NATIONAL UNIVERSITY OF SINGAPORE
Department of Mathematics

Module: MA1508E Linear Algebra for Engineering
Year/Semester: 2018-2019 (Semester 2)
Tutorial: 9

1. For each of the following linear system $\boldsymbol{Ax} = \boldsymbol{b}$,

   (i) Show that the system is inconsistent;

   (ii) Find a least squares solution $\boldsymbol{x'}$ to the system. Is there a unique least squares solution or infinitely many?

   (iii) Compute the least squares error, defined as $||\boldsymbol{b} - \boldsymbol{Ax'}||$. If there are infinitely many least squares solution and $\boldsymbol{x'_1}$, $\boldsymbol{x'_2}$ are any two of them, would the least squares error $||\boldsymbol{b} - \boldsymbol{Ax'_1}||$ and $||\boldsymbol{b} - \boldsymbol{Ax'_2}||$ be the same?

   (a) $\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}$.

   (b) $\boldsymbol{A} = \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$.

(a) (i)

$$\left(\begin{array}{cc|c} 1 & -1 & 4 \\ 3 & 2 & 1 \\ -2 & 4 & 3 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right).$$

So the system is inconsistent.

(ii) We solve $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$:

$$\left(\begin{array}{cc|c} 14 & -3 & 1 \\ -3 & 21 & 10 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 0 & \frac{17}{95} \\ 0 & 1 & \frac{143}{285} \end{array}\right).$$

So the least squares solution is $\boldsymbol{x'} = (x_1, x_2) = (\frac{17}{95}, \frac{143}{285})$, and it is unique.

(iii) The least squares error is

$$||\boldsymbol{b} - \boldsymbol{Ax'}|| = || \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} \frac{17}{95} \\ \frac{143}{285} \end{pmatrix} || = || \begin{pmatrix} \frac{1232}{285} \\ -\frac{154}{285} \\ \frac{77}{57} \end{pmatrix} || \approx 4.5611.$$

(b) (i)

$$\left(\begin{array}{ccc|c} 3 & 2 & -1 & 2 \\ 1 & -4 & 3 & -2 \\ 1 & 10 & -7 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{7} & 0 \\ 0 & 1 & -\frac{5}{7} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

So the system is inconsistent.

(ii) We solve $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$:

$$\left(\begin{array}{ccc|c} 11 & 12 & -7 & 5 \\ 12 & 120 & -84 & 22 \\ -7 & -84 & 59 & -15 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{7} & \frac{2}{7} \\ 0 & 1 & -\frac{5}{7} & \frac{13}{84} \\ 0 & 0 & 0 & 0 \end{array}\right).$$

So there are infinitely many least squares solutions. One of them is $\boldsymbol{x'} = (x_1, x_2, x_3) = (\frac{2}{7},\ \frac{13}{84},\ 0)$.

(iii) The least squares error is

$$||\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x'}|| = ||\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} - \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix}\begin{pmatrix} \frac{2}{7} \\ \frac{13}{84} \\ 0 \end{pmatrix}|| = ||\begin{pmatrix} \frac{5}{6} \\ -\frac{5}{3} \\ -\frac{5}{6} \end{pmatrix}|| \approx 2.0412.$$

The least squares errors for different least squares solution would be the same since $\boldsymbol{A}\boldsymbol{x'_1}$ and $\boldsymbol{A}\boldsymbol{x'_2}$ are equal.

2. For each of the following, compute the orthogonal projection of $\boldsymbol{u}$ onto the subspace spanned by $\boldsymbol{v_1}, \cdots, \boldsymbol{v_k}$.

   (a) $\boldsymbol{u} = (1, -6, 1)$, $\boldsymbol{v_1} = (-1, 2, 1)$, $\boldsymbol{v_2} = (2, 2, 4)$.

   (b) $\boldsymbol{u} = (6, 12, 3, 6)$, $\boldsymbol{v_1} = (1, 1, 0, 0)$, $\boldsymbol{v_2} = (1, 0, 1, 0)$, $\boldsymbol{v_3} = (3, 1, 1, 1)$.

