OSM Boot Camp: **Math ProbSet2**

Harrison Beard

2 July 2018

Exercise 3.1.

Question. Verify the polarization and parallelogram identities on a real inner product space, with the usual norm $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ arising from the inner product:

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

(ii) $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \frac{1}{2}\left(\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2\right)$.

It can be shown that in any normed linear space over $\mathbb{R}$ for which (ii) holds, one can define an inner product by using (i).

Answer.

(i)

$$\begin{aligned}
\frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right) &= \frac{1}{4}\left(\left(\sqrt{\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle}\right)^2 - \left(\sqrt{\langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle}\right)^2\right) \\
&= \frac{1}{4}\left(\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle\right) \\
&= \frac{1}{4}\left(\langle \mathbf{x}+\mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}+\mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}-\mathbf{y}, \mathbf{y} \rangle\right) \\
&= \frac{1}{4}\left(\overline{\langle \mathbf{x}, \mathbf{x}+\mathbf{y} \rangle} + \overline{\langle \mathbf{y}, \mathbf{x}+\mathbf{y} \rangle} - \overline{\langle \mathbf{x}, \mathbf{x}-\mathbf{y} \rangle} + \overline{\langle \mathbf{y}, \mathbf{x}-\mathbf{y} \rangle}\right) \\
&= \frac{1}{4}\left(\langle \mathbf{x}, \mathbf{x}+\mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x}+\mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{x}-\mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x}-\mathbf{y} \rangle\right) \\
&= \frac{1}{4}\left(\cancel{\langle \mathbf{x}, \mathbf{x} \rangle} + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \cancel{\langle \mathbf{y}, \mathbf{y} \rangle} - \cancel{\langle \mathbf{x}, \mathbf{x} \rangle} + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle - \cancel{\langle \mathbf{y}, \mathbf{y} \rangle}\right) \\
&= \frac{1}{4}\left(\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle\right) \\
&= \frac{1}{4}\left(\langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} + \langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle}\right) \\
&= \frac{1}{4}\left(\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle\right) \\
&= \frac{1}{\cancel{4}}\left(\cancel{4} \cdot \langle \mathbf{x}, \mathbf{y} \rangle\right) \\
&= \langle \mathbf{x}, \mathbf{y} \rangle .
\end{aligned}$$

■

(ii)

$$\begin{aligned}
\frac{1}{2}\left(\|\mathbf{x}+\mathbf{y}\|^2+\|\mathbf{x}-\mathbf{y}\|^2\right) &= \frac{1}{2}\left(\left(\sqrt{\langle\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}\rangle}\right)^2+\left(\sqrt{\langle\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}\rangle}\right)^2\right)\\
&= \frac{1}{2}\left(\langle\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}\rangle+\langle\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}\rangle\right)\\
&= \frac{1}{2}\left(\langle\mathbf{x}+\mathbf{y},\mathbf{x}\rangle+\langle\mathbf{x}+\mathbf{y},\mathbf{y}\rangle+\langle\mathbf{x}-\mathbf{y},\mathbf{x}\rangle-\langle\mathbf{x}-\mathbf{y},\mathbf{y}\rangle\right)\\
&= \frac{1}{2}\left(\overline{\langle\mathbf{x},\mathbf{x}+\mathbf{y}\rangle}+\overline{\langle\mathbf{y},\mathbf{x}+\mathbf{y}\rangle}+\overline{\langle\mathbf{x},\mathbf{x}-\mathbf{y}\rangle}-\overline{\langle\mathbf{y},\mathbf{x}-\mathbf{y}\rangle}\right)\\
&= \frac{1}{2}\left(\langle\mathbf{x},\mathbf{x}+\mathbf{y}\rangle+\langle\mathbf{y},\mathbf{x}+\mathbf{y}\rangle+\langle\mathbf{x},\mathbf{x}-\mathbf{y}\rangle-\langle\mathbf{y},\mathbf{x}-\mathbf{y}\rangle\right)\\
&= \frac{1}{2}\left(\langle\mathbf{x},\mathbf{x}\rangle+\cancel{\langle\mathbf{x},\mathbf{y}\rangle}+\cancel{\langle\mathbf{y},\mathbf{x}\rangle}+\langle\mathbf{y},\mathbf{y}\rangle+\langle\mathbf{x},\mathbf{x}\rangle-\cancel{\langle\mathbf{x},\mathbf{y}\rangle}-\cancel{\langle\mathbf{y},\mathbf{x}\rangle}+\langle\mathbf{y},\mathbf{y}\rangle\right)\\
&= \frac{1}{2}\left(\langle\mathbf{x},\mathbf{x}\rangle+\langle\mathbf{y},\mathbf{y}\rangle+\langle\mathbf{x},\mathbf{x}\rangle+\langle\mathbf{y},\mathbf{y}\rangle\right)\\
&= \frac{1}{2}\left(2\cdot\langle\mathbf{x},\mathbf{x}\rangle+2\cdot\langle\mathbf{y},\mathbf{y}\rangle\right)\\
&= \langle\mathbf{x},\mathbf{x}\rangle+\langle\mathbf{y},\mathbf{y}\rangle\\
&= \|\mathbf{x}\|^2+\|\mathbf{y}\|^2.
\end{aligned}$$

■

Exercise 3.2.

Question. Verify the polarization identity on a complex inner product space, with the usual norm $\|\mathbf{x}\| = \sqrt{\langle\mathbf{x},\mathbf{x}\rangle}$ arising from the inner product:

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

A nice consequence of the polarization identity on a real or complex inner product space is that if two inner products induce the same norm, then the inner products are equal.

Answer.

