Maths Problem Set 2

Shekhar Kumar[*,†]

July 01 , 2018

Ex 3.1 **(i)** $\langle x,y\rangle = \frac{1}{4}(||x+y||^2 - ||x-y||^2)$
$= \frac{1}{4}(\langle x+y, x+y\rangle - \langle x-y, x-y\rangle)$
We note we are on a real inner product space so we can write:
$= \frac{1}{4}(\langle x,x\rangle + \langle y,y\rangle + 2\langle x,y\rangle - \langle x,x\rangle - \langle y,y\rangle + 2\langle x,y\rangle)$
$= \frac{1}{4}(4\langle x,y\rangle)$
$= \langle x,y\rangle$

(ii)
$||x||^2 + ||y||^2 = \frac{1}{2}(||x+y||^2 + ||x-y||^2)$
Again because we in a real space we can write:
$= \frac{1}{2}(\langle x,x\rangle + \langle y,y\rangle + 2\langle y,x\rangle + \langle x,x\rangle + \langle y,y\rangle - 2\langle y,x\rangle)$
$= \langle x,x\rangle + \langle y,y\rangle$
$= ||x||^2 + ||y||^2$

Ex 3.2
$\langle x,y\rangle = \frac{1}{4}(||x+y||^2 - ||x-y||^2 + i||x-iy||^2 - i||x+iy||^2)$
Using the proof from above we can write this as:
$= \mathcal{R}\langle x,y\rangle + \frac{1}{4}i(\langle x-iy, x-iy\rangle - \langle x+iy, x+iy\rangle)$
$= \mathcal{R}\langle x,y\rangle + \frac{1}{4}4(\mathcal{I}\langle x,y\rangle)$
$= \langle x,y\rangle$

Ex 3.3

(i)
$cos(\theta) = \frac{\langle x,y\rangle}{||x||||y||}$
Subbing in we have:
$= \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx}\sqrt{\int_0^1 x^10 dx}}$

*University of Chicago, OSM Boot Camp 2018, +91-801-702-544, shekharkhetan@gmail.com.
†I thank Yung-Hsu Tsui for sharing the basic latex file as well as Navneeraj and Keshav for helping understand the problems and Matt for his generous help.

$= \frac{1/7}{\sqrt{1/33}}$

Therefore the angle is 34.84 degrees.

(ii)

$cos(\theta) = \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^4 dx}\sqrt{\int_0^1 x^8 dx}}$

$= \frac{1/7}{\sqrt{1/45}}$

Therefore the angle is 16.6 degrees.

Ex 3.8

(i)

A set is orthonormal if the inner products of the combinations of elements of the set satisfy:

- $\langle x_i, x_j \rangle = 1$ if $i = j$
- $\langle x_i, x_j \rangle = 0$ if $i \neq j$

Checking the first condition:

Firstly for cos(t), cos(t)

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

$= \frac{1}{\pi}[\frac{x}{2} + \frac{sin(2x)}{4}]_{-\pi}^{\pi}$

$= \frac{1}{\pi}(\pi)$

$= 1$

We can also see that this result will hold for cos(2t), cos(2t) as well. (The evaluated sin functions in the integral will still be zero).

Now checking sin(t), sin(t), and by virtue of the argument above, sin(2t), sin(2t) as well.

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

$= \frac{1}{\pi}[\frac{x}{2} - \frac{1}{4} sin(2x)]_{-\pi}^{\pi}$

$= 1$

Now we need to check the cross terms, and verify that their inner product is zero.

$\langle cos(t), sin(t) \rangle = \frac{1}{\pi}[sin(t)^2]_{-\pi}^{\pi}$

$= 0$

And we note that this also holds for the combinations of cos(2t), sin(t) and also cos(t), sin(2t).

$\langle cos(t), cos(2t) \rangle = \frac{1}{\pi}[\frac{sin(t)}{2} + \frac{sin(3t)}{6}]_{-\pi}^{\pi}$

$= 0$

$\langle sin(t), sin(2t)\rangle = \frac{1}{\pi}[\frac{sin(t)^3}{1.5}]_{-\pi}^{\pi}$

$= 0$

Therefore the set is orthonormal.

