

Chapter 3 Math

Lehner White

October 20, 2015

Exercise 3.1

i)

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

ii)

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

Exercise 3.2

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

Exercise 3.3

i)

$$\begin{aligned}
\cos(\theta) &= \frac{\langle x, x^5 \rangle}{\|x\| \|x^5\|} \\
\langle x, x^5 \rangle &= \int_0^1 x^6 dx = \frac{1}{7} x^7 \Big|_0^1 = \frac{1}{7} \\
\|x\| &= \sqrt{\langle x, x \rangle} = \int_0^1 x^2 dx = \frac{1}{3} x^3 \Big|_0^1 = \frac{1}{\sqrt{3}} \\
\|x^5\| &= \sqrt{\langle x^5, x^5 \rangle} = \int_0^1 x^{10} dx = \frac{1}{11} x^{11} \Big|_0^1 = \frac{1}{\sqrt{11}} \\
\theta &= \cos^{-1}\left(\frac{\sqrt{33}}{7}\right) \approx .60824
\end{aligned}$$

ii)

$$\begin{aligned}
\cos(\theta) &= \frac{\langle x^2, x^4 \rangle}{\|x^2\| \|x^4\|} \\
\langle x^2, x^4 \rangle &= \int_0^1 x^5 dx = \frac{1}{6} x^6 \Big|_0^1 = \frac{1}{6} \\
\|x^2\| &= \sqrt{\langle x^2, x^2 \rangle} = \int_0^1 x^4 dx = \frac{1}{5} x^5 \Big|_0^1 = \frac{1}{\sqrt{5}} \\
\|x^4\| &= \sqrt{\langle x^4, x^4 \rangle} = \int_0^1 x^8 dx = \frac{1}{9} x^9 \Big|_0^1 = \frac{1}{\sqrt{9}} \\
\theta &= \cos^{-1}\left(\frac{\sqrt{45}}{6}\right) \approx .48121
\end{aligned}$$

Exercise 3.4

Exercise 3.5

$$f = e^x \quad g = x - 1$$

$$\text{proj}_u(f) = \frac{\langle g, f \rangle}{\langle g, g \rangle} \cdot g$$

$$\langle g, f \rangle = \langle e^x, x - 1 \rangle = \int_0^1 x e^x - e^x dx = x e^x - 2e^x \Big|_0^1 = (e - 2e) + (2) = -e + 2$$

$$\langle g, g \rangle = \int_0^1 (x - 1)^2 = \int_0^1 x^2 - 2x + 1 dx = \frac{1}{3} x^3 - x^2 + x \Big|_0^1 = \frac{1}{3}$$

$$\frac{-e + 2}{\frac{1}{3}} \cdot (x - 1) = (-3e + 6)(x - 1)$$

Exercise 3.6

Starting with $0 \leq \|x - \lambda y\|^2$ we define $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$ and we continue as follows:

$$\|x - \lambda y\|^2 = \langle x - \lambda y, x - \lambda y \rangle$$

Note that $x - \lambda y$ is orthogonal to y and thus λy , meaning that $\langle x - \lambda y, -\lambda y \rangle = 0$ and $\langle -\lambda y, x - \lambda y \rangle = 0$.

$$\begin{aligned} 0 &\leq \|x - \lambda y\|^2 = \langle x - \lambda y, x - \lambda y \rangle \\ &= \langle x, x \rangle - \langle \lambda y, x \rangle \\ &= \|x\|^2 - \frac{\langle x, y \rangle^2}{\langle y, y \rangle^2} \\ &= \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2} \\ 0 &\leq \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2} \\ &\Rightarrow \frac{\langle x, y \rangle^2}{\|y\|^2} \leq \|x\|^2 \\ &\Rightarrow \frac{\langle x, y \rangle}{\|y\|} \leq \|x\| \\ &\Rightarrow \langle x, y \rangle \leq \|x\| \|y\| \end{aligned}$$

Exercise 3.7

Starting with Bessel's inequality saying that $\|v\|^2 \geq \sum_{i=1}^m |\langle x_i, v \rangle|^2 = \|\text{proj}_x(v)\|^2$ and we define $x_i = \{\frac{x}{\|x\|}\}$.

