

LCE 5861 - MODELOS LINEARES I

Prof. César Gonçalves de Lima (FZEA/USP)

cegdlima@usp.br

Baseado no livro:

Rencher, A. C.; Schaalje, G. B. (2008). *Linear Models in Statistics*,
2nd edition. New York: Wiley

CAPÍTULO 4. DISTRIBUIÇÃO NORMAL MULTIVARIADA

Nas inferências feitas em modelos lineares assumimos que o vetor aleatório de interesse tem **distribuição normal multivariada**.

4.1. FUNÇÃO DENSIDADE NORMAL UNIVARIADA

A função densidade de probabilidades (f.d.p.) de uma variável aleatória normal univariada y , de média $E(y) = \mu$ e $var(y) = \sigma^2$ é dada por:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(y-\mu)^2}{2\sigma^2} \right\} \quad (4.3)$$

Lembrando que um vetor aleatório $\mathbf{y} = [y_1, y_2, \dots, y_p]'$ envolve p variáveis aleatórias avaliadas em um mesmo indivíduo.

4.2. FUNÇÃO DE DENSIDADE NORMAL MULTIVARIADA

A f.d.p. do vetor aleatório normal multivariado $\mathbf{y} = [y_1, y_2, \dots, y_p]'$, com $E(\mathbf{y}) = \boldsymbol{\mu}$ e $cov(\mathbf{y}) = \boldsymbol{\Sigma}$, em que $\boldsymbol{\mu}$ é um vetor $p \times 1$ e $\boldsymbol{\Sigma}$ é uma matriz $p \times p$, positiva definida, é dada por:

$$f(\mathbf{y}) = \frac{1}{(\sqrt{2\pi})^p |\boldsymbol{\Sigma}|^{1/2}} e^{\{-(\mathbf{y}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y}-\boldsymbol{\mu})/2\}} \quad (4.9)$$

Quando $\mathbf{y}$ tem a densidade (4.9) dizemos que “o vetor aleatório $\mathbf{y}$ tem distribuição normal p -variada com vetor de médias $\boldsymbol{\mu}$ e matriz de variâncias e covariâncias $\boldsymbol{\Sigma}$ ”, ou simplesmente, indicamos que

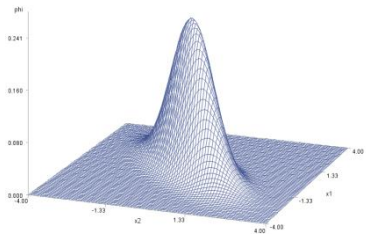
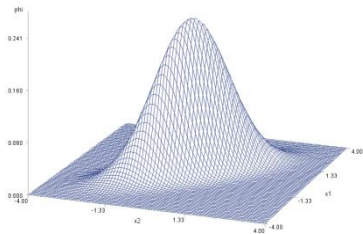
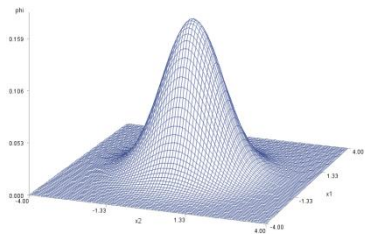
$$\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Construção do gráfico da normal bivariada no SAS:

```
%let r=-0.75;      * Fixa o coeficiente de correlação entre y1 e y2;

data a;
  pi=3.1416;
  do y1=-4 to 4 by 0.1;
    do y2=-4 to 4 by 0.1;
      phi=exp(-(y1*y1-2*&r*y1*y2+y2*y2)/2/(1-&r*&r))/2/pi/sqrt(1-&r*&r);
      output;
    end;
  end;
run;

options reset=all border;
proc g3d;
  title 'Densidade Normal Bivariada (r = ' &r ')';
  plot y1*y2=phi / rotate=-20;
run;
```

Densidade Normal Bivariada ($r = -0.75$)Densidade Normal Bivariada ($r = 0.75$)Densidade Normal Bivariada ($r = 0$)

4.3 FUNÇÕES GERADORAS DE MOMENTOS

Para maiores detalhes: ver Rencher (2008), pág. 90-92.

Momentos são valores que expressam características importantes de uma variável aleatória e são calculados por meio do valor esperado de potências dessa variável.

