In retrospect

Stochastic vector $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_n)$

Vector of expectations $E\underset{\sim}{Y} = (EY_1, EY_2, \ldots, EY_n) = \underset{\sim}{\mu}$

Covariance matrix, variance-covariance matrix,
variance matrix, dispersion matrix

$$\text{Var}\,\underset{\sim}{Y} = E\left[(\underset{\sim}{Y} - E\underset{\sim}{Y})(\underset{\sim}{Y} - E\underset{\sim}{Y})^T\right] = \Sigma$$

Linear transformation $A\underset{\sim}{Y} + \underset{\sim}{b}$

$$E[A\underset{\sim}{Y} + \underset{\sim}{b}] = A\,E\underset{\sim}{Y} + \underset{\sim}{b}$$

$$\text{Var}(A\underset{\sim}{Y} + \underset{\sim}{b}) = A\,\text{Var}\,\underset{\sim}{Y}\,A^T$$

spec. $E[\underset{\sim}{a}^T \underset{\sim}{Y} + b] = \underset{\sim}{a}^T E\underset{\sim}{Y} + b$

$$\text{Var}(\underset{\sim}{a}^T \underset{\sim}{Y} + b) = \underset{\sim}{a}^T\,\text{Var}\,\underset{\sim}{Y}\,\underset{\sim}{a}$$

Normal distribution assumption

$\underset{\sim}{Y} \sim N_n(\underset{\sim}{\mu}, \Sigma)$, $E\underset{\sim}{Y} = \underset{\sim}{\mu}$, $\text{Var}\,\underset{\sim}{Y} = \Sigma$

Joint density function for $Y_1, \ldots, Y_n$:

$$f(y_1, \ldots, y_n) = (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}(\underset{\sim}{y} - \underset{\sim}{\mu})^T \Sigma^{-1} (\underset{\sim}{y} - \underset{\sim}{\mu})\right)$$

$A\underset{\sim}{Y} + \underset{\sim}{b} \sim N_k(A\underset{\sim}{\mu} + \underset{\sim}{b}, A\Sigma A^T)$, when rank $A = k$

spec. $\underset{\sim}{a}^T \underset{\sim}{Y} + b \sim N(\underset{\sim}{a}^T \underset{\sim}{\mu} + b, \underset{\sim}{a}^T \Sigma \underset{\sim}{a})$

Linear normal model

$\underset{\sim}{Y} \sim N_n(\underset{\sim}{\mu}, \sigma^2 I_n)$

$\underset{\sim}{\mu} \in L \subseteq R^n$, $L = \{ X\underset{\sim}{\beta} \mid \underset{\sim}{\beta} \in R^k \}$

design matrix

short notation: $\underset{\sim}{\mu} = X\underset{\sim}{\beta}$

Projection on L

$$\hat{\mu} = P y = X(X^T X)^{-1} X^T y \quad , \quad X \ n \times k \quad , \quad \text{rank } X = k$$

normal equations : $X^T X \beta = X^T y$

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

$$\hat{\beta} \sim N_k(\beta, \sigma^2 (X^T X)^{-1})$$

residuals : $r = y - \hat{y} = y - \hat{\mu} = y - Py = (I_n - P) y$

residual sum of squares : $\|r\|^2 = \|y - \hat{y}\|^2 = y^T (I_n - P) y$

$$Y^T (I_n - P) Y \sim \sigma^2 \chi^2 (n-k)$$

$$s^2 = \frac{1}{n-k} \|r\|^2 \quad , \quad E S^2 = \sigma^2$$

$\hat{\beta}$ og S^2 independent

Maximum likelihood estimation

$\hat{\mu} = P y$ (same as least squares' estimate)

