Math Homework #1

Eric C. Miller

July 5, 2017

Question #2

Part A

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

Part B

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

Question #2

$$\begin{aligned}\langle x,y\rangle &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x-iy\|^2 - i\|x+iy\|^2\right)\\ &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 - i[\|x-iy\|^2 + \|x+iy\|^2]\right)\\ &= \frac{1}{4}\left(\langle y,x\rangle + \langle x,y\rangle - \langle -y,x\rangle - \langle x,-y\rangle + i[-\langle -iy,x\rangle - \langle x,-iy\rangle + \langle x,iy\rangle + \langle iy,x\rangle]\right)\\ &= \frac{1}{4}\left(\langle y,x\rangle + \langle x,y\rangle - \langle -y,x\rangle - \langle x,-y\rangle + i[\langle iy,x\rangle + \langle x,iy\rangle + \langle x,iy\rangle + \langle iy,x\rangle]\right)\end{aligned} \tag{3}$$

Question #3

Part A

$$cos\theta = \frac{\langle x, x^5\rangle}{\sqrt{\langle x,x\rangle}\sqrt{\langle x^5,x^5\rangle}} = \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^2 dx}\sqrt{\int_0^1 x^1 0 dx}} = \frac{\frac{1}{7}}{\frac{1}{9}\frac{1}{121}} = 155 \tag{4}$$

Part B

$$cos\theta \;=\; \frac{\langle x^2, x^4\rangle}{\sqrt{\langle x^2, x^2\rangle}\sqrt{\langle x^4, x^4\rangle}} = \frac{\int_0^1 x^6 dx}{\sqrt{\int_0^1 x^4 dx}\sqrt{\int_0^1 x^8 dx}} = \frac{\frac{1}{7}}{\frac{1}{25}\frac{1}{81}} = 289 \tag{5}$$

Question #8

To prove that S is an orthonormal set, it is sufficient to show that $x, y \in S, \langle x, y\rangle = 0$ where $x \neq y$. We narrow the set of integrals necessary down to 6: $\langle sin(x), cos(x)\rangle$, $\langle sin(x), sin(x)\rangle$, $\langle sin(x), cos(2x)\rangle$, $\langle cos(x), sin(2x)\rangle$, $\langle cos(x), cos(2x)\rangle$, and $\langle sin(2x), cos(2x)\rangle$. We compute these below:

$$\begin{aligned}
\langle sin(x), cos(x)\rangle =& \frac{1}{\pi}\int_{-\pi}^{\pi} sin(x)cos(x)dx = 0\\
\langle sin(x), sin(2x)\rangle =& \frac{1}{\pi}\int_{-\pi}^{\pi} sin(x)sin(2x)dx = 0\\
\langle sin(x), cos(2x)\rangle =& \frac{1}{\pi}\int_{-\pi}^{\pi} sin(x)cos(2xdx = 0\\
\langle cos(x), sin(2x)\rangle =& \frac{1}{\pi}\int_{-\pi}^{\pi} cos(x)sin(2x)dx = 0\\
\langle cos(x), cos(2x)\rangle =& \frac{1}{\pi}\int_{-\pi}^{\pi} cos(x)cos(2x)dx = 0\\
\langle sin(2x), cos(2x)\rangle =& \frac{1}{\pi}\int_{-\pi}^{\pi} sin(2x)cos(2x)dx = 0
\end{aligned} \tag{6}$$

Hence, S is orthonormal.

$$\|t\| = \sqrt{\langle t, t\rangle} = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} t^2 dt} = \frac{2\pi^2}{3} \tag{7}$$

$$proj_x(cos(3t)) \;=\; \sum_{x_i \in S}\Big[\langle x_i, cos(3t)\rangle x_i\Big] \;=\; \sum_{x_i \in S}\left[\frac{x_i}{\pi}\int_{-\pi}^{\pi} x_i cos(3t)dt\right] = 0 + 0 + 0 + 0 = 0 \tag{8}$$

This stems from the fact that $cos(3t)$ is orthogonal to $s \in S, \forall s$.

$$proj_x(t) \;=\; \sum_{x_i \in S}\Big[\langle x_i, t\rangle x_i\Big] = \sum_{x_i \in S}\left[\frac{x_i}{\pi}\int_{-\pi}^{\pi} t x_i dt\right] = 2sin(t) - sin(2t) \tag{9}$$

Question #9

Let $A = M_2(\mathbb{R})$ be operated on by the 2x2 rotation matrix R. We use the Frobenius inner product to calculate the following:

$$\langle R(A), R(A)\rangle = tr(R(A)^T R(A)) = tr((A^T)^T A^T) = tr(AA^T) = tr(A^T A) = \langle A, A\rangle \tag{10}$$

Question #10

Recall the definition of an orthonormal matrix Assume the usual inner product in $\mathbb{F}^n$. Show:

