

Exercício

Seja X variável aleatória com $\text{Var}(X) = 6^2$

$$\text{e } Z = \begin{bmatrix} X \\ 3X \end{bmatrix}.$$

Determinar $\text{Var}(Z)$ e mostrar que é positiva semidefinida.

Lembrando: Seja $\tilde{x}' A \tilde{x}$ uma forma quadrática. A matriz A é positiva ^{semi} definida se

$$\tilde{x}' A \tilde{x} \geq 0 \quad \text{e} \quad \tilde{x}' A \tilde{x} = 0 \quad \text{para pelo menos um} \\ \text{vetor } \tilde{x} \neq 0.$$

Resultado: Se A $n \times n$ é simétrica e positiva semidefinida de posto r então r raízes características são estritamente positivas e $n-r$ são nulas.

Resultado: Toda matriz de Variância e Covariância é positiva definida ou positiva semidefinida.

$$\text{Var}(a_1 X_1 + a_2 X_2 + \dots + a_p X_p) = \text{Var}(a'X) \geq 0$$

$$a' = [a_1 \ a_2 \ \dots \ a_p] \quad X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

$$\text{Var}(a'X) = a' \text{Var}(X) a \geq 0$$

↓
matriz de var e cov do vetor X

∴ $\text{Var}(X)$ é positiva definida ou positiva semidefinida.

No exercício

$$\text{Var}(3X) = 3^2 \text{Var}(X) = 96^2$$

$$\text{Cov}(X, 3X) = 3 \text{Cov}(X, X) = 3 \text{Var}(X) = 36^2$$

$$\therefore \text{Var}(Z) = \begin{bmatrix} 6^2 & 36^2 \\ 36^2 & 96^2 \end{bmatrix} = V \quad \text{ou}$$

$$Z = \underbrace{\begin{bmatrix} 1 \\ 3 \end{bmatrix}}_C X \quad \text{Var}(Z) = C \text{Var}(X) C' = \begin{bmatrix} 1 \\ 3 \end{bmatrix} [6^2] \begin{bmatrix} 1 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 6^2 & 36^2 \\ 36^2 & 96^2 \end{bmatrix}$$

Cálculo das raízes características

3

$$V - \lambda I = \begin{bmatrix} 6^2 - \lambda & 36^2 \\ 36^2 & 96^2 - \lambda \end{bmatrix}$$

$$\det(V - \lambda I) = (6^2 - \lambda)(96^2 - \lambda) - 96^4 = 0$$

$$96^4 - \lambda 6^2 - 9\lambda 6^2 + \lambda^2 - 96^4 = 0$$

$$\lambda^2 - 106^2 \lambda = 0$$

$$\lambda(\lambda - 106^2) = 0$$

$$\Rightarrow \lambda = 0 \quad \lambda = 106^2 > 0$$

V só pode ser positiva definida ou positiva semi-definida.

Não pode ser positiva definida porque uma das raízes características é nula.

Portanto, é positiva semidefinida.

Resultados Adicionais

4

$$\text{Var}(X) = \Sigma_X$$

a) $a'X$ e $b'X$ duas v.a. unidimensionais

$$\text{Cov}(a'X, b'X) = \sum_{i=1}^p \sum_{j=1}^p a_i b_j \sigma_{ij} = a' \Sigma_X b$$

b) Sejam $Y' = [Y_1 \ Y_2 \ \dots \ Y_p]$ $Z' = [Z_1 \ Z_2 \ \dots \ Z_r]$ vetores aleatórios. A matriz $\text{Cov}(Y, Z)$ é definida como

$$\text{Cov}(Y, Z) = \begin{bmatrix} \text{Cov}(Y_1, Z_1) & \text{Cov}(Y_1, Z_2) & \dots & \text{Cov}(Y_1, Z_r) \\ \text{Cov}(Y_2, Z_1) & \text{Cov}(Y_2, Z_2) & \dots & \text{Cov}(Y_2, Z_r) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(Y_p, Z_1) & \text{Cov}(Y_p, Z_2) & \dots & \text{Cov}(Y_p, Z_r) \end{bmatrix}_{p \times r}$$

Obs: $\text{Cov}(Y, Z)$ não é simétrica (até porque $p \neq r$).

