Problem Set 2

OSM Lab-Math
Sophia Mo

Problem 1 (3.1)

In a real product space,
$||x+y||^2 =< x+y, x+y >=< x,x > + < x,y > + < y,x > + < y,y >=$
$||x||^2 + ||y||^2 + 2 < x,y >$

$||x-y||^2 =< x-y, x-y >=< x,x > - < x,y > - < y,x > + < y,y >=$
$||x||^2 + ||y||^2 - 2 < x,y >$

$\Rightarrow \frac{1}{2}(||x+y||^2 + ||x-y||^2) =< x,x > + < y,y >= ||x||^2 + ||y||^2$
$\Rightarrow \frac{1}{4}(||x+y||^2 - ||x-y||)^2 =< x,y >$

Problem 2 (3.2)

In a complex product space,
$||x+y||^2 =< x+y, x+y >=< x,x > + < y,x > + < x,y > + < y,y >$
$||x-y||^2 =< x-y, x-y >=< x,x > - < y,x > - < x,y > + < y,y >$
$||x+y||^2 - ||x-y||^2 = 2 < x,y > +2 < y,x >= 2 < x,y > +2\overline{< x,y >} = 4Re < x,y >$

$||x+iy||^2 =< x+iy, x+iy >=< x,x > +i < y,x > +i < x,y > - < y,y >$
$||x-iy||^2 =< x-iy, x-iy >=< x,x > -i < y,x > -i < x,y > - < y,y >$
$i(||x-iy||^2 - ||x+iy||^2) = i(-2i < x,y > -2i < y,x >= -i(2i < x,y >$
$-2i\overline{< x,y >}) = -4iIm < x,y >$

$\Rightarrow \frac{1}{4}(||x+y||^2 - ||x-y||^2 + i||x-iy||^2 - i||x+iy||^2) = Re < x,y > +iIm <$
$x,y >=< x,y >$

Problem 3 (3.3)

(i) $cos(\theta) = \frac{<x,x^5>}{||x||||x^5||} = \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx \int_0^1 x^{10} dx}} = \frac{\sqrt{33}}{7}, \theta \approx 0.608$

(ii) $cos(\theta) = \frac{<x^2,x^4>}{||x^2||||x^4||} = \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^4 dx \int_0^1 x^8 dx}} = \frac{3\sqrt{5}}{7}, \theta \approx 0.290$

Problem 4 (3.8)

(i)
$< cos(t), cos(t) >= \frac{1}{\pi}\int_{-\pi}^{\pi} cos^2(t)dt = 1; < sin(t), sin(t) >= \frac{1}{\pi}\int_{-\pi}^{\pi} sin^2(t)dt = 1$
$< cos(2t), cos(2t) >= 1; < sin(2t), sin(2t) >= 1$
$< sin(t), cos(t) >= \frac{1}{\pi}\int_{-\pi}^{\pi} sin(t)cos(t)dt = 0$
$< sin(t), sin(2t) >= \frac{1}{\pi}\int_{-\pi}^{\pi} sin(t)sin(2t)dt = 0$
$< sin(t), cos(2t) >= \frac{1}{\pi}\int_{-\pi}^{\pi} sin(t)cos(2t)dt = 0$
$< cos(t), sin(2t) >= \frac{1}{\pi}\int_{-\pi}^{\pi} cos(t)sin(2t)dt = 0$
$< cos(t), cos(2t) >= \frac{1}{\pi}\int_{-\pi}^{\pi} cos(t)cos(2t)dt = 0$

$< sin(2t), cos(2t) >= \frac{1}{\pi}\int_{-\pi}^{\pi} cos(t)cos(2t)dt = 0$
So S is an orthonormal basis.

