Chapter 1: Vector Spaces

Linear Algebra Done Right, by Sheldon Axler

A: $\mathbb{R}$ and $\mathbb{C}$

Problem 1

Suppose a and b are real numbers, not both 0. Find real numbers c and d such that
$$\frac{1}{a+bi} = c + di$$

Proof. We have
$$\frac{1}{a+bi} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i,$$
and hence let
$$c = \frac{a}{a^2+b^2}, \quad d = -\frac{b}{a^2+b^2},$$
and we're done. □

Problem 3

Find two distinct square roots of i.

Proof. Suppose $a, b \in \mathbb{R}$ are such that $(a+bi)^2 = i$. Then
$$(a^2 - b^2) + (2ab)i = i.$$
Since the real and imaginary part of both sides must be equal, respectively, we have a system of two equations in two variables
$$\begin{aligned} a^2 - b^2 &= 0 \\ ab &= \frac{1}{2}. \end{aligned}$$
The first equation implies $b = \pm a$. Plugging $b = -a$ into the second equation would imply $-a^2 = 1/2$, which is impossible, and hence we must have $a = b$. But this in turn tells us $a = \pm 1/\sqrt{2}$, and hence our two roots are
$$\pm\left(\frac{1}{\sqrt{2}}\right)(1+i),$$
as desired. □

Problem 5

Show that $(\alpha+\beta)+\lambda = \alpha+(\beta+\lambda)$ for all $\alpha, \beta, \lambda \in \mathbb{C}$.

Proof. Suppose $\alpha = a_1 + a_2 i$, $\beta = b_1 + b_2 i$, and $\lambda = c_1 + c_2 i$ for $a_k, b_k, c_k \in \mathbb{R}$, where $k = 1, 2$. Then

$$\begin{aligned}
(\alpha+\beta)+\lambda &= \big[(a_1+a_2 i)+(b_1+b_2 i)\big]+(c_1+c_2 i) \\
&= \big[(a_1+b_1)+(a_2+b_2)i\big]+(c_1+c_2 i) \\
&= \big[(a_1+b_1)+c_1\big]+\big[(a_2+b_2)+c_2\big]\,i \\
&= \big[a_1+(b_1+c_1)\big]+\big[a_2+(b_2+c_2)\big]\,i \\
&= (a_1+a_2 i)+\big[(b_1+c_1)+(b_2+c_2)i\big] \\
&= (a_1+a_2 i)+\big[(b_1+b_2 i)+(c_1+c_2)i\big] \\
&= \alpha+(\beta+\lambda),
\end{aligned}$$

as desired. □

Problem 7

Show that for every $\alpha \in \mathbb{C}$, there exists a unique $\beta \in \mathbb{C}$ such that $\alpha+\beta = 0$.

Proof. Suppose $\alpha = a_1 + a_2 i$ for some $a_1, a_2 \in \mathbb{R}$, and define $\beta = -a_1 - a_2 i$. Then

$$\begin{aligned}
\alpha+\beta &= (a_1+a_2 i)+(-a_1-a_2 i) \\
&= (a_1-a_1)+(a_2-a_2)i \\
&= 0+0i \\
&= 0,
\end{aligned}$$

proving existence. To see that β is unique, suppose $\lambda \in \mathbb{C}$ such that $\alpha+\lambda = 0$. Then

$$\lambda = \lambda+(\alpha+\beta) = (\lambda+\alpha)+\beta = 0+\beta = \beta,$$

and thus β is unique. □

Problem 9

Show that $\lambda(\alpha+\beta) = \lambda\alpha+\lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbb{C}$.

Proof. Suppose $\alpha = a_1 + a_2 i$, $\beta = b_1 + b_2 i$, and $\lambda = c_1 + c_2 i$ for $a_k, b_k, c_k \in \mathbb{R}$, where $k = 1, 2$. Then

$$\begin{aligned}
\lambda(\alpha+\beta) &= (c_1 + c_2 i)\big[(a_1 + a_2 i) + (b_1 + b_2 i)\big] \\
&= (c_1 + c_2 i)\big[(a_1 + b_1) + (a_2 + b_2) i\big] \\
&= \big[c_1(a_1 + b_1) - c_2(a_2 + b_2)\big] + \big[c_1(a_2 + b_2) + c_2(a_1 + b_1)\big]\, i \\
&= \big[(c_1 a_1 + c_1 b_1) - (c_2 a_2 + c_2 b_2)\big] + \big[(c_1 a_2 + c_1 b_2) + (c_2 a_1 + c_2 b_1)\big]\, i \\
&= \big[(c_1 a_1 - c_2 a_2) + (c_1 b_1 - c_2 b_2)\big] + \big[(c_1 a_2 + c_2 a_1) + (c_1 b_2 + c_2 b_1)\big]\, i \\
&= \big[(c_1 a_1 - c_2 a_2) + (c_1 a_2 + c_2 a_1) i\big] + \big[(c_1 b_1 - c_2 b_2) + (c_1 b_2 + c_2 b_1) i\big] \\
&= (c_1 + c_2 i)(a_1 + a_2 i) + (c_1 + c_2 i)(b_1 + b_2 i) \\
&= \lambda\alpha + \lambda\beta,
\end{aligned}$$

as desired. ☐

Problem 11

Explain why there does not exist $\lambda \in \mathbb{C}$ such that

$$\lambda(2 - 3i, 5 + 4i, -6 + 7i) = (12 - 5i, 7 + 22i, -32 - 9i).$$

Proof. Suppose such a $\lambda \in \mathbb{C}$ exists, say $\lambda = a + bi$ for some $a, b \in \mathbb{R}$. Then

$$\begin{aligned}
(a + bi)(2 - 3i) &= 12 - 5i \\
(a + bi)(5 + 4i) &= 7 + 22i \\
(a + bi)(-6 + 7i) &= -32 - 9i,
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
(2a + 3b) + (-3a + 2b)i &= 12 - 5i && (1) \\
(5a - 4b) + (4a + 5b)i &= 7 + 22i && (2) \\
(-6a - 7b) + (7a - 6b)i &= -32 - 9i. && (3)
\end{aligned}$$

For each equation above, the real part of the LHS must equal the real part of the RHS, and similarly for their imaginary parts. In particular, the following two equations hold by comparing the real parts of Equations (1) and (3)

$$\begin{aligned}
2a + 3b &= 12 \\
-6a - 7b &= -32.
\end{aligned}$$

Multiplying the first equation by 3 and adding it to the second, we find $b = 2$. Substituting this value back into the first equation yields $a = 2$. However, comparing the imaginary parts of Equation (3) tells us we must have

$$7a - 6b = -9,$$

a contradiction, since $a = 3$ and $b = 2$ yields $7a - 6b = 9$. Thus no such $\lambda \in \mathbb{C}$ exists, as was to be shown. ☐

Problem 13

Show that $(ab)x = a(bx)$ for all $x \in \mathbb{F}^n$ and all $a, b \in \mathbb{F}$.

Proof. We may write $x = (x_1, \ldots, x_n)$ for $x_1, \ldots, x_n \in \mathbb{F}$. It follows

$$\begin{aligned}(ab)x &= \big((ab)x_1, \ldots, (ab)x_n\big) \\ &= \big(a(bx_1), \ldots, a(bx_n)\big) \\ &= a\,(bx_1, \ldots, bx_n) \\ &= a(bx),\end{aligned}$$

as desired. □

Problem 15

Show that $\lambda(x + y) = \lambda x + \lambda y$ for all $\lambda \in \mathbb{F}$ and all $x, y \in \mathbb{F}^n$.

Proof. We may write $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ for $x_k, y_k \in \mathbb{F}^n$, where $k = 1, \ldots, n$. It follows

$$\begin{aligned}\lambda(x + y) &= \lambda((x_1, \ldots, x_n) + (y_1, \ldots, y_n)) \\ &= \lambda((x_1 + y_1) + \cdots + (x_n + y_n)) \\ &= (\lambda(x_1 + y_1) + \cdots + \lambda(x_n + y_n)) \\ &= ((\lambda x_1 + \lambda y_1) + \cdots + (\lambda x_n + \lambda y_n)) \\ &= (\lambda x_1 + \cdots + \lambda x_n) + (\lambda y_1 + \cdots + \lambda y_n) \\ &= \lambda x + \lambda y,\end{aligned}$$

as desired. □

B: Definition of a Vector Space

Problem 1

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

Proof. We wish to show that v is the additive inverse of $(-v)$. We have

$$(-v) + v = (-1)v + 1v = (-1 + 1)v = 0v = 0,$$

as desired. □

Problem 3

Suppose $v, w \in V$. Explain why there exists a unique $x \in V$ such that $v + 3x = w$.

Proof. First we prove existence. Define $x \in V$ by

$$x = \frac{1}{3}(w - v).$$

Then

$$\begin{aligned} v + 3x &= v + 3\left(\frac{1}{3}(w - v)\right) \\ &= v + \left(3 \cdot \frac{1}{3}\right)(w - v) \\ &= v + (w - v) \\ &= w, \end{aligned}$$

and so such an x exists. To see that it's unique, suppose $y \in V$ such that $v + 3y = w$. Then

$$v + 3y = v + 3x \iff 3y = 3x \iff y = x,$$

proving uniqueness. $\square$

Problem 5

Show that in the definition of a vector space (1.19), the additive inverse condition can be replaced with the condition that

$$0v = 0 \text{ for all } v \in V.$$

Here the 0 on the left side is the number 0, and the 0 on the right side is the additive identity of V. (The phrase "a condition can be replaced" in a definition means that the collection of objection satisfying the definition is unchanged if the original condition is replaced with the new definition.)

Proof. We show that the two statements are equivalent.

First suppose that every $v \in V$ has an additive inverse. Since we have

$$0v + 0v = (0 + 0)v = 0v,$$

adding the additive inverse to both sides yields $0v = 0$.

Conversely, suppose that $0v = 0$ for all $v \in V$. Then

$$v + (-1)v = (1 + (-1))v = 0v = 0,$$

and hence every element has an additive inverse, as desired. $\square$

C: Subspaces

Problem 1

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

1. $\{(x_1, x_2, x_3) \in \mathbb{F}^3 \mid x_1 + 2x_2 + 3x_3 = 0\}$
2. $\{(x_1, x_2, x_3) \in \mathbb{F}^3 \mid x_1 + 2x_2 + 3x_3 = 4\}$
3. $\{(x_1, x_2, x_3) \in \mathbb{F}^3 \mid x_1x_2x_3 = 0\}$
4. $\{(x_1, x_2, x_3) \in \mathbb{F}^3 \mid x_1 = 5x_3\}$

Proof. (a) Let S denote the specified subset. We claim S is a subspace. To see this, note that $0 + 2 \cdot 0 + 3 \cdot 0 = 0$, and hence $0 \in S$. Now suppose $x = (x_1, x_2, x_3) \in S$ and $y = (y_1, y_2, y_3) \in S$. Then

$$x_1 + 2x_2 + 3x_3 = 0 \quad \text{and} \quad y_1 + 2y_2 + 3y_3 = 0,$$

and hence

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

and so $x + y \in S$ and S is closed under addition. Now letting $a \in \mathbb{F}$, we have

$$a(x_1 + 2x_2 + 3x_3) = ax_1 + 2(ax_2) + 3(ax_3) = 0,$$

and hence $ax \in S$ as well, and so S is closed under scalar multiplication, thus proving S is a subspace, as claimed.

(b) Let S denote the specified subset. Then S is not a subspace, for $0 + 2 \cdot 0 + 3 \cdot 0 = 0$, and hence S does not contain the additive identity.

(c) Let S denote the specified subset. We claim S is not a subspace since it is not closed under addition. To see this, let $x = (1, 0, 0)$ and $y = (0, 1, 1)$. Then $x, y \in S$, but $x + y = (1, 1, 1) \notin S$ since $(x_1 + y_1)(x_2 + y_2)(x_3 + y_3) = 1 \cdot 1 \cdot 1 \neq 0$.

(d) Let S denote the specified subset. We claim S is a subspace. To see this, note that $0 = 5 \cdot 0$, and hence $0 \in S$. Now suppose $x = (x_1, x_2, x_3) \in S$ and $y = (y_1, y_2, y_3) \in S$. Then

$$x_1 = 5x_3 \quad \text{and} \quad y_1 = 5y_3,$$

and hence

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

and so $x + y \in S$ and S is closed under addition. Now letting $a \in \mathbb{F}$, we have

$$a(x_1) = a(5x_3)$$

and thus

$$(ax_1) = 5(ax_3),$$

showing $ax \in S$ as well. Therefore S is closed under scalar multiplication as well, proving S is a subspace, as claimed. $\square$

Problem 3

Show that the set of differentiable real-valued functions f on the interval $(-4, 4)$ such that $f'(-1) = 3f(2)$ is a subspace of $\mathbb{R}^{(-4,4)}$.

Proof. Let S denote the set of differentiable real-valued functions f on the interval $(-4, 4)$ such that $f'(-1) = 3f(2)$. Denote the zero-function (the additive identity of $\mathbb{R}^{(-4,4)}$) by f_0. Then $f_0' = f_0$ and $f_0'(-1) = 0$, and hence $f_0'(-1) = 3f_0(2)$ — both sides are 0 — showing $f_0 \in S$. Now suppose $f, g \in S$. Then

$$(f + g)' = f' + g',$$

and hence

$$\begin{aligned}(f + g)'(-1) &= f'(-1) + g'(-1)\\ &= 3f(2) + 3g(2)\\ &= 3(f(2) + g(2))\\ &= 3(f + g)(2),\end{aligned}$$

showing $(f + g) \in S$ and S is closed under addition. Now letting $a \in \mathbb{R}$, we have

$$a(f'(-1)) = a(3f(2)) \implies (af')(-1) = 3(af)(2),$$

and hence $(af) \in S$ as well, and S is closed under scalar multiplication. Therefore, S is a subspace. $\square$

Problem 5

Is $\mathbb{R}^2$ a subspace of the complex vector space $\mathbb{C}^2$?

Proof. The set $\mathbb{R}^2$ is not a subspace of $\mathbb{C}^2$ over the field $\mathbb{C}$ since $\mathbb{R}^2$ is not closed under scalar multiplication. In particular, we have $ix \notin \mathbb{R}^2$ for all $x \in \mathbb{R}^2 - \{0\}$. $\square$

Problem 7

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

Proof. Consider the set $\mathbb{Z} \times \mathbb{Z}$. Then for all $(a, b) \in \mathbb{Z} \times \mathbb{Z}$, we have $(-a, -b) \in \mathbb{Z} \times \mathbb{Z}$, and so it's closed under additive inverses. Similarly, for any $(c, d) \in \mathbb{Z} \times \mathbb{Z}$, we have $(a, b) + (c, d) = (a + c, b + d) \in \mathbb{Z} \times \mathbb{Z}$, and so it's closed under addition. But $\mathbb{Z} \times \mathbb{Z}$ is not a subspace of $\mathbb{R}^2$, since it is not closed under scalar multiplication. In particular, $\frac{1}{2}(1, 1) = \left(\frac{1}{2}, \frac{1}{2}\right) \notin \mathbb{Z} \times \mathbb{Z}$. $\square$

Problem 9

A function $f : \mathbb{R} \to \mathbb{R}$ is called ***periodic*** if there exists a positive number p such that $f(x) = f(x + p)$ for all $x \in \mathbb{R}$. Is the set of periodic functions from $\mathbb{R}$ to $\mathbb{R}$ a subspace of $\mathbb{R}^{\mathbb{R}}$? Explain.

Proof. Let P denote the set of periodic functions from $\mathbb{R}$ to $\mathbb{R}$. We claim P is not a subspace of $\mathbb{R}^{\mathbb{R}}$, since it is not closed under addition. To see this, define

$$f(x) = \cos\left(\frac{2\pi}{\sqrt{2}}x\right) \quad \text{and} \quad g(x) = \cos(2\pi x).$$

Then

$$\begin{aligned} f(x + \sqrt{2}) &= \cos\left(\frac{2\pi}{\sqrt{2}}(x + \sqrt{2})\right) \\ &= \cos\left(\frac{2\pi}{\sqrt{2}}x + 2\pi\right) \\ &= \cos\left(\frac{2\pi}{\sqrt{2}}x\right) \\ &= f(x), \end{aligned}$$

and so f has period $\sqrt{2}$, and

$$g(x + 1) = \cos(2\pi(x + 1)) = \cos(2\pi x + 2\pi) = \cos(2\pi x) = g(x),$$

so that g has period 1.

Suppose by way of contradiction that $f + g$ were periodic with respect to some $p \in \mathbb{R}^+$. Then, since

$$(f + g)(0) = \cos\left(\frac{2\pi}{\sqrt{2}} \cdot 0\right) + \cos(2\pi \cdot 0) = \cos(0) + \cos(0) = 2,$$

by periodicity of $f+g$ we must also have

$$(f+g)(p) = \cos\left(\frac{2\pi}{\sqrt{2}}p\right) + \cos(2\pi p) = 2.$$

The maximum of cosine is 1, and hence both f and g must have maxima at p. But the maxima of cosine occur at the integer multiples of 2π, and hence we must have

$$\frac{2\pi}{\sqrt{2}}p = 2\pi n \quad \text{and} \quad 2\pi p = 2\pi m$$

for some $n, m \in \mathbb{Z}^+$. But this implies

$$p = \sqrt{2}n \quad \text{and} \quad p = m.$$

In other words,

$$\frac{m}{n} = \sqrt{2},$$

a contradiction since $\sqrt{2}$ is irrational. Thus $f+g$ cannot be periodic, and indeed P is not closed under addition, as claimed. $\square$

Problem 11

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

Proof. Let $\mathfrak{C}$ denote a collection of subspaces of V, and let

$$U = \bigcap_{W \in \mathfrak{C}} W.$$

Then, since $0 \in W$ for all $W \in \mathfrak{C}$, we have $0 \in U$ and so U contains the additive identity. Now suppose $u, v \in U$. Then $u, v \in W$ for all $W \in \mathfrak{C}$, and hence $u+v \in W$ for all $W \in \mathfrak{C}$. Therefore, $u+v \in U$ and U is closed under addition. Next let $a \in \mathbb{F}$. Then $au \in W$ for all $W \in \mathfrak{C}$, and hence $au \in U$, showing U is closed under scalar multiplication. Therefore, U is indeed a subspace of V. $\square$

Problem 12

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.

Proof. Let U_1, U_2 be subspaces of V.

First suppose that one of the subspaces is contained in the other. Then either $U_1 \cup U_2 = U_1$ or $U_1 \cup U_2 = U_2$, and in both cases $U_1 \cup U_2$ is indeed a subspace of V.

Conversely, suppose by way of contradiction that $U_1 \cup U_2$ is a subspace of V, but neither subspace is contained in the other. That is, the sets $U_1 \setminus U_2$ and $U_2 \setminus U_1$ are both nonempty. Let $x \in U_1 \setminus U_2$ and $y \in U_2 \setminus U_1$. We claim $x+y \notin U_1$ and $x+y \notin U_2$, so that $U_1 \cup U_2$ is not closed under addition, a contradiction. To see this, suppose $x+y \in U_1$. Then $(x+y)-x \in U_1$ by closure of addition in U_1, but this is absurd since this implies $y \in U_1$, contrary to assumption. Similarly, suppose $x+y \in U_2$. Then $(x+y)-y \in U_2$, which is also absurd since this implies $x \in U_2$. Therefore $U_1 \cup U_2$ is not closed under addition, producing a contradiction as claimed. Thus we must have one of the subspaces contained in the other, as desired. □

Problem 13

Prove that the union of three subspaces of V is a subspace of V if and only if one of the subspaces contains the other two.

Proof. Let U_1, U_2, U_3 be subspaces of V.

($\Leftarrow$) Suppose that one of the subspaces contains the other two. Without loss of generality, assume $U_1 \subseteq U_3$ and $U_2 \subseteq U_3$. Then $U_1 \cup U_2 \cup U_3 = U_3$, and so $U_1 \cup U_2 \cup U_3$ is indeed a subspace of V.

($\Rightarrow$) Now suppose $U_1 \cup U_2 \cup U_3$ is a subspace. If U_2 contains U_3 (or conversely), let $W = U_2 \cup U_3$. Then applying Problem 12 to the union $U_1 \cup W$, we have that either U_1 contains W or W contains U_1, showing that one of the three subspaces contains the other two, as desired. So assume U_2 and U_3 are such that neither contains the other. Let

$$x \in U_2 \setminus U_3 \quad \text{and} \quad y \in U_3 \setminus U_2,$$

and choose $a, b \in \mathbb{F} \setminus \{0\}$ such that $a - b = 1$ (such a, b exist since we assume $\mathbb{F}$ is not finite).

We claim that $ax + y$ and $bx + y$ are both in U_1. To see that $ax + y \in U_1$, suppose not. Then either $ax + y \in U_2$ or $ax + y \in U_3$. If $ax + y \in U_2$, then we have $(ax + y) - ax = y \in U_2$, a contradiction. And if $ax + y \in U_3$, we have $(ax + y) - y = ax \in U_3$, another contradiction, and so $ax + y \in U_1$. Similarly for $bx + y$, suppose $bx + y \in U_2$. Then $(bx + y) - bx = y \in U_2$, a contradiction. And if $bx + y \in U_3$, then $(bx + y) - y = bx \in U_3$, also a contradiction. Thus $bx + y \in U_1$ as well. Therefore

$$(ax + y) - (bx + y) = (a - b)x = x \in U_1.$$

Now, since $x \in U_2 \setminus U_3$ implies $x \in U_1$, we have $U_2 \setminus U_3 \subseteq U_1$. A similar argument shows that $x + ay$ and $x + by$ must be in U_1 as well, and hence

$$(x + ay) - (x + by) = (a - b)y = y \in U_1,$$

and therefore $U_3 \setminus U_2 \subseteq U_1$. If $U_2 \cap U_3 = \emptyset$, we're done, so assume otherwise.

Now for any $u \in U_2 \cap U_3$, choose $v \in U_3 \setminus U_2 \subseteq U_1$. Then $u + v \notin U_2 \cap U_3$, for otherwise $(u+v) - u = v \in U_2$, a contradiction. But this implies $u + v$ must be in U_1, and hence so is $(u+v) - v = u$. In other words, if $u \in U_2 \cap U_3$, then $u \in U_1$, and hence $U_2 \cap U_3 \subseteq U_1$, as was to be shown. $\square$

Problem 15

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

Proof. We claim $U + V = V$. First suppose $x \in V$. Then $x = 0 + x \in U + V$, and hence $V \subseteq U + V$. Now suppose $y \in U + V$. Then there exist $u \in U$ and $v \in V$ such that $y = u + v$. But since U is a subspace of V, we have $u \in V$, and hence $u + v \in V$. Therefore $U + V \subseteq V$, proving the claim. $\square$

Problem 17

Is the operation of addition on the subspaces of V associative? In other words, if U_1, U_2, U_3 are subspaces of V, is

$$(U_1 + U_2) + U_3 = U_1 + (U_2 + U_3)?$$

Proof. Let U_1, U_2, U_3 be subspaces of V, and let $V_1 = U_1 + U_2$ and $V_2 = U_2 + U_3$. We claim

$$V_1 + U_3 = U_1 + V_2.$$

To see this, suppose $x \in V_1 + U_3$. Then there exist $v_1 \in V_1$ and $u_3 \in U_3$ such that $x = v_1 + u_3$. But since $v_1 \in V_1 = U_1 + U_2$, there exist $u_1 \in U_1$ and $u_2 \in U_2$ such that $v_1 = u_1 + u_2$. Then $x = u_1 + u_2 + u_3$, and since $u_2 + u_3 \in U_2 + U_3 = V_2$, we have $x \in U_1 + V_2$ and hence $V_1 + U_3 \subseteq U_1 + V_2$. Now suppose $y \in U_1 + V_2$. Then there exist $u_1' \in U_1$ and $v_2 \in V_2$ such that $y = u_1' + v_2$. But since $v_2 \in V_2 = U_2 + U_3$, there exist $u_2' \in U_2$ and $u_3' \in U_3$ such that $v_2 = u_2' + u_3'$. Then $y = u_1' + u_2' + u_3'$, and since $u_1' + u_2' \in U_1 + U_2 = V_1$, have $y \in V_1 + U_3$ and hence $U_1 + V_2 \subseteq V_1 + U_3$. Thus $V_1 + U_3 = U_1 + V_2$, as claimed. $\square$

Problem 19

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

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

then $U_1 = U_2$.

Proof. The statement is false. To see this, let $V = U_1 = W = \mathbb{R}^2$ and $U_2 = \mathbb{R} \times \{0\}$. Then $U_1 + W = \mathbb{R}^2$ and $U_2 + W = \mathbb{R}^2$, but clearly $U_1 \neq U_2$. $\square$

Problem 21

Suppose

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbb{F}^5 \mid x, y \in \mathbb{F}\}.$$

Find a subspace W of $\mathbb{F}^5$ such that $\mathbb{F}^5 = U \oplus W$.

Proof. Let $v_1 = (1, 0, 1, 1, 2,), v_2 = (0, 1, 1, -1, 0)$, so that we may instead write V as

$$V = \{\alpha_1 v_1 + \alpha_2 v_2 \in \mathbb{F}^5 \mid \alpha_1, \alpha_2 \in \mathbb{F}\}.$$

Now let $w_1 = (0, 0, 1, 0, 0), w_2 = (0, 0, 0, 1, 0), w_3 = (0, 0, 0, 0, 1)$ and define

$$W = \{\alpha_1 w_1 + \alpha_2 w_2 + \alpha_3 w_3 \in \mathbb{F}^5 \mid \alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}\}.$$

We claim $U \oplus W = \mathbb{F}^5$. There are three things to prove: (1) W is a subspace of $\mathbb{F}^5$, (2) $U + W = \mathbb{F}^5$, and (3) this sum is direct.

To see that W is a subspace of $\mathbb{F}^5$, note that $0 \cdot w_1 + 0 \cdot w_2 + 0 \cdot w_3 = 0$, and hence $0 \in W$. Next suppose $a, b \in W$. Then there exist some $\alpha_k, \beta_k \in \mathbb{F}$, where $k = 1, 2, 3$, such that $a = \alpha_1 w_1 + \alpha_2 w_2 + \alpha_3 w_3$ and $b = \beta_1 w_1 + \beta_2 w_2 + \beta_3 w_3$. But then $a + b = (\alpha_1 + \beta_1)w_1 + (\alpha_2 + \beta_2)w_2 + (\alpha_3 + \beta_3)w_3$, which is again in W, and hence W is closed under addition. Finally, let $\gamma \in \mathbb{F}$. Then $\gamma a = \gamma(\alpha_1 w_1 + \alpha_2 w_2 + \alpha_3 w_3) = (\gamma\alpha_1)w_1 + (\gamma\alpha_2)w_2 + (\gamma\alpha_3)w_3$ which is again in W, and hence W is closed under scalar multiplication. So W is indeed a subspace.

We next show that $U + W = \mathbb{F}^5$. First notice that $U + W \subseteq \mathbb{F}^5$ since U, W are both subspaces of $\mathbb{F}^5$. To see the that $\mathbb{F}^5 \subseteq U + W$, let $a = (a_1, a_2, a_3, a_4, a_5) \in \mathbb{F}^5$. Recalling our definition of the vectors v_1, v_2, w_1, w_2, w_3, consider the linear combination

$$(a_1 v_1 + a_2 v_2) + \left[(a_3 - a_1 - a_2)w_1 + (a_4 - a_1 + a_2)w_2 + (a_5 - 2a_1)w_3\right].$$

Note that the sum above is an element of $U + W$. And after reducing, we find that the sum above equals $(a_1, a_2, a_3, a_4, a_5)$, and hence $a \in U + W$ and so in fact $\mathbb{F}^5 = U + W$.

Lastly we show that the sum is direct. Every element of $U + W$ has the form $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 + \alpha_5 v_5$ for some $\alpha_k \in \mathbb{F}$ with $k = 1, \ldots, 5$, so suppose $0 = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 + \alpha_5 v_5$. Simplifying yields

$$(\alpha_1, \alpha_2, \alpha_1 + \alpha_2 + \alpha_3, \alpha_1 - \alpha_2 + \alpha_4, 2\alpha_1 + \alpha_5) = 0.$$

Clearly $\alpha_1 = \alpha_2 = 0$. But now this equation simplifies to

$$(0, 0, \alpha_3, \alpha_4, \alpha_5) = 0,$$

and so $\alpha_3 = \alpha_4 = \alpha_5 = 0$ as well, and hence the sum is indeed direct. $\square$

Problem 23

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 $U_1 = U_2$.

Proof. The statement is false. Let $V = \mathbb{R}^2$, $W = \mathbb{R} \times \{0\}$, $U_1 = \{0\} \times \mathbb{R}$, and $U_2 = \{(x, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$. Then clearly
$$V = U_1 + W = U_2 + W.$$
Moreover, $U_1 \cap W = \{0\}$ and $U_2 \cap W = \{0\}$, and hence the sums are direct. That is,
$$V = U_1 \oplus W = U_2 \oplus W,$$
but $U_1 \neq U_2$. $\square$

Chapter 2: Finite-Dimensional Vector Spaces

Linear Algebra Done Right, by Sheldon Axler

A: Span and Linear Independence

Problem 1

Suppose v_1, v_2, v_3, v_4 spans V. Prove that the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

also spans V.

Proof. Let $w \in V$. Then there exist $a_1, a_2, a_3, a_4 \in \mathbb{F}$ such that

$$w = a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4.$$

We wish to find $b_1, b_2, b_3, b_4 \in \mathbb{F}$ such that

$$b_1(v_1 - v_2) + b_2(v_2 - v_3) + b_3(v_3 - v_4) + b_4v_4 = a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4.$$

Simplifying the LHS, we have

$$b_1v_1 + (b_2 - b_1)v_2 + (b_3 - b_2)v_3 + (b_4 - b_3)v_4 = a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4.$$

Hence we may choose

$$\begin{aligned} b_1 &= a_1 \\ b_2 &= a_1 + a_2 \\ b_3 &= a_1 + a_2 + a_3 \\ b_4 &= a_1 + a_2 + a_3 + a_4, \end{aligned}$$

so that w is given as a linear combination of the list $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$, and thus the list spans V as well. $\square$

Problem 3

Find a number t such that

$$(3, 1, 4), (2, -3, 5), (5, 9, t)$$

is not linearly independent in $\mathbb{R}^3$.

Proof. Let $t = 2$. Then

$$3(3, 1, 4) - 2(2, -3, 5) = (5, 9, 2),$$

and hence the vectors are not linearly independent since one of the vectors can be written as a linear combination of the other two. □

Problem 5

(a) Show that if we think of $\mathbb{C}$ as a vector space over $\mathbb{R}$, then the list $(1 + i, 1 - i)$ is linearly independent.

(b) Show that if we think of $\mathbb{C}$ as a vector space over $\mathbb{C}$, then the list $(1 + i, 1 - i)$ is linearly dependent.

Proof. (a) Suppose

$$a(1 + i) + b(1 - i) = 0$$

for some $a, b \in \mathbb{R}$. Then

$$(a + b) + (a - b)i = 0.$$

Comparing imaginary parts, this implies $a - b = 0$ and hence $a = b$. But now substituting for b and comparing real parts, this implies $2a = 0$, and hence $a = b = 0$. Thus the vectors are linearly independent over $\mathbb{R}$.

(b) Note that

$$-i(1 + i) = 1 - i,$$

so that $1 - i$ is a scalar multiple of $1 + i$ and hence the vectors are linearly dependent over $\mathbb{C}$. □

Problem 7

Prove or give a counterexample: If $v_1, v_2, \ldots, v_m$ is a linearly independent list of vectors in V, then

$$5v_1 - 4v_2, v_2, v_3, \ldots, v_m$$

is linearly independent.

Proof. Let $u = 5v_1 - 4v_2$. We claim the list $u, v_2, \ldots, v_m$ is linearly independent. To see this, suppose not. Then there exists some $j \in \{2, \ldots, m\}$ such that $v_j \in \operatorname{span}(u, v_2, \ldots, v_{j-1})$. But then v_j is also in $\operatorname{span}(v_1, v_2, \ldots, v_{j-1})$, since $u = 5v_1 - 4v_2$ is a linear combination of v_1 and v_2, a contradiction. □

Problem 9

Prove or give a counterexample: If $v_1, \ldots, v_m$ and $w_1, \ldots, w_m$ are linearly independent lists of vectors in V, then $v_1 + w_1, \ldots, v_m + w_m$ is linearly independent.

Proof. The statement is false. To see this, let $w_k = -v_k$ for $k = 1, \ldots, m$. Then $w_1, \ldots, w_m$ are also linearly independent, but $v_1 + w_1 = \cdots = v_m + w_m = 0$. $\square$

Problem 11

Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Show that $v_1, \ldots, v_m, w$ is linearly independent if and only if

$$w \notin \operatorname{span}(v_1, \ldots, v_m).$$

Proof. ($\Rightarrow$) First suppose $v_1, \ldots, v_m, w$ is linearly independent. If $w \in \operatorname{span}(v_1, \ldots, v_m)$, then there exist $a_1, \ldots, a_m \in \mathbb{F}$ such that

$$w = a_1 v_1 + \cdots + a_m v_m.$$

But then

$$-w + a_1 v_1 + \cdots + a_m v_m = 0,$$

a contradiction. Therefore we must have $w \notin \operatorname{span}(v_1, \ldots, v_m)$.

($\Leftarrow$) Now suppose $w \notin \operatorname{span}(v_1, \ldots, v_m)$ and consider the list $v_1, \ldots, v_m, w$. Suppose the list were linearly dependent. Then there exists a vector in the list which is in the span of its predecessors. Since this vector cannot be w by assumption, there exists some $j \in \{1, \ldots, m\}$ such that $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$, contradicting the hypothesis that $v_1, \ldots, v_m$ is linearly independent (and hence all sublists are). Thus $v_1, \ldots, v_m, w$ must be linearly independent. $\square$

Problem 13

Explain why no list of four polynomials spans $\mathcal{P}_4(\mathbb{F})$.

Proof. Note that the list $1, z, \ldots, z^4$ spans $\mathcal{P}_4(\mathbb{F})$, is linearly independent, and has length 5. Since the length of every spanning list must be at least as long as every linearly independent list, there exist no spanning lists of vectors in $\mathcal{P}(\mathbb{F})$ of length less than 5. $\square$

Problem 14

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 $v_1, \ldots, v_m$ is linearly independent for every positive integer m.

Proof. ($\Rightarrow$) First suppose V is infinite-dimensional. We will prove by induction that there exists a sequence $v_1, v_2, \ldots$ of vectors in V such that for every $m \in \mathbb{Z}^+$, the first m vectors are linearly independent.

Base Case: Since V is infinite-dimensional, V contains some nonzero vector v_1. The list containing only this vector is clearly linearly independent.

Inductive Step: Suppose the list of vectors $v_1, \ldots, v_k$ is linearly independent for some $k \in \mathbb{Z}^+$. Since V is infinite-dimensional, these vectors cannot span V, and hence there exists some $v_{k+1} \in V \setminus \operatorname{span}(v_1, \ldots, v_k)$. In particular, note that $v_{k+1} \neq 0$. But then $v_1, \ldots, v_k, v_{k+1}$ is linearly independent by the Linear Dependence Lemma (for if it were linearly dependent, the Lemma guarantees there would exist a vector in the list which could be written as a linear combination of its predecessors, which is impossible by our construction).

By induction, we have shown there exists a list $v_1, v_2, \ldots$ such that $v_1, \ldots, v_m$ is linearly independent for every $m \in \mathbb{Z}^+$.

($\Leftarrow$) Now suppose there is a sequence $v_1, v_2, \ldots$ of vectors in V such that $v_1, \ldots, v_m$ is linearly independent for every $m \in \mathbb{Z}^+$. If V were finite-dimensional, there would exist a list $v_1, \ldots, v_n$ for some $n \in \mathbb{Z}^+$ such that $V = \operatorname{span}(v_1, \ldots, v_n)$. But then, by our assumption, the list $v_1, \ldots, v_{n+1}$ is linearly independent. Since every linearly independent list must have length no longer than every spanning list, this is a contradiction. Thus V is infinite-dimensional. □

Problem 15

Prove that $\mathbb{F}^\infty$ is infinite-dimensional.

Proof. For each $k \in \mathbb{Z}$, define the vector e_k such that it has a 1 in coordinate k and 0 everywhere else. Then for the sequence $e_1, e_2, \ldots$, the list $e_1, \ldots, e_m$ is linearly independent for any choice of $m \in \mathbb{Z}^+$. By Problem 14, $\mathbb{F}^\infty$ must be infinite-dimensional. □

Problem 17

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

Proof. Suppose it were. We will show that this implies $p_0, p_1, \ldots, p_m$ spans $\mathcal{P}_m(\mathbb{F})$ and that this in turn leads to a contradiction by explicitly constructing a polynomial that is not in this span.

Note that the list $1, z, \ldots, z^{m+1}$ spans $\mathcal{P}_m(\mathbb{F})$ and has length $m+1$, hence every linearly independent list must have length $m+1$ or less. If $\operatorname{span}(p_0, p_1, \ldots, p_m) \neq \mathcal{P}_m(\mathbb{F})$, there exists some $p \notin \operatorname{span}(p_0, p_1, \ldots, p_m)$, and thus the list $p_0, p_1, \ldots, p_m, p$ is linearly independent and of length $m+2$, a contradiction. And so we must have $\operatorname{span}(p_0, p_1, \ldots, p_m) = \mathcal{P}_m(\mathbb{F})$.

Now define the polynomial $q \equiv 1$. Then $q \in \operatorname{span}(p_0, p_1, \ldots, p_m)$, and hence there exist $a_0, \ldots, a_m \in \mathbb{F}$ such that

$$q = a_0 p_0 + a_1 p_1 + \cdots + a_m p_m,$$

which in turn implies

$$q(2) = a_0 p_0(2) + a_1 p_1(2) + \cdots + a_m p_m(2).$$

But this is absurd, since this implies $1 = 0$. Therefore $p_0, p_1, \ldots, p_m$ cannot be linearly independent, as desired. □

B: Bases

Problem 1

Find all vector spaces that have exactly one basis.

Proof. We claim that only the trivial vector space has exactly one basis. We first consider finite-dimensional vector spaces. Let V be a nontrivial vector space with basis $v_1, \ldots, v_n$. We claim that for any $c \in \mathbb{F}^\times$, the list $cv_1, \ldots, cv_n$ is a basis as well. Clearly the list is still linearly independent, and to see that it still spans V, let $u \in V$. Then, since $v_1, \ldots, v_n$ spans V, there exist $a_1, \ldots, a_n \in \mathbb{F}$ such that

$$u = a_1 v_1 + \cdots + a_n v_n.$$

But then we have

$$u = \frac{a_1}{c}(cv_1) + \cdots + \frac{a_n}{c}(cv_n)$$

and so $cv_1, \ldots, cv_n$ span V as well. Thus we have more than one basis for all finite-dimensional vector spaces.

Essentially the same proof shows the same thing for infinite-dimensional vector spaces. So let W be an infinite-dimensional vector space with basis $w_1, w_2, \ldots$. We claim that for any $c \in \mathbb{F}$, the list $cw_1, cw_2, \ldots$ is a basis as well. Clearly the list is again linearly independent, and to see that it still spans V, let $u \in V$. Then, since $w_1, w_2, \ldots$ spans W, there exist $a_1, a_2, \cdots \in \mathbb{F}$ such that

$$u = a_1 w_1 + a_2 w_2 + \ldots$$

But then we have

$$u = \frac{a_1}{c}(cw_1) + \frac{a_2}{c}(cw_2) + \ldots$$

and so $cw_1, cw_2, \ldots$ span W as well. Thus we have more than one basis for all infinite-dimensional vector spaces as well, proving our original claim. □

Problem 3

(a) Let U be the subspace of $\mathbb{R}^5$ defined by

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

Find a basis of U.

(b) Extend the basis in part (a) to a basis of $\mathbb{R}^5$.

(c) Find a subspace W of $\mathbb{R}^5$ such that $\mathbb{R}^5 = U \oplus W$.

Proof. (a) We claim the list of vectors

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

is a basis of U. We first show they span U. So let $u \in U$. Then there exist $x_1, \ldots, x_5 \in \mathbb{R}$ such that

$$u = (x_1, x_2, x_3, x_4, x_5)$$

and such that $x_1 = 3x_2$ and $x_3 = 7x_4$. Substitution yields

$$u = (3x_2, x_2, 7x_4, x_4, x_5),$$

and hence we have

$$u = x_2(3, 1, 0, 0, 0) + x_4(0, 0, 7, 1, 0) + x_5(0, 0, 0, 0, 1)$$

and indeed they span U. Now suppose $a_1, a_2, a_3 \in \mathbb{R}$ are such that

$$a_1(3, 1, 0, 0, 0) + a_2(0, 0, 7, 1, 0) + a_3(0, 0, 0, 0, 1) = 0.$$

Then we have

$$(3a_1, a_1, 0, 0, 0) + (0, 0, 7a_2, a_2, 0) + (0, 0, 0, 0, a_3) = 0$$

which clearly implies $a_1 = a_2 = a_3 = 0$. Thus they are also linearly independent, and hence a basis.

