Linear Algebra and Its Applications

A Solution Manual

Yan Zeng

Linear Algebra and Its Applications, 2ed.
A Solution Manual

Yan Zeng

Version 1.0.5, last revised on 2014-08-20

This book is for sale at https://leanpub.com/lax_la_solman

This version was published on 2023-02-21.

This is a Leanpub book. Leanpub empowers authors and publishers with the Lean Publishing process. Lean Publishing is the act of publishing an in-progress ebook using lightweight tools and many iterations to get reader feedback, pivot until you have the right book and build traction once you do.

Abstract

This is a solution manual for *Linear Algebra and Its Applications*, 2nd edition, by Peter Lax [8]. This version omits the following problems: exercise 2, 9 of Chapter 8; exercise 3, 4 of Chapter 17; exercises of Chapter 18; exercise 3 of Appendix 3; exercises of Appendix 4, 5, 8 and 11.

If you find any typos/errors, please email me at quantsummaries@gmail.com.

Contents

Chapter 1

Fundamentals

The book's own solution gives answers to Ex 1, 3, 7, 10, 13, 14, 16, 19, 20, 21.

▶ 1. (page 2)

Proof. Suppose 0 and $0'$ are two zeros of vector addition, then by the definition of zero and commutativity, we have $0' = 0' + 0 = 0 + 0' = 0$. □

▶ 2. (page 3)

Proof. For any $x = (x_1, \cdots, x_n) \in K^n$, we have

$$x + 0 = (x_1, \cdots, x_n) + (0, \cdots, 0) = (x_1 + 0, \cdots, x_n + 0) = (x_1, \cdots, x_n) = x.$$

So $0 = (0, \cdots, 0)$ is *the* zero element of classical vector addition. □

▶ 3. (page 3)

Proof. The isomorphism T can be defined as $T((a_1, \cdots, a_n)) = a_1 + a_2 x + \cdots + a_n x^{n-1}$. □

▶ 4. (page 3)

Proof. Suppose $S = \{s_1, \cdots, s_n\}$. The isomorphism T can be defined as $T(f) = (f(s_1), \cdots, f(s_n))$,$\forall f \in K^S$. □

▶ 5. (page 4)

Proof. For any $p(x) = a_1 + a_2 x + \cdots + a_n x^{n-1}$, we define

$$T(p) = p(x),$$

where p on the left side of the equation is regarded as a polynomial over $\mathbb{R}$ while $p(x)$ on the right side of the equation is regarded as a function defined on $S = \{s_1, \cdots, s_n\}$. To prove T is an isomorphism, it suffices to prove T is one-to-one. This is seen through the observation that

$$\begin{bmatrix} 1 & s_1 & s_1^2 & \cdots & s_1^{n-1} \\ 1 & s_2 & s_2^2 & \cdots & s_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & s_n & s_n^2 & \cdots & s_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} p(s_1) \\ p(s_2) \\ \cdots \\ p(s_n) \end{bmatrix}$$

and the Vandermonde matrix
$$\begin{bmatrix} 1 & s_1 & s_1^2 & \cdots & s_1^{n-1} \\ 1 & s_2 & s_2^2 & \cdots & s_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & s_n & s_n^2 & \cdots & s_n^{n-1} \end{bmatrix}$$
is invertible for distinct s_1, s_2, $\cdots$, s_n. $\square$

▶ 6. (page 4)

Proof. For any $y, y' \in Y$, $z, z' \in Z$ and $k \in K$, we have (by commutativity and associative law)

$$\begin{aligned} (y+z)+(y'+z') &= (z+y)+(y'+z') = z+(y+(y'+z')) = z+((y+y')+z') = z+(z'+(y+y')) \\ &= (z+z')+(y+y') = (y+y')+(z+z') \in Y+Z, \end{aligned}$$

and
$$k(y+z) = ky + kz \in Y + Z.$$
So $Y+Z$ is a linear subspace of X if Y and Z are. $\square$

▶ 7. (page 4)

Proof. For any $x_1, x_2 \in Y \cap Z$, since Y and Z are linear subspaces of X, $x_1+x_2 \in Y$ and $x_1+x_2 \in Z$. Therefore, $x_1 + x_2 \in Y \cap Z$. For any $k \in K$ and $x \in Y \cap Z$, since Y and Z are linear subspaces of X, $kx \in Y$ and $kx \in Z$. Therefore, $kx \in Y \cap Z$. Combined, we conclude $Y \cap Z$ is a linear subspace of X. $\square$

