Problem Set 2

Fiona Fan

With reference to Zushun Zong and Winston Xu.

1

i)

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

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

In $\mathbb{R}$, $\langle x, y\rangle = \sum_i a_i b_i = \sum_i b_i a_i = \langle y, x\rangle$,

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

ii)

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

$$\frac{1}{2}(\|x+y\|^2 + \|x-y\|^2) = \|x\|^2 + \|y\|^2$$

2

$$\begin{aligned}
\text{RHS} &= \frac{1}{4}[\langle \vec{x}+\vec{y}, \vec{x}+\vec{y}\rangle - \langle \vec{x}-\vec{y}, \vec{x}-\vec{y}\rangle + i\langle \vec{x}-i\vec{y}, \vec{x}-i\vec{y}\rangle - i\langle \vec{x}+i\vec{y}, \vec{x}+i\vec{y}\rangle] \\
&= \frac{1}{4}[\langle \vec{x}, \vec{x}\rangle + \langle \vec{x}, \vec{y}\rangle + \langle \vec{y}, \vec{x}\rangle + \langle \vec{y}, \vec{y}\rangle - \langle \vec{x}, \vec{x}\rangle + \langle \vec{x}, \vec{y}\rangle + \langle \vec{y}, \vec{x}\rangle \\
&- \langle \vec{y}, \vec{y}\rangle + i\langle \vec{x}, \vec{x}\rangle + \langle \vec{x}, \vec{y}\rangle - \langle \vec{y}, \vec{x}\rangle - i\langle \vec{y}, \vec{y}\rangle - i\langle \vec{x}, \vec{x}\rangle + \langle \vec{x}, \vec{y}\rangle \\
&- \langle \vec{y}, \vec{x}\rangle + i\langle \vec{y}, \vec{y}\rangle] \\
&= \frac{1}{4}(4\langle \vec{x}, \vec{y}\rangle) = \langle \vec{x}, \vec{y}\rangle = \text{LHS}
\end{aligned}$$

3 We need the following computation:

$\langle x, x^5\rangle = \int_0^1 x^6 dx = \frac{1}{7}$

$\sqrt{\langle x, x\rangle} = \sqrt{\int_0^1 x^2 dx} = \sqrt{\frac{1}{3}}$

$\sqrt{\langle x^5, x^5\rangle} = \sqrt{\int_0^1 x^1 0 dx} = \sqrt{\frac{1}{11}}$

$\langle x^2, x^4\rangle = \int_0^1 x^6 dx = \frac{1}{7}$

$\sqrt{\langle x^2, x^2\rangle} = \sqrt{\int_0^1 x^4 dx} = \sqrt{\frac{1}{5}}$

$\sqrt{\langle x^4, x^4\rangle} = \sqrt{\int_0^1 x^8 dx} = \sqrt{\frac{1}{9}}$
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. $$\theta_2 = \arccos(\frac{\langle x^2, x^4\rangle}{\|x^2\|\,\|x^4\|}) = \arccos(\frac{\frac{1}{7}}{\sqrt{\frac{1}{5}}\sqrt{\frac{1}{9}}}) = \arccos(\frac{\sqrt{45}}{7})$$

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 $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(2t)\sin(2$
   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, $\mathrm{proj}_X(t) = 2\sin(t) - \sin(2t)$.

9 we can convert the rotation transformation into a matrix in the standard basis Q. If we can show that $Q^TQ = I$,then the transformation is orthonormal.

$$Q = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$\Rightarrow$

$$Q^TQ = \begin{bmatrix} \cos(\theta)^2 + \sin(\theta)^2 & 0 \\ 0 & \cos(\theta)^2 + \sin(\theta)^2 \end{bmatrix}$$

$\Rightarrow$

$$Q^TQ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Hence, the transformation is orthonormal.

