

## 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  $\hat{x} = (A^T A)^{-1} A^T \mathbf{b}$  to  $A\mathbf{x} = \mathbf{b}$ .

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

In general, there is no reason for  $A^T A$  to be invertible, so  $(A^T A)^{-1}$  need not even exist. If it does exist, then  $\hat{x} = (A^T A)^{-1} A^T \mathbf{b}$  is unique.

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

- (b) \_\_\_\_\_ For  $A$  an  $m \times n$  matrix:

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

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

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:

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

Multiply both sides by  $(A^T A)^{-1}$  and use  $(\ddagger)$  with  $\text{NS}(A^T A) = \{\mathbf{0}\}$  since  $(A^T A)^{-1}$  exists.

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

Multiply both sides by  $A$  and use  $(\ddagger)$  with  $\text{NS}(A) = \text{NS}(A^T A) = \{\mathbf{0}\}$ .

$$\begin{aligned} \hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} &\iff A\hat{\mathbf{x}} = A(A^T A)^{-1} A^T \mathbf{b} \\ &\iff \boxed{A\hat{\mathbf{x}} = \hat{\mathbf{b}}} \end{aligned}$$

Note that

$$\mathbf{x} = \mathbf{y} \iff C\mathbf{x} = C\mathbf{y}. \tag{\ddagger}$$

for all  $\mathbf{x}$  and  $\mathbf{y}$  iff  $\text{NS}(C) = \{\mathbf{0}\}$ .

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

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

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

Here

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

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

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

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

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

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

**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  $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{1, x, x^2\}$ .
- $\mathbf{u}_1 = \mathbf{v}_1$
- $\mathbf{q}_1 = \mathbf{u}_1 / \|\mathbf{u}_1\|$
- $\mathbf{u}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{q}_1 \rangle \mathbf{q}_1$
- $\mathbf{q}_2 = \mathbf{u}_2 / \|\mathbf{u}_2\|$
- $\mathbf{u}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{q}_1 \rangle \mathbf{q}_1 - \langle \mathbf{v}_3, \mathbf{q}_2 \rangle \mathbf{q}_2$
- $\mathbf{q}_3 = \mathbf{u}_3 / \|\mathbf{u}_3\|$

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

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

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

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

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

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

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

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

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

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

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

So the projection of  $\mathbf{p}$  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.