Ejercicio 1. Suponga que $Y \in \{0, \ldots, K-1\}$ con $K \geq 2$. Si $f(x)$ es gaussiano: $X \mid Y = k \sim N(\mu_k, \Sigma_k)$, la regla de Bayes se escribe como

$$h^*(x) = \arg\max_k \delta_k(x),$$

donde

$$\delta_k = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k) + \log(\pi_k).$$

estimador puntual

Sea $n_k = \sum_i \mathbf{1}_{\{y_i=k\}}$ para $k = 0, \ldots, K-1$. Demuestra que los estimadores puntuales de π_k, μ_k, Σ_k, y Σ son:

$$\hat{\pi}_k = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{y_i=k\}}, \qquad \hat{\mu}_k = \frac{1}{n_k}\sum_{\{i:y_i=k\}} X_i, \qquad \hat{\Sigma}_k = \frac{1}{n_k - 1}\sum_{\{i:y_i=k\}}(X_i - \hat{\mu}_k)(X_i - \hat{\mu}_k)^T$$

y

$$\hat{\Sigma} = \frac{\sum_{k=0}^{K-1}(n_k - 1)\hat{\Sigma}_k}{n-K}.$$

Toma un conjunto de puntos y calcula tales valores estimados.

$$h^*(x) = \arg\max_k \left[-\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k) + \log(\pi_k)\right]$$

En general, el estimador para la media es

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Y en este caso, para cada k

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} x_i$$

El estimador de la covarianza

$$\Sigma_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

entonces para cada k

$$\hat{\Sigma}_k = \frac{1}{n_k - 1}\sum (x_i - \mu_k)^2$$

$$\hat{\Sigma}_k = \frac{1}{n_k - 1}\sum [(x_i - \hat{\mu}_k)(x_i - \hat{\mu}_k)]^T$$

F. clasificadora

$\pi_k = P(x \in \pi_k)$ prob de que un punto esté en π_k

$$\mathbb{1}_A(x) = \begin{cases} 1, & x \in A \\ 0 & x \notin A \end{cases}$$

$$\Rightarrow \quad \hat{\pi}_k = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\pi_k}(x)$$

Tomamos puntos y calculamos

```python
#ejercicio 1
import numpy as np

n=1000
K=3
d=2 #dimensión

#generar datos
X=np.zeros((n, d))
y=np.zeros(n)

#medias y matrices de covarianza para cada clase
mus=np.random.rand(K, d)*10
sigmas=np.zeros((K,d,d))
for k in range(K):
    A=np.random.rand(d, d)
    sigmas[k]=np.dot(A, A.T) #garantizar que la matriz sea semidefinida positiva

for i in range(n):
    k=np.random.randint(0, K) #clase aleatoria
    X[i]=np.random.multivariate_normal(mus[k],sigmas[k])
    y[i]=k

n_k=np.array([np.sum(y == k) for k in range(K)])
pi_hat=n_k / n
mu_hat=np.array([np.sum(X[y == k], axis=0)/n_k[k] for k in range(K)])
sigma_hat=np.array([np.dot((X[y == k]-mu_hat[k]).T, X[y == k]-mu_hat[k])/(n_k[k] - 1) for k in range(K)])
#sigma_total = np.sum(n_k * sigma_hat, axis=0) / (n - K)
sigma_total=np.sum([n_k[k]*sigma_hat[k] for k in range(K)],axis=0)/(n - K)
```

```
Número de puntos por clase: [343 320 337]
pi_k: [0.343 0.32  0.337]
mu_k: [[1.88419301 7.26684056]
 [7.12873531 5.10725298]
 [1.73688305 0.7719954 ]]
Sigma_k: [[[0.57812033 0.61633774]
  [0.61633774 0.83995534]]

 [[0.03756978 0.23310139]
  [0.23310139 1.70494588]]

 [[0.94557216 0.75686213]
  [0.75686213 0.60597589]]]
Sigma: [[0.53056712 0.54268689]
 [0.54268689 1.04102431]]
```

