

ME731 - Métodos em Análise Multivariada

– Análise de Correlação Canônica I –

Prof. Carlos Trucíos
ctrucios@unicamp.br
[ctruciosm.github.io](https://github.com/ctruciosm)

Instituto de Matemática, Estatística e Computação Científica,
Universidade Estadual de Campinas

Aula 19



Agenda I

- 1 Motivação
- 2 Introdução
- 3 Análise de Correlação Canônica
- 4 Propriedades
- 5 Implementação

Motivação

Motivação

- 1 Suponha que estamos interessados em estudar a relação entre os resultados de um teste de capacidade intelectual (o resultado do teste são vários scores para diferentes *atributos*) e as medidas físicas de um grupo de indivíduos.
- 2 Uma empresa de venda de autos está interessada em estudar a relação entre *Preço* e *Valor* do auto com as variáveis *Economia*, *Serviço*, *Desing*, *Carro esportivo*, *Segurança* e *Facilidade para dirigir*, respectivamente.

Motivação

- 1 Suponha que estamos interessados em estudar a relação entre os resultados de um teste de capacidade intelectual (o resultado do teste são vários scores para diferentes *atributos*) e as medidas físicas de um grupo de indivíduos.
- 2 Uma empresa de venda de autos está interessada em estudar a relação entre *Preço* e *Valor* do auto com as variáveis *Economia*, *Serviço*, *Desing*, *Carro esportivo*, *Segurança* e *Facilidade para dirigir*, respectivamente.

Como faria as análises?

Introdução

Introdução

- Em análise de regressão múltipla temos um conjunto de covariáveis \mathbf{x} e uma única variável resposta y . A solução implica em encontrar \mathbf{a} tal que $\mathbf{a}'\mathbf{x}$ tenha a maior correlação possível com y .

Introdução

- Em análise de regressão multipla temos um conjunto de covariáveis \mathbf{x} e uma única variável resposta y . A solução implica em encontrar \mathbf{a} tal que $\mathbf{a}'\mathbf{x}$ tenha a maior correlação possível com y .
- Suponha agora que y não é mais de dimensão $p = 1$, mas de dimensão $p > 1$.

Introdução

- Em análise de regressão multipla temos um conjunto de covariáveis \mathbf{x} e uma única variável resposta y . A solução implica em encontrar \mathbf{a} tal que $\mathbf{a}'\mathbf{x}$ tenha a maior correlação possível com y .
- Suponha agora que y não é mais de dimensão $p = 1$, mas de dimensão $p > 1$.
- Análise de correlação canônica implica encontrar vetores \mathbf{a} e \mathbf{b} tal que $\eta = \mathbf{a}'\mathbf{x}$ e $\phi = \mathbf{b}'\mathbf{y}$ tenham a maior correlação possível.

Análise de Correlação Canônica

Análise de Correlação Canônica

Proposto por Hotelling (1935, 1936), tem como objetivo o estudo das relações lineares existentes entre dois conjuntos de variáveis (\mathbf{x} e \mathbf{y}).

Análise de Correlação Canônica

Proposto por Hotelling (1935, 1936), tem como objetivo o estudo das relações lineares existentes entre dois conjuntos de variáveis (\mathbf{x} e \mathbf{y}).

- As combinações lineares ($\eta = \mathbf{a}'\mathbf{x}$ e $\phi = \mathbf{b}'\mathbf{y}$) são chamadas variáveis canônicas.
- A correlação entre $\eta = \mathbf{a}'\mathbf{x}$ e $\phi = \mathbf{b}'\mathbf{y}$ é chamada de correlação canônica e mede o grau de associação entre os dois conjuntos de variáveis.

Análise de Correlação Canônica

Sejam $\mathbf{X} \in \mathbb{R}^q$ e $\mathbf{Y} \in \mathbb{R}^p$ dois vetores aleatorios tais que:

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \left(\begin{pmatrix} \mu \\ \nu \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix} \right)$$

Análise de Correlação Canônica

Sejam $\mathbf{X} \in \mathbb{R}^q$ e $\mathbf{Y} \in \mathbb{R}^p$ dois vetores aleatórios tais que:

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \left(\begin{pmatrix} \mu \\ \nu \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix} \right)$$

Sejam os vetores \mathbf{a} e \mathbf{b} tais que $\eta = \mathbf{a}'\mathbf{X}$ e $\phi = \mathbf{b}'\mathbf{Y}$.