(ii)

$||t||^2 = \frac{1}{\pi}\int_{-\pi}^{\pi} t^2 dt$

$= [\frac{t^3}{3}]_{-\pi}^{\pi}$

$= [\frac{\pi^3}{3} + \frac{\pi^3}{3}]$

$= 2\frac{\pi^3}{3}$

Therefore $||t|| = (\frac{2\pi^3}{3})^{0.5}$

(iii)

Because we are dealing with an orthnormal set we can write:

$Proj_x(cos(3t)) = \Sigma_i \langle S_i, cos3t\rangle s_i$

$= \langle cos(t), cos(3t)\rangle cos(t) + \langle cos(2t), cos(3t)\rangle cos(2t) + \langle sin(t), cost(3t)\rangle sin(t) + \langle sin(2t), cos(3t)\rangle sin(2t)$

After substituting in the integrals we get

$= 0$

i.e. cos(3t) is orthogonal to all the elements in S, as its projection matrix is a zero matrix.

(iv)

$Proj_x(t) = \Sigma_i \langle S_i, t\rangle s_i$

$= \langle cos(t), t\rangle cos(t) + \langle cos(2t), t\rangle cos(2t) + \langle sin(t), t\rangle sin(t) + \langle sin(2t), t\rangle sin(2t)$

$= 0 + 0 + 2sin(t) - sin(2t)$

$= 2sin(t) - sin(2t)$

Ex 3.9

We use the fact that we can convert the rotation transformation into a matrix in the standard basis, which we call Q. Then, we know that if $Q^TQ = I$ then the transformation is orthonormal.

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

So,

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

$$QQ^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Ex 3.10

(i)

First we show that if Q is orthonormal then $QQ^H = I$.

If Q is an orthonormal matrix, then it preserves the inner product of two vectors. i.e.

$\langle m, n \rangle = \langle Qm, Qn \rangle$

Which we can rewrite as:

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

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

Therefore, since this has to hold for all m and n:

$Q^h Q = I$

Now we can show that if $QQ^H = I$, then Q is orthonormal.

If $QQ^H = I$

Then:

$\langle Qm, Qn \rangle = (Qm)^H (Qn)$

$= m^H Q^H Q n$

$= \langle m, n \rangle$

(ii)

$||Qx|| = \sqrt{\langle Qx, Qx \rangle}$

By the definition of what a orthonrmal matrix is (it preserves the inner product), we can write:

$= \sqrt{\langle x, x \rangle}$

$= ||x||$

(iii)

If Q is orthonormal we can write:

$QQ^H = I$

i.e. $Q^H = Q^{-1}$

Q^H is clearly orthonormal because $(Q^H)^H = Q$, therefore so is Q^{-1}.

(iv)

If Q is orthonormal we know that $G = Q^H Q = I$

For some element of G, we can write that:

$G_{i,j} = \langle q_i, q_j \rangle$

Where q_i is the i'th column of Q.

By the definition of orthornomality, we know that:

$\langle q_i, q_j \rangle = 1$ if $i = j$

and

$\langle q_i, q_j \rangle = 0$ if $i \neq j$

So we can see that when $i = j$ we are on the diagonal of Q, so clearly $\langle q_i, q_j \rangle = 1$ if $i = j$. And similarly, everywhere else $i \neq j$, and have zero entries, so $\langle q_i, q_j \rangle = 0$ if $i \neq j$.

(v)

We can find a counterexample to show that not all matrices with determinant equal to 1 are orthonormal.

