Solutions to Linear Algebra

(Friedberg; Insel; Spence 4/e)

Author:
Cheng-Mao Lee

Email:
chengmao.lee@gmail.com

2016/02/04 (Version 1.0)

Contents

Abstract

These solutions were done as a student who is fresh out of university. Now I'm studying the book when doing my mandatory military service.

The solutions might be brief. However, I state them as clearly as possible I can. For the computation part, I might omit details because it is tedious but simple.

This manual now only contains some selected solutions which I can afford. If there is something vague or incredible, it is possible that it doesn't make sense since it is wrong. If you have any suggestions or corrections to solutions, please direct to email "chengmao.lee@gmail.com". Your intelligence will be highly appreciated.

Hope this manual could share some constructive ideas or help you that given important hints to solve problems by yourselves, or just verify whether the answers are consistent or not.

Any way, I hope everybody has good luck and fun in solving problems!

Chapter 1

Vector Spaces

1.1 Introduction

Exercise 1.1.1

(a) **Answer.** No.

Solution. Since there is no real number $k \neq 0$ such that $(3,1,2) = k(6,4,2)$. ■

(b) **Answer.** Yes.

Solution. Since $(9,-3,-21) = -3(-3,1,7)$. ■

(c) **Answer.** Yes.

Solution. Since $(5,-6,7) = -1(-5,6,-7)$. ■

(d) **Answer.** No.

Solution. Since there is no real number $k \neq 0$ such that $(2,0,-5) = k(5,0,-2)$. ■

Exercise 1.1.2

(a) **Answer.** $(3,-2,4) + t(-8,9,-3)$ for $t \in \mathbb{R}$.

Solution. Notice that $(-8,9,-3) = (-5,7,1) - (3,-2,4)$. ■

(b) **Answer.** $(2,4,0) + t(-5,-10,0)$ for $t \in \mathbb{R}$.

Solution. Notice that $(-5,-10,0) = (-3,-6,0) - (2,4,0)$. ■

(c) **Answer.** $(3,7,2) + t(0,0,-10)$ for $t \in \mathbb{R}$.

Solution. Notice that $(0,0,-10) = (3,7,-8) - (3,7,2)$. ■

(d) **Answer.** $(-2,-1,5) + t(5,10,2)$ for $t \in \mathbb{R}$.

Solution. Notice that $(5,10,-2) = (3,9,7) - (-2,-1,5)$. ■

Exercise 1.1.3

(a) **Answer.** $(2,-5,-1) + s(-2,9,7) + t(-5,12,2)$ for $s,t \in \mathbb{R}$.

Solution. Let $A := (2,-5,-1), B := (0,4,6), C := (-3,7,1)$.
Since $\overrightarrow{AB} = (-2,9,7)$ and $\overrightarrow{AC} = (-5,12,2)$, then the equation of the plane is $A + s\overrightarrow{AB} + t\overrightarrow{AC} = (2,-5,-1) + s(-2,9,7) + t(-5,12,2)$ for $s,t \in \mathbb{R}$. ■

(b) **Answer.** $(3,-6,7) + s(-5,6,-11) + t(2,-3,-9)$ for $s,t \in \mathbb{R}$.

Solution. Let $A := (3,-6,7), B := (-2,0,-4), C := (5,-9,-2)$.
Since $\overrightarrow{AB} = (-5,6,-11)$ and $\overrightarrow{AC} = (2,-3,-9)$, then the equation of the plane is $A + s\overrightarrow{AB} + t\overrightarrow{AC} = (3,-6,7) + s(-5,6,-11) + t(2,-3,-9)$ for $s,t \in \mathbb{R}$. ■

(c) **Answer.** $(-8,2,0) + s(9,1,0) + t(14,-7,0)$ for $s,t \in \mathbb{R}$.

Solution. Let $A := (-8,2,0), B := (1,3,0), C := (6,-5,0)$.
Since $\overrightarrow{AB} = (9,1,0)$ and $\overrightarrow{AC} = (14,-7,0)$, then the equation of the plane is $A + s\overrightarrow{AB} + t\overrightarrow{AC} = (-8,2,0) + s(9,1,0) + t(14,-7,0)$ for $s,t \in \mathbb{R}$. ■

(d) **Answer.** $(1,1,1) + s(4,4,4) + t(-7,3,1)$ for $s,t \in \mathbb{R}$.

Solution. Let $A := (1,1,1), B := (5,5,5), C := (-6,4,2)$.
Since $\overrightarrow{AB} = (4,4,4)$ and $\overrightarrow{AC} = (-7,3,1)$, then the equation of the plane is $A + s\overrightarrow{AB} + t\overrightarrow{AC} = (1,1,1) + s(4,4,4) + t(-7,3,1)$ for $s,t \in \mathbb{R}$. ■

Exercise 1.1.4

Answer. $(0,0,\cdots,0) \in \mathbb{R}^n$.

Proof. Let $x := (a_1, a_2, \cdots, a_n) \in \mathbb{R}^n$, and $\mathbf{0} := (0, 0, \cdots, 0) \in \mathbb{R}^n$, then

$$x + \mathbf{0} = (a_1 + 0, a_2 + 0, \cdots, a_n + 0) = (a_1, a_2, \cdots, a_n) = x.$$

Hence $\mathbf{0}$ is the zero vector. ■

Exercise 1.1.5

Proof. Since the vector $x = (a_1, a_2) - (0, 0) = (a_1, a_2)$, then $tx = t(a_1, a_2) = (ta_1, ta_2)$. Because $(ta_1, ta_2) - (0, 0) = (ta_1, ta_2) = tx$, we conclude the vector tx that emanates from the origin terminates at the point with coordinates (ta_1, ta_2). ■

Exercise 1.1.6

Proof. The vector emanates from (a, b) terminates at the midpoint $\frac{1}{2}(c - a, d - b)$, so the coordinates of the midpoint is

$$(a, b) + \frac{1}{2}(c - a, d - b) = \left(\frac{a + c}{2}, \frac{b + d}{2}\right).$$

■

Exercise 1.1.7

Proof. Consider $\square ABCD$. Notice that $\overrightarrow{BC} = \overrightarrow{AD}$ since the property of the parallelogram. Then

$$\overline{AC} = \left\{t\overrightarrow{AC} : 0 \le t \le 1\right\} = \left\{t\left(\overrightarrow{AB} + \overrightarrow{BC}\right) : 0 \le t \le 1\right\} = \left\{t\left(\overrightarrow{AB} + \overrightarrow{AD}\right) : 0 \le t \le 1\right\};$$
$$\overline{BD} = \left\{\overrightarrow{AB} + s\left(\overrightarrow{AD} - \overrightarrow{AB}\right) : 0 \le s \le 1\right\}.$$

Hence the intersection M of $\overline{AC}$ and $\overline{BD}$ is

$$t\left(\overrightarrow{AB} + \overrightarrow{AD}\right) = \overrightarrow{AB} + s\left(\overrightarrow{AD} - \overrightarrow{AB}\right)$$
$$\implies (1 - s - t)\overrightarrow{AB} = (t - s)\overrightarrow{AD}.$$

Since $\overrightarrow{AB} \nparallel \overrightarrow{AD}$, then

$$1 - s - t = 0 \text{ and } t - s = 0;$$

therefore

$$s = t = \frac{1}{2}.$$

Notice that $0 \le s, t \le 1$, so M exists.

Denote $A = (x_a, y_a)$, $B = (x_b, y_b)$, $D = (x_d, y_d)$. We know the coordinates of M is

$$A + \frac{1}{2}\overrightarrow{AC} = \left(\frac{x_c - x_a}{2}, \frac{y_c - y_a}{2}\right).$$

By Exercise 1.1.6, we know M bisects $\overline{AC}$. A similar argument establishes M bisects $\overline{BD}$.

Finally, we conclude the diagonals of a parallelogram bisect each other. ■

1.2 Vector Spaces

Exercise 1.2.1

(a) **Answer.** True.

Solution. This follows from the definition $(VS\ 3)$. ■

(b) **Answer.** False.

Solution. This is by corollary 1 of Theorem 1.1. The proof could reference Exercise 1.2.9. ■

(c) **Answer.** False.

Solution. If $x = 0$, then a might not equal to b. ■

(d) **Answer.** False.

Solution. If $a = 0$, then x might not equal to y. ■

(e) **Answer.** True.

Solution. This is the definition. ■

(f) **Answer.** False.

Solution. It has m rows and n columns. ■

(g) **Answer.** False.

Solution. Since x and x^2 are members of $P(F)$, then the sum of them is $x^2 + x$; however, x has degree 1 and x^2 has degree 2 which implies x and x^2 have different degrees. ■

(h) **Answer.** False.

Solution. x and $-x$ are polynomials of degree 1, but $x + (-x) = 0$ is a polynomial of degree -1. ■

(i) **Answer.** True.

Proof. Let $f(x) = a_n x^n + \cdots + a_0$ is a polynomial of degree n, and c is a nonzero scalar. Notice that $a_n \neq 0$. Consider

$$cf(x) = ca_n x^n + \cdots + ca_0.$$

Since $ca_n \neq 0$, then $cf(x)$ is also a polynomial of degree n by definition. ■

(j) **Answer.** True.

Solution. This is by definition directly. ■

(k) **Answer.** True.

Solution. This is the definition. ■

Exercise 1.2.2

Answer. $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Exercise 1.2.3

Solution. $M_{13} = 3$, $M_{21} = 4$, $M_{22} = 5$. ■

Exercise 1.2.4

(a) **Answer.** $\begin{pmatrix} 6 & 3 & 2 \\ -4 & 3 & 9 \end{pmatrix}$.

(b) **Answer.** $\begin{pmatrix} 1 & -1 \\ 3 & -5 \\ 3 & 8 \end{pmatrix}$.

(c) **Answer.** $\begin{pmatrix} 8 & 20 & -12 \\ 4 & 0 & 28 \end{pmatrix}$.

(d) **Answer.** $\begin{pmatrix} 30 & -20 \\ -15 & 10 \\ -5 & -40 \end{pmatrix}$.

(e) **Answer.** $2x^4 + x^3 + 2x^2 - 2x + 10$.

(f) **Answer.** $-x^3 + 7x^2 + 4$.

(g) **Answer.** $10x^7 - 30x^4 + 40x^2 - 15x$.

(h) **Answer.** $3x^5 - 6x^3 + 12x + 6$.

Exercise 1.2.5

Solution. The upstream crossing matrix is $\begin{pmatrix} 8 & 3 & 1 \\ 3 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix}$.

The downstream crossing matrix is $\begin{pmatrix} 9 & 1 & 4 \\ 3 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}$.

The total of crossing matrix is $\begin{pmatrix} 8 & 3 & 1 \\ 3 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 9 & 1 & 4 \\ 3 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 17 & 4 & 5 \\ 6 & 0 & 0 \\ 4 & 1 & 0 \end{pmatrix}$. ■

Exercise 1.2.6

Solution.

$$2M - A = 2\begin{pmatrix} 4 & 2 & 1 & 3 \\ 5 & 1 & 1 & 4 \\ 3 & 1 & 2 & 6 \end{pmatrix} - \begin{pmatrix} 5 & 3 & 1 & 2 \\ 6 & 2 & 1 & 5 \\ 1 & 0 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 & 4 \\ 4 & 0 & 1 & 3 \\ 5 & 2 & 1 & 9 \end{pmatrix}.$$

Sum of all entries is 34; hence 34 suites were sold during the June Sale. ■

Exercise 1.2.7

Proof. Since $f(0) = 1 = g(0)$ and $f(1) = 3 = g(1)$, then $f = g$ in $\mathcal{F}(S, R)$ by definition.

Similarly, since $f(0) + g(0) = 2 = h(0)$ and $f(1) + g(1) = 6 = h(1)$, then $f + g = h$ in $\mathcal{F}(S, R)$ by definition. ■

Exercise 1.2.8

Proof. Since $x, y \in V$ and $a, b \in F$, by definition $(VS\ 7)$, we know

$$ax + ay + bx + by = a(x + y) + b(x + y).$$

Notice that $x + y$ is unique in V. Then by definition $(VS\ 8)$, we know

$$a(x + y) + b(x + y) = (a + b)(x + y).$$

We together the results, then conclude

$$(a + b)(x + y) = ax + bx + ay + by.$$

■

Exercise 1.2.9

- **Prove Corollary** 1 **of Theorem** 1.1.

Proof. We have known the zero vector always exists by definition $(VS\,3)$. Suppose to contrary that a vector space contains more than one zero vector, so at least two zero vectors in this vector space.

Suppose x, y are arbitrary distinct zero vectors, then by $(VS\,3)$, we know $x = x + y = y$ which means x, y are identical. This contradicts x, y are distinct.

Hence we conclude the zero vector in a vector space is unique. ∎

- **Prove Corollary** 2 **of Theorem** 1.1.

Proof. We have known for every $x \in V$, there always exists an element $y \in V$ such that $x + y = 0$ by definition $(VS\,4)$.

Suppose to contrary that at least two distinct elements y, z such that $x + y = 0$ and $x + z = 0$, then we have $y = z$ which contradicts $y \neq z$ by the supposition.

Hence the element y mentioned from definition $(VS\,4)$ is unique. ∎

- **Prove Theorem 1.2(c).**

Proof. By definition $(VS\,1)$, $(VS\,3)$, $(VS\,7)$, we know for all $a \in F$,

$$a0 + a0 = a(0 + 0) = a0 = a0 + 0 = 0 + a0.$$

Hence $a0 = 0$ for all $a \in F$. ∎

Exercise 1.2.16

Answer. Yes.

Proof. Since $\mathbb{Q} \subset \mathbb{R}$, and we already known V is a vector space over $\mathbb{R}$ by hypothesis; hence V is a vector space over $\mathbb{Q}$. ∎

Exercise 1.2.21

Proof. It suffices to prove Z is compatible with the definition from $(VS\,1)$ to $(VS\,8)$. We denoted the zero vector in V is 0_V and W is 0_W. Notice that we have known V, W are compatible with the definitions since they are vector spaces by hypothesis.

Suppose $x := (x_1, x_2)$ and $y := (y_1, y_2)$ and $z := (z_1, z_2)$ and the zero vector in Z is $0_Z := (0_V, 0_W)$. Here we go by verifying patiently.

- (**VS 1**) For all $x, y \in Z$, we have

$$x + y = (x_1 + y_1, x_2 + y_2) = (y_1 + x_1, y_2 + x_2) = y + x.$$

- (**VS 2**) For all $x, y, z \in Z$, we have

$$(x + y) + z = ((x_1 + y_1) + z_1, (x_2 + y_2) + z_2) = (x_1 + (y_1 + z_1), x_2 + (y_2 + z_2)) = x + (y + z).$$

- (**VS 3**) For all $x \in Z$, we have
$$x + 0_Z = (x_1 + 0_V, x_2 + 0_W) = (x_1, x_2) = x.$$
- (**VS 4**) For all $x \in Z$, we pick $y' = (-x_1, -x_2)$ since $-x_1 \in V$ and $-x_2 \in W$. So $y' \in Z$. Then we have
$$x + y' = (x_1 - x_1, x_2 - x_2) = 0.$$
- (**VS 5**) For all $x \in Z$, we have
$$1x = (1x_1, 1x_2) = (x_1, x_2) = x.$$
- (**VS 6**) For all $a, b \in F$ and $x \in Z$, we have
$$\begin{aligned}(ab)x &= ((ab)x_1, (ab)x_2) \\ &= (a(bx_1), a(bx_2)) \\ &= a(bx_1, bx_2) \\ &= a(b(x_1, x_2)) \\ &= a(bx).\end{aligned}$$
- (**VS 7**) For all $a \in F$ and $x, y \in Z$, we have
$$\begin{aligned}a(x+y) &= (a(x_1+y_1), a(x_2+y_2)) \\ &= (ax_1 + ay_1, ax_2 + ay_2) \\ &= (ax_1, ax_2) + (ay_1, ay_2) \\ &= a(x_1, x_2) + a(y_1, y_2) \\ &= ax + ay.\end{aligned}$$
- (**VS 8**) For all $a, b \in F$ and $x \in Z$, we have
$$\begin{aligned}(a+b)x &= ((a+b)x_1, (a+b)x_2) \\ &= (ax_1 + bx_1, ax_2 + bx_2) \\ &= (ax_1, ax_2) + (bx_1 + bx_2) \\ &= a(x_1, x_2) + b(x_1, x_2) \\ &= ax + bx.\end{aligned}$$

We conclude Z is a vector space over F as promised. ■

1.3 Subspaces

Exercise 1.3.1

(a) **Answer.** False.

Solution. For example, let $V := \mathbb{R}$ and $W := \mathbb{Q}$, then W is a subset of V and V is a vector space over $\mathbb{R}$; however, W is not a vector space over $\mathbb{R}$. ■

(b) **Answer.** False.

Solution. 0 is not an element of $\varnothing$ and hence $\varnothing$ is not a subspace of every vector space. ■

(c) **Answer.** True.

Solution. We pick W such as a zero vector space. ■

(d) **Answer.** False.

Solution. Let $V = \mathbb{R}$, then we pick $W_0 = \{0\}$ and $W_1 = \{1\}$. We know W_0, W_1 are subsets of V, however $W_0 \cap W_1 = \varnothing$ is not a subspace of V form part (b). ■

(e) **Answer.** True.

Solution. We know only diagonal entries could be 0 by definition of a diagonal matrix. Since this is $n \times n$ diagonal matrix, it contains at most n zero entries. ■

(f) **Answer.** False.

Solution. The trace of a square matrix is the **sum** of its diagonal entries. ■

(g) **Answer.** False.

Solution. Since $(0,0,0) \in W$, however $(0,0,0) \notin \mathbb{R}^2$. ■

Exercise 1.3.8

(a) **Answer.** Yes.

Solution. Let $x := (3t, t, -t) \in W_1$ for $t \in \mathbb{R}$, and $y := (3k, k, -k) \in W_1$ for $k \in \mathbb{R}$, and c is a arbitrary scalar. Then

$$x + y = (3t + 3k, t + k, -(t + k)) = (3(t + k), t + k, -(t + k)) \in W_1.$$

Also,

$$cx = c(3t, t, -t) = (3(ct), ct, -ct) \in W_1.$$

Moreover,

$$(0, 0, 0) \in W_1$$

Hence W_1 is a subspace of $\mathbb{R}^3$. ■

(b) **Answer.** No.

Solution. Let $x := (x_3 + 2, x_2, x_3) \in W_2$ and $y := (y_3 + 2, y_2, y_3) \in W_2$. Then

$$x + y = (x_3 + y_3 + 4, x_2 + y_2, x_3 + y_2) \notin W_2.$$

Hence W_2 is not a subspace of $\mathbb{R}^3$. ∎

(c) **Answer.** Yes.

Solution. Let $x := (x_1, x_2, -2x_1 + 7x_2) \in W_3$ and $y := (y_1, y_2, -2y_1 + 7y_2) \in W_3$ and c is a arbitrary scalar. Then

$$x + y = (x_1 + y_1, x_2 + y_2, -2(x_1 + y_1) + 7(x_2 + y_2)) \in W_3.$$

And

$$cx = (cx_1, cx_2, c(-2x_1 + 7x_2)) \in W_3.$$

Moreover,

$$(0, 0, 0) \in W_3.$$

Hence W_3 is a subspace of $\mathbb{R}^3$. ∎

(d) **Answer.** Yes.

Solution. Let $x := (x_1, x_2, -x_1 + 4x_2) \in W_4$ and $y := (y_1, y_2, -y_1 + 4y_2) \in W_4$ and c is a arbitrary scalar. Then

$$x + y = (x_1 + y_1, x_2 + y_2, -(x_1 + y_1) + 4(x_2 + y_2)) \in W_4.$$

And

$$cx = (cx_1, cx_2, c(-x_1 + 4x_2)) \in W_4.$$

Moreover,

$$(0, 0, 0) \in W_4.$$

Hence W_4 is a subspace of $\mathbb{R}^3$. ∎

(e) **Answer.** No.

Solution. Since $(0, 0, 0) \notin W_5$, then W_5 is not a subspace of $\mathbb{R}^3$. ∎

(f) **Answer.** No.

Solution. Let $x := (1, \frac{\sqrt{33}}{3}, 1) \in W_6$ and $y := (1, 1, \frac{\sqrt{3}}{3}) \in W_6$, however $x + y \notin W_6$. Since there exists two elements of W_6 such that the sum of them is not an element of W_6; hence W_6 is not a subspace of $\mathbb{R}^3$. ∎

Exercise 1.3.9

- $W_1 \cap W_3$ is a subspace of $\mathbb{R}^3$.

Solution. Consider

$$W_1 \cap W_3 = \left\{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 = a_2 = a_3 = 0\right\} = \{(0,0,0)\}.$$

Since the zero vector space is a subspace of $\mathbb{R}^3$ trivially, then $W_1 \cap W_3$ is a subspace of $\mathbb{R}^3$. ■

- $W_1 \cap W_4$ is a subspace of $\mathbb{R}^3$.

Solution. Since $W_1 = W_4$, then $W_1 \cap W_4 = W_1$. By Exercise 1.3.8(a), we know W_1 is a subspace of $\mathbb{R}^3$ and so is $W_1 \cap W_4$. ■

- $W_3 \cap W_4$ is a subspace of $\mathbb{R}^3$.

Solution. Consider

$$\begin{aligned} W_3 \cap W_4 &= \left\{(a_1, a_2, a_3) \in \mathbb{R}^3 : 2a_1 - 7a_2 + a_3 = 0 \text{ and } a_1 - 4a_2 - a_3 = 0\right\} \\ &= \left\{(a_1, a_2, a_3) \in \mathbb{R}^3 : 3a_1 = 11a_2 \text{ and } a_2 = -3a_3\right\} \end{aligned}$$

Let $x := (11x_1, 3x_1, -x_1) \in W_3 \cap W_4$ and $y := (11y_1, 3y_1, -y_1) \in W_3 \cap W_4$ and c is a arbitrary scalar. Then

$$x + y = (11(x_1 + y_1), 3(x_1 + y_1), -(x_1 + y_1)) \in W_3 \cap W_4.$$

And

$$cx = (c(11x_1), c(3x_1), c(-x_1)) \in W_3 \cap W_4.$$

Moreover,

$$(0,0,0) \in W_3 \cap W_4.$$

Hence $W_3 \cap W_4$ is a subspace of $\mathbb{R}^3$. ■

Exercise 1.3.19

Proof. ($\Longrightarrow$)

Suppose to contrary that $W_1 \nsubseteq W_2$ and $W_2 \nsubseteq W_1$. Notice that $W_1 \cup W_2$ is a subspace of V. Let $x \in W_1 \setminus W_2$ and $y \in W_2 \setminus W_1$, then $x + y \in W_1$ or $x + y \in W_2$.

If $x + y \in W_1$, then since $y = (x + y) - x \in W_1$, this contradicts that $y \notin W_1$.

Otherwise, if $x + y \in W_2$, then since $x = (x + y) - y \in W_2$, this contradicts that $x \notin W_2$.

In either of two cases, it always leads to a contradiction. Hence $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

($\Longleftarrow$)

Since $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$, then $W_1 \cup W_2 = W_1$ or $W_1 \cup W_2 = W_2$. In either of two cases, $W_1 \cup W_2$ is a subspace of V by hypothesis that W_1 and W_2 are subspaces of V. ■

Exercise 1.3.23

(a) **Proof.** Let $a := x_1 + y_1, b := x_2 + y_2 \in W_1 + W_2$ where $x_1, x_2 \in W_1$ and $y_1, y_2 \in W_2$. Also let c be a arbitrary scalar. Let $0_{W_1}, 0_{W_2}$ be zero vectors in W_1, W_2, respectively. Then

$$a + b = (x_1 + y_1) + (x_2 + y_2) = (x_2 + y_2) + (x_1 + y_1) = b + a \in W_1 + W_2.$$

Also,

$$c(a + b) = c(x_1 + y_1 + x_2 + y_2) = c(x_1 + y_1) + c(x_2 + y_2) = ca + cb \in W_1 + W_2.$$

Moreover,

$$0_{W_1} + 0_{W_2} \in W_1 + W_2.$$

Hence we obtain $W_1 + W_2$ is a subspace of V.

