Math Problem Set #2

Charlie Walker

3.1

(i)

$$\langle x, y\rangle = \|x\|\|y\|\cos\theta$$

Law of cosines:

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos(\pi-\theta)\\ \|x-y\|^2 &= \|x\|^2 - \|y\|^2 - 2\|x\|\|y\|\cos\theta\end{aligned}$$

Subtracting and substituting gives:

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

(ii)

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

3.2

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

3.3
(i)

$$\begin{aligned}
\langle x, x^5\rangle &= \int_0^1 x x^5 dx = \frac{1}{7} \\
\|x\|^2 = \langle x, x\rangle &= \int_0^1 x^2 dx = \frac{1}{3} \\
\|x^5\|^2 &= \int_0^1 x^{10} dx = \frac{1}{11} \\
\implies \cos\theta &= \frac{\frac{1}{7}}{\frac{1}{\sqrt{3}} \cdot \frac{1}{11}}
\end{aligned}$$

(ii)

$$\begin{aligned}
\langle x^2, x^4\rangle &= \int_0^1 x^2 x^4 dx = \frac{1}{7} \\
\|x^2\|^2 &= \int_0^1 x^4 dx = \frac{1}{5} \\
\|x^4\|^2 &= \int_0^1 x^8 dx = \frac{1}{9} \\
\implies \cos\theta &= \frac{\frac{1}{7}}{\frac{1}{5} \cdot \frac{1}{3}}
\end{aligned}$$

3.8
(i) For a set to be orthonormal, the inner product of any distinct pair of elements in the set must be equal to zero (orthogonality), and the norm of any element must be 1 (normality):

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

For norms,

$$\begin{aligned}\langle \cos(t), \cos(t) \rangle &= \|\cos(t)\|^2 \\ &= \|\sin(t)\|^2 \\ &= \|\cos(2t)\|^2 \\ &= \|\sin(2t)\|^2 \\ &= 1\end{aligned}$$

(ii)
(iii)
(iv)

3.9
A matrix R is an orthonormal if and only if it satisfies $R^H R = I$:

$$\begin{aligned}R(\theta) &= \begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix} \\ R^H(\theta) &= \begin{bmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{bmatrix} \\ R^H R(\theta) &= \begin{bmatrix}\cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \cos\theta\sin\theta \\ \cos\theta\sin\theta - \cos\theta\sin\theta & \cos^2\theta + \sin^2\theta\end{bmatrix} \\ &= \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} = I_2\end{aligned}$$

$\implies$ the rotation matrix in $\mathbb{R}^2$ is an orthonormal transformation.

3.10
(i) Orthonormal transformations preserve inner products, i.e. $\langle x, y \rangle = \langle Qx, Qy \rangle$, where $x, y \in \mathbb{F}^n$ and $Q \in M_n(\mathbb{F})$. Thus, $\langle x, y \rangle = \langle Qx, Qy \rangle = \langle Q^H Qx, Qy \rangle \implies x = Q^H Qx \implies Q^H Q = I$.

(ii) $\|Qx\|^2 = \langle Qx, Qx \rangle = \langle Q^H Qx, Qx \rangle = \langle x, x \rangle = \|x\|^2$. Since $\|\cdot\|$ is strictly positive $\forall x \neq 0$, we take square roots to complete the desired result.

(iii) Q orthonormal $\implies Q^H Q = I \implies Q^H = Q^{-1}$. It is thus equivalent to show that Q^H is orthonormal: $Q^H Q = I = Q^{HH} Q^H = Q^H Q^{HH} \implies Q^H = Q^{-1}$ is orthonormal.

(iv) For orthonormal matrix $Q \in M_n(\mathbb{F}), QQ^H = I$, the matrix with ones along the diagonal and zeros everywhere else. Thus, from the rules of matrix multiplication, $q_i q_j^T = 1, \forall i = j$ columns in Q. If $i \neq j$, the product is equal to zero. In terms of inner products, we have $\langle q_i, q_i \rangle = \|q_1\|^2 = 1$, and $\langle q_i, q_j \rangle = 0, i \neq j \implies$ every column in Q is orthonormal.

