Math 225 (Q1) Solution to Homework Assignment 5

1.

(a)

$$||\underline{u_1}|| = \sqrt{(2/3)^2 + (1/3)^2 + (2/3)^2} = 4/9 + 1/9 + 4/9 = 1,$$
$$||\underline{u_2}|| = \sqrt{(-2/3)^2 + (2/3)^2 + (1/3)^2} = 1.$$
$$\underline{u_1} \cdot \underline{u_2} = (2/3)(-2/3) + (1/3)(2/3) + (2/3)(1/3) = -4/9 + 2/9 + 2/9 = 0.$$

Therefore $\{\underline{u_1}, \underline{u_2}\}$ is an orthonormal set.

(b)

$$\underline{\hat{y}} = \text{proj}_W(\underline{y}) = (\underline{y} \cdot \underline{u_1})\underline{u_1} + (\underline{y} \cdot \underline{u_2})\underline{u_2} = (11/3)\underline{u_1} + (-23/3)\underline{u_2} = \begin{pmatrix} 68/9 \\ -35/9 \\ -1/9 \end{pmatrix}.$$

(c) Let $\underline{z} = \underline{y} - \underline{\hat{y}} = \begin{pmatrix} -23/9 \\ -46/9 \\ 46/9 \end{pmatrix}$. Then the distance from $\underline{y}$ to W is:

$$d(\underline{y}, W) = ||\underline{z}|| = ||\underline{y} - \underline{\hat{y}}|| = \sqrt{(-23/9)^2 + (-46/9)^2 + (46/9)^2} = 23/3.$$

2.

(a) Let $\underline{a_1} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$, $\underline{a_2} = \begin{pmatrix} 2 \\ 1 \\ 4 \\ -4 \\ 2 \end{pmatrix}$ and $\underline{a_3} = \begin{pmatrix} 5 \\ -4 \\ -3 \\ 7 \\ 1 \end{pmatrix}$ be the column vectors of the matrix A. Let $W = \text{Span}\{\underline{a_1}, \underline{a_2}, \underline{a_3}\} = \text{Col}(A)$. According to the Gram–Schmidt process, define

$$\underline{u_1} = \underline{a_1},$$

$$\underline{u_2} = \underline{a_2} - \frac{\underline{a_2} \cdot \underline{u_1}}{||\underline{u_1}||^2}\underline{u_1} = \underline{a_2} + \underline{u_1} = \begin{pmatrix} 3 \\ 0 \\ 3 \\ -3 \\ 3 \end{pmatrix}$$

$$\underline{u_3} = \underline{a_3} - \left[\frac{\underline{a_3} \cdot \underline{u_1}}{||\underline{u_1}||^2}\underline{u_1} + \frac{\underline{a_3} \cdot \underline{u_2}}{||\underline{u_2}||^2}\underline{u_2}\right] = \underline{a_3} - [4\underline{u_1} - \frac{1}{3}\underline{u_2}] = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 2 \\ -2 \end{pmatrix}.$$

Then $\{\underline{u_1}, \underline{u_2}, \underline{u_3}\}$ is an orthogonal basis of $\mathrm{Col}(A)$.

(b) Next, we normalize and let

$$\underline{v_1} = \frac{1}{||\underline{u_1}||}\underline{u_1} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}, \quad \underline{v_2} = \frac{1}{||\underline{u_2}||}\underline{u_2} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \underline{v_3} = \frac{1}{||\underline{u_3}||}\underline{u_3} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}.$$

Then $\{\underline{v_1}, \underline{v_2}, \underline{v_3}\}$ is an orthonormal basis of $\mathrm{Col}(A)$. Let $Q = \begin{pmatrix} 1/\sqrt{5} & 1/2 & 1/2 \\ -1/\sqrt{5} & 0 & 0 \\ -1/\sqrt{5} & 1/2 & 1/2 \\ 1/\sqrt{5} & -1/2 & 1/2 \\ 1/\sqrt{5} & 1/2 & -1/2 \end{pmatrix}$.

Then the matrix Q has orthonormal columns. From $A = QR$, we see that $Q^TA = Q^TQR = R$, since $Q^TQ = I$. Thus,

$$R = Q^TA = \begin{pmatrix} 1/\sqrt{5} & -1/\sqrt{5} & -1/\sqrt{5} & 1/\sqrt{5} & 1/\sqrt{5} \\ 1/2 & 0 & 1/2 & -1/2 & 1/2 \\ 1/2 & 0 & 1/2 & 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{5} & -\sqrt{5} & 4\sqrt{5} \\ 0 & 6 & -2 \\ 0 & 0 & 4 \end{pmatrix}$$

which is an upper triangular matrix with positive diagonal entries.

3.