Claim: The matrix $Q \in M_n(\mathbb{F})$ is an orthonormal matrix $\iff Q^H Q = QQ^H = I$.
Proof: ($\Rightarrow$) Let $Q \in M_n(\mathbb{F})$. Hence, Q^{-1} exists and is also an orthonormal matrix. As a matrix inverse is unique, it is sufficient to show that $Q^{-1} = Q^H$. This fact follows from Q having orthonormal columns in $\mathbb{F}$, which implies the rows of Q^H are orthonormal in $\mathbb{F}$. Hence, it follows from Definition 3.2.1 in Humphreys, Jarvis, & Evans (95) that $Q^H Q = QQ^H = I$.
($\Leftarrow$) Let $Q^H Q = QQ^H = I$. From this, we have that $Q^{-1} = Q^H$ (as matrix inverses are unique), and hence $\sum_{i=1}^{n} q_i h_i = 1$ and 0 otherwise (where q_i and h_i are rows or columns in Q and Q^H respectively, as Q^H is a valid inverse this is valid regardless of the direction.) However, this is exactly the definition of an orthonormal set; hence, Q is orthonormal.

Claim: If $Q \in M_n\mathbb{F}$ is an orthonormal matrix, then $\|Qx\| = \|x\| \ \forall x \in V$.
Proof: $\|Qx\| = \sqrt{\langle Qx, Qx\rangle} = \sqrt{(Qx)^T Qx} = \sqrt{x^T Q^H Q x} = \sqrt{x^T x} = \sqrt{\langle x, x\rangle} = \|x\|$

Claim: If $Q \in M_n\mathbb{F}$ is an orthonormal matrix, then so is Q^{-1}.
Proof: By Theorem 3.2.15, and that the matrix inverse is unique, $Q^{-1} = Q^H$. However, as $[Q^H]^H = Q$, we see that $[Q^H]^{-1} = Q$, which is an if and only if condition for being an orthonormal matrix. Thus, Q^{-1} is orthonormal.

Claim: The columns of an orthonormal matrix $Q \in M_n(\mathbb{F})$ are orthonormal.
Proof: Assume the columns of $Q \in M_n(\mathbb{F})$ are not orthonormal. This implies that $Q^H Q = I$ is not equal in general, as we could then have non-zero elements not along the diagonal. Hence we encounter a contradiction to the necessary condition of an orthonormal matrix $Q^H Q = QQ^H = I$.

Claim: if $Q \in M_n(\mathbb{F})$ is an orthonormal matrix, then $|det(Q)| = 1$. Is the converse true?.
Proof: Assume the $Q \in M_n(\mathbb{F})$, $|det(Q)| \neq 1$, or that Q is not unimodular. However, as $Q^{-1} = Q^H$ has the same elements as Q, this forms a contradiction. The converse is also not true, as for a unimodular matrix the inverse is not equal to the transpose in general.

Claim: If $Q_1, Q_2 \in M_n(\mathbb{F})$ are orthonormal matricies, then the product $Q_1 Q_2$ is also an orthonormal matrix.
Proof:

$$Q_1 Q_2 (Q_1 Q_2)^H = Q_1 Q_1 Q_2^H Q_1^H = Q_1 Q_1^H = I \tag{11}$$

The same argument follows for the inverse multiplication of the other side.

Question #11

Describe what happens when we apply the Gram-Schmidt orthonormalization process to a collection of linearly dependent vectors.
Let $X = \{x_i\}_{i=1}^n$ be a linearly dependent set, where $\exists k,\ 1 < k < n$ s.t. $x_{k+1} \in span\{X\}$. Observe that $p_{k+1} = \sum_{j=1}^{k+1} \langle q_j, x_{k+2}\rangle q_j = 0$, which then leads to a division by 0.

Question #16

Prove the following results about the QR decomposition:

Claim: The QR decomposition is not unique.
Proof: Where D is a diagonal matrix, $A = QR = QDD^{-1}R$. By letting $D = -I$, we see that $QD \neq Q$ and $D^{-1}R \neq R$. As $-Q$ is still orthonormal and $-R$ is still upper triangular, we conclude that this forms another QR decomposition.

Claim: If A is invertible, then there is a unique QR decomposition of A such that R has only positive diagonal elements..
Proof: Assume $A = Q_1R_1 = Q_2R_2$. Note that:.

$$\begin{aligned} R_1R_1^H =& R_1^HQ_1^HQ_1R_1 = A^HA = R_2^HQ_2^HQ_2^HR_2^H = R_2R_2^H \\ &\text{and thus} \\ (R_2^H)^{-1}R_1^H &= R_2R_1^{-1} \\ (R_2^{-1})^HR_1^H &= R_2R_1^{-1} \\ (R_1R_2^{-1})^H &= R_2R_1^{-1} \end{aligned} \tag{12}$$

Notice that the left side above is a lower-triangular matrix, and the right an upper-triangular matrix; hence, $R_2R_1^{-1}$ is a diagonal matrix. By restricting their diagonals to the positive reals, we can retain a positive diagonal for $R_2R_1^{-1}$. As A is nonsingular, R_1 and R_2 must have non-zero diagonals, and hence so does $R_2R_1^{-1}$.

