathbf{T}\left(f(x)\right) = f'(x)$. 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, $\mathbf{dim}(E_{\lambda_1}) = 2 = m(\lambda_1)$.

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

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

Direct Sum

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$. We can decompose $\mathbb{V}$ into simpler subspaces which offers more insight on the behavior of $\mathbf{T}$.

In the case that $\mathbf{T}$ is diagonalizable operator, the simpler subspaces are the eigenspaces of the operator. However, This approach is of a great interest when the operator is not diagonalizable. This will be discussed in more details in Chapter 7.

Definition 5.2.4 Sum of Subspaces

Let $\mathbb{W}_1, \mathbb{W}_2, \cdots, \mathbb{W}_k$ be subspaces of a vector space $\mathbb{V}$. We define the **sum** of these subspaces to be the set

$$\{ v_1 + v_2 + \cdots + v_k : v_i \in \mathbb{W}_i \text{ for } 1 \leq i \leq k \},$$

which we denote by $\mathbb{W}_1 + \mathbb{W}_2 + \cdots + \mathbb{W}_k$ or $\sum_{i=1}^{k} \mathbb{W}_i$.

Example 5.2.7

Let $\mathbb{W}_1$ denote the xy-plane and $\mathbb{W}_2$ denote the yz-plane. Both $\mathbb{W}_1$ and $\mathbb{W}_2$ are subspaces of $\mathbb{R}^3$. Show that $\mathbb{R}^3 = \mathbb{W}_1 + \mathbb{W}_2$.

Solution:

Let $(a, b, c) \in \mathbb{R}^3$. Then, clearly

$$(a, b, c) = (a, 0, 0) + (0, b, c),$$

where $(a, 0, 0) \in \mathbb{W}_1$ and $(0, b, c) \in \mathbb{W}_2$. Therefore,

$$\mathbb{R}^3 = \mathbb{W}_1 + \mathbb{W}_2.$$

It is clear that the representation of (a, b, c) in Example 5.2.7 is not unique. For example, $(a, b, c) = (a, b, 0) + (0, 0, c)$ is another representation.

Definition 5.2.5 Direct Sum of Subspaces

Let $\mathbb{W}_1, \mathbb{W}_2, \cdots, \mathbb{W}_k$ be subspaces of a vector space $\mathbb{V}$. We define the **direct sum** of these subspaces $\mathbb{W}_1, \mathbb{W}_2, \cdots, \mathbb{W}_k$ and write $\mathbb{W}_1 \oplus \mathbb{W}_2 \oplus \cdots \oplus \mathbb{W}_k$, if

$$\mathbb{V} = \sum_{i=1}^{k} \mathbb{W}_i,$$

and

$$\mathbb{W}_j \cap \sum_{i \neq j}^{k} \mathbb{W}_i = \{ 0 \},$$

for each j $(1 \leq j \leq k)$.

In Example 5.2.7, one can see that $\mathbb{R}^3 \neq \mathbb{W}_1 \oplus \mathbb{W}_2$ since $\mathbb{W}_1 \cap \mathbb{W} = \{\,(\,0, x, 0\,) : x \in \mathbb{R}\,\} \neq \{\,0\,\}$.

Example 5.2.8

Let

$$\begin{aligned} \mathbb{W}_1 &= \{\,(\,a, b, 0, 0\,) : a, b \in \mathbb{R}\,\} \\ \mathbb{W}_2 &= \{\,(\,0, 0, c, 0\,) : c \in \mathbb{R}\,\} \\ \mathbb{W}_3 &= \{\,(\,0, 0, 0, d\,) : d \in \mathbb{R}\,\} \end{aligned}$$

be three subspaces of $\mathbb{R}^4$. Show that $\mathbb{R}^4$ is a direct sum of $\mathbb{W}_1$, $\mathbb{W}_2$ and $\mathbb{W}_3$.

Solution:

Let $(\,a, b, c, d\,) \in \mathbb{R}^4$. Then, clearly

$$(\,a, b, c, d\,) = (\,a, b, 0, 0\,) + (\,0, 0, c, 0\,) + (\,0, 0, 0, d\,),$$

where $(\,a, b, 0, 0\,) \in \mathbb{W}_1$, $(\,0, 0, c, 0\,) \in \mathbb{W}_2$ and $(\,0, 0, 0, d\,) \in \mathbb{W}_3$. Therefore,

$$\mathbb{R}^4 = \mathbb{W}_1 + \mathbb{W}_2 + \mathbb{W}_3.$$

We also need to show that

$$\mathbb{W}_1 \cap (\,\mathbb{W}_2 + \mathbb{W}_3\,) = \mathbb{W}_2 \cap (\,\mathbb{W}_1 + \mathbb{W}_3\,) = \mathbb{W}_3 \cap (\,\mathbb{W}_1 + \mathbb{W}_3\,) = \{\,0\,\}.$$

Clearly these equalities are clear. Therefore,

$$\mathbb{R}^4 = \mathbb{W}_1 \oplus \mathbb{W}_2 \oplus \mathbb{W}_3.$$

Theorem 5.2.6

Let $\mathbb{W}_1, \mathbb{W}_2, \cdots, \mathbb{W}_k$ be subspaces of a finite-dimensional vector space $\mathbb{V}$. The following conditions are equivalent.

1. $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2 \oplus \cdots \oplus \mathbb{W}_k$.
2. $\mathbb{V} = \sum_{i=1}^{k} \mathbb{W}_i$ and, for any vectors $v_1, v_2, \cdots, v_k$ such that $v_i \in \mathbb{W}_i$ $(1 \leq i \leq k)$, if $v_1 + v_2 + \cdots + v_k = 0$, then $v_i = 0$ for all i.
3. Each vector $v \in \mathbb{V}$ can be uniquely written as $v = v_1 + v_2 + \cdots + v_k$, where $v_i \in \mathbb{W}_i$.
4. If γ_i is an ordered basis for $\mathbb{W}_i$ $(1 \leq i \leq k)$, then $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is an ordered basis

for $\mathbb{V}$.

5. For each $i = 1, 2, \cdots, k$, there exists an ordered basis γ_i for $\mathbb{W}_i$ such that $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is an ordered basis for $\mathbb{V}$.

Theorem 5.2.7

A linear operator $\mathbf{T}$ on a finite-dimensional vector space $\mathbb{V}$ is diagonalizable if and only if $\mathbb{V}$ is the direct sum of the eigenspaces of $\mathbf{T}$.

Exercise 5.2.1

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

Exercise 5.2.2

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

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

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

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

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

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 $(a,0,0) \in \mathbb{W}_2$. 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

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

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 $i = 1, 2, \cdots$.

$$\begin{aligned}\mathbf{T}(e_1) &= \mathbf{T}(1, 0, 0) = (0, 1, 0) = e_2,\\ \mathbf{T}^2(e_1) &= \mathbf{T}\left(\mathbf{T}\left(e_1\right)\right) = \mathbf{T}\left(e_2\right) = (-1, 0, 0) = -e_1.\end{aligned}$$

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. Find a basis for $\mathbb{W}$.

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

$$\begin{aligned}\mathbf{T}(e_1) &= (1, -2) = e_1 + (-2)e_2\\ \mathbf{T}(e_2) &= (2, 1) = 2e_1 + e_2.\end{aligned}$$

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,

$$\begin{aligned} f(\mathbf{T}) &= (\mathbf{T}^2 - 2\mathbf{T} + 5\mathbf{I}) \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \mathbf{T}^2 \begin{pmatrix} a \\ b \end{pmatrix} - 2\mathbf{T} \begin{pmatrix} a \\ b \end{pmatrix} + 5\mathbf{I} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \mathbf{T} \begin{pmatrix} a+2b \\ -2a+b \end{pmatrix} - 2 \begin{pmatrix} a+2b \\ -2a+b \end{pmatrix} + 5 \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} (a+2b) + 2(-2a+b) \\ -2(a+2b) + (-2a+b) \end{pmatrix} + \begin{pmatrix} -2a-4b \\ 4a-2b \end{pmatrix} + \begin{pmatrix} 5a \\ 5b \end{pmatrix} \\ &= \begin{pmatrix} -3a+4b \\ -4a-3b \end{pmatrix} + \begin{pmatrix} -2a-4b \\ 4a-2b \end{pmatrix} + \begin{pmatrix} 5a \\ 5b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{T}_0 \begin{pmatrix} a \\ b \end{pmatrix}. \end{aligned}$$

Note that

$$f(A) = A^2 - 2A + 5I = \begin{pmatrix} -3 & 4 \\ -4 & -3 \end{pmatrix} + \begin{pmatrix} -2 & -4 \\ 4 & -2 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

Now we try to decompose a finite-dimensional vector space $\mathbb{V}$ into a direct sum of as many $\mathbf{T}$-invariant subspaces as possible. For that we start to collect a few facts about direct sums of $\mathbf{T}$-invariant subspaces. These facts will be considered in Section 5.4.