$$
\begin{aligned}
& \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2-\|\mathbf{x}-\mathbf{y}\|^2+i\,\|\mathbf{x}-i\mathbf{y}\|^2-i\,\|\mathbf{x}+i\mathbf{y}\|^2\right) \\
= \; & \frac{1}{4}\left(\left(\sqrt{\langle\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}\rangle}\right)^2-\left(\sqrt{\langle\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}\rangle}\right)^2\right. \\
& \left.+i\left(\sqrt{\langle\mathbf{x}-i\mathbf{y},\mathbf{x}-i\mathbf{y}\rangle}\right)^2-i\left(\sqrt{\langle\mathbf{x}+i\mathbf{y},\mathbf{x}+i\mathbf{y}\rangle}\right)^2\right) \\
= \; & \frac{1}{4}\left(\langle\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}\rangle-\langle\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}\rangle\right. \\
& \left.+i\cdot\langle\mathbf{x}-i\mathbf{y},\mathbf{x}-i\mathbf{y}\rangle-i\cdot\langle\mathbf{x}+i\mathbf{y},\mathbf{x}+i\mathbf{y}\rangle\right) \\
= \; & \frac{1}{4}\left(\langle\mathbf{x}+\mathbf{y},\mathbf{x}\rangle+\langle\mathbf{x}+\mathbf{y},\mathbf{y}\rangle-\langle\mathbf{x}-\mathbf{y},\mathbf{x}\rangle+\langle\mathbf{x}-\mathbf{y},\mathbf{y}\rangle\right. \\
& \left.+i\cdot\langle\mathbf{x}-i\mathbf{y},\mathbf{x}\rangle-i^2\cdot\langle\mathbf{x}-i\mathbf{y},\mathbf{y}\rangle-i\cdot\langle\mathbf{x}+i\mathbf{y},\mathbf{x}\rangle-i^2\cdot\langle\mathbf{x}+i\mathbf{y},\mathbf{y}\rangle\right) \\
= \; & \frac{1}{4}\left(\overline{\langle\mathbf{x},\mathbf{x}+\mathbf{y}\rangle}+\overline{\langle\mathbf{y},\mathbf{x}+\mathbf{y}\rangle}-\overline{\langle\mathbf{x},\mathbf{x}-\mathbf{y}\rangle}+\overline{\langle\mathbf{y},\mathbf{x}-\mathbf{y}\rangle}\right. \\
& \left.+i\cdot\overline{\langle\mathbf{x},\mathbf{x}-i\mathbf{y}\rangle}-i^2\cdot\overline{\langle\mathbf{y},\mathbf{x}-i\mathbf{y}\rangle}-i\cdot\overline{\langle\mathbf{x},\mathbf{x}+i\mathbf{y}\rangle}-i^2\cdot\overline{\langle\mathbf{y},\mathbf{x}+i\mathbf{y}\rangle}\right) \\
= \; & \frac{1}{4}\left(\langle\mathbf{x},\mathbf{x}+\mathbf{y}\rangle+\langle\mathbf{y},\mathbf{x}+\mathbf{y}\rangle-\langle\mathbf{x},\mathbf{x}-\mathbf{y}\rangle+\langle\mathbf{y},\mathbf{x}-\mathbf{y}\rangle\right. \\
& \left.+i\cdot\langle\mathbf{x},\mathbf{x}+i\mathbf{y}\rangle-i^2\cdot\langle\mathbf{y},\mathbf{x}+i\mathbf{y}\rangle-i\cdot\langle\mathbf{x},\mathbf{x}-i\mathbf{y}\rangle-i^2\cdot\langle\mathbf{y},\mathbf{x}-i\mathbf{y}\rangle\right) \\
= \; & \frac{1}{4}\left(\langle\mathbf{x},\mathbf{x}\rangle+\langle\mathbf{x},\mathbf{y}\rangle+\langle\mathbf{y},\mathbf{x}\rangle+\langle\mathbf{y},\mathbf{y}\rangle-\langle\mathbf{x},\mathbf{x}\rangle+\langle\mathbf{x},\mathbf{y}\rangle+\langle\mathbf{y},\mathbf{x}\rangle-\langle\mathbf{y},\mathbf{y}\rangle\right. \\
& +i\cdot\langle\mathbf{x},\mathbf{x}\rangle+i^2\cdot\langle\mathbf{x},\mathbf{y}\rangle-i^2\cdot\langle\mathbf{y},\mathbf{x}\rangle-i^3\cdot\langle\mathbf{y},\mathbf{y}\rangle-i\cdot\langle\mathbf{x},\mathbf{x}\rangle+i^2\cdot\langle\mathbf{x},\mathbf{y}\rangle \\
& \left.-i^2\cdot\langle\mathbf{y},\mathbf{x}\rangle+i^3\cdot\langle\mathbf{y},\mathbf{y}\rangle\right) \\
= \; & \frac{1}{4}\left(\cancel{\langle\mathbf{x},\mathbf{x}\rangle}+\cancel{\langle\mathbf{x},\mathbf{y}\rangle}+\langle\mathbf{y},\mathbf{x}\rangle+\cancel{\langle\mathbf{y},\mathbf{y}\rangle}-\cancel{\langle\mathbf{x},\mathbf{x}\rangle}+\cancel{\langle\mathbf{x},\mathbf{y}\rangle}+\langle\mathbf{y},\mathbf{x}\rangle-\cancel{\langle\mathbf{y},\mathbf{y}\rangle}\right. \\
& \left.+\cancel{i\cdot\langle\mathbf{x},\mathbf{x}\rangle}-\cancel{\langle\mathbf{x},\mathbf{y}\rangle}+\langle\mathbf{y},\mathbf{x}\rangle+\cancel{i\cdot\langle\mathbf{y},\mathbf{y}\rangle}-\cancel{i\cdot\langle\mathbf{x},\mathbf{x}\rangle}-\cancel{\langle\mathbf{x},\mathbf{y}\rangle}+\langle\mathbf{y},\mathbf{x}\rangle-\cancel{i\cdot\langle\mathbf{y},\mathbf{y}\rangle}\right) \\
= \; & \frac{1}{4}\left(\langle\mathbf{y},\mathbf{x}\rangle+\langle\mathbf{y},\mathbf{x}\rangle+\langle\mathbf{y},\mathbf{x}\rangle+\langle\mathbf{y},\mathbf{x}\rangle\right) \\
= \; & \frac{1}{4}\left(\overline{\langle\mathbf{x},\mathbf{y}\rangle}+\overline{\langle\mathbf{x},\mathbf{y}\rangle}+\overline{\langle\mathbf{x},\mathbf{y}\rangle}+\overline{\langle\mathbf{x},\mathbf{y}\rangle}\right) \\
= \; & \frac{1}{4}\left(\langle\mathbf{x},\mathbf{y}\rangle+\langle\mathbf{x},\mathbf{y}\rangle+\langle\mathbf{x},\mathbf{y}\rangle+\langle\mathbf{x},\mathbf{y}\rangle\right) \\
= \; & \frac{1}{\cancel{4}}\left(\cancel{4}\cdot\langle\mathbf{x},\mathbf{y}\rangle\right) \\
= \; & \langle\mathbf{x},\mathbf{y}\rangle\,.
\end{aligned}
$$

■

Exercise 3.3.

Question. Let $\mathbb{R}[x]$ have the inner product

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x.$$

Using (3.8), find the angle θ between the following sets of vectors:
(i) x and x^5.
(ii) x^2 and x^4.

Answer.

(i) In (3.8), we define $\cos\theta$ for the angle θ between vectors x and y as

$$\frac{\langle x, y\rangle}{\|x\|\,\|y\|}.$$

Solving in our case, we have

$$\begin{aligned}
\cos\theta &= \frac{\left\langle x, x^5\right\rangle}{\|x\|\,\|x^5\|} \\
&= \frac{\left\langle x, x^5\right\rangle}{\sqrt{\langle x, x\rangle}\sqrt{\langle x^5, x^5\rangle}} \\
&= \frac{\int_0^1 x^6\,\mathrm{d}x}{\sqrt{\int_0^1 x^2\,\mathrm{d}x}\sqrt{\int_0^1 x^{10}\,\mathrm{d}x}} \\
&= \frac{\frac{1}{7}x^7\big|_0^1}{\frac{1}{3}x^3\big|_0^1\cdot\frac{1}{11}x^{11}\big|_0^1} \\
&= \frac{\frac{1}{7}}{\frac{1}{3}\cdot\frac{1}{11}} \\
&= \boxed{\frac{33}{7}}.
\end{aligned}$$

(ii)

$$\begin{aligned}
\cos\theta &= \frac{\left\langle x^2, x^4\right\rangle}{\|x^2\|\,\|x^4\|} \\
&= \frac{\left\langle x^2, x^4\right\rangle}{\sqrt{\langle x^2, x^2\rangle}\sqrt{\langle x^4, x^4\rangle}} \\
&= \frac{\int_0^1 x^6\,\mathrm{d}x}{\sqrt{\int_0^1 x^4\,\mathrm{d}x}\sqrt{\int_0^1 x^8\,\mathrm{d}x}} \\
&= \frac{\frac{1}{7}x^7\big|_0^1}{\frac{1}{5}x^5\big|_0^1\cdot\frac{1}{9}x^9\big|_0^1} \\
&= \frac{\frac{1}{7}}{\frac{1}{5}\cdot\frac{1}{9}} \\
&= \boxed{\frac{45}{7}}.
\end{aligned}$$

Exercise 3.8.

Question. Let V be the inner product space $\mathcal{C}\left([-\pi,\pi];\mathbb{R}\right)$ with inner product

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

Let $X = \operatorname{span}(S) \subset V$, where $S = \{\cos(t), \sin(t), \cos(2t), \sin(2t)\}$.
(i) Prove that S is an orthonormal set.
(ii) Compute $\|t\|$.
(iii) Compute the projection $\operatorname{proj}_X(\cos(3t))$.
(iv) Compute the projection $\operatorname{proj}_X(t)$.

Answer.

(i) Normality:

$$\begin{aligned}
\|\cos(t)\| = \langle\cos(t),\cos(t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\cos(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi}\cos^2(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi}\frac{1}{2}\left(\cos(2t)+\cos(0)\right)\,\mathrm{d}t \\
&= \underbrace{\frac{1}{\pi}\int_{-\pi}^{\pi}\frac{1}{2}\cos(2t)\,\mathrm{d}t}_{2t=:u\implies\begin{cases}\frac{\mathrm{d}u}{\mathrm{d}t}=2 \implies \mathrm{d}t=\frac{1}{2}\mathrm{d}u\\ u\in 2\cdot(-\pi,\pi)=(-2\pi,2\pi)\end{cases}} + \frac{1}{\pi}\int_{-\pi}^{\pi}\frac{1}{2}\,\mathrm{d}t \\
&= \frac{1}{4\pi}\int_{-2\pi}^{2\pi}\cos(u)\,\mathrm{d}u + \frac{1}{2\pi}\cdot t\big|_{-\pi}^{\pi} \\
&= \frac{1}{4\pi}\cdot\sin(u)\big|_{-2\pi}^{2\pi} + \frac{1}{2\pi}(\pi+\pi) \\
&= \frac{1}{4\pi}\left(\sin(2\pi)-\sin(-2\pi)\right)+1 \\
&= \frac{1}{4\pi}(0)+1 \\
&= 1.
\end{aligned}$$

$$\begin{aligned}
\|\sin(t)\| = \langle \sin(t), \sin(t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\sin(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin^2(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\cos(0) - \cos(2t)\right)\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{d}t - \frac{1}{4\pi}\int_{-2\pi}^{2\pi}\cos(u)\,\mathrm{d}u \\
&= \frac{1}{2\pi}(2\pi) - \frac{1}{4\pi}\cdot \sin(u)\big|_{-2\pi}^{2\pi} \\
&= 1.
\end{aligned}$$

$$\begin{aligned}
\|\cos(2t)\| = \langle \cos(2t), \cos(2t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(2t)\cos(2t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2(2t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\cos(4t) + \cos(0)\right)\,\mathrm{d}t \\
&= \frac{1}{8\pi}\int_{-4\pi}^{4\pi}\cos(u)\,\mathrm{d}u + \frac{1}{2\pi}\cdot t\big|_{-\pi}^{\pi} \\
&= \frac{1}{8\pi}\cdot \sin(u)\big|_{-4\pi}^{4\pi} + \frac{1}{2\pi}(2\pi) \\
&= 1.
\end{aligned}$$

$$\begin{aligned}
\|\sin(2t)\| = \langle \sin(2t), \sin(2t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(2t)\sin(2t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin^2(2t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\cos(0) - \cos(4t)\right)\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{d}t - \frac{1}{8\pi}\int_{-4\pi}^{4\pi}\cos(u)\,\mathrm{d}u \\
&= \frac{1}{2\pi}(2\pi) - \frac{1}{8\pi}\cdot \sin(u)\big|_{-4\pi}^{4\pi} \\
&= 1.
\end{aligned}$$