Question #17

Let $A \in M_{mxn}$ have rank $n < m$, and let $A = \hat{Q}\hat{R}$ be a reduced QR decomposition.

Claim: Solving the system $A^HAx = A^Hb$ is equivalent to solving the system $\hat{R}x = \hat{Q}^Hb$
Proof:

$$\begin{aligned} \hat{R}x &= \hat{Q}^Hb \\ \hat{Q}^H\hat{Q}\hat{R}x &= \hat{Q}^Hb \\ \hat{R}^H\hat{Q}^H\hat{Q}\hat{R}x &= \hat{R}^H\hat{Q}^Hb \\ \left[\hat{R}\hat{Q}\right]^H\hat{Q}\hat{R}x &= \left[\hat{Q}\hat{R}\right]^Hb \\ A^HAx &= A^Hb \end{aligned} \tag{13}$$

Question #23

Let $(V, \|\cdot\|)$ be a normed linear space

Claim: $|\|x\| - \|y\|| \leq \|x - y\| \; \forall x, y \in V$.
Proof: $|\|x\| - \|y\|| \leq \|x - y\| \Rightarrow \|x\| - \|y\| \leq \|x - y\|$ and $\|y\| - \|x\| \leq \|x - y\|$. Notice, as a valid norm is strictly positive, we may prove the same problem by verifying $\|x\| - \|y\|^2 < \|x - y\|^2$ and $\|y\| - \|x\|^2 < \|x - y\|^2$ instead; this allows us to test both cases at once, as the square function is also strictly positive. We then

can show the following:

$$
\begin{aligned}
\|x-y\|^2 &= \|x\|^2 + \langle -y, x\rangle + \langle x, -y\rangle + \|-y\|^2 \\
&= \|x\|^2 - \langle x, y\rangle - \langle x, y\rangle + \|y\|^2 \\
&= \|x\|^2 - \big[\langle x, y\rangle + \langle x, y\rangle\big] + \|y\|^2 \\
&\geq \|x\|^2 - 2|\langle x, y\rangle| + \|y\|^2 \\
&\geq \|x\|^2 - 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2
\end{aligned}
\tag{14}
$$

Question #24

Let $C([a,b];\mathbb{F})$ be the vector space of all continuous functions from $[a,b] \subset$ to $\mathbb{F}$. Prove the following are valid norms in this space:

Claim: $\|f\|_{L^1} = \int_a^b |f(t)|dt$ **is a valid norm.**
Proof: We seek to verify the three properties of a valid norm:

1. Positivity: As $|f(t)| \geq 0$, it also has a non-negative integral over $[a,b]$, with $\|f(t)\| = 0 \iff f(t) = 0$ as an integral over non-negative functional support can only be zero if there is no area to integrate, or that $f(t) = 0$.

2. Scale Preservation: $\|kf(t)\| = \int_a^b |kf(t)|dt = |k| \int_a^b |f(t)|dt = |k|\|f(t)\|$

3. Triangle Inequality: $\|f(t) + g(t)\| = \int_a^b |f(t) + g(t)|dt \leq \int_a^b |f(t)| + |g(t)|dt = \int_a^b |f(t)|dt + \int_a^b |g(t)|dt = \|f(t)\| + \|g(t)\|$

Claim: $\|f\|_{L^2} = \left(\int_a^b |f(t)|^2 dt\right)^{\frac{1}{2}}$ **is a valid norm.**
Proof: We seek to verify the three properties of a valid norm:

1. Positivity: As $|f(t)|^2 \geq 0$, it also has a non-negative integral over $[a,b]$, with $\|f(t)\| = 0 \iff f(t) = 0$ as an integral over non-negative functional support can only be zero if there is no area to integrate, or that $f(t) = 0$ These properties are preserved, the square root is a continuous function over the non-negatives.

2. Scale Preservation:

$$
\begin{aligned}
\|kf(t)\| = \left(\int_a^b |kf(t)|^2 dt\right)^{\frac{1}{2}} &= \left(\int_a^b |k|^2 |f(t)|^2 dt\right)^{\frac{1}{2}} \\
= \left(|k|^2 \int_a^b |f(t)|^2 dt\right)^{\frac{1}{2}} &= (|k|^2)^{\frac{1}{2}} \left(\int_a^b |f(t)|^2 dt\right)^{\frac{1}{2}} \\
= |k| \left(\int_a^b |f(t)|^2 dt\right)^{\frac{1}{2}} &= |k|\|f(t)\|
\end{aligned}
\tag{15}
$$

