Problem Set 2

Natasha Watkins

Exercise 3.1

Part i)

We can write

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2 \tag{1}$$

as $\langle \mathbf{y}, \mathbf{x}\rangle = \overline{\langle \mathbf{x}, \mathbf{y}\rangle} = \langle \mathbf{x}, \mathbf{y}\rangle$ in $\mathbb{R}^n$.

Using $\langle \mathbf{x}, -\mathbf{y}\rangle = (-1)\langle \mathbf{x}, \mathbf{y}\rangle$, we can write

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

Combining,

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

Part ii)

Combining equations 1 and 2, we find

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

Exercise 3.2

$$
\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 + i\|\mathbf{x}-i\mathbf{y}\|^2 - i\|\mathbf{x}+i\mathbf{y}\|^2) \\
&= \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) + \frac{1}{4}(i\|\mathbf{x}-i\mathbf{y}\|^2 - i\|\mathbf{x}+i\mathbf{y}\|^2) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i\|\mathbf{x}-i\mathbf{y}\|^2 - i\|\mathbf{x}+i\mathbf{y}\|^2) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i\langle \mathbf{x}-i\mathbf{y}, \mathbf{x}-i\mathbf{y} \rangle - i\langle \mathbf{x}+i\mathbf{y}, \mathbf{x}+i\mathbf{y} \rangle) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i\langle \mathbf{x}, \mathbf{x}-i\mathbf{y} \rangle + i^2\langle \mathbf{y}, \mathbf{x}-i\mathbf{y} \rangle - i\langle \mathbf{x}, \mathbf{x}+i\mathbf{y} \rangle + (-i)^2\langle \mathbf{y}, \mathbf{x}+i\mathbf{y} \rangle) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{4}(i\langle \mathbf{x}, \mathbf{x} \rangle + i^2\langle \mathbf{x}, \mathbf{y} \rangle + i^2\langle \mathbf{y}, \mathbf{x} \rangle + i^3\langle \mathbf{y}, \mathbf{y} \rangle \\
&\qquad - i\langle \mathbf{x}, \mathbf{x} \rangle + (-i)^2\langle \mathbf{x}, \mathbf{y} \rangle + (-i)^2\langle \mathbf{y}, \mathbf{x} \rangle + (-i^3)\langle \mathbf{y}, \mathbf{y} \rangle) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle
\end{aligned}
$$

Exercise 3.3

Part i)

$\langle \mathbf{x}, \mathbf{x}^5 \rangle = \int_0^1 x^6 dx = \frac{x^7}{7}\Big|_0^1 = \frac{1}{7}$
$\langle \mathbf{x}, \mathbf{x} \rangle = \int_0^1 x^2 dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3} \implies \|\mathbf{x}\| = \sqrt{\frac{1}{3}}$
$\langle \mathbf{x}^5, \mathbf{x}^5 \rangle = \int_0^1 x^{10} dx = \frac{x^{11}}{11}\Big|_0^1 = \frac{1}{11} \implies \|\mathbf{x}^5\| = \sqrt{\frac{1}{11}}$

$\cos(\theta) = \frac{\sqrt{3}\sqrt{11}}{7} = \frac{\sqrt{33}}{7} \implies \theta \approx 35°$

Part ii)

$\langle \mathbf{x}^2, \mathbf{x}^4 \rangle = \int_0^1 x^6 dx = \frac{x^7}{7}\Big|_0^1 = \frac{1}{7}$
$\langle \mathbf{x}^2, \mathbf{x}^2 \rangle = \int_0^1 x^4 dx = \frac{x^5}{5}\Big|_0^1 = \frac{1}{5} \implies \|\mathbf{x}^2\| = \sqrt{\frac{1}{5}}$
$\langle \mathbf{x}^4, \mathbf{x}^4 \rangle = \int_0^1 x^8 dx = \frac{x^9}{9}\Big|_0^1 = \frac{1}{9} \implies \|\mathbf{x}^4\| = \sqrt{\frac{1}{9}}$

$\cos(\theta) = \frac{\sqrt{9}\sqrt{5}}{7} = \frac{\sqrt{45}}{7} \implies \theta \approx 17°$

Exercise 3.8

i)

$\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\sin(t)dt = 0$
$\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\cos(2t)dt = 0$
$\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\sin(2t)dt = 0$
$\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(2t)\sin(t)dt = 0$
$\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\cos(t)dt = 1$
$\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(t)\sin(t)dt = 1$
$\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(2t)\sin(2t)dt = 1$
Therefore, S is an orthonormal set.

ii) $\|t\| = \sqrt{\langle t,t\rangle} = \int_{-\pi}^{\pi} t^2 dt = \frac{t^3}{3}\Big|_{-\pi}^{\pi} = \frac{2}{3}\pi^2$

iii) $\text{proj}_X(\cos(3t)) = \sum_{i=1}^{m}\langle \mathbf{x}_i, \cos(3t)\rangle \mathbf{x_i} = 0$

iv) $\text{proj}_X(t) = \sum_{i=1}^{m}\langle \mathbf{x}_i, t\rangle \mathbf{x_i} = 1$

Exercise 3.9

By Theorem 3.2.15, a matrix Q is orthonormal if and only if $Q^H Q = QQ^H = 1$.

