3.1

(i)

$$\begin{aligned}
(||\mathbf{x}+\mathbf{y}||^2 - ||\mathbf{x}-\mathbf{y}||^2) &= \langle \mathbf{x},\mathbf{y}\rangle + \langle \mathbf{y},\mathbf{x}\rangle + \langle \mathbf{x},\mathbf{y}\rangle + \langle \mathbf{y},\mathbf{x}\rangle \\
&= 2\langle \mathbf{x},\mathbf{y}\rangle + 2\langle \mathbf{y},\mathbf{x}\rangle \\
&= 4\langle \mathbf{x},\mathbf{y}\rangle \\
\frac{1}{4}(||\mathbf{x}+\mathbf{y}||^2 - ||\mathbf{x}-\mathbf{y}||^2) &= \langle \mathbf{x},\mathbf{y}\rangle
\end{aligned}$$

as desired.

(ii)

$$\begin{aligned}
(||\mathbf{x}+\mathbf{y}||^2 + ||\mathbf{x}-\mathbf{y}||^2) &= ||\mathbf{x}||^2 + ||\mathbf{x}||^2 + ||\mathbf{y}||^2 + ||\mathbf{y}||^2 \\
&= 2||\mathbf{x}||^2 + 2||\mathbf{y}||^2 \\
&= 2(||\mathbf{x}||^2 + ||\mathbf{y}||^2) \\
\frac{1}{2}(||\mathbf{x}+\mathbf{y}||^2 + ||\mathbf{x}-\mathbf{y}||^2) &= ||\mathbf{x}||^2 + ||\mathbf{y}||^2
\end{aligned}$$

as desired.

3.2

$$\begin{aligned}
&\frac{1}{4}(||\mathbf{x}+\mathbf{y}||^2 - ||\mathbf{x}-\mathbf{y}||^2 + i||\mathbf{x}-i\mathbf{y}||^2 - i||\mathbf{x}+i\mathbf{y}||^2) \\
&\qquad = \frac{1}{4}(4\langle \mathbf{x},\mathbf{y}\rangle)
\end{aligned}$$

3.3

(i) $\cos\theta = \frac{\int_0^1 (x)(x^5)dx}{\sqrt{\int_0^1 x^2 dx}\sqrt{\int_0^1 x^{10}dx}} = \frac{1/7}{\sqrt{1/3}\sqrt{1/11}} = \frac{1/7}{\sqrt{1/33}} = \frac{\sqrt{33}}{7};$ $\boxed{\theta = \cos^{-1}\left(\frac{\sqrt{33}}{7}\right)}$

(ii) $\cos\theta = \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^4 dx}\sqrt{\int_0^1 x^8 dx}} = \frac{1/7}{\sqrt{1/5}\sqrt{1/9}} = \frac{3\sqrt{5}}{7};$ $\boxed{\theta = \cos^{-1}\left(\frac{3\sqrt{5}}{7}\right)}$

3.8

(i) *Proof.* Let V be the inner product space $C([-\pi,\pi];\mathbb{R})$ with inner product

$$\langle f,g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)dt$$

Let $X = \operatorname{span}(S) \subset V$, where $S = \{\cos(t), \sin(t), \cos(2t), \sin(2t)\}$. Observe that

$$\begin{aligned}
\langle \cos(t), \sin(t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(t)\sin(t)dt = 0\\
\langle \cos(t), \cos(2t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(t)\cos(2t)dt = 0\\
\langle \cos(t), \sin(2t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(t)\sin(2t)dt = 0\\
\langle \sin(t), \cos(2t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\cos(2t)dt = 0\\
\langle \sin(t), \sin(2t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\sin(2t)dt = 0\\
\langle \cos(2t), \sin(2t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(2t)\sin(2t)dt = 0
\end{aligned}$$

Thus, each function is orthogonal to each other. Now, observe that

$$\begin{aligned}
\frac{1}{\pi}\int_{-\pi}^{\pi} \cos(t)\cos(t)dt &= 1\\
\frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\sin(t)dt &= 1\\
\frac{1}{\pi}\int_{-\pi}^{\pi} \cos(2t)\cos(2t)dt &= 1\\
\frac{1}{\pi}\int_{-\pi}^{\pi} \sin(2t)\sin(2t)dt &= 1
\end{aligned}$$

