Selected Solutions to *Linear Algebra Done Wrong*

HUANG Jing

August 2018

Introduction

Linear Algebra Done Wrong by Sergei Treil is a well-known linear algebra reference book for students. While when I read this book and did the exercises, I found no solution manual was available online. In contrast, the solution manual for the other famous book *Linear Algebra Done Right* can be easily found as it is much more popular than *Dong Wrong* (without any doubt, *Dong Wrong* is an excellent book worth reading). Insights tend to hide behind mistakes and ambiguity. Without a good way to check our answers, we may miss them. The reference solution also helps to give inspiration and save time when there is no hint or the hint is not sufficient.

Anyway, I scanned all and did most of the problems of the first 6 chapters (2014 version) and share those I think valuable (those are relatively hard or insightful from my perspective, I read this book for kind of reviewing and deeper mathematical understanding). The rest problems should be solvable even for new learners. Of course, there is no guarantee of the correctness of these solutions and they also may not be the best, either. One important thing is that the material is meant for reference to facilitate learning, especially self-learning. DO NOT copy the contents for the homework as it will violate the academic codes of your college. This really matters.

As I'm now a PhD student with limited time for this project, the update is sporadic and your advice and contributions are most welcomed. Contact me at `huangjingonly@gmail.com` and we can make this project better and helpful for those in need.

Chapter 1. Basic Notations

1.4. Prove that a zero vector of a vector space V is unique.
Proof Suppose there exist 2 different zero vectors $\mathbf{0}_1$ and $\mathbf{0}_2$. So for any $\mathbf{v} \in V$, we have

$$\mathbf{v} + \mathbf{0}_1 = \mathbf{v}$$
$$\mathbf{v} + \mathbf{0}_2 = \mathbf{v}$$

Find the difference of the equations above, we get

$$\mathbf{0}_2 - \mathbf{0}_1 = \mathbf{v} - \mathbf{v} = \mathbf{0}_1 \,/(\text{or})\, \mathbf{0}_2$$

then

$$\mathbf{0}_2 = \mathbf{0}_1 + \mathbf{0}_1/\mathbf{0}_2 = \mathbf{0}_1$$

So the zero vector is unique.

1.6. Prove that the additive inverse, defined in Axiom 4 of a vector space is unique.
Proof Assume there exist 2 different additive inverses $\mathbf{w}_1$ and $\mathbf{w}_2$ of vector $\mathbf{v} \in V$. Then

$$\mathbf{v} + \mathbf{w}_1 = \mathbf{0}$$
$$\mathbf{v} + \mathbf{w}_2 = \mathbf{0}$$

Obtain the difference of the two equations, we get

$$\mathbf{w}_1 - \mathbf{w}_2 = \mathbf{0}$$

then

$$\mathbf{w}_1 = \mathbf{w}_2$$

So the additive inverse is unique.

1.7. Prove that $0\mathbf{v} = \mathbf{0}$ for any vector $\mathbf{v} \in V$.
Proof $0\mathbf{v} = (0\alpha)\mathbf{v} = \alpha(0\mathbf{v})$ for any scalar α, so $(1-\alpha)0\mathbf{v} = 0$ for all scalar $(1-\alpha)$. Then $0\mathbf{v} = \mathbf{0}$.

1.8. Prove that for any vector $\mathbf{v}$ its additive inverse $-\mathbf{v}$ is given by $(-1)\mathbf{v}$.
Proof $\mathbf{v} + (-1)\mathbf{v} = (1-1)\mathbf{v} = 0\mathbf{v} = \mathbf{0}$ and we know form Problem 1.6 that the additive inverse is unique. So $-\mathbf{v} = (-1)\mathbf{v}$.

2.5. Let a system of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$ be linearly independent but not generating. Show that it is possible to find a vector $\mathbf{v}_{r+1}$ such that the system $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r, \mathbf{v}_{r+1}$ is linear independent.
Proof Take $\mathbf{v}_{r+1}$ that can not be represented as $\sum_{k=1}^{r} \alpha_k \mathbf{v}_k$. It is possible because $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$ is not generating. Now we need to show $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r, \mathbf{v}_{r+1}$ is linear independent. Suppose that $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r, \mathbf{v}_{r+1}$ is linear dependent, i.e.

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_r\mathbf{v}_r + \alpha_{r+1}\mathbf{v}_{r+1} = \mathbf{0}$$

and $\sum_{k=1}^{r+1} |\alpha_k| \neq 0$. If $\alpha_{r+1} = 0$, then

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_r\mathbf{v}_r = \mathbf{0}$$

and $\sum_{k=1}^{r} |\alpha_k| \neq 0$. This contradicts that $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$ is linearly independent. So $\alpha_{r+1} \neq 0$. Thus $\mathbf{v}_{r+1}$ can be represented as

$$\mathbf{v}_{r+1} = -\frac{1}{\alpha_{r+1}} \sum_{k=1}^{r} \alpha_k \mathbf{v}_k$$

This contradicts the premise that $\mathbf{v}_{r+1}$ can not be represented by $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$. Thus, the system $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r, \mathbf{v}_{r+1}$ is linearly independent.

3.2. Let a linear transformation in $\mathbb{R}^2$ be in the line $x_1 = x_2$. Find its matrix.
Solution 1. Reflection is a linear transformation. It is completely defined on the standard basis. And $\mathbf{e}_1 = [1\ 0]^\mathsf{T} \overset{T}{\Rightarrow} \mathbf{r}_1 = [0\ 1]^\mathsf{T}$, $\mathbf{e}_2 = [0\ 1]^\mathsf{T} \overset{T}{\Rightarrow} \mathbf{r}_2 = [1\ 0]^\mathsf{T}$. So the matrix is the combination of the two transformed standard basis as its first and second column. i.e.

$$T = [\mathbf{r}_1\ \mathbf{r}_2] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Solution 2. (A more general method) Let α be the angle between the x-axis and the line. The reflection can be achieved through following steps: first, rotate the line around the origin (z-axis in 3D space) $-\alpha$ so the line aligns with the x-axis (This line happen to pass through the origin, if not, translation is needed in advance to make the line pass through the origin and we need to use *homogeneous coordinates* since translation is not a linear transformation if represented in standard coordinates). Secondly, perform reflection about the x-axis. Lastly, we need to rotate the current frame back to its original location or perform other corresponding inverse transformation. So

$$T = Rotz(-\alpha) \cdot Ref \cdot Rotz(\alpha)$$

That is

$$\begin{aligned} T &= \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \\ &= \begin{bmatrix} \cos(-\frac{\pi}{4}) & -\sin(-\frac{\pi}{4}) \\ \sin(-\frac{\pi}{4}) & \cos(-\frac{\pi}{4}) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{aligned}$$

3.7. Show that any linear transformation in $\mathbb{C}$ (treated as a complex vector space) is a multiplication by $\alpha \in \mathbb{C}$.
Proof Suppose a linear transformation $T : \mathbb{C} \Rightarrow \mathbb{C}$. $T(1) = a + ib$ and then

$T(-1) = -T(1) = -a - ib$. Note that $i^2 = -1$. Then $T(-1) = T(i^2) = iT(i)$. Thus

$$T(i) = \frac{-a - ib}{i} = i(a + ib)$$

So for any $\omega = x + iy \in \mathbb{C}$,

$$\begin{aligned} T(\omega) &= T(x + iy) = xT(1) + yT(i) \\ &= x(a + ib) + yi(a + ib) \\ &= (x + iy)(a + ib) \\ &= \omega T(1) \\ &= \omega\alpha \end{aligned}$$

and $\alpha = T(1)$.

