13 marks total

MAST30026 Metric and Hilbert Spaces

Assignment 1 Solutions — 2018 Semester 2

Notation and conventions

The set $\mathbb{N}$ is the set of positive integers $\{1, 2, 3, \dots\}$.

The set $\mathbb{N}_0$ is the set of nonnegative integers $\{0, 1, 2, \dots\}$.

The transpose of a matrix X is denoted X'.

Question 1

(2 marks)

Lemma L2-2. The pair (S^1, d_a) is a metric space.

Exercise L2-3. Give a direct proof of Lemma L2-2 by dividing into cases as follows: Given $x, y, z \in S^1$, set $\theta := \Phi^{-1}(x)$, $\theta' := \Phi^{-1}(y)$, and $\theta'' := \Phi^{-1}(z)$. Consider the following three statements:

P_1 $|\theta - \theta'| \leqslant \pi$

P_2 $|\theta' - \theta''| \leqslant \pi$

P_3 $|\theta - \theta''| \leqslant \pi$

Each is either true or false for a particular triple (x, y, z), and this means there are $2^3 = 8$ cases (e.g. P_1, P_2 true but P_3 false). Prove each case individually, and in this way prove the lemma.

Recall that $d_a \colon S^1 \times S^1 \longrightarrow \mathbb{R}_{\geqslant 0}$ is defined as

$$d_a(\Phi(\omega), \Phi(\omega')) = \min\{|\omega - \omega'|, 2\pi - |\omega - \omega'|\} = \begin{cases} |\omega - \omega'|, & |\omega - \omega'| \leqslant \pi, \\ 2\pi - |\omega - \omega'|, & |\omega - \omega'| > \pi. \end{cases}$$

for every $\omega, \omega' \in [0, 2\pi)$. We note that

(i) d_a is nonnegative: If $\omega, \omega' \in [0, 2\pi)$ then $0 \leqslant |\omega - \omega'| < 2\pi$ so that $|\omega - \omega'|$ and $2\pi - |\omega - \omega'|$ are both nonnegative;

(ii) d_a is symmetric: If $\omega, \omega' \in [0, 2\pi)$ then $|\omega - \omega'| = |\omega' - \omega|$.

(iii) d_a separates distinct elements: If $\omega, \omega' \in [0, 2\pi)$ and $d_a(\Phi(\omega), \Phi(\omega')) = 0$ then $|\omega - \omega'| = 0$ (since $2\pi - |\omega - \omega'| > 0$) and $\omega = \omega'$. Conversely if $\omega \in [0, 2\pi)$ then $d_a(\Phi(\omega), \Phi(\omega)) = \min\{0, 2\pi\} = 0$.

To prove that (S^1, d_a) is a metric space, it remains to establish that d_a satisfies the triangle equality: For our given $x, y, z \in S^1$ we need to prove that

$$d_a(x, y) + d_a(y, z) \geqslant d_a(x, z).$$

We proceed by considering cases in the following order:

(a) P_1, P_2, and P_3 are all false.

(b) P_1 is true, while P_2 and P_3 are both false.

(c) P_2 is true, while P_1 and P_3 are both false.

(d) P_1 and P_2 are true, while P_3 is false.

(e) P_3 is true, while P_1 and P_2 are both false.

(f) P_1 and P_3 are true, while P_2 is false.

(g) P_2 and P_3 are true, while P_1 is false.

(h) P_1, P_2, and P_3 are all true.

For $a, b \in \mathbb{R}$, we will use $T(a, b)$ to denote the triangle inequality for $|\cdot|$ in $\mathbb{R}$, i.e. that $|a| + |b| \geqslant |a + b|$.

1(a) All false

1(a) This case is not actually possible: Writing $[0, 2\pi)$ as the disjoint union $[0, \pi) \cup [\pi, 2\pi)$, the pigeonhole principle tells us that either $[0, \pi)$ or $[\pi, 2\pi)$ contains two of θ, θ', and θ''. The difference of those two must therefore be strictly less than π, so that P_1, P_2, and P_3 cannot simultaneously all be false.

(b) P_1 true; P_2, P_3 false

In this case we have

$$\begin{aligned} d_a(x, y) + d_a(y, z) &= |\theta - \theta'| + (2\pi - |\theta' - \theta''|), \\ d_a(x, z) &= 2\pi - |\theta - \theta''|. \end{aligned}$$

Observe:

$$\begin{aligned} & d_a(x, y) + d_a(y, z) \geqslant d_a(x, z) \\ \iff \quad & |\theta - \theta'| + (2\pi - |\theta' - \theta''|) \geqslant 2\pi - |\theta - \theta''| \\ \iff \quad & |\theta - \theta'| + |\theta - \theta''| \geqslant |\theta' - \theta''| \\ \iff \quad & |\theta' - \theta| + |\theta - \theta''| \geqslant |\theta' - \theta''|. \end{aligned}$$

The final inequality is precisely $T(\theta' - \theta, \theta - \theta'')$, so we have shown that the triangle inequality holds for d_a when only P_1 is true.

