LINEAR ALGEBRA III

Least-Squares Problems:
Suppose the system $Ax = b$ is inconsistent. i.e, the solution doesnot exist. The best one can do is to find an x that makes Ax as close as possible to b.
If A is an $m \times n$ matrix and b is in $\mathbb{R}^m$, a least-squares solution of $Ax = b$ is an $\hat{x}$ in $\mathbb{R}^n$ such that $|b - A\hat{x}| \leq |b - Ax|$ for all x in $\mathbb{R}^n$.
Note:
The set of least-squares solution of $Ax = b$ coincides with the non-empty set of solutions of the normal equations $A^TAx = A^Tb$.
ex.1.
Find a least-squares solution of the inconsistent system $Ax = b$ for $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$.

Soln:

$$A^TA = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^TAx = A^Tb$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

$$R_2 = 17R_2 - R_1 \implies \begin{bmatrix} 17 & 1 \\ 0 & 84 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 168 \end{bmatrix}$$

$\implies 17x_1 + x_2 = 19, 84x_2 = 168 \implies x_1 = 1, x_2 = 2$

$\therefore \hat{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is the least-squares solution.

ex.2.

Find a least-squares solution of the inconsistent system $Ax = b$ for $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$, $b = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$.

Soln:

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

Then the equation $A^TAx = A^Tb$ becomes

$$\begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

The augmented matrix $[A^TA : A^Tb]$ is

$$\begin{bmatrix} 6 & 2 & 2 & 2 & : & 4 \\ 2 & 2 & 0 & 0 & : & -4 \\ 2 & 0 & 2 & 0 & : & 2 \\ 2 & 0 & 0 & 2 & : & 6 \end{bmatrix}$$

$$R_2 = 3R_2 - R_1, R_3 = 3R_3 - R_1, R_4 = 3R_4 - R1 \implies \begin{bmatrix} 6 & 2 & 2 & 2 & : & 4 \\ 0 & 4 & -2 & -2 & : & -16 \\ 0 & -2 & 4 & -2 & : & 2 \\ 0 & -2 & -2 & 4 & : & 14 \end{bmatrix}$$

$$R_3 = 2R_3 + R_2, R_4 - 2R_4 + R_2 \implies \begin{bmatrix} 6 & 2 & 2 & 2 & : & 4 \\ 0 & 4 & -2 & -2 & : & -16 \\ 0 & 0 & 6 & -6 & : & -12 \\ 0 & 0 & -6 & 6 & : & 12 \end{bmatrix}$$

$R_4 = R_4 + R_3 \implies$

$$\begin{bmatrix} 6 & 2 & 2 & 2 & : & 4 \\ 0 & 4 & -2 & -2 & : & -16 \\ 0 & 0 & 6 & -6 & : & -12 \\ 0 & 0 & 0 & 0 & : & 0 \end{bmatrix}$$

$\implies 6x_1 + 2x_2 + x_3 + 2x_4 = 4, 4x_2 - 2x_3 - 2x_4 = -16, 6x_3 - 6x_4 = -12$

$\implies x_3 = -2 + x_4, x_2 = -5 + x_4, x_1 = 3 - x_4$

Let $x_4 = k$, then $x_3 = -2 + k, x_2 = -5 + k, x_1 = 3 - k$

$$\therefore \hat{x} = \begin{bmatrix} 3-k \\ -5+k \\ -2+k \\ k \end{bmatrix}.$$

Application to Linear models:

Suppose we want to fit a linear equation of the form $y = \beta_0 + \beta_1 x$ to the data points $(x_1, y_1), (x_2, y_2), ..., (x_n, y$ that when graphed , seem to lie close to a line. If the data points were on the line, the parameters β_0 and β_1, would satisfy the equations

predicted y-value	observed y-value
$\beta_0 + \beta_1 x_1$	y_1
$\beta_0 + \beta_1 x_2$	y_2
.	.
.	.
.	.
$\beta_0 + \beta_1 x_n$	y_n

We can write the system as $X\beta = y$, where $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}$.

If the data points don't lie on a line, then there are no parameters β_0, β_1 for which the predicted $y-$values in $X\beta$ equal the observed $y-$values in y and $X\beta = y$ has no solution. Thus we have a least-squares solution to this problem. Computing the least-squares solution $X\beta = y$ is equivalent to finding the β that determines the least-squares line $y = \beta_0 + \beta_1 x$.

ex.1.

Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points $(2,1),(5,2),(7,3),(8,3)$.

Soln:

Using the $x-$coordinates and $y-$coordinates of the data points, we can write

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix},\ y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}.$$

For the least-squares solution $X\beta = y$, we obtain the normal equations by $X^T X\beta = X^T y$

i.e, $X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$

$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$

The normal equations are $\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$

Hence $\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$

Thus the least-squares line has the equation $y = \frac{2}{7} + \frac{5}{14}x$.

ex.2.

A healthy child's blood pressure P(in millimeters of mercury) and weight W(in pounds) are approximately related by the equation $\beta_0 + \beta_1 lnW = P$. Use the following data to estimate the blood pressure of a healthy child weighing 100 pounds.

W	44	61	81	113	131
ln W	3.78	4.11	4.39	4.73	4.88
P	91	98	103	110	112

Soln: $X = \begin{bmatrix} 1 & 3.78 \\ 1 & 4.11 \\ 1 & 4.39 \\ 1 & 4.73 \\ 1 & 4.88 \end{bmatrix}$ $y = \begin{bmatrix} 91 \\ 98 \\ 103 \\ 110 \\ 112 \end{bmatrix}$ The normal equations are given by $X^T X\beta = X^T y$

i.e, $X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3.78 & 4.11 & 4.39 & 4.73 & 4.88 \end{bmatrix} \begin{bmatrix} 1 & 3.78 \\ 1 & 4.11 \\ 1 & 4.39 \\ 1 & 4.73 \\ 1 & 4.88 \end{bmatrix} = \begin{bmatrix} 5 & 21.89 \\ 21.89 & 96.6399 \end{bmatrix}$

$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3.78 & 4.11 & 4.39 & 4.73 & 4.88 \end{bmatrix} \begin{bmatrix} 91 \\ 98 \\ 103 \\ 110 \\ 112 \end{bmatrix} = \begin{bmatrix} 514 \\ 2265.8 \end{bmatrix}$

The normal equations are $\begin{bmatrix} 5 & 21.89 \\ 21.89 & 96.6399 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 514 \\ 2265.8 \end{bmatrix}$

Hence $\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = \begin{bmatrix}5 & 21.89\\ 21.89 & 96.6399\end{bmatrix}^{-1}\begin{bmatrix}514\\ 2265.8\end{bmatrix} = \dfrac{1}{4.0274}\begin{bmatrix}96.6399 & -21.89\\ -21.89 & 5\end{bmatrix}\begin{bmatrix}514\\ 2265.8\end{bmatrix} = \begin{bmatrix}18.5642\\ 19.2407\end{bmatrix}$

ex.3.

Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points $(-1, 0), (0, 1), (1, 2), (2, 4)$.

ex.4.

Find the line of best fit for the below data.
$b = 2$ at $t = -1$, $b = 0$ at $t = 0$, $b = -3$ at $t = 1$, $b = -5$ at $t = 2$.