

# Math Pset #2: Inner Product Spaces

Alex Weinberg\*  
OSM Boot Camp 2018

July 1, 2018

## Problem 1.

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

PF:

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

As we are in  $\mathbb{R}$ ,

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

## Problem 2.

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

**Problem 3.**  $\langle x, x^5 \rangle = \int_0^1 x^6 dx = x^7/7|_0^1 = 1/7$ ,  $\|x\| = \int_0^1 x^2 dx = x^3/3|_0^1 = 1/3$   
and  $\|x^5\| = \int_0^1 x^{10} dx = x^{11}/11|_0^1 = 1/11$ . Therefore  $\cos \theta = \sqrt{33}/7$  implies  $\theta = 34.5$ .

---

\*Alberto did most of the Latexing

**Problem 8.**

(i)

$$\|\cos(t)\| = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(t) dt = \frac{1}{\pi} \left. \frac{\cos(x) \sin(x) - x}{2} \right|_{-\pi}^{\pi} i = \frac{\pi}{\pi} = 1,$$

and similarly  $\|\sin(t)\| = 1$ . Also

$$\|\cos(2t)\| = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(2t) dt = \frac{1}{\pi} \left. \frac{\sin(4t) + 4t}{8} \right|_{-\pi}^{\pi} i = \frac{\pi}{\pi} = 1,$$

and similarly  $\|\sin(2t)\| = 1$ . Therefore the basis is normalized.

The following integrals:

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

$$\langle \cos(t), \cos(2t) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(t) \cos(2t) dt = \frac{1}{\pi} \left. \frac{3 \sin(t) - 2 \sin^3(t)}{3} \right|_{-\pi}^{\pi} i = 0,$$

$$\langle \cos(t), \sin(2t) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(t) \sin(2t) dt = \frac{1}{\pi} \left. \frac{-2 \cos^3(t)}{3} \right|_{-\pi}^{\pi} i = 0,$$

$$\langle \cos(2t), \sin(2t) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(2t) \sin(2t) dt = \frac{1}{\pi} \left. \frac{-\cos^2(2t)}{4} \right|_{-\pi}^{\pi} i = 0,$$

and so on, shows that  $S$  is an orthonormal basis.

(ii)

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

(iii) Since  $\langle x, \cos(3x) \rangle = 0$  for any  $x \in S$ ,  $\text{proj}_X(\cos(3x)) = 0$ .

(iv)

$$\begin{aligned} \langle \sin(t), t \rangle &= \sin(t) - t \cos(t) \Big|_{-\pi}^{\pi} = 2\pi, \\ \langle \cos(t), t \rangle &= t \sin(t) - \cos(t) \Big|_{-\pi}^{\pi} = 0, \\ \langle \cos(2t), t \rangle &= (2t \sin(2t) + \cos(2t)) / 4 \Big|_{-\pi}^{\pi} = 0, \text{ and finally} \\ \langle \sin(2t), t \rangle &= \sin(2x) - 2x \cos(2x) / 4 \Big|_{-\pi}^{\pi} = -\pi. \end{aligned}$$

Therefore,  $\text{proj}_X(t) = 2\pi \sin(t) - \pi \sin(2t)$

**Problem 9.**

A rotation of angle  $\theta$  in  $\mathbb{R}^2$  represented as a matrix  $R$  in the standard basis is an orthonormal transformation since

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}^T \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & 0 \\ 0 & \cos^2(\theta) + \sin^2(\theta) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Similarly, one shows that  $RR^T = I$ . Therefore, a rotation in  $\mathbb{R}^2$  is an orthonormal transformation.

**Problem 10.**

(i) Suppose  $Q$  represents an orthonormal operator on  $\mathbb{F}^n$ . Then  $\langle x, y \rangle = \langle Q(x), Q(y) \rangle$  for each  $x, y \in \mathbb{F}^n$ . Since  $\langle Q(x), Q(y) \rangle = (Qx)^H(Qy) = x^H Q^H Q y$ , it equals  $x^H y$  for all  $x, y \in \mathbb{F}^n$  only if  $Q^H Q = I$ . On the other hand if  $Q^H Q = Q Q^H = I$ , then  $\langle Q(x), Q(y) \rangle = (Qx)^H(Qy) = x^H Q^H Q y = x^H y = \langle x, y \rangle$ .