Se $Y = Z$, $\text{Cov}(Y, Z) = \text{Var}(Y)$.

c) Se $Y = AX$ e $Z = BX$, A $n \times p$ e B $s \times p$, então

$$\text{Cov}(Y, Z) = \text{Cov}(AX, BX) = A \Sigma_X B'$$

d) Para Y e Z vetores aleatórios quaisquer,

$$\text{Cov}(AY, BZ) = A \text{Cov}(Y, Z) B'$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \underline{x_q} \\ x_{q+1} \\ \vdots \\ x_p \end{bmatrix} \quad q = \begin{bmatrix} X^{(1)} \\ \hline X^{(2)} \end{bmatrix}$$

$p-q$

Partições da Matriz de Covariância 5

Pag 73 Johnson e Wichern

$$\mu = E(X) = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \underline{\mu_q} \\ \mu_{q+1} \\ \vdots \\ \mu_p \end{bmatrix} = \begin{bmatrix} \mu^{(1)} \\ \hline \mu^{(2)} \end{bmatrix}$$

$$\Sigma = \text{Var}(X) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1q} & | & \sigma_{1q+1} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & & \sigma_{2q} & | & \sigma_{2q+1} & \dots & \sigma_{2p} \\ \vdots & & & & | & & & \\ \sigma_{q1} & \sigma_{q2} & & \sigma_{qq} & | & \sigma_{qq+1} & \dots & \sigma_{qp} \\ \hline \sigma_{q+11} & & & \sigma_{q+1q} & | & \sigma_{q+1q+1} & & \sigma_{q+1p} \\ & & & & | & & & \\ & & & & | & & & \\ \sigma_{p1} & & & \sigma_{pq} & | & \sigma_{pq+1} & & \sigma_{pp} \end{bmatrix}$$

$$= \begin{bmatrix} \overset{q \times q}{\Sigma_{11}} & \overset{q \times (p-q)}{\Sigma_{12}} \\ \underset{(p-q) \times q}{\Sigma_{21}} & \underset{(p-q) \times (p-q)}{\Sigma_{22}} \end{bmatrix}$$

$$\Sigma_{11} = \text{Var}(X^{(1)})$$

$$\Sigma_{12} = \text{Cov}(X^{(1)}, X^{(2)})$$

$$\Sigma_{21} = \text{Cov}(X^{(2)}, X^{(1)})$$

$$\Sigma_{22} = \text{Var}(X^{(2)})$$

Σ_{11} e Σ_{22} matrizes quadradas simétricas

$$\Sigma_{12} = \Sigma_{21}'$$

Exemplo

$$X' = [X_1 \ X_2 \ X_3 \ X_4 \ X_5] \quad \mu'_X = [2 \ 4 \ -1 \ 3 \ 0]$$

$$\Sigma_X = \begin{bmatrix} 4 & -1 & 1/2 & -1/2 & 0 \\ -1 & 3 & 1 & -1 & 0 \\ 1/2 & 1 & 6 & 1 & -1 \\ -3/2 & -1 & 1 & 4 & 0 \\ 0 & 0 & -1 & 0 & 2 \end{bmatrix}$$

A ordem é importante!

Particionando X como $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{bmatrix} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix}$ determinar

$$E(X^{(1)}) \quad E(X^{(2)}) \quad \text{Var}(X^{(1)}) \quad \text{Var}(X^{(2)}) \\ \text{Cov}(X^{(1)}, X^{(2)}).$$

$$E(X^{(1)}) = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad E(X^{(2)}) = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} \quad \text{Var}(X^{(1)}) = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix}$$

$$\text{Var}(X^{(2)}) = \begin{bmatrix} 6 & 1 & -1 \\ 1 & 4 & 0 \\ -1 & 0 & 2 \end{bmatrix} \quad \text{Cov}(X^{(1)}, X^{(2)}) = \begin{bmatrix} 1/2 & -1/2 & 0 \\ 1 & -1 & 0 \end{bmatrix}.$$

