

# Math Problem Set 2

Matthew Brown  
OSM Boot Camp 2018

June 30, 2018

**Problem 3.1.** There are two parts:

- (i)

$$\begin{aligned} ||x + y||^2 - ||x - y||^2 &= \langle x + y, x + y \rangle - \langle x - y, x - y \rangle \\ &= \langle x + y, x \rangle + \langle x + y, y \rangle - (\langle x - y, x \rangle + \langle x - y, -y \rangle) \\ &= \langle x, x + y \rangle + \langle y, x + y \rangle - (\langle x, x - y \rangle + \langle -y, x - y \rangle) \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle - \langle x, x \rangle - \langle x, -y \rangle - \langle -y, x \rangle - \langle -y, -y \rangle \\ &= 4\langle x, y \rangle \text{ after some easy manipulations} \end{aligned}$$

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

- (ii) As above,

$$\begin{aligned} ||x + y||^2 + ||x - y||^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= \langle x + y, x \rangle + \langle x + y, y \rangle + (\langle x - y, x \rangle + \langle x - y, -y \rangle) \\ &= \langle x, x + y \rangle + \langle y, x + y \rangle + (\langle x, x - y \rangle + \langle -y, x - y \rangle) \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle + \langle x, x \rangle + \langle x, -y \rangle + \langle -y, x \rangle + \langle -y, -y \rangle \\ &= 2\langle x, x \rangle + 2\langle y, y \rangle \text{ after some easy manipulations} \\ &= 2(||x||^2 + ||y||^2) \end{aligned}$$

So it is clear that  $\frac{1}{2}(|x + y|^2 + |x - y|^2) = ||x||^2 + ||y||^2$

**Problem 3.2.** I proceed similarly: just start from the left side, expand, and simplify.

**Problem 3.3.** Let  $\theta$  be the angle in question. Recall that

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

- (i)

$$\begin{aligned}\langle f, g \rangle &= \int_0^1 f g dx = \int_0^1 x^6 dx = \left( \frac{x^7}{7} \right) \Big|_0^1 = \frac{1}{7} \\ \|f\|^2 &= \int_0^1 f^2 dx = \int_0^1 x^2 dx = \left( \frac{x^3}{3} \right) \Big|_0^1 = \frac{1}{3} \\ \|g\|^2 &= \int_0^1 g^2 dx = \int_0^1 x^{10} dx = \left( \frac{x^{11}}{11} \right) \Big|_0^1 = \frac{1}{11}\end{aligned}$$

Thus, we see that

$$\cos(\theta) = \frac{\frac{1}{7}}{\left(\frac{1}{3} \cdot \frac{1}{11}\right)^{\frac{1}{2}}} = \frac{33^{\frac{1}{2}}}{7} \quad (1)$$

And (1) implies that  $\theta \approx 35^\circ$ .

- (ii)

$$\begin{aligned}\langle f, g \rangle &= \int_0^1 f g dx = \int_0^1 x^4 dx = \left( \frac{x^5}{5} \right) \Big|_0^1 = \frac{1}{5} \\ \|f\|^2 &= \int_0^1 f^2 dx = \int_0^1 x^2 dx = \left( \frac{x^3}{3} \right) \Big|_0^1 = \frac{1}{3} \\ \|g\|^2 &= \int_0^1 g^2 dx = \int_0^1 x^8 dx = \left( \frac{x^9}{9} \right) \Big|_0^1 = \frac{1}{9}\end{aligned}$$

Thus, we see that

$$\cos(\theta) = \frac{\frac{1}{5}}{\left(\frac{1}{3} \cdot \frac{1}{9}\right)^{\frac{1}{2}}} = \frac{45^{\frac{1}{2}}}{7} \quad (2)$$

And (1) implies that  $\theta \approx 17^\circ$

**Problem 3.8.** There are four parts

- (i)

*Proof.* This is just a matter of checking all the relevant details.