(c) P_2 true; P_1, P_3 false

In this case we have

$$\begin{aligned} d_a(x, y) + d_a(y, z) &= (2\pi - |\theta - \theta'|) + |\theta' - \theta''|, \\ d_a(x, z) &= 2\pi - |\theta - \theta''|. \end{aligned}$$

Observe:

$$\begin{aligned} & d_a(x, y) + d_a(y, z) \geqslant d_a(x, z) \\ \iff \quad & (2\pi - |\theta - \theta'|) + |\theta' - \theta''| \geqslant 2\pi - |\theta - \theta''| \\ \iff \quad & |\theta' - \theta''| + |\theta - \theta''| \geqslant |\theta - \theta'| \\ \iff \quad & |\theta' - \theta''| + |\theta'' - \theta| \geqslant |\theta' - \theta|. \end{aligned}$$

The final inequality is precisely $T(\theta' - \theta'', \theta'' - \theta)$, so we have shown that the triangle inequality holds for d_a when only P_2 is true.

(d) P_1, P_2 true; P_3 false

In this case we have $d_a(x, y) + d_a(y, z) = |\theta - \theta'| + |\theta' - \theta''|$. By $T(\theta - \theta', \theta' - \theta'')$, this means that $d_a(x, y) + d_a(y, z) \geqslant |\theta - \theta''|$. By definition, $d_a(x, z) \leqslant |\theta - \theta''|$, so we have shown that the triangle inequality holds for d_a when P_1 and P_2 are true. (Note that we did not actually use the fact that P_3 is false.)

(e) P_3 true; P_1, P_2 false

In this case we have

$$\begin{aligned} d_a(x, y) + d_a(y, z) &= (2\pi - |\theta - \theta'|) + (2\pi - |\theta' - \theta''|) = 4\pi - (|\theta - \theta'| + |\theta' - \theta''|), \\ d_a(x, z) &= |\theta - \theta''|. \end{aligned}$$

Observe:

$$\begin{aligned} & d_a(x, y) + d_a(y, z) \geqslant d_a(x, z) \\ \iff \quad & 4\pi - (|\theta - \theta'| + |\theta' - \theta''|) \geqslant |\theta - \theta''| \\ \iff \quad & |\theta - \theta'| + |\theta' - \theta''| + |\theta - \theta''| \leqslant 4\pi. \end{aligned}$$

1(e) Now, as a function of $(\theta, \theta', \theta'')$, the left-hand side of the last inequality is invariant under permutations, so without loss of generality we may assume that $0 \leqslant \theta \leqslant \theta' \leqslant \theta'' < 2\pi$. The left-hand side then becomes

$$(\theta' - \theta) + (\theta'' - \theta') + (\theta'' - \theta) = 2(\theta'' - \theta) < 4\pi,$$

where the last inequality is because $\theta'' - \theta < 2\pi$. Thus we have shown that the triangle inequality holds for d_a when only P_3 is true.

(f) P_1, P_3 true; P_2 false

In this case we have

$$\begin{aligned} d_a(x,y) + d_a(y,z) &= |\theta - \theta'| + (2\pi - |\theta' - \theta''|), \\ d_a(x,z) &= |\theta - \theta''|. \end{aligned}$$

Observe:

$$\begin{aligned} & d_a(x,y) + d_a(y,z) \geqslant d_a(x,z) \\ \iff & |\theta - \theta'| + (2\pi - |\theta' - \theta''|) \geqslant |\theta - \theta''| \\ \iff & |\theta - \theta''| + |\theta' - \theta''| - |\theta - \theta'| \leqslant 2\pi \\ \Longleftarrow & |\theta - \theta''| + |\theta' - \theta''| - |\theta - \theta'| \leqslant 2|\theta - \theta''| \quad \text{since } |\theta - \theta''| \leqslant \pi \\ \iff & |\theta' - \theta''| - |\theta - \theta'| \leqslant |\theta - \theta''| \\ \iff & |\theta - \theta'| + |\theta - \theta''| \geqslant |\theta' - \theta''| \\ \iff & |\theta' - \theta| + |\theta - \theta''| \geqslant |\theta' - \theta''|. \end{aligned}$$

The final inequality is precisely $T(\theta' - \theta, \theta - \theta'')$, so we have shown that the triangle inequality holds for d_a when P_1 and P_3 are true but P_2 is false.

