

Problem Set 2- Inner Product Spaces

Shekhar kumar *

2018-07-02

Problem 3.1

In problem 3.1 and 3.2 the inner product spaces are defined on real numbers. Hence

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

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

We note we are on a real inner product space so we can write:

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

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

Problem 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:

$$\begin{aligned} &= \mathcal{R}\langle x, y \rangle + \frac{1}{4}i(\langle x - iy, x - iy \rangle - \langle x + iy, x + iy \rangle) \\ &= \mathcal{R}\langle x, y \rangle + \frac{1}{4}4(\mathcal{I}\langle x, y \rangle) \\ &= \langle x, y \rangle \end{aligned}$$

*I thank Jayhyung Kim for his base Latex file

Problem 3.3

i. $\cos(\theta) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$

Subbing in we have:

$$= \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx} \sqrt{\int_0^1 x^4 dx}}$$

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

Therefore the angle is $= \cos^{-1} \left(\frac{1/7}{\sqrt{1/33}} \right)$ 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 $= \cos^{-1} \left(\frac{1/7}{\sqrt{1/45}} \right)$ degrees.

Problem 3.8

i. A set is orthonormal if the inner products 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 $\cos(t)$, $\cos(t)$

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

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

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

$$= 1$$

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.

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

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

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

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

$$\begin{aligned}\langle \cos(t), \cos(2t) \rangle &= \frac{1}{\pi} \left[\frac{\sin(t)}{2} + \frac{\sin(3t)}{6} \right]_{-\pi}^{\pi} \\ &= 0\end{aligned}$$

$$\begin{aligned}\langle \sin(t), \sin(2t) \rangle &= \frac{1}{\pi} \left[\frac{\sin(t)^3}{1.5} \right]_{-\pi}^{\pi} \\ &= 0\end{aligned}$$

Therefore the set is orthonormal

$$\begin{aligned}
\text{ii. } ||t||^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 dt \\
&= \left[\frac{t^3}{3} \right]_{-\pi}^{\pi} \\
&= \left[\frac{\pi^3}{3} + \frac{\pi^3}{3} \right] \\
&= 2 \frac{\pi^3}{3}
\end{aligned}$$

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

iii. Because we are dealing with an orthonormal sets we can write:

$$\begin{aligned}
Proj_x(cos(3t)) &= \sum_i \langle S_i, cos3t \rangle s_i \\
&= \langle cos(t), cos(3t) \rangle cos(t) + \langle cos(2t), cos(3t) \rangle cos(2t) + \langle sin(t), cos(3t) \rangle sin(t) + \\
&\quad \langle sin(2t), cos(3t) \rangle sin(2t)
\end{aligned}$$

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.

$$\begin{aligned}
\text{iv. } Proj_x(t) &= \sum_i \langle S_i, t \rangle s_i \\
&= \langle cos(t), t \rangle cos(t) + \langle cos(2t), t \rangle cos(2t) + \langle sin(t), t \rangle sin(t) + \langle sin(2t), t \rangle sin(2t) \\
&= 0 + 0 + 2sin(t) - sin(2t) \\
&= 2sin(t) - sin(2t)
\end{aligned}$$

Problem 3.9

We use the fact that for an orthonormal transformation, if for the corresponding transformation matrix Q, if $Q^T Q = I$ then the transformation is orthonormal.

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

So,

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

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

Problem 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 orthonormal 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 orthonormality, we know that:

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

and

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

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.

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

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

$$DD^H = \begin{bmatrix} n^2 & 0 \\ 0 & (\frac{1}{n})^2 \end{bmatrix} \neq I$$

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

$$(Q_1 Q_2)(Q_1 Q_2)^H = Q_1 Q_2 Q_2^H Q_1^H$$

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

$$Q_1 Q_2 Q_2^H Q_1^H = Q_1 Q_1^H = I$$

So the product of the matrices is orthonormal.

Problem 3.11

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

However, if $\{x_i\}_{i=1}^m$ are linearly dependent, then $x_m \in \text{Span}(\{x_i\}_{i=1}^{m-1})$, and $q_m = 0$.

This is contradictory to the assumption that $\{q_1, \dots, q_N\}$ are linearly dependent.

