Introdução, notação e revisão sobre matrizes

Prof. Caio Azevedo

Motivação

- Até o momento, essencialmente, estudou-se metodologias para a análise de uma (única) variável resposta por vez.
- Em muitos casos, pode-se obter resultados mais precisos quando se analisa duas ou mais variáveis resposta simultaneamente.
- Outras vezes, há a necessidade (pelo próprio problema em si) de analisar duas ou mais variáveis resposta.
- Estudaremos métodos para a análise de duas ou mais variáveis resposta simultaneamente, podendo haver ou não variáveis explicativas (ou covariáveis).

Motivação

- Situação hipótetica: a altura e o peso de dois grupos podem ser comparados separadamente (análise univariada) ou simultaneamente (análise multivariada). Por exemplo, na primeira situação, as alturas médias podem não ser diferentes e os pesos médios sim, enquanto que na segunda as duas médias podem ser diferentes, entre os grupos.
- Caracterizar melhor as unidades experimentais (pessoas, animais, cidades, objetos etc) (várias medidas precisam ser consideradas).
- Charles Spearman, Thomson, Thurstone e Burt, buscaram obter uma melhor compreensão para "inteligência". Este conceito está relacionado à várias variáveis cognitivas (Análise Fatorial).

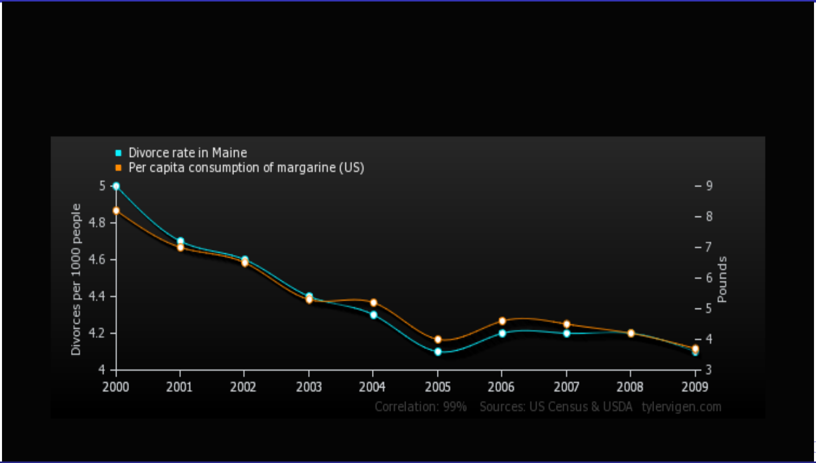
Idéias gerais sobre análise multivariada

- Toda metodologia de análise multivariada busca estudar e/ou compreender e/ou utilizar, quando existem, estruturas de dependência e/ou correlação.
  - Estrutura de dependência: modelo probabilístico ou modelo estatístico (modelo de regressão), ou seja, a distribuição conjunta das variáveis (resposta) de interesse.
  - Estrutura de correlação: Pearson (variáveis contínuas e/ou discretas com suporte correspondendo à um conjunto infinito), Spearman, Kendall (não paramétricas), policórica, tetracórica (variáveis discretas com suporte correspondendo à um conjunto finito).

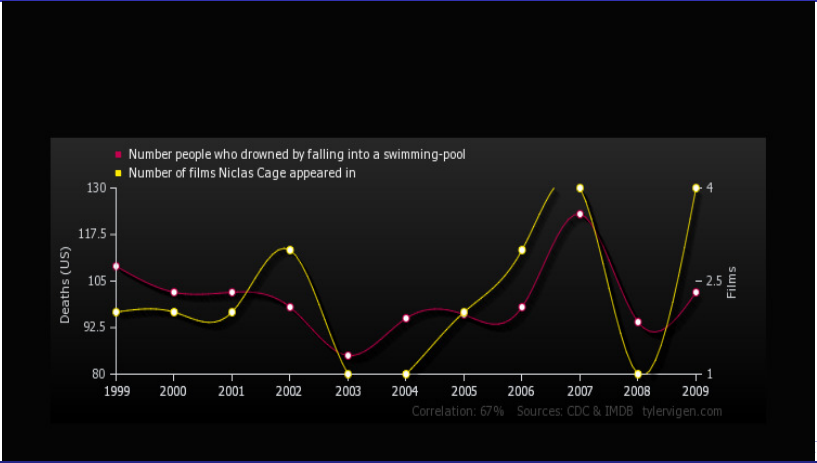
Correlação

- Existência de correlação (correlação significativa) não implica numa relação de causalidade (do ponto de vista do problema).
  - Em geral, altura e peso são positivamente correlacionados, mas a altura não é determinada (biologicamente) pelo peso e vice-versa.
  - Outros fatores: genética, alimentação, meio-ambiente, de fato determinam (simultaneamente) a altura e peso.
- Os dois gráficos a seguir foram extraídos do site `http://www.fastcodesign.com/3030529/infographic-of-the-day/hilarious-graphs-prove-that-correlation-isnt-causation`

