CE071 - Análise de Regressão Linear

Cesar Augusto Taconeli

26 de março, 2018

Revisão de matrizes

Definição e propriedades básicas

- Uma matriz é um conjunto de números ou variáveis dispostos em linhas e colunas.
- Uma matriz **A** de n linhas e p colunas (dimensão $n \times p$) pode ser representada, genericamente, por:

$$\boldsymbol{A}_{n \times p} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix},$$

- A matriz **A** pode ser denotada ainda por $\mathbf{A} = \{a_{ij}\}$, onde o primeiro índice indica linha, o segundo coluna e a_{ij} é o termo geral da matriz.

Definição e propriedades básicas

- Um vetor **x**, de dimensão n, é representado, genericamente, por:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- Numa análise de regressão com n indivíduos e p variáveis, as linhas da matriz de dados (observações dos indivíduos) podem ser consideradas n vetores de tamanho p: $\boldsymbol{x}_i' = (x_{i1}, x_{i2}, ..., x_{ip}), i = 1, 2, ..., n$;
- As colunas da matriz de dados (observações referentes às variáveis) podem ser consideradas p vetores de tamanho n: $\boldsymbol{x}_j' = (x_{1j}, x_{2j}, ..., x_{nj}), j = 1, 2, ..., p$.

Definição e propriedades básicas

- A multiplicação de um vetor $\boldsymbol{x} = (x_1, x_2, ..., x_p)'$ por uma constante c resulta em um vetor $\boldsymbol{y} = c\boldsymbol{x} = (cx_1, cx_2, ..., cx_p)'$, de igual dimensão em relação ao vetor original;
    - Geometricamente, a multiplicação de um vetor por um escalar pode mudar seu tamanho e sentido, mas não sua direção.
- A soma de dois vetores $\boldsymbol{x}$ e $\boldsymbol{y}$, de iguais dimensões, resulta em um terceiro vetor dado por:

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, ..., x_p + y_p)'$$

Definição e propriedades básicas

- A diferença de dois vetores $\boldsymbol{x}$ e $\boldsymbol{y}$, de iguais dimensões, resulta em um terceiro vetor dado por:

$$\boldsymbol{w} = \boldsymbol{x} - \boldsymbol{y} = (x_1 - y_1, x_2 - y_2, ..., x_p - y_p)'.$$

- O produto interno de dois vetores $\boldsymbol{x}$ e $\boldsymbol{y}$, é definido por:

$$\boldsymbol{v} = \boldsymbol{x}'\boldsymbol{y} = \sum_{i=1}^{p} x_i y_i = x_1 y_1 + x_2 y_2 + ... + x_p y_p.$$

- O tamanho do vetor $\boldsymbol{x} = (x_1, x_2, ..., x_p)'$ é definido pela distância do ponto p-dimensional, determinado por suas coordenadas, à origem:

$$L_{\boldsymbol{x}} = \sqrt{\boldsymbol{x}'\boldsymbol{x}} = \sqrt{x_1^2 + x_2^2 + ... + x_p^2}.$$

Definição e propriedades básicas

- O co-seno do ângulo θ entre os vetores $\boldsymbol{x}$ e $\boldsymbol{y}$ definidos numa mesma dimensão é dado por:

$$cos(\theta) = \frac{\boldsymbol{x}'\boldsymbol{y}}{\sqrt{\boldsymbol{x}'\boldsymbol{x}}\sqrt{\boldsymbol{y}'\boldsymbol{y}}}.$$

- Dois vetores $\boldsymbol{x}$ e $\boldsymbol{y}$ são **ortogonais** entre si se o ângulo θ entre eles é 90º, de tal forma que $cos(\theta) = 0$, ou, de forma equivalente, $\boldsymbol{x}'\boldsymbol{y} = 0$.
- A normalização de um vetor $\boldsymbol{x}$ corresponde à divisão de $\boldsymbol{x}$ por $L_{\boldsymbol{x}}$, de tal forma que o vetor resultante tenha comprimento unitário:

$$\boldsymbol{x}^* = \frac{\boldsymbol{x}}{L_{\boldsymbol{x}}}.$$

Definição e propriedades básicas

- Igualdade de matrizes: Dizemos que duas matrizes $\boldsymbol{A}$ e $\boldsymbol{B}$ são iguais se elas tem iguais dimensões e $\{a_{ij}\} = \{b_{ij}\}$ para todo i e para todo j.
- Matriz transposta: A transposta de uma matriz $\boldsymbol{A}_{\boldsymbol{n}\times\boldsymbol{p}}$ é a matriz $\boldsymbol{A}'_{\boldsymbol{p}\times\boldsymbol{n}}$ tal que $\boldsymbol{a}'_{\boldsymbol{ij}} = \boldsymbol{a}_{\boldsymbol{ji}}$ para todo i e para todo j:

$$\mathbf{A}' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ldots & \vdots \\ a_{1p} & a_{2p} & \cdots & a_{np} \end{bmatrix}$$

