Godinez Bravo Diego

Estadística Multivariada

Tarea 1

Ejercicio 1. Demuestre que la matriz de centrado $P = I - \frac{1}{n}1'$ cumple con las siguientes propiedades:

a) Tiene rango $(n-1)$, es decir, tiene $n-1$ columnas o renglones linealmente independientes.

b) Sus valores propios son 1 o 0.

Solución

a) Sea la matriz de centrado P definida como

$$P = I - \frac{1}{n}11' = \begin{pmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{pmatrix}$$

$$PP = \begin{pmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{pmatrix} \begin{pmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{pmatrix}$$

Desarrollando el término x_{11}

$$x_{11} = (1-\frac{1}{n})(1-\frac{1}{n}) + (-\frac{1}{n})(-\frac{1}{n}) + (-\frac{1}{n})(-\frac{1}{n}) + ... + (-\frac{1}{n})(-\frac{1}{n})$$

$$= 1 - \frac{2}{n} + \frac{1}{n^2} + (-\frac{1}{n^2}) + ... + (-\frac{1}{n^2})$$

$$= 1 - \frac{2}{n} + \frac{1}{n^2} + (n-1)(-\frac{1}{n})^2$$

$$= 1 - \frac{2}{n} + \frac{1}{n^2} + (n-1)(\frac{1}{n^2})$$

$$= 1 - \frac{2}{n} + \frac{1}{n^2} + \frac{1}{n} - \frac{1}{n^2}$$

$$x_{11} = 1 - \frac{1}{n}$$

Por lo tanto se puede generalizar que para cada uno de los términos de la diagonal de la matriz PP $x_{ii} = 1 - \frac{1}{n}$.

Desarrollando el término x_{12}

$$x_{12} = (1-\frac{1}{n})(-\frac{1}{n}) + (-\frac{1}{n})(1-\frac{1}{n}) + (-\frac{1}{n})(-\frac{1}{n}) + ... + (-\frac{1}{n})(-\frac{1}{n})$$

$$= (1-\frac{1}{n})(-\frac{1}{n}) + (-\frac{1}{n})(1-\frac{1}{n}) + (n-2)(-\frac{1}{n})^2$$

$$= -\frac{1}{n} + \frac{1}{n^2} - \frac{1}{n} + \frac{1}{n^2} + \frac{1}{n} - \frac{2}{n^2}$$

$$= -\frac{1}{n} + \frac{1}{n^2} - \frac{1}{n} + \frac{1}{n^2} + \frac{1}{n} - \frac{2}{n^2}$$

$$x_{12} = -\frac{1}{n}$$

De manera similar, por medio de propiedades algebraicas podemos generalizar que los elementos fuera de la diagonal $x_{ij} = -\frac{1}{n}$.

Por lo tanto, observamos que los elementos de la diagonal estarían dados por $1-\frac{1}{n}$, mientras que los elementos fuera de la diagonal estarían definidos como $-\frac{1}{n}$. Demostrando que la matriz P es idempotente, esto es

$$PP = P$$

b) Sea P una matriz idempotente, entonces

$$Px = \lambda x$$

$$P^2x = P\lambda x$$

$$= \lambda Px$$

$$= \lambda(\lambda x)$$

$$= \lambda^2 x$$

$$Px = \lambda^2 x$$

$$\lambda x = \lambda^2 x$$

De manera que $\lambda = \lambda^2$, por lo que esto solo sería posible cuando λ toma los valores $\lambda = 0$ y $\lambda = 1$.

Ejercicio 2. Dado los siguientes datos:

Promotora	X_1 = Duración media (hipoteca años)	X_2 = Precio medio (millones euros)	X_3 = Superficie media (m^2 de cocina)
1	8.7	0.3	3.1
2	14.3	0.9	7.4
3	18.9	1.8	9.0
4	19.0	0.8	9.4
5	20.5	0.9	8.3
6	14.7	1.1	7.6
7	18.8	2.5	12.6
8	37.3	2.7	18.1
9	12.6	1.3	5.9
10	25.7	3.4	15.9

a) Dibújese el diagrama de dispersión múltiple y coméntese el aspecto del gráfico.

b) Para X_1 y X_2, calcúlense, respectivamente, las medias muestrales $\bar{x}_1$ y $\bar{x}_2$, las varianzas muestrales s_{11} y s_{22}, la covarianza entre X_1 y X_2, s_{12}, y la correlación entre ambas, r_{12}. Interprétese el valor obtenido de r_{12}.

c) Utilizando la matriz de datos X y la de centrado P, calcúlense el vector de medias muestrales $\hat{x}$ y la matriz de covarianzas muestrales S. A partir de ésta obténgase la matriz de correlaciones R.