(b) We claim the list

$$v_1 = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ v_2 = \begin{pmatrix} 0 \\ 0 \\ 7 \\ 1 \\ 0 \end{pmatrix},\ v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ v_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

is a basis of $\mathbb{R}^5$ expanding the basis from (a). To see that it spans $\mathbb{R}^5$, let $u = (u_1, u_2, u_3, u_4, u_5) \in \mathbb{R}^5$. Notice

$$u_2v_1 + u_4v_2 + u_5v_3 + (u_1 - 2u_2)v_4 + (u_3 - 6u_4)v_5 =$$
$$\begin{pmatrix} 3u_2 \\ u_2 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 7u_4 \\ u_4 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ u_5 \end{pmatrix} + \begin{pmatrix} u_1 - 2u_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ u_3 - 6u_4 \\ 0 \\ 0 \end{pmatrix}.$$

Simplifying the RHS, we have

$$\begin{pmatrix} 3u_2 \\ u_2 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 7u_4 \\ u_4 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ u_5 \end{pmatrix} + \begin{pmatrix} u_1 - 2u_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ u_3 - 6u_4 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{pmatrix},$$

and so indeed $v_1, \ldots, v_5$ span $\mathbb{R}^5$. To see that they are linearly independent, suppose $a_1, \ldots, a_5 \in \mathbb{R}$ are such that

$$a_1 \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 0 \\ 7 \\ 1 \\ 0 \end{pmatrix} + a_3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + a_4 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + a_5 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We have the equivalent system of linear equations

$$\begin{aligned} 3a_1 + a_4 &= 0 \\ a_1 &= 0 \\ 7a_2 + a_5 &= 0 \\ a_2 &= 0 \\ a_3 &= 0, \end{aligned}$$

which clearly implies each of the a_k are 0. Hence $v_1, \ldots, v_5$ are linearly independent as well, and thus a basis.

(c) Let $W = \operatorname{span}(v_4, v_5)$, where v_4 and v_5 are defined as in (b). We claim $\mathbb{R}^5 = U \oplus W$. To see $\mathbb{R}^5 = U + W$, let $v \in \mathbb{R}^5$. Then, because we've already shown $v_1, \ldots, v_5$ span $\mathbb{R}^5$, there exist $a_1, \ldots, a_5 \in \mathbb{R}$ such that

$$u = (a_1v_1 + a_2v_2 + a_3v_3) + (a_4v_4 + a_5v_5).$$

The first term in parentheses is an element of U, and the second is an element of W, and thus $V = U + W$.

To prove the sum is direct, it suffices to show $U \cap W = \{0\}$. So suppose $u \in U \cap W$. Then there exist $a_1, a_2, a_3, b_1, b_2 \in \mathbb{R}$ such that

$$v = a_1v_1 + a_2v_2 + a_3v_3 = b_1v_4 + b_2v_5.$$

Thus

$$a_1v_1 + a_2v_2 + a_3v_3 - b_1v_4 - b_2v_5 = 0.$$

Since $v_1, \ldots, v_5$ are linearly independent, this implies each of the a's and b's are 0, and so indeed $U \cap W = \{0\}$. Therefore the sum is direct, proving our claim that $\mathbb{R}^5 = U \oplus W$. $\square$

Problem 5

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

Proof. Consider the list

$$p_0 = 1, p_1 = X, p_2 = X^3 + X^2, p_3 = X^3$$

which contains no polynomial of degree 2. We claim this list is a basis. First we prove $\operatorname{span}(p_0, p_1, p_2, p_3) = \mathcal{P}_3(\mathbb{F})$. Let $q \in \mathcal{P}_3(\mathbb{F})$. Then there exist $a_0, \ldots, a_3 \in \mathbb{F}$ (some of which may be 0) such that

$$q = a_0 + a_1X + a_2X^2 + a_3X^3.$$

But notice

$$\begin{aligned} a_0p_0 + a_1p_1 + a_2p_2 + (a_3 - a_2)p_3 &= a_0 + a_1X + a_2(X^3 + X^2) + (a_3 - a_2)X^3 \\ &= a_0 + a_1X + a_2X^2 + a_3X^3 \\ &= q, \end{aligned}$$

and so indeed p_0, p_1, p_2, p_3 spans $\mathcal{P}_3(\mathbb{F})$. To see the list is linearly independent, suppose $b_0, \ldots, b_3 \in \mathbb{F}$ are such that

$$b_0p_0 + b_1p_1 + b_2p_2 + b_3p_3 = 0.$$

It follows that

$$b_0 + b_1X + b_2X^2 + (b_2 + b_3)X^3 = 0$$

which is true iff all coefficients are zero. Hence we must have $b_0 = b_1 = b_2 = b_3 = 0$, and so $p_0, \ldots, p_3$ is linearly independent. Thus it is a basis, as claimed. $\square$

Problem 7

Prove or give a counterexample: If v_1, v_2, v_3, v_4 is a basis of V and U is a subspace of V such that $v_1, v_2 \in U$ and $v_3 \notin U$ and $v_4 \notin U$, then v_1, v_2 is a basis of U.

Proof. The statement is false. To see this, let $V = \mathbb{R}^4$ and let

$$v_1 = (1,0,0,0), v_2 = (0,1,0,0), v_3 = (0,0,1,0), v_4 = (0,0,0,1).$$

Define

$$U = \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 \mid x_3 = x_4\}.$$

We have $v_1, v_2 \in U$ and $v_3, v_4 \notin U$. But since no linear combination of v_1, v_2 yields $(0,0,1,1)$, v_1, v_2 do not span U, and hence they cannot form a basis. □

C: Dimension

Problem 1

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

Proof. Let $n = \dim U = \dim V$, and let $u_1, \ldots, u_n$ be a basis for U. Since this list is linearly independent and has length equal to the dimension of V, it must be a basis for V as well (by Theorem 2.39). Clearly we have $U \subseteq V$, so it remains to show $V \subseteq U$. Let $v \in V$. Then there exist $a_1, \ldots, a_n \in \mathbb{F}$ such that

$$v = a_1u_1 + \cdots + a_nu_n.$$

But now v is expressed as a linear combination of vectors in U and hence is in U as well. Thus $U = V$, as desired. □

Problem 3

Show that the subspaces of $\mathbb{R}^3$ are precisely $\{0\}$, $\mathbb{R}^3$, all lines in $\mathbb{R}^2$ through the origin, and all planes in $\mathbb{R}^3$ through the origin.

Proof. A subspace of $\mathbb{R}^3$ can have a basis of length $0, 1, 2$ or 3. We consider each in turn:

0: The only basis of length 0 is the empty basis, which generates $\{0\}$.

1: Any basis of length 1 contains a single $x \in \mathbb{R}^\times$. Notice $\operatorname{span}(x) = \{ax \in \mathbb{R} \mid a \in \mathbb{R}\}$, and hence bases of length 1 generate lines through the origin.

2: Any basis of length 2 consists of two linearly independent $x, y \in \mathbb{R}^\times$. Notice $\operatorname{span}(x,y) = \{ax + by \in \mathbb{R}^2 \mid a, b \in \mathbb{R}\}$, and hence bases of length 2 generate planes through the origin.

3: Any basis of length 3 is simply a basis of $\mathbb{R}^3$ and hence generates all of $\mathbb{R}^3$.

Since we've exhausted all possibilities, all subspaces of $\mathbb{R}^3$ have been classified as one of these four types. □

Problem 4

(a) Let $U = \{p \in \mathcal{P}_4(\mathbb{F}) \mid p(6) = 0\}$. Find a basis of U.

(b) Extend the basis in part (a) to a basis of $\mathcal{P}_4(\mathbb{F})$.

(c) Find a subspace W of $\mathcal{P}_4(\mathbb{F})$ such that $\mathcal{P}_4(\mathbb{F}) = U \oplus W$.

We first prove a helpful lemma that we will use repeatedly.

Lemma 1. *Any list of nonzero polynomials in $\mathcal{P}(\mathbb{F})$, no two of which have the same degree, is linearly independent.*

Proof of the lemma. Let $p_1, \dots, p_n \in \mathcal{P}(\mathbb{F})$ be nonzero and each of unique degree, and without loss of generality suppose they are ordered from smallest degree to largest. Denote their degrees by $d_1, \dots, d_n$. Now suppose $a_1, \dots, a_n \in \mathbb{F}$ are such that

$$a_1 p_1 + \cdots + a_n p_n = 0.$$

Without explicitly expanding the LHS, we see that it must have an X^{d_n} term with a nonzero coefficient (since each polynomial is assumed to have unique degree). Since the RHS is identically 0, this implies $a_n = 0$. But now by repeating this same argument $n-1$ times, we see that in fact each of $a_1, \dots, a_{n-1}$ must be zero as well, and hence the list is indeed linearly independent. □

Proof. (a) We claim the list of polynomials

$$(X-6), (X-6)^2, (X-6)^3, (X-6)^4$$

is a basis of U. By Lemma 1, since each polynomial in the list has unique degree, the list is linearly independent. Thus $\dim U$ must be at least 4, since we've demonstrated a linearly independent list of length 4. Since U is a subspace of $\mathcal{P}_4(\mathbb{F})$, which has dimension 5, this implies $\dim U \in \{4, 5\}$. But notice U is a *proper* subset of $\mathcal{P}_4(\mathbb{F})$ since, in particular, it excludes the monomial X. Thus $\dim U$ cannot be 5, and we conclude $\dim U = 4$. Since our list is linearly independent and of length equal to $\dim U$, it must be a basis.

(b) We claim

$$1, (X-6), (X-6)^2, (X-6)^3, (X-6)^4$$

is an extension of our basis of U to $\mathcal{P}_4(\mathbb{F})$. Since this list is of length equal to $\dim \mathcal{P}_4(\mathbb{F})$, it suffices to show it is linearly independent. But this follows immediate by Lemma 1, since each polynomial in the list has unique degree.

(c) Let $W = \mathbb{F}$. We claim $\mathcal{P}_4(\mathbb{F}) = U \oplus W$. Label our basis from (b) as

$$p_0 = 1, p_1 = (X-6), p_2 = (X-6)^2, p_3 = (X-6)^3, p_4 = (X-6)^4.$$

In this notation, we have $W = \operatorname{span}(p_0)$ and $U = \operatorname{span}(p_1, \ldots, p_4)$. Clearly $\mathcal{P}_4(\mathbb{F}) = U + W$ since $p_0, \ldots, p_4$ is a basis of $\mathcal{P}_4(\mathbb{F})$, so it suffices to show $U \cap W = \{0\}$. Suppose $q \in U \cap W$. Then q must be a scalar by inclusion in W. If q were nonzero, there would exist $a_0, \ldots, a_3 \in \mathbb{F}$ such that

$$a_0(X-6) + a_1(X-6)^2 + a_2(X-6)^3 + a_3(X-6)^4 \neq 0$$

for all $X \in \mathbb{F}$. But this is absurd, since the LHS evaluates to 0 for $X = 6$. Thus q cannot be nonzero, and the sum is indeed direct. □

Problem 7

(a) Let $U = \{p \in \mathcal{P}_4(\mathbb{F}) \mid p(2) = p(5) = p(6)\}$. Find a basis of U.

(b) Extend the basis in part (a) to a basis of $\mathcal{P}_4(\mathbb{F})$.

(c) Find a subspace W of $\mathcal{P}_4(\mathbb{F})$ such that $\mathcal{P}_4(\mathbb{F}) = U \oplus W$.

Proof. (a) We claim the list of polynomials

$$1, (X-2)(X-5)(X-6), (X-2)(X-5)(X-6)^2 \tag{†}$$

is a basis of U. Linear independence follows from Lemma 1, and so $\dim U$ must be at least 3. We will exhibit a proper subspace V of $\mathcal{P}_4(\mathbb{F})$ of dimension 4 such that U is a proper subspace of V. This will in turn imply that $3 \leq \dim U < 4$. Since all dimensions are of course integers, this will imply $\dim U = 3$. Since our list of polynomials is a linearly independent list of length equal to $\dim U$, this will prove it to be a basis. So consider the subspace

$$V = \{p \in \mathcal{P}_4(\mathbb{F}) \mid p(2) = p(5)\}$$

of $\mathcal{P}_4(\mathbb{F})$. Clearly U is a subspace of V, and moreover it is a proper subspace since $(X-2)(X-5)$ is in V but not in U. So it only remains to show $\dim V = 4$. Note that the list of polynomials

$$1, (X-2)(X-5), (X-2)^2(X-5), (X-2)^2(X-5)^2$$

is linearly independent in V (again by Lemma 1). Note also that V a proper subspace of $\mathcal{P}_4(\mathbb{F})$ since it does not contain the monomial X. Since this implies $4 \leq \dim V < 5$, we must have $\dim V = 4$, completing the proof that (†) is indeed a basis of U.

(b) We claim

$$1, X, X^2, (X-2)(X-5)(X-6), (X-2)(X-5)(X-6)^2$$

is an extension of our basis of U to $\mathcal{P}_4(\mathbb{F})$. Since this list is of length equal to $\dim \mathcal{P}_4(\mathbb{F})$, it suffices to show it is linearly independent. But this follows immediately from Lemma 1.

(c) Label our basis from (b) as

$$\begin{aligned} p_0 &= 1, \\ p_1 &= X, \\ p_2 &= X^2, \\ p_3 &= (X-2)(X-5)(X-6), \\ p_4 &= (X-2)(X-5)(X-6)^2 \end{aligned}$$

and let $W = \operatorname{span}(p_1, p_2)$. We claim $\mathcal{P}_4(\mathbb{F}) = U \oplus W$. That $\mathcal{P}_4(\mathbb{F}) = U + W$ follows from the fact that $p_0, \ldots, p_4$ is a basis of $\mathcal{P}_4(\mathbb{F})$ and since $U = \operatorname{span}(p_0, p_3, p_4)$. To prove the sum is direct, it suffices to show $U \cap W = \{0\}$. So suppose $q \in U \cap W$. Then there exist $a_0, a_1, b_0, b_1, b_2 \in \mathbb{F}$ such that

$$q = a_0 p_1 + a_1 p_2 = b_0 p_0 + b_1 p_3 + b_2 p_4.$$

But then

$$a_0 p_1 + a_1 p_2 - b_0 p_0 - b_1 p_3 - b_2 p_4 = 0,$$

and since the $p_0, \ldots, p_4$ are linearly independent, this implies each of the a's and b's are zero. Thus $q = 0$ and the sum is indeed direct. □

Problem 9

Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Prove that

$$\dim \operatorname{span}(v_1 + w, \ldots, v_m + w) \geq m - 1.$$

Proof. Let $W = \operatorname{span}(v_1 + w, \ldots, v_m + w)$, and consider the list

$$v_2 - v_1, v_3 - v_2, \ldots, v_m - v_{m-1},$$

which has length $m-1$. Note that $v_k - v_{k-1} = (v_k + w) - (v_{k-1} + w)$, so that each vector in this list is indeed in W. Since the dimension of W must be greater than the length of any linearly independent list, if we prove this list is linearly independent, we will have proved $\dim W \geq m - 1$. So suppose $a_1, \ldots a_{m-1} \in \mathbb{F}$ are such that

$$a_1(v_2 - v_1) + \cdots + a_{m-1}(v_m - v_{m-1}) = 0.$$

Expanding, we see

$$(-a_1)v_1 + (a_1 - a_2)v_2 + \cdots + (a_{m-2} - a_m)v_{m-1} = 0.$$

But since $v_1, \ldots, v_{m-1}$ is linearly independent by hypothesis, each of the coefficients must be zero. Thus $a_1 = 0$ and $a_{k-1} = a_k$ for $k = 2, \ldots, m-1$, and hence we must have $a_2 = \cdots = a_{m-1} = 0$ as well. Therefore, our list is linearly independent, and indeed $\dim W \geq m - 1$. □

Problem 11

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

Proof. We have

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

and thus since $U + W = \mathbb{R}^8$, $\dim U = 3$, and $\dim W = 5$, it follows

$$8 = 3 + 5 - \dim(U \cap W),$$

and hence $\dim(U \cap W) = 0$. Therefore we must have $U \cap W = \{0\}$, and hence $\mathbb{R}^8 = U \oplus W$. $\square$

Problem 13

Suppose U and W are both 4-dimensional subspaces of $\mathbb{C}^6$. Prove that there exist two vectors in $U \cap W$ such that neither of these vectors is a scalar multiple of the other.

Proof. Note that we view $\mathbb{C}^6$ as a vector space over $\mathbb{C}$. We have

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

and thus since $\dim U = \dim W = 4$, it follows

$$\dim(U + W) = 8 - \dim(U \cap W). \tag{1}$$

Since $U + W$ is a subspace of $\mathbb{C}^6$ and $\dim \mathbb{C}^6 = 6$, and since $\dim(U + W) \geq \max\{\dim U, \dim W\} = 4$, we have

$$4 \leq \dim(U + W) \leq 6. \tag{2}$$

Combining (1) and (2) yields

$$-4 \leq -\dim(U \cap W) \leq -2,$$

and hence

$$2 \leq \dim(U \cap W) \leq 4.$$

Thus $U \cap W$ has a basis of length at least two, and thus there exist two vectors in $U \cap W$ such that neither is a scalar multiple of the other (namely, two vectors in the basis). $\square$

Problem 14

Suppose $U_1, \ldots, U_m$ are finite-dimensional subspace of V. Prove that $U_1 + \cdots + U_m$ is finite-dimensional and

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

Proof. For each $j = 1, \ldots, m$, choose a basis for U_j. Combine these bases to form a single list of vectors in V. Clearly this list spans $U_1 + \cdots + U_m$ by construction. Hence $U_1 + \cdots + U_m$ is finite-dimensional with dimension less than or equal to the number of vectors in this list, which is equal to $\dim U_1 + \cdots + \dim U_m$. That is,

$$\dim\left(U_1 + \cdots + U_m\right) \leq \dim U_1 + \cdots + \dim U_m,$$

as desired. □

Problem 15

Suppose V is finite-dimensional, with $\dim V = n \geq 1$. Prove that there exist 1-dimensional subspaces $U_1, \ldots, U_n$ of V such that

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

Proof. Since $\dim V = n$, there exists a basis $v_1, \ldots, v_n$ of V. Let $U_k = \operatorname{span}(v_k)$ for $k = 1, \ldots, n$, so that each U_k has dimension 1. Clearly

$$V = U_1 + \cdots + U_n,$$

so it remains to show this sum is direct. If $u \in U_1 + \cdots + U_n$, there exist $a_1, \ldots, a_n \in \mathbb{F}$ such that

$$u = a_1 v_1 + \cdots + a_n v_n.$$

But since $v_1, \ldots, v_n$ is a basis, this representation of u as a linear combination of $v_1, \ldots, v_n$ is unique, and thus the sum is direct, as desired. □

Problem 16

Suppose $U_1, \ldots, U_m$ are finite-dimensional subspaces of V such that $U_1 + \cdots + U_m$ is a direct sum. Prove that $U_1 \oplus \cdots \oplus U_m$ is finite-dimensional and

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

Proof. For each $j = 1, \ldots, m$, choose a basis for U_j. Combine these bases to form a single list of vectors in V. Clearly this list spans $U_1 + \cdots + U_m$ by construction, so that $U_1 + \cdots + U_m$ is finite-dimensional. We claim this list must be linearly independent, hence it will be a basis of length $\dim U_1 + \cdots + \dim U_m$, and thus

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

So suppose some linear combination of the vectors in this list equals 0. For $k = 1, \ldots, m$, denote by u_k the sum of all terms in that linear combination which are formed from our chosen basis of U_k. Then we have

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

Since $U_1 + \cdots + U_m = U_1 \oplus \cdots \oplus U_m$, each u_k must equal 0. But then, since u_k is a linear combination of a basis of U_k, each of the coefficients in that linear combination must equal 0. Thus all coefficients in our original linear combination must be 0. That is, our basis is linearly independent, justifying our claim and completing the proof. $\square$

Problem 17

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

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

Prove this or give a counterexample.

Proof. The statement is false. Consider

$$U_1 = \mathbb{R} \times \{0\},\ U_2 = \{(x, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\},\ U_3 = \{0\} \times \mathbb{R}.$$

We have

$$\begin{aligned}&\dim(U_1 + U_2 + U_3) = \dim \mathbb{R}^2 = 2 \\ &\dim U_1 = \dim U_2 = \dim U_3 = 1 \\ &\dim(U_1 \cap U_2) = \dim(U_2 \cap U_3) = 1 \\ &\dim(U_1 \cap U_3) = \dim(U_1 \cap U_2 \cap U_3) = 0,\end{aligned}$$

and therefore

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

since the LHS is 2, whereas the RHS is 1 in this case. $\square$

Chapter 3: Linear Maps

Linear Algebra Done Right, by Sheldon Axler

A: The Vector Space of Linear Maps

Problem 1

Suppose $b, c \in \mathbb{R}$. Define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by

$$T(x, y, z) = (2x - 4y + 3z + b, 6x + cxyz).$$

Show that T is linear if and only if $b = c = 0$.

Proof. ($\Leftarrow$) Suppose $b = c = 0$. Then

$$T(x, y, z) = (2x - 4y + 3z, 6x).$$

Let $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$. Then

$$\begin{aligned}
T((x_1, y_1, z_1) + (x_2, y_2, z_2)) &= T(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\
&= (2(x_1 + x_2) - 4(y_1 + y_2) + 3(z_1 + z_2), 6(x_1 + x_2)) \\
&= (2x_1 + 2x_2 - 4y_1 - 4y_2 + 3z_1 + 3z_2, 6x_1 + 6x_2) \\
&= (2x_1 - 4y_1 + 3z_1, 6x_1) + (2x_2 - 4y_2 + 3z_2, 6x_2) \\
&= T(x_1, y_1, z_1) + T(x_2, y_2, z_2).
\end{aligned}$$

Now, for $\lambda \in \mathbb{F}$ and $(x, y, z) \in \mathbb{R}^3$, we have

$$\begin{aligned}
T(\lambda(x, y, z)) &= T(\lambda x, \lambda y, \lambda z) \\
&= (2(\lambda x) - 4(\lambda y) + 3(\lambda z), 6(\lambda x)) \\
&= (\lambda(2x - 4y + 3z), \lambda(6x)) \\
&= \lambda(2x - 4y + 3z, 6x) \\
&= \lambda T(x, y, z),
\end{aligned}$$

and thus T is a linear map.

($\Rightarrow$) Supose T is a linear map. Then

$$T(x_1 + x_2, y_1 + y_2, z_1 + z_2) = T(x_1, y_1, z_1) + T(x_2, y_2, z_2) \tag{$\dagger$}$$

for all $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$. In particular, by applying the definition of T and comparing first coordinates of both sides of ($\dagger$), we have

$$\begin{aligned}
&2(x_1 + x_2) - 4(y_1 + y_2) + 3(z_1 + z_2) + b = \\
&\qquad (2x_1 - 4y_1 + 3z_1 + b) + (2x_2 - 4y_2 + 3z_2 + b),
\end{aligned}$$

and after simplifying, we have $b = 2b$, and hence $b = 0$. Now by applying the definition of T and comparing second coordinates of both sides of (†), we have

$$\begin{aligned} 6(x_1 + x_2) + c(x_1 + x_2)(y_1 + y_2)(z_1 + z_2) &= 6x_1 + c(x_1 y_1 z_1) + 6x_2 + c(x_2 y_2 z_2) \\ &= 6(x_1 + x_2) + c(x_1 y_1 z_1 + x_2 y_2 z_2), \end{aligned}$$

which implies

$$c(x_1 + x_2)(y_1 + y_2)(z_1 + z_2) = c(x_1 y_1 z_1 + x_2 y_2 z_2).$$

Now suppose $c \neq 0$. Then choosing $(x_1, y_1, z_1) = (x_2, y_2, z_2) = (1, 1, 1)$, the equation above implies $8 = 2$, a contradiction. Thus $c = 0$, completing the proof. □

Problem 3

Suppose $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. Show that there exist scalars $A_{j,k} \in \mathbb{F}$ for $j = 1, \ldots, m$ and $k = 1, \ldots, n$ such that

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

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

Proof. Given $x \in \mathbb{F}^n$, we may write

$$x = x_1 e_1 + \cdots + x_n e_n,$$

where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{F}^n$. Since T is linear, we have

$$Tx = T(x_1 e_1 + \cdots + x_n e_n) = x_1 T e_1 + \cdots + x_n T e_n.$$

Now for each $Te_k \in \mathbb{F}^m$, where $k = 1, \ldots, n$, there exist $A_{1,k}, \ldots, A_{m,k} \in \mathbb{F}$ such that

$$\begin{aligned} Te_k &= A_{1,k} e_1 + \cdots + A_{m,k} e_m \\ &= \big(A_{1,k}, \ldots, A_{m,k}\big) \end{aligned}$$

and thus

$$x_k T e_k = \big(A_{1,k} x_k, \ldots, A_{m,k} x_k\big).$$

Therefore, we have

$$\begin{aligned} Tx &= \sum_{k=1}^{n} \big(A_{1,k} x_k, \ldots, A_{m,k} x_k\big) \\ &= \left(\sum_{k=1}^{n} A_{1,k} x_k, \ldots, \sum_{k=1}^{n} A_{m,k} x_k \right), \end{aligned}$$

and thus there exist scalars $A_{j,k} \in \mathbb{F}$ for $j = 1, \ldots, m$ and $k = 1, \ldots, n$ of the desired form. □

Problem 5

Prove that $\mathcal{L}(V, W)$ is a vector space.

Proof. We check each property in turn.
Commutative: Given $S, T \in \mathcal{L}(V, W)$ and $v \in V$, we have

$$(T + S)(v) = Tv + Sv = Sv + Tv = (S + T)(v)$$

and so addition is commutative.
Associative: Given $R, S, T \in \mathcal{L}(V, W)$ and $v \in V$, we have

$$\begin{aligned}((R + S) + T)(v) = (R + S)(v) + Tv &= Rv + Sv + Tv \\ &= R + (S + T)(v) = (R + (S + T))(v)\end{aligned}$$

and so addition is associative. And given $a, b \in \mathbb{F}$, we have

$$((ab)T)(v) = (ab)(Tv) = a(b(Tv)) = (a(bT))(v)$$

and so scalar multiplication is associative as well.
Additive identity: Let $0 \in \mathcal{L}(V, W)$ denote the zero map, let $T \in \mathcal{L}(V, W)$, and let $v \in V$. Then

$$(T + 0)(v) = Tv + 0v = Tv + 0 = Tv$$

and so the zero map is the additive identity.
Additive inverse: Let $T \in \mathcal{L}(V, W)$ and define $(-T) \in \mathcal{L}(V, W)$ by $(-T)v = -Tv$. Then

$$(T + (-T))(v) = Tv + (-T)v = Tv - Tv = 0,$$

and so $(-T)$ is the additive inverse for each $T \in \mathcal{L}(V, W)$.
Multiplicative identity: Let $T \in \mathcal{L}(V, W)$. Then

$$(1T)(v) = 1(Tv) = Tv$$

and so the multiplicative identity of $\mathbb{F}$ is the multiplicative identity of scalar multiplication.
Distributive properties: Let $S, T \in \mathcal{L}(V, W)$, $a, b \in \mathbb{F}$, and $v \in V$. Then

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

and

$$((a + b)T)(v) = (a + b)(Tv) = a(Tv) + b(Tv) = (aT)(v) + (bT)(v)$$

and so the distributive properties hold.

Since all properties of a vector space hold, we see $\mathcal{L}(V, W)$ is in fact a vector space, as desired. $\square$

Problem 7

Show that every linear map from a 1-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 $\lambda \in \mathbb{F}$ such that $Tv = \lambda v$ for all $v \in V$.

Proof. Since $\dim V = 1$, a basis of V consists of a single vector. So let $w \in V$ be such a basis. Then there exists $\alpha \in \mathbb{F}$ such that $v = \alpha w$ and $\lambda \in \mathbb{F}$ such that $Tw = \lambda w$. It follows

$$Tv = T(\alpha w) = \alpha T w = \alpha \lambda w = \lambda(\alpha w) = \lambda v,$$

as desired. $\square$

Problem 9

Give an example of a function $\varphi : \mathbb{C} \to \mathbb{C}$ such that

$$\varphi(w + z) = \varphi(w) + \varphi(z)$$

for all $w, z \in \mathbb{C}$ but φ is not linear. (Here $\mathbb{C}$ is thought of as a complex vector space.)

Proof. Define

$$\begin{aligned} \varphi : \mathbb{C} &\to \mathbb{C} \\ x + yi &\mapsto x - yi. \end{aligned}$$

Then for $x_1 + y_1 i, x_2 + y_2 i \in \mathbb{C}$, it follows

$$\begin{aligned} \varphi((x_1 + y_1 i) + (x_2 + y_2 i)) &= \varphi((x_1 + x_2) + (y_1 + y_2)i) \\ &= (x_1 + x_2) - (y_1 + y_2)i \\ &= (x_1 - y_1)i + (x_2 - y_2)i \\ &= \varphi(x_1 + y_1 i) + \varphi(x_2 + y_2 i) \end{aligned}$$

and so φ satisfies the additivity requirement. However, we have

$$\varphi(i \cdot i) = \varphi(-1) = -1$$

and

$$i \cdot \varphi(i) = i(-i) = 1$$

and hence φ fails the homogeneity requirement of a linear map. $\square$

Problem 11

Suppose V is finite-dimensional. Prove that every 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 $u \in U$.

Proof. Suppose U is a subspace of V and $S \in \mathcal{L}(U, W)$. Let $v_1, \ldots, u_m$ be a basis of U and let $v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ be an extension of this basis to V. For any $z \in V$, there exist $a_1, \ldots, a_n \in \mathbb{F}$ such that $z = \sum_{k=1}^n a_k v_k$, and so we define

$$\begin{aligned} T : V &\to W \\ \sum_{k=1}^n a_k v_k &\mapsto \sum_{k=1}^m a_k S v_k + \sum_{k=m+1}^n a_k v_k. \end{aligned}$$

Since every $v \in V$ has a unique representation as a linear combination of elements of our basis, the map is well-defined. We first show T is a linear map. So suppose $z_1, z_2 \in V$. Then there exist $a_1, \ldots a_n \in \mathbb{F}$ and $b_1, \ldots, b_n \in \mathbb{F}$ such that

$$z_1 = a_1 v_1 + \cdots + a_n v_n \quad \text{and} \quad z_2 = b_1 v_1 + \cdots + b_n v_n.$$

It follows

$$\begin{aligned} T(z_1 + z_2) &= T\left(\sum_{k=1}^n a_k v_k + \sum_{k=1}^n b_k v_k\right) \\ &= T\left(\sum_{k=1}^n (a_k + b_k) v_k\right) \\ &= \sum_{k=1}^m (a_k + b_k) S v_k + \sum_{k=m+1}^n (a_k + b_k) v_k \\ &= \left(\sum_{k=1}^m a_k S v_k + \sum_{k=m+1}^n a_k v_k\right) + \left(\sum_{k=1}^m b_k S v_k + \sum_{k=m+1}^n b_k v_k\right) \\ &= T\left(\sum_{k=1}^n a_k v_k\right) + T\left(\sum_{k=1}^n b_k v_k\right) \\ &= Tz_1 + Tz_2 \end{aligned}$$

and so T is additive. To see that T is homogeneous, let $\lambda \in \mathbb{F}$ and $z \in V$, so that we may write $z = \sum_{k=1}^n a_k v_k$ for some $a_1, \ldots, a_n \in \mathbb{F}$. We have

$$\begin{aligned}
T(\lambda z) &= T\left(\lambda \sum_{k=1}^n a_k v_k\right) \\
&= T\left(\sum_{k=1}^n (\lambda a_k) v_k\right) \\
&= S\left(\sum_{k=1}^m (\lambda a_k) v_k\right) + \sum_{k=m+1}^n (\lambda a_k) v_k \\
&= \lambda S\left(\sum_{k=1}^m a_k v_k\right) + \lambda \sum_{k=m+1}^n a_k v_k \\
&= \lambda\left(S\left(\sum_{k=1}^m a_k v_k\right) + \sum_{k=m+1}^n \lambda a_k v_k\right) \\
&= \lambda T\left(\sum_{k=1}^n a_k v_k\right) \\
&= \lambda T z
\end{aligned}$$

and so T is homogeneous as well hence $T \in \mathcal{L}(V, W)$. Lastly, to see that $T\mid_U = S$, let $u \in U$. Then there exist $a_1, \ldots, a_m \in \mathbb{F}$ such that $u = \sum_{k=1}^m a_k v_k$, and hence

$$\begin{aligned}
Tu &= T\left(\sum_{k=1}^m a_k v_k\right) \\
&= S\left(\sum_{k=1}^m a_k v_k\right) \\
&= Su,
\end{aligned}$$

and so indeed T agrees with S on U, completing the proof. $\square$

Problem 13

Suppose $v_1, \ldots, v_m$ is a linearly dependent list of vectors in V. Suppose also that $W \neq \{0\}$. Prove that there exist $w_1, \ldots, w_m \in W$ such that no $T \in \mathcal{L}(V, W)$ satisfies $Tv_k = w_k$ for each $k = 1, \ldots, m$.

Proof. Since $v_1, \ldots, v_m$ is linearly dependent, one of them may be written as a linear combination of the others. Without loss of generality, suppose this is v_m.

Then there exist $a_1, \ldots, a_{m-1} \in \mathbb{F}$ such that

$$v_m = a_1 v_1 + \cdots + a_{m-1} v_{m-1}.$$

Since $W \neq \{0\}$, there exists some nonzero $z \in W$. Define $w_1, \ldots, w_m \in W$ by

$$w_k = \begin{cases} z & \text{if } k = m \\ 0 & \text{otherwise.} \end{cases}$$

Now suppose there exists $T \in \mathcal{L}(V, W)$ such that $Tv_k = w_k$ for $k = 1, \ldots, m$. It follows

$$\begin{aligned} T(0) &= T(v_m - a_1 v_1 - \cdots - a_{m-1} v_{m-1}) \\ &= Tv_m - a_1 Tv_1 - \cdots - a_{m-1} Tv_{m-1} \\ &= z. \end{aligned}$$

But $z \neq 0$, and thus $T(0) \neq 0$, a contradiction, since linear maps take 0 to 0. Therefore, no such linear map can exist. $\square$

B: Null Spaces and Ranges

Problem 1

Give an example of a linear map T such that $\dim \operatorname{null} T = 3$ and $\dim \operatorname{range} T = 2$.

Proof. Define the map

$$\begin{aligned} T : \mathbb{R}^5 &\to \mathbb{R}^5 \\ (x_1, x_2, x_3, x_4, x_5) &\mapsto (0, 0, 0, x_4, x_5). \end{aligned}$$

First we show T is a linear map. Suppose $x, y \in \mathbb{R}^5$. Then

$$\begin{aligned} T(x + y) &= T((x_1, x_2, x_3, x_4, x_5) + (y_1, y_2, y_3, y_4, y_5)) \\ &= T(x_1 + y_1, x_2 + y_2, x_3 + y_3, x_4 + y_4, x_5 + y_5) \\ &= (0, 0, 0, x_4 + y_5, x_5 + y_5) \\ &= (0, 0, 0, x_4, x_5) + (0, 0, 0, y_4, y_5) \\ &= T(x) + T(y). \end{aligned}$$

Next let $\lambda \in \mathbb{R}$. Then

$$\begin{aligned} T(\lambda x) &= T(\lambda x_1, \lambda x_2, \lambda x_3, \lambda x_4, \lambda x_5) \\ &= (0, 0, 0, \lambda x_4, \lambda x_5) \\ &= \lambda(0, 0, 0, x_4, x_5) \\ &= \lambda T(x), \end{aligned}$$

and so T is in fact a linear map. Now notice that

$$\operatorname{null} T = \{(x_1, x_2, x_3, 0, 0) \in \mathbb{R}^5 \mid x_1, x_2, x_3 \in \mathbb{R}\}.$$

This space clearly has as a basis $e_1, e_2, e_3 \in \mathbb{R}^5$ and hence has dimension 3. Now, by the Fundamental Theorem of Linear Maps, we have

$$\dim \mathbb{R}^5 = 3 + \dim \operatorname{range} T$$

and hence $\dim \operatorname{range} T = 2$, as desired. □

Problem 3

Suppose $v_1, \ldots, v_m$ is a list of vectors in V. Define $T \in \mathcal{L}(\mathbb{F}^m, V)$ by

$$T(z_1, \ldots, z_m) = z_1 v_1 + \cdots + z_m v_m.$$

(a) What property of T corresponds to $v_1, \ldots, v_m$ spanning V?

(b) What property of T corresponds to $v_1, \ldots, v_m$ being linearly independent?

Proof. (a) We claim surjectivity of T corresponds to $v_1, \ldots, v_m$ spanning V. To see this, suppose T is surjective, and let $w \in V$. Then there exists $z \in \mathbb{F}^m$ such that $Tz = w$. This yields

$$z_1 v_1 + \cdots + z_m v_m = w,$$

and hence every $w \in V$ can be expressed as a linear combination of $v_1, \ldots, v_n$. That is, $\operatorname{span}(v_1, \ldots, v_n) = V$.

(b) We claim injectivity of T corresponds to $v_1, \ldots, v_m$ being linearly independent. To see this, suppose T is injective, and let $a_1, \ldots, a_n \in \mathbb{F}$ such that

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Then

$$T(a) = T(a_1, \ldots, a_n) = a_1 v_1 + \cdots + a_n v_n = 0$$

which is true iff $a = 0$ since T is injective. That is, $a_1 = \cdots = a_n = 0$ and hence $v_1, \ldots, v_n$ is linearly independent. □

Problem 5

Give an example of a linear map $T : \mathbb{R}^4 \to \mathbb{R}^4$ such that

$$\operatorname{range} T = \operatorname{null} T.$$

Proof. Define

$$T : \mathbb{R}^4 \to \mathbb{R}^4$$
$$(x_1, x_2, x_3, x_4) \mapsto (x_3, x_4, 0, 0).$$

Clearly T is a linear map, and we have

$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \mid x_3 = x_4 = 0 \in \mathbb{R}\} = \mathbb{R}^2 \times \{0\}^2$$

and

$$\operatorname{range} T = \{(x, y, 0, 0) \mid x, y \in \mathbb{R}\} = \mathbb{R}^2 \times \{0\}^2.$$

Hence $\operatorname{range} T = \operatorname{null} T$, as desired. □

Problem 7

Suppose V and W are finite-dimensional with $2 \leq \dim V \leq \dim W$. Show that $\{T \in \mathcal{L}(V, W) \mid T \text{ is not injective}\}$ is not a subspace of $\mathcal{L}(V, W)$.

Proof. Let $Z = \{T \in \mathcal{L}(V, W) \mid T \text{ is not injective}\}$, let $v_1, \ldots, v_m$ be a basis of V, where $m \geq 2$, and let $w_1, \ldots, w_n$ be a basis of W, where $n \geq m$. We define $T \in \mathcal{L}(V, W)$ by its behavior on the basis

$$Tv_k := \begin{cases} 0 & \text{if } k = 1 \\ w_2 & \text{if } k = 2 \\ \frac{1}{2} w_k & \text{otherwise} \end{cases}$$

so that clearly T is not injective since $Tv_1 = 0 = T(0)$, and hence $T \in Z$. Similarly, define $S \in \mathcal{L}(V, W)$ by its behavior on the basis

$$Sv_k := \begin{cases} w_1 & \text{if } k = 1 \\ 0 & \text{if } k = 2 \\ \frac{1}{2} w_k & \text{otherwise} \end{cases}$$

and note that S is not injective either since $Sv_2 = 0 = S(0)$, and hence $S \in Z$. However, notice

$$(S + T)(v_k) = w_k \text{ for } k = 1, \ldots, n$$

and hence $\operatorname{null}(S + T) = \{0\}$ since it takes the basis of V to the basis of W, so that $S + T$ is in fact injective. Therefore $S + T \notin Z$, and Z is not closed under addition. Thus Z is not a subspace of $\mathcal{L}(V, W)$. □

Problem 9

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

Proof. Suppose $a_1, \ldots, a_n \in \mathbb{F}$ are such that

$$a_1 T v_1 + \cdots + a_n T v_n = 0.$$

Since T is a linear map, it follows

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

But since $\operatorname{null} T = \{0\}$ (by virtue of T being a linear map), this implies $a_1 v_1 + \cdots + a_n v_n = 0$. And since $v_1, \ldots, v_n$ are linearly independent, we must have $a_1 = \cdots = a_n = 0$, which in turn implies $Tv_1, \ldots, Tv_n$ is indeed linearly independent in W. $\square$

Problem 11

Suppose $S_1, \ldots, S_n$ are injective linear maps such that $S_1 S_2 \ldots S_n$ makes sense. Prove that $S_1 S_2 \ldots S_n$ is injective.

Proof. For $n \in \mathbb{Z}_{\geq 2}$, let $P(n)$ be the statement: $S_1, \ldots, S_n$ are injective linear maps such that $S_1 S_2 \ldots S_n$ makes sense, and the product $S_1 S_2 \ldots S_n$ is injective. We induct on n.

Base case: Suppose $n = 2$, and assume $S_1 \in \mathcal{L}(V_0, V_1)$ and $S_2 \in \mathcal{L}(V_1, V_2)$, so that the product $S_1 S_2$ is defined, and assume that both S_1 and S_2 are injective. Suppose $v_1, v_2 \in V_0$ are such that $v_1 \neq v_2$, and let $w_1 = S_2 v_1$ and $w_2 = S_2 v$. Since S_2 is injective, $w_1 \neq w_2$. And since S_1 is injective, this in turn implies that $S_1(w_1) \neq S_1(w_2)$. In other words, $S_1(S_2(v_1)) \neq S_1(S_2(v_2))$, so that $S_1 S_2$ is injective as well, and hence $P(2)$ is true.

Inductive step: Suppose $P(k)$ is true for some $k \in \mathbb{Z}^+$, and consider the product $(S_1 S_2 \ldots S_k) S_{k+1}$. The term in parentheses is injective by hypothesis, and the product of this term with S_{k+1} is injective by our base case. Thus $P(k+1)$ is true.

By the principle of mathematical induction, the statement $P(n)$ is true for all $n \in \mathbb{Z}_{\geq 2}$, as was to be shown. $\square$

Problem 13

Suppose T is a linear map from $\mathbb{F}^4$ to $\mathbb{F}^2$ such that

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

Prove that T is surjective.

Proof. We claim the list

$$(5, 1, 0, 0), (0, 0, 7, 1)$$

is a basis of $\operatorname{null} T$. This implies

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

and hence T is surjective (since the only 2-dimensional subspace of $\mathbb{F}^2$ is the space itself). So let's prove our claim that this list is a basis.