(ii)
$||t|| = \sqrt{< t, t >} = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} t^2 dt} = \sqrt{\frac{2}{3}}\pi$

(iii)
$proj_X(cos(3t)) = \sum_{s\in S} < s, cos(3t) > \frac{s}{||s||^2} =$
$\frac{1}{\pi}[cos(t)\int_{-\pi}^{\pi} cos(t)cos(3t)dt + sin(t)\int_{-\pi}^{\pi} sin(t)cos(3t)dt + cos(2t)\int_{-\pi}^{\pi} cos(2t)cos(3t)dt + sin(2t)\int_{-\pi}^{\pi} sin(2t)cos(3t)dt] = 0$

(iv)
$proj_X t = \sum_{s\in S} < s, t > \frac{s}{||s||^2} =$
$\frac{1}{\pi}[cos(t)\int_{-\pi}^{\pi} cos(t)tdt + sin(t)\int_{-\pi}^{\pi} sin(t)tdt + cos(2t)\int_{-\pi}^{\pi} cos(2t)tdt + sin(2t)\int_{-\pi}^{\pi} sin(2t)tdt] = 2sin(t) - sin(2t)$

Problem 5 (3.9)
Let $x = (x_1, x_2), y = (y_1, y_2)$ be any two elements in the inner product space $\mathbb{R}^2$. Denote the transformation under rotation by L. We know $< x, y >= x_1y_1 + x_2y_2$
$\Rightarrow L(x) = (x_1cos\theta - x_2sin\theta, x_1sin\theta + x_2cos\theta), L(y) = (y_1cos\theta - y_2sin\theta, y_1sin\theta + y_2cos\theta)$
$\Rightarrow < L(x), L(y) >= (x_1cos\theta - x_2sin\theta)(y_1cos\theta - y_2sin\theta) + (x_1sin\theta + x_2cos\theta)(y_1sin\theta + y_2cos\theta) = x_1y_1 + x_2y_2 =< x, y >$
So by definition, a rotation is an orthonormal transformation.

Problem 6 (3.10)
(i)
Suppose $Q^HQ = QQ^H = I$, then $\forall x, y \in M_n(\mathbb{F}), < Qx, Qy >= (Qx)^HQy = x^HQ^HQy = x^H(Q^HQ)y = x^HIy = x^Hy =< x, y >$
So Q is an orthonormal matrix.

Suppose Q is an orthonormal matrix, then $\forall x, y \in M_n(\mathbb{F}), < Q^HQx, y >= x^HQ^HQy =< QQ^HQx, Qy >= x^HQ^HQQ^HQy$
Since the equation holds for any elements, cancellative property holds.
$\Rightarrow QQ^H = I$
Similarly, $< QQ^Hx, y >= x^HQQ^Hy =< QQQ^Hx, Qy >= x^HQQ^HQ^HQy$
$\Rightarrow Q^HQ = I$

(ii)
$||x||^2 =< x, x >, ||Qx||^2 =< Qx, Qx >$.
By the definition of orthonormal basis, $< x, x >=< Qx, Qx >$, so$||x|| = ||Qx||$

(iii)
By (i), we know that $Q^HQ = QQ^H = I$
$\Rightarrow (Q^H)^{-1}Q^HQ = (Q^{-1})^HQ^HQ = (Q^{-1})^HQQ^H$

$\Rightarrow Q = (Q^{-1})^H(QQ^H) \Rightarrow Q^{-1}Q = Q^{-1}(Q^{-1})^H, QQ^{-1} = (Q^{-1})^H Q^{-1}$
$\Rightarrow I = Q^{-1}(Q^{-1})^H = (Q^{-1})^H Q^{-1}$
By (i), Q^{-1} is an orthonormal matrix.

(iv)
Denote the columns of Q by q_1, q_2, q_3. Since Q is an orthonormal matrix, by (i), we know that $Q^H Q = QQ^H = I$
$\Rightarrow q_1^2 = q_2^2 = q_3^2 = 0, q_i q_j = 0, \forall i \neq j$
q_1, q_2, q_3 are orthonormal, by definition.