$$\begin{aligned}\|v\|^2 &\geq \sum_{i=1}^m |\langle \frac{x}{\|x\|}, v \rangle|^2 \geq |\langle \frac{x}{\|x\|}, v \rangle|^2 \\ \|v\| &\geq \frac{1}{\|x\|} |\langle x, v \rangle| \\ \|v\| \|x\| &\geq |\langle x, v \rangle|\end{aligned}$$

Exercise 3.8

i)

We show that the set is normal by showing that for $s \in S$ $\|s\| = 1$, and show that it is orthogonal by showing that $\langle s_1, s_2 \rangle = 0 \forall s_1, s_2 \in S$.

$$\begin{aligned}\langle \cos(x), \cos(x) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(x) dx = \frac{1}{\pi} [x + \sin(x)\cos(x)]_0^{\pi} = 1 \\ \langle \sin(x), \sin(x) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(x) dx = \frac{1}{\pi} [x - \sin(x)\cos(x)]_0^{\pi} = 1\end{aligned}$$

Our other two elements follow from the first as they are simply a change in period within our trigonometric functions.

$$\begin{aligned}\langle \cos(x), \sin(x) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(x)\sin(x) dx = (\frac{1}{\pi})(\frac{-1}{2})\cos^2(x)_{-\pi}^{\pi} = 0 \\ \langle \cos(x), \cos(2x) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(x)\cos(2x) dx = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(x)[\cos^2(x) - \sin^2(x)] dx \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^3(x) dx - \int_{-\pi}^{\pi} \sin^2(x)\cos(x) dx = \frac{1}{\pi}(0) = 0 \\ \langle \cos(x), \sin(2x) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(x)\sin(2x) dx = \frac{1}{\pi} [\frac{-2}{3}\cos^3(x)]_{-\pi}^{\pi} = 0 \\ \langle \sin(x), \cos(2x) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(x)\sin(x) dx - \int_{-\pi}^{\pi} \sin^3(x) dx = -\frac{1}{3\pi}\cos^3(x)_{-\pi}^{\pi} = 0 \\ \langle \sin(x), \sin(2x) \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(x)\sin(2x) dx = \frac{2}{3\pi}\sin^3(x)_{-\pi}^{\pi} = 0\end{aligned}$$

Again because we solved for $\langle \cos(x), \sin(x) \rangle$ we know that the same is true of $\langle \cos(2x), \sin(2x) \rangle$.

ii)

$$\|t\| = \langle t, t \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 dt = \frac{1}{3\pi} [t^3]_{-\pi}^{\pi} = \frac{2\pi^2}{3}$$

iii)

$$\begin{aligned} \text{proj}_X(\cos(3t)) &= \langle \cos(t), \cos(3t) \rangle \cos(t) + \langle \sin(t), \cos(3t) \rangle \sin(t) + \\ &\quad \langle \cos(2t), \cos(3t) \rangle \cos(2t) + \langle \sin(2t), \cos(3t) \rangle \sin(2t) \\ &= \frac{1}{\pi} \left(\int_{-\pi}^{\pi} \cos(t) \cos(3t) dt \cos(t) + \int_{-\pi}^{\pi} \sin(t) \cos(3t) dt \sin(t) + \right. \\ &\quad \left. \int_{-\pi}^{\pi} \cos(2t) \cos(3t) dt \cos(2t) + \int_{-\pi}^{\pi} \sin(2t) \cos(3t) dt \sin(2t) \right) \\ &= \frac{1}{\pi} (0 \cdot \cos(t) + \frac{-3\pi}{2} \sin(t) + 0 + 4\pi \sin(2t)) = \frac{-3}{2} \sin(t) + 4 \sin(2t) \end{aligned}$$

This is a linear combination of the starting equations.

iv)

$$\begin{aligned} \text{proj}_X(t) &= \langle \cos(t), t \rangle \cos(t) + \langle \sin(t), t \rangle \sin(t) + \langle \cos(2t), t \rangle \cos(2t) + \langle \sin(2t), t \rangle \sin(2t) \\ &= \frac{1}{\pi} \left(\int_{-\pi}^{\pi} \cos(t) t dt \cos(t) + \int_{-\pi}^{\pi} \sin(t) t dt \sin(t) + \right. \\ &\quad \left. \int_{-\pi}^{\pi} \cos(2t) t dt \cos(2t) + \int_{-\pi}^{\pi} \sin(2t) t dt \sin(2t) \right) \\ &= \frac{1}{\pi} (2\pi \sin(t) - \pi \sin(2t)) = 2 \sin(t) - \sin(2t) \end{aligned}$$

This is a linear combination of the starting equations.