Clearly the list is linearly independent. To see that it spans $\operatorname{null} T$, suppose $x = (x_1, x_2, x_3, x_4) \in \operatorname{null} T$, so that $x_1 = 5x_2$ and $x_3 = 7x_4$. We may write

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5x_2 \\ x_2 \\ 7x_4 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} 5 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ 7 \\ 1 \end{pmatrix},$$

and indeed x is in the span of our list, so that our list is in fact a basis, completing the proof. $\square$

Problem 15

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

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

Proof. Suppose such a $T \in \mathcal{L}(\mathbb{F}^5, \mathbb{F}^2)$ did exist. We claim

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

is a basis of $\operatorname{null} T$. This implies

$$\begin{aligned}\dim \operatorname{range} T &= \dim \mathbb{F}^5 - \dim \operatorname{null} T \\ &= 5 - 2 \\ &= 3,\end{aligned}$$

which is absurd, since the codomain of T has dimension 2. Hence such a T cannot exist. So, let's prove our claim that this list is a basis.

Clearly $(3, 1, 0, 0, 0), (0, 0, 1, 1, 1)$ is linearly independent. To see that it spans $\operatorname{null} T$, suppose $x = (x_1, \ldots, x_5) \in \operatorname{null} T$, so that $x_1 = 3x_2$ and $x_3 = x_4 = x_5$. We may write

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3x_2 \\ x_2 \\ x_3 \\ x_3 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

and indeed x is in the span of our list, so that our list is in fact a basis, completing the proof. $\square$

Problem 17

Suppose V and W are both finite-dimensional. Prove that there exists an injective linear map from V to W if and only if $\dim V \leq \dim W$.

Proof. ($\Rightarrow$) Suppose $T \in \mathcal{L}(V, W)$ is injective. If $\dim V > \dim W$, Thereom 3.23 tells us that no map from V to W would be injective, a contradiction, and so we must have $\dim V \leq \dim W$.

($\Leftarrow$) Suppose $\dim V \leq \dim W$. Then the inclusion map $\iota : V \to W$ is both a linear map and injective. $\square$

Problem 19

Suppose 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 $\dim U \geq \dim V - \dim W$.

Proof. ($\Leftarrow$) Suppose $\dim U \geq \dim V - \dim W$. Since U is a subspace of V, there exists a subspace U' of V such that

$$V = U \oplus U'.$$

Let $u_1, \dots u_m$ be a basis for U, let $u_1', \dots, u_n'$ be a basis for U', and let $w_1, \dots, w_p$ be a basis for W. By hypothesis, we have

$$m \geq (m + n) - p,$$

which implies $p \geq n$. Thus we may define a linear map $T \in \mathcal{L}(V, W)$ by its values on the basis of $V = U \oplus U'$ by taking $Tu_k = 0$ for $k = 1, \dots m$ and $Tu_j' = w_j$ for $j = 1, \dots, n$ (since $p \geq n$, there is a w_j for each u_j'). The map is linear by Theorem 3.5, and its null space is U by construction.

($\Rightarrow$) Suppose U is a subspace of V, $T \in \mathcal{L}(V, W)$, and $\operatorname{null} T = U$. Then, since $\operatorname{range} T$ is a subspace of W, we have $\dim \operatorname{range} T \leq \dim W$. Combining this inequality with the Fundamental Theorem of Linear Maps yields

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

Since $\dim \operatorname{null} T = \dim U$, we have the desired inequality. $\square$

Problem 21

Suppose 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.

Proof. ($\Rightarrow$) Suppose $T \in \mathcal{L}(V, W)$ is surjective, so that W is necessarily finite-dimensional as well. Let $v_1, \dots, v_m$ be a basis of V and let $n = \dim W$, where $m \geq n$ by surjectivity of T. Note that

$$Tv_1, \dots, Tv_m$$

span W. Thus we may reduce this list to a basis by removing some elements (possibly none, if $n = m$). Suppose this reduced list were $Tv_{i_1}, \dots, Tv_{i_n}$ for some $i_1, \dots, i_n \in \{1, \dots, m\}$. We define $S \in \mathcal{L}(W, V)$ by its behavior on this basis

$$S(Tv_{i_k}) := v_{i_k} \text{ for } k = 1, \dots, n.$$

Suppose $w \in W$. Then there exist $a_1, \dots, a_n \in \mathbb{F}$ such that

$$w = a_1 T v_{i_1} + \cdots + a_n T v_{i_n}$$

and thus

$$\begin{aligned} TS(w) &= TS\left(a_1 T v_{i_1} + \cdots + a_n T v_{i_n}\right) \\ &= T\big(S\left(a_1 T v_{i_1} + \cdots + a_n T v_{i_n}\right)\big) \\ &= T\big(a_1 S(Tv_{i_1}) + \cdots + a_n S(Tv_{i_n})\big) \\ &= T(a_1 v_{i_1} + \cdots + a_n v_{i_n}) \\ &= a_1 T v_{i_1} + \cdots + a_n T v_{i_n} \\ &= w, \end{aligned}$$

and so TS is the identity map on W.

($\Leftarrow$) Suppose there exists $S \in \mathcal{L}(W, V)$ such that $TS \in \mathcal{L}(W, W)$ is the identity map, and suppose by way of contradiction that T is not surjective, so that $\dim \operatorname{range} TS < \dim W$. By the Fundamental Theorem of Linear Maps, this implies

$$\begin{aligned} \dim W &= \dim \operatorname{null} TS + \dim \operatorname{range} TS \\ &< \dim \operatorname{null} TS + \dim W \end{aligned}$$

and hence $\dim \operatorname{null} TS > 0$, a contradiction, since the identity map can only have trivial null space. Thus T is surjective, as desired. ☐

Problem 23

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

$$\dim \operatorname{range} ST \leq \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

Proof. We will show that both $\dim \operatorname{range} ST \leq \dim \operatorname{range} S$ and $\dim \operatorname{range} ST \leq \dim \operatorname{range} T$, since this implies the desired inequality.

We first show that $\dim \operatorname{range} ST \leq \dim \operatorname{range} S$. Suppose $w \in \operatorname{range} ST$. Then there exists $u \in U$ such that $ST(u) = w$. But this implies that $w \in \operatorname{range} S$ as well, since $Tu \in S^{-1}(w)$. Thus $\operatorname{range} ST \subseteq \operatorname{range} S$, which implies $\dim \operatorname{range} ST \leq \dim \operatorname{range} S$.

We now show that $\dim \operatorname{range} ST \leq \dim \operatorname{range} T$. Note that if $v \in \operatorname{null} T$, so that $Tv = 0$, then $ST(v) = 0$ (since linear maps take zero to zero). Thus we have $\operatorname{null} T \subseteq \operatorname{null} ST$, which implies $\dim \operatorname{null} T \leq \dim \operatorname{null} ST$. Combining this inequality with the Fundamental Theorem of Linear Maps applied to T yields

$$\dim U \leq \dim \operatorname{null} ST + \dim \operatorname{range} T. \tag{1}$$

Similarly, we have

$$\dim U = \dim \operatorname{null} ST + \dim \operatorname{range} ST. \tag{2}$$

Combining (1) and (2) yields

$$\dim \operatorname{null} ST + \dim \operatorname{range} ST \leq \dim \operatorname{null} ST + \dim \operatorname{range} T$$

and hence $\dim \operatorname{range} ST \leq \dim \operatorname{range} T$, completing the proof. □

Problem 25

Suppose V is finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} T_1 \subseteq \operatorname{range} T_2$ if and only if there exists $S \in \mathcal{L}(V, V)$ such that $T_1 = T_2 S$.

Proof. ($\Leftarrow$) Suppose there exists $S \in \mathcal{L}(V, V)$ such that $T_1 = T_2 S$, and let $w \in \operatorname{range} T_1$. Then there exists $v \in V$ such that $T_1 v = w$, and hence $T_2 S(v) = w$. But then $w \in \operatorname{range} T_2$ as well, and hence $\operatorname{range} T_1 \subseteq \operatorname{range} T_2$.

($\Rightarrow$) Suppose $\operatorname{range} T_1 \subseteq \operatorname{range} T_2$, and let $v_1, \ldots, v_n$ be a basis of V. Let $w_k = T v_k$ for $k = 1, \ldots, n$. Then there exist $u_1, \ldots, u_n \in V$ such that $T_2 u_k = w_k$ for $k = 1, \ldots, n$ (since $w_k \in \operatorname{range} T_1$ implies $w_k \in \operatorname{range} T_2$). Define $S \in \mathcal{L}(V, V)$ by its behavior on the basis

$$S v_k := u_k \text{ for } k = 1, \ldots, n.$$

It follows that $T_2 S(v_k) = T_2 u_k = w_k = T_1 v_k$. Since $T_2 S$ and T_1 are equal on the basis, they are equal as linear maps, as was to be shown. □

Problem 27

Suppose $p \in \mathcal{P}(\mathbb{R})$. Prove that there exists a polynomial $q \in \mathcal{P}(\mathbb{R})$ such that $5q'' + 3q' = p$.

Proof. Suppose $\deg p = n$, and consider the linear map

$$\begin{aligned} D : \mathcal{P}_{n+1}(\mathbb{R}) &\to \mathcal{P}_n(\mathbb{R}) \\ q &\mapsto 5q'' + 3q'. \end{aligned}$$

If we can show D is surjective, we're done, since this implies that there exists some $q \in \mathcal{P}_{n+1}(\mathbb{R})$ such that $Dq = 5q'' + 3q' = p$. To that end, suppose $r \in \operatorname{null} D$. Then we must have $r'' = 0$ and $r' = 0$, which is true if and only if r is constant. Thus any $\alpha \in \mathbb{R}^\times$ is a basis of $\operatorname{null} D$, and so $\dim \operatorname{null} D = 1$. By the Fundamental Theorem of Linear Maps, we have

$$\dim \operatorname{range} D = \dim \mathcal{P}_{n+1}(\mathbb{R}) - \dim \operatorname{null} D,$$

and hence

$$\dim \operatorname{range} D = (n+2) - 1 = n + 1.$$

Since the only subspace of $\mathcal{P}_n(\mathbb{R})$ with dimension $n+1$ is the space itself, D is surjective, as desired. □

Problem 29

Suppose $\varphi \in \mathcal{L}(V, \mathbb{F})$. Suppose $u \in V$ is not in $\operatorname{null} \varphi$. Prove that

$$V = \operatorname{null} \varphi \oplus \{au \mid a \in \mathbb{F}\}.$$

Proof. First note that since $u \in V - \operatorname{null} \varphi$, there exists some nonzero $\varphi(u) \in \operatorname{range} \varphi$ and hence $\dim \operatorname{range} \varphi \geq 1$. But since $\operatorname{range} \varphi \subseteq \mathbb{F}$, and $\dim \mathbb{F} = 1$, we must have $\dim \operatorname{range} \varphi = 1$. Thus, letting $n = \dim V$, it follows

$$\begin{aligned} \dim \operatorname{null} \varphi &= \dim V - \dim \operatorname{range} \varphi \\ &= n - 1. \end{aligned}$$

Let $v_1, \ldots, v_{n-1}$ be a basis for $\operatorname{null} \varphi$. We claim $v_1, \ldots, v_{n-1}, u$ is an extension of this basis to a basis of V, which would then imply $V = \operatorname{null} \varphi \oplus \{au \mid a \in \mathbb{F}\}$, as desired.

To show $v_1, \ldots, v_{n-1}, u$ is a basis of V, it suffices to show linearly independence (since it has length $n = \dim V$). So suppose $a_1, \ldots, a_n \in \mathbb{F}$ are such that

$$a_1 v_1 + \cdots + a_{n-1} v_{n-1} + a_n u = 0.$$

We may write

$$a_n u = -a_1 v_1 - \cdots - a_{n-1} v_{n-1},$$

which implies $a_n u \in \operatorname{null} \varphi$. By hypothesis, $u \notin \operatorname{null} \varphi$, and thus we must have $a_n = 0$. But now each of the $a_1, \ldots, a_{n-1}$ must be 0 as well (since $v_1, \ldots, v_{n-1}$ form a basis of $\operatorname{null} \varphi$ and thus are linearly independent). Therefore, $v_1, \ldots, v_{n-1}, u$ is indeed linearly independent, proving our claim. □

Problem 31

Give an example of two linear maps T_1 and T_2 from $\mathbb{R}^5$ to $\mathbb{R}^2$ that have the same null space but are such that T_1 is not a scalar multiple of T_2.

Proof. Let $e_1, \dots, e_5$ be the standard basis of $\mathbb{R}^5$. We define $T_1, T_2 \in \mathcal{L}(\mathbb{R}^5, \mathbb{R}^2)$ by their behavior on the basis (using the standard basis for $\mathbb{R}^2$ as well)

$$\begin{aligned} T_1 e_1 &:= e_2 \\ T_1 e_2 &:= e_1 \\ T_1 e_k &:= 0 \text{ for } k = 3, 4, 5 \end{aligned}$$

and

$$\begin{aligned} T_2 e_1 &:= e_1 \\ T_2 e_2 &:= e_2 \\ T_2 e_k &:= 0 \text{ for } k = 3, 4, 5. \end{aligned}$$

Clearly $\operatorname{null} T_1 = \operatorname{null} T_2$. We claim T_2 is not a scalar multiple of T_1. To see this, suppose not. Then there exists $\alpha \in \mathbb{R}$ such that $T_1 = \alpha T_2$. In particular, this implies $T_1 e_1 = \alpha T_2 e_1$. But this is absurd, since $T_1 e_1 = e_2$ and $T_2 e_1 = e_1$, and of course e_1, e_2 is linearly independent. Thus no such α can exist, and T_1, T_2 are as desired. $\square$

C: Matrices

Problem 1

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that with respect to each choice of bases of V and W, the matrix of T has at least $\dim \operatorname{range} T$ nonzero entries.

Proof. Let $v_1, \dots, v_n$ be a basis of V, let $w_1, \dots, w_m$ be a basis of W, let $r = \dim \operatorname{range} T$, and let $s = \dim \operatorname{null} T$. Then there are s basis vectors of V which map to zero and r basis vectors of V with nontrivial representation as linear combinations of $w_1, \dots, w_m$. That is, suppose $T v_k \neq 0$, where $k \in \{1, \dots, n\}$. Then there exist $a_1, \dots, a_m \in \mathbb{F}$, not all zero, such that

$$T v_k = a_1 w_1 + \dots + a_m w_m.$$

The coefficients form column k of $\mathcal{M}(T)$, and there are r such vectors in the basis of V. Hence there are r columns of $\mathcal{M}(T)$ with at least one nonzero entry, as was to be shown. $\square$

Problem 3

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ are 0 except the entries in row j, column j, equal 1 for $1 \leq j \leq \dim \operatorname{range} T$.

Proof. Let R be the subspace of V such that

$$V = R \oplus \operatorname{null} T,$$

let $r_1, \dots, r_m$ be a basis of R (where $m = \dim \operatorname{range} T$), and let $v_1, \dots, v_n$ be a basis of $\operatorname{null} T$ (where $n = \dim \operatorname{null} T$). Then $r_1, \dots, r_m, v_1, \dots, v_n$ is a basis of V. It follows that $Tr_1, \dots, Tr_m$ is a basis of $\operatorname{range} T$, and hence there is an extension of this list to a basis of W. Suppose $Tr_1, \dots, Tr_m, w_1, \dots, w_p$ is such an extension (where $p = \dim W - m$). Then, for $j = 1, \dots m$, we have

$$Tr_j = \left(\sum_{i=1}^{m} \delta_{i,j} \cdot Tr_t \right) + \left(\sum_{k=1}^{p} 0 \cdot w_k \right),$$

where $\delta_{i,j}$ is the Kronecker delta function. Thus, column j of $\mathcal{M}(T)$ is has an entry of 1 in row j and 0's elsewhere, where j ranges over 1 to $m = \dim \operatorname{range} T$. Since $Tv_1 = \cdots = Tv_n = 0$, the remaining columns of $\mathcal{M}(T)$ are all zero. Thus $\mathcal{M}(T)$ has the desired form. $\square$

Problem 5

Suppose $w_1, \dots, w_n$ is a basis of W and V is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $v_1, \dots, v_m$ of V such that all the entries in the first row of $\mathcal{M}(T)$ (with respect to the bases $v_1, \dots, v_m$ and $w_1, \dots, w_n$) are 0 except for possibly a 1 in the first row, first column.

Proof. First note that if $\operatorname{range} T \subseteq \operatorname{span}(w_2, \dots, w_n)$, the first row of $\mathcal{M}(T)$ will be all zeros regardless of choice of basis for V.

So suppose $\operatorname{range} T \not\subseteq \operatorname{span}(w_2, \dots, w_n)$ and let $u_1 \in V$ be such that $Tu_1 \notin \operatorname{span}(w_2, \dots, w_n)$. There exist $a_1, \dots, a_n \in \mathbb{F}$ such that

$$Tu_1 = a_1 w_1 + \cdots + a_n w_n,$$

and notice $a_1 \neq 0$ since $Tu_1 \notin \operatorname{span}(w_2, \dots, w_n)$. Hence we may define

$$z_1 := \frac{1}{a_1} u_1.$$

It follows

$$Tz_1 = w_1 + \frac{a_2}{a_1} w_2 + \cdots + \frac{a_n}{a_1} w_n. \tag{3}$$

Now extend z_1 to a basis $z_1, \ldots, z_m$ of V. Then for $k = 2, \ldots, m$, there exist $A_{1,k}, \ldots, A_{n,k} \in \mathbb{F}$ such that

$$Tz_k = A_{1,k}w_1 + \cdots + A_{n,k}w_n,$$

and notice

$$\begin{aligned} T(z_k - A_{1,k}z_1) &= Tz_k - A_{1,k}Tz_1 \\ &= \big(A_{1,k}w_1 + \cdots + A_{n,k}w_n\big) - A_{1,k}\left(w_1 + \frac{a_2}{a_1}w_2 + \cdots + \frac{a_n}{a_1}w_n\right) \\ &= \big(A_{2,k} - A_{1,k}\big)\frac{a_2}{a_1}w_2 + \cdots + \big(A_{n,k} - A_{1,k}\big)\frac{a_n}{a_1}w_n. \qquad (4) \end{aligned}$$

Now we define a new list in V by

$$v_k := \begin{cases} z_1 & \text{if } k = 1 \\ z_k - A_{1,k}z_1 & \text{otherwise} \end{cases}$$

for $k = 1, \ldots, m$. We claim $v_1, \ldots, v_m$ is a basis. To see this, it suffices to prove the list is linearly independent, since its length equals $\dim V$. So suppose $b_1, \ldots, b_m \in \mathbb{F}$ are such that

$$b_1v_1 + \cdots + b_mv_m = 0.$$

By definition of the v_k, it follows

$$b_1z_1 + b_2(z_2 - A_{1,k}z_1) + \cdots + b_m(z_m - A_{1,k}z_1) = 0.$$

But since $z_1, \ldots, z_m$ is a basis of V, the expression on the LHS above is simply a linear combination of vectors in a basis. Thus we must have $b_1 = \cdots = b_m = 0$, and indeed $v_1, \ldots, v_m$ are linearly independent, as claimed.

Finally, notice (3) tells us the first column of $\mathcal{M}(T, v_k, w_k)$ is all 0's except a 1 in the first entry, and (4) tells us the remaining columns have a 0 in the first entry. Thus $\mathcal{M}(T, v_k, w_k)$ has the desired form, completing the proof. ☐

Problem 7

Suppose $S, T \in \mathcal{L}(V, W)$. Prove that $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.

Proof. Let $v_1, \ldots, v_m$ be a basis of V and let $w_1, \ldots, w_n$ be a basis of W. Also, let $A = \mathcal{M}(S)$ and $B = \mathcal{M}(T)$ be the matrices of these linear transformations with respect to these bases. It follows

$$\begin{aligned} (S + T)v_k &= Sv_k + Tv_k \\ &= \big(A_{1,k}w_1 + \cdots + A_{n,k}w_n\big) + \big(B_{1,k}w_1 + \cdots + B_{n,k}w_n\big) \\ &= (A_{1,k} + B_{1,k})w_1 + \cdots + (A_{n,k} + B_{n,k})w_n. \end{aligned}$$

Hence $\mathcal{M}(S+T)_{j,k} = A_{j,k}+B_{j,k}$, and indeed we have $\mathcal{M}(S+T) = \mathcal{M}(S)+\mathcal{M}(T)$, as desired. ☐

Problem 9

Suppose A is an m-by-n matrix and $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ is an n-by-1 matrix.

Prove that
$$Ac = c_1 A_{\cdot,1} + \cdots + c_n A_{\cdot,n}.$$

Proof. By definition, it follows

$$\begin{aligned}
Ac &= \begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ A_{2,1} & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \dots & A_{m,n} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \\
&= \begin{pmatrix} A_{1,1}c_1 + A_{1,2}c_2 + \cdots + A_{1,n}c_n \\ A_{2,1}c_1 + A_{2,2}c_2 + \cdots + A_{2,n}c_n \\ \vdots \\ A_{m,1}c_1 + A_{m,2}c_2 + \cdots + A_{m,n}c_n \end{pmatrix} \\
&= c_1 \begin{pmatrix} A_{1,1} \\ A_{2,1} \\ \vdots \\ A_{m,1} \end{pmatrix} + c_2 \begin{pmatrix} A_{1,2} \\ A_{2,2} \\ \vdots \\ A_{m,2} \end{pmatrix} + \cdots + c_n \begin{pmatrix} A_{1,n} \\ A_{2,n} \\ \vdots \\ A_{m,n} \end{pmatrix} \\
&= c_1 A_{\cdot,1} + \cdots + c_n A_{\cdot,n},
\end{aligned}$$

as desired. □

Problem 11

Suppose $a = (a_1, \dots, a_n)$ is a 1-by-n matrix and C is an n-by-p matrix. Prove that
$$aC = a_1 C_{1,\cdot} + \cdots + a_n C_{n,\cdot}.$$

Proof. By definition, it follows

$$\begin{aligned}
aC &= (a_1, \ldots, a_n) \begin{pmatrix} C_{1,1} & C_{1,2} & \ldots & C_{1,p} \\ C_{2,1} & C_{2,2} & \ldots & C_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n,1} & C_{n,2} & \ldots & C_{n,p} \end{pmatrix} \\
&= \left(\sum_{k=1}^{n} a_k C_{k,1}, \sum_{k=1}^{n} a_k C_{k,2}, \ldots, \sum_{k=1}^{n} a_k C_{k,p} \right) \\
&= \sum_{k=1}^{n} \big(a_k C_{k,1}, \ldots, a_k C_{k,p} \big) \\
&= \sum_{k=1}^{n} a_k \big(C_{k,1}, \ldots, C_{k,p} \big) \\
&= \sum_{k=1}^{n} a_k C_{k,\cdot},
\end{aligned}$$

as desired. ☐

Problem 13

Prove that the distributive property holds for matrix addition and matrix multiplication. In other words, suppose A, B, C, D, E, and F are matrices whose sizes are such that $A(B+C)$ and $(D+E)F$ make sense. Prove that $AB+AC$ and $DF+EF$ both make sense and that $A(B+C) = AB+AC$ and $(D+E)F = DF+EF$.

Proof. First note that if $A(B+C)$ makes sense, then the number of columns of A must equal the number of rows of $B+C$. But the sum of two matrices is only defined if their dimensions are equal, and hence the number of rows of both B and C must equal the number of columns of A. Thus $AB+AC$ makes sense. So suppose $A \in \mathbb{F}^{m,n}$ and $B, C \in \mathbb{F}^{n,p}$. It follows

$$\begin{aligned}
\big(A(B+C)\big)_{j,k} &= \sum_{r=1}^{n} A_{j,r}(B+C)_{r,k} \\
&= \sum_{r=1}^{n} A_{j,r}(B_{r,k} + C_{r,k}) \\
&= \sum_{r=1}^{n} \big(A_{j,r} B_{r,k} + A_{j,r} C_{r,k} \big) \\
&= \sum_{r=1}^{n} A_{j,r} B_{r,k} + \sum_{r=1}^{n} A_{j,r} C_{r,k} \\
&= (AB)_{j,k} + (AC)_{j,k},
\end{aligned}$$

proving the first distributive property.

Now note that if $(D+E)F$ makes sense, then the number of columns of $D+E$ must equal the number of rows of F. Hence the number of columns of both D and E must equal the number of rows of F, and thus $DF+EF$ makes sense as well. So suppose $D, E \in \mathbb{F}^{m,n}$ and $F \in \mathbb{F}^{n,p}$. It follows

$$\begin{aligned}\big((D+E)F\big)_{j,k} &= \sum_{r=1}^{n}(D+E)_{j,r}F_{r,k} \\ &= \sum_{r=1}^{n}(D_{j,r}+E_{j,r})F_{r,k} \\ &= \sum_{r=1}^{n} D_{j,r}F_{r,k}+E_{j,r}F_{r,k} \\ &= \sum_{r=1}^{n} D_{j,r}F_{r,k}+\sum_{r=1}^{n} E_{j,r}F_{r,k} \\ &= (DF)_{j,k}+(EF)_{j,k},\end{aligned}$$

proving the second distributive property. □

Problem 15

Suppose A is an n-by-n matrix and $1 \leq j, k \leq n$. show that the entry in row j, column k, of A^3 (which is defined to mean AAA) is

$$\sum_{p=1}^{n}\sum_{r=1}^{n} A_{j,p}A_{p,r}A_{r,k}.$$

Proof. For $1 \leq p, k \leq n$, we have

$$(A^2)_{p,k} = \sum_{r=1}^{n} A_{p,r}A_{r,k}.$$

Thus, for $1 \leq j, k \leq n$, it follows

$$\begin{aligned}(A^3)_{j,k} &= \sum_{p=1}^{n} A_{j,p}(A^2)_{p,k} \\ &= \sum_{p=1}^{n} A_{j,p}\sum_{r=1}^{n} A_{p,r}A_{r,k} \\ &= \sum_{p=1}^{n}\sum_{r=1}^{n} A_{j,p}A_{p,r}A_{r,k},\end{aligned}$$

as desired. □

D: Invertibility and Isomorphic Vector Spaces

Problem 1

Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$ are both invertible linear maps. Prove that $ST \in \mathcal{L}(U, W)$ is invertible and that $(ST)^{-1} = T^{-1}S^{-1}$.

Proof. For all $u \in U$, we have

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

and hence $T^{-1}S^{-1}$ is a left inverse of ST. Similarly, for all $w \in W$, we have

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

and hence $T^{-1}S^{-1}$ is a right inverse of ST. Therefore, ST is invertible, as desired. □

Problem 3

Suppose V is finite-dimensional, U is a subspace of V, and $S \in \mathcal{L}(U, V)$. Prove there exists an invertible operator $T \in \mathcal{L}(V)$ such that $Tu = Su$ for every $u \in U$ if and only if S is injective.

Proof. ($\Leftarrow$) Suppose S is injective, and let W be the subspace of V such that $V = U \oplus W$. Let $u_1, \dots, u_m$ be a basis of U and let $w_1, \dots, w_n$ be a basis of W, so that $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis of V. Define $T \in \mathcal{L}(V)$ by its behavior on this basis of V

$$\begin{aligned}Tu_k &:= Su_k \\ Tw_j &:= w_j\end{aligned}$$

for $k = 1, \dots, m$ and $j = 1, \dots, n$. Since S is injective, so too is T. And since V is finite-dimensional, this implies that T is invertible, as desired.

($\Rightarrow$) Suppose there exists an invertible operator $T \in \mathcal{L}(V)$ such that $Tu = Su$ for every $u \in U$. Since T is invertible, it is also injective. And since T is injective, so to is $S = T\,|_U$, completing the proof. □

Problem 5

Suppose V is finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} T_1 = \operatorname{range} T_2$ if and only if there exists an invertible operator $S \in \mathcal{L}(V)$ such that $T_1 = T_2 S$.

Proof. ($\Rightarrow$) Suppose $\operatorname{range} T_1 = \operatorname{range} T_2 := R$, so that $\operatorname{null} T_1 = \operatorname{null} T_2 := N$ as well. Let Q be the unique subspace of V such that

$$V = N \oplus Q,$$

and let $u_1, \ldots, u_m$ be a basis of N and $v_1, \ldots, v_n$ be a basis of Q. We claim there exists a unique $q_k \in Q$ such that $T_2 q_k = T_1 v_k$ for $k = 1, \ldots, n$. To see this, suppose $q_k, q_k' \in Q$ are such that $T_2 q_k = T_2 q_k' = T_1 v_k$. Then $T_2(q_k - q_k') = 0$, and hence $q_k - q_k' \in N$. But since $N \cap Q = \{0\}$, this implies $q_k - q_k' = 0$ and thus $q_k = q_k'$. And so the choice of q_k is indeed unique. We now define $S \in \mathcal{L}(V)$ by its behavior on the basis

$$\begin{aligned} Su_k &= u_k \text{ for } k = 1, \ldots, m \\ Sv_j &= q_j \text{ for } j = 1, \ldots, n. \end{aligned}$$

Let $v \in V$, so that there exist $a_1, \ldots, a_m, b_1, \ldots, b_n \in \mathbb{F}$ such that

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

It follows

$$\begin{aligned} (T_2 S)(v) &= T_2(S(a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n)) \\ &= T_2(a_1 S u_1 + \cdots + a_m S u_m + b_1 S v_1 + \cdots + b_n S v_n) \\ &= T_2(a_1 u_1 + \cdots + a_m u_m + b_1 q_1 + \cdots + b_n q_n) \\ &= a_1 T_2 u_1 + \cdots + a_m T_2 u_m + b_1 T_2 q_1 + \cdots + b_n T_2 q_n \\ &= b_1 T_1 v_1 + \cdots + b_n T_1 v_n. \end{aligned}$$

Similarly, we have

$$\begin{aligned} T_1 v &= T_1(a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n) \\ &= a_1 T_1 u_1 + \cdots + a_m T_1 u_m + b_1 T_1 v_1 + \cdots + b_n T_1 v_n \\ &= b_1 T_1 v_1 + \cdots + b_n T_1 v_n, \end{aligned}$$

and so indeed $T_1 = T_2 S$. To see that S is invertible, it suffices to prove it is injective. So let $v \in V$ be as before, and suppose $Sv = 0$. It follows

$$\begin{aligned} Sv &= S(a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n) \\ &= (a_1 u_1 + \cdots + a_m u_m) + (b_1 S v_1 + \cdots + b_n S v_n) \\ &= 0. \end{aligned}$$

By the proof of Theorem 3.22, $Sv_1, \dots, Sv_n$ is a basis of R, and thus the list $u_1, \dots, u_m, Sv_1, \dots, Sv_n$ is a basis of V, and each of the a's and b's must be zero. Therefore S is indeed injective, completing the proof in this direction.

($\Leftarrow$) Suppose there exists an invertible operator $S \in \mathcal{L}(V)$ such that $T_1 = T_2 S$. If $w \in \operatorname{range} T_1$, then there exists $v \in V$ such that $T_1 v = w$, and hence $(T_2 S)(v) = T_2(S(v)) = w$, so that $w \in \operatorname{range} T_2$ and we have $\operatorname{range} T_1 \subseteq \operatorname{range} T_2$. Conversely, suppose $w' \in \operatorname{range} T_2$, so that there exists $v' \in V$ such that $T_2 v' = w'$. Then, since $T_2 = T_1 S^{-1}$, we have $(T_1 S^{-1})(v') = T_1(S^{-1}(v')) = w'$, so that $w' \in \operatorname{range} T_1$. Thus $\operatorname{range} T_2 \subseteq \operatorname{range} T_1$, and we have shown $\operatorname{range} T_1 = \operatorname{range} T_2$, as desired. □

Problem 7

Suppose V and W are finite-dimensional. Let $v \in V$. Let

$$E = \{T \in \mathcal{L}(V, W) \mid Tv = 0\}.$$

(a) Show that E is a subspace of $\mathcal{L}(V, W)$.

(b) Suppose $v \neq 0$. What is $\dim E$?

Proof. (a) First note that the zero map is clearly an element of E, and hence E contains the additive identity of $\mathcal{L}(V, W)$. Now suppose $T_1, T_2 \in E$. Then

$$(T_1 + T_2)(v) = T_1 v + T_2 v = 0$$

and hence $T_1 + T_2 \in E$, so that E is closed under addition. Finally, suppose $T \in E$ and $\lambda \in \mathbb{F}$. Then

$$(\lambda T)(v) = \lambda T v = \lambda 0 = 0,$$

and so E is closed under scalar multiplication as well. Thus E is indeed a subspace of $\mathcal{L}(V, W)$.

(b) Suppose $v \neq 0$, and let $\dim V = m$ and $\dim W = n$. Extend v to a basis $v, v_2, \dots, v_m$ of V, and endow W with any basis. Let $\mathcal{E}$ denote the subspace of $\mathbb{F}^{m,n}$ of matrices whose first column is all zero.

We claim $T \in E$ if and only if $\mathcal{M}(T) \in \mathcal{E}$, so that $\mathcal{M} : E \to \mathcal{E}$ is an isomorphism. Clearly if $T \in E$ (so that $Tv = 0$), then $\mathcal{M}(T)_{\cdot,1}$ is all zero, and hence $T \in \mathcal{E}$. Conversely, suppose $\mathcal{M}(T) \in \mathcal{E}$. It follows

$$\begin{aligned}
\mathcal{M}(Tv) &= \mathcal{M}(T)\mathcal{M}(v) \\
&= \begin{pmatrix} 0 & A_{1,2} & \dots & A_{1,n} \\ 0 & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & A_{m,2} & \dots & A_{m,n} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},
\end{aligned}$$

and thus we must have $Tv = 0$ so that $T \in E$, proving our claim. So indeed $E \cong \mathcal{E}$.

Now note that $\mathcal{E}$ has as a basis the set of all matrices with a single 1 in a column besides the first, and zeros everywhere else. There are $mn - n$ such matrices, and hence $\dim \mathcal{E} = mn - n$. Thus we have $\dim E = mn - n$ as well, as desired. ☐

Problem 9

Suppose 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.

Proof. ($\Leftarrow$) Suppose S and T are both invertible. Then by Problem 1, ST is invertible.

($\Rightarrow$) Suppose ST is invertible. We will show T is injective and S is surjective. Since V is finite-dimensional, this implies that both S and T are invertible. So suppose $v_1, v_2 \in V$ are such that $Tv_1 = Tv_2$. Then $(ST)(v_1) = (ST)(v_2)$, and since ST is invertible (and hence injective), we must have $v_1 = v_2$, so that T is injective. Next, suppose $v \in V$. Since T^{-1} is surjective, there exists $w \in V$ such that $T^{-1}w = v$. And since ST is surjective, there exists $p \in V$ such that $(ST)(p) = w$. It follows that $(STT^{-1})(p) = T^{-1}(w)$, and hence $Sp = v$. Thus S is surjective, completing the proof. ☐

Problem 11

Suppose V is finite-dimensional and $S, T, U \in \mathcal{L}(V)$ and $STU = I$. Show that T is invertible and that $T^{-1} = US$.

Proof. Notice STU is invertible since $STU = I$ and I is invertible. By Problem 9, we have that $(ST)U$ is invertible if and only if ST and U are invertible. By a second application of the result, ST is invertible if and only if S and T are invertible. Thus S, T, and U are all invertible. To see that $T^{-1} = US$, notice

$$\begin{aligned} US &= (T^{-1}T)US \\ &= T^{-1}(S^{-1}S)TUS \\ &= T^{-1}S^{-1}(STU)S \\ &= T^{-1}S^{-1}S \\ &= T^{-1}, \end{aligned}$$

as desired. ☐

Problem 13

Suppose V is a finite-dimensional vector space and $R, S, T \in \mathcal{L}(V)$ are such that RST is surjective. Prove that S is injective.

Proof. Since V is finite-dimensional and RST is surjective, RST is also invertible. By Problem 9, we have that $(RS)T$ is invertible if and only if RS and T are invertible. By a second application of the result, RS is invertible if and only if R and S are invertible. Thus R, S, and T are all invertible, and hence injective. In particular, S is injective, as desired. ☐

Problem 15

Prove that every linear map from $\mathbb{F}^{n,1}$ to $\mathbb{F}^{m,1}$ is given by a matrix multiplication. In other words, prove that if $T \in \mathcal{L}(\mathbb{F}^{n,1}, \mathbb{F}^{m,1})$, then there exists an m-by-n matrix A such that $Tx = Ax$ for every $x \in \mathbb{F}^{n,1}$.

Proof. Endow both $\mathbb{F}^{n,1}$ and $\mathbb{F}^{m,1}$ with the standard basis, and let $T \in \mathcal{L}(\mathbb{F}^{n,1}, \mathbb{F}^{m,1})$. Let $A = \mathcal{M}(T)$ with respect to these bases, and note that if $x \in \mathbb{F}^{n,1}$, then $\mathcal{M}(x) = x$ (and similarly if $y \in \mathbb{F}^{m,1}$, then $\mathcal{M}(y) = y$). Hence

$$\begin{aligned} Tx &= \mathcal{M}(Tx) \\ &= \mathcal{M}(T)\mathcal{M}(x) \\ &= Ax, \end{aligned}$$

as desired. ☐

Problem 16

Suppose 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 $S \in \mathcal{L}(V)$.

Proof. ($\Rightarrow$) Suppose $T = \lambda I$ for some $\lambda \in \mathbb{F}$, and let $S \in \mathcal{L}(V)$ be arbitrary. For any $v \in V$, we have $STv = S(\lambda I)v = \lambda Sv$ and $TSv = (\lambda I)Sv = \lambda Sv$, and hence $ST = TS$. Since S was arbitrary, we have the desired result.

($\Leftarrow$) Suppose $ST = TS$ for every $S \in \mathcal{L}(V)$, and let $v \in V$ be arbitrary. Consider the list v, Tv. We claim it is linearly dependent. To see this, suppose not. 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

$$S(\alpha v + \beta Tv + \gamma_1 u_1 + \cdots + \gamma_n u_n) = \beta v,$$

where $\alpha, \beta, \gamma_1, \ldots, \gamma_n$ are the unique coefficients of our basis for the given input vector. In particular, notice $S(Tv) = v$ and $Sv = 0$. Thus $STv = TSv$ implies $v = T(0) = 0$, contradicting our assumption that v, TV is linearly independent. So v, Tv must be linearly dependent, and so for for all $v \in V$ there exists $\lambda_v \in \mathbb{F}$ such that $Tv = \lambda_v v$ (where λ_0 can be any nonzero element of $\mathbb{F}$, since $T0 = 0$). We claim λ_v is independent of the choice of v for $v \in V - \{0\}$, hence $Tv = \lambda v$ for all $v \in V$ (including $v = 0$) and some $\lambda \in \mathbb{F}$, and thus $T = \lambda I$.

So suppose $w, z \in V - \{0\}$ are arbitrary. We want to show $\lambda_w = \lambda_z$. If w and z are linearly dependent, then there exists $\alpha \in \mathbb{F}$ such that $w = \alpha z$. It follows

$$\begin{aligned}
\lambda_w w &= Tw \\
&= T(\alpha z) \\
&= \alpha T z \\
&= \alpha \lambda_z z \\
&= \lambda_z (\alpha z) \\
&= \lambda_z w.
\end{aligned}$$

Since $w \neq 0$, this implies $\lambda_w = \lambda_z$. Next suppose w and z are linearly independent. Then we have

$$\begin{aligned}
\lambda_{w+z}(w + z) &= T(w + z) \\
&= Tw + Tz \\
&= \lambda_w w + \lambda_z z,
\end{aligned}$$

and hence

$$(\lambda_{w+z} - \lambda_w)w + (\lambda_{w+z} - \lambda_z)z = 0.$$

Since w and z are assumed to be linearly independent, we have $\lambda_{w+z} = \lambda_w$ and $\lambda_{w+z} = \lambda_z$, and hence again we have $\lambda_w = \lambda_z$, completing the proof. $\square$

Problem 17

Suppose V is finite-dimensional and $\mathcal{E}$ is a subspace of $\mathcal{L}(V)$ such that $ST \in \mathcal{E}$ and $TS \in \mathcal{E}$ for all $S \in \mathcal{L}(V)$ and all $T \in \mathcal{E}$. Prove that $\mathcal{E} = \{0\}$ or $\mathcal{E} = \mathcal{L}(V)$.

Proof. If $\mathcal{E} = \{0\}$, we're done. So suppose $\mathcal{E} \neq \{0\}$. If $\dim V = n$, then $\mathcal{L}(V) \cong \mathbb{F}^{n,n}$, and so there exists an isomorphic subspace $\mathfrak{E} := \mathcal{M}(\mathcal{E}) \subseteq \mathbb{F}^{n,n}$ with the property that $AB \in \mathfrak{E}$ and $BA \in \mathfrak{E}$ for all $A \in \mathbb{F}^{n,n}$ and all $B \in \mathfrak{E}$. It suffices to show $\mathfrak{E} = \mathbb{F}^{n,n}$.

Define $E^{i,j}$ to be the matrix which is 1 in row i and column j and 0 everywhere else, and let $A \in \mathbb{F}^{n,n}$ be nonzero. Then there exists some $1 \leq j, k \leq n$ such that $A_{j,k} \neq 0$. Notice for $1 \leq i, j, r, s \leq n$, we have $E^{i,j}A \in \mathfrak{E}$, and hence $E^{i,j}AE^{r,s} \in \mathfrak{E}$. This product has the form

$$E^{i,j}AE^{k,\ell} = A_{j,k} \cdot E^{i,\ell}.$$

In other words, $E^{i,j}AE^{k,\ell}$ takes $A_{j,k}$ and puts it in the i^{th} row and ℓ^{th} column of a matrix which is 0 everywhere else. Since $\mathfrak{E}$ is closed under addition, this implies

$$E^{1,j}AE^{k,1} + E^{2,j}AE^{k,2} + \cdots + E^{n,j}AE^{k,n} = A_{j,k} \cdot I \in \mathfrak{E}.$$

But since $\mathfrak{E}$ is closed under scalar multiplication, and $A_{j,k} \neq 0$, we have

$$\left(\frac{1}{A_{j,k}} \cdot A_{j,k} \right) \cdot I = I \in \mathfrak{E}.$$

Since $\mathfrak{E}$ contains I, by our characterization of $\mathfrak{E}$ it must also contain every element of $\mathbb{F}^{n,n}$. Thus $\mathfrak{E} = \mathbb{F}^{n,n}$, and since $\mathfrak{E} \cong \mathcal{E}$, we must have $\mathcal{E} = \mathcal{L}(V)$, as desired. □

Problem 19

Suppose $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ is such that T is injective and $\deg Tp \leq \deg p$ for every nonzero polynomial $p \in \mathcal{P}(\mathbb{R})$.