3. Triangle Inequality:

$$\begin{aligned} \|f(t)+g(t)\|^2 &= \int_a^b \big[|f(t)+g(t)|\big]^2 dt \leq \int_a^b \big[|f(t)|+|g(t)|\big]^2 dt \\ &= \int_a^b |f(t)|^2 + 2|f(t)||g(t)| + |g(t)|^2 dt \\ &= \int_a^b |f(t)|^2 dt + 2\int_a^b |f(t)||g(t)|dt + \int_a^b |g(t)|^2 dt \\ &= \big[\|f(t)\| + \|g(t)\|\big]^2 \end{aligned} \tag{16}$$

Claim: $\|f\|_{L^\infty} = sup_{t\in[a,b]}|f(t)|$ **is a valid norm.**
Proof: We seek to verify the three properties of a valid norm:

1. Positivity: We observe that the sup norm takes the largest possible absolute value of a given function of a closed interval. This is strictly positive, and only can equal 0 if $\exists t$ such that $f(t) = 0$, as otherwise it could not be the sup.

2. Scale: $sup_{t\in[a,b]}|kf(t)| = sup_{t\in[a,b]}|k||f(t)| = |k|sup_{t\in[a,b]}|f(t)|$

3. Triangle Inequality:

$$\begin{aligned} \|f(t)+g(t)\| &= sup_{t\in[a,b]}|f(t)+g(t)| \\ &\leq sup_{t\in[a,b]}\big[|f(t)|+|g(t)|\big] \\ &= sup_{t\in[a,b]}|f(t)| + sup_{t\in[a,b]}|g(t)| \\ &= \|f(t)\| + \|g(t)\| \end{aligned} \tag{17}$$

Question #26

Two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are *topologically equivalent* if there exist constants $0 < m \leq M$ such that:

$$m\|x\|_a \leq \|x\|_a \leq M\|x\|_a, \quad \forall x \in X \tag{18}$$

Claim: Topological equivalence is an equivalence relation.
Proof: A relation is an equivalence relation $\iff$ it is reflexive, symmetric and transitive. Let us redefine topological equivalence to be $\forall\|\cdot\|$ attributable to $\mathscr{C}$, we have

1. Reflexive: $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a$, where $m = M = 1$

2. Symmetric: $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a \iff M\|x\|_a \geq \|x\|_b \geq m\|x\|_a$

3. Transitive: $m\|x\|_a \leq \|x\|_b \leq M\|x\|_a$ and $m\|x\|_c \leq \|x\|_a \leq M\|x\|_c$ imply $m\|x\|_c \leq \|x\|_b \leq M\|x\|_c$.

Claim: The following following inequalities are valid:

$$\begin{aligned} m\|x\|_2 \leq \|x\|_1 &\leq \sqrt{n}\ \|x\|_2 \\ m\|x\|_\infty\ \leq \|x\|_2 &\leq \sqrt{n}\ \|x\|_\infty \end{aligned} \tag{19}$$

Proof: For these inequalities to be valid, $\|\cdot\|_1,\ \|\cdot\|_2,$ and $\|\cdot\|_\infty$ must be topologically equivalent. For the first:

$$m\Big[\sum_{i=1}^n |x_i|^2\Big]^{\frac{1}{2}} \leq \sum_{i=1}^n |x_i| \leq M\Big[\sum_{i=1}^n |x_i|^2\Big]^{\frac{1}{2}} \tag{20}$$

for $m = 1$ and some $M = \sqrt{n}$ by the generalized Pythagorean Theorem, and that the square of the sum of all values cannot be more than n copies of the sum of the squared values. For the second,

$$\begin{aligned} m sup\{\{|x_i|\}_{i=1}^{n}\} &\leq \Big[\sum_{i=1}^{n} |x_i|^2\Big]^{\frac{1}{2}} \leq M sup\{\{|x_i|\}_{i=1}^{n}\} \\ m sup\{\{|x_i|\}_{i=1}^{n}\}^2 &\leq \sum_{i=1}^{n} |x_i|^2 \leq M^2 sup\{\{|x_i|\}_{i=1}^{n}\}^2 \end{aligned} \tag{21}$$

for $m = 1$ and some $M = \sqrt{n}$ by noticing that $sup\{\{|x_i|\}_{i=1}^{n}\}^2$ is contained within $\sum_{i=1}^{n} |x_i|^2$, and that the middle term cannot be more than n copies of the supremum.

Question #28

Let A be an n x n matrix.