(g) P_2, P_3 true; P_1 false

In this case we have

$$\begin{aligned} d_a(x,y) + d_a(y,z) &= (2\pi - |\theta - \theta'|) + |\theta' - \theta''|, \\ d_a(x,z) &= |\theta - \theta''|. \end{aligned}$$

Observe:

$$\begin{aligned} & d_a(x,y) + d_a(y,z) \geqslant d_a(x,z) \\ \iff & (2\pi - |\theta - \theta'|) + |\theta' - \theta''| \geqslant |\theta - \theta''| \\ \iff & |\theta - \theta''| + |\theta - \theta'| - |\theta' - \theta''| \leqslant 2\pi \\ \Longleftarrow & |\theta - \theta''| + |\theta - \theta'| - |\theta' - \theta''| \leqslant 2|\theta - \theta''| \quad \text{since } |\theta - \theta''| \leqslant \pi \\ \iff & |\theta - \theta'| - |\theta' - \theta''| \leqslant |\theta - \theta''| \\ \iff & |\theta - \theta'| - |\theta' - \theta''| \leqslant |\theta - \theta''| \\ \iff & |\theta - \theta''| + |\theta' - \theta''| \geqslant |\theta - \theta'| \\ \iff & |\theta - \theta''| + |\theta'' - \theta'| \geqslant |\theta - \theta'|. \end{aligned}$$

The final inequality is precisely $T(\theta - \theta'', \theta'' - \theta')$, so we have shown that the triangle inequality holds for d_a when P_2 and P_3 are true but P_1 is false.

(h) All true

Please refer to the case where only P_1 and P_2 are true.

Question 2 (3 marks)

Exercise L3-3. Prove that any element of $\mathrm{Isom}(S^1, d_a)$ of the form

$$F = g_1 g_2 \dots g_r, \quad r \geqslant 0,$$

where each g_i is either R_θ for some $\theta \in \mathbb{R}$ or T, may be proven equal to $R_\psi T^n$ for some $\psi \in [0, 2\pi)$ and $n \in \{0, 1\}$, using relations (R1), (R2), and (R3).

Recall that the relations were (R1) $R_{\theta_1} R_{\theta_2} = R_{\theta_1+\theta_2}$, (R2) $R_\theta T = T R_{-\theta}$, and (R3) $T^2 = \mathrm{id}$.

We proceed by induction. For each $m \in \mathbb{N}_0$, let $P(m)$ be the following proposition: For every m-tuple $(g_1, g_2, \dots, g_m)$ where each g_i is either R_θ for some real θ or T, there exist $\psi \in [0, 2\pi)$ and $n \in \{0, 1\}$ such that $g_1 g_2 \dots g_m = R_\psi T^n$.

Note that $P(0)$ is simply the proposition that $\mathrm{id} \in \mathrm{Isom}(S^1, d_a)$ may be written as $R_\psi T^n$ for some $\psi \in [0, 2\pi)$ and $n \in \{0, 1\}$. Taking $\psi := 0$ and $n := 0$, we see that $P(0)$ holds.

Suppose $k \in \mathbb{N}_0$ is such that $P(k)$ holds. We will show that $P(k+1)$ holds. For a given $(g_1, g_2, \dots, g_{k+1})$, we wish to show that there exist $\psi \in [0, 2\pi)$ and $n \in \{0, 1\}$ such that $g_1 g_2 \dots g_{k+1} = R_\psi T^n$. Since $P(k)$ holds, there exist $\omega \in [0, 2\pi)$ and $p \in \{0, 1\}$ such that $g_1 g_2 \dots g_k = R_\omega T^p$. Hence it is sufficient to show that there exist $\psi \in [0, 2\pi)$ and $n \in \{0, 1\}$ such that $R_\omega T^p g_{k+1} = R_\psi T^n$.

