Module – 3.3

A GEOMETRIC VIEW OF LEAST SQUARES - PROJECTIONS

AN OVERVIEW

S. Lakshmivarahan

School of Computer Science

University of Oklahoma

Norman, Ok – 73069, USA

varahan@ou.edu

PROJECTIONS

- Let Z = $(Z_1, Z_2)^T$, $P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $P_2 = \begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix}$, $a > 0$
- $\widehat{Z_1} = P_1 Z = \begin{pmatrix} Z_1 \\ 0 \end{pmatrix}$, $\widehat{Z_2} = P_2 Z = \begin{pmatrix} Z_1 + aZ_2 \\ 0 \end{pmatrix}$

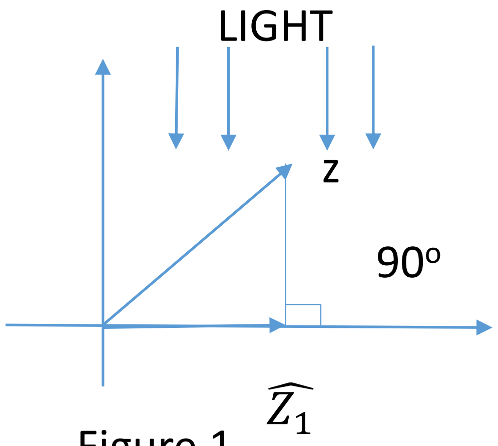


Figure 1

$\widehat{Z_1}$ - orthogonal projection of Z on the x_1 - axis

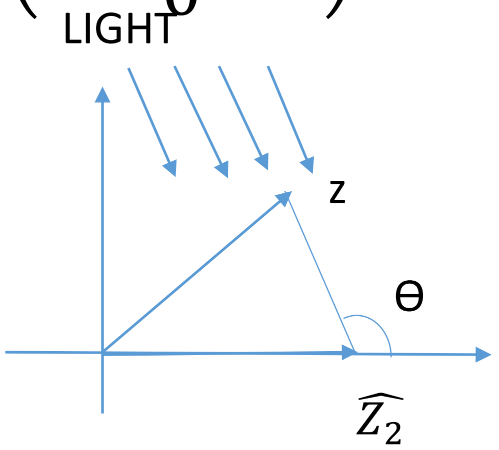


Figure 2

$\widehat{Z_2}$ - oblique projection of Z on the x_1 - axis

$\hat{Z}$ - Shadow of Z on the x_1- axis

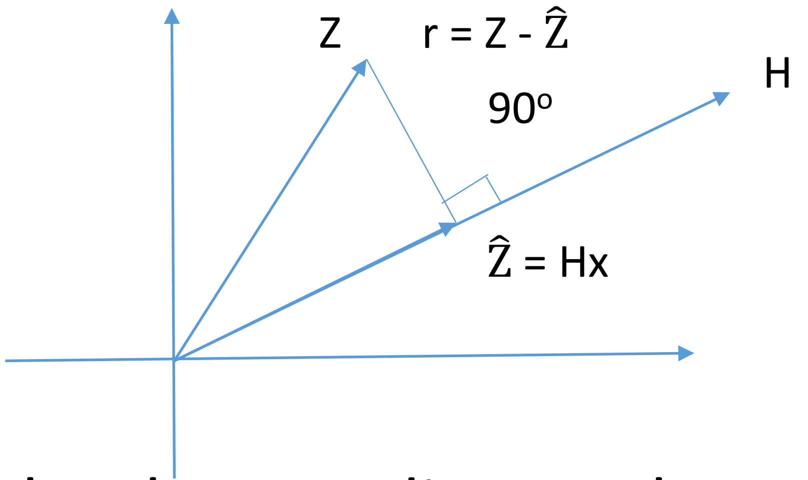
PROJECTION MATRICES/OPERATORS

- P_1 – orthogonal projection matrix
- P_2 – oblique projection matrix
- Every projection matrix is idempotent: $P_1^2 = P_1$

  $P_2^2 = P_2$
- Every orthogonal projection matrix is symmetric: $P_1^T = P_1$
- Every oblique projection matrix is not symmetric: $P_2^T \neq P_2$
- Every projection matrix is singular, that is, rank deficient: $\det(P_1) = 0$, $\det(P_2) = 0$