Norms = 1:

$$\begin{aligned}\| \cos(t) \|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(t) dt = \frac{1}{\pi} \frac{\cos(x)\sin(x) + x}{2} \Big|_{-\pi}^{\pi} = 1 \\ \| \sin(t) \|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(t) dt = \frac{1}{\pi} \frac{-\sin(2x) + 2x}{4} \Big|_{-\pi}^{\pi} = 1 \\ \| \cos(2t) \|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(2t) dt = \frac{1}{\pi} \frac{\sin(4x) + 4x}{8} \Big|_{-\pi}^{\pi} = 1 \\ \| \sin(2t) \|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(2t) dt = \frac{1}{\pi} \frac{-\sin(4x) + 4x}{8} \Big|_{-\pi}^{\pi} = 1\end{aligned}$$

Inner Products = 1:

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

The proof is completed by checking all the other inner products similarly.  $\square$

- (ii)

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

- (iii)

$$\begin{aligned} \text{proj}_X(\cos(3t)) &= \langle \sin(t), \cos(3t) \rangle \sin(t) + \langle \cos(t), \cos(3t) \rangle \cos(t) \\ &\quad + \langle \sin(2t), \cos(3t) \rangle \sin(2t) + \langle \cos(2t), \cos(3t) \rangle \cos(2t) \\ &= 0 + 0 + 0 + 0 = 0 \end{aligned}$$

(We see that  $\cos(3t)$  is orthogonal to  $X$ )

- (iv)

$$\text{proj}_X(t) = 0 + 2\sin(t) + 0 - \sin(2t)$$

**Problem 3.9.** *Proof.* In  $\mathbb{R}^2$ , a rotation about the origin by arbitrary angle  $\theta$  can be described by the matrix

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

So that  $M \left( \begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} x\cos(\theta) - y\sin(\theta) \\ x\sin(\theta) + y\cos(\theta) \end{pmatrix}$ . Let  $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$ . Then

$$\begin{aligned} \langle M(a), M(b) \rangle &= \left\langle \begin{pmatrix} a_1\cos(\theta) - a_2\sin(\theta) \\ a_1\sin(\theta) + a_2\cos(\theta) \end{pmatrix}, \begin{pmatrix} b_1\cos(\theta) - b_2\sin(\theta) \\ b_1\sin(\theta) + b_2\cos(\theta) \end{pmatrix} \right\rangle \\ &= a_1b_1\cos^2(\theta) + a_2b_2\sin^2(\theta) + a_1b_1\sin^2(\theta) + a_2b_2\cos^2(\theta) \\ &= a_1b_1 + a_2b_2 = \langle a, b \rangle \end{aligned}$$

So we see that  $M$  is an orthonormal operator. (N.b that this is a terribly inefficient way to prove this - I should have just shown that the columns of  $M$  were orthonormal!  $\square$ )

**Problem 3.10.** Recall that taking the Hermitian is flipping rows and columns and taking the conjugate.

- (i)

*Proof.* We need to show both directions.

$\Rightarrow$ : Let  $Q$  be an orthonormal matrix. Then  $\langle m, n \rangle = \langle Qm, Qn \rangle \implies m^H n = (Qm)^H Qn = m^H Q^H Qn$ . And because  $m$  and  $n$  were arbitrarily chosen, the only way that this equality holds is if  $Q^H Q = I$ , and this gives us that  $QQ^H = Q$  since left inverse  $\implies$  right inverse.

$\Leftarrow$ : Let  $Q$  be a matrix so that  $Q^H Q = I$ . Then consider  $\langle Qm, Qn \rangle = (Qm)^H Qn = m^H Q^H Qn = m^H n = \langle m, n \rangle$ .  $\square$

- (ii)

*Proof.* This is pretty easy:

$$\|x\| = \sqrt[2]{\langle x, x \rangle} = \sqrt[2]{\langle Qx, Qx \rangle} = \|Qx\|$$

$\square$

- (iii)

*Proof.* Assume  $Q$  is orthonormal. Then  $QQ^H = Q^H Q = I \implies Q^H = Q^{-1}$ . I'll prove the following short lemma.

**Lemma 0.1.** For  $Q$  orthonormal,  $Q^H$  is orthonormal.

Recall that  $(Q^H)^H = Q$ , and see that

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

which proves the lemma.

And since  $Q^{-1} = Q^H$ ,  $Q^{-1}$  is orthonormal.  $\square$

- (iv)

*Proof.* We'll examine the elements of the identity matrix element by element. First note that:

$$I_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Then we'll compare this to what we get when we multiply  $QQ^H$ , which we know is equal to  $I$  in all its coordinates. First, though, for any matrix  $A$ , define  $A^i$  to be the "ith row" of  $A$  and  $A_j$  to be the jth column. Then

$$\delta_{ij} = (Q^H Q)_{ij} = (Q^H)^i Q_j =$$

Recall now that  $(Q^H)^i = \bar{Q}_i$ , by definition of the Hermitian. But now we see that

$$\langle \bar{Q}_i, q_j \rangle = \delta_{ij}$$

and the columns of  $Q$  are orthonormal.  $\square$

- (v) A counterexample would be the matrix  $M = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ .  $\det B = 1$ , but  $Be_1 = 2e_1$  and  $\|e_1\| \neq \|2e_1\| = \|Be_1\|$  which violates what we proved in (ii).
- (vi)

