Linear Algebra Done Right
Solutions Manual

Sheldon Axler

7 July 1997

This solutions manual provides answers to all the exercises in my book *Linear Algebra Done Right.* It is distributed (without charge) only to instructors who are using *Linear Algebra Done Right* as a textbook. If you are such an instructor, please contact me (preferably via e-mail) to obtain the appropriate version of the solutions manual. This version of the solutions manual is intended to accompany the second edition (1997) of *Linear Algebra Done Right.*

This solutions manual is copyrighted by me, but instructors using *Linear Algebra Done Right* as a textbook have permission to modify and/or abridge this solutions manual and to distribute copies only to students enrolled in their courses (if students are charged for copies, the price may not exceed the cost of materials and reproduction). If you want to modify and/or abridge this solutions manual, you can request (via e-mail) the LaTeX file from me (or from Springer if I am no longer around).

Some instructors may not want to distribute any of these solutions to students; other instructors may want to save class time by distributing to students solutions to some or even many of the exercises. Obviously students should not see the solution to an exercise until they have attempted the exercise themselves. Please do not distribute solutions needlessly, especially if there is a danger that many solutions will get passed down to next year's class. I will not distribute this solutions manual to students, so if neither you nor colleagues in your department make copies for students, you can be reasonably confident that your students are working the exercises without benefit of these solutions.

This solutions manual has not been subjected to the same amount of scrutiny as the book, so errors are more likely. I would be grateful for information about any errors that you notice. If you know nicer solutions to any of the exercises than the solutions given here, please let me know so that I can improve future versions of this solutions manual.

Please check my web site for errata and other information about *Linear Algebra Done Right.* I welcome comments about either the book or the solutions manual.

Have fun!

Sheldon Axler
Mathematics Department
San Francisco State University
San Francisco, CA 94132, USA

e-mail: `axler@math.sfsu.edu`
www home page: `http://math.sfsu.edu/axler`

CHAPTER 1

Vector Spaces

1. Suppose a and b are real numbers, not both 0. Find real numbers c and d such that

$$1/(a+bi) = c+di.$$

SOLUTION: Multiplying both the numerator and the denominator of the left side of the equation above by $a-bi$ gives

$$\frac{a-bi}{a^2+b^2} = c+di.$$

Thus we must have

$$c = \frac{a}{a^2+b^2} \quad \text{and} \quad d = \frac{-b}{a^2+b^2};$$

because a and b are not both 0, we are not dividing by 0.

COMMENT: Note that these formulas for c and d are derived under the assumption that $a+bi$ has a multiplicative inverse. However, we can forget about the derivation and verify (using the definition of complex multiplication) that

$$(a+bi)\Big(\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i\Big) = 1,$$

which shows that every nonzero complex number does indeed have a multiplicative inverse.

2. Show that

$$\frac{-1+\sqrt{3}i}{2}$$

is a cube root of 1 (meaning that its cube equals 1).

SOLUTION: Using the definition of complex multiplication, we have

$$\left(\frac{-1+\sqrt{3}i}{2}\right)^2 = \frac{-1-\sqrt{3}i}{2}.$$

Thus

$$\begin{aligned}\left(\frac{-1+\sqrt{3}i}{2}\right)^3 &= \left(\frac{-1-\sqrt{3}i}{2}\right)\left(\frac{-1+\sqrt{3}i}{2}\right)\\ &= 1.\end{aligned}$$

3. Prove that $-(-v) = v$ for every $v \in V$.

SOLUTION: Let $v \in V$. By the definition of additive inverse, we have

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

The additive inverse of $-v$, which by definition is $-(-v)$, is the unique vector that when added to $-v$ gives 0. The equation above shows that v has this property. Thus $-(-v) = v$.

COMMENT: Using 1.6 twice leads to another proof that $-(-v) = v$. However, the proof given above uses only the additive structure of V, whereas a proof using 1.6 also uses the multiplicative structure.

4. Prove that if $a \in \mathbf{F}$, $v \in V$, and $av = 0$, then $a = 0$ or $v = 0$.

SOLUTION: Suppose that $a \in \mathbf{F}$, $v \in V$, and

$$av = 0.$$

We want to prove that $a = 0$ or $v = 0$. If $a = 0$, then we are done. So suppose that $a \neq 0$. Multiplying both sides of the equation above by $1/a$ gives

$$\frac{1}{a}(av) = \frac{1}{a}0.$$

The associative property shows that the left side of the equation above equals $1v$, which equals v. The right side of the equation above equals 0 (by 1.5). Thus $v = 0$, completing the proof.

5. For each of the following subsets of $\mathbf{F}^3$, determine whether it is a subspace of $\mathbf{F}^3$:

(a) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}$;

(b) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}$;

(c) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 x_2 x_3 = 0\}$;

(d) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 = 5x_3\}$.

SOLUTION: (a) Let

$$U = \{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}.$$

To show that U is a subspace of $\mathbf{F}^3$, first note that $(0, 0, 0) \in U$, so U is nonempty.

Next, suppose that $(x_1, x_2, x_3) \in U$ and $(y_1, y_2, y_3) \in U$. Then

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ y_1 + 2y_2 + 3y_3 &= 0. \end{aligned}$$

Adding these equations, we have

$$(x_1 + y_1) + 2(x_2 + y_2) + 3(x_3 + y_3) = 0,$$

which means that $(x_1 + y_1, x_2 + y_2, x_3 + y_3) \in U$. Thus U is closed under addition.

Next, suppose that $(x_1, x_2, x_3) \in U$ and $a \in \mathbf{F}$. Then

$$x_1 + 2x_2 + 3x_3 = 0.$$

Multiplying this equation by a, we have

$$(ax_1) + 2(ax_2) + 3(ax_3) = 0,$$

which means that $(ax_1, ax_2, ax_3) \in U$. Thus U is closed under scalar multiplication.

Because U is a nonempty subset of $\mathbf{F}^3$ that is closed under addition and scalar multiplication, U is a subspace of $\mathbf{F}^3$.

(b) Let

$$U = \{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}.$$

Then $(4, 0, 0) \in U$ but $0(4, 0, 0)$, which equals $(0, 0, 0)$, is not in U. Thus U is not closed under scalar multiplication. Thus U is not a subspace of $\mathbf{F}^3$.

(c) Let

$$U = \{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 x_2 x_3 = 0\}.$$

Then $(1, 1, 0) \in U$ and $(0, 0, 1) \in U$, but the sum of these two vectors, which equals $(1, 1, 1)$, is not in U. Thus U is not closed under addition. Thus U is not a subspace of $\mathbf{F}^3$.

(d) Let

$$U = \{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 = 5x_3\}.$$

To show that U is a subspace of $\mathbf{F}^3$, first note that $(0, 0, 0) \in U$, so U is nonempty.

Next, suppose that $(x_1, x_2, x_3) \in U$ and $(y_1, y_2, y_3) \in U$. Then

$$\begin{aligned} x_1 &= 5x_3 \\ y_1 &= 5y_3. \end{aligned}$$

Adding these equations, we have

$$x_1 + y_1 = 5(x_3 + y_3),$$

which means that $(x_1 + y_1, x_2 + y_2, x_3 + y_3) \in U$. Thus U is closed under addition.

Next, suppose that $(x_1, x_2, x_3) \in U$ and $a \in \mathbf{F}$. Then

$$x_1 = 5x_3.$$

Multiplying this equation by a, we have

$$ax_1 = 5(ax_3),$$

which means that $(ax_1, ax_2, ax_3) \in U$. Thus U is closed under scalar multiplication.

Because U is a nonempty subset of $\mathbf{F}^3$ that is closed under addition and scalar multiplication, U is a subspace of $\mathbf{F}^3$.

6. Give an example of a nonempty subset U of $\mathbf{R}^2$ such that U is closed under addition and under taking additive inverses (meaning $-u \in U$ whenever $u \in U$), but U is not a subspace of $\mathbf{R}^2$.

SOLUTION: Let $U = \{(m,n) : m \text{ and } n \text{ are integers}\}$. Then clearly U is closed under addition and under taking additive inverses. However, $(1,1) \in U$ but $\frac{1}{2}(1,1)$, which equals $(\frac{1}{2},\frac{1}{2})$, is not in U, so U is not closed under scalar multiplication. Thus U is not a subspace of $\mathbf{R}^2$.

Of course there are also many other examples.

7. Give an example of a nonempty subset U of $\mathbf{R}^2$ such that U is closed under scalar multiplication, but U is not a subspace of $\mathbf{R}^2$.

SOLUTION: Let U be the union of the two coordinate axes in $\mathbf{R}^2$. More precisely, let

$$U = \{(x,0) : x \in \mathbf{R}\} \cup \{(0,y) : y \in \mathbf{R}\}.$$

Then clearly U is closed under scalar multiplication. However, $(1,0)$ and $(0,1)$ are in U but their sum, which equals $(1,1)$ is not in U, so U is not closed under addition. Thus U is not a subspace of $\mathbf{R}^2$.

Of course there are also many other examples.

8. Prove that the intersection of any collection of subspaces of V is a subspace of V.

SOLUTION: Suppose $\{U_\alpha\}_{\alpha\in\Gamma}$ is a collection of subspaces of V; here Γ is an arbitrary index set. We need to prove that $\bigcap_{\alpha\in\Gamma} U_\alpha$, which equals the set of vectors that are in U_α for every $\alpha \in \Gamma$, is a subspace of V.

The additive identity 0 is in U_α for every $\alpha \in \Gamma$ (because each U_α is a subspace of V). Thus $0 \in \bigcap_{\alpha\in\Gamma} U_\alpha$. In particular, $\bigcap_{\alpha\in\Gamma} U_\alpha$ is a nonempty subset of V.

Suppose $u, v \in \bigcap_{\alpha\in\Gamma} U_\alpha$. Then $u, v \in U_\alpha$ for every $\alpha \in \Gamma$. Thus $u + v \in U_\alpha$ for every $\alpha \in \Gamma$ (because each U_α is a subspace of V). Thus $u + v \in \bigcap_{\alpha\in\Gamma} U_\alpha$. Thus $\bigcap_{\alpha\in\Gamma} U_\alpha$ is closed under addition.

Suppose $u \in \bigcap_{\alpha\in\Gamma} U_\alpha$ and $a \in \mathbf{F}$. Then $u \in U_\alpha$ for every $\alpha \in \Gamma$. Thus $au \in U_\alpha$ for every $\alpha \in \Gamma$ (because each U_α is a subspace of V). Thus $au \in \bigcap_{\alpha\in\Gamma} U_\alpha$. Thus $\bigcap_{\alpha\in\Gamma} U_\alpha$ is closed under scalar multiplication.

Because $\bigcap_{\alpha\in\Gamma} U_\alpha$ is a nonempty subset of V that is closed under addition and scalar multiplication, $\bigcap_{\alpha\in\Gamma} U_\alpha$ is a subspace of V.

COMMENT: For many students, the hardest part of this exercise is understanding the meaning of an arbitrary intersection of sets. Instructors who do not want to deal with this issue should change the exercise to "Prove that the intersection of any finite collection of subspaces of V is a subspace of V." Many students will then prove that the intersection of two subspaces of V is a subspace of V and use induction to get the result for finite collections of subspaces.

9. Prove that the union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.

SOLUTION: Suppose U and W are subspaces of V such that $U \cup W$ is a subspace of V. We will use proof by contradiction to show that $U \subset W$ or $W \subset U$. Suppose that our desired result is false. Then $U \not\subset W$ and $W \not\subset U$. This means that there exists $u \in U$ such that $u \notin W$ and there exists $w \in W$ such that $w \notin U$. Because u and w are both in $U \cup W$, which is a subspace of V, we can conclude that $u + w \in U \cup W$. Thus $u + w \in U$ or $u + w \in W$.

First consider the possibility that $u+w \in U$. In this case w, which equals $(u + w) + (-u)$, would be in the sum of two elements of U and hence we would have $w \in U$, contradicting our assumption that $w \notin U$.

Now consider the possibility that $u+w \in W$. In this case u, which equals $(u + w) + (-w)$, would be in the sum of two elements of W and hence we would have $u \in W$, contradicting our assumption that $u \notin W$.

The two paragraphs above show that $u + w \notin U$ and $u + w \notin W$, contradicting the final sentence of the first paragraph of this solution. This contradiction completes our proof that $U \subset W$ or $W \subset U$.

The other direction of this exercise is trivial: if we have two subspaces of V, one of which is contained in the other, then the union of these two subspaces equals the larger of them, which is a subspace of V.

10. Suppose that U is a subspace of V. What is $U + U$?

SOLUTION: By definition, $U+U = \{u+v : u, v \in U\}$. Clearly $U \subset U+U$ because if $u \in U$, then u equals $u + 0$, which expresses u as a sum of two elements of U. Conversely, $U+U \subset U$ because the sum of two elements of U is an element of U (because U is a subspace of V). Conclusion: $U+U = U$.

11. Is the operation of addition on the subspaces of V commutative? Associative? (In other words, if U_1, U_2, U_3 are subspaces of V, is $U_1+U_2 = U_2+U_1$? Is $(U_1 + U_2) + U_3 = U_1 + (U_2 + U_3)$?)

SOLUTION: Suppose U_1, U_2, U_3 are subspaces of V.

A typical element of U_1+U_2 is a vector of the form u_1+u_2, where $u_1 \in U_1$ and $u_2 \in U_2$. Because addition of vectors is commutative, $u_1 + u_2$ equals

$u_2 + u_1$, which is a typical element of $U_2 + U_1$. Thus $U_1 + U_2 = U_2 + U_1$. In other words, the operation of addition on the subspaces of V is commutative.

A typical element of $(U_1 + U_2) + U_3$ is a vector of the form $(u_1 + u_2) + u_3$, where $u_1 \in U_1$, $u_2 \in U_2$, and $u_3 \in U_3$. Because addition of vectors is associative, $(u_1 + u_2) + u_3$ equals $u_1 + (u_2 + u_3)$, which is a typical element of $U_1 + (U_2 + U_3)$. Thus $(U_1 + U_2) + U_3 = U_1 + (U_2 + U_3)$. In other words, the operation of addition on the subspaces of V is associative.

12. Does the operation of addition on the subspaces of V have an additive identity? Which subspaces have additive inverses?

SOLUTION: The subspace $\{0\}$ is an additive identity for the operation of addition on the subspaces of V. More precisely, if U is a subspace of V, then $U + \{0\} = \{0\} + U = U$.

For a subspace U of V to have an additive inverse, there would have to be another subspace W of V such that $U + W = \{0\}$. Because both U and W are contained in $U + W$, this is possible only if $U = W = \{0\}$. Thus $\{0\}$ is the only subspace of V that has an additive inverse.

13. Prove or give a counterexample: if U_1, U_2, W are subspaces of V such that

$$U_1 + W = U_2 + W,$$

then $U_1 = U_2$.

SOLUTION: To construct a counterexample for the assertion above, choose V to be any nonzero vector space. Let $U_1 = \{0\}$, $U_2 = V$, and $W = V$. Then $U_1 + W$ and $U_2 + W$ are both equal to V, but $U_1 \neq U_2$.

Of course there are also many other examples.

14. Suppose U is the subspace of $\mathcal{P}(\mathbf{F})$ consisting of all polynomials p of the form

$$p(z) = az^2 + bz^5,$$

where $a, b \in \mathbf{F}$. Find a subspace W of $\mathcal{P}(\mathbf{F})$ such that $\mathcal{P}(\mathbf{F}) = U \oplus W$.

SOLUTION: Let W be the set of all polynomials (with coefficients in $\mathbf{F}$) whose z^2-coefficient and z^5-coefficient both equal 0. Then every polynomial in $\mathcal{P}(\mathbf{F})$ can be written uniquely in the form $p + q$, where $p \in U$ and $q \in W$. Thus $\mathcal{P}(\mathbf{F}) = U \oplus W$.

COMMENT: There are other possible choices for W that give a correct solution to this exercise, but the choice for W made above is certainly the most natural one.

15. Prove or give a counterexample: if U_1, U_2, W are subspaces of V such that

$$V = U_1 \oplus W \quad \text{and} \quad V = U_2 \oplus W,$$

then $U_1 = U_2$.

SOLUTION: To construct a counterexample for the assertion above, let $V = \mathbf{F}^2$, let $U_1 = \{(x, 0) : x \in \mathbf{F}\}$, let $U_2 = \{(0, y) : y \in \mathbf{F}\}$, and let $W = \{(z, z) : z \in \mathbf{F}\}$. Then

$$\mathbf{F}^2 = U_1 \oplus W \quad \text{and} \quad \mathbf{F}^2 = U_2 \oplus W,$$

as is easy to verify, but $U_1 \neq U_2$.

Of course there are also many other examples.

Chapter 2

Finite-Dimensional Vector Spaces

1. Prove that if $(v_1, \ldots, v_n)$ spans V, then so does the list

$$(v_1 - v_2, v_2 - v_3, \ldots, v_{n-1} - v_n, v_n)$$

obtained by subtracting from each vector (except the last one) the following vector.

SOLUTION: Suppose $(v_1, \ldots, v_n)$ spans V. Let $v \in V$. To show that $v \in \operatorname{span}(v_1 - v_2, v_2 - v_3, \ldots, v_{n-1} - v_n, v_n)$, we need to find $a_1, \ldots, a_n \in \mathbf{F}$ such that

$$v = a_1(v_1 - v_2) + a_2(v_2 - v_3) + \ldots a_{n-1}(v_{n-1} - v_n) + a_n v_n.$$

Rearranging terms of the equation above, we see that we need to find $a_1, \ldots, a_n \in \mathbf{F}$ such that

(a) $$v = a_1 v_1 + (a_2 - a_1)v_2 + (a_3 - a_2)v_3 + \cdots + (a_n - a_{n-1})v_n.$$

Because $(v_1, \ldots, v_n)$ spans V, there exist $b_1, \ldots, b_n \in \mathbf{F}$ such that

(b) $$v = b_1 v_1 + b_2 v_2 + b_3 v_3 + \cdots + b_n v_n.$$

Comparing equations (a) and (b), we see that (a) will be satisfied if we choose a_1 to equal b_1 and then choose a_2 to equal $b_2 + a_1$ and then choose a_3 to equal $b_3 + a_2$, and so on.

2. Prove that if $(v_1, \ldots, v_n)$ is linearly independent in V, then so is the list

$$(v_1 - v_2, v_2 - v_3, \ldots, v_{n-1} - v_n, v_n)$$

obtained by subtracting from each vector (except the last one) the following vector.

SOLUTION: Suppose $(v_1, \ldots, v_n)$ is linearly independent in V. To prove that the list displayed above is linearly independent, suppose $a_1, \ldots, a_n \in \mathbf{F}$ are such that

$$a_1(v_1 - v_2) + a_2(v_2 - v_3) + \cdots + a_{n-1}(v_{n-1} - v_n) + a_n v_n = 0.$$

Rearranging terms, the equation above can be rewritten as

$$a_1 v_1 + (a_2 - a_1)v_2 + (a_3 - a_2)v_3 + \cdots + (a_n - a_{n-1})v_n = 0.$$

Because $(v_1, \ldots, v_n)$ is linearly independent, the equation above implies that

$$\begin{aligned} a_1 &= 0 \\ a_2 - a_1 &= 0 \\ a_3 - a_2 &= 0 \\ &\vdots \\ a_n - a_{n-1} &= 0. \end{aligned}$$

The first equation above tells us that $a_1 = 0$. That information, combined with the second equation, tells us that $a_2 = 0$. That information, combined with the third equation, tells us that $a_3 = 0$. Continue in this fashion, getting $a_1 = \cdots = a_n = 0$. Thus $(v_1 - v_2, v_2 - v_3, \ldots, v_{n-1} - v_n, v_n)$ is linearly independent.

3. Suppose $(v_1, \ldots, v_n)$ is linearly independent in V and $w \in V$. Prove that if $(v_1 + w, \ldots, v_n + w)$ is linearly dependent, then $w \in \operatorname{span}(v_1, \ldots, v_n)$.

SOLUTION: Suppose $(v_1 + w, \ldots, v_n + w)$ is linearly dependent. Then there exist scalars $a_1, \ldots, a_n$, not all 0, such that

$$a_1(v_1 + w) + \cdots + a_n(v_n + w) = 0.$$

Rearranging this equation, we have

$$a_1 v_1 + \cdots + a_n v_n = -(a_1 + \cdots + a_n)w.$$

If $a_1 + \cdots + a_n$ were 0, then the equation above would contradict the linear independence of $(v_1, \ldots, v_n)$. Thus $a_1 + \cdots + a_n \neq 0$. Hence we can divide both sides of the equation above by $-(a_1 + \cdots + a_n)$, showing that $w \in \operatorname{span}(v_1, \ldots, v_n)$.

4. Suppose m is a positive integer. Is the set consisting of 0 and all polynomials with coefficients in $\mathbf{F}$ and with degree equal to m a subspace of $\mathcal{P}(\mathbf{F})$?

SOLUTION: The set consisting of 0 and all polynomials with coefficients in $\mathbf{F}$ and with degree equal to m is not a subspace of $\mathcal{P}(\mathbf{F})$ because it is not closed under addition. Specifically, the sum of two polynomials of degree m may be a polynomial with degree less than m. For example, suppose $m = 2$. Then $7 + 4z + 5z^2$ and $1 + 2z - 5z^2$ are both polynomials of degree 2 but their sum, which equals $8 + 6z$, is a polynomial of degree 1.

5. Prove that $\mathbf{F}^\infty$ is infinite dimensional.

SOLUTION: For each positive integer m, let e_m be the element of $\mathbf{F}^\infty$ whose m^{th} coordinate equals 1 and whose other coordinates equal 0:

$$e_m = (0, \ldots, 0, \underset{\substack{\uparrow \\ m^{\text{th}} \text{ coordinate}}}{1}, 0, \ldots).$$

Then $(e_1, \ldots, e_m)$ is a linearly independent list of vectors in $\mathbf{F}^\infty$, as is easy to verify. This implies, by the marginal comment attached to 2.6, that $\mathbf{F}^\infty$ is infinite dimensional.

6. Prove that the real vector space consisting of all continuous real-valued functions on the interval $[0, 1]$ is infinite dimensional.

SOLUTION: Let V denote the real vector space of all continuous real-valued functions on the interval $[0, 1]$. For each positive integer m, the list $(1, x, \ldots, x^m)$ is linearly independent in V (because if $a_0, \ldots, a_m \in \mathbf{R}$ are such that

$$a_0 + a_1 x + \cdots + a_m x^m = 0$$

for every $x \in [0, 1]$, then the polynomial above has infinitely many roots and hence all its coefficients must equal 0). This implies, by the marginal comment attached to 2.6, that V is infinite dimensional.

7. Prove that V is infinite dimensional if and only if there is a sequence $v_1, v_2, \ldots$ of vectors in V such that $(v_1, \ldots, v_n)$ is linearly independent for every positive integer n.

SOLUTION: First suppose that V is infinite dimensional. Choose v_1 to be any nonzero vector in V. Choose $v_2, v_3, \ldots$ by the following inductive process: suppose that $v_1, \ldots, v_{n-1}$ have been chosen; choose any vector $v_n \in V$ such that $v_n \notin \operatorname{span}(v_1, \ldots, v_{n-1})$—because V is not finite dimensional, $\operatorname{span}(v_1, \ldots, v_{n-1})$ cannot equal V so choosing v_n in this fashion is possible. The linear dependence lemma (2.4) implies that $(v_1, \ldots, v_n)$ is linearly independent for every positive integer n, as desired.

Conversely, suppose there is a sequence $v_1, v_2, \ldots$ of vectors in V such that $(v_1, \ldots, v_n)$ is linearly independent for every positive integer n. This implies, by the marginal comment attached to 2.6, that V is infinite dimensional.

8. Let U be the subspace of $\mathbf{R}^5$ defined by

$$U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{R}^5 : x_1 = 3x_2 \text{ and } x_3 = 7x_4\}.$$

Find a basis of U.

SOLUTION: Obviously

$$U = \{(3x_2, x_2, 7x_4, x_4, x_5) : x_2, x_4, x_5 \in \mathbf{R}\}.$$

From this representation of U, we see easily that

$$((3,1,0,0,0), (0,0,7,1,0), (0,0,0,0,1))$$

is a basis of U.

Of course there are also other possible choices of bases of U.

9. Prove or disprove: there exists a basis (p_0, p_1, p_2, p_3) of $\mathcal{P}_3(\mathbf{F})$ such that none of the polynomials p_0, p_1, p_2, p_3 has degree 2.

SOLUTION: Define $p_0, p_1, p_2, p_3 \in \mathcal{P}_3(\mathbf{F})$ by

$$\begin{aligned} p_0(z) &= 1, \\ p_1(z) &= z, \\ p_2(z) &= z^2 + z^3, \\ p_3(z) &= z^3. \end{aligned}$$

None of the polynomials p_0, p_1, p_2, p_3 has degree 2, but (p_0, p_1, p_2, p_3) is a basis of $\mathcal{P}_3(\mathbf{F})$, as is easy to verify.

Of course there are also other possible choices of bases of $\mathcal{P}_3(\mathbf{F})$ without using polynomials of degree 2.

10. Suppose that V is finite dimensional, with $\dim V = n$. Prove that there exist one-dimensional subspaces $U_1, \ldots, U_n$ of V such that

$$V = U_1 \oplus \cdots \oplus U_n.$$

SOLUTION: Let $(v_1, \ldots, v_n)$ be a basis of V. For each j, let U_j equal $\operatorname{span}(v_j)$; in other words, $U_j = \{av_j : a \in \mathbf{F}\}$. Because $(v_1, \ldots, v_n)$ is a basis of V, each vector in V can be written uniquely in the form

$$a_1 v_1 + \cdots + a_n v_n,$$

where $a_1, \ldots, a_n \in \mathbf{F}$ (see 2.8). By definition of direct sum, this means that $V = U_1 \oplus \cdots \oplus U_n$.

11. Suppose that V is finite dimensional and U is a subspace of V such that $\dim U = \dim V$. Prove that $U = V$.

SOLUTION: Let $(u_1, \ldots, u_n)$ be a basis of U. Thus $n = \dim U$, and by hypothesis we also have $n = \dim V$. Thus $(u_1, \ldots, u_n)$ is a linearly independent (because it is a basis of U) list of vectors in V with length $\dim V$. From 2.17, we see that $(u_1, \ldots, u_n)$ is a basis of V. In particular every vector in V is a linear combination of $(u_1, \ldots, u_n)$. Because each $u_j \in U$, this implies that $U = V$.

12. Suppose that $p_0, p_1, \ldots, p_m$ are polynomials in $\mathcal{P}_m(\mathbf{F})$ such that $p_j(2) = 0$ for each j. Prove that $(p_0, p_1, \ldots, p_m)$ is not linearly independent in $\mathcal{P}_m(\mathbf{F})$.

SOLUTION: Because $p_j(2) = 0$ for each j, the constant polynomial 1 is not in $\operatorname{span}(p_0, \ldots, p_m)$. Thus $(p_0, \ldots, p_m)$ is not a basis of $\mathcal{P}_m(\mathbf{F})$. Because $(p_0, \ldots, p_m)$ is a list of length $m+1$ and $\mathcal{P}_m(\mathbf{F})$ has dimension $m+1$, this implies (by 2.17) that $(p_0, \ldots, p_m)$ is not linearly independent.

13. Suppose U and W are subspaces of $\mathbf{R}^8$ such that $\dim U = 3$, $\dim W = 5$, and $U + W = \mathbf{R}^8$. Prove that $U \cap W = \{0\}$.

SOLUTION: We know (from 2.18) that

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Because $\dim(U + W) = 8$, $\dim U = 3$, and $\dim W = 5$, this implies that $\dim(U \cap W) = 0$. Thus $U \cap W = \{0\}$.

14. Suppose that U and W are both five-dimensional subspaces of $\mathbf{R}^9$. Prove that $U \cap W \neq \{0\}$.

SOLUTION: Using 2.18 we have

$$\begin{aligned} 9 &\geq \dim(U + W) \\ &= \dim U + \dim W - \dim(U \cap W) \\ &= 10 - \dim(U \cap W). \end{aligned}$$

Thus $\dim(U \cap W) \geq 1$. In particular, $U \cap W \neq \{0\}$.

15. You might guess, by analogy with the formula for the number of elements in the union of three subsets of a finite set, that if U_1, U_2, U_3 are subspaces of a finite-dimensional vector space, then

$$\begin{aligned} &\dim(U_1 + U_2 + U_3) \\ &\qquad = \dim U_1 + \dim U_2 + \dim U_3 \\ &\qquad\quad - \dim(U_1 \cap U_2) - \dim(U_1 \cap U_3) - \dim(U_2 \cap U_3) \\ &\qquad\quad + \dim(U_1 \cap U_2 \cap U_3). \end{aligned}$$

Prove this or give a counterexample.

SOLUTION: To give a counterexample, let $V = \mathbf{R}^2$, and let

$$\begin{aligned} U_1 &= \{(x, 0) : x \in \mathbf{R}\}, \\ U_2 &= \{(0, y) : y \in \mathbf{R}\}, \\ U_3 &= \{(x, x) : x \in \mathbf{R}\}. \end{aligned}$$

Then $U_1 + U_2 + U_3 = \mathbf{R}^2$, so $\dim(U_1 + U_2 + U_3) = 2$. However,

$$\dim U_1 = \dim U_2 = \dim U_3 = 1$$

and

$$\dim(U_1 \cap U_2) = \dim(U_1 \cap U_3) = \dim(U_2 \cap U_3) = \dim(U_1 \cap U_2 \cap U_3) = 0.$$

Thus in this case our guess would reduce to the formula $2 = 3$, which obviously is false.

Of course there are also many other examples.

16. Prove that if V is finite dimensional and $U_1, \ldots, U_m$ are subspaces of V, then

$$\dim(U_1 + \cdots + U_m) \leq \dim U_1 + \cdots + \dim U_m.$$

SOLUTION: For each $j = 1, \ldots m$, choose a basis for U_j. Put these bases together to form a single list of vectors in V. Clearly this list spans $U_1 + \cdots + U_m$. Hence the dimension of $U_1 + \cdots + U_m$ is less than or equal to the number of vectors in this list (by 2.10), which equals $\dim U_1 + \cdots + \dim U_m$. In other words,

$$\dim(U_1 + \cdots + U_m) \leq \dim U_1 + \cdots + \dim U_m.$$

17. Suppose V is finite dimensional. Prove that if $U_1, \ldots, U_m$ are subspaces of V such that $V = U_1 \oplus \cdots \oplus U_m$, then

$$\dim V = \dim U_1 + \cdots + \dim U_m.$$

COMMENT: This exercise deepens the analogy between direct sums of subspaces and disjoint unions of subsets. Specifically, compare this exercise to the following obvious statement: if a finite set is written as a disjoint union of subsets, then the number of elements in the set equals the sum of the number of elements in the disjoint subsets.

SOLUTION: Suppose that $U_1, \ldots, U_m$ are subspaces of V such that $V = U_1 \oplus \cdots \oplus U_m$. For each $j = 1, \ldots m$, choose a basis for U_j. Put these bases together to form a single list B of vectors in V. Clearly B spans $U_1 + \cdots + U_m$, which equals V. If we show that B is also linearly independent, then it will be a basis of V. Thus the dimension of V will equal the number of vectors B. In other words, we will have

$$\dim V = \dim U_1 + \cdots + \dim U_m,$$

as desired.

We still need to show that B is linearly independent. To do this, suppose that some linear combination of B equals 0. Write this linear combination as $u_1 + \cdots + u_m$, where we have grouped together the terms that come from the basis vectors of U_1 and called their sum u_1, and similarly up to u_m. Thus we have

$$u_1 + \cdots + u_m = 0,$$

where each $u_j \in U_j$. Because $V = U_1 \oplus \cdots \oplus U_m$, this implies that each u_j equals 0. Because each u_j is a linear combination of our basis of U_j, all the coefficients in the linear combination defining u_j must equal 0. Thus all the coefficients in our original linear combination of B must equal 0. In other words, B is linearly independent, completing our proof.

CHAPTER 3

Linear Maps

1. Show that every linear map from a one-dimensional vector space to itself is multiplication by some scalar. More precisely, prove that if $\dim V = 1$ and $T \in \mathcal{L}(V, V)$, then there exists $a \in \mathbf{F}$ such that $Tv = av$ for all $v \in V$.

SOLUTION: Suppose $\dim V = 1$ and $T \in \mathcal{L}(V, V)$. Let u be any nonzero vector in V. Then every vector in V is a scalar multiple of u. In particular, $Tu = au$ for some $a \in \mathbf{F}$.

Now consider a typical vector $v \in V$. There exists $b \in \mathbf{F}$ such that $v = bu$. Thus

$$\begin{aligned} Tv &= T(bu) \\ &= bT(u) \\ &= b(au) \\ &= a(bu) \\ &= av. \end{aligned}$$

2. Give an example of a function $f\colon \mathbf{R}^2 \to \mathbf{R}$ such that

$$f(av) = af(v)$$

for all $a \in \mathbf{R}$ and all $v \in \mathbf{R}^2$ but f is not linear.

SOLUTION: Define $f\colon \mathbf{R}^2 \to \mathbf{R}$ by

$$f(x, y) = (x^3 + y^3)^{1/3}.$$

Then $f(av) = af(v)$ for all $a \in \mathbf{R}$ and all $v \in \mathbf{R}^2$. However, f is not linear because $f(1, 0) = 1$ and $f(0, 1) = 1$ but

$$\begin{aligned} f\big((1,0)+(0,1)\big) &= f(1,1) \\ &= 2^{1/3} \\ &\neq f(1,0)+f(0,1). \end{aligned}$$

Of course there are also many other examples.

COMMENT: This exercise shows that homogeneity alone is not enough to imply that a function is a linear map. Additivity alone is also not enough to imply that a function is a linear map, although the proof of this involves advanced tools that are beyond the scope of this book.

3. Suppose that V is finite dimensional. Prove that any linear map on a subspace of V can be extended to a linear map on V. In other words, show that if U is a subspace of V and $S \in \mathcal{L}(U, W)$, then there exists $T \in \mathcal{L}(V, W)$ such that $Tu = Su$ for all $u \in U$.

