# Mínimos Quadrados Ponderados e Mínimos Quadrados Generalizados

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**, novembro de 2022

# Conteúdo

# Presuposições

- Para se ajustar e fazer inferência na classe de modelos de regressão lineares é necessário, como bem sabemos, que algumas presuposições sejam válidas.
- Uma delas versa sobre condições a respeito da distribuição da fonte de variação $\boldsymbol{e}$, em especial, é considerado que

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

- O que ocorre caso a fonte de variação não seja homoscedástica e continuemos a usar o método de mínimos quadrados usual, de agora em diante, denominado de método de Mínimos Quadrados Ordinário (MQO)? 😇
- Em breve, responderemos esta pergunta.

# Presuposições

- Para se ajustar e fazer inferência na classe de modelos de regressão lineares é necessário, como bem sabemos, que algumas presuposições sejam válidas.
- Uma delas versa sobre condições a respeito da distribuição da fonte de variação $\boldsymbol{e}$, em especial, é considerado que

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

- O que ocorre caso a fonte de variação não seja homoscedástica e continuemos a usar o método de mínimos quadrados usual, de agora em diante, denominado de método de Mínimos Quadrados Ordinário (MQO)?
- Em breve, responderemos esta pergunta.

# Presuposições

- Para se ajustar e fazer inferência na classe de modelos de regressão lineares é necessário, como bem sabemos, que algumas presuposições sejam válidas.
- Uma delas versa sobre condições a respeito da distribuição da fonte de variação $\boldsymbol{e}$, em especial, é considerado que

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

- O que ocorre caso a fonte de variação não seja homoscedástica e continuemos a usar o método de mínimos quadrados usual, de agora em diante, denominado de método de Mínimos Quadrados Ordinário (MQO)? 😇
- Em breve, responderemos esta pergunta.

# Presuposições

- Para se ajustar e fazer inferência na classe de modelos de regressão lineares é necessário, como bem sabemos, que algumas presuposições sejam válidas.
- Uma delas versa sobre condições a respeito da distribuição da fonte de variação $\boldsymbol{e}$, em especial, é considerado que

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

- O que ocorre caso a fonte de variação não seja homoscedástica e continuemos a usar o método de mínimos quadrados usual, de agora em diante, denominado de método de Mínimos Quadrados Ordinário (MQO)? 😇
- Em breve, responderemos esta pergunta.

# Mínimos Quadrados Ponderados

- Consideremos um MRLM com a seguinte forma funcional

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

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2 \mathrm{diag}(w_1, \ldots, w_n))$ com todas as demais pressuposições válidas, i.e., a única pressuposição que não se verifica é a de homoscedasticidade.

- Assume-se que $w_i$ é conhecido, $\forall i = 1, \ldots, n$, ou seja, tem-se um modelo heteroscedástico, com heteroscedasticidade conhecida.

- O respectivo MRLS associado é

$$y_i = \beta_0 + \beta_1 x_i + e_i, e_i \overset{\text{ind}}{\sim} (0, \sigma^2 w_i), w_i \text{ conhecido}, i = 1, \ldots, n.$$

- Como $w_i$ é conhecido, podemos considerar

$$\begin{aligned} \frac{y_i}{\sqrt{w_i}} &= \frac{\beta_0}{\sqrt{w_i}} + \beta_1 \frac{x_i}{\sqrt{w_i}} + \frac{e_i}{\sqrt{w_i}} \\ y_i^* &= \beta_0^* + \beta_1 x_i^* + e_i^*, e_i^* \overset{\text{ind}}{\sim} (0, \sigma^2). \end{aligned}$$

e este último modelo pode ser ajustado via metodologia de MRLS usual.

# Mínimos Quadrados Ponderados

- Consideremos um MRLM com a seguinte forma funcional

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

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2 \mathrm{diag}(w_1, \ldots, w_n))$ com todas as demais pressuposições válidas, i.e., a única pressuposição que não se verifica é a de homoscedasticidade.

- Assume-se que $w_i$ é conhecido, $\forall i = 1, \ldots, n$, ou seja, tem-se um modelo heteroscedástico, com heteroscedasticidade conhecida.

- O respectivo MRLS associado é

$$y_i = \beta_0 + \beta_1 x_i + e_i, e_i \overset{\text{ind}}{\sim} (0, \sigma^2 w_i), w_i \text{ conhecido}, i = 1, \ldots, n.$$

- Como $w_i$ é conhecido, podemos considerar

$$\begin{aligned} \frac{y_i}{\sqrt{w_i}} &= \frac{\beta_0}{\sqrt{w_i}} + \beta_1 \frac{x_i}{\sqrt{w_i}} + \frac{e_i}{\sqrt{w_i}} \\ y_i^* &= \beta_0^* + \beta_1 x_i^* + e_i^*, e_i^* \overset{\text{ind}}{\sim} (0, \sigma^2). \end{aligned}$$

e este último modelo pode ser ajustado via metodologia de MRLS usual.

# Mínimos Quadrados Ponderados

- Consideremos um MRLM com a seguinte forma funcional

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

em que $\boldsymbol{e} \sim (\mathbf{0}, \sigma^2 \mathrm{diag}(w_1, \ldots, w_n))$ com todas as demais pressuposições válidas, i.e., a única pressuposição que não se verifica é a de homoscedasticidade.

- Assume-se que $w_i$ é conhecido, $\forall i = 1, \ldots, n$, ou seja, tem-se um modelo heteroscedástico, com heteroscedasticidade conhecida.

