Math 54: Worksheet #18, Solutions

Name: ______________________________ Date: November 4, 2021

Fall 2021

Problem 1 (True/False). If $\underline{b}$ is in the column space of A, then every solution of $A\underline{x} = \underline{b}$ is a least-squares solution.

Solution. **True.** If $\underline{b}$ is in the column space of A, then $A\underline{x} = \underline{b}$ is consistent. Every solution $\hat{\underline{x}}$ of $A\underline{x} = \underline{b}$ then has $\|A\hat{\underline{x}} - \underline{b}\| = 0$, which clearly minimizes $\|A\underline{x} - \underline{b}\|$. Thus, it is a least-squares solution.

Another way to see this is that if $\underline{b}$ is the column space of A, then $\underline{b} = \operatorname{proj}_{\operatorname{Col} A} \underline{b}$, so any solution of $A\underline{x} = \underline{b}$ satisfies $A\underline{x} = \operatorname{proj}_{\operatorname{Col} A} \underline{b}$, meaning that it is a least-squares solution.

Problem 2 (True/False). The least-squares solution of $A\underline{x} = \underline{b}$ is the point in the column space of A closest to $\underline{b}$.

Solution. **False.** The least-squares solution of $A\underline{x} = \underline{b}$ is a point $\hat{\underline{x}}$ such that $A\hat{\underline{x}} = \operatorname{proj}_{\operatorname{Col} A} \underline{b}$, the point in the column space of A closest to $\underline{b}$. So the least-squares solution is simply a vector of weights for the linear combination of the columns of A that makes $\operatorname{proj}_{\operatorname{Col} A} \underline{b}$.

Problem 3 (True/False). The function $\langle f, g\rangle = f(0)^2 + f(1)^2 + g(0)^2 + g(1)^2$ is an inner product on the vector space $V = \mathbb{P}_2$, the space of polynomials of degree at most 2.

Solution. **False.** This function is not bilinear. Consider $f(x) = 1$ and $g(x) = 0$. Then,

$$\begin{aligned} \langle f, g\rangle &= 1^2 + 1^2 + 0^2 + 0^2 = 2, \\ \langle 2f, g\rangle &= 2^2 + 2^2 + 0^2 + 0^2 = 8 \neq 2\langle f, g\rangle. \end{aligned}$$

Problem 4 (True/False). The function $\langle f, g\rangle = f(-2)g(-2) + f(0)g(0) + f(2)g(2)$ is an inner product on the vector space $V = \mathbb{P}_2$, the space of polynomials of degree at most 2.

Solution. **True.** This function is clearly bilinear. Indeed, $\langle f, g\rangle = \langle g, f\rangle$ since multiplication commutes. Also,

$$\begin{aligned} \langle f + h, g\rangle &= (f(-2) + h(-2))g(-2) + (f(0) + h(0))g(0) + (f(2) + h(2))g(2) \\ &= (f(-2)g(-2) + f(0)g(0) + f(2)g(2)) + (h(-2)g(-2) + h(0)g(0) + h(2)g(2)) = \langle f, g\rangle + \langle h, g\rangle, \\ \langle cf, g\rangle &= (cf(-2))g(-2) + (cf(0))g(0) + (cf(2))g(2) = c(f(-2)g(-2) + f(0)g(0) + f(2)g(2)) = c\langle f, g\rangle. \end{aligned}$$

Finally, we have that

$$\langle f, f\rangle = f(-2)^2 + f(0)^2 + f(2)^2 \geq 0.$$

Also, $\langle f, f\rangle = 0$ if and only if $f(-2) = f(0) = f(2) = 0$. This means that f has at least 3 distinct roots. Since f is a polynomial of degree at most 2, this means that f must be the zero polynomial, $f = 0$.

Problem 5 (6.5 #10)**.** Consider the following matrix A and vector $\underline{b}$:

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \\ 1 & 2 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}.$$

(a) Find the orthogonal projection of $\underline{b}$ onto $\operatorname{Col} A$.

(b) Find a least squares solution of $A\underline{x} = \underline{b}$.

Solution. (a) To find the orthogonal projection of $\underline{b}$ onto $\operatorname{Col} A$, we first want an orthogonal basis of $\operatorname{Col} A$. A quick check verifies that the columns of A are orthogonal, so they themselves form this orthogonal basis. Then, we can find the orthogonal projection:

$$\operatorname{proj}_{\operatorname{Col} A} \underline{b} = \frac{\underline{b} \cdot \underline{a}_1}{\underline{a}_1 \cdot \underline{a}_1}\underline{a}_1 + \frac{\underline{b} \cdot \underline{a}_2}{\underline{a}_2 \cdot \underline{a}_2}\underline{a}_2 = \frac{3+1+5}{1+1+1}\underline{a}_1 + \frac{6-4+10}{4+16+4}\underline{a}_2 = 3\underline{a}_1 + \frac{1}{2}\underline{a}_2 = 3\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 4 \end{bmatrix}.$$

(b) Finding a least squares solution of $A\underline{x} = \underline{b}$ amounts to solving $A\hat{\underline{x}} = \operatorname{proj}_{\operatorname{Col} A} \underline{b}$. From above, we see that $\operatorname{proj}_{\operatorname{Col} A} \underline{b} = 3\underline{a}_1 + \frac{1}{2}\underline{a}_2$, which means that $\hat{\underline{x}} = \begin{bmatrix} 3 \\ 1/2 \end{bmatrix}$ is a solution to $A\hat{\underline{x}} = \operatorname{proj}_{\operatorname{Col} A} \underline{b}$.

