Problem Set #2

Boot Camp 2018

Name: Zunda Xu

Email: zunda@uchicago.edu

Exercise 3.1.

Proof:

(i)

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

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

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

$= \frac{1}{4}(4\langle x,y\rangle)$

$= \langle x,y\rangle$

(ii)

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

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

$= \langle x,x\rangle + \langle y,y\rangle$

$= ||x||^2 + ||y||^2$

Exercise 3.2.

Proof:

$\frac{1}{4}(||x+y||^2 - ||x-y||^2 + i||x-iy||^2 - i||x+iy||^2)$

$= \langle x,y\rangle + \frac{1}{4}(i\langle x-iy, x-iy\rangle - i\langle x+iy, x+iy\rangle)$

(The first item of the equation came from proof 3.1(i))

$= \langle x,y\rangle + \frac{1}{4}(i(\langle x,x\rangle + \langle x,-iy\rangle + \langle -iy,x\rangle + \langle -iy,-iy\rangle) - i(\langle x,x\rangle + \langle x,iy\rangle + \langle iy,x\rangle +$

$\langle iy,iy\rangle))$

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

$= \langle x, y \rangle$

Exercise 3.3.

Proof:

(i)

$cos(\theta) = \frac{\langle f,g \rangle}{||f||||g||}$

$= \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx}\sqrt{\int_0^1 x^10 dx}}$

$= \frac{1/7}{\sqrt{(1/3)\cdot(1/11)}} = \frac{\sqrt{33}}{7} \Rightarrow$ The angle is around 35 degrees.

(ii) $cos(\theta) = \frac{\langle f,g \rangle}{||f||||g||}$

$= \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^4 dx}\sqrt{\int_0^1 x^8 dx}}$

$= \frac{1/7}{\sqrt{(1/5)\cdot(1/9)}} = \frac{\sqrt{45}}{7} \Rightarrow$ The angle is around 17 degrees.

Exercise 3.8.

Proof:

(i)

A set is orthonormal if the inner products of the combinations of elements of the set satisfy: (1) $\langle x_i, x_j \rangle = 1$, if $i = j$; (2) $\langle x_i, x_j \rangle = 0$, if $i \neq j$.

We first check the first condition:

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

$\langle \sin(t), \sin(t) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(t) dt = \frac{1}{\pi} \frac{-\sin(2x)+2x}{4} \Big|_{-\pi}^{\pi} = 1$

$\langle \cos(2t), \cos(2t) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(2t) dt = \frac{1}{\pi} \frac{\sin(4x)+4x}{8} \Big|_{-\pi}^{\pi} = 1$

$\langle \sin(t), \sin(t) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(2t) dt = \frac{1}{\pi} \frac{-\sin(4x)+4x}{8} \Big|_{-\pi}^{\pi} = 1$

Then we check the second condition:

$\langle \cos(t), \sin(t) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(t)\sin(t) dt = \frac{\sin^2(x)}{2} \Big|_{-\pi}^{\pi} = \sin^2(\pi) - \sin^2(-\pi) = 0$

We can also prove other cross terms similarly.

Hence, we have shown that the set is orthonormal.

(ii)

$|t||^2 = \int_{-\pi}^{\pi} t^2 dt = \frac{t^3}{3}\Big|_{-\pi}^{\pi} = \frac{\pi^3}{3} - \frac{(-\pi)^3}{3} = \frac{2\pi^3}{3} \Rightarrow ||t|| = \left(\frac{2\pi^3}{3}\right)^{\frac{1}{2}}$

(iii)

$Proj_X(\cos(3t)) = \langle \sin(t), \cos(3t)\rangle \sin(t) + \langle \cos(t), \cos(3t)\rangle \cos(t) + \langle \sin(2t), \cos(3t)\rangle \sin(2t) + \langle \cos(2t), \cos(3t)\rangle \cos(2t) = 0 + 0 + 0 + 0 = 0$

(iv)

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

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

Exercise 3.9.

Proof:

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.

Exercise 3.10.

Proof:

(i)

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

If Q is an orthonormal matrix, then for $\forall$ two vectors m, n, we have:

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

$\Rightarrow m^H n = (Qm)^H(Qn) = m^H(Q^H Q)n$

$\Rightarrow Q^H Q = I$

Next we will show that if $QQ^H = I \Rightarrow Q$ is orthonormal.

