Math Problem Set # 2 *

Suleyman Gozen

Exercises from **Chapter 3** at "Humpherys, Jeffrey, Tyler J. Jarvis, and Emily J. Evans, Foundations of Applied Mathematics, Volume I: Mathematical Analysis, 2017. SIAM, Philadelphia, PA."

Problem 1. We have that $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|\cos(\theta)$ and $\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos(\theta)$ where $\cos(\theta) = \dfrac{\langle x,y\rangle}{\|x\|\|y\|}$.

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

(ii) $\frac{1}{2}(\|x+y\|^2+\|x-y\|^2) = \frac{1}{2}(\|x\|^2+\|y\|^2+2\|x\|\|y\|\cos(\theta)+\|x\|^2+\|y\|^2-2\|x\|\|y\|\cos(\theta)) = \frac{1}{2}(2(\|x\|^2+\|y\|^2)) = \|x\|^2+\|y\|^2$

Problem 2.

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

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

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

Problem 3.

(i) Let $f(x) = x$ and $g(x) = x^5$. Then

$$\theta = \cos^{-1}\left(\frac{\langle f,g\rangle}{\|f\|\|g\|}\right) = \cos^{-1}\left(\frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx}\sqrt{\int_0^1 x^{10} dx}}\right) = \cos^{-1}\left(\frac{1/7}{\sqrt{1/3}\sqrt{1/11}}\right)$$

(ii) Let $f(x) = x^2$ and $g(x) = x^4$. Then

$$\theta = \cos^{-1}\left(\frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^4 dx}\sqrt{\int_0^1 x^8 dx}}\right) = \cos^{-1}\left(\frac{1/7}{\sqrt{1/5}\sqrt{1/9}}\right)$$

*I thank Yung-Hsu Tsui for his valuable comments.

Problem 8.

(i)

$$\langle \cos(t), \sin(t) \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(t)\sin(t)dt = \frac{1}{\pi} * 0 = 0$$

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

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

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

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

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

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

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

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

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

(ii)

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

(iii)

$$proj_X(\cos(3t)) = \langle \cos(3t), \cos(t) \rangle \cos(t) + \langle \cos(3t), \sin(t) \rangle \sin(t) + \langle \cos(3t), \cos(2t) \rangle \cos(2t)$$
$$+ \langle \cos(3t), \sin(2t) \rangle \sin(2t) = \frac{\cos(t)}{\pi}\int_{-\pi}^{\pi} \cos(t)\cos(3t)dt + \frac{\sin(t)}{\pi}\int_{-\pi}^{\pi} \sin(t)\cos(3t)dt +$$
$$\frac{\cos(2t)}{\pi}\int_{-\pi}^{\pi} \cos(2t)\cos(3t)dt + \frac{\sin(2t)}{\pi}\int_{-\pi}^{\pi} \sin(2t)\cos(3t)dt = 0 + 0 + 0 + 0 = 0$$

(iv)

$$proj_X(t) = \frac{\cos(t)}{\pi}\int_{-\pi}^{\pi} t\cos(t)dt + \frac{\sin(t)}{\pi}\int_{-\pi}^{\pi} t\sin(t)dt + \frac{\cos(2t)}{\pi}\int_{-\pi}^{\pi} t\cos(2t)dt$$
$$+ \frac{\sin(2t)}{\pi}\int_{-\pi}^{\pi} t\sin(2t)dt = \frac{1}{\pi}(0 + 2\pi\sin(t) + 0 - \pi\sin(2t)) = 2\sin(t) - \sin(2t)$$