Solución

a) Se generó el gráfico de dispersión múltiple de la siguiente manera, utilizando el lenguage de programación R.

```r
data <- matrix(c(
  1, 8.7, 0.3, 3.1,
  2, 14.3, 0.9, 7.4,
  3, 18.9, 1.8, 9.0,
  4, 19.0, 0.8, 9.4,
  5, 20.5, 0.9, 8.3,
  6, 14.7, 1.1, 7.6,
  7, 18.8, 2.5, 12.6,
  8, 37.3, 2.7, 18.1,
  9, 12.6, 1.3, 5.9,
  10, 25.7, 3.4, 15.9
), ncol = 4, byrow = TRUE)

colnames(data) <- c("id", "X1", "X2", "X3")
X <- data[, c("X1", "X2", "X3")]

ggpairs(X) # the pairwise scatter plot help us to visualize the distribution of single variables as well as relationships between two variable
```

El gráfico de pares muestra los histogramas de las variables en la diagonal principal y gráficos de dispersión en los cuadrantes restantes. Nos ayuda a visualizar las distribuciones univariadas y las relaciones bivariadas entre las variables (**Figura 1**). Observamos que cada una de las variables se encuentran correlacionadas (positivamente) a pares.

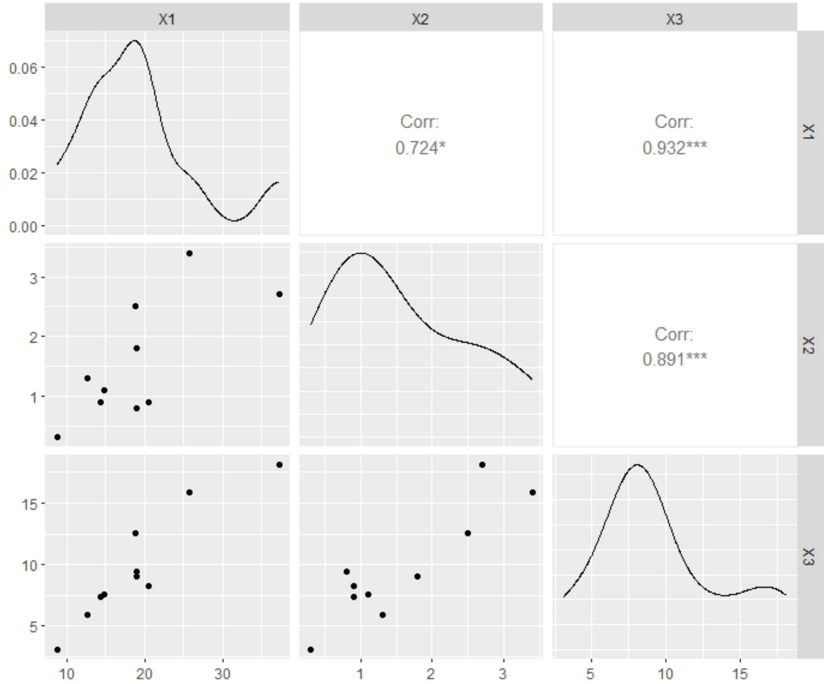

Figura 1. Diagrama de dispersión múltiple.

b) Los resultados obtenidos fueron los siguientes

Medias muestrales obtenidas

$$\bar{x_1} = 19{,}05 \qquad \bar{x_2} = 1{,}57$$

Varianzas muestrales obtenidas

$$s_{11} = 56{,}968 \qquad s_{22} = 0{,}894$$

Covarianza entre las variables X_1 y X_2

$$s_{12} = 5{,}745$$

Por último, el coeficiente de correlación entre las variables X_1 y X_2 es de $r_{12} = 0{,}724.$ Lo que indica una correlación positiva entre estas dos variables, variables que corresponden a la duración media de la hipoteca en años y al precio promedio en millones de euros.

c) La dependencia lineal entre todos los pares de variables de mide por la matriz de correlacioón, denotada por $R.$ Es una matriz simetrica de tamaño pxp, la cual contiene unos en la diagonal y los coeficientes de correlación entre los pares de variables fuera de la diagonal.