Furthermore, by definition, we know

$$\begin{aligned} W_1 + W_2 &= \{x + y : x \in W_1 \text{ and } y \in W_2\} \\ &\supseteq \{x + 0_{W_2} : x \in W_1 \text{ and } 0_{W_2} \in W_2\} \\ &= W_1. \end{aligned}$$

Also,

$$\begin{aligned} W_1 + W_2 &= \{x + y : x \in W_1 \text{ and } y \in W_2\} \\ &\supseteq \{0_{W_1} + y : 0_{W_1} \in W_1 \text{ and } y \in W_2\} \\ &= W_2. \end{aligned}$$

Hence $W_1 + W_2$ contains both W_1 and W_2. ■

(b) **Proof.** Let W be an arbitrary subspace of V such that $W \supseteq W_1$ and $W \supseteq W_2$. Therefore, for all $x \in W_1$ and $y \in W_2$ imply $x, y \in W$.

Consider

$$\begin{aligned} & W_1 + W_2 = \{x + y : x \in W_1 \text{ and } y \in W_2\} \\ \Longrightarrow\ & W_1 + W_2 \subseteq \{x + y : x \in W \text{ and } y \in W\} \\ \Longrightarrow\ & W_1 + W_2 \subseteq W. \end{aligned}$$

By the arbitrary choice of W, we know any subspace of V that contains W_1 and W_2 must also contain $W_1 + W_2$ as promised. ■

Exercise 1.3.29

Proof. Let $A \in M_{n \times n}(F)$ arbitrarily. And then pick $B \in W_2$ arbitrarily. We know $A - B \in W_1$. Consider

$$\begin{aligned} A &= (A - B) + B \\ &\in \{x + y : x \in W_1 \text{ and } y \in W_2\} \\ &= W_1 + W_2. \end{aligned}$$

By the arbitrary choice of A, we know $M_{n\times n}(F) \subseteq W_1 + W_2$.

On the other hand, $W_1 + W_2 \subseteq M_{n\times n}(F)$ trivially. Hence

$$W_1 + W_2 = M_{n\times n}(F).$$

Notice that

$$\begin{aligned} W_1 \cap W_2 &= \{A \in M_{n\times n}(F) : A_{ij} = 0, i \leq j, A_{ij} = A_{ji}\} \\ &= \{\mathbf{O}_{n\times n}\} \end{aligned}$$

By definition of direct sum, we conclude $M_{n\times n}(F) = W_1 \oplus W_2$. ■

Exercise 1.3.30

Proof. ($\Longrightarrow$) We have known $V = W_1 \oplus W_2$. Let x be an arbitrary vector in V. Suppose to contrary that x cannot be uniquely written. i.e.,

$$x = x_1 + x_2 = x_1' + x_2'$$

where $x_1, x_1' \in W_1$ and $x_2, x_2' \in W_2$.

Notice that $x_1 - x_1' = -(x_2 - x_2')$. Since $x_1 - x_1' \in W_1$ and $-(x_2 - x_2') \in W_2$ from the property of vector spaces in W_1 and W_2, we know $x_1 - x_1' \in W_2$ and $-(x_2 - x_2') \in W_1$. Because $W_1 \cap W_2 = \{0\}$, $x_1 - x_1' = 0$ and $-(x_2 - x_2') = 0$. It follows that

$$x_1 = x_1' \text{ and } x_2 = x_2'.$$

This contradicts the supposition. Hence x can be uniquely written as what we need to show.

By the arbitrary choice of x, we conclude for each vector in V can be uniquely written as $x_1 + x_2$ where $x_1 \in W_1$ and $x_2 \in W_2$.

($\Longleftarrow$) We have known for each vector in V can be uniquely written as $x_1 + x_2$ where $x_1 \in W_1$ and $x_2 \in W_2$. This follows $V = W_1 + W_2$ directly.

Suppose to contrary that $W_1 \cap W_2 \neq \{0\}$. Let $x \in V$ such that $x \neq 0$ and $x \in W_1 \cap W_2$. Then

$$x = x + 0_{W_1} = 0_{W_2} + x$$

where $0_{W_1} \in W_1$ and $0_{W_2} \in W_2$. This contradict the hypothesis; hence $W_1 \cap W_2 = \{0\}$.

We together the results $W_1 + W_2 = V$ and $W_1 \cap W_2 = \{0\}$ to obtain $V = W_1 \oplus W_2$. ■

Exercise 1.3.31

(a) **Proof.** ($\Longrightarrow$) Consider

$$0 = v + (-v) \in v + W.$$

Then $-v \in W$ by definition from $v + W$. Since W is a subspace of V over F by hypothesis, then $v \in W$.

($\Longleftarrow$) Since $v \in W$, then $v + W = W$ is a subspace of V trivially. ■

(b) **Proof.** Consider

$$\begin{aligned} & v_1 + W = v_2 + W \\ \iff & v_1 - v_2 + W = W \\ \iff & v_1 - v_2 \in W. \end{aligned}$$

Hence we complete the proof. ■

(c) **Proof.** Since $v_1 + W = v_1' + W$ and $v_2 + W = v_2' + W$, then from part (b), this implies $v_1 - v_1' \in W$ and $v_2 - v_2' \in W$. Since W is a subspace of V over F, then

$$(v_1 - v_1') + (v_2 - v_2') = (v_1 + v_2) - (v_1' + v_2') \in W.$$

Now from part (b) again, we have

$$(v_1 + v_2) + W = (v_1' + v_2') + W.$$

By definition of addition in cosets of W, we have

$$(v_1 + W) + (v_2 + W) = (v_1' + W) + (v_2' + W).$$

On the other hand, we need to prove $a(v_1 + W) = a(v_1' + W)$.

Since $v_1 + W$ and $v_1' + W$ are subspaces of W, then $v_1, v_1' \in W$ from part (a). Moreover, $av_1, av_1' \in W$ since W is a subspace of V over F and $a \in F$. It follows that $av_1 - av_1' \in W$. From part (b), we know

$$av_1 + W = av_1' + W. \tag{1.1}$$

Consider for all $a \in F$, then by definition of scalar multiplication by scalars of F and equation (1.1), we have

$$a(v_1 + W) = av_1 + W = av_1' + W = a(v_1' + W).$$

Hence we conclude the two operations are well defined. ■

(d) **Proof.** Notice that $S = V/W = \{v + W : v \in V\}$. Now we need to prove S is compatible with eight definitions.

Suppose the zero vector in S is $0 + W$. Here we go by verifying patiently.

- (**VS 1**) For all $v_1 + W, v_2 + W \in S$, we have

$$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W = (v_2 + v_1) + W = (v_2 + W) + (v_1 + W).$$

- (**VS 2**) For all $v_1 + W, v_2 + W, v_3 + W \in S$, we have

$$\begin{aligned} (v_1 + W + v_2 + W) + (v_3 + W) &= (v_1 + v_2) + W) + (v_3 + W) \\ &= (v_1 + v_2 + v_3) + W) \\ &= (v_1 + W) + ((v_2 + v_3) + W) \\ &= (v_1 + W) + (v_2 + W + v_3 + W). \end{aligned}$$

- (**VS 3**) For all $v_1 + W \in S$, there is a zero vector $0 + W$ so that

$$v_1 + W + (0 + W) = (v_1 + 0) + W = v_1 + W.$$

- (**VS 4**) For all $v_1 + W \in S$, we pick $-v_1 + w \in S$ since $v_1 \in V$ and $-v_1 \in V$. So $-v_1 + w \in S$. Then we have

$$(v_1 + W) + (-v_1 + W) = (v_1 - v_1) + W = 0 + W.$$

- (**VS 5**) For all $v_1 + W \in S$, we have

$$1(v_1 + W) = v_1 + W.$$

- (**VS 6**) For all $a, b \in F$ and $v_1 + W \in S$, we have

$$\begin{aligned}(ab)(v_1 + W) &= abv_1 + W \\ &= a(bv_1 + W) \\ &= a(b(v_1 + W)).\end{aligned}$$

- (**VS 7**) For all $a \in F$ and $v_1 + W, v_2 + W \in S$, we have

$$\begin{aligned}a(v_1 + W + v_2 + W) &= a((v_1 + v_2) + W) \\ &= (av_1 + av_2) + W \\ &= (av_1 + W) + (av_2 + W) \\ &= a(v_1 + W) + a(v_2 + W).\end{aligned}$$

- (**VS 8**) For all $a, b \in F$ and $v_1 + W \in S$, we have

$$\begin{aligned}(a + b)(v_1 + W) &= ((a + b)v_1) + W \\ &= (av_1 + bv_1) + W \\ &= (av_1 + W) + (bv_1 + W) \\ &= a(v_1 + W) + b(v_1 + W).\end{aligned}$$

Hence S is a vector space with the operations defined in part (c). ■

1.4 Linear Combinations and Systems of Linear Equations

Exercise 1.4.1

(a) **Answer.** True.

Solution. Pick all scalars which equal zero. ■

(b) **Answer.** False.

Solution. By definition, $span(S) = \{0\}$. ■

(c) **Answer.** True.

Solution. Let W_i be a subspace of V that contains S for any $i \in \mathbb{N}$. By Theorem 1.5, we know $span(S) \subseteq W_i$ for each i. Hence $span(S) = \cap_{i=1}^{\infty} W_i$. ■

(d) **Answer.** False.

Solution. The scalar must be nonzero. ■

(e) **Answer.** True.

Solution. This is trivial from operations. ■

(f) **Answer.** False.

Solution. Consider

$$\begin{cases} 2x + 3y = 1 \\ 4x + 6y = 3 \end{cases}.$$

This system of linear equations have no solution. ■

Exercise 1.4.5

(a) **Answer.** Yes.

Solution. Consider

$$a(1,0,2) + b(-1,1,1) = (2,-1,1).$$

This implies

$$\begin{cases} a - b = 2 \\ b = -1 \\ 2a + b = 1 \end{cases}.$$

Solve to obtain

$$\begin{cases} a = 1 \\ b = -1 \end{cases}.$$

Hence $(2,-1,1) = (1,0,2) - (-1,1,1) \in span(S)$. ■

(b) **Answer.** No.

Solution. Consider

$$a(1,0,2) + b(-1,1,1) = (-1,2,1).$$

This implies

$$\begin{cases} a - b = -1 \\ b = 2 \\ 2a + b = 1 \end{cases}.$$

This systems of linear equations has no solution. Hence $(-1,2,1) \notin span(S)$. ■

(c) **Answer.** No.

Solution. Consider

$$a(1,0,1,-1)+b(0,1,1,1)=(-1,1,1,2).$$

This implies

$$\begin{cases} a=-1 \\ b=1 \\ a+b=1 \\ -a+b=2 \end{cases}.$$

This systems of linear equations has no solution. Hence $(-1,1,1,2)\notin span(S)$. ■

(d) **Answer.** Yes.

Solution. Consider

$$a(1,0,1,-1)+b(0,1,1,1)=(2,-1,1,-3).$$

This implies

$$\begin{cases} a=2 \\ b=-1 \\ a+b=1 \\ -a+b=-3 \end{cases}.$$

Hence $(2,-1,1,-3)=2(1,0,1,-1)-(0,1,1,1)\in span(S)$. ■

(e) **Answer.** Yes.

Solution. Consider

$$a(x^3+x^2+x+1)+b(x^2+x+1)+c(x+1)=-x^3+2x^2+3x+3.$$

This implies

$$\begin{cases} a=-1 \\ a+b=2 \\ a+b+c=3 \\ a+b+c=3 \end{cases}.$$

Solve to obtain $a=-1, b=3, c=1$. Hence

$$-x^3+2x^2+3x+3=-(x^3+x^2+x+1)+3(x^2+x+1)+(x+1)\in span(S).$$

■

(f) **Answer.** No.

Solution. Consider

$$a(x^3+x^2+x+1)+b(x^2+x+1)+c(x+1)=2x^3-x^2+x+3.$$

This implies

$$\begin{cases} a=2 \\ a+b=-1 \\ a+b+c=1 \\ a+b+c=3 \end{cases}.$$

This system of linear equations have no solution. Hence $2x^3-x^2+x+3\notin \operatorname{span}(S)$. ■

(g) **Answer.** Yes.

Solution. Consider

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

This implies

$$\begin{cases} a+c=1 \\ b+c=2 \\ -a=-3 \\ b=4 \end{cases}.$$

Solve to obtain $a=3, b=4, c=-2$. Hence $\begin{pmatrix}1 & 2\\ -3 & 4\end{pmatrix}\in \operatorname{span}(S)$. ■

(h) **Answer.** No.

Solution. Consider

$$a\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix}+b\begin{pmatrix}0 & 1\\ 0 & 1\end{pmatrix}+c\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

This implies

$$\begin{cases} a+c=1 \\ b+c=0 \\ -a=0 \\ b=1 \end{cases}.$$

This system of linear equations have no solution. Hence $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\notin \operatorname{span}(S)$. ■

Exercise 1.4.15

Proof. Let $v \in \text{span}(S_1 \cap S_2)$ arbitrarily, then we pick scalars a_i and vectors $x_i \in S_1 \cap S_2$ for each $i = 1, 2, \cdots, n$. It follows that

$$v = \sum_{i=1}^{n} a_i x_i.$$

So $v \in span(S_1)$ and $v \in span(S_2)$. Then $v \in span(S_1) \cap span(S_2)$.
By the arbitrary choice of v, we conclude $span(S_1 \cap S_2) \subseteq span(S_1) \cap span(S_2)$. ■

- *Example of* $\mathbf{span}(\mathbf{S_1} \cap \mathbf{S_2}) = \mathbf{span}(\mathbf{S_1}) \cap \mathbf{span}(\mathbf{S_2})$.

Answer. $S_1 = S_2 = (1, 1)$

Solution. Since

$$span(S_1 \cap S_2) = \{(x, y) : x = y\},$$

and

$$span(S_1) \cap span(S_2) = \{(x, y) : x = y)\} \cap \{(x, y) : x = y)\} = \{(x, y) : x = y)\}.$$

Hence

$$span(S_1 \cap S_2) = span(S_1) \cap span(S_2).$$

■

- *Example of* $\mathbf{span}(\mathbf{S_1} \cap \mathbf{S_2}) \neq \mathbf{span}(\mathbf{S_1}) \cap \mathbf{span}(\mathbf{S_2})$.

Answer. $S_1 = (0, 1)$ and $S_2 = (1, 0)$

Solution. Since

$$span(S_1 \cap S_2) = span(\emptyset) = \{(0, 0)\},$$

and

$$span(S_1) \cap span(S_2) = \mathbb{R}^2.$$

Hence

$$span(S_1 \cap S_2) \neq span(S_1) \cap span(S_2).$$

■

1.5 Linear Dependence and Linear Independence

Exercise 1.5.1

(a) **Answer.** False.

Solution. Let $S = \{(0, 1), (1, 0), (2, 0)\}$ be linearly dependent, however $(0, 1)$ is not a linear combination of $(1, 0), (2, 0)$. ■

(b) **Answer.** True.

Solution. The nonzero scalar a satisfies $a\vec{0} = \vec{0}$. ■

(c) **Answer.** False.

Solution. Any linearly dependent sets are nonempty. ■

(d) **Answer.** False.

Solution. Let $S = \{(0,1),(1,0),(2,0)\}$ be linearly dependent, however $S' = \{(0,1),(1,0)\} \subset S$ is linearly independent. ■

(e) **Answer.** True.

Solution. It follows from the corollary of Theorem 1.6. ■

(f) **Answer.** True.

Solution. This is by the definition. ■

Exercise 1.5.3

Proof. Consider

$$a\begin{pmatrix}1&1\\0&0\\0&0\end{pmatrix}+b\begin{pmatrix}0&0\\1&1\\0&0\end{pmatrix}+c\begin{pmatrix}0&0\\0&0\\1&1\end{pmatrix}+d\begin{pmatrix}1&0\\1&0\\1&0\end{pmatrix}+e\begin{pmatrix}0&1\\0&1\\0&1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\\0&0\end{pmatrix}. \tag{1.2}$$

This means

$$\begin{cases}a+d=0\\a+e=0\\b+d=0\\b+e=0\\c+d=0\\c+e=0\end{cases}.$$

Hence $a=b=c$ and $d=e$ and $a+d=0$. Now we can pick $a=b=c=1$ and $d=e=-1$ so that the equation (1.2) holds. Since a,b,c,d,e are not all zeroes, by definition, we conclude the set

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

is linearly dependent. ■

Exercise 1.5.9

Proof. ($\Longrightarrow$) Consider
$$au + bv = 0.$$

If either u or v is a zero vector, W.L.O.G, we assume $u = 0$. Then $u = 0v$ holds. A similar argument proves if $v = 0$, then $v = 0u$ also holds.

Otherwise, if u and v are nonzero vectors, we know a and b are both nonzero since u, v is linearly dependent. Then
$$u = -\frac{b}{a}v \text{ and } v = -\frac{a}{b}u.$$
All conditions are discussed. It follows that u or v is a multiple of the other.

($\Longleftarrow$) W.L.O.G., if $u = kv$ where k is a scalar, then $u - kv = 0$. Since 1 and $-k$ are not both zero, we know $\{u, v\}$ is linearly dependent.

A similar argument establishes $v = ku$ where k is a scalar, then $\{u, v\}$ is also linearly dependent. ■

Exercise 1.5.11

Answer. 2^n.

Proof. Suppose $u_i \in \{0, 1\}$ for each $i = 1, 2, \cdots, n$. Since a set is linearly independent if and only if the only representations of 0 as linear combinations of its vectors are trivial representations, then for any $v \in span(S)$,
$$v = \sum_{i=1}^{n} a_i u_i = 0$$
is a unique representation. There are 2^n representations and hence $|span(S)| = 2^n$. ■

Exercise 1.5.13

Proof. ($\Longrightarrow$) Since $\{u, v\}$ is linearly independent, we consider
$$au + bv = 0.$$
where $a = b = 0$. Then we let
$$a' := \frac{a+b}{2} \text{ and } b' := \frac{a-b}{2}.$$
Notice that $a' = b' = 0$. Consider
$$\begin{aligned} a'(u+v) + b'(u-v) &= \frac{a+b}{2}(u+v) + \frac{a-b}{2}(u-v) \\ &= au + bv \\ &= 0. \end{aligned}$$

This means $\{u + v, u - v\}$ is linearly independent.

($\Longleftarrow$) Since $\{u+v, u-v\}$ is linearly independent, we consider

$$a(u+v)+b(u-v)=0.$$

where $a=b=0$. Then we let

$$a' := a+b \text{ and } b' := a-b.$$

Notice that $a'=b'=0$. Consider

$$\begin{aligned} a'u+b'v &= (a+b)u+(a-b)v \\ &= a(u+v)+b(u-v) \\ &= 0. \end{aligned}$$

This means $\{u, v\}$ is linearly independent. ■

1.6 Bases and Dimension

Exercise 1.6.1

(a) **Answer.** False.

Solution. Since $span(\emptyset)=\{0\}$, then $\emptyset$ is a basis for the zero vector space. ■

(b) **Answer.** True.

Solution. It follows from Theorem 1.9. Moreover, bases must be finite. ■

(c) **Answer.** False.

Solution. The counter-example is $P(F)$, which has a infinite basis. ■

(d) **Answer.** False.

Solution. Here we give a counter-example. $\mathbb{R}^2$ has a basis $\{(0,1),(1,0)\}$; also, it has a basis $\{(1,1),(0,1)\}$. ■

(e) **Answer.** True.

Solution. It follows from the Corollary of Theorem 1.10. ■

(f) **Answer.** False.

Solution. $\dim(P_n(F)) = \mathbf{n+1}$. ■

(g) **Answer.** False.

Solution. $\dim\left(M_{n\times n(F)}\right) = \mathbf{mn}$. ■

(h) **Answer.** True.

Solution. It follows from the Replacement Theorem directly. ■

(i) **Answer.** False.

Solution. Here we give a counter-example. Let $S := \{(0,1),(0,2),(1,1)\}$, then $span(S) = \mathbb{R}^2$. However, $(0,5) = 3(0,1) + (0,2) = (0,1) + 2(0,2) \in \mathbb{R}^2$. ■

(j) **Answer.** True.

Solution. Let W is a subspace of V where $\dim(V) < \infty$, then by Theorem 1.11, we know $\dim(W) \leq \dim(V)$; hence $\dim(W) < \infty$. ■

(k) **Answer.** True.

Solution. We have known $\dim(V) = n$ by hypothesis. Then $\{0\}$ is the only subspace of V such that $\dim(\{0\}) = 0$, and V is the only subspace of V such that $\dim(V) = n$. ■

(l) **Answer.** True.

Solution. It follows from the Corollary 2 of Theorem 1.10. ■

Exercise 1.6.4

Answer. No.

Proof. By the Corollary 2(a) of Theorem 1.10, we know any finite generating set for a vector space with dimension n contains at least n vectors.

Notice that the set

$$\left\{x^3 - 2x^2 + 1, 4x^2 - x + 3, 3x - 2\right\}$$

has only 3 vectors which is less than $\dim(P_3(\mathbb{R})) = 4$, then the set cannot generate $P_3(\mathbb{R})$. ■

Exercise 1.6.9

Answer. $(a_1, a_2, a_3, a_4) = a_1u_1 + (a_2 - a_1)u_2 + (a_3 - a_2)u_3 + (a_4 - a_3)u_4$.

Solution. Let $(a_1, a_2, a_3, a_4) := c_1u_1 + c_2u_2 + c_3u_3 + c_4u_4$ where $c_1, c_2, c_3, c_4 \in F$. By hypothesis, we can compute to obtain

$$c_1u_1 + c_2u_2 + c_3u_3 + c_4u_4 = (c_1, c_1 + c_2, c_1 + c_2 + c_3, c_1 + c_2 + c_3 + c_4).$$

Then

$$\begin{cases} a_1 = c_1 \\ a_2 = c_1 + c_2 \\ a_3 = c_1 + c_2 + c_3 \\ a_4 = c_1 + c_2 + c_3 + c_4 \end{cases}.$$

Solve it to know $c_1 = a_1, c_2 = a_2 - a_1, c_3 = a_3 - a_2, c_4 = a_4 - a_3$. Hence we conclude

$$(a_1, a_2, a_3, a_4) = a_1u_1 + (a_2 - a_1)u_2 + (a_3 - a_2)u_3 + (a_4 - a_3)u_4.$$

■

Exercise 1.6.13

Answer. $\{(1, 1, 1)\}$.

Solution. First, solve the systems of linear equations

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_1 - 3x_2 + x_3 = 0 \end{cases}.$$

Then we know $x_1 = x_2 = x_3$. Let $x_1 := t$ where $t \in \mathbb{R}$, so $x_2 = x_3 = t$. That is, the solution set is $\{(t, t, t) : t \in \mathbb{R}\}$ which is a subspace of $\mathbb{R}^3$. We just pick a basis $\{(1, 1, 1)\}$ trivially. ■

Exercise 1.6.17

Answer. $\{E_{ij} - E_{ji} : i < j \text{ and } 1 \le i, j \le n\}$ is a basis for W, and $\dim(W) = \frac{n(n-1)}{2}$.

Solution. By the definition of skew-symmetric matrix, we suppose a skew-symmetric A is

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{pmatrix}.$$

Then we pick a basis for A such that

$$\{E_{ij} - E_{ji} : i < j \text{ and } 1 \le i, j \le n\}.$$

Notice that the set of vectors is linearly independent.

Since the dimension of W is the amount of vectors in a basis, then

$$\dim(W) = \sum_{j=1}^{n}\sum_{i=1}^{j-1} 1 = \sum_{j=1}^{n}(j-1) = \frac{n(n+1)}{2} - n = \frac{n(n-1)}{2}.$$

■

Exercise 1.6.20

(a) **Proof.** By the Maximal Principle, we can pick a set B which is a subset of S such that B is linearly independent and is the only linearly independent subset of S that contains B is itself. It suffices to say that B is a maximal linearly independent subset of S.