$$\hat{\sigma}^2 = \frac{1}{n} \|r\|^2 = \frac{n-k}{n} s^2$$

$$L(\hat{\mu}, \hat{\sigma}^2) = (2\pi e \hat{\sigma}^2)^{-\frac{n}{2}}$$

Confidence interval for $a^T \beta$

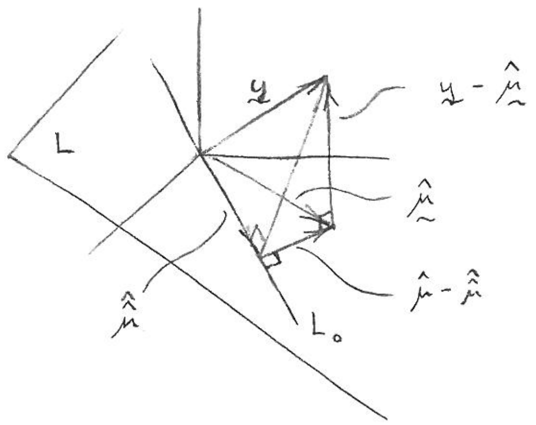
$$a^T \beta = a^T \hat{\beta} \pm t_{1-\frac{\alpha}{2}}(n-k) \, s \sqrt{a^T (X^T X)^{-1} a}$$

Prediction interval for y_{n+1}, where $E Y_{n+1} = a^T \beta$

$$y_{n+1} = a^T \hat{\beta} \pm t_{1-\frac{\alpha}{2}}(n-k) \, s \sqrt{1 + a^T (X^T X)^{-1} a}$$

Test $H_0 : \mu \in L_0$, $\dim L_0 = k_0 < k$

$H_1 : \mu \in L \setminus L_0$, $\dim L = \text{rank } X = k$



Likelihood ratio test statistic

$$Q(\underset{\sim}{y}) = \frac{L(\hat{\hat{\underset{\sim}{\mu}}}, \hat{\hat{\sigma}}^2)}{L(\hat{\underset{\sim}{\mu}}, \hat{\sigma}^2)} ,$$

small values critical for the hypothesis

Equivalent to Q is

$$F(\underset{\sim}{y}) = \frac{\frac{1}{k-k_0} \| \hat{\underset{\sim}{\mu}} - \hat{\hat{\underset{\sim}{\mu}}} \|^2}{\frac{1}{n-k} \| \underset{\sim}{y} - \hat{\underset{\sim}{\mu}} \|^2} ,$$

large values critical for the hypothesis

distribution: $F(Y) \sim F(k-k_0, n-k)$

Checking the model

The residuals:

$$\underset{\sim}{r} = \underset{\sim}{y} - \hat{\underset{\sim}{y}} = (I_n - P)\underset{\sim}{y}$$

Distribution:

$$\underset{\sim}{R} \sim N_n(\underset{\sim}{0}, \sigma^2 (I_n - P)) \quad \text{singular}$$

Standardized residuals:

$$S_i = \frac{R_i}{\sqrt{1 - p_{ii}}} , \qquad S_i \sim N(0, \sigma^2) , \qquad i = 1, \dots, n$$

The S_i's are often assumed to be approximately independent

Notation

$\underset{\sim}{a}$ column vector

A matrix

A^- generalised invers , $AA^-A = A$

$|A| = \det A$

$\det A \neq 0$ ~ A non singular (invertible)

$\mathcal{R}(A)$ column space corresponding to A

$\mathcal{N}(A)$ zero space (kernel) corresponding to A

x stochastic variable

$\underset{\sim}{x}$ stochastic vector

X stochastic matrix , $X = [\underset{\sim}{x}_1 \ldots \underset{\sim}{x}_n]^T$

$\mathcal{E}$ ~ E expectation operator

$\mathcal{D}$ ~ Var variance operator

$\mathcal{C}$ ~ Cov covariance operator

$\underset{\sim}{\mu}, \underset{\sim}{\theta}$ vector of expectations

M matrix of expectations, $M = [\underset{\sim}{\mu}_1 \ldots \underset{\sim}{\mu}_n]^T$