(v)
Since $Q^H Q = QQ^H = I, Q^H = Q^{-1}$
We know $det(Q^H) = det(Q), det(Q)det(Q^{-1}) = det(QQ^{-1}) = det(I) = 1$
$\Rightarrow (det(Q))^2 = 1 \Rightarrow ||det(Q)|| = 1$
The inverse of the statement is not true. For example, consider a counterexample:
$P = \begin{pmatrix} 1 & 2 \\ \frac{1}{2} & 1 \end{pmatrix}, det(P) = 1, PP^H \neq I.$

(vi)
We know $Q_1^H Q_1 = Q_1 Q_1^H = Q_2^H Q_2 = Q_2 Q_2^H = I$
$\Rightarrow Q_1 Q_2 (Q_1 Q_2)^H = Q_1 Q_2 Q_2^H Q_1^H = Q_1 (Q_2 Q_2^H) Q_1^H = Q_1 Q_1^H = I$
$(Q_1 Q_2)^H Q_1 Q_2 = Q_2^H Q_1^H Q_1 Q_2 = Q_2^H (Q_1^H Q_1) Q_2 = Q_2^H Q_2 = I$
By (i), $Q_1 Q_2$ is an orthonormal matrix.

Problem 7 (3.11)
Let $\{x_i\}_{i=1}^n$ be a collection of linearly dependent vectors. This means that $\exists x_j \in \{x_i\}$ such that $x_j \in span\{x_1, x_2, ..., x_{j-1}\}$, so by projecting x_j onto $span\{q_1, q_2, ...q_{j-1}\}$, the process will map x_j into 0, so we can drop it and move on to x_{j+1}. In general, if the collection of linearly dependent vectors have at most m linearly independent vectors, the Gram-Schmidt Process will yield an orthonormal set with m elements.

Problem 8 (3.16)
(i)
Consider a 2-by-2 matrix A with linearly independent columns. By Theorem 3.3.9, we know A can be decomposed into QR. Let $D = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. Let $Q' = QD = -Q$ (still an orthonormal matrix), $R' = DR = -R$ (still upper-triangular). We know that $Q'R'$ is another decomposition of A. So QR decomposition is not unique.

(ii)
Suppose that there are two different decompositions Q_1R_1, Q_2R_2 of A, and R_1, R_2 both have only positive diagonal elements. Then $Q_1R_1 = A = Q_2R_2 \Rightarrow Q_2^H Q_1 = Q_2^H A R_1^{-1} = R_2 R_1^{-1}$.
Since the set of upper triangular matrices is closed under multiplication, we know that $Q_2 H R_1^{-1}$ is upper triangular. Since R_1, R_2 have only positive entries, $Q_2 H R_1^{-1}$

must have only positive entries. Since Q_2, Q_1 are orthonormal matrices, by 3.10 we know that $Q_2 H R_1^{-1}$ must be an orthonormal matrix. For these conditions to hold, $Q_2 H R_1^{-1}$ must be the identity matrix. $\Rightarrow Q_2^H Q_1 = I \Rightarrow Q_2 = Q_1$, so the QR decomposition of A such that R has onyly positive diagonal elements is unique.

Problem 9 (3.17)

Let $A = \hat{Q}\hat{R}$ be a reduced QR decomposition. Then
$A^H Ax = \hat{R}^H \hat{Q}^H \hat{Q}\hat{R}x = \hat{R}^H(\hat{Q}^H\hat{Q})\hat{R}x = \hat{R}^H\hat{R}x$
$A^H b = \hat{R}^H \hat{Q}^H b.$
So $A^H Ax = A^H b \Leftrightarrow \hat{R}x = \hat{Q}^H b$

Problem 10 (3.23)

By the definition of a norm, $||x-y|| + ||y|| \geq ||(x-y)+y|| = ||x||, ||x-y|| + ||x|| = ||y-x|| + ||x|| \geq ||(y-x)+x|| = |||y|$
$\Rightarrow ||x-y|| \geq ||x|| - ||y||, ||x-y|| \geq ||y|| - ||x||$
$\Rightarrow |||x|| - ||y||| \leq ||x-y||$

Problem 11 (3.24)