Problem 9. Let $x = (x_1, x_2)$ and let $y = (y_1, y_2)$. Then $R_\theta x = (\cos(\theta)x_1 - \sin(\theta)x_2, \sin(\theta)x_1 + \cos(\theta)x_2)$ and $R_\theta y = (\cos(\theta)y_1 - \sin(\theta)y_2, \sin(\theta)y_1 + \cos(\theta)y_2)$. If we expand:

$$\begin{aligned}\langle R_\theta x, R_\theta y\rangle =&(\cos(\theta)x_1 - \sin(\theta)x_2) * (\cos(\theta)y_1 - \sin(\theta)y_2)+\\ &(\sin(\theta)x_1 + \cos(\theta)x_2) * (\sin(\theta)y_1 + \cos(\theta)y_2) =\\ &\cos(\theta)^2 x_1y_1 - \sin(\theta)\cos(\theta)x_2y_1 - \cos(\theta)\sin(\theta)x_1y_2 + \sin(\theta)^2 x_2y_2+\\ &\sin(\theta)^2 x_1y_1 + \cos(\theta)^2 x_2y_2 + \cos(\theta)\sin(\theta)x_1y_2 + \cos(\theta)\sin(\theta)x_2y_1 =\\ &(x_1y_1 + x_2y_2)(\cos(\theta)^2 + \sin(\theta)^2) = x_1y_1 + x_2y_2 = \langle x, y\rangle\end{aligned}$$

Problem 10. Let $Q \in M_n(\mathbb{F})$ be an orthonormal matrix.

(i) ($\Rightarrow$) Let $Q \in M_n(\mathbb{F})$ be an orthonormal matrix. Then $\langle Qx, Qy\rangle = \langle x, y\rangle$ which expanding gives $\langle Qx, Qy\rangle = x^H Q^H Q y = x^H y$. This implies that $Q^H Q = I$. By **Proposition 3.2.12**, since Q is an orthonormal operator and $\mathbb{F}^n$ is finite dimensional, Q is invertible. Since inverses are unique, $Q^{-1} = Q^H$ so $Q^H Q = QQ^H = I$.
($\Leftarrow$) Let $Q^H Q = QQ^H = I$. Then $\langle Qx, Qy\rangle = x^H Q^H Q y = x^H I y = x^H y = \langle x, y\rangle$.

(ii) This follows directly from part (i). $\|Qx\| = \sqrt{\langle Qx, Qx\rangle} = \sqrt{\langle x, x\rangle} = \|x\|$.

(iii) By part i we have $Q^{-1} = Q^H$. Then $\langle Q^H x, Q^H y\rangle = x^H QQ^H y = x^H I y = x^H y = \langle x, y\rangle$.

(iv) Consider $\langle Qe_i, Qe_j\rangle = \langle e_i, e_j\rangle = \delta_{ij}$ which is the Kronecker delta. Qe_i is column i of matrix Q, so the dot product of column i with itself is 1 and 0 when $i \neq j$, implying the columns of an orthonormal matrix are orthonormal.

(v) $\det(QQ^H) = \det(I) = 1$. Since $\det(Q) = \det(Q^H)$ we have $\det(Q)^2 = 1$ so $\det(Q) = 1$. The converse is not necessarily true. Consider:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The matrix A is upper triangular and therefore has $\det(A) = 1$ but is not orthonormal because column 2 dot product itself is not equal to 1 which needs to be true by part (iv).

(vi) Let Q_1, Q_2 be orthonormal. Then $\langle Q_1Q_2x, Q_1Q_2y\rangle = x^H Q_2^H Q_1^H Q_1 Q_2 y = x^H Q_2^H I Q_2 y = x^H Q_2^H Q_2 y = x^H y = \langle x, y\rangle$ so the product Q_1Q_2 is orthonormal.