Se $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$ Calcular $E(AX^{(1)})$

$\text{Var}(AX^{(1)}) \quad E(BX^{(2)}) \quad \text{Var}(BX^{(2)}) \quad \text{Cov}(AX^{(1)}, BX^{(2)})$

$$E(AX^{(1)}) = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \end{bmatrix} \quad AX^{(1)} = \begin{bmatrix} X_1 - X_2 \\ X_1 + X_2 \end{bmatrix}$$

$$\text{Var}(AX^{(1)}) = A \text{Var}(X^{(1)}) A' = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 1 \\ 1 & 5 \end{bmatrix}$$

$$\text{Cov}(AX^{(1)}, BX^{(2)}) = A \text{Cov}(X^{(1)}, X^{(2)}) B' = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$BX^{(2)} = \begin{bmatrix} X_3 + X_4 + X_5 \\ X_3 + X_4 - 2X_5 \end{bmatrix}$

2×2 2×3 3×2

Amostras Aleatórias e Matriz de Covariância Amostral

$X' = [X_1 \ X_2 \ \dots \ X_p]$ vetor aleatório p dimensional com densidade $f(\underline{x}) = f(x_1, x_2, \dots, x_p)$.

Sejam $\underline{X}'_i = [X_{i1} \ X_{i2} \ \dots \ X_{ip}]$, $i = 1, 2, \dots, n$, vetores aleatórios com a mesma distribuição de probabilidades de X' .

Se $\underline{X}'_1, \underline{X}'_2, \dots, \underline{X}'_n$ são independentes então $\underline{X}'_1, \underline{X}'_2, \dots, \underline{X}'_n$ é uma amostra aleatória de $f(\underline{x})$.

$\underline{X}'_1, \underline{X}'_2, \dots, \underline{X}'_n$ são vetores aleatórios independentes $\Leftrightarrow$

$$f_c(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) = f(\underline{x}_1) f(\underline{x}_2) \dots f(\underline{x}_n)$$

↓
conjunta

onde $f(\underline{x}) = f(x_{i1}, x_{i2}, \dots, x_{ip})$ $i = 1, 2, \dots, n$.

Esses elementos amostrais podem ser dispostos na matriz aleatória

$$X_{n \times p} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & & X_{2p} \\ \vdots & & & \\ X_{n1} & X_{n2} & & X_{np} \end{bmatrix} = \begin{bmatrix} \underline{X}'_1 \\ \underline{X}'_2 \\ \vdots \\ \underline{X}'_n \end{bmatrix} \quad \leadsto \text{um elemento amostral por linha.}$$

9

Obs: Medidas das p variáveis no mesmo elemento amostral usualmente são correlacionadas. Medidas em diferentes unidades amostrais são, em geral, independentes

Observada a amostra de tamanho n , ficam definidos

Vetor de Médias amostrais

$$\bar{X} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix} \quad \bar{X}_j = \frac{\sum_{i=1}^n X_{ij}}{n} \quad j = 1, 2, \dots, p$$

Obs: $\bar{X} = \frac{1}{n} \sum_{i=1}^n \tilde{x}_i$

Matriz de Variâncias e Covariâncias Amostral

$$S = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1p} \\ s_{21} & s_{22} & & s_{2p} \\ \vdots & & & \\ s_{p1} & s_{p2} & & s_{pp} \end{bmatrix}_{p \times p}$$

$$s_{jk} = \frac{1}{n} \sum_{i=1}^n (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k) \quad j, k = 1, 2, \dots, p$$

Obs: s_{jk} - covariância amostral entre as variáveis X_j e X_k .

Matriz de Correlações Amostral

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & & r_{2p} \\ & & \ddots & \vdots \\ r_{p1} & r_{p2} & & 1 \end{bmatrix}$$

$$r_{jk} = \frac{s_{jk}}{\sqrt{s_{jj}} \sqrt{s_{kk}}} \quad j, k = 1, 2, \dots, p$$

Obs: r_{jk} - coeficiente de correlação amostral entre X_j e X_k

Formas Matriciais Alternativas para o vetor de médias e matriz de covariância amostral.