If g_{k+1} is R_θ for some $\theta \in \mathbb{R}$ then, using (R2) p times, we have

$$T^p g_{k+1} = T^p R_\theta = \begin{cases} R_\theta, & p = 0, \\ R_{-\theta} T, & p = 1, \end{cases}$$

which we may write as $R_{(-1)^p \theta} T^p$. Hence $R_\omega T^p g_{k+1} = R_\omega R_{(-1)^p \theta} T^p$. Using (R1), we have $R_\omega T^p g_{k+1} = R_{\omega + (-1)^p \theta} T^p$. Finally, taking ψ to be $(\omega + (-1)^p \theta) \bmod 2\pi$ and $n := p$, we have

$$g_1 g_2 \dots g_k g_{k+1} = R_{\omega + (-1)^p \theta} T^p = R_\psi T^n.$$

If g_{k+1} is T then using (R3)

$$T^p g_{k+1} = T^p T = \begin{cases} T, & p = 0, \\ \mathrm{id}, & p = 1, \end{cases}$$

which we may write as T^{1-p}. Hence $R_\omega T^p g_{k+1} = R_\omega T^{1-p}$. Taking $\psi := \omega$ and $n := 1 - p$, we have

$$g_1 g_2 \dots g_k g_{k+1} = R_\omega T^{1-p} = R_\psi T^n.$$

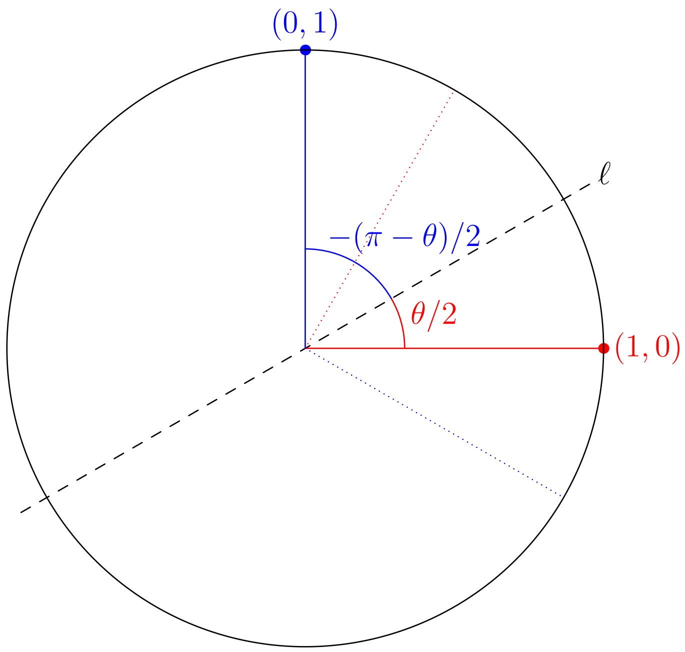
In both cases we have produced $\psi \in [0, 2\pi)$ and $n \in \{0, 1\}$ such that $R_\omega T^p g_{k+1} = R_\psi T^n$. Thus, we have shown that $P(k+1)$ holds if $P(k)$ holds. By the principle of mathematical induction, we therefore have that $P(m)$ holds for every $m \in \mathbb{N}_0$.

Question 3 (1 mark)

Exercise L3-4. Prove that $R_\theta T\colon S^1 \longrightarrow S^1$ is the reflection of S^1 through the straight line which passes through the origin and $(\cos(\theta/2), \sin(\theta/2))$.

Let ℓ denote the line passing through the origin and $(\cos(\theta/2), \sin(\theta/2))$.

Since R_θ, T, and reflection in ℓ are all linear transformations, it suffices to show that the image of $(1,0)$ and $(0,1)$ under $R_\theta T$ are their respective images under reflection in ℓ.



Let us first determine these images: The (directed and counterclockwise) angles subtended at the origin are

- from $(1,0)$ to $(\cos(\theta/2), \sin(\theta/2))$: $\theta/2$, and
- from $(0,1)$ to $(\cos(\theta/2), \sin(\theta/2))$: $-(\pi-\theta)/2$.

Thus, the image of $(1,0)$ should be a counterclockwise rotation of $(\cos(\theta/2), \sin(\theta/2))$ around the origin by $\theta/2$, while the image of $(1,0)$ should be a counterclockwise rotation of $(\cos(\theta/2), \sin(\theta/2))$ around the origin by $-(\pi-\theta)/2$. That is, $(1,0)$ is sent to $(\cos\theta, \sin\theta)$, and $(0,1)$ is sent to

$$(\cos(\theta/2 - (\pi-\theta)/2), \sin(\theta/2 - (\pi-\theta)/2)) = (\cos(\theta - \pi/2), \sin(\theta - \pi/2)) = (\sin\theta, -\cos\theta).$$

The images of $(1,0)$ and $(0,1)$ under $R_\theta T$ are given by

$$\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix} \quad\text{and}\quad \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix} = \begin{bmatrix}\sin\theta\\ -\cos\theta\end{bmatrix}$$

respectively. These agree with their images under reflection in ℓ. Hence the linear transformations $R_\theta T$ and reflection through the line passing through the origin and $(\cos(\theta/2), \sin(\theta/2))$ are the same transformations.