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix}$$

We can see that: $det(D) = 1$ But, if we test for orthonormality,

$$DS^H = \begin{bmatrix} 4 & 0 \\ 0 & 0.25 \end{bmatrix} \neq I$$

(vi)

Checking if the product of the two matrices is an orthonormal matrix:

$(Q_1Q_2)(Q_1Q_2)^H = Q_1Q_2Q_2^HQ_1^H$

Then using the fact that Q_1 and Q_2 are orthonormal we can write:

$Q_1Q_2Q_2^HQ_1^H = Q_1Q_1^H = I$

So the product of the matrices is orthonormal.

Ex 3.11

Proof. Suppose, WLOG, that for $2 < k < N$, $\{x_i\}_{i=1}^{k-1}$, linearly independent. That also means that $\{q_i\}_{i=1}^{k-1}$ are linearly independent(Showing this is very trivial!)

However, if $\{x_i\}_{i=1}^{k}$ are linearly dependent, then $x_k \in Span(\{x_i\}_{i=1}^{k-1})$, and $q_k = 0$. This is contradictory to the assumption that $\{q_1, ..., q_N\}$ are linearly dependent.

Ex 3.16

(i) Let $A \in \mathbb{M}_{mxn}$ where rank$(A) = n \leq m$. Then there exist orthonormal $Q \in \mathbb{M}_{mxm}$ and upper triangular $R \in \mathbb{M}_{mxn}$ such that $A = QR$. Since $\tilde{Q} = -Q$ is still orthonormal $(-Q(-Q)^H = -Q(-Q^H) = QQ^H = I$ and similarly one shows $(-Q)^H(-Q) = I)$ and $\tilde{R} = -R$ is still upper triangular, $A = QR = \tilde{Q}\tilde{R}$. Therefore QR-decomposition is not unique.

(ii) Now take a reduced QR-decomposition $A = \hat{Q}\hat{R}$, where $\hat{Q} \in \mathbb{M}_{mxn}$ is orthonormal and $\hat{R} \in \mathbb{M}_{nxn}$ is upper triangular. Since A has full column rank, $\hat{R}$ has full rank and is therefore nonsingular. Then,

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

and premultiplying both LHS and RHS of the last equation by $\hat{R}^{-1}$ gives $\hat{R}x = \hat{Q}^H b$.

Ex 3.17

Proof. $A^H AX = A^H b \leftrightarrow (\hat{Q}\hat{R})^H(\hat{Q}\hat{R})x = (\hat{Q}\hat{R})^H b \leftrightarrow \hat{R}^H(\hat{Q}^H\hat{Q})\hat{R}x = \hat{R}^H\hat{Q}^H b$
$\leftrightarrow \hat{R}^H\hat{R}x = \hat{R}^H\hat{Q}^H b$
Note that $\hat{R}^H$ is invertible. Thus, $\hat{R}^H x = \hat{Q}^H b$. You can proceed in the other way around to complete your proof. Q.E.D

Ex 3.23

Proof. Let $z := x - y$. Then, by triangle inequality, the following is satisfied:

$$\|z+y\| \leq \|z\| + \|y\| \leftrightarrow \|x\| \leq \|x-y\| + \|y\| \leftrightarrow \|x\| - \|y\| \leq \|x-y\|$$

In the same way, $\|y\| - \|x\| \leq \|x-y\|$. Thus, the statement holds. Q.E.D

Ex 3.24

(i)

$\|f\|_{L^1} = \int_a^b |f(t)|dt > 0$ for $f(t) \neq 0$

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

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

(ii)

$\|f\|_{L^2} = (\int_a^b |f(t)|^2 dt)^{1/2} > 0$ for $f(t) \neq 0$

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

$\|f+g\|_{L^2} = (\int_a^b |f(t)+g(t)|^2 dt)^{1/2} \leq (\int_a^b |f(t)|^2 + |g(t)|^2 dt)^{1/2} \leq (\|f\|^2 + \|g\|^2)^{1/2} \leq \sqrt{\|f\|^2} + \sqrt{\|g\|^2} = \|f\| + \|g\|$

(iii)

$\|f\|_{L^\infty} = \sup_{x\in[a,b]} |f(x)| > 0$ for $f(t) \neq 0$

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

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

Ex 3.26

Proof. Let $a \sim b$ if $\exists m, M > 0$ and $m \leq M$ for vector space X s.t $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a$ $\forall x \in X$

1. Reflexivity

If $m = M = 1$, then $m\|x\|_a \leq \|x\|_a \leq M\|x\|_a$

2. Symmetry

Suppose $a \sim b$. Then, $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a$. This leads to the following inequalities; $\frac{1}{M}\|x\|_b \leq \|x\|_a \leq \frac{1}{m}\|x\|_b$. So, as long as $m = M$, the symmetry property can be satisfied!

