Quiz 5

Problem 1 (15 points; 3 points each). Decide if each of the following are true or false and provide a justification or counterexample in each case. A justification could consist of a theorem from the text. All vector spaces are assumed to be finite-dimensional here.

(a) ______ There is a unique least squares solution $\hat{x} = (A^T A)^{-1} A^T \boldsymbol{b}$ to $A\boldsymbol{x} = \boldsymbol{b}$.

This is false, you had a DQ where you showed that the set of least square solutions to $A\boldsymbol{x} = \boldsymbol{b}$ is exactly $\hat{\boldsymbol{x}} + \mathrm{NS}(A)$, where $\hat{\boldsymbol{x}}$ is any fixed least squares solution.

There is a unique $\hat{\boldsymbol{b}}$ so that $\hat{\boldsymbol{b}}$ is the closest thing of the form $A\boldsymbol{x}$ to $\boldsymbol{b}$, in other words, $\|\hat{\boldsymbol{b}} - \boldsymbol{b}\|_2^2 = \min\{\|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 \mid \boldsymbol{x} \in \mathbb{R}^n\}$ and a least-square solution is a a solution to $A\boldsymbol{x} = \hat{\boldsymbol{b}}$.

(b) ______ For A and $m \times n$ matrix of rank n (this is assumed for now in all of our least-square problems),

$$\hat{\boldsymbol{x}} \text{ is a least squares solution to } A\boldsymbol{x} = \boldsymbol{b} \text{ iff } A\hat{\boldsymbol{x}} = \hat{b}$$

where $\hat{\boldsymbol{b}} = A(A^T A)^{-1} A^T \boldsymbol{b}$ is the unique vector satisfying $\|\boldsymbol{b} - \hat{\boldsymbol{b}}\|^2 = \min\{\|A\boldsymbol{x} - \boldsymbol{b}\|^2 \mid \boldsymbol{x} \in \mathbb{R}^n\}$.

This is true.

This is an "if and only if" ($\iff$) argument. Here one we can actually do this as a series of equivalences:

By the definition of "least-square" solution:

$$\hat{\boldsymbol{x}} \text{ is a least-squares solution to } A\boldsymbol{x} = \boldsymbol{b} \iff A^T A\hat{\boldsymbol{x}} = A^T \boldsymbol{b}$$

Multiply both sides by $(A^T A)^{-1}$ and use (‡) with $\mathrm{NS}(A^T A) = \{\boldsymbol{0}\}$.

$$\iff \hat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b}$$

Multiply both sides by A and use (‡) with $\mathrm{NS}(A) = \{\boldsymbol{0}\}$.

$$\iff A\hat{\boldsymbol{x}} = A(A^T A) A^T \boldsymbol{b} = \hat{\boldsymbol{b}}$$

Note that

$$\boldsymbol{x} = \boldsymbol{y} \iff C\boldsymbol{x} = C\boldsymbol{y} \tag{‡}$$

is only true for all $\boldsymbol{x}$ and $\boldsymbol{y}$ if $\mathrm{NS}(C) = \{0\}$.

(c) ______ If $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ is an orthonormal basis for V with respect to an inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ and $\boldsymbol{v} = \sum_{i=1}^n \alpha_i \boldsymbol{u}_i$, then $\|\boldsymbol{v}\|_2^2 = \sum_{i=1}^n |\alpha_i|^2$.

This is true and is essentially the Pythagorean Theorem. This is a computation

$$\begin{aligned}
\|\boldsymbol{v}\|_2^2 = \langle \boldsymbol{v}, \boldsymbol{v} \rangle &= \left\langle \sum_{i=1}^n \alpha_i \boldsymbol{u}_i, \sum_{i=1}^n \alpha_i \boldsymbol{u}_i \right\rangle \\
&= \sum_{i=1}^n \alpha_i \left\langle \boldsymbol{u}_i, \sum_{i=1}^n \alpha_i \boldsymbol{u}_i \right\rangle \\
&= \sum_{i=1}^n \alpha_i \left(\sum_{j=1}^n \bar{\alpha}_j \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle \right) \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \bar{\alpha}_j \delta_{i,j} \\
&= \sum_{i=1}^n \alpha_i \bar{\alpha}_i = \sum_{i=1}^n |\alpha_i|^2
\end{aligned}$$

Here

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

(d) ______ All norms $\|\cdot\| : \mathbb{R}^n \to [0, \infty)$ on $\mathbb{R}^n$ come from an inner product by $\|\boldsymbol{x}\|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle$.