▶ 8. (page 4)

Proof. By definition of zero vector, $0+0=0 \in \{0\}$. For any $k \in K$, $k0 = k(0+0) = k0 + k0$. So $k0 = 0 \in \{0\}$. Combined, we can conclude $\{0\}$ is a linear subspace of X. $\square$

▶ 9. (page 4)

Proof. Define $Y = \{k_1x_1 + \cdots + k_jx_j : k_1, \cdots, k_j \in K\}$. Then clearly $x_1 = 1x_1 + 0x_2 + \cdots + 0x_j \in Y$. Similarly, we can show $x_2, \cdots, x_j \in Y$. Since for any $k_1, \cdots, k_j, k_1', \cdots, k_j' \in K$,

$$(k_1x_1 + \cdots + k_jx_j) + (k_1'x_1 + \cdots + k_j'x_j) = (k_1 + k_1')x_1 + \cdots + (k_j + k_j')x_j \in Y$$

and for any $k_1, \cdots, k_j, k \in K$,

$$k(k_1x_1 + \cdots + k_jx_j) = (kk_1)x_1 + \cdots + (kk_j)x_j \in Y,$$

we can conclude Y is a linear subspace of X containing $x_1, \cdots, x_j$. Finally, if Z is any linear subspace of X containing $x_1, \cdots, x_j$, it is clear that $Y \subset Z$ as Z must be closed under scalar multiplication and vector addition. Combined, we have proven Y is the smallest linear subspace of X containing $x_1, \cdots, x_j$. $\square$

▶ 10. (page 5)

Proof. We prove by contradiction. Without loss of generality, assume $x_1 = 0$. Then $1x_1 + 0x_2 + \cdots + 0x_j = 0$. This shows $x_1, \cdots, x_j$ are linearly dependent, a contradiction. So $x_1 \neq 0$. We can similarly prove $x_2, \cdots, x_j \neq 0$. $\square$

▶ 11. (page 7)

Proof. Suppose Y_i has a basis $y_1^i, \cdots, y_{n_i}^i$. Then it suffices to prove $y_1^1, \cdots, y_{n_1}^1, \cdots, y_1^m, \cdots, y_{n_m}^m$ form a basis of X. By definition of direct sum, these vectors span X, so we only need to show they are linearly independent. In fact, if not, then 0 has two distinct representations: $0 = 0 + \cdots + 0$ and $0 = \sum_{i=1}^m (a_1^i y_1^i + \cdots + a_{n_i}^i y_{n_i}^i)$ for some $a_1^1, \cdots, a_{n_1}^1, \cdots, a_1^m, \cdots, a_{n_m}^m$, where not all a_j^i are zero. This is contradictory with the definition of direct sum. So we must have linear independence, which imply $y_1^1, \cdots, y_{n_1}^1, \cdots, y_1^m, \cdots, y_{n_m}^m$ form a basis of X. Consequently, $\dim X = \sum \dim Y_i$. □

▶ 12. (page 7)

Proof. Fix a basis $x_1, \cdots, x_n$ of X, any element $x \in X$ can be uniquely represented as $\sum_{i=1}^n \alpha_i(x) x_i$ for some $\alpha_i(x) \in K$, $i = 1, \cdots, n$. We define the isomorphism as $x \mapsto (\alpha_1(x), \cdots, \alpha_n(x))$. Clearly this isomorphism depends on the basis and by varying the choice of basis, we can have different isomorphisms. □

▶ 13. (page 7)

Proof. For any $x_1, x_2 \in X$, if $x_1 \equiv x_2$, i.e. $x_1 - x_2 \in Y$, then $x_2 - x_1 = -(x_1 - x_2) \in Y$, i.e. $x_2 \equiv x_1$. This is symmetry. For any $x \in X$, $x - x = 0 \in Y$. So $x \equiv x$. This is reflexivity. Finally, if $x_1 \equiv x_2$, $x_2 \equiv x_3$, then $x_1 - x_3 = (x_1 - x_2) + (x_2 - x_3) \in Y$, i.e. $x_1 \equiv x_3$. This is transitivity. □

▶ 14. (page 7)

Proof. For any $x_1, x_2 \in X$, we can find $y \in \{x_1\} \cap \{x_2\}$ if and only if $x_1 - y \in Y$ and $x_2 - y \in Y$. Then

$$x_1 - x_2 = (x_1 - y) - (x_2 - y) \in Y.$$

So $\{x_1\} \cap \{x_2\} \neq \emptyset$ if and only if $\{x_1\} = \{x_2\}$. □

▶ 15. (page 8)