- O respectivo MRLS associado é

$$y_i = \beta_0 + \beta_1 x_i + e_i, e_i \overset{\text{ind}}{\sim} (0, \sigma^2 w_i), w_i \text{ conhecido}, i = 1, \ldots, n.$$

- Como $w_i$ é conhecido, podemos considerar

$$\begin{aligned} \frac{y_i}{\sqrt{w_i}} &= \frac{\beta_0}{\sqrt{w_i}} + \beta_1 \frac{x_i}{\sqrt{w_i}} + \frac{e_i}{\sqrt{w_i}} \\ y_i^* &= \beta_0^* + \beta_1 x_i^* + e_i^*, e_i^* \overset{\text{ind}}{\sim} (0, \sigma^2). \end{aligned}$$

e este último modelo pode ser ajustado via metodologia de MRLS usual.

# Mínimos Quadrados Ponderados

- Consideremos um MRLM com a seguinte forma funcional

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

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2 \mathrm{diag}(w_1, \ldots, w_n))$ com todas as demais pressuposições válidas, i.e., a única pressuposição que não se verifica é a de homoscedasticidade.

- Assume-se que $w_i$ é conhecido, $\forall i = 1, \ldots, n$, ou seja, tem-se um modelo heteroscedástico, com heteroscedasticidade conhecida.

- O respectivo MRLS associado é

$$y_i = \beta_0 + \beta_1 x_i + e_i, e_i \overset{\text{ind}}{\sim} (0, \sigma^2 w_i), w_i \text{ conhecido}, i = 1, \ldots, n.$$

- Como $w_i$ é conhecido, podemos considerar

$$\begin{aligned} \frac{y_i}{\sqrt{w_i}} &= \frac{\beta_0}{\sqrt{w_i}} + \beta_1 \frac{x_i}{\sqrt{w_i}} + \frac{e_i}{\sqrt{w_i}} \\ y_i^* &= \beta_0^* + \beta_1 x_i^* + e_i^*, e_i^* \overset{\text{ind}}{\sim} (0, \sigma^2). \end{aligned}$$

e este último modelo pode ser ajustado via metodologia de MRLS usual.

# Mínimos Quadrados Ponderados

- A função objetivo para determinar o EMQ é dada por

$$
\begin{aligned}
S(\beta_0, \beta_1) &= \sum_{i=1}^{n} (y_i^* - \beta_0^* - \beta_1 x_i^*)^2 \\
&= \sum_{i=1}^{n} \frac{1}{w_i} (y_i - \beta_0 - \beta_1 x_i)^2,
\end{aligned}
$$

ou seja as observações agora não possuem o mesmo peso no processo de estimação, neste caso o peso é inversamente proporcional a variância associada a fonte de variação daquela observação.

- Por esta razão este método é denominado de método de Mínimos Quadrados Ponderados (MQP).

# Mínimos Quadrados Ponderados

- A função objetivo para determinar o EMQ é dada por

$$\begin{aligned} S(\beta_0, \beta_1) &= \sum_{i=1}^{n} (y_i^* - \beta_0^* - \beta_1 x_i^*)^2 \\ &= \sum_{i=1}^{n} \frac{1}{w_i} (y_i - \beta_0 - \beta_1 x_i)^2, \end{aligned}$$

  ou seja as observações agora não possuem o mesmo peso no processo de estimação, neste caso o peso é inversamente proporcional a variância associada a fonte de variação daquela observação.

- Por esta razão este método é denominado de método de Mínimos Quadrados Ponderados (MQP).

# Mínimos Quadrados Ponderados

- Esta ideia pode ser utilizada também para obter um processo de estimação robusto.
- Primeiramente, ajusta-se um MRLS usual e posteriomente se faz uso desta ideia utilizando os pesos como $w_i = \hat{e}_i{}^2$, ou seja, dando um peso menor para observações com resíduos elevados. Essencialmente é o que se acontece quando consideramos uma distribuição $t$ ou Laplace ao invés da Normal.

# Mínimos Quadrados Ponderados

- Esta ideia pode ser utilizada também para obter um processo de estimação robusto.
- Primeiramente, ajusta-se um MRLS usual e posteriomente se faz uso desta ideia utilizando os pesos como $w_i = {\hat{e}_i}^2$, ou seja, dando um peso menor para observações com resíduos elevados. Essencialmente é o que se acontece quando consideramos uma distribuição $t$ ou Laplace ao invés da Normal.

# Mínimos Quadrados Ponderados

- Agora, fazendo uso da notação matricial, tem-se a seguinte forma funcional do MRLM

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \quad (1)$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2 \boldsymbol{W})$, em que $\boldsymbol{W} = \mathrm{diag}(w_1, \ldots, w_n) = \oplus_{i=1}^n w_i$.

- Como $\boldsymbol{W}$ é uma matriz diagonal, então $\boldsymbol{W}^{-1/2} = \mathrm{diag}(w_1^{-1/2}, \ldots, w_n^{-1/2})$, de forma que $\boldsymbol{W}^{-1/2}\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$, logo pré-multiplicando os elementos de (5) por $\boldsymbol{W}^{-1/2}$ obtemos

$$\boldsymbol{W}^{-1/2}\boldsymbol{y} = \boldsymbol{W}^{-1/2}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}^{-1/2}\boldsymbol{e}, \quad (2)$$

que representa um MRLM homoscedástico, com vetor de variável resposta $\boldsymbol{W}^{-1/2}\boldsymbol{y}$ e matriz de especificação $\boldsymbol{W}^{-1/2}\boldsymbol{X}$.

- A função objetivo correspondente é dada por:

$$S(\boldsymbol{\beta}) = (\boldsymbol{W}^{-1/2}\boldsymbol{y} - \boldsymbol{W}^{-1/2}\boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{W}^{-1/2}\boldsymbol{y} - \boldsymbol{W}^{-1/2}\boldsymbol{X}\boldsymbol{\beta}),$$

e lembrando que $(\boldsymbol{A}\boldsymbol{B})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top$ e que $\boldsymbol{W}$ é uma matriz simétrica, implicando que $\boldsymbol{W}^{-1/2}$ também é, então $S(\boldsymbol{\beta})$ consegue ser reescrita como

$$S(\boldsymbol{\beta}) = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top \boldsymbol{W}^{-1} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}).$$

# Mínimos Quadrados Ponderados

- Agora, fazendo uso da notação matricial, tem-se a seguinte forma funcional do MRLM

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}, \tag{1}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2 \boldsymbol{W})$, em que $\boldsymbol{W} = \mathrm{diag}(w_1, \ldots, w_n) = \oplus_{i=1}^{n} w_i$.

