5.8) (i) $y = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

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

Prediction: x_0^T

$\hat{y}_0 = \underline{x_0^T \hat{\beta}_{(i)} > 0.5}$

(ii) $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} y, 1-y \end{bmatrix} = [y, y_1]$

$\hat{\beta}_{(ii)} = (X^T X)^{-1} X^T Y = \Big[\hat{\beta}_{(i)}, \underbrace{(X^T X)^{-1} X^T (\mathbb{1} - y)}_{\underbrace{(X^T X)^{-1} X^T \mathbb{1} - \hat{\beta}_{(i)}}_{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \hat{\beta}_{(i)}}}\Big]$

$0 < \hat{y}_1 - \hat{y}_0 = x_0^T \hat{\beta}_{(i)} - x_0^T \left(\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \hat{\beta}_{(i)} \right)$

$= 2 x_0^T \hat{\beta}_{(i)} - 1$

$\Rightarrow \underline{x_0^T \hat{\beta}_{(i)} > 0.5}$

mar. 18-08:05

4.2) a) This is just (4.11) and (4.10)

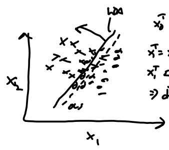
$\delta_2(x) > \delta_1(x)$, or you can get this from (4.9).

c) $\hat{\Sigma}_B \hat{\beta} = (\hat{\mu}_2 - \hat{\mu}_1) \underbrace{(\hat{\mu}_2 - \hat{\mu}_1)^T \hat{\beta}}_{\lambda} = \underline{\lambda (\hat{\mu}_2 - \hat{\mu}_1)}$

(4.56) $\Rightarrow (N-2) \hat{\Sigma} \hat{\beta} + \frac{N_1 N_2}{N} \lambda (\hat{\mu}_2 - \hat{\mu}_1) = N (\hat{\mu}_2 - \hat{\mu}_1)$

$\Rightarrow \hat{\Sigma} \hat{\beta} = (\hat{\mu}_2 - \hat{\mu}_1) \left(N - \frac{N_1 N_2}{N} \lambda\right) \frac{1}{N-2} = \gamma (\hat{\mu}_2 - \hat{\mu}_1)$

$\underline{\hat{\beta} = \gamma \hat{\Sigma}^{-1} (\hat{\mu}_2 - \hat{\mu}_1)}$



x_0 on line

$x_0^T \Sigma^{-1}(\mu_2 - \mu_1) = c$

$x_1^T = x_0^T + d^T$

$x_1^T \Sigma^{-1}(\mu_2 - \mu_1) = c$

$\Rightarrow d^T \Sigma^{-1}(\mu_2 - \mu_1) = 0$

mar. 18-08:34