Σ variance matrix , $\Sigma > 0$

Expectation , variance, covariance

$$EX = \{E[x_{ij}]\} = [\underset{\sim}{\mu}_1 \ldots \underset{\sim}{\mu}_n]^T = M$$

$$E[AXB + C] = A\,EX\,B + C$$

$$\begin{aligned} \mathrm{Cov}(\underset{\sim}{x}, \underset{\sim}{y}) &= \{\mathrm{Cov}(x_i, y_j)\} \\ &= \{E[(x_i - Ex_i)(y_j - Ey_j)]\} \\ &= E[(\underset{\sim}{x} - E\underset{\sim}{x})(\underset{\sim}{y} - E\underset{\sim}{y})^T] \\ &= E[\underset{\sim}{x}\underset{\sim}{y}^T] - E\underset{\sim}{x}\,E\underset{\sim}{y}^T \end{aligned}$$
cf. ex. 1.1 b

$\underset{\sim}{x}, \underset{\sim}{y}$ indep. $\Rightarrow$ $\mathrm{Cov}(\underset{\sim}{x}, \underset{\sim}{y}) = E\underset{\sim}{x}\,E\underset{\sim}{y}^T - E\underset{\sim}{x}\,E\underset{\sim}{y}^T = 0$

$$\operatorname{Cov}(A\underset{\sim}{x}, B\underset{\sim}{y}) = A \operatorname{Cov}(\underset{\sim}{x}, \underset{\sim}{y}) B^T \quad \text{cf. ex. 1.1 a}$$

$$\operatorname{Var}\underset{\sim}{x} = \operatorname{Cov}(\underset{\sim}{x}, \underset{\sim}{x}) = E[(\underset{\sim}{x} - E\underset{\sim}{x})(\underset{\sim}{x} - E\underset{\sim}{x})^T]$$

$$= E[\underset{\sim}{x}\underset{\sim}{x}^T] - E\underset{\sim}{x}\, E\underset{\sim}{x}^T = \begin{bmatrix} \operatorname{Var} x_1 & \operatorname{Cov}(x_1, x_2) & \dots & \operatorname{Cov}(x_1, x_n) \\ \vdots & \vdots & & \vdots \\ \operatorname{Cov}(x_n, x_1) & \operatorname{Cov}(x_n, x_2) & \dots & \operatorname{Var} x_n \end{bmatrix}$$

$$\operatorname{Var}(A\underset{\sim}{x}) = \operatorname{Cov}(A\underset{\sim}{x}, A\underset{\sim}{x}) = A \operatorname{Cov}(\underset{\sim}{x}, \underset{\sim}{x}) A^T = A \operatorname{Var}\underset{\sim}{x}\, A^T$$

$$\operatorname{Cov}\Big(\sum_i a_i \underset{\sim}{x}_i, \sum_j b_j \underset{\sim}{y}_j\Big) = \Big\{\operatorname{Cov}\Big(\sum_i a_i x_{ir}, \sum_j b_j y_{js}\Big)\Big\} = \Big\{\sum_i \sum_j a_i b_j \operatorname{Cov}(x_{ir}, y_{js})\Big\}$$

$$= \sum_i \sum_j a_i b_j \{\operatorname{Cov}(x_{ir}, y_{js})\} = \sum_i \sum_j a_i b_j \operatorname{Cov}(\underset{\sim}{x}_i, \underset{\sim}{y}_j)$$

Quadratic form

$$\sum_i \sum_j a_{ij} x_i x_j = \sum_i x_i \sum_j a_{ij} x_j = \sum_i x_i (A\underset{\sim}{x})_i = \underset{\sim}{x}^T A \underset{\sim}{x},$$

$$a_{ij} := \tfrac{1}{2}(a_{ij} + a_{ji}) \Rightarrow A \text{ symmetric}$$

Generalized quadratic form

$$\sum_i \sum_j a_{ij} \underset{\sim}{x}_i \underset{\sim}{x}_j^T = X^T A X, \quad A \text{ symmetric}$$

cf. extra exercise

lemma $\underset{\sim}{x}_1, \dots, \underset{\sim}{x}_n$ indep., $E\underset{\sim}{x}_i = \underset{\sim}{\mu}_i$, $\operatorname{Var}\underset{\sim}{x}_i = \Sigma_i$

$$E[X^T A X] = \sum_i a_{ii} \Sigma_i + M^T A M, \quad M = EX$$

proof: $E[X^T A X] = E\Big[\sum_i \sum_j a_{ij} \underset{\sim}{x}_i \underset{\sim}{x}_j^T\Big] = \sum_i \sum_j a_{ij} E[\underset{\sim}{x}_i \underset{\sim}{x}_j^T]$