SOLUTION: Suppose U is a subspace of V and $S \in \mathcal{L}(U, W)$. Let $(u_1, \dots, u_m)$ be a basis of U. Then $(u_1, \dots, u_m)$ is a linearly independent list of vectors in V, and so can be extended to a basis $(u_1, \dots, u_m, v_1, \dots, v_n)$ of V (by 2.12). Define $T \in \mathcal{L}(V, W)$ by

$$T(a_1u_1 + \dots a_mu_m + b_1v_1 + \dots b_nv_n) = a_1Su_1 + \cdots + a_mSu_m.$$

Then $Tu = Su$ for all $u \in U$.

COMMENT: Defining $T\colon V \to W$ by

$$Tv = \begin{cases} Sv & \text{if } v \in U; \\ 0 & \text{if } v \notin U. \end{cases}$$

does not work because this map is not linear.

4. Suppose that T is a linear map from V to $\mathbf{F}$. Prove that if $u \in V$ is not in $\operatorname{null} T$, then

$$V = \operatorname{null} T \oplus \{au : a \in \mathbf{F}\}.$$

SOLUTION: Suppose $u \in V$ is not in $\operatorname{null} T$. If $a \in \mathbf{F}$ and $au \in \operatorname{null} T$, then $0 = T(au) = aTu$, which implies that $a = 0$ (because $Tu \neq 0$). Thus

$$\operatorname{null} T \cap \{au : a \in \mathbf{F}\} = \{0\}.$$

If $v \in V$, then

$$v = \left(v - \frac{Tv}{Tu}u\right) + \frac{Tv}{Tu}u.$$

Note that $T\left(v - \frac{Tv}{Tu}u\right) = Tv - \frac{Tv}{Tu}Tu = 0$. Thus the equation above expresses an arbitrary vector $v \in V$ as the sum of a vector in $\operatorname{null} T$ and a scalar multiple of u. Hence $V = \operatorname{null} T + \{au : a \in \mathbf{F}\}$. Using 1.9, we conclude that $V = \operatorname{null} T \oplus \{au : a \in \mathbf{F}\}$.

5. Suppose that $T \in \mathcal{L}(V, W)$ is injective and $(v_1, \dots, v_n)$ is linearly independent in V. Prove that $(Tv_1, \dots, Tv_n)$ is linearly independent in W.

SOLUTION: To show that $(Tv_1, \dots, Tv_n)$ is linearly independent, suppose $a_1, \dots, a_n \in \mathbf{F}$ are such that

$$a_1Tv_1 + \cdots + a_nTv_n = 0.$$

Because T is a linear map, this equation can be rewritten as

$$T(a_1v_1 + \cdots + a_nv_n) = 0.$$

Because T is injective, this implies that

$$a_1v_1 + \cdots + a_nv_n = 0.$$

Because $(v_1, \dots, v_n)$ is linearly independent, the equation above implies that $a_1 = \cdots = a_n = 0$. Thus $(Tv_1, \dots, Tv_n)$ is linearly independent.

6. Prove that if $S_1, \dots, S_n$ are injective linear maps such that $S_1 \dots S_n$ makes sense, then $S_1 \dots S_n$ is injective.

SOLUTION: Suppose that $S_1, \dots, S_n$ are injective linear maps such that $S_1 \dots S_n$ makes sense (which means that the domains of $S_1, \dots, S_n$ are such that $S_1 \dots S_n$ is well defined). Suppose v is a vector in the domain of $S_1 \dots S_n$ (which equals the domain of S_n) such that

$$(S_1 \dots S_n)v = 0.$$

To show that $S_1 \dots S_n$ is injective, we need to show that $v = 0$ (see 3.2). To do this, rewrite the equation above as

$$S_1\big((S_2 \dots S_n)v\big) = 0.$$

Because S_1 is injective, this implies that

$$(S_2 \dots S_n)v = 0.$$

The same argument, now applied to the equation above, shows that

$$(S_3 \dots S_n)v = 0.$$

Repeat this process until reaching the equation $S_n v = 0$, which implies (because S_n is injective) that $v = 0$, as desired.

7. Prove that if $(v_1, \dots, v_n)$ spans V and $T \in \mathcal{L}(V, W)$ is surjective, then $(Tv_1, \dots, Tv_n)$ spans W.

SOLUTION: Suppose that $(v_1, \dots, v_n)$ spans V and $T \in \mathcal{L}(V, W)$ is surjective. Let $w \in W$. Because T is surjective, there exists $v \in V$ such that $Tv = w$. Because $(v_1, \dots, v_n)$ spans V, there exist $a_1, \dots, a_n \in \mathbf{F}$ such that

$$v = a_1 v_1 + \dots + a_n v_n.$$

Applying T to both sides of this equation, we get

$$Tv = a_1 T v_1 + \dots + a_n T v_n.$$

Because $Tv = w$, the equation above implies that $w \in \operatorname{span}(Tv_1, \dots, Tv_n)$. Because w was an arbitrary vector in W, this implies that $(Tv_1, \dots, Tv_n)$ spans W.

8. Suppose that V is finite dimensional and that $T \in \mathcal{L}(V, W)$. Prove that there exists a subspace U of V such that $U \cap \operatorname{null} T = \{0\}$ and $\operatorname{range} T = \{Tu : u \in U\}$.

SOLUTION: There exists a subspace U of V such that

$$V = \operatorname{null} T \oplus U;$$

this follows from 2.13 (with null T playing the role of U and U playing the role of W).

From the definition of direct sum, we have $U \cap \operatorname{null} T = \{0\}$.

Obviously $\operatorname{range} T \supset \{Tu : u \in U\}$. To prove the inclusion in the other direction, suppose $v \in V$. Then there exist $w \in \operatorname{null} T$ and $u' \in U$ such that

$$v = w + u'.$$

Applying T to both sides of this equation, we have $Tv = Tw + Tu' = Tu'$. Thus $Tv \in \{Tu : u \in U\}$. Because v was an arbitrary vector in V (and thus Tv is an arbitrary vector in range T), this implies that

$$\operatorname{range} T \subset \{Tu : u \in U\}.$$

Thus $\operatorname{range} T = \{Tu : u \in U\}$, as desired.

9. Prove that if T is a linear map from $\mathbf{F}^4$ to $\mathbf{F}^2$ such that

$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_1 = 5x_2 \text{ and } x_3 = 7x_4\},$$

then T is surjective.

SOLUTION: Suppose $T \in \mathcal{L}(\ F4, \mathbf{F}^2)$ is such that $\operatorname{null} T$ is as above. Then $\big((5,1,0,0),(0,0,7,1)\big)$ is a basis of $\operatorname{null} T$, and hence $\dim \operatorname{null} T = 2$. From 3.4 we have

$$\begin{aligned} \dim \operatorname{range} T &= \dim \mathbf{F}^4 - \dim \operatorname{null} T \\ &= 4 - 2 \\ &= 2. \end{aligned}$$

Because $\operatorname{range} T$ is a two-dimensional subspace of $\mathbf{R}^2$, we have $\operatorname{range} T = \mathbf{R}^2$. In other words, T is surjective.

10. Prove that there does not exist a linear map from $\mathbf{F}^5$ to $\mathbf{F}^2$ whose null space equals

$$\{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2 \text{ and } x_3 = x_4 = x_5\}.$$

SOLUTION: Suppose U is the subspace of $\mathbf{F}^5$ displayed above. Then $\big((3,1,0,0,0),(0,0,1,1,1)\big)$ is a basis of U, and hence $\dim U = 2$.

If $T \in \mathcal{L}(\mathbf{F}^5, \mathbf{F}^2)$ then from 3.4 we have

$$\begin{aligned} \dim \operatorname{null} T &= \dim \mathbf{F}^5 - \dim \operatorname{range} T \\ &= 5 - \dim \operatorname{range} T \\ &\geq 3 \\ &> \dim U, \end{aligned}$$

where the first inequality holds because $\operatorname{range} T \subset \mathbf{F}^2$. The inequality above shows that if $T \in \mathcal{L}(\mathbf{F}^5, \mathbf{F}^2)$, then $\operatorname{null} T \neq U$, as desired.

11. Prove that if there exists a linear map on V whose null space and range are both finite dimensional, then V is finite dimensional.

SOLUTION: Suppose there exists a linear map T from V into some vector space such that $\operatorname{null} T$ and $\operatorname{range} T$ are both finite dimensional. Thus there exist vectors $u_1, \ldots, u_m \in V$ and $w_1, \ldots, w_n \in \operatorname{range} T$ such that $(u_1, \ldots, u_m)$ spans $\operatorname{null} T$ and $(w_1, \ldots, w_n)$ spans $\operatorname{range} T$. Because each $w_j \in \operatorname{range} T$, there exists $v_j \in V$ such that $w_j = Tv_j$.

Suppose $v \in V$. Then $Tv \in \operatorname{range} T$, so there exist $b_1, \ldots, b_n \in \mathbf{F}$ such that

$$\begin{aligned} Tv &= b_1 w_1 + \cdots + b_n w_n \\ &= b_1 T v_1 + \cdots + b_n T v_n \\ &= T(b_1 v_1 + \ldots b_n v_n). \end{aligned}$$

The equation above implies that $T(v - b_1 v_1 - \cdots - b_n v_n) = 0$. In other words, $v - b_1 v_1 - \cdots - b_n v_n \in \operatorname{null} T$. Thus there exist $a_1, \ldots, a_m \in \mathbf{F}$ such that

$$v - b_1 v_1 - \cdots - b_n v_n = a_1 u_1 + \cdots + a_m u_m.$$

The equation above can be rewritten as

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n.$$

The equation above shows that an arbitrary vector $v \in V$ is a linear combination of $(u_1, \ldots, u_m, v_1, \ldots, v_n)$. In other words, $(u_1, \ldots, u_m, v_1, \ldots, v_n)$ spans V. Thus V is finite dimensional.

COMMENT: The hypothesis of 3.4 is that V is finite dimensional (which is what we are trying to prove in this exercise), so 3.4 cannot be used in this exercise.

12. Suppose that V and W are both finite dimensional. Prove that there exists a surjective linear map from V onto W if and only if $\dim W \leq \dim V$.

SOLUTION: First suppose that there exists a surjective linear map T from V onto W. Then

$$\begin{aligned} \dim W &= \dim \operatorname{range} T \\ &= \dim V - \dim \operatorname{null} T \\ &\leq \dim V, \end{aligned}$$

where the second equality comes from 3.4.

To prove the other direction, now suppose that $\dim W \leq \dim V$. Let $(w_1, \ldots, w_m)$ be a basis of W and let $(v_1, \ldots, v_n)$ be a basis of V. For $a_1, \ldots, a_n \in \mathbf{F}$ define $T(a_1 v_1 + \cdots + a_n v_n)$ by

$$T(a_1v_1 + \cdots + a_nv_n) = a_1w_1 + \cdots + a_mw_m.$$

Because $\dim W \le \dim V$, we have $m \le n$ and so a_m on the right side of the equation above makes sense. Clearly T is a surjective linear map from V onto W.

13. Suppose that V and W are finite dimensional and that U is a subspace of V. Prove that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = U$ if and only if $\dim U \ge \dim V - \dim W$.

SOLUTION: First suppose that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = U$. Then

$$\begin{aligned}\dim U &= \dim \operatorname{null} T\\ &= \dim V - \dim \operatorname{range} T\\ &\ge \dim V - \dim W,\end{aligned}$$

where the second equality comes from 3.4.

To prove the other direction, now suppose that $\dim U \ge \dim V - \dim W$. Let $(u_1, \dots, u_m)$ be a basis of U. Extend to a basis $(u_1, \dots, u_m, v_1, \dots, v_n)$ of V. Let $(w_1, \dots, w_p)$ be a basis of W. For $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{F}$ define $T(a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n)$ by

$$T(a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n) = b_1w_1 + \cdots + b_nw_n.$$

Because $\dim W \ge \dim V - \dim U$, we have $p \ge n$ and so w_n on the right side of the equation above makes sense. Clearly $T \in \mathcal{L}(V, W)$ and $\operatorname{null} T = U$.

14. Suppose that W is finite dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is injective if and only if there exists $S \in \mathcal{L}(W, V)$ such that ST is the identity map on V.

SOLUTION: First suppose that T is injective. Define $S'\colon \operatorname{range} T \to V$ by

$$S'(Tv) = v;$$

because T is injective, each element of $\operatorname{range} T$ can be represented in the form Tv in only one way, so T is well defined. As can be easily checked, S' is a linear map on $\operatorname{range} T$. By Exercise 3 of this chapter, S' can be extended to a linear map $S \in \mathcal{L}(W, V)$. If $v \in V$, then $(ST)v = S(Tv) = S'(Tv) = v$. Thus ST is the identity map on V, as desired.

To prove the implication in the other direction, now suppose that there exists $S \in \mathcal{L}(W, V)$ such that ST is the identity map on V. If $u, v \in V$ are such that $Tu = Tv$, then

$$u = (ST)(u) = S(Tu) = S(Tv) = (ST)v = v$$

and hence $u = v$. Thus T is injective, as desired.

15. Suppose that V is finite dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is surjective if and only if there exists $S \in \mathcal{L}(W, V)$ such that TS is the identity map on W.

SOLUTION: First suppose that T is surjective. Thus W, which equals $\operatorname{range} T$, is finite dimensional (by 3.4). Let $(w_1, \ldots, w_m)$ be a basis of W. Because T is surjective, for each j there exists $v_j \in V$ such that $w_j = Tv_j$. Define $S \in \mathcal{L}(W, V)$ by

$$S(a_1 w_1 + \cdots + a_m w_m) = a_1 v_1 + \cdots + a_m v_m.$$

Then

$$\begin{aligned}(TS)(a_1 w_1 + \cdots + a_m w_m) &= T(a_1 v_1 + \cdots + a_m v_m)\\ &= a_1 T v_1 + \cdots + a_m T v_m\\ &= a_1 w_1 + \cdots + a_m w_m.\end{aligned}$$

Thus TS is the identity map on W.

To prove the implication in the other direction, now suppose that there exists $S \in \mathcal{L}(W, V)$ such that TS is the identity map on W. If $w \in W$, then $w = T(Sw)$, and hence $w \in \operatorname{range} T$. Thus $\operatorname{range} T = W$. In other words, T is surjective, as desired.

16. Suppose that U and V are finite-dimensional vector spaces and that $S \in \mathcal{L}(V, W)$, $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{null} ST \leq \dim \operatorname{null} S + \dim \operatorname{null} T.$$

SOLUTION: Define a linear map $T' \colon \operatorname{null} ST \to V$ by $T'u = Tu$. If $u \in \operatorname{null} ST$, then $S(Tu) = 0$, which means that $Tu \in \operatorname{null} S$. In other words, $\operatorname{range} T' \subset \operatorname{null} S$. Now

$$\begin{aligned}\dim \operatorname{null} ST &= \dim \operatorname{null} T' + \dim \operatorname{range} T'\\ &\leq \dim \operatorname{null} T' + \dim \operatorname{null} S\\ &\leq \dim \operatorname{null} T + \dim \operatorname{null} S,\end{aligned}$$

where the first line follows from 3.4 (applied to T'), the second line holds because range $T' \subset$ null S, and the third line holds because of the obvious inclusion null $T' \subset$ null T.

17. Prove that the distributive property holds for matrix addition and matrix multiplication. In other words, suppose A, B, and C are matrices whose sizes are such that $A(B+C)$ makes sense. Prove that $AB+AC$ makes sense and that $A(B+C) = AB + AC$.

SOLUTION: Because $A(B+C)$ makes sense, B and C must have the same size. Furthermore, the number of columns of A (let's call this number n) must equal the number of rows of B and C. All this means that $AB + AC$ makes sense.

To prove that $A(B+C) = AB + AC$, just use the definition of matrix addition, the definition of matrix multiplication, and the usual distributive property for elements of $\mathbf{F}$. Specifically, let $a_{j,k}$, $b_{j,k}$, and $c_{j,k}$ denote the entries in row j, column k of A, B, and C, respectively. The entry in row j, column k of $B + C$ is $b_{j,k} + c_{j,k}$. Thus the entry in row j, column k of $A(B+C)$ is

$$\sum_{r=1}^{n} a_{j,r}(b_{r,k} + c_{r,k}),$$

which equals

$$\sum_{r=1}^{n} a_{j,r}b_{r,k} + \sum_{r=1}^{n} a_{j,r}c_{r,k},$$

which equals the entry in row j, column k of $AB + AC$, as desired.

18. Prove that matrix multiplication is associative. In other words, suppose A, B, and C are matrices whose sizes are such that $(AB)C$ makes sense. Prove that $A(BC)$ makes sense and that $(AB)C = A(BC)$.

SOLUTION: This exercise can be done by a brute force calculation, in the style of the solution to the previous exercise. Here is a solution that uses only the associativity of the product of linear maps (which is easy to verify because composition of functions is clearly associative) and the nice property that the matrix of the product of two linear maps equals the product of the matrices of the two linear maps (see 3.11).

Suppose A is an m-by-n matrix, B is an n-by-p matrix, and C is a p-by-q matrix; the sizes much match up like this in order for $(AB)C$ to make sense.

Let $R \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m), S \in \mathcal{L}(\mathbf{F}^p, \mathbf{F}^n), T \in \mathcal{L}(\mathbf{F}^q, \mathbf{F}^p)$ be such that, with respect to the standard bases, $\mathcal{M}(R) = A, \mathcal{M}(S) = B, \mathcal{M}(T) = C$; 3.19 insures that such linear maps exist. Now

$$\begin{aligned}(AB)C &= \big(\mathcal{M}(R)\mathcal{M}(S)\big)\mathcal{M}(T) \\ &= \mathcal{M}(RS)\mathcal{M}(T) \\ &= \mathcal{M}\big((RS)T\big) \\ &= \mathcal{M}\big(R(ST)\big) \\ &= \mathcal{M}(R)\mathcal{M}(ST) \\ &= \mathcal{M}(R)\big(\mathcal{M}(S)\mathcal{M}(T)\big) \\ &= A(BC).\end{aligned}$$

19. Suppose $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ and that

$$\mathcal{M}(T) = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix},$$

where we are using the standard bases. Prove that

$$T(x_1, \dots, x_n) = (a_{1,1}x_1 + \cdots + a_{1,n}x_n, \dots, a_{m,1}x_1 + \cdots + a_{m,n}x_n)$$

for every $(x_1, \dots, x_n) \in \mathbf{F}^n$.

COMMENT: This exercise shows T has the form promised on page 39.

SOLUTION: Let $x = (x_1, \dots, x_n) \in \mathbf{F}^n$. Using the standard bases, we then have

$$\begin{aligned}\mathcal{M}(Tx) &= \mathcal{M}(T)\mathcal{M}(x) \\ &= \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{bmatrix},\end{aligned}$$

where the first equality comes from 3.14. The last equation implies that

$$Tx = (a_{1,1}x_1 + \cdots + a_{1,n}x_n, \ldots, a_{m,1}x_1 + \cdots + a_{m,n}x_n),$$

as desired.

20. Suppose $(v_1, \ldots, v_n)$ is a basis of V. Prove that the function $T\colon V \to \mathrm{Mat}(n, 1, \mathbf{F})$ defined by

$$Tv = \mathcal{M}(v)$$

is an invertible linear map of V onto $\mathrm{Mat}(n, 1, \mathbf{F})$; here $\mathcal{M}(v)$ is the matrix of $v \in V$ with respect to the basis $(v_1, \ldots, v_n)$.

SOLUTION: Suppose $u, w \in V$. We can write

$$u = a_1v_1 + \cdots + a_nv_n \quad \text{and} \quad w = b_1v_1 + \cdots + b_nv_n$$

for some $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbf{F}$. Thus

$$u + w = (a_1 + b_1)v_1 + \cdots + (a_n + b_n)v_n.$$

Hence

$$\begin{aligned} T(u + w) &= \mathcal{M}(u + w) \\ &= \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{bmatrix} \\ &= \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \\ &= \mathcal{M}(u) + \mathcal{M}(w) \\ &= Tu + Tw, \end{aligned}$$

which shows that T satisfies that additivity property required for linearity.

If $c \in \mathbf{F}$, then

$$cu = ca_1v_1 + \cdots + ca_nv_n.$$

Hence

$$\begin{aligned}
T(cu) &= \mathcal{M}(cu) \\
&= \begin{bmatrix} ca_1 \\ \vdots \\ ca_n \end{bmatrix} \\
&= c \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \\
&= c\mathcal{M}(u) \\
&= cTu,
\end{aligned}$$

which shows that T satisfies the homogeneity property required for linearity. Thus T is linear.

If $Tu = 0$, then $a_1 = \cdots = a_n = 0$, which implies that $u = 0$. Thus T is injective.

If $c_1, \ldots, c_n \in \mathbf{F}$, then

$$T(c_1 v_1 + \cdots + c_n v_n) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

which implies that T is surjective.

Because the linear map T is injective and surjective, it is invertible (see 3.17).

21. Prove that every linear map from $\text{Mat}(n, 1, \mathbf{F})$ to $\text{Mat}(m, 1, \mathbf{F})$ is given by a matrix multiplication. In other words, prove that if

$$T \in \mathcal{L}(\text{Mat}(n, 1, \mathbf{F}), \text{Mat}(m, 1, \mathbf{F})),$$

then there exists an m-by-n matrix A such that $TB = AB$ for every $B \in \text{Mat}(n, 1, \mathbf{F})$.

SOLUTION: The vector spaces $\text{Mat}(n, 1, \mathbf{F})$ and $\text{Mat}(m, 1, \mathbf{F})$ have obvious bases (consisting of matrices that have 0 in all entries except for a 1 in one entry). Let A be the matrix of T with respect to these bases. Note that if $B \in \text{Mat}(n, 1, \mathbf{F})$, then $\mathcal{M}(B) = B$ and $\mathcal{M}(TB) = TB$. Thus

$$\begin{aligned} TB &= \mathcal{M}(TB) \\ &= \mathcal{M}(T)\mathcal{M}(B) \\ &= AB, \end{aligned}$$

where the second equality comes from 3.14.

22. Suppose that V is finite dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST is invertible if and only if both S and T are invertible.

SOLUTION: First suppose ST is invertible. Thus there exists $R \in \mathcal{L}(V)$ such that $R(ST) = (ST)R = I$. If $v \in V$ is such that $Tv = 0$, then

$$\begin{aligned} v &= Iv \\ &= R(ST)v \\ &= 0. \end{aligned}$$

Because v was an arbitrary vector in $\operatorname{null} T$, this shows that $\operatorname{null} T = \{0\}$. Thus T is injective (by 3.2), and hence T is invertible (by 3.21), as desired.

If $u \in V$, then

$$\begin{aligned} u &= Iu \\ &= (ST)Ru \\ &= S(TRu), \end{aligned}$$

which shows that $u \in \operatorname{range} S$. Because u was an arbitrary vector in V, this implies that $\operatorname{range} S = V$. Thus V is surjective, and hence V is invertible (by 3.21), as desired.

To prove the implication in the other direction, now suppose that both S and T are invertible. Then

$$\begin{aligned} (ST)(T^{-1}S^{-1}) &= S(TT^{-1})S^{-1} \\ &= SS^{-1} \\ &= I \end{aligned}$$

and

$$\begin{aligned} (T^{-1}S^{-1})(ST) &= T^{-1}(S^{-1}S)T \\ &= T^{-1}T \\ &= I. \end{aligned}$$

Thus $T^{-1}S^{-1}$ satisfies the properties required for an inverse of ST. Thus ST is invertible and $(ST)^{-1} = T^{-1}S^{-1}$.

23. Suppose that V is finite dimensional and $S, T \in \mathcal{L}(V)$. Prove that $ST = I$ if and only if $TS = I$.

SOLUTION: First suppose that

$$ST = I.$$

Because I is invertible, the previous exercise implies that S and T are both invertible. Multiply both sides of the equation above by T^{-1} on the right, getting

$$S = T^{-1}.$$

Now multiply both sides of the equation above by T on the left, getting

$$TS = I,$$

as desired.

To prove the implication in the other direction, simply reverse the roles of S and T in the direction we have already proved, showing that if $TS = I$, then $ST = I$.

24. Suppose that V is finite dimensional and $T \in \mathcal{L}(V)$. Prove that T is a scalar multiple of the identity if and only if $ST = TS$ for every $S \in \mathcal{L}(V)$.

SOLUTION: First suppose that $T = aI$ for some $a \in \mathbf{F}$. Let $S \in \mathcal{L}(V)$. Then

$$\begin{aligned} ST &= S(aI) \\ &= aS \\ &= (aI)S \\ &= TS. \end{aligned}$$

To prove the implication in the other direction, suppose now that $ST = TS$ for all $S \in \mathcal{L}(V)$. We begin by proving that (v, Tv) is linearly dependent for every $v \in V$. To do this, fix $v \in V$, and suppose that (v, Tv) is linearly independent. Then (v, Tv) can be extended to a basis $(v, Tv, u_1, \ldots, u_n)$ of V. Define $S \in \mathcal{L}(V)$ by

$$S(av + bTv + c_1u_1 + \cdots + c_nu_n) = bv.$$

Thus $S(Tv) = v$ and $Sv = 0$. Thus the equation $S(Tv) = T(Sv)$ becomes the equation $v = 0$, a contradiction because (v, Tv) was assumed to be linearly independent. This contradiction shows that (v, Tv) is linearly dependent for every $v \in V$. This implies that for each $v \in V \setminus \{0\}$, there exists $a_v \in \mathbf{F}$ such that

$$Tv = a_v v.$$

To show that T is a scalar multiple of the identity, we must show that a_v is independent of v. To do this, suppose $v, w \in V \setminus \{0\}$. We want to show that $a_v = a_w$. First consider the case where (v, w) is linearly dependent. Then there exists $b \in \mathbf{F}$ such that $w = bv$. We have

$$\begin{aligned} a_w w &= Tw \\ &= T(bv) \\ &= bTv \\ &= b(a_v v) \\ &= a_v w, \end{aligned}$$

which shows that $a_v = a_w$, as desired.

Finally, consider the case where (v, w) is linearly independent. We have

$$\begin{aligned} a_{v+w}(v + w) &= T(v + w) \\ &= Tv + Tw \\ &= a_v v + a_w w, \end{aligned}$$

which implies that

$$(a_{v+w} - a_v)v + (a_{v+w} - a_w)w = 0.$$

Because (v, w) is linearly independent, this implies that $a_{v+w} = a_v$ and $a_{v+w} = a_w$, so again we have $a_v = a_w$, as desired.

25. Prove that if V is finite dimensional with $\dim V > 1$, then the set of noninvertible operators on V is not a subspace of $\mathcal{L}(V)$.

SOLUTION: Suppose that V is finite dimensional with $\dim V > 1$. Let $n = \dim V$ and let $(v_1, \ldots, v_n)$ be a basis of V. Define $S, T \in \mathcal{L}(V)$ by

$$S(a_1 v_1 + \cdots + a_n v_n) = a_1 v_1$$

and

$$T(a_1v_1 + \cdots + a_nv_n) = a_2v_2 + \cdots + a_nv_n.$$

Then S is not injective because $Sv_2 = 0$ (this is where we use the hypothesis that $\dim V > 1$), and T is not injective because $Tv_1 = 0$. Thus both S and T are not invertible. However, $S + T$ equals I, which is invertible. Thus the set of noninvertible operators on V is not closed under addition, and hence it is not a subspace of $\mathcal{L}(V)$.

COMMENT: If $\dim V = 1$, then the set of noninvertible operators on V equals $\{0\}$, which is a subspace of $\mathcal{L}(V)$.

26. Suppose n is a positive integer and $a_{i,j} \in \mathbf{F}$ for $i, j = 1, \ldots, n$. Prove that the following are equivalent:

(a) The trivial solution $x_1 = \cdots = x_n = 0$ is the only solution to the homogeneous system of equations

$$\sum_{k=1}^{n} a_{1,k}x_k = 0$$

$$\vdots$$

$$\sum_{k=1}^{n} a_{n,k}x_k = 0.$$

(b) For every $c_1, \ldots, c_n \in \mathbf{F}$, there exists a solution to the system of equations

$$\sum_{k=1}^{n} a_{1,k}x_k = c_1$$

$$\vdots$$

$$\sum_{k=1}^{n} a_{n,k}x_k = c_n.$$

Note that here we have the same number of equations as variables.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{F}^n)$ by

$$T(x_1, \ldots, x_n) = \Big(\sum_{k=1}^{n} a_{1,k}x_k, \ldots, \sum_{k=1}^{n} a_{n,k}x_k\Big).$$

Then (a) above is the assertion that T is injective, and (b) above is the assertion that T is surjective. By 3.21, these two assertions are equivalent.

CHAPTER 4

Polynomials

1. Suppose m and n are positive integers with $m \leq n$. Prove that there exists a polynomial $p \in \mathcal{P}_n(\mathbf{F})$ with exactly m distinct roots.

SOLUTION: Define $p \in \mathcal{P}_n(\mathbf{F})$ by

$$p(z) = (z-1)^{n-m+1}(z-2)(z-3)\dots(z-m).$$

Then p is a polynomial of degree n with exactly m distinct roots (which are $1, \dots, m$).

2. Suppose that $z_1, \dots, z_{m+1}$ are distinct elements of $\mathbf{F}$ and $w_1, \dots, w_{m+1} \in \mathbf{F}$. Prove that there exists a unique polynomial $p \in \mathcal{P}_m(\mathbf{F})$ such that

$$p(z_j) = w_j$$

for $j = 1, \dots, m+1$.

SOLUTION: Define $T\colon \mathcal{P}_m(\mathbf{F}) \to \mathbf{F}^{m+1}$ by

$$Tp = \big(p(z_1), \dots, p(z_{m+1})\big).$$

We need to prove that T is injective (which implies that at most one polynomial p satisfies the condition required by the exercise) and surjective (which implies that at least one polynomial p satisfies the condition required by the exercise).

Clearly T is a linear map. If $p \in \operatorname{null} T$, then

$$p(z_1) = \dots = p(z_{m+1}) = 0,$$

which means that p is a polynomial of degree m with at least $m+1$ distinct roots, which means that $p = 0$ (by 4.3). Thus p is injective, as desired.

Now

$$\begin{aligned}\dim \operatorname{range} T &= \dim \mathcal{P}_m(\mathbf{F}) - \dim \operatorname{null} T \\ &= (m+1) - 0 \\ &= \dim \mathbf{F}^{m+1},\end{aligned}$$

where the first equality comes from 3.4 and the second equality holds because $\operatorname{null} T = \{0\}$. The last equality above implies that $\operatorname{range} T = \mathbf{F}^{m+1}$. Thus T is surjective, as desired.

COMMENT: Surjectivity of T can also be proved by using an explicit construction. But linear algebra, specifically 3.4, gives us surjectivity easily once we get injectivity.

3. Prove that if $p, q \in \mathcal{P}(\mathbf{F})$, with $p \neq 0$, then there exist unique polynomials $s, r \in \mathcal{P}(\mathbf{F})$ such that

$$q = sp + r$$

and $\deg r < \deg p$. In other words, add a uniqueness statement to the division algorithm (4.5).

SOLUTION: Suppose $p, q \in \mathcal{P}(\mathbf{F})$, with $p \neq 0$. We know from the division algorithm (4.5) that there exist $s, r \in \mathcal{P}(\mathbf{F})$, with $\deg r < \deg p$, such that

$$q = sp + r.$$

To prove that s and r are unique, suppose that $\tilde{s}, \tilde{r}$ are in $\mathcal{P}(\mathbf{F})$, with $\deg \tilde{r} < \deg p$ and

$$q = \tilde{s}p + \tilde{r}.$$

Subtracting the last two equations are rearranging, we have

$$(\tilde{s} - s)p = r - \tilde{r}.$$

The right side of the equation above is a polynomial whose degree is less than $\deg p$. If $\tilde{s}$ were not equal to s, then the left side of the equation above would be a polynomial whose degree is at least $\deg p$. Thus we must have $\tilde{s} = s$, which, from the equation above, implies that $\tilde{r} = r$. Thus the choices of s and r were indeed unique.

4. Suppose $p \in \mathcal{P}(\mathbf{C})$ has degree m. Prove that p has m distinct roots if and only if p and its derivative p' have no roots in common.

SOLUTION: First suppose that p has m distinct roots. Because p has degree m, this implies that p can be written in the form

$$p(z) = c(z - \lambda_1) \dots (z - \lambda_m),$$

where $\lambda_1, \dots, \lambda_m$ are distinct. To prove that p and p' have no roots in common, we must show that $p'(\lambda_j) \neq 0$ for each j. To do this, fix j. The expression above for p shows that we can write p in the form

$$p(z) = (z - \lambda_j)q(z),$$

where q is a polynomial such that $q(\lambda_j) \neq 0$. Differentiating both sides of this equation, we have

$$p'(z) = (z - \lambda_j)q'(z) + q(z).$$

Thus

$$\begin{aligned} p'(\lambda_j) &= q(\lambda_j) \\ &\neq 0, \end{aligned}$$

as desired.

To prove the other direction, we will proved the contrapositive, meaning that we will prove that if p has less than m distinct roots, then p and p' have at least one root in common. To do this, suppose that p has less than m distinct roots. Then for some root λ of p, we can write p in the form

$$p(z) = (z - \lambda)^n q(z),$$

where $n \geq 2$ and q is a polynomial. Differentiating both sides of this equation, we have

$$p'(z) = (z - \lambda)^n q'(z) + n(z - \lambda)^{n-1} q(z).$$

Thus $p'(\lambda) = 0$, and so λ is a common root of p and p', as desired.

5. Prove that every polynomial with odd degree and real coefficients has a real root.

SOLUTION: Suppose that p is a polynomial with odd degree and real coefficients. By 4.14, p is a constant times the product of factors of the form

$x - \lambda$ and/or $x^2 + \alpha x + \beta$, where $\lambda, \alpha, \beta \in \mathbf{R}$. Not all the factors can be of the form $x^2 + \alpha x + \beta$, because otherwise p would have even degree. Thus at least one factor must be of the form $x - \lambda$. Any such λ is a real root of p.

COMMENT: Here is another proof, using calculus but not using 4.14. Suppose p is a polynomial with odd degree m. We can write p in the form

$$p(x) = a_0 + a_1 x + \cdots + a_m x^m,$$

where $a_0, \ldots, a_m \in \mathbf{R}$ and $a_m \neq 0$. Replacing p with $-p$ if necessary, we can assume that $a_m > 0$. Now

$$p(x) = x^m \Big(\frac{a_0}{x^m} + \frac{a_1}{x^{m-1}} + \cdots + \frac{a_{m-1}}{x} + a_m\Big).$$

This implies that

$$\lim_{x \to -\infty} p(x) = -\infty \quad \text{and} \quad \lim_{x \to \infty} p(x) = \infty.$$

The intermediate value theorem now implies that there is a real number λ such that $p(\lambda) = 0$. In other words, p has a real root.