Therefore, S is an orthonormal set. □

(ii) $||t|| = \frac{1}{\pi}\int_{-\pi}^{\pi} t^2 dt = \frac{1}{\pi}(\frac{2\pi^3}{3}) = \frac{2\pi^2}{3}$

(iii) Recall from part (i) that each component of S is orthogonal to one another. Using this fact, $\text{proj}_X(\cos(3t)) = 0$

(iv) Note that $\langle\cos(t), t\rangle = 0$ and $\langle\cos(2t), t\rangle = 0$.
Now, $\langle\sin(t), t\rangle \sin(t) = 2\sin(t)$ and $\langle\sin(2t), t\rangle \sin(2t) = -\sin(2t)$
Take the sum of each of these yields $\boxed{\text{proj}_X(t) = 2\sin(t) - \sin(2t)}$

3.9

Proof. $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ Now observe that

$$\begin{aligned} R_\theta^H R_\theta &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta - \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= I \end{aligned}$$

By Theorem 3.2.15, the rotation is an orthonormal transformation. □

3.10

(i) Let $Q \in M_n(\mathbb{F})$ be orthonormal. So, $\langle m, n\rangle = \langle Q_m, Q_n\rangle$. Observe that

$$(Qm)^H Qn = m^H n \Leftrightarrow m^H Q^H Qn = m^H n$$

Hence, $Q^H Q = I$. Note also that $\langle Qm, Qn\rangle = \langle m, n\rangle$. So,

$$m^H n = (Qm)^H Qn = m^H Q^H Qn$$

(ii) *Proof.* Let $Q \in M_n(\mathbb{F})$ be an orthonormal matrix. Observe that

$$||x|^2| = \langle x, x\rangle = \langle Qx, Qx\rangle = ||Qx||^2$$

□

(iii) *Proof.* Let $Q \in M_n(\mathbb{F})$ be an orthonormal matrix. Observe that

$$\begin{aligned} QQ^H = Q^H Q = I \\ \Rightarrow Q^H = Q^{-1} \end{aligned}$$

Now observe that

$$\begin{aligned} (Q^H)^H = Q \\ (Q^H)(Q^H)^H = (Q)^H Q = I \end{aligned}$$

So $Q^H = Q^{-1}$ is an orthonormal matrix. □

(iv)

(v) No. Let $Q = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ so $Q^T = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$. Note that $\det(Q) = 1$. Observe that

$$\begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \neq I$$

and $Q \in M_n(\mathbb{F})$

(vi) Let $Q_1, Q_2 \in M_n(\mathbb{F})$ be orthonormal matrices. Hence, $Q_1^H Q_1 = I$ and $Q_2^H Q_2 = I$. Let $Q = Q_1 Q_2$. Observe that

$$\begin{aligned} Q^H Q &= (Q_1 Q_2)^H (Q_1 Q_2) \\ &= Q_2^H Q_1^H Q_1 Q_2 \\ &= Q_2^H Q_2 \\ &= I \end{aligned}$$

Therefore, $Q_1 Q_2$ is an orthonormal matrix.

3.11 When we apply the Gram-Schmidt orthonormalization process to a collection of linearly *dependent* vectors then the kth step will produce $\mathbf{0}$

since $\mathbf{x}_k$ is a linear combination of $\mathbf{x}_1, ..., \mathbf{x}_{k-1}$. Note also that by producing the zero vector or any multiple of it, the length of it will not be equal to one, which is a necessary condition to be orthonormal.

3.16

(i)

(ii)

3.17

Proof. Let $A \in M_{m\times n}$ have rank $n \leq m$, and let $A = \hat{Q}\hat{R}$ be a reduced QR decomposition. Note that $\hat{Q}$ is an $m \times n$ orthonormal matrix and $\hat{R}$ is an $n \times n$ upper-triangular matrix (see Remark 3.3.10). Observe that

$$\begin{aligned} A^H A\mathbf{x} &= A^H\mathbf{b} \\ (\hat{Q}\hat{R})^H\hat{Q}\hat{R}\mathbf{x} &= (\hat{Q}\hat{R})^H\mathbf{b} \\ \hat{R}^H\hat{Q}^H\hat{Q}\hat{R}\mathbf{x} &= \hat{R}^H\hat{Q}^H\mathbf{b} \\ \hat{R}^H\hat{R}\mathbf{x} &= \hat{R}^H\hat{Q}^H\mathbf{b} \\ \hat{R}\mathbf{x} &= \hat{Q}^H\mathbf{b} \end{aligned}$$