Orthogonality:

$$\begin{aligned}
\langle \cos(t), \sin(t)\rangle = \langle \sin(t), \cos(t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\cos(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\sin(2t) - \sin(0)\right)\,\mathrm{d}t \\
&= \frac{1}{4\pi}\int_{-2\pi}^{2\pi}\sin(u)\,\mathrm{d}u - 0 \\
&= \frac{1}{4\pi}\left(-\cos(u)\right)\big|_{-2\pi}^{2\pi} \\
&= \frac{1}{4\pi}(-1+1) \\
&= 0.
\end{aligned}$$

$$\begin{aligned}
\langle \cos(t), \cos(2t)\rangle = \langle \cos(2t), \cos(t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(2t)\cos(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\cos(3t)+\cos(-t)\right)\,\mathrm{d}t \\
&= \frac{1}{6\pi}\int_{-3\pi}^{3\pi} \cos(u)\,\mathrm{d}u - \frac{1}{2\pi}\cdot \sin(v)|_{\pi}^{-\pi} \\
&= \frac{1}{6\pi}\sin(u)|_{-3\pi}^{3\pi} - \frac{1}{2\pi}(\sin(-\pi)-\sin(\pi)) \\
&= \frac{1}{6\pi}\left(\sin(\pi)-\sin(-\pi)\right) - \frac{1}{2\pi}(\sin(-\pi)-\sin(\pi)) \\
&= \frac{1}{6\pi}\left(\sin(\pi)-\sin(-\pi)\right) - \frac{1}{2\pi}(\sin(-\pi)-\sin(\pi \\
&= 0.
\end{aligned}$$

$$\begin{aligned}
\langle \cos(t), \sin(2t)\rangle = \langle \sin(2t), \cos(t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(2t)\cos(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\sin(3t)-\sin(-t)\right)\,\mathrm{d}t \\
&= \frac{1}{6\pi}\int_{-3\pi}^{3\pi} \sin(u)\,\mathrm{d}u - \frac{1}{2\pi}\cdot \cos(-t)|_{-\pi}^{\pi} \\
&= \frac{1}{6\pi}\cdot -\cos(u)|_{-3\pi}^{3\pi} - \frac{1}{2\pi}\left(\cos(-\pi)-\cos(\pi)\right) \\
&= \frac{1}{6\pi}\left(-\cos(\pi)+\cos(-\pi)\right) - \frac{1}{2\pi}(-1+1) \\
&= \frac{1}{6\pi}(1-1) - 0 \\
&= 0.
\end{aligned}$$

$$\begin{aligned}
\langle \sin(t), \cos(2t)\rangle = \langle \cos(2t), \sin(t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(2t)\sin(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\sin(3t)-\sin(t)\right)\,\mathrm{d}t \\
&= \frac{1}{6\pi}\int_{-3\pi}^{3\pi} \sin(u)\,\mathrm{d}u - \frac{1}{2\pi}\cdot \cos(t)|_{-\pi}^{\pi} \\
&= \frac{1}{6\pi}\cdot -\cos(u)|_{-3\pi}^{3\pi} - \frac{1}{2\pi}\left(\cos(\pi)-\cos(-\pi)\right) \\
&= \frac{1}{6\pi}\left(-\cos(\pi)+\cos(-\pi)\right) - \frac{1}{2\pi}(1-1) \\
&= \frac{1}{6\pi}(1-1) - 0 \\
&= 0.
\end{aligned}$$

$$
\begin{aligned}
\langle \sin(t), \sin(2t)\rangle = \langle \sin(2t), \sin(t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(2t)\sin(t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\cos(-t)+\cos(3t)\right)\,\mathrm{d}t \\
&= -\frac{1}{2\pi}\int_{3\pi}^{-3\pi}\cos(u)\,\mathrm{d}u - \frac{1}{6\pi}\cdot\int_{-3\pi}^{3\pi}\cos(v)\,\mathrm{d}v \\
&= -\frac{1}{2\pi}\cdot \sin(u)\big|_{3\pi}^{-3\pi} - \frac{1}{6\pi}\cdot\sin(v)\big|_{-3\pi}^{3\pi} \\
&= -\frac{1}{2\pi}(\sin(-\pi)-\sin(\pi)) - \frac{1}{6\pi}(\sin(\pi)-\sin(-\pi)) \\
&= 0.
\end{aligned}
$$

$$
\begin{aligned}
\langle \cos(2t), \sin(2t)\rangle = \langle \sin(2t), \cos(2t)\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(2t)\cos(2t)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\sin(3t)-\sin(0)\right)\,\mathrm{d}t \\
&= \frac{1}{6\pi}\int_{-3\pi}^{3\pi}\sin(u)\,\mathrm{d}u - 0 \\
&= \frac{1}{6\pi}\cdot -\cos(u)\big|_{-3\pi}^{3\pi} \\
&= \frac{1}{6\pi}\left(-\cos(\pi)+\cos(-\pi)\right) \\
&= \frac{1}{6\pi}(1-1) \\
&= 0.
\end{aligned}
$$

■

(ii)

$$
\begin{aligned}
\|t\| &= \sqrt{\langle t, t\rangle} \\
&= \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} t^2\,\mathrm{d}t} \\
&= \sqrt{\frac{1}{\pi}\cdot\frac{1}{3}t^3\Big|_{-\pi}^{\pi}} \\
&= \sqrt{\frac{1}{\pi}\cdot\left(\frac{1}{3}\pi^3 - \frac{1}{3}(-\pi)^3\right)} \\
&= \sqrt{\frac{2}{3}\pi^2} \\
&= \boxed{\sqrt{\frac{2}{3}}\pi}.
\end{aligned}
$$

(iii)

$$
\begin{aligned}
\operatorname{proj}_X(\cos(3t)) &\triangleq \sum_{i=1}^{m} \langle \mathbf{x}_i, \cos(3t)\rangle \, \mathbf{x}_i \text{ for } \mathbf{x} \in \operatorname{basis}(X) \text{ because } X \subset V \text{ is orthonormal} \\
&= \langle \cos(t), \cos(3t)\rangle \cdot \cos(t) \\
&\quad + \langle \sin(t), \cos(3t)\rangle \cdot \sin(t) \\
&\quad + \langle \cos(2t), \cos(3t)\rangle \cdot \cos(2t) \\
&\quad + \langle \sin(2t), \cos(3t)\rangle \cdot \sin(2t) \text{ because } X \coloneqq \operatorname{span}\left(\{\cos(t), \sin(t), \cos(2t), \sin(2t)\}\right) \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(t)\cos(3t)\,\mathrm{d}t \cdot \cos(t) \\
&\quad + \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)\cos(3t)\,\mathrm{d}t \cdot \sin(t) \\
&\quad + \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(2t)\cos(3t)\,\mathrm{d}t \cdot \cos(2t) \\
&\quad + \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(2t)\cos(3t)\,\mathrm{d}t \cdot \sin(2t) \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\cos(4t) + \cos(2t)\right)\,\mathrm{d}t \cdot \cos(t) \\
&\quad + \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\sin(4t) - \sin(2t)\right)\,\mathrm{d}t \cdot \sin(t) \\
&\quad + \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\cos(5t) + \cos(t)\right)\,\mathrm{d}t \cdot \cos(2t) \\
&\quad + \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\left(\sin(5t) - \sin(t)\right)\,\mathrm{d}t \cdot \sin(2t)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\cos(t)}{2\pi}\left(\underbrace{\int_{-\pi}^{\pi} \cos(4t)\,\mathrm{d}t}_{4t \eqqcolon u_1 \implies \begin{cases} \frac{\mathrm{d}u_1}{\mathrm{d}t} = 4 \implies \mathrm{d}t = \frac{1}{4}\mathrm{d}u_1 \\ u_1 \in 4\cdot(-\pi,\pi) = (-4\pi, 4\pi) \end{cases}} + \int_{-\pi}^{\pi} \cos(2t)\,\mathrm{d}t \right) \\
&\quad + \frac{\sin(t)}{2\pi}\left(\int_{-\pi}^{\pi} \sin(4t)\,\mathrm{d}t - \int_{-\pi}^{\pi} \sin(2t)\,\mathrm{d}t \right) \\
&\quad + \frac{\cos(2t)}{2\pi}\left(\int_{-\pi}^{\pi} \cos(5t)\,\mathrm{d}t + \int_{-\pi}^{\pi} \cos(t)\,\mathrm{d}t \right) \\
&\quad + \frac{\sin(2t)}{2\pi}\left(\int_{-\pi}^{\pi} \sin(5t)\,\mathrm{d}t - \int_{-\pi}^{\pi} \sin(t)\,\mathrm{d}t \right) \\
&= \frac{\cos(t)}{2\pi}\left(\frac{1}{4}\int_{-4\pi}^{4\pi} \cos(u_1)\,\mathrm{d}u_1 + \frac{1}{2}\int_{-2\pi}^{2\pi} \cos(u_2)\,\mathrm{d}u_2 \right) \\
&\quad + \frac{\sin(t)}{2\pi}\left(\frac{1}{4}\int_{-4\pi}^{4\pi} \sin(u_3)\,\mathrm{d}u_3 - \frac{1}{2}\int_{-2\pi}^{2\pi} \sin(u_4)\,\mathrm{d}u_4 \right) \\
&\quad + \frac{\cos(2t)}{2\pi}\left(\frac{1}{5}\int_{-5\pi}^{5\pi} \cos(u_5)\,\mathrm{d}u_5 + \int_{-\pi}^{\pi} \cos(t)\,\mathrm{d}t \right) \\
&\quad + \frac{\sin(2t)}{2\pi}\left(\frac{1}{5}\int_{-5\pi}^{5\pi} \sin(u_6)\,\mathrm{d}u_6 - \int_{-\pi}^{\pi} \sin(t)\,\mathrm{d}t \right)
\end{aligned}
$$

