SME0822 Análise Multivariada e Aprendizado Não-Supervisionado

Aula 3: **Vetores aleatórios**

Prof. Cibele Russo

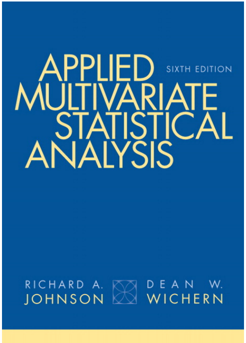
cibele@icmc.usp.br

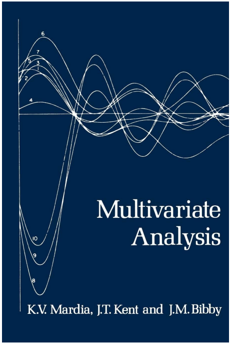
http://www.icmc.usp.br/~cibele

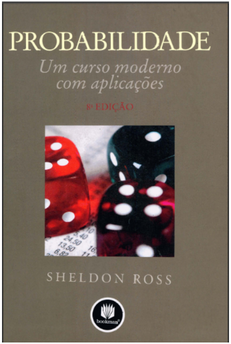
Johnson, R. A., & Wichern, D. W. (2007) Applied Multivariate Statistical Analysis. Prentice Hall.

Mardia, K. V., Kent, J. T. & Bibby J. M. (1979) Multivariate analysis. Academic Press Inc.

Referências indicadas







Vetor aleatório

Def: $\underset{\sim}{X}$ é um **vetor aleatório** de dimensão $p \times 1$

$$\underset{\sim}{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix},$$

em que cada X_i é uma variável aleatória com **média** μ_i e **variância** $\sigma_i^2 = \sigma_{ii}$, para $i = 1, \ldots, p$, em que

$$\mu_i = E(X_i) \quad \text{(média populacional)}$$
$$\sigma_i^2 = Var(X_i) = E[(X_i - \mu_i)^2] \quad \text{(variância populacional)}$$

Vetor aleatório

Mais especificamente,

$$\mu_i = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x_i f_i(x_i) dx_i, & \text{se } X_i \text{ é v.a. contínua com f.d.p. } f_i(x_i) \\ \displaystyle\sum_{\text{todo } x_i} x_i p_i(x_i), & \text{se } X_i \text{ é v.a. discreta com f.p. } p_i(x_i) \end{cases}$$

$$\sigma_i^2 = \begin{cases} \displaystyle\int_{-\infty}^{\infty} (x_i - \mu_i)^2 f_i(x_i) dx_i, & \text{se } X_i \text{ é v.a. contínua com f.d.p. } f_i(x_i) \\ \displaystyle\sum_{\text{todo } x_i} (x_i - \mu_i)^2 p_i(x_i), & \text{se } X_i \text{ é v.a. discreta com f.p. } p_i(x_i) \end{cases}$$

Covariância e correlação (populacionais)

A **covariância** e a **correlação** (populacionais) entre duas variáveis X_i e X_k, com $i, k = 1, \ldots, p$ e $i \neq k$ são dadas, respectivamente, por

$$\text{Cov}(X_i, X_k) = \sigma_{ik} = E[(X_i - \mu_i)(X_k - \mu_k)] \text{ e}$$

$$\text{Cor}(X_i, X_k) = \rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_i^2}\sqrt{\sigma_k^2}}$$

Covariância e correlação (populacionais)

Mais especificamente,

$$\sigma_{ik} = \begin{cases} \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_i-\mu_i)(x_k-\mu_k)f_{ik}(x_i,x_k)dx_i dx_k, & \text{se } X_i, X_k \text{ são v.a.'s contínuas com f.d.p. conjunta } f_{ik}(x_i,x_k) \\ \displaystyle\sum_{\text{todo } x_i}\sum_{\text{todo } x_k}(x_i-\mu_i)(x_k-\mu_k)p_{ik}(x_i,x_k), & \text{se } X_i, X_k \text{ são v.a.'s discretas com f.p. conjunta } p_{ik}(x_i,x_k) \end{cases}$$

Distribuição conjunta de $X_1, \ldots, X_p$

A distribuição conjunta das p variáveis aleatórias (v.a.'s) $X_1, \ldots, X_p$ ou do vetor $\underset{\sim}{X} = (X_1, \ldots, X_p)^\top$ é descrito pela **função densidade de probabilidade conjunta** (ou função de probabilidade conjunta no caso discreto) de $X_1, \ldots, X_p$, dada por

$$f(x_1, \ldots, x_p) = f(\underset{\sim}{x}),$$

com

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(u_1, \ldots, u_p) du_1 \ldots du_p = 1$$

(ou somatória no caso discreto).