11

$$\bar{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix}_{n \times 1} = \frac{1}{n} \sum_{i=1}^n \underset{\sim}{X}_i \quad \underset{\sim}{X}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{bmatrix}$$

$$S = \frac{1}{n} \sum_{i=1}^n (\underset{\sim}{X}_i - \bar{X})(\underset{\sim}{X}_i - \bar{X})' = \frac{1}{n} \left[\sum_{i=1}^n \underset{\sim}{X}_i \underset{\sim}{X}_i' - n \bar{X} \bar{X}' \right] \quad \text{pag 13}$$

alternativamente, lembrando que $X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$

e depois

$$\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}_{1 \times n} X = \begin{bmatrix} \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} & \dots & \sum_{i=1}^n x_{ip} \end{bmatrix} = n \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \dots & \bar{x}_p \end{bmatrix}$$

$$\boxed{\bar{X} = \frac{1}{n} X' \underset{\sim}{1}}$$

$$\underset{\sim}{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n \times 1}$$

$$= n \bar{X}'$$

$$\bar{X}' = \frac{\underset{\sim}{1}' X}{n} \Rightarrow \bar{X} = \frac{X' \underset{\sim}{1}}{n}$$

$$nS = \sum_{i=1}^n (\underset{\sim}{X}_i - \bar{X})(\underset{\sim}{X}_i - \bar{X})' = \sum_{i=1}^n \underset{\sim}{X}_i \underset{\sim}{X}_i' - n \bar{X} \bar{X}' =$$

$$= \sum_{i=1}^n \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{bmatrix} [x_{i1} \ x_{i2} \ \dots \ x_{ip}] - \frac{1}{n} X' \underset{\sim}{1} \underset{\sim}{1}' X =$$

$$= \sum_{i=1}^n \begin{bmatrix} X_{i1}^2 & X_{i1}X_{i2} & \dots & X_{i1}X_{ip} \\ X_{i2}X_{i1} & X_{i2}^2 & & X_{i2}X_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ X_{ip}X_{i1} & X_{ip}X_{i2} & \dots & X_{ip}^2 \end{bmatrix} - \frac{X' \mathbf{1} \mathbf{1}' X}{n}$$

Temos que

$$X'X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{i=1}^n X_{i1}^2 & \sum_{i=1}^n X_{i1}X_{i2} & \dots & \sum_{i=1}^n X_{i1}X_{ip} \\ \sum_{i=1}^n X_{i2}X_{i1} & \sum_{i=1}^n X_{i2}^2 & \dots & \sum_{i=1}^n X_{i2}X_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n X_{ip}X_{i1} & \sum_{i=1}^n X_{ip}X_{i2} & \dots & \sum_{i=1}^n X_{ip}^2 \end{bmatrix}$$

Portanto

$$MS = X'X - \frac{X' \mathbf{1} \mathbf{1}' X}{n} = X' \left[I - \frac{\mathbf{1} \mathbf{1}'}{n} \right] X$$

MS é chamada matriz de produtos cruzados.

$$S = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' = \frac{1}{n} \left[\sum_{i=1}^n X_i X_i' - n \bar{X} \bar{X}' \right]$$

Demonstrações:

$$\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' = \sum_{i=1}^n (X_i - \bar{X}) X_i' - \sum_{i=1}^n (X_i - \bar{X}) \bar{X}'$$

$$= \sum_{i=1}^n X_i X_i' - \bar{X} \underbrace{\sum_{i=1}^n X_i'}_{n \bar{X}'} - \left[\sum_{i=1}^n (X_i - \bar{X}) \right] \bar{X}'$$

*

$$* \sum_{i=1}^n (X_i - \bar{X}) = \sum_{i=1}^n \left\{ \begin{bmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{ip} \end{bmatrix} - \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix} \right\} = \sum_{i=1}^n \begin{bmatrix} X_{i1} - \bar{X}_1 \\ \vdots \\ X_{ip} - \bar{X}_p \end{bmatrix} = \mathbf{O}_{p \times 1}$$

$$= \sum_{i=1}^n X_i X_i' - n \bar{X} \bar{X}'$$