Ejercicio 2. Bajo el supuesto de que las observaciones en la k-ésima clase se toma a partir de una distribución normal, el clasificador de Bayes asigna una observación a la clase para la cual se maximiza la función discriminante. ¿Cuál es la función discriminante en este caso?

$$f_k(x) = p(x|y=k)$$

Pero si los datos se distribuyen normal multivariante

$\rightarrow f$ de densidad de prob multivariante

$$p(\mathrm{x}|\mu,\Sigma) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\{ -\frac{1}{2}(\mathrm{x}-\mu)^T\Sigma^{-1}(\mathrm{x}-\mu) \}$$

$$f_k(x) = p(x|y=k) = \frac{1}{(2\pi)^{p/2}|\Sigma_k|^{1/2}} \exp\{-\tfrac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)\}$$

p - dimensión datos

F discriminante

$$\log\left(\frac{f_1(x)\pi_1}{f(x)\pi}\right) = \log(f_1(x)\pi_1) - \log(f_k(x)\pi_k)$$

$$= -\tfrac{1}{2}\log|\Sigma_1| - \tfrac{1}{2}(x-\mu_1)^t\Sigma_1^{-1}(x-\mu_1) + \log(\pi_1)$$
$$+ \tfrac{1}{2}\log|\Sigma_k| + \tfrac{1}{2}(x-\mu_k)^t\Sigma_2^{-1}(x-\mu_k) - \log(\pi_k)$$

$$= -\tfrac{1}{2}\log\left(\frac{|\Sigma_1|}{|\Sigma_k|}\right) - \tfrac{1}{2}\left((x-\mu_1)^t\Sigma_1^{-1}(x-\mu_1) - (x-\mu_k)^t\Sigma_k^{-1}(x-\mu_k)\right) + \log\left(\frac{\pi_1}{\pi_k}\right)$$

Ejercicio 3. Sean $\mu_1 = (1,0)$. Considere la siguiente matriz:

$$\Sigma = \begin{pmatrix} 4 & 8 \\ 8 & 4 \end{pmatrix}$$

¿Existe un vector gaussiano con media μ_1 y matriz de covarianza Σ?

Eigenvalores

$$\begin{vmatrix} 4-\lambda & 8 \\ 8 & 4-\lambda \end{vmatrix} = (4-\lambda)^2 - 64 = 0$$

$$16 - 8\lambda + \lambda^2 - 64 = 0$$

$$\lambda^2 - 8\lambda - 48 = 0$$

$$\lambda_1 = 12 \quad \lambda_2 = -4$$

Notemos que no todos los eigenvalores son positivos, por lo que la matriz no es positiva definida

$\Rightarrow$ no existe vector gaussiano

Ejercicio 4. Sea $u \in (1, \infty)$ y (X_1, X_2, X_3) un vector aleatorio con densidad:

$$f(x_1, x_2, x_3) = \frac{\sqrt{u^2-1}}{(2\pi)^{3/2}} \exp\left(-\frac{1}{2}(ux_1^2 + x_2^2 + ux_3^2 + 2x_1x_2)\right),$$

con $(x_1, x_2, x_3) \in \mathbb{R}^3$. Calcula la distribución de (X_1, X_2, X_3).