(a) An orthogonal basis for span$\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is $\{\boldsymbol{w_1}, \boldsymbol{w_2}\}$ where

$$\begin{aligned} \boldsymbol{w_1} &= \boldsymbol{v_1} \\ \boldsymbol{w_2} &= \boldsymbol{v_2} - \frac{\boldsymbol{v_2}\cdot\boldsymbol{w_1}}{\boldsymbol{w_1}\cdot\boldsymbol{w_1}}\boldsymbol{w_1} \\ &= (2, 2, 4) - \frac{6}{6}(-1, 2, 1) = (3, 0, 3). \end{aligned}$$

So the orthogonal projection of $\boldsymbol{u}$ onto span$\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ = span$\{\boldsymbol{w_1}, \boldsymbol{w_2}\}$ is

$$\frac{\boldsymbol{u}\cdot\boldsymbol{w_1}}{\boldsymbol{w_1}\cdot\boldsymbol{w_1}}\boldsymbol{w_1} + \frac{\boldsymbol{u}\cdot\boldsymbol{w_2}}{\boldsymbol{w_2}\cdot\boldsymbol{w_2}}\boldsymbol{w_2} = -2\boldsymbol{w_1} + \frac{6}{18}\boldsymbol{w_2} = -2(-1, 2, 1) + \frac{1}{3}(3, 0, 3) = (3, -4, -1).$$

(b) Applying Gram-Schmidt Process, we obtain an orthogonal basis $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_3}\}$ for span$\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ where

$$\boldsymbol{w_1} = (1, 1, 0, 0), \quad \boldsymbol{w_2} = (\frac{1}{2},\ -\frac{1}{2},\ 1,\ 0), \quad \boldsymbol{w_3} = (\frac{1}{3},\ -\frac{1}{3},\ -\frac{1}{3},\ 1).$$

So the orthogonal projection of $\boldsymbol{u}$ onto span$\{\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_3}\}$ is

$$\frac{\boldsymbol{u}\cdot\boldsymbol{w_1}}{\boldsymbol{w_1}\cdot\boldsymbol{w_1}}\boldsymbol{w_1} + \frac{\boldsymbol{u}\cdot\boldsymbol{w_2}}{\boldsymbol{w_2}\cdot\boldsymbol{w_2}}\boldsymbol{w_2} + \frac{\boldsymbol{u}\cdot\boldsymbol{w_3}}{\boldsymbol{w_3}\cdot\boldsymbol{w_3}}\boldsymbol{w_3}$$

$$= 9(1, 1, 0, 0) + 0(\frac{1}{2},\ -\frac{1}{2},\ 1,\ 0) + \frac{9}{4}(\frac{1}{3},\ -\frac{1}{3},\ -\frac{1}{3},\ 1) = (\frac{39}{4},\ \frac{33}{4},\ -\frac{3}{4},\ \frac{9}{4}).$$

3. A series of experiments were performed to investigate the relationship between two physical quantities x and y. The results of the experiments are shown in the table below.

x	0	1	2	3
y	3	2	4	4

(a) Find a least squares solution $\boldsymbol{x} = (\hat{a}, \hat{b})$ if it is believed that x and y are related linearly, that is, $y = ax + b$.

(b) Find a least squares solution $\boldsymbol{x} = (\hat{a}, \hat{b}, \hat{c})$ if it is believed that x and y are related by the quadratic polynomial $y = ax^2 + bx + c$.

(c) Which model (linear or quadratic) would produce a smaller least squares error?

(a) We find a least squares solution to

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 4 \\ 4 \end{pmatrix}.$$

Solving $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$,

$$\left(\begin{array}{cc|c} 14 & 6 & 22 \\ 6 & 4 & 13 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{5}{2} \end{array}\right)$$

So a least squares solution is $(\hat{a}, \hat{b}) = (\frac{1}{2}, \frac{5}{2})$.

(b) We find a least squares solution to

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 4 \\ 4 \end{pmatrix}.$$

Solving $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$,

$$\left(\begin{array}{ccc|c} 98 & 36 & 14 & 54 \\ 36 & 14 & 6 & 22 \\ 14 & 6 & 4 & 13 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{11}{4} \end{array}\right).$$

So a least squares solution is $(\hat{a}, \hat{b}, \hat{c}) = (\frac{1}{4}, -\frac{1}{4}, \frac{11}{4})$.

(c) For the linear model, the least squares error is

$$|| \begin{pmatrix} 3 \\ 2 \\ 4 \\ 4 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{5}{2} \end{pmatrix} || = || \begin{pmatrix} \frac{1}{2} \\ -1 \\ \frac{1}{2} \\ 0 \end{pmatrix} || \approx 1.2247.$$

For the quadratic model, the least squares error is

$$\left|\left| \begin{pmatrix} 3 \\ 2 \\ 4 \\ 4 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ \frac{11}{4} \end{pmatrix} \right|\right| = \left|\left| \begin{pmatrix} \frac{1}{4} \\ -\frac{3}{4} \\ \frac{3}{4} \\ -\frac{1}{4} \end{pmatrix} \right|\right| \approx 1.118.$$

So the quadratic model has a smaller least squares error.