(a) Prove that T is surjective.

(b) Prove that $\deg Tp = \deg p$ for every nonzero $p \in \mathcal{P}(\mathbb{R})$.

Proof. (a) Let $q \in \mathcal{P}(\mathbb{R})$, and suppose $\deg q = n$. Let $T_n = T\,|_{\mathcal{P}_n(\mathbb{R})}$, so that T_n is the restriction of T to a linear operator on $\mathcal{P}_n(\mathbb{R})$. Since T is injective, so is T_n. And since T_n is an injective linear operator over a finite-dimensional vector space, T_n is surjective as well. Thus there exists $r \in \mathcal{P}_n(\mathbb{R})$ such that $T_n r = q$, and so we have $Tr = q$ as well. Therefore T is surjective.

(b) We induct on the degree of the restriction maps $T_n \in \mathcal{L}(\mathcal{P}_n(\mathbb{R}))$, each of which is bijective by (a). Let $P(k)$ be the statement: $\deg T_k p = k$ for every nonzero $p \in \mathcal{P}_k(\mathbb{R})$.
Base case: Suppose $p \in \mathcal{P}_0(\mathbb{R})$ is nonzero. Since T_0 is a bijective , $T_0 p = 0$ iff $p = 0$ (the zero polynomial), which is the only polynomial with degree

< 0. Since p is nonzero by hypothesis, we must have $\deg T_0 p = 0$. Hence $P(0)$ is true.
Inductive step: Let $n \in \mathbb{Z}^+$, and suppose $P(k)$ is true for all $0 \leq k < n$. Let $p \in \mathcal{P}_n(\mathbb{R})$ be nonzero. If $\deg T_n p < n$, then for some $k < n$ there exists $q \in \mathcal{P}_k(R)$ and $T_k \in \mathcal{P}(\mathbb{R})$ such that $T_k q = p$ (since T_k is surjective). Hence $Tq = Tp$, a contradiction since $\deg p \neq \deg q$ and T is injective. Thus we must have $\deg T_n p = n$, and $P(n)$ is true.

By the principle of mathematical induction, $P(k)$ is true for all $k \in \mathbb{Z}_{\geq 0}$. Hence $\deg Tp = \deg p$ for all nonzero $p \in \mathcal{P}(\mathbb{R})$, since $Tp = T_k p$ for $k = \deg p$. $\square$

E: Products and Quotients of Vector Spaces

Problem 1

Suppose T is a function from V to W. The **graph** of T is the subset of $V \times W$ defined by

$$\text{graph of } T = \{(v, Tv) \in V \times W \mid v \in V\}.$$

Prove that T is a linear map if and only if the graph of T is a subspace of $V \times W$.

Proof. Define $G := \{(v, Tv) \in V \times W \mid v \in V\}$.

($\Rightarrow$) Suppose T is a linear map. Since T is linear, $T0 = 0$, and hence $(0, 0) \in G$, so that G contains the additive identity. Next, let $(v_1, Tv_1), (v_2, Tv_2) \in G$. Then

$$(v_1, Tv_1) + (v_2, Tv_2) = (v_1 + v_2, Tv_1 + Tv_2) = (v_1 + v_2, T(v_1 + v_2)) \in G,$$

and G is closed under addition. Lastly, let $\lambda \in \mathbb{F}$ and $(v, Tv) \in G$. It follows

$$\lambda(v, Tv) = (\lambda v, \lambda Tv) = (\lambda v, T(\lambda v)) \in G,$$

and G is closed under scalar multiplication. Thus G is a subspace of $V \times W$.

($\Leftarrow$) Suppose G is a subspace of $V \times W$, and let $(v_1, Tv_1), (v_2, Tv_2) \in G$. Since G is closed under addition, it follows

$$(v_1 + v_2, Tv_1 + Tv_2) \in G,$$

and hence we must have $Tv_1 + Tv_2 = T(v_1 + v_2)$, so that T is additive. And since G is closed under scalar multiplication, for $\lambda \in \mathbb{F}$ and $(v, Tv) \in G$, it follows

$$(\lambda v, \lambda Tv) \in G,$$

and hence we must have $\lambda Tv = T(\lambda v)$, so that T is homogeneous. Therefore, T is a linear map, as desired. $\square$

Problem 3

Give an example of a vector space V and subspaces U_1, U_2 of V such that $U_1 \times U_2$ is isomorphic to $U_1 + U_2$ but $U_1 + U_2$ is not a direct sum.

Proof. Define the following two subspaces of $\mathcal{P}(\mathbb{R})$

$$U_1 := \mathcal{P}(\mathbb{R})$$
$$U_2 := \mathbb{R},$$

so that $U_1 \cap U_2 = \mathbb{R}$ and the sum $U_1 + U_2 = \mathcal{P}(\mathbb{R})$ is not direct. Endow $\mathcal{P}(\mathbb{R})$ and $\mathbb{R}$ with their standard bases, and define φ by its behavior on the basis of $U_1 \times U_2$

$$\begin{aligned} \varphi : U_1 \times U_2 &\to U_1 + U_2 \\ \left(X^k, 0\right) &\mapsto X^{k+1} \\ (0, 1) &\mapsto 1. \end{aligned}$$

We claim φ is an isomorphism. To see that φ is injective, suppose

$$(a_0 + a_1 X + \cdots + a_m X^m, \alpha), (b_0 + b_1 X + \cdots + b_n X^n, \beta) \in U_1 \times U_2$$

and

$$(a_0 + a_1 X + \cdots + a_m X^m, \alpha) \neq (b_0 + b_1 X + \cdots + b_n X^n, \beta).$$

We have

$$\varphi (a_0 + a_1 X + \cdots + a_m X^m, \alpha) = \alpha + a_0 X + a_1 X^2 + \cdots + a_m X^{m+1} \tag{5}$$

and

$$\varphi (b_0 + b_1 X + \cdots + b_n X^n, \beta) = \beta + b_0 X + b_1 X^2 + \cdots + b_n X^{n+1}. \tag{6}$$

Since $\alpha \neq \beta$, this implies (5) does not equal (6) and hence φ is injective. To see that φ is surjective, suppose $c_0 + c_1 X + \cdots + c_p X^p \in U_1 + U_2$. Then

$$\varphi\left(c_1 + c_2 X + \cdots + c_p X^{p-1}, c_0\right) = c_0 + c_1 X + \cdots + c_p X^p$$

and φ is indeed surjective.

Since φ an injective and surjective linear map, it is an isomorphism. Thus $U_1 \times U_2 \cong U_1 + U_2$, as was to be shown. $\square$

Problem 5

Suppose $W_1, \ldots, W_m$ are vector spaces. Prove that $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ and $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ are isomorphic vector spaces.

Proof. Define the projection map π_k for $k = 1, \ldots, m$ by

$$\begin{aligned} \pi_k : W_1 \times \cdots \times W_m &\to W_k \\ (w_1, \ldots, w_m) &\mapsto w_k. \end{aligned}$$

Clearly π_k is linear. Now define

$$\begin{aligned} \varphi : \mathcal{L}(V, W_1 \times \cdots \times W_m) &\to \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m) \\ T &\mapsto (\pi_1 T, \ldots, \pi_m T). \end{aligned}$$

To see that φ is linear, let $T_1, T_2 \in \mathcal{L}(V, W_1 \times \cdots \times W_m)$. It follows

$$\begin{aligned} \varphi(T_1 + T_2) &= (\pi_1(T_1 + T_2), \ldots, \pi_m(T_1 + T_2)) \\ &= (\pi_1 T_1 + \pi_1 T_2, \ldots, \pi_m T_1 + \pi_m T_2) \\ &= (\pi_1 T_1, \ldots, \pi_m T_1) + (\pi_1 T_2, \ldots, \pi_m T_2) \\ &= \varphi(T_1) + \varphi(T_2), \end{aligned}$$

and hence φ is additive. Now for $\lambda \in \mathbb{F}$ and $T \in \mathcal{L}(V, W_1 \times \cdots \times W_m)$, we have

$$\begin{aligned} \varphi(\lambda T) &= (\pi_1(\lambda T), \ldots, \pi_m(\lambda T)) \\ &= (\lambda(\pi_1 T), \ldots, \lambda(\pi_m T)) \\ &= \lambda(\pi_1 T, \ldots, \pi_m T), \end{aligned}$$

and thus φ is homogenous. Therefore, φ is linear.

We now show φ is an isomorphism. To see that it is injective, suppose $T \in \mathcal{L}(V, W_1 \times \cdots \times W_m)$ and $\varphi(T) = 0$. Then

$$(\pi_1 T, \ldots, \pi_m T) = (0, \ldots, 0)$$

which is true iff T is the zero map. Thus φ is injective. To see that φ is surjective, suppose $(S_1, \ldots, S_m) \in \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$. Define

$$\begin{aligned} S : V &\to W_1 \times \cdots \times W_m \\ v &\mapsto (S_1 v, \ldots, S_m v), \end{aligned}$$

so that $\varphi_k S = S_k$ for $k = 1, \ldots, m$. Then

$$\begin{aligned} \varphi(S) &= (\pi_1 S, \ldots, \pi_m S) \\ &= (S_1, \ldots, S_n) \end{aligned}$$

and S is indeed surjective. Therefore, φ is an isomorphism, and we have

$$\mathcal{L}(V, W_1 \times \cdots \times W_m) \cong \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m),$$

as desired. $\square$

Problem 7

Suppose v, x are vectors in V and U, W are subspaces of V such that $v + U = x + W$. Prove that $U = W$.

Proof. First note that since $v + 0 = v \in v + U$, there exists $w_0 \in W$ such that $v = x + w_0$, and hence $v - x = w_0 \in W$. Similarly, there exists $u_0 \in U$ such that $x - v = u_0 \in U$.

Suppose $u \in U$. Then there exists $w \in W$ such that $v + u = x + w$, and hence

$$u = (x - v) + w = -w_0 + w \in W,$$

and we have $U \subseteq W$. Conversely, suppose $w' \in W$. Then there exists $u' \in U$ such that $x + w' = v + u'$, and hence

$$w' = (v - x) + u' = -u_0 + u' \in U,$$

and we have $W \subseteq U$. Therefore $U = W$, as desired. $\square$

Problem 8

Prove that a nonempty subset A of V is an affine subset of V if and only if $\lambda v + (1 - \lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbb{F}$.

Proof. ($\Rightarrow$) Suppose $A \subseteq V$ is an affine subset of V. Then there exists $x \in V$ and a subspace $U \subseteq V$ such that $A = x + U$. Suppose $v, w \in A$. Then there exist $u_1, u_2 \in U$ such that $v = x + u_1$ and $w = x + u_2$. Thus, for all $\lambda \in \mathbb{F}$, we have

$$\begin{aligned} \lambda v + (1 - \lambda)w &= \lambda(x + u_1) + (1 - \lambda)(x + u_2) \\ &= x + \lambda u_1 + (1 - \lambda)u_2. \end{aligned}$$

Since $\lambda u_1 + (1 - \lambda)u_2 \in U$, this implies $v + (1 - \lambda)w \in x + U = A$, as desired.

($\Leftarrow$) Suppose $\lambda v + (1 - \lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbb{F}$. Choose $a \in A$ and define

$$U := -a + A.$$

We claim U is a subspace of V. Clearly $0 \in U$ since $a \in A$. Let $x \in U$, so that $x = -a + a_0$ for some $a_0 \in A$, and let $\lambda \in \mathbb{F}$. It follows

$$\lambda a_0 + (1 - \lambda)a \in A \Rightarrow -\lambda a + \lambda a_0 + a \in A \Rightarrow \lambda(-a + a_0) \in -a + A = U,$$

and thus $\lambda x = \lambda(-a + a_0) \in U$, and U is closed under scalar multiplication. Now let $x, y \in U$. Then there exist $a_1, a_2 \in A$ such that $x = -a + a_1$ and $y = -a + a_2$. Notice

$$\frac{1}{2}a_1 + \left(1 - \frac{1}{2}\right)a_2 = \frac{1}{2}a_1 + \frac{1}{2}a_2 \in A,$$

and hence

$$-a + \frac{1}{2}a_1 + \frac{1}{2}a_2 \in U.$$

It follows

$$\begin{aligned} x + y &= -2a + a_1 + a_2 \\ &= 2\left(-a + \frac{1}{2}a_1 + \frac{1}{2}a_2\right) \in U, \end{aligned}$$

using the fact that U has already been shown to be closed under scalar multiplication. Thus U is also closed under addition, and so U is a subspace of V. Now, since $A = a + U$, we have that A is indeed an affine subset of V, as desired. $\square$

Problem 9

Suppose A_1 and A_2 are affine subsets of V. Prove that the intersection $A_1 \cap A_2$ is either an affine subset of V or the empty set.

Proof. If $A_1 \cap A_2 = \emptyset$, we're done, so suppose $A_1 \cap A_2$ is nonempty and let $v \in A_1 \cap A_2$. Then we may write

$$A_1 = v + U_1 \quad \text{and} \quad A_2 = v + U_2$$

for some subspaces $U_1, U_2 \subseteq V$.

We claim $A_1 \cap A_2 = v + (U_1 \cap U_2)$, which is an affine subset of V. To see this, suppose $x \in v + (U_1 \cap U_2)$. Then there exists $u \in U_1 \cap U_2$ such that $x = v + u$. Since $u \in U_1$, we have $x \in v + U_1 = A_1$. And since $u \in U_2$, we have $x \in v + U_2 = A_2$. Thus $x \in A_1 \cap A_2$ and $v + (U_1 \cap U_2) \subseteq A_1 \cap A_2$. Conversely, suppose $y \in A_1 \cap A_2$. Then there exist $u_1 \in U_1$ and $u_2 \in U_2$ such that $y = v + u_1$ and $y = v + u_2$. But this implies $u_1 = u_2$, and hence $u_1 = u_2 \in U_1 \cap U_2$, thus $y \in v + (U_1 \cap U_2)$. Therefore $A_1 \cap A_2 \subseteq v + (U_1 \cap U_2)$, and hence we have $A_1 \cap A_2 = v + (U_1 \cap U_2)$, as claimed. $\square$

Problem 11

Suppose $v_1, \ldots, v_m \in V$. Let

$$A = \{\lambda_1 v_1 + \cdots + \lambda_m v_m \mid \lambda_1, \ldots, \lambda_m \in \mathbb{F} \text{ and } \lambda_1 + \cdots + \lambda_m = 1\}.$$

(a) Prove that A is an affine subset of V.

(b) Prove that every affine subset of V that contains $v_1, \ldots, v_m$ also contains A.

(c) Prove that $A = v + U$ for some $v \in V$ and some subspace U of V with $\dim U \leq m - 1$.

Proof. (a) Let $v, w \in A$, so that there exist $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$ and $\beta_1, \ldots, \beta_m \in \mathbb{F}$ such that

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_m v_m \\ w &= \beta_1 v_1 + \cdots + \beta_m v_m, \end{aligned}$$

where $\sum \alpha_k = 1$ and $\sum \beta_k = 1$. Given $\lambda \in \mathbb{F}$, it follows

$$\begin{aligned} \lambda v + (1-\lambda) w &= \lambda \sum_{k=1}^{m} \alpha_k v_k + (1-\lambda) \sum_{k=1}^{m} \beta_k v_k \\ &= \sum_{k=1}^{m} \left[\lambda \alpha_k + (1-\lambda)\beta_k\right] v_k. \end{aligned}$$

But notice

$$\sum_{k=1}^{m} \left[\lambda \alpha_k + (1-\lambda)\beta_k\right] = \lambda + (1-\lambda) = 1,$$

and hence $\lambda v + (1-\lambda) w \in A$ by the way we defined A. By Problem 8, this implies that A is an affine subset of V, as was to be shown.

(b) We induct on m.
Base case: When $m = 1$, the statement is trivially true, since $A = \{v_1\}$, and hence any affine subset of V that contains v_1 of course contains A.
Inductive step: Let $k \in \mathbb{Z}^+$, and suppose the statement is true for $m = k$. Suppose A' is an affine subset of V that contains $v_1, \ldots, v_{k+1}$, and let $x \in A$. Then there exist $\lambda_1, \ldots, \lambda_{k+1} \in \mathbb{F}$ such that $\sum_j \lambda_j = 1$ and

$$x = \lambda_1 v_1 + \cdots + \lambda_{k+1} v_{k+1}.$$

Now, if $\lambda_{k+1} = 1$, then $x = v_{k+1} \in A'$. Otherwise, we have

$$\frac{\lambda_1}{1-\lambda_{k+1}} + \cdots + \frac{\lambda_k}{1-\lambda_{k+1}} = 1,$$

and hence by our inductive hypothesis, this implies

$$\frac{\lambda_1}{1-\lambda_{k+1}} v_1 + \cdots + \frac{\lambda_k}{1-\lambda_{k+1}} v_k \in A'.$$

By Problem 8, we know

$$(1-\lambda_{k+1}) \left(\frac{\lambda_1}{1-\lambda_{k+1}} v_1 + \cdots + \frac{\lambda_k}{1-\lambda_{k+1}} v_k \right) + \lambda_{k+1} v_{k+1} \in A'.$$

But after simplifying, this tells us

$$\lambda_1 v_1 + \cdots + \lambda_{k+1} v_{k+1} = x \in A'.$$

Hence $A \subseteq A'$, and the statement is true for $m = k+1$.

By the principal of mathematical induction, the statement is true for all $m \in \mathbb{Z}^+$. Thus any affine subset of V that contains $v_1, \ldots, v_m$ also contains A, as was to be shown.

(c) Define $U := \operatorname{span}(v_2 - v_1, \ldots, v_m - v_1)$. Let $x \in A$, so that there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$ with $\sum_k \lambda_k = 1$ such that

$$x = \lambda_1 v_1 + \cdots + \lambda_m v_m.$$

Notice

$$\begin{aligned} v_1 + \lambda_2(v_2 - v_1) + \cdots + \lambda_m(v_m - v_1) &= \left(1 - \sum_{k=2}^{m} \lambda_k\right) v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m \\ &= \lambda_1 v_1 + \cdots + \lambda_m v_m \\ &= x, \end{aligned}$$

and hence $x \in v_1 + U$, so that $A \subseteq v_1 + U$. Next suppose $y \in v_1 + U$, so that there exist $\alpha_1, \ldots, \alpha_{m-1} \in \mathbb{F}$ such that

$$y = v_1 + \alpha_1(v_2 - v_1) + \cdots + \alpha_{m-1}(v_m - v_1).$$

Expanding the RHS yields

$$y = \left(1 - \sum_{k=1}^{m-1} \alpha_k\right) v_1 + \alpha_1 v_2 + \cdots + \alpha_{m-1} v_m.$$

But since

$$\left(1 - \sum_{k=1}^{m-1} \alpha_k\right) + \sum_{k=1}^{m-1} \alpha_k = 1,$$

this implies $y \in A$, and hence $v_1 + U \subseteq A$. Therefore $A = v_1 + U$, and since $\dim U \leq m - 1$, we have the desired result. $\square$

Problem 13

Suppose U is a subspace of V and $v_1 + U, \ldots, v_m + U$ is a basis of V/U and $u_1, \ldots, u_n$ is a basis of U. Prove that $v_1, \ldots, v_m, u_1, \ldots, u_n$ is a basis of V.

Proof. Since

$$\begin{aligned} \dim V &= \dim V/U + \dim U \\ &= m + n, \end{aligned}$$

it suffices to show $v_1, \ldots, v_m, u_1, \ldots, u_n$ spans V. Suppose $v \in V$. Then there exist $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$ such that

$$v + U = \alpha_1(v_1 + U) + \cdots + \alpha_m(v_m + U).$$

But then

$$v + U = (\alpha_1 v_1 + \cdots + \alpha_m v_m) + U$$

and hence

$$v - (\alpha_1 v_1 + \cdots + \alpha_m v_m) \in U.$$

Thus there exist $\beta_1, \ldots, \beta_n \in U$ such that

$$v - (\alpha_1 v_1 + \cdots + \alpha_m v_m) = \beta_1 u_1 + \cdots + \beta_n u_n,$$

and we have

$$v = \alpha_1 v_1 + \cdots + \alpha_m v_m + \beta_1 u_1 + \cdots + \beta_n u_n,$$

so that indeed $v_1, \ldots, v_m, u_1, \ldots, u_n$ spans V. $\square$

Problem 15

Suppose $\varphi \in \mathcal{L}(V, \mathbb{F})$ and $\varphi \neq 0$. Prove that $\dim V/(\operatorname{null} \varphi) = 1$.

Proof. Since $\varphi \neq 0$, we must have $\dim \operatorname{range} \varphi = 1$, so that $\operatorname{range} \varphi = \mathbb{F}$. Since $V/(\operatorname{null} \varphi) \cong \operatorname{range} \varphi$, this implies $V/(\operatorname{null} \varphi) \cong \mathbb{F}$, and hence $\dim V/(\operatorname{null} \varphi) = 1$, as desired. $\square$

Problem 17

Suppose U is a subspace of V such that V/U is finite-dimensional. Prove that there exists a subspace W of V such that $\dim W = \dim V/U$ and $V = U \oplus W$.

Proof. Suppose $\dim V/U = n$, and let $v_1 + U, \ldots, v_n + U$ be a basis of V/U. Define $W := \operatorname{span}(v_1, \ldots, v_n)$. We claim $v_1, \ldots, v_n$ must be linearly independent, so that $v_1, \ldots, v_n$ is a basis of W. To see this, suppose $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ are such that

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0.$$

Then

$$(\alpha_1 v_1 + \cdots + \alpha_n v_n) + U = \alpha_1(v_1 + U) + \cdots + \alpha_n(v_n + U),$$

and hence we must have $\alpha_1 = \cdots = \alpha_n = 0$. Thus $v_1, \ldots, v_n$ is indeed linearly independent, as claimed.

We now claim $V = U \oplus W$. To see that $V = U + W$, suppose $v \in V$. Then there exist $\beta_1, \ldots, \beta_n \in \mathbb{F}$ such that

$$v + U = \beta_1(v_1 + U) + \cdots + \beta_n(v_n + U).$$

It follows

$$v - \sum_{k=1}^{n} \beta_k v_k \in U,$$

and hence

$$v = \left(v - \sum_{k=1}^{n} \beta_k v_k \right) + \left(\sum_{k=1}^{n} \beta_k v_k \right).$$

Since first term in parentheses is in U and the second term in parentheses is in W, we have $v \in U + W$, and hence $V \subseteq U + W$. Clearly $U + W \subseteq V$, since U and W are each subspaces of V, and hence $V = U + W$. To see that the sum is direct, suppose $w \in U \cap W$. Since $w \in W$, there exist $\lambda_1, \ldots, \lambda_n$ such that $w = \lambda_1 v_1 + \cdots + \lambda_n v_n$, and hence

$$\begin{aligned} w + U &= (\lambda_1 v_1 + \cdots + \lambda_n v_n) + U \\ &= \lambda_1 (v_1 + U) + \cdots + \lambda_n (v_n + U). \end{aligned}$$

Since $w \in U$, we have $w + U = 0 + U$. Thus $\lambda_1 = \cdots = \lambda_n = 0$, which implies $w = 0$. Since $U \cap W = \{0\}$, the sum is indeed direct. Thus $V = U \oplus W$, with $\dim W = n = \dim V/U$, as desired. □

Problem 19

Find a correct statement analogous to 3.78 that is applicable to finite sets, with unions analogous to sums of subspaces and disjoint unions analogous to direct sums.

Theorem. *Suppose $|V| < \infty$ and $U_1, \ldots, U_n \subseteq V$. Then $U_1, \ldots, U_n$ are pairwise disjoint if and only if*

$$|U_1 \cup \cdots \cup U_n| = |U_1| + \cdots + |U_n|.$$

Proof. We induct on n.
Base case: Let $n = 2$. Since $|U_1 \cup U_2| = |U_1| + |U_2| - |U_1 \cap U_2|$, we have $U_1 \cap U_2 = \emptyset$ iff $|U_1 \cup U_2| = |U_1| + |U_2|$.
Inductive hypothesis: Let $k \in \mathbb{Z}_{\geq 2}$, and suppose the statement is true for $n = k$. Let $U_{k+1} \subseteq V$. Then

$$|U_1 \cup \cdots \cup U_{k+1}| = |U_1 \cup \cdots \cup U_k| + |U_{k+1}|$$

iff $U_{k+1} \cap (U_1 \cup \cdots \cup U_k) = \emptyset$ by our base case. Combining this with our inductive hypothesis, we have

$$|U_1 \cup \cdots \cup U_{k+1}| = |U_1| + \cdots + |U_k| + |U_{k+1}|$$

iff $U_1, \ldots, U_{k+1}$ are pairwise disjoint, and the statement is true for $n = k + 1$.

By the principal of mathematical induction, the statement is true for all $n \in \mathbb{Z}_{\geq 2}$. □

F: Duality

Problem 1

Explain why every linear functional is either surjective or the zero map.

Proof. Since $\dim \mathbb{F} = 1$, the only subspaces of $\mathbb{F}$ are $\mathbb{F}$ itself and $\{0\}$. Let V be a vector space (not necessarily finite-dimensional) and suppose $\varphi \in V'$. Since $\operatorname{range} \varphi$ is a subspace of $\mathbb{F}$, it must be either $\mathbb{F}$ itself (in which case φ is surjective) or $\{0\}$ (in which case φ is the zero map). $\square$

Problem 3

Suppose V is finite-dimensional and U is a subspace of V such that $U \neq V$. Prove that there exists $\varphi \in V'$ such that $\varphi(u) = 0$ for every $u \in U$ but $\varphi \neq 0$.

Proof. Suppose $\dim U = m$ and $\dim V = n$ for some $m, n \in \mathbb{Z}^+$ such that $m < n$. Let $u_1, \ldots, u_m$ be a basis of U. Expand this to a basis $u_1, \ldots, u_m, u_{m+1}, \ldots, u_n$ of V, and let $\varphi_1, \ldots, \varphi_n$ be the corresponding dual basis of V'. For any $u \in U$, there exist $\alpha_1, \ldots, \alpha_m$ such that $u = \alpha_1 u_1 + \cdots + \alpha_m u_m$. Now notice

$$\begin{aligned} \varphi_{m+1}(u) &= \varphi_{m+1}(\alpha_1 u_1 + \cdots + \alpha_m u_m) \\ &= \alpha_1 \varphi_{m+1}(u_1) + \cdots + \alpha_m \varphi_{m+1}(u_m) \\ &= 0, \end{aligned}$$

but $\varphi_{m+1}(u_{m+1}) = 1$. Thus $\varphi_{m+1}(u) = 0$ for every $u \in U$ but $\varphi_{m+1} \neq 0$, as desired. $\square$

Problem 5

Suppose $V_1, \ldots, V_m$ are vector spaces. Prove that $(V_1 \times \cdots \times V_m)'$ and $V_1' \times \cdots \times V_m'$ are isomorphic vector spaces.

Proof. For $i = 1, \ldots, m$, let

$$\begin{aligned} \xi_i : V_i &\to V_1 \times \cdots \times V_m \\ v_i &\mapsto (0, \ldots, v_i, \ldots, 0). \end{aligned}$$

Now define

$$\begin{aligned} T : (V_1 \times \cdots \times V_m)' &\to V_1' \times \cdots \times V_m' \\ \varphi &\mapsto (\varphi \circ \xi_1, \ldots, \varphi \circ \xi_m). \end{aligned}$$

We claim T is an isomorphism. We must show three things: (1) that T is a linear map; (2) that T is injective; and (3) that T is surjective.

To see that T is a linear map, first suppose $\varphi_1, \varphi_2 \in (V_1 \times \cdots \times V_m)'$. It follows

$$\begin{aligned} T(\varphi_1 + \varphi_2) &= \big((\varphi_1 + \varphi_2) \circ \xi_1, \ldots, (\varphi_1 + \varphi_2) \circ \xi_m\big) \\ &= (\varphi_1 \circ \xi_1 + \varphi_2 \circ \xi_1, \ldots, \varphi_1 \circ \xi_m + \varphi_2 \circ \xi_m) \\ &= (\varphi_1 \circ \xi_1, \ldots, \varphi_1 \circ \xi_m) + (\varphi_2 \circ \xi_1, \ldots, \varphi_2 \circ \xi_m) \\ &= T(\varphi_1) + T(\varphi_2), \end{aligned}$$

thus T is additive. To see that it is also homogeneous, suppose $\lambda \in \mathbb{F}$ and $\varphi \in (V_1 \times \cdots \times V_m)'$. We have

$$\begin{aligned} T(\lambda\varphi) &= \big((\lambda\varphi) \circ \xi_1, \ldots, (\lambda\varphi) \circ \xi_m\big) \\ &= \big(\lambda(\varphi \circ \xi_1), \ldots, \lambda(\varphi \circ \xi_m)\big) \\ &= \lambda\,(\varphi \circ \xi_1, \ldots, \varphi \circ \xi_m) \\ &= \lambda T(\varphi), \end{aligned}$$

and thus T is homogeneous as well and therefore it is a linear map.

To see that T is injective, suppose $\varphi, \psi \in (V_1 \times \cdots \times V_m)'$ but $\varphi \neq \psi$. Then there exists some $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$ such that $\varphi(v_1, \ldots, v_m) \neq \psi(v_1, \ldots, v_m)$. Since φ and ψ are linear, this means that there exists some index $k \in \{1, \ldots, m\}$ such that $\varphi(0, \ldots, v_k, \ldots, 0) \neq \psi(0, \ldots, v_k, \ldots, 0)$. But then $\varphi \circ \xi_k \neq \psi \circ \xi_k$, and hence $T(\varphi) \neq T(\psi)$, so that T is injective.

To see that T is surjective, suppose $(\varphi_1, \ldots, \varphi_m) \in V_1' \times \cdots \times V_m'$ and define

$$\begin{aligned} \theta : V_1 \times \cdots \times V_m &\to \mathbb{F} \\ (v_1, \ldots, v_m) &\mapsto \sum_{k=1}^{m} \varphi_k(v_k). \end{aligned}$$

We claim $T(\theta) = (\varphi_1, \ldots, \varphi_m)$. To see this, let $k \in \{1, \ldots, m\}$. We will show that the map in the k-th component of $T(\theta)$ is equal to φ_k. Given $v_k \in V_k$, we have

$$\begin{aligned} T(\theta)_k(v_k) &= (\theta \circ \xi_k)(v_k) \\ &= \theta(\xi_k(v_k)) \\ &= \theta(0, \ldots, v_k, \ldots, 0) \\ &= \varphi_1(0) + \cdots + \varphi_k(v_k) + \cdots + \varphi_m(0) \\ &= \varphi_k(v_k), \end{aligned}$$

as desired. Thus $T(\theta) = (\varphi_1, \ldots, \varphi_m)$, and T is indeed surjective. Since T is both injective and surjective, it's an isomorphism. $\square$

Problem 7

Suppose m is a positive integer. Show that the dual basis of the basis $1, \ldots, x^m$ of $\mathcal{P}_m(\mathbb{R})$ is $\varphi_0, \varphi_1, \ldots, \varphi_m$, where $\varphi_j(p) = \frac{p^{(j)}(0)}{j!}$. Here $p^{(j)}$ denotes the j^{th} derivative of p, with the understanding that the 0^{th} derivative of p is p.

Proof. For $j = 0, \ldots, m$, we have by direct computation of the j-th derivative

$$\varphi_j\left(x^k\right) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise,} \end{cases}$$

so that $\varphi_0, \varphi_1, \ldots, \varphi_m$ is indeed the dual basis of $1, \ldots, x^m$. Note the uniqueness of the dual basis follows by uniqueness of a linear map (including the linear functionals in the dual basis) whose values on a basis are specified. □

Problem 9

Suppose $v_1, \ldots, v_n$ is a basis of V and $\varphi_1, \ldots, \varphi_n$ is the corresponding dual basis of V'. Suppose $\psi \in V'$. Prove that

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

Proof. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ be such that

$$\psi = \alpha_1\varphi_1 + \cdots + \alpha_n\varphi_n.$$

For $k = 1, \ldots, n$, we have

$$\begin{aligned} \psi(v_k) &= \alpha_1\varphi_1(v_k) + \cdots + \alpha_k\varphi_k(v_k) + \cdots + \alpha_n\varphi_n(v_k) \\ &= \alpha_1 \cdot 0 + \cdots + \alpha_k \cdot 1 + \cdots + \alpha_n \cdot 0 \\ &= \alpha_k. \end{aligned}$$

Thus we have

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n,$$

as desired □

Problem 11

Suppose A is an m-by-n matrix with $A \neq 0$. Prove that the rank of A is 1 if and only if there exist $(c_1, \ldots, c_m) \in \mathbb{F}^m$ and $(d_1, \ldots, d_n) \in \mathbb{F}^n$ such that $A_{j,k} = c_j d_k$ for every $j = 1, \ldots, m$ and every $k = 1, \ldots, n$.

Proof. ($\Rightarrow$) Suppose the rank of A is 1. By the assumption that $A \neq 0$, there exists a nonzero entry $A_{i.,j}$ for some $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$. Thus $\operatorname{span}\{A_{\cdot,1}, \dots, A_{\cdot,n}\} = \operatorname{span}\{A_{\cdot,j}\}$, and hence there exist $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ such that $A_{\cdot,c} = \alpha_c A_{\cdot,j}$ for $c = 1, \dots, n$. Expanding out each out these columns, we have

$$A_{r,c} = \alpha_c A_{r,j} \tag{7}$$

for $r = 1, \dots, m$. Similarly for the rows, we have $\operatorname{span}\{A_{1,\cdot}, \dots, A_{m,\cdot}\} = \operatorname{span}\{A_{i,\cdot}\}$, and hence there exist $\beta_1, \dots, \beta_m \in \mathbb{F}$ such that $A_{r',\cdot} = \beta_r A_{i,\cdot}$ for $r' = 1, \dots, m$. Expanding out each of these rows, we have

$$A_{r',c'} = \beta_{r'} A_{i,c'} \tag{8}$$

for $c' = 1, \dots, n$. Now by replacing the $A_{r,j}$ term in (7) according to (8), we have $A_{r,j} = \beta_r A_{i,j}$, and hence (7) may be rewritten

$$A_{r,c} = \alpha_c \beta_r A_{i,j},$$

and the result follows by defining $c_r = \beta_r A_{i,j}$ and $d_c = \alpha_c$ for $r = 1, \dots, m$ and $c = 1, \dots, n$.

($\Leftarrow$) Suppose there exist $(c_1, \dots, c_m) \in \mathbb{F}^m$ and $(d_1, \dots, d_n) \in \mathbb{F}^n$ such that $A_{j,k} = c_j d_k$ for every $j = 1, \dots, m$ and every $k = 1, \dots, n$. Then each of the columns is a scalar multiple of $(d_1, \dots, d_n)^t \in \mathbb{F}^{n,1}$ and the column rank is 1. Since the rank of a matrix equals its column rank, the rank of A is 1 as well. $\square$

Problem 13

Define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by $T(x, y, z) = (4x + 5y + 6z, 7x + 8y + 9z)$. Suppose φ_1, φ_2 denotes the dual basis of the standard basis of $\mathbb{R}^2$ and ψ_1, ψ_2, ψ_3 denotes the dual basis of the standard basis of $\mathbb{R}^3$.

(a) Describe the linear functionals $T'(\varphi_1)$ and $T'(\varphi_2)$.

(b) Write $T'(\varphi_1)$ and $T'(\varphi_2)$ as a linear combination of ψ_1, ψ_2, ψ_3.

Proof. (a) Endowing $\mathbb{R}^3$ and $\mathbb{R}^2$ with their respective standard basis, we have

$$\begin{aligned}(T'(\varphi_1))(x, y, z) &= (\varphi_1 \circ T)(x, y, z) \\ &= \varphi_1(T(x, y, z)) \\ &= \varphi_1(4x + 5y + 6z, 7x + 8y + 9z) \\ &= 4x + 5y + 6z\end{aligned}$$

and similarly

$$\begin{aligned}(T'(\varphi_2))(x, y, z) &= \varphi_2(4x + 5y + 6z, 7x + 8y + 9z) \\ &= 7x + 8y + 9z.\end{aligned}$$

(b) Notice

$$\begin{aligned}(4\psi_1 + 5\psi_2 + 6\psi_3)(x, y, z) &= 4\psi_1(x, y, z) + 5\psi_2(x, y, z) + 6\psi_3(x, y, z)\\ &= 4x + 5y + 6z\\ &= T'(\varphi_1)(x, y, z)\end{aligned}$$

and

$$\begin{aligned}(7\psi_1 + 8\psi_2 + 9\psi_3)(x, y, z) &= 7\psi_1(x, y, z) + 8\psi_2(x, y, z) + 9\psi_3(x, y, z)\\ &= 7x + 8y + 9z\\ &= T'(\varphi_2)(x, y, z),\end{aligned}$$

as desired. □

Problem 15

Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that $T' = 0$ if and only if $T = 0$.

Proof. ($\Rightarrow$) Suppose $T' = 0$. Let $\varphi \in W'$ and $v \in V$ be arbitrary. We have

$$0 = (T'(\varphi))(v) = \varphi(Tv).$$

Since φ is arbitrary, we must have $Tv = 0$. But now since v is arbitrary, this implies $T = 0$ as well.

($\Leftarrow$) Suppose $T = 0$. Again let $\varphi \in W'$ and $v \in V$ be arbitrary. We have

$$(T'(\varphi))(v) = \varphi(Tv) = \varphi(0) = 0,$$

and hence $T' = 0$ as well. □

Problem 17

Suppose $U \subseteq V$. Explain why $U^0 = \{\varphi \in V' \mid U \subseteq \operatorname{null}\varphi\}$.

Proof. It suffices to show that, for arbitrary $\varphi \in V'$, we have $U \subseteq \operatorname{null}\varphi$ if and only if $\varphi(u) = 0$ for all $u \in U$. So suppose $U \subseteq \operatorname{null}\varphi$. Then for all $u \in U$, we have $\varphi(u) = 0$ (simply by definition of $\operatorname{null}\varphi$). Conversely, suppose $\varphi(u) = 0$ for all $u \in U$. Then if $u' \in U$, we must have $u' \in \operatorname{null}\varphi$. That is, $U \subseteq \operatorname{null}\varphi$, completing the proof. □

Problem 19

Suppose V is finite-dimensional and U is a subspace of V. Show that $U = V$ if and only if $U^0 = \{0\}$.

Proof. ($\Rightarrow$) Suppose $U = V$. Then

$$\begin{aligned} U^0 &= \{\varphi \in V' \mid U \subseteq \operatorname{null}\varphi\} \\ &= \{\varphi \in V' \mid V \subseteq \operatorname{null}\varphi\} \\ &= \{0\}, \end{aligned}$$

since only the zero functional can have all of V in its null space.

($\Leftarrow$) Suppose $U^0 = \{0\}$. It follows

$$\begin{aligned} \dim V &= \dim U + \dim U^0 \\ &= \dim U + 0 \\ &= \dim U. \end{aligned}$$

Since the only subspace of V with dimension $\dim V$ is V itself, we have $U = V$, as desired. $\square$

Problem 20

Suppose U and W are subsets of V with $U \subseteq W$. Prove that $W^0 \subseteq U^0$.

Proof. Suppose $\varphi \in W^0$. Then $\varphi(w) = 0$ for all $w \in W$. If $\varphi \notin U^0$, then there exists some $u \in U$ such that $\varphi(u) \neq 0$. But since $U \subseteq W$, $u \in W$. This is absurd, hence we must have $\varphi \in U^0$. Thus $W^0 \subseteq U^0$, as desired. $\square$

Problem 21

Suppose V is finite-dimensional and U and W are subspaces of V with $W^0 \subseteq U^0$. Prove that $U \subseteq W$.

Proof. Suppose not. Then there exists a nonzero vector $u \in U$ such that $u \notin W$. There exists some basis of U containing u. Define $\varphi \in V'$ such that, for any vector v in this basis, we have

$$\varphi(v) = \begin{cases} 1 & \text{if } v = u \\ 0 & \text{otherwise.} \end{cases}$$

By construction, $\varphi \in W^0$, and hence $\varphi \in U^0$. But this implies $\varphi(u) = 0$, a contradiction. $\square$

Problem 22

Suppose U, W are subspaces of V. Show that $(U + W)^0 = U^0 \cap W^0$.

Proof. Since $U \subseteq U + W$ and $W \subseteq U + W$, Problem 20 tells us that $(U+W)^0 \subseteq U^0$ and $(U+W)^0 \subseteq W^0$. Thus $(U+W)^0 \subseteq U^0 \cap W^0$. Conversely, suppose $\varphi \in U^0 \cap W^0$. Let $x \in U + W$. Then there exist $u \in U$ and $w \in W$ such that $x = U + W$. Then

$$\begin{aligned}\varphi(x) &= \varphi(u+w)\\ &= \varphi(u) + \varphi(w)\\ &= 0,\end{aligned}$$

where the second equality follows since $\varphi \in U^0$ and $\varphi \in W^0$ by assumption. Hence $\varphi \in (U+W)^0$ and we have $U^0 + W^0 \subseteq (U+W)^0$. Thus $(U+W)^0 = U^0 \cap W^0$, as desired. ☐

Problem 23

Suppose V is finite-dimensional and U and W are subspaces of V. Prove that $(U \cap W)^0 = U^0 + W^0$.

