Capítulo 2

Componentes Principales

2.1. Distribución normal bivariante

1) Considere una distribución normal bivariante con $\mu_1 = 1$, $\mu_2 = 3$, $\sigma_{11} = 2$, $\sigma_{22} = 1$ y $\rho_{12} = -0.8$ (correlación).

- **a)** Escribir la densidad normal bivariante
- **b)** Contruir la ecuación de la elipse
- **c)** Graficar los contornos que contienen el 50 %, 75 % y 95 % de confianza.
- **d)** Escriba la densidad normal bivariante si se considera la matriz de correlaciones
- **e)** Contruir la ecuación de la elipse utilizando los datos del inciso d).
- **f)** Graficar la región al 95 % de confianza, tomando en cuenta los resultados en e).

2) Considere una distribución normal bivariante con $\mu_1 = 1$, $\mu_2 = 3$, $\sigma_{11} = 2$, $\sigma_{22} = 1$ y $\rho_{12} = 0.8$

- **a)** Escribir la densidad normal bivariante
- **b)** Contruir la ecuación de la elipse
- **c)** Graficar los contornos que contienen el 50 %, 75 % y 95 % de confianza.
- **d)** Escriba la densidad normal bivariante si se considera la matriz de correlaciones
- **e)** Contruir la ecuación de la elipse utilizando los datos del inciso d).
- **f)** Graficar la región al 95 % de confianza, tomando en cuenta los resultados en e).

2.2. Combinaciones Lineales

3) Mostrar que si $x_1, x_2, .., x_n$ es una muestra aleatoria proveniente de una población $N_p(\mu, \Sigma)$, luego:

$$\sum_{i=1}^{n} a_i\, x_i \sim N_p\left(\sum_{i=1}^{n} a_i\mu, \sum_{i=1}^{n} a_i^2\Sigma\right)$$

Cabe señalar que: $a_1, a_2, .., a_n$ es un conjunto de escalares.

4) Sea x que tiene distribución $N_3(\mu, \Sigma)$ con $\mu' = [2, -3, 1]$ y

$$\Sigma = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

a) Encontrar la distribución de $3x_1 - 2x_2 + x_3$

b) Encontrar el vector a tal que x_2 y

$$x_2 - a' \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

son independientes

2.3. Evaluación de normalidad multivariante

5) La Tabla 4.4 (página 190 - Johnson y Wichern) presenta 30 vectores de realizaciones medidas en cuatro variables x_1, x_2, x_3 y x_4. Evaluar la normalidad multivariante analítica y gráficamente (plot chi-cuadrado), para determinar la normalidad:

a) bivariante entre pares de variables (x_i, x_j), para $i, j = 1, 2, 3, 4$ con $i \neq j$. En este caso en particular construir además las elipses y graficar las observaciones en estas elipses.

b) trivariante para ternas (x_i, x_j, x_k), para $i, j, k = 1, 2, 3, 4$

c) tetravariante para (x_1, x_2, x_3, x_4)

2.4. Contexto Teórico

6. Sea la matriz de varianzas y covarianzas:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Además calcular:

a) Hallar los eigen valores y eigen vectores

b) Definir las componentes principales y_1 y y_2 .

7. Utilizando la matriz de covarianzas del ejercicio anterior, realizar los siguientes:

a) Convertir la matriz Σ en una matriz de correlaciones ρ

b) Hallar los eigen valores y eigen vectores de ρ.

c) Determinar las componentes principales de variables estandarizadas: y_1 y y_2.

8. Dada la matriz de covarianzas:

$$S = \begin{bmatrix} c^2 + d & c & c \\ c & 1 + d & 1 \\ c & 1 & 1 + d \end{bmatrix}$$

a) Encontrar las componentes principales.

b) Determinar las correlaciones entre las componentes y las variables.

c) Interpretar las componentes en función del tamaño de d.

2.5. Aplicación Práctica

9. Se dispone de 3 indicadores económicos x_1, x_2 y x_3, que se miden en cuatro paises, con los resultados siguientes:

x_1	x_2	x_3
2	3	-1
1	5	-2
2	2	1
2	3	1

a) Construir las componentes principales utilizando la matriz de covarianzas

b) Estimar las componentes principales utilizando la matriz de datos

c) Verificar si las varianzas de las componentes principales corresponden a los eigen valores de la matriz de covarianzas.

d) Verificar si la covarianza entre pares de componentes principales son nulas.

10. Utilizando el conjunto de datos anterior realizar lo siguiente:

a) Estandarizar los datos, es decir determinar los elementos de la matriz Z.

b) Construir las componentes principales utilizando la matriz de covarianzas (o correlaciones).

c) Estimar las componentes principales utilizando la matriz de datos estandarizada

d) Verificar si las varianzas de las componentes principales corresponden a los eigen valores de la matriz de covarianzas (o correlaciones).

e) Verificar si la covarianza entre pares de componentes principales son nulas.