Problem Set #2

Reiko Laski

Exercise 1

(i) *Proof:*

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

(ii) *Proof:*

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

Exercise 2
Proof:

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= \frac{1}{4}(||\mathbf{x}+\mathbf{y}||^2 - ||\mathbf{x}-\mathbf{y}||^2 + i||\mathbf{x}-i\mathbf{y}||^2 - i||\mathbf{x}+i\mathbf{y}||^2) \\
&= \frac{1}{4}(||\mathbf{x}+\mathbf{y}||^2 - ||\mathbf{x}-\mathbf{y}||^2) + \frac{1}{4}(i||\mathbf{x}-i\mathbf{y}||^2 - i||\mathbf{x}+i\mathbf{y}||^2) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i||\mathbf{x}-i\mathbf{y}||^2 - i||\mathbf{x}+i\mathbf{y}||^2) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i\langle \mathbf{x}-i\mathbf{y}, \mathbf{x}-i\mathbf{y} \rangle - i\langle \mathbf{x}+i\mathbf{y}, \mathbf{x}+i\mathbf{y} \rangle) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i\langle \mathbf{x}, \mathbf{x}-i\mathbf{y} \rangle + i^2\langle \mathbf{y}, \mathbf{x}-i\mathbf{y} \rangle - i\langle \mathbf{x}, \mathbf{x}+i\mathbf{y} \rangle + (-i)^2\langle \mathbf{y}, \mathbf{x}+i\mathbf{y} \rangle) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i\langle \mathbf{x}, \mathbf{x} \rangle + i^2\langle \mathbf{x}, \mathbf{y} \rangle + i^2\langle \mathbf{y}, \mathbf{x} \rangle + i^3\langle \mathbf{y}, \mathbf{y} \rangle \\
&\qquad - i\langle \mathbf{x}, \mathbf{x} \rangle + (-i)^2\langle \mathbf{x}, \mathbf{y} \rangle + (-i)^2\langle \mathbf{y}, \mathbf{x} \rangle + (-i^3)\langle \mathbf{y}, \mathbf{y} \rangle) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i^2\langle \mathbf{x}, \mathbf{y} \rangle + i^2\langle \mathbf{y}, \mathbf{x} \rangle + (-i)^2\langle \mathbf{x}, \mathbf{y} \rangle + (-i)^2\langle \mathbf{y}, \mathbf{x} \rangle) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(-\langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle
\end{aligned}$$

Exercise 3
(i) $\cos\ \theta = \frac{\langle x, x^5 \rangle}{||x|| \cdot ||x^5||}$
$\langle x, x^5 \rangle = \int_0^1 x \cdot x^5\ dx = \frac{1}{7}$
$\langle x, x \rangle = \int_0^1 x \cdot x\ dx = \frac{1}{3} \implies ||x|| = \frac{1}{\sqrt{3}}$
$\langle x^5, x^5 \rangle = \int_0^1 x^5 \cdot x^5\ dx = \frac{1}{11} \implies ||x^5|| = \frac{1}{\sqrt{11}}$
$\implies \theta = \cos^{-1}(\frac{\sqrt{33}}{7})$

(ii) $\cos\ \theta = \frac{\langle x^2, x^4 \rangle}{||x^2|| \cdot ||x^4||}$
$\langle x^2, x^4 \rangle = \int_0^1 x^2 \cdot x^4\ dx = \frac{1}{7}$
$\langle x^2, x^2 \rangle = \int_0^1 x^2 \cdot x^2\ dx = \frac{1}{5} \implies ||x^2|| = \frac{1}{\sqrt{5}}$
$\langle x^4, x^4 \rangle = \int_0^1 x^4 \cdot x^4\ dx = \frac{1}{9} \implies ||x^4|| = \frac{1}{3}$
$\implies \theta = \cos^{-1}(\frac{3\sqrt{5}}{7})$

Exercise 8
(i) $\langle \cos t, \sin t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos t \sin t\ dt = 0$
$\langle \cos t, \cos 2t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos t \cos 2t\ dt = 0$
$\langle \cos t, \sin 2t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos t \sin 2t\ dt = 0$
$\langle \sin t, \cos 2t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin t \cos 2t\ dt = 0$
$\langle \sin t, \sin 2t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos t \sin 2t\ dt = 0$
$\langle \cos 2t, \sin 2t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos 2t \sin 2t\ dt = 0$