Análise de Correlação Canônica

Sejam $\mathbf{X} \in \mathbb{R}^q$ e $\mathbf{Y} \in \mathbb{R}^p$ dois vetores aleatorios tais que:

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \left(\begin{pmatrix} \mu \\ \nu \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix} \right)$$

Sejam os vetores \mathbf{a} e \mathbf{b} tais que $\eta = \mathbf{a}'\mathbf{X}$ e $\phi = \mathbf{b}'\mathbf{Y}$.

$$\rho = \text{Cor}(\eta, \phi) = \frac{\text{Cov}(\mathbf{a}'\mathbf{X}, \mathbf{b}'\mathbf{Y})}{\sqrt{\mathbb{V}(\mathbf{a}'\mathbf{X})}\sqrt{\mathbb{V}(\mathbf{b}'\mathbf{Y})}} = \frac{\mathbf{a}'\Sigma_{XY}\mathbf{b}}{\sqrt{\mathbf{a}'\Sigma_{XX}\mathbf{a}} \times \sqrt{\mathbf{b}'\Sigma_{YY}\mathbf{b}}}$$

Análise de Correlação Canônica

Dado que $\mathbb{C}or(\eta, \phi) = \mathbb{C}or(c\eta, d\phi)$, para $c, d \in \mathbb{R}$ (a correlação não depende da escala), maximizar $\mathbb{C}or(\eta, \phi)$ é equivalente a maximizar

$$\mathbf{a}'\Sigma_{XY}\mathbf{b},$$

sujeito às restrições

- $\mathbf{a}'\Sigma_{XX}\mathbf{a} = 1$
- $\mathbf{b}'\Sigma_{YY}\mathbf{b} = 1$

Análise de Correlação Canônica

Dado que $\mathbb{C}or(\eta, \phi) = \mathbb{C}or(c\eta, d\phi)$, para $c, d \in \mathbb{R}$ (a correlação não depende da escala), maximizar $\mathbb{C}or(\eta, \phi)$ é equivalente a maximizar

$$\mathbf{a}'\Sigma_{XY}\mathbf{b},$$

sujeito às restrições

- $\mathbf{a}'\Sigma_{XX}\mathbf{a} = 1$
- $\mathbf{b}'\Sigma_{YY}\mathbf{b} = 1$

Como nos interessa a magnitude da correlação e não o sinal, estamos, na verdade, interessados no seguinte problema de otimização

$$\text{Max: } (\mathbf{a}'\Sigma_{XY}\mathbf{b})^2 - \lambda(\mathbf{a}'\Sigma_{XX}\mathbf{a} - 1) - \gamma(\mathbf{b}'\Sigma_{YY}\mathbf{b} - 1)$$

Análise de Correlação Canônica

$$\text{Max: } M = (\mathbf{a}'\Sigma_{XY}\mathbf{b})^2 - \lambda(\mathbf{a}'\Sigma_{XX}\mathbf{a} - 1) - \gamma(\mathbf{b}'\Sigma_{YY}\mathbf{b} - 1)$$

Derivando w.r.t \mathbf{a} e \mathbf{b} temos que:

$$\frac{\partial M}{\partial \mathbf{a}} = 2\Sigma_{XY}\mathbf{b} - 2\lambda\Sigma_{XX}\mathbf{a}, \quad \frac{\partial M}{\partial \mathbf{b}} = 2\underbrace{\Sigma'_{XY}}_{\Sigma_{YX}}\mathbf{a} - 2\gamma\Sigma_{YY}\mathbf{b}$$

Análise de Correlação Canônica

$$\text{Max: } M = (\mathbf{a}'\Sigma_{XY}\mathbf{b})^2 - \lambda(\mathbf{a}'\Sigma_{XX}\mathbf{a} - 1) - \gamma(\mathbf{b}'\Sigma_{YY}\mathbf{b} - 1)$$

Derivando w.r.t \mathbf{a} e \mathbf{b} temos que:

$$\frac{\partial M}{\partial \mathbf{a}} = 2\Sigma_{XY}\mathbf{b} - 2\lambda\Sigma_{XX}\mathbf{a}, \quad \frac{\partial M}{\partial \mathbf{b}} = 2\underbrace{\Sigma'_{XY}}_{\Sigma_{YX}}\mathbf{a} - 2\gamma\Sigma_{YY}\mathbf{b}$$

Igualando a zero:

$$\Sigma_{XY}\mathbf{b} = \lambda\Sigma_{XX}\mathbf{a} \quad \text{e} \quad \Sigma_{YX}\mathbf{a} = \gamma\Sigma_{YY}\mathbf{b} \quad (1)$$

Análise de Correlação Canônica

$$\text{Max: } M = (\mathbf{a}'\Sigma_{XY}\mathbf{b})^2 - \lambda(\mathbf{a}'\Sigma_{XX}\mathbf{a} - 1) - \gamma(\mathbf{b}'\Sigma_{YY}\mathbf{b} - 1)$$

Derivando w.r.t \mathbf{a} e \mathbf{b} temos que:

$$\frac{\partial M}{\partial \mathbf{a}} = 2\Sigma_{XY}\mathbf{b} - 2\lambda\Sigma_{XX}\mathbf{a}, \quad \frac{\partial M}{\partial \mathbf{b}} = 2\underbrace{\Sigma'_{XY}}_{\Sigma_{YX}}\mathbf{a} - 2\gamma\Sigma_{YY}\mathbf{b}$$

