Problem Set #2

Inner Product Space
Zeshun Zong

Exercise 1

$$\text{RHS} = \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) = \frac{1}{4}(\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle)$$
$$= \frac{1}{4}(\langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle - \langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle - \langle \mathbf{y}, \mathbf{y}\rangle)$$
$$= \frac{1}{4}(4\langle \mathbf{x}, \mathbf{y}\rangle) = \langle \mathbf{x}, \mathbf{y}\rangle = \text{LHS}$$

$$\text{RHS} = \frac{1}{2}(\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle + \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle)$$
$$= \frac{1}{2}(\langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{x}, \mathbf{y}\rangle - \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle)$$
$$= \frac{1}{2}(2\langle \mathbf{x}, \mathbf{x}\rangle + 2\langle \mathbf{y}, \mathbf{y}\rangle) = \langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle$$
$$= \|x\|^2 + \|y\|^2 = \text{LHS}$$

Exercise 2

$$\text{RHS} = \frac{1}{4}[\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle + i\langle \mathbf{x}-i\mathbf{y}, \mathbf{x}-i\mathbf{y}\rangle - i\langle \mathbf{x}+i\mathbf{y}, \mathbf{x}+i\mathbf{y}\rangle]$$
$$= \frac{1}{4}[\langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle - \langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle$$
$$- \langle \mathbf{y}, \mathbf{y}\rangle + i\langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle - \langle \mathbf{y}, \mathbf{x}\rangle - i\langle \mathbf{y}, \mathbf{y}\rangle - i\langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle$$
$$- \langle \mathbf{y}, \mathbf{x}\rangle + i\langle \mathbf{y}, \mathbf{y}\rangle]$$
$$= \frac{1}{4}(4\langle \mathbf{x}, \mathbf{y}\rangle) = \langle \mathbf{x}, \mathbf{y}\rangle = \text{LHS}$$

Exercise 3

We need the following computation:
$\langle x, x^5\rangle = \int_0^1 x^6 dx = \frac{1}{7}$
$\sqrt{\langle x, x\rangle} = \sqrt{\int_0^1 x^2 dx} = \sqrt{\frac{1}{3}}$
$\sqrt{\langle x^5, x^5\rangle} = \sqrt{\int_0^1 x^10 dx} = \sqrt{\frac{1}{11}}$
$\langle x^2, x^4\rangle = \int_0^1 x^6 dx = \frac{1}{7}$
$\sqrt{\langle x^2, x^2\rangle} = \sqrt{\int_0^1 x^4 dx} = \sqrt{\frac{1}{5}}$
$\sqrt{\langle x^4, x^4\rangle} = \sqrt{\int_0^1 x^8 dx} = \sqrt{\frac{1}{9}}$
We now have the following:

1.

$$\theta_1 = \arccos(\frac{\langle x, x^5 \rangle}{\|x\| \, \|x^5\|}) = \arccos(\frac{\frac{1}{7}}{\sqrt{\frac{1}{3}}\sqrt{\frac{1}{11}}}) = \arccos(\frac{\sqrt{33}}{7})$$

2.

$$\theta_2 = \arccos(\frac{\langle x^2, x^4 \rangle}{\|x^2\| \, \|x^4\|}) = \arccos(\frac{\frac{1}{7}}{\sqrt{\frac{1}{5}}\sqrt{\frac{1}{9}}}) = \arccos(\frac{\sqrt{45}}{7})$$

Exercise 8

1. Observe that $\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\sin(t)dt = 0$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\cos(2t)dt = 0$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\sin(2t)dt = 0$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(t)\cos(2t)dt = 0$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(t)\sin(2t)dt = 0$, and $\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(2t)\sin(2t)dt = 0$.
   Moreover, we have $\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\cos(t)dt = 1$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(t)\sin(t)dt = 1$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(2t)\cos(2t)dt = 1$, and $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(2t)\sin(2t)dt = 1$.
   Hence S is an orthonormal set.

2.

$$\|t\| = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} t^2 dt} = \sqrt{\frac{2}{3}\pi^2} = \frac{\sqrt{6}\pi}{3}.$$

3. Observe that $\langle \cos(t), \cos(3t) \rangle = 0$, $\langle \sin(t), \cos(3t) \rangle = 0$, $\langle \cos(2t), \cos(3t) \rangle = 0$, $\langle \sin(2t), \cos(3t) \rangle = 0$. Hence we have $\mathrm{proj}_X(\cos(3t)) = 0$.

4. Note that $\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)tdt = 0$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(t)tdt = 2$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(2t)tdt$, $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(2t)tdt = -1$. Hence, $\mathrm{proj}_X(t) = 2\sin(t) - \sin(2t)$.