Proof. If $\{x\} = \{x'\}$ and $\{y\} = \{y'\}$, then $x - x', y - y' \in Y$. So $(x+y) - (x'+y') = (x - x') + (y - y') \in Y$. This shows $\{x + y\} = \{x' + y'\}$. Also, for any $k \in K$, $kx - kx' = k(x - x') \in Y$. So $k\{x\} = \{kx\} = \{kx'\} = k\{x'\}$. □

▶ 16. (page 9)

Proof. By theory of polynomials, we have

$$Y = \left\{ q(t) \prod_{i=1}^{j} (t - t_i) : q(t) \text{ is a polynomial of degree} < n - j \right\}.$$

Then it's easy to see $\dim Y = n - j$ and $\dim X/Y = \dim X - \dim Y = j$. □

▶ 17. (page 10)

Proof. By Theorem 6, $\dim X/Y = \dim X - \dim Y = 0$, which implies $X/Y = \{\{0\}\}$. So $X = Y$. □

▶ 18. (page 11)

Proof. Define $Y_1 = \{(x, 0) : x \in X_1, 0 \in X_2\}$ and $Y_2 = \{(0, x) : 0 \in X_1, x \in X_2\}$. Then Y_1 and Y_2 are linear subspaces of $X_1 \oplus X_2$. It is easy to see Y_1 is isomorphic to X_1, Y_2 is isomorphic to X_2, and $Y_1 \cap Y_2 = \{(0, 0)\}$. So by Theorem 7, $\dim X_1 \oplus X_2 = \dim Y_1 + \dim Y_2 - \dim(Y_1 \cap Y_2) = \dim X_1 + \dim X_2 - 0 = \dim X_1 + \dim X_2$. □

▶ 19. (page 11)

Proof. By Exercise 18 and Theorem 6, $\dim(Y \oplus X/Y) = \dim Y + \dim(X/Y) = \dim Y + \dim X - \dim Y = \dim X$. Since linear spaces of same finite dimension are isomorphic (by one-to-one mapping between their bases), $Y \oplus X/Y$ is isomorphic to X. □

▶ 20. (page 12)

Proof. (a) is not since $\{x : x_1 \geq 0\}$ is not closed under the scalar multiplication by -1. (b) is. (c) is not since $x_1 + x_2 + 1 = 0$ and $x_1' + x_2' + 1 = 0$ imply $(x_1 + x_1') + (x_2 + x_2') + 1 = -1$. (d) is. (e) is not since x_1 being an integer does not guarantee rx_1 is an integer for any $r \in \mathbb{R}$. □

▶ 21. (page 12)

Proof. (From the textbook's solutions, page 279.) The statement is false; here is an example to the contrary:

$$
\begin{aligned}
&X = \mathbb{R}^2 = (x, y) \text{ space} \\
&U = \{y = 0\}, V = \{x = 0\}, W = \{x = y\}. \\
&U + V + W = \mathbb{R}^2, U \cap V = \{0\}, U \cap W = \{0\} \\
&V \cap W = \{0\}, U \cap V \cap W = 0.
\end{aligned}
$$

□

Chapter 2

Duality

The book's own solution gives answers to Ex 4, 5, 6, 7.

► 1. (page 15)

Proof. We let $Y = \{kx_1 : k \in K\}$. Then Y is a 1-dimensional linear subspace of X. By Theorem 2 and Theorem 4,

$$\dim Y^\perp = \dim X - \dim Y < \dim X = \dim X'$$

So there must exist some $l \in X' \setminus Y^\perp$ such that $l(x_1) \neq 0$. □

Remark 2.1. *When K is $\mathbb{R}$ or $\mathbb{C}$, the proof can be constructive. Indeed, assume $e_1, \cdots, e_n$ is a basis for X and $x_1 = \sum_{i=1}^n a_i e_i$. In the case of $K = \mathbb{R}$, define l by setting $l(e_i) = a_i$, $i = 1, \cdots, n$; in the case of $K = \mathbb{C}$, define l by setting $l(e_i) = a_i^*$ (the conjugate of a_i), $i = 1, \cdots, n$. Then in both cases, $l(x_1) = \sum_{i=1}^n \|a_i\|^2 > 0$.*

► 2. (page 15)

