3rd Seminar

Review Questions

3.11. In a linear least squares problem $Ax \cong b$, where A is an $m \times n$ matrix, if $\operatorname{rank}(A) < n$, then which of the following situations are possible?
(*a*) There is no solution.
(*b*) There is a unique solution.
(c) There is a solution, but it is not unique.

Solution:

(c) is possible since A is rank-deficient

3.12. In solving an overdetermined least squares problem $Ax \cong b$, which would be a more serious difficulty: that the rows of A are linearly dependent, or that the columns of A are linearly dependent? Explain.

Solution:

Assume A is a $m \times n$ matrix.

The number of equation is greater than the number of unknown numbers in solving an overdetermined least squares problem, which means $m > n$.

① If rows of A are linearly dependent, indicating that there exists redudant equations. Therefore, we can eliminate them by rescalling so that it will not cause any effect on the solution.

② If columns of A are linearly dependent, indicating that $\operatorname{rank}(A) < n$. In this case, the solutions exist, but they are not unique.

Thus, the 2nd case would be a more serious difficulty.

3.13. In an overdetermined linear least squares problem with model function $f(t, x) = x_1\phi_1(t) + x_2\phi_2(t) + x_3\phi_3(t)$, what will be the rank of the resulting least squares matrix A if we take $\phi_1(t) = 1$, $\phi_2(t) = t$, and $\phi_3(t) = 1 - t$?

Solution:

$f(t, x) = x_1 + x_2 t + x_3(1-t)$

Assume that $A \in \mathbb{R}^{m \times n}$ where $m > n$

since there are 3 unknowns x_1, x_2 and x_3,

assume that $n = 3$

$$\Rightarrow A = \begin{pmatrix} 1 & t_1 & 1-t_1 \\ 1 & t_2 & 1-t_2 \\ \vdots & \vdots & \vdots \\ 1 & t_m & 1-t_m \end{pmatrix} \Rightarrow \vec{v_1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \vec{v_2} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{pmatrix}, \vec{v_3} = \begin{pmatrix} 1-t_1 \\ 1-t_2 \\ \vdots \\ 1-t_m \end{pmatrix}$$

$\Rightarrow \vec{v_3} = \vec{v_1} + \vec{v_2}$ $\Rightarrow$ only two independent columns $\Rightarrow \operatorname{rank}(A) = 2$

3.17. Which of the following properties of an $m \times n$ matrix A, with $m > n$, indicate that the minimum residual solution of the least squares problem $Ax \cong b$ is *not* unique?
(*a*) The columns of A are linearly dependent.
(*b*) The rows of A are linearly dependent.
(c) The matrix $A^T A$ is singular.

Solution:

(a) Yes (b) No (c) Yes

Exercise

3.1. If a vertical beam has a downward force applied at its lower end, the amount by which it stretches will be proportional to the magnitude of the force. Thus, the total length y of the beam is given by the equation

$$y = x_1 + x_2 t,$$

where x_1 is its original length, t is the force applied, and x_2 is the proportionality constant. Suppose that the following measurements are taken:

t	10	15	20
y	11.60	11.85	12.25

(*a*) Set up the overdetermined 3×2 system of linear equations corresponding to the data collected.
(*b*) Is this system consistent? If not, compute each possible pair of values for (x_1, x_2) obtained by selecting any two of the equations from the system. Is there any reason to prefer any one of these results?
(*c*) Set up the system of normal equations and solve it to obtain the least squares solution to the overdetermined system. Compare your result with those obtained in part *b*.