- Como $\boldsymbol{W}$ é uma matriz diagonal, então $\boldsymbol{W}^{-1/2} = \mathrm{diag}(w_1^{-1/2}, \ldots, w_n^{-1/2})$, de forma que $\boldsymbol{W}^{-1/2}\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$, logo pré-multiplicando os elementos de (5) por $\boldsymbol{W}^{-1/2}$ obtemos

$$\boldsymbol{W}^{-1/2}\boldsymbol{y} = \boldsymbol{W}^{-1/2}\boldsymbol{X}\beta + \boldsymbol{W}^{-1/2}\boldsymbol{e}, \tag{2}$$

que representa um MRLM homoscedástico, com vetor de variável resposta $\boldsymbol{W}^{-1/2}\boldsymbol{y}$ e matriz de especificação $\boldsymbol{W}^{-1/2}\boldsymbol{X}$.

- A função objetivo correspondente é dada por:

$$S(\beta) = (\boldsymbol{W}^{-1/2}\boldsymbol{y} - \boldsymbol{W}^{-1/2}\boldsymbol{X}\beta)^\top (\boldsymbol{W}^{-1/2}\boldsymbol{y} - \boldsymbol{W}^{-1/2}\boldsymbol{X}\beta),$$

e lembrando que $(\boldsymbol{AB})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top$ e que $\boldsymbol{W}$ é uma matriz simétrica, implicando que $\boldsymbol{W}^{-1/2}$ também é, então $S(\beta)$ consegue ser reescrita como

$$S(\beta) = (\boldsymbol{y} - \boldsymbol{X}\beta)^\top \boldsymbol{W}^{-1} (\boldsymbol{y} - \boldsymbol{X}\beta).$$

# Mínimos Quadrados Ponderados

- Agora, fazendo uso da notação matricial, tem-se a seguinte forma funcional do MRLM

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}, \tag{1}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{W})$, em que $\boldsymbol{W} = \mathrm{diag}(w_1, \ldots, w_n) = \oplus_{i=1}^{n} w_i$.

- Como $\boldsymbol{W}$ é uma matriz diagonal, então $\boldsymbol{W}^{-1/2} = \mathrm{diag}(w_1^{-1/2}, \ldots, w_n^{-1/2})$, de forma que $\boldsymbol{W}^{-1/2}\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{I}_n)$, logo pré-multiplicando os elementos de (5) por $\boldsymbol{W}^{-1/2}$ obtemos

$$\boldsymbol{W}^{-1/2}\boldsymbol{y} = \boldsymbol{W}^{-1/2}\boldsymbol{X}\beta + \boldsymbol{W}^{-1/2}\boldsymbol{e}, \tag{2}$$

que representa um MRLM homoscedástico, com vetor de variável resposta $\boldsymbol{W}^{-1/2}\boldsymbol{y}$ e matriz de especificação $\boldsymbol{W}^{-1/2}\boldsymbol{X}$.

- A função objetivo correspondente é dada por:

$$S(\beta) = (\boldsymbol{W}^{-1/2}\boldsymbol{y} - \boldsymbol{W}^{-1/2}\boldsymbol{X}\beta)^\top(\boldsymbol{W}^{-1/2}\boldsymbol{y} - \boldsymbol{W}^{-1/2}\boldsymbol{X}\beta),$$

e lembrando que $(\boldsymbol{AB})^\top = \boldsymbol{B}^\top\boldsymbol{A}^\top$ e que $\boldsymbol{W}$ é uma matriz simétrica, implicando que $\boldsymbol{W}^{-1/2}$ também é, então $S(\beta)$ consegue ser reescrita como

$$S(\beta) = (\boldsymbol{y} - \boldsymbol{X}\beta)^\top\boldsymbol{W}^{-1}(\boldsymbol{y} - \boldsymbol{X}\beta).$$

# Mínimos Quadrados Ponderados

- Como é bem conhecido, o estimador de **MQO** de $\boldsymbol{\beta}$ é dado por

$$\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y}. \tag{3}$$

- Já o estimador de **MQP** é dado por

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= ([\boldsymbol{W}^{-1/2}\boldsymbol{X}]^\top \boldsymbol{W}^{-1/2}\boldsymbol{X})^{-1}[\boldsymbol{W}^{-1/2}\boldsymbol{X}]^\top \boldsymbol{W}^{-1/2}\boldsymbol{y} \\ &= (\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{y}. \end{aligned} \tag{4}$$

- Pode-se provar, que mesmo sob a má especificação, i.e. considerar o ajuste de um modelo homoscedástico, sendo ele heteroscedático, o EMQO continua sendo não viesado, basta perceber que

$$\begin{aligned}\mathbb{E}\left[\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}\right] &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbb{E}[\boldsymbol{y}] \\ &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta}, \forall \boldsymbol{\beta} \in \mathbb{R}^p.\end{aligned}$$

- Então, se $\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}$ continua não viesado, por qual razão devemos considerar o EMQP? 🙂

# Mínimos Quadrados Ponderados

- Como é bem conhecido, o estimador de **MQO** de $\boldsymbol{\beta}$ é dado por

$$\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y}. \tag{3}$$

- Já o estimador de **MQP** é dado por

$$\begin{aligned} \widehat{\boldsymbol{\beta}} &= ([\boldsymbol{W}^{-1/2}\boldsymbol{X}]^\top \boldsymbol{W}^{-1/2}\boldsymbol{X})^{-1}[\boldsymbol{W}^{-1/2}\boldsymbol{X}]^\top \boldsymbol{W}^{-1/2}\boldsymbol{y} \\ &= (\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{y}. \end{aligned} \tag{4}$$