Exercise 3.11

i)

As we take the Hermetian the Q's cancel out due to orthogonality.

$$\|Qx\|_2 = \langle Qx, Qx \rangle = (Qx)^H Qx = x^H x = \langle x, x \rangle = \|x\|_2$$

ii)

By definition if we can multiply this product by it's Hermetian and obtain the identity matrix it is orthonormal.

$$\langle Q_1, Q_2 \rangle = (Q_1 Q_2)(Q_1 Q_2)^H = Q_1 Q_2 Q_2^H Q_1^H = Q_1 Q_1^H = I$$

iii)

Because Q is orthonormal, when we take the Hermetian of the inverse this results in the original Q allowing us to perform the following proof:

$$\langle Q^{-1}, Q^{-1} \rangle = Q^{-1}(Q^{-1})^H = Q^{-1}Q = I$$

iv)

\Rightarrow If Q is orthonormal the previous exercises show that the resulting statement is true. \Leftarrow If $Q^H Q = Q Q^H = I$ then by definition Q is orthogonal and further examination will show that $\|Q\| = 1$ and is orthonormal.

v)

Because we know that if Q is orthonormal then either the rows or the columns must be orthonormal. Suppose that the rows are orthonormal. Because we know that Q^T is also orthonormal then the columns of Q are orthonormal.

vi)

If $|\det(Q)| = 1$ this does not imply that Q is orthonormal. An example is $\begin{bmatrix} \frac{1}{2} & 2 \\ 0 & 1 \end{bmatrix}$

Exercise 3.12

If this process is applied to a linearly dependent then this process will yield zero vectors because they are linear combinations of each other.

Exercise 3.14

The set of orthogonal vectors for this set are:

$$E = \left\{ \frac{1}{\sqrt{\pi}}, \frac{x}{\sqrt{\frac{\pi}{2}}}, \frac{x^2 - \frac{1}{2}}{\sqrt{\frac{\pi}{8}}}, \frac{x^3 - \frac{3}{4}x}{\sqrt{\frac{7\pi}{8}}} \right\}$$

These are all orthogonal to each other.

Exercise 3.15

i)

If we say that $A = QR$ then:

$$(QR)^H QRx = R^H Q^H QRx = R^H Rx = x$$

$$x = (QR)^H b = R^H Q^H b \Rightarrow Rx = Q^H b$$

ii)

$$u_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \quad u_2 = \begin{bmatrix} 3\sqrt{2} + 6\sqrt{3} \\ -3\sqrt{2} + 6\sqrt{3} \\ 6\sqrt{3} \end{bmatrix} - \begin{bmatrix} 6\sqrt{3} \\ 6\sqrt{3} \\ 6\sqrt{3} \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} \\ -3\sqrt{2} \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 \end{bmatrix} \quad R = Q^T A = \begin{bmatrix} 12 & 18 \\ 0 & 6 \end{bmatrix}$$

Exercise 3.16

$$u_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \quad u_2 = \begin{bmatrix} -1 \\ 4 \\ -1 \\ 4 \end{bmatrix} - \begin{bmatrix} -3 \\ -3 \\ -3 \\ -3 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \\ -4 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{1}{2} \\ \frac{-1}{2} \\ \frac{1}{2} \\ \frac{-1}{2} \end{bmatrix}$$

$$Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} \end{bmatrix} \quad R = Q^T A = \begin{bmatrix} 2 & 3 \\ 0 & -5 \end{bmatrix}$$

$$x = R^{-1}Q^{-1}b \quad x = \begin{bmatrix} 4.6 \\ .9 \end{bmatrix}$$

Exercise 3.17

i)

