Chapter 7 - Multivariate Linear Regression

Classical Model $Y = \beta_0 + \beta_1 z_1 + \dots + \beta_r z_r + \varepsilon$

Data: $Y_j = \beta_0 + \beta_1 z_{j1} + \cdots + \beta_r z_{jr} + \varepsilon_j \qquad j = 1, \dots, n$

1) $E\{\varepsilon_j\} = 0$ 2) $V\{\varepsilon_j\} = \sigma^2$ 3) $Cov\{\varepsilon_j, \varepsilon_k\} = 0 \quad j \neq k$

Matrix form:

$$\underset{\sim}{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & z_{11} & \cdots & z_{1r} \\ \vdots & \vdots & & \vdots \\ 1 & z_{n1} & & z_{nr} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_r \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_r \end{bmatrix}$$

Random Y — $\underset{n\times 1}{Y} = \underset{n\times(r+1)}{Z} \; \underset{(r+1)\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$ $\qquad Z =$ Design (aka Model) Matrix

$$E\{\underset{\sim}{\varepsilon}\} = \underset{\sim}{0} \qquad Cov\{\underset{\sim}{\varepsilon}\} = E\{\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}'\} = \sigma^2 I_n$$

observed Z

ANOVA as Regression with dummy variables (Rank issue)

7.3 Least Squares Estimations

~~Choose~~ $\underset{\sim}{b} = \begin{bmatrix} b_0 \\ \vdots \\ b_r \end{bmatrix}$ $S(\underset{\sim}{b}) = \sum_{j=1}^{n} (y_j - b_0 - b_1 x_{j1} - \cdots - b_r x_{jr})^2$

$= (\underset{\sim}{y} - Z\underset{\sim}{b})'(\underset{\sim}{y} - Z\underset{\sim}{b})$ (Assume $Z \equiv$ full rank)

Goal: Choose $\hat{\beta}$ to minimize $S(\underset{\sim}{b})$

$\hat{\underset{\sim}{\varepsilon}} = \underset{\sim}{y} - Z\hat{\underset{\sim}{\beta}} = \underset{\sim}{y} - \hat{\underset{\sim}{y}}$ where $\hat{\beta} = (Z'Z)^{-1}Z'\underset{\sim}{y}$

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

$\Rightarrow \hat{\underset{\sim}{\varepsilon}} = (I - H)y = (I - Z(Z'Z)^{-1}Z')y$

<u>Note: symmetric</u> $[I - Z(Z'Z)^{-1}Z']' = I' - Z\overbrace{[(Z'Z)^{-1}]'}^{\text{symmetric}}Z' = I - Z(Z'Z)^{-1}Z'$

<u>idempotent</u> $(I - Z(Z'Z)^{-1}Z')(I - Z(Z'Z)^{-1}Z')$

$= II - IZ(Z'Z)^{-1}Z' - Z(Z'Z)^{-1}Z' + Z(Z'Z)^{-1}Z'Z(Z'Z)^{-1}Z'$

$= I - Z(Z'Z^{-1})Z' - Z(Z'Z)^{-1}Z' + Z(Z'Z)^{-1}Z' = I - Z(Z'Z)^{-1}Z'$

$Z'\hat{\varepsilon} = Z'(I - \cancel{Z}(Z'Z)^{-1}Z')y = (Z'I - Z'Z(Z'Z)^{-1}Z')y$

$= (Z' - Z')\underset{\sim}{y} = 0$

$S(\underset{\sim}{b}) = (\underset{\sim}{y} - Z\underset{\sim}{b})'(\underset{\sim}{y} - Z\underset{\sim}{b}) = (y - Z\hat{\beta})'(y - Z\hat{\beta}) + (\hat{\beta} - b)'Z'Z(\hat{\beta} - b)$

which is minimized @ $b = \hat{\beta}$ when $Z'Z$ is full rank.

$SSE = \underset{\sim}{y}'(I - H)\underset{\sim}{y}$

Alternative Derivation

$$S(b) = (y - Zb)'(y - Zb) = y'y - y'Zb - b'Z'y + b'Z'Zb$$

$$= \underset{\sim}{y}'\underset{\sim}{y} - 2\underset{\sim}{y}'Z\underset{\sim}{b} + \underset{\sim}{b}'Z'Z\underset{\sim}{b}$$

$$\frac{\partial}{\partial \underset{\sim}{x}} \underset{\sim}{a}'\underset{\sim}{x} = \frac{\partial}{\partial \underset{\sim}{x}} (a_1 x_1 + \dots + a_k x_k) = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} = \underset{\sim}{a}$$

A symmetric: $\frac{\partial}{\partial \underset{\sim}{x}} \underset{\sim}{x}' A \underset{\sim}{x} = 2A\underset{\sim}{x}$ $\quad$ $(= (A + A')x$ when not symmetric$)$

$$\frac{\partial S(b)}{\partial \underset{\sim}{b}} = -2(y'Z)' + 2Z'Zb$$

$$\overset{set}{=} 0 \Rightarrow Z'Z\hat{\beta} = Z'\underset{\sim}{y} \Rightarrow \hat{\underset{\sim}{\beta}} = (Z'Z)^{-1}Z'y$$

Sums of Squares Decomposition

$$y = \hat{\underset{\sim}{y}} + \hat{\underset{\sim}{\varepsilon}} = H\underset{\sim}{y} + (I - H)\underset{\sim}{y} = \hat{y} + (y - \hat{y})$$

$$\hat{\underset{\sim}{y}}'\hat{\underset{\sim}{\varepsilon}} = (H\underset{\sim}{y})'(I - H)\underset{\sim}{y} = \underset{\sim}{y}'H(I - H)\underset{\sim}{y} = 0$$

$$\underset{\sim}{y}'\underset{\sim}{y} = (\hat{\underset{\sim}{y}} + \hat{\underset{\sim}{\varepsilon}})'(\hat{\underset{\sim}{y}} + \hat{\underset{\sim}{\varepsilon}}) = \hat{\underset{\sim}{y}}'\hat{\underset{\sim}{y}} + \underbrace{\hat{\underset{\sim}{y}}'\hat{\underset{\sim}{\varepsilon}}}_{0} + \underbrace{\hat{\underset{\sim}{\varepsilon}}'\hat{\underset{\sim}{y}}}_{0} + \hat{\underset{\sim}{\varepsilon}}'\hat{\underset{\sim}{\varepsilon}} = \hat{y}'\hat{y} + \hat{\varepsilon}'\hat{\varepsilon}$$