CHAPTER 5

Eigenvalues and Eigenvectors

1. Suppose $T \in \mathcal{L}(V)$. Prove that if $U_1, \dots, U_m$ are subspaces of V invariant under T, then $U_1 + \cdots + U_m$ is invariant under T.

SOLUTION: Suppose $U_1, \dots, U_m$ are subspaces of V invariant under T. Consider a vector $u \in U_1 + \cdots + U_m$. There exist $u_1 \in U_1, \dots, u_m \in U_m$ such that

$$u = u_1 + \cdots + u_m.$$

Applying T to both sides of this equation, we get

$$Tu = Tu_1 + \cdots + Tu_m.$$

Because each U_j is invariant under T, we have $Tu_1 \in U_1, \dots, Tu_m \in U_m$. Thus the equation above shows that $Tu \in U_1 + \cdots + U_m$, which implies that $U_1 + \cdots + U_m$ is invariant under T.

2. Suppose $T \in \mathcal{L}(V)$. Prove that the intersection of any collection of subspaces of V invariant under T is invariant under T.

SOLUTION: Suppose $\{U_\alpha\}_{\alpha \in \Gamma}$ is a collection of subspaces of V invariant under T; here Γ is an arbitrary index set. We need to prove that $\bigcap_{\alpha \in \Gamma} U_\alpha$, which equals the set of vectors that are in U_α for every $\alpha \in \Gamma$, is invariant under T. To do this, suppose $u \in \bigcap_{\alpha \in \Gamma} U_\alpha$. Then $u \in U_\alpha$ for every $\alpha \in \Gamma$. Thus $Tu \in U_\alpha$ for every $\alpha \in \Gamma$ (because every U_α is invariant under T). Thus $Tu \in \bigcap_{\alpha \in \Gamma} U_\alpha$, which implies that $\bigcap_{\alpha \in \Gamma} U_\alpha$ is invariant under T.

3. Prove or give a counterexample: if U is a subspace of V that is invariant under every operator on V, then $U = \{0\}$ or $U = V$.

SOLUTION: We will prove that if U is a subspace of V that is invariant under every operator on V, then $U = \{0\}$ or $U = V$. Actually we will prove the (logically equivalent) contrapositive, meaning that we will prove that if U is a subspace of V such that $U \neq \{0\}$ and $U \neq V$, then there exists $T \in \mathcal{L}(V)$ such that U is not invariant under T. To do this, suppose U is a subspace of V such that $U \neq \{0\}$ and $U \neq V$. Choose $u \in U \setminus \{0\}$ (this is possible because $U \neq \{0\}$) and $w \in V \setminus U$ (this is possible because $U \neq V$). Extend the list (u), which is linearly independent because $u \neq 0$, to a basis $(u, v_1, \dots, v_n)$ of V. Define $T \in \mathcal{L}(V)$ by

$$T(au + b_1v_1 + \dots b_nv_n) = aw.$$

Thus $Tu = w$. Because $u \in U$ but $w \notin U$, this shows that U is not invariant under T, as desired.

4. Suppose that $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Prove that $\operatorname{null}(T - \lambda I)$ is invariant under S for every $\lambda \in \mathbf{F}$.

SOLUTION: Fix $\lambda \in \mathbf{F}$. Suppose $v \in \operatorname{null}(T - \lambda I)$. Then

$$\begin{aligned}(T - \lambda I)(Sv) &= TSv - \lambda Sv\\ &= STv - \lambda Sv\\ &= S(Tv - \lambda v)\\ &= 0.\end{aligned}$$

Thus $Sv \in \operatorname{null}(T - \lambda I)$. Hence $\operatorname{null}(T - \lambda I)$ is invariant under S.

5. Define $T \in \mathcal{L}(\mathbf{F}^2)$ by

$$T(w, z) = (z, w).$$

Find all eigenvalues and eigenvectors of T.

SOLUTION: Suppose λ is an eigenvalue of T. For this particular operator, the eigenvalue-eigenvector equation $T(w, z) = \lambda(w, z)$ becomes the system of equations

$$\begin{aligned}z &= \lambda w\\ w &= \lambda z.\end{aligned}$$

Substituting the value for z from the first equation into the second equation gives $w = \lambda^2 w$. Thus $1 = \lambda^2$ (we can ignore the possibility that $w = 0$

because if $w = 0$, then the first equation above implies that $z = 0$). Thus $\lambda = 1$ or $\lambda = -1$. The set of eigenvectors corresponding to the eigenvalue 1 is

$$\{(w, w) : w \in \mathbf{F}\};$$

The set of eigenvectors corresponding to the eigenvalue -1 is

$$\{(w, -w) : w \in \mathbf{F}\}.$$

6. Define $T \in \mathcal{L}(\mathbf{F}^3)$ by

$$T(z_1, z_2, z_3) = (2z_2, 0, 5z_3).$$

Find all eigenvalues and eigenvectors of T.

SOLUTION: Suppose λ is an eigenvalue of T. For this particular operator, the eigenvalue-eigenvector equation $T(z_1, z_2, z_3) = \lambda(z_1, z_2, z_3)$ becomes the system of equations

$$\begin{aligned} 2z_2 &= \lambda z_1 \\ 0 &= \lambda z_2 \\ 5z_3 &= \lambda z_3. \end{aligned}$$

If $\lambda \neq 0$, then the second equation implies that $z_2 = 0$, and the first equation then implies that $z_1 = 0$. Because an eigenvalue must have a nonzero eigenvector, there must be a solution to the system above with $z_3 \neq 0$. The third equation then shows that $\lambda = 5$. In other words, 5 is the only nonzero eigenvalue of T. The set of eigenvectors corresponding to the eigenvalue 5 is

$$\{(0, 0, z_3) : z_3 \in \mathbf{F}\}.$$

If $\lambda = 0$, the first and third equations above show that $z_2 = 0$ and $z_3 = 0$. With these values for z_2, z_3, the equations above are satisfied for all values of z_1. Thus 0 is an eigenvalue of T. The set of eigenvectors corresponding to the eigenvalue 0 is

$$\{(z_1, 0, 0) : z_1 \in \mathbf{F}\}.$$

7. Suppose n is a positive integer and $T \in \mathcal{L}(\mathbf{F}^n)$ is defined by

$$T(x_1, \ldots, x_n) = (x_1 + \cdots + x_n, \ldots, x_1 + \cdots + x_n);$$

in other words, T is the operator whose matrix (with respect to the standard basis) consists of all 1's. Find all eigenvalues and eigenvectors of T.

SOLUTION: Suppose λ is an eigenvalue of T. For this particular operator, the eigenvalue-eigenvector equation $Tx = \lambda x$ becomes the system of equations

$$\begin{gathered} x_1 + \cdots + x_n = \lambda x_1 \\ \vdots \\ x_1 + \cdots + x_n = \lambda x_n. \end{gathered}$$

Thus

$$\lambda x_1 = \cdots = \lambda x_n.$$

Hence either $\lambda = 0$ or $x_1 = \cdots = x_n$.

Consider first the possibility that $\lambda = 0$. In this case all the equations in the eigenvector-eigenvalue system of equations above become the equation $x_1 + \cdots + x_n = 0$. Thus we see that 0 is an eigenvalue of T and that the corresponding set of eigenvectors equals

$$\{(x_1, \ldots, x_n) \in \mathbf{F}^n : x_1 + \cdots + x_n = 0\}.$$

Now consider the possibility that $x_1 = \cdots = x_n$; let t denote the common value of $x_1, \ldots, x_n$. In this case all the equations in the eigenvector-eigenvalue system of equations above become the equation $nt = \lambda t$. Hence λ must equal n (an eigenvalue must have a nonzero eigenvector, so we can take $t \neq 0$). Thus we see that n is an eigenvalue of T and that the corresponding set of eigenvectors equals

$$\{(x_1, \ldots, x_n) \in \mathbf{F}^n : x_1 = \cdots = x_n\}.$$

Because the eigenvector-eigenvalue system of equations above implies that $\lambda = 0$ or $x_1 = \cdots = x_n$, we see that T has no eigenvalues other than 0 and n.

8. Find all eigenvalues and eigenvectors of the backward shift operator $T \in \mathcal{L}(\mathbf{F}^\infty)$ defined by

$$T(z_1, z_2, z_3, \ldots) = (z_2, z_3, \ldots).$$

SOLUTION: Suppose λ is an eigenvalue of T. For this particular operator, the eigenvalue-eigenvector equation $Tz = \lambda z$ becomes the system of equations

$$\begin{aligned} z_2 &= \lambda z_1 \\ z_3 &= \lambda z_2 \\ z_4 &= \lambda z_3 \\ &\vdots \end{aligned}$$

From this we see that we can choose z_1 arbitrarily and then solve for the other coordinates:

$$\begin{aligned} z_2 &= \lambda z_1 \\ z_3 &= \lambda z_2 = \lambda^2 z_1 \\ z_4 &= \lambda z_3 = \lambda^3 z_1 \\ &\vdots \end{aligned}$$

Thus each $\lambda \in \mathbf{F}$ is an eigenvalue of T and the set of corresponding eigenvectors is

$$\{(w, \lambda w, \lambda^2 w, \lambda^3 w \ldots) : w \in \mathbf{F}\}.$$

9. Suppose $T \in \mathcal{L}(V)$ and $\dim \operatorname{range} T = k$. Prove that T has at most $k+1$ distinct eigenvalues.

SOLUTION: Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, and let $v_1, \ldots, v_m$ be corresponding nonzero eigenvectors. If $\lambda_j \neq 0$, then

$$T(v_j/\lambda_j) = v_j.$$

Because at most one of $\lambda_1, \ldots, \lambda_m$ equals 0, this implies that at least $m-1$ of the vectors $v_1, \ldots, v_m$ are in $\operatorname{range} T$. These vectors are linearly independent (by 5.6), which implies that

$$m - 1 \leq \dim \operatorname{range} T = k.$$

Thus $m \leq k + 1$, as desired.

10. Suppose $T \in \mathcal{L}(V)$ is invertible and $\lambda \in \mathbf{F} \backslash \{0\}$. Prove that λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.

SOLUTION: First suppose that λ is an eigenvalue of T. Thus there exists a nonzero vector $v \in V$ such that

$$Tv = \lambda v.$$

Applying T^{-1} to both sides of the equation above, we get $v = \lambda T^{-1}v$, which is equivalent to the equation $T^{-1}v = \frac{1}{\lambda}v$. Thus $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.

To prove the implication in the other direction, replace T by T^{-1} and λ by $\frac{1}{\lambda}$ and then apply the result from the paragraph above.

11. Suppose $S, T \in \mathcal{L}(V)$. Prove that ST and TS have the same eigenvalues.

SOLUTION: Suppose that $\lambda \in \mathbf{F}$ is an eigenvalue of ST. We want to prove that λ is an eigenvalue of TS. Because λ is an eigenvalue of ST, there exists a nonzero vector $v \in V$ such that

$$(ST)v = \lambda v.$$

Now

$$\begin{aligned}(TS)(Tv) &= T(STv)\\ &= T(\lambda v)\\ &= \lambda Tv.\end{aligned}$$

If $Tv \neq 0$, then the equation above shows that λ is an eigenvalue of TS, as desired.

If $Tv = 0$, then $\lambda = 0$ (because $S(Tv) = \lambda v$) and furthermore T is not invertible, which implies that TS is not invertible (by Exercise 22 in Chapter 3), which implies that λ (which equals 0) is an eigenvalue of TS.

Regardless of whether or not $Tv = 0$, we have shown that λ is an eigenvalue of TS. Because λ was an arbitrary eigenvalue of ST, we have shown that every eigenvalue of ST is an eigenvalue of TS.

Reversing the roles of S and T, we conclude that every eigenvalue of TS is also an eigenvalue of ST. Thus ST and TS have the same eigenvalues.

12. Suppose $T \in \mathcal{L}(V)$ is such that every vector in V is an eigenvector of T. Prove that T is a scalar multiple of the identity operator.

SOLUTION: For each $v \in V$, there exists $a_v \in \mathbf{F}$ such that

$$Tv = a_v v.$$

Because $T0 = 0$, we can choose a_0 to be any number in $\mathbf{F}$, but for $v \in V\backslash\{0\}$ the value of a_v is uniquely determined by the equation above.

To show that T is a scalar multiple of the identity, we must show that a_v is independent of v for $v \in V \setminus \{0\}$. To do this, suppose $v, w \in V \setminus \{0\}$. We want to show that $a_v = a_w$. First consider the case where (v, w) is linearly dependent. Then there exists $b \in \mathbf{F}$ such that $w = bv$. We have

$$\begin{aligned} a_w w &= Tw \\ &= T(bv) \\ &= bTv \\ &= b(a_v v) \\ &= a_v w, \end{aligned}$$

which shows that $a_v = a_w$, as desired.

Finally, consider the case where (v, w) is linearly independent. We have

$$\begin{aligned} a_{v+w}(v + w) &= T(v + w) \\ &= Tv + Tw \\ &= a_v v + a_w w, \end{aligned}$$

which implies that

$$(a_{v+w} - a_v)v + (a_{v+w} - a_w)w = 0.$$

Because (v, w) is linearly independent, this implies that $a_{v+w} = a_v$ and $a_{v+w} = a_w$, so again we have $a_v = a_w$, as desired.

13. Suppose $T \in \mathcal{L}(V)$ is such that every subspace of V with dimension

$$\dim V - 1$$

is invariant under T. Prove that T is a scalar multiple of the identity operator.

SOLUTION: Suppose that T is not a scalar multiple of the identity operator. By the previous exercise, there exists $u \in V$ such that u is not an eigenvector of T. Thus (u, Tu) is linearly independent. Extend (u, Tu) to a basis $(u, Tu, v_1, \ldots, v_n)$ of V. Let

$$U = \operatorname{span}(u, v_1, \ldots, v_n).$$

Then U is a subspace of V and $\dim U = \dim V - 1$. However, U is not invariant under T because $u \in U$ but $Tu \notin U$. This contradiction to our hypothesis about T shows that our assumption that T is not a scalar multiple of the identity must have been false.

14. Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Prove that if $p \in \mathcal{P}(\mathbf{F})$ is a polynomial, then

$$p(STS^{-1}) = Sp(T)S^{-1}.$$

SOLUTION: First suppose m is a positive integer. Then

$$\begin{aligned}(STS^{-1})^m &= (STS^{-1})(STS^{-1})\dots(STS^{-1})\\ &= ST(S^{-1}S)T(S^{-1}S)\dots(S^{-1}S)TS^{-1}\\ &= ST^mS^{-1},\end{aligned}$$

which is our desired equation in the special case when $p(z) = z^m$. Multiplying both sides of the equation above by a scalar and then summing a finite number of equations of the resulting form shows that $p(STS^{-1}) = Sp(T)S^{-1}$ for every polynomial $p \in \mathcal{P}(\mathbf{F})$.

15. Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, $p \in \mathcal{P}(\mathbf{C})$, and $a \in \mathbf{C}$. Prove that a is an eigenvalue of $p(T)$ if and only if $a = p(\lambda)$ for some eigenvalue λ of T.

SOLUTION: First suppose that a is an eigenvalue of $p(T)$. Thus $p(T) - aI$ is not injective. Write the polynomial $p(z) - a$ in factored form:

$$p(z) - a = c(z - \lambda_1)\dots(z - \lambda_m),$$

where $c, \lambda_1, \dots, \lambda_m \in \mathbf{C}$. We can assume that $c \neq 0$ (otherwise p is a constant polynomial, in which case the desired result clearly holds). The equation above implies that

$$p(T) - aI = c(T - \lambda_1 I)\dots(T - \lambda_m I).$$

Because $p(T) - aI$ is not injective, this implies that $T - \lambda_j I$ is not injective for some j. In other words, some λ_j is an eigenvalue of T. The formula above for $p(z) - a$ shows that $p(\lambda_j) - a = 0$. Hence $a = p(\lambda_j)$, as desired.

For the other direction, now suppose that $a = p(\lambda)$ for some eigenvalue λ of T. Thus there exists a nonzero vector $v \in V$ such that

$$Tv = \lambda v.$$

Repeatedly applying T to both sides of this equation shows that $T^k v = \lambda^k v$ for every positive integer k. Thus

$$\begin{aligned}p(T)v &= p(\lambda)v\\ &= av.\end{aligned}$$

Thus a is an eigenvalue of $p(T)$.

16. Show that the result in the previous exercise does not hold if $\mathbf{C}$ is replaced with $\mathbf{R}$.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{R}^2)$ by $T(x, y) = (-y, x)$. Define $p \in \mathcal{P}(\mathbf{R})$ by $p(x) = x^2$. Then $p(T) = T^2 = -I$, and hence -1 is an eigenvalue of $p(T)$. However, T has no eigenvalues (as we saw on page 78 of the textbook; the point here is that eigenvalues are required to be real because we are working on a real vector space), so there does not exist an eigenvalue λ of T such that $-1 = p(\lambda)$.

Of course there are also many other examples.

17. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Prove that T has an invariant subspace of dimension j for each $j = 1, \dots, \dim V$.

SOLUTION: There is a basis $(v_1, \dots, v_{\dim V})$ with respect to which T has an upper-triangular matrix (see 5.13). For each $j = 1, \dots, \dim V$, the span of $(v_1, \dots, v_j)$ is a j-dimensional subspace of V that is invariant under T (by 5.12).

18. Give an example of an operator whose matrix with respect to some basis contains only 0's on the diagonal, but the operator is invertible.

SOLUTION: Let $T \in \mathcal{L}(\mathbf{F}^2)$ be the operator whose matrix (with respect to the standard basis) is

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

Obviously this matrix has only 0's on the diagonal, but T is invertible (because $TT = I$, as is clear from squaring the matrix above).

Of course there are also many other examples.

COMMENT: This exercise and the next one show that 5.16 fails without the hypothesis that an upper-triangular matrix is under consideration.

19. Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{F}^2)$ to be the operator whose matrix (with respect to the standard basis) is

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Then $T(1, 0) = T(0, 1) = (1, 1)$, so T is not injective, so T is not invertible, even though the diagonal of the matrix above contains only nonzero numbers.

Of course there are also many other examples.

20. Suppose that $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues and that $S \in \mathcal{L}(V)$ has the same eigenvectors as T (not necessarily with the same eigenvalues). Prove that $ST = TS$.

SOLUTION: Let $n = \dim V$. There is a basis $(v_1, \dots, v_n)$ of V consisting of eigenvectors of T (see 5.20 and 5.21). Letting $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues, we have

$$Tv_j = \lambda_j v_j$$

for each j. Each v_j is also an eigenvector of S, so

$$Sv_j = \alpha_j v_j$$

for some $\alpha_j \in \mathbf{F}$.

For each j, we have

$$(ST)v_j = S(Tv_j) = \lambda_j Sv_j = \alpha_j \lambda_j v_j$$

and

$$(TS)v_j = T(Sv_j) = \alpha_j Tv_j = \alpha_j \lambda_j v_j.$$

Because the operators ST and TS agree on a basis, they are equal.

21. Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$. Prove that $V = \operatorname{null} P \oplus \operatorname{range} P$.

SOLUTION: First suppose $u \in \operatorname{null} P \cap \operatorname{range} P$. Then $Pu = 0$, and there exists $w \in V$ such that $u = Pw$. Applying P to both sides of the last equation, we have $Pu = P^2 w = Pw$. But $Pu = 0$, so this implies that $Pw = 0$. Because $u = Pw$, this implies that $u = 0$. Because u was an arbitrary vector in $\operatorname{null} P \cap \operatorname{range} P$, this implies that $\operatorname{null} P \cap \operatorname{range} P = \{0\}$.

Now suppose $v \in V$. Then obviously

$$v = (v - Pv) + Pv.$$

Note that $P(v - Pv) = Pv - P^2 v = 0$, so $(v - Pv) \in \operatorname{null} P$. Clearly $Pv \in \operatorname{range} P$. Thus the equation above shows that $v \in \operatorname{null} P + \operatorname{range} P$. Because v was an arbitrary vector in V, this implies that $V = \operatorname{null} P + \operatorname{range} P$.

We have shown that $\operatorname{null} P \cap \operatorname{range} P = \{0\}$ and $V = \operatorname{null} P + \operatorname{range} P$. Thus $V = \operatorname{null} P \oplus \operatorname{range} P$ (by 1.9).

22. Suppose $V = U \oplus W$, where U and W are nonzero subspaces of V. Find all eigenvalues and eigenvectors of $P_{U,W}$.

SOLUTION: Because $V = U \oplus W$, each vector $v \in V$ can be written uniquely in the form

$$v = u + w,$$

where $u \in U$ and $w \in W$. Recall that if v is represented as above, then $P_{U,W}v = u$.

Suppose $\lambda \in \mathbf{F}$ is an eigenvalue of $P_{U,W}$. Then there exists a nonzero vector $v \in V$ such that $P_{U,W}v = \lambda v$. Writing this equation using the representation of v given above, we have $u = \lambda(u + w)$. Thus

$$(1 - \lambda)u - \lambda w = 0.$$

Because $V = U \oplus W$, if 0 is written as the sum of a vector in U and a vector in W, then both vectors must be 0. Thus the equation above implies that $(1 - \lambda)u = \lambda w = 0$. Because u and w are not both 0 (because $v \neq 0$), this implies that $\lambda = 1$ or $\lambda = 0$.

For $v \in V$ with representation as above, the equation $P_{U,W}v = 0$ is equivalent to the equation $u = 0$, which is equivalent to the equation $v = w$, which is equivalent to the statement that $v \in W$. This means that 0 is an eigenvalue of $P_{U,W}$ (because W is a nonzero subspace of V) and that W equals the set of eigenvectors corresponding to the eigenvalue 0.

For $v \in V$ with representation as above, the equation $P_{U,W}v = v$ is equivalent to the equation $v = u$, which is equivalent to the statement that $v \in U$. This means that 1 is an eigenvalue of $P_{U,W}$ (because U is a nonzero subspace of V) and that U equals the set of eigenvectors corresponding to the eigenvalue 1.

23. Give an example of an operator $T \in \mathcal{L}(\mathbf{R}^4)$ such that T has no (real) eigenvalues.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{R}^4)$ by

$$T(x_1, x_2, x_3, x_4) = (-x_2, x_1, -x_4, x_3).$$

Suppose $\lambda \in \mathbf{R}$. For this particular operator, the eigenvalue-eigenvector equation $T(x_1, x_2, x_3, x_4) = \lambda(x_1, x_2, x_3, x_4)$ becomes the system of equations

$$\begin{aligned} -x_2 &= \lambda x_1 \\ x_1 &= \lambda x_2 \\ -x_4 &= \lambda x_3 \\ x_3 &= \lambda x_4. \end{aligned}$$

Multiplying together the first two equations and also multiplying together the last two equations gives $-x_1x_2 = \lambda^2 x_1x_2$ and $-x_3x_4 = \lambda^2 x_3x_4$. If either x_1 or x_2 does not equal 0, then the first two equations show that neither of x_1, x_2 equals 0. Similarly, if either x_3 or x_4 does not equal 0, then the last two equations show that neither of x_3, x_4 equals 0. Thus if λ is an eigenvalue of T, then there is a solution to the system of equations above with $x_1x_2 \neq 0$ or $x_3x_4 \neq 0$. Either way, we conclude that $-1 = \lambda^2$, which is impossible for any real number λ. Thus T has no real eigenvalues.

24. Suppose V is a real vector space and $T \in \mathcal{L}(V)$ has no eigenvalues. Prove that every subspace of V invariant under T has even dimension.

SOLUTION: Suppose U is a subspace of V that is invariant under T. Thus $T|_U \in \mathcal{L}(U)$. If $\dim U$ were odd, then $T|_U$ would have an eigenvalue $\lambda \in \mathbf{R}$ (by 5.26), so there would exist a nonzero vector $u \in U$ such that

$$T|_U u = \lambda u.$$

Obviously this would imply that $Tu = \lambda u$, which would imply that λ is an eigenvalue of T. But T has no eigenvalues, so $\dim U$ must be even, as desired.

CHAPTER 6

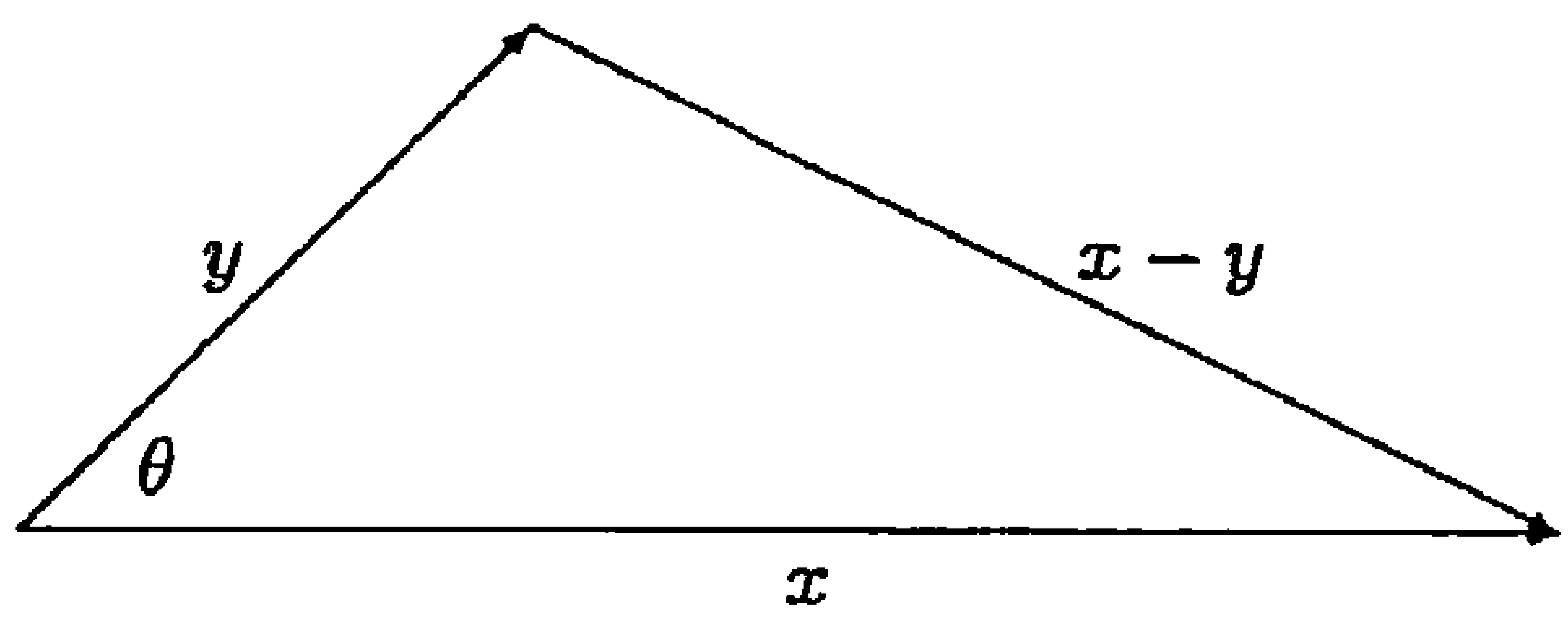
Inner-Product Spaces

1. Prove that if x, y are nonzero vectors in $\mathbf{R}^2$, then

$$\langle x, y\rangle = \|x\| \|y\| \cos\theta,$$

where θ is the angle between x and y (thinking of x and y as arrows with initial point at the origin). *Hint:* draw the triangle formed by x, y, and $x-y$; then use the law of cosines.

SOLUTION: Suppose that x, y are nonzero vectors in $\mathbf{R}^2$ and θ is the angle between x and y. Consider the triangle formed by x, y, and $x - y$:



The law of cosines states that

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \cos\theta.$$

As usual, we compute the norm of a vector squared by taking the inner product of the vector with itself:

$$\begin{aligned}\|x - y\|^2 &= \langle x - y, x - y\rangle \\ &= \langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle.\end{aligned}$$

Substitute the last expression for $\|x-y\|^2$ into the left side of the law of cosines, obtaining

$$\|x\|^2+\|y\|^2-2\langle x,y\rangle=\|x\|^2+\|y\|^2-2\|x\|\|y\|\cos\theta.$$

Now subtract $\|x\|^2+\|y\|^2$ from both sides of the equation above, and then divide both sides by -2, obtaining

$$\langle x,y\rangle=\|x\|\|y\|\cos\theta.$$

2. Suppose $u,v\in V$. Prove that $\langle u,v\rangle=0$ if and only if

$$\|u\|\leq\|u+av\|$$

for all $a\in\mathbf{F}$.

SOLUTION: First suppose that $\langle u,v\rangle=0$. Let $a\in\mathbf{F}$. Then u,av are orthogonal. The Pythagorean theorem thus implies that

$$\begin{aligned}\|u+av\|^2&=\|u\|^2+\|av\|^2\\&\geq\|u\|^2.\end{aligned}$$

Taking square roots gives $\|u\|\leq\|u+av\|$, as desired.

To prove the implication in the other direction, now suppose that $\|u\|\leq\|u+av\|$ for all $a\in\mathbf{F}$. Squaring this inequality, we get

$$\begin{aligned}\|u\|^2&\leq\|u+av\|^2\\&=\langle u+av,u+av\rangle\\&=\langle u,u\rangle+\langle u,av\rangle+\langle av,u\rangle+\langle av,av\rangle\\&=\|u\|^2+\bar{a}\langle u,v\rangle+a\overline{\langle u,v\rangle}+|a|^2\|v\|^2\\&=\|u\|^2+2\operatorname{Re}\bar{a}\langle u,v\rangle+|a|^2\|v\|^2\end{aligned}$$

for all $a\in\mathbf{F}$. Thus

$$-2\operatorname{Re}\bar{a}\langle u,v\rangle\leq|a|^2\|v\|^2$$

for all $a\in\mathbf{F}$. In particular, we can let a equal $-t\langle u,v\rangle$ for $t>0$. Substituting this value for a into the inequality above gives

$$2t|\langle u,v\rangle|^2\leq t^2|\langle u,v\rangle|^2\|v\|^2$$

for all $t>0$. Divide both sides of the inequality above by t, getting

$$2|\langle u,v\rangle|^2 \leq t|\langle u,v\rangle|^2\|v\|^2$$

for all $t > 0$. If $v = 0$, then $\langle u,v\rangle = 0$, as desired. If $v \neq 0$, set t equal to $1/\|v\|^2$ in the inequality above, getting

$$2|\langle u,v\rangle|^2 \leq |\langle u,v\rangle|^2,$$

which implies that $\langle u,v\rangle = 0$, as desired.

3. Prove that

$$\Big(\sum_{j=1}^{n} a_j b_j\Big)^2 \leq \Big(\sum_{j=1}^{n} j{a_j}^2\Big)\Big(\sum_{j=1}^{n} \frac{{b_j}^2}{j}\Big)$$

for all real numbers $a_1,\dots,a_n$ and $b_1,\dots,b_n$.

SOLUTION: Suppose $a_1,\dots,a_n,b_1,\dots,b_n \in \mathbf{R}$. Using the usual inner product on $\mathbf{R}^n$, we have

$$\begin{aligned}
\Big(\sum_{j=1}^{n} a_j b_j\Big)^2 &= \Big(\sum_{j=1}^{n} (\sqrt{j}a_j)(b_j/\sqrt{j})\Big)^2 \\
&= \langle (a_1, \sqrt{2}a_2, \dots, \sqrt{n}a_n), (b_1, b_2/\sqrt{2}, \dots, b_n/\sqrt{n})\rangle^2 \\
&\leq \|(a_1, \sqrt{2}a_2, \dots, \sqrt{n}a_n)\|^2 \|(b_1, b_2/\sqrt{2}, \dots, b_n/\sqrt{n})\|^2 \\
&= \Big(\sum_{j=1}^{n} j{a_j}^2\Big)\Big(\sum_{j=1}^{n} {b_j}^2/j\Big),
\end{aligned}$$

where the inequality above comes from the Cauchy-Schwarz inequality.

4. Suppose $u, v \in V$ are such that

$$\|u\| = 3, \quad \|u+v\| = 4, \quad \|u-v\| = 6.$$

What number must $\|v\|$ equal?

SOLUTION: From the parallelogram equality, we have

$$\begin{aligned}
\|v\|^2 &= \frac{\|u+v\|^2 + \|u-v\|^2 - 2\|u\|^2}{2} \\
&= \frac{16 + 36 - 18}{2} \\
&= 17.
\end{aligned}$$

Thus $\|v\| = \sqrt{17}$.

5. Prove or disprove: there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$\|(x_1, x_2)\| = |x_1| + |x_2|$$

for all $(x_1, x_2) \in \mathbf{R}^2$.

SOLUTION: We will show that there does not exist an inner product on $\mathbf{R}^2$ such that the associated norm is given by the formula above by showing that the parallelogram equality is violated. Let

$$u = (3, 2) \quad \text{and} \quad v = (1, 3).$$

Then

$$u + v = (4, 5) \quad \text{and} \quad u - v = (2, -1).$$

Using the formula above, we then have

$$\begin{aligned} \|u + v\|^2 + \|u - v\|^2 &= 81 + 9 \\ &= 90 \end{aligned}$$

and

$$\begin{aligned} 2(\|u\|^2 + \|v\|^2) &= 2(25 + 16) \\ &= 82. \end{aligned}$$

Thus the parallelogram equality fails, as desired.

6. Prove that if V is a real inner-product space, then

$$\langle u, v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2}{4}$$

for all $u, v \in V$.

SOLUTION: Suppose V is a real inner-product space and $u, v \in V$. Then

$$\begin{aligned} \frac{\|u + v\|^2 - \|u - v\|^2}{4} &= \frac{\langle u + v, u + v \rangle - \langle u - v, u - v \rangle}{4} \\ &= \frac{\|u\|^2 + 2\langle u, v \rangle + \|v\|^2 - (\|u\|^2 - 2\langle u, v \rangle + \|v\|^2)}{4} \\ &= \frac{4\langle u, v \rangle}{4} \\ &= \langle u, v \rangle, \end{aligned}$$

as desired.