Exercise 9

Exercise 10

Proof. Suppose $Q = [\mathbf{q_1}, \mathbf{q_2}, ..., \mathbf{q_n}]$ and $\mathbf{x} = [x_1, x_2, ..., x_n]^T, \mathbf{y} = [y_1, y_2, ..., y_n]^T$. Then

$$< Q\mathbf{x}, Q\mathbf{y} >= \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{x_i} y_j < \mathbf{q_i}, \mathbf{q_j} > .$$

By definition, this equals $< \mathbf{x}, \mathbf{y} >= \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{x_i} y_j$ only when $< \mathbf{q_i}, \mathbf{q_j} >= 0$ if $i \neq j$ and $< \mathbf{q_i}, \mathbf{q_j} >= 0$ if $i = j$. This indicates that $Q^H Q = I$ and $QQ^H = I$.
For the other direction, observe that $Q^H Q = I$ and $QQ^H = I$ imply $< \mathbf{q_i}, \mathbf{q_j} >= 0$ if $i \neq j$ and $< \mathbf{q_i}, \mathbf{q_j} >= 0$ if $i = j$. The result then follows immediately.

$$
\begin{aligned}
\|Q\mathbf{x}\| &= \sqrt{< x_1\mathbf{q_1} + x_2\mathbf{q_2} + ... + x_n\mathbf{q_n}, x_1\mathbf{q_1} + x_2\mathbf{q_2} + ... + x_n\mathbf{q_n} >} \\
&= \sqrt{\sum_{i,j} \overline{x_i} x_j < \mathbf{q_i}, \mathbf{q_j} >} \\
&= \sqrt{\sum_{i} \overline{x_i} x_i \times 1} = \|\mathbf{x}\|
\end{aligned}
$$

To show that Q^{-1} is an orthonormal matrix, observe that $QQ^H = I$ and $QQ^{-1} = I$. This implies that $Q^{-1} = Q^H$. Then it is trivially true that ${Q^{-1}}^H Q^{-1} = I$ and $Q^{-1}{Q^{-1}}^H = I$.

The columns of Q are orthonormal have been shown in part 1.

Since Q is an orthonormal matrix, we know $Q^{-1} = Q^H$. Hence, $\det(Q)\det(Q^H) = \det(QQ^H) = \det(I) = 1$. Since $\det(Q) = \det(Q^H)$, it follows that $|\det(Q)| = 1$. The converse is not true.

Observe that $(Q_1Q_2)(Q_1Q_2)^H = Q_1Q_2Q_2^HQ_1^H = Q_1IQ_1^H = Q_1Q_1^H = I$. Also, $(Q_1Q_2)^H(Q_1Q_2) = Q_2^HQ_1^HQ_1Q_2 = Q_2^HIQ_2 = I$. Hence Q_1Q_2 is also an orthonormal matrix. $\square$

Exercise 11

Suppose there are only r independent vectors. Then we would first get r orthonormal vectors and then get $n - r$ zero vectors.

Sorry I don't have time to TeX out the rest of the problems.

3.16

1) Let $D = \begin{bmatrix} -1 & & 0 \\ & 1 & \\ 0 & & \ddots \end{bmatrix}$

then $QD = [q_1\ q_2 \cdots q_n]\begin{bmatrix} -1 & & \\ & 1 & \\ & & \ddots \end{bmatrix}$

$= [-q_1\ q_2 \cdots q_n]$

is still an orthonormal matrix

and $D^{-1}R$ is still an upper-triangular matrix.

Observe that

$$QD \cdot D^{-1}R = Q(DD^{-1})R = QR = A$$

Hence $A = (QD)(D^{-1}R)$ is another form of QR decomposition, so it is not unique.

2) We need the following lemmata:

① If Q_1 and Q_2 are orthonormal matrices, then so does $Q_1^T Q_2$.

proof: $(Q_1^TQ_2)(Q_1^TQ_2)^T = (Q_1^TQ_2)(Q_2^TQ_1)$

$= Q_1^T(Q_2Q_2^T)Q_1 = Q_1^TQ_1 = I$

$(Q_1^TQ_2)^T(Q_1^TQ_2) = Q_2^TQ_1Q_1^TQ_2 = I$ ☐

② If U is an invertible upper-triangular matrix, then so does U^{-1}.

proof easily follows from induction

③ if U_1 and U_2 are upper-triangular matrix, then so does U_1U_2. ($n\times n$ matrix)

proof: Let $U_1 = [a_{ij}]$ $U_2 = [b_{ij}]$

Since both are upper triangular, it follows that

$a_{ij} = 0$ if $i>j$, $b_{ij} = 0$ if $i>j$.

Let $C = [c_{ij}] = U_1U_2$, then $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

Fix some $i>j$

$c_{ij} = (a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{i(i-1)}b_{(i-1)j})$

$+ (a_{ii}b_{ij} + a_{i(i+1)}b_{(i+1)j} + \cdots + a_{in}b_{nj})$

Note that in the first term, all $a_{ik} = 0$

in the second term, all $b_{ki} = 0$

Hence $c_{ij} = 0$ when $i>j$.

$\Rightarrow C$ is upper triangular. ☐

Now, by contradiction, assume that

$A = Q_1R_1$ and $A = Q_2R_2$, where both R_1 and R_2 have positive diagonal elements.

