Lecture - 21 & 22

Class Note

Orthogonal Complement: let V be an inner product space over $\mathbb{R}$. let S be a subset of V.
The **orthogonal complement** $S^{\perp}$ of S is the set of all elements in V that are orthogonal to each element of S. So

$$S^{\perp} = \{v \in V \mid v \perp w \ \forall \ w \in S\}$$

Note:

(i) $S^{\perp}$ is a subspace of V

proof: let $v_1, v_2 \in S^{\perp}$ $\forall$ $w \in S$.

$\Rightarrow$ $v_1 \perp w$ & $v_2 \perp w$

Now $\langle c v_1 + d v_2 . w \rangle$

$= c \langle v_1 . w + d \langle v_2 . w \rangle$

$= c \times 0 + d \times 0$ $\left(\because \begin{array}{l} v_1 \perp w \Rightarrow \langle v_1 . w \rangle = 0 \\ v_2 \perp w \Rightarrow \langle v_2 . w \rangle = 0 \end{array} \right)$

$= 0$

$\Rightarrow$ $S^{\perp}$ is a subspace of V.

(2) $\{\theta\}^{\perp} = V$

(3) $V^{\perp} = \{\theta\}$

Theorem: The Orthogonal Decomposition theorem:

Let W be a subspace of $\mathbb{R}^n$. Then each y in $\mathbb{R}^n$ can be written uniquely in the form

$$y = \hat{y} + z$$

where $\hat{y}$ is in W (= is called orthogonal projection of y onto W).

an z is in $W^\perp$.

In fact, if $\{u_1, u_2, \dots u_n\}$ is any orthogonal basis of W, then

$$\hat{y} = \frac{\langle y . u_1 \rangle}{\langle u_1 . u_1 \rangle} u_1 + \dots + \frac{\langle y . u_n \rangle}{\langle u_n . u_n \rangle} u_n .$$

and $z = y - \hat{y}$.

proof:- let $\{u_1, u_2, \dots, u_n\}$ be any orthogonal basis for W.

$\therefore \hat{y} = c_1 u_1 + \dots + c_n u_n$

$\Rightarrow \langle \hat{y} . u_1 \rangle = c_1 \langle u_1 . u_1 \rangle + \dots + c_n \langle u_n . u_1 \rangle$

$= c_1 \langle u_1 . u_1 \rangle$ $(\because \langle u_i . u_j \rangle = 0 \ \forall \ i \neq j)$

Similarly $c_2 = \dfrac{\langle y . u_2 \rangle}{\langle u_2 . u_2 \rangle}, \dots, c_n = \dfrac{\langle y . u_n \rangle}{\langle u_n . u_n \rangle}$.

$$\therefore \hat{y} = \frac{\langle y . u_1 \rangle}{\langle u_1 . u_1 \rangle} u_1 + \frac{\langle y . u_2 \rangle}{\langle u_2 . u_2 \rangle} u_2 + \dots + \frac{\langle y . u_n \rangle}{\langle u_n . u_n \rangle} u_n$$

$\Rightarrow \hat{y}$ can be written as linear combination of $\{u_1, u_2 \dots u_n\}$ $\Rightarrow \hat{y} \in W$.

Now let $z = y - \hat{y}$

$\because$ u_1 is orthogonal to $u_2, \dots u_n$

it follows that:-

$$\langle z . u_1 \rangle = \langle (y - \hat{y}) . u_1 \rangle$$

$$= \langle y . u_1 \rangle - \langle \hat{y} . u_1 \rangle$$

$$= \langle y . u_1 \rangle - \left\langle \frac{\langle y . u_1 \rangle}{\langle u_1 . u_1 \rangle} u_1 + \dots + \frac{\langle y . u_n \rangle}{\langle u_n . u_n \rangle} . u_1 \right\rangle$$

$$= \langle y . u_1 \rangle - \frac{\langle y . u_1 \rangle}{\cancel{\langle u_1 . u_1 \rangle}} \cancel{\langle u_1 . u_1 \rangle} - 0 - \dots - 0$$

$$= \langle y . u_1 \rangle - \langle y . u_1 \rangle = 0$$

$\Rightarrow$ z is orthogonal to u_1

|||^{ly} z is " to u_j in the basis for W

Hence z is orthogonal to every vector in W.

$\Rightarrow$ z is in $W^{\perp}$

Best Approximation theorem:

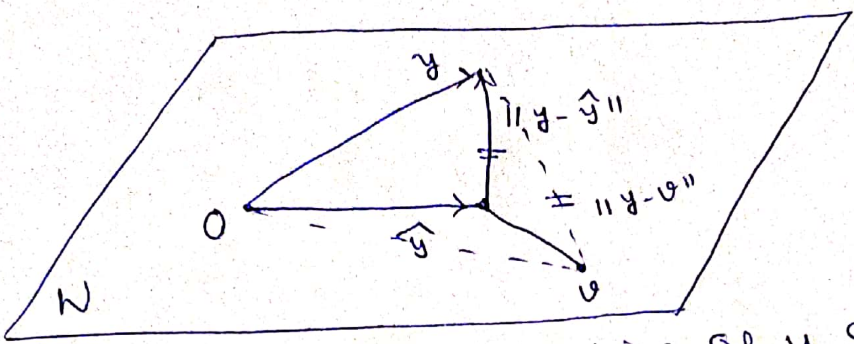
Let W be a subspace of $\mathbb{R}^n$. Let y be any vector in $\mathbb{R}^n$. and let $\hat{y}$ be the orthogonal projection of y onto W. Then $\hat{y}$ is the closest point in W to y, in the sense that

$$\| y - \hat{y} \| < \| y - v \| \text{ destinct from } \hat{y}$$

$$\forall \ v \text{ in } W$$

Note: For u and v in $\mathbb{R}^n$, the distance between u and is the length of the vector $u - v$ i.e.