3. Transitivity

Suppose $a \sim b$ and $b \sim c$

Then, $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a$ and $m^*\|x\|_b \leq \|x\|_c \leq M^*\|x\|_b$

Then, this leads to: $m\|x\|_a \leq \|x\|_b \leq \frac{1}{m^*}\|x\|_c \leq \frac{M}{m^*}\|x\|_b \leq \frac{M^2}{m^*}\|x\|_a \to m\|x\|_a \leq \frac{1}{m^*}\|x\|_c \leq \frac{M^2}{m^*}\|x\|_a$

Thus, $a \sim c$

Take $x \in \mathbb{R}^n$ Notice that

$$\sum_{i=1}^n |x_i|^2 \leq \left(\sum_{i=1}^n |x_i|^2 + 2\sum_{i\neq j}|x_i||x_j|\right) = \left(\sum_{i=1}^n |x_i|\right)^2$$

and that

$$\sum_{i=1}^n |x_i|\cdot 1 \leq \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}\left(\sum_{i=1}^n 1^2\right)^{1/2} = \sqrt{n}\left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$$

prove that $||x||_2 \leq ||x||_1 \leq \sqrt{n}||x||_2$.

Also notice that

$$\max_i |x_i| = \left(\max_i |x_i|^2\right)^{1/2} \leq \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} =$$

and

$$\sum_{i=1}^n |x_i|^2 \leq n \cdot \max_i |x_i|^2$$

prove that $||x||_\infty \leq ||x||_2 \leq \sqrt{n}||x||_\infty$.

Ex 3.28

(i) Notice that (applying the results of the previous exercise)

$$\sup_{x\neq 0} \frac{||Ax||_1}{||x||_1} \leq \sup_{x\neq 0} \frac{||Ax||_1}{||x||_1} \leq \sqrt{n} \sup_{x\neq 0} \frac{||Ax||_2}{||x||_2},$$

and

$$\sup_{x\neq 0} \frac{||Ax||_1}{||x||_1} \geq \sup_{x\neq 0} \frac{||Ax||_2}{||x||_1} \geq \frac{1}{\sqrt{n}} \sup_{x\neq 0} \frac{||Ax||_2}{||x||_2}$$

imply that $\frac{1}{\sqrt{n}}||A||_2 \leq ||A||_1 \leq ||A||_2$.

(ii) Notice that

$$\sup_{x\neq 0} \frac{||Ax||_2}{||x||_2} \leq \sup_{x\neq 0} \frac{\sqrt{n}||Ax||_\infty}{||x||_\infty},$$

and

$$\sup_{x\neq 0} \frac{||Ax||_2}{||x||_2} \geq \sup_{x\neq 0} \frac{||Ax||_\infty}{\sqrt{n}||x||_\infty}.$$

Ex 3.29

Pf of 1. $\|Q\|_p := \sup_{x\neq 0} \frac{\|Qx\|_p}{\|x\|_p} = \sup_{x\neq 0} \frac{\|x\|_p}{\|x\|_p} = \sup 1 = 1$(By orthonormal transformation).

Pf of 2.

