Problem Set : Measure Theory

OSM 2018 Dr. Evans
Cache Ellsworth

Problem 3.1:

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

Notice that $\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{y}, \mathbf{x}\rangle$ on the real inner product.

ii) $\frac{1}{2}(\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle + \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle) = \frac{1}{2}(\langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{x}\rangle - \langle \mathbf{y}, \mathbf{x}\rangle - \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle) = \frac{1}{2}(2\langle \mathbf{x}, \mathbf{x}\rangle + 2\langle \mathbf{y}, \mathbf{y}\rangle) = \langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle = ||\mathbf{x}||^2 + ||\mathbf{y}||^2$

Problem 3.2:

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

Problem 3.3:

i) The angle between $\mathbf{x}$ and $\mathbf{x}^5$.

$$cos\theta = \frac{\langle \mathbf{x}, \mathbf{x}^5\rangle}{||\mathbf{x}||\,||\mathbf{x}^5||} \implies \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx}\sqrt{\int_0^1 x^{10} dx}} = \frac{\sqrt{33}}{7} = cos\theta \implies \theta = 0.60825 \text{ radians}$$

ii) The angle between $\mathbf{x}^2$ and $\mathbf{x}^4$.

$$cos\theta = \frac{\langle \mathbf{x}^2, \mathbf{x}^4\rangle}{||\mathbf{x}^2||\,||\mathbf{x}^4||} \implies \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^4 dx}\sqrt{\int_0^1 x^8 dx}} = \frac{\sqrt{45}}{7} = cos\theta \implies \theta = 0.2898 \text{ radians}$$

Problem 3.8:

i)

Proof. Let $\mathbf{x} = \text{span}(S) \subset V$, where $S = \{\cos(t), \sin(t), \cos(2t), \sin(2t)\}$. We notice that the following integrals equal 0:

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

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

We notice that the following integrals equal 1:

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

Therefore by definition, this is a orthonormal set. □

ii) Compute $||t||$.

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

iii)

$$\text{proj}_{\mathbf{x}}(\cos(3t)) = \sum_{i=1}^{m}\langle x_i, \cos(3t)\rangle x_i = \sum_{i=1}^{m}\frac{1}{\pi}\int_{\pi}^{\pi} x_i\cos(3t)dt x_i = 0+0+0+0 = 0$$

iv)

$$\text{proj}_{\mathbf{x}}(t) = \sum_{i=1}^{m}\langle x_i, t\rangle x_i = \sum_{i=1}^{m}\frac{1}{\pi}\int_{\pi}^{\pi} x_i t dt x_i =$$
$$0 + \frac{1}{\pi}\int_{-\pi}^{\pi}\sin(t)\cdot t\ dt\cdot\sin(t) + 0 + \frac{1}{\pi}\int_{-\pi}^{\pi}\sin(2t)\cdot tdt\cdot\sin(2t) = 2\sin(t) - \sin(2t)$$

Problem 3.9

Notice that the considered function is $P_\theta(x,y) = (x\cos(\theta) - y\sin(\theta), x\sin(\theta) + y\cos(\theta))$. This is an orthonormal transformation because:

$$\langle(x_1\cos(\theta) - x_2\sin(\theta), x_1\sin(\theta) + x_2\cos(\theta)), (y_1\cos(\theta) - y_2\sin(\theta), y_1\sin(\theta) + y_2\cos(\theta))\rangle =$$
$$(x_1\cos(\theta) - x_2\sin(\theta)(y_1\cos(\theta) - y_2\sin(\theta))) + (x_1\sin(\theta) + x_2\cos(\theta)(y_1\sin(\theta) + y_2\cos(\theta))) =$$
$$x_1y_1 + x_2y_2 = \langle x, y\rangle$$

Problem 3.10

i)

