Exercice 4.2.a

Ex. 4.2 Suppose we have features $x \in \mathbb{R}^p$, a two-class response, with class sizes N_1, N_2, and the target coded as $-N/N_1, N/N_2$.

(a) Show that the LDA rule classifies to class 2 if

$$x^T\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\mu}_2-\hat{\mu}_1) > \frac{1}{2}\hat{\mu}_2^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\mu}_2 - \frac{1}{2}\hat{\mu}_1^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\mu}_1 + \log\left(\frac{N_1}{N}\right) - \log\left(\frac{N_2}{N}\right),$$

and class 1 otherwise.

1) Fonctions discriminantes de Fisher:

Pour une observation $x \in \mathbb{R}^p$, la fct discriminante pour la classe k est:

$$\delta_k(x) = x^T \hat{\Sigma}^{-1} \hat{\mu}_k - \frac{1}{2} \hat{\mu}_k + \log(\pi_k).$$

2) Comparaison des fonctions discriminantes:

La règle de LDA attribue x à la classe 2 si

$\delta_2(x) > \delta_1(x)$

$$\delta_2(x) - \delta_1(x) = \underbrace{x^T \hat{\Sigma}^{-1}(\hat{\mu}_2 - \hat{\mu}_1)}_{\text{Terme linéaire}}$$

$$\underbrace{-\frac{1}{2}\left(\hat{\mu}_2^T \hat{\Sigma}^{-1}\hat{\mu}_2 - \hat{\mu}_1^T \hat{\Sigma}^{-1}\hat{\mu}_1\right)}_{\text{Terme quadratique}} + \underbrace{\log\left(\frac{\pi_2}{\pi_1}\right)}_{\text{Terme des priors}}$$

$$\hat{\mu}_2^T \hat{\Sigma}^{-1} \hat{\mu}_2 - \hat{\mu}_1^T \hat{\Sigma}^{-1} \hat{\mu}_1 = (\hat{\mu}_2 + \hat{\mu}_1)^T \hat{\Sigma}^{-1} (\hat{\mu}_2 - \hat{\mu}_1)$$

Utilisation de l'identité $a^Ta - b^Tb = (a+b)^T(a-b)$

La différence devient :

$$\delta_2(x) - \delta_1(x) = x^T \hat{\Sigma}^{-1}(\hat{\mu}_2 - \hat{\mu}_1) - \frac{1}{2}(\hat{\mu}_2 - \hat{\mu}_1)^T \hat{\Sigma}^{-1}(\hat{\mu}_2 - \hat{\mu}_1)$$

$$+ \log\left(\frac{N_2}{N_1}\right).$$

3) Condition de classification :

On classe x dans la classe 2 si $\delta_2(x) - \delta_1(x) > 0$:

$$\boxed{x^T \hat{\Sigma}^{-1}(\hat{\mu}_2 - \hat{\mu}_1) > \frac{1}{2}(\hat{\mu}_2 + \hat{\mu}_1)^T \hat{\Sigma}^{-1}(\hat{\mu}_2 - \hat{\mu}_1) - \log\left(\frac{N_2}{N_1}\right)}$$