Theorem 5.4.4

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and suppose that $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2 \oplus \cdots \oplus \mathbb{W}_k$, where $\mathbb{W}_i$ is a $\mathbf{T}$-invariant subspace of $\mathbb{V}$ for each i $(1 \le i \le k)$. Suppose that $f_i(t)$ is the characteristic polynomial of $\mathbf{T}_{\mathbb{W}_i}$ $(1 \le i \le k)$. Then $f_1(t) \cdot f_2(t) \cdot \cdots \cdot f_k(t)$ is the characteristic polynomial of $\mathbf{T}$.

Let $\mathbf{T}$ be a diagonalizable linear operator on a finite-dimensional vector space $\mathbb{V}$ with distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_k$. Theorem 5.2.7 shows that $\mathbb{V}$ is a direct sum of the eigenspaces of $\mathbf{T}$.

Since each eigenspace is $\mathbf{T}$-invariant, Theorem 5.4.4 shows that for each eigenvalue λ_i, the restriction of $\mathbf{T}$ to E_{λ_i} has characteristic polynomial $(\lambda_i - t)^{m_i}$, where m_i is the dimension of E_{λ_i}. That is, the characteristic polynomial $f(t)$ of $\mathbf{T}$ is

$$f(t) = (\lambda_1 - t)^{m_1}(\lambda_2 - t)^{m_2} \cdots (\lambda_k - t)^{m_k}.$$

It follows that the multiplicity of each eigenvalue is equal to the dimension of the corresponding eigenspace.

Definition 5.4.3

Let $B_1 \in M_{m\times m}(\mathbb{F})$ and $B_2 \in M_{n\times n}(\mathbb{F})$. We define the **direct sum** of B_1 and B_2, denoted $B_1 \oplus B_2$, as the $(m+n)\times(m+n)$ matrix A such that

$$A = \begin{cases} (B_1)_{ij} & \text{for } 1 \le i, j \le m \\ (B_2)_{(i-m),(j-m)} & \text{for } m+1 \le i, j \le m+n \\ 0 & \text{otherwise.} \end{cases}$$

If $B_1, B_2, \cdots, B_k$ are square matrices with entries from $\mathbb{F}$, then we define the **direct sum** of $B_1, B_2, \cdots, B_k$ recursively by

$$B_1 \oplus B_2 \oplus \cdots \oplus B_k = (B_1 \oplus B_2 \oplus \cdots \oplus B_{k-1}) \oplus B_k.$$

If $A = B_1 \oplus B_2 \oplus \cdots \oplus B_k$, then we often write

$$A = \begin{pmatrix} B_1 & O & \cdots & O \\ O & B_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & B_k \end{pmatrix}.$$

Theorem 5.4.5

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and let $\mathbb{W}_1, \mathbb{W}_2, \cdots, \mathbb{W}_k$ be $\mathbf{T}$-invariant subspaces of $\mathbb{V}$ such that $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2 \oplus \cdots \oplus \mathbb{W}_k$. For each i, let β_i be an ordered basis for $\mathbb{W}_i$, and let $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$. Let $A = [\mathbf{T}]_\beta$ and $B_i = \left[\mathbf{T}_{\mathbb{W}_i}\right]_{\beta_i}$ for $i = 1, 2, \cdots, k$. Then, $A = B_1 \oplus B_2 \oplus \cdots \oplus B_k$.

Example 5.4.7

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^4$ defined by

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

and let $\mathbb{W}_1 = \{ (s, t, 0, 0) : s, t \in \mathbb{R} \}$ and $\mathbb{W}_2 = \{ (0, 0, s, t) : s, t \in \mathbb{R} \}$.
Note that $\mathbb{W}_1$ and $\mathbb{W}_2$ are each $\mathbf{T}$-invariant and that $\mathbb{R}^4 = \mathbb{W}_1 \oplus \mathbb{W}_2$. Let $\beta_1 = \{ e_1, e_2 \}$ and $\beta_2 = \{ e_3, e_4 \}$ and hence $\beta = \beta_1 \cup \beta_2 = \{ e_1, e_2, e_3, e_4 \}$.
Then β_1, β_2 and β are bases for $\mathbb{W}_1$, $\mathbb{W}_2$ and $\mathbb{R}^4$, respectively.
Let $A = [\mathbf{T}]_\beta$, $B_1 = \left[\mathbf{T}_{\mathbb{W}_1}\right]_{\beta_1}$, and $B_2 = \left[\mathbf{T}_{\mathbb{W}_2}\right]_{\beta_2}$. Then,

$$B_1 = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

and

$$A = \begin{pmatrix} B_1 & O \\ O & B_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Let $f(t)$, $f_1(t)$, and $f_2(t)$ denote the characteristic polynomials of $\mathbf{T}$, $\mathbf{T}_{\mathbb{W}_1}$, and $\mathbf{T}_{\mathbb{W}_2}$, respectively. Then,

$$f(t) = \det(A - tI) = \det(B_1 - tI) \cdot \det(B_2 - tI) = f_1(t) \cdot f_2(t).$$

Example 5.4.8

Find the direct sum of the following matrices.

$$B_1 = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \; B_2 = \begin{pmatrix} 3 \end{pmatrix}, \quad \text{and } B_3 = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}.$$

Solution:

Clearly,

$$B_1 \oplus B_2 \oplus B_3 = \left[\begin{array}{cc|c|ccc} 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right].$$

Exercise 5.4.1

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 $\mathbf{T}$.

Exercise 5.4.2

Use Cayley-Hamilton Theorem to find A^{-1} if $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix}$.

Chapter

6 Bilinear Forms

Section 6.1: Bilinear Forms

Definition 6.1.1

Let $\mathbb{V}$ be a vector space over a field $\mathbb{F}$. A function $H : \mathbb{V} \times \mathbb{V} \to \mathbb{F}$ is called a **bilinear form** on $\mathbb{V}$ if H is linear in each variable when the other variable is held fix; that is, H is a bilinear form on $\mathbb{V}$ if

1. $H(x + y, z) = H(x, z) + H(y, z)$ and $H(ax, y) = aH(x, y)$ for all $x, y, z \in \mathbb{V}$ and $a \in \mathbb{F}$.
2. $H(x, y + z) = H(x, y) + H(x, z)$ and $H(x, ay) = aH(x, y)$ for all $x, y, z \in \mathbb{V}$ and $a \in \mathbb{F}$.

The set of all bilinear forms is denoted by $\mathcal{B}(\mathbb{V})$.

Remark 6.1.1

The Definition 6.1.1 can be restated as: H is a bilinear form on $\mathbb{V}$ if

1. $H(x + ay, z) = H(x, z) + aH(y, z)$ for all $x, y, z \in \mathbb{V}$ and $a \in \mathbb{F}$.
2. $H(x, y + az) = H(x, y) + aH(x, z)$ for all $x, y, z \in \mathbb{V}$ and $a \in \mathbb{F}$.

An obvious example of a bilinear form is the following: $H : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by $H(x, y) = xy$. Note the difference here: $H(x, y) = x + y$ is linear and $H(x, y) = xy$ is bilinear.

An inner product on a real vector space is a bilinear form, but an inner product on a complex vector space is not, since it is conjugate-linear in the second component rather than (actually) linear.

Theorem 6.1.1 Bilinear Forms on $\mathbb{R}^n$

Every bilinear form on $\mathbb{R}^n$ has the form

$$H(x, y) = x^t A\, y$$

for some $n \times n$ matrix A.

Example 6.1.1

Define a mapping $H : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ by

$$H\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right) = 2a_1b_1 + 3a_1b_2 + 4a_2b_1 - a_2b_2,$$

for $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$. Show that H is a bilinear form on $\mathbb{R}^2$.

Solution:

Note that if $A = \begin{pmatrix} 2 & 3 \\ 4 & -1 \end{pmatrix}$, $x = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, and $y = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, then $H(x, y) = x^t Ay$.
Then H is a bilinear form on $\mathbb{R}^2$ by the distributive property of matrix multiplication over matrix addition. That is,

$$H(x + y, z) = (x^t + y^t)Az = x^t Az + y^t Az = H(x, z) + H(y, z),$$

and

$$H(ax, y) = ax^t Ay = aH(x, y).$$

The same thing applies for $H(x, y + z)$ and $H(x, ay)$. Therefore, H is a bilinear form on $\mathbb{R}^2$.

