Forma matricial do MRLS

Prof. Juvêncio Santos Nobre

Departamento de Estatística e Matemática Aplicada

Universidade Federal do Ceará-Brasil

http://www.dema.ufc.br/∼juvencio

DEMA-UFC

Capital do **Ceará**, setembro de 2022

Conteúdo

MRLS

Considere o MRLS com forma funcional

$$y_i = \beta_0 + \beta_1 x_i + e_i,\ i = 1, \ldots, n. \tag{1}$$

organizando lexicograficamente, temos que

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_1 + e_1 \\ y_2 &= \beta_0 + \beta_1 x_2 + e_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_n + e_n. \end{aligned}$$

MRLS

Perceba que podemos reescrever a estrutura do modelo da seguinte forma

$$
\begin{aligned}
\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} &= \begin{pmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \\
&= \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}. \qquad (2)
\end{aligned}
$$

MRLS

Denotando

- $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$: vetor $(n \times 1)$ de variáveis respostas.
- $\boldsymbol{X}$: matriz $(n \times 2)$ de especificação do modelo, dada por

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

- $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$: vetor (2×1) de parâmetros de regressão.
- $\boldsymbol{e} = (e_1, \ldots, e_n)^\top$: vetor $(n \times 1)$ da fonte de variação/erro de medida.
- Então, podemos reescrever o MRLS (1) da seguinte forma

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}.$$

MRLS

Denotando

- $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$: vetor $(n \times 1)$ de variáveis respostas.
- $\boldsymbol{X}$: matriz $(n \times 2)$ de especificação do modelo, dada por

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

- $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$: vetor (2×1) de parâmetros de regressão.
- $\boldsymbol{e} = (e_1, \ldots, e_n)^\top$: vetor $(n \times 1)$ da fonte de variação/erro de medida.
- Então, podemos reescrever o MRLS (1) da seguinte forma

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}.$$

MRLS

Denotando

- $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$: vetor $(n \times 1)$ de variáveis respostas.
- $\boldsymbol{X}$: matriz $(n \times 2)$ de especificação do modelo, dada por

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

- $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$: vetor (2×1) de parâmetros de regressão.
- $\boldsymbol{e} = (e_1, \ldots, e_n)^\top$: vetor $(n \times 1)$ da fonte de variação/erro de medida.
- Então, podemos reescrever o MRLS (1) da seguinte forma

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}.$$

MRLS

Denotando

- $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$: vetor $(n \times 1)$ de variáveis respostas.
- $\boldsymbol{X}$: matriz $(n \times 2)$ de especificação do modelo, dada por

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

- $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$: vetor (2×1) de parâmetros de regressão.
- $\boldsymbol{e} = (e_1, \ldots, e_n)^\top$: vetor $(n \times 1)$ da fonte de variação/erro de medida.
- Então, podemos reescrever o MRLS (1) da seguinte forma

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}.$$

MRLS

Denotando

- $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$: vetor $(n \times 1)$ de variáveis respostas.
- $\boldsymbol{X}$: matriz $(n \times 2)$ de especificação do modelo, dada por

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

- $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$: vetor (2×1) de parâmetros de regressão.
- $\boldsymbol{e} = (e_1, \ldots, e_n)^\top$: vetor $(n \times 1)$ da fonte de variação/erro de medida.
- Então, podemos reescrever o MRLS (1) da seguinte forma

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}.$$

MRLS

- Em geral, para ajustar o MRLS assume-se que

$$\boldsymbol{e} \sim (\mathbf{0}, \sigma^2 \boldsymbol{I}_n).$$

- Tal suposição implica que $e_1, \ldots, e_n$ representa uma sequência de variáveis aleatórias de média zero, homoscedásticos com variância σ^2 e não-correlacionados.
- Para se realizar inferências de segunda ordem (construção de intervalos de confiança e testes de hipóteses, por exemplo) é comum assumir normalidade, i.e.,

$$\boldsymbol{e} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \boldsymbol{I}_n).$$

MRLS

- Em geral, para ajustar o MRLS assume-se que

$$\boldsymbol{e} \sim (\mathbf{0}, \sigma^2 \boldsymbol{I}_n).$$

- Tal suposição implica que $e_1, \ldots, e_n$ representa uma sequência de variáveis aleatórias de média zero, homoscedásticos com variância σ^2 e não-correlacionados.
- Para se realizar inferências de segunda ordem (construção de intervalos de confiança e testes de hipóteses, por exemplo) é comum assumir normalidade, i.e.,

$$\boldsymbol{e} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \boldsymbol{I}_n).$$

MRLS

- Em geral, para ajustar o MRLS assume-se que

$$\boldsymbol{e} \sim (\mathbf{0}, \sigma^2 \boldsymbol{I}_n).$$

- Tal suposição implica que $e_1, \ldots, e_n$ representa uma sequência de variáveis aleatórias de média zero, homoscedásticos com variância σ^2 e não-correlacionados.
- Para se realizar inferências de segunda ordem (construção de intervalos de confiança e testes de hipóteses, por exemplo) é comum assumir normalidade, i.e.,

$$\boldsymbol{e} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \boldsymbol{I}_n).$$

Exercícios (entregar próxima aula)

Exercício 1: Faça um breve ensaio sobre distribuição normal multivariada e suas propriedades apresentadas de forma matricial. Em especial, apresente as 3 definições equivalentes da distribuição normal multivariada (algumas sem a necessidade da fdp).