Problem 11. We can apply the Gram-Schmidt orthonormalization process to a collection of linearly dependent vectors so that we are ultimately forced to divide by zero when trying to form the orthonormal vector for the first dependent vector in the set. We can see this by setting the first dependent vector to be the second vector in the set. Then

$$q_1 = \frac{x_1}{\|x_1\|}$$

$$q_2 = \frac{x_2 - p_2}{\|x_2 - p_2\|}$$

where $p_2 = \langle x_2, q_1 \rangle q_1$. However, assuming $x_2 = a_1 x_1$, we have that

$$p_2 = \langle a_1 x_1, \frac{x_1}{\|x_1\|} \rangle \frac{x_1}{\|x_1\|} = \langle \frac{x_1}{\|x_1\|}, \frac{x_1}{\|x_1\|} \rangle a_1 x_1 = a_1 x_1 = x_2$$

so $x_2 - p_2 = 0$.

Problem 16.

(i) The QR factorization is not unique. Consider $QR = QIR = QDD^{-1}R$ where D is a diagonal matrix with -1 on the diagonal. Because D is diagonal, D^{-1} is also diagonal. Diagonal matrices are upper triangular, and upper triangular matrices are closed under multiplication, so $R' = D^{-1}R$ is upper triangular. Additionally, $Q' = QD$ is also orthonormal because it is equivalent to Q multiplied by -1. Therefore $Q'R' = QR$ and the QR decomposition is not unique.

(ii) Let A be invertible and let there be a QR decompositions $A = Q_1 R_1 = Q_2 R_2$. Then $Q_2^H Q^1 = R_2 R_1^{-1}$. Both orthonormal and upper triangular matrices are closed under multiplication so the matrix $Q_2^H Q^1$ is both upper triangular and orthonormal. And upper triangular matrix that is orthonormal is a diagonal matrix with either 1 or -1 on the diagonal. Both R_2 and R_1 have positive diagonal elements so the matrix $Q_2^H Q^1$ must positive 1 on the diagonal. Therefore $I = Q_2^H Q^1 = R_2 R_1^{-1}$ so $R_2 = R_1$.

Problem 17. Let $A = \widehat{Q}\widehat{R}$ be reduced QR decomposition. Then the system $A^H A x = A^H b$ can be rewritten as

$$\begin{aligned}
&(\widehat{Q}\widehat{R})^H (\widehat{Q}\widehat{R}) x = (\widehat{Q}\widehat{R})^H b \\
&\Leftrightarrow \widehat{R}^H \widehat{Q}^H \widehat{Q} \widehat{R} x = \widehat{R}^H \widehat{Q}^H b \\
&\Leftrightarrow \widehat{R}^H \widehat{R} x = \widehat{R}^H \widehat{Q}^H b \\
&\Leftrightarrow \widehat{R} x = \widehat{Q}^H b
\end{aligned}$$

Problem 23. We have that $\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$ so $\|x\| - \|y\| \leq \|x - y\|$. Similarly, $\|y\| - \|x\| \leq \|x - y\|$ since $\|x - y\| = \|y - x\|$. Putting these together implies that $|\|x\| - \|y\|| \leq \|x - y\|$ because if $\|x\| > \|y\|$ then $|\|x\| - \|y\|| = \|x\| - \|y\|$ and else $|\|x\| - \|y\|| = \|y\| - \|x\|$.

Problem 24. We need to show positivity, scale preservation, and the triangle inequality. Note that the absolute value $|f(t)|$ is nonnegative and only equal to 0 if the function $f(t) = 0$.

(1) (i) By the note above, the integrand $|f(t)|$ is nonnegative. Therefore the integral is non negative. The integral of a non-negative function is only equal to 0 on $[a, b]$ if the function is equal to 0 on $[a, b]$.

(ii) $\|af\|_{L_1} = \int_a^b |af(t)| dt = \int_a^b |a||f(t)| dt = |a| \int_a^b |f(t)| dt = |a| \|f\|_{L_1}$

(iii) $\|f + g\|_{L_1} = \int_a^b |f(t) + g(t)| dt \leq \int_a^b (|f(t)| + |g(t)|) dt = \|f\|_{L_1} + \|g\|_{L_1}$

(2) (i) By the note above, the integrand $|f(t)|^2$ is nonnegative. Therefore the integral is non negative. The integral of a non-negative function is only equal to 0 on $[a, b]$ if the function is equal to 0 on $[a, b]$. Furthermore, the square root is also non-negative.

