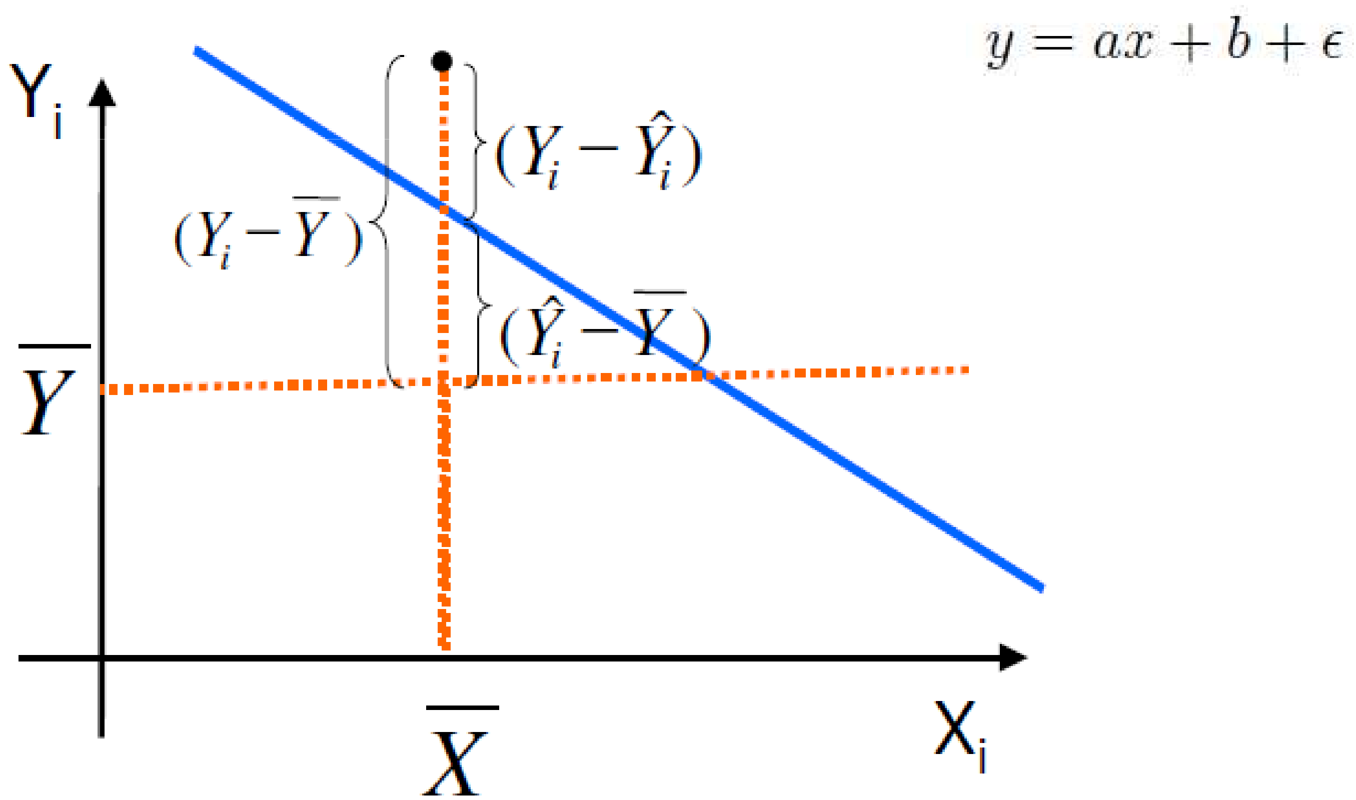
Modelos Lineares
Regressão Linear Simples

Professora Ariane Ferreira

Instituto Politécnico
IPRJ
Campus Regional da UERJ
Nova Friburgo - RJ

Decomposição dos residuos



Somas Quadradas

Soma dos Quadrados totais (SST): $$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Decompondo o termo: $$y_i - \bar{y} = (y_i - \hat{y}_i) - (\bar{y} - \hat{y}_i)$$

SST fica:

$$SST = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{i=1}^{n} (\bar{y} - \hat{y}_i)^2 - 2\sum_{i=1}^{n} (y_i - \hat{y}_i)(\bar{y} - \hat{y}_i)$$

Somas Quadradas

Usando os residuos: $y_i - \hat{y}_i = e_i$

E a propriedade de ortogonalidade: $\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} e_i x_i = \sum_{i=1}^{n} e_i \hat{y}_i = 0$

Temos:

$$
\begin{aligned}
SST &= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{i=1}^{n} (\bar{y} - \hat{y}_i)^2 - 2\sum_{i=1}^{n} e_i(\bar{y} - \hat{y}_i) \\
&= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{i=1}^{n} (\bar{y} - \hat{y}_i)^2 - 2\bar{y}\underbrace{\sum_{i=1}^{n} e_i}_{0} - 2\underbrace{\sum_{i=1}^{n} e_i\hat{y}_i}_{0}
\end{aligned}
$$

Resultando em:

$$SST = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

Somas Quadradas

Soma dos Quadrados dos Erros (SSE): $$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Soma dos Quadrados da Regressão: $$SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

Temos então: $SST = SSE + SSR$

Estimação de σ^2

Precisamos estimar σ^2 para calcularmos $V(\hat{a})$ e $V(\hat{b})$.