La matriz R está relacionada con la matriz de varianzas y covarianzas S mediante

$$R = D^{-\frac{1}{2}} S D^{-\frac{1}{2}}$$

considerando esta relación, se calculó la matriz S con el objetivo de emplearla para el calculó de la matriz R.

Para esto se calculó en primera instancia la matriz de centrado P y la matriz de varianzas y covarianzas S, considerando que

$$P = I - \frac{1}{n} 11' \qquad S = \frac{1}{n} X' P X$$

```r
n <- nrow(X) # no. of rows
P <- diag(n) - (1/n)*matrix(1, n, n) # P matrix; centering matrix plays an important role since we will use it to remove the column means from a matrix, centering the matrix

x_bar <- colMeans(X) # the mean vector consists of the means of each variable; x_bar dimension 1 row by 3 columns

S <- (1/n)*t(X)%*%P%*%X
print(S) # the variance-covariance matrix consists of the variances of the variables along the main diagonal and the covariances between each pair of variables in the other matrix positions

D <- diag(diag(S))
D_inv <- diag(1 / sqrt(diag(D)))

R <- D_inv %*% S %*% D_inv # correlation coefficients between variables
```

De esta manera se obtuvo la matriz de correlaciones R.

$$R = \begin{pmatrix} 1{,}0000 & 0{,}7244 & 0{,}9321 \\ 0{,}7244 & 1{,}0000 & 0{,}8906 \\ 0{,}9321 & 0{,}8906 & 1{,}0000 \end{pmatrix}$$

Ejercicio 3. Considérese la muestra $x_1, ..., x_n$ de vectores en R^p. Pruébese que la matriz de covarianzas $S = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})'$, se puede expresar como $\frac{1}{n}\sum_{i=1}^{n} x_i {x_i}' - \bar{x}\bar{x}'$.

Solución

Para una muestra de observaciones X que provienen de una variable multivariada se define la matriz de varianzas y covarianzas como

$$S = \frac{1}{n}\sum_{i=n}^{n}(x_i - \bar{x})(x_i - \bar{x})'$$

Desarrollando la expresión

$$\begin{aligned}
S &= \frac{1}{n}\sum_{i=n}^{n}\left(x_i{x_i}' - x_i\bar{x}' - \bar{x}{x_i}' + \bar{x}\bar{x}'\right) \\
&= \frac{1}{n}\Big(\sum_{i=n}^{n} x_i{x_i}' - \sum_{i=n}^{n} x_i\bar{x}' - \sum_{i=n}^{n}\bar{x}{x_i}' + \sum_{i=n}^{n}\bar{x}\bar{x}'\Big) \\
&= \frac{1}{n}\Big(\sum_{i=n}^{n} x_i{x_i}' - \sum_{i=n}^{n} x_i\bar{x}' - \sum_{i=n}^{n} x_i\bar{x}' + \sum_{i=n}^{n}\bar{x}\bar{x}'\Big) \\
&= \frac{1}{n}\sum_{i=n}^{n} x_i{x_i}' - 2\frac{1}{n}\sum_{i=n}^{n} x_i\bar{x}' + \frac{1}{n}\sum_{i=n}^{n}\bar{x}\bar{x}' \\
&= \frac{1}{n}\sum_{i=n}^{n} x_i{x_i}' - 2\bar{x}\bar{x}' + \bar{x}\bar{x}' \\
&= \frac{1}{n}\sum_{i=n}^{n} x_i{x_i}' - \bar{x}\bar{x}'
\end{aligned}$$

De esta manera demostramos que la matriz de covarianzas S se puede expresar como

$$S = \frac{1}{n}\sum_{i=n}^{n} x_i{x_i}' - \bar{x}\bar{x}'$$

Ejercicio 4. Considere una población normal bivariada con $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_{11} = 2$, $\sigma_{22} = 1$ y $\rho_{12} = 0{,}5$.

a) Escriba la densidad normal bivariada explicitamente.

b) Escriba la expresión de distancia cuadrada generalizada $(x-\mu)'\Sigma^{-1}(x-\mu)$ como función de x_1 y x_2.

c) Determine y grafique el contorno de densidad constante que contiene el 50 % de la probabilidad.

d) Especifique la distribución condicional de X_1 dado que $X_2 = x_2$.