Question 4 (5 marks)

Exercise L3-5. Let $F\colon V \longrightarrow V$ be an invertible linear operator on a finite-dimensional vector space.

3

(a) Prove that precisely one of the following two possibilities is realised:

(I) $\forall \mathcal{B}\ (F(\mathcal{B}) \sim \mathcal{B})$

(II) $\forall \mathcal{B}\ (F(\mathcal{B}) \not\sim \mathcal{B})$

where $\mathcal{B}$ ranges over all ordered bases and $F(\mathcal{B})$ denotes $(F(\boldsymbol{b}_1), \dots, F(\boldsymbol{b}_n))$ if $\mathcal{B} = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$. In the first case we say F is *orientation preserving*, and in the latter case we say F is *orientation reversing*.

(b) Prove that F is orientation preserving iff $\det(F) > 0$ and orientation reversing iff $\det(F) < 0$.

(c) Define

$$O(n) := \{X \in M_n(\mathbb{R}) \mid X \text{ is orthogonal, i.e. } X'X = I_n\}$$
$$SO(n) := \{X \in O(n) \mid \det(X) = 1\}.$$

1

Prove that $X \in O(n)$ if and only if for all $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$

$$(X\boldsymbol{v}) \cdot (X\boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{w}.$$

By part (b), $SO(n)$ are precisely the matrices in $O(n)$ that give rise to orientation preserving linear transformations $\mathbb{R}^n \longrightarrow \mathbb{R}^n$.

1

(d) Prove that $O(n)$ is a group under multiplication, and $SO(n)$ is a subgroup. Produce an element $T \in O(n)$ such that $T^2 = \text{id}$ and every element of $O(n)$ not in $SO(n)$ may be written as XT for some $X \in SO(n)$. Thus prove $SO(n) \subseteq O(n)$ is a normal subgroup and that there is a group isomorphism

$$O(n)/SO(n) \cong \mathbb{Z}/2\mathbb{Z}.$$

4(a) Fix an ordered basis $\mathcal{C}$ of V. Since F is a linear operator on a finite-dimensional vector space, for some matrix $A \in M_n(\mathbb{R})$ (where $n = \dim V$) we have $F(\mathcal{C}) = A\mathcal{C}$. Note that because F is invertible, A must be an invertible matrix. That is, $\det A \neq 0$.

Now, our definition of A above means that $A = [\text{id}]^{\mathcal{C}}_{F(\mathcal{C})}$, where $[\text{id}]^{\mathcal{C}}_{F(\mathcal{C})} \in M_n(\mathbb{R})$ is the matrix which changes $F(\mathcal{C})$-coordinates to $\mathcal{C}$-coordinates. For our basis $\mathcal{C}$ we have by definition (from Tutorial 1)

$$F(\mathcal{C}) \sim \mathcal{C} \quad \iff \quad \det A > 0.$$

By negating (and noting that $\det A \neq 0$ as mentioned above) we also have for our basis $\mathcal{C}$ that

$$F(\mathcal{C}) \not\sim \mathcal{C} \quad \iff \quad \det A < 0.$$

Next, consider another arbitrary ordered basis $\mathcal{B}$ of V. Then $[\text{id}]^{\mathcal{B}}_{F(\mathcal{B})}$ is the matrix which changes $F(\mathcal{B})$-coordinates to $\mathcal{B}$-coordinates. We can see that this matrix is the same as $[F]^{\mathcal{B}}_{\mathcal{B}}$, the matrix corresponding to applying F under $\mathcal{B}$-coordinates. We now carry out a change of basis to relate $[F]^{\mathcal{B}}_{\mathcal{B}}$ and $[F]^{\mathcal{C}}_{\mathcal{C}} = [\text{id}]^{\mathcal{C}}_{F(\mathcal{C})} = A$: We write

$$[F]^{\mathcal{B}}_{\mathcal{B}} = [\text{id}]^{\mathcal{B}}_{\mathcal{C}}\,[F]^{\mathcal{C}}_{\mathcal{C}}\,[\text{id}]^{\mathcal{C}}_{\mathcal{B}} = [\text{id}]^{\mathcal{B}}_{\mathcal{C}}\,A\,[\text{id}]^{\mathcal{C}}_{\mathcal{B}}.$$

From this we can see that

$$F(\mathcal{B}) \sim \mathcal{B} \quad \iff \quad \det([\text{id}]^{\mathcal{B}}_{F(\mathcal{B})}) > 0 \quad \iff \quad \det([\text{id}]^{\mathcal{B}}_{\mathcal{C}}\,A\,[\text{id}]^{\mathcal{C}}_{\mathcal{B}}) > 0.$$

