



ME613 - Análise de Regressão

Parte 5

Samara F. Kiihl - IMECC - UNICAMP

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMP.github.io/nulas/slides/parte05/parte05.html#1

1/25

Inferência simultânea

Suponha que tenhamos interesse tanto em fazer inferência sobre β_0 quanto para β_1 .

Vimos como construir um intervalo de, por exemplo, 95% de confiança para $\hat{\beta}_0$ e um intervalo de 95% de confiança para $\hat{\beta}_1$.

Mas, se utilizarmos esses intervalos de confiança individuais, não temos 95% de confiança para β_0 e β_1 conjuntamente.

Por exemplo, se essas inferências fossem independentes, teríamos $0.95^2 = 0.9025$ de confiança.

No caso, já vimos que $\hat{\beta}_0$ e $\hat{\beta}_1$ não são independentes, o que dificulta a determinação do verdadeiro nível de confiança.

3/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMP.github.io/nulas/slides/parte05/parte05.html#1

3/25

Intervalo de confiança simultâneo

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMP.github.io/nulas/slides/parte05/parte05.html#1

2/25

Intervalo de confiança conjunto de Bonferroni

Relembrando:

1. Intervalo de confiança $100 \times (1 - \alpha)\%$ para β_0 :

$$\hat{\beta}_0 \pm t_{n-2, \alpha/2} \sqrt{\widehat{Var}(\hat{\beta}_0)}$$

2. Intervalo de confiança $100 \times (1 - \alpha)\%$ para β_1 :

$$\hat{\beta}_1 \pm t_{n-2, \alpha/2} \sqrt{\widehat{Var}(\hat{\beta}_1)}$$

4/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMP.github.io/nulas/slides/parte05/parte05.html#1

4/25

Intervalo de confiança conjunto de Bonferroni

Seja A_1 o evento de que o intervalo 1 não contenha β_0 . $P(A_1) = \alpha$.

Seja A_2 o evento de que o intervalo 2 não contenha β_1 . $P(A_2) = \alpha$.

Qual a probabilidade de que ambos intervalos estejam corretos, isto é, $P(A_1^c \cap A_2^c)$?

$$P(A_1^c \cap A_2^c) = 1 - P(A_1 \cup A_2) = 1 - P(A_1) - P(A_2) + P(A_1 \cap A_2)$$

Desigualdade de Bonferroni:

$$\begin{aligned} P(A_1^c \cap A_2^c) &\geq 1 - P(A_1) - P(A_2) \\ &\geq 1 - 2\alpha \end{aligned}$$

5/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/parte05/parte05.html#1

5/25

Notação Matricial para Regressão

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/parte05/parte05.html#1

7/25

Intervalo de confiança conjunto de Bonferroni

Para obter $100 \times (1 - \alpha)\%$ de confiança para β_0 e β_1 conjuntamente, segundo o procedimento de Bonferroni:

- $\hat{\beta}_0 \pm t_{n-2, \alpha/4} \sqrt{\widehat{Var}(\hat{\beta}_0)}$
- $\hat{\beta}_1 \pm t_{n-2, \alpha/4} \sqrt{\widehat{Var}(\hat{\beta}_1)}$

De maneira que:

$$\begin{aligned} P(A_1^c \cap A_2^c) &\geq 1 - P(A_1) - P(A_2) \\ &\geq 1 - \alpha/2 - \alpha/2 \\ &\geq 1 - \alpha \end{aligned}$$

6/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/parte05/parte05.html#1

6/25

Vetores e Matrizes Aleatórios

Suponha que tenhamos um vetor aleatório \mathbf{Y} em que $n = 3$:

$$\mathbf{Y}_{3 \times 1} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$$

$E(\mathbf{Y})$ é definida como:

$$E(\mathbf{Y})_{3 \times 1} = \begin{pmatrix} E(Y_1) \\ E(Y_2) \\ E(Y_3) \end{pmatrix}$$

Se \mathbf{Y} é uma matriz aleatória de dimensão $n \times p$:

$$E(\mathbf{Y})_{n \times p} = [E(Y_{ij})] \quad i = 1, \dots, n; j = 1, \dots, p$$