In order to prove B is a basis for V, it suffices to prove B generates V. We claim $S \subseteq span(B)$.

Suppose to contrary that $S \nsubseteq span(B)$. There exists a $v \in S$ such that $v \notin span(B)$. Since Theorem 1.7 implies that $B \cup \{v\}$ is linearly independent. This contradicts that B is maximal.

Since $span(S) = V$ by hypothesis, it follows from Theorem 1.5 and $B \subseteq S \subseteq span(B)$ that $span(B) = V$. ■

(b) **Proof.** Let G be any generating set for V, then from part (a), we know some subset of H of G is a basis for V. By the Corollary 1 of Theorem 1.10, we know H contains exactly n vectors. Since a subset of G contains n vectors, we conclude G must contain at least n vectors. ■

Exercise 1.6.26

Answer. The dimension is n.

Solution. Let $W := \{f \in P_n(\mathbb{R}) : f(a) = 0\}$. We know $f_k(x) = (x-a)(a_{k-1}x^{k-1} + \cdots + a_1x + a_0)$ for $k = 1, 2, \cdots, n$; also, $f_k(x) \in W$ for each k. Moreover, $\{f_1(x), f_2(x), \cdots, f_n(x)\}$ is linearly independent. Hence $\dim(W) = n$. ■

Exercise 1.6.29

(a) **Proof.** Follow the hint. First, we claim $\beta = \{u_1, \cdots, u_k, v_1, \cdots, v_m, w_1, \cdots, w_p\}$ is a basis for $W_1 + W_2$.

- **β generates $\mathbf{W_1 + W_2}$.**

Let

$$t_1 := \sum_{i=1}^{k} a_i u_i + \sum_{i=1}^{m} b_i v_i \in W_1;$$
$$t_2 := \sum_{i=1}^{k} a_i' u_i + \sum_{i=1}^{p} c_i w_i \in W_2.$$

Then pick $t = t_1 + t_2 \in W_1 + W_2$, we have

$$t = t_1 + t_2 = \sum_{i=1}^{k}(a_i + a_i')u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{p} c_i w_i \in span(\beta).$$

Since t_1 and t_2 are arbitrary, this implies $W_1 + W_2 = span(\beta)$.

- **β is linearly independent.**

Consider

$$\sum_{i=1}^{k} a_i u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{p} c_i w_i = 0. \tag{1.3}$$

Let

$$v := \sum_{i=1}^{k} a_i u_i \sum_{i=1}^{m} b_i v_i = -\sum_{i=1}^{p} c_i w_i. \tag{1.4}$$

We know $v \in W_1$ and $v \in W_2$; hence $v \in W_1 \cap W_2$. It follows that

$$v = \sum_{i=1}^{k} a_i' u_i.$$

Since $\{u_1, \cdots, u_k, w_1, \cdots, w_p\}$ is a basis for W_2 and $v \in W_2$, then

$$\sum_{i=1}^{k} a_i' u_i + \sum_{i=1}^{p} c_i w_i = 0 \implies a_1' = \cdots = a_k' = c_1 = \cdots = c_p = 0.$$

Also since $\{u_1, \cdots, u_k, v_1, \cdots, v_m\}$ is a basis for W_1 and $v \in W_1$, from equation (1.4), we know

$$v = \sum_{i=1}^{k} a_i u_i \sum_{i=1}^{m} b_i v_i = 0 \implies a_1 = \cdots = a_k = b_1 = \cdots = b_m = 0.$$

Hence from equation (1.3), we obtain β is linearly independent.

We prove the claim holds as promised. Finally, we compute

$$\begin{aligned} \dim(W_1 + W_2) &= k + m + p \\ &= (k+m) + (k+p) - k \\ &= \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2). \end{aligned} \tag{1.5}$$

By hypothesis that W_1, W_2 are finite-dimensional vector spaces, we know $W_1 \cap W_2$ is also finite-dimensional; hence we obtain $k+m, k+p, k < \infty$. From formula (1.5), we conclude $W_1 + W_2$ is also a finite-dimensional vector space since $k + m + p < \infty$. ■

(b) **Proof.** From part (a) and by hypothesis, we know

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

Hence

$$\begin{aligned} & V = W_1 \oplus W_2 \\ \iff \quad & V = W_1 + W_2 \text{ and } W_1 \cap W_2 = \{0\} \\ \iff \quad & \dim(V) = \dim(W_1) + \dim(W_2) - \dim(\{0\}) \\ \iff \quad & \dim(V) = \dim(W_1) + \dim(W_2) \end{aligned}$$

■

Exercise 1.6.34

(a) **Proof.** Since V is a finite-dimensional vector space and W_1 is a subspace of V, by Theorem 1.11, W_1 is also a finite-dimensional vector space and $\dim(W_1) < \dim(V)$.

Suppose $\dim(V) = n < \infty$. Let $\dim(W_1) = k \leq n$ and $\beta_1 = \{u_1, \cdots, u_k\}$ is a basis for W_1. By the Corollary 2(c) of Theorem 1.10, we can extend β_1 to β so that $\beta = \{u_1, \cdots, u_k, u_{k+1}, \cdots, u_n\}$ is a basis for V. We pick $\beta_2 = \{u_{k+1}, \cdots, u_n\}$ such that β_2 is a basis for W_2.

Notice that $\beta_1 \cup \beta_2 = \beta$. Since β_1, β_2, β are bases for W_1, W_2, V, respectively; by Exercise 1.6.33(b), we conclude $V = W_1 \oplus W_2$. ■

(b) **Answer.**

$$W_2 = \{(0, a_2) : a_2 \in \mathbb{R}\} \text{ and } W_2' = \{(a_2, a_2) : a_2 \in \mathbb{R}\}.$$

($\star$ Notice that answers are not unique)

Solution. $\beta_1 = \{(1, 0)\}$ is a basis for W_1, and $\beta_2 = \{(0, 1)\}$ is a basis for W_2. Since $\beta_1 \cap \beta_2 = \varnothing$ and $\beta_1 \cup \beta_2$ is a basis for $\mathbb{R}^2$, by Exercise 1.6.33(b), we know $V = W_1 \oplus W_2$.

A similar argument shows that $\beta_2' = \{(1, 1)\}$ is a basis for W_2' and hence $V = W_1 \oplus W_2'$. ■

1.7 Maximal Linearly Independent Subsets

Exercise 1.7.1

(a) **Answer.** False.

Solution. For a counter-example, the family $\{1, 2, \cdots, n, \cdots\}$ ordered by $\leq$ has no maximal element. ■

(b) **Answer.** False.

Solution. For a counter-example, the chain $\{1, 2, \cdots, n, \cdots\}$ ordered by $\leq$ has no maximal element. ■

(c) **Answer.** False.

Solution. The family

$$S = \{\{a\}, \{b\}\}$$

is ordered by containment. S has two maximal elements $\{a\}$ and $\{b\}$. ■

(d) **Answer.** True.

Solution. Suppose to contrary that there exists at least two distinct maximal elements A, B. Since they are in a chain, so by definition, $A \subseteq B$ or $B \subseteq A$ must holds. Moreover, they are maximal, so $A \subseteq B$ and $B \subseteq A$. It follows that $A = B$ contradicts that A, B are distinct. ■

(e) **Answer.** True.

Solution. Let β be a basis for a vector space V. By definition of bases, β is linearly independent. Also $\text{span}(\beta) = V$, by Theorem 1.7, if $v \in V$ and $v \notin \beta$, then $\beta \cup \{v\}$ is linearly dependent. That is, β is a maximal linearly independent subset by definition. ■

(f) **Answer.** True.

Solution. It follows from Theorem 1.12. ■

Chapter 2

Linear Transformations and Matrices

2.1 Linear Transformations, Null Spaces, and Ranges

Exercise 2.1.1

(a) **Answer.** True.

Solution. This is by definition. ∎

(b) **Answer.** False.

Solution. Let $T : \mathbb{C} \to \mathbb{C}$ such that $T(a+bi) = a$, and $x := a_1 + b_1 i$ and $y := a_2 + b_2 i$. Notice that $T(x+y) = T(a_1 + b_1 i + a_2 + b_2 i) = a_1 + a_2 = T(x) + T(y)$. However, $T(ci) = 0$ and $iT(c) = ci$ are distinct. So T is not linear. ∎

(c) **Answer.** False.

Solution. We will give a counter-example for each direction.
($\Longrightarrow$) Let $T : \mathbb{R} \to \mathbb{R}$ such that $T(x) = x + 1$. We know T is one-to-one, however $T(x) = 0$ implies $x = -1$.

($\Longleftarrow$) Let $T(x) = |x|$. Since $T(x) = 1$ implies $x = 1$ or $x = -1$, then T is not one-to-one. ∎

(d) **Answer.** True.

Solution. Since

$$T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V),$$

then we conclude $T(0_V) = 0_W$. ∎

(e) **Answer.** False.

Solution. It follows from Theorem 2.3. Notice that it equals to $\dim(V)$, not $\dim(W)$. ■

(f) **Answer.** False.

Solution. Consider $T : \mathbb{R} \to \mathbb{R}$ such that $T(x) = 0$. ■

(g) **Answer.** True.

Solution. It follows from the Corollary of Theorem 2.6. ■

(h) **Answer.** False.

Solution. It is precise to say this statement might hold or not hold. In other words, it doesn't hold always.

For example, we set $x_2 = 2x_1$, then $T(x_2) = T(2x_1) = 2T(x_1) = 2y_1$. Hence if $y_2 \neq 2y_1$, there doesn't exist such a linear transformation. ■

Exercise 2.1.6

Proof. Consider $T(A + cB)$ where $A, B \in M_{n\times n}(F)$ and $c \in F$. Then

$$T(A + cB) = \sum_{i=1}^{n} (A_{ii} + cB_{ii}) = \sum_{i=1}^{n} A_{ii} + c\sum_{i=1}^{n} B_{ii} = T(A) + cT(B).$$

Hence T is linear.

Now we find $N(T)$ and $R(T)$ with their bases. $N(T) = \{A : \mathrm{tr}(A) = 0\}$ with a basis $\{E_{ij}\}_{i\neq j} \cup \{E_{11} - E_{ii}\}_{i=2,3,\cdots,n}$; $R(T) = F$ with a basis $\{1\}$.

It follows that $\mathrm{nullity}(T) = n^2 - 1$ and $\mathrm{rank}(T) = 1$. We can verify

$$\dim(M_{n\times n}(F)) = n^2 = (n^2 - 1) + 1 = \mathrm{nullity}(T) + \mathrm{rank}(T).$$

Since $\mathrm{nullity}(T) \neq 0$, then T is not one-to-one. But $\mathrm{rank}(T) = \dim(F)$, then T is onto. ■

Exercise 2.1.11

Answer. $T(8, 11) = (5, -3, 16)$.

Proof. Since $\{(1, 1), (2, 3)\}$ is a basis for $\mathbb{R}^2$, by Theorem 2.6, we know the linear transformation T exists. Moreover, $T(8, 11) = 2T(1, 1) + 3T(2, 3) = 2(1, 0, 2) + 3(1, -1, 4) = (5, -3, 16)$. ■

Exercise 2.1.12

Answer. No.

Proof. Since $T(-2,0,-6) = (2,1)$ and $2T(1,0,3) = 2(1,1) = (2,2)$, however $T(-2,0,-6) \neq 2T(1,0,3)$. So there doesn't exist the linear transformation T. ∎

Exercise 2.1.20

Proof. Let $w_1, w_2 \in T(V_1)$ and $c \in F$, then there exists $v_1, v_2 \in V_1$ such that $w_1 = T(v_1)$ and $w_2 = T(v_2)$. Consider

$$cw_1 + w_2 = cT(v_1) + T(v_2) = T(cv_1 + v_2) \in T(V_1).$$

Hence $T(V_1)$ is a subspace of W.

Moreover, let $B = \{x \in V : T(x) \in W_1\}$ and $x_1, x_2 \in B$ such that $T(x_1) = w_1' \in W_1$ and $T(x_2) = w_2' \in W_1$. Consider

$$T(cx_1 + x_2) = cT(x_1) + T(x_2) = cw_1' + w_2' \in W_1.$$

Hence $cx_1 + x_2 \in B$. This means B is a subspace of V. ∎

Exercise 2.1.22

- $T : \mathbb{R}^3 \to \mathbb{R}$ is linear

Proof. Let $\beta := \{e_1, e_2, e_3\}$ be a basis for $\mathbb{R}^3$. Set $T(e_1) = a$, $T(e_2) = b$, $T(e_3) = c$. For any $(x, y, z) \in \mathbb{R}^3$, we have

$$T(x, y, z) = T(xe_1 + ye_2 + ze_3) = xT(e_1) + yT(e_2) + zT(e_3) = ax + by + cz.$$

∎

- $T : F^n \to F$ is linear

Proof. Let $\beta := \{e_1, \cdots, e_n\}$ be a basis for F^n. Set $T(e_i) = c_i \in F$ for each $i = 1, \cdots, n$. For any $(x_1, \cdots, x_n) \in F^n$, we have

$$T(x_1, \cdots, x_n) = T(x_1e_1 + \cdots, x_ne_n) = x_1T(e_1) + \cdots + x_nT(e_n) = c_1x_1 + \cdots c_nx_n.$$

∎

- $T : F^n \to F^m$ is linear

Proof. Let $\beta := \{e_1, \cdots, e_n\}$ be a basis for F^n. Set $T(e_i) = \sum_{j=1}^m c_{ij}w_j \in F$ for each $i = 1, \cdots, n$. For any $(x_1, \cdots, x_n) \in F^n$, we have

$$\begin{aligned} T(x_1, \cdots, x_n) &= T\left(\sum_{i=1}^n x_i e_i\right) \\ &= \sum_{i=1}^n x_i T(e_i) \\ &= \sum_{i=1}^n x_i \sum_{j=1}^m c_{ij} w_j \\ &= \sum_{i=1}^n \sum_{j=1}^m c_{ij} x_i w_j \\ &= \sum_{j=1}^m \left(\sum_{i=1}^n c_{ij} x_i\right) w_j \\ &= (\sum_{i=1}^n c_{i1} x_1, \cdots, \sum_{i=1}^n c_{im} x_m). \end{aligned}$$

■

Exercise 2.1.24

(a) **Answer.** $T(a, b) = (0, b)$.

Solution. Set $W_1 = \{(0, b) : b \in \mathbb{R}\}$ and $W_2 = \{(a, 0) : a \in \mathbb{R}\}$. We know $W_1 \cup W_2 = \mathbb{R}^2$ and $W_1 \cap W_2 = \{(0, 0)\}$; hence $\mathbb{R}^2 = W_1 \oplus W_2$. Since $(a, b) = (0, b) + (a, 0)$ with $(0, b) \in W_1$ and $(a, 0) \in W_2$. So $T(a, b) = (0, b)$. ■

(b) **Answer.** $T(a, b) = (0, b - a)$.

Solution. Set $W_1 = \{(0, b - a) : a, b \in \mathbb{R}\}$ and $W_2 = \{(a, a) : a \in \mathbb{R}\}$. We know $W_1 \cup W_2 = \mathbb{R}^2$ and $W_1 \cap W_2 = \{(0, 0)\}$; hence $\mathbb{R}^2 = W_1 \oplus W_2$. Since $(a, b) = (0, b - a) + (a, a)$ with $(0, b - a) \in W_1$ and $(a, a) \in W_2$. So $T(a, b) = (0, b - a)$. ■

Exercise 2.1.35

(a) **Proof.** Notice that V is finite-dimensional vector space. then both nullity(T) and rank(T) are finite. By hypothesis $V = R(T) + N(T)$, it suffices to prove $R(T) \cap N(T) = \{0\}$.

By Exercise 1.6.29(a), we know

$$\dim(R(T) + N(T)) = \dim(R(T)) + \dim(N(T)) - \dim(R(T) \cap N(T)).$$

By hypothesis, we have

$$\dim(R(T) \cap N(T)) = \dim(R(T)) + \dim(N(T)) - \dim(V).$$

Since $\dim(V)$ is finite, by the Dimension Theorem, we have

$$\dim(V) = \operatorname{rank}(T) + \operatorname{nullity}(T) = \dim(R(T)) + \dim(N(T)).$$

Finally we obtain

$$\dim(R(T) \cap N(T)) = 0$$

which implies

$$R(T) \cap N(T) = \{0\}.$$

Hence

$$V = R(T) \oplus N(T).$$

■

(b) **Proof.** This is similar to part (a). Notice that $\dim(V)$ is finite, then both $\operatorname{nullity}(T)$ and $\operatorname{rank}(T)$ are finite. Consider

$$\dim(R(T) + N(T)) = \dim(R(T)) + \dim(N(T)) - \dim(R(T) \cap N(T)).$$

By hypothesis, we know

$$\dim(R(T) + N(T)) = \dim(R(T)) + \dim(N(T)) - 0.$$

By the Dimension Theorem, we have

$$\dim(R(T) + N(T)) = \dim(R(T)) + \dim(N(T)) = \dim(V).$$

Hence $R(T) + N(T) = V$. Then following $R(T) \cap N(T) = \{0\}$ by hypothesis, we conclude

$$V = R(T) \oplus N(T).$$

■

2.2 The Matrix Representation of a Linear Transformation

Exercise 2.2.1

(a) **Answer.** True.

Solution. It follows from Theorem 2.7(a). ■

(b) **Answer.** True.

Solution. It follows from the Corollary of Theorem 2.6. ■

(c) **Answer.** False.

Solution. It should be $n \times m$ matrix. ■

(d) **Answer.** True.

Solution. It follows from Theorem 2.8(a). ■

(e) **Answer.** True.

Solution. It follows from definition immediately. ■

(f) **Answer.** False.

Solution. Suppose $M_{m\times n}(F)$ is a subspace of $\mathcal{L}(V, W)$, then $M_{m\times n}(F)$ is not a subspace of $\mathcal{L}(W, V)$ unless $m = n$. ■

Exercise 2.2.2

(a) **Answer.** $[T]_\beta^\gamma = \begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 1 & 0 \end{pmatrix}$.

Solution. Since $T(1,0) = (2,3,1)$ and $T(0,1) = (-1,4,0)$. ■

(b) **Answer.** $[T]_\beta^\gamma = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 0 & 1 \end{pmatrix}$.

Solution. Since $T(1,0,0) = (2,1)$ and $T(0,1,0) = (3,0)$ and $T(0,0,1) = (-1,1)$. ■

(c) **Answer.** $[T]_\beta^\gamma = \begin{pmatrix} 2 & 1 & -3 \end{pmatrix}$.

Solution. Since $T(1,0,0) = 2$ and $T(0,1,0) = 1$ and $T(0,0,1) = -3$. ■

(d) **Answer.** $[T]_\beta^\gamma = \begin{pmatrix} 0 & 2 & 1 \\ -1 & 4 & 5 \\ 1 & 0 & 1 \end{pmatrix}$.

Solution. Since $T(1,0,0) = (0,-1,1)$ and $T(0,1,0) = (2,4,0)$ and $T(0,0,1) = (1,5,1)$. ■

(e) **Answer.** $[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}$.

Solution. Since $T(e_i) = e_i$ for each $i = 1, 2, \cdots, n$. ■

(f) **Answer.** $[T]_\beta^\gamma = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{pmatrix}.$

Solution. Since $T(e_i) = e_{n+1-i}$ for each $i = 1, 2, \cdots, n$. ■

(g) **Answer.** $[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \end{pmatrix}_{1\times n}$.

Solution. Since $[T]_\beta^\gamma \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 + a_n$. ■

Exercise 2.2.6

Proof. It suffices to prove $\mathcal{L}(V, W)$ is compatible with the definition from $(VS\ 1)$ to $(VS\ 8)$. We denoted the zero vector in W is 0_W. Notice that we have known V, W are compatible with the definitions since they are vector spaces by hypothesis.

Suppose an arbitrary vector v in V. T_0 is a zero transformation that $T_0(v) = 0_W$. Here we go by verifying patiently.

- (**VS 1**) For all $S, T \in \mathcal{L}(V, W)$, we have

$$(S+T)(v) = S(v) + T(v) = T(v) + S(v) = (T+S)(v).$$

- (**VS 2**) For all $S, T, U \in \mathcal{L}(V, W)$, we have

$$\begin{aligned}((S+T)+U)(v) &= (S+T)(v) + U(v) = S(v) + T(v) + U(v) \\ &= S(v) + (T+U)(v) = (S+(T+U))(v).\end{aligned}$$

- (**VS 3**) For all $T \in \mathcal{L}(V, W)$, we have

$$(T+T_0)(v) = T(v) + T_0(v) = T(v) + 0_W = T(v).$$

- (**VS 4**) For all $T \in \mathcal{L}(V, W)$, we pick $-T$ such that $(-T)(v) = -T(v)$. Then we have

$$(T+(-T))(v) = T(v) + (-T)(v) = T(v) - T(v) = 0_W.$$

- (**VS 5**) For all $T \in \mathcal{L}(V, W)$, we have

$$(1 \cdot T)(v) = 1 \cdot T(v) = T(v).$$

- (**VS 6**) For all $a, b \in F$ and $T \in \mathcal{L}(V, W)$, we have

$$\begin{aligned}((ab)T)(v) &= (ab)T(v)\\ &= a(bT(v))\\ &= a(bT)(v)\end{aligned}$$

- (**VS 7**) For all $a \in F$ and $S, T \in \mathcal{L}(V, W)$, we have

$$\begin{aligned}(a(S+T))(v) &= a((S+T))(v))\\ &= a(S(v)+T(v))\\ &= aS(v)+aT(v)\\ &= (aS+aT)(v).\end{aligned}$$

- (**VS 8**) For all $a, b \in F$ and $T \in \mathcal{L}(V, W)$, we have

$$\begin{aligned}((a+b)T)(v) &= (a+b)T(v)\\ &= (aT)(v)+(bT)(v)\\ &= (aT+bT)(v).\end{aligned}$$

We conclude $\mathcal{L}(V, W)$ is a vector space over F as promised. ■

Exercise 2.2.8

Proof. Let $\beta := (v_1, \cdots, v_n)$ be a basis for V. Then $x = \sum_{i=1}^n a_i v_i$ and $y = \sum_{i=1}^n b_i v_i$ where $a_i, b_i \in F$. For $c \in F$, we consider

$$\begin{aligned}T(cx+y) &= T\left(c + \sum_{i=1}^n a_i v_i + \sum_{i=1}^n b_i v_i\right)\\ &= T\left(\sum_{i=1}^n (ca_i + b_i) v_i\right)\\ &= (ca_1 + b_1, \cdots, ca_n + b_n)\\ &= c(a_1, \cdots, a_n) + (b_1, \cdots, b_n)\\ &= cT(x) + T(y).\end{aligned}$$

Hence T is linear. ■

Exercise 2.2.12

Answer.

$$[T]_\beta = \begin{pmatrix} I_k & O \\ O & O \end{pmatrix}.$$

Proof. Suppose V has the dimension n. Let $\{v_1, \cdots, v_k\}, \{v_{k+1}, \cdots, v_n\}$ be bases for W, W' respectively. By Exercise 1.6.33(b), we know $\{v_1, \cdots, v_n\}$ is a basis for $V = W \oplus W'$.

By hypothesis and definition of projection, we know

$$T(v_j) = \begin{cases} v_j & \text{for } 1 \leq j \leq k \\ 0 & \text{otherwise} \end{cases}.$$

Then $x = \sum_{i=1}^{n} a_i v_i \in V$ for $a_i \in F$. It follows that

$$T(x) = T\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i T(v_i) = \sum_{i=1}^{k} a_i v_i = (a_1, a_2, \cdots, a_k, 0, \cdots, 0)_{1\times n} \in W.$$

■

Exercise 2.2.16

Answer.

$$[T]_\beta^\gamma = \begin{pmatrix} O & O \\ O & I_{n-k} \end{pmatrix}.$$

where $n = \dim(V) = \dim(W)$ and $k = \text{nullity}(T)$.