- Pode-se provar, que mesmo sob a má especificação, i.e. considerar o ajuste de um modelo homoscedástico, sendo ele heteroscedático, o EMQO continua sendo não viesado, basta perceber que

$$\begin{aligned} \mathbb{E}\left[\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}\right] &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbb{E}[\boldsymbol{y}] \\ &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta}, \forall \boldsymbol{\beta} \in \mathbb{R}^p. \end{aligned}$$

- Então, se $\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}$ continua não viesado, por qual razão devemos considerar o EMQP? 🙂

# Mínimos Quadrados Ponderados

- Como é bem conhecido, o estimador de **MQO** de $\boldsymbol{\beta}$ é dado por

$$\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y}. \tag{3}$$

- Já o estimador de **MQP** é dado por

$$\begin{aligned} \widehat{\boldsymbol{\beta}} &= ([\boldsymbol{W}^{-1/2}\boldsymbol{X}]^\top \boldsymbol{W}^{-1/2}\boldsymbol{X})^{-1}[\boldsymbol{W}^{-1/2}\boldsymbol{X}]^\top \boldsymbol{W}^{-1/2}\boldsymbol{y} \\ &= (\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{y}. \end{aligned} \tag{4}$$

- Pode-se provar, que mesmo sob a má especificação, i.e. considerar o ajuste de um modelo homoscedástico, sendo ele heteroscedático, o EMQO continua sendo não viesado, basta perceber que

$$\begin{aligned} \mathbb{E}\left[\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}\right] &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbb{E}[\boldsymbol{y}] \\ &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta}, \forall \boldsymbol{\beta} \in \mathbb{R}^p. \end{aligned}$$

- Então, se $\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}$ continua não viesado, por qual razão devemos considerar o EMQP? 🙂

# Mínimos Quadrados Ponderados

- Como é bem conhecido, o estimador de **MQO** de $\boldsymbol{\beta}$ é dado por

$$\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y}. \tag{3}$$

- Já o estimador de **MQP** é dado por

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= ([\boldsymbol{W}^{-1/2}\boldsymbol{X}]^\top \boldsymbol{W}^{-1/2}\boldsymbol{X})^{-1}[\boldsymbol{W}^{-1/2}\boldsymbol{X}]^\top \boldsymbol{W}^{-1/2}\boldsymbol{y} \\ &= (\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{y}. \end{aligned} \tag{4}$$

- Pode-se provar, que mesmo sob a má especificação, i.e. considerar o ajuste de um modelo homoscedástico, sendo ele heteroscedático, o EMQO continua sendo não viesado, basta perceber que

$$\begin{aligned}\mathbb{E}\left[\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}\right] &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbb{E}[\boldsymbol{y}] \\ &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta}, \forall \boldsymbol{\beta} \in \mathbb{R}^p.\end{aligned}$$

- Então, se $\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}$ continua não viesado, por qual razão devemos considerar o EMQP? 🙂

# Mínimos Quadrados Ponderados

- A matriz de variância-covariâncias do estimador de MQO é dada por

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}} = \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{W}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1},$$

e pode-se mostrar que sob condições bem suaves que o EMQO continua consistente.

- Já a matriz de variância-covariâncias do estimador de **MQP** é dada por

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\beta}}} = \sigma^2(\boldsymbol{X}^\top\boldsymbol{W}^{-1}\boldsymbol{X})^{-1},$$

e pelo teorema de Gauss-Markov, $\widehat{\beta}$ é o **BLUE** de $\beta$, implicando que

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}} - \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\beta}}} \geq 0,$$

ou seja é não-negativa definida, implicando por exemplo, que o estimador de MQP de $\beta$ é consistente e mais eficiente que o estimador de MQO.

# Mínimos Quadrados Ponderados

- A matriz de variância-covariâncias do estimador de MQO é dada por

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}} = \sigma^2(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{W}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1},$$

e pode-se mostrar que sob condições bem suaves que o EMQO continua consistente.

- Já a matriz de variância-covariâncias do estimador de **MQP** é dada por

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\beta}}} = \sigma^2(\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1},$$

e pelo teorema de Gauss-Markov, $\widehat{\boldsymbol{\beta}}$ é o **BLUE** de $\boldsymbol{\beta}$, implicando que

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\beta}}_{\mathrm{MQO}}} - \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\beta}}} \geq 0,$$

ou seja é não-negativa definida, implicando por exemplo, que o estimador de MQP de $\boldsymbol{\beta}$ é consistente e mais eficiente que o estimador de MQO.

# Comentários

- Podemos utilizar **todos** os resultados da classe de MRLM, basta considerar $\boldsymbol{W}^{-1/2}\boldsymbol{y}$ ao invés de $\boldsymbol{y}$ e $\boldsymbol{W}^{-1/2}\boldsymbol{X}$ ao invés de $\boldsymbol{X}$. Por exemplo, o MINQUE de $\sigma^2$ é dado por:

$$\hat{\sigma}^2 = \mathrm{QMRes} = \frac{\boldsymbol{y}^\top \boldsymbol{W}^{-1/2}[\boldsymbol{I} - \boldsymbol{W}^{-1/2}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{W}^{-1/2}]\boldsymbol{W}^{-1/2}\boldsymbol{y}}{n-p}$$

$$= \frac{\boldsymbol{y}^\top[\boldsymbol{W}^{-1} - \boldsymbol{W}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{W}^{-1}]\boldsymbol{y}}{n-p}.$$

- Podemos através das técnicas de diagnóstico já estudadas ter ideia sob a forma da heteroscedasticidade, por exemplo, ter indícios de que $\mathrm{Var}[e_i] \propto \sigma^2 x_i^2$.
- No software `R` podemos usar a função `'weight'` em `lm` para utilizar o método de mínimos quadrados ponderados. Olhar com calma a forma no qual ele considera a função 'peso' pois é diferente do formato utilizado aqui.