If we consider the matrices $Q' = QD$ and $R' = D^{-1}R$ where D is a diagonal matrix:

$$A = Q'R' = (QD)(D^{-1}R) = QR$$

Showing that it is not unique in this circumstance.

ii)

Exercise 3.23

i)

A hyperplane can be defined by a unit vector $v \in \mathbb{F}^n$. Reflecting x across the hyperplane has a matrix representation $H_v = I - 2vv^H$. The vector v can be written in polar coordinates as $(\cos(\phi), \sin(\phi))^T$ for some $\phi \in [0, 2\pi)$. We will use the equation $H_v = I - 2(\frac{vv^H}{V^H V})$ in the proof.

We first start by noting that,

$$V = \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \end{bmatrix}$$

and we also have

$$V^H = [r\cos(\theta) \quad r\sin(\theta)]$$

continuing we get,

$$\begin{aligned} V^H V &= r\cos(\theta)r\cos(\theta) + r\sin(\theta)r\sin(\theta) \\ &= r^2\cos^2(\theta) + r^2\sin^2(\theta) \\ &= r^2(\cos^2(\theta) + \sin^2(\theta)) \\ &= r^2 \end{aligned}$$

also we have the numerator,

$$\begin{aligned} VV^H &= \begin{bmatrix} r\cos(\theta)r\cos(\theta) & r\sin(\theta)r\cos(\theta) \\ r\cos(\theta)r\sin(\theta) & r\sin(\theta)r\sin(\theta) \end{bmatrix} \\ &= \begin{bmatrix} r\cos^2(\theta) & r\sin(\theta)r\cos(\theta) \\ r\cos(\theta)r\sin(\theta) & r\sin^2(\theta) \end{bmatrix} \end{aligned}$$

using the three above equations we get

$$\begin{aligned} H_v &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2\left(\frac{r^2}{r^2}\right) \begin{bmatrix} r\cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & r\sin^2(\theta) \end{bmatrix} \\ &= \begin{bmatrix} 1 - 2\cos^2(\theta) & -2\cos(\theta)\sin(\theta) \\ 2\cos(\theta)\sin(\theta) & 1 - 2\sin^2(\theta) \end{bmatrix} \\ &= \begin{bmatrix} -\cos(2\theta) & -\sin(2\theta) \\ -\sin(2\theta) & 1 - 1 + \cos(2\theta) \end{bmatrix} \\ &= \begin{bmatrix} -\cos(2\theta) & -\sin(2\theta) \\ -\sin(2\theta) & \cos(2\theta) \end{bmatrix} \\ &= \begin{bmatrix} -\cos(-2\theta) & \sin(-2\theta) \\ \sin(-2\theta) & \cos(-2\theta) \end{bmatrix} \end{aligned}$$

Since θ was an arbitrary angle chosen at the beginning of the proof, as long as the interior of the functions are uniform we can redefine this value to be our new θ value.

ii)

In part one we proved that any reflection can be represented by a matrix:

$$\begin{bmatrix} -\cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

So we can say that two reflections can be represented as:

$$\begin{bmatrix} -\cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} -\cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Which when multiplied out gives:

$$\begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & -\cos(\theta)\sin(\theta) + \cos(\theta)\sin(\theta) \\ -\cos(\theta)\sin(\theta) + \cos(\theta)\sin(\theta) & \sin^2(\theta) + \cos^2(\theta) \end{bmatrix}$$

Which can be simplified down to:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus showing that two reflections is simply a rotation around the origin.

iii)

Let $(\rho\cos(\phi), \rho\sin(\phi))^T$ be some arbitrary vector in the plane \mathbb{R}^2

And so :

$$\begin{aligned} & \begin{bmatrix} -\cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \rho\cos(\phi) \\ \rho\sin(\phi) \end{bmatrix} \\ &= \begin{bmatrix} -\rho(\cos(\phi)\cos(\theta) + \sin(\theta)\cos(\phi)) \\ \rho(\sin(\phi)\cos(\theta) + \sin(\theta)\cos(\phi)) \end{bmatrix} = \begin{bmatrix} \rho(\sin(\theta)\cos(\phi)) - \cos(\phi)\cos(\theta) \\ \rho(\sin(\phi)\cos(\theta) + \sin(\theta)\cos(\phi)) \end{bmatrix} \end{aligned}$$