(v) $1 = \det(I_n) = \det(Q^H Q) = \det(Q^H)\det(Q) = \det(Q)\det(Q) = \det(Q)^2 \implies |\det(Q)| = 1$. The converse is not always true.

3.11

Suppose $v_1, \ldots, v_n$ is a linearly independent list of vectors in V. Applying the Gram-Schmidt procedure gives an orthonormal list of vectors $e_1, \ldots, e_n$ such that $span(e_1, \ldots, e_n) = span(v_1, \ldots, v_n)$. Take some vector $v_{n+1} \in span(v_1, \ldots, v_n)$ and apply the Gram-Schmidt procedure:

$$e_{n+1} = \frac{v_{n+1} - \langle v_{n+1}, e_1\rangle e_1 - \ldots - \langle v_{n+1}, e_n\rangle e_n}{\|v_{n+1} - \langle v_{n+1}, e_1\rangle e_1 - \ldots - \langle v_{n+1}, e_n\rangle e_n\|}$$

Because any $v \in V$ can be written as $v = \langle v, e_1\rangle e_1 + \ldots + \langle v, e_n\rangle e_n$, the above expression is undefined (numerator and denominator both reduce to 0).

3.16

(i) Any matrix $A \in M_{m\times n}$ with $m \geq n$ can be reduced to the product of an $m \times m$ orthonormal matrix Q and an $m \times n$ upper triangular matrix R. Since the bottom $(m-n)$ rows of any $m \times n$ matrix are entirely zeros, R and Q can be partitioned:

$$A = QR = Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix} = \begin{bmatrix} Q_1, Q_2 \end{bmatrix}\begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1$$

where R_1 is an $n \times n$ upper triangular matrix, 0 is an $(m-n) \times n$ zero matrix, Q_1 is $m \times n, Q_2$ is $m \times (m-n)$ and Q_1 and Q_2 are both orthonormal. For some $Q_2' \neq Q_2$,

$$\begin{bmatrix} Q_1, Q_2 \end{bmatrix}\begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1 = \begin{bmatrix} Q_1, Q_2' \end{bmatrix}\begin{bmatrix} R_1 \\ 0 \end{bmatrix} = A$$

(ii) Suppose that $A = Q_1 R_1 = Q_2 R_2$ where Q_1, Q_2 are orthonormal and R_1, R_2 are upper triangular with positive diagonal entries. Then,

$$M := R_1 R_2^{-1} = Q_1^H Q_2.$$

Since M is orthonormal and upper triangular, it must be diagonal. Further, the diagonal entries of M are positive, because the upper triangular matrices R_1 and R_2^-1 have positive diagonal entries, and of modulus one, because M is a diagonal orthonormal matrix. Thus, $M = I$, implying $R_1 = R_2$ and $Q_1 = Q_2$.

3.16

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

3.17

$$
\begin{aligned}
A^H A x &= A^H b \\
(\widehat{Q}\widehat{R})^H \widehat{Q}\widehat{R} x &= \widehat{R}^H \widehat{Q}^H b \\
\widehat{R}^H \widehat{Q}^H \widehat{Q}\widehat{R} x &= \widehat{R}^H \widehat{Q}^H b \\
\widehat{R}^H \widehat{R} x &= \widehat{R}^H \widehat{Q}^H b \\
\widehat{R} x &= (\widehat{R}^H)^{-1} \widehat{R}^H \widehat{Q}^H b \\
\widehat{R} x &= \widehat{Q}^H b
\end{aligned}
$$

As desired.

3.23

This result follows directly from the triangle inequality because $\|x\| = \|x-y+y\| \leq \|x-y\| + \|y\| \implies \|x\| - \|y\| \leq \|x-y\|$ for all $x, y \in V$.