□

3.23 Let $(V, ||\cdot||)$ be a normed linear space. Observe that

$$\begin{aligned} ||x - y + y|| &\leq ||x - y|| + ||y|| \\ ||x|| &\leq ||x - y|| + ||y|| \\ ||x|| - ||y|| &\leq ||x - y|| \end{aligned}$$

Also observe that

$$\begin{aligned} ||y - x + x|| &\leq ||y - x|| + ||x|| \\ ||y|| &\leq ||x - y|| + ||x|| \\ ||y|| - ||x|| &\leq ||x - y|| \end{aligned}$$

Therefore, $\big|||x| - ||y|||\big| \leq ||x-y||$

3.24 Let $C([a,b];\mathbb{F})$ be the vector space of all continuous functions from $[a,b] \subset \mathbb{R}$ to $\mathbb{F}$.

(i) $||f||_{L^1} = \int_a^b |f(t)|dt$

- positivity: $|f(t)| \geq 0$ by definition. If $f(t) = 0$, then $|f(t) = 0$ and if $f(t) \neq 0$ then $|f(t) \neq 0$. Thus $\int_a^b |f(t)|dt$ is positive.
- scale preservation: Let $\alpha \in \mathbb{R}$ be a scalar. Note that $||\alpha f(t)||_{L^1} = \int_a^b |\alpha f(t)|dt = \int_a^b |\alpha||f(t)|dt = |\alpha| \int_a^b |f(t)|dt$
- triangle inequality:

$$\begin{aligned} ||f+g||_{L^1} &= \int_a^b |f(t)+g(t)|dt \\ &\leq \int_a^b |f(t)| + |g(t)|dt \\ &= \int_a^b |f(t)|dt + \int_a^b |g(t)|dt \end{aligned}$$

Therefore, $||f||_{L^1}$ is a norm on $C([a,b];\mathbb{F})$.

(ii) $||f||_{L^2} = \big(\int_a^b |f(t)|^2 dt\big)^{\frac{1}{2}}$

- positivity: $|f(t)|$ is positive (see exercise 3.24 (i)). It follows that $|f(t)|^2$ is also positive. Hence $\int_a^b |f(t)|^2 dt$ is positive. Now, if $\int_a^b |f(t)|^2 dt = 0$ then $\big(\int_a^b |f(t)|^2 dt\big)^{\frac{1}{2}} = 0$. But if $\int_a^b |f(t)|^2 dt \neq 0$ then $\big(\int_a^b |f(t)|^2 dt\big)^{\frac{1}{2}} \neq 0$ and so $\big(\int_a^b |f(t)|^2 dt\big)^{\frac{1}{2}}$ is positive.

- scalar preservation: For some scalar $\alpha \in \mathbb{R}$ observe that

$$
\begin{aligned}
||\alpha f||_{L^2} &= \left(\int_a^b |\alpha f(t)|^2 dt \right)^{\frac{1}{2}} \\
&= \left(\int_a^b |\alpha|^2 |f(t)|^2 dt \right)^{\frac{1}{2}} \\
&= \left(|\alpha|^2 \int_a^b |f(t)|^2 dt \right)^{\frac{1}{2}} \\
&= |\alpha| \left(\int_a^b |f(t)|^2 dt \right)^{\frac{1}{2}}
\end{aligned}
$$

- triangle inequality: $||f+g||_{L^2} = \left(\int_a^b |f(t)+g(t)|^2 \right)^{\frac{1}{2}}$. Observe that

$$
\begin{aligned}
(||f+g||_{L^2})^2 &= \int_a^b |f(t)+g(t)|^2 dt \\
&= \int_a^b |f(t)|^2 + 2|f(t)g(t)| + |g(t)|^2 dt \\
&= ||f||_{L^2} + 2|f(t)g(t)| + ||g||_{L^2} \\
&\leq ||f||_{L^2} + 2||f||_{L^2}||g||_{L^2} + ||g||_{L^2} \\
&= (||f||_{L^2} + ||g||_{L^2})^2
\end{aligned}
$$

  Hence $||f+g||_{L^2} \leq ||f||_{L^2} + ||g||_{L^2}$

Therefore $||f||_{L^2}$ is a norm on $C([a,b];\mathbb{F})$.