7. Prove that if V is a complex inner-product space, then

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i}{4}$$

for all $u, v \in V$.

SOLUTION: Suppose V is a complex inner-product space and $u, v \in V$. Then

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \|u\|^2 + \langle u, v\rangle + \langle v, u\rangle + \|v\|^2\end{aligned}$$

and

$$\begin{aligned}-\|u-v\|^2 &= -\langle u-v, u-v\rangle \\ &= -\|u\|^2 + \langle u, v\rangle + \langle v, u\rangle - \|v\|^2\end{aligned}$$

and

$$\begin{aligned}i\|u+iv\|^2 &= i\langle u+iv, u+iv\rangle \\ &= i\|u\|^2 + \langle u, v\rangle - \langle v, u\rangle + i\|v\|^2\end{aligned}$$

and

$$\begin{aligned}-i\|u-iv\|^2 &= -i\langle u-iv, u-iv\rangle \\ &= -i\|u\|^2 + \langle u, v\rangle - \langle v, u\rangle - i\|v\|^2.\end{aligned}$$

Adding the four equations, we have

$$\|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2 = 4\langle u, v\rangle,$$

as desired.

8. A norm on a vector space U is a function $\|\ \|: U \to [0, \infty)$ such that $\|u\| = 0$ if and only if $u = 0$, $\|\alpha u\| = |\alpha|\|u\|$ for all $\alpha \in \mathbf{F}$ and all $u \in U$, and $\|u+v\| \le \|u\| + \|v\|$ for all $u, v \in U$. Prove that a norm satisfying the parallelogram equality comes from an inner product (in other words, show that if $\|\ \|$ is a norm on U satisfying the parallelogram equality, then there is an inner product $\langle\ ,\ \rangle$ on U such that $\|u\| = \langle u, u\rangle^{1/2}$ for all $u \in U$).

COMMENT: This is among the hardest exercises in the book. Instructors may want to simplify this exercise slightly by allowing students to consider only the case where $\mathbf{F} = \mathbf{R}$.

SOLUTION: Suppose that U is a vector space and $\| \ \|$ is a norm on U satisfying the parallelogram equality. We want to find an inner product $\langle \ , \ \rangle$ on U such that $\|u\| = \langle u, u\rangle^{1/2}$ for all $u \in U$.

First consider the case where $\mathbf{F} = \mathbf{R}$. For $u, v \in U$, define $\langle u, v\rangle$ by

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}.$$

This definition is motivated by Exercise 6 of this chapter, which gives a formula for the inner product in terms of norms.

For $u \in U$ we have

$$\begin{aligned}\langle u, u\rangle &= \frac{\|u+u\|^2 - \|u-u\|^2}{4}\\ &= \frac{\|2u\|^2 - \|0\|^2}{4}\\ &= \|u\|^2.\end{aligned}$$

Thus $\|u\| = \langle u, u\rangle^{1/2}$, as desired. However, we still must show that $\langle \ , \ \rangle$ satisfies the properties required of an inner product.

Because $\langle u, u\rangle = \|u\|^2$ (as shown above), we have $\langle u, u\rangle \geq 0$ for all $u \in U$, with equality if and only if $u = 0$; these properties follow from the properties of a norm. Thus $\langle \ , \ \rangle$ satisfies the positivity and definiteness properties required of an inner product.

To prove that $\langle \ , \ \rangle$ is additive in the first slot, let $u, v, w \in U$. Then

$$\begin{aligned}&4(\langle u+v, w\rangle - \langle u, w\rangle - \langle v, w\rangle)\\ &\quad= \|u+v+w\|^2 - \|u+v-w\|^2 - \|u+w\|^2\\ &\qquad+ \|u-w\|^2 - \|v+w\|^2 + \|v-w\|^2\\ &\quad= \|u+v+w\|^2 + (\|u-w\|^2 + \|v-w\|^2)\\ &\qquad- \|u+v-w\|^2 - (\|u+w\|^2 + \|v+w\|^2),\end{aligned}$$

where the first equality comes from the definition of $\langle \ , \ \rangle$. In the last equality above, the parentheses indicate groupings to which we will apply the parallelogram equality, which asserts that the sum of the norms squared of two vectors $x, y \in U$ can be computed from the formula

$$\|x\|^2 + \|y\|^2 = \frac{\|x+y\|^2}{2} + \frac{\|x-y\|^2}{2}.$$

Applying the parallelogram equality to the two terms in parentheses above (take $x = u - w, y = v - w$ for the first sum in parentheses, then $x = u + w, y = v + w$ for the second term in parentheses) gives

$$\begin{aligned}
&4(\langle u+v, w\rangle - \langle u, w\rangle - \langle v, w\rangle) \\
&= \|u+v+w\|^2 + \frac{\|u+v-2w\|^2}{2} + \frac{\|u-v\|^2}{2} \\
&\quad - \|u+v-w\|^2 - \frac{\|u+v+2w\|^2}{2} - \frac{\|u-v\|^2}{2} \\
&= (\|u+v+w\|^2 + \|w\|^2) + \frac{\|u+v-2w\|^2}{2} \\
&\quad - (\|u+v-w\|^2 + \|w\|^2) - \frac{\|u+v+2w\|^2}{2}.
\end{aligned}$$

Applying the parallelogram equality to the two terms in parentheses above (take $x = u + v + w, y = w$ for the first sum in parentheses, then $x = u + v - w, y = w$ for the second term in parentheses) gives

$$\begin{aligned}
&4(\langle u+v, w\rangle - \langle u, w\rangle - \langle v, w\rangle) \\
&= \frac{\|u+v+2w\|^2}{2} + \frac{\|u+v\|^2}{2} + \frac{\|u+v-2w\|^2}{2} \\
&\quad - \frac{\|u+v\|^2}{2} - \frac{\|u+v-2w\|^2}{2} - \frac{\|u+v+2w\|^2}{2} \\
&= 0.
\end{aligned}$$

Thus

$$\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle,$$

completing the proof that $\langle\ ,\ \rangle$ is additive in the first slot.

To prove that $\langle\ ,\ \rangle$ is homogeneous in the first slot, let $u, v \in U$. If n is a positive integer, then

$$\begin{aligned}\langle nu, v\rangle &= \langle \underbrace{u + \cdots + u}_{n \text{ times}}, v\rangle \\ &= \underbrace{\langle u, v\rangle + \cdots + \langle u, v\rangle}_{n \text{ times}} \\ &= n\langle u, v\rangle,\end{aligned}$$

where the second equality comes from additivity in the first slot, which we have already verified. Replacing u with u/n in the equality above gives $\langle u, v\rangle = n\langle u/n, v\rangle$, which implies that

$$\langle \frac{u}{n}, v\rangle = \frac{1}{n}\langle u, v\rangle.$$

Let m be another positive integer, and replace u with mu in the equality above, getting

$$\begin{aligned}\langle \frac{m}{n}u, v\rangle &= \frac{1}{n}\langle mu, v\rangle \\ &= \frac{m}{n}\langle u, v\rangle,\end{aligned}$$

where the second equality holds because we have already shown that $\langle\ ,\ \rangle$ is homogeneous in the first slot with respect to positive integers. We have now shown that $\langle\ ,\ \rangle$ is homogeneous in the first slot with respect to positive rational numbers.

From the definition of $\langle\ ,\ \rangle$, we have

$$\begin{aligned}\langle -u, v\rangle &= \frac{\|-u + v\|^2 - \|-u - v\|^2}{4} \\ &= -\frac{\|u + v\|^2 - \|u - v\|^2}{4} \\ &= -\langle u, v\rangle.\end{aligned}$$

Combining this with the result from the previous paragraph, we can now conclude that $\langle\ ,\ \rangle$ is homogeneous in the first slot with respect to all rational numbers.

Now suppose that $\lambda \in \mathbf{R}$. There exists a sequence $r_1, r_2, \ldots$ of rational numbers such that $\lim_{n\to\infty} r_n = \lambda$. Thus

$$\begin{aligned}\lambda\langle u,v\rangle &= \lim_{n\to\infty} r_n\langle u,v\rangle\\ &= \lim_{n\to\infty}\langle r_nu,v\rangle\\ &= \lim_{n\to\infty}\frac{\|r_nu+v\|^2-\|r_nu-v\|^2}{4}.\end{aligned}$$

In the next paragraph we will show that $\lim_{n\to\infty}\|r_nu+v\| = \|\lambda u+v\|$ and $\lim_{n\to\infty}\|r_nu-v\| = \|\lambda u-v\|$. Combining this with the last equation above we can conclude that

$$\begin{aligned}\lambda\langle u,v\rangle &= \frac{\|\lambda u+v\|^2-\|\lambda u-v\|^2}{4}\\ &= \langle\lambda u,v\rangle,\end{aligned}$$

which will complete the proof that $\langle\ ,\ \rangle$ is homogeneous in the first slot.

If $x,y\in U$, then

$$\begin{aligned}\|x\| &= \|y+(x-y)\|\\ &\le \|y\|+\|x-y\|\end{aligned}$$

and thus

$$\|x\|-\|y\|\le\|x-y\|.$$

Interchanging the roles of x and y, we get

$$\|y\|-\|x\|\le\|x-y\|.$$

Because $\big|\|x\|-\|y\|\big|$ equals $\|x\|-\|y\|$ or $\|y\|-\|x\|$, we can now conclude that

$$\big|\|x\|-\|y\|\big|\le\|x-y\|.$$

With $x=r_nu+v$ and $y=\lambda u+v$, this inequality gives

$$\begin{aligned}\big|\|r_nu+v\|-\|\lambda u+v\|\big| &\le \|r_nu-\lambda u\|\\ &= |r_n-\lambda|\|u\|.\end{aligned}$$

Because $\lim_{n\to\infty} r_n=\lambda$, this shows that

$$\lim_{n\to\infty}\|r_nu+v\| = \|\lambda u+v\|.$$

Replacing v with $-v$, we have

$$\lim_{n \to \infty} \|r_n u - v\| = \|\lambda u - v\|.$$

The last two equations are the promised ingredients that were needed for the proof that $\langle\ ,\ \rangle$ is homogeneous in the first slot.

Finally, we must show that $\langle u, v\rangle = \langle v, u\rangle$ (recall that we are considering the case where $\mathbf{F} = \mathbf{R}$). This last step is easy:

$$\begin{aligned} \langle u, v\rangle &= \frac{\|u+v\|^2 - \|u-v\|^2}{4} \\ &= \frac{\|v+u\|^2 - \|v-u\|^2}{4} \\ &= \langle v, u\rangle. \end{aligned}$$

This completes the proof that $\langle\ ,\ \rangle$ is an inner product when $\mathbf{F} = \mathbf{R}$. Whew!

Now consider the case where $\mathbf{F} = \mathbf{C}$. For $u, v \in U$, define $\langle u, v\rangle$ by

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i}{4}.$$

This definition is motivated by Exercise 7 of this chapter, which gives a formula for the inner product in terms of norms.

For $u \in U$ we have

$$\begin{aligned} \langle u, u\rangle &= \frac{\|u+u\|^2 - \|u-u\|^2 + \|u+iu\|^2 i - \|u-iu\|^2 i}{4} \\ &= \frac{\|2u\|^2 + |1+i|^2\|u\|^2 i - |1-i|^2\|u\|^2 i}{4} \\ &= \frac{4\|u\|^2 + 2\|u\|^2 i - 2\|u\|^2 i}{4} \\ &= \|u\|^2. \end{aligned}$$

Thus $\|u\| = \langle u, u\rangle^{1/2}$, as desired. However, we still must show that $\langle\ ,\ \rangle$ satisfies the properties required of an inner product.

Because $\langle u, u\rangle = \|u\|^2$ (as shown above), we have $\langle u, u\rangle \geq 0$ for all $u \in U$, with equality if and only if $u = 0$; these properties follow from the properties of a norm. Thus $\langle\ ,\ \rangle$ satisfies the positivity and definiteness properties required of an inner product.

For convenience, let's define $\langle\ ,\ \rangle_{\mathbf{R}}$ by

$$\langle u, v\rangle_{\mathbf{R}} = \frac{\|u+v\|^2 - \|u-v\|^2}{4}.$$

Here the subscript $\mathbf{R}$ reminds us that $\langle\, , \rangle_{\mathbf{R}}$ was the inner product we defined when considering the case $\mathbf{F} = \mathbf{R}$. Now we are assuming that $\mathbf{F} = \mathbf{C}$, but $\langle\, , \rangle_{\mathbf{R}}$ is still well defined. Note that

$$\langle u, v\rangle = \langle u, v\rangle_{\mathbf{R}} + \langle u, iv\rangle_{\mathbf{R}} i.$$

We have already proved that $\langle\, , \rangle_{\mathbf{R}}$ is additive in the first slot, and now we use that information. Let $u, v, w \in U$. Then

$$\begin{aligned}
\langle u+v, w\rangle &= \langle u+v, w\rangle_{\mathbf{R}} + \langle u+v, iw\rangle_{\mathbf{R}} i \\
&= \langle u, w\rangle_{\mathbf{R}} + \langle v, w\rangle_{\mathbf{R}} + \langle u, iw\rangle_{\mathbf{R}} i + \langle v, iw\rangle_{\mathbf{R}} i \\
&= (\langle u, w\rangle_{\mathbf{R}} + \langle u, iw\rangle_{\mathbf{R}} i) + (\langle v, w\rangle_{\mathbf{R}} + \langle v, iw\rangle_{\mathbf{R}} i) \\
&= \langle u, w\rangle + \langle v, w\rangle.
\end{aligned}$$

Thus $\langle\, , \rangle$ is additive in the first slot.

To prove that $\langle\, , \rangle$ is homogeneous in the first slot, let $u, v \in U$. If $\lambda \in \mathbf{R}$, then

$$\begin{aligned}
\langle \lambda u, v\rangle &= \langle \lambda u, v\rangle_{\mathbf{R}} + \langle \lambda u, iv\rangle_{\mathbf{R}} i \\
&= \lambda\langle u, v\rangle_{\mathbf{R}} + \lambda\langle u, iv\rangle_{\mathbf{R}} i \\
&= \lambda\langle u, v\rangle,
\end{aligned}$$

where we have used the homogeneity of $\langle\, , \rangle_{\mathbf{R}}$ in the first slot. The last equation above shows that $\langle\, , \rangle$ is homogeneous in the first slot with respect to all real numbers. We must still extend this result to complex numbers.

Note that

$$\begin{aligned}
\langle iu, v\rangle &= \frac{\|iu+v\|^2 - \|iu-v\|^2 + \|iu+iv\|^2 i - \|iu-iv\|^2 i}{4} \\
&= \frac{\|i(u+v)\|^2 i - \|i(u-v)\|^2 i - \|i(u+iv)\|^2 + \|i(u-iv)\|^2}{4} \\
&= \frac{\|u+v\|^2 i - \|u-v\|^2 i - \|u+iv\|^2 + \|u-iv\|^2}{4} \\
&= i\langle u, v\rangle.
\end{aligned}$$

Combining this result with additivity and homogeneity with respect to real numbers, we get that

$$\langle (a+bi)u, v\rangle = (a+bi)\langle u, v\rangle$$

for all $a, b \in \mathbf{R}$. In other words, $\langle\ ,\ \rangle$ is homogeneous in the first slot with respect to all complex numbers.

Finally, we must show that $\langle u, v\rangle = \overline{\langle v, u\rangle}$. This last step is easy:

$$
\begin{aligned}
\langle u, v\rangle &= \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i}{4} \\
&= \frac{\|v+u\|^2 - \|v-u\|^2 + \|i(-iu+v)\|^2 i - \|(-i)(iu+v)\|^2 i}{4} \\
&= \frac{\overline{\|v+u\|^2 - \|v-u\|^2 + \|v+iu\|^2 i - \|v-iu\|^2 i}}{4} \\
&= \overline{\langle v, u\rangle}.
\end{aligned}
$$

This completes the proof that $\langle\ ,\ \rangle$ is an inner product when $\mathbf{F} = \mathbf{C}$.

9. Suppose n is a positive integer. Prove that

$$
\Bigl(\frac{1}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}\Bigr)
$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$
\langle f, g\rangle = \int_{-\pi}^{\pi} f(x) g(x)\, dx.
$$

COMMENT: This orthonormal list is often used for modeling periodic phenomena such as tides.

SOLUTION: First we need to show that each element of the list above has norm 1. This follows easily from the following formulas:

$$
\begin{aligned}
\int (\sin jt)^2\, dt &= \frac{2jt - \sin 2jt}{4j} \\
\int (\cos jt)^2\, dt &= \frac{2jt + \sin 2jt}{4j}.
\end{aligned}
$$

Next we need to show that any two distinct elements of the list above are orthogonal. This follows easily from the following formulas, valid when $j \neq k$:

$$\int (\sin jt)(\sin kt)\,dt = \frac{j\sin(j-k)t + k\sin(j-k)t - j\sin(j+k)t + k\sin(j+k)t}{2(j-k)(j+k)}$$

$$\int (\sin jt)(\cos kt)\,dt = \frac{j\cos(j-k)t + k\cos(j-k)t + j\cos(j+k)t - k\cos(j+k)t}{2(k-j)(j+k)}$$

$$\int (\cos jt)(\cos kt)\,dt = \frac{j\sin(j-k)t + k\sin(j-k)t + j\sin(j+k)t - k\sin(j+k)t}{2(j-k)(j+k)}$$

$$\int (\sin jt)(\cos jt)\,dt = -\frac{(\cos jt)^2}{2j}.$$

10. On $\mathcal{P}_2(\mathbf{R})$, consider the inner product given by

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx.$$

Apply the Gram-Schmidt procedure to the basis $(1, x, x^2)$ to produce an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$.

SOLUTION: Applying the Gram-Schmidt procedure to $(1, x, x^2)$ produces (using elementary calculus and some arithmetic) the following orthonormal basis of $\mathcal{P}_2(\mathbf{R})$:

$$\bigl(1, \sqrt{3}(-1+2x), \sqrt{5}(1-6x+6x^2)\bigr).$$

11. What happens if the Gram-Schmidt procedure is applied to a list of vectors that is not linearly independent?

SOLUTION: Suppose $(v_1, \dots, v_m)$ is a linearly dependent list of vectors in V.

If $v_1 = 0$, then at the first step of applying the Gram-Schmidt procedure to $(v_1, \dots, v_m)$ we will be dividing by 0 when trying to set $e_1 = v_1/\|v_1\|$.

If $v_1 \neq 0$, then by the linear dependence lemma (2.4), some v_j is in $\operatorname{span}(v_1, \dots, v_{j-1})$; here we choose j to be the smallest positive integer with this property. If we apply the Gram-Schmidt procedure to produce $(e_1, \dots, e_{j-1})$ at the end of step j, then

$$\operatorname{span}(v_1, \dots, v_{j-1}) = \operatorname{span}(e_1, \dots, e_{j-1}).$$

Thus $v_j \in \operatorname{span}(e_1, \dots, e_{j-1})$. By 6.17, this implies that

$$v_j = \langle v_j, e_1 \rangle e_1 + \cdots + \langle v_j, e_{j-1} \rangle e_{j-1}.$$

Thus the Gram-Schmidt formula 6.23 for e_j includes a division by 0, which is not allowed.

12. Suppose V is a real inner-product space and $(v_1, \dots, v_m)$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $(e_1, \dots, e_m)$ of vectors in V such that

$$\operatorname{span}(v_1, \dots, v_j) = \operatorname{span}(e_1, \dots, e_j)$$

for all $j \in \{1, \dots, m\}$.

SOLUTION: For $j = 1$, the condition above states that

$$\operatorname{span}(v_1) = \operatorname{span}(e_1).$$

Because there are only two vectors in $\operatorname{span}(v_1)$ with norm 1 (these two vectors are $v_1/\|v_1\|$ and $-v_1/\|v_1\|$), we have only these two choices for e_1.

Now suppose that $j > 1$ and that an orthonormal list $(e_1, \dots, e_{j-1})$ has been chosen such that

$$\operatorname{span}(v_1, \dots, v_{j-1}) = \operatorname{span}(e_1, \dots, e_{j-1}).$$

The Gram-Schmidt procedure produces $e_j \in V$ such that $(e_1, \dots, e_j)$ is an orthonormal list and

$$\operatorname{span}(v_1, \dots, v_j) = \operatorname{span}(e_1, \dots, e_j).$$

Suppose $e_j{}' \in V$ is another vector with these properties, meaning that $(e_1, \dots, e_{j-1}, e_j{}')$ is an orthonormal list and

$$\operatorname{span}(v_1, \dots, v_j) = \operatorname{span}(e_1, \dots, e_{j-1}, e_j{}').$$

The last two equations show that $\operatorname{span}(e_1, \dots, e_{j-1}, e_j{}') = \operatorname{span}(e_1, \dots, e_j)$. In particular, $e_j{}' \in \operatorname{span}(e_1, \dots, e_j)$, which implies that

$$\begin{aligned} e_j{}' &= \langle e_j{}', e_1 \rangle e_1 + \cdots + \langle e_j{}', e_j \rangle e_j \\ &= \langle e_j{}', e_j \rangle e_j, \end{aligned}$$

where the first equality comes from 6.17 (with $\operatorname{span}(e_1, \dots, e_j)$ replacing V and j replacing n) and the second equality holds because $(e_1, \dots, e_{j-1}, e_j{}')$

is an orthonormal list. Taking norms of both sides of the last equation, and recalling that e_j and e_j' both have norm 1, we see that $|\langle e_j', e_j\rangle| = 1$. Thus $\langle e_j', e_j\rangle = 1$ or $\langle e_j', e_j\rangle = -1$. Hence the last equation above implies that $e_j' = e_j$ or $e_j' = -e_j$.

We have shown that there are exactly two possible choices for each e_j. As j ranges from 1 to m, this gives us exactly 2^m possible choices for $(e_1, \dots, e_m)$.

13. Suppose $(e_1, \dots, e_m)$ is an orthonormal list of vectors in V. Let $v \in V$. Prove that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \dots + |\langle v, e_m\rangle|^2$$

if and only if $v \in \operatorname{span}(e_1, \dots, e_m)$.

SOLUTION: Extend $(e_1, \dots, e_m)$ to an orthonormal basis $(e_1, \dots, e_n)$ of V. Then

$$v = \langle v, e_1\rangle e_1 + \dots + \langle v, e_n\rangle e_n$$

and

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \dots + |\langle v, e_n\rangle|^2;$$

see 6.17. From the last equation, we see that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \dots + |\langle v, e_m\rangle|^2$$

if and only if $\langle v, e_{m+1}\rangle = \dots = \langle v, e_n\rangle = 0$. From the first equation above, this happens if and only if

$$v = \langle v, e_1\rangle e_1 + \dots + \langle v, e_m\rangle e_m,$$

which happens if and only if $v \in \operatorname{span}(e_1, \dots, e_m)$.

14. Find an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ (with inner product as in Exercise 10) such that the differentiation operator (the operator that takes p to p') on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to this basis.

SOLUTION: Because $1' = 0$, $x' = 1$, and $(x^2)' = 2x$, the differentiation operator on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to the basis $(1, x, x^2)$. However, $(1, x, x^2)$ is not an orthonormal basis. But, as can be seen from the proof of 6.27, if the Gram-Schmidt procedure is applied to this basis, we will get an orthonormal basis with respect to which the differentiation operator has an upper-triangular matrix. As we saw in Exercise 10 of this chapter, the Gram-Schmidt procedure applied to $(1, x, x^2)$ gives

$$\big(1, \sqrt{3}(-1+2x), \sqrt{5}(1-6x+6x^2)\big).$$

which is our desired orthonormal basis of $\mathcal{P}_2(\mathbf{R})$.

15. Suppose U is a subspace of V. Prove that

$$\dim U^\perp = \dim V - \dim U.$$

SOLUTION: From 6.29, we know that

$$V = U \oplus U^\perp.$$

Thus by Exercise 17 in Chapter 2, we have

$$\dim V = \dim U + \dim U^\perp,$$

which implies that $\dim U^\perp = \dim V - \dim U$.

16. Suppose U is a subspace of V. Prove that $U^\perp = \{0\}$ if and only if $U = V$.

SOLUTION: From 6.29, we know that

$$V = U \oplus U^\perp.$$

This clearly implies that $U^\perp = \{0\}$ if and only if $U = V$.

17. Prove that if $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in $\operatorname{null} P$ is orthogonal to every vector in $\operatorname{range} P$, then P is an orthogonal projection.

SOLUTION: Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in $\operatorname{null} P$ is orthogonal to every vector in $\operatorname{range} P$. Let $U = \operatorname{range} P$. We will show that P equals the orthogonal projection P_U. To do this, suppose $v \in V$. Then

$$v = Pv + (v - Pv).$$

Clearly $Pv \in \operatorname{range} P = U$. Also, $P(v - Pv) = Pv - P^2v = 0$, which means that $v - Pv \in \operatorname{null} P$. Thus $v - Pv$ is orthogonal to every vector in U. In other words, $v - Pv \in U^\perp$. Thus the equation above writes v as the sum of a vector in U and a vector in $U^\perp$. In this decomposition, the vector in U equals, by definition, $P_U v$. Hence $Pv = P_U v$, as desired.

18. Prove that if $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and

$$\|Pv\| \leq \|v\|$$

for every $v \in V$, then P is an orthogonal projection.

SOLUTION: Suppose $u \in \operatorname{range} P$ and $w \in \operatorname{null} P$. If we can show that $\langle u, w \rangle = 0$, then by the previous exercise we can conclude that P is an orthogonal projection.

Because $u \in \operatorname{range} P$, there exists $u' \in V$ such that

$$u = Pu'.$$

Applying P to both sides of this equation, we have

$$\begin{aligned} Pu &= P^2 u' \\ &= Pu' \\ &= u. \end{aligned}$$

Because $w \in \operatorname{null} P$, this implies that

$$P(u + aw) = u$$

for every $a \in \mathbf{F}$. Thus

$$\begin{aligned} \|u\|^2 &= \|P(u + aw)\|^2 \\ &\leq \|u + aw\|^2 \end{aligned}$$

for every $a \in \mathbf{F}$, where the second line follows from our hypothesis that $\|Pv\| \leq \|v\|$ for every $v \in V$. The inequality above implies (see Exercise 2 of this chapter) that $\langle u, w \rangle = 0$, as desired.

19. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that U is invariant under T if and only if $P_U T P_U = T P_U$.

SOLUTION: First suppose that U is invariant under T. Let $v \in V$. Then $P_U v \in U$ and hence $T(P_U v) \in U$ (because U is invariant under T). Thus $P_U\big(T(P_U v)\big) = T(P_U v)$. Because v was an arbitrary vector in V, this implies that $P_U T P_U = T P_U$, as desired.

To prove the implication in the other direction, now suppose that

$$P_U T P_U = T P_U.$$

Suppose $u \in U$. Then $P_U u = u$, so applying both sides of the equation above to u gives $P_U(Tu) = Tu$, which implies that $Tu \in U$. Because u was an arbitrary vector in U, this implies that T is invariant under U, as desired.

20. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that U and $U^\perp$ are both invariant under T if and only if $P_U T = T P_U$.

SOLUTION: First suppose that U and $U^\perp$ are both invariant under T. By the previous exercise, this implies that

$$P_U T P_U = T P_U$$

and

$$P_{U^\perp} T P_{U^\perp} = T P_{U^\perp}.$$

But $P_{U^\perp} = I - P_U$, so the last equation becomes

$$(I - P_U)T(I - P_U) = T(I - P_U).$$

Expanding both sides of the equation above and rearranging terms, we get

$$P_U T P_U = P_U T.$$

Combining this with the first equation above, we get $P_U T = T P_U$, as desired.

To prove the implication in the other direction, suppose now that

$$P_U T = T P_U.$$

Then

$$\begin{aligned} P_U T P_U &= (P_U T) P_U \\ &= (T P_U) P_U \\ &= T {P_U}^2 \\ &= T P_U, \end{aligned}$$

which implies (by the previous exercise) that U is invariant under T, as desired. Also,

$$\begin{aligned} P_{U^\perp} T P_{U^\perp} &= \big((I - P_U)T\big) P_{U^\perp} \\ &= (T - P_U T) P_{U^\perp} \\ &= (T - T P_U) P_{U^\perp} \\ &= T(1 - P_U) P_{U^\perp} \\ &= T {P_{U^\perp}}^2 \\ &= T P_{U^\perp}, \end{aligned}$$

which implies (by the previous exercise) that $U^\perp$ is invariant under T, as desired.

21. In $\mathbf{R}^4$, let

$$U = \operatorname{span}\big((1,1,0,0),(1,1,1,2)\big).$$

Find $u \in U$ such that $\|u - (1,2,3,4)\|$ is as small as possible.

SOLUTION: First we find an orthonormal basis of U by applying the Gram-Schmidt procedure to $\big((1,1,0,0),(1,1,1,2)\big)$, getting

$$e_1 = \Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\Big)$$

$$e_2 = \Big(0, 0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\Big).$$

Thus with e_1, e_2 as above, (e_1, e_2) is an orthonormal basis of U. By 6.36 and 6.35, the closest point $u \in U$ to $(1,2,3,4)$ is

$$\langle(1,2,3,4), e_1\rangle e_1 + \langle(1,2,3,4), e_2\rangle e_2,$$

which equals

$$\Big(\frac{3}{2}, \frac{3}{2}, \frac{11}{5}, \frac{22}{5}\Big).$$

22. Find $p \in \mathcal{P}_3(\mathbf{R})$ such that $p(0) = 0$, $p'(0) = 0$, and

$$\int_0^1 |2 + 3x - p(x)|^2\, dx$$

is as small as possible.

SOLUTION: Define an inner product on $\mathcal{P}_3(\mathbf{R})$ by

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\, dx.$$

Let $q(x) = 2 + 3x$, and let

$$U = \{p \in \mathcal{P}_3(\mathbf{R}) : p(0) = 0, p'(0) = 0\}.$$

With this notation, our problem is to find the closest point $p \in U$ to q. To do this, first we find an orthonormal basis of U.

A polynomial p satisfying $p(0) = 0$, $p'(0) = 0$ has constant term 0 and first degree term also equal to 0. Thus a basis of U is

$$(x^2, x^3).$$

Apply the Gram-Schmidt procedure to this basis, getting

$$e_1 = \sqrt{5}x^2$$

$$e_2 = \sqrt{7}(-5x^2 + 6x^3).$$

Thus with e_1, e_2 as above, (e_1, e_2) is an orthonormal basis of U. By 6.36 and 6.35, the closest point $p \in U$ to q is given by the formula

$$p = \langle q, e_1\rangle e_1 + \langle q, e_2\rangle e_2.$$

A short computation now shows that

$$p(x) = 24x^2 - \frac{203}{10}x^3.$$

23. Find $p \in \mathcal{P}_5(\mathbf{R})$ that makes

$$\int_{-\pi}^{\pi} |\sin x - p(x)|^2\, dx$$

as small as possible. (The polynomial 6.40 is an excellent approximation to the answer to this exercise, but here you are asked to find the exact solution, which involves powers of π. A computer that can perform symbolic integration will be useful.)

SOLUTION: Let $C[-\pi, \pi]$ denote the real vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\, dx.$$

Let $v \in C[-\pi, \pi]$ be the function defined by $v(x) = \sin x$. Let U denote the subspace of $C[-\pi, \pi]$ consisting of the polynomials with real coefficients and degree at most 5. We need to find $p \in U$ such that $\|v - p\|$ is as small as possible.

First find an orthonormal basis of U by applying the Gram-Schmidt procedure (using the inner product above) to the basis $(1, x, x^2, x^3, x^4, x^5)$ of U, producing the orthonormal basis $(e_1, e_2, e_3, e_4, e_5, e_6)$, where

$$e_1 = \frac{1}{\sqrt{2\pi}},$$

$$e_2 = \frac{\sqrt{\frac{3}{2}}x}{\pi^{3/2}},$$

$$e_3 = -\frac{\sqrt{\frac{5}{2}}(\pi^2 - 3x^2)}{2\pi^{5/2}},$$

$$e_4 = -\frac{\sqrt{\frac{7}{2}}(3\pi^2 x - 5x^3)}{2\pi^{7/2}},$$

$$e_5 = \frac{3(3\pi^4 - 30\pi^2 x^2 + 35x^4)}{8\sqrt{2}\pi^{9/2}},$$

$$e_6 = -\frac{\sqrt{\frac{11}{2}}(15\pi^4 x - 70\pi^2 x^3 + 63x^5)}{8\pi^{11/2}}.$$

Now compute $P_U v$ using 6.35 (with $m = 6$), getting

$$P_U v = \frac{105(1485 - 153\pi^2 + \pi^4)}{8\pi^6}x - \frac{315(1155 - 125\pi^2 + \pi^4)}{4\pi^8}x^3$$
$$+ \frac{693(945 - 105\pi^2 + \pi^4)}{8\pi^{10}}x^5.$$

Finally, 6.36 and the discussion following 6.42 show that the function above is the one we seek.

24. Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$p(\tfrac{1}{2}) = \int_0^1 p(x)q(x)\,dx$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

SOLUTION: We will need an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$, where the inner product of two polynomials in $\mathcal{P}_2(\mathbf{R})$ is defined to be the integral from 0 to 1 of the product of the two polynomials. An orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ was already computed in Exercise 10 of this chapter. Specifically, let

$$e_1(x) = 1$$
$$e_2(x) = \sqrt{3}(-1 + 2x)$$
$$e_3(x) = \sqrt{5}(1 - 6x + 6x^2).$$

Then (e_1, e_2, e_3) is an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$.

Define a linear functional φ on $\mathcal{P}_2(\mathbf{R})$ by

$$\varphi(p) = p(\tfrac{1}{2}).$$

We seek $q \in \mathcal{P}_2(\mathbf{R})$ such that $\varphi(p) = \langle p, q\rangle$ for every $p \in \mathcal{P}_2(\mathbf{R})$. By the formula given in the proof of 6.45, we have

$$q = \varphi(e_1)e_1 + \varphi(e_2)e_2 + \varphi(e_3)e_3.$$

Evaluate the right side of the equation above to get

$$q(x) = -\frac{3}{2} + 15x - 15x^2.$$

25. Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

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

for every $p \in \mathcal{P}_2(\mathbf{R})$.

