Math Problem Set 2
Open Source Macroeconomics Laboratory Boot Camp

Ruby Zhang

July 3, 2017

3.1 (i)

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

(ii)

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

3.2

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

3.3 (i)

$$
\begin{aligned}
\cos\theta &= \frac{\langle x, x^5\rangle}{\|x\|\|x^5\|} = \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx \int_0^1 x^{10} dx}} = \frac{\frac{1}{7}}{\sqrt{\frac{1}{3}\frac{1}{11}}} \\
\therefore \theta &= \arccos\frac{\sqrt{33}}{7} = 0.608
\end{aligned}
$$

(ii)

$$
\begin{aligned}
\cos\theta &= \frac{\langle x^2, x^4\rangle}{\|x^2\|\|x^4\|} = \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^4 dx \int_0^1 x^8 dx}} = \frac{\frac{1}{7}}{\sqrt{\frac{1}{5}\frac{1}{9}}} \\
\therefore \theta &= \arccos\frac{\sqrt{45}}{7} = 0.29
\end{aligned}
$$

3.8 (i) We must prove that $S = \{\cos(t), \sin(t), \cos(2t), \sin(2t)\}$ is an orthnormal set. First let us prove that all pairs of the basis are orthogonal:

$$\begin{aligned}
\langle \cos(t), \sin(t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\cos(t)dt = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t) d\sin(t) \\
&= \frac{1}{2\pi}\sin^2(t)]_{-\pi}^{\pi} = 0 \\
\langle \cos(t), \cos(2t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(2t)\cos(t)dt = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\cos(2t+t)+\cos(2t-t)}{2}dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(3t)+\cos(t)dt = \frac{1}{2\pi}[\frac{1}{3}\sin(3t)+\sin(t)]_{-\pi}^{\pi} = 0 \\
\langle \cos(t), \sin(2t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(2t)\cos(t)dt = \frac{2}{\pi}\int_{-\pi}^{\pi} \sin(t)\cos^2(t)dt \\
&= \frac{-2}{\pi}\int_{-\pi}^{\pi} \cos^2(t) d\cos(t) = \frac{-2}{3\pi}[\cos^3(t)]_{-\pi}^{\pi} = 0 \\
\langle \cos(2t), \sin(t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\cos(2t)dt = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin(2t+t)-\sin(2t-t)}{2}dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \sin(3t)-\sin(t)dt = \frac{1}{2\pi}[-\frac{1}{3}\cos(3t)+\cos(t)]_{-\pi}^{\pi} = 0 \\
\langle \sin(2t), \sin(t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\sin(2t)dt = \frac{2}{\pi}\int_{-\pi}^{\pi} \sin^2(t)\cos(t)dt \\
&= \frac{2}{\pi}\int_{-\pi}^{\pi} \sin^2(t) d\sin(t) = \frac{2}{3\pi}[\sin^3(t)]_{-\pi}^{\pi} = 0 \\
\langle \sin(2t), \cos(2t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(2t)\sin(2t)dt = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(2t) d\cos(2t) \\
&= -\frac{1}{4\pi}[\cos^2(2t)]_{-\pi}^{\pi} = 0
\end{aligned}$$

Now we will prove that the norm of each basis element is equal to 1:

$$
\begin{aligned}
\langle \cos(t), \cos(t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2(t)dt = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1+cos(2t)}{2}dt \\
&= \frac{1}{2\pi}[t + \frac{1}{2}\sin(2t)]_{-\pi}^{\pi} = \frac{1}{2\pi}2\pi = 1 \\
\langle \sin(t), \sin(t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin^2(t)dt = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1-cos(2t)}{2}dt \\
&= \frac{1}{2\pi}[t - \frac{1}{2}\sin(2t)]_{-\pi}^{\pi} = \frac{1}{2\pi}2\pi = 1 \\
\langle \cos(2t), \cos(2t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2(2t)dt = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1+cos(4t)}{2}dt \\
&= \frac{1}{2\pi}[t + \frac{1}{4}\sin(4t)]_{-\pi}^{\pi} = \frac{1}{2\pi}2\pi = 1 \\
\langle \sin(2t), \sin(2t) \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin^2(2t)dt = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1-cos(4t)}{2}dt \\
&= \frac{1}{2\pi}[t - \frac{1}{4}\sin(4t)]_{-\pi}^{\pi} = \frac{1}{2\pi}2\pi = 1
\end{aligned}
$$

