

Contents

1	Vector Spaces	3
1.1	Vector Spaces	3
1.2	Subspaces	14
1.3	Linear Combinations	27
1.4	Linear Dependence and Linear Independence	31
1.5	Bases and Dimension	40
2	Linear Transformations and Matrices	49
2.1	Linear Transformations, Null spaces, and Ranges	49
2.2	The Matrix Representation of a Linear Transformation	69
2.3	Composition of Linear Transformations and Matrix Multiplication	75
2.4	Invertibility and Isomorphisms	86
2.5	The Change of Coordinate Matrix	95
3	Matrix Operations and Elementary Matrices	101
3.1	The Rank of a Matrix and Matrix Inverses	101
4	Determinants	109
4.1	Determinants of Order 2	109
4.2	Determinants of Order n	112
5	Diagonalization	117
5.1	Eigenvalues and Eigenvectors	117
5.2	Diagonalization	127
6	Inner Product Spaces	135
6.1	Inner Products and Norms	135
6.2	Gram-Schmidt Orthogonalization Process	156
6.3	The Adjoint of a Linear Operator	163

Linear Algebra Exercises

Lance Remigio

April 23, 2024

Chapter 1

Vector Spaces

1.1 Vector Spaces

Exercise 1.2.1

Label the following statements as true or false.

- (a) Every vector space contains a zero vector.
- (b) A vector space may have more than one zero vector.
- (c) In any vector space, $ax = bx$ implies that $a = b$.
- (d) In any vector space, $ax = ay$ implies that $x = y$.

Exercise 1.2.7

Let $S = \{0, 1\}$ and $F = \mathbb{R}$. In $\mathcal{F}(S, \mathbb{R})$, show that $f = g$ and where $f(t) = 2t+1$, $g(t) = 1+4t-2t^2$, and $h(t) = 5t + 1$.

Proof. To show that $f = g$, we have to show that for each $s \in S$ that $f(s) = g(s)$. Since $S = \{0, 1\}$, we can just evaluate both f and g for elements in S . Note that

$$f(0) = 2(0) + 1 = 1$$

and likewise,

$$g(0) = 1 + 4(0) - 2(0)^2.$$

Hence, $f(0) = g(0)$. Now let us evaluate both functions f and g at $s = 1$. Hence, we have

$$f(1) = 2(1) + 1 = 3$$

and

$$g(1) = 1 + 4(1) - 2(1)^2 = 3.$$

Thus, we must have $f(s) = g(s)$ for all $s \in S$.

Now, we need to show that $f + g = h$. Like we did above, we have to show that this is the case for all $s \in S$. Note that

$$(f + g)(s) = f(s) + g(s).$$

Hence, we have

$$f(s) + g(s) = 2 + 6s - 2s^2.$$

Evaluating at $s = 0$, we have

$$f(0) + g(0) = 2.$$

and likewise,

$$h(0) = 5^0 + 1 = 2.$$

Hence, $(f + g)(0) = h(0)$. Now let us evaluate $f + g$ at $s = 1$

$$f(1) + g(1) = 2 + 6(1) - 2(1)^2 = 6$$

and likewise, we have

$$h(1) = 5^1 + 1 = 6.$$

Hence, we have $(f + g)(1) = h(1)$. Thus, we have that $f + g = h$ for all $s \in S$. ■

Exercise 1.2.8

In any vector space V , we have

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

for any $x, y \in V$ and any $a, b \in F$.

Proof. Observe the following set equalities:

$$(a + b)(x + y) = a(x + y) + b(x + y) \quad (\text{VS 8})$$

$$= ax + ay + bx + by. \quad (\text{VS 7})$$

Hence, we have that

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

for any $x, y \in V$ and any $a, b \in F$. ■

Exercise 1.2.10

Let V denote the set of all differentiable real-valued functions defined on the real line. Prove that V is a vector space with the operations of addition and scalar multiplication defined in Example 3.

Proof. Let V denote the set of all differentiable real-valued functions defined on the real line. We need to show that V is a vector space over \mathbb{R} with addition and scalar multiplication defined in Example 3.

(VS 1) Let $f, g \in V$. We need to show that $f + g = g + f$ for all $x \in \mathbb{R}$. Let $x \in \mathbb{R}$. Since $f(x)$ and $g(x)$ are also real numbers, we have

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x).$$

Hence, we have $f + g = g + f$.

(VS 2) Let $f, g, h \in V$. We need to show that $f + (g + h) = (f + g) + h$. Let $x \in \mathbb{R}$. Since

$f(x), g(x), h(x) \in \mathbb{R}$, we can see that

$$\begin{aligned}(f + (g + h))(x) &= f(x) + (g + h)(x) \\ &= f(x) + g(x) + h(x) \\ &= (f + g)(x) + h(x) \\ &= ((f + g) + h)(x).\end{aligned}$$

Hence, we have $f + (g + h) = (f + g) + h$.

(VS 3) Observe that $f_0(x) = 0$ for all $x \in \mathbb{R}$ is also a real-valued function that is differentiable. We need to show that $f + f_0 = f$. Hence, we have

$$(f + f_0)(x) = f(x) + f_0(x) = f(x) + 0 = f(x).$$

Hence, $f + f_0 = f$ for all $x \in \mathbb{R}$.

(VS 4) Take $c = -1$ and perform a scalar operation with a $f \in V$. Observe that $(-1 \cdot f)(x) = (-1)f(x) = -f(x)$. Denote $g = -f$. Since additive inverses exists in \mathbb{R} , we have

$$\begin{aligned}(f - g)(x) &= (f - f)(x) \\ &= f(x) - f(x) \\ &= 0 \\ &= f_0(x).\end{aligned}$$

Hence, $f - g = f_0$ for all $x \in \mathbb{R}$.

(VS 5) Let $f \in V$ and let $x \in \mathbb{R}$. We need to show that $1 \cdot f = f$. Since $f(x) \in \mathbb{R}$, we can see that multiplicative identities in \mathbb{R} are also preserved in V ; that is, we have

$$(1f)(x) = 1 \cdot f(x) = f(x).$$

Hence, we have $1f = f$ for all $x \in \mathbb{R}$.

(VS 6) Let $a, b \in \mathbb{R}$ and $f \in V$. Let $x \in \mathbb{R}$ be arbitrary. Then observe that

$$((ab)f)(x) = (ab)f(x) = a(bf(x)) = a(bf)(x).$$

Hence, we have $(ab)f = a(bf)$ for all $x \in \mathbb{R}$.

(VS 7) Let $a \in \mathbb{R}$ and let $f, g \in V$. We need to show that $a(f + g) = af + ag$. Let $x \in \mathbb{R}$. Observe that $f(x), g(x) \in \mathbb{R}$ imply

$$\begin{aligned}(a(f + g))(x) &= a(f + g)(x) \\ &= a(f(x) + g(x)) \\ &= af(x) + ag(x) \\ &= (af)(x) + (ag)(x).\end{aligned}$$

Hence, we conclude that $a(f + g) = af + ag$ for all $x \in \mathbb{R}$.

(VS 8) Let $a, b \in \mathbb{R}$ and let $f \in V$. Let $x \in \mathbb{R}$. Then we have

$$\begin{aligned} ((a+b)f)(x) &= (a+b)f(x) \\ &= af(x) + bf(x) \\ &= (af)(x) + (bf)(x). \end{aligned}$$

Hence, we have $(a+b)f = af + bf$ for all $x \in \mathbb{R}$.

Since all the properties of a vector space have been satisfied, we conclude that V is a vector space. ■

Exercise 1.2.11

Let $V = \{0\}$ consist of a single vector O and define $O + O = O$ and $cO = O$ for each scalar $c \in F$. Prove that V is a vector space over F .

Proof. Let $x, y \in O$. Since $V = \{0\}$, we know that x and y are both the 0 vector. By using the addition defined on V , we can see that

$$x + y = O + O = y + x.$$

Note that $x + y \in V$ implies that $x + y = O + O$ and likewise $y + z = O + O$ for every $x, y, z \in V$. Hence, we have that

$$x + (y + z) = O + (O + O) = (O + O) + O = (x + y) + z.$$

Since V consists of only the zero vector O , we know that (VS 3) and (VS 4) are satisfied. By the same reasoning, (VS 5) is satisfied because

$$1 \cdot O = O.$$

Let $a, b \in F$. Then (VS 6) and (VS 7) are satisfied because

$$(ab)O = 0 = a \cdot 0 = a(bO) \text{ and } a(O + O) = 0 = O + O = aO + aO \text{ respectively.}$$

Let $a, b \in F$ again. Then we have

$$\begin{aligned} (a+b)O &= 0 \\ &= O + O \\ &= aO + bO. \end{aligned}$$

Hence, (VS 8) is satisfied. ■

Exercise 1.2.12

A real-valued function f defined on the real line is called an **even function** if $f(-t) = f(t)$ for all $t \in \mathbb{R}$. Prove that the set of even functions defined on \mathbb{R} with the operations of addition and scalar multiplication defined in Example 3 is a vector space.

Proof. First, let us show that for every $t \in \mathbb{R}$ that $(f + g)(t)$ is also an even function for every pair of even functions f, g and likewise $(cf)(t)$ is an even function for every $c \in \mathbb{R}$.

Observe that

$$\begin{aligned}(f + g)(-t) &= f(-t) + g(-t) \\ &= f(t) + g(t) \\ &= (f + g)(t).\end{aligned}$$

Hence, the function $f + g$ is also even. Now observe that

$$\begin{aligned}(cf)(-t) &= cf(-t) \\ &= cf(t) \\ &= (cf)(t).\end{aligned}$$

Hence, cf is also even function.

Now we can show that the set of even functions is a vector space.

(VS 1) Let f, g be a pair of real-valued even functions. Let $t \in \mathbb{R}$ be arbitrary. We need to show that $f + g = g + f$. Since $f(t)$ and $g(t)$ are real numbers, observe that

$$(f + g)(t) = f(t) + g(t) = g(t) + f(t) = (g + f)(t).$$

Hence, we have that $f + g = g + f$.

(VS 2) Let f, g, h be even functions and let $t \in \mathbb{R}$ be arbitrary. We need to show that $f + (g + h) = (f + g) + h$. By the same reasoning we used to prove (VS 1), observe that

$$\begin{aligned}f(t) + ((g + h)(t)) &= f(t) + (g(t) + h(t)) \\ &= (f(t) + g(t)) + h(t) \\ &= ((f + g)(t)) + h(t)\end{aligned}$$

Hence, we have that $f + (g + h) = (f + g) + h$.

(VS 3) Let f be an even function. Note that the zero function is an even function. Denote the zero function as f_0 . For every $t \in \mathbb{R}$, we have $f_0(x) = 0$. We need to show that $f + f_0 = f$. Since $f_0(x)$ is just a real number, we have

$$(f + f_0)(t) = f(t) + f_0(t) = f(t) + 0 = f(t).$$

(VS 4) Let f be an even function and let f_0 be the zero function defined above. We need to show that there exists an even function g such that $f + g = f_0$. Let $t \in \mathbb{R}$ be arbitrary. Note that

$$(f + g)(t) = f(t) + g(t)$$

Since $f(t)$ and $g(t)$ are real numbers and there exists an additive identity in the real numbers, we have that $f(t) + g(t) = f_0(t)$. Hence, (VS 5) is satisfied.

(VS 5) Let $x \in \mathbb{R}$ and f an even function. Then we immediately have $(1 \cdot f)(t) = 1f(t) = f(t)$.

(VS 6) Let $a, b \in \mathbb{R}$. We need to show that $(ab)f = a(bf)$. Hence, let $t \in \mathbb{R}$. Then we have

$$(abf)(t) = a(bf)(t).$$

(VS 7) Let $a \in \mathbb{R}$ and let f, g be even functions. Let $t \in \mathbb{R}$. We need to show that $a(f + g) = af + ag$. Then we have

$$\begin{aligned} a(f + g)(t) &= a[f(t) + g(t)] \\ &= af(t) + ag(t). \end{aligned}$$

Hence, (VS 7) is satisfied.

(VS 8) Now let $a, b \in \mathbb{R}$ and let f be an even function. We need to show that $(a + b)f = af + bf$. Let $t \in \mathbb{R}$ be arbitrary. Observe that

$$\begin{aligned} (a + b)f(t) &= af(t) + bf(t) \\ &= (af)(t) + (bf)(t) \end{aligned}$$

Hence, (VS 8) is satisfied. ■

Exercise 1.2.13

Let V denote the set of ordered pairs of real numbers. If (a_1, a_2) and (b_1, b_2) are elements of V and $c \in \mathbb{R}$, define

$$(a_1, a_2) + (b_1, b_2) = (a_1 + a_2, a_2 b_2) \text{ and } c(a_1, a_2) = (ca_1, a_2).$$

Is V a vector space over \mathbb{R} with these operations? Justify your answer.

Proof. We have that V is not a vector space of \mathbb{R} . To see why, let $(2, 1), (4, 2) \in V$ where $x = (2, 1)$ and $y = (4, 2)$. We will show that (VS 1) does not hold; that is, $x + y \neq y + x$. Hence, observe that

$$(2, 1) + (4, 2) = (2 + 1, 2) = (3, 2)$$

and

$$(4, 2) + (2, 1) = (4 + 2, 2) = (6, 2)$$

Hence, we have $x + y \neq y + x$ and so V is **NOT** a vector space. ■

Exercise 1.2.14

Let $V = \{(a_1, a_2, \dots, a_n) : a_i \in C \text{ for } i = 1, 2, \dots, n\}$; so V is a vector space over C by Example 1. Is V a vector space over the field of real numbers with the operations of coordinate-wise addition and multiplication?

Proof. Yes, V where

$$V = \{(a_1, a_2, \dots, a_n) : a_i \in \mathbb{R} \text{ for } i = 1, 2, \dots, n\}$$

is a vector space of \mathbb{R} .

(VS 1) Let $x, y \in V$ such that $x = (a_1, a_2, \dots, a_n)$ and $y = (b_1, b_2, \dots, b_n)$. Since addition is

entry-wise in V and each entry in both x and y are elements of \mathbb{R} (where \mathbb{R} is a field), we have $a_i + b_i = b_i + a_i$ for all $i = 1, 2, \dots, n$. Hence, $x + y = y + x$.

- (VS 2) Let $x, y, z \in V$ with x and y as defined as before where z contains entries c_i for all $i = 1, 2, \dots, n$. We can see that the entries of x, y, z are elements of \mathbb{R} so associativity is preserved; that is, $a_i + (b_i + c_i) = (a_i + b_i) + c_i$ for all $i = 1, 2, \dots, n$. Hence, we have $x + (y + z) = (x + y) + z$.
- (VS 3) Since \mathbb{R} contains the zero element 0 and V is the set of n -tuples, there exists an element denoted by O such that this element consisting of entries that only have the zero element 0; that is, $O = (0, 0, \dots, 0)$. Take $x \in V$. Hence, we have $a_i + 0 = a_i$ for every $i = 1, 2, \dots, n$. Thus, we must have $x + O = x$.
- (VS 4) Let $x \in V$ be arbitrary as defined before. Since every entry in x is an element of \mathbb{R} ; that is, every $a_i \in \mathbb{R}$ for all $i = 1, 2, \dots, n$, we know that every entry contains an element c_i such that $a_i + c_i = 0$ for every $i = 1, 2, \dots, n$. Denote $x' = (c_1, c_2, \dots, c_n)$. Hence, we have $x + x' = O$.
- (VS 5) Let $x \in V$. Every entry $a_i \in \mathbb{R}$ for all $i = 1, 2, \dots, n$, we have $1 \cdot a_i = a_i$ which holds for all i . Denote this identity element as I with entries consisting only of 1. Hence, we have $I \cdot x = x$.
- (VS 6) Let $e, r \in \mathbb{R}$ and let $x \in V$. For every entry $a_i \in \mathbb{R}$ for all $i = 1, 2, \dots, n$, we have $(er)a_i = e(ra_i)$ for all $i = 1, 2, \dots, n$. By using the operations of scalar multiplication for n -tuples, this tells us that $(er)x = e(rx)$. Hence, (VS 6) is satisfied.
- (VS 7) Let $e \in \mathbb{R}$ and let $x, y \in V$. We need to show that $e(x + y) = ex + ey$. Note that $a_i, b_i \in \mathbb{R}$ implies that $e(a_i + b_i) = ea_i + eb_i$. Hence, we have $e(x + y) = ex + ey$.
- (VS 8) Let $e, r \in \mathbb{R}$ and $x \in V$ as defined before. We need to show that $(e + r)x = ex + rx$. Since $a_i \in \mathbb{R}$ with $e, r \in \mathbb{R}$, we are guaranteed to have $(e + r)a_i = ea_i + ra_i$ for all $i = 1, 2, \dots, n$. Hence, we have $(e + r)x = ex + rx$.

■

Exercise 1.2.16

Let V denote the set of all $m \times n$ matrices with real entries; so V is a vector space over \mathbb{R} by Example 2. Let F be the field of rational numbers. Is V a vector space over F with the usual definitions of matrix addition and scalar multiplication?

Proof(VS 1) Let $A, B \in V$. Since A, B consist of elements $A_{ij}, B_{ij} \in \mathbb{R}$, we know that $A_{ij} + B_{ij} = B_{ij} + A_{ij}$. Hence, $A + B = B + A$.

(VS 2) Let $A, B, C \in V$ with A, B defined as before and C containing real entries C_{ij} . With the same reasoning used to prove (VS 1), we know that $A_{ij} + (B_{ij} + C_{ij}) = (A_{ij} + B_{ij}) + C_{ij}$. Hence, we have $A + (B + C) = (A + B) + C$.

(VS 3) Let $A \in V$ once again. Since the entries of A imply that there exists an element O such that $A_{ij} + 0 = A_{ij}$, we know that $A + O = A$ where O is the **zero matrix** of V .

(VS 4) Since the real entries of A also consists of an element A'_{ij} such that $A_{ij} + A'_{ij} = 0$, this implies that $A + A' = O$ where A' is the additive inverse matrix of V .

(VS 5) Let $x \in V$ as defined as before. Every entry of A , $A_{ij} \in \mathbb{R}$, has the following property:

$1 \cdot A_{ij} = A_{ij}$ for all for all $1 \leq i \leq m$ and for all $1 \leq j \leq n$. The matrix whose entries consists of only one we can define as the **identity matrix** denoted by I where $I_{ij} = 1$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$. Hence, we have $A \cdot I = A$.

(VS 6) Let $r, t \in \mathbb{Q}$ and let $A \in V$ as defined before. We need to show that $(rt)A = r(tA)$. Since $A_{ij} \in \mathbb{R}$, entry-wise scalar multiplication implies that $(rt)A_{ij} = r(tA_{ij})$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$. Hence, we must have $(rt)A = r(tA)$.

(VS 7) Let $r \in \mathbb{Q}$ and $A, B \in V$ as defined before. Since $A_{ij}, B_{ij} \in \mathbb{R}$, we know that $r(A_{ij} + B_{ij}) = rA_{ij} + rB_{ij}$. Hence, we have $r(A + B) = rA + rB$.

(VS 8) Let $r, t \in \mathbb{Q}$ and let $A \in V$ as defined before. Since $A_{ij} \in \mathbb{R}$ for all i, j , we must have $(r + t)A_{ij} = rA_{ij} + tA_{ij}$. Hence, we have $(r + t)A = rA + tA$.

Hence, V is a vector space over \mathbb{Q} . ■

Exercise 1.2.17

Let $V = \{(a_1, a_2) : a_1, a_2 \in F\}$, where F is a field. Define addition of elements of V coordinate-wise, and for $c \in F$ and $(a_1, a_2) \in V$, define

$$c(a_1, a_2) = (ca_1, 0).$$

Is V a vector space over F with these operations? Justify your answer.

Proof. We claim that V is not a vector space over F because V fails to satisfy (VS 5). To see why, let $(1, 2) \in V$. Using (VS 5), we have

$$1 \cdot (1, 2) = (1, 0) \neq (1, 2).$$

Hence, V cannot be a vector space. ■

Exercise 1.2.18

Let $V = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\}$. For $(a_1, a_2), (b_1, b_2) \in V$ and $c \in \mathbb{R}$, define

$$(a_1, a_2) + (b_1, b_2) = (a_1 + 2b_1, a_2 + 3b_2) \text{ and } c(a_1, a_2) = (ca_1, ca_2).$$

Is V a vector space over $F = \mathbb{R}$ with these operations? Justify your answer?

Proof. We claim that V is not a vector space over \mathbb{R} and we will use (VS 1) to show this. Let $x, y \in V$ be defined by $x = (1, 2)$ and $y = (3, 4)$. Observe that

$$x + y = (1, 2) + (3, 4) = (7, 14)$$

and

$$y + x = (3, 4) + (1, 2) = (5, 10).$$

Clearly, we have $x + y = (7, 14) \neq (5, 10) = y + x$ and so (VS 1) does not hold. ■

Exercise 1.2.19

Let $V = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\}$. Define addition of elements of V coordinate-wise, and for $(a_1, a_2) \in V$ and $c \in \mathbb{R}$, define

$$c(a_1, a_2) = \begin{cases} (0, 0) & \text{if } c = 0 \\ (ca_1, \frac{a_2}{c}) & \text{if } c \neq 0. \end{cases}$$

Is V a vector space over \mathbb{R} with these operations? Justify your answer.

Proof. We claim that V is not a vector space over \mathbb{R} . To see why, consider (VS 8). If we let $(0, 1) \in V$ with $c = 2 + 1 = 3$. Observe that

$$(2 + 1)(0, 1) = \left(0, \frac{1}{2 + 1}\right) = \left(0, \frac{1}{3}\right).$$

Likewise, we have

$$2(0, 1) + 1(0, 1) = \left(0, \frac{1}{2}\right) + (0, 1) = \left(0, \frac{3}{2}\right).$$

Notice that $(2 + 1)(0, 1) \neq 2(0, 1) + 1(0, 1)$. Hence, V cannot be a vector space over \mathbb{R} . ■

Exercise 1.2.20

Let V denote the set of all real-valued functions f defined on the real line such that $f(1) = 0$. Prove that V is a vector space with the operations of addition and scalar multiplication defined in Example 3.

Proof. Define V as a vector space with the operations of addition and scalar multiplication defined in Example 3. We must show that V is a vector space.

(VS 1) Let $f, g \in V$. This means that $f(1) = 0$ and $g(1) = 0$. We need to show that $f + g = g + f$. Since $f(1), g(1) \in \mathbb{R}$ and commutativity holds in \mathbb{R} , we can write

$$\begin{aligned} (f + g)(1) &= f(1) + g(1) \\ &= g(1) + f(1) \\ &= (g + f)(1) \end{aligned}$$

Hence, we have $f + g = g + f$.

(VS 2) Let $f, g, h \in V$ then $f(1) = g(1) = h(1) = 0$. We need to show that $f + (g + h) = (f + g) + h$. Observe that

$$\begin{aligned} (f + (g + h))(1) &= f(1) + (g + h)(1) \\ &= f(1) + g(1) + h(1) \\ &= (f + g)(1) + h(1) \\ &= ((f + g) + h)(1). \end{aligned}$$

Hence, we have $f + (g + h) = (f + g) + h$.

(VS 3) Let $f \in V$. We need to show that $f + f_0 = f$ for some $f_0 \in V$. Since V contains elements of $f \in V$ such that $f(1) = 0$, we can choose f_0 such that $f_0(1) = 0$. We can show that this is indeed the additive inverse of V by writing the following:

$$(f + f_0)(1) = f(1) + f_0(1) = f(1) + 0 = f(1).$$

Hence, we have $f + f_0 = f$.

(VS 4) Let $f \in V$. We need to find an element $g \in V$ such that $f + g = f_0$. By definition of V , $f(1) = 0$. We need to show that $f + g = f_0$ with f_0 defined as before. Choose $g = -f$

as our additive inverse and observe that

$$\begin{aligned}
 (f + g)(1) &= (f - f)(1) \\
 &= f(1) - f(1) \\
 &= 0 - 0 \\
 &= 0 \\
 &= f_0(1).
 \end{aligned}$$

Hence, $g = -f$ an element such that $f + g = f_0$.

(VS 5) Let $f \in V$. By definition of V , we have $f(1) = 0$. Since $f(1) \in \mathbb{R}$, we know that $1 \cdot f(1) = f(1)$. We need to show that $1f = f$. Observe that

$$(1f)(1) = 1 \cdot f(1) = f(1).$$

Hence, (VS 5) holds.

(VS 6) Let $a, b \in \mathbb{R}$ and let $x \in V$. We need to show that $(ab)f = a(bf)$. By using scalar multiplication, we can see that

$$(ab \cdot f)(1) = (ab)f(1) = a(bf(1)) = a(b \cdot f)(1).$$

Hence, (VS 6) must hold.

(VS 7) Let $a \in \mathbb{R}$ and let $f, g \in V$. We need to show that $a(f + g) = af + ag$. Observe that

$$\begin{aligned}
 a(f + g)(1) &= a(f(1) + g(1)) \\
 &= af(1) + ag(1) \\
 &= (af)(1) + (ag)(1).
 \end{aligned}$$

Hence, $a(f + g) = af + ag$ and so (VS 7) is satisfied.

(VS 8) Let $a, b \in \mathbb{R}$ and let $f \in V$. Observe that

$$\begin{aligned}
 (a + b)f(1) &= af(1) + bf(1) \\
 &= (af)(1) + (bf)(1)
 \end{aligned}$$

Hence, $(a + b)f = af + bf$. Thus, V must be a vector space over \mathbb{R} . ■

Exercise 1.2.21

Let V and W be vector spaces over a field F . Let

$$Z = \{(v, w) : v \in V \text{ and } w \in W\}.$$

Prove that Z is a vector space over F with the operations

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2) \text{ and } c(v_1, w_1) = (cv_1, cw_1).$$

Proof. Let V and W be vector spaces over a field F .

(VS 1) Let $x, y \in Z$ where $x = (v_1, w_1)$ and $y = (v_2, w_2)$. Since V and W are vector spaces, commutativity holds. Since the addition defined on both V and W is entry-wise, we

can see that

$$\begin{aligned}
 x + y &= (v_1, w_1) + (v_2, w_2) \\
 &= (v_1 + v_2, w_1 + w_2) \\
 &= (v_2 + v_1, w_2 + w_1) \\
 &= (v_2, w_2) + (v_1, w_1) \\
 &= y + x.
 \end{aligned}$$

Hence, we can see that (VS 1) is satisfied.

(VS 2) Let $x, y, z \in Z$ with x and y as defined before as well as $z = (z_3, w_3)$. Using the entry-wise addition defined for Z , we can see that

$$\begin{aligned}
 x + (y + z) &= (v_1, w_1) + ((v_2, w_2) + (v_3, w_3)) \\
 &= (v_1, w_1) + (v_2 + v_3, w_2 + w_3) \\
 &= (v_1 + (v_2 + v_3), w_1 + (w_2 + w_3)) \\
 &= ((v_1 + v_2) + v_3, (w_1 + w_2) + w_3) \\
 &= (v_1 + v_2, w_1 + w_2) + (v_3, w_3) \\
 &= (x + y) + z.
 \end{aligned}$$

(VS 3) Since V and W are vector spaces, we know that there exists an additive identity O_V and O_W respectively. Hence, we have $(O_V, O_W) \in Z$ and denote $O_Z = (O_V, O_W)$. Now, let $x \in Z$ as defined before. Observe that

$$\begin{aligned}
 x + O_Z &= (v_1, w_1) + (O_V, O_W) \\
 &= (v_1 + O_V, w_1 + O_W) \\
 &= (v_1, w_1) \\
 &= x.
 \end{aligned}$$

(VS 4) Observe that V and W contain additive inverses for each $v \in V$ and $w \in W$ respectively. Since $x = (v_1, w_1)$ with $v_1 \in V$ and $w_1 \in W$, there exists an additive inverse $v'_1 \in V$ and $w'_1 \in W$ such that $v_1 + v'_1 = O_V$ and $w_1 + w'_1 = O_W$. This implies that $(v'_1, w'_1) \in Z$ which we will denote by x' such that

$$\begin{aligned}
 x + x' &= (v_1, w_1) + (O_V, O_W) \\
 &= (v_1 + v'_1, w_1 + w'_1) \\
 &= (O_V, O_W) \\
 &= O_Z.
 \end{aligned}$$

(VS 5) Let $x \in Z$. Since V and W are vector spaces (VS 5) implies that $1 \cdot v_1 = v_1$ and $1 \cdot w_1 = w_1$ respectively. Then observe that

$$\begin{aligned}
 1 \cdot x &= 1 \cdot (v_1, w_1) \\
 &= (1 \cdot v_1, 1 \cdot w_1) \\
 &= (v_1, w_1) \\
 &= x.
 \end{aligned}$$

(VS 6) Let $x \in Z$ and $a, b \in F$. Then

$$\begin{aligned}
 (ab)x &= (ab)(v_1, w_1) \\
 &= ((ab)v_1, (ab)w_1) \\
 &= (a(bv_1), a(bw_1)) && (V, W \text{ vector space}) \\
 &= a(bv_1, bw_1) \\
 &= a(bx)
 \end{aligned}$$

(VS 7) Let $x, y \in Z$ as defined before. Let $a \in F$. Then

$$\begin{aligned}
 a(x + y) &= a((v_1, w_1) + (v_2, w_2)) \\
 &= a((v_1 + v_2, w_1 + w_2)) \\
 &= (a(v_1 + v_2), a(w_1 + w_2)) \\
 &= (av_1 + av_2, aw_1 + aw_2) \\
 &= (av_1, av_2) + (aw_1, aw_2) \\
 &= a(v_1, v_2) + a(w_1, w_2) \\
 &= ax + ay.
 \end{aligned}$$

(VS 8) Let $a, b \in F$ and let $x \in V$ as defined before. Since V and W are vector spaces, we know that (VS 8) holds for bot entries $v_1 \in V$ and $w_1 \in W$. Hence, observe that

$$\begin{aligned}
 (a + b)x &= (a + b)(v_1, w_1) \\
 &= ((a + b)v_1, (a + b)w_1) \\
 &= (av_1 + bv_1, aw_1 + bw_1) \\
 &= (av_1, aw_1) + (bv_1, bw_1) \\
 &= a(v_1, w_1) + b(v_1, w_1) \\
 &= ax + bx.
 \end{aligned}$$

Hence, Z is a vector space. ■

1.2 Subspaces

Exercise 1.3.3

Prove that $(aA + bB)^t = aA^t + bB^t$ for any $A, B \in M_{m \times n}(F)$ and any $a, b \in F$.

Proof. Let $A, B \in M_{m \times n}(F)$ and let $a, b \in F$ be arbitrary. Using scalar multiplication defined on $M_{m \times n}(F)$, we have

$$\begin{aligned}
 (aA + bB)^t &= (aA)^t + (bB)^t \\
 &= aA^t + bB^t.
 \end{aligned}$$

Hence, we are done. ■

Exercise 1.3.4

Prove that $(A^t)^t = A$ for each $A \in M_{n \times n}(F)$.

Proof. Let $A \in M_{n \times n}(F)$. By definition of transpose, we have

$$\left((A^t)^t\right)_{ij} = (A^t)_{ji} = A_{ij}$$

for all $1 \leq i, j \leq n$. Hence, $(A^t)^t = A$. ■

Exercise 1.3.5

Prove that $A + A^t$ is symmetric for any square matrix A .

Proof. Let A be an arbitrary square matrix. Since square matrices are symmetric, we have that $A^t = A$. We need to show that $(A + A^t)^t$. Observe that

$$\begin{aligned}(A + A^t)^t &= A^t + (A^t)^t \\ &= A + A^t.\end{aligned}$$

Hence, we have $A + A^t$ is symmetric. ■

Exercise 1.3.

Prove that $\text{tr}(aA + bB) = a\text{tr}(A) + b\text{tr}(B)$ for any $A, B \in M_{n \times n}(F)$.

Proof. Let $A, B \in M_{n \times n}(F)$ and let $a, b \in F$ be arbitrary. Now, let $i = j$ and observe that

$$\begin{aligned}\text{tr}(aA + bB) &= \sum_{i,j \in \mathbb{N}}^n (aA + bB)_{ij} \\ &= \sum_{i,j \in \mathbb{N}}^n (aA)_{ij} + (bB)_{ij} \\ &= \sum_{i,j \in \mathbb{N}}^n (aA)_{ij} + \sum_{i,j \in \mathbb{N}}^n (bB)_{ij} \\ &= \sum_{i,j \in \mathbb{N}}^n aA_{ij} + \sum_{i,j \in \mathbb{N}}^n bB_{ij} \\ &= a \sum_{i,j \in \mathbb{N}}^n A_{ij} + b \sum_{i,j \in \mathbb{N}}^n B_{ij} \\ &= a\text{tr}(A) + b\text{tr}(B).\end{aligned}$$

Hence, we conclude

$$\text{tr}(aA + bB) = a\text{tr}(A) + b\text{tr}(B)$$

for any $A, B \in M_{n \times n}(F)$. ■

Exercise 1.3.7

Prove that diagonal matrices are symmetric matrices.

Proof. Let $A \in M_{n \times n}(F)$ be diagonal. Let $i \neq j$ where $A_{ij} = 0$. We need to show that $A^t = A$. If we apply a transpose on A , we get that $A_{ji} = 0$ since A^t is also diagonal and square. Since $A_{ij} = A_{ji} = 0$ for all $1 \leq i \leq n$ and $1 \leq j \leq n$. Hence, $A^t = A$ ■

Exercise 1.3.8

Determine whether the following sets are subspaces of \mathbb{R}^3 under the operations of addition and scalar multiplication defined on \mathbb{R}^3 . Justify your answers.

- (a) $W_1 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 = 3a_2 \text{ and } a_3 = -a_2\}$

Proof. We claim that W_1 is a subspace of \mathbb{R}^3 .

- (a) Note that $O_{\mathbb{R}^3} \in W_1$ where $O_{\mathbb{R}^3} = (0, 0, 0)$ because $0 = 3 \cdot 0$ and $0 = -1 \cdot 0$.
- (b) Let $x, y \in W_1$ where $x = (a_1, a_2, a_3)$ and $y = (b_1, b_2, b_3)$. We need to show that $x + y \in W_1$. Since $a_1 = 3a_2$ and $a_3 = -a_2$ as well as $b_1 = 3b_2$ and $b_3 = -b_2$, we can write $a_1 + b_1 = 3(a_2 + b_2)$ and $a_3 + b_3 = -(a_2 + b_2)$. Hence, $x + y \in W_1$.
- (c) Let $c \in \mathbb{R}$ and $x \in W_1$ with x defined as before. Then observe that $ca_1 = c(3a_2) = 3(ca_2)$ and $ca_3 = c(-a_2) = -(ca_2)$. Hence, $cx \in W_1$.

Since all the properties of a Theorem 3 have been satisfied, we can conclude that W_1 is a subspace of \mathbb{R}^3 . ■

- (b) $W_2 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 = a_3 + 2\}$
- (c) $W_3 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : 2a_1 - 7a_2 + a_3 = 0\}$
- (d) $W_4 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 - 4a_2 - 3a_3 = 1\}$
- (e) $W_5 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 + 2a_2 - 3a_3 = 1\}$
- (f) $W_6 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : 5a_1^2 - 3a_2^2 + 6a_3^2 = 0\}$

Exercise 1.3.9

Let W_1, W_3, W_4 be as in Exercise 8. Describe $W_1 \cap W_3$, $W_1 \cap W_4$, and $W_3 \cap W_4$ and observe that each is a subspace of \mathbb{R}^3 .

Proof. ■

Exercise 1.3.11

Prove that the set $W_1 = \{(a_1, a_2, \dots, a_n) \in F^n : a_1 + a_2 + \dots + a_n = 0\}$ is a subspace of F^n , but $W_2 = \{(a_1, a_2, \dots, a_n) \in F^n : a_1 + a_2 + \dots + a_n = 1\}$ is not.

Proof. We need to show that W_1 is a subspace of F^n . We proceed by satisfying the properties of Theorem 3 to do this.

- (a) Note that $O_{F^n} \in W_1$ since $0 + 0 + \dots + 0 = 0$ n times.
- (b) Let $x, y \in W_1$ with $x = (a_1, a_2, \dots, a_n)$ and $y = (b_1, b_2, \dots, b_n)$. By definition of W_1 ,

we can see that

$$\begin{aligned}\sum_{i=1}^n (a_i + b_i) &= \sum_{i=1}^n a_i + \sum_{i=1}^n b_i \\ &= 0 + 0 \\ &= 0.\end{aligned}$$

Hence, $x + y \in W_1$ which tells us that W_1 is closed under addition.

(c) Let $x \in W_1$ and $c \in F$. Then observe that

$$\sum_{i=1}^n (ca_i) = c \sum_{i=1}^n a_i = c \cdot 0 = 0.$$

Hence, we have $cx \in F^n$.

We claim that W_2 is not a subspace because W_2 is not closed under addition. Let $(0, 1), (1, 0) \in F^2$. Observe that $0+1 = 1$ and $1+0 = 1$, but $(0+1)+(1+0) = 1+1 = 2$. Hence, $(0, 1) + (1, 0) \notin W_2$. ■

Exercise 1.3.11

Is the set $W = \{f(x) \in P(F) : f(x) = 0 \text{ or } f(x) \text{ has degree } n\}$ a subspace of $P(F)$ if $n \geq 1$? Justify your answer.

Exercise 1.3.12

Prove that the set of $m \times n$ upper triangular matrices is a subspace of $M_{m \times n}(F)$.

Proof. Let V denote the set of $m \times n$ upper triangular matrices. We will show that V is a subspace of $M_{m \times n}(F)$ using Theorem 3.

- (a) The zero matrix O from $M_{m \times n}(F)$ contains entries $O_{ij} = 0$ whenever $i > j$. Hence, $O \in V$.
- (b) Let $A, B \in V$. By definition of V , A and B are upper triangular where $A_{ij} = 0$ and $B_{ij} = 0$ whenever $i > j$. Observe that

$$(A + B)_{ij} = A_{ij} + B_{ij} = 0 + 0 = 0$$

whenever $i > j$. Hence, $A + B \in V$.

- (c) Let $c \in F$ and $A \in V$ as defined before. Let $i > j$ and observe that $(cA)_{ij} = cA_{ij} = c \cdot 0 = 0$. Hence, $cA \in V$.

Since all the properties of theorem 3 have been satisfied, we conclude that V is indeed a subspace of $M_{m \times n}(F)$. ■

Exercise 1.3.13

Let S be nonempty set and F is a field. Prove that for any $s_0 \in S$, the set $\{f \in \mathcal{F}(S, F) : f(s_0) = 0\}$, is a subspace of $\mathcal{F}(S, F)$.

Proof. We will proceed to prove that S is a subspace of $\mathcal{F}(S, F)$ over the field F by satisfying the properties of Theorem 3. Let $V = \{f \in \mathcal{F}(S, F) : f(s_0) = 0\}$.

(a) Note that the zero function $f_0 \in \mathcal{F}(S, F)$ where $f_0(s_0) = 0$ for any $s_0 \in S$ implies that $f_0 \in V$.

(b) Let $f, g \in V$. By definition of V , $f(s_0) = 0$ and $g(s_0) = 0$ for any $s_0 \in S$. We have $f + g \in V$ since

$$(f + g)(s_0) = f(s_0) + g(s_0) = 0 + 0 = 0.$$

Hence, V is closed under addition.

(c) Let $f \in V$ and $c \in F$. We have $cf \in V$ since

$$(cf)(s_0) = cf(s_0) = c \cdot 0 = 0$$

for any $s_0 \in S$. Hence, V is closed under scalar multiplication. ■

Exercise 1.3.14

Let S be a nonempty set and F a field. Let $\mathcal{C}(S, F)$ denote the set of all functions $f \in \mathcal{F}(S, F)$ such that $f(s) = 0$ for all but a finite number of elements of S . Prove that $\mathcal{C}(S, F)$ is a subspace of $\mathcal{F}(S, F)$.

Proof. We proceed by using Theorem 3 to prove that $\mathcal{C}(S, F)$ is a subspace where S is a nonempty set and F is a field.

(a) Note that the zero vector $f_0 \in \mathcal{F}(S, F)$ is in $\mathcal{C}(S, F)$ because $f_0(x_n) = 0$ where $x_n \in S$ for finitely many n .

(b) Let $f, g \in \mathcal{C}(S, F)$. We need to show that $f + g \in \mathcal{C}(S, F)$. Let $x_n \in S$ for finitely many n . Then using the addition defined on $\mathcal{F}(S, F)$, we can write

$$(f + g)(x_n) = f(x_n) + g(x_n) = 0 + 0 = 0.$$

(c) Let $f \in \mathcal{C}(S, F)$ and $c \in F$. We need to show that $cf \in \mathcal{C}(S, F)$. Let $x_n \in S$ for finitely many n . Using the scalar operation defined on $\mathcal{F}(S, F)$, we can write

$$(cf)(x_n) = cf(x_n) = c \cdot 0 = 0.$$

Since all the properties of Theorem 3 have been satisfied, we conclude that $\mathcal{C}(S, F)$ is indeed a subspace of $\mathcal{F}(S, F)$. ■

Exercise 1.3.15

Is the set of all differentiable real-valued functions defined on \mathbb{R} a subspace of $C(\mathbb{R})$?

Proof. We claim that the set of all differentiable real-valued functions defined on \mathbb{R} is a subspace of $C(\mathbb{R})$. Denote this set as V .

(a) Note that the zero function f_0 is differentiable for all $x \in \mathbb{R}$ and continuous for all $x \in \mathbb{R}$. Hence, $f_0 \in V$.

- (b) Let $f, g \in V$. Using the addition operation defined on $C(\mathbb{R})$, we get that the sum $(f + g)(x) = f(x) + g(x)$ is differentiable which implies that the sum of functions f, g is also continuous. Hence, $f + g \in V$.
- (c) Let $f \in V$ and let $c \in \mathbb{R}$. Then $(cf)(x) = cf(x)$ is differentiable for all $x \in \mathbb{R}$ which means that cf is also continuous. Hence, $cf \in V$.

Hence, V is a subspace of $C(\mathbb{R})$. ■

Exercise 1.3.16

Let $C^n(\mathbb{R})$ denote the set of all real-valued functions defined on the real line that have a continuous n th derivative. Prove that $C^n(\mathbb{R})$ is a subspace of $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Proof. ■

Exercise 1.3.17

Prove that a subset W of a vector space V is a subspace of V if and only if $W \neq \emptyset$ and, whenever $a \in F$ and $x, y \in W$, then $ax \in W$ and $x + y \in W$.

Proof. For the forwards direction, let $W \subseteq V$ where V is a vector space and W is a subspace of V . Let $a \in F$ and $x, y \in W$. Since W is a subspace, we know that $O_V \in W$. So, W is nonempty. Since W is closed under addition and multiplication, we get that $x + y \in W$ and $ax \in W$ and we are done.

For the backwards direction, let $W \neq \emptyset$ and $W \subseteq V$. Let $a \in F$ and $x, y \in W$ be arbitrary such that $ax \in W$ and $x + y \in W$. We need to show that W is a subspace of V . We need only show that $O_V \in W$ since W is closed under addition and scalar multiplication. Let $x \in W$. We can pick any $c \in F$ such that $c = 0$. So, we have $c \cdot x = 0 \cdot x = O_w$. Since the zero vector $O_W \in W$ is unique, we must have $O_V = O_W$. Hence, $O_V \in W$ and we conclude that W is a subspace of V . ■

Exercise 1.3.18

Prove that a subset W of a vector space V is a subspace of V if and only if $O \in W$ and $ax + y \in W$ whenever $a \in F$ and $x, y \in W$.

Proof. (\Rightarrow) Let $W \subseteq V$ where W is a subspace of V . Since W is a subspace of V , we know that W is closed under addition and scalar multiplication. Let $a \in F$ and $x, y \in W$. Using the third property of Theorem 3, we can see that $ax \in W$. Since W is closed under addition, we can take $y \in W$ and $ax \in W$ such that $ax + y \in W$. Since W is also a vector space by definition, we know that $O_W \in W$. But $O_W = O_V$ so $O_V \in W$.

(\Leftarrow) Let $a \in F$ and $x, y \in W$. We want to show that $W \subseteq V$ is a subspace of V . We can do this by using Theorem 3.

- (a) By assumption, the zero vector $O_V \in W$.
- (b) Let $x, y \in W$. Choose $a = 1$ such that $ax + y = x + y$. Since $ax + y \in W$ and $ax + y = x + y$, we also have $x + y \in W$. Hence, W is closed under addition.
- (c) Let $x \in W$ and $O_V \in W$. Let $a \in F$. Then we have $ax + O_V = ax \in W$.

Hence, W is a subspace of V by Theorem 3. ■

Exercise 1.3.19

Let W_1 and W_2 be subspaces of a vector space V . Prove that $W_1 \cup W_2$ is a subspace of V if and only if $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Proof. (\Rightarrow) Let $W_1 \cup W_2$ is a subspace of V . We need to show that $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$. We proceed by showing the contrapositive. Assume $W_2 \not\subseteq W_1$ and $W_1 \not\subseteq W_2$. We need to show that $W_1 \cup W_2$ is **NOT** a subspace of V . By assumption, $x \in W_1$ is not contained in W_2 as well as $y \in W_2$ is not contained in W_1 . This implies that $W_1 \cup W_2 \neq \emptyset$. Since $W_1 \cup W_2$ is empty where $W_1 \cup W_2$ does not contain O_V , it cannot possibly be a subspace of V .

(\Leftarrow) Let $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$. We need to show that $W_1 \cup W_2$ is a subspace of V . We proceed by using Theorem 3 to do this. Without loss of generality, assume $W_1 \subseteq W_2$. The proof will be the same if we use $W_2 \subseteq W_1$.

- (a) Since W_1 is a subspace of V , we get that $O_V \in W_1$. Furthermore, $W_1 \subseteq W_2$ implies that $O_V \in W_2$. Since $O_V \in W_1$ and $O_V \in W_2$, we get that $O_V \in W_1 \cup W_2$ by definition of union.
- (b) Let $x, y \in W_1$. Since W_1 is a subspace, we get that $x + y \in W_1$. Since $W_1 \subseteq W_2$, we also get that $x + y \in W_2$. Since both $x + y \in W_1$ and $x + y \in W_2$, we know that $x + y \in W_1 \cup W_2$ by definition of the union.
- (c) Let $x \in W_1$ and $c \in F$. Since W_1 is closed under scalar multiplication, we have that $cx \in W_1$. But $W_1 \subseteq W_2$ so W_2 also contains $cx \in W_1$. So we must have $cx \in W_1 \cup W_2$.

Hence, $W_1 \cup W_2$ is a subspace of a vector space V . ■

Exercise 1.3.20

Prove that if W is a subspace of a vector space V and w_1, w_2, \dots, w_n are in W , then $a_1w_1 + a_2w_2 + \dots + a_nw_n \in W$.

Proof. Let W be a subspace of a vector space V . Our goal is to show that the following statement: $a_1w_1 + a_2w_2 + \dots + a_nw_n \in W$ for all $1 \leq i \leq n$. We proceed by induction on $i \geq 1$. Let $i = 1$. Since W is closed under scalar multiplication, we know that $a_1 \in F$ and $w_1 \in W$ implies that $a_1w_1 \in W$. Now let $i = 2$, then $w_1, w_2 \in W$ and $a_1, a_2 \in F$ implies that $a_1w_1 + a_2w_2 \in W$ since W is closed under scalar multiplication and addition. Now, assume that our result holds for all $1 \leq i \leq n$. We want to show that it also holds for $i = n + 1$. By our inductive hypothesis, we know that $a_1w_1 + a_2w_2 + \dots + a_nw_n \in W$. Let $a_{n+1} \in F$ and $w_{n+1} \in W$. Using vector addition and scalar multiplication, we know that

$$(a_1w_1 + a_2w_2 + \dots + a_nw_n) + a_{n+1}w_{n+1} \in W.$$
■

Exercise 1.3.21

Let V denote the vector space of sequences in \mathbb{R} , as defined in Example 5 of section 1.2. Show that the set of convergent sequences (a_n) (that is, those for which $\lim_{n \rightarrow \infty} a_n$ exists) is a subspace of V .

Proof. We will show that W (the set of convergent sequences in \mathbb{R}) is a subspace of V by using Theorem 3.

- (a) Note that the zero sequence O_n is zero for all $n = 1, 2, \dots$. Hence, the limit of O_n converges to 0 and so we have $O_n \in W$.
- (b) Let (a_n) and (b_n) be two convergent sequences in \mathbb{R} . Then we get that $\lim_{n \rightarrow \infty} a_n$ and $\lim_{n \rightarrow \infty} b_n$ exists. Observe that

$$\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n.$$

Since the a_n and b_n are both convergent sequences, we also get that the sum $(a_n + b_n)$ also converges. Hence, $(a_n + b_n) \in W$.

- (c) Let $t \in F$ and $a_n \in W$ as before. Then we have

$$\lim_{n \rightarrow \infty} ta_n = t \lim_{n \rightarrow \infty} a_n.$$

Since any constant $t \in F$ multiplied by a convergent sequence is convergent, we also get that the sequence (ta_n) is also convergent. Hence, $ta_n \in W$.

■

Exercise 1.3.22

Let F_1 and F_2 be fields. A function $g \in \mathcal{F}(F_1, F_2)$ is called an **even function** if $g(-t) = g(t)$ for each $t \in F_1$ and is called an **odd function** if $g(-t) = -g(t)$ for each $t \in F_1$. Prove that the set of all even functions in $\mathcal{F}(F_1, F_2)$ and the set of all odd functions in $\mathcal{F}(F_1, F_2)$ are subspaces of $\mathcal{F}(F_1, F_2)$.

Proof. Let V be the set of all even functions and let W be the set of all odd functions. We will first prove that V is a subspace of $\mathcal{F}(F_1, F_2)$ and the same with W .

- (a) Note that the zero function f_0 from $\mathcal{F}(F_1, F_2)$ is even since $f_0(-t) = 0 = f_0(t)$ for all $t \in F_1$. Hence, $f_0 \in V$.
- (b) Let $f, g \in V$. We need to show that $f + g \in V$; that is, we need to show that it is even. Let $t \in F_1$. Then observe that

$$(f + g)(-t) = f(-t) + g(-t) = f(t) + g(t) = (f + g)(t)$$

Hence, $(f + g)(-t) = (f + g)(t)$ for all $t \in F_1$.

- (c) Let $f \in V$ and let $c \in F_2$. Then observe that

$$(cf)(-t) = cf(-t) = cf(t) = (cf)(t).$$

Hence, $cf \in V$.

Since all the properties of V are satisfied, we have that V is a subspace of $\mathcal{F}(F_1, F_2)$.

Now we will prove W is a subspace of $\mathcal{F}(F_1, F_2)$.

- (a) Note that the zero function f_0 is in W because for any $t \in F_1$ we have $f_0(-t) = 0 = -1 \cdot 0 = -f_0(t)$. Hence, $f_0 \in W$.

(b) Let $f, g \in W$ and Let $t \in F_1$. Observe that

$$(f + g)(-t) = f(-t) + g(-t) = -f(t) - g(t) = -(f + g)(t).$$

Hence, $f + g \in W$.

(c) Let $c \in F_2$ and $f \in W$. Let $t \in F_1$ such that $f(-t) = -f(t)$. Then observe that

$$(cf)(-t) = cf(-t) = cf(t) = (cf)(t).$$

Hence, $cf \in W$.

Hence, W is a subspace of $\mathcal{F}(F_1, F_2)$ ■

Definition 1.2.1 (Sum of Two Sets). If S_1 and S_2 are nonempty subsets of a vector spaces V , then the **sum** S_1 and S_2 , denoted $S_1 + S_2$, is the set $\{x + y : x \in S_1 \text{ and } y \in S_2\}$.

Definition 1.2.2 (Direct Sum). A vector space V is called the **direct sum** of W_1 and W_2 if W_1 and W_2 are subspaces of V such that $W_1 \cap W_2 = \{0\}$ and $W_1 + W_2 = V$. We denote that V is the direct sum of W_1 and W_2 by writing $V = W_1 \oplus W_2$.

Exercise 1.3.23

Let $W_1 + W_2$ is a subspace of a vector space V .

(a) Prove that $W_1 + W_2$ is a subspace of V that contains both W_1 and W_2 .

Proof. Let W_1 and W_2 be subspaces of a vector space V . Define $W_1 + W_2$ as the set

$$\{x_1 + y_2 : w_1 \in W_1 \text{ and } w_2 \in W_2\}.$$

(a) Since W_1 and W_2 are subspaces of V , we know that the zero vector O_V is contained in both W_1 and W_2 . Hence, the sum $O_V = O_V + O_V \in W_1 + W_2$.