5.4. Find the matrix of the orthogonal projection in $\mathbb{R}^2$ onto the line $x_1 = -2x_2$.

Solution

$$\begin{aligned} T &= R(\alpha)PR(-\alpha) \\ &= R(\alpha)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} R(-\alpha) \end{aligned}$$

and $\alpha = \tan^{-1}(-\frac{1}{2})$ so we can get the matrix is

$$T = \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix}$$

5.7. Find the matrix of the reflection through the line $y = -2x/3$. Perform all the multiplications.

Solution *(Similar to 5.4 though not exactly the same.)*

$$\begin{aligned} T &= R(\alpha)RefR(-\alpha) \\ &= R(\alpha)\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} R(-\alpha) \end{aligned}$$

and $\alpha = \tan^{-1}(-\frac{2}{3})$ so we can get the matrix is

$$T = \begin{bmatrix} \frac{5}{13} & -\frac{12}{13} \\ -\frac{12}{13} & -\frac{5}{13} \end{bmatrix}$$

6.1. Prove that if $A : V \to W$ is an isomorphism (i.e. an invertible linear transformation) and $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is a basis in V, then $A\mathbf{v}_1, A\mathbf{v}_2, ..., A\mathbf{v}_n$ is

a basis in W.
Proof Any $w \in W$, $A^{-1}w = v \in V$ and

$$\begin{aligned} w = Av &= A[\mathbf{v}_1\ \mathbf{v}_2\ ...\ \mathbf{v}_n][v_1\ v_2\ ...\ v_n]^\mathsf{T} \\ &= [A\mathbf{v}_1\ A\mathbf{v}_2\ ...\ A\mathbf{v}_n][v_1\ v_2\ ...\ v_n]^\mathsf{T} \end{aligned}$$

So we can see $[A\mathbf{v}_1\ A\mathbf{v}_2\ ...\ A\mathbf{v}_n]$ is in the form of a basis in W. Next we show that $A\mathbf{v}_1\ A\mathbf{v}_2\ ...\ A\mathbf{v}_n$ is linearly independent. If not, suppose $A\mathbf{v}_1$ can be expressed as a linear combination of $A\mathbf{v}_2\ A\mathbf{v}_3\ ...\ A\mathbf{v}_n$ without loss of generality. Multiplying them with A in the left side, it results in that $\mathbf{v}_1$ can be expressed by $\mathbf{v}_2\ ...\ \mathbf{v}_n$, which contradicts the fact that $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is a basis in V. So the proposition is proved.

7.4. Let $\mathbf{X}$ and $\mathbf{Y}$ be subspaces of a vector space $\mathbf{V}$. Using the previous exercise, show that $\mathbf{X} \cup \mathbf{Y}$ is a subspace if and only if $\mathbf{X} \subset \mathbf{Y}$ or $\mathbf{Y} \subset \mathbf{X}$.
Proof The sufficiency is obvious and easy to verify. For the necessity, suppose $\mathbf{X} \not\subseteq \mathbf{Y}$ nor $\mathbf{Y} \not\subseteq \mathbf{X}$ and $\mathbf{X} \cup \mathbf{Y}$ is a subspace of $\mathbf{V}$. Then there are vectors $\mathbf{x} \in \mathbf{X}$, $\mathbf{y} \in \mathbf{Y}$ and $\mathbf{x} \notin \mathbf{Y}$, $\mathbf{y} \notin \mathbf{X}$. According to Problem 7.3, $\mathbf{x} + \mathbf{y} \notin \mathbf{X}$, $\mathbf{x} + \mathbf{y} \notin \mathbf{Y}$. So, $\mathbf{x} + \mathbf{y} \notin \mathbf{X} \cup \mathbf{Y}$. i.e., $\mathbf{x} \in \mathbf{X} \cup \mathbf{Y}$, $\mathbf{y} \in \mathbf{X} \cup \mathbf{Y}$, but $\mathbf{x} + \mathbf{y} \notin \mathbf{X} \cup \mathbf{Y}$, which contradicts $\mathbf{X} \cup \mathbf{Y}$ is a subspace. Thus, $\mathbf{X} \subset \mathbf{Y}$ or $\mathbf{Y} \subset \mathbf{X}$.

8.5. A transformation T in $\mathbb{R}^3$ is a rotation about the line $y = x + 3$ in the x-y plane through an angle γ. Write a 4×4 matrix corresponding to this transformation.
You can leave the result as a product of matrices.
Solution For a general spatial rotation around a given direction (suppose the direction is given by a vector) through an angle γ, the 3×3 rotation matrix can be given by:

$$R = R_x^{-1} R_y^{-1} R_z(\gamma) R_y R_x,$$

where the rotation by γ is assumed to be performed around z-axis. R_x and R_y are rotations used to align the direction with z-axis and can be determined by simple trigonometry.
For the problem given, the line $y = x + 3$ doesn't go through the origin, so extra step T_0 is needed to translate the line to make it pass the origin and homogeneous coordinates are applied:

$$R = T_0^{-1} R_x^{-1} R_y^{-1} R_z(\gamma) R_y R_x T_0.$$

According to the description,

$$\begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix} = R_x^{-1} R_y^{-1} R_z(\gamma) R_y R_x.$$

The corresponding matrix is then

$$R = T_0^{-1} \begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix} T_0.$$

T_0 is not unique for the translation to make two parallel lines align.

Chapter 2. Systems of Linear Equations

3.8. Show that if the equation $A\mathbf{x} = \mathbf{0}$ has unique solution (i.e. if echelon form of A has pivot in every column), then A is left invertible.
Proof $A\mathbf{x} = \mathbf{0}$ has unique solution, then the solution is trivial solution. The echelon form of A has pivot at every column. A is $m \times n$ matrix, then $m \geq n$. The row number is greater or equal to the column number. The reduced echelon form of A is denoted as

$$A_{re} = \begin{bmatrix} I_{n\times n} \\ \mathbf{0}_{(m-n)\times n} \end{bmatrix}.$$

And suppose A_{re} is obtained by a sequence of elementary row operation $E_1, E_2, ..., E_k$,

$$A_{re} = E_k \;...\; E_2 E_1 A$$

E_i is $m \times m$. The left inverse of A is the first n rows of the product of E_i. i.e.

$$E_{left} = I_{n\times m} E_k \;...\; E_2 E_1,$$

where

$$I_{n\times m} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \dots \end{bmatrix}_{n\times m},$$

is used to extract the $I_{n\times n}$ identity matrix in A_{re}. $E_{left}A = I$ so A is left invertible.

5.5. Let vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ be a basis in V. Show that $\mathbf{u}+\mathbf{v}+\mathbf{w}$, $\mathbf{u}+\mathbf{v}$, $\mathbf{w}$ is also a basis in V.
Solution For any vector $\mathbf{x} \in V$, suppose $\mathbf{x} = x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w}$. Then it is easy to figure out the $\mathbf{x} = x_1(\mathbf{u}+\mathbf{v}+\mathbf{w}) + (x_2 - x_1)(\mathbf{v}+\mathbf{w}) + (x_3 - x_2 - x_1)\mathbf{w}$. Thus $\mathbf{u}+\mathbf{v}+\mathbf{w}$, $\mathbf{u}+\mathbf{v}$, $\mathbf{w}$ is also a basis in V.

7.4. Prove that is $A : X \to Y$ and V is a subspace of X then dim $AV \leq$ rank A. (AV here means the subspace V transformed by the transformation A, i.e., any vector in AV can be represented as $A\mathbf{v}, \mathbf{v} \in V$). Deduce from here that rank$(AB) \leq$ rank$A$.

Proof $\dim AV \le \dim AX \le \dim \operatorname{Ran} A = \operatorname{rank} A$
Suppose that the column vectors of A compose a basis of space V. Then $\operatorname{rank}(AB) \le \operatorname{rank} A$.

7.5. Prove that if $A : X \to Y$ and V is a subspace of X then $\dim AV \le \dim V$. Deduce from here that $\operatorname{rank}(AB) \le \operatorname{rank} B$.
Proof Suppose $\dim V = k$ and let $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ be a basis of V. AV is defined by $A\mathbf{v}_1, A\mathbf{v}_2, ..., A\mathbf{v}_k$. $\dim AV = \operatorname{rank}\,[A\mathbf{v}_1, A\mathbf{v}_2, ..., A\mathbf{v}_k] \le k = \dim V$.