*Proof.* This is also quite short:

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

And we also get the left inverse by properties of inverses.  $\square$

**Problem 3.11.** Suppose that  $x_1, \dots, x_n$  is a set of linearly *dependent* vectors. Let's apply Gram-Schmidt. Eventually, we will arrive at a vector  $x_k$  which is linearly dependent upon  $x_1, \dots, x_{k-1}$ . But then, if  $X = \text{span}(x_1, \dots, x_{k-1})$ , then  $x_k \in X$  and  $p_{k-1} = \text{proj}_X(x_k) = x_k$ , which forces  $q_k = 0$ . In the end, if we throw out all these zeroes, we still get an orthonormal basis  $q_1, \dots, q_m$  of  $X$  where  $m = \dim X$ .

**Problem 3.16.** Done with Albi

**Problem 3.17.** *Proof.*

$$\begin{aligned} \hat{R}x &= \hat{Q}^H b \\ \iff \hat{Q} \hat{R}x &= b \\ \iff Ax &= b \\ \iff A^H Ax &= A^H b \end{aligned}$$

Note that THIS IS NOT THE CORRECT PROOF AND I NEED TO FIX THIS  $\square$

**Problem 3.23.** doesn't look bad

**Problem 3.24.** Procedural (show things are norms)

**Problem 3.26.** First I must show that topological equivalence is an equivalence relation.

*Proof.* REST OF THIS DONE WITH ALBI SEE GIT I must show three things: (i)  $x \sim x$ .

(ii)  $x \sim y \implies y \sim x$ , (iii)  $x \sim y$  and  $y \sim z \implies x \sim z$ .

(i)  $\|\cdot\|_1 \sim \|\cdot\|_1$  trivially. Let  $M \geq m$ , then  $m\|x\|_1 \leq \|x\|_1 \leq M\|x\|_1$  for all  $x$ .

(ii) Also trivial: Suppose  $\|\cdot\|_1 \sim \|\cdot\|_2$ . Then  $m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$  for all  $x$ , which implies that  $M^{-1}\|x\|_2 \leq \|x\|_1 \leq m^{-1}\|x\|_2$ .

(iii) Suppose  $\|\cdot\|_1 \sim \|\cdot\|_2$ , and  $\|\cdot\|_2 \sim \|\cdot\|_3$ . Then  $m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$  and  $n\|x\|_2 \leq \|x\|_3 \leq N\|x\|_2$ . But we get from this that  $mn\|x\|_1 \leq \|x\|_3 \leq MN\|x\|_1$ .  $\square$

Now I'll show that the 1, 2, and  $\infty$  norms are topologically equivalent.

*Proof.* (i)  $\|\cdot\|_1 \sim \|\cdot\|_2$ :

If we think about the inner product as the standard dot-product, then we have

$$(\|x\|_1)^2 = \sum_{i=1}^n \sum_{j=1}^n |x_i| |x_j| \geq \sum_{i=1}^n x_i^2 = \langle x, x \rangle = (\|x\|_2)^2$$

(the inequality comes because we simply threw out some positive terms on the left side). This implies that  $\|x\|_1 \geq \|x\|_2$ . Moreover,

$$(\sqrt[2]{n} \|x\|_2)^2 = n \sum_{i=1}^n x_i^2$$

(ii)  $\|\cdot\|_\infty \sim \|\cdot\|_2$

$$\|x\|_\infty = \max_{1 \leq i \leq n} \{x_i\} = \sqrt[2]{(\max_{1 \leq i \leq n} \{x_i\})^2} \leq \sqrt[2]{\sum_{i=1}^n x_i^2} = \|x\|_2$$

□

**Problem 3.28.** Related to Previous

**Problem 3.29.** I will prove two statements.

**The norm of an orthonormal matrix is 1:**

*Proof.* Let  $Q$  be an orthonormal matrix. Then

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

□

**If  $R_x : M_n(\mathbb{F}) \rightarrow \mathbb{F}$ ,  $R_x(A) = Ax$ , then  $\|R_x\| = \|x\|$ :**

*Proof.* The first step is to show  $\|R_x\| \leq \|x\|$ .