10

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

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

By definition, this equals $< \vec{x}, \vec{y} >= \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{x_i} y_j$ only when $< \vec{q_i}, \vec{q_j} >= 0$ if $i \neq j$ and $< \vec{q_i}, \vec{q_j} >= 0$ 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 $< \vec{q_i}, \vec{q_j} >= 0$ if $i \neq j$ and $< \vec{q_i}, \vec{q_j} >= 0$ if $i = j$. The result then follows immediately.

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

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. □

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

16

1) Let $D =$

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

Then $QD = \begin{bmatrix} -q_1 & q_2 & \dots & q_n \end{bmatrix}$, which 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 lemmas:

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

Proof.

$$(Q_1^T Q_2)(Q_1^T Q_2)^T = (Q_1^T Q_2)(Q_2^T Q_1) = Q_1^T(Q_2 Q_2^T)Q_1 = Q_1^T Q_1 = I$$

$$(Q_1^T Q_2)^T(Q_1^T Q_2) = Q_2^T(Q_1 Q_1^T)Q_2 = I$$

□

2. If U is an invertible upper-triangular matrix, then so is U^{-1}. The proof follows induction.
3. If U_1 and U_2 are upper-triangular matrices, then so is $U_1 U_2$.

Proof. Let $U_1 = [a_{ij}]$, $U_2 = [b_{ij}]$. Since bothe are upper triangular, it follows that $a_{ij}, b_{i,j} = 0$ if $i > j$. □

Let $C = [c_{i,j}] = U_1 U_2$, then $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.
Fix some $i > j$, then

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

Note that in the first term, all $a_{ik} = 0$, and in the second term, all $b_{ki} = 0$.
Hence $c_{ij} = 0$ where $i > j$, $\Rightarrow C$ is upper-triangular.
Now by contradiction, assume that $A = Q_1 R_1$ and $A = Q_2 R_2$, where both R_1 and R_2 have positive diagonal elements.
Then $Q_2 R_2 = Q_1 R_1$, $\Rightarrow Q_1^T Q_2 = R_1 R_2^{-1}$.
Let $M = Q_1^T Q_2 = R_1 R_2^{-1}$, by lemma 1, M is orthonormal. By lemma 2 and 3, M is upper triangular with positive diagonals. It follows that M must be the identity matrix I.
Hence, $R_1 R_2^{-1} = I \Rightarrow R_1 = R_2$, and therefore, $Q_1 = Q_2$.
Thus, the decomposition is unique.

17 Since $\hat{R}$ is an n by n upper-triangular matrix, $\hat{R}$ is invertible, so is $\hat{R}^H$.
Since $A = \hat{Q}\hat{R}$,

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

Hence the two systems are equivalent.

23

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

24

1). 1. $\|f\|)_{L^1} \geq 0$ is trivial.

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

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

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

Observe that since $|f(t)| \geq 0$,
$original = 0 \iff f(t) = 0$ on $[a,b]$.

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

In $\mathbb{L}^2, [a,b]$, $\langle f,g\rangle = \int_a^b (\bar{f}g)^2 dt$.
By Cauchy-Schwarz,
$|\langle f,g\rangle| \leq \|f\|\,\|g\|$
i.e. $|\int_a^b \bar{f}g dt|^2 \leq \int_a^b |\bar{f}f|dt \cdot \int_a^b |\bar{g}g|dt$
$\Rightarrow |\int_a^b |f||g|dt|^2 \leq \int_a^b |f|^2 dt \cdot \int_a^b |g|^2 dt$
Hence, $(\int_a^b |f|^2 + |g|^2 + 2|f||g|dt)^{\frac{1}{2}} \leq \int_a^b |f|^2 dt + \int_a^b |g|^2 dt + 2(\int_a^b f^2 dt \int_a^b g^2 dt)^{\frac{1}{2}}$
$\int_a^b f^2 + g^2 + |f||g|dt \leq (\sqrt{\int_a^b f^2 dt} + \sqrt{\int_a^b g^2 dt})^2$
$\Rightarrow \|f+g\|_{L^2} \leq \|f\|_{L^2} + \|g\|_{L^2}$

3). 1. $\|f\|)_{L^\infty} \geq 0$ is trivial.