(ii) $\|af\|_{L_2} = \left(\int_a^b |af(t)|^2 dt\right)^{\frac{1}{2}} = \left(\int_a^b a^2|f(t)|^2 dt\right)^{\frac{1}{2}} = |a|\left(\int_a^b |f(t)|^2 dt\right)^{\frac{1}{2}} = |a|\|f\|_{L_2}$

(iii) $\|f+g\|_{L_2}^2 = \int_a^b |f(t)+g(t)|^2 dt \leq \int_a^b (|f(t)| + |g(t)|)^2 =$

$= \int_a^b (|f(t)|^2 + |g(t)|^2 + 2|f(t)||g(t)|)dt = (\|f\|_{L_2} + \|g\|_{L_2})^2$

Therefore $\|f+g\|_{L_2} \leq \|f\|_{L_2} + \|g\|_{L_2}$.

(3) (i) By the note above, $|f(t)|$ is nonnegative. Then $\sup_{x\in[a,b]} |f(x)| = 0$ if and only if $|f(x)| = 0$ because if $|f(x)| > 0$ for some $x \in [a, b]$ then by definition $\sup_{x\in[a,b]} |f(x)| > 0$.

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

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

Problem 26. To show that topological equivalence is an equivalence relation we show three things: reflexivity, symmetry, and transitivity. Let $0 < m \leq M$.

(i) Let $m = \frac{1}{2}$ and $M = 2$. Then $m\|\cdot\|_a \leq \|\cdot\|_a \leq M\|\cdot\|_a$ so $\|\cdot\|_a \sim \|\cdot\|_a$.

(ii) Let $\|\cdot\|_a \sim \|\cdot\|_b$ so $m\|\cdot\|_a \leq \|\cdot\|_b \leq M\|\cdot\|_a$. Then it follows form the previous inequality that $\frac{1}{M}\|\cdot\|_b \leq \|\cdot\|_a \leq \frac{1}{m}\|\cdot\|_b$ so $\|\cdot\|_a \sim \|\cdot\|_b$ if and only if $\|\cdot\|_b \sim \|\cdot\|_a$.

(iii) Let $\|\cdot\|_a \sim \|\cdot\|_b$ and Let $\|\cdot\|_b \sim \|\cdot\|_c$. Then $m_1\|\cdot\|_a \leq \|\cdot\|_b \leq M_1\|\cdot\|_a$ and $m_2\|\cdot\|_b \leq \|\cdot\|_c \leq M_2\|\cdot\|_b$. It follows then that $m_1m_2\|\cdot\|_a \leq m_2\|\cdot\|_b \leq M_2\|\cdot\|_b \leq M_1M_2\|\cdot\|_a$ and substituting in the second inequality we have that $m_1m_2\|\cdot\|_a \leq m_2\|\cdot\|_b \leq \|\cdot\|_c \leq M_2\|\cdot\|_b \leq M_1M_2\|\cdot\|_a$ so $\|\cdot\|_a \sim \|\cdot\|_c$.

We now show that the p-norms for $p = 1, 2, \infty$ on $\mathbb{F}^n$ are topologically equivalent.

(i) The first inequality follows from $\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \leq \sqrt{\sum_{i=1}^n |x_i|^2 + 2\sum_{i<j} |x_i||x_j|} = \sum_i^n \sqrt{x_i^2} = \sum_i^n |x_i| = \|x\|_1$. For the second inequality consider a vector of 1's. Then by Cauchy-Schwarz: $\|x\|_1 = \sum_i^n |1x_i| \leq \sqrt{\sum_i^n |1| \sum_i^n |x_i|} = \sqrt{n}\sqrt{\sum_i^n |x_i|} = \sqrt{n}\|x\|_2$ so $\|x\|_1 \sim \|x\|_2$