In general

$$\cos(u)\cos(v) - \sin(u)\sin(v) = \cos(u+v)$$

$$\sin(u)\cos(v) + \sin(v)\cos(u) = \sin(u+v)$$

So,

$$\rho \begin{bmatrix} (\sin(\theta)\cos(\phi)) - \cos(\phi)\cos(\theta) \\ (\sin(\phi)\cos(\theta) + \sin(\theta)\cos(\phi)) \end{bmatrix} = \rho \begin{bmatrix} (-1)\cos(\theta + \phi) \\ \sin(\theta + \phi) \end{bmatrix}$$

In order to be a complete rotation we would need

$$\begin{bmatrix} (-1)\cos(\theta + \phi) \\ \sin(\theta + \phi) \end{bmatrix} = \begin{bmatrix} \rho\cos(\phi) \\ \rho\sin(\phi) \end{bmatrix}$$

For contradiction's sake suppose $\rho \sin(\theta + \phi) = \rho \sin(\theta)$ Which is true if and only if $\phi = 2k\pi$ for some integer k. Which implies $-\rho \cos(\theta + 2k\pi) = \rho \cos(\theta)$ However $\cos(\theta + 2k\pi) = \cos(\theta)$ is only true if $\cos(\theta) = 0 \Rightarrow \theta = \frac{1}{2}\pi + n\pi$ for some integer n. Hence, $\rho \sin(\theta + \phi) = \rho \sin(\frac{1}{2}\pi + n\pi + 2k\pi) = \pm\rho$

And so

$$\begin{bmatrix} -\cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{bmatrix}$$

And so

$$\begin{bmatrix} -\cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \rho \cos(\phi) \\ \rho \sin(\phi) \end{bmatrix} = \begin{bmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \pm \rho \end{bmatrix} = \begin{bmatrix} \rho \\ 0 \end{bmatrix}$$

Yet for nonzero ρ

$$\begin{bmatrix} \rho \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ \pm \rho \end{bmatrix}$$

In which case we have a contradiction, and so the initially given matrix does not represent a complete rotation.

Exercise 3.24

In order to prove this, we need to prove that these norms do in fact exhibit the three characteristics of a norm. These being: (1) positivity (2) preservation of scalar multiplication, and (3) preservation of the triangle inequality.

i)

$\int_a^b |f(t)| dt = \int_a^b |-f(t)| dt$ showing it is always positive.

$\int_a^b |af(t)| dt = \int_a^b |a| |f(t)| dt$ preserving scalar multiplication.

$\int_a^b |f(t) \cdot g(t)| dt = \int_a^b |f(t)| |g(t)| dt$ preserving the triangle inequality.

ii)

$(\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} = (\int_a^b |-f(t)|^2 dt)^{\frac{1}{2}}$ showing it is always positive.

$(\int_a^b |af(t)|^2 dt)^{\frac{1}{2}} = (\int_a^b |a|^2 |f(t)|^2 dt)^{\frac{1}{2}} = |a| (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}}$ preserving scalar multiplication.

$(\int_a^b |f(t) + g(t)|^2 dt)^{\frac{1}{2}} \leq (\int_a^b |f(t)|^2 |g(t)|^2 dt)^{\frac{1}{2}} \leq (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} + (\int_a^b |g(t)|^2 dt)^{\frac{1}{2}}$ preserving the triangle inequality.

iii)

$\sup_{x \in [a,b]} |-f(x)| = \sup_{x \in [a,b]} |f(x)|$ showing it is always positive.

$\sup_{x \in [a,b]} |af(x)| = |a| (\sup_{x \in [a,b]} |f(x)|)$ preserving scalar multiplication.

$\sup_{x \in [a,b]} |f(x) + g(x)| \leq \sup_{x \in [a,b]} |f(x)| + \sup_{x \in [a,b]} |g(x)| \leq \sup_{x \in [a,b]} |f(x)| + \sup_{x \in [a,b]} |g(x)|$
preserving the triangle inequality.