Igualando a zero:

$$\Sigma_{XY}\mathbf{b} = \lambda\Sigma_{XX}\mathbf{a} \quad \text{e} \quad \Sigma_{YX}\mathbf{a} = \gamma\Sigma_{YY}\mathbf{b} \quad (1)$$

Se premultiplicarmos a primeira equação por \mathbf{a}' e a segunda por \mathbf{b}' , temos:

Análise de Correlação Canônica

$$\mathbf{a}'\Sigma_{XY}\mathbf{b} = \lambda \underbrace{\mathbf{a}'\Sigma_{XX}\mathbf{a}}_1 = \lambda \quad e \quad \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma \underbrace{\mathbf{b}'\Sigma_{YY}\mathbf{b}}_1 = \gamma$$

Análise de Correlação Canônica

$$\mathbf{a}'\Sigma_{XY}\mathbf{b} = \lambda \underbrace{\mathbf{a}'\Sigma_{XX}\mathbf{a}}_1 = \lambda \quad e \quad \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma \underbrace{\mathbf{b}'\Sigma_{YY}\mathbf{b}}_1 = \gamma$$

Note que $\lambda = \mathbf{a}'\Sigma_{XY}\mathbf{b} = (\mathbf{a}'\Sigma_{XY}\mathbf{b})' = \mathbf{b}'\Sigma'_{XY}\mathbf{a} = \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma$.

Análise de Correlação Canônica

$$\mathbf{a}'\Sigma_{XY}\mathbf{b} = \lambda \underbrace{\mathbf{a}'\Sigma_{XX}\mathbf{a}}_1 = \lambda \quad e \quad \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma \underbrace{\mathbf{b}'\Sigma_{YY}\mathbf{b}}_1 = \gamma$$

Note que $\lambda = \mathbf{a}'\Sigma_{XY}\mathbf{b} = (\mathbf{a}'\Sigma_{XY}\mathbf{b})' = \mathbf{b}'\Sigma'_{XY}\mathbf{a} = \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma$.

Então, de (1), temos que

$$\Sigma_{XY}\mathbf{b} = \lambda\Sigma_{XX}\mathbf{a} \quad e \quad \Sigma_{YX}\mathbf{a} = \lambda\Sigma_{YY}\mathbf{b} \quad (2)$$

Análise de Correlação Canônica

$$\mathbf{a}'\Sigma_{XY}\mathbf{b} = \lambda \underbrace{\mathbf{a}'\Sigma_{XX}\mathbf{a}}_1 = \lambda \quad e \quad \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma \underbrace{\mathbf{b}'\Sigma_{YY}\mathbf{b}}_1 = \gamma$$

Note que $\lambda = \mathbf{a}'\Sigma_{XY}\mathbf{b} = (\mathbf{a}'\Sigma_{XY}\mathbf{b})' = \mathbf{b}'\Sigma'_{XY}\mathbf{a} = \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma$.

Então, de (1), temos que

$$\Sigma_{XY}\mathbf{b} = \lambda\Sigma_{XX}\mathbf{a} \quad e \quad \Sigma_{YX}\mathbf{a} = \lambda\Sigma_{YY}\mathbf{b} \quad (2)$$

Colocando \mathbf{b} em evidência na segunda equação e substituindo na primeira:

$$\mathbf{b} = \lambda^{-1}\Sigma_{YY}^{-1}\Sigma_{YX}\mathbf{a} \quad e \quad \Sigma_{XY}\lambda^{-1}\Sigma_{YY}^{-1}\Sigma_{YX}\mathbf{a} = \lambda\Sigma_{XX}\mathbf{a}.$$

Análise de Correlação Canônica

$$\mathbf{a}'\Sigma_{XY}\mathbf{b} = \lambda \underbrace{\mathbf{a}'\Sigma_{XX}\mathbf{a}}_1 = \lambda \quad \text{e} \quad \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma \underbrace{\mathbf{b}'\Sigma_{YY}\mathbf{b}}_1 = \gamma$$

Note que $\lambda = \mathbf{a}'\Sigma_{XY}\mathbf{b} = (\mathbf{a}'\Sigma_{XY}\mathbf{b})' = \mathbf{b}'\Sigma'_{XY}\mathbf{a} = \mathbf{b}'\Sigma_{YX}\mathbf{a} = \gamma$.