O **momento de ordem k** , para $k = 1, 2, \dots$, da v.a. Y é definido como $E(Y^k)$, desde que esta quantidade exista.

Exemplo: $\mu = E(Y)$ é momento de primeira ordem da v.a. Y .

O **momento central de ordem k** é definido como $E[(Y - \mu)^k]$ sempre que essa quantidade existir e $E(Y) < \infty$.

Exemplo: $\sigma^2 = \text{var}(y) = E(y - \mu)^2$ é o momento central de segunda ordem da variável Y .

Mais exemplos:

- **Coefficiente de assimetria (*skewness*):** é o quociente entre o 3º momento central e a raiz quadrada do 2º momento central da distribuição.
- **Coefficiente de achatamento (*kurtosis*):** é o quociente entre o 4º momento central e o quadrado do 2º momento central da distribuição.

A **função geradora de momentos** (*f. g. m.*) de uma variável aleatória Y é definida como:

$$M_y(t) = E(e^{ty}) = \int_{-\infty}^{\infty} e^{ty} f(y) dy \quad (4.10)$$

desde que $E(e^{ty})$ exista na vizinhança $-h < t < h$ para algum $h \in R$, positivo e próximo de zero.

Detalhes em: <https://www.teses.usp.br/teses/disponiveis/11/11134/tde-20220207-184408/publico/CorreaAngelaMariaCassaviaJorge.pdf>

Supondo que a f.g.m. de Y exista e seja conhecida, então o seu k -ésimo momento pode ser obtido calculando a derivada de ordem k ($k = 1, 2, 3, \dots$) em $t = 0$, ou seja:

$$M_y^{(k)}(t = 0) = \left. \frac{\partial^k M_y(t)}{\partial t^k} \right|_{t=0} = E(Y^k) \quad (4.14)$$

Exemplo: A f.g.m. de uma v.a. com distribuição normal de média μ e variância σ^2 , ou seja, $Y \sim N(\mu, \sigma^2)$, é dada por:

$$M_Y(t) = e^{t\mu + \frac{t^2\sigma^2}{2}}$$

Então:

$$\frac{\partial M_Y(t)}{\partial t} = (\mu + t\sigma^2) e^{t\mu + t^2\sigma^2/2} \quad \Rightarrow M_Y^{(1)}(t=0) = E(Y) = \mu$$

$$\frac{\partial^2 M_Y(t)}{\partial t^2} = M_Y^{(2)}(t) = [(\mu + t\sigma^2)^2 + \sigma^2] e^{t\mu + t^2\sigma^2/2}$$

$$\Rightarrow M_Y^{(2)}(t=0) = E(Y^2) = \mu^2 + \sigma^2$$

Mas

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2 = M_Y^{(2)}(t=0) - [M_Y^{(1)}(t=0)]^2$$

$$\Rightarrow \text{var}(Y) = (\mu^2 + \sigma^2) - (\mu)^2 = \sigma^2$$

Vale observar que:

- A f.g.m. facilita o cálculo dos momentos e caracteriza a distribuição em alguns aspectos muito importantes.
- A f.g.m. determina completamente uma distribuição de probabilidades: Se duas variáveis aleatórias possuem funções geradoras de momentos iguais, elas têm a mesma distribuição de probabilidades.

Por analogia ao caso univariado a f.g.m. de um vetor aleatório $\mathbf{y} = [y_1, y_2, \dots, y_p]'$ é definida como:

$$M_{\mathbf{y}}(\mathbf{t}) = E(e^{t_1 y_1 + \dots + t_p y_p}) = E(e^{\mathbf{t}' \mathbf{y}}) \quad (4.15)$$

Por analogia a (4.13) temos que:

$$\frac{dM_{\mathbf{y}}(\mathbf{t}=\mathbf{0})}{dt} = E(\mathbf{y}) \quad (4.16)$$