Exercise 3.26

$$||x| - |y|| = |\sqrt{\langle x, x \rangle} - \sqrt{\langle y, y \rangle}| \leq |\sqrt{\langle x, x \rangle - \langle y, y \rangle}| = \|x - y\|$$

Exercise 3.27

We want to show that $a \equiv a$, remember that $\forall m \in (0, 1), M \in (0, \infty), m\|x\|_a \leq \|x\|_a \leq M\|x\|_a$

Proving that $b \equiv a$ if $a \equiv b$, we want to keep in mind that

$$\exists m_1, M_1 : m_1\|x\|_a \leq \|x\|_b \leq M_1\|x\|_a$$

If we let $m_2 = \frac{1}{m_1}$ and $M_2 = \frac{1}{M_1}$. We can say that:

$$m_2\|x\|_b \leq \|x\|_a \leq M_2\|x\|_b$$

Now to show transitivity we suppose $\exists M_1, M_2, m_1, m_2$ with

$$m_1\|x\|_a \leq \|x\|_b \leq M_1\|x\|_a \quad \& \quad m_2\|x\|_b \leq \|x\|_c \leq M_2\|x\|_b$$

Let $M_3 = M_1M_2$, and $m_3 = m_1m_2$

$$m_3\|x\|_a = m_1m_2\|x\|_a \leq m_2\|x\|_b \leq \|x\|_c \leq M_2\|x\|_b \leq M_1M_2\|x\|_a = M_3\|x\|_a$$

Showing that this is an equivalence relation.

Exercise 3.31

\Leftarrow Suppose that $a^p = b^q$. Then $(a^p/p + b^q/q) = a^p(1/p + 1/q) = a^p$ and $ab = (a^p)^{1/p}(b^q)^{1/q} = (a^p)^{1/p}(a^p)^{1/q} = a^p$, showing that $ab = a^p/p + b^q/q$.

\Rightarrow Suppose that $ab = (1/p)a^p + (1/q)b^q$ Upon dividing by ab and using the fact that $a^p/b = (a^p/b^q)^{1/q}$ and $b^q/a = (b^q/a^p)^{1/p}$, we see that

$$\frac{1}{p} \left(\frac{a^p}{b^q} \right)^{1/q} + \frac{1}{q} \left(\frac{b^q}{a^p} \right)^{1/p} = 1$$

Let $x = a^p/b^q$. Multiplying by $x^{1/p}$, we obtain

$$\frac{1}{p} x^{1/p+1/q} + \frac{1}{q} = 1$$

$$\rightarrow \frac{1}{p}x + \frac{1}{q} = 1$$

This implies that $x = a^p/b^q = 1$, so $a^p = b^q$, as desired.

Exercise 3.34

Letting $\theta = \frac{1}{2}$, we get :

$$\begin{aligned} a^{\frac{1}{2}}b^{\frac{1}{2}} &\leq \frac{1}{2}(a+b) \\ (ab)^{\frac{1}{2}} &\leq \frac{1}{2}(a+b) \\ (area)^{\frac{1}{2}} &\leq \frac{1}{4}Perimeter \\ P &\geq 4\sqrt{A} \end{aligned}$$

The minimum lies at $P = 4\sqrt{A}$ which holds only for

$$2(a+b) = 4\sqrt{ab} \Rightarrow 4(a+b)^2 = 16(ab) \Rightarrow 4(a-b)^2 = 0 \Rightarrow a = b$$

Exercise 3.37

The D operator is given by the matrix

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

The four fundamental subspaces of D are given by the column space of D, Null space of D, and column and null spaces of D adjoint.

column space of $D = span(1, 0, 0)^T, (0, 1, 0)^T$

nullspace of $D = span(1, 0, 0)^T$ column space of $D^* = (0, 0, 1)^T, (0, 1, 0)^T$

nullspace of $D^* = (0, 0, 1)^T$

Exercise 3.38

Null of $D = span(0, 0, 0, 1), (1, 0, -1, 0), (-1, 1, 0, 0)$

Column of $D = span(1, 0, 2)$

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

Column space of $D^T = span(1, 1, 1, 0)$

Null space of $D^T = span(0, -1, 0), (2, 0, -1)$

Exercise 3.41

i)