Observe that since $|f(t)| \geq 0$,
$original = 0 \iff f(t) = 0$ on $[a,b]$.

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

26 To prove this is an equivalence relationship:

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

2. If $m\|x\|_a \le \|x\|_b \le M\|x\|_a, \quad \forall\vec{x}$,
then $\frac{1}{M}\|x\|_b \le \|x\|_a \le \frac{1}{m}\|x\|_b, \quad \forall\vec{x}$,
so it is symmetric.

3. If $m\|x\|_a \le \|x\|_b \le M\|x\|_a, \quad \forall\vec{x}$, and if $n\|x\|_b \le \|x\|_c \le N\|x\|_b, \quad \forall\vec{x}$,
then $mn\|x\|_a \le \|x\|_c \le MN\|x\|_a \quad \forall\vec{x}$.
Thus it is transitive.
$\Rightarrow$ Thus this is an equivalence relationship.

□

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

$$\|x\|_1^2 = \sum_{i=1}^n |x_i|^2 \tag{1}$$
$$= \sum_{i=1}^n x_i^2 + \sum_{i\neq j} |x_i||x_j| \tag{2}$$
$$\ge \sum_{i=1}^n x_i^2 = \|\vec{x}\|_2^2 \tag{3}$$

Thus, $\|\vec{x}\|_1 \ge \|\vec{x}\|_2$
Let $\vec{u} = [sgn(x_1), \dots, sgn(x_n)]^T$, $\|\vec{x}\|_1 = \sum_{i=1}^n x_i \cdot sgn(x_i) = |\langle\vec{u}, \vec{x}\rangle|$.
By Cauchy-Schwarz, $|\langle\vec{u}, \vec{x}\rangle| \le \|\vec{u}\|_2 \|\vec{x}\|_2 = \sqrt{n}\|\vec{x}\|_2$
Hence, $\|\vec{x}\|_1 \le \sqrt{n}\|\vec{x}\|_2$

2). Let $|x_k| = \|\vec{x}\|_\infty = \max_{i=1}^n |x_i|$
Then, $\|x\|_2^2 = \sum_{i=1}^n |x_i|^2 = x_K^2 = \|\vec{x}\|_\infty^2$
Hence, $\|x\|_2 \ge \|x\|_\infty$
Moreover, $\|x\|_2^2 \le n x_k^2$
$\Rightarrow \|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$

28

i. From the previous exercise, we can get $\sup_{x\neq 0} \frac{||Ax||_1}{||x||_1} \le \sup_{x\neq 0} \frac{||Ax||_1}{||x||_1} \le \sqrt{n}\sup_{x\neq 0} \frac{||Ax||_2}{||x||_2}$, and

$\sup_{x\neq 0} \frac{||Ax||_1}{||x||_1} \ge \sup_{x\neq 0} \frac{||Ax||_2}{||x||_1} \ge \frac{1}{\sqrt{n}} \sup_{x\neq 0} \frac{||Ax||_2}{||x||_2}$, which imply that
$\frac{1}{\sqrt{n}}||A||_2 \le ||A||_1 \le ||A||_2$.

ii. From previous we can get: $\sup_{x\neq 0} \frac{||Ax||_2}{||x||_2} \le \sup_{x\neq 0} \frac{\sqrt{n}||Ax||_\infty}{||x||_\infty}$, and $\sup_{x\neq 0} \frac{||Ax||_2}{||x||_2} \ge \sup_{x\neq 0} \frac{||Ax||_\infty}{\sqrt{n}||x||_\infty}$.

29

Proof. First, NTS: $\|Q\|_2 = 1$, where Q is orthonormal.