SOLUTION: Define a linear functional φ on $\mathcal{P}_2(\mathbf{R})$ by

$$\varphi(p) = \int_0^1 p(x)(\cos \pi x)\, dx.$$

We seek $q \in \mathcal{P}_2(\mathbf{R})$ such that $\varphi(p) = \langle p, q\rangle$ for every $p \in \mathcal{P}_2(\mathbf{R})$, where the inner product on $\mathcal{P}_2(\mathbf{R})$ is defined as in the previous exercise. Letting e_1, e_2, e_3 be as in the previous exercise, but using our new definition of φ, we again have

$$q = \varphi(e_1)e_1 + \varphi(e_2)e_2 + \varphi(e_3)e_3.$$

Evaluate the right side of the equation above to get

$$q(x) = \frac{12 - 24x}{\pi^2}.$$

26. Fix a vector $v \in V$ and define $T \in \mathcal{L}(V, \mathbf{F})$ by $Tu = \langle u, v\rangle$. For $a \in \mathbf{F}$, find a formula for T^*a.

SOLUTION: Because $T \in \mathcal{L}(V, \mathbf{F})$, we know that $T^* \in \mathcal{L}(\mathbf{F}, V)$. Fix $a \in \mathbf{F}$. Then T^*a is the unique vector in V such that

(a) $$\langle Tu, a\rangle = \langle u, T^*a\rangle$$

for all $u \in U$. The inner product on the right is the inner product in V, but the inner product on the left is the usual inner product on $\mathbf{F}$: the product of the entry in the first slot with the complex conjugate of the entry in the second slot. Thus

$$\begin{aligned}\langle Tu, a\rangle &= (Tu)\bar{a}\\ &= \langle u, v\rangle\bar{a}\\ &= \langle u, av\rangle. \qquad \text{(b)}\end{aligned}$$

Comparing (a) with (b) gives

$$\langle u, T^*a\rangle = \langle u, av\rangle$$

for all $u \in U$. Thus $T^*a = av$.

27. Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by

$$T(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1}).$$

Find a formula for $T^*(z_1, \dots, z_n)$.

SOLUTION: Fix $(z_1, \dots, z_n) \in \mathbf{F}^n$. Then for every $(w_1, \dots, w_n) \in \mathbf{F}^n$, we have

$$\begin{aligned}\langle (w_1, \dots, w_n), T^*(z_1, \dots, z_n)\rangle &= \langle T(w_1, \dots, w_n), (z_1, \dots, z_n)\rangle\\ &= \langle (0, w_1, \dots, w_{n-1}), (z_1, \dots, z_n)\rangle\\ &= w_1\overline{z_2} + \cdots + w_{n-1}\overline{z_n}\\ &= \langle (w_1, \dots, w_n), (z_2, \dots, z_n, 0)\rangle.\end{aligned}$$

Thus

$$T^*(z_1, \dots, z_n) = (z_2, \dots, z_n, 0).$$

28. Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Prove that λ is an eigenvalue of T if and only if $\bar{\lambda}$ is an eigenvalue of T^*.

SOLUTION: We have

$$\begin{aligned}
\lambda \text{ is an not eigenvalue of } T &\iff T - \lambda I \text{ is invertible}\\
&\iff S(T-\lambda I) = (T-\lambda I)S = I\\
&\qquad \text{for some } S \in \mathcal{L}(V)\\
&\iff (T-\lambda I)^* S^* = S^*(T-\lambda I)^* = I\\
&\qquad \text{for some } S \in \mathcal{L}(V)\\
&\iff (T-\lambda I)^* \text{ is invertible}\\
&\iff T^* - \bar{\lambda} I \text{ is invertible}\\
&\iff \bar{\lambda} \text{ is not an eigenvalue of } T^*.
\end{aligned}$$

Thus λ is an eigenvalue of T if and only if $\bar{\lambda}$ is an eigenvalue of T^*.

29. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that U is invariant under T if and only if $U^\perp$ is invariant under T^*.

SOLUTION: First suppose that U is invariant under T. To prove that $U^\perp$ is invariant under T^*, let $v \in U^\perp$. We need to show that $T^*v \in U^\perp$. But

$$\begin{aligned}
\langle u, T^*v\rangle &= \langle Tu, v\rangle\\
&= 0
\end{aligned}$$

for every $u \in U$ (because if $u \in U$, then $Tu \in U$ and hence Tu is orthogonal to v, an element of $U^\perp$). Thus $T^*v \in U^\perp$, and hence $U^\perp$ is invariant under T^*, as desired.

To prove the other direction, now suppose that $U^\perp$ is invariant under T^*. Then by the first direction, we know that $(U^\perp)^\perp$ is invariant under $(T^*)^*$. But $(U^\perp)^\perp = U$ (by 6.33) and $(T^*)^* = T$, so U is invariant under T, completing the proof.

30. Suppose $T \in \mathcal{L}(V, W)$. Prove that

(a) T is injective if and only if T^* is surjective;

(b) T is surjective if and only if T^* is injective.

SOLUTION: First we prove (a):

$$\begin{aligned}
T \text{ is injective } &\iff \operatorname{null} T = \{0\}\\
&\iff (\operatorname{range} T^*)^\perp = \{0\}\\
&\iff \operatorname{range} T^* = W\\
&\iff T^* \text{ is surjective },
\end{aligned}$$

where the second line comes from 6.46(c).

Now that (a) has been proved, (b) follows immediately by replacing T with T^* in (a).

31. Prove that

$$\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$$

and

$$\dim \operatorname{range} T^* = \dim \operatorname{range} T$$

for every $T \in \mathcal{L}(V, W)$.

SOLUTION: Let $T \in \mathcal{L}(V, W)$. Then

$$\begin{aligned}\dim \operatorname{null} T^* &= \dim(\operatorname{range} T)^\perp \\ &= \dim W - \dim \operatorname{range} T \\ &= \dim \operatorname{null} T + \dim W - \dim V,\end{aligned}$$

where the first equality comes from 6.46(a), the second equality comes from Exercise 15 of this chapter, and the third equality comes from 3.4. This proves the first equality that we seek.

To prove the second equality, note that

$$\begin{aligned}\dim \operatorname{range} T^* &= \dim W - \dim \operatorname{null} T^* \\ &= \dim V - \dim \operatorname{null} T \\ &= \dim \operatorname{range} T,\end{aligned}$$

where the first and third equalities come from 3.4 and the second equality comes from the first part of this exercise. This proves the second equality that we seek.

32. Suppose A is an m-by-n matrix of real numbers. Prove that the dimension of the span of the columns of A (in $\mathbf{R}^m$) equals the dimension of the span of the rows of A (in $\mathbf{R}^n$).

SOLUTION: Let $T \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ be such that the matrix of T (with respect to the standard bases) equals A. Then $\operatorname{range} T$ equals the span of the columns of A. Thus

$$
\begin{aligned}
\text{dimension of the span of the columns of } A
&= \dim \operatorname{range} T \\
&= \dim \operatorname{range} T^* \\
&= \text{dimension of the span of the columns of } \mathcal{M}(T^*) \\
&= \text{dimension of the span of the columns of the transpose of } A \\
&= \text{dimension of the span of the rows of } A,
\end{aligned}
$$

where the second equality comes from the previous exercise and the fourth equality comes from 6.47.

CHAPTER 7

Operators on Inner-Product Spaces

1. Make $\mathcal{P}_2(\mathbf{R})$ into an inner-product space by defining

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx.$$

Define $T \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ by $T(a_0 + a_1x + a_2x^2) = a_1x$.

(a) Show that T is not self-adjoint.

(b) The matrix of T with respect to the basis $(1, x, x^2)$ is

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

This matrix equals its conjugate transpose, even though T is not self-adjoint. Explain why this is not a contradiction.

SOLUTION: (a): Note that

$$\begin{aligned}\langle T1, x\rangle &= \langle 0, x\rangle \\ &= 0\end{aligned}$$

but

$$\begin{aligned}\langle 1, Tx\rangle &= \langle 1, x\rangle \\ &= \frac{1}{2}.\end{aligned}$$

Thus $\langle T1, x\rangle \neq \langle 1, Tx\rangle$, which shows that T is not self-adjoint.

(b): The result stating that the matrix of T^* is the conjugate transpose of the matrix of T has as a hypothesis that we are working with orthonormal bases (see 6.47). Because $(1, x, x^2)$ is not an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$, we cannot compute the matrix of T^* with respect to this basis by taking the conjugate transpose of the matrix of T.

2. Prove or give a counterexample: the product of any two self-adjoint operators on a finite-dimensional inner-product space is self-adjoint.

SOLUTION: Let $S, T \in \mathcal{L}(\mathbf{F}^2)$ be the operators whose matrices (with respect to the standard basis) are given by

$$\mathcal{M}(S) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \mathcal{M}(T) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Each of these matrices obviously equals its conjugate transpose, and hence S, T are self-adjoint. Now

$$\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T) = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}.$$

Because $\mathcal{M}(ST)$ does not equal its conjugate transpose, ST is not self-adjoint. Thus we have an example of two self-adjoint operators whose product is not self-adjoint.

Of course there are also many other examples.

COMMENT: Suppose $S, T \in \mathcal{L}(V)$ are self-adjoint. Then ST is self-adjoint if and only if $ST = TS$ (as is easy to see).

3. (a) Show that if V is a real inner-product space, then the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.

   (b) Show that if V is a complex inner-product space, then the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

SOLUTION: (a): Suppose V is a real inner-product space. Obviously the zero operator is self-adjoint. Furthermore, if $S, T \in \mathcal{L}(V)$ are self-adjoint, then

$$\begin{aligned}(S+T)^* &= S^* + T^* \\ &= S + T,\end{aligned}$$

and thus $S+T$ is self-adjoint. Finally, if $T \in \mathcal{L}(V)$ is self-adjoint and $a \in \mathbf{R}$, then

$$\begin{aligned}(aT)^* &= aT^* \\ &= aT,\end{aligned}$$

and thus aT is self-adjoint. We have shown that the set of self-adjoint operators on V contains the zero operator and that it is closed under addition and scalar multiplication. Thus the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.

(b): Suppose now that V is a complex vector space. The identity operator I is self-adjoint, but $(iI)^* = -iI$ so iI is not self-adjoint. Thus the set of self-adjoint operators on V is not closed under scalar multiplication and hence it is not a subspace of $\mathcal{L}(V)$.

4. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that P is an orthogonal projection if and only if P is self-adjoint.

SOLUTION: First suppose that P is an orthogonal projection. Thus there is a subspace U of V such that $P = P_U$. Suppose $v_1, v_2 \in V$. Write

$$v_1 = u_1 + w_1, \quad v_2 = u_2 + w_2,$$

where $u_1, u_2 \in U$ and $w_1, w_2 \in U^\perp$ (see 6.29). Now

$$\begin{aligned}\langle Pv_1, v_2\rangle &= \langle u_1, u_2 + w_2\rangle \\ &= \langle u_1, u_2\rangle + \langle u_1, w_2\rangle \\ &= \langle u_1, u_2\rangle \\ &= \langle u_1, u_2\rangle + \langle w_1, u_2\rangle \\ &= \langle u_1 + w_1, u_2\rangle \\ &= \langle v_1, Pv_2\rangle.\end{aligned}$$

Thus $P = P^*$, and hence P is self-adjoint.

To prove the implication in the other direction, now suppose that P is self-adjoint. Let $v \in V$. Because $P(v - Pv) = Pv - P^2v = 0$, we have

$$v - Pv \in \operatorname{null} P = (\operatorname{range} P^*)^\perp = (\operatorname{range} P)^\perp,$$

where the first equality comes from 6.46(c). Writing

$$v = Pv + (v - Pv),$$

we have $Pv \in \operatorname{range} P$ and $(v - Pv) \in (\operatorname{range} P)^{\perp}$. Thus $Pv = P_{\operatorname{range} P} v$. Because this holds for all $v \in V$, we have $P = P_{\operatorname{range} P}$, which shows that P is an orthogonal projection.

5. Show that if $\dim V \geq 2$, then the set of normal operators on V is not a subspace of $\mathcal{L}(V)$.

SOLUTION: Suppose $\dim V \geq 2$. Let $(e_1, \dots, e_n)$ be an orthonormal basis of V. Define $S, T \in \mathcal{L}(V)$ by

$$S(a_1 e_1 + \dots + a_n e_n) = a_2 e_1 - a_1 e_2$$

and

$$T(a_1 e_1 + \dots + a_n e_n) = a_2 e_1 + a_1 e_2.$$

A simple calculation verifies that

$$S^*(a_1 e_1 + \dots + a_n e_n) = -a_2 e_1 + a_1 e_2.$$

From this formula, another simple calculation shows that $SS^* = S^*S$. Yet another simple calculation shows that T is self-adjoint. Thus both S and T are normal. However, $S + T$ is given by the formula

$$(S + T)(a_1 e_1 + \dots + a_n e_n) = 2a_2 e_1.$$

A simple calculation verifies that

$$(S + T)^*(a_1 e_1 + \dots + a_n e_n) = 2a_1 e_2.$$

A final simple calculation shows that $(S + T)(S + T)^* \neq (S + T)^*(S + T)$. In other words, $S + T$ is not normal. Thus the set of normal operators on V is not closed under addition and hence is not a subspace of $\mathcal{L}(V)$.

6. Prove that if $T \in \mathcal{L}(V)$ is normal, then

$$\operatorname{range} T = \operatorname{range} T^*.$$

SOLUTION: Suppose T is normal. Then

$$\begin{aligned}
\operatorname{range} T &= (\operatorname{null} T^*)^{\perp} \\
&= (\operatorname{null} T)^{\perp} \\
&= \operatorname{range} T^*,
\end{aligned}$$

where the first equality comes from 6.46(d), the second equality comes from 7.6 (see especially the marginal comment at 7.6), and the third equality comes from 6.46(b).

7. Prove that if $T \in \mathcal{L}(V)$ is normal, then

$$\operatorname{null} T^k = \operatorname{null} T \quad \text{and} \quad \operatorname{range} T^k = \operatorname{range} T$$

for every positive integer k.

SOLUTION: Suppose $T \in \mathcal{L}(V)$ is normal and that k is a positive integer. Obviously we can assume that $k \geq 2$.

First we will prove that $\operatorname{null} T^k = \operatorname{null} T$. If $v \in \operatorname{null} T$, then

$$\begin{aligned} T^k v &= T^{k-1}(Tv) \\ &= T^{k-1}0 \\ &= 0, \end{aligned}$$

and so $v \in \operatorname{null} T^k$. Thus $\operatorname{null} T \subset \operatorname{null} T^k$.

To prove an inclusion in the other direction, suppose now that $v \in \operatorname{null} T^k$. Then

$$\begin{aligned} \langle T^*T^{k-1}v, T^*T^{k-1}v\rangle &= \langle TT^*T^{k-1}v, T^{k-1}v\rangle \\ &= \langle T^*T^k v, T^{k-1}v\rangle \\ &= \langle 0, T^{k-1}v\rangle \\ &= 0, \end{aligned}$$

where the second equality holds because $T^*T = TT^*$. The last equality above implies that $T^*T^{k-1}v = 0$. Thus

$$\begin{aligned} 0 &= \langle T^*T^{k-1}v, T^{k-2}v\rangle \\ &= \langle T^{k-1}v, T^{k-1}v\rangle. \end{aligned}$$

Hence $T^{k-1}v = 0$. In other words, $v \in \operatorname{null} T^{k-1}$. The same argument, with k replaced with $k-1$, shows that $v \in \operatorname{null} T^{k-2}$. Repeat this process until reaching the conclusion that $v \in \operatorname{null} T$. This shows that $\operatorname{null} T^k \subset \operatorname{null} T$, completing that proof that $\operatorname{null} T^k = \operatorname{null} T$.

Now we will show that $\operatorname{range} T^k = \operatorname{range} T$. If $v \in \operatorname{range} T^k$, then there exists $u \in V$ such that $v = T^k u = T(T^{k-1})u$, which implies that $v \in \operatorname{range} T$. Thus $\operatorname{range} T^k \subset \operatorname{range} T$. Note that

$$\begin{aligned}\dim \operatorname{range} T^k &= \dim V - \dim \operatorname{null} T^k \\ &= \dim V - \dim \operatorname{null} T \\ &= \dim \operatorname{range} T,\end{aligned}$$

where the first and third equalities come from 3.4 and the second equality comes from the first part of this exercise. Because $\operatorname{range} T^k$ and $\operatorname{range} T$ have the same dimension and one of them is contained in the other, these two subspaces of V must be equal, completing the proof.

8. Prove that there does not exist a self-adjoint operator $T \in \mathcal{L}(\mathbf{R}^3)$ such that $T(1,2,3) = (0,0,0)$ and $T(2,5,7) = (2,5,7)$.

SOLUTION: Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ is such that $T(1,2,3) = (0,0,0)$ and $T(2,5,7) = (2,5,7)$. Obviously $(1,2,3)$ is an eigenvector of T with eigenvalue 0 and $(2,5,7)$ is an eigenvector of T with eigenvalue 1. If T were self-adjoint, then eigenvectors corresponding to distinct eigenvalues would be orthogonal (see 7.8). Because $(1,2,3)$ and $(2,5,7)$ are not orthogonal, T cannot be self-adjoint.

9. Prove that a normal operator on a complex inner-product space is self-adjoint if and only if all its eigenvalues are real.

COMMENT: This exercise strengthens the analogy (for normal operators) between self-adjoint operators and real numbers.

SOLUTION: Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is normal.

If T is self-adjoint, then by 7.1 all its eigenvalues are real.

Conversely, suppose that all the eigenvalues of T are real. By the complex spectral theorem (7.9), there is an orthonormal basis $(e_1, \dots, e_n)$ of V consisting of eigenvectors of T. Thus there exist real numbers $\lambda_1, \dots, \lambda_n$ such that $Te_j = \lambda_j e_j$ for $j = 1, \dots, n$. The matrix of T with respect to the basis $(e_1, \dots, e_n)$ is the diagonal matrix with $\lambda_1, \dots, \lambda_n$ on the diagonal. This matrix equals its conjugate transpose. Thus $T = T^*$. In other words, T is self-adjoint, as desired.

10. Suppose V is a complex inner-product space and $T \in \mathcal{L}(V)$ is a normal operator such that $T^9 = T^8$. Prove that T is self-adjoint and $T^2 = T$.

SOLUTION: By the complex spectral theorem (7.9), there is an orthonormal basis $(e_1, \dots, e_n)$ of V consisting of eigenvectors of T. Let $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues. Thus

$$Te_j = \lambda_j e_j$$

for $j = 1, \ldots, n$. Applying T repeatedly to both sides of the equation above, we get $T^9 e_j = \lambda_j{}^9 e_j$ and $T^8 e_j = \lambda_j{}^8 e_j$. Thus $\lambda_j{}^9 = \lambda_j{}^8$, which implies that λ_j equals 0 or 1. In particular, all the eigenvalues of T are real. This implies (by the previous exercise) that T is self-adjoint.

Applying T to both sides of the equation above, we get

$$\begin{aligned} T^2 e_j &= \lambda_j{}^2 e_j \\ &= \lambda_j e_j \\ &= Te_j, \end{aligned}$$

where the second equality holds because λ_j equals 0 or 1. Because T^2 and T agree on a basis, they must be equal.

11. Suppose V is a complex inner-product space. Prove that every normal operator on V has a square root. (An operator $S \in \mathcal{L}(V)$ is called a *square root* of $T \in \mathcal{L}(V)$ if $S^2 = T$.)

SOLUTION: Suppose $T \in \mathcal{L}(V)$ is normal. By the complex spectral theorem (7.9), there is an orthonormal basis $(e_1, \ldots, e_n)$ of V consisting of eigenvectors of T. Thus there exist complex numbers $\lambda_1, \ldots, \lambda_n$ such that $Te_j = \lambda_j e_j$ for $j = 1, \ldots, n$. Define S to be the operator on V such that $Se_j = \lambda_j{}^{1/2} e_j$ for $j = 1, \ldots, n$; here $\lambda_j{}^{1/2}$ denotes a complex square root of λ_j (every nonzero complex number has two square roots—it does not matter which one is chosen). Then, as is easy to verify, $S^2 = T$. Thus S is a square root of T.

12. Give an example of a real inner-product space V and $T \in \mathcal{L}(V)$ and real numbers α, β with $\alpha^2 < 4\beta$ such that $T^2 + \alpha T + \beta I$ is not invertible.

COMMENT: This exercise shows that the hypothesis that T is self-adjoint is needed in 7.11, even for real vector spaces.

SOLUTION: Let $T \in \mathcal{L}(\mathbf{R}^2)$ be the counterclockwise rotation on $\mathbf{R}^2$; so $T(x, y) = (-y, x)$ for $(x, y) \in \mathbf{R}^2$. Thus $T^2 = -I$. Taking $\alpha = 0$ and $\beta = 1$, we have $\alpha^2 < 4\beta$ and

$$\begin{aligned} T^2 + \alpha T + \beta I &= T^2 + I \\ &= 0. \end{aligned}$$

In particular, $T^2 + \alpha T + \beta I$ is not invertible.

13. Prove or give a counterexample: every self-adjoint operator on V has a cube root. (An operator $S \in \mathcal{L}(V)$ is called a *cube root* of $T \in \mathcal{L}(V)$ if $S^3 = T$.)

SOLUTION: Suppose $T \in \mathcal{L}(V)$ is self-adjoint. By the spectral theorem (7.13), there is an orthonormal basis $(e_1, \ldots, e_n)$ of V consisting of eigenvectors of T. The corresponding eigenvalues must be real (by 7.1). Thus there exist real numbers $\lambda_1, \ldots, \lambda_n$ such that $Te_j = \lambda_j e_j$ for $j = 1, \ldots, n$. Define S to be the operator on V such that $Se_j = {\lambda_j}^{1/3} e_j$ for $j = 1, \ldots, n$. Then, as is easy to verify, $S^3 = T$. Thus S is a cube root of T, completing the proof that every self-adjoint operator on V has a cube root.

14. Suppose $T \in \mathcal{L}(V)$ is self-adjoint, $\lambda \in \mathbf{F}$, and $\epsilon > 0$. Prove that if there exists $v \in V$ such that $\|v\| = 1$ and

$$\|Tv - \lambda v\| < \epsilon,$$

then T has an eigenvalue λ' such that $|\lambda - \lambda'| < \epsilon$.

SOLUTION: By the spectral theorem (7.13), there is an orthonormal basis $(e_1, \ldots, e_n)$ of V consisting of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues.

Suppose $v \in V$ is such that $\|v\| = 1$ and $\|Tv - \lambda v\| < \epsilon$. From 6.17 we have

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n,$$

and so

$$Tv = \lambda_1 \langle v, e_1 \rangle e_1 + \cdots + \lambda_n \langle v, e_n \rangle e_n.$$

Thus

$$\begin{aligned}
\epsilon^2 &> \|Tv - \lambda v\|^2 \\
&= \|(\lambda_1 - \lambda)\langle v, e_1 \rangle e_1 + \cdots + (\lambda_n - \lambda)\langle v, e_n \rangle e_n\|^2 \\
&= |\lambda_1 - \lambda|^2 |\langle v, e_1 \rangle|^2 + \cdots + |\lambda_n - \lambda|^2 |\langle v, e_n \rangle|^2 \\
&\geq \left(\min\{|\lambda_1 - \lambda|^2, \ldots, |\lambda_n - \lambda|^2\}\right) \left(|\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2\right) \\
&= \min\{|\lambda_1 - \lambda|^2, \ldots, |\lambda_n - \lambda|^2\}
\end{aligned}$$

Thus $\epsilon > |\lambda_j - \lambda|$ for some j. In other words, there is an eigenvalue whose distance from λ is less than ϵ, as desired.

15. Suppose U is a finite-dimensional real vector space and $T \in \mathcal{L}(U)$. Prove that U has a basis consisting of eigenvectors of T if and only if there is an inner product on U that makes T into a self-adjoint operator.

SOLUTION: First suppose that U has a basis $(e_1, \dots, e_n)$ of eigenvectors of T. Because $(e_1, \dots, e_n)$ is a basis of U, every element of U can be uniquely written as a linear combination of $(e_1, \dots, e_n)$. Thus we can define an inner product on U by

$$\langle a_1e_1 + \cdots + a_ne_n, b_1e_1 + \cdots + b_ne_n\rangle = a_1b_1 + \cdots + a_nb_n.$$

It is easy to verify that this is indeed an inner product on U and that $(e_1, \dots, e_n)$ is on orthonormal basis of U with respect to this inner product. Because each e_j is an eigenvector of T, the operator T has a diagonal matrix with respect to the orthonormal basis $(e_1, \dots, e_n)$. Thus T is self-adjoint.

Conversely, now suppose that there is an inner product on U that makes T into a self-adjoint operator. Then by the spectral theorem (7.13), U has a basis consisting of eigenvectors of T.

16. Give an example of an operator T on an inner product space such that T has an invariant subspace whose orthogonal complement is not invariant under T.

COMMENT: This exercise shows that 7.18 can fail without the hypothesis that T is normal.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(w, z) = (z, 0)$. Then $T(w, 0) = (0, 0)$ for all $w \in \mathbf{F}$. Thus the subspace U defined by $U = \{(w, 0) : w \in \mathbf{F}\}$ is invariant under T. However, $U^\perp = \{(0, z) : \in \mathbf{F}\}$, which is not invariant under T because $(0, 1) \in U^\perp$ but $T(0, 1) = (1, 0) \notin U^\perp$.

Of course there are also many other examples.

17. Prove that the sum of any two positive operators on V is positive.

SOLUTION: Suppose S and T are positive operators on V. Because S and T are self-adjoint, so is $S + T$. Furthermore,

$$\begin{aligned}\langle (S+T)v, v\rangle &= \langle Sv, v\rangle + \langle Tv, v\rangle \\ &\geq 0.\end{aligned}$$

Thus $S + T$ is a positive operator, as desired.

18. Prove that if $T \in \mathcal{L}(V)$ is positive, then so is T^k for every positive integer k.

SOLUTION: Suppose $T \in \mathcal{L}(V)$ is positive and k is a positive integer. Then T^k is self-adjoint (because T is self-adjoint).

First consider the case where k is an even integer. Then we can write $k = 2m$ for some positive integer m. Now

$$\begin{aligned}\langle T^k v, v\rangle &= \langle T^{2m} v, v\rangle \\ &= \langle T^m v, T^m v\rangle \\ &\geq 0\end{aligned}$$

for every $v \in V$, where the second equality holds because T is self-adjoint. The inequality above shows that T^k is positive, as desired.

Now consider the case where k is an odd integer. Then we can write $k = 2m + 1$ for some nonnegative integer m. Now

$$\begin{aligned}\langle T^k v, v\rangle &= \langle T^{2m+1} v, v\rangle \\ &= \langle T(T^m v), T^m v\rangle \\ &\geq 0\end{aligned}$$

for every $v \in V$, where the second equality holds because T is self-adjoint and the inequality holds because T is positive. The inequality above shows that T^k is positive, as desired.

19. Suppose that T is a positive operator on V. Prove that T is invertible if and only if

$$\langle Tv, v\rangle > 0$$

for every $v \in V \setminus \{0\}$.

SOLUTION: First suppose that T is invertible. By 7.27, there exists an operator $S \in \mathcal{L}(V)$ such that $T = S^*S$. Suppose $v \in V \setminus \{0\}$. Then $Sv \neq 0$ because otherwise we would have $Tv = S^*Sv = 0$, which would contradict the invertibility of T. Now

$$\begin{aligned}\langle Tv, v\rangle &= \langle S^*Sv, v\rangle \\ &= \langle Sv, Sv\rangle \\ &> 0,\end{aligned}$$

as desired.

Now suppose that $\langle Tv, v\rangle > 0$ for every $v \in V \setminus \{0\}$. In particular, this means that $Tv \neq 0$ for every $v \in V \setminus \{0\}$. Thus T is injective, and hence T is invertible (see 3.21), as desired.

20. Prove or disprove: the identity operator on $\mathbf{F}^2$ has infinitely many self-adjoint square roots.

SOLUTION: For each $t \in [-1, 1]$, the operator whose matrix (with respect to the standard basis) equals

$$\begin{bmatrix} t & \sqrt{1-t^2} \\ \sqrt{1-t^2} & -t \end{bmatrix}$$

is self-adjoint and a square root of the identity operator, as can be verified by squaring the matrix above. Thus the identity operator has infinitely many self-adjoint square roots.

21. Prove or give a counterexample: if $S \in \mathcal{L}(V)$ and there exists an orthonormal basis $(e_1, \ldots, e_n)$ of V such that $\|Se_j\| = 1$ for each e_j, then S is an isometry.

SOLUTION: Define $S \in \mathcal{L}(\mathbf{F}^2)$ by

$$S(w, z) = (w + z, 0).$$

With the usual inner product on $\mathbf{F}^2$, the standard basis $\big((1,0),(0,1)\big)$ is an orthonormal basis of $\mathbf{F}^2$. Note that $\|S(1,0)\| = \|S(0,1)\| = 1$. However, S is not an isometry because $\|S(1,-1)\| = 0$.

Of course there are also many other examples.

22. Prove that if $S \in \mathcal{L}(\mathbf{R}^3)$ is an isometry, then there exists a nonzero vector $x \in \mathbf{R}^3$ such that $S^2x = x$.

SOLUTION: Suppose $S \in \mathcal{L}(\mathbf{R}^3)$ is an isometry. Then there is a basis of $\mathbf{R}^3$ with respect to which S has a block diagonal matrix, where each block on the diagonal is a 1-by-1 matrix containing 1 or -1 or is a 2-by-2 matrix (see 7.38). Because $\mathbf{R}^3$ has odd dimension, at least one of these blocks must be a 1-by-1 matrix. In other words, either 1 or -1 must be an eigenvalue of S. Thus there is a nonzero vector $x \in \mathbf{R}^3$ such that $Sx = \lambda x$, where $\lambda = \pm 1$. Hence

$$S^2x = S(Sx) = S(\lambda x) = \lambda Sx = \lambda^2 x = x.$$

23. Define $T \in \mathcal{L}(\mathbf{F}^3)$ by

$$T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2).$$

Find (explicitly) an isometry $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T = S\sqrt{T^*T}$.

SOLUTION: With respect to the standard basis of $\mathbf{F}^3$, we have

$$\mathcal{M}(T) = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}.$$

Thus

$$\mathcal{M}(T^*) = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{bmatrix}.$$

Computing the product $\mathcal{M}(T^*)\mathcal{M}(T)$, which equals $\mathcal{M}(T^*T)$, we get

$$\mathcal{M}(T^*T) = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

From the matrix above, we see that $(T^*T)(z_1, z_2, z_3) = (4z_1, 9z_2, z_3)$. Thus $\sqrt{T^*T}(z_1, z_2, z_3) = (2z_1, 3z_2, z_3)$. Hence if we define $S \in \mathcal{L}(\mathbf{F}^3)$ by

$$S(z_1, z_2, z_3) = (z_3, z_1, z_2),$$

then S is an isometry and $T = S\sqrt{T^*T}$.

24. Suppose $T \in \mathcal{L}(V)$, $S \in \mathcal{L}(V)$ is an isometry, and $R \in \mathcal{L}(V)$ is a positive operator such that $T = SR$. Prove that $R = \sqrt{T^*T}$.

COMMENT: This exercise shows that if we write T as the product of an isometry and a positive operator (as in the polar decomposition), then the positive operator must equal $\sqrt{T^*T}$.

SOLUTION: Taking adjoints of both sides of the equation $T = SR$, we have

$$\begin{aligned} T^* &= R^*S^* \\ &= RS^*, \end{aligned}$$

where the last equation holds because R is positive (and hence self-adjoint). Multiplying together our formulas for T^* and T, we get

$$\begin{aligned} T^*T &= RS^*SR \\ &= R^2, \end{aligned}$$

where the last equation holds because S is an isometry (and hence $S^*S = I$ by 7.36). The equation above asserts that R is a square root of T^*T; because R is positive, this implies that $R = \sqrt{T^*T}$.

25. Suppose $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if there exists a unique isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.

SOLUTION: First suppose that T is invertible. The polar decomposition (7.41) states that there exists an isometry $S \in \mathcal{L}(V)$ such that

$$T = S\sqrt{T^*T}.$$

Because T is invertible, this implies that $\sqrt{T^*T}$ is invertible (see Exercise 22 in Chapter 3). Thus the equation above implies that $S = T(\sqrt{T^*T})^{-1}$. Because S must be given by this formula, we see that there is a unique operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$, as desired.

Now suppose that there exists a unique isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$. This means that the linear map S_2 in the proof of the polar decomposition (7.41) must be 0 because otherwise we could replace S_2 with $-S_2$ and get another choice for S. But $\operatorname{range} S_2$ equals $(\operatorname{range} T)^\perp$, and hence $(\operatorname{range} T)^\perp = \{0\}$. This implies that $\operatorname{range} T = V$, which implies that T is invertible (by 3.21), as desired.

26. Prove that if $T \in \mathcal{L}(V)$ is self-adjoint, then the singular values of T equal the absolute values of the eigenvalues of T (repeated appropriately).

SOLUTION: Suppose $T \in \mathcal{L}(V)$ is self-adjoint. There exists an orthonormal basis $(e_1, \dots, e_n)$ of V consisting of eigenvectors of T. Thus

$$Te_j = \lambda_j e_j$$

for each j, where $\lambda_1, \dots, \lambda_n \in \mathbf{R}$ are the eigenvalues of T. Thus

$$\begin{aligned} T^*Te_j &= T^2 e_j \\ &= (\lambda_j)^2 e_j \end{aligned}$$

for each j. The equation above implies that $\sqrt{T^*T}e_j = |\lambda_j| e_j$ for each j. Thus the singular values of T are $|\lambda_1|, \dots, |\lambda_n|$, as desired.

27. Prove or give a counterexample: if $T \in \mathcal{L}(V)$, then the singular values of T^2 equal the squares of the singular values of T.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{F}^2)$ by

$$T(z_1, z_2) = (z_2, 0).$$

Then $T^*T(z_1, z_2) = (0, z_2)$ and hence $\sqrt{T^*T}(z_1, z_2) = (0, z_2)$. Thus the eigenvalues of $\sqrt{T^*T}$ are $0, 1$. Hence the singular values of T are $0, 1$.

However, $T^2 = 0$, so the singular values of T^2 are $0, 0$. Thus for this operator T, the singular values of T^2 do not equal the squares of the singular values of T.

Of course there are also many other examples.

28. Suppose $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if 0 is not a singular value of T.

SOLUTION: If $S \in \mathcal{L}(V)$ and $ST = TS = I$, then taking adjoints we get $T^*S^* = S^*T^* = I$. Thus if T is invertible, then so is T^*.