Número de divórcios em Maine $\times$ Consumo per capita de margarina (EUA)



Número de pessoas que se afogaram em piscinas $\times$ número de filmes em que o Nicolas Cage apareceu



Notação

- X: variável aleatória, x: valor observado da variável aleatória.
- $\boldsymbol{X}$: vetor aleatório, $\boldsymbol{x}$: valor observado do vetor aleatório.
- $f_X(x)$: fdp (função densidade de probabilidade, ou simplesmente densidade, no caso contínuo, ou função de probabilidade, no caso discreto), relativa à uma variável aleatória.
- $f_{\boldsymbol{X}}(\boldsymbol{x})$: fdp (função densidade de probabilidade, ou simplesmente densidade, no caso contínuo, ou função de probabilidade, no caso discreto), relativa à um vetor aleatório.
- $\mathcal{E}(X)$ média, $\mathcal{V}(X)$ (variância), $Cov(X_1, X_2)$ (covariância), $Corre(X_1, X_2)$ (correlação).

Matriz de dados

Indivíduo	Variável 1	Variável 2	...	Variável p
1	X_{11}	X_{12}	...	X_{1p}
2	X_{21}	X_{22}	...	X_{2p}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	X_{n1}	X_{n2}	...	X_{np}

Em geral matrizes e vetores serão escritos com letra maiúscula (se forem objetos aleatórios) e com letra minúscula (se forem objetos não aleatórios).

Revisão de álgebra de matrizes

- Uma matriz é um arranjo retangular de elementos (números, variáveis aleatórias, letras):

$$\boldsymbol{A}_{(2\times 3)} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

- Um vetor é uma matriz com somente uma linha (vetor linha) ou somente uma coluna (vetor coluna).

$$\boldsymbol{A}_{(1\times 3)} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \end{bmatrix}; \boldsymbol{A}_{(3\times 1)} = \begin{bmatrix} A_{11} \\ A_{21} \\ A_{31} \end{bmatrix}$$

- Se a matriz tiver apenas uma linha e uma coluna teremos um escalar.

- Denotaremos uma matriz ou vetor por uma letra maiúscula ou minúscula em negrito.
- Uma matriz é dita ser quadrada se o número de linhas for igual ao número de colunas.

$$\boldsymbol{A}_{(3\times 3)} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

- Matriz diagonal: matriz quadrada em que todos os elementos fora da diagonal principal são iguais à 0.

$$\boldsymbol{A} = \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{bmatrix}$$

- Matriz diagonal inferior: matriz quadrada em que todos os elementos acima da diagonal principal são iguais à 0.

$$\boldsymbol{A} = \begin{bmatrix} A_{11} & 0 & 0 \\ A_{12} & A_{22} & 0 \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

- Matriz diagonal superior: matriz quadrada em que todos os elementos abaixo da diagonal principal são iguais à 0.

$$\boldsymbol{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}$$

- Matriz identidade: matriz diagonal cujos elementos da diagonal principal são todos iguais à 1.

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Matriz simétrica $(A_{ij} = A_{ji}), \forall i, j; i \neq j$.

$$\boldsymbol{A} = \begin{bmatrix} 5 & -2 & 3 \\ -2 & 7 & 1 \\ 3 & 1 & 4 \end{bmatrix}$$

Operações com matrizes

- Soma de matrizes :

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix} = \begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22} \\ A_{31}+B_{31} & A_{32}+B_{32} \end{bmatrix}$$

- Multiplicação de matrizes:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} \\ A_{31}B_{11}+A_{32}B_{21} & A_{31}B_{12}+A_{32}B_{22} \end{bmatrix}$$

- Traço de uma matriz quadrada: é a soma dos elementos de sua diagonal principal.

$$tr(\boldsymbol{A}) = tr\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = A_{11} + A_{22} + A_{33}$$

- O operador *vec* cria um vetor coluna, a partir de uma matriz, pela concatenação de suas colunas:

$$vec(\boldsymbol{A}) = vec\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ A_{31} \\ A_{12} \\ A_{22} \\ A_{32} \end{bmatrix}$$

- Sejam $\boldsymbol{A}_{(n \times p)}$ e $\boldsymbol{B}_{(m \times q)}$ duas matrizes quaisquer. O produto de Kronecker entre elas é definido por:

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} A_{11}\boldsymbol{B} & A_{12}\boldsymbol{B} & \dots & A_{1p}\boldsymbol{B} \\ A_{21}\boldsymbol{B} & A_{22}\boldsymbol{B} & \dots & A_{2p}\boldsymbol{B} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1}\boldsymbol{B} & A_{n2}\boldsymbol{B} & \dots & A_{np}\boldsymbol{B} \end{bmatrix}$$

- A resultante do produto de Kronecker será uma matriz $\boldsymbol{C}_{(nm \times pq)}$.
- O posto ou "rank" de uma matriz é o mínimo entre o número de linhas (posto linha) ou colunas (posto coluna) linearmente independentes.

- Matriz transposta (cada coluna da matriz original é transformada em uma linha) :

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 4 & -1 \\ 3 & 2 \end{bmatrix} \rightarrow \boldsymbol{A}' = \begin{bmatrix} 1 & 4 & 3 \\ 2 & -1 & 2 \end{bmatrix}$$

- Determinante de uma matriz quadrada: é uma função que associa um escalar para uma dada matriz quadrada. Existem alguns métodos para se calcular o determinante de uma matriz. Exemplo:

$$|\boldsymbol{A}| = det(\boldsymbol{A}) = det \begin{bmatrix} 1 & 2 \\ 4 & -1 \end{bmatrix} = -1 - 8 = -9$$

- Inversa de uma matriz : a inversa de uma matriz quadrada $\boldsymbol{A}$ denotada por $\boldsymbol{A}^{-1}$ é tal que:

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$$

Existem diversos métodos para se obter (numericamente) a inversa de uma matriz. A solução analítica é dada por:

$$\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{C}$$

em que $\boldsymbol{C}$ é a matriz adjunta de $\boldsymbol{A}$, ou seja, a transposta da matriz que se obtem substituindo cada termo A_{ij} pelo determinante da matriz resultante ao se retirar de $\boldsymbol{A}$ a linha i e a coluna j multiplicado por $(-1)^{i+j}$. Naturalmente, se $det(\boldsymbol{A}) = 0$ sua inversa não existe.

- Inversa generalizada de uma matriz (não necessariamente quadrada), é uma matriz $\boldsymbol{A}^{-}$ tal que

$$\boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} = \boldsymbol{A}$$

- Em princípio, uma matriz pode ter infinitas inversas generalizadas.
- Mesmo que o $det(\boldsymbol{A}) = 0$ sua inversa generalizada pode ser obtida.

- Autovalores de uma matriz quadrada de dimensão n: são os n números, digamos $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_n)'$, que satisfazem:

$$|\boldsymbol{A} - \boldsymbol{\lambda}\boldsymbol{I}| = 0$$

- Autovetores de uma matriz quadrada de dimensão n: são os n vetores que satisfazem a seguinte relação:

$$\boldsymbol{A}\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i; i = 1, ..., n$$

Os autovalores e autovetores são de extrema importância na decomposição e na verificação de certas propriedas de matrizes.

Tipos de matrizes

- Uma matriz (quadrada) é dita ser idempotente se $\boldsymbol{AA} = \boldsymbol{A}$.
- Uma matriz (quadrada) é dita ser ortogonal $\boldsymbol{A}^{-1} = \boldsymbol{A}'$.
- Uma matriz (quadrada) é dita ser:
  - Positiva definida: se $\lambda_i > 0, \forall i$.
  - Positiva semi-definida: se $\lambda_i \geq 0, \forall i$ e $\exists$ pelo menos um $\lambda_i = 0$.
  - Negativa semi-definida: se $\lambda_i \leq 0, \forall i$ e $\exists$ pelo menos um $\lambda_i = 0$.
  - Negativa definida: se $\lambda_i < 0, \forall i$.

Propriedades

- Se todas as operações estiverem bem definidas, teremos:
- $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$.
- $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$.
- $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.
- $tr(\boldsymbol{A} + \boldsymbol{B}) = tr(\boldsymbol{A}) + tr(\boldsymbol{B})$.
- $tr(a\boldsymbol{A}) = atr(\boldsymbol{A})$, a um escalar.
- $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$.