ORDINARY LEAST SQUARES AND ORTHOGONAL PROJECTION

- Let $H \in R^m$ and $Z \in R^m$ and $Z \neq H$



- $r = (Z - \hat{Z}) \perp H$
- $\hat{Z}$ is the orthogonal projection Z onto H

- A geometric fact: The shortest distance between a line and a point not on the line, is the length of the perpendicular from the point to the line
- Referring to the figure, let $\hat{Z}$ be the point where the perpendicular line from the point Z (tip of the vector Z) intersects the vector H
- Then, $r = Z - \hat{Z}$ is perpendicular to H

ORTHOGONAL PROJECTION

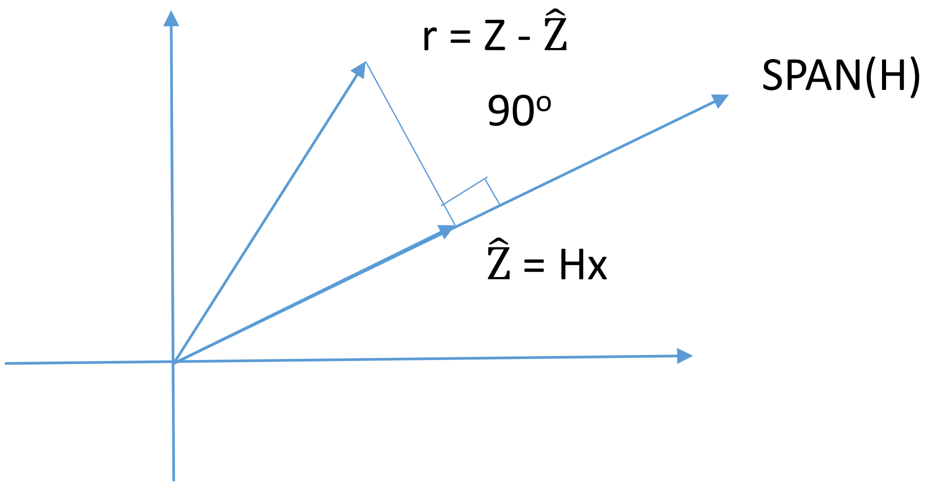
- Since $\hat{Z}$ is a vector in the direction H, there is a scalar $x \in R$ such that

  $\hat{Z} = Hx$
- Combining: $r = Z - \hat{Z} = (Z - Hx) \perp H$
- That is: $H^T(Z - Hx) = 0$ leads to the least square solution

  $\Rightarrow (H^TH)x = H^TZ$ or $x_{LS} = (H^TH)^{-1}H^TZ$
- Then $\hat{Z} = Hx_{LS} = H(H^TH)^{-1}H^TZ = P_HZ$
- $P_H = H(H^TH)^{-1}H^T \in R^{mxm}$ is called the orthogonal projection matrix induced by H

GENERALIZATION

- Let $H \in R^{m\times n}$, $z \in R^m$ and $m > n \geq 1$



- $r = (Z - \hat{Z}) \perp$ SPAN(H)
- $\hat{Z}$ is the orthogonal projection Z onto the SPAN(H)
- $x \in R^n$

- Referring to the figure:

$$r = (Z - \hat{Z}) \perp \text{columns of } H$$

- Since $\hat{Z} \in$ SPAN(H), there exist $x \in R^n$: $\hat{Z} = Hx$
- Combining: $r = (Z - Hx) \perp H$

GENERALIZATION

- That is, $H^T(Z - Hx) = 0$

  $\Rightarrow (H^T H)x = H^T Z$ – Normal equation [refer to Module 3.1]

- Therefore: $x_{LS} = H^T(H^T H)^{-1}Z$, the least square solution

- $\hat{Z} = Hx_{LS} = H(H^T H)^{-1}H^T Z = P_H Z$

- $P_H = H(H^T H)^{-1}H^T = HH^+ \in R^{mxm}$ is an orthogonal projection matrix

- $H^+ = (H^T H)^{-1}H^T$ is the generalized inverse of H

PROPERTIES OF P_H

- $P_H^2 = P_H$ – idempotent
- $P_H^T = P_H$ – symmetric
- P_H is the orthogonal projection operator from R^m to R^n = SPAN(H) where $m > n \geq 1$
- $\det(P_H) = 0$ and P_H is non-singular