Função distribuição acumulada

Se $\underset{\sim}{X} = (X_1, \ldots, X_p)^\top$ é um vetor aleatório e tem f.d.p. $f(\underset{\sim}{x}) = f(x_1, \ldots, x_p)$, então a **função distribuição acumulada** de $\underset{\sim}{x}$ é dada por

$$F(\underset{\sim}{x}) = P(\underset{\sim}{X} \leq \underset{\sim}{x}) = P(X_1 \leq x_1, \ldots, X_p \leq x_p)$$

que pode ser calculada no nosso contexto como

$$F(\underset{\sim}{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(u_1, \ldots, u_p) du_1 \ldots du_p$$

(no caso discreto substituir por somatórias adequadas)

Função distribuição acumulada

Se $\underset{\sim}{X} = (X_1, \ldots, X_p)^\top$ é um vetor aleatório e tem f.d.p. $f(\underset{\sim}{x}) = f(x_1, \ldots, x_p)$, então a **função distribuição acumulada** de $\underset{\sim}{x}$ é dada por

$$F(\underset{\sim}{x}) = P(\underset{\sim}{X} \leq \underset{\sim}{x}) = P(X_1 \leq x_1, \ldots, X_p \leq x_p)$$

que pode ser calculada no nosso contexto como

$$F(\underset{\sim}{x}) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_p} f(u_1, \ldots, u_p) du_1 \ldots du_p$$

(no caso discreto substituir por somatórias adequadas)

F.d.a. e f.d.p. marginais de X_i

Sem perda de generalidade, suponha $i = 1$.

A função distribuição acumulada marginal de X_1 é dada por

$$F_1(x_1) = P(X_1 \leq x_1) = \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(u_1, \dots, u_p) du_1 du_2 \dots du_p$$

A função densidade marginal de X_1 é dada por

$$f_1(x_1) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(u_1, \dots, u_p) du_2 \dots du_p$$

Note que na última equação temos $p - 1$ integrais somente.

Da mesma forma poderíamos obter outras f.d.a.'s e f.d.p.'s, e analogamente para o caso discreto.

F.d.a. e f.d.p. marginais de X_i

Sem perda de generalidade, suponha $i = 1$.

A função distribuição acumulada marginal de X_1 é dada por

$$F_1(x_1) = P(X_1 \leq x_1) = \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(u_1, \dots, u_p) du_1 du_2 \dots du_p$$

A função densidade marginal de X_1 é dada por

$$f_1(x_1) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(u_1, \dots, u_p) du_2 \dots du_p$$

Note que na última equação temos $p - 1$ integrais somente.

Da mesma forma poderíamos obter outras f.d.a.'s e f.d.p.'s, e analogamente para o caso discreto.

Independência estatística

Dada a distribuição conjunta de X_i e X_k, se

$$P(X_i \leq x_i, X_k \leq x_k) = P(X_i \leq x_i)P(X_k \leq x_k)$$

para todos os pares possíveis (x_i, x_k), então dizemos que X_i e X_k são estatisticamente independentes.

Quando X_i e X_k são variáveis aleatórias contínuas com f.d.p. $f_{ik}(x_i, x_k)$, a condição para a independência é

$$f_{ik}(x_i, x_k) = f_i(x_i)f_k(x_k)$$

para todo par (x_i, x_k).

Independência estatística

Dada a distribuição conjunta de X_i e X_k, se

$$P(X_i \leq x_i, X_k \leq x_k) = P(X_i \leq x_i)P(X_k \leq x_k)$$

para todos os pares possíveis (x_i, x_k), então dizemos que X_i e X_k são estatisticamente independentes.