(ii)

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

(iii)

$$
\begin{aligned}
\text{proj}_X(\cos(3t)) &= \langle \cos(3t), \cos(t) \rangle \cos(t) + \langle \cos(3t), \sin(t) \rangle \sin(t) \\
&+ \langle \cos(3t), \cos(2t) \rangle \cos(2t) + \langle \cos(3t), \sin(2t) \rangle \sin(2t) \\
&= \frac{1}{\pi}(\cos(t)\int_{-\pi}^{\pi} \cos(3t)\cos(t)dt + \sin(t)\int_{-\pi}^{\pi} \cos(3t)\sin(t)dt \\
&+ \cos(2t)\int_{-\pi}^{\pi} \cos(3t)\cos(2t)dt + \sin(2t)\int_{-\pi}^{\pi} \cos(3t)\sin(2t)dt) \\
&= 0
\end{aligned}
$$

(iv)

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

3.9 Let $L_\theta : \mathbf{R}^2 \to \mathbf{R}^2$ be the rotation transformation around the origina counterclockwise by angle θ:

$$L_\theta(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$$

Now let us calculate the inner product after the transformation for vectors $\mathbf{u} = (x_1, y_1)$ and $\mathbf{v} = (x_2, y_2)$:

$$
\begin{aligned}
\langle L_\theta\mathbf{u}, L_\theta\mathbf{v}\rangle &= (x_1\cos\theta - y_1\sin\theta)(x_2\cos\theta - y_2\sin\theta) + (x_1\sin\theta + y_1\cos\theta)(x_2\sin\theta + y_2\cos\theta) \\
&= x_1x_2(\cos^2\theta + \sin^2\theta) + y_1y_2(\sin^2\theta + \cos^2\theta) \\
&- (x_1y_2 + x_2y_1)\cos\theta\sin\theta + (x_1y_2 + x_2y_1)\cos\theta\sin\theta \\
&= x_1x_2 + y_1y_2 \\
&= \langle\mathbf{u}, \mathbf{v}\rangle
\end{aligned}
$$

Therefore, the rotation transformation is an orthonormal transformation.

3.10 (i) $\Longrightarrow$

$$\langle Qx, Qy\rangle = (Qx)^H(Qy) = (x^HQ^H)Qy = x^H(Q^HQy) = \langle x, Q^HQy\rangle$$

Since Q is an orthonormal transformation, we know that $\langle x, Q^HQy\rangle = \langle x, y\rangle = \langle x, Iy\rangle$. Therefore, $Q^HQ = I$.

$\Longleftarrow$

We know that $Q^HQ = I$, thus:

$$\langle Qx, Qy\rangle = (Qx)^H(Qy) = (x^HQ^H)Qy = x^H(Q^HQ)y = x^Hy = \langle x, y\rangle$$

(ii) By the orthonormality of matrix Q, we have that

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

Since the norm is nonnegative, then $\|Qx\| = \|x\|$.

(iii) Since Q is orthonormal, $QQ^H = I$ so $Q^H = Q^{-1}$. Note that $I = QQ^H = (Q^H)^HQ^H = (Q^{-1})^HQ^{-1}$. By (i), Q^{-1} is also orthnormal.

(iv) Suppose we have the orthonormal matrix $Q = [a_1, ..., a_n]$ where a_i's are column vectors. Then the element in the ith row and jth column of matrix $Q^H Q = I$ is given by the entry $a_i^H a_j = \delta_{ij}$. Therefore, $a_i^H a_j = \langle a_i, a_j \rangle = 0$ for all $i \neq j$ and 0 otherwise, which makes $\{a_i\}$ a collection of orthnormal vectors.

(v) If Q is orthogonal, then $\det^2(Q) = \det(Q)\det(Q^H) = \det(QQ^H) = \det(I) = 1$. Therefore, $\sqrt{\det^2(Q)} = |\det(Q)| = 1$. However, the converse is not true. Take the following matrix:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

Note that $\det(A) = 2 - 1 = 1$ but the columns are not orthonormal.

(vi)

$$(Q_1Q_2)(Q_1Q_2)^H = (Q_1Q_2)(Q_2^H Q_1^H) = Q_1(Q_2Q_2^H)Q_1^H) = Q_1Q_1^H = I$$

According to (i), Q_1Q_2 is orthnormal.