# Comentários

- Podemos utilizar **todos** os resultados da classe de MRLM, basta considerar $\boldsymbol{W}^{-1/2}\boldsymbol{y}$ ao invés de $\boldsymbol{y}$ e $\boldsymbol{W}^{-1/2}\boldsymbol{X}$ ao invés de $\boldsymbol{X}$. Por exemplo, o MINQUE de $\sigma^2$ é dado por:

$$
\begin{aligned}
\hat{\sigma}^2 = \mathrm{QMRes} &= \frac{\boldsymbol{y}^\top \boldsymbol{W}^{-1/2}[\boldsymbol{I} - \boldsymbol{W}^{-1/2}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{W}^{-1/2}]\boldsymbol{W}^{-1/2}\boldsymbol{y}}{n-p} \\
&= \frac{\boldsymbol{y}^\top[\boldsymbol{W}^{-1} - \boldsymbol{W}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{W}^{-1}]\boldsymbol{y}}{n-p}.
\end{aligned}
$$

- Podemos através das técnicas de diagnóstico já estudadas ter ideia sob a forma da heteroscedasticidade, por exemplo, ter indícios de que $\mathrm{Var}[e_i] \propto \sigma^2 x_i^2$.

- No software `R` podemos usar a função `'weight'` em `lm` para utilizar o método de mínimos quadrados ponderados. Olhar com calma a forma no qual ele considera a função 'peso' pois é diferente do formato utilizado aqui. 😐

# Comentários

- Podemos utilizar **todos** os resultados da classe de MRLM, basta considerar $\boldsymbol{W}^{-1/2}\boldsymbol{y}$ ao invés de $\boldsymbol{y}$ e $\boldsymbol{W}^{-1/2}\boldsymbol{X}$ ao invés de $\boldsymbol{X}$. Por exemplo, o MINQUE de $\sigma^2$ é dado por:

$$\begin{aligned}\hat{\sigma}^2 = \mathrm{QMRes} &= \frac{\boldsymbol{y}^\top \boldsymbol{W}^{-1/2}[\boldsymbol{I} - \boldsymbol{W}^{-1/2}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{W}^{-1/2}]\boldsymbol{W}^{-1/2}\boldsymbol{y}}{n-p} \\ &= \frac{\boldsymbol{y}^\top [\boldsymbol{W}^{-1} - \boldsymbol{W}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{W}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{W}^{-1}]\boldsymbol{y}}{n-p}.\end{aligned}$$

- Podemos através das técnicas de diagnóstico já estudadas ter ideia sob a forma da heteroscedasticidade, por exemplo, ter indícios de que $\mathrm{Var}[e_i] \propto \sigma^2 x_i^2$.
- No software `R` podemos usar a função `'weight'` em `lm` para utilizar o método de mínimos quadrados ponderados. Olhar com calma a forma no qual ele considera a função 'peso' pois é diferente do formato utilizado aqui. 😐

# Comentários

- Se o objetivo for somente em inferência de primeira ordem, podemos continuar usando o estimador de MQO sem problemas, dado que ele é não viesado e consistente. Todavia, na prática se tem interesse em se fazer inferência de segunda ordem, construção de intervalos de confiança, testes de hipóteses, etc... e neste caso é preferível utilizar o estimador de MQP e sua respectiva matriz de variância-covariância estimada.
- Quando a matriz $\boldsymbol{W}$ não for conhecida? Como devemos proceder?
- Quando $\boldsymbol{W}$ não for diagonal, i.e., quando tivermos uma estrutura de correlação associada, muito comum em análise de série temporais, análise de dados longitudinais, processos estocásticos, por exemplo, como proceder?
- Agora vamos apresentar algumas propostas associados ao primeiro caso, i.e., quando tivermos um modelo heteroscedástico com heteroscedasticidade de forma desconhecida.

# Comentários

- Se o objetivo for somente em inferência de primeira ordem, podemos continuar usando o estimador de MQO sem problemas, dado que ele é não viesado e consistente. Todavia, na prática se tem interesse em se fazer inferência de segunda ordem, construção de intervalos de confiança, testes de hipóteses, etc... e neste caso é preferível utilizar o estimador de MQP e sua respectiva matriz de variância-covariância estimada.
- Quando a matriz $\boldsymbol{W}$ não for conhecida? Como devemos proceder?
- Quando $\boldsymbol{W}$ não for diagonal, i.e., quando tivermos uma estrutura de correlação associada, muito comum em análise de série temporais, análise de dados longitudinais, processos estocásticos, por exemplo, como proceder?
- Agora vamos apresentar algumas propostas associados ao primeiro caso, i.e., quando tivermos um modelo heteroscedástico com heteroscedasticidade de forma desconhecida.

# Comentários

- Se o objetivo for somente em inferência de primeira ordem, podemos continuar usando o estimador de MQO sem problemas, dado que ele é não viesado e consistente. Todavia, na prática se tem interesse em se fazer inferência de segunda ordem, construção de intervalos de confiança, testes de hipóteses, etc... e neste caso é preferível utilizar o estimador de MQP e sua respectiva matriz de variância-covariância estimada.
- Quando a matriz $\boldsymbol{W}$ não for conhecida? Como devemos proceder?
- Quando $\boldsymbol{W}$ não for diagonal, i.e., quando tivermos uma estrutura de correlação associada, muito comum em análise de série temporais, análise de dados longitudinais, processos estocásticos, por exemplo, como proceder?
- Agora vamos apresentar algumas propostas associados ao primeiro caso, i.e., quando tivermos um modelo heteroscedástico com heteroscedasticidade de forma desconhecida.