$\langle \cos t, \cos t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos t \cos t\ dt = 1$
$\langle \sin t, \sin t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin t \sin t\ dt = 1$
$\langle \cos 2t, \cos 2t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos 2t \cos 2t\ dt = 1$
$\langle \sin 2t, \sin 2t \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin 2t \sin 2t\ dt = 1$

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

(iii)

$$\begin{aligned}
proj_X(\cos 3t) &= \langle \cos t, \cos 3t\rangle \cdot \cos t + \langle \sin t, \cos 3t\rangle \cdot \sin t \\
&\qquad + \langle \cos 2t, \cos 3t\rangle \cdot \cos 2t + \langle \sin 2t, \cos 3t\rangle \cdot \sin 2t \\
&= \left(\frac{1}{\pi}\int_{-\pi}^{\pi} \cos t \cos 3t \ dt\right) \cdot \cos t + \left(\frac{1}{\pi}\int_{-\pi}^{\pi} \sin t \cos 3t \ dt\right) \cdot \sin t \\
&\qquad + \left(\frac{1}{\pi}\int_{-\pi}^{\pi} \cos 2t \cos 3t \ dt\right) \cdot \cos 2t + \left(\frac{1}{\pi}\int_{-\pi}^{\pi} \sin 2t \cos 3t \ dt\right) \cdot \sin 2t \\
&= 0
\end{aligned}$$

(iv)

$$\begin{aligned}
proj_X(t) &= \langle \cos t, t\rangle \cdot \cos t + \langle \sin t, t\rangle \cdot \sin t \\
&\qquad + \langle \cos 2t, t\rangle \cdot \cos 2t + \langle \sin 2t, t\rangle \cdot \sin 2t \\
&= \left(\frac{1}{\pi}\int_{-\pi}^{\pi} (\cos t) \ t \ dt\right) \cdot \cos t + \left(\frac{1}{\pi}\int_{-\pi}^{\pi} (\sin t) \ t \ dt\right) \cdot \sin t \\
&\qquad + \left(\frac{1}{\pi}\int_{-\pi}^{\pi} (\cos 2t) \ t \ dt\right) \cdot \cos 2t + \left(\frac{1}{\pi}\int_{-\pi}^{\pi} (\sin 2t) \ t \ dt\right) \cdot \sin 2t \\
&= 2\sin t - \sin(2t)
\end{aligned}$$

Exercise 9

The rotation matrix R_θ is orthonormal since $R_\theta R_\theta^H = I$:

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

Exercise 10

(i) $Q \in M_n(\mathbb{F}^n)$ is an orthonormal matrix if and only if $Q^H Q = QQ^H = I$.

Proof:

Let Q be an orthonormal matrix. Then we have that $\langle \mathbf{x}, \mathbf{y}\rangle = \langle Q\mathbf{x}, Q\mathbf{y}\rangle = (Q\mathbf{x})^H Q\mathbf{y} = \mathbf{x}^H Q^H Q\mathbf{y}$. Since we know that $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^H\mathbf{y}$, it must be that $Q^H Q = I$. Also, $QQ^H = I$ since Q is invertible.

Now suppose that for a matrix Q, we have that $Q^H Q = QQ^H = I$. Consider the inner product $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = (Q\mathbf{x})^H Q\mathbf{y} = \mathbf{x}^H Q^H Q\mathbf{y} = \mathbf{x}^H\mathbf{y} = \langle \mathbf{x}, \mathbf{y}\rangle$. Therefore, Q must be an orthonormal matrix.

(ii) If $Q \in M_n(\mathbb{F})$ is an orthonormal matrix, then $||Q\mathbf{x}|| = ||\mathbf{x}||$ for all $\mathbf{x} \in \mathbb{F}^n$.

Proof:

$||Q\mathbf{x}|| = \sqrt{\langle Q\mathbf{x}, Q\mathbf{y}\rangle} = \sqrt{\langle \mathbf{x}, \mathbf{y}\rangle} = ||\mathbf{x}||$.

(iii) If $Q \in M_n(\mathbb{F}^n)$ is an orthonormal matrix, then so is Q^{-1}.

Proof:

Assume Q is an orthonormal matrix. Then we know that $Q^H Q = QQ^H = I$, so $Q^{-1} = Q^H$. To prove that Q^H is orthonormal, we can show that $(Q^H)^H Q^H = Q^H (Q^H)^H = I$.

(iv) The columns of an orthonormal matrix $Q \in M_n(\mathbb{F}^n)$ are orthonormal.
Proof:
Assume that Q is orthonormal. Consider standard basis vectors $e_i, e_j \in \mathbb{F}^n$. The ith column of Q is $Q\mathbf{e_i}$. Since Q is orthonormal, we have that $\langle Q\mathbf{e_i}, Q\mathbf{e_j}\rangle = \langle e_i, e_j\rangle = \delta_{ij}$. Therefore the columns of Q must be orthonormal.