3.11 Suppose we have the vectors $\{v_1, ..., v_n\}$ where vector v_n is linearly dependent on the vectors $\{v_1, ..., v_{n-1}\}$, which are linearly independent. Then the process would output the 0 vector for vector q_n since $p_{n-1} = v_k$ as the projection of the vector onto the space is itself since it is linearly dependent with the basis of the space.

3.16 (i) For any diagnal matrix D and QR decomposition, we have that $QR = QIR = QDD^{-1}R = (QD)(D^{-1}R)$. Note that all diagonal matrices (and its inverse) are both orthonormal $D = D^{-1} = D^H$ and upper triangular. Since orthonormal and triangular matrices are closed under multiplication, then $Q'R'$, where $Q' = QD$ and $R' = D^{-1}R$, is another QR decomposition of the same matrix.

(ii) Suppose there exist two QR decompositions of A: $Q_1R_1 = A = Q_2R_2$. Then we have that $B = Q_2^H Q^1 = R_2R_1^{-1}$. Since orthnormal and upper triangular matrices are closed under multiplication and inverses, then B is an orthonormal, upper triangular matrix. Then B must be a diagonal matrix since $B_H = B_{-1}$ which would not hold if there were nonzero, nondiagonal entries as that would produce a lower triangular matrix which violate the closing of upper triangular matrices under the inverse. In addition, since the columns of B must be orthonormal, then all the diagonal entries are ± 1. But we know that R_1, R_2 have positive diagonal entries so $B = I$. Thus, $I = Q_2^H Q^1 = R_2R_1^{-1}$ so we have that $Q_2 = Q_1$ and $R_1 = R_2$.

3.17

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

3.23 Note that using the triangle inequality property of the norm, we have that:

$$\begin{aligned}
\|y\| &= \|x + (y-x)\| \le \|x\| + \|y-x\| \\
\therefore \|y\| - \|x\| &\le \|y-x\| = \|-1\|\|x-y\| = \|x-y\| \\
\|x\| &= \|y + (x-y)\| \le \|y\| + \|x-y\| \\
\therefore \|x\| - \|y\| &\le \|x-y\|
\end{aligned}$$

$$\therefore |\|x\| - \|y\|| \le \|x-y\|$$

3.24 We must prove that each of the following satisfies positivity (and equality), scale preservation, and triangle inequality:

(i) 1. Since $|f(t)|$ is a nonnegative function, its integral is also nonnegative. If $\|f\|_{L^1} = \int_a^b |f(t)|dt = 0$, then $f(t) = 0$ on $[a,b]$ since f is continuous on $[a,b]$.

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

3.

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

(ii) 1. Since $|f(t)|^2$ is a nonnegative function, its integral and the square root of it are also non-negative. If $\|f\|_{L^2} = (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} = 0$, then $f(t) = 0$ on $[a,b]$ since f is continuous on $[a,b]$.

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

3.

$$\begin{aligned}
\|f+g\|_{L^2}^2 &= \int_a^b |f(t)+g(t)|^2 dt) = \int_a^b |f(t)|^2 dt + \int_a^b |g(t)|^2 dt + 2\int_a^b f(t)g(t)dt \\
&\le \int_a^b |f(t)|^2 dt + \int_a^b |g(t)|^2 dt + 2\sqrt{\int_a^b |f(t)g(t)|^2 dt} \text{ by the Schwarz inequality} \\
&= \left(\sqrt{\int_a^b |f(t)|^2 dt} + \sqrt{\int_a^b |g(t)|^2 dt}\right)^2 \\
\therefore \|f+g\|_{L^2} &\le \sqrt{\int_a^b |f(t)|^2 dt} + \sqrt{\int_a^b |g(t)|^2 dt} = \|f\|_{L^2} + \|g\|_{L^2}
\end{aligned}$$

(iii) 1. Since $|f(t)|$ is a nonnegative function, its supremum is also nonnegative. If $\sup_{x\in[a,b]} |f(x)| = 0$, then $f(t) = 0$ on $[a,b]$ since any non-zero value of $f(t)$ on $[a,b]$ would contradict $\sup_{x\in[a,b]} |f(x)| = 0$.

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

3.