(b) Let $u, v \in W_1 + W_2$ with $u = x_1 + y_1$ and $v = x_2 + y_2$. Since W_1 and W_2 are subspaces of V , we know that addition is closed in both subsets W_1 and W_2 . Hence, $x_1 + x_2 \in W_1$ and $y_1 + y_2 \in W_2$. Observe that

$$\begin{aligned} u + v &= (x_1 + y_1) + (x_2 + y_2) \\ &= (x_1 + x_2) + (y_1 + y_2) \in W_1 + W_2. \end{aligned}$$

Hence, addition is closed in $W_1 + W_2$.

(c) Let $c \in F$ where F is a field and let $u \in W_1 + W_2$ with $u = x_1 + y_1$. Since W_1 and W_2 are subspaces of V , we know that $cx_1 \in W_1$ and $cy_1 \in W_2$. Observe that

$$\begin{aligned} cu &= c(x_1 + y_1) \\ &= cx_1 + cy_1 \in W_1 + W_2. \end{aligned}$$

Hence, $W_1 + W_2$ is a subspace of V . ■

(b) Prove that any subspace of V that contains both W_1 and W_2 must also contain $W_1 + W_2$.

Proof. Let X be a subspace of V . Suppose X contains both W_1 and W_2 ; that is, $W_1 \subseteq X$ and $W_2 \subseteq X$. We must show that $W_1 + W_2 \subseteq X$. Let $u \in W_1 + W_2$ with $u = x_1 + y_1$. By definition, we have $x_1 \in W_1$ and $y_1 \in W_2$. Since $W_1 \subseteq X$ and $W_2 \subseteq X$, we have $x_1 \in X$ and $y_1 \in X$. Since X is a subspace of V and X is closed under addition, we have $x_1 + y_1 \in X$. Hence, $u \in X$ and we conclude $W_1 + W_2 \subseteq X$. ■

Exercise 1.3.24

Show that F^n is the direct sum of the subspaces

$$W_1 = \{(a_1, a_2, \dots, a_n) \in F^n : a_n = 0\}$$

and

$$W_2 = \{(a_1, a_2, \dots, a_n) \in F^n : a_1 = a_2 = \dots = a_{n-1} = 0\}.$$

Proof. Let W_1 and W_2 be subspaces of F^n . Since W_1 and W_2 are subspaces, they both contain the zero vector O_{F^n} . Observe that for any element $x \in W_1$, we have the n th element $a_n = 0$. On the other hand, $y \in W_2$ implies that $a_1 = a_2 = \dots = a_{n-1} = 0$ except for the n th element. Thus, the only element that W_1 and W_2 have in common is the zero vector O_{F^n} . Hence, we have $W_1 \cap W_2 = \{O_{F^n}\}$.

Now, we show $W_1 + W_2 = F^n$. To do this, we need to show the following containments:

(i) $W_1 + W_2 \subseteq F^n$ and

(ii) $F^n \subseteq W_1 + W_2$.

Let $u \in W_1 + W_2$ with $u = x_1 + x_2$ where $x_1 \in W_1$ and $y_1 \in W_2$. Since $W_1 \subseteq F^n$, we know that $x_1 \in F^n$. Likewise, $W_2 \subseteq F^n$ implies that $y_1 \in F^n$. Since F^n is a vector space where addition is closed, we have that $u = x_1 + y_1 \in F^n$. Hence, $W_1 + W_2 \subseteq F^n$.

Now, let $u \in F^n$. Since F^n is a vector space over F , each entry in U (where each entry is an element of a field F) can be written and separated using the addition defined in F^n in the following way:

$$\begin{aligned} u &= (a_1, a_2, \dots, a_n) \\ &= (a_1 + 0, a_2 + 0, \dots, 0 + a_n) \\ &= (a_1, a_2, \dots, 0) + (0, 0, \dots, a_n) \end{aligned}$$

where the first term is an element of W_1 and the second term is an element of W_2 . Hence, $u \in W_1 + W_2$.

Since both containments are satisfied, we have that $W_1 + W_2 = F^n$. Thus, $W_1 \oplus W_2 = F^n$. ■

Exercise 1.3.25

Let W_1 denote the set of all polynomials $f(x)$ in $P(F)$ such that in the representation

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0,$$

we have $a_i = 0$ whenever i is even. Likewise, let W_2 denote the set of all polynomials $g(x)$ in $P(F)$ such that in the representation

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \dots + b_1 x + b_0,$$

we have $b_i = 0$ whenever i is odd. Prove that $P(F) = W_1 \oplus W_2$.

Proof. Let $f(x) \in W_1$. Then we have

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n$$

where $a_i = 0$ where i is even. Likewise, let $g(x) \in W_2$. Hence,

$$g(x) = b_0 + b_1x + b_2x^2 + \cdots + b_{n-1}x^{n-1} + b_nx^n$$

where $b_i = 0$ for i odd. This tells us that the only representation that W_1 and W_2 have in common is the zero polynomial $f(x) = 0$ where $a_i = 0$ for all $0 \leq i \leq n$. Hence, $W_1 \cap W_2 = \{0\}$.

Now, we want to show that $W_1 + W_2 = P(F)$; that is, we need to show $P(F) \subseteq W_1 + W_2$ and $W_1 + W_2 \subseteq P(F)$. Starting with the former, let $f \in P(F)$. Observe that for a_i where $1 \leq i \leq n$

$$\begin{aligned} f(x) &= a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n \\ &= \underbrace{(a_1x + a_3x^3 + \cdots + a_nx^n)}_{\in W_1} + \underbrace{(a_0 + a_2x^2 + a_4x^4 + \cdots + a_nx^n)}_{\in W_2} \end{aligned}$$

The first term of the last equality contains coefficients $a_i = 0$ for even i and the second term contains coefficients $a_i = 0$ for odd i . This implies that $f(x) \in W_1 + W_2$. Hence, $P(F) \subseteq W_1 + W_2$.

Now, let $u(x) \in W_1 + W_2$ with $u(x) = f(x) + g(x)$ with $f(x) \in W_1$ and $g(x) \in W_2$. Since W_1 and W_2 are subsets of $P(F)$, we have that $f(x), g(x) \in P(F)$. Since addition is closed in $P(F)$, we have that $u(x) = f(x) + g(x) \in P(F)$. Hence, $W_1 + W_2 \subseteq P(F)$. Since $W_1 + W_2 = P(F)$ and $W_1 \cap W_2 = \{0\}$, we have $W_1 \oplus W_2 = P(F)$. ■

Exercise 1.3.26

In $M_{m \times n}(F)$ define $W_1 = \{A \in M_{m \times n}(F) : A_{ij} = 0 \text{ whenever } i > j\}$ and $W_2 = \{A \in M_{m \times n}(F) : A_{ij} \text{ whenever } i \leq j\}$. Show that $M_{m \times n}(F) = W_1 \oplus W_2$. (W_1 is the set of all triangular matrices as defined in the previous section.)

Proof. TO DO. ■

Exercise 1.3.30

Let W_1 and W_2 be subspaces of a vector space V . Prove that V is the direct sum of W_1 and W_2 if and only if each vector in V can be *uniquely* written as $x_1 + x_2$, where $x_1 \in W_1$ and $x_2 \in W_2$.

Proof. Let W_1 and W_2 be subspaces of V . Let $v \in V$. Let $x_1, x'_1 \in W_1$ and $x_2, x'_2 \in W_2$ such that $v = x_1 + x_2 = x'_1 + x'_2$. Since W_1 and W_2 is closed under addition, we know that $x_1 - x'_1 \in W_1$ and $x_2 - x'_2 \in W_2$. But observe that $x_1 - x'_1 = x_2 - x'_2 \in W_1 \cap W_2$. Since V is a direct sum of the two subspaces W_1 and W_2 , we know that $W_1 \cap W_2 = \{0_V\}$ which implies that $x_1 = x'_1$ and $x_2 = x'_2$. This tells us that every v can be expressed uniquely as $x_1 + x_2$ with $x_1 \in W_1$ and $x_2 \in W_2$.

Conversely, suppose every vector $v \in V$ can be *uniquely* written as $x_1 + x_2$ with $x_1 \in W_1$ and $x_2 \in W_2$. We need to show that $W_1 \oplus W_2 = V$. Since $v \in V$ is uniquely expressed in terms of $x_1 + x_2$, the only vector that the two subspaces W_1 and W_2 share is the zero vector.

Hence, $W_1 \cap W_2 = \{O_V\}$. Now, we need to show that $V = W_1 + W_2$; that is, we need to show $V \subseteq W_1 + W_2$ and $W_1 + W_2 \subseteq V$. Suppose $v \in V$. Since $x_1 \in W_1$ and $x_2 \in W_2$, we have that $v \in W_1 + W_2$. Hence, $V \subseteq W_1 + W_2$. Now, let $v \in W_1 + W_2$. Since $v = x_1 + x_2$ where $x_1 \in W_1$ and $x_2 \in W_2$ and $W_1 \subseteq V$ and $W_2 \subseteq V$, we know that x_1 and x_2 are contained within V . Since V is vector space, we know that sum of x_1 and x_2 are contained in V . Hence, $v \in V$ and so $V \subseteq W_1 + W_2$. Thus, $V = W_1 \oplus W_2$. ■

Definition 1.2.3 (Cosets). Let W be a subspace of a vector space V over a field F . For any $v \in V$ the set $\{v\} + W = \{v + w : w \in W\}$ is called the **coset** of W **containing** v . It is customary to denote this coset by $v + W$ rather than $\{v\} + W$.

Definition 1.2.4 (Addition of Cosets). The **addition** of two cosets $v_1 + W$ and $v_2 + W$ is defined as

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

for all $v_1, v_2 \in V$.

Definition 1.2.5 (Scalar Multiplication of Cosets). The **scalar multiplication** of cosets by scalars of F can be defined as

$$a(v + W) = av + W$$

for all $v \in V$ and $a \in F$.

Exercise 1.3.31

The following exercises require the application of the definitions above.

- (a) Prove that $v + W$ is a subspace of V if and only if $v \in W$.

Proof. Suppose $v + W$ is a subspace of V . Then $0_V \in v + W$ which implies that $0_V = v + w$. Hence, $v = -w$ where $-w \in W$ and thus $v \in W$. Conversely, suppose $v \in W$.

- (a) Since W is a subspace of V , we know that $0_V \in W$. Since $v \in W$, there exists an element $v' \in W$ such that $v + v' = 0_V$. But this tell us that $0_V \in v + W$.
 (b) Let $x, y \in v + W$. By definition, $x = v + w$ and $y = v + w'$ for $w, w' \in W$. Then observe that

$$x + y = (v + w) + (v + w') = v + (v + w + w')$$

where $v + w + w' \in W$ since W is a subspace of V . Hence, $v + W$ is closed under addition.

- (c) Let $x \in v + W$. Choose $c \in F$ for which we will show that $cx \in v + W$. Then

$$cx = c(v + w) = cv + cw = v(cv + cw - v)$$

where $cv + cw - v$ since W is a subspace of V .

Hence, $v + W$ is a subspace of V . ■

- (b) Prove that $v_1 + W = v_2 + W$ if and only if $v_1 - v_2 \in W$.

Proof. Let $x \in v_1 + W$. Then for $\alpha \in W$, we have $x = v_1 + \alpha$. Since $v_1 + W = v_2 + W$, then $x \in v_2 + W$ implies that for $\beta \in W$, we have $x = v_2 + \beta$. Then observe that

$$v_1 + \alpha = v_2 + \beta \Rightarrow v_1 - v_2 = \beta - \alpha.$$

Since W is a subspace, we have that $\beta - \alpha = v_1 - v_2 \in W$.

Conversely, suppose $v_1 - v_2 \in W$. By part (a), we know that $(v_1 - v_2) + W$ is a subspace of V . Hence, $0_V \in (v_1 - v_2) + W$. This means that there exists $w' \in W$ such that

$$\begin{aligned} (v_1 - v_2) + w' &= 0_V \\ \Rightarrow v_1 + w' &= v_2 + (w - w) \\ \Rightarrow v_1 + \underbrace{(w' + w)}_{\in W} &= v_2 + \underbrace{w}_{\in W}. \end{aligned}$$

This tells us that $v_1 + W = v_2 + W$ and we are done. ■

- (c) Prove that the preceding operations are well defined; that is, show that if $v_1 + W = v'_1 + W$ and $v_2 + W = v'_2 + W$, then

$$(v_1 + W) + (v_2 + W) = (v'_1 + W) + (v'_2 + W)$$

and

$$a(v_1 + W) = a(v'_1 + W)$$

for all $a \in F$.

Proof. Suppose $v_1 + W = v'_1 + W$ and $v_2 + W = v'_2 + W$. Then $v_1 - v'_1 \in W$ and $v_2 - v'_2 \in W$ by part (b). Consequently, we have $(v_1 - v'_1) + (v_2 - v'_2) \in W$ if and only if $(v_1 + v_2) - (v'_1 + v'_2) \in W$ since W is subspace. Using part (b) again, we have that

$$(v_1 + v_2) + W = (v'_1 + v'_2) + W.$$

Using the addition defined on cosets, we get

$$(v_1 + W) + (v_2 + W) = (v'_1 + W) + (v'_2 + W).$$

Since W is a subspace, we have $c(v_1 - v'_1) \in W$ if and only if $cv_1 - cv'_1 \in W$ for some $a \in F$. By part (b), we get that

$$av_1 + W = av'_1 + W$$

which can be re-written to

$$a(v_1 + W) = a(v'_1 + W)$$

as our desired result. ■

- (d) Prove that the set $S = \{v + W : v \in V\}$ is a vector space with the operations defined in (c). This vector space is called the **quotient space of V modulo W** by $V \setminus W$.

Proof. It suffices to show that S is a subspace of V .

- (a) Since W is a subspace of V , we know that $0_V \in W$. Then $0_V + 0_V = 0_V \in S$.

(b) Let $x, y \in S$. Then using the operations defined in part (c), we write

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

This tells us that $x + y \in S$.

(c) Let $x \in S$. Our goal is to show that $c \in F$ implies $cx \in S$. Using the scalar multiplication in part (c), we get that

$$cx = c(v_1 + W) = cv_1 + W.$$

Thus, S is a vector space. ■

1.3 Linear Combinations

Exercise 1.4.7

In F^n , let e_j denote the vector whose j th coordinate is 1 and whose other coordinates are 0. Prove that $\{e_1, e_2, \dots, e_n\}$ generates F^n .

Proof. Denote the set $V = \{e_1, e_2, \dots, e_n\}$. Our goal is to find scalars $\delta_1, \delta_2, \dots, \delta_n \in F$ such that for all $e_j \in V$ for $1 \leq j \leq n$,

$$\delta_1 e_1 + \delta_2 e_2 + \dots + \delta_n e_n = (a_1, a_2, \dots, a_n). \quad (1)$$

Since $e_j = 1$ for the j th coordinate and the rest of the entries are zeros, we have that

$$\delta_j e_j = \delta_j (0, 0, \underbrace{1}_{j\text{th entry}}, \dots, 0) = (0, 0, \delta_j, \dots, 0)$$

for all $1 \leq j \leq n$. Summing up each term on the left side of (1) and equating each coordinate to the right side of (1), we get that $a_j = \delta_j$ for each $1 \leq j \leq n$. Hence, the set V can span the vector space F^n . ■

Exercise 1.4.8

Show that $P_n(F)$ is generated by $\{1, x, \dots, x^n\}$.

Proof. Let $V = \{1, x, \dots, x^n\}$. We need to find scalars $\delta_1, \delta_2, \dots, \delta_n \in F$ such that

$$\delta_0 + \delta_1 x + \delta_2 x^2 + \dots + \delta_n x^n = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n. \quad (1)$$

Immediately, we have that equating coefficients in (1) gives us $a_i = \delta_i$ for all $1 \leq i \leq n$. Hence, V generates $P_n(F)$. ■

Exercise 1.4.9

Show that the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

generate $M_{2 \times 2}(F)$.

Proof. Define V with the given 2×2 matrices above. We need to find scalars $\delta_i \in F$ for all $1 \leq i \leq 4$ such that

$$\delta_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \delta_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \delta_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \delta_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \quad (1)$$

Distributing each scalar δ_i for all $1 \leq i \leq 4$ for each term in (1), summing up each matrix in (1), and equating each term entry-wise, we have that $a_{11} = \delta_1$, $a_{12} = \delta_2$, $a_{21} = \delta_3$, and $a_{22} = \delta_4$. Hence, V spans $M_{2 \times 2}(F)$. ■

Exercise 1.4.10

Show that if

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \text{ and } M_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

then the span of $\{M_1, M_2, M_3\}$ is the set of all symmetric 2×2 matrices.

Proof. We need to show that the set $\{M_1, M_2, M_3\}$ spans the set of all symmetric 2×2 matrices. We need to find scalars a_1, a_2, a_3 such that

$$a_1 M_1 + a_2 M_2 + a_3 M_3 = \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} = A \quad (1)$$

where A is any 2×2 symmetric matrix. Observe that for $i = j$, we have $\delta_{12} = \delta_{21}$. Performing scalar multiplication on matrices M_1, M_2 , and M_3 , addition of all three terms on (1), and equating entry-wise, we get that $\delta_{12} = \delta_{21} = \delta_3$ and $\delta_{11} = \delta_1$ and $\delta_2 = a_{22}$. Hence, we have

$$\begin{pmatrix} a_1 & a_3 \\ a_3 & a_2 \end{pmatrix}^t = \begin{pmatrix} a_1 & a_3 \\ a_3 & a_2 \end{pmatrix}.$$

Thus, the set $\{M_1, M_2, M_3\}$ spans the set of all symmetric 2×2 matrices. ■

Exercise 1.4.11

Prove that $\text{span}(\{x\}) = \{ax : a \in F\}$ for any vector x in a vector space V . Interpret this result geometrically in \mathbb{R}^3 .

Proof. We need to show that $\text{span}(\{x\}) = \{ax : a \in F\}$, we need to show two containments; that is, $\text{span}(\{x\}) \subseteq \{ax : a \in F\}$ and $\{ax : a \in F\} \subseteq \text{span}(\{x\})$. Let $v \in \text{span}(\{x\})$. Then observe that we can find $\delta \in F$ such that multiplying by $x \in V$ leads $v = \delta x$. But this means that $v \in \{ax : a \in F\}$ by definition. Hence, $\text{span}(\{x\}) \subseteq \{ax : a \in F\}$. Let $v \in \{ax : a \in F\}$. Then $v = ax$ for some $a \in F$. But this is a linear combination of x that makes v . So $v \in \text{span}(\{x\})$ and hence, $\{ax : a \in F\} \subseteq \text{span}(\{x\})$. This result can be viewed as the scaling of vectors in \mathbb{R}^3 . ■

Exercise 1.4.12

Show that a subset W of a vector space V is a subspace of V if and only if $\text{span}(W) = W$.

Proof. (\Rightarrow) Let W be a subspace of V . To show that $\text{span}(W) = W$, we need to show two containments; that is, $\text{span}(W) \subseteq W$ and $W \subseteq \text{span}(W)$. Clearly, W contains itself. Hence, $\text{span}(W) \subseteq W$ Theorem 5. Let $v \in W$. Since W is a subspace, we can find scalars $a_1, a_2, \dots, a_n \in F$ and vectors $w_1, w_2, \dots, w_n \in W$ such that

$$v = a_1x_1 + a_2x_2 + \cdots + a_nx_n \in W$$

using the result in Exercise 1.3.20. But this tells us that $v \in \text{span}(W)$. Hence, $W \subseteq \text{span}(W)$.

(\Leftarrow) Since the span of any subset of W is a subspace and $W = \text{span}(W)$, we have that W is a subspace as well by Theorem 5. ■

Exercise 1.4.13

Show that if S_1 and S_2 are subsets of a vector space V such that $S_1 \subseteq S_2$, then $\text{span}(S_1) \subseteq \text{span}(S_2)$. In particular, if $S_1 \subseteq S_2$ and $\text{span}(S_1) = V$, deduce that $\text{span}(S_2) = V$.

Proof. Let S_1 and S_2 be subsets of a vector space V . Let $v \in \text{span}(S_1)$. We can find scalars $\delta_1, \delta_2, \dots, \delta_n \in F$ and $x_1, x_2, \dots, x_n \in S_1$ such that

$$v = \delta_1x_1 + \delta_2x_2 + \cdots + \delta_nx_n.$$

Since $S_1 \subseteq S_2$, we know that $x_1, x_2, \dots, x_n \in S_2$ so we must have $v \in \text{span}(S_2)$. Hence, $\text{span}(S_1) \subseteq \text{span}(S_2)$.

Now, let $\text{span}(S_1) = V$. We need to show that $\text{span}(S_2) = V$; that is, we need to show $\text{span}(S_2) \subseteq V$ and $V \subseteq \text{span}(S_2)$. By assumption, $S_2 \subseteq V$ and $\text{span}(S_2)$ is a subspace. Clearly, $\text{span}(S_2) \subseteq V$. Since $\text{span}(S_1) = V$ and $\text{span}(S_1) \subseteq \text{span}(S_2)$, we have $V \subseteq \text{span}(S_2)$. Hence, $\text{span}(S_2) = V$. ■

Exercise 1.4.14

Show that if S_1 and S_2 are arbitrary subsets of a vector space V , then $\text{span}(S_1 \cup S_2) = \text{span}(S_1) + \text{span}(S_2)$. (The sum of two subsets is defined in the exercises of Section 1.3.)

Proof. Let S_1 and S_2 be subsets of a vector space V . We need to show $\text{span}(S_1 \cup S_2) = \text{span}(S_1) + \text{span}(S_2)$; that is, $\text{span}(S_1 \cup S_2) \subseteq \text{span}(S_1) + \text{span}(S_2)$ and $\text{span}(S_1) + \text{span}(S_2) \subseteq \text{span}(S_1 \cup S_2)$. Let $v \in \text{span}(S_1 \cup S_2)$. We can find $\delta_1, \delta_2, \dots, \delta_n \in F$ such that $x_1, x_2, \dots, x_n \in S_1 \cup S_2$ implies that

$$v = \delta_1x_1 + \delta_2x_2 + \cdots + \delta_nx_n.$$

Hence, either $x_1, x_2, \dots, x_n \in S_1$ or $x_1, x_2, \dots, x_n \in S_2$. If $x_1, x_2, \dots, x_n \in S_1$, then $v \in \text{span}(S_1)$. Since $\text{span}(S_2)$ is a subspace, we know that $0_V \in \text{span}(S_2)$. Hence, $0_V \in \text{span}(S_1)$ and $v \in \text{span}(S_1)$ imply that $v + 0_V = v \in \text{span}(S_1) + \text{span}(S_2)$. The other case follows a similar process. Hence, $\text{span}(S_1 \cup S_2) \subseteq \text{span}(S_1) + \text{span}(S_2)$.

Let $s \in \text{span}(S_1) + \text{span}(S_2)$. Hence, $s = u + v$ where $u \in \text{span}(S_1)$ and $v \in \text{span}(S_2)$. The former implies that we can find scalars $a_1, a_2, \dots, a_n \in F$ such that $x_1, x_2, \dots, x_n \in S_1$ where

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

and the latter implies that there exists scalars $b_1, b_2, \dots, b_n \in F$ such that $y_1, y_2, \dots, y_n \in S_2$

where

$$v = \sum_{i=1}^n b_i y_i.$$

Since both $x_i \in S_1$ and $y_i \in S_2$ for all $1 \leq i \leq n$, we have $x_i, y_i \in S_1 \cup S_2$ for all $1 \leq i \leq n$. So we must have $s \in \text{span}(S_1 \cup S_2)$. Hence, $\text{span}(S_1) + \text{span}(S_2) \subseteq \text{span}(S_1 \cup S_2)$. ■

Exercise 1.4.15

Let S_1 and S_2 be subsets of a vector space V . Prove that $\text{span}(S_1 \cap S_2) \subseteq \text{span}(S_1) \cap \text{span}(S_2)$. Give an example in which $\text{span}(S_2 \cap S_2)$ and $\text{span}(S_1) \cap \text{span}(S_2)$ are equal and one in which they are not unequal.

Proof. Let S_1 and S_2 be subsets of a vector space V . Let $v \in \text{span}(S_1 \cap S_2)$. Then we can find scalars $a_i \in F$ and vectors $x_i \in S_1 \cap S_2$ for all $1 \leq i \leq n$ such that

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

If $x_i \in S_1 \cap S_2$ for all $1 \leq i \leq n$, then $x_i \in S_1$ and $x_i \in S_2$ for all $1 \leq i \leq n$. This implies that $v \in \text{span}(S_1)$ and $v \in \text{span}(S_2)$. Hence, $v \in \text{span}(S_1) \cap \text{span}(S_2)$. Thus, we conclude that $\text{span}(S_1 \cap S_2) \subseteq \text{span}(S_1) \cap \text{span}(S_2)$. ■

Example 1.3.1. Define S_1 as the set

$$\{(1, 1, 0) \in \mathbb{R}^3\}$$

and S_2 as the set

$$\{(1, 1, 0), (1, 0, 1), (0, 1, 1) \in \mathbb{R}^3\}.$$

Observe that $S_1 \cap S_2 = \{(1, 1, 0)\}$ and thus the of this set yields $\{a(1, 1, 0) : a \in F \text{ and } (1, 1, 0) \in \mathbb{R}^3\}$. The span of S_1 yields the following set $\{a(1, 1, 0) : a \in F \text{ and } (1, 1, 0) \in \mathbb{R}^3\}$ and the span of S_2 yields the following set

$$\{a(1, 1, 0) + b(1, 0, 1) + c(0, 1, 1) : a, b, c \in F \text{ and } (1, 1, 0), (1, 0, 1), (0, 1, 1) \in \mathbb{R}^3\}.$$

Note that $\text{span}(S_1 \cap S_2) \neq \text{span}(S_1) \cap \text{span}(S_2)$ since $\text{span}(S_1) \cap \text{span}(S_2) \neq \emptyset$.

Now define $S_1 = \{O_V\} = S_2$. Clearly, $S_1 \cap S_2 = \{O_V\}$, $\text{span}(S_1) \cap \text{span}(S_2) = \{O_V\}$, and that $\text{span}(S_1 \cap S_2)$ is also equal to this set. $\text{span}(S_1 \cap S_2) = \text{span}(S_1) \cap \text{span}(S_2)$

Exercise 1.4.16

Let V be a vector space and S a subset of V with the property that whenever $v_1, v_2, \dots, v_n \in S$ and $a_1 v_1 + a_2 v_2 + \dots + a_n v_n = 0$, then $a_1 = a_2 = \dots = a_n = 0$. Prove that every vector in the span of S can be *uniquely* written as a linear combination of vectors of S .

Proof. Let $x \in \text{span}(S)$. Suppose there exists two sets of scalars $a_1, a_2, \dots, a_n \in F$ and $b_1, b_2, \dots, b_n \in F$ such that whenever $v_1, v_2, \dots, v_n \in S$ such that

$$a_1 v_1 + a_2 v_2 + \dots + a_n v_n = x \tag{1}$$

and

$$b_1 v_1 + b_2 v_2 + \dots + b_n v_n = x \tag{2}$$

implies that a_i . Setting (1) and (2) together, we have

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = b_1v_1 + b_2v_2 + \cdots + b_nv_n \quad (3)$$

which implies that

$$(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n = 0.$$

Since $v_1, v_2, \dots, v_n \in S$, we have that $a_i - b_i = 0$ and thus $a_i = b_i$ for all $1 \leq i \leq n$. Hence, every vector in the span of S can be *uniquely* written as a linear combination of vectors of S . ■

1.4 Linear Dependence and Linear Independence

Exercise 1.5.1

Label the following statements as true or false.

- (a) If S is a linearly dependent set, then each vector in S is a linear combination of other vector in S .

Proof. True ■

- (b) Any set containing the zero vector is linearly dependent.

Proof. True ■

- (c) The empty set is linearly dependent.

Proof. False. It is linearly independent. ■

- (d) Subsets of linearly dependent sets are linearly dependent.

Proof. This is **False**. We can have a linearly independent subset of a set that is linearly dependent. ■

- (e) Subsets of linearly independent sets are linearly independent.

Proof. True by corollary to Theorem 6. ■

- (f) If $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ and x_1, x_2, \dots, x_n are linearly independent, then all the scalars a_i are zero.

Proof. True this is by definition. ■

Exercise 1.5.4

In F^n , let e_j denote the vector whose j th coordinate is 1 and whose other coordinates are 0. Prove that $\{e_1, e_2, \dots, e_n\}$ is linearly independent.

Proof. Choose a finite amount of scalars $a_1, a_2, \dots, a_n \in F$ to create the following linear combination:

$$a_1e_1 + a_2e_2 + \cdots + a_ne_n = (0, 0, \dots, 0). \quad (1)$$

To show that the set $\{e_1, e_2, \dots, e_n\}$ is linearly independent, we need to show that the scalars $a_1, a_2, \dots, a_n \in F$ have the trivial representation; that is, $a_1 = a_2 = \cdots = a_n = 0$.

Since the j th coordinate of e_j is 1 but 0 in all the other entries, we have that

$$\begin{aligned} & a_1(1, 0, \dots, 0) + a_2(0, 1, \dots, 0) + \dots + a_n(0, 0, \dots, 1) \\ &= (a_1, 0, \dots, 0) + (0, a_2, \dots, 0) + \dots + (0, 0, \dots, a_n) \\ &= (a_1, a_2, \dots, a_n). \end{aligned}$$

Hence, we have

$$(a_1, a_2, \dots, a_n) = (0, 0, \dots, 0).$$

Equating each entry of the left side of the equation above to 0, we find that $a_i = 0$ for all $1 \leq j \leq n$. Hence, the set $\{e_1, e_2, \dots, e_n\}$ is linearly independent. ■

Exercise 1.5.4

Show that the set $\{1, x, x^2, \dots, x^n\}$ is linearly independent in $P_n(F)$.

Proof. Just like the prior exercise, we need to show that we can find scalars $a_0, a_1, \dots, a_n \in F$ such that

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n = 0$$

where $a_i = 0$ for all $0 \leq i \leq n$. Note that the 0 polynomial is just

$$0 + 0x + 0x^2 + \dots + 0x^n = 0.$$

Hence, equating coefficients we immediately get that $a_i = 0$ for all $0 \leq i \leq n$. Thus, the set $\{1, x, x^2, \dots, x^n\}$ is linearly independent. ■

Exercise 1.5.6

In $M_{m \times n}(F)$, let E^{ij} denote the matrix whose only nonzero entry is 1 in the i th row and j th column. Prove that $\{E^{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ is linearly independent.

Proof. First, we create a linear combination of a finite amount vectors in $E = \{E^{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ with scalars δ_k for $1 \leq k \leq N$ with $N = mn$ as the number of total entries in each matrix in $\{E^{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$. Note that after doing our scalar multiplication and summing up each term, we find that each $\delta_k E_{ij} = \delta_k$ in our linear combination can equated with a corresponding i and j entry in the zero matrix such that $\delta_k = 0$ for all $1 \leq k \leq N$. Hence, E is a linearly independent set. ■

Exercise 1.5.7

Recall from Example 3 in Section 1.3 that the set of diagonal matrices in $M_{2 \times 2}(F)$ is a subspace. Find a linearly independent set that generates this subspace.

Proof. Define W as the linearly independent spanning set of the set of diagonal matrices in $M_{2 \times 2}$ where

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

To see why W is a linearly independent set, choose scalars $\delta_1, \delta_2 \in F$ such that

$$\delta_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \delta_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Performing scalar multiplication and vector addition gives us the following equation

$$\begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Since the zero matrix is a diagonal matrix, we know that equation entries where $i = j$ yields $\delta_1 = \delta_2 = 0$. Hence, W is a linearly independent set that generates the set of diagonal matrices of $M_{2 \times 2}(F)$. ■

Exercise 1.5.8

Let $S = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ be a subset of the vector space F^3 .

- (a) Prove that if $F = \mathbb{R}$, then S is linearly independent.

Proof. ■

- (b) Prove that if F has characteristic two, then S is linearly dependent.

Proof. ■

Exercise 1.5.9

Let u and v be distinct vectors in a vector space V . Show that $\{u, v\}$ is linearly dependent if and only if u or v is a multiple of the other. I have written two proofs for this:

Proof. Let u and v be distinct vectors in a vector space V .

(\Rightarrow) Since $\{u, v\}$ is a linearly dependent set, we can find scalars $a_1, a_2 \in F$ such that

$$a_1 u + a_2 v = 0 \tag{1}$$

Suppose v is not a multiple of u and choose $a_1 \neq 0$ since $\{u, v\}$ is linearly dependent. We need to show that u is a multiple of v . Solving for u , we get that

$$u = -\frac{a_2}{a_1} v.$$

Hence, u is a multiple of v .

(\Leftarrow) Suppose u or v is a scalar multiple of the other. Assume u is the scalar multiple of v . Then for some $c \neq 0 \in F$, we have $u = cv$. Hence, we have $u - cv = 1u - cv = 0$. This tells us that $\{u, v\}$ is linearly dependent. ■

Exercise 1.5.12

Prove Theorem 1.6 and its corollary.

Proof. See proof in notes. ■

Exercise 1.5.13

Let V be a vector space over a field of characteristic not equal to two.

- (a) Let u and v be distinct vectors in V . Prove that $\{u, v\}$ is linearly independent if and only if $\{u + v, u - v\}$ is linearly independent.

Proof. Let u and v be distinct vectors in V .

For the forwards direction, assume $\{u, v\}$ is a linearly independent set. We need to show that $\{u + v, u - v\}$ is linearly independent. Hence, we need to find $a, b \in F$ such that

$$a(u + v) + b(u - v) = 0. \quad (1)$$

Note that (1) leads to

$$\begin{aligned} a(u + v) + b(u - v) &= au + av + bu - bv \\ &= au - bv + av + bu. \end{aligned}$$

Since $\{u, v\}$ is a linearly independent set, we have that

$$au - bv = 0$$

and

$$av + bu = 0$$

for $a = b = 0$. Hence,

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

for $a = b = 0$ and so $\{u - v, u + v\}$ is a linearly independent set.

For the backwards direction, suppose $\{u + v, u - v\}$ is linearly independent. We need to show that $\{u, v\}$ is linearly independent. Note that $a, b \in F$ such that

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

for $a = b = 0$ since $\{u - v, u + v\}$ is linearly independent. Note that

$$\begin{aligned} a(u + v) + b(u - v) &= au + av + bu - bv \\ &= au - bv + av + bu \\ &= 0 + av + bu \\ &= 0. \end{aligned}$$

Thus, $av + bu = 0$ where a, b both zero. Thus, the set $\{u, v\}$ is linearly independent. ■

- (b) Let u, v , and w be distinct vectors in V . Prove that $\{u, v, w\}$ is linearly independent if and only if $\{u + v, u + w, v + w\}$ is linearly independent.

Proof. For the forwards direction, suppose $\{u, v, w\}$ is linearly independent. Then choose scalars $a_1, a_2, a_3 \in F$ such that

$$a_1u + a_2v + a_3w = 0$$

with $a_1 = a_2 = a_3 = 0$. We need to show that $\{u + v, u + w, v + w\}$ is linearly independent; that is, we need to show that we can find scalars $a_1, a_2, a_3 \in F$ such that

$$a_1(u + v) + a_2(u + w) + a_3(v + w) = 0 \quad (1)$$

for $a_1 = a_2 = a_3 = 0$. Observe that (1) can be written in the following way

$$(a_1u + a_3v + a_2w) + (a_1v + a_2u + a_3w) = 0 \quad (2)$$

Since $\{u, v, w\}$ is linearly independent, we know that $a_1 = a_2 = a_3 = 0$. But this also has to mean that $\{u + v, u + w, v + w\}$ is a linearly independent set.

For the backwards direction, suppose $\{u + v, u + w, v + w\}$ is linearly independent. Then choose scalars $a_1, a_2, a_3 \in F$ such that

$$a_1(u + v) + a_2(u + w) + a_3(v + w) = 0. \quad (1)$$

We need to show that $\{u, v, w\}$ is linearly independent. Observe that (1) can be re-written as

$$\begin{aligned} (a_1u + a_3v + a_2w) + (a_1v + a_2u + a_3w) &= 0 \\ \Rightarrow 0 + (a_1v + a_2u + a_3w) &= 0 \\ \Rightarrow a_1v + a_2u + a_3w &= 0 \end{aligned}$$

where $a_1 = a_2 = a_3 = 0$. Hence, $\{u, v, w\}$ is linearly independent. ■

Exercise 1.5.14

Prove that a set S is linearly dependent if and only if $S = \{0\}$ or there exists distinct vectors $v, u_1, u_2, \dots, u_n \in S$ such that v is a linear combination of u_1, u_2, \dots, u_n .

Proof. For the forwards direction, Let S be a linearly dependent. Then we need to show that either $S = \{0\}$ or S contains distinct vectors $v, u_1, u_2, \dots, u_n \in S$ such that $v \in \text{span}(S)$. Suppose there does not exist distinct vectors $v, u_1, u_2, \dots, u_n \in S$ such that $v \in \text{span}(S)$. This tells us that S only contains the singleton $S = \{v\}$. Furthermore, we must require $v = 0$ since S is linearly dependent. Otherwise, $v \neq 0$ would imply that S is linearly independent. Hence, we have $S = \{0\}$. Now suppose $S \neq \{0\}$. Since S is linearly dependent, there exists scalars a_1, a_2, \dots, a_{n+1} and vectors v, u_1, u_2, \dots, u_n such that

$$a_1v + a_2u_1 + a_3u_2 + \dots + a_{n+1}u_n = 0$$

with all $a_1, a_2, \dots, a_n, a_{n+1}$ not all equal to zero. Solving for v , we get that

$$v = -(a_1^{-1}a_2)u_1 - (a_1^{-1}a_3)u_2 - \dots - (a_1^{-1}a_{n+1})u_n.$$

Since v is a linear combination of u_1, u_2, \dots, u_n , we have that $v \in \text{span}(S)$.

Conversely, suppose that either $S = \{0\}$ or there exists distinct vectors $v, u_1, u_2, \dots, u_n \in S$ such that v is a linear combination of u_1, u_2, \dots, u_n . Assume $S = \{0\}$. Then S is linearly dependent because the singleton is the zero vector. Now suppose v is a linear combination of vectors u_1, u_2, \dots, u_n . Then there exists scalars $a_1, a_2, \dots, a_n \in F$ and distinct vectors $u_1, u_2, \dots, u_n \in S$ such that

$$v = a_1u_1 + a_2u_2 + \dots + a_nu_n.$$

Subtracting v from both sides of this equation yields the following equation

$$a_1u_1 + a_2u_2 + \dots + a_nu_n - 1v = 0.$$

Since not all scalars in the equation above are zero and $v, u_1, u_2, \dots, u_n \in S$, we must have that S is a linearly dependent set. ■

Exercise 1.5.15

Prove that a set $S = \{u_1, u_2, \dots, u_n\}$ be a finite set of vectors. Prove that S is linearly dependent if and only if $u_1 = 0$ or $u_{k+1} \in \text{span}(\{u_1, u_2, \dots, u_k\})$ for some k where $1 \leq k < n$.

Proof. Suppose $u_{k+1} \notin \text{span}(\{u_1, u_2, \dots, u_k\})$ for all $1 \leq k < n$. Since S is linearly dependent, we know that the zero vector is contained in S . Choose $k = 1$ such that $u_1 = 0$ and we are done. On the other hand, suppose $u_1 \neq 0$. We need to show that $u_{k+1} \in \text{span}(\{u_1, u_2, \dots, u_k\})$. Choose $k = n - 1$. Then clearly $n = k + 1$. Since S is linearly independent, choose scalars $a_1, a_2, \dots, a_n \in F$ such that

$$a_1 u_1 + a_2 u_2 + \dots + a_k u_k + a_{k+1} u_{k+1} = 0. \quad (1)$$

where $a_1, a_2, \dots, a_k, a_{k+1}$ not all zero. Solving for u_{k+1} by subtracting $a_{k+1} u_{k+1}$ on both sides of (1) and multiplying $-a_{k+1}^{-1}$ on both sides of (1), we end up with the following equation:

$$u_{k+1} = -(a_{k+1}^{-1} a_1) u_1 - (a_{k+1}^{-1} a_2) u_2 - \dots - (a_{k+1}^{-1} a_k) u_k.$$

This tells us that u_{k+1} can be written as a linear combination of vectors u_1, u_2, \dots, u_k . Hence, $u_{k+1} \in \text{span}(\{u_1, u_2, \dots, u_k\})$.

Conversely, either $u_1 = 0$ or $u_{k+1} \in \text{span}(u_1, u_2, \dots, u_k)$. Suppose $u_1 = 0$. Then S contains the zero vector so S must be linearly dependent. On the other hand, choose $k = n - 1$ where $1 \leq k < n$ such that $u_{k+1} \in \text{span}(\{u_1, u_2, \dots, u_k\})$ implies that there exists scalars a_1, a_2, \dots, a_k such that

$$\begin{aligned} u_{k+1} &= a_1 u_1 + a_2 u_2 + \dots + a_k u_k \\ \Rightarrow u_n &= a_1 u_1 + a_2 u_2 + \dots + a_{n-1} u_{n-1}. \end{aligned} \quad (1)$$

Subtracting u_n on both sides of (1) implies that

$$(a_1 u_1 + a_2 u_2 + \dots + a_{n-1} u_{n-1}) - 1 u_n = 0.$$

Since not all scalars in the linear combination above are zero, we know that S must be linearly dependent. ■

Exercise 1.5.16

Prove that a set S of vectors is linearly independent if and only if each finite subset of S is linearly independent.

Proof. (\Rightarrow) Suppose S is a linearly independent set. Let S' be any finite subset of S . By corollary to Theorem 6, we can see that $S' \subseteq S$ implies that S' is also linearly independent. (\Leftarrow) We will proceed by proving the contrapositive. Let $S' \subseteq S$ be a finite subset that is linearly dependent set. We will prove that S is a linearly dependent set. Since $S' \subseteq S$, we have that S must be a linearly dependent set by Theorem 6. ■

Exercise 1.5.17

Let M be a square upper triangular matrix (as defined in Section 1.3) with nonzero diagonal entries. Prove that the columns of M are linearly independent.

Proof. Let M be a square upper triangular matrix. Note that M have the following form:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \ddots & a_{n-1n} \\ 0 & 0 & \cdots & \cdots & a_{nn} \end{pmatrix}.$$

We can define the following columns as vectors v_1, v_2, \dots, v_n where

$$v_1 = \begin{pmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ 0 \end{pmatrix}, \dots, v_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}.$$

Choose scalars $\delta_1, \delta_2, \dots, \delta_n \in F$ such that

$$\delta_1 v_1 + \delta_2 v_2 + \cdots + \delta_n v_n = 0$$

where the zero vector 0 is denoted as

$$0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Using scalar multiplication and addition, we arrive at the following system of equations:

$$\delta_1 a_{11} + \delta_2 a_{12} + \cdots + \delta_{n-1} a_{1n-1} + \delta_n a_{1n} = 0$$

$$\delta_2 a_{22} + \delta_3 a_{23} + \cdots + \delta_n a_{2n} = 0$$

$$\vdots$$

$$\delta_{n-1} a_{n-1n-1} + \delta_n a_{n-1n} = 0$$

$$\delta_n a_{nn} = 0.$$

Since all the diagonal entries of M are non-zero, we can see from the equation above that $\delta_n = 0$ which subsequently tells us that $\delta_{n-1} = 0$. We claim that $\delta_1 = \cdots = \delta_{n-1} = \delta_n = 0$ for all $n \geq 1$. We can prove this via induction. Let our base case be $n = 1$. Then we have $\delta_1 a_{11} = 0$ with $a_{11} \neq 0$ implies $\delta_1 = 0$. Now let $n = 2$. Then observe that we have an upper triangular 2×2 matrix such that the linear combination of the columns lead to the following system of linear equations:

$$\delta_1 a_{11} + \delta_2 a_{12} = 0$$

$$0 + \delta_2 a_{22} = 0.$$

Observe that $\delta_2 = 0$ which also implies that $\delta_1 = 0$. Now suppose our claim that $\delta_1 = \cdots = \delta_{n-1} = \delta_n = 0$ holds for all $n \geq 1$. We want to show that our claim still holds for the $n + 1$ case. Observe that $\delta_{n+1} = 0$ derived from an $(n + 1) \times (n + 1)$ matrix. Using the same process that proved the base case, we find that $\delta_n = \delta_{n+1} = 0$. By our inductive hypothesis, we know that $\delta_1 = \cdots = \delta_{n-1} = \delta_n = 0$ for all $n \geq 1$. This tells us that $\delta_1 = \cdots = \delta_{n+1}$ for all $n \geq 1$. Hence, the columns of M are linearly independent. ■

Exercise 1.5.18

Let S be a set of nonzero polynomials in $P(F)$ such that no two have the same degree. Prove that S is linearly independent.

Proof. Let $0 \leq n \leq k$ such that $S = \{p_0(x), p_1(x), p_2(x), \dots, p_k(x)\}$ where every $p_n(x)$ non-zero such that no two polynomials in this set have the same degree. Define $p_n(x) = x^n + x^{n+1} + \dots + x^k$. Choose scalars $a_1, a_2, \dots, a_k \in F$ such that

$$a_1 p_1(x) + a_2 p_2(x) + \dots + a_k p_k(x) = 0 \quad (1)$$

We need to show that $\delta_1 = \delta_2 = \dots = \delta_k = 0$. Observe that (1) implies that

$$a_0 + (a_0 + a_1 + a_2)x^2 + \dots + (a_0 + a_1 + \dots + a_k)x^k = 0. \quad (2)$$

Setting each x^n to both sides of the equation above leads to the following system of linear equations:

$$\begin{aligned} a_0 &= 0 \\ a_0 + a_1 &= 0 \\ a_0 + a_1 + a_2 &= 0 \\ &\vdots \\ a_0 + a_1 + a_2 + \dots + a_k &= 0. \end{aligned}$$

It can be proved via induction that $a_1 = a_2 = \dots = a_k = 0$ for all $1 \leq n \leq k$. Hence, S is a linearly independent set. ■

Exercise 1.5.19

Prove that if $\{A_1, A_2, \dots, A_k\}$ is a linearly independent subset of $M_{n \times n}(F)$, then $\{A_1^t, A_2^t, A_3^t, \dots, A_k^t\}$ is also linearly independent.

Proof. Suppose $\{A_1, A_2, \dots, A_k\}$ is a linearly independent set. Choose a finite set of scalars $\delta_1, \delta_2, \dots, \delta_k \in F$ such that

$$\delta_1 A_1 + \delta_2 A_2 + \dots + \delta_k A_k = 0$$

with $\delta_1 = \delta_2 = \dots = \delta_k = 0$. Apply the transpose to both sides, we know that $0^t = 0$ and $(\delta_i A_i)^t = \delta_i (A_i)^t$ for all $1 \leq i \leq k$. Hence, we have

$$\begin{aligned} (\delta_1 A_1 + \delta_2 A_2 + \dots + \delta_k A_k)^t &= 0^t \\ \Rightarrow (\delta_1 A_1)^t + (\delta_2 A_2)^t + \dots + (\delta_k A_k)^t &= 0 \\ \Rightarrow \delta_1 A_1^t + \delta_2 A_2^t + \dots + \delta_k A_k^t &= 0. \end{aligned}$$

Hence, the set $\{A_1^t, A_2^t, \dots, A_k^t\}$ is also linearly independent. ■

Exercise 1.5.20

Let $f, g \in \mathcal{F}(\mathbb{R}, \mathbb{R})$ be the functions defined by $f(t) = e^{rt}$ and $g(t) = e^{st}$, where $r \neq s$. Prove that f and g are linearly independent in $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Proof. Let $f, g \in \mathcal{F}(\mathbb{R}, \mathbb{R})$ where f and g are defined by $f(t) = e^{rt}$ and $g(t) = e^{st}$ with

$r \neq s$. Suppose for sake of contradiction that the set $\{f, g\}$ is a linearly dependent set of vectors. By Exercise 1.5.9, we know that either f is a multiple of g or g is a multiple of f . Assume f is a multiple of g . Hence, there exists a $c \in R$ such that $f(t) = cg(t)$. In other words,

$$f(t) = cg(t) \Rightarrow e^{rt} = ce^{st}.$$

To solve for c , let $t = 0$. Then we have

$$f(0) = cg(0) \Rightarrow 1 = c.$$

Hence, we have

$$e^{rt} = e^{st}.$$

This equality is valid only when $r = s$ because otherwise f would not be a scalar multiple of g . But note that $r \neq s$ by assumption. Hence, we have a contradiction and thus $\{f, g\}$ must be linearly independent. ■

Exercise 1.5.21

Let S_1 and S_2 be disjoint linearly independent subsets of V . Prove that $S_1 \cup S_2$ is linearly dependent if and only if $\text{span}(S_1) \cap \text{span}(S_2) \neq \{0\}$.

Proof. (\Rightarrow) We will proceed via contrapositive. Since $\text{span}(S_1) \cap \text{span}(S_2) = \{0\}$, let $v \in \text{span}(S_1)$ such that we can find a $w \in \text{span}(S_2)$ such that both $v = w$ where both $v = 0$ and $w = 0$. Since $v \in \text{span}(S_1)$, we can find a finite set of vectors $x_1, x_2, \dots, x_n \in S_1$ and scalars $a_1, a_2, \dots, a_n \in F$ such that

$$v = \sum_{i=1}^n a_i x_i = 0.$$

Likewise, $w \in \text{span}(S_2)$ implies that we can find $y_1, y_2, \dots, y_n \in S_2$ and scalars $b_1, b_2, \dots, b_n \in F$ such that

$$w = \sum_{j=1}^n b_j y_j = 0.$$

Observe that

$$v = w \Rightarrow \sum_{i=1}^n a_i x_i = \sum_{j=1}^n b_j y_j.$$

Then we have

$$\sum_{i=1}^n a_i x_i - \sum_{j=1}^n b_j y_j = 0.$$

Since S_1 and S_2 are disjoint linearly independent sets, we know that $x_i \notin S_2$ and $y_j \notin S_1$ and that $a_i = 0$ and $b_j = 0$ for all $1 \leq i \leq n$ and $1 \leq j \leq n$ respectively. Hence, $S_1 \cup S_2$ is a linearly independent set.

(\Leftarrow) We will proceed via contrapositive for this direction as well. Suppose $S_1 \cup S_2$ is linearly independent. Then choose a finite number of distinct vectors $x_1, x_2, \dots, x_n \in S_1 \cup S_2$ and scalars $a_1, a_2, \dots, a_n \in F$ such that

$$\underbrace{\sum_{i=1}^n a_i x_i}_{\in \text{span}(S_1)} = 0$$

with $a_i = 0$ for all $1 \leq i \leq m$. Since $x_i \in S_1 \cup S_2$, then either $x_i \in S_1$ or $x_i \in S_2$. Without loss of generality, suppose $x_i \in S_1$. Then we know that $x_i \notin S_2$ since $S_1 \cap S_2 = \emptyset$. Since S_2 is linearly independent, choose a finite number of vectors $y_1, y_2, \dots, y_m \in S_2$ and scalars $b_1, b_2, \dots, b_m \in F$ such that

$$\sum_{\substack{j=1 \\ \in \text{span}(S_2)}}^m b_j y_j = 0$$

with $b_j = 0$ for all $1 \leq j \leq m$. Observe that

$$\sum_{i=1}^n a_i x_i = \sum_{j=1}^m b_j y_j = 0.$$

Hence, we have $\text{span}(S_1) \cap \text{span}(S_2) = \{0\}$. ■

1.5 Bases and Dimension

Exercise 1.6.11

Let u and v be distinct vectors of a vector space V . Show that if $\{u, v\}$ is a basis for V and a and b are nonzero scalars, then both $\{u + v, au\}$ and $\{au, bv\}$ are also bases for V .

Proof. (\Rightarrow) We want to show that $\{u + v, au\}$ and $\{au, bv\}$ is a basis for V ; that is, we want to show that $\{u + v, au\}$ and $\{au, bv\}$ is both linearly independent and generates V . We will start by showing that $\{u + v, au\}$ is linearly independent. Choose scalars δ_1, δ_2 such that

$$\delta_1(u + v) + \delta_2(au) = 0 \tag{1}$$

with $\delta_1 = \delta_2 = 0$. Let us algebraically manipulate (1) into the following form:

$$\delta_1 u + \delta_1 v + (\delta_2 a)u = 0.$$

Since $\{u, v\}$ is linearly independent and $a \neq 0$, we get that

$$\delta_1 v + (\delta_2 a)u = 0$$

implies $\delta_1 = 0$ and $\delta_2 a = 0$ such that $\delta_2 = 0$. But this implies that $\{u + v, au\}$ is also linearly independent. To show that $\{u + v, au\}$ spans V , it suffices to show that $V \subseteq \text{span}\{u + v, au\}$ since the other containment $\text{span}\{u + v, au\} \subseteq V$ follows immediately. Let $v \in V$. By Theorem 1.7, we know that adjoining an arbitrary vector $w \in V$ but not in $\text{span}\{u + v, au\}$ creates a linearly dependent set. So, we must have $w \in \text{span}\{u + v, au\}$ and thus $V \subseteq \text{span}\{u + v, au\}$.