The rotation matrix is given by

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Calculating $R_\theta R$, we find

$$\begin{aligned} R_\theta R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta - \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathrm{I} \end{aligned}$$

So R_θ is an orthonormal transformation.

Exercise 3.10

<u>Part i)</u>

Assume Q is orthonormal, which implies $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \langle Q\mathbf{e}_i, Q\mathbf{e}_j \rangle$.

$$\begin{aligned}
\mathbf{e}_i^H \mathbf{e}_j &= \langle \mathbf{e}_i, \mathbf{e}_j \rangle \\
&= \langle Q\mathbf{e}_i, Q\mathbf{e}_j \rangle \\
&= (Q\mathbf{e}_i)^H (Q\mathbf{e}_j) \\
&= \mathbf{e}_i^H Q^H Q \mathbf{e}_j
\end{aligned}$$

$\mathbf{e}_i^H \mathbf{e}_j = \mathbf{e}_i^H Q^H Q \mathbf{e}_j$ only if $Q^H Q = I$.

Part ii)

$$\begin{aligned}
\|Q\mathbf{x}\| &= \sqrt{\langle Q\mathbf{x}, Q\mathbf{x} \rangle} \\
&= \sqrt{\mathbf{x}^H Q^H Q \mathbf{x}} \\
&= \sqrt{\mathbf{x}^H \mathbf{x}} \qquad \text{as } Q^H Q = 1 \\
&= \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \\
&= \|\mathbf{x}\|
\end{aligned}$$

Part iii)

$Q^H Q = I$ so $Q^H = Q^{-1}$.

$$\begin{aligned}
\langle Q^{-1}\mathbf{x}, Q^{-1}\mathbf{x} \rangle &= \mathbf{x}^H (Q^{-1})^H Q^{-1} \mathbf{x} \\
&= \mathbf{x}^H (Q^H)^H Q^H \mathbf{x} \qquad \text{as } Q^H = Q^{-1} \\
&= \mathbf{x}^H Q Q^H \mathbf{x} \\
&= \mathbf{x}^H \mathbf{x} \qquad \text{as } QQ^H = 1 \\
&= \langle \mathbf{x}, \mathbf{x} \rangle
\end{aligned}$$

So Q^{-1} is orthonormal.

Part iv)

$Q_{n,m}$ is an orthonormal matrix. We can write Q as

$$Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \end{bmatrix}$$

Where $\mathbf{q}_i$ is a $(n \times 1)$ column of Q, and $\{\mathbf{q}_i\}_{i=1}^m$ is a set of columns vectors.

As Q is orthonormal, we know $Q^H Q = I_m$. So we can write

$$\begin{aligned}
Q^H Q &= \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \end{bmatrix}^H \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{q}_1^H \\ \mathbf{q}_2^H \\ \vdots \\ \mathbf{q}_m^H \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{q}_1^H \mathbf{q}_1 & \mathbf{q}_2^H \mathbf{q}_1 & \cdots & \mathbf{q}_m^H \mathbf{q}_1 \\ \mathbf{q}_1^H \mathbf{q}_2 & \mathbf{q}_2^H \mathbf{q}_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \\ \mathbf{q}_1^H \mathbf{q}_2 & \cdots & & \mathbf{q}_m^H \mathbf{q}_m \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \\ 0 & \cdots & & 1 \end{bmatrix}
\end{aligned}$$

which satisfies the definition of an orthonormal set, ie. $\langle \mathbf{q}_i, \mathbf{q}_j \rangle = \delta_{i,j}$ where

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Part v)

It is not true that $|\det(Q)| = 1$ implies Q is orthonormal. For example

$$Q = \begin{bmatrix} 2 & \sqrt{3} \\ \sqrt{3} & 2 \end{bmatrix}$$

has $|\det(Q)| = 1$ but

$$Q^H Q = \begin{bmatrix} 2 & \sqrt{3} \\ \sqrt{3} & 2 \end{bmatrix} \begin{bmatrix} 2 & \sqrt{3} \\ \sqrt{3} & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2\sqrt{3} - 3\sqrt{2} \\ 3\sqrt{2} - 2\sqrt{3} & -1 \end{bmatrix} \neq I$$

So Q is not orthonormal.

Part vi)

Q_1 and Q_2 are orthonormal, so $Q_1^H Q_1 = I$ and $Q_2^H Q_2 = I$ Given $Q = Q_1 Q_2$

$$\begin{aligned} Q^H Q &= (Q_1 Q_2)^H Q_1 Q_2 \\ &= Q_2^H Q_1^H Q_1 Q_2 \\ &= Q_2^H Q_2 \\ &= I \end{aligned}$$

Q is therefore orthonormal.

Exercise 3.11

If the Gram–Schmidt orthonormalisation process is applied to a collection of linearly dependent vectors, it outputs the $\mathbf{0}$ vector for some $\mathbf{x}_k - \mathbf{p}_{k-1}$, and therefore $\mathbf{q}_k$ is undefined.

Exercise 3.16

Part i)

Consider a diagonal matrix D with elements on the diagonal equal to -1. We can then write $Q^* = QD$ and $R^* = D^{-1}R$, and deduce that $A = Q^*R^* = QDD^{-1}R = QR$, showing that the QR decomposition is not unique.

Part ii)

Suppose $A = Q_1 R_1$ and $A = Q_2 R_2$, , where R_1, R_2 are upper triangular matrices with positive elements along the diagonal.

Then $Q_1 R_1 = Q_2 R_2 \implies Q_2^{-1} Q_1 = R_1^{-1} R_2$, which means that $R_1^{-1} R_2$ is orthonormal. As $R_1^{-1} R_2$ is an upper triangular matrix that is orthonormal, its diagonal elements must be equal to 1 (as the diagonal of R is positive). So $Q_2^{-1} Q_1 = I$ and $R_1^{-1} R_2 = I$, and thus $Q_1 = Q_2$ and $R_1 = R_2$, showing that the decomposition is unique in this case.

Exercise 3.17

$$
\begin{aligned}
A^H A\mathbf{x} &= (\hat{Q}\hat{R})^H \hat{Q}\hat{R}\mathbf{x} \\
&= \hat{R}^H \hat{Q}^H \hat{Q}\hat{R}\mathbf{x} \\
&= \hat{R}^H \hat{R}\mathbf{x} \qquad \text{as } \hat{Q}^H\hat{Q} = 1 \\
\hat{R}^H \hat{R}\mathbf{x} &= \hat{R}^H \hat{Q}^H \mathbf{b} \qquad \text{as } A^H = \hat{R}^H\hat{Q}^H \\
\hat{R}\mathbf{x} &= \hat{Q}^H \mathbf{b} \qquad \text{as required}
\end{aligned}
$$

Exercise 3.23

By the triangle inequality,

$$\|\mathbf{x}\| = \|(\mathbf{x}-\mathbf{y}) + \mathbf{y}\| \leq \|\mathbf{x}-\mathbf{y}\| + \|\mathbf{y}\| \Leftrightarrow \|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x}-\mathbf{y}\|$$

By scale preservation,

$$\|\mathbf{y}\| = \|-1\| \cdot \|\mathbf{y}\| = \|-\mathbf{y}\| = \|\mathbf{x}-\mathbf{y}+\mathbf{x}\| \leq \|\mathbf{x}-\mathbf{y}\| + \|\mathbf{y}\| \Leftrightarrow \|\mathbf{y}\| - \|\mathbf{x}\| \leq \|\mathbf{x}-\mathbf{y}\|$$

Exercise 3.24

Part i)

1. **Positivity:**

   Consider $f, g \in C[a,b]$, with $f(t) \geq g(t)$ for all $t \in [a,b]$, then

   $$\int_a^b f(t)dt \geq \int_a^b g(t)dt$$

   With $g(t) = 0$, and $|f(t)| \geq 0$

   $$\int_a^b f(t)dt \geq \int_a^b 0dt = 0$$

2. **Scale preservation:**

Given $h \in \mathbb{F}$ and $f \in C[a,b]$,

$$\|hf\| = \int_a^b |hf(t)|dt = \int_a^b |h||f(t)|dt = |h| \int_a^b |f(t)|dt = |h|\|f\|$$

3. **Triangle inequality:**

   Consider $f, g \in C[a,b]$.

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

Part ii)

Results follow from part i).

Part iii)

1. **Positivity:**

   $|f(x)| \geq 0$ for all x, so it follows that $\sup_{x\in[a,b]} |f(x)| \geq 0$.