It is sufficient to show $\langle Y, AX \rangle = \langle A^H Y, X \rangle$
 $\langle Y, AX \rangle = \text{tr}(Y^H AX) = \text{tr}(A^H Y)^H X = \langle A^H Y, X \rangle$

ii)

$\langle V, WA \rangle = \text{tr}(V^H WA) = \text{tr}(AV^H W) = \text{tr}((VA^H)^H W) = \langle VA^*, W \rangle$

iii)

$\langle (T_A(B))^*, C \rangle = \langle B, T_A(C) \rangle = \langle B, AC - CA \rangle = \text{tr}(B^H(AC - CA)) =$
 $\text{tr}(B^H AC) - \text{tr}(B^H CA) = \langle B, AC \rangle - \langle B, CA \rangle = \langle A^* B, C \rangle - \langle BA^*, C \rangle$
 $= \text{tr}(B^H (A^*)^H C) - \text{tr}((A^*)^H B^H C) = \text{tr}(B^H (A^*)^H C - (A^*)^H B^H C) =$
 $\langle A^* B - BA^*, C \rangle = \langle T_{A^*}(B), C \rangle$

Thus, $(T_A)^* = T_{A^*}$

Exercise 3.43

Let A and B be arbitrary matrices from $\text{Sym}_n(\mathbb{R})$ and $\text{Skew}_n(\mathbb{R})$ respectively. So $A^T = A$ and $B^T = -B$. Remember that $\text{tr}(CD) = \text{tr}(DC)$ and $\text{tr}(C) = \text{tr}(C^T)$

$$\langle A, B \rangle = \text{tr}(A^T B) = \text{tr}((-1)A^T B^T) = -\text{tr}(A^T B^T) = -\text{tr}((BA)^T) = -\text{tr}(AB)$$

but the only way that $\text{tr}(AB) = -\text{tr}(AB)$ is if $\text{tr}(AB) = 0$. Therefore, the set of symmetric matrices in M is orthogonal to the set of skew matrices.

Exercise 3.44

i)

We notice that since $\mathbf{x} \in N(A^H A)$ we know that $\mathbf{x} \in R^n$. From previous theorems, we know that $R(A) = \{\mathbf{b} \in R^m : \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \in R^n\}$. So it's clear that $A\mathbf{x} \in R(A)$ because $A\mathbf{x}$ generates the elements in $R(A)$ by definition. Also $N(A^H)$ has all elements (y) such that $A^H \mathbf{y} = \mathbf{0}$. So clearly $A\mathbf{x} \in N(A^H)$ because $\mathbf{x} \in N(A^H A)$ which implies that $A^H(A\mathbf{x}) = \mathbf{0}$ for all $A\mathbf{x}$

ii)

Consider that $N(A)$ is all \mathbf{x} such that $A\mathbf{x} = \mathbf{0}$. Multiplying both sides by A^H , we get:

$$A^H A\mathbf{x} = A^H \mathbf{0} = \mathbf{0}$$

For all \mathbf{x} . Since $N(A^H A)$ is all \mathbf{x} such that $A^H A\mathbf{x} = \mathbf{0}$, it must be that $N(A^H A)$ is a subset of $N(A)$.

All we need to show is that $N(A)$ is also a subset of $N(A^H A)$

We know that $N(A^H A)$ is all \mathbf{x} such that $A^H A\mathbf{x} = \mathbf{0}$. We know that $\mathbf{x} \in N(A)$, then $A\mathbf{x}$ is in both $R(A)$ and $N(A^H)$

A most useful fact is that $R(A)$ and $N(A^H)$ are orthogonal, so the only element they can have in common is the zero vector. Thus, $A\mathbf{x} = \mathbf{0}$. So clearly $N(A)$ is a subset of $N(A^H A)$. So $N(A^H A) = N(A)$

iii)

From the previous part, we know that $N(A^H A) = N(A)$, their dimensions are equivalent. So $\dim(N(A^H A)) = \dim(N(A)) = a$ for some integer a .