(ii) Let $\max_i |x_i| = |x_j|$. Then $\|x\|_\infty = \max_i |x_i| = |x_j| \leq \sqrt{\sum_{i=1}^{n} |x_i|^2} = \|x\|_2 \leq \sqrt{n|x_j|^2} = \sqrt{n}\|x\|_\infty$ so $\|x\|_2 \sim \|x\|_\infty$

Problem 28. Applying the properties proven in problem 26 we have that:

(i)

$$\frac{1}{\sqrt{n}}\|A\|_2 = \sup_{x\neq 0} \frac{1}{\sqrt{n}} \frac{\|Ax\|_2}{\|x\|_2} \leq \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_1} \leq \sup_{x\neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \|A\|_1$$

$$\leq \sup_{x\neq 0} \sqrt{n}\frac{\|Ax\|_2}{\|x\|_2} = \sqrt{n}\|A\|_2$$

(ii)

$$\frac{1}{\sqrt{n}}\|A\|_\infty = \sup_{x\neq 0} \frac{1}{\sqrt{n}} \frac{\|Ax\|_\infty}{\|x\|_\infty} \leq \sup_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_2} \leq \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \|A\|_2$$

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

Problem 29.

(i) We proved in problem 10 that $\|Qx\|_2 = \|x\|_2$ if Q is orthonormal. Then $\|Q\|_2 = \sup_{x\neq 0} \frac{\|Qx\|_2}{\|x\|_2} = 1$.

(ii) The induced norm of the transformation R_x is $\|R_x\|_2 = \sup_{A\neq 0} \frac{\|Ax\|_2}{\|A\|_2} \leq \sup_{A\neq 0} \frac{\|A\|_2\|x\|_2}{\|A\|_2} = \|x\|_2$. Now consider $\|R_x\|_2 = \sup_{A\neq 0} \frac{\|Ax\|_2}{\|A\|_2} \geq \frac{\|Ix\|_2}{\|I\|} = \|x\|_2$ and hence $\|Rx\|_2 = \|x\|_2$.

Problem 30. To show that $\|\cdot\|_S$ is a matrix norm we need to show the following properties:

(i) Positivity: $\|A\|_S = \|S(A)S^{-1}\|$ where $\|\cdot\|$ is norm, which being a norm is nonnegative.

(ii) Scale: $\|\alpha A\|_S = \|S(\alpha A)S^{-1}\| = |\alpha|S(A)S^{-1}\| = |\alpha|\|A\|_S$

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

(iv) Submultiplicativity: $\|AB\|_S = \|S(AB)S^{-1}\| = \|SAIBS^{-1}\| = \|SAS^{-1}SBS^{-1}\| \leq \|SAS^{-1}\|\|SBS^{-1}\| = \|A\|_S\|B\|_S$

Problem 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,

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

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

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

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

Problem 38. Let $D : V \to V$ be the derivative operator where $V = \mathbb{F}[x; 2]$ is a subspace of the inner product space $L^2([0, 1]; \mathbb{R})$. Given the basis $[1, x, x^2]$ we can write the matrix representation of D:

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

Furthermore, the adjoint is the Hermitian, so we can write the matrix representation of D^*:

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

Problem 39.

(i) (1) $\langle v, (S + T)^*w \rangle = \langle (S + T)v, w \rangle = \langle Sv, w \rangle + \langle Tv, w \rangle = \langle v, S^*w \rangle + \langle v, T^*w \rangle = \langle v, (S^* + T^*)w \rangle$

(2) $\langle v, (\alpha T)w \rangle = \alpha \langle (T^*)v, w \rangle = \langle \bar{\alpha}(T^*)v, w \rangle$

(ii) $\langle v, S^*w \rangle = \langle Sv, w \rangle$ so $S^{*^*} = S$.