Now, we want to show that $\{au, bv\}$ is a basis. Choose scalars δ_1, δ_2 such that

$$\delta_1(au) + \delta_2(bv) = 0 \tag{2}$$

such that $\delta_1 = \delta_2 = 0$. We can manipulate (2) by rewriting it in the following form:

$$(\delta_1 a)u + (\delta_2 b)v = 0. \tag{3}$$

Since $\{u, v\}$ is a linearly independent set, we know that $\delta_1 a = \delta_2 b = 0$. Since $a, b \neq 0$, this implies that $\delta_1 = \delta_2 = 0$. Hence, the representation in (2) is trivial and thus the

set $\{au, bv\}$ is linearly independent. Since adjoining any $w \in V$ not in $\{au, bv\}$ creates a linearly dependent set, we get that $w \in \text{span}(\{au, bv\})$ by Theorem 1.7. Hence, $\{au, bv\}$ generates V . ■

Exercise 1.6.12

Let u, v , and w be distinct vectors of a vector space V . Show that if $\{u, v, w\}$ is a basis for V , then $\{u + v + w, v + w, w\}$ is also a basis for V .

Proof. First, we prove that $\{u + v + w, v + w, w\}$ is linearly independent. Choose $\delta_1, \delta_2, \delta_3 \in F$ such that

$$\delta_1(u + v + w) + \delta_2(v + w) + \delta_3w = 0. \quad (1)$$

We can rewrite (1) in the following way:

$$(\delta_1u + \delta_2v + \delta_3w) + \delta_1(v + w) + \delta_2w = 0.. \quad (2)$$

Since $\{u, v, w\}$ is also a basis, we know that $\{u, v, w\}$ is also linearly independent. Hence, $\delta_1 = \delta_2 = \delta_3 = 0$. Thus, (1) contains the trivial representation and so $\{u + v + w, v + w, w\}$ is linearly independent.

Now, to prove that $S = \{u + v + w, v + w, w\}$ generates V , it suffices to show that $V \subseteq \text{span}(S)$. Adjoining a vector $x \in V$ but not in S produces a linearly independent set. Hence, Theorem 1.7 implies that $s \in \text{span}(S)$. Hence, S generates V and that S is a basis for V . ■

Exercise 1.6.19

Complete the proof of Theorem 1.8.

Proof. See proof in notes. ■

Exercise 1.6.20

Let V be a vector space having dimension n , and let S be a subset of V that generates V .

- (a) Prove that there is a subset of S that is a basis for V . (Be careful not to assume that S is finite.)

Proof. Let $\dim(V) = n$. Suppose S is a subset of V such that S generates V . Then S could be either $S = \{0\}$ or $S = \emptyset$. In either case, we find that $\text{span}(\emptyset) = \{0\} = V$ or $\text{span}(\{0\}) = \{0\} = V$. Now, suppose S contains a non-zero vector u_1 . Thus, the set $\{u_1\}$ is linearly independent. Suppose we continue adding vectors inductively u_2, u_3, \dots, u_k into this set such that this process stops at exactly k vectors. We claim that our constructed set

$$L = \{u_1, u_2, \dots, u_k\}$$

is linearly independent for $k \geq 1$. Suppose we assume that L holds for the k th case. We want to show that it also holds for the $k + 1$ case. Observe that

$$L = \{u_1, u_2, \dots, u_{k+1}\} = \{u_1, u_2, \dots, u_k\} \cup \{u_{k+1}\}.$$

If $u_{k+1} = 0$, then L would be linearly dependent. Otherwise $u_{k+1} \neq 0$ and so $\{u_{k+1}\}$ is linearly independent. By inductive hypothesis, we also know that $\{u_1, u_2, \dots, u_k\}$ is

linearly independent. Since $\text{span}(\{u_1, u_2, \dots, u_k\}) \cap \text{span}(\{u_{k+1}\}) = \{0\}$ and that the two sets are disjoint, we know that $\{u_1, u_2, \dots, u_k\} \cup \{u_{k+1}\}$ is linearly independent. This ends our induction proof.

Note that we cannot have $S \neq L$ since S could be an infinite set. Since L is a subset of V where $\dim(V) = n$, L can be extended into a basis for V by Corollary 2 of the Replacement Theorem that contains exactly n vectors. ■

(b) Prove that S contains at least n vectors.

Proof. Denote the basis constructed from part (a) as β . Since β is a basis for V , β must contain exactly n vectors. Since $\beta \subseteq S$ and S is a generating set for V , then S must contain at least n vectors. ■

Exercise 1.6.21

Prove that a vectors space is infinite-dimensional if and only if it contains an infinite linearly independent subset.

Proof. Let V be a vector space. For the forwards direction, suppose V is an infinite-dimensional vector space. By definition, V contains a basis β that is infinite-dimensional. By definition, β is also linearly independent. Thus, V contains an infinite linearly independent set.

For the backwards direction, we proceed using the converse. Suppose V is a finite-dimensional vector space. Let $\dim(V) = n$. By definition, V contains a basis β that contains exactly n vectors. Since β is also linearly independent, β is a finite linearly independent subset. ■

Exercise 1.6.22

Let W_1 and W_2 be subspaces of a finite-dimensional vector space V . Determine the necessary and sufficient conditions on W_1 and W_2 so that $\dim(W_1 \cap W_2) = \dim(W_1)$.

Proof. We must have $W_1 \subseteq W_2$ in order for $\dim(W_1 \cap W_2) = \dim(W_1)$. Let W_1 and W_2 be subspaces of a finite dimensional vector space V . Since W_1 and W_2 are subspaces, we must also have $W_1 \cap W_2$ as a subspace. Hence, $W_1 \cap W_2$ is finite-dimensional by Theorem 1.11. This implies that $W_1 \cap W_2$ contains a basis β containing exactly $\dim(W_1 \cap W_2)$ vectors. Since β is a linearly independent subset of W_1 , we know that β must contain at most $\dim(W_1)$ vectors. Hence, we have $\dim(W_1 \cap W_2) \leq \dim(W_1)$. Since $W_1 \cup W_2$, then $W_1 \subseteq W_1 \cap W_2$. Since W_1 is finite-dimensional, let α be a basis containing exactly $\dim(W_1)$ vectors. Since $\alpha \subseteq W_1 \cap W_2$ and α is a linearly independent set, α must contain at most $\dim(W_1 \cap W_2)$ amount of vectors. Hence, $\dim(W_1) \leq \dim(W_1 \cap W_2)$. Thus, we have $\dim(W_1) = \dim(W_1 \cap W_2)$.

Conversely, we have $\dim(W_1 \cap W_2) = \dim(W_1)$. By Theorem 1.11, we have $W_1 \cap W_2 = W_1$. Since $W_1 \cap W_2 \subseteq W_2$, we know that $W_1 \subseteq W_2$. ■

Exercise 1.6.23

Let v_1, v_2, \dots, v_k, v be vectors in a vector space V , and define $W_1 = \text{span}(\{v_1, v_2, \dots, v_k\})$, and $W_2 = \text{span}(\{v_1, v_2, \dots, v_k, v\})$.

(a) Find necessary and sufficient conditions on v such that $\dim(W_1) = \dim(W_2)$.

Proof. The condition we need is $v \in W_1$. Since W_1 and W_2 are subspaces, we also have $W_1 \cap W_2$ is a subspace. Hence, theorem 1.11 tells us that $W_1 \cap W_2$ is also finite-dimensional. Suppose $v \in W_1$. Since $v \in W_2$ as well, we have that $W_1 \subseteq W_2$. Now let $v \in W_2$. Then choose scalars a_1, a_2, \dots, a_k such that

$$a_1 v_1 + a_2 v_2 + \dots + a_k v_k = v.$$

But this tells us that $v \in W_1$. So, $W_2 \subseteq W_1$ and thus $W_1 = W_2$. By theorem 1.11, $\dim(W_1) = \dim(W_2)$.

Conversely, $\dim(W_1) = \dim(W_2)$. Since $v \in W_2$, this also means that $v \in W_1$ since $W_1 = W_2$ by theorem 1.11. ■

- (b) State and prove a relationship involving $\dim(W_1)$ and $\dim(W_2)$ in the case that $\dim(W_1) \neq \dim(W_2)$.

Proof. If $\dim(W_1) \neq \dim(W_2)$, then $v \notin W_1$. This is just the contrapositive of the statement above. ■

Exercise 1.6.24

Let $f(x)$ be a polynomial of degree n in $P_n(\mathbb{R})$. Prove that for any $g(x) \in P_n(\mathbb{R})$ there exists scalars c_0, c_1, \dots, c_n such that

$$g(x) = c_0 f(x) + c_1 f^{(1)}(x) + c_2 f^{(2)}(x) + \dots + c_n f^{(n)}(x),$$

where $f^{(n)}(x)$ denotes the n th derivative of $f(x)$.

Proof. Since f is differentiable n times, we can construct the set

$$W = \{f(x), f^{(1)}(x), f^{(2)}(x), \dots, f^{(n)}(x)\}$$

containing $n + 1$ polynomials such that no two polynomials contain the same degree (with each derivative of $f(x)$, the degree decreases by one). Since W is a subset of $P_n(\mathbb{R})$ with no two polynomials having the same degree, we see that following the process seen in example 4 in section 1.5 shows that W is a linearly independent set containing $n + 1$ vectors. Hence, W is a basis for $P_n(\mathbb{R})$ such that any $g(x) \in P_n(\mathbb{R})$ by Theorem 1.11. Consequently, $g(x)$ can be expressed in terms of the vectors in W such that

$$g(x) = c_0 f(x) + c_1 f^{(1)}(x) + c_2 f^{(2)}(x) + \dots + c_n f^{(n)}(x)$$

for unique scalars c_0, c_1, \dots, c_n by Theorem 1.8. ■

Exercise 1.6.29

- (a) Prove that if W_1 and W_2 are finite-dimensional subspaces of a vector space V , then the subspace $W_1 + W_2$ is finite-dimensional, and $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$.

Proof. Let W_1 and W_2 be subspaces of V . Since W_1 and W_2 are finite-dimensional, we also know that $W_1 + W_2$ is finite-dimensional. Now, we will show that

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

Consider $W_1 \cap W_2$ and note that $W_1 \cap W_2$ being finite-dimensional implies that it

contains a basis $\beta_0 = \{u_1, u_2, \dots, u_k\}$. We can extend β_0 into a basis for W_1 by adding vectors v_1, v_2, \dots, v_m into β_0 . Denote this new set as β_1 . Likewise, we add vectors w_1, w_2, \dots, w_p into β_0 to make a basis β_2 for W_2 . We claim that $\beta = \beta_0 \cup \beta_1 \cup \beta_2$ is a basis for $W_1 + W_2$. First, we will show that β is linearly independent. To do this, we need to show that

$$\sum_{i=1}^k a_i u_i + \sum_{j=1}^m b_j v_j + \sum_{\ell=1}^p \gamma_\ell w_\ell = 0. \quad (1)$$

Subtracting the third term on both sides of (1) produces the following equation:

$$\sum_{i=1}^k a_i u_i + \sum_{j=1}^m b_j v_j = - \sum_{\ell=1}^p \gamma_\ell w_\ell.$$

Observe that the left-hand side is an element of W_1 while the other side is an element of W_2 . Hence, we know that the term on the right-hand side of (1) is also an element of $W_1 \cap W_2$. This implies that

$$- \sum_{\ell=1}^p \gamma_\ell w_\ell = \sum_{i=1}^k \delta_i u_i$$

which can be re-written as

$$\sum_{i=1}^k \delta_i u_i + \sum_{\ell=1}^p \gamma_\ell w_\ell = 0.$$

Since β_2 is a basis for W_2 , we know that $\delta_i = 0$ and $\gamma_\ell = 0$ implying that $a_i = 0$ and $\gamma_\ell = 0$. We can re-write (1) in the following form:

$$\sum_{i=1}^k a_i u_i + \sum_{j=1}^m b_j v_j = 0. \quad (2)$$

Since β_1 is a linearly independent set, we get that $a_i = 0$ and $b_j = 0$. Hence, (1) contains the trivial-representation which implies that $\beta = \beta_0 \cup \beta_1 \cup \beta_2$ is a linearly independent set.

Now, we will show that β spans $W_1 + W_2$. Observe that $\text{span}(\beta) \subseteq W_1 + W_2$. Now, we will show $W_1 + W_2 \subseteq \text{span}(\beta)$. Suppose we take a vector $v \in W_1 + W_2$ that is not in β and adjoin this vector in β . Note that $\beta \cup \{v\}$ produces a linearly dependent set that by which Theorem 1.7 implies that $v \in \text{span}(\beta)$ and we are done. Hence, β spans $W_1 + W_2$ and thus β is a basis.

Note that β contains exactly $m+p+k$ vectors. Hence, denote $\dim(W_1 + W_2) = m+p+k$ which can be re-written as

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

■

- (b) Let W_1 and W_2 be finite-dimensional subspaces of a vector space V , and let $V = W_1 + W_2$. Deduce that V is the direct sum of W_1 and W_2 if and only if $\dim(V) = \dim(W_1) + \dim(W_2)$.

Proof. Suppose V is a direct sum of W_1 and W_2 . Then $V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$. We need to show that $\dim(V) = \dim(W_1) + \dim(W_2)$. Since $W_1 \cap W_2 = \{0\}$, we know that it contains the empty set \emptyset as the basis for $W_1 \cap W_2$. Hence, $\dim(W_1 \cap W_2) = 0$. Using the formula derived in part (a), we can write

$$\begin{aligned}\dim(V) &= \dim(W_1 + W_2) \\ &= \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) \\ &= \dim(W_1) + \dim(W_2)\end{aligned}$$

and we are done.

Conversely, $\dim(V) = \dim(W_1) + \dim(W_2)$ implies that $V = W_1 + W_2$. Using part (a) again, we see that the sum $\dim(V) = \dim(W_1) + \dim(W_2) - 0$ implies that $\dim(W_1 \cap W_2) = 0$ and hence $W_1 \cap W_2$ must be equal to the zero set $\{0\}$ (which we know by definition that $\text{span}(\emptyset) = \{0\}$). Hence, V is a direct sum of W_1 and W_2 . ■

Exercise 1.6.31

Let W_1 and W_2 be subspaces of a vector space V having dimensions m and n , respectively, where $m \geq n$.

- (a) Prove that $\dim(W_1 \cap W_2) \leq n$.

Proof. Observe that W_1 and W_2 being subspaces of V implies that $W_1 \cap W_2$ is a subspace of V . Hence, $W_1 \cap W_2$ is finite-dimensional. Denote $\dim(W_1 \cap W_2) = k$ and let β be a basis for $W_1 \cap W_2$. Since $W_1 \cap W_2 \subseteq W_2$, we know that β must contain at most $\dim(W_2) = n$. Hence, $\dim(W_1 \cap W_2) \leq n$. ■

- (b) Prove that $\dim(W_1 + W_2) \leq m + n$.

Proof. Using the formula found in part (a) of Exercise 1.3.29, part (a) of this exercise, and $\dim(W_1) \geq \dim(W_2)$, we find that

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

Exercise 1.6.33

- (a) Let W_1 and W_2 be subspaces of a vector space V such that $V = W_1 \oplus W_2$. If β_1 and β_2 are bases for W_1 and W_2 , respectively, show that $\beta_1 \cup \beta_2$ is a basis for V .

Proof. Let W_1 and W_2 be subspaces of V . Assume β_1 and β_2 are bases for W_1 and W_2 respectively. We need to show that $\beta_1 \cap \beta_2 = \emptyset$ and $\beta_1 \cup \beta_2$ is a basis for V .

Since β_1 and β_2 contain distinct linearly independent vectors, we must have $\beta_1 \cap \beta_2 = \emptyset$. Since V is a direct sum of the W_1 and W_2 , we know that $W_1 \cap W_2 = \{0\}$ by definition. Since β_1 and β_2 generate W_1 and W_2 respectively, we must have $\text{span}(\beta_1) \cap \text{span}(\beta_2) = \{0\}$. Now, we have the set $\beta_1 \cup \beta_2$ as a linearly independent set by exercise 1.5.21. Observe that $\text{span}(\beta_1 \cup \beta_2) \subseteq V$ follows immediately. Now, take any $v \in V$ that is not in $\beta_1 \cup \beta_2$ such that adjoining this vector $v \in V$ produces a linearly dependent set. By Theorem 1.7, we have $v \in \text{span}(\beta_1 \cup \beta_2)$. Thus, we have $V \subseteq \text{span}(\beta_1 \cup \beta_2)$. Hence, $\beta_1 \cup \beta_2$ is a generating set for V and we are done. ■

- (b) Conversely, let β_1 and β_2 be disjoint bases for subspaces W_1 and W_2 , respectively, of a vector space V . Prove that if $\beta_1 \cup \beta_2$ is a basis for V , then $V = W_1 \oplus W_2$.

Proof. Let β_1 and β_2 be disjoint bases for subspaces W_1 and W_2 respectively. Suppose $\beta_1 \cup \beta_2$ is a basis for V . This tells us that $\beta_1 \cup \beta_2$ is linearly independent. Thus, $\text{span}(\beta_1) \cap \text{span}(\beta_2) = \{0\}$ and hence $W_1 \cap W_2 = \{0\}$ since $\text{span}(\beta_1) = W_1$ and $\text{span}(\beta_2) = W_2$. This tells us that $\dim(W_1 \cap W_2) = 0$. Using the fact that $\beta_1 \cup \beta_2$ is a basis for V that contains exactly $\dim(W_1) + \dim(W_2)$, we get that

$$\dim(V) = \dim(W_1) + \dim(W_2).$$

By part (b) of Exercise 1.6.29, we get that $V = W_1 \oplus W_2$. ■

Exercise 1.6.34

- (a) Prove that if W_1 is any subspace of a finite-dimensional vector space V , then there exists a subspace W_2 of V such that $V = W_1 \oplus W_2$.

Proof. Since W_1 is a subspace of a finite-dimensional vector space V , we know that W_1 is also finite-dimensional and $\dim(W_1) \leq \dim(V)$ by Theorem 1.11. Thus, let β be a basis for W_1 and let α be a basis for V . Since α is a generating set consisting of $\dim(V)$ vectors and β is a linearly independent subset of V , we can find a subset σ of α consisting of $\dim(V) - \dim(W_1)$ vectors such that $\beta \cup \sigma$ generates V by the Replacement Theorem. Suppose σ is a basis for a subspace of V denoted by W_2 for which $\dim(W_2) = \dim(V) - \dim(W_1)$. Note that $\beta \cup \sigma$ contains exactly $\dim(V)$ vectors so it is also a basis for V and that $\beta \cap \sigma = \emptyset$. Hence, $\dim(V) = \dim(W_1) + \dim(W_2)$ for which it implies that $V = W_1 \oplus W_2$. ■

- (b) Let $V = \mathbb{R}^2$ and $W_1 = \{(a_1, 0) : a_1 \in \mathbb{R}\}$. Give examples of two different subspaces W_2 and W'_2 such that $V = W_1 \oplus W_2$ and $V = W_1 \oplus W'_2$.

Proof. TO DO. ■

Exercise 1.6.35

Let W be a subspace of a finite-dimensional vector space V , and consider the basis $\beta_0 = \{u_1, u_2, \dots, u_k\}$ for W . Let $\beta_1 = \{u_1, u_2, \dots, u_k, u_{k+1}, \dots, u_n\}$ be an extension of this basis to a basis for V .

- (a) Prove that $\beta_2 = \{u_{k+1} + W, u_{k+2} + W, \dots, u_n + W\}$ is a basis for V/W .

Proof. To show that β_2 is a basis, we need to show that β_2 is a linearly independent set and a generating set for V . Observe that

$$\sum_{j=k+1}^n \delta_j (u_j + W) = W \tag{1}$$

for scalars δ_j for $k+1 \leq j \leq n$. Note that W , in this case, is the zero vector of V/W . Hence, (1) implies that

$$\left[\sum_{j=k+1}^n \delta_j u_j \right] + W = W. \tag{2}$$

This implies that

$$\sum_{j=k+1}^n \delta_j u_j \in W. \quad (3)$$

Since W contains β_0 as a basis, we can write

$$\sum_{j=k+1}^n \delta_j u_j = \sum_{i=1}^k \gamma_i u_i$$

for scalars γ_i for all $1 \leq i \leq k$ which can be re-written to

$$\sum_{j=k+1}^n \delta_j u_j - \sum_{i=1}^k \gamma_i u_i = 0$$

where all $\delta_j = 0$ and $\gamma_i = 0$ since β_1 is a basis for V . Since all $\delta_j = 0$, we get that β_2 is a linearly independent set.

To show that β_2 is a generating set for V/W , we need to show that $\text{span}(\beta_2) = V/W$. Note that the containment $\text{span}(\beta_2) \subseteq V/W$. To show that other containment, let $v \in V$ not in β_2 . Observe that adjoining v to β_2 creates a linearly dependent set. By Theorem 1.7, we have $v \in \text{span}(\beta_2)$. Hence, $V \subseteq \text{span}(\beta_2)$. ■

- (b) Derive a formula relating $\dim(V)$, $\dim(W)$, and $\dim(V/W)$.

Proof. In part (a), we see that β_2 contains n amount of vectors. Unioning this set with β_1 creates a basis for V that contains $\dim(W) + \dim(V/W)$ vectors. Hence, we must have

$$\dim(V/W) = \dim(V) - \dim(W).$$
■

Chapter 2

Linear Transformations and Matrices

2.1 Linear Transformations, Null spaces, and Ranges

Exercise 2.1.1

Label the following statements as true or false. In each part, V and W are finite-dimensional vector spaces (over F), and T is a function from V to W .

- (a) If T is linear, then T preserves sums and scalar products.

Solution. True. This is by definition. ■

- (b) If $T(x + y) = T(x) + T(y)$, then T is linear.

Solution. False. We also need to have $T(cx) = cT(x)$ where $x \in V$ and $c \in F$ in order for T to be a linear map. ■

- (c) If T is one-to-one if and only if the only vector x such that $T(x) = 0$ is $x = 0$.

Solution. True. This is Theorem 2.2. ■

- (d) If T is linear, then $T(0_V) = 0_W$.

Solution. True by properties of linear maps. ■

- (e) If T is linear, then $\text{nullity}(T) + \text{rank}(T) = \dim(W)$.

Solution. False. This only happens when V and W have equal dimensions. ■

- (f) If T is linear, then T carries linearly independent subsets of V onto linearly independent subsets of W .

Solution. False. We need T to be injective in order to make this valid. ■

- (g) If $T, U : V \rightarrow W$ are both linear and agree on a basis for V , then $T = U$.

Solution. True by Corollary to Theorem 2.6. ■

- (h) Given $x_1, x_2 \in V$ and $y_1, y_2 \in W$, there exists a linear transformation $T : V \rightarrow W$ such that $T(x_1) = y_1$ and $T(x_2) = y_2$.

Solution. False. We need to have a basis for V . ■

For Exercises 2 through 6, prove that T is a linear transformation, and find bases for both $N(T)$ and $R(T)$. Then compute the nullity and rank of T , and verify the dimension theorem. Finally, use the appropriate theorems in this section to determine whether T is injective or surjective.

Exercise 2.1.2

$T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ defined by $T(a_1, a_2, a_3) = (a_1 - a_2, 2a_3)$.

Solution. Our first goal is to show that T is linear. Let $x, y \in \mathbb{R}^3$ and $c \in \mathbb{R}$ where $x = (a_1, a_2, a_3)$ and $y = (b_1, b_2, b_3)$. Then observe that

$$\begin{aligned} T(cx + y) &= T(ca_1 + b_1, ca_2 + b_2, ca_3 + b_3) \\ &= ([ca_1 + b_1] - [ca_2 + b_2], 2(ca_3 + b_3)) \\ &= ([ca_1 - ca_2] + [b_1 - b_2], 2ca_3 + 2b_3) \\ &= (c(a_1 - a_2), 2ca_3) + (b_1 - b_2, 2b_3) \\ &= c(a_1 - a_2, 2a_3) + (b_1 - b_2, 2b_3) \\ &= cT(x) + T(y). \end{aligned}$$

Hence, T is linear. Now, let's compute the bases for both $N(T)$ and $R(T)$. To compute $N(T)$, we need to have all the solutions needed for

$$T(a_1, a_2, a_3) = (a_1 - a_2, 2a_3) = 0$$

to hold. Solving the equation above gives us the following solutions

$$a_1 = a_2 \text{ and } a_3 = 0.$$

Hence, for any $a \in \mathbb{R}$ we have

$$N(T) = \{(a, a, 0) : a \in \mathbb{R}\}.$$

Note that $N(T)$ is generated by the linearly independent set $\{(1, 1, 0)\}$. Thus, $\{(1, 1, 0)\}$ is a basis for $N(T)$ and contains 1 vector. So, $\text{nullity}(T) = 1$ by the Dimension Theorem. This immediately tells us that T is not injective since $\text{nullity}(T) \neq 0$ and hence $\text{rank}(T) = 2$. But $\text{rank}(T) = \dim(\mathbb{R}^2)$. Hence, T must be surjective. Since $R(T) = \mathbb{R}^2$, the basis is just $\{(1, 0), (0, 1)\}$. ■

Exercise 2.1.8

$T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ defined by $T(a_1, a_2) = (a_1 + a_2, 0, 2a_1 - a_2)$.

Solution. First, we show that T is linear. Let $x, y \in \mathbb{R}^2$ defined by $x = (a_1, a_2)$ and $y = (b_1, b_2)$. Let $c \in \mathbb{R}$. Note that

$$cx + y = (ca_1 + b_1, ca_2 + b_2).$$

Then observe that

$$\begin{aligned}
T(cx + y) &= ((ca_1 + b_1) + (ca_2 + b_2), 0, 2(ca_1 + b_1) - (ca_2 + b_2)) \\
&= (c(a_1 + a_2) + (b_1 + b_2), 0, c(2a_1 - a_2) + (b_1 - b_2)) \\
&= (c(a_1 + a_2), c(2a_1 - a_2)) + (b_1 + b_2, 2b_1 - b_2) \\
&= c(a_1 + a_2, 2a_1 - a_2) + (b_1 + b_2, 2b_1 - b_2) \\
&= cT(x) + T(y).
\end{aligned}$$

Hence, T is linear. ■

Exercise 2.1.7

Prove properties 1, 2, 3, and 4 on page 65.

Proof. See proof in notes. ■

Exercise 2.1.8

Prove that the transformations in Example 2 and 3 are linear.

Proof. First we prove that $T_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$$T_\theta(a_1, a_2) = (a_1 \cos \theta - a_2 \sin \theta, a_1 \cos \theta + a_2 \sin \theta)$$

is linear. Let $x, y \in \mathbb{R}^2$ defined by $x = (a_1, a_2)$ and $y = (b_1, b_2)$. Let $c \in F$ such that $cx = (ca_1, ca_2)$. To make the computation less difficult, we have

$$cx + y = (ca_1 + b_1, ca_2 + b_2).$$

Using the definition T_θ now, we have that

$$\begin{aligned}
T_\theta(cx + y) &= ((ca_1 + b_1) \cos \theta - (ca_2 + b_2) \sin \theta, \\
&\quad (ca_1 + b_1) \cos \theta + (ca_2 + b_2) \sin \theta) \\
&= (c(a_1 \cos \theta - a_2 \sin \theta) + (b_1 \cos \theta - b_2 \sin \theta), \\
&\quad c(a_1 \cos \theta + a_2 \sin \theta) + (b_1 \cos \theta + b_2 \sin \theta)) \\
&= c(a_1 \cos \theta - a_2 \sin \theta, a_1 \cos \theta + a_2 \sin \theta) \\
&\quad + (b_1 \cos \theta - b_2 \sin \theta, b_1 \cos \theta + b_2 \sin \theta) \\
&= cT_\theta(a_1, a_2) + T_\theta(b_1, b_2) \\
&= cT_\theta(x) + T_\theta(y).
\end{aligned}$$

Hence, we get that T_θ is linear.

Using the same process, we show that $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(a_1, a_2) = (a_1, -a_2)$ is linear. That is, we have

$$\begin{aligned}
T(cx + y) &= (ca_1 + b_1, -(ca_2 + b_2)) \\
&= (ca_1 + b_1, -ca_2 - b_2) \\
&= c(a_1, -a_2) + (b_1, -b_2) \\
&= cT(a_1, a_2) + T(b_1, b_2) \\
&= cT(x) + T(y).
\end{aligned}$$

Hence, T is also linear.

Now, let's compute $N(T)$. We need to find solutions to the scalars a_1 and a_2 such that

$$(a_1 + a_2, 0, 2a_1 - a_2) = (0, 0, 0).$$

Equating entries, we get the following system of linear equations:

$$\begin{aligned}a_1 + a_2 &= 0 \\ 2a_1 - a_2 &= 0.\end{aligned}$$

Using elimination, we get that $a_1 = 0$ and $a_2 = 0$. This tells us that $N(T) = \{0\}$ and that T is injective by Theorem 2.4. Note that \emptyset is a basis for $N(T)$. Furthermore, $N(T) = \{0\}$ tells us that $\text{nullity}(T) = 0$. Using the dimension theorem, this means that $\text{rank}(T) = 2$.

To find the basis of $R(T)$, note that $\beta = \{(1, 0), (0, 1)\}$ is a basis for \mathbb{R}^2 and that $T(\beta)$ generates $R(T)$. Thus,

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

which solving the following equation

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

for scalars a and b yields $a = 0$ and $b = 0$. Hence, $T(\beta)$ is a basis for $R(T)$. ■

Exercise 2.1.10

Suppose that $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is linear, $T(1, 0) = (1, 4)$, and $T(1, 1) = (2, 5)$.

Solution. First, observe that $(2, 3) \in \mathbb{R}^2$ can be expressed in terms of a linear combination of $(1, 0)$ and $(1, 1)$. Hence, we need to find scalars $a, b \in \mathbb{R}$ such that

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

Solving for the scalars gives us the solutions $a = -1$ and $b = 3$. Since T is linear, we can now compute $T(2, 3)$ by doing the following:

$$\begin{aligned}T(2, 3) &= T(-(1, 0) + 3(1, 1)) \\ &= -T(1, 0) + 3T(1, 1) \\ &= -(1, 4) + 3(2, 5) \\ &= (5, 11).\end{aligned}$$

Hence, we get that $T(2, 3) = (5, 11)$.

Note that the domain and codomain have equal finite dimensions. Hence, Theorem 2.5 tells us that T must be one-to-one (injective). ■

Exercise 2.1.11

Prove that there exists a linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ such that $T(1, 1) = (1, 0, 2)$ and $T(2, 3) = (1, -1, 4)$. What is $T(8, 11)$?

Solution. Note that $\beta = \{(1, 0), (0, 1)\}$ is a basis for \mathbb{R}^2 . By Theorem 2.6, T being linear implies that there exists a unique linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ such that $T(1, 1) =$

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

To compute $T(8, 11)$, we need to find scalars $a, b \in \mathbb{R}$ such that

$$a(1, 1) + b(2, 3) = (8, 11).$$

Solving for a and b gives us the following linear combination:

$$2(1, 1) + 3(2, 3) = (8, 11).$$

Since T is linear, we find that

$$\begin{aligned} T(8, 11) &= T(2(1, 1) + 3(2, 3)) \\ &= 2T(1, 1) + 3T(2, 3) \\ &= 2(1, 0, 2) + 3(1, -1, 4) \\ &= (5, -3, 16) \end{aligned}$$

■

Exercise 2.1.13

Let V and W be vector spaces, let $T : V \rightarrow W$ be linear, and let $\{w_1, w_2, \dots, w_k\}$ be a linearly independent subset of $R(T)$. If $S = \{v_1, v_2, \dots, v_k\}$ is chosen so that $T(v_i) = w_i$ for $i = 1, 2, \dots, k$ then S is linearly independent.

Proof. Consider

$$\sum_{i=1}^k a_i v_i = 0 \tag{1}$$

for some scalars $a_1, a_2, \dots, a_k \in F$. In order to show that S is linearly independent, we need to show that $a_i = 0$ for all $1 \leq i \leq k$. Since T is linear, we get that $T(0) = 0$ implies

$$T\left(\sum_{i=1}^k a_i v_i\right) = 0.$$

Since T is linear and S is chosen so that $T(v_i) = w_i$ for $1 \leq i \leq k$, we get that

$$\sum_{i=1}^k a_i T(v_i) = 0 \Leftrightarrow \sum_{i=1}^k a_i w_i = 0.$$

Since $\{w_1, w_2, \dots, w_k\}$ is linearly independent, we must have $a_i = 0$ for all $1 \leq i \leq k$. But this tells us that (1) must have the trivial representation. Hence, S must also be linearly independent. ■

Exercise 2.1.14

Let V and W be vector spaces and $T : V \rightarrow W$ be linear.

- (a) Prove that T is injective if and only if T carries linearly independent subsets of V onto linearly independent subsets of W .

Proof. (\Rightarrow) Let T be injective. Since \mathcal{W} is a linearly independent subset of W , we

have $w_1, w_2, \dots, w_n \in \mathcal{W}$ and scalars $a_1, a_2, \dots, a_n \in F$ such that

$$\sum_{i=1}^n a_i w_i = 0.$$

Since $T(v_i) = w_i$ and T is linear, we have

$$\sum_{i=1}^n a_i T(v_i) = 0 \Leftrightarrow T\left(\sum_{i=1}^n a_i v_i\right) = 0$$

with $v_i \in S$ being linearly independent. Since T is injective, we must have

$$\sum_{i=1}^n a_i v_i = 0$$

with $v_i \in S$ being linearly independent. Hence, T carries S onto \mathcal{W} .

(\Leftarrow) Suppose T carries linearly independent subsets of V onto linearly independent subsets of W . Our goal is to show that T is injective. Suppose

$$T\left(\sum_{i=1}^n a_i v_i\right) = T\left(\sum_{i=1}^n b_i v_i\right) \quad (1)$$

for some scalars a_i, b_i and vectors $v_i \in S$ for all $1 \leq i \leq k$. Since T is linear and $T : S \rightarrow \mathcal{W}$ is onto, we can re-write (1) into the following form:

$$\sum_{i=1}^k a_i T(v_i) = \sum_{i=1}^k b_i T(v_i) \quad (2)$$

which manipulating again, we get that

$$\sum_{i=1}^k (a_i - b_i) T(v_i) = 0. \quad (3)$$

Since $T(v_i) \in W$ and W is linearly independent, we must have $a_i - b_i = 0$ if and only if $a_i = b_i$. Hence, we have

$$\sum_{i=1}^n a_i v_i = \sum_{i=1}^n b_i v_i$$

and so T must be injective. ■

- (b) Suppose that T is injective and that S is a subset of V . Prove that S is linearly independent if and only if $T(S)$ is linearly independent.

Proof. (\Rightarrow) In order to show that $T(S)$ is linearly independent, we must show that

$$\sum_{i=1}^n a_i T(v_i) = 0 \quad (1)$$

contains the trivial-representation. Since T is linear, we can write (1) into the following form

$$T\left(\sum_{i=1}^n a_i v_i\right) = 0.$$

This implies that

$$\sum_{i=1}^n a_i v_i \in N(T).$$

Since T is injective, we know that $N(T) = \{0\}$ (by Theorem 2.4), and so we must have

$$\sum_{i=1}^n a_i v_i = 0. \quad (2)$$

But $v_i \in S$ for $1 \leq i \leq k$ is linearly independent, and so $a_i = 0$ for all $1 \leq i \leq k$. This tells us that (1) contains the trivial-representation. Hence, $T(S)$ is linearly independent.

(\Rightarrow) Suppose $T(S)$ is linearly independent. Then observe that

$$\sum_{i=1}^n a_i T(v_i) = 0$$

for some scalars a_i and $T(v_i) \in T(S)$ for $1 \leq i \leq k$ such that $a_i = 0$. Since T is linear and T is injective, we can write

$$T\left(\sum_{i=1}^n a_i v_i\right) = 0 \Leftrightarrow \sum_{i=1}^n a_i v_i = 0.$$

Since $a_i = 0$ and $v_i \in S$, we also find that S is a linearly independent set. ■

- (c) Suppose $\beta = \{v_1, v_2, \dots, v_n\}$ is a basis for V and T is injective and surjective. Prove that $T(\beta) = \{T(v_1), T(v_2), \dots, T(v_n)\}$ is a basis for W .

Proof. Since β is a basis for V , we get that $\text{span}(T(\beta)) = R(T)$ by Theorem 2.2. Since T is surjective, we have $R(T) = W$, so $\text{span}(T(\beta)) = W$. Hence, $T(\beta)$ generates W . Since β is a linearly independent subset of V and T is an injective linear map, we must also have $T(\beta)$ as a linearly independent subset of W by part (b). Hence, $T(\beta)$ is a basis for W . ■

Exercise 2.1.15

Recall the definition of $P(\mathbb{R})$ on page 10. Define

$$T : P(\mathbb{R}) \rightarrow P(\mathbb{R}) \text{ by } T(f(x)) = \int_0^x f(t) \, dt.$$

Prove that T is linear and injective, but not surjective.

Proof. First, we show that $T : P(\mathbb{R}) \rightarrow P(\mathbb{R})$ defined by

$$T(f(x)) = \int_0^x f(t) \, dt.$$

Let $cf(x) + g(x) \in P(\mathbb{R})$ where $c \in F$. Then observe that

$$\begin{aligned} T(cf(x) + g(x)) &= \int_0^x [cf(t) + g(t)] dt \\ &= \int_0^x cf(t) dt + \int_0^x g(t) dt \\ &= c \int_0^x f(t) dt + \int_0^x g(t) dt \\ &= cT(f(x)) + T(g(x)). \end{aligned}$$

Hence, $T : P(\mathbb{R}) \rightarrow P(\mathbb{R})$ is a linear map. Let $f(x), g(x) \in P(\mathbb{R})$, then

$$\begin{aligned} T(f(x)) &= T(g(x)) \\ \int_0^x f(t) dt &= \int_0^x g(t) dt \\ \frac{d}{dx} \left[\int_0^x f(t) dt \right] &= \frac{d}{dx} \left[\int_0^x g(t) dt \right] \\ f(x) &= g(x). \end{aligned}$$

Hence, T is an injective map. To see why T fails to be surjective, observe that $1 \in P(\mathbb{R})$ but we cannot find a polynomial $f(x)$ such that $T(f(x)) = 1$; that is, the integration of any polynomial can never yield 1. ■

Exercise 2.1.16

Let $T : P(\mathbb{R}) \rightarrow P(\mathbb{R})$ be defined by $T(f(x)) = f'(x)$. Recall that T is linear. Prove that T is surjective, but not injective.

Proof. Let $T : P(\mathbb{R}) \rightarrow P(\mathbb{R})$ be defined by $T(f(x)) = f'(x) = \frac{d}{dx}[f(x)]$. For $cf(x) + g(x) \in P(\mathbb{R})$ where $c \in F$, observe that

$$\begin{aligned} T(cf(x) + g(x)) &= \frac{d}{dx}[cf(x) + g(x)] \\ &= \frac{d}{dx}[cf(x)] + \frac{d}{dx}[g(x)] \\ &= c \frac{d}{dx}[f(x)] + \frac{d}{dx}[g(x)] \\ &= cT(f(x)) + T(g(x)). \end{aligned}$$

Hence, T is a linear map. Now, we want to show that T is surjective. Define

$$f(x) = \int_0^x g(t) dt.$$

Then

$$T(f(x)) = \frac{d}{dx}[f(x)] = \frac{d}{dx} \left[\int_0^x g(t) dt \right] = g(x)$$

by the Second Fundamental Theorem of Calculus. Hence, T is surjective.

To show that T is not injective, let $x^n + a, x^n + b \in P(\mathbb{R})$ where $a, b \in \mathbb{R}$ such that $a \neq b$ and $n > 0$. Clearly, we have $x^n + a \neq x^n + b$. But, we have

$$T(x^n + a) = \frac{d}{dx}[x^n + a] = nx^{n-1}$$

and

$$T(x^n + b) = \frac{d}{dx}[x^n + b] = nx^{n-1}.$$

Note that $x^n + a \neq x^n + b$ yet $T(x^n + a) = T(x^n + b)$. Hence, T is not injective. ■

Exercise 2.1.17

Let V and W be finite-dimensional vector spaces and $T : V \rightarrow W$ be linear.

- (a) Prove that if $\dim(V) < \dim(W)$, then T cannot be surjective.

Proof. Using the Dimension Theorem, we find that

$$\text{rank}(T) = \dim(V) - \text{nullity}(T).$$

Since $\dim(V) < \dim(W)$, we find that

$$\text{rank}(T) < \dim(W) - \text{nullity}(T) < \dim(W).$$

Hence, T cannot be surjective in this case. ■

- (b) Prove that if $\dim(V) > \dim(W)$, then T cannot be injective.

Proof. Using the Dimension Theorem again and $\dim(V) > \dim(W)$, we find that

$$\text{nullity}(T) = \dim(V) - \text{rank}(T) > \dim(W) - \text{rank}(T) > 0.$$

This means that $\text{nullity}(T)$ can never be zero, otherwise T is injective. Hence, T can never be injective if $\dim(V) > \dim(W)$. ■

Exercise 2.1.20

Let V and W be vector spaces with subspaces V_1 and W_1 , respectively. If $T : V \rightarrow W$ is linear, prove that $T(V_1)$ is a subspace of W and that $\{x \in V : T(x) \in W_1\}$ is a subspace of V .

Proof. First, we prove that $T(V_1)$ is a subspace of W . Let $x, y \in T(V_1)$ and $c \in F$. Since V_1 is a subspace of V , we know that $0_V \in V_1$, $x + y \in V_1$, and $cx \in V_1$. Observe that $T(0_V) = 0_W$ since T is linear. Hence, $0_W \in T(V_1)$. Let $x, y \in T(V_1)$. There exists $z, w \in V_1$ such that $x = T(z)$ and $y = T(w)$. Since $z + w \in V_1$ and T is linear, we must have $T(z + w) = T(z) + T(w) = x + y$. This tells us that $x + y \in T(V_1)$. Now, let $c \in F$. Since $cz \in V_1$ and T is linear, we must have $T(cz) = cT(z) = cx$. Hence, $cx \in T(V_1)$. Thus, $T(V_1)$ is a subspace of W .

Now, we show $S = \{x \in V : T(x) \in W_1\}$ is a subspace of V . Since $0_W \in W_1$ (because W_1 is a subspace of W) and T is linear, we have that $T(0_V) = 0_W$. Hence, $0_V \in S$. Now, let $x, y \in S$. Hence, $T(x), T(y) \in W_1$ implies $T(x) + T(y) \in W_1$ since W_1 is a subspace of W . Since T is linear, we have $T(x) + T(y) = T(x + y)$, and so $x + y \in S$. Now, let $c \in F$ and $x \in S$. Again, W_1 is a subspace so $cT(x) \in W_1$. Thus, T being linear implies that $cT(x) = T(cx)$. Hence, $cx \in S$. Thus, S is a subspace of V . ■

Exercise 2.1.21

Let V be the vector space of sequences described in Example 5 of Section 1.2. Define the functions $T, U : V \rightarrow V$ by

$$T(a_1, a_2, \dots) = (a_2, a_3, \dots) \quad \text{and} \quad U(a_1, a_2, \dots) = (0, a_1, a_2, \dots).$$

T and U are called the **left shift** and **right shift** operators on V , respectively.

(a) Prove that T and U are linear.

Proof. Let $(x_n), (y_n) \in V$ with $(x_n) = (a_1, a_2, \dots)$ and $(y_n) = (b_1, b_2, \dots)$. Let $c \in F$. Then we have

$$\begin{aligned} T(cx_n + y_n) &= (ca_2 + b_2, ca_3 + b_3, \dots) \\ &= (ca_2, ca_3, \dots) + (b_2, b_3, \dots) \\ &= c(a_2, a_3, \dots) + (b_2, b_3, \dots) \\ &= cT(x_n) + T(y_n). \end{aligned}$$

Hence, $T : V \rightarrow V$ is a linear map.

Now with $U : V \rightarrow V$ observe that

$$\begin{aligned} U(cx_n + y_n) &= (0, ca_1 + b_1, ca_2 + b_2, \dots) \\ &= (0, ca_1, ca_2, \dots) + (0, b_1, b_2, \dots) \\ &= c(0, a_1, a_2, \dots) + (0, b_1, b_2, \dots) \\ &= cU(x_n) + U(y_n). \end{aligned}$$

Hence, $U : V \rightarrow V$ is a linear map. ■

(b) Prove that T is surjective, but not injective.

Proof. Let $\{b_n\} \in V$. Define this sequence as

$$\{b_n\} = (b_1, b_2, \dots).$$

We can construct a sequence such that $b_i = a_{i+1}$ for $i \in \mathbb{N}$. Hence,

$$(b_1, b_2, \dots) = (a_2, a_3, \dots).$$

By definition of T , we get that

$$T(a_1, a_2, \dots) = (a_2, a_3, \dots) = (b_1, b_2, \dots).$$

Since $\{b_n\} \in V$, we get that T is surjective.

To see why T is not injective, suppose we take two sequences distinct $\{x_n\}, \{y_n\} \in V$ defined by

$$\{x_n\} = (a, 0, 0, \dots) \text{ and } \{y_n\} = (b, 0, 0, \dots)$$

respectively. But we have $T(\{x_n\}) = 0_n = T(\{y_n\})$ where 0_n is the sequence with all terms being zero. Hence, T cannot be injective. ■

(c) Prove that U is injective, but not surjective.

Proof. Let $\{a_n\}, \{b_n\} \in V$ defined by

$$\{a_n\} = (a_1, a_2, \dots) \text{ and } \{b_n\} = (b_1, b_2, \dots).$$

Then observe that

$$\begin{aligned} U(a_1, a_2, \dots) &= U(b_1, b_2, \dots) \\ (0, a_1, a_2, \dots) &= (0, b_1, b_2, \dots). \end{aligned}$$

Equating entries, we get that $a_i = b_i$ for all $i \in \mathbb{N}$ and so, $\{a_n\} = \{b_n\}$. Hence, U is injective.

Observe that $(x_1, 0, 0, \dots) \in V$ but there does not exist a sequence $\{x_n\} \in V$ such that $U(x_1, x_2, \dots) = (x_1, 0, 0)$. Hence, U is not surjective. ■

Exercise 2.1.22

Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}$ be linear. Show that there exist scalars a, b , and c such that $T(x, y, z) = ax + by + cz$ for all $(x, y, z) \in \mathbb{R}^3$. Can you generalize this result for $T : F^n \rightarrow F$? State and prove an analogous result for $T : F^n \rightarrow F^m$.

Proof. Let $(x, y, z) \in \mathbb{R}^3$ arbitrary. Observe that

$$(x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z) = xe_1 + ye_2 + ze_3.$$

since $\beta = \{e_1, e_2, e_3\}$ is a basis for \mathbb{R}^3 . Since T is linear, we know, by Theorem 2.2, that $\text{span}(T(\beta)) = R(T)$. So, we have

$$\begin{aligned} T(x, y, z) &= T(xe_1 + ye_2 + ze_3) \\ &= T(e_1)x + T(e_2)y + T(e_3)z. \end{aligned}$$

Since $T : \mathbb{R}^3 \rightarrow \mathbb{R}$, we know that $T(e_1), T(e_2), T(e_3) \in \mathbb{R}$ are just scalars, so denote $T(e_1) = a, T(e_2) = b, T(e_3) = c$. Hence, we have

$$T(x, y, z) = ax + by + cz.$$

Now for the n th case, our basis β for F^n now contains n . So, we have

$$T(x_1, x_2, x_3, \dots, x_n) = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

for some scalars $a_1, a_2, \dots, a_n \in F$ using the same argument above. ■

Exercise 2.1.23

Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}$ be linear. Describe geometrically the possibilities for the null space of T .

Proof. By definition of T described in Exercise 2.1.22, the null space is described as the set

$$N(T) = \{ax + by + cz = 0 : a, b, c \in \mathbb{R} \text{ and } (x, y, z) \in \mathbb{R}^3\}.$$

This represents two vectors in \mathbb{R}^3 where we fix $(x, y, z) \in \mathbb{R}^3$ such that we find a $(a, b, c) \in \mathbb{R}^3$ such that the two vectors are perpendicular to each other. ■

Definition 2.1.1 (Projections). Let V be a vector space and W_1 and W_2 be subspaces of V such that $V = W_1 \oplus W_2$. A function $T : V \rightarrow V$ is called the **projection on W_1 along W_2** if, for $x = x_1 + x_2$ with $x_1 \in W_1$ and $x_2 \in W_2$, we have $T(x) = x_1$.

Exercise 2.1.24

Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Include figures for each of the following parts.

- (a) Find a formula for $T(a, b)$, where T represents the projection on the y -axis along the x -axis.

Solution. Since $(a, b) = (a, 0) + (0, b)$ and T represents the projection on the y -axis along the x -axis, we must have $T(a, b) = (0, b)$ by definition of projections. ■

- (b) Find a formula for $T(a, b)$, where T represents the projection on the y -axis along the line $L = \{(s, s) : s \in \mathbb{R}\}$.

Solution. Observe that $(a, b) \in \mathbb{R}^2$ can be written in the following form:

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

So, $T(a, b) = (0, b - a)$ by definition of projection. ■

Exercise 2.1.25

Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

- (a) If $T(a, b, c) = (a, b, 0)$, show that T is the projection on the xy -plane along the z -axis.

Solution. Observe that $(a, b, c) \in \mathbb{R}^3$ can be written in the following form:

$$(a, b, c) = \underbrace{(a, b, 0)}_{\in \mathbb{R}^2} + \underbrace{(0, 0, c)}_{z\text{-axis}}. \quad (1)$$

Since $T(a, b, c) = (a, b, 0)$, we know that T represents the projection on the xy -plane along the z -axis. ■

- (b) Find a formula for $T(a, b, c)$, where T represents the projection on the z -axis along the xy -plane.

Solution. Let $(a, b, c) \in \mathbb{R}^3$. Then the projection on the z -axis along the xy -plane should be $T(a, b, c) = (0, 0, c)$ since $(a, b, c) = (a, b, 0) + (0, 0, c)$ where $(a, b, 0) \in \mathbb{R}^2$ and $(0, 0, c)$ is a point on the z -axis. ■

- (c) If $T(a, b, c) = (a - c, b, 0)$, show that T is the projection on the xy -plane along the line $L = \{(a, 0, a) : a \in \mathbb{R}\}$.

Solution. Let $(a, b, c) \in \mathbb{R}^3$. Then we can write this vector in the following way:

$$\begin{aligned} (a, b, c) &= (a, b, 0) + (0, 0, c) \\ &= (a - c + c, b, 0) + (0, 0, c) \\ &= (a - c, b, 0) + (c, 0, c) \end{aligned}$$

where $(a - c, b, 0) \in \mathbb{R}^2$ and $(c, 0, c) \in L$. Since $T(a, b, c) = (a - c, b, 0)$, T must be the projection on the xy -plane along the line L . ■

Exercise 2.1.26

Using the notation in the definition above, assume that $T : V \rightarrow V$ is the projection on W_1 along W_2 .

- (a) Prove that T is linear and $W_1 = \{x \in V : T(x) = x\}$.

Proof. Let $x, y \in V$. We need to show that T is linear. Since T is the projection on W_1 along W_2 , we must have $x = x_1 + x_2$ and $y = y_1 + y_2$ with $T(x) = x_1$ and $T(y) = y_1$ respectively. Observe that

$$\begin{aligned} cx + y &= c(x_1 + x_2) + (y_1 + y_2) \\ &= (cx_1 + y_1) + (cx_2 + y_2). \end{aligned}$$

Let $c \in F$. Since $T : V \rightarrow V$ is the projection on W_1 along W_2 , we must have

$$T(cx + y) = cx_1 + y_1 = cT(x) + T(y).$$

Hence, T is linear.

Now, let's show that $W_1 = \{x \in V : T(x) = x\}$. Denote $S = \{x \in V : T(x) = x\}$. Let $x \in W_1$. Since T is the projection on W_1 along W_2 and $x = x + 0_V$, we have $T(x) = x$. Hence, $W_1 \subseteq S$. Conversely, let $x \in S$. Then $T(x) = x$. Since T is the projection on W_1 along W_2 , this would mean that $x \in W_1$. Hence,

$$W_1 = \{x \in V : T(x) = x\}.$$

■

- (b) Prove that $W_1 = R(T)$ and $W_2 = N(T)$.