$$\begin{aligned}
&= \frac{\cos(t)}{2\pi}\left(\frac{1}{4}\cdot\sin(u_1)\big|_{-4\pi}^{4\pi}+\frac{1}{2}\cdot\sin(u_2)\big|_{-2\pi}^{2\pi}\right)\\
&\quad+\frac{\sin(t)}{2\pi}\left(\frac{1}{4}\cdot(-\cos(u_3))\big|_{-4\pi}^{4\pi}-\frac{1}{2}\cdot(-\cos(u_4))\big|_{-2\pi}^{2\pi}\right)\\
&\quad+\frac{\cos(2t)}{2\pi}\left(\frac{1}{5}\cdot\sin(u_5)\big|_{-5\pi}^{5\pi}+\left(\sin(t)\big|_{-\pi}^{\pi}\right)\right)\\
&\quad+\frac{\sin(2t)}{2\pi}\left(\frac{1}{5}\cdot(-\cos(u_6))\big|_{-5\pi}^{5\pi}-\left((-\cos(t))\big|_{-\pi}^{\pi}\right)\right)\\
&= \frac{\cos(t)}{2\pi}\left(\cancel{\frac{1}{4}(\sin(0)-\sin(0))}+\cancel{\frac{1}{2}(\sin(0)-\sin(0))}\right)\\
&\quad+\frac{\sin(t)}{2\pi}\left(\cancel{\frac{1}{4}(-\cos(0)+\cos(0))}-\cancel{\frac{1}{2}(-\cos(0)+\cos(0))}\right)\\
&\quad+\frac{\cos(2t)}{2\pi}\left(\frac{1}{5}(\sin(\pi)-\sin(-\pi))+(\sin(\pi)-\sin(-\pi))\right)\\
&\quad+\frac{\sin(2t)}{2\pi}\left(\frac{1}{5}(-\cos(\pi)+\cos(-\pi))-(-\cos(\pi)+\cos(-\pi))\right)\\
&= \frac{\cos(2t)}{2\pi}\left(\cancel{\frac{1}{5}(0-0)+(0-0)}\right)+\frac{\sin(2t)}{2\pi}\left(\cancel{\frac{1}{5}(1-1)-(1-1)}\right)\\
&= \boxed{0}.
\end{aligned}$$

(iv)

$$\begin{aligned}
\mathrm{proj}_X(t) &\triangleq \sum_{i=1}^{m}\langle \mathbf{x}_i, t\rangle\,\mathbf{x}_i \text{ for } \mathbf{x}\in\mathrm{basis}(X) \text{ because } X\subset V \text{ is orthonormal}\\
&= \langle\cos(t),t\rangle\cdot\cos(t)\\
&\quad+\langle\sin(t),t\rangle\cdot\sin(t)\\
&\quad+\langle\cos(2t),t\rangle\cdot\cos(2t)\\
&\quad+\langle\sin(2t),t\rangle\cdot\sin(2t) \text{ because } X\coloneqq\mathrm{span}\left(\{\cos(t),\sin(t),\cos(2t),\sin(2t)\}\right)\\
&= \underbrace{\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(t)\cdot t\,\mathrm{d}t}_{\begin{cases}u_1=t\\ \mathrm{d}v_1=\cos(t)\end{cases}\Rightarrow\begin{cases}\mathrm{d}u_1=\mathrm{d}t\\ v_1=\sin(t)\end{cases}}\cdot\cos(t) + \underbrace{\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(t)\cdot t\,\mathrm{d}t}_{\begin{cases}u_2=t\\ \mathrm{d}v_2=\sin(t)\end{cases}\Rightarrow\begin{cases}\mathrm{d}u_2=\mathrm{d}t\\ v_2=-\cos(t)\end{cases}}\cdot\sin(t)\\
&\quad+\underbrace{\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(2t)\cdot t\,\mathrm{d}t}_{\begin{cases}u_3=t\\ \mathrm{d}v_3=\cos(2t)\end{cases}\Rightarrow\begin{cases}\mathrm{d}u_3=\mathrm{d}t\\ v_3=\frac{1}{2}\sin(2t)\end{cases}}\cdot\cos(2t) + \underbrace{\frac{1}{\pi}\int_{-\pi}^{\pi}\sin(2t)\cdot t\,\mathrm{d}t}_{\begin{cases}u_4=t\\ \mathrm{d}v_4=\sin(2t)\end{cases}\Rightarrow\begin{cases}\mathrm{d}u_4=\mathrm{d}t\\ v_4=-\frac{1}{2}\cos(2t)\end{cases}}\cdot\sin(2t)\\
&= \frac{\cos(t)}{\pi}\left(t\sin(t)-\int_{-\pi}^{\pi}\sin(t)\,\mathrm{d}t\right)+\frac{\sin(t)}{\pi}\left(-t\cos(t)+\int_{-\pi}^{\pi}\cos(t)\,\mathrm{d}t\right)\\
&\quad+\frac{\cos(2t)}{\pi}\left(\frac{1}{2}t\sin(2t)-\frac{1}{2}\int_{-\pi}^{\pi}\sin(2t)\,\mathrm{d}t\right)+\frac{\sin(2t)}{\pi}\left(-\frac{1}{2}t\cos(2t)+\frac{1}{2}\int_{-\pi}^{\pi}\cos(2t)\,\mathrm{d}t\right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{\cos(t)}{\pi}\left(t\sin(t)+\cos(t)\right)\Big|_{t=-\pi}^{\pi} + \frac{\sin(t)}{\pi}\left(-t\cos(t)+\sin(t)\right)\Big|_{t=-\pi}^{\pi} \\
&\quad + \frac{\cos(2t)}{\pi}\left(\frac{1}{2}t\sin(2t)+\frac{1}{4}\cos(w_1)\right)\Big|_{(t,w_1)=(-\pi,-2\pi)}^{(\pi,2\pi)} + \frac{\sin(2t)}{\pi}\left(-\frac{1}{2}t\cos(2t)+\frac{1}{4}\sin(w_2)\right)\Big|_{(t,w_2)=(-\pi,-2\pi)}^{(\pi,2\pi)} \\
&= \frac{\cos(\pi)}{\pi}\left(\pi\sin(\pi)+\cos(\pi)\right) - \frac{\cos(-\pi)}{\pi}\left(-\pi\sin(-\pi)+\cos(-\pi)\right) \\
&\quad + \frac{\sin(\pi)}{\pi}\left(-\pi\cos(\pi)+\sin(\pi)\right) - \frac{\sin(-\pi)}{\pi}\left(\pi\cos(-\pi)+\sin(-\pi)\right) \\
&\quad + \frac{\cos(2\pi)}{\pi}\left(\frac{1}{2}\pi\sin(2\pi)+\frac{1}{4}\cos(2\pi)\right) - \frac{\cos(-2\pi)}{\pi}\left(-\frac{1}{2}\pi\sin(-2\pi)+\frac{1}{4}\cos(-2\pi)\right) \\
&\quad + \frac{\sin(2\pi)}{\pi}\left(-\frac{1}{2}\pi\cos(2\pi)+\frac{1}{4}\sin(2\pi)\right) - \frac{\sin(-2\pi)}{\pi}\left(\frac{1}{2}\pi\cos(-2\pi)+\frac{1}{4}\sin(-2\pi)\right) \\
&= -\frac{1}{\pi}(-1) + \frac{1}{\pi}(-1) + \frac{1}{\pi}\left(\frac{1}{4}\right) - \frac{1}{\pi}\left(\frac{1}{4}\right) \\
&= \cancel{\frac{1}{\pi} - \frac{1}{\pi}} + \cancel{\frac{1}{4\pi} - \frac{1}{4\pi}} \\
&= \boxed{0}.
\end{aligned}$$

Exercise 3.9.