Similarly, assume $\operatorname{rank} B = k$ and $\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_k$ are linearly independent column vectors in B. Then $\operatorname{rank} AB = \operatorname{rank}\,[A\mathbf{b}_1, A\mathbf{b}_2, ..., A\mathbf{b}_k] \le k = \operatorname{rank} B$.

7.6. Prove that if the product AB of two $n \times n$ matrices is invertible, then both A and B are invertible. Do not use determinant for this problem.
Proof AB is invertible, $\operatorname{rank}(AB) = n$. From Problem 7.5, we have $\operatorname{rank}(AB) = n \leqslant \operatorname{rank}(A) \leqslant n$. Thus $\operatorname{rank}(A) = n$. So is B. A, B have full rank and are invertible.

7.7. Prove that if $A\mathbf{x} = \mathbf{0}$ has unique solution, then the equation $A^{\mathsf{T}}\mathbf{x} = \mathbf{b}$ has a solution for every right side $\mathbf{b}$. (*Hint:* count pivots)
Proof Suppose $A \in \mathbb{R}^m \times \mathbb{R}^n$. Note that for $A\mathbf{x} = \mathbf{0}$, there is always a trivial solution $\mathbf{x} = \mathbf{0} \in \mathbb{R}^n$. And we know the trivial solution is unique, which also indicates that the echelon form of A has a pivot at every column. Accordingly, the echelon form of A^{T} has a pivot at every row (Think that the echelon form of A^{T} is completed by column reduction that corresponds to the row reduction of A). So $A\mathbf{x} = \mathbf{b}$ is consistent for any $\mathbf{b}$.

7.14. Is it possible for a real matrix A that $\operatorname{Ran} A = \operatorname{Ker} A^{\mathsf{T}}$? Is it possible for a complex A?
Solution Both are not possible. Suppose A is $m \times n$ and $\operatorname{Ran} A = \operatorname{Ker} A^{\mathsf{T}}$. Then $\operatorname{Ran} A \subset \operatorname{Ker} A^{\mathsf{T}}$, i.e., $A^{\mathsf{T}}A\mathbf{v} = \mathbf{0}$ for any $\mathbf{v} \in \mathbb{R}^n$. This holds only when $A^{\mathsf{T}}A = 0_{n\times n}$. Then $A = 0_{m\times n}$. (Use the row vectors of A^{T} and check the diagonal entries of $A^{\mathsf{T}}A$ equal to 0. It will lead to the conclusion that the row vectors are all zero vector.)

On the other hand, if $\operatorname{Ran} A = \operatorname{Ker} A^{\mathsf{T}}$, $\operatorname{Ker} A^{\mathsf{T}} \subset \operatorname{Ran} A$. i.e., if $A^{\mathsf{T}}\mathbf{b} = \mathbf{0}$, then the function $A\mathbf{x} = \mathbf{b}$ has a solution. But we have $A = 0_{m\times n}$, then for arbitrary $b \in \mathbb{R}^m$, $A^{\mathsf{T}}\mathbf{b} = \mathbf{0}$ holds. But for $\mathbf{b} \neq \mathbf{0}$, $A\mathbf{x} = \mathbf{b}$ does not have a solution. This is contradictory. So it is not possible for the real or complex matrix A that $\operatorname{Ran} A = \operatorname{Ker} A^{\mathsf{T}}$.

8.5. Prove that if A and B are similar matrices then trace A = trace B. (*Hint:* recall how trace(XY) and trace(YX) are related.)
Proof $\operatorname{trace}(A) = \operatorname{trace}(Q^{-1}BQ) = \operatorname{trace}(Q^{-1}QB) = \operatorname{trace}(B)$. (Note that

trace(AB) = trace(BA) as long as AB, BA can be performed.)

Chapter 3. Determinants

3.4. A square matrix ($n \times n$) is called skew-symmetric (or antisymmetric) if $A^\mathsf{T} = A$. Prove that if A is skew-symmetric and n is odd, then $\det A = 0$. Is this true for even n?
Proof $\det A = \det A^\mathsf{T} = \det(-A) = (-1)^n \det A$ by using the the properties of determinant and skew-symmetric matrices. If n is odd, $(-1)^n = -1$, we have $\det A = -\det A$, thus $\det A = 0$.

If n is even, we just have $\det A = \det A$ so this conclusion generally is not true.

3.5. A square matrix is called *nilpotent* if $A^k = \mathbf{0}$ for some positive integer k. Show that for a nilpotent matrix A, $\det A = 0$.
Proof $\det A^k = (\det A)^k = \det \mathbf{0} = 0$, thus $\det A = 0$.

3.6. Prove that if A and B are similar, then $\det A = \det B$.
proof A and B are similar, then $A = Q^{-1}BQ$ for an invertible matrix Q.Then

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

3.7. A real square matrix Q is called orthogonal if $Q^\mathsf{T}Q = I$. Prove that if Q is an orthogonal matrix then $\det Q = \pm 1$.
Proof $\det Q^\mathsf{T}Q = (\det Q^\mathsf{T})(\det Q) = (\det Q)^2 = \det I = 1$. Thus $\det Q = \pm 1$.

3.9. Let points A, B and C in the plane $\mathbb{R}^2$ have coordinates (x_1, y_1), (x_2, y_2) and (x_3, y_3) respectively. Show that the area of triangle ABC is the absolute value of

$$\frac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}.$$

Hint: use row operation and geometric interpretation of 2×2 determinants (area).

Proof The area of triangle ABC is half of the parallelogram defined by neighbouring sides AB, AC. which also can be computed by

$$\begin{aligned} S_{\triangle ABC} &= \frac{1}{2}\begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix} \\ &= \frac{1}{2}|(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)|, \end{aligned}$$

In the same time, if we use row reduction to check the determinant

$$\begin{aligned} \frac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} &= \frac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 0 & x_2 - x_1 & y_2 - y_1 \\ 0 & x_3 - x_1 & y_3 - y_1 \end{vmatrix} \\ &= \frac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 0 & x_2 - x_1 & y_2 - y_1 \\ 0 & 0 & y_3 - y_1 - (y_2 - y_1)\frac{x_3 - x_1}{x_2 - x_1} \end{vmatrix} . \\ &= \frac{1}{2}((x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)), \end{aligned}$$

We assume that $x_2 - x_1 \neq 0$ and it can be verified if $x_2 - x_1 = 0$, the result still holds. Thus we can see the conclusion holds.

3.10. Let A be a square matrix. Show that block triangular matrices

$$\begin{bmatrix} I & * \\ \mathbf{0} & A \end{bmatrix} \quad \begin{bmatrix} A & * \\ \mathbf{0} & I \end{bmatrix} \quad \begin{bmatrix} I & \mathbf{0} \\ * & A \end{bmatrix} \quad \begin{bmatrix} A & \mathbf{0} \\ * & I \end{bmatrix}$$

all have determinant equal to $\det A$. Here $*$ can be anything.
Proof Considering performing row reduction to make A be triangular, the whole matrix will also be triangular and the rest part on the diagonal is just I. Thus the determinant of the block matrix equals to $\det A$.
(Problem 3.11 and 3.12 are just applications of the conclusion of Problem 3.10. The hint just tells the answer.)