Então, de (1), temos que

$$\Sigma_{XY}\mathbf{b} = \lambda\Sigma_{XX}\mathbf{a} \quad \text{e} \quad \Sigma_{YX}\mathbf{a} = \lambda\Sigma_{YY}\mathbf{b} \quad (2)$$

Colocando \mathbf{b} em evidência na segunda equação e substituindo na primeira:

$$\mathbf{b} = \lambda^{-1}\Sigma_{YY}^{-1}\Sigma_{YX}\mathbf{a} \quad \text{e} \quad \Sigma_{XY}\lambda^{-1}\Sigma_{YY}^{-1}\Sigma_{YX}\mathbf{a} = \lambda\Sigma_{XX}\mathbf{a}.$$

Então

$$\mathbf{b} = \lambda^{-1}\Sigma_{YY}^{-1}\Sigma_{YX}\mathbf{a} \quad \text{e} \quad \Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}\mathbf{a} = \lambda^2\mathbf{a}.$$

Análise de Correlação Canônica

\mathbf{a} é o autovetor associado ao autovalor λ^2 da matriz $\Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$!

Análise de Correlação Canônica

a é o autovetor associado ao autovalor λ^2 da matriz $\Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$!

De forma semelhante, obtemos que **b** é o autovetor associado ao autovalor λ^2 da matriz $\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}$ (**Provar!**)

Análise de Correlação Canônica

a é o autovetor associado ao autovalor λ^2 da matriz $\Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$!

De forma semelhante, obtemos que **b** é o autovetor associado ao autovalor λ^2 da matriz $\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}$ (**Provar!**)

Voilà

Note que $\rho^2 = \lambda^2 = \eta^2$, ou seja, para maximizar a correlação entre as variáveis canônicas, os vetores **a** e **b** devem ser os autovetores associados a λ^2 de $\Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$ e $\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}$, respectivamente.

Análise de Correlação Canônica

Algoritmo

- 1 Identificar as matrizes Σ_{XX} , Σ_{YY} e Σ_{XY} .
- 2 Calcular $\mathbf{A} = \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}$ e $\mathbf{B} = \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}$
- 3 Fazer a decomposição espectral $\mathbf{A}\mathbf{a} = \lambda^2\mathbf{a}$ e $\mathbf{B}\mathbf{b} = \lambda^2\mathbf{b}$

Análise de Correlação Canônica

Algoritmo

- 1 Identificar as matrizes Σ_{XX} , Σ_{YY} e Σ_{XY} .
- 2 Calcular $\mathbf{A} = \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}$ e $\mathbf{B} = \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}$
- 3 Fazer a decomposição espectral $\mathbf{A}\mathbf{a} = \lambda^2 \mathbf{a}$ e $\mathbf{B}\mathbf{b} = \lambda^2 \mathbf{b}$

Note que de (2) temos que:

$$\mathbf{a} = \Sigma_{XX}^{-1} \Sigma_{XY} \mathbf{b} \lambda^{-1} \quad \text{e} \quad \mathbf{b} = \Sigma_{YY}^{-1} \Sigma'_{XY} \mathbf{a} \lambda^{-1}$$

Obtendo \mathbf{a} (\mathbf{b}) teremos \mathbf{b} (\mathbf{a}). Só precisamos obter os autovetores de uma das matrizes (\mathbf{A} ou \mathbf{B}).

Análise de Correlação Canônica

Observação:

- É possível que, uma vez encontrados **a** e **b**, não exista mais relação entre ambos os conjuntos de dados. Neste caso dizemos que toda a relação entre ambos os conjuntos se resume em uma dimensão.

Análise de Correlação Canônica

Observação:

- É possível que, uma vez encontrados **a** e **b**, não exista mais relação entre ambos os conjuntos de dados. Neste caso dizemos que toda a relação entre ambos os conjuntos se resume em uma dimensão.
- Para verificar isto, podemos calcular as $2r$ combinações lineares ($r = \min\{p, q\}$), extraíndo os r autovetores associados aos r maiores autovalores de **A** e **B**.

Propriedades

Propriedades

- ① Os autovalores de \mathbf{A} e \mathbf{B} são reais e não negativos.
- ② $\mathbf{a}_i' \mathbf{X}$ e $\mathbf{a}_j' \mathbf{X}$ são não correlacionadas.
- ③ $\mathbf{b}_i' \mathbf{Y}$ e $\mathbf{b}_j' \mathbf{Y}$ são não correlacionadas.
- ④ $\mathbf{a}_i' \mathbf{X}$ e $\mathbf{b}_j' \mathbf{Y}$ são não correlacionadas.
- ⑤ Se $\mathbf{a}_i' \mathbf{X}$ é uma variável canônica, também é $-\mathbf{a}_i' \mathbf{X}$
- ⑥ As correlações canônicas λ_i^2 são o quadrado do coeficiente de correlação entre $\mathbf{a}_i' \mathbf{X}$ e $\mathbf{b}_i' \mathbf{Y}$

Propriedades: P1

Lema

As matrizes $\mathbf{R}^{-1}\mathbf{Q}$ e $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}$ tem os mesmos autovalores.