Quando X_i e X_k são variáveis aleatórias contínuas com f.d.p. $f_{ik}(x_i, x_k)$, a condição para a independência é

$$f_{ik}(x_i, x_k) = f_i(x_i)f_k(x_k)$$

para todo par (x_i, x_k).

Independência estatística

Se tivermos p variáveis aleatórias contínuas $X_1, \ldots, X_p$, elas são estatisticamente independentes se a sua f.d.p. conjunta pode ser fatorada como

$$f(\underset{\sim}{x}) = f_1(x_1) \ldots f_p(x_p), \forall (x_1, \ldots, x_p).$$

A independência tem uma implicação importante:
Se X_i e X_k são independentes, então

$$Cov(X_i, X_k) = 0$$

A volta não vale necessariamente.

Independência estatística

Se tivermos p variáveis aleatórias contínuas $X_1, \ldots, X_p$, elas são estatisticamente independentes se a sua f.d.p. conjunta pode ser fatorada como

$$f(\underset{\sim}{x}) = f_1(x_1) \ldots f_p(x_p), \forall (x_1, \ldots, x_p).$$

A independência tem uma implicação importante:
Se X_i e X_k são independentes, então

$$Cov(X_i, X_k) = 0$$

A volta não vale necessariamente.

Momentos populacionais

Se $\underset{\sim}{X}_{p\times 1}$ é um vetor aleatório com f.d.p. $f(\underset{\sim}{x})$, o **valor esperado** de uma função escalar $g(\underset{\sim}{x})$ é definido por

$$E(g(\underset{\sim}{x})) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(\underset{\sim}{u}) f(\underset{\sim}{u}) du_1 \dots du_p.$$

Mais geralmente, o valor esperado de uma função vetorial de $\underset{\sim}{x}$, $\underset{\sim}{g}(\underset{\sim}{x})$, é dado por

$$E(\underset{\sim}{g}(\underset{\sim}{x})) = (E(g_i(\underset{\sim}{x}))),$$

ou seja, o valor esperado é calculado elemento a elemento da função vetorial $\underset{\sim}{g}(\underset{\sim}{x})$.

Momentos populacionais

Se $\utilde{X}_{p\times 1}$ é um vetor aleatório com f.d.p. $f(\utilde{x})$, o **valor esperado** de uma função escalar $g(\utilde{x})$ é definido por

$$E(g(\utilde{x})) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(\utilde{u}) f(\utilde{u}) du_1 \dots du_p.$$

Mais geralmente, o valor esperado de uma função vetorial de $\utilde{x}$, $\utilde{g}(\utilde{x})$, é dado por

$$E(\utilde{g}(\utilde{x})) = (E(g_i(\utilde{x}))),$$

ou seja, o valor esperado é calculado elemento a elemento da função vetorial $\utilde{g}(\utilde{x})$.

Vetor de médias e matriz de variâncias e covariâncias (populacionais)

O **vetor de médias** (populacionais) de $\underset{\sim}{X}$ é $\underset{\sim}{\mu}_{p\times 1}$ e a **matriz de variâncias e covariâncias** (populacionais) de $\underset{\sim}{X}$ é $\Sigma_{p\times p}$, em que

$$\underset{\sim}{\mu} = E(\underset{\sim}{X}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \begin{pmatrix} \mathrm{E}(X_1) \\ \mathrm{E}(X_2) \\ \vdots \\ \mathrm{E}(X_p) \end{pmatrix} \text{ e } \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ & \sigma_2^2 & \dots & \sigma_{2p} \\ & & \ddots & \vdots \\ \text{sim.} & & & \sigma_p^2 \end{pmatrix}$$

Obs: Note que Σ é simétrica pois $\mathrm{Cov}(X_i, X_k) = \mathrm{Cov}(X_k, X_i)$, para $i, j = 1, \dots, p$.

Matriz de correlações (populacionais)

A **matriz de correlações** (populacionais) de $\underset{\sim}{X}$ é $P = P_{p\times p}$, em que

$$P = Cor(\underset{\sim}{X}) = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ & 1 & \dots & \rho_{2p} \\ & & \ddots & \vdots \\ \text{sim.} & & & 1 \end{pmatrix}$$

Obs 1: Pelo mesmo motivo que Σ é simétrica, P também é.

