

Álgebra Linear Aplicada a Ciência de Dados

Professor Dr. Marcelo Guterres
Supervisor acadêmico Prof. Dr. Ernesto C. Marujo

25 de junho de 2018

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

MODELOS DE REGRESSÃO

Prática R

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- Calcular a reta de quadrados mínimos:

Prática R

MODELOS DE REGRESSÃO

- Prática R

- Prática R

- Prática R

- Forma Matricial

- Forma Matricial

- Forma Matricial

- Estimativa do vetor β

- Estimativa do vetor β

- Exemplo RLM

- Prática R

- Prática Python

- Calcular a reta de quadrados mínimos:

| | | | | | | | | |
|---|-----|-----|-----|-----|------|------|------|------|
| x | 1.2 | 2.5 | 3.0 | 4.1 | 6.2 | 7.1 | 8.8 | 9.5 |
| y | 6.8 | 6.1 | 9.9 | 9.7 | 12.1 | 17.9 | 18.0 | 21.5 |

- Pelo procedimento demonstrado anteriormente;

Prática R

MODELOS DE REGRESSÃO

- Prática R

- Prática R

- Prática R

- Forma Matricial

- Forma Matricial

- Forma Matricial

- Estimativa do vetor β

- Estimativa do vetor β

- Exemplo RLM

- Prática R

- Prática Python

- Calcular a reta de quadrados mínimos:

| | | | | | | | | |
|---|-----|-----|-----|-----|------|------|------|------|
| x | 1.2 | 2.5 | 3.0 | 4.1 | 6.2 | 7.1 | 8.8 | 9.5 |
| y | 6.8 | 6.1 | 9.9 | 9.7 | 12.1 | 17.9 | 18.0 | 21.5 |

- Pelo procedimento demonstrado anteriormente;
- Realizar o procedimento usando a função `lm` do R;

Prática R

MODELOS DE REGRESSÃO

- Prática R

- Prática R

- Prática R

- Forma Matricial

- Forma Matricial

- Forma Matricial

- Estimativa do vetor β

- Estimativa do vetor β

- Exemplo RLM

- Prática R

- Prática Python

- Calcular a reta de quadrados mínimos:

| | | | | | | | | |
|---|-----|-----|-----|-----|------|------|------|------|
| x | 1.2 | 2.5 | 3.0 | 4.1 | 6.2 | 7.1 | 8.8 | 9.5 |
| y | 6.8 | 6.1 | 9.9 | 9.7 | 12.1 | 17.9 | 18.0 | 21.5 |

- Pelo procedimento demonstrado anteriormente;
- Realizar o procedimento usando a função `lm` do R;