Propriedades: P1

Lema

As matrizes $\mathbf{R}^{-1}\mathbf{Q}$ e $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}$ tem os mesmos autovalores.

Demonstração:

- $\mathbf{R}^{-1}\mathbf{Q}\mathbf{a} = \lambda\mathbf{a}$
- $\mathbf{R}^{1/2}\mathbf{R}^{-1}\mathbf{Q}\mathbf{a} = \lambda\mathbf{R}^{1/2}\mathbf{a}$
- $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{R}^{1/2}\mathbf{a} = \lambda\mathbf{R}^{1/2}\mathbf{a}$
- $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}\underbrace{\mathbf{a}^*}_{\mathbf{R}^{1/2}\mathbf{a}} = \lambda\underbrace{\mathbf{a}^*}_{\mathbf{R}^{1/2}\mathbf{a}}$

Propriedades: P1

Lema

As matrizes $\mathbf{R}^{-1}\mathbf{Q}$ e $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}$ tem os mesmos autovalores.

Demonstração:

- $\mathbf{R}^{-1}\mathbf{Q}\mathbf{a} = \lambda\mathbf{a}$
- $\mathbf{R}^{1/2}\mathbf{R}^{-1}\mathbf{Q}\mathbf{a} = \lambda\mathbf{R}^{1/2}\mathbf{a}$
- $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{R}^{1/2}\mathbf{a} = \lambda\mathbf{R}^{1/2}\mathbf{a}$
- $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2} \underbrace{\mathbf{a}^*}_{\mathbf{R}^{1/2}\mathbf{a}} = \lambda \underbrace{\mathbf{a}^*}_{\mathbf{R}^{1/2}\mathbf{a}}$

$\mathbf{R}^{-1}\mathbf{Q}$ e $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}$ tem os mesmos autovalores. Note que se \mathbf{a} é o autovetor de $\mathbf{R}^{-1}\mathbf{Q}$, então $\mathbf{R}^{1/2}\mathbf{a}$ é o autovetor de $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}$.

Propriedades: P1

- Se fizermos $\mathbf{R} = \Sigma_{XX}$ e $\mathbf{Q} = \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$, temos que

$$\mathbf{R}^{-1}\mathbf{Q} = \underbrace{\Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}}_{\mathbf{A}} \text{ e } \mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2} =$$

$$\Sigma_{XX}^{-1/2}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1/2} = \underbrace{\Sigma_{XX}^{-1/2}\Sigma_{XY}\Sigma_{YY}^{-1/2}}_{\mathbf{H}} \underbrace{\Sigma_{YY}^{-1/2}\Sigma_{YX}\Sigma_{XX}^{-1/2}}_{\mathbf{H}'} \text{ tem}$$

os mesmo autovalores.

Propriedades: P1

- Se fizermos $\mathbf{R} = \Sigma_{XX}$ e $\mathbf{Q} = \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}$, temos que

$$\mathbf{R}^{-1} \mathbf{Q} = \underbrace{\Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}}_{\mathbf{A}} \text{ e } \mathbf{R}^{-1/2} \mathbf{Q} \mathbf{R}^{-1/2} =$$

$$\Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1/2} = \underbrace{\Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}}_{\mathbf{H}} \underbrace{\Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1/2}}_{\mathbf{H}'} \text{ tem}$$

os mesmo autovalores.

- Se fizermos $\mathbf{R} = \Sigma_{YY}$ e $\mathbf{Q} = \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}$, temos que

$$\mathbf{R}^{-1} \mathbf{Q} = \underbrace{\Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}}_{\mathbf{B}} \text{ e } \mathbf{R}^{-1/2} \mathbf{Q} \mathbf{R}^{-1/2} =$$

$$\Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1/2} = \underbrace{\Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1/2}}_{\mathbf{H}'} \underbrace{\Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}}_{\mathbf{H}} \text{ tem os}$$

mesmos autovalores

Propriedades: P1

- **A** tem os mesmo autovalores do que **HH'** e **B** tem os mesmos autovalores do que **H'H**.

Propriedades: P1

- **A** tem os mesmo autovalores do que **HH'** e **B** tem os mesmos autovalores do que **H'H**.
- Como **HH'** e **H'H** são simétricas e semidefinidas positivas, os autovalores são reais e não negativos (**Propriedade**)

Propriedades: P2

- Sejam \mathbf{a}_1 e \mathbf{a}_2 autovetores de \mathbf{A} (associados a diferentes autovalores, digamos λ_1 e λ_2).