Claim: Operator p-norms are topologically equivalent for $p = 1, 2, \infty$. Specifically:

$$\begin{aligned} \frac{1}{\sqrt{n}}\|A\|_2 \ &\leq \|A\|_1 \leq \sqrt{n}\|A\|_2 \\ \frac{1}{\sqrt{n}}\|A\|_\infty &\leq \|A\|_2 \leq \sqrt{n}\|A\|_\infty \end{aligned} \tag{22}$$

Proof: For the first case:

$$\begin{aligned} &\frac{1}{\sqrt{n}}\|A\|_2 \leq \|A\|_1 \leq \sqrt{n}\|A\|_2 \\ &\frac{\sigma_{max}(A)}{\sqrt{n}} \ \leq sup_i \sum_{i=1}^{m} |a_{ij}| \ \leq \sqrt{n}\sqrt{\Big[sup_i \sum_{i=1}^{m} |a_{ij}|\Big]\Big[sup_i \sum_{i=1}^{m} |a_{ij}\Big]} \end{aligned} \tag{23}$$

Notice that $\sigma_{max}(A) = \sqrt{\lambda_{max} A^H A}$, where $A^H A$ is a positive semi-definite matrix. Hence, it is bounded by the maximum column sum, which is precisely the definition of the matrix 1-norm. This argument follows for the second inequality, however after much deliberation with myself I am still confused on how to put it in notation.

Question #29

Take $\mathbb{F}$ with the 2-norm, and let the norm on $M_n(\mathbb{F})$ be the corresponding induced norm. Further, for any $x \in \mathbb{F}^n$, let $R_x \ : \ M_n(\mathbb{F}) \to \mathbb{F}^n$ be the linear transformation $A \mapsto Ax$

Claim: Any orthonormal matrix $Q \in M_n(\mathbb{F})$ has $\|Q\| = 1$.
Proof: Let $\|A\|$ be the induced 2-norm on a matrix A, defined as $\|A\| = sup\frac{\|Ax\|_2}{\|x\|_2}$, where $\|\cdot\|_2$ is the vector 2-norm. This is equal to $sup\sqrt{\lambda_i}$, where λ_i are the eigenvalues of $A^H A$. However, given that Q is by hypothesis orthonormal, and hence $Q^H Q = I$, the eigenvalues of the identity matrix are all 1. Thus, we have that $\|Q\| = 1$.

Claim: The induced norm of the transformation R_x is equal to $\|x\|_2$.
Proof: We have the following:

$$\|R_x\|_{A,Ax} = sup_A \frac{\|Ax\|_2}{\|A\|_{2,A}} = sup\frac{\|Ax\|_2}{\sigma_{max}(A)} \tag{24}$$

From this, we see that if A is an orthonormal matrix, $\sigma_{max}(A) = 1$, and thus $sup\frac{\|Ax\|_2}{\sigma_{max}(A)} = sup(\|Ax\|_2)$. As the function is bounded and x is the transformer of the function, we finally have $sup(\|Ax\|_2) = \|x\|_2$.

Question #30

Let $S \in M_n(\mathbb{F})$ be an invertible matrix. Given any matrix norm $\|\cdot\|$ on M_n, define $\|\cdot\|$ by $\|A\|_S = \|SAS^{-1}\|$.

Claim: $\|\cdot\|_S$ is a matrix norm on M_n.
Proof: We seek to show that $\|\cdot\|_S$ demonstrates the submultiplicative property across its elements, namely that $\|SAS^{-1}\| \leq \|S\| \cdot \|A\| \cdot \|S^{-1}\|$. This is as follows:

$$\|SA\| = sup\frac{\|SAx\|}{\|x\|} = sup\frac{\|SAx\|}{\|Ax\|}\frac{\|Ax\|}{\|x\|} \leq sup\frac{\|Sy\|}{\|y\|} \cdot sup\frac{\|Ax\|}{\|x\|} = \|S\| \cdot \|A\|$$

Let $Y = SA$. As this has demonstrated above is a matrix norm, we have the following: (25)

$$\|YS^{-1}\| = sup\frac{\|YS^{-1}x\|}{\|x\|} = sup\frac{\|YS^{-1}x\|}{\|S^{-1}x\|}\frac{\|S^{-1}x\|}{\|x\|} \leq sup\frac{\|Yz\|}{\|z\|} \cdot sup\frac{\|S^{-1}x\|}{\|x\|} = \|Y\| \cdot \|S^{-1}\|$$

Question #3.37

Let $V = \mathbb{R}[x; 2]$ be the space of polynomials of degree at most two, which is a subspace of the inner product space $L^2([0, 1]; \mathbb{R})$. Let $L : V \to \mathbb{R}$ be the linear functional given by $L[0] = p'(1)$.