Prática R

MODELOS DE REGRESSÃO

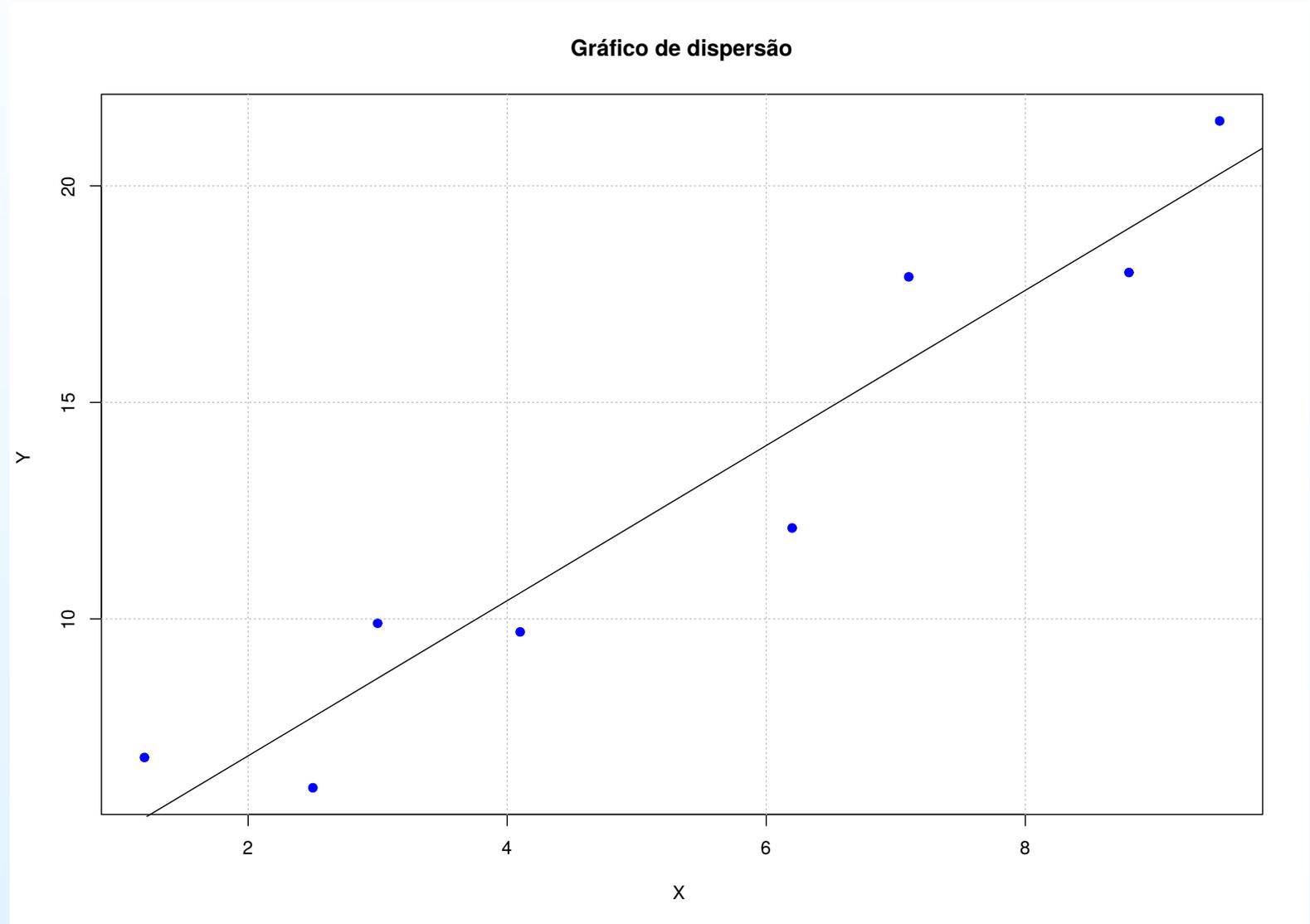
- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

```
rm(list = ls())
x<-c(1.2,2.5,3.0,4.1,6.2,7.1,8.8,9.5)
y<-c(6.8,6.1,9.9,9.7,12.1,17.9,18.0,21.5)
df<-data.frame(y,x)
plot(df$x,df$y,main="Gráfico de dispersão",xlab="X ",
ylab="Y ", pch=19,col="blue")
grid(nx = NULL, ny = NULL, col = "grey", lty = "dotted")
rls<-lm(y~x,data = df)
abline(rls)
summary(rls)
```

Prática R

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python



Forma Matricial

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- **Forma Matricial**
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- Digamos que queremos ajustar o modelo

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$ aos pontos $(y_1, x_{11}, x_{12}, \dots, x_{1p})$, $(y_2, x_{21}, x_{22}, \dots, x_{2p}), \dots, (y_n, x_{n1}, x_{n2}, \dots, x_{np})$, determinados experimentalmente.

Forma Matricial

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- **Forma Matricial**
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- Digamos que queremos ajustar o modelo

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$ aos pontos $(y_1, x_{11}, x_{12}, \dots, x_{1p})$, $(y_2, x_{21}, x_{22}, \dots, x_{2p})$, ..., $(y_n, x_{n1}, x_{n2}, \dots, x_{np})$, determinados experimentalmente.

- Se estes pontos de dados fossem coplanares, o modelo passaria por todos os n pontos e, então os coeficientes $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ desconhecidos satisfariam:

Forma Matricial

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- **Forma Matricial**
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- Digamos que queremos ajustar o modelo

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$ aos pontos $(y_1, x_{11}, x_{12}, \dots, x_{1p})$, $(y_2, x_{21}, x_{22}, \dots, x_{2p})$, ..., $(y_n, x_{n1}, x_{n2}, \dots, x_{np})$, determinados experimentalmente.