Remark 6.1.2

For any bilinear form H on a vector space $\mathbb{V}$ over a field $\mathbb{F}$, the following properties hold.

1. If, for any $x \in \mathbb{V}$, the functions $L_x, R_x : \mathbb{V} \to \mathbb{F}$ are defined by

$$L_x(y) = H(x, y) \quad \text{and} \quad R_x(y) = H(y, x) \quad \text{for all } y \in \mathbb{V},$$

   then L_x and R_x are linear.
2. $H(0, x) = H(x, 0) = 0$ for all $x \in setV$.
3. For all $x, y, z, w \in \mathbb{V}$,

$$H(x + y, z + w) = H(x, z) + H(x, w) + H(y, z) + H(y, w).$$

4. If $J : \mathbb{V} \times \mathbb{V} \to \mathbb{F}$ is defined by $J(x, y) = H(y, x)$, then J is a bilinear form.

Matrix Representation of Bilinear Forms

Definition 6.1.2

Let $\mathbb{V}$ be a finite-dimensional vector space over a field $\mathbb{F}$ with a basis $\beta = \{\beta_1, \beta_2, \cdots, \beta_n\}$ and let H be a bilinear form on $\mathbb{V}$.

The associated matrix of H with respect to β is the matrix $[H]_\beta \in M_{n\times n}(\mathbb{F})$ whose (i, j)-entry is the value $H(\beta_i, \beta_j)$.

Example 6.1.2

Define a mapping $H : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ by

$$H((a, b), (c, d)) = 2ac + 4ad - bc,$$

for $(a, b), (c, d) \in \mathbb{R}^2$. Find $[H]_\beta$ for the standard basis $\beta = \{(1, 0), (0, 1)\}$.

Solution:

We simply calculate the values $H(\beta_i, \beta_j)$ for $i, j \in \{1, 2\}$, where $\beta_1 = (1, 0)$ and $\beta_2 = (0, 1)$. Clearly, $H(\beta_1, \beta_1) = 2$, $H(\beta_1, \beta_2) = 4$, $H(\beta_2, \beta_1) = -1$, and $H(\beta_2, \beta_2) = 0$. That is,

$$[H]_\beta = \begin{pmatrix} 2 & 4 \\ -1 & 0 \end{pmatrix}.$$

Theorem 6.1.2

Let $\mathbb{V}$ be a finite-dimensional vector space over a field $\mathbb{F}$ with two bases β and γ, and let H be a bilinear form on $\mathbb{V}$. Then, if $Q = [\mathbf{I}]_\gamma^\beta$ is the change of basis matrix from γ to β, then

$$[H]_\gamma = Q^t\ [H]_\beta\ Q.$$

Example 6.1.3

Define a mapping $H : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ by

$$H((a, b), (c, d)) = 2ac + 4ad - bc,$$

for $(a, b), (c, d) \in \mathbb{R}^2$. Find the change of basis matrix Q from γ to β and use it to find $[H]_\gamma$, where $\beta = \{ (1, 0), (0, 1) \}$ and $\gamma = \{ (1, 0), (1, 1) \}$.

Solution:

We first compute $Q = [\mathbf{I}]_{\gamma}^{\beta}$ whose columns are the vectors of γ in the β coordinates. That is,

$$Q = \left[[(1,0)]_\beta \quad [(1,1)]_\beta \right] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Then, we calculate $[H]_\beta$:

$$[H]_\beta = \begin{pmatrix} 2 & 4 \\ -1 & 0 \end{pmatrix}.$$

Therefore,

$$[H]_\gamma = Q^t \, [H]_\beta \, Q = \begin{pmatrix} 2 & 6 \\ 1 & 5 \end{pmatrix}.$$

This can be verified by a direct computation as well.

Exercise 6.1.1

Define a mapping $H : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ by

$$H(\,(a,b),(c,d)\,) = ac + ad - bd,$$

for $(a,b),(c,d) \in \mathbb{R}^2$. Show that H is a bilinear form on $\mathbb{R}^2$.

Exercise 6.1.2

Define a mapping $H : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ by

$$H(\,(\,a,b\,),(\,c,d\,)\,) = 2ac + 4ad - bc,$$

for $(\,a,b\,),(\,c,d\,) \in \mathbb{R}^2$. Find $[H]_\beta$ for the basis $\beta = \{\,(\,2,1\,),(\,-1,4\,)\,\}$.

Exercise 6.1.3

Define a mapping $H : \mathbb{P}_2(\mathbb{R}) \times \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}$ by

$$H(\,p,q\,) = \int_0^1 p(x)q(x)\,dx,$$

for $p,q \in \mathbb{P}_2(\mathbb{R})$. Find $[H]_\beta$ for the standard basis $\beta = \{\,1,x,x^2\,\}$.

Exercise 6.1.4

Define a mapping $H : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ by $H(x,y) = x^t\,A\,y$ for $x,y \in \mathbb{R}^2$ and $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$. Find the change of basis matrix Q from γ to β and use it to find $[H]_\gamma$, where $\beta = \{\,(\,1,0\,),(\,0,1\,)\,\}$ and $\gamma = \{\,(\,1,-1\,),(\,1,2\,)\,\}$. Moreover, use Q to find $[(1,7)]_\beta$.

Chapter

7 Canonical Forms

Recall that a diagonalizable linear operator has a diagonal matrix representation. That is, there is an ordered basis consisting of eigenvectors of the operator.

However, not every operator is diagonalizable, even if its characteristic polynomial splits.

In this chapter, we consider alternative matrix representation for nondiagonalizable operators. These representations are called **canonical forms**.

We mainly consider two common canonical forms. The first one is called **Jordan canonical form** which requires that the characteristic polynomial splits. That is, every polynomial with coefficients from the underlying field is algebraically closed.

The other canonical form is the **rational canonical form** which does not require such a factorization.

Section 7.1: Jordan Canonical Form I

Recall the following example.

Example 7.1.1

Let $\mathbf{T}$ be a linear operator on $\mathbb{R}^3$ defined by $\mathbf{T}(a,b,c) = (4a+c, 2a+3b+2c, a+4c)$. Is $\mathbf{T}$ diagonalizable? Explain.

Solution:

Using the standard ordered basis for $\mathbb{R}^3$, $\beta = \{e_1, e_2, e_3\}$, we get

$$A = [\mathbf{T}]_\beta = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}.$$

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 $m(5) = 1$.

$$E_{\lambda_1} = \left\{ r \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} : t, r \in \mathbb{R} \right\}; \ \mathbf{dim}(E_{\lambda_1}) = 2 = m(\lambda_1).$$

$$E_{\lambda_2} = \left\{ t \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\}.; \ \mathbf{dim}(E_{\lambda_2}) = 1 = m(\lambda_2).$$

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

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

On the other hand,

Example 7.1.2

Let $\mathbf{T}$ be a linear operator on $\mathbb{R}^3$ defined by $\mathbf{T}(a, b, c) = (3a + b - 2c, -a + 5c, -a - b + 4c)$.

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

Solution:

Using the standard ordered basis for $\mathbb{R}^3$, $\beta = \{ e_1, e_2, e_3 \}$, we get

$$A = [\mathbf{T}]_\beta = \begin{pmatrix} 3 & 1 & -2 \\ -1 & 0 & 5 \\ -1 & -1 & 4 \end{pmatrix}.$$

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

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

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

$$E_{\lambda_1} = \left\{ t \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\}; \ \mathbf{dim}(E_{\lambda_1}) = 1 = m(\lambda_1).$$

$$E_{\lambda_2} = \left\{ t \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\}.; \ \mathbf{dim}(E_{\lambda_2}) = 1 \neq m(\lambda_2) = 2.$$

Therefore, $\mathbf{T}$ is **NOT** diagonalizable.

In what follows, we try to work things out to make the matrix A to be "almost" a diagonal matrix.

The plan is to extend the definition of eigenspace to **generalized eigenspace** by considering **generalized eigenvectors** of the operator $\mathbf{T}$. From these subspaces, we select ordered bases whose union is an ordered basis β for $\mathbb{V}$ such that

$$J = [\mathbf{T}]_\beta = \begin{pmatrix} J_1 & O & \dots & O \\ O & J_2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & J_k \end{pmatrix},$$

where each O is a zero matrix and each J_i is a square matrix of the form (λ) or

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

for some eigenvalue λ for $\mathbf{T}$.

Such a matrix J_i is called a **Jordan block** corresponding to λ, and the matrix $J = [\mathbf{T}]_\beta$ is called a **Jordan canonical form** of $\mathbf{T}$.