Question: Find the unique $q \in V$ such that $L[p] = \langle q, p\rangle$, as guaranteed by the Riesz representation.
Proof: For $p \in V$, we have that $p = ax^2 + bx + c$. Hence, we have that $L(p) = p'(1) = 2a(1) + b = 2a + b$. Using the definition of the L_2 inner product space:

$$\begin{aligned} L(p) &= 2a + b = \langle ax^2 + bx + c, q(x)\rangle = \int_0^1 [ax^2 + bx + c][q(x)]dx \\ q(x) &= \sum_{i=1}^{n} L(x_i)x_i \quad \text{ where } \{x\}_{i=1}^{n} \text{ is an orthonormal basis in } V \end{aligned} \tag{26}$$

We use the Gram-Schmidt process to compute a basis in $\mathbb{R}[x; 2]$. We begin with a linearly independent set $f_1(x) = 1,\ f_2(x) = x,\ f_3(x) = x^2$:

$$\begin{aligned} q_1(x) &= 1 \\ q_2(x) &= x - \frac{\int_0^1 x \cdot 1dx}{\int_0^1 1 \cdot 1dx}(1) = x - \frac{1}{2} \\ q_3(x) &= x^2 - \frac{\int_0^1 x^2 \cdot 1dx}{\int_0^1 1 \cdot 1dx}(1) - \frac{\int_0^1 x^2 \cdot \left(x - \frac{1}{2}\right)dx}{\int_0^1 \left(x - \frac{1}{2}\right)^2 dx}\left(x - \frac{1}{2}\right) \\ &= x^2 - \frac{1}{3} - \frac{\frac{1}{12}}{\frac{1}{12}}\left(x - \frac{1}{2}\right) = x^2 - x + \frac{1}{6} \end{aligned} \tag{27}$$

We then use this orthonormal basis to achieve $q(x)$:

$$q(x) = \sum_{i=1}^{3} L(q_i(x))q_i(x) = (0)(1) + (1)\left(x - \frac{1}{2}\right) + (1)\left(x^2 - x + \frac{1}{6}\right) = x^2 - \frac{1}{3} \tag{28}$$

Question #3.38

Let $V = \mathbb{F}[x;]$, which is a subspace of the inner product space $L^2([0,1],\mathbb{R})$. Let D be the derivative operator $D : V \to V$; that is, $D[p](x) = p'(x)$.

Question: Write the matrix representation of D with respect to the power basis $[1, x, x^2]$ of $V = \mathbb{F}[x;]$.
Proof: We seek to find a matrix D such that $D[p(x)] = p'(x)$, where $p(x) = ax^2 + bx + c$. We then have the following:

$$B = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} ax^2 \\ bx \\ c \end{bmatrix}, \quad D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \begin{bmatrix} \frac{2}{x} & 0 & 0 \\ 0 & \frac{1}{x} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We then have:

$$B^* = B^H = \begin{bmatrix} \frac{2}{x} & 0 & 0 \\ 0 & \frac{1}{x} & 0 \\ 0 & 0 & 0 \end{bmatrix} = B$$

Question #3.39

Let V and W be finite-dimensional inner product spaces.

Claim: if $S, T \in \mathscr{L}(V, W)$, then $(S+T)^* = S^* + T^*$ and $(\alpha T)^* = \overline{\alpha} T^*$, $\alpha \in \mathbb{F}$.
Proof: For some $w \in W$, $v \in V$, $\langle (S+T)^*, v\rangle_V = \langle w, S+T\rangle_W = \langle w, S\rangle_W + \langle w, T\rangle_W = \langle S^*, v\rangle_V + \langle T^*_v\rangle_V = \langle S^* + T^*, v\rangle_V$

If $S \in \mathscr{L}(V; W)$, **then** $(S^*)^* = S$
Proof: For some $w \in W$, $v \in V$, $\langle w, S\rangle_W == \langle S^*, v\rangle_V = \langle w, (S^*)^*\rangle_W$

Claim: If $S, T \in \mathscr{L}(V, W)$, then $(ST)^* = T^* S^*$
Proof: For some $w \in W$, $v \in V$, $\langle T^* S^*, v\rangle_V = \langle w, (T^* S^*)^*\rangle_W = \langle w, ST\rangle_W = \langle (ST)^*, v\rangle_V$

Claim: If $T \in \mathscr{L}(V, W)$ and T is invertible, then $(T^*)^{-1} = (T^{-1})^*$
Proof: For some $w \in W$, $v \in V$, $\langle T^*(T^{-1})^*, v\rangle_V = \langle (T^{-1}T)^*, v\rangle_V = \langle w, T^{-1}T\rangle_W = \langle w, I\rangle_W$, which implies $(T^{-1})^*$ is the inverse of T^*, hence showing $(T^{-1})^* = (T^*)^{-1}$.

Question #3.40

Let $M_n(\mathbb{F})$ be endowed with the Frobenius inner product. Any $A \in M_n(\mathbb{F})$ defines a linear operator on $M_n(\mathbb{F})$ by left multiplication: $B \mapsto AB$.

Claim: $A^* = A^H$.
Proof: $\langle (AB)^H, C\rangle = tr\left(((AB)^H)^H C\right) = tr\,(ABC) = tr\,(CAB) = \langle C^*, AB\rangle$.