Proof. Since $U \cap W \subseteq U$ and $U \cap W \subseteq W$, Problem 20 tells us that $U^0 \subseteq (U \cap W)^0$ and $W^0 \subseteq (U \cap W)^0$. Thus $U^0 + W^0 \subseteq (U \cap W)^0$. Now, notice (using Problem 22 to deduce the second equality)

$$\begin{aligned}\dim(U^0 + W^0) &= \dim(U^0) + \dim(W^0) - \dim(U^0 \cap W^0)\\ &= \dim(U^0) + \dim(W^0) - \dim((U+W)^0)\\ &= (\dim V - \dim U) + (\dim V - \dim W) - [\dim V - \dim(U+W)]\\ &= \dim V - \dim U - \dim W + \dim(U+W)\\ &= \dim V - [\dim U + \dim W - \dim(U+W)]\\ &= \dim V - \dim(U \cap W)\\ &= \dim((U \cap W)^0).\end{aligned}$$

Hence we must have $U^0 + W^0 = (U \cap W)^0$, as desired. ☐

Problem 25

Suppose V is finite-dimensional and U is a subspace of V. Show that

$$U = \{v \in V \mid \varphi(v) = 0 \text{ for every } \varphi \in U^0\}.$$

Proof. Let $A = \{v \in V \mid \varphi(v) = 0 \text{ for every } \varphi \in U^0\}$. Suppose $u \in U$. Then $\varphi(u) = 0$ for all $\varphi \in U^0$, and hence $u \in A$, showing $U \subseteq A$.

Conversely, suppose $v \in A$ but $v \notin U$. Since $0 \in U$, we must have $v \neq 0$. Thus there exists a basis $u_1, \ldots, u_m, v, v_1, \ldots, v_n$ of V such that $u_1, \ldots, u_m$ is a basis of U. Let $\psi_1, \ldots, \psi_m, \varphi, \varphi_1, \ldots, \varphi_n$ be the dual basis of V', and consider for a moment the functional φ. Clearly we have both $\varphi \in U^0$ and $\varphi(v) = 1$ by construction, but this is a contradiction, since we assumed $v \in A$. Thus $A \subseteq U$, and we conclude $U = A$, as was to be shown. ☐

Problem 27

Suppose $T \in \mathcal{L}(\mathcal{P}_5(\mathbb{R}), \mathcal{P}_5(\mathbb{R}))$ and $\operatorname{null} T' = \operatorname{span}(\varphi)$, where φ is the linear functional on $\mathcal{P}_5(\mathbb{R})$ defined by $\varphi(p) = p(8)$. Prove that $\operatorname{range} T = \{p \in \mathcal{P}_5(\mathbb{R}) \mid p(8) = 0\}$.

Proof. By Theorem 3.107, we know $\operatorname{null} T' = (\operatorname{range} T)^0$, and hence $(\operatorname{range} T)^0 = \{\alpha\varphi \mid \alpha \in \mathbb{R}\}$. It follows by Problem 25

$$\begin{aligned} \operatorname{range} T &= \{p \in \mathcal{P}_5(\mathbb{R}) \mid \psi(p) = 0 \text{ for all } \psi \in (\operatorname{range} T)^0\} \\ &= \{p \in \mathcal{P}_5(\mathbb{R}) \mid (\alpha\varphi)(p) = 0 \text{ for all } \alpha \in \mathbb{R}\} \\ &= \{p \in \mathcal{P}_5(\mathbb{R}) \mid \varphi(p) = 0\} \\ &= \{p \in \mathcal{P}_5(\mathbb{R}) \mid p(8) = 0\}, \end{aligned}$$

as desired. ☐

Problem 29

Suppose V and W are finite-dimensional, $T \in \mathcal{L}(V, W)$, and there exists $\varphi \in V'$ such that $\operatorname{range} T' = \operatorname{span}(\varphi)$. Prove that $\operatorname{null} T = \operatorname{null} \varphi$.

Proof. By Theorem 3.107, we know $\operatorname{range} T' = (\operatorname{null} T)^0$, and hence $(\operatorname{null} T)^0 = \{\alpha\varphi \mid \alpha \in \mathbb{R}\}$. It follows by Problem 25

$$\begin{aligned} \operatorname{null} T &= \{v \in V \mid \psi(v) = 0 \text{ for all } \psi \in (\operatorname{null} T)^0\} \\ &= \{v \in V \mid \alpha\varphi(v) = 0 \text{ for all } \alpha \in \mathbb{F}\} \\ &= \{v \in V \mid \varphi(v) = 0\} \\ &= \operatorname{null} \varphi, \end{aligned}$$

as desired. ☐

Problem 31

Suppose V is finite-dimensional and $\varphi_1, \ldots, \varphi_n$ is a basis of V'. Show that there exists a basis of V whose dual basis is $\varphi_1, \ldots, \varphi_n$.

Proof. To prove this, we will first show $V \cong V''$. We will then take an existing basis of V', map it to its dual basis in V'', and then use the inverse of the isomorphism to take this basis of V'' to a basis in V. This basis of V will have the known basis of V' as its dual.

So, for any $v \in V$, define $E_v \in V''$ by $E_v(\varphi) = \varphi(v)$. We claim the map $\hat{\cdot} : V \to V''$ given by $\hat{v} = E_v$ is an isomorphism. To do so, it suffices to show it to be both linear and injective, since $\dim(V'') = \dim((V')') = \dim(V') = \dim(V)$.

We first show $\hat{\cdot}$ is linear. So suppose $u, v \in V$. Then for any $\varphi \in V'$, we have

$$\begin{aligned}
(\widehat{u+v})(\varphi) &= E_{u+v}(\varphi) \\
&= \varphi(u+v) \\
&= \varphi(u) + \varphi(v) \\
&= E_u(\varphi) + E_v(\varphi) \\
&= \hat{u}(\varphi) + \hat{v}(\varphi)
\end{aligned}$$

so that $\hat{\cdot}$ is indeed linear. Next we show it to be homogeneous. So suppose $\lambda \in \mathbb{F}$, and again let $v \in V$. Then for any $\varphi \in V'$, we have

$$\begin{aligned}
(\widehat{\lambda v})(\varphi) &= E_{\lambda v}(\varphi) \\
&= \varphi(\lambda v) \\
&= \lambda \varphi(v) \\
&= \lambda E_v(\varphi) \\
&= \lambda \hat{v},
\end{aligned}$$

so that $\hat{\cdot}$ is homogenous as well. Being both linear and homogenous, it is a linear map.

Next we show $\hat{\cdot}$ is injective. So suppose $\hat{v} = 0$ for some $v \in V$. We want to show $v = 0$. Let $v_1, \ldots, v_n$ be a basis of V. Then there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Then, for all $\varphi \in V'$, we have

$$\begin{aligned}
\hat{v} = 0 &\implies (\alpha_1 v_1 + \cdots + \alpha_n v_n)^\wedge = 0 \\
&\implies \alpha_1 \widehat{v_1} + \cdots + \alpha_n \widehat{v_n} = 0 \\
&\implies (\alpha_1 \widehat{v_1} + \cdots + \alpha_n \widehat{v_n})(\varphi) = 0 \\
&\implies \alpha_1 \widehat{v_1}(\varphi) + \cdots + \alpha_n \widehat{v_n}(\varphi) = 0 \\
&\implies \alpha_1 \varphi(v_1) + \cdots + \alpha_n \varphi_n(v_n) = 0.
\end{aligned}$$

Since this last equation holds for all $\varphi \in V'$, it holds in particular for each element of the dual basis $\varphi_1, \ldots, \varphi_n$. That is, for $k = 1, \ldots, n$, we have

$$\alpha_1 \varphi_k(v_1) + \cdots + \alpha_k \varphi_k(v_k) + \cdots + \alpha_n \varphi_k(v_n) = 0 \implies \alpha_k = 0,$$

and therefore $v = 0 \cdot v_1 + \cdots + 0 \cdot v_n = 0$, as desired. Thus $\hat{\cdot}$ is indeed an isomorphism.

We now prove the main result. Suppose $\varphi_1, \ldots, \varphi_n$ is a basis of V', and let $\Phi_1, \ldots, \Phi_n$ be the dual basis in V''. For each Φ_k, let v_k be the inverse of Φ_k under the isomorphism $\hat{\cdot}$. Since the inverse of an isomorphism is an isomorphism, and isomorphisms take bases to bases, $v_1, \ldots, v_n$ is a basis of V. Let us now check that its dual basis is $\varphi_1, \ldots, \varphi_n$. For $j, k = 1, \ldots, n$, we have

$$\begin{aligned}\varphi_j(v_k) &= \widehat{v_k}(\varphi_j) \\ &= \Phi_k(\varphi_j) \\ &= \begin{cases} 1 & \text{if } k = j \\ 0 & \text{otherwise,} \end{cases}\end{aligned}$$

so indeed there exists a basis of V whose dual basis is $\varphi_1, \ldots, \varphi_n$, as was to be shown. ☐

Problem 32

Suppose $T \in \mathcal{L}(V)$ and $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are bases of V. Prove that the following are equivalent:

(a) T is invertible.

(b) The columns of $\mathcal{M}(T)$ are linearly independent in $\mathbb{F}^{n,1}$.

(c) The columns of $\mathcal{M}(T)$ span $\mathbb{F}^{n,1}$.

(d) The rows of $\mathcal{M}(T)$ are linearly independent in $\mathbb{F}^{1,n}$.

(e) The rows of $\mathcal{M}(T)$ span $F^{1,n}$.

Here $\mathcal{M}(T)$ means $\mathcal{M}(T, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$.

Proof. We prove the following: $(a) \iff (b) \iff (c) \iff (e) \iff (d)$.

$(a) \iff (b)$. Suppose T is invertible. That is, for any $w \in V$, there exists a unique $x \in V$ such that $w = Tx$. It follows

$$\begin{aligned}\mathcal{M}(w) &= \mathcal{M}(Tx) \\ &= \mathcal{M}(T)\mathcal{M}(x) \\ &= \mathcal{M}(x)_1\mathcal{M}(T)_{\cdot,1} + \cdots + \mathcal{M}(x)_n\mathcal{M}(T)_{\cdot,n}.\end{aligned}$$

That is, every vector in $\mathbb{F}^{1,n}$ can be exhibited as a unique linear combination of the columns of $\mathcal{M}(T)$. This is true if and only if the columns of $\mathcal{M}(T)$ are linearly independent.

$(b) \iff (c)$. Suppose the columns of $\mathcal{M}(T)$ are linearly independent in $\mathbb{F}^{n,1}$. Since they form a linearly independent list of length $\dim(\mathbb{F}^{n,1})$, they are a basis. But this is true if and only if they span $\mathbb{F}^{n,1}$ as well.

$(c) \iff (e)$. Suppose the columns of $\mathcal{M}(T)$ span $\mathbb{F}^{n,1}$, so that the column rank is n. Since the row rank equals the column rank, so too must the rows of $\mathcal{M}(T)$ span $\mathbb{F}^{1,n}$.

$(e) \iff (d)$. Suppose the rows of $\mathcal{M}(T)$ span $\mathbb{F}^{1,n}$. Since they form a spanning list of length $\dim(\mathbb{F}^{1,n})$, they are a basis. But this is true if and only if they are linearly independent in $\mathbb{F}^{1,n}$ as well. ☐

Problem 33

Suppose m and n are positive integers. Prove that the function that takes A to A^t is a linear map from $F^{m,n}$ to $\mathbb{F}^{n,m}$. Furthermore, prove that this linear map is invertible.

Proof. We first show taking the transpose is linear. So suppose $A, B \in \mathbb{F}^{m,n}$ and let $j = 1, \ldots, n$ and $k = 1, \ldots, m$. It follows

$$\begin{aligned}(A+B)^t_{j,k} &= (A+B)_{k,j} \\ &= A_{k,j} + B_{k,j} \\ &= A^t_{j,k} + B^t_{j,k},\end{aligned}$$

so that taking the transpose is additive. Next, let $\lambda \in \mathbb{F}$. It follows

$$\begin{aligned}(\lambda A)^t_{j,k} &= (\lambda A)_{k,j} \\ &= \lambda A_{k,j} \\ &= \lambda A^t_{j,k},\end{aligned}$$

so that taking the transpose is homogenous. Since it is both additive and homogeneous, it is a linear map. To see that taking the transpose is invertible, note that $(A^t)^t = A$, so that the inverse of the transpose is the transpose itself. $\square$

Problem 34

The ***double dual space*** of V, denoted V'', is defined to be the dual space of V'. In other words, $V'' = (V')'$. Define $\Lambda : V \to V''$ by

$$(\Lambda v)(\varphi) = \varphi(v)$$

for $v \in V$ and $\varphi \in V'$.

(a) Show that Λ is a linear map from V to V''.

(b) Show that if $T \in \mathcal{L}(V)$, then $T'' \circ \Lambda = \Lambda \circ T$, where $T'' = (T')'$.

(c) Show that if V is finite-dimensional, then Λ is an isomorphism from V onto V''.

Proof. We proved (a) and (c) in Problem 31 (where we defined $\hat{\cdot}$ in precisely the same way as Λ). So it only remains to prove (b). So suppose $v \in V$ and $\varphi \in V'$ are arbitrary. Evaluating $T'' \circ \Lambda$, notice

$$\begin{aligned}((T'' \circ \Lambda)(v))(\varphi) &= (T''(\Lambda v))(\varphi) \\ &= (\Lambda v)(T'\varphi) \\ &= (T'\varphi)(v) \\ &= \varphi(Tv),\end{aligned}$$

where the second and fourth equalities follow by definition of the dual map, and the third equality follows by definition of Λ. And evaluating $\Lambda \circ T$, we have

$$
\begin{aligned}
((\Lambda \circ T)(v))(\varphi) &= (\Lambda(Tv))(\varphi) \\
&= \varphi(Tv),
\end{aligned}
$$

so that the two expressions evaluate to the same thing. Since the choice of both v and φ was arbitrary, we have $T'' \circ \Lambda = \Lambda \circ T$, as desired. ☐

Problem 35

Show that $(\mathcal{P}(\mathbb{R}))'$ and $\mathbb{R}^\infty$ are isomorphic.

Proof. For any sequence $\alpha = (\alpha_0, \alpha_1, \dots) \in \mathbb{R}^\infty$, let φ_α be the unique linear functional in $(\mathcal{P}(\mathbb{R}))'$ such that $\varphi_\alpha(X^k) = \alpha_k$ for all $k \in \mathbb{Z}^+$ (note that since the list $1, X, X^2, \dots$ is a basis of $\mathcal{P}(\mathbb{R})$, this description of φ_α is sufficient). We claim

$$
\begin{aligned}
\Phi : \mathbb{R}^\infty &\to (\mathcal{P}(\mathbb{R}))' \\
\alpha &\mapsto \varphi_\alpha
\end{aligned}
$$

is an isomorphism. There are three things to show: that Φ is a linear map, that it's injective, and that it's surjective.

We first show Φ is linear. Suppose $\alpha, \beta \in \mathbb{R}^\infty$. For any $k \in \mathbb{Z}^+$, it follows

$$
\begin{aligned}
(\Phi(\alpha + \beta))(X^k) &= \varphi_{\alpha+\beta}(X^k) \\
&= (\alpha + \beta)_k \\
&= \alpha_k + \beta_k \\
&= \varphi_\alpha(X^k) + \varphi_\beta(X^k) \\
&= (\Phi(\alpha))(X^k) + (\Phi(\beta))(X^k),
\end{aligned}
$$

so that Φ is additive. Next suppose $\lambda \in \mathbb{R}$. Then we have

$$
\begin{aligned}
\Phi(\lambda\alpha)(X^k) &= \varphi_{\lambda\alpha}(X^k) \\
&= (\lambda\alpha)_k \\
&= \lambda\alpha_k \\
&= \lambda\Phi(\alpha),
\end{aligned}
$$

so that Φ is homogenous. Being both additive and homogeneous, Φ is indeed linear.

Next, to see that Φ is injective, suppose $\Phi(\alpha) = 0$ for some $\alpha \in \mathbb{R}^\infty$. Then $\varphi_\alpha(X^k) = \alpha_k = 0$ for all $k \in \mathbb{Z}^+$, and hence $\alpha = 0$. Thus Φ is injective.

Lastly, to see that Φ is surjective, suppose $\varphi \in (\mathcal{P}(\mathbb{R}))'$. Define $\alpha_k = \varphi(X^k)$ for all $k \in \mathbb{Z}^+$ and let $\alpha = (\alpha_0, \alpha_1, \dots)$. By construction, we have $(\Phi(\alpha))(X^k) = \alpha_k$ for all $k \in \mathbb{Z}^+$, and hence $\Phi(\alpha) = \varphi_\alpha$. Thus Φ is surjective.

Since Φ is linear, injective, and surjective, it's an isomorphism, as desired. ☐

Problem 37

Suppose U is a subspace of V. Let $\pi : V \to V/U$ be the usual quotient map. Thus $\pi' \in \mathcal{L}((V/U)', V')$.

(a) Show that π' is injective.

(b) Show that $\operatorname{range} \pi' = U^0$.

(c) Conclude that π' is an isomorphism from $(V/U)'$ onto U^0.

Proof. (a) Let $\varphi \in (V/U)'$, and suppose $\pi'(\varphi) = 0$. Then $(\varphi \circ \pi)(v) = \varphi(v + U) = 0$ for all $v \in V$. This is true only if $\varphi = 0$, and hence π' is indeed injective.

(b) First, suppose $\varphi \in \operatorname{range} \pi'$. Then there exists $\psi \in (V/U)'$ such that $\pi'(\psi) = \varphi$. So for all $u \in U$, we have

$$\begin{aligned} \varphi(u) &= (\pi'(\psi))(u) \\ &= \psi(\pi(u)) \\ &= \psi(u + U) \\ &= \psi(0 + U) \\ &= 0, \end{aligned}$$

and thus $\varphi \in U^0$, showing $\operatorname{range} \pi' \subseteq U^0$. Conversely, suppose $\varphi \in U^0$, so that $\varphi(u) = 0$ for all $u \in U$. Define $\psi \in (V/U)'$ by $\psi(v + U) = \varphi(v)$ for all $v \in V$. Then $(\pi'(\psi))(v) = \psi(\pi(v)) = \psi(v + U) = \varphi(v)$, and so indeed $\varphi \in \operatorname{range} \pi'$, showing $U^0 \subseteq \operatorname{range} \pi'$. Therefore, we have $\operatorname{range} \pi' = U^0$, as desired.

(c) Notice that (b) may be interpreted as saying $\pi' : (V/U)' \to U^0$ is surjective. Since π' was shown to be injective in (a), we conclude π' is an isomorphism from $(V/U)'$ onto U^0, as desired. $\square$

Chapter 4: Polynomials

Linear Algebra Done Right, by Sheldon Axler

Problem 1

Verify all the assertions in 4.5 except the last one.

Proof. Suppose $w, z \in \mathbb{C}$, and let $a, b, c, d \in \mathbb{R}$ be such that $w = a + bi$ and $z = c + di$.

- Notice $z + \overline{z} = (c + di) + (c - di) = 2c = 2\Re(z)$.
- We have $z - \overline{z} = (c + di) - (c - di) = 2di = 2\Im(z)i$.
- Notice $z\,\overline{z} = (c + di)(c - di) = c^2 + d^2 = \left(\sqrt{c^2 + d^2}\right)^2 = |z|^2$.
- We have $\overline{w + z} = \overline{(a + c) + (b + d)i} = (a - bi) + (c - di) = \overline{w} + \overline{z}$. Also, $\overline{wz} = \overline{(ac - bd) + (ad + bc)i} = (ac - bd) - (ad + bc)i$ and $\overline{w}\,\overline{z} = (a - bi)(c - di) = (ac - bd) - (ad + bc)i$, so that $\overline{wz} = \overline{w}\,\overline{z}$.
- Notice $\overline{\overline{z}} = \overline{c - di} = c + di = z$.
- We have $\left|\Re(z)\right| = |c| = \sqrt{c^2} \leq \sqrt{c^2 + d^2} = |z|$, and similarly $|\Im z| = |d| = \sqrt{d^2} \leq \sqrt{c^2 + d^2} = |z|$.
- Notice $|\overline{z}| = |c - di| = \sqrt{c^2 + (-d)^2} = \sqrt{c^2 + d^2} = |z|$.
- We have
$$\begin{aligned}
|wz| &= \left|(ac - bd) + (ad + bc)i\right| \\
&= \sqrt{(ac - bd)^2 + (ad + bc)^2} \\
&= \sqrt{a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2} \\
&= \sqrt{(a^2 + b^2)(c^2 + d^2)} \\
&= \sqrt{a^2 + b^2}\sqrt{c^2 + d^2} \\
&= |w|\,|z|\,,
\end{aligned}$$
as desired. $\square$

Problem 3

Is the set
$$\{0\} \cup \{p \in \mathcal{P}(\mathbb{F}) \mid \deg p \text{ is even}\}$$
a subspace of $\mathcal{P}(\mathbb{F})$?

Proof. Let $E = \{0\} \cup \{p \in \mathcal{P}(\mathbb{F}) \mid \deg p \text{ is even}\}$. Then E is not a subspace of $\mathcal{P}(\mathbb{F})$. To see this, notice $p(x) = x^2 + x \in E$ and $q(x) = -x^2 + x \in E$, but $p + q = 2x \notin E$, so that E is not closed under addition. $\square$

Problem 5

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

$$p(z_j) = w_j$$

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

Proof. Define

$$\begin{aligned} T : \mathcal{P}_m(\mathbb{F}) &\to \mathbb{F}^{m+1} \\ p &\mapsto (p(z_1), \ldots, p(z_{m+1})). \end{aligned}$$

It suffices to show that T is an isomorphism, since injectivity implies uniqueness of such a $p \in \mathcal{P}_m(\mathbb{F})$, and surjectivity implies its existence. So we first show that T is a linear map. Suppose $p, q \in \mathcal{P}_m(\mathbb{F})$. Then

$$\begin{aligned} T(p+q) &= \big((p+q)(z_1), \ldots, (p+q)(z_{m+1})\big) \\ &= \big(p(z_1) + q(z_1), \ldots, p(z_{m+1}) + q(z_{m+1})\big) \\ &= \big(p(z_1), \ldots, p(z_{m+1})\big) + \big(q(z_1), \ldots, q(z_{m+1})\big) \\ &= Tp + Tq, \end{aligned}$$

so that T is additive. Next suppose $\lambda \in \mathbb{F}$. Then

$$\begin{aligned} T(\lambda p) &= \big((\lambda p)(z_1), \ldots, (\lambda p)(z_{m+1})\big) \\ &= \big(\lambda p(z_1), \ldots, \lambda p(z_{m+1})\big) \\ &= \lambda \big(p(z_1), \ldots, p(z_{m+1})\big) \\ &= \lambda(Tp), \end{aligned}$$

so that T is also homogenous. Hence T is a linear map. To see that T is an isomorphism, it's enough to show T is injective. So suppose $Tp = 0$ for some $p \in \mathcal{P}_m(\mathbb{F})$. Then

$$Tp = \big(p(z_1), \ldots, p(z_{m+1})\big) = (0, \ldots, 0),$$

and hence p has $m+1$ zeros. Since it has degree at most m, p must therefore be the zero polynomial, completing the proof. $\square$

Problem 7

Prove that every polynomial of odd degree with real coefficients has a real zero.

Proof. Suppose not. Then there exists some $p \in \mathcal{P}(\mathbb{R})$ of odd degree with no real zeros. By Theorem 4.17, p must be of the form

$$p(x) = c(x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M),$$

where $c, b_1, \ldots, b_M, c_1, \ldots, c_M \in \mathbb{R}$ and $M \in \mathbb{Z}^+$. But then p has even degree, a contradiction. Thus every polynomial of odd degree with real coefficients must indeed have a real zero. □

Problem 9

Suppose $p \in \mathcal{P}(\mathbb{C})$. Define $q : \mathbb{C} \to \mathbb{C}$ by

$$q(z) = p(z)\, \overline{p(\overline{z})}\,.$$

Prove that q is a polynomial with real coefficients.

Proof. Suppose p has degree n. Then there exist $c, \lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that

$$p(z) = c(z_1 - \lambda_1) \cdots (z_n - \lambda_n).$$

Thus we have

$$\begin{aligned} q(z) &= c(z_1 - \lambda_1) \cdots (z_n - \lambda_n)\, \overline{c\,(\overline{z_1} - \lambda_1) \cdots (\overline{z_n} - \lambda_n)} \\ &= c\overline{c}(z_1 - \lambda_1)\left(z_1 - \overline{\lambda_1}\right) \cdots (z_n - \lambda_n)\left(z_n - \overline{\lambda_n}\right) \\ &= |c|^2 \left({z_1}^2 - 2\Re(\lambda_1) z_1 + |\lambda_1|^2\right) \cdots \left({z_n}^2 - 2\Re(\lambda_n) z_n + |\lambda_n|^2\right), \end{aligned}$$

so that $q(z)$ is the product of polynomials with real coefficients. Thus q is itself a polynomial with real coefficients, as was to be shown. □

Problem 11

Suppose $p \in \mathcal{P}(\mathbb{F})$ with $p \neq 0$. Let $U = \{pq \mid q \in \mathcal{P}(\mathbb{F})\}$.

(a) Show that $\dim \mathcal{P}(\mathbb{F})/U = \deg p$

(b) Find a basis of $\mathcal{P}(\mathbb{F})/U$.

Proof. Suppose $\dim p = n$ for some $n \in \mathbb{Z}^+$.

(a) Consider the map

$$\begin{aligned} T : \mathcal{P}(\mathbb{F}) &\mapsto \mathcal{P}_{n-1}(\mathbb{F}) \\ f &\mapsto r(f), \end{aligned}$$

where $r(f)$ is the unique remainder when f is divided by p. We will show that T is linear, $\operatorname{null} T = U$, and $\operatorname{range} T = \mathcal{P}_{n-1}(\mathbb{F})$, so that $V/U \cong \mathcal{P}_{n-1}$.

Since $\mathcal{P}_{n-1}(\mathbb{F}) \cong \mathbb{F}^n$ and $\dim \mathbb{F}^n = n = \deg p$, this gives the desired result.

First we show T is a linear map. To see this, suppose $f, g \in \mathcal{P}(F)$. Then there exist unique $q_1, q_2 \in \mathcal{P}(F)$ such that $f = q_1 p + r(f)$ and $g = q_2 p + r(g)$. But then $f + g = (q_1 + q_2)p + r(f) + r(g)$, and hence $r(f + g) = r(f) + r(g)$. Thus

$$T(f + g) = r(f) + r(g) = T(f) + T(g),$$

and so T is additive. To see that T is also homogenous, suppose $\lambda \in \mathbb{F}$. Then $\lambda f = (\lambda q_1)p + \lambda r(f)$, and since both the quotient and remainder are unique, we must have $\lambda r(f) = r(\lambda f)$. Therefore

$$T(\lambda f) = \lambda r(f) = \lambda T f,$$

and so T is homogeneous. Thus T is a linear map, as claimed.

Next we show $\operatorname{null} T = U$. Suppose $f \in \operatorname{null} T$. Then $Tf = 0$, and hence $r(f) = 0$. That is, there exists $q_1 \in \mathcal{P}(\mathbb{F})$ such that $f = pq_1$, and thus $f \in U$. Conversely, if $g \in U$, then there exists $q_2 \in \mathcal{P}(\mathbb{F})$ such that $g = pq_2$. But then $r(g) = 0$, and hence $Tg = 0$ and $g \in \operatorname{null} T$.

Lastly we show $\operatorname{range} T = \mathcal{P}_{n-1}$. Of course $\operatorname{range} T \subseteq \mathcal{P}_{n-1}$. So suppose $r \in \mathcal{P}_{n-1}$. Then $r = 0p + r$ (where 0 denotes the zero polynomial), and hence $Tr = r$. Thus $\operatorname{range} T = U$.

(b) We claim $1 + U, x + U, \ldots, x^{n-1} + U$ is a basis of $\mathcal{P}(\mathbb{F})/U$. Notice none of these vectors is the zero vector since all elements of U have degree at least n. Clearly the list is linearly independent. Since it has the right length, it's indeed a basis. $\square$

Chapter 5: Eigenvalues, Eigenvectors, and Invariant Subspaces

Linear Algebra Done Right, by Sheldon Axler

A: Invariant Subspaces

Problem 1

Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V.

(a) Prove that if $U \subseteq \operatorname{null} T$, then U is invariant under T.

(b) Prove that if $\operatorname{range} T \subseteq U$, then U is invariant under T.

Proof. (a) Suppose $u \in U$. Since $U \subseteq \operatorname{null} T$, we must have $Tu = 0$. And since $0 \in U$, this implies $Tu \in U$, and so U is indeed invariant under T.

(b) Suppose $u \in U$. Since $Tu \in \operatorname{range} T$ (by definition of $\operatorname{range} T$) and $\operatorname{range} T \subseteq U$, we have $Tu \in U$. Thus U is invariant under T. □

Problem 3

Suppose $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Prove that $\operatorname{range} S$ is invariant under T.

Proof. Suppose $w \in \operatorname{range} S$. Then there exists $v \in V$ such that

$$Sv = w.$$

It follows

$$Tw = TSv = STv,$$

and thus $Tw \in \operatorname{range} S$, so that $\operatorname{range} S$ is indeed invariant under T. □

Problem 5

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

Proof. Let $\mathfrak{U}$ be a collection of subspaces of V invariant under T, and let

$$W = \bigcap_{U \in \mathfrak{U}} U.$$

By Problem 11 of Section 1.C, W is a subspace of V. Assume $u \in W$. Then $u \in U$ for every $U \in \mathfrak{U}$. Since each such U is invariant under T, we have $Tu \in U$ for all $U \in \mathfrak{U}$ as well. This implies $Tu \in W$, and hence W is invariant under T also, as desired. □

Problem 7

Suppose $T \in \mathcal{L}(\mathbb{R}^2)$ is defined by $T(x, y) = (-3y, x)$. Find the eigenvalues of T.

Proof. Suppose $T(x, y) = \lambda(x, y)$, where $(x, y) \in \mathbb{R}^2$ is nonzero and $\lambda \in \mathbb{F}$. Then

$$-3y = \lambda x \tag{1}$$
$$x = \lambda y. \tag{2}$$

Substituting the value for x given by the Equation 2 into Equation 1 gives

$$-3y = \lambda^2 y.$$

Now, y cannot be 0, for otherwise $x = 0$ (by Equation 2), contrary to our assumption that (x, y) is nonzero. Hence $-3 = \lambda^2$. Thus, if $\mathbb{F} = \mathbb{C}$, T two eigenvalues: $\lambda = \pm\sqrt{3}i$. If $\mathbb{F} = \mathbb{R}$, T has no eigenvalues. □

Problem 9

Define $T \in \mathcal{L}\left(\mathbb{F}^3\right)$ by

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

Proof. Suppose $T(z_1, z_2, z_3) = \lambda(z_1, z_2, z_3)$, where $(z_1, z_2, z_3) \in \mathbb{F}^3$ is nonzero and $\lambda \in \mathbb{F}$. Then

$$2z_2 = \lambda z_1 \tag{3}$$
$$0 = \lambda z_2 \tag{4}$$
$$5z_3 = \lambda z_3. \tag{5}$$

First notice that $\lambda = 0$ satisfies the above equations if either z_1 or z_3 is nonzero, and thus 0 is an eigenvalue with corresponding eigenvectors

$$\{(s, 0, t) \mid s, t \in \mathbb{F}, s \text{ and } t \text{ are not both } 0\}.$$

If $\lambda \neq 0$, then we must have $z_2 = 0$ by Equation 4, and hence $z_1 = 0$ by Equation 3. Since $(z_1, z_2, z_3) \neq (0, 0, 0)$, we conclude z_3 must be nonzero. Thus Equation 5 implies $\lambda = 5$ is the only other eigenvalue with corresponding eigenvectors

$$\{(0, 0, t) \mid t \in \mathbb{F} - \{0\}\},$$

and we're done. □

Problem 11

Define $T : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ by $Tp = p'$. Find all eigenvalues and eigenvectors of T.

Proof. Let $p \in \mathcal{P}(\mathbb{R})$ be nonzero and suppose $Tp = \lambda p$ for some $\lambda \in \mathbb{R}$. Note that $\deg p$ must be 0, for otherwise, since $\deg p = \deg(Tp) = \deg p'$, we have a contradiction. Thus the only eigenvalue of T is $\lambda = 0$, and the corresponding eigenvectors are the constant, nonzero polynomials in $\mathcal{P}(\mathbb{R})$. $\square$

Problem 13

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. Prove that there exists $\alpha \in \mathbb{F}$ such that $|\alpha - \lambda| < \frac{1}{1000}$ and $T - \alpha I$ is invertible.

Proof. Suppose not. Then for any $\alpha \in \mathbb{F}$ such that $|\alpha - \lambda| < \frac{1}{1000}$, $T - \alpha I$ is not invertible. But then, by Theorem 5.6, α is an eigenvalue of T. This is a contradiction, since there are infinitely many such α, but T can have at most $\dim V$ eigenvalues by Theorem 5.13. $\square$

Problem 15

Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible.

(a) Prove that T and $S^{-1}TS$ have the same eigenvalues.

(b) What is the relationship between the eigenvectors of T and the eigenvectors of $S^{-1}TS$?

Proof. (a) Suppose $\lambda \in \mathbb{F}$ is an eigenvalue of $S^{-1}TS$. Then there exists a nonzero $v \in V$ such that $(S^{-1}TS)v = \lambda v$. This equation is true if and only if $TSv = \lambda(Sv)$, which is in turn true if and only if $Tw = \lambda w$, where $w = Sv$. Note that since S is invertible, $w \neq 0$. Thus T and $S^{-1}TS$ indeed have the same eigenvalues.

(b) As shown in the proof of (a), $v \in V$ is an eigenvector of $S^{-1}TS$ if and only if Sv is an eigenvector of T. $\square$

Problem 17

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

Proof. Consider the following operator

$$\begin{aligned} T : \mathbb{R}^4 &\to \mathbb{R}^4 \\ (x_1, x_2, x_3, x_4) &\mapsto (-x_4, x_1, x_2, x_3). \end{aligned}$$

We claim T has no real eigenvalues. To see this, suppose $T(x_1, x_2, x_3, x_4) = \lambda(x_1, x_2, x_3, x_4)$ for some $\lambda \in \mathbb{R}$ and nonzero $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$. It follows

$$(-x_4, x_1, x_2, x_3) = \lambda(x_1, x_2, x_3, x_4),$$

and hence

$$-x_4 = \lambda x_1 \tag{6}$$

$$x_1 = \lambda x_2 \tag{7}$$

$$x_2 = \lambda x_3 \tag{8}$$

$$x_3 = \lambda x_4. \tag{9}$$

This implies $-x_4 = \lambda^4 x_4$. Notice λ cannot be 0, for otherwise (x_1, x_2, x_3, x_4) is the zero vector, a contradiction. Hence we must have $x_4 = 0$. But then Equation 6 implies $x_1 = 0$, which in turn implies $x_2 = 0$ by Equation 7, and which thus implies $x_3 = 0$ by Equation 8. But now we have that (x_1, x_2, x_3, x_4) is the zero vector, another contradiction. So we conclude T indeed has no real eigenvalues, as claimed. □

Problem 19

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

Proof. Suppose $T(x_1, \ldots, x_n) = \lambda(x_1, \ldots, x_n)$ for some $\lambda \in \mathbb{F}$ and some nonzero $(x_1, \ldots, x_n) \in \mathbb{F}^n$. Then

$$(x_1 + \cdots + x_n, \ldots, x_1 + \cdots + x_n) = \lambda(x_1, \ldots, x_n).$$

It follows

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

Thus, our first eigenvalue is $\lambda = 0$ with corresponding eigenvectors

$$\{(x_1, \ldots, x_n) \in \mathbb{F}^n - \{0\} \mid x_1 + \cdots + x_n = 0\}.$$

Next, if $\lambda \neq 0$, notice the equations above imply $\lambda x_1 = \cdots = \lambda x_n$, and thus $x_1 = \cdots = x_n$. Denote the common value of the x_k's by y. Then any of the above equations is now equivalent to $ny = \lambda y$. Thus our second eigenvalue is $\lambda = n$ with corresponding eigenvectors

$$\{(x_1, \dots, x_n) \in \mathbb{F}^n - \{0\} \mid x_1 = \cdots = x_n\},$$

and we're done. ☐

Problem 21

Suppose $T \in \mathcal{L}(V)$ is invertible.

(a) Suppose $\lambda \in \mathbb{F}$ with $\lambda \neq 0$. Prove that λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.

(b) Prove that T and T^{-1} have the same eigenvectors.

Proof. (a) By definition, $\lambda \neq 0$ is an eigenvalue of T if and only if there exists $v \in V - \{0\}$ such that $Tv = \lambda v$. Since T is invertible, this is true if and only if $v = T^{-1}(\lambda v)$, which is itself true if and only if (after simplification) $\left(\frac{1}{\lambda}\right) v = T^{-1}v$. Thus λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}, as was to be shown.

(b) First notice that $\lambda = 0$ cannot be an eigenvalue of T or T^{-1} since they are both injective. Now, suppose v is an eigenvector of T corresponding to $\lambda \neq 0$. By the proof of (a), v is an eigenvector of T^{-1} corresponding to $\frac{1}{\lambda}$. Thus all eigenvectors of T are eigenvectors of T^{-1}. Now, reversing the roles of T and T^{-1} and applying the same argument yields the reverse inclusion, completing the proof. ☐

Problem 23

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

Proof. Let $\lambda \in \mathbb{F}$ be an eigenvalue of ST and $v \in V - \{0\}$ be a corresponding eigenvector, so that $STv = \lambda v$. First, if $Tv \neq 0$, it follows

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

so that λ is an eigenvalue of TS. Next, if $Tv = 0$, then we must have $\lambda = 0$ (since $STv = \lambda v$). Moreover, T is not invertible (since $v \neq 0$). Thus TS is not invertible (by Problem 9 of Chapter 3.D). Since TS is not invertible, there exists a nonzero $w \in V$ such that $TSw = 0$, and hence $\lambda = 0$ is an eigenvalue of TS as well.

Since λ is an eigenvalue of TS in both cases, we conclude that every eigenvalue of ST is also an eigenvalue of TS. Reversing the roles of S and T and applying the same argument yields the reverse inclusion, completing the proof. □

Problem 25

Suppose $T \in \mathcal{L}(V)$ and u, v are eigenvectors of T such that $u+v$ is also an eigenvector of T. Prove that u and v are eigenvectors of T corresponding to the same eigenvalue.

Proof. Suppose λ_1 is the eigenvalue associated to u, λ_2 is the eigenvalue associated to v, and λ_3 is the eigenvalue associated to $u+v$, so that

$$Tu = \lambda_1 u \tag{10}$$

$$Tv = \lambda_2 v \tag{11}$$

$$T(u+v) = \lambda_3(u+v). \tag{12}$$

It follows that

$$Tu + Tv = \lambda_1 u + \lambda_2 v,$$

and hence, by Equation 12, we have

$$\lambda_3 u + \lambda_3 v = \lambda_1 u + \lambda_2 v.$$

Thus

$$(\lambda_1 - \lambda_3)u + (\lambda_2 - \lambda_3)v = 0.$$

Since u and v are both eigenvectors of T, they are linearly independent. Thus $\lambda_1 = \lambda_3$ and $\lambda_2 = \lambda_3$, and hence $\lambda_1 = \lambda_2 = \lambda_3$, showing that u and v indeed correspond to the same eigenvalue. □

Problem 26

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

Proof. By hypothesis, for all $v \in V$ there exists $\lambda_v \in \mathbb{F}$ such that $Tv = \lambda_v v$ (where λ_0 can be any nonzero element of $\mathbb{F}$, since $T0 = 0$). We claim λ_v is independent of the choice of v for $v \in V - \{0\}$, hence $Tv = \lambda v$ for all $v \in V$ (including $v = 0$) and some $\lambda \in \mathbb{F}$, and thus $T = \lambda I$.

So suppose $w, z \in V - \{0\}$ are arbitrary. We want to show $\lambda_w = \lambda_z$. If w and

z are linearly dependent, then there exists $\alpha \in \mathbb{F}$ such that $w = \alpha z$. It follows

$$\begin{aligned}
\lambda_w w &= Tw \\
&= T(\alpha z) \\
&= \alpha T z \\
&= \alpha \lambda_z z \\
&= \lambda_z(\alpha z) \\
&= \lambda_z w.
\end{aligned}$$

Since $w \neq 0$, this implies $\lambda_w = \lambda_z$. Next suppose w and z are linearly independent. Then we have

$$\begin{aligned}
\lambda_{w+z}(w+z) &= T(w+z) \\
&= Tw + Tz \\
&= \lambda_w w + \lambda_z z,
\end{aligned}$$

and hence

$$(\lambda_{w+z} - \lambda_w)w + (\lambda_{w+z} - \lambda_z)z = 0.$$

Since w and z are assumed to be linearly independent, we have $\lambda_{w+z} = \lambda_w$ and $\lambda_{w+z} = \lambda_z$, and hence again we have $\lambda_w = \lambda_z$, completing the proof. □

Problem 27

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ 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.

