Math Problem Set #2: Inner Product Spaces

OSM Lab - University of Chicago
Geoffrey Kocks

Problem 1: HJE 3.1.

Solution.

(i). The right side of the polarization identity is equivalent to:

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

Therefore we have verified the polarization identity that:

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

(ii). The left side of the parallelogram identity is equivalent to:

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \langle \mathbf{x},\mathbf{x}\rangle + \langle \mathbf{y},\mathbf{y}\rangle.$$

The right side of the parallelogram identity is equivalent to:

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

Therefore we have verified the parallelogram identity that:

$$\boxed{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \tfrac{1}{2}(\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2)}.$$

Problem 2: HJE 3.2.

Solution. The right side of the polarization identity for complex numbers is equivalent to:

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

Therefore we have verified the complex inner product space identity that:

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

Problem 3: HJE 3.3.

Solution.

(i). The angle θ between x and x^5 is given by:

$$\cos\theta = \frac{\langle x, x^5\rangle}{\|x\|\|x^5\|} = \frac{\int_0^1 x^6 dx}{\sqrt{(\int_0^1 x^2 dx)(\int_0^1 x^{10} dx)}} = \frac{\frac{1}{7}}{\sqrt{(\frac{1}{3})(\frac{1}{11})}} = \frac{\sqrt{33}}{7}.$$

$$\boxed{\theta = \arccos \tfrac{\sqrt{33}}{7}.}$$

(ii). The angle θ between x^2 and x^4 is given by:

$$\cos\theta = \frac{\langle x^2, x^4\rangle}{\|x^2\|\|x^4\|} = \frac{\int_0^1 x^6 dx}{\sqrt{(\int_0^1 x^4 dx)(\int_0^1 x^8 dx)}} = \frac{\frac{1}{7}}{\sqrt{(\frac{1}{5})(\frac{1}{9})}} = \frac{\sqrt{45}}{7}.$$

$$\boxed{\theta = \arccos \tfrac{\sqrt{45}}{7}.}$$

Problem 4: HJE 3.8.

Solution.

(i). A collection S of $\{x_i\}_{i\in J}$ is an orthonormal set if for all i, $j \in J$ we have $\langle \mathbf{x}_i, \mathbf{x}_j\rangle = 1$ if $i = j$ and 0 if $i \neq j$. With the given inner product definition, we verify this definition as follows (all integrals solved using Wolfram Alpha):

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

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

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

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

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

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

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

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

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

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

(ii). The norm of t is given by:

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

(iii). The projection is given by:

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

(iv). The projection is given by:

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

Problem 5: HJE 3.9.

Solution. Define $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$. Then the standard inner product space of the column vectors is $x_1y_1 + x_2y_2$. The rotation transformation is orthonormal if the standard inner product space of the transformation preserves the original inner product space. The transformations result in the following:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1\cos\theta - x_2\sin\theta \\ x_1\sin\theta + x_2\cos\theta \end{pmatrix}$$

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1\cos\theta - y_2\sin\theta \\ y_1\sin\theta + y_2\cos\theta \end{pmatrix}$$

Taking the standard inner product space (dot product) of the two resulting column vectors gives:

$$\begin{aligned} &(x_1\cos\theta - x_2\sin\theta)(y_1\cos\theta - y_2\sin\theta) + (x_1\sin\theta + x_2\cos\theta)(y_1\sin\theta + y_2\cos\theta) = \\ &x_1y_1(\cos\theta)^2 + x_2y_2(\sin\theta)^2 - \sin\theta\cos\theta(x_1y_2 + x_2y_1) + x_1y_1(\sin\theta)^2 \\ &+ x_2y_2(\cos\theta)^2 + \sin\theta\cos\theta(x_1y_2 + x_2y_1) = \\ &x_1y_1((\sin\theta)^2 + (\cos\theta)^2) + x_2y_2((\sin\theta)^2 + (\cos\theta)^2) = x_1y_1 + x_2y_2. \end{aligned}$$

The resulting inner product space is the same as the original inner product space so the rotation transformation is orthonormal.

Problem 6: HJE 3.10.

Solution.