If $QQ^H = I \Rightarrow \langle Qm, Qn\rangle = (Qm)^H(Qn) = m^H Q^H Q n = \langle m, n\rangle$

Hence, we have shown that Q is orthonormal i.f.f $QQ^H = I$.

(ii)

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

(iii)

If Q is orthonormal $\Rightarrow QQ^H = I \Rightarrow Q^H = Q^{-1}$.

Since $(Q^H)^H = Q \Rightarrow Q^H$ is orthonormal $\Rightarrow Q^{-1}$ is also orthonormal.

(iv)

Since Q is orthonormal $\Rightarrow QQ^H = I$

Let A = $QQ^H \Rightarrow A_{i,j} = \langle q_i, q_j\rangle$ (Where q_i is the i'th column of Q).

By the definition of orthonormality $\Rightarrow \langle q_i, q_j\rangle = 1$ if $i = j$ and $\langle q_i, q_j\rangle = 0$ if $i \neq j$.

Therefore, when $i = j$, we are on the diagonal of $Q \Rightarrow \langle q_i, q_j\rangle = 1$.

While for $i \neq j$, $\langle q_i, q_j\rangle = 0 \Rightarrow$ The columns of Q are orthonormal.

(v)

We can find a matrix:

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

We note that $det(A) = 1$.

However,

$$AA^H = \begin{bmatrix} 4 & 0 \\ 0 & 1/4 \end{bmatrix} \neq I$$

Therefore, the converse is not ture.

(vi)

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

So the product of the matrices is orthonormal.

Exercise 3.11.

Proof:

Let $a_1, a_2..., a_n$ is a collection of linearly dependent vectors, then applying Gram-Schmidt orthonormalization process. We will definitely find a a vector a_k which is linearly dependent upon $a_1, ..., a_{k-1}$. Let $A = span(a_1, ..., a_{k-1}) \Rightarrow a_k \in A$ and $p_{k-1} = Proj_A(a_k) = a_k \Rightarrow q_k = 0$. Therefore, by the end, if throwing all zeros, we will get an orthonormal basis $q_1, ...q_m$ of A where $m = \dim A$.

(The above proof refer to Matthew)

Exercise 3.16.

Proof:

(i)

Let $D \in \mathbb{M}_{m\times n}$ where rank$(D) = n \leq m \Rightarrow \exists Q \in \mathbb{M}_{m\times m}$ and upper triangular $R \in \mathbb{M}_{m\times n}$ s.t. $D = QR$. Since $-Q(-Q)^H = -Q(-Q^H) = QQ^H = I \Rightarrow -Q$ is still orthonormal. Also $-R$ is still upper triangular $\Rightarrow A = QR = (-Q)\cdot(-R) \Rightarrow$ QR-decomposition is not unique.

(ii)

Assume A is invertible and can be decomposed into two different QR decompositions: QR and $Q^{'}R^{'}$, and the diagonal entries of R and $R^{'}$ are strictly positive.$\Rightarrow$ Both R and $R^{'}$ are invertible and we conclude that $R^{'-1}R = Q^HQ^{'}$. Since R and $R^{'}$ are upper triangular, so is the LHS of the previous equation. Meanwhile, since Q and $Q^{'}$ are orthonormal, so is the RHS.$\Rightarrow$ $R^{'-1}R = I \Rightarrow R = R^{'}$, and $Q = Q^{'}$.

Exercise 3.17.

Proof:

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

Also since $A = \hat{Q}\hat{R}$, we have:

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

Hence the two systems are equivalent.

Exercise 3.23.

Proof:

(Latex Code from Fiona Fan)

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

Exercise 3.24.

Proof:

(Latex Code from Fiona Fan)

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

Exercise 3.26.

Proof:

(Latex Code from Fiona Fan)

To prove this is an equivalence relationship:

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

2. If $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a, \quad \forall\vec{x}$,

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

so it is symmetric.

3. If $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a, \quad \forall\vec{x}$, and if $n\|x\|_b \leq \|x\|_c \leq N\|x\|_b, \quad \forall\vec{x}$,

then $mn\|x\|_a \leq \|x\|_c \leq MN\|x\|_a \quad \forall\vec{x}$.

Thus it is transitive.

$\Rightarrow$ Thus this is an equivalence relationship.

□

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

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

Thus, $\|\vec{x}\|_1 \geq \|\vec{x}\|_2$

Let $\vec{u} = [sgn(x_1), \ldots, sgn(x_n)]^T$, $\|\vec{x}\|_1 = \sum_{i=1}^n x_i \cdot sgn(x_i) = |\langle\vec{u},\vec{x}\rangle|$.