Proof. For any l_1 and $l_2 \in Y^\perp$, we have $(l_1 + l_2)(y) = l_1(y) + l_2(y) = 0 + 0 = 0$ for any $y \in Y$. So $l_1 + l_2 \in Y^\perp$. For any $k \in K$, $(kl)(y) = k(l(y)) = k0 = 0$ for any $y \in Y$. So $kl \in Y^\perp$. Combined, we conclude $Y^\perp$ is a subspace of X'. □

► 3. (page 17)

Proof. Since $S \subset Y$, $Y^\perp \subset S^\perp$. For "$\supset$", let $x_1, \cdots, x_m$ be a maximal linearly independent subset of S. Then $S = \text{span}(x_1, \cdots, x_m)$ and $Y = \{\sum_{i=1}^m \alpha_i x_i : \alpha_1, \cdots, \alpha_m \in K\}$ by Exercise 9 of Chapter 1. By the definition of annihilator, for any $l \in S^\perp$ and $y = \sum_{i=1}^m \alpha_i x_i \in Y$, we have

$$l(y) = \sum_{i=1}^m \alpha_i l(x_i) = 0.$$

So $l \in Y^\perp$. By the arbitrariness of l, $S^\perp \subset Y^\perp$. Combined, we have $S^\perp = Y^\perp$. □

► 4. (page 18)

Proof. Suppose three linearly independent polynomials p_1, p_2 and p_3 are applied to formula (9). Then m_1, m_2 and m_3 must satisfy the linear equations

$$\begin{bmatrix} p_1(t_1) & p_1(t_2) & p_1(t_3) \\ p_2(t_1) & p_2(t_2) & p_2(t_3) \\ p_3(t_1) & p_3(t_2) & p_3(t_3) \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} \int_{-1}^{1} p_1(t)dt \\ \int_{-1}^{1} p_2(t)dt \\ \int_{-1}^{1} p_3(t)dt \end{bmatrix}$$

We take $p_1(t) = 1$, $p_2(t) = t$ and $p_3(t) = t^2$. The above equation becomes

$$\begin{bmatrix} 1 & 1 & 1 \\ -a & 0 & a \\ a^2 & 0 & a^2 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ \frac{2}{3} \end{bmatrix}$$

So

$$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -a & 0 & a \\ a^2 & 0 & a^2 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 0 \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{2a} & \frac{1}{2a^2} \\ 1 & 0 & -\frac{1}{a^2} \\ 0 & \frac{1}{2a} & \frac{1}{2a^2} \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 0 \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3a^2} \\ 2 - \frac{2}{3a^2} \\ \frac{1}{3a^2} \end{bmatrix}$$

Then it's easy to see that for $a > \sqrt{1/3}$, all three weights are positive.

To show formula (9) holds for all polynomials of degree < 6 when $a = \sqrt{3/5}$, we note for any odd $n \in \mathbb{N}$,

$$\int_{-1}^{1} x^n dx = 0, \; m_1 p(-a) + m_3 p(a) = 0 \text{ since } m_1 = m_2 \text{ and } p(-x) = -p(x), \text{ and } m_2 p(0) = 0.$$

So (9) holds for any x^n of odd degree n. In particular, for $p(x) = x^3$ and $p(x) = x^5$. For $p(x) = x^4$, we have

$$\int_{-1}^{1} x^4 dx = \frac{2}{5}, \; m_1 p(t_1) + m_2 p(t_2) + m_3 p(t_3) = 2m_1 a^4 = \frac{2}{3} a^2.$$

So formula (9) holds for $p(x) = x^4$ when $a = \sqrt{3/5}$. Combined, we conclude for $a = \sqrt{3/5}$, (9) holds for all polynomials of degree < 6. □

Remark 2.2. *In this exercise problem and Exercise 5 below, "Theorem 6" is corrected to "Theorem 7".*

► 5. (page 18)

Proof. We take $p_1(t) = 1$, $p_2(t) = t$, $p_3(t) = t^2$, and $p_4(t) = t^3$. Then m_1, m_2, m_3, and m_4 solve the following equation:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -a & -b & b & a \\ a^2 & b^2 & b^2 & a^2 \\ -a^3 & -b^3 & b^3 & a^3 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2/3 \\ 0 \end{bmatrix}$$