Note that each Jordan block J_i is "almost" a diagonal matrix. In fact, J is a diagonal matrix if and only if each of J_i is of the form (λ).

Definition 7.1.1

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$, and let λ be a scalar. A nonzero vector x in $\mathbb{V}$ is called a **generalized eigenvectors** of $\mathbf{T}$ corresponding to λ if

$$(\mathbf{T} - \lambda I)^p(x) = \mathbf{0}$$

for some positive integer p.

Note that if x is a generalized eigenvector for $\mathbf{T}$ corresponding to λ, and p is the smallest positive integer for which $(\mathbf{T} - \lambda I)^p x = \mathbf{0}$, then $(\mathbf{T} - \lambda I)^{p-1} x$ is an eigenvector of $\mathbf{T}$ corresponding to λ. That is, λ is an eigenvalue of $\mathbf{T}$.

Definition 7.1.2

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$, and let λ be an eigenvalue of $\mathbf{T}$. The **generalized eigenspace** of $\mathbf{T}$ corresponding to λ, denoted by $K_\lambda(\mathbf{T})$, is the subset of $\mathbb{V}$ defined by

$$K_\lambda(\mathbf{T}) = \{\, x \in \mathbb{V} : (\mathbf{T} - \lambda I)^p(x) = \mathbf{0} \text{ for some positive integer } p \,\}.$$

Note that $K_\lambda(\mathbf{T})$ consists of the zero vector and all generalized eigenvectors corresponding to λ.

Theorem 7.1.1

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$ such that the characteristic polynomial of $\mathbf{T}$ splits. Then

1. If λ is an eigenvalue of $\mathbf{T}$ with (algebraic) multiplicity m, then

$$K_\lambda(\mathbf{T}) = \mathcal{N}((\mathbf{T} - \lambda I)^m) \text{ and } \mathbf{dim}(K_\lambda(\mathbf{T})) = m.$$

2. If $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of $\mathbf{T}$, then

$$\mathbb{V} = K_{\lambda_1}(\mathbf{T}) \oplus K_{\lambda_2}(\mathbf{T}) \oplus \ldots \oplus K_{\lambda_k}(\mathbf{T}).$$

Note that if $m(\lambda) = 1$, then we simply have $K_\lambda(\mathbf{T}) = E_\lambda$.

Theorem 7.1.2

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$ such that the characteristic polynomial of $\mathbf{T}$ splits, and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues of $\mathbf{T}$ with corresponding multiplicities $m_1, m_2, \cdots, m_k$. For $1 \leq i \leq k$, let β_i be an ordered basis for $K_{\lambda_i}(\mathbf{T})$. Then,

1. $\beta_i \cap \beta_j = \varnothing$ for $i \neq j$.
2. $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is an ordered basis for $\mathbb{V}$.
3. $\mathbf{dim}(K_{\lambda_i}(\mathbf{T})) = m_i$ for all i.

Definition 7.1.3

Let $\mathbf{T}$ be a linear operator on a vector space $\mathbb{V}$, and let x be a generalized eigenvector of $\mathbf{T}$ corresponding to the eigenvalue λ. Suppose that p is the smallest positive integer for which $(\mathbf{T}-\lambda I)^p x = \mathbf{0}$. Then the ordered sets

$$\left\{ (\mathbf{T}-\lambda I)^{p-1}x, (\mathbf{T}-\lambda I)^{p-2}x, \ldots, (\mathbf{T}-\lambda I)x, x \right\}$$

is called a cycle of generalized eigenvectors of $\mathbf{T}$ corresponding to λ. The vectors $(\mathbf{T}-\lambda I)^{p-1}x$ and x are called the **initial vector** and **end vector** of the cycle, respectively. We also say that the length of the cycle is p.

Note that the initial vector $(\mathbf{T}-\lambda I)^{p-1}x$ of a cycle of generalized eigenvectors of a linear operator $\mathbf{T}$ on a finite-dimensional vector space $\mathbb{V}$ is the only eigenvector of $\mathbf{T}$ in the cycle. On the other hand, a vector x is the end vector of such a cycle if and only if $(\mathbf{T}-\lambda I)^{p-1}x \neq \mathbf{0}$, but $(\mathbf{T}-\lambda I)^p x = \mathbf{0}$.

Moreover, if x is an eigenvector of $\mathbf{T}$ corresponding to the eigenvalue λ, then the set $\{x\}$ is a cycle of generalized eigenvectors of $\mathbf{T}$ corresponding to λ of length 1. If β is a disjoint union of cycles of generalized eigenvectors of $\mathbf{T}$, then β is a **Jordan canonical basis** for $\mathbb{V}$.

Theorem 7.1.3

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and let λ be an eigenvalue of $\mathbf{T}$. Then $K_\lambda(\mathbf{T})$ has an ordered basis consisting of a union of disjoint cycles of generalized eigenvectors corresponding to λ.

Corollary 7.1.1

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$ whose characteristic polynomial splits. Then $\mathbf{T}$ has a Jordan canonical form.

Remark 7.1.1

Let $A \in M_{n\times n}(\mathbb{F})$ such that the characteristic polynomial of A (and hence of L_A) splits. Then

1. The Jordan canonical form of A is defined to be the Jordan canonical form of the linear operator L_A on $\mathbb{F}^n$.
2. A has a Jordan canonical form J, and A is similar to J.

Section 7.2: Jordan Canonical Form II

Let $\mathbf{T}$ be a linear operator on an n-dimensional vector space $\mathbb{V}$ such that the characteristic polynomial of $\mathbf{T}$ splits. Let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be the distinct eigenvalues of $\mathbf{T}$.

By Theorem 7.1.3, each K_{λ_i} contains an ordered basis β_i consisting of a union of disjoint cycles of generalized eigenvectors corresponding to λ_i. Therefore, Theorem 7.1.2(2) implies that $\beta = \bigcup_{i=1}^{k} \beta_i$ is a Jordan canonical basis for $\mathbf{T}$.

For each i, let $\mathbf{T}_i$ be the restriction of $\mathbf{T}$ to K_{λ_i} and let $J_i = [\mathbf{T}_i]_{\beta_i}$. Then J_i is the Jordan canonical form of $\mathbf{T}_i$, and

$$J = [\mathbf{T}]_\beta = \begin{pmatrix} J_1 & O & \dots & O \\ O & J_2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & J_k \end{pmatrix}$$

is the Jordan canonical form of $\mathbf{T}$, where O is a zero matrix of appropriate size.

Remark 7.2.1

If β_i is a disjoint union of cycles (boxes) $\gamma_1, \gamma_2, \cdots, \gamma_{n_i}$ each of which of length p_i, then we index the cycles so that $p_1 \geq p_2 \geq \cdots \geq p_{n_i}$.

In this way, the number n_i of cycles for β_i, and the length p_j, $j = 1, \dots n_i$ of each cycle is completely determined by $\mathbf{T}$.

In other words, each Jordan block is decomposed into smaller boxes (cycles) each of which is of length p_i.

Example 7.2.1

Suppose that for some i, the ordered basis β_i for K_{λ_i} is the union of four cycles, i.e. $\beta_i = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4$ with respective lengths $p_1 = 3, p_2 = 3, p_3 = 2$ and $p_4 = 1$. Then

$$J_i = \left(\begin{array}{ccc|ccc|cc|c} \lambda_i & 1 & & & & & & & \\ & \lambda_i & 1 & & & & & & \\ & & \lambda_i & & & & & & \\ \hline & & & \lambda_i & 1 & & & & \\ & & & & \lambda_i & 1 & & & \\ & & & & & \lambda_i & & & \\ \hline & & & & & & \lambda_i & 1 & \\ & & & & & & & \lambda_i & \\ \hline & & & & & & & & \lambda_i \end{array}\right).$$

Remark 7.2.2

To visualize J_i and β_i, we use an array of dots called a **dot diagram** of $\mathbf{T}_i$, where $\mathbf{T}_i$ is the restriction of $\mathbf{T}$ to K_{λ_i}.

Let β_i be a disjoint union of cycles of generalized eigenvectors $\gamma_1, \gamma_2, \cdots, \gamma_{n_i}$ with lengths $p_1 \geq p_2 \geq \cdots \geq p_{n_i}$, respectively.

The dot diagram of $\mathbf{T}_i$ contains one dot for each vector in β_i, and the dots are configured with respect to the following reules.

1. The diagram consists of n_i column (one column for each cycle).
2. Counting from left to right, the j^{th} column consists of the p_j dots that corresponds to the vectors of γ_j starting with the initial vector at the top and continuing down to the end vector.