Proof. Let's show that $W_1 = R(T)$. Let $x \in W_1$. Since $W_1 \subseteq V$, we know that $x \in V$ where $x = x + 0_V$. Since T is the projection on W_1 along W_2 , we have that $T(x) = x$. Hence, $x \in R(T)$. Conversely, $x \in R(T)$. Then for some $z \in V$, we have $T(x) = z$. Since T is the projection on W_1 along W_2 , we have that $x = z + 0_V$ implies that $z \in W_1$. Thus, $x \in W_1$ and so,

$$W_1 = R(T).$$

Now, let's show that $W_2 = N(T)$. Let $x \in W_2$. Since T is a projection on W_1 along W_2 , we have $x = 0_V + x$ implies $T(x) = 0_V$. This tells us that $x \in N(T)$. On the other hand, assume $x \in N(T)$. Then $T(x) = 0_V$. Since $x \in V$ and $W_1 + W_2 = V$, we have $x = x_1 + x_2$ with $x_1 \in W_1$ and $x_2 \in W_2$. Since $T(x) = 0_V$, either $x_1 = 0$ or $x_2 = 0$. Assume $x_2 = 0$, then $x = x_1$ and $x \in W_1$ which is a contradiction. If $x_1 = 0$, then $x_2 = x$, so $T(x_2) = 0_V$. Hence, $x \in W_2$. Thus,

$$W_2 = N(T).$$

■

- (c) Describe T if $W_1 = V$.

Solution. If $W_1 = V$, then we simply have $T(x) = x$ as described in part (a). Moreover, $N(T) = \{0\}$ since $V = W_1 + W_2$ implies $N(T) = W_2 = \{0\}$. Thus, T must be injective. ■

- (d) Describe T if W_1 is the zero subspace.

Solution. If W_1 is the zero subspace, then $\dim(W_1) = 0$. This implies that $\dim(V) = \dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) + \dim(W_1 \cap W_2) = \dim(W_2)$ (note that $\dim(W_1 \cap W_2) = 0$). So, $\dim(V) = \dim(W_2)$ implies that $V = W_2$. Since T is a projection on W_1 along W_2 , then $T(x) = 0_V$ for all $x \in V$. ■

Exercise 2.1.27

Suppose that W is a subspace of a finite-dimensional vector space V .

- (a) Prove that there exists a subspace W' and a function $T : V \rightarrow V$ such that T is a projection on W along W' .

Proof. Since V is finite-dimensional and W is a subspace of V , we know by Exercise 1.6.34 that there exists a subspace W' such that $W_1 \oplus W_2 = V$. Note that any $x \in V$ can be defined by $x = x_1 + x_2$ where $x_1 \in W$ and $x_2 \in W'$. Define the function

$$T : V \rightarrow V \text{ by } T(x) = x_1.$$

Since $W + W' = V$ and $W \cap W' = \{0\}$, we get a unique representation of any $x \in V$. Thus, we have that T is well-defined and that T is a projection on W along W' . ■

Remark. I am not quite sure how to show that T is a projection here. I am not sure if showing that T is well-defined is enough to do it.

- (b) Give an example of a subspace W of a vector space V such that there are two projections on W along two (distinct) subspaces.

Solution. ■

Definition 2.1.2 (Invariance). Let V be a vector space, and let $T : V \rightarrow V$ be linear. A subspace W of V is said to be **T -invariant** if $T(x) \in W$ for every $x \in W$, that is, $T(W) \subseteq W$. If W is T -invariant, we define the **restriction of T on W** to be the function $T_W : W \rightarrow W$ defined by $T_W(x) = T(x)$ for all $x \in W$.

Exercises 28-32 assume that W is a subspace of a vector space V and that $T : V \rightarrow V$ is linear. *Warning:* Do not assume that W is T -invariant or that T is a projection unless explicitly stated.

Exercise 2.1.28

Prove that the subspaces $\{0\}$, V , $R(T)$, and $N(T)$ are all T -invariant.

- Proof.** (a) Let $x \in \{0\}$. Since T is linear, we must have $x = 0_V$ such that $T(0_V) = 0_V$. Hence, $\{0\}$ is T -invariant.
- (b) Since V is a subspace of itself, and $T(x) \in V$ for all $x \in V$, we have that V is T -invariant.
- (c) Let $x \in R(T)$. Then there exists $w \in V$ such that $T(x) = w \in R(T)$. Hence, $R(T)$ is T -invariant.
- (d) Let $x \in N(T)$. Then by definition of $N(T)$, we have $T(x) = 0_V \in N(T)$. Hence, $N(T)$ is T -invariant.

Exercise 2.1.29

If W is T -invariant, prove that T_W is linear.

Proof. Since W is T -invariant, we have a restriction of T on W such that $T_W : W \rightarrow W$ is defined by $T_W(x) = T(x)$ for all $x \in W$. Let $cx + y \in W$. Since T is linear, we know that

$$\begin{aligned} T_W(cx + y) &= T(cx + y) \\ &= cT(x) + T(y) \\ &= cT_W(x) + T_W(y). \end{aligned}$$

Hence, T_W is linear. ■

Exercise 2.1.30

Suppose that T is the projection on W along some subspace W' . Prove that W is T -invariant and that $T_W = I_W$.

Proof. Let $x \in W$. Since T is the projection on W along some subspace W' , we must have $T(x) = x$. Since $x \in W$, this must imply that $T(x) \in W$ as well and so, we have that W is T -invariant. In fact, the linearity of T and W being T -invariant implies that T is linear (by exercise 29) and that

$$T_W(x) = T(x) = x = I_W(x)$$

for any arbitrary $x \in W$. Hence, we also have that $T_W = I_W$. ■

Exercise 2.1.31

Suppose that $V = R(T) \oplus W$ and W is T -invariant.

(a) Prove that $W \subseteq N(T)$.

Proof. Let $v \in V$. Since V is T -invariant, we must have $T(v) \in V$. Using the fact that $V = R(T) \oplus W$, and that W is T -invariant, $T(v)$ can be expressed in the following way:

$$T(v) = T(y) + T(w) \tag{1}$$

with $T(y) \in R(T)$ and $T(w) \in W$. Subtracting $T(y)$ on both sides and using the linearity of T , we get that

$$T(v) - T(y) = T(w) \Leftrightarrow T(v - y) = T(w).$$

Note that $T(v - y) \in R(T)$ and $T(w) \in W$. Since $R(T) \cap W = \{0\}$ and $T(w) \in R(T) \cap W$, this means that $T(w) = 0$. Hence, $W \subseteq N(T)$. ■

(b) Show that if V is finite-dimensional, then $W = N(T)$.

Proof. Since V is finite-dimensional, we can use the dimension theorem to write

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

Furthermore, we can use Exercise 1.6.29 and that $\dim(R(T) \cap W) = 0$ to write that

$$\begin{aligned}\dim(V) &= \dim(R(T) + W) \\ &= \text{rank}(T) + \dim(W) - \dim(R(T) \cap W) \\ &= \text{rank}(T) + \dim(W).\end{aligned}$$

Equating both equations we find that

$$\text{nullity}(T) + \text{rank}(T) = \text{rank}(T) + \dim(W)$$

which implies that

$$\text{nullity}(T) = \dim(W).$$

By theorem 1.11, we get that $N(T) = W$. ■

- (c) Show by example that the conclusion of (b) is not necessarily true if V is not finite-dimensional.

Proof. ■

Exercise 2.1.32

Suppose that W is T -invariant. Prove that $N(T_W) = N(T) \cap W$ and $R(T_W) = T(W)$.

Proof. First, we show that $N(T_W) = N(T) \cap W$. Let $x \in N(T_W)$. Then $x \in W$ implies that $T_W(x) = 0_W$ by definition of $N(T_W)$. Since W is T -invariant, we get that $T_W(x) = T(x)$. Since $T(W) \subseteq W$, this tells us that $x \in W$. But note that $T(x) = 0_W$. Hence, we must also have $x \in N(T)$. So, $x \in N(T)$ and $x \in W$ implies that $x \in N(T) \cap W$ and thus $N(T_W) \subseteq N(T) \cap W$. Now, let $x \in N(T) \cap W$. Then $x \in N(T)$ and $x \in W$. Since $x \in N(T)$, we must have $T(x) = 0_W$. Since $x \in W$ and W is T -invariant, we must also have $T(x) = T_W(x) = 0_W$. Thus, $x \in N(T_W)$ and we get that $N(T) \cap W \subseteq N(T_W)$. We conclude that

$$N(T_W) = N(T) \cap W.$$

Now, let us show that $R(T_W) = T(W)$. Let $w \in R(T_W)$. Then there exists $x \in W$ such that $T_W(x) = w$. Since W is T -invariant, we have $T_W(x) = T(x)$ and thus $w \in T(W)$. Thus, $R(T_W) \subseteq T(W)$. Conversely, let $w \in T(W)$. It follows immediately that $w \in R(T)$. Then there exists $x \in W$ such that $T(x) = w$. Since W is T -invariant, we must have $T(x) = T_W(x)$. Hence, $w \in R(T_W)$ and thus $T(W) \subseteq R(T_W)$. Thus, we have

$$R(T_W) = T(W).$$
■

Exercise 2.1.33

Prove Theorem 2.2 for the case that β is infinite, that is, $R(T) = \text{span}(\{T(v) : v \in \beta\})$.

Proof. We will show that

$$R(T) = \text{span}(\{T(v) : v \in \beta\});$$

that is, we will show that both $R(T) \subseteq \text{span}(\{T(v) : v \in \beta\})$ and $\text{span}(\{T(v) : v \in \beta\}) \subseteq R(T)$. Since $T(v) \in R(T)$ for every $v \in \beta$, we have that $R(T)$ must also contain the span of $\{T(v) : v \in \beta\}$ by Theorem 1.5. Hence, we have $\text{span}(\{T(v) : v \in \beta\}) \subseteq R(T)$.

On the other hand, let $x \in R(T)$. Then there exists some $v \in R(T)$ such that $T(v) = x$.

Since β is a basis for V , we can write $x \in V$ as a linear combination of vectors in β ; that is, we can choose scalars a_1, a_2, a_3, \dots such that

$$v = \sum_{i \geq 1} a_i v_i$$

where $v_i \in \beta$ for $i \geq 1$. Since T is linear, we can write

$$\begin{aligned} T(v) &= T\left(\sum_{i \geq 1} a_i v_i\right) \\ &= \sum_{i \geq 1} a_i T(v_i). \end{aligned}$$

But this tells us that $w \in \text{span}(\{T(v) : v \in \beta\})$ and thus we have $R(T) \subseteq \text{span}(\{T(v) : v \in \beta\})$. ■

Exercise 2.1.34

Prove the following generalization of Theorem 2.6: Let V and W be vector spaces over a common field, and let β be a basis for V . Then for any function $f : \beta \rightarrow W$ there exists exactly one linear transformation $T : V \rightarrow W$ such that $T(x) = f(x)$ for all $x \in \beta$.

Proof. Let $x \in V$. Since β is a basis for V , we can find $a_1, a_2, a_3, \dots \in F$ such that

$$x = \sum_{i \geq 1} a_i v_i$$

for v_1, v_2, v_3, \dots in β . Let $f : \beta \rightarrow W$. Define the map $T : V \rightarrow W$ such that

$$T(x) = \sum_{i \geq 1} a_i f(v_i)$$

with $f(v_i) \in W$ for $i \geq 1$. We claim that T is linear. Let $cx + y \in V$ with $c \in F$. Then using β as a basis for V , we get that

$$x = \sum_{i \geq 1} a_i v_i$$

and

$$y = \sum_{i \geq 1} b_i v_i$$

for $a_i, b_i \in F$ for $i \geq 1$. Then observe that

$$\begin{aligned} T(cx + y) &= \sum_{i \geq 1} (ca_i + b_i) f(v_i) \\ &= \sum_{i \geq 1} (ca_i) f(v_i) + \sum_{i \geq 1} b_i f(v_i) \\ &= c \sum_{i \geq 1} a_i f(v_i) + \sum_{i \geq 1} b_i f(v_i) \\ &= cT(x) + T(y). \end{aligned}$$

Hence, T is linear. Now, we claim that T is a unique linear transformation. Suppose there exists $U : V \rightarrow W$ defined by

$$U(x) = \sum_{i \geq 1} a_i w_i$$

with $w_i \in W$. Note $w_i \in W$ implies that $f(v_i) = w_i$ for $i \geq 1$. Thus,

$$U(x) = \sum_{i \geq 1} a_i w_i = \sum_{i \geq 1} a_i f(v_i) = T(x).$$

Since $x \in V$ is arbitrary, we must have $U = T$. Hence, T is unique. ■

Exercises 35 and 36 assume the definition of *direct sum* given in the exercises of Section 1.3.

Exercise 2.1.35

Let V be a finite-dimensional vector space $T : V \rightarrow V$ be linear.

- (a) Suppose that $V = R(T) + N(T)$. Prove that $V = R(T) \oplus N(T)$.

Proof. Since V is finite-dimensional, the subspaces $R(T)$ and $N(T)$ are finite-dimensional by Theorem 1.11. Using the Dimension Theorem and the fact that T is linear, we can write

$$\begin{aligned} \dim(V) &= \text{rank}(T) + \text{nullity}(T) \\ &= \dim(R(T)) + \dim(N(T)). \end{aligned}$$

Since $V = R(T) + N(T)$, we can use Exercise 1.6.29 to conclude that

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

■

- (b) Suppose that $R(T) \cap N(T) = \{0\}$. Prove that $V = R(T) \oplus N(T)$.

Proof. Since V is finite-dimensional and $R(T)$ and $N(T)$ are subspaces of V , we know by Theorem 1.11 that $R(T)$ and $N(T)$ are also finite-dimensional. Thus, $R(T) + N(T)$ is subspace and is also finite-dimensional by part (a) of Exercise 1.6.29. Since $R(T) \cap N(T) = \{0\}$, we can write

$$\begin{aligned} \dim(R(T) + N(T)) &= \dim(R(T)) + \dim(N(T)) - \dim(R(T) \cap N(T)) \\ &= \text{rank}(T) + \text{nullity}(T). \end{aligned}$$

Since V is finite-dimensional and T is linear, the Dimension Theorem states that

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

But this tells us that

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

This implies that $R(T) + N(T) = V$ by Theorem 1.11. Together with $R(T) \cap N(T) = \{0\}$, we have that

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

■

Be careful to say in each part where finite-dimensional is used.

Exercise 2.1.36

Let V and T be as defined in Exercise 21.

- (a) Prove that $V = R(T) + N(T)$, but V is not a direct sum of these two spaces. Thus the result of Exercise 35(a) above cannot be proved without assuming that V is finite-dimensional.
- (b) Find a linear operator T_1 on V such that $R(T_1) \cap N(T_1) = \{0\}$ but V is not a direct sum of $R(T_1)$ and $N(T_1)$. Conclude that V being finite-dimensional is also essential in Exercise 35(b).

Definition 2.1.3 (Additive). A function $T : V \rightarrow W$ between vector spaces V and W is called **additive** if $T(x + y) = T(x) + T(y)$ for all $x, y \in V$.

Exercise 2.1.37

Prove that if V and W are vector spaces over the field of rational numbers, then any additive function from V into W is a linear transformation.

Proof. Let $T : V \rightarrow W$ be an additive function from V into W where V and W are vector spaces over \mathbb{Q} . Let $c \in \mathbb{Q}$ and let $x, y \in V$. Then we can find $u = T(cx)$ and $v = T(x)$ such that

$$c = \frac{u}{v} = \frac{T(cx)}{T(x)} \Leftrightarrow cT(x) = T(cx)$$

with $v \neq 0$. Using the additivity of T , we can write

$$\begin{aligned} T(cx + y) &= T(cx) + T(y) \\ &= cT(x) + T(y). \end{aligned}$$

Hence, T is linear. ■

Exercise 2.1.38

Let $T : \mathbb{C} \rightarrow \mathbb{C}$ be the function defined by $T(z) = \bar{z}$. Prove that T is additive but not linear.

Proof. Let $x, y \in \mathbb{C}$ such that $x = \delta_1 + \delta_2 i$ and $y = \varepsilon_1 + \varepsilon_2 i$ for $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2 \in \mathbb{R}$. By definition of T , we know that

$$T(x) = \bar{x} = \delta_1 - \delta_2 i$$

and

$$T(y) = \bar{y} = \varepsilon_1 - \varepsilon_2 i.$$

Note that

$$x + y = (\delta_1 + \varepsilon_1) + (\delta_2 + \varepsilon_2)i.$$

By definition of T , we have that

$$\begin{aligned} T(x + y) &= (\delta_1 + \varepsilon_1) - (\delta_2 + \varepsilon_2)i \\ &= (\delta_1 - \delta_2 i) + (\varepsilon_1 - \varepsilon_2 i) \\ &= T(x) + T(y). \end{aligned}$$

Hence, T is an additive function.

Let $z \in \mathbb{C}$ where $z = a + bi$. Observe that $\bar{z} = a - bi$. Then

$$f(z\bar{z}) = f(a^2 + b^2) = a^2 + b^2$$

but

$$zf(\bar{z}) = (a + bi)^2 = a^2 + 2abi - b^2.$$

Hence, $f(z\bar{z}) \neq zf(\bar{z})$ implies that f is not linear. ■

Exercise 2.1.40

Let V be a vector space and W be a subspace of V . Define the mapping $\eta : V \rightarrow V/W$ by $\eta(v) = v + W$ for $v \in V$.

- (a) Prove that η is a linear transformation from V onto V/W and that $N(\eta) = W$.

Proof. Let us first show that $\eta : V \rightarrow V/W$ is a linear transformation. Let $cx + y \in V$. By the operations defined on η , we must have

$$\begin{aligned} \eta(cx + y) &= (cx + y) + W \\ &= (cx + W) + (y + W) \\ &= c(x + W) + (y + W) \\ &= c\eta(x) + \eta(y). \end{aligned}$$

Hence, η is a linear transformation.

Now, let us show that η is onto. Observe that $R(\eta) \subseteq V/W$. Let $x \in V/W$. Then $x = v + W$ for $v \in V$. By definition of $\eta : V \rightarrow V/W$, we have $\eta(v) = v + W$. This means that $\eta(v) = x$ which tells us that $x \in R(\eta)$. Hence, we conclude that $R(\eta) = V/W$ and hence η is onto.

Lastly, we will prove that $N(\eta) = W$. Let $x \in N(\eta)$. Then $\eta(x) = W$. By definition of η , we must have that $\eta(x) = x + W$. Hence, $x + W = W$. By Exercise 1.3.31, we get that $x \in W$. Thus, $N(\eta) \subseteq W$. On the other hand, $x \in W$ implies that $W = x + W$. By definition of η , we get that $\eta(x) = W$. Since W is the zero vector defined on V/W , we must have $x \in N(\eta)$ and hence $W \subseteq N(\eta)$. Thus, $N(\eta) = W$. ■

- (b) Suppose that V is finite-dimensional. Use (a) and the dimension theorem to derive a formula relating $\dim(V)$, $\dim(W)$, and $\dim(V/W)$.

Proof. Since V is finite-dimensional, we can write

$$\dim(V) = \text{rank}(\eta) + \text{nullity}(\eta). \tag{1}$$

Since η is an onto function by part (a), we have $R(\eta) = V/W$ implies $\text{rank}(\eta) = \dim(V/W)$ by Theorem 1.11. Furthermore, $N(\eta) = W$ implies that $\text{nullity}(\eta) = \dim(W)$. Hence, (1) can be re-written in the following way:

$$\dim(V) = \dim(V/W) + \dim(W).$$

Thus, we have our desired result

$$\dim(V/W) = \dim(V) - \dim(W).$$
■

2.2 The Matrix Representation of a Linear Transformation

Exercise 2.2.1

Label the following statements as true or false. Assume that V and W are finite-dimensional vector spaces with ordered bases β and γ , respectively, and $T, U : V \rightarrow W$ are linear transformations.

- (a) For any scalar a , $aT + U$ is a linear transformation from V to W .

Solution. True by Theorem 2.7. ■

- (b) $[T]_{\beta}^{\gamma} = [U]_{\beta}^{\gamma}$ implies $T = U$.

Solution. True by corollary to Theorem 2.6. ■

- (c) If $m = \dim(V)$ and $n = \dim(W)$, then $[T]_{\beta}^{\gamma}$ is an $m \times n$ matrix.

Solution. True. ■

- (d) $[T + U]_{\beta}^{\gamma} = [T]_{\beta}^{\gamma} + [U]_{\beta}^{\gamma}$.

Solution. True by Theorem 2.8. ■

- (e) $\mathcal{L}(V, W)$ is a vector space.

Solution. True. ■

- (f) $\mathcal{L}(V, W) = \mathcal{L}(W, V)$.

Solution. Unless $V = W$, this statement is false. ■

Exercise 2.2.6

Complete the proof of part (b) of Theorem 2.7.

Proof. See proof in notes. ■

Exercise 2.2.7

Prove part (b) of Theorem 2.8.

Proof. See proof in notes. ■

Exercise 2.2.8

Let V be an n -dimensional vector space with an ordered basis β . Define $T : V \rightarrow F^n$ by $T(x) = [x]_{\beta}$. Prove that T is linear.

Proof. Let $x, y \in V$. Since β is an ordered basis for V , we have distinct vectors $v_1, v_2, \dots, v_n \in$

β where

$$x = \sum_{i=1}^n a_i v_i$$

and

$$y = \sum_{i=1}^n b_i v_i$$

for scalars a_1, a_2, \dots, a_n and b_1, b_2, \dots, b_n , respectively. Let $c \in F$. Then we have $cx + y \in V$ implies

$$cx + y = \sum_{i=1}^n (ca_i + b_i) v_i.$$

By definition T , we must have

$$\begin{aligned} T(cx + y) &= [cx + y]_\beta \\ &= \begin{pmatrix} ca_1 + b_1 \\ ca_2 + b_2 \\ \vdots \\ ca_n + b_n \end{pmatrix} \\ &= c \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\ &= c[x]_\beta + [y]_\beta \\ &= cT(x) + T(y). \end{aligned}$$

Hence, T is a linear map. ■

Exercise 2.2.9

Let V be the vector space of complex numbers over the field \mathbb{R} . Define $T : V \rightarrow V$ by $T(z) = \bar{z}$, where \bar{z} is the complex conjugate of z . Prove that T is linear, and compute $[T]_\beta$, where $\{1, i\}$. (Compare this to Exercise 2.1.38)

Proof. Let $cx + y \in \mathbb{C}$ where $c \in \mathbb{R}$ and $x, y \in \mathbb{C}$. Observe that

$$x = a + bi \text{ and } y = v + wi$$

for $a, b, v, w \in \mathbb{R}$. So,

$$cx + y = (ca + v) + (cb + w)i$$

Furthermore, By definition of T and definition of conjugate, we write

$$\begin{aligned} T(cx + y) &= \overline{cx + y} \\ &= (ca + v) - (cb + w)i \\ &= c(a - bi) + (v - wi) \\ &= c\bar{x} + \bar{y} \\ &= cT(x) + T(y). \end{aligned}$$

Hence, T is linear. Now, let's compute $[T]_\beta$ with $\beta = \{1, i\}$ as our ordered basis for \mathbb{C} . So,

$$T(1) = 1 = 1 \cdot 1 + 0 \cdot i \text{ and } T(i) = -i = 0 \cdot 1 - 1 \cdot i.$$

Hence,

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

■

Exercise 2.2.10

Let V be a vector space with the ordered basis $\beta = \{v_1, v_2, \dots, v_n\}$. Define $v_0 = 0$. By Theorem 2.6, there exists a linear transformation $T : V \rightarrow V$ such that $T(v_j) = v_j + v_{j-1}$ for $j = 1, 2, \dots, n$. Compute $[T]_\beta$.

Solution. For $1 \leq j \leq n$, we see that

$$T(v_1) = v_1 + v_0 = v_1$$

$$T(v_2) = v_2 + v_1$$

$$T(v_3) = v_3 + v_2$$

$$\vdots$$

$$T(v_n) = v_n + v_{n-1}.$$

Then we have

$$[T]_\beta = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \ddots & 1 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

■

Exercise 2.2.11

Let V be an n -dimensional vector space, and let $T : V \rightarrow V$ be a linear transformation. Suppose that W is a T -invariant subspace of V having dimension k . Show that there is a basis β for V such that $[T]_\beta$ has the form

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix},$$

where A is a $k \times k$ matrix and O is the $(n - k) \times k$ zero matrix.

Proof. Let $\dim(V) = n$. Since W is a subspace of V , let $\alpha = \{w_1, w_2, \dots, w_k\}$ be an ordered basis for W . By corollary to Theorem 1.11, we extend α to a basis for V by adding distinct and linear independent vectors $w_{k+1}, w_{k+2}, \dots, w_n$. Denote this basis for V as β with

$$\beta = \{w_1, w_2, \dots, w_n\}.$$

Hence, for $1 \leq i \leq n$ we have

$$T(w_j) = \sum_{i=1}^n a_{ij} w_i \quad \text{for } 1 \leq j \leq n.$$

Since W is T -invariant, we know that $T(w_j)$ for every $w_j \in W$ with $1 \leq j \leq k$. Hence, $T(w_j)$ for $1 \leq j \leq k$ can be written as a linear combination of vectors in α ; that is,

$$T(w_j) = \sum_{i=1}^k a_{ij} w_i \quad \text{for } 1 \leq j \leq k.$$

Since β is a basis for V , we know that for $k+1 \leq i \leq n$, we know that

$$T(w_j) = \sum_{i=1}^k a_{ij} w_i + \sum_{i=k+1}^n a_{ij} w_i = \sum_{i=1}^k a_{ij} w_i + 0.$$

where $a_{ij} = 0$ for $k+1 \leq i \leq n$. Notice that the second term above, represents the O matrix that $(n-k) \times k$. On the other hand, for $1 \leq i \leq k$ and $1 \leq j \leq k$ we can see that $T : W \rightarrow W$ (since W is T -invariant) can be represented as the matrix A that is $k \times k$. Then for $1 \leq i \leq n$ and $1 \leq j \leq k$, we get that

$$[T(w_j)]_{\beta} = \begin{pmatrix} A \\ O \end{pmatrix}.$$

Now, if we let $k+1 \leq j \leq n$, then we would see that $T(w_j)$ can be written as

$$T(w_j) = \sum_{i=1}^n a_{ij} w_i.$$

Together with $1 \leq i \leq n$ and $k+1 \leq j \leq n$ where

$$[T(w_j)]_{\beta} = (T(w_{k+1}) \quad T(w_{k+2}) \quad \cdots \quad T(w_n))$$

which can be denoted with B as a $k \times (n-k)$ matrix and C as a $(n-k) \times (n-k)$ matrix where

$$[T(w_j)]_{\beta} = \begin{pmatrix} B \\ C \end{pmatrix}$$

Hence, $[T]_{\beta}$ has the following matrix form

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix}.$$

■

Exercise 2.2.12

Let V be a finite-dimensional vector space and T be the projection on W along W' , where W and W' are subspaces of V . Find an ordered basis β for V such that $[T]_{\beta}$ is a diagonal matrix.

Solution. Let $\dim(V) = n$ since V is a finite-dimensional vector space. Since $W \subseteq V$ is a subspace, we know that W must be finite-dimensional as well. Let $\dim(W) = k$ and let

$\lambda = \{w_1, w_2, \dots, w_k\}$ be a basis for W . By corollary to Theorem 1.11, we can extend λ to be a basis for V by adding $w_{k+1}, w_{k+2}, \dots, w_n \in W$ into λ . Denote this ordered basis for V as β where

$$\beta = \{w_1, w_2, \dots, w_n\}.$$

Since W is a $T : V \rightarrow V$ is a projection on W along W' (W' is also a subspace), then

$$w_j = T(w_j) = \sum_{i=1}^n a_{ij} w_i \quad \text{for } 1 \leq j \leq n$$

for every $w_j \in W$. This tells us that each w_j can be expressed a linear combination if and only if $a_{ij} = 1$ whenever $i = j$ and $a_{ij} = 0$ otherwise. Hence, $[T]_\beta$ is a diagonal matrix. ■

Exercise 2.2.13

Let V and W be vector spaces, and let T and U be nonzero linear transformations from V to W . If $R(T) \cap R(U) = \{0\}$, prove that $\{T, U\}$ is a linearly independent subset of $\mathcal{L}(V, W)$.

Proof. Let T and U be nonzero linear transformations from V to W . For $a, b \in F$, we need to show that

$$aT + bU = T_0$$

with a and b both zero. Note that T_0 is the zero linear transformation. Let $x \in V$. Then we have

$$\begin{aligned} (aT + bU)(x) &= T_0(x) \\ (aT)(x) + (bU)(x) &= 0. \end{aligned}$$

Hence, we have

$$(aT)(x) = -(bU)(x).$$

Since $(aT)(x) \in R(T) \cap R(U)$ where $R(T) \cap R(U) = \{0\}$ by assumption, we get that

$$(aT)(x) = 0 \Leftrightarrow aT(x) = 0.$$

Since $T(x) \neq 0$, we must have that $a = 0$ when dividing $T(x)$ on both sides on the equation above. This also implies that $b = 0$ since $U(x) \neq 0$ and so $\{T, U\}$ must be linearly independent. ■

Exercise 2.2.14

Let $V = P(\mathbb{R})$, and for $j \geq 1$ define $T_j(f(x)) = f^{(j)}(x)$, where $f^{(j)}(x)$ is the j th derivative of $f(x)$. Prove that the set $\{T_1, T_2, \dots, T_n\}$ is a linearly independent subset of $\mathcal{L}(V)$ for any positive integer n .

Proof. Let $j \geq 1$. We want to show that for scalars $a_1, a_2, \dots, a_n \in \mathbb{R}$, we have

$$a_1 T_1 + a_2 T_2 + \dots + a_n T_n = T_0.$$

Let $f(x) \in P(\mathbb{R})$ be arbitrary. Since each T_j is linear, we have that

$$a_1 T_1(f(x)) + a_2 T_2(f(x)) + \dots + a_n T_n(f(x)) = 0$$

which can be re-written to be

$$a_1 f^{(1)}(x) + a_2 f^{(2)}(x) + \cdots + a_n f^{(n)}(x) = 0 \quad (1)$$

with each $f^{(j)}(x)$ being the j th derivative of $f(x)$. Since no two polynomials in the set $\{f^{(1)}(x), f^{(2)}(x), \dots, f^{(n)}(x)\}$ have the same degree, we know that the representation in (1) contains the trivial solution; that is, $a_n = 0$ for any $n \in \mathbb{N}$. Hence, the set

$$\{T_1, T_2, \dots, T_n\}$$

must be linearly independent by Exercise 1.5.18. ■

Exercise 2.2.15

Let V and W be vector spaces, and let S be a subset of V . Define $S^0 = \{T \in \mathcal{L}(V, W) : T(x) = 0 \text{ for all } x \in S\}$. Prove the following statements.

(a) S^0 is a subspace of $\mathcal{L}(V, W)$.

Proof. (i) Note that $T_0 \in S^0$ since $T_0(x) = 0$ for all $x \in S$.

(ii) Let $T, U \in S^0$. Then $T(x) = 0$ and $U(x) = 0$ for all $x \in S$. Then

$$(T + U)(x) = T(x) + U(x) = 0 + 0 = 0.$$

So $T + U \in S^0$.

(iii) Let $c \in F$ and $T \in S^0$. Then $T(x) = 0$ for all $x \in S$. Thus,

$$(cT)(x) = cT(x) = c \cdot 0 = 0.$$

So, $cT \in S^0$.

Hence, S^0 is a subspace of $\mathcal{L}(V, W)$. ■

(b) If S_1 and S_2 are subsets of V and $S_1 \subseteq S_2$, then $S_2^0 \subseteq S_1^0$.

Proof. Let $x_1 \in S_1$. Since $S_1 \subseteq S_2$, we have that $x_1 \in S_2$. If $T \in S_2^0$, then $T(x_1) = 0$ for $x_1 \in S_2$. Since $x_1 \in S_1$, we must also have $T \in S_1^0$. Hence, $S_2^0 \subseteq S_1^0$. ■

(c) If V_1 and V_2 are subspaces of V , then $(V_1 + V_2)^0 = V_1^0 \cap V_2^0$.

Proof. Since $V_1 + V_2$ and $V_1 \cap V_2$ are subsets of V and that $V_1 \cap V_2 \subseteq V_1 + V_2$, we know that $(V_1 + V_2)^0 \subseteq V_1^0 \cap V_2^0$.

Now, let $T \in V_1^0 \cap V_2^0$. Then $T \in V_1^0$ and $T \in V_2^0$ implies $T(x_1) = 0$ and $T(x_2) = 0$ for all $x_1 \in V_1$ and $x_2 \in V_2$. Since T is linear, we must have

$$0 = 0 + 0 = T(x_1) + T(x_2) = T(x_1 + x_2).$$

Hence, $T \in (V_1 + V_2)^0$ and so $V_1^0 \cap V_2^0 \subseteq (V_1 + V_2)^0$. Thus,

$$(V_1 + V_2)^0 = V_1^0 \cap V_2^0. \quad \blacksquare$$

Exercise 2.2.16

Let V and W be vector spaces such that $\dim(V) = \dim(W)$, and let $T : V \rightarrow W$ be linear. Show that there exists ordered bases β and γ for V and W , respectively, such that $[T]_{\beta}^{\gamma}$ is a diagonal matrix.

Proof. Using the same process found in the dimension theorem, let $\lambda = \{v_1, v_2, \dots, v_k\}$ be a basis for $N(T)$. We can extend λ into a basis for V by adding distinct linearly independent vectors $v_{k+1}, v_{k+2}, \dots, v_n$ into λ . Denote this new basis as β where

$$\beta = \{v_1, v_2, \dots, v_n\}.$$

Since $\dim(V) = \dim(W)$, we get that

$$\dim(W) = \dim(N(T)) + \dim(R(T)). \quad (1)$$

Note that $\{T(v_{k+1}), T(v_{k+2}), \dots, T(v_n)\}$ is a basis for $R(T)$ which can be extended to be a basis for W . Hence, we have γ defined by

$$\gamma = \{T(v_1), T(v_2), \dots, T(v_n)\}$$

with $T(v_i) = w_i$ for $1 \leq i \leq n$ to a basis for W . Now, note that for $1 \leq j \leq n$, we have

$$T(v_j) = \sum_{i=1}^n a_{ij}w_i = \sum_{i=1}^k a_{ij}w_i + \sum_{i=k+1}^n a_{ij}w_i.$$

By definition of $N(T)$, we know that for $1 \leq j \leq k$ and $1 \leq i \leq k$ that

$$0 = \sum_{i=1}^k a_{ij}w_i.$$

For $k+1 \leq i \leq n$ and $k+1 \leq j \leq n$,

$$w_j = T(v_j) = \sum_{i=k+1}^n a_{ij}w_i$$

so we have $a_{ij} = 1$ whenever $i = j$ and $a_{ij} = 0$ otherwise. So, the matrix representation of T can be written as

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} O & O \\ O & I \end{pmatrix}$$

where I is the $(n-k) \times (n-k)$ matrix. Thus, we can see that $[T]_{\beta}^{\gamma}$ is a diagonal matrix. ■

2.3 Composition of Linear Transformations and Matrix Multiplication

Exercise 2.3.5

Complete the proof of Theorem 2.12 and its corollary.

Proof. Let A be an $m \times n$ matrix, B and C be $n \times p$ matrices, and D and E be $q \times m$ matrices. Then

(a) Let $1 \leq i \leq m$ and $1 \leq j \leq p$. By definition of the product of two matrices, we have

$$\begin{aligned}(A(B + C))_{ij} &= \sum_{k=1}^n A_{ik}(B + C)_{kj} \\ &= \sum_{k=1}^n A_{ik}(B_{kj} + C_{kj}) \\ &= \sum_{k=1}^n A_{ik}B_{kj} + \sum_{k=1}^n A_{ik}C_{kj} \\ &= (AB)_{ij} + (AC)_{ij}.\end{aligned}$$

Hence, $A(B + C) = AB + AC$.

Now, let $1 \leq i \leq q$ and $1 \leq j \leq n$. For the second formula, we can use the same definition to write

$$\begin{aligned}((D + E)A)_{ij} &= \sum_{k=1}^m (D + E)_{ik}A_{kj} \\ &= \sum_{k=1}^m (D_{ik} + E_{ik})A_{kj} \\ &= \sum_{k=1}^m D_{ik}A_{kj} + \sum_{k=1}^m E_{ik}A_{kj} \\ &= (DA)_{ij} + (EA)_{ij}.\end{aligned}$$

Hence, $(D + E)A = DA + EA$.

(b) Let $1 \leq i \leq m$ and $1 \leq j \leq p$. Let $a \in F$. Then using the definition of the product once again, we have

$$\begin{aligned}a(AB)_{ij} &= a \sum_{k=1}^n A_{ik}B_{kj} \\ &= \sum_{k=1}^n a(A_{ik}B_{kj}) \\ &= \sum_{k=1}^n (aA_{ik})B_{kj} \\ &= \sum_{k=1}^n (aA)_{ik}B_{kj} \\ &= ((aA)B)_{ij}.\end{aligned}$$

Then observe that

$$\begin{aligned}
 ((aA)B)_{ij} &= \sum_{k=1}^n (aA_{ik})B_{kj} \\
 &= \sum_{k=1}^n (A_{ik}a)B_{kj} \\
 &= \sum_{k=1}^n A_{ik}(aB_{kj}) \\
 &= \sum_{k=1}^n A_{ik}(aB)_{kj} \\
 &= (A(aB))_{ij}.
 \end{aligned}$$

Hence, $(aA)B = A(aB)$. Thus, we conclude that

$$a(AB) = (aA)B = A(aB).$$

- (c) Let $1 \leq i \leq m$ and $1 \leq j \leq n$. Since $\delta_{ik} = 1$ only when $i = k$ and 0 otherwise, we must have

$$(I_m A)_{ij} = \sum_{k=1}^m \delta_{ik} A_{kj} = A_{ij}.$$

Likewise,

$$(A I_n)_{ij} = \sum_{k=1}^n A_{ik} \delta_{kj} = A_{ij}$$

by the same reasoning. Hence, we have

$$I_m A = A = A I_n.$$

- (d) Let V be an n -dimensional vector space with $\beta = \{v_1, v_2, \dots, v_n\}$ as an ordered basis. Since $I_V(v_j) = v_j$ for all $1 \leq j \leq n$, we must have

$$v_j = I_V(v_j) = \sum_{i=1}^n a_{ij} v_i$$

which holds only if $a_{ij} = 1$ for all $i = j$ and 0 otherwise. But this means that $a_{ij} = \delta_{ij}$, so $[I_V]_\beta = I_n$. ■

Proof. Let A be an $m \times n$ matrix and B_1, B_2, \dots, B_k be $n \times p$ matrices. Let $1 \leq \ell \leq m$

and $1 \leq s \leq p$. Then

$$\begin{aligned}
\left[A \left(\sum_{i=1}^k a_i B_i \right) \right]_{\ell s} &= \sum_{\lambda=1}^n A_{\ell \lambda} \left(\sum_{i=1}^k a_i B_i \right)_{\lambda s} \\
&= \sum_{\lambda=1}^n A_{\ell \lambda} \left(\sum_{i=1}^k (a_i B_i)_{\lambda s} \right) \\
&= \sum_{\lambda=1}^n A_{\ell \lambda} \left(\sum_{i=1}^k a_i (B_i)_{\lambda s} \right) \\
&= \sum_{i=1}^k a_i \left(\sum_{\lambda=1}^n A_{\ell \lambda} (B_i)_{\lambda s} \right) && \text{(part (a) of Theorem 2.12)} \\
&= \sum_{i=1}^k a_i (AB_i)_{\ell s}.
\end{aligned}$$

Hence, we have

$$A \left(\sum_{i=1}^k a_i B_i \right) = \sum_{i=1}^k a_i AB_i.$$

To show the second formula, let $1 \leq \ell \leq q$ and $1 \leq s \leq n$. Then

$$\begin{aligned}
\left[\left(\sum_{i=1}^k a_i C_i \right) A \right]_{\ell s} &= \sum_{\lambda=1}^n \left(\sum_{i=1}^k a_i C_i \right)_{\ell \lambda} A_{\lambda s} \\
&= \sum_{\lambda=1}^n \left(\sum_{i=1}^k (a_i C_i)_{\ell \lambda} \right) A_{\lambda s} \\
&= \sum_{\lambda=1}^n \left(\sum_{i=1}^k a_i (C_i)_{\ell \lambda} \right) A_{\lambda s} \\
&= \sum_{i=1}^k a_i \left(\sum_{\lambda=1}^n (C_i)_{\ell \lambda} A_{\lambda s} \right) && \text{(part (a) of Theorem 2.12)} \\
&= \sum_{i=1}^k a_i (C_i A)_{\ell s}.
\end{aligned}$$

Hence, we have

$$\left(\sum_{i=1}^k a_i C_i \right) A = \sum_{i=1}^k a_i C_i A.$$

■

Exercise 2.3.6

Prove (b) of Theorem 2.13.

Proof. To show the other equation, we apply part (c) of Theorem 2.12, to write

$$v_j = \begin{pmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{nj} \end{pmatrix} = \begin{pmatrix} (BI_p)_{1j} \\ (BI_p)_{2j} \\ \vdots \\ (BI_p)_{nj} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n B_{1k}\delta_{kj} \\ \sum_{k=1}^n B_{2k}\delta_{kj} \\ \vdots \\ \sum_{k=1}^n B_{nk}\delta_{kj} \end{pmatrix} = B \begin{pmatrix} \delta_{1j} \\ \delta_{2j} \\ \vdots \\ \delta_{pj} \end{pmatrix} = Be_j$$

where δ_{1j} are the Kronecker delta constants. ■

Exercise 2.3.7

Prove (c) and (f) of Theorem 2.15.

Proof. (c) Using the sum rule for matrices, we must have

$$L_{A+B}(x) = (A+B)(x) = A(x) + B(x) = L_A(x) + L_B(x).$$

Hence, $L_{A+B} = L_A + L_B$. Now, let $a \in F$. Using the same reasoning, we have

$$L_{aA}(x) = (aA)(x) = a(A(x)) = aL_A(x).$$

Hence, $L_{aA} = aL_A$.

(f) Let $1 \leq j \leq n$. Then

$$\begin{aligned} L_{I_n}(e_j) &= I_n(e_j) = e_j \\ &= I_{F^n}(e_j). \end{aligned} \quad \text{(Part (d) of Theorem 2.3.4)}$$

Hence, $L_{I_n} = I_{F^n}$. ■

Exercise 2.3.8

Prove Theorem 2.10. Now state and prove a more general result involving linear transformations with domains unequal to their codomains.

Proposition 2.3.1. Let V, W, Y , and Z be vector spaces. Then we have the following properties:

(a) Let $U, U_1, U_2 \in \mathcal{L}(V, W)$ and $T, T_1, T_2 \in \mathcal{L}(W, Z)$. Then we have

$$T(U_1 + U_2) = TU_1 + TU_2 \text{ and } (T_1 + T_2)U = T_1U + T_2U.$$

(b) Let $U \in \mathcal{L}(Y, Z), T_1 \in \mathcal{L}(W, Y), T_2 \in \mathcal{L}(V, W)$. Then we have

$$U(T_1T_2) = (UT_1)T_2.$$

(c) Let $I \in \mathcal{L}(V, V)$ and $I \in \mathcal{L}(W, W)$. Then $T \in \mathcal{L}(V, W)$ implies that

$$TI = IT = T.$$

(d) Let $a \in F$ and let $U_1, U_2 \in \mathcal{L}(V, W)$. Then

$$a(U_1 U_2) = (aU_1)U_2 = U_1(aU_2).$$

Proof. The proof is similar in the notes. ■

Exercise 2.3.10

Let A be an $n \times n$ matrix. Prove that A is a diagonal matrix if and only if $A_{ij} = \delta_{ij} A_{ij}$ for all i and j .

Proof. For the forwards direction, suppose A is a diagonal matrix. Let $1 \leq j \leq n$ and $1 \leq i \leq n$. Since A is a diagonal matrix, we know that $A_{ij} = 0$ whenever $i \neq j$. Furthermore, we have that $\delta_{ij} = 1$ whenever $i = j$ and 0 otherwise. Hence, we have $A = IA$ by Theorem 2.12 which implies

$$A_{ij} = (I_n A)_{ij} = \sum_{k=1}^n \delta_{ik} A_{kj} = \delta_{ij} A_{ij}$$

for $1 \leq j \leq n$ and $1 \leq i \leq n$.

For the backwards direction, assume $A_{ij} = \delta_{ij} A_{ij}$. Since $\delta_{ij} = 1$ whenever $i = j$ and 0 whenever $i \neq j$, we get that $A_{ij} = A_{ij}$ and $A_{ij} = 0$ respectively. Hence, A is a diagonal matrix. ■

Exercise 2.3.11

Let V be a vector space, and let $T : V \rightarrow V$ be linear. Prove that $T^2 = T_0$ if and only if $R(T) \subseteq N(T)$.

Proof. For the forwards direction, let $T^2 = T_0$. Let $y \in R(T)$. Then for some $x \in V$, we have $y = T(x)$. Then we have

$$T(y) = T(T(x)) = T^2(x) = T_0(x) = 0.$$

Hence, $y \in N(T)$.

For the backwards direction, let $x \in V$. Then we have

$$T^2(x) = T(T(x)).$$

Note that $T(x) \in R(T)$ and $R(T) \subseteq N(T)$ implies that $T^2(x) = 0$. But this also means that $T_0(x) = 0$. Hence, we have $T^2(x) = T_0(x)$. ■

Exercise 2.3.12

Let V, W , and Z be vector spaces, and let $T : V \rightarrow W$ and $U : W \rightarrow Z$ be linear.

(a) Prove that if UT is injective, then T is injective. Must U also be injective?

Proof. Suppose for sake of contradiction that T is not injective. Then there exists $x, y \in V$ such that $x \neq y$ implies $T(x) = T(y)$. But UT being injective must imply that for any $x, y \in V$, we have $UT(x) = UT(y)$ implies $x = y$ which is a contradiction. Hence, T must be injective. Note that U need not be injective in this case. ■

(b) Prove that if UT is surjective, then U is surjective. Must T also be surjective?

Proof. Suppose that UT is surjective. Let $y \in R(UT)$. By default, we know that $R(U) \subseteq Z$. Then for some $x \in V$, we have

$$y = UT(x) = U(T(x)).$$

Hence, $y \in R(U)$ as well which implies that $Z \subseteq R(U)$. Thus, U must be surjective. Notice that T need not be surjective for the equation above to be true. ■

(c) Prove that if U and T are injective and surjective, then UT is also.

Proof. First, we show that UT is injective. Let $x, y \in V$. Then

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

But U being injective, implies that $T(x) = T(y)$. Since T is also injective, we must have $x = y$. Hence, UT is injective.

Now, we show that UT is surjective. Let $z \in R(U)$. Since U is surjective, we have that for some $y \in W$,

$$z = U(y).$$

Since T is also surjective, we have that $y \in W$ implies that $T(x) = y$ for some $x \in V$. Hence, we have

$$z = U(y) = U(T(x)) = UT(x)$$

and so we have UT surjective. ■

Exercise 2.3.13

Let A and B be $n \times n$ matrices. Recall that the trace of A is defined by

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

Prove that $\text{tr}(AB) = \text{tr}(BA)$ and $\text{tr}(A) = \text{tr}(A^t)$.

Proof. Let $1 \leq i \leq n$. Observe that

$$\begin{aligned} \text{tr}(AB) &= \sum_{i=1}^n (AB)_{ii} \\ &= \sum_{i=1}^n \left(\sum_{k=1}^n A_{ik} B_{ki} \right) \\ &= \sum_{k=1}^n \left(\sum_{i=1}^n B_{ki} A_{ik} \right) \\ &= \sum_{k=1}^n (BA)_{kk} \\ &= \text{tr}(BA). \end{aligned}$$

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

For the second formula, observe that

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

■

Exercise 2.3.14

Assume the notation in Theorem 2.13.

- (a) Suppose that z is a (column) vector in F^p . Use Theorem 2.13(b) to prove that Bz is a linear combination of the columns of B . In particular, if $z = (a_1, a_2, \dots, a_p)^t$, then show that

$$Bz = \sum_{j=1}^p a_j v_j.$$

Proof. Note that B is an $n \times p$ matrix and that

$$z = (a_1, a_2, \dots, a_p)^t = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}.$$

Observe that z can be re-written in the following way:

$$z = a_1 e_1 + a_2 e_2 + \dots + a_p e_p = \sum_{j=1}^p a_j e_j.$$

Using theorem 2.13, we find that

$$Bz = B\left(\sum_{j=1}^p a_j e_j\right) = \sum_{j=1}^p a_j (Be_j) = \sum_{j=1}^p a_j v_j.$$

Hence, we have

$$Bz = \sum_{j=1}^p a_j v_j.$$

■

- (b) Extend (a) to prove that column j of AB is a linear combination of the columns of A with coefficients in the linear combination being entries of column j of B .

Proof. Denote the column vectors of A as x_i with $1 \leq i \leq n$ and note that A is an $m \times n$ matrix. Observe that the j th column of B can be written as

$$v_j = \sum_{i=1}^n B_{ij} e_i$$

for $1 \leq j \leq p$. Using Theorem 2.13 again, we find that

$$\begin{aligned} u_j &= Av_j = A\left(\sum_{i=1}^n B_{ij}e_i\right) \\ &= \sum_{i=1}^n B_{ij}(Ae_i) \\ &= \sum_{i=1}^n B_{ij}x_i \end{aligned}$$

where x_i is a column vector in F^m with entries in A . Hence, we have

$$u_j = \sum_{i=1}^n B_{ij}x_i \quad \text{for } 1 \leq j \leq p.$$

■

- (c) For any row vector $w \in F^m$, prove that wA is a linear combination of the rows of A with the coefficients in the linear combination being the coordinates of w .

Proof. Let $w \in F^m$ be a row vector with entries b_j for $1 \leq i \leq m$. Denote x_i as the i th row of A which is an $m \times n$ matrix. Using the properties of transpose, we write

$$\begin{aligned} wA &= (A^t w^t)^t = \left(\sum_{i=1}^m b_i x_i^t\right)^t \\ &= \sum_{i=1}^m (b_i x_i^t)^t \\ &= \sum_{i=1}^m b_i x_i. \end{aligned}$$

Hence,

$$wA = \sum_{i=1}^m b_i x_i.$$

■

- (d) Prove the analogous result to (b) about rows: Row i of AB is a linear combination of the rows of B with the coefficients in the linear combination being the entries of row i of A .

Proof. Let u_i denote the i th row of the matrix AB . Note that AB is an $m \times p$ matrix. By taking a similar approach to proving part (a) of Theorem 2.13, we have

$$u_i = \begin{pmatrix} (AB)_{i1} \\ (AB)_{i2} \\ \vdots \\ (AB)_{ip} \end{pmatrix}^t = \begin{pmatrix} \sum_{k=1}^n A_{ik}B_{k1} \\ \sum_{k=1}^n A_{ik}B_{k2} \\ \vdots \\ \sum_{k=1}^n A_{ik}B_{kp} \end{pmatrix}^t = \begin{pmatrix} A_{i1} \\ A_{i2} \\ \vdots \\ A_{in} \end{pmatrix}^t B = x_i B$$

where x_i is the i th row of A . Apply part (c), we have

$$u_i = x_i B = \sum_{j=1}^n A_{ij} v_j \quad \text{for } 1 \leq i \leq m$$

where v_j is the j th row of B that has dimensions $1 \times p$. ■

Exercise 2.3.15

Let M and A be matrices for which the product matrix MA is defined. If the j th column of A is a linear combination of a set of columns of A , prove that the j th column of MA is a linear combination of the corresponding columns of MA with the same corresponding coefficients.

Proof. Let M be an $m \times n$ matrix. Let x_j be the j th column of A where A is an $n \times p$ matrix. Let u_j be the j th column of MA where MA is an $m \times p$ matrix. Define this j th column vector as

$$u_j = Mx_j$$

by Theorem 2.13. Note that

$$x_j = \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{nj} \end{pmatrix} = \sum_{i=1}^n A_{ij} e_i$$

where e_i is the standard basis vector for F^n . We can see by Corollary to Theorem 2.12 that

$$u_j = Mx_j = M \left(\sum_{i=1}^n A_{ij} e_i \right) = \sum_{i=1}^n A_{ij} (Me_i).$$

Note that e_i is an $n \times 1$ matrix and M is an $m \times n$ matrix. This means that Me_i is in F^m . Denote this column vector as $v_i = Me_i$ where i is the i th column of M . Hence, we have

$$u_j = \sum_{i=1}^n A_{ij} v_i \quad \text{for } 1 \leq j \leq p.$$
■

Exercise 2.3.16

Let V be finite-dimensional vector space, and let $T : V \rightarrow V$ be linear.