Definição e propriedades básicas

- Matriz simétrica: Dizemos que uma matriz $\boldsymbol{A}_{p\times p}$ é simétrica se $\{a_{ij}\} = \{a_{ji}\}$ para todo i e para todo j, ou seja, $\boldsymbol{A}' = \boldsymbol{A}$.
- Diagonal de uma matriz: A diagonal de uma matriz quadrada $\boldsymbol{A}_{p\times p}$ corresponde ao conjunto de elementos $a_{11}, a_{22}, ..., a_{pp}$.
- Matriz diagonal: Dizemos que a matriz quadrada $\boldsymbol{A}_{p\times p}$ é diagonal se todos os elementos fora da diagonal são iguais a zero:

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & a_{pp} \end{bmatrix}$$

Definição e propriedades básicas

- Matriz identidade: Dizemos que a matriz quadrada $\boldsymbol{I_{p\times p}}$ é uma matriz identidade se ela é uma matriz diagonal com todos os elementos da diagonal iguais a 1:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Definição e propriedades básicas

- Matriz triangular superior: Dizemos que a matriz quadrada $\boldsymbol{A}_{\boldsymbol{p}\times\boldsymbol{p}}$ é uma matriz triangular superior se todos os elementos abaixo da diagonal são iguais a zero:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ 0 & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{pp} \end{bmatrix},$$

- Uma matriz triangular inferior é definida de forma semelhante.

Operações envolvendo matrizes

- A soma de duas matrizes $\boldsymbol{A}_{n\times p}$ e $\boldsymbol{B}_{n\times p}$ **de mesma dimensão** é a matriz resultante das somas dos elementos nas posições correspondentes:

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1p}+b_{1p} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2p}+b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{np}+b_{np} \end{bmatrix}$$

Operações envolvendo matrizes

- A diferença de duas matrizes $\boldsymbol{A}_{n\times p}$ e $\boldsymbol{B}_{n\times p}$ **de mesma dimensão** é a matriz resultante das diferenças dos elementos nas posições correspondentes:

$$\mathbf{A}-\mathbf{B}=\begin{bmatrix} a_{11}-b_{11} & a_{12}-b_{12} & \cdots & a_{1p}-b_{1p} \\ a_{21}-b_{21} & a_{22}-b_{22} & \cdots & a_{2p}-b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}-b_{n1} & a_{n2}-b_{n2} & \cdots & a_{np}-b_{np} \end{bmatrix}$$

Operações envolvendo matrizes

- Sejam $\boldsymbol{A}_{\boldsymbol{n\times k}}$ e $\boldsymbol{B}_{\boldsymbol{k\times p}}$ duas matrizes, tais que o número de linhas da segunda é igual ao número de colunas da primeira. O **produto** $\boldsymbol{AB}$ é definido por:

$$\mathbf{AB} = \begin{bmatrix} \sum_{l=1}^{k} a_{1l}b_{l1} & \sum_{l=1}^{k} a_{1l}b_{l2} & \cdots & \sum_{l=1}^{k} a_{1l}b_{lp} \\ \sum_{l=1}^{k} a_{2l}b_{l1} & \sum_{l=1}^{k} a_{2l}b_{l2} & \cdots & \sum_{l=1}^{k} a_{2l}b_{lp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{l=1}^{k} a_{nl}b_{l1} & \sum_{l=1}^{k} a_{nl}b_{l2} & \cdots & \sum_{l=1}^{k} a_{nl}b_{lp} \end{bmatrix}$$

- Dizemos que uma matriz quadrada $\boldsymbol{Q}$ é **ortogonal** se $\boldsymbol{QQ'} = \boldsymbol{Q'Q} = \boldsymbol{I}$.
- Como consequência, se $\boldsymbol{Q}$ é ortogonal então $\boldsymbol{Q}^{-1} = \boldsymbol{Q'}$.

Operações envolvendo matrizes

- Sejam $\boldsymbol{A}_{\boldsymbol{n}\times\boldsymbol{p}}$ e c uma constante. O produto $c\boldsymbol{A}$ resulta no produto de cada elemento de $\boldsymbol{A}$ por c:

$$\mathbf{cA} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1p} \\ ca_{21} & ca_{22} & \cdots & ca_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{np} \end{bmatrix}$$

Operações envolvendo matrizes

Sejam $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ matrizes com dimensões compatíveis para as operações consideradas. Então:

- $(\boldsymbol{A}')' = \boldsymbol{A}$;
- $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$;
- $(\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$;
- $(\boldsymbol{A} - \boldsymbol{B})' = \boldsymbol{A}' - \boldsymbol{B}'$;
- $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$.