Question. Prove that a rotation (2.17) in $\mathbb{R}^2$ is an orthonormal transformation (with respect to the usual inner product).

Answer.

We can prove the orthonormality of this transformation by showing the following two identities:

$$\begin{aligned}
&\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}^{\mathsf{T}}\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix} \\
&= \begin{pmatrix}\cos^2(\theta)+\sin^2(\theta) & 0\\ 0 & \cos^2(\theta)+\sin^2(\theta)\end{pmatrix} \\
&= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \\
&= I.
\end{aligned}$$

$$\begin{aligned}
&\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}^{\mathsf{T}} \\
&= \begin{pmatrix}\cos^2(\theta)+\sin^2(\theta) & 0\\ 0 & \cos^2(\theta)+\sin^2(\theta)\end{pmatrix} \\
&= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \\
&= I.
\end{aligned}$$

■

Exercise 3.10.

Question. Recall the definition of an orthonormal matrix given in Definition 3.2.14. Assume the usual inner product on $\mathbb{F}^n$. Prove the following statements:
(i) The matrix $Q \in M_n(\mathbb{F})$ is an orthonormal matrix if and only if $Q^{\mathsf{H}}Q = QQ^{\mathsf{H}} = I$.
(ii) If $Q \in M_n(\mathbb{F})$ is an orthonormal matrix, then $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{F}^n$.
(iii) If $Q \in M_n(\mathbb{F})$ is an orthonormal matrix, then so is Q^{-1}.
(iv) The columns of an orthonormal matrix $Q \in M_n(\mathbb{F})$ are orthonormal.
(v) If $Q \in M_n(\mathbb{F})$ is an orthonormal matrix, then $|\det(Q)| = 1$. Is the converse true?
(vi) If $Q_1, Q_2 \in M_n(\mathbb{F})$ are orthonormal matrices, then the product Q_1Q_2 is also an orthonormal matrix.

Answer.

(i) If Q is orthonormal, then $\langle \mathbf{x}, \mathbf{y} \rangle = \langle Q\mathbf{x}, Q\mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$. Moreover, if $Q^{\mathsf{H}}Q = QQ^{\mathsf{H}} = I$, then note that

$$\begin{aligned} \langle Q\mathbf{x}, Q\mathbf{y} \rangle &= (Q\mathbf{x})^{\mathsf{H}}(Q\mathbf{y}) \\ &= x^{\mathsf{H}}Q^{\mathsf{H}}Q\mathbf{y} \\ &\overset{Q^{\mathsf{H}}Q=I}{=} \mathbf{x}^{\mathsf{H}}\mathbf{y}, \end{aligned}$$

again for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$. It follows that (again if $Q^{\mathsf{H}}Q = QQ^{\mathsf{H}} = I$)

$$\begin{aligned} \langle Q\mathbf{x}, Q\mathbf{y} \rangle &= (Q\mathbf{x})^{\mathsf{H}}(Q\mathbf{y}) \\ &= \mathbf{x}^{\mathsf{H}}Q^{\mathsf{H}}Q\mathbf{y} \\ &= \mathbf{x}^{\mathsf{H}}\mathbf{y} \\ &= \langle \mathbf{x}, \mathbf{y} \rangle . \end{aligned}$$

■

(ii)

$$\begin{aligned} \|Qx\| &= \sqrt{\langle Qx, Qx \rangle} \\ &= \sqrt{x^{\mathsf{H}}Q^{\mathsf{H}}Qx} \\ &= \sqrt{\langle x, x \rangle} \\ &= \|x\| . \end{aligned}$$

■

(iii) We can observe that if $QQ^{\mathsf{H}} = Q^{\mathsf{H}}Q = I$, then $Q^{-1} = Q^{\mathsf{H}}$. Moreover, we know Q^{H} is orthonormal because of the fact that $(Q^{\mathsf{H}})^{\mathsf{H}} = Q$. Then,

$$\begin{aligned} (Q^{\mathsf{H}})^{\mathsf{H}}Q^{\mathsf{H}} &= QQ^{\mathsf{H}} \\ &= I \\ &= Q^{\mathsf{H}}Q \\ &= Q^{\mathsf{H}}(Q^{\mathsf{H}})^{\mathsf{H}}. \end{aligned}$$

■

(iv) Let $\mathbf{v}_i$ be the i-th column (vector) of Q, which itself is orthonormal. Then, $(Q^{\mathsf{H}}Q)_{ij} = \mathbf{v}_i^{\mathsf{H}}\mathbf{v}_j = \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{i,j}$ where $\delta_{i,j}$ is the Kronecker delta. Thus, the columns of Q are orthonormal. ■

(v) No. Counterexample:

$$\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

■

(vi) Consider orthonormal matrices Q and P. Then,

$$\begin{aligned} (QP)^{\mathsf{H}}QP &= P^{\mathsf{H}}Q^{\mathsf{H}}QP \\ &= P^{\mathsf{H}}P \\ &= I \end{aligned}$$

and

$$\begin{aligned} QP(QP)^{\mathsf{H}} &= QPP^{\mathsf{H}}Q^{\mathsf{H}} \\ &= QQ^{\mathsf{H}} \\ &= I. \end{aligned}$$

■

Exercise 3.11.

Question. Describe what happens when we apply the Gram-Schmidt orthonormalization process to a collection of linearly *dependent* vectors.

Answer.

Consider a set $\{x_i\}_{i=1}^n$ of linearly dependent vectors in V, for $n \in \mathbb{N}$. Then let $\{x_i\}_{i=1}^{k-1}$ for $k \in (2, N)$ be a linearly *independent* set of vectors, also in V. Then, $\{q_i\}_{i=1}^{k-1}$ is a linearly independent set. But Gram Schmidt does not work in this situation; $q_k = 0$ because $x_k \in \operatorname{span}\left(\{x_i\}_{i=1}^{k-1}\right)$.

Exercise 3.16.

Question. Prove the following results about the QR decomposition:

(i) The QR decomposition is not unique. *Hint*: Consider matrices of the form QD and $D^{-1}R$, where D is a diagonal matrix.

(ii) If A is invertible, then there is a unique QR decomposition of A such that R has only positive diagonal elements.

Answer.

(i) Claim: $-Q$ for the Q in a resultant QR decomposition of $m \times n$ matrix A is orthonormal:

$$\begin{aligned} -Q(-Q)^{\mathsf{H}} &= -Q(-Q^{\mathsf{H}}) \\ &= QQ^{\mathsf{H}} \\ &= I, \end{aligned}$$

and
$$(-Q)^{\mathsf{H}}(-Q) = I.$$
Next, note that $-R$ is clearly upper triangular. Then,
$$\begin{aligned} A &= QR \\ &= (-Q)(-R). \end{aligned}$$
Thus, there are A, Q, R that satisfy both $A = QR$ and $A = (-Q)(-R)$.

(ii) If A is invertible and can be written as two different QR decompositions (QR and $\hat{Q}\hat{R}$) in which the diagonal entries of R and $\hat{R}$ are strictly positive. This means that R and $\hat{R}$ are both invertible. Thus, $\hat{R}^{-1}R = Q^{\mathsf{H}}\hat{Q}$. Because R and $\hat{R}$ are both upper triangular matrices, we know $\hat{R}^{-1}R$ is upper triangular. Moreover, since Q and $\hat{Q}$ are both orthonormal, then $Q^{\mathsf{H}}\hat{Q}$ is orthonormal. It follows that $\hat{R}^{-1}R = I$ and $R = \hat{R}$ and $Q = \hat{Q}$ (since the matrix inverse is always unique for a given matrix). ■

Exercise 3.17.

Question. Let $A \in M_{m\times n}$ have rank $n \leq m$, and let $A = \hat{Q}\hat{R}$ be a reduced QR decomposition. Prove that solving the system $A^{\mathsf{H}}A\mathbf{x}$

Answer.
Let A have full column rank where $A = \hat{Q}\hat{R}$ is of reduced form; then, $\hat{R}$ has full rank, and so $\hat{R}$ is invertible. Then,
$$\begin{aligned} A^{\mathsf{H}}A\mathbf{x} &= A^{\mathsf{H}}b \\ \Downarrow \quad & \quad \Downarrow \\ (\hat{Q}\hat{R})^{\mathsf{H}}\hat{Q}\hat{R}\mathbf{x} &= (\hat{Q}\hat{R})^{\mathsf{H}}b \\ \Downarrow \quad & \quad \Downarrow \\ \hat{R}^{\mathsf{H}}\hat{Q}^{\mathsf{H}}\hat{Q}\hat{R}\mathbf{x} &= \hat{R}^{\mathsf{H}}\hat{Q}^{\mathsf{H}}b \\ &\Downarrow \\ \hat{R}\mathbf{x} &= \hat{Q}^{\mathsf{H}}b \end{aligned}$$
■

Exercise 3.23.