(v) If $Q \in M_n(\mathbb{F}^n)$ is an orthonormal matrix, then $|det(Q)| = 1$. *Proof:*
Since Q is orthonormal, we know that $QQ^H = I$. Then $det(QQ^H) = det(I) = 1$. By the properties of determinants, we have that $det(QQ^H) = det(Q) \cdot det(Q^H) = 1$. Using the fact that Q and Q^H have the same determinants, we conclude that $|det(Q)| = 1$. No. Consider the matrix

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

(vi) If $Q_1, Q_2 \in M_n(\mathbb{F}^n)$ are orthonormal matrices, then the product $Q_1 Q_2$ is also an orthonormal matrix.
Proof:
$(Q_1 Q_2)^H Q_1 Q_2 = Q_2^H Q_1^H Q_1 Q_2 = I$
$Q_1 Q_2 (Q_1 Q_2)^H = Q_1 Q_2 Q_2^H Q_1^H = I$.
Therefore $Q_1 Q_2$ is an orthonormal matrix.

Exercise 11
Let $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\}$ be a linearly dependent set. Assume that the vector $\mathbf{x}_k$ is linearly dependent on $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{k-1}\}$, i.e. $\mathbf{x}_k \in X = Span(\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{k-1}\})$. When we apply the Gram-Schmidt orthonormalization process, we have that $\mathbf{p}_{k-1} = proj_X(\mathbf{x}_k) = \mathbf{x}_k \implies \mathbf{q}_k = \mathbf{0}$. If we remove all the zero vectors from the set, we have an orthonormal set of linearly dependent vectors.

Exercise 16
(i) Let $Q_1 R_1$ and $Q_2 R_2$ be distinct QR decompositions of a matrix A s.t. $Q_2 = Q_1 D$ and $R_2 = D^{-1} R_2$ where D is a diagonal matrix with all its diagonal entries ± 1. Then Q_2 is orthonormal since Q_1 is orthonormal. Also, R_2 is upper triangular since R_1 is upper triangular.

(ii) Assume that $A = Q_1 R_1 = Q_2 R_2$ s.t. Q_i is orthonormal and R_i is upper triangular with only positive diagonal entries. Thus R_i is invertible and $Q_i Q_i^H = I$. Then we have that $R_1 R_2^{-1} = Q_1^H Q_2$. Since the R_i are upper triangular with only positive entries on the diagonal, $R_1 R_2^{-1}$ must be upper triangular with only positive entries on the diagonal as well. Also, since the Q_i are orthonormal, $Q_1^H Q_2$ must be orthonormal. Then, $R_1 R_2^{-1} = Q_1^H Q_2 = I$, so $R_1 = R_2$ and $Q_1 = Q_2$.

Exercise 17
Let $A = \hat{Q}\hat{R}$.
$A^H A\mathbf{x} = A^H\mathbf{b}$
$\implies (\hat{Q}\hat{R})^H\hat{Q}\hat{R}\mathbf{x} = (\hat{Q}\hat{R})^H\mathbf{b}$
$\implies \hat{R}^H\hat{Q}^H\hat{Q}\hat{R}\mathbf{x} = \hat{R}^H\hat{Q}^H\mathbf{b}$
$\implies \hat{Q}\hat{R}\mathbf{x} = \mathbf{b}$
$\implies \hat{R}\mathbf{x} = \hat{Q}^H\mathbf{b}$

Exercise 23
By the triangle inequality,

$$||\mathbf{x}|| = ||(\mathbf{x}-\mathbf{y})+\mathbf{y}|| \leq ||\mathbf{x}-\mathbf{y}|| + ||\mathbf{y}|| \Leftrightarrow ||\mathbf{x}|| - ||\mathbf{y}|| \leq ||\mathbf{x}-\mathbf{y}||.$$

Similarly and by scale preservation,

$$||\mathbf{y}|| = |-1|\cdot||\mathbf{y}|| = ||-\mathbf{y}|| = ||(\mathbf{x}-\mathbf{y})+\mathbf{x}|| \leq ||\mathbf{x}-\mathbf{y}||+||\mathbf{y}|| \Leftrightarrow ||\mathbf{y}||-||\mathbf{x}|| \leq ||\mathbf{x}-\mathbf{y}||$$
$$\implies |||\mathbf{x}|| - ||\mathbf{y}||| \leq ||\mathbf{x}-\mathbf{y}||$$

Exercise 24
(i) $||f||_{L^1} = \int_a^b |f(t)|dt$
1. Positivity: $|f(t)| \geq 0$ for all $x \implies \int_a^b |f(t)|dt \geq 0$.
Also, $\int_a^b |f(t)|dt = 0$ if and only if $|f(t)| = 0$.