This is false. The book provides several norms. For a norm $\|\cdots\|$ to be given by an inner product it must satisfy the parallelogram law $\|\boldsymbol{u} - \boldsymbol{v}\|^2 + \|\boldsymbol{u} + \boldsymbol{v}\|^2 = 2\|\boldsymbol{u}\|^2 + 2\|\boldsymbol{v}\|^2$.

Of all of the norms $\|\cdot\|_p$ for $1 \leq p \leq \infty$, the only one that satisfies the parallelogram law is $p = 2$, this is the only one given by an inner product.

For example, $\|(a, b)\|_\infty = \max\{|a|, |b|\}$ and clearly we can choose a, b, c, and d so that

$$\max\{|a - c|, |b - d|\} + \max\{|a + c|, |b + d|\} \neq 2\max\{|a|, |b|\} + 2\max\{|c|, |d|\}$$

Let $(a, b) = (1, 3)$ and $(c, d) = (2, 1)$, then

$$\begin{aligned}
\max\{|1 - 2|, |3 - 1|\} + &\max\{|1 + 2|, |3 + 1|\} = 2 + 4 \\
&\neq 2\max\{|1|, |3|\} + 2\max\{|2|, |1|\} = 6 + 4
\end{aligned}$$

(e) ______ If $\mathcal{C} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ is an orthonormal basis for V with respect to an inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ and $\boldsymbol{v} \in V$, then for any $(c_1, \ldots, c_n) = [\boldsymbol{v}]_{\mathcal{C}}$, $c_i = \langle v, u_i \rangle$.

This is another computation. Say $(c_1, \ldots, c_n) = [\boldsymbol{v}]_{\mathcal{C}}$, then $\boldsymbol{v} = \sum_{i=1}^n c_i \boldsymbol{u}_i$. Now just compute

$$\langle \boldsymbol{v}, \boldsymbol{u}_j \rangle = \left\langle \sum_{i=1}^n c_i \boldsymbol{u}_i, \boldsymbol{u}_j \right\rangle = \sum_{i=1}^n c_i \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \sum_{i=1}^n c_i \delta_{i,j} = c_j$$

Problem 2 (10 points)**.** Using the inner product

$$\langle p, q\rangle = \int_0^1 pq\,dx$$

use Gram-Schmidt to find an orthonormal basis for $\mathbb{P}_2[x]$, the space of all polynomials of degree 2 or less.

Use this to find the projection, q, of $p = x^{2/3}$ onto $\mathbb{P}_2[x]$.

Note q is the "closest point in $\mathbb{P}_2[x]$ to p in the sense that $||p-q||_2$ is as small as possible.

The strategy here is simple:

- Start with columns of $\mathcal{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\} = \{1, x, x^2\}$.
- $\boldsymbol{u}_1 = v_1$
- $\boldsymbol{q}_1 = \boldsymbol{u}_1/||u_1||$
- $\boldsymbol{u}_2 = \boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{q}_1\rangle\, \boldsymbol{q}_1$
- $\boldsymbol{q}_2 = \boldsymbol{u}_2/||\boldsymbol{u}_2||$
- $\boldsymbol{u}_3 = \boldsymbol{v}_3 - \langle \boldsymbol{v}_3, \boldsymbol{q}_1\rangle\, \boldsymbol{q}_1 - \langle \boldsymbol{v}_3, \boldsymbol{q}_2\rangle\, \boldsymbol{q}_2$
- $\boldsymbol{q}_3 = \boldsymbol{u}_3/||\boldsymbol{u}_3||$

That is the idea, here are the details. This is how all simple Gram-Schmidt proceeds.

Choose $\boldsymbol{u}_1 = 1$, then $\langle 1, 1\rangle = \int_0^1 1\,dx = 1$ so this is already normalized and so set

$$\boldsymbol{q}_1 = \boldsymbol{u}_1.$$

Set $\boldsymbol{u}_2 = x - \langle x, \boldsymbol{q}_1\rangle \boldsymbol{q}_1 = x - \int_0^1 x\,dx = x - \frac{1}{2}$. Now $||\boldsymbol{u}_2||^2 = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \int_0^1 x^2 - x + \frac{1}{4}\,dx = \frac{1}{3}x^3 - \frac{1}{2}x^2 + \frac{1}{4}x\Big|_{x=0}^{x=1} = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12}$. So