Solution:

(a) The system is given below

$$\begin{pmatrix} 1 & 10 \\ 1 & 15 \\ 1 & 20 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 11.6 \\ 11.85 \\ 12.25 \end{pmatrix}$$

(b)

let $A = \begin{pmatrix} 1 & 10 \\ 1 & 15 \\ 1 & 20 \end{pmatrix}$, $\vec{y} = \begin{pmatrix} 11.6 \\ 11.85 \\ 12.25 \end{pmatrix}$

$$\Rightarrow [A \ \vec{y}] = \left(\begin{array}{cc|c} 1 & 10 & 11.6 \\ 1 & 15 & 11.85 \\ 1 & 20 & 12.25 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 10 & 11.6 \\ 0 & 5 & 0.25 \\ 0 & 10 & 0.65 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 10 & 11.6 \\ 0 & 5 & 0.25 \\ 0 & 0 & 0.15 \end{array}\right)$$

Thus, this system is inconsistent since $\vec{y}$ cannot be expressed as linear conbinations of columns of A since they are linearly independent.

Choose equations 1, 2

$$\begin{cases} x_1 + 10x_2 = 11.6 \\ x_1 + 15x_2 = 11.85 \end{cases} \Rightarrow x_1 = 11.1,\ x_2 = 0.05$$

$\Rightarrow$ 3rd formula equals to 12.1 $\Rightarrow$ difference $= 0.15$

Choose equations 1, 3

$$\begin{cases} x_1 + 10x_2 = 11.6 \\ x_1 + 20x_2 = 12.25 \end{cases} \Rightarrow x_1 = 10.95,\ x_2 = 0.065$$

$\Rightarrow$ 2nd formula equals to 11.925 $\Rightarrow$ difference $= 0.075$

Choose equations 2, 3

$$\begin{cases} x_1 + 15x_2 = 11.85 \\ x_1 + 20x_2 = 12.25 \end{cases} \Rightarrow x_1 = 10.65,\ x_2 = 0.08$$

$\Rightarrow$ 1st formula equals to 11.45 $\Rightarrow$ difference $= 0.15$

The 2nd pair is preferred since it has the minimum difference and the most accurate precision.

(c)

$$A^T = \begin{pmatrix} 1 & 1 & 1 \\ 10 & 15 & 20 \end{pmatrix} \Rightarrow A^T A = \begin{pmatrix} 1 & 1 & 1 \\ 10 & 15 & 20 \end{pmatrix} \begin{pmatrix} 1 & 10 \\ 1 & 15 \\ 1 & 20 \end{pmatrix} = \begin{pmatrix} 3 & 45 \\ 45 & 725 \end{pmatrix},\ A^T\vec{y} = \begin{pmatrix} 1 & 1 & 1 \\ 10 & 15 & 20 \end{pmatrix} \begin{pmatrix} 11.6 \\ 11.85 \\ 12.25 \end{pmatrix} = \begin{pmatrix} 35.7 \\ 538.75 \end{pmatrix}$$

$$\Rightarrow A^T A \vec{x} = A^T \vec{b} \Leftrightarrow \begin{pmatrix} 3 & 45 \\ 45 & 725 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 35.7 \\ 538.75 \end{pmatrix}$$

$$\Rightarrow [A^T A \ A^T\vec{b}] = \left(\begin{array}{cc|c} 3 & 45 & 35.7 \\ 45 & 725 & 538.75 \end{array}\right) \sim \begin{pmatrix} 3 & 45 & 35.7 \\ 0 & 50 & 3.25 \end{pmatrix} \Rightarrow \vec{x} = (10.925, 0.065)^T$$

This solution is closest to the 2nd solution obtained in (b), using the first and third equations.

3.2. Suppose you are fitting a straight line to the three data points (0,1), (1,2), (3,3).
(*a*) Set up the overdetermined linear system for the least squares problem.
(*b*) Set up the corresponding normal equations.
(*c*) Compute the least squares solution by Cholesky factorization.