2. Scale preservation: $||af||_{L^1} = \int_a^b |af(t)|dt = |a|\int_a^b |f(t)|dt = |a|||f||_{L^1}$

3. Triangle inequality: Consider $f, g \in C[a,b]$.

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

(ii) $||f||_{L^2} = (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}}$
1. Positivity: $|f(t)|^2 \geq 0$ for all $x \implies (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} \geq 0$.
Also, $(\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} = 0$ if and only if $|f(t)| = 0$.

2. Scale preservation: $||af||_{L^2} = (\int_a^b |af(t)|^2 dt)^{\frac{1}{2}} = |a|(\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} = |a|||f||_{L^2}$

3. Triangle inequality: Consider $f, g \in C[a,b]$.

$$\begin{aligned}||f+g|| &= \Big(\int_a^b |f(t)+g(t)|^2 dt\Big)^{\frac{1}{2}} \\ &\leq \Big(\int_a^b |f(t)|^2 dt + \int_a^b |g(t)|^2 dt\Big)^{\frac{1}{2}} \\ &\leq ||f||_{L^2} + ||g||_{L^2}\end{aligned}$$

(iii) $||f||_{L^\infty} = \sup_{x\in[a,b]|f(x)|}$
1. Positivity: $|f(x)| \geq 0$ for all $x \implies \sup_{x\in[a,b]} |f(x)| \geq 0$.
Also, $\sup_{x\in[a,b]} |f(x)| = 0$ if and only if $|f(x)| = 0$.

2. Scale preservation: $||af||_{L^\infty} = \sup_{x\in[a,b]} |a||f(x)| = |a| \sup_{x\in[a,b]} |f(x)| = |a|||f||_{L^\infty}$

3. Triangle inequality: For any $x \in [a,b]$

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

Exercise 26
Prove that topological equivalence is an equivalence relation. We check that it satisfies the three conditions:

1. Reflexivity: $||\cdot||_a \sim ||\cdot||_a$
We see that $m||\mathbf{x}||_a \leq ||\mathbf{x}||_a \leq M||\mathbf{x}||_a$ holds for a choice of constants $m = 1$ and $M = 2$.

2. Symmetry: $||\cdot||_a \sim ||\cdot||_b \implies ||\cdot||_b \sim ||\cdot||_a$
If $m||\mathbf{x}||_a \leq ||\mathbf{x}||_b \leq M||\mathbf{x}||_a$, then the relation holds for $\frac{1}{M}||\mathbf{x}||_b \leq ||\mathbf{x}||_a \leq \frac{1}{m}||\mathbf{x}||_b$.

3. Transitivity: $||\cdot||_a \sim ||\cdot||_b$ and $||\cdot||_b \sim ||\cdot||_c \implies ||\cdot||_a \sim ||\cdot||_c$
If $m_1||\mathbf{x}||_a \leq ||\mathbf{x}||_b \leq M_1||\mathbf{x}||_a$ and $m_2||\mathbf{x}||_b \leq ||\mathbf{x}||_c \leq M_2||\mathbf{x}||_b$, then the relation holds for $m_1m_2||\mathbf{x}||_a \leq ||\mathbf{x}||_c \leq M_1M_2||\mathbf{x}||_a$.

(i) Show that $||\mathbf{x}||_2 \leq ||\mathbf{x}||_1 \leq \sqrt{n}||\mathbf{x}||_2$.

$$\begin{aligned}||\mathbf{x}||_1^2 = (|x_1| + |x_2| + \cdots + |x_n|)^2 \geq |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 \\ \implies ||\mathbf{x}||_1 \geq (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{\frac{1}{2}} = ||\mathbf{x}||_2\end{aligned}$$

Also,

$$||\mathbf{x}||_1 = |\langle \mathbf{x}, \mathbf{1} \rangle| \le ||\mathbf{x}||_2 \cdot ||\mathbf{1}||_2 \text{ where } ||\mathbf{1}||_2 = \left(\sum_{j=1}^{n} 1^2 \right)^{\frac{1}{2}} = \sqrt{n}$$
$$\implies ||\mathbf{x}||_2 \le ||\mathbf{x}||_1 \le \sqrt{n}||\mathbf{x}||_2$$