4(a) We simplify the last condition by writing

$$\begin{aligned}
\det([\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}\, A\, [\mathrm{id}]^{\mathcal{C}}_{\mathcal{B}}) &= \det([\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}})(\det A)\det([\mathrm{id}]^{\mathcal{C}}_{\mathcal{B}}),\\
&= (\det A)\det([\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}})\det([\mathrm{id}]^{\mathcal{C}}_{\mathcal{B}}),\\
&= (\det A)\det([\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}\,[\mathrm{id}]^{\mathcal{C}}_{\mathcal{B}}),\\
&= \det A \qquad\qquad \text{since } [\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}\,[\mathrm{id}]^{\mathcal{C}}_{\mathcal{B}} = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{B}} = I_n.
\end{aligned}$$

This allows us to conclude that

$$F(\mathcal{B}) \sim \mathcal{B} \quad\Longleftrightarrow\quad \det(A) > 0.$$

Since $\mathcal{B}$ was an arbitrary ordered basis of V, we have in fact established that

$$F(\mathcal{B}) \sim \mathcal{B} \quad \forall \text{bases } \mathcal{B} \text{ of } V \quad\Longleftrightarrow\quad \det A > 0,$$

as well as

$$F(\mathcal{B}) \not\sim \mathcal{B} \quad \forall \text{bases } \mathcal{B} \text{ of } V \quad\Longleftrightarrow\quad \det A < 0.$$

by negating. Precisely one of $\det A > 0$ and $\det A < 0$ is true, so we may conclude that either (I) $F(\mathcal{B}) \sim \mathcal{B}$ for every basis $\mathcal{B}$ of V, or (II) $F(\mathcal{B}) \not\sim \mathcal{B}$ for every basis $\mathcal{B}$ of V.

(b) In the previous part we proved that it makes sense to classify the invertible linear operator F itself as being orientation-preserving or orientation-reversing, since applying F either preserves the orientation of all bases or preserves the orientation of no bases at all. As demonstrated above, since $\det F \neq 0$, F is orientation-preserving if and only if $\det F > 0$, while F is orientation-reversing if and only if $\det F < 0$.

(c) First observe that for every $X \in M_n(\mathbb{R})$ and for every $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ we have

$$(X\boldsymbol{v}) \cdot (X\boldsymbol{w}) = (X\boldsymbol{v})'(X\boldsymbol{w}) = \boldsymbol{v}'X'X\boldsymbol{w}.$$

($\Rightarrow$) Suppose $X \in O(n)$. Then $X'X = I_n$, so for every $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ we have

$$(X\boldsymbol{v}) \cdot (X\boldsymbol{w}) = \boldsymbol{v}'I_n\boldsymbol{w} = \boldsymbol{v} \cdot \boldsymbol{w}.$$

($\Leftarrow$) Suppose that $(X\boldsymbol{v}) \cdot (X\boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{w}$ for every $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$. For every $i \in \{1, 2, \ldots, n\}$, let $\boldsymbol{e}_i \in \mathbb{R}^n$ denote the column vector consisting of 1 as the ith coordinate and 0 in every other coordinate. Then for every $i, j \in \{1, 2, \ldots, n\}$ we have

$$(X\boldsymbol{e}_i) \cdot (X\boldsymbol{e}_j) = \boldsymbol{e}_i'X'X\boldsymbol{e}_j = (X'X)_{i,j},$$

where $(X'X)_{i,j}$ is the (i, j)th entry of $X'X$. By hypothesis, this must be equal to $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \mathbf{1}(i = j)$. That is, the (i, j)th entry of $X'X$ must be 1 if $i = j$ and 0 if $i \neq j$. Thus $X'X = I_n$ and $X \in O(n)$.

(d) We begin by showing that $O(n)$ is a group under multiplication. The identity element is I_n (which is orthogonal).

We claim that the inverse of $X \in O(n)$ is $X' \in O(n)$. Recall that X and X' are inverses in the group $GL(n)$ if X is orthogonal (i.e. $X'X = XX' = I_n$), so it remains to show that $X' \in O(n)$. Observe that $X''X' = (XX')' = (XX^{-1})' = I_n$, so indeed $X' \in O(n)$.

Next, if $X, Y \in O(n)$ then

$$(XY)'(XY) = Y'\underbrace{(X'X)}_{=I_n}Y = Y'Y = I_n,$$

so that $XY \in O(n)$. This shows that $O(n)$ is a group under matrix multiplication.