Similarmente, $\frac{\partial^2 M_{\mathbf{y}}(\mathbf{t})}{\partial t_r \partial t_s}$ avaliada em $t_r = t_s = 0$ fornece $E(y_r y_s)$ que pode ser usada para calcular a **covariância** entre y_r e y_s , utilizando a fórmula:

$$\text{cov}(y_r, y_s) = E(y_r y_s) - E(y_r)E(y_s)$$

Teorema 4.3A. Se $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ então a sua f.g.m. é dada por

$$M_{\mathbf{y}}(\mathbf{t}) = e^{\mathbf{t}'\boldsymbol{\mu} + \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2} \quad (4.18)$$

Corolário 1. A função geradora de momentos para $(\mathbf{y} - \boldsymbol{\mu})$ é

$$M_{\mathbf{y}}(\mathbf{t}) = e^{\mathbf{t}' \boldsymbol{\Sigma} \mathbf{t} / 2} \quad (4.22)$$

Importante:

1. Se dois vetores aleatórios têm a mesma função geradora de momentos $\Rightarrow$ eles têm a mesma função densidade.
2. Dois vetores aleatórios são **independentes** se e somente se a sua f.g.m. conjunta puder ser **fatorada** no produto das suas f.g.m.'s.

Para $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ e $\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$, tem-se que:

$$y_1 \text{ e } y_2 \text{ são } \underline{\text{independentes}} \Leftrightarrow M_{\mathbf{y}}(\mathbf{t}) = M_{y_1}(t_1) M_{y_2}(t_2) \quad (4.23)$$

4.4. PROPRIEDADES DA DISTRIBUIÇÃO NORMAL MULTIVARIADA

Teorema 4.4A. Sejam $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\mathbf{a}$ um vetor $p \times 1$ de constantes e $\mathbf{A}$ uma matriz $k \times p$ de constantes e de posto $k \leq p$. Então:

i) a variável aleatória $z = \mathbf{a}'\mathbf{y} \sim N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})$

ii) o vetor aleatório $\mathbf{z} = \mathbf{A}\mathbf{y} \sim N_k(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$

Corolário 1. Se $\mathbf{b}$ é um vetor $k \times 1$ de constantes então

$$\mathbf{z} = \mathbf{A}\mathbf{y} + \mathbf{b} \sim N_k(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$$

Obs.: As provas de diversos Teoremas deste capítulo são baseadas na função geradora de momentos da normal uni ou multivariada.

Teorema 4.4B. Se $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ então **qualquer subvetor** $r \times 1$ de $\mathbf{y}$ tem uma distribuição normal r -variada com médias, variâncias e covariâncias iguais às da distribuição normal p -variada original.

Corolário 1. Se $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ então **qualquer variável individual** y_i em $\mathbf{y}$ é distribuída como $N(\mu_i, \sigma_{ii})$, ou seja, $y_i \sim N(\mu_i, \sigma_i^2)$.

Exemplo: $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, com $\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ e $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 5 \end{bmatrix}$ então:

- $\begin{bmatrix} y_1 \\ y_3 \end{bmatrix} \sim N_2\left(\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}; \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}\right)$
- $y_2 \sim N(\mu = 1; \sigma^2 = 2)$

Se o vetor aleatório $\mathbf{v}$ é particionado em dois vetores denotados por $\mathbf{y}$ e $\mathbf{x}$, de dimensões $p \times 1$ e $q \times 1$, respectivamente. O vetor de médias e a matriz de covariâncias do vetor aleatório $\mathbf{v} = \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix}$ ficam:

$$\boldsymbol{\mu} = E(\mathbf{v}) = \begin{bmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{bmatrix} \quad \boldsymbol{\Sigma} = cov(\mathbf{v}) = \begin{bmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{bmatrix}$$

Teorema 4.4C. Se o vetor particionado $\mathbf{v} = \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} \sim N_{p+q}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ então os subvetores aleatórios $\mathbf{y}$ e $\mathbf{x}$ são independentes se $\boldsymbol{\Sigma}_{xy} = \mathbf{0}$.

Corolário 1. Se $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ então quaisquer duas variáveis individuais y_i e y_j são independentes se $\sigma_{ij} = 0$.

Corolário 2. Se $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e se $\text{cov}(\mathbf{A}\mathbf{y}, \mathbf{B}\mathbf{y}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{B}' = \mathbf{0}$ então $\mathbf{A}\mathbf{y}$ e $\mathbf{B}\mathbf{y}$ são dois conjuntos de variáveis independentes.