Using the rank nullity theorem, we know that:

$$\text{rank}(A^H A) + \dim(N(A^H A)) = n$$

We also know that:

$$\text{rank}(A) + \dim(N(A)) = n$$

Working with these two equations

$$\text{rank}(A^H A) = n - \dim(N(A^H A))$$

$$\text{rank}(A) = n - \dim(N(A))$$

So both ranks are equal to $n - a$, since we defined the dimensions as equal to a . Therefore their ranks must be equal.

iv)

If A has linearly independent columns, then its rank is n . In the last part, we showed that A and $A^H A$ have the same rank. This implies that they both have rank n . Since $A^H A$ has rank n in this case its columns must be linearly independent, which means that $A^H A$ is nonsingular because it's $n \times n$.

Exercise 3.45

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

i)

$$P^2 = P$$

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

ii)

$$P^H = P$$

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

iii)

Since P is symmetric and idempotent:

$$\text{rank}(P) = \text{Trace}(P)$$

$$\text{rank}(A(A^H A)^{-1} A^H) = \text{Trace}(A(A^H A)^{-1} A^H) = \text{Trace}(A^H A(A^H A)^{-1}) = \text{Trace}(I)$$

Based on the rules of matrix multiplication, this version of the identity matrix is $n \times n$. So the sum of the main diagonal elements of P is n . So therefore:

$$\text{rank}(P) = n$$

Exercise 3.46

$$(QR)^H QR\hat{x} = (QR)^H b$$

$$R^H Q^H QR\hat{x} = R^H Q^H b$$

$$R^H R\hat{x} = R^H Q^H b$$

Since R is invertable,

$$R\hat{x} = Q^H b$$

Exercise 3.49

i)

Show P is linear. Note that $P(\alpha A + \beta B) = \frac{(\alpha A + \beta B) + (\alpha A + \beta B)^T}{2} = \frac{\alpha A + \alpha(A)^T}{2} + \frac{\beta B + \beta(B)^T}{2} = \alpha P(A) + \beta P(B)$, so P is linear.

ii)

$$\begin{aligned} P^2 &= P(P(A)) = \frac{\frac{A+A^T}{2} + \frac{(A+A^T)^T}{2}}{2} \\ &= \frac{\frac{A+A^T}{2} + \frac{A+A^T}{2}}{2} = \frac{A+A^T}{2} = P(A) \end{aligned}$$

iii)

Show $P^* = P$. We showed in exercise 41 that $A^* = A^H$ for any matrix A under the Frobenius inner product. Since we are in the real numbers, we have that $A^* = A^H = A^T$. So we have only to show that $P^T = P$. Well,

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

iv)

Suppose A is in the null space of P . So $\frac{A+A^T}{2} = 0$. Then $A/2 = -A^T/2$.

$$\Rightarrow A = -A^T$$

So $A \in \text{skew}_n(\mathbb{R})$

Now suppose that $A \in \text{skew}_n(\mathbb{R})$. Then $A = -A^T$, and therefore $P(A) = \frac{A+A^T}{2} = 0$, so A is in the null space of P . Therefore $N(P) = \text{skew}_n(\mathbb{R})$.

Extra Exercise

We measure the weight of a radioactive substance over time to get n observation points $\{t_i, w_i\}_{i=1}^n$. The substance decays exponentially according to the curve $w(t) = ae^{kt}$. By taking the log of both side, the equation becomes $\log w(t) = \log(a) + kt$. We can express this in matrix form $Ax = b$ where

$$A = \begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{bmatrix}, x = \begin{bmatrix} x \\ \log(a) \end{bmatrix}, \text{ and } b = \begin{bmatrix} \log(w_1) \\ \log(w_2) \\ \vdots \\ \log(w_n) \end{bmatrix}$$

If we have the data points (3.0,7.3), (4.0,3.5), (5.0,1.2), and (6.0,0.8), where the first point is time measured in years and the second point is the weight of the radioactive substance measured by grams. We might notice that this is similar to OLS and by using the familiar equation of $x = (A^T A)^{-1} A^T b$ to get $k = -0.7703$ and $a = 71.2736$. Therefore, the half life in this case is $-(\log 2)/k = 0.8998$.