Then $Q_1R_1 = Q_2R_2 \Rightarrow Q_1^TQ_2 = R_1R_2^{-1}$.

Let $M = Q_1^TQ_2 = R_1R_2^{-1}$

By Lemma 1 M is orthonormal

By Lemma 2 & 3 M is upper triangular with positive diagonals.

It follows that M must be the identity matrix I.

$\therefore R_1R_2^{-1} = I \Rightarrow R_1 = R_2$

and therefore $Q_1 = Q_2$.

Hence the decomposition is unique. ☐

3.17. Since $\hat{R}$ is an $(n\times n)$ upper-triangular matrix, $\hat{R}$ is invertible, so does $\hat{R}^H$.

~~Lemma 1: $\forall A \in M_{m\times n}$, A^HA is invertible.~~

~~Now, from $A^HA\vec{x} = A^Hb$ and Lemma 1, we know $\vec{x} = (A^HA)^{-1}A^Hb$.~~

Since $A = \hat{Q}\hat{R}$,

$$A^HA\vec{x} = A^Hb$$

$\Rightarrow (\hat{Q}\hat{R})^H(\hat{Q}\hat{R})\vec{x} = (\hat{Q}\hat{R})^Hb$

$\Rightarrow \hat{R}^H\cancel{\hat{Q}^H\hat{Q}}\hat{R}\vec{x} = \hat{R}^H\hat{Q}^Hb$

$\Rightarrow \hat{R}^H\hat{R}\vec{x} = \hat{R}^H\hat{Q}^Hb \Rightarrow \hat{R}\vec{x} = \hat{Q}^Hb$

Hence the two systems are equivalent. ☐

3.23

$$\|x-y\|^2 = \langle x-y, x-y\rangle$$
$$= \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle$$
$$\geq \langle x,x\rangle - |\langle x,y\rangle| - |\langle y,x\rangle| + \langle y,y\rangle$$
$$\geq \langle x,x\rangle - \|x\|\|y\| - \|y\|\|x\| + \langle y,y\rangle$$
$$= \|x\|^2 - 2\|x\|\|y\| + \|y\|^2$$
$$= (\|x\| - \|y\|)^2$$

This implies $\|x-y\| \geq \big|\|x\| - \|y\|\big|$ ▨

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

① $\|f\|_{L^1} \geq 0$ is trivial.

Observe that since $|f(t)| \geq 0$,

$\int_a^b |f(t)|\,dt = 0$ iff $f(t) \equiv 0$ on $[a,b]$.

② $\|\alpha f\|_{L_1} = \int_a^b |\alpha f|\,dt$

$= |\alpha| \int_a^b |f|\,dt = |\alpha|\,\|f\|_{L_1}$

③ $\|f+g\|_{L_1} = \int_a^b |f+g|\,dt$

$\leq \int_a^b |f| + |g|\,dt = \int_a^b |f|\,dt + \int_a^b |g|\,dt$

$= \|f\|_{L_1} + \|g\|_{L_1}$

(2) $\|f\|_{L^2} = \left[\int_a^b |f(t)|^2\,dt\right]^{\frac{1}{2}}$

① $\|f\|_{L^2} \geq 0$ is trivial. $= 0$ iff $f \equiv 0$.

② $\|\alpha f\|_{L^2} = \left[\int_a^b |\alpha f|^2\,dt\right]^{\frac{1}{2}}$

$= \left[\int_a^b |\alpha|^2 |f|^2\,dt\right]^{\frac{1}{2}}$

$= |\alpha| \left[\int_a^b |f|^2\,dt\right]^{\frac{1}{2}} = |\alpha|\,\|f\|_{L^2}$.

③ $\|f+g\|_{L^2} = \left[\int_a^b |f+g|^2\,dt\right]^{\frac{1}{2}}$

$= \left[\int_a^b |f|^2 + |g|^2 + 2|f||g|\,dt\right]^{\frac{1}{2}}$

In the inner product space $L^2[a,b]$, the inner product is defined as $\langle f, g\rangle = \int_a^b (\bar{f}\,g)^2\,dt$.

Then by Cauchy-Schwarz,

$$|\langle f, g\rangle| \leq \|f\|\,\|g\|$$

i.e. $\left|\int_a^b \bar{f} g\,dt\right|^2 \leq \int_a^b |\bar{f} f|\,dt \cdot \int_a^b |\bar{g} g|\,dt$

$\Rightarrow \left|\int_a^b |f||g|\,dt\right|^2 \leq \int_a^b |f|^2\,dt \cdot \int_a^b |g|^2\,dt$

This implies $\int_a^b |f||g|\,dt \leq \sqrt{\int_a^b |f|^2\,dt \int_a^b |g|^2\,dt}$

Hence

$\int_a^b f^2\,dt + \int_a^b g^2\,dt + 2\int_a^b |f||g|\,dt \leq \int_a^b f^2\,dt + \int_a^b g^2\,dt$
$+ 2\sqrt{\int_a^b f^2\,dt \int_a^b g^2\,dt}$.