Proof. Let $Q \in M_n(\mathbb{F})$ be orthonormal. Thus $\langle m,n\rangle = \langle Qm, Qn\rangle$. Then $(Qm)^H Qn = m^H n \implies m^H Q^H Qn = m^H n \implies Q^H Q = I$ From the other side, $Q^H Q = QQ^H = I, then \langle Q(x), Q(y)\rangle = (Qx)^H(Qy) = x^H Q^H Qy = x^H y = \langle x,y\rangle$. □

ii)

$$||x||^2 = \langle x,x\rangle = \langle Qx, Qx\rangle = ||Qx||^2 \implies ||x|| = ||Qx||$$

iii)

Assume Q is orthonormal matrix. This implies that $QQ^H = Q^HQ = I \implies Q^H = Q^{-1} \implies (Q^H)^H = Q \implies (Q^H)(Q^H)^H = (Q^H)^HQ^H = I \implies Q^H = Q^{-1}$ is orthonormal.

iv) Let q_i be the i^{th} column of Q. We know that Q is orthonormal, therefore, $(Q^HQ)_{i,j} = q_i^Hq_j = \langle q_i, q_j\rangle$. If $i = j$ then this is equal to one and if $i \neq j$ then this is equal to zero. Therefore by defintion the columns of Q are orthonormal.

v)

No, the converse isn't true. Here is a counterexample. $\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}^H \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0\frac{1}{4} \end{pmatrix}$ which is not equal to I.

Problem 3.11

This will give a $\mathbf{0}$ vector for one of the steps or one of the $\boldsymbol{q}_k$. Particularly it will be the zero vector on the $\boldsymbol{x_i}$ where it is a linear combination or dependent on $\boldsymbol{x_1}, \ldots, \boldsymbol{x}_{i-1}$. This causes problems because $\boldsymbol{q_{k+1}}$ will have errors in the calculation because $\boldsymbol{p}_k$ will be multiplied by the zero vector. Thus it will not give us a set that is orthonormal in V with the same span. (Unless these zero vectors are discarded when realized they are a linear combination.)

Problem 3.16

i) Let $A \in \mathbb{M}_{mxn}$ where rank$(A) = n \leq m$.Then there exist orthonormal $Q \in \mathbb{M}_{mxm}$ and upper triangular $R \in \mathbb{M}_{mxn}$ 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 $Q = \tilde{Q}$.

**(I thank Albi for typing this problem out.)

Problem 3.17

Proof. Let $A \in M_{mxn}$ have rank $n \leq m$, and let $A = \hat{Q}\hat{R}$ be a reduced QR decomposition. Then $A^HA\boldsymbol{x} = A^H\boldsymbol{b} \implies (\hat{Q}\hat{R})^H(\hat{Q}\hat{R})\boldsymbol{x} = (\hat{Q}\hat{R})^H\boldsymbol{b} \implies \hat{R}^H\hat{Q}^H\hat{Q}\hat{R}\boldsymbol{x} = \hat{R}^H\hat{Q}^H\boldsymbol{b}$. Because we have a QR decomposition, we know that Q is orthonormal.

Thus $\hat{Q}^H\hat{Q} = I$. We also know that R is invertible because it is a upper triangular matrix. Taking the inverse of R on the LHS, we then get $\hat{R}\boldsymbol{x} = \hat{Q}^H\boldsymbol{b}$. $\square$

Problem 3.23

Proof. Notice $||\mathbf{x}-\mathbf{y}||^2 = ||\mathbf{x}||^2 - \langle\mathbf{x},\mathbf{y}\rangle - \langle\mathbf{x},\mathbf{y}\rangle + ||\mathbf{y}||^2 \geq ||\mathbf{x}||^2 - 2\langle\mathbf{x},\mathbf{y}\rangle + ||\mathbf{y}||^2 \geq ||\mathbf{x}||^2 - 2||\mathbf{x}||||\mathbf{y}|| + ||\mathbf{y}||^2 = (||\mathbf{x}|| - ||\mathbf{y}||)^2$. Thus $||\mathbf{x}|| - ||\mathbf{y}|| \leq ||\mathbf{x}-\mathbf{y}||$. Switch x and y in this equation to get the other inequality.

$\square$

Problem 3.24
Let $C([a,b];\mathbb{F})$ be the vector space of all continuous functions from $[a,b] \subset \mathbb{R}$ to $\mathbb{F}$
i).
1. Positivity. $0 \leq |\int_a^b |f(t)+g(t)|dt \leq \int_a^b |f(t)dt$. So it's positive. The equality will hold if and only if $f(t) = 0$. $\rightarrow$) Assume $f(t) = 0$ then $\int_a^b |0|dt = 0$. $\leftarrow$) We assume $0 = \int_a^b |f(t)|dt$. We prove by contradiction. $f(t) \neq 0$. Because $b > a$ then $0 < \int_a^b |f(t)|dt$. This is a contradiction. So $f(t) = 0$.
2. Scale Preservation. Let c be a scaler. $||cf||_{L^1} = \int_a^b |cf(t)|dt = \int_a^b |c||f(t)|dt = |c|\int_a^b |f(t)|dt = |c|||f||_{L^1}$.
3. Triangle inequality. $||f+g||_{L^1} = \int_a^b |f(t)+g(t)|dt \leq \int_a^b (|f(t)|+|g(t)|)dt = \int_a^b |f(t)|dt + \int_a^b |g(t)|dt = ||f||_{L^1} + ||g||_{L^1}$. So $||f+g||_{L^1} \leq ||f||_{L^1} + ||g||_{L^1}$.
So this is a norm on $C([a,b];\mathbb{F})$
ii).

1. Positivity. $||f||_{L^2} = (\int_a^b |f(t)|^2)^{\frac{1}{2}}$. By the thereom $0 \geq \int_a^b |f(t)|^2$ Then $0 \leq (\int_a^b |f(t)|^2)^{\frac{1}{2}}$. So it's positive. The equality will hold iff $f(t) = 0$. This is essentially the same as *i*.
2. Scale Preservation. Let c be a scaler. $||cf||_{L^2} = (\int_a^b |f(t)|^2)^{\frac{1}{2}} = (\int_a^b |c|^2|f(t)|^2)^{\frac{1}{2}} = (|c|^2\int_a^b |f(t)|^2)^{\frac{1}{2}} = |c|||f||_{L^2}$.
3. Triangle inequality. $||f+g||_{L^2}^2 = (\int_a^b |f+g||_{L^2}^2) = \int_a^b |f^2+2fg+g|dt \leq \int_a^b |f(t)|^2dt + 2\int_a^b |g(t)|^2dt \int_a^b |f(t)|^2dt + \int_a^b |g(t)|dt = ||f||_{L^2}^2 + 2||f||_{L^2}||g||_{L^2} + ||g||_{L^2}^2 = (||f(t)|| + ||g(t)||)^2$. So $||f+g||_{L^2} \leq ||f(t)||_{L^2} + ||g(t)||_{L^2}$.
iii)

1. Positivity. This is positive because we take the supremum of a positive function. The equality will hold iff $f(x) = 0.(\rightarrow)$: Assume $\sup_{x\in[a,b]}|f(x)| = 0.$. By contradiction we assume $f(x) \neq 0.$. Case 1: $x < 0$. so $|f(x)| > 0$ and so $\sup_{x\in[a,b]}|f(x)| > 0$. This is a contradiction. Case 2: $f(x) > 0\sup_{x\in[a,b]}|f(x)| > 0$. Contradiction. So $f(x) = 0$.
($\leftarrow$:) We assume $f(x) = 0$. Then $\sup_{x\in[a,b]}|f(x)| = 0$. Therefore the inequality holds

iff $f(x) = 0$.
2. Scalar Preservation. $||\lambda f||_{L^\infty} = \sup_{x\in[a,b]}|\lambda f(x)| = |\lambda|\sup_{x\in[a,b]}|f(x)| = |\lambda|||f||_{L^\infty}$.
3. Triangle inequality. $||f+g||_\infty = \sup_{x\in[a,b]}|f(x)+g(x)| \le \sup_{x\in[a,b]}(|f(x)|+|g(x)|) \le \sup_{x\in[a,b]}|f(x)| + \sup_{x\in[a,b]}|g(x)| = ||f||_\infty + ||g||_\infty$.

Problem 3.26

i)
1. Reflexive: Let $||\cdot||_a \in V$. Let $m = \frac{1}{3}$ and $M = 3$. So $\frac{1}{3}||\cdot||_a \le ||\cdot||_a \le 3||\cdot||_a$. Since $||\cdot||_a \ge 0$ by norm. properties. So $||\cdot||_a \; ||\cdot||_a$
2. Symmetric: Suppose $||\cdot||_a \; ||\cdot||_b$ so $\exists 0 \le m \le M \text{s.t.} m||\cdot||_a \le ||\cdot||_b \le M||\cdot||_a$. So $m||\cdot||_a \le ||\cdot||_b \implies ||\cdot||_a \le \frac{1}{m}||\cdot||_b$. Also $||\cdot||_b \le M||\cdot||_a \implies \frac{1}{M}||\cdot||_b \le ||\cdot||_a$. We let $N = \frac{1}{m}$ and $n = \frac{1}{M}$. Then $n||\cdot||_b \le ||\cdot||_a \le N||\cdot||_b$. Therefore $||\cdot||_b \; ||\cdot||_a$.
3. Transitivity: Assume $||\cdot||_a \; ||\cdot||_b$ and $||\cdot||_b \; ||\cdot||_c$. So $\exists 0 \le m \le M \text{s.t.} m||\cdot||_a \le ||\cdot||_b \le M||\cdot||_a$ and $\exists 0 \le n \le N \text{s.t.} n||\cdot||_b \le ||\cdot||_c \le N||\cdot||_b$. We times n to our first inequality. $nm||\cdot||_a \le n||\cdot||_b \le ||\cdot||_c \le N||\cdot||_b \le NM||\cdot||_a$. So $nm||\cdot||_a \le ||\cdot||_c \le NM||\cdot||_a$ where $nm = g$ and $NM = G$. Therefore $||\cdot||_a \; ||\cdot||_c$. This is an equivalence relation.
We prove that the p-norms for $p = 1, 2, \infty$ on $\mathbb{F}^n$ are topologically equivalent by
a) $||x||_2^2 = |x_1|^2 + \cdots + |x_n|^2 \le |x_1|^2 + \cdots + |x_n|^2 + |x_i||x_j| + \cdots = (|x_1| + \cdots + |x_n|)^2 = ||x||_1^2$. So $||x||_2 \le ||x||_1$. $||x||_1 = \sum_{i=1}^n |x_i| = \sum_{i=1}^n |x_i| * 1 \le (\sum_{i=1}^n |x_i| * 1)^{\frac{1}{2}}(\sum |1|)^{\frac{1}{2}}$ by cauchy schwartz. Then this is equal to $||x||_2 n^{\frac{1}{2}}$. Thus $||x||_2 \le ||x||_1 \le sqrtn||x||_2$.
b) $||x||_\infty^2 = (\sup\{|x_1|, |x_2|, \ldots, |x_n|\})^2 = \sup\{|x_1|^2, \ldots, |x_n|^2\} \le \sum_{i=1}^n |x_i|^2 = ||x||_2^2$. So $||x||_\infty \le ||x||_2$. Also $||x||_2^2 = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2) \le n\sup\{|x_1|^2, |x_2|^2, \ldots, |x_n|^2\} = n||x||_\infty$. So, $||x||_2^2 \le n||x||_\infty^2$. Therefore the ∞ and 2 norm are topologically equivalent by $||x||_\infty \le ||x||_2 \le \sqrt{2}||x||_\infty$.
Therefore, $||\cdot||_2$ and $||\cdot||_\infty$ are topologically equivalent and $||\cdot||_1$ and $||\cdot||_2$ are topologically equivalent. Adding it all together, $||x||_\infty \le ||x||_2 \le ||x||_1 \le \sqrt{n}||x||_2 \le n||x||_\infty$. So —$|x||_\infty \le ||x||_1 \le n||x||_\infty$. So ∞ and 1 are topologically equivalent. So ∞ and 1 and 2 are TE.

Problem 3.28

i) Using the previous exercise we can see that $\sup_{x\neq 0}\frac{||Ax||_1}{||x||_1} \le \sup_{x\neq 0}\frac{||Ax||_1}{||x||_2}$ because the 2 norm is smaller. Then this is less than $\sqrt{n}\sup_{x\neq 0}\frac{||Ax||_2}{||x||_2}$ because Ax is a vector and the 2 norm is bigger with a square rooted n. Also, $\frac{1}{sqrtn}\sup_{x\neq 0}\frac{||Ax||_2}{||x||_2} \le \sup_{x\neq 0}\frac{||Ax||_2}{||x||_1}$ because the denominator is larger using the fact proved in problem 26. This is also less than $\sup_{x\neq 0}\frac{||Ax||_1}{||x||_1}$ because $||\cdot||_2 \le ||\cdot||_1$. Thus the inequality follows by putting these together.
ii) Using the previous exercise and skipping the intermediate step shown in part i) we can see that $\frac{1}{\sqrt{n}}\sup_{x\neq 0}\frac{||Ax||_\infty}{||x||_\infty} \le \sup_{x\neq 0}\frac{||Ax||_2}{||x||_2}$. Also we see that $\sup_{x\neq 0}\frac{||Ax||_2}{||x||_2} \le \sqrt{n}\sup_{x\neq 0}\frac{||Ax||_\infty}{||x||_\infty}$. Putting these together gives us the inequality.

Problem 3.29
Let $\mathbf{x} \neq 0$ and $||\cdot||$ be the standard inner product. Then

$||Qx|| = \sqrt{\langle Qx, Qx\rangle} = \sqrt{x^H Q^H Qx} = \sqrt{(Q^H Qx)^H x} = \sqrt{\langle Q^H Qx, x\rangle} = \sqrt{\langle x, x\rangle} = ||x||$. Thus, $||Q|| = \sup_{x\neq 0} \frac{||Qx||}{||x||} = 1$. Now $||R_x|| = \sup_{A\neq 0} \frac{||Ax||}{||A||} = \sup_{A\neq 0} \frac{||Ax||||x||}{||A||||x||}$ and then by the sup-multiplicative property this is less than $\sup_{A\neq 0} \frac{||Ax||||x||}{||Ax||} = ||x||$.

Problem 3.30
Let $S, A, B \in M_n\mathbb{F}$ and S be an invertible matrix.
1.) Positivity. So $||A||_S = ||SAS^{-1}|| \geq 0 \forall A$. We prove equality, $||0||_S = ||S0S^{-1} = ||0|| = 0$ because $||\cdot||$ is a norm. Also if $0 = ||A||_S = ||SAS^{-1}|| \implies A = SAS^{-1} \implies A = 0$
2.) Scalar Preservation. Let $a \in \mathbb{F}$

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

3.) Triangle Inequality:

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

4.) Sub-multiplicative Property:

$$||AB||_S = ||SABS^{-1}|| = ||SAS^{-1}SBS^{-1}|| \leq ||SAS^{-1}||||SBS^{-1}|| = ||A||_S||B||_S.$$

Problem 3.37
Let $V = \mathbb{R}[x; 2]$ be the space of polynomials of degree at most two. Let $L : V \to \mathbb{R}$ be the linear functional given by $\mathrm{L}[p] = p'(1)$. Let $p(x), q(x) \in V$ where $p(x) = ax^2+bx+c$ and $q(x) = a'x^2+b'x+c'$ with $a, b, c, a', b', c' \in \mathbb{R}$ and a vector such that $p = (a, b, c)$ and $q = (a', b', c')$. Also we let q be the unique vector so that $\langle q, p\rangle = 2a+b = p'(1) = L[p]$. So $a' = 2, b' = 1, c' = 0$ and $q = (2, 1, 0)$.

Problem 3.38
Let $V = \mathbb{F}[x; 2]$ and $D : V \to V$ be the derivative operator. Then the matrix representation of D with respect to the power basis $[1, x, x^2]$ is $\begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Also the adjoint of D is $\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Problem 3.39
Let V and W be finite-dimensional inner product spaces. Let $S < T \in \mathcal{L}(V : W)$.

(i)

Proof. $\langle (S+T)^*w, v\rangle_V = \langle w, (S+T)v\rangle_W = \langle w, Sv+Tv\rangle_W = \langle w, Sv\rangle_W + \langle w, Tv\rangle_W = \langle S^*w, v\rangle_V + \langle T^*w, v\rangle_V = \langle S^*w+T^*w, v\rangle_V$ So $(S+T) = S^*+T^*$. Also, $\langle (\alpha T)^*w, v\rangle_V = \langle w, (\alpha T)v\rangle_W = \langle w, \alpha Tv\rangle_W = \alpha\langle w, Tv\rangle = \alpha\langle T^*w, v\rangle = \langle \bar{\alpha}T^*w, v\rangle$. So $(\alpha T)^* = \bar{\alpha}T$. □

(ii)

Proof. $\langle w, Sv\rangle_W = \langle S^*w, v\rangle_V = \overline{\langle v, S^*w\rangle_V} = \overline{\langle S^{**}v, w\rangle_W} = \langle w, S^{**}v\rangle_W$. Therefore $(S^*)^* = S$ □

(iii)

Proof. $\langle (ST)^*v', v\rangle_V = \langle v', (ST)v\rangle_V = \langle v', S(Tv)\rangle_V = \langle S^*v', Tv\rangle_V = \langle T^*S^*v', v\rangle_V$ Therefore $(ST)^* = T^*S^*$ □

(iv)

Proof. By part (iii) : $(TT^{-1})^* = T^*(T^*)^{-1} = I^* = I$. Thus $(T^*)^{-1} = (T^{-1})^*$. □

Problem 3.40
(i)

Proof. Let $S, T \in M_n(\mathbb{F})$. Then $\langle A^*S, T\rangle_F = \langle S, AT\rangle_F = \mathrm{tr}(S^H AT) = \mathrm{tr}((A^H S)^H T) = \langle A^H S, T\rangle$. Therefore, $A^* = A^H$. □

(ii)

Proof. Let $A_1, A_2, A_3 \in M_n(\mathbb{F})$. Then $\langle A_2, A_3A_1\rangle_F = \mathrm{tr}(A_2^H A_3 A_1) = \mathrm{tr}(A_1 A_2^H A_3) = \mathrm{tr}((A_2 A_1^H)^H A_3) = \langle A_2 A_1^H, A_3\rangle_F = \langle A_2 A_1^*, A_3\rangle_F$ by part (i). Therefore, $\langle A_2, A_3A_1\rangle_F = \langle A_2A_1^*, A_3\rangle_F$. □

(iii)

Proof. Let $A, B, C \in M_n(\mathbb{F})$ and define the linear operator $T_A : M_n(\mathbb{F}) \to M_n(\mathbb{F})$ by $T_A(X) = AX - XA$. Then $\langle T_A^* B, C\rangle_F = \langle B, AC - CA\rangle_F = \langle B, AC\rangle - \langle B, CA\rangle_F$. We apply part (ii) and get $\langle B, AC\rangle - \langle BA^*, C\rangle$ Also, $\langle B, AC\rangle = \mathrm{tr}(B^A C) = \mathrm{tr}((A^H B)^H C) = \langle A^H B, C\rangle = \langle A * B, C\rangle$. So no we have $\langle A^*B, C\rangle - \langle BA^*, C\rangle = \langle AB - BA, C\rangle$. Therefore $T_A^* = T_{A^*}$.. □

Problem 3.44
We suppose that $y \in \mathcal{N}(A^H)$ such that $\langle y, b\rangle \neq 0$. Then $b \notin \mathcal{N}(A^H)^\perp = \mathcal{R}(A)$ Then there is no $x \in \mathbb{F}^n, Ax = b$. Now suppose $\exists x \in \mathbb{F}^n$ s.t. $Ax = b$. Then, $\forall y \in \mathcal{N}(A^H)$, $\langle y, b\rangle = \langle y, Ax\rangle = \langle A^H y, x\rangle = \langle 0, x\rangle = 0$.

Problem 3.45
Let $A \in \mathrm{Sym}_n(\mathbb{R})$ and $B \in \mathrm{Skew}_n(\mathbb{R})$. Then $\langle B, A\rangle = \mathrm{Tr}(B^T A) = \mathrm{Tr}(AB^T) = \mathrm{Tr}(A^T(-B)) = -\langle A, B\rangle$ Therefore $\langle A, B\rangle$ must equal 0 because the inner product is

positive. Also, $\text{Skew}_n(\mathbb{R}) \subset \text{Sym}_n(\mathbb{R})^\perp$. Let $B \in \text{Sym}_n(\mathbb{R})^\perp$ and $B + B^T \in \text{Sym}_n(\mathbb{R})$. So,$0 = \rangle B + B^T, B\langle = \text{Tr}((B + B^T)B) = \text{Tr}(BB + B^TB) = \text{Tr}(BB) + \text{Tr}(B^TB) \implies \langle B^T, B\rangle = \langle -B, B\langle$. So, $B^T = -B$. Therefore, $\text{Sym}_n(\mathbb{R})^\perp = \text{Skew}_n(\mathbb{R})$

Problem 3.46

i)

$$x \in \mathcal{N}(A^HA) \text{ So, } 0 = (A^HA)x = A^H(Ax) \text{ so } Ax \in \mathcal{N}(A^H) \text{ then } Ax \in \mathcal{R}(A) \text{ by definition.}$$

ii) Suppose $x \in \mathcal{N}(A^HA)$. Then $0 = A^HAx \implies x^H0 = x^HA^HAx = ||Ax||$. Then $Ax = 0$ and so $x \in \mathcal{N}(A)$. Now, Suppose $x \in \mathcal{N}(A)$. So $Ax = 0 \implies A^HAx = A^H0 = 0 \implies x \in \mathcal{N}(A^HA)$. Thus $\mathcal{N}(A^HA) = \mathcal{N}(A)$.

iii)

$$n = \text{rank}(A^HA) + \text{dim}\mathcal{N}(A^HA). \text{ Since } \mathcal{N}(A^HA) = \mathcal{N}(A). \text{ So then, } n = \text{rank}(A^HA) + \text{dim}\mathcal{N}(A), \text{ Thus rank}(A^HA) = \text{rank}(A).$$

iv) Assume A has linearly independent columns. So, $\text{rank}(A) = n = \text{rank}(A^HA)$ by part three. Then $A^HA \in \mathbb{M}_n$ so it is nonsingular.

Problem 3.47

i)

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

ii)

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

iii) We know that $\text{rank}(A) = n$

$$textrank(A(A^HA)^{-1}A^H) \leq \min(\text{rank}(A), \text{rank}(A^HA^{-1}), \text{rank}(A^H)). \text{ This is by a matrix property}$$

.

Thus, we know that an invertible matrix is full rank and so all of these ranks are n. So the minimum is n. So the rank is n.

Problem 3.48

i) Let $A, B \in \mathbb{M}_n(\mathbb{R})$ and $s, t \in \mathbb{R}$. Then $P(sA + tB) = \frac{sA+tB+(sA+tB)^T}{2} = s\frac{A+A^2}{2} + t\frac{B+B^T}{2} = sP(A) + tP(B)$.

ii)

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

iii) Adjoint is defined as $< P^*(A), B >=< A, P(B) >$. Thus,$< A, P(B) >=< A, \frac{(B+B^T)}{2} >=< A, \frac{B}{2} > + < A, \frac{B^T}{2} >= \text{Tr}(\frac{A^TB}{2}) + \text{Tr}(\frac{A^TB^T}{2}) = \text{Tr}(\frac{A^T}{2B}) + \text{Tr}(\frac{BA}{2}) = \text{Tr}(\frac{A^T}{2B}) + \text{Tr}(\frac{A}{2B}) =< \frac{(A+A^T)}{2}, B >=< P(A), B >$. So $P = P^*$.
iv) Let $A \in \mathcal{N}(P)$. Then $0 = P(A) = \frac{A+A^T}{2} \implies A^T = -A$, so $\mathcal{N}(P) \subset \text{Skew}(\mathbb{R})$. Now, let $A \in \text{Skew}(\mathbb{R})$. Then $A^T = -A$ and so $P(A) = \frac{A+A^T}{2} = 0$. Therefore, $\text{Skew}(\mathbb{R}) \subset \mathcal{N}(P)$
v) Let $A \in \mathbb{M}_n(\mathbb{R})$. Then $P(A) = \frac{A+A^T}{2} = \frac{A^T+A}{2} = P(A)^T$ and so $\mathcal{R}(P) = \text{Sym}(\mathbb{R})$. Now let $A = \text{Sym}(\mathbb{R})$. Then $A = A^T$ and $P(A) = \frac{A+A^T}{2} = \frac{A+A}{2} = A$ and so $A \in \mathcal{R}(P)$. Therefore by proof, $\mathcal{R}(P) = \text{Sym}(\mathbb{R})$.
vi) Notice that

$$||A - P(A)||_F^2 =< A - P(A), A - P(A) >=< A - \frac{A + A^T}{2}, A - \frac{A + A^T}{2} >=$$
$$< \frac{A - A^T}{2}, \frac{A - A^T}{2} >= \text{Tr}\left(\left(\frac{A - A^T}{2}\right)^T \frac{A - A^T}{2}\right) =$$
$$\text{Tr}\left(\frac{A^T - A}{2}\frac{A - A^T}{2}\right) = \text{Tr}\left(\frac{A^TA - A^2 - (A^T)^2 + AA^T}{4}\right) =$$
$$\text{Tr}\left(\frac{A^TA - A^2 - A^2 + A^TA}{4}\right) = \text{Tr}\left(\frac{A^TA - A^2}{2}\right) = \frac{\text{Tr}(A^TA) - \text{Tr}(A^2)}{2}.$$

Therefore $||A - P(A)||_F = \sqrt{\frac{\text{Tr}(A^TA) - \text{Tr}(A^2)}{2}}$.
**(I thank Albi for typing most of this problem out.)

Problem 3.50
Let $A = \begin{bmatrix} x_1^2 & 1 \\ x_2^2 & 1 \\ \vdots & \vdots \\ x_n^2 & 1 \end{bmatrix}$. $\mathbf{x} = \begin{bmatrix} \frac{1}{s} \\ \frac{-r}{s} \end{bmatrix}$ $\mathbf{b} = \begin{bmatrix} y_1^2 \\ y_2^2 \\ \vdots \\ y_n^2 \end{bmatrix}$
So $Ax = b$ model will solve the least squares approximation for r and s. Then the normal equation in terms are $A^HAx = A^Hb$ where $A^HAx = \begin{bmatrix} n/s - \frac{r}{s}\sum_i^n x_i^2 \\ \frac{1}{s}\sum_i^n x_i^2 - \sum_i^n x_i^4 \end{bmatrix} = A^Hb = \begin{bmatrix} \sum_i^n y_i^2 \\ \sum_i^2 x_i^2 y_i^2 \end{bmatrix}$