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

3.26 First, we will prove that topological equivalence (denote by $\sim$) is an equivalence relation:

1. Reflexivity: We have that $0.5\|x\|_a \leq \|x\|_a \leq 2\|x\|_a$ for all $x \in X$. Thus, $\|x\|_a \sim \|x\|_a$.
2. Symmetry: Suppose $\|x\|_a \sim \|x\|_b$. Then there exists constants $0 < m \leq M$ such that $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a$. We then have that $\frac{1}{M}\|x\|_b \leq \|x\|_a \leq \frac{1}{m}\|x\|_b$ where $0 < \frac{1}{M} \leq \frac{1}{m}$. Thus, $\|x\|_b \sim \|x\|_a$ as well.
3. Transitivity: Suppose $\|x\|_a \sim \|x\|_b$ and $\|x\|_b \sim \|x\|_c$. Then there exist constants $0 < m_1 \leq M_1$ and $0 < m_2 \leq M_2$ such that $m_1\|x\|_a \leq \|x\|_b \leq M_1\|x\|_a$ and $m_2\|x\|_b \leq \|x\|_c \leq M_2\|x\|_b$. Then we have that $m_1 m_2\|x\|_a \leq \|x\|_c \leq M_1 M_2\|x\|_a$ where $0 < m_1 m_2 \leq M_1 M_2$. Therefore, $\|x\|_a \sim \|x\|_c$.

Now we will show that the p-norms for $p = 1, 2, \infty$ are topologically equivalent by establishing the following inequalities:

(i) First, note that we have:

$$\begin{aligned}\Big(\sum_{i=1}^n |x_i|\Big)^2 &= \sum_{i=1}^n |x_i|^2 + \sum_{i\neq j}^n |x_i||x_j| \\ &\geq \sum_{i=1}^n |x_i|^2 \text{ since all terms are nonnegative} \\ \sum_{i=1}^n |x_i| &\geq \sqrt{\sum_{i=1}^n |x_i|^2} \\ \therefore \|x\|_2 &\leq \|x\|_1\end{aligned}$$

From the Cauchy-Schwarz inequality, we have that $\left|\sum_{i=1}^n x_i y_i\right|^2 \leq \sum_{j=1}^n |x_j|^2 \sum_{k=1}^n |y_j|^2$. Letting $y_i = 1$ for all $1 \leq i \leq n$, we have that

$$\begin{aligned}\Big(\sum_{i=1}^n 1|x_i|\Big)^2 &\leq \sum_{i=1}^n |x_i|^2 \sum_{i=1}^n 1 = n\sum_{i=1}^n |x_i|^2 \\ \sum_{i=1}^n 1|x_i| &\leq \sqrt{n\sum_{i=1}^n |x_i|^2} \\ \therefore \|x\|_1 &\leq \sqrt{n}\|x\|_2\end{aligned}$$

Therefore, $\|x\|_1 \sim \|x\|_2$ since $\|x\|_2 \leq \|x\|_1 \leq \|x\|_2$.

(ii) Note that $\|x\|_\infty = \max_i |x_i|$. Suppose $\max_i |x_i| = |x_j|$. Therefore, $|x_j|^2 \leq |x_j|^2 + \sum_{i\neq j}^n |x_i|^2 \leq \sum_{i=1}^n |x_j|^2 = n|x_j|^2$. Then we have that $(\max_i |x_i|)^2 \leq \sum_{=1}^n |x_i|^2 \leq n(\max_i |x_i|)^2$. Therefore, $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$ so $\|x\|_2 \sim \|x\|_\infty$.

Since $\sim$ is an equivalence relation, transitivity holds so $\|x\|_1 \sim \|x\|_\infty$.

3.28 (i) From question 3.26(i), we have that $\frac{1}{\|x\|_2} \leq \frac{\sqrt{n}}{\|x\|_1}$ and $\frac{1}{\|x\|_1} \leq \frac{1}{\|x\|_2}$ for all x. Also using the direct inequality proved in 3.26(i), we have that:

$$\begin{aligned}\frac{1}{\sqrt{n}}\|A\|_2 &= \frac{1}{\sqrt{n}}\sup_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} \leq \frac{1}{\sqrt{n}}\sup_{x\neq 0}\sqrt{n}\frac{\|Ax\|_1}{\|x\|_1} \\ &= \|A\|_1 = \sup_{x\neq 0}\frac{\|Ax\|_1}{\|x\|_1} \leq \sup_{x\neq 0}\sqrt{n}\frac{\|Ax\|_2}{\|x\|_2} \\ &= \sqrt{n}\|A\|_2 \\ \therefore \frac{1}{\sqrt{n}}\|A\|_2 &\leq \|A\|_1 \leq \sqrt{n}\|A\|_2\end{aligned}$$