$\Rightarrow \int_a^b f^2 + g^2 + |f||g|\,dt \leq \left(\sqrt{\int_a^b f^2\,dt} + \sqrt{\int_a^b g^2\,dt}\right)^2$

$\Rightarrow \left[\int_a^b f^2 + g^2 + |f||g|\,dt\right]^{\frac{1}{2}} \leq \left[\int_a^b f^2\,dt\right]^{\frac{1}{2}} + \left[\int_a^b g^2\,dt\right]^{\frac{1}{2}}$

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

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

① $\|f\|_{L^\infty} \geq 0$ is trivial. $= 0$ iff $f \equiv 0$.

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

$\leq \sup_{x\in[a,b]} |f(x)| + |g(x)|$

$= \sup_{x\in[a,b]} |f(x)| + \sup_{x\in[a,b]} |g(x)|$

$= \|f\|_{L^\infty} + \|g\|_{L^\infty}$

▨

3.26

Proof that this is an equivalence relationship:

① Obviously $\|\cdot\|_a$ is topologically equivalent to $\|\cdot\|_a$ by choosing $m = M = 1$

② If $m\|x\|_a \le \|x\|_b \le M\|x\|_a$, $\forall \vec{x}$,

then $\frac{1}{M}\|x\|_b \le \|x\|_a \le \frac{1}{m}\|x\|_b$, $\forall \vec{x}$

So it is symmetric

③ if $m\|x\|_a \le \|x\|_b \le M\|x\|_a$, and

$n\|x\|_b \le \|x\|_c \le N\|x\|_b \quad \forall \vec{x}$

then $mn\|x\|_a \le \|x\|_c \le MN\|x\|_a$, $\forall \vec{x}$

So it is transitive

Hence it is an equivalence relationship.

(1) $\|\vec{x}\|_2 \le \|\vec{x}\|_1 \le \sqrt{n}\|\vec{x}\|_2$

Since $\|\vec{x}\|_2^2 = \sum_{i=1}^{n} x_i^2$

$$\|\vec{x}\|_1^2 = \sum_{i=1}^{n}|x_i|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}|x_i||x_j|$$

$$= \sum_{i=1}^{n} x_i^2 + \sum_{i\ne j}|x_i||x_j|$$

$$\ge \sum_{i=1}^{n} x_i^2 = \|\vec{x}\|_2^2$$

Hence $\|\vec{x}\|_1 \ge \|\vec{x}\|_2$

Now take the inner product on $\mathbb{F}^n$ to be

$$\langle \vec{x}, \vec{y}\rangle = \sum_{i=1}^{n} x_i y_i.$$

Let $\vec{u} = [\,\mathrm{sgn}(x_1), \mathrm{sgn}(x_2), \dots \mathrm{sgn}(x_n)]^T$

where $\mathrm{sgn}(x_i) = x_i/|x_i|$ when $x_i \ne 0$ and $\mathrm{sgn}(0) = 1$.

Observe that $\|\vec{x}\|_1 = \sum_{i=1}^{n}|x_i| = \sum_{i=1}^{n} x_i\cdot \mathrm{sgn}(x_i)$

$$= \sum_{i=1}^{n}\mathrm{sgn}(x_i)\, x_i$$

$$= |\langle \vec{u}, \vec{x}\rangle|$$

By Cauchy - Schwarz,

$$|\langle \vec{u}, \vec{x}\rangle| \le \|\vec{u}\|_2\|\vec{x}\|_2 = \sqrt{\sum_{i=1}^{n} 1^2}\cdot\|\vec{x}\|_2$$

$$= \sqrt{n}\|\vec{x}\|_2.$$

Hence $\|\vec{x}\|_1 \le \sqrt{n}\|\vec{x}\|_2$.

and therefore $\|\vec{x}\|_2 \le \|\vec{x}\|_1 \le \sqrt{n}\|\vec{x}\|_2$.

(2) $\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$.

Suppose x_k is the largest term in magnitude.

i.e. $\|\vec{x}\|_\infty = \max_{i=1}^{n}|x_i| = |x_k|$.

Then, $\|x\|_2^2 = \sum_{i=1}^{n}|x_i|^2 = x_1^2 + x_2^2 + \dots + x_k^2 + \dots + x_n^2$

$$\ge 0 + 0 + \dots + 0 + x_k^2 + 0 + \dots + 0$$

$$= x_k^2 = \|\vec{x}\|_\infty^2$$

Hence $\|x\|_2 \ge \|x\|_\infty$.

Moreover,

$$\|x\|_2^2 = x_1^2 + x_2^2 + \dots + x_k^2 + \dots + x_n^2$$

$$\le x_k^2 + x_k^2 + \dots + x_k^2 + \dots + x_k^2$$

$$= n x_k^2$$

$\Rightarrow \|x\|_2^2 \le n\|x\|_\infty^2 \Rightarrow \|x\|_2 \le \sqrt{n}\|x\|_\infty$

Hence, $\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$ ▨

3.28.