By Cauchy-Schwarz, $|\langle\vec{u},\vec{x}\rangle| \leq \|\vec{u}\|_2\|\vec{x}\|_2 = \sqrt{n}\|\vec{x}\|_2$

Hence, $\|\vec{x}\|_1 \leq \sqrt{n}\|\vec{x}\|_2$

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, $\|x\|_2 \geq \|x\|_\infty$

Moreover, $\|x\|_2^2 \leq nx_k^2$

Exercise 3.28.

Proof:

(Latex Code from Fiona Fan)

i. From previous exercise, we can get

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

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

which imply that

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

ii. From previous we can get:

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

Exercise 3.29.

Proof:

(i):

Let Q be an orthonormal matrix. Then $||Qx|| = ||x|| \implies sup_{x\neq 0}\frac{||Qx||}{||x||} = ||Q|| = 1$

(ii):

First we will show $||R_x|| < ||x||$.

Since $||R_x|| = sup_{A\neq 0}\frac{||R_x(A)||}{||A||} = sup_{A\neq 0}\frac{||Ax||}{||A||} = sup_{A\neq 0}\frac{||Ax||\cdot||x||}{||A||\cdot||x||}$,

Also we have $||Ax|| \leqslant ||A|| \cdot ||x|| \quad \forall x \in \mathbb{F}^n$

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

Next we will show the equality.

Let $q_1 = e_1$. By using the gram-schmidt algorithm, we can construct an orthonormal basis $q_1, ...q_n$ for $\mathbb{F}^n$. Let Q be the matrix with these basis vectors as its columns. Then Q is an orthonormal matrix.$\Rightarrow ||Q|| = 1$ and $\frac{||Qx||}{||x||} = ||Q|| = 1$ at all nonzero x

$\Rightarrow ||R_x|| = ||x||$

Exercise 3.30.

Proof:

(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.

Exercise 3.38.

Proof:

(Latex Code from Fiona Fan)

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 $\forall p \in V, L[p] = \langle q \cdot p \rangle$

Exercise 3.38.

Proof:

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:

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

Exercise 3.39.

Proof:

(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 $\forall x, y$

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

Exercise 3.40.

Proof:

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

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

Exercise 3.44.

Proof:

(Latex code from Fiona Fan)

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 $b \neq 0$, $\quad \langle y, b\rangle = \langle b, b\rangle \neq 0$

Exercise 3.45.

Proof:

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

$\operatorname{tr}(-AB^H) = -\overline{\langle A, B\rangle}$

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 $A \in \text{Sym}_n(\mathbb{R})$,

$\langle B + B^T, A\rangle = \langle B, A\rangle + \langle B^T, A\rangle = 0 + \langle B^T, A\rangle$

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

But $B + B^T \in \text{Sym}_n(\mathbb{R}) \Rightarrow ||B + B^T|| = 0 \Rightarrow B + B^T = 0 \Rightarrow B^T = -B$

Exercise 3.46.

Proof:

(Latex code from Fiona Fan)

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.

Exercise 3.47.

Proof:

(Latex code from Fiona Fan)

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:

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

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

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

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 $p_y = A(A^H A)^{-1}A^H y = Ax = y$,

$\Rightarrow y \in R(p)$

It follows that $R(A) \subset R(p)$, so $n = rank(A) \leq rank(p)$

We can now conclude that $rank(p) = n$.

Exercise 3.48.

Proof:

(i)

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

$P(\alpha(A+B))$

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

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

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

$= \alpha(P(A)+P(B))$

(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)

$\langle P(A), B\rangle = \operatorname{tr}(P(A)^T B) = \operatorname{tr}(\frac{A+A^T}{2} \cdot B) = \frac{\operatorname{tr}(A^T B + AB)}{2} = \operatorname{tr}(AB) = \frac{\operatorname{tr}(AB + AB^T)}{2} = \operatorname{tr}(A \cdot \frac{B+B^T}{2}) = \operatorname{tr}(AP(B)) = \langle A, P(B)\rangle$

(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)

$A \in \operatorname{Range}(P) \iff \exists B : A = P(B) \iff \exists B : B + B^T = 2A \iff A \in \operatorname{Sym}_n(\mathbb{R})$

(vi)

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

Exercise 3.50.

Proof:

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:

$AA^Tx = A^Tb$