(ii) Similarly, using 3.26(ii), we have that $\frac{1}{\|x\|_\infty} \leq \frac{\sqrt{n}}{\|x\|_2}$ and $\frac{1}{\|x\|_2} \leq \frac{1}{\|x\|_\infty}$ for all x. Also using the direct inequality proved in 3.26(ii), we have that:

$$\begin{aligned}\frac{1}{\sqrt{n}}\|A\|_\infty &= \frac{1}{\sqrt{n}}\sup_{x\neq 0}\frac{\|Ax\|_\infty}{\|x\|_\infty} \leq \frac{1}{\sqrt{n}}\sup_{x\neq 0}\sqrt{n}\frac{\|Ax\|_2}{\|x\|_2} \\ &= \|A\|_2 = \sup_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} \leq \sup_{x\neq 0}\sqrt{n}\frac{\|Ax\|_\infty}{\|x\|_\infty} \\ &= \sqrt{n}\|A\|_\infty \\ \therefore \frac{1}{\sqrt{n}}\|A\|_\infty &\leq \|A\|_2 \leq \sqrt{n}\|A\|_\infty\end{aligned}$$

3.29 From exercise 3.10(ii), we proved that $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{F}^n$ and orthnormal matrix Q. Therefore, $\|Q\| = \sup_{x\neq 0}\frac{\|Q\mathbf{x}\|}{\|\mathbf{x}\|} = 1$. The induced norm of $R_\mathbf{x}$ is given by: $\|R_\mathbf{x}\|_2 = \sup_{A\neq 0}\frac{\|R_\mathbf{x}A\|_2}{\|A\|_2} = \sup_{A\neq 0}\frac{\|A\mathbf{x}\|_2}{\|A\|_2}$. Since $\|A\|_2 = \sup_{\mathbf{y}\neq 0}\frac{\|A\mathbf{y}\|_2}{\|\mathbf{y}\|_2} \geq \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$, then $\|\mathbf{x}\|_2 \geq \frac{\|A\mathbf{x}\|_2}{\|A\|_2}$ for all matrices A so $\|\mathbf{x}\|_2 \geq \|R_\mathbf{x}\|_2$. Note that equality is possible when A is orthonormal since $\|A\mathbf{x}\|_2 = \|\mathbf{x}\|_2 = \|A\|_2\|\mathbf{x}\|_2$, thus $\frac{\|A\mathbf{x}\|_2}{\|A\|_2} = \|\mathbf{x}\|_2$. Therefore, $\|R_\mathbf{x}\|_2 = \sup_{A\neq 0}\frac{\|R_\mathbf{x}A\|_2}{\|A\|_2} = \sup_{A\neq 0}\frac{\|A\mathbf{x}\|_2}{\|A\|_2} = \|\mathbf{x}\|_2$.

3.30 We will first show that $\|\dot{|}|_S$ satisfies the properties of a norm, then that it satisfies the submultiplicative property of the matrix norm.

1. Positivity is obviously satisfied since $\|\cdot\|$ is a matrix norm.
2. Scale preservation: $\|\alpha A\|_S = \|S(\alpha A)S^{-1}\| = |\alpha|\|SAS^{-1}\| = |\alpha|\|A\|_S$.
3. Triangle inequality: $\|A + B\|_S = \|S(A + B)S^{-1}\| = \|SAS^{-1} + SBS^{-1}\| \leq \|A\|_S + \|B\|_S$ since matrix multiplication obeys distributive properties from both left and right.