(i.) The matrix Q being orthonormal is equivalent to $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. This implies that $Q^H Q\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. Because these are equal, then $Q^H Q = I$. By the properties of matrix inverses, $AB = I$ implies $BA = I$, so we also have that $QQ^H = I$. The same reasoning applies for the converse, so that if the product is I, then it directly follows that $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. Therefore a matrix Q is orthonormal if and only if $Q^H Q = QQ^H = I$.

(ii.) $\|Q\mathbf{x}\| = \sqrt{\langle \mathbf{x}^H Q^H, Q\mathbf{x}\rangle} = \sqrt{(QQ^H)\langle \mathbf{x}^H, \mathbf{x}\rangle} = \sqrt{\langle \mathbf{x}^H, \mathbf{x}\rangle}$ by the property proved in part (i). The last square root is equivalent to $\|\mathbf{x}\|$. Therefore if Q is an orthonormal matrix, then $\|Q\mathbf{x}\| = \|\mathbf{x}\|$.

(iii.) By part (i) of this problem $Q^H Q = QQ^H = I$. By the definition of the matrix inverse: $Q^{-1}Q = QQ^{-1} = I$. Therefore, for an orthonormal matrix Q: $Q^{-1} = Q^H$. We also have that for any Q by definition, $(Q^{-1})^{-1} = Q$ and $(Q^H)^H = Q$. Thus: $(Q^{-1})^{-1} = (Q^H)^H = (Q^{-1})^H$. Because the H and inverse of Q^{-1} are equal, Q^{-1} is also an orthogonal matrix.

(iv.) By part (i) of this problem, $QQ^H = I$. Each entry in the product QQ^H is an inner product space of columns of Q, and this inner product space equals 1 along the diagonals and 0 elsewhere. This implies that $\langle \mathbf{x}_i, \mathbf{x}_j\rangle = 1$ only when $i = j$ and 0 any other time, satisfying the definition of an orthonormal set.

(v.) By definition, $\det I = 1$. By part (i), since $Q^H Q = I$, $\det Q^H Q = 1$. By the properties of determinants, this implies that $\det Q^2 = 1 \implies |\det Q| = 1$. However the converse is not true because there are matrices that have a determinant of 1 that are not orthonormal. One example is the matrix: $\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}$ which has a determinant of 1 but does not satisfy the property proved in part (i) of this problem.

(vi.) Let Q_1 and Q_2 both be orthonormal matrices. Then $(Q_1 Q_2)^{-1} = (Q_2)^{-1}(Q_1)^{-1}$. By part (i) this is equivalent to $(Q_2)^H(Q_1)^H = (Q_1 Q_2)^H$. Since the H is the same as the transpose, the product $Q_1 Q_2$ is also orthonormal.

Problem 7: HJE 3.11.

Solution. We show below that attempting to apply the Gram-Schmidt orthonormalization process to a collection of linearly dependent vectors will not work because we will end up dividing by 0.

Denote the set of linearly dependent vectors as $\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n}$. Let $\mathbf{x_j}$ be the first vector that is a linear combination of preceding vectors. The beginning of the Gram-Schmidt process gives:

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

Projecting $\mathbf{x_2}$ onto $\mathbf{q_1}$ gives:

$$\mathbf{p_1} = \langle \mathbf{q_1}, \mathbf{x_2}\rangle \mathbf{q_1}.$$

Define $\mathbf{r_i} = \mathbf{x_i} - \mathbf{p_{i-1}}$. Then the next vector in the construction of the orthonormal set is defined by:

$$\mathbf{q_2} = \frac{\mathbf{r_2}}{\|\mathbf{r_2}\|}.$$

We continue this process until we get to $\mathbf{r_j} = \mathbf{x_j} - \mathbf{p_{j-1}} = \mathbf{x_j} - \langle \mathbf{q_{j-1}}, \mathbf{x_j} \rangle \mathbf{q_{j-1}} = 0$. Then in the next step, we will be dividing by 0, and cannot construct an orthonormal set.

Problem 8: HJE 3.16.

Solution.