Solución

a) La función de densidad de una distribución normal bivariada se encuentra dada de la siguiente manera

$$f(x) = \frac{1}{(2\pi)|\Sigma|^{\frac{1}{2}}} e^{-(x-\mu)'\Sigma^{-1}(x-\mu)/2} \qquad -\infty \leq x_1 \geq \infty, \text{ con } i = 1, ..., 2$$

donde $-(x-\mu)'\Sigma^{-1}(x-\mu)/2$ es la distancia cuadrada generalizada de x a μ.

Sea

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

Sustituyendo en $-(x-\mu)'\Sigma^{-1}(x-\mu)/2$, podemos reescribir la función de densidad de probabilidad normal bivariada como

$$f(x) = \frac{1}{(2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)})} e^{-\frac{1}{2(1-\rho_{12}^2)}[(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}})^2 + (\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}})^2 - 2\rho_{12}(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}})(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}})]}$$

Reemplazando los valores proporcionados de $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_{11} = 2$, $\sigma_{22} = 1$ y $\rho_{12} = 0{,}5$

$$= \frac{1}{2\pi\sqrt{\frac{3}{2}}} e^{\left(-\frac{2}{3}\left[(\frac{x_1}{\sqrt{2}})^2 + (x_2-2)^2 - (\frac{x_1}{\sqrt{2}})(x_2-2)\right]\right)}$$

$$= \frac{1}{2\pi\sqrt{\frac{3}{2}}} e^{\left(-\frac{2}{3}\left[\frac{x_1^2}{2} + (x_2-2)^2 - \frac{x_1(x_2-2)}{\sqrt{2}}\right]\right)}$$

De manera que la función de densidad se encuentra definida de la siguiente manera

$$f(x) = \frac{1}{2\pi\sqrt{\frac{3}{2}}} e^{\left(-\frac{2}{3}\left[\frac{x_1^2}{2} + (x_2-2)^2 - \frac{x_1(x_2-2)}{\sqrt{2}}\right]\right)}$$

b) La distancia cuadrada generalizada se puede escribir como

$$(x-\mu)'\Sigma^{-1}(x-\mu) = \frac{1}{1-\rho_{12}^2}\big[\big(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\big)^2 + \big(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\big)^2 - 2\rho_{12}\big(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\big)\big(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\big)\big]$$

Escribiendo la expresión en función de x_1 y x_2

$$= \frac{1}{1-(\frac{1}{2})^2}\big[\big(\frac{x_1}{\sqrt{2}}\big)^2 + (x_2-2)^2 - \big(\frac{x_1}{\sqrt{2}}\big)(x_2-2)\big]$$

$$= \frac{1}{\frac{3}{4}}\big[\frac{x_1^2}{2} + (x_2-2)^2 - \frac{x_1(x_2-2)}{\sqrt{2}}\big]$$

$$= \frac{4}{3}\big[\frac{x_1^2}{2} + (x_2-2)^2 - \frac{x_1(x_2-2)}{\sqrt{2}}\big]$$

Por lo tanto, la distancia cuadrada generalizada queda definida de la siguiente manera

$$(x-\mu)'\Sigma^{-1}(x-\mu) = \frac{4}{3}\left[\frac{x_1^2}{2} + (x_2-2)^2 - \frac{x_1(x_2-2)}{\sqrt{2}}\right]$$

c) Sabemos que la distancia generalizada $(x-\mu)'\Sigma^{-1}(x-\mu)$ se distribuye como una distribución χ^2 con p grados de libertad. Por lo que el elipsoide satisface

$$(x-\mu)'\Sigma^{-1}(x-\mu) \leq \chi^2(\alpha)$$

con una probabilidad $1-\alpha$; donde $\chi^2(\alpha)$ denota el percentil superior $(100\alpha\,\%)$ de la distribución χ^2.

De esta manera consideramos un valor $\alpha = 0{,}5$, esperando que el 50 % de los datos se encuentren dentro del contorno estimado.