Since $\|\cdot\|$ is a matrix norm with the submultiplicative property, we have that:

$$\|AB\|_S = \|S(AB)S^{-1}\| = \|(SAS^{-1})(SBS^{-1})\| \leq \|SAS^{-1}\|\|SBS^{-1}\| = \|A\|_S\|B\|_S$$

3.37 Let $q = 180x^2 - 168x + 24$. Then for any $p = ax^2 + bx + c$, we have that

$$\begin{aligned}
\langle q, p\rangle &= \int_0^1 qpdx = \int_0^1 (180x^2 - 168x + 24)(ax^2 + bx + c)dx \\
&= \int_0^1 180ax^4 + 180bx^3 + 180cx^2 - 168ax^3 - 168bx^2 - 168cx + 24ax^2 + 24bx + 24c \\
&= \frac{180a}{5} + \frac{180b}{4} + \frac{180c}{3} - \frac{168a}{4} - \frac{168b}{3} - \frac{168c}{2} + \frac{24a}{3} + \frac{24b}{2} + 24c \\
&= (36 - 42 + 8)a + (45 - 56 + 12)b + (60 - 84 + 24)c \\
&= 2a + b = L[p]
\end{aligned}$$

3.38 Letting the basis be $[1, x, x^2]$, then the coordinates of the basis are $[1, 0, 0], [0, 1, 0], [0, 0, 1]$. Thus $p(x) = a + bx + cx^2$ can be represented as $[a, b, c]$. Then the differentiation matrix is the following:

$$D[p](x) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b \\ 2c \\ 0 \end{bmatrix} = b + 2cx = p'(x)$$

The adjoint of D is the map D^* such that (using integration by parts):

$$\begin{aligned}
\langle f, D^*g\rangle &= \langle Df, g\rangle \\
\int_0^1 f(x)D^*[g](x)dx &= \int_0^1 D[f](x)g(x)dx \\
&= \Big[f(x)g(x)\Big]_0^1 - \int_0^1 f(x)g'(x)dx
\end{aligned}$$

Restricting to polynomials with $f(0) = f(1)$, then we have that $D^* = -D$, which gives us the following adjoint matrix:

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

3.39 (i)

$$\begin{aligned}
\langle (S + \alpha T)^*(\mathbf{w}), \mathbf{v}\rangle_V &= \langle \mathbf{w}, (S + \alpha T)(\mathbf{v})\rangle_W \\
&= \langle \mathbf{w}, S\mathbf{v} + \alpha T\mathbf{v}\rangle_W \\
&= \langle \mathbf{w}, S\mathbf{v}\rangle_W + \langle \mathbf{w}, \alpha T\mathbf{v}\rangle_W \\
&= \langle S^*\mathbf{w}, \mathbf{v}\rangle_V + \langle (T^*(\mathbf{w}), \alpha\mathbf{v}\rangle_V \\
&= \langle S^*\mathbf{w}, \mathbf{v}\rangle_V + \langle (\bar{\alpha}T^*(\mathbf{w}), \mathbf{v}\rangle_V \\
&= \langle (S^* + \bar{\alpha}T^*)(\mathbf{w}), \mathbf{v}\rangle_V \\
\therefore (S + \alpha T)^* &= S^* + (\alpha T)^* = S^* + \bar{\alpha}T^*
\end{aligned}$$

(ii)

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

Therefore, $(S^*)^* = S$.

(iii)

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

(iv) Since the identity matrix is its own adjoint, we have that

$$(T^*)(T^*)^{-1} = I = I^* \implies (T^*)(T^*)^{-1} = (T^{-1}T)^* = T^*(T^{-1})^* \implies (T^*)^{-1} = (T^{-1})^*$$

3.40 (i) For matrices B, C and linear operator (also a matrix) A, we have that:

$$\langle A^*B, C\rangle = \langle B, AC\rangle = \mathrm{tr}(B^HAC) = \mathrm{tr}((A^HB)^HC) = \langle A^HB, C\rangle$$

Therefore, $A^* = A^H$.

(ii)

$$\langle A_2A_1^*, A_3\rangle = \langle A_2A_1^H, A_3\rangle = \mathrm{tr}(A_1A_2^HA_3) = \mathrm{tr}(A_1(A_2^HA_3)) = \mathrm{tr}(A_2^HA_3A_1) = \langle A_2, A_3A_1\rangle$$

(iii) Using the fact that trace is a linear mapping (i.e. $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$), we have that

$$\begin{aligned}
\langle T_A^*X, Y\rangle &= \langle X, T_AY\rangle = \langle X, AY - YA\rangle = \mathrm{tr}(X^HAY - X^HYA) \\
&= \mathrm{tr}(X^HAY) - \mathrm{tr}(X^HYA) = \mathrm{tr}(X^HAY) - \mathrm{tr}(AX^HY) \\
&= \langle A^HX, Y\rangle - \langle XA^H, Y\rangle = \langle A^*X - XA^*, Y\rangle \\
&= \langle T_{A^*}X, Y\rangle \\
\therefore T_A^* &= T_{A^*}
\end{aligned}$$