Question. Let $(V, \|\cdot\|)$ be a normed linear space. Prove that $|\|\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x}-\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in V$. *Hint*: Prove $\|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x}-\mathbf{y}\|$ and $\|\mathbf{y}\| - \|\mathbf{x}\| \leq \|\mathbf{x}-\mathbf{y}\|$.

Answer.
Let $\mathbf{x}, \mathbf{y} \in V$. By definition, norms satisfy nonnegativity and the triangle inequality, so note that
$$\begin{aligned} \|\mathbf{x}\| - \|-\mathbf{y}\| &\leq \|\mathbf{x}\| + \|-\mathbf{y}\| \\ &\leq \|\mathbf{x}-\mathbf{y}\|, \end{aligned}$$

which implies by homogeneity that

$$\begin{aligned} \|\mathbf{y}\| - \|\mathbf{x}\| &\leq \|\mathbf{y} - \mathbf{x}\| \\ &= \|-(\mathbf{y} - \mathbf{x})\| \\ &= \|\mathbf{x} - \mathbf{y}\| . \end{aligned}$$

■

Exercise 3.24.

Question. Let $\mathcal{C}([a,b];\mathbb{F})$ be the vector space of all continuous functions from $[a,b] \subset \mathbb{R}$ to $\mathbb{F}$. Prove that each of the following is a norm on $\mathcal{C}([a,b];\mathbb{F})$:

(i) $\|f\|_{L^1} = \int_a^b |f(t)| \, \mathrm{d}t$.

(ii) $\|f\|_{L^2} = \left(\int_a^b |f(t)|^2 \, \mathrm{d}t\right)^{1/2}$.

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

I will omit this answer. ■

Exercise 3.26.

Question. Two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ on the vector space X are *topologically equivalent* if there exist constants $0 < m \leq M$ such that

$$m \|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq M \|\mathbf{x}\|_a \qquad \forall \mathbf{x} \in X.$$

Prove that topological equivalence is an equivalence relation. Then prove that the p-norms for $p = 1, 2, \infty$ on $\mathbb{F}^n$ are topologically equivalent by establishing the following inequalities:

(i) $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n} \|\mathbf{x}\|_2$.

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

Hint: Use the Cauchy-Schwarz inequality.
The idea of topological equivalence is especially important in Chapter 5.

Answer.

Let $\|\cdot\|_{\mathfrak{p}}$ be a norm on X, for $\mathfrak{p} \in \{a, b, c\}$. Clearly, $\|\cdot\|_{\mathfrak{p}}$ is topologically equivalent to itself for any $m \in (0,1]$ and $M \geq 1$.

Suppose that $\|\cdot\|_a$ is topologically equivalent to $\|\cdot\|_b$ with constants $0 < m \leq M$. Then, $\|\cdot\|_b$ is topologically equivalent to $\|\cdot\|_a$ with constants $0 < \frac{1}{M'} \leq \frac{1}{m'}$. This leads us to: If $\|\cdot\|_a$ is topologically equivalent to $\|\cdot\|_b$ with constants $0 < m \leq M$, and $\|\cdot\|_b$ is topologically equivalent with with $\|\cdot\|_c$ with constants $0 < m' \leq M'$, then $\|\cdot\|_a$ is topologically equivalent to $\|\cdot\|_b$ with constants $0 < mm' \leq MM'$.

Consider arbitrary $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$. Then, we can show $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n}\,\|\mathbf{x}\|_2$ and $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\,\|\mathbf{x}\|_\infty$ by the following:

$$\begin{aligned} \sum_{i=1}^n |x_i|^2 &\leq \left(\sum_{i=1}^n |x_i|^2 + 2\sum_{i\neq j} |x_i|\,|x_j|\right) \\ &= \left(\sum_{i=1}^n |x_i|\right)^2 \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^n |x_i| \cdot 1 &\leq \left(\sum_{i=1}^n |x_i^2|\right)^{1/2} \left(\sum_{i=1}^n 1^2\right)^{1/2} \\ &= \sqrt{n}\left(\sum_{i=1}^n |x_i|^2\right)^{1/2}. \end{aligned}$$

$$\begin{aligned} \max_{i\in[1,n]} |x_i| &= \left(\max_{i\in[1,n]} |x_i|^2\right)^{1/2} \\ &\leq \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} \end{aligned}$$

$$\sum_{i=1}^n |x_i|^2 \leq n \cdot \max_{i\in[1,n]} |x_i|^2.$$

$\blacksquare$

Exercise 3.28.

Question. Let A be an $n \times n$ matrix. Prove that the operator p-norms are topologically equivalent for $p = 1, 2, \infty$ by establishing the following inequalities:

(i) $\frac{1}{\sqrt{n}}\|A\|_2 \leq \|A\|_1 \leq \sqrt{n}\,\|A\|_2$.

(ii) $\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{n}\,\|A\|_\infty$.

Answer.

(i) From above, we know that

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

and

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

Thus, $\frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_1 \leq \|A\|_2$. ■

(ii)

$$\sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\sqrt{n}\,\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty}$$

$$\sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \geq \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_\infty}{\sqrt{n}\,\|\mathbf{x}\|_\infty}.$$

■

Exercise 3.29.

Question. Take $\mathbb{F}^n$ with the 2-norm, and let the norm on $M_n(\mathbb{F})$ be the corresponding induced norm. Prove that any orthonormal matrix $Q \in M_n(\mathbb{F})$ has $\|Q\| = 1$. For any $\mathbf{x} \in \mathbb{F}^n$, let $R_\mathbf{x} : M_n(\mathbb{F}) \to \mathbb{F}^n$ be the linear transformation $A \mapsto A\mathbf{x}$. Prove that the induced norm of the transformation $R_\mathbf{x}$ is equal to $\|\mathbf{x}\|_2$. *Hint*: First prove $\|R_\mathbf{x}\| \leq \|\mathbf{x}\|_2$. Then recall that by Gram-Schmidt, any vector $\mathbf{x}$ with norm $\|\mathbf{x}\|_2 = 1$ is part of an orthonormal basis, and hence is the first column of an orthonormal matrix. Use this to prove equality.

Answer.

Consider arbitrary $\mathbf{x} \neq \mathbf{0}$ and let $\|\cdot\|$ be the norm induced by the inner product. Then

$$\begin{aligned}\|Q\mathbf{x}\| &= (\langle Q\mathbf{x}, Q\mathbf{x}\rangle)^{1/2} \\ &= \left(\langle Q^{\mathsf{H}}Q\mathbf{x}, \mathbf{x}\rangle\right)^{1/2} \\ &= (\langle \mathbf{x}, \mathbf{x}\rangle)^{1/2} \\ &= \|\mathbf{x}\|,\end{aligned}$$

so

$$\begin{aligned}\|Q\| &= \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|Q\mathbf{x}\|}{\|\mathbf{x}\|} \\ &= 1.\end{aligned}$$

Then, for the $R_\mathbf{x}$ considered in the question, we have that

$$\begin{aligned}\|R_\mathbf{x}\| &= \sup_{A\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|A\|} \\ &= \sup_{A\neq\mathbf{0}} \frac{\|A\mathbf{x}\|\,\|\mathbf{x}\|}{\|A\|\,\|\mathbf{x}\|} \\ &\leq \sup_{A\neq\mathbf{0}} \left(\frac{\|A\mathbf{x}\|\,\|\mathbf{x}\|}{\|A\mathbf{x}\|}\right) \\ &= \|\mathbf{x}\|.\end{aligned}$$

■

Exercise 3.30.

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

Answer.

- $\|A\|_S = \left\|SAS^{-1}\right\| \geq 0$ for $A \in M_n(\mathbb{F})$ because $\|\cdot\|$ is a norm on $M_n(\mathbb{F})$ and $SAS^{-1} \in M_n(\mathbb{F})$.
- $\|\mathbf{0}\|_S = \left\|S\mathbf{0}S^{-1}\right\| = \|\mathbf{0}\| = 0$. If $0 = \|A\|_S = \left\|SAS^{-1}\right\|$, then $SAS^{-1} = \mathbf{0}$ implying $A = \mathbf{0}$.
- For arbitrary $a \in \mathbb{F}$,

$$\begin{aligned} \|aA\|_S &= \left\|SaAS^{-1}\right\| \\ &= \left\|aSAS^{-1}\right\| \\ &= |a| \left\|SAS^{-1}\right\| \\ &= |a| \, \|A\|_S . \end{aligned}$$

- For arbitrary $B \in M_n(\mathbb{F})$,

$$\begin{aligned} \|A+B\|_S &= \left\|S(A+B)S^{-1}\right\| \\ &= \left\|SAS^{-1} + SBS^{-1}\right\| \\ &\leq \left\|SAS^{-1}\right\| + \left\|SBS^{-1}\right\| \\ &= \|A\|_S + \|B\|_S . \end{aligned}$$

- So, $\|\cdot\|_S$ is a norm on $M_n(\mathbb{F})$. To show matrix norm:

$$\begin{aligned} \|AB\|_S &= \left\|SABS^{-1}\right\| \\ &= \left\|SAS^{-1}ABS^{-1}\right\| \\ &\leq \left\|SAS^{-1}\right\| \left\|SBS^{-1}\right\| , \end{aligned}$$

  so $\|AB\|_S \leq \|A\|_S \, \|B\|_S$. ■

Exercise 3.37.