Problem 3.16

- i. Let $A \in \mathbb{M}_{m \times n}$ where $\text{rank}(A) \leq n$. Then there exist orthonormal $Q \in \mathbb{M}_{m \times n}$ and upper triangular $R \in \mathbb{M}_{m \times n}$ 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 $\hat{Q} \in \mathbb{M}_{n \times n}$ is orthonormal and $\hat{R} \in \mathbb{M}_{n \times n}$ is upper triangular. Since A has full column rank, \hat{R} has full rank and is therefore nonsingular. Then,

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

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

Problem 3.17

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

Note that \hat{R}^H is invertible. Thus, $\hat{R}^H x = \hat{Q}^H b$.

Problem 3.23

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, $\|y\| - \|x\| \leq \|x - y\|$.

Problem 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 $f(t) \neq 0$

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

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

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

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

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

Problem 3.26

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

i. Reflexivity

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

ii. 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!

iii. Transitivity

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

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

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

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

Problem 3.28

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

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

and

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

Problem 3.29

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

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

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

Problem 3.30

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

$$\|A\|_S \|x\| = \|SAS^{-1}\| \|x\| = \left(\sup_{x \neq 0} \frac{\|SAS^{-1}x\|}{\|x\|} \right) \|x\| \geq \|S\| \|A\| \|S^{-1}\| \|x\| = \|A\| \|x\| \geq \|Ax\|$$

Problem 3.37

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,

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

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

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

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

Problem 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;

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

Problem 3.39

- i. By definition of adjoint and linearity of inner products,

$$\begin{aligned} \langle (S + T)^* w, v \rangle_V &= \langle w, (S + T)v \rangle_W = \\ \langle w, Sv + Tv \rangle_W &= \langle w, Sv \rangle_W + \langle w, Tv \rangle_W = \\ \langle S^* w, v \rangle_V + \langle T^* w, v \rangle_V &= \langle S^* w + T^* w, v \rangle_V. \end{aligned}$$

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

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

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

- ii. By the definition of adjoint of S and the properties of inner products we have that

$$\langle w, Sv \rangle_W = \langle S^* w, v \rangle_V = \overline{\langle v, S^* w \rangle_V} = \overline{\langle S^{**} v, w \rangle_W} = \langle w, S^{**} v \rangle_W$$

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

iii. By the definition of adjoint we have

$$\begin{aligned} \langle (ST)^* v', v \rangle_V &= \langle v', (ST)v \rangle_V = \langle v', S(Tv) \rangle_V = \\ &= \langle S^* v', Tv \rangle_V = \langle T * S * v', v \rangle_V, \end{aligned}$$

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

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

Problem 3.40

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

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

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

$$\begin{aligned} \langle A_2, A_3 A_1 \rangle_F &= \text{tr}(A_2^H A_3 A_1) = \text{tr}(A_1 A_2^H A_3) = \text{tr}((A_2 A_1^H)^H A_3) = \langle A_2 A_1^H, A_3 \rangle_F \\ &= \langle A_2 A_1^*, A_3 \rangle. \end{aligned}$$

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

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

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

Problem 3.44

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

$$\langle y, b \rangle = \langle y, Ax \rangle = \langle A^H y, x \rangle = \langle 0, x \rangle = 0.$$

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

Problem 3.45

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

$$\langle B, A \rangle = \text{Tr}(B^T A) = \text{Tr}(AB^T) = \text{Tr}(A^T(-B)) = -\langle A, B \rangle.$$

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

$$0 = \langle B + B^T, B \rangle = \text{Tr}((B + B^T)B) = \text{Tr}(BB + B^T B) = \text{Tr}(BB) + \text{Tr}(B^T B),$$

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

Problem 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 $n = \text{Rank}(A) + \text{Dim}\mathcal{N}(A)$ and $n = \text{Rank}(A^H A) + \text{Dim}\mathcal{N}(A^H A)$. Then by (ii) it follows that $\text{Rank}(A) = \text{Rank}(A^H A)$.

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

Problem 3.47

- i. Notice that

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

- ii. Notice that

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

- 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

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

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

Problem 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 $\langle P^*(A), B \rangle = \langle A, P(B) \rangle$. Then, notice that

$$\begin{aligned} \langle A, P(B) \rangle &= \langle A, (B + B^T)/2 \rangle = \langle A, B/2 \rangle + \langle A, B^T/2 \rangle = \\ \text{Tr}(A^T B/2) + \text{Tr}(A^T B^T/2) &= \text{Tr}(A^T/2 B) + \text{Tr}(B A/2) = \\ \text{Tr}(A^T/2 B) + \text{Tr}(A/2 B) &= \langle (A + A^T)/2, B \rangle = \langle P(A), B \rangle. \end{aligned}$$

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

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

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

Problem 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 $A^H A \hat{x} = A^H b$, where

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

and

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