(1) $\frac{1}{\sqrt{n}}\|A\|_2 \le \|A\|_1 \le \sqrt{n}\|A\|_2$

$\forall \vec{x} \ne \vec{0}$, $\dfrac{\|Ax\|_1}{\|x\|_1} \ge \dfrac{\|Ax\|_2}{\sqrt{n}\|x\|_2}$ by 3.26(i)

Take sup on both sides and we get

$$\|A\|_1 \ge \frac{1}{\sqrt{n}}\|A\|_2.$$

$\forall \vec{x} \ne \vec{0}$ $\dfrac{\sqrt{n}\|Ax\|_2}{\|x\|_2} \ge \dfrac{\|Ax\|_1}{\|x\|_1}$ by 3.26(i)

Take sup on both sides and we get

$$\sqrt{n}\|A\|_2 \ge \|A\|_1$$

Hence $\frac{1}{\sqrt{n}}\|A\|_2 \le \|A\|_1 \le \sqrt{n}\|A\|_2$

(ii) $\frac{1}{\sqrt{n}}\|A\|_\infty \le \|A\|_2 \le \sqrt{n}\|A\|_\infty$

$\forall \vec{x}\neq\vec{0}$, $\frac{\|Ax\|_2}{\|x\|_2} \geq \frac{\|Ax\|_\infty}{\sqrt{n}\|x\|_\infty}$ by 3.26(ii)

Take sup on both sides,

$\Rightarrow \|A\|_2 \geq \frac{1}{\sqrt{n}}\|A\|_\infty$.

$\forall \vec{x}\neq\vec{0}$, $\frac{\|Ax\|_2}{\|x\|_2} \le \frac{\sqrt{n}\|Ax\|_\infty}{\|x\|_\infty}$ by 3.26(ii)

Take sup on both sides and we have

$\|A\|_2 \le \sqrt{n}\|A\|_\infty$

Hence $\frac{1}{\sqrt{n}}\|A\|_\infty \le \|A\|_2 \le \sqrt{n}\|A\|_\infty$ ∎

3.29. We first show that $\|Q\|_2 = 1$, where Q is any orthonormal matrix.

Observe that $\forall \vec{x}\neq\vec{0}$,

$\|Q\vec{x}\|_2^2 = \langle Q\vec{x}, Q\vec{x}\rangle = (Qx)^T Qx$

$= x^T Q^T Q x = x^T x = \langle x, x\rangle = \|x\|_2^2$.

Hence, $\forall \vec{x}\neq\vec{0}$ $\frac{\|Qx\|_2}{\|x\|_2} = 1$

$\Rightarrow \|Q\|_2 = \sup_{\vec{x}\neq\vec{0}} \frac{\|Qx\|_2}{\|x\|_2} = 1$.

By def. $\|A\|_2 = \sup_{\vec{x}\neq 0} \frac{\|Ax\|_2}{\|x\|_2}$

$\Rightarrow \|Ax\|_2 \le \|A\|_2\|x\|_2$, $\forall\vec{x}$

Now $\frac{\|Ax\|_2}{\|A\|_2} \le \frac{\|A\|_2\|x\|_2}{\|A\|_2} = \|x\|_2$, $\forall\vec{x}$

Take sup on both sides, we have

$\|Rx\|_2 = \sup_{\|A\|_2\neq 0} \frac{\|Ax\|_2}{\|A\|_2} \le \|x\|_2$

I don't know what's next.

3.30. To show that $\|\cdot\|_S$ matrix norm, it suffices to show nonnegativity, homogeneity and triangular inequality.

① $\forall A$, $\|A\|_S = \|SAS^{-1}\| \geq 0$ is trivially true.

Moreover, if $A = O$, $\|A\|_S = \|O\| = 0$

if $\|A\|_S = 0$, then $\|SAS^{-1}\| = 0 \Rightarrow A = O$.

② $\|\alpha A\|_S = \|S\alpha AS^{-1}\| = \alpha\|SAS^{-1}\| = \alpha\|A\|_S$

③ $\|A+B\|_S = \|S(A+B)S^{-1}\|$

$= \|SAS^{-1} + SBS^{-1}\|$

$\le \|SAS^{-1}\| + \|SBS^{-1}\| = \|A\|_S + \|B\|_S$

Hence $\|\cdot\|_S$ is a matrix norm. ∎

3.37 We first find a set of orthonormal basis for V.

Let $p_1 = 1$ $\quad q_1 = \frac{p_1}{\|p_1\|} = \frac{1}{\int_0^1 1\,dx} = 1$

Let $p_2 = x - \text{proj}_1 x = x - \frac{1}{2}$

$q_2 = \frac{p_2}{\|p_2\|} = \sqrt{12}(x - \frac{1}{2})$

Let $p_3 = x^2 - \text{proj}_1 x^2 - \text{proj}_{\sqrt{12}(x-\frac{1}{2})} x^2 = x^2 - x + \frac{1}{6}$

$q_3 = \frac{p_3}{\|p_3\|} = \sqrt{180}(x^2 - x + \frac{1}{6})$

Then. $\vec{q} = \sum_{i=1}^{3} L(\vec{q_i})\vec{q_i}$

$= L(1)\cdot 1 + L(\sqrt{12}(x-\frac{1}{2}))\sqrt{12}(x-\frac{1}{2})$

$+ L(\sqrt{180}(x^2-x+\frac{1}{6}))\sqrt{180}(x^2-x+\frac{1}{6})$

$= 0 + 12(x-\frac{1}{2}) + 180(x^2-x+\frac{1}{6})$

$= 180x^2 - 168x + 24$.