$$\Rightarrow \underbrace{\underset{\sim}{y}'\underset{\sim}{y}}_{\substack{\text{SSTOTAL} \\ \text{uncorrected}}} = \underbrace{\hat{\underset{\sim}{y}}'\hat{\underset{\sim}{y}}}_{\substack{\text{SS} \\ \text{Model}}} + \underbrace{\hat{\underset{\sim}{\varepsilon}}'\hat{\underset{\sim}{\varepsilon}}}_{\substack{\text{SS} \\ \text{Residual} \\ \text{(Error)}}}$$

$$Z'\hat{\varepsilon} = 0 \Rightarrow 1'\hat{\underset{\sim}{\varepsilon}} = 0 \Rightarrow \Sigma y_j - \Sigma \hat{y}_j = 0$$
$$\Rightarrow \bar{y} = \bar{\hat{y}}$$

$$\Rightarrow \underset{\sim}{y}'\underset{\sim}{y} - n\bar{y}^2 = \hat{\underset{\sim}{y}}'\hat{\underset{\sim}{y}} - n\bar{y}^2 + \hat{\underset{\sim}{\varepsilon}}'\hat{\underset{\sim}{\varepsilon}}$$

$$\Rightarrow \underbrace{\sum_j (y_j - \bar{y})^2}_{\substack{\text{SSTOTAL} \\ \text{Corrected}}} = \underbrace{\sum_j (\hat{y}_j - \bar{y})^2}_{\text{SSRegression}} + \underbrace{\sum_j (y_j - \hat{y}_j)^2}_{\substack{\text{SSResidual} \\ \text{(Error)}}}$$

Geometry

$Y \equiv$ response vector in n Dim space

$Z\beta \equiv$ vector in model Plane (Straight line when $r+1=2$)

$\varepsilon \equiv$ error vector

$\hat{y} = Z\hat{\beta}$ Projects $\underset{\sim}{y}$ onto plane consisting of all linear combinations of columns of Z

$\underset{\sim}{\hat{\varepsilon}} = \underset{\sim}{y} - \underset{\sim}{\hat{y}}$ is perpendicular to the plane

Spectral Decomposition of $Z'Z$ (Z = full rank)

$$Z'Z = \lambda_1 \underset{\sim}{e}_1 \underset{\sim}{e}_1' + \dots + \lambda_{r+1} \underset{\sim}{e}_{r+1} \underset{\sim}{e}_{r+1}' \qquad \lambda_i \geq 0 \;\forall i$$

$$(Z'Z)^{-1} = \frac{1}{\lambda_1} \underset{\sim}{e}_1 \underset{\sim}{e}_1' + \dots + \frac{1}{\lambda_{r+1}} \underset{\sim}{e}_{r+1} \underset{\sim}{e}_{r+1}'$$

$$Z(Z'Z)^{-1}Z' = \sum_{i=1}^{r+1} \lambda_i^{-1} Z \cdot \underset{\sim}{e}_i \underset{\sim}{e}_i' Z' = \sum_{i=1}^{r+1} \underset{\sim}{q}_i \underset{\sim}{q}_i'$$

Result 2A.2
Def 2A.12: Projection of y on a linear combination of $\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_{r+1}$:

$$\sum_{i=1}^{r+1} (q_i' y) q_i = \left(\sum_{i=1}^{r+1} \underset{\sim}{q}_i \underset{\sim}{q}_i' \right) \underset{\sim}{y} = Z(Z'Z)^{-1} Z' \underset{\sim}{y} = Z\hat{\beta}$$

$I - Z(Z'Z)^{-1}Z'$ projects $\underset{\sim}{y}$ onto plane perpendicular to plane spaned by columns of Z.

Sampling Properties of Least Squares Estimators

$\hat{\underset{\sim}{\beta}} = (Z'Z)^{-1}Z'\underset{\sim}{Y}$ $\qquad E\{\underset{\sim}{Y}\} = Z\beta \quad V\{\underset{\sim}{Y}\} = \sigma^2 I$

$\hat{\underset{\sim}{\varepsilon}} = [I - Z(Z'Z)^{-1}Z']\underset{\sim}{Y}$

$E\{\hat{\underset{\sim}{\beta}}\} = (Z'Z)^{-1}Z'Z\underset{\sim}{\beta} = \underset{\sim}{\beta}$

$V\{\hat{\underset{\sim}{\beta}}\} = (Z'Z)^{-1}Z'\sigma^2 I\left((Z'Z)^{-1}Z'\right)' = (Z'Z)^{-1}Z'\sigma^2 I Z(Z'Z)^{-1}$

$= \sigma^2(Z'Z)^{-1}Z'Z(Z'Z)^{-1} = \sigma^2(Z'Z)^{-1}$

$E\{\hat{\underset{\sim}{\varepsilon}}\} = Z\beta - Z(Z'Z)^{-1}Z'Z\beta = Z\beta - Z\beta = 0$

$V\{\hat{\underset{\sim}{\varepsilon}}\} = \left[I - Z(Z'Z)^{-1}Z'\right]\sigma^2 I\left[I - Z(Z'Z)^{-1}Z'\right]'$

$= \sigma^2\left(I - Z(Z'Z)^{-1}Z'\right)$ $\qquad$ (symmetric, idempotent)

$E\{\hat{\underset{\sim}{\varepsilon}}'\hat{\underset{\sim}{\varepsilon}}\} = [n-(r+1)]\sigma^2 \Rightarrow s^2 = \dfrac{\hat{\underset{\sim}{\varepsilon}}'\hat{\varepsilon}}{n-(r+1)} = \dfrac{\underset{\sim}{Y}'(I-H)\underset{\sim}{Y}}{n-(r+1)}$

$\mathrm{Cov}\{\hat{\beta}, \hat{\varepsilon}\} = \mathrm{Cov}\{HY, (I-H)Y\} = H\sigma^2 I(I-H)'$

$= \sigma^2 H(I-H) = \sigma^2(H-H) = 0$

<u>Gauss Least Squares Theorem</u> $\quad$ for any vector $\underset{\sim}{c}$ of fixed constants

$c'\hat{\beta} = c_0\hat{\beta}_0 + c_1\hat{\beta}_1 + \cdots + c_r\hat{\beta}_r$ $\quad$ has smallest variance

of ~~all~~ linear estimators of for $\underset{\sim}{a}'\underset{\sim}{Y} = a_1Y_1 + \cdots + a_nY_n$

7.4 Inferences Regarding Model

$\underset{\sim}{\varepsilon} \sim MVN(\underset{\sim}{0}, \sigma^2 I)$ — Same previous assumptions with normality added

$$E\{Y\} = \beta_0 + \beta_1 z_1 + \dots + \beta_r z_r$$

$$\Rightarrow \hat{\beta} = (Z'Z)^{-1} Z' \underset{\sim}{Y} \sim N_{r+1}\left(\underset{\sim}{\beta}, \sigma^2 (Z'Z)^{-1}\right)$$

$$n\tilde{\sigma}^2 = \hat{\varepsilon}'\hat{\varepsilon} \sim \sigma^2 \chi^2_{n-(r-1)}$$