Question. 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 $L[p] = p'(1)$. Find the unique $q \in V$ such that $L[p] = \langle q, p \rangle$, as guaranteed by the Riesz representation theorem. *Hint*: Look at the discussion just before Theorem 3.7.1.

Answer.

Note that $V := \mathbb{R}[x; 2] \cong \mathbb{R}^3$, so $V \ni p = ax^2 + bx + c$ can be represented as a vector in $\mathbb{R}^3$. Let $p = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$.

Take $q = \boxed{(2, 1, 0)}$ where $p'q = 2a + b = p'(1) = L[p]$.

Exercise 3.38.

Question. Let $V = \mathbb{F}[x;2]$, 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, $D[p](x) = p'(x)$. Write the matrix representation of D with respect to the power basis $[1, x, x^2]$ of $\mathbb{F}[x;2]$. Write the matrix representation of the adjoint of D with respect to this basis.

Answer.

Let $p = ax^2 + vx + c$ be an arbitrary element of $V = \mathbb{F}[x;2]$. Because $p = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ and $p' = D(p) = \begin{pmatrix} 0 \\ 2a \\ b \end{pmatrix}$, we know the matrix representation of D is

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

and

$$D^{\mathsf{H}} = D^{\mathsf{T}} = \boxed{\begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}}.$$

Exercise 3.39.

Question. Prove Proposition 3.7.12.
Proposition 3.7.12. Let V and W by fin-dim inner product spaces. The adjoint has the following properties:

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

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

(iii) If $S, T \in \mathscr{L}(V)$, then $(ST)^* = T^* S^*$.

(iv) If $T \in \mathscr{L}(V)$ and T is invertible, then $(T^*)^{-1} = (T^{-1})^*$.

Answer.

(i)

$$\begin{aligned} \langle (S+T)^* \mathbf{w}, \mathbf{v} \rangle_V &= \langle \mathbf{w}, (S+T)\mathbf{v} \rangle_W \\ &= \langle \mathbf{w}, S\mathbf{v} + T\mathbf{v} \rangle_W \\ &= \langle \mathbf{w}, S\mathbf{v} \rangle_W + \langle \mathbf{w}, T\mathbf{v} \rangle_W \\ &= \langle S^* \mathbf{w}, \mathbf{v} \rangle_V + \langle T^* \mathbf{w}, \mathbf{v} \rangle_V \\ &= \langle S^* \mathbf{w} + T^* \mathbf{w}, \mathbf{v} \rangle_V . \end{aligned}$$

From above, we have that $(S+T)^* = S^* + T^*$. Also,

$$\begin{aligned} \langle (\alpha T)^* \mathbf{w}, \mathbf{v} \rangle_V &= \langle \mathbf{w}, (\alpha T)\mathbf{v} \rangle_W \\ &= \langle \mathbf{w}, \alpha T \mathbf{v} \rangle_W \\ &= \alpha \langle \mathbf{w}, T\mathbf{v} \rangle \\ &= \alpha \langle T^* \mathbf{w}, \mathbf{v} \rangle \\ &= \langle \bar{\alpha} T^* \mathbf{w}, \mathbf{v} \rangle . \end{aligned}$$

■

(ii)

$$\begin{aligned} \langle \mathbf{w}, S\mathbf{v} \rangle_W &= \langle S^* \mathbf{w}, \mathbf{v} \rangle_V \\ &= \overline{\langle \mathbf{v}, S^* \mathbf{w} \rangle_V} \\ &= \overline{\langle S^{**} \mathbf{v}, \mathbf{w} \rangle_W} \\ &= \langle \mathbf{w}, S^{**} \mathbf{v} \rangle_W \end{aligned}$$

for $\mathbf{v} \in V$ and $\mathbf{w} \in W$, so it follows that $S = S^{**}$. ■

(iii)

$$\begin{aligned} \langle (ST)^* \mathbf{v}', \mathbf{v} \rangle_V &= \langle \mathbf{v}', (ST)\mathbf{v} \rangle_V \\ &= \langle \mathbf{v}', S(T\mathbf{v}) \rangle_V \\ &= \langle S^* \mathbf{v}', T\mathbf{v} \rangle_V \\ &= \langle T^* S^* \mathbf{v}', \mathbf{v} \rangle_V , \end{aligned}$$

■

(iv) From (iii), we have that

$$\begin{aligned} T^*(T^*)^{-1} &= (TT^{-1})^* \\ &= I^* \\ &= I. \end{aligned}$$

■

Exercise 3.40.

Question. Let $M_n(\mathbb{F})$ be endowed with the Frobenius inner product (see Example 3.1.7). Any $A \in M_n(\mathbb{F})$ defines a linear operator on $M_n(\mathbb{F})$ by left multiplication: $B \mapsto AB$.

(i) Show that $A^* = A^{\mathsf{H}}$.

(ii) Show that for any $A_1, A_2, A_3 \in M_n(\mathbb{F})$ we have $\langle A_2, A_3 A_1 \rangle = \langle A_2 A_1^*, A_3 \rangle$. *Hint*: Recall $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

(iii) Let $A \in M_n(\mathbb{F})$. Define the linear operator $T_A : M_n(\mathbb{F}) \to M_n(\mathbb{F})$ by $T_A(X) = AX - XA$, and show that $(T_A)^* = T_{A^*}$.

Answer.

(i) Consider arbitrary $B, C \in M_n(\mathbb{F})$. Then, by definition of Frobenius inner product, we have:

$$\begin{aligned} \langle B, AC \rangle_F &= \operatorname{tr}\left(B^{\mathsf{H}} AC\right) \\ &= \operatorname{tr}\left((A^{\mathsf{H}} B)^{\mathsf{H}} C\right) \\ &= \left\langle A^{\mathsf{H}} B, C \right\rangle_F . \end{aligned}$$

■

(ii)

$$\begin{aligned} \langle A_2, A_3 A_1 \rangle_F &= \operatorname{tr}\left(A_2^{\mathsf{H}} A_3 A_1\right) \\ &= \operatorname{tr}\left(A_1 A_2^{\mathsf{H}} A_3\right) \\ &= \operatorname{tr}\left((A_2 A_1^{\mathsf{H}})^{\mathsf{H}} A_3\right) \\ &= \left\langle A_2 A_1^{\mathsf{H}}, A_3 \right\rangle_F \\ &= \langle A_2 A_1^*, A_3 \rangle . \end{aligned}$$

■

(iii) For arbitrary $B, C \in M_n(\mathbb{F})$, we have $\langle B, AC - CA \rangle = \langle B, AC \rangle - \langle B, CA \rangle$. Then, we showed above, we then have $\langle B, CA \rangle = \langle BA^*, C \rangle$. Moreover,

$$\begin{aligned} \langle B, AC \rangle &= \operatorname{tr}\left(B^{\mathsf{H}} AC\right) \\ &= \operatorname{tr}\left((A^{\mathsf{H}} B)^{\mathsf{H}} C\right) \\ &= \left\langle A^{\mathsf{H}} B, C \right\rangle \\ &= \langle A^* B, C \rangle . \end{aligned}$$

Thus, it follows that $T_A^* = T_A$. ■

Exercise 3.44.

Question. Given $A \in M_{m\times n}(\mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^m$, prove the *Fredholm alternative*: Either $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x} \in \mathbb{F}^n$ or there exists $\mathbf{y} \in \mathscr{N}\left(A^{\mathsf{H}}\right)$ such that $\langle \mathbf{y}, \mathbf{b} \rangle \neq 0$.

Answer.

- If $\exists \mathbf{x} \in \mathbb{F}^n : A\mathbf{x} = \mathbf{b}$ then $\forall \mathbf{y} \in \mathcal{N}(A^{\mathsf{H}})$, so we have:

$$\begin{aligned} \langle \mathbf{y}, \mathbf{b} \rangle &= \langle \mathbf{y}, A\mathbf{x} \rangle \\ &= \left\langle A^{\mathsf{H}} \mathbf{y}, \mathbf{x} \right\rangle \\ &= \langle \mathbf{0}, \mathbf{x} \rangle \\ &= 0. \end{aligned}$$

- If $\exists \mathbf{y} \in \mathscr{N}(A^{\mathsf{H}}) : \langle y, b \rangle \neq 0$ then $\mathbf{b} \notin \mathscr{N}(A^{\mathsf{H}})^{\perp} = \mathscr{R}(A)$.
- Therefore, $\exists \mathbf{x} \in \mathbb{F}^n : A\mathbf{x} = \mathbf{b}$. ■

Exercise 3.45.