# Comentários

- Se o objetivo for somente em inferência de primeira ordem, podemos continuar usando o estimador de MQO sem problemas, dado que ele é não viesado e consistente. Todavia, na prática se tem interesse em se fazer inferência de segunda ordem, construção de intervalos de confiança, testes de hipóteses, etc... e neste caso é preferível utilizar o estimador de MQP e sua respectiva matriz de variância-covariância estimada.
- Quando a matriz $\boldsymbol{W}$ não for conhecida? Como devemos proceder?
- Quando $\boldsymbol{W}$ não for diagonal, i.e., quando tivermos uma estrutura de correlação associada, muito comum em análise de série temporais, análise de dados longitudinais, processos estocásticos, por exemplo, como proceder?
- Agora vamos apresentar algumas propostas associados ao primeiro caso, i.e., quando tivermos um modelo heteroscedástico com heteroscedasticidade de forma desconhecida.

# Heteroscedasticidade de forma desconhecida

- Até o momento, consideramos o modelo heteroscedástico em que a forma da heteroscedasticidade é conhecida. Entretanto, em problemas práticos, freqüentemente desconhecemos se as fontes de variação são homoscedásticas ou heteroscedásticas, e se forem heteroscedásticas não temos ideia de que forma.
- Em Hoffman (2016) é apresentado um teste para verificar a existência de heteroscedasticidade, todavia, já vimos procedimentos baseados na análise de diagnóstico para avaliar isto. Uma excelente referência nesta linha é o livro:
- Davidson, R. and MacKinnon, J.G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press: New York.
- Hoje em dia, isso não é de se preocupar, tendo em vista que conseguimos modelar heteroscedasticidade de forma extremamente simples através de modelos lineares generalizados, modelos de regressão simétricos, modelos de regressão beta, etc... tendo em vista que uma das poucas variáveis aleatórias em que o parâmetro de escala é funcionamente independente do parâmetro de posição é a Normal.

# Heteroscedasticidade de forma desconhecida

- Até o momento, consideramos o modelo heteroscedástico em que a forma da heteroscedasticidade é conhecida. Entretanto, em problemas práticos, freqüentemente desconhecemos se as fontes de variação são homoscedásticas ou heteroscedásticas, e se forem heteroscedásticas não temos ideia de que forma.
- Em Hoffman (2016) é apresentado um teste para verificar a existência de heteroscedasticidade, todavia, já vimos procedimentos baseados na análise de diagnóstico para avaliar isto. Uma excelente referência nesta linha é o livro:
- Davidson, R. and MacKinnon, J.G. (1993). *Estimation and Inference in Econometrics.* Oxford University Press: New York.
- Hoje em dia, isso não é de se preocupar, tendo em vista que conseguimos modelar heteroscedasticidade de forma extremamente simples através de modelos lineares generalizados, modelos de regressão simétricos, modelos de regressão beta, etc... tendo em vista que uma das poucas variáveis aleatórias em que o parâmetro de escala é funcionamente independente do parâmetro de posição é a Normal.

# Heteroscedasticidade de forma desconhecida

- Até o momento, consideramos o modelo heteroscedástico em que a forma da heteroscedasticidade é conhecida. Entretanto, em problemas práticos, freqüentemente desconhecemos se as fontes de variação são homoscedásticas ou heteroscedásticas, e se forem heteroscedásticas não temos ideia de que forma.
- Em Hoffman (2016) é apresentado um teste para verificar a existência de heteroscedasticidade, todavia, já vimos procedimentos baseados na análise de diagnóstico para avaliar isto. Uma excelente referência nesta linha é o livro:
- Davidson, R. and MacKinnon, J.G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press: New York.
- Hoje em dia, isso não é de se preocupar, tendo em vista que conseguimos modelar heteroscedasticidade de forma extremamente simples através de modelos lineares generalizados, modelos de regressão simétricos, modelos de regressão beta, etc... tendo em vista que uma das poucas variáveis aleatórias em que o parâmetro de escala é funcionamente independente do parâmetro de posição é a Normal.

# Heteroscedasticidade de forma desconhecida

- Até o momento, consideramos o modelo heteroscedástico em que a forma da heteroscedasticidade é conhecida. Entretanto, em problemas práticos, freqüentemente desconhecemos se as fontes de variação são homoscedásticas ou heteroscedásticas, e se forem heteroscedásticas não temos ideia de que forma.
- Em Hoffman (2016) é apresentado um teste para verificar a existência de heteroscedasticidade, todavia, já vimos procedimentos baseados na análise de diagnóstico para avaliar isto. Uma excelente referência nesta linha é o livro:
- Davidson, R. and MacKinnon, J.G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press: New York.
- Hoje em dia, isso não é de se preocupar, tendo em vista que conseguimos modelar heteroscedasticidade de forma extremamente simples através de modelos lineares generalizados, modelos de regressão simétricos, modelos de regressão beta, etc... tendo em vista que uma das poucas variáveis aleatórias em que o parâmetro de escala é funcionamente independente do parâmetro de posição é a Normal.

# Heteroscedasticidade de forma desconhecida

- Considere o MRLM

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \tag{5}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{\Omega})$, em que $\boldsymbol{\Omega} = \mathrm{diag}(w_1, \ldots, w_n)$ desconhecida.

- O estimador de Mínimos Quadrados Ordinários é tal que

$$\widehat{\beta}_{\mathrm{MQO}} \xrightarrow{\mathbb{P}} \boldsymbol{\beta}$$
$$\widehat{\beta}_{\mathrm{MQO}} \xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}),$$

todavia o mesmo não é eficiente.