$\hookrightarrow \hat{\sigma}^2 \equiv$ MLE for σ^2: $\hat{\sigma}^2 = \dfrac{\hat{\varepsilon}'\hat{\varepsilon}}{n} = \dfrac{n-(r+1)}{n} s^2$

Estimated Variance-Covariance matrix for $\hat{\beta}$:

$$s^2_{\hat{\beta}} = \hat{V}\{\hat{\beta}\} = s^2 (Z'Z)^{-1}$$

100(1−α)% Confidence Ellipsoid for β:

$$\beta: \quad \frac{(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})' \left[s^2 (Z'Z)^{-1}\right]^{-1} (\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})}{r+1} \le F_{r+1, n-(r+1)}(\alpha)$$

$$\Rightarrow (\beta - \hat{\beta})' (Z'Z) (\beta - \hat{\beta}) \le s^2 (r+1) F_{r+1, n-(r+1)}(\alpha)$$

Simultaneous 100(1−α)% CI's for β_i

$$\hat{\beta}_i \pm \sqrt{\hat{V}\{\hat{\beta}_i\}} \sqrt{(r+1) F_{r+1, n-(r+1)}(\alpha)}$$

$\hookrightarrow$ diagonal elements of $s^2 (Z'Z)^{-1}$

One-at-a-time Intervals for β_i

$$\hat{\beta}_i \pm t_{n-(r+1)}^{(\alpha/2)} \sqrt{\hat{V}\{\hat{\beta}_i\}}$$ (Standard computer output)

Likelihood Ratio Tests for Regression Parameters

$$\underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_q \\ \hline \beta_{q+1} \\ \vdots \\ \beta_r \end{bmatrix} = \begin{bmatrix} \underset{\sim}{\beta}_{(1)} \\ \hline \underset{\sim}{\beta}_{(2)} \end{bmatrix} \quad Z = \begin{bmatrix} \underset{n\times(q+1)}{Z_1} & \vdots & \underset{n\times(r-q)}{Z_2} \end{bmatrix}$$

$H_0: \beta_{q+1} = \cdots \beta_r = 0 \qquad H_A: \beta_{(2)} \neq \underset{\sim}{0}$

$\beta_{(2)} = \underset{\sim}{0}$

$H_0: \quad \underset{\sim}{Y} = Z_1 \beta_{(1)} + \underset{\sim}{\varepsilon} \Rightarrow \hat{\beta}_{(1)} = (Z_1' Z_1)^{-1} Z_1' \underset{\sim}{Y}$

$SS_{Res}(Z_1) = (\underset{\sim}{Y} - Z_1 \hat{\beta}_{(1)})'(\underset{\sim}{Y} - Z_1 \hat{\beta}_{(1)})$

$HA: \quad Y = Z\beta + \varepsilon \Rightarrow \hat{\beta} = (Z'Z)^{-1} Z' \underset{\sim}{Y}$

$SS_{Res}(Z) = (\underset{\sim}{Y} - Z\hat{\beta})'(\underset{\sim}{Y} - Z\hat{\beta})$

Extra Sum of Squares due to Z_2: $\quad SS_{Res}(Z_1) - SS_{Res}(Z)$

$df(Z_1) = n-(q+1) \qquad df(Z) = n-(r+1)$

$df(Z_1) - df(Z_2) = r-q$

LR Test Stat: $$F_{obs} = \frac{\dfrac{SS_{Res}(Z_1) - SS_{Res}(Z)}{r-q}}{S^2}$$

Reject H_0 if $F_{obs} \geq F_{r-q,\, n-(r+1)}(\alpha)$

General Linear Test

$H_0: C'\underset{\sim}{\beta} = \underset{\sim}{m} \Rightarrow C'\underset{\sim}{\beta} - \underset{\sim}{m} = \underset{\sim}{0}$

T.S. $F_{obs} = \dfrac{(C'\hat{\underset{\sim}{\beta}} - \underset{\sim}{m})'(C'(Z'Z)^{-1}C)^{-1}(C'\hat{\beta} - m)}{s^2 \times r(C)}$

(# of rows of C)
"restrictions"

Reject H_0 if $F_{obs} \geq F_{r(C), n-(r+1)}^{(\alpha)}$

7.5 Inferences Involving the Estimated Reg function

Estimating mean response @ $\underset{\sim}{z} = \underset{\sim}{z_0} = \begin{bmatrix} 1 \\ z_{01} \\ \vdots \\ z_{0r} \end{bmatrix}$

Parameter: $\underset{\sim}{z_0}'\underset{\sim}{\beta}$ Estimator: $\underset{\sim}{z_0}'\hat{\beta}$

$V\{\underset{\sim}{z_0}'\hat{\underset{\sim}{\beta}}\} = z_0'\sigma^2(Z'Z)^{-1}z_0 = \sigma^2 z_0'(Z'Z)^{-1}z_0$

$\hat{V}\{z_0'\hat{\beta}\} = s^2 z_0'(Z'Z)^{-1}z_0$

(1-α)100% CI for $z_0'\beta$:

$z_0'\hat{\beta} \pm t_{n-(r+1)}^{(\alpha/2)}\sqrt{s^2 z_0'(Z'Z)^{-1}z_0}$

Forecasting ~~a~~ new observation @ z_0

New response $\equiv Y_0 = z_0'\beta + \varepsilon_0$

forecast $\equiv z_0'\hat{\beta}$ forecast error $\equiv Y_0 - z_0'\hat{\beta}$

$$V\{Y_0 - z_0'\hat{\beta}\} = V\{Y_0\} + V\{z_0'\hat{\beta}\} - 2\underbrace{\text{Cov}\{Y_0, z_0'\hat{\beta}\}}_{0}$$

$$= \sigma^2 + \sigma^2 \underset{\sim}{z_0}'(Z'Z)^{-1}\underset{\sim}{z_0} = \sigma^2\left[1 + z_0'(Z'Z)^{-1}z_0\right]$$

$\Rightarrow$ $(1-\alpha)100\%$ Prediction Interval for Y_0:

$$\underset{\sim}{z_0}'\underset{\sim}{\hat{\beta}} \pm t_{n-(r+1)}\left(\tfrac{\alpha}{2}\right)\sqrt{s^2\left(1 + z_0'(Z'Z)^{-1}z_0\right)}$$

