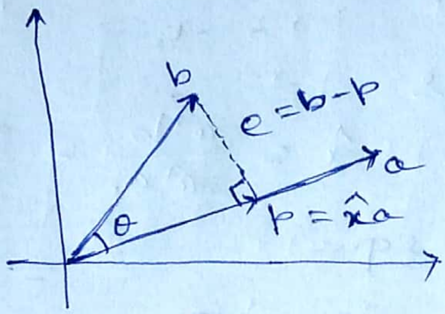
3.3: Projections and Least Squares

Course Outcomes: Students will have understanding about projection of vectors onto lines as well as onto column space of a matrix, least square solution of single as well as several variable systems and Least-squares fitting of data.

Least square problems with single variable:



The least-squares solution to a problem $ax = b$ in one unknown is

$$\boxed{\hat{x} = \frac{a^T b}{a^T a}}$$

Ex:
$$\begin{aligned} 2x &= b_1 \\ 3x &= b_2 \\ 4x &= b_3 \end{aligned} \Rightarrow ax = b,$$

where $a = (2,3,4)$ and $b = (b_1, b_2, b_3)$.

This is solvable when b_1, b_2, b_3 are in the ratio $2:3:4$. The solution x will exist only if b is on the same line as the column $a = (2,3,4)$.

For $b = (4,6,8)$, $x = 2$, which is an exact solution.

For $b = (6,9,12)$, $x = 3$, which is an exact solution also.

But for $b=(4,5,8)$, there is no exact solution of the system $ax=b$ as the components of b are not obeying the ratio $2:3:4$.

In this case b is not on the line passing through a. So, we have to find the least square solution. The least square solution is

$$\hat{x} = \frac{a^Tb}{a^Ta} = \frac{2\times4+3\times5+4\times8}{2^2+3^2+4^2}$$

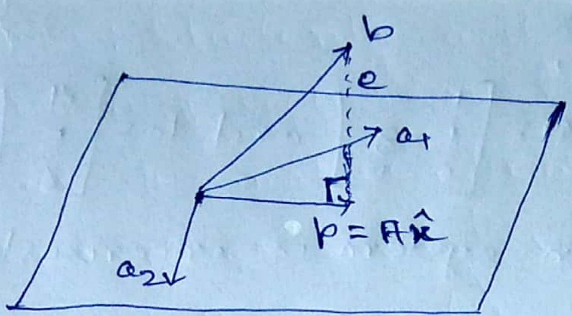
$$= \frac{55}{29}.$$

Least-Square Problems with Several Variables:

Given: $Ax=b$

The problem is to choose $\hat{x}$ so as to minimize the error and again this minimization will be done in the least square sense. The error is $E=\|Ax-b\|$ and this is exactly the distance from b to the point Ax in the column space. Searching for the least-squares solution $\hat{x}$, which minimizes E, is the same as locating the point $p=A\hat{x}$ that is closer to b than any other point in the column space. The error vector $e=b-p=b-A\hat{x}$ must be perpendicular to the column space.

To find: 1. To find the least square soln $\hat{x}$.
2. The projection $p=A\hat{x}$ onto $C(A)$.



Given: $Ax = b$ (inconsistent i.e. $b \notin C(A)$)

All vectors perpendicular to the column space lie in the left null space. So, the error vector $e = b - p = b - A\hat{x}$ must lie in the nullspace of A^T.

$$A^T(b - A\hat{x}) = 0$$

$$\Rightarrow A^T b - A^T A\hat{x} = 0$$

$$\Rightarrow A^T A\hat{x} = A^T b$$

$$\Rightarrow \boxed{\hat{x} = (A^T A)^{-1} A^T b}$$ is the least-squares solution.

This least-squares solution is also known as the best estimate.

The projection of b onto the column space is the nearest point $A\hat{x}$.

$$p = A\hat{x}$$

$$\Rightarrow \boxed{p = A(A^T A)^{-1} A^T b}$$

Ex: Solve $Ax = b$ by least squares and find the projection of b onto the column space of A, where $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$ and $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$.

Solⁿ: $Ax = b$,

where $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$ and $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

The given system is inconsistent as $b \notin C(A)$.

$$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}$$

$$(A^TA)^{-1} = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 23 \end{bmatrix}$$

The least-squares solⁿ is

$$\hat{x} = (A^TA)^{-1}A^Tb$$
$$= \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 9 \\ 23 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

The projection of b onto the column space of A is

$$p = A\hat{x}$$
$$= \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}.$$

$p \in C(A)$.

Projection Matrices:

$$p = A\hat{x}$$
$$\Rightarrow p = A(A^TA)^{-1}A^Tb$$
$$\Rightarrow p = Pb,$$

where $\boxed{P = A(A^TA)A^T}$ is the projection matrix that projects any vector b onto the column space of A.

No. 12. $a_1 = (1,0,1)$, $a_2 = (1,1,-1)$

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

Projection matrix $P = A(A^TA)^{-1}A^T$

$$A^TA = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad (A^TA)^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

$$(A^TA)^{-1}A^T = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

$$P = A(A^TA)^{-1}A^T$$

$$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{5}{6} \end{bmatrix}$$