- (a) If $\text{rank}(T) = \text{rank}(T^2)$, prove that $R(T) \cap N(T) = \{0\}$. Deduce that $V = R(T) \oplus N(T)$ (see the exercises of Section 1.3).

Proof. Since V is finite-dimensional, we know that

$$\dim(V) = \text{nullity}(T) + \text{rank}(T) \tag{1}$$

by the dimension theorem. By the same reasoning, we also get that

$$\dim(V) = \text{nullity}(T^2) + \text{rank}(T^2). \tag{2}$$

Equating (1) and (2) together, we get that

$$\text{nullity}(T) + \text{rank}(T) = \text{nullity}(T^2) + \text{rank}(T^2).$$

Since $\text{rank}(T) = \text{rank}(T^2)$, we find that $\text{nullity}(T) = \text{nullity}(T^2)$. Let $x \in N(T)$ be arbitrary. Then we have that $T(x) = 0$ with $x \in N(T)$ implying $x = 0$. But since $x \in N(T^2)$ as well, we have that

$$T^2(x) = 0.$$

Note that

$$T^2(x) = 0 \Leftrightarrow T(T(x)) = 0$$

with $T(x) \in N(T)$ implying $T(x) = 0$ since T is injective. Note that

$$T(x) = x = 0$$

where $T(x) \in R(T)$ and $x \in N(T)$. Hence, $R(T) \cap N(T) = \{0\}$.

By the dimension theorem and Exercise 1.6.29, we have

$$\begin{aligned} \dim(R(T) + N(T)) &= \dim(R(T)) + \dim(N(T)) - \dim(R(T) \cap N(T)) \\ &= \dim(R(T)) + \dim(N(T)) \\ &= \dim(V). \end{aligned}$$

Hence, $V = R(T) + N(T)$ by Theorem 1.11. Thus, we have

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

■

(b) Prove that $V = R(T^k) \oplus N(T^k)$ for some positive integer k .

Proof. Pick a positive integer k such that $\text{rank}(T^k) = \text{rank}(T^{k+1})$. By dimension theorem, we can write

$$\dim(V) = \text{rank}(T^k) + \text{nullity}(T^k). \quad (1)$$

By the same reasoning, we also have

$$\dim(V) = \text{rank}(T^{k+1}) + \text{nullity}(T^{k+1}). \quad (2)$$

By setting (1) and (2) equal to each other and using the fact that $\text{rank}(T^k) = \text{rank}(T^{k+1})$, we have

$$\text{nullity}(T^k) = \text{nullity}(T^{k+1}).$$

Thus, $N(T^k) = N(T^{k+1})$ by Theorem 1.11. Now, let $x \in N(T^{k+1})$. Then we have $T^{k+1}(x) = 0$. Since $x \in N(T^k)$, we also have $T^k(x) = 0$. But note that

$$T^{k+1}(x) = T^k(T(x)) = 0.$$

So, we have

$$T^k(T(x)) = T^k(x) = 0.$$

Since T_k is injective, we know that

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

But note that $T(x) \in R(T^k)$ and $x \in N(T^k)$. Hence, $R(T^k) \cap N(T^k) = \{0\}$. To show that $V = R(T^k) + N(T^k)$, we can just follow the same process shown in part (a). Hence, we have

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

■

Exercise 2.3.18

Using only the definition of matrix multiplication, prove that multiplication of matrices is associative.

Proof. Define A, B and C as $m \times n$, $n \times p$, and $p \times \ell$ matrices respectively. This establishes that $(AB)C$ is an $m \times \ell$ matrix. Using the definition of matrix multiplication, we get that

$$\begin{aligned} ((AB)C)_{ij} &= \sum_{k=1}^p (AB)_{ik} C_{kj} \\ &= \sum_{k=1}^p \left(\sum_{\lambda=1}^n A_{i\lambda} B_{\lambda k} \right) C_{kj} \\ &= \sum_{\lambda=1}^n A_{i\lambda} \left(\sum_{k=1}^p B_{\lambda k} C_{kj} \right) \\ &= \sum_{\lambda=1}^n A_{i\lambda} (BC)_{\lambda j} \\ &= (A(BC))_{ij} \end{aligned}$$

where $1 \leq i \leq m$ and $1 \leq j \leq \ell$. Hence, we have $(AB)C = A(BC)$. ■

2.4 Invertibility and Isomorphisms

Exercise 2.4.4

Let A and B be $n \times n$ invertible matrices. Prove that AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. Let A and B be $n \times n$ invertible matrices. We need to show that $(AB)B^{-1}A^{-1} = I_n$ and $B^{-1}A^{-1}(AB) = I_n$. Observe that

$$\begin{aligned} BB^{-1} = I_n &\Leftrightarrow A(BB^{-1}) = A \\ &\Leftrightarrow (AB)B^{-1} = A \\ &\Leftrightarrow (AB)B^{-1}A^{-1} = AA^{-1} \\ &\Leftrightarrow (AB)B^{-1}A^{-1} = I_n. \end{aligned}$$

Similarly, we have

$$\begin{aligned} A^{-1}A = I_n &\Leftrightarrow B^{-1}(A^{-1}A) = B^{-1} \\ &\Leftrightarrow (B^{-1}A^{-1})A = B^{-1} \\ &\Leftrightarrow (B^{-1}A^{-1})AB = B^{-1}B \\ &\Leftrightarrow (B^{-1}A^{-1})AB = I_n. \end{aligned}$$

This tells us that AB is invertible and that

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

■

Exercise 2.4.5

Let A be invertible. Prove that A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$.

Proof. Let A be an invertible. Observe that $(A^t)^t = A$. Thus, we have

$$I_n = AA^{-1} = (A^t)^t A^{-1} = \left((A^{-1})^t A^t \right)^t.$$

Taking the transpose on both sides gives us the following equation

$$(A^{-1})^t A^t = I_n$$

where $(I_n)^t = I_n$. Similarly, we have

$$I_n = A^{-1}A = A^{-1}(A^t)^t = \left(A^t(A^{-1})^t \right)^t$$

which implies

$$\left(A^t(A^{-1})^t \right)^t = I_n.$$

Taking the transpose again then gives us

$$A^t(A^{-1})^t = I_n.$$

Therefore, A^t is invertible and that

■

Exercise 2.4.6

Prove that if A is invertible and $AB = O$, then $B = O$.

Proof. Suppose that A is an invertible matrix. Let AB be a defined matrix product where $AB = O$. Since A is invertible, we have $A^{-1}A = AA^{-1} = I$. Thus, we have

$$\begin{aligned} AB = O &\Rightarrow A^{-1}(AB) = A^{-1}O \\ &\Rightarrow (A^{-1}A)B = O \\ &\Rightarrow IB = O \\ &\Rightarrow B = O. \end{aligned}$$

Hence, we have $B = O$.

■

Exercise 2.4.7

Let A be an $n \times n$ matrix.

- (a) Suppose that $A^2 = O$. Prove that A is not invertible.

Proof. Let $A^2 = O$ where A is an $n \times n$ matrix. Suppose for sake of contradiction that A is invertible. Thus, we have $AA^{-1} = A^{-1}A = I_n$. Now, observe that multiplying

A^{-1} on the left side of $A^2 = O$ produces the following

$$\begin{aligned} A^{-1}(A^2) &= A^{-1}O \Rightarrow A^{-1}(AA) = O \\ &\Rightarrow (A^{-1}A)A = O \\ &\Rightarrow I_n A = O \\ &\Rightarrow A = O. \end{aligned}$$

■

- (b) Suppose that $AB = O$ for some nonzero $n \times n$ matrix B . Could A be invertible? Explain.

Solution. The matrix A cannot be invertible in this case since the nonzero matrix B implies that $A = O$ for $AB = O$ to hold and that we know that the zero matrix O cannot be invertible. ■

Exercise 2.4.8

Prove Corollaries 1 and 2 of Theorem 2.18.

Proof. Let $T : V \rightarrow V$ be linear and let $\beta = \{v_1, v_2, \dots, v_n\}$ be an ordered basis for V . Suppose T is invertible. Then there exists a unique linear transformation denoted by $T^{-1} : V \rightarrow V$ such that

$$TT^{-1} = T^{-1}T = I_V.$$

By using the Corollary to Theorem 2.11 and part (d) of Theorem 2.12, we must have that

$$[T]_{\beta}[T^{-1}]_{\beta} = [TT^{-1}]_{\beta} = [I_V]_{\beta} = I_n$$

and similarly,

$$[T^{-1}]_{\beta}[T]_{\beta} = [T^{-1}T]_{\beta} = [I_V]_{\beta} = I_n.$$

This tells us that $[T]_{\beta}$ is invertible and that

$$[T^{-1}]_{\beta} = ([T]_{\beta})^{-1}.$$

For the backwards direction, the ordered basis β defined earlier implies that there exists a unique linear transformation $U : V \rightarrow V$ defined by

$$U(v_j) = \sum_{i=1}^n A_{ij}v_i \text{ for } 1 \leq j \leq n.$$

We need to show that $U = T^{-1}$. Using the fact that $[T]_{\beta}$ is invertible, we can write

$$[TU]_{\beta} = [T]_{\beta}[U]_{\beta} = I_n = [I_V]_{\beta}$$

and similarly

$$[UT]_{\beta} = [U]_{\beta}[T]_{\beta} = I_n = [I_V]_{\beta}.$$

But this tells us that $UT = TU = I_V$. So, $U = T^{-1}$ and that T is invertible. ■

Proof. Let $L_A : F^n \rightarrow F^n$. Let $\beta = \{v_1, v_2, \dots, v_n\}$ be a basis for F^n . By part (a) of Theorem 2.15, we have that $[L_A]_{\beta} = A$. Since A is invertible, we know that L_A must also be invertible by Corollary to Theorem 2.18. Furthermore, we have that $L_A L_A^{-1} = I$ implies

that

$$L_{A^{-1}} = (L_A)^{-1}.$$

Conversely, L_A invertible implies that $[L_A]_\beta$ is invertible by Corollary to Theorem 2.18. By Theorem 2.15, we must have that $[L_A]_\beta = A$. But this means that A is invertible. ■

Exercise 2.4.9

Let A and B be $n \times n$ matrices such that AB is invertible. Prove that A and B are invertible. Given an example to show that arbitrary matrices A and B need not be invertible if AB is invertible.

Proof. ■

Exercise 2.4.10

Let A and B be $n \times n$ matrices such that $AB = I_n$.

- (a) Use Exercise 9 to conclude that A and B are invertible.

Solution. Apply Exercise 9. ■

- (b) Prove $A = B^{-1}$ (and hence $B = A^{-1}$).

Proof. Since A is an $n \times n$ invertible matrix, we can write that

$$\begin{aligned} AB = I_n &\Rightarrow A^{-1}(AB) = A^{-1}I_n \\ &\Rightarrow (A^{-1}A)B = A^{-1}I_n \\ &\Rightarrow I_n B = A^{-1}I_n \\ &\Rightarrow B = A^{-1}. \end{aligned}$$

Likewise, B being an $n \times n$ invertible matrix implies that

$$\begin{aligned} AB = I_n &\Rightarrow (AB)B^{-1} = I_n B^{-1} \\ &\Rightarrow A(BB^{-1}) = I_n B^{-1} \\ &\Rightarrow AI_n = I_n B^{-1} \\ &\Rightarrow A = B^{-1}. \end{aligned}$$

- (c) State and prove analogous results for linear transformations defined on finite-dimensional vector spaces.

Proof. Define $T : V \rightarrow V$ and $U : V \rightarrow V$ with V being an arbitrary finite-dimensional vector space. Let $\beta = \{v_1, v_2, \dots, v_n\}$ be an ordered basis for V such that $[T]_\beta$ and $[U]_\beta$ are properly defined and that $[T]_\beta[U]_\beta = I_n$. Then we have

$$[T]_\beta = ([U]_\beta)^{-1} \text{ and } [U]_\beta = ([T]_\beta)^{-1}.$$

To prove this, we can let $A = [T]_\beta$ and $B = [U]_\beta$ and use parts (a) and (b) to get our desired result. ■

Exercise 2.4.11

Verify that the transformation in Example 5 is injective.

Proof. Since $\dim(P_3(\mathbb{R})) = \dim(M_{2 \times 2}(\mathbb{R}))$, we must have that $P_3(\mathbb{R})$ is isomorphic to $M_{2 \times 2}(\mathbb{R})$ by Theorem 2.19. By definition, this means that T is invertible which further implies that T is injective. ■

Exercise 2.4.12

Prove Theorem 2.21.

Proof. Let $\dim(V) = n$. Let $\beta = \{v_1, v_2, \dots, v_n\}$ be an ordered basis for V . We can show that $\phi_\beta(y)$ is an isomorphism by showing that ϕ_β is a surjective and injective linear map. Suppose $\phi_\beta(x) = \phi_\beta(y)$. Then by definition of ϕ_β , we must have that $[x]_\beta = [y]_\beta$ which further implies that $x = y$. Now, let $y \in V$. Since β is an ordered basis for V , we can find scalars $\delta_1, \delta_2, \dots, \delta_n$ such that

$$y = \sum_{i=1}^n \delta_i v_i.$$

This implies that we have constructed a coordinate vector such that $[y]_\beta = \phi_\beta(y)$. Hence, ϕ_β is surjective. Thus, we find that ϕ_β is an isomorphism. ■

Exercise 2.4.13

Let \sim mean "is isomorphic to." Prove that \sim is an equivalence relation on the class of vector spaces over F .

Proof. Let S be the class of vector spaces over F . Let $V \in S$. Notice that V is isomorphic to itself since the identity linear transformation $I_V : V \rightarrow V$ is invertible. Thus, $V \sim V$.

Let $V, W \in S$. Suppose that $V \sim W$. Then there exists an invertible linear map $T : V \rightarrow W$. This means that T contains an inverse $T^{-1} : W \rightarrow V$ such that $TT^{-1} = I_W$ and $T^{-1}T = I_V$. Note that T^{-1} is linear and that T^{-1} is also invertible since T is its inverse. Thus, $W \sim V$.

Let $V, W, Z \in S$. Now, suppose that $V \sim W$ and $W \sim Z$. This means that there exists invertible linear maps $T : V \rightarrow W$ and $U : W \rightarrow Z$. Let $L : V \rightarrow Z$ be defined by $L = UT$. Let $x, y \in V$. Suppose $L(x) = L(y)$. Then

$$\begin{aligned} L(x) &= L(y) \\ UT(x) &= UT(y) \\ U(T(x)) &= U(T(y)). \end{aligned}$$

Notice that U is an injective map which tells us that $T(x) = T(y)$. But T is also injective, so we must have $x = y$. Hence, L is an injective map.

Now, let $z \in Z$. Since U is surjective, we must have $z = U(y)$ for some $y \in W$. But note that T is surjective implies that $y = T(x)$ for some $x \in V$. But this tells us that

$$z = U(y) = U(T(x)) = UT(x) = L(x).$$

Hence, L is a surjective map and that L is an invertible map. Thus, $V \sim Z$. ■

Exercise 2.4.15

Let V and W be finite-dimensional vector spaces, and let $T : V \rightarrow W$ be a linear transformation. Suppose that β is a basis for V . Prove that T is an isomorphism if and only if $T(\beta)$ is a basis for W .

Proof. For the forwards direction, suppose $T : V \rightarrow W$ is an isomorphism where V and W are finite-dimensional vector spaces. Hence, T is invertible by definition. Let $\beta = \{v_1, v_2, \dots, v_n\}$ be an ordered basis for V . Since T is also injective and that β is a linearly independent subset of V , we know that $T(\beta)$ is a linearly independent subset of W by part (b) of Exercise 2.1.14. Furthermore, β being a basis for V implies that $\text{span}(T(\beta)) = R(T)$ by Theorem 2.2. But notice that T is also surjective since it is an isomorphism. Hence, we must have that $R(T) = W$ and that $T(\beta)$ spans W . Thus, we have that $T(\beta)$ is a basis for W .

For the backwards direction, assume $T(\beta)$ is a basis for W . In order to show that T is an isomorphism, we must show that T is a bijective linear map. Let $x, y \in V$. Since β is a basis for V , we must have

$$x = \sum_{i=1}^n a_i v_i \text{ and } y = \sum_{i=1}^n b_i v_i$$

for some scalars a_1, a_2, \dots, a_n and b_1, b_2, \dots, b_n . Suppose $T(x) = T(y)$. Then observe that T being linear implies that

$$\begin{aligned} T(x) &= T(y) \\ T\left(\sum_{i=1}^n a_i v_i\right) &= T\left(\sum_{i=1}^n b_i v_i\right) \\ \sum_{i=1}^n a_i T(v_i) &= \sum_{i=1}^n b_i T(v_i) \end{aligned}$$

which subsequently leads to

$$\sum_{i=1}^n (a_i - b_i) T(v_i) = 0 \Leftrightarrow a_i = b_i \text{ for all } i$$

since $T(\beta)$ is a linearly independent subset of W . Hence, $x = y$ and that T is injective.

Now, let $y \in W$. Since $T(\beta)$ is a basis for W and T linear, we know that there exists scalars a_1, a_2, \dots, a_n such that

$$y = \sum_{i=1}^n a_i T(v_i) = T\left(\sum_{i=1}^n a_i v_i\right).$$

But this means that

$$x = \sum_{i=1}^n a_i v_i$$

for some $x \in V$. Hence, T is surjective and thus T is an isomorphism. ■

Exercise 2.4.16

Let B be an $n \times n$ invertible matrix. Define $\Phi : M_{n \times n}(F) \rightarrow M_{n \times n}(F)$ by $\Phi(A) = B^{-1}AB$. Prove that Φ is an isomorphism.

Proof. Let B be an $n \times n$ matrix. First, we show that $\Phi : M_{n \times n}(F) \rightarrow M_{n \times n}(F)$ is linear. Let $a \in F$ and $D, C \in M_{n \times n}(F)$. Then by using the algebraic properties of matrices, we must have

$$\begin{aligned}\Phi(aC + D) &= B^{-1}(aC + D)B \\ &= (B^{-1}(aC) + B^{-1}D)B \\ &= B^{-1}(aC)B + B^{-1}DB \\ &= a(B^{-1}CB) + B^{-1}DB \\ &= a\Phi(C) + \Phi(D)\end{aligned}$$

Hence, Φ is a linear map. Let $A \in M_{n \times n}(F)$. By definition of Φ , we know that

$$\Phi(A) = B^{-1}AB.$$

So, define the linear map $U : M_{n \times n}(F) \rightarrow M_{n \times n}(F)$ by $U(A) = BAB^{-1}$. Notice that B is invertible, so we can write

$$B(B^{-1}AB)B^{-1} = B^{-1}(BAB^{-1})B = IA = AI = A$$

where I is the identity linear map for $M_{n \times n}$. Hence, we have $\Phi U = U \Phi = I$. Hence, Φ is an invertible map and thus Φ is an isomorphism. ■

Exercise 2.4.17

Let V and W be finite-dimensional vector spaces and $T : V \rightarrow W$ be an isomorphism. Let V_0 be a subspace of V .

- (a) Prove that $T(V_0)$ is a subspace of W .

Proof. Let V_0 be a subspace of V . Observe that $T(0_V) = 0_W \in T(V_0)$ since $0_V \in V_0$. Let $z, y \in T(V_0)$. Then $T(x) = y$ and $T(w) = z$ for $x, w \in V_0$ implies that

$$z + y = T(x) + T(w) = T(x + w)$$

since $x + w \in V_0$. Now, let $c \in F$. Hence, $z + y \in T(V_0)$. Then

$$cy = cT(x) = T(cx)$$

with $cx \in V_0$. Hence, $cy \in T(V_0)$. ■

- (b) Prove that $\dim(V_0) = \dim(T(V_0))$.

Proof. Since V and W are finite-dimensional vector spaces and that V_0 and $T(V_0)$ are subspaces of V and W respectively, we must have that V_0 and $T(V_0)$ be both finite-dimensional by Theorem 1.11. Let $T(V_0) = W_0$. Since T is an isomorphism, we know that T must be injective and surjective. So, the Dimension Theorem implies that

$$\dim(V_0) = \dim(N(T)) + \dim(R(T)) = \dim(W_0).$$

Hence, $\dim(V_0) = \dim(W_0)$. ■

Exercise 2.4.20

Let $T : V \rightarrow W$ be a linear transformation from an n -dimensional vector space V to an m -dimensional vector space W . Let β and γ be ordered bases for V and W , respectively. Prove that $\text{rank}(T) = \text{rank}(L_A)$ and that $\text{nullity}(T) = \text{nullity}(L_A)$, where $A = [T]_{\beta}^{\gamma}$.

Proof. Observe that $R(T)$ is a subspace of W and $R(L_A)$ is a subspace F^m . Since $\phi_{\gamma} : W \rightarrow F^m$ is an isomorphism by Theorem 2.21 (that is, $\dim(W) = \dim(F^m)$), we can use Exercise 17 to state that $\text{rank}(T) = \text{rank}(L_A)$. Similarly, $N(T)$ and $N(L_A)$ are subspaces of V and F^n , respectively. Hence, $\phi_{\beta} : V \rightarrow F^n$ (that is, $\dim(V) = \dim(F^n)$) being an isomorphism implies that $\text{nullity}(T) = \text{nullity}(L_A)$. ■

Exercise 2.4.21

Let V and W be finite-dimensional vector spaces with ordered bases $\beta = \{v_1, v_2, \dots, v_n\}$ and $\gamma = \{w_1, w_2, \dots, w_m\}$, respectively. By Theorem 2.6, there exists linear transformations $T_{ij} : V \rightarrow W$ such that

$$T_{ij}(v_k) = \begin{cases} w_i & \text{if } k = j \\ 0 & \text{if } k \neq j. \end{cases}$$

First prove that $\{T_{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ is a basis for $\mathcal{L}(V, W)$. Then let M^{ij} be the $m \times n$ matrix with 1 in the i th row and j th column and 0 elsewhere, and prove that $[T_{ij}]_{\beta}^{\gamma} = M^{ij}$. Again by Theorem 2.6, there exists a linear transformation $\Phi : \mathcal{L}(V, W) \rightarrow M_{m \times n}(F)$ such that $\Phi(T_{ij}) = M^{ij}$. Prove that Φ is an isomorphism.

Proof. First, we need to show that $S = \{T_{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ is a basis for $\mathcal{L}(V, W)$; that is, we need to show that

$$\sum_{i=1}^m a_{ij} T_{ij}(v_k) = 0 \quad \text{for } 1 \leq j \leq n. \quad (1)$$

for some scalars a_{ij} . By definition of T_{ij} , observe that for $1 \leq j \leq n$ we have $j = k$ such that $T_{ij}(v_k) = w_i$. So, (1) can be re-written as

$$\sum_{i=1}^m a_{ij} w_i = 0.$$

Now, we can use the linear independence of $\gamma = \{w_1, w_2, \dots, w_m\}$ to conclude that $a_{ij} = 0$ for all $1 \leq i \leq m$. But this means that S must be linearly independent. Let $T_{ij} \in \mathcal{L}(V, W)$ but not in S . Then adjoining T_{ij} to S produces a linearly dependent set such that $T_{ij} \in \text{span}(S)$ by Theorem 1.7. Hence, S is a basis for $\mathcal{L}(V, W)$.

Using the fact that S is a basis for $\mathcal{L}(V, W)$, we can now write that

$$w_i = T_{ij}(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{for } 1 \leq j \leq n. \quad (2)$$

We can see that for $1 \leq j \leq n$ that the matrix representation $[T_{ij}]_{\beta}^{\gamma}$ contains entries $a_{ij} = 1$ whenever $i = j$ and 0 otherwise. But observe that this is just M^{ij} and hence $[T_{ij}]_{\beta}^{\gamma} = M^{ij}$. By Theorem 2.6, we can see that there exists a linear transformation $\Phi : \mathcal{L}(V, W) \rightarrow M_{m \times n}(F)$ such that $\Phi(T_{ij}) = M^{ij}$. Our goal now is to show that Φ is an isomorphism. That is, we will show that Φ is both injective and surjective. Let $T_{ij}, U_{ij} \in \mathcal{L}(V, W)$.

Assume $\Phi(T_{ij}) = \Phi(U_{ij})$. Then we can write

$$\begin{aligned}\Phi(T_{ij}) &= \Phi(U_{ij}) \\ M^{ij} &= N^{ij} \\ [T_{ij}]_{\beta}^{\gamma} &= [U_{ij}]_{\beta}^{\gamma}.\end{aligned}$$

Notice that the last equality implies that $T_{ij} = U_{ij}$ by corollary to Theorem 2.6. Hence, Φ is an injective map.

Let $U_{ij} \in \mathcal{L}(V, W)$. Using ordered bases β and γ and the fact that S is a basis for $\mathcal{L}(V, W)$, we can construct $[U_{ij}]_{\beta}^{\gamma}$ such that

$$w_i = U_{ij}(v_j) = \sum_{i=1}^m B_{ij} w_i \quad \text{for } 1 \leq j \leq n.$$

Observe that $[U_{ij}]_{\beta}^{\gamma} = N^{ij} = \Phi(U_{ij})$ and thus Φ is surjective. We can conclude that Φ is an isomorphism. ■

Exercise 2.4.24

Let $T : V \rightarrow Z$ be a linear transformation of a vector space V onto a vector space Z . Define the mapping

$$\bar{T} : V/N(T) \rightarrow Z \text{ by } \bar{T}(v + N(T)) = T(v)$$

for any coset $v + N(T)$ in $V/N(T)$.

- (a) Prove that \bar{T} is well-defined; that is, prove that if $v + N(T) = v' + N(T)$, then $T(v) = T(v')$.

Proof. Suppose $v + N(T) = v' + N(T)$. Then

$$\begin{aligned}T(v) &= \bar{T}(v + N(T)) \\ &= \bar{T}(v' + N(T)) \\ &= T(v').\end{aligned}$$

Hence, \bar{T} is well-defined map. ■

- (b) Prove that \bar{T} is linear.

Proof. Let $a \in F$ and $v + N(T), u + N(T) \in V/N(T)$. Then using the operations defined in Exercise 1.3.31 and the fact that T is linear, we have

$$\begin{aligned}\bar{T}(a(v + N(T)) + (u + N(T))) &= \bar{T}((av + N(T)) + (u + N(T))) \\ &= \bar{T}((av + u) + N(T)) \\ &= T(av + u) \\ &= aT(v) + T(u) \\ &= a\bar{T}(v + N(T)) + \bar{T}(u + N(T)).\end{aligned}$$

Thus, \bar{T} is linear. ■

- (c) Prove that \bar{T} is an isomorphism.

Proof. First, we show that \bar{T} is an injective map. Let $x + N(T), y + N(T) \in V/N(T)$. Then observe that

$$\begin{aligned}\bar{T}(x + N(T)) &= \bar{T}(y + N(T)) \\ T(x) &= T(y).\end{aligned}$$

Since T is linear, we can write

$$T(x) = T(y) \Leftrightarrow T(x - y) = 0.$$

Hence, $x - y \in N(T)$. Since $N(T)$ is a subspace of V , we have $x + N(T) = y + N(T)$ by Exercise 1.3.31. This implies that \bar{T} is injective.

Now, we will show that \bar{T} is a surjective map. Since T is a linear transformation from V onto Z , we know that any arbitrary $y \in Z$ can be written as $T(x) = y$ for some $x \in V$. By definition of \bar{T} , we have

$$y = T(x) = \bar{T}(x + N(T))$$

where $x + N(T) \in V/N(T)$. Hence, \bar{T} is a surjective linear map. Thus, \bar{T} is an isomorphism. ■

- (d) Prove that the diagram shown in Figure 2.3 commutes; that is, prove that $T = \bar{T}\eta$.

Proof. Let $v \in V$. Since $\eta : V \rightarrow V/N(T)$ is defined by $\eta(v) = v + N(T)$, we can write that

$$\begin{aligned}T(v) &= \bar{T}(v + N(T)) \\ &= \bar{T}(\eta(v)) \\ &= \bar{T}\eta(v).\end{aligned}$$

Since $v \in V$ is arbitrary, we know that $T = \bar{T}\eta$. ■

2.5 The Change of Coordinate Matrix

Exercise 2.5.8

Prove the following generalization of Theorem 2.23. Let $T : V \rightarrow W$ be a linear transformation from a finite-dimensional vector space V to a finite-dimensional vector space W . Let β and β' be ordered bases for V , and let γ and γ' be ordered bases for W . Then $[T]_{\beta'}^{\gamma'} = P^{-1}[T]_{\beta}^{\gamma}Q$ where Q is the matrix that changes β' -coordinates into β -coordinates and P is the matrix that changes γ' -coordinates into γ -coordinates.

Proof. Our goal is to show that

$$P[T]_{\beta'}^{\gamma'} = [T]_{\beta}^{\gamma}Q \tag{1}$$

where $P = [I_W]_{\gamma'}^{\gamma}$ and $Q = [I_V]_{\beta'}^{\beta}$. Using Theorem 2.11, we can write that

$$\begin{aligned} P[T]_{\beta'}^{\gamma'} &= [I_W]_{\gamma'}^{\gamma} [T]_{\beta'}^{\gamma'} = [I_W T]_{\beta'}^{\gamma} \\ &= [T]_{\beta'}^{\gamma} \\ &= [T I_V]_{\beta'}^{\gamma} \\ &= [T]_{\beta}^{\gamma} [I_V]_{\beta'}^{\beta} \\ &= [T]_{\beta}^{\gamma} Q. \end{aligned}$$

Hence, we have

$$P[T]_{\beta'}^{\gamma'} = [T]_{\beta}^{\gamma} Q.$$

Since P is invertible (by Theorem 2.22), we can do a left-multiplication of P^{-1} on both sides of (1) to get our desired result

$$[T]_{\beta'}^{\gamma'} = P^{-1} [T]_{\beta}^{\gamma} Q.$$

■

Exercise 2.5.9

Prove that "is similar to" is an equivalence relation on $M_{n \times n}(F)$.

Proof. Let $A \in M_{n \times n}(F)$. Note that

$$AI = IA = A$$

and that I is invertible. Thus, we have

$$A = I^{-1}AI$$

and that $A \sim A$.

Let $A, B \in M_{n \times n}(F)$. Suppose $A \sim B$. Then there exists an invertible matrix Q such that

$$A = Q^{-1}BQ.$$

Using the invertibility of Q , we get our desired result

$$B = QAQ^{-1} = P^{-1}AP$$

where $Q = P^{-1}$ and $Q^{-1} = P$. Hence, $B \sim A$.

Now, let $A, B, C \in M_{n \times n}(F)$. Suppose $A \sim B$ and $B \sim C$. We want to show that $A \sim C$; that is, we want to show that there exists an invertible $Q \in M_{n \times n}(F)$ such that

$$A = Q^{-1}CQ.$$

Since $A \sim B$, there exists an invertible matrix P such that

$$A = P^{-1}BP.$$

Likewise, $B \sim C$ implies that there exists an invertible matrix L such that

$$B = L^{-1}CL.$$

Using Exercise 2.4.4, we can write that

$$\begin{aligned} A &= P^{-1}BP \\ &= P^{-1}(L^{-1}CL)P \\ &= (LP)^{-1}C(LP) \\ &= Q^{-1}CQ \end{aligned}$$

where $(LP)^{-1} = Q^{-1}$ and $LP = Q$. Hence, $A \sim C$. ■

Exercise 2.5.10

Prove that if A and B are similar $n \times n$ matrices, then $\text{tr}(A) = \text{tr}(B)$.

Proof. Suppose A and B are similar $n \times n$ matrices. Then there exists an invertible matrix Q such that

$$A = Q^{-1}BQ.$$

Using the fact that $\text{tr}(AB) = \text{tr}(BA)$ and that matrix multiplication is associative, we can write that

$$\begin{aligned} \text{tr}(A) &= \text{tr}((Q^{-1}B)Q) \\ &= \text{tr}(Q(Q^{-1}B)) \\ &= \text{tr}((QQ^{-1})B) \\ &= \text{tr}(I_n B) \\ &= \text{tr}(B). \end{aligned}$$

Hence, we have that $\text{tr}(A) = \text{tr}(B)$. ■

Exercise 2.5.11

Let V be a finite-dimensional vector space with ordered bases α, β and γ .

- (a) Prove that if Q and R are the change of coordinate matrices that change α -coordinates into β -coordinates and β -coordinates into γ -coordinates, respectively, then RQ is the change of coordinate matrix that changes α -coordinates into γ -coordinates.

Proof. Since Q is the matrix that changes α -coordinates into β -coordinates, we have $Q = [I_V]_{\alpha}^{\beta}$. Likewise, R is the matrix that changes β -coordinates into γ -coordinates. So, $R = [I_V]_{\beta}^{\gamma}$. Thus, we have

$$\begin{aligned} RQ &= [I_V]_{\alpha}^{\beta} [I_V]_{\beta}^{\gamma} \\ &= [I_V]_{\alpha}^{\gamma} \end{aligned}$$

by Theorem 2.11. Thus, RQ is the change of coordinate matrix that changes α -coordinates into γ -coordinates ■

- (b) Prove that if Q changes α -coordinates into β -coordinates, then Q^{-1} changes β -coordinates into α -coordinates.

Proof. Suppose Q changes α -coordinates into β -coordinates, we have for any $v \in V$ that

$$[v]_{\alpha} = Q[v]_{\beta}. \tag{1}$$

Since Q is invertible, we have

$$\begin{aligned} Q^{-1}(Q[v]_{\alpha}) &= Q^{-1}[v]_{\alpha} \Rightarrow (Q^{-1}Q)[v]_{\beta} = Q^{-1}[v]_{\alpha} \\ &\Rightarrow [v]_{\beta} = Q^{-1}[v]_{\alpha} \end{aligned}$$

Hence, Q^{-1} changes β -coordinates into α -coordinates. ■

Exercise 2.5.12

Prove the corollary to Theorem 2.23.

Proof. Observe that part (a) of Theorem 2.15 implies that $[L_A]_{\gamma} = A$. Note that $Q = [I_{F^n}]_{\gamma}$ and that $I_{F^n}L_A = L_AI_{F^n}$ so we write

$$\begin{aligned} Q[L_A]_{\gamma} &= [I_{F^n}]_{\gamma}[L_A]_{\gamma} \\ &= [I_{F^n}L_A]_{\gamma} \\ &= [L_AI_{F^n}]_{\gamma} \\ &= [L_A]_{\gamma}[I_{F^n}]_{\gamma} \\ &= AQ. \end{aligned}$$

Hence, we have

$$Q[L_A]_{\gamma} = AQ \tag{1}$$

. Since Q is invertible, we can do left-multiplication of Q^{-1} on both sides of (1) to get our desired result

$$[L_A]_{\gamma} = Q^{-1}AQ. \quad \blacksquare$$

Exercise 2.5.13

Let V be a finite-dimensional vector space over a field F , and let $\beta = \{x_1, x_2, \dots, x_n\}$ be an ordered basis for V . Let Q be an $n \times n$ invertible matrix with entries from F . Define

$$x'_j = \sum_{i=1}^n Q_{ij}x_i \quad \text{for } 1 \leq j \leq n,$$

and set $\beta' = \{x'_1, x'_2, \dots, x'_n\}$. Prove that β' is a basis for V and hence that Q is the change of coordinate matrix changing β' -coordinates into β -coordinates.

Proof. Let $1 \leq j \leq n$. First, we need to show that β' is linearly independent; that is, there exists scalars a_1, a_2, \dots, a_n such that

$$\sum_{j=1}^n a_j x_j = 0 \tag{1}$$

where $a_j = 0$ for all j . Since

$$x'_j = \sum_{i=1}^n Q_{ij}x_i,$$

we can re-write the left side of (1) into

$$\sum_{j=1}^n a_j \left(\sum_{i=1}^n Q_{ij} x_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n a_j Q_{ij} \right) x_i.$$

But since $\beta = \{x_1, x_2, \dots, x_n\}$ is linearly independent, we have that

$$\sum_{j=1}^n a_j Q_{ij} = 0 \Leftrightarrow a_j Q = O$$

where O is the zero-matrix. Since Q is an invertible $n \times n$ matrix, we can multiply Q^{-1} on both sides of $a_j Q = O$, to write

$$\begin{aligned} (a_j Q) Q^{-1} &= a_j (Q Q^{-1}) \\ &= a_j I_n \end{aligned}$$

Thus, $a_j I_n = O$ implies that $a_j = 0$ for all $1 \leq j \leq n$. Hence, β' is linearly independent. Now, let $v \in V$ but not in β' . Then adjoining v into β' will produce a linearly dependent set. Thus, Theorem 2.2 implies that $\text{span}(\beta') = V$. So, β' is a basis for V . ■

Exercise 2.5.14

Prove the converse of Exercise 8: If A and B are each $m \times n$ matrices with entries from a field F , and if there exists invertible $m \times m$ and $n \times n$ matrices P and Q , respectively, such that $B = P^{-1} A Q$, then there exist an n -dimensional vector space V and an m -dimensional vector space W (both over F), ordered bases β and β' for V and γ and γ' for W , and a linear transformation $T : V \rightarrow W$ such that

$$A = [T]_{\beta}^{\gamma} \quad \text{and} \quad B = [T]_{\beta'}^{\gamma'}.$$

Proof. Since A and B are $m \times n$ with entries from a field F , let $V = F^n$ and $W = F^m$. Since V and W are finite-dimensional, they contain ordered bases $\beta = \{v_1, v_2, \dots, v_n\}$ and $\gamma = \{w_1, w_2, \dots, w_m\}$. By Theorem 2.6, there must exist a linear transformation $T : V \rightarrow W$ such that

$$T(v_j) = \sum_{i=1}^m A_{ij} w_i \quad \text{for } 1 \leq j \leq n.$$

Since $V = F^n$ and $W = F^m$, let $T = L_A$. Hence, we have

$$[T]_{\beta}^{\gamma} = [L_A]_{\beta}^{\gamma} = A$$

by part (a) of Theorem 2.15.

Now, define ordered bases $\beta' = \{x'_1, x'_2, \dots, x'_n\}$ and $\gamma' = \{w'_1, w'_2, \dots, w'_m\}$ for V and W , respectively. Since there exists invertible $n \times n$ and $m \times m$ matrices Q and P respectively, we can apply the result from Exercise 2.5.13 to imply that

$$x'_j = \sum_{i=1}^n Q_{ij} x_i \quad \text{for } 1 \leq j \leq n$$

and

$$w'_j = \sum_{i=1}^m P_{ij} w_i \quad \text{for } 1 \leq j \leq n,$$

implying that Q and P are change of coordinate matrices that changes β' -coordinates to β -coordinates and γ' -coordinates to γ -coordinates, respectively.

Finally, we show that $B = [T]_{\beta'}^{\gamma'}$. Using Theorem 2.11, we can write

$$\begin{aligned} B = P^{-1}AQ &= [I_V]_{\gamma'}^{\gamma} [T]_{\beta}^{\gamma} [I_W]_{\beta'}^{\beta} \\ &= [I_V T]_{\beta}^{\gamma'} [I_W]_{\beta'}^{\beta} \\ &= [T]_{\beta}^{\gamma'} [I_W]_{\beta'}^{\beta} \\ &= [T I_W]_{\beta'}^{\gamma'} \\ &= [T]_{\beta'}^{\gamma'}. \end{aligned}$$

Hence, $B = [T]_{\beta'}^{\gamma'}$ and we are done. ■

Chapter 3

Matrix Operations and Elementary Matrices

3.1 The Rank of a Matrix and Matrix Inverses

Exercise 3.2.3

Prove that for any $m \times n$ matrix A , $\text{rank}(A) = 0$ if and only if A is the zero matrix.

Proof. Let A be an $m \times n$ matrix. Notice that $\text{rank}(A) = \dim(R(L_A))$. By the Dimension Theorem, we can see that $\text{rank}(A) = 0$ implies that

$$\dim(F^n) = \dim(R(L_A)) + \dim(N(L_A)) = \dim(N(L_A)).$$

Let $\beta = \{e_1, e_2, \dots, e_n\}$ is the standard ordered basis of F^n . Since $\dim(F^n) = \dim(N(L_A)) = n$, we know that

$$L_A(e_j) = a_j = 0 \quad \text{for all } 1 \leq j \leq n$$

where a_j is the j th column of A . Thus, A must be the zero matrix.

Conversely, suppose that A is the zero matrix. Then we can see that

$$a_j = 0 \quad \text{for all } 1 \leq j \leq n.$$

Using Theorem 3.5, we can see that

$$R(L_A) = \text{span}(\{a_j : 1 \leq j \leq n\}) = \text{span}(\{0\}).$$

Hence, we have

$$\text{rank}(A) = \dim(R(L_A)) = \dim(\text{span}(\{0\})) = 0.$$

Thus, $\text{rank}(A) = 0$. ■

Exercise 3.2.7

Express the invertible matrix

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

as a product of elementary matrices.

Proof. ■

Exercise 3.2.8

Let A be an $m \times n$ matrix. Prove that if c is any nonzero scalar, then $\text{rank}(cA) = \text{rank}(A)$.

Proof. Let A be an $m \times n$ matrix. Our goal is to show that

$$R(L_{cA}) = R(L_A).$$

Let $y \in R(L_{cA})$. Then for some $x \in F^n$, we have that

$$y = L_{cA}(x) = cL_A(x) = A(cx) = L_A(cx).$$

via part (c) of Theorem 2.15. Hence, $y \in R(L_A)$. Now, let $y \in R(L_A)$. Then for some $x \in F^n$, we have $y = L_A(x)$. But note that this can be re-written in the following form

$$y = L_A(x) = cL_A\left(\frac{1}{c}x\right) = L_{cA}\left(\frac{1}{c}x\right)$$

via part (c) of Theorem 2.15. Thus, we can see that $y \in R(L_A)$. This tells us that $R(L_{cA}) = R(L_A)$ and that we have

$$\text{rank}(cA) = \dim(R(L_{cA})) = \dim(R(L_A)) = \text{rank}(A)$$

by theorem 1.11. ■

Exercise 3.2.9

Complete the proof of the corollary to Theorem 3.4 by showing that elementary column operations preserve rank.

Proof. If B is obtained from a matrix A via an elementary column operation, then there exists an elementary matrix E such that $B = AE$. Using Theorem 3.2, we can see that E is invertible and that

$$\text{rank}(B) = \text{rank}(AE) = \text{rank}(A)$$

by part (a) of Theorem 3.4. Hence, elementary column operations preserve rank. ■

Exercise 3.2.10

Prove Theorem 3.6 for the case that A is an $m \times 1$ matrix.

Proof. Fix $n = 1$. Using at most one type 1 row operation and at most one type 2 row operation, we can transform A to have a 1 in the 1, 1 position. By means of at most $m - 1$ type 3 column operations, A can be transformed into the following matrix

$$D = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where the only linearly independent row is the first row. Hence, $\text{rank}(D) = \text{rank}(A) = 1$. ■

Exercise 3.2.11

Let

$$B = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{pmatrix}$$

where B' is an $m \times n$ submatrix of B . Prove that if $\text{rank}(B) = r$, then $\text{rank}(B') = r - 1$.

Proof. Let $\beta = \{e_1, e_2, \dots, e_n\}$ be the standard ordered basis of F^n . Observe by Theorem 3.5 that

$$R(L_B) = \text{span}(L_B(\beta)) = \text{span}(\{b_j : 1 \leq j \leq n\}) \quad (1)$$

where b_j is the j th column of B . Note that

$$L_B(\beta) = \{b_j : 1 \leq j \leq n\} = \{b_1\} \cup \{b_j : 1 < j \leq n\}.$$

So, (1) can be re-written as

$$\begin{aligned} R(L_B) &= \text{span}(\{b_1\} \cup \{b_j : 1 < j \leq n\}) \\ &= \text{span}(\{b_1\}) + \text{span}(\{b_j : 1 < j \leq n\}) \end{aligned}$$

by Exercise 14 of Section 1.4. Note that $\{b_1\} \cap \{b_j : 1 < j \leq n\}$ is disjoint. By the formula found in Exercise 29 in Section 1.6, we find that

$$\begin{aligned} \text{rank}(B) &= \dim(R(L_B)) \\ &= \dim(\text{span}(b_1) + \text{span}(\{b_j : 1 < j \leq n\})) \\ &= \dim(\text{span}(\{b_1\})) + \dim(\text{span}(\{b_j : 1 < j \leq n\})) \\ &= 1 + \dim(R(L_{B'})) \quad (\text{Theorem 3.5}) \\ &= 1 + \text{rank}(B'). \end{aligned}$$

Solving for $\text{rank}(B')$, we get our desired result that

$$\text{rank}(B') = \text{rank}(B) - 1 = r - 1.$$

■

Exercise 3.2.12

Let B' and D' be $m \times n$ matrices, and let B and D be $(m+1) \times (n+1)$ matrices respectively defined by

$$B = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & D' & \\ 0 & & & \end{pmatrix}.$$

Prove that if B' can be transformed into D' by an elementary row [column] operation, then B can be transformed into D by an elementary row [column] operation.

Proof. If B' can be transformed into D' by an elementary row operation, then there exists

an invertible matrix E' such that $D' = E'B'$. Observe that

$$D = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & D' & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & E'B' & \\ 0 & & & \end{pmatrix}.$$

Now, let's apply an elementary row operation on D so that D can be written as

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{pmatrix}.$$

But this tells us that there exists some invertible matrix E such that

$$B = ED. \quad (1)$$

Since E is invertible, we can apply its inverse E^{-1} on the left side of (1) to get that

$$D = E^{-1}B.$$

Thus, D can be obtained from B via an elementary row operation. ■

Exercise 3.2.13

Prove (b) and (c) of Corollary 2 to Theorem 3.6.

Proof. Let $A \in M_{m \times n}(F)$ be arbitrary.

- (b) Let $\gamma = \{e_1, e_2, \dots, e_m\}$ be the standard ordered basis for F^m . Since $L_{A^t} : F^m \rightarrow F^n$ is linear, we can see by Theorem 2.2 that

$$R(L_{A^t}) = \text{span}(L_{A^t}(\gamma)) = \text{span}(\{L_{A^t}(e_j) : 1 \leq j \leq m\}).$$

Furthermore, we have that $L_{A^t}(e_j) = A^t e_j = a_j^t$ where a_j^t is the j th column of A^t (or the j th row of A). Since $\text{rank}(A^t) = \text{rank}(A)$ by part (a), we can write that

$$\begin{aligned} \text{rank}(A) &= \text{rank}(A^t) = \dim(R(L_{A^t}(\gamma))) \\ &= \dim(\text{span}(\{L_{A^t}(e_j) : 1 \leq j \leq m\})). \end{aligned}$$

Thus, the rank of A is generated by its rows.

- (c) Part (b) and Theorem 3.5 tells us that the rows and columns generate subspaces of the same dimension that are numerically equal to the rank of A . ■

Exercise 3.2.14

Let $T, U : V \rightarrow W$ be linear transformations.

- (a) Prove that $R(T + U) \subseteq R(T) + R(U)$. (See the definition of the sum of subsets of a vector space in Section 1.3).

Proof. Let $y \in R(T+U)$. Then for some $x \in V$, we know that $(T+U)(x) = y$. Since U and T are both linear, we have that

$$y = (T+U)(x) = T(x) + U(x).$$

But note that $T(x) \in R(T)$ and $U(x) \in R(U)$. Thus, $y \in R(T) + R(U)$ and so $R(T+U) \subseteq R(T) + R(U)$. ■

- (b) Prove that if W is finite-dimensional, then $\text{rank}(T+U) \leq \text{rank}(T) + \text{rank}(U)$.

Proof. Let W be a finite-dimensional vector space. Since $R(T+U)$ and $R(T) + R(U)$ are subspaces of W , we know that these subspaces are also finite-dimensional vector spaces by Theorem 1.11. By part (a), we can see that

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

Thus, we have that

$$\text{rank}(T+U) \leq \text{rank}(T) + \text{rank}(U). \quad \blacksquare$$

- (c) Deduce from (b) that $\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$ for any $m \times n$ matrices A and B .

Proof. Observe that

$$\begin{aligned} \text{rank}(A+B) &= \text{rank}(L_{A+B}) \\ &= \text{rank}(L_A + L_B) && \text{(part (c) of Theorem 2.15)} \\ &\leq \text{rank}(L_A) + \text{rank}(L_B) && \text{(part (b))} \\ &= \text{rank}(A) + \text{rank}(B). \end{aligned}$$

Hence, we can see that $\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$. ■

Exercise 3.2.15

Suppose that A and B are matrices having n rows. Prove that $M(A|B) = (MA|MB)$ for any $m \times n$ matrices M .

Proof. Let A and B be $n \times p$ and $n \times \ell$ matrices, respectively. Suppose M be is an arbitrary $m \times n$ matrix. Then define the product $M(A|B)$ as

$$M(A|B) = \sum_{k=1}^n M_{ik}(A|B)_{kj} \quad (1)$$

for $1 \leq i \leq m$ and $1 \leq j \leq p + \ell$. For $1 \leq j \leq p$, we can see that product in (1) can be re-written as

$$M(A|B) = \sum_{k=1}^n M_{ik} A_{kj} = MA. \quad (2)$$

For $p \leq j \leq \ell$, (1) can be re-written into

$$M(A|B) = \sum_{k=1}^n M_{ik} B_{kj} = MB. \quad (3)$$

So, with (2) and (3) we can write that

$$M(A|B) = (MA|MB).$$

■

Exercise 3.2.16

Supply the details to the proof of (b) of Theorem 3.4.

Proof. Observe that

$$\begin{aligned} R(L_{PA}) &= R(L_P L_A) \\ &= L_P L_A(F^n) \\ &= L_P(L_A(F^n)) \\ &= L_P(R(L_A)). \end{aligned} \quad (1)$$

Note that $R(L_A)$ is a subspace of F^m . By exercise 17 of Section 2.4, we can see that the invertibility of L_P also implies that $L_P(R(L_A))$ is also a subspace of F^m . Thus, we have that $\dim(R(L_A)) = \dim(L_P(R(L_A)))$ implies $R(L_A) = L_P(R(L_A))$ by Theorem 1.11. So (1) implies that $R(L_{PA}) = R(L_A)$ and thus

$$\text{rank}(PA) = \text{rank}(A).$$

■

Exercise 3.2.17

Prove that if B is a 3×1 matrix and C is a 1×3 matrix, then the 3×3 matrix BC has rank at most 1. Conversely, show that if A is any 3×3 matrix having rank 1, then there exists a 3×1 matrix B and a 1×3 matrix C such that $A = BC$.

Proof. Suppose B is a 3×1 matrix and C is a 1×3 matrix where

$$B = \begin{pmatrix} B_{11} \\ B_{21} \\ B_{31} \end{pmatrix}$$

and

$$C = (C_{11} \ C_{12} \ C_{13}).$$

Computing BC gives us

$$BC = \begin{pmatrix} B_{11}C_{11} & B_{11}C_{12} & B_{11}C_{13} \\ B_{21}C_{11} & B_{21}C_{12} & B_{21}C_{13} \\ B_{31}C_{11} & B_{31}C_{12} & B_{31}C_{13} \end{pmatrix}.$$

Observe that each column vector of this matrix is just B scaled by entries in C ; that is,

$u_j = BC_{1j}$ for $1 \leq j \leq n$ where u_j has a rank of 1 and u_j is j th column vector of BC . Thus, BC has a rank of at most 1.

Conversely, let $A \in M_{3 \times 3}(F)$ with $\text{rank}(A) = 1$ and let u_j for $1 \leq j \leq 3$ denote the j th column vector of A . Since $\text{rank}(A) = 1$, there exists only one linearly independent column, say u_1 . This tells us that the other two columns, namely u_2 and u_3 , are scalar multiples of u_1 . Hence, for some nonzero k and c in F , we have that

$$u_2 = ku_1 \text{ and } u_3 = cu_1.$$

Now we can write A in the following way

$$A = (u_1 \quad ku_1 \quad cu_1) = u_1 (1 \quad k \quad c)$$

where

$$u_1 = B \text{ and } C = (1 \quad k \quad c)$$

both having rank 1. ■

Exercise 3.2.18

Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Prove that AB can be written as a sum of n matrices of rank one.

Proof. Let u_j be the j th column vector of AB . By part (b) Exercise 2.3.14, we can see that u_j can be written in terms of a linear combination of column vectors of A and entries from B as coefficients. In other words, we have

$$u_j = \sum_{i=1}^n B_{ij}x_i \text{ for } 1 \leq j \leq p$$

where $1 \leq i \leq n$ and x_i is the i th column vector of A and that each x_i has a rank of 1. ■

Exercise 3.2.19

Let A be an $m \times n$ matrix with rank m and B be an $n \times p$ matrix with rank n . Determine the rank of AB . Justify your answer.