Now

$$\begin{aligned}
T \text{ is invertible} &\iff T \text{ and } T^* \text{ are invertible}\\
&\iff T^*T \text{ is invertible}\\
&\iff \sqrt{T^*T}\sqrt{T^*T} \text{ is invertible}\\
&\iff \sqrt{T^*T} \text{ is invertible}\\
&\iff 0 \text{ is not an eigenvalue of } \sqrt{T^*T}\\
&\iff 0 \text{ is not a singular value of } T,
\end{aligned}$$

where the second and fourth equivalences follow from Exercise 22 in Chapter 2.

29. Suppose $T \in \mathcal{L}(V)$. Prove that $\dim \operatorname{range} T$ equals the number of nonzero singular values of T.

SOLUTION: By the singular value decomposition (7.46), there exist orthonormal bases $(u_1, \ldots, u_n)$ and $(w_1, \ldots, w_n)$ of V such that

$$Tv = s_1\langle v, u_1\rangle w_1 + \cdots + s_n\langle v, u_n\rangle w_n$$

for every $v \in V$, where $s_1, \ldots, s_n$ are the singular values of T. For each j, we have $Tu_j = s_j w_j$. Thus each w_j corresponding to a nonzero s_j is in $\operatorname{range} T$. The equation above also shows that the w_j's corresponding to nonzero s_j's span $\operatorname{range} T$. Thus $\dim \operatorname{range} T$ equals the number of nonzero singular values of T.

30. Suppose $S \in \mathcal{L}(V)$. Prove that S is an isometry if and only if all the singular values of S equal 1.

SOLUTION: We have

$$\begin{aligned} S \text{ is an isometry} &\iff S^*S = I \\ &\iff \sqrt{S^*S} = I \\ &\iff \text{all the eigenvalues of } \sqrt{S^*S} \text{ equal } 1 \\ &\iff \text{all the singular values of } S \text{ equal } 1, \end{aligned}$$

where the first equivalence comes from 7.36 and the third equivalence comes from the spectral theorem (7.9 or 7.13) applied to the self-adjoint operator $\sqrt{S^*S}$.

31. Suppose $T_1, T_2 \in \mathcal{L}(V)$. Prove that T_1 and T_2 have the same singular values if and only if there exist isometries $S_1, S_2 \in \mathcal{L}(V)$ such that $T_1 = S_1 T_2 S_2$.

SOLUTION: First suppose that T_1 and T_2 have the same singular values $s_1, \dots, s_n$. By the singular-value decomposition (7.46), there exist orthonormal bases $(e_1, \dots, e_n), (f_1, \dots, f_n), (e'_1, \dots, e'_n), (f'_1, \dots, f'_n)$ of V such that

$$\begin{aligned} T_1 v &= s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n, \\ T_2 v &= s_1 \langle v, e'_1 \rangle f'_1 + \cdots + s_n \langle v, e'_n \rangle f'_n \end{aligned}$$

for every $v \in V$. Define $S_1, S_2 \in \mathcal{L}(V)$ by

$$\begin{aligned} S_1(a_1 f'_1 + \cdots + a_n f'_n) &= a_1 f_1 + \cdots + a_n f_n, \\ S_2(a_1 e_1 + \cdots + a_n e_n) &= a_1 e'_1 + \cdots + a_n e'_n. \end{aligned}$$

Then

$$\begin{aligned} \|S_1(a_1 f'_1 + \cdots + a_n f'_n)\|^2 &= \|a_1 f_1 + \cdots + a_n f_n\|^2 \\ &= |a_1|^2 + \cdots + |a_n|^2 \\ &= \|a_1 f'_1 + \cdots + a_n f'_n\|^2, \end{aligned}$$

and thus S_1 is an isometry. Similarly, S_2 is an isometry. This implies that ${S_2}^* = {S_2}^{-1}$ (see 7.36). In particular, ${S_2}^* e'_j = e_j$. Now for $v \in V$ we have

$$\begin{aligned} T_2(S_2 v) &= s_1 \langle S_2 v, e'_1 \rangle f'_1 + \cdots + s_n \langle S_2 v, e'_n \rangle f'_n \\ &= s_1 \langle v, {S_2}^* e'_1 \rangle f'_1 + \cdots + s_n \langle v, {S_2}^* e'_n \rangle f'_n \\ &= s_1 \langle v, e_1 \rangle f'_1 + \cdots + s_n \langle v, e_n \rangle f'_n. \end{aligned}$$

Thus

$$\begin{aligned} S_1(T_2S_2v) &= s_1\langle v, e_1\rangle S_1 f_1' + \cdots + s_n\langle v, e_n\rangle S_1 f_n' \\ &= s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n \\ &= T_1 v \end{aligned}$$

for every $v \in V$. Hence $S_1T_2S_2 = T_1$, as desired.

To prove the implication in the other direction, now suppose that there exist isometries $S_1, S_2 \in \mathcal{L}(V)$ such that $T_1 = S_1T_2S_2$. Using 7.36, we have

$$\begin{aligned} T_1{}^*T_1 &= S_2{}^*T_2{}^*S_1{}^*S_1T_2S_2 \\ &= S_2{}^{-1}T_2{}^*T_2S_2. \end{aligned}$$

This implies that $T_1{}^*T_1$ and $T_2{}^*T_2$ have the same eigenvalues (and that the corresponding spaces of eigenvectors have the same dimensions). Thus T_1 and T_2 have the same singular values.

32. Suppose $T \in \mathcal{L}(V)$ has singular-value decomposition given by

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for every $v \in V$, where $s_1, \ldots, s_n$ are the singular values of T and $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$ are orthonormal bases of V.

(a) Prove that

$$T^*v = s_1\langle v, f_1\rangle e_1 + \cdots + s_n\langle v, f_n\rangle e_n$$

for every $v \in V$.

(b) Prove that if T is invertible, then

$$T^{-1}v = \frac{\langle v, f_1\rangle e_1}{s_1} + \cdots + \frac{\langle v, f_n\rangle e_n}{s_n}$$

for every $v \in V$.

SOLUTION: (a): Fix $v \in V$. Then

$$\begin{aligned} \langle w, T^*v\rangle &= \langle Tw, v\rangle \\ &= \langle s_1\langle w, e_1\rangle f_1 + \cdots + s_n\langle w, e_n\rangle f_n, v\rangle \\ &= s_1\langle w, e_1\rangle\langle f_1, v\rangle + \cdots + s_n\langle w, e_n\rangle\langle f_n, v\rangle \\ &= \langle w, s_1\langle v, f_1\rangle e_1 + \cdots + s_n\langle v, f_n\rangle e_n\rangle \end{aligned}$$

for all $w \in V$. This implies that

$$T^*v = s_1\langle v, f_1\rangle e_1 + \cdots + s_n\langle v, f_n\rangle e_n,$$

as desired.

(b): Suppose T is invertible. Let $v \in V$ and let

$$w = \frac{\langle v, f_1\rangle e_1}{s_1} + \cdots + \frac{\langle v, f_n\rangle e_n}{s_n};$$

none of the singular values $s_1, \ldots, s_n$ equals 0 (see Exercise 28 of this chapter), so this makes sense. Now

$$\begin{aligned}
Tw &= \frac{\langle v, f_1\rangle Te_1}{s_1} + \cdots + \frac{\langle v, f_n\rangle Te_n}{s_n} \\
&= \frac{\langle v, f_1\rangle s_1 f_1}{s_1} + \cdots + \frac{\langle v, f_n\rangle s_n f_n}{s_n} \\
&= \langle v, f_1\rangle f_1 + \cdots + \langle v, f_n\rangle f_n \\
&= v.
\end{aligned}$$

Thus $w = T^{-1}v$, as desired.

33. Suppose $T \in \mathcal{L}(V)$. Let $\hat{s}$ denote the smallest singular value of T, and let s denote the largest singular value of T. Prove that

$$\hat{s}\|v\| \le \|Tv\| \le s\|v\|$$

for every $v \in V$.

SOLUTION: Let $v \in V$. By the singular value decomposition (7.46), there exist orthonormal bases $(u_1, \ldots, u_n)$ and $(w_1, \ldots, w_n)$ of V such that

$$Tv = s_1\langle v, u_1\rangle w_1 + \cdots + s_n\langle v, u_n\rangle w_n,$$

where $s_1, \ldots, s_n$ are the singular values of T. Because $(u_1, \ldots, u_n)$ and $(w_1, \ldots, w_n)$ are both orthonormal bases of V, we have

$$\begin{aligned}
\hat{s}^2\|v\|^2 &= \hat{s}^2(|\langle v, u_1\rangle|^2 + \cdots + |\langle v, u_n\rangle|^2) \\
&\le {s_1}^2|\langle v, u_1\rangle|^2 + \cdots + {s_n}^2|\langle v, u_n\rangle|^2 \\
&= \|Tv\|^2,
\end{aligned}$$

giving the first desired inequality. Also,

$$\begin{aligned}\|Tv\|^2 &= {s_1}^2|\langle v, u_1\rangle|^2 + \cdots + {s_n}^2|\langle v, u_n\rangle|^2 \\ &\le s^2(|\langle v, u_1\rangle|^2 + \cdots + |\langle v, u_n\rangle|^2) \\ &= s^2\|v\|^2,\end{aligned}$$

giving the second desired inequality.

34. Suppose $T', T'' \in \mathcal{L}(V)$. Let s' denote the largest singular value of T', let s'' denote the largest singular value of T'', and let s denote the largest singular value of $T' + T''$. Prove that $s \le s' + s''$.

SOLUTION: Let $T = T' + T''$. Because s is a singular value of T, we know that s is an eigenvalue of $\sqrt{T^*T}$. Thus there exists a vector $v \in V$ such that $\|v\| = 1$ and $\sqrt{T^*T}v = sv$. Now

$$\begin{aligned}s &= \|sv\| \\ &= \|\sqrt{T^*T}v\| \\ &= \|Tv\| \\ &= \|T'v + T''v\| \\ &\le \|T'v\| + \|T''v\| \\ &\le s'\|v\| + s''\|v\| \\ &= s' + s'',\end{aligned}$$

where the third line above comes from 7.42 and the sixth line above comes from the previous exercise.

CHAPTER 8

Operators on Complex Vector Spaces

1. Define $T \in \mathcal{L}(\mathbf{C}^2)$ by

$$T(w, z) = (z, 0).$$

Find all generalized eigenvectors of T.

SOLUTION: Suppose λ is an eigenvalue of T. For this particular operator, the eigenvalue-eigenvector equation $T(w, z) = \lambda(w, z)$ becomes the system of equations

$$\begin{aligned} z &= \lambda w \\ 0 &= \lambda z. \end{aligned}$$

If $\lambda \neq 0$, then the second equation implies that $z = 0$, and the first equation then implies that $w = 0$. Because an eigenvalue must have a nonzero eigenvector, this shows that 0 is the only possible eigenvalue of T. For $\lambda = 0$, the equations above show that z must equal 0, but w can be arbitrary. Thus 0 is indeed an eigenvalue of T, and the set of eigenvectors corresponding to this eigenvalue is

$$\{(w, 0) : w \in \mathbf{F}\}.$$

Note that $T^2 = 0$. Thus every vector in $\mathbf{C}^2$ is a generalized eigenvector of T (corresponding to the eigenvalue 0).

2. Define $T \in \mathcal{L}(\mathbf{C}^2)$ by

$$T(w, z) = (-z, w).$$

Find all generalized eigenvectors of T.

SOLUTION: On page 78 of the textbook we saw that the eigenvalues of T are i and $-i$. Note that $T^2 = -I$.

The set of generalized eigenvectors of T corresponding to the eigenvalue i equals $\operatorname{null}(T - iI)^2$ (by 8.7). To compute this, note that

$$\begin{aligned}(T - iI)^2 &= T^2 - 2iT - I \\ &= -2I - 2iT \\ &= -2i(T - iI).\end{aligned}$$

Thus $\operatorname{null}(T - iI)^2$ equals the set of eigenvectors of T corresponding to the eigenvalue i. On page 78 of the textbook we noted that this equals $\{(a, -ia) : a \in \mathbf{C}\}$.

The set of generalized eigenvectors of T corresponding to the eigenvalue $-i$ equals $\operatorname{null}(T + iI)^2$ (by 8.7). To compute this, note that

$$\begin{aligned}(T + iI)^2 &= T^2 + 2iT - I \\ &= -2I + 2iT \\ &= 2i(T + iI).\end{aligned}$$

Thus $\operatorname{null}(T + iI)^2$ equals the set of eigenvectors of T corresponding to the eigenvalue $-i$. On page 78 of the textbook we noted that this equals $\{(a, ia) : a \in \mathbf{C}\}$.

3. Suppose $T \in \mathcal{L}(V)$, m is a positive integer, and $v \in V$ is such that $T^{m-1}v \neq 0$ but $T^m v = 0$. Prove that

$$(v, Tv, T^2v, \ldots, T^{m-1}v)$$

is linearly independent.

SOLUTION: Suppose $a_0, a_1, a_2, \ldots, a_{m-1} \in \mathbf{F}$ are such that

$$a_0 v + a_1 Tv + a_2 T^2 v + \cdots + a_{m-1} T^{m-1} v = 0.$$

Because $T^m v = 0$, if we apply T^{m-1} to both sides of the equation above, we get $a_0 T^{m-1} v = 0$. Because $T^{m-1}v \neq 0$, this implies that $a_0 = 0$. Thus the equation above can be rewritten as

$$a_1Tv + a_2T^2v + \cdots + a_{m-1}T^{m-1}v = 0.$$

Applying T^{m-2} to both sides of this equation, we get $a_1T^{m-1}v = 0$. Thus $a_1 = 0$. Continuing in this fashion, we have $a_0 = a_1 = a_2 = \cdots = a_{m-1} = 0$, which means that $(v, Tv, T^2v, \ldots, T^{m-1}v)$ is linearly independent.

4. Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is defined by $T(z_1, z_2, z_3) = (z_2, z_3, 0)$. Prove that T has no square root. More precisely, prove that there does not exist $S \in \mathcal{L}(\mathbf{C}^3)$ such that $S^2 = T$.

SOLUTION: Note that $T^3 = 0$. Suppose there exists $S \in \mathcal{L}(\mathbf{C}^3)$ such that $S^2 = T$. Then $S^6 = T^3 = 0$, so S is nilpotent. By 8.8, this implies that $S^3 = 0$. Thus

$$\begin{aligned} T^2 &= S^4 \\ &= SS^3 \\ &= 0. \end{aligned}$$

But $T^2(z_1, z_2, z_3) = (z_3, 0, 0)$, so T^2 is not the 0 operator, contradicting the equation above. This contradiction shows that our supposition that there exists $S \in \mathcal{L}(\mathbf{C}^3)$ such that $S^2 = T$ must have been false.

5. Suppose $S, T \in \mathcal{L}(V)$. Prove that if ST is nilpotent, then TS is nilpotent.

SOLUTION: Suppose ST is nilpotent. Thus there exists a positive integer n such that $(ST)^n = 0$. Now

$$\begin{aligned} (TS)^{n+1} &= (TS)(TS)\ldots(TS) \\ &= T(ST)(ST)\ldots(ST)S \\ &= T(ST)^nS \\ &= (T)(0)(S) \\ &= 0, \end{aligned}$$

and thus TS is nilpotent.

6. Suppose $N \in \mathcal{L}(V)$ is nilpotent. Prove (without using 8.26) that 0 is the only eigenvalue of N.

SOLUTION: There is a positive integer m such that $N^m = 0$. This implies that N is not injective, so 0 is an eigenvalue of N.

Conversely, suppose λ is an eigenvalue of N. Then there exists a nonzero vector $v \in V$ such that

$$\lambda v = Nv.$$

Repeatedly applying N to both sides of this equation shows that

$$\begin{aligned}\lambda^m v &= N^m v\\ &= 0.\end{aligned}$$

Thus $\lambda = 0$, as desired.

7. Suppose V is an inner-product space. Prove that if $N \in \mathcal{L}(V)$ is self-adjoint and nilpotent, then $N = 0$.

SOLUTION: Suppose $N \in \mathcal{L}(V)$ is self-adjoint and nilpotent. Because N is self-adjoint, there is an orthonormal basis $(e_1, \ldots, e_n)$ of V consisting of eigenvectors of N (by the spectral theorem). Because N is nilpotent, 0 is the only eigenvalue of N (see Exercise 6 of this chapter). Thus the eigenvalue corresponding to each e_j must equal 0. In other words, $Ne_j = 0$ for each j. Because $(e_1, \ldots, e_n)$ is a basis of V, this implies that $N = 0$.

8. Suppose $N \in \mathcal{L}(V)$ is such that $\operatorname{null} N^{\dim V - 1} \neq \operatorname{null} N^{\dim V}$. Prove that N is nilpotent and that

$$\dim \operatorname{null} N^j = j$$

for every integer j with $0 \leq j \leq \dim V$.

SOLUTION: Because $\operatorname{null} N^{\dim V-1} \neq \operatorname{null} N^{\dim V}$, we know (by 8.5) that $\operatorname{null} N^{j-1} \neq \operatorname{null} N^j$ whenever $0 \leq j \leq \dim V$. Thus

$$\{0\} = \operatorname{null} N^0 \subsetneq \operatorname{null} N^1 \subsetneq \ldots \subsetneq \operatorname{null} N^{\dim V-1} \subsetneq \operatorname{null} N^{\dim V}.$$

At each of the strict inclusions in the chain above, the dimension must increase by at least 1. However, if the dimension increases by more than 1 at any step, we would end up with $\dim \operatorname{null} N^{\dim V} > \dim V$, a contradiction because a subspace of V cannot have dimension larger than $\dim V$. Thus the dimension increases by exactly one at each step. In other words, $\dim \operatorname{null} N^j = j$ for every integer j with $0 \leq j \leq \dim V$. In particular, taking $j = \dim V$, we have $\dim \operatorname{null} N^{\dim V} = \dim V$. This means that $\operatorname{null} N^{\dim V} = V$. Thus $N^{\dim V} = 0$, and so N is nilpotent.

9. Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer such that

$$\operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

Prove that $\operatorname{range} T^k = \operatorname{range} T^m$ for all $k > m$.

SOLUTION: Suppose $u \in \operatorname{range} T^{m+1}$. Thus there exists a vector v in V (the domain of T) such that $u = T^{m+1}v$. Now $T^m v$ is in $\operatorname{range} T^m$, which by our hypothesis equals $\operatorname{range} T^{m+1}$. Thus there exists $w \in V$ such that $T^m v = T^{m+1}w$. Putting all this together, we have

$$\begin{aligned} u &= T^{m+1}v \\ &= T(T^m v) \\ &= T(T^{m+1}w) \\ &= T^{m+2}w. \end{aligned}$$

Thus $u \in \operatorname{range} T^{m+2}$. Because u was an arbitrary vector in $\operatorname{range} T^{m+1}$, we have shown that $\operatorname{range} T^{m+1} \subset \operatorname{range} T^{m+2}$. We also have an easy inclusion in the other direction, so we conclude that $\operatorname{range} T^{m+1} = \operatorname{range} T^{m+2}$.

In the paragraph above, we showed that $\operatorname{range} T^m = \operatorname{range} T^{m+1}$ implies $\operatorname{range} T^{m+1} = \operatorname{range} T^{m+2}$. Apply that result, with m replaced with $m+1$, to conclude that $\operatorname{range} T^{m+2} = \operatorname{range} T^{m+3}$. Continuing in this fashion, we see that

$$\operatorname{range} T^m = \operatorname{range} T^{m+1} = \operatorname{range} T^{m+2} = \dots,$$

as desired.

10. Prove or give a counterexample: if $T \in \mathcal{L}(V)$, then

$$V = \operatorname{null} T \oplus \operatorname{range} T.$$

SOLUTION: Define $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(w, z) = (z, 0)$. Thus

$$\operatorname{null} T = \operatorname{range} T = \{(w, 0) : w \in \mathbf{F}\},$$

which clearly implies that $\mathbf{F}^2$ is not the direct sum of $\operatorname{null} T$ and $\operatorname{range} T$.

Of course there are also many other examples.

11. Prove that if $T \in \mathcal{L}(V)$, then

$$V = \operatorname{null} T^n \oplus \operatorname{range} T^n,$$

where $n = \dim V$.

SOLUTION: Let $T \in \mathcal{L}(V)$. First we show that

$$V = \operatorname{null} T^n + \operatorname{range} T^n.$$

To do this, let $v \in V$. Then

$$v = (v - T^n u) + T^n u$$

for any vector $u \in V$. Obviously $T^n u \in \operatorname{range} T^n$. Thus looking at the equation above, we see that we need to show that there exists $u \in V$ such that $v - T^n u \in \operatorname{null} T^n$. In other words, we want a vector $u \in V$ such that $T^n(v - T^n u) = 0$, which is equivalent to $T^n v = T^{2n} u$. But $T^n v \in \operatorname{range} T^n$, and $\operatorname{range} T^n = \operatorname{range} T^{2n}$ (by 8.9), so $T^n v \in \operatorname{range} T^{2n}$. Thus there indeed exists $u \in V$ such that $T^n v = T^{2n} u$, completing our proof $v \in \operatorname{null} T^n + \operatorname{range} T^n$. Because v was an arbitrary vector in V, this implies that $V = \operatorname{null} T^n + \operatorname{range} T^n$.

For any linear map (and in particular for T^n), the dimension of the domain equals the sum of the dimensions of the null space and range (by 3.4). In other words,

$$\dim V = \dim \operatorname{null} T^n + \dim \operatorname{range} T^n.$$

This equation, along with the equation $V = \operatorname{null} T^n + \operatorname{range} T^n$, implies that $V = \operatorname{null} T^n \oplus \operatorname{range} T^n$ (by 2.19).

12. Suppose V is a complex vector space, $N \in \mathcal{L}(V)$, and 0 is the only eigenvalue of N. Prove that N is nilpotent. Give an example to show that this is not necessarily true on a real vector space.

SOLUTION: Because 0 is the only eigenvalue of N, 8.23(a) implies that every vector in V is a generalized eigenvector of T corresponding to the eigenvalue 0. This implies that N is nilpotent.

Define $T \in \mathcal{L}(\mathbf{R}^3)$ by

$$T(x, y, z) = (-y, x, 0).$$

Then 0 is an eigenvalue of T because $T(0,0,1) = (0,0,0)$. As can be verified from the definition of eigenvalue, T has no other eigenvalues (which must be in $\mathbf{R}$, because T is an operator on a real vector space). However, $T^3(x, y, z) = (y, -x, 0)$. In particular, $T^3 \neq 0$. Thus T is not nilpotent.

Of course there are also many other examples.

13. Suppose that V is a complex vector space with $\dim V = n$ and $T \in \mathcal{L}(V)$ is such that

$$\operatorname{null} T^{n-2} \neq \operatorname{null} T^{n-1}.$$

Prove that T has at most two distinct eigenvalues.

SOLUTION: Because $\operatorname{null} T^{n-2} \neq \operatorname{null} T^{n-1}$, we see that $\dim \operatorname{null} T^j$ is at least 1 more than $\dim \operatorname{null} T^{j-1}$ for $j = 1, \dots, n-1$ (by 8.5). Thus $\dim \operatorname{null} T^{n-1} \geq n-1$. In particular, 0 is an eigenvalue of T with multiplicity at least $n-1$. Because the sum of the multiplicities of all the eigenvalues of T equals n (by 8.18), this implies that T can have at most one additional eigenvalue.

14. Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $(z-7)^2(z-8)^2$.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{C}^4)$ by

$$T(z_1, z_2, z_3, z_4) = (7z_1, 7z_2, 8z_3, 8z_4).$$

Then $\operatorname{null}(T - 7I)$ is the two-dimensional subspace

$$\{(z_1, z_2, 0, 0) : z_1, z_2 \in \mathbf{C}\}$$

and $\operatorname{null}(T - 8I)$ is the two-dimensional subspace

$$\{(0, 0, z_3, z_4) : z_3, z_4 \in \mathbf{C}\}.$$

Thus 7 is an eigenvalue of T with multiplicity at least 2 and 8 is an eigenvalue of T with multiplicity at least 2. Because $2+2 = 4 = \dim \mathbf{C}^4$, there can be no other eigenvalues of T and the eigenvalues 7 and 8 must have multiplicity 2 (by 8.18). Thus the characteristic polynomial of T equals $(z-7)^2(z-8)^2$.

Of course there are also many other examples.

15. Suppose V is a complex vector space. Suppose $T \in \mathcal{L}(V)$ is such that 5 and 6 are eigenvalues of T and that T has no other eigenvalues. Prove that

$$(T - 5I)^{n-1}(T - 6I)^{n-1} = 0,$$

where $n = \dim V$.

SOLUTION: Because 5 and 6 are eigenvalues of T and T has no other eigenvalues, the characteristic polynomial of T must be of the form

$$(z-5)^{d_1}(z-6)^{d_2},$$

where $1 \leq d_1$ and $1 \leq d_2$. Because $d_1 + d_2 = n$, we must also have $d_1 \leq n-1$ and $d_2 \leq n-1$. The Cayley-Hamilton theorem implies that

$$(T - 5I)^{d_1}(T - 6I)^{d_2} = 0.$$

Because $d_1 \leq n-1$ and $d_2 \leq n-1$, we can multiply the equation above by appropriate powers of $T-5I$ and $T-6I$ to get $(T-5I)^{n-1}(T-6I)^{n-1} = 0$.

16. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Prove that V has a basis consisting of eigenvectors of T if and only if every generalized eigenvector of T is an eigenvector of T.

COMMENT: For complex vector spaces, this exercise adds another equivalence to the list given by 5.21.

SOLUTION: First suppose that V has a basis consisting of eigenvectors of T. Thus there exists a basis $(v_1, \ldots, v_n)$ of V and $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$ such that $Tv_j = \lambda_j v_j$ for each j. Suppose $v \in V$ is a generalized eigenvector of T corresponding to an eigenvalue λ. Because $(v_1, \ldots, v_n)$ is a basis of V, there exist $a_1, \ldots, a_n \in \mathbf{C}$ such that

$$v = a_1 v_1 + \cdots + a_n v_n.$$

Thus

$$(T - \lambda I)v = (\lambda_1 - \lambda)a_1 v_1 + \cdots + (\lambda_n - \lambda)a_n v_n.$$

Applying $T - \lambda I$ repeatedly to both sides of this equation, we get

$$(T - \lambda I)^n v = (\lambda_1 - \lambda)^n a_1 v_1 + \cdots + (\lambda_n - \lambda)^n a_n v_n.$$

The left side of the equation above equals 0 (because v is a generalized eigenvector of T corresponding to the eigenvalue λ). Thus the right side of the equation above equals 0. Thus $(\lambda_j - \lambda)^n a_j = 0$ for each j. This implies that the indices j such that $a_j \neq 0$ must all have all satisfy $\lambda_j = \lambda$, which means that v is a linear combination of the v_j's that correspond to eigenvalue λ, which means that v is itself an eigenvector corresponding to eigenvalue λ, as desired.

To prove the other direction, now suppose that every generalized eigenvector of T is an eigenvector of T. By 8.25, there exists a basis of V consisting of generalized eigenvectors of T. Because every generalized eigenvector of T is an eigenvector of T, this gives a basis of V consisting of eigenvectors of T, as desired.

17. Suppose V is an inner-product space and $N \in \mathcal{L}(V)$ is nilpotent. Prove that there exists an orthonormal basis of V with respect to which N has an upper-triangular matrix.

SOLUTION: By 8.26, there is a basis of V with respect to which N has an upper-triangular matrix. By 6.27, this basis can be chosen to be orthonormal. (If V is a complex inner-product space, then we can just use 6.28 directly, so the hypothesis that N is nilpotent is not needed.)

18. Define $N \in \mathcal{L}(\mathbf{F}^5)$ by

$$N(x_1, x_2, x_3, x_4, x_5) = (2x_2, 3x_3, -x_4, 4x_5, 0).$$

Find a square root of $I + N$.

SOLUTION: Note that

$$\begin{aligned}
N^2(x_1, x_2, x_3, x_4, x_5) &= (6x_3, -3x_4, -4x_5, 0, 0)\\
N^3(x_1, x_2, x_3, x_4, x_5) &= (-6x_4, -12x_5, 0, 0, 0)\\
N^4(x_1, x_2, x_3, x_4, x_5) &= (-24x_5, 0, 0, 0, 0)\\
N^5(x_1, x_2, x_3, x_4, x_5) &= (0, 0, 0, 0, 0).
\end{aligned}$$

Because $N^5 = 0$, the proof of 8.30 shows that

$$I + \frac{1}{2}N - \frac{1}{8}N^2 + \frac{1}{16}N^3 - \frac{5}{128}N^4$$

is a square root of $I + N$. Using the formulas above, we calculate that the operator $S \in \mathcal{L}(\mathbf{F}^5)$ defined by

$$\begin{aligned}
&S(x_1, x_2, x_3, x_4, x_5) =\\
&\Big(x_1 + x_2 - \frac{3x_3}{4} - \frac{3x_4}{8} + \frac{15x_5}{16},\ x_2 + \frac{3x_3}{2} + \frac{3x_4}{8} - \frac{3x_5}{4},\\
&x_3 - \frac{x_4}{2} + \frac{x_5}{2},\ x_4 + 2x_5,\ x_5\Big)
\end{aligned}$$

is a square root of $I + N$.

19. Prove that if V is a complex vector space, then every invertible operator on V has a cube root.

SOLUTION: First suppose that $N \in \mathcal{L}(V)$ is nilpotent. We will show that $I + N$ has a cube root by imitating the proof of 8.30. Specifically, we guess that there is a cube root of $I + N$ of the form

$$I + a_1 N + a_2 N^2 + \cdots + a_{m-1} N^{m-1},$$

where m is such that $N^m = 0$. Having made this guess, we can try to choose $a_1, a_2, \dots, a_{m-1}$ so that the operator above has its cube equal to $I + N$. Now

$$\begin{aligned}(I + a_1N + a_2N^2 + a_3N^3 + \cdots + a_{m-1}N^{m-1})^3 \\ = I + 3a_1N + (3a_2 + 3a_1{}^2)N^2 + (3a_3 + 6a_1a_2 + a_1{}^3)N^3 + \cdots \\ + (3a_{m-1} + \text{terms involving } a_1, \dots, a_{m-2})N^{m-1}.\end{aligned}$$

We want the right side of the equation above to equal $I + N$. Hence choose a_1 so that $3a_1 = 1$ (thus $a_1 = 1/3$). Next, choose a_2 so that $3a_2 + 3a_1{}^2 = 0$ (thus $a_2 = -1/9$). Then choose a_3 so that the coefficient of N^3 on the right side of the equation above equals 0 (thus $a_3 = 5/81$). Continue in this fashion for $j = 4, \dots, m-1$, at each step solving for a_j so that the coefficient of N^j on the right side of the equation above equals 0. Actually we don't care about the explicit formula for the a_j's. We need only know that some choice of the a_j's gives a cube root of $I + N$.

Having shown that the identity plus a nilpotent always has a cube root, we now look at the proof that every invertible operator on a complex vector space has a square root (see 8.32). In that proof, if we replace the words "square root" with "cube root", we get a proof that every invertible operator on a complex vector space has a cube root.

20. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbf{F})$ such that $T^{-1} = p(T)$.

SOLUTION: Let $a_0 + a_1z + \cdots + a_{m-1}z^{m-1} + z^m$ denote the minimal polynomial of T. This is the monic polynomial of smallest degree such that

$$a_0I + a_1T + \cdots + a_{m-1}T^{m-1} + T^m = 0.$$

If a_0 were equal to 0, then we could multiply both sides of the equation above by T^{-1} to get

$$a_1I + a_2T + \cdots + a_{m-1}T^{m-2} + T^{m-1} = 0,$$

which would give a monic polynomial q of smaller degree such that $q(T) = 0$, contradicting the definition of the minimal polynomial. Thus $a_0 \neq 0$.

Because $a_0 \neq 0$, we can solve the first equation above for the identity operator I, getting

$$I = -\frac{a_1}{a_0}T - \cdots - \frac{a_{m-1}}{a_0}T^{m-1} - \frac{1}{a_0}T^m.$$

Now multiply both sides of the equation above by T^{-1}, getting

$$T^{-1} = -\frac{a_1}{a_0}I - \frac{a_2}{a_0}T - \dots - \frac{a_{m-1}}{a_0}T^{m-2} - \frac{1}{a_0}T^{m-1}.$$

Setting

$$p(z) = -\frac{a_1}{a_0} - \frac{a_2}{a_0}z - \dots - \frac{a_{m-1}}{a_0}z^{m-2} - \frac{1}{a_0}z^{m-1},$$

we thus have $T^{-1} = p(T)$.

21. Give an example of an operator on $\mathbf{C}^3$ whose minimal polynomial equals z^2.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{C}^3)$ by

$$T(w_1, w_2, w_3) = (w_3, 0, 0).$$

Clearly $T^2 = 0$. In other words, the polynomial z^2 when applied to T gives 0. Thus the minimal polynomial of T is a divisor of z^2 (by 8.34). But the only monic polynomials that divide z^2 are 1, z, and z^2. The polynomial 1 applied to T gives the identity operator, which is not 0, and the polynomial z applied to T gives T, which is also not 0. Thus the minimal polynomial of T must be z^2.

Of course there are also many other examples.

22. Give an example of an operator on $\mathbf{C}^4$ whose minimal polynomial equals $z(z-1)^2$.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{C}^4)$ by

$$T(w_1, w_2, w_3, w_4) = (0, w_2 + w_4, w_3, w_4).$$

Then

$$(T - I)(w_1, w_2, w_3, w_4) = (-w_1, w_4, 0, 0),$$

which implies that

$$(T - I)^2(w_1, w_2, w_3, w_4) = (w_1, 0, 0, 0),$$

which implies that

$$T(T - I)^2 = 0.$$

In other words, the polynomial $z(z-1)^2$ when applied to T gives 0. Thus the minimal polynomial of T is a divisor of $z(z-1)^2$ (by 8.34).

Note that 0 is an eigenvalue of T because $T(1,0,0,0) = (0,0,0,0)$ and 1 is an eigenvalue of T because $T(0,1,0,0) = (0,1,0,0)$. Thus 0 and 1 must both be roots of the minimal polynomial of T (by 8.36).