Obs 2: $-1 \leq \rho_{ik} \leq 1$ para $i \neq k$ e $i, k = 1 \dots, p$.

Matriz de variâncias e covariâncias (populacionais)

Matricialmente,

$$\text{Var}(\underset{\sim}{X}) = \Sigma = \text{E}[(\underset{\sim}{X} - \underset{\sim}{\mu})(\underset{\sim}{X} - \underset{\sim}{\mu})^\top].$$

Além disso,

$$\text{Cor}(\underset{\sim}{X}) = P = (V^{1/2})^{-1}\Sigma(V^{1/2})^{-1} = V^{-1/2}\Sigma V^{-1/2},$$

em que

$$V^{1/2} = \begin{pmatrix} \sqrt{\sigma_{11}} & 0 & \ldots & 0 \\ 0 & \sqrt{\sigma_{22}} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \sqrt{\sigma_{pp}} \end{pmatrix}$$

Variância e covariância em forma matricial

Sejam $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ vetores aleatórios p-dimensionais com $\mathrm{E}(\underset{\sim}{X}) = \underset{\sim}{\mu}_X$ e $\mathrm{E}(\underset{\sim}{Y}) = \underset{\sim}{\mu}_Y$.

Define-se

1. $\mathrm{Var}(\underset{\sim}{X}) = \mathrm{E}[(\underset{\sim}{X} - \underset{\sim}{\mu}_X)(\underset{\sim}{X} - \underset{\sim}{\mu}_X)^\top]$,
2. $\mathrm{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \mathrm{E}[(\underset{\sim}{X} - \underset{\sim}{\mu}_X)(\underset{\sim}{Y} - \underset{\sim}{\mu}_Y)^\top]$

Variância e covariância em forma matricial

Propriedades.

Sejam $\underset{\sim}{X}$, $\underset{\sim}{Y}$, $\underset{\sim}{X}_1$ e $\underset{\sim}{X}_2$ vetores aleatórios p-dimensionais

1. $\text{Var}(\underset{\sim}{X}) = \text{E}[\underset{\sim}{X}\underset{\sim}{X}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_X^\top$,
2. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \text{E}[\underset{\sim}{X}\underset{\sim}{Y}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_Y^\top$
3. $\text{Var}(A\underset{\sim}{X} + \underset{\sim}{b}) = A\,\text{Var}(\underset{\sim}{X})A^\top$
4. $\text{Cov}(A\underset{\sim}{X} + \underset{\sim}{b}, C\underset{\sim}{Y} + \underset{\sim}{d}) = A\,\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y})C^\top$
5. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{X}) = \text{Var}(\underset{\sim}{X})$
6. $\text{Cov}(\underset{\sim}{X}_1 + \underset{\sim}{X}_2, \underset{\sim}{Y}) = \text{Cov}(\underset{\sim}{X}_1, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}_2, \underset{\sim}{Y})$
7. $\text{Var}(\underset{\sim}{X} + \underset{\sim}{Y}) = \text{Var}(\underset{\sim}{X}) + \text{Var}(\underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{Y}, \underset{\sim}{X})$
8. Se $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ são independentes, então $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = 0$. A volta nem sempre vale.

Variância e covariância em forma matricial

Propriedades.

Sejam $\underset{\sim}{X}$, $\underset{\sim}{Y}$, $\underset{\sim}{X}_1$ e $\underset{\sim}{X}_2$ vetores aleatórios p-dimensionais

1. $\text{Var}(\underset{\sim}{X}) = \text{E}[\underset{\sim}{X}\underset{\sim}{X}^\top] - \underset{\sim}{\mu}_X\underset{\sim}{\mu}_X^\top$,
2. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \text{E}[\underset{\sim}{X}\underset{\sim}{Y}^\top] - \underset{\sim}{\mu}_X\underset{\sim}{\mu}_Y^\top$
3. $\text{Var}(A\underset{\sim}{X} + \underset{\sim}{b}) = A\,\text{Var}(\underset{\sim}{X})A^\top$
4. $\text{Cov}(A\underset{\sim}{X} + \underset{\sim}{b}, C\underset{\sim}{Y} + \underset{\sim}{d}) = A\,\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y})C^\top$
5. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{X}) = \text{Var}(\underset{\sim}{X})$
6. $\text{Cov}(\underset{\sim}{X}_1 + \underset{\sim}{X}_2, \underset{\sim}{Y}) = \text{Cov}(\underset{\sim}{X}_1, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}_2, \underset{\sim}{Y})$
7. $\text{Var}(\underset{\sim}{X} + \underset{\sim}{Y}) = \text{Var}(\underset{\sim}{X}) + \text{Var}(\underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{Y}, \underset{\sim}{X})$
8. Se $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ são independentes, então $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = 0$. A volta nem sempre vale.