Problem 6 (6.6 #4)**.** Find the line of best fit, $y = \beta_0 + \beta_1 x$, which minimizes the square of the difference in y-values for the following data points:

$$(2,3),\ (3,2),\ (5,1)\ (6,0).$$

Solution. We want to find a least-squares solution fo $A\underline{\beta} = \underline{b}$, where

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix}.$$

We solve using the normal equations $A^T A\underline{\beta} = A^T\underline{b}$:

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 4 & 16 \\ 16 & 74 \end{bmatrix},$$

$$A^T\underline{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 5 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 17 \end{bmatrix},$$

$$\left[\begin{array}{cc|c} 4 & 16 & 6 \\ 16 & 74 & 17 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 4 & 16 & 6 \\ 0 & 10 & -7 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 2 & 8 & 3 \\ 0 & 10 & -7 \end{array}\right].$$

Solving for β_0, β_1, we get:

$$\beta_1 = -7/10 = -0.7,$$
$$2\beta_0 = 3 - 8\beta_1 = 3 + 5.6 = 8.6 \Rightarrow \beta_0 = 4.3.$$

Thus, we get the line $y = 4.3 - 0.7x$.

Problem 7 (6.7 #9)**.** Let $\mathbb{P}_3$ have the inner product given by evaluation at -3, -1, 1, and 3: $\langle f, g\rangle = f(-3)g(-3) + f(-1)g(-1) + f(1)g(1) + f(3)g(3)$. Let $p_0(t) = 1$, $p_1(t) = t$, and $p_2(t) = t^2$.

(a) Compute the orthogonal projection of p_2 onto the subspace spanned by p_0 and p_1.

(b) Find a polynomial q that is orthogonal to p_0 and p_1 such that $\{p_0, p_1, q\}$ is orthogonal basis for $\operatorname{span}\{p_0, p_1, p_2\} = \mathbb{P}_2$. Scale the polynomial q so that its vector of values at $(-3, -1, 1, 3)$ is $(1, -1, -1, 1)$.

Solution. First, we notice that the inner product only depends on the values of the polynomials at the points -3, -1, 1, and 3. Thus, we list each polynomial as a vector in $\mathbb{R}^4$ that contains the value of the polynomial at those four points:

$$[p_0]_{\mathcal{B}} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \quad [p_1]_{\mathcal{B}} = \begin{bmatrix}-3\\-1\\1\\3\end{bmatrix}, \quad [p_2]_{\mathcal{B}} = \begin{bmatrix}9\\1\\1\\9\end{bmatrix}.$$

Then, the inner product between any two polynomials f and g is given by the usual dot product between these coordinate vectors.

Note: This isn't a required step, but it puts this inner product in context of the usual dot product that we see, so some students find it easier for calculation. As suggested by my notation above, these are in fact coordinate vectors, with respect to a special basis $\mathcal{B}$. The basis is formed by Lagrange polynomials, which you don't need to understand the details of. However, it is important to understand that every polynomial in $\mathbb{P}_3$ is completely determined by its value at any four points (such as -3, -1, 1, and 3), so these coordinate vectors correspond to exactly one polynomial, as expected. This forms an isomorphism between $\mathbb{P}_3$ and $\mathbb{R}^4$.

(a) Let W be the subspace spanned by p_0 and p_1. We notice that $\langle p_0, p_1\rangle = (1)(-3) + (1)(-1) + (1)(1) + (1)(3) = 0$, so p_0 and p_1 are orthogonal and form a basis of W. This means that we can use our orthogonal projection formula:

$$\begin{aligned}\operatorname{proj}_W p_2 &= \frac{\langle p_2, p_0\rangle}{\langle p_0, p_0\rangle}p_0 + \frac{\langle p_2, p_1\rangle}{\langle p_1, p_1\rangle}p_1 \\ &= \frac{(9)(1) + (1)(1) + (1)(1) + (9)(1)}{1^2 + 1^2 + 1^2 + 1^2}p_0 + \frac{(9)(-3) + (1)(-1) + (1)(1) + (9)(3)}{(-3)^2 + (-1)^2 + 1^2 + 3^2}p_1 \\ &= \frac{20}{4}p_0 + \frac{0}{20}p_1 = 5p_0 = 5.\end{aligned}$$

(b) To find an orthogonal basis $\{p_0, p_1, q\}$ for $\mathbb{P}_2 = \operatorname{span}\{p_0, p_1, p_2\}$, we want to apply Gram-Schmidt to $\{p_0, p_1, p_2\}$. We already saw that p_0 and p_1 are orthogonal, so the first two vectors will remain unchanged. Then, for the third vector, we get that

$$q = p_2 - \frac{\langle p_2, p_0\rangle}{\langle p_0, p_0\rangle}p_0 - \frac{\langle p_2, p_1\rangle}{\langle p_1, p_1\rangle}p_1 = p_2 - \operatorname{proj}_W p_2 = t^2 - 5.$$

We see that $q(1) = 1^2 - 5 = -4$, so we want to scale q by a factor of $1/4$. This gives us $q(t) = \frac{1}{4}(t^2 - 5)$. Then, the corresponding coordinate vector is:

$$[q]_{\mathcal{B}} = \begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}.$$

We can easily check that q is indeed orthogonal to p_0 and p_1, as desired.