Now let $R_x : \mathbb{M}_n(\mathbb{F}) \to \mathbb{F}^n, A \mapsto Ax$ for every $x \in \mathbb{F}^n$.

$$||R_x|| = \sup_{x\neq 0} \frac{||Ax||}{||A||} = \sup_{x\neq 0} \frac{||Ax||||x||}{||A||||x||}$$

Ex 3.30

(i) $\|AB\|_S = \|SABS^{-1}\| = \|SAS^{-1}SBS^{-1}\| \leq \|SAS^{-1}\|\|SBS^{-1}\| = \|A\|_S\|B\|_S$
(ii) $\|A\|_S\|x\| = \|SAS^{-1}\|\|x\| = (\sup_{x\neq 0} \frac{\|SAS^{-1}x\|}{\|x\|})\|x\| \geq \|S\|\|A\|\|S^{-1}\|\|x\| = \|A\|\|x\| \geq \|Ax\|$ Other basic properties of norm can be proved in the exactly same way as vector norm! Q.E.D

Ex 3.37

Answer Note that according to the Riesz Representation theorem, $L[q] = \langle q, q\rangle = \int_0^1 q^2(x)dx = q^{'}(1)$. Let $q(x) := a + bx + cx^2$. Then,

$$L[1] = 0 = \langle q, 1\rangle = \int_0^1 q(x)dx = a + \frac{1}{2}b + \frac{1}{3}c$$

$$L[x] = 1 = \langle q, x\rangle = \int_0^1 xq(x)dx = \frac{1}{2}a + \frac{1}{3}b + \frac{1}{4}c$$

$$L[x^2] = 2 = \langle q, x^2\rangle = \int_0^1 x^2q(x)dx = \frac{1}{3}a + \frac{1}{4}b + \frac{1}{5}c$$

If we solve this 3 equations with three unknowns(a, b, c), then $q(x) = 24 - 168x + 180x^2$

Ex 3.38

The matrix representation of differential operator is

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

The adjoint matrix is the following;

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

Ex 3.39

(i) By definition of adjoint and linearity of inner products,

$$\begin{aligned} &< (S+T)^* w, v >_V =< w, (S+T)v >_W = \\ &< w, Sv + Tv >_W =< w, Sv >_W + < w, Tv >_W = \\ &< S^* w, v >_V + < T^* w, v >_V =< S^* w + T^* w, v >_V . \end{aligned}$$

Then $(S+T)^* = S^* + T^*$. Also,

$$\begin{aligned} &< (\alpha T)^* w, v >_V =< w, (\alpha T)v >_W = \\ &< w, \alpha T v >_W = \alpha < w, Tv >= \\ &\alpha < T^* w, v >=< \bar{\alpha} T^* w, v > \end{aligned}$$

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

(ii) By the definition of adjoint of S and the properties of inner products we have that

$$< w, Sv >_W =< S^* w, v >_V = \overline{< v, S^* w >_V} = \overline{< S^{**} v, w >_W} =< w, S^{**} v >_W$$

for all $v \in V$ and $w \in W$. Therefore $S = S**$.

(iii) By the definition of adjoint we have

$$< (ST)^*v', v >_V = < v', (ST)v >_V = < v', S(Tv) >_V =$$
$$< S^*v', Tv >_V = < T * S * v', v >_V$$

thereby proving that $(ST)^* = T^*S^*$.

(iv) Using (iii) we have $T^*(T^*)^{-1} = (TT^{-1})^* = I^* = I$.

Ex 3.40

(i) Let $B, C \in \mathbb{M}_n(\mathbb{F})$. By definition of Frobenious inner product

$$< B, AC >_F = \mathrm{tr}(B^H AC) = \mathrm{tr}((A^H B)^H C) = < A^H B, C >_F .$$

(ii) By definition of Frobenious norm and the properties of the trace we have

$$< A_2, A_3A_1 >_F = \mathrm{tr}(A_2^H A_3 A_1) = \mathrm{tr}(A_1 A_2^H A_3) = \mathrm{tr}((A_2 A_1^H)^H A_3) = < A_2 A_1^H, A_3 >_F = < A_2 A_1^*, A_3 > .$$

(iii) Given $B, C \in \mathbb{M}_n(\mathbb{F})$, we have $< B, AC - CA > = < B, AC > - < B, CA >$. Applying (ii) to the second term we get $< B, CA > = < BA^*, C >$. On the other hand,

$$< B, AC > = \mathrm{tr}(B^H AC) = \mathrm{tr}((A^H B)^H C) = < A^H B, C > = < A^* B, C > .$$

Putting all together we obtain that $T_A^* = T_{A^*}$.