Operações envolvendo matrizes

- $\boldsymbol{AB} \neq \boldsymbol{BA}$, a menos de situações bem específicas;
- $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC}$, valendo o mesmo ao substituir a soma pela diferença;
- $(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC}+\boldsymbol{BC}$, valendo o mesmo ao substituir a soma pela diferença;
- $(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C} \neq \boldsymbol{CA}+\boldsymbol{CB}$, a menos de situações bem específicas.
- $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$, para qualquer $\mathbf{A}$.

Definição e propriedades básicas

- O **traço** de uma matriz de uma matriz $\boldsymbol{A}_{p\times p}$, denotado por $tr(\boldsymbol{A})$, corresponde à soma dos elementos da diagonal de $\boldsymbol{A}$:

$$tr(\boldsymbol{A}) = \sum_{i=1}^{p} a_{ii} \tag{1}$$

Sejam $\boldsymbol{A}$ e $\boldsymbol{B}$ matrizes quadradas. Então:

- $tr(\boldsymbol{A}+\boldsymbol{B}) = tr(\boldsymbol{A}) + tr(\boldsymbol{B})$;
- $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$;
- $tr(a\boldsymbol{A}+b\boldsymbol{B}) = a \times tr(\boldsymbol{A}) + b \times tr(\boldsymbol{B})$.

Combinações lineares e formas quadráticas

- Para um conjunto de constantes $a_1, a_2, ..., a_p$, o vetor $\boldsymbol{y} = a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + ... + a_p\boldsymbol{x}_p$ é uma **combinação linear** dos vetores $\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_p$.
- O conjunto de vetores $\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_p$ é dito **linearmente dependente** se há um conjunto de constantes $a_1, a_2, ..., a_p$, nem todas nulas, tal que:

$$a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + ... + a_p\boldsymbol{x}_p = \boldsymbol{0}.$$

- Caso contrário os vetores são linearmente independentes.

Combinações lineares e formas quadráticas

- Uma **forma quadrática**, definida a partir de uma matriz simétrica $\boldsymbol{A}$ $p \times p$, é definida como:

$$Q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{p}\sum_{i=1}^{p} a_{ik}x_i x_k,$$

para $\boldsymbol{x} \neq \mathbf{0}$ definido em R^p.

- Classificamos a matriz $\boldsymbol{A}$, e a consequente forma quadrática $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$, como **positiva definida** se $Q(\boldsymbol{x}) > 0$ para qualquer $\boldsymbol{x} \neq \mathbf{0}$.

Matriz inversa

- Matriz inversa: Considere uma matriz $\boldsymbol{A}_{p\times p}$. Caso exista uma matriz $\boldsymbol{A}^{-1}_{p\times p}$ tal que

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}, \qquad (2)$$

dizemos que $\boldsymbol{A}^{-1}$ é a **matriz inversa** de $\boldsymbol{A}$.

- Quando uma matriz possui uma matriz inversa, dizemos que ela é **não-singular**. Caso contrário, ela é classificada como **singular**.

Matriz inversa

- A condição fundamental para que uma matriz tenha inversa é que suas colunas sejam linearmente independentes (matriz de rank completo).

- O **rank** de uma matriz $\boldsymbol{A}_{n\times p}$, denotado por $rank(\boldsymbol{A})$, é definido como o número de linhas (ou colunas) linearmente independentes de $\boldsymbol{A}$.

- Dizemos que a matriz quadrada $\boldsymbol{A}_{p\times p}$ tem rank completo se $rank(\boldsymbol{A}) = p$, configurando uma matriz não singular.

- Uma matriz quadrada $\boldsymbol{A}$ admite uma inversa se, e só se, $det(\boldsymbol{A}) \neq 0$.