$$\frac{1}{\sqrt{(2\pi)^n |\det J|}} e^{-\frac{1}{2}(x-\mu)^t J (x-\mu)}$$

matriz cov

μ - grodiente de exp de f. exp. $\mu = (0,0,0)$

$$\Sigma = \begin{pmatrix} Var(x_1) & Cov(x_1,x_2) & \cdots \\ Cov(x_2,x_1) & Var(x_2) & \\ \vdots & & \ddots \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 0 \\ 0 & 0 & u^2/4 \end{pmatrix} \qquad \Sigma = 1/4 \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & u^2 \end{pmatrix}$$

$$f_{x_1,x_2,x_3} = \frac{1}{(2\pi)^{3/2}\sqrt{|\Sigma|}} \exp\left(-1/2 (x-\mu)^t \Sigma_1^{-1} (x-\mu)\right) \qquad |\Sigma| = 1/16 (1 \cdot 1 \cdot u^2) = u^2/16$$

$$= \frac{1}{(2\pi)^{3/2}\sqrt{\frac{u^2}{16}}} \exp\left(-1/2 (x-\mu)^t \Sigma_1^{-1} (x-\mu)\right)$$

$$= \frac{1}{\sqrt{2\pi}\, u} \exp\left(-u^2/32 \,(x_1^2 + 2x_1x_2 + x_2^2)\right)$$

Hacemos $U_1 = X_1 \quad U_2 = X_2 \quad U_3 = X_3$

$$\to f_{x_1 x_2 x_3} = \frac{1}{(2\pi)^{3/2} |\det I|} \exp\left(-1/2 (x-\mu)^t (x-\mu)\right) \qquad |\det I| = 1$$

$$= \frac{1}{(2\pi)^{3/2}} \exp\left(-1/2 (x_1^2 + x_2^2 + x_3^2)\right)$$

Ejercicio 5. $\mu_1 = (1,0)$ y $\mu_2 = (-2,2)$. Considere las siguientes matrices:

$$\Sigma_1 = \begin{pmatrix} 7 & 2 \\ 2 & 1 \end{pmatrix}, \qquad \Sigma_2 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

- ¿Existe un vector gaussiano bivariado con media μ_1 y matriz de covarianza Σ_1?
- Si $\pi_k = 0.5$ para $k = 1, 2$ y si $f_1(x_1, x_2) \sim N(\mu_1, \Sigma_1)$ y $f_2(x_1, x_2) \sim N(\mu_2, \Sigma_1)$; calcule

$$L(x_1, x_2) = \log\left(\frac{f_1(x_1,x_2)\pi_1}{f_2(x_1,x_2)\pi_2}\right).$$

Recuerda que asignamos (x_1, x_2) a Π_1 si: $L(x_1, x_2) > 0$.

- ¿A qué grupo asignas $(2, 1)$?

- Si $\pi_k = 0.5$ para $k = 1, 2$; calcule

$$Q(x_1, x_2) = \log\left(\frac{f_1(x_1,x_2)\pi_1}{f_2(x_1,x_2)\pi_2}\right).$$

Recuerda que asignamos (x_1, x_2) a Π_1 si: $Q(x_1, x_2) > 0$.

- ¿A qué grupo asignas $(2, 1)$?

Eigenvalores

$$\begin{vmatrix} 7-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = (7-\lambda)(1-\lambda) - 4 = 0$$

$$7 - 7\lambda - \lambda + \lambda^2 - 4 = 0$$

$$\lambda^2 - 8\lambda + 3 = 0$$

$$\lambda_1 = 7.6 \qquad \lambda_2 \approx 0.39$$

la matriz es definida positiva
existe un vector gaussiano bivariado
con μ_1 y Σ_1

$$\begin{vmatrix} 1-\lambda & 0 \\ 0 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) = 0$$

$$\lambda_1 = 1 \qquad \lambda_2 = 3$$

la matriz es definida positiva
existe un vector gaussiano bivariado
con μ_1 y Σ_1

$$L(x_1, x_2) = \log\left(\frac{f_1(x_1,x_2)\pi_1}{f_2(x_1,x_2)\pi_2}\right).$$

$\pi_k = 0.5 \quad k = 1, 2. \quad f_1(x_1, x_2) \sim N(\mu_1, \Sigma_1) \quad f_2(x_1, x_2) \sim N(\mu_1, \Sigma_1)$