4. Prove that if $\boldsymbol{A}$ has linearly independent column vectors, and if $\boldsymbol{b}$ is orthogonal to the column space of $\boldsymbol{A}$, then the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$ is $\boldsymbol{x} = \boldsymbol{0}$.

Let $\boldsymbol{A}$ be a $m \times n$ matrix. Since $\boldsymbol{A}$ has linearly independent columns, we have $\text{rank}(\boldsymbol{A}) = n$. By dimension theorem for matrices, $\text{nullity}(\boldsymbol{A}) = 0$. From Tutorial 8, we know that the nullspace of $\boldsymbol{A}$ is equal to the nullspace of $\boldsymbol{A}^T\boldsymbol{A}$, so we know that $\text{nullity}(\boldsymbol{A}^T\boldsymbol{A}) = 0$, in other words, $\boldsymbol{A}^T\boldsymbol{A}$ (which has n columns) has rank n and thus is invertible.

Write the vectors in this question as column vectors. Let $\boldsymbol{a_1}, \boldsymbol{a_2}, \cdots, \boldsymbol{a_n}$ be the columns of $\boldsymbol{A}$. Since $\boldsymbol{b}$ is orthogonal to the column space of $\boldsymbol{A}$, we have $\boldsymbol{a_i} \cdot \boldsymbol{b} = 0$ for all $i = 1, 2, \cdots, n$. Thus $\boldsymbol{A}^T\boldsymbol{b} = \boldsymbol{0}$.

So a least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$ is a solution of $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$, which implies

$$(\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b} \Rightarrow (\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \Rightarrow \boldsymbol{x} = \boldsymbol{0}$$

since $(\boldsymbol{A}^T\boldsymbol{A})$ is invertible.

5. (**$\boldsymbol{QR}$-factorisation**) Let $\boldsymbol{A} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $\boldsymbol{u_1} = (1, 1, 1, 0)^T$, $\boldsymbol{u_2} = (-1, 0, -1, 0)^T$, $\boldsymbol{u_3} = (-1, 0, 0, -1)^T$.

   (a) Use Gram-Schmidt Process to transform $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ into an orthonormal basis $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_3}\}$ for the column space of $\boldsymbol{A}$. (Do not change the order of $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$ when applying the Gram-Schmidt Process.)

   (b) Write each of $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$ as a linear combination of $\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_3}$.

   (c) Hence or otherwise, write $\boldsymbol{A} = \boldsymbol{QR}$ where $\boldsymbol{Q}$ is a 4×3 matrix with orthonormal columns and $\boldsymbol{R}$ is a 3×3 upper triangular matrix with positive entries along its diagonal.

Remark: $\boldsymbol{QR}$-factorisation is widely used in computer algorithms for various computations concerning matrices.

(a) $\boldsymbol{w_1} = \frac{1}{\sqrt{3}}(1, 1, 1, 0)^T$.

$$\boldsymbol{w_2} = \frac{1}{\sqrt{6}}(-1, 2, -1, 0), \quad \boldsymbol{w_3} = \frac{1}{\sqrt{6}}(-1, 0, 1, -2).$$

(b)

$$\begin{aligned}
\boldsymbol{u_1} &= \sqrt{3}\boldsymbol{w_1} + 0\boldsymbol{w_2} + 0\boldsymbol{w_3} \\
\boldsymbol{u_2} &= -\frac{2}{\sqrt{3}}\boldsymbol{w_1} + \frac{2}{\sqrt{6}}\boldsymbol{w_2} + 0\boldsymbol{w_3} \\
\boldsymbol{u_3} &= -\frac{1}{\sqrt{3}}\boldsymbol{w_1} + \frac{1}{\sqrt{6}}\boldsymbol{w_2} + \frac{\sqrt{6}}{2}\boldsymbol{w_3}
\end{aligned}$$

(c) $\boldsymbol{A} = \boldsymbol{QR}$ where

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{w_1} & \boldsymbol{w_2} & \boldsymbol{w_3} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix} \quad \boldsymbol{R} = \begin{pmatrix} \sqrt{3} & -\frac{2}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{6}}{2} \end{pmatrix}.$$