

## Math, Problem Set #2, Inner Product Spaces

Instructor: Zachary Boyd

Due Wednesday, July 5 at 8:00am

**Homework:** 1, 2, 3, 8, 9, 10, 11, 16, 17, 23, 24, 26, 28, 29, 30, 37, 38, 39, 40, 44, 45, 46, 47, 48, 50 at the end of Chapter 3 of Humpherys et al. (2017)

3.1) i)

$$\begin{aligned} & \frac{1}{4} (||x + y||^2 - ||x - y||^2) \\ &= \frac{1}{4} (||x||^2 + ||y||^2 + 2||x|| ||y|| \cos(\theta) - [||x||^2 + ||y||^2 - 2||x|| ||y|| \cos(\theta)]) \\ &= \frac{1}{4} (4||x|| ||y|| \cos(\theta)) \quad \text{by definition of } \cos(\theta) \\ &= ||x|| ||y|| \frac{\langle x, y \rangle}{||x|| ||y||} \\ &= \langle x, y \rangle \end{aligned}$$

ii)

$$\begin{aligned} & \frac{1}{2} (||x + y||^2 + ||x - y||^2) \\ &= \frac{1}{2} (||x||^2 + ||y||^2 + 2||x|| ||y|| \cos(\theta) + ||x||^2 + ||y||^2 - 2||x|| ||y|| \cos(\theta)) \\ &= \frac{1}{2} 2 (||x||^2 + ||y||^2) \\ &= ||x||^2 + ||y||^2 \end{aligned}$$

3.2)

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

3.3)

$$\theta = \cos^{-1} \left( \frac{\langle f, g \rangle}{||f|| ||g||} \right)$$

for i) and ii) we have

$$\langle f, g \rangle = \int_0^1 x^6 dx = \frac{1}{7}$$

and for i) we further have:

$$\begin{aligned}
||g|| &= ||x|| = \sqrt{\langle g, g \rangle} = \left( \frac{1}{3} \right)^{\frac{1}{2}} \\
||f|| &= ||x^5|| = \sqrt{\langle f, f \rangle} = \left( \frac{1}{11} \right)^{\frac{1}{2}} \\
\theta &= \cos^{-1} \left( \frac{\sqrt{33}}{7} \right)
\end{aligned}$$

and for ii)

$$\|g\| = \|x^2\| = \sqrt{\langle g, g \rangle} = \left(\frac{1}{5}\right)^{\frac{1}{2}}$$

$$\|f\| = \|x^4\| = \sqrt{\langle f, f \rangle} = \left(\frac{1}{9}\right)^{\frac{1}{2}}$$

$$\theta = \cos^{-1} \left( \frac{\sqrt{45}}{7} \right)$$

3.8) i)

$$\begin{aligned} \langle \cos^2(t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(t) dt = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1 + \cos(2t)}{2} dt = \frac{1}{2\pi} \left( \int_{-\pi}^{\pi} 1 dt + \int_{-\pi}^{\pi} \cos(t) dt \right) \\ &= \frac{1}{2\pi} [t]_{-\pi}^{\pi} = \frac{1}{2\pi} [\pi + \pi] = 1 \\ \langle \cos^2(2t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(2t) dt = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1 + \cos(4t)}{2} dt \quad \text{and because } \cos(4t) = \cos(2t) \\ &= \frac{1}{2\pi} \left( \int_{-\pi}^{\pi} 1 dt + \int_{-\pi}^{\pi} \cos(t) dt \right) \\ &= \frac{1}{2\pi} [t]_{-\pi}^{\pi} = \frac{1}{2\pi} [\pi + \pi] = 1 \\ \langle \sin^2(t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(t) dt = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1 - \cos(2t)}{2} dt = \frac{1}{2\pi} \left( \int_{-\pi}^{\pi} 1 dt - \int_{-\pi}^{\pi} \cos(t) dt \right) \\ &= \frac{1}{2\pi} [t]_{-\pi}^{\pi} = \frac{1}{2\pi} [\pi + \pi] = 1 \\ \langle \sin^2(2t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(2t) dt = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1 - \cos(4t)}{2} dt \quad \text{and because } \sin(4t) = \cos(2t) \\ &= \frac{1}{2\pi} \left( \int_{-\pi}^{\pi} 1 dt + \int_{-\pi}^{\pi} \cos(t) dt \right) \\ &= \frac{1}{2\pi} [t]_{-\pi}^{\pi} = \frac{1}{2\pi} [\pi + \pi] = 1 \end{aligned}$$