3.24

(i) Positivity is trivial. Scale preservation: for some scalar $x \in \mathbb{R}$,

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

Triangle inequality:

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

(ii) Positivity is once again trivial. For some scalar $x \in \mathbb{R}$,

$$\begin{aligned}
\|xf\|_{L^2} &= (\int_a^b |x|^2 |f(t)|^2 dt)^{1/2} \\
&= |x|(\int_a^b |f(t)|^2 dt)^{1/2} \\
&= |x| \|f\|_{L^2}
\end{aligned}$$

Triangle inequality:

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

(iii) Positivity is trivial. Scale preservation:

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

Triangle inequality:

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

3.26

(i) $\|x\|_2 \leq \|x\|_1$ follows from the triangle inequality. To show $\|x\|_1 \leq \sqrt{n}$, let y be a vector of all ones, where the sign is equal to the sign of x. Then $\|x\|_1 = |\langle x, y \rangle| \leq \|y\|_2 \|x\|_2 = \sqrt{n} \|x\|_2$.

(ii) $\|x\|_2^2 = \sum_{i=1}^m x_i^2 \leq m \cdot \sup_{1 \leq m} x_i^2 = m(sup_{1\ \leq m})^2 = m\|x\|_\infty^2 \implies \|x\|_2 \leq \sqrt{m} \|x\|_\infty$

3.28

(i) $\|A\|_2 \leq \|A\|_F = \sqrt{tr(A^H A)}$, which can be interpreted as the square root of the sum of all squared entries in the A. Thus, $\frac{\|A\|_2^2}{n}$ is the average squared column sum, and $\frac{\|A\|_2}{\sqrt{n}}$ is the average column sum. Since $\|A\|_1$ is the maximum column sum in A, $\frac{\|A\|_2}{\sqrt{n}}$ is clearly less than or equal to $\|A\|_1$.

3.29

Orthonormal transformations preserve inner products and thus preserve norms, i.e. $\|Qx\| = \|x\| \implies \|Q\| = 1$. For $R_x : M_n(\mathbb{F}) \to \mathbb{F}^n$,

$$\begin{aligned}
\|R_x\| &= \sup \frac{\|R_x A\|}{\|A\|} \\
&= \sup \frac{\|Ax\|}{\|A\|} \\
&= \sup \frac{\|Ax\|}{\sup \frac{\|Ax\|}{\|x\|_2}} \\
&\leq \|x\|_2
\end{aligned}$$

By Gram-Schmidt, any vector x with norm $\|x\|_2 = 1$ is part of an orthonormal basis, and hence is the first column of an orthonormal matrix.

3.37

$$\begin{aligned}
q &= \sum_{i=1}^{n} L(x_1)e_1 \\
&= L(1)\begin{bmatrix}1\\0\\0\end{bmatrix} + L(x)\begin{bmatrix}0\\1\\0\end{bmatrix} + L(x^2)\begin{bmatrix}0\\0\\1\end{bmatrix} \\
&= 0 \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} + 1 \cdot \begin{bmatrix}0\\1\\0\end{bmatrix} + 2 \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} \\
&= \begin{bmatrix}0\\0\\0\end{bmatrix} + \begin{bmatrix}0\\1\\0\end{bmatrix} + \begin{bmatrix}0\\0\\2\end{bmatrix} \\
&= \begin{bmatrix}0\\1\\2\end{bmatrix}
\end{aligned}$$

3.38

The matrix of the derivative operator with respect to the power basis is:

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

The adjoint of D is its Hermitian conjugate:

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

3.44

If $Ax = b$ has a solution, then for any vector y we have $\langle y, Ax \rangle = \langle y, b \rangle$. Equivalently, $\langle y, Ax \rangle = \langle A^H y, x \rangle$. It follows that for any vector y we must have $\langle A^* y, x \rangle = \langle y, b \rangle$. If y satisfies $A^* y = 0$, we must also have $\langle y, b \rangle = 0$.