- $det(\boldsymbol{A}) = det(\boldsymbol{A}')$.
- $det(\boldsymbol{A}^{-1}) = [det(\boldsymbol{A})]^{-1}$.
- $det(\boldsymbol{AB}) = det(\boldsymbol{A})det(\boldsymbol{B})$ se A e B forem ambas quadradas.
- $\boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C}) = (\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C}$.
- $a \otimes \boldsymbol{A} = \boldsymbol{A} \otimes a = a\boldsymbol{A}$, a escalar.
- $a\boldsymbol{A} \otimes b\boldsymbol{B} = ab(\boldsymbol{A} \otimes \boldsymbol{B})$.
- $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = \boldsymbol{AC} \otimes \boldsymbol{BD}$.
- $(\boldsymbol{A} \otimes \boldsymbol{B})' = (\boldsymbol{A}' \otimes \boldsymbol{B}')$.
- $(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = (\boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1})$.
- $tr(\boldsymbol{A} \otimes \boldsymbol{B}) = tr(\boldsymbol{A})tr(\boldsymbol{B})$.
- Para $\boldsymbol{A}_{(m \times m)}$ e $\boldsymbol{B}_{(n \times n)}$, temos que $|\boldsymbol{A} \otimes \boldsymbol{B}| = |\boldsymbol{A}|^m |\boldsymbol{B}|^n$

- Forma quadrática. Sejam $\boldsymbol{Y}$ um vetor e $\boldsymbol{A}$ uma matriz quadrada. Dizemos que $\boldsymbol{Z} = \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ é uma forma quadrática em $\boldsymbol{Y}$ com matriz núcleo $\boldsymbol{A}$.

$$
\begin{aligned}
\boldsymbol{Z} &= \begin{bmatrix} Y_1 & Y_2 & \ldots & Y_p \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1p} \\ A_{21} & A_{22} & \ldots & A_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \ldots & A_{pp} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \ldots \\ Y_p \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^{p} A_{1i} Y_1 & \sum_{i=1}^{p} A_{2i} Y_2 & \ldots & \sum_{i=1}^{p} A_{pi} Y_p \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \ldots \\ Y_p \end{bmatrix} \\
&= \sum_{i=1}^{p} \sum_{j=1}^{p} A_{ij} Y_i Y_j, \text{Se } \boldsymbol{A} = \boldsymbol{I}, \text{então } \boldsymbol{Z} = \sum_{i=1}^{p} Y_i^2 .
\end{aligned}
$$

Decomposição de Cholesky

- Seja $\boldsymbol{\Sigma}_{p \times p}$, simétrica e positiva definida (todos os seus auto-valores são positivos).
- Exemplo: Matriz de covariâncias e matriz de correlações.
- Definição: Dada uma matriz $\boldsymbol{\Sigma}$ existe uma matriz $\boldsymbol{L}$, diagonal inferior com os valores da diagonal estritamente positivos, $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}'$.
- Tal decomposição é única.
- Existem alguns algoritmos que possibilitam a obtenção de $\boldsymbol{L}$.

Exemplo

- Exemplo:

$$\mathbf{\Sigma} = \begin{bmatrix} 3 & 1 & 0.5 \\ 1 & 2 & 0.8 \\ 0.5 & 0.8 & 4 \end{bmatrix}$$

$$\boldsymbol{L} \approx \begin{bmatrix} 1,732 & 0 & 0 \\ 0,577 & 1,291 & 0 \\ 0,289 & 0,491 & 1,917 \end{bmatrix}$$

- Seja $\boldsymbol{\Sigma}_{p \times p}$. Então, $\boldsymbol{\Sigma}$ pode ser fatorizada como

$$\boldsymbol{\Sigma} = \boldsymbol{E}\boldsymbol{\Lambda}\boldsymbol{E}'$$

em que $\boldsymbol{\Lambda} = diag(\lambda_1, ...\lambda_p)$ (autovalores) e as colunas da matriz $\boldsymbol{E}$ são formadas pelos respectivos autovetores orto-normalizados.

- A decomposição acima também vale se autovetores não estiverem ortonormalizados, neste caso, ela é dada por:

$$\boldsymbol{\Sigma} = \boldsymbol{E}\Lambda\boldsymbol{E}^{-1}$$

- **Observação: basicamente todas as operações e decomposições matriciais apresentadas estão implementadas no programa R.**

Alguns comandos para operações matriciais em R

Operação	Comando
$\boldsymbol{A}'$	t($\boldsymbol{A}$)
$\boldsymbol{A}^{-1}$	solve($\boldsymbol{A}$)
Cholesky($\boldsymbol{A}$)	chol($\boldsymbol{A}$)
$\boldsymbol{AB}$	$\boldsymbol{A}$%*%$\boldsymbol{B}$
$\boldsymbol{A} \otimes \boldsymbol{B}$	kronecker($\boldsymbol{A}$,$\boldsymbol{B}$)
Autovalores e autovetores de $\boldsymbol{A}$	eigen($\boldsymbol{A}$)