(i.) The QR decomposition is not unique because if there are two matrices Q and R that multiply to equal another matrix A, then we can obtain the same product by multiplying every value in Q by α and every value in R by $\frac{1}{\alpha}$. As an example, we take the Matrix A from Example 3.3.11 of the textbook. We can obtain a decomposition of the same matrix by multiplying Q and R each by -1, obtaining:

$$Q = \begin{pmatrix} -1/2 & 1/2 & -1/2 \\ -1/2 & -1/2 & 1/2 \\ -1/2 & -1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \end{pmatrix}$$

$$R = \begin{pmatrix} -2 & -1 & -3 \\ 0 & -5 & 1 \\ 0 & 0 & -3 \end{pmatrix}$$

(ii.) Assume that A is invertible and has two different QR decompositions with only positive diagonal elements. Then $A = Q_1R_1 = Q_2R_2 \implies Q_2^{-1}Q_1 = R_2R_1^{-1}$. The matrix $Q_2^{-1}Q_1$ is orthonormal and $R_2R_1^{-1}$ is upper triangular. This means that each of the matrix products will be upper triangular and orthonormal, which occurs if the product is I or $-I$. We are assuming that there are positive diagonals so the product must be I. Then: $Q_2^{-1}Q_1 = I \implies Q_2 = Q_1$. $R_2R_1^{-1} = I \implies R_1 = R_2$.

Therefore there is a unique QR decomposition.

Problem 9: HJE 3.17.

Solution. We start with the assumption that $A^HA\mathbf{x} = A^Hb$. When A can be decomposed as $A = \hat{Q}\hat{R}$:

$$\begin{aligned} A^HA\mathbf{x} = A^Hb &\implies (\hat{Q}\hat{R})^H(\hat{Q}\hat{R})\mathbf{x} = (\hat{Q}\hat{R})^H\mathbf{b} \\ &\implies \hat{R}^H\hat{Q}^H\hat{Q}\hat{R}\mathbf{x} = \hat{R}^H\hat{Q}^H\mathbf{b} \\ &\implies \hat{R}^H\hat{R}\mathbf{x} = \hat{R}^H\hat{Q}^H\mathbf{b} \\ &\implies \hat{R}\mathbf{x} = \hat{Q}^H\mathbf{b}. \end{aligned}$$

Problem 10: HJE 3.23.

Solution. By the triangle inequality:

$$\|x\| = \|x - y + y\| \le \|x - y\| + \|y\| \implies \|x\| - \|y\| \le \|x - y\|.$$

$$\|y\| = \|y - x + x\| \le \|x - y\| + \|x\| \implies \|y\| - \|x\| \le \|x - y\|.$$

Therefore $\boxed{|\|x\| - \|y\|| \le \|x - y\|.}$

Problem 11: HJE 3.24.

Solution. There are three necessary conditions for a map to be considered a norm:

(1) Positivity: $\|\mathbf{x}\| \ge 0$ and $\|\mathbf{x}\| = 0$ only if $\mathbf{x} = 0$.
(2) Scale preservation: $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$.
(3) Triangle Inequality: $\|\mathbf{x} + \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$.

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

(1) $|f(t)| \ge 0$ for all $f(t)$ so the integral from a to b will always be at least 0. The integral will also only equal 0 when $f(t) = 0$ for all t in the domain.

(2) $\|\alpha\mathbf{x}\|_{L^1} = \int_a^b |\alpha f(t)|dt = |\alpha| \int_a^b |f(t)|dt = |\alpha|\|\mathbf{x}\|$.

(3) $\|f + g\|_{L^1} = \int_a^b |f(t) + g(t)|dt \le \int_a^b |f(t)|dt| + \int_a^b |g(t)|dt = \|f\| + \|g\|$.

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

(1) Once again, $|f(t)|$ is non-negative, so its square will be non-negative, as well as the integral of this value and its square root. $\|f\|_{L^2} = 0$ implies that $|f(t)| = 0$ which only occurs when $f(t) = 0$ for all t in the domain.

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

(3)

$$\begin{aligned}
\|f + g\|_{L^2}^2 = \int_a^b |f + g|^2 dt &\le \int_a^b (|f|^2 + 2|f||g| + |g|^2) dt \\
&\le \|f\|_{L^2}^2 + \|g\|_{L^2}^2 + 2\|f\|_{L^2}\|g\|_{L^2} = \|f\|_{L^2}^2 + \|g\|_{L^2}^2 \\
&\implies \|f + g\|_{L^2}^2 \le \|f\|_{L^2}^2 + \|g\|_{L^2}^2.
\end{aligned}$$

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

(1) Because each norm $|f(x)|$ is always non-negative, its supremum will also be non-negative. The norm will only equal 0 when $f(x)$ is always 0 in the domain.