(ii)

$$\|Qx\| = \sqrt{\langle Qx, Qx \rangle} = \sqrt{x^H Q^H Q x} = \sqrt{\langle x, x \rangle} = \|x\|.$$

(iii) If  $Q^H Q = Q Q^H = I$ , then  $Q^{-1} = Q^H$ . Since  $(Q^H)^H = Q$ ,  $Q^H$  is also orthonormal:

$$(Q^H)^H Q^H = Q Q^H = I = Q^H Q = Q^H (Q^H)^H.$$

(iv) Let  $q_i$  denote the  $i^{\text{th}}$  column of  $Q$ . Since  $Q$  is orthonormal,  $(Q^H Q)_{ij} = q_i^H q_j = \langle q_i, q_j \rangle$  is 1 if  $i = j$  and 0 if  $i \neq j$ . Thus, the columns of  $Q$  are orthonormal.

(v) The matrix

$$\begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}$$

shows that the converse is not true.

(vi)

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

and

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

Therefore,  $Q_1 Q_2$  is orthonormal.

**Problem 11.**

Fix  $N \in \mathbb{N}$ ,  $N > 0$ , and suppose  $\{x_i\}_{i=1}^N$  is a set of linearly dependent vectors in  $V$ . Also, suppose, without loss of generality, that for  $2 < k < N$ ,  $\{x_i\}_{i=1}^{k-1}$  is a linearly independent set and  $\{x_i\}_{i=1}^k$  is a linearly dependent set. Then  $\{q_i\}_{i=1}^{k-1}$  (as they are defined in the book) is also a linearly independent set. However, since  $x_k \in \text{span}(\{x_i\}_{i=1}^{k-1})$ , we have that  $q_k = 0$ . Therefore the Gram-Schmidt orthonormalization process brakes down.

**Problem 16.**

(i) Let  $A \in \mathbb{M}_{m \times n}$  where  $\text{rank}(A) = n \leq m$ . Then there exist orthonormal  $Q \in \mathbb{M}_{m \times m}$  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) Suppose now that  $A$  is invertible and can be decomposed into two different QR decompositions:  $QR$  and  $\tilde{Q}\tilde{R}$ , where the diagonal entries of  $R$  and  $\tilde{R}$  are strictly positive. Then both  $R$  and  $\tilde{R}$  are invertible and we conclude that  $\tilde{R}^{-1}R = Q^H\tilde{Q}$ . Since  $R$  and  $\tilde{R}$  are upper triangular, so is the LHS of the previous equation. On the other hand, since  $Q$  and  $\tilde{Q}$  are orthonormal, so is the RHS. Therefore  $\tilde{R}^{-1}R = I$  and, by unicity of the inverse, we conclude that  $R = \tilde{R}$ , and so  $Q = \tilde{Q}$ .

**Problem 17.**

Suppose,  $A = \hat{Q}\hat{R}$  is reduced.

WTS:

$$A^T Ax = A^T b \iff \hat{R}x = \hat{Q}^T b$$

PF:

$$\begin{aligned} (A^T A)x &= A^T b \\ \implies (\hat{R}^T \hat{Q}^T \hat{Q} \hat{R})x &= A^T b \\ \implies \hat{R}^T \hat{R}x &= A^T b \\ \implies \hat{R}^T \hat{R}x &= (\hat{Q} \hat{R})^T b \\ \implies \hat{R}^T \hat{R}x &= \hat{R}^T \hat{Q}^T b \\ \implies \hat{R}x &= \hat{Q}^T b \quad \square \end{aligned}$$

**Problem 23.** Let  $x, y \in V$  and define  $v := -y$ . Since a norm is nonnegative and satisfies the triangular property,  $\|x\| - \|v\| \leq \|x\| + \|v\| \leq \|x + v\|$ . Then our definition of  $v$  implies  $\|x\| - \|y\| = \|x\| - \|-y\| \leq \|x - y\|$ . Interchanging the role of  $x$  and  $y$  and using the homogeneity property of norms we have  $\|y\| - \|x\| \leq \|y - x\| = \|- (y - x)\| = \|x - y\|$ , and the result follows.

