Assignment 8 - Solutions

Introduction to Linear Algebra

Material from Lectures 14 and 15

Due Thursday, March 9, 2023

14.2 (✠2.13) Let $A \in \mathcal{M}_{m\times n}$.

(a) Suppose that $\mathbf{x} \in \text{null}(A)$. Show that $\mathbf{x} \in \text{null}(A^T A)$ as well

(b) Suppose that $\mathbf{y} \in \text{null}(A^T A)$. Show that $\mathbf{y} \in \text{null}(A)$ as well.

(c) The above two points imply that $\text{null}(A) = \text{null}(A^T A)$. In the case that the columns of A are linearly independent, use this fact to show that $A^T A$ has full rank.

(a) Recal the nullspace of A is all the vectors $\mathbf{x}$ for which $A\mathbf{x} = 0$. Suppose that $\mathbf{x} \in \text{null}(A)$, which means that $A\mathbf{x} = 0$. Multiplying this equation by A^T on both sides gives $A^T A\mathbf{x} = 0$, which means $\mathbf{x} \in \text{null}(A^T A)$.

(b) Suppose that $\mathbf{y} \in \text{null}(A^T A)$. That is, $A^T A\mathbf{y} = 0$, and multiplying by $\mathbf{y}^T$ on both sides gives

$$0 = \mathbf{y}^T(A^T A\mathbf{y}) = (\mathbf{y}^T A^T)(A\mathbf{y}) = (A\mathbf{y})^T(A\mathbf{y}) = (A\mathbf{y}) \bullet (A\mathbf{y}) = \|A\mathbf{y}\|^2.$$

Hence $\|A\mathbf{y}\| = 0$ and since the norm is positive definite, it follows that $A\mathbf{y} = 0$. That is, $\mathbf{y} \in \text{null}(A)$.

(c) Since the columns of A are linearly independent, it follows that $\text{null}(A) = \{0\}$, so $\dim(\text{null}(A)) = 0$. By the two points above, also $\dim(\text{null}(A^T A)) = 0$. Since $A^T A \in \mathcal{M}_{n\times n}$, by the rank-nullity theorem we have that

$$\begin{aligned} \text{rank}(A^T A) + \dim(\text{null}(A^T A)) = n \quad &\iff \quad \text{rank}(A^T A) + 0 = n \\ &\iff \quad \text{rank}(A^T A) = n. \end{aligned}$$

Since A is an $m \times n$ matrix, $A^T A$ is an $n \times n$ matrix. The matrix $A^T A$ having rank n means it has full rank.

14.3 (✠2.14) Consider the two points $p_1 = (1, 0)$, $p_2 = (2, 1)$ in $\mathbf{R}^2$.

(a) Let $m \in \mathbf{R}$. Find a point $p_3 \in \mathbf{R}^2$ so that the least squares approximation to $\{p_1, p_2, p_3\}$ has slope m.

(b) Let $a \geqslant 3 \in \mathbf{R}$. Find points $p_3, p_4 \in \mathbf{R}^2$ so that the degree 2 least squares approximation to $\{p_1, p_2, p_3, p_4\}$ has its vertex on the line $x = a$.

(a) We begin with $p_3 = (3, y)$, fixing p_3 already on the line $x = 3$. This is not necessary, but we are limiting the choices so that the formulas are simpler. If we knew what y was, the least squares solution would find the solution to

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ y \end{bmatrix},$$

which, by Theorem 14.7 in the lecture notes, would be the same as solving

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ y \end{bmatrix} \qquad \Longleftrightarrow \qquad \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 1 + y \\ 2 + 3y \end{bmatrix}.$$

The solution to this equation is found by row reduction of the augmented matrix

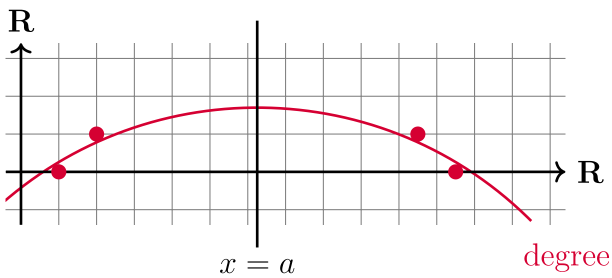
$$\begin{bmatrix} 3 & 6 & 1 + y \\ 6 & 14 & 2 + 3y \end{bmatrix} \xrightarrow{G.E.} \begin{bmatrix} 1 & 0 & \frac{1-2y}{3} \\ 0 & 1 & \frac{y}{2} \end{bmatrix},$$

meaning the optimal solution is the function $ax + b$, with $a = \frac{1-2y}{3}$ and $b = \frac{y}{2}$. We already know that the solution has slope m, so

$$m = \frac{1 - 2y}{3} \quad \Longleftrightarrow \quad 3m = 1 - 2y \quad \Longleftrightarrow \quad y = \frac{1 - 3m}{2}.$$

Therefore, the set $\left\{(1, 0), (2, 1), \left(2, \frac{1-3m}{2}\right)\right\}$ has least squares solution with slope m.