Teorema 4.4D. Se $\mathbf{y}$ e $\mathbf{x}$ têm distribuição conjunta normal multivariada com $\boldsymbol{\Sigma}_{yx} \neq \mathbf{0}$ então a **distribuição condicional** de $\mathbf{y}$ dado $\mathbf{x}$, $f(\mathbf{y}|\mathbf{x})$, é **normal multivariada** com vetor de média e matriz de covariâncias dados por:

$$E(\mathbf{y}|\mathbf{x}) = \boldsymbol{\mu}_y + \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}(\mathbf{x} - \boldsymbol{\mu}_x) \quad (4.26)$$

$$\text{cov}(\mathbf{y}|\mathbf{x}) = \boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xy} \quad (4.27)$$

(*Prova:* ver Rencher, pág. 95-96)

Veremos exemplos a seguir.

- Como (4.26) é uma função linear de $\mathbf{x}$, qualquer par de variáveis y_i e y_j em um vetor normal multivariado exibe uma **tendência linear**:

$$E(y_i|y_j) = \mu_i + (\sigma_{ij}/\sigma_{jj})(y_j - \mu_j)$$

- A covariância σ_{ij} está relacionada com a inclinação da reta que associa y_i e y_j e é uma medida útil para quantificar a relação linear entre as duas variáveis.
- A matriz de covariâncias condicionais em (4.26) não envolve valores de $\mathbf{x}$.
- Para algumas distribuições não normais, $cov(\mathbf{y}|\mathbf{x})$ é função de $\mathbf{x}$.

Corolário 1. Se o vetor $\mathbf{v} = [y, x_1, \dots, x_q]' = \begin{bmatrix} y \\ \mathbf{x}' \end{bmatrix}$ com $\boldsymbol{\mu} = \begin{bmatrix} \mu_y \\ \boldsymbol{\mu}_x \end{bmatrix}$ e $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_y^2 & \boldsymbol{\sigma}_{yx}^t \\ \boldsymbol{\sigma}_{yx} & \boldsymbol{\Sigma}_{xx} \end{bmatrix}$, então $y|\mathbf{x}$ tem distribuição normal univariada com:

$$E(y|\mathbf{x}) = \mu_y + \boldsymbol{\sigma}_{yx}^t \boldsymbol{\Sigma}_{xx}^{-1} (\mathbf{x} - \boldsymbol{\mu}_x) \quad (4.33)$$

$$var(y|\mathbf{x}) = \sigma_y^2 - \boldsymbol{\sigma}_{yx}^t \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\sigma}_{yx} \quad (4.34)$$

Como (4.34), $\boldsymbol{\sigma}_{yx}^t \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\sigma}_{yx} \geq 0$ porque $\boldsymbol{\Sigma}_{xx}^{-1}$ é positiva definida, então:

$$var(y|\mathbf{x}) \leq var(y) \quad (4.35)$$

Exemplo 4.4(a). Para ilustrar os Teoremas 4.4.A, 4.4B e 4.4C, vamos

usar o vetor aleatório $\mathbf{y} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ com $\boldsymbol{\mu} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$ e $\boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix}$.

i) Para $z = y_1 - 2y_2 + y_3 = [1 \ -2 \ 1]\mathbf{y} = \mathbf{a}'\mathbf{y}$ tem-se:

$$E(z) = \mathbf{a}'\boldsymbol{\mu} = [1 \ -2 \ 1] \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = 3$$

$$var(z) = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a} = [1 \ -2 \ 1] \begin{bmatrix} 4 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 19 \Rightarrow z \sim N(3, 19)$$

ii) As combinações lineares $z_1 = y_1 - y_2 + y_3$ e $z_2 = 3y_1 + y_2 - 2y_3$ podem ser escritas como: $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 3 & 1 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \mathbf{A}\mathbf{y}$. Pelos Teoremas 3.6B(i) e 3.6D(i) tem-se:

$$E(\mathbf{z}) = \mathbf{A}\boldsymbol{\mu} = \begin{bmatrix} 1 & -1 & 1 \\ 3 & 1 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

$$\text{cov}(\mathbf{z}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}' = \begin{bmatrix} 1 & -1 & 1 \\ 3 & 1 & -2 \end{bmatrix} \begin{bmatrix} 4 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -1 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 4 \\ 4 & 29 \end{bmatrix}$$