(iii) $\langle v, (ST)^*w \rangle = \langle (ST)v, w \rangle = \langle (T)v, (S^*)w \rangle = \langle v, (T^*S^*)w \rangle$

(iv) First, $\langle T^*(T^{-1})^*v, w \rangle = \langle v, T^{-1}Tw \rangle = \langle v, w \rangle$. Second, $\langle (T^{-1})^*T^*v, w \rangle = \langle v, TT^{-1}w \rangle = \langle v, w \rangle$. Therefore, $T^*(T^{-1})^* = (T^{-1})^*T^* = I$ so $(T^*)^{-1} = (T^{-1})^*$.

Problem 40.

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

$$< B, AC >_F = \text{tr}(B^H AC) = \text{tr}((A^H B)^H C) =< A^H B, C >_F .$$

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

$$< A_2, A_3A_1 >_F = \text{tr}(A_2^H A_3A_1) = \text{tr}(A_1A_2^H A_3) = \text{tr}((A_2A_1^H)^H A_3) =< A_2A_1^H, A_3 >_F =< A_2A_1^*, A_3 > .$$

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

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

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 $y \in \mathcal{N}(A^H)$,

$$< y, b >=< y, Ax >=< A^H y, x >=< 0, x >= 0.$$

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

Problem 45. We define $Sym_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) | A^T = A\}$ and $Skew_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) | A^T = -A\}$. Let $A \in Sym_n(\mathbb{R})$ and $B \in Skew_n(\mathbb{R})$. Then $\langle A, B\rangle = tr(A^T B) = tr(BA^T) = -tr(B^T A) = -\langle B, A\rangle = -\langle A, B\rangle$ implying that $\langle A, B\rangle = 0$. Furthermore, any matrix $A \in M_N(\mathbb{R})$ can be written as the sum of a Skew and Sym matrix where $A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$ which shows that $Sym_n(\mathbb{R})$ and $Skew_n(\mathbb{R})$ are orthogonal complements and $Sym_n(\mathbb{R})^\perp = Skew_n(\mathbb{R})$.

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) Let A have linearly independent columns. Then $rank(A) = n = rank(A^H A)$ and by part (iii). Since $A^H A$ is a square n-by-n matrix of rank n, it is nonsingular.

Problem 47. Let $P = A(A^H A)^{-1} A^H$.

(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

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

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

Problem 48.

(i) To show linearity we show two things:

(1) $P(A+B) = \dfrac{A+B+(A+B)^T}{2} = \dfrac{A+B+A^T+B^T}{2} = \dfrac{A+A^T}{2} + \dfrac{B+B^T}{2} = P(A) + P(B)$

(2) $P(\alpha A) = \dfrac{\alpha A + \alpha A^T}{2} = \alpha \dfrac{A+A^T}{2} = \alpha P(A)$

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

(iii) $\langle P(A), B\rangle = \dfrac{1}{2}tr(AB + A^T B) = \dfrac{1}{2}tr(A^T B + A^T B^T) = \langle A, P(B)\rangle$

(iv) Let $A \in Skew_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) | A^T = -A\}$. Then $P(A) = \dfrac{1}{2}(A+A^T) = \dfrac{1}{2}(A-A) = 0$ so $A \in N(P)$.

(v) Let $A \in Sym_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) | A^T = A\}$. Then $P(A) = \dfrac{1}{2}(A+A^T) = \dfrac{1}{2}(A+A) = A$ so $A \in R(P)$.

(vi) $\|A - P(A)\|_F = \sqrt{\langle \frac{1}{2}(A - A^T), \frac{1}{2}(A - A^T)\rangle} = \sqrt{\frac{1}{4}tr((A^T - A)(A - A^T)} =$
$= \sqrt{\frac{1}{4}tr(A^T A - A^2 + AA^T - A^T A^T)} = \sqrt{\frac{1}{4}tr(2A^T A - A^2 + AA^T - A^2)} =$
$= \sqrt{\frac{1}{2}(tr(A^T A) - tr(A^2))}$

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

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