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

(3)

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

Problem 12: HJE 3.26.

Solution. First we note that topological equivalence is an equivalence relation. Topological equivalence is an equivalence relation because it if norm $\|.\|_a$ and $\|.\|_b$ are topologically equivalent, and $\|.\|_b$ and $\|.\|_c$ are topologically equivalent, then it is also true that $\|.\|_a$ and $\|.\|_c$ are topologically equivalent. The equivalences imply:

$$m_1\|x\|_b \leq \|x\|_a \leq M_1\|x\|_b.$$

$$m_2\|x\|_b \leq \|x\|_c \leq M_2\|x\|_b.$$

$$\frac{m_1}{m_2} \leq \frac{\|x\|_a}{\|x\|_c} \leq \frac{M_1}{M_2}.$$

$$\frac{m_1\|x\|_c}{m_2} \leq \|x\|_a \leq \frac{M_1\|x\|_c}{M_2}.$$

Therefore we also will have an equivalence between $\|.\|_a$ and $\|.\|_c$.

We now show that the p-norms for $p = 1, 2, \infty$ are topologically equivalent:

(i.) Let $\|x\|_2$ denote the p-norm for p = 2 and let $\|x\|_1$ denote the p-norm for p=1 on $\mathbb{F}^n$. First, by the Cauchy-Schwartz inequality:

$$(\sum |x_j|)^2 \leq n \sum |x_j^2|.$$

$$\sum |x_j| \leq \sqrt{n \sum |x_j^2|}.$$

$$\|x\|_1 \leq \sqrt{n}\|x\|_2.$$

It also holds that:

$$(\sum x_i)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j = \sum_{i=1}^{n} x_i^2 + \sum_{i}\sum_{j\neq i} x_i x_j \geq \sum x_i^2.$$

$$\sum |x_i| \geq \sqrt{\sum |x_i|^2}.$$

$$\|x\|_1 \geq \|x\|_2.$$

Therefore the p-norms for p=1 and p=2 are topologically equivalent.

(ii.) Let $\|x\|_2$ denote the p-norm for p = 2 and let $\|x\|_\infty$ denote the p-norm for $p = \infty$ on $\mathbb{F}^n$. First, we have:

$$\|x\|_2^2 = \sum |x_i|^2 \leq n \sup |x_i|^2 = n(\sup |x_i|)^2 = n\|x\|_\infty^2.$$

$$\|x\|_2 \leq \sqrt{n}\|x\|_\infty.$$

where $\sup |x_i|$ is shorthand for $\sup(|x_1|, |x_2|, ..., |x_n|)$. We also have:

$$(\sup |x_i|)^2 = (\sup |x_i|^2) \leq |x_i^2| + \sum_{j\neq i} |x_j^2| = \|x\|_2^2.$$

$$\|x\|_\infty \leq \|x\|_2.$$

Therefore the p-norms for $p = \infty$ and $p = 2$ are also topologically equivalent.

Problem 13: HJE 3.28.

Solution.

(i.) Let $\|A\|_2$ denote the operator 2-norm and let $\|A\|_1$ denote the operator 1-norm. Then by the properties proven in Exercise 3.26:

$$\frac{1}{\sqrt{n}}\|A\|_2 = \frac{1}{\sqrt{n}} sup \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq sup \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_1} \leq sup \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \|A\|_1$$

Similarly:

$$\|A\|_1 \leq sup \frac{\sqrt{n}\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_1} \leq \sqrt{n} sup \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sqrt{n}\|A\|_2$$

(ii.) Let $\|A\|_2$ denote the operator 2-norm and let $\|A\|_\infty$ denote the operator ∞-norm. Then by the properties proved in Exercise 3.26:

$$\frac{1}{\sqrt{n}}\|A\|_\infty = \frac{1}{\sqrt{n}} sup \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \leq sup \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_2} \leq sup \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \|A\|_2$$

Similarly:

$$\|A\|_2 \leq sup \frac{\sqrt{n}\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_2} \leq \sqrt{n} sup \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \sqrt{n}\|A\|_\infty$$

Problem 14: HJE 3.29.