$(x_1, x_2) : \Pi_1 \quad$ si $\quad L(x_1, x_2) > 0$

$\mu_1 = (1, 0) \qquad \mu_2 = (-2, 2)$

$$L(x) = \log\left(\frac{f_1(x_1,x_2)\pi_1}{f_2(x_1,x_2)\pi_2}\right) = \beta_0 + \beta^t X$$

$$\Sigma_1^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -2 \\ 2 & 7 \end{pmatrix} = \begin{pmatrix} 1/3 & -2/3 \\ 2/3 & 7/3 \end{pmatrix} \qquad \Sigma_2^{-1} = 1/3 \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix}$$

$$\beta = \Sigma_1^{-1}(\mu_1 - \mu_2) = \begin{pmatrix} 1/3 & -2/3 \\ 2/3 & 7/3 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 7/3 \\ -20/3 \end{pmatrix}$$

$$\beta_0 = -\tfrac{1}{2}\left(\mu_1^t \Sigma_1^{-1} \mu_1 - \mu_2^t \Sigma_1^{-1} \mu_2\right) + \log\left(\frac{\pi_2}{\pi_1}\right)$$

$$= -\tfrac{1}{2}\left((1,0)\begin{pmatrix} 1/3 & -2/3 \\ -2/3 & 7/3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} - (-2 \;\; 2)\begin{pmatrix} 1/3 & -2/3 \\ -2/3 & 7/3 \end{pmatrix}\begin{pmatrix} -2 \\ 2 \end{pmatrix}\right) + \log(1)$$

$$= -\tfrac{1}{2}\left(1/3 - 16\right) = -\tfrac{1}{2}\left(-47/3\right) = 47/6$$

$$L(x) = 47/6 + 7/3\,(2) - 20/3\,(1) = 35/6$$

Como $L(x) > 0$

$(2,1)$ es asignado a π_1

$$Q(x) = \beta_0 + \beta_1^t x + x^t \Omega x$$

$$\Omega = \tfrac{1}{2}\left(\Sigma_1^{-1} - \Sigma_2^{-1}\right) = -\tfrac{1}{2}\begin{pmatrix} -2/3 & -2/3 \\ -2/3 & 2 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 \\ 1/3 & -1 \end{pmatrix}$$

$$\beta = \Sigma_1^{-1}\mu_1 - \Sigma_1^{-1}\mu_2 = \begin{pmatrix} 1/3 & -2/3 \\ -2/3 & 7/3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix}\begin{pmatrix} -2 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1/3 \\ -2/3 \end{pmatrix} - \begin{pmatrix} -2 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 7/3 \\ -4/3 \end{pmatrix}$$

$$\beta_0 = -\tfrac{1}{2}\left[\log\frac{|\Sigma_1|}{|\Sigma_2|} + \mu_1^t \Sigma_1^{-1}\mu_1 - \mu_2^t \Sigma_2^{-1}\mu_2\right] - \log\left(\frac{\pi_2}{\pi_1}\right)$$

$$= -\tfrac{1}{2}\left[\log(1) + 10\begin{pmatrix} 1/3 & -2/3 \\ -2/3 & 7/3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} - (-2 \;\; 2)\begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix}\begin{pmatrix} -2 \\ 2 \end{pmatrix}\right] - \log(1)$$

$$= -\tfrac{1}{2}\left[\log(1) + \begin{pmatrix} 10/3 \\ -20/3 \end{pmatrix} - \frac{16}{3}\right] - \log(1)$$

$$= -\tfrac{1}{2}\left[1/3 - 16/3\right] = 5/2$$

$$Q(x) = 5/2 + (7/3 \;\; -4/3)\begin{pmatrix} 2 \\ 1 \end{pmatrix} + (2 \;\; 1)\begin{pmatrix} 1/3 & 1/3 \\ 1/3 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$= 5/2 + 2/3 + 5/3 = 29/6 \approx 4.83$$

Como $Q(x) > 0$

asignamos a $(2,1)$ a π_1