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

Next, by definition, $\|A\|_2 = \sup_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2}$,
$\Rightarrow \|Ax\|_2 \leq \|A\|_2\,\|x\|_2$.
Now,
$$\frac{\|Ax\|_2}{\|A\|_2} \leq \frac{\|A\|_2\,\|x\|_2}{\|A\|_2} = \|x\|_2$$

Take sup on both sides, we have
$$\|Rx\|_2 = \sup_{\|A\|_2\neq 0}\frac{\|Ax\|_2}{\|A\|_2} \leq \|x\|_2$$

$\square$

30 (i) Positivity:
Since $||A||_S = ||SAS^{-1}|| \geqslant 0$ and $||A||_S = ||SAS^{-1}|| = 0$ if and only if $SAS^{-1} = 0$.
(ii) Scalar Preservation:
$||kA||_S = ||SkAS^{-1}|| = ||kSAS^{-1}|| = k||SAS^{-1}|| = k||A||_S$
(iii) Triangle Inequality:
$||(A+B)||_S = ||S(A+B)S^{-1}|| = ||SAS^{-1} + SBS^{-1}|| \leqslant ||||SAS^{-1}| + ||SBS^{-1}|| = ||A||_S + ||B||_S$
(iv) Submultiplicative:
$||AB||_S = ||SABS^{-1}|| = ||SAS^{-1}SBS^{-1}|| \leqslant ||SAS^{-1}|| \cdot ||SBS^{-1}|| = ||A||_S \cdot ||B||_S$
Therefore, we have shown that $|| \cdot ||_S$ is a matrix norm,

37 We first find a set of orthonormal basis for V.
Let $p_1 = 1$, $q_1 = \frac{p_1}{\|p_1\|} = \frac{1}{\int_0^1 1dx} = 1$.
let $p_2 = x - proj_1 x = x - \frac{1}{2}$, $q_2 == \frac{p_2}{\|p_2\|} = \sqrt{12}(x - \frac{1}{2})$.
Let $p_3 = x^2 - proj_1 x^2 - proj_{x-\frac{1}{2}}x^2 = x^2 - x + \frac{1}{6}$, $q_3 = \frac{p_3}{\|p_3\|} = \sqrt{180}(x^2 - x + \frac{1}{6})$.
Then, $q = \sum_{i=1}^{3} L(q_i)q_i = 0 + 12(x - \frac{1}{2}) + 180(x^2 - x + \frac{1}{6}) = 180x^2 - 168x + 24$.
It can be referred that $\forall p \in V, L[p] = \langle q \cdot p\rangle$

38 Let $\mathcal{B} = \{1, x, x^2\}$, then:

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

Also we have:

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

39 (i)

$\langle (S+T)v, w\rangle = \langle Sv, w\rangle + \langle Tv, w\rangle = \langle v, S^*w\rangle + \langle v, T^*w\rangle = \langle v, (S^*+T^*)w\rangle$

$\langle \alpha T^* v, w\rangle = \alpha\langle Tv, w\rangle = \alpha\langle v, T^*w\rangle = \langle v, \overline{\alpha}T^*w\rangle$

(ii)

$\langle S^*v, w\rangle = \overline{\langle w, S^*v\rangle} = \overline{\langle Sw, v\rangle} = \langle v, Sw\rangle$

(iii)

$\langle STv, w\rangle = \langle Tv, S^*w\rangle = \langle v, T^*S^*w\rangle$

(iv)

Since $\langle T^*(T^{-1})^*x, y\rangle = \langle (T^{-1})^*x, Ty\rangle = \langle x, (T^{-1})Ty\rangle = \langle x, y\rangle$ for $\forall x, y$

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

40 (i)

View A as the operator,

since $\langle AB, C\rangle = \text{ tr } (AB)^H C = \text{ tr } B^H A^H C = \langle B, A^H C\rangle \Rightarrow A^* = A^H$

(ii)

$\langle A_2, A_3A_1\rangle = \text{ tr}(A_2^H A_3 A_1) = \text{ tr}(A_1 A_2^H A_3) = \text{ tr}(A_2 A_1^H A_3) = \langle A_2 A_1^*, A_3\rangle$

(iii)

For some $B, C \in \mathbb{M}_n(\mathbb{F})$, we have $\langle B, AC - CA\rangle = \langle B, AC\rangle - \langle B, CA\rangle$.