Matriz inversa

- A inversa de uma matriz diagonal é dada pela matriz diagonal composta pelos inversos dos elementosa matriz original:

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{pp} \end{bmatrix} ; \mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{a_{pp}} \end{bmatrix}$$

Matriz inversa

- $\boldsymbol{A}$ e $\boldsymbol{B}$ não singulares $(p \times p)$, $\boldsymbol{(AB)}^{-1} = \boldsymbol{(B)}^{-1}\boldsymbol{(A)}^{-1}$;

- Para c uma constante real diferente de zero, $(c\boldsymbol{B})^{-1} = c^{-1}\boldsymbol{(A)}^{-1}$;

- $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$;

- Se $rank(\boldsymbol{A}) = p$ então $\boldsymbol{A}^{-1}$ existe;

- Se $\boldsymbol{A}$ é ortogonal, então $\boldsymbol{A}^{-1}$ existe, além do que $\boldsymbol{A}^{-1} = \boldsymbol{A}'$;

- Se $\boldsymbol{B}$ é não singular, $\boldsymbol{AB} = \boldsymbol{CB}$ implica $\boldsymbol{A} = \boldsymbol{C}$.

Matriz idempotente

- Uma matriz $\boldsymbol{A}_{p\times p}$ é chamada **idempotente** se:

$$\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}$$

- Se $\boldsymbol{A}$ é também simétrica, então $\boldsymbol{A}$ é chamada **simétrica idempotente**;

- Se $\boldsymbol{A}$ é simétrica idempotente, então $\boldsymbol{I} - \boldsymbol{A}$ também é simétrica idempotente.

- O rank de uma matriz $\boldsymbol{A}$ idempotente é igual ao traço de $\boldsymbol{A}$.

Resultados sobre partição de matrizes

- Seja uma matrix $\boldsymbol{X}$ particionada tal que:

$$\boldsymbol{X} = [\boldsymbol{X}_1 \boldsymbol{X}_2]$$

.

- Então, valem os seguintes resultados:

$$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{X}'\boldsymbol{X_1} = \boldsymbol{X_1} \text{ e } \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{X}'\boldsymbol{X_2} = \boldsymbol{X_2}$$

$$\boldsymbol{X_1'}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{X}' = \boldsymbol{X_1'} \text{ e } \boldsymbol{X_2'}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{X}' = \boldsymbol{X_2'}$$

Derivadas matriciais

- Seja $\boldsymbol{A}$ uma matriz de constantes $k \times k$, $\boldsymbol{a}$ um vetor de constantes $k \times 1$ e $\boldsymbol{y}$ um vetor $k \times 1$ de variáveis.

- Se $\boldsymbol{z} = \boldsymbol{a}'\boldsymbol{y}$, então:

$$\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{y}} = \frac{\partial \boldsymbol{a}'\boldsymbol{y}}{\partial \boldsymbol{y}} = \boldsymbol{a}$$

- Se $\boldsymbol{z} = \boldsymbol{y}'\boldsymbol{y}$, então:

$$\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{y}} = \frac{\partial \boldsymbol{y}'\boldsymbol{y}}{\partial \boldsymbol{y}} = 2\boldsymbol{y}$$

Derivadas matriciais

- Se $\boldsymbol{z} = \boldsymbol{a}'\boldsymbol{A}\boldsymbol{y}$, então:

$$\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{y}} = \frac{\partial \boldsymbol{a}'\boldsymbol{A}\boldsymbol{y}}{\partial \boldsymbol{y}} = \boldsymbol{A}'\boldsymbol{a}$$

- Se $\boldsymbol{z} = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$, então:

$$\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{y}} = \frac{\partial \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}}{\partial \boldsymbol{y}} = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{A}'\boldsymbol{y}$$

- Se $\boldsymbol{z} = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ e A é simétrica, então:

$$\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{y}} = 2\boldsymbol{A}\boldsymbol{y}$$

Esperanças

- Seja $\boldsymbol{A}$ uma matriz de constantes $k \times k$, $\boldsymbol{a}$ um vetor de constantes $k \times 1$ e $\boldsymbol{y}$ um vetor aleatório $k \times 1$ com vetor de médias $\boldsymbol{\mu}$ e matriz de variâncias e covariâncias $\boldsymbol{\Sigma}$.
  Então:

1. $E(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{a}'\boldsymbol{\mu}$;
2. $E(\boldsymbol{Ay}) = \boldsymbol{A\mu}$;
3. $Var(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{a}'\boldsymbol{\Sigma a}$;
4. $Var(\boldsymbol{Ay}) = \boldsymbol{A\Sigma A}'$; Nota: Se $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$ então $Var(\boldsymbol{Ay}) = \sigma^2\boldsymbol{AA}'$.
5. $E(\boldsymbol{y}'\boldsymbol{Ay}) = \text{tr}(\boldsymbol{A\Sigma}) + \boldsymbol{\mu}'\boldsymbol{A\mu}$; Nota: Se $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$, então $E(\boldsymbol{y}'\boldsymbol{Ay}) = \sigma^2\text{tr}(\boldsymbol{A}) + \boldsymbol{\mu}'\boldsymbol{A\mu}$.