(i)
$\forall f \in C, \int_a^b |f(t)|dt \geq 0$ because $|f(t)| > 0$. $\int_a^b |f(t)|dt = 0$ iff $f(t) = 0, \forall t \in [a,b]$
$\forall c \in \mathbb{F}, \int_a^b |cf(t)|dt = |c| \int_a^b |f(t)|dt$
$\forall f, g \in C, \int_a^b |f(t)|dt + \int_a^b |g(t)|dt = \int_a^b |f(t)| + |g(t)|dt \geq \int_a^b |f(t) + g(t)|dt$
So by definition, this is a norm.

(ii)
$\forall f \in C, (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} \geq 0$ because $|f(t)| > 0$. $(\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} = 0$ iff $f(t) = 0$.
$\forall c \in \mathbb{F}, (\int_a^b |cf(t)|^2 dt)^{\frac{1}{2}} = |c|(\int_a^b |f(t)|^2 dt)^{\frac{1}{2}}$
$\forall f, g \in C, (\int_a^b |f(t) + g(t)|^2 dt)^{\frac{1}{2}} = (\int_a^b |f(t)|^2 + |g(t)|^2 + 2f(t)g(t)dt)^{\frac{1}{2}}$
By Cauchy-Schwartz Inequality, with respect to the inner product $< f, g >= \int fg$
$\int_a^b f(t)g(t)dt \leq (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} (\int_a^b |g(t)|^2 dt)^{\frac{1}{2}}$
$\Rightarrow (\int_a^b |f(t) + g(t)|^2 dt)^{\frac{1}{2}} \leq (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} + (\int_a^b |g(t)|^2 dt)^{\frac{1}{2}}$
So by definition, this is a norm.

(iii)
$\forall f \in C, sup_{x \in [a,b]}|f(x)| \geq 0$ because $|f(t)| > 0$. $sup_{x \in [a,b]}|cf(x)| = 0$ iff $f(t) = 0$.
$\forall c \in \mathbb{F}, sup_{x \in [a,b]}|cf(x)| = |c| sup_{x \in [a,b]}|f(x)|$
$\forall f, g \in C, sup_{x \in [a,b]}|f(x) + g(x)| \leq sup_{x \in [a,b]}|f(x)| + sup_{x \in [a,b]}|g(x)|$
So by definition, this is a norm.

Problem 12 (3.26)

Let $||.||_a, ||.||_b, ||.||_c$ be three norms on the vector space X.
1) $\exists m = M = 1$s.t. $m||x||_a \leq ||x||_a \leq M||x||_a, \forall x \in X$, so reflexivity is satisfied.

2) Suppose $||.||_a$ is topologically equivalent to $||.||_b$,then $\exists 0 < m \leq M$ s.t. $m||.||_a \leq ||x||_b \leq M||x||_a, \forall x \in X$
$\exists 0 < \frac{1}{M} \leq \frac{1}{m}$ s.t. $\frac{1}{M}||.||_b \leq ||x||_a \leq \frac{1}{m}||x||_b, \forall x \in X$.
$||.||_b$ is topologically equivalent to $||.||_a$, so symmetry is satisfied.

3) Suppose $||.||_a$ and $||.||_b$ are topologically equivalent, and $||.||_b, ||.||_c$ are topologically equivalent. Then $\exists 0 < m_1 \leq M_1, 0 < m_2 \leq M_2$ s.t. $m_1||.||_a \leq ||x||_b \leq M_1||x||_a, m_2||.||_b \leq ||x||_c \leq M_2||x||_b, \forall x \in X$.
$\Rightarrow \exists 0 < m_1m_2 \leq M_1M_2$ s.t. $m_1m_2||x||_a \leq m_2||x||_b \leq ||x||_c \leq M_2||x||_b \leq M_1M_2||x||_a$
$\Rightarrow ||.||_a, ||.||_c$ are topologically equivalent. So transitivity is satisfied.
Topological equivalence is thus an equivalence relation.