(a) Let $\underline{a_1} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$ and $\underline{a_2} = \begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix}$ be the two column vectors of A. We use the Gram–Schmidt process to turn $\{\underline{a_1}, \underline{a_2}\}$ into an orthogonal basis for $\mathrm{Col}(A)$. Let

$\underline{u_1} = \underline{a_1} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$ and

$$\underline{u_2} = \underline{a_2} - \mathrm{proj}_{\mathrm{Span}\{\underline{u_1}\}}(\underline{a_2}) = \underline{a_2} - \frac{\underline{a_2} \cdot \underline{u_1}}{\underline{u_1} \cdot \underline{u_1}}\underline{u_1} = \begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix} - \frac{0}{14}\begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix}.$$

Thus, $\left\{ \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix} \right\}$ is an orthogonal basis of $\mathrm{Col}(A)$.

$$\hat{\underline{b}} = \mathrm{proj}_{\mathrm{Col}(A)}(\underline{b}) = \frac{\underline{b} \cdot \underline{u_1}}{\underline{u_1} \cdot \underline{u_1}}\underline{u_1} + \frac{\underline{b} \cdot \underline{u_2}}{\underline{u_2} \cdot \underline{u_2}}\underline{u_2} = \frac{4}{14}\begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + \frac{6}{42}\begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

(b) Least squares solution is found by solving the equation $A\underline{x} = \hat{\underline{b}}$, that is, $\begin{pmatrix} 1 & 5 \\ 3 & 1 \\ -2 & 4 \end{pmatrix} \underline{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Since the augmented coefficient matrix $\left(\begin{array}{cc|c} 1 & 5 & 1 \\ 3 & 1 & 1 \\ -2 & 4 & 0 \end{array}\right)$ has the reduced row echelon form $\left(\begin{array}{cc|c} \boxed{1} & 0 & 2/7 \\ 0 & \boxed{1} & 1/7 \\ 0 & 0 & 0 \end{array}\right)$, therefore $x_1 = 2/7$ and $x_2 = 1/7$. Thus, $\underline{x} = \begin{pmatrix} 2/7 \\ 1/7 \end{pmatrix}$ is the least squares solution.

(c) The normal equation is $A^T A\underline{x} = A^T\underline{b}$, which is, $\begin{pmatrix} 14 & 0 \\ 0 & 42 \end{pmatrix} \underline{x} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$.

(d) By solving the normal equation in (c), we see that the least squares solution is $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2/7 \\ 1/7 \end{pmatrix}$.

4. Let $\underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$ be the parameter vector. The design matrix is $X = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$. The observation vector is $\underline{y} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 4 \end{pmatrix}$. We need to find a least squares solution to the least square problem $X\underline{\beta} = \underline{y}$. $\underline{\beta}$ can be found using the normal equation $X^T X\underline{\beta} = X^T\underline{y}$. Note that

$$X^T X = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix}$$

and

$$X^T\underline{y} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 10 \end{pmatrix}.$$

Solving $\begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 7 \\ 10 \end{pmatrix}$, we get $\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 11/10 \\ 13/10 \end{pmatrix}$. Thus, the least squares line is $y = \frac{11}{10} + \frac{13}{10}x$.

5.

(a) The observation vector is $\underline{p} = \begin{pmatrix} 91 \\ 98 \\ 103 \\ 110 \\ 112 \end{pmatrix}$. The prediction vector is

$$\begin{pmatrix} \beta_0 + \beta_1 \ln(44) \\ \beta_0 + \beta_1 \ln(61) \\ \beta_0 + \beta_1 \ln(81) \\ \beta_0 + \beta_1 \ln(113) \\ \beta_0 + \beta_1 \ln(131) \end{pmatrix} = \begin{pmatrix} 1 & \ln(44) \\ 1 & \ln(61) \\ 1 & \ln(81) \\ 1 & \ln(113) \\ 1 & \ln(113) \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = X\underline{\beta},$$

where $X = \begin{pmatrix} 1 & \ln(44) \\ 1 & \ln(61) \\ 1 & \ln(81) \\ 1 & \ln(113) \\ 1 & \ln(113) \end{pmatrix} = \begin{pmatrix} 1 & 3.784189634 \\ 1 & 4.110873864 \\ 1 & 4.394449155 \\ 1 & 4.727387819 \\ 1 & 4.875197323 \end{pmatrix}$ is the design matrix and $\underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$ is the unknown parameter vector. The least square problem is $X\underline{\beta} = \underline{p}$ and the corresponding normal equation is $X^T X\underline{\beta} = X^T\underline{p}$, which is

$$\begin{pmatrix} 5 & 21.89209779 \\ 21.89209779 & 96.64630303 \end{pmatrix} \underline{\beta} = \begin{pmatrix} 514 \\ 2265.889918 \end{pmatrix}.$$

Solving this, we get $\underline{\beta} = \begin{pmatrix} 17.92435 \\ 19.384998 \end{pmatrix}$. Thus, the equation is

$$p = 17.92435 + 19.384998 \ln(w).$$

(b) For $w = 100$, the estimated systolic blood pressure is

$$p = 17.92435 + 19.384998 \ln(100) = 107.1955649.$$