Ex 3.44

Suppose there exists an $x \in \mathbb{F}^n$ such that $Ax = b$. Then, for every $y \in \mathcal{N}(A^H)$,

$$< y, b > = < y, Ax > = < A^H y, x > = < 0, x > = 0.$$

Now suppose that there exists a $y \in \mathcal{N}(A^H)$ such that $< y, b > \neq 0$. Then $b \notin \mathcal{N}(A^H)^\perp = \mathcal{R}(A)$. Therefore for no $x \in \mathbb{F}^n$, $Ax = b$.

Ex 3.45

Let $A \in \mathrm{Sym}_n(\mathbb{R})$ and $B \in \mathrm{Skew}_n(\mathbb{R})$. Then

$$< B, A >= \mathrm{Tr}(B^T A) = \mathrm{Tr}(AB^T) = \mathrm{Tr}(A^T(-B)) = - < A, B > .$$

We conclude that $< A, B >= 0$ and $\mathrm{Skew}_n(\mathbb{R}) \subset \mathrm{Sym}_n(\mathbb{R})^{\perp}$. Now suppose $B \in \mathrm{Sym}_n(\mathbb{R})^{\perp}$. As for any other matrix, $B + B^T \in \mathrm{Sym}_n(\mathbb{R})$. Thus,

$$0 =< B + B^T, B >= \mathrm{Tr}((B + B^T)B) = \mathrm{Tr}(BB + B^T B) = \mathrm{Tr}(BB) + \mathrm{Tr}(B^T B),$$

which implies $< B^T, B >=< -B, B >$ and so $B^T = -B$. Therefore $\mathrm{Sym}_n(\mathbb{R})^{\perp} = \mathrm{Skew}_n(\mathbb{R})$.

Ex 3.46

(i) if $x \in \mathcal{N}(A^H A)$, $0 = (A^H A)x = A^H(Ax)$ and $Ax \in \mathcal{N}(A^H)$. Also, Ax is in the range of A by definition.

(ii) Suppose $x \in \mathcal{N}(A)$. Then $Ax = 0$. Premultiplying by A^H both sides of the equation we obtain $A^H Ax = A^H 0 = 0$ and so $x \in \mathcal{N}(A^H A)$. On the other hand, suppose $x \in \mathcal{N}(A^H A)$. Then $||Ax|| = x^H A^H Ax = x^H 0 = 0$, so that $Ax = 0$ and $x \in \mathcal{N}(A)$

(iii) By the rank-nullity theorem we have $n = \mathrm{Rank}(A) + \mathrm{Dim}\mathcal{N}(A)$ and $n = \mathrm{Rank}(A^H A) + \mathrm{Dim}\mathcal{N}(A^H A)$. Then by (ii) it follows that $\mathrm{Rank}(A) = \mathrm{Rank}(A^H A)$.

(iv) By (iii) and the assumption on A we have that $n = \mathrm{Rank}(A) = \mathrm{Rank}(A^H A)$. Since $A^H A \in \mathbb{M}_n$, it is nonsingular.

Ex 3.47

(i) Notice that

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

(ii) Notice that

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

(iii) A has rank n, therefore P has at most rank n. Take y in the range of A. Then there exists an $x \in \mathbb{F}^n$ such that $y = Ax$. Then

$$Py = A(A^HA)A^Hy = A(A^HA)^{-1}A^HAx = Ax = y$$

shows that y is also in the range of P. Therefore $\text{Rank}(P) \geq \text{Rank}(A)$ and so P has rank p

Ex 3.48

(i) Let $A, B \in \mathbb{M}_n(\mathbb{R})$ and $x \in \mathbb{R}$. Then

$$P(A + xB) = \frac{(A + xB) + (A + xB)^T}{2} = \frac{A + A^T + x(B + B^T)}{2} = P(A) + xP(B).$$

Thus P is a linear transformation.