(ii) Show that $||\mathbf{x}||_\infty \le ||\mathbf{x}||_2 \le \sqrt{n}||\mathbf{x}||_\infty$.

$$||\mathbf{x}||_\infty = \sup_i |x_i| = (\sup_i |x_i|^2)^{\frac{1}{2}} \le (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{\frac{1}{2}} = ||\mathbf{x}||_2$$

Also,

$$||\mathbf{x}||_2 = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{\frac{1}{2}} \le (n|x_n|^2)^{\frac{1}{2}} = \sqrt{n}||\mathbf{x}||_2$$
$$\implies ||\mathbf{x}||_\infty \le ||\mathbf{x}||_2 \le \sqrt{n}||\mathbf{x}||_\infty$$

Exercise 28
(i) $\frac{1}{\sqrt{n}}||A||_2 \le ||A||_1 \le \sqrt{n}||A||_2$
Proof:
From Exercise 26(i), we know that

$$||A\mathbf{x}||_2 \le ||A\mathbf{x}||_1 \le \sqrt{n}||A\mathbf{x}||_2 \quad \text{and} \quad \frac{1}{\sqrt{n}}||\mathbf{x}||_2 \le ||\mathbf{x}||_1 \le ||\mathbf{x}||_2$$
$$\implies \frac{1}{\sqrt{n}} \frac{||A\mathbf{x}||_2}{||\mathbf{x}||_2} \le \frac{||A\mathbf{x}||_1}{||\mathbf{x}||_1} \le \sqrt{n}\frac{||A\mathbf{x}||_2}{||\mathbf{x}||_2}$$
$$\implies \frac{1}{\sqrt{n}}||A||_2 \le ||A||_1 \le \sqrt{n}||A||_2$$

(ii) $\frac{1}{\sqrt{n}}||A||_\infty \le ||A||_2 \le \sqrt{n}||A||_\infty$
Proof:
From Exercise 26(ii), we know that

$$||A\mathbf{x}||_\infty \le ||A\mathbf{x}||_2 \le \sqrt{n}||A\mathbf{x}||_\infty \quad \text{and} \quad \frac{1}{\sqrt{n}}||\mathbf{x}||_\infty \le ||\mathbf{x}||_2 \le ||\mathbf{x}||_\infty$$
$$\implies \frac{1}{\sqrt{n}} \frac{||A\mathbf{x}||_\infty}{||\mathbf{x}||_\infty} \le \frac{||A\mathbf{x}||_2}{||\mathbf{x}||_2} \le \sqrt{n}\frac{||A\mathbf{x}||_\infty}{||\mathbf{x}||_\infty}$$
$$\implies \frac{1}{\sqrt{n}}||A||_\infty \le ||A||_2 \le \sqrt{n}||A||_\infty$$

Exercise 29
Any orthonormal matrix $Q \in M_n(\mathbb{F})$ has $||Q|| = 1$.
Proof:
$||Q|| = \sup_{\mathbf{x} \ne \mathbf{0}} \frac{||Q\mathbf{x}||_2}{||\mathbf{x}||_2} = \frac{||\mathbf{x}||_2}{||\mathbf{x}||_2} = 1$

For any $\mathbf{x} \in \mathbb{F}^n$, let $R_x : M_n(\mathbb{F} \to \mathbb{F}^n)$ be the linear transformation $A \mapsto A\mathbf{x}$. The induced norm of the transformation R_x is equal to $||\mathbf{x}||_2$.

Proof:
$||R_x|| = \sup_{||A||_2 \neq 0} \frac{||A\mathbf{x}||_2}{||A||} = \sup_{||A||_2 \neq 0} \frac{||A\mathbf{x}||_2 ||\mathbf{x}||_2}{||A|| ||\mathbf{x}||_2} \leq \sup_{||A||_2 \neq 0} \frac{||A\mathbf{x}||_2 ||\mathbf{x}||_2}{||A\mathbf{x}||_2} = ||\mathbf{x}_2||$
Since A is orthonormal, we have that $||A|| = 1$ and $||A\mathbf{x}||_2 = ||\mathbf{x}||_2 \implies ||R_x|| = ||\mathbf{x}_2||$.

Exercise 30
Let $S \in M_n(\mathbb{F})$ be an invertible matrix. Given any matrix norm $|| \cdot ||$ on M_n, define $|| \cdot ||_S$ by $||A||_S = ||SAS^{-1}||$. Then $|| \cdot ||_S$ is a matrix norm on M_n.
Proof:
1. Positivity: $||SAS^{-1}|| \geq 0$ by definition.
Also, $||SAS^{-1}|| = 0$ if and only if $A = 0$.