4.2. Let P be a *permutation matrix*, i.e., an $n \times n$ matrix consisting of zeros and ones and such that there is exactly one 1 in every row and every column.

a) Can you describe the corresponding linear transformation? That will explain the name.

b) Show that P is invertible. Can you describe P^{-1}?

c) Show that for some $N > 0$

$$P^N := \underbrace{PP \dots P}_{N \text{ times}} = I.$$

Use the fact that there are only finitely many permutations.
Solution a) Consider the linear transformation $\mathbf{y} = P\mathbf{x}$ and each row of P. There is only one 1 in each row of P. Suppose in the first row of P, $P_{1,j} = 1$, then $y_1 = \mathbf{p_1}\mathbf{x} = x_j$, where $\mathbf{p_1}$ is the first row of P. Namely x_j is moved to the 1st place after the linear transformation. Similarly, for the 2nd row of P, suppose $P_{2,k} = 1$, then $y_2 = x_k$, x_k is moved to the 2nd place, so on and so forth. There is also only 1 for each column, then we know after multiplying by the permutation matrix P, the elements in $\mathbf{x}$ change their order.
b) Suppose P is invertible, by multiplying P^{-1}, $\mathbf{x} = P^{-1}\mathbf{y}$. But we know $y_1 = x_j$, then we have $P^{-1}_{j,1} = 1$ so that x_j can return to its original position. Similarly, $y_2 = x_k$, then $P^{-1}_{k,2} = 1$. Following this we can see that $P^{-1}_{i,j} = P_{j,i}$ if $P_{j,i} = 1$ and the rest are all 0. So we can see P is invertible and $P^{-1} = P^{\mathsf{T}}$.
c) Note that $P\mathbf{x}, P^2\mathbf{x}, P^3\mathbf{x} \dots P^N\mathbf{x}$ are all permutations of $(x_1, x_2, ..., x_n)$. If P^N can never equal to I, $P\mathbf{x}, P^2\mathbf{x}, P^3\mathbf{x} \dots P^N\mathbf{x}$ will be different permutations. And N can be infinitely big, so there will be infinitely many permutations of $(x_1, x_2, ..., x_n)$, which is impossible. Thus there must be some $N > 0$, $P^N = I$.

Exercises Prat 5 and Part 7 in this chapter are normal. So I try to give some ideas and answers for reference:

- Problem 5.3, we can use the last column expansion and the left matrix $(A+tI)_{i,j}$ is a triangular matrix. The final expression is $\det(A+tI) = a_0 + a_1 t + a_2 t^2 + ... + a_{n-1}t^{n-1}$. The order of -1 in each term is even.

- Problem 5.7, $n!$ multiplications is needed. We can use induction to prove it.

- Problem 7.4 and Problem 7.5, consider $\det RA = (\det R)(\det A) = \det A$, where R is the rotation matrix with its determinant equal to 1. For proof of the parallelogram area, we can also use parameter angle, i.e., $\mathbf{v_1} = [x_1, y_1]^{\mathsf{T}} = [v_1\cos\alpha,\ v_1\sin\alpha]^{\mathsf{T}}$, $\mathbf{v_2} = [x_2, y_2]^{\mathsf{T}} = [v_2\cos\beta,\ v_2\sin\beta]^{\mathsf{T}}$. v_1, v_2 are the lengths of $\mathbf{v_1}, \mathbf{v_2}$, respectively. α, β represents the angle between the vector and x-axis positive direction. Then

  $$\begin{aligned} \det A = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} &= x_1y_2 - x_2y_1 \\ &= v_1v_2(\cos\alpha\sin\beta - \cos\beta\sin\alpha) \\ &= v_1v_2\sin(\beta - \alpha). \end{aligned}$$

  $\beta - \alpha$ is the angle from $\mathbf{v_1}$ to $\mathbf{v_2}$.

Chapter 4. Introduction to Spectral Theory (Eigenvalues and Eigenvectors)

1.1. (Part) True or false:

b) If a matrix has one eigenvector, it has infinitely many eigenvectors;
True, if $A\mathbf{x} = \lambda\mathbf{x}$, $A(\alpha\mathbf{x}) = \lambda(\alpha\mathbf{x})$, α is an arbitrary scalar.

c) There exists a square matrix with no real eigenvalues;
True, like the 2D rotation matrix $R_\alpha, \alpha \neq n\pi$.

d) There exists a square matrix with no (complex) eigenvectors;
False, when discussing in complex space, there are always eigenvalues and as a result $A - \lambda I$ has nonempty null space.

f) Similar matrices always have the same eigenvectors;
False, if A, B are similar and $A = SBS^{-1}$. If $A\mathbf{x} = \lambda\mathbf{x}$, then $SBS^{-1}\mathbf{x} = \lambda\mathbf{x}$. i.e., $B(S^{-1}\mathbf{x}) = \lambda(S^{-1}\mathbf{x})$, $S^{-1}\mathbf{x}$ is an eigenvector of B, not $\mathbf{x}$.

g) The sum of two eigenvectors of a matrix A is always an eigenvector;
False

1.6. An operator A is called *nilpotent* if $A^k = \mathbf{0}$ for some K. Prove that if A is nilpotent, then $\sigma(A) = \{0\}$ (i.e. that 0 is the only eigenvalue of A).
Proof Note that if λ is a nonzero eigenvalue of A and $A\mathbf{x} = \lambda\mathbf{x}$. Then $A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda^2\mathbf{x}$, $A^3\mathbf{x} = A(\lambda^2\mathbf{x}) = \lambda^3\mathbf{x}$ ... $A^k\mathbf{x} = \lambda^k\mathbf{x}$. That is to say if $\lambda \in \sigma(A), \lambda^k \in \sigma(A^k)$. Now $A^k = \mathbf{0}$, $\sigma(A^k) = \{0\}$. Then 0 is the only eigenvalue of A.

1.8. Let $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ be a basis in a vector space V. Assume also that the first k vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ of the basis are eigenvectors of an operator A, corresponding to an eigenvalue λ (i.e. that $A\mathbf{v}_j = \lambda\mathbf{v}_j, j = 1, 2, ..., k$). Show that in this basis the matrix of the operator A has block triangular form

$$\begin{bmatrix} \lambda I_k & * \\ \mathbf{0} & B \end{bmatrix}$$

Proof $A_{VV} = [I]_{VS}A_{SS}[I]_{SV}$, where S represents the standard basis. $[I]$ is the coordinate change matrix and $[I]_{SV} = [\mathbf{v}_1^\mathsf{T}, \mathbf{v}_2^\mathsf{T}, ..., \mathbf{v}_n^\mathsf{T}]$. $[I]_{SV}A_{VV} = A_{SS}[I]_{SV} = A_{SS}[\mathbf{v}_1^\mathsf{T}, \mathbf{v}_2^\mathsf{T}, ..., \mathbf{v}_n^\mathsf{T}] = [\lambda\mathbf{v}_1^\mathsf{T}, \lambda\mathbf{v}_2^\mathsf{T}, ..., \lambda\mathbf{v}_k^\mathsf{T}, ..., A_{SS}\mathbf{v}_n^\mathsf{T}]$. Denote the i-th column of A_{VV} with $\mathbf{a}_i$. Consider $\mathbf{a}_1$, then $[I]_{SV}\mathbf{a}_1 = \lambda\mathbf{v}_1^\mathsf{T}$. Since $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is a basis, then $\mathbf{a}_1$ can only be the form $\mathbf{a}_1 = [\lambda, 0, 0, ..., 0]^\mathsf{T}$. Similarly, check the first k columns of A_{VV}, they are λ times the first k standard base vector. So A_{SS} has the block triangular form above.

1.9. Use the two previous exercises to prove that geometric multiplicity

of an eigenvalue cannot exceed its algebraic multiplicity.
Proof We consider the problem in the basis $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ and A has the block triangular form shown in Problem 1.8. Note k is the number of linearly independent eigenvectors corresponding to λ_k, which is also the dimension of $\mathrm{Ker}(A - \lambda_k I)$ (consider the queation $(A - \lambda_k I)\mathbf{x} = \mathbf{0}$). Namely, k is the geometric multiplicity of λ_k.

For the algebraic multiplicity, consider the determinant

$$\begin{aligned}\det(A - \lambda I) &= \begin{vmatrix} (\lambda_k - \lambda)I_k & * \\ \mathbf{0} & B - \lambda I_{n-k} \end{vmatrix} \\ &= (\lambda_k - \lambda)^k \det(B - \lambda I_{n-k}).\end{aligned}$$

So the algebraic multiplicity of λ_k is at least k. It is further possible that λ_k is a root of the polynomial $\det(B - \lambda I_{n-k})$, then in this case the algebraic multiplicity will just exceed k. Thus geometric multiplicity of an eigenvalue cannot exceed its algebraic multiplicity.