7.6 Model Checking and Other Aspects

Residuals, Studentized Residuals

Residual Plots - Linearity, Constant variance, unused predictors

NP Plots, residuals vs Time D-W test

Leverage, Influence, Model Building, Collinearity

Bias from misspecified models ($\underset{\sim}{\beta}, \sigma^2$)

7.7 – Multivariate Multiple Regression

m Response Variables, single set of r predictors

Characteristic

$$1 \quad Y_1 = \beta_{01} + \beta_{11} z_1 + \cdots + \beta_{r1} z_r + \varepsilon_1$$

$$\vdots$$

$$m \quad Y_m = \beta_{0m} + \beta_{1m} z_1 + \cdots + \beta_{rm} z_r + \varepsilon_m$$

$$\underset{\sim}{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_m \end{bmatrix} \quad E\{\underset{\sim}{\varepsilon}\} = \underset{\sim}{0} \quad V\{\underset{\sim}{\varepsilon}\} = \Sigma$$

j^{th} Trial: $z_{j0}, z_{j1}, \ldots, z_{jr}$ $\quad$ ($z_{j0} = 1$ for intercept)

$$\underset{\sim}{Y}_j' = [Y_{j1}, Y_{j2}, \ldots, Y_{jm}]$$

$$\underset{n\times(r+1)}{Z} = \begin{bmatrix} z_{10} & z_{11} & \cdots & z_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n0} & z_{n1} & \cdots & z_{nr} \end{bmatrix} \quad \underset{n\times m}{Y} = \begin{bmatrix} Y_{11} & \cdots & Y_{1m} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \cdots & Y_{nm} \end{bmatrix}$$

$$Y = \left[\underset{\sim}{Y}_{(1)} \;\vdots\; \cdots \;\vdots\; \underset{\sim}{Y}_{(m)} \right]$$

$$\underset{(r+1)\times m}{\beta} = \begin{bmatrix} \beta_{01} & \cdots & \beta_{0m} \\ \beta_{11} & \cdots & \beta_{1m} \\ \vdots & & \\ \beta_{r1} & \cdots & \beta_{rm} \end{bmatrix} = \left[\beta_{(1)} \;\vdots\; \cdots \;\vdots\; \beta_{(m)} \right]$$

$$\underset{\substack{\sim \\ n\times m}}{\varepsilon} = \begin{bmatrix} \varepsilon_{11} & \cdots & \varepsilon_{1m} \\ \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \cdots & \varepsilon_{nm} \end{bmatrix} = \left[\underset{\sim}{\varepsilon}_{(1)} \;\vdots\; \cdots \;\vdots\; \varepsilon_{(m)} \right]$$

$$\underset{n\times m}{Y} = \underset{n\times(r+1)}{Z} \; \underset{(r+1)\times m}{\beta} + \underset{n\times m}{\varepsilon}$$

$E\{\varepsilon_{(i)}\} = 0 \quad Cov\{\varepsilon_{(i)}, \varepsilon_{(k)}\} = \sigma_{ik} I$

$i, k = 1, \ldots, m$

$$\underset{\sim}{Y}_{(i)} = Z \underset{\sim}{\beta}_{(i)} + \underset{\sim}{\varepsilon}_{(i)} \quad i=1,\dots,m$$

$$\mathrm{Cov}\{\varepsilon_{(i)}\} = \sigma_{ii} I$$

Different responses on the same trial/unit can be correlated.

$$\hat{\beta}_{(i)} = (Z'Z)^{-1} Z' Y_{(i)}$$

$$\hat{\beta} = [\hat{\beta}_{(1)} \vdots \cdots \vdots \hat{\beta}_{(m)}] = (Z'Z)^{-1} Z' [\underset{\sim}{Y}_{(1)} \vdots \cdots \vdots \underset{\sim}{Y}_{(m)}]$$

$$= (Z'Z)^{-1} Z' Y$$

For any choice of Parameters: $B = [b_{(1)} \vdots \cdots \vdots b_{(m)}]$,

matrix of Errors $\equiv Y - ZB$

Error Sums of Squares and Crossproducts Matrix:

$$(Y-ZB)'(Y-ZB) = \begin{bmatrix} (Y_{(1)}-Zb_{(1)})'(Y_{(1)}-Zb_{(1)}) & \cdots & (Y_{(1)}-Zb_{(1)})'(Y_{(m)}-Zb_{(m)}) \\ \vdots & & \\ (Y_{(m)}-Zb_{(m)})'(Y_{(1)}-Zb_{(1)}) & \cdots & (Y_{(m)}-Zb_{(m)})'(Y_{(m)}-Zb_{(m)}) \end{bmatrix}$$

$b_{(i)} = \hat{\beta}_{(i)}$ minimizes i^{th} diagonal sum of squares

$\Rightarrow tr[(Y-ZB)'(Y-ZB)]$ minimized at $B = \hat{\beta}$

Also Generalized Variance $|(Y-ZB)'(Y-ZB)|$ minimized.

$$\hat{\beta} = (Z'Z)^{-1}Z'Y = (Z'Z)^{-1}Z'\left[Y_{(1)} \,\vdots\, \cdots \,\vdots\, Y_{(m)}\right]$$

<u>Predicted Values</u>: $\hat{Y} = Z\hat{\beta} = Z(Z'Z)^{-1}Z'Y$

$$= \left[\hat{Y}_{(1)} \,\vdots\, \cdots \,\vdots\, \hat{Y}_{(m)}\right]$$

<u>Residuals</u>: $\hat{\varepsilon} = Y - \hat{Y} = (I - Z(Z'Z)^{-1}Z')Y = \left(\hat{\varepsilon}_{(1)} \,\vdots\, \cdots \,\vdots\, \hat{\varepsilon}_{(m)}\right)$

$$Z'\hat{\varepsilon} = \hat{Y}'\hat{\varepsilon} = 0 \Rightarrow \hat{Y}_{(i)} \perp \hat{\varepsilon}_{(k)} \quad \forall\, i,k$$

$$Y = \hat{Y} + \hat{\varepsilon} \Rightarrow \underbrace{Y'Y}_{m \times m} = (\hat{Y} + \hat{\varepsilon})'(\hat{Y} + \hat{\varepsilon}) = \hat{Y}'\hat{Y} + \hat{\varepsilon}'\hat{\varepsilon} + 0 + 0'$$

$$Y'Y = \hat{Y}'\hat{Y} + \hat{\varepsilon}'\hat{\varepsilon}$$

Total sum of squares and cross-products	Predicted SSCP	Residual (Error) SSCP

$$\hat{\varepsilon}'\hat{\varepsilon} = Y'Y - \hat{Y}'\hat{Y} = Y'Y - \hat{\beta}'Z'Z\hat{\beta}$$