Exercício 2: Usando um software de sua preferência, plote a função densidade de probabilidade e a respectiva curva de nível (mapa de contorno) da $\mathcal{N}_2(0, 0, 1, 1, \rho)$, para $\rho = 0, \pm 0.1, \pm 0.5, \pm 0.9$. Interprete os gráficos.

Exercícios (entregar próxima aula)

Exercício 3: Faça um breve ensaio sobre distribuição t-Student multivariada e suas propriedades apresentadas de forma matricial.

Exercício 4: Usando um software de sua preferência, plote a função densidade de probabilidade e a respectiva curva de nível (mapa de contorno) da $t_{\nu}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, para $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \rho \boldsymbol{I}_2$, $\rho = 0, \pm 0.1, \pm 0.5, \pm 0.9$ e $\nu = 1, 2, 10, 30$ e 100. Interprete os gráficos.

Exercício 5: Apresente um resumo sobre:

i) Distribuições χ^2 e F não centrais.

ii) Distribuições de formas lineares e quadráticas (sob suposição de normalidade).

Método de Mínimos Quadrados

- Sob as suposições usuais associadas ao MRLS é possível obter o Estimador de Mínimos Quadrados (EMQ) de $\boldsymbol{\beta}$ através da minimização da seguinte função objetivo

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2. \tag{3}$$

- Lembrando que qualquer soma de quadrados pode ser reescrita como (Lista 0) uma forma quadrática

$$\sum_{i=1}^{n} a_i^2 = \boldsymbol{a}^\top \boldsymbol{a} = \|\boldsymbol{a}\|^2$$

com $\boldsymbol{a} = (a_1, \ldots, a_n)^\top$.

- Então podemos reescrever (3) como

$$Q(\beta) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = (\boldsymbol{y} - \boldsymbol{X}\beta)^\top(\boldsymbol{y} - \boldsymbol{X}\beta) = \|\boldsymbol{y} - \boldsymbol{X}\beta\|^2,$$

i.e., $Q(\beta)$ é uma forma quadrática.

Método de Mínimos Quadrados

- Sob as suposições usuais associadas ao MRLS é possível obter o Estimador de Mínimos Quadrados (EMQ) de $\boldsymbol{\beta}$ através da minimização da seguinte função objetivo

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2. \tag{3}$$

- Lembrando que qualquer soma de quadrados pode ser reescrita como (Lista 0) uma forma quadrática

$$\sum_{i=1}^{n} a_i^2 = \boldsymbol{a}^\top \boldsymbol{a} = \|\boldsymbol{a}\|^2$$

com $\boldsymbol{a} = (a_1, \ldots, a_n)^\top$.

- Então podemos reescrever (3) como

$$Q(\beta) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = (\boldsymbol{y} - \boldsymbol{X}\beta)^\top(\boldsymbol{y} - \boldsymbol{X}\beta) = \|\boldsymbol{y} - \boldsymbol{X}\beta\|^2,$$

i.e., $Q(\beta)$ é uma forma quadrática.

Método de Mínimos Quadrados

- Sob as suposições usuais associadas ao MRLS é possível obter o Estimador de Mínimos Quadrados (EMQ) de $\boldsymbol{\beta}$ através da minimização da seguinte função objetivo

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2. \tag{3}$$

- Lembrando que qualquer soma de quadrados pode ser reescrita como (Lista 0) uma forma quadrática

$$\sum_{i=1}^{n} a_i^2 = \boldsymbol{a}^\top \boldsymbol{a} = \|\boldsymbol{a}\|^2$$

com $\boldsymbol{a} = (a_1, \ldots, a_n)^\top$.

- Então podemos reescrever (3) como

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) = \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2,$$

i.e., $Q(\boldsymbol{\beta})$ é uma forma quadrática.