**Problem 24.** (i) Since  $|f(t)| \geq 0$  for every  $t$ , so is  $\int_a^b |f(t)| dt$ . In addition, if  $f = 0$ , then  $\int_a^b |f(t)| dt = 0$ . On the other hand, if  $\int_a^b |f(t)| dt = 0$  and  $|f(t)| \geq 0$ , it must be that  $|f(t)| = 0$  for all  $t$ , implying that  $f = 0$ . Now take a constant  $c \in \mathbb{F}$ , then  $\int_a^b |cf(t)| dt = \int_a^b |c| |f(t)| dt = |c| \int_a^b |f(t)| dt$ , since  $c$  does not depend on  $t$ . Finally, take  $g \in C([a, b]; \mathbb{F})$ . Since  $|f(t) + g(t)| \leq |f(t)| + |g(t)|$  for all  $t$  and the integral is a linear operator, we have that  $\int_a^b |f(t) + g(t)| dt \leq \int_a^b |f(t)| dt + \int_a^b |g(t)| dt$ .

(ii) Since  $|f(t)|^2 \geq 0$  for every  $t$ , so is  $\int_a^b |f(t)|^2 dt$  and its square root. In addition, if  $f = 0$ , then  $|f(t)|^2 = 0$  for all  $t$  and  $\sqrt{\int_a^b |f(t)|^2 dt} = 0$ . On the other hand, if  $\sqrt{\int_a^b |f(t)|^2 dt} = 0$ , then  $\int_a^b |f(t)|^2 dt = 0$  and since  $|f(t)|^2 \geq 0$  for all  $t$ , it must be

that  $|f(t)|^2 = 0$  for all  $t$ , implying that  $f = 0$ . Now take a constant  $c \in \mathbb{F}$ , then  $\sqrt{\int_a^b |cf(t)|^2 dt} = \sqrt{\int_a^b |c|^2 |f(t)|^2 dt} = |c| \sqrt{\int_a^b |f(t)|^2 dt}$ , since  $c$  does not depend on  $t$ . Finally, take  $g \in C([a, b]; \mathbb{F})$ . Since  $|f(t) + g(t)| \leq |f(t)| + |g(t)|$  for all  $t$ ,  $x \mapsto x^2$  and  $x \mapsto \sqrt{x}$  are monotonically increasing for nonnegative  $x$  and the integral is a linear operator, we have that  $\sqrt{\int_a^b |f(t) + g(t)|^2 dt} \leq \sqrt{\int_a^b |f(t)|^2 dt + \int_a^b |g(t)|^2 dt} \leq \|f\|_{L^2} + \|g\|_{L^2}$ .

(iii) Since  $|f(x)| \geq 0$  for all  $x$ , so is the  $\sup_{x \in [a, b]} |f(x)|$ . In addition, if  $f = 0$ , then  $\sup_{x \in [a, b]} |f(x)|$  is also zero. On the other hand, since  $|f(x)| \geq 0$  for all  $x$ ,  $0 \leq \sup_{x \in [a, b]} |f(x)| = 0$  implies that we must have  $f = 0$ . Now take a constant  $c \in \mathbb{F}$ , then  $\sup_{x \in [a, b]} |cf(x)| = \sup_{x \in [a, b]} |c| |f(x)| = |c| \sup_{x \in [a, b]} |f(x)|$ . Finally, take  $g \in C([a, b]; \mathbb{F})$ . Since  $|f(x) + g(x)| \leq |f(x)| + |g(x)|$  for all  $x$ , we have that  $\sup_{x \in [a, b]} |f(x) + g(x)| \leq \sup_{x \in [a, b]} \{|f(x)| + |g(x)|\} \leq \sup_{x \in [a, b]} |f(x)| + \sup_{x \in [a, b]} |g(x)|$ .