Applying (ii), we have $\langle B, CA\rangle = \langle BA^*, C\rangle$.

Meanwhile, $\langle B, AC\rangle = \text{tr}(B^H AC) = \text{tr}((A^H B)^H C) = \langle A^H B, C\rangle = \langle A^* B, C\rangle$

$\Rightarrow T_A^* = T_{A^*}$

44 If $b = 0$, then $b \in R(A)$, and $x = 0$ is a solution to $Ax = 0$.
Now if $b \neq 0$, since $\mathbb{F}^n = R(A) + N(A^H)$,
then either $b \in R(A)$ or $b \in N(A^H)$.
If $b \in R(A)$, then $\exists x$ as a solution.
If $b \in N(A^H)$, let $y = b$, since $b \neq 0$, $\quad \langle y, b\rangle = \langle b, b\rangle \neq 0$

45 (i)

First we will show that $\text{Skew}_n(\mathbb{R}) \subset \text{ Sym}_n^\perp(\mathbb{R})$.

Let $A \in Skew_n(\mathbb{R})$ Then, $\forall B \in \text{ Sym}_n(\mathbb{R})$, $\langle A, B\rangle = \text{ tr}(A^H B) = \text{ tr}(-AB) = \text{ tr}(-AB^H) = -\overline{\langle A, B\rangle}$

Also $\langle A, B\rangle = -\overline{\langle A, B\rangle} \implies \langle A, B\rangle = 0$ for all $B \in \text{ Sym}_n(\mathbb{R})$

$\Rightarrow A \in \text{ Sym}_n(\mathbb{R})^\perp$

(ii)

Then we will show that $\text{ Sym}_n^\perp(\mathbb{R}) \subset \text{Skew}_n(\mathbb{R})$.

Let $B \in \text{ Sym}_n(\mathbb{R})^\perp$. Then for $\forall A \in \text{ Sym}_n(\mathbb{R})$,

$\langle B + B^T, A\rangle = \langle B, A\rangle + \langle B^T, A\rangle = 0 + \langle B^T, A\rangle$
and $\langle B^T, A\rangle = \operatorname{tr}(BA) = \operatorname{tr}(BA^T) = \operatorname{tr}(A^T B) = \operatorname{tr}(B^T A) = \langle B, A\rangle = 0$
$\Rightarrow \langle B + B^T, A\rangle = 0$ for all $A \in \operatorname{Sym}_n(\mathbb{R})$
But $B + B^T \in \operatorname{Sym}_n(\mathbb{R}) \Rightarrow ||B + B^T|| = 0 \Rightarrow B + B^T = 0 \Rightarrow B^T = -B$

46

1). $\because x \in N(A^H A), \therefore Ax \in R(A)$.
Since $x \in N(A^H A), A^H Ax = 0 \Rightarrow A^H * (Ax) = 0 \Rightarrow Ax \in N(A^H)$.

2). i). NTS:$N(A^H A) \subset N(A)$.
Pick $x \in N(A^H A)$, then $A^H Ax = 0$.
iF $x = 0$, then $x = 0 \in N(A)$
If $x \neq 0$, NTS: $Ax = 0$
By contradiction, assum $Ax \neq 0$.
Then, $A^H(Ax) = 0$ implies that $Ax \in N(A^H)$
Since $Ax \in R(A)$ and $Ax \neq 0$, this contradicts with the fact that $R(A)^\perp = N(A^H)$.
Hence $Ax = 0$ and $x \in N(A)$.
Therefore, $N(A^H A) \subset N(A)$
ii). NTS:$N(A) \subset N(A^H A)$.
Pick $x \in N(A)$, then $Ax = 0$. It follows that $A^H Ax = A^H(Ax) = A^H \cdot 0 = 0$.
Hence, $x \in N(A^H A)$ and $N(A^H A) = N(A)$.
$\Rightarrow N(A) = N(A^H A)$