Proof. We determine that $\text{rank}(AB)$ can be at most n by Theorem 3.7. ■

Exercise 3.2.21

Let A be an $m \times n$ matrix with $\text{rank}(A) = m$. Prove that there exists an $n \times m$ matrix B such that $AB = I_m$.

Proof. Observe that $\text{rank}(A) = \text{rank}(A^t) = m$. This tells us that A^t is an invertible matrix. By Theorem 3.4, we can write that $\text{rank}(AA^t) = \text{rank}(A) = m$. Hence, we have that AA^t is an invertible matrix and thus there must exist a matrix B' such that

$$(AA^t)B' = I_m \Leftrightarrow A(A^tB') = I_m$$

where $A^tB' = B$ is an $n \times m$ matrix. ■

Exercise 3.2.22

Let B be an $n \times m$ matrix with $\text{rank}(B) = m$. Prove that there exists an $m \times n$ matrix A such that $AB = I_m$.

Proof. Observe that $\text{rank}(B) = \text{rank}(B^t) = m$ by Corollary 2 to Theorem 3.6. By Corollary 2 to Theorem 2.18 and fact 3 found in page 103, $\text{rank}(B^t) = m$ implies that B^t is invertible. This implies that

$$\text{rank}(B^t B) = \text{rank}(B) = m$$

by Corollary 2 to Theorem 3.6. Again, $B^t B$ is also invertible and thus there exist a matrix A' such that

$$A'(B^t B) = I_m \Leftrightarrow (A' B^t) B = I_m$$

where $A = A' B^t$ is an $m \times n$ matrix. ■

Chapter 4

Determinants

4.1 Determinants of Order 2

Exercise 4.1.5

Prove that if B is the matrix obtained by interchanging the rows of a 2×2 matrix A , then $\det(B) = -\det(A)$.

Proof. Suppose $B \in M_{2 \times 2}(F)$ can be obtained by interchanging the rows of $A \in M_{2 \times 2}(F)$ where

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{pmatrix}.$$

Applying the definition of determinant for 2×2 matrices, we get that

$$\begin{aligned} \det(B) &= A_{12}A_{21} - A_{11}A_{22} \\ &= -(A_{11}A_{22} - A_{12}A_{21}) \\ &= -\det(A). \end{aligned}$$

Hence, we have that $\det(B) = -\det(A)$. ■

Exercise 4.1.6

Prove that if the two columns of $A \in M_{2 \times 2}(F)$ are identical, then $\det(A) = 0$.

Proof. Let $A \in M_{2 \times 2}(F)$ be defined as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

with $A_{11} = A_{12}$ and $A_{21} = A_{22}$ by assumption. Applying the definition of the determinant gives us $\det(A) = 0$. ■

Exercise 4.1.7

Prove that $\det(A^t) = \det(A)$ for any $A \in M_{2 \times 2}(F)$.

Proof. Let $A \in M_{2 \times 2}(F)$ where

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad A^t = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix}.$$

Applying the determinant to A and A^t , we can see that

$$\det(A) = A_{11}A_{22} - A_{12}A_{21} = \det(A^t).$$

■

Exercise 4.1.8

Prove that if $A \in M_{2 \times 2}(F)$ is upper triangular, then $\det(A)$ equals the product of diagonal entries of A .

Proof. Let $A \in M_{2 \times 2}(F)$ be an upper triangular matrix. Then we have

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}.$$

Using the definition of determinant, we have

$$\det(A) = A_{11}A_{22} - 0A_{12} = A_{11}A_{22}.$$

Hence, $\det(A)$ equals the product of diagonal entries of A .

■

Exercise 4.1.9

Prove that $\det(AB) = \det(A)\det(B)$ for any $A, B \in M_{2 \times 2}(F)$.

Proof. Let $A, B \in M_{2 \times 2}(F)$ with

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

Using the definition of the matrix product, we have that

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

By definition of the determinant, we have

$$\begin{aligned} \det(AB) &= (A_{11}B_{11} + A_{12}B_{21})(A_{21}B_{12} + A_{22}B_{22}) \\ &\quad - (A_{21}B_{11} + A_{22}B_{21})(A_{11}B_{12} + A_{12}B_{22}) \\ &= A_{11}A_{21}B_{11}B_{12} + A_{12}A_{21}B_{12}B_{21} \\ &\quad + A_{12}A_{22}B_{21}B_{22} + A_{12}A_{21}B_{11}B_{22} \\ &= (A_{11}A_{22} - A_{12}A_{21})B_{11}B_{22} - (A_{11}A_{22} - A_{12}A_{21})B_{12}B_{21} \\ &= (A_{11}A_{22} - A_{12}A_{21})(B_{11}B_{22} - B_{12}B_{21}) \\ &= \det(A)\det(B). \end{aligned}$$

Hence, we have that $\det(AB) = \det(A)\det(B)$.

■

Exercise 4.1.10

The **classical adjoint** of a 2×2 matrix $A \in M_{2 \times 2}(F)$ is the matrix

$$C = \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}.$$

Prove that

- (a) $CA = AC = [\det(A)]I$.
- (b) $\det(C) = \det(A)$.
- (c) The classical adjoint of A^t is C^t .
- (d) If A is invertible, then $A^{-1} = [\det(A)]^{-1}C$.

Proof. Let $A, C \in M_{2 \times 2}(F)$ where C is the classical adjoint of A .

- (a) We show that $CA = [\det(A)]I$. Applying the definition of the matrix product, we get that

$$\begin{aligned} CA &= \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \\ &= \begin{pmatrix} A_{11}A_{22} - A_{12}A_{21} & A_{22}A_{12} - A_{12}A_{22} \\ A_{11}A_{21} - A_{11}A_{21} & A_{11}A_{22} - A_{12}A_{21} \end{pmatrix} \\ &= \begin{pmatrix} \det(A) & 0 \\ 0 & \det(A) \end{pmatrix} \\ &= [\det(A)]I. \end{aligned}$$

Note that proving $AC = [\det(A)]I$ is a similar process as above. Hence, we have $AC = CA = [\det(A)]I$

- (b) Applying the definition of determinant gives us

$$\det(C) = A_{11}A_{22} - A_{12}A_{21} = \det(A).$$

- (c) Note that the transpose of $A \in M_{2 \times 2}(F)$ is

$$A^t = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix}$$

where the classical adjoint of A^t is

$$W = \begin{pmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{pmatrix}.$$

But note that this is just the transpose of C which was defined earlier. So, we have $W = C^t$ is the classical adjoint of A^t .

- (d) Since A is an invertible 2×2 matrix, we can write

$$\begin{aligned} A^{-1} &= \frac{1}{\det(A)} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} \\ &= [\det(A)]^{-1}C \end{aligned}$$

by Theorem 4.2. ■

Exercise 4.1.11

Let $\delta : M_{2 \times 2}(F) \rightarrow F$ be a function with the following three properties.

- (i) δ is a linear function of each row of the matrix when the other row is held fixed.
- (ii) If the two rows of $A \in M_{2 \times 2}(F)$ are identical $\delta(A) = 0$.
- (iii) If I is the 2×2 identity matrix, then $\delta(I) = 1$.

Prove that $\delta(A) = \det(A)$ for all $A \in M_{2 \times 2}(F)$.

Proof. Let $A \in M_{2 \times 2}(F)$. Let the vectors $u, v \in F^2$ for an ordered basis β for F^2 . Define the row vectors of A by $u = (A_{11}, A_{12})$ and $v = (A_{21}, A_{22})$. Note that

$$A = \begin{pmatrix} u \\ v \end{pmatrix}$$

and that

$$\delta \begin{pmatrix} u \\ v \end{pmatrix} = O \begin{pmatrix} u \\ v \end{pmatrix} \cdot A \begin{pmatrix} u \\ v \end{pmatrix}$$

where O is the **orientation** of β and A is the area of the parallelogram formed by u and v . Since

$$A = \left| \det \begin{pmatrix} u \\ v \end{pmatrix} \right|,$$

and definition of O , we find that

$$\begin{aligned} \delta(A) &= \delta \begin{pmatrix} u \\ v \end{pmatrix} = O \begin{pmatrix} u \\ v \end{pmatrix} \cdot A \begin{pmatrix} u \\ v \end{pmatrix} \\ &= \frac{\det \begin{pmatrix} u \\ v \end{pmatrix}}{\left| \det \begin{pmatrix} u \\ v \end{pmatrix} \right|} \cdot \left| \det \begin{pmatrix} u \\ v \end{pmatrix} \right| \\ &= \det \begin{pmatrix} u \\ v \end{pmatrix} \\ &= \det(A). \end{aligned}$$

Hence, $\delta(A) = \det(A)$. ■

4.2 Determinants of Order n

Exercise 4.2.23

Prove that the determinant of an upper triangular matrix is the product of its diagonal entries.

Proof. We proceed by inducting on n . The result easily follows from cases $n = 1$ and $n = 2$, with the former leading to just the single entry itself and the latter being easily shown by applying the determinant formula found in Section 4.1 for 2×2 upper triangular matrices.

Let $A \in M_{n \times n}(F)$ be upper triangular. Now, assume the result holds for $(n-1) \times (n-1)$ upper triangular matrices and that $n \geq 3$ and $1 \leq j \leq n$. We can obtain a new matrix D by performing a finite number of type 3 row operations on A to convert the entries found in the first row and columns $1 < j \leq n$ into zeroes. Using cofactor expansion along the first

row, we find that

$$\begin{aligned}
 \det(D) &= \sum_{j=1}^n (-1)^{1+j} D_{1j} \cdot \det(\tilde{D}_{1j}) \\
 &= (-1)^{1+1} D_{11} \cdot \det(\tilde{D}_{11}) + (-1)^{1+2} D_{12} \cdot \det(\tilde{D}_{12}) + \cdots \\
 &\quad + (-1)^{1+n} D_{1n} \cdot \det(\tilde{D}_{1n}) \\
 &= (-1)^{1+1} D_{11} \cdot \det(\tilde{D}_{11}) + (-1)^{1+2}(0) \cdot \det(\tilde{D}_{12}) + \cdots \\
 &\quad + (-1)^{1+n}(0) \cdot \det(\tilde{D}_{1n}) \\
 &= (-1)^{1+1} D_{11} \cdot \det(\tilde{D}_{11}).
 \end{aligned}$$

Since \tilde{D}_{11} is an $(n-1) \times (n-1)$ matrix, we can apply the induction hypothesis to write that

$$\det(\tilde{D}_{11}) = \prod_{i=2}^n D_{ii}.$$

Thus, we must have that

$$\det(D) = (-1)^{1+1} D_{11} \cdot \det(\tilde{D}_{11}) = \prod_{i=1}^n D_{ii}.$$

Note that $\det(A) = \det(D)$ by Theorem 4.6 and that $D_{ii} = A_{ii}$ since we only changed the entries in the first row (that were not A_{11}). Thus, the result holds for $n \times n$ upper triangular matrices. \blacksquare

Exercise 4.2.24

Prove the corollary to Theorem 4.3.

Proof. We proceed by inducting on n . Let $A \in M_{n \times n}(F)$. If $n = 1$, then we simply have $A_{11} = 0$ and the result follows immediately. If $n = 2$, then the result follows from applying the formula for 2×2 determinants and choosing either the $i = 1$ or $i = 2$ as the rows with all zeros. Now, assume that the corollary holds for $(n-1) \times (n-1)$ matrices. Let $n \geq 3$ and $1 \leq j \leq n$. Let r represent the selected row of all zero entries. Suppose that for some r ($1 \leq r \leq n$), we have $a_r = u + kv$ for some $u, v \in F^n$ and $k \in F$. Let $u = (b_1, b_2, \dots, b_n)$ and $v = (c_1, c_2, \dots, c_n)$, and let B and C be the matrices obtained from A by replacing row r of A by u and v , respectively. Furthermore, let the row $r-1$ be defined as in the proof for Theorem 4.3. If $r = 1$, then by cofactor expansion along the first row, we have

$$\begin{aligned}
 \det(A) &= \sum_{j=1}^n (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j}) \\
 &= (-1)^{1+1}(0) \cdot \det(\tilde{A}_{11}) + (-1)^{1+2}(0) \cdot \det(\tilde{A}_{12}) + \cdots \\
 &\quad + (-1)^{1+n}(0) \cdot \det(\tilde{A}_{1n}) \\
 &= 0.
 \end{aligned}$$

Now, suppose $r > 1$. Then by Theorem 4.3, we get

$$\begin{aligned}\det(A) &= \det(B) + k\det(C) \\ &= \sum_{j=1}^n (-1)^{1+j} B_{1j} \cdot \det(\tilde{B}_{1j}) + k \sum_{j=1}^n (-1)^{1+j} C_{1j} \cdot \det(\tilde{C}_{1j}).\end{aligned}$$

Note that $A_{1j} = B_{1j} = C_{1j}$ and that $\tilde{A}_{1j} = \tilde{B}_{1j} = \tilde{C}_{1j}$ except for row $r - 1$. Since \tilde{B}_{1j} and \tilde{C}_{1j} are $(n - 1) \times (n - 1)$ matrices, we find that $\det(\tilde{B}_{1j}) = 0$ and $\det(\tilde{C}_{1j}) = 0$ by our induction hypothesis. So, we get that

$$\det(A) = \det(B) + k\det(C) = 0 + 0 = 0$$

and thus $\det(A) = 0$. Hence, we can see that the corollary holds for all $n \times n$ matrices. ■

Exercise 4.2.25

Prove that $\det(kA) = k^n \det(A)$ for any $A \in M_{n \times n}(F)$.

Proof. Let $A \in M_{n \times n}(F)$ and $k \in F$ be nonzero. We proceed by inducting on n . If $n = 1$, then the result follows immediately. If $n = 2$, then apply the formula found in section 4.1 for 2×2 matrices to A to get $\det(kA) = k^2 \det(A)$.

Suppose $n > 2$ and suppose the result holds for $(n - 1) \times (n - 1)$ matrices. Since \tilde{A}_{1j} is an $(n - 1) \times (n - 1)$ matrix, we must have $\det(k\tilde{A}_{1j}) = k^{n-1} \det(\tilde{A}_{1j})$. Using cofactor expansion along the first row, we get that

$$\begin{aligned}\det(A) &= \sum_{j=1}^n (-1)^{1+j} kA_{1j} \cdot \det(k\tilde{A}_{1j}) \\ &= k \sum_{j=1}^n (-1)^{1+j} A_{1j} \cdot k^{n-1} \det(\tilde{A}_{1j}) \\ &= k \cdot k^{n-1} \sum_{j=1}^n (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j}) \\ &= k^n \det(A).\end{aligned}$$

■

Exercise 4.2.26

Let $A \in M_{n \times n}(F)$. Under what conditions is $\det(-A) = \det(A)$.

Proof. In order for $\det(-A) = \det(A)$, A must be an even matrix. ■

Exercise 4.2.27

Prove that if $A \in M_{n \times n}(F)$ has two identical columns, then $\det(A) = 0$.

Proof. Let $A \in M_{n \times n}(F)$. Suppose that A contains two identical columns at row r and row s where $r \neq s$. Let u_1, u_2, \dots, u_n be the columns of A . Note that these are the rows of A^t . Observe that $u_r = u_s$ for $r \neq s$, we can apply the Corollary to Theorem 4.4 to write

$$\det(A^t) = 0.$$

Since $\text{rank}(A) = \text{rank}(A^t)$, we get that $\det(A) = \det(A^t) = 0$. Hence, $\det(A) = 0$ for all $A \in M_{n \times n}(F)$. ■

Proof. Let $A \in M_{n \times n}(F)$. We proceed by inducting on $n \geq 2$. For $n = 2$, let $u_1 = (A_{11}, A_{21})$ and $u_2 = (A_{12}, A_{22})$ denote the two columns that make up A . By Exercise 4.1.6, we get that $\det(A) = 0$.

Now, suppose $n > 2$ and that the result holds for $(n - 1) \times (n - 1)$ matrices. Let $1 \leq j \leq n$. Using cofactor expansion along any row i , we get that

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

by Theorem 4.4. Since \tilde{A}_{1j} is an $(n - 1) \times (n - 1)$ matrix, we know that \tilde{A}_{1j} contains identical columns u_r and u_s where $r \neq s$. Thus, we have that $\det(\tilde{A}_{1j}) = 0$ by our induction hypothesis. Thus, (1) implies that $\det(A) = 0$ which ends our induction argument. ■

Exercise 4.2.28

Compute $\det(E_i)$ if E_i is an elementary matrix of type i .

Solution. Let E_i be an elementary matrix of type i where $1 \leq i \leq n$. Computing $\det(E_i)$ depends on the type of elementary operations we perform on E_i . If we interchange any two rows or columns of E_i , then we would find that $\det(E_i) = -1 = -\det(I_i)$. If we add a scalar multiple k to E_i , then we would find that $\det(E_i) = c = c\det(I_i)$. Lastly, performing any type 3 elementary operation on E_i gives us $\det(E_i) = 1 = \det(I_i)$. ■

Exercise 4.2.29

Prove that if E is an elementary matrix, then $\det(E^t) = \det(E)$.

Proof. Suppose that E is an $n \times n$ elementary matrix. Observe that $E_{1j} = E^t_{j1}$ and so using cofactor expansion along the first row, of E , we see that

$$\begin{aligned} \det(E) &= \sum_{j=1}^n (-1)^{1+j} E_{1j} \cdot \det(\tilde{E}_{1j}) \\ &= \sum_{j=1}^n (-1)^{1+j} (E^t)_{j1} \cdot \det(\tilde{E}_{j1}^t) \\ &= \det(E^t). \end{aligned}$$

Note that the second equality above is just a cofactor expansion along the first column. Hence,

$$\det(E) = \det(E^t). \quad \blacksquare$$

Chapter 5

Diagonalization

5.1 Eigenvalues and Eigenvectors

Exercise 5.1.5

Prove Theorem 5.4.

Proof. Let $v \in V$ be an eigenvector of T corresponding to λ . Since λ is an eigenvalue of T , we know that $T(v) = \lambda v$ where $v \neq 0$. Hence,

$$\begin{aligned} T(v) = \lambda v &\Leftrightarrow T(v) - \lambda v = 0 \\ &\Leftrightarrow (T - \lambda I)(v) = 0. \end{aligned}$$

Note that this is true if and only if $v \in N(T - \lambda I)$. We can reverse this argument to show that $v \in V$ is an eigenvector of T corresponding to λ . ■

Exercise 5.1.6

Let T be a linear operator on a finite-dimensional vector space V , and let β be an ordered basis for V . Prove that λ is an eigenvalue of T if and only if λ is an eigenvalue of $[T]_\beta$.

Proof. Suppose λ is an eigenvalue of T . Since β is an ordered basis for V (suppose $\dim(V) = n$) and $T(v) = \lambda v$ for $v \neq 0$, we have that

$$[T(v)]_\beta = [\lambda v]_\beta = \lambda[v]_\beta$$

by Theorem 2.8. Using Theorem 2.14, we can write

$$\begin{aligned} [T(v)]_\beta = \lambda[v]_\beta &\Rightarrow [T]_\beta[v]_\beta = \lambda[v]_\beta \\ &\Rightarrow ([T]_\beta - \lambda I_n)[v]_\beta = 0 \end{aligned}$$

This implies that $[T]_\beta - \lambda I_n$ is not invertible for all $[v]_\beta$. So, $\det([T]_\beta - \lambda I_n) = 0$ which is true if and only if λ is an eigenvalue of $[T]_\beta$ by Theorem 5.2. This argument can be reversed to show that λ is an eigenvalue of T . ■

Exercise 5.1.7

Let T be a linear operator on a finite-dimensional vector space V . We define the **determinant** of T , denoted $\det(T)$, as follows: Choose any ordered basis β for V , and define $\det(T) = \det([T]_\beta)$.

- (a) Prove that the preceding definition is independent of the choice of an ordered basis for V , then $\det([T]_\beta) = \det([T]_\gamma)$.

Proof. If β and γ are two ordered bases for V , then we must have $\det(T) = \det([T]_\beta)$ and $\det(T) = \det([T]_\gamma)$. This implies that $\det([T]_\beta) = \det([T]_\gamma)$. ■

- (b) Prove that T is invertible if and only if $\det(T) \neq 0$.

Proof. We have T is invertible if and only if $[T]_\beta$ is invertible by Corollary 1 to Theorem 2.18. This is true if and only if $\det([T]_\beta) \neq 0$ by Corollary to Theorem 4.7. Thus, we have $\det(T) = \det([T]_\beta) \neq 0$. Note that this argument is reversible. ■

- (c) Prove that if T is invertible, then $\det(T^{-1}) = [\det(T)]^{-1}$.

Proof. Suppose T is invertible. By part (b) and Theorem 2.18, we can see that $\det([T]_\beta) \neq 0$ and that $([T]_\beta)^{-1} = [T^{-1}]_\beta$. Thus,

$$\det(T^{-1}) = \det([T^{-1}]_\beta) = \det\left([T]_\beta^{-1}\right).$$

■

- (d) Prove that if U is also a linear operator on V , then $\det(TU) = \det(T) \cdot \det(U)$.

Proof. Let β be an ordered basis for V and let TU and $[TU]_\beta$ and $[T]_\beta[U]_\beta$ be defined. Then we have

$$\begin{aligned} \det(TU) &= \det([TU]_\beta) \\ &= \det([T]_\beta[U]_\beta) && \text{(Corollary to Theorem 2.11)} \\ &= \det([T]_\beta)\det([U]_\beta) && \text{(Corollary to Theorem 4.7)} \\ &= \det(T)\det(U). \end{aligned}$$

Thus, we have $\det(TU) = \det(T)\det(U)$. ■

- (e) Prove that $\det(T - \lambda I_V) = \det([T]_\beta - \lambda I)$ for any scalar λ and any ordered basis β for V .

Proof. Let $\lambda \in F$ and let β be an ordered basis for V . Then we have

$$\begin{aligned} \det(T - \lambda I_V) &= \det([T - \lambda I_V]_\beta) \\ &= \det([T]_\beta - \lambda[I_V]_\beta) && \text{(parts (a) and (b) of Theorem 2.8)} \\ &= \det([T]_\beta - \lambda I) \end{aligned}$$

where I is the matrix representation of I_V . ■

Exercise 5.1.8

- (a) Prove that a linear operator T on a finite-dimensional vector space is invertible if and only if zero is not an eigenvalue of T .

Proof. Suppose T is an invertible linear operator. Suppose for sake of contradiction that $\lambda = 0$ is an eigenvalue of T . Using Theorem 5.2, we write

$$\det([T]_\beta - \lambda I) = \det([T]_\beta) = 0.$$

Note that $\det([T]_\beta) \neq 0$ since T is invertible which is a contradiction. Thus, $\lambda = 0$

cannot be an eigenvalue of T if T is invertible.

For the backwards direction, we proceed via proving the contrapositive. Let $\lambda = 0$. Observe that $\det([T]_\beta)$ can be written in the following way:

$$\det([T]_\beta) = \det([T]_\beta - \lambda I).$$

Since T is not invertible, we have $\det([T]_\beta) = 0$ implies $\det([T]_\beta - \lambda I) = 0$. Thus, $\lambda = 0$ is an eigenvalue of T . ■

- (b) Let T be an invertible linear operator. Prove that a scalar λ is an eigenvalue of T if and only if λ^{-1} is an eigenvalue of T^{-1} .

Proof. Suppose λ is an eigenvalue of T . Then for $v \neq 0$ we have $T(v) = \lambda v$. Since T is an invertible linear operator, we have $\lambda \neq 0$ (by part (a)) implies

$$\begin{aligned} T(v) = \lambda v &\Rightarrow T^{-1}(T(v)) = T^{-1}(\lambda v) \\ &\Rightarrow T^{-1}T(v) = T^{-1}(\lambda v) \\ &\Rightarrow I_V(v) = \lambda T^{-1}(v) \\ &\Rightarrow \lambda^{-1}v = T^{-1}(v). \end{aligned}$$

Thus, λ^{-1} is an eigenvalue of T^{-1} . To prove the backwards direction, we can just reverse the argument written above. ■

- (c) State and prove results analogous to (a) and (b) for matrices.

Proof. Let $A \in M_{n \times n}(F)$. The analogous results to (a) and (b) are

A is invertible if and only if 0 is not an eigenvalue of A

and

If L_A is an invertible linear operator, then λ is an eigenvalue of A if and only if λ^{-1} is an eigenvalue of L_A^{-1} .

To prove the two results above, apply parts (a) and (b) to the linear operator L_A . ■

Exercise 5.1.9

Prove that the eigenvalues of an upper triangular matrix M are the diagonal entries of M .

Proof. Let M be an upper triangular matrix. Let $\lambda_j \in F$ be the eigenvalues of M and v_j be the corresponding eigenvectors. By Theorem 5.4, we see that ■

Exercise 5.1.10

Let V be a finite-dimensional vector space, and let λ be any scalar.

- (a) For any ordered basis β for V , prove that $[\lambda I_V]_\beta = \lambda I$.

Proof. Let β be an ordered basis for V . Using part (b) of Theorem 2.8, we obtain

$$[\lambda I_V]_\beta = \lambda [I_V]_\beta = \lambda I$$

where I is the identity matrix. ■

- (b) Compute the characteristic polynomial of λI_V .

Solution. The characteristic polynomial of λI_V is

$$f(t) = \det(\lambda I_V - tI_V) = \det(O) = 0$$

where O is the zero matrix. ■

- (c) Show that λI_V is diagonalizable and has only one eigenvalue.

Proof. Since V is a finite-dimensional vector space, let $\beta = \{v_1, v_2, \dots, v_n\}$ be an ordered basis for V . Thus, we have

$$\lambda I_V(v_j) = \lambda v_j \text{ for all } 1 \leq j \leq n.$$

Since each v_j is linearly independent, we have that each column of $[\lambda I_V]_\beta$ is linearly independent where $\lambda_j = ([\lambda I_V]_\beta)_{ij} = \lambda$ for $i = j$ and 0 elsewhere. Thus, we have that λI_V is diagonalizable by Theorem 5.1 and that λ is its only eigenvalue. ■

Exercise 5.1.11

A **scalar matrix** is a square matrix of the form λI for some scalar λ ; that is, a scalar matrix is a diagonal matrix in which all the diagonal entries are equal.

- (a) Prove that if a square matrix A is similar to a scalar matrix λI , then $A = \lambda I$.

Proof. Suppose A is a square matrix that is similar to a scalar matrix λI . Then there exists an invertible square matrix Q such that $A = Q^{-1}\lambda IQ$. Observe that

$$A = \lambda(Q^{-1}IQ) = \lambda(Q^{-1}Q) = \lambda I$$

by part (b) of Theorem 2.15. ■

- (b) Show that a diagonalizable matrix having only one eigenvalue is a scalar matrix.

Proof. Let β be an ordered basis for F^n . Since $A = \lambda I$ by part (a), and that λI_{F^n} is a diagonalizable and has only one eigenvalue, then $A = [\lambda I_{F^n}]_\beta$ must be a diagonal matrix where the all the diagonal entries are equal to each other. ■

- (c) Prove that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable.

Proof. Note that $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not a diagonal matrix since $A_{12} = 1 \neq 0$. Thus, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable. ■

Exercise 5.1.12

- (a) Prove that similar matrices have the same characteristic polynomial.

Proof. Let A, B be square matrices. Let the characteristic polynomials of A and B be

$$f(t) = \det(A - tI) \text{ and } g(t) = \det(B - tI)$$

respectively. We claim that $A - tI_n$ is similar to $B - tI_n$ which will allow us to show that $\det(A - tI) = \det(B - tI)$. Since $A \sim B$, there exists an invertible square matrix Q such that $A = Q^{-1}BQ$. Then observe using Theorem 2.12, we can write

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

Thus, $A - tI \sim B - tI$ and so, $f(t) = \det(A - tI) = \det(B - tI) = g(t)$. ■

- (b) Show that the definition of the characteristic polynomial of a linear operator on a finite-dimensional vector space V is independent of the choice of basis for V .

Proof. Let β and γ be two ordered bases for V and let $[T]_\beta$ and $[T]_\gamma$ be defined. Using part (e) of Exercise 7, we obtain

$$\det([T]_\beta - \lambda I) = \det(T - \lambda I_V) = \det([T]_\gamma - \lambda I_V). \quad \blacksquare$$

Exercise 5.1.13

Let T be a linear operator on a finite-dimensional vector space V over a field F , let β be an ordered basis for V , and let $A = [T]_\beta$. In reference, to figure 5.1, prove the following.

- (a) If $v \in V$ and $\phi_\beta(v)$ is an eigenvector of A corresponding to the eigenvalue λ , then v is an eigenvector of T corresponding to λ .

Proof. Suppose $v \in V$ and $\phi_\beta(v)$ is an eigenvector of A corresponding to λ . By the linearity of ϕ_β , we have that

$$L_A \phi_\beta(v) = \lambda \phi_\beta(v) = \phi_\beta(\lambda v).$$

Since $L_A \phi_\beta = \phi_\beta T$ (by Figure 5.1) and ϕ_β is an injective map, we obtain

$$\begin{aligned} L_A \phi_\beta(v) &= \phi_\beta T(v) \\ \lambda \phi_\beta(v) &= \phi_\beta T(v) \\ \phi_\beta(\lambda v) &= \phi_\beta(T(v)) \end{aligned}$$

which implies that $T(v) = \lambda v$ where $v \neq 0$. Hence, v is an eigenvector of T corresponding to λ . ■

- (b) If λ is an eigenvalue of A (and hence of T), then a vector $y \in F^n$ is an eigenvector of A corresponding to λ if and only if $\phi_\beta^{-1}(y)$ is an eigenvector of T corresponding to λ .

Proof. Suppose that $y \in F^n$ is an eigenvector of A corresponding to λ . By definition, we have

$$L_A(y) = \lambda y. \quad (1)$$

Since ϕ_β is a surjective mapping, we know that $y = \phi_\beta(x)$ for some $x \in F^n$. This tells us that x is an eigenvector of T corresponding to λ by part (a). Since ϕ_β is also invertible, we have that $x = \phi_\beta^{-1}(y)$ implies

$$T(\phi_\beta^{-1}(y)) = \lambda \phi_\beta^{-1}(y).$$

Thus, $\phi_\beta^{-1}(y)$ is the eigenvector of T corresponding to λ .

Conversely, if $\phi_\beta^{-1}(y)$ is an eigenvector of T corresponding to λ , then we have

$$T(\phi_\beta^{-1}(y)) = \lambda \phi_\beta^{-1}(y).$$

Using the linearity of ϕ_β^{-1} , we obtain

$$T(\phi_\beta^{-1}(y)) = \phi_\beta^{-1}(\lambda y). \quad (2)$$

Now, applying ϕ_β on both sides of (2), we obtain

$$\phi_\beta T(\phi_\beta^{-1}(y)) = \lambda y. \quad (3)$$

Since $L_A \phi_\beta = \phi_\beta T$, we get that

$$L_A(\phi_\beta \phi_\beta^{-1}(y)) = \lambda y \Leftrightarrow L_A(y) = \lambda y$$

with $y \in F^n$. Thus, we have that y is an eigenvector of A corresponding to λ . ■

Exercise 5.1.14

For any square matrix A , prove that A and A^t have the same characteristic polynomial (and hence the same eigenvalues).

Proof. Let A be an arbitrary square matrix. Let $f(t) = \det(A - \lambda I)$ and $g(t) = \det(A^t - \lambda I)$. Using the properties of the transpose, observe that

$$(A - \lambda I)^t = A^t - (\lambda I)^t = A^t - \lambda I.$$

By Theorem 4.8, we obtain that $\det(A - \lambda I) = \det(A - \lambda I)^t = \det(A^t - \lambda I)$. Thus, we have $g(t) = f(t)$ and we are done. ■

Exercise 5.1.15

- (a) Let T be a linear operator on a vector space V , and let x be an eigenvector of T corresponding to the eigenvalue λ . For any positive integer m , prove that x is an eigenvalue of T^m corresponding to the eigenvalue λ^m .

Proof. We proceed via induction on m . Let $m = 1$. Then x is an eigenvalue of T which implies $T(x) = \lambda x$. Now, assume this result holds for $m - 1$ case. We will show that m th case. By induction hypothesis, we have that

$$\begin{aligned} T^m(x) &= T^{(m-1)+1}(x) = T^{m-1}T(x) \\ &= T^{m-1}(T(x)) \\ &= \lambda^{m-1}T(x) \\ &= \lambda^{m-1}\lambda x \\ &= \lambda^m x. \end{aligned}$$

Hence, $T^m(x) = \lambda^m x$ which ends our induction argument. ■

- (b) State and prove the analogous result for matrices.

Proof. The analogous result for matrices is

For any positive integer m , prove that x is an eigenvalue of A corresponding to the eigenvalue λ^m .

This can be proven by applying part (a) to L_A . ■

Exercise 5.1.16

- (a) Prove that similar matrices have the same trace.

Proof. See Exercise 2.5.10. ■

- (b) How would you define the trace of a linear operator on a finite-dimensional vector space? Justify that your definition is well-defined.

Proof. I would define the trace of a linear operator by the following notation:

$$\operatorname{tr}(T) = \operatorname{tr}([T]_{\beta}).$$

Since V is a finite-dimensional vector space, let β and γ be two ordered bases such that $[T]_{\beta}$ and $[T]_{\gamma}$ are defined. Observe that

$$\operatorname{tr}([T]_{\beta}) = \operatorname{tr}(T) = \operatorname{tr}([T]_{\gamma}).$$

■

Exercise 5.1.17

Let T be the linear operator on $M_{n \times n}(\mathbb{R})$ defined by $T(A) = A^t$.

- (a) Show that ± 1 are the only eigenvalues of T .

Proof. ■

- (b) Describe the eigenvectors corresponding to each eigenvalue of T .

Solution. ■

- (c) Find an ordered basis β for $M_{2 \times 2}(\mathbb{R})$ such that $[T]_{\beta}$ is a diagonal matrix.

Solution. ■

- (d) Find an ordered basis β for $M_{n \times n}(\mathbb{R})$ such that $[T]_{\beta}$ is a diagonal matrix.

Solution. ■

Exercise 5.1.19

Let A and B be similar $n \times n$ matrices. Prove that there exists an n -dimensional vector space V , a linear operator T on V , and ordered bases β and γ for V such that $A = [T]_{\beta}$ and $B = [T]_{\gamma}$.

Proof. Let $A, B \in M_{n \times n}(F)$. Since A and B are similar matrices, there exists an invertible

$n \times n$ matrix Q such that

$$A = Q^{-1}BQ.$$

By Exercise 2.5.14, there exists an n dimensional vector space V , ordered bases β and γ for V , and a linear operator $T : V \rightarrow V$ such that

$$A = [T]_{\beta} \text{ and } B = [T]_{\gamma}.$$

■

Exercise 5.1.20

Let A be an $n \times n$ matrix with characteristic polynomial

$$f(t) = (-1)^n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0.$$

Prove that $f(0) = a_0 = \det(A)$. Deduce that A is invertible if and only if $a_0 \neq 0$.

Proof. The characteristic polynomial of A is just

$$f(t) = \det(A - tI) = (-1)^n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0.$$

Letting $t = 0$, we obtain

$$\begin{aligned} f(0) &= \det(A - 0 \cdot I) = \det(A) \\ &= (-1)^n (0)^n + a_{n-1} (0)^{n-1} + \cdots + a_1 (0) + a_0 \\ &= a_0. \end{aligned}$$

Suppose that A is an invertible matrix. Using the corollary to Theorem 4.7, we know that this is true if and only if $\det(A) \neq 0$. So, $\det(A) = a_0 \neq 0$. ■

Exercise 5.1.21

Let A and $f(t)$ be as in Exercise 20.

- (a) Prove that $f(t) = (A_{11} - t)(A_{22} - t) \cdots (A_{nn} - t) + q(t)$ where $q(t)$ is a polynomial of degree at most $n - 2$.

Proof. We proceed via induction on $n \geq 2$. If $n = 2$, then we have

$$\begin{aligned} f(t) &= \det(A - tI) = \det \begin{pmatrix} A_{11} - t & A_{12} \\ A_{21} & A_{22} - t \end{pmatrix} \\ &= (A_{11} - t)(A_{22} - t) - A_{12}A_{21} \end{aligned}$$

where $q(t) = A_{21}A_{12}$ is a polynomial of degree 0.

Assume the result holds for polynomials with degree less than n . Note that $(A_{nn} - t)$ divides $f(t)$ is true if and only if $f(A_{nn}) = 0$ by Corollary to the Division Algorithm for polynomials. Thus, there exists a polynomial $g(t)$ such that $f(t) = (A_{nn} - t)g(t)$. Since $g(t)$ is a polynomial with degree less than n , we have

$$g(t) = (A_{11} - t)(A_{22} - t) \cdots (A_{n-1, n-1} - t) + \phi(t)$$

with $\phi(t)$ with degree $n - 1$ by the induction hypothesis. Thus, we can see that

$$\begin{aligned} f(t) &= (A_{nn} - t)g(t) = (A_{11} - t)(A_{22} - t) \cdots (A_{nn} - t) + (A_{nn} - t)\phi(t) \\ &= (A_{11} - t)(A_{22} - t) \cdots (A_{nn} - t) + q(t) \end{aligned}$$

with $q(t) = (A_{nn} - t)\phi(t)$ having a degree less than $n - 2$. This ends our induction argument. ■

(b) Show that $\text{tr}(A) = (-1)^{n-1}a_{n-1}$.

Proof. Let A be defined as in Exercise 20. If $n = 2$, then part (a) implies that

$$f(t) = \det(A - tI) = t^2 - (A_{11} + A_{22})t - (A_{11}A_{22} + A_{12}A_{21})$$

where $\text{tr}(A) = A_{11} + A_{22}$. Then equating coefficients with the way $f(t)$ is defined in Exercise 20, we obtain that $\text{tr}(A) = (-1)^{2-1}a_1$. A similar process is done when $n = 3$ where we get $\text{tr}(A) = (-1)^{3-1}a_2$. If we continue this pattern then we will find that $\text{tr}(A) = (-1)^{n-1}a_{n-1}$.

We can show this via an induction argument. We claim that $f(t)$ can be expanded into

$$f(t) = (-1)^n t^n + (-1)^{n-1} \text{tr}(A) t^{n-1} + \phi(t)$$

where $\phi(t)$ is a polynomial of degree less than n . Suppose the argument above holds for $(n-1) \times (n-1)$ matrices. If $(A_{nn} - t)$ divides $f(t)$, we get that $f(t) = (A_{nn} - t)g(t)$ where $g(t)$ is some polynomial of degree $n - 1$. Using our induction hypothesis, $g(t)$ can be expanded into

$$g(t) = (-1)^{n-1} t^{n-1} + (-1)^{n-2} \text{tr}(A) t^{n-2} + \phi(t).$$

Distributing through by $(A_{nn} - t)$, we get that

$$\begin{aligned} f(t) &= (A_{nn} - t)[(-1)^{n-1} t^{n-1} + (-1)^{n-2} \text{tr}(A) t^{n-2} + \phi(t)] \\ &= (A_{nn} - t)(-1)^{n-1} t^{n-1} + (-1)^{n-2} \text{tr}(A) t^{n-2} (A_{nn} - t) + (A_{nn} - t)\phi(t) \\ &= (-1)^n t^n + (\text{tr}(A) + A_{nn})(-1)^{n-1} t^{n-1} + (-1)^{n-2} A_{nn} \text{tr}(A) t^{n-2} + (A_{nn} - t)\phi(t) \\ &= (-1)^n t^n + (-1)^{n-1} \left(\sum_{i=1}^n A_{ii} \right) t^{n-1} + (-1)^{n-2} \text{tr}(A) A_{nn} t^{n-2} + (A_{nn} - t)\phi(t) \\ &= (-1)^n t^n + (-1)^{n-1} \left(\sum_{i=1}^n A_{ii} \right) t^{n-1} + r(t) \end{aligned}$$

where

$$r(t) = (-1)^{n-2} \text{tr}(A) t^{n-2} + (A_{nn} - t)\phi(t)$$

is a polynomial of degree less than n . Equating coefficients with $f(t)$, we get that

$$a_{n-1} = (-1)^{n-1} \left(\sum_{i=1}^n A_{ii} \right)$$

which implies that

$$\sum_{i=1}^n A_{ii} = \text{tr}(A) = (-1)^{n-1} a_{n-1}$$

which ends our induction argument. ■

Exercise 5.1.22

- (a) Let T be a linear operator on a vector space V over the field F , and let $g(t)$ be a polynomial with coefficients from F . Prove that if x is an eigenvector of T with corresponding eigenvalue λ , then $g(T)(x) = g(\lambda)(x)$. That is, x is an eigenvector of $g(T)$ with corresponding eigenvalue $g(\lambda)$.

Proof. Suppose x is an eigenvector of T with corresponding eigenvalue λ . Using part (a) of Exercise 5.1.15, we obtain

$$\begin{aligned} g(T)(x) &= (a_n T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I)(x) \\ &= a_n T^n(x) + a_{n-1} T^{n-1}(x) + \cdots + a_1 T(x) + a_0 I(x) \\ &= a_n \lambda^n(x) + a_{n-1} \lambda^{n-1}(x) + \cdots + a_1 \lambda x + a_0 x \\ &= (a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0)(x) \\ &= g(\lambda)(x). \end{aligned}$$

Thus, we have $g(T)(x) = g(\lambda)(x)$. ■

- (b) State and prove a comparable result for matrices.

Proof. Replace T with L_A and apply part (a). ■

- (c) Verify (b) for the matrix A in Exercise 4(a) with polynomial $g(t) = 2t^2 - t + 1$, eigenvector $x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, and corresponding eigenvalue $\lambda = 4$.

Solution. Let $\lambda = 4$,

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \text{ and } x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

First, we will compute $g(\lambda)(x)$. Observe that

$$\begin{aligned} g(4)(x) &= (2(4)^2 - 4 + 1)(x) \\ &= 29 \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 58 \\ 87 \end{pmatrix}. \end{aligned}$$

Now, we will compute $g(L_A)(x)$. Observe that $A^2 = \begin{pmatrix} 32 \\ 48 \end{pmatrix}$. Thus, we have

$$\begin{aligned} g(L_A)(x) &= (2L_A^2 - L_A + 1I)(x) \\ &= 2A^2(x) - A(x) - I(x) \\ &= 2 \begin{pmatrix} 32 \\ 48 \end{pmatrix} - \begin{pmatrix} 8 \\ 12 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 56 \\ 84 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 58 \\ 87 \end{pmatrix}. \end{aligned}$$

Thus, we get that $g(A)(x) = g(\lambda)(x)$. ■

Exercise 5.1.23

Use Exercise 22 to prove that if $f(t)$ is the characteristic polynomial of a diagonalizable linear operator T , then $f(T) = T_0$, the zero operator.

Proof. Since the linear operator T is diagonalizable, there exists an ordered basis $\beta = \{v_1, v_2, \dots, v_n\}$ for V consisting of eigenvectors v_j of T corresponding to eigenvalues λ_j . Denote the characteristic polynomial $f(t)$ of a diagonalizable linear operator T by

$$f(t) = \det([T]_{\beta} - tI_n).$$

Observe that for every λ_j , we have $f(\lambda_j) = 0$ by Theorem 5.2. Thus, using Exercise 22 we obtain the following:

$$f(T)(v_j) = f(\lambda_j)(v_j) = T_0(v_j).$$

Therefore, we have $f(T) = T_0$. ■

Exercise 5.1.24

Use Exercise 21 (a) to prove Theorem 5.3

Proof. Let $f(t)$ be the characteristic polynomial of A and let A be defined as in Exercises 20 and 21. Observe that exercise 20 and exercise 21 part (a) proves part (a) of Theorem 5.3. Since the characteristic polynomial of A is of degree n , its characteristic polynomial contains at most n distinct zeros by Corollary 1 to the Division Algorithm for Polynomials. But this can only happen if and only if t is an eigenvalue of A . Thus, A must have n distinct eigenvalues. This proves part (b). ■

5.2 Diagonalization

Exercise 5.2.5

State and prove the matrix version of Theorem 5.6.

Proof. Let $A \in M_{n \times n}(F)$ and let $f(t) = \det(A - tI)$ is the characteristic polynomial of A . Suppose L_A is diagonalizable. The proof is the same except T is replaced with L_A instead. Thus, the characteristic polynomial $f(t)$ of A splits over F . ■

Exercise 5.2.8

Suppose that $A \in M_{n \times n}(F)$ has two distinct eigenvalues λ_1 and λ_2 and that $\dim(E_{\lambda_1}) = n - 1$. Prove that A is diagonalizable.

Proof. Let $A \in M_{n \times n}(F)$ and λ_1 and λ_2 be two distinct eigenvalues where E_{λ_1} and E_{λ_2} are the eigenspaces associated with these eigenvalues (related to the linear operator L_A). Let β_1 be a basis for E_{λ_1} . Since $\dim(E_{\lambda_1}) = n - 1$, we have

$$\beta_1 = \{v_1, v_2, \dots, v_{n-1}\}.$$

Note that the dimensions of E_{λ_1} and E_{λ_2} must sum to n . Thus, we must have $\dim(E_{\lambda_2}) = 1$

and let $\beta_2 = \{v'\}$ be the basis for E_{λ_2} where $v' \neq 0$. So, extend β_1 to a basis β for V taking the union of β_2 with this basis. Since $\beta = \beta_1 \cup \beta_2$ is a set that contains exactly n linearly independent vectors that are eigenvectors of L_A corresponding to eigenvalues λ_1 and λ_2 , we have that L_A must be diagonalizable by part (b) of Theorem 5.9. Thus, A must be diagonalizable. ■

Exercise 5.2.9

Let T be a linear operator on a finite-dimensional vector space V , and suppose there exists an ordered basis for V such that $[T]_\beta$ is an upper triangular matrix.

- (a) Prove that the characteristic polynomial for T splits.

Proof. Let

$$f(t) = \det([T]_\beta - tI)$$

be the characteristic polynomial of $[T]_\beta$. By Exercise 23 of Section 4.2, we can see that the determinant of this upper triangular matrix is just the product of its diagonal entries. Thus, we have

$$f(t) = (-1)^n ([T]_{\beta_{11}} - t)([T]_{\beta_{22}} - t) \cdots ([T]_{\beta_{nn}} - t).$$

Hence, $f(t)$ splits. ■

- (b) State and prove an analogous result for matrices.

Proof. Substitute L_A in place of T and apply part (a). ■

Exercise 5.2.10

Let T be a linear operator on a finite-dimensional vector space V with the distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ and corresponding multiplicities m_1, m_2, \dots, m_k . Suppose that β is a basis for V such that $[T]_\beta$ is an upper triangular matrix. Prove that the diagonal entries of $[T]_\beta$ are $\lambda_1, \lambda_2, \dots, \lambda_k$ and that each λ_i occurs m_i times ($1 \leq i \leq k$).

Proof. Let $[T]_\beta = A$ and $\dim(V) = n$. Since A is an upper triangular matrix, we know that the characteristic polynomial of A splits by Exercise 5.2.10. Thus, we have

$$\begin{aligned} f(t) &= \det(A - tI) \\ &= (-1)^n (A_{11} - t)(A_{22} - t) \cdots (A_{nn} - t). \end{aligned}$$

By Theorem 5.3, $f(t)$ contains at most n distinct eigenvalues. So, we must have $f(t) = 0$ if and only if $\lambda = A_{11}, A_{22}, \dots, A_{nn}$. But this means that each A_{ii} for $1 \leq i \leq n$ is an eigenvalue of A by Theorem 5.2. Thus, the eigenvalues of T must occur on the diagonals of A . But we only have distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ of T with corresponding multiplicities m_1, m_2, \dots, m_k where $k \leq n$. So, each m_i corresponding to each λ_i for $1 \leq i \leq k$ must sum up to n since $f(t)$ is an n th degree polynomial and there must be n solutions to $f(t) = 0$. Thus, each λ_i must occur m_i times along the diagonal of A . ■

Exercise 5.2.11

Let A be an $n \times n$ matrix that is similar to an upper triangular matrix and has the distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ with corresponding multiplicities m_1, m_2, \dots, m_k . Prove the following statements.

- (a) $\text{tr}(A) = \sum_{i=1}^k m_i \lambda_i$.

Proof. Let $A \in M_{n \times n}(F)$. Suppose A is similar to some upper triangular matrix B . Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be the eigenvalues with corresponding multiplicities m_1, m_2, \dots, m_k of B . Since these eigenvalues must occur on the diagonal of A (by Exercise 5.2.10) where each λ_i occurs m_i times along with $\text{tr}(A) = \text{tr}(B)$ (since A is similar to B), we can take the sum of our diagonal entries to get

$$\text{tr}(A) = \text{tr}(B) = \sum_{i=1}^k m_i \lambda_i.$$

■

- (b) $\det(A) = (\lambda_1)^{m_1} (\lambda_2)^{m_2} \dots (\lambda_k)^{m_k}$.

Proof. Since A is similar to B , we know that $\det(A) = \det(B)$ by Exercise 2.5.10. Since B is upper triangular, we know that by Exercise 4.2.23 that the determinant of B is just the product of its diagonal entries. Since the diagonal entries of B consists of eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ corresponding to multiplicities m_1, m_2, \dots, m_k , we have

$$\det(A) = \det(B) = \prod_{i=1}^k (\lambda_i)^{m_i}$$

which is our desired result. ■

Exercise 5.2.12

Let T be an invertible linear operator on a finite-dimensional vector space V .

- (a) Recall that for any eigenvalue of T , λ^{-1} is an eigenvalue of T^{-1} (Exercise 8 of Section 5.1). Prove that the eigenspace of T corresponding to λ is the same as the eigenspace of T^{-1} corresponding to λ^{-1} .

Proof. Let λ be an eigenvalue of T . Denote the eigenspaces E_λ and $E_{\lambda^{-1}}$ of T and T^{-1} respectively. We need to show that $E_\lambda = E_{\lambda^{-1}}$. Note that $E_\lambda = N(T - \lambda I)$ by definition. Let $x \in E_\lambda$. Then this is true if and only if $x \neq 0$ is an eigenvector of T corresponding to λ by Theorem 5.4. By Exercise 5.1.8, this is true if and only if x is an eigenvector of T^{-1} corresponding to λ^{-1} . By Theorem 5.4 again, this is true if and only if $x \in N(T^{-1} - \lambda^{-1}I)$. So, $x \in E_{\lambda^{-1}}$ and thus $E_\lambda \subseteq E_{\lambda^{-1}}$. We can reverse this argument to get the other containment. Thus, $E_{\lambda^{-1}} \subseteq E_\lambda$ and so we conclude that $E_\lambda = E_{\lambda^{-1}}$. ■

- (b) Prove that if T is diagonalizable, then T^{-1} is diagonalizable.

Proof. Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be the eigenvalues of T corresponding to multiplicities m_1, m_2, \dots, m_k . Suppose that T is diagonalizable. By Theorem 5.9, the multiplicity of each λ_i is equal to $\dim(E_{\lambda_i})$ for all $1 \leq i \leq k$. Using part (a), we obtain

$$m_i = \dim(E_{\lambda_i}) = \dim(E_{\lambda_i^{-1}}) \text{ for all } 1 \leq i \leq k.$$

Since each λ_i^{-1} is an eigenvalue of T^{-1} and $\dim(E_{\lambda_i^{-1}}) = m_i$, T^{-1} must be diagonalizable by part (a) of Theorem 5.9. ■

Exercise 5.2.13

Let $A \in M_{n \times n}(F)$. Recall from Exercise 14 of Section 5.1 that A and A^t have the same characteristic polynomial and hence share the same eigenvalues with the same multiplicities. For any eigenvalue λ of A and A^t , let E_λ and E'_λ denote the corresponding eigenspaces for A and A^t , respectively.

- (a) Show by way of example that for a given common eigenvalue, these two eigenspaces need not be the same.

Solution. ■

- (b) Prove that for any eigenvalue λ , $\dim(E_\lambda) = \dim(E'_\lambda)$.

Proof. Let $A \in M_{n \times n}(F)$ and denote $E_\lambda = N(L_A - \lambda I)$. Let $x \in E_\lambda$. By Theorem 5.9, $x \neq 0$ is an eigenvector of L_A corresponding to the eigenvalue λ . So, λ must be an eigenvalue of A by definition. By Theorem 5.2, we must have $\det(A - \lambda I) = 0$. By Exercise 5.1.14, we have

$$\det(A - \lambda I) = \det(A^t - \lambda I) = 0.$$

Thus, λ is an eigenvalue of A^t by Theorem 5.2 and so λ is corresponding to the eigenvector x of L_{A^t} . By Theorem 5.4 again, we can see that $x \in E'_\lambda$. Thus, $E_\lambda \subseteq E'_\lambda$. We can reverse this argument to show the other containment. Thus, $E'_\lambda \subseteq E_\lambda$ and so we conclude that $E_\lambda = E'_\lambda$. ■

- (c) Prove that if A is diagonalizable, then A^t is diagonalizable.