Solution:

(a)

The straight line can be expressed as $y(t) = x_1 + x_2 t \quad (x_1, x_2 \in R)$

Thus, the system is given below

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

(b)

Let $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{pmatrix}$, $\vec{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, the normal equation is

$A^T A \vec{x} = A^T \vec{y} \Rightarrow A^T = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix} \Rightarrow A^T A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 4 & 10 \end{pmatrix}$

$A^T \vec{y} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 11 \end{pmatrix}$

Thus, the normal equations are $\begin{pmatrix} 3 & 4 \\ 4 & 10 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 11 \end{pmatrix}$

(c)

Since $A^T A$ is square and symmetric $\Rightarrow A^T A = LU = LL^T$

Let $L = \begin{pmatrix} a_{11} & 0 \\ a_{12} & a_{22} \end{pmatrix} \Rightarrow U = L^T = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$

$\Rightarrow LU = LL^T = \begin{pmatrix} a_{11} & 0 \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}^2 & a_{11}a_{12} \\ a_{12}a_{11} & a_{12}^2 + a_{22}^2 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 4 & 10 \end{pmatrix}$

$\Rightarrow a_{11} = \sqrt{3},\ a_{12} = \frac{4}{\sqrt{3}},\ a_{22} = \sqrt{\frac{14}{3}}$

Thus the normal equation $A^T A \vec{x} = A^T \vec{y} \Leftrightarrow LL^T \vec{x} = A^T \vec{y} \Leftrightarrow L\vec{z} = A^T \vec{y}$

where $\vec{z} = L^T \vec{x}$

$\Rightarrow \begin{pmatrix} \sqrt{3} & 0 \\ \frac{4}{\sqrt{3}} & \sqrt{\frac{14}{3}} \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 11 \end{pmatrix} \Rightarrow \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 2\sqrt{3} \\ \frac{3\sqrt{3}}{\sqrt{14}} \end{pmatrix}$

Thus, $\begin{pmatrix} \sqrt{3} & \frac{4}{\sqrt{3}} \\ 0 & \sqrt{\frac{14}{3}} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2\sqrt{3} \\ \frac{3\sqrt{3}}{\sqrt{14}} \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{8}{7} \\ \frac{9}{14} \end{pmatrix}$

3.3. Set up the linear least squares system $\boldsymbol{Ax} \cong \boldsymbol{b}$ for fitting the model function $f(t, \boldsymbol{x}) = x_1 t + x_2 e^t$ to the three data points (1,2), (2,3), (3,5).

Solution:

Since $f(t, \vec{x}) = x_1 t + x_2 e^t$

$\Rightarrow f(1, \vec{x}) = x_1 + x_2 e = 2$

$f(2, \vec{x}) = x_1 + x_2 e^2 = 3$

$f(3, \vec{x}) = x_1 + x_3 e^3 = 5$

Thus, the linear least squares system is

$$\begin{pmatrix} 1 & e \\ 1 & e^2 \\ 1 & e^3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

3.4. In fitting a straight line $y = x_0 + x_1 t$ to the three data points (t_i, y_i) = (0,0), (1,0), (1,1), is the least squares solution unique? Why?

Solution: Since $y = x_0 + x_1 t$

$\Rightarrow \begin{cases} x_0 + 0 = 0 \\ x_0 + x_1 = 0 \\ x_0 + x_1 = 1 \end{cases} \Rightarrow \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$

Thus, this least squares solution is unique since the columns of A are linearly independent

3.11. Suppose that the partitioned matrix

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{C} \end{bmatrix}$$

is orthogonal, where the submatrices $\boldsymbol{A}$ and $\boldsymbol{C}$ are square. Prove that $\boldsymbol{A}$ and $\boldsymbol{C}$ must be orthogonal, and $\boldsymbol{B} = \boldsymbol{O}$.

proof

Denote the given matrix by M

Since M is orthogonal $\Rightarrow M^T M = \begin{pmatrix} A^T & O^T \\ B^T & C^T \end{pmatrix}\begin{pmatrix} A & B \\ O & C \end{pmatrix} = \begin{pmatrix} A^T A & A^T B \\ B^T A & B^T B + C^T C \end{pmatrix}$

$= I$

$\Rightarrow A^T A = I \Rightarrow A$ is orthogonal

$A^T B = 0 \Rightarrow B = 0$

$B^T B + C^T C = I = C^T C \Rightarrow C$ is orthogonal