

OSM Lab 2017: Math Pset 2

Wei Han Chia

Due: 5 July 2017

Problems from the Book

3.1

We can derive each of the identities simply by using the properties of inner products in a real Inner Product Space (IPS).

(i) Polarization identity.

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

(ii) Parallelogram identity.

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

3.2

We perform a similar expansion as in 3.1, accounting for conjugates since we are in complex IPS.

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

3.3

We define the inner product on $\mathbb{R}[x]$ to be:

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx$$

We can then compute the angle θ between any pair of vectors x, y by $\theta = \arccos \frac{\langle x, y \rangle}{\|x\|\|y\|}$

(i)

$$\begin{aligned}\theta &= \arccos \frac{\int_0^1 x x^5 dx}{\sqrt{\int_0^1 x^2 dx} \sqrt{\int_0^1 x^{10} dx}} \\ &= \arccos \frac{1/7}{1/\sqrt{33}} \\ &= 0.608\end{aligned}$$

(ii)

$$\begin{aligned}\theta &= \arccos \frac{\langle x^2, x^4 \rangle}{\|x^2\| \|x^4\|} \\ &= \arccos \frac{\int_0^1 x^2 x^4 dx}{\sqrt{\int_0^1 x^4 dx} \sqrt{\int_0^1 x^8 dx}} \\ &= \arccos \frac{1/7}{1/\sqrt{45}} \\ &= 0.289\end{aligned}$$

3.8

Let V be the IPS $C([-\pi, \pi]; \mathbb{R})$ with inner product

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

(i) To show that S is an orthonormal set, we show that elements in S are all pairwise orthogonal, and have norm 1 (equivalently, the square of their norm is 1).

$$\begin{aligned}
\|\cos(t)\|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} + \cos(2t)/2 dt \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\|\sin(t)\|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} - \cos(2t)/2 dt \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\|\cos(2t)\|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(2t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} + \cos(4t)/2 dt \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\|\sin(2t)\|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(2t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} - \cos(4t)/2 dt \\
&= 1
\end{aligned}$$

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

$$\begin{aligned}
\langle \cos(2t), \sin(2t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(2t) \sin(2t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} \sin(4t) dt \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\langle \cos(t), \cos(2t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(t) \cos(2t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} (\cos(t) + \cos(3t)) dt \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\langle \cos(t), \sin(2t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(t) \sin(2t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} (\sin(t) + \sin(3t)) dt \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\langle \sin(t), \cos(2t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(t) \cos(2t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} (\sin(3t) - \sin(t)) dt \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\langle \sin(t), \sin(2t) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(t) \sin(2t) dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} (\cos(t) - \cos(3t)) dt \\
&= 0
\end{aligned}$$

(ii)

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

(iii)

$$\begin{aligned}
\text{proj}_{\text{span}X}(\cos(3t)) &= \sum_{x \in S} \langle x, \cos(3t) \rangle \frac{x}{\|x\|^2} \\
&= 0
\end{aligned}$$

We note that this is true since for every $x \in S$, the integral of the product of x and $\cos(3t)$ is equivalent to an integral of a symmetric function about 0, which will give us 0.

(iv)

$$\begin{aligned}
\text{proj}_{\text{span}X}(t) &= \sum_{x \in S} \langle x, \cos(3t) \rangle \frac{x}{\|x\|^2} \\
&= \frac{1}{\pi} \left[\int_{-\pi}^{\pi} t \sin(2t) dt + \int_{-\pi}^{\pi} \sin(t) dt \right]
\end{aligned}$$

We note that the terms involving integration of $t \cos(t)$ are equal to zero, since this is an odd function and we are integrating over a symmetric interval.

Next, we want to integrate by parts to solve for the two other integrals. We get that $\int_{-\pi}^{\pi} t \sin(2t) = -\pi$ and $\int_{-\pi}^{\pi} t \sin(t) = 2\pi$. Now combining this with our result in (ii) yields

$$\text{proj}_{\text{span}X}(t) = \sqrt{\frac{3}{2}} t$$

3.9

Consider the rotation matrix R . We will show that $\langle Rx, Ry \rangle = \langle x, y \rangle$. Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $y = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix}$.

$$\begin{aligned}
\langle x, y \rangle &= x_1 x_3 + x_2 x_4 \\
A &= \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \\
Ax &= \begin{bmatrix} x_1 \cos \theta - x_2 \sin \theta \\ x_1 \sin \theta + x_2 \cos \theta \end{bmatrix} \\
Ay &= \begin{bmatrix} x_3 \cos \theta - x_4 \sin \theta \\ x_3 \sin \theta + x_4 \cos \theta \end{bmatrix} \\
\langle Ax, Ay \rangle &= x_1 x_3 \cos^2 \theta - x_1 x_4 \sin \theta \cos \theta - x_2 x_3 \sin \theta \cos \theta + x_1 x_4 \sin^2 \theta + \\
&\quad x_1 x_3 \sin^2 \theta + x_1 x_4 \cos^2 \theta + x_1 x_4 \sin \theta \cos \theta + x_2 x_3 \sin \theta \cos \theta \\
&= (x_1 x_3 + x_2 x_4)(\sin^2 \theta + \cos^2 \theta) \\
&= x_1 x_3 + x_2 x_4 = \langle x, y \rangle
\end{aligned}$$

3.10

We will use the definition that a matrix A is orthonormal if $\langle Ax, Ay \rangle = \langle x, y \rangle$.

(i) Assume Q is orthonormal.

$$\begin{aligned}
\langle Qx, Qy \rangle &= \langle x, y \rangle \\
&= \langle Q^H Qx, y \rangle &= \langle x, QQ^H y \rangle
\end{aligned}$$

This implies that $Q^H Q = Q Q^H = I$.

Now assume $Q^H Q = Q Q^H = I$.

$$\begin{aligned}\langle Qx, Qy \rangle &= \langle Q^H Qx, y \rangle \\ &= \langle x, y \rangle\end{aligned}$$

So Q is orthonormal.

(ii) Since Q is orthonormal, $\|Q\mathbf{x}\|^2 = \langle Q\mathbf{x}, Q\mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$. Since norm is always non-negative, $\|Q\mathbf{x}\| = \|\mathbf{x}\|$.

(iii) We know that Q^{-1} is a matrix such that $QQ^{-1} = I$. Now it is clear that $Q^{-1} = Q^H$ in (i), and so it follows that Q^{-1} is also an orthonormal matrix, since we can switch the order of Q and Q^H .

(iv) From (i), we know that $Q^H Q = I$. We also know from definition of matrix multiplication, that each i, j entry of $Q^H Q = \sum_{k=1}^n q_{ki} q_{kj}$. Now this is clearly $\langle \mathbf{q}_i, \mathbf{q}_j \rangle$. Since $Q^H Q$ is the identity matrix, it follows that for all $i \neq j$, $\langle \mathbf{q}_i, \mathbf{q}_j \rangle = 0$ and $\langle \mathbf{q}_i, \mathbf{q}_i \rangle = 1$, or that the columns of Q are orthonormal.

(v) Consider any orthogonal matrix Q .

$$\begin{aligned}1 = \det(I) &= \det(Q^H Q) = \det(Q^H) \det(Q) = \det(Q)^2 \\ &\implies |\det(Q)| = 1\end{aligned}$$

The converse is not true. Consider the following symmetric real matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. Now clearly $A^2 \neq I$, but $\det(A) = 1$.

3.11

Lets consider a set of linearly dependent vectors, i.e. given $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, there exists $\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{j-1})$. Now our Gram-Schmidt orthonormalization process (GS) applied to $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}$ yields an orthonormal-basis for the span of that set $\mathbf{q}_1, \dots, \mathbf{q}_{j-1}$ where $\mathbf{v}_j = \sum_{i=1}^{j-1} \alpha_i \mathbf{q}_i$. Now applying GS yields

$$\begin{aligned}\mathbf{q}_j &= \frac{\mathbf{v}_j - \sum_{i=1}^{j-1} \langle \mathbf{q}_i, \mathbf{v}_j \rangle \mathbf{q}_i}{\|\mathbf{v}_j - \sum_{i=1}^{j-1} \langle \mathbf{q}_i, \mathbf{v}_j \rangle \mathbf{q}_i\|} \\ &= \frac{\mathbf{v}_j - \sum_{i=1}^{j-1} \alpha_i \mathbf{q}_i}{\|\mathbf{v}_j - \sum_{i=1}^{j-1} \alpha_i \mathbf{q}_i\|}\end{aligned}$$

However, we note that $\mathbf{v}_j - \sum_{i=1}^{j-1} \alpha_i \mathbf{q}_i = \mathbf{0}$, and so dividing by zero on the denominator will yield an error in the algorithm.

3.16

(i) Let $A \in M_{2 \times 2}$, $A = QR$. Now consider $D = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. Note that $DD = I$. Now $A = QDD^{-1}R$. Now note that QD is still orthonormal, since $QDD^H Q^H = I$. And $D^{-1}R$ is also still upper triangular. Thus we have another QR decomposition for A .

(ii) Let A have two QR decompositions with R having all positive diagonal elements.

$$\begin{aligned}Q_1 R_1 &= Q_2 R_2 \\ M &= Q_2^H Q_1 = R_2 R_1^{-1}\end{aligned}$$

Now note that M is orthonormal and upper triangular. It follows that M must be diagonal. Further, since we know that R_2 and R_1 have all positive diagonal entries, it follows that all entries of M must be positive as well. Finally, since M is orthonormal, it follows that all the diagonal entries are 1, i.e. that $M = I$. Then, we have

$$Q_2 = Q_1, R_2 = R_1$$

So the QR decomposition with R havin gall positive diagonal elements is unique.

3.17

Let $A \in M_{m \times n}$ have rank $n \leq m$. Let $A = \hat{Q} \hat{R}$ be a reduce QR decomposition. We want to solve $A^H A \mathbf{x} = A^H \mathbf{b}$.

$$\begin{aligned}A^H A \mathbf{x} &= A^H \mathbf{b} \\ \hat{R}^H \hat{Q}^H \hat{Q} \hat{R} \mathbf{x} &= \hat{R}^H \hat{Q}^H \mathbf{b} \\ \hat{R}^{-H} \hat{R}^H \hat{R} \mathbf{x} &= \hat{R}^{-H} \hat{R}^H \hat{Q}^H \mathbf{b} \\ \hat{R} \mathbf{x} &= \hat{Q}^H \mathbf{b}\end{aligned}$$

3.23

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

$$\begin{aligned}
& \|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\| \quad \text{Triangle Inequality} \\
\implies & \|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| \\
& \|\mathbf{y}\| \leq \|\mathbf{y} - \mathbf{x}\| + \|\mathbf{x}\| \quad \text{Triangle Inequality} \\
\implies & \|\mathbf{y}\| - \|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\| \\
& \text{Combining the two yields } \|\|\mathbf{x}\| - \|\mathbf{y}\|\| \leq \|\mathbf{x} - \mathbf{y}\|
\end{aligned}$$

3.24

Consider $C([a, b]; \mathbb{F})$ be the space of all continuous functions from $[a, b] \subset \mathbb{R}$ to \mathbb{F} . In order to prove that each operation is a norm, we will prove the following 3 properties:

- (1) $\|x\| \geq 0$, and $\|x\| = 0 \iff x = 0$
- (2) $\|\alpha x\| = |\alpha| \|x\|$
- (3) $\|x + y\| \leq \|x\| + \|y\|$

(i) $\|f\|_{L1} = \int_a^b |f(t)| dt.$

(1) Follows from the definition of $|\cdot|$

(2) $\|af\|_{L1} = |a| \int_a^b |f(t)| dt = |a| \|f\|_{L1}$

(3) $\|f + g\|_{L1} = \int_a^b |f + g| dt \leq \int_a^b |f| + |g| dt = \|f\|_{L1} + \|g\|_{L1}$

(ii) $\|f\|_{L2} = (\int_a^b |f(t)|^2 dt)^{1/2}.$

(1) Once again since $|x| = 0 \iff x = 0$, this holds

(2) $\|af\|_{L2} = (|a|^2 \int_a^b |f(t)|^2 dt)^{1/2} = |a| (\int_a^b |f(x)|^2 dt)^{1/2} = |a| \|f\|_{L2}$

(3) $\|f + g\|_{L2} = (\int_a^b |f(t) + g(t)|^2 dt)^{1/2} = (\int_a^b |f(t)|^2 dt + \int_a^b |g(t)|^2 dt + 2 \int_a^b |f(t)g(t)| dt)^{1/2}$
 $\leq (\int_a^b |f(t)|^2 dt)^{1/2} + (\int_a^b |g(t)|^2 dt)^{1/2} = \|f\|_{L2} + \|g\|_{L2}$

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

(1) Once again since $|x| = 0 \iff x = 0$, this holds

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

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

3.26

Topological equivalence between two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ is defined as $\exists m, M, 0 < m \leq M$ such that

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

To show that this is an equivalence relation, we will show that it is (1) reflexive, (2) symmetric and (3) transitive.

(1) $m = 1, M = 1$ then $\|\mathbf{x}\|_a = \|\mathbf{x}\|_a = \|\mathbf{x}\|_a$

(2) Given m, M such that $m\|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq M\|\mathbf{x}\|_a$

Consider $1/M < 1/m, 1/M\|\mathbf{x}\|_b \leq \|\mathbf{x}\|_a \leq 1/m\|\mathbf{x}\|_b$

(3) $a \sim b, b \sim c$ Consider $m_a m_b < M_a M_b$. Then

$$m_a m_b \|\mathbf{x}\|_a \leq \|\mathbf{x}\|_c \leq M_a M_b \|\mathbf{x}\|_a$$

(i) $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n}\|\mathbf{x}\|_2$. To save time on notation, all sums below are over n .

$$\begin{aligned}\|\mathbf{x}\|_2^2 &= \left(\sum (x_i)^2\right) = \sum |x_i|^2 \leq \left(\sum |x_i|\right)^2 = \|\mathbf{x}\|_1^2 \\ \Rightarrow \|\mathbf{x}\|_2 &\leq \|\mathbf{x}\|_1 \\ \|\mathbf{x}\|_1 &= \sum |x_i| = \sum |x_i| \cdot 1 \leq \left(\sum |x_i|^2\right)^{1/2} \left(\sum 1^2\right)^{1/2} = \sqrt{n}\|\mathbf{x}\|_2 \text{ By Cauchy Schwarz}\end{aligned}$$

(ii) $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\|\mathbf{x}\|_\infty$

$$\begin{aligned}\|\mathbf{x}\|_\infty &= \left(\max_i |x_i|\right)^2 \leq \sum x_i^2 = \|\mathbf{x}\|_2 \\ \Rightarrow \|\mathbf{x}\|_\infty &\leq \|\mathbf{x}\|_2 \\ \|\mathbf{x}\|_2 &= \left(\sum (x_i)^2\right)^{1/2} \leq \left(\sum \max_i x_i^2\right)^{1/2} = \sqrt{n}(\max_i (x_i)^2)^{1/2} = \sqrt{n} \max_i |x_i|\end{aligned}$$

3.28

Let A be an $n \times n$ matrix. We will show the topological equivalence of the operator p -norms. We will use the results from 3.27.

(i) From 3.26(i), we have $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n}\|\mathbf{x}\|_2$. We can also manipulate this inequality to get $\frac{1}{\sqrt{n}}\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$.

$$\begin{aligned}\frac{1}{\sqrt{n}}\|A\|_2 &= \frac{1}{\sqrt{n}} \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \\ &\leq \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_1}{\sqrt{n}\|\mathbf{x}\|_2} \\ &\leq \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \|A\|_1\end{aligned}$$

$$\begin{aligned}\|A\|_1 &= \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \\ &\leq \sqrt{n} \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_1} \\ &\leq \sqrt{n} \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sqrt{n}\|A\|_2\end{aligned}$$

(ii) From 3.26(ii), we have $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\|\mathbf{x}\|_\infty$. Manipulating this expression also yields $\frac{1}{\sqrt{n}}\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2$.

$$\begin{aligned}\frac{1}{\sqrt{n}}\|A\|_\infty &= \frac{1}{\sqrt{n}} \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \\ &\leq \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\sqrt{n}\|\mathbf{x}\|_\infty} \\ &\leq \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \|A\|_2\end{aligned}$$

$$\begin{aligned}\|A\|_2 &= \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \\ &\leq \sqrt{n} \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_2} \\ &\leq \sqrt{n} \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \sqrt{n}\|A\|_\infty\end{aligned}$$

3.29

Lets consider \mathbb{F}^n with the 2-norm, and let the norm on $M_n(\mathbb{F})$ be the induced norm. Note that given any orthonormal matrix $Q \in M_n(\mathbb{F})$, $\|Q\| = \sup_{\mathbf{x}=1} \|Q\mathbf{x}\|_2 = \sup_{\mathbf{x}=1} 1 = 1$.

Now consider the linear map $R_{\mathbf{x}} : M_n(\mathbb{F}) \rightarrow \mathbb{F}^n$ be given by $A \rightarrow A\mathbf{x}$.

$$\begin{aligned}\|R_{\mathbf{x}}\| &= \sup_{\|A\|_2 \neq 0} \frac{\|A\mathbf{x}\|_2}{\|A\|_2} \\ &\leq \frac{\|A\mathbf{x}\|_2 \|\mathbf{x}\|_2}{\|A\mathbf{x}\|_2} \\ &= \|\mathbf{x}\|_2\end{aligned}$$

Next, consider the alternate definition of $\|R_{\mathbf{x}}\|$. Given any vector \mathbf{x} , we can write it as $\|\mathbf{x}\|_2 \mathbf{x}/\|\mathbf{x}\|_2$.

$$\begin{aligned}\|R_{\mathbf{x}}\| &= \sup_{\|A\|_2=1} \|A\mathbf{x}\|_2 \\ &= \|\mathbf{x}\|_2 \sup_{\|A\|_2=1} \|A\mathbf{x}/\|\mathbf{x}\|_2\|_2\end{aligned}$$

Now since $\mathbf{v} = \mathbf{x}/\|\mathbf{x}\|_2$ has norm 1, it follows that we can find some orthonormal matrix with \mathbf{v} as its first row. This matrix will have norm 1, as we have proven above. Now this implies that:

$$\begin{aligned}\|R_{\mathbf{x}}\| &= \|\mathbf{x}\|_2 \sup_{\|A\|_2=1} \|A\mathbf{x}/\|\mathbf{x}\|_2\|_2 \\ &\geq \|\mathbf{x}\|_2\end{aligned}$$

Combining these two inequalities shows that $\|R_{\mathbf{x}}\|_2 = \|\mathbf{x}\|_2$.

3.30

Let $S \in M_n(\mathbb{F})$ be an invertible matrix. Given any norm $\|\cdot\|$, we define $\|\cdot\|_S$ by $\|A\|_S = \|SAS^{-1}\|$. We will prove that $\|\cdot\|_S$ is also a matrix norm, i.e. it fulfills the same properties we explored in 3.24.

- (1a) $\|A\|_S = \|SAS^{-1}\| \geq 0$
- (1b) $\|A\|_S = 0 \iff SAS^{-1} = 0 \iff A = 0$ Note here that $S \neq 0$ since S is invertible
- (2) $\|kA\|_S = \|SkAS^{-1}\| = |k|\|SAS^{-1}\| = |k|\|A\|_S$
- (3) $\|A + B\|_S = \|S(A + B)S^{-1}\| \leq \|SAS^{-1}\| + \|SBS^{-1}\| \leq \|A\|_S + \|B\|_S$ Triangle inequality for $\|\cdot\|$

3.37

Let $V = \mathbb{R}[x; 2]$ be the space of polynomials of at least two. We can consider the basis of V given by $[1 \ x \ x^2]$. Now we define $L : V \rightarrow \mathbb{R}$ to be given by $L[p] = p'(1)$.

Now given any polynomial $p(x) = ax^2 + bx + c$, we know that $p(1) = 2a + b$. So, the required $q \in V$ such that

$$L(p) = \langle q, p \rangle \text{ is } [0 \ 1 \ 2]. \text{ Since } [0 \ 1 \ 2] \begin{bmatrix} c \\ b \\ a \end{bmatrix} = 2a + b.$$

3.38

Let $V = \mathbb{F}[x; 2]$, the subspace of the inner product space $L^2([0, 1]; \mathbb{R})$. We want to get the derivative operator $D : V \rightarrow V$ given by $D[p(x)] = p'(x)$. Once again lets consider the power basis $[1, x, x^2]$.

Given any polynomial $p = ax^2 + bx + c$, $p' = 2ax + b$. So we want a matrix such that $A \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 2a \\ b \end{bmatrix}$.

It is easy to see that

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

3.39

We will prove proposition 3.7.12 on the properties of adjoints using the properties of inner products. Here, we also use the fact that if the adjoint exists, it is unique.

(i) If $S, T \in L(V; W)$, then $(S + T)^* = S^* + T^*$ and $(\alpha T)^* = \bar{\alpha} T^*$.

$$\begin{aligned}\langle y, (S + T)x \rangle &= \langle y, Sx \rangle + \langle y, Tx \rangle = \langle S^* y, x \rangle + \langle T^* y, x \rangle = \langle (S^* + T^*) y, x \rangle \\ \langle y, \alpha T x \rangle &= \alpha \langle y, T x \rangle = \alpha \langle T^* y, x \rangle = \langle \bar{\alpha} T^* y, x \rangle\end{aligned}$$

(ii) If $S \in L(V; W)$, then $(S^*)^* = S$.

$$\langle y, S^* x \rangle = \overline{\langle S^* x, y \rangle} = \overline{\langle x, S y \rangle} = \langle S y, x \rangle$$

(iii) If $S, T \in L(V; W)$, then $(ST)^* = T^* S^*$.

$$\langle y, STx \rangle = \langle S^* y, Tx \rangle = \langle T^* S^* y, x \rangle$$

(iv) If $T \in L(V; W)$, then $(T^*)^{-1} = (T^{-1})^*$.

$$S = T^{-1}$$

$$\text{From (iii) } (ST)^* = I^* = I = T^* S^*$$

$$T^* (T^{-1})^* = I \implies (T^{-1})^* = (T^*)^{-1}$$

3.40

Let $M_n(\mathbb{F})$ be endowed with the Frobenius inner product. Given any $A \in M_n(\mathbb{F})$, we can define a linear operator by left multiplication. The following results follow from the properties of trace.

(i) $A^* = A^H$.

$$\langle Y, AX \rangle = \text{tr}(Y^H AX) = \langle A^H Y, X \rangle$$

(ii) Given $A_1, A_2, A_3 \in M_n(\mathbb{F})$, $\langle A_2, A_3 A_1 \rangle = \langle A_2 A_1^* A_3 \rangle$. From (i), we know that $A^* = A^H$.

$$\langle A_2, A_3 A_1 \rangle = \text{tr}(A_2^* A_3 A_1) = \text{tr}(A_1 A_2^* A_3) = \langle A_2 A_1^*, A_3 \rangle$$

(iii) Define $T_A : M_n(\mathbb{F}) \rightarrow M_n(\mathbb{F})$ by $T_A(X) = AX - XA$. We want to show that $T_A^* = T_{A^*}$.

$$\langle Y, T_A(X) \rangle = \langle Y, AX - XA \rangle = \text{tr}(Y^* AX) - \text{tr}(Y^* XA) = \langle A^* Y, X \rangle - \langle Y A^*, X \rangle = \langle T_{A^*}(Y), X \rangle$$

3.44

Given $A \in M_{m \times n}(\mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^m$, we will prove that either $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x} \in \mathbb{F}^n$ or there exists $\mathbf{y} \in N(A^H)$ such that $\langle \mathbf{y}, \mathbf{b} \rangle \neq 0$.

Proof. We will use the fundamental subspaces theorem (FST). If there is a solution \mathbf{x} such that $A\mathbf{x} = \mathbf{b}$, then $\mathbf{b} \in R(A)$. Then by the FST, $\mathbf{b} \in N(A^H)^\perp$, i.e. $\forall \mathbf{y} \in N(A^H)$, $\langle \mathbf{y}, \mathbf{b} \rangle = 0$.

Now if there does not exist a solution, then $\mathbf{b} \notin N(A^H)^\perp$. I.e. $\exists \mathbf{y} \in N(A^H)$ such that $\langle \mathbf{y}, \mathbf{b} \rangle \neq 0$ □

3.45

Consider any $A \in \text{Skew}_n(\mathbb{R})$, we will first show that for any $B \in \text{Sym}_n(\mathbb{R})$, $\langle B, A \rangle = 0$.

$$\begin{aligned}\langle B, A \rangle &= \text{tr}(BA) = \text{tr}(A^H B^H) = -\text{tr}(BA) \\ \implies \text{tr}(BA) &= 0\end{aligned}$$

So it follows from our definition that $\text{Skew}_n(\mathbb{R}) \subset \text{Sym}_n(\mathbb{R})^\perp$.

Now consider any $X \in \text{Sym}_n(\mathbb{R})^\perp$ and $A \in \text{Sym}_n(\mathbb{R})$

$$\begin{aligned}\langle A, X \rangle &= \text{tr}(AX) = 0 = \text{tr}(-X^H A) \\ \implies X^H &= -X\end{aligned}$$

So $\text{Sym}_n(\mathbb{R})^\perp \subset \text{Skew}_n(\mathbb{R})$, and so $\text{Sym}_n(\mathbb{R})^\perp = \text{Skew}_n(\mathbb{R})$

3.46

Consider an $m \times n$ matrix A :

- (i) If $\mathbf{x} \in N(A^H A)$. Now we know that $A^H A \mathbf{x} = \mathbf{0}$, i.e. $A \mathbf{x} \in N(A^H)$. It is also clear that $A \mathbf{x} \in R(A)$.
- (ii) $N(A^H A) = N(A)$. We will prove that $N(A^H A) \subset N(A)$ and $N(A) \subset N(A^H A)$. Consider $\mathbf{x} \in N(A)$, then $A^H A \mathbf{x} = A^H \mathbf{0} = \mathbf{0}$, so $\mathbf{x} \in N(A^H A)$. Next consider $\mathbf{x} \in N(A^H A)$. Then $A^H A \mathbf{x} = \mathbf{0}$. Now $\mathbf{x}^T A^H A \mathbf{x} = 0 \implies \langle A \mathbf{x}, A \mathbf{x} \rangle = 0 \implies A \mathbf{x} = \mathbf{0}$. So $\mathbf{x} \in N(A)$.
- (iii) From rank nullity, $\dim(N(A^H A)) + \text{rank}(A^H A) = n$, and $\text{rank}(A) + \dim(N(A)) = n$. Now since $N(A^H A) = N(A)$, then it follows that they have the same dimension, and so A and $A^H A$ have the same rank.
- (iv) If A has linearly independent columns, then $\text{rank}(A) = n$. Now note that $A^H A$ is a $n \times n$ matrix with rank n . Since this matrix is full rank, therefore it is invertible, and hence non-singular.

3.47

Assume A is a $m \times n$ matrix of rank n . Let $P = A(A^H A)^{-1} A^H$. We will show the following:

- (i) $P^2 = P$.

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

$$P^H = (A(A^H A)^{-1} A^H)^H = A(A^H A)^{-1} A^H = A(A^H A)^{-1} A^H$$
- (iii) $\text{rank}(P) = n$. From rank nullity, we know that $\text{rank}(P) = n - \dim(N(P))$. Now we will show that $N(P) = \{\mathbf{0}\}$. Consider $\mathbf{b} \neq \mathbf{0} \in N(P)$.

$$A(A^H A)^{-1} A^H \mathbf{b} = \mathbf{0}$$

From our Least Squares Framework, we know that $A \hat{\mathbf{x}} = A(A^H A)^{-1} A^H \mathbf{b}$. Now we also know that \mathbf{x} is a unique solution to $A \mathbf{x} = \mathbf{b}$. However, since $A \hat{\mathbf{x}} = \mathbf{0}$, and A is injective, then it follows that $\mathbf{x} = \mathbf{0}$, which implies that $A \mathbf{x} = \mathbf{b} = \mathbf{0}$, a contradiction.

So the null space of P must be the zero vector, and so we are done.

3.48

Consider the vector space $M_n(\mathbb{R})$. Let $P(A) = \frac{A+A^T}{2}$ be the map from $M_n(\mathbb{R})$ to itself.

- (i) $P(X + Y) = \frac{(X+Y)+(X+Y)^T}{2} = \frac{X+X^T}{2} + \frac{Y+Y^T}{2} = P(X) + P(Y)$.
- (ii) $P^2 = \frac{(X+X^T)/2 + (X+X^T)/2}{2} = \frac{X+X^T}{2} = P(X)$.
- (iii)

$$\begin{aligned} \langle Y, P(X) \rangle &= \text{tr}\left(\frac{YX + YX^T}{2}\right) = \text{tr}\left(\frac{YX}{2}\right) + \text{tr}\left(\frac{X^T Y}{2}\right) \\ &= \text{tr}\left(\frac{Y + Y^T}{2} X\right) \\ &= \langle P(Y), X \rangle = \langle P^*(Y), X \rangle \end{aligned}$$

(iv) $N(P)$ is the set of all \mathbf{x} such that $P(\mathbf{x}) = \mathbf{0}$. Now from the definition of A , this is true if $f A^T = -A$, which implies that $A \in \text{Skew}_n(\mathbb{R})$.

(v) Consider any $\frac{A+A^T}{2}$. Then note that $\frac{(A+A^T)^T}{2} = \frac{A+A^T}{2}$. So $R(P) \subset \text{Sym}_n(\mathbb{R})$. Now consider any $A \in \text{Sym}_n(\mathbb{R})$. Now $A = A^T = \frac{A+A^T}{2}$. So $\text{Sym}_n(\mathbb{R}) \subset R(P)$.

(vi)

$$\begin{aligned} \|A - P(A)\|_F &= \sqrt{\text{tr}\left(\frac{(A^T - A)(A^T - A)}{4}\right)} = \sqrt{\frac{2\text{tr}(A^T A) - \text{tr}(A^2) - \text{tr}(((A^T)^2))}{4}} \\ &\quad \text{Note that } \text{tr}(A^T A^T) = \text{tr}(A^2) \\ &= \sqrt{\frac{\text{tr}(A^T A) - \text{tr}(A^2)}{2}} \end{aligned}$$

3.50

Let $(x_i, y_i)_{i=1}^n$ be a collection of data points of the form $rx^2 + sy^2 = 1$. We can manipulate the equation to give $y^2 = \frac{1}{s} - \frac{r}{s}x^2$. Now we can write the normal equation.

$$A = \begin{bmatrix} x_1^2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} \frac{r}{s} \\ \frac{1}{s} \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} y_1^2 \\ \vdots \\ y_n^2 \end{bmatrix}$$