Proof. Suppose $\dim(W) = \dim(V) =: n$. Let $\{v_1, \cdots, v_k\}$ be a basis for $N(T)$. Now extend it to a basis $\beta = \{v_1, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ for V. For $j = k+1, \cdots, n$, we set $w_j = T(v_j)$ where $T = [T]_\beta^\gamma$ is mentioned in the answer, then $\{w_{k+1}, \cdots, w_n\}$ is linear independent. Again, we extend it to a basis $\gamma = \{w_1, \cdots, w_k, w_{k+1}, \cdots, w_n\}$ for W. ■

2.3 Composition of Linear Transformations and Matrix Multiplication

Exercise 2.3.1

(a) **Answer.** False.

Solution. It should be $[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$. That is, it follows from Theorem 2.11. ■

(b) **Answer.** True.

Solution. It follows from Theorem 2.14. ■

(c) **Answer.** False.

Solution. It should be $[U(w)]_\beta = [U]_\alpha^\beta [w]_\alpha$. That is, it follows from Theorem 2.14. ■

(d) **Answer.** True.

Solution. It follows from Theorem 2.12(d). ■

(e) **Answer.** False.

Solution. Since $\dim(V)$ might not equal to $\dim(W)$. It only holds on $T : V \to V$. Since $[T^2]_\alpha = [T \cdot T]_\alpha = [T]_\alpha [T]_\alpha = [T]_\alpha^2$. ■

(f) **Answer.** False.

Solution. Let

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

Then $A^2 = I$; however, $A \neq I$ and $A \neq -I$. ■

(g) **Answer.** False.

Solution. It follows from Theorem 2.15(d). $T = L_A$ for some $A \in M_{m\times n}(F)$ if and only if $T : F_n \to F_m$. ■

(h) **Answer.** False.

Solution. Let

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

Then $A^2 = O$; however, $A \neq O$. ■

(i) **Answer.** True.

Solution. It follows from Theorem 2.15(c). ■

(j) **Answer.** True.

Solution. It follows from definition of δ_{ij} immediately. ■

Exercise 2.3.2

- Compute $A(2B + 3C)$.

  Answer.

$$A(2B + 3C) = \begin{pmatrix} 20 & -9 & 18 \\ 5 & 10 & 8 \end{pmatrix}.$$

Solution.

$$\begin{aligned} A(2B+3C) &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \cdot \left(2 \begin{pmatrix} 1 & 0 & -3 \\ 4 & 1 & 2 \end{pmatrix} + 3 \begin{pmatrix} 1 & 1 & 4 \\ -1 & -2 & 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 5 & 3 & 6 \\ 5 & -4 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 20 & -9 & 18 \\ 5 & 10 & 8 \end{pmatrix}. \end{aligned}$$

■

- Compute $(AB)D$.

Answer.

$$(AB)D = \begin{pmatrix} 29 \\ -26 \end{pmatrix}.$$

Solution.

$$\begin{aligned} (AB)D &= \left(\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -3 \\ 4 & 1 & 2 \end{pmatrix} \right) \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 13 & 3 & 3 \\ -2 & -1 & -8 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 29 \\ -26 \end{pmatrix}. \end{aligned}$$

■

- Compute $A(BD)$.

Answer.

$$A(BD) = \begin{pmatrix} 29 \\ -26 \end{pmatrix}.$$

Solution.

$$\begin{aligned} A(BD) &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \left(\begin{pmatrix} 1 & 0 & -3 \\ 4 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix} \right) \\ &= \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} -7 \\ 12 \end{pmatrix} \\ &= \begin{pmatrix} 29 \\ -26 \end{pmatrix}. \end{aligned}$$

■

Exercise 2.3.12

(a) **Proof.** Pick arbitrary $v_1, v_2 \in V$ such that $T(v_1) = T(v_2)$. Then $U(T(v_1)) = U(T(v_2))$ implies $(UT)(v_1) = (UT)(v_2)$. Since UT is one-to-one, we know $v_1 = v_2$. It follows that T is also one-to-one. ■

- Must U also be one-to-one?

Answer. No.

Proof. We pick $V = \mathbb{R}^2$, $W = Z = \mathbb{R}^3$. Suppose $T(a, b) = (a, b, 0)$ and $U(a, b, c) = (a, b, 0)$. Then UT and T are one-to-one. However, U is not one-to-one. ■

(b) **Proof.** Notice that $UT : V \to Z$ is onto. Let z is an arbitrary element of Z, then exists $v \in V$ such that $UT(v) = z$. This implies $U(T(v)) = z$. We pick $w = T(v)$, then we know $U(w) = U(T(v)) = z$. Hence U is also onto. ■

- Must T also be onto?

Answer. No.

Proof. We pick $V = Z = \mathbb{R}^2$ and $W = \mathbb{R}^3$. Suppose $T(a, b) = (a, b, 0)$ and $U(a, b, c) = (a, b)$. Then UT and U are onto. However, T is not onto. ■

(c) **Proof.** First we prove UT is one-to-one, then we prove UT is onto.

Pick arbitrary $v_1, v_2 \in V$ such that $(UT)(v_1) = (UT)(v_2)$. This implies $U(T(v_1)) = U(T(v_2))$. Since U is one-to-one, then $T(v_1) = T(v_2)$; also T is one-to-one, so $v_1 = v_2$. This means UT is one-to-one.

Pick arbitrary $z \in Z$. Since U is onto, then there exists $w \in W$ such that $U(w) = z$. Also since T is onto, then there exists $v \in V$ such that $T(v) = w$. They follow that $U(T(v)) = U(w) = z$. Then $UT(v) = z$. Hence UT is onto.

Finally, we conclude UT is one-to-one and onto (bijective). ■

Exercise 2.3.13

- $\text{tr}(AB) = \text{tr}(BA)$.

Proof. Consider

$$\begin{aligned}\text{tr}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij} B_{ji} \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} A_{ij} B_{ji} \right) \\ = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} B_{ji} A_{ij} \right) = \sum_{j=1}^{n} (BA)_{jj} = \text{tr}(BA).\end{aligned}$$

Hence we obtain what to prove. ■

- $\text{tr}(A) = \text{tr}(A^t)$.

Proof. Notice that $A_{ij} = A^t_{ji}$ for all i, j. So $A_{ii} = A^t_{ii}$. Consider

$$\text{tr}(A) = \sum_{i=1}^{n} A_{ii} = \sum_{i=1}^{n} A^t_{ii} = \text{tr}(A^t).$$

Hence we obtain what to prove. ■

2.4 Invertibility and Isomorphisms

Exercise 2.4.1

(a) **Answer.** False.

Solution. It follows from Theorem 2.18. So it should be $([T]_\alpha^\beta)^{-1} = [T^1]_\beta^\alpha$. ■

(b) **Answer.** True.

Solution. It follows from Appendix B. ■

(c) **Answer.** False.

Solution. Notice that L_A can only map F_n into F_m. However, T maps V into W. ■

(d) **Answer.** False.

Solution. It follows from Theorem 2.19. So it should be $M_{2\times 3}(F) = F^6$. ■

(e) **Answer.** True.

Solution. It follows from Theorem 2.19. ■

(f) **Answer.** False.

Solution. Pick

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \text{ and } B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

We know $AB = I$, but A, B are not invertible trivially. ■

(g) **Answer.** True.

Solution. It follows from Appendix B. ■

(h) **Answer.** True.

Solution. It follows from the Corollary 2 of Theorem 2.18. ■

(i) **Answer.** True.

Solution. It follows from definition immediately. ■

Exercise 2.4.2

(a) **Answer.** No.

Solution. Since T maps $\mathbb{R}^2$ into $\mathbb{R}^3$ where $\dim(\mathbb{R}^2) = 2 \neq 3 = \dim(\mathbb{R}^3)$. ■

(b) **Answer.** No.

Solution. Since T maps $\mathbb{R}^2$ into $\mathbb{R}^3$ where $\dim(\mathbb{R}^2) = 2 \neq 3 = \dim(\mathbb{R}^3)$. ■

(c) **Answer.** Yes.

Solution. Since

$$T = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 0 & 1 \\ 3 & 4 & 0 \end{pmatrix},$$

then we have

$$T^{-1} = \frac{1}{12}\begin{pmatrix} 4 & 8 & 0 \\ -3 & -6 & 3 \\ 0 & 12 & 0 \end{pmatrix}.$$

Hence $T^{-1}(a_1, a_2, a_3) = \frac{1}{12}(4a_1 + 8a_2, -3a_1 - 6a_2 + 3a_3, 12a_2)$. ■

(d) **Answer.** No.

Solution. Since T maps $P_3(\mathbb{R})$ into $P_2(\mathbb{R})$ where $\dim(P_3(\mathbb{R})) = 4 \neq 3 = \dim(P_2(\mathbb{R}))$. ■

(e) **Answer.** No.

Solution. Since T maps $M_{2\times 2}(\mathbb{R})$ into $P_2(\mathbb{R})$ where $\dim(M_{2\times 2}(\mathbb{R})) = 4 \neq 3 = \dim(P_2(\mathbb{R}))$. ■

(f) **Answer.** Yes.

Solution. Let $\beta := \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$. Put $\phi_\beta : M_{2\times 2}(\mathbb{R}) \to F^4$ such that $\phi_\beta \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a, b, c, d)$. And define $U : F_4 \to F_4$ by $U(a, b, c, d) = (a + b, a, c, c + d)$. We know

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

This implies

$$U^{-1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

That is, $U^{-1}(a, b, c, d) = (b, a - b, c, d - c)$. Notice that $T^{-1} = \phi_\beta^{-1} U^{-1} \phi_\beta$. So

$$T^{-1} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} b & a - b \\ c & d - c \end{pmatrix}.$$

■

Exercise 2.4.3

They follow from Theorem 2.19 immediately.

(a) **Answer.** No.

Solution. Since $\dim(F^3) = 3 \neq 4 = \dim(P_3(F))$. ■

(b) **Answer.** Yes.

Solution. Since $\dim(F^4) = 4 = \dim(P_3(F))$. ■

(c) **Answer.** Yes.

Solution. Since $\dim(M_{2\times 2}(\mathbb{R})) = 4 = \dim(P_3(\mathbb{R}))$. ■

(d) **Answer.** No.

Solution. Notice that $\{E_{11} + E_{22}, E_{12}, E_{21}\}$ is a basis for V. So $\dim(V) = 3 \neq 4 = \dim(\mathbb{R}^4)$. ■

Exercise 2.4.5

Proof. Since A is invertible, then $A \cdot A^{-1} = A^{-1} \cdot A = I$. Consider

$$A^t(A^{-1})^t = (A^{-1} \cdot A)^t = I^t = I;$$
$$(A^{-1})^t A^t = (A \cdot A^{-1})^t = I^t = I.$$

Hence A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$. ■

Exercise 2.4.6

Proof. Since A is invertible, then $A^{-1}A = I$. Consider

$$AB = O.$$

This implies

$$A^{-1}AB = A^{-1}O.$$

Then

$$B = IB = A^{-1}AB = A^{-1}O = O.$$

■

2.5 The Change of Coordinate Matrix

Exercise 2.5.1

(a) **Answer.** False.

Solution. The jth column of Q is $[x'_j]_{\beta'}$. ■

(b) **Answer.** True.

Solution. It follows from Theorem 2.22. ■

(c) **Answer.** False.

Solution. It follows from Theorem 2.23. ■

(d) **Answer.** False.

Solution. That means $B = Q^{-1}AQ$, not $B = Q^tAQ$. ■

(e) **Answer.** True.

Solution. It follows from Theorem 2.23 and definition of similarity. ■

Exercise 2.5.2

We denote the change of coordinate matrix by Q.

(a) **Answer.** $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$.

Solution. Since $(a_1, a_2) = a_1e_1 + a_2e_2$ and $(b_1, b_2) = b_1e_1 + b_2e_2$. So $[(a_1, a_2)]_\beta = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ and $[(b_1, b_2)]_\beta = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. Then we know $Q = [I]_{\beta'}^{\beta} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$. ■

(b) **Answer.** $\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$.

Solution. Since $(0, 10) = 4(-1, 3) + 2(2, -1)$ and $(5, 0) = (-1, 3) + 3(2, -1)$. So $[(0, 10)]_\beta = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ and $[(5, 0)]_\beta = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$. Then we know $Q = [I]_{\beta'}^{\beta} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$. ■

(c) **Answer.** $\begin{pmatrix} 3 & -1 \\ 5 & -2 \end{pmatrix}$.

Solution. Since $(1, 0) = 3(2, 5) + 5(-1, -3)$ and $(0, 1) = -(2, 5) - 2(-1, -3)$. So $[e_1]_\beta = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ and $[e_2]_\beta = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$. Then we know $Q = [I]_{\beta'}^{\beta} = \begin{pmatrix} 3 & -1 \\ 5 & -2 \end{pmatrix}$. ■

(d) **Answer.** $\begin{pmatrix} 2 & -1 \\ 5 & -4 \end{pmatrix}$.

Solution. Since $(2, 1) = 2(-4, 3) + 5(2, -1)$ and $(-4, 1) = -(-4, 3) - 4(2, -1)$. So $[(2, 1)]_\beta = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$ and $[(-4, 1)]_\beta = \begin{pmatrix} -1 \\ -4 \end{pmatrix}$. Then we know $Q = [I]_{\beta'}^{\beta} = \begin{pmatrix} 2 & -1 \\ 5 & -4 \end{pmatrix}$. ■

Exercise 2.5.7

(a) **Answer.** $T(x, y) = \frac{1}{m^2+1}\left((1 - m^2)x + 2my, 2mx + (m^2 - 1)y\right)$.

Solution. Consider the line $L : y = mx$ where $m \neq 0$. Since $(1, m) \in L$, so the reflection of $(1, m)$ over L is $(1, m)$ itself. Then consider another line $L' : y' = -m^{-1}x'$ such that $L \perp L'$. We observe $(-m, 1) \in L'$. Notice that the reflection of $(-m, 1)$ over L is $(m, -1)$.

Pick $\beta' = \{(1, m), (-m, 1)\}$. Also we have

$$T(1, m) = (1, m), \text{ and } T(-m, 1) = (m, -1).$$

Since

$$(1,m) = (1,m) + 0(m,-1), \text{ and } (m,-1) = 0(1,m) - (-m,1).$$

It follows that $[T]_{\beta'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Now pick $\beta = \{e_1, e_2\}$ which is a standard basis for $\mathbb{R}^2$. So we know $[I]_{\beta'}^{\beta} = \begin{pmatrix} 1 & -m \\ m & 1 \end{pmatrix}$. Compute

$$[I]_{\beta}^{\beta'} = ([I]_{\beta'}^{\beta})^{-1} = \frac{1}{m^2+1}\begin{pmatrix} 1 & m \\ -m & 1 \end{pmatrix}.$$

Hence

$$\begin{aligned} [T]_\beta &= [I]_{\beta'}^{\beta}[T]_{\beta'}[I]_{\beta}^{\beta'} \\ &= \begin{pmatrix} 1 & -m \\ m & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \frac{1}{m^2+1}\begin{pmatrix} 1 & m \\ -m & 1 \end{pmatrix} \\ &= \frac{1}{m^2+1}\begin{pmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{pmatrix}. \end{aligned}$$

Finally, the answer follows from $[T]_\beta \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{m^2+1}\begin{pmatrix} (1-m^2)x + 2my \\ 2mx + (m^2-1)y \end{pmatrix}$. ■

(b) **Answer.** $T(x,y) = \frac{1}{m^2+1}(x + my, mx + m^2y)$.

Solution. Define L and L' be the same as part (a). Notice that $\mathbb{R} = L \oplus L'$. If $p \in \mathbb{R}^2$ such that $p = p_1 + p_2$ where $p_1 \in L$ and $p_2 \in L'$, then $T(p) = p_1$. Since $(1,m) \in L$, then $T(1,m) = T(1,m)$; on the other hand, since $(-m,1) \in L'$, then $T(-m,1) = (0,0)$.

Now we pick β, β' which are the same as part (a). It follows that $[T]_{\beta'} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Of course,

$$[I]_{\beta'}^{\beta} = \begin{pmatrix} 1 & -m \\ m & 1 \end{pmatrix}, \text{ and } [I]_{\beta}^{\beta'} = \frac{1}{m^2+1}\begin{pmatrix} 1 & m \\ -m & 1 \end{pmatrix}.$$

Hence $[T]_\beta = [I]_{\beta'}^{\beta}[T]_{\beta'}[I]_{\beta}^{\beta'} = \frac{1}{m^2+1}\begin{pmatrix} 1 & m \\ m & m^2 \end{pmatrix}$.

Finally, the answer follows from $[T]_\beta \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{m^2+1}\begin{pmatrix} x + my \\ mx + m^2y \end{pmatrix}$. ■

2.6 Dual Spaces

Exercise 2.6.1

(a) **Answer.** False.

Solution. The codomain of linear functional must be F. It should be that every linear functional is a linear transformation. ■

(b) **Answer.** True.

Solution. The linear functional maps F into F, so we can use a 1×1 matrix to represent. ■

(c) **Answer.** True.

Solution. Since $\dim(V) = \dim(V^*)$. ■

(d) **Answer.** False.

Solution. Here we interpret "is" as "equals". Since the codomain of a dual space must be F. ■

(e) **Answer.** False.

Solution. Let $\beta := \{x_1, x_2, \cdots, x_n\}$ and $\beta^* := \{f_1, f_2, \cdots, f_n\}$. Pick $T(x_i) = 2f_i$ for each i, then we observe $T(\beta) = 2\beta^*$. ■

(f) **Answer.** True.

Solution. Since $T^t : W^* \to V^*$, then $(T^t)^t : V^{**} \to W^{**}$. So the domain of $(T^t)^t$ is V^{**}. ■

(g) **Answer.** True.

Solution. Since $\dim(V) = \dim(V^*)$ and $\dim(W) = \dim(W^*)$. By hypothesis that $\dim(V) = \dim(W)$, we have $\dim(V^*) = \dim(W^*)$ immediately. ■

(h) **Answer.** False.

Solution. The codomain of linear functional must be F. However, the domain of the derivative of a function might not always be F. ■

Exercise 2.6.3

Let $\dim(V) = n$, then we suppose $\beta^* = \{f_1, f_2, \cdots f_n\}$.

(a) **Answer.** $f_1(x,y,z) = x - \frac{y}{2}, f_2(x,y,z) = \frac{y}{2}, f_3(x,y,z) = x - z$.

Solution. We should solve the systems of linear equations such that

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_1(e_1) & f_2(e_1) & f_3(e_1) \\ f_1(e_2) & f_2(e_2) & f_3(e_2) \\ f_1(e_3) & f_2(e_3) & f_3(e_3) \end{pmatrix} = I_3.$$

Hence we have

$$\begin{pmatrix} f_1(e_1) & f_2(e_1) & f_3(e_1) \\ f_1(e_2) & f_2(e_2) & f_3(e_2) \\ f_1(e_3) & f_2(e_3) & f_3(e_3) \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The answer follows from the change of notations. ■

(b) **Answer.** $f_1(a_0 + a_1x + a_2x^2) = a_0, f_2(a_0 + a_1x + a_2x^2) = a_1, f_3(a_0 + a_1x + a_2x^2) = a_2$.

Solution. We should solve the systems of linear equations such that

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_1(1) & f_2(1) & f_3(1) \\ f_1(x) & f_2(x) & f_3(x) \\ f_1(x^2) & f_2(x^2) & f_3(x^2) \end{pmatrix} = I_3.$$

Hence we have

$$\begin{pmatrix} f_1(1) & f_2(1) & f_3(1) \\ f_1(x) & f_2(x) & f_3(x) \\ f_1(x^2) & f_2(x^2) & f_3(x^2) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The answer follows from the change of notations. ■

Exercise 2.6.5

Proof. Suppose $p(x) = ax + b$. Then we have

$$f_1(p(x)) = \int_0^1 (at + b)dt = \frac{1}{2}a + b;$$
$$f_2(p(x)) = \int_0^2 (at + b)dt = 2a + 2b.$$

We can observe that

$$ax + b = \left(\frac{1}{2}a + b\right)(-2x + 2) + (2a + 2b)\left(-\frac{1}{2} + x\right).$$

Consider

$$k(-2x + 2) + t\left(x - \frac{1}{2}\right) = 0.$$

This implies $t = k = 0$. It follows that $\left\{-2x + 2, x - \frac{1}{2}\right\}$ is a basis for V. By Theorem 2.24, we know $\{f_1, f_2\}$ is basis for V^*. ■

Exercise 2.6.9

Proof. Follow the hint. Let $\{e_1, e_2, \cdots, e_m\}$ be the standard basis for F^m.

($\Longrightarrow$) Since $g_i \in (F^m)^*$ for each $i = 1, 2, \cdots, m$, then we know for all $1 \le i, j \le m$,

$$g_i(e_j) = \delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Given arbitrary $x \in F^n$, then

$$T(x) = (c_1, c_2, \cdots, c_m) = \sum_{j=1}^{m} c_j e_j$$

where $c_j \in F$ for each j. Since T, g_i are linear, we consider

$$f_i(x) = (g_i T)(x) = g_i(T(x)) = g_i(\sum_{j=1}^{m} c_j e_j) = \sum_{j=1}^{m} c_j g_i(e_j) = c_i.$$

It follows that $(f_1(x), f_2(x), \cdots, f_m(x)) = (c_1, c_2, \cdots, c_m) = T(x)$.
Since x was arbitrary, we obtain

$$T(x) = (f_1(x), f_2(x), \cdots, f_m(x)), \forall x \in F^n.$$

$(\Longleftarrow)$ Notice that f_i is linear. For $x, y \in F^n$ and $c \in F$, we consider

$$\begin{aligned} T(x+cy) &= (f_1(x+cy), f_2(x+cy), \cdots, f_m(x+cy)) \\ &= (f_1(x) + cf_1(y), f_2(x) + cf_2(y), \cdots, f_m(x) + cf_m(y)) \\ &= (f_1(x), f_2(x), \cdots, f_m(x)) + c(f_1(y), f_2(y), \cdots, f_m(y)) \\ &= T(x) + cT(y). \end{aligned}$$

Hence T is linear. ■

Exercise 2.6.19

Proof. Since $W \subset V$ but $W \neq V$ by hypothesis, we can pick $x_0 \in V \setminus W$. Let β be a basis for W. Since $x_0 \notin W$, then we have $\beta \cup \{x_0\}$ is linearly independent. Now we extend $\beta \cup \{x_0\}$ to β' which is a basis for V.

Define a function $g : \beta' \to F$ by

$$g(x) = \begin{cases} 1 & \text{for} x = x_0 \\ 0 & \text{otherwise} \end{cases}.$$

Hence we know $g(x) = 0$ for all $x \in \beta$.

By Exercise 2.1.34, we know there exists the unique linear transformation $f : \beta' \to F$ such that $f(x) = g(x)$ for all $x \in \beta'$. Then, we can write $y \in W$ as $y = \sum_{j=1}^{n} c_j y_j$ where $y_j \in \beta$ and $c_j \in F$. Consider

$$f(y) = f(\sum_{j=1}^{n} c_j y_j) = \sum_{j=1}^{n} c_j f(x_j) = \sum_{j=1}^{n} c_j g(x_j) = 0$$

for all $y \in W$. Notice that $f(x_0) = g(x_0) = 1$ implies that f is nonzero. In particular, f maps β' to F, we know f is a linear functional. i.e., f is what we desire. ■

Exercise 2.6.20

(a) **Proof.** ($\Longrightarrow$) Suppose $T^t(f) = 0$ for some $f \in W^*$. Then $T^t(f) = f(T(x)) = 0$ for all $x \in V$. Since T is onto by assumption, then for $y \in W$, we have $y = T(x)$ for some $x \in V$. i.e., $f(y) = f(T(x)) = 0$. Hence T^t is one-to-one.

($\Longleftarrow$) Suppose to contrary that T is not onto. This is equivalent to $R(T) \neq W$. Then by Exercise 2.6.19, there exists a linear functional $f \in W^*$ such that $f(x) = 0, \forall x \in R(T)$ with $f \neq 0$.