$$= Y'Y - Y'Z(Z'Z)^{-1}Z Z(Z'Z)^{-1}Z'Y = Y'Y - Y'Z(Z'Z)^{-1}Z'Y$$

Result 7.9 $\quad \hat{\beta} = [\hat{\beta}_{(1)} \mid \cdots \mid \hat{\beta}_{(m)}]$

$$E\{\hat{\beta}_{(i)}\} = \beta_{(i)} \qquad E\{\hat{\beta}\} = \beta$$

$$\mathrm{Cov}\{\hat{\beta}_{(i)}, \hat{\beta}_{(k)}\} = \sigma_{ik}(Z'Z)^{-1} \qquad i,k = 1,\dots,m$$

$$\hat{\varepsilon} = [\hat{\varepsilon}_1 \mid \cdots \mid \hat{\varepsilon}_m] = Y - Z\hat{\beta}$$

$$E\{\hat{\varepsilon}_{(i)}\} = 0 \qquad E\{\hat{\varepsilon}_{(i)}' \hat{\varepsilon}_{(k)}\} = (n-(r+1))\,\sigma_{ik}$$

$$\Rightarrow E\{\hat{\varepsilon}\} = 0 \qquad E\left\{\frac{1}{n-(r+1)} \hat{\varepsilon}'\hat{\varepsilon}\right\} = \Sigma$$

Proof: $Y_{(i)} = Z\beta_{(i)} + \varepsilon_{(i)} \qquad E\{\varepsilon_{(i)}\} = 0 \quad E\{\varepsilon_{(i)}\varepsilon_{(i)}'\} = \sigma_{ii} I$

$$\hat{\beta}_{(i)} - \beta_{(i)} = (Z'Z)^{-1}Z'Y_{(i)} - \beta_{(i)}$$

$$= (Z'Z)^{-1}Z'[Z\beta_{(i)} + \varepsilon_{(i)}] - \beta_{(i)} = (Z'Z)^{-1}Z'\varepsilon_{(i)}$$

$$\hat{\varepsilon}_{(i)} = Y_{(i)} - \hat{Y}_{(i)} = [I - Z(Z'Z)^{-1}Z']Y_{(i)} = [I - Z(Z'Z)^{-1}Z']\varepsilon_{(i)}$$

$$\Rightarrow E\{\hat{\beta}_{(i)}\} = \beta_{(i)} \qquad E\{\hat{\varepsilon}_{(i)}\} = 0$$

$$\mathrm{Cov}\{\hat{\beta}_{(i)}, \hat{\beta}_{(k)}\} = E\{(\hat{\beta}_{(i)} - \beta_{(i)})(\hat{\beta}_{(k)} - \beta_{(k)})'\}$$

$= E\left\{\left((Z'Z)^{-1}Z'\varepsilon_{(i)}\right)\left((Z'Z)^{-1}Z'\varepsilon_{(k)}\right)'\right\}$

$= E\left\{(Z'Z)^{-1}Z'\varepsilon_{(i)}\varepsilon_{(k)}'Z(Z'Z)^{-1}\right\}$

$= (Z'Z)^{-1}Z'\underbrace{E\left\{\varepsilon_{(i)}\varepsilon_{(k)}'\right\}}_{\sigma_{ik}I}Z(Z'Z)^{-1}$

$= \sigma_{ik}(Z'Z)^{-1}Z'Z(Z'Z)^{-1} = \sigma_{ik}(Z'Z)^{-1}$

$U \equiv$ Random vector $A \equiv$ fixed matrix

$E\{U'AU\} = E\{tr(AUU')\} = tr(A\,E\{UU'\})$

$\Rightarrow E\{\hat{\underset{\sim}{\varepsilon}}_{(i)}'\hat{\underset{\sim}{\varepsilon}}_{(k)}\} = E\{\underset{\sim}{\varepsilon}_{(i)}'(I - Z(Z'Z)^{-1}Z')\varepsilon_{(k)}\}$

$= tr\left[(I - Z(Z'Z)^{-1}Z')\,\sigma_{ik}I\right]$

$= \sigma_{ik}\,tr\left(I - Z(Z'Z)^{-1}Z'\right) = \sigma_{ik}\left\{tr(I) - tr(Z(Z'Z)^{-1}Z')\right\}$

$= \sigma_{ik}\left[n - tr((Z'Z)(Z'Z)^{-1})\right] = \sigma_{ik}\left[n - tr(I_{r+1})\right]$

$= \sigma_{ik}\left(n-(r+1)\right)$

$\Rightarrow E\left\{\frac{1}{n-(r+1)}\hat{\underset{\sim}{\varepsilon}}'\hat{\underset{\sim}{\varepsilon}}\right\} = \Sigma$

$$\mathrm{Cov}\{\hat{\beta}_{(i)}, \hat{\varepsilon}_{(k)}\}$$

$$= E\{(\hat{\beta}_{(i)} - \beta_{(i)})(\hat{\varepsilon}_{(k)} - \varepsilon_{(k)})'\}$$

$$= E\{((Z'Z)^{-1}Z'\varepsilon_{(i)})((I - Z(Z'Z)^{-1}Z')\varepsilon_{(k)})'\}$$

$$= E\{(Z'Z)^{-1}Z' \varepsilon_{(i)} \varepsilon_{(k)}' (I - Z(Z'Z)^{-1}Z')\}$$

$$= (Z'Z)^{-1}Z' \sigma_{ik} I (I - Z(Z'Z)^{-1}Z')$$

$$= \sigma_{ik}\{(Z'Z)^{-1}Z' - (Z'Z)^{-1}Z'Z(Z'Z)^{-1}Z'\} = \sigma_{ik}\{(Z'Z)^{-1}Z' - (Z'Z)^{-1}Z'\}$$

$$= 0$$

Estimating mean vector @ $\underset{\sim}{z_0}$

Parameter : $z_0'\beta_{(i)}$ $\quad i = 1, \dots, m$

Estimator : $z_0'\hat{\beta}_{(i)}$ $\quad E\{z_0'\hat{\beta}_{(i)}\} = z_0'\beta_{(i)}$

$$V\{z_0'\hat{\beta}_{(i)} - z_0'\hat{\beta}_{(i)}\} = E\{z_0'(\hat{\beta}_{(i)} - \hat{\beta}_{(i)})(\hat{\beta}_{(k)} - \hat{\beta}_{(k)})'z_0\}$$

$$= z_0'[E\{(\hat{\beta}_{(i)} - \hat{\beta}_{(i)})(\hat{\beta}_{(k)} - \hat{\beta}_{(k)})'\}]z_0$$