(b) Many solutions are possible, a geometric solution is presented here. Adding every new point to the set which the least squares solution will approximate pulls the approximation closer to that new point. Since we are given the condition that the vertex must be on the vertical line $x = a$, it is reasonable to choose two points that are the same as p_1, p_2, but reflected across the vertical line $x = a$, so everything is balanced and symmetric around $x = a$.



degree 2 least squares solution

So $p_3 = (2a - 2, 1)$ and $p_4 = (2a - 1, 0)$ will work. We check to make sure the least squares does indeed have its vertex on $x = a$. Here we solve the equation $A\mathbf{x} = \mathbf{b}$ as

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 2a - 2 & (2a - 2)^2 \\ 1 & 2a - 1 & (2a - 1)^2 \end{bmatrix} \begin{bmatrix} d \\ c \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix},$$

which used the Vandermonde matrix for a degree 2 approximation bx^2+cx+d. This equation does not have a solution, so instead we solve $A^TA\mathbf{x} = A^T\mathbf{b}$, which has solution

$$\mathbf{x} = \begin{bmatrix} d \\ c \\ b \end{bmatrix} = \begin{bmatrix} (16a^4+4a^2-24a+3)/(32a^4-4a^2-20a-9) \\ (-4a^2+6a)/(32a^4-4a^2-20a-9) \\ (2a-3)/(32a^4-4a^2-20a-9) \end{bmatrix}.$$

While this does look quite unappealing, all we are intereseted in is the x-value of the vertex. For a quadratic function bx^2+cx+d, the vertex is in between the roots, and hs x-coordinate $-\frac{c}{2b}$. For this situation, since the denominators are the same, that becomes

$$\frac{-c}{2b} = \frac{4a^2-6a}{2(2a-3)} = \frac{4a^2-6a}{4a-6} = \frac{(4a-6)a}{4a-6} = a,$$

as expected.

15.2 (✠2.15) Apply the Gram-Schmidt process to the vectors $\begin{bmatrix}0\\1\\1\\1\end{bmatrix}, \begin{bmatrix}2\\0\\2\\0\end{bmatrix}, \begin{bmatrix}2\\1\\1\\2\end{bmatrix}, \begin{bmatrix}2\\0\\0\\1\end{bmatrix}, \begin{bmatrix}1\\2\\0\\1\end{bmatrix}$.

Step 1: Set $\mathbf{w}_1 = \mathbf{v}_1 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix}$.

Step 2: Project $\mathbf{v}_2$ onto $\mathbf{w}_1$, and subtract this from $\mathbf{v}_2$ to ensure the new vector will be orthogonal to the previous vector. That is, set $\mathbf{w}_2$ to be the error vector when projecting to $\mathbf{w}_1$. So we get

$$\mathbf{w}_2 = \mathbf{v}_2 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_2) = \mathbf{v}_2 - \frac{\mathbf{w}_1^T\mathbf{v}_2}{\mathbf{w}_1^T\mathbf{w}_1}\mathbf{w}_1 = \begin{bmatrix}2\\0\\2\\0\end{bmatrix} - \frac{2}{3}\begin{bmatrix}0\\1\\1\\1\end{bmatrix} = \frac{1}{3}\begin{bmatrix}6\\-2\\4\\-2\end{bmatrix}.$$

Step 3: Project $\mathbf{v}_3$ onto $\mathbf{w}_1$ and $\mathbf{w}_2$, and subtract these from $\mathbf{v}_3$ to make sure everything is still orthogonal. The formula is

$$\mathbf{w}_3 = \mathbf{v}_3 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{w}_2}(\mathbf{v}_3) = \mathbf{v}_3 - \frac{\mathbf{w}_1^T\mathbf{v}_3}{\mathbf{w}_1^T\mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{w}_2^T\mathbf{v}_3}{\mathbf{w}_2^T\mathbf{w}_2}\mathbf{w}_2 = \begin{bmatrix}1\\0\\-1\\1\end{bmatrix}.$$

Step 4: Repeat the same for $\mathbf{v}_4$ to get

$$\mathbf{w}_4 = \mathbf{v}_4 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{w}_2}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{w}_3}(\mathbf{v}_4) = \mathbf{v}_4 - \frac{\mathbf{w}_1^T\mathbf{v}_4}{\mathbf{w}_1^T\mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{w}_2^T\mathbf{v}_4}{\mathbf{w}_2^T\mathbf{w}_2}\mathbf{w}_2 - \frac{\mathbf{w}_3^T\mathbf{v}_4}{\mathbf{w}_3^T\mathbf{w}_3}\mathbf{w}_3 = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}.$$

This means that $\mathbf{v}_4$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, and that we do not have to subtract the projection of $\mathbf{v}_5$ onto $\mathbf{w}_4$. Or rather, that the projection will be 0.