Let $g = (T^t)f \in V^*$. For arbitrary $x \in V$, since $T(x) \in R(T)$, we have

$$g(x) = ((T^t)f)(x) = f(T(x)) = 0.$$

Because x was arbitrary, this means $g(x) = 0$ for all $x \in V$. Then $(T^t)f = 0$. However $f \neq 0$, this contradicts T^t is one-to-one by hypothesis. Hence T is onto. ■

(b) **Proof.** ($\Longrightarrow$) Let $x_0 \in V$. So if T is one-to-one, then $T(x_0) = 0$ implies $x_0 = 0$. Suppose to contrary that $x_0 \neq 0$. We pick a basis β for V with $x_0 \in \beta$.

Define the function $h : \beta \to F$ by

$$h(x) = \begin{cases} 1 & \text{for} x = x_0 \\ 0 & \text{otherwise} \end{cases}.$$

By exercise 2.6.19, there exists a linear functional $f \in V^*$ such that $f(x) = h(x)$ for all $x \in \beta$ with $f \neq 0$.

Since T^t is onto by hypothesis, then $(T^t)g = f$ for some $g \in W^*$; that is, for all $x \in V$, $g(T(x)) = f(x)$.

Now pick $x \in \beta$ with $x \neq x_0$, then $f(x) = h(x) = 0$. On the other hand, notice that g is a linear functional, so $g(0) = 0$. It follows that $f(x_0) = g(T(x_0)) = g(0) = 0$.

We have known $f(x) = 0$ for all $x \in \beta$ and β is a basis for V, so $f = 0$. This contradicts that $f \neq 0$. Hence $x_0 = 0$, and it follows T is one-to-one.

($\Longleftarrow$) Let β be a basis for V. Since T is one-to-one, then $T(\beta)$ is a linearly independent set in W. We can extend $T(\beta)$ to a basis β' for W.

Define $h : \beta' \to F$ by $(h \circ T)(x) = g(x)$ for $x \in \beta$, and $h(x) = 0$ otherwise. Notice that $g \in V^*$. By Exercise 2.1.34, since β' is a basis for W, there exists a linear transformation $f : W \to F$ such that $f(x) = h(x)$ for all $x \in \beta'$; that is, $f \in W^*$.

Consider arbitrary $x \in \beta \subseteq \beta'$, we have

$$g(x) = h(T(x)) = f(T(x)) = T^t f(x).$$

Because x was arbitrary, then this means for every $g \in V^*$, there exists $f \in W^*$ such that $T^t = g$. Hence T^t is onto. ■

Chapter 3

Elementary Matrix Operations and Systems of Linear Equations

3.1 Elementary Matrix Operations and Elementary Matrices

Exercise 3.1.1

(a) **Answer**. True.

Solution. It follows from definition immediately. ■

(b) **Answer**. False.

Solution. $2I$ is an elementary matrix of type 2. ■

(c) **Answer**. True.

Solution. I_n is an elementary matrix of type 2 with scalars 1. ■

(d) **Answer**. False.

Solution. Consider $E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ and $E_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. We have $E_1 E_2 = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ which is not an elementary matrix. ■

(e) **Answer**. True.

Solution. It follows from Theorem 3.2. ■

(f) **Answer**. False.

Solution. Consider $E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $E_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. We have $E_1 + E_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ which is not an elementary matrix. ■

(g) **Answer.** True.

Solution. It follows from Exercise 3.1.5. ■

(h) **Answer.** False.

Solution. It fails trivially. ■

(i) **Answer.** True.

Solution. Since $B = EA$ where E is an elementary matrix of row operation, then we have $A = E^{-1}B$ where E^{-1} is also an elementary matrix of row operation. ■

Exercise 3.1.12

Proof. For $m = 1$, we are done. Suppose the assertion holds for $m - 1$, then for m, we pick b such that $A_{a,b'} = 0$ for all $1 \leq a \leq m$ and for all $b' < b$. Then pick a' such that $A_{a',b} \neq 0$, we can use a sequence of elementary row operations of type 1 and 3 to transform $A_{1,b} \neq 0$. Let B be the $(m-1) \times n$ matrix such that B comes from deleting the first row of A. By supposition, we know B is an upper triangular matrix. Now we recover the first row of B. This implies A is also an upper triangular matrix.

By induction, since m is finite, we can find a finite sequence of such row operations to make A be an upper triangular matrix. ■

3.2 The Rank of a Matrix and Matrix Inverses

Exercise 3.2.1

(a) **Answer.** False.

Solution. Consider $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has two nonzero column vectors. However, its rank is 1, not 2. ■

(b) **Answer.** False.

Solution. Let $A := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, then $\operatorname{rank}(A) = 1$. But we consider $A \times A = O$ where $\operatorname{rank}(O) = 0$. This means $\operatorname{rank}(O) \neq \operatorname{rank}(A)$. ■

(c) **Answer.** True.

Solution. It follows from Exercise 3.2.3. ■

(d) **Answer.** True.

Solution. It follows from the Corollary of Theorem 3.4. ■

(e) **Answer.** False.

Solution. It follows from the Corollary of Theorem 3.4. ■

(f) **Answer.** True.

Solution. It follows from the Corollary 2(b) of Theorem 3.6. ■

(g) **Answer.** True.

Solution. See the page 162. ■

(h) **Answer.** True.

Solution. Since the rank of a $m \times n$ matrix must be less than or equal to m and n. ■

(i) **Answer.** True.

Solution. Let A be the $n \times n$ matrix with rank n. We know L_A maps F^n into F^n and L_A is onto. It follows that L_A is one-to-one. Hence L_A^{-1} exists so that A is invertible. ■

Exercise 3.2.3

Proof. ($\Longrightarrow$) Suppose A is not a zero matrix. Then $A_{ij} \neq 0$ for some i, j, This means $\operatorname{rank}(A) \neq 0$. This argument is equivalent to $\operatorname{rank}(A) = 0$ implies A is a zero matrix.

($\Longleftarrow$) Since A is a zero matrix, then by definition, $\operatorname{rank}(A) = 0$ trivially. ■

Exercise 3.2.4

(a) **Answer.** $D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. $\operatorname{rank}(D) = 2$.

Solution. $\begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 0 & -1 & 2 \\ 1 & 1 & 1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & -2 & -3 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -\frac{1}{2} & 2 \\ 0 & 1 & \frac{3}{2} & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow D.$ ■

(b) **Answer.** $D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. $\operatorname{rank}(D) = 2$.

Solution. $\begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 2 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -2 \\ 2 & 1 \\ 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -2 \\ 0 & 5 \\ 0 & 0 \end{pmatrix} \longrightarrow D.$ ■

Exercise 3.2.5

We denoted the matrix by A.

(a) **Answer.** $A^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$. $\operatorname{rank}(A) = 2$.

Solution. $\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{array}\right).$ ■

(b) **Answer.** A^{-1} doesn't exist. $\operatorname{rank}(A) = 1$.

Solution. $\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right).$ ■

(c) **Answer.** A^{-1} doesn't exist. $\operatorname{rank}(A) = 2$.

Solution.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 3 & 4 & 0 & 1 & 0 \\ 2 & 3 & -1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & -1 & -3 & -2 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 0 & -3 & 1 & 1 \end{array}\right).$$

■

Exercise 3.2.6

We pick the standard basis α for domain of T, and the standard basis β for codomain of T.

(a) **Answer.** $[T]_\alpha^\beta$ is invertible; $[T^{-1}]_\beta^\alpha = \begin{pmatrix} -1 & -2 & -10 \\ 0 & -1 & -4 \\ 0 & 0 & -1 \end{pmatrix}$.

Solution. Notice that $T(ax^2 + bx + c) = -ax^2 + (4a - b)x + (2a + 2b - c)$.

Then we know $[T]_\alpha^\beta = \begin{pmatrix} -1 & 2 & 2 \\ 0 & -1 & 4 \\ 0 & 0 & -1 \end{pmatrix}$. ■

(b) **Answer.** $[T]_\alpha^\beta$ is not invertible.

Solution. Notice that $T(ax^2+bx+c) = 2ax^2 + (2a+b)x + b$.

Then we know $[T]_\alpha^\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$. But $\operatorname{rank}(T) = 2 \neq 3$. ■

(c) **Answer.** $[T]_\alpha^\beta$ is invertible; $[T^{-1}]_\beta^\alpha = \frac{1}{6}\begin{pmatrix} 1 & -2 & 3 \\ 3 & 0 & -3 \\ -1 & 2 & 3 \end{pmatrix}$.

Solution. We know $[T]_\alpha^\beta = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$. ■

Exercise 3.2.7

Solution. Let $A := \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. Consider

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} \xrightarrow{E_1} \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} \xrightarrow{E_2} \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{E_3} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$
$$\xrightarrow{E_4} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{E_5} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{E_6} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Put

$$E_1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}; E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix};$$
$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}; E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}; E_6 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It follows that $E_6E_5E_4E_3E_2E_1A = I$. Then $A = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}E_5^{-1}E_6^{-1}$ where

$$E_1^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}; E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$
$$E_4^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}; E_5^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}; E_6^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

■

Exercise 3.2.14

(a) **Proof.** Let arbitrary $w \in R(T+U)$, then there exists $v \in V$ such that $w = (T+U)(v) = T(v) + U(v)$. Since $T(v) \subseteq R(T)$ and $U(v) \subseteq R(U)$, we have $T(v) + U(v) \subseteq R(T) + R(U)$. Because w was arbitrary, we conclude $R(T+U) \subseteq R(T) + R(U)$. ■

(b) **Proof.** Since $\dim(W)$ is finite, by Exercise 1.6.29(a), we have

$$\begin{aligned}\operatorname{rank}(T+U) &\leq \dim(R(T) + R(U)) \\ &= \dim(R(T)) + \dim(R(U)) - \dim(R(T) \cap R(U)) \\ &\leq \dim(R(T)) + \dim(R(U)) \\ &= \operatorname{rank}(T) + \operatorname{rank}(U).\end{aligned}$$

■

(c) **Proof.** Form part (b), we consider $\operatorname{rank}(A+B) = \operatorname{rank}(L_{A+B}) = \operatorname{rank}(L_A + L_B) \leq \operatorname{rank}(L_A) + \operatorname{rank}(L_B) = \operatorname{rank}(A) + \operatorname{rank}(B)$. ■

Exercise 3.2.17

Proof. ($\Longrightarrow$) Since $\operatorname{rank}(B) \leq \min\{3, 1\} = 1$ and $\operatorname{rank}(C) \leq \min\{1, 3\} = 1$, then by Theorem 3.7, we have

$$\operatorname{rank}(BC) \leq \min\{\operatorname{rank}(B), \operatorname{rank}(C)\} \leq \min\{1, 1\} = 1.$$

($\Longleftarrow$) Since $A \in M_{3\times 3}(F)$ with $\operatorname{rank}(A) = 1$. This means $R(A) \cong F^1$. Therefore we can map F^3 into F^1 by $B \in M_{1\times 3}(F)$, and then map F into F^3 by $C \in M_{3\times 1}(F)$. Hence $A = BC$. ■

3.3 Systems of Linear Equations - Theoretical Aspects

Exercise 3.3.1

(a) **Answer.** False.

Solution. Consider $\begin{cases} x = 0 \\ x = 1 \end{cases}$. Then it has no solution. ■

(b) **Answer.** False.

Solution. Consider $\begin{cases} x_1 + x_2 = 2 \\ 2x_1 + 2x_2 = 4 \end{cases}$. Then it has infinitely many solutions. ■

(c) **Answer.** True.

Solution. The solutions at least contain zero solution trivially. ■

(d) **Answer.** False.

Solution. Consider $\begin{cases} x_1 + x_2 = 2 \\ 2x_1 + 2x_2 = 4 \end{cases}$. Then it has infinitely many solutions. ■

(e) **Answer.** False.

Solution. Consider $\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 = 1 \end{cases}$. Then it has no solution. ■

(f) **Answer.** False.

Solution. Pick $0x = 0$. Then it has infinitely many solutions. However, $0x = 1$ has no solution. ■

(g) **Answer.** True.

Solution. Let A be a $n \times n$ matrix. Then we know $Ax = 0$. It follows that $x = A^{-1}0 = 0$. So the solution must be zero. ■

(h) **Answer.** False.

Solution. Since $x = 1$ has the solution set $\{1\}$, it has no zero vector in F^1, then it violate the definition of subspace. ■

Exercise 3.3.2

Let the solution set be S.

(a) **Answer.** $\dim(S) = 1$; $\left\{ \begin{pmatrix} -3 \\ 1 \end{pmatrix} \right\}$ is a basis for S.

Solution. Put $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$. Since $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$. Then $\text{rank}(A) = 1$ and $S = \{t(-3,1) : t \in \mathbb{R}\}$. ■

(b) **Answer.** $\dim(S) = 1$; $\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$ is a basis for S.

Solution. Put $A = \begin{pmatrix} 1 & 1 & -1 \\ 4 & 1 & -2 \end{pmatrix}$. Since $\begin{pmatrix} 1 & 1 & -1 \\ 4 & 1 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -1 \\ 0 & -3 & 2 \end{pmatrix}$. Then $\text{rank}(A) = 1$ and $S = \{t(1,2,3) : t \in \mathbb{R}\}$. ■

Exercise 3.3.3

(a) **Answer.** $\left\{\begin{pmatrix}5\\0\end{pmatrix}+t\begin{pmatrix}-3\\1\end{pmatrix} : t\in\mathbb{R}\right\}$.

Solution. Since $\begin{cases}x_1=5\\x_2=0\end{cases}$ is a particular solution. Then the answer follows from Exercise 3.3.2(a). ■

(b) **Answer.** $\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}+t\begin{pmatrix}1\\2\\3\end{pmatrix} : t\in\mathbb{R}\right\}$.

Solution. Since $\begin{cases}x_1=1\\x_2=1\\x_3=1\end{cases}$ is a particular solution. Then the answer follows from Exercise 3.3.2(b). ■

Exercise 3.3.8

(a) **Answer.** No.

Solution. This means whether $\begin{cases}a+b=1\\b-2c=3\\a+2c=-2\end{cases}$ has a solution or not. After computing, we have $\left(\begin{array}{ccc|c}1&1&0&1\\0&1&-2&3\\0&0&0&4\end{array}\right)$, it implies the systems of linear equations has no solution. Hence $v\notin R(T)$. ■

(b) **Answer.** Yes.

Solution. This means whether $\begin{cases}a+b=2\\b-2c=1\\a+2c=1\end{cases}$ has a solution or not. After computing, we have $\left(\begin{array}{ccc|c}1&1&0&2\\0&1&-2&1\\0&0&0&0\end{array}\right)$, it implies $T(-1,3,1)=(2,1,1)=v$. Hence $v\in R(T)$. ■

Exercise 3.3.10

Proof. Let A be the coefficient matrix, then the augmented matrix is $A|b$. Since $A|b$ is a $m \times (n+1)$ matrix, we know $\text{rank}(A|b) \leq m$. Notice that $\text{rank}(A|b) \geq \text{rank}(A) = m$. Hence $\text{rank}(A|b) = m$. Moreover, because $\text{rank}(A) = \text{rank}(A|b) = m$, by Theorem 3.11, $Ax = b$ must have at least one solution. ■

3.4 Systems of Linear Equations - Computational Aspects

Exercise 3.4.1

(a) **Answer.** False.

Solution. It follows from the Corollary of Theorem 3.13. ■

(b) **Answer.** True.

Solution. It follows from the Corollary of Theorem 3.13. ■

(c) **Answer.** True.

Solution. It follows from Theorem 3.16. ■

(d) **Answer.** True.

Solution. It follows from Theorem 3.14. ■

(e) **Answer.** False.

Solution. Consider $0x = 1$, then the augmented matrix $(0|1)$ has no solution. ■

(f) **Answer.** True.

Solution. It follows from Theorem 3.15. ■

(g) **Answer.** True.

Solution. It follows from Theorem 3.16. ■

Exercise 3.4.2

(a) **Answer.** $(x_1, x_2, x_3) = (4, -3, -1)$.

Solution. $\left(\begin{array}{ccc|c} 1 & 2 & -1 & -1 \\ 2 & 2 & 1 & 1 \\ 3 & 5 & -2 & -1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & -1 & -1 \\ 0 & -2 & 3 & 3 \\ 0 & -1 & 1 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & -1 & -1 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right).$

The solution follows from $\begin{cases} x_1 + 2x_2 - x_3 = -1 \\ \quad - x_2 + x_3 = 2 \\ \quad\quad x_3 = -1 \end{cases}.$ ■

(f) **Answer.** $(x_1, x_2, x_3, x_4) \in \{(-3, 3, 1, 0) + t(1, -2, 0, 1) : t \in \mathbb{R}\}$.

Solution. $\left(\begin{array}{cccc|c} 1 & 2 & -1 & 3 & 2 \\ 2 & 4 & -1 & 6 & 5 \\ 0 & 1 & 0 & 2 & 3 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 2 & -1 & 3 & 2 \\ 0 & 1 & 0 & 2 & 3 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right)$. Consider

$$\begin{cases} x_1 + 2x_2 - x_3 + 3x_4 = 2 \\ \quad x_2 + \quad + 2x_4 = 3 \\ \quad\quad x_3 \quad = 1 \end{cases}.$$

Put $x_4 = t \in \mathbb{R}$, then the solution is $\begin{cases} x_1 = t - 3 \\ x_2 = -2t + 3 \\ x_3 = 1 \\ x_4 = t \end{cases}.$

■

Exercise 3.4.12

(a) **Proof.** Since $(0, -1, 0, 1, 1, 0), (1, 0, 1, 1, 1, 0) \in V$ and consider

$$s(0, -1, 0, 1, 1, 0) + t(1, 0, 1, 1, 1, 0) = (0, 0, 0, 0, 0, 0).$$

This leads to $s, t = 0$, we know S is a linearly independent subset of V. ■

(b) **Solution.** Consider $\begin{pmatrix} 1 & -1 & 0 & 2 & -3 & 1 \\ 2 & -1 & -1 & 3 & -4 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 0 & 2 & -3 & 1 \\ 0 & 1 & -1 & -1 & 2 & 2 \end{pmatrix}.$

We put $x_3 = a, x_4 = b, x_5 = c, x_6 = d$, then

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \left\{ a \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c \begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + d \begin{pmatrix} -3 \\ -2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} : a, b, c, d, \in \mathbb{R} \right\}.$$

Consider the matrix

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

Now reduce it to obtain

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

We pick the first, second, forth, sixth column to know that we can extend S to a basis

$$\{(0,-1,0,1,1,0),(1,0,1,1,1,0),(-1,1,0,1,0,0),(-3,-2,0,0,0,1)\}$$

for V as desired. ■

Chapter 4

Determinants

4.1 Determinants of Order 2

Exercise 4.1.1

(a) **Answer**. False.

Solution. $\det(2I_2) = 4 \neq 2\det(I_2)$ where $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. ■

(b) **Answer**. True.

Solution. Consider

$$\begin{aligned} \begin{vmatrix} a_1 + ka_2 & b_1 + kb_2 \\ c & d \end{vmatrix} &= (a_1 + ka_2)d - (b_1 + kb_2)c \\ &= (a_1 d - b_1 c) + k(a_2 d - b_2 c) \\ &= \begin{vmatrix} a_1 & b_1 \\ c & d \end{vmatrix} + k \begin{vmatrix} a_2 & b_2 \\ c & d \end{vmatrix}. \end{aligned}$$

A similar argument shows that $\begin{vmatrix} a & b \\ c_1 + kc_2 & d_1 + kd_2 \end{vmatrix} = \begin{vmatrix} a & b \\ c_1 & d_1 \end{vmatrix} + k \begin{vmatrix} a & b \\ c_2 & d_2 \end{vmatrix}$. ■

(c) **Answer**. False.

Solution. Suppose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, if A is invertible, then

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

It is easy to observe when $\det(A) = 0$, then A^{-1} is undefined. ■

(d) **Answer.** False.

Solution. The area should equal to $\left|\det\begin{pmatrix}u\\v\end{pmatrix}\right|$. i.e., the area can not be negative. ■

(e) **Answer.** True.

Solution. It follows from Exercise 4.1.12. ■

Exercise 4.1.11

Proof. First, we need to know the values of $\delta\begin{pmatrix}1&0\\1&0\end{pmatrix}, \delta\begin{pmatrix}1&0\\0&1\end{pmatrix}, \delta\begin{pmatrix}0&1\\1&0\end{pmatrix}, \delta\begin{pmatrix}0&1\\0&1\end{pmatrix}$.

By property (ii), we know

$$\delta\begin{pmatrix}1&0\\1&0\end{pmatrix} = \delta\begin{pmatrix}0&1\\0&1\end{pmatrix} = 0;$$

also by property (iii), we know $\delta\begin{pmatrix}1&0\\0&1\end{pmatrix} = 1$. Since

$$\begin{aligned}0 = \delta\begin{pmatrix}1&1\\1&1\end{pmatrix} &= \delta\begin{pmatrix}1&0\\1&1\end{pmatrix} + \delta\begin{pmatrix}0&1\\1&1\end{pmatrix}\\ &= \left(\delta\begin{pmatrix}1&0\\1&0\end{pmatrix} + \delta\begin{pmatrix}1&0\\0&1\end{pmatrix}\right) + \left(\delta\begin{pmatrix}0&1\\1&0\end{pmatrix} + \delta\begin{pmatrix}0&1\\0&1\end{pmatrix}\right)\\ &= (0+1) + \left(\delta\begin{pmatrix}0&1\\1&0\end{pmatrix} + 0\right),\end{aligned}$$

this implies $\delta\begin{pmatrix}0&1\\1&0\end{pmatrix} = -1$.

Let $A = \begin{pmatrix}a&b\\c&d\end{pmatrix} \in M_{2\times 2}(F)$. Now consider

$$\begin{aligned}\delta(A) &= \delta\begin{pmatrix}a&b\\c&d\end{pmatrix}\\ &= a\delta\begin{pmatrix}1&0\\c&d\end{pmatrix} + b\delta\begin{pmatrix}0&1\\c&d\end{pmatrix}\\ &= a\left(c\delta\begin{pmatrix}1&0\\1&0\end{pmatrix} + d\delta\begin{pmatrix}1&0\\0&1\end{pmatrix}\right) + b\left(c\delta\begin{pmatrix}0&1\\1&0\end{pmatrix} + d\delta\begin{pmatrix}0&1\\0&1\end{pmatrix}\right)\\ &= ad - bc\\ &= \det(A).\end{aligned}$$

Hence we complete the proof. ■

4.2 Determinants of Order n

Exercise 4.2.1

(a) **Answer.** False.

Solution. $\det(kI_n) = k^n \det(I_n) \neq k \det(I_n)$ where I_n is the $n \times n$ identity matrix, and k is a scalar. ■

(b) **Answer.** True.

Solution. It follows from Theorem 4.4. ■

(c) **Answer.** True.

Solution. It follows from the Corollary of Theorem 4.4. ■

(d) **Answer.** True.

Solution. It follows from Theorem 4.5. ■

(e) **Answer.** False.

Solution. $\det \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = 2$ and $\det(I) = 1$; however, they do not equal. ■

(f) **Answer.** False.

Solution. $\det \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} = 1$ and $k \det(I) = k$; however, $\det \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \neq k \det(I)$. ■

(g) **Answer.** False.

Solution. $\text{rank}(I_2) = 2$, but $\det(I_2) = 1$. ■

(h) **Answer.** True.

Solution. It follows from Exercise 4.2.23. ■

4.3 Properties of Determinants

Exercise 4.3.1

(a) **Answer.** False.

Solution. Consider an elementary matrix E of type 2 where $E = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. However, $\det(E) = 2 \neq \pm 1$. ■

(b) **Answer.** True.

Solution. It follows from Theorem 4.7. ■

(c) **Answer.** False.

Solution. It follows from the Corollary of Theorem 4.7. ■

(d) **Answer.** True.

Solution. Since $\text{rank}(M) = n$ if and only if M is invertible, then if follows from the Corollary of Theorem 4.7. ■

(e) **Answer.** False.

Solution. $\det(A^t) = \det(A)$. This follows from Theorem 4.8. ■

(f) **Answer.** True.