3.44 From 3.40(i) and the fundamental subspaces theorem (3.8.9), we know that $\mathscr{R}(A)^\perp = \mathscr{N}(A^H)$. If $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x} \in \mathbb{F}^n$, then $\mathbf{b} \in \mathscr{R}(A)$, in which case by definition $\langle \mathbf{b}, \mathbf{y}\rangle = 0$ for all $\mathbf{y} \in \mathscr{R}(A)^\perp = \mathscr{N}(A^H)$. On the other hand, if there exists $\mathbf{y} \in \mathscr{N}(A^H) = \mathscr{R}(A)^\perp$ such that $\langle \mathbf{y}, \mathbf{b}\rangle \neq 0$, then by definition $\mathbf{b} \notin \mathscr{R}(A)$ since $\mathbf{b} \neq \mathbf{0}$ so there does not exist a solution to $A\mathbf{x} = \mathbf{b}$.

3.45 From exercise 1.18, we know that:

$$\begin{aligned}
\mathrm{Sym}_n(\mathbb{R}) &= \{A \in M_n(\mathbb{R}) | A^T = A\} \\
\mathrm{Skew}_n(\mathbb{R}) &= \{A \in M_n(\mathbb{R}) | A^T = -A\}
\end{aligned}$$

Note that $A^H = A^T$ and $\langle A, B\rangle = \langle B, A\rangle$ for all $A, B \in M_n(\mathbb{R})$ since the conjugate of a real number is itself. Then for all $A \in \mathrm{Sym}_n(\mathbb{R})$ and $B \in \mathrm{Skew}_n(\mathbb{R})$:

$$\begin{aligned}
&\langle B, A\rangle = \mathrm{tr}(B^T A) = -\mathrm{tr}(BA) = -\mathrm{tr}(BA^T) = -\mathrm{tr}(A^T B) = -\langle A, B\rangle = -\langle B, A\rangle \\
&\therefore \langle B, A\rangle = 0
\end{aligned}$$

Then by definition, $\mathrm{Skew}_n(\mathbb{R}) \subset \mathrm{Sym}_n(\mathbb{R})^\perp$.

Now we will show the other way. Let $E_{ij} = e_i e_j^T$, (the zero matrix except for 1 in position (i,j)). Let $A = E_ij + E_ji \in \mathrm{Sym}_n(\mathbb{R})$. If $C \in \mathrm{Sym}_n(\mathbb{R})^\perp$, then $\langle C, E_{ij}\rangle + \langle C, E_{ji}\rangle = 0 \Longrightarrow [C]_{ij} + [C]_{ji} = 0$ for all i, j. Therefore, $C^T = -C$ so $C \in \mathrm{Skew}_n(\mathbb{R})$.

Therefore, $\mathrm{Skew}_n(\mathbb{R}) = \mathrm{Sym}_n(\mathbb{R})^\perp$.

3.46 (i) If $\mathbf{x} \in \mathscr{N}(A^H A)$, then $A^H(A\mathbf{x}) = \mathbf{0}$. Therefore, $A\mathbf{x} \in \mathscr{N}(A^H)$ and obviously $\mathscr{R}(A)$ by definition.

(ii) If $\mathbf{x} \in \mathscr{N}(A)$, then $A^H(A\mathbf{x}) = A^H\mathbf{0} = \mathbf{0}$ so $\mathbf{x} \in \mathscr{N}(A^H A)$. If $\mathbf{x} \in \mathscr{N}(A)$, then by previous part and the fundamental subspaces theorem, $A\mathbf{x} \in \mathscr{N}(A^H) = \mathscr{R}(A)^\perp$ and $\mathscr{R}(A)$. Thus, $\langle A\mathbf{x}, A\mathbf{x}\rangle = \mathbf{0}$ so $A\mathbf{x} = 0 \Longrightarrow \mathbf{x} \in \mathscr{N}(A)$.

(iii) By the rank nullity theorem, $\mathrm{rank}(A) = n - \dim(\mathscr{N}(A)) = n - \dim(\mathscr{N}(A^H A)) = \mathrm{rank}(A^H A)$.