Proof. Apply the same argument in part (b) of Exercise 5.2.12 and part (a) of this exercise to show that A^t is diagonalizable. ■

Definition 5.2.1 (Simultaneously Diagonalizable). Two linear operators T and U on a finite-dimensional vector space V are called **simultaneously diagonalizable** if there exists an ordered basis β for V such that both $[T]_\beta$ and $[U]_\beta$ are diagonal matrices. Similarly, $A, B \in M_{n \times n}(F)$ are called **simultaneously diagonalizable** if there exists an invertible matrix $Q \in M_{n \times n}(F)$ such that both $Q^{-1}AQ$ and $Q^{-1}BQ$ are diagonal matrices.

Exercise 5.2.17

- (a) Prove that if T and U are simultaneously diagonalizable linear operators on a finite-dimensional vector space V , then the matrices $[T]_\beta$ and $[U]_\beta$ are simultaneously diagonalizable for any ordered basis β .

Proof. Let T, U be simultaneously diagonalizable linear operators. Since V is finite-dimensional let β' be another ordered basis for V (that is distinct from β). Using Theorem 2.23, we have

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q$$

where Q^{-1}, Q are invertible and diagonal where

$$Q = [I_V]_{\beta'}^\beta \text{ and } Q^{-1} = [I_V]_\beta^{\beta'}.$$

Since $[T]_\beta$ is diagonal, $[T]_{\beta'}$ is diagonal as well. We can apply this same argument to U to show that $Q^{-1}[U]_\beta Q$ for some invertible matrix Q . ■

- (b) Prove that if A and B are simultaneously diagonalizable matrices, then L_A and L_B are simultaneously diagonalizable linear operators.

Proof. Let $T = L_A$ and $U = L_B$. Apply part (a) to show that L_A and L_B are also simultaneously diagonalizable linear operators. ■

Exercise 5.2.18

- (a) Prove that if T and U are simultaneously diagonalizable operators, then T and U commute (i.e., $TU = UT$).

Proof. Let $\beta = \{v_1, v_2, \dots, v_n\}$ be a basis for a finite dimensional vector space V . First, we need to ensure that given T and U are simultaneously diagonalizable that both TU and UT are simultaneously diagonalizable as well; that is, $[TU]_\beta$ and $[UT]_\beta$ are diagonal matrices. Since T and U are simultaneous diagonalizable, there exists a invertible matrix Q such that

$$Q^{-1}[T]_\beta Q \text{ and } Q^{-1}[U]_\beta Q$$

are diagonal matrices. Since V is finite-dimensional, let β' be a basis such that $[T]_{\beta'} = Q^{-1}[T]_\beta Q$ and $[U]_{\beta'} = Q^{-1}[U]_\beta Q$. Using Theorem 2.11, we have

$$\begin{aligned} [TU]_\beta &= [T]_\beta [U]_\beta \\ &= (Q[T]_{\beta'} Q^{-1})(Q[U]_{\beta'} Q^{-1}) \\ &= Q[T]_{\beta'} [U]_{\beta'} Q^{-1} \\ &= Q[TU]_{\beta'} Q^{-1}. \end{aligned}$$

But this implies that TU is simultaneously diagonalizable. A similar argument shows UT being simultaneously diagonalizable. Hence, $[TU]_\beta$ and $[U]_\beta$ are diagonal matrices. Let $\lambda_1, \lambda_2, \dots, \lambda_n$ and $\lambda'_1, \lambda'_2, \dots, \lambda'_n$ be the eigenvalues of T and U respectively. Thus, for $1 \leq j \leq n$ we have

$$\begin{aligned} (TU)(v_j) &= \sum_{i=1}^n D_{ij} v_j = D_{jj} = (\lambda_j \lambda'_j) v_j = (\lambda'_j \lambda_j) v_j = C_{jj} \\ &= \sum_{i=1}^n C_{ij} v_j = (UT)(v_j). \end{aligned}$$

Thus, TU and UT commute. ■

- (b) Show that if A and B are simultaneously diagonalizable matrices, then A and B commute.

Proof. Let $T = L_A$ and $U = L_B$, then apply part (a). ■

Exercise 5.2.19

Let T be a diagonalizable linear operator on a finite-dimensional vector space, and let m be any positive integer. Prove that T and T^m are simultaneously diagonalizable.

Proof. Let V be a finite-dimensional vector space. Let β and β' be two ordered bases for V . We need to show that $[T]_\beta$ and $[T^m]_\beta$ are diagonal matrices. The former is shown via part (a) of Exercise 5.2.17. We will induct on m . The base case is taken care of. Let's assume that the result holds for $m - 1$ case. That is, there exists an invertible matrix Q such that $Q^{-1}[T^{m-1}]_\beta Q = [T]_{\beta'}$. Similarly, we have $Q^{-1}[T]_\beta Q = [T]_{\beta'}$. We will show that the result holds for the m th case. By the induction hypothesis and Theorem 2.11, we can write

$$\begin{aligned} [T^m]_\beta &= [T^{(m-1)+1}]_\beta = [T^{m-1}T]_\beta = [T^{m-1}]_\beta [T]_\beta \\ &= (Q[T^{m-1}]_{\beta'} Q^{-1})(Q[T]_{\beta'} Q^{-1}) = Q[T^m]_{\beta'} Q^{-1}. \end{aligned}$$

Thus, we have $Q^{-1}[T^m]_\beta Q = [T^m]_{\beta'}$. Notice how T^{m-1} and T commute via part (a) of Exercise 5.2.18. Thus, we get that both T and T^m are simultaneously diagonalizable linear operators. ■

Exercise 5.2.20

Let W_1, W_2, \dots, W_k be subspaces of a finite-dimensional vector space V such that

$$\sum_{i=1}^k W_i = V.$$

Prove that V is the direct sum of W_1, W_2, \dots, W_k if and only if

$$\dim(V) = \sum_{i=1}^k \dim(W_i).$$

Proof. For the forwards direction, suppose V is the direct sum of W_1, W_2, \dots, W_k . By definition, we have

$$V = \sum_{i=1}^n W_i$$

and

$$W_j \cap \sum_{i \neq j} W_i = \{0\}$$

for some $1 \leq j \leq k$. Using repeated applications of part (a) of Exercise 1.6.29, we get that

$$\dim(V) = \dim\left(\sum_{i=1}^n W_i\right) = \sum_{i=1}^n \dim(W_i) + \dim\left(W_j \cap \sum_{i \neq j} W_i\right) = \sum_{i=1}^n \dim(W_i).$$

Thus, we get that

$$\dim(V) = \sum_{i=1}^n \dim(W_i).$$

Conversely, each W_i is finite-dimensional, there exists a basis β_i for each W_i for $1 \leq i \leq k$. Thus, we have $\beta = \beta_1 \cup \beta_2 \cup \dots \cup \beta_k$ is a basis for V which implies that

$$\bigoplus_{i=1}^k W_i = V$$

by part (e) of Theorem 5.9. ■

Definition 5.2.2 (Partitions). We call the sets $\beta_1, \beta_2, \dots, \beta_k$ a partition of β if $\beta_1, \beta_2, \dots, \beta_k$ are subsets of β such that

$$\beta = \beta_1 \cup \beta_2 \cup \dots \cup \beta_k \text{ and } \beta_i \cap \beta_j = \emptyset \text{ if } i \neq j.$$

Exercise 5.2.21

Let V be a finite-dimensional vector space with a basis β , and let $\beta_1, \beta_2, \dots, \beta_k$ be a **partition** of β . Prove that

$$V = \text{span}(\beta_1) \oplus \text{span}(\beta_2) \oplus \dots \oplus \text{span}(\beta_k).$$

Proof. Let β be a basis for V and let $\beta_1, \beta_2, \dots, \beta_k$ be a partition of β . By definition, $\beta_1, \beta_2, \dots, \beta_k$ are subsets of β such that

$$\beta = \beta_1 \cup \beta_2 \cup \dots \cup \beta_k \text{ and } \beta_i \cap \beta_j = \emptyset \text{ if } i \neq j.$$

Since β is a basis, we must have $\text{span}(\beta) = V$. By repeated applications of Exercise 1.4.14, we must have

$$V = \text{span}(\beta) = \text{span}\left(\bigcup_{i=1}^k \beta_i\right) = \sum_{i=1}^k \text{span}(\beta_i).$$

Thus, we establish that

$$V = \sum_{i=1}^k \text{span}(\beta_i).$$

Now suppose for sake of contradiction that there exists a $1 \leq j \leq k$ such that

$$v \in \text{span}(\beta_j) \cap \sum_{i \neq j} \text{span}(\beta_i).$$

But this implies that $\beta_j \cap \beta_i \neq \emptyset$ which is a contradiction. Thus, we must have that

$$\text{span}(\beta_j) \cap \sum_{i=1}^k \text{span}(\beta_i) = \{0\}.$$

Hence, we must have

$$V = \bigoplus_{i=1}^k \text{span}(\beta_i).$$

Exercise 5.2.22

Let T be a linear operator on a finite-dimensional vector space V , and suppose that the distinct eigenvalues of T are $\lambda_1, \lambda_2, \dots, \lambda_k$. Prove that

$$\text{span}(\{x \in V : x \text{ is an eigenvector of } T\}) = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_k}.$$

Proof. Let $S = \{x \in V : x \text{ is an eigenvector of } T\}$. Our goal is to show that

$$\text{span}(S) = \sum_{i=1}^k E_{\lambda_i} \quad \text{and} \quad E_{\lambda_j} \cap \sum_{i \neq j} E_{\lambda_i} = \{0\}.$$

Since V is finite-dimensional and each E_{λ_i} is a subspace of V , we have that each E_{λ_i} is finite-dimensional by Theorem 1.11. Thus, there exists a basis of eigenvectors γ_i of T for each E_{λ_i} . Using repeated applications of Exercise 14 of section 1.4, we must have

$$\sum_{i=1}^k E_{\lambda_i} = \sum_{i=1}^k \text{span}(\gamma_i) = \text{span}\left(\bigcup_{i=1}^k \gamma_i\right).$$

Observe that each γ_i is a linearly independent set containing eigenvectors of T . Thus, we can see that

$$\bigcup_{i=1}^k \gamma_i = S$$

and so, we must have

$$\sum_{i=1}^k E_{\lambda_i} = \text{span}(S).$$

Since each eigenspace E_{λ_i} is distinct, we have for all $i \neq j$ that

$$E_{\lambda_j} \cap \sum_{i \neq j} E_{\lambda_i} = \{0\}.$$

Thus, we conclude that

$$\text{span}(S) = \bigoplus_{i=1}^k E_{\lambda_i}.$$

■

Chapter 6

Inner Product Spaces

6.1 Inner Products and Norms

Exercise 6.1.6

Complete the proof of Theorem 6.1.

Proof. (b) Let $x, y \in V$. Using the linearity of the first component, we must have

$$\begin{aligned}\langle x, cy \rangle &= \overline{\langle cy, x \rangle} \\ &= \overline{c \langle y, x \rangle} \\ &= \bar{c} \overline{\langle y, x \rangle} \\ &= \bar{c} \langle x, y \rangle.\end{aligned}$$

(c) Let $x \in V$. Note that for any $v \in V$, we have $0 \cdot v = 0$. So, we have

$$\langle x, 0 \rangle = \langle x, 0 \cdot v \rangle = \bar{0} \langle x, v \rangle = 0 \langle x, v \rangle = 0$$

Likewise, we have

$$\langle 0, x \rangle = \langle 0 \cdot v, x \rangle = 0 \langle v, x \rangle = 0.$$

Thus, we have $\langle x, 0 \rangle = \langle 0, x \rangle = 0$.

(d) Suppose $\langle x, x \rangle = 0$. By part (c), we can see that

$$\langle x, x \rangle = \langle 0, x \rangle = \langle x, 0 \rangle = 0$$

which is true if and only if $x = 0$. The converse is trivial.

(e) Suppose $\langle x, y \rangle = \langle x, z \rangle$ for all $x \in V$. Then observe that

$$\begin{aligned}\langle x, y \rangle = \langle x, z \rangle &\Rightarrow \langle x, y \rangle - \langle x, z \rangle = 0 \\ &\Rightarrow \langle 0, y - z \rangle = 0.\end{aligned}$$

By part (c), we know that the above is true if and only if $y - z = 0$. So, we have $y = z$. ■

Exercise 6.1.7

Complete the proof of Theorem 6.2.

Proof. (a) Let $x \in V$ and $c \in F$. Observe that

$$\begin{aligned}\|cx\| &= \sqrt{\langle cx, cx \rangle} = \sqrt{c\langle x, cx \rangle} \\ &= \sqrt{c\bar{c}\langle x, x \rangle} \\ &= \sqrt{|c|^2 \cdot \langle x, x \rangle} \\ &= |c| \cdot \sqrt{\langle x, x \rangle} \\ &= |c| \cdot \|x\|.\end{aligned}$$

Thus, we have that $\|cx\| = |c| \cdot \|x\|$.

(b) Let $x \in V$. Suppose $\|x\| = 0$. Then by definition of norm, we have

$$\|x\| = \sqrt{\langle x, x \rangle} = 0.$$

Squaring both sides, we can see that

$$\|x\|^2 = \langle x, x \rangle = 0.$$

This is true if and only if $x = 0$ by part (d) of Theorem 6.1. Conversely, suppose $x = 0$. So by definition of norm and part (d) of Theorem 6.1, we have

$$\sqrt{\langle x, x \rangle} = 0 \Leftrightarrow \|x\| = 0.$$

■

Exercise 6.1.9

Let β be a basis for a finite-dimensional inner product space.

(a) Prove that if $\langle x, z \rangle = 0$ for all $z \in \beta$, then $x = 0$.

Proof. Let $z \in \beta$ where $z \neq 0$. By part (c) of the definition of inner product, $\langle x, z \rangle = 0$ if and only if $x = 0$. ■

(b) Prove that if $\langle x, z \rangle = \langle y, z \rangle$ for all $z \in \beta$, then $x = y$.

Proof. Let $z \in \beta$. Observe that

$$\langle x, z \rangle = \langle y, z \rangle \Leftrightarrow \langle x - y, z \rangle = 0.$$

By part (a), we find that $x - y = 0$ if and only if $x = y$. ■

Exercise 6.1.10

Let V be an inner product space, and suppose that x and y are orthogonal vectors in V . Prove that $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. Deduce the Pythagorean Theorem in \mathbb{R}^2 .

Proof. Let $x, y \in V$ be orthogonal. Thus, $\langle x, y \rangle = 0$ implies that

$$\begin{aligned}\|x + y\|^2 &= \|x\|^2 + \langle y, x \rangle + \langle x, y \rangle + \|y\|^2 \\ &= \|x\|^2 + \overline{\langle x, y \rangle} + \langle x, y \rangle + \|y\|^2 \\ &= \|x\|^2 + \|y\|^2\end{aligned}$$

which is our desired result. Let $x' = (x, 0)$ and $y' = (0, y)$ where $x', y' \in \mathbb{R}^2$. Denote $\|x' + y'\| = z$. Using the Euclidean definition of length along with our result, we find that

$$\begin{aligned} z^2 &= \|x' + y'\|^2 = \|x'\|^2 + \|y'\|^2 \\ &= |x|^2 + |y|^2 \\ &= x^2 + y^2. \end{aligned}$$

Thus, we have obtained the Pythagorean Theorem in \mathbb{R}^2 . ■

Exercise 6.1.11

Prove the *parallelogram law* on an inner product space V ; that is, show that

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \text{for all } x, y \in V.$$

Proof. Let $x, y \in V$. Using the definition of norm and properties of inner product, we get that

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= \langle x + y, x \rangle + \langle x + y, y \rangle + \langle x - y, x \rangle + \langle x - y, -y \rangle \\ &= \langle x, x \rangle + \langle y, x \rangle + \langle x, y \rangle + \langle y, y \rangle + \langle x, x \rangle - \langle y, x \rangle \\ &\quad + -\langle x, y \rangle + \langle y, y \rangle \\ &= 2\langle x, x \rangle + 2\langle y, y \rangle \\ &= 2\|x\|^2 + 2\|y\|^2. \end{aligned}$$

Thus, we obtain

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2. \quad \text{■}$$

Exercise 6.1.12

Let $\{v_1, v_2, \dots, v_k\}$ be an orthogonal set in V , and let a_1, a_2, \dots, a_k be scalars. Prove that

$$\left\| \sum_{i=1}^k a_i v_i \right\|^2 = \sum_{i=1}^k |a_i|^2 \|v_i\|^2.$$

Proof. Let $\{v_1, v_2, \dots, v_k\}$ be an orthogonal set in V and let a_1, a_2, \dots, a_k be scalars. By the linearity in the first component and conjugate linearity in the second component, we obtain

$$\begin{aligned} \left\| \sum_{i=1}^k a_i v_i \right\|^2 &= \left\langle \sum_{i=1}^k a_i v_i, \sum_{j=1}^k a_j v_j \right\rangle \\ &= \sum_{i=1}^k a_i \left\langle v_i, \sum_{j=1}^k a_j v_j \right\rangle \\ &= \sum_{i=1}^k a_i \left(\sum_{j=1}^k \overline{a_j} \langle v_i, v_j \rangle \right). \end{aligned}$$

Since $\{v_1, v_2, \dots, v_k\}$ is orthogonal in V , we have that $\langle v_i, v_j \rangle = \delta_{ij}$ where $\delta_{ij} = 1$ or $\delta_{ij} = 0$ if $i = j$ or $i \neq j$, respectively. Thus, we have

$$\left\| \sum_{i=1}^k a_i v_i \right\|^2 = \sum_{i=1}^k a_i \overline{a_i} \langle v_i, v_i \rangle = \sum_{i=1}^k |a_i|^2 \|v_i\|^2$$

which is our desired result. ■

Exercise 6.1.13

Suppose that $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ is another inner product on V . Show that $\langle \cdot, \cdot \rangle$ is an inner product on V .

Proof. Suppose that $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ is another inner product on V . Let $x, y, z \in V$ and $c \in F$.

(a) By linearity of the first component, we obtain

$$\begin{aligned} \langle x + z, y \rangle &= \langle x + z, y \rangle_1 + \langle x + z, y \rangle_2 \\ &= \langle x, y \rangle_1 + \langle z, y \rangle_1 + \langle x, y \rangle_2 + \langle z, y \rangle_2 \\ &= (\langle x, y \rangle_1 + \langle x, y \rangle_2) + (\langle z, y \rangle_1 + \langle z, y \rangle_2) \\ &= \langle x, y \rangle + \langle z, y \rangle. \end{aligned}$$

(b) Observe that

$$\begin{aligned} \langle cx, y \rangle &= \langle cx, y \rangle_1 + \langle cx, y \rangle_2 \\ &= c\langle x, y \rangle_1 + c\langle x, y \rangle_2 \\ &= c(\langle x, y \rangle_1 + \langle x, y \rangle_2) \\ &= c\langle x, y \rangle. \end{aligned}$$

(c)

$$\begin{aligned} \overline{\langle x, y \rangle} &= \overline{\langle x, y \rangle_1 + \langle x, y \rangle_2} \\ &= \overline{\langle x, y \rangle_1} + \overline{\langle x, y \rangle_2} \\ &= \langle y, x \rangle_1 + \langle y, x \rangle_2 \\ &= \langle y, x \rangle. \end{aligned}$$

(d) For any $x \neq 0$, we have

$$\langle x, x \rangle = \langle x, x \rangle_1 + \langle x, x \rangle_2.$$

Note that $\langle x, x \rangle_1 > 0$ and $\langle x, x \rangle_2 > 0$. Thus, we must also have $\langle x, x \rangle > 0$.

We conclude that $\langle \cdot, \cdot \rangle$ is a inner product on V . ■

Exercise 6.1.14

Let A and B be $n \times n$ matrices, and let c be a scalar. Prove that $(A + cB)^* = A^* + \bar{c}B^*$.

Proof. Let A and B be $n \times n$ matrices and let $c \in F$. By using the definition of adjoint,

we get that for all i, j we get

$$\begin{aligned}(A + cB)_{ij}^* &= \overline{(A + cB)_{ji}} \\ &= \overline{A_{ji} + cB_{ji}} \\ &= \overline{A_{ji}} + \overline{cB_{ji}} \\ &= (A^*)_{ij} + \bar{c}(B^*)_{ij}.\end{aligned}$$

Thus, we conclude that $(A + cB)^* = A^* + \bar{c}B^*$. ■

Exercise 6.1.15

- (a) Prove that if V is an inner product space, then $|\langle x, y \rangle| = \|x\| \cdot \|y\|$ if and only if one of the vectors x or y is a multiple of the other. *Hint:* If the identity holds and $y \neq 0$, let

$$a = \frac{\langle x, y \rangle}{\|y\|^2},$$

and let $z = x - ay$. Prove that y and z are orthogonal and

$$|a| = \frac{\|x\|}{\|y\|}.$$

Then apply Exercise 10 to $\|x\|^2 = \|ay + z\|^2$ to obtain $\|z\| = 0$.

Proof. Suppose V is an inner product space. For the forwards direction, assume $|\langle x, y \rangle| = \|x\| \cdot \|y\|$. We will show that either x or y is a multiple of the other. Observe that $y \neq 0$ implies that

$$|\langle x, y \rangle| = \|x\| \cdot \|y\| \Leftrightarrow \frac{|\langle x, y \rangle|}{\|y\|^2} \cdot \|y\| = \|x\|.$$

We can see a similar process proves the other case

$$|\langle x, y \rangle| = \|x\| \cdot \|y\| \Leftrightarrow \frac{|\langle x, y \rangle|}{\|x\|^2} \cdot \|x\| = \|y\|.$$

For the backwards direction, suppose either x or y is a multiple of the other. Thus, for some $c \neq 0$, we have

$$\begin{aligned}|\langle x, y \rangle| &= \sqrt{\langle x, y \rangle \cdot \langle x, y \rangle} \\ &= \sqrt{\langle cy, y \rangle \cdot \langle x, \frac{1}{c}x \rangle} \\ &= \sqrt{\bar{c} \cdot \frac{1}{c} \cdot \langle y, y \rangle \langle x, x \rangle} \\ &= \sqrt{\langle y, y \rangle \cdot \langle x, x \rangle} \\ &= \sqrt{\langle y, y \rangle} \cdot \sqrt{\langle x, x \rangle} \\ &= \|y\| \cdot \|x\| = \|x\| \cdot \|y\|.\end{aligned}$$

Note that the other case is similar. Thus, we have

$$|\langle x, y \rangle| = \|x\| \cdot \|y\|.$$

Now, let $y \neq 0$ and $z = x - ay$. Then set

$$a = \frac{\langle x, y \rangle}{\|y\|^2}.$$

We will show that y and z are orthogonal and that

$$|a| = \frac{\|x\|}{\|y\|}.$$

It suffices to show that $\langle y, z \rangle = 0$. Observe

$$\begin{aligned} \langle y, z \rangle &= \langle y, x - ay \rangle \\ &= \langle y, x \rangle - \bar{a} \langle y, y \rangle \\ &= \overline{\langle x, y \rangle} - \bar{a} \|y\|^2. \end{aligned}$$

Note that we have

$$a = \frac{\langle x, y \rangle}{\|y\|^2} \Leftrightarrow \langle x, y \rangle = a \|y\|^2$$

and

$$\overline{\|y\|^2} = \overline{\langle y, y \rangle} = \langle y, y \rangle.$$

So, we must have that

$$\langle y, z \rangle = \overline{\langle x, y \rangle} - \bar{a} \|y\|^2 = \bar{a} \|y\|^2 - \bar{a} \|y\|^2 = 0.$$

Thus, y and z are orthogonal. Using

$$a = \frac{\langle x, y \rangle}{\|y\|^2},$$

we see that

$$\begin{aligned} |\langle x, y \rangle| &= \|x\| \cdot \|y\| \Leftrightarrow \langle x, y \rangle \cdot \overline{\langle x, y \rangle} = \|x\|^2 \cdot \|y\|^2 \\ &\Leftrightarrow |a|^2 \|y\|^4 = \|x\|^2 \cdot \|y\|^2 \\ &\Leftrightarrow |a|^2 = \frac{\|x\|^2}{\|y\|^2} \\ &\Leftrightarrow |a| = \frac{\|x\|}{\|y\|}. \end{aligned}$$

Finally, we show that $\|z\| = 0$ using Exercise 10. Since $\|x\|^2 = \|ay + z\|^2$, we see that

$$\|x\|^2 = \|ay + z\|^2 = a^2 \|y\|^2 + \|z\|^2.$$

Since $|a| = \|x\|/\|y\|$, we can re-write the statement above into

$$a^2 \|y\|^2 = a^2 \|y\|^2 + \|z\|^2 \Leftrightarrow \|z\|^2 = 0.$$

Thus, we get $\|z\| = 0$ which is our desired result. ■

- (b) Derive a similar result for the equality $\|x + y\| = \|x\| + \|y\|$, and generalize it to the case of n vectors.

Proof. ■

Exercise 6.1.16

- (a) Show that the vector space H with $\langle \cdot, \cdot \rangle$ defined on page 332 is an inner product space.

Proof. Let $f, g, h \in H$. Then

(a)

$$\begin{aligned}
 \langle f + h, g \rangle &= \frac{1}{2\pi} \int_0^{2\pi} (f + h)(t) \overline{g(t)} \, dt \\
 &= \frac{1}{2\pi} \int_0^{2\pi} (f(t) + h(t)) \overline{g(t)} \, dt \\
 &= \frac{1}{2\pi} \int_0^{2\pi} (f(t) \overline{g(t)} + h(t) \overline{g(t)}) \, dt \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \overline{g(t)} \, dt + \frac{1}{2\pi} \int_0^{2\pi} h(t) \overline{g(t)} \, dt \\
 &= \langle f, g \rangle + \langle h, g \rangle.
 \end{aligned}$$

(b) For $c \in F$, we have

$$\begin{aligned}
 \langle cf, g \rangle &= \frac{1}{2\pi} \int_0^{2\pi} (cf)(t) \overline{g(t)} \, dt \\
 &= \frac{1}{2\pi} \int_0^{2\pi} cf(t) \overline{g(t)} \, dt \\
 &= c \cdot \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)} \, dt \\
 &= c \cdot \langle f, g \rangle.
 \end{aligned}$$

(c)

$$\begin{aligned}
 \overline{\langle f, g \rangle} &= \overline{\frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)} \, dt} \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t) \overline{g(t)}} \, dt \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t) \, dt & (\overline{\overline{g(t)}} = g(t)) \\
 &= \frac{1}{2\pi} \int_0^{2\pi} g(t) \overline{f(t)} \, dt \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t) \, dt \\
 &= \langle g, f \rangle.
 \end{aligned}$$

(d) For $f > 0$, we know that $\int_0^{2\pi} f \, dt > 0$ by our integration properties. Thus, we

have

$$\langle f, f \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{f(t)} dt > 0.$$

We conclude that $\langle \cdot, \cdot \rangle$ satisfies all properties of an inner product. ■

(b) Let $V = C([0, 1])$, and define

$$\langle f, g \rangle = \int_0^{1/2} f(t)g(t) dt.$$

Is this an inner product on V ?

Proof. This is not an inner product because it fails property (c) of the definition of inner product. ■

Exercise 6.1.17

Let T be a linear operator on an inner product space V , and suppose that $\|T(x)\| = \|x\|$ for all x . Prove that T is injective.

Proof. Let $x, y \in V$ and suppose $T(x) = T(y)$. Taking advantage of the linearity of T , we can write

$$T(x) = T(y) \Leftrightarrow T(x - y) = 0.$$

Using our assumption, we see that

$$0 = \|T(x - y)\| = \|x - y\|.$$

By using the definition of norm, we can see further that

$$\|x - y\|^2 = 0 \Leftrightarrow \langle x - y, x - y \rangle = 0$$

Since V is an inner product space, we can see that the above is true if and only if $x - y = 0$ and thus $x = y$. Therefore, we conclude that T is injective. ■

Exercise 6.1.18

Let V be a vector space over F , where $F = \mathbb{R}$ or $F = \mathbb{C}$, and let W be an inner product space over F with inner product $\langle \cdot, \cdot \rangle$. If $T : V \rightarrow W$ is linear, prove that $\langle x, y \rangle' = \langle T(x), T(y) \rangle$ defines an inner product on V if and only if T is one-to-one.

Proof. Let $x, y \in V$. For the forwards direction, suppose that $\langle x, y \rangle' = \langle T(x), T(y) \rangle$ defines an inner product on V . Suppose that $T(x) = T(y)$. Then we see that $T(x - y) = 0$ since T is linear. Consider the norm $\|T(x - y)\| = 0$. By definition of $\langle \cdot, \cdot \rangle'$, observe that

$$\begin{aligned} 0 &= \|T(x - y)\|^2 = \langle T(x - y), T(x - y) \rangle \\ &= \langle x - y, x - y \rangle'. \end{aligned}$$

Since $\langle \cdot, \cdot \rangle'$ is an inner product, we must have $x - y = 0$ implies $x = y$. For the backwards direction, suppose T is one-to-one. We need to show that $\langle \cdot, \cdot \rangle'$ defines an inner product on V . Let $x, y, z \in V$ and let $c \in F$. Suppose T is linear and $\langle \cdot, \cdot \rangle$ defines an inner product over W . Then

(a)

$$\begin{aligned}\langle x + y, z \rangle' &= \langle T(x + y), T(z) \rangle \\ &= \langle T(x) + T(y), T(z) \rangle \\ &= \langle T(x), T(z) \rangle + \langle T(y), T(z) \rangle \\ &= \langle x, y \rangle' + \langle z, y \rangle' .\end{aligned}$$

(b)

$$\begin{aligned}\langle cx, y \rangle' &= \langle T(cx), T(y) \rangle \\ &= \langle cT(x), T(y) \rangle \\ &= c\langle T(x), T(y) \rangle \\ &= c\langle x, y \rangle' .\end{aligned}$$

(c)

$$\begin{aligned}\overline{\langle x, y \rangle'} &= \overline{\langle T(x), T(y) \rangle} \\ &= \langle T(y), T(x) \rangle \\ &= \langle y, x \rangle' .\end{aligned}$$

(d) Suppose $x \neq 0$. Since $\langle \cdot, \cdot \rangle$ is an inner product on W , we have

$$\langle x, x \rangle' = \langle T(x), T(x) \rangle > 0.$$

Thus, we have that $\langle \cdot, \cdot \rangle'$ is an inner product on V . ■

Exercise 6.1.19

Let V be an inner product space. Prove that

- (a) $\|x \pm y\|^2 = \|x\|^2 \pm 2\Re\langle x, y \rangle + \|y\|^2$ for all $x, y \in V$, where $\Re\langle x, y \rangle$ denotes the real part of the complex number $\langle x, y \rangle$.

Proof. Let $x, y \in V$ where V is an inner product space. Note that

$$\|x + y\|^2 = \|x\|^2 + 2\Re\langle x, y \rangle + \|y\|^2$$

by the proof for part (d) of Theorem 6.2. We will show the other case. By the properties found in Theorem 6.1, we have

$$\begin{aligned}\|x - y\|^2 &= \langle x - y, x - y \rangle \\ &= \langle x, x \rangle - \langle x, y \rangle - (\langle y, x \rangle) \\ &= \langle x, x \rangle - (\langle x, y \rangle + \overline{\langle x, y \rangle}) + \langle y, y \rangle \\ &= \langle x, x \rangle - 2\Re\langle x, y \rangle + \|y\|^2 .\end{aligned}$$
■

- (b) $|||x| - |y||| \leq \|x - y\|$ for all $x, y \in V$.

Proof. Let $x, y \in V$. Using part (a) and part (c) of Theorem 6.2, we can see that

$$\begin{aligned}\|x - y\|^2 &= \|x\|^2 - 2\Re\langle x, y \rangle + \|y\|^2 \\ &\geq \|x\|^2 - 2|\langle x, y \rangle| + \|y\|^2 \\ &\geq \|x\|^2 - 2\|x\| \cdot \|y\| + \|y\|^2 & (-2|\langle x, y \rangle| \geq -2\|x\| \cdot \|y\|) \\ &= (\|x\| - \|y\|)^2.\end{aligned}$$

Taking the square root of both sides of our inequality, we obtain

$$\|x - y\| \geq |\|x\| - \|y\||.$$

Note that

$$\sqrt{(\|x\| - \|y\|)^2} = |\|x\| - \|y\||.$$

■

Exercise 6.1.20

Let V be an inner product space over F . Prove the *polar identities*: For all $x, y \in V$,

- (a) $\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2$ if $F = \mathbb{R}$.

Proof. Let $x, y \in V$ and let $F = \mathbb{R}$. Using part (a) of Exercise 19, we can write

$$\begin{aligned}\frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2 &= \frac{1}{4}(\|x\|^2 + 2\langle x, y \rangle + \|y\|^2) \\ &\quad - \frac{1}{4}(\|x\|^2 - 2\langle x, y \rangle + \|y\|^2) \\ &= \frac{1}{4} \cdot 4\langle x, y \rangle = \langle x, y \rangle.\end{aligned}$$

Thus, we have that

$$\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2.$$

■

- (b) $\langle x, y \rangle = \frac{1}{4} \sum_{k=1}^4 i^k \|x + i^k y\|^2$ if $F = \mathbb{C}$, where $i^2 = -1$.

Proof. Let $x, y \in V$ and $F = \mathbb{C}$. Using part (a), we can write

$$\begin{aligned}\frac{1}{4} \sum_{k=1}^4 i^k \|x + i^k y\|^2 &= \frac{1}{4} \left[i\|x + iy\|^2 + i^2\|x + i^2 y\|^2 \right. \\ &\quad \left. + i^3\|x + i^3 y\|^2 + i^4\|x + i^4 y\|^2 \right] \\ &= \frac{1}{4} \left[i\|x + iy\|^2 - \|x - y\|^2 - i\|x - iy\|^2 + \|x + y\|^2 \right] \\ &= \frac{1}{4} \left[\|x + y\|^2 - \|x - y\|^2 \right] + \frac{i}{4} \left[\|x + iy\|^2 - \|x - iy\|^2 \right]\end{aligned}$$

where

$$\Re\langle x, y \rangle = \frac{1}{4} \left[\|x + y\|^2 - \|x - y\|^2 \right] \quad (\text{part(a)})$$

and

$$\Im\langle x, y \rangle = \frac{1}{4} \left[\|x + iy\|^2 - \|x - iy\|^2 \right].$$

Thus, we have that

$$\langle x, y \rangle = \frac{1}{4} \sum_{k=1}^4 i^k \|x + i^k y\|^2.$$

■

Exercise 6.1.21

Let A be an $n \times n$ matrix. Define

$$A_1 = \frac{1}{2}(A + A^*) \quad \text{and} \quad A_2 = \frac{1}{2i}(A - A^*).$$

- (a) Prove that $A_1^* = A_1$, $A_2^* = A_2$, and $A = A_1 + iA_2$. Would it be reasonable to define A_1 and A_2 to be the real and imaginary parts, respectively, of the matrix A ?

Proof. We will first show that $A_1^* = A_1$ and $A_2^* = A_2$. For all i, j , we see that

$$\begin{aligned} (A_1^*)_{ij} &= \overline{(A_1)_{ji}} = \frac{1}{2} \overline{(A + A^*)_{ji}} \\ &= \frac{1}{2} \overline{(A_{ji} + A_{ji}^*)} \\ &= \frac{1}{2} (\overline{A_{ji}} + \overline{A_{ji}^*}) \\ &= \frac{1}{2} ((A^*)_{ij} + A_{ij}) & (\overline{\overline{A_{ij}}} = A_{ij}) \\ &= \frac{1}{2} (A^* + A)_{ij} \\ &= (A_1)_{ij}. \end{aligned}$$

Thus, $A_1^* = A_1$. Similarly, we have

$$\begin{aligned} (A_2^*)_{ij} &= \overline{(A_2)_{ji}} = \frac{1}{2i} \overline{(A - A^*)_{ji}} \\ &= -\frac{1}{2i} \overline{(A_{ji} - A_{ji}^*)} \\ &= -\frac{1}{2i} (\overline{A_{ji}} - \overline{A_{ji}^*}) \\ &= -\frac{1}{2i} (A_{ij}^* - A_{ij}) & (\overline{\overline{A_{ij}}} = A_{ij}) \\ &= \frac{1}{2i} (A_{ij} - A_{ij}^*) \\ &= \frac{1}{2i} (A - A^*)_{ij} = (A_2)_{ij}. \end{aligned}$$

Thus, we have that $A_2^* = A_2$.

Now, we will show that $A = A_1 + iA_2$. Observe that

$$\begin{aligned} A_1 + iA_2 &= \frac{1}{2}(A + A^*) + i \cdot \frac{1}{2i}(A - A^*) \\ &= \frac{1}{2}A + \frac{1}{2}A^* + \frac{1}{2}A - \frac{1}{2}A^* \\ &= A. \end{aligned}$$

Thus, we have $A = A_1 + iA_2$. This is a reasonable definition since we can just define the corresponds parts of the complex entries of A as entries in \mathbb{R} corresponding to two

different matrices with each dedicated to the real and imaginary parts, respectively. ■

- (b) Let A be an $n \times n$ matrix. Prove that the representation in (a) is unique. That is, prove that if $A = B_1 + iB_2$, where $B_1^* = B_1$ and $B_2^* = B_2$, then $B_1 = A_1$ and $B_2 = A_2$.

Proof. Let A be an $n \times n$ matrix. Suppose there exists another representation of A such that $A = B_1 + iB_2$ where $B_1 = B_1^*$ and $B_2 = B_2^*$. Equating this representation to that of the one found in part (a), we obtain that

$$A_1 + iA_2 = B_1 + iB_2 \Leftrightarrow A_1 = A_2, B_1 = B_2.$$

Thus, the representation in (a) is unique. ■

Exercise 6.1.22

Let V be a real or complex vector space (possibly infinite-dimensional), and let β be a basis for V . For $x, y \in V$ there exists $v_1, v_2, \dots, v_n \in \beta$ such that

$$x = \sum_{i=1}^n a_i v_i \text{ and } y = \sum_{i=1}^n b_i v_i.$$

Define

$$\langle x, y \rangle = \sum_{i=1}^n a_i \bar{b}_i.$$

- (a) Prove that $\langle \cdot, \cdot \rangle$ is an inner product on V and that β is an orthonormal basis for V . Thus every real or complex vector space may be regarded as an inner product space.

Proof. Let β be an orthonormal basis for V . We will show that $\langle \cdot, \cdot \rangle$ is an inner product on V . Let $x, y, z \in V$ and $c \in F$.

- (a) We can find scalars $a_i, b_i, c_i \in F$ and $v_i \in \beta$ for $1 \leq i \leq n$ such that

$$x = \sum_{i=1}^n a_i v_i, \quad y = \sum_{i=1}^n b_i v_i, \quad z = \sum_{i=1}^n c_i v_i.$$

Using our definition of $\langle \cdot, \cdot \rangle$, we obtain that

$$\begin{aligned} \langle x + y, z \rangle &= \sum_{i=1}^n (a_i + b_i) \bar{c}_i = \sum_{i=1}^n (a_i \bar{c}_i + b_i \bar{c}_i) \\ &= \sum_{i=1}^n a_i \bar{c}_i + \sum_{i=1}^n b_i \bar{c}_i \\ &= \langle x, z \rangle + \langle y, z \rangle. \end{aligned}$$

Thus, part (a) is satisfied.

- (b) Using the representation of $x, y \in V$, we obtain

$$\langle cx, y \rangle = \sum_{i=1}^n ca_i \bar{b}_i = c \sum_{i=1}^n a_i \bar{b}_i = c \langle x, y \rangle.$$

Thus, part (b) is satisfied.

(c)

$$\begin{aligned}\overline{\langle x, y \rangle} &= \overline{\sum_{i=1}^n a_i \overline{b_i}} = \sum_{i=1}^n \overline{a_i \overline{b_i}} = \sum_{i=1}^n \overline{a_i} b_i & (\overline{\overline{b_i}} = b_i) \\ &= \sum_{i=1}^n b_i \overline{a_i} \\ &= \langle y, x \rangle.\end{aligned}$$

Thus, part (d) is satisfied.

(d) Suppose $x \neq 0$. Then by definition of $\langle \cdot, \cdot \rangle$, we must have

$$\langle x, x \rangle = \sum_{i=1}^n a_i \overline{a_i} = \sum_{i=1}^n |a_i|^2 > 0.$$

Thus, part (d) is satisfied. ■

(b) Prove that if $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$ and β is the standard ordered basis, then the inner product defined above is the standard inner product.

Proof. Let $F = \mathbb{R}^n$ or $F = \mathbb{C}^n$. Let $x, y, z \in F^n$ and $c \in F$. Repeat the same process above but with fixing $x = (a_1, a_2, \dots, a_n)$, $y = (b_1, b_2, \dots, b_n)$, and $z = (c_1, c_2, \dots, c_n)$. Thus, the inner product above is the standard inner product. ■

Exercise 6.1.23

Let $V = F^n$, and let $A \in M_{n \times n}(F)$.

(a) Prove that $\langle x, Ay \rangle = \langle A^*x, y \rangle$ for all $x, y \in V$.

Proof. Let $x, y \in V$. By definition of the adjoint for matrices, we obtain, for all i, j , through the linearity of the first component and the conjugate linearity of the second component that

$$\begin{aligned}\langle x, A_{ij}y \rangle &= \overline{A_{ij}} \langle x, y \rangle = (A^*)_{ji} \langle x, y \rangle \\ &= \langle (A^*)_{ji} x, y \rangle.\end{aligned}$$

Thus, we have that $\langle x, Ay \rangle = \langle A^*x, y \rangle$ for all $x, y \in V$. ■

(b) Suppose that for some $B \in M_{n \times n}(F)$, we have $\langle x, Ay \rangle = \langle Bx, y \rangle$ for all $x, y \in V$. Prove that $B = A^*$.

Proof. Let $x, y \in V$. By part (a), we have

$$\langle x, Ay \rangle = \langle A^*x, y \rangle.$$

So, by linearity of the first component, we find that

$$\langle Bx, y \rangle = \langle A^*x, y \rangle \Leftrightarrow B \langle x, y \rangle = A^* \langle x, y \rangle.$$

by assumption. If $\langle x, y \rangle \neq 0$, then $B = A^*$. Otherwise, B is just the zero matrix.

Thus, we must have $B = A^*$. ■

- (c) Let α be the standard ordered basis for V . For any orthonormal basis β for V , let Q be the $n \times n$ matrix whose columns are the vectors in β . Prove that $Q^* = Q^{-1}$.

Proof. Let $V = F^n$ and α as the standard ordered basis for V . It suffices to show that $QQ^* = QQ^* = I$. Let $\beta = \{v_1, v_2, \dots, v_n\}$ be an orthonormal basis for V . Since the columns of Q are just the vectors of β , we can write

$$\begin{aligned} (QQ^*)_{ij} &= \sum_{k=1}^n Q_{ik} Q_{kj}^* \\ &= \sum_{k=1}^n Q_{ik} \overline{Q_{jk}} \\ &= \langle v_i, v_j \rangle \\ &= \delta_{ij} \end{aligned}$$

Note that the above is equal to 1 whenever $i = j$ and 0 otherwise. So, we must have $QQ^* = I$. Showing $Q^*Q = I$ is done similarly as above. Thus, we conclude that $Q^* = Q^{-1}$. ■

- (d) Define linear operators T and U on V by $T(x) = Ax$ and $U(x) = A^*x$. Show that $[U]_\beta = [T]_\beta^*$ for any orthonormal basis β for V .

Proof. Observe that $T = L_A$ and $U = L_{A^*}$ since $T(x) = Ax$ and $U(x) = A^*x$. By Theorem 2.15, we can see that

$$[T]_\beta = A \Leftrightarrow [T]_\beta^* = A^*.$$

But note that $[U]_\beta = A^*$. Thus, we conclude that

$$[T]_\beta^* = [U]_\beta.$$
■

Definition 6.1.1 (Norm Spaces). Let V be a vector space over F , where F is either \mathbb{R} or \mathbb{C} . Regardless of whether V is or not an inner product space, we may still define a **norm** $\|\cdot\|$ as a real-valued function on V satisfying the following three conditions for all $x, y \in V$ and $a \in F$.

- (1) $\|x\| \geq 0$, and $\|x\| = 0$ if and only if $x = 0$.
- (2) $\|ax\| = |a|\|x\|$.
- (3) $\|x + y\| \leq \|x\| + \|y\|$.

Exercise 6.1.24

Prove that the following are norms on the given vector spaces V .

- (a) $V = M_{m \times n}(F)$; $\|A\| = \max_{i,j} |A_{ij}|$ for all $A \in V$.

Proof. Let $A \in M_{m \times n}(F)$. For (1), let $A \neq O$ where O is the zero matrix. Thus, $|A_{ij}| > 0$ and so $\|A\| = \max_{i,j} |A_{ij}| > 0$. On the other hand, $A = O$ implies that $A_{ij} = 0$ for all i, j . This is true if and only if $\|A\| = 0$. Hence, (1) is satisfied.

Let $a \in F$. By properties of $\max_{i,j}$

$$\begin{aligned} \|aA\| &= \max_{i,j} |(aA)_{ij}| = \max_{i,j} |aA_{ij}| \\ &= \max_{i,j} |a| |A_{ij}| = |a| \max_{i,j} |A_{ij}| \\ &= |a| \|A\|. \end{aligned}$$

Thus, part (2) is satisfied.

Now, let $A, B \in M_{m \times n}(F)$. Then by triangle inequality of the absolute value, we have for all i, j

$$\begin{aligned} \|A + B\| &= \max_{i,j} |(A + B)_{ij}| \\ &= \max_{i,j} |A_{ij} + B_{ij}| \\ &\leq \max_{i,j} |A_{ij}| + \max_{i,j} |B_{ij}| = \|A\| + \|B\|. \end{aligned}$$

Thus, (3) is satisfied. We conclude that $\|A\| = \max_{i,j} |A_{ij}|$ is a norm for all $A \in V$. ■

(b) $V = C([0, 1]); \quad \|f\| = \max_{t \in [0, 1]} |f(t)|$ for all $f \in V$.

Proof. Let $V = C([0, 1])$ and let $f \in V$. Suppose $f \neq 0$, then $f(t) > 0$ for all $t \in [0, 1]$. Thus, $\|f\| = \max_{t \in [0, 1]} |f(t)| > 0$. Clearly, $f = 0$ implies that for all $t \in [0, 1]$, $f(t) = 0$.

This is true if and only if $\|f\| = 0$.

Let $a \in F$. Then observe that

$$\begin{aligned} \|af\| &= \max_{t \in [0, 1]} |(af)(t)| = \max_{t \in [0, 1]} |af(t)| \\ &= \max_{t \in [0, 1]} |a| |f(t)| \\ &= |a| \max_{t \in [0, 1]} |f(t)| = |a| \|f\|. \end{aligned}$$

Thus, (2) is satisfied.

Now, let $f, g \in C([0, 1])$. By using the properties of the absolute value, we have

$$\begin{aligned} \|f + g\| &= \max_{t \in [0, 1]} |(f + g)(t)| = \max_{t \in [0, 1]} |f(t) + g(t)| \\ &\leq \max_{t \in [0, 1]} |f(t)| + \max_{t \in [0, 1]} |g(t)| \\ &= \|f\| + \|g\|. \end{aligned}$$

Thus, we have that $\|f + g\| \leq \|f\| + \|g\|$ for all $t \in [0, 1]$. Thus, (3) is satisfied. Hence, $\|f\| = \max_{t \in [0, 1]} |f(t)|$ is a norm over V . ■

(c) $V = C([0, 1]); \quad \|f\| = \int_0^1 |f(t)| dt$ for all $f \in V$.

Proof. If $f \geq 0$, then $|f| \geq 0$. Thus,

$$\|f\| = \int_0^1 |f(t)| dt > 0$$

by our integration properties. Otherwise $f = 0$ implies that

$$\|f\| = \int_0^1 |f(t)| dt = 0.$$

So, property (1) is satisfied. Conversely, $f = 0$. Since this is course on linear algebra, I will not bother with the technicalities of the argument. Thus, property (1) is satisfied.

Now, for property (2), let $f \in C([0, 1])$ and $a \in F$. Then by our properties of integration, we must have

$$\begin{aligned} \|af\| &= \int_0^1 |(af)(t)| dt = \int_0^1 |af(t)| dt \\ &= \int_0^1 |a||f(t)| dt \\ &= |a| \int_0^1 |f(t)| dt \\ &= |a|\|f\|. \end{aligned}$$

Thus, property (2) is satisfied.

Now, let $f, g \in C([0, 1])$. Then using the triangle inequality, we obtain

$$\begin{aligned} \|f + g\| &= \int_0^1 |(f + g)(t)| dt = \int_0^1 |f(t) + g(t)| dt \\ &\leq \int_0^1 |f(t)| + |g(t)| dt \\ &= \int_0^1 f(t) dt + \int_0^1 g(t) dt \\ &= \|f\| + \|g\|. \end{aligned}$$

Thus, $\|f + g\| \leq \|f\| + \|g\|$ for all $f, g \in C([0, 1])$ and hence we conclude that

$$\|f\| = \int_0^1 |f(t)| dt$$

defines a norm on $V = C([0, 1])$. ■

Exercise 6.1.25

Use Exercise 20 to show that there is no inner product $\langle \cdot, \cdot \rangle$ on \mathbb{R}^2 such that $\|x\|^2 = \langle x, x \rangle$ for all $x \in \mathbb{R}^2$ if the norm is defined as in Exercise 24(d).

Proof. There is no inner product on \mathbb{R}^2 because the first property of the inner product fails. ■

Exercise 6.1.26

Let $\|\cdot\|$ be a norm on a vector space V , and define, for each ordered pair of vectors, the scalar $d(x, y) = \|x - y\|$, called the **distance** between x and y . Prove the following results for all $x, y, z \in V$.

(a) $d(x, y) \geq 0$.

Proof. Let $x, y \in V$. Then by definition of norm and $d(x, y)$, we have

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle} \geq 0.$$

■

(b) $d(x, y) = d(y, x)$.

Proof. Let $x, y \in V$. Then we have

$$d(x, y) = \|x - y\| = \|y - x\| = d(y, x).$$

■

(c) $d(x, y) \leq d(x, z) + d(z, y)$.

Proof. Let $x, y, z \in V$. Then by the triangle inequality of the norm, we have

$$\begin{aligned} d(x, y) &= \|x - y\| = \|x - z + z - y\| \\ &\leq \|x - z\| + \|z - y\| \\ &= d(x, z) + d(z, y). \end{aligned}$$

Thus, we have

$$d(x, y) \leq d(x, z) + d(z, y).$$

■

(d) $d(x, x) = 0$.

Proof. Let $x \in V$. Then

$$d(x, x) = \|x - x\| = \|0\| = 0.$$

■

(e) $d(x, y) \neq 0$ if $x \neq y$.

Proof. Let $x, y \in V$ be distinct ($x \neq y$). Thus, $x - y \neq 0$ implies that

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle} > 0.$$

■

Exercise 6.1.27

Let $\|\cdot\|$ be a norm on a real vector space V satisfying the parallelogram law given in Exercise 11. Define

$$\langle x, y \rangle = \frac{1}{4}[\|x + y\|^2 - \|x - y\|^2].$$

Prove that $\langle \cdot, \cdot \rangle$ defines an inner product on V such that $\|x\|^2 = \langle x, x \rangle$ for all $x \in V$.

(a) Prove $\langle x, 2y \rangle = 2\langle x, y \rangle$ for all $x, y \in V$.

Proof. Let $x, y \in V$. Then

$$\langle x, 2y \rangle = \frac{1}{4} [\|x + 2y\|^2 - \|x - 2y\|^2].$$

Observe that the expression inside the brackets of the equation above can be written as, using the parallelogram law,

$$\begin{aligned} \|x + 2y\|^2 - \|x - 2y\|^2 &= \|(x + y) + y\|^2 - \|(x - y) - y\|^2 \\ &= \|(x + y) + y\|^2 + \|(x + y) - y\|^2 \\ &\quad - (\|(x + y) - y\|^2 + \|(x - y) - y\|^2) \\ &= 2\|x + y\|^2 + 2\|y\|^2 - (2\|x - y\|^2 + 2\|y\|^2) \\ &= 2(\|x + y\|^2 - \|x - y\|^2). \end{aligned}$$

This tells us that

$$\langle x, 2y \rangle = 2 \cdot \frac{1}{4} [\|x + y\|^2 - \|x - y\|^2] = 2\langle x, y \rangle$$

which is our desired result. ■

(b) Prove $\langle x + u, y \rangle = \langle x, y \rangle + \langle u, y \rangle$ for all $x, u, y \in V$.

