St. Xavier's College (Autonomous), Kolkata

Department of Statistics

MDTS 4113/SEM I

Linear Algebra

Notes on Orthogonal Projections

Let V be an inner product space and W be a finite-dimensional subspace of V.

Then any vector $\underline{x} \in$ V is uniquely represented as $\underline{x} = \underline{y} + \underline{e}$, where $\underline{y} \in W$ and $\underline{e} \perp$ W.

The component $\underline{y}$ is the orthogonal projection of the vector $\underline{x}$ onto the subspace W. The distance from $\underline{x}$ to the subspace W is $\|\underline{e}\|$.

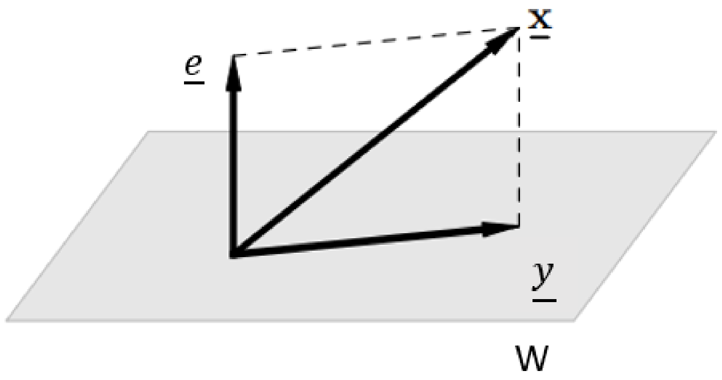
If $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ is an orthogonal basis for W then

$$\underline{y} = \frac{\langle \underline{\mathbf{x}}, \underline{\mathbf{v}}_1 \rangle}{\langle \underline{\mathbf{v}}_1, \underline{\mathbf{v}}_1 \rangle}\underline{\mathbf{v}}_1 + \frac{\langle \underline{\mathbf{x}}, \underline{\mathbf{v}}_2 \rangle}{\langle \underline{\mathbf{v}}_2, \underline{\mathbf{v}}_2 \rangle}\underline{\mathbf{v}}_2 + \cdots + \frac{\langle \underline{\mathbf{x}}, \underline{\mathbf{v}}_n \rangle}{\langle \underline{\mathbf{v}}_n, \underline{\mathbf{v}}_n \rangle}\underline{\mathbf{v}}_n$$

Projection of $\underline{x}$ on $\underline{v}_1$

Projection of $\underline{x}$ on $\underline{v}_2$

Projection of $\underline{x}$ on $\underline{v}_n$



- The orthogonal projection of any vector $\underline{x}$ onto $\underline{v}_i$: $Proj_{\underline{v}_i}(\underline{x}) = (\underline{x}.\ \underline{v}_i)\ \underline{v}_i$, i= 1, 2,…n.

- The orthogonal projection of any vector $\underline{x}$ onto W: $Proj_W(\underline{x})$ = $\sum_{i=1}^{n} (\underline{x}.\, \underline{v_i})\, \underline{v_i} \;=\; \sum_{i=1}^{n} Proj_{\underline{v_i}}(\underline{x})$
- $\underline{x} = Proj_W(\underline{x}) + \underline{e}$, $\qquad \underline{e} \perp$W
- Define a matrix P =[$\underline{v_1}\ \underline{v_2}\ \ldots.\underline{v_n}$] then, $Proj_W(\underline{x})$ = P$P^T\underline{x}$

Best Approximation:

Let W be a subspace.

- Then, for any vector $\underline{y}$, $Proj_W(\underline{y}) \in$ W is the best approximation to $\underline{y}$ by vectors in W.
- More precisely, $Proj_W(\underline{y})$ is the closest point in W to $\underline{y}$,

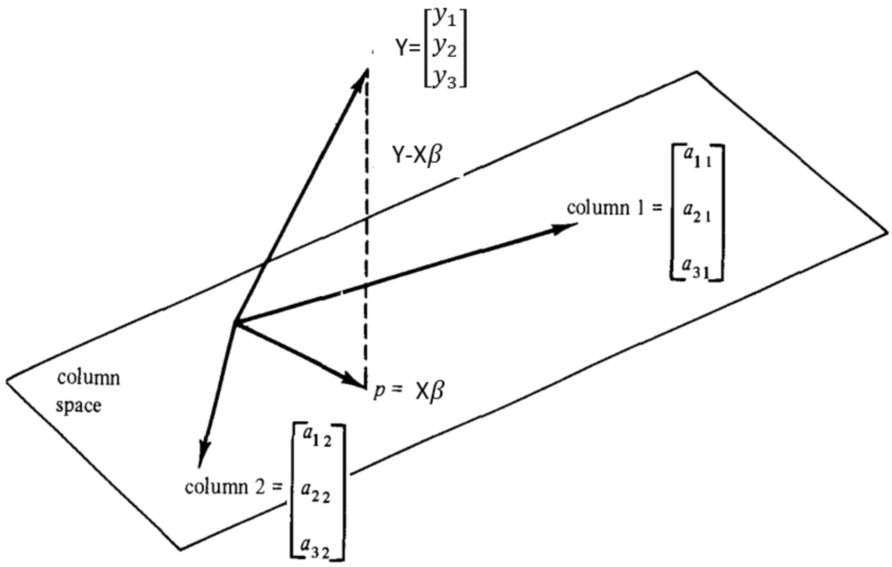
$$\text{i.e., } \left\|\underline{y} - \mathrm{Proj}_W(\underline{y})\right\| \leq \left\|\underline{y} - \underline{w}\right\|, \text{ for any } \underline{w} \in W$$

- This closest distance is defined as the distance from $\underline{y}$ to W :

$$\mathrm{dist}(\underline{y}, W) = \min \left\{ \left\|\underline{y} - \underline{w}\right\| : \underline{w} \in W \right\} = \left\|\underline{y} - \mathrm{Proj}_W(\underline{y})\right\|$$

Method of Least Squares observations:

$\underline{y}$(response)
Consider a Linear Model: $\underline{y}$=X$\underline{\beta} + \underline{\varepsilon}$
To estimate the parameter $\underline{\beta}$.
The error is E=|| $\underline{y}$ −X$\underline{\beta}$||: distance of $\underline{y}$ to the point X$\underline{\beta}$ in the column space of X.
So, searching for the least square solution $\underline{\beta}$ minimizing E, is same as locating the point p= X$\underline{\beta}$, that is closer to $\underline{y}$ than any other point in the column space.


Y= $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$
Y-Xβ
column 1 = $\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}$
column space
p = Xβ
column 2 = $\begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix}$