Solution. It follows from Theorem 4.4 and 4.8. ■

(g) **Answer.** False.

Solution. If the determinant equals to 0, then it will fail for using Cramer's Rule. ■

(h) **Answer.** False.

Solution. The statement holds if M_k is the $n \times n$ matrix obtained from A by replacing **column** k of A by b^t. ■

Exercise 4.3.15

Proof. Since A is similar to B, then there exists $Q \in M_{n\times n}(F)$ such that $A = Q^{-1}BQ$. It follows that

$$\begin{aligned}\det(A) &= \det(Q^{-1}BQ) = \det(Q^{-1})\det(B)\det(Q) = (\det(Q^{-1})\det(Q))\det(B) \\ &= \det(Q^{-1}Q)\det(B) = \det(I)\det(B) = \det(IB) = \det(B).\end{aligned}$$

■

Exercise 4.3.21

Proof. For $n = 2$, we know $M = \begin{pmatrix} A & B \\ O_1 & C \end{pmatrix}$ where $A, B, C, O_1 \in M_{1\times 1}(F)$. Then $\det(M) = \det(A)\det(C)$ by definition. Suppose for $n-1$, the statement holds; then for n, we know the cofactor expansion of M along the first column gives

$$\det(M) = \sum_{i=1}^{n}(-1)^{i+1}M_{i1}\det(\tilde{M}_{i1}).$$

Assume A is a $k \times k$ matrix where $k < n$. Since $M_{i1} = 0$ for $i > k$. The formula leads to

$$\det(M) = \sum_{i=1}^{k}(-1)^{i+1}M_{i1}\det(\tilde{M}_{i1}).$$

Suppose $\tilde{M}_{i1} = \begin{pmatrix} \tilde{A}_{i1} & \tilde{B} \\ O_{n-k-1} & C \end{pmatrix}$ where $\tilde{B}$ is obtained from deleting the i-th row of B. Since $\tilde{M}_{i1}$ is a $(n-1)\times(n-1)$ matrix, by supposition of induction, we have $\det(\tilde{M}_{i1}) = \det(\tilde{A}_{i1})\det(C)$. Hence

$$\det(M) = \sum_{i=1}^{k}(-1)^{i+1}M_{i1}\det(\tilde{A}_{i1})\det(C).$$

Notice that $M_{i1} = A_{i1}$ for each $i = 1, 2, \cdots, k$, then it follows that

$$\det(M) = \left(\sum_{i=1}^{k}(-1)^{i+1}A_{i1}\det(\tilde{A}_{i1})\right)\det(C) = \det(A)\det(C).$$

This means the statement also holds for n.

By induction, we conclude the statement holds for $n \geq 2$. ■

4.4 Summary - Important Facts about Determinants

4.5 A Characterization of the Determinant

Chapter 5

Diagonalization

5.1 Eigenvalues and Eigenvectors

Exercise 5.1.1

(a) **Answer.** False.

Solution. I_2 has duplicate eigenvalues 1. ∎

(b) **Answer.** True.

Solution. Since a real matrix A has one eigenvector v, by Theorem 5.4, we know $(A - \lambda I)(v) = 0$ for $v \neq 0$. It follows that $(A - \lambda I)(tv) = t(A - \lambda I)(v) = t \cdot 0 = 0$ for $t \in \mathbb{R}$. This means tv is also an eigenvector. Since $\mathbb{R}$ is an infinite set, then we have infinite eigenvectors.

Notice that this statement will fail if the field is finite. ∎

(c) **Answer.** True.

Solution. Consider the square matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Since it has no eigenvalues, and hence has no eigenvectors. ∎

(d) **Answer.** False.

Solution. The zero matrix $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ has duplicate eigenvalues 0. ∎

(e) **Answer.** False.

Solution. The vectors $(1,0),(2,0)$ are eigenvectors of I_2. However they are linearly dependent. ■

(f) **Answer.** False.

Solution. Consider I_2 which has two eigenvalues $1,1$; however $1+1=2$, and 2 is not an eigenvalue of I_2 obviously. ■

(g) **Answer.** False.

Solution. Put a linear operator T on $P(\mathbb{R})$. Then T has an eigenvalue 1. ■

(h) **Answer.** True.

Solution. Put $A = Q^{-1}DQ$ where D is a diagonal matrix and $P,Q \in M_{n\times n}(F)$. Let α be a basis and β be the standard basis for F_n. Since Q is invertible, then $Q = [I]_\beta^\alpha$. Since A is diagonalizable, by Theorem 5.1, we know β can be consisted of the eigenvectors of A. ■

(i) **Answer.** True.

Solution. Suppose A is similar to B, then there exists Q which is invertible such that $A = Q^{-1}BQ$. If $Av = \lambda v$ for any eigenvalues λ corresponding to some eigenvectors v, then we have

$$\begin{aligned} & Av = \lambda v \\ \Longrightarrow\ & (Q^{-1}BQ)v = \lambda v \\ \Longrightarrow\ & (BQ)v = Q\lambda v \\ \Longrightarrow\ & B(Qv) = \lambda(Qv). \end{aligned}$$

This means A, B have the same eigenvalues. ■

(j) **Answer.** False.

Solution. Suppose A is similar to B, then there exists Q which is invertible such that $A = Q^{-1}BQ$. Put $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$ and $Q = Q^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. It follows that $(1,0)$ is an eigenvector of A, but so is not B. ■

(k) **Answer.** False.

Solution. Put $T = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}$. Then $(1,1)$ and $(1,-3)$ are eigenvectors of T; however, $(1,1) + (1,-3) = (2,-2)$ is not an eigenvector of T. ■

Exercise 5.1.3

(a) **Answer.** (i) $\lambda_1 = 4, \lambda_2 = -1$

(ii) $v_1 = t\begin{pmatrix}2\\3\end{pmatrix}, t \neq 0; v_2 = t\begin{pmatrix}1\\-1\end{pmatrix}, t \neq 0.$

(iii) $\begin{pmatrix}2 & 1\\3 & -1\end{pmatrix}.$

(iv) $Q = \begin{pmatrix}2 & 1\\3 & -1\end{pmatrix}, D = \begin{pmatrix}4 & 0\\0 & -1\end{pmatrix}.$

Solution. We want to know all eigenvalues of A, we need to solve $\det(A - \lambda I) = 0$. This implies

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

Hence $\lambda_1 = 4, \lambda_2 = -1$.

Then we want to know all eigenvectors v_1 corresponding to λ_1. Let

$$B_1 = A - \lambda_1 I = \begin{pmatrix}-3 & 2\\3 & -2\end{pmatrix}.$$

Since $x \neq 0$ and $x \in N(L_{B_1})$, we have $B_1 x = 0$. Then we obtain $x \in \left\{t\begin{pmatrix}2\\3\end{pmatrix} : t \in \mathbb{R}\right\}$. Hence

$$v_1 = t\begin{pmatrix}2\\3\end{pmatrix}, t \neq 0.$$

It turn to know eigenvectors v_2 corresponding to λ_2. Let

$$B_2 = A - \lambda_2 I = \begin{pmatrix}2 & 2\\3 & 3\end{pmatrix}.$$

Since $x \neq 0$ and $x \in N(L_{B_2})$, we have $B_2 x = 0$. Then we obtain $x \in \left\{t\begin{pmatrix}1\\-1\end{pmatrix} : t \in \mathbb{R}\right\}$. Hence

$$v_2 = t\begin{pmatrix}1\\-1\end{pmatrix}, t \neq 0.$$

They follow that $\left\{\begin{pmatrix}2\\3\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\}$ is a basis for $\mathbb{R}^2$ consisting of eigenvectors of A. Hence A is diagonalizable. We just pick $Q = \begin{pmatrix}2 & 1\\3 & -1\end{pmatrix}$, then we have

$$Q^{-1}AQ = \frac{1}{5}\begin{pmatrix}1 & 1\\3 & -2\end{pmatrix}\begin{pmatrix}1 & 2\\3 & 2\end{pmatrix}\begin{pmatrix}2 & 1\\3 & -1\end{pmatrix} = \begin{pmatrix}4 & 0\\0 & -1\end{pmatrix} =: D.$$

Finally, we complete the computation. ■

Exercise 5.1.5

Proof. Following the hypothesis of Theorem 5.4, we have

$$(T - \lambda I)v = 0 \iff Tv - \lambda Iv = 0 \iff Tv = \lambda v.$$

We complete the proof of Theorem 5.4. ∎

Exercise 5.1.8

(a) **Proof.** ($\Longrightarrow$) Suppose to contrary that $\lambda = 0$ is an eigenvalue of T. Since T is invertible, then $\det(T) \neq 0$. Consider $\det(T - \lambda I) = 0$. It follows that $\det(T - 0I = \det(T) = 0$ which contradicts the supposition. Hence 0 is not an eigenvalue of T.

($\Longleftarrow$) Suppose T is **not** invertible, then $\det(T) = 0$. Then if $\lambda = 0$ is an eigenvalue of T, it follows that $\det(T - \lambda I) = \det(T) = 0$ holds. We have proven the contraposition. ∎

(b) **Proof.** Let v be some eigenvectors corresponding to λ. Notice that T is invertible. Consider

$$\begin{aligned} & Tv = \lambda v \\ \iff \quad & v = T^{-1}\lambda v \\ \iff & \lambda^{-1}v = T^{-1}v. \end{aligned}$$

This means λ^{-1} is an eigenvalue of T^{-1}. ∎

(c) **Proof.** We restate part(a) as a matrix A is singular if and only if 0 is not an eigenvalue of L_A. Here is the proof:

$$A \text{ is singular } \iff L_A \text{ is not invertible } \iff 0 \text{ is an eigenvalue of } L_A.$$

We restate part(b) as if A is invertible, then λ is an eigenvalue of L_A if and only if λ^{-1} is an eigenvalue of $(L_A)^{-1}$. Here is the proof:

Since A is invertible, then L_A is invertible. from part (b), we obtain the statement holds. ∎

Exercise 5.1.9

Proof. Let M be $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$. Then we consider $\det(M - \lambda I) = \prod_{i=1}^{n}(a_{ii} - \lambda) = 0$. Hence $\lambda = a_{ii}$ for each $i = 1, 2, \cdots, n$. ∎

Exercise 5.1.11

(a) **Proof.** Since A is similar to λI, there exists Q which is invertible such that $A = Q^{-1}\lambda I Q$. It follows that

$$A = \lambda Q^{-1} I Q = \lambda Q^{-1} Q = \lambda I.$$

■

(b) **Proof.** Let $\beta = \{v_1, v_2, \cdots, v_n\}$ is a basis for A and v_i is an eigenvector of A for each i. By Theorem 5.1, we know $[A]_\beta$ is a diagonal matrix and hence is a upper triangular matrix. From Exercise 5.1.9, and by hypothesis that all eigenvalues are the same called λ, then $[A]_\beta = \lambda I$. This implies A is similar to λI, from part (a), we conclude $A = \lambda I$. ■

(c) **Proof.** Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Since A has the same eigenvalues 1, from part (b), if A is diagonalizable, then A must be a scalar matrix; however, A is not a scalar matrix and hence is not diagonalizable. ■

Exercise 5.1.12

(a) **Proof.** Suppose A is similar to B, there exists Q which is invertible such that $A = Q^{-1}BQ$. Then consider

$$\begin{aligned}
\det(A - \lambda I) &= \det(Q^{-1}BQ - \lambda I) \\
&= \det(Q^{-1}BQ - Q^{-1}\lambda I Q) \\
&= \det(Q^{-1}(B - \lambda I)Q) \\
&= \det(Q^{-1})\det(B - \lambda I)\det(Q) \\
&= \det(B - \lambda I).
\end{aligned}$$

Hence A, B have the same characteristic polynomial. ■

(b) **Proof.** Let α and β be arbitrary different bases for V, and T be a linear operator on V. Since $[T]_\alpha$ is similar to $[T]_\beta$, from part (a), we know they have the same characteristic polynomial. Because α, β were arbitrary, this leads to that the statement holds. ■

Exercise 5.1.14

Proof. Since $\lambda I = (\lambda I)^t$ and hence $\det(A - \lambda I) = \det((A - \lambda I)^t) = \det(A^t - \lambda I)$. ■

Exercise 5.1.20

Proof. By hypothesis, we know $f(0) = a_0$ trivially. Notice that $f(t) = \det(A - \lambda I)$ and hence $f(0) = \det(A - 0I) = \det(A)$. They follow that $f(0) = a_0 = \det(A)$. Moreover,

$$A \text{ is invertible} \iff \det(A) \neq 0 \iff a_0 \neq 0.$$

■

5.2 Diagonalizability

Exercise 5.2.1

(a) **Answer**. False.

Solution. Since $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is diagonalizable. However, it has two distinct eigenvalues 0 and 1. ■

(b) **Answer**. False.

Solution. Consider the matrix $\begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}$. We know $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ are eigenvectors corresponding to the eigenvalue 3; however, they are linearly independent. ■

(c) **Answer**. False.

Solution. The zero vector can not be an eigenvector. ■

(d) **Answer**. True.

Solution. Let $x \in E_{\lambda_1} \cap E_{\lambda_2}$, then $Tx = \lambda_1 x = \lambda_2 x$. It follows that $x = 0$ since $\lambda_1 \neq \lambda_2$. ■

(e) **Answer**. True.

Solution. It follows from Theorem 5.1. ■

(f) **Answer**. False.

Solution. It doesn't satisfy Theorem 5.9. Besides we also need the characteristic polynomial of T can split. ■

(g) **Answer**. True.

Solution. Since the vector space is nonzero, then its characteristic polynomial must have degree more than 0. Hence if it can diagonalize, then it must have an eigenvalue. ■

(h) **Answer**. True.

Solution. Since $W_i \cap \sum_{i \neq j} W_j = \{0\}$ by definition. We also know $W_j \in \sum_k W_k$. It follows that $W_i \cap W_j = \{0\}$ where $i \neq j$. ■

(i) **Answer**. False.

Solution. Consider $W_1 = \text{span}\{(0,1)\}$, $W_2 = \text{span}\{1,1\}$, $W_3 = \text{span}\{(1,0)\}$, then we have $W_i \cap W_j = \{0\}$ for $i \neq j$. However $W_1 \cap (W_2 + W_3) \neq \{0\}$. ■

Exercise 5.2.3

(f) **Answer.** T is diagonalizable. $\beta = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right\}$.

Solution. Let α be the standard basis for V, then $[T]_\alpha = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Solve $\det(T - \lambda I) = 0$, we have $(\lambda - 1)^3(\lambda + 1) = 0$. Pick $\lambda_1 = 1$, $\lambda_2 = -1$ with multiplicities 3, 1, respectively. We obtain

$$N_{\lambda_1} = \left\{ \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \in \mathbb{R}^4 : \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Hence $\beta_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$ is a basis for N_{λ_1}. On the other hand, we also obtain

$$N_{\lambda_2} = \left\{ \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \in \mathbb{R}^4 : \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Hence $\beta_2 = \left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ is a basis for N_{λ_2}.

We have shown that $\dim(N_{\lambda_1}) = 3$, $\dim(N_{\lambda_2}) = 1$. It is consistent with the multiplicities 3, 1. By Theorem 5.9, we know T is diagonalizable. Moreover, we can construct a basis $\beta = \beta_1 \cup \beta_2$ so that

$$[I]_\beta^\alpha = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \text{ and } [I]_\alpha^\beta = \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

It follows that

$$[T]_\beta = [I]_\alpha^\beta [T]_\alpha [I]_\beta^\alpha = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

■

Exercise 5.2.7

Answer. $A^n = \frac{1}{3}\begin{pmatrix} 5^n + 2(-1)^n & 2 \cdot 5^n - 2(-1)^n \\ 5^n - (-1)^n & 2 \cdot 5^n + (-1)^n \end{pmatrix}$.

Solution. Solve $\det(A - \lambda I) = 0$, we have $(\lambda - 5)(\lambda + 1) = 0$. Pick $\lambda_1 = 5$, $\lambda_2 = -1$. For $\lambda_1 = 5$, we know $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ is a basis for N_{λ_1}. For $\lambda_2 = -1$, we know $\left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$ is a basis for N_{λ_2}. Hence we pick $Q = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$. It follows that

$$\begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} = Q^{-1} A Q.$$

This implies

$$\begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}^n = (Q^{-1} A Q)^n = Q^{-1} A^n Q.$$

So we conclude

$$\begin{aligned} A^n &= Q \begin{pmatrix} 5^n & 0 \\ 0 & (-1)^n \end{pmatrix} Q^{-1} \\ &= \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5^n & 0 \\ 0 & (-1)^n \end{pmatrix} \left(\frac{1}{3} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \right) \\ &= \frac{1}{3} \begin{pmatrix} 5^n + 2(-1)^n & 2 \cdot 5^n - 2(-1)^n \\ 5^n - (-1)^n & 2 \cdot 5^n + (-1)^n \end{pmatrix}. \end{aligned}$$

■

Exercise 5.2.12

(a) **Proof.** We denoted the eigenspace of T corresponding to λ by $E_{\lambda,T}$.

Let $x \in E_{\lambda,T}$, then $T(x) = \lambda x$. Since T is invertible, we have $x = \lambda T^{-1} x$. Moreover, by Exercise 5.1.8(a), 0 is not an eigenvalue of T. Hence we can divide λ on both sides to obtain $\lambda^{-1} x = T^{-1}(x)$. This means $x \in E_{\lambda^{-1},T^{-1}}$. Conversely, a similar argument establishes $x \in E_{\lambda^{-1},T^{-1}}$ implies $x \in E_{\lambda,T}$.

Since x was arbitrary, we conclude $E_{\lambda,T} = E_{\lambda^{-1},T^{-1}}$. ■

(b) **Proof.** Since T is diagonalizable, then 0 is not an eigenvalue of T. Hence there exists Q which is invertible such that $D = Q^{-1} T Q$ where D is a diagonal matrix; also, D^{-1} exists. It follows that $T = QDQ^{-1}$. This implies that $T^{-1} = (QDQ^{-1})^{-1} = QD^{-1}Q^{-1}$. This means T^{-1} is diagonalizable. ■

Exercise 5.2.13

(a) **Proof.** Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Then $\lambda = 2$ is an eigenvalue of A. We know $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ is a basis for E_λ.

Consdier A^t, $\lambda' = 2$ is also an eigenvalue of A. We know $\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is a basis for $E_{\lambda'}$.

However $\text{span}\{(1,1)\} \neq \text{span}\{(0,1)\}$ since $\text{span}\{(1,1)\} = \mathbb{R}^2$, $\text{span}\{(0,1)\} = \mathbb{R}^1$. ■

(b) **Proof.** Pick an arbitrary eigenvalue λ. By the Dimension Theorem, we have

$$\dim(E_\lambda) = \dim(N(A - \lambda I)) = n - \text{rank}(A - \lambda I);$$
$$\dim(E'_\lambda) = \dim(N(A^t - \lambda I)) = n - \text{rank}(A^t - \lambda I).$$

Since $(A - \lambda I)^t = A^t - \lambda I$, then

$$\text{rank}(A - \lambda I) = \text{rank}((A - \lambda I)^t) = \text{rank}(A^t - \lambda I).$$

Hence we know for a particular eigenvalue λ,

$$\dim(E_\lambda) = \dim(E'_\lambda).$$

Since λ was arbitrary, the statement follows. ■

(c) **Proof.** Since A is diagonalizable, then $\det(A - \lambda I)$ can split and $\dim(E_\lambda)$ equals to the multiplicity of λ for each λ.

Because $\det(A - \lambda I) = \det((A - \lambda I)^t) = \det(A^t - \lambda I)$, A^t can split, too. From part (b), we know $\dim(E_\lambda) = \dim(E'_\lambda)$; hence $\dim(E'_\lambda)$ equals to the multiplicity of λ for each λ.

They follows A^t is diagonalizable. ■

Exercise 5.2.18

(a) **Proof.** Notice that if D_T, D_U are diagonal matrices, then $D_T D_U = D_U D_T$.

Suppose T, U are $n \times n$ matrices. By hypothesis and definition, we know there exists a $n \times n$ matrix Q which is invertible such that

$$D_T = Q^{-1} T Q;$$
$$D_U = Q^{-1} U Q$$

where D_T, D_U are diagonal matrices. It follows that

$$D_T D_U = (Q^{-1} T Q)(Q^{-1} U Q) = Q^{-1} T U Q;$$
$$D_U D_T = (Q^{-1} U Q)(Q^{-1} T Q) = Q^{-1} U T Q.$$

Observe $D_T D_U = D_U D_T$, then $Q^{-1} T U Q = Q^{-1} U T Q$. Hence $TU = UT$. ■

5.3 Matrix Limits and Markov Chains

5.4 Invariant Subspace and the Cayley-Hamilton Theorem

Exercise 5.4.1

(a) **Answer.** False.

Solution. Since $\{0\}$ is subspace of T trivially, and $\{0\}$ is T-invariant subspace. ■

(b) **Answer.** True.

Solution. It follows from Theorem 5.21. ■

(c) **Answer.** False.

Solution. Define $T : \mathbb{R} \to \mathbb{R}$ by $T(x) = x$. Pick $v = \{1\}, w = \{-1\}$. This implies $W = W' = \mathbb{R}$. However, $v \neq w$. ■

(d) **Answer.** False.

Solution. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x, y) = x$. Pick $v = (1, 1)$, then T-cyclic subspace generated by v is $\mathbb{R}^2$. However, one generated by $T(v)$ is y-axis which in $\mathbb{R}$. Clearly, $\mathbb{R}^2 \neq \mathbb{R}$. ■

(e) **Answer.** True.

Solution. It follows from Cayley-Hamilton Theorem (Theorem 5.23). That is, $g(t)$ is the characteristic polynomial of T. ■

(f) **Answer.** True.

Solution. It follows from Exercise 5.1.29(a). ■

(g) **Answer.** True.

Solution. It follow from Theorem 5.25. ■

Exercise 5.4.3

(a) **Proof.** Since $T(0) \in \{0\}$, then $\{0\}$ is a T-invariant subspace of V.

Since $T(V) \subseteq V$ by definition, then V is also a T-invariant subspace of V. ■

(b) **Proof.** Let $x \in N(T)$, then $T(x) = 0 \in N(T)$. Since x was arbitrary, $T(N(T)) \subseteq N(T)$. Hence $N(T)$ is T-invariant.

Let $x \in R(T)$, then $T(x) \in R(T)$ by definition. Since x was arbitrary, $T(R(T)) \subseteq R(T)$. Hence $R(T)$ is also T-invariant. ■

(c) **Proof.** Let $x \in E_\lambda$, this means $x \in N(T - \lambda I)$. It follows that

$$\begin{aligned} (T - \lambda I)x &= 0 \\ \implies \quad Tx &= \lambda I x \\ \implies \quad Tx &= \lambda x. \end{aligned}$$

Then $T(x) \in E_\lambda$ and hence E_λ is T-invariant. ■

Exercise 5.4.6

(a) **Answer.** $\{(1,0,0,0),(1,0,1,1),(1,-1,2,2)\}$.