Variância e covariância em forma matricial

Propriedades.

Sejam $\underset{\sim}{X}$, $\underset{\sim}{Y}$, $\underset{\sim}{X}_1$ e $\underset{\sim}{X}_2$ vetores aleatórios p-dimensionais

1. $\text{Var}(\underset{\sim}{X}) = \text{E}[\underset{\sim}{X}\underset{\sim}{X}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_X^\top$,
2. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \text{E}[\underset{\sim}{X}\underset{\sim}{Y}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_Y^\top$
3. $\text{Var}(A\underset{\sim}{X} + \underset{\sim}{b}) = A\ \text{Var}(\underset{\sim}{X})A^\top$
4. $\text{Cov}(A\underset{\sim}{X} + \underset{\sim}{b}, C\underset{\sim}{Y} + \underset{\sim}{d}) = A\ \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y})C^\top$
5. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{X}) = \text{Var}(\underset{\sim}{X})$
6. $\text{Cov}(\underset{\sim}{X}_1 + \underset{\sim}{X}_2, \underset{\sim}{Y}) = \text{Cov}(\underset{\sim}{X}_1, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}_2, \underset{\sim}{Y})$
7. $\text{Var}(\underset{\sim}{X} + \underset{\sim}{Y}) = \text{Var}(\underset{\sim}{X}) + \text{Var}(\underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{Y}, \underset{\sim}{X})$
8. Se $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ são independentes, então $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = 0$. A volta nem sempre vale.

Variância e covariância em forma matricial

Propriedades.

Sejam $\underset{\sim}{X}$, $\underset{\sim}{Y}$, $\underset{\sim}{X}_1$ e $\underset{\sim}{X}_2$ vetores aleatórios p-dimensionais

1. $\text{Var}(\underset{\sim}{X}) = \text{E}[\underset{\sim}{X}\underset{\sim}{X}^\top] - \underset{\sim}{\mu}_X\underset{\sim}{\mu}_X^\top$,
2. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \text{E}[\underset{\sim}{X}\underset{\sim}{Y}^\top] - \underset{\sim}{\mu}_X\underset{\sim}{\mu}_Y^\top$
3. $\text{Var}(A\underset{\sim}{X} + \underset{\sim}{b}) = A\ \text{Var}(\underset{\sim}{X})A^\top$
4. $\text{Cov}(A\underset{\sim}{X} + \underset{\sim}{b}, C\underset{\sim}{Y} + \underset{\sim}{d}) = A\ \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y})C^\top$
5. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{X}) = \text{Var}(\underset{\sim}{X})$
6. $\text{Cov}(\underset{\sim}{X}_1 + \underset{\sim}{X}_2, \underset{\sim}{Y}) = \text{Cov}(\underset{\sim}{X}_1, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}_2, \underset{\sim}{Y})$
7. $\text{Var}(\underset{\sim}{X} + \underset{\sim}{Y}) = \text{Var}(\underset{\sim}{X}) + \text{Var}(\underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{Y}, \underset{\sim}{X})$
8. Se $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ são independentes, então $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = 0$. A volta nem sempre vale.

Variância e covariância em forma matricial

Propriedades.