It can be verified that $\forall \vec{p} \in V$,

$L[p] = \langle q, p \rangle$.

3.38

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

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

~~3.39~~ 3.39

1) By definition,

$\langle (S+T)^*(w), v \rangle_V = \langle w, (S+T)(v) \rangle_W$

$= \langle w, S(v) \rangle_W + \langle w, T(v) \rangle_W$

$= \langle S^*(w), v \rangle_V + \langle T^*(w), v \rangle_V$.

$= \langle S^*(w) + T^*(w), v \rangle_V$

Hence $(S+T)^* = S^* + T^*$.

$\langle (\alpha T)^*(w), v \rangle_V = \langle w, (\alpha T)(v) \rangle_W$

$= \bar{\alpha} \langle w, T(v) \rangle_W$

$= \bar{\alpha} \langle T^*(w), v \rangle_V$

$= \langle \bar{\alpha} T^*(w), v \rangle_V$

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

2) $\langle (S^*)^*(w), v \rangle_V = \langle w, S^*(v) \rangle_W$

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

$= \overline{\langle v, S(w) \rangle_V}$

$= \langle S(w), v \rangle_V$

Hence $(S^*)^* = S$.

Let $\vec{w}, \vec{v} \in V$

3) $\langle (ST)^*(w), v \rangle_V = \langle w, ST(v) \rangle_V$

$= \langle w, S[T(v)] \rangle_V$

$= \langle S^*(w), T(v) \rangle_V$

$= \langle T^*[S^*(w)], v \rangle_V$

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

4) By 3)

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

$\Rightarrow (T^{-1})^* = (T^*)^{-1}$.

3.40

1) By definition, let $B_1, B_2 \in M_n(\mathbb{F})$

$< A^* B_2, B_1 >_F = < B_2, AB_1 >_F$.

Since $< A^H B_2, B_1 >_F = tr[(A^H B_2)^H B_1]$

$= tr[B_2^H A B_1] = tr[B_2^H (AB_1)]$

$= < B_2, AB_1 >_F$,

we have $< A^* B_2, B_1 >_F = < A^H B_2, B_1 >_F$,

$\forall B_1, B_2$

$\Rightarrow A^* = A^H$. ∎

2) By 1), we know $A^* = A^H$

Hence, $< A_2 A_1^*, A_3 >_F$

$= tr((A_2 A_1^*)^H A_3)$

$= tr((A_1^*)^H A_2^H A_3)$

$= tr((A_1^H)^H A_2^H A_3)$

$= tr(A_1 A_2^H A_3)$

$= tr(A_2^H A_3 A_1)$

$= < A_2, A_3 A_1 >_F$ ∎

3.44.

If $\vec{b}=0$, then $\vec{b}\in R(A)$

and $\vec{x}=\vec{0}$ is a solution to $A\vec{x}=\vec{b}$.

Now if $\vec{b}\neq 0$

Since $\mathbb{F}^m = R(A)\oplus N(A^H)$,

Either $\vec{b}\in R(A)$ or $\vec{b}\in N(A^H)$

If $\vec{b}\in R(A)$, $\exists\,\vec{x}$ a solution;

If $\vec{b}\in N(A^H)$, let $\vec{y}=\vec{b}$,

since $\vec{b}\neq\vec{0}$, $\langle\vec{y},\vec{b}\rangle=\langle\vec{b},\vec{b}\rangle\neq 0$. $\blacksquare$

3.45.

We need the following lemmata.

1) $tr(AB)=tr(BA)$, $A,B\in M_n(\mathbb{R})$

2) $tr(A)=tr(A^T)$, $A\in M_n(\mathbb{R})$

3) $tr(A+B)=tr(A)+tr(B)$, $A,B\in M_n(\mathbb{R})$

We first show that $Skew_n(\mathbb{R})\subset Sym_n(\mathbb{R})^\perp$

Pick $B\in Skew_n(\mathbb{R})$, $\forall A\in Sym_n(\mathbb{R})$,

$\langle B,A\rangle = tr(B^TA) = tr(-BA) = -tr(BA)$

on the other hand,

$\langle B,A\rangle = tr(B^TA) = tr((B^TA)^T) = tr(A^TB)$

$= tr(AB) = tr(BA)$

Hence, $tr(BA) = -tr(BA) = 0$

$\Rightarrow \langle B,A\rangle = 0$, $\forall A\in Sym_n(\mathbb{R})^\perp$

$\Rightarrow Skew_n(\mathbb{R})\subset Sym_n(\mathbb{R})^\perp$

Next we show $Sym_n(\mathbb{R})^\perp\subset Skew_n(\mathbb{R})$

Pick any $A\in Sym_n(\mathbb{R})^\perp$,

observe that $(A+A^T)^T = A+A^T \Rightarrow A+A^T\in Sym_n(\mathbb{R})$

Hence, $\langle A^T+A, A\rangle = 0$.