8/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/parte05/parte05.html#1

8/25

Matriz de Variância-Covariância de um Vetor Aleatório

Suponha que tenhamos um vetor aleatório \mathbf{Y} em que $n = 3$:

$$\mathbf{Y}_{3 \times 1} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$$

A matriz de variância-covariância de \mathbf{Y} é definida como:

$$Var(\mathbf{Y})_{3 \times 3} = \begin{pmatrix} Var(Y_1) & Cov(Y_1, Y_2) & Cov(Y_1, Y_3) \\ Cov(Y_2, Y_1) & Var(Y_2) & Cov(Y_2, Y_3) \\ Cov(Y_3, Y_1) & Cov(Y_3, Y_2) & Var(Y_3) \end{pmatrix}$$

9/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/part05/part05.html#1

9/25

Propriedades básicas

Seja \mathbf{W} um vetor aleatório obtido pela multiplicação do vetor aleatório \mathbf{Y} e da matriz de constantes \mathbf{A} :

$$\mathbf{W} = \mathbf{A}\mathbf{Y}.$$

Temos as seguintes propriedades:

$$E(\mathbf{A}) = \mathbf{A}$$

$$E(\mathbf{W}) = E(\mathbf{A}\mathbf{Y}) = \mathbf{A}E(\mathbf{Y})$$

$$Var(\mathbf{Y}) = Var(\mathbf{A}\mathbf{Y}) = \mathbf{A}Var(\mathbf{Y})\mathbf{A}^T$$

11/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/part05/part05.html#1

11/25

Matriz de variância-covariância de um vetor aleatório

Em geral:

$$Var(\mathbf{Y}) = E[(\mathbf{Y} - E(\mathbf{Y}))(\mathbf{Y} - E(\mathbf{Y}))^T]$$

$$Var(\mathbf{Y})_{n \times n} = \begin{pmatrix} Var(Y_1) & Cov(Y_1, Y_2) & \dots & Cov(Y_1, Y_n) \\ Cov(Y_2, Y_1) & Var(Y_2) & \dots & Cov(Y_2, Y_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(Y_n, Y_1) & Cov(Y_n, Y_2) & \dots & Var(Y_n) \end{pmatrix}$$

10/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/part05/part05.html#1

10/25

Distribuição Normal Multivariada

$$\mathbf{Y}_{p \times 1} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{pmatrix}$$

$$\boldsymbol{\mu}_{p \times 1} = E(\mathbf{Y})_{p \times 1} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \begin{pmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_p) \end{pmatrix}$$

12/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/part05/part05.html#1

12/25

Distribuição Normal Multivariada

$$\Sigma_{p \times p} = \begin{pmatrix} \text{Var}(Y_1) & \text{Cov}(Y_1, Y_2) & \dots & \text{Cov}(Y_1, Y_p) \\ \text{Cov}(Y_2, Y_1) & \text{Var}(Y_2) & \dots & \text{Cov}(Y_2, Y_p) \\ \vdots & \vdots & & \vdots \\ \text{Cov}(Y_p, Y_1) & \text{Cov}(Y_p, Y_2) & \dots & \text{Var}(Y_p) \end{pmatrix}$$

Função de densidade da normal p -variada:

$$f(\mathbf{Y}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{Y} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{Y} - \boldsymbol{\mu}) \right]$$

13/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMP.github.io/nulas/slides/part05/part05.html#1

13/25

Regressão Linear Simples com notação matricial

$$E(\boldsymbol{\varepsilon})_{n \times 1} = \mathbf{0}_{n \times 1}$$

$$\text{Var}(\boldsymbol{\varepsilon})_{n \times n} = \sigma^2 \mathbf{I}_{n \times n}$$

$$E(\mathbf{Y}) = E(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = \mathbf{X}\boldsymbol{\beta}$$

$$\text{Var}(\mathbf{Y}) = \sigma^2 \mathbf{I}$$

15/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMP.github.io/nulas/slides/part05/part05.html#1

15/25

Regressão Linear Simples com notação matricial

$$Y_i = \beta_0 + \beta_1 X_i, \quad \varepsilon_i \stackrel{iid}{\sim} \mathbf{N}(0, \sigma^2), \quad i = 1, 2, \dots, n$$

$$\mathbf{Y}_{n \times 1} = \mathbf{X}_{n \times 2} \boldsymbol{\beta}_{2 \times 1} + \boldsymbol{\varepsilon}_{n \times 1}, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$$

$$\mathbf{Y}_{n \times 1} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}$$

$$\mathbf{X}_{n \times 1} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix} \quad \boldsymbol{\beta}_{2 \times 1} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \quad \boldsymbol{\varepsilon}_{n \times 1} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

14/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMP.github.io/nulas/slides/part05/part05.html#1

14/25

Mínimos Quadrados

Queremos encontrar $\hat{\boldsymbol{\beta}}$ que minimiza:

$$\begin{aligned} S(\boldsymbol{\beta}) &= \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{Y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{Y}^T \mathbf{Y} - 2\boldsymbol{\beta}^T \mathbf{X}^T \mathbf{Y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \end{aligned}$$

$$\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^T \mathbf{Y} + 2\mathbf{X}^T \mathbf{X}\boldsymbol{\beta}$$

Equação normal: $\mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^T \mathbf{Y}$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

16/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMP.github.io/nulas/slides/part05/part05.html#1

16/25

Mínimos Quadrados

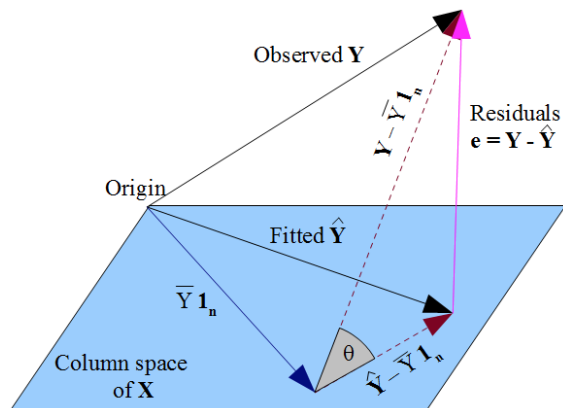
$$\begin{aligned} \text{Var}(\hat{\beta}) &= \text{Var}[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}] \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{Var}(\mathbf{Y}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \sigma^2 \mathbf{I} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \end{aligned}$$

\mathbf{H} é a matriz de projeção ortogonal no espaço coluna de \mathbf{X} .

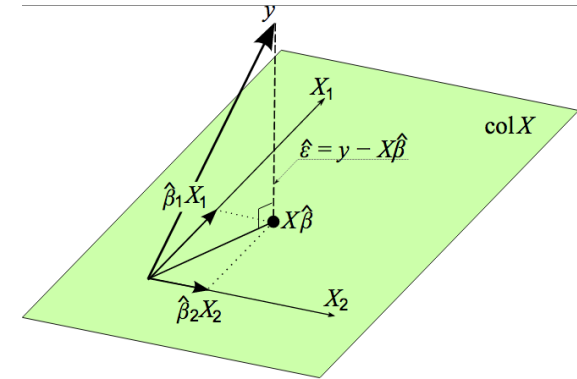
$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = \underbrace{\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T}_{\mathbf{H}} \mathbf{Y}$$

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{H}\mathbf{Y} = (\mathbf{I} - \mathbf{H})\mathbf{Y}$$

Interpretação geométrica



Interpretação geométrica



Interpretação geométrica

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Pitágoras:

$$\|\mathbf{Y} - \bar{Y} \mathbf{1}_n\|^2 = \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 + \|\hat{\mathbf{Y}} - \bar{Y} \mathbf{1}_n\|^2$$

$$R^2 = 1 - \frac{SQE}{SQT} = 1 - \frac{\|\mathbf{Y} - \hat{\mathbf{Y}}\|^2}{\|\mathbf{Y} - \bar{Y} \mathbf{1}_n\|^2} = \frac{SQReg}{SQT} = \frac{\|\hat{\mathbf{Y}} - \bar{Y} \mathbf{1}_n\|^2}{\|\mathbf{Y} - \bar{Y} \mathbf{1}_n\|^2}$$

$$R = \cos(\theta) = \frac{\|\hat{\mathbf{Y}} - \bar{Y} \mathbf{1}_n\|}{\|\mathbf{Y} - \bar{Y} \mathbf{1}_n\|}$$

Exercício

X : temperatura ($^{\circ}F$).

Y : semanas até a deterioração do sabor.

```
##      x      y
## 1  8    7.8
## 2  4    9.0
## 3  0   10.2
## 4 -4   11.0
## 5 -8   11.7
```

Utilizando a notação matricial para o modelo de regressão, obtenha:

- $\mathbf{Y}^T \mathbf{Y}$, $\mathbf{X}^T \mathbf{X}$ e $\mathbf{X}^T \mathbf{Y}$, $(\mathbf{X}^T \mathbf{X})^{-1}$.
- $\hat{\boldsymbol{\beta}}$
- vetor de resíduos
- matriz de variância-covariância estimada para $\hat{\boldsymbol{\beta}}$.

21/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/parte05/parte05.html#1

21/25

ANOVA

$$\underbrace{\sum_{i=1}^n (Y_i - \bar{Y})^2}_{SQT} = \underbrace{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}_{SQReg} + \underbrace{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}_{SQE}$$

$$SQT = \sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n Y_i^2 - \frac{(\sum_{i=1}^n Y_i)^2}{n}$$

$$SQT = \mathbf{Y}^T \mathbf{Y} - \frac{1}{n} \mathbf{Y}^T \mathbf{1} \mathbf{1}^T \mathbf{Y} = \mathbf{Y}^T \left[\mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \right] \mathbf{Y}$$

22/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/parte05/parte05.html#1

22/25

ANOVA

$$\begin{aligned} SQE &= \mathbf{e}^T \mathbf{e} = (\mathbf{Y} - \hat{\mathbf{Y}})^T (\mathbf{Y} - \hat{\mathbf{Y}}) \\ &= (\mathbf{Y} - \mathbf{H}\mathbf{Y})^T (\mathbf{Y} - \mathbf{H}\mathbf{Y}) \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{H}\mathbf{Y} - \mathbf{Y}^T \mathbf{H}^T \mathbf{Y} + \mathbf{Y}^T \mathbf{H}^T \mathbf{H}\mathbf{Y} \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{H}\mathbf{Y} \\ &= \mathbf{Y}^T (\mathbf{I} - \mathbf{H}) \mathbf{Y} \end{aligned}$$

Obs: $\mathbf{H}^T = \mathbf{H}$ e $\mathbf{H}^T \mathbf{H} = \mathbf{H}$

23/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/parte05/parte05.html#1

23/25

ANOVA

$$SQReg = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = \sum_{i=1}^n \hat{Y}_i^2 - \frac{(\sum_{i=1}^n Y_i)^2}{n}$$

$$\begin{aligned} SQReg &= (\mathbf{H}\mathbf{Y})^T \mathbf{H}\mathbf{Y} - \frac{1}{n} \mathbf{Y}^T \mathbf{1} \mathbf{1}^T \mathbf{Y} \\ &= \mathbf{Y}^T \mathbf{H}\mathbf{Y} - \frac{1}{n} \mathbf{Y}^T \mathbf{1} \mathbf{1}^T \mathbf{Y} \\ &= \mathbf{Y}^T \left[\mathbf{H} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \right] \mathbf{Y} \end{aligned}$$

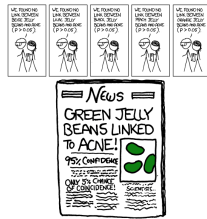
24/25

file:///Users/amac/Documents/GitHub/ME613-UNICAMP/ME613-UNICAMPgithub.io/nulas/slides/parte05/parte05.html#1

24/25

Leitura

- Applied Linear Statistical Models: 4.1-4.3, Capítulo 5.
- Draper & Smith - [Applied Regression Analysis](#): Capítulos 4, 5 e 20.
- Faraway - [Linear Models with R](#): Capítulo 2.
- [The Matrix Cookbook](#)



<https://xkcd.com/882/>