2. Scale preservation: Let $a \in \mathbb{R}$. Then $||aA||_S = ||aSAS^{-1}|| = |a|||SAS^{-1}|| = |a|||A||_S$.

3. Triangle inequality: Let $A_1, A_2 \in M_n$. Then $||A_1 + A_2||_S = ||S(A_1 + A_2)S^{-1}|| = ||SA_1S^{-1} + SA_2S^{-1}|| \leq ||SA_1S^{-1}|| + ||SA_2S^{-1}|| = ||A_1||_S + ||A_2||_S$.

4. Submultiplicative property: $||A_1A_2||_S = ||SA_1A_2S^{-1}|| = ||SA_1S^{-1}SA_2S^{-1}|| \leq ||SA_1S^{-1}|| \cdot ||SA_2S^{-1}|| = ||A_1||_S \cdot ||A_2||_S$.

Exercise 37
We define $\mathcal{S} = \{1, x, x^2\}$ to be the basis of the space V. Then we can evaluate L on the basis vectors: $L[1] = 0, L[x] = 1, L[x^2] = 2$. Every function $p \in V$ can be written as a linear combination of the basis vectors:

$$p = a_1 + a_2x + a_3x^2 \implies L[p] = a_1L[1] + a_2L[x] + a_3L[x^2]$$

$$L[p] = \begin{bmatrix} L[1] & L[x] & L[x^2] \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \langle (0, 1, 2), x \rangle$$

Exercise 38
We again define $\mathcal{S} = \{1, x, x^2\}$ to be the basis of the space V, and we evaluate D on the basis vectors: $D[1](x) = 0, D[x](x) = 1, D[x^2](x) = 2x$. The matrix representation of D with respect to this basis is

$$D = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Using intergration by parts,

$$\langle q, D[p] \rangle = \int_{-\infty}^{\infty} q(x)p'(x)dx = -\int_{-\infty}^{\infty} q'(x)p(x)dx = -\langle D[q], [p] \rangle.$$

Therefore the matrix representation of the adjoint of D with respect to the basis is

$$D^* = -D = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Exercise 39

Let V and W be finite-dimensional inner product spaces. The adjoint has the following properties:

(i) If $S, T \in \mathscr{L}(V; W)$, then $(S+T)^* = S^* + T^*$ and $(\alpha T)^* = \bar{\alpha} T^*, \alpha \in \mathbb{F}$.

Proof:

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

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

(ii) If $S \in \mathscr{L}(V; W)$, then $(S^*)^* = S$.

Proof:

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

(iii) If $S, T \in \mathscr{L}(V)$, then $(ST)^* + T^* S^*$.

Proof:

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

(iv) If $T \in \mathscr{L}(V)$ and T is invertible, then $(T^*)^{-1} = (T^{-1})^*$.

Proof:

By (iii), $T^*(T^*)^{-1} = (TT^{-1})^* = I^* = I$.

Exercise 40

Let $M_n(\mathbb{F})$ be endowed with the Frobenius inner product. Any $A \in M_n(\mathbb{F})$ defines a linear operator on $M_n(\mathbb{F})$ by left multiplication: $B \mapsto AB$.

(i) Show that $A^* = A^H$.

Proof:

Let $B, C \in M_n$. Then the Frobenius inner product is $\langle B, AC \rangle = tr(B^H AC) = tr((A^H B)^H C) = \langle A^H B, C \rangle$.

(ii) Show that for any $A_1, A_2, A_3 \in M_n(\mathbb{F})$ we have $\langle A_2, A_3 A_1 \rangle = \langle A_2 A_1^*, A_3 \rangle$.

Proof:

$\langle A_2, A_3 A_1 \rangle = tr(A_2^H A_3 A_1) = tr(A_1 A_2^H A_3) = tr((A_2 A_1^H)^H A_3) = \langle A_2 A_1^H, A_3 \rangle = \langle A_2 A_1^*, A_3 \rangle$ by the result of (i).