The only monic polynomials that divide $z(z-1)^2$ and have $0,1$ as roots are $z(z-1)$ and $z(z-1)^2$. Because $T(T-I) \neq 0$, as is easy to check, this implies that $z(z-1)^2$ is the minimal polynomial of T.

Of course there are also many other examples.

23. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Prove that V has a basis consisting of eigenvectors of T if and only if the minimal polynomial of T has no repeated roots.

COMMENT: For complex vector spaces, this exercise adds another equivalence to the list given by 5.21.

SOLUTION: First suppose that there is a basis $(v_1, \dots, v_n)$ of V consisting of eigenvectors of T. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T. Then for each v_j, there exists λ_k with $(T - \lambda_k I)v_j = 0$. Thus

$$(T - \lambda_1 I) \dots (T - \lambda_m I)v_j = 0$$

for each j (because all the operators in sight commute, for each j the appropriate $T - \lambda_k I$ can be moved to the last position in the product above). An operator that sends each vector in a basis to the 0 vector is the 0 operator, so

$$(T - \lambda_1 I) \dots (T - \lambda_m I) = 0.$$

Thus the polynomial $(z - \lambda_1) \dots (z - \lambda_m)$ when applied to T gives 0. Thus the minimal polynomial of T is a divisor of $(z - \lambda_1) \dots (z - \lambda_m)$ (by 8.34). Because $(z - \lambda_1) \dots (z - \lambda_m)$ has no repeated roots, this implies that the minimal polynomial of T has no repeated roots, as desired. (Because each eigenvalue of T must be a root of the minimal polynomial of T, the minimal polynomial of T actually equals $(z - \lambda_1) \dots (z - \lambda_m)$.)

To prove the implication in the other direction, now suppose that the minimal polynomial of T has no repeated roots. Letting $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T, this means that the minimal polynomial of T equals

$$(z - \lambda_1) \dots (z - \lambda_m).$$

Thus

$$(T - \lambda_1 I) \dots (T - \lambda_m I) = 0.$$

Let U_m be the subspace of generalized eigenvectors corresponding to the eigenvalue λ_m. Recall that U_m is invariant under T (see 8.23(b)). Suppose $v \in U_m$. Let $u = (T - \lambda_m I)v$, so $u \in U_m$. The equation above implies that

$$\begin{aligned}(T|_{U_m} - \lambda_1 I) \dots (T|_{U_m} - \lambda_{m-1} I)u &= (T - \lambda_1 I) \dots (T - \lambda_m I)v \\ &= 0.\end{aligned}$$

Because $(T - \lambda_m I)|_{U_m}$ is nilpotent (see see 8.23(c)), 0 is the only eigenvalue of $(T - \lambda_m I)|_{U_m}$ (this follows from 8.26 and 5.18). Thus $T|_{U_m} - \lambda_j I$ is invertible (as an operator on U_m) for $j = 1, \dots, m-1$. The equation above thus implies that $u = 0$. In other words, v is an eigenvector of T.

We have shown that every generalized eigenvector of T corresponding to the eigenvalue λ_m is an eigenvector of T. There is nothing special about the eigenvalue λ_m—we could have relabeled the eigenvalues so that any of them was called λ_m. Thus we can conclude that every generalized eigenvector of T is actually an eigenvector of T. Because there is a basis of V consisting of generalized eigenvectors of T (see 8.25), this means that there is a basis of V consisting of eigenvectors of T, as desired.

24. Suppose V is an inner-product space. Prove that if $T \in \mathcal{L}(V)$ is normal, then the minimal polynomial of T has no repeated roots.

SOLUTION: Suppose $T \in \mathcal{L}(V)$ is normal. Let p denote the minimal polynomial of T. Suppose $\lambda \in \mathbf{F}$ is an eigenvalue of T. We can write

$$p(z) = (z - \lambda)^m q(z)$$

where m is a positive integer and q is a monic polynomial such that $q(\lambda) \neq 0$. Our goal is to prove that $m = 1$, which implies that p has no repeated roots.

Because $p(T) = 0$, we know that $(T - \lambda I)^m q(T) = 0$. This is equivalent to the statement that

$$\operatorname{range} q(T) \subset \operatorname{null}(T - \lambda I)^m.$$

Because T is normal, so is $T - \lambda I$, and thus $\operatorname{null}(T - \lambda I)^m = \operatorname{null}(T - \lambda I)$ (by Exercise 7 in Chapter 7). Hence the set inclusion above becomes

$$\operatorname{range} q(T) \subset \operatorname{null}(T - \lambda I),$$

which implies that $(T-\lambda I)q(T) = 0$. Letting $p_1(z) = (z-\lambda)q(z)$, this means that p_1 is a monic polynomial with the property that $p_1(T) = 0$. If $m > 1$, then the degree of p_1 would be less than the degree of p, contradicting the definition of minimal polynomial. Thus we can conclude that $m = 1$, as desired.

COMMENT: The proof given above works on both real and complex vector spaces. If V is a complex vector space, then this exercise can be done by using the complex spectral theorem (7.9) and Exercise 23 of this chapter.

25. Suppose $T \in \mathcal{L}(V)$ and $v \in V$. Let p be the monic polynomial of smallest degree such that

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

Prove that p divides the minimal polynomial of T.

SOLUTION: Let q denote the minimal polynomial of T. By the division algorithm (4.5), there exist polynomials $s, r \in \mathcal{P}(\mathbf{F})$ such that

$$q = sp + r$$

and $\deg r < \deg p$. Thus

$$q(T)v = s(T)p(T)v + r(T)v.$$

Because $q(T) = 0$ and $p(T)v = 0$, the equation above shows that $r(T)v = 0$. This implies that $r = 0$ (otherwise we could multiply r by a scalar to get a monic polynomial with degree smaller than $\deg p$ that when applied to T gives an operator having v in its null space, which would contract the definition of p as the monic polynomial of smallest degree with this property).

Using the information that $r = 0$, rewrite the formula above for q as

$$q = sp.$$

Thus p divides q, the minimal polynomial of T.

26. Give an example of an operator on $\mathbf{C}^4$ whose characteristic and minimal polynomials both equal $z(z-1)^2(z-3)$.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{C}^4)$ by

$$T(w_1, w_2, w_3, w_4) = (0, w_2 + w_3, w_3, 3w_4).$$

An easy computation shows that $T(T - I)^2(T - 3I) = 0$. Thus the minimal polynomial of T is a divisor of $z(z-1)^2(z-3)$ (by 8.34).

Note that 0 is an eigenvalue of T because $T(1,0,0,0) = (0,0,0,0)$ and 1 is an eigenvalue of T because $T(0,1,0,0) = (0,1,0,0)$ and 3 is an eigenvalue of T because $T(0,0,0,1) = (0,0,0,3)$. Thus 0, 1, and 3 must be roots of the minimal polynomial of T (by 8.36).

The only monic polynomials that divide $z(z-1)^2(z-3)$ and have $0, 1, 3$ as roots are $z(z-1)(z-3)$ and $z(z-1)^2(z-3)$. Because $T(T-1)(T-3) \neq 0$, as is easy to check, this implies that $z(z-1)^2(z-3)$ is the minimal polynomial of T, as desired.

Because T is an operator on a four-dimensional complex vector space and the minimal polynomial of T has degree 4, the characteristic polynomial of T (which is a monic polynomial of degree 4 that is divisible by the minimal polynomial of T) must equal the minimal polynomial of T.

Of course there are also many other examples.

27. Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $z(z-1)^2(z-3)$ and whose minimal polynomial equals $z(z-1)(z-3)$.

SOLUTION: Define $T \in \mathcal{L}(\mathbf{C}^4)$ by

$$T(w_1, w_2, w_3, w_4) = (0, w_2, w_3, 3w_4).$$

An easy computation shows that $T(T-I)(T-3I) = 0$. Thus the minimal polynomial of T is a divisor of $z(z-1)(z-3)$ (by 8.34).

Note that 0 is an eigenvalue of T because $T(1,0,0,0) = (0,0,0,0)$ and 1 is an eigenvalue of T because $T(0,1,0,0) = (0,1,0,0)$ and 3 is an eigenvalue of T because $T(0,0,0,1) = (0,0,0,3)$. Thus 0, 1, and 3 must be roots of the minimal polynomial of T (by 8.36).

The only monic polynomial that is a divisor of $z(z-1)(z-3)$ and has $0, 1, 3$ as roots is $z(z-1)(z-3)$. Thus $z(z-1)(z-3)$ is the minimal polynomial of T, as desired.

Note that every vector in $\{(0, w_2, w_3, 0) : w_2, w_3 \in \mathbf{C})$ is an eigenvector of T corresponding to the eigenvalue 1. Thus the eigenvalue 1 of T has multiplicity at least 2. The eigenvalues 0 and 3 of T have multiplicity at least 1.

Because T is an operator on a four-dimensional complex vector space, the sum of the multiplicities of all the eigenvalues equals 4 (by 8.18). Thus the each use of the phrase "at least" in the previous paragraph can be replaced by "equal" because if any of the eigenvalues had larger multiplicity, the sum of the multiplicities of all the eigenvalues would exceed 4.

Because 0 and 3 are eigenvalues of T with multiplicity 1 and 1 is an eigenvalue of T with multiplicity 2 and T has no other eigenvalues (the multiplicities of the eigenvalues mentioned already sum to 4), the characteristic polynomial of T equals $z(z-1)^2(z-3)$, as desired.

Of course there are also many other examples.

28. Suppose $a_0, \ldots, a_{n-1} \in \mathbf{C}$. Find the minimal and characteristic polynomials of the operator on $\mathbf{C}^n$ whose matrix (with respect to the standard basis) is

$$\begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{bmatrix}.$$

COMMENT: This exercise shows that every monic polynomial is the characteristic polynomial of some operator.

SOLUTION: Suppose that $T \in \mathcal{L}(\mathbf{C}^n)$ has matrix as above with respect to the standard basis $(e_1, \ldots, e_n)$ of $\mathbf{C}^n$. Thus

$$\begin{aligned} Te_1 &= e_2 \\ T^2 e_1 &= Te_2 = e_3 \\ &\vdots \\ T^{n-1} e_1 &= Te_{n-1} = e_n \\ T^n e_1 &= Te_n = -a_0 e_1 - a_1 e_2 - \cdots - a_{n-1} e_n. \end{aligned}$$

Thus

$$(e_1, Te_1, T^2 e_1, \ldots, T^{n-1} e_1) = (e_1, e_2, e_3, \ldots, e_n).$$

In particular, $(e_1, Te_1, T^2 e_1, \ldots, T^{n-1} e_1)$ is linearly independent. Thus if p is a monic polynomial with degree less than n, then $p(T)e_1 \neq 0$. Thus the minimal polynomial of T must have degree n.

Writing $T^n e_1$ as a linear combination of $(e_1, Te_1, T^2 e_1, \ldots, T^{n-1} e_1)$ is possible in only one way:

$$T^n e_1 = -a_0 e_1 - a_1 Te_1 - \cdots - a_{n-1} T^{n-1} e_1.$$

Thus setting

$$p(z) = a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n,$$

we see that p is the only monic polynomial of degree n such that $p(T)e_1 = 0$. Hence p must equal the minimal polynomial of T.

Because T is an operator on an n-dimensional complex vector space and the minimal polynomial of T has degree n, the characteristic polynomial of T (which is a monic polynomial of degree n that is divisible by the minimal polynomial of T) must equal the minimal polynomial of T. Thus the characteristic polynomial of T also equals p.

29. Suppose $N \in \mathcal{L}(V)$ is nilpotent. Prove that the minimal polynomial of N is z^{m+1}, where m is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the matrix of N with respect to any Jordan basis for N.

SOLUTION: Suppose $(v_1, \dots, v_n)$ is a Jordan basis for N and that m equals the length of the longest consecutive string of 1's that appears on the line directly above the diagonal of $\mathcal{M}\big(N, (v_1, \dots, v_n)\big)$. The diagonal of $\mathcal{M}\big(N, (v_1, \dots, v_n)\big)$ contains only 0's (by Exercise 6 of this chapter). Thus $\mathcal{M}\big(N, (v_1, \dots, v_n)\big)$ is a block diagonal matrix whose blocks have the form

$$\begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix},$$

and the largest such block in an $(m+1)$-by-$(m+1)$ matrix. For each v_j, we see that $N^{m+1}v_j = 0$. Because N^{m+1} equals 0 on a basis of V, we conclude that $N^{m+1} = 0$. Thus the minimal polynomial of N must divide z^{m+1} (by 8.34) and hence must be of the form z^k for some $k \le m+1$. But there is a basis vector v_j such that $N^m v_j = v_{j-m} \neq 0$. Thus $N^m \neq 0$, which implies that the minimal polynomial of N equals z^{m+1}.

30. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Prove that there does not exist a direct sum decomposition of V into two proper subspaces invariant under T if and only if the minimal polynomial of T is of the form $(z - \lambda)^{\dim V}$ for some $\lambda \in \mathbf{C}$.

SOLUTION: First suppose that there does not exist a direct sum decomposition of V into two proper subspaces invariant under T. Thus T has only one eigenvalue (by 8.23), which we will call λ.

There is a Jordan basis of T (by 8.47), meaning that with respect to this basis T has a block diagonal matrix

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is an upper-triangular matrix of the form

$$A_j = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}.$$

If $m > 1$, then we could let U denote the span of the basis vectors corresponding to A_1 and let W denote the span of the basis vectors corresponding to $U_2, \dots, U_m$; we would have $V = U \oplus W$, where U and W would be proper subspaces of V invariant under T. This contradiction shows that $m = 1$. Thus the matrix of T with respect to our Jordan basis is just A_1. In other words, there is a basis $(v_1, \dots, v_{\dim V})$ of V such that the matrix of $T - \lambda I$ with respect to this basis is

$$\begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}.$$

The previous problem now implies that the minimal polynomial of $T - \lambda I$ equals $z^{\dim V}$. This clearly implies that the minimal polynomial of T equals $(z - \lambda)^{\dim V}$, as desired.

To prove the implication in the other direction, suppose that the minimal polynomial of T equals $(z - \lambda)^{\dim V}$. Suppose that there exist two proper subspaces U_1, U_2 of V, each invariant under T, such that $V = U_1 \oplus U_2$.

Let p_1 denote the minimal polynomial of $T|_{U_1}$ and p_2 denote the minimal polynomial of $T|_{U_2}$. If $u_1 \in U_1$, then

$$\begin{aligned} (p_1 p_2)(T)u_1 &= p_2(T)p_1(T)u_1 \\ &= 0. \end{aligned}$$

Similarly, if $u_2 \in U_2$, then

$$\begin{aligned}(p_1p_2)(T)u_2 &= p_1(T)p_2(T)u_2\\ &= 0.\end{aligned}$$

Because every vector in V can be written in the form u_1+u_2, where $u_1 \in U_1$ and $u_2 \in U_2$, the equations above imply that $(p_1p_2)(T) = 0$. Thus the minimal polynomial of T, which equals $(z-\lambda)^{\dim V}$, is a divisor of p_1p_2 (by 8.34).

The degree of the monic polynomial p_1p_2 equals the degree of p_1 plus the degree of p_2, which is less than or equal to $\dim U_1 + \dim U_2$, which equals $\dim V$. Because $(z-\lambda)^{\dim V}$ is a divisor of p_1p_2, this implies that

$$p_1(z)p_2(z) = (z-\lambda)^{\dim V}.$$

This implies that

$$p_1(z) = (z-\lambda)^{\dim U_1} \quad \text{and} \quad p_2(z) = (z-\lambda)^{\dim U_2}.$$

Let $m = \max\{\dim U_1, \dim U_2\}$. Let $p(z) = (z-\lambda)^m$. The equations above show that $p(T)|_{U_1} = 0$ and $p(T)|_{U_2} = 0$. Thus $p(T) = 0$. Because the degree of p is less than $\dim V$, this contradicts our hypothesis that $(z-\lambda)^{\dim V}$ is the minimal polynomial of T. This contradiction means that our assumption that there exists a direct sum decomposition of V into two proper subspaces invariant under T must have been false, as desired.

31. Suppose $T \in \mathcal{L}(V)$ and $(v_1, \ldots, v_n)$ is a basis of V that is a Jordan basis for T. Describe the matrix of T with respect to the basis $(v_n, \ldots, v_1)$ obtained by reversing the order of the v's.

SOLUTION: The matrix of T with respect to $(v_1, \ldots, v_n)$ is a block diagonal matrix

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is an upper-triangular matrix of the form

$$A_j = \begin{bmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_j \end{bmatrix}.$$

Temporarily fix j, and let $(u_1, \dots, u_k)$ be the part of $(v_1, \dots, v_n)$ corresponding to the block A_j. Thus

$$\begin{aligned} Tu_1 &= \lambda_j u_1 \\ Tu_2 &= u_1 + \lambda_j u_2 \\ Tu_3 &= u_2 + \lambda_j u_3 \\ &\vdots \\ Tu_k &= u_{k-1} + \lambda_j u_k. \end{aligned}$$

Write the equations above in the form

$$\begin{aligned} Tu_k &= \lambda_j u_k + u_{k-1} \\ Tu_{k-1} &= \lambda_j u_{k-1} + u_{k-2} \\ &\vdots \\ Tu_2 &= \lambda_j u_2 + u_1 \\ Tu_1 &= \lambda_j u_1. \end{aligned}$$

Thus the matrix of $T|_{\mathrm{span}(u_1,\dots,u_k)}$ with respect to $(u_k, \dots, u_1)$ is

$$B_j = \begin{bmatrix} \lambda_j & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda_j \end{bmatrix}.$$

In other words, B_j is obtained from A_j by reflection across the diagonal, so in B_j the 1's lie below the diagonal instead of above it. Now we see that the matrix of T with respect to $(v_n, \dots, v_1)$ is the block diagonal matrix

$$\begin{bmatrix} B_m & & 0 \\ & \ddots & \\ 0 & & B_1 \end{bmatrix}.$$

CHAPTER 9

Operators on Real Vector Spaces

1. Prove that 1 is an eigenvalue of every square matrix with the property that the sum of the entries in each row equals 1.

SOLUTION: Suppose that A is an n-by-n matrix such that the sum of the entries in each row of A equals 1. Let x be the n-by-1 matrix all of whose entries equal 1. Then from the definition of matrix multiplication we see that the entry in row j, column 1 of Ax equals that sum of the entries in row j of A, which equals 1. Thus $Ax = x$, which implies that 1 is an eigenvalue of A.

2. Consider a 2-by-2 matrix of real numbers

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

Prove that A has an eigenvalue (in $\mathbf{R}$) if and only if

$$(a-d)^2 + 4bc \geq 0.$$

SOLUTION: A number $\lambda \in \mathbf{R}$ is an eigenvalue of A if and only if there exist numbers $x, y \in \mathbf{R}$, not both 0, such that

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}.$$

The left side of this equation equals $\left[\begin{smallmatrix} ax + cy \\ bx + dy \end{smallmatrix} \right]$, so the equation above is equivalent to the system of equations

$$\begin{aligned}(a-\lambda)x + \qquad cy &= 0\\ bx + (d-\lambda)y &= 0.\end{aligned}$$

It is easy to see that this system of equations has a solution other than $x = y = 0$ if and only if

$$(a-\lambda)(d-\lambda) = bc,$$

which is equivalent to the equation

$$\lambda^2 - (a+d)\lambda + (ad - bc) = 0.$$

There is a real number λ satisfying the equation above if and only if

$$(a+d)^2 - 4(ad-bc) \geq 0.$$

The left side of the inequality above equals $(a-d)^2 + 4bc$, and thus we conclude that A has a real eigenvalue if and only if

$$(a-d)^2 + 4bc \geq 0.$$

3. Suppose A is a block diagonal matrix

$$A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a square matrix. Prove that the set of eigenvalues of A equals the union of the eigenvalues of $A_1, \dots, A_m$.

SOLUTION: Suppose that A_j has size n_j-by-n_j. Let

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix},$$

where each x_j is an n_j-by-1 matrix. Then

$$Ax = \begin{bmatrix} A_1x_1 \\ \vdots \\ A_mx_m \end{bmatrix}.$$

Thus the equation $Ax = \lambda x$ is equivalent to the system of equations

$$A_1x_1 = \lambda x_1$$
$$\vdots$$
$$A_m x_m = \lambda x_m.$$

Suppose that λ is an eigenvalue of A. Then there is a nonzero vector x satisfying the system of equations above. Because x is nonzero, there exists k such that x_k is nonzero. Because $A_k x_k = \lambda x_k$, this implies that λ is an eigenvalue of A_k. Thus λ is in the union of the eigenvalues of $A_1, \dots, A_m$, as desired.

To prove the implication in the other direction, suppose now that λ is in the union of the eigenvalues of $A_1, \dots, A_m$. Then there exists k such that λ is an eigenvalue of A_k. Thus there exists a nonzero n_k-by-1 vector x_k such that $A_k x_k = \lambda x_k$. For $j \neq k$, define x_j to be the n_j-by-1 matrix all of whose entries equal 0, and define x to be the matrix determined by $x_1, \dots, x_m$ as above. Then x is nonzero (because x_k is nonzero) and $Ax = \lambda x$, which shows that λ is an eigenvalue of A, as desired.

4. Suppose A is a block upper-triangular matrix

$$A = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a square matrix. Prove that the set of eigenvalues of A equals the union of the eigenvalues of $A_1, \dots, A_m$.

COMMENT: Clearly Exercise 4 is a stronger statement than Exercise 3. Even so, students may want to do Exercise 3 first because it is easier than Exercise 4.

SOLUTION: We will prove that 0 is an eigenvalue of A if and only if 0 is an eigenvalue of at least one of the A_k's. This is all we need to do, because for arbitrary $\lambda \in \mathbf{F}$, we can replace A with $A - \lambda I$ and each A_k with $A_k - \lambda I$, concluding that λ is an eigenvalue of A if and only if λ is an eigenvalue of at least one of the A_k's. This last statement implies that the set of eigenvalues of A equals the union of the eigenvalues of $A_1, \dots, A_m$.

Suppose that A has size n-by-n and each A_j has size n_j-by-n_j. We can write a typical n-by-1 matrix x in the form

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix},$$

where each x_j is an n_j-by-1 matrix. The product Ax can be computed by multiplying together the block matrices of A and x given above, using the same formula as one would use when multiplying matrices of numbers (this follows from the definition of matrix multiplication).

First suppose that 0 is an eigenvalue of A. Thus there exists a nonzero n-by-1 matrix x such that $Ax = 0$. Write x in the form above, and let k be the largest index with nonzero x_k; thus we can write

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

(If $k = m$, then the 0's shown above at the tail of x do not appear.) If we break Ax into blocks of the same size as was done for x, then the k^{th} block of Ax will equal $A_k x_k$; this follows from the block upper-triangular form of A and the 0's that appear in x after the k^{th} block. But $Ax = 0$, so the k^{th} block of Ax equals 0, so $A_k x_k = 0$. Because $x_k \neq 0$, this implies that 0 is an eigenvalue of A_k, as desired.

To prove the implication in the other direction, suppose now that 0 is an eigenvalue of some A_k. This means that the operator on $\mathrm{Mat}(n, 1, \mathbf{F})$ (the vector space of n-by-1 matrices) that sends $x_k \in \mathrm{Mat}(n, 1, \mathbf{F})$ to $A_k x_k$ is not injective. Thus this operator is not surjective (by 3.21). Thus the operator on $\mathrm{Mat}(n_1 + \cdots + n_k, 1, \mathbf{F})$ that sends

$$\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}$$

to

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_k \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}$$

is not surjective (because the last block in the product above will be $A_k x_k$, which cannot be an arbitrary n_k-by-1 matrix). Again using 3.21, this means that the last operator is not injective. In other words, there exists a nonzero vector

$$\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \in \operatorname{Mat}(n_1 + \cdots + n_k, 1, \mathbf{F})$$

such that

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_k \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} = 0.$$

Adjoining an appropriate number of 0's, this implies that

$$\begin{bmatrix} A_1 & & & & & * \\ & \ddots & & & & \\ & & A_k & & & \\ & & & A_{k+1} & & \\ & & & & \ddots & \\ 0 & & & & & A_m \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0.$$

In other words, 0 is an eigenvalue of A, as desired.

5. Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose $\alpha, \beta \in \mathbf{R}$ are such that $T^2 + \alpha T + \beta I = 0$. Prove that T has an eigenvalue if and only if $\alpha^2 \geq 4\beta$.

SOLUTION: First suppose that T has an eigenvalue $\lambda \in \mathbf{R}$. Thus there exists a nonzero vector $v \in V$ such that $Tv = \lambda v$. Applying T to both sides of the last equation, we get $T^2 v = \lambda^2 v$. Thus

$$\begin{aligned} 0 &= (T^2 + \alpha T + \beta I)v \\ &= \lambda^2 v + \alpha\lambda v + \beta v \\ &= (\lambda^2 + \alpha\lambda + \beta)v. \end{aligned}$$

Because $v \neq 0$, the last equation implies that

$$\lambda^2 + \alpha\lambda + \beta = 0,$$

which implies (recall that λ, α, and β are all real) that $\alpha^2 \geq 4\beta$, as desired.

To prove the implication in the opposite direction, suppose now that $\alpha^2 \geq 4\beta$. Then the polynomial $x^2 + \alpha x + \beta$ has two real roots, which means that we can write

$$x^2 + \alpha x + \beta = (x - r_1)(x - r_2)$$

for some $r_1, r_2 \in \mathbf{R}$. Thus

$$\begin{aligned} 0 &= T^2 + \alpha T + \beta I \\ &= (T - r_1 I)(T - r_2 I). \end{aligned}$$

In particular, $(T - r_1 I)(T - r_2 I)$ is not injective, which implies that at least one of $T - r_1 I$ and $T - r_2 I$ is not injective. In other words, at least one of r_1, r_2 must be an eigenvalue of T. Thus T has an eigenvalue, as desired.

6. Suppose V is a real inner-product space and $T \in \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which T has a block upper-triangular matrix

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues.

SOLUTION: We know that there is a basis $(v_1, \ldots, v_n)$ of V with respect to which the matrix of T has the block upper-triangular form above, where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues (see 9.4).

Apply the Gram-Schmidt procedure to $(v_1, \ldots, v_n)$, getting an orthonormal basis $(e_1, \ldots, e_n)$ of V such that

$$\operatorname{span}(v_1, \ldots, v_j) = \operatorname{span}(e_1, \ldots, e_j)$$

for $j = 1, \ldots, n$ (see 6.20). This condition on the spans implies that the matrix of T with respect to $(e_1, \ldots, e_n)$ is also a block upper-triangular of the form above, where each A_j is a 1-by-1 matrix or a 2-by-2 matrix (these A_j's may differ from the A_j's corresponding to the matrix of T with respect to $(v_1, \ldots, v_n)$). All that remains is to show that, if necessary, we can modify our orthonormal basis so that none of the 2-by-2 blocks on the diagonal have eigenvalues.

Suppose that A_j is a 2-by-2 block on the diagonal of the matrix of T with respect to $(e_1, \ldots, e_n)$ and that A_j has an eigenvalue λ. Thus there exist $x, y \in \mathbf{R}$, not both 0, such that

$$A_j \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Let e_k, e_{k+1} denote the basis vectors corresponding to A_j. Then, as is easy to verify,

$$T(xe_k + ye_{k+1}) = u + \lambda(xe_k + ye_{k+1})$$

for some $u \in \operatorname{span}(e_1, \ldots, e_{k-1})$. Let

$$f_k = \frac{xe_k + ye_{k+1}}{\|xe_k + ye_{k+1}\|}$$

and choose $f_{k+1} \in V$ such that (f_k, f_{k+1}) is orthonormal and

$$\operatorname{span}(f_k, f_{k+1}) = \operatorname{span}(e_k, e_{k+1}).$$

Then the matrix of T with respect to the orthonormal basis

$$(e_1, \ldots, e_{k-1}, f_k, f_{k+1}, e_{k+2}, \ldots, e_n)$$

will be a block upper-triangular matrix with the same entries on the diagonal as previously, except that A_j will be replaced by the 2-by-2 matrix

$$\begin{bmatrix} \lambda & * \\ 0 & * \end{bmatrix}.$$

In other words, where we had A_j on the diagonal, we can now think of having two 1-by-1 matrices on the diagonal (and we still have a block upper-triangular matrix because of the 0 in the lower-left entry of the matrix above).

Repeating, when necessary, the procedure described above, we obtain an orthonormal basis of V with respect to which T has a block upper-triangular matrix

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues.

7. Prove that if $T \in \mathcal{L}(V)$ and j is a positive integer such that $j \leq \dim V$, then T has an invariant subspace whose dimension equals $j - 1$ or j.

SOLUTION: Suppose $T \in \mathcal{L}(V)$ and j is a positive integer such that $j \leq \dim V$. If V is a complex vector space, then T has an invariant subspace whose dimension equals j (by Exercise 17 in Chapter 5). So we can assume that V is a real vector space.

By 9.4, there is a basis $(v_1, \dots, v_n)$ of V with respect to which T has a block upper-triangular matrix

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_k is a 1-by-1 matrix or a 2-by-2 matrix. Either v_{j-1} or v_j must be the last vector in a block of vectors corresponding to some A_k. Thus either $\operatorname{span}(v_1, \dots, v_{j-1})$ or $\operatorname{span}(v_1, \dots, v_j)$ must be invariant under T.

8. Prove that there does not exist an operator $T \in \mathcal{L}(\mathbf{R}^7)$ such that T^2+T+I is nilpotent.

SOLUTION: Suppose $T \in \mathcal{L}(\mathbf{R}^7)$. From part (b) of 9.9 (combined with 9.4, which insures that T has a matrix of the form 9.10 with respect to some basis of $\mathbf{R}^7$), we see that

$$\dim \operatorname{null}(T^2 + T + I)^7$$

must be an even integer. In particular, $\dim \operatorname{null}(T^2 + T + I)^7 \neq 7$, which implies that $\operatorname{null}(T^2 + T + I)^7 \neq \mathbf{R}^7$, which implies that $(T^2 + T + I)^7 \neq 0$, which implies that T is not nilpotent (by 8.8).

9. Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^7)$ such that T^2+T+I is nilpotent.

SOLUTION: Let $\lambda = (-1 + 3i)/2$. Define $T \in \mathcal{L}(\mathbf{C}^7)$ by $T = \lambda I$. Then

$$\begin{aligned} T^2 + T + I &= (\lambda^2 + \lambda + 1)I \\ &= 0. \end{aligned}$$

In particular, $T^2 + T + I$ is nilpotent.

10. Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose $\alpha, \beta \in \mathbf{R}$ are such that $\alpha^2 < 4\beta$. Prove that

$$\operatorname{null}(T^2 + \alpha T + \beta I)^k$$

has even dimension for every positive integer k.

SOLUTION: Let k be a positive integer, and let $U = \operatorname{null}(T^2+\alpha T+\beta I)^k$. We need to prove that $\dim U$ is even.

Because U is invariant under T (by 8.22, with $p(x) = (x^2 + \alpha x + \beta)^k$), we can define $S \in \mathcal{L}(U)$ by $S = T|_U$. Clearly $(S^2 + \alpha S + \beta I)^k$ is the 0 operator on U. Thus $S^2 + \alpha S + \beta I$ is a nilpotent operator on U, which implies that $\operatorname{null}(S^2 + \alpha S + \beta I)^{\dim U} = U$ (by 8.8). Now part (b) of 9.9, applied to S and U instead of T and V, shows that $\dim U$ is an even integer, as desired.

11. Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose $\alpha, \beta \in \mathbf{R}$ are such that $\alpha^2 < 4\beta$ and $T^2 + \alpha T + \beta I$ is nilpotent. Prove that $\dim V$ is even and

$$(T^2 + \alpha T + \beta I)^{\dim V/2} = 0.$$

SOLUTION: Let $S = T^2 + \alpha T + \beta I$. Because S is nilpotent, there is a smallest positive integer m such that $S^m = 0$. Thus

$$\{0\} = \operatorname{null} S^0 \subsetneq \operatorname{null} S \subsetneq \operatorname{null} S^2 \subsetneq \dots \subsetneq \operatorname{null} S^m = V,$$

where the proper inclusions come from 8.5. The previous exercise states that each $\operatorname{null} S^k$ has even dimension (in particular V, which equals $\operatorname{null} S^m$, has even dimension). Hence the dimension must increase by at least 2 in all the proper inclusions above. Thus $\dim S^m \geq 2m$, which clearly implies that $m \leq (\dim V)/2$. Because $S^m = 0$, this implies that $S^{\dim V/2} = 0$, as desired.

12. Prove that if $T \in \mathcal{L}(\mathbf{R}^3)$ and $5, 7$ are eigenvalues of T, then T has no eigenpairs.

SOLUTION: Suppose that $T \in \mathcal{L}(\mathbf{R}^3)$ and $5, 7$ are eigenvalues of T. Of course each of these two eigenvalues must have multiplicity at least 1. By 9.17, the sum of the multiplicities of all the eigenvalues of T plus twice the sum of the multiplicities of all the eigenpairs of T equals 3. Because the sum of the multiplicities of all the eigenvalues of T is at least 2, there is no room for an eigenpair, which would add at least 2 more to the sum (because we take twice the sum of the multiplicities of all the eigenpairs). Thus T has no eigenpairs.

13. Suppose V is a real vector space with $\dim V = n$ and $T \in \mathcal{L}(V)$ is such that

$$\operatorname{null} T^{n-2} \neq \operatorname{null} T^{n-1}.$$

Prove that T has at most two distinct eigenvalues and that T has no eigenpairs.

SOLUTION: Because $\operatorname{null} T^{n-2} \neq \operatorname{null} T^{n-1}$, we see that $\dim \operatorname{null} T^j$ is at least 1 more than $\dim \operatorname{null} T^{j-1}$ for $j = 1, \dots, n-1$ (by 8.5). Thus

$\dim \operatorname{null} T^{n-1} \geq n - 1$. In particular, 0 is an eigenvalue of T with multiplicity at least $n - 1$. Because the sum of the multiplicities of all the eigenvalues of T plus the sum of twice the multiplicities of all the eigenpairs of T equals n (by 9.17), this implies that T can have no eigenpairs and at most one additional eigenvalue.

14. Suppose V is a vector space with dimension 2 and $T \in \mathcal{L}(V)$. Prove that if

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

is the matrix of T with respect to some basis of V, then the characteristic polynomial of T equals $(z - a)(z - d) - bc$.

