Akshay
Lec-8

Date:
Page No.

# Regression Analysis

$Z_1, Z_2, Z_3, \ldots, Z_r$ ; $Y \rightarrow$ Random

$Y =$ Current market value of home.

mean depends on these

$Z_1 =$ square feet of living area

$Z_2 =$ location

$Z_3 =$ last year valuation

$Z_4 =$ Quality of the construction

$Y =$ mean + random error
$\epsilon$

$$Y = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 Z_4 + \epsilon$$

$$Y = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \cdots + \beta_r Z_r + \epsilon$$

$$Y_1 = \beta_0 + \beta_1 Z_{11} + \cdots + \beta_r Z_{1r} + \epsilon_1$$

$$Y_2 = \beta_0 + \beta_1 Z_{21} + \cdots + \beta_{2r} Z_{2r} + \epsilon_2$$

$$Y_n = \beta_0 + \beta_1 Z_{n1} + \beta_2 Z_{n2} + \cdots + \beta_r Z_{nr} + \epsilon_n$$

$\epsilon_j \sim N(0, \sigma^2)$

$E(\epsilon_j) = 0$, $Var(\epsilon_j) = \sigma^2$

$Cov(\epsilon_j, \epsilon_k) = 0$, $j \neq k$

Date:
Page No.

$$\underset{\sim}{Y}_{n\times 1} = Z_{n\times(r+1)} \; \underset{\sim}{\beta}_{(r+1)\times 1} + \underset{\sim}{\epsilon}_{n\times 1}$$

$$Z = \begin{bmatrix} 1 & z_{11} & z_{12} & z_{1r} \\ 1 & z_{21} & z_{22} & z_{2r} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & z_{n1} & z_{n2} & z_{nr} \end{bmatrix} \qquad \underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_r \end{bmatrix}$$

① $E(\underset{\sim}{\epsilon}) = \underset{\sim}{0}$

② $Cov(\underset{\sim}{\epsilon}) = \sigma^2 I$
$= E(\epsilon\epsilon')$

$$\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}$$

$$\boxed{\begin{array}{l} \underset{n\times 1}{\underset{\sim}{Y}} = \underset{n\times(r+1)}{Z}\,\underset{(r+1)\times 1}{\underset{\sim}{\beta}} + \underset{n\times 1}{\underset{\sim}{\epsilon}} \rightarrow \text{random error component} \qquad \underset{\sim}{\beta}, \sigma^2 = \text{unknown} \\ E(\underset{\sim}{\epsilon}) = \underset{\sim}{0}, \; Cov(\underset{\sim}{\epsilon}) = \sigma^2 I_{n\times n} \end{array}}$$

$\hat{\underset{\sim}{\beta}}, \underset{\sim}{y}, Z$

$\hat{\underset{\sim}{\epsilon}} = \underset{\sim}{y} - \hat{\underset{\sim}{y}}$
↓
residual

estimated mean responses $\leftarrow \hat{\underset{\sim}{y}} = Z\hat{\underset{\sim}{\beta}}$

$SS_{res} = \sum \hat{\epsilon}_j^2 = \sum \hat{\epsilon}'\hat{\epsilon}$ | $\sum_{j=1}^{n} (y_j - \hat{y}_j)^2$

$\hat{\epsilon} = y - Z\hat{\beta}$

Scanned with CamScanner

Date:
Page No.

$$S(\hat{\beta}) = \hat{e}'\hat{e} = (y - Z\hat{\beta})'(y - Z\hat{\beta})$$

$$S(\hat{\beta}) = y'y + \hat{\beta}'Z'Z\hat{\beta} - 2\hat{\beta}'Z'y$$

$$\frac{\partial S(\hat{\beta})}{\partial \hat{\beta}} = 2Z'Z\hat{\beta} - 2Z'y = 0 \Rightarrow Z'Z = Z'y$$

$$\Rightarrow \hat{\beta} = (Z'Z)^{-1}Z'y$$

$$\frac{\partial^2 S(\beta)}{\partial \hat{\beta}^2} = 2Z'Z$$

$$\Rightarrow \boxed{\hat{\beta} = (Z'Z)^{-1}Z'y}$$

$$\boxed{\begin{aligned}\hat{y} = Z\hat{\beta} &= \underbrace{Z(Z'Z)^{-1}Z'}_{H}y \\ &= Hy\end{aligned}}$$

$H \to$ Hat Matrix

$$\hat{e} = y - \hat{y} = y - Z\hat{\beta} = [I - H]y$$

$I - H$ Symmetric & Idempotent Matrix

$(I-H)' = I - H$

$(I-H)^2 = I - H$

~~Least~~ /

Least Square estimation

$\hat{\beta}$ ?

$$Z'\hat{\epsilon} = Z'[I - H]y = Z'[I - Z(Z'Z)^{-1}Z']y$$

$$= [Z' - Z'Z(Z'Z)^{-1}Z']y$$

$$= [Z' - Z']y = \underset{\sim}{0}$$

$$\hat{y}'\hat{\epsilon} = (Z\hat{\beta})'\hat{\epsilon} = \underset{\sim}{\hat{\beta}}'Z'\hat{\epsilon} = 0$$

Example Find $\hat{\beta}$, $\hat{\epsilon}$

$$y_j = \beta_0 + \beta_1 z_{j1} + \epsilon_j$$

| $z_1$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $y$ | 1 | 4 | 3 | 8 | 9 |

$$Z = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \quad y = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 8 \\ 9 \end{bmatrix} \quad \hat{\beta} = (Z'Z)^{-1}Z'y$$

$$Z'Z = \begin{bmatrix} 5 & 10 \\ 10 & 30 \end{bmatrix} \quad (Z'Z)^{-1} = \begin{bmatrix} 0.6 & -0.2 \\ -0.2 & 0.1 \end{bmatrix}$$

$$Z'\underset{\sim}{y} = \begin{bmatrix} 25 \\ 70 \end{bmatrix} \quad \underset{\sim}{\hat{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$\hat{\underset{\sim}{y}} = 1 + 2z$ $\qquad$ $\hat{y} = Z\hat{\underset{\sim}{\beta}} = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \\ 9 \end{bmatrix}$

$\hat{\underset{\sim}{\epsilon}} = \underset{\sim}{y} - \hat{\underset{\sim}{y}} = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}$

$\hat{\epsilon}'\hat{\epsilon} = 6$

Scanned with CamScanner