$$= \sum_i \sum_j a_{ij} E[(\underset{\sim}{x}_i - \underset{\sim}{\mu}_i)(\underset{\sim}{x}_j - \underset{\sim}{\mu}_j)^T + \underset{\sim}{x}_i \underset{\sim}{\mu}_j^T + \underset{\sim}{\mu}_i \underset{\sim}{x}_j^T - \underset{\sim}{\mu}_i \underset{\sim}{\mu}_j^T]$$

$$= \sum_i \sum_j a_{ij} \big(E[(\underset{\sim}{x}_i - \underset{\sim}{\mu}_i)(\underset{\sim}{x}_j - \underset{\sim}{\mu}_j)^T] + \underset{\sim}{\mu}_i \underset{\sim}{\mu}_j^T + \underset{\sim}{\mu}_i \underset{\sim}{\mu}_j^T - \underset{\sim}{\mu}_i \underset{\sim}{\mu}_j^T\big)$$

$$= \sum_i a_{ii} \operatorname{Var}\underset{\sim}{x}_i + 0 + \sum_i \sum_j a_{ij} \underset{\sim}{\mu}_i \underset{\sim}{\mu}_j^T$$

$$= \sum_i a_{ii} \Sigma_i + M^T A M$$

corollary $\Sigma_1 = \dots = \Sigma_n \Rightarrow E[X^T A X] = (\operatorname{tr} A)\Sigma + M^T A M$

Kronecker product (direct product, tensor product)

A $m \times m$ B $n \times n$

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mm}B \end{bmatrix}$$

$$\sim \begin{bmatrix} Ab_{11} & Ab_{12} & \cdots & Ab_{1n} \\ Ab_{21} & Ab_{22} & & Ab_{2n} \\ \vdots & \vdots & & \vdots \\ Ab_{n1} & Ab_{n2} & \cdots & Ab_{nn} \end{bmatrix}$$ (alternative arrangement of the elements)

ex. $\underset{\sim}{y} = [\underset{\sim}{x}_1^T \ \underset{\sim}{x}_2^T \ \ldots \ \underset{\sim}{x}_n^T]^T$ $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_n$ indep.

$\operatorname{Var} \underset{\sim}{x}_i = \Sigma$

$$\operatorname{Var} \underset{\sim}{y} = \begin{bmatrix} \Sigma & 0 & \cdots & 0 \\ 0 & \Sigma & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \Sigma \end{bmatrix} = I_n \otimes \Sigma$$

Sample $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_n$ $E\underset{\sim}{x}_i = \underset{\sim}{\mu}$ $\operatorname{Var} \underset{\sim}{x}_i = \Sigma$

$\underset{\sim}{x}_i$'s indep.

$$\bar{\underset{\sim}{x}} = \frac{1}{n} \sum_i \underset{\sim}{x}_i$$

$$S = \frac{1}{n-1} \sum_i (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})^T = \frac{1}{n-1} Q$$

$$E\bar{\underset{\sim}{x}} = \frac{1}{n} \sum_i E\underset{\sim}{x}_i = \frac{1}{n} n \underset{\sim}{\mu} = \underset{\sim}{\mu}$$

$$\operatorname{Var} \bar{\underset{\sim}{x}} = \frac{1}{n^2} \sum_i \operatorname{Var} \underset{\sim}{x}_i = \frac{1}{n^2} n \Sigma = \frac{1}{n} \Sigma$$

cf. ex. 1.3 a

Let $\underset{\sim}{y_i} = \underset{\sim}{x_i} - \underset{\sim}{\mu}$ (hence $\bar{\underset{\sim}{y}} = \bar{\underset{\sim}{x}} - \underset{\sim}{\mu}$)

$$Q = \sum_i (\underset{\sim}{x_i} - \bar{\underset{\sim}{x}})(\underset{\sim}{x_i} - \bar{\underset{\sim}{x}})^T = \sum_i (\underset{\sim}{y_i} - \bar{\underset{\sim}{y}})(\underset{\sim}{y_i} - \bar{\underset{\sim}{y}})^T$$

$$= \sum_i \underset{\sim}{y_i}\underset{\sim}{y_i}^T - n\bar{\underset{\sim}{y}}\bar{\underset{\sim}{y}}^T \quad \text{cf. ex. 1.2}$$

$$EQ = \sum_i E[\underset{\sim}{y_i}\underset{\sim}{y_i}^T] - nE[\bar{\underset{\sim}{y}}\bar{\underset{\sim}{y}}^T]$$

$$= n \operatorname{Var} \underset{\sim}{y_i} - n \operatorname{Var} \bar{\underset{\sim}{y}} \quad , \quad E\underset{\sim}{y_i} = \underset{\sim}{0} \quad E\bar{\underset{\sim}{y}} = \underset{\sim}{0}$$

$$= n\Sigma - n\tfrac{1}{n}\Sigma$$

$$= (n-1)\Sigma$$

$$ES = \frac{1}{n-1} EQ = \frac{1}{n-1}(n-1)\Sigma = \Sigma$$

Alternative calculations:

Let $\tilde{X}^T = [\underset{\sim}{x_1} - \bar{\underset{\sim}{x}} \quad \underset{\sim}{x_2} - \bar{\underset{\sim}{x}} \quad \dots \quad \underset{\sim}{x_n} - \bar{\underset{\sim}{x}}]$

$$= X^T - \bar{\underset{\sim}{x}}[1 \ 1 \ \dots \ 1]$$

$$= X^T - \bar{\underset{\sim}{x}}\underset{\sim}{1}^T = X^T - \underbrace{\tfrac{1}{n} X^T \underset{\sim}{1_n}}_{\bar{\underset{\sim}{x}}} \underset{\sim}{1_n}^T$$

$$= X^T (I_n - \tfrac{1}{n}\underset{\sim}{1_n}\underset{\sim}{1_n}^T)$$

$$= X^T (I_n - P)$$

Note that $P^T = \frac{1}{n}(\underset{\sim}{1_n}^T)^T \underset{\sim}{1_n}^T = \frac{1}{n}\underset{\sim}{1_n}\underset{\sim}{1_n}^T = P$

$$P^2 = \frac{1}{n^2}\underset{\sim}{1_n}\underset{\sim}{1_n}^T\underset{\sim}{1_n}\underset{\sim}{1_n}^T = \frac{n}{n^2}\underset{\sim}{1_n}\underset{\sim}{1_n}^T = P$$

rank $P = 1$

$$Q = \tilde{X}^T\tilde{X} = X^T(I_n - P)(I_n - P)X = X^T(I_n - P)X$$

$$EQ = (\operatorname{tr}(I_n - P))\Sigma + \underset{\sim}{\mu}\underset{\sim}{1_n}^T \underbrace{(I_n - P)\underset{\sim}{1_n}}_{\underset{\sim}{1_n} - \frac{1}{n}\underset{\sim}{1_n}\underset{\sim}{1_n}^T\underset{\sim}{1_n} = \underset{\sim}{1_n} - \frac{n}{n}\underset{\sim}{1_n}} \underset{\sim}{\mu}^T \quad \text{cf. corollary to lemma 1.1}$$

$$= (n-1)\Sigma + 0$$

Empirical correlation $r_{jk} = \dfrac{s_{jk}}{\sqrt{s_{jj}}\sqrt{s_{kk}}}$, $j,k = 1,\dots,d$ *

Let $D_s = \lceil s_{11} \; s_{22} \; \cdots \; s_{dd} \rfloor$ (diagonal matrix)

Empirical correlation matrix $R = D_s^{-\frac{1}{2}} S D_s^{-\frac{1}{2}}$

* $s_{jk} = \frac{1}{n-1} \sum_i (x_{ij} - x_{\cdot j})(x_{ik} - x_{\cdot k})$

$s_{jj} = \frac{1}{n-1} \sum_i (x_{ij} - x_{\cdot j})^2$

$s_{kk} = \frac{1}{n-1} \sum_i (x_{ik} - x_{\cdot k})^2$