Proof. Let $x, y, u \in V$. Observe that the parallelogram law can be written as

$$\frac{1}{2} [\|x + y\|^2 + \|x - y\|^2] = \|x\|^2 + \|y\|^2.$$

Observe the sum

$$\langle x, y \rangle + \langle u, y \rangle = \frac{1}{4} [\|x + y\|^2 - \|x - y\|^2] + \frac{1}{4} [\|u + y\|^2 - \|u - y\|^2].$$

The first term can be written further as

$$\begin{aligned} \frac{1}{4} [\|x + y\|^2 - \|x - y\|^2] &= \frac{1}{4} \left[\frac{1}{2} [\|(x + y) + u\|^2 + \|(x + y) - u\|^2] \right. \\ &\quad \left. - \frac{1}{2} [\|(x - y) + u\|^2 + \|(x - y) - u\|^2] \right] \\ &= \frac{1}{4} \left[\frac{1}{2} [\|(x + u) + y\|^2 + \|(x + y) - u\|^2] \right. \\ &\quad \left. - \frac{1}{2} [\|(x + u) - y\|^2 + \|(x - y) - u\|^2] \right]. \end{aligned}$$

Likewise, the second term can be written as

$$\begin{aligned} \frac{1}{4} [\|u + y\|^2 - \|u - y\|^2] &= \frac{1}{4} \left[\frac{1}{2} [\|(u + y) + x\|^2 + \|(u + y) - x\|^2] \right. \\ &\quad \left. - \frac{1}{2} [\|(u - y) + x\|^2 + \|(u - y) - x\|^2] \right] \\ &= \frac{1}{4} \left[\frac{1}{2} [\|(u + x) + y\|^2 + \|(x - y) - u\|^2] \right. \\ &\quad \left. - \frac{1}{2} [\|(x + u) - y\|^2 + \|(x + y) - u\|^2] \right]. \end{aligned}$$

As a whole, we have

$$\begin{aligned}\langle x, y \rangle + \langle u, y \rangle &= \frac{1}{4} [\|(x+u) + y\|^2 - \|(x+u) - y\|^2] \\ &= \langle x+u, y \rangle.\end{aligned}$$

■

- (c) Prove $\langle nx, y \rangle = n\langle x, y \rangle$ for every positive integer n and every $x, y \in V$.

Proof. Let $x, y \in V$. Let us proceed via induction to show the result. Note that the base case is trivial; that is, if $n = 1$, then we have $\langle x, y \rangle = \langle x, y \rangle$. Suppose that the result holds for all positive integers less than n . Then observe that $\langle (n-1)x, y \rangle = (n-1)\langle x, y \rangle$. Using part (b), we can write

$$\begin{aligned}\langle nx, y \rangle &= \langle ((n-1) + 1)x, y \rangle = \langle (n-1)x, y \rangle + \langle x, y \rangle \\ &= (n-1)\langle x, y \rangle + \langle x, y \rangle \\ &= ((n-1) + 1)\langle x, y \rangle \\ &= n\langle x, y \rangle.\end{aligned}$$

Thus, we see that $\langle nx, y \rangle = n\langle x, y \rangle$ which is our desired result. ■

- (d) Prove $m\langle \frac{1}{m}x, y \rangle = \langle x, y \rangle$ for every positive integer m and every $x, y \in V$.

Proof. Let $x, y \in V$ and let m be any positive integer. Then set $z = \frac{1}{m}x$. By using part (c), we can write

$$\begin{aligned}m\left\langle \frac{1}{m}x, y \right\rangle &= m\langle z, y \rangle \\ &= \langle mz, y \rangle \\ &= \left\langle m \cdot \frac{1}{m}x, y \right\rangle = \langle x, y \rangle.\end{aligned}$$

Thus, we have

$$m\left\langle \frac{1}{m}x, y \right\rangle = \langle x, y \rangle.$$

■

- (e) Prove $\langle rx, y \rangle = r\langle x, y \rangle$ for every rational number r and every $x, y \in V$.

Proof. Let $x, y \in V$ and let $r \in \mathbb{Q}$. Observe that, by definition of rational numbers, we have $r = p/q$ with $p, q \in \mathbb{Z}$ and $q \neq 0$. Then we write

$$\begin{aligned}\langle rx, y \rangle &= \left\langle \frac{p}{q} \cdot x, y \right\rangle \\ &= p\left\langle \frac{1}{q}x, y \right\rangle && \text{(Part (c))} \\ &= \frac{p}{q}\langle x, y \rangle && \text{(Part (d))} \\ &= r\langle x, y \rangle.\end{aligned}$$

Thus, we have that $\langle rx, y \rangle = r\langle x, y \rangle$. ■

- (f) Prove $|\langle x, y \rangle| \leq \|x\|\|y\|$ for every $x, y \in V$.

Proof. Let $x, y \in V$. Using the triangle inequality for norms, we must have

$$\begin{aligned} |\langle x, y \rangle| &= \frac{1}{4} [\|x + y\|^2 - \|x - y\|^2] \\ &\leq \frac{1}{4} [(\|x\| + \|y\|)^2 - \|x - y\|^2] \\ &= \frac{1}{4} [2\|x\|\|y\| - (\|x - y\|^2 - \|x\|^2 - \|y\|^2)] \\ &\leq \|x\|\|y\|. \end{aligned}$$

Thus, we have that $|\langle x, y \rangle| \leq \|x\|\|y\|$. ■

(g) Prove that for every $c \in \mathbb{R}$, every rational number r , and every $x, y \in V$,

$$|c\langle x, y \rangle - \langle cx, y \rangle| = |(c - r)\langle x, y \rangle - \langle (c - r)x, y \rangle| \leq 2|c - r|\|x\|\|y\|.$$

Proof. Let $c \in \mathbb{R}$, $r \in \mathbb{Q}$, and $x, y \in V$. Using part (e), we can write

$$\begin{aligned} |c\langle x, y \rangle - \langle cx, y \rangle| &= |c\langle x, y \rangle - r\langle x, y \rangle + r\langle x, y \rangle - \langle cx, y \rangle| \\ &= |(c - r)\langle x, y \rangle + \langle (r - c)x, y \rangle| \\ &= |(c - r)\langle x, y \rangle - \langle (c - r)x, y \rangle|. \end{aligned}$$

Working with the second equality, we can use the triangle inequality to write

$$\begin{aligned} |(c - r)\langle x, y \rangle + \langle (r - c)x, y \rangle| &\leq |(c - r)\langle x, y \rangle| + |\langle (r - c)x, y \rangle| \\ &= |c - r||\langle x, y \rangle| + |\langle (r - c)x, y \rangle| \\ &\leq |c - r|\|x\|\|y\| + \|(r - c)x\|\|y\| \\ &= |c - r|\|x\|\|y\| + |c - r|\|x\|\|y\| \\ &= 2|c - r|\|x\|\|y\|. \end{aligned}$$

Hence, we conclude that

$$|c\langle x, y \rangle - \langle cx, y \rangle| = |(c - r)\langle x, y \rangle - \langle (c - r)x, y \rangle| \leq 2|c - r|\|x\|\|y\|.$$
■

(h) Use the fact that for any $c \in \mathbb{R}$, $|c - r|$ can be made arbitrarily small, where r varies over the set of rational numbers, to establish item (b) of the definition of inner product.

Proof. Let $x, y \in V$ be non-zero. Let $c \in \mathbb{R}$ and let $\varepsilon > 0$. Suppose that

$$|c - r| < \frac{\varepsilon}{2\|x\|\|y\|}.$$

Then we have

$$|c\langle x, y \rangle - \langle cx, y \rangle| \leq 2|c - r|\|x\|\|y\| < \varepsilon$$

which is our desired result. ■

Exercise 6.1.28

Let V be a complex inner product space with an inner product $\langle \cdot, \cdot \rangle$. Let $[\cdot, \cdot]$ be the real-valued function such that $[x, y]$ is the real part of the complex number $\langle x, y \rangle$ for all $x, y \in V$. Prove that

$[\cdot, \cdot]$ is an inner product for V , where V is regarded as a vector space over \mathbb{R} . Prove, furthermore, that $[x, ix] = 0$ for all $x \in V$.

Proof. Let $x, y, z \in V$ and $c \in \mathbb{R}$.

(a) Since $\langle \cdot, \cdot \rangle$ is linear in the first component, we must have that

$$\begin{aligned} [x + y, z] &= \Re \langle x + y, z \rangle = \Re(\langle x, z \rangle + \langle y, z \rangle) \\ &= \Re \langle x, z \rangle + \Re \langle y, z \rangle \\ &= [x, z] + [y, z]. \end{aligned}$$

(b)

$$\begin{aligned} [cx, y] &= \Re \langle cx, y \rangle = \Re c \langle x, y \rangle \\ &= c \Re \langle x, y \rangle \\ &= c[x, y]. \end{aligned}$$

(c)

$$\overline{[x, y]} = \Re \overline{\langle x, y \rangle} = \Re \langle y, x \rangle = [y, x].$$

(d) Suppose that $x \neq 0$. Since $\langle x, x \rangle > 0$, we also have that

$$[x, x] = \Re \langle x, x \rangle > 0.$$

So, property (d) is satisfied. Now, let us show that $[x, ix] = 0$ for any $x \in V$. Now we can see that $[\cdot, \cdot]$ is a inner product. Since $\Re(i) = 0$, we obtain that

$$\begin{aligned} [x, ix] &= \Re \langle x, ix \rangle = \Re(\bar{i} \langle x, x \rangle) \\ &= \Re(-i) \Re \langle x, x \rangle \\ &= 0 \cdot \langle x, x \rangle = 0. \end{aligned}$$

So, $[x, ix] = 0$ for all $x \in V$. ■

Exercise 6.1.29

Let V be a vector space over \mathbb{C} , and suppose that $[\cdot, \cdot]$ is a real inner product on V , where V is regarded as a vector space over \mathbb{R} , such that $[x, ix] = 0$ for all $x \in V$. Let $\langle \cdot, \cdot \rangle$ be the complex-valued function defined by

$$\langle x, y \rangle = [x, y] + i[x, y] \quad \text{for } x, y \in V.$$

Prove that $\langle \cdot, \cdot \rangle$ is a complex inner product on V .

Proof. Let $x, y, z \in V$ and $c \in F$.

(a) Using the linearity of the first component of the inner product $[\cdot, \cdot]$, we must have

$$\begin{aligned}\langle x + y, z \rangle &= [x + y, z] + i[x + y, iz] \\ &= ([x, z] + [y, z]) + i([x, iz] + [y, iz]) \\ &= ([x, z] + i[x, iz]) + ([y, z] + i[y, iz]) \\ &= \langle x, z \rangle + \langle y, z \rangle.\end{aligned}$$

So we have $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

(b) By the same reasoning, we also have

$$\begin{aligned}\langle cx, y \rangle &= [cx, y] + i[cx, iy] \\ &= c[x, y] + ci[x, iy] \\ &= c([x, y] + i[x, iy]) \\ &= c\langle x, y \rangle.\end{aligned}$$

So, part (b) is satisfied.

(c)

$$\begin{aligned}\overline{\langle x, y \rangle} &= \overline{[x, y] + i[x, iy]} \\ &= \overline{[x, y]} + \overline{i[x, iy]} \\ &= [y, x] + \bar{i}[iy, x] \\ &= [y, x] + i[y, ix] \\ &= \langle y, x \rangle.\end{aligned}$$

(d) Suppose $x \neq 0$. Since $[x, ix] = 0$ for any $x \in V$ and that $[x, x] > 0$ (Since $[\cdot, \cdot]$ is a real inner product for V), we can write

$$\langle x, x \rangle = [x, x] + i[x, ix] = [x, x] > 0.$$

Thus, property (d) is satisfied. Hence, we conclude that $\langle \cdot, \cdot \rangle$ is a complex inner product on V . ■

6.2 Gram-Schmidt Orthogonalization Process

Exercise 6.2.6

Let V be an inner product space, and let W be a finite-dimensional subspace of V . If $x \notin W$, prove that there exists $y \in V$ such that $y \in W^\perp$, but $\langle x, y \rangle \neq 0$.

Proof. Let $x \in V$ but not in W . By Theorem 6.6, there exists unique scalars $u \in W$ and $y \in W^\perp$ such that $x = u + y$ if and only if $y = x - u$. Then observe that

$$\langle x, y \rangle = \langle x, x - u \rangle = \langle x, x \rangle - \langle x, u \rangle.$$

If $x = 0$, then we would find that $x \in W$ since W is a finite-dimensional subspace of V which is a contradiction. Does x must be non-zero as well as $x - u \neq 0$ since U is a unique scalar. Otherwise, we gain the same contradiction. Thus, we have that $\langle x, y \rangle \neq 0$. ■

Exercise 6.2.7

Let β be a basis for a subspace W an inner product space V , and let $z \in V$. Prove that $z \in W^\perp$ if and only if $\langle z, v \rangle = 0$ for every $v \in \beta$.

Proof. Let $z \in W^\perp$ and $v \in \beta$. Then by definition of W^\perp and the fact that $\beta \subseteq W$, we must have $\langle z, v \rangle = 0$.

Conversely, suppose $\langle z, v \rangle = 0$ for all $v \in \beta$. Since $\beta \subseteq W$, we immediately have $z \in W^\perp$. ■

Exercise 6.2.8

Prove that if $\{w_1, w_2, \dots, w_n\}$ is an orthogonal set of nonzero vectors, then the vectors v_1, v_2, \dots, v_n derived from the Gram-Schmidt process satisfy $v_i = w_i$ for $i = 1, 2, \dots, n$.

Proof. We proceed by mathematical induction on n . Let $k = 1, 2, \dots, n$ and let $S = \{w_1, w_2, \dots, w_k\}$ be an orthogonal set of non-zero vectors. Let $S' = \{v_1, v_2, \dots, v_k\}$ be constructed as in with the setup in Theorem 6.4. Let $k = 1$ be our base case. Then $v_1 = w_1$ as in Theorem 6.4. Now, suppose the result holds for the $k-1$ case. Since each $v_{k-1} = w_{k-1}$ for $k = 1, 2, \dots, n$ and that S is an orthogonal set of non-zero vectors, we have

$$\begin{aligned} v_k &= w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{\|v_j\|^2} v_j \\ &= w_k - \left[\frac{\langle w_k, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle w_k, v_2 \rangle}{\|v_2\|^2} v_2 + \dots + \frac{\langle w_k, v_{k-1} \rangle}{\|v_{k-1}\|^2} v_{k-1} \right] \\ &= w_k - \left[\frac{\langle w_k, w_1 \rangle}{\|w_1\|^2} w_1 + \frac{\langle w_k, w_2 \rangle}{\|w_2\|^2} w_2 + \dots + \frac{\langle w_k, w_{k-1} \rangle}{\|w_{k-1}\|^2} w_{k-1} \right] \\ &= w_k \end{aligned}$$

which ends our induction argument. Thus, we conclude that $v_k = w_k$ for all $k = 1, 2, \dots, n$. ■

Exercise 6.2.10

Let W be a finite-dimensional subspace of an inner product space V . Prove that $V = W \oplus W^\perp$. Prove that there exists a projection T on W along W^\perp that satisfies $N(T) = W^\perp$. In addition, prove that $\|T(x)\| \leq \|x\|$ for all $x \in V$.

Proof. Let $S = \{v_1, v_2, \dots, v_k\}$ be an orthonormal basis for W . Let $x \in V$. By Theorem 6.6, we can find $u \in W$ and $z \in W^\perp$ such that $x = u + z$. Thus, $V = W + W^\perp$. Note that $W \cap W^\perp = \{0\}$. So, $V = W \oplus W^\perp$. Now, define the map

$$T : V \rightarrow V \text{ by } T(x) = u$$

with $x \in V$. We will show that this map is a projection on W along W^\perp . Note that it can be easily shown that this map is linear. Since $V = W \oplus W^\perp$, we have $x = u + z$ where $u \in W$ and $z \in W^\perp$. Thus, T must be a projection on W along W^\perp .

Next, we will show that T satisfies $N(T) = W^\perp$. Let $x \in W^\perp$. Since T is a projection on W along W^\perp , we have $T(x) = 0$. So, $x \in N(T)$. For the other containment, let $x \in N(T)$. Then by definition of $N(T)$, we have $T(x) = 0$. With T being a projection on W along W^\perp , we have $x = 0 + z$ where $z \in W^\perp$. But this tells us that $x \in W^\perp$. Thus, we conclude that $N(T) = W^\perp$.

Let $x \in V$. Then $x = u + z$ since $V = W \oplus W^\perp$. By Exercise 10 from Section 6.1, we have

$$\|x\|^2 = \|u + z\|^2 = \|u\|^2 + \|z\|^2 \geq \|u\|^2 = \|T(x)\|^2.$$

Thus, we have $\|T(x)\| \leq \|x\|$ for all $x \in V$. ■

Exercise 6.2.13

Let V be an inner product space, S and S_0 be subsets of V , and W be a finite-dimensional subspace of V . Prove the following results.

- (a) $S_0 \subseteq S$ implies that $S^\perp \subseteq S_0^\perp$.

Proof. Suppose $S_0 \subseteq S$. Let $s_0 \in S_0$. Then $s_0 \in S$ as well. If we let $x \in S^\perp$, then we see that $\langle x, s_0 \rangle = 0$. But s_0 is an element of S_0 , so $x \in S_0^\perp$ and we are done. ■

- (b) $S \subseteq (S^\perp)^\perp$; so $\text{span}(S) \subseteq (S^\perp)^\perp$.

Proof. Let $v_1, v_2, \dots, v_k \in S$. If $y \in \text{span}(S)$, then

$$y = \sum_{i=1}^k a_i v_i$$

for some scalars a_1, a_2, \dots, a_k . Since $S \subseteq (S^\perp)^\perp$, we see that $v_1, v_2, \dots, v_k \in (S^\perp)^\perp$. So, y must also be an element of $(S^\perp)^\perp$ and we are done. ■

- (c) $W = (W^\perp)^\perp$.

Proof. By part (b), we have $W \subseteq (W^\perp)^\perp$. We need only show the other containment now, namely, $(W^\perp)^\perp \subseteq W$. Suppose for sake of contradiction that $x \notin W$. Now, let $x \in (W^\perp)^\perp$. By definition, we have

$$\langle x, w \rangle = 0 \text{ for all } w \in W^\perp.$$

However, we have $x \notin W$ implies that $w \in W^\perp$ such that $\langle x, w \rangle \neq 0$ for all $w \in W^\perp$ which contradicts the assumption that $x \in (W^\perp)^\perp$. Thus, we must have $(W^\perp)^\perp \subseteq W$. Hence, we conclude that

$$W = (W^\perp)^\perp. \quad \text{■}$$

- (d) $V = W \oplus W^\perp$.

Proof. Since W is a subspace and finite-dimensional, let $\beta = \{v_1, v_2, \dots, v_k\}$ be an orthonormal basis for W . By Theorem 6.7, we have $\dim(V) = \dim(W) + \dim(W^\perp)$. Note that $W \cap W^\perp = \{0\}$. By Exercise 29 (a), we have

$$\begin{aligned} \dim(V) &= \dim(W) + \dim(W^\perp) - \dim(W \cap W^\perp) \\ &= \dim(W + W^\perp). \end{aligned}$$

Hence, $V = W + W^\perp$. Thus, we conclude that

$$V = W \oplus W^\perp.$$

Exercise 6.2.14

Let W_1 and W_2 be subspaces of a finite-dimensional inner product space. Prove that $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$ and $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

Proof. Let V be a finite-dimensional vector space. We will first show that $(W_1 + W_2)^\perp \subseteq W_1^\perp \cap W_2^\perp$. Since $W_1 \cap W_2 \subseteq W_1 + W_2$, we can see by part (a) that $(W_1 + W_2)^\perp \subseteq W_1^\perp \cap W_2^\perp$. Let $x \in W_1^\perp \cap W_2^\perp$. Then $x \in W_1^\perp$ and $x \in W_2^\perp$. So, $\langle x, w_1 \rangle = 0$ and $\langle x, w_2 \rangle = 0$ for all $w_1 \in W_1$ and $w_2 \in W_2$, respectively. Since

$$\langle x, w_1 \rangle + \langle x, w_2 \rangle = \langle x, w_1 + w_2 \rangle = 0,$$

we can see that $x \in (W_1 + W_2)^\perp$. So, $W_1^\perp \cap W_2^\perp \subseteq (W_1 + W_2)^\perp$. Thus,

$$(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp. \quad (1)$$

Now, we can use the first equation to show the second equation. Then by using part(c) of Exercise 13 and (1), we can write that

$$\begin{aligned} (W_1 \cap W_2)^\perp &= ((W_1^\perp)^\perp \cap (W_2^\perp)^\perp) \\ &= ((W_1^\perp + W_2^\perp)^\perp)^\perp \\ &= W_1^\perp + W_2^\perp \end{aligned}$$

which is our desired result. ■

Exercise 6.2.15

Let V be a finite-dimensional inner product space over F .

- (a) *Parseval's Identity.* Let $\{v_1, v_2, \dots, v_n\}$ be an orthonormal basis for V . For any $x, y \in V$ prove that

$$\langle x, y \rangle = \sum_{i=1}^n \langle x, v_i \rangle \overline{\langle y, v_i \rangle}.$$

Proof. Let $x, y \in V$. Since $\{v_1, v_2, \dots, v_n\}$ is an orthonormal basis for V , write

$$x = \sum_{i=1}^n \langle x, v_i \rangle v_i \quad \text{and} \quad y = \sum_{j=1}^n \langle y, v_j \rangle v_j.$$

Since $\langle v_i, v_j \rangle = \delta_{ij}$ where $\delta_{ij} = 1$ whenever $i = j$ and $\delta_{ij} = 0$ otherwise, we have

$$\begin{aligned}\langle x, y \rangle &= \left\langle \sum_{i=1}^n \langle x, v_i \rangle v_i, \sum_{j=1}^n \langle y, v_j \rangle v_j \right\rangle \\ &= \sum_{i=1}^n \langle x, v_i \rangle \left\langle v_i, \sum_{j=1}^n \langle y, v_j \rangle v_j \right\rangle \\ &= \sum_{i=1}^n \langle x, v_i \rangle \sum_{j=1}^n \overline{\langle y, v_j \rangle} \langle v_i, v_j \rangle \\ &= \sum_{i=1}^n \langle x, v_i \rangle \overline{\langle y, v_i \rangle}.\end{aligned}$$

Thus, we have

$$\langle x, y \rangle = \sum_{i=1}^n \langle x, v_i \rangle \overline{\langle y, v_i \rangle}.$$

■

- (b) Use (a) to prove that if β is an orthonormal basis for V with inner product $\langle \cdot, \cdot \rangle$, then for any $x, y \in V$

$$\langle \phi_\beta(x), \phi_\beta(y) \rangle' = \langle [x]_\beta, [y]_\beta \rangle' = \langle x, y \rangle.$$

Proof. Let $x, y \in V$. Since β is an orthonormal basis for V , write

$$x = \sum_{i=1}^n \langle x, v_i \rangle v_i \quad \text{and} \quad y = \sum_{j=1}^n \langle y, v_j \rangle v_j.$$

Since $\langle \cdot, \cdot \rangle'$ is the standard inner product on F^n , we have

$$\langle \phi_\beta(x), \phi_\beta(y) \rangle' = \langle [x]_\beta, [y]_\beta \rangle' = \sum_{i=1}^n \langle x, v_i \rangle \overline{\langle y, v_i \rangle} = \langle x, y \rangle$$

which is our desired result. ■

Exercise 6.2.16

- (a) *Bessel's Inequality.* Let V be an inner product space, and let $S = \{v_1, v_2, \dots, v_n\}$ be an orthonormal subset of V . Prove that for any $x \in V$ we have

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

Proof. Suppose that the orthonormal subset $S = \{v_1, v_2, \dots, v_n\}$ generates some subspace W of V . Note that this implies that W is a finite-dimensional subspace of V . Let $x \in V$. Using Theorem 6.6, there exists unique vectors $u \in W$ and $z \in W^\perp$ such that $x = u + z$. Since S is an orthonormal basis for W , we must have that

$$u = \sum_{i=1}^k \langle x, v_i \rangle v_i.$$

By using Exercise 10 from Section 6.1 and the fact that S is an orthonormal set, we can see that

$$\begin{aligned}\|x\|^2 &= \|u + z\|^2 \geq \|u\|^2 = \left\| \sum_{i=1}^k \langle x, v_i \rangle v_i \right\|^2 \\ &= \sum_{i=1}^n |\langle x, v_i \rangle|^2 \|v_i\|^2 \\ &= \sum_{i=1}^n |\langle x, v_i \rangle|^2\end{aligned}$$

which establishes our result. \blacksquare

(b) In the context of (a), prove that Bessel's inequality is an equality if and only if $x \in \text{span}(S)$.

Proof. Suppose that Bessel's inequality is an equality. Since S is an orthonormal subset of V and Exercise 10 from Section 6.1, we see that

$$\|x\|^2 = \sum_{i=1}^k |\langle x, v_i \rangle|^2 = \sum_{i=1}^n |\langle x, v_i \rangle|^2 \|v_i\|^2 = \left\| \sum_{i=1}^n \langle x, v_i \rangle v_i \right\|^2.$$

This implies that

$$x = \sum_{i=1}^n \langle x, v_i \rangle v_i$$

and thus $x \in \text{span}(S)$.

Conversely, suppose $x \in \text{span}(S)$. Then

$$x = \sum_{i=1}^n \langle x, v_i \rangle v_i.$$

Using Exercise 10 again, we have

$$\|x\|^2 = \left\| \sum_{i=1}^n \langle x, v_i \rangle v_i \right\|^2 = \sum_{i=1}^n |\langle x, v_i \rangle|^2 \|v_i\|^2 = \sum_{i=1}^n |\langle x, v_i \rangle|^2$$

which achieves the equality of (a). \blacksquare

Exercise 6.2.17

Let T be a linear operator on an inner product space V . If $\langle T(x), y \rangle = 0$ for all $x, y \in V$, prove that $T = T_0$. In fact, prove this result if the equality holds for all x and y in some basis for V .

Proof. Let $x, y \in V$. Our goal is to show that $T_0 = T$. Consider T_0 . Then V being an inner product space implies that we have $\langle T_0(x), y \rangle = \langle 0, y \rangle = 0$ for all $x \in V$ where $T_0(x) = 0$. Since $\langle T(x), y \rangle = 0$, this must imply that $T(x) = T_0(x)$. Thus, $T = T_0$. Now, suppose β is a basis for V where $x, y \in \beta$. Using Exercise 9 from the last section, we see that $\langle T_0(x), y \rangle = \langle T(x), y \rangle$ implies that $T_0(x) = T(x)$. So, $T = T_0$. \blacksquare

Exercise 6.2.18

Let $V = C([-1, 1])$. Suppose that W_e and W_o denote the subspaces of V consisting of the even and odd functions, respectively. Prove that $W_e^\perp = W_o$, where the inner product on V is defined by

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt.$$

Proof. Our goal is to show that $W_o \subseteq W_e^\perp$ and $W_e^\perp \subseteq W_o$. Let $f \in W_o$. Consider $\langle f, g \rangle$. Then by definition

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt$$

where $g \in W_e$. We need to show that $\langle f, g \rangle = 0$ for all $g \in W_e$. Since g is an even function and f is an odd function, then the product fg is an odd function. By our integration properties, we have that

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt = \int_{-1}^1 fg(t) dt = 0.$$

Hence, $f \in W_e^\perp$.

Let $f \in W_e^\perp$. Suppose for sake of contradiction that $f \notin W_o$; that is, f is an even function. By definition of W_e^\perp , we have

$$0 = \langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt = \int_{-1}^1 fg(t) dt = 2 \int_0^1 fg(t) dt.$$

But we have fg is a continuous even function where

$$\int_{-1}^1 fg(t) dt - 2 \int_0^1 fg(t) dt \neq 0$$

which is a contradiction since we have assumed that $f \in W_e^\perp$. Hence, we must have $f \in W_o$. Thus, we conclude that $W_o = W_e^\perp$. ■

Exercise 6.2.23

Let V be the vector space defined in Example 5 of Section 1.2, the space of all sequences σ in F (where $F = \mathbb{R}$ or $F = \mathbb{C}$) such that $\sigma(n) \neq 0$ for only finitely many positive integers n . For $\sigma, \mu \in V$, we define $\langle \sigma, \mu \rangle = \sum_{n=1}^{\infty} \sigma(n)\mu(n)$. Since all but a finite number of terms of the series are zero, the series converges.

- (a) Prove that $\langle \cdot, \cdot \rangle$ is an inner product on V , and hence V is an inner product space.

Proof. ■

- (b) For each positive integer n , let e_n be the sequence defined by $e_n(k) = \delta_{nk}$, where δ_{nk} is the Kronecker delta. Prove that $\{e_1, e_2, \dots\}$ is an orthonormal basis for V .

Proof. ■

- (c) Let $\sigma_n = e_1 + e_n$ and $W = \text{span}(\{\sigma_n : n \geq 2\})$.

- (i) Prove that $e_1 \notin W$, so $W \neq V$.

Proof. ■

- (ii) Prove that $W^\perp = \{0\}$, and conclude that $W \neq (W^\perp)^\perp$. Thus the assumption in Exercise 13(c) that W is finite-dimensional is essential.

Proof. ■

6.3 The Adjoint of a Linear Operator

Exercise 6.3.4

Complete the proof of Theorem 6.11.

Proof. Let T and U be linear operators on V whose adjoints exist. Let $x, y \in V$ and $c \in F$. Then

- (a) Since

$$\begin{aligned}\langle x, (T + U)^*(y) \rangle &= \langle (T + U)(x), y \rangle \\ &= \langle T(x) + U(x), y \rangle \\ &= \langle T(x), y \rangle + \langle U(x), y \rangle \\ &= \langle x, T^*(y) \rangle + \langle x, U^*(y) \rangle \\ &= \langle x, (T^* + U^*)(y) \rangle,\end{aligned}$$

we have $(T + U)^* = T^* + U^*$.

- (b) Since

$$\begin{aligned}\langle x, (cT)^*(y) \rangle &= \langle cT(x), y \rangle \\ &= c\langle T(x), y \rangle \\ &= c\langle x, T^*(y) \rangle \\ &= \langle x, \bar{c}T^*(y) \rangle,\end{aligned}$$

we must have $(cT)^* = \bar{c}T^*$.

- (c) Observe that

$$\begin{aligned}\langle x, (TU)^*(y) \rangle &= \langle (TU)(x), y \rangle \\ &= \langle T(U(x)), y \rangle \\ &= \langle U(x), T^*(y) \rangle \\ &= \langle x, U^*(T^*(y)) \rangle \\ &= \langle x, (U^*T^*)(y) \rangle\end{aligned}$$

which implies that $(TU)^* = U^*T^*$.

- (d) Since

$$\langle x, (T^*)^*(y) \rangle = \langle T^*(x), y \rangle = \langle x, T(y) \rangle.$$

1. (e) Observe that

$$\langle x, I^*(y) \rangle = \langle I(x), y \rangle = \langle x, I(y) \rangle$$

■

Exercise 6.3.5

- (a) Complete the proof of the corollary to Theorem 6.11 by using Theorem 6.11, as in the proof of (c).

Proof. Let $A, B \in M_{n \times n}(F)$ and let β be an orthonormal basis for $V = F^n$. Using Theorem 6.10, we have

- (a)

$$\begin{aligned} (A + B)^* &= [L_{A+B}]_\beta^* = [L_A^* + L_B^*]_\beta \\ &= [L_A^*]_\beta + [L_B^*]_\beta \\ &= [L_A]_\beta^* + [L_B]_\beta^* \\ &= A^* + B^*. \end{aligned}$$

- (b) Let $c \in F$. Then

$$\begin{aligned} (cA)^* &= [L_{cA}]_\beta^* = [cL_A^*]_\beta \\ &= c[L_A^*]_\beta \\ &= c[L_A]_\beta^* \\ &= cA^*. \end{aligned}$$

- (c) Using Theorem 2.15, we have

$$\begin{aligned} (AB)^* &= [L_{AB}]_\beta^* = [L_{AB}^*]_\beta \\ &= [(L_A L_B)^*]_\beta \\ &= [L_B^* L_A^*]_\beta \\ &= [L_B^*]_\beta [L_A^*]_\beta \\ &= [L_B]_\beta^* [L_A]_\beta^* \\ &= B^* A^*. \end{aligned}$$

- (d) Since $L_A^{**} = L_A$, we must have $A^{**} = A$.

■

- (b) State a result for nonsquare matrices that is analogous to the corollary to Theorem 6.11, and prove it using a matrix argument.

Proof. For nonsquare matrices, we can just use the definition of the conjugate. Let $1 \leq i \leq n$ and $1 \leq j \leq n$. Thus, we have

(a)

$$\begin{aligned}\left((A+B)^*\right)_{ij} &= \overline{(A+B)_{ji}} = \overline{A_{ji} + B_{ji}} \\ &= \overline{A_{ji}} + \overline{B_{ji}} \\ &= (A^*)_{ij} + (B^*)_{ij}.\end{aligned}$$

Thus, we have $(A+B)^* = A^* + B^*$.

(b)

$$\begin{aligned}\left((cA)^*\right)_{ij} &= \overline{(cA)_{ji}} = \overline{cA_{ji}} \\ &= \overline{c}\overline{A_{ji}} \\ &= \overline{c}(A^*)_{ij}.\end{aligned}$$

Thus, we have $(cA)^* = \overline{c}(A^*)_{ij}$.

(c) Using our definition of matrix multiplication found in Section 2.2, we have

$$\begin{aligned}\left((AB)^*\right)_{ij} &= \overline{(AB)_{ji}} = \overline{\sum_{k=1}^n A_{jk}B_{ki}} = \sum_{k=1}^n \overline{A_{jk}B_{ki}} \\ &= \sum_{k=1}^n \overline{A_{jk}}\overline{B_{ki}} \\ &= \sum_{k=1}^n (B^*)_{ik}(A^*)_{kj} \\ &= (B^*A^*)_{ij}.\end{aligned}$$

(d)

$$(A^{**})_{ij} = \overline{(A^*)_{ji}} = \overline{\overline{A_{ij}}} = A_{ij}.$$

So, $A^{**} = A$.

(e)

$$(I^*)_{ij} = \overline{I_{ji}} = I_{ij}.$$

So, we have $I^* = I$. ■

Exercise 6.3.5

Let T be a linear operator on an inner product space V . Let $U_1 = T + T^*$ and $U_2 = TT^*$. Prove that $U_1 = U_1^*$ and $U_2 = U_2^*$.

Proof. Let $U_1 = T + T^*$ and $U_2 = TT^*$. Using Theorem 6.11, we have

$$U_1^* = (T + T^*)^* = T^* + T^{**} = T^* + T = U_1.$$

Likewise, we have

$$U_2^* = (TT^*)^* = T^{**}T^* = TT^* = U_2. \quad \blacksquare$$

Exercise 6.3.8

Let V be a finite-dimensional inner product space, and let T be a linear operator on V . Prove that if T is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

Proof. Let $x, y \in V$. Our goal is to show that $T^*(T^{-1})^* = (T^{-1})^*T^* = I$. Consider the inner product $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$. Since T is invertible, we have

$$\begin{aligned}\langle T(x), I(y) \rangle &= \langle x, T^*(y) \rangle = \langle T^{-1}T(x), T^*(y) \rangle \\ &= \langle T(x), (T^{-1})^*T^*(y) \rangle.\end{aligned}$$

Using Theorem 6.1, we must have $(T^*)^{-1}T^* = I$. Now, let us show $T^*(T^{-1})^* = I$. Observe that

$$\begin{aligned}\langle T(x), I(y) \rangle &= \langle T(x), T^{-1}T(y) \rangle \\ &= \langle T^*(T^{-1})^*T(x), I(y) \rangle\end{aligned}$$

This implies that $T^*(T^{-1})^*T = T$. Since T is invertible, we have $T^*(T^{-1})^* = I$. Thus, T^* is invertible and therefore $(T^*)^{-1} = (T^{-1})^*$. ■

Exercise 6.3.9

Prove that if $V = W \oplus W^\perp$ and T is the projection on W along W^\perp then $T = T^*$.

Proof. Let $x \in V$ such that $x = u + z$ where $u \in W$ and $z \in W^\perp$. Since $z \in W^\perp$ and that $W^\perp = N(T)$, we have $T(z) = 0$. Furthermore, $\langle z, u \rangle = 0$. Since T is a projection on W along W^\perp , we must have that

$$\langle z, T(u) \rangle = \langle z, T(u) \rangle = 0.$$

But notice that given $u, z \in V$, the fact that

$$\langle T(z), u \rangle = \langle z, T^*(u) \rangle = 0$$

must also be satisfied. This tells us that $\langle z, T(u) \rangle = \langle z, T^*(u) \rangle$ implies $T(u) = T^*(u)$ for all $u \in W$ by Theorem 6.1 and that $T^* = T$. ■

Exercise 6.3.10

Let T be a linear operator on an inner product space V . Prove that $\|T(x)\| = \|x\|$ for all $x \in V$ if and only if $\langle T(x), T(y) \rangle = \langle x, y \rangle$ for all $x, y \in V$.

Proof. For the forwards direction, let $x, y \in V$. Suppose $\|T(x)\| = \|x\|$. Using Exercise 10 from Section 6.1, we have

$$\begin{aligned}\langle x, y \rangle &= \frac{1}{4} [\|x + y\|^2 - \|x - y\|^2] \\ &= \frac{1}{4} [\|T(x + y)\|^2 - \|T(x - y)\|^2] \\ &= \frac{1}{4} [\|T(x) + T(y)\|^2 - \|T(x) - T(y)\|^2].\end{aligned}$$

Consider $\|T(x) + T(y)\|^2$. If $F = \mathbb{R}$, then

$$\begin{aligned}\|T(x) + T(y)\|^2 &= \|T(x)\|^2 + 2\Re\langle T(x), T(y) \rangle + \|T(y)\|^2 \\ &= \|T(x)\|^2 + 2\langle T(x), T(y) \rangle + \|T(y)\|^2\end{aligned}$$

by Exercise 19 from Section 6.1. Likewise, we have

$$\|T(x) - T(y)\|^2 = \|T(x)\|^2 - 2\langle T(x), T(y) \rangle + \|T(y)\|^2.$$

Thus, we conclude that

$$\langle x, y \rangle = \langle T(x), T(y) \rangle.$$

■

Exercise 6.3.11

For a linear operator T on an inner product space V , prove that $T^*T = T_0$ implies $T = T_0$. Is the same result true if we assume that $TT^* = T_0$?

Proof. Let $x \in V$. Consider the norm $\|T(x)\|$. Then observe that

$$\|T(x)\|^2 = \langle T(x), T(x) \rangle = \langle x, T^*T(x) \rangle = \langle x, T_0(x) \rangle = 0$$

since $T^*T = T_0$. By part (d) of Theorem 6.1, $T(x) = 0$ and so we conclude that $T_0 = T$. ■

Exercise 6.3.12

Let V be an inner product space, and let T be a linear operator on V . Prove the following results.

(a) $R(T^*)^\perp = N(T)$.

Proof. Let $y \in R(T^*)^\perp$. Then $\langle y, z \rangle = 0$ for all $z \in R(T^*)$. By definition of $R(T^*)$, we have $T^*(x) = z$ for some $x \in V$. Then we have

$$\begin{aligned}\langle y, z \rangle &= \langle y, T^*(x) \rangle = 0 \\ &= \langle T(y), x \rangle\end{aligned}$$

which implies $\langle T(y), x \rangle = 0$. Using Exercise 17 from Section 6.1, $T = T^0$. So, $T(y) = 0$ for all $y \in V$ and thus $y \in N(T)$. Now, let $y \in N(T)$. Then $T(y) = 0$ for all $y \in V$. Fix $z \in V$. Then

$$\langle T(y), z \rangle = \langle y, T^*(z) \rangle = 0.$$

Observe that $T^*(z) \in R(T^*)$. So, $y \in R(T^*)^\perp$ and we conclude that $R(T^*)^\perp = N(T)$. ■

(b) If V is finite-dimensional, then $R(T^*) = N(T)^\perp$.

Proof. Suppose V is finite-dimensional. Using Exercise 13 (c) from Section 6.2 and part (a), we have

$$R(T^*) = \left(R(T^*)^\perp\right)^\perp = \left(N(T)\right)^\perp.$$

Thus, $R(T^*) = N(T)^\perp$. ■

Exercise 6.3.13

Let T be a linear operator on a finite-dimensional inner product space V . Prove the following results.

- (a) $N(T^*T) = N(T)$. Deduce that $\text{rank}(T^*T) = \text{rank}(T)$.

Proof. Let $x \in N(T^*T)$. Then we have $T^*T(x) = 0$ for all $x \in V$. But note that $T^*T = T_0$, so we have $T = T_0$ by Exercise 6.3.11. Thus, $T(x) = 0$ for all $x \in V$ which implies $x \in N(T)$. Now, let $x \in N(T)$. Then $T(x) = 0$ for all $x \in V$. Then

$$T^*T(x) = T^*(0) = 0$$

for all $x \in V$. Thus, $x \in N(T^*T)$ and so we conclude that $N(T^*T) = N(T)$. Using the Dimension Theorem, we can see that

$$\dim(R(T)) + \dim(N(T)) = \dim(R(T^*T)) + \dim(N(T^*T))$$

implies $\dim(R(T)) = \dim(R(T^*T))$ so $\text{rank}(T) = \text{rank}(T^*T)$. ■

- (b) $\text{rank}(T) = \text{rank}(T^*)$. Deduce from (a) that $\text{rank}(TT^*) = \text{rank}(T)$.

Proof. Since $\text{rank}(T^*T) = \text{rank}(T)$ and that $\text{rank}(T^*T) = \text{rank}(T^*)$, we have $\text{rank}(T) = \text{rank}(T^*)$. From (a), we get that

$$\text{rank}(TT^*) = \text{rank}((T^*T)^*) = \text{rank}(T^*T) = \text{rank}(T).$$

■

- (c) For any $n \times n$ matrix A , $\text{rank}(A^*A) = \text{rank}(AA^*) = \text{rank}(A)$.

Proof. Using parts (a) and (b) as well as the corollary to Theorem 6.10, we have

$$R(L_{A^*A}) = R(L_{A^*}L_A) = R((L_A)^*L_A) = R(L_A)$$

and similarly, we have

$$R(L_{AA^*}) = R(L_AL_{A^*}) = R(L_A(L_A)^*) = R(L_A).$$

So, we conclude that

$$\text{rank}(A^*A) = \text{rank}(AA^*) = \text{rank}(A).$$

■

Exercise 6.3.14

Let V be an inner product space, and let $y, z \in V$. Define $T : V \rightarrow V$ by $T(x) = \langle x, y \rangle z$ for all $x \in V$. First prove that T is linear. Then show that T^* exists, and find an explicit expression for it.

Proof. Fix $y, z \in V$. Let $x_1, x_2 \in V$ and $\varepsilon \in F$. Our first goal is to show that $T : V \rightarrow V$

is linear. Observe that

$$\begin{aligned} T(\varepsilon x_1 + x_2) &= \langle \varepsilon x_1 + x_2, y \rangle z \\ &= \langle \varepsilon x_1, y \rangle z + \langle x_2, y \rangle z \\ &= \varepsilon T(x_1) + T(x_2). \end{aligned}$$

Thus, T is linear. Using Theorem 6.9, there exists a function $T^* : V \rightarrow V$ such that $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$ for $x, y \in V$. Consider the inner product $\langle x, T^*(x) \rangle$. Then observe that

$$\begin{aligned} \langle x, T^*(x) \rangle &= \langle T(x), x \rangle = \langle \langle x, y \rangle z, x \rangle \\ &= \langle x, y \rangle \langle z, x \rangle \\ &= \langle x, \langle x, z \rangle y \rangle. \end{aligned}$$

Using Theorem 6.1, we see that $T^*(x) = \langle x, z \rangle y$. ■

Definition 6.3.1. Let $T : V \rightarrow W$ be a linear transformation, where V and W are finite-dimensional inner product spaces with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$, respectively. A function $T^* : W \rightarrow V$ is called an **adjoint** of T if $\langle T(x), y \rangle_2 = \langle x, T^*(y) \rangle_1$ for all $x \in V$ and $y \in W$.

Exercise 6.3.15

Let $T : V \rightarrow W$ be a linear transformation, where V and W are finite-dimensional inner product spaces with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$, respectively. Prove the following results.

- (a) There is a unique adjoint T^* of T , and T^* is linear.

Proof. Define the function $f : W \rightarrow F$ by $f(y) = \langle T(x), y \rangle_2$. By Theorem 6.8, there exists a unique vector $y' \in V$ such that $f(y) = \langle x, y' \rangle_1$. Define $T^* : W \rightarrow V$ by $T^*(y) = y'$ for $y \in W$. Thus, we have $\langle T(x), y \rangle_2 = \langle x, T^*(y) \rangle_1$. Let $\varepsilon \in F$ and $y_1, y_2 \in W$. Then observe that

$$\begin{aligned} \langle x, T^*(\varepsilon y_1 + y_2) \rangle_1 &= \langle T(x), \varepsilon y_1 + y_2 \rangle_2 \\ &= \langle T(x), \varepsilon y_1 \rangle_2 + \langle T(x), y_2 \rangle_2 \\ &= \varepsilon \langle T(x), y_1 \rangle_2 + \langle T(x), y_2 \rangle_2 \\ &= \langle x, \varepsilon T^*(y_1) \rangle_1 + \langle x, T^*(y_2) \rangle_1 \\ &= \langle x, \varepsilon T^*(y_1) + T^*(y_2) \rangle_1. \end{aligned}$$

By Theorem 6.1, we conclude that $T^*(\varepsilon y_1 + y_2) = \varepsilon T^*(y_1) + T^*(y_2)$ and thus T^* is linear. To prove uniqueness, suppose there exists the linear map $U : W \rightarrow V$ such that $\langle T(x), y \rangle_2 = \langle x, U(y) \rangle_1$. Then we have $\langle x, T^*(y) \rangle_1 = \langle x, U(y) \rangle_1$ implies $T^* = U$. Thus, T^* is a unique linear map. ■

- (b) If β and γ are orthonormal bases for V and W , respectively, then $[T^*]_\gamma^\beta = ([T]_\beta^\gamma)^*$.

Proof. Let $\beta = \{v_1, v_2, \dots, v_n\}$ and $\gamma = \{w_1, w_2, \dots, w_m\}$ be orthonormal bases for V and W , respectively. Define $A = [T]_\beta^\gamma$ and $B = [T^*]_\gamma^\beta$. Using the corollary to Theorem 6.5, we have $B_{ij} = \langle T^*(w_j), v_i \rangle_2$ and $A_{ij} = \langle T(v_j), w_i \rangle_1$. Our goal is to show that

$B_{ij} = (A^*)_{ij}$. For $1 \leq i \leq m$ and $1 \leq j \leq n$, we have

$$\begin{aligned} B_{ij} &= \langle T^*(w_j), v_i \rangle_2 = \overline{\langle v_i, T^*(w_j) \rangle_2} \\ &= \overline{\langle T(v_i), w_j \rangle_1} \\ &= \overline{A_{ji}} \\ &= (A^*)_{ij}. \end{aligned}$$

Thus, we have $[T^*]_\gamma^\beta = ([T]_\beta^\gamma)^*$. ■

(c) $\text{rank}(T^*) = \text{rank}(T)$.

Proof. Apply part Exercise 6.3.13 part (b). ■

(d) $\langle T^*(x), y \rangle_1 = \langle x, T(y) \rangle_2$ for all $x \in W$ and $y \in V$.

Proof. Let $x \in W$ and $y \in V$. Then observe that

$$\langle T^*(x), y \rangle_1 = \overline{\langle y, T^*(x) \rangle_1} = \overline{\langle T(y), x \rangle_2} = \langle x, T(y) \rangle_2$$

which is our desired result. ■

(e) For all $x \in V$, $T^*T(x) = 0$ if and only if $T(x) = 0$.

Proof. Let $x \in V$ and suppose $T^*T(x) = 0$. Consider the norm $\|T(x)\|$. Using part (a), we have

$$\|T(x)\|^2 = \langle T(x), T(x) \rangle_2 = \langle x, T^*T(x) \rangle_1 = 0.$$

By Theorem 6.1, this is true if and only if $T(x) = 0$. Conversely, let $T(x) = 0$ for all $x \in V$. Then

$$T^*T(x) = T^*(T(x)) = T^*(0) = 0$$

which is our desired result. ■

Exercise 6.3.16

State and prove a result that extends the first four parts of Theorem 6.11 using the preceding definition.

Proof. Define $T : V \rightarrow W$ and $U : V \rightarrow W$. Let $x \in V$ and $y \in W$.

(a) Since

$$\begin{aligned} \langle x, (T + U)^*(y) \rangle_1 &= \langle (T + U)(x), y \rangle_2 \\ &= \langle T(x) + U(x), y \rangle_2 \\ &= \langle T(x), y \rangle_2 + \langle U(x), y \rangle_2 \\ &= \langle x, T^*(y) \rangle_1 + \langle x, U^*(y) \rangle_1 \\ &= \langle x, (T^* + U^*)(y) \rangle_1, \end{aligned}$$

we have $(T + U)^* = T^* + U^*$.

(b) Let $c \in F$. Thus, we have

$$\begin{aligned}\langle x, (cT)^*(y) \rangle_1 &= \langle cT(x), y \rangle_2 \\ &= c\langle T(x), y \rangle_2 \\ &= c\langle x, T^*(y) \rangle_1 \\ &= \langle x, \bar{c}T^*(y) \rangle_1,\end{aligned}$$

Thus, we conclude that $(cT)^* = \bar{c}T^*$.

(c) Define $T : W \rightarrow Z$ and $U : V \rightarrow W$. Define $TU : V \rightarrow Z$.

$$\begin{aligned}\langle x, (TU)^*(y) \rangle_1 &= \langle (TU)(x), y \rangle_2 \\ &= \langle T(U(x)), y \rangle_2 \\ &= \langle U(x), T^*(y) \rangle_1 \\ &= \langle x, U^*(T^*(y)) \rangle_1 \\ &= \langle x, (U^*T^*)(y) \rangle_1\end{aligned}$$

which implies that $(TU)^* = U^*T^*$.

(d) Since

$$\langle x, (T^*)^*(y) \rangle_1 = \langle T^*(x), y \rangle_2 = \langle x, T(y) \rangle_1.$$

■

Exercise 6.3.17

Let $T : V \rightarrow W$ be a linear transformation, where V and W are finite-dimensional inner product spaces. Prove that $(R(T^*))^\perp = N(T)$, using the preceding definition.

Proof. Let $y \in R(T^*)^\perp$. Then $\langle y, z \rangle_2 = 0$ for all $z \in R(T^*)$. Since $z \in R(T^*)$, we have $T^*(x) = z$ for some $x \in V$. Then we have

$$0 = \langle y, z \rangle_2 = \langle y, T^*(x) \rangle_2 = \langle T(y), x \rangle_1.$$

By Exercise 17 from Section 6.2, we see that $T = T_0$. So, we must have $T(y) = 0$ for all $y \in V$. Thus, $y \in N(T)$.

Conversely, $y \in N(T)$ implies that

$$\langle T(y), z \rangle_1 = \langle y, T^*(z) \rangle_2.$$

But $T(y) = 0$ for all $y \in V$ which means $\langle y, T^*(z) \rangle_2 = 0$. Since $T^*(z) \in R(T^*)$, we have $y \in R(T^*)^\perp$. Thus, we conclude that $(R(T^*))^\perp = N(T)$. ■

Exercise 6.3.18

Let A be an $n \times n$ matrix. Prove that $\det(A^*) = \overline{\det(A)}$.

Proof. We proceed by mathematical induction on $n \geq 1$. Let A be an $n \times n$ matrix. Note that the result immediately follows for $n = 1$. Now, assume that the result holds for $n - 1 \times n - 1$ matrices. Thus, $\det(\tilde{A}_{1j}) = \det(\tilde{A}_{j1}^*)$. By cofactor expansion along the first

row (or the first column), we have

$$\begin{aligned}
 \overline{\det(A)} &= \overline{\sum_{j=1}^n (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j})} \\
 &= \sum_{j=1}^n (-1)^{1+j} \overline{A_{1j}} \cdot \overline{\det(\tilde{A}_{1j})} \\
 &= \sum_{j=1}^n (-1)^{1+j} (A^*)_{j1} \cdot \det(\tilde{A}_{j1}^*) \\
 &= \det(A^*).
 \end{aligned}$$

Thus, we conclude that $\overline{\det(A)} = \det(A^*)$. ■