WEIGHTED LEAST SQUARES

- Consider $Z = Hx$, $Z \in R^m$, $H \in R^{mxn}$, $x \in R^n$, $W \in R^{mxm}$ - SPD
- $r(x) = Z - H(x)$ – residual vector
- $f(x) = r^T(x)Wr(x)$ – weighted sum of squared residuals
- $X_{Ls} = (H^TWH)^{-1}H^TWZ$
- $\hat{Z} = Hx_{Ls} = H(H^TWH)^{-1}H^TWZ = P_H(W)Z$
- $P_H(W) = H(H^TWH)^{-1}H^TW = HH^+(W) \in R^{mxn}$ – Projection matrix
- $H^+(W) = (H^TWH)^{-1}H^TW$ – Weighted generalized inverse

$P_H(W)$ – OBLIQUE PROJECTION MATRIX

- Verify $P_H^2(W) = P_H(W)$ – idempotent
- Verify $P_H^T(W) \neq P_H(W)$ – not symmetric
- Hence, $P_H(W)$ is an oblique projection matrix

ILLUSTRATION: m = 2, n =1

- Let H = $\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$, Z = $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, x ∈ R, W = $\begin{bmatrix} W_1 & a \\ a & W_2 \end{bmatrix}$
- $\mathrm{H^TWH} = (W_1 h_1^2 + 2ah_1h_2 + W_2 h_2^2) \in R$
- $P_H(W) = \frac{1}{(H^TWH)} HH^TW$

$$= \frac{1}{(H^TWH)} \begin{bmatrix} W_1 h_1^2 + ah_1h_2 & ah_1^2 + W_2 h_1 h_2 \\ ah_2^2 + W_1 h_1 h_2 & W_2 h_2^2 + ah_1h_2 \end{bmatrix}$$

- Set $h_1 = 1$, $h_2 = 0$ => h = $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$$\Rightarrow P_H(W) = \frac{1}{W_1} \begin{bmatrix} W_1 & a \\ 0 & 0 \end{bmatrix}$$

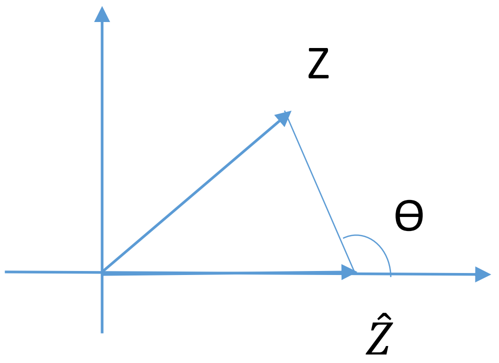
$$\hat{Z} = P_H(W)Z = \begin{bmatrix} Z_1 + \bar{a} Z_2 \\ 0 \end{bmatrix}, \bar{a} = \frac{a}{W_1}$$

ILLUSTRATION - CONTINUED

- $r(x) = Z - \hat{Z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{bmatrix} Z_1 + \bar{a}Z_2 \\ 0 \end{bmatrix} = \begin{pmatrix} -\bar{a} \\ 1 \end{pmatrix} z_2$
- $r^T(x)H = \|r(x)\|_2 \|H\|_2 \cos\theta$

  $-\bar{a}z_2 = z_2(1 + \bar{a}^2)^{½} \cos\theta$

  $\Rightarrow \cos\theta = -\frac{\bar{a}}{(1+\bar{a}^2)^{\frac{1}{2}}} = -\frac{a}{(a^2 + w_1^2)^{½}}$



- That is, $\theta > 90^o$ and r(x) makes an obtuse angle θ with H - see the illustration
- When a = 0, $\cos\theta = 0$ and $\theta = 90 \Rightarrow$ Projection is orthogonal

EXERCISES

8.1) Recall the formula

$\cos\theta = -\dfrac{\overline{a}}{(1+\overline{a}^2)^{\frac{1}{2}}}$. Plot the value of θ as $\overline{a}$ ranges in the interval [-1, 1]

8.2) Let $H^{+}(W) = (H^{T}WH)^{-1}H^{T}W$ is the expression for the weighted generalized inverse. Check if satisfies the Moore – Penrose condition in Module 2.2

REFERENCES

- This module follows Chapter 6 of LLD (2006)