If $v_1, v_2, \cdots, v_{n_i}$ are the end vectors of the cycles $\gamma_1, \gamma_2, \cdots, \gamma_{n_i}$, then we get:

$$\begin{array}{llll}
\bullet(\mathbf{T}-\lambda_i I)^{p_1-1}v_1 & \bullet(\mathbf{T}-\lambda_i I)^{p_2-1}v_2 & \cdots & \bullet(\mathbf{T}-\lambda_i I)^{p_{n_i}-1}v_{n_i} \\
\bullet(\mathbf{T}-\lambda_i I)^{p_1-2}v_1 & \bullet(\mathbf{T}-\lambda_i I)^{p_2-2}v_2 & \cdots & \bullet(\mathbf{T}-\lambda_i I)^{p_{n_i}-2}v_{n_i} \\
\vdots & \vdots & & \vdots \\
 & & & \bullet(\mathbf{T}-\lambda_i I)v_{n_i} \\
 & & & \bullet v_{n_i} \\
 & \bullet(\mathbf{T}-\lambda_i I)v_2 & & \\
 & \bullet v_2 & & \\
\bullet(\mathbf{T}-\lambda_i I)v_1 & & & \\
\bullet v_1 & & &
\end{array}$$

Note that the dot diagram of $\mathbf{T}_i$ has n_i columns (one for each cycle) and p_1 rows.

Let r_j denote the number of dots in the j^{th} row of the dot diagram. Note that $r_1 \geq r_2 \geq \cdots \geq r_{p_1}$. Moreover, the dot diagram can be reconstructed from the values of the r_j's.

Back to Example 7.2.1

We have $n_i = 4$; $p_1 = p_2 = 3$, $p_3 = 2$ and $p_4 = 1$. The dot diagram of $\mathbf{T}_i$ is

$$\begin{array}{cccc}
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \\
\bullet & \bullet & &
\end{array}$$

Moreover, $r_1 = 4$, $r_2 = 3$, and $r_3 = 2$.

Remark 7.2.3

We compute the (unique) dot diagram of $\mathbf{T}_i$ using the ranks of linear operators determined by $\mathbf{T}$ and λ_i. In this way, the dot diagram is uniquely determined by $\mathbf{T}$.

To determine the dot diagram of $\mathbf{T}_i$, we devise a method for computing each r_i, the number of dots in the j^{th} row of the dot diagram, using only $\mathbf{T}$ and λ_i.

Theorem 7.2.1

For any positive integer r, the vectors in β_i associated with the dots in the first r rows of the dot diagram of $\mathbf{T}_i$ forms a basis for $\mathcal{N}((\mathbf{T} - \lambda_i I)^r)$.

That is, the number of dots in the first r rows of the dot diagram equals nullity$((\mathbf{T} - \lambda_i I)^r)$.

Theorem 7.2.2

Let r_j denote the number of dots in the j^{th} row of the dot diagram of $\mathbf{T}_i$, the restriction of $\mathbf{T}$ to K_{λ_i}. Then

1. $r_1 = \mathbf{dim}(\mathbb{V}) - \mathbf{rank}((\mathbf{T} - \lambda_i I))$.
2. $r_j = \mathbf{rank}((\mathbf{T} - \lambda_i I)^{j-1}) - \mathbf{rank}((\mathbf{T} - \lambda_i I)^j)$ for $j > 1$.

Remark 7.2.4 Computing Jordan Canonical Forms

Given a matrix $A \in M_{n\times n}(\mathbb{C})$, we use the following method to compute the Jordan canonical form of A:

1. Calculate eigenvalues of A: $\lambda_1, \lambda_2, \cdots, \lambda_k$.
2. For each λ_i $(i = 1, 2, \ldots, k)$: determine $m(\lambda_i)$ and r_j using Theorem 7.2.2.

Example 7.2.2 See Example 7.1.2

Let

$$A = \begin{pmatrix} 3 & 1 & -2 \\ -1 & 0 & 5 \\ -1 & -1 & 4 \end{pmatrix}.$$

Find the Jordan canonical form of A.

Solution:

1. **Eigenvalues**: We consider the characteristic polynomial of A to find all eigenvalues.

$$f_A(\lambda) = |\,A - \lambda I\,| = \begin{vmatrix} 3-\lambda & 1 & -2 \\ -1 & -\lambda & 5 \\ -1 & -1 & 4-\lambda \end{vmatrix}. = \cdots = -(3-\lambda)(2-\lambda)^2.$$

Therefore, we have $\lambda_1 = 3$ and $\lambda_2 = 2$ with multiplicities $m(\lambda_1) = 1$ and $m(\lambda_2) = 2$, respectively. Therefore, we can deduce that there are two Jordan blocks. One for λ_1 of size 1×1 and the other for λ_2 of size 2×2. That is,

$$J = \begin{pmatrix} J_1 & O \\ O & J_2 \end{pmatrix}.$$

2. **Multiplicities and r_j**: Next we compute r_j for λ_1 and λ_2:

$\boxed{\lambda_1 = 3}$ Here, we have a cycle of length 1. Clearly $m(\lambda_1) = 1$ and hence $\mathbf{dim}(K_{\lambda_1}) = 1$. Thus,

$$J_1 = (3).$$

$\boxed{\lambda_2 = 2}$ Here, we have a cycle of length 2. So, we start with $r_1 = 3 - \mathbf{rank}(A - 2I) = 3 - 2 = 1$, where

$$\mathbf{rank}(A - 2I) = \mathbf{rank}(\begin{pmatrix} 1 & 1 & -2 \\ -1 & -2 & 5 \\ -1 & -1 & 2 \end{pmatrix}) = \mathbf{rank}(\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}) = 2.$$

That is, the dot diagram has one dot in the first row. That leaves us with one more dot for r_2. But we compute it using Theorem 7.2.2 anyway. We have, $r_2 = \mathbf{rank}(A - 2I) - \mathbf{rank}((A - 2I)^2) = 2 - 1 = 1$, where

$$\mathbf{rank}((A - 2I)^2) = \mathbf{rank}(\begin{pmatrix} 2 & 1 & -1 \\ -4 & -2 & 2 \\ -2 & -1 & 1 \end{pmatrix})\mathbf{rank}(\begin{pmatrix} 1 & 0.5 & -0.5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}) = 1.$$

Therefore, our dot diagram is

$$\begin{matrix} \bullet \\ \bullet \end{matrix}$$

and we have one cycle of lengths 2. That is,

$$J_2 = \begin{pmatrix} 2 & 1 \\ & 2 \end{pmatrix}$$

Therefore,

$$J = \begin{pmatrix} J_1 & O \\ O & J_2 \end{pmatrix} = \left(\begin{array}{c|cc} 3 & & \\ \hline & 2 & 1 \\ & & 2 \end{array}\right).$$

Example 7.2.3

Let

$$A = \begin{pmatrix} 3 & 1 & 0 & 1 \\ -1 & 5 & 4 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Find the Jordan canonical form of A.

Solution:

1. **Eigenvalues**: We consider the characteristic polynomial of A to find all eigenvalues. Hint: Use the determinant of blocks!

$$f_A(\lambda) = |\,A - \lambda I\,| = \begin{vmatrix} 3-\lambda & 1 & 0 & 1 \\ -1 & 5-\lambda & 4 & 1 \\ 0 & 0 & 2-\lambda & 0 \\ 0 & 0 & 0 & 4-\lambda \end{vmatrix} = \cdots = (\,4-\lambda\,)^3(\,2-\lambda\,).$$

Therefore, we have $\lambda_1 = 2$ and $\lambda_2 = 4$ with multiplicities $m(\lambda_1) = 1$ and $m(\lambda_2) = 3$, respectively. Therefore, we can deduce that there are two Jordan blocks. One for λ_1 of size 1×1 and the other for λ_2 of size 3×3. That is,

$$J = \begin{pmatrix} J_1 & O \\ O & J_2 \end{pmatrix}.$$

2. **Multiplicities and** r_j: Next we compute r_j for λ_1 and λ_2:

$\boxed{\lambda_1 = 2}$ Here, we have a cycle of length 1. Clearly $m(\lambda_1) = 1$ and hence $\mathbf{dim}(K_{\lambda_1}) = 1$.