1.10. Prove that determinant of a matrix A is the product of its eigenvalues (counting multiplicity).
Proof (Just use the hint) The characteristic polynomial of $n \times n$ square matrix A is $\det(A - \lambda I)$ and we consider the roots of it in complex space. According to the fundamental theorem of algebra, $\det(A - \lambda I)$ has n roots counting multiplicity and can be factorized as $\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda - \lambda_n)$. (Recall the formal definition of determinant, the highest order term of λ, λ^n, is generated by the diagonal product $\Pi_{i=1}^n (a_{ii} - \lambda)$. Thus the sign of the factorization is correct.) Then let $\lambda = 0$, we will get $\det A = \lambda_1 \lambda_2 ... \lambda_n$.

1.11. Prove that the trace a matrix equals the sum of eigenvalues in three steps. First, compute the coefficient of λ^{n-1} in the right side of the equality

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda_n - \lambda).$$

Then show that $\det(A - \lambda I)$ can be represented as

$$\det(A - \lambda I) = (a_{1,1} - \lambda)(a_{2,2} - \lambda)...(a_{n,n} - \lambda) + q(\lambda).$$

where $q(\lambda)$ is polynomial of degree at most $n - 2$. And finally, comparing the coefficients of λ^{n-1} get the conclusion.
Proof First, recall the binomial theorem, the coefficient of λ^{n-1} in $\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda_n - \lambda)$ is $C(\lambda^{n-1}) = (-1)^{n-1}(\lambda_1 + \lambda_2 + ... + \lambda_n)$. Because to get the term λ^{n-1}, we need to pick $-\lambda$ n times from the total n factors $\lambda_i - \lambda$. Then the last one pick is λ_j from the factor whose $-\lambda$ is not picked. The resulting term is then $(-1)^{n-1}\lambda_j \lambda^{n-1}$. There are n combinations and the sum of each term is $C(\lambda^{n-1})\lambda^{n-1}$.

Then, we show $\det(A - \lambda I)$ can be represented as

$$\det(A - \lambda I) = (a_{1,1} - \lambda)(a_{2,2} - \lambda)...(a_{n,n} - \lambda) + q(\lambda).$$

That is to say in $\det(A - \lambda I)$, the term λ^{n-1} are all from $(a_{1,1} - \lambda)(a_{2,2} - \lambda)...(a_{n,n} - \lambda)$. This holds because λ only appears on the diagonal of $A - \lambda I$. Using the formal definition of determinant, if we pick $n - 1$ diagonal term with λ, then the last pick must also be on the diagonal. There is no other way to generate λ^{n-1}. Thus $q(\lambda)$ is a polynomial of degree at most $n - 2$. Then we know the coefficient of λ^{n-1} also equals to $C(\lambda^{n-1}) = (-1)^{n-1}(a_{1,1} + a_{2,2} + ... + a_{n,n})$.

The coefficients derived from two different ways are identical, so we have $\sum_{i=1}^{n} a_{i,i} = \sum_{i=1}^{n} \lambda_i$, namely, the trace a matrix equals the sum of eigenvalues.

2.1. Let A be $n \times n$ matrix. True or false:

a) A^{T} has the same eigenvalues as A.
True, $\det(A - \lambda I) = \det(A - \lambda I)^{\mathsf{T}} = \det(A^{\mathsf{T}} - \lambda I)$

b) A^{T} has the same eigenvectors as A.
False.

c) If A is diagonalizable, then so is A^{T}.
True, $A = SDS^{-1}, A^{\mathsf{T}} = (SDS^{-1})^{\mathsf{T}} = (S^{-1})^{\mathsf{T}} D^{\mathsf{T}} S^{\mathsf{T}} = (S^{\mathsf{T}})^{-1} D S^{\mathsf{T}}$.

2.2. Let A be a square matrix with real entries, and let λ be its complex eigenvalue. Suppose $\mathbf{v} = (v_1, v_2, ..., v_n)^{\mathsf{T}}$ is a corresponding eigenvector, $A\mathbf{v} = \lambda\mathbf{v}$. Prove that the $\bar{\lambda}$ is an eigenvalue of A and $A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$. Here $\bar{\mathbf{v}}$ is the complex conjugate of the vector $\mathbf{v}$, $\bar{\mathbf{v}} := (\bar{v_1}, \bar{v_2}, ..., \bar{v_n})^{\mathsf{T}}$.
Proof A is real matrix. Then $\bar{A}\bar{\mathbf{v}} = A\bar{\mathbf{v}}$. In the same time $\bar{A}\bar{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$. Thus $A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$.

Chapter 5. Inner Product Spaces

1.4. Prove that for vectors in an inner product space

$$||\mathbf{x} \pm \mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{y}||^2 \pm 2\mathrm{Re}(\mathbf{x}, \mathbf{y}).$$

Recall that $\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z})$.
Proof

$$\begin{aligned} ||\mathbf{x} - \mathbf{y}||^2 &= (\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y}) \\ &= (\mathbf{x}, \mathbf{x} - \mathbf{y}) - (\mathbf{y}, \mathbf{x} - \mathbf{y}) \\ &= (\mathbf{x}, \mathbf{x}) - (\mathbf{x}, \mathbf{y}) - (\mathbf{y}, \mathbf{x}) + (\mathbf{y}, \mathbf{y}) \\ &= ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - (\mathbf{x}, \mathbf{y}) - \overline{(\mathbf{x}, \mathbf{y})} \\ &= ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - 2\mathrm{Re}(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Similarly $||\mathbf{x}+\mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{y}||^2 + 2\mathrm{Re}(\mathbf{x},\mathbf{y})$.

1.5. *Hint:* a) Check conjugate symmetry. b) Check linearity. c) Check conjugate symmetry.

1.7. Prove the parallelogram identity for an inner product space V,

$$||\mathbf{x}+\mathbf{y}||^2 + ||\mathbf{x}-\mathbf{y}||^2 = 2(||\mathbf{x}||^2 + ||\mathbf{y}||^2).$$

Proof

$$\begin{aligned}
||\mathbf{x}+\mathbf{y}||^2 + ||\mathbf{x}-\mathbf{y}||^2 &= (\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) + (\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}) \\
&= (\mathbf{x},\mathbf{x}) + (\mathbf{x},\mathbf{y}) + (\mathbf{y},\mathbf{x}) + (\mathbf{y},\mathbf{y}) + \\
&\quad\ (\mathbf{x},\mathbf{x}) - (\mathbf{x},\mathbf{y}) - (\mathbf{y},\mathbf{x}) + (\mathbf{y},\mathbf{y}) \\
&= 2(\mathbf{x},\mathbf{x}) + 2(\mathbf{y},\mathbf{y}) \\
&= 2(||\mathbf{x}||^2 + ||\mathbf{y}||^2).
\end{aligned}$$

1.8. *Proof sketch:* a) Let $\mathbf{v} = \mathbf{x}$, then $(\mathbf{x},\mathbf{x}) = 0$, $\mathbf{x} = \mathbf{0}$.
b) $(\mathbf{x}, \mathbf{v}_k = 0), \forall\, k$, then $(\mathbf{x},\mathbf{v}) = 0$, form conclusion in a), $\mathbf{x} = \mathbf{0}$.
c) $(\mathbf{x}-\mathbf{y}, \mathbf{v}_k), \forall\, k$, form b) $\mathbf{x}-\mathbf{y} = \mathbf{0}$, then $\mathbf{x} = \mathbf{y}$.

2.3. Let $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ be an orthonormal basis in V.

a) Prove that for any $\mathbf{x} = \sum_{k=1}^n \alpha_k \mathbf{v}_k$, $\mathbf{y} = \sum_{k=1}^n \beta_k \mathbf{v}_k$

$$(\mathbf{x},\mathbf{y}) = \sum_{k=1}^n \alpha_k \overline{\beta}_k.$$

b) Deduce from this *Parseval's identity*

$$(\mathbf{x},\mathbf{y}) = \sum_{k=1}^n (\mathbf{x},\mathbf{v}_k)\overline{(\mathbf{y},\mathbf{v}_k)}.$$

c) Assume now that $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is only an orthogonal basis, not an orthonormal one. Can you write down Parseval's identity in this case?