Solution.

We first show that the induced norm of the transformation is equal to $\|\mathbf{x}\|_2$. In Exercise 3.10(ii), we showed that if Q is an orthonormal matrix, then $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$. Therefore, when Q is orthonormal, the 2-norm is given by:

$$\|Q\|_2 = sup \frac{\|Q\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = sup \frac{\|\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = 1.$$

We now show that the induced norm of the transformation R_x is equal to $\|\mathbf{x}\|_2$. By definition, $\|\mathbf{x}\|_2^2 = \mathbf{x}^H\mathbf{x}$. Then $\|A\mathbf{x}\|_2^2 = (A\mathbf{x})^H(A\mathbf{x}) = \mathbf{x}^H A^H A\mathbf{x} = \mathbf{x}^H\mathbf{x} = \|\mathbf{x}\|_2^2$. Therefore the induced norm of the transformation is equal to $\|\mathbf{x}\|_2$.

Problem 15: HJE 3.30.

Solution. There are three necessary conditions for a map to be considered a norm:

(1) Positivity: $\|\mathbf{x}\| \geq 0$ and $\|\mathbf{x}\| = 0$ only if $\mathbf{x} = 0$.
(2) Scale preservation: $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$.
(3) Triangle Inequality: $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

(1) We already know that $\|.\|$ is a matrix norm so it satisfies the conditions of a norm. Therefore $\|\mathbf{x}\|_S \geq 0$. We also know that it will only equal 0 if $SAS^{-1} = 0$, which only occurs when $A = 0$. Therefore positivity holds.

(2)

$$\|\alpha A\|_S = \|\alpha SAS^{-1}\| = |\alpha|\|SAS^{-1}\| = |\alpha|\|A\|_S$$

The steps above follow from the fact that $\|.\|$ is a matrix norm. Therefore scale preservation holds.

(3)

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

The steps above hold by the triangle inequality of $\|.\|$ because it is already a matrix norm. Therefore the triangle inequality also holds for $\|.\|_S$. Because all three properties hold, $\|.\|_S$ is considered a matrix norm.

Problem 16: HJE 3.37.

Solution. The Riesz representation theorem tells us that there will be a unique $q \in V$ such that $L(p) = \langle q, p\rangle$ but it does not tell us what this q will be. Define a generic $q = ex^2 + fx + g$ and $p = ax^2 + bx + c$. We want to find e, f, g such that $\langle q, p\rangle = \int_0^1 qpdx = L(p) = 2a + b$. This gives us:

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

$$\int_0^1 aex^4 + (be + af)x^3 + (ce + bf + ag)x^2 + (cf + bg)x + cg =$$
$$\frac{ae}{5} + \frac{be+af}{4} + \frac{ce+bf+ag}{3} + \frac{cf+bg}{2} + cg = a(\frac{e}{5} + \frac{f}{4} + \frac{g}{3}) + b(\frac{e}{4} + \frac{f}{3} + \frac{g}{2}) + c(\frac{e}{3} + \frac{f}{2} + g)$$

We then set this equation equal to $2a + b$ and solve the following system of equations:

$$0.2e + 0.25f + (1/3)g = 2$$
$$0.25e + (1/3)f + 0.5g = 1$$
$$(1/3)e + 0.5f + g = 0$$

This yields solution $e = 180, f = -168, g = 24$. Therefore, the unique q guaranteed by the Riesz representation theorem is: $\boxed{180x^2 - 168x + 24.}$

Problem 17: HJE 3.38.

Solution. With respect to the power basis, we seek a D such that:

$$D(1, 0, 0) = (0, 0, 0).$$
$$D(0, 1, 0) = (1, 0, 0).$$
$$D(0, 0, 1) = (0, 2, 0).$$

This is satisfied by:

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

Following example 3.7.8 in the textbook, the adjoint is given by the Hermitian conjugate D^H so:

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

Problem 18: HJE 3.39.

Solution.