i.e. $tr((A^T+A)^TA) = tr((A+A^T)A) = tr(A^2+A^TA) = 0$

Then, $tr(A^2)+tr(A^TA) = 0$

$\Rightarrow \langle A,A\rangle = -\langle A^T,A\rangle$

$\Rightarrow A^T = -A$

Hence $A\in Skew_n(\mathbb{R})$ and $Sym_n(\mathbb{R})^\perp\subset Skew_n(\mathbb{R})$.

Finally we conclude that

$Sym_n(\mathbb{R})^\perp = Skew_n(\mathbb{R})$ $\blacksquare$

3.46

(i) $x\in N(A^HA)$

$Ax\in R(A)$ is trivial since A maps x to Ax,

so Ax is in $R(A)$.

Since $x\in N(A^HA)$,

$A^HAx=0 \Rightarrow A^H(Ax)=0$

$\Rightarrow Ax\in N(A^H)$.

(ii) $N(A^HA) = N(A)$

① $N(A^HA)\subset N(A)$

pick $\vec{x}\in N(A^HA)$, then $A^HAx=0$.

if $\vec{x}=0$, then $x=\vec{0}\in N(A)$.

if $\vec{x}\neq 0$, we want to show $Ax=0$

By contradiction, assume $Ax\neq 0$.

Then $A^H(Ax)=0$ implies that $Ax\in N(A^H)$

Since $Ax\in R(A)$ and $Ax\neq 0$, this contradicts the

fact that $R(A)^\perp = N(A^H)$.

Hence $Ax=0$ and $x\in N(A)$.

Therefore $N(A^HA)\subset N(A)$.

② $N(A)\subset N(A^HA)$

pick $\vec{x}\in N(A)$, then $Ax=0$.

It follows that $A^HAx = A^H(Ax) = A^H\vec{0} = \vec{0}$

Hence $\vec{x}\in N(A^HA)$ and $N(A)\subset N(A^HA)$

We conclude that $N(A^HA) = N(A)$.

(iii) Observe that both A and A^HA are linear

transformations from $\mathbb{R}^n \mapsto \mathbb{R}^m$.

By rank-nullity Thm, $\dim(V) = rank(L)+\dim(N(L))$

where $L: V\mapsto W$.

Since $N(A^HA) = N(A)$ by (ii),

we have $\dim(N(A^HA)) = \dim(N(A))$.

It follows that

$$\operatorname{rank}(A^H A) = \dim(\mathbb{R}^n) - \dim(N(A^H A))$$
$$= \dim(\mathbb{R}^n) - \dim(N(A))$$
$$= \operatorname{rank}(A).$$

(iv.) Since $A \in M_{m\times n}(\mathbb{R})$

$A^T A \in M_{n\times n}(\mathbb{R})$

if A has linearly indep. columns, then $\operatorname{rank}(A) = n$.

It follows that $\operatorname{rank}(A^T A) = \operatorname{rank}(A) = n$.

Since $A^T A$ is an $(n\times n)$ matrix, it is non-singular. ▨

3.47.

(i) $P^2 = [A(A^H A)^{-1}A^H][A(A^H A)^{-1}A^H]$

$= A\cancel{(A^H A)^{-1}}\cancel{(A^H A)}(A^H A)^{-1}A^H$

$= A(A^H A)^{-1}A^H$

$= P.$ ▨

(ii) Lemma: $(A^{-1})^H = (A^T)^{-1}$

proof of lemma:

$(A^{-1})^H A^H = (AA^{-1})^H = I^H = I$

$A^T(A^{-1})^H = (A^{-1}A)^H = I^H = I$ ▨

$P^H = [A(A^H A)^{-1}A^H]^H$

$= A[(A^H A)^{-1}]^H A^H$

$= A[(A^H A)^H]^{-1} A^H$

$= A(A^H A)^{-1} A^H = P.$ ▨

(iii) Since we know that rank will not increase in matrix multiplication, we know that

$\operatorname{rank}(P) \le \operatorname{rank}(A) = n.$

Now, $\forall y \in R(A)$ $\exists x$, s.t. $Ax = y$

Observe that $Py = A(A^H A)^{-1}A^H y$

$= A\cancel{(A^H A)^{-1}}\cancel{(A^H A)}x = Ax = y$

Hence $y \in R(P)$

It follows that $R(A) \subset R(P)$

So $n = \operatorname{rank}(A) \le \operatorname{rank}(P)$

We now conclude that $\operatorname{rank}(P) = n$. ▨

3.48

(1) $P(\alpha A + \beta B) = \dfrac{\alpha A + \beta B + (\alpha A+\beta B)^T}{2}$