(ii) Now notice that

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

(iii) By definition of adjoint we have $< P^*(A), B >=< A, P(B) >$. Then, notice that

$$\begin{aligned}
&< A, P(B) >=< A, (B + B^T)/2 >=< A, B/2 > + < A, B^T/2 >= \\
&\text{Tr}(A^TB/2) + \text{Tr}(A^TB^T/2) = \text{Tr}(A^T/2B) + \text{Tr}(BA/2) = \\
&\text{Tr}(A^T/2B) + \text{Tr}(A/2B) =< (A + A^T)/2, B >=< P(A), B > .
\end{aligned}$$

Thus $P = P^*$.

(iv) Suppose $A \in \mathcal{N}(P)$. Then $0 = P(A) = (A + A^T)/2$ implies $A^T = -A$, thus $\mathcal{N}(P) \subset \text{Skew}(\mathbb{R})$. Now suppose $A \in \text{Skew}(\mathbb{R})$. Then $A^T = -A$ and so $P(A) = (A + A^T)/2 = 0$. Thus $\text{Skew}(\mathbb{R}) \subset \mathcal{N}(P)$.

(v) Let $A \in \mathbb{M}_n(\mathbb{R})$. Then $P(A) = (A + A^T)/2 = (A^T + A)/2 = P(A)^T$ and so $\mathcal{R}(P) = \text{Sym}(\mathbb{R})$. Now let $A = \text{Sym}(\mathbb{R})$. Thus $A = A^T$ and $P(A) = (A + A^T)/2 = (A + A)/2 = A$ and so $A \in \mathcal{R}(P)$. This shows that $\mathcal{R}(P) = \text{Sym}(\mathbb{R})$.

(vi) Notice that

$$||A - P(A)||_F^2 =< A - P(A), A - P(A) >=< A - \frac{A + A^T}{2}, A - \frac{A + A^T}{2} >=$$
$$< \frac{A - A^T}{2}, \frac{A - A^T}{2} >= \text{Tr}\left(\left(\frac{A - A^T}{2}\right)^T \frac{A - A^T}{2}\right) =$$
$$\text{Tr}\left(\frac{A^T - A}{2}\frac{A - A^T}{2}\right) = \text{Tr}\left(\frac{A^TA - A^2 - (A^T)^2 + AA^T}{4}\right) =$$
$$\text{Tr}\left(\frac{A^TA - A^2 - A^2 + A^TA}{4}\right) = \text{Tr}\left(\frac{A^TA - A^2}{2}\right) = \frac{\text{Tr}(A^TA) - \text{Tr}(A^2)}{2}.$$

Therefore $||A - P(A)||_F = \sqrt{\frac{\text{Tr}(A^TA) - \text{Tr}(A^2)}{2}}$.

Ex 3.50

We want to estimate $y^2 = 1/s + rx^2/s$ via OLS. We rewrite the model in the form $Ax = b$ where $b_i = y_i^2$, $A_i = (1\ x_i)$ and $x = (\beta_1\ \beta_2)^T$ where $\beta_1 = 1/s$ and $\beta_2 = r/s$. Then the normal equations are $A^HA\hat{x} = A^Hb$, where

$$A^HA\hat{x} = \begin{bmatrix} \sum_i 1 & \sum_i x_i^2 \\ \sum_i x_i^2 & \sum_i x_i^4 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} n\hat{\beta}_1 - \hat{\beta}_2 \sum_i x_i^2 \\ \hat{\beta}_1 \sum_i x_i^2 - \hat{\beta}_2 \sum_i x_i^4 \end{bmatrix}$$

and

$$A^Hb = \begin{bmatrix} \sum_i y_i^2 \\ \sum_i x_i^2 y_i^2 \end{bmatrix}.$$