To show that $SO(n)$ is a subgroup of $O(n)$, we need to show that (i) $SO(n)$ contains the identity I_n; (ii) $SO(n)$ is closed under inversion; and (iii) $SO(n)$ is closed under matrix multiplication.

4(d) Since $\det I_n = 1$, we know that $I_n \in SO(n)$, so $SO(n)$ contains the identity. Next, if $X \in O(n)$ and $\det X = 1$ then $X' \in O(n)$ and $\det(X') = \det(X^{-1}) = (\det X)^{-1} = 1$, so $X' \in SO(n)$. Hence $SO(n)$ is closed under inversion. Finally, if $X, Y \in O(n)$ and $\det X = \det Y = 1$, then $XY \in O(n)$ and $\det(XY) = (\det X)(\det Y) = 1$, so $XY \in SO(n)$. Thus, $SO(n)$ is a subgroup of $O(n)$.

We claim that a suitable choice of $T \in O(n)$ is

$$\begin{bmatrix} -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix},$$

the $n \times n$ diagonal matrix with -1 in the upper left entry and 1 everywhere else along the diagonal. We verify that T is orthogonal: Note that $T' = T$ and $T^2 = I_n$, so that $T'T = T^2 = I_n$ and $T \in O(n)$. In our discussion below we will use the fact that $\det T = -1$.

Next, consider $Y \in O(n)$. Both Y and YT are elements in $O(n)$. Since $Y'Y = I_n$, we know that $(\det Y)^2 = 1$, so that $\det Y = 1$ or $\det Y = -1$. If $Y \notin SO(n)$, we necessarily have that $\det Y = -1$, so that $\det(YT) = (\det Y)(\det T) = 1$ and $YT \in SO(n)$. With the choice of $X := YT$, we have $X \in SO(n)$ and $XT = YT^2 = Y$. Thus our T satisfies the required criteria.

Before showing that $SO(n)$ is a normal subgroup of $O(n)$, we will first show that T in fact satisfies a further similar property. In particular, we will show that if $Y \in O(n) \setminus SO(n)$, then $Y = TZ$ for some $Z \in SO(n)$: We simply take $Z := TY \in O(n)$ and note that $\det Y = -1$ since $Y \in O(n) \setminus SO(n)$, so that $\det Z = (\det T)(\det Y) = 1$ and $Z \in SO(n)$. Then we verify that $TZ = T^2 Y = Y$.

The above results show that the right cosets of $O(n)$ with respect to $SO(n)$ is the set $\{SO(n), SO(n)T\}$, while the left cosets of $O(n)$ with respect to $SO(n)$ is the set $\{SO(n), T\,SO(n)\}$. Since each is a bipartition of $SO(n)$, we have that $SO(n)T = T\,SO(n)$.

We are now ready to show normality of $SO(n)$ as a subgroup of $O(n)$: If $Y \in SO(n)$ then $Y\,SO(n) = SO(n) = SO(n)\,Y$. If $Y \in O(n) \setminus SO(n)$ then for some $X \in SO(n)$ and $Z \in SO(n)$ we have $Y = XT = TZ$ and

$$Y\,SO(n) = TZ\,SO(n) = T\,SO(n) = SO(n)\,T = SO(n)\,XT = SO(n)Y.$$

This shows that $SO(n)$ is a normal subgroup of $O(n)$. Since there are only two cosets, the quotient group $O(n)/SO(n)$ is a group of order 2, so $O(n)/SO(n) \cong \mathbb{Z}/2\mathbb{Z}$.

Question 5 (1 mark)

Exercise L4-0. Prove that

$$d_1(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} |x_i - y_i|$$
$$d_\infty(\boldsymbol{x}, \boldsymbol{y}) = \max\{|x_i - y_i|\}_{i=1}^{n}$$

define metrics on $\mathbb{R}^n$.

Since $|\cdot|$ is nonnegative, and the sum and the maximum of nonnegative numbers are nonnegative, both d_1 and d_∞ are nonnegative. Furthermore, since $|a-b| = |b-a|$ for all real a and b, we have $d_1(\boldsymbol{x}, \boldsymbol{y}) = d_1(\boldsymbol{y}, \boldsymbol{x})$ and $d_\infty(\boldsymbol{x}, \boldsymbol{y}) = d_\infty(\boldsymbol{y}, \boldsymbol{x})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. That is, both d_1 and d_∞ are symmetric.

It remains to show that d_1 and d_∞ both separate distinct points and both satisfy the triangle inequality.

