Solution

Exerc. 4.2

(a) This is the result from (4.11), see also (4.9)

(b) Note first that $\sum_{i=1}^{N} y_i = 0$

(To see this, let the N_1 from class 1 be the first entries of $\underset{\sim}{y}$, and:

$$\sum_{i=1}^{N} y_i = \sum_{i=1}^{N_1} y_i + \sum_{i=N_1+1}^{N_1+N_2} y_i = N_1\left(-\frac{N}{N_1}\right) + N_2\left(\frac{N}{N_2}\right) = \underline{\underline{0}} \quad)$$

Denote $g(\beta_0, \beta) = \sum_i (y_i - \beta_0 - \beta^T x_i)^2$. Then

$$\frac{\partial}{\partial \beta_0} g(\beta_0, \beta) = -2 \sum_i (y_i - \beta_0 - \beta^T x_i)$$

and utilising that $\sum_i y_i = 0$, we get $\hat{\beta}_0 = -\beta^T \bar{x}$. Inserting this we get

$$g(\hat{\beta}_0, \beta) = \sum_i (y_i - \beta^T (x_i - \bar{x}))^2.$$

Now

$$\frac{\partial}{\partial \beta} g(\hat{\beta}_0, \beta) = -2 \sum_i (y_i - \beta^T (x_i - \bar{x}))(x_i - \bar{x})^T$$

which, when put to zero, gives

$$\sum y_i (x_i - \bar{x})^T = \hat{\beta}^T \sum (x_i - \bar{x})(x_i - \bar{x})^T$$

Here the left hand side equals (since $\sum y_i = 0$)

$$\sum y_i x_i =$$

$$= \sum_{i=1}^{N_1} \left(-\frac{N}{N_1}\right) x_i + \sum_{i=N_1+1}^{N_1+N_2} \frac{N}{N_2} x_i$$

$$= -N\hat{\mu}_1 + N\hat{\mu}_2 = N(\hat{\mu}_2 - \hat{\mu}_1)$$

Thus

$$\hat{\beta}^T T = N(\hat{\mu}_2 - \hat{\mu}_1)$$

where

$$T = \sum_i (x_i - \bar{x})(x_i - \bar{x})^T$$

Now

$$T = \sum_i (x_i - \bar{x})(x_i - \bar{x})^T$$

$N_1 \rightarrow N_1$

$$= \sum_{i=1}^{N_1} \left(x_i - \frac{N_1\hat{\mu}_1 + N_2\hat{\mu}_2}{N}\right)\left(x_i - \frac{N_1\hat{\mu}_1 + N_2\hat{\mu}_2}{N}\right)^T$$

$$= \sum_{i=1}^{N_1} + \sum_{i=N_1+1}^{N_1+N_2}$$

Consider the first sum, $\sum_{i=1}^{N_1}$:

$$= \sum_{i=1}^{N_1} \left(x_i - \hat{\mu}_1 + \cancel{N}\hat{\mu}_1 - \frac{N_1}{N}\hat{\mu}_1 - \frac{N_2\hat{\mu}_2}{N}\right)(\ \cdot\ \cdot\ )^T$$

$+\cancel{\sum}$

$$= \sum_{i=1}^{N_1} \left(x_i - \hat{\mu}_1 + \frac{N_2}{N}(\hat{\mu}_1 - \hat{\mu}_2)\right)\left(x_i - \hat{\mu}_1 + \frac{N_2}{N}(\hat{\mu}_1 - \hat{\mu}_2)\right)^T$$

$$= \sum_{i=1}^{N_1} (x_i - \hat{\mu}_1)(x_i - \hat{\mu}_1)^T + \frac{N_2^2}{N^2} N_1 (\hat{\mu}_1 - \hat{\mu}_2)(\hat{\mu}_1 - \hat{\mu}_2)^T$$

(where the rest of the terms are 0)

By symmetry, the second sum is

$$\sum_{i=N_1+1}^{N_1+N_2} \quad :$$

$$\sum_{i=N_1+1}^{N_1+N_2} (x_i - \hat{\mu}_2)(x_i - \hat{\mu}_2)^T + \frac{N_1^2}{N^2} N_2 (\hat{\mu}_1 - \hat{\mu}_2)(\hat{\mu}_1 - \hat{\mu}_2)^T$$

so summing the two gives:

$$(N-2)\hat{\Sigma} + \underbrace{\frac{N_1 N_2}{N} (\hat{\mu}_1 - \hat{\mu}_2)(\hat{\mu}_1 - \hat{\mu}_2)^T}_{= \hat{\Sigma}_B}$$

which gives (4.50)

(c) It follows from (4.56) that

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

$$= (\hat{\mu}_2 - \hat{\mu}_1)\underbrace{\left[-\frac{N_1 N_2}{N} (\hat{\mu}_2 - \hat{\mu}_1)^T \hat{\beta} + N \right]}_{\text{this is a number}}$$

Thus $\hat{\Sigma}\hat{\beta} \propto (\hat{\mu}_2 - \hat{\mu}_1)$

or $\hat{\beta} \propto \hat{\Sigma}^{-1} (\hat{\mu}_2 - \hat{\mu}_1)$

which is (4.57).

[We did not explicitly show that "$\hat{\Sigma}_B \hat{\beta}$ is in the direction of ..."

but this is implicit in the above calculations

(d) Drop it!

(e)

We get

$$\begin{aligned}\hat{f} =& \hat{\beta}_0 + \hat{\beta}^T x = -\beta^T \bar{x} + \hat{\beta}^T x = \hat{\beta}^T (x - \bar{x}) \\ =& (x - \bar{x})^T \hat{\beta} \propto (x - \bar{x})^T \hat{\Sigma}^{-1} (\hat{\mu}_2 - \hat{\mu}_1)\end{aligned}$$

For $N_1 = N_2, \bar{x} = 0.5(\hat{\mu}_1 + \hat{\mu}_2)$ and by inserting this linto the expression in (a), we obtain

$$(x - \bar{x})^T \hat{\Sigma}^{-1} (\hat{\mu}_2 - \hat{\mu}_1) > 0$$

which is equivalent to that $\hat{f} > 0$.