Then

$$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -a & -b & b & a \\ a^2 & b^2 & b^2 & a^2 \\ -a^3 & -b^3 & b^3 & a^3 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 0 \\ 2/3 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{b^2}{-2a^2+2b^2} & \frac{b^2}{2a^3-2ab^2} & \frac{1}{2a^2-2b^2} & \frac{1}{-2a^3+2ab^2} \\ \frac{a^2}{2a^2-2b^2} & \frac{a^2}{-2a^2b+2b^3} & \frac{1}{-2a^2+2b^2} & \frac{1}{2a^2b-2b^3} \\ \frac{a^2}{2a^2-2b^2} & \frac{a^2}{2a^2b-2b^3} & \frac{1}{-2a^2+2b^2} & \frac{1}{-2a^2b+2b^3} \\ \frac{b^2}{-2a^2+2b^2} & \frac{b^2}{-2a^3+2ab^2} & \frac{1}{2a^2-2b^2} & \frac{1}{2a^3-2ab^2} \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 2/3 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-3b^2+1}{3(a^2-b^2)} \\ \frac{3a^2-1}{3(a^2-b^2)} \\ \frac{3a^2-1}{3(a^2-b^2)} \\ \frac{-3b^2+1}{3(a^2-b^2)} \end{bmatrix}$$

So the weights are positive if and only if one of the following two mutually exclusive cases hold
1) $b^2 > \frac{1}{3}$, $a^2 < b^2$, $a^2 > \frac{1}{3}$;
2) $b^2 < \frac{1}{3}$, $a^2 > b^2$, $a^2 < \frac{1}{3}$.

□

► 6. (page 18)
(a)

Proof. (From the textbook's solutions, page 280) Suppose there is a linear relation

$$al_1(p) + bl_2(p) + cl_3(p) = 0.$$

Set $p = p(x) = (x - \xi_2)(x - \xi_3)$. Then $p(\xi_2) = p(\xi_3) = 0$, $p_1(\xi_1) \neq 0$; so we get from the above relation that $a = 0$. Similarly $b = 0$, $c = 0$. □

(b)

Proof. Since $\dim \mathcal{P}_2 = 3$, $\dim \mathcal{P}_2' = 3$. Since l_1, l_2, l_3 are linearly independent, they span $\mathcal{P}_2'$. □

(c1)

Proof. We define l_1 by setting

$$l_1(e_j) = \begin{cases} 1, & \text{if } j = 1 \\ 0, & \text{if } j \neq 1 \end{cases}$$

and extending l_1 to V by linear combination, i.e. $l_1(\sum_{j=1}^n \alpha_j e_j) := \sum_{j=1}^n \alpha_j l_1(e_j) = \alpha_1$. l_2, $\cdots$, l_n can be constructed similarly. If there exist a_1, $\cdots$, a_n such that $a_1 l_1 + \cdots + a_n l_n = 0$, we have

$$0 = a_1 l_1(e_j) + \cdots a_n l_n(e_j) = a_j, \ j = 1, \cdots, n.$$

So l_1, $\cdots$, l_n are linearly independent. Since $\dim V' = \dim V = n$, $\{l_1, \cdots, l_n\}$ is a basis of V'. □

(c2)

Proof. We define

$$p_1(x) = \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)},\ p_2(x) = \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)},\ p_3(x) = \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}.$$

□

► 7. (page 18)

Proof. (From the textbook's solutions, page 280) $l(x)$ has to be zero for $x = (1, 0, -1, 2)$ and $x = (2, 3, 1, 1)$. These yield two equations for $c_1, \cdots, c_4$:

$$c_1 - c_3 + 2c_4 = 0,\ 2c_1 + 3c_2 + c_3 + c_4 = 0.$$

We express c_1 and c_2 in terms of c_3 and c_4. From the first equation, $c_1 = c_3 - 2c_4$. Setting this into the second equation gives $c_2 = -c_3 + c_4$. □

Chapter 3

Linear Mappings

The book's own solution gives answers to Ex 1, 2, 4, 5, 6, 7, 8, 10, 11, 13.

★ **Comments**: To memorize Theorem 5 ($R_T^{\perp} = N_{T'}$), recall for a given $l \in U'$, $(l, Tx) = 0$ for any $x \in X$ if and only if $T'l = 0$.

▶ 1. (page 20) Prove Theorem 1.