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

ii)

$$||t|| = \langle t, t \rangle^{0.5} = \left( \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 dt \right)^{0.5} = \left( \frac{1}{\pi} \times \frac{2\pi^3}{3} \right)^{0.5} = \sqrt{\frac{2}{3}}\pi$$

iii)

$$\begin{aligned}
proj_X(\cos(3t)) &= \sum_{i=1}^4 \langle x_i, \cos(3t) \rangle x_i \\
&= \langle \cos(t), \cos(3t) \rangle \cos(t) + \langle \sin(t), \cos(3t) \rangle \sin(t) \\
&\quad + \langle \cos(2t), \cos(3t) \rangle \cos(2t) + \langle \sin(2t), \cos(3t) \rangle \sin(2t) \\
&= 0 + 0 + 0 + 0 = 0
\end{aligned}$$

iv)

$$\begin{aligned}
proj_X(t) &= \sum_{i=1}^4 \langle x_i, t \rangle x_i \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} t \cos(t) dt \cos(t) + \frac{1}{\pi} \int_{-\pi}^{\pi} t \sin(t) dt \sin(t) \\
&\quad + \frac{1}{\pi} \int_{-\pi}^{\pi} t \cos(2t) dt \cos(2t) + \frac{1}{\pi} \int_{-\pi}^{\pi} t \sin(2t) dt \sin(2t) \\
&= 0 + \frac{2\pi}{\pi} \sin(t) + 0 + -\frac{\pi}{\pi} \sin(2t) \\
&= 2\sin(t) - \sin(2t)
\end{aligned}$$

3.9) If  $R_\theta$  is an orthonormal transformation it holds that  $\langle R_\theta x, R_\theta y \rangle = \langle x, y \rangle = x'y$ .

Here we have:

$$\begin{aligned}
\langle R_\theta x, R_\theta y \rangle &= (R_\theta x)' R_\theta y = x' R_\theta' R_\theta y \\
&= x' \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} y \\
&= x' \begin{pmatrix} \cos(\theta)^2 + \sin(\theta)^2 & \cos(\theta)\sin(\theta) - \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) - \cos(\theta)\sin(\theta) & \sin(\theta)^2 + \cos(\theta)^2 \end{pmatrix} y \\
&= x' I_2 y = x' y
\end{aligned}$$

3.10) i)

$$\begin{aligned}
\langle Qx, Qy \rangle &= \langle x, y \rangle \\
(xQ)^H Qy &= x'y \\
x^H Q^H Qy &= x'y \\
Q^H Qy &= y \\
Q^H Q &= I
\end{aligned}$$

- ii) Be  $Q$  an orthonormal matrix as in i) then  $\|Qx\| = \sqrt{\langle Qx, Qx \rangle} = \sqrt{\langle x, x \rangle} = \|x\|$
- iii)  $\langle Q^{-1}x, Q^{-1}y \rangle = (Q^{-1}x)^H Q^{-1}y = x^H (Q^{-1})^H Q^{-1}y = x^H (Q^{-1}Q^{-1})^H y = x^H (QQ^H)^{-1}y = x^H (Q^H Q)y = x^H y$
- iv) Be  $Q = (q_1 \ q_2 \ \dots \ q_n)$  is an  $(n \times n)$  matrix with orthonormal, that is  $\|q_k\| = 1 = \langle q_k, q_k \rangle$  and  $\langle q_k, q_j \rangle = 0 \ \forall k \neq j$  columns  $q_1$  to  $q_n$  then we know from i) that

$$\begin{aligned}
Q^H Q &= I_n \\
\begin{pmatrix} q_1^H \\ q_2^H \\ \vdots \\ q_n^H \end{pmatrix} (q_1 \ q_2 \ \dots \ q_n) &= I_n \\
\begin{pmatrix} q_1^H q_1 & q_1^H q_2 & \dots & q_1^H q_n \\ q_2^H q_1 & q_2^H q_2 & \dots & q_2^H q_n \\ \vdots & \vdots & & \vdots \\ q_n^H q_1 & q_n^H q_2 & \dots & q_n^H q_n \end{pmatrix} &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}
\end{aligned}$$

v)  $\det(Q)^2 = \det(Q) \det(Q) = \det(Q^T) \det(Q) = \det(Q^T Q) = \det(I) = 1$

vi)

$$\begin{aligned}
\langle Q_1 Q_2 x, Q_1 Q_2 y \rangle &= (Q_1 Q_2 x)^H Q_1 Q_2 y = x^H Q_2^H Q_1^H Q_1 Q_2 y \\
&= x^H Q_2^H I_n Q_2 y = x^H I_n y = x^H y
\end{aligned}$$