$$= z_0'[E\{(Z'Z)^{-1}Z'\varepsilon_{(i)} \varepsilon_{(k)}' Z(Z'Z)^{-1}\}]z_0$$

$$= \sigma_{ik}\, \underline{z_0}'(Z'Z)^{-1}\underline{z_0}$$

Forecasting Y_0 @ $z = z_0$

forecast error $Y_{0i} - z_0'\hat{\beta}_{(i)}$

$$= \left(Y_{0i} - z_0'\beta_{(i)}\right) + \left(z_0'\beta_{(i)} - z_0'\hat{\beta}_{(i)}\right)$$

$$= \varepsilon_{0i} - z_0'\left(\hat{\beta}_{(i)} - \beta_{(i)}\right)$$

$$E\{Y_{0i} - z_0'\hat{\beta}_{(i)}\} = 0$$

$$E\left\{\left(Y_{0i} - z_0'\hat{\beta}_{(i)}\right)\left(Y_{0k} - z_0'\hat{\beta}_{(k)}\right)\right\}$$

$$= E\left\{\left(\varepsilon_{0i} - z_0'(\hat{\beta}_{(i)} - \beta_{(i)})\right)\left(\varepsilon_{0k} - z_0'(\hat{\beta}_{(k)} - \beta_{(k)})\right)\right\}$$

$$= E\{\varepsilon_{0i}\varepsilon_{0k}\} + z_0' E\{(\hat{\beta}_{(i)} - \beta_{(i)})(\hat{\beta}_{(k)} - \beta_{(k)})'\} z_0$$

$$- z_0' \underbrace{E\{(\hat{\beta}_{(i)} - \beta_{(i)})\varepsilon_{0k}\}}_{0} - E\{\underbrace{\varepsilon_{0i}(\hat{\beta}_{(k)} - \beta_{(k)})'}_{0}\} z_0$$

$$= \sigma_{ik}\left[1 + z_0'(Z'Z)^{-1}z_0\right]$$

Result 7.10 multivariate multiple linear reg model w/ full rank $(Z) = r+1$, $n \geq (r+1) + m$

$\underset{\sim}{\varepsilon} \sim$ Normal

$$\Rightarrow \hat{\beta} = (Z'Z)^{-1} Z' Y \equiv \text{MLE of } \beta$$

$$\hat{\beta} \sim N\left(\beta, \ \mathrm{Cov}\{\hat{\beta}_{(i)}, \hat{\beta}_{(k)}\} = \sigma_{ik}(Z'Z)^{-1}\right)$$

$$\hat{\Sigma} = \frac{1}{n} \hat{\underset{\sim}{\varepsilon}}' \hat{\underset{\sim}{\varepsilon}} = \frac{1}{n}\left(Y - Z\hat{\beta}\right)'\left(Y - Z\hat{\beta}\right) \equiv \text{MLE of } \Sigma$$

$$n\hat{\Sigma} \sim \text{Wishart}_{p,\, n-(r+1)}(\Sigma) \qquad \hat{\beta}, \hat{\Sigma}$$

Maximized Likelihood: $L(\hat{\mu}, \hat{\Sigma}) = (2\pi)^{-mn/2} |\hat{\Sigma}|^{-n/2} e^{-mn/2}$

Likelihood Ratio Tests for Parameters

$$\beta = \begin{bmatrix} \underset{(q+1)\times m}{\beta_{(1)}} \\ \hline \underset{(r-q)\times m}{\beta_{(2)}} \end{bmatrix} \qquad H_0: \beta_{(2)} = 0 \qquad Z = \left[\underset{n\times(q+1)}{Z_1} \ \vdots \ \underset{n\times(r-q)}{Z_2}\right]$$

$$E\{Y\} = Z\beta = \left[Z_1 \ \vdots \ Z_2\right] \begin{bmatrix} \beta_1 \\ \hline \beta_2 \end{bmatrix} = Z_1\beta_1 + Z_2\beta_2$$

Extra SSCP: $\left(Y - Z_1\hat{\beta}_{(1)}\right)'\left(Y - Z_1\hat{\beta}_{(1)}\right) - \left(Y - Z\hat{\beta}\right)'\left(Y - Z\hat{\beta}\right)$

$$= n\left(\hat{\Sigma}_1 - \hat{\Sigma}\right) \qquad \hat{\beta}_{(1)} = (Z_1'Z_1)^{-1} Z_1' Y$$

$$\hat{\Sigma}_1 = n^{-1}\left(Y - Z_1\hat{\beta}_{(1)}\right)'\left(Y - Z_1\hat{\beta}_{(1)}\right)$$

Likelihood Ratio: $\Lambda = \dfrac{\max_{\beta_{(1)}, \Sigma} L(\beta_{(1)}, \Sigma)}{\max_{\beta, \Sigma} L(\beta, \Sigma)} = \dfrac{L(\hat{\beta}_{(1)}, \hat{\Sigma}_1)}{L(\hat{\beta}, \hat{\Sigma})} = \left(\dfrac{|\hat{\Sigma}|}{|\hat{\Sigma}_1|}\right)^{n/2}$

Wilk's Lambda Statistic: $\Lambda^{2/n} = \dfrac{|\hat{\Sigma}|}{|\hat{\Sigma}_1|}$

$H_0: \beta_{(2)} = 0$

T.S. $-2\ln\Lambda = -n\ln\left(\dfrac{|\hat{\Sigma}|}{|\hat{\Sigma}_1|}\right)$ (Reject if Large)

$n-r$, $n-m$ large

$\Rightarrow -\left[n-r-1-\frac{1}{2}(m-r+q+1)\right]\ln\left(\dfrac{|\hat{\Sigma}|}{|\hat{\Sigma}_1|}\right) \,\dot\sim\, \chi^2_{m(r-q)}$

Information criteria

Model w/ d predictors: $\hat{\Sigma}_d = \frac{1}{n} SSCP_{Res}$

$AIC = n\ln(|\hat{\Sigma}_d|) - 2pd$

Other Test Statistics for $H_0: \beta_{(2)} = 0$

$p = m = \#$ response vars

$E = n\hat{\Sigma}$ from full model

$H = n(\hat{\Sigma}_1 - \hat{\Sigma})$ difference between Reduced and full model.

