Ejercicios Análisis Discriminante

Daniel Czarnievicz

May 16, 2018

Ejercicio 1

$$\vec{X}_1 \sim N_1 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \; \begin{pmatrix} 2 & 1.6 \\ 1.6 & 2 \end{pmatrix} \right]$$

$$\vec{X}_2 \sim N_2 \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} ; \; \begin{pmatrix} 2 & 1.6 \\ 1.6 & 2 \end{pmatrix} \right]$$

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \, \sigma_2} \Rightarrow \sigma_{12} = 1.6$$

Dado que ambas poblaciones son normaes y con igual matriz de varianzas y covarianzas, se construye un discriminante lineal.

$$f_{\vec{X}_i}(\vec{x}_i) = \frac{1}{2\pi \, |\mathbf{\Sigma}|^{1/2}} \;\; \exp\Big\{ -\frac{1}{2}(\vec{x} - \vec{\mu})' \mathbf{\Sigma}^{-1} (\vec{x} - \vec{\mu}) \Big\}$$

La partición óptima es calsificar en la población $P_2 \Leftrightarrow \frac{f_{\vec{X}_2}(\vec{x}_2) \;\; \pi_2}{c(2|1)} > \frac{f_{\vec{X}_1}(\vec{x}_2) \;\; \pi_1}{c(1|2)}$, donde π_i es la probabilidad a priori de pertenecer a la población i, y $c(i|j)$ es el costo de clasificar en el grupo i a una observación proveniende de la población j. Dado que no se informan probabilidades a priori distintas para cada caso, ni costos distintos para cada caso, se asumen iguales. De esta forma, la partición óptima será clasificar en $P_2 \Leftrightarrow f_{\vec{X}_2}(\vec{x}_2) \;\; \pi_2 > f_{\vec{X}_1}(\vec{x}_1) \;\; \pi_1$. Dado que ambos términos son estrictamente positivos y ambas poblacions tienen igual matriz de varianzas, podemos concentrarnos en el kernel de las distribuciones. La regla será entonces: clasificar en $P_2 \Leftrightarrow \pi_1 \;\; D_1^2(\vec{x}) > \pi_2 \;\; D_2^1(\vec{x})$.

Las funciones discriminantes serán:

$$P(i \in g | \vec{X} = \vec{x}) = \max \Big\{ \pi_g \;\; \exp\big\{ -\frac{1}{2} \, D_{ig}^2 \big\} \Big\} \Rightarrow$$

$$\Rightarrow \log(\pi_g) - \frac{1}{2} \, D_{ig}^2 = \log(\pi_g) - \frac{1}{2} \Big[(\vec{x} - \vec{\mu})' \mathbf{\Sigma}^{-1} (\vec{x} - \vec{\mu}) \Big]$$

Ejercicio 2

Ejercicio 3

Ejericio 4