COMMENT: As usual unless otherwise specified, here V may be a real or complex vector space.

SOLUTION: Let $q(z) = (z - a)(z - d) - bc$. If (v_1, v_2) is the basis of V with respect to which T has the matrix above, then

$$Tv_1 = av_1 + bv_2 \quad \text{and} \quad Tv_2 = cv_1 + dv_2.$$

From these equations you can easily verify that $q(T)v_1 = 0$ and $q(T)v_2 = 0$. Because $q(T)$ is 0 on a basis of V, we conclude that $q(T) = 0$.

Because q is a monic polynomial of degree 2 and $q(T) = 0$, we conclude that the minimal polynomial of T has degree 1 or 2.

Suppose first that the minimal polynomial of T has degree 2. Because the minimal polynomial of T is a divisor of q (by 8.34), and because a monic polynomial of degree 2 can be a divisor of another monic polynomial of degree 2 only if the two polynomials are equal, we conclude that q is the minimal polynomial of T. The Cayley-Hamilton theorem now implies that q is a divisor of the characteristic polynomial of T, which is also a monic polynomial of degree 2. This implies that q is the characteristic polynomial of T, as desired.

Now consider the only remaining possible case, which is that the minimal polynomial of T has degree 1, meaning that it equals $z - \lambda$ for some $\lambda \in \mathbf{F}$. This implies that $T = \lambda I$, which implies that the characteristic polynomial of T equals $(z-\lambda)^2$. Because $T = \lambda I$, we must have $a = d = \lambda$ and $b = c = 0$. Thus $q(z) = (z - \lambda)^2$. In particular, q is the characteristic polynomial of T, as desired.

COMMENT: Note that we did not need to find the eigenvalues of T to do this exercise.

15. Suppose V is a real inner-product space and $S \in \mathcal{L}(V)$ is an isometry. Prove that if (α, β) is an eigenpair of S, then $\beta = 1$.

SOLUTION: There is a basis of V with respect to which S has a block diagonal matrix, where each block on the diagonal is a 1-by-1 matrix or a 2-by-2 matrix of the form

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

with $\theta \in (0, \pi)$ (see 7.38). The characteristic polynomial of the matrix above is $(x - \cos\theta)^2 + \sin^2\theta$, which equals $x^2 - 2(\cos\theta)x + 1$.

If (α, β) is an eigenpair of S, then

$$\dim \operatorname{null}(S^2 + \alpha S + \beta I)^{\dim V} > 0,$$

which implies that $x^2 + \alpha x + \beta$ is the characteristic polynomial of a 2-by-2 matrix of the form displayed above (see 9.9). Thus $\beta = 1$.

Chapter 10

Trace and Determinant

1. Suppose that $T \in \mathcal{L}(V)$ and $(v_1, \dots, v_n)$ is a basis of V. Prove that $\mathcal{M}(T, (v_1, \dots, v_n))$ is invertible if and only if T is invertible.

SOLUTION: First suppose that $\mathcal{M}(T)$ is an invertible matrix (because the only basis is sight is $(v_1, \dots, v_n)$, we can leave the basis out of the notation). Thus there exists an n-by-n matrix B such that

$$\mathcal{M}(T)B = B\mathcal{M}(T) = I.$$

There exists an operator $S \in \mathcal{L}(V)$ such that $\mathcal{M}(S) = B$ (see 3.19). Thus the equation above becomes

$$\mathcal{M}(T)\mathcal{M}(S) = \mathcal{M}(S)\mathcal{M}(T) = I,$$

which we can rewrite as

$$\mathcal{M}(TS) = \mathcal{M}(ST) = \mathcal{M}(I),$$

which implies that

$$TS = ST = I.$$

Thus T is invertible, as desired, with inverse S.

To prove the implication in the other direction, suppose now that T is invertible. Thus there exists $S \in \mathcal{L}(V)$ such that

$$TS = ST = I.$$

This implies that

$$\mathcal{M}(TS) = \mathcal{M}(ST) = \mathcal{M}(I),$$

which implies that

$$\mathcal{M}(T)\mathcal{M}(S) = \mathcal{M}(S)\mathcal{M}(T) = I.$$

Thus $\mathcal{M}(T)$ is invertible, as desired, with inverse $\mathcal{M}(S)$.

2. Prove that if A and B are square matrices of the same size and $AB = I$, then $BA = I$.

SOLUTION: Suppose that A and B are n-by-n matrices and $AB = I$. There exist $S, T \in \mathcal{L}(\mathbf{F}^n)$ such that

$$\mathcal{M}(S) = A \quad \text{and} \quad \mathcal{M}(T) = B;$$

here we are using the standard basis of $\mathbf{F}^n$ (the existence of $S, T \in \mathcal{L}(\mathbf{F}^n)$ satisfying the equations above follows from 3.19). Because $AB = I$, we have $\mathcal{M}(S)\mathcal{M}(T) = I$, which implies that $\mathcal{M}(ST) = \mathcal{M}(I)$, which implies that $ST = I$, which implies that $TS = I$ (by Exercise 23 in Chapter 3). Thus

$$\begin{aligned} BA &= \mathcal{M}(T)\mathcal{M}(S) \\ &= \mathcal{M}(TS) \\ &= \mathcal{M}(I) \\ &= I. \end{aligned}$$

3. Suppose $T \in \mathcal{L}(V)$ has the same matrix with respect to every basis of V. Prove that T is a scalar multiple of the identity operator.

SOLUTION: We begin by proving that (v, Tv) is linearly dependent for every $v \in V$. To do this, fix $v \in V$, and suppose that (v, Tv) is linearly independent. Then (v, Tv) can be extended to a basis $(v, Tv, u_1, \ldots, u_n)$ of V. The first column of the matrix of T with respect to this basis is

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Clearly $(2v, Tv, u_1, \ldots, u_n)$ is also a basis of V. The first column of the matrix of T with respect to this basis is

$$\begin{bmatrix} 0 \\ 2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus T has different matrices with respect to the two bases we have considered. This contradiction shows that (v, Tv) is linearly dependent for every $v \in V$. This implies that for every vector in V is an eigenvector of T. This implies that T is a scalar multiple of the identity operator (by Exercise 12 in Chapter 5).

4. Suppose that $(u_1, \dots, u_n)$ and $(v_1, \dots, v_n)$ are bases of V. Let $T \in \mathcal{L}(V)$ be the operator such that $Tv_k = u_k$ for $k = 1, \dots, n$. Prove that

$$\mathcal{M}\big(T, (v_1, \dots, v_n)\big) = \mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big).$$

SOLUTION: Fix k. Write

$$u_k = a_1 v_1 + \cdots + a_n v_n,$$

where $a_1, \dots, a_n \in \mathbf{F}$. Because $Tv_k = u_k$, the k^{th} column of the matrix $\mathcal{M}\big(T, (v_1, \dots, v_n)\big)$ consists of the numbers $a_1, \dots, a_n$. Because $Iu_k = u_k$, the k^{th} column of $\mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)$ also consists of the numbers $a_1, \dots, a_n$.

Because $\mathcal{M}\big(T, (v_1, \dots, v_n)\big)$ and $\mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)$ have the same columns, these two matrices must be equal.

5. Prove that if B is a square matrix with complex entries, then there exists an invertible square matrix A with complex entries such that $A^{-1}BA$ is an upper-triangular matrix.

SOLUTION: Suppose B is an n-by-n matrix with complex entries. Let $(e_1, \dots, e_n)$ denote the standard basis of $\mathbf{C}^n$. There exists $T \in \mathcal{L}(\mathbf{C}^n)$ such that $\mathcal{M}\big(T, (e_1, \dots, e_n)\big) = B$ (see 3.19).

There is a basis $(v_1, \dots, v_n)$ of V such that $\mathcal{M}\big(T, (v_1, \dots, v_n)\big)$ is an upper-triangular matrix (see 5.13). Let $A = \mathcal{M}\big((v_1, \dots, v_n), (e_1, \dots, e_n)\big)$. Then A is invertible (by 10.2) and

$$\begin{aligned} A^{-1}BA &= A^{-1}\mathcal{M}\big(T, (e_1, \dots, e_n)\big)A \\ &= \mathcal{M}\big(T, (v_1, \dots, v_n)\big), \end{aligned}$$

where the second equality comes from 10.3. Thus $A^{-1}BA$ is an upper-triangular matrix.

6. Give an example of a real vector space V and $T \in \mathcal{L}(V)$ such that

$$\operatorname{trace}(T^2) < 0.$$

SOLUTION: Define $T \in \mathcal{L}(\mathbf{R}^2)$ by

$$T(x, y) = (-y, x).$$

Then $T^2 = -I$, so $\operatorname{trace}(T^2) = -2$.

7. Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and V has a basis consisting of eigenvectors of T. Prove that $\operatorname{trace}(T^2) \geq 0$.

SOLUTION: Let $(v_1, \dots, v_n)$ be a basis of V consisting of eigenvectors of T. Thus there exist $\lambda_1, \dots, \lambda_n \in \mathbf{R}$ such that $Tv_j = \lambda_j v_j$ for each j. Clearly the matrix of T^2 with respect to the basis $(v_1, \dots, v_n)$ is the diagonal matrix

$$\begin{bmatrix} {\lambda_1}^2 & & 0 \\ & \ddots & \\ 0 & & {\lambda_n}^2 \end{bmatrix}.$$

Thus $\operatorname{trace} T^2 = {\lambda_1}^2 + \cdots + {\lambda_n}^2 \geq 0$.

8. Suppose V is an inner-product space and $v, w \in \mathcal{L}(V)$. Define $T \in \mathcal{L}(V)$ by $Tu = \langle u, v \rangle w$. Find a formula for $\operatorname{trace} T$.

SOLUTION: First suppose that $v \neq 0$. Extend $(\frac{v}{\|v\|})$ to an orthonormal basis $(\frac{v}{\|v\|}, e_1, \dots, e_n)$ of V. Note that for each j, we have $Te_j = 0$ (because $\langle e_j, v \rangle = 0$). The trace of T equals the sum of the diagonal entries in the matrix of T with respect to the basis $(\frac{v}{\|v\|}, e_1, \dots, e_n)$. Thus

$$\begin{aligned} \operatorname{trace} T &= \langle T(\frac{v}{\|v\|}), \frac{v}{\|v\|} \rangle + \langle Te_1, e_1 \rangle + \cdots + \langle Te_n, e_n \rangle \\ &= \langle \langle \frac{v}{\|v\|}, v \rangle w, \frac{v}{\|v\|} \rangle \\ &= \langle w, v \rangle. \end{aligned}$$

If $v = 0$, then $T = 0$ and so $\operatorname{trace} T = 0 = \langle w, v \rangle$. Thus we have the formula

$$\operatorname{trace} T = \langle w, v \rangle$$

regardless of whether or not $v = 0$.

9. Prove that if $P \in \mathcal{L}(V)$ satisfies $P^2 = P$, then $\operatorname{trace} P$ is a nonnegative integer.

SOLUTION: Suppose that $T \in \mathcal{L}(V)$ satisfies $P^2 = P$. Let $(u_1, \dots, u_m)$ be a basis of $\operatorname{range} P$ and let $(v_1, \dots, v_n)$ be a basis of $\operatorname{null} P$. Then

$$(u_1, \dots, u_m, v_1, \dots, v_n)$$

is a basis of V (this holds because $V = \operatorname{range} T \oplus \operatorname{null} T$; see Exercise 21 in Chapter 5). For each u_j we have $Pu_j = u_j$ and for each v_k we have $Pv_k = 0$. Thus the matrix of P with respect to the basis above of V is a diagonal matrix whose diagonal contains m 1's followed by n 0's. Thus $\operatorname{trace} P = m$, which is a nonnegative integer, as desired. In fact, we have shown that

$$\operatorname{trace} P = \dim \operatorname{range} P.$$

10. Prove that if V is an inner-product space and $T \in \mathcal{L}(V)$, then

$$\operatorname{trace} T^* = \overline{\operatorname{trace} T}.$$

SOLUTION: Suppose that V is an inner-product space and $T \in \mathcal{L}(V)$. Let $(e_1, \dots, e_n)$ be an orthonormal basis of V. The trace of any operator on V equals the sum of the diagonal entries on the matrix of the operator with respect to this basis. Thus

$$\begin{aligned}
\operatorname{trace} T^* &= \langle T^* e_1, e_1 \rangle + \cdots + \langle T^* e_n, e_n \rangle \\
&= \langle e_1, Te_1 \rangle + \cdots + \langle e_n, Te_n \rangle \\
&= \overline{\langle Te_1, e_1 \rangle} + \cdots + \overline{\langle Te_n, e_n \rangle} \\
&= \overline{\langle Te_1, e_1 \rangle + \cdots + \langle Te_n, e_n \rangle} \\
&= \overline{\operatorname{trace} T}.
\end{aligned}$$

11. Suppose V is an inner-product space. Prove that if $T \in \mathcal{L}(V)$ is a positive operator and $\operatorname{trace} T = 0$, then $T = 0$.

SOLUTION: Suppose $T \in \mathcal{L}(V)$ is a positive operator and $\operatorname{trace} T = 0$. There exists an operator $S \in \mathcal{L}(V)$ such that $T = S^* S$ (by 7.27). Let $(e_1, \dots, e_n)$ be an orthonormal basis of V. Then

$$\begin{aligned}
0 &= \operatorname{trace} T \\
&= \langle Te_1, e_1\rangle + \cdots + \langle Te_n, e_n\rangle \\
&= \langle S^*Se_1, e_1\rangle + \cdots + \langle S^*Se_n, e_n\rangle \\
&= \|Se_1\|^2 + \cdots + \|Se_n\|^2.
\end{aligned}$$

The equation above implies that $Se_j = 0$ for each j. Because S is 0 on a basis of V, we have $S = 0$. Because $T = S^*S$, this implies that $T = 0$.

12. Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator whose matrix is

$$\begin{bmatrix} 51 & -12 & -21 \\ 60 & -40 & -28 \\ 57 & -68 & 1 \end{bmatrix}.$$

Someone tells you (accurately) that -48 and 24 are eigenvalues of T. Without using a computer or writing anything down, find the third eigenvalue of T.

SOLUTION: The sum of the eigenvalues of T equals the sum of the diagonal terms of the matrix above (both quantities equal $\operatorname{trace} T$). The sum of the diagonal terms of the matrix above equals 12. The sum of two of the eigenvalues of T, -48 and 24, equals -24. Because the sum of all three eigenvalues of T must equal 12, the third eigenvalue of T must be 36.

13. Prove or give a counterexample: if $T \in \mathcal{L}(V)$ and $c \in \mathbf{F}$, then $\operatorname{trace}(cT) = c \operatorname{trace} T$.

SOLUTION: Suppose $T \in \mathcal{L}(V)$ and $c \in \mathbf{F}$. To prove that $\operatorname{trace}(cT) = c \operatorname{trace} T$, consider a basis of V. Then $\operatorname{trace} T$ equals the sum of the diagonal terms of the matrix of T with respect to this basis. The matrix of cT, with respect to the same basis, equals c times the matrix of T. Thus the sum of the diagonal terms of the matrix of cT equals c times the sum of the diagonal terms of the matrix of T. In other words, $\operatorname{trace}(cT) = c \operatorname{trace} T$.

14. Prove or give a counterexample: if $S, T \in \mathcal{L}(V)$, then

$$\operatorname{trace}(ST) = (\operatorname{trace} S)(\operatorname{trace} T).$$

SOLUTION: Define $S, T \in \mathcal{L}(\mathbf{F}^2)$ by $S(x, y) = T(x, y) = (-y, x)$. Then with respect to the standard bases the matrix of S (which of course equals the matrix of T) is

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

Thus $\operatorname{trace} S = \operatorname{trace} T = 0$. However, $ST = -I$, so $\operatorname{trace} ST = -2$. Thus for this choice of S and T, we have $\operatorname{trace}(ST) \neq (\operatorname{trace} S)(\operatorname{trace} T)$.

Of course there are also many other examples.

15. Suppose $T \in \mathcal{L}(V)$. Prove that if $\operatorname{trace}(ST) = 0$ for all $S \in \mathcal{L}(V)$, then $T = 0$.

SOLUTION: Suppose that $\operatorname{trace}(ST) = 0$ for all $S \in \mathcal{L}(V)$. Then $\operatorname{trace}(TS) = 0$ for all $S \in \mathcal{L}(V)$ (by 10.12). Suppose that there exists $v \in V$ such that $Tv \neq 0$. Then (Tv) can be extended to a basis $(Tv, u_1, \dots, u_n)$ of V. Define $S \in \mathcal{L}(V)$ by

$$S(aTv + b_1u_1 + \cdots + b_nu_n) = av.$$

Thus $S(Tv) = v$ and $Su_j = 0$ for each j. Hence $(TS)(Tv) = T(S(Tv)) = Tv$ and $(TS)(u_j) = 0$ for each j. This implies that with respect to the basis $(Tv, u_1, \dots, u_n)$, the matrix of TS consists of all 0's except for a 1 in the upper-left corner. Thus $\operatorname{trace}(TS) = 1$. This contradiction shows that our assumption that $Tv \neq 0$ must have been false. Thus $Tv = 0$ for every $v \in V$, which means that $T = 0$.

16. Suppose V is an inner-product space and $T \in \mathcal{L}(V)$. Prove that if $(e_1, \dots, e_n)$ is an orthonormal basis of V, then

$$\operatorname{trace}(T^*T) = \|Te_1\|^2 + \cdots + \|Te_n\|^2.$$

Conclude that the right side of the equation above is independent of which orthonormal basis $(e_1, \dots, e_n)$ is chosen for V.

SOLUTION: Suppose that $(e_1, \dots, e_n)$ is an orthonormal basis of V. Then

$$\begin{aligned} \operatorname{trace} T^*T &= \langle T^*Te_1, e_1\rangle + \cdots + \langle T^*Te_n, e_n\rangle \\ &= \langle Te_1, Te_1\rangle + \cdots + \langle Te_n, Te_n\rangle \\ &= \|Te_1\|^2 + \cdots + \|Te_n\|^2. \end{aligned}$$

Because $\operatorname{trace} T^*T$ does not depend upon the choice of a basis of V, the formula above shows that $\|Te_1\|^2 + \cdots + \|Te_n\|^2$ is independent of the orthonormal basis $(e_1, \dots, e_n)$.

17. Suppose V is a complex inner-product space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, repeated according to multiplicity. Suppose

$$\begin{bmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \ldots & a_{n,n} \end{bmatrix}$$

is the matrix of T with respect to some orthonormal basis of V. Prove that

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \le \sum_{k=1}^{n} \sum_{j=1}^{n} |a_{j,k}|^2.$$

SOLUTION: Suppose that $(e_1, \ldots, e_n)$ is the orthonormal basis with respect to which T has the matrix above. Thus for each k, we have

$$Te_k = a_{1,k}e_1 + \cdots + a_{n,k}e_n,$$

which implies that

$$\|Te_k\|^2 = |a_{1,k}|^2 + \cdots + |a_{n,k}|^2.$$

Thus

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \sum_{k=1}^{n} \sum_{j=1}^{n} |a_{j,k}|^2.$$

By the previous exercise, the left side of this equation equals trace(T^*T). This reduces the exercise at hand to proving that

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \le \operatorname{trace}(T^*T).$$

There is an orthonormal basis $(f_1, \ldots, f_n)$ with respect to which T has an upper-triangular matrix (by 6.28). The diagonal entries of the matrix of T with respect to $(f_1, \ldots, f_n)$ are precisely $\lambda_1, \ldots, \lambda_n$ (by 8.10), where we can relabel the eigenvalues of T so that they appear in the order $\lambda_1, \ldots, \lambda_n$ along the diagonal. In other words,

$$\mathcal{M}(T, (f_1, \ldots, f_n)) = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

From the matrix above, we see that

$$|\lambda_k|^2 \le \|Tf_k\|^2$$

for each k. Thus

$$\begin{aligned}|\lambda_1|^2 + \dots + |\lambda_n|^2 &\le \|Tf_1\|^2 + \dots + \|Tf_n\|^2\\ &= \operatorname{trace}(T^*T),\end{aligned}$$

as desired (here the last equality comes from the previous exercise).

18. Suppose V is an inner-product space. Prove that

$$\langle S, T\rangle = \operatorname{trace}(ST^*)$$

defines an inner product on $\mathcal{L}(V)$.

SOLUTION: Suppose that $\langle\cdot,\cdot\rangle$ is defined as above and $R, S, T \in \mathcal{L}(V)$. Then $\langle T, T\rangle = \operatorname{trace}(TT^*)$, and thus $\langle T, T\rangle \ge 0$ (by the formula given in Exercise 16 of this chapter, with T replaced with T^*). Because TT^* is a positive operator (see 7.27), we also see that $\langle T, T\rangle = 0$ if and only if $T = 0$ (by Exercise 11 of this chapter).

Now

$$\begin{aligned}\langle R+S, T\rangle &= \operatorname{trace}\big((R+S)T^*\big)\\ &= \operatorname{trace}(RT^* + ST^*)\\ &= \operatorname{trace}(RT^*) + \operatorname{trace}(ST^*)\\ &= \langle R, T\rangle + \langle S, T\rangle,\end{aligned}$$

where the third equality comes from 10.12.

For $c \in \mathbf{F}$, we have

$$\begin{aligned}\langle cS, T\rangle &= \operatorname{trace}(cST^*)\\ &= c\operatorname{trace}(ST^*)\\ &= c\langle S, T\rangle,\end{aligned}$$

where the second equality comes from Exercise 13 of this chapter.

Finally,

$$\begin{aligned}\langle S, T\rangle &= \operatorname{trace}(ST^*)\\ &= \overline{\operatorname{trace}\big((ST^*)^*\big)}\\ &= \overline{\operatorname{trace}(TS^*)}\\ &= \overline{\langle S, T\rangle},\end{aligned}$$

where the second equality comes from Exercise 10 of this chapter.

We have shown that $\langle \cdot, \cdot \rangle$ satisfies all the properties required of an inner product.

COMMENT: Suppose $(e_1, \dots, e_n)$ is an orthonormal basis of V and

$$\begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix}$$

is the matrix of T with respect to this basis. Then

$$\begin{aligned} \langle T, T \rangle &= \operatorname{trace}(TT^*) \\ &= \operatorname{trace}(T^*T) \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} |a_{j,k}|^2, \end{aligned}$$

where the second equality comes from 10.12 and the third equality comes from Exercise 16 of this chapter. Thus the norm on $\mathcal{L}(V)$ induced by $\langle \cdot, \cdot \rangle$ is the same as the standard norm on $\mathbf{F}^{n^2}$ (here we are identifying each operator with its matrix, which has n^2 entries). Because norms determine the inner product (see Exercises 6 and 7 in Chapter 6), this means that the inner product $\langle \cdot, \cdot \rangle$ is the same as the standard inner product of $\mathbf{F}^{n^2}$ (again using the identification via matrices with respect to the orthonormal basis $(e_1, \dots, e_n)$).

19. Suppose V is an inner-product space and $T \in \mathcal{L}(V)$. Prove that if

$$\|T^*v\| \leq \|Tv\|$$

for every $v \in V$, then T is normal.

SOLUTION: Suppose that $\|T^*v\| \leq \|Tv\|$ for every $v \in V$. Suppose $u \in V$ with $\|u\| = 1$. Extend (u) to an orthonormal basis $(u, e_1, \dots, e_n)$ of V. Then

$$\begin{aligned} \operatorname{trace}(TT^*) &= \|T^*u\|^2 + \|T^*e_1\|^2 + \dots + \|T^*e_n\|^2 \\ &\leq \|Tu\|^2 + \|Te_1\|^2 + \dots + \|Te_n\|^2 \\ &= \operatorname{trace}(T^*T) \\ &= \operatorname{trace}(TT^*), \end{aligned}$$

where the first and third lines come from Exercise 16 of this chapter and the last line comes from 10.12. Because the first and last lines above are equal, we must have equality throughout. Thus $\|T^*u\| = \|Tu\|$. This clearly implies that

$$\|T^*(au)\| = \|T(au)\|$$

for every $a \in \mathbf{F}$. Because every vector in V can be written in the form au for some $a \in \mathbf{F}$ and some $u \in V$ with $\|u\| = 1$, this implies that $\|T^*v\| = \|Tv\|$ for every $v \in V$. This implies that T is normal (by 7.6).

COMMENT: This exercise fails on infinite-dimensional inner-product spaces, leading to what are called hyponormal operators, which have a well-developed theory.

20. Prove or give a counterexample: if $T \in \mathcal{L}(V)$ and $c \in \mathbf{F}$, then $\det(cT) = c^{\dim V} \det T$.

SOLUTION: Let $n = \dim V$. If A is an n-by-n matrix, then

$$\det(cA) = c^n \det A$$

for every $c \in \mathbf{F}$ (this follows immediately from the definition 10.25). Now suppose that $T \in \mathcal{L}(V)$. Because $\det T$ equals the determinant of the matrix of T with respect to any basis (see 10.33), the equation above implies that for every $c \in \mathbf{F}$ we have

$$\begin{aligned}\det(cT) &= \det \mathcal{M}(cT)\\ &= \det(c\mathcal{M}(T))\\ &= c^n \det \mathcal{M}(T)\\ &= c^n \det T,\end{aligned}$$

as desired.

21. Prove or give a counterexample: if $S, T \in \mathcal{L}(V)$, then

$$\det(S + T) = \det S + \det T.$$

SOLUTION: Define $S, T \in \mathcal{L}(\mathbf{F}^2)$ by

$$S(x, y) = (x, 0) \quad \text{and} \quad T(x, y) = (0, y).$$

Then S and T are both not invertible. Thus $\det S = \det T = 0$ (by 10.14). However, $S + T = I$ and $\det I = 1$, so $\det(S + T) \neq \det S + \det T$.

Of course there are also many other examples.

22. Suppose A is a block upper-triangular matrix

$$A = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j along the diagonal is a square matrix. Prove that

$$\det A = (\det A_1) \dots (\det A_m).$$

SOLUTION: First consider the case where $m = 2$, so we can write A in the form

$$A = \begin{bmatrix} B & * \\ 0 & C \end{bmatrix},$$

where B is an n-by-n matrix

$$B = \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,n} \end{bmatrix}$$

and C is a p-by-p matrix

$$C = \begin{bmatrix} c_{1,1} & \dots & c_{1,p} \\ \vdots & & \vdots \\ c_{p,1} & \dots & c_{p,p} \end{bmatrix}.$$

Let $a_{j,k}$ denote the entry in row j, column k of A. Note that $a_{j,k} = 0$ if $j > n$ and $k \leq n$ (this follows from the block upper-triangular form of A). Because A is an $(n+p)$-by-$(n+p)$ matrix, to compute $\det A$ we need to consider a typical permutation $(m_1, \dots, m_{n+p}) \in \operatorname{perm}(n+p)$. If any of $m_1, \dots, m_n$ is greater than n, then

$$a_{m_1,1} \dots a_{m_{n+p},n+p} = 0.$$

Thus in computing $\det A$ we need only consider permutations in which the first n coordinates all come from $\{1, \dots, n\}$, which means that the last p coordinates all come from $\{n+1, \dots, n+p\}$. We can break any such permutation $(m_1, \dots, m_{n+p})$ into two pieces: a permutation of $\{1, \dots, n\}$ and (relabeling $n+1, \dots, n+p$ as $1, \dots, p$ to correspond to the labeling of the entries of C) a permutation of $\{1,\ldots,p\}$. Clearly the sign of $(m_1,\ldots,m_{n+p})$ will equal the product of the signs of these two permutations. Putting all this together, we have

$$\begin{aligned}
\det A &= \sum_{(m_1,\ldots,m_{n+p})\in \operatorname{perm}(n+p)} \big(\operatorname{sign}(m_1,\ldots,m_{n+p})\big) a_{m_1,1}\ldots a_{m_{n+p},n+p} \\
&= \sum_{\substack{(j_1,\ldots,j_n)\in \operatorname{perm} n\\ (k_1,\ldots,k_p)\in \operatorname{perm} p}} \Big[\big(\operatorname{sign}(j_1,\ldots,j_n)\big)\big(\operatorname{sign}(k_1,\ldots,k_p)\big)\cdot \\
&\qquad\qquad b_{j_1,1}\ldots b_{j_n,n} c_{k_1,1}\ldots c_{k_p,p}\Big] \\
&= \sum_{(j_1,\ldots,j_n)\in \operatorname{perm} n} \big(\operatorname{sign}(j_1,\ldots,j_n)\big) b_{j_1,1}\ldots b_{j_n,n}\cdot \\
&\qquad\qquad \sum_{(k_1,\ldots,k_p)\in \operatorname{perm} p} \big(\operatorname{sign}(k_1,\ldots,k_p)\big) c_{k_1,1}\ldots c_{k_p,p} \\
&= (\det B)(\det C),
\end{aligned}$$

completing the proof when $m = 2$.

Suppose now that $m > 2$ and

$$A = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix}.$$

Writing

$$B = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_{m-1} \end{bmatrix},$$

we have

$$A = \begin{bmatrix} B & * \\ 0 & A_m \end{bmatrix}.$$

Thus

$$\begin{aligned}
\det A &= (\det B)(\det A_m) \\
&= (\det A_1)\ldots(\det A_{m-1})(\det A_m),
\end{aligned}$$

where the first equality holds by the $m = 2$ case proved above and the second equality comes from induction on m (meaning that we can assume the desired result holds when m is replaced with $m-1$).

23. Suppose A is an n-by-n matrix with real entries. Let $S \in \mathcal{L}(\mathbf{C}^n)$ denote the operator on $\mathbf{C}^n$ whose matrix equals A, and let $T \in \mathcal{L}(\mathbf{R}^n)$ denote the operator on $\mathbf{R}^n$ whose matrix equals A. Prove that $\operatorname{trace} S = \operatorname{trace} T$ and $\det S = \det T$.

SOLUTION: The formulas defining the trace and determinant of a matrix do not depend upon whether we think of the matrix entries as real or complex numbers. We have $\operatorname{trace} A = \operatorname{trace} S$ and $\operatorname{trace} A = \operatorname{trace} T$ (by 10.11). Thus $\operatorname{trace} S = \operatorname{trace} T$. Similarly, we have $\det A = \det S$ and $\det A = \det T$ (by 10.33). Thus $\det S = \det T$.

24. Suppose V is an inner-product space and $T \in \mathcal{L}(V)$. Prove that

$$\det T^* = \overline{\det T}.$$

Use this to prove that $|\det T| = \det \sqrt{T^*T}$, giving a different proof than was given in 10.37.

SOLUTION: Let $n = \dim V$. If $\lambda \in \mathbf{F}$, then $\big((T - \lambda I)^n\big)^* = (T^* - \bar{\lambda} I)^n$, which implies that

$$\dim \operatorname{null}(T - \lambda I)^n = \dim \operatorname{null}(T^* - \bar{\lambda} I)^n,$$

where we have used Exercise 31 in Chapter 7. The equation above shows that the eigenvalues of T^* are precisely the complex conjugates of the eigenvalues of T, with the same multiplicities. If $\mathbf{F} = \mathbf{C}$, the determinant equals the product of the eigenvalues, counting multiplicity, so we have $\det T^* = \overline{\det T}$ (so far just on complex vector spaces).

Now suppose $\mathbf{F} = \mathbf{R}$, so we must consider eigenpairs. If $\alpha, \beta \in \mathbf{R}$, then

$$\dim \operatorname{null}(T^2 + \alpha T + \beta I)^n = \dim \operatorname{null}((T^*)^2 + \alpha T^* + \beta I)^n,$$

again by Exercise 31 in Chapter 7. Thus T and T^* have the same eigenpairs with the same multiplicities. From the paragraph above, T and T^* have the same eigenvalues with the same multiplicities. Because the determinant equals the product of the eigenvalues (counting multiplicity) times the product of the second coordinates of the eigenpairs (counting multiplicity), this implies that $\det T^* = \det T = \overline{\det T}$.

At this point, we know that $\det T^* = \overline{\det T}$ regardless of whether $\mathbf{F} = \mathbf{C}$ or $\mathbf{F} = \mathbf{R}$. Thus

$$\begin{aligned}(\det \sqrt{T^*T})^2 &= (\det \sqrt{T^*T})(\det \sqrt{T^*T}) \\ &= \det(T^*T) \\ &= (\det T^*)(\det T) \\ &= (\overline{\det T})(\det T) \\ &= |\det T|^2.\end{aligned}$$

Taking square roots (and recalling that the positive operator $\sqrt{T^*T}$ has a nonnegative determinant), we have $\det \sqrt{T^*T} = |\det T|$, as desired.

25. Let a, b, c be positive numbers. Find the volume of the ellipsoid

$$\left\{(x, y, z) \in \mathbf{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\right\}$$

by finding a set $\Omega \subset \mathbf{R}^3$ whose volume you know and an operator $T \in \mathcal{L}(\mathbf{R}^3)$ such that $T(\Omega)$ equals the ellipsoid above.

SOLUTION: Let E denote the ellipsoid defined above and let Ω be the ball with radius 1 defined by

$$\Omega = \{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 < 1\}.$$

Of course volume $\Omega = \frac{4}{3}\pi$. Define $T \in \mathcal{L}(\mathbf{R}^3)$ by

$$T(x, y, z) = (ax, by, cz).$$

If $(x, y, z) \in \Omega$, then

$$\begin{aligned}\frac{(ax)^2}{a^2} + \frac{(by)^2}{b^2} + \frac{(cz)^2}{c^2} &= x^2 + y^2 + z^2 \\ &< 1,\end{aligned}$$

which shows that $T(x, y, z) \in E$. Thus $T(\Omega) \subset E$.

Conversely, if $(x, y, z) \in E$, then

$$\begin{aligned}\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 &= \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \\ &< 1,\end{aligned}$$

which shows that $(\frac{x}{a}, \frac{y}{b}, \frac{z}{c}) \in \Omega$. Because $T(\frac{x}{a}, \frac{y}{b}, \frac{z}{c}) = (x, y, z)$, this implies that $E \subset T(\Omega)$.

The last two paragraphs show that $E = T(\Omega)$. Thus

$$\begin{aligned} \operatorname{volume} E &= \operatorname{volume} T(\Omega) \\ &= |\det T|(\operatorname{volume} \Omega) \\ &= \frac{4\pi abc}{3}, \end{aligned}$$

where the second equality comes from 10.38.