(i)
$||x||_1 = \sum_{i=1}^n |x_i|, ||x||_2 = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}}$
$||x||_1 = \sum_{i=1}^n |x_i|^2 \leq (\sum_{i=1}^n |x_i|)^2 = ||x||_2$ because the right hand side contains the terms of the left hand side, and all terms are positive.
By Cauchy-Schwartz Inequality,
$||x||_1 = \sum_{i=1}^n |x_i| = \sum_{i=1}^n |x_i| \times 1 \leq (\sum_1^n |x_i^2|)^{\frac{1}{2}}(\sum_1^n 1)^{\frac{1}{2}} = \sqrt{n}||x||_2$

(ii)
$||x||_\infty = max_i(|x_i|) \leq (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}} = ||x||_2$, because the square root of the sum contains the term $max_i(|x_i|)$
$||x||_2 = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}} \leq (\sum_{i=1}^n max_i|x_i|)^{\frac{1}{2}} = \sqrt{n}||x||_\infty$

Therefore, $p = 1, 2, \infty$ on $\mathbb{F}^n$ are topologically equivalent.

Problem 13 (3.28)
Let $A = [\alpha_{ij}]$ be an $n \times n$ matrix. Then $||A||_p = sup_{x \neq 0} \frac{||Ax||_p}{||x||_p}$
(i)
By 3.26 (i), we know
$||A||_2 = sup_{x \neq 0} \frac{||Ax||_2}{||x||_2} \leq sup_{x \neq 0} \frac{||Ax||_1}{||x||_2} \leq sup_{x \neq 0} \frac{||Ax||_1}{\frac{1}{\sqrt{n}}||x||_1} = \sqrt{n} sup_{x \neq 0} \frac{||Ax||_1}{||x||_1}$
$\Rightarrow \frac{1}{\sqrt{n}}||A||_2 \leq ||A||_1$
$||A||_1 = sup_{x \neq 0} \frac{||Ax||_1}{||x||_1} \leq sup_{x \neq 0} \frac{\sqrt{n}||Ax||_2}{||x||_1} \leq sup_{x \neq 0} \frac{||Ax||_2}{||x||_2}$
$\Rightarrow ||A||_1 \leq \sqrt{n}||A||_2$

(ii)
By 3.26 (ii), we know
$||A||_\infty = sup_{x \neq 0} \frac{||Ax||_\infty}{||x||_\infty} \leq sup_{x \neq 0} \frac{||Ax||_2}{||x||_\infty} \leq sup_{x \neq 0} \frac{||Ax||_2}{\frac{1}{\sqrt{n}}||x||_2} = \sqrt{n} sup_{x \neq 0} \frac{||Ax||_2}{||x||_2}$
$\Rightarrow \frac{1}{\sqrt{n}}||A||_\infty \leq ||A||_2$
$||A||_2 = sup_{x \neq 0} \frac{||Ax||_2}{||x||_2} \leq sup_{x \neq 0} \frac{\sqrt{n}||Ax||_\infty}{||x||_2} \leq sup_{x \neq 0} \frac{||Ax||_\infty}{||x||_\infty}$
$\Rightarrow ||A||_2 \leq \sqrt{n}||A||_\infty$

Problem 14 (3.29)
$\forall$ orthonormal matrix Q, as proved in 3.10, $||Qx|| = ||x|| = ||Q||||x|| \Rightarrow ||Q|| = 1$. $\forall x \in \mathbb{F}^n$, let $R_x : M_n(\mathbb{F}) \to \mathbb{F}^n$ be the transformation $A \mapsto Ax$.
The induced norm is $||R_x|| = sup_{A\neq 0}\frac{||Ax||_2}{||A||_2}$, where $||A||_2 = sup_{y\neq 0}\frac{||Ay||_2}{||y||_2} \geq \frac{||Ax||_2}{||x||_2}, \forall x$.
$\Rightarrow ||R_x|| = sup_{x\neq 0}\frac{||Ax||_2}{||A||_2} \leq \frac{||Ax||_2||x||_2}{||Ax||_2} = ||x||_2$
Moreover, $||R_x|| = ||x||_2 sup_{x\neq 0}\frac{||A(x/||x||_2)||_2}{||A||_2}$. According to the hint, since $x/||x||_2$ has norm 1, it is the first column of some orthonormal matrix B. We know then that $\frac{||B^H(x/||x||_2)||_2}{||B^H||_2} = 1$, because B^H is also an orthonormal matrix, with norm 1. It follows that $||R_x|| = ||x||_2 sup_{x\neq 0}\frac{||A(x/||x||_2)||_2}{||A||_2} \geq ||x||_2$
$\Rightarrow ||R_x|| = ||x||_2$