**Problem 26.** We show that topological equivalence is an equivalence relation. Let  $\|\cdot\|_r$  be a norm on  $X$  for  $r \in \{a, b, c\}$ . Clearly  $\|\cdot\|_r$  is topologically equivalent with itself, just pick any  $0 < m \leq 1$  and any  $M \geq 1$  to show this. Also, suppose that  $\|\cdot\|_a$  is topologically equivalent to  $\|\cdot\|_b$  with constants  $0 < m \leq M$ . Then,  $\|\cdot\|_b$  is topologically equivalent to  $\|\cdot\|_a$  with constants  $0 < 1/M' \leq 1/m'$ . Finally, if  $\|\cdot\|_a$  is topologically equivalent to  $\|\cdot\|_b$  with constants  $0 < m \leq M$  and so is  $\|\cdot\|_b$  with  $\|\cdot\|_c$  with constants  $0 < m' \leq M'$ , then  $\|\cdot\|_a$  is topologically equivalent to  $\|\cdot\|_b$  with constants  $0 < mm' \leq MM'$ .

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 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

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

**Problem 29.** Take an arbitrary  $x \neq 0$  and suppose  $\|\cdot\|$  is an inner product induced norm. Since

$$\|Qx\| = (\langle Qx, Qx \rangle)^{1/2} = (\langle Q^H Qx, x \rangle)^{1/2} = (\langle x, x \rangle)^{1/2} = \|x\|,$$

then

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

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

$$\|R_x\| = \sup_{A \neq 0} \frac{\|Ax\|}{\|A\|} = \sup_{A \neq 0} \frac{\|Ax\| \|x\|}{\|A\| \|x\|} \leq \sup_{A \neq 0} \left( \frac{\|Ax\| \|x\|}{\|Ax\|} \right) = \|x\|.$$

**Problem 30.** Take arbitrary matrices  $A, B \in \mathbb{M}_n(\mathbb{F})$ . First,  $\|A\|_S = \|SAS^{-1}\| \geq 0$  for any  $A$  because  $\|\cdot\|$  is a norm on  $\mathbb{M}_n(\mathbb{F})$  and  $SAS^{-1} \in \mathbb{M}_n(\mathbb{F})$ . In addition,  $\|0\|_S = \|S0S^{-1}\| = \|0\| = 0$  and if  $0 = \|A\|_S = \|SAS^{-1}\|$ , then  $SAS^{-1} = 0$  which implies  $A = 0$ . Second, take  $a \in \mathbb{F}$ , then

$$\|aA\|_S = \|SaAS^{-1}\| = \|aSAS^{-1}\| = |a|\|SAS^{-1}\| = |a|\|A\|_S.$$

Finally, let  $B \in \mathbb{M}_n(\mathbb{F})$  and notice that

$$\|A+B\|_S = \|S(A+B)S^{-1}\| = \|SAS^{-1} + SBS^{-1}\| \leq \|SAS^{-1}\| + \|SBS^{-1}\| = \|A\|_S + \|B\|_S.$$

Therefore  $\|\cdot\|_S$  is a norm on  $\mathbb{M}_n(\mathbb{F})$ . To show that it is a matrix norm, notice that

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

and so  $\|AB\|_S \leq \|A\|_S \|B\|_S$ .

**Problem 37.**

FIND:  $q$  s.t.  $p'q = L[p] = p'(1)$

$p$  can be written as

$$\begin{aligned} p(x) &= a_0 + a_1x + a_2x^2. \\ \implies p &= [1 \quad x_1 \quad x_2^2][a_0 \quad a_1 \quad a_2]^T \\ &\implies p'(1) = a_1 + 2a_2 \\ &\iff [a_0 \quad a_1 \quad a_2][0 \quad 1 \quad 2]^T \end{aligned}$$

So choose  $q = [0 \quad 1 \quad 2]^T \quad \square$

**Problem 38.** Let  $p = ax^2 + vx + c$  be an arbitrary element of  $V = \mathbb{F}[x; 2]$ . Since we can represent  $p = (a, b, c)^T$ , and  $p' = D(p) = (0, 2a, b)^T$ , we that the matrix representation of  $D$  is

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

and the hermitian is just the transpose

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

**Problem 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 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

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

(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 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 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 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 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 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

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

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

**Problem 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

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