Thus,

$$J_1 = (2).$$

$\boxed{\lambda_2 = 4}$ Here, we have a cycle of length 3. So, we start with $r_1 = 4 - \textbf{rank}(A - 4I) = 4 - 2 = 2$, where

$$\textbf{rank}(A - 4I) = \textbf{rank}\left(\begin{pmatrix} -1 & 1 & 0 & 1 \\ -1 & 1 & 4 & 1 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right) = \cdots = 2.$$

That is, the dot diagram has two dots in the first row. That leaves us with one more dot for r_2. But we compute it using Theorem 7.2.2 anyway. We have, $r_2 = \textbf{rank}(A - 4I) - \textbf{rank}((A - 4I)^2) = 2 - 1 = 1$, where

$$\textbf{rank}((A - 4I)^2) = \textbf{rank}\left(\begin{pmatrix} 0 & 0 & 4 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right) = 1.$$

Therefore, our dot diagram is

$$\begin{array}{cc} \bullet & \bullet \\ \bullet & \end{array}$$

and we have two cycles of lengths 2 and 1. That is,

$$J_2 = \left(\begin{array}{cc|c} 4 & 1 & \\ & 4 & \\ \hline & & 4 \end{array}\right).$$

Therefore,

$$J = \begin{pmatrix} J_1 & O \\ O & J_2 \end{pmatrix} = \left(\begin{array}{c|cc|c} 2 & & & \\ \hline & 4 & 1 & \\ & & 4 & \\ \hline & & & 4 \end{array}\right).$$

Remark 7.2.5 Computing Jordan Canonical Basis

Given a matrix $A \in M_{n\times n}(\mathbb{C})$, we use the following method to compute the Jordan canonical basis of A so that A is decomposed as $A = Q\,J\,Q^{-1}$:

1. Compute eigenvalues of A: $\lambda_1, \lambda_2, \cdots, \lambda_k$.
2. For each λ_i $(i = 1, 2, \ldots, k)$:
   (a) determine $m(\lambda_i)$ and r_j using Theorem 7.2.2.
   (b) Calculate eigenspace and generalized eigenspace
3. Construct Q.

Example 7.2.4

Let

$$A = \begin{pmatrix} 3 & 1 & -2 \\ -1 & 0 & 5 \\ -1 & -1 & 4 \end{pmatrix}.$$

Find the Jordan canonical form of A and find a Jordan canonical basis for the linear operator $\mathbf{T} = L_A$.

Solution:

1. **Eigenvalues**: We consider the characteristic polynomial of A to find all eigenvalues.

$$f_A(\lambda) = |\,A - \lambda I\,| = \begin{pmatrix} 3-\lambda & 1 & -2 \\ -1 & -\lambda & 5 \\ -1 & -1 & 4-\lambda \end{pmatrix} = \cdots = -(\,2-\lambda\,)^2(\,3-\lambda\,).$$

Therefore, we have $\lambda_1 = 2$ and $\lambda_2 = 3$ with multiplicities $m(\lambda_1) = 2$ and $m(\lambda_2) = 1$, respectively. Therefore, we can deduce that there are two Jordan blocks. One for λ_1 of size 2×2 and the other for λ_2 of size 1×1. That is,

$$J = \begin{pmatrix} J_1 & O \\ O & J_2 \end{pmatrix}.$$

Also, let $\mathbf{T}_1$ and $\mathbf{T}_2$ denote the restrictions of L_A to the generalized eigenspaces K_{λ_1} and K_{λ_2}, respectively.

2. $\boxed{\lambda_1 = 2}$

(a) **Multiplicities and** r_j: We have $m(\lambda_1) = 2$ and hence if β_1 is a Jordan canonical basis for $\mathbf{T}_1$, then $\mathbf{dim}(K_{\lambda_1}) = 2$ and hence the dot diagram for λ_1 has two dots. Computing $r_1 = 3 - \mathbf{rank}(A - 2I) = 3 - 2 = 1$, where

$$\mathbf{rank}(A-2I) = \mathbf{rank}(\begin{pmatrix} 1 & 1 & -2 \\ -1 & -2 & 5 \\ -1 & -1 & 2 \end{pmatrix}) = \mathbf{rank}(\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}) = 2.$$

That is, the dot diagram has one dot in the first row. That leaves us with one more dot for r_2. But we compute it using Theorem 7.2.2 anyway. We have, $r_2 = \mathbf{rank}(A - 2I) - \mathbf{rank}((A - 2I)^2) = 2 - 1 = 1$, where

$$\mathbf{rank}((A-2I)^2) = \mathbf{rank}(\begin{pmatrix} 2 & 1 & -1 \\ -4 & -2 & 2 \\ -2 & -1 & 1 \end{pmatrix}) = \mathbf{rank}(\begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}) = 1.$$

Therefore, our dot diagram is

$$\begin{matrix} \bullet \\ \bullet \end{matrix}$$

and we have one cycle of lengths 2. That is,

$$J_1 = \begin{pmatrix} 2 & 1 \\ & 2 \end{pmatrix}$$

(b) **Eigenspace and generalized eigenspace** Here we determine a Jordan canonical basis β_1 for $\mathbf{T}_1$. The dot diagram of $\mathbf{T}_1$ has one column which corresponds to the cycle of generalized eigenvectors. Let v_1 denote the end vector of the this cycle. Then we have

$$\begin{matrix} \bullet\ (\mathbf{T} - 2I)v_1 \\ \bullet\ v_1 \end{matrix}$$

Note that $(\mathbf{T} - 2I)^2 v_1 = \mathbf{0}$ but $(\mathbf{T} - 2I)v_1 \neq \mathbf{0}$. That is, $v_1 \in \mathcal{N}(\,(\mathbf{T} - 2I)^2\,)$ but $v_1 \notin \mathcal{N}(\,\mathbf{T} - 2I\,)$. Thus, we compute bases for $\mathcal{N}(\,\mathbf{T} - 2I\,)$ and $\mathcal{N}(\,(\mathbf{T} - 2I)^2\,)$. That is, we solve the systems $(A - 2I)x = \mathbf{0}$ and $(A - 2I)^2 x = \mathbf{0}$. For $\mathcal{N}(\,\mathbf{T} - 2I\,)$, we have

$$\mathcal{N}(\,A - 2I\,) = \left\{ (a,b,c) : \begin{pmatrix} 1 & 1 & -2 \\ -1 & -2 & 5 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ -1 & -2 & 5 & 0 \\ -1 & -1 & 2 & 0 \end{array}\right] \xrightarrow{\text{r.r.e.f}} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus, $a = -c$ and $b = 3c$. Hence,

$$\gamma_1 = \left\{ \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix} \right\}$$

is a basis for $\mathcal{N}(A - 2I)$. Note that $\mathcal{N}(A - 2I) \subseteq \mathcal{N}(\,(A - 2I)^2\,)$. Now, We do the same thing for $\mathcal{N}(\,(A - 2I)^2\,)$.

$$\mathcal{N}\Big(\,(A - 2I)^2\Big) = \left\{ (a, b, c) : \begin{pmatrix} 2 & 1 & -1 \\ -4 & -2 & 2 \\ -2 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

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

Thus, $2a + b - c = 0$. Hence, $\gamma_1 \cup \gamma_2$ is a basis for $\mathcal{N}(\,(A - 2I)^2\,)$, where

$$\gamma_1 \cup \gamma_2 = \gamma_1 \cup \left\{ \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\}$$

Now, we choose v_1 from the basis produced in $\mathcal{N}(\,(A - 2I)^2\,)$; namely from γ_2. That is, $v_1 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$. Hence,

$$(\,A - 2I\,)v_1 = \begin{pmatrix} 1 & 1 & -2 \\ -1 & -2 & 5 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix}.$$

Thus, we have constructed the Jordan canonical basis β_1 consisting of one cycles for the Jordan block of $\lambda_1 = 2$.

$$\beta_1 = \{ (A - 2I)v_1, v_1 \} = \left\{ \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\}.$$

$\boxed{\lambda_2 = 3}$

(a) **Multiplicities and** r_j: We have $m(\lambda_2) = 1$. Here, we have a cycle of length 1. Clearly $m(\lambda_2) = 1$ and hence $\mathbf{dim}(K_{\lambda_2}) = 1$. Thus,

$$J_2 = (3).$$

(b) **Eigenspace and generalized eigenspace**: Note that $\mathbf{dim}(K_{\lambda_2}) = 1 = \mathbf{dim}(E_{\lambda_2})$. Therefore, $K_{\lambda_2} = E_{\lambda_2}$ and hence any eigenvector of L_A corresponding to λ_2 would form the basis β_2. For $\mathcal{N}(\mathbf{T} - 3I)$, we have

$$\mathcal{N}(A - 3I) = \left\{ (a, b, c, d) : \begin{pmatrix} 0 & 1 & -2 \\ -1 & -3 & 5 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{ccc|c} 0 & 1 & -2 & 0 \\ -1 & -3 & 5 & 0 \\ -1 & -1 & 1 & 0 \end{array} \right] \xrightarrow{\text{r.r.e.f.}} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Thus, $a = -c$ and $b = c$. Hence,

$$\beta_2 = \left\{ \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

is a basis for $\mathcal{N}(A - 3I)$.