2. **Scale preservation:**

$$\|hf\| = \sup_{x\in[a,b]} |h||f(x)| = |h| \sup_{x\in[a,b]} |f(x)| = |h|\|f\|$$

3. **Triangle inequality:**

   For any $x \in [a,b]$

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

As this holds for all $x \in [a,b]$,

$$\sup_{x\in[a,b]} |f(x)+g(x)| = \|f+g\| \leq \|f\| + \|g\|$$

Exercise 3.26

$\underline{\|\mathbf{x}\|_a \sim \|\mathbf{x}\|_a}$

$$\begin{aligned} m\|\mathbf{x}\|_a &\leq \|\mathbf{x}\|_a \leq M\|\mathbf{x}\|_a \\ \implies m &\leq 1 \leq M \end{aligned}$$

with $m \in [0,1]$, and $M \in [1,\infty]$.

$\underline{\|\mathbf{x}\|_a \sim \|\mathbf{x}\|_b \implies \|\mathbf{x}\|_b \sim \|\mathbf{x}\|_a}$

$$\begin{aligned} &m\|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq M\|\mathbf{x}\|_a \\ &\implies \frac{1}{m} \leq \frac{\|\mathbf{x}\|_b}{\|\mathbf{x}\|_a} \leq \frac{1}{M} \\ &\implies \frac{1}{M} \geq \frac{\|\mathbf{x}\|_a}{\|\mathbf{x}\|_b} \geq \frac{1}{m} \\ &\implies Y \geq \frac{\|\mathbf{x}\|_a}{\|\mathbf{x}\|_b} \geq y \\ &\implies y\|\mathbf{x}\|_b \leq \|\mathbf{x}\|_a \leq Y\|\mathbf{x}\|_b \end{aligned}$$

where $0 \leq y \leq Y$.

$\underline{\|\mathbf{x}\|_a \sim \|\mathbf{x}\|_b \text{ and } \|\mathbf{x}\|_b \sim \|\mathbf{x}\|_c \implies \|\mathbf{x}\|_a \sim \|\mathbf{x}\|_c}$

$$m\|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq M\|\mathbf{x}\|_a \tag{3}$$

$$w\|\mathbf{x}\|_b \leq \|\mathbf{x}\|_c \leq W\|\mathbf{x}\|_b \tag{4}$$

Multiply (3) by w, to get

$$mw\|\mathbf{x}\|_a \leq w\|\mathbf{x}\|_b \leq wM\|\mathbf{x}\|_a \tag{5}$$

Combining (5) with (4), we see

$$mw\|\mathbf{x}\|_a \leq \|\mathbf{x}\|_c \leq W\|\mathbf{x}\|_b$$

Similarily,

$$mw\|\mathbf{x}\|_a \leq \|\mathbf{x}\|_c \leq MW\|\mathbf{x}\|_a$$

where $0 \leq mw \leq MW$, as required.

<u>Part i)</u>

$$\begin{aligned}
\|\mathbf{x}\|_2 &= (|\mathbf{x}_1|^2 + |\mathbf{x}_2|^2 + \cdots + |\mathbf{x}_n|^2)^{1/2} \\
&\leq |\mathbf{x}_1| + |\mathbf{x}_2| + \cdots + |\mathbf{x}_n| \qquad \text{by Cauchy-Schwarz} \\
&= \|\mathbf{x}\|_1
\end{aligned}$$

Exercise 3.29

The induced norm on $M_n(\mathbb{F})$ is given by

$$\|A\|_2 = \sup \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

Given an orthonormal matrix Q,

$$\begin{aligned}
\|Q\|_2 &= \sup \frac{\|Q\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \\
&= \sup \frac{\|\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \\
&= 1
\end{aligned}$$

Given a matrix norm, we know $\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|$. So

$$\begin{aligned}
\|R_{\mathbf{x}}\| &= \sup \frac{\|A\mathbf{x}\|}{\|A\|} \\
&= \sup \frac{\|A\mathbf{x}\|\|\mathbf{x}\|}{\|A\|\|\mathbf{x}\|} \\
&\leq \sup \frac{\|A\mathbf{x}\|\|\mathbf{x}\|}{\|A\mathbf{x}\|} \\
&= \|\mathbf{x}\|
\end{aligned}$$

Consider $\mathbf{x}$ such that $\|\mathbf{x}\| = 1$. If $\|R_{\mathbf{x}}\| < \|\mathbf{x}\| = 1$, it must be that $\|A\|\|\mathbf{x}\| < \|A\mathbf{x}\|$. This can be written as

$$\|A\| < \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$$

Exercise 3.47

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

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

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

iii. P is an idempotent matrix by i), which means $rank(P) = tr(P)$.

$$tr(P) = tr(A(A^H A)^{-1} A^H) = tr(A^H A(A^H A)^{-1}) = tr(I_n) = n = rank(P)$$