Note 1:

$$p(y) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\|y - X\beta\|^2\right\}$$

$$\log p(y) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\|y - X\beta\|^2.$$

Note 2:

$$\frac{\partial \log p(y)}{\partial \beta} = -\frac{1}{2\sigma^2}\left(-2X^T y + 2X^T X\beta\right) = 0$$

$$X'X\beta = X^T y \longrightarrow \text{normal equation.}$$

Note 3:

Assignment Project Exam Help

$$H = X(X^T X)^{-1} X^T$$

https://powcoder.com

$$H^2 = X(X^T X)^{-1} X^T X (X^T X)^{-1} X^T = X(X^T X)^{-1} X^T = H$$

$$H^T = X(X^T X)^{-1} X^T = H$$

Add WeChat powcoder

H orthogonal projector.

why:

$$\operatorname{cov}(\hat\beta, y - X\hat\beta) = \operatorname{cov}\left((X'X)^{-1}X'y, (I-H)y\right)$$

$$= (X'X)^{-1}X' \operatorname{cov}(y)(I-H)'$$

$$= \sigma^2 (X'X)^{-1}X'(I-H) = 0$$

$$y - X\hat\beta = y - X(X'X)^{-1}X'y = (I-H)y$$

$$(X'X)^{-1}X'H = (X'X)^{-1}X'X(X'X)^{-1}X' = (X'X)^{-1}X'$$

Note 4:

$$RSS_1 = (N - p_1 - 1)\hat{\sigma}_1^2 \sim \sigma^2 \chi^2_{N-p_1-1}$$

$$RSS_0 = (N - p_0 - 1)\hat{\sigma}_0^2 \sim \sigma^2 \chi^2_{N-p_0-1}$$

$$RSS_0 - RSS_1 \sim \sigma^2 \chi^2_{p_1-p_0}$$

$RSS_0 - RSS_1$ independent of RSS_1.

$$\frac{(RSS_0 - RSS_1)/(p_1-p_0)}{RSS_1/(N-p_1-1)} \sim F_{p_1-p_0,\, N-p_1-1}.$$

Note 4:

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

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

$$\frac{(X'X)^{\frac{1}{2}}}{\sigma}(\hat{\beta} - \beta) \sim N(0, I_{p+1})$$

$$(\hat{\beta}-\beta)^T \frac{X'X}{\sigma}(\hat{\beta}-\beta) \sim \chi^2_{p+1}$$

$$\frac{\hat{\beta}_j - \beta_j}{\sigma\sqrt{v_j}} \sim N(0,1)$$

$$\frac{N-p-1}{\sigma^2} \sim \chi^2_{N-p-1}$$

$$H_0: \beta_j = 0 \longrightarrow \frac{\hat{\beta}_j}{\sigma\sqrt{v_j}} \sim t_{N-p-1}$$

Note 5:

$$y_0 = f(x_0) + \xi_0$$

$$E[y_0 - \tilde{f}(x_0)] = E[(f(x_0) + \xi - \tilde{f}(x_0))^2]$$

$$= E[\xi^2] + E[\xi(f(x_0) - \tilde{f}(x_0))] + E[(f(x_0) - \tilde{f}(x_0))]$$

$$= \sigma^2 + E[(f(x_0) - \tilde{f}(x_0))^2] = \sigma^2 + MSE[\tilde{f}(x_0)]$$

$E[\xi(f(x_0) - \tilde{f}(x_0))] = 0$ by independence of the error.

Note 6:

$$p(y) = \frac{1}{\sqrt{|2\pi||\Sigma|}} \exp\left\{-\frac{1}{2}(y - X\beta)^T \Sigma^{-1} (y - X\beta)\right\}$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

→ taking the log and ignoring the terms independent of β.

Note 7:

$$Var(\hat{\gamma}_j) = \frac{\sigma^2}{1 - R_j^2}$$

where γ_j are estimated from the centered & scaled predictors

$$VIF(\hat{\beta}_j) = \frac{Var(\hat{\gamma}_j)}{\sigma^2}$$

$$Var\,\hat{\gamma}_j = \sigma^2 (X'X)^{-1}_{jj} = \sigma^2 \sum_{l=1}^{p} t_{jl}^2 d_l^{-1}$$

where $\sum_{l=1}^{p} t_{jl}^2 = 1$, $X'X = T\Lambda T'$, $T = [t_1, \dots, t_p]$.

$$VIF_j \leq \sum_l t_{jl}^2 d_{min}^{-1} \leq d_{min}^{-1} \leq d_{max} d_{min}^{-1} = K(X)^2$$

Note 8:

For a success there is an increase of B_1.

Note 9:

$$y_i = f(x_i, \beta) + \varepsilon_i \quad i = 1, \dots, N$$

$$L(\beta) = \sum_{i=1}^{n} (y_i - f(x_i, \beta))^2 =$$

Note 10:

H: orthogonal projector $HX = X$

$$E[e] = E[(I-H)y] = (I-H)E[y] = (I-H)(X\beta + \delta)$$

$$= (I-H)\delta$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

Note 11:

$$y_i = \alpha_0 + \beta_1(x_{i1} - \bar{x}_1) + \dots + \beta_p(x_{ip} - \bar{x}_p) + \varepsilon_i$$

$$\alpha_0 = \beta_0 + \beta_1 \bar{x}_1 + \dots + \beta_p \bar{x}_p \qquad \bar{x}_j = \sum_i x_{ij}/n$$

The model becomes: $Y = X_c \alpha + \varepsilon$, $X_c = (1_n, \tilde{X})$

$\tilde{X}$ has element $\tilde{x}_{ij} = x_{ij} - \bar{x}_j$ and $\tilde{X}' 1_n = 0$ (mean).

$$H = X_c (X_c' X_c)^{-1} X_c' = (1_n \tilde{X}) \begin{pmatrix} n & 0' \\ 0 & \tilde{X}'\tilde{X} \end{pmatrix}^{-1} (1_n, \tilde{X})'$$

$$= \frac{1}{n} 1_n 1_n' + \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'$$

$$h_i = \frac{1}{n} + \frac{1}{n-1}(x_i - \bar{x})' S^{-1} (x_i - \bar{x})$$

h_i increase with the distance of x_i from $\bar{x}$

Note 11:

Using the multivariate gaussian density → take the log → ignore the terms independent of β.

Note 12:

This leads to a computational complexity problem.

Note 14:

Similar to Note 5 where $p_1 = i+1$ & $p_0 = i$, $i = k+1$.

Note 15

$$F = \frac{RSS(\hat\beta) - RSS(\hat\beta_1)}{RSS(\hat\beta_1)/(N-k-2)} \quad i = k.$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

Note 16:

After standardizing

$$L(\beta) = (y - X\beta)^T(y - X\beta) + \lambda\beta^T\beta$$

$$\frac{\partial L(\beta)}{\partial \beta} = -X^T(y - X\beta) + \lambda\beta = -X^Ty + (X^TX + \lambda I)\beta$$

$$\beta = (X^TX + \lambda I)^{-1}X^Ty.$$

Note 17:

orthogonal predictors: $X^TX = I$, $\beta_{LS} = X^Ty$, $\gamma = \frac{1}{1+\lambda}$.

Note 19:

By decomposing the right term of the last equation we can relate the solution to PCA.