Solution. Compute

$$\begin{aligned} e_1 &= (1,0,0,0) \\ T(e_1) &= (1,0,1,1) \\ T^2(e_1) &= T(1,0,1,1) = (1,-1,2,2) \\ T^3(e_1) &= T(1,-1,2,2) = (0,-3,3,3). \end{aligned}$$

We know $\beta = \{e_1, T(e_1), T^2(e_1)\}$ is linearly independent. However $T^3(e_1) = -3T(e_1) + 3T^2(e_1)$. This means $T^3(e_1) \in \operatorname{span}(\beta)$. Hence

$$\{(1,0,0,0),(1,0,1,1),(1,-1,2,2)\}$$

is a basis for T-cyclic subspace generated by e_1. ■

Exercise 5.4.18

(a) **Proof.** Consider

$$f(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 = \det(A - tI). \tag{5.1}$$

We know $f(0) = a_0 = \det(A)$. Hence

$$A \text{ is invertible} \iff \det(A) \neq 0 \iff a_0 \neq 0.$$

■

(b) **Proof.** Consider formula (5.1), we have

$$f(A) = (-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 = \det(O).$$

Since $\det(O) = 0$, then $(-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 = 0$. Notice that A is invertible and hence from part (a), $a_0 \neq 0$. Then divide $-a_0$ on both sides and rearrange it to obtain

$$\frac{-1}{a_0}\left[(-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A\right] = 1.$$

Now multiply A^{-1} on both sides to conclude

$$A^{-1} = \frac{-1}{a_0}\left[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n\right].$$

■

(c) **Proof.** Compute $\det(A - tI) = -t^3 + 2t^2 + t - 2$ and $A^2 = \begin{pmatrix} 1 & 6 & 6 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix}$. From part (b), we know

$$\begin{aligned} A^{-1} &= \frac{-1}{a_0}\left[(-1)A^2 + 2A + I\right] \\ &= \frac{-1}{-2}\left[(-1)\begin{pmatrix} 1 & 6 & 6 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix} + 2\begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] \\ &= \frac{1}{2}\begin{pmatrix} 2 & -2 & -4 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned}$$

■

Exercise 5.4.19

Proof. For $k = 1$, $\det(A - tI) = -a_0 - t = (-1)^1(a_0 + t)$, it follows the statement holds.
Suppose the statement holds for k. Then for $k + 1$, we have

$$\begin{aligned} \det(A - tI) &= \det\begin{pmatrix} -t & 0 & \cdots & 0 & -a_0 \\ 1 & -t & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -a_{k-1} \\ 0 & 0 & \cdots & 1 & -a_k \end{pmatrix} \\ &= (-t)\det\begin{pmatrix} -t & \cdots & 0 & -a_1 \\ 1 & \cdots & 0 & -a_2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & -a_k \end{pmatrix} + (-1)^{k+2}(-a_0)\det\begin{pmatrix} 1 & -t & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_k \end{pmatrix} \\ &= (-t)\left[(-1)^k(a_1 + a_2 t + \cdots + a_k t^k + t^{k+1})\right] + (-1)^{k+2}(-a_0)(1) \\ &= (-1)^{k+1}(a_1 t + \cdots + a_k t^k + t^{k+1}) + (-1)^{k+1}(a_0) \\ &= (-1)^{k+1}(a_0 + a_1 t + \cdots + a_k t^k + t^{k+1}). \end{aligned}$$

Hence the statement also holds for $k + 1$.
By induction, we conclude the statement holds for $k \in \mathbb{N}$. ■

Exercise 5.4.42

Answer. $\begin{cases} (-1)^n t^{n-1}(t-n) & \text{if } n > 1 \\ t-1 & \text{if } n = 1 \end{cases}$.

Proof. For $n = 1$, the characteristic polynomial is $t-1$ trivially.

For $n > 1$, since all columns of A are the same as $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{n\times 1}$, we know $A = \text{span}\left\{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{n\times 1}\right\}$. So rank$(A) = 1$. By the Dimension Theorem, we know nullity$(A) = n-1$. That is, A is not invertible, by Exercise 5.1.8(a), we know 0 is an eigenvalue of A. Also notice that

$$A\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} n \\ n \\ \vdots \\ n \end{pmatrix} = n\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

It follows that n is also an eigenvalue of A. If λ has multiplicity m_λ, by Theorem 5.7, we have $1 \leq \dim(E_\lambda) \leq m_\lambda$. Hence $m_0 \geq n-1$, $m_n \geq 1$. Moreover, the degree of characteristic polynomial is n, then $m_0 + m_n \leq n$. It follows that $m_0 = n-1$, $m_n = 1$. So the characteristic polynomial of A is $(-1)^n t^{n-1}(t-n)$. ■

Chapter 6

Inner Product Spaces

6.1 Inner Products and Norms

Exercise 6.1.1

(a) **Answer.** False.

Solution. It follows from definition immediately. ■

(b) **Answer.** True.

Solution. Since the field F means $\mathbb{R}$ or $\mathbb{C}$. ■

(c) **Answer.** False.

Solution. It is conjugate linear in the second component by Theorem 6.1. ■

(d) **Answer.** False.

Solution. An inner product can be defined by our own. e.g., $\langle u, v \rangle := uv$. ■

(e) **Answer.** False.

Solution. It follows from Theorem 6.2. ■

(f) **Answer.** False.

Solution. It follows from definition immediately. ■

(g) **Answer.** False.

Solution. Let $x = (1,1)$, $y = (1,0)$, $z = (0,1)$ in $\mathbb{R}^2$. Then $\langle x,y\rangle = \langle x,z\rangle = 1$. However, $y \neq z$. ■

(h) **Answer.** True.

Solution. Suppose to contrary that $y \neq 0$. By Theorem 6.1, $\langle x,y\rangle = 0$ implies $x = 0$ or $x = y$. Since $\langle x,y\rangle = 0$ for all x, this contradicts. Hence $y = 0$. ■

Exercise 6.1.4

(a) **Proof.** Let $A, B, C \in V$ and $c \in F$. We prove it is compatible with definitions.

- $\langle A+B, C\rangle = \langle A,C\rangle + \langle B,C\rangle$.

$$\langle A+B, C\rangle = \operatorname{tr}(C^*(A+B)) = \operatorname{tr}(C^*A) + \operatorname{tr}(C^*B) = \langle A,C\rangle + \langle B,C\rangle.$$

- $\langle cA, B\rangle = c\langle A,B\rangle$.

$$\langle cA, B\rangle = \operatorname{tr}(B^*(cA)) = c\operatorname{tr}(B^*A) = c\langle A,B\rangle.$$

- $\overline{\langle A,B\rangle} = \langle B,A\rangle$.

$$\overline{\langle A,B\rangle} = \operatorname{tr}(\overline{B^*A}) = \operatorname{tr}(A^*B) = \langle B,A\rangle.$$

- $\langle A,A\rangle > 0 \iff A \neq O$.

$$\begin{aligned}\langle A,A\rangle &= tr(A^*A)\\ &= \sum_{i=1}^{n}\sum_{k=1}^{n}(A^*)_{ik}A_{ki}\\ &= \sum_{i=1}^{n}\sum_{k=1}^{n}\overline{A_{ki}}A_{ki}\\ &= \sum_{i=1}^{n}\sum_{k=1}^{n}|A_{ki}|^2.\end{aligned}$$

Hence $A \neq 0 \iff \exists i,k : A_{ki} \neq 0 \iff \langle A,A\rangle > 0$.

■

Exercise 6.1.8

(a) **Proof.** Consider $(1,1) \in \mathbb{R}^2$ where $(1,1)$ is nonzero. However, $\langle (1,1),(1,1)\rangle = 1 - 1 = 0$. This violates the definition. ■

Exercise 6.1.9

(a) **Proof.** Let $\beta = \{z_1, z_2, \cdots, z_n\}$, then we know $\langle x, z_i \rangle = 0$ for each i. Since $z \in \beta$, we have $z = \sum_{i=1}^{n} a_i z_i$ for $a_i \in F$. Consider

$$\langle x, x \rangle = \langle x, \sum_{i=1}^{n} a_i z_i \rangle = \sum_{i=1}^{n} \overline{a_i} \langle x, z_i \rangle = 0.$$

Hence $x = 0$. ■

(b) **Proof.** Notice that $\langle x, z \rangle = \langle y, z \rangle$ implies $\langle x - y, z \rangle = 0$. Let $\beta = \{z_1, z_2, \cdots, z_n\}$, then we know $\langle x - y, z_i \rangle = 0$ for each i. From part (a), we conclude $x - y = 0$ implies $x = y$. ■

Exercise 6.1.11

Proof. Consider for all $x, y \in V$,

$$\begin{aligned}
\|x+y\|^2 + \|x-y\|^2 &= \left(\sqrt{\langle x+y, x+y \rangle}\right)^2 + \left(\sqrt{\langle x-y, x-y \rangle}\right)^2 \\
&= \langle x+y, x+y \rangle + \langle x-y, x-y \rangle \\
&= (\langle x, x+y \rangle + \langle y, x+y \rangle) + (\langle x, x-y \rangle - \langle y, x-y \rangle) \\
&= (\langle x, x+y \rangle + \langle x, x-y \rangle) + (\langle y, x+y \rangle + \langle y, x-y \rangle) \\
&= \langle x, 2x \rangle + \langle y, 2y \rangle \\
&= 2\langle x, x \rangle + 2\langle y, y \rangle \\
&= 2\|x\|^2 + 2\|y\|^2.
\end{aligned}$$

This means that the sum of square of the four edges is the sum of square of the two diagonals in the same parallelogram. ■

Exercise 6.1.12

Proof. Since $\{v_1, v_2, \cdots, v_k\}$ is orthogonal, then $\langle v_i, v_j \rangle = 0$ for all $i \neq j$. Follow the hypothesis, then we fix i to consider

$$\left\langle v_i, \sum_{j=1}^{k} a_j v_j \right\rangle = \sum_{j=1}^{k} \overline{a_j} \langle v_i, v_j \rangle = \overline{a_i} \langle v_i, v_i \rangle.$$

It follows that

$$\begin{aligned}\left\|\sum_{i=1}^{k} a_i v_i\right\|^2 &= \left\langle \sum_{i=1}^{k} a_i v_i, \sum_{j=1}^{k} a_j v_j \right\rangle \\ &= \sum_{i=1}^{k} a_i \left\langle v_i, \sum_{j=1}^{k} a_j v_j \right\rangle \\ &= \sum_{i=1}^{k} a_i \sum_{j=1}^{k} \overline{a_j} \left\langle v_i, v_j \right\rangle \\ &= \sum_{i=1}^{k} a_i \overline{a_i} \left\langle v_i, v_i \right\rangle \\ &= \sum_{i=1}^{k} |a_i|^2 \|v_i\|^2.\end{aligned}$$

We obtain the desired result. ■

6.2 The Gram-Schmidt Orthogonalization Process and Orthogonal Complements

Exercise 6.2.1

(a) **Answer.** False.

Solution. Sets of vectors must be linearly independent. ■

(b) **Answer.** True.

Solution. It follows from Theorem 6.5. ■

(c) **Answer.** True.

Solution. Let V be an arbitrary inner product space and W be a subspace of V. Then let $x, y \in W^\perp$ and c be a scalar. For all $w \in W$,

$$\langle x+y, w\rangle = \langle x, w\rangle + \langle y, w\rangle = 0 + 0 = 0;$$

also

$$\langle cx, w\rangle = c\langle x, w\rangle = c \cdot 0 = 0.$$

And $\langle 0, w\rangle = 0$ trivially.
Hence $W^\perp$ is a subspace of V. ■

(d) **Answer.** False.

Solution. The basis must be orthonormal. ■

(e) **Answer.** True.

Solution. It follows from definition immediately. ■

(f) **Answer.** False.

Solution. An orthogonal set must be nonzero. ■

(g) **Answer.** True.

Solution. It follows from the Corollary 2 of Theorem 6.3. ■

Exercise 6.2.2

(b) **Solution.** First, we try to find an orthogonal basis for span(S). Pick $w_1 = (1,1,1)$, $w_2 = (0,1,1)$, $w_3 = (0,0,1)$. Then

$$\begin{aligned} v_1 &= w_1 = (1,1,1); \\ v_2 &= w_2 - \frac{\langle w_2, v_1\rangle}{\|v_1\|^2}v_1 = \frac{1}{3}(-2,1,1); \\ v_3 &= w_3 - \frac{\langle w_3, v_1\rangle}{\|v_1\|^2}v_1 - \frac{\langle w_3, v_2\rangle}{\|v_2\|^2}v_2 = \frac{1}{2}(0,-1,1). \end{aligned}$$

Hence we let $\{v_1, v_2, v_3\}$ is an orthogonal basis for span(S), then we normalize them.

$$\begin{aligned} u_1 &= \frac{1}{\|v_1\|}v_1 = \frac{\sqrt{3}}{3}(1,1,1); \\ u_2 &= \frac{1}{\|v_2\|}v_2 = \frac{\sqrt{6}}{6}(-2,1,1); \\ u_3 &= \frac{1}{\|v_3\|}v_3 = \frac{\sqrt{2}}{2}(0,-1,1). \end{aligned}$$

Hence $\beta = \{u_1, u_2, u_3\}$ is an orthonormal basis for span(S).

Secondly, we find the Fourier coefficients relative to β.

$$\begin{aligned} c_1 &= \langle x, u_1\rangle = \frac{2\sqrt{3}}{3}; \\ c_2 &= \langle x, u_2\rangle = -\frac{\sqrt{6}}{6}; \\ c_3 &= \langle x, u_3\rangle = \frac{\sqrt{2}}{2}. \end{aligned}$$

Thirdly, we verify Theorem 6.5.

$$x = \langle x, u_1 \rangle u_1 + \langle x, u_2 \rangle u_2 + \langle x, u_3 \rangle u_3.$$

■

Exercise 6.2.6

Proof. Let $u \in W$, $y \in W^\perp$, then by Theorem 6.6, $x = u + y \in V$ is unique. Since $x \notin W$, then $y \neq 0$. Notice that $\langle u, y \rangle = 0$. Consider

$$\langle x, y \rangle = \langle u + y, y \rangle = \langle u, y \rangle + \langle y, y \rangle = \langle y, y \rangle > 0.$$

Hence $\langle x, y \rangle \neq 0$ follows. ■

Exercise 6.2.13

(a) **Proof.** Let $x \in S^\perp$. Then $\langle x, y \rangle$ for all $y \in S$. Since $S_0 \subseteq S$, $\langle x, y \rangle$ for all $y \in S_0$. Hence $x \in S_0^\perp$. Because x was arbitrary, it follows that $S^\perp \subseteq S_0^\perp$. ■

(b) **Proof.** Let $x \in S$, then for any $y \in S^\perp$, we have $\langle x, y \rangle = 0$. This means $x \in (S^\perp)^\perp$. Since x was arbitrary, $S \subseteq (S^\perp)^\perp$. Because $\mathrm{span}(S)$ is the smallest subspace containing S, we conclude $\mathrm{span}(S) \subseteq (S^\perp)^\perp$. ■

(c) **Proof.** To prove $W = (W^\perp)^\perp$, we need to prove $W \subseteq (W^\perp)^\perp$ and $(W^\perp)^\perp \subseteq W$.

From part (b), we know $W \subseteq (W^\perp)^\perp$. On the other hand, we suppose to contrary that $(W^\perp)^\perp \nsubseteq W$. Let $x \in (W^\perp)^\perp$, then $x \notin W$. By Exercise 6.2.6, there exists $y \in V$ such that $y \in W^\perp$ implies $\langle x, y \rangle \neq 0$. However this contradicts $x \in (W^\perp)^\perp$. Hence $(W^\perp)^\perp \subseteq W$. ■

(d) **Proof.** To prove $V = W \oplus W^\perp$, it suffice to prove $V = W + W^\perp$ and $W \cap W^\perp = \{0\}$.

By Theorem 6.6, for any $x \in V$, we can find unique $y \in W$ and $z \in W^\perp$ such that $x = y + z$. This means $V = W + W^\perp$. Let $w \in W \cap W^\perp$, then $\langle w, w \rangle = 0$. By definition, we know $w = 0$. So $W \cap W^\perp = \{0\}$. ■

Exercise 6.2.19

(b) **Answer.** $\frac{1}{14}(29, 17, 40)$.

Solution. We pick $\{(-3, 1, 0), (2, 0, 1)\}$ as a basis for W. Pick $w_1 = (-3, 1, 0)$ and $w_2 = (2, 0, 1)$. Then use Gram-Schmidt process to compute

$$v_1 = w_1 = (-3, 1, 0);$$
$$v_2 = w_2 - \frac{\langle w_2, v_1 \rangle}{\|v_1\|^2} v_1 = \frac{1}{5}(1, 3, 5).$$

And then normalize them to obtain

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{10}}(-3,1,0);$$
$$u_2 = \frac{v_2}{\|v_2\|} = \frac{1}{\sqrt{35}}(1,3,5).$$

So we can find the projection p of u on W is

$$p = \langle u, u_1 \rangle u_1 + \langle u, u_2 \rangle u_2 = \frac{1}{14}(29,17,40).$$

■

6.3 The Adjoint of a Linear Operator

Exercise 6.3.1

(a) **Answer.** True.

Solution. It follows from Theorem 6.9. ■

(b) **Answer.** False.

Solution. The form is mapping V to F. This violates the definition of linear operator. ■

(c) **Answer.** False.

Solution. It violate the hypothesis of Theorem 6.10. β must be an orthonormal basis. ■

(d) **Answer.** True.

Solution. It follows from Theorem 6.9. ■

(e) **Answer.** False.

Solution. Under the hypothesis, it should be $(aT + bU)^* = \bar{a}T^* + \bar{b}U^*$. ■

(f) **Answer.** True.

Solution. It follows from the Corollary of Theorem 6.10. ■

(g) **Answer.** True.

Solution. It follows from Theorem 6.11. ■

Exercise 6.3.2

(b) **Answer.** $y = (1, -2)$.

Solution. Since $g(z_1, z_2) = z_1 - 2z_2 = \langle (z_1, z_2), (1, -2) \rangle$, then we pick $y = (1, -2)$. ■

Exercise 6.3.7

Solution. We pick $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. Then we know $T^* = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. Observe $N(T) = \text{span}\{(0, 1)\}$, $N(T^*) = \text{span}\{(1, -1)\}$. Notice that $(0, 1) \in N(T)$; however $(0, 1) \notin N(T^*)$. Hence $N(T) \neq N(T^*)$. ■

Exercise 6.3.8

Proof. Under the hypothesis, we consider

$$T^*(T^{-1})^* = (T^{-1}T)^* = I^* = I;$$
$$(T^{-1})^*T^* = (TT^{-1})^* = I^* = I.$$

Hence $(T^*)^{-1} = (T^{-1})^*$. ■

Exercise 6.3.12

(a) **Proof.** It suffice to prove $R(T^*)^\perp \subseteq N(T)$ and $N(T) \subseteq R(T^*)^\perp$.

First, we claim $N(T) \subseteq R(T^*)^\perp$. Let $x \in N(T)$, then $Tx = 0$. For any $y \in V$, we consider $0 = \langle 0, y \rangle = \langle Tx, y \rangle = \langle x, T^*y \rangle$. Hence $x \in R(T^*)^\perp$. Because x was arbitrary, the claim holds.

Secondly, we claim $R(T^*)^\perp \subseteq N(T)$. Let $x \in R(T^*)^\perp$, then for any $y \in V$, we consider $0 = \langle x, T^*y \rangle = \langle Tx, y \rangle$. This implies $Tx = 0$ and hence $x \in N(T)$. Because x was arbitrary, the claim holds.

Finally, we conclude $R(T^*)^\perp = N(T)$. ■

(b) **Proof.** Since V is finite-dimensional, then both $R(T^*)$ and $N(T)$ are finite-dimensional. From part (a), $(R(T^*)^\perp)^\perp = N(T)^\perp$ follows. Hence we obtain $R(T^*) = N(T)^\perp$ as desired. ■

6.4 Normal and Self-Adjoint Operators

Exercise 6.4.1

(a) **Answer.** True.

Solution. Since $T = T^*$, then $TT^* = T^2 = T^*T$. This means T is normal. ■

(b) **Answer.** False.

Solution. Consider $T = \begin{pmatrix} 3 & 4 \\ 1 & 0 \end{pmatrix}$, this implies T have eigenvectors $\begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. On the other hand, $T^* = \begin{pmatrix} 3 & 1 \\ 4 & 0 \end{pmatrix}$. This implies T^* have eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \end{pmatrix}$. Hence their eigenvectors are different. ■

(c) **Answer.** False.

Solution. β should be an orthonormal basis. It follows from the definition of normal operator and Theorem 6.10. A counter-example could see Exercise 6.4.3. ■

(d) **Answer.** True.

Solution. It follows from Theorem 6.10. ■

(e) **Answer.** True.

Solution. It follows from the Lemma in page 373. ■

(f) **Answer.** True.

Solution. Since $I^* = I$ and $O^* = O$. ■

(g) **Answer.** False.

Solution. Consider $T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, then T is a normal operator. However, T is not diagonalizable. ■

(h) **Answer.** True.

Solution. If the inner product space is over $\mathbb{R}$, then it follows from Theorem 6.17; if it is over $\mathbb{C}$, it follows from Theorem 6.16. ■

Exercise 6.4.2

(a) **Solution.** Observe $T = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix} = T^*$. Hence T is self-adjoint and normal. Solve $\det(T - \lambda I) = \lambda^2 - 7\lambda + 6 = 0$ to get $\lambda_1 = 6, \lambda_2 = -1$. Then we have eigenvector $(1, -2)$ corresponding to λ_1, and eigenvector $(2, 1)$ corresponding to λ_2. And normalize them to obtain

$$\left\{ \frac{1}{\sqrt{3}}(2,1), \frac{1}{\sqrt{3}}(1,-2) \right\}$$

is an orthonormal basis for V. ■

Exercise 6.4.3

Answer. $T(x, y) = (x, 2y)$ and $\{(1,1), (0,1)\}$ is a basis for $\mathbb{R}^2$.

Solution. Pick $T : \mathbb{R}^2 \to \mathbb{R}^2$ where $T(x, y) = (x, 2y)$. Then $T = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Observe $T = T^*$ and hence T is normal.

Now pick $\beta = \{(1,1), (0,1)\}$, then

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

Consequetly, we verify whether it is normal or not.

$$[T]_\beta [T]^*_\beta = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix};$$
$$[T]^*_\beta [T]_\beta = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}.$$

Hence $[T]_\beta$ is not normal. ■

Exercise 6.4.6

(a) **Proof.** Consider

$$T_1^* = \left[\frac{1}{2}(T + T^*) \right]^* = \frac{1}{2}(T + T^*)^* = \frac{1}{2}(T^* + T) = \frac{1}{2}(T + T^*) = T_1.$$

This means T_1 is self-adjoint.

Consider

$$T_2^* = \left[\frac{1}{2}\mathrm{i}(T - T^*) \right]^* = \frac{1}{-2\,\mathrm{i}}(T - T^*)^* = -\frac{1}{2\,\mathrm{i}}(T^* - T) = \frac{1}{2\,\mathrm{i}}(T - T^*) = T_2.$$

This means T_2 is self-adjoint.

Consider

$$T_1 + \mathrm{i}\,T_2 = \frac{1}{2}(T + T^*) + \mathrm{i} \cdot \frac{1}{2\mathrm{i}}(T - T^*) = \frac{1}{2}T + \frac{1}{2}T^* + \frac{1}{2}T - \frac{1}{2}T^* = T.$$

We obtain the desired result. ■

(b) **Proof.** Suppose $T = U_1 + \mathrm{i}\,U_2$ with $U_1 = U_1^*$ and $U_2 = U_2^*$.
Consider

$$T_1 = \frac{1}{2}(T + T^*) = \frac{1}{2}(U_1 + \mathrm{i}\,U_2 + U_1^* - \mathrm{i}\,U_2^*) = \frac{1}{2}(U_1 + \mathrm{i}\,U_2^* + U_1 - \mathrm{i}\,U_2^*) = U_1.$$

On the other hand,

$$T_2 = \frac{1}{2\,\mathrm{i}}(T - T^*) = \frac{1}{2\,\mathrm{i}}(U_1 + \mathrm{i}\,U_2 - U_1^* + \mathrm{i}\,U_2^*) = \frac{1}{2\,\mathrm{i}}(U_1^* + \mathrm{i}\,U_2 - U_1^* + \mathrm{i}\,U_2) = U_2.$$

As a result, we obtain the desired results. ■

(c) **Proof.** ($\Longrightarrow$) Since T is normal, then $TT^* = T^*T$. Consider

$$T_1T_2 = \frac{1}{2}(T + T^*) \cdot \frac{1}{2\,\mathrm{i}}(T - T^*) = \frac{1}{4\,\mathrm{i}}(T^2 + T^*T - TT^* - T) = \frac{1}{4\,\mathrm{i}}(T^2 - T);$$
$$T_2T_1 = \frac{1}{2\,\mathrm{i}}(T - T^*) \cdot \frac{1}{2}(T + T^*) = \frac{1}{4\,\mathrm{i}}(T^2 - T^*T + TT^* - T) = \frac{1}{4\,\mathrm{i}}(T^2 - T).$$

It follows that $T_1T_2 = T_2T_1$.