(a)

Proof. For any $y, y' \in T(X)$, there exist $x, x' \in X$ such that $T(x) = y$ and $T(x') = y'$. So $y + y' = T(x) + T(x') = T(x + x') \in T(X)$. For any $k \in K$, $ky = kT(x) = T(kx) \in T(X)$. Combined, we conclude $T(X)$ is a linear subspace of U. □

(b)

Proof. Suppose V is a linear subspace of U. For any $x, x' \in T^{-1}(V)$, there exist $y, y' \in V$ such that $T(x) = y$ and $T(x') = y'$. Since $T(x + x') = T(x) + T(x') = y + y' \in V$, $x + x' \in T^{-1}(V)$. For any $k \in K$, since $T(kx) = kT(x) = ky \in V$, $kx \in T^{-1}(V)$. Combined, we conclude $T^{-1}(V)$ is a linear subspace of X. □

▶ 2. (page 24)

Proof. (From the textbook's solution, page 280) Suppose we drop the ith equation; if the remaining equations do not determine x uniquely, there is an $x \neq 0$ that is mapped into a vector whose components except the ith are zero. If this were true for all $i = 1, \cdots, m$, the range of the mapping $x \to u$ would be m-dimensional; but according to Theorem 2, the dimension of the range is $\leq n < m$. Therefore one of the equations may be dropped without losing uniqueness; by induction $m - n$ of the equations may be omitted.

Alternative solution: Uniqueness of the solution x implies the column vectors of the matrix $T = (t_{ij})$ are linearly independent. Since the column rank of a matrix equals its row rank (see Chapter 3, Theorem 6 and Chapter 4, Theorem 2), it is possible to select a subset of n of these equations which uniquely determine the solution. □

Remark 3.1. *The textbook's solution is a proof that the column rank of a matrix equals its row rank.*

▶ 3. (page 25) Prove Theorem 3.

(i)

Proof. $S \circ T(ax + by) = S(T(ax + by)) = S(aT(x) + bT(y)) = aS(T(x)) + bS(T(y)) = aS \circ T(x) + bS \circ T(y)$. So $S \circ T$ is also a linear mapping. □

(ii)

Proof. $(R+S) \circ T(x) = (R+S)(T(x)) = R(T(x)) + S(T(x)) = (R \circ T + S \circ T)(x)$ and $S \circ (T+P)(x) = S((T + P)(x)) = S(T(x) + P(x)) = S(T(x)) + S(P(x)) = (S \circ T + S \circ P)(x)$. □

► 4. (page 25)

Proof. For Example 8, the linearity of S and T is easy to see. To see the non-commutativity, consider the polynomial $p(s) = s$. We have $TS(s) = T(s^2) = 2s \neq s = S(1) = ST(s)$. So $ST \neq TS$.

For Example 9, $\forall x = (x_1, x_2, x_3) \in X$, $S(x) = (x_1, x_3, -x_2)$ and $T(x) = (x_3, x_2, -x_1)$. So it's easy to see S and T are linear. To see the non-commutativity, note $ST(x) = S(x_3, x_2, -x_1) = (x_3, -x_1, -x_2)$ and $TS(x) = T(x_1, x_3, -x_2) = (-x_2, x_3, -x_1)$. So $ST \neq TS$ in general. □

Remark 3.2. *Note the problem does not specify the direction of the rotation, so it is also possible that $S(x) = (x_1, -x_3, x_2)$ and $T(x) = (-x_3, x_2, x_1)$. There are total of four choices of (S, T), and each of the corresponding proofs is similar to the one presented above.*

► 5. (page 25)

Proof. $TT^{-1}(x) = T(T^{-1}(x)) = x$ by definition. So $TT^{-1} = id$. □

► 6. (page 25) Prove Theorem 4.

(i)

Proof. Suppose $T : X \to U$ is invertible. Then for any $y, y' \in U$, there exist a unique $x \in X$ and a unique $x' \in X$ such that $T(x) = y$ and $T(x') = y'$. So $T(x + x') = T(x) + T(x') = y + y'$ and by the injectivity of T, $T^{-1}(y + y') = x + x' = T^{-1}(y) + T^{-1}(y')$. For any $k \in K$, since $T(kx) = kT(x) = ky$, injectivity of T implies $T^{-1}(ky) = kx = kT^{-1}(y)$. Combined, we conclude T^{-1} is linear. □

(ii)

Proof. Suppose $T : X \to U$ and $S : U \to V$. First, by the definition of multiplication, ST is a linear map. Second, if $x \in X$ is such that $ST(x) = 0 \in V$, the injectivity of S implies $T(x) = 0 \in U$ and the injectivity of T further implies $x = 0 \in X$. So, ST is one-to-one. For any $z \in V$, there exists $y \in U$ such that $S(y) = z$. Also, we can find $x \in X$ such that $T(x) = y$. So $ST(x) = S(y) = z$. This shows ST is onto. Combined, we conclude ST is invertible.