3. **Construct** J **and** Q We have already computed J_1 and J_2 and hence

$$J = \begin{pmatrix} J_1 & O \\ O & J_2 \end{pmatrix} = \left(\begin{array}{cc|c} 2 & 1 & \\ & 2 & \\ \hline & & 3 \end{array} \right).$$

Moreover, we compute the invertible matrix Q whose columns are the vectors of the **ordered** basis $\beta = \beta_1 \cup \beta_2$. That is,

$$\beta = \left\{ \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \right\},$$

and

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

would satisfies $J = Q^{-1}AQ$ or $QJ = AQ$.

Example 7.2.5

Let

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

Find the Jordan canonical form of A and find a Jordan canonical basis for the linear operator $\mathbf{T} = L_A$.

Solution:

1. **Eigenvalues**: We consider the characteristic polynomial of A to find all eigenvalues.

$$f_A(\lambda) = |\, A - \lambda I \,| = \begin{vmatrix} 2-\lambda & -1 & 0 & 1 \\ 0 & 3-\lambda & -1 & 0 \\ 0 & 1 & 1-\lambda & 0 \\ 0 & -1 & 0 & 3-\lambda \end{vmatrix} = \cdots = (\,2-\lambda\,)^3(\,3-\lambda\,).$$

Therefore, we have $\lambda_1 = 2$ and $\lambda_2 = 3$ with multiplicities $m(\lambda_1) = 3$ and $m(\lambda_2) = 1$, respectively. Therefore, we can deduce that there are two Jordan blocks. One for λ_1 of size 3×3 and the other for λ_2 of size 1×1. That is,

$$J = \begin{pmatrix} J_1 & O \\ O & J_2 \end{pmatrix}.$$

Also, let $\mathbf{T}_1$ and $\mathbf{T}_2$ denote the restrictions of L_A to the generalized eigenspaces K_{λ_1} and K_{λ_2}, respectively.

2. $\boxed{\lambda_1 = 2}$

   (a) **Multiplicities and** r_j: We have $m(\lambda_1) = 3$ and hence if β_1 is a Jordan canonical basis for $\mathbf{T}_1$, then $\mathbf{dim}(K_{\lambda_1}) = 3$ and hence the dot diagram for λ_1 has three dots. Computing $r_1 = 4 - \mathbf{rank}(A - 2I) = 4 - 2 = 2$, where

$$\mathbf{rank}(A - 2I) = \mathbf{rank}(\begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}) = 2.$$

   Computing $r_2 = \mathbf{rank}(A - 2I) - \mathbf{rank}((A - 2I)^2) = 2 - 1 = 1$, where

$$\mathbf{rank}((A - 2I)^2) = \mathbf{rank}(\begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}) = 1.$$

   Therefore, our dot diagram is

$$\begin{matrix} \bullet & \bullet \\ \bullet & \end{matrix}$$

   and we have two cycles (correspond to vertical dots) of lengths 2 and 1. That is,

$$J_1 = [\mathbf{T}_1]_{\beta_1} = \left(\begin{array}{cc|c} 2 & 1 & \\ & 2 & \\ \hline & & 2 \end{array}\right).$$

   (b) **Eigenspace and generalized eigenspace** Here we determine a Jordan canonical basis β_1 for $\mathbf{T}_1$. The dot diagram of $\mathbf{T}_1$ has two columns, each of which corresponds to a cycle of generalized eigenvectors. Let v_1 and v_2 denote the end vectors of the first and second cycles, respectively. Then we have

$$\begin{matrix} \bullet\ (\mathbf{T} - 2I)v_1 & \bullet\ v_2 \\ \bullet\ v_1 & \end{matrix}$$

   Note that $(\mathbf{T} - 2I)^2 v_1 = \mathbf{0}$ but $(\mathbf{T} - 2I)v_1 \neq \mathbf{0}$. That is, $v_1 \in \mathcal{N}(\,(\mathbf{T} - 2I)^2\,)$ but $v_1 \notin \mathcal{N}(\,\mathbf{T} - 2I\,)$. Thus, we compute bases for $\mathcal{N}(\,\mathbf{T} - 2I\,)$ and $\mathcal{N}(\,(\mathbf{T} - 2I)^2\,)$.

That is, we solve the systems $(A-2I)x = \mathbf{0}$ and $(A-2I)^2 x = \mathbf{0}$. For $\mathcal{N}(\,\mathbf{T}-2I\,)$, we have

$$\mathcal{N}(\,A-2I\,) = \left\{ (a,b,c,d) : \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

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

Thus, $b = d$ and $c = d$. Hence,

$$\gamma_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

is a basis for $\mathcal{N}(A-2I)$. Note that $\mathcal{N}(A-2I) \subseteq \mathcal{N}(\,(A-2I)^2\,)$. Now, We do the same thing for $\mathcal{N}(\,(A-2I)^2\,)$.

$$\mathcal{N}\Big(\,(A-2I)^2\,\Big) = \left\{ (a,b,c,d) : \begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

$$\left[\begin{array}{cccc|c} 0 & -2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 & 0 \end{array}\right] \xrightarrow{\text{r.r.e.f}} \left[\begin{array}{cccc|c} 0 & 1 & -0.5 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Thus, $b = \frac{1}{2}c + \frac{1}{2}d$. Therefore, $\gamma_1 \cup \gamma_2$ is a basis for $\mathcal{N}(\,(A-2I)^2\,)$, where

$$\gamma_1 \cup \gamma_2 = \left\{ \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\2\\0\end{pmatrix} \right\}$$

Now, we choose v_1 from the basis produced in $\mathcal{N}(\,(A-2I)^2\,)$; namely from γ_2.

That is, $v_1 = \begin{pmatrix}0\\1\\2\\0\end{pmatrix}$. Hence,

$$(\,A-2I\,)v_1 = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix}0\\1\\2\\0\end{pmatrix} = \begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}.$$

For v_2, we simply choose a vector in $\mathcal{N}(A-2I)$ but linearly independent of $(\,A-2I\,)v_1$. One choice from γ_1 is

$$v_2 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \qquad \text{, another option is } v_2 = \begin{pmatrix}0\\1\\1\\1\end{pmatrix}.$$

Thus, we have constructed the Jordan canonical basis β_1 consisting of two cycles for the Jordan block of $\lambda_1 = 2$.

$$\beta_1 = \{\,(\,A-2I\,)v_1, v_1, v_2\,\} = \left\{ \begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}, \begin{pmatrix}0\\1\\2\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \right\}.$$

$\boxed{\lambda_2 = 3}$

(a) **Multiplicities and** r_j: We have $m(\lambda_2) = 1$. Here, we have a cycle of length 1. Clearly $m(\lambda_2) = 1$ and hence $\mathbf{dim}(K_{\lambda_2}) = 1$. Thus,

$$J_2 = (3).$$

(b) **Eigenspace and generalized eigenspace**: Note that $\mathbf{dim}(K_{\lambda_2}) = 1 = \mathbf{dim}(E_{\lambda_2})$. Therefore, $K_{\lambda_2} = E_{\lambda_2}$ and hence any eigenvector of L_A corresponding to λ_2 would form the basis β_2. For $\mathcal{N}(\,\mathbf{T} - 3I\,)$, we have

$$\mathcal{N}(\,A - 3I\,) = \left\{ (a,b,c,d) : \begin{pmatrix} -1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This is can be solved as follows:

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

Thus, $a = d$ and $b = c = 0$. Hence,

$$\beta_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is a basis for $\mathcal{N}(A - 3I)$.

3. **Construct J and Q** We have already computed J_1 and J_2 and hence

$$J = \begin{pmatrix} J_1 & O \\ O & J_2 \end{pmatrix} = \left(\begin{array}{cc|c|c} 2 & 1 & & \\ & 2 & & \\ \hline & & 2 & \\ \hline & & & 3 \end{array}\right).$$

Moreover, we compute the invertible matrix Q whose columns are the vectors of the **ordered** basis $\beta = \beta_1 \cup \beta_2$. That is,

$$\beta = \left\{ \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

and

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

would satisfies $J = Q^{-1}AQ$ or $QJ = AQ$.

Exercise 7.2.1

Let $\mathbf{T}$ be the linear operator on $\mathbb{P}_2(\mathbb{R})$ defined by $\mathbf{T}(f(x)) = -f(x) - f'(x)$. Find a Jordan canonical form of $\mathbf{T}$.