3.11 Given a linearly independent set  $\mathbf{X}$ , the Gram Schmidt process produces a new set  $\mathbf{Q}$  that is orthonormal in the same vector space  $\mathbf{V}$  and with respect to the same inner product that constitutes linear independence of  $\mathbf{X}$ .  $q_1$ , the first element of  $\mathbf{Q}$  is simply the first element of  $\mathbf{X}$ ,  $x_1$  divided by its norm so. If  $\mathbf{V} = \mathbb{R}$ ,  $q_1$  has unit length and points in the same direction as  $x_1$ . Staying in this example case, next elements of  $\mathbf{Q}$ ,  $q_k$  are defined as the difference of  $x_k$  and its projection of the first  $k-1$  elements of  $\mathbf{Q}$ , such that  $q_k$  will be orthogonal to the first  $(k-1)$   $q$ 's, similar like the residual of a least squares regression being orthogonal to the regressors. Put differently,  $q_k$  adds a new direction to the set of the first  $(k-1)$   $q$ 's.

3.16) i) We have  $A = QR$  and suppose that  $D = -I$  with  $D = D^{-1}$  such that  $A = \underbrace{QD}_{Q_{neg}} \underbrace{D^{-1}R}_{R_{neg}} = Q_{neg}R_{neg}$  where the subscript *neg* indicates that the diagonal

of these matrices has been multiplied by  $-1$  while all off-diagonal elements are the same as in the original matrices  $Q$  and  $R$ . The QR-decomposition yields unique results up to the sign of the diagonal.

ii) Suppose there exist two QR decompositions  $A = Q_1R_1 = Q_2R_2$  such with positive diagonal elements of  $R_1$  and  $R_2$  then

$$\begin{aligned} Q_1R_1 &= Q_2R_2 \\ Q_2^{-1}Q_1R_1 &= R_2 \\ Q_2^{-1}Q_1 &= R_1^{-1}R_2 \end{aligned}$$

so  $X = R_1^{-1}R_2$  must be an orthonormal matrix because the product of two orthonormal matrices is also orthonormal. Moreover, it must be upper triangular because the product of two upper triangular matrices is also upper triangular. Let  $x_1, x_2, \dots, x_n$  be the columns of that matrix. From orthonormality it follows that  $x_i^T x_j = 0 \quad \forall i \neq j$  and the last  $n-i$  elements of  $x_i$  are zero (upper triangular). For these properties to hold  $X$  must be equal to the identity matrix  $I$  this can be seen from the inner product of the of the first two columns:

$$x_1^T x_2 = \begin{pmatrix} x_{11} & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = x_{11}x_{12} + x_{22}0 = 0 \quad \Longleftrightarrow \quad x_{12} = 0$$

Given  $x_{12} = 0$  it follows that  $x_{13} = x_{23} = 0$  and because of the triangularity of  $X$  this holds for all following inner products, hence  $X = I$  and so  $I = R_1^{-1}R_2 \Leftrightarrow R_1 = R_2 \Rightarrow Q_1 = Q_2$ .

3.17)

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

3.23)

$$\begin{aligned}
||x - y||^2 &= ||x||^2 - \langle x, y \rangle - \langle y, x \rangle + ||y||^2 \\
&= ||x||^2 - \langle x, y \rangle - \overline{\langle x, y \rangle} + ||y||^2 \\
&= ||x||^2 - (\langle x, y \rangle + \overline{\langle x, y \rangle}) + ||y||^2 \\
&\geq ||x||^2 - 2|\langle x, y \rangle| + ||y||^2 \quad (\text{by definition of an inner product}) \\
&\geq ||x||^2 - 2||x|| ||y|| + ||y||^2 \quad (\text{by the Cauchy-Schwartz Inequality}) \\
&\geq (||x|| - ||y||)^2 = (||y|| - ||x||)^2
\end{aligned}$$

and hence  $||x - y|| \geq ||x|| - ||y|| = ||y|| - ||x||$

3.24) For (i), (ii) and (iii) to be norms they have to fulfill a) positivity, b) scale preservation and the c) triangle inequality.