Step 5: Repeat the same for $\mathbf{v}_5$ to get

$$\mathbf{w}_5 = \mathbf{v}_5 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_5) - \mathrm{proj}_{\mathbf{w}_2}(\mathbf{v}_5) - \mathrm{proj}_{\mathbf{w}_3}(\mathbf{v}_5) = \frac{1}{3}\begin{bmatrix}1\\3\\-1\\-2\end{bmatrix}.$$

Normalization of these vectors to get the final output of the Gram-Schmidt process is straightforward:

$$\mathbf{q}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}0\\1\\1\\1\end{bmatrix}, \qquad \mathbf{q}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \frac{1}{2\sqrt{15}}\begin{bmatrix}6\\-2\\4\\-2\end{bmatrix},$$

$$\mathbf{q}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\0\\-1\\1\end{bmatrix}, \qquad \mathbf{q}_5 = \frac{\mathbf{w}_5}{\|\mathbf{w}_5\|} = \frac{1}{\sqrt{15}}\begin{bmatrix}1\\3\\-1\\-2\end{bmatrix},$$

and $\mathbf{q}_4 = 0$.

15.4 (✠2.16) Let V be the vector space of polynomials of degree at most 3 with domain $[0, 1]$. The "dot product" on V is defined by $f \bullet g = \int_0^1 f(x)g(x)\,dx$, which helps to define length and angle. You may assume that $\{1, x, x^2, x^3\}$ is a basis for V.

(a) Is the given basis orthogonal? Find the lengths of the elements in the basis.

(b) Are the two functions $2x, x^2 - 1$ linearly independent in V? Are they orthogonal?

(c) Extend $\{2x, x^2 - 1\}$ to an orthonormal basis of V.

(a) The basis is not orthogonal, as we can quickly compute the dot product on pairs:

$$1 \bullet x = \int_0^1 x\,dx = \frac{1}{2} \qquad 1 \bullet x^2 = \int_0^1 x^2\,dx = \frac{1}{3}$$
$$1 \bullet x^3 = \int_0^1 x^3\,dx = \frac{1}{4} \qquad x \bullet x^2 = \int_0^1 x^3\,dx = \frac{1}{4}$$
$$x \bullet x^3 = \int_0^1 x^4\,dx = \frac{1}{5} \qquad x^2 \bullet x^3 = \int_0^1 x^5\,dx = \frac{1}{6}$$

None of these are 0, so the set is not orthogonal. In a similar manner we find

$$\|1\| = 1, \qquad \|x\| = \frac{1}{\sqrt{2}}, \qquad \|x^2\| = \frac{1}{\sqrt{5}}, \qquad \|x^3\| = \frac{1}{\sqrt{7}}.$$

(b) If the set $\{2x, x^2 - 1\}$ was linearly dependent, there would be some $c \in \mathbf{R}$ with $2x = c(x^2 - 1)$ for all x. This is not possible, as c must be $\frac{2x}{x^2-1}$, but that is not a constant number for all $x \in [0, 1]$. These are not orthogonal, as

$$(2x) \bullet (x^2 - 1) = \int_0^1 2x^3 - 2x\,dx = -\frac{1}{2} \neq 0.$$

(c) We know $\{1, x, x^2, x^3\}$ is a basis of V, so we apply the Gram-Schmidt process to $\{2x, x^2 - 1, 1, x, x^2, x^3\}$. We go step by step:

$$\begin{aligned}
f_1 &= 2x,\\
f_2 &= (x^2 - 1) - \mathrm{proj}_{f_1}(x^2 - 1) = x^2 + \frac{3}{4}x - 1,\\
f_3 &= 1 - \mathrm{proj}_{f_1}(1) - \mathrm{proj}_{f_2}(1) = \frac{70}{83}x^2 - \frac{72}{83}x + \frac{13}{83},\\
f_4 &= x - \mathrm{proj}_{f_1}(x) - \mathrm{proj}_{f_2}(x) - \mathrm{proj}_{f_3}(x) = 0,\\
f_5 &= x^2 - \mathrm{proj}_{f_1}(x^2) - \mathrm{proj}_{f_2}(x^2) - \mathrm{proj}_{f_3}(x^2) = 0,\\
f_6 &= x^3 - \mathrm{proj}_{f_1}(x^3) - \mathrm{proj}_{f_2}(x^3) - \mathrm{proj}_{f_3}(x^3) = x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}.
\end{aligned}$$

Hence $\{2x, x^2 - 1\}$ extended to an orthonormal basis is

$$\left\{2x, \quad x^2 + \frac{3}{4}x - 1, \quad \frac{70}{83}x^2 - \frac{72}{83}x + \frac{13}{83}, \quad x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}\right\}.$$