3). Observe that both A and $A^H A$ are both map to the n-dimensional spaces.
By rank-nullity, $dim(V) = rank(L) + dim(N(L))$, where $L : V \to W$.
Since $N(A^H A) = N(A)$ by 2)., we have $dim(N(A^H A)) = dim(N(A))$
It follows that $rank(A^H A) = dim(\mathbb{R}^n) - dim(N(A^H A)) = dim(\mathbb{R}^n) - Dim(N(A)) = rank(A)$

4). Since $A \in M_{m*n}(\mathbb{R}), A^T A \in M_{m*n}(\mathbb{R})$
If A has linearly independent columns, then $rank(A) = n$
Since $A^T A$ is an n by n matrix, it is non-singular.

47

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

ii). lemma: $(A^{-1})^H = (A^T)^{-1}$
proof of lemma:
$(A^{-1} A^H = (AA^{-1})^H) = I^H = I$

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

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

iii. Since we know that rank will not increase in matrix multiplication, we can infer that $rank(p) \leq rank(A) = n$.
Now, $\forall y \in R(A), \exists x \quad s.t \quad Ax = y$.
Observe that $p_y = A(A^HA)^{-1}A^Hy = Ax = y$,
$\Rightarrow y \in R(p)$
It follows that $R(A) \subset R(p)$, so $n = rank(A) \leq rank(p)$
We can now conclude that $rank(p) = n$.

48 (i)
let $\alpha \in \mathbb{R}, A, B \in M_n(\mathbb{R})$, then we have:

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

(ii)
$P^2(A) = \frac{P(A)+P(A)^T}{2} = \frac{\frac{A+A^T}{2}+\frac{A+A^T}{2}}{2} = \frac{A+A^T}{2} = P(A)$
(iii)
$\langle P(A), B\rangle = \operatorname{tr}(P(A)^TB) = \operatorname{tr}(\frac{A+A^T}{2} \cdot B) = \frac{\operatorname{tr}(A^TB+AB)}{2} = \operatorname{tr}(AB) = \frac{\operatorname{tr}(AB+AB^T)}{2} = \operatorname{tr}(A \cdot \frac{B+B^T}{2}) = \operatorname{tr}(AP(B)) = \langle A, P(B)\rangle$
(iv)
$A \in \operatorname{Ker}(P) \iff P(A) = 0 \iff A + A^T = 0 \iff A = -A^T \iff A \in \operatorname{Skew}_n(\mathbb{R})$
(v)

$$\begin{aligned} A \in \text{ Range}(P) &\iff \exists B : A = P(B) \\ &\iff \exists B : B + B^T = 2A \\ &\iff A \in \text{ Sym}_n(\mathbb{R}) \end{aligned}$$

(vi)

$$\begin{aligned} &||A - P(A)||_F^2 = \langle A - P(A), A - P(A)\rangle = \langle A - \frac{A + A^T}{2}, A - \frac{A + A^T}{2}\rangle \\ =&\langle \frac{A - A^T}{2}, \frac{A - A^T}{2}\rangle = \text{Tr}\left(\left(\frac{A - A^T}{2}\right)^T \frac{A - A^T}{2}\right) \\ =&\text{Tr}\left(\frac{A^T - A}{2}\frac{A - A^T}{2}\right) = \text{Tr}\left(\frac{A^TA - A^2 - (A^T)^2 + AA^T}{4}\right) \\ =&\text{Tr}\left(\frac{A^TA - A^2 - A^2 + A^TA}{4}\right) = \text{Tr}\left(\frac{A^TA - A^2}{2}\right) = \frac{\text{Tr}(A^TA) - \text{Tr}(A^2)}{2}. \end{aligned}$$

50 Let

$$A = \begin{bmatrix} x_1^2 & y_1^2 \\ . & . \\ . & . \\ . & . \\ x_n^2 & y_n^2 \end{bmatrix}, x = \begin{bmatrix} r \\ s \end{bmatrix}, b = \begin{bmatrix} 1 \\ . \\ . \\ . \\ 1 \end{bmatrix}$$

Then the normal equation to solve is:
$AA^Tx = A^Tb$