Exercise 44
Given $A \in M_{m\times n}(\mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^m$, either $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x} \in \mathbb{F}^n$ or there exists $\mathbf{y} \in \mathscr{N}(A^H)$ such that $\langle \mathbf{y}, \mathbf{b}\rangle \neq 0$ (Fredholm alternative).
Proof:
First, we must show that if $A\mathbf{x} = \mathbf{b}$ has a solution, then $\forall \mathbf{y} \in \mathscr{N}(A^H)$ we have that $\langle \mathbf{y}, \mathbf{b}\rangle = 0$. We know that $\mathscr{N}(A^H)$ is orthogonal to $\mathscr{R}(A)$ so $\forall \mathbf{y} \in \mathscr{N}(A^H)$ and $\forall \mathbf{b} \in \mathscr{R}(A)$, we have that $\langle \mathbf{y}, \mathbf{b}\rangle = 0$ by definition.
On the other hand, we want to show that if $\exists \mathbf{y} \in \mathscr{N}(A^H)$ s.t. $\langle \mathbf{y}, \mathbf{b}\rangle \neq 0$, then $A\mathbf{x} = \mathbf{b}$ has no solution. This follows similarly to the proof of the first claim; if $\exists \mathbf{y} \in \mathscr{N}(A^H)$ s.t. $\langle \mathbf{y}, \mathbf{b}\rangle \neq 0$ then $b \notin \mathscr{R}(A)$, so $A\mathbf{x} = \mathbf{b}$ must have no solution.

Exercise 45
Consider the vector space $M_n(\mathbb{R})$ with the Frobenius inner product. Show that $Sym_n(\mathbb{R})^\perp = Skew_n(\mathbb{R})$.
Proof:

Exercise 46
Let A be an $m \times n$ matrix.

(i) If $\mathbf{x} = \mathscr{N}(A^H A)$, then $A\mathbf{x}$ is in both $\mathscr{R}(A)$ and $\mathscr{N}(A^H)$.
Proof:
By definition, $A\mathbf{x} \in \mathscr{R}(A)$. Additionally, $A^H A\mathbf{x} = \mathbf{0}$, so $A\mathbf{x} \in \mathscr{N}(A^H)$.

(ii) $\mathscr{N}(A^H A) = \mathscr{N}(A)$.
Proof:
Let $\mathbf{x} \in \mathscr{N}(A)$. Then $A\mathbf{x} = \mathbf{0} \implies A^H A\mathbf{x} = \mathbf{0}$. Now let $\mathbf{x} \in \mathscr{N}(A^H A)$. Then $A^H A\mathbf{x} = \mathbf{0}$. To show that $A\mathbf{x} = \mathbf{0}$, we take its norm $||A\mathbf{x}||^2 = \langle A\mathbf{x}, A\mathbf{x}\rangle = \mathbf{x}^H A^H A\mathbf{x} = 0$. Recall that this happens if and only if $A\mathbf{x} = \mathbf{0}$.

(iii) A and $A^H A$ have the same rank.
Proof:
Let $rank(A) = r$. Then we know that $rank(A^H) = r$. Since $rank(A^H A) \leq \min\{rank(A), rank(A^H)\} = r$. By part (ii), we know that $dim(\mathscr{N}(A^H A)) = dim(\mathscr{N}(A))$, so $rank(A) = rank(A^H)$.

(iv) If A has linearly independent columns, the $A^H A$ is nonsingular.
Proof:
If A has linearly independent columns, then $rank(A) = rank(A^H A) = n$. Since we know that $A^H A$ has full rank, it must be nonsingular.

Exercise 47
Let $P = A(A^H A)^{-1} A^H$
(i) $P^2 = P$

Proof:

$$
\begin{aligned}
P^2 &= A(A^H A)^{-1} A^H A (A^H A)^{-1} A^H \\
&= A I_n (A^H A)^{-1} A^H \\
&= A(A^H A)^{-1} A^H
\end{aligned}
$$

(ii) $P^H = P$
Proof:

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

(iii) $rank(P) = n$
Proof:
Since P is idempotent, so $tr(P) = tr(A(A^H A)^{-1} A^H) = tr((A^H A)^{-1} A^H A) = tr(I_n) = n = rank(P)$

Exercise 48
Consider the vector space $M_n(\mathbb{R})$ with the Frobenius inner product. Let $P(A) = \frac{A+A^T}{2}$ be the map $P : M_n(\mathbb{R}) \to M_n(\mathbb{R})$. Then

(i) P is linear.
Proof:
Let $A, B \in M_n(\mathbb{R}), \alpha, \beta \in \mathbb{R}$.

$$
\begin{aligned}
P(\alpha A + \beta B) &= \frac{(\alpha A + \beta B) + (\alpha A + \beta B)^T}{2} \\
&= \frac{\alpha A + \beta B + \alpha A^T + \beta B^T}{2} \\
&= \alpha \frac{A + A^T}{2} + \beta \frac{B + B^T}{2}
\end{aligned}
$$