Proof
a)

$$(\mathbf{x},\mathbf{y}) = (\sum_{k=1}^n \alpha_k \mathbf{v}_k, \sum_{k=1}^n \beta_k \mathbf{v}_k) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\beta}_j (\mathbf{v}_i, \mathbf{v}_j) = \sum_{k=1}^n \alpha_k \overline{\beta}_k.$$

Because $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is an orthonormal basis, $(\mathbf{v}_i, \mathbf{v}_j) = 0, i \neq j, (\mathbf{v}_i, \mathbf{v}_j) = 1, i = j$.

b) Use $(\mathbf{x}, \mathbf{v}_k) = \alpha_k, (\mathbf{y}, \mathbf{v}_k) = \beta_k$ and conclusion in a).
c) Use equation in a),

$$(\mathbf{x}, \mathbf{y}) = (\sum_{k=1}^{n} \alpha_k \mathbf{v}_k, \sum_{k=1}^{n} \beta_k \mathbf{v}_k) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \overline{\beta}_j (\mathbf{v}_i, \mathbf{v}_j) = \sum_{k=1}^{n} \alpha_k \overline{\beta}_k (\mathbf{v}_k, \mathbf{v}_k)$$
$$= \sum_{k=1}^{n} \alpha_k \overline{\beta}_k ||\mathbf{v}_k||^2 = \sum_{k=1}^{n} \frac{(\mathbf{x}, \mathbf{v}_k)\overline{(\mathbf{y}, \mathbf{v}_k)}}{||\mathbf{v}_k||^2}.$$

As the basis is only orthogonal, not orthonomal, then $(\mathbf{x}, \mathbf{v}_k) = (\alpha_k \mathbf{v}_k, \mathbf{v}_k) = \alpha_k ||\mathbf{v}_k||^2$.

3.3 Complete an orthogonal system obtained in the previous problem to an orthogonal basis in $\mathbb{R}^3$, i.e., add to the system some vectors (how many?) to get an orthogonal basis.

Can you describe how to complete an orthogonal system to an orthogonal basis in general situation of $\mathbb{R}^n$ or $\mathbb{C}^n$?

Solution For 3D space, we already have 2 orthogonal vectors $\mathbf{v}_1, \mathbf{v}_2$ as the basis components. Then we just need another basis vector $\mathbf{v}_3$. The computation of $\mathbf{v}_3$ exploits the orthogonality, i.e., $(\mathbf{v}_1, \mathbf{v}_3) = 0, (\mathbf{v}_2, \mathbf{v}_3) = 0$. Expressed in matrix form, let $A = [\mathbf{v}_1, \mathbf{v}_2]$. Then solve $A^\mathsf{T} \mathbf{v}_3 = \mathbf{0}$. (Since it is in 3D space, using cross product is also simple.)

Generally, to complete an orthogonal system of $\mathbf{v}_1, \mathbf{v}_2 ... \mathbf{v}_r$. Consider $A = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r]$, using the orthogonality, we compute the rest basis vectors by solving $A^\mathsf{T} \mathbf{v} = \mathbf{0}$, i.e., the rest basis vectors compose an basis of Ker A^T or Null A^T.

3.9 (Using eigenvalues to compute determinants).

a) Find the matrix of the orthogonal projection onto the one-dimensional subspace in $\mathbb{R}^n$ spanned by the vector $(1, 1, ..., 1)^\mathsf{T}$;

b) Let A be the $n \times n$ matrix with all entries equal 1. Compute its eigenvalues and their multiplicities (use the previous problem);

c) Compute eigenvalues (and multiplicities) of the matrix $A - I$, i.e., of the matrix with zeros on the main diagonal and ones everywhere else;

d) Compute $\det(A - I)$.

Solution a) Note that from Remark 3.5, we know $P_E = \sum_{k=1}^{n} \frac{1}{||\mathbf{v}_k||^2} \mathbf{v}_k \mathbf{v}_k^*$. For this one-dimensional subspace, it is

$$P_E = \frac{1}{n} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}_{n \times n}.$$

b) Note that $A = nP_E$. Suppose $A\mathbf{x} = \lambda\mathbf{x}$, then $nP_E\mathbf{x} = \lambda\mathbf{x}$. i.e., n times of the eigenvector's projection on the 1D subspace equals λ times of itself. It also shows that the eigenvector's orthogonal projection is parallel to itself. Then there are two possibilities:

- One is the eigenvector is parallel with basis of the 1D subspace $\mathbf{v}$, i.e., $\mathbf{x} = \alpha\mathbf{v}, \alpha \neq 0$. In this case, $P_E\mathbf{x} = \mathbf{x}$, then $\lambda = n$. The geometric multiplicity is 1 for 1D eigenspace.

- The eigenvector is orthogonal to $\mathbf{v}$, i.e., $\mathbf{x} \perp \mathbf{v}$ and $P_E\mathbf{x} = \mathbf{0}, \lambda = 0$. We can totally find $n-1$ linearly independent eigenvectors so the geometric multiplicity of eigenvalue $\lambda = 0$ is $n-1$.

c) $\det(A - I - \lambda I) = \det(A - (\lambda+1)I)$. i.e., an eigenvalue of $A - I$ plus 1 is an eigenvalue of A. Then the eigenvalues of $A - I$ equal to the eigenvalues of A minus 1. Thus the eigenvalues of $A - I$ are $n-1$ with multiplicity 1 and -1 with multiplicity $n-1$.
d) $\det(A - I) = (n-1)(-1)^{n-1}$, which equals $n-1$ is n is odd, $1-n$ if n is even.

3.10. (Legendre's polynomials) *Hint:* Using the Gram-Schmidt orthogonalization algorithm is sufficient. But remember to use the inner product defined in the problem, e.g., $||1||^2 = (1,1) = \int_{-1}^{1} 1 \cdot \bar{1} dt = 2$.

3.11. Let $P = P_E$ be the matrix of an orthogonal projection onto a subspace E. Show that

a) The matrix P is *self-adjoint*, meaning that $P^* = P$.

b) $P^2 = P$.

Remark The above 2 properties completely characterize orthogonal projection.
Proof a) From the orthogonality, we have $(\mathbf{x}, \mathbf{x} - P\mathbf{x}) = (\mathbf{x} - P\mathbf{x}, \mathbf{x}) = 0$, $\forall\mathbf{x}$. $(\mathbf{x}, \mathbf{x} - P\mathbf{x}) = (\mathbf{x} - P\mathbf{x})^*\mathbf{x} = (\mathbf{x}^* - \mathbf{x}^*P^*)\mathbf{x} = \mathbf{x}^*\mathbf{x} - \mathbf{x}^*P^*\mathbf{x} = 0$. On the other hand, $(\mathbf{x} - P\mathbf{x}, \mathbf{x}) = \mathbf{x}^*(\mathbf{x} - P\mathbf{x}) = \mathbf{x}^*\mathbf{x} - \mathbf{x}^*P\mathbf{x} = 0$. Subtract two equalities, $\mathbf{x}^*(P - P^*)\mathbf{x} = 0$, $\forall\mathbf{x}$. Then $P - P* = 0_{n\times n}, P = P^*$.
b) Consider $(P\mathbf{x}, \mathbf{x} - P\mathbf{x}) = (\mathbf{x} - P\mathbf{x})^*P\mathbf{x} = (\mathbf{x}^* - \mathbf{x}^*P^*)P\mathbf{x} = \mathbf{x}^*(P - P^*P)\mathbf{x} = 0$. Thus $P = P^*P = P^2$ since $P = P^*$.

3.13 Suppose P is the orthogonal projection onto an subspace E, and Q is the orthogonal projection onto the orthogonal complement $E^\perp$.

a) What are $P + Q$ and PQ?

b) Show that $P - Q$ is its inverse.