(iii) $||f||_{L^\infty} = \sup_{x\in[a,b]} |f(x)|$

- positivity: $|f(x)| \geq 0$ by definition. If $f(x) = 0$ then $|f(x)| = 0$, and if $f(x) \neq 0$ then $|f(x) \neq 0$. Thus $\sup_{x\in[a,b]} |f(x)|$ is positive.

- scalar preservation: For some scalar $\alpha \in \mathbb{R}$ observe that

$$\begin{aligned} ||\alpha f||_{L^\infty} &= \sup_{x\in[a,b]} |\alpha f(x)| \\ &= \sup_{x\in[a,b]} |\alpha||f(x)| \\ &= |\alpha| \sup_{x\in[a,b]} |f(x)| \end{aligned}$$

- triangle inequality: Observe that

$$\begin{aligned} ||f+g||_{L^\infty} &= \sup_{x\in[a,b]} |f(x)+g(x)| \\ &\le \sup_{x\in[a,b]} (|f(x)|+|g(x)|) \\ &= \sup_{x\in[a,b]} |f(x)| + \sup_{x\in[a,b]} |g(x)| \end{aligned}$$

Therefore, $||f||_{L^\infty}$ is a norm on $C([a,b]);\mathbb{F}$.

3.26 Reflexive: Note that $||x||_a = ||x||_a$. Given $0 \le m \le M$ we have $m||x||_a \le ||x||_a \le M||x||_a$ where $||x||_a \le ||x||_a \le ||x||_a$ for $m, M = 1$. Thus, $||x||_a \sim ||x||_a$.
Symmetric: Suppose $||x||_a \sim ||x||_b$. There exists $m, M \in R$ with $0 < m \le M$ such that $m||x||_a \le ||x||_b \le M||x||_a$. Hence $\frac{1}{M}||x||_b \le ||x||_a \le \frac{1}{m}||x||_b$. Thus $||x||_b \sim ||x||_a$
Transitive: Suppose $||x||_a \sim ||x||_b$ and $||x||_b \sim ||x||_c$. There exists $m_1, M_1, m_2, M_2 \in \mathbb{R}$ with $0 < m_1 \le M_1$ and $0 < m_2 \le M_2$ such that $m_1||x||_a \le ||x||_b \le M_1||x||_a$ and $m_2||x||_b \le ||x||_c \le M_2||x||_b$. Now $m_2||x||_b \le ||x||_c$ so $m_1m_2||x||_a \le ||x||_c$ and $||x||_c \le M_2||x||_b$. So $||x||_c \le M_1M_2||x||_a$. Putting it all together we have $m_1m_2||x||_a \le ||x||_c \le M_1M_2||x||_a$. Thus $||x||_a \sim ||x||_c$.
Therefore, topological equivalence is an equivalence relation.

(i)

$$|x_1|^2 + |x_2|^2 + .. + |x_n|^2 \leq (|x_1| + |x_2| + ... + |x_n|)^2$$
$$(|x_1|^2 + |x_2|^2 + .. + |x_n|^2)^{\frac{1}{2}} \leq |x_1| + |x_2| + ... + |x_n|$$
$$||x||_2 \leq ||x||_1$$

Now, $|\langle x, \mathbf{1}\rangle| \leq ||x||_2||\mathbf{1}||_2$ (Cauchy-Schwartz) where $||\mathbf{1}||_2 = (1^2 + 1^2 + ... + 1^2)^{\frac{1}{2}} = \sqrt{n}$. So, $||x||_1 \leq \sqrt{n}||x||_2$.
Therefore, $||x||_2 \leq ||x||_1 \leq \sqrt{n}||x||_2$

(ii)

3.28 Let A be an $n \times n$ matrix.

(i)

3.29
3.30

Proof. Let $S \in M_n(\mathbb{F})$ be an invertible matrix. Given any matrix norm $|| \cdot ||$ on M_n define $|| \cdot ||_S$ by $||A||_S = ||SAS^{-1}||$. By definition 3.5.15, any norm $|| \cdot ||$ on the finite-dimensional vector space $M_n(\mathbb{F})$ that satisfies the submultiplicative property is called a *matrix norm.* □

3.37
3.38
3.39

Proof. let V and W be finite-dimensional inner-product spaces.