Problem 15 (3.30)
1) $||A||_S = ||SAS^{-1}|| \geq 0$. $||A||_S = 0 \Leftrightarrow ||SAS^{-1}|| = 0 \Leftarrow A = 0$, since S is invertible ($\neq 0$), and $||.||$ is a norm.
2) $\forall c \in \mathbb{F}, ||cA||_S = ||ScAS^{-1}|| = |c|||SAS^{-1}|| = |c|||A||_S$
3) $\forall A, B \in M_n(\mathbb{F}), ||A+B||_S = ||S(A+B)S^{-1}||_S = ||SAS^{-1}+SBS^{-1}|| \leq ||SAS^{-1}||+ ||SBS^{-1}|| = ||A||_S + ||B||_S$
So $||.||_S$ is a norm.

Problem 16 (3.37)
$\forall p \in V$, we can write $p = ax^2 + bx + c$. Need to find $q = dx^2 + ex + f \in V$ s.t. $\int_0^1 pq = 2a + b, \forall p$.
$\Rightarrow \int_0^1 pq = a(\frac{d}{5} + \frac{e}{4} + \frac{f}{3}) + b(\frac{d}{4} + \frac{e}{3} + \frac{f}{2}) + c(a(\frac{d}{3} + \frac{e}{2} + \frac{f}{1}))$
$\Rightarrow (\frac{d}{5} + \frac{e}{4} + \frac{f}{3}) = 2, (\frac{d}{4} + \frac{e}{3} + \frac{f}{2}) = 1$
$\Rightarrow d = 180, e = -168, f = 24, q = 180x^2 - 168x + 24$

Problem 17 (3.38)
Let $p = ax^2 + bx + c$, so $p' = 2ax + b$
Let $D = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. We could check that $D \times (c, b, a) = \begin{pmatrix} b \\ 2a \\ c \end{pmatrix}$ satisfies the condition.
Denote the adjoint by D^*. By definition, $\forall f, g \in V, \int_0^1 gDf = \int_0^1 gf' = \int_0^1 D^*gf$
$\Rightarrow \int_0^1 D^*gf = gf|_0^1 - \int_0^1 fg'$ (integration by parts).
If we require that the function yields the same values for 0 and 1, then:
$$D^* = -D = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

Problem 18 (3.39)
$\forall v, v' \in V, w \in W$
(i)
$< w, (S + T)v >=< w, Sv > + < w, Tv >=< S^*w, v > + < T^*w, v >=< (S^* + T^*)w, v >$

$\Rightarrow (S+T)^* = S^* + T^*$
$< w, (\alpha T)v >= \alpha < w, Tv >= \alpha < T^*w, v >=< \bar{\alpha}T^*w, v >$
$\Rightarrow (\alpha T)^* = \bar{\alpha}T^*$

(ii)
$< w, Sv >=< S^*w, v >=< w, (S^*)^*v >$
$\Rightarrow S = (S^*)^*$

(iii)
$< v', STv >=< S^*v', Tv >=< T^*S^*v', v >$
$\Rightarrow (ST)^* = T^*S^*$

(iv)
From (iii), we know:
$(TT^{-1})^* = I^* = I = (T^{-1})^*T^*$
$\Rightarrow (T^{-1})^* = (T^*)^{-1}$

Problem 19 (3.40)
(i)
$\forall B, B' \in M_n(\mathbb{F}), < B', AB >= tr(B'^H AB) =< A^H B', B >$
$\Rightarrow A^* = A^H$

(ii)
$< A_2, A_3A_1 >= tr(A_2^H A_3A_1), < A_2A_1^*, A_3 >= tr((A_2A_1^*)^H A_3) = tr((A_2A_1^H)^H A_3) = tr(A_1A_2^H A_3)$
$\Rightarrow < A_2, A_3A_1 >=< A_2A_1^*, A_3 >$, according to the hint.