Propriedades: P2

- Sejam \mathbf{a}_1 e \mathbf{a}_2 autovetores de \mathbf{A} (associados a diferentes autovalores, digamos λ_1 e λ_2).
- Então os correspondentes autovetores de $\mathbf{H}\mathbf{H}'$ (associados aos mesmos λ_1 e λ_2) são, respectivamente $\Sigma_{XX}^{1/2}\mathbf{a}_1$ e $\Sigma_{XX}^{1/2}\mathbf{a}_2$ (**ver Lema**)

Propriedades: P2

- Sejam \mathbf{a}_1 e \mathbf{a}_2 autovetores de \mathbf{A} (associados a diferentes autovalores, digamos λ_1 e λ_2).
- Então os correspondentes autovetores de $\mathbf{H}\mathbf{H}'$ (associados aos mesmos λ_1 e λ_2) são, respectivamente $\Sigma_{XX}^{1/2}\mathbf{a}_1$ e $\Sigma_{XX}^{1/2}\mathbf{a}_2$ (**ver Lema**)
- Como $\mathbf{H}\mathbf{H}'$ é simétrica, então $\Sigma_{XX}^{1/2}\mathbf{a}_1$ e $\Sigma_{XX}^{1/2}\mathbf{a}_2$ são ortogonais, o que implica que $(\Sigma_{XX}^{1/2}\mathbf{a}_1)' \Sigma_{XX}^{1/2}\mathbf{a}_2 = \mathbf{a}_1' \Sigma_{XX}\mathbf{a}_2 = 0$

Propriedades: P2

- Sejam \mathbf{a}_1 e \mathbf{a}_2 autovetores de \mathbf{A} (associados a diferentes autovalores, digamos λ_1 e λ_2).
- Então os correspondentes autovetores de $\mathbf{H}\mathbf{H}'$ (associados aos mesmos λ_1 e λ_2) são, respectivamente $\Sigma_{XX}^{1/2}\mathbf{a}_1$ e $\Sigma_{XX}^{1/2}\mathbf{a}_2$ (**ver Lema**)
- Como $\mathbf{H}\mathbf{H}'$ é simétrica, então $\Sigma_{XX}^{1/2}\mathbf{a}_1$ e $\Sigma_{XX}^{1/2}\mathbf{a}_2$ são ortogonais, o que implica que $(\Sigma_{XX}^{1/2}\mathbf{a}_1)' \Sigma_{XX}^{1/2}\mathbf{a}_2 = \mathbf{a}_1' \Sigma_{XX}\mathbf{a}_2 = 0$
- Então, $\text{Cov}(\mathbf{a}_1'\mathbf{X}, \mathbf{a}_2'\mathbf{X}) = \mathbf{a}_1' \Sigma_{XX}\mathbf{a}_2 = 0$ (as variáveis canônicas são não correlacionadas).

Propriedades: P4

$$\text{Cov}(\mathbf{a}'_1 \mathbf{X}, \mathbf{b}'_2 \mathbf{Y}) = \mathbf{a}'_1 \underbrace{\Sigma_{XY} \mathbf{b}_2}_{\lambda \Sigma_{XX} \mathbf{a}_2 \quad (\text{ver (2)})} = \lambda \underbrace{\mathbf{a}'_1 \Sigma_{XX} \mathbf{a}_2}_0 = 0$$

Implementação

Implementação

Algoritmo I

- Calcular as versões amostrais de **A** e **B** dadas por $\hat{\mathbf{A}} = \mathbf{S}_{XX}^{-1} \mathbf{S}_{XY} \mathbf{S}_{YY}^{-1} \mathbf{S}_{YX}$ e $\hat{\mathbf{B}} = \mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{S}_{XX}^{-1} \mathbf{S}_{XY}$
- Decompor a matriz de menor dimensão, digamos $\hat{\mathbf{A}}$.
- Seja **a** o vetor próprio associado a λ^2 ($\mathbf{S}_{XX}^{-1} \mathbf{S}_{XY} \mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{a} = \lambda^2 \mathbf{a}$).
Então o vetor próprio associado a λ^2 mas obtido com a decomposição de $\hat{\mathbf{B}}$ será $\mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{a}$ ($\mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{S}_{XX}^{-1} \mathbf{S}_{XY} \mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{a} = \lambda^2 \mathbf{S}_{YY}^{-1} \underbrace{\mathbf{S}_{YX} \mathbf{a}}_{\mathbf{b}}$)

Implementação

Algoritmo I

- Calcular as versões amostrais de **A** e **B** dadas por $\hat{\mathbf{A}} = \mathbf{S}_{XX}^{-1} \mathbf{S}_{XY} \mathbf{S}_{YY}^{-1} \mathbf{S}_{YX}$ e $\hat{\mathbf{B}} = \mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{S}_{XX}^{-1} \mathbf{S}_{XY}$
- Decompor a matriz de menor dimensão, digamos $\hat{\mathbf{A}}$.
- Seja **a** o vetor próprio associado a λ^2 ($\mathbf{S}_{XX}^{-1} \mathbf{S}_{XY} \mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{a} = \lambda^2 \mathbf{a}$).
Então o vetor próprio associado a λ^2 mas obtido com a decomposição de $\hat{\mathbf{B}}$ será $\mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{a}$ ($\mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{S}_{XX}^{-1} \mathbf{S}_{XY} \mathbf{S}_{YY}^{-1} \mathbf{S}_{YX} \mathbf{a} = \lambda^2 \mathbf{S}_{YY}^{-1} \underbrace{\mathbf{S}_{YX} \mathbf{a}}_{\mathbf{b}}$)

Mas, *peraí*...