Proof a) $P+Q=I$ since $(P+Q)\mathbf{x}=P\mathbf{x}+Q\mathbf{x}=P_E\mathbf{x}+Q_{E^\perp}\mathbf{x}=\mathbf{x}$.
$PQ=0_{n\times n}$ as $\mathbf{x}^*PQ\mathbf{x}=\mathbf{x}^*P^*Q\mathbf{x}=(Q\mathbf{x},P\mathbf{x})=0$, $\forall\mathbf{x}$ (using P is self-adjoint shown in Problem 3.11).
b) $(P-Q)^2=(P-Q)(P-Q)=P^2-PQ-QP+Q^2=P^2+Q^2=P^2+Q^2+PQ+QP=(P+Q)^2=I^2=I$ (using $PQ=QP=0$). i.e., $(P-Q)^{-1}=P-Q$.

4.5. *Minimal norm solution.* Let an equation $A\mathbf{x}=\mathbf{b}$ has a solution, and let A has non-trivial kernel (so the solution is not unique). Prove that

a) There exists a unique solution $\mathbf{x}_0$ of $A\mathbf{x}=\mathbf{b}$ minimizing the norm $||\mathbf{x}||$, i.e., that there exists unique $\mathbf{x}_0$ such that $A\mathbf{x}_0=\mathbf{b}$ and $||\mathbf{x}_0||\leqslant||\mathbf{x}||$ for any $\mathbf{x}$ satisfying $A\mathbf{x}=\mathbf{b}$.

b) $\mathbf{x}_0=P_{(\operatorname{Ker}A)^\perp}\mathbf{x}$ for any $\mathbf{x}$ satisfying $A\mathbf{x}=\mathbf{0}$.

Proof a) Suppose $\mathbf{x}_0,\mathbf{x}_1$ are solutions of $A\mathbf{x}=\mathbf{b}$. Then $A(\mathbf{x}_1-\mathbf{x}_0)=\mathbf{0}$. i.e., $\mathbf{x}_1-\mathbf{x}_0\in\operatorname{Ker}A$. As a result, $P_{(\operatorname{Ker}A)^\perp}(\mathbf{x}_1-\mathbf{x}_0)=\mathbf{0}=P_{(\operatorname{Ker}A)^\perp}\mathbf{x}_1-P_{(\operatorname{Ker}A)^\perp}\mathbf{x}_0$. So we have $P_{(\operatorname{Ker}A)^\perp}\mathbf{x}_1=P_{(\operatorname{Ker}A)^\perp}\mathbf{x}_0=\mathit{const}:=\mathbf{h}$.

Note that $||\mathbf{x}||^2=||P_{(\operatorname{Ker}A)^\perp}\mathbf{x}||^2+||\mathbf{x}-P_{(\operatorname{Ker}A)^\perp}\mathbf{x}||^2\geqslant||\mathbf{h}||^2$ for any $\mathbf{x}$ satisfying $A\mathbf{x}=\mathbf{b}$. When $\mathbf{x}_0-P_{(\operatorname{Ker}A)^\perp}\mathbf{x}_0=\mathbf{0}$, $\mathbf{x}_0=\mathbf{h}$, such a $\mathbf{x}_0$ has the smallest norm among all the solutions. The existence and uniqueness of $\mathbf{x}_0$ are guaranteed by $\mathbf{h}$.
b) It is shown above.

5.1. Show that for a square matrix A the equality $\det(A^*)=\overline{\det(A)}$ holds.
Proof $\det(A^*)=\det(\overline{A}^{\mathsf{T}})=\det(\overline{A})=\overline{\det(A)}$.

5.3. Let A be an $m\times n$ matrix. Show that Ker $A=$ Ker (A^*A).
Proof It is easy to see Ker $A\subset$ Ker (A^*A). Next we show Ker $(A^*A)\subset$ Ker A. Consider $||A\mathbf{x}||^2=(A\mathbf{x},A\mathbf{x})=\mathbf{x}^*A^*A\mathbf{x}$. Thus if $A^*A\mathbf{x}=\mathbf{0}$, we have $\mathbf{x}^*A^*A\mathbf{x}=||A\mathbf{x}||^2=0$, i.e., $A\mathbf{x}=\mathbf{0}$. Thus Ker $(A^*A)\subset$ Ker A. As a result, we can conclude Ker $A=$ Ker (A^*A).

6.4. Show that a product of unitary (orthogonal) matrices is unitary (orthogonal) as well.
Proof Suppose U_1,U_2 are unitary (orthogonal), then

$$(U_1U_2)^*U_1U_2=U_2^*U_1^*U_1U_2=U_2^*IU_2=I.$$

From Lemma 6.2, we know the product is unitary (orthogonal).

Chapter 6. Structure of Operators in Inner Product Spaces

1.1. Use the upper triangular representations of an operator to give an alternative proof of the fact that the determinant is the product and the trace is the sum of eigenvalues counting multiplicities.
Proof (The proof use the fact that the entries on the diagonal of T are the eigenvalues of A, counting multiplicity, which seems not be explicitly stated in the book and can be found like in the Wiki.) $A = UTU^*$. $\det A = (\det U)(\det T)(\det U^*) = \det U = \Pi_{i=1}^n \lambda_i$ because U is unitary and T is upper triangular with eigenvalues of A on its diagonal.

To consider the trace, suppose $U = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n]$. Then A can be represented by

$$A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ & \lambda_2 & \dots & t_{2n} \\ & & \ddots & \vdots \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^\intercal \\ \mathbf{u}_2^\intercal \\ \vdots \\ \mathbf{u}_n^\intercal \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \mathbf{u}_1 & t_{12}\mathbf{u}_1 + \lambda_2 \mathbf{u}_2 & \dots & t_{1n}\mathbf{u}_1 + t_{2n}\mathbf{u}_2 + ... + \lambda_n \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^\intercal \\ \mathbf{u}_2^\intercal \\ \vdots \\ \mathbf{u}_n^\intercal \end{bmatrix}$$

$$= \lambda_1 \mathbf{u}_1 \mathbf{u}_1^\intercal + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^\intercal + ... + \lambda_n \mathbf{u}_n \mathbf{u}_n^\intercal,$$

where we exploit the orthogonality of $\mathbf{u}_1, \mathbf{u}_2, ...\mathbf{u}_n$. Then note that the trace of matrix $\mathbf{u}_i \mathbf{u}_i^\intercal = [u_{i1}\mathbf{u}_i \ u_{i2}\mathbf{u}_i \ ... \ u_{in}\mathbf{u}_i]$ (outer product) is $u_{i1}^2 + u_{i2}^2 + ... + u_{in}^2 = ||\mathbf{u}_i||^2 = 1$. Thus trace$A$ = trace$(\lambda_1 \mathbf{u}_1 \mathbf{u}_1^\intercal)$ + trace$(\lambda_2 \mathbf{u}_2 \mathbf{u}_2^\intercal) + ... +$ trace$(\lambda_n \mathbf{u}_n \mathbf{u}_n^\intercal) = \lambda_1 + \lambda_2 + ... + \lambda_n = \sum_{i=1}^n \lambda_i$.

2.2. True or false: The sum of normal operators is normal? Justify your conclusion.
Solution True. Suppose two normal operators are $N_1 = U_1 D_1 U_1^*, N_2 = U_2 D_2 U_2^*$. U_1, U_2 are unitary and D_1, D_2 are diagonal.

$$(N_1 + N_2)^*(N_1 + N_2) = N_1^* N_1 + N_1^* N_2 + N_2^* N_1 + N_2^* N_2$$
$$(N_1 + N_2)(N_1 + N_2)^* = N_1 N_1^* + N_1 N_2^* + N_2 N_1^* + N_2 N_2^*$$

N_1, N_2 are normal, we need to prove $N_1^* N_2 + N_2^* N_1 = N_1 N_2^* + N_2 N_1^*$. In face, $N_1^* N_2 = U_1 D_1^* U_1^* U_2 D_2 U_2^*$. $N_1 N_2^* = U_1 D_1 U_1^* U_2 D_2^* U_2^*$. As can be shown, $D_1^* U_1^* U_2 D_2 = D_1 U_1^* U_2 D_2^*$ because D_1, D_2 are diagonal matrices, $D_1^* = \overline{D_1}, D_2^* = \overline{D_2}$. $D_1^* D_2 = D_1 D_2^*$ (for complex numbers $c_1, c_2, \overline{c_1} c_2 = c_1 \overline{c_2}$). By checking the entries of the product, one can conclude $D_1^* U_1^* U_2 D_2 =$

$D_1U_1^*U_2D_2^*$ and $N_1^*N_2 = N_1N_2^*, N_2^*N_1 = N_2N_1^*$. So the statement is true.