(i) a)  $\int_a^b |f(t)| dt$  obviously satisfies positivity because of the absolute value.

$$b) \int_a^b |af(t)| dt = \int_a^b |a| |f(t)| dt = |a| \int_a^b |f(t)| dt$$

c)

$$\begin{aligned}
||f + g||_{L1} &\leq ||f||_{L1} + ||g||_{L1} \\
\int_a^b |f(t) + g(t)| dt &\leq \int_a^b |f(t)| dt + \int_a^b |g(t)| dt = \int_a^b |f(t)| + |g(t)| dt \\
\Leftrightarrow |f + g| &\leq |f| + |g| \quad \text{which is the actual triangle inequality}
\end{aligned}$$

(ii) a) is obviously given by the absolute value and the square inside the integral.  
b) scale preservation is given by homogeneity of degree one of  $||\cdot||_{L2}$ . c) Similar to ii) c)

$$\begin{aligned}
||f + g||_{L2} &\leq ||f||_{L1} + ||g||_{L2} \\
\left( \int_a^b |f(t) + g(t)|^2 dt \right)^{\frac{1}{2}} &\leq \left( \int_a^b |f(t)|^2 dt \right)^{\frac{1}{2}} + \left( \int_a^b |g(t)|^2 dt \right)^{\frac{1}{2}} \\
&= \int_a^b |f(t)| + |g(t)| dt \\
\Leftrightarrow |f + g| &\leq |f| + |g| \quad \text{which is the actual triangle inequality}
\end{aligned}$$

(iii) a)

$$\begin{aligned}\|f(x)\|_{L^\infty} &= \sup_{x \in [a,b]} |f(x)| = 0 \\ |f(x)| &= 0 \quad \Rightarrow \quad \text{iff} \quad f(x) = 0\end{aligned}$$

b)

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

c)

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

3.26 i)  $\|x\|_2^2 = \sum^n |x_i|^2 \leq (\sum^n |x_i|)^2 = \|x\|_1^2 \Rightarrow \|x\|_2^2 \leq \|x\|_1^2$ . To show the second inequality suppose that  $\mathbf{1}$  be a vector of only ones such that  $\|x\|_1 = |\langle \mathbf{1}, x \rangle| \leq \|\mathbf{1}\|_2 \|x\|_2 = \sqrt{\sum^n 1} \|x\|_2 = \sqrt{n} \|x\|_2$

ii) To show the left inequality let  $|x_1| = \max(x)$  then  $\|x\|_2^2 = \sum^n |x_i|^2 \geq |x_1|^2 = \|x\|_\infty^2 \Rightarrow \|x\|_\infty \leq \|x\|_2$ . For the right inequality suppose, again, that  $x_k$  be the greatest element in  $x$  such that  $\|x\|_\infty = x_k$ . Then consider  $\|x\|_2 = \sqrt{\sum x_i^2} \leq \sqrt{\sum x_k^2} = \sqrt{nx_k^2} = \sqrt{n} x_k = \sqrt{n} \|x\|_\infty$

3.30 Let  $\|A\|_S = \|SAS^{-1}\|$ . To show that this is a norm we show the three properties of a norm hold. a) positivity:  $\text{sign}(\|A\|_S) = \text{sign}(\|SAS^{-1}\|) = +$  because  $\|\cdot\|$  is a norm.

b) homogeneity:  $\|aA\|_S = \|SaAS^{-1}\| = |a| \|SAS^{-1}\|$

c) triangle inequality:  $\|A + B\|_S = \|S(A + B)S^{-1}\| = \|(SA + SB)S^{-1}\| = \|(SAS^{-1} + SBS^{-1})\| \leq \|SAS^{-1}\| + \|SBS^{-1}\|$  because the triangle inequality holds for the norm  $\|\cdot\| \Rightarrow \|A + B\|_S \leq \|A\|_S + \|B\|_S$

3.38

$$\begin{pmatrix} 0 \\ 1 \\ 2x \end{pmatrix} = D(p) \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} \Rightarrow D(p) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$



3.40 i) The *Frobenius inner product* is defined as  $\langle A, B \rangle = \text{tr}(A^H B)$  and for the adjoint of  $A$   $A^*$  is defined by  $\langle B, AC \rangle = \text{tr}(B^H AC) = \text{tr}((A^* B)^H C) = \langle A^* B, C \rangle$ . From  $\text{tr}(B^H AC) = \text{tr}((A^* B)^H C)$  it follows that  $A = (A^*)^H$  and taking the Hermetian conjugate on both sides yields  $A^H = A^*$

ii)