Proof. Suppose not. Then by the contrapositive of Problem 26, there exists some nonzero $v \in V$ which is not an eigenvector of T. Thus the list v, Tv is linearly independent, and, assuming $\dim V = n$, we may extend it to some basis $v, Tv, u_1, \ldots, u_{n-2}$ of V. Let $U = \operatorname{span}(v, u_1, \ldots, u_{n-2})$. Since $\dim U = \dim V - 1$, U must be invariant under T. But this is a contradiction, since $Tv \notin U$. Thus T must be a scalar multiple of the identity operator, as desired. □

Problem 29

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

Proof. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T, and let $v_1, \ldots, v_m$ be corresponding eigenvectors. For $k \in \{1, \ldots, m\}$, if $\lambda_k \neq 0$, then

$$T\left(\frac{1}{\lambda_k} v_k\right) = v_k.$$

Since at most one of the $\lambda_1, \ldots, \lambda_m$ can be 0, at least $m-1$ of our eigenvectors are in range T. Thus, since lists of distinct eigenvectors are linearly independent by Theorem 5.10, we have

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

which implies $m \leq k + 1$, as desired. □

Problem 31

Suppose V is finite-dimensional and $v_1, \ldots, v_m$ is a list of vectors in V. Prove that $v_1, \ldots, v_m$ is linearly independent if and only if there exists $T \in \mathcal{L}(V)$ such that $v_1, \ldots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues.

Proof. ($\Leftarrow$) If $T \in \mathcal{L}(V)$ is such that $v_1, \ldots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues, then $v_1, \ldots, v_m$ is linearly independent by Theorem 5.10.

($\Rightarrow$) Suppose $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Define $T \in \mathcal{L}(V)$ by $Tv_k = kv_k$ for $k = 1, \ldots, m$. The existence (and uniqueness) of T is guaranteed by Theorem 3.5, and clearly $v_1, \ldots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues. □

Problem 33

Suppose $T \in \mathcal{L}(V)$. Prove that $T/(\operatorname{range} T) = 0$.

Proof. Let $v + \operatorname{range} T \in V/(\operatorname{range} T)$. Then

$$\begin{aligned}(T/(\operatorname{range} T))(v + \operatorname{range} T) &= Tv + \operatorname{range} T \\ &= 0 + \operatorname{range} T.\end{aligned}$$

Thus $T/(\operatorname{range} T)$ is indeed the zero map, as was to be shown. □

Problem 35

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is invariant under T. Prove that each eigenvalue of T/U is an eigenvalue of T.

Proof. Suppose $\lambda \in \mathbb{F}$ is an eigenvalue of T/U. Then there exists some nonzero $v + U \in V/U$ such that

$$(T/U)(v + U) = \lambda(v + U),$$

which implies

$$Tv + U = \lambda v + U,$$

and hence $Tv - \lambda v \in U$. If λ is an eigenvalue of $T|_U$, we're done. So suppose not. Then, since V is finite-dimensional, Theorem 5.6 tells us $T|_U - \lambda I : U \to U$ is invertible. Hence there exists some $u \in U$ such that $(T|_U - \lambda I)(u) = Tv - \lambda v$, and thus

$$Tu - \lambda u = Tv - \lambda v.$$

Simplifying, we have $T(u - v) = \lambda(u - v)$. Since $v \notin U$ by assumption, this implies $u - v \neq 0$ and hence λ is an eigenvalue of T, completing the proof. $\square$

B: Eigenvectors and Upper-Triangular Matrices

Problem 1

Suppose $T \in \mathcal{L}(V)$ and there exists a positive integer n such that $T^n = 0$.

(a) Prove that $I - T$ is invertible and that

$$(I - T)^{-1} = I + T + \cdots + T^{n-1}.$$

(b) Explain how you would guess the formula above.

Proof. (a) We will show that $S := I + T + \cdots + T^{n-1}$ is both a left and right inverse of $I - T$. Suppose $v \in V$. We have

$$\begin{aligned}
(I - T)Sv &= (I - T)\left(v + Tv + \cdots + T^{n-1}v\right) \\
&= \left(v + Tv + \cdots + T^{n-1}v\right) - T\left(v + Tv + \cdots + T^{n-1}v\right) \\
&= v + \left(Tv + \cdots + T^{n-1}v - Tv - T^2v - \cdots - T^{n-1}\right) + T^n v \\
&= v
\end{aligned}$$

and

$$\begin{aligned}
S(I - T)v &= \left(v + Tv + \cdots + T^{n-1}v\right)(I - T) \\
&= \left(v + Tv + \cdots + T^{n-1}v\right) - \left(Tv + T^2v + \cdots + T^n v\right) \\
&= v + \left(Tv + \cdots + T^{n-1}v - Tv - T^2v - \cdots - T^{n-1}\right) + T^n v \\
&= v.
\end{aligned}$$

Thus $I - T$ is indeed invertible, and S is its inverse.

(b) Recall the power series expansion for $(1 - x)^{-1}$ when $|x| < 1$:

$$(1 - x)^{-1} = \sum_{k=0}^{\infty} x^k.$$

Substituting T for x and supposing $T^k = 0$ for $k \geq n$, we have the formula from (a). $\square$

Problem 3

Suppose $T \in \mathcal{L}(V)$ and $T^2 = I$ and -1 is not an eigenvalue of T. Prove that $T = I$.

Proof. Since -1 is not an eigenvalue of T, Theorem 5.6 implies $T + I$ is invertible. Hence for all $w \in V$, there exists $v \in V$ such that $(T + I)v = w$. Thus

$$Tv + v = w. \tag{13}$$

Since $T^2 = I$, applying T to both sides yields

$$v + Tv = Tw. \tag{14}$$

Combining Equations 13 and 14, we see $Tw = w$. Therefore it must be that $T = I$, as was to be shown. $\square$

Problem 4

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

Proof. First notice $0 = P^2 - P = P(P - I)$, hence $(P - I)v \in \operatorname{null} P$ for all $v \in V$. Next notice we can write $v = Pv - (P - I)v$. Since of course $Pv \in \operatorname{range} P$, this yields

$$V = \operatorname{null} P + \operatorname{range} P.$$

To see this sum is direct, suppose $w \in \operatorname{null} P \cap \operatorname{range} P$. Then $Pw = 0$ (since $w \in \operatorname{null} P$) and there exists $u \in V$ such that $w = Pu$ (since $w \in \operatorname{range} P$). Combining these two equations with the hypothesis that $P^2 = P$, we now have

$$w = Pu = P^2 u = P(Pu) = Pw = 0,$$

and thus the sum is indeed direct, as desired. $\square$

Problem 5

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

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

Proof. For $k \in \mathbb{Z}^+$, notice

$$\left(STS^{-1}\right)^k = ST^kS^{-1}.$$

Since $p \in \mathcal{P}(\mathbb{F})$, there exist $n \in \mathbb{Z}^+$ and $\alpha_0, \ldots, \alpha_n \in \mathbb{F}$ such that

$$p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n.$$

It follows

$$p(T) = \alpha_0 I + \alpha_1 T + \cdots + \alpha_n T^n,$$

and hence

$$Sp(T) = \alpha_0 S + \alpha_1 ST + \cdots + \alpha_n ST^n,$$

and thus we have

$$Sp(T)S^{-1} = \alpha_0 I + \alpha_1 STS^{-1} + \cdots + \alpha_n ST^nS^{-1} = p\left(STS^{-1}\right),$$

as was to be shown. □

Problem 7

Suppose $T \in \mathcal{L}(V)$. Prove that 9 is an eigenvalue of T^2 if and only if 3 or -3 is an eigenvalue of T.

Proof. ($\Leftarrow$) Suppose 3 or -3 is an eigenvalue of T. Then there exists a nonzero $v \in V$ such that either

$$Tv = 3v \quad \text{or} \quad Tv = -3v.$$

In the former case, we have $T^2v = 3Tv = 9v$, and in the latter we have $T^2 = -3Tv = 9v$. In both cases, 9 is an eigenvalue of T^2.

($\Rightarrow$) If 9 is an eigenvalue of T^2, then there exists a nonzero $w \in V$ such that $T^2w = 9w$. Hence $T^2 - 9I$ is not invertible, whereby $(T - 3I)(T + 3I)$ is not invertible. Thus, by Problem 9 of Section 3.D, either $T - 3I$ or $T + 3I$ is not invertible. This implies either 3 or -3 is an eigenvalue of T, as desired. □

Problem 9

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$ with $v \neq 0$. Let p be a nonzero polynomial of smallest degree such that $p(T)v = 0$. Prove that every zero of p is an eigenvalue of T.

Proof. Suppose $\lambda \in \mathbb{F}$ is a zero of p. Then there exists $q \in \mathcal{P}(\mathbb{F})$ with $\deg q = \deg p - 1$ such that

$$p(X) = (X - \lambda)q(X).$$

Then, since $p(T)v = 0$ by hypothesis, we have

$$(T - \lambda I)q(T)v = 0.$$

Since $\deg q < \deg p$, $q(T)v \neq 0$, and hence λ is indeed an eigenvalue of T. □

Problem 11

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

Proof. ($\Rightarrow$) Suppose α is an eigenvalue of $p(T)$. Then $p(T) - \alpha I$ is not injective. By the Fundamental Theorem of Algebra, there exist $c, \lambda_1, \ldots, \lambda_m \in \mathbb{C}$ such that

$$p(z) - \alpha = c(z - \lambda_1) \ldots (z - \lambda_m).$$

If $c = 0$, then $p(z) = \alpha$ and p is constant, a contradiction. So we must have $c \neq 0$. By the above equation, we have

$$p(T) - \alpha I = c(T - \lambda_1 I) \ldots (T - \lambda_m I).$$

Since $p(T) - \alpha I$ is not injective, there exists $j \in \{1, \ldots, m\}$ such that $T - \lambda_j I$ is not injective. In other words, λ_j is an eigenvalue of T. Moreover, notice $p(\lambda_j) - \alpha = 0$, and hence $\alpha = p(\lambda_j)$, as desired.

($\Leftarrow$) Suppose $\alpha = p(\lambda)$ for some eigenvalue λ of T. Let $v \in V - \{0\}$ be a corresponding eigenvector, and let $\alpha_0, \ldots, \alpha_n \in \mathbb{C}$ be such that

$$p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n.$$

Notice $T^k v = \lambda^k v$ for any $k \in \mathbb{Z}^+$. It follows

$$\alpha_0 + \alpha_1 \lambda + \cdots + \alpha_n \lambda^n = \alpha,$$

and hence

$$\begin{aligned} p(T)v &= \alpha_0 v + \alpha_1 T v + \cdots + \alpha_n T^n v \\ &= \alpha_0 v + \alpha_1 \lambda v + \cdots + \alpha_n \lambda^n v \\ &= (\alpha_0 + \alpha_1 \lambda + \cdots + \alpha_n \lambda^n)\, v \\ &= \alpha v. \end{aligned}$$

Thus α is an eigenvalue of $p(T)$, completing the proof. $\square$

Problem 13

Suppose W is a complex vector space and $T \in \mathcal{L}(W)$ has no eigenvalues. Prove that every subspace of W invariant under T is either $\{0\}$ or infinite-dimensional.

Proof. Suppose $U \subseteq W$ is invariant under T. If $U = \{0\}$ the result holds, so suppose otherwise. Now, if U were finite-dimensional, then $T|_U$ would have an eigenvalue by Theorem 5.21. Thus T would have an eigenvalue as well, a contradiction. So U must be infinite-dimensional. $\square$

Problem 14

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.

Proof. Consider the operator

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto (y, x). \end{aligned}$$

With respect to the standard basis, we have

$$\mathcal{M}(T) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Clearly T is invertible (it's its own inverse), but its matrix with respect to the standard basis has only 0's on the diagonal. □

Problem 15

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.

Proof. Consider the operator

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto (x + y, x + y). \end{aligned}$$

With respect to the standard basis, we have

$$\mathcal{M}(T) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Notice that T is not invertible, since $T(0, 0) = (0, 0) = T(-1, 1)$, and yet its matrix with respect to the standard basis has only nonzero numbers on the diagonal. Combining this result with Problem 14, we see that Theorem 5.30 fails without the hypothesis that an upper-triangular matrix is under consideration. □

Problem 17

Rewrite the proof of 5.21 using the linear map that sends $p \in \mathcal{P}_{n^2}(\mathbb{C})$ to $p(T) \in \mathcal{L}(V)$ (and use 3.23).

Proof. We will show that every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue. Suppose V is a complex vector space with dimension $n > 0$ and $T \in \mathcal{L}(V)$. Consider the linear map

$$\begin{aligned} M : \mathcal{P}_{n^2}(\mathbb{C}) &\to \mathcal{L}(V) \\ p &\mapsto p(T). \end{aligned}$$

Since $\dim\big(\mathcal{P}_{n^2}(\mathbb{C})\big) = n^2+1$ but $\dim\big(\mathcal{L}(V)\big) = n^2$, M is not injective by Theorem 3.23. Thus there exists a nonzero $p \in \mathcal{P}_{n^2}(\mathbb{C})$ such that $Mp = p(T) = 0$. By the Fundamental Theorem of Algebra, p has a factorization

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

where c is a nonzero complex number, each λ_j is in $\mathbb{C}$, and the equation holds for all $z \in \mathbb{C}$. Now choose any $v \in V - \{0\}$. It follows

$$\begin{aligned} 0 &= p(T)v \\ &= c(T - \lambda_1 I)\dots(T - \lambda_m I)v. \end{aligned}$$

Since $v \neq 0$, $T - \lambda_j$ is not injective for at least one j. In other words, T has an eigenvalue. □

Problem 19

Suppose V is finite-dimensional with $\dim V > 1$ and $T \in \mathcal{L}(V)$. Prove that

$$\{p(T) \mid p \in \mathcal{P}(\mathbb{F})\} \neq \mathcal{L}(V).$$

Proof. Let $\mathcal{U} = \{p(T) \mid p \in \mathcal{P}(\mathbb{F})\}$, and suppose by way of contradiction that $\mathcal{U} = \mathcal{L}(V)$. Let $p \in \mathcal{P}(\mathbb{F})$, and let $\alpha_0, \dots, \alpha_n \in \mathbb{F}$ be such that $p(z) = \alpha_0 + \alpha_1 z \dots + \alpha_n z^n$ for all $z \in \mathbb{F}$. Notice

$$\begin{aligned} Tp(T) &= T\left(\alpha_0 I + \alpha_1 T + \dots + \alpha_n T^n\right) \\ &= \alpha_0 T + \alpha_1 T^2 + \dots + \alpha_n T^{n+1} \\ &= \left(\alpha_0 I + \alpha_1 T + \dots + \alpha_n T^n\right) T \\ &= p(T)T, \end{aligned}$$

so that T commutes with all elements of U. By Problem 16 of Chapter 3.D, this implies $T = \lambda I$ for some $\lambda \in \mathbb{F}$. It follows

$$\begin{aligned} \mathcal{U} &= \{p(T) \mid p \in \mathcal{P}(\mathbb{F})\} \\ &= \{p(\lambda I) \mid p \in \mathcal{P}(\mathbb{F})\} \\ &= \left\{\alpha_0 I + \alpha_1(\lambda I) + \dots + \alpha_1(\lambda I)^n \mid \alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{F} \text{ and } n \in \mathbb{Z}^+\right\} \\ &= \{p(\lambda)I \mid p \in \mathcal{P}(\mathbb{F})\} \\ &= \{\alpha I \mid \alpha \in \mathbb{F}\}, \end{aligned}$$

and thus $\dim \mathcal{U} = 1$. Since $\dim \mathcal{L}(V) = (\dim V)^2$ and $\dim V > 1$ by hypothesis, we have $\dim \mathcal{L}(V) > 1$, a contradiction. Thus our assumption that $\mathcal{U} = \mathcal{L}(V)$ must be false, as was to be shown. □

C: Eigenspaces and Diagonal Matrices

Problem 1

Suppose $T \in \mathcal{L}(V)$ is diagonalizable. Prove that $V = \operatorname{null} T \oplus \operatorname{range} T$.

Proof. By Theorem 5.41, there exists a basis $v_1, \dots, v_n$ of V consisting of eigenvectors of T. Let $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ be corresponding eigenvalues, respectively. Let m denote the number of eigenvalues λ_j such that $\lambda_j = 0$. After relabeling, we may assume $\lambda_j = 0$ for $j = 1, \dots, m$ and $\lambda_j \neq 0$ for $j = m+1, \dots, n$. It follows

$$V = \operatorname{span}(v_1, \dots, v_m) \oplus \operatorname{span}(v_{m+1}, \dots, v_n).$$

Note that if $m = 0$, the left hand term in the direct sum becomes the span of the empty list, which is defined to be $\{0\}$. We claim $\operatorname{null} T = \operatorname{span}(v_1, \dots, v_m)$ and $\operatorname{range} T = \operatorname{span}(v_{m+1}, \dots, v_n)$, which provides the desired result.

First we prove $\operatorname{null} T = \operatorname{span}(v_1, \dots, v_m)$. This result is trivially true if $m = 0$, so suppose otherwise. Since each of $v_1, \dots, v_m$ is an eigenvector corresponding to 0, we have $v_1, \dots, v_m \in E(0, T)$, and hence $\operatorname{span}(v_1, \dots, v_m) \subseteq E(0, T) = \operatorname{null} T$. For the reverse inclusion, suppose $v \in \operatorname{null} T$. Let $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ be such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. It follows

$$\begin{aligned} 0 &= Tv \\ &= \alpha_1 T v_1 + \cdots + \alpha_n T v_n \\ &= \alpha_{m+1} T v_{m+1} + \cdots + \alpha_n T v_n \\ &= (\alpha_{m+1}\lambda_{m+1}) v_{m+1} + \cdots + (\alpha_n \lambda_n) v_n. \end{aligned}$$

Since $\lambda_{m+1}, \dots, \lambda_n$ are all nonzero, the linear independence of $v_{m+1}, \dots, v_n$ implies $\alpha_{m+1} = \cdots = \alpha_n = 0$. Thus $v = \alpha_1 v_1 + \cdots + \alpha_m v_m$, and indeed $v \in \operatorname{span}\{v_1, \dots, v_m\}$. We conclude $\operatorname{null} T = \operatorname{span}(v_1, \dots, v_m)$.

Now we prove $\operatorname{range} T = \operatorname{span}(v_{m+1}, \dots, v_n)$. Clearly we have $v_{m+1}, \dots, v_n \in \operatorname{range} T$, since $T(v_k/\lambda_k) = v_k$ for $k = m+1, \dots, n$, and hence $\operatorname{span}(v_{m+1}, \dots, v_n) \subseteq \operatorname{range} T$. For the reverse inclusion, suppose $w \in \operatorname{range} T$. Then there exists $z \in V$ such that $Tz = w$. Let $\beta_1, \dots, \beta_n \in \mathbb{F}$ be such that $z = \beta_1 v_1 + \cdots + \beta_n v_n$. It follows

$$\begin{aligned} w &= Tz \\ &= \beta_1 T v_1 + \cdots + \beta_n T v_n \\ &= (\beta_{m+1}\lambda_{m+1}) v_{m+1} + \cdots + (\beta_n \lambda_n) v_n. \end{aligned}$$

Thus $w \in \operatorname{span}(v_{m+1}, \dots, v_n)$, and we conclude $\operatorname{range} T = \operatorname{span}(v_{m+1}, \dots, v_n)$, completing the proof of our claim. $\square$

Problem 3

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent:

(a) $V = \operatorname{null} T \oplus \operatorname{range} T$.

(b) $V = \operatorname{null} T + \operatorname{range} T$.

(c) $\operatorname{null} T \cap \operatorname{range} T = \{0\}$.

Proof. Let $N = \operatorname{null} T$ and $R = \operatorname{range} T$.

($a \Rightarrow b$) If $V = N \oplus R$, then $V = N + R$ by the definition of direct sum.

($b \Rightarrow c$) Suppose $V = N + R$. By Theorem 2.43, we know

$$\dim(N + R) = \dim N + \dim R - \dim(N \cap R), \tag{15}$$

and by hypothesis, the LHS of Equation 15 equals $\dim V$. Hence we have

$$\dim V = \dim N + \dim R - \dim(N \cap R). \tag{16}$$

Now, by the Fundamental Theorem of Linear Maps, we have

$$\dim V = \dim N + \dim R. \tag{17}$$

Combining Equations 16 and 17 yields $\dim(N \cap R) = 0$, and hence $N \cap R = \{0\}$.

($c \Rightarrow a$) Suppose $N \cap R = \{0\}$. Again by Theorem 2.43, we have

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

By hypothesis, $\dim(N \cap R) = 0$. Thus

$$\dim(N + R) = \dim N + \dim R. \tag{18}$$

By another application of the Fundamental Theorem of Linear Maps, the RHS of Equation 18 equals $\dim V$. Hence we have $\dim V = \dim(N + R)$, and therefore $V = N + R$. Since $N \cap R = \{0\}$ by hypothesis, this sum is direct. $\square$

Problem 5

Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

for every $\lambda \in \mathbb{C}$.

Proof. ($\Rightarrow$) Suppose T is diagonalizable. Then there exists a basis such that $\mathcal{M}(T)$ is diagonal. Letting $\lambda \in \mathbb{C}$, it follows

$$\begin{aligned}\mathcal{M}(T - \lambda I) &= \mathcal{M}(T) - \lambda\mathcal{M}(I)\\ &= \mathcal{M}(T) - \lambda I,\end{aligned}$$

where we abuse notation and use I to denote both the identity operator on V and the identity matrix in $\mathbb{F}^{\dim V, \dim V}$. Since λI is diagonal, so too is $\mathcal{M}(T) - \lambda I$, and hence $T - \lambda I$ is diagonalizable. The desired result now follows by Problem 1.

($\Leftarrow$) Conversely, suppose

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

for every $\lambda \in \mathbb{C}$. We induct on $n = \dim V$. If $n = 1$, the result clearly holds, since every matrix in $\mathbb{F}^{1,1}$ is diagonal. Now assume $n \in \mathbb{Z}^+$ and that the assertion holds for all vector spaces of dimension $k < n$. Let $\lambda_1 \in \mathbb{C}$ be an eigenvalue of T (such an eigenvalue must exist by Theorem 5.21). By hypothesis, we have

$$V = E(\lambda_1, T) \oplus \operatorname{range}(T - \lambda_1 I). \tag{19}$$

Let $R = \operatorname{range}(T - \lambda_1 I)$. We claim

$$R = \operatorname{null}(T|_R - \lambda I) \oplus \operatorname{range}(T|_R - \lambda I)$$

for all $\lambda \in \mathbb{C}$. By Problem 3c, it suffices to show $\operatorname{null}(T|_R - \lambda I) \cap \operatorname{range}(T|_R - \lambda I) = \{0\}$. Notice

$$\operatorname{null}(T|_R - \lambda I) \subseteq \operatorname{null}(T - \lambda I) \quad \text{and} \quad \operatorname{range}(T|_R - \lambda I) \subseteq \operatorname{range}(T - \lambda I).$$

It follows

$$\operatorname{null}(T|_R - \lambda I) \cap \operatorname{range}(T|_R - \lambda I) \subseteq \operatorname{null}(T - \lambda I) \cap \operatorname{range}(T - \lambda I) = \{0\},$$

proving our claim. Now, let $v_1, \ldots, v_k$ be a basis of $E(\lambda_1, T)$. Since $T|_R$ is diagonalizable, R has a basis of eigenvectors by Theorem 5.41. Call them $v_{k+1}, \ldots, v_n$. By Equation 19, the list $v_1, \ldots, v_n$ is a basis of V consisting of eigenvectors of T. By another application of Theorem 5.41, this implies T is diagonalizable, as desired. $\square$

Problem 7

Suppose $T \in \mathcal{L}(V)$ has a diagonal matrix A with respect to some basis of V and that $\lambda \in \mathbb{F}$. Prove that λ appears on the diagonal of A precisely $\dim E(\lambda, T)$ times.

Proof. Let $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$ be the distinct eigenvalues of T, let $v_1, \ldots, v_n$ be a basis consisting of eigenvectors of T (such a basis is guaranteed by Theorem 5.41), and let $A = \mathcal{M}(T)$ with respect to this basis. Denote by s_k the number of our basis vectors contained in $E(\lambda_k, T)$ for $k \in \{1, \ldots, m\}$, so that the eigenvalue λ_k appears on the diagonal of A exactly s_k times. We will show $s_k = \dim E(\lambda_k, T)$.

Since any subset of the basis contained in $E(\lambda_k, T)$ is of course linearly independent, we first note $s_k \leq \dim E(\lambda_k, T)$. So we have

$$\begin{aligned} s_1 + \cdots + s_m &\leq \dim E(\lambda_1, T) + \cdots + \dim(\lambda_m, T) \\ &= n. \end{aligned}$$

Since $E(\lambda_i, T) \cap E(\lambda_j, T) = \{0\}$ for $i \neq j$, each element of our basis is contained in at most one $E(\lambda_k, T)$. Hence the LHS of the equation above equals n as well, and the inequality is in fact an equality. This implies

$$s_1 - \dim(E_1, T) = (\dim E(\lambda_2, T) - s_2) + \cdots + (\dim E(\lambda_m, T) - s_m).$$

Each term in parentheses on the RHS is nonnegative, and hence $s_1 - \dim(E_1, T) \geq 0$, which implies $s_1 \geq \dim(E_1, T)$. Since we've already shown $s_1 \leq \dim(E_1, T)$, we conclude $s_1 = \dim(E_1, T)$. An analogous argument shows $s_\ell = E(\lambda_\ell, T)$ for all $\ell \in \{2, \ldots, m\}$.

Therefore, if $\lambda \in \mathbb{C}$ is an eigenvalue of T, then λ indeed appears on the diagonal of A precisely $\dim E(\lambda, T)$ times. And if λ is not an eigenvalue of T, then it appears on the diagonal zero times, which also equals $\dim E(\lambda, T)$. In both cases, the desired result holds, completing the proof. $\square$

Problem 9

Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that $E(\lambda, T) = E\left(\frac{1}{\lambda}, T^{-1}\right)$ for every $\lambda \in \mathbb{F}$ with $\lambda \neq 0$.

Proof. Let $\lambda \in \mathbb{F} - \{0\}$, and suppose $v \in E(\lambda, T)$. Then

$$\begin{aligned} Tv = \lambda v &\implies v = T^{-1}(\lambda v) \\ &\implies \frac{1}{\lambda} v = T^{-1} v \\ &\implies v \in E\left(\frac{1}{\lambda}, T^{-1}\right), \end{aligned}$$

and thus $E(\lambda, T) \subseteq E\left(\frac{1}{\lambda}, T^{-1}\right)$. Conversely, suppose $w \in E\left(\frac{1}{\lambda}, T^{-1}\right)$. It follows

$$\begin{aligned} T^{-1} w = \frac{1}{\lambda} w &\implies w = T\left(\frac{1}{\lambda} w\right) \\ &\implies \lambda w = Tw \\ &\implies w \in E(\lambda, T), \end{aligned}$$

and so $E\left(\frac{1}{\lambda}, T^{-1}\right) \subseteq E(\lambda, T)$. Therefore, we conclude $E(\lambda, T) = E\left(\frac{1}{\lambda}, T^{-1}\right)$, as was to be shown. $\square$

Problem 11

Verify the assertion in Example 5.40.

Proof. Define $T \in \mathcal{L}(\mathbb{R}^2)$ by

$$T(x,y) = (41x + 7y, -20x + 74y).$$

Example 5.40 asserts that T is diagonalizable, because the matrix of T with respect to the basis $(1,4),(7,5)$ is

$$\begin{bmatrix} 69 & 0 \\ 0 & 46 \end{bmatrix}.$$

To see this, first notice

$$\begin{aligned} T(1,4) &= (69, 276) \\ &= 69 \cdot (1,4) + 0 \cdot (7,5), \end{aligned}$$

and so the first column of the matrix is correct. Next notice

$$\begin{aligned} T(7,5) &= (322, 230) \\ &= 0 \cdot (1,4) + 46 \cdot (7,5), \end{aligned}$$

and so the second column of the matrix is correct as well. □

Problem 12

Suppose $R, T \in \mathcal{L}(\mathbb{F}^3)$ each have $2, 6, 7$ as eigenvalues. Prove that there exists an invertible operator $S \in \mathcal{L}(\mathbb{F}^3)$ such that $R = S^{-1}TS$.

Proof. Since R and T each have 3 eigenvalues and $\dim \mathbb{F}^3 = 3$, they are both diagonalizable by Theorem 5.44. Letting $\lambda_1 = 2, \lambda_2 = 6$, and $\lambda_3 = 7$, there exist (again by Theorem 5.44) bases v_1, v_2, v_3 and w_1, w_2, w_3 of $\mathbb{F}^3$ such that

$$Rv_k = \lambda_k v_k \quad \text{and} \quad Tw_k = \lambda_k w_k$$

for $k = 1, 2, 3$. Define the operator $S \in \mathcal{L}(\mathbb{F}^3)$ by its behavior on the v_k's

$$Sv_k = w_k.$$

Since S takes one basis to another basis, it's invertible. Now notice

$$\begin{aligned} S^{-1}TSv_k &= S^{-1}Tw_k \\ &= S^{-1}(\lambda_k w_k) \\ &= \lambda_k S^{-1} w_k \\ &= \lambda_k v_k \\ &= Rv_k, \end{aligned}$$

and thus $R = S^{-1}TS$, as desired. □

Problem 13

Find $R, T \in \mathcal{L}(\mathbb{F}^4)$ such that R and T each have $2, 6, 7$ as eigenvalues, R and T have no other eigenvalues, and there does not exist an invertible operator $S \in \mathcal{L}(\mathbb{F}^4)$ such that $R = S^{-1}TS$.

Proof. For $x = (x_1, x_2, x_3, x_4) \in \mathbb{F}^4$, define $R, T \in \mathcal{L}(\mathbb{F}^4)$ by

$$Rx = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad \text{and} \quad Tx = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

By Theorem 5.32, R and T each have precisely $2, 6, 7$ as eigenvalues and no others. We claim T is diagonalizable, and we will use this fact to derive a contradiction from which the result will follow. To see this, first notice

$$T - 2I = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$

Since $T - 2I$ is in echelon form and has three pivots, $\dim \operatorname{range}(T - 2I) = 3$, and thus $\dim E(2, T) = \dim \operatorname{null}(T - 2I) = 1$. Similarly, we have

$$T - 6I = \begin{bmatrix} -4 & 1 & 1 & 1 \\ 0 & -4 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

so that $T - 6I$ has three pivots as well and hence $\dim E(6, T) = 1$. Lastly, notice

$$T - 7I = \begin{bmatrix} -5 & 1 & 1 & 1 \\ 0 & -5 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and $T - 6I$ also has three pivots and so $\dim E(7, T) = 1$. Since $\dim E(2, T) + \dim E(6, T) + \dim E(7, T) < \dim \mathbb{F}^4$, T is not diagonalizable by Theorem 5.41.

Now, by way of contradiction, suppose there exists an invertible $S \in \mathcal{L}(\mathbb{F}^4)$ such that $R = S^{-1}TS$. Then the list $Se_1, \ldots, Se_4$ is a basis of $\mathbb{F}^4$. Notice

$$\begin{aligned} T(Se_1) &= S(Re_1) \\ &= S(2e_1) \\ &= 2Se_1, \end{aligned}$$

and similarly we have

$$T(Se_2) = 2Se_2, \quad T(Se_3) = 6Se_3, \quad \text{and} \quad T(Se_4) = 7Se_4.$$

Thus $\mathcal{M}\left(T,(Se_1,\ldots,Se_4)\right)$ is diagonal, a contradiction. Therefore, no such S exists, and R and T are operators of the desired form. $\square$

Problem 15

Suppose $T \in \mathcal{L}(\mathbb{C}^3)$ is such that 6 and 7 are eigenvalues of T. Furthermore, suppose T does not have a diagonal matrix with respect to any basis of $\mathbb{C}^3$. Prove that there exists $(x,y,z) \in \mathbb{C}^3$ such that $T(x,y,z) = (17+8x, \sqrt{5}+8y, 2\pi+8z)$.

Proof. By hypothesis, T is not diagonalizable. Hence by Theorem 5.44, 6 and 7 are the only eigenvalues of T. In particular, 8 is not an eigenvalue. Thus

$$\dim E(8,T) = \dim \operatorname{null}(T-8I) = 0,$$

and hence $T-8I$ is surjective. So there exists $(x,y,z) \in \mathbb{C}^3$ such that $(T-8I)(x,y,z) = (17,\sqrt{5},2\pi)$. It follows

$$\begin{aligned} T(x,y,z) &= (17,\sqrt{5},2\pi) + 8(x,y,z) \\ &= (17+8x, \sqrt{5}+8y, 2\pi+8z), \end{aligned}$$

as was to be shown. $\square$

Chapter 6: Inner Product Spaces

Linear Algebra Done Right, by Sheldon Axler

A: Inner Products and Norms

Problem 1

Show that the function that takes $\big((x_1, x_2), (y_1, y_2)\big) \in \mathbb{R}^2 \times \mathbb{R}^2$ to $|x_1 y_1| + |x_2 y_2|$ is not an inner product on $\mathbb{R}^2$.

Proof. Suppose it were. First notice

$$\begin{aligned}\langle (1,1) + (-1,-1), (1,1) \rangle &= \langle (0,0), (1,1) \rangle \\ &= |0 \cdot 1| + |0 \cdot 1| \\ &= 0.\end{aligned}$$

Next, since inner products are additive in the first slot, we also have

$$\begin{aligned}\langle (1,1) + (-1,-1), (1,1) \rangle &= \langle (1,1), (1,1) \rangle + \langle (-1,-1), (1,1) \rangle \\ &= |1 \cdot 1| + |1 \cdot 1| + |(-1) \cdot 1| + |(-1) \cdot 1| \\ &= 4.\end{aligned}$$

But this implies $0 = 4$, a contradiction. Hence we must conclude that the function does not in fact define an inner product. $\square$

Problem 3

Suppose $\mathbb{F} = \mathbb{R}$ and $V \neq \{0\}$. Replace the positivity condition (which states that $\langle v, v \rangle \geq 0$ for all $v \in V$) in the definition of an inner product (6.3) with the condition that $\langle v, v \rangle > 0$ for some $v \in V$. Show that this change in the definition does not change the set of functions from $V \times V$ to $\mathbb{R}$ that are inner products on V.

Proof. Let V be a nontrivial vector space over $\mathbb{R}$, let A denote the set of functions $V \times V \to \mathbb{R}$ that are inner products on V in the standard definition, and let B denote the set of functions $V \times V \to \mathbb{R}$ under the modified definition. We will show $A = B$.

Suppose $\langle \cdot, \cdot \rangle_1 \in A$. Since $V \neq \{0\}$, there exists $v \in V - \{0\}$. Then $\langle v, v \rangle_1 > 0$, and so $\langle \cdot, \cdot \rangle_1 \in B$. Thus $A \subseteq B$.

Conversely, suppose $\langle \cdot, \cdot \rangle_2 \in B$. Then there exists some $v' \in V$ such that

$\langle v', v' \rangle_2 > 0$. Suppose by way of contradiction there exists $u \in V$ is such that $\langle u, u \rangle_2 < 0$. Define $w = \alpha u + (1-\alpha)v'$ for $\alpha \in \mathbb{R}$. It follows

$$\begin{aligned}\langle w, w \rangle_2 &= \langle \alpha u + (1-\alpha)v', \alpha u + (1-\alpha)v' \rangle_2 \\ &= \langle \alpha u, \alpha u \rangle_2 + 2\langle \alpha u, (1-\alpha)v' \rangle_2 + \langle (1-\alpha)v', (1-\alpha)v' \rangle_2 \\ &= \alpha^2 \langle u, u \rangle_2 + 2\alpha(1-\alpha)\langle u, v' \rangle_2 + (1-\alpha)^2 \langle v', v' \rangle_2.\end{aligned}$$

Notice the final expression is a polynomial in the indeterminate α, call it p. Since $p(0) = \langle v', v' \rangle_2 > 0$ and $p(1) = \langle u, u \rangle_2 < 0$, by Bolzano's theorem there exists $\alpha_0 \in (0, 1)$ such that $p(\alpha_0) = 0$. That is, if $w = \alpha_0 u + (1-\alpha_0)v'$, then $\langle w, w \rangle_2 = 0$. In particular, notice $\alpha_0 \neq 0$, for otherwise $w = v'$, a contradiction since $\langle v', v' \rangle_2 > 0$. Now, since $\langle w, w \rangle_2 = 0$ iff $w = 0$ (by the definiteness condition of an inner product), it follows

$$u = \frac{\alpha_0 - 1}{\alpha_0} v.$$

Letting $t = \frac{\alpha_0 - 1}{\alpha_0}$, we now have

$$\begin{aligned}\langle u, u \rangle_2 &= \langle tv', tv' \rangle_2 \\ &= t^2 \langle v', v' \rangle_2 \\ &> 0,\end{aligned}$$

where the inequality follows since $t \in (-1, 0)$ and $\langle v', v' \rangle_2 > 0$. This contradicts our assumption that $\langle u, u \rangle_2 < 0$, and so we have $\langle \cdot, \cdot \rangle_2 \in A$. Therefore, $B \subseteq A$. Since we've already shown $A \subseteq B$, this implies $A = B$, as desired. $\square$

Problem 5

Let V be finite-dimensional. Suppose $T \in \mathcal{L}(V)$ is such that $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that $T - \sqrt{2}I$ is invertible.

Proof. Let $v \in \operatorname{null}(T - \sqrt{2}I)$, and suppose by way of contradiction that $v \neq 0$. Then

$$\begin{aligned}Tv - \sqrt{2}v = 0 &\implies Tv = \sqrt{2}v \\ &\implies \|\sqrt{2}v\| \leq \|v\| \\ &\implies \sqrt{2} \cdot \|v\| \leq \|v\| \\ &\implies \sqrt{2} \leq 1,\end{aligned}$$

a contradiction. Hence $v = 0$ and $\operatorname{null}(T - \sqrt{2}I) = \{0\}$, so that $T - \sqrt{2}I$ is injective. Since V is finite-dimensional, this implies $T - \sqrt{2}I$ is invertible, as desired. $\square$

Problem 7

Suppose $u, v \in V$. Prove that $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbb{R}$ if and only if $\|u\| = \|v\|$.

Proof. ($\Rightarrow$) Suppose $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbb{R}$. Then this equation holds when $a = 1$ and $b = 0$. But then we must have $\|u\| = \|v\|$, as desired.

($\Leftarrow$) Conversely, suppose $\|u\| = \|v\|$. Let $a, b \in \mathbb{R}$ be arbitrary, and notice

$$\begin{aligned} \|au + bv\| &= \langle au + bv, au + bv \rangle \\ &= \langle au, au \rangle + \langle au, bv \rangle + \langle bv, au \rangle + \langle bv, bv \rangle \\ &= a^2 \|u\|^2 + ab \left(\langle u, v \rangle + \langle v, u \rangle \right) + b^2 \|v\|^2. \end{aligned} \tag{1}$$

Also, we have

$$\begin{aligned} \|bu + av\| &= \langle bu + av, bu + av \rangle \\ &= \langle bu, bu \rangle + \langle bu, av \rangle + \langle av, bu \rangle + \langle av, av \rangle \\ &= b^2 \|u\|^2 + ab \left(\langle u, v \rangle + \langle v, u \rangle \right) + a^2 \|v\|^2. \end{aligned} \tag{2}$$

Since $\|u\| = \|v\|$, (1) equals (2), and hence $\|au + bv\| = \|bu + av\|$. Since a, b were arbitrary, the result follows. $\square$

Problem 9

Suppose $u, v \in V$ and $\|u\| \leq 1$ and $\|v\| \leq 1$. Prove that

$$\sqrt{1 - \|u\|^2} \sqrt{1 - \|v\|^2} \leq 1 - |\langle u, v \rangle|.$$

Proof. By the Cauchy-Schwarz Inequality, we have $|\langle u, v \rangle| \leq \|u\| \|v\|$. Since $\|u\| \leq 1$ and $\|v\| \leq 1$, this implies

$$0 \leq 1 - \|u\| \|v\| \leq 1 - |\langle u, v \rangle|,$$

and hence it's enough to show

$$\sqrt{1 - \|u\|^2} \sqrt{1 - \|v\|^2} \leq 1 - \|u\| \|v\|.$$

Squaring both sides, it suffices to prove

$$\left(1 - \|u\|^2\right) \left(1 - \|v\|^2\right) \leq \left(1 - \|u\| \|v\|\right)^2. \tag{3}$$

Notice

$$\begin{aligned} \left(1 - \|u\| \|v\|\right)^2 - \left(1 - \|u\|^2\right) \left(1 - \|v\|^2\right) &= \|u\|^2 - 2\|u\| \|v\| + \|v\|^2 \\ &= \left(\|u\| - \|v\|\right)^2 \\ &\geq 0, \end{aligned}$$

and hence inequality (3) holds, completing the proof. $\square$

Problem 11

Prove that
$$16 \le (a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)$$
for all positive numbers a, b, c, d.

Proof. Define
$$x = \left(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}\right) \quad \text{and} \quad y = \left(\sqrt{\frac{1}{a}}, \sqrt{\frac{1}{b}}, \sqrt{\frac{1}{c}}, \sqrt{\frac{1}{d}}\right).$$
Then the Cauchy-Schwarz Inequality implies
$$\begin{aligned}(a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right) &\ge \left(\sqrt{a}\sqrt{\frac{1}{a}}+\sqrt{b}\sqrt{\frac{1}{b}}+\sqrt{c}\sqrt{\frac{1}{c}}+\sqrt{d}\sqrt{\frac{1}{d}}\right)^2 \\ &= (1+1+1+1)^2 \\ &= 16,\end{aligned}$$
as desired. □

Problem 13

Suppose u, v are nonzero vectors in $\mathbb{R}^2$. Prove that
$$\langle u, v\rangle = \|u\|\|v\|\cos\theta,$$
where θ is the angle between u and v (thinking of u and v as arrows with initial point at the origin).