Let us proceed first for d_1. Suppose $d_1(\boldsymbol{x}, \boldsymbol{y}) = 0$. Since each $|x_i - y_i|$ is nonnegative, the only way for the sum $\sum_{i=1}^{n} |x_i - y_i|$ to be 0 is to have $|x_i - y_i|$ be exactly 0 for every i. That is, we must have $x_i = y_i$ for every i, so that $\boldsymbol{x} = \boldsymbol{y}$. Hence d_1 separates distinct points. For the triangle inequality, fix $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$. Then

$$d_1(\boldsymbol{x}, \boldsymbol{y}) + d_1(\boldsymbol{y}, \boldsymbol{z}) = \sum_{i=1}^{n} |x_i - y_i| + \sum_{i=1}^{n} |y_i - z_i| = \sum_{i=1}^{n} (|x_i - y_i| + |y_i - z_i|) \geqslant \sum_{i=1}^{n} |x_i - z_i|,$$

where the inequality comes from the triangle inequality for $|\cdot|$ on $\mathbb{R}$. The last term is precisely $d_1(\boldsymbol{x}, \boldsymbol{z})$, so we have established that $d_1(\boldsymbol{x}, \boldsymbol{y}) + d_1(\boldsymbol{y}, \boldsymbol{z}) \geqslant d_1(\boldsymbol{x}, \boldsymbol{z})$. Therefore d_1 satisfies the triangle inequality, meaning that d_1 meets all the conditions of being a metric on $\mathbb{R}^n$.

We now turn to d_∞. Suppose $d_\infty(\boldsymbol{x}, \boldsymbol{y}) = 0$. Then for every i we must have $|x_i - y_i| \leqslant 0$, so that $x_i = y_i$. Hence $\boldsymbol{x} = \boldsymbol{y}$ if $d_\infty(\boldsymbol{x}, \boldsymbol{y}) = 0$, and d_∞ separates distinct points. For the triangle inequality, fix $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$. Observe that for every i we necessarily have

$$d_\infty(\boldsymbol{x}, \boldsymbol{y}) + d_\infty(\boldsymbol{y}, \boldsymbol{z}) \geqslant |x_i - y_i| + |y_i - z_i| \geqslant |x_i - z_i|,$$

where the first inequality comes from the definition of d_∞ and the second inequality comes from the triangle inequality for $|\cdot|$ on $\mathbb{R}$. As this holds for every i, we may conclude that

$$d_\infty(\boldsymbol{x}, \boldsymbol{y}) + d_\infty(\boldsymbol{y}, \boldsymbol{z}) \geqslant \max\{|x_i - z_i|\}_{i=1}^{n} = d_\infty(\boldsymbol{x}, \boldsymbol{z}).$$

Hence d_∞ satisfies the triangle inequality, and d_∞ too meets all the conditions of being a metric on $\mathbb{R}^n$.

Question 6 (1 mark)

Exercise L4-4. Prove that if $P_1 = Q^{-1}P_2Q$ for some orthogonal matrix Q then multiplication by Q gives an isometry (assume P_1, P_2 positive-definite)

$$(\mathbb{R}^n, d_{P_1}) \longrightarrow (\mathbb{R}^n, d_{P_2})$$

That is, the metric we get on $\mathbb{R}^n$ from P_1 is *essentially the same* as the one we get from P_2.

Since Q is orthogonal, Q is invertible, and multiplication by Q is a bijection from $\mathbb{R}^n$ to $\mathbb{R}^n$ (in particular, it is surjective). It remains to show that multiplication preserves distance. Fix $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ (treated as column vectors). We wish to show that $d_{P_2}(Q\boldsymbol{x}, Q\boldsymbol{y}) = d_{P_1}(\boldsymbol{x}, \boldsymbol{y})$. We write

$$d_{P_2}(Q\boldsymbol{x}, Q\boldsymbol{y}) = (Q\boldsymbol{x})' P_2 (Q\boldsymbol{y}) = \boldsymbol{x}'(Q' P_2 Q)\boldsymbol{y}.$$

Since Q is orthogonal, we have $Q' = Q^{-1}$, so that $Q'P_2Q = Q^{-1}P_2Q = P_1$, and

$$d_{P_2}(Q\boldsymbol{x}, Q\boldsymbol{y}) = \boldsymbol{x}' P_1 \boldsymbol{y} = d_{P_1}(\boldsymbol{x}, \boldsymbol{y}).$$

This shows that multiplication by Q is distance-preserving. As mentioned, multiplication by Q is also surjective, so altogether multiplication by Q is an isometry from $(\mathbb{R}^n, d_{P_1})$ to $(\mathbb{R}^n, d_{P_2})$.