Método de Mínimos Quadrados

Note que

$$\begin{aligned} Q(\beta) &= (\boldsymbol{y} - \boldsymbol{X}\beta)^\top(\boldsymbol{y} - \boldsymbol{X}\beta) \\ &= \boldsymbol{y}^\top\boldsymbol{y} - \boldsymbol{y}^\top\boldsymbol{X}\beta - \beta^\top\boldsymbol{X}^\top\boldsymbol{y} + \beta^\top\boldsymbol{X}^\top\boldsymbol{X}\beta. \end{aligned}$$

Dado que $\boldsymbol{y}^\top\boldsymbol{X}\beta = \beta^\top\boldsymbol{X}^\top\boldsymbol{y}$ (Por que isso é verdade?), então a função objetivo se simplifica a

$$Q(\beta) = \boldsymbol{y}^\top\boldsymbol{y} - 2\boldsymbol{y}^\top\boldsymbol{X}\beta + \beta^\top(\boldsymbol{X}^\top\boldsymbol{X})\beta.$$

Resultados - Q15 Lista 0

Lembrete: Considere $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ e $\boldsymbol{a} = (a_1, \ldots, a_n)^\top$ vetores reais de dimensão $n \times 1$ e $\boldsymbol{A}$ uma matriz quadrada de ordem n. Então,

i) $\dfrac{\partial \boldsymbol{a}^\top \boldsymbol{x}}{\partial \boldsymbol{x}} = \dfrac{\partial \boldsymbol{x}^\top \boldsymbol{a}}{\partial \boldsymbol{x}} = \boldsymbol{a}$.

ii) $\dfrac{\partial \boldsymbol{A}\boldsymbol{x}}{\partial \boldsymbol{x}} = \boldsymbol{A}^\top$.

iii) $\dfrac{\partial \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}{\partial \boldsymbol{x}} = (\boldsymbol{A} + \boldsymbol{A}^\top)\boldsymbol{x}$.

Equações normais

Logo, dado que a matriz $\boldsymbol{X}^\top \boldsymbol{X}$ é simétrica, tem-se que

$$\frac{\partial Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\boldsymbol{X}^\top \boldsymbol{y} + 2(\boldsymbol{X}^\top \boldsymbol{X})\boldsymbol{\beta},$$

de forma que o sistema de equações normais (equação de estimação) é dado por

$$\left.\frac{\partial Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = \boldsymbol{0},$$

implicando que

$$-2\boldsymbol{X}^\top \boldsymbol{y} + 2(\boldsymbol{X}^\top \boldsymbol{X})\widehat{\boldsymbol{\beta}} = \boldsymbol{0} \Leftrightarrow (\boldsymbol{X}^\top \boldsymbol{X})\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^\top \boldsymbol{y}.$$

Considerando que $\boldsymbol{X}$ é de **posto completo**, implicando que $\boldsymbol{X}^\top \boldsymbol{X}$ é não singular, tem-se que a solução do sistema de equações, e respectivo candidato a EMQ de $\boldsymbol{\beta}$, é

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y}. \tag{4}$$

Matriz de segundas derivadas

A matriz de segundas derivadas, com relação ao vetor de parâmetros, da função objetivo é dada por

$$\frac{\partial^2 Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = 2(\boldsymbol{X}^\top \boldsymbol{X})$$

que é uma matriz positiva definida (exercício - fazer quadro). Portanto,

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y}$$

corresponde ao estimador de Mínimos Quadrados de $\boldsymbol{\beta}$.

Exercícios (entregar próxima aula)

Exercício 6: Discuta sob que condição a matriz de especificação $\boldsymbol{X}$ associada ao MRLS (2) é de posto completo.

Exercício 7: Mostre que $\boldsymbol{A}^\top \boldsymbol{A}$ é não-singular, **se e somente se**, $\boldsymbol{A}$ é de posto completo, sendo $\boldsymbol{A} \in \mathbb{R}^{n \times p}$, i.e., uma matriz de dimensão $n \times p$.

Exercício 8: Se $\boldsymbol{A} \in \mathbb{R}^{n \times p}$, mostre que $\boldsymbol{A}^\top \boldsymbol{A}$ é positiva definida, i.e., $\forall \boldsymbol{x} \in \mathbb{R}^p$

$$\boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} \geq 0.$$

Exercício 9: Mostre que

$$Q(\boldsymbol{\beta}) = \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 = \left\|\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right\|^2 + \left\|\boldsymbol{X}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right)\right\|^2.$$

Observações

Obs1. Note que

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y} = \boldsymbol{A}_x \boldsymbol{y},$$

i.e., $\widehat{\boldsymbol{\beta}}$ é uma transformação linear do vetor de variáveis respostas, implicando que $\widehat{\beta}_0$ e $\widehat{\beta}_1$ são combinações lineares de $\boldsymbol{y}$ como já provamos.