Proof. Let A denote the line segment from the origin to u, let B denote the line segment from the origin to v, and let C denote the line segment from v to u. Then A has length $\|u\|$, B has length $\|v\|$ and C has length $\|u-v\|$. Letting θ denote the angle between A and B, by the Law of Cosines we have
$$C^2 = A^2 + B^2 - 2BC\cos\theta,$$
or equivalently
$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos\theta.$$
It follows
$$\begin{aligned}2\|u\|\|v\|\cos\theta &= \|u\|^2 + \|v\|^2 - \|u-v\|^2 \\ &= \langle u, u\rangle + \langle v, v\rangle - \langle u-v, u-v\rangle \\ &= \langle u, u\rangle + \langle v, v\rangle - \big(\langle u, u\rangle - 2\langle u, v\rangle + \langle v, v\rangle\big) \\ &= 2\langle u, v\rangle.\end{aligned}$$
Dividing both sides by 2 gives the desired result. □

Problem 15

Prove that
$$\left(\sum_{j=1}^{n} a_j b_j\right)^2 \leq \left(\sum_{j=1}^{n} j{a_j}^2\right)\left(\sum_{j=1}^{n} \frac{{b_j}^2}{j}\right)$$
for all real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$.

Proof. Let
$$u = \left(a_1, \sqrt{2}a_2, \ldots, \sqrt{n}a_n\right) \quad \text{and} \quad v = \left(b_1, \frac{1}{\sqrt{2}}b_2, \ldots, \frac{1}{\sqrt{n}}b_n\right).$$
Since $\langle u, v\rangle = \sum_{k=1}^{n} a_k b_k$, the Cauchy-Schwarz Inequality yields
$$\begin{aligned}(a_1b_1 + \cdots + a_nb_n)^2 &\leq \|u\|^2\|v\|^2 \\ &= \left({a_1}^2 + 2{a_2}^2 + \cdots + n{a_n}^2\right)\left({b_1}^2 + \frac{{b_2}^2}{2} + \cdots + \frac{{b_n}^2}{n}\right),\end{aligned}$$
as desired. □

Problem 17

Prove or disprove: there is an inner product on $\mathbb{R}^2$ such that the associated norm is given by
$$\|(x, y)\| = \max\{|x|, |y|\}$$
for all $(x, y) \in \mathbb{R}^2$.

Proof. Suppose such an inner product existed. Then by the Parallelogram Equality, it follows
$$\|(1,0) + (0,1)\|^2 + \|(1,0) - (0,1)\|^2 = 2\left(\|(1,0)\|^2 + \|(0,1)\|^2\right).$$
After simplification, this implies $2 = 4$, a contradiction. Hence no such inner product exists. □

Problem 19

Suppose V is a real inner product space. Prove that
$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$$
for all $u, v \in V$.

Proof. Suppose V is a real inner product space and let $u, v \in V$. It follows

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

as desired. $\square$

Problem 20

Suppose V is a complex inner product space. Prove that

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

for all $u, v \in V$.

Proof. Notice we have

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

and

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

Also, we have

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

and

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

Thus it follows

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

Dividing both sides by 4 yields the desired result. $\square$

Problem 23

Suppose $V_1, \dots, V_m$ are inner product spaces. Show that the equation

$$\langle (u_1, \dots, u_m), (v_1, \dots, v_m)\rangle = \langle u_1, v_1\rangle + \cdots + \langle u_m, v_m\rangle$$

defines an inner product on $V_1 \times \cdots \times V_m$.

Proof. We prove that this definition satisfies each property of an inner product in turn.
Positivity: Let $(v_1, \dots, v_m) \in V_1 \times \dots V_m$. Since $\langle v_k, v_k\rangle$ is an inner product on V_k for $k = 1, \dots, m$, we have $\langle v_k, v_k\rangle \geq 0$. Thus

$$\langle (v_1, \dots, v_m), (v_1, \dots, v_m)\rangle = \langle v_1, v_1\rangle + \cdots + \langle v_m, v_m\rangle \geq 0.$$

Definiteness: First suppose $\langle (v_1, \dots, v_m), (v_1, \dots, v_m)\rangle = 0$ for $(v_1, \dots, v_m) \in V_1 \times \cdots \times V_m$. Then

$$\langle v_1, v_1\rangle + \cdots + \langle v_m, v_m\rangle = 0.$$

By positivity of each inner product on V_k (for $k = 1, \dots, m$), we must have $\langle v_k, v_k\rangle \geq 0$. Thus the equation above holds only if $\langle v_k, v_k\rangle = 0$ for each k, which is true iff $v_k = 0$ (by definiteness of the inner product on V_k). Hence $(v_1, \dots, v_m) = (0, \dots, 0)$. Conversely, suppose $(v_1, \dots, v_m) = (0, \dots, 0)$. Then

$$\begin{aligned}
\langle (v_1, \dots, v_m), (v_1, \dots, v_m)\rangle &= \langle v_1, v_1\rangle + \cdots + \langle v_m, v_m\rangle \\
&= \langle 0, 0\rangle + \cdots + \langle 0, 0\rangle \\
&= 0 + \cdots + 0 \\
&= 0,
\end{aligned}$$

where the third equality follows from definiteness of the inner product on each V_k, respectively.
Additivity in first slot: Let

$$(u_1, \dots, u_m), (v_1, \dots, v_m), (w_1, \dots, w_m) \in V_1 \times \cdots \times V_m.$$

It follows

$$\begin{aligned}
\langle (u_1, \dots, u_m) + &(v_1, \dots, v_m)), (w_1, \dots, w_m)\rangle \\
&= \langle (u_1 + v_1, \dots, u_m + v_m), (w_1, \dots, w_m)\rangle \\
&= \langle u_1 + v_1, w_1\rangle + \cdots + \langle u_m + v_m, w_m\rangle \\
&= \langle u_1, w_1\rangle + \langle v_1, w_1\rangle + \cdots + \langle u_m, w_m\rangle + \langle v_m, w_m\rangle \\
&= \langle (u_1, \dots, u_m), (w_1, \dots, w_m)\rangle + \langle (v_1, \dots, v_m), (w_1, \dots, w_m)\rangle,
\end{aligned}$$

where the third equality follows from additivity in the first slot of each inner product on V_k, respectively.
Homogeneity in the first slot: Let $\lambda \in \mathbb{F}$ and

$$(u_1, \ldots, u_m), (v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m.$$

It follows

$$\begin{aligned}
\langle \lambda(u_1, \ldots, u_m), (v_1, \ldots, v_m) \rangle &= \langle (\lambda u_1, \ldots, \lambda u_m), (v_1, \ldots, v_m) \rangle \\
&= \langle \lambda u_1, v_1 \rangle + \cdots + \langle \lambda u_m, v_m \rangle \\
&= \lambda \langle u_1, v_1 \rangle + \cdots + \lambda \langle u_m, v_m \rangle \\
&= \lambda (\langle u_1, v_1 \rangle + \cdots + \langle u_m, v_m \rangle) \\
&= \lambda \langle (u_1, \ldots, u_m), (v_1, \ldots, v_m) \rangle,
\end{aligned}$$

where the third equality follows from homogeneity in the first slot of each inner product on V_k, respectively.
Conjugate symmetry: Again let

$$(u_1, \ldots, u_m), (v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m.$$

It follows

$$\begin{aligned}
\langle (u_1, \ldots, u_m), (v_1, \ldots, v_m) \rangle &= \langle u_1, v_1 \rangle + \cdots + \langle u_m, v_m \rangle \\
&= \overline{\langle v_1, u_1 \rangle} + \cdots + \overline{\langle v_m, u_m \rangle} \\
&= \overline{\langle u_1, v_1 \rangle + \cdots + \langle u_m, v_m \rangle} \\
&= \overline{\langle (v_1, \ldots, v_m), (u_1, \ldots, u_m) \rangle},
\end{aligned}$$

where the second equality follows from conjugate symmetry of each inner product on V_k, respectively. $\square$

Problem 24

Suppose $S \in \mathcal{L}(V)$ is an injective operator on V. Define $\langle \cdot, \cdot \rangle_1$ by

$$\langle u, v \rangle_1 = \langle Su, Sv \rangle$$

for $u, v \in V$. Show that $\langle \cdot, \cdot \rangle_1$ is an inner product on V.

Proof. We prove that this definition satisfies each property of an inner product in turn.
Positivity: Let $v \in V$. Then $\langle v, v \rangle_1 = \langle Sv, Sv \rangle \geq 0$.
Definiteness: Suppose $\langle v, v \rangle = 0$ for some $v \in V$. This is true iff $\langle Sv, Sv \rangle = 0$ (by definition) which is true iff $Sv = 0$ (by definiteness of $\langle \cdot, \cdot \rangle$), which is true iff

$v = 0$ (since S is injective).
Additivity in first slot: Let $u, v, w \in V$. Then

$$\begin{aligned}\langle u+v, w\rangle_1 &= \langle S(u+v), Sw\rangle \\ &= \langle Su + Sv, Sw\rangle \\ &= \langle Su, Sw\rangle + \langle Sv, Sw\rangle \\ &= \langle u, w\rangle_1 + \langle v, w\rangle_1.\end{aligned}$$

Homogeneity in first slot: Let $\lambda \in \mathbb{F}$ and $u, v \in V$. Then

$$\begin{aligned}\langle \lambda u, v\rangle_1 &= \langle S(\lambda u), Sv\rangle \\ &= \langle \lambda Su, Sv\rangle \\ &= \lambda\langle Su, Sv\rangle \\ &= \lambda\langle u, v\rangle_1.\end{aligned}$$

Conjugate symmetry Let $u, v \in V$. Then

$$\begin{aligned}\langle u, v\rangle_1 &= \langle Su, Sv\rangle \\ &= \overline{\langle Sv, Su\rangle} \\ &= \overline{\langle v, u\rangle_1}.\end{aligned}$$

□

Problem 25

Suppose $S \in \mathcal{L}(V)$ is not injective. Define $\langle\cdot,\cdot\rangle_1$ as in the exercise above. Explain why $\langle\cdot,\cdot\rangle_1$ is not an inner product on V.

Proof. If S is not injective, then $\langle\cdot,\cdot\rangle_1$ fails the definiteness requirement in the definition of an inner product. In particular, there exists $v \neq 0$ such that $Sv = 0$. Hence $\langle v, v\rangle_1 = \langle Sv, Sv\rangle = 0$ for a nonzero v. □

Problem 27

Suppose $u, v, w \in V$. Prove that

$$\left\|w - \frac{1}{2}(u+v)\right\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}.$$

Proof. We have

$$\begin{aligned}
\left\| w - \frac{1}{2}(u+v) \right\|^2 &= \left\| \left(\frac{w-u}{2} \right) + \left(\frac{w-v}{2} \right) \right\|^2 \\
&= 2\left\| \frac{w-u}{2} \right\|^2 + 2\left\| \frac{w-v}{2} \right\|^2 - \left\| \left(\frac{w-u}{2} \right) - \left(\frac{w-v}{2} \right) \right\|^2 \\
&= \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \left\| \frac{-u+v}{2} \right\|^2 \\
&= \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4},
\end{aligned}$$

where the second equality follows by the Parallelogram Equality. □

The next problem requires some extra work to prove. We first include a definition and prove a theorem.

Definition. *Suppose* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *are norms on vector space* V*. We say* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *are* equivalent *if there exist* $0 < C_1 \leq C_2$ *such that*

$$C_1\|v\|_1 \leq \|v\|_2 \leq C_2\|v\|_1$$

for all $v \in V$*.*

Theorem. *Any two norms on a finite-dimensional vector space are equivalent.*

Proof. Let V be finite-dimensional with basis $e_1, \ldots, e_n$. It suffices to prove that every norm on V is equivalent to the ℓ_1-style norm $\|\cdot\|_1$ defined by

$$\|v\|_1 = |\alpha_1| + \cdots + |\alpha_n|$$

for all $v = \alpha_1 e_1 + \cdots + \alpha_n e_n \in V$.

Let $\|\cdot\|$ be a norm on V. We wish to show $C_1\|v\|_1 \leq \|v\| \leq C_2\|v\|_1$ for all $v \in V$ and some choice of C_1, C_2. Since this is trivially true for $v = 0$, we need only consider $v \neq 0$, in which case we have

$$C_1 \leq \|u\| \leq C_2, \tag{*}$$

where $u = v/\|v\|_1$. Thus it suffices to consider only vectors $v \in V$ such that $\|v\|_1 = 1$.

We will now show that $\|\cdot\|$ is continuous under $\|\cdot\|_1$ and apply the Extreme Value Theorem to deduce the desired result. So let $\epsilon > 0$ and define $M = \max\{\|e_1\|, \ldots, \|e_n\|\}$ and

$$\delta = \frac{\epsilon}{M}.$$

It follows that if $u, v \in V$ are such that $\|u - v\|_1 < \delta$, then

$$\begin{aligned}
\big| \|u\| - \|v\| \big| &\leq \|u - v\| \\
&\leq M\|u - v\|_1 \\
&\leq M\delta \\
&= \epsilon,
\end{aligned}$$

and $\|\cdot\|$ is indeed continuous under the topology induced by $\|\cdot\|_1$. Let $\mathcal{S} = \{u \in V \mid \|u\|_1 = 1\}$ (the unit sphere with respect to $\|\cdot\|_1$). Since $\mathcal{S}$ is compact and $\|\cdot\|$ is continuous on it, by the Extreme Value Theorem we may define

$$C_1 = \min_{u \in \mathcal{S}} \|u\| \quad \text{and} \quad C_2 = \max_{u \in \mathcal{S}} \|u\|.$$

But now C_1 and C_2 satisfy (*), completing the proof. $\square$

Problem 29

For $u, v \in V$, define $d(u, v) = \|u - v\|$.

(a) Show that d is a metric on V.

(b) Show that if V is finite-dimensional, then d is a complete metric on V (meaning that every Cauchy sequence converges).

(c) Show that every finite-dimensional subspace of V is a closed subset of V (with respect to the metric d).

Proof. (a) We show that d satisfies each property of the definition of a metric in turn.

Identity of indiscernibles: Let $u, v \in V$. It follows

$$\begin{aligned} d(u,v) = 0 &\iff \sqrt{\langle u - v, u - v\rangle} = 0 \\ &\iff \langle u - v, u - v, =\rangle 0 \\ &\iff u - v = 0 \\ &\iff u = v. \end{aligned}$$

Symmetry: Let $u, v \in V$. We have

$$\begin{aligned} d(u,v) &= \|u - v\| \\ &= \big\|(-1)(u - v)\big\| \\ &= \|v - u\| \\ &= d(v, u). \end{aligned}$$

Triangle inequality: Let $u, v, w \in V$. Notice

$$\begin{aligned} d(u,v) + d(v,w) &= \|u - v\| + \|v - w\| \\ &\leq \big\|(u - v) + (v - w)\big\| \\ &= \|u, w\| \\ &= d(u, w). \end{aligned}$$

(b) Suppose V is a p-dimensional vector space with basis $e_1, \ldots, e_p$. Assume $\{v_k\}_{k=1}^{\infty}$ is Cauchy. Then for $\epsilon > 0$, there exists $N \in \mathbb{Z}^+$ such that

$\|v_m - v_n\| < \epsilon$ whenever $m, n > N$. Given any v_i in our Cauchy sequence, we adopt the notation that $\alpha_{i,1}, \ldots, \alpha_{i,p} \in \mathbb{F}$ are always defined such that

$$v_i = \alpha_{i,1}e_1 + \cdots + \alpha_{i,p}e_p.$$

By our previous theorem, $\|\cdot\|$ is equivalent to $\|\cdot\|_1$ (where $\|\cdot\|_1$ is defined in that theorem's proof). Thus there exists some $c > 0$ such that, whenever $m, n > N$, we have

$$c\|v_m - v_n\|_1 \leq \|v_m - v_n\| < \epsilon,$$

and hence

$$c\left(\sum_{i=1}^{p} \left|\alpha_{m,i} - \alpha_{n,i}\right|\right) < \epsilon.$$

This implies that $\{\alpha_{k,i}\}_{k=1}^{\infty}$ is Cauchy in $\mathbb{R}$ for each $i = 1, \ldots, p$. Since $\mathbb{R}$ is complete, these sequences converge. So let $\alpha_i = \lim_{k\to\infty} \alpha_{k,i}$ for each i, and define $v = \alpha_1 e_1 + \cdots + \alpha_p e_p$. It follows

$$\begin{aligned} \left\|v_j - v\right\| &= \left\|(\alpha_{j,1} - \beta_1)e_1 + \cdots + (\alpha_{j,p} - \beta_p)e_p\right\| \\ &\leq \left|\alpha_{j,1} - \alpha_1\right|\left\|e_1\right\| + \cdots + \left|\alpha_{j,p} - \alpha_p\right|\left\|e_p\right\|. \end{aligned}$$

Since $\alpha_{j,i} \to \alpha_i$ for $i = 1, \ldots, p$, the RHS can be made arbitrarily small by choosing sufficiently large $M \in \mathbb{Z}^+$ and considering $j > M$. Thus $\{v_k\}_{k=1}^{\infty}$ converges to v, and V is indeed complete with respect to $\|\cdot\|$.

(c) Suppose U is a finite-dimensional subspace of V, and suppose $\{u_k\}_{k=1}^{\infty} \subseteq U$ is Cauchy. By (b), $\lim_{k\to\infty} u_k \in U$, hence U contains all its limit points. Thus U is closed. □

Problem 31

Use inner products to prove Apollonius's Identity: In a triangle with sides of length a, b, and c, let d be the length of the line segment from the midpoint of the side of length c to the opposite vertex. Then

$$a^2 + b^2 = \frac{1}{2}c^2 + 2d^2.$$

Proof. Consider a triangle formed by vectors $v, w \in \mathbb{R}^2$ and the origin such that $\|w\| = a$, $\|v\| = c$, and $\|w - v\| = b$. The identity follows by applying Problem 27 with $u = 0$. □

B: Orthonormal Bases

Problem 1

(a) Suppose $\theta \in \mathbb{R}$. Show that $(\cos\theta, \sin\theta)$, $(-\sin\theta, \cos\theta)$ and $(\cos\theta, \sin\theta)$, $(\sin\theta, -\cos\theta)$ are orthonormal bases of $\mathbb{R}^2$.

(b) Show that each orthonormal basis of $\mathbb{R}^2$ is of the form given by one of the two possibilities of part (a).

Proof. (a) Notice

$$\langle(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\rangle = -\sin\theta\cos\theta + \sin\theta\cos\theta = 0$$

and

$$\langle(\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)\rangle = \sin\theta\cos\theta - \sin\theta\cos\theta = 0,$$

hence both lists are orthonormal. Clearly the three distinct vectors contained in the two lists all have norm 1 (following from the identity $\cos^2\theta + \sin^2\theta = 1$). Since both lists have length 2, by Theorem 6.28 both lists are orthonormal bases.

(b) Suppose e_1, e_2 is an orthonormal basis of $\mathbb{R}^2$. Since $\|e_1\| = \|e_2\| = 1$, there exist $\theta, \varphi \in [0, 2\pi)$ such that

$$e_1 = (\cos\theta, \sin\theta) \quad \text{and} \quad e_2 = (\cos\varphi, \sin\varphi).$$

Next, since $\langle e_1, e_2\rangle = 0$, we have

$$\cos\theta\cos\varphi + \sin\theta\sin\varphi = 0.$$

Since $\cos\theta\cos\varphi = \frac{1}{2}(\cos(\theta+\varphi) + \cos(\theta-\varphi))$ and $\sin\theta\sin\varphi = \cos(\theta-\varphi) - \cos(\theta+\varphi)$, the above implies

$$\cos(\theta - \varphi) = 0$$

and thus $\varphi = \theta + \frac{3\pi}{2} - n\pi$, for $n \in \mathbb{Z}$. Since $\theta, \varphi \in [0, 2\pi)$, this implies $\varphi = \theta \pm \frac{\pi}{2}$. If $\varphi = \theta + \frac{\pi}{2}$, then

$$\begin{aligned} e_2 &= \left(\cos\left(\theta + \frac{\pi}{2}\right), \sin\left(\theta + \frac{\pi}{2}\right)\right) \\ &= (-\sin\theta, \cos\theta), \end{aligned}$$

and if $\varphi = \theta - \frac{\pi}{2}$, then

$$\begin{aligned} e_2 &= \left(\cos\left(\theta - \frac{\pi}{2}\right), \sin\left(\theta - \frac{\pi}{2}\right)\right) \\ &= (\sin\theta, -\cos\theta). \end{aligned}$$

Thus all orthonormal bases of $\mathbb{R}^2$ have one of the two forms from (a). $\square$

Problem 3

Suppose $T \in \mathcal{L}(\mathbb{R}^3)$ has an upper-triangular matrix with respect to the basis $(1,0,0)$, $(1,1,1)$, $(1,1,2)$. Find an orthonormal basis of $\mathbb{R}^3$ (use the usual inner product on $\mathbb{R}^3$) with respect to which T has an upper-triangular matrix.

Proof. Let $v_1 = (1,0,0), v_2 = (1,1,1)$, and $v_3 = (1,1,2)$. By the proof of 6.37, T has an upper-triangular matrix with respect to the the basis e_1, e_2, e_3 generated by applying the Gram-Schmidt Procedure to v_1, v_2, v_3. Since $\|v_1\| = 1$, $e_1 = v_1$. Next, we have

$$\begin{aligned} e_2 &= \frac{v_2 - \langle v_2, e_1\rangle e_1}{\big\|v_2 - \langle v_2, e_1\rangle e_1\big\|} \\ &= \frac{(1,1,1) - \langle (1,1,1),(1,0,0)\rangle (1,0,0)}{\big\|(1,1,1) - \langle (1,1,1),(1,0,0)\rangle (1,0,0)\big\|} \\ &= \frac{(1,1,1)-(1,0,0)}{\big\|(1,1,1)-(1,0,0)\big\|} \\ &= \frac{(0,1,1)}{\big\|(0,1,1)\big\|} \\ &= \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \end{aligned}$$

and

$$\begin{aligned} e_3 &= \frac{v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2}{\big\|v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2\big\|} \\ &= \frac{(1,1,2) - \langle (1,1,2),(1,0,0)\rangle (1,0,0) - \left\langle (1,1,2), \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right\rangle \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)}{\left\|(1,1,2) - \langle (1,1,2),(1,0,0)\rangle (1,0,0) - \left\langle (1,1,2), \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right\rangle \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right\|} \\ &= \frac{(1,1,2)-(1,0,0) - \frac{3}{\sqrt{2}}\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)}{\left\|(1,1,2)-(1,0,0) - \frac{3}{\sqrt{2}}\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right\|} \\ &= \frac{\left(0, -\frac{1}{2}, \frac{1}{2}\right)}{\left\|\left(0, -\frac{1}{2}, \frac{1}{2}\right)\right\|} \\ &= \left(0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \end{aligned}$$

and we're done. $\square$

Problem 4

Suppose n is a positive integer. Prove that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \ldots, \frac{\sin nx}{\sqrt{\pi}}$$

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

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

Proof. First we show that all vectors in the list have norm 1. Notice

$$\begin{aligned}
\left\| \frac{1}{\sqrt{2\pi}} \right\| &= \sqrt{\int_{-\pi}^{\pi} \frac{1}{2\pi} dx} \\
&= \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} dx} \\
&= 1.
\end{aligned}$$

And for $k \in \mathbb{Z}^+$, we have

$$\begin{aligned}
\left\| \frac{\cos(kx)}{\sqrt{\pi}} \right\| &= \sqrt{\frac{1}{\pi} \int_{-\pi}^{\pi} \cos(kx)^2 dx} \\
&= \sqrt{\frac{1}{\pi} \left[\frac{\sin(2kx)}{4k} + \frac{x}{2} \right]_{-\pi}^{\pi}} \\
&= \sqrt{\frac{1}{\pi} \left[\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right]} \\
&= 1,
\end{aligned}$$

and

$$\begin{aligned}
\left\| \frac{\sin(kx)}{\sqrt{\pi}} \right\| &= \sqrt{\frac{1}{\pi} \int_{-\pi}^{\pi} \sin(kx)^2 dx} \\
&= \sqrt{\frac{1}{\pi} \left[\frac{x}{2} - \frac{\cos(2kx)}{4k} \right]_{-\pi}^{\pi}} \\
&= \sqrt{\frac{1}{\pi} \left[\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right]} \\
&= 1,
\end{aligned}$$

so indeed all vectors have norm 1. Now we show them to be pairwise orthogonal. Suppose $j, k \in \mathbb{Z}$ are such that $j \neq k$. It follows from basic calculus

$$\begin{aligned}\left\langle \frac{\sin(jx)}{\sqrt{\pi}}, \frac{\sin(kx)}{\sqrt{\pi}} \right\rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(jx)\sin(kx)dx \\ &= \frac{1}{\pi} \left[\frac{k\sin(jx)\cos(kx) + j\cos(jx)\sin(kx)}{j^2 - k^2} \right]_{-\pi}^{\pi} \\ &= 0,\end{aligned}$$

$$\begin{aligned}\left\langle \frac{\sin(jx)}{\sqrt{\pi}}, \frac{\cos(kx)}{\sqrt{\pi}} \right\rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(jx)\cos(kx)dx \\ &= -\frac{1}{\pi} \left[\frac{k\sin(jx)\sin(kx) + j\cos(jx)\cos(kx)}{j^2 - k^2} \right]_{-\pi}^{\pi} \\ &= -\frac{1}{\pi} \left[\left(\frac{j\cos(j\pi)\cos(k\pi)}{j^2 - k^2} \right) - \left(\frac{j\cos(-j\pi)\cos(-k\pi)}{j^2 - k^2} \right) \right] \\ &= 0,\end{aligned}$$

$$\begin{aligned}\left\langle \frac{\cos(jx)}{\sqrt{\pi}}, \frac{\cos(kx)}{\sqrt{\pi}} \right\rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(jx)\cos(kx)dx \\ &= \frac{1}{\pi} \left[\frac{j\sin(jx)\cos(kx) - k\cos(jx)\sin(kx)}{j^2 - k^2} \right]_{-\pi}^{\pi} \\ &= 0,\end{aligned}$$

$$\begin{aligned}\left\langle \frac{\sin(jx)}{\sqrt{\pi}}, \frac{\cos(jx)}{\sqrt{\pi}} \right\rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(jx)\cos(jx)dx \\ &= \left[-\frac{\cos^2(jx)}{2j} \right]_{-\pi}^{\pi} \\ &= 0,\end{aligned}$$

$$\begin{aligned}\left\langle \frac{1}{\sqrt{2\pi}}, \frac{\cos(jx)}{\sqrt{\pi}} \right\rangle &= \frac{1}{\sqrt{2}\pi} \int_{-\pi}^{\pi} \cos(jx)dx \\ &= \frac{1}{\sqrt{2}\pi} \left[\frac{\sin(jx)}{j} \right]_{-\pi}^{\pi} \\ &= 0,\end{aligned}$$

and

$$\begin{aligned}\left\langle \frac{1}{\sqrt{2\pi}}, \frac{\sin(jx)}{\sqrt{\pi}} \right\rangle &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \sin(jx) dx \\ &= \frac{1}{\sqrt{2\pi}} \left[-\frac{\cos(jx)}{j} \right]_{-\pi}^{\pi} \\ &= \frac{1}{\sqrt{2\pi}} \left[-\frac{\cos(j\pi) - \cos(-j\pi)}{j} \right] \\ &= 0.\end{aligned}$$

Thus the list is indeed an orthonormal list in $\mathcal{C}[-\pi, \pi]$. □

Problem 5

On $\mathcal{P}_2(\mathbb{R})$, 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 $\mathcal{P}_2(\mathbb{R})$.

Proof. First notice $\|1\| = 1$, hence $e_1 = 1$. Next notice

$$\begin{aligned} v_2 - \langle v_1, e_1 \rangle e_1 &= x - \langle x, 1 \rangle \\ &= x - \int_0^1 x\, dx \\ &= x - \frac{1}{2} \end{aligned}$$

and

$$\begin{aligned} \left\| x - \frac{1}{2} \right\| &= \sqrt{\left\langle x - \frac{1}{2}, x - \frac{1}{2} \right\rangle} \\ &= \sqrt{\int_0^1 \left(x - \frac{1}{2} \right) \left(x - \frac{1}{2} \right) dx} \\ &= \sqrt{\int_0^1 \left(x^2 - x + \frac{1}{4} \right) dx} \\ &= \sqrt{\frac{1}{3} - \frac{1}{2} + \frac{1}{4}} \\ &= \frac{1}{2\sqrt{3}}, \end{aligned}$$

and therefore we have

$$e_2 = 2\sqrt{3}\left(x - \frac{1}{2}\right).$$

To compute e_3, first notice

$$\begin{aligned}
v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2 &= x^2 - \int_0^1 x^2\,dx - \left[2\sqrt{3}\int_0^1 x^2\left(x - \frac{1}{2}\right)\,dx\right] e_2 \\
&= x^2 - \frac{1}{3} - \left[2\sqrt{3}\int_0^1 \left(x^3 - \frac{x^2}{2}\right)\,dx\right]\left[2\sqrt{3}\left(x - \frac{1}{2}\right)\right] \\
&= x^2 - \frac{1}{3} - 12\left(\frac{1}{4} - \frac{1}{6}\right)\left(x - \frac{1}{2}\right) \\
&= x^2 - \frac{1}{3} - \left(x - \frac{1}{2}\right) \\
&= x^2 - x + \frac{1}{6}
\end{aligned}$$

and

$$\begin{aligned}
\left\|x^2 - x + \frac{1}{6}\right\| &= \sqrt{\left\langle x^2 - x + \frac{1}{6}, x^2 - x + \frac{1}{6}\right\rangle} \\
&= \sqrt{\int_0^1 \left(x^2 - x + \frac{1}{6}\right)\left(x^2 - x + \frac{1}{6}\right)\,dx} \\
&= \sqrt{\int_0^1 \left(x^4 - 2x^3 + \frac{4}{3}x^2 - \frac{x}{3} + \frac{1}{36}\right)\,dx} \\
&= \sqrt{\frac{1}{5} - \frac{1}{2} + \frac{4}{9} - \frac{1}{6} + \frac{1}{36}} \\
&= \frac{1}{\sqrt{180}} \\
&= \frac{1}{6\sqrt{5}}.
\end{aligned}$$

Thus

$$e_3 = 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right),$$

and we're done. □

Problem 7

Find a polynomial $q \in \mathcal{P}_2(\mathbb{R})$ such that

$$p\left(\frac{1}{2}\right) = \int_0^1 p(x)q(x)\,dx$$

for every $p \in \mathcal{P}_2(\mathbb{R})$.

Proof. Consider the inner product $\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx$ on $\mathcal{P}_2(\mathbb{R})$. Define $\varphi \in \mathcal{L}(\mathcal{P}_2(\mathbb{R}))$ by $\varphi(p) = p\left(\frac{1}{2}\right)$ and let e_1, e_2, e_3 be the orthonormal basis found in Problem 5. By the Riesz Representation Theorem, there exists $q \in \mathcal{P}_2(\mathbb{R})$ such that $\varphi(p) = \langle p, q\rangle$ for all $p \in \mathcal{P}_2(\mathbb{R})$. That is, such that

$$p\left(\frac{1}{2}\right) = \int_0^1 p(x)q(x)\,dx.$$

Equation 6.43 in the proof of the Riesz Representation Theorem fashions a way to find q. In particular, we have

$$\begin{aligned}
q(x) &= \overline{\varphi(e_1)}\,e_1 + \overline{\varphi(e_2)}\,e_2 + \overline{\varphi(e_3)}\,e_3 \\
&= e_1 + 2\sqrt{3}\left(\frac{1}{2} - \frac{1}{2}\right)e_2 + 6\sqrt{5}\left(\frac{1}{4} - \frac{1}{2} + \frac{1}{6}\right)e_3 \\
&= 1 + 6\sqrt{5}\left(\frac{-1}{12}\right)\left[6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right)\right] \\
&= -15(x^2 - x) - \frac{3}{2},
\end{aligned}$$

as desired. □

Problem 9

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

Proof. Suppose $v_1, \ldots, v_m$ are linearly dependent. Let j be the smallest integer in $\{1, \ldots, m\}$ such that $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$. Then $v_1, \ldots, v_{j-1}$ are linearly independent. The first $j - 1$ steps of the Gram-Schmidt Procedure will produce an orthonormal list $e_1, \ldots, e_{j-1}$. At step j, however, notice

$$v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1} = v_j - v_j = 0,$$

and we are left trying to assign e_j to $\frac{0}{0}$, which is undefined. Thus the procedure cannot be applied to a linearly dependent list. □

Problem 11

Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V such that $\langle v, w \rangle_1 = 0$ if and only if $\langle v, w \rangle_2 = 0$. Prove that there is a positive number c such that $\langle v, w \rangle_1 = c\langle v, w \rangle_2$ for every $v, w \in V$.

Proof. Let $v, w \in V$ be arbitrary. By hypothesis, if v and w are orthogonal relative to one of the inner products, they're orthogonal relative to the other. Hence any choice of $c \in \mathbb{R}$ would satisfy $\langle v, w \rangle_1 = c\langle v, w \rangle_2$. So suppose v and w are not orthogonal relative to either inner product. Then both v and w must be nonzero (by Theorem 6.7, parts b and c, respectively). Thus $\langle v, v \rangle_1$, $\langle w, w \rangle_1$, $\langle v, v \rangle_2$, and $\langle w, w \rangle_2$ are all nonzero as well. It now follows

$$
\begin{aligned}
0 &= \langle v, w \rangle_1 - \frac{\langle v, w \rangle_1}{\langle v, v \rangle_1} \langle v, v \rangle_1 \\
&= \langle v, w \rangle_1 - \left\langle v, \overline{\left(\frac{\langle v, w \rangle_1}{\langle v, v \rangle_1} \right)} v \right\rangle_1 \\
&= \left\langle v, w - \overline{\left(\frac{\langle v, w \rangle_1}{\langle v, v \rangle_1} \right)} v \right\rangle_1 \\
&= \left\langle v, w - \overline{\left(\frac{\langle v, w \rangle_1}{\langle v, v \rangle_1} \right)} v \right\rangle_2 \\
&= \langle v, w \rangle_2 - \left\langle v, \overline{\left(\frac{\langle v, w \rangle_1}{\langle v, v \rangle_1} \right)} v \right\rangle_2 \\
&= \langle v, w \rangle_2 - \frac{\langle v, w \rangle_1}{\langle v, v \rangle_1} \langle v, v \rangle_2 \\
&= \langle v, w \rangle_2 - \frac{\langle v, v \rangle_2}{\langle v, v \rangle_1} \langle v, w \rangle_1,
\end{aligned}
$$

where the fifth equality follows by our hypothesis. Thus

$$
\langle v, w \rangle_1 = \frac{\|v\|_1^{\,2}}{\|v\|_2^{\,2}} \langle v, w \rangle_2. \tag{4}
$$

By a similar computation, notice

$$\begin{aligned}
0 &= \langle v, w \rangle_1 - \frac{\langle v, w \rangle_1}{\langle w, w \rangle_1} \langle w, w \rangle_1 \\
&= \langle v, w \rangle_1 - \left\langle \frac{\langle v, w \rangle_1}{\langle w, w \rangle_1} w, w \right\rangle_1 \\
&= \left\langle v - \frac{\langle v, w \rangle_1}{\langle w, w \rangle_1} w, w \right\rangle_1 \\
&= \left\langle v - \frac{\langle v, w \rangle_1}{\langle w, w \rangle_1} w, w \right\rangle_2 \\
&= \langle v, w \rangle_2 - \left\langle \frac{\langle v, w \rangle_1}{\langle w, w \rangle_2} w, w \right\rangle_2 \\
&= \langle v, w \rangle_2 - \frac{\langle v, w \rangle_1}{\langle w, w \rangle_2} \langle w, w \rangle_2 \\
&= \langle v, w \rangle_2 - \frac{\langle w, w \rangle_2}{\langle w, w \rangle_1} \langle v, w \rangle_1,
\end{aligned}$$

and thus

$$\langle v, w \rangle_1 = \frac{\|w\|_1^{\,2}}{\|w\|_2^{\,2}} \langle v, w \rangle_2 \tag{5}$$

as well. By combining Equations (4) and (5), we conclude

$$\frac{\langle v, v \rangle_1}{\langle v, v \rangle_2} = \frac{\langle w, w \rangle_1}{\langle w, w \rangle_2}.$$

Since v and w were arbitrary nonzero vectors in V, choosing $c = \|u\|_1^{\,2} / \|u\|_2^{\,2}$ for any $u \neq 0$ guarantees $\langle v, w \rangle_1 = c \langle v, w \rangle_2$ for every $v, w \in V$, as was to be shown. ☐

Problem 13

Suppose $v_1, \ldots, v_m$ is a linearly independent list in V. Show that there exists $w \in V$ such that $\langle w, v_j \rangle > 0$ for all $j \in \{1, \ldots, m\}$.

Proof. Let $W = \operatorname{span}(v_1, \ldots, v_m)$. Given $v \in W$, let $a_1, \ldots, a_m \in \mathbb{F}$ be such that $v = a_1 v_1 + \cdots + a_m v_m$. Define $\varphi \in \mathcal{L}(W)$ by

$$\varphi(v) = a_1 + \cdots + a_m.$$

By the Riesz Representation Theorem, there exists $w \in W$ such that $\varphi(v) = \langle v, w \rangle$ for all $v \in W$. But then $\varphi(v_j) = 1$ for $j \in \{1, \ldots, m\}$, and indeed such a $w \in V$ exists. ☐

Problem 15

Suppose $C_{\mathbb{R}}([-1,1])$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$$

for $f, g \in C_{\mathbb{R}}([-1,1])$. Let φ be the linear functional on $C_{\mathbb{R}}([-1,1])$ defined by $\varphi(f) = f(0)$. Show that there does not exist $g \in C_{\mathbb{R}}([-1,1])$ such that

$$\varphi(f) = \langle f, g\rangle$$

for every $f \in C_{\mathbb{R}}([-1,1])$.

Proof. Suppose not. Then there exists $g \in C_{\mathbb{R}}([-1,1])$ such that

$$\varphi(f) = \langle f, g\rangle$$

for every $f \in C_{\mathbb{R}}([-1,1])$. Choose $f(x) = x^2 g(x)$. Then $f(0) = 0$, and hence

$$\int_{-1}^{1} f(x)g(x)dx = \int_{-1}^{1} [xg(x)]^2 dx = 0.$$

Now, let $h(x) = xg(x)$. Since h is continuous on $[-1,1]$, there exists an interval $[a,b] \subseteq [-1,1]$ such that $h(x) \neq 0$ for all $x \in [a,b]$. By the Extreme Value Theorem, $h(x)^2$ has a minimum at some $m \in [a,b]$. Thus $h(m)^2 > 0$, and we now conclude

$$0 = \int_{-1}^{1} h(x)^2 dx = \int_{a}^{b} h(x)^2 dx \geq (b-a)h(m)^2 > 0,$$

which is absurd. Thus it must be that no such g exists. $\square$

Problem 17

For $u \in V$, let Φ_u denote the linear functional on V defined by

$$(\Phi_u)(v) = \langle v, u\rangle$$

for $v \in V$.

(a) Show that if $\mathbb{F} = \mathbb{R}$, then Φ is a linear map from V to V'.

(b) Show that if $\mathbb{F} = \mathbb{C}$ and $V \neq \{0\}$, then Φ is not a linear map.

(c) Show that Φ is injective.

(d) Suppose $\mathbb{F} = \mathbb{R}$ and V is finite-dimensional. Use parts (a) and (c) and a dimension-counting argument (but without using 6.42) to show that Φ is an isomorphism from V to V'.

Proof. (a) Suppose $\mathbb{F} = \mathbb{R}$. Let $u, u' \in V$ and $\alpha \in \mathbb{R}$. Then, for all $v \in V$, we have

$$\Phi_{u+u'}(v) = \langle v, u + u' \rangle = \langle v, u \rangle + \langle v, u' \rangle = \Phi_u(v) + \Phi_{u'}(v)$$

and

$$\Phi_{\alpha u}(v) = \langle v, \alpha u \rangle = \overline{\alpha} \langle v, u \rangle = \alpha \langle v, u \rangle = \alpha \Phi_u(v).$$

Thus Φ is indeed a linear map.

(b) Suppose $\mathbb{F} = \mathbb{C}$ and $V \neq \{0\}$. Let $u \in V$. Then, given $v \in V$, we have

$$\Phi_{iu}(v) = \langle v, iu \rangle = \bar{i} \langle v, u \rangle,$$

whereas

$$i\Phi_u(v) = i\langle v, u \rangle.$$

Thus $\Phi_{iu} \neq i\Phi_u$, and indeed Φ is not a linear map, since is is not homogeneous.

(c) Suppose $u, u' \in V$ are such that $\Phi_u = \Phi_{u'}$. Then, for all $v \in V$, we have

$$\begin{aligned}
&\Phi_u(v) = \Phi_{u'}(v) \\
\implies &\langle v, u \rangle = \langle v, u' \rangle \\
\implies &\langle v, u \rangle - \langle v, u' \rangle = 0 \\
\implies &\langle v, u - u' \rangle = 0.
\end{aligned}$$

In particular, choosing $v = u - u'$, the above implies $\langle u - u', u - u' \rangle = 0$, which is true iff $u - u' = 0$. Thus we conclude $u = u'$, so that Φ is indeed injective.