Eigenvalues of HE^{-1}: $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_s$ $s = \min(p, r-q)$
$\uparrow$
$m=p$

Wilks' Lambda $= \prod_{i=1}^{s} \frac{1}{1+\eta_i} = \frac{|E|}{|E+H|}$

Pillai's Trace $= \sum_{i=1}^{s} \frac{\eta_i}{1+\eta_i} = tr\left[H(H+E)^{-1}\right]$

Hotelling-Lawley Trace $= \sum_{i=1}^{s} \eta_i = tr(HE^{-1})$

Roy's Greatest root $= \frac{\eta_1}{1+\eta_1}$

Predictions from Multivariate Multiple Regression

$$\hat{\beta}' z_0 \sim N_m\left(\beta' z_0,\ z_0'(Z'Z)^{-1} z_0 \Sigma\right)$$

$n\hat{\Sigma}$ independently $\sim \text{Wishart}_{n-r-1}(\Sigma)$

$$T^2 = \left(\frac{\hat{\beta}' z_0 - \beta' z_0}{\sqrt{z_0'(Z'Z)^{-1} z_0}}\right)' \left(\frac{n}{n-r-1}\hat{\Sigma}\right)^{-1} \left(\frac{\hat{\beta}' z_0 - \beta' z_0}{\sqrt{z_0'(Z'Z)^{-1} z_0}}\right)$$

$(1-\alpha)100\%$ Confidence ellipsoid for $\beta' z_0$

$$\beta' z_0 : \quad T^2 \leq \frac{m(n-r-1)}{n-r-m} F_{m, n-r-m}(\alpha)$$

$(1-\alpha)100\%$ Simultaneous CI's for $E[Y_i] = z_0'\beta_{(i)}$:

$$z_0'\hat{\beta}_{(i)} \pm \sqrt{\frac{m(n-r-1)}{n-r-m} F_{m,n-r-m}(\alpha)} \sqrt{z_0'(Z'Z)^{-1}z_0\left(\frac{n}{n-r-1}\right)\hat{\sigma}_{ii}}$$

$i = 1, \ldots, m$

$\hat{\beta}_{(i)} \equiv i^{th}$ column of $\hat{\beta}$, $\hat{\sigma}_{ii} \equiv i^{th}$ diagonal element of $\hat{\Sigma}$

$(1-\alpha)100\%$ Prediction ellipsoid for Y_0

$$Y_0 : \left(Y_0 - \hat{\beta}'z_0\right)'\left(\frac{n}{n-r-1}\hat{\Sigma}\right)^{-1}\left(Y_0 - \hat{\beta}'z_0\right)$$

$$\leq \left(1 + z_0'(Z'Z)^{-1}z_0\right)\left[\left(\frac{m(n-r-1)}{n-r-m}\right) F_{m,n-r-m}(\alpha)\right]$$

$(1-\alpha)100\%$ Simultaneous prediction intervals for Y_{0i}

$$z_0'\hat{\beta}_{(i)} \pm \sqrt{\left(\frac{m(n-r-1)}{n-r-m}\right) F_{m,n-r-m}(\alpha)} \sqrt{\left(1 + z_0'(Z'Z)^{-1}z_0\right)\left(\frac{n}{n-r-1}\hat{\sigma}_{ii}\right)}$$

$i = 1, \ldots, m$

7.8 Concept of Linear Regression

$Y \equiv$ random variable $\quad z_1, \dots, z_r \equiv$ random (not fixed)

$$\underset{\sim}{\mu} = \begin{bmatrix} \underset{1\times 1}{\mu_Y} \\ \hline \underset{r\times 1}{\mu_{\underset{\sim}{z}}} \end{bmatrix} \qquad \Sigma = \left[\begin{array}{c|c} \underset{1\times 1}{\sigma_{YY}} & \underset{1\times r}{\sigma'_{\underset{\sim}{z}Y}} \\ \hline \underset{r\times 1}{\sigma_{\underset{\sim}{z}Y}} & \underset{r\times r}{\Sigma_{\underset{\sim}{z}\underset{\sim}{z}}} \end{array}\right]$$

Predicting Y: $\quad b_0 + b_1 z_1 + \dots + b_r z_r = b_0 + \underset{\sim}{b}'\underset{\sim}{z}$

prediction Error: $\quad Y - (b_0 + b_1 z_1 + \dots + b_r z_r) = Y - b_0 - \underset{\sim}{b}'\underset{\sim}{z}$

$MSE = \quad E\{(Y - b_0 - \underset{\sim}{b}'\underset{\sim}{z})^2\}$ $\quad$ depends on joint dist'n of Y and $\underset{\sim}{z}$ only through $\underset{\sim}{\mu}, \Sigma$

Result 7.12

Linear Predictor of Y: $\beta_0 + \beta' z$ w/ Coefficients:

$$\beta = \Sigma_{\underset{\sim}{z}\underset{\sim}{z}}^{-1}\sigma_{\underset{\sim}{z}Y}, \quad \beta_0 = \mu_Y - \underset{\sim}{\beta}'\underset{\sim}{z}$$

has minimum MSE of all linear predictors of Y.

w/ $E\{(Y - \beta_0 - \beta' z)^2\} = \sigma_{YY} - \sigma'_{zY}\Sigma_{zz}^{-1}\sigma_{zY}$

Also $\beta_0 + \beta_0' z = \mu_Y + \sigma'_{zY}\Sigma_{zz}^{-1}\sigma_{zY}$ is linear predictor w/ max correlation w/ Y

$$\text{Corr}(Y, \beta_0 + \underset{\sim}{\beta}'\underset{\sim}{z}) = \max_{b_0, \underset{\sim}{b}} \text{Corr}(Y, b_0 + \underset{\sim}{b}'\underset{\sim}{z})$$

$$= \sqrt{\frac{\beta'\Sigma_{zz}\beta}{\sigma_{YY}}} = \sqrt{\frac{\sigma'_{zY}\Sigma_{zz}^{-1}\sigma_{zY}}{\sigma_{YY}}}$$

Population ~~mean~~ multiple Coefficient of Correlation

$$\rho_{Y(z)} = +\sqrt{\frac{\sigma_{ZY}' \Sigma_{ZZ}^{-1} \sigma_{ZY}}{\sigma_{YY}}} \qquad 0 \le \rho_{Y(z)} \le 1$$