- White (1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica*) propõe utilizar $\widehat{\boldsymbol{\Omega}} = \mathrm{diag}(\hat{e}_1^2, \ldots, \hat{e}_n^2)$ e estimar a matriz de variância-covariâncias por

$$\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\beta}}} = \mathrm{QMRes}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\widehat{\boldsymbol{\Omega}}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}.$$

- Tal procedimento gera estimadores consistentes para os erros-padrão dos estimadores dos coeficientes de regressão. Este estimador é denominado HC0.

# Heteroscedasticidade de forma desconhecida

- Considere o MRLM

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}, \tag{5}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{\Omega})$, em que $\boldsymbol{\Omega} = \mathrm{diag}(w_1, \ldots, w_n)$ desconhecida.

- O estimador de Mínimos Quadrados Ordinários é tal que

$$\widehat{\beta}_{\mathrm{MQO}} \xrightarrow{\mathbb{P}} \boldsymbol{\beta}$$
$$\widehat{\beta}_{\mathrm{MQO}} \xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}),$$

todavia o mesmo não é eficiente.

- White (1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica*) propõe utilizar $\widehat{\Omega} = \mathrm{diag}(\hat{e}_1^2, \ldots, \hat{e}_n^2)$ e estimar a matriz de variância-covariâncias por

$$\widehat{\Sigma}_{\widehat{\beta}} = \mathrm{QMRes}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\widehat{\boldsymbol{\Omega}}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}.$$

- Tal procedimento gera estimadores consistentes para os erros-padrão dos estimadores dos coeficientes de regressão. Este estimador é denominado HC0.

# Heteroscedasticidade de forma desconhecida

- Considere o MRLM

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \tag{5}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{\Omega})$, em que $\boldsymbol{\Omega} = \mathrm{diag}(w_1, \ldots, w_n)$ desconhecida.

- O estimador de Mínimos Quadrados Ordinários é tal que

$$\begin{aligned} \widehat{\beta}_{\mathrm{MQO}} &\xrightarrow{\mathbb{P}} \boldsymbol{\beta} \\ \widehat{\beta}_{\mathrm{MQO}} &\xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}), \end{aligned}$$

todavia o mesmo não é eficiente.

- White (1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica*) propõe utilizar $\widehat{\boldsymbol{\Omega}} = \mathrm{diag}(\hat{e}_1^2, \ldots, \hat{e}_n^2)$ e estimar a matriz de variância-covariâncias por

$$\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\beta}}} = \mathrm{QMRes}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\widehat{\boldsymbol{\Omega}}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}.$$

- Tal procedimento gera estimadores consistentes para os erros-padrão dos estimadores dos coeficientes de regressão. Este estimador é denominado HC0.

# Heteroscedasticidade de forma desconhecida

- Considere o MRLM
$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}, \tag{5}$$
em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{\Omega})$, em que $\boldsymbol{\Omega} = \mathrm{diag}(w_1, \ldots, w_n)$ desconhecida.
- O estimador de Mínimos Quadrados Ordinários é tal que
$$\begin{aligned} \widehat{\beta}_{\mathrm{MQO}} &\xrightarrow{\mathbb{P}} \boldsymbol{\beta} \\ \widehat{\beta}_{\mathrm{MQO}} &\xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}), \end{aligned}$$
todavia o mesmo não é eficiente.
- White (1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica*) propõe utilizar $\widehat{\boldsymbol{\Omega}} = \mathrm{diag}(\hat{e}_1^2, \ldots, \hat{e}_n^2)$ e estimar a matriz de variância-covariâncias por
$$\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\beta}}} = \mathrm{QMRes}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\widehat{\boldsymbol{\Omega}}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}.$$
- Tal procedimento gera estimadores consistentes para os erros-padrão dos estimadores dos coeficientes de regressão. Este estimador é denominado $\mathrm{HC0}$.

# Algumas propostas de refinamento

- Mackkinnon and White (1985, Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties, *Journal of Econometrics*):

$$\begin{aligned} \mathrm{HC1}: w_i &= \frac{n}{n-p}\hat{e}_i^2, \\ \mathrm{HC2}: w_i &= \frac{\hat{e}_i^2}{1-h_{ii}}, \\ \mathrm{HC3}: w_i &= \frac{\hat{e}_i^2}{(1-h_{ii})^2}, \end{aligned}$$

  em que $h_{ii}$ representa a alavancagem da $i$-ésima observação. Tais estimadores melhoram a perfomance em pequenas amostras.

- Cribari-Neto (2004, Asymptotic inference under heteroskedasticity of unknown form, *Computational Statistics and Data Analysis*):

$$\mathrm{HC4}: w_i = \frac{\hat{e}_i^2}{(1-h_{ii})^{\delta_i}}, \ \delta_i := \min\left\{4, \frac{h_{ii}}{\bar{h}}\right\},$$

  tal estimador melhora a perfomance em pequenas amostras, especialmente quando se tem observações influentes.

# Algumas propostas de refinamento

- Mackkinnon and White (1985, Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties, *Journal of Econometrics*):

$$\begin{aligned} \mathrm{HC1}: w_i &= \frac{n}{n-p}\hat{e}_i^2, \\ \mathrm{HC2}: w_i &= \frac{\hat{e}_i^2}{1-h_{ii}}, \\ \mathrm{HC3}: w_i &= \frac{\hat{e}_i^2}{(1-h_{ii})^2}, \end{aligned}$$

em que $h_{ii}$ representa a alavancagem da $i$-ésima observação. Tais estimadores melhoram a perfomance em pequenas amostras.

- Cribari-Neto (2004, Asymptotic inference under heteroskedasticity of unknown form, *Computational Statistics and Data Analysis*):

$$\mathrm{HC4}: w_i = \frac{\hat{e}_i^2}{(1-h_{ii})^{\delta_i}},\ \delta_i := \min\left\{4, \frac{h_{ii}}{\bar{h}}\right\},$$

tal estimador melhora a perfomance em pequenas amostras, especialmente quando se tem observações influentes.