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

By Remark 3.5.12,  $\|Ax\| \leq \|A\| \cdot \|x\| \forall x \in \mathbb{F}^n$ , so

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

Now I'll show equality. For the  $\leq$  above to be strict, we must have  $\|Ax\| < \|A\| \cdot \|x\|$  for all operators  $A$  (because we're taking the supremum).  $\|x\| > 0$ , so I can rearrange for the condition:

$$\frac{\|Ax\|}{\|x\|} < \|A\|, \text{ for all operators } A, \text{ vectors } x$$

In other words, no  $x$  achieves the supremum which is encoded in the definition of  $\|A\|$ . I will use the previous result to show that this will never hold.

Let  $q_1 = e_1$  (or some other vector with norm 1). I can use the gram-schmidt algorithm to construct an orthonormal basis  $q_1, \dots, q_n$  for  $\mathbb{F}^n$ . Let  $Q$  be the matrix with these basis vectors as its columns. Then  $Q$  is an orthonormal matrix. Specifically,  $\|Q\| = 1$  and it achieves  $\frac{\|Qx\|}{\|x\|} = \|Q\| = 1$  at all nonzero  $x$ .

This shows that the inequality can never be strict, so we have  $\|R_x\| = \|x\|$   $\square$

**Problem 3.30.** *Proof.* To show something is a norm, I must show three properties:

- *Positivity:* This follows immediately from the positivity of the underlying matrix norm:  $\|A\|_S = \|SAS^{-1}\| \geq 0$ , with equality iff  $SAS^{-1} = 0$ , and there are no elements that are conjugate to 0 other than itself so  $A = 0$  in this case.

- *Scalar Preservation:* I use linearity of  $S$  and the corresponding property for the matrix norm.

$$\|kA\|_S = \|SkAS^{-1}\| = \|kSAS^{-1}\| = k\|SAS^{-1}\| = k\|A\|_S$$

- *Triangle Inequality:* This also follows from the linearity of  $S$ .

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

Finally, to see that it is a matrix norm, I must show that it is submultiplicative. And indeed

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

$\square$

**Problem 3.37.** The first thing is to define the standard basis, which is  $\mathcal{B} = \{1, x, x^2\}$ . Evaluate  $L$  on the basis vectors:

$$L(1) = 0, L(x) = 1, L(x^2) = 2$$

Now, for  $p \in V$ ,  $p$  can be written as a linear combination of these basis vectors. So

$$L(p) = L(a_1 1 + a_2 x + a_3 x^2) = a_1 L(1) + a_2 L(x) + a_3 L(x^2) = \langle (L(1) \cdot 1, L(x), L(x^2)), (a_1, a_2, a_3) \rangle$$

which is the idea behind the Riesz Representation theorem. So we see that in this case,  $q = (0, 1, 2) \dots$  which squares with what we know about derivatives.

**Problem 3.38.** Let  $\mathcal{B}$  as above.

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

By following the same method as exercise 3.7.9, I see that

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

**Problem 3.39.** There are 4 things to show.

- (i)

*Proof.*

$$\begin{aligned}\langle (S + T)v, w \rangle &= \langle Sv, w \rangle + \langle Tv, w \rangle = \langle v, S^*w \rangle + \langle v, T^*w \rangle = \langle v, (S^* + T^*)w \rangle \\ \langle \alpha T^*v, w \rangle &= \alpha \langle Tv, w \rangle = \alpha \langle v, T^*w \rangle = \langle v, \bar{\alpha} T^*w \rangle\end{aligned}$$

□

- (ii)

*Proof.*

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

□

- (iii)

*Proof.*

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

□

- (iv)

*Proof.* Consider the composition  $T^*(T^{-1})^*$ .

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

Since the above is true for all  $x, y$ , we must have  $T^*(T^{-1})^* = I$

□

**Problem 3.40.**

- (i) (Considering  $A$  as the operator)

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

- (ii) (seems easy!)

**Problem 3.44.** *Proof.* By the fundamental subspaces theorem,  $\text{Ker}(A^H) = \text{Range}(A)$ . So we can reformulate the second possibility to: there exists  $y \in \text{Range}(A)^\perp : \langle y, b \rangle \neq 0$ .

Consider now  $p = \text{proj}_{\text{Range}(A)} b$ . If  $p = b$ , then  $b \in \text{Range}(A)$  and we have the first case. Otherwise, the procedure creates a residual vector  $r$ ,  $r = b - p$ .  $r \in \text{Range}(A)^\perp$ , and

$$\langle r, b \rangle = \overline{\langle p + r, r \rangle} = \overline{\langle p, r \rangle} + \overline{\langle r, r \rangle} = \langle p, b \rangle + \langle r, r \rangle = \langle r, r \rangle \neq 0$$

which is the second case.