By associativity, we have $(ST)(T^{-1}S^{-1}) = ((ST)T^{-1})S^{-1} = (S(TT^{-1}))S^{-1} = SS^{-1} = \mathrm{id}_V$. Replace S with T^{-1} and T with S^{-1}, we also have $(T^{-1}S^{-1})(ST) = \mathrm{id}_X$. Therefore, we can conclude $(ST)^{-1} = T^{-1}S^{-1}$. □

► 7. (page 26)

(i)

Proof. Suppose $T : X \to U$ and $S : U \to V$ are linear maps. Then for any given $l \in V'$, $((ST)'l, x) = (l, STx) = (S'l, Tx) = (T'S'l, x)$, $\forall x \in X$. Therefore, $(ST)'l = T'S'l$. Let l run through every element of V', we conclude $(ST)' = T'S'$. □

(ii)

Proof. Suppose T and R are both linear maps from X to U. For any given $l \in U'$, we have $((T+R)'l, x) = (l, (T+R)x) = (l, Tx+Rx) = (l, Tx)+(l, Rx) = (T'l, x)+(R'l, x) = ((T'+R')l, x)$, $\forall x \in X$. Therefore $(T+R)'l = (T'+R')l$. Let l run through every element of V', we conclude $(T+R)' = T' + R'$. □

(iii)

Proof. Suppose T is an isomorphism from X to U, then T^{-1} is a well-defined linear map. We first show T' is an isomorphism from U' to X'. Indeed, if $l \in U'$ is such that $T'l = 0$, then for any $x \in X$, $0 = (T'l, x) = (l, Tx)$. As x varies and goes through every element of X, Tx goes through every element of U. By considering the identification of U with U'', we conclude $l = 0$. So T' is one-to-one. For any given $m \in X'$, define $l = mT^{-1}$, then $l \in U'$. For any $x \in X$, we have $(m, x) = (m, T^{-1}(Tx)) = (l, Tx) = (T'l, x)$. Since x is arbitrary, $m = T'l$ and T' is therefore onto. Combined, we conclude T' is an isomorphism from U' to X' and $(T')^{-1}$ is hence well-defined.

By part (i), $(T^{-1})'T' = (TT^{-1})' = (\mathrm{id}_U)' = \mathrm{id}_{U'}$ and $T'(T^{-1})' = (T^{-1}T)' = (\mathrm{id}_X)' = \mathrm{id}_{X'}$. This shows $(T^{-1})' = (T')^{-1}$. □

▶ 8. (page 26)

Proof. Suppose $\xi : X \to X''$ and $\eta : U \to U''$ are the isomorphisms defined in Chapter 2, formula (5), which identify X with X'' and U with U'', respectively. Then for any $x \in X$ and $l \in U'$, we have

$$(T''\xi_x, l) = (\xi_x, T'l) = (T'l, x) = (l, Tx) = (\eta_{Tx}, l).$$

Since l is arbitrary, we must have $T''\xi_x = \eta_{Tx}$, $\forall x \in X$. Hence, $T'' \circ \xi = \eta \circ T$, which is the precise interpretation of $T'' = T$. □

▶ 9. （page 28)

Proof. If $Bx = 0$, by applying A to both sides of the equation and $AB = I$, we conclude $x = 0$. So B is injective. By Corollary B of Theorem 2, B is surjective. Therefore the inverse of B, denoted by B^{-1}, always exists, and $A = A(BB^{-1}) = (AB)B^{-1} = IB^{-1} = B^{-1}$, which implies $BA = I$. □

Remark 3.3. *For a general algebraic structure, e.g. a ring with unit, it's not always the case that an element's right inverse equals to its left inverse. In the proof above, we used the fact that for finite dimensional linear vector space, a linear mapping is injective if and only if it's surjective.*

▶ 10. (page 30)

Proof. Suppose $K = M_S$. Then $K(M^{-1})_S = SMS^{-1}SM^{-1}S^{-1} = I$. By Exercise 9, K is also invertible and $K^{-1} = (M^{-1})_S$. □

▶ 11. (page 30) Prove Theorem 9.