(iv) If A has linearly independent columns, then it is injective and has rank n. Thus, $n = \mathrm{rank}(A) = \mathrm{rank}(A^H A)$ so $A^H A$ is an $n \times n$ matrix with full rank of n, making it nonsingular and bijective.

3.47 (i)

$$P^2 = (A(A^HA)^{-1}A^H)(A(A^HA)^{-1}A^H) = A(A^HA)^{-1}(A^HA)(A^HA)^{-1}A^H = A(A^HA)^{-1}A^H = P$$

(ii) Using the fact that $(B^H)^{-1} = (B^{-1})^H$, we have that

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

(iii) Let $\mathbf{x} \in \mathscr{N}(P)$. Then we have that $P\mathbf{x} = A((A^HA)^{-1}A^H\mathbf{x}) = 0 \Longrightarrow (A^HA)^{-1}A^H\mathbf{x} \in \mathscr{N}(A)$. Since $\mathrm{rank}(A) = n$, then $\dim(\mathscr{N}(A)) = 0$ so $(A^HA)^{-1}(A^H\mathbf{x}) = 0$. Similarly $A^H\mathbf{x} \in \mathscr{N}((A^HA)^{-1})$. From question 3.46(iv), we know that A^HA is nonsingular so $\mathrm{rank}((A^HA)^{-1}) = n$ and $\dim(\mathscr{N}((A^HA)^{-1})) = 0 \Longrightarrow A^H\mathbf{x} = 0$. Similarly, $\mathbf{x} \in \mathscr{N}(A^H)$ and $\dim(\mathscr{N}(A^H)) = \dim(\mathscr{R}(A)^\perp) = 0$ so $\mathbf{x} = 0$. Thus, $\dim(\mathscr{N}(P)) = 0$ and $\mathrm{rank}(P) = n$ by rank nullity theorem.

3.48 (i) For any $A, B \in M_n(\mathbb{R})$ and $\alpha \in \mathbb{R}$, we have that

$$P(A + \alpha B) = \frac{(A + \alpha B) + (A + \alpha B)^T}{2} = \frac{A + A^T}{2} + \alpha\frac{B + B^T}{2} = P(A) + \alpha P(B)$$

which follows from the fact that the transpose is linear since the (i, j) entry of $A^T + B^T$ is the sum of the (i, j) entries of A^T and B^T, which is the sum of the (j, i) entries of A and B, which is equal to (i, j) entry of $(A + B)^T$.

(ii)

$$
\begin{aligned}
P^2(A) &= P(\frac{P(A)+P(A)^T}{2}) = \frac{1}{2}P(\frac{A+A^T}{2}+\frac{A^T+A}{2}) \\
&= \frac{1}{2}P(A+A^T) = \frac{1}{2}\frac{A+A^T+A^T+A}{2} = \frac{A+A^T}{2} = P(A)
\end{aligned}
$$

(iii)

$$
\begin{aligned}
\langle P(A), B\rangle &= \operatorname{tr}((\frac{A+A^T}{2})^T B) = \operatorname{tr}(\frac{AB}{2}) + \operatorname{tr}(\frac{A^TB}{2}) = \operatorname{tr}(\frac{A^TB^T}{2}) + \operatorname{tr}(\frac{A^TB}{2}) \\
&= \operatorname{tr}(A^T\frac{B^T+B}{2}) = \langle A, P(B)\rangle \\
\therefore P^* &= P
\end{aligned}
$$

(iv) Note that $P(A) = \frac{A^T+A}{2} = 0$ if and only if $A^T = -A$, in other words if $A \in \mathrm{Skew}_n(\mathbb{R})$.

(v) If $B \in \mathscr{R}(P)$, then there exists matrix A such that $B = \frac{A+A^T}{2} \implies B^T = (\frac{A+A^T}{2})^T = \frac{A+A^T}{2} = B$. Therefore, $B \in \mathrm{Sym}_n(\mathbb{R})$. If $B \in \mathrm{Sym}_n(\mathbb{R})$, then $B = B^T$ so $P(B) = \frac{B+B^T}{2} = \frac{2B}{2} = B$ so $B \in \mathscr{R}(P)$.

(vi)

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

3.50 We want to solve the normal equation $A^HA\mathbf{x} = A^H\mathbf{b}$ with the following $A, \mathbf{x}, \mathbf{b}$:

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