MAT 22A Problem Set 9 Solutions

1. Find the projection of matrix P and projection $\mathbf{p}$ of $\mathbf{b} = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ onto the subspace spanned by $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right\}$. What is the length of the error vector?

   Solution.
   Let
   $$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix}.$$
   The projection matrix onto the subspace is $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$. Computing P, we find that
   $$P = \begin{bmatrix} \frac{2}{5} & -\frac{1}{5} & -\frac{1}{5} & \frac{2}{5} \\ -\frac{1}{5} & \frac{14}{15} & -\frac{1}{15} & \frac{2}{15} \\ -\frac{1}{5} & -\frac{1}{15} & \frac{14}{15} & \frac{2}{15} \\ \frac{2}{5} & \frac{2}{15} & \frac{2}{15} & \frac{11}{15} \end{bmatrix}$$
   and so
   $$\mathbf{p} = \begin{bmatrix} \frac{2}{5} & -\frac{1}{5} & -\frac{1}{5} & \frac{2}{5} \\ -\frac{1}{5} & \frac{14}{15} & -\frac{1}{15} & \frac{2}{15} \\ -\frac{1}{5} & -\frac{1}{15} & \frac{14}{15} & \frac{2}{15} \\ \frac{2}{5} & \frac{2}{15} & \frac{2}{15} & \frac{11}{15} \end{bmatrix} \begin{bmatrix} 5 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -1 \\ 2 \end{bmatrix}.$$
   The length of the error vector is
   $$\|\mathbf{e}\| = \|\mathbf{b} - A\hat{\mathbf{x}}\| = 3.87298.$$
   □

2. Prove the following properties of the projection matrix P:

   (a) $P^2 = P$

   (b) $P^{\mathrm{T}} = P$

   Solution.

(a) We know that $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$, so

$$\begin{aligned} P^2 &= A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} \\ &= A(A^{\mathrm{T}}A)^{-1}(A^{\mathrm{T}}A)(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} \\ &= A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} \\ &= P \end{aligned}$$

(b) We have

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

Now, suppose B is an invertible matrix. Then

$$\begin{aligned} BB^{-1} &= I \\ (BB^{-1})^{\mathrm{T}} &= I \\ (B^{-1})^{\mathrm{T}}B^{\mathrm{T}} &= I \end{aligned}$$

so $(B^{\mathrm{T}})^{-1} = (B^{-1})^{\mathrm{T}}$. Then $((A^{\mathrm{T}}A)^{-1})^{\mathrm{T}} = ((A^{\mathrm{T}}A)^{\mathrm{T}})^{-1}$. Now,

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

□

3. Let S_1 and S_2 be subspaces of $\mathbb{R}^4$. S_1 is spanned by $\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ and S_2 spanned by $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}$. Show that S_1 and S_2 are orthogonal subspaces. Find the orthogonal complement of $S_1 + S_2 = \operatorname{span}\left(\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\} \right)$. Let P_1 be the projection matrix onto S_1, P_2 the projection matrix onto S_2 and P_3 the projection matrix onto $(S_1+S_2)^{\perp}$. Show that $P_1+P_2+P_3 = I$.

Solution.

Let $\mathbf{x}_1 \in S_1$ and $\mathbf{x}_2 \in S_2$. Since the two given vectors that span S_1 are linealry independent, they form a basis of S_1, so we can find c_1 and c_2 so that

$$\mathbf{x}_1 = c_1 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

and we may find c_3 so that

$$\mathbf{x}_2 = c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

Now, we see that

$$\begin{aligned} \mathbf{x}_2^{\mathrm{T}}\mathbf{x}_1 &= c_3 \begin{bmatrix} 0 & 0 & 1 & 2 \end{bmatrix} \left(c_1 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right) \\ &= c_3 c_1 (0) + c_3 c_2 (0) \\ &= 0. \end{aligned}$$

Thus, we see that S_1 and S_2 are orthogonal subspaces.

Let

$$A = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

Since $C(A) = S_1 + S_2$, we see that $(S_1 + S_2)^{\perp} = (\mathrm{C}(A))^{\perp} = \mathrm{N}(A^{\mathrm{T}})$. To determine $\mathrm{N}(A^{\mathrm{T}})$, we solve

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \mathbf{0}.$$

We see that $x_1 = x_2$, $x_1 = -x_2$, and $x_3 = -2x_4$, so $x_1 = x_2 = 0$. Then, $\mathbf{x} = x_4 \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}$. Thus, $(S_1 + S_2)^{\perp} = \mathrm{span}\left(\begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right)$. Now, we compute

P_1, P_2, and P_3.

$$P_1 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \left(\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

$$P_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \left(\begin{bmatrix} 0 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right)^{-1} \begin{bmatrix} 0 & 0 & 1 & 2 \end{bmatrix}$$

$$= \frac{1}{5} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 2 \end{bmatrix}$$

$$= \frac{1}{5} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix}.$$

$$P_3 = \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \end{bmatrix} \left(\begin{bmatrix} 0 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 0 & 0 & -2 & 1 \end{bmatrix}$$

$$= \frac{1}{5} \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 1 \end{bmatrix}$$

$$= \frac{1}{5} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix}.$$

Now,

$$P_1 + P_2 + P_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \frac{1}{5} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix} + \frac{1}{5} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix} = I.$$

□

4. Consider the following data points: $(-2, -6)$, $(-1, 0)$, $(0, 0)$, $(1, 0)$, and $(2, 6)$. Interpolate the data using the following polynomials:

   (a) Line

   (b) Parabola

   (c) Cubic

   Which is the best fit to the data? Why?

   Solution.

   (a) We want a line of the form $c_0 + c_1 x = y$, so we get the system

   $$\begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \mathbf{c} = \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \\ 6 \end{bmatrix}.$$

   Multiplying the by transpose of the LHS and multiplying the inverse of the resulting matrix, we find that

   $$\mathbf{c} = \begin{bmatrix} 0 \\ \frac{24}{10} \end{bmatrix}.$$

The length of the error vector is

$$\|\mathbf{e}\| = \left\| \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \\ 6 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{24}{10} \end{bmatrix} \right\| = 3.7947.$$

(b) We want a parabola of the form $c_0 + c_1 x + c_2 x^2 = y$, so we get the system

$$\begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \mathbf{c} = \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \\ 6 \end{bmatrix}.$$

Multiplying the by transpose of the LHS and multiplying the inverse of the resulting matrix, we find that

$$\mathbf{c} = \begin{bmatrix} 0 \\ \frac{24}{10} \\ 0 \end{bmatrix}$$

The legnth of the error vector is

$$\|\mathbf{e}\| = \left\| \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \\ 6 \end{bmatrix} - \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{24}{10} \\ 0 \end{bmatrix} \right\| = 3.7947.$$

(c) We want a parabola of the form $c_1 + c_1 x + c_2 x^2 + c_3 x^3 = y$, so we get the system

$$\begin{bmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{bmatrix} \mathbf{c} = \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \\ 6 \end{bmatrix}.$$

Multiplying the by transpose of the LHS and multiplying the inverse of the resulting matrix, we find that

$$\mathbf{c} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

The legnth of the error vector is

$$\|\mathbf{e}\| = \left\| \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \\ 6 \end{bmatrix} - \begin{bmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\| = 0.$$

Of the line, parabola, and cubic function, we see that the cubic function has the smallest error, so it is likely the best choice to fit the data. □