$= \dfrac{\alpha A + \beta B + \alpha A^T + \beta B^T}{2} = \dfrac{\alpha}{2}(A+A^T) + \dfrac{\beta}{2}(B+B^T)$

$= \alpha P(A) + \beta P(B).$ ▨

(2) $\forall A$, $P(P(A)) = \dfrac{P(A) + P(A)^T}{2}$

$= \frac{1}{2}\left[\dfrac{A+A^T}{2} + \left(\dfrac{A+A^T}{2}\right)^T\right]$

$= \frac{1}{2}\left[\dfrac{A+A^T}{2} + \dfrac{A+A^T}{2}\right] = \dfrac{A+A^T}{2} = P(A)$

$\Rightarrow P^2 = P.$ ▨

(3) By def, $\langle P^*(A), B\rangle = \langle A, P(B)\rangle$

Now, $\langle A, P(B)\rangle = \langle A, \frac{B+B^T}{2}\rangle$

$= \frac{1}{2}\langle A,B\rangle + \frac{1}{2}\langle A, B^T\rangle$

$= \frac{1}{2}\operatorname{tr}(A^T B) + \frac{1}{2}\operatorname{tr}(A^T B^T)$

$= \frac{1}{2}\operatorname{tr}(A^T B) + \frac{1}{2}\operatorname{tr}[(BA)^T]$

$= \frac{1}{2}\operatorname{tr}(A^T B) + \frac{1}{2}\operatorname{tr}(BA)$

$= \frac{1}{2}\operatorname{tr}(A^T B) + \frac{1}{2}\operatorname{tr}(AB)$

Also, $\langle P(A), B\rangle = \langle \frac{A+A^T}{2}, B\rangle$

$= \frac{1}{2}\langle A,B\rangle + \frac{1}{2}\langle A^T, B\rangle$

$= \frac{1}{2}\operatorname{tr}(A^T B) + \frac{1}{2}\operatorname{tr}(AB)$

~~Hence, $\langle P^*(A)$~~

Hence, $\langle P^*(A), B \rangle = \langle A, P(B) \rangle$

$$= \frac{1}{2}tr(A^TB) + \frac{1}{2}tr(AB)$$

$$= \langle P(A), B \rangle.$$

$\Rightarrow P^* = P$ ∎

(4) Suppose $A \in N(P)$

then $P(A) = \dfrac{A + A^T}{2} = 0$

$\Rightarrow A + A^T = 0 \Rightarrow A^T = -A$

$\Rightarrow A \in Skew_n(\mathbb{R})$

Hence $N(P) \subset Skew_n(\mathbb{R})$

Now, $\forall B \in Skew_n(\mathbb{R})$, $-B = B^T$

$\therefore P(B) = \dfrac{B + B^T}{2} = 0$

$\Rightarrow B \in N(P)$

Hence $Skew_n(\mathbb{R}) \subset N(P)$

We conclude that

$Skew_n(\mathbb{R}) = N(P)$ ∎

(5) Suppose $B \in R(P)$, then

$B = \dfrac{A+A^T}{2}$ for some A.

Observe that $B^T = \frac{1}{2}(A+A^T)^T = \frac{1}{2}(A^T+A) = B$

Hence $B \in Sym_n(\mathbb{R})$ and $R(P) \subset Sym_n(\mathbb{R})$

Conversely, if $B \in Sym_n(\mathbb{R})$, then $B^T = B$.

Let $A = B$, then

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

$\Rightarrow B \in R(P) \Rightarrow Sym_n(\mathbb{R}) \subset R(P)$.

Finally we have $Sym_n(\mathbb{R}) = R(P)$ ∎

(6) $\|A - P(A)\|_F$

$$= \left\| A - \frac{A+A^T}{2} \right\|_F$$

$$= \left\| \frac{A-A^T}{2} \right\|_F$$

$$= \frac{1}{2}\|A - A^T\|_F$$

$$= \frac{1}{2}\sqrt{tr((A-A^T)^T(A-A^T))}$$

$$= \frac{1}{2}\sqrt{tr(A^TA - A^TA^T - AA + AA^T)}$$

$$= \frac{1}{2}\sqrt{2tr(A^TA) - 2tr(A^2)}$$

$$= \sqrt{\frac{tr(A^TA) - tr(A^2)}{2}}$$ ∎

50. $A = \begin{bmatrix} x_1^2 & y_1^2 \\ x_2^2 & y_2^2 \\ \vdots & \vdots \\ x_n^2 & y_n^2 \end{bmatrix}$ $\vec{x} = \begin{bmatrix} r \\ s \end{bmatrix}$ $\vec{b} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

normal eq. $A^TA\hat{x} = A^Tb$