Sejam $\underset{\sim}{X}$, $\underset{\sim}{Y}$, $\underset{\sim}{X}_1$ e $\underset{\sim}{X}_2$ vetores aleatórios p-dimensionais

1. $\text{Var}(\underset{\sim}{X}) = \text{E}[\underset{\sim}{X}\underset{\sim}{X}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_X^\top$,
2. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \text{E}[\underset{\sim}{X}\underset{\sim}{Y}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_Y^\top$
3. $\text{Var}(A\underset{\sim}{X} + \underset{\sim}{b}) = A\,\text{Var}(\underset{\sim}{X})A^\top$
4. $\text{Cov}(A\underset{\sim}{X} + \underset{\sim}{b}, C\underset{\sim}{Y} + \underset{\sim}{d}) = A\,\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y})C^\top$
5. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{X}) = \text{Var}(\underset{\sim}{X})$
6. $\text{Cov}(\underset{\sim}{X}_1 + \underset{\sim}{X}_2, \underset{\sim}{Y}) = \text{Cov}(\underset{\sim}{X}_1, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}_2, \underset{\sim}{Y})$
7. $\text{Var}(\underset{\sim}{X} + \underset{\sim}{Y}) = \text{Var}(\underset{\sim}{X}) + \text{Var}(\underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{Y}, \underset{\sim}{X})$
8. Se $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ são independentes, então $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = 0$. A volta nem sempre vale.

Variância e covariância em forma matricial

Propriedades.

Sejam $\underset{\sim}{X}$, $\underset{\sim}{Y}$, $\underset{\sim}{X}_1$ e $\underset{\sim}{X}_2$ vetores aleatórios p-dimensionais

1. $\text{Var}(\underset{\sim}{X}) = \text{E}[\underset{\sim}{X}\underset{\sim}{X}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_X^\top$,
2. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \text{E}[\underset{\sim}{X}\underset{\sim}{Y}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_Y^\top$
3. $\text{Var}(A\underset{\sim}{X} + \underset{\sim}{b}) = A\,\text{Var}(\underset{\sim}{X})A^\top$
4. $\text{Cov}(A\underset{\sim}{X} + \underset{\sim}{b}, C\underset{\sim}{Y} + \underset{\sim}{d}) = A\,\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y})C^\top$
5. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{X}) = \text{Var}(\underset{\sim}{X})$
6. $\text{Cov}(\underset{\sim}{X}_1 + \underset{\sim}{X}_2, \underset{\sim}{Y}) = \text{Cov}(\underset{\sim}{X}_1, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}_2, \underset{\sim}{Y})$
7. $\text{Var}(\underset{\sim}{X} + \underset{\sim}{Y}) = \text{Var}(\underset{\sim}{X}) + \text{Var}(\underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{Y}, \underset{\sim}{X})$
8. Se $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ são independentes, então $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = 0$. A volta nem sempre vale.

Variância e covariância em forma matricial

Propriedades.

Sejam $\underset{\sim}{X}$, $\underset{\sim}{Y}$, $\underset{\sim}{X}_1$ e $\underset{\sim}{X}_2$ vetores aleatórios p-dimensionais

1. $\text{Var}(\underset{\sim}{X}) = \text{E}[\underset{\sim}{X}\underset{\sim}{X}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_X^\top$,
2. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \text{E}[\underset{\sim}{X}\underset{\sim}{Y}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_Y^\top$
3. $\text{Var}(A\underset{\sim}{X} + \underset{\sim}{b}) = A\ \text{Var}(\underset{\sim}{X})A^\top$
4. $\text{Cov}(A\underset{\sim}{X} + \underset{\sim}{b}, C\underset{\sim}{Y} + \underset{\sim}{d}) = A\ \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y})C^\top$
5. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{X}) = \text{Var}(\underset{\sim}{X})$
6. $\text{Cov}(\underset{\sim}{X}_1 + \underset{\sim}{X}_2, \underset{\sim}{Y}) = \text{Cov}(\underset{\sim}{X}_1, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}_2, \underset{\sim}{Y})$
7. $\text{Var}(\underset{\sim}{X} + \underset{\sim}{Y}) = \text{Var}(\underset{\sim}{X}) + \text{Var}(\underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{Y}, \underset{\sim}{X})$
8. Se $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ são independentes, então $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = 0$. A volta nem sempre vale.

Variância e covariância em forma matricial

Propriedades.