(i.)

$$\begin{aligned} \langle \mathbf{v}, (S+T)^*, \mathbf{w}\rangle = \langle (S+T)\mathbf{x}, \mathbf{w}\rangle &= \langle S\mathbf{v}, \mathbf{w}\rangle + \langle T\mathbf{v}, \mathbf{w}\rangle \\ &= \langle \mathbf{v}, T^*\mathbf{w}\rangle + \langle \mathbf{v}, S^*\mathbf{w}\rangle = \langle \mathbf{v}, (T^* + S^*)\mathbf{w}\rangle. \end{aligned}$$

Therefore, we have that $\boxed{(S+T)^* = S^* + T^*.}$

Additionally, $\langle Q_1, \alpha T(Q_2)\rangle = \alpha\langle T^*(Q_1), Q_2\rangle = \langle \overline{\alpha} T^*(Q_1), Q_2\rangle$.

(ii.) We have that $\langle \mathbf{v}, S\mathbf{w}\rangle = \langle S^*\mathbf{v}, \mathbf{w}\rangle$. By reflexive properties, we therefore also have that $\langle S\mathbf{w}, \mathbf{v}\rangle = \langle \mathbf{w}, S^*\mathbf{w}\rangle$. Finally, by the definition of the adjoint: $\langle \mathbf{w}, S^*\mathbf{v}\rangle = \langle (S^*)^*\mathbf{w}, \mathbf{v}\rangle$. Therefore, $\boxed{(S^*)^* = S.}$

(iii.)

$$\langle Q_1, (ST)^*Q_2\rangle = \langle STQ_1, Q_2\rangle = \langle TQ_1, S^*Q_2\rangle = \langle Q_1, T^*S^*Q_2\rangle.$$

Therefore $\boxed{(ST)^* = T^*S^*.}$

(iv.) We have that:

$$\langle T^*(T^{-1})^*Q_1, Q_2\rangle = \langle Q_1, T^{-1}TQ_2\rangle = \langle Q_1, Q_2\rangle = \langle Q_1, Q_2\rangle.$$

Therefore, because the product of T^* and $(T^{-1})^*$ is the identity matrix, they are each the inverse of each other, so: $\boxed{(T^{-1})^* = (T^*)^{-1}.}$

Problem 19: HJE 3.40.

Solution.

(i.) $\langle A^*Q_1, Q2\rangle = \langle Q_1, AQ_2\rangle = tr(Q_1^H AQ_2)$. Equivalently, $\langle A^H Q_1, Q_2\rangle = tr(Q_1^H AQ_2)$. Therefore $A^* = A^H$.

(ii.) For any A_1, A_2, A_3 :

$$\langle A_2, A_3A_1\rangle = tr(A_2^H A_3 A_1) = tr(A_1 A_2^H A_3) = \langle A_2 A_1^H, A_3\rangle = \langle A_2 A_1^*, A_3\rangle.$$

Therefore $\langle A_2, A_3A_1\rangle = \langle A_2A_1^*, A_3\rangle$.

(iii.) Define $T_A(X) = AX - XA$. First we have:

$$\langle T_{A^*}(Q_1), Q_2\rangle = \langle A^*Q_1 - Q_1A^*, Q_2\rangle = \langle A^HQ_1 - Q_1A^H, Q_2\rangle = tr(Q_1^HAQ_2 - Q_1^HQ_2A).$$

We also have:

$$\langle T_A^*(Q_1), Q_2\rangle = \langle Q_1, T_A(Q_2)\rangle = \langle Q_1, AQ_2 - Q_2A\rangle = tr(Q_1^HAQ_2 - Q_1^HQ_2A).$$

Therefore $(T_A)^* = T_{A^*}$.

Problem 20: HJE 3.44.

Solution. To prove the Fredholm alternative, we frst show that if $A\mathbf{x} = \mathbf{b}$ has no solution, then there is no $\mathbf{y}$ in the null of A^H s.t. $\langle \mathbf{y}, \mathbf{b}\rangle = 0$. Assume that there is a solution to the equation and let $\mathbf{y}$ be part of the null space as described above. Then:

$$\langle \mathbf{y}, \mathbf{b}\rangle = \langle \mathbf{y}, \mathbf{Ax}\rangle = tr(\mathbf{y}^HA\mathbf{x}) = \langle A^H\mathbf{y}, \mathbf{x}\rangle = \langle 0, \mathbf{x}\rangle = 0.$$

We next show that if $\langle \mathbf{y}, \mathbf{b}\rangle \neq 0$, then there is no solution to the equation. $\mathbf{y}$ is in the null space of A^H which also means that $\mathbf{y}$ is in $R(A)^\perp$. However, if there is a solution to the equation, then $\mathbf{b}$ is in $R(A)$. There is a contradiction since for $\langle \mathbf{y}, \mathbf{b}\rangle$ to not equal 0, $\mathbf{b}$ cannot be in $R(A)$. There can be no solution to the equation. We have then proven the Fredholm alternative.

