Tema: Geometria de la muestra y muestreo aleatorio

1. Dada la matriz de datos

$$\mathbf{X} = \begin{bmatrix} 9 & 1 \\ 5 & 3 \\ 1 & 2 \end{bmatrix}$$

   (a) Grafique el diagrama de dispersión en $p = 2$ dimensiones. Localice la media de la muestra en su diagrama.

   (b) Dibuje la representación $n = 3$ -dimensional de los datos y trace los vectores de desviación $\mathbf{y}_1 - \bar{x}_1\mathbf{1}$ y $\mathbf{y}_2 - \bar{x}_2\mathbf{1}$

   (c) Dibuje los vectores de desviación en (b) que emanan del origen. Calcula las longitudes de estos vectores y el coseno del ángulo entre ellos. Relacione estas cantidades con $\mathbf{S}_n$ y $\mathbf{R}$

   (d) Calcular la varianza muestral generalizada $|\mathbf{S}|$

2. Dibuje los elipsoides sólidos $(\mathbf{x} - \overline{\mathbf{x}})'\mathrm{S}^{-1}(\mathbf{x} - \overline{\mathbf{x}}) \leq 1$ para las tres matrices

$$\mathbf{S} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

   (Tenga en cuenta que estas matrices tienen la misma varianza generalizada $|\mathbf{S}|$.)

3. Demuestre que $|\mathbf{S}| = (s_{11}s_{22}\cdots s_{pp})\,|\mathbf{R}|$

4. Considere la matriz de datos $\mathbf{X}$ del ejercicio 1. Tenemos $n = 3$ observaciones sobre $p = 2$ variables X_1 y X_2. Forman las combinaciones lineales

$$\mathbf{c}'\mathbf{X} = \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = -X_1 + 2X_2$$
$$\mathbf{b}'\mathbf{X} = \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 2X_1 + 3X_2$$

   (a) Evalúe las medias, varianzas y covarianzas muestrales de $\mathbf{b}'\mathbf{X}$ y $\mathbf{c}'\mathbf{X}$ a partir de los primeros principios. Es decir, calcule los valores observados de $\mathbf{b}'\mathbf{X}$ y $\mathbf{c}'\mathbf{X}$, y luego utilice las fórmulas de media, varianza y covarianza de la muestra.

   (b) Calcule las medias, las varianzas y la covarianza muestrales de $\mathbf{b}'\mathbf{X}$ y $\mathbf{c}'\mathbf{X}$. Compare los resultados en (a) y (b).

Tema: Densidad normal multivariante y sus propiedades

5. Sea $\mathbf{V}$ una variable aleatoria vectorial con un vector medio $E(\mathbf{V}) = \boldsymbol{\mu}_{\mathbf{v}}$ y una matriz de covarianza $E(\mathbf{V} - \boldsymbol{\mu}_{\mathbf{V}})(\mathbf{V} - \boldsymbol{\mu}_{\mathbf{V}})' = \boldsymbol{\Sigma}_{\mathbf{v}}$. Demuestre que $E(\mathbf{V}\mathbf{V}') = \boldsymbol{\Sigma}_{\mathbf{V}} + \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}'$

6. Considere una distribución normal bivariada con $\mu_1 = 1, \mu_2 = 3, \sigma_{11} = 2, \sigma_{22} = 1$ y $\rho_{12} = -.8$

   (a) Escriba la densidad normal bivariada.

(b) Escriba la expresión de distancia estadística al cuadrado $(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ como una función cuadrática de x_1 y x_2.

7. Sea $\mathbf{X}$ $N_3(\boldsymbol{\mu},\boldsymbol{\Sigma})$ con $\boldsymbol{\mu}'=[-3,1,4]$ y

$$\boldsymbol{\Sigma}=\begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

¿Cuáles de las siguientes variables aleatorias son independientes? Explique.

(a) X_1 y X_2
(b) X_2 y X_3
(c) (X_1,X_2) y X_3
(d) $\frac{X_1+X_2}{2}$ y X_3
(e) X_2 y $X_2-\frac{5}{2}X_1-X_3$
(f) Especifique la distribución condicional de X_2, dado que $X_1=x_1$ y $X_3=x_3$

8. Sea $\mathbf{X}$ $N_3(\boldsymbol{\mu},\boldsymbol{\Sigma})$ con $\boldsymbol{\mu}'=[2,-3,1]$ y

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

(a) Encuentre la distribución de $3X_1-2X_2+X_3$
(b) Vuelva a etiquetar las variables si es necesario y encuentre un vector 2×1 tal que X_2 y

$$X_2-\mathbf{a}'\begin{bmatrix} X_1 \\ X_3 \end{bmatrix} \text{ are independent.}$$

9. Sea $\mathbf{X}$ distribuido como $N_3(\boldsymbol{\mu},\boldsymbol{\Sigma})$, donde $\boldsymbol{\mu}'=[1,-1,2]$ y

$$\boldsymbol{\Sigma}=\begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

especifique cada uno de los siguientes:

(a) La distribución condicional de X_1, dado que $X_3=x_3$
(b) La distribución condicional de X_1, dado que $X_2=x_2$ y $X_3=x_3$

10. Sean $\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3$, y $\mathbf{X}_4$ vectores aleatorios independientes $N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$.

(a) Encuentre las distribuciones marginales para cada uno de los vectores aleatorios

$$\mathbf{V}_1=\frac{1}{4}\mathbf{X}_1-\frac{1}{4}\mathbf{X}_2+\frac{1}{4}\mathbf{X}_3-\frac{1}{4}\mathbf{X}_4$$

y

$$\mathbf{V}_2=\frac{1}{4}\mathbf{X}_1+\frac{1}{4}\mathbf{X}_2-\frac{1}{4}\mathbf{X}_3-\frac{1}{4}\mathbf{X}_4$$

(b) Encuentre la densidad conjunta de los vectores aleatorios $\mathbf{V}_1$ y $\mathbf{V}_2$ definidos en (a).