Sejam $\underset{\sim}{X}$, $\underset{\sim}{Y}$, $\underset{\sim}{X}_1$ e $\underset{\sim}{X}_2$ vetores aleatórios p-dimensionais

1. $\text{Var}(\underset{\sim}{X}) = \text{E}[\underset{\sim}{X}\underset{\sim}{X}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_X^\top$,
2. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = \text{E}[\underset{\sim}{X}\underset{\sim}{Y}^\top] - \underset{\sim}{\mu}_X \underset{\sim}{\mu}_Y^\top$
3. $\text{Var}(A\underset{\sim}{X} + \underset{\sim}{b}) = A \text{ Var}(\underset{\sim}{X})A^\top$
4. $\text{Cov}(A\underset{\sim}{X} + \underset{\sim}{b}, C\underset{\sim}{Y} + \underset{\sim}{d}) = A \text{ Cov}(\underset{\sim}{X}, \underset{\sim}{Y})C^\top$
5. $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{X}) = \text{Var}(\underset{\sim}{X})$
6. $\text{Cov}(\underset{\sim}{X}_1 + \underset{\sim}{X}_2, \underset{\sim}{Y}) = \text{Cov}(\underset{\sim}{X}_1, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}_2, \underset{\sim}{Y})$
7. $\text{Var}(\underset{\sim}{X} + \underset{\sim}{Y}) = \text{Var}(\underset{\sim}{X}) + \text{Var}(\underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) + \text{Cov}(\underset{\sim}{Y}, \underset{\sim}{X})$
8. Se $\underset{\sim}{X}$ e $\underset{\sim}{Y}$ são independentes, então $\text{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = 0$. A volta nem sempre vale.

Amostra aleatória

Suponha que $\underset{\sim}{X}$ tem uma **distribuição de probabilidade multivariada,** e que vamos observar uma **amostra aleatória** dessa distribuição, ou seja, *n* observações independentes e que vem todas da mesma distribuição:

$$\textbf{amostra aleatória}: \underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n,$$

em que $\underset{\sim}{X}_i^\top = (X_{i1}, X_{i2}, \ldots\ X_{ip})$.

Matriz aleatória

Suponha que $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ ainda não foram observados, mas podemos organiza-las em uma **matriz aleatória**

$$X = \begin{pmatrix} X_{11} & X_{12} & \ldots & X_{1p} \\ X_{21} & X_{22} & \ldots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \ldots & X_{np} \end{pmatrix}_{n \times p} = \begin{pmatrix} \underset{\sim}{X}_1^\top \\ \underset{\sim}{X}_2^\top \\ \vdots \\ \underset{\sim}{X}_n^\top \end{pmatrix}$$

Estimadores para $\underset{\sim}{\mu}$ e Σ

Seja $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ a.a. de uma distribuição com $E(\underset{\sim}{X}) = \underset{\sim}{\mu}$ e $\text{Var}(\underset{\sim}{X}) = \Sigma$. Então

- $\bar{\underset{\sim}{X}}$ é um estimador não-viesado para $\underset{\sim}{\mu}$, com $Var(\bar{\underset{\sim}{X}}) = \dfrac{\Sigma}{n}$ e
- $\dfrac{n}{n-1}S$ é um estimador não-viesado para Σ, ou seja $E\left(\dfrac{n}{n-1}S\right) = \Sigma$.

Obs: Vimos na Aula 1 que os elementos de S são dados por

$$s_{ik} = \frac{1}{n}\sum_{j=1}^{n}(x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)$$

Estimadores para $\underset{\sim}{\mu}$ e Σ

Seja $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ a.a. de uma distribuição com $E(\underset{\sim}{X}) = \underset{\sim}{\mu}$ e $\text{Var}(\underset{\sim}{X}) = \Sigma$.

Então

- $\bar{\underset{\sim}{X}}$ é um estimador não-viesado para $\underset{\sim}{\mu}$, com $Var(\bar{\underset{\sim}{X}}) = \dfrac{\Sigma}{n}$ e
- $\dfrac{n}{n-1}S$ é um estimador não-viesado para Σ, ou seja $E\left(\dfrac{n}{n-1}S\right) = \Sigma$.

Obs: Vimos na Aula 1 que os elementos de S são dados por

$$s_{ik} = \frac{1}{n}\sum_{j=1}^{n}(x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)$$