Claim: For any $A_1, A_2, A_3 \in M_n(\mathbb{F})$ we have $\langle A_2, A_3 A_1\rangle = \langle A_2 A_1^*, A_3\rangle$.
Proof: $\langle A_2, A_3 A_1\rangle = tr(A_2^* A_3 A_1) = tr(A_2^* A_3 A_1) = tr(A_3 A_1 A_2^*) = \langle A_3^*, A_1 A_2^*\rangle = \langle (A_1 A_2^*)^*, A_3\rangle = \langle A_2 A_1^*, A_3\rangle$.

Claim: Let $A \in M_n(\mathbb{F})$. Define the linear operator $T_A : M_n(\mathbb{F}) \to M_n(\mathbb{F})$ by $T_A(X) = AX - XA$. Hence, $(T_A)^* = T_A$.
Proof:

$$\begin{aligned} \langle A^*X - XA^*, X\rangle &= \langle A^*X, X\rangle - \langle XA^*, X\rangle \\ &= tr((A^*X)^*X) - tr(XA^*)^*X \\ &= tr(X^*AX) - tr(AX^*X) \\ &= tr(X^*AX) - tr(X^*XA) \\ &= \langle X, AX\rangle - \langle X, XA\rangle \\ &= \langle X, AX - XA\rangle \\ &= \langle (AX - XA)^*, X\rangle \end{aligned} \tag{29}$$

Question #3.44

Let $A \in M_{mxm}(\mathbb{F})$ and $b \in \mathbb{F}^m$.

Claim: The Fredholm alternative, or that either $Ax = b$ has a solution $x \in \mathbb{F}^n$ or there exists $y \in \mathscr{N}(A^H)$ such that $\langle y, b\rangle \neq 0$, is valid.
Proof: We see that by necessity either $Ax = b$ has a solution $x \in \mathbb{F}$ or it does not. If it doesn't, we have that $b \notin \mathscr{R}(L)$, which is analogous to for some $y \in \mathscr{R}^\perp$, $\langle y, b\rangle \neq 0$. However, by the Fundamental Subspaces Theorem, $\mathscr{R}(L)^\perp = \mathscr{N}(L^*) = \mathscr{N}(L^H)$.

Question #3.45

Consider the vector space $M_n(\mathbb{R})$ with the Frobenius inner product.

Claim: $Sym_n(\mathbb{R})^\perp = Skew_n(\mathbb{R})$
Proof: We seek to demonstrate that as $Sym_n(\mathbb{R})^\perp$ and $Skew_n(\mathbb{R})$ have the exact same composition, they are they same spaces. Let $L : Sym_n(\mathbb{R})^\perp \to Skew_n(\mathbb{R})$. We then have the following:

$$\begin{aligned} \mathscr{N}(L)^\perp + \mathscr{R}(L^*)^\perp &= \mathscr{N}(L^*) + \mathscr{R}(L) \\ \mathscr{N}(L)^\perp + \mathscr{R}(L)^\perp &= \mathscr{N}(L^*) + \mathscr{R}(L^*) \\ \mathscr{N}(L)^\perp + \mathscr{N}(L^*) &= \mathscr{N}(L^*) + \mathscr{N}(L)^\perp \end{aligned} \tag{30}$$

Hence, we have the desired outcome.

Question #3.46

Prove the following for an mxn matrix A:

Claim: If $x \in \mathscr{N}(A^HA)$, then Ax is in both $\mathscr{R}(A)$ and $\mathscr{N}(A^H)$.
Proof: We have $A^HAx = 0$, and x exists. We seek to show x is identified in such a way that also $A^HAx = 0$ and $Ay = Ax$, for some y. The second equality follows trivially from the definition. For the first, notice that by using the Frobenius IP, we have $\langle Ax, Ax\rangle = tr(x^HA^HAx) = tr(x^H(A^Ax)) = tr(x^H(0)) = 0$. Hence, this $Ax \in \mathscr{N}(A^H)^\perp$, which is equal to $\mathscr{R}(A)$ by the Fundamental Subspaces Theorem.

Claim: $\mathscr{N}(A^H A) = \mathscr{N}(A^H)$.
Proof: We have x such that $A^H Ax = 0$. Let $Ax = b$. If $b = 0$, we have that the claim holds trivially. If $b \neq 0$, then we have $A^H Ax = A^H b = 0$. This is by definition $b \in \mathscr{N}(A^H)$. Hence, we have $\mathscr{N}(A^H A) = \mathscr{N}(A^H)$.

Claim: A and $A^H A$ have the same rank.
Proof: We assume $rank(A) \neq rank(A^H A)$. Thus, $rank(A^H A) \leq min\{rank(A^H), rank(A)\} \neq rank(A)$. This implies that $rank(A^H A) \neq rank(A)$, which forms a contradiction.