Question. Consider the vector space $M_n(\mathbb{R})$ with the Frobenius inner product (3.5). Show that $\mathrm{Sym}_n(\mathbb{R})^\perp = \mathrm{Skew}_n(\mathbb{R})$. (See Exercise 1.18 for the definition of Sym and Skew.)

Answer.

- Let $A \in \mathrm{Sym}_n(\mathbb{R})$ and $B \in \mathrm{Skew}_n(\mathbb{R})$. Then,

$$\begin{aligned} \langle B, A\rangle &= \operatorname{tr}\left(B^\mathsf{T} A\right) \\ &= \operatorname{tr}\left(AB^\mathsf{T}\right) \\ &= \operatorname{tr}\left(A^\mathsf{T}(-B)\right) \\ &= -\langle A, B\rangle . \end{aligned}$$

  (It follows that $\langle A, B\rangle = 0$ and $\mathrm{Skew}_n(\mathbb{R}) \subset \mathrm{Sym}_n(\mathbb{R})^\perp$.)

- Let $B \in \mathrm{Sym}_n(\mathbb{R})^\perp$, so $B + B^\mathsf{T} \in \mathrm{Sym}_n(\mathbb{R})$. Thus,

$$\begin{aligned} 0 &= \left\langle B + B^\mathsf{T}, B\right\rangle \\ &= \operatorname{tr}\left((B + B^\mathsf{T})B\right) \\ &= \operatorname{tr}\left(BB + B^\mathsf{T} B\right) \\ &= \operatorname{tr}(BB) + \operatorname{tr}\left(B^\mathsf{T} B\right), \end{aligned}$$

  so $\left\langle B^\mathsf{T}, B\right\rangle = \langle -B, B\rangle$ so $B^\mathsf{T} = -B$. Thus $\mathrm{Sym}_n(\mathbb{R})^\perp = \mathrm{Skew}_n(\mathbb{R})$. ∎

Exercise 3.46.

Question. Prove the following for an $m \times n$ matrix A:

(i) If $\mathbf{x} \in \mathscr{N}\left(A^\mathsf{H} A\right)$, then $A\mathbf{x}$ is in both $\mathscr{R}(A)$ and $\mathscr{N}\left(A^\mathsf{H}\right)$.

(ii) $\mathscr{N}\left(A^\mathsf{H} A\right) = \mathscr{N}(A)$.

(iii) A and $A^\mathsf{H} A$ have the same rank.

(iv) If A has linearly independent columns, then $A^\mathsf{H} A$ is nonsingular.

Answer.

(i) If $\mathbf{x} \in \mathscr{N}(A^\mathsf{H} A)$, then $\mathbf{0} = (A^\mathsf{H} A)\mathbf{x} = A^\mathsf{H}(A\mathbf{x})$ and $Ax \in \mathscr{N}(A^\mathsf{H})$. $A\mathbf{x} \in \operatorname{im}(A)$ by definition. ∎

(ii) If $\mathbf{x} \in \mathscr{N}(A)$, then $A\mathbf{x} = \mathbf{0}$. Then, $A^\mathsf{H} A\mathbf{x} = A^\mathsf{H}\mathbf{0} = \mathbf{0}$, so $\mathbf{x} \in \mathscr{N}(A^\mathsf{H} A)$. If $\mathbf{x} \in \mathcal{N}(A^\mathsf{H} A)$, then $\|A\mathbf{x}\| = \mathbf{x}^\mathsf{H} A^\mathsf{H} A\mathbf{x} = \mathbf{x}^\mathsf{H}\mathbf{0} = \mathbf{0}$, so $A\mathbf{x} = \mathbf{0}$ and $\mathbf{x} \in \mathscr{N}(A)$. ∎

(iii) $n = \operatorname{rank}(A) + \dim(\mathscr{N}(A))$ by rank-nullity, and $n = \operatorname{rank}\left(A^\mathsf{H} A\right) + \dim(\mathscr{N}(A^\mathsf{H} A))$. Then, by (ii), it follows that $\operatorname{rank}(A) = \operatorname{rank}\left(A^\mathsf{H} A\right)$. ∎

(iv) By (iii), we have that $n = \operatorname{rank}(A) = \operatorname{rank}\left(A^\mathsf{H} A\right)$. We know it is invertible by $A^\mathsf{H} A \in M_n$. ∎

Exercise 3.47.

Question. Assume A is an $m \times n$ matrix of rank n. Let $P = A\left(A^{\mathsf{H}}A\right)^{-1}A^{\mathsf{H}}$. Prove the following:

(i) $P^2 = P$.

(ii) $P^{\mathsf{H}} = P$.

(iii) $\operatorname{rank}(P) = n$.

Whenever a linear operator satisfies $P^2 = P$, it is called a *projection*. Projections are treated in detail in Section 12.1.

Answer.

(i)

$$\begin{aligned} P^2 &= (A(A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}})(A(A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}}) \\ &= A(A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}}A(A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}} \\ &= A(A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}} \\ &= P. \end{aligned}$$

■

(ii)

$$\begin{aligned} P^{\mathsf{H}} &= (A(A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}})^{\mathsf{H}} \\ &= (A^{\mathsf{H}})^{\mathsf{H}}(A^{\mathsf{H}}A)^{-\mathsf{H}}A^{\mathsf{H}} \\ &= A(A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}} \\ &= P. \end{aligned}$$

■

(iii) $\operatorname{rank}(A) = n \implies \operatorname{rank}(P) \leq n$. Thus, $\forall \mathbf{y} \in \operatorname{im}(A) : \exists \mathbf{x} \in \mathbb{F}^n : \mathbf{y} = A\mathbf{x}$. So,

$$\begin{aligned} P\mathbf{y} &= A(A^{\mathsf{H}}A)A^{\mathsf{H}}\mathbf{y} \\ &= A(A^{\mathsf{H}}A)^{-1}A^{\mathsf{H}}A\mathbf{x} \\ &= A\mathbf{x} \\ &= \mathbf{y}, \end{aligned}$$

so $\mathbf{y} \in \operatorname{im}(P)$. Thus, $\operatorname{rank}(P) \geq \operatorname{rank}(A)$, so $\operatorname{rank}(P) = p$. ■

Exercise 3.48.

Question. Consider the vector space $M_n(\mathbb{R})$ with the Frobenius inner product (3.5). Let $P(A) = \frac{A+A^{\mathsf{T}}}{2}$ be the map $P : M_n(\mathbb{R}) \to M_n(\mathbb{R})$. Prove the following:

(i) P is linear.

(ii) $P^2 = P$.

(iii) $P^* = P$ (note that $*$ here means the adjoint with respect to the Frobenius inner product).

(iv) $\mathscr{N}(P) = \mathrm{Skew}_n(\mathbb{R})$.

(v) $\mathscr{R}(P) = \mathrm{Sym}_n(\mathbb{R})$.

(vi) $\|A - P(A)\|_F = \sqrt{\frac{\mathrm{tr}(A^{\mathsf{T}}A) - \mathrm{tr}(A^2)}{2}}$. Here $\|\cdot\|_F$ is the norm with respect to the Frobenius inner product.

Hint: Recall that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ and $\mathrm{tr}(A) = \mathrm{tr}\left(A^{\mathsf{T}}\right)$.

I will omit this answer. ■

Exercise 3.50.

Question. 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 $rx^2 + sy^2 = 1$. We wish to find the least squares approximation for r and s. Write A, $\mathbf{x}$, and $\mathbf{b}$ for the corresponding normal equation in terms of the data x_i and y_i and the unknowns r and s.

Answer.

Consider the regression $\mathbf{y}^2 = \frac{1}{s} + \frac{r\mathbf{x}^2}{s}$ in the form $A\mathbf{x} = \mathbf{b}$. Let $b_i := y_i^2$ for each component $b_i, y_i \in \mathbf{b}, \mathbf{y}$, and let $\mathbf{x} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ where $\beta_1 = \frac{1}{s}$ and $\beta_2 = \frac{r}{s}$. Then the corresponding normal equation is $A^{\mathsf{H}}A\hat{x} = A^{\mathsf{H}}b$, where

$$\begin{aligned} A^{\mathsf{H}}A\hat{x} &= \begin{pmatrix} \sum_i 1 & \sum_i x_i^2 \\ \sum_i x_i^2 & \sum_i x_i^4 \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \\ &= \begin{pmatrix} n\hat{\beta}_1 - \hat{\beta}_2 \sum_i x_i^2 \\ \hat{\beta}_1 \sum_i x_i^2 - \hat{\beta}_2 \sum_i x_i^4 \end{pmatrix} \end{aligned}$$

and

$$A^{\mathsf{H}}b = \begin{pmatrix} \sum_i y_i^2 \\ \sum_i x_i^2 y_i^2 \end{pmatrix}.$$