$$dis(u, v) = \| u - v \|$$



The orthogonal projection of y onto W is the closest point in W to y.

Least - Squares Problems :

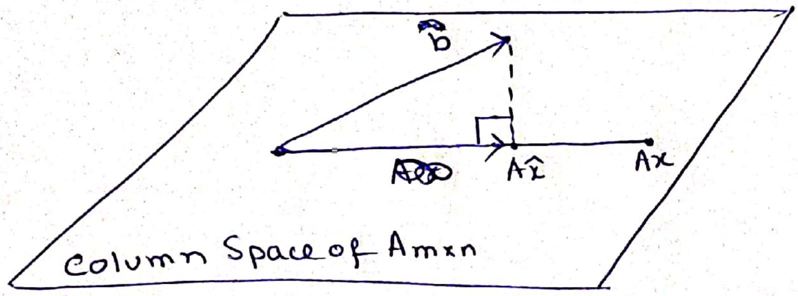
If $Ax = b$ had no solution i.e system is inconsistent When a solution is demanded and none exists, the best one can do is to find an x that makes Ax as close as possible to b.

Think of Ax as an approximation to b. The smaller the distance between b and Ax given by $\|b - Ax\|$, the better the approximation. The general least-squares problem is to find an x that makes $\|b - Ax\|$ as small as possible. The adjective "least squares" arises from the fact that $\|b - Ax\|$ is the square root of a sum of squares.

Definition: If A is $m \times n$ matrix and b is $m \times 1$ matrix, a least-squares solution of $Ax = b$ is an $\hat{x}$ in $\mathbb{R}^n$ ($A\hat{x}$ is the orthogonal projection of b on Ax) such that

$$\|b - A\hat{x}\| \leq \|b - Ax\|$$

$\forall$ x in $\mathbb{R}^n$.



The vector b is closer to $A\hat{x}$ than to Ax for other x.

Solution of the General least-squares Problem:

Theorem: The set of least-squares solutions of $Ax = b$ coincides with the non-empty set of solutions of the normal equations

$$A^T A x = A^T b$$

proof: Column Space of $A_{m \times n}$ = {set of all column vectors} $\subset \mathbb{R}^n$

let $\hat{b}$ = orthogonal projection of b onto column space of $A_{m \times n}$

then by best approximation theorem in column space $\hat{b}$ is the closest point to b.

$\Rightarrow$ $\hat{b} \in$ column space of A

$\therefore$ the equation $Ax = \hat{b}$ is consestent and there is an $\hat{x}$ in $\mathbb{R}$ such that

$$A\hat{x} = \hat{b} \longrightarrow (1)$$

$\because$ $\hat{b}$ is the closest point in column space to b, a vector $\hat{x}$ is a least-squares solution of $Ax = b$ iff $\hat{x}$ satisfies (1)

Then by orthogonal decomposition theorem the projection $\hat{b}$ has the property that $b - \hat{b}$ is orthogonal to column space A.

So $b - A\hat{x}$ ($\because \hat{b} = A\hat{x}$) is orthogonal to each column of A. then.

If a_j is any column of A.

$$\langle a_j \cdot (b - A\hat{x}) \rangle = 0$$

and $a_j^T (b - A\hat{x}) = 0$ ($\because$ each a_j^T is a row of A^T).

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

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

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

$\Rightarrow$ least square solution of $Ax = b$ satisfies

$$A^T A x = A^T b.$$

Ex 1: Find a least-squares solutions of the inconsestent system $Ax = b$ for.

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}.$$

Soln:-

$$A^T A = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}.$$

$$A^T b = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 6 \\ 11 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix}.$$

Then the eqn $A^T A x = A^T b$ becomes.

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

$$17x_1 + x_2 = 19.$$
$$x_1 + 5x_2 = 11$$

$$\begin{array}{r} 17 \\ \times 5 \\ \hline 85 \end{array} \quad \begin{array}{r} 17 \\ 11 \\ \hline 187 \\ 19 \\ \hline 168 \end{array}$$

$$\Rightarrow \quad \begin{array}{rcl} \cancel{17x_1} + x_2 &=& 19. \\ \cancel{17x_1} + 5 \times 17 x_2 &=& 11 \times 17 \\ - \quad\quad - & & - \\ \hline 84x_2 &=& 168 \end{array}$$

$$x_2 = 2.$$

$$x_1 = \frac{19 - 2}{17} = 1$$

$$\therefore \hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$