Implementação

Lembrete: Definição SVD

Qualquer matriz $\mathbf{M}_{m \times n}$ de posto r pode ser decomposta como:

$$\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{V}^t,$$

em que $\mathbf{U}_{m \times r}$, $\mathbf{V}_{n \times r}$ e $\mathbf{D}_{r \times r} = \text{Diag}\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}\}$. $\lambda_1, \dots, \lambda_r$ são os autovalores (em ordem decrescente) de $\mathbf{M}\mathbf{M}'$.

- Os elementos D_{ij} são chamados de valores singulares da matriz \mathbf{M} .
- As colunas de \mathbf{U} são os r autovetores (normalizados) associados a $\mathbf{M}\mathbf{M}'$.
- As colunas de \mathbf{V} são os r autovetores (normalizados) associados a $\mathbf{M}'\mathbf{M}$.

Implementação

Seja $\mathbf{M} = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}$, então:

- $\mathbf{M}\mathbf{M}' = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1/2}$
- $\mathbf{M}'\mathbf{M} = \Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1/2}$

Implementação

Seja $\mathbf{M} = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}$, então:

- $\mathbf{M}\mathbf{M}' = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1/2}$
- $\mathbf{M}'\mathbf{M} = \Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1/2}$

Pelo Lema 1, as matrizes $\mathbf{R}^{-1}\mathbf{Q}$ e $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}$ tem os mesmos autovalores, com respectivos autovalores δ e $\mathbf{R}^{1/2}\delta$.

Implementação

Seja $\mathbf{M} = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}$, então:

- $\mathbf{MM}' = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1/2}$
- $\mathbf{M}'\mathbf{M} = \Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1/2}$

Pelo Lema 1, as matrizes $\mathbf{R}^{-1}\mathbf{Q}$ e $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}$ tem os mesmos autovalores, com respectivos autovalores δ e $\mathbf{R}^{1/2}\delta$.

- Se fizermos $\mathbf{R} = \Sigma_{XX}$ e $\mathbf{Q} = \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}$, os autovetores de \mathbf{MM}' são da forma $\mathbf{R}^{1/2}\mathbf{a}$ em que \mathbf{a} são os autovetores de $\mathbf{R}^{-1}\mathbf{Q} = \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} = \mathbf{A}$.

Implementação

Seja $\mathbf{M} = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}$, então:

- $\mathbf{MM}' = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1/2}$
- $\mathbf{M}'\mathbf{M} = \Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1/2}$

Pelo Lema 1, as matrizes $\mathbf{R}^{-1}\mathbf{Q}$ e $\mathbf{R}^{-1/2}\mathbf{Q}\mathbf{R}^{-1/2}$ tem os mesmos autovalores, com respectivos autovalores δ e $\mathbf{R}^{1/2}\delta$.

- Se fizermos $\mathbf{R} = \Sigma_{XX}$ e $\mathbf{Q} = \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}$, os autovetores de \mathbf{MM}' são da forma $\mathbf{R}^{1/2}\mathbf{a}$ em que \mathbf{a} são os autovetores de $\mathbf{R}^{-1}\mathbf{Q} = \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} = \mathbf{A}$.
- Se fizermos $\mathbf{R} = \Sigma_{YY}$ e $\mathbf{Q} = \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}$, os autovetores de $\mathbf{M}'\mathbf{M}$ são da forma $\mathbf{R}^{1/2}\mathbf{b}$ em que \mathbf{b} são os autovetores de $\mathbf{R}^{-1}\mathbf{Q} = \Sigma_{YY}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} = \mathbf{B}$.

Implementação

Algoritmo II

- Defina $\mathbf{M} = \mathbf{S}_{XX}^{-1/2} \mathbf{S}_{YX} \mathbf{S}_{YY}^{-1/2}$
- Aplique a decomposição SVD ($\mathbf{M} = \mathbf{U} \mathbf{D} \mathbf{V}'$)
- Faça $\mathbf{a}_i = \mathbf{S}_{XX}^{-1/2} \mathbf{u}_i$ e $\mathbf{b}_i = \mathbf{S}_{YY}^{-1/2} \mathbf{v}_i$