$\langle A_2, A_3 A_1 \rangle = \text{tr}(A_2^H A_3 A_1) = \text{tr}(A_2^H A_1 A_3) = \langle A_2, A_1 A_3 \rangle$  Furthermore, by the definition of the adjoint of  $A_1$  we have  $\langle A_2, A_1 A_3 \rangle = \langle A_1^* A_2, A_3 \rangle = \text{tr}((A_1^* A_2)^H A_3) = \text{tr}((A_2 A_1^*)^H A_3) = \langle A_2 A_1^*, A_3 \rangle$

iii) From the definition of the adjoint we know that  $\langle Y, T_A(X) \rangle = \langle T_A^*(Y), X \rangle$ . It remains to show that  $\langle Y, T_A(X) \rangle = \langle T_{A^*}(Y), X \rangle$ :

$$\begin{aligned} \langle Y, T_A(X) \rangle &= \langle Y, AX - XA \rangle \\ &= \langle Y, AX \rangle - \langle Y, XA \rangle \\ &= \langle A^* Y, X \rangle - \langle Y A^*, X \rangle \\ &= \langle A^* Y - Y A^*, X \rangle \\ &= \langle T_{A^*}(Y), X \rangle \end{aligned}$$

Hence,  $\langle T_A^*(Y), X \rangle = \langle T_{A^*}(Y), X \rangle \iff T_A^* = T_{A^*}$

3.44 Suppose that  $y \in \mathbf{N}(A^H)$  then, by the fundamental theorem of sub spaces we have  $y \in \mathbf{R}(A)^\perp$ . Furthermore, suppose that  $b \notin \mathbf{R}(A)$ . Then  $y$  and  $x$  are not orthogonal and hence  $\langle y, b \rangle \neq 0$

3.45 Let  $X = A + A^T$  such that  $X^T = (A + A^T)^T = X = A + A^T = X$  so  $X \in \text{Sym}(\mathbb{R})$  and  $Y = B - B^T$  such that  $Y^T = (B - B^T)^T = B^T - B = -Y$  so  $Y \in \text{Skew}(\mathbb{R})$  then  $\langle X, Y \rangle = \text{tr}(X^T Y) = \text{tr}((A + A^T)(B - B^T)) = \text{tr}(AB - AB^T + A^T B - A^T B^T) = 0$

3.46 i)  $x \in \mathbf{N}(A^H A)$  implies that  $A^H A x = 0$  from there it follows that  $Ax$  is also in  $\mathbf{N}(A^H)$

ii) From part i) we know that  $Ax \in \mathbf{N}(A^H)$  and  $Ax \in \mathbf{R}(A)$ . From the fundamental theorem of sub-spaces it follows that  $\mathbf{N}(A^H) = \mathbf{R}(A)^\perp$ . Hence,  $\mathbf{R}(A)$  and  $\mathbf{N}(A^H)$  are orthogonal and do not share any common elements except for  $\{0\}$ . Therefore  $Ax = 0$  or  $x \in \mathbf{N}(A)$ . Starting from the left hand side we get the following:  $x \in \mathbf{N}(A) \Rightarrow Ax = 0$ . Premultiplying by  $A^H$  gives  $A^H Ax = 0 \Rightarrow x \in \mathbf{N}(A^H A)$

iii) Since  $A^H$  and  $A$  have the same rank and matrix multiplication is purely linear operation that does not change linear (in)dependence of columns or rows, the matrix product of the two  $A^H A$  must have the same rank as  $A^H$  and  $A$ .

iv) If  $\text{rank}(A) = n = \text{rank}(A^H A)$  and since  $A$  is of dimension  $(m \times n)$  then  $A^H A$  is  $(n \times n)$  that is it has full rank and is invertible. Hence, it is nonsingular.

3.47 i) Show the first inequality  $P^2 = A \underbrace{(A^H A)^{-1} A^H A (A^H A)^{-1} A^H}_{I_n} = A(A^H A)^{-1} A^H =$

$P$ .

ii)  $P^H = (A(A^H A)^{-1} A^H)^H = A [(A^H A)^{-1}]^H A^H = A [(A^H A)^H]^{-1} A^H = A(A^H A)^{-1} A^H = P$

iii)  $\text{rank}(A) = n$  and consists of linear combination of the columns and rows of  $A$  so the rank must be the same.