(iii)
$< X', T_AX >=< X', AX - XA >= tr(X'^H(AX - XA)) = tr(X'^H AX) - tr(X'^H XA) = tr(X'^* AX) - tr(X'^* XA) = tr(X'^* AX) - tr(AX'^* X) =< A^*X', X > - < X'A^*, X >=< A^*X' - X'A^*, X >=< T_{A^*}X, X' >$
$\Rightarrow (T_A)^* = T_{A^*}$

Problem 20 (3.44)
Given $A \in M_{m\times n}(\mathbb{F}), b \in \mathbb{F}^m$,
Suppose we can find x s.t. $Ax = b$, then $b \in \mathscr{R}(A)$.
By Fundamental Subspaces Theorem, and 3.40, we know $b \in \mathscr{N}(A^H)^\perp \Rightarrow \forall y \in \mathscr{N}(A^H), < y, b >= 0$
Suppose there is no such x, then $b \notin \mathscr{N}(A^H)^\perp \Rightarrow \exists y \in \mathscr{N}(A^H), < y, b >\neq 0$

Problem 21 (3.45)
$\forall Y \in Skew_n(\mathbb{R}), X \in Sym_n(\mathbb{R}), X' \in (Sym_n(\mathbb{R}))^\perp$
$< X, Y >= tr(Y^TX) = -tr(YX^T) = -tr(XY^T) = -tr(Y^TX) = 0$
$\Rightarrow Y \in (Sym_n(\mathbb{R}))^\perp \Rightarrow Skew_n(\mathbb{R}) \subseteq (Sym_n(\mathbb{R}))^\perp$

Denote the matrix with 1 on the (i,j)th entry, and 0's on all other entries by E_{ij}.
$< X, X' >= 0 \Rightarrow < (E_{ij} + E_{ji}), X' >=< E_{ij}, X' > + < E_{ji}, X' >= 0$
$\Rightarrow (X')_{ij} + (X')_{ji} = 0, \forall i, j \Rightarrow X' = -X'^T$
$\Rightarrow X' \in Skew_n(\mathbb{R}) \Rightarrow (Sym_n(\mathbb{R}))^\perp \subseteq Skew_n(\mathbb{R})$
$(Sym_n(\mathbb{R}))^\perp = Skew_n(\mathbb{R})$

Problem 22 (3.46)
(i)
$x \in \mathscr{N}(A^H A) \Rightarrow (A^H A)x = A^H(Ax) = 0 \Rightarrow Ax \in \mathscr{N} A^H$
Moreover, Ax is a linear combination of the columns of A, so clearly, $Ax \in \mathscr{R} A$.

(ii)
$\forall x \in \mathscr{N}(A), Ax = 0 \Rightarrow A^H Ax = A^H(Ax) = 0 \Rightarrow x \in \mathscr{N}(A^H A) \Rightarrow \mathscr{N}(A) \subseteq \mathscr{N}(A^H A)$
$\forall x \in \mathscr{N}(A^H A), A^H Ax = 0 \Rightarrow x^H A^H Ax = 0 \Rightarrow (Ax)^H(Ax) =< Ax, Ax >= 0 \Rightarrow Ax = 0 \Rightarrow x \in \mathscr{N}(A) \Rightarrow \mathscr{N}(A^H A) \subseteq \mathscr{N}(A)$
$\Rightarrow \mathscr{N}(A^H A) = \mathscr{N}(A)$

(iii)
By rank-nullity Theorem, we know $n = dim(\mathscr{N}(A^H A)) + rank(A^H A) = dim(\mathscr{N}(A) + rank(A))$
Since the nullity spaces are the same, $rank(A^H A) = rank(A)$

(iv)
Since A has linearly independent columns, $rank(A) = n$.
$\Rightarrow rank(A^H A) = rank(A) = n$.
Since $A^H A$ has n columns, we know that all columns must be linearly independent. So the determinant is not zero, and $A^H A$ is nonsingular.