- Se estes pontos de dados fossem coplanares, o modelo passaria por todos os n pontos e, então os coeficientes $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ desconhecidos satisfariam:

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_p x_{1p}$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_p x_{2p}$$

⋮

$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_p x_{np}$$

Forma Matricial

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

Nós podemos escrever este sistema em forma matricial como:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{1p} \\ 1 & x_{21} & x_{22} & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

$$y = X\beta + \epsilon$$

Forma Matricial

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- **Forma Matricial**
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- y : vetor ($n \times 1$) contendo as n observações da variável resposta;

Forma Matricial

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- **Forma Matricial**
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- y : vetor ($n \times 1$) contendo as n observações da variável resposta;
- X : matriz ($n \times (p + 1)$), $n \geq p + 1$, contendo os n valores das p variáveis explicativas, além da primeira coluna de 1's relativa à β_0 ;

Forma Matricial

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- **Forma Matricial**
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- y : vetor $(n \times 1)$ contendo as n observações da variável resposta;
- X : matriz $(n \times (p + 1))$, $n \geq p + 1$, contendo os n valores das p variáveis explicativas, além da primeira coluna de 1's relativa à β_0 ;
- β :vetor $((p + 1) \times 1)$ dos parâmetros a serem estimados e;

Forma Matricial

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- **Forma Matricial**
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- y : vetor $(n \times 1)$ contendo as n observações da variável resposta;
- X : matriz $(n \times (p + 1))$, $n \geq p + 1$, contendo os n valores das p variáveis explicativas, além da primeira coluna de 1's relativa à β_0 ;
- β :vetor $((p + 1) \times 1)$ dos parâmetros a serem estimados e;
- ϵ : vetor $(n \times 1)$ dos erros aleatórios;

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{1p} \\ 1 & x_{21} & x_{22} & x_{2p} \\ & & \vdots & \\ 1 & x_{n1} & x_{n2} & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Estimativa do vetor β

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

Se os pontos de dados não são coplanares, é impossível encontrar coeficientes β_s que satisfaçam $y = X\beta$ exatamente, ou seja, o sistema é inconsistente. Neste caso, vamos procurar uma solução, do que chamamos mínimos quadrados.

$$\beta^* = \begin{bmatrix} \beta_0^* \\ \beta_1^* \\ \vdots \\ \beta_n^* \end{bmatrix}$$

a equação

$$y = \beta_0^* + \beta_1^* x_1 + \beta_2^* x_2 + \dots + \beta_p^* x_p$$

é chamada de modelo de regressão linear múltipla dos dados ou um ajuste linear de mínimos quadrados aos dados se os coeficientes provém de uma solução de mínimos quadrados.

Estimativa do vetor β

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

- Minimizar a função:

$$f(\beta) = \|y - X\beta\|^2 = (y - X\beta)^T (y - X\beta)$$

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

Exemplo RLM

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

| i | y Consumo | x_1 Renda | x_2 Taxa de Juros |
|-----|----------------|----------------|------------------------|
| 1 | 122 | 139 | 11,5% |
| 2 | 114 | 126 | 12,0% |
| 3 | 86 | 90 | 10,5% |
| 4 | 134 | 144 | 9,0% |
| 5 | 146 | 163 | 10,0% |
| 6 | 107 | 136 | 12,0% |
| 7 | 68 | 61 | 10,5% |
| 8 | 117 | 62 | 8,0% |
| 9 | 71 | 41 | 10,0% |
| 10 | 98 | 120 | 11,5% |

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

Prática R

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python

```
rm(list = ls())
y<-c(122,114,86,134,146,107,68,117,71,98)
x1<-c(139,126,90,144,163,136,61,62,41,120)
x2<-c(11.5,12,10.5,9,10,12,10.5,8,10,11.5)
x2<-x2/100 v<-rep(1,length(y))
X<-cbind(v,x1,x2)
Xt<-t(X)
A<-solve(Xt*%*%X)
beta<-A*%*%Xt*%*%y
print(beta)
```

Prática Python

MODELOS DE REGRESSÃO

- Prática R
- Prática R
- Prática R
- Forma Matricial
- Forma Matricial
- Forma Matricial
- Estimativa do vetor β
- Estimativa do vetor β
- Exemplo RLM
- Prática R
- Prática Python