Claim: If A has linearly independent columns, then $A^H A$ is nonsingular.
Proof: This holds somewhat trivially, as by the previous problem $rank(A) = rank(A^H A)$, and thus if A is of full rank (equal to A having linearly independent columns), $A^H A$ is also full rank, has an inverse, and is thus non-singular.

Problem #3.48

Consider the vector space $M_n(\mathbb{R})$ with the Frobenius inner product. Let $P(A) = \frac{A+A^T}{2}$ be a valid mapping.

Claim: P is linear.
Proof: We must verify scalar multiplication and matrix addition:

1. Scalar Multiplication: $P(\alpha A) = \frac{\alpha A + \alpha A^T}{2} = \alpha \frac{A+A^T}{2} = \alpha P(A)$.
2. Matrix Addition: $P(A+B) = \frac{(A+B)+(A+B)^T}{2} = \frac{A+B+A^T+B^T}{2} = \frac{A+A^T}{2} + \frac{B+B^T}{2} = P(A) + P(B)$

Claim: $P^2 = P$
Proof: $P(A)^2 = \left(\frac{A+A^T}{2}\right)^2 = \left(\frac{AA+AA^T+A^TA+A^TA^T}{4}\right) = \left(\frac{A(A+A^T)+A^T(A+A^T)}{4}\right)$

Claim: $P^* = P$.
Proof: $\langle P^*, A\rangle = tr\left((P^*)^H A\right) = tr\left(PA\right) = tr\left(\frac{AA+A^TA}{2}\right) = tr\left(\frac{A^TA+AA}{2}\right) = tr(P^*A) = \langle P, A\rangle$

Claim: $\mathscr{N}(P) = Skew_n(\mathbb{R})$.
Proof:$\mathscr{N}(P) = \mathscr{N}(P^*) = \mathscr{R}(P)^\perp$. Hence, $0 = \langle P, Y\rangle = tr(P^H Y)$

Claim: $\mathscr{R}(P) = Sym_n(\mathbb{R})$.
Proof:

Claim: $\|A - P(A)\|_F = \sqrt{\frac{tr(A^TA)+tr(AA)}{2}}$

Proof: We have:

$$\begin{aligned}
\|A - P(A)\|_F &= \|A - \frac{A + A^T}{2}\| \\
&= \|\frac{2A - A - A^T}{2}\| \\
&= \|\frac{A - A^T}{2}\| \\
&= \sqrt{\langle \frac{A - A^T}{2}, \frac{A - A^T}{2} \rangle} \\
&= \sqrt{tr\left(\left(\frac{A - A^T}{2}\right)^T \left(\frac{A - A^T}{2}\right)\right)} \\
&= \sqrt{tr\left(\left(\frac{A - A^T}{2}\right) \left(\frac{A - A^T}{2}\right)\right)} \\
&= \sqrt{tr\left(\frac{AA - AA^T - A^TA + A^TA^T}{4}\right)} \\
&= \sqrt{\frac{tr(AA - AA^T) - tr(A^TA + A^TA^T)}{4}} \\
&= \sqrt{\frac{tr(A^TA - A^TA^T) - tr(AA + AA^T)}{4}} \\
&= \sqrt{\frac{tr(A^TA) - tr(A^TA^T) - tr(AA) + tr(AA^T)}{4}} \\
&= \sqrt{\frac{tr(A^TA) - tr(AA) - tr(AA) + tr(A^TA)}{4}} \\
&= \sqrt{\frac{tr(A^TA) - tr(A^2)}{2}}
\end{aligned} \tag{31}$$

Question #3.49

($\Rightarrow$) We have $\hat{x} \in \mathbb{F}^n$ is a least squares solution. From this, we see that $A(A^HA)^{-1}A^Hb + b - proj_{\mathscr{R}(A)}b = p + b - p = b$ and $A^H(b - proj_{\mathscr{R}(A)}b) = A^H(b - p) = 0$, with the last equality following from Proposition 3.9.1.

($\Leftarrow$) For the first inequality, we have $A\hat{x} + b - p = p$, hence requiring $\hat{x}$ to be exactly defined such that $A\hat{x} = p$, which is $\hat{x} = (A^HA)^{-1}A^Hb$ (shown by Theorem 3.9.3).

Question #3.50

Let $(x_i, y_i)_{i=1}^n$ be a collection of data points that we have reason to believe should lie (roughly) on an ellipse of the form $rx^2 + sy = 1$. We wish to find the least-squares approximation for r and s. Write A, x, and b for the corresponding normal equation in terms of the data x_i and y_i and the unknowns r and s.

Solution: We have:

$$A = \begin{bmatrix} x_1^2 & y_1 \\ x_2^2 & y_2 \\ \vdots & \vdots \\ x_n^2 & y_n \end{bmatrix}, x = \begin{bmatrix} r \\ s \end{bmatrix}, b = \begin{bmatrix} 1_1 \\ 1_2 \\ \vdots \\ 1_n \end{bmatrix}$$