Problem 23 (3.47)
(i)
Since matrix multiplication is associative,
$P^2 = A(A^H A)^{-1} A^H A(A^H A)^{-1} A^H = A[(A^H A)^{-1}(A^H A)](A^H A)^{-1} A^H = AI(A^H A)^{-1} A^H$
$A(A^H A)^{-1} A^H = P$

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

(iii)
$\forall b \in \mathscr{R}(A), \exists x$ s.t. $Ax = b$
$\Rightarrow Pb = A(A^H A)^{-1} A^H Ax = A[(A^H A)^{-1}(A^H A)]x = Ax = b \Rightarrow b \in \mathcal{R}(p) \Rightarrow \mathscr{R}(A) \subseteq \mathscr{R}(P)$
So $rank(P) \geq rank(A) = n$. Since P has n columns, its rank is at most n. We then know that $rank(P) = n$.

Problem 24 (3.48)
(i)
$\forall x, y \in \mathbb{R}, A, B \in M_n\mathbb{R}$
$P(xA + yB) = \frac{(xA+yB)+(xA+yB)^T}{2} = x\frac{A+A^T}{2} + y\frac{B+B^T}{2} = xP(A) + yP(B)$
By definition 2.1.1, P is a linear transformation.

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

(iii)
$< B, P(A) >= tr((\frac{A+A^T}{2})^T B) = tr(\frac{AB}{2}) + tr(\frac{A^T B}{2}) = tr(\frac{AB}{2}) + tr(\frac{AB^T}{2}) = tr(\frac{B+B^T}{2}A)$
$=< P(B), A >$
$\Rightarrow P^* = P$

(iv)
$A \in \mathscr{N}(P) \Leftrightarrow P(A) = 0 \Leftrightarrow \frac{A+A^T}{2} = 0 \Leftrightarrow A = -A^T \Leftrightarrow A \in skew_n(\mathbb{R})$

(v)
$A \in \mathscr{R}(P) \Leftrightarrow \exists B$ s.t. $P(B) = A \Leftrightarrow \frac{B+B^T}{2} = A \Leftrightarrow A^T = (\frac{B+B^T}{2})^T = \frac{B+B^T}{2} = A \Leftrightarrow$
$A \in sym_n(\mathbb{R})$

(vi)
$||A - P(A)||_F = \sqrt{< A - P(A), A - P(A) >} = \sqrt{tr[(A - P(A))^T(A - P(A))]}$
$= \sqrt{tr(\frac{(A-A^T)(A-A^T)}{4})} = \sqrt{\frac{tr(A^2)-2tr(A^T A)+tr((A^T)^2)}{4}} = \sqrt{\frac{2tr(A^2)-2tr(A^T A)}{4}}$
$= \sqrt{\frac{tr(A^2)-tr(A^T A)}{2}}$

Problem 25 (3.50)
$rx^2 + sy^2 = 1 \Leftrightarrow y^2 = \frac{1}{s} - \frac{r}{s}x^2$
So the normal equation could be written as:
$A = \begin{pmatrix} -x_1^2 & 1 \\ -x_2^2 & 1 \\ \vdots & \vdots \\ -x_n^2 & 1 \end{pmatrix}, x = \begin{pmatrix} \frac{r}{s} \\ \frac{1}{s} \end{pmatrix}, b = \begin{pmatrix} y_1^2 \\ y_2^2 \\ \vdots \\ y_n^2 \end{pmatrix}$