Usaremos a soma quadrada:

$$\sum_{i=1}^{n}(y_i - ax_i - b)^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i - ax_i - b + \hat{y}_i)^2$$

Trocando $\hat{y}_i$ por $\hat{a}x_i + \hat{b}$.

$$\sum_{i=1}^{n}(y_i - ax_i - b)^2 = \sum_{i=1}^{n}\{y_i - \hat{y}_i + (\hat{a} - a)x_i + (\hat{b} - b)\}^2$$

Estimação de σ^2

Desenvolvendo a soma quadrada e sabendo que

Usaremos a soma quadrada: $y_i - \hat{y}_i = e_i$

Temos:

$$\sum_{i=1}^{n}(y_i - ax_i - b)^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + (\hat{a} - a)^2 \sum_{i=1}^{n} x_i^2 + n(\hat{b} - b)^2$$

$$+2(\hat{a} - a) \underbrace{\sum_{i=1}^{n} e_i x_i}_{0} + 2(\hat{b} - b) \underbrace{\sum_{i=1}^{n} e_i}_{0} + (\hat{a} - a)(\hat{b} - b) \sum_{i=1}^{n} x_i$$

Propriedade de ortogonalidade dos residuos

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} e_i x_i = \sum_{i=1}^{n} e_i \hat{y}_i = 0$$

Estimação de σ^2

Temos:

$$\sum_{i=1}^{n}(y_i-ax_i-b)^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+(\hat{a}-a)^2\sum_{i=1}^{n}x_i^2+n(\hat{b}-b)^2+2n(\hat{a}-a)(\hat{b}-b)\bar{x}$$

Trocamos na expressão acima $\sum_{i=1}^{n}x_i^2$ por $\sum_{i=1}^{n}(x_i-\bar{x})^2+n\bar{x}^2=S_{xx}+n\bar{x}^2$

Temos:

$$\begin{aligned}\sum_{i=1}^{n}(y_i-ax_i-b)^2 &= \sum_{i=1}^{n}(y_i-\hat{y}_i)^2+(\hat{a}-a)^2S_{xx}\\ &+n\bar{x}^2(\hat{a}-a)^2+n(\hat{b}-b)^2+2n(\hat{a}-a)(\hat{b}-b)\bar{x}\end{aligned}$$

$$\sum_{i=1}^{n}(y_i-ax_i-b)^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+(\hat{a}-a)^2S_{xx}+n\{(\hat{a}-a)\bar{x}+(\hat{b}-b)\}^2$$

Estimação de σ^2

Usando: $\hat{a}\bar{x} + \hat{b} = \bar{y},$

$$\sum_{i=1}^{n}(y_i - ax_i - b)^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + (\hat{a} - a)^2 S_{xx} + n\{\bar{y} - (a\bar{x} + b)\}^2$$

Dividindo ambos os lados por σ^2 :

$$\underbrace{\sum_{i=1}^{n}\left(\frac{y_i - ax_i - b}{\sigma}\right)^2}_{f_{\chi^2(n)}} = \sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i}{\sigma}\right)^2 + \underbrace{\left(\frac{(\hat{a} - a)\sqrt{S_{xx}}}{\sigma}\right)^2}_{f_{\chi^2(1)}} + \underbrace{\left(\frac{\{\bar{y} - (a\bar{x} + b)\}\sqrt{n}}{\sigma}\right)^2}_{f_{\chi^2(1)}}$$

v.a. à n graus de liberdade

v.a. à 1 grau de liberdade

v.a. à 1 grau de liberdade

Estimação de σ^2

Então: $$\frac{SSE}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{y_i - \hat{y}_i}{\sigma} \right)^2 \sim f_{\chi^2}(n-2)$$

Calculamos a esperança matematica: $$E(SSE) = (n-2)\sigma^2$$

Em conclusão o estimador não tendencioso de σ^2 é:

$$\widehat{\sigma^2} = \frac{SSE}{n-2} \stackrel{\text{def}}{=} MSE$$

MSE= Média quadrada dos erros (residuos)