Implementação

```
library(expm)
cc_me731 = function(x, y) {
  S_xx <- cov(x)
  S_yy <- cov(y)
  S_xy <- cov(x, y)
  M <- sqrtm(solve(S_xx)) %*% S_xy %*% sqrtm(solve(S_yy))
  decomposicao_svd <- svd(M)
  a <- sqrtm(solve(S_xx)) %*% decomposicao_svd$u
  b <- sqrtm(solve(S_yy)) %*% decomposicao_svd$v
  lambda <- decomposicao_svd$d
  return(list(a, b, lambda))
}
```

Como exemplo, utilizaremos o *data set* disponível [aqui](#)

Implementação

```
library(dplyr)
library(CCA)
dados <- read.table("https://raw.githubusercontent.com/ctruciosm/...",
colnames(dados) <- c("name", "economy",
                      "service", "value",
                      "price", "design",
                      "sporty", "safety",
                      "handling")
X <- dados %>% dplyr::select(price, value)
Y <- dados %>% dplyr::select(-price, -value, -name)
cor_canonica_na_mao <- cc_me731(X, Y)
cc_results <- stats::cancor(X,Y)
cca_results <- cc(X, Y)
```

Implementação

```
# Autovalores
```

```
cca_results$cor # Pacote CCA
```

```
## [1] 0.9791972 0.8851224
```

```
cc_results$cor # Pacote stats
```

```
## [1] 0.9791972 0.8851224
```

```
cor_canonica_na_mao[[3]] # Nossa implementacao
```

```
## [1] 0.9791972 0.8851224
```

Implementação

```
# Autovetores
```

```
round(cca_results$xcoef, 3)    # Pacote CCA
```

```
##           [,1]    [,2]
```

```
## price -0.333 -1.602
```

```
## value  0.587 -1.686
```

```
round(cc_results$xcoef, 3)    # Pacote stats
```

```
##           [,1]    [,2]
```

```
## price  0.071 -0.342
```

```
## value -0.125 -0.360
```

Implementação

```
# Autovetores
```

```
round(cca_results$xcoef, 3) # Pacote CCA
```

```
##           [,1]    [,2]
```

```
## price -0.333 -1.602
```

```
## value  0.587 -1.686
```

```
round(cor_canonica_na_mao[[1]], 3) # Nossa implementacao
```

```
##           [,1]    [,2]
```

```
## [1,] -0.333  1.602
```

```
## [2,]  0.587  1.686
```

Implementação

```
# Autovetores
```

```
round(cca_results$ycoef, 3) # Pacote CCA
```

```
##           [,1]    [,2]  
## economy  -0.433 -0.568  
## service   0.191 -0.544  
## design    0.005  0.012  
## sporty    0.458  0.096  
## safety    0.223  0.014  
## handling  0.376 -0.915
```

Implementação

```
# Autovetores
```

```
round(cc_results$ycoef, 3)      # Pacote stats
```

##	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
## economy	0.092	-0.121	-0.255	-0.008	-0.037	0.340
## service	-0.041	-0.116	0.294	0.007	-0.730	0.329
## design	-0.001	0.003	-0.478	-0.516	-0.009	-0.051
## sporty	-0.098	0.020	0.014	0.439	-0.027	-0.256
## safety	-0.047	0.003	0.006	-0.006	0.549	0.208
## handling	-0.080	-0.195	-0.007	0.004	-0.016	-0.811

Implementação

```
# Autovetores
```

```
round(cor_canonica_na_mao[[2]], 3) # Nossa implementacao
```

```
##           [,1]    [,2]  
## [1,] -0.433    0.568  
## [2,]  0.191    0.544  
## [3,]  0.005   -0.012  
## [4,]  0.458   -0.096  
## [5,]  0.223   -0.014  
## [6,]  0.376    0.915
```

Implementação

- Nossa implementação a implementação do pacote CCA são iguais já os resultados obtidos com o pacote stats são diferentes.

Implementação

- Nossa implementação a implementação do pacote CCA são iguais já os resultados obtidos com o pacote stats são diferentes.
- Nossa implementação a implementação do pacote CCA batem com os resultados do livro (ver página 437 do livro *Applied Multivariate Statistical Analysis* do Hardle e Simar (2019, 6ed)).

Implementação

- Nossa implementação a implementação do pacote CCA são iguais já os resultados obtidos com o pacote stats são diferentes.
- Nossa implementação a implementação do pacote CCA batem com os resultados do livro (ver página 437 do livro *Applied Multivariate Statistical Analysis* do Hardle e Simar (2019, 6ed)).
- Na próxima aula continuaremos utilizando as três implementações para descobrir o que está acontecendo.

Referências

Referências

- Härdle, W. K., & Simar, L. (2019). Applied Multivariate Statistical Analysis. Fifth Editon. Springer Nature. Capítulo 16.
- Mardia, K. V., Kent, J. T., & Bibby, J, M. (1979). Multivariate Analysis. Academic Press. Capítulo 10.
- Peña, D. (2002). Análisis de Datos Multivariantes. Mc Graw Hill. Capítulo 16.