(ii) $P^2 = P$
Proof:
Let $A \in M_n(\mathbb{R})$.
$P^2 = \frac{\frac{A+A^T}{2} + (\frac{A+A^T}{2})^T}{2} = \frac{A+A^T}{2} = P$

(iii) $P^* = P$
Proof:

Let $A, B \in M_n(\mathbb{R})$.

$$\begin{aligned}
\langle A, P(B)\rangle &= tr(A^T P(B)) \\
&= tr\Big(A^T \frac{B + B^T}{2}\Big) \\
&= tr\Big(A^T \frac{B}{2} + A^T \frac{B^T}{2}\Big) \\
&= tr\Big(\frac{A^T}{2} B + \frac{A^T}{2} B^T\Big) \\
&= tr\Big(\frac{A^T}{2} B\Big) + tr(\frac{A^T}{2} B^T\Big) \\
&= tr\Big(\frac{A}{2} B^T\Big) + tr(\frac{A^T}{2} B^T\Big) \\
&= tr\Big(\frac{A + A^T}{2} B^T\Big) \\
&= tr\Big(\Big(\frac{A + A^T}{2}\Big)^T B) \\
&= \langle P(A), B\rangle
\end{aligned}$$

(iv) $\mathscr{N}(P) = Skew_n(\mathbb{R})$
Proof:
First suppose that $A \in \mathscr{N}(P)$. Then we know that
$P(A) = \frac{A+A^T}{2} = 0 \implies A^T = -A$. Therefore $A \in Skew_n(\mathbb{R})$.
Now assume that $A \in Skew_n(\mathbb{R})$. Then $A^T = A \implies \frac{A+A^T}{2} = \frac{A-A}{2} = 0$. Therefore $A \in \mathscr{N}(P)$, and so $\mathscr{N}(P) = Skew_n(\mathbb{R})$.

(v) $\mathscr{R}(P) = Sym_n(\mathbb{R})$
Proof:
First suppose that $A \in \mathscr{R}(P)$. Then $\exists B \in M_n(\mathbb{R})$ s.t. $P(B) = \frac{B+B^T}{2} = A$. We know that $A^T = (\frac{B+B^T}{2})^T = \frac{B+B^T}{2} = A$, and therefore, $A \in Sym_n(\mathbb{R})$.
Now suppose that $A \in Sym_n(\mathbb{R})$. Then we know that $A = A^T$. Then $P(A) = \frac{A+A^T}{2} = \frac{2A}{2} = A, so A \in \mathscr{R}(P)$, and therefore $\mathscr{R}(P) = Sym_n(\mathbb{R})$.

(v) $||A - P(A)||_F = \sqrt{\frac{tr(A^T A) - tr(A^2)}{2}}$
Proof:

$$
\begin{aligned}
||A - P(A)||_F &= \sqrt{tr((A - P(A))^T(A - P(A)))} \\
&= \sqrt{tr((A^T - P(A))(A - P(A)))} \\
&= \sqrt{tr\Big(\Big(A^T - \frac{A + A^T}{2}\Big)\Big(A - \frac{A + A^T}{2}\Big)\Big)} \\
&= \sqrt{tr\Big(\Big(\frac{A^T - A}{2}\Big)\Big(\frac{A - A^T}{2}\Big)\Big)} \\
&= \sqrt{tr\Big(\Big(\frac{A^T A - A^2 + AA^T - (A^T)^2}{4}\Big)\Big)} \\
&= \sqrt{\frac{tr(A^T A) - tr(A^2)}{2}}
\end{aligned}
$$

Exercise 50

Let $(x_i, y_i)_{i=1}^n$ be a collection of data points lying roughly on an ellipse of the form $rx^2 + sy^2 = 1$. Find the least squares approximation for r and s. Write $A, \mathbf{x}$, and $\mathbf{b}$ for the corresponding normal equation in terms of the data x_i and y_i and the unknowns r and s.

Rewriting the ellipse equation, we get $rx^2 + sy^2 = 1 \implies \frac{1}{s} - \frac{r}{s}x^2 = y^2$. We can estimate this system using OLS on the following matrices:

$$
A = \begin{bmatrix} 1 & x_1^2 \\ 1 & x_2^2 \\ \vdots & \vdots \\ 1 & x_n^2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \frac{1}{s} \\ -\frac{r}{s} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}
$$

The cooresponding normal equation for the system is $A^H A\hat{\mathbf{x}} = A^H \mathbf{b}$.