Proof. Suppose A is invertible, we have $AB = AB(AA^{-1}) = A(BA)A^{-1}$. So AB and BA are similar. The case of B being invertible can be proved similarly. □

▶ 12. (page 31)

Proof. For any $\alpha, \beta \in K$ and $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n)$, we have

$$\begin{aligned}
P(\alpha x + \beta y) &= P((\alpha x_1 + \beta y_1, \cdots, \alpha x_n + \beta y_n)) \\
&= (0, 0, \alpha x_3 + \beta y_3, \cdots, \alpha x_n + \beta y_n) \\
&= (0, 0, \alpha x_3, \cdots, \alpha x_n) + (0, 0, \beta y_3, \cdots, \beta y_n) \\
&= \alpha(0, 0, x_3, \cdots, x_n) + \beta(0, 0, y_3, \cdots, y_n) \\
&= \alpha P(x) + \beta P(y).
\end{aligned}$$

This shows P is a linear map. Furthermore, we have

$$P^2(x) = P((0, 0, x_3, \cdots, x_n)) = (0, 0, x_3, \cdots, x_n) = P(x).$$

So P is a projection. □

▶ 13. (page 31)

Proof. For any $\alpha, \beta \in K$ and $f, g \in C[-1, 1]$, we have

$$\begin{aligned}
P(\alpha f + \beta g)(x) &= \frac{1}{2}[(\alpha f + \beta g)(x) + (\alpha f + \beta g)(-x)] \\
&= \frac{\alpha}{2}[f(x) + f(-x)] + \frac{\beta}{2}[g(x) + g(-x)] \\
&= \alpha P(f)(x) + \beta P(g)(x).
\end{aligned}$$

This shows P is a linear map. Furthermore, we have

$$\begin{aligned}
(P^2 f)(x) &= (P(Pf))(x) = P\left(\frac{f(\cdot) + f(-\cdot)}{2}\right)(x) = \frac{1}{2}\left[\frac{f(x) + f(-x)}{2} + \frac{f(-x) + f(x)}{2}\right] \\
&= \frac{1}{2}(f(x) + f(-x)) = (Pf)(x).
\end{aligned}$$

So P is a projection. □

▶ 14. (page 31)

(a)

Proof. Since $\dim R_T = 1$, it suffices to prove the following claim: *if T is a linear map on a 1-dimensional linear vector space X, there exists a unique number c such that $T(x) = cx$, $\forall x \in X$.* We assume the underlying filed K is either $\mathbb{R}$ or $\mathbb{C}$. We further assume $S : X \to K$ is an isomorphism. Then $S \circ T \circ S^{-1}$ is a linear map on K. Define $c = S \circ T \circ S^{-1}(1)$, we have

$$S \circ T \circ S^{-1}(k) = S \circ T \circ S^{-1}(k \cdot 1) = k \cdot c, \ \forall k \in K.$$

So $T \circ S^{-1}(k) = S^{-1}(c \cdot k) = cS^{-1}(k)$, $\forall k \in K$. This shows T is a scalar multiplication. □

(b)

Proof. If $c \neq 1$, it's easy to verify $I + \frac{1}{1-c}T$ is the inverse of $I - T$. □

▶ 15. (page 31)

Proof. Because $R_{ST} \subset R_S$, $\text{rank}(ST) = \dim(R_{ST}) \leq \dim R_S = \text{rank}(S)$. Moreover, since the column rank of a matrix equals its row rank (see Chapter 3, Theorem 6 and Chapter 4, Theorem 2), we have $\text{rank}(ST) = \text{rank}(T'S') \leq \text{rank}(T') = \text{rank}(T)$. Combined, we conclude $\text{rank}(ST) \leq \min\{\text{rank}(S), \text{rank}(T)\}$.

Also, we note N_{ST}/N_T is isomorphic to $N_S \cap R_T$, with the isomorphism defined by $\phi(\{x\}) = Tx$, where $\{x\} := x + N_T$. It's easy to see ϕ is well-defined, is linear, and is both injective and surjective. So by Theorem 6 of Chapter 1,

$$\dim N_{ST} = \dim N_T + \dim N_{ST}/N_T = \dim N_T + \dim(N_S \cap R_T) \leq \dim N_T + \dim N_S.$$

$\square$

Remark 3.4. *The result* $\textit{rank}(ST) \leq \min\{\textit{rank}(S), \textit{rank}(T)\}$ *is used in econometrics. Cf. Greene [4, page 985] Appendix A.*