$\rho^2_{Y(z)} \equiv$ Population coefficient of determination

$$MSE(\beta_0 + \beta' z) = \sigma_{YY} - \sigma_{ZY}' \Sigma_{ZZ}^{-1} \sigma_{ZY} = \sigma_{YY} - \sigma_{YY}\left(\frac{\sigma_{ZY}' \Sigma_{ZZ}^{-1} \sigma_{ZY}}{\sigma_{YY}}\right)$$

$$= \sigma_{YY}\left(1 - \rho^2_{Y(z)}\right)$$

Result 7.13 Random Sample of Size n from $Y, \underset{\sim}{z}$

$$Y, \underset{\sim}{z} \sim N_{r+1}(\underset{\sim}{\mu}, \Sigma)$$

$$\hat{\mu} = \begin{bmatrix} \bar{Y} \\ \hline \bar{\underset{\sim}{z}} \end{bmatrix} \qquad S = \left[\begin{array}{c|c} S_{YY} & S_{ZY}' \\ \hline S_{ZY} & S_{ZZ} \end{array}\right]$$

$$\hat{\underset{\sim}{\beta}} = S_{ZZ}^{-1} S_{ZY} \qquad \hat{\beta}_0 = \bar{Y} - S_{ZY}' S_{ZZ}^{-1} \bar{\underset{\sim}{z}} = \bar{Y} - \hat{\underset{\sim}{\beta}}' \bar{\underset{\sim}{z}}$$

MLE of Regression Function: $\hat{\beta}_0 + \hat{\beta}' \underset{\sim}{z} = \bar{Y} + S_{ZY}' S_{ZZ}^{-1}(\underset{\sim}{z} - \bar{\underset{\sim}{z}})$

MLE of $MSE = E\{(Y - \beta_0 - \beta' z)^2\} =$

$$\hat{\sigma}_{YY.z} = \frac{n-1}{n}\left(S_{YY} - S_{ZY}' S_{ZZ}^{-1} S_{ZY}\right)$$

unbiased estimator: $\dfrac{n-1}{n-(r+1)}\left(S_{YY} - S_{ZY}' S_{ZZ}^{-1} S_{ZY}\right) = \dfrac{\sum_{j=1}^{n} (Y_j - \hat{\beta}_0 - \hat{\underset{\sim}{\beta}}' \underset{\sim}{z}_j)^2}{n-r-1}$

Predicting Several Variables

$$\begin{bmatrix} \underset{m\times 1}{Y} \\ --- \\ \underset{r\times 1}{Z} \end{bmatrix} \sim N_{m+r}(\underset{\sim}{\mu}, \Sigma) \qquad \mu = \begin{bmatrix} \underset{\sim}{\mu}_Y \\ -- \\ \underset{\sim}{\mu}_Z \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{bmatrix}$$

$$E\{Y | z_1, \ldots, z_r\} = \mu_Y + \Sigma_{YZ} \Sigma_{ZZ}^{-1} (z - \underset{\sim}{\mu}_Z)$$

$$\Sigma_{YY\cdot Z} = E\{(\underset{\sim}{Y} - \underset{\sim}{\mu}_Y - \Sigma_{YZ}\Sigma_{ZZ}^{-1}(\underset{\sim}{Z} - \underset{\sim}{\mu}_Z))[\underset{\sim}{Y} - \underset{\sim}{\mu}_Y - \Sigma_{YZ}\Sigma_{ZZ}^{-1}(\underset{\sim}{Z} - \mu_Z)]'\}$$

$$= \Sigma_{YY} - \Sigma_{YZ} \Sigma_{ZZ}^{-1} \Sigma_{ZY}$$

Result 7.14 $\underset{\sim}{Y}, \underset{\sim}{Z} \sim N_{m+r}(\underset{\sim}{\mu}, \Sigma)$

The Regression of Y on Z is:

$$\beta_0 + \beta Z = \underset{\sim}{\mu}_Y - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\mu_Z + \Sigma_{YZ}\Sigma_{ZZ}^{-1} \underset{\sim}{z} = \underset{\sim}{\mu}_Y + \Sigma_{YZ}\Sigma_{ZZ}^{-1}(z - \underset{\sim}{\mu}_Z)$$

$$E\{(Y - \beta_0 - \beta Z)(Y - \beta_0 - \beta Z)'\} = \Sigma_{YY\cdot Z} = \Sigma_{YY} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Sigma_{ZY}$$

Based on ~~a~~ random Sample of Size n,

ML estimator: $\hat{\beta}_0 + \hat{\beta} z = \overline{Y} + S_{YZ} S_{ZZ}^{-1} (z - \overline{Z})$

$$\hat{\Sigma}_{YY\cdot Z} = \frac{n-1}{n}\left(S_{YY} - S_{YZ} S_{ZZ}^{-1} S_{ZY}\right)$$

Unbiased estimator: $\dfrac{1}{n-r-1} \sum_{j=1}^{n} (Y_j - \hat{\beta}_0 - \hat{\beta} z_j)(Y_j - \hat{\beta}_0 - \hat{\beta} z_j)'$

Comments

Joint normal Assumption $\Rightarrow$

$$\left.\begin{aligned} \hat{Y}_1 &= \hat{\beta}_{01} + \hat{\beta}_{11} z_1 + \cdots + \hat{\beta}_{r1} z_r \\ \hat{Y}_2 &= \hat{\beta}_{02} + \hat{\beta}_{12} z_1 + \cdots + \hat{\beta}_{r2} z_r \\ &\vdots \\ \hat{Y}_m &= \hat{\beta}_{0m} + \hat{\beta}_{1m} z_1 + \cdots + \hat{\beta}_{rm} z_r \end{aligned}\right\}$$

Same z's for each Y_i

$\hat{\beta}_{ik}$ are estimates of (i,k)th entry of $\beta = \Sigma_{YZ} \Sigma_{ZZ}^{-1}$

Partial Correlation Coefficient

$$\left.\begin{aligned} &Y_1 - \mu_{Y_1} - \Sigma_{Y_1 Z} \Sigma_{ZZ}^{-1}(Z - \mu_Z) \\ &Y_2 - \mu_{Y_2} - \Sigma_{Y_2 Z} \Sigma_{ZZ}^{-1}(Z - \mu_Z) \end{aligned}\right\}$$

errors from BLUPs for Y_1, Y_2

Correlation determined from error covariance matrix

$$\Sigma_{YY \cdot Z} = \Sigma_{YY} - \Sigma_{YZ} \Sigma_{ZZ}^{-1} \Sigma_{ZY}$$

$\Rightarrow$ Association between Y_1, Y_2 controlling for $z_1, \ldots, z_r$

$$\rho_{Y_1 Y_2 \cdot Z} = \frac{\sigma_{Y_1 Y_2 \cdot Z}}{\sqrt{\sigma_{Y_1 Y_1 \cdot Z}}\sqrt{\sigma_{Y_2 Y_2 \cdot Z}}}$$

elements of $\Sigma_{YY \cdot Z}$

$$r_{Y_1 Y_2 \cdot Z} = \frac{s_{Y_1 Y_2 \cdot Z}}{\sqrt{s_{Y_1 Y_1 \cdot Z}}\sqrt{s_{Y_2 Y_2 \cdot Z}}}$$

elements of $S_{YY \cdot Z}$