□



**Problem 3.45.** Props to Albi for this idea - I think this is really slick (especially the last part)!

*Proof.* Double inclusion. First show that  $\text{Skew}_n(\mathbb{R}) \subset \text{Sym}_n^\perp(\mathbb{R})$ .

Let  $A \in \text{Skew}_n(\mathbb{R})$  Then, recalling the definitions of the spaces and the properties of trace,

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

But since we are over  $\mathbb{R}$ ,  $\langle A, B \rangle = -\overline{\langle A, B \rangle} \implies \langle A, B \rangle = 0$  for all  $B \in \text{Sym}_n(\mathbb{R})$ , so we see that  $A \in \text{Sym}_n(\mathbb{R})^\perp$

Now I'll show that  $\text{Sym}_n^\perp(\mathbb{R}) \subset \text{Skew}_n(\mathbb{R})$ .

Let  $A \in \text{Sym}_n(\mathbb{R})^\perp$ . The key insight here is that  $A + A^T \in \text{Sym}_n(\mathbb{R})$  Then

$$0 = \langle A, A + A^T \rangle = \text{tr}(A, A + A^T)$$

...Actually I have some qualms about the final step of this \*Remember to ask albi. □

**Problem 3.46.** Four Parts

*Proof.* • (i)  $Ax \in \text{Range}(A)$  by definition, and  $x \in \text{Ker}(A^H A) \implies A^H(Ax) = 0 \implies Ax \in \text{Ker} A^H$ .

• (ii) Clearly  $\text{Ker}(A) \subset \text{Ker}(A^H A)$  since  $A^H(0) = 0$ . It remains to show that

$$A^H Ax = 0 \implies Ax = 0$$

Suppose  $A^H Ax = 0$ . Then if we apply the operator  $x^T$  as a row vector, we see that

$$x^H A^H Ax = (Ax)^H Ax = \langle Ax, Ax \rangle = 0$$

and  $\langle Ax, Ax \rangle = 0 \implies Ax = 0$  which is what we wanted to show.

• (iii)  $A$  and  $A^H A$  both map to the  $n$ -dimensional spaces and  $\text{Ker}(A) = \text{Ker}(A^H A)$  by the above, so by Rank-Nullity,

$$\begin{aligned} n - \dim \text{Ker}(A) &= \dim \text{Range}(A) \\ n - \dim \text{Ker}(A^H A) &= \dim \text{Range}(A^H A) \end{aligned}$$

The left sides of the two equations are equal so the right sides must also be equal!

• (iv)  $A$  has linearly independent columns  $\implies$   
 $A$  has rank  $n \implies$   
 $A^H A$  has (full) rank  $n \implies$   
 $A^H A$  is non singular.

□

**Problem 3.47.** *Proof.* • (i)

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

- (ii)

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

- (iii) Er... I'm not sure this is even true. Suppose  $m < n$ , then  $P$  is  $m$  by  $m$  and cannot possibly have rank  $n$ ? What am I missing here?

□

**Problem 3.48.** SO MANY PARTS AAAH!

*Proof.* • (i) I'll do it all at once... let  $k \in \mathbb{R}, A, B \in M_n(\mathbb{R})$

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

- (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) First see that  $P(A) = P(A^T)$  (this is trivial). Then

$$\begin{aligned} \langle P(A), B \rangle &= \text{tr}(P(A)^T B) = \text{tr}\left(\frac{A+A^T}{2} \cdot B\right) = \frac{\text{tr}(A^T B + AB)}{2} = \text{tr}(AB) \\ &= \frac{\text{tr}(AB + AB^T)}{2} = \text{tr}\left(A \cdot \frac{B+B^T}{2}\right) = \text{tr}(AP(B)) = \langle A, P(B) \rangle \end{aligned}$$

- (iv)

$$A \in \text{Ker}(P) \iff P(A) = 0 \iff A + A^T = 0 \iff A = -A^T \iff A \in \text{Skew}_n(\mathbb{R})$$

- (v)

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

- (vi) Copied latex code from Albi as this is just long and mechanical.

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

□

**Problem 3.50.** I want to estimate the least squares solution for  $Ax = b$  where:

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

The normal equation to solve is:

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i y_i & \sum_{i=1}^n y_i^2 \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i \end{bmatrix}$$