$\Rightarrow$ Pelo Teorema 4.4A(ii) tem-se que: $\mathbf{z} \sim N_2 \left(\begin{bmatrix} 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 14 & 4 \\ 4 & 29 \end{bmatrix} \right)$

iii) Para ilustrar a distribuição marginal no Teorema 4.4B, tem-se que $y_1 \sim N(3, 4)$, $y_2 \sim N(1, 1)$ e $y_3 \sim N(2, 3)$ e que:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \sim N_2 \left(\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \right) \quad \text{e} \quad \begin{bmatrix} y_1 \\ y_3 \end{bmatrix} \sim N_2 \left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} \right)$$

iv) Para ilustrar o Teorema 4.4.C note que $\text{cov}(y_1, y_2) = \sigma_{12} = 0 \Rightarrow y_1$ e y_2 são independentes.

Exemplo 4.4(b). Para ilustrar o Teorema 4.4D, consideremos o vetor aleatório $\mathbf{v} \sim N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, em que:

$$\boldsymbol{\mu} = \begin{bmatrix} 2 \\ 5 \\ -2 \\ 1 \end{bmatrix} \text{ e } \boldsymbol{\Sigma} = \begin{bmatrix} 9 & 0 & 3 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & -1 & 6 & -3 \\ 3 & 2 & -3 & 7 \end{bmatrix}$$

i) Se $\mathbf{v}$ é particionado como $\mathbf{v} = [y_1, y_2, x_1, x_2]'$ então:

$$\boldsymbol{\mu}_y = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \boldsymbol{\mu}_x = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \boldsymbol{\Sigma}_{yy} = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{\Sigma}_{yx} = \begin{bmatrix} 3 & 3 \\ -1 & 2 \end{bmatrix} \text{ e } \boldsymbol{\Sigma}_{xx} = \begin{bmatrix} 6 & -3 \\ -3 & 7 \end{bmatrix}$$

De (4.26) e (4.27) obtemos:

$$\begin{aligned} E(\mathbf{y} | \mathbf{x}) &= \boldsymbol{\mu}_y + \boldsymbol{\Sigma}_{yx} \boldsymbol{\Sigma}_{xx}^{-1} (\mathbf{x} - \boldsymbol{\mu}_x) \\ &= \begin{bmatrix} 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 3 & 3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 6 & -3 \\ -3 & 7 \end{bmatrix}^{-1} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right) = \dots \\ \Rightarrow E(\mathbf{y} | \mathbf{x}) &= \begin{bmatrix} 3 + \frac{10}{11}x_1 + \frac{9}{11}x_2 \\ \frac{14}{3} - \frac{1}{33}x_1 + \frac{3}{11}x_2 \end{bmatrix} \end{aligned}$$

Por (4.27) tem-se que:

$$\begin{aligned}
 cov(\mathbf{y}|\mathbf{x}) &= \mathbf{\Sigma}_{yy} - \mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy} \\
 &= \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 6 & -3 \\ -3 & 7 \end{bmatrix}^{-1} \begin{bmatrix} 3 & -1 \\ 3 & 2 \end{bmatrix} \\
 &= \frac{1}{33} \begin{bmatrix} 126 & -24 \\ -24 & 14 \end{bmatrix} \\
 \therefore \mathbf{y}|\mathbf{x} &\sim N_2 \left(\begin{bmatrix} 3 + \frac{10}{11}x_1 + \frac{9}{11}x_2 \\ \frac{14}{3} - \frac{1}{33}x_1 + \frac{3}{11}x_2 \end{bmatrix}, \frac{1}{33} \begin{bmatrix} 126 & -24 \\ -24 & 14 \end{bmatrix} \right)
 \end{aligned}$$

Note que:

$$var(y_1) = 9 \quad \text{e} \quad var(y_1|x_1, x_2) = 126/33 \cong 3,82$$

$$var(y_2) = 1 \quad \text{e} \quad var(y_2|x_1, x_2) = 14/33 \cong 0,42$$

Exemplo 4.4(c). Para ilustrar o Corolário 1 do Teorema 4.4D, continuamos com um vetor $\mathbf{v} \sim N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, em que $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ são apresentados no Exemplo 4.4(b).

$$\boldsymbol{\mu} = \begin{bmatrix} 2 \\ 5 \\ -2 \\ 1 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 9 & 0 & 3 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & -1 & 6 & -3 \\ 3 & 2 & -3 & 7 \end{bmatrix}$$

Se $\mathbf{v} = [y, x_1, x_2, x_3]'$ então:

$$\mu_y = 2 \text{ e } \text{var}(y) = 9 \quad \boldsymbol{\mu}_x = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} \text{ e } \boldsymbol{\Sigma}_{xx} = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 6 & -3 \\ 2 & -3 & 7 \end{bmatrix}$$

i) De (4.33) temos que:

$$\begin{aligned}
 E(y|x_1, x_2, x_3) &= \mu_y + \boldsymbol{\sigma}_{yx}^t \boldsymbol{\Sigma}_{xx}^{-1} (\mathbf{x} - \boldsymbol{\mu}_x) \\
 &= 2 + [0 \ 3 \ 3] \begin{bmatrix} 1 & -1 & 2 \\ -1 & 6 & -3 \\ 2 & -3 & 7 \end{bmatrix}^{-1} \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} \right) \\
 &= \frac{95}{7} - \frac{12}{7}x_1 + \frac{6}{7}x_2 + \frac{9}{7}x_3
 \end{aligned}$$

ii) De (4.34) nós obtemos

$$\begin{aligned}
 var(y|x_1, x_2, x_3) &= \sigma_y^2 - \boldsymbol{\sigma}_{yx}^t \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\sigma}_{yx} \\
 &= 9 - [0 \ 3 \ 3] \begin{bmatrix} 1 & -1 & 2 \\ -1 & 6 & -3 \\ 2 & -3 & 7 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix} = \frac{18}{7} \cong 2,57
 \end{aligned}$$

Assim, temos que:

$$y|x_1, x_2, x_3 \sim N\left(\frac{95}{7} - \frac{12}{7}x_1 + \frac{6}{7}x_2 + \frac{9}{7}x_3, \frac{18}{7}\right)$$

e que

$$\text{var}(y|x_1, x_2, x_3) = \frac{18}{7} \cong 2,57 < \text{var}(y) = 9$$

o que ilustra (4.35).

4.5. CORRELAÇÃO PARCIAL

Por conveniência, usaremos a notação dos Teoremas 4.4C e 4.4D, admitindo que o vetor $\mathbf{v}$ é formado por um subconjunto de y 's que inclui as variáveis y_i e y_j , é denotado por $\mathbf{y}$ e o outro subconjunto de y 's é denotado por $\mathbf{x}$.

Seja $\mathbf{v} \sim N_{p+q}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e seja $\mathbf{v}$, $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ particionadas como no Teorema 4.4C e 4.4D, ou seja:

$$\mathbf{v} = \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} \quad \boldsymbol{\mu} = E(\mathbf{v}) = \begin{bmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{bmatrix} \quad \boldsymbol{\Sigma} = cov(\mathbf{v}) = \begin{bmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{bmatrix}$$

A covariância entre y_i e y_j na **distribuição condicional** de $\mathbf{y}$ dado $\mathbf{x}$ será denotada por $\sigma_{ij.rs\dots q}$ em que y_i e y_j são duas variáveis do vetor de variáveis $\mathbf{y}$ e $\mathbf{x} = [y_r, y_s, \dots, y_q]'$ é o vetor com as variáveis que terão seus valores fixados.

Desse modo $\sigma_{ij.rs\dots q}$ é o (ij) -ésimo elemento da matriz

$$\text{cov}(\mathbf{y}|\mathbf{x}) = \mathbf{\Sigma}_{yy} - \mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}$$

Por exemplo:

- $\sigma_{13.567}$: representa a **covariância** entre y_1 e y_3 na distribuição condicional de $\mathbf{y} = [y_1, y_2, y_3, y_4]'$ dado $\mathbf{x} = [y_5, y_6, y_7]'$.
- $\sigma_{22.567}$: representa a **variância condicional** de y_2 dado y_5, y_6 e y_7 .

O **coeficiente de correlação parcial** (populacional) entre y_i e y_j , na distribuição condicional de $\mathbf{y}$ dado $\mathbf{x}$, onde $\mathbf{x} = [y_r, y_s, \dots, y_q]'$ é definido como:

$$\rho_{ij.rs\dots q} = \frac{\sigma_{ij.rs\dots q}}{\sqrt{(\sigma_{ii.rs\dots q})(\sigma_{jj.rs\dots q})}} \quad (4.36)$$

O coeficiente de correlação parcial amostral, $r_{ij.rs\dots q}$, será discutido posteriormente, na seção 10.7.

A matriz de correlações parciais de $\mathbf{y}$ dado $\mathbf{x}$ pode ser obtida a partir de (3.28) e de (4.27) como:

$$\mathbf{P}_{\mathbf{y}.\mathbf{x}} = \mathbf{D}_{\mathbf{y}.\mathbf{x}}^{-1} \mathbf{\Sigma}_{\mathbf{y}.\mathbf{x}} \mathbf{D}_{\mathbf{y}.\mathbf{x}}^{-1} \quad (4.37)$$

em que

$$\mathbf{\Sigma}_{\mathbf{y}.\mathbf{x}} = \text{cov}(\mathbf{y}|\mathbf{x}) = \mathbf{\Sigma}_{\mathbf{y}\mathbf{y}} - \mathbf{\Sigma}_{\mathbf{y}\mathbf{x}} \mathbf{\Sigma}_{\mathbf{x}\mathbf{x}}^{-1} \mathbf{\Sigma}_{\mathbf{x}\mathbf{y}} \text{ e } \mathbf{D}_{\mathbf{y}.\mathbf{x}} = [\text{diag}(\mathbf{\Sigma}_{\mathbf{y}.\mathbf{x}})]^{1/2}$$

A menos que os vetores $\mathbf{y}$ e $\mathbf{x}$ sejam independentes ($\mathbf{\Sigma}_{\mathbf{y}.\mathbf{x}} = \mathbf{0}$) a correlação parcial $\rho_{ij.rs\dots q}$ em (4.16) é diferente da correlação usual

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} \sigma_{jj}}}, \text{ podendo até ter sinal contrário.}$$

Exemplo 4.5. Vamos comparar ρ_{12} com $\rho_{12.34}$ usando a matriz de variâncias e covariâncias, Σ , do Exemplo 4.4(b):

$$\Sigma = \begin{bmatrix} 9 & 0 & 3 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & -1 & 6 & -3 \\ 3 & 2 & -3 & 7 \end{bmatrix}$$

A correlação linear entre as variáveis y_1 e y_2 é igual a:

$$\rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11} \sigma_{22}}} = \frac{0}{\sqrt{(9)(1)}} = 0$$

ou seja, as variáveis y_1 e y_2 são não correlacionadas.

Com a partição sugerida, podemos identificar:

$$\Sigma_{yy} = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}, \Sigma_{xx} = \begin{bmatrix} 6 & -3 \\ -3 & 7 \end{bmatrix} \text{ e } \Sigma_{xy} = \begin{bmatrix} 3 & 3 \\ -1 & 2 \end{bmatrix}$$

Usando (4.27), obtemos a matriz de covariâncias condicionais de y_1 e y_2 dado y_3 e y_4 :

$$\text{cov}(\mathbf{y}|\mathbf{x}) = \mathbf{\Sigma}_{yy} - \mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy} = \frac{1}{33} \begin{bmatrix} 126 & -24 \\ -24 & 14 \end{bmatrix}$$

Daí

$$\rho_{12.34} = \frac{\sigma_{12.34}}{\sqrt{\sigma_{11.34} \sigma_{22.34}}} = \frac{-24/33}{\sqrt{(126/33)(14/33)}} = -0,571$$

Ou seja, dados (conhecendo/fixando) os valores das variáveis y_3 e y_4 , a correlação parcial entre y_1 e y_2 é negativa ($\rho_{12.34} = -0,571$) e diferente da correlação linear simples ($\rho_{12} = 0$).

EXERCÍCIOS: Problemas das páginas 101-103 do livro-texto.