2.9. Give a proof if the statement is true, or give a counterexample if it is false:

a) If $A = A^*$ then $A + iI$ is invertible.
True. The eigenvalues of $A + iI$ are $\lambda_i + i$ where λ_i are eigenvalues of A and are real. Then $\det(A + iI) = \Pi_{i=1}n(\lambda_i + i) \neq 0$. (If $c_1, c_2 \in \mathbb{C}, c_1c_2 = 0$, then at least one of c_1, c_2 is 0.)

b) If U is unitary, $U + \frac{3}{4}I$ is invertible.
True. If $(U + \frac{3}{4}I)\mathbf{x} = U\mathbf{x} + \frac{3}{4}\mathbf{x} = \mathbf{0}$, note that $||U\mathbf{x}|| = ||\mathbf{x}||$. Then $||U\mathbf{x} + \frac{3}{4}\mathbf{x}|| \geqslant ||U\mathbf{x}|| - ||\frac{3}{4}\mathbf{x}|| = \frac{1}{4}||\mathbf{x}||$. So the homogeneous equation only has the trivial solution, $U + \frac{3}{4}I$ is invertible.

c) If a matrix is real, $A - iI$ is invertible.
False. A can have an eigenvalue i.

3.1. Show that the number of non-zero singular values of a matrix A coincides with its rank.
Proof Suppose the dimension of A is $m \times n$. It is known that Ker A = Ker (A^*A). Then Rank $A = n-$ dim Ker A = n - dim Ker (A^*A) = Rank (A^*A). The SVD of A is $A = W\Sigma V^*$, then $A^*A = V\Sigma^*W^*W\Sigma V^* = V\Sigma^2V^*$. Then Rank A = Rank (A^*A) = Rank $(V\Sigma^2V^*)$ = Rank Σ^2 = number of non-zero singular values, because V is an orthogonal matrix (full rank).

3.5. Find singular value decomposition of the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}.$$

Use it to find

a) $\max_{||\mathbf{x}|| \leqslant 1} ||A\mathbf{x}||$ and the vector where the maximum is attained;

b) $\max_{||\mathbf{x}||=1} ||A\mathbf{x}||$ and the vector where the minimum is attained;

c) the image $A(B)$ of the closed unit ball in $\mathbb{R}^2$, $B = \{\mathbf{x} \in \mathbb{R}^2 : ||\mathbf{x}|| \leqslant 1\}$. Describe $A(B)$ geometrically.

Solution (The SVD steps is ignored here.) a) Suppose $A = W\Sigma V^*$, then $(A\mathbf{x}, A\mathbf{x}) = \mathbf{x}^*A^*A\mathbf{x} = \mathbf{x}^*V\Sigma^2V^*\mathbf{x} = (V^*\mathbf{x})^*\Sigma^2(V^*\mathbf{x})$. Define $\mathbf{y} = [y_1\ y_2]^\mathsf{T} = V^*\mathbf{x}$. Because V is orthogonal, then $\mathbf{y}$ also lies in the unit ball. Thus

$$\begin{aligned}(A\mathbf{x}, A\mathbf{x}) &= \mathbf{y}^*\Sigma^2\mathbf{y} \\ &= \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 16 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= 16y_1^2 + y_2^2.\end{aligned}$$

$y_1^2 + y_2^2 \leqslant 1$. Hence the maximum is 16 attained when $\mathbf{y} = [1\ 0]^\mathsf{T}$. Corresponding $\mathbf{x}$ can be solved by $\mathbf{x} = V\mathbf{y}$.
b) Similarly, the minimum is 1 attained when $\mathbf{y} = [0\ 1]^\mathsf{T}$.
c) Ellipse.

3.8. Let A be an $m \times n$ matrix. Prove that *non-zero* eigenvalues of the matrices A^*A and AA^* (counting multiplicities) coincide.
Proof Suppose $\mathbf{v}$ is an eigenvector of A^*A corresponding to a non-zero eigenvalue λ, i.e., $A^*A\mathbf{v} = \lambda\mathbf{v}$. Then $AA^*A\mathbf{v} = A(\lambda\mathbf{v}) = \lambda(A\mathbf{v})$, i.e., λ is an eigenvalue of AA^* with corresponding eigenvector $A\mathbf{v}$. Similarly, we can show that the non-zero eigenvalues of AA^* are also eigenvalues of A^*A. Thus non-zero eigenvalues of the matrices $A * A$ and AA^* coincide.

4.2. Let A be a normal operator, and let $\lambda_1, \lambda_2, ..., \lambda_n$ be its eigenvalues (counting multiplicities). Show that singular values of A are $|\lambda_1|, |\lambda_2|, ..., |\lambda_n|$.
Proof First, we show for normal operator A, $||A\mathbf{x}|| = ||A^*\mathbf{x}||$. Note that $AA^* = A^*A$, then

$$\begin{aligned}((AA^* - A^*A)\mathbf{x}, \mathbf{x}) &= (\mathbf{0}, \mathbf{x}) \\ &= (AA^*\mathbf{x}, \mathbf{x}) - (A^*A\mathbf{x}, \mathbf{x}) \\ &= (A^*\mathbf{x}, A^*\mathbf{x}) - (A\mathbf{x}, A\mathbf{x}) \\ &= ||A^*\mathbf{x}||^2 - ||A\mathbf{x}||^2 \\ &= 0\end{aligned}$$

Thus $||A\mathbf{x}|| = ||A^*\mathbf{x}||$.

Suppose $\mathbf{v}$ is an eigenvector of A corresponding to the eigenvalue λ. Note that $A - \lambda I$ is also normal (see Problem 2.2). Thus we have $||(A - \lambda I)\mathbf{v}|| = ||(A - \lambda I)^*\mathbf{v}|| = ||(A^* - \overline{\lambda} I)\mathbf{v}|| = 0$, i.e., $A^*\mathbf{v} = \overline{\lambda}\mathbf{v}$. So $A^*A\mathbf{v} = A^*(\lambda\mathbf{v}) = \lambda\overline{\lambda}\mathbf{v} = |\lambda|^2\mathbf{v}$. $|\lambda|^2$ is an eigenvalue of A^*A, then $|\lambda|$ is a singular value of A.

4.4. Let $A = \widetilde{W}\widetilde{\Sigma}\widetilde{V}^*$ be a *reduced* singular value decomposition of A. Show that Ran A = Ran $\widetilde{W}$, and then by taking adjoint that Ran A^* = Ran $\widetilde{W}$.
Proof Suppose A is an $m \times n$ matrix. $\widetilde{\Sigma} = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_r)$. $\widetilde{W}$ is $m \times r$ and $\widetilde{V}^*$ is $r \times n$. To show Ran A = Ran $\widetilde{W}$, we just need to show Ran $\widetilde{\Sigma}\widetilde{V}^*$ $= \mathbb{R}^r =$ Ran $\widetilde{V}^*$ ($\widetilde{\Sigma}$ has full rank r), which holds because Rank $\widetilde{V}^* = r$ and $r \leqslant n$. Thus Ran $\widetilde{\Sigma}\widetilde{V}^* = \mathbb{R}^r$ and Ran A = Ran $\widetilde{W}$.

By taking adjoint, we can prove Ran A^* = Ran $\widetilde{W}$.

End