3.48 i) Let  $c$  be constant then  $P(cA) = \frac{cA+cA^T}{2} = \frac{c(a+A^T)}{s} = cP(A)$  Furthermore,  $P(A+B) = \frac{A+B+(A+B)^T}{2} = \frac{A+B+A^T+B^T}{2} = P(A) + P(B)$  so  $P$  is linear.

ii)  $P(A)^2 = P(P(A)) = \frac{P(A)+P(A)^T}{2} = \frac{\frac{A+A^T}{2} + \frac{A+A^T}{2}}{2} = \frac{A+A^T+(A+A^T)^T}{4} = \frac{2A+2A^T}{4} = \frac{A+A^T}{2} = P(A)$

iii) If from the definition of the adjoint with respect to the Frobenius inner product we know that  $\langle A, P(B) \rangle = \langle P^*(A), B \rangle$ . So  $P(A) = P^*(A) \iff \langle A, P(B) \rangle = \langle P(A), B \rangle$ .

$$\begin{aligned} \langle A, P(B) \rangle &= \left\langle A, \frac{B+B^T}{2} \right\rangle \\ &= \text{tr}\left(\frac{1}{2} A^T (B+B^T)\right) \\ &= \frac{1}{2} \text{tr}(A^T B + A^T B^T) \\ &= \frac{1}{2} (\text{tr}(A^T B) + \text{tr}(B^T A^T)) \\ &= \frac{1}{2} (\text{tr}((A^T B)^T) + \text{tr}(B^T A^T)) \\ &= \frac{1}{2} (\text{tr}(B A^T) + \text{tr}(B^T A^T)) \\ &= \frac{1}{2} (\text{tr}(A^T B^T + A B^T)) \\ &= \frac{1}{2} \text{tr}(B^T (A + A^T)) \\ &= \frac{1}{2} \text{tr}((A + A^T)^T B) \\ &= \left\langle \frac{A + A^T}{2}, B \right\rangle \\ &= \langle P(A), B \rangle = \langle P^*(A), B \rangle \quad (\text{by definition of adjoint}) \\ &\iff P(A) = P^*(A) \end{aligned}$$

iv) Let  $B \in \text{Skew}(\mathbb{R})$  such that  $B^T = -B$  then  $\mathbf{N}(P) = \text{Skew}(\mathbb{R})$  iff  $P(B)^T = 0$ .  $P(B)^T = \left(\frac{B+B^T}{2}\right)^T = \frac{B^T-B}{2} = 0$

v) Let  $C \in \text{Sym}(\mathbb{R})$  that is  $C^T = C$  then  $\mathbf{R}(P) = \text{Sym}(\mathbb{R})$  iff  $P(C)^T = P(C)$ .  $P(C)^T = \frac{(C+C^T)^T}{2} = \frac{C^T+C}{2} = P(C)$

vi)

3.50 Suppose  $r$  and  $s$  are scalars and  $y$  and  $x$  are  $n \times 1$  containing all the data. Be  $\mathbf{1}$  a  $n \times 1$  vector of only ones. Moreover, then let  $x^2$  and  $y^2$  denote the element-wise square, then

$$\begin{aligned}
 sy^2 + rx^2 &= \mathbf{1} \\
 y^2 &= \frac{1}{s}\mathbf{1} - \frac{r}{s}x^2 \\
 \underbrace{y^2}_b &= \underbrace{\begin{pmatrix} \mathbf{1} & x^2 \end{pmatrix}}_A \underbrace{\begin{pmatrix} \frac{1}{s} \\ -\frac{r}{s} \end{pmatrix}}_c \quad (\text{changing notation}) \\
 b &= Ac
 \end{aligned}$$

And the normal equations of such a system are given by  $A^H A \hat{c} = A^H b$  where  $\hat{c}$  is the least squares solution of the system. Since  $x$  and  $y$  are real the normal equations in this case are:

$$\begin{aligned}
 A^T A \hat{c} &= A^T b \\
 \begin{bmatrix} 1 & x^2 \\ x^2 & x^4 \end{bmatrix} \hat{c} &= \begin{pmatrix} \mathbf{1}^T \\ (x^2)^T \end{pmatrix} y^2
 \end{aligned}$$