Problem 21: HJE 3.45.

Solution. We define a linear transformation $L : M_n(\mathbb{R}) \to M_n(\mathbb{R})$. as $L(A) = A + A^T$. The definition of Skew gives: $Skew_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) | A^T = -A\}$. For linear transformation L this is exactly equivalent to the null space because $L(A) = 0$ when $A = -A^T$.

Next, we show that $Sym_n(\mathbb{R}) = R(L)$. The definition of Sym gives: $Sym_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) | A^T = A\}$. Because $(A + A^T)^T = A + A^T$ the kernel of L is part of $Sym_n(\mathbb{R})$. We also note that for any matrix A in $Sym_n(\mathbb{R})$: $A = A^T$ so $A = 0.5A^T + 0.5A$ so $Sym_n(\mathbb{R})$ is also part of the kernel of L. Therefore $Sym_n(\mathbb{R}) = R(L)$.

Finally, we seek to prove that $L = L^*$. We note that $tr(BA) = tr(AB) \implies tr(A^TB^T) = tr(AB) \implies tr(A^T(B + B^T)) = tr((A + A^T)B) \implies \langle A, L(B)\rangle = \langle L(A), B\rangle$. Therefore $L = L^*$.

We already have that $Sym_n(\mathbb{R})^\perp = R(L)^\perp$. By the Fundamental Subspaces Theorem, this is then equal to $N(L^*) = N(L)$ since $L = L^*$. Since $N(L) = Skew_n(\mathbb{R})$, we have that $Sym_n(\mathbb{R})^\perp = Skew_n(\mathbb{R})$.

Problem 22: HJE 3.46.

Solution.

(i.) Since $\mathbf{x}$ is in the null space of A^HA: $A^HA\mathbf{x} = 0 \implies A^H(A\mathbf{x}) = 0 \implies A\mathbf{x}$ is in the null space of A^H. By definition $A\mathbf{x}$ is also in the kernel of A.

(ii.) By definition (and from the result of part (i): $N(A) \subseteq N(A^HA)$. Now let $\mathbf{x} \in N(A^HA)$. Then $\mathbf{x}^HA^HA\mathbf{x} = 0 \implies (A\mathbf{x})^H(A\mathbf{x}) = 0 \implies A\mathbf{x} = 0$. Therefore $N(A^HA) \subseteq N(A)$. Then $N(A^HA) = N(A)$.

(iii.) From part (ii) above, $dimN(A) = dimN(A^HA) \implies n - rank(A) = n - rank(A^HA)$. Therefore, A and A^HA have the same rank.

(iv.) If A has linearly independent columns then $N(A) = 0$ and by part (ii), $N(A) = N(A^HA)$. Because the null space is 0, A^HA is invertible.

Problem 23: HJE 3.47.

Solution.

(i.) Let $P = A(A^HA)^{-1}A^H$. First we note that A^HA is a symmetric matrix. Therefore we can proceed to show the lemma that $(A^HA)^{-1} = ((A^HA)^{-1})^H$. Let $Q = (A^H)A$. By the properties of the matrix inverse, $QQ^{-1} = I$. Then:

$$I = I^H.$$

$$(QQ^{-1}) = (QQ^{-1})^H.$$
$$QQ^{-1} = (Q^{-1})^H Q^H.$$

And since $Q = Q^H$:

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

Having verified this, it is easy to calculate $P^2 = A((A^HA)^{-1})^H A^H A((A^HA)^{-1})A^H = A(A^HA)^{-1}A^H = P$. Therefore $P^2 = P$ so the operator P is a projection.

(ii.) Using the fact proved in part (i), we have $P^H = A((A^HA)^{-1})^H A^H = A(A^HA)^{-1}A^H = P$. Therefore $P^H = P$.

(iii.) First we note that by the result of Exercise 2.14(i), $rank(KL) \leq min(rank(L), rank(K))$ so we also have:

$$rank(P) \leq min(rank(A), rank((A^HA)^{-1}A^H)) \leq rank(A).$$

Then, let $\mathbf{b} \in R(A)$. This implies that $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x}$. We then have that: $P(\mathbf{b}) = A(A^HA)^{-1}A^HA\mathbf{x} = A\mathbf{x}$ so $\mathbf{b}$ is also in the range of P. Therefore the range of A is a subset of the range of P so:

$$rank(A) \leq rank(P).$$

Since $rank(A) \leq rank(P)$ and $rank(P) \leq rank(A)$, $rank(P) = rank(A) = n$.

Problem 24: HJE 3.48.

Solution.

(i.) P is linear if it preserves addition, preserves multiplication, and preserves 0. We first show that P preserves addition:

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

We next show that P preserves multiplication:

$$P(cA) = \frac{cA + (cA)^T}{2} = \frac{c(A + A^T)}{2} = cP(A).$$

Finally, P preserves 0 so it is linear:

$$P(0) = \frac{0 + 0}{2} = 0.$$

(ii.) We define P^2 as applying the linear transformation P twice consecutively. Then:

$$P^2 = \frac{1}{2}(\frac{A + A^T}{2} + (\frac{A + A^T}{2})^T) = \frac{1}{2}(\frac{A}{2} + \frac{A^T}{2} + \frac{A^T}{2} + \frac{A}{2}) = \frac{A + A^T}{2} = P.$$

Therefore $P^2 = P$.

(iii.)

$$\begin{aligned}
\langle A_1, P(A_2)\rangle &= \langle A_1, \frac{A_2 + A_2^T}{2}\rangle = \frac{1}{2}tr(A_1^T(A_2 + A_2^T)) \\
&= \frac{1}{2}tr(A_1^T A_2 + A_1^T A_2^T) = \frac{1}{2}tr(A_1^T A_2) + \frac{1}{2}tr(A_1^T A_2^T) \\
&= \frac{1}{2}tr(A_1^T A_2) + \frac{1}{2}tr(A_2 A_1) = \frac{1}{2}tr(A_1 A_2 + A_1^T A_2) \\
&= \langle P(A_1), A_2\rangle
\end{aligned}$$

Therefore $P^* = P$.

(iv.) $Skew_n(R)$ is defined as $\{A \in M_n(\mathbb{R}) | A^T = -A\}$. When $A^T = -A$, then $P(A) = 0$ so this is equivalent to being in the null space of P.

(v.) $Sym_n(R)$ is defined as $\{A \in M_n(\mathbb{R}) | A^T = A\}$. When $A^T = A$, then $\frac{1}{2}A = \frac{1}{2}A^T \implies A = \frac{1}{2}(A + A^T)$. so $A \in R(P)$.

(vi.)

$$\begin{aligned}
\|A - P(A)\| &= \sqrt{(\langle A - P(A), A - P(A)\rangle)} \\
&= \sqrt{\langle A/2 - A^T/2, A/2 - A^T/2\rangle} = \sqrt{tr((A^T - 2 - A/2)(A/2 - A^T/2))} \\
&= \sqrt{tr(A^T A/4 - A^2/4 + AA^T/4 - A^T A^T/4)} \\
&= \sqrt{\frac{tr(A^T A) - tr(A^2)}{2}}
\end{aligned}$$

Problem 25: HJE 3.50.

Solution. If we believe that data points $(x_i, y_i)_{i=1}^n$ lie roughly on an ellipse of form $rx^2 + sy^2 = 1$, we would solve the normal equation $A\mathbf{x} = b$ where:

$$A = \begin{pmatrix} 1 & x_1^2 \\ 1 & x_2^2 \\ ... & ... \\ 1 & x_n^2 \end{pmatrix}$$

$$b = \begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} \frac{1}{s} \\ \frac{-r}{s} \end{pmatrix}$$

We would then use this equation and the data points to solve for the unknowns r and s.