(i) Let $S, T \in \mathscr{L}(V; W)$. Observe that

$$\langle (S + T)(\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, (S + T)^*(\mathbf{w})\rangle$$

which follows from the definition of adjoint. Now,

$$\begin{aligned}\langle (S+T)(\mathbf{v}), \mathbf{w}\rangle &= \langle S(\mathbf{v}) + T(\mathbf{v}), \mathbf{w}\rangle \\ &= \langle S(\mathbf{v}), \mathbf{w}\rangle + \langle T(\mathbf{v}), \mathbf{w}\rangle \\ &= \langle \mathbf{v}, S^*(\mathbf{w})\rangle + \langle \mathbf{v}, T^*(\mathbf{w})\rangle \\ &= \langle \mathbf{v}, (S^* + T^*)(\mathbf{w})\rangle\end{aligned}$$

Hence, $(S+T)^* = S^* + T^*$ Now, observe that for $\alpha \in \mathbb{F}$

$$\langle (\alpha T)(\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, (\alpha T)^*(\mathbf{w})\rangle$$

Additionally,

$$\begin{aligned}\langle (\alpha T)(\mathbf{v}), \mathbf{w}\rangle &= \langle \alpha T(\mathbf{v}), \mathbf{w}\rangle \\ &= \alpha\langle \mathbf{v}, T^*(\mathbf{w})\rangle \\ &= \langle \mathbf{v}, \bar{\alpha} T^*(\mathbf{w})\rangle\end{aligned}$$

Hence, $(\alpha T^*) = \bar{\alpha} T^*$

(ii) Let $S \in \mathscr{L}(V; W)$. Observe that

$$\langle (S^*)^*(\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, (S^*)^*(\mathbf{w})\rangle$$

Additionally,

$$\begin{aligned}\langle S^*(\mathbf{w}), \mathbf{v}\rangle &= \overline{\langle \mathbf{v}, S^*(\mathbf{w})\rangle} \\ &= \overline{\langle S(\mathbf{v}), \mathbf{w}\rangle} \\ &= \langle \mathbf{w}, S(\mathbf{v})\rangle\end{aligned}$$

Hence, $(T^*)^* = T$

(iii) Let $S, T \in \mathscr{L}(V)$. Observe that

$$\langle (ST)(\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, (ST)^*(\mathbf{w})\rangle$$

Additionally,

$$\begin{aligned}\langle (ST)(\mathbf{v}), \mathbf{w}\rangle &= \langle S(T(\mathbf{v})), \mathbf{w}\rangle \\ &= \langle T(\mathbf{v}), S^*(\mathbf{w})\rangle \\ &= \langle \mathbf{v}, T^*(S^*(\mathbf{w}))\rangle \\ &= \langle \mathbf{v}, (T^*S^*)(\mathbf{w})\rangle\end{aligned}$$

Hence, $(ST)^* = T^*S^*$

(iv) Let $T \in \mathscr{L}(V)$ and suppose T is invertible. Observe that

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

But $I^* = I$. From property (iii) we have that $T^{**}((T^*)^{-1})^* = I$. Note that $T^* * = T$, so we have that

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

□

3.40

3.44

3.45

3.46

3.47 Assume A is an $m \times n$ matrix of rank n. Let $P = A(A^HA)^{-1}A^H$

(i)

$$\begin{aligned}P^2 &= A(A^HA)^{-1}A^HA(A^HA)^{-1}A^H && (0.1)\\ &= A(A^HA)^{-1}A^H && (0.2)\\ & && (0.3)\end{aligned}$$

(ii)

$$\begin{aligned}
P^H &= (A(A^H A)^{-1} A^H)^H \\
&= A((A^H A)^{-1})^H A^H \\
&= A((A^H A)^H)^{-1} A^H \\
&= A(A^H A)^{-1} A^H \\
&= P
\end{aligned}$$

(iii)

$$\operatorname{rank}(P) = \operatorname{rank}(A(A^H A)^{-1} A^H)$$

Now because P is idempotent by (i) we have that

$$\begin{aligned}
\operatorname{tr}(A(A^H A)^{-1} A^H) &= \operatorname{tr}(A^H A(A^H A)^{-1}) \\
&= \operatorname{tr}(I)
\end{aligned}$$

where I is $n \times n$, so it has rank n. Therefore, rank(P)=n.

3.48
3.50