Obs2. Perceba que

$$\boldsymbol{X}^\top \boldsymbol{X} = \begin{pmatrix} n & n\bar{x}_n \\ n\bar{x}_n & \sum_{i=1}^n x_i^2 \end{pmatrix}, \quad \boldsymbol{X}^\top \boldsymbol{y} = \begin{pmatrix} n\bar{y}_n \\ \sum_{i=1}^n x_i y_i \end{pmatrix}$$

$$\text{e} \quad (\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \frac{1}{S_{xx}} \begin{pmatrix} \sum_{i=1}^n x_i^2/n & -\bar{x}_n \\ -\bar{x}_n & 1 \end{pmatrix}.$$

Observações

Portanto,

$$(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y} = \frac{1}{S_{xx}}\begin{pmatrix} \sum_{i=1}^{n} x_i^2/n & -\bar{x}_n \\ -\bar{x}_n & 1 \end{pmatrix}\begin{pmatrix} n\bar{y}_n \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}.$$

Lembrando que

$$\sum_{i=1}^{n} x_i^2 = S_{xx} + n\bar{x}_n^2 \text{ e } \sum_{i=1}^{n} x_i y_i = S_{xy} + n\bar{x}_n\bar{y}_n,$$

então

$$(\sum_{i=1}^{n} x_i^2/n)n\bar{y}_n - \bar{x}_n\sum_{i=1}^{n} x_i y_i = S_{xx}\bar{y}_n - \bar{x}_n S_{xy}$$

e

$$-\bar{x}_n n\bar{y}_n + \sum_{i=1}^{n} x_i y_i = S_{xy}.$$

Observações

Logo, pode-se concluir que

$$\therefore (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y} = \frac{1}{S_{xx}} \begin{pmatrix} S_{xx}\bar{y}_n - \bar{x}_n S_{xy} \\ S_{xy} \end{pmatrix} = \begin{pmatrix} \bar{y}_n - \bar{x}_n S_{xy}/S_{xx} \\ S_{xy}/S_{xx} \end{pmatrix} = \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \end{pmatrix}$$

Obs3. Sob o MRLS tem-se que

$$\boldsymbol{y} \sim (\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}_n),$$

em que $\boldsymbol{I}_n$ denota a matriz identidade de ordem, n. Adicionalmente, sob a suposição de normalidade

$$\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}_n).$$

Observações

Obs4. Tem-se que

$$\begin{aligned}
\mathbb{E}[\widehat{\boldsymbol{\beta}}] &= \mathbb{E}[(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y}] = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\mathbb{E}[\boldsymbol{y}] \\
&= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta},
\end{aligned}$$

i.e., $\widehat{\boldsymbol{\beta}}$ é um estimador não viesado de $\boldsymbol{\beta}$, como já havíamos provado de forma marginal.

Por outro lado, a matriz de variâncias-covariâncias do EMQ de $\boldsymbol{\beta}$ é dada por

$$\begin{aligned}
\mathrm{Var}[\widehat{\boldsymbol{\beta}}] &= \mathrm{Var}[\boldsymbol{A_x}\boldsymbol{y}] = \boldsymbol{A_x}\mathrm{Var}[\boldsymbol{y}]\boldsymbol{A_x}^\top = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\sigma^2\boldsymbol{I}_n[(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top]^\top \\
&= \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1} = \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1} \\
&= \frac{\sigma^2}{S_{xx}}\begin{pmatrix} \sum_{i=1}^n x_i^2/n & -\bar{x}_n \\ -\bar{x}_n & 1 \end{pmatrix} = \sigma^2\begin{pmatrix} \frac{1}{n} + \frac{\bar{x}_n^2}{S_{xx}} & -\frac{\bar{x}_n}{S_{xx}} \\ -\frac{\bar{x}_n}{S_{xx}} & \frac{1}{S_{xx}} \end{pmatrix},
\end{aligned}$$

dado que $\sum_{i=1}^n x_i^2 = S_{xx} + n\bar{x}_n^2$.

Observações

Obs5. Sob a suposição de normalidade, tem-se que

$$\widehat{\boldsymbol{\beta}} \sim \mathcal{N}_2 \left[\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\right].$$

Obs6. (Teorema de Gauss-Markov) Sob as suposições do MRLS, $\widehat{\boldsymbol{\beta}}$ é o BLUE de $\boldsymbol{\beta}$.

A versão **multivariada** do teorema de Gauss-Markov será demonstrada nas próximas aulas.

Obs7. O vetor de valores preditos é dado por

$$\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y} = \boldsymbol{H}\boldsymbol{y},$$

que é uma transformação linear das observações. A matriz $\boldsymbol{H} := \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top$ é uma matriz simétrica e idempotente, que representa a matriz de projeção ortogonal do subespaço gerado pelas colunas de $\boldsymbol{X}$ ($\mathcal{C}(\boldsymbol{X})$).