(d) Suppose $\mathbb{F} = \mathbb{R}$ and $\dim V = n$. Notice that since $\Phi : V \hookrightarrow V'$, we have

$$\dim V = \dim \operatorname{null} \Phi + \dim \operatorname{range} \Phi = \dim \operatorname{range} \Phi.$$

Thus Φ is surjective as well, and we have $V \cong V'$, as was to be shown. $\square$

C: Orthogonal Complements and Minimization Problems

Problem 1

Suppose $v_1, \ldots, v_m \in V$. Prove that

$$\{v_1, \ldots, v_m\}^\perp = (\operatorname{span}(v_1, \ldots, v_m))^\perp.$$

Proof. Suppose $v \in \{v_1, \ldots, v_m\}^\perp$. Then $\langle v, v_k \rangle = 0$ for $k = 1, \ldots, m$. Let $u \in \text{span}(v_1, \ldots, v_m)$ be arbitrary. We want to show $\langle v, u \rangle = 0$, since this implies $v \in (\text{span}(v_1, \ldots, v_m))^\perp$ and hence $\{v_1, \ldots, v_m\}^\perp \subseteq (\text{span}(v_1, \ldots, v_m))^\perp$. To see this, notice

$$\begin{aligned}\langle v, u \rangle &= \langle v, \alpha_1 v_1 + \cdots + \alpha_m v_m \rangle \\ &= \alpha_1 \langle v, v_1 \rangle + \cdots + \alpha_m \langle v, v_m \rangle \\ &= 0,\end{aligned}$$

as desired. Next suppose $v' \in (\text{span}(v_1, \ldots, v_m))^\perp$. Since $v_1, \ldots, v_m$ are all clearly elements of $\text{span}(v_1, \ldots, v_m)$, clearly $v' \in \{v_1, \ldots, v_m\}^\perp$, and thus $(\text{span}(v_1, \ldots, v_m))^\perp \subseteq \{v_1, \ldots, v_m\}^\perp$. Therefore we conclude $\{v_1, \ldots, v_m\}^\perp = (\text{span}(v_1, \ldots, v_m))^\perp$. $\square$

Problem 3

Suppose U is a subspace of V with basis $u_1, \ldots, u_m$ and

$$u_1, \ldots, u_m, w_1, \ldots, w_n$$

is a basis of V. Prove that if the Gram-Schmidt Procedure is applied to the basis of V above, producing a list $e_1, \ldots, e_m, f_1, \ldots, f_n$, then $e_1, \ldots, e_m$ is an orthonormal basis of U and $f_1, \ldots, f_n$ is an orthonormal basis of $U^\perp$.

Proof. By 6.31, $\text{span}(u_1, \ldots, u_m) = \text{span}(e_1, \ldots, e_m)$. Since $e_1, \ldots, e_m$ is an orthonormal list by construction (and linearly independent by 6.26), $e_1, \ldots, e_m$ is indeed an orthonormal basis of U. Next, since each of f_i is orthogonal to each e_j, so too is each f_i orthogonal to any element of U. Thus $f_k \in U^\perp$ for $k = 1, \ldots, n$. Now, since $\dim U^\perp = \dim V - \dim U = n$ by 6.50, we conclude $f_1, \ldots, f_n$ is an orthonormal list of length $\dim U^\perp$ and hence an orthonormal basis of $U^\perp$. $\square$

Problem 5

Suppose V is finite-dimensional and U is a subspace of V. Show that $P_{U^\perp} = I - P_U$, where I is the identity operator on V.

Proof. For $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U^\perp$. It follows

$$\begin{aligned}P_{U^\perp}(v) &= w \\ &= (u + w) - u \\ &= Iv - P_U v,\end{aligned}$$

and therefore $P_{U^\perp} = I - P_U$, as desired. $\square$

Problem 7

Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in $\operatorname{null} P$ is orthogonal to every vector in $\operatorname{range} P$. Prove that there exists a subspace U of V such that $P = P_U$.

Proof. By Problem 4 of Chapter 5B, we know $V = \operatorname{null} P \oplus \operatorname{range} P$. Let $v \in V$. Then there exist $u \in \operatorname{null} P$ and $w \in \operatorname{range} P$ such that $v = u + w$ and hence

$$\begin{aligned} Pv &= P(u + w) \\ &= Pu + Pw \\ &= Pw. \end{aligned}$$

Let $U = \operatorname{range} P$ and notice that $\operatorname{null} P \subseteq \operatorname{null} P_U = U^{\perp}$ by 6.55e. Then $Pv = Pw = P_U(v)$, and so U is the desired subpace. $\square$

Problem 9

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

Proof. ($\Leftarrow$) Suppose $P_U T P_U = T P_U$ and let $u \in U$. It follows

$$TP_u(u) = P_U T P_U(v)$$

and thus

$$Tu = P_U T u.$$

Since $\operatorname{range} P_U = U$ by 6.55d, this implies $Tu \in U$. Thus U is indeed invariant under T.

($\Rightarrow$) Now suppose U is invariant under T and let $v \in V$. Since $P_U(v) \in U$, it follows that $TP_U(v) \in U$. And thus $P_U T P_U(v) = T P_U(v)$, as desired. $\square$

Problem 11

In $\mathbb{R}^4$, let

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

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

Proof. We first apply the Gram-Schmidt Procedure to $v_1 = (1,1,0,0)$ and $v_2 = (1,1,1,2)$. This yields

$$\begin{aligned} e_1 &= \frac{v_1}{\|v_1\|} \\ &= \frac{(1,1,0,0)}{\|(1,1,0,0)\|} \\ &= \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right) \end{aligned}$$

and

$$\begin{aligned} e_2 &= \frac{v_2 - \langle v_2, e_1\rangle e_1}{\|v_2 - \langle v_2, e_1\rangle e_1\|} \\ &= \frac{(1,1,1,2) - \left\langle (1,1,1,2), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)\right\rangle \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)}{\left\|(1,1,1,2) - \left\langle (1,1,1,2), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)\right\rangle \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)\right\|} \\ &= \frac{(1,1,1,2) - \frac{2}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)}{\left\|(1,1,1,2) - \frac{2}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)\right\|} \\ &= \frac{(0,0,1,2)}{\|(0,0,1,2)\|} \\ &= \left(0, 0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right). \end{aligned}$$

Now, with our orthonormal basis e_1, e_2, it follows by 6.55(i) and 6.56 that $\|u - (1,2,3,4)\|$ is minimized by the vector

$$\begin{aligned} u &= P_U(1,2,3,4) \\ &= \langle (1,2,3,4), e_1\rangle e_1 + \langle (1,2,3,4), e_2\rangle e_2 \\ &= \frac{3}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right) + \frac{11}{\sqrt{2}}\left(0, 0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right) \\ &= \left(\frac{3}{2}, \frac{3}{2}, 0, 0\right) + \left(0, 0, \frac{11}{5}, \frac{22}{5}\right) \\ &= \left(\frac{3}{2}, \frac{3}{2}, \frac{11}{5}, \frac{22}{5}\right), \end{aligned}$$

completing the proof. $\square$

Problem 13

Find $p \in \mathcal{P}_5(\mathbb{R})$ that makes

$$\int_{-\pi}^{\pi} \left|\sin x - p(x)\right|^2 dx$$

as small as possible.

Proof. Let $\mathcal{C}_{\mathbb{R}}[-\pi, \pi]$ denote the real inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

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

and let U denote the subspace of $\mathcal{C}_{\mathbb{R}}[-\pi, \pi]$ consisting of the polynomials with real coefficients and degree at most 5. In this inner product space, observe that

$$\left\|\sin x - p(x)\right\| = \sqrt{\int_{-\pi}^{\pi} \left(\sin x - p(x)\right)^2 dx} = \sqrt{\int_{-\pi}^{\pi} \left|\sin x - p(x)\right|^2 dx}.$$

Notice also that $\sqrt{\int_{-\pi}^{\pi} \left|\sin x - p(x)\right|^2 dx}$ is minimized if and only if $\int_{-\pi}^{\pi} \left|\sin x - p(x)\right|^2 dx$ is minimized. Thus by 6.56, we may conclude $p(x) = P_U(\sin x)$ minimizes $\int_{-\pi}^{\pi} \left|\sin x - p(x)\right|^2 dx$. To compute $P_U(\sin x)$, we first find an orthonormal basis of $\mathcal{C}_{\mathbb{R}}[-\pi, \pi]$ by applying the Gram-Schmidt Procedure to the basis $1, x, x^2, x^3, x^4, x^5$ of U. A lengthy computation yields the orthonormal basis

$$\begin{aligned}
e_1 &= \frac{1}{\sqrt{2\pi}} \\
e_2 &= \frac{\sqrt{\frac{3}{2}}x}{x^{3/2}} \\
e_3 &= -\frac{\sqrt{\frac{5}{2}}\left(\pi^2 - 3x^2\right)}{2\pi^{5/2}} \\
e_4 &= -\frac{\sqrt{\frac{7}{2}}\left(3\pi^2 x - 5x^3\right)}{2\pi^{7/2}} \\
e_5 &= \frac{3\left(3\pi^4 - 30\pi^2 x^2 + 35x^4\right)}{8\sqrt{2}\pi^{9/2}} \\
e_6 &= -\frac{\sqrt{\frac{11}{2}}\left(15\pi^4 x - 70\pi^2 x^3 + 63x^5\right)}{8\pi^{11/2}}.
\end{aligned}$$

Now we compute $P_U(\sin x)$ using 6.55(i), yielding

$$P_U(\sin x) = \frac{105\left(1485 - 153\pi^2 + \pi^4\right)}{8\pi^6}x - \frac{315\left(1155 - 125\pi^2 + \pi^4\right)}{4\pi^8}x^3 + \frac{693\left(945 - 105\pi^2 + \pi^4\right)}{8\pi^{10}}x^5,$$

which is the desired polynomial. $\square$

Chapter 7: Operators on Inner Product Spaces

Linear Algebra Done Right, by Sheldon Axler

A: Self-Adjoint and Normal Operators

Problem 1

Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbb{F}^n)$ by

$$T(z_1, \ldots, z_n) = (0, z_1, \ldots, z_{n-1}).$$

Find a formula for $T^*(z_1, \ldots, z_n)$.

Proof. Fix $(y_1, \ldots, y_n) \in \mathbb{F}^n$. Then for all $(z_1, \ldots, z_n) \in \mathbb{F}^n$, we have

$$\begin{aligned}
\langle (z_1, \ldots, z_n), T^*(y_1, \ldots, y_n) \rangle &= \langle T(z_1, \ldots, z_n), (y_1, \ldots, y_n) \rangle \\
&= \langle (0, z_1, \ldots, z_{n-1}), (y_1, \ldots, y_n) \rangle \\
&= z_1 y_2 + z_2 y_3 + \cdots + z_{n-1} y_n \\
&= \langle (z_1, \ldots, z_{n-1}, z_n), (y_2, \ldots, y_n, 0) \rangle.
\end{aligned}$$

Thus T^* is the left-shift operator. That is, for all $(z_1, \ldots, z_n) \in \mathbb{F}^n$, we have

$$T^*(z_1, \ldots, z_n) = (z_2, \ldots, z_n, 0),$$

as desired. $\square$

Problem 2

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Prove that λ is an eigenvalue of T if and only if $\bar{\lambda}$ is an eigenvalue of T^*.

Proof. Suppose λ is an eigenvalue of T. Then there exists $v \in V$ such that $Tv = \lambda v$. It follows

$$\begin{aligned}
\lambda \text{ is not an eigenvalue of } T &\iff T - \lambda I \text{ is invertible} \\
&\iff S(T - \lambda I) = (T - \lambda I)S = I \\
&\qquad \text{for some } S \in \mathcal{L}(V) \\
&\iff S^*(T^* - \lambda I)^* = (T - \lambda I)^* S^* = I^* \\
&\qquad \text{for some } S^* \in \mathcal{L}(V) \\
&\iff (T - \lambda I)^* \text{ is invertible} \\
&\iff T^* - \bar{\lambda} I \text{ is invertible} \\
&\iff \bar{\lambda} \text{ is not an eigenvalue of } T^*.
\end{aligned}$$

Since the first statement and the last statement are equivalent, so too are their contrapositives. Hence λ is an eigenvalue of T if and only if $\bar{\lambda}$ is an eigenvalue of T^*, as was to be shown. $\square$

Problem 3

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^*.

Proof. ($\Rightarrow$) First suppose U is invariant under T, and let $x \in U^\perp$. For any $u \in U$, it follows

$$\begin{aligned} \langle T^* x, u \rangle &= \langle x, Tu \rangle \\ &= 0, \end{aligned}$$

where the second equality follows since $Tu \in U$ (by hypothesis). Thus $T^* x \in U^\perp$ for all $x \in U^\perp$. That is, $U^\perp$ is invariant under T^*.

($\Leftarrow$) Now suppose $U^\perp$ is invariant under T^*, and let $y \in U$. For any $u' \in U^\perp$, it follows

$$\begin{aligned} \langle Ty, u' \rangle &= \langle y, T^* u' \rangle \\ &= 0, \end{aligned}$$

where the second equality follows since $T^* u' \in U^\perp$ (by hypothesis). Thus $Ty \in U$ for all $y \in U$. That is, U is invariant under T, completing the proof. $\square$

Problem 5

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 $T \in \mathcal{L}(V, W)$.

Proof. Let $T \in \mathcal{L}(V, W)$. Notice

$$\begin{aligned} \dim \operatorname{null} T^* &= \dim (\operatorname{range} T)^\perp \\ &= \dim W - \dim \operatorname{range} T \\ &= \dim W + \dim \operatorname{null} T - \dim V, \end{aligned}$$

where the first equality follows by 7.7(a), the second equality follows by 6.50, and the third equality follows by the Fundamental Theorem of Linear Maps. Next notice

$$\begin{aligned}\dim \operatorname{range} T^* &= \dim (\operatorname{null} T)^{\perp}\\ &= \dim V - \dim \operatorname{null} T\\ &= \dim \operatorname{range} T,\end{aligned}$$

where the first equality follows by 7.7(b), and the second and third equalities follow again by the same theorems above. □

Problem 7

Suppose $S, T \in \mathcal{L}(V)$ are self-adjoint. Prove that ST is self-adjoint if and only if $ST = TS$.

Proof. ($\Rightarrow$) Suppose ST is self-adjoint. We have

$$\begin{aligned}ST &= (ST)^*\\ &= T^*S^*\\ &= TS,\end{aligned}$$

where the second equality follows by 7.6(e).

($\Leftarrow$) Conversely, suppose $ST = TS$. It follows

$$\begin{aligned}(ST)^* &= (TS)^*\\ &= S^*T^*,\end{aligned}$$

where the second equality again follows by 7.6(e), completing the proof. □

Problem 9

Suppose V is a complex inner product space with $V \neq \{0\}$. Show that the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

Proof. Let $\mathcal{A}$ denote the set of self-adjoint operators on V, and suppose $T \in \mathcal{A}$. By 7.6(b), notice $(iT)^* = -iT^*$, so that $\mathcal{A}$ is not closed under scalar multiplication. Thus $\mathcal{A}$ is not a subspace of $\mathcal{L}(V)$. □

Problem 11

Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that there is a subspace U of V such that $P = P_U$ if and only if P is self-adjoint.

Proof. ($\Rightarrow$) First suppose there is a subspace $U \subseteq V$ such that $P = P_U$, and let $v_1, v_2 \in V$. It follows

$$\begin{aligned}
\langle Pv_1, v_2 \rangle &= \langle u_1, u_2 + w_2 \rangle \\
&= \langle u_1, u_2 \rangle + \langle u_1, w_2 \rangle \\
&= \langle u_1, u_2 \rangle \\
&= \langle u_1, u_2 \rangle + \langle w_1, u_2 \rangle \\
&= \langle u_1 + w_1, u_2 \rangle \\
&= \langle v_1, Pv_2 \rangle,
\end{aligned}$$

and thus $P = P^*$.

($\Leftarrow$) Conversely, suppose $P = P^*$. Let $v \in V$. Notice $P(v - Pv) = Pv - P^2v = 0$, and hence $v - Pv \in \operatorname{null} P$. By 7.7(c), we know $\operatorname{null} P = (\operatorname{range} T^*)^{\perp}$. By hypothesis, P is self-adjoint, and hence we have $v - Pv \in (\operatorname{range} T)^{\perp}$. Notice we may write

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

where $Pv \in \operatorname{range} P$ and $v - Pv \in (\operatorname{range} T)^{\perp}$. Let $U = \operatorname{range} P$. Since the above holds for all $v \in V$, we conclude $P = P_U$, and the proof is complete. $\square$

Problem 13

Give an example of an operator $T \in \mathcal{L}(\mathbb{C}^4)$ such that T is normal but not self-adjoint.

Proof. Let T be the operator on $\mathbb{C}^4$ whose matrix with respect to the standard basis is

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

We claim T is normal and not self-adjoint. To see that T is not self-adjoint, notice that the entry in row 2, column 1 does not equal the complex conjugate of the entry in row 1 column 2.

Next, notice

$$\mathcal{M}(TT^*) = \begin{bmatrix} 2 & -3 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 & 0 & 0 \\ -3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 13 & 0 & 0 & 0 \\ 0 & 13 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$\mathcal{M}(T^*T) = \begin{bmatrix} 2 & 3 & 0 & 0 \\ -3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -3 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 13 & 0 & 0 & 0 \\ 0 & 13 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and hence TT^* and T^*T have the same matrix. Thus $TT^* = T^*T$, and T is normal. □

Problem 15

Fix $u, x \in V$. Define $T \in \mathcal{L}(V)$ by

$$Tv = \langle v, u \rangle x$$

for every $v \in V$.

(a) Suppose $\mathbb{F} = \mathbb{R}$. Prove that T is self-adjoint if and only if u, x is linearly dependent.

(a) Prove that T is normal if and only if u, x is linearly dependent.

Proof. We first derive a useful formula for T^* which we'll use in both (a) and (b). Let $w_1, w_2 \in V$ and notice

$$\begin{aligned} \langle w_1, T^* w_2 \rangle &= \langle T w_1, w_2 \rangle \\ &= \langle \langle w_1, u \rangle x, w_2 \rangle \\ &= \langle w_1, u \rangle \langle x, w_2 \rangle \\ &= \langle w_1, \overline{\langle x, w_2 \rangle}\, u \rangle \\ &= \langle w_1, \langle w_2, x \rangle u \rangle, \end{aligned}$$

and thus $T^* w_2 = \langle w_2, x \rangle u$. Since w_2 was arbitrary, we may rewrite this as $T^* v = \langle v, x \rangle u$ for all $v \in V$.

(a) ($\Rightarrow$) Suppose T is self-adjoint. Then we have

$$\langle v, u \rangle x - \langle v, x \rangle u = Tv - T^* v = 0$$

for all $v \in V$. In particular, we have

$$\langle u, u \rangle x - \langle u, x \rangle u = 0.$$

We may assume both u and x are nonzero, for otherwise there is nothing to prove. Hence $\langle u, u \rangle \neq 0$, which forces $\langle u, x \rangle$ to be nonzero as well, and thus the equation above shows u, x is linearly dependent.

($\Leftarrow$) Conversely, suppose u, x is linearly dependent. We may again assume both u and x are nonzero, for otherwise $T = 0$, which is self-adjoint. Thus there exists a nonzero $\alpha \in \mathbb{R}$ such that $u = \alpha x$. It follows

$$\begin{aligned} Tv &= \langle v, u \rangle x \\ &= \langle v, \alpha x \rangle \frac{1}{\alpha} u \\ &= \langle v, x \rangle u \\ &= T^*, \end{aligned}$$

and thus T is self-adjoint, completing the proof.

(b) ($\Rightarrow$) Suppose T is normal and let $v \in V$. It follows

$$\begin{aligned} \langle \langle v, u \rangle x, x \rangle u &= T^*(\langle v, u \rangle x) \\ &= T^*Tv \\ &= TT^*v \\ &= T(\langle v, x \rangle u) \\ &= \langle \langle v, x \rangle u, u \rangle x. \end{aligned}$$

We may assume both u and x are nonzero, for otherwise there is nothing to prove. Since the above holds for $v = u$, we may conclude $\langle \langle v, u \rangle x, x \rangle \neq 0$, which also forces $\langle \langle v, x \rangle u, u \rangle \neq 0$. Thus u, x is linearly dependent.

($\Leftarrow$) Conversely, suppose u, x is linearly dependent. We may again assume both u and x are nonzero, for otherwise $T = 0$, which is normal. Thus there exists a nonzero $\alpha \in \mathbb{R}$ such that $u = \alpha x$. It follows

$$\begin{aligned} T^*Tv &= T^*(\langle v, u \rangle x) \\ &= \langle \langle v, u \rangle x, x \rangle u \\ &= \left\langle \langle v, \alpha x \rangle \frac{1}{\alpha} u, \frac{1}{\alpha} u \right\rangle \alpha x \\ &= \langle \langle v, x \rangle u, u \rangle x \\ &= T(\langle v, x \rangle u) \\ &= TT^*v, \end{aligned}$$

and thus T is normal, completing the proof. ☐

Problem 16

Suppose $T \in \mathcal{L}(V)$ is normal. Prove that

$$\operatorname{range} T = \operatorname{range} T^*.$$

Proof. Suppose $T \in \mathcal{L}(V)$ is normal. We first prove $\operatorname{null} T = \operatorname{null} T^*$. It follows

$$\begin{aligned} v \in \operatorname{null} T &\iff Tv = 0 \\ &\iff \|Tv\| = 0 \\ &\iff \|T^*v\| = 0 \\ &\iff T^*v = 0 \\ &\iff v \in \operatorname{null} T^*, \end{aligned}$$

where the third equivalence follows by 7.20, and indeed we have $\operatorname{null} T = \operatorname{null} T^*$. This implies $(\operatorname{null} T)^\perp = (\operatorname{null} T^*)^\perp$, and by 7.7(b) and 7.7(c), this is equivalent to $\operatorname{range} T^* = \operatorname{range} T$, as desired. ☐

Problem 17

Suppose $T \in \mathcal{L}(V)$ is normal. Prove that

$$\operatorname{null} T^k = \operatorname{null} T \quad \text{and} \quad \operatorname{range} T^k = \operatorname{range} T$$

for every positive integer k.

Proof. To show $\operatorname{null} T^k = \operatorname{null} T$, we first prove $\operatorname{null} T^k = \operatorname{null} T^{k+1}$ for all $k \in \mathbb{Z}^+$. Let $m \in \mathbb{Z}^+$. If $m = 1$, there's nothing to prove, so we may assume $m > 1$. Clearly, if $v \in \operatorname{null} T^m$, then $v \in \operatorname{null} T^{m+1}$, and hence $\operatorname{null} T^m \subseteq \operatorname{null} T^{m+1}$. Next, suppose $v \in \operatorname{null} T^{m+1}$. Then $T(T^m v) = 0$, and hence $T^m v \in \operatorname{null} T$. By Problem 16, this implies $T^m v \in \operatorname{null} T^*$, and by 7.7(a) we have $T^m \in (\operatorname{range} T)^\perp$. Since of course $T^m v \in \operatorname{range} T$ as well, we must have $T^m v = 0$. Thus $v \in \operatorname{null} T^m$, and therefore $\operatorname{null} T^{m+1} \subseteq \operatorname{null} T^m$. Thus $\operatorname{null} T^m = \operatorname{null} T^{m+1}$. Since m was arbitrary, this implies $\operatorname{null} T^k = \operatorname{null} T$ for all $k \in \mathbb{Z}^+$, as desired.

Now we will show $\operatorname{range} T^k = \operatorname{range} T$ for all $k \in \mathbb{Z}^+$. Let $n \in \mathbb{Z}^+$. If $n = 1$, there's nothing to prove, so we may assume $n > 1$. Suppose $w \in \operatorname{range} T^n$. Then there exists $v \in V$ such that $T^n v = w$, and hence $T(T^{n-1}v) = w$, so that $w \in \operatorname{range} T$ as well and we have $\operatorname{range} T^n \subseteq \operatorname{range} T$. Next, notice

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

where the second equality follows from the previous paragraph. Since $\operatorname{range} T^n$ is a subspace of $\operatorname{range} T$ of the same dimension, it must equal $\operatorname{range} T$. And since n was arbitrary, we conclude $\operatorname{range} T^k = \operatorname{range} T$ for all $k \in \mathbb{Z}^+$, completing the proof. ☐

Problem 19

Suppose $T \in \mathcal{L}(\mathbb{C}^3)$ is normal and $T(1,1,1) = (2,2,2)$. Suppose $(z_1, z_2, z_3) \in \operatorname{null} T$. Prove that $z_1 + z_2 + z_3 = 0$.

Proof. By Problem 16, $\operatorname{null} T = \operatorname{null} T^*$, hence $T^*(z_1, z_2, z_3) = 0$. Therefore, we have

$$\begin{aligned} 2(z_1 + z_2 + z_3) &= \langle (2,2,2), (z_1, z_2, z_3) \rangle \\ &= \langle T(1,1,1), (z_1, z_2, z_3) \rangle \\ &= \langle (1,1,1), T^*(z_1, z_2, z_3) \rangle \\ &= \langle (1,1,1), (0,0,0) \rangle \\ &= 0, \end{aligned}$$

and so $z_1 + z_2 + z_3 = 0$, as was to be shown. $\square$

Problem 21

Fix a positive integer n. In the inner product 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

$$V = \operatorname{span}(1, \cos x, \cos 2x, \dots, \cos nx, \sin x, \sin 2x, \dots, \sin nx).$$

(a) Define $D \in \mathcal{L}(V)$ by $Df = f'$. Show that $D^* = -D$. Conclude that D is normal but not self-adjoint.

(b) Define $T \in \mathcal{L}(V)$ by $Tf = f''$. Show that T is self-adjoint.

Proof. From Problem 4 of 6B, recall that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list, and hence it is an orthonormal basis of V.

(a) For $k = 1, \dots, n$, define

$$e_k = \frac{\cos(kx)}{\sqrt{\pi}} \quad \text{and} \quad f_k = \frac{\sin(kx)}{\sqrt{\pi}}.$$

Notice

$$De_k = -\frac{k\sin(kx)}{\sqrt{\pi}} = -kf_k \quad \text{and} \quad Df_k = \frac{k\cos(kx)}{\sqrt{\pi}} = ke_k,$$

and thus, for any $v, w \in V$, it follows

$$\begin{aligned}
\langle v, D^* w\rangle &= \langle Dv, w\rangle \\
&= \left\langle D\left(\left\langle v, \frac{1}{\sqrt{2\pi}}\right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{k=1}^{n}\left(\langle v, e_k\rangle e_k + \langle v, f_k\rangle f_k\right)\right), w\right\rangle \\
&= \left\langle \sum_{k=1}^{n}\left(-k\langle v, e_k\rangle f_k + k\langle v, f_k\rangle e_k\right), w\right\rangle \\
&= -\sum_{k=1}^{n} k\langle v, e_k\rangle\langle f_k, w\rangle + \sum_{k=1}^{n} k\langle v, f_k\rangle\langle e_k, w\rangle \\
&= -\sum_{k=1}^{n} k\langle w, f_k\rangle\langle v, e_k\rangle + \sum_{k=1}^{n} k\langle w, e_k\rangle\langle v, f_k\rangle \\
&= \sum_{k=1}^{n} k\langle w, e_k\rangle\langle v, f_k\rangle - \sum_{k=1}^{n} k\langle w, f_k\rangle\langle v, e_k\rangle \\
&= \left\langle v, \sum_{k=1}^{n} k\langle w, e_k\rangle f_k\right\rangle - \left\langle v, \sum_{k=1}^{n} k\langle w, f_k\rangle e_k\right\rangle \\
&= \left\langle v, \sum_{k=1}^{n}\left(k\langle w, e_k\rangle f_k - k\langle w, f_k\rangle e_k\right)\right\rangle \\
&= \left\langle v, -D\left(\left\langle w, \frac{1}{\sqrt{2\pi}}\right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{k=1}^{n}\left(\langle w, e_k\rangle e_k + \langle w, f_k\rangle f_k\right)\right)\right\rangle \\
&= \langle v, -Dw\rangle,
\end{aligned}$$

and thus $D^* = -D$, showing that D is not self-adjoint. Moreover, notice that this implies

$$DD^* = D(-D) = -DD = (D^*)D = D^*D,$$

so that D is normal, completing the proof.

(b) Notice $T = D^2$, and hence

$$T^* = (DD)^* = D^*D^* = (-D)(-D) = D^2 = T.$$

Thus T is self-adjoint. □

B: The Spectral Theorem

Problem 1

True or false (and give a proof of your answer): There exists $T \in \mathcal{L}(\mathbb{R}^3)$ such that T is not self-adjoint (with respect to the usual inner product) and such that there is a basis of $\mathbb{R}^3$ consisting of eigenvectors of T.

Proof. The statement is true. To see this, consider the linear operator T defined by its action on the basis $(1,0,0), (0,1,0), (0,1,1)$:

$$\begin{aligned} T(1,0,0) &= (0,0,0) \\ T(0,1,0) &= (0,0,0) \\ T(0,1,1) &= (0,1,1). \end{aligned}$$

Notice $T(1,0,0) = 0 \cdot (1,0,0)$ and $T(0,1,0) = 0 \cdot (0,1,0)$, so that $(1,0,0)$ and $(0,1,0)$ are eigenvectors with eigenvalue 0. Also, $(0,1,1)$ is an eigenvector with eigenvalue 1. Thus $(1,0,0), (0,1,0), (0,1,1)$ is a basis of $\mathbb{R}^3$ consisting of eigenvectors of T. That T is not self-adjoint follows from the contrapositive of 7.22, since $(0,1,0)$ and $(0,1,1)$ correspond to distinct eigenvalues yet they are not orthogonal. $\square$

Problem 3

Give an example of an operator $T \in \mathcal{L}(\mathbb{C}^3)$ such that 2 and 3 are the only eigenvalues of T and $T^2 - 5T + 6I \neq 0$.

Proof. Define $T \in \mathcal{L}(\mathbb{C}^3)$ by its action on the standard basis:

$$\begin{aligned} Te_1 &= 2e_2 \\ Te_2 &= e_1 + 2e_2 \\ Te_3 &= 3e_3. \end{aligned}$$

Then

$$\mathcal{M}(T) = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

By 5.32, the only eigenvalues of T are the entries on the diagonal: 2 and 3. Now notice

$$\begin{aligned}(T^2 - 5T + 6I)e_2 &= (T - 3I)(T - 2I)e_2 \\ &= (T - 3I)(Te_2 - 2e_2) \\ &= (T - 3I)(e_1 + 2e_2 - 2e_2) \\ &= (T - 3I)e_1 \\ &= Te_1 - 3e_1 \\ &= -e_1,\end{aligned}$$

so that $T^2 - 5T + 6I \neq 0$. Thus T is an operator of the desired form. $\square$

Problem 5

Suppose $\mathbb{F} = \mathbb{R}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Proof. ($\Leftarrow$) Suppose all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. By 5.41, V has a basis consisting of eigenvectors of T. Dividing each element of the basis by its norm produces an orthonormal basis consisting of eigenvectors of T. By the Real Spectral Theorem, T is self-adjoint, as desired.

($\Rightarrow$) Conversely, suppose T is self-adjoint as suppose $v_1, v_2 \in V$ are eigenvectors of T corresponding to eigenvalues $\lambda_1, \lambda_2 \in \mathbb{R}$ such that $\lambda_1 \neq \lambda_2$. It follows

$$\begin{aligned}0 &= \langle Tv_1, v_2\rangle - \langle v_1, Tv_2\rangle \\ &= \langle \lambda_1 v_1, v_2\rangle - \langle v_1, \lambda_2 v_2\rangle \\ &= \lambda_1\langle v_1, v_2\rangle - \overline{\lambda_2}\langle v_1, v_2\rangle \\ &= \lambda_1\langle v_1, v_2\rangle - \lambda_2\langle v_1, v_2\rangle \\ &= (\lambda_1 - \lambda_2)\langle v_1, v_2\rangle.\end{aligned}$$

Since $\lambda_1 \neq \lambda_2$, it must be that $\langle v_1, v_2\rangle = 0$. Thus all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal. By the Real Spectral Theorem, since T is self-adjoint, T is diagonalizable. And by 5.34, this implies

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T, completing the proof. $\square$

Problem 6

Prove that a normal operator on a complex vector space is self-adjoint if and only if all its eigenvalues are real.

Proof. Let T be a normal operator on a complex vector space, V.

($\Rightarrow$) Suppose T is self-adjoint. Then by 7.13, all eigenvalues of T are real.

($\Leftarrow$) Conversely, suppose all eigenvalues of T are real. By the Complex Spectral Theorem, there exists an orthonormal basis $v_1, \ldots, v_n$ of V consisting of eigenvectors of T. Thus there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that $Tv_k = \lambda_k v_k$ for $k = 1, \ldots, n$. Thus $\mathcal{M}(T)$ is diagonal, and all entries along the diagonal are real. Therefore $\mathcal{M}(T)$ equals the conjugate transpose of $\mathcal{M}(T)$. By 7.10, this implies $\mathcal{M}(T) = \mathcal{M}(T^*)$, and we conclude $T = T^*$, so that T is indeed self-adjoint. □

Problem 7

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 $T^2 = T$.

Proof. By the Complex Spectral Theorem, since T is normal, V has an orthonormal basis $v_1, \ldots, v_n$ consisting of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ be the corresponding eigenvalues, so that

$$Tv_k = \lambda_k v_k$$

for $k = 1, \ldots, n$. Repeatedly applying T to both sides of the equation above 8 times yields

$$T^9 v_k = (\lambda_k)^9 v_k \quad \text{and} \quad T^8 v_k = (\lambda_k)^8 v_k.$$

Since $T^9 = T^8$, we conclude $(\lambda_k)^9 = (\lambda_k)^8$ and thus $\lambda_k \in \{0, 1\}$. In particular, all eigenvalues of T are real, hence by Problem 6 we have that T is self-adjoint.

To see that $T^2 = T$, notice

$$\begin{aligned} T^2 v_k &= (\lambda_k)^2 v_k \\ &= \lambda_k v_k \\ &= Tv_k, \end{aligned}$$

where the second equality follows from the fact that $\lambda_k \in \{0, 1\}$, and the proof is complete. □

Problem 9

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 $S^2 = T$.)

Proof. Suppose $T \in \mathcal{L}(V)$ is normal. By the Complex Spectral Theorem, V has an orthonormal basis $v_1, \ldots, v_n$ consisting of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ be the corresponding eigenvalues, so that

$$Tv_k = \lambda_k v_k$$

for $k = 1, \ldots, n$. Define $S \in \mathcal{L}(V)$ by its action on this basis:

$$Sv_k = \sqrt{\lambda_k} v_k,$$

choosing the complex square root $\sqrt{\lambda_k}$ by some definite rule. Let $v \in V$. Then there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. It follows

$$\begin{aligned} S^2 v &= S^2(\alpha_1 v_1 + \cdots + \alpha_n v_n) \\ &= S\left(\alpha_1 \sqrt{\lambda_1} v_1 + \cdots + \alpha_n \sqrt{\lambda_n} v_n\right) \\ &= \alpha_1 \lambda_1 v_1 + \cdots + \alpha_n \lambda_n v_n \\ &= \alpha_1 T v_1 + \cdots + \alpha_n T v_n \\ &= T(\alpha_1 v_1 + \cdots + \alpha_n v_n) \\ &= Tv. \end{aligned}$$

Thus $S^2 = T$, and indeed T has a square root, as was to be shown. $\square$

Problem 11

Prove or give a counterexample: every self-adjoint operator on V has a cube root. (An operator $T \in \mathcal{L}(V)$ is called a ***cube root*** of $T \in \mathcal{L}(V)$ if $S^3 = T$.)

Proof. Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Regardless of whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{R}$, both Spectral Theorems imply that V has an orthonormal basis $v_1, \ldots, v_n$ consisting of eigenvectors of T. By 7.13, all eigenvalues of T are real. So let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ be the eigenvalues corresponding to $v_1, \ldots, v_n$, so that

$$Tv_k = \lambda_k v_k$$

for $k = 1, \ldots, n$. Define $S \in \mathcal{L}(V)$ by its action on this basis:

$$Sv_k = (\lambda_k)^{\frac{1}{3}} v_k,$$

Let $v \in V$. Then there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. It follows

$$\begin{aligned}
S^3 v &= S^3(\alpha_1 v_1 + \cdots + \alpha_n v_n) \\
&= S^2 \left(\alpha_1 (\lambda_1)^{\frac{1}{3}} v_1 + \cdots + \alpha_n (\lambda_n)^{\frac{1}{3}} v_n \right) \\
&= S \left(\alpha_1 (\lambda_1)^{\frac{2}{3}} v_1 + \cdots + \alpha_n (\lambda_n)^{\frac{2}{3}} v_n \right) \\
&= \alpha_1 \lambda_1 v_1 + \cdots + \alpha_n \lambda_n v_n \\
&= \alpha_1 T v_1 + \cdots + \alpha_n T v_n \\
&= T(\alpha_1 v_1 + \cdots + \alpha_n v_n) \\
&= T v.
\end{aligned}$$

Thus $S^3 = T$, and indeed T has a cube root. Thus, all self-adjoint operators on a finite-dimensional inner product space have a cube root. □

Problem 13

Give an alternative proof of the Complex Spectral Theorem that avoids Schur's Theorem and instead follows the pattern of the proof of the Real Spectral Theorem.

Proof. Suppose (c) holds, so that T has a diagonal matrix with respect to some orthonormal basis of V. The matrix of T^* (with respect to the same basis) is obtained by taking the conjugate transpose of the matrix of T; hence T^* also has a diagonal matrix. Any two diagonal matrices commute; thus T commutes with T^*, which means that T is normal. That is, (a) holds.

We will prove that (a) implies (b) by induction on $\dim V$. For our base case, suppose $\dim V = 1$. Since 5.21 guarantees the existence of an eigenvector of T, clearly (b) is true in this case. Next assume that $\dim V > 1$ and that (a) implies (b) for all complex inner product spaces of smaller dimension.

Suppose (a) holds, so that T is normal. Let u be an eigenvector of T with $\|u\| = 1$, and set $U = \operatorname{span}(u)$. Clearly U is invariant under T. By Problem 3 of 7A, this implies that $U^\perp$ is invariant under T^* as well. But of course T^* is also normal, and since $\dim U^\perp = \dim V - 1$, our inductive hypothesis implies that there exists an orthonormal basis of $U^\perp$ consisting of eigenvectors of $T\,|_{U^\perp}$. Adjoining u to this basis gives an orthonormal basis of V consisting of eigenvectors of T, completing the proof that (a) implies (b).

We have proved that (c) implies (a) and that (a) implies (b). Clearly (b) implies (c), and the proof is complete. □

Problem 15

Find the value of x such that the matrix

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

is normal.

Proof. Let M be the above matrix. We wish to find $x \in \mathbb{F}$ such that $MM^* = M^*M$. Notice

$$\begin{aligned} MM^* &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & x \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & x \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & x \\ 1 & x & 1+x^2 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} M^*M &= \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & x \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & x \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 & x \\ 1 & 2 & 1 \\ x & 1 & 1+x^2 \end{bmatrix}. \end{aligned}$$

Thus it must be that $x = 1$. ☐

C: Positive Operators and Isometries

Problem 1

Prove or give a counterexample: If $T \in \mathcal{L}(V)$ is self-adjoint and there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $\langle Te_j, e_j \rangle \geq 0$ for each j, then T is a positive operator.

Proof. The statement is false. To see this, let $e_1, e_2 \in \mathbb{R}^2$ be the standard basis and consider $T \in \mathcal{L}(\mathbb{R}^2)$ defined by

$$\begin{aligned} Te_1 &= e_1 \\ Te_2 &= -e_2. \end{aligned}$$

Then

$$\mathcal{M}(T) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

and since $\mathcal{M}(T)$ is diagonal, T must be self-adjoint by the Real Spectral Theorem. But notice that the basis

$$v_1 = \frac{1}{\sqrt{2}}(e_1 + e_2)$$
$$v_2 = \frac{1}{\sqrt{2}}(e_1 - e_2)$$

is orthonormal and that

$$\langle Tv_1, v_1 \rangle = \langle v_2, v_1 \rangle = 0$$

and

$$\langle Tv_2, v_2 \rangle = \langle v_1, v_2 \rangle = 0.$$

Thus T is of the desired form, but T is not a positive operator, since

$$\langle Te_2, e_2 \rangle = \langle -e_2, e_2 \rangle = -1,$$

completing the proof. □

Problem 3

Suppose T is a positive operator on V and U is a subspace of V invariant under T. Prove that $T\mid_U \in \mathcal{L}(U)$ is a positive operator on U.

Proof. That $T\mid_U$ is self-adjoint follows by 7.28. Let $u \in U$. Then, since

$$\langle T\mid_U (u), u \rangle = \langle Tu, u \rangle > 0,$$

$T\mid_U$ is a positive operator on U, as was to be shown. □

Problem 5

Prove that the sum of two positive operators on V is positive.

Proof. Let $S, T \in \mathcal{L}(V)$ be positive operators. Notice

$$(S+T)^* = S^* + T^* = S + T,$$

hence $S + T$ is self-adjoint. Next, let $v \in V$. It follows

$$\begin{aligned} \langle (S+T)v, v \rangle &= \langle Sv + Tv, v \rangle \\ &= \langle Sv, v \rangle + \langle Tv, v \rangle \\ &\geq 0, \end{aligned}$$

and thus $S + T$ is a positive operator as well. □

Problem 7

Suppose T is a positive operator on V. Prove that T is invertible if and only if

$$\langle Tv, v\rangle > 0$$

for every $v \in V$ with $v \neq 0$.

Proof. Let T be a positive operator on V.

($\Rightarrow$) Suppose T is invertible and let $v \in V \setminus \{0\}$. Since T is a positive operator, by 7.35(e) there exists $R \in \mathcal{L}(V)$ such that $T = R^2$. Since T is invertible, so is R. In particular, R is injective, and thus $Rv \neq 0$. It follows

$$\begin{aligned}\langle Tv, v\rangle &= \langle R^2, v\rangle \\ &= \langle Rv, R^*v\rangle \\ &= \langle Rv, Rv\rangle \\ &= \|Rv\|^2 \\ &> 0,\end{aligned}$$

completing the proof in one direction.

($\Leftarrow$) Now suppose $\langle Tv, v\rangle > 0$ for every $v \in V \setminus \{0\}$. Assume by way of contraction that T is not invertible, so that there exists $w \in V \setminus \{0\}$ such that $Tw = 0$. But then $\langle Tw, w\rangle = \langle 0, w\rangle = 0$, a contradiction. Thus T must be invertible, completing the proof. $\square$

Problem 9

Prove or disprove: the identity operator on $\mathbb{F}^2$ has infinitely many self-adjoint square roots.