Exercise 7.2.2

Let

$$A = \begin{pmatrix} 2 & 1 & -1 & 8 & -3 \\ 0 & 2 & 0 & 7 & 5 \\ 0 & 0 & 2 & 7 & 5 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Find the Jordan canonical form of A.

Exercise 7.2.3

Let

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

Find the Jordan canonical form of A.

Exercise 7.2.4

Let

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

Find the Jordan canonical form of A and find a Jordan canonical basis for the linear operator $\mathbf{T} = L_A$.

Section 7.3: Minimal Polynomials

By the Cayley-Hamilton Theorem, for any linear operator $\mathbf{T}$ on an n-dimensional vector space $\mathbb{V}$ there is a polynomial $f(t)$ of degree n such that $f(\mathbf{T}) = \mathbf{T}_0$; namely, the characteristic polynomial of $\mathbf{T}$.

Hence there is a polynomial of least degree with this property. This degree is at most n.

If $g(t)$ is such a polynomial, we can derive another polynomial $p(t)$ of the same degree with leading coefficient 1, by dividing $g(t)$ by its leading coefficient. In that case, we say that $p(t)$ is a monic polynomial.

Definition 7.3.1

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space. A polynomial $p(t)$ is called a **minimal polynomial** of $\mathbf{T}$ if $p(t)$ is a monic polynomial of least positive degree for which $p(\mathbf{T}) = \mathbf{T}_0$.

Theorem 7.3.1

Let $p(t)$ be a minimal polynomial of a linear operator $\mathbf{T}$ on a finite-dimensional vector space $\mathbb{V}$. Then

1. For any polynomial $g(t)$, if $g(\mathbf{T}) = \mathbf{T}_0$, then $p(t)$ divides $g(t)$. In particular, $p(t)$ divides the characteristic polynomial of $\mathbf{T}$.
2. The minimal polynomial of $\mathbf{T}$ is unique.

The minimal polynomial of a linear operator has an obvious analog for a matrix.

Definition 7.3.2

Let $A \in M_{n\times n}(\mathbb{F})$. The **minimal polynomial** of A is the monic polynomial of least positive degree for which $p(A) = O$, where O is the zero matrix.

Theorem 7.3.2

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and let β be an ordered basis for $\mathbb{V}$. Then, the minimal polynomial of $\mathbf{T}$ is the same as the minimal polynomial of the matrix $[\mathbf{T}]_\beta$.

In particular, for any matrix $A \in M_{n\times n}(\mathbb{F})$, the minimal polynomial of A is the same as the minimal polynomial of L_A.

Theorem 7.3.3

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$, and let $p(t)$ bt the minimal polynomial of $\mathbf{T}$. A scalar λ is an eigenvalue of $\mathbf{T}$ if and only if $p(\lambda) = 0$. Hence the characteristic polynomial and the minimal polynomial of $\mathbf{T}$ have the same size.

Proof:

„ $\Rightarrow$" Suppose that λ is an eigenvalue of $\mathbf{T}$ corresponding to eigenvector x. That is, $\mathbf{T}(x) = \lambda x$. Then,

$$0 = \mathbf{T}_0(x) = p(\mathbf{T})(x) = p(\lambda)(x).$$

Since $x \neq \mathbf{0}$, we get $p(\lambda) = 0$ and so λ is a zero of $p(t)$.

„ $\Leftarrow$" Let $f(t)$ be the characteristic polynomial of $\mathbf{T}$. Since $p(t)$ divides $f(t)$, there exists a polynomial $g(t)$ such that $f(t) = g(t)p(t)$. Now, if λ is a zero of $p(t)$, then

$$f(\lambda) = g(\lambda)p(\lambda) = g(\lambda)\cdot 0 = 0.$$

That is, λ is a zero of $f(t)$ and hence it is an eigenvalue of $\mathbf{T}$.

Corollary 7.3.1

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$ with minimal polynomial $p(t)$ and a characteristic polynomial $f(t)$. Suppose that $f(t)$ factors as

$$f(t) = (\lambda_1 - t)^{n_1}(\lambda_2 - t)^{n_2}\cdots(\lambda_k - t)^{n_k},$$

where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the distinct eigenvalues of $\mathbf{T}$. Then there exist integers $m_1, m_2, \cdots, m_k$ such that $1 \leq m_i \leq n_i$ for all i and

$$p(t) = (\lambda_1 - t)^{m_1}(\lambda_2 - t)^{m_2}\cdots(\lambda_k - t)^{m_k}.$$

Example 7.3.1

Find the minimal polynomial of

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

Solution:

Clearly the characteristic polynomial of A is

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

Thus, the minimal polynomial of A must be either $(\,t-2\,)(\,t-3\,)$ or $(\,t-2\,)^2(\,t-3\,)$. Substituting A into $p(t) = (\,t-2\,)(\,t-3\,)$, we find that

$$p(A) = (\,A - 2I\,)(\,A - 3I\,) = 0.$$

Hence $p(t) = (\,t-2\,)(\,t-3\,)$ is the minimal polynomial of A.

Example 7.3.2

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^2$ defined by

$$\mathbf{T}(a, b) = (\,2a + 5b, 6a + b\,),$$

and β be the standard ordered basis for $\mathbb{R}^2$. Find the minimal polynomial of $\mathbf{T}$.

Solution:

Clearly, $A := [\mathbf{T}]_\beta = \begin{pmatrix} 2 & 5 \\ 6 & 1 \end{pmatrix}$ and hence the characteristic polynomial of $\mathbf{T}$ is

$$f(t) = |\, A - tI_2 \,| = \begin{vmatrix} 2-t & 5 \\ 6 & 1-t \end{vmatrix} = \cdots = (\,t-7\,)(\,t+4\,).$$

Therefore, the minimal polynomial of $\mathbf{T}$ is clearly $p(t) = (\,t-7\,)(\,t+4\,)$.

Example 7.3.3

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

$$\mathbf{T}(g(t)) = g'(t).$$

Find the minimal polynomial of $\mathbf{T}$.

Solution:

Using the standard ordered basis $\beta = \{1, x, x^2\}$, we clearly find

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

Then

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

That is, $p(t)$ is either t, t^2, or t^3.

Note that

$$p(A^2) = A^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0.$$

But $p(A^3) = 0$ and hence the minimal polynomial of $\mathbf{T}$ is t^3.

Theorem 7.3.4

Let $\mathbf{T}$ be a linear operator on a finite-dimensional vector space $\mathbb{V}$. Then $\mathbf{T}$ is diagonalizable if and only if the minimal polynomial of $\mathbf{T}$ is of the form

$$p(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_k),$$

where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the distinct eigenvalues of $\mathbf{T}$.

Example 7.3.4

Determine all matrices $A \in M_{2\times 2}(\mathbb{R})$ for which $A^2 - 3A + 2I_2 = 0$.

Solution:

Let $g(t) = t^2 - 3t + 2 = (t-1)(t-2)$. Since $g(A) = 0$, the minimal polynomial $p(t)$ of A divides $g(t)$.

Therefore, the only possible candidates for $p(t)$ are $t-1$, $t-2$, and $(t-1)(t-2)$.

If $p(t) = t-1$ or $p(t) = t-2$, then $A = I$ or $A = 2I$, respectively. If $p(t) = (t-1)(t-2)$, then A is diagonalizable with eigenvalues 1 and 2, and hence A is similar to $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

Example 7.3.5

Let $A \in M_{n\times n}(\mathbb{R})$ with $A^3 = A$. Show that A is diagonalizable.

Solution:

Let $g(t) = t^3 - t = t(t+1)(t-1)$. Then $g(A) = 0$ and hence the minimal polynomial $p(t)$ of A divides $g(t)$.

Since $g(t)$ has distinct zeros, so $p(t)$ has distinct zeros as well. Thus, A is diagonalizable.

Remark 7.3.1

The operator $\mathbf{T}$ on $\mathbb{P}_2(\mathbb{R})$ of Example 7.3.3 is not diagonalizable. This is because its minimal polynomial is t^3 and hence Theorem 7.3.4 implies that $\mathbf{T}$ is not diagonalizable.

Exercise 7.3.1

Find the minimal polynomial of

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Exercise 7.3.2

Find the minimal polynomial of

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

Exercise 7.3.3

Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^2$ defined by

$$\mathbf{T}((a,b)) = (a+b, a-b).$$

Find the minimal polynomial of $\mathbf{T}$.

Exercise 7.3.4

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

$$\mathbf{T}(g(t)) = g'(t) + 2g(x).$$

Find the minimal polynomial of $\mathbf{T}$.

The Index