Para esto se obtuvieron los valores y vectores propios.

```r
sigma <- matrix(c(2, rho*sqrt(2), rho*sqrt(2), 1), 2)  # variance-covariance matrix

eigenv <- eigen(sigma) # eigenvalues and eigenvectors
```

Valores propios

$$(\lambda_1, \lambda_2) = (2{,}366, 0{,}633)$$

Vectores propios

$$(e_1, e_2) = \begin{pmatrix} -0{,}888 & 0{,}459 \\ -0{,}459 & -0{,}888 \end{pmatrix}$$

Los vectores propios determinan la dirección de los ejes y los valores propios su longitud.

Se realizó una simulación de la distribución normal bivariada con los parámetros correspondientes y se graficó la región definida por el elipsoide.

```r
# Simulate from a Multivariate Normal Distribution

n <- 10000 # no. of samples
mu <- c(0, 2) # vector giving the means of the variables
sigma <- sigma # variance-covariance matrix

data <- mvrnorm(n, mu, sigma)
distances <- mahalanobis(data, colMeans(data), cov(data)) # measure between a sample point and a distribution; D^{2} = (x - \mu){'}\Sigma^{-1}(x - \mu)

plot(data, pch = ".", xlab = "x1", ylab = "x2", main = "Title")
points(data[distances > 1.386,], pch = 20, cex = 0.9, col = 'Steel Blue 2')
points(data[(distances - 0.6) < distances & distances < (1.386+0.6),], pch = 20, cex = 0.9, col = 'Orange 2')
```

El gráfico muestra los puntos que se sitúan fuera de la región del 50 % definida por el elipsoide. Los puntos azules representan aquellos que están fuera de esta región, mientras que los puntos naranjas indican aquellos que están dentro (**Figura 2**).

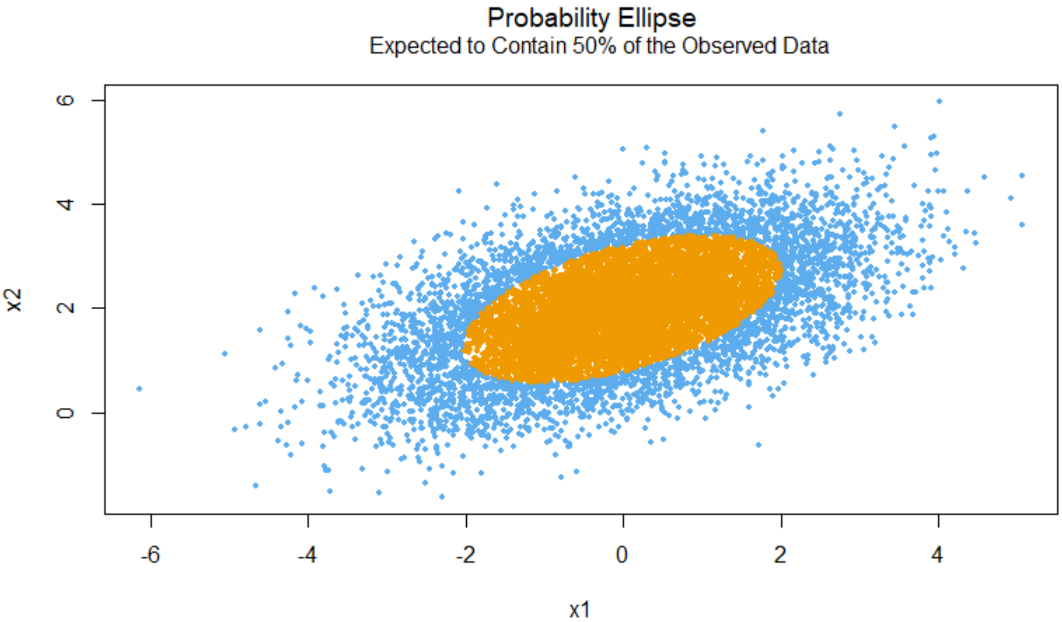

Figura 2. Elipsoide de Probabilidad. Contorno de densidad conteniendo el 50 % de los datos.

d) Sea $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ con distribución $N_p(\mu, \Sigma)$ con $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ con $|\Sigma_{22}| > 0$. Entonces la distribución condicional de X_1 dado $X_2 = x_2$, es normal con

$$Mean = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$$

y

$$Covariance = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Nótese que la covarianza no depende del valor x_2 de la variable condicionante.

En este caso tenemos $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N_p(\mu, \Sigma)$ con $\mu = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 2 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 1 \end{bmatrix}$, donde $|\Sigma_{22} = 1| > 0$. Por lo tanto la distribución condicional de X_1 dado $X_2 = x_2$ es normal con

$$Mean = \frac{1}{\sqrt{2}}(x_2 - 2) = \frac{x_2 - 2}{\sqrt{2}}$$

y

$$Covariance = 2 - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} = \frac{3}{2}$$

De manera que la distribución condicional de X_1 dado $X_2 = x_2$, $X_1|X_2 = x_2 \sim N(\frac{x_2-2}{\sqrt{2}}, \frac{3}{2})$.

Ejercicio 5. Sea X un vector aleatorio de distribución normal con media $\mu = (-1, 1, 0)'$ y matriz de covarianzas

$$\Sigma = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

a) Hállese la distribución de $X_1 + 2X_2 - 3X_3$.

b) Hállese un vector a_{2x1} tal que las variables X_1 y $X_1 - a' \begin{pmatrix} X_2 \\ X_3 \end{pmatrix}$ sean independientes.

c) Calcúlese la distribución de X_3 condicionada a $X_1 = x_1$ y $X_2 = x_2$.

Solución

a) Si A es una matrix qxp, $x \sim N_p(\mu, \Sigma)$, entonces cualquier conjunto de q combinaciones lineales

$$Ax = \begin{bmatrix} \sum_{i=1}^{p} a_{1i}x_i \\ \sum_{i=1}^{p} a_{2i}x_i \\ \vdots \\ \sum_{i=1}^{p} a_{qi}x_i \end{bmatrix} \sim N_q(A\mu, A\Sigma A')$$

Además, si d es un vector conformado de constantes, entonces $x + d \sim N_p(\mu + d, \Sigma)$.

Considerando lo anterior, podemos expresar la combinación lineal de la siguiente manera

$$X_1 + 2X_2 - 3X_3 = (1 \ 2 \ -3) \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

de manera que $A = (1 \ 2 \ -3)$.

Entonces

$$A\mu = (1 \ 2 \ -3) \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = -1 + 2 + 0 = 1$$

$$A\Sigma' A = (1 \ 2 \ -3) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = 13$$

Por lo que $X_1 + 2X_2 - 3X_3$ sigue una distribución $N(-1, 13)$

c) Sea $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ con distribución $N_p(\mu, \Sigma)$ con $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ con $|\Sigma_{22}| > 0$. Entonces la distribución condicional de X_1 dado $X_2 = x_2$, es normal con

$$Mean = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$$

y

$$Covariance = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Tenemos que $\mu_1 = 0$, $\mu_2 = (-1 \ 1)'$, $\Sigma_{11} = 2$, $\Sigma_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$, $\Sigma_{12} = (1 \ 1)'$. Por lo tanto la distribución de X_3 condicionada a $X_1 = x_1$ y $X_2 = x_2$ es normal con

$$Mean = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) = (-1 \ 1) \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}^{-1} \begin{pmatrix} x_1 + 1 \\ x_2 - 1 \end{pmatrix}$$

$$= (-1\ 1)\begin{pmatrix}1 & 0\\ 0 & \frac{1}{3}\end{pmatrix}\begin{pmatrix}x_1+1\\ x_2-1\end{pmatrix}$$

$$= (-1\ \frac{1}{3})\begin{pmatrix}x_1+1\\ x_2-1\end{pmatrix}$$

$$= -x_1 - 1 + \frac{1}{3}X_2 - \frac{1}{3}$$

$$Mean = -x_1 + \frac{1}{3}X_2 - \frac{4}{3}$$

$$Covariance = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = 2 - (1\ 1)\begin{pmatrix}1 & 0\\ 0 & 3\end{pmatrix}^{-1}\begin{pmatrix}1\\ 1\end{pmatrix}$$

$$= 2 - (1\ 1)\begin{pmatrix}1 & 0\\ 0 & \frac{1}{3}\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}$$

$$= 2 - (1\ \frac{1}{3})\begin{pmatrix}1\\ 1\end{pmatrix}$$

$$= 2 - \frac{4}{3}$$

$$Covariance = \frac{2}{3}$$

La distribución de X_3 condicionada a $X_1 = x_1$ y $X_2 = x_2$, $X_3|X_1 = x_1, X_2 = x_2 \sim N(-x_1+\frac{1}{3}x_2-\frac{4}{3}, \frac{2}{3})$