($\Longleftarrow$) Compute T_1T_2 and T_2T_1, since $T_1T_2 = T_2T_1$, we have

$$\frac{1}{4\,\mathrm{i}}(T^2 + T^*T - TT^* - T) = \frac{1}{4\,\mathrm{i}}(T^2 - T^*T + TT^* - T).$$

This implies

$$2T^*T = 2TT^*.$$

Then $T^*T = TT^*$ follows and hence T is normal. ■

Exercise 6.4.9

Proof. Since T is normal, if $\|Tx\| = 0$ for arbitrary $x \in V$, then $\|T^*x\| = 0$. This converse also holds. Hence $N(T) = N(T^*)$. By Exercise 6.3.12, we know

$$R(T^*) = N(T)^\perp = N(T^*)^\perp = \left(R(T)^\perp\right)^\perp = R(T).$$

■

6.7 The Singular Value Decomposition and the Pseudoinverse

Exercise 6.7.1

(a) **Answer.** False.

Solution. Consider $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ where eigenvalues are 2. However, the singular value of A is 4. ■

(b) **Answer.** False.

Solution. See part(a). ■

(c) **Answer.** True.

Solution. If σ is the singular value of A, then σ^2 is the eigenvalue of A^*A. We know $(cA)^*(cA) = (c^*c)A^*A = c^2A^*A$ and hence $|c|^2\sigma^2$ is the eigenvalue of cA. It follows that $|c|\sigma$ is the singular value of cA. ■

(d) **Answer.** True.

Solution. It follows from definition. ■

(e) **Answer.** False.

Solution. See part(a). ■

(f) **Answer.** False.

Solution. It follows from Theorem 6.30. ■

(g) **Answer.** True.

Solution. It follows from definition. ■

Exercise 6.7.2

(a) **Answer.** $\left\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\}$ is a basis for V.

$\left\{\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-1\end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix}2\\-1\\-1\end{pmatrix}\right\}$ is a basis for W.

$\sigma_1 = \sqrt{3}$, $\sigma_2 = \sqrt{2}$.

Solution. Pick the standard basis $\beta = \{e_1, e_2\}$. We have $[T]_\beta = \begin{pmatrix}1 & 0\\1 & 1\\1 & -1\end{pmatrix}$. Then we calculate $T^*T = \begin{pmatrix}1 & 1 & 1\\1 & 2 & 0\\1 & 0 & 2\end{pmatrix}$. Hence we know the eigenvalues are

$$\lambda_1 = 3,\ \lambda_2 = 2,\ \lambda_3 = 0.$$

It follows that

$$v_1' = \begin{pmatrix}1\\1\\1\end{pmatrix},\ v_2' = \begin{pmatrix}0\\1\\-1\end{pmatrix},\ v_3' = \begin{pmatrix}2\\-1\\-1\end{pmatrix}$$

are corresponding eigenvectors. Normalize them to obtain

$$v_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix},\ v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-1\end{pmatrix},\ v_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}2\\-1\\-1\end{pmatrix}.$$

Finally, we know the singular values are

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3},\ \sigma_2 = \sqrt{\lambda_2} = \sqrt{2}.$$

■

Exercise 6.7.3

(a) **Answer.** $A = \left[\frac{1}{\sqrt{6}}\begin{pmatrix}\sqrt{2} & \sqrt{3} & 1\\ \sqrt{2} & -\sqrt{3} & 1\\ -\sqrt{2} & 0 & 2\end{pmatrix}\right]\begin{pmatrix}\sqrt{6} & 0\\ 0 & 0\\ 0 & 0\end{pmatrix}\left[\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\right]^*.$

Solution. Let $A = \begin{pmatrix}1 & 1\\ 1 & 1\\ -1 & -1\end{pmatrix}$. Calculate $A^*A = \begin{pmatrix}3 & 3\\ 3 & 3\end{pmatrix}$. Hence we know the eigenvalues are

$$\lambda_1 = 6,\ \lambda_2 = 0.$$

It follows that

$$v_1' = \begin{pmatrix}1\\1\end{pmatrix},\ v_2' = \begin{pmatrix}1\\-1\end{pmatrix}$$

are corresponding eigenvectors. Normalize them to obtain

$$v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix},\ v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}.$$

We know the singular value is

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{6}.$$

This means

$$u_1 = \frac{1}{\sigma_1}Av_1 = \frac{1}{\sqrt{6}}\begin{pmatrix}1 & 1\\ 1 & 1\\ -1 & -1\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\-1\end{pmatrix}.$$

Put $u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}$ and $u_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\2\end{pmatrix}$ such that $\{u_1, u_2, u_3\}$ is an orthonormal basis. We obtain the singular value decomposition $A = U\Sigma V^*$ where

$$U = \frac{1}{\sqrt{6}}\begin{pmatrix}\sqrt{2} & \sqrt{3} & 1\\ \sqrt{2} & -\sqrt{3} & 1\\ -\sqrt{2} & 0 & 2\end{pmatrix},\ \Sigma = \begin{pmatrix}\sqrt{6} & 0\\ 0 & 0\\ 0 & 0\end{pmatrix},\ V = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}.$$

■

(b) **Answer.** $A = \left[\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\right]\begin{pmatrix}\sqrt{2} & 0 & 0\\ 0 & \sqrt{2} & 0\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}^*$.

Solution. Let $A = \begin{pmatrix}1 & 0 & 1\\ 1 & 0 & -1\end{pmatrix}$. Calculate $A^*A = \begin{pmatrix}2 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 2\end{pmatrix}$. Hence we know the eigenvalues are

$$\lambda_1 == \lambda_2 = 2,\ \lambda_3 = 0.$$

It follows that

$$v_1 = \begin{pmatrix}0\\0\\1\end{pmatrix},\ v_2 = \begin{pmatrix}0\\0\\1\end{pmatrix},\ v_3 = \begin{pmatrix}0\\1\\0\end{pmatrix}$$

are corresponding eigenvectors. We know the singular values are

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{2},\ \sigma_2 = \sqrt{\lambda_2} = \sqrt{2}$$

This means

$$u_1 = \frac{1}{\sigma_1}Av_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0 & 1\\ 1 & 0 & -1\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}.$$

Put $u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$ such that $\{u_1, u_2\}$ is an orthonormal basis. We obtain the singular value decomposition $A = U\Sigma V^*$ where

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix},\ \Sigma = \begin{pmatrix}\sqrt{2} & 0 & 0\\ 0 & \sqrt{2} & 0\end{pmatrix},\ V = \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}.$$

■

Exercise 6.7.4

(a) **Answer.** $A = \left[\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\right]\left[\frac{1}{2}\begin{pmatrix}3\sqrt{2} & -\sqrt{2}\\ -\sqrt{2} & 3\sqrt{2}\end{pmatrix}\right]$

Solution. Let $A = \begin{pmatrix}1 & 1\\ 2 & -2\end{pmatrix}$. Calculate $A^*A = \begin{pmatrix}5 & -3\\ -3 & 5\end{pmatrix}$. Hence we know the eigenvalues are

$$\lambda_1 = 8,\ \lambda_2 = 2.$$

It follows that

$$v_1' = \begin{pmatrix}1\\-1\end{pmatrix},\ v_2' = \begin{pmatrix}1\\1\end{pmatrix}$$

are corresponding eigenvectors. Normalize them to obtain

$$v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix},\ v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}.$$

We know the singular values are

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{8},\ \sigma_2 = \sqrt{\lambda_2} = \sqrt{2}.$$

This means

$$u_1 = \frac{1}{\sigma_1}Av_1 = \frac{1}{\sqrt{8}}\begin{pmatrix}1 & 1\\2 & -2\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix};$$
$$u_2 = \frac{1}{\sigma_2}Av_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\2 & -2\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$$

such that $\{u_1, u_2\}$ is an orthonormal basis. So we have

$$U = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix},\ \Sigma = \begin{pmatrix}\sqrt{8} & 0\\0 & \sqrt{2}\end{pmatrix},\ V = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}.$$

Hence we get the polar decomposition $A = WP$ where

$$W = UV^* = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\left[\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}\right] = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix};$$
$$P = V\Sigma V^* = \left[\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}\right]\begin{pmatrix}\sqrt{8} & 0\\0 & \sqrt{2}\end{pmatrix}\left[\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}\right] = \frac{1}{2}\begin{pmatrix}3\sqrt{2} & -\sqrt{2}\\-\sqrt{2} & 3\sqrt{2}\end{pmatrix}.$$

■

Exercise 6.7.6

Follow Exercise 6.7.3 and use Theorem 6.29 that $A^\dagger = V\Sigma^\dagger U^*$.

(a) **Answer.** $A^\dagger = \left[\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\right]\begin{pmatrix}\sqrt{6} & 0 & 0\\0 & 0 & 0\end{pmatrix}\left[\frac{1}{\sqrt{6}}\begin{pmatrix}\sqrt{2} & \sqrt{3} & 1\\\sqrt{2} & -\sqrt{3} & 1\\-\sqrt{2} & 0 & 2\end{pmatrix}\right]^*.$

(b) **Answer.** $A^\dagger = \begin{pmatrix}1 & 0 & 0\\0 & 0 & 1\\0 & 1 & 0\end{pmatrix}\begin{pmatrix}\sqrt{2} & 0\\0 & \sqrt{2}\\0 & 0\end{pmatrix}\left[\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\right]^*.$

6.8 Bilinear and Quadratic Forms

Exercise 6.8.3

Let $H_1, H_2 : V \times V \to \mathbb{R}$ be bilinear forms and a, c be scalars.

(a) **Proof.** Check

$$\begin{aligned}(H_1+H_2)(ax_1+x_2,y) &= H_1(ax_1+x_2,y)+H_2(ax_1+x_2,y)\\ &= aH_1(x_1,y)+H_1(x_2,y)+aH_2(x_1,y)+H_2(x_2,y)\\ &= a(H_1+H_2)(x_1,y)+(H_1+H_2)(x_2,y).\end{aligned}$$

A similar argument shows that

$$(H_1+H_2)(x,ay_1+y_2) = a(H_1+H_2)(x,y_1)+(H_1+H_2)(x,y_2).$$

■

(b) **Proof.** Check

$$\begin{aligned}cH(ax_1+x_2,y) &= c\big(aH(x_1,y)+H(x_2,y)\big)\\ &= acH(x_1,y)+cH(x_2,y);\\ cH(x,ay_1+y_2) &= c\big(aH(x,y_1)+H(x,y_2)\big)\\ &= acH(x,y_1)+cH(x,y_2).\end{aligned}$$

■

(c) **Proof.** Put $H_0(x,y)=0$. Check properties from $(VS\ 1)$ to $(VS\ 8)$. It is so tedious that we skip verification here. ■

Chapter 7

Canonical Forms

7.1 The Jordan Canonical Form I

Exercise 7.1.1

(a) **Answer**. True.

Solution. It follows from definition. ■

(b) **Answer**. False.

Solution. Let x be a generalized eigenvector of T, then by definition, there exists p is the smallest integer such that $(T - \lambda I)^p(x) = 0$. However $y = (T - \lambda I)^{p-1} \neq 0$ is an eigenvector corresponding to λ. This means λ is an eigenvalue. ■

(c) **Answer**. False.

Solution. It also need to satisfy the characteristic polynomial can split. It follows from the Corollary 1 of Theorem 7.7. ■

(d) **Answer**. True.

Solution. It follows from the Corollary of Theorem 7.6. ■

(e) **Answer**. False.

Solution. Consider $A = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & 3 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 3 \end{pmatrix}$. Corresponding to the eigenvalue 2, we have two cycle of generalized eigenvectors of A. ■

(f) **Answer.** False.

Solution. Consider $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Observe the eigenvalue is 1. Then $K_1 = \mathbb{R}^2$. We pick $\beta = \{(1,-1),(1,1)\}$ to know that $[T]_\beta = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$ is not a Jordan canonical form. ■

(g) **Answer.** True.

Solution. We just pick β be the standard basis. Then $[L_{\mathcal{J}}]_\beta = \mathcal{J}$. ■

(h) **Answer.** True.

Solution. It follows from Theorem 7.2(b). ■

Exercise 7.1.3

(a) **Answer.** For $\lambda = 2$, $\{2, -2x, x^2\}$ is a basis. $\mathcal{J} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

Solution. Let β be the standard basis for $P_2(\mathbb{R})$, then $[T]_\beta = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 2 \end{pmatrix}$. We observe the eigenvalue is 2 with multiplicity 3. Now we should pick $v = (0,0,1)$ such that $v \neq 0$, $(T-2I)(v) \neq 0$, $(T-2I)^2(v) \neq 0$, and $(T-2I)^3(v) = 0$. That is, $\{2, -2x, x^2\}$ is a basis for T. By doing so, we obtain $Q = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $Q^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$. We get the Jordan canonical form of T is

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

■

Exercise 7.1.13

Proof. Let $V = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \cdots \oplus K_{\lambda_k}$ by Theorem 7.8. We pick a basis β_i which is a union of cycles of generalized eigenvectors for $T_{K_{\lambda_i}}$ for each $i = 1, 2, \cdots, k$. This means $\mathcal{J}_i = [T_{K_{\lambda_i}}]_{\beta_i}$ is a Jordan canonical form for $T_{K_{\lambda_i}}$. Put $\beta = \cup_{i=1}^k \beta_i$. Then β is a basis for V where β is a union of cycles of generalized eigenvectors in T. Hence $\mathcal{J} = [T]_\beta$ is a Jordan canonical form for T. Moreover, by Theorem 5.25, $\mathcal{J} = \mathcal{J}_1 \oplus \mathcal{J}_2 \oplus \cdots \oplus \mathcal{J}_k$ as promised. ■

7.2 The Jordan Canonical Form II

Exercise 7.2.1

(a) **Answer.** True.

Solution. Since the Jordan form of a diagonal matrix is itself. And the Jordan canonical form is unique. ■

(b) **Answer.** True.

Solution. Since T is similar to $[T]_\beta$ for arbitrary β, then by Theorem 7.11, we know they have the same Jordan canonical form $\mathcal{J}$. ■

(c) **Answer.** False.

Solution. Consider two matrices $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, they have the same characteristic polynomial $(\lambda - 1)^2$; however, they are not similar. ■

(d) **Answer.** True.

Solution. It follows from Theorem 7.11. ■

(e) **Answer.** True.

Solution. If we use a Jordan canonical basis β to represent a matrix T, then we get the Jordan canonical form $\mathcal{J}$. Hence T is similar to $\mathcal{J} = [T]_\beta$. ■

(f) **Answer.** False.

Solution. Consider a particular case for which $t = 1$ and $n = 2$. Then we consider two matrices $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. They have different Jordan canonical forms. ■

(g) **Answer.** False.

Solution. Consider the identity linear operator T, then $T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

We can pick $\{(0,1),(1,0)\}$ or $\{(1,1),(-1,1)\}$ to be a Jordan canonical basis for T. ■

(h) **Answer.** True.

Solution. It follows from the Corollary of Theorem 7.10. ■

Exercise 7.2.4

(a) **Answer.** $\mathcal{J} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. $Q = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}$.

Solution. Firstly, we need to find eigenvalues. Solve $\det(A - \lambda I) = -(\lambda - 1)(\lambda - 2)^2 = 0$. Then $\lambda_1 = 1, \lambda_2 = 2$.

Secondly, we need to find a basis. For $\lambda_1 = 1$, since it is with multiplicity 1, then we just pick eigenvector $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ to be a basis. For $\lambda_2 = 2$, we know $T - 2I = \begin{pmatrix} -5 & 3 & -2 \\ -7 & 4 & -3 \\ 1 & -1 & 0 \end{pmatrix}$ and $(T - 2I)^2 = \begin{pmatrix} 2 & -1 & 1 \\ 4 & -2 & 2 \\ 2 & -1 & 0 \end{pmatrix}$. Since λ_2 is with multiplicity 2 and

$$\begin{aligned} &\dim(A) - \operatorname{rank}(T - 2I) = 3 - 2 = 1; \\ &\operatorname{rank}(T - 2I) - \operatorname{rank}((T - 2I)^2) = 2 - 1 = 1, \end{aligned}$$

we should find a vector v such that $v \in N\big((T - 2I)^2\big)$ and $v \notin N\big(T - 2I\big)$. Here we pick $v = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, then $(T - 2I)(v) = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

Hence this means $Q = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}$. $\mathcal{J} = Q^{-1}AQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ follows. ■

Exercise 7.2.5

Notice that a Jordan canonical basis β is not unique.

(a) **Answer.** $\beta = \{e^t, te^t, t^2e^t, e^{2t}\}$. $\mathcal{J} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

Solution. Observe

$$\begin{aligned} T(e^t) &= e^t; \\ T(te^t) &= e^t + te^t; \\ T(t^2e^t) &= 2te^t + t^2e^t; \\ T(e^{2t}) &= 2e^{2t}. \end{aligned}$$

Pick $\beta = \{e^t, te^t, t^2e^t, e^{2t}\}$ which is a basis for V. Then we know

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

Solve $\det(T - \lambda I) = (\lambda - 1)^3(\lambda - 2) = 0$, we obtain $\lambda_1 = 1$, $\lambda_2 = 2$ with multiplicities 3,1, respectively.
For $\lambda_1 = 1$, we know

$$T - I = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix};\ (T-I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix};(T-I)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This implies

$$\begin{aligned} &\dim(T) - \mathrm{rank}(T - I) = 4 - 3 = 1; \\ &\mathrm{rank}(T - I) - \mathrm{rank}\left((T-I)^2\right) = 3 - 1 = 2; \\ &\mathrm{rank}\left((T-I)^2\right) - \mathrm{rank}\left((T-I)^3\right) = 2 - 1 = 1. \end{aligned}$$

We should find a vector v such that $v \in N\big((T-I)^3\big)$, $v \notin N\big((T-I)^k\big)$ for $k = 1, 2$. Hence we pick $v = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. By doing so, $(T - I)(v) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ and $(T-I)^2(v) = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}$ follow. For $\lambda_2 = 2$, since it has multiplicity 1, we just consider $T - 2I = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, then pick an eigenvector $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

Hence we have $Q = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ such that $\mathcal{J} = Q^{-1}[T]_\beta Q$, then the Jordan canonical form $\mathcal{J} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ follows.

■

Exercise 7.2.6

Proof. Follow the hint. At first, since $\det(A - \lambda I) = \det(A^t - \lambda I)$, this means A and A^t have the same characteristic polynomial, and hence have the same eigenvalues. We pick arbitrary eigenvalue λ of both A and A^t, fix it and for any $r \in \mathbb{N}$, we have

$$\operatorname{rank}\left(A - \lambda I\right)^r = \operatorname{rank}\left(\left(A - \lambda I\right)^t\right)^r = \operatorname{rank}\left(A^t - \lambda I\right)^r.$$

Hence they have the same dot diagram by Theorem 7.9. Then they have the same Jordan canonical form since the Corollary of Theorem 7.10. Moreover, by Theorem 7.11, they are similar. ■

7.3 The Minimal Polynomial

Exercise 7.3.1

(a) **Answer.** False.

Solution. Consider $T = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix}$. We know the characteristic polynomial $f(t) = -(t-2)^2(t-3)$. However, we can pick $p(t) = (t-2)(t-3)$ such that $p(T) = T_0$ where the degree of $p(t)$ is less than $f(t)$. ■

(b) **Answer.** True.

Solution. It follows from Theorem 7.12(b). ■

(c) **Answer.** False.

Solution. It should be that the minimal polynomial divides the characteristic polynomial. This follows from Theorem 7.12(a). ■

(d) **Answer.** False.

Solution. Consider the identity linear operator T which is diagonalizable. Its characteristic polynomial $(t-1)^2$; however its minimal polynomial is $t-1$. Hence they are different. ■

(e) **Answer.** True.

Solution. By the Corollary of Theorem 7.14, we suppose

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

where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are different eigenvalues of T. Then there exists $m_1, m_2, \cdots, m_k$ where $1 \le m_i \le n_i$ for each i such that

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

Notice that $n_1 + n_2 + \cdots + n_k = n$. Hence $f(t)$ divides $\left[p(t)\right]^n$. ■

(f) **Answer.** False.

Solution. The counter-example could see part (d). ■

(g) **Answer.** False.

Solution. Consider $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. We know its minimal polynomial $(t-1)^2$ could split. However, T is not diagonalizable. ■

(h) **Answer.** True.

Solution. It follows from Theorem 7.15. ■

(i) **Answer.** True.

Solution. We know the degree of the minimal polynomial must be greater than or equal to n by the Corollary of Theorem 7.14. By Cayley-Hamilton Theorem, its degree must be less than or equal to n. Hence the degree of the minimal polynomial must be equal to n. ■

Exercise 7.3.2

In spite of checking all possibilities, we can find Jordan canonical form in an alternative way by Exercise 7.3.13.

(a) **Answer.** $p(t) = (t-1)(t-3)$.

Solution. Find the Jordan canonical form $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$. Hence we know the minimal polynomial is $(t-1)(t-3)$. ■

(b) **Answer.** $p(t) = (t-1)^2$.

Solution. Find the Jordan canonical form $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Hence we know the minimal polynomial is $(t-1)^2$. ■

(c) **Answer.** $p(t) = (t-1)^2(t-2)$.

Solution. Find the Jordan canonical form $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. Hence we know the minimal polynomial is $(t-1)^2(t-2)$. ■

(d) **Answer.** $p(t) = (t-2)^2$.

Solution. Let the matrix be A. We need to find the characteristic polynomial at first. Solve $\det(A - tI) = 0$ to obtain $-(t-2)^3$.

Since the minimal polynomial divides the characteristic polynomial, we just check all possibilities one by one. Check $(t-2)$, then $A - 2I = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ -1 & 0 & -1 \end{pmatrix} \neq O$. Then check $(t-2)^2$, it follows that $(A-2I)^2 = O$. Hence the minimal polynomial is $(t-2)^2$ ■

Exercise 7.3.3

Find $[T]_\beta$ and use a similar argument from previous exercise. Here we use the standard basis β.

(a) **Answer.** $p(t) = (t - \sqrt{2})(t + \sqrt{2})$.

Solution. Calculate $[T]_\beta = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Then the Jordan canonical form is $\begin{pmatrix} -\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$. Hence the minimal polynomial is $(t - \sqrt{2})(t + \sqrt{2})$. ■

(b) **Answer.** $p(t) = (t-2)^3$.

Solution. Calculate $[T]_\beta = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$. Then the Jordan canonical form is $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Hence the minimal polynomial is $(t-2)^3$. ■

(c) **Answer.** $p(t) = (t-2)^2$.

Solution. Calculate $[T]_\beta = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ which is the Jordan canonical form coincidentally. Hence we know the minimal polynomial is $(t-2)^2$. ■

(d) **Answer.** $p(t) = (t-1)(t+1)$.

Solution. Following the hint, we have $(T - I)(T + I) = O$. Put $f(t) = (t-1)(t+1)$ with $f(T) = O$. By the Cayley-Hamilton Theorem, f is the characteristic polynomial. Since $T - I \neq O$ and $T + I \neq O$, then the minimal polynomial must be $(t-1)(t+1)$. ■

Exercise 7.3.4

Answer. Exercise 7.3.2(a), 7.3.3(a)(d).

Solution. It follows from Theorem 7.16. ■

Exercise 7.3.5

Solution. Define $g(t) = t^3 - 2t^2 + t = t(t-1)^2$. Since T is diagonalizable, we know the minimal polynomial $p(t)$ could split into distinct linear factors, and $\deg\big(p(t)\big) \leq 2$ There three posibilities:

1. $p(t) = t$.

   In this case, $T = T_0$ trivially.

2. $p(t) = t - 1$.

   In this case, $T = I$ trivially.

3. $p(t) = t(t-1)$.

   In this case, T is similar to $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

■