$$\boldsymbol{q}_2 = \sqrt{12}\left(x - \frac{1}{2}\right) = \sqrt{3}(2x-1).$$

Finally, $\boldsymbol{u}_3 = x^2 - \langle x^2, \boldsymbol{q}_2\rangle \boldsymbol{q}_2 - \langle x^2, \boldsymbol{q}_1\rangle \boldsymbol{q}_1$. We have $\langle x^2, \boldsymbol{q}_2\rangle = \int_0^1 \sqrt{3}(2x-1)\,x^2\,dx = \sqrt{3}\left(\frac{1}{2}x^4 - \frac{1}{3}x^3\right)\Big|_{x=0}^{x=1} = \frac{1}{2\sqrt{3}}$. So $\langle x^2, \boldsymbol{q}_2\rangle\, \boldsymbol{q}_2 = \left(x - \frac{1}{2}\right)$. Also, $\langle x^2, \boldsymbol{q}_1\rangle = \int_0^1 x^2\,dx = \frac{1}{3}$, so $\boldsymbol{u}_3 = x^2 - \left(x - \frac{1}{2}\right) - \frac{1}{3} = x^2 - x + \frac{1}{6}$.

We have $||\boldsymbol{u}_3||^2 = \int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 dx = \int_0^1 x^4 - 2x^3 + \frac{4}{3}x^2 - \frac{1}{3}x + \frac{1}{36}\,dx = \frac{1}{180}$ and so

$$\boldsymbol{q}_3 = \sqrt{180}\left(x^2 - x + \frac{1}{6}\right) = \sqrt{5}(6x^2 - 6x + 1).$$

The projection of p onto $\mathbb{P}_2[x]$ is

$$q = \langle p, q_1\rangle\, \boldsymbol{q}_1 + \langle p, \boldsymbol{q}_2\rangle\, \boldsymbol{q}_2 + \langle p, \boldsymbol{q}_3\rangle\, \boldsymbol{q}_3$$

$$
\begin{aligned}
\langle \boldsymbol{p}, \boldsymbol{q}_1 \rangle &= \int_0^1 (x^{2/3})(1)\, dx = \frac{3}{5} x^{5/3} \big|_0^1 = \frac{3}{5} \\
\langle \boldsymbol{p}, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 &= \frac{5}{3} \cdot 1 = \frac{3}{5} \\
\langle \boldsymbol{p}, \boldsymbol{q}_2 \rangle &= \int_0^1 (x^{2/3})(\sqrt{3})(2x-1)\sqrt{3} \int_0^1 (2x^{5/3} - x^{2/3}) = \sqrt{3}\left(2 \cdot \frac{3}{8} x^{8/3} - \frac{3}{5} x^{5/3}\right)_{x=0}^{x=1} \\
&= \sqrt{3}\left(\frac{6}{8} - \frac{3}{5}\right) = \sqrt{3} \cdot \frac{3}{20} \\
\langle \boldsymbol{p}, \boldsymbol{q}_2 \rangle \boldsymbol{q}_2 &= \sqrt{3}\frac{3}{20}\sqrt{3}(2x-1) = \frac{9}{20}(2x-1) \\
\langle p, q_3 \rangle &= \int_0^1 (x^{2/3})\sqrt{5}(6x^2 - 6x + 1) = \sqrt{5} \int_0^1 (6x^{8/3} - 6x^{5/3} + x^{2/3})\, dx \\
&= \sqrt{5}\left(6 \cdot \frac{3}{11} x^{11/3} - 6 \cdot \frac{3}{8} x^{8/3} + \frac{3}{5} x^{5/3}\right)_{x=0}^{x=1} = \sqrt{5}\left(\frac{18}{11} - \frac{9}{4} + \frac{3}{5}\right) = -\sqrt{5} \cdot \frac{3}{220} \\
\langle \boldsymbol{p}, \boldsymbol{q}_3 \rangle \boldsymbol{q}_3 &= -\sqrt{5} \cdot \frac{3}{220} \sqrt{5}(6x^2 - 6x + 1) = -\frac{3}{44}(6x^2 - 6x + 1)
\end{aligned}
$$

So the projection of $\boldsymbol{p}$ onto $\mathbb{P}_2[x]$ is

$$
-\frac{3}{44}(6x^2 - 6x + 1) + \frac{9}{20}(2x - 1) + \frac{3}{5} = -\frac{9}{22}x^2 + \frac{72}{55}x + \frac{9}{110} = \boxed{-\frac{9}{110}(5x^2 - 16x - 1)}
$$

A SageCell page that does computations

Problem 3 (10 points)**.** Submit your Linear Algebra Tutorial MATLAB Certificate to the shared MATLAB drive.