# Algumas propostas de refinamento

- Outras propostas de refinamento podem ser vistas em:
- HC5 :Cribari-Neto, F., Souza, T. and Vasconcelos, K. (2007). Inference Under Heteroskedasticity and Leveraged Data. *Communications Communications in Statistics - Theory and Methods*, **36**, 1877–1888. doi:10.1080/03610920601126589.

$$\mathrm{HC5}: w_i = \frac{\hat{e}_i^2}{(1-h_{ii})^{\alpha_i}},\ \alpha_i := \min\left\{\frac{h_{ii}}{\bar{h}}, \max\left\{\frac{kh_{\max}}{\bar{h}}\right\}\right\},$$

em que $0 < k < 1$ é uma constante pré-definida (os autores sugerem usar $k = 0,7$) e $h_{\max} = \max\{h_{11}, \ldots, h_{nn}\}$.

- HC4m :Cribari-Neto, F. and da Silva, W.B. (2011). A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model. *AStA Advances in Statistical Analysis*, **95**, 129–146.

$$\mathrm{HC4m}: w_i = \frac{\hat{e}_i^2}{(1-h_{ii})^{\delta_i}},\ \delta_i := \min\left\{\gamma_1, \frac{nh_{ii}}{p}\right\} + \min\left\{\gamma_2, \frac{nh_{ii}}{p}\right\}.$$

em que $\gamma_1, \gamma_2$ são constantes positivas pré-definidas. Os autores sugerem usar $\gamma_1 = 1$ e $\gamma_2 = 1,5$.

# Algumas propostas de refinamento

- Outras propostas de refinamento podem ser vistas em:
- HC5 :Cribari-Neto, F., Souza, T. and Vasconcelos, K. (2007). Inference Under Heteroskedasticity and Leveraged Data. *Communications Communications in Statistics - Theory and Methods*, **36**, 1877–1888. doi:10.1080/03610920601126589.

$$\text{HC5} : w_i = \frac{\hat{e}_i^2}{(1-h_{ii})^{\alpha_i}},\ \alpha_i := \min\left\{\frac{h_{ii}}{\bar{h}}, \max\left\{\frac{kh_{\max}}{\bar{h}}\right\}\right\},$$

em que $0 < k < 1$ é uma constante pré-definida (os autores sugerem usar $k = 0,7$) e $h_{\max} = \max\{h_{11}, \ldots, h_{nn}\}$.

- HC4m :Cribari-Neto, F. and da Silva, W.B. (2011). A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model. *AStA Advances in Statistical Analysis*, **95**, 129–146.

$$\text{HC4m} : w_i = \frac{\hat{e}_i^2}{(1-h_{ii})^{\delta_i}},\ \delta_i := \min\left\{\gamma_1, \frac{nh_{ii}}{p}\right\} + \min\left\{\gamma_2, \frac{nh_{ii}}{p}\right\}.$$

em que $\gamma_1, \gamma_2$ são constantes positivas pré-definidas. Os autores sugerem usar $\gamma_1 = 1$ e $\gamma_2 = 1,5$.

# Algumas propostas de refinamento

- Outras propostas de refinamento podem ser vistas em:
- $\mathrm{HC5}$ :Cribari-Neto, F., Souza, T. and Vasconcelos, K. (2007). Inference Under Heteroskedasticity and Leveraged Data. *Communications Communications in Statistics - Theory and Methods*, **36**, 1877–1888. doi:10.1080/03610920601126589.

$$\mathrm{HC5} : w_i = \frac{\hat{e}_i^2}{(1-h_{ii})^{\alpha_i}},\ \alpha_i := \min\left\{\frac{h_{ii}}{\bar{h}}, \max\left\{\frac{kh_{\max}}{\bar{h}}\right\}\right\},$$

em que $0 < k < 1$ é uma constante pré-definida (os autores sugerem usar $k = 0,7$) e $h_{\max} = \max\{h_{11}, \ldots, h_{nn}\}$.

- $\mathrm{HC4m}$ :Cribari-Neto, F. and da Silva, W.B. (2011). A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model. *AStA Advances in Statistical Analysis*, **95**, 129–146.

$$\mathrm{HC4m} : w_i = \frac{\hat{e}_i^2}{(1-h_{ii})^{\delta_i}},\ \delta_i := \min\left\{\gamma_1, \frac{nh_{ii}}{p}\right\} + \min\left\{\gamma_2, \frac{nh_{ii}}{p}\right\}.$$

em que $\gamma_1, \gamma_2$ são constantes positivas pré-definidas. Os autores sugerem usar $\gamma_1 = 1$ e $\gamma_2 = 1,5$.

# Aspectos computacionais

- No software `R` podemos usar a library `'sandwich'` para obter os estimadores $\mathrm{HC0}, \mathrm{HC1}, \mathrm{HC2}, \mathrm{HC3},\ \mathrm{HC4},\ \mathrm{HC5}$ e $\mathrm{HC4m}$.
- Para detalhes consultar `?sandwich`. 😉
- `https://cran.r-project.org/web/packages/sandwich/vignettes/sandwich.pdf`

# Aspectos computacionais

- No software R podemos usar a library 'sandwich' para obter os estimadores $\mathrm{HC0}, \mathrm{HC1}, \mathrm{HC2}, \mathrm{HC3}$, $\mathrm{HC4}$, $\mathrm{HC5}$ e $\mathrm{HC4m}$.
- Para detalhes consultar `?sandwich`.
- https://cran.r-project.org/web/packages/sandwich/vignettes/sandwich.pdf

# Aspectos computacionais

- No software `R` podemos usar a library `'sandwich'` para obter os estimadores $\mathrm{HC0}, \mathrm{HC1}, \mathrm{HC2}, \mathrm{HC3}$, $\mathrm{HC4}$, