Problem 1

Part (i)

Proof.

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

□

Part (ii)

Proof.

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

□

Problem 2

Proof. Note that:

$$\begin{aligned}
\|x+y\|^2 - \|x-y\|^2 &= \langle x+y, x+y\rangle - \langle x-y, x-y\rangle \\
&= \overline{\langle x, x+y\rangle} + \overline{\langle y, x+y\rangle} - \overline{\langle x, x-y\rangle} + \overline{\langle y, x-y\rangle} \\
&= \overline{\langle x,x\rangle} + \overline{\langle x,y\rangle} + \overline{\langle y,x\rangle} + \overline{\langle y,y\rangle} - \overline{\langle x,x\rangle} + \overline{\langle x,y\rangle} + \overline{\langle y,x\rangle} - \overline{\langle y,y\rangle} \\
&= 2\langle y,x\rangle + 2\langle x,y\rangle
\end{aligned}$$

Also, note that:

$$\begin{aligned}
i\|x-iy\|^2 - i\|x+iy\|^2 &= i\langle x-iy, x-iy\rangle - i\langle x+iy, x+iy\rangle \\
&= \langle x-iy, y\rangle + \langle x-iy, ix\rangle + \langle x+iy, y\rangle - \langle x+iy, ix\rangle \\
&= \overline{\langle y, x\rangle} - \overline{\langle y, iy\rangle} + \overline{\langle ix, x\rangle} - \overline{\langle ix, iy\rangle} + \overline{\langle y, x\rangle} + \overline{\langle y, iy\rangle} - \overline{\langle ix, x\rangle} - \overline{\langle ix, iy\rangle} \\
&= 2\langle x, y\rangle - 2\langle iy, ix\rangle \\
&= 2\langle x, y\rangle - 2\langle y, x\rangle
\end{aligned}$$

So, we have that:

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

$\square$

Problem 3

Part (i)

$$\begin{aligned}
\theta &= \cos^{-1}\left(\frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx} \cdot \sqrt{\int_0^1 x^{10} dx}}\right) \\
&= \cos^{-1}\left(\frac{\sqrt{33}}{7}\right) \\
&\approx 0.608
\end{aligned}$$

Part (ii)

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

$$\begin{aligned}
&= \cos^{-1}\left(\frac{\sqrt{45}}{7}\right) \\
&\approx 0.29
\end{aligned}$$

Problem 8

Part (i)

Proof. Consider $\langle v_i, v_j \rangle$ for $v_i, v_j \in S$, $i \neq j$. Clearly, if v_i is some form of sine and v_j is some form of cosine (or vice versa), then their product is odd and therefore the integral of their product over the interval $[-\pi, \pi]$ is 0, so $\langle v_i, v_j \rangle = 0$. If $v_i = cos(t)$ and $v_j = cos(2t)$ (or vice versa), we have that:

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

If $v_i = sin(t)$ and $v_j = sin(2t)$ (or vice versa), we have that:

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

Consider $\langle v_i, v_j \rangle$ for $v_i, v_j \in S$, $i = j$. Note that:

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

Thus, by our definition of the inner product over V, $\langle v_i, v_j \rangle = 1$. Thus, S is an orthonormal set. $\square$

Part (ii)

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

Part (iii)

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

Part (iv)

$$\begin{aligned} proj_X(t) &= \langle t, cos(t) \rangle \cdot cos(t) + \langle t, sin(t) \rangle \cdot sin(t) + \langle t, cos(2t) \rangle \cdot cos(2t) + \langle t, sin(2t) \rangle \cdot sin(2t) \\ &= 0 + 2sin(t) + 0 - sin(2t) \\ &= 2sin(t) - sin(2t) \end{aligned}$$

Problem 9

Proof. We define a rotation $r : \mathbb{R}^2 \to \mathbb{R}^2$ as:

$$r\left(\theta, \begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \cdot cos(\theta) - y \cdot sin(\theta) \\ x \cdot sin(\theta) + y \cdot cos(\theta) \end{bmatrix}$$

Consider two vectors $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$ in $\mathbb{R}^2$. Then we have that:

$$\left\langle \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix} \right\rangle = ac + bd$$

Rotating, we get:

$$\begin{aligned}
&\left\langle \begin{bmatrix} a \cdot cos(\theta) - b \cdot sin(\theta) \\ a \cdot sin(\theta) + b \cdot cos(\theta) \end{bmatrix}, \begin{bmatrix} c \cdot cos(\theta) - d \cdot sin(\theta) \\ c \cdot sin(\theta) + d \cdot cos(\theta) \end{bmatrix} \right\rangle = \\
&= (a \cdot cos(\theta) - b \cdot sin(\theta)) \cdot (c \cdot cos(\theta) - d \cdot sin(\theta)) + (a \cdot sin(\theta) + b \cdot cos(\theta)) \cdot (c \cdot sin(\theta) + d \cdot cos(\theta)) \\
&= ac \cdot cos^2(\theta) - ad \cdot sin(\theta)cos(\theta) - bc \cdot sin(\theta)cos(\theta) + bd \cdot sin^2(\theta) + \\
&\quad + ac \cdot sin^2(\theta) + ad \cdot sin(\theta)cos(\theta) + bc \cdot sin(\theta)cos(\theta) + bd \cdot cos^2(\theta) \\
&= ac \cdot (cos^2(\theta) + sin^2(\theta)) + bd \cdot (cos^2(\theta) + sin^2(\theta)) \\
&= ac + bd
\end{aligned}$$

So, a rotation in $\mathbb{R}^2$ is an orthonormal transformation (with respect to the usual inner product). □

Problem 10

Part (i)

Proof. Suppose $Q \in M_n(\mathbb{F})$ is an orthonormal matrix. Then we must have $\langle Qu, Qv \rangle = \langle u, v \rangle \ \forall u, v \in \mathbb{F}^n$. By the definition of the standard inner product:

$$\begin{aligned}
(Qu)^H Qv &= u^H v \\
u^H Q^H Qv &= u^H v \\
Q^H Q &= I
\end{aligned}$$

Thus, by the uniqueness of inverses, $QQ^H = Q^HQ = I$ if $Q \in M_n(\mathbb{F})$ is an orthonormal matrix. Suppose $QQ^H = Q^HQ = I$. Then for any $u, v \in \mathbb{F}^n$,

$$\begin{aligned}
Q^HQ &= I \\
Q^HQv &= v \\
u^HQ^HQv &= u^Hv \\
(Qu)^HQv &= u^Hv \\
\langle Qu, Qv\rangle &= \langle u, v\rangle
\end{aligned}$$

So, $\langle Qu, Qv\rangle = \langle u, v\rangle\ \forall u, v \in \mathbb{F}^n$, i.e. $Q \in M_n(\mathbb{F})$ is an orthonormal matrix. □

Part (ii)

Proof. Consider $x \in \mathbb{F}^n$. Then, we have that:

$$\begin{aligned}
\|Qx\| &= \sqrt{\langle Qx, Qx\rangle} \\
&= \sqrt{\langle x, x\rangle} && \text{Since } Q \in M_n(\mathbb{F}) \text{ is an orthonormal matrix} \\
&= \|x\|
\end{aligned}$$

□

Part (iii)

Proof. From part (i), we know that if $Q \in M_n(\mathbb{F})$ is an orthonormal matrix, then $Q^{-1} = Q^H$. Also, we know that if $QQ^H = Q^HQ = I$, then Q is an orthonormal matrix. Also, note that $Q^H(Q^H)^H = Q^HQ$ and $(Q^H)^HQ^H = QQ^H$. So, we have that:

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

So $Q^{-1} = Q^H$ is an orthonormal matrix if $Q \in M_n(\mathbb{F})$ is an orthonormal matrix. □

Part (iv)

Proof. Suppose an orthonormal matrix $Q \in M_n(\mathbb{F})$ has column vectors $\{v_1, v_2, \ldots, v_n\}$. Then we have that:

$$Q = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} \qquad Q^H = \begin{bmatrix} v_1^H \\ v_2^H \\ \vdots \\ v_n^H \end{bmatrix}$$

Since Q is orthonormal, we know that:

$$
\begin{aligned}
Q^H Q &= \begin{bmatrix} v_1^H \\ v_2^H \\ \vdots \\ v_n^H \end{bmatrix} \cdot \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \\
&= \begin{bmatrix} v_1^H v_1 & v_1^H v_2 & \dots & v_1^H v_n \\ v_2^H v_1 & v_2^H v_2 & \dots & v_2^H v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n^H v_1 & v_n^H v_2 & \dots & v_n^H v_n \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \\
&= I
\end{aligned}
$$

So, we have that the column vectors of an orthonormal matrix form an orthonormal set, since $\langle v_i, v_j \rangle = v_i^H v_j = 0$ if $i \neq j$, and $\langle v_i, v_j \rangle = v_i^H v_j = 1$ if $i = j$. $\square$

Part (v)

Proof. For $Q \in M_n(\mathbb{F})$ an orthonormal matrix, we have that:

$$1 = det(I) = det(Q^H Q) = det(Q^H)det(Q) = (det(Q))^2$$

So, $\mid det(Q) \mid = 1$. The converse is not true; consider $P = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ $\square$

Part (vi)

Proof. Suppose $Q_1, Q_2 \in M_n(\mathbb{F})$ are orthonormal matrices. Consider their product, $Q_1 Q_2$. We have that:

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

Thus, by part (i), $Q_1 Q_2$ is an orthonormal matrix. $\square$

Problem 11

If a collection $\{v_1, v_2, \dots, v_n\}$ of vectors is linearly dependent, then $\exists i \in \mathbb{N}$ s.t. $1 \leq i \leq n$ and $v_i \in span(\{v_1, v_2, \dots, v_{i-1}\})$. WLOG, say that $\{v_1, v_2, \dots, v_{i-1}\}$ are linearly independent. Note that we define:

$$q_i = \frac{x_i - proj_{span(\{q_1, \dots, q_{i-1}\})}(x_i)}{\|x_i - proj_{span(\{q_1, \dots, q_{i-1}\})}(x_i)\|}$$

Note that since $\{v_1, v_2, \ldots, v_{i-1}\}$ are linearly independent, then $span(\{q_1, \ldots, q_{i-1}\}) = span(\{v_1, v_2, \ldots, v_{i-1}\})$ by construction. So, we get that:

$$\begin{aligned} proj_{span(\{q_1,\ldots,q_{i-1}\})}(x_i) &= proj_{span(\{v_1,v_2,\ldots,v_{i-1}\})}(x_i) \\ &= x_i & \text{since } x_i \in span(\{v_1, v_2, \ldots, v_{i-1}\}) \end{aligned}$$

So, the denominator in the expression for q_i is 0, so q_i is undefined, and thus every q_j with $j \geq i$ is undefined.

Problem 16

Part (i)

Proof. Suppose we have factored a matrix $A \in M_{m\times n}$ into a product $A = QR$, where Q is an $m \times m$ orthonormal matrix and R is an $m \times n$ upper-triangular matrix. Consider the $m \times m$ matrix:

$$D := \begin{bmatrix} -1 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & -1 \end{bmatrix}$$

Then the matrix $QD \neq Q$ is still an orthonormal matrix, and the matrix $DR \neq R$ is still an upper-triangular matrix, and:

$$(QD)(DR) = QR = A$$

Thus, the QR decomposition is not unique. □

Part (ii)

Proof. Suppose $A = Q_1R_1 = Q_2R_2$ with Q_1, Q_2 orthonormal matrices and R_1, R_2 upper-triangular matrices. Then we have that:

$$\begin{aligned} R_1^H R_1 &= R_1^H (Q_1^H Q_1) R_1 \\ &= A^H A \\ &= R_2^H (Q_2^H Q_2) R_2 \\ &= R_2^H R_2 \end{aligned}$$

So, we have that:

$$(R_2^H)^{-1} R_1^H = R_2 (R_1)^{-1}$$

Note that the left hand side of the above equation must be a lower-triangular matrix, since the inverse of a lower-triangular matrix is lower-triangular and the product of two lower-triangular matrices is lower-triangular. Also, the right hand side of the equation must be an upper-triangular matrix, since the inverse of an upper-triangular matrix is upper-triangular and the product of two upper-triangular matrices is upper-triangular. Thus, we must have that both sides of the equation are in fact diagonal matrices.

Note that the notion of "positive" is not well-defined on $\mathbb{C}\backslash\mathbb{R}$, so we must assert that the diagonal elements of R_1 and R_2 have no imaginary component in order for them to be positive. Say that the diagonal elements of R_1 are $\alpha_i > 0$ for $1 \leq i \leq n$ and the diagonal elements of R_2 are $\beta_i > 0$ for $1 \leq i \leq n$. Note then that the diagonal elements of $(R_1)^{-1}$ are $1/\alpha_i > 0$ for $1 \leq i \leq n$, and that the diagonal elements of $(R_2^H)^{-1}$ are $1/\beta_i > 0$ for $1 \leq i \leq n$. So, since the diagonal of the product of two triangular matrices of the same kind (i.e. both upper or both lower) is the element-wise product of their diagonals, we have that:

$$\frac{\alpha_i}{\beta_i} = \frac{\beta_i}{\alpha_i} \text{ for } 1 \leq i \leq n$$

Since $\alpha_i > 0$ and $\beta_i > 0$ for $1 \leq i \leq n$, we have that $\alpha_i = \beta_i$ for $1 \leq i \leq n$. Consequently,

$$(R_2^H)^{-1} R_1^H = R_2 (R_1)^{-1} = I$$

So, by the uniqueness of inverses, we must have that $R_1 = R_2$. Since $Q_1 R_1 = Q_2 R_2$, it follows that $Q_1 = Q_2$. So, There is a unique QR decomposition of A s.t. R has only positive diagonal elements. □

Problem 17

Proof.

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

□

Problem 23

Proof. Note that:

$$\begin{aligned}\|x\| &= \|x-y+y\| \\ &\leq \|x-y\|+\|y\|\end{aligned}$$

So, we have that $\|x\|-\|y\| \leq \|x-y\|$ Exchanging x and y, we get that $\|y\|-\|x\| \leq \|x-y\|$ since $\|y-x\| = \|x-y\|$. Thus, since $|\|x\|-\|y\|| \in \{\|x\|-\|y\|, \|y\|-\|x\|\}$, we have that:

$$|\|x\|-\|y\|| \leq \|x-y\| \ \forall\ x,y \in V$$

□

Problem 24

Part (i)

Proof. Consider the norm:

$$\|f\|_{L^1} = \int_a^b |f(t)|dt$$

Then, we have that $\|f\|_{L^1} \geq 0$ and $\|f\|_{L^1} = 0$ iff f is the zero function. This is a direct result of the fact that $|\cdot|$ has this property on $\mathbb{F}$. We also have that, for some constant $c \in \mathbb{F}$:

$$\begin{aligned}\|cf\|_{L^1} &= \int_a^b |cf(t)|dt \\ &= \int_a^b |c|\cdot|f(t)|dt \\ &= |c|\int_a^b |f(t)|dt \\ &= |c|\cdot\|f\|_{L^1}\end{aligned}$$

Finally, we have that:

$$\begin{aligned}\|f\|_{L^1}+\|g\|_{L^1} &= \int_a^b |f(t)|dt + \int_a^b |g(t)|dt \\ &= \int_a^b |f(t)|+|g(t)|dt \\ &\geq \int_a^b |f(t)+g(t)|dt \\ &= \|f+g\|_{L^1}\end{aligned}$$

$\square$

Proof. Consider the norm:

$$\|f\|_{L^2} = \left(\int_a^b |f(t)|^2 dt \right)^{1/2}$$

Then, we have that $\|f\|_{L^1} \geq 0$ and $\|f\|_{L^1} = 0$ iff f is the zero function. This is a direct result of the fact that $|\cdot|$ has this property on $\mathbb{F}$. We also have that, for some constant $c \in \mathbb{F}$:

$$\begin{aligned} \|cf\|_{L^2} &= \left(\int_a^b |cf(t)|^2 dt \right)^{1/2} \\ &= \left(\int_a^b |c|^2 |f(t)|^2 dt \right)^{1/2} \\ &= |c| \cdot \left(\int_a^b |f(t)|^2 dt \right)^{1/2} \\ &= |c| \cdot \|f\|_{L^2} \end{aligned}$$

Finally, we have that:

$$\begin{aligned} (\|f+g\|_{L^2})^2 &= \int_a^b |f(t)+g(t)|^2 dt \\ &\leq \int_a^b (|f(t)| + |g(t)|) \cdot |f(t)+g(t)| dt \\ &= \int_a^b |f(t)| \cdot |f(t)+g(t)| dt + \int_a^b |g(t)| \cdot |f(t)+g(t)| dt \\ &\leq \left(\left(\int_a^b |f(t)|^2 dt \right)^{1/2} + \left(\int_a^b |g(t)|^2 dt \right)^{1/2} \right) \cdot \left(\int_a^b |f(t)+g(t)|^2 dt \right)^{1/2} \text{ by Hölder's inequality} \\ &= (\|f\|_{L^2} + \|g\|_{L^2}) \cdot \|f+g\|_{L^2} \end{aligned}$$

Thus, $\|f+g\|_{L^2} \leq \|f\|_{L^2} + \|g\|_{L^2}$ $\square$

Proof. Consider the norm:

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

Then, we have that $\|f\|_{L^\infty} \geq 0$ and $\|f\|_{L^\infty} = 0$ iff f is the zero function. This is a direct result of the fact that $|\cdot|$ has this property on $\mathbb{F}$. We also have that, for some constant $c \in \mathbb{F}$:

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

Finally, we have that:

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

□

Problem 26

Part (i)

Proof. $\|\cdot\|_a \sim \|\cdot\|_a$ by choosing $m = M = 1$

Suppose $\|\cdot\|_a \sim \|\cdot\|_b$. Then $\exists$ constants $0 < m \leq M$ such that:

$$m\|x\|_a \leq \|x_b\| \leq M\|x\|_a, \quad \forall x \in X$$

So, $\exists$ constants $0 < \frac{1}{M} \leq \frac{1}{m}$ such that:

$$\frac{1}{M}\|x\|_b \leq \|x_a\| \leq \frac{1}{m}\|x\|_b, \quad \forall x \in X$$

i.e. $\|\cdot\|_b \sim \|\cdot\|_a$

Suppose $\|\cdot\|_a \sim \|\cdot\|_b$ and $\|\cdot\|_b \sim \|\cdot\|_c$. Then $\exists$ constants $0 < m \leq M$ and $0 < n \leq N$ such that:

$$m\|x\|_a \leq \|x_b\| \leq M\|x\|_a, \quad \forall x \in X$$

$$n\|x\|_b \leq \|x_c\| \leq N\|x\|_b, \quad \forall x \in X$$

So, $\exists$ constants $0 < nm \leq NM$ such that:

$$nm\|x\|_a \leq \|x_c\| \leq NM\|x\|_a, \quad \forall x \in X$$

i.e. $\|\cdot\|_a \sim \|\cdot\|_c$
Thus, $\sim$ is an equivalence relation.

Define the following p-norms for a vector $x \in \mathbb{F}^n$ with components $\{x_1, x_2, \ldots, x_n\}$:

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|$$

$$\|x\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2}$$

$$\|x\|_\infty = sup\{|x_1|, |x_2|, \ldots, |x_n|\}$$

Then, by the triangle inequality on the 2-norm, we have that:

$$\|x\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2} \leq \sum_{j=1}^{n} (|x_j|^2)^{1/2} = \sum_{j=1}^{n} |x_j| = \|x\|_1$$

Also, by Hölder's inequality, we have that:

$$\|x\|_1 = \sum_{j=1}^{n} |x_j| \leq \left(\sum_{j=1}^{n} |1|^2 \right)^{1/2} \cdot \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2} = \sqrt{n}\|x\|_2$$

It follows that $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$

Part (ii)

Since n finite by assumption, we have that:

$$\begin{aligned} \|x\|_\infty &= sup\{|x_1|, |x_2|, \ldots, |x_n|\} \\ &= max\{|x_1|, |x_2|, \ldots, |x_n|\} \\ &:= |x_i| \quad \text{for some } i \in [1, n] \end{aligned}$$

It follows that:

$$\begin{aligned} \|x\|_\infty &= |x_i| \\ &= (|x_i|^2)^{1/2} \\ &\leq \left(\sum_{j=1}^{i-1} |x_j|^2 + |x_i|^2 + \sum_{k=i+1}^{n} |x_k|^2 \right)^{1/2} \\ &= \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2} \\ &= \|x\|_2 \end{aligned}$$

Finally, we have that:

$$\|x\|_2 = \left(\sum_{j=1}^{n}|x_j|^2\right)^{1/2} \leq \left(\sum_{j=1}^{n}|x_i|^2\right)^{1/2} = (n|x_i|^2)^{1/2} = \sqrt{n}|x_i| = \sqrt{n}\|x\|_\infty$$

It follows that $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$

So, by part (i), we have that $\|\cdot\|_1 \sim \|\cdot\|_2 \sim \|\cdot\|_\infty$ □

Problem 28

Part (i)

Proof. By problem 26, we have that:

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

Also, we have that:

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

It follows that $\frac{1}{\sqrt{n}}\|A\|_2 \leq \|A\|_1 \leq \sqrt{n}\|A\|_2$ □

Part (ii)

Proof. Replacing $\|A\|_1$ with $\|A\|_2$ and $\|A\|_2$ with $\|A\|_\infty$, we have that $\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{n}\|A\|_\infty$ by part (i) and the results of problem 26. Thus, we have that the operator p-norms are topologically equivalent. □

Problem 29

Proof. Note that by defining $\|\cdot\|_2$ on $\mathbb{F}^n$, we have implicitly defined an inner product $\langle\cdot\rangle$ on $\mathbb{F}^n$ s.t. for vectors $x, y \in \mathbb{F}^n$ with components $\{x_1, x_2, \ldots, x_n\}$ and $\{y_1, y_2, \ldots, y_n\}$, we have that:

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

Thus, for $x \in \mathbb{F}^n$, we have that $\|x\|_2 = \sqrt{\langle x, x\rangle}$

$$\|Q\| = \sup_{x\neq 0}\frac{\|Qx\|_2}{\|x\|_2} = \sup_{x\neq 0}\frac{\sqrt{\langle Qx, Qx\rangle}}{\sqrt{\langle x, x\rangle}} = \sup_{x\neq 0}\frac{\sqrt{\langle x, x\rangle}}{\sqrt{\langle x, x\rangle}} = 1$$

Note that:

$$\|R_x\| = \sup_{A\neq 0} \frac{\|Ax\|_2}{\|A\|} = \sup_{A\neq 0} \frac{\|Ax\|_2}{\left(\sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}\right)} \leq \sup_{A\neq 0} \frac{\|Ax\|_2}{\frac{\|Ax\|_2}{\|x\|_2}} = \sup_{A\neq 0} \|x\|_2 = \|x\|_2$$

Don't know how to prove equality. □

Problem 30

Proof. Clearly, $\|A\|_S \geq 0$ $\forall A \in M_n(\mathbb{F})$ since $SAS^{-1} \in M_n(\mathbb{F})$ and $\|A\| \geq 0$ $\forall A \in M_n(\mathbb{F})$. Suppose $\|A\|_S = 0$. Then we have that:

$$\begin{aligned}
\|SAS^{-1}\| &= 0 \\
SAS^{-1} &= 0 \\
SAS^{-1}S &= 0S \\
SA &= 0 \\
S^{-1}SA &= S^{-1}0 \\
A &= 0
\end{aligned}$$

Thus, if $\|A\|_S = 0$, then $A = 0$.

Consider $\|cA\|_S$ for some constant $c \in \mathbb{F}$. Then,

$$\begin{aligned}
\|cA\|_S &= \|S(cA)S^{-1}\| \\
&= \|cSAS^{-1}\| \\
&= |c| \cdot \|SAS^{-1}\| \\
&= |c| \cdot \|A\|_S
\end{aligned}$$

Finally, consider $\|A+B\|_S$ for $A, B \in M_n(\mathbb{F})$:

$$\begin{aligned}
\|A+B\|_S &= \|S(A+B)S^{-1}\| \\
&= \|(SA+SB)S^{-1}\| \\
&= \|SAS^{-1} + SBS^{-1}\| \\
&\leq \|SAS^{-1}\| + \|SBS^{-1}\| \\
&= \|A\|_S + \|B\|_S
\end{aligned}$$

Thus, $\|\cdot\|_S$ is a matrix norm on $M_n(\mathbb{F})$ □

Problem 37

Proof. We want to find $d, e, f \in \mathbb{R}$ s.t. $\forall a, b, c \in \mathbb{R}$:

$$\int_0^1 (dx^2 + ex + f)(ax^2 + bx + c)dx = 2a + b$$

By evaluating the above integral and solving a system of 3 equations, 3 variables, we obtain:

$$d = 180, e = -168, f = 24$$

Thus, we have the required $q \in V$, namely $q = 180x^2 - 168x + 24$ □

Problem 38

Proof. The matrix representation of D with respect to the power basis $[1, x, x^2]$ of $\mathbb{F}[x; 2]$ is:

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

The matrix representation of the adjoint of D is:

$$\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

I can't figure out the boundary conditions (how to get rid of $f(1)g(1) - f(0)g(0)$).

□

Problem 39

Part (i)

Proof.

$$\begin{aligned} \langle (S+T)^*(w), v \rangle_V &= \langle w, (S+T)(v) \rangle_W \\ &= \langle w, S(v) \rangle_W + \langle w, T(v) \rangle_W \\ &= \langle S^*(w), v \rangle_V + \langle T^*(w), v \rangle_V \\ &= \overline{\langle v, S^*(w) \rangle}_V + \overline{\langle v, T^*(w) \rangle}_V \\ &= \overline{\langle v, S^*(w) + T^*(w) \rangle}_V \\ &= \langle S^*(w) + T^*(w), v \rangle_V \end{aligned}$$

By uniqueness of the adjoint, then, we must have $(S+T)^* = S^* + T^*$.

$$\begin{aligned}\langle(\alpha T)^*(w), v\rangle_V &= \langle w, (\alpha T)(v)\rangle_W \\ &= \alpha\langle w, T(v)\rangle_W \\ &= \alpha\langle T^*(w), v\rangle_V \\ &= \langle\overline{\alpha}T^*(w), v\rangle_V\end{aligned}$$

By uniqueness of the adjoint, then, we must have $(\alpha T)^* = \overline{\alpha}T^*$ ☐

Part (ii)

Proof.

$$\langle w, (S^*)^*(v)\rangle_W = \langle S^*(w), v\rangle_V = \langle w, S(v)\rangle_W$$

☐

Part (iii)

Proof.

$$\begin{aligned}\langle(ST)^*(v_2), v_1\rangle &= \langle v_2, ST(v_1)\rangle \\ &= \langle S^*(v_2), T(v_1)\rangle \\ &= \langle T^*S^*(v_2), v_1\rangle\end{aligned}$$

By uniqueness of the adjoint, then, we must have $(ST)^* = T^*S^*$ ☐

Part (iv)

Proof.

$$\begin{aligned}\langle(T^{-1})^*(v_2), v_1\rangle &= \langle(T^{-1})^*(v_2), T(T^{-1}(v_1))\rangle \\ &= \langle T^*(T^{-1})^*(v_2), T^{-1}(v_1)\rangle \\ &= \langle(T^{-1})^*T^*(T^{-1})^*(v_2), v_1\rangle\end{aligned}$$

By uniqueness of the adjoint, then, we must have $(T^{-1})^*T^* = I$, i.e. $(T^*)^{-1} = (T^{-1})^*$ ☐

Problem 40

Part (i)

Proof. Consider three matrices $A, B, C \in M_n(\mathbb{F})$. Then, we have that:

$$\begin{aligned}
\langle C, AB \rangle &= \langle A^* C, B \rangle \\
tr(C^H AB) &= tr((A^* C)^H B) \\
&= tr(C^H (A^*)^H B)
\end{aligned}$$

Note that setting $A^* = A^H$ solves the above equation since $(A^H)^H = A$. Thus, by uniqueness of the adjoint we must have $A^* = A^H$ $\square$

Part (ii)

Proof. Consider three matrices $A_1, A_2, A_3 \in M_n(\mathbb{F})$. Then, we have that:

$$\begin{aligned}
\langle A_2, A_3 A_1 \rangle &= tr((A_2)^H A_3 A_1) \\
&= tr(A_1 (A_2)^H A_3) \\
&= tr((A_2 (A_1)^H)^H A_3) \\
&= tr((A_2 A_1^*)^H A_3) \\
&= \langle A_2 A_1^*, A_3 \rangle
\end{aligned}$$

Note that setting $A^* = A^H$ solves the above equation since $(A^H)^H = A$. Thus, by uniqueness of the adjoint we must have $A^* = A^H$ $\square$

Part (iii)

Proof. Consider three matrices $A, X, Y \in M_n(\mathbb{F})$. Then, we have that:

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

Thus, by uniqueness of the adjoint we must have $(T_A)^* = T_{A^*}$ $\square$

Problem 44

Proof. By the fundamental subspaces theorem,

$$\mathcal{R}(A)^{\perp} = \mathcal{N}(A^*) = \mathcal{N}(A^H)$$

Note that if $Ax = b$ has no solution $x \in \mathbb{F}^n$, then $b \notin \mathcal{R}(A)$. In this case, we must show that $\exists\ y \in \mathcal{N}(A^H) = \mathcal{R}(A)^{\perp}$ s.t. $\langle y, b\rangle \neq 0$.

Suppose not. Then $\forall y \in \mathcal{R}(A)^{\perp}$, we must have $\langle y, b\rangle = 0$. Then by definition, $b \in (\mathcal{R}(A)^{\perp})^{\perp} = \mathcal{R}(A)$ But $b \notin \mathcal{R}(A)$, so we have reached a contradiction.

Thus, if $Ax = b$ has no solution $x \in \mathbb{F}^n$, then $\exists\ y \in \mathcal{N}(A^H)$ s.t. $\langle y, b\rangle \neq 0$. For the case in which $Ax = b$ has some solution $x \in \mathbb{F}^n$, then we know that $b \in \mathcal{R}(A)$. So, $\forall y \in \mathcal{N}(A^H) = \mathcal{R}(A)^{\perp}$, we must have $\langle y, b\rangle = 0$ by the definition of $\mathcal{R}(A)^{\perp}$. So, we have established the Fredholm alternative. □

Problem 45

Proof. By definition,

$$Sym_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid A^T = A\}$$

$$Skew_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid A^T = -A\}$$

We must show that:

$$Skew_n(\mathbb{R}) = Sym_n(\mathbb{R})^{\perp} = \{A \in M_n(\mathbb{R}) \mid \langle B, A\rangle = 0\ \forall\ B \in Sym_n(\mathbb{R})\}$$

Consider $A \in Sym_n(\mathbb{R})^{\perp}$, $B \in Sym_n(\mathbb{R})$ defined as:

$$A := \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \dots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \dots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \dots & \alpha_{n,n} \end{bmatrix}$$

$$B := \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \dots & \beta_{1,n} \\ \beta_{2,1} & \beta_{2,2} & \dots & \beta_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n,1} & \beta_{n,2} & \dots & \beta_{n,n} \end{bmatrix}$$

Then, we have that:

$$\langle B, A\rangle = tr(B^T A) = tr(BA) = 0$$

Note that:

$$tr(BA) = \sum_{j=i}(\beta_{i,j}\alpha_{i,j}) + \sum_{j>i}(\beta_{i,j}\alpha_{j,i} + \beta_{j,i}\alpha_{i,j}) \text{ for } 1 \leq j \leq n,\ 1 \leq i \leq n$$

In order for $tr(BA) = 0 \ \forall \ B \in Sym_n(\mathbb{R})$, we must have that for $1 \leq j \leq n$, $1 \leq i \leq n$, $j = i$, $(\beta_{i,j}\alpha_{i,j}) = 0$, and for $1 \leq j \leq n$, $1 \leq i \leq n$, $j > i$, $(\beta_{i,j}\alpha_{j,i} + \beta_{j,i}\alpha_{i,j}) = 0$ for all choices of $B \in Sym_n(\mathbb{R})$ with B defined as above. This is easily proved by contradiction.

It follows that for $1 \leq j \leq n$, $1 \leq i \leq n$, $j = i$, $\alpha_{i,j} = 0$, and for $1 \leq j \leq n$, $1 \leq i \leq n$, $j > i$, $\alpha_{j,i} = -\alpha_{i,j}$, since $\beta_{i,j} = \beta_{j,i}$. If A satisfies these conditions, then $A \in Skew_n(\mathbb{R})$, and if $A \in Skew_n(\mathbb{R})$, then A satisfies these conditions. Thus, we have that

$$Skew_n(\mathbb{R}) = Sym_n(\mathbb{R})^\perp$$

□

Problem 46

Part (i)

Proof. Clearly, $Ax \in \mathcal{R}(A)$, by definition. Suppose $x \in \mathcal{N}(A^H A)$. Then $A^H(Ax) = 0$, so $Ax \in \mathcal{N}(A^H)$. □

Part (ii)

Proof. Suppose $x \in \mathcal{N}(A^H A)$. Then:

$$\begin{aligned} A^H Ax &= 0 \\ x^H A^H Ax &= 0 \\ (Ax)^H Ax &= 0 \\ Ax &= 0 \\ x &\in \mathcal{N}(A) \end{aligned}$$

Suppose $x \in \mathcal{N}(A)$. Then $Ax = 0$, so $A^H Ax = 0$, so $x \in \mathcal{N}(A^H A)$. Thus, $\mathcal{N}(A^H A) = \mathcal{N}(A)$ □

Part (iii)

Proof. For an $m \times n$ matrix A, by the rank-nullity theorem we have that:

$$n = dim(\mathcal{R}(A)) + dim(\mathcal{N}(A))$$

For an $n \times n$ matrix $A^H A$, by the rank-nullity theorem we have that:

$$n = dim(\mathcal{R}(A^H A)) + dim(\mathcal{N}(A^H A))$$

B part (ii), we have that $\mathcal{N}(A^H A) = \mathcal{N}(A)$, and in particular that $dim(\mathcal{N}(A^H A)) = dim(\mathcal{N}(A))$. So, we have that $dim(\mathcal{R}(A)) = dim(\mathcal{R}(A^H A))$, i.e. A and $A^H A$ have the same rank. □

Part (iv)

Proof. If A has linearly independent columns, then it has rank n. By part (iii), $A^H A$ must also have rank n. Since $A^H A$ is an $n \times n$ matrix then (i.e. it has n columns), it must be nonsingular. □

Problem 47

Part (i)

Proof. Define:

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

Then we have that:

$$\begin{aligned} P^2 &= A(A^H A)^{-1} A^H A (A^H A)^{-1} A^H \\ &= A(A^H A)^{-1}((A^H A)(A^H A)^{-1}) A^H \\ &= A(A^H A)^{-1} A^H \\ &= P \end{aligned}$$

□

Part (ii)

Proof.

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

□

Part (iii)

Proof. Note that $rank(A^H) = n$ since row-rank = column-rank. Also, note that $(A^H A)^{-1}$ is an $n \times n$ matrix of rank n since by problem 46, $A^H A$ has rank n, and a matrix has the same rank as its inverse. Then we must have $rank(A(A^H A)^{-1}) = n$ since A is an $m \times n$ matrix. We know that $rank(A(A^H A)^{-1} A^H) \leq min\{rank(A(A^H A)^{-1}), rank(A^H)\}$, with equality holding since $rank(A(A^H A)^{-1}) = rank(A^H) = n$. Thus, $rank(P) = rank(A(A^H A)^{-1} A^H) = n$. □

Problem 48

Part (i)

Proof.

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

$$\begin{aligned}
P(cA) &= \frac{cA+(cA)^T}{2} \\
&= \frac{cA+c(A^T)}{2} \\
&= c \cdot \frac{A+A^T}{2} \\
&= cP(A)
\end{aligned}$$

Thus, P is linear. □

Part (ii)

Proof.

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

□

Part (iii)

Proof.

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

□

Part (iv)

Proof. If $P(A) = 0$, then $\frac{A+A^T}{2} = 0$, so $A + A^T = 0$, i.e. $A^T = -A$. Thus, $A \in Skew_n(\mathbb{R})$. Also, if $A \in Skew_n(\mathbb{R})$, it is clear that $P(A) = 0$. Thus $\mathcal{N}(P) = Skew_n(\mathbb{R})$ □

Part (v)

Proof.

$$\begin{aligned} \mathcal{R}(P)^\perp &= \mathcal{R}(P^T)^\perp && \text{by part (iii)} \\ &= \mathcal{N}(P) && \text{by the fundamental subspaces theorem} \\ &= Skew_n(\mathbb{R}) && \text{by part (iv)} \\ &= Sym_n(\mathbb{R})^\perp && \text{by problem 45} \end{aligned}$$

Thus, we have that $\mathcal{R}(P)^\perp = Sym_n(\mathbb{R})^\perp$, so $\mathcal{R}(P) = Sym_n(\mathbb{R})$ □

Part (vi)

Proof.

$$\begin{aligned}
\|A - P(A)\|_F &= \sqrt{tr((A - P(A))^T(A - P(A)))} \\
&= \sqrt{tr\left(\left(A - \frac{A + A^T}{2}\right)^T \left(A - \frac{A + A^T}{2}\right)\right)} \\
&= \sqrt{tr\left(\left(\frac{A - A^T}{2}\right)^T \left(\frac{A - A^T}{2}\right)\right)} \\
&= \sqrt{tr\left(\frac{A^T - A}{2} \cdot \frac{A - A^T}{2}\right)} \\
&= \sqrt{tr\left(\frac{A^TA - A^TA^T - AA + AA^T}{4}\right)} \\
&= \sqrt{\frac{tr(A^TA) + tr(AA^T) - tr((AA)^T) - tr(AA)}{4}} \\
&= \sqrt{\frac{tr(A^TA) + tr(A^TA) - tr(AA) - tr(AA)}{4}} \\
&= \sqrt{\frac{2tr(A^TA) - 2tr(AA)}{4}} \\
&= \sqrt{\frac{tr(A^TA) - tr(A^2)}{2}}
\end{aligned}$$

□

Problem 50

Define the following (note that b is an n-vector):

$$A \coloneqq \begin{bmatrix} x_1^2 & y_1^2 \\ x_2^2 & y_2^2 \\ \vdots & \vdots \\ x_n^2 & y_n^2 \end{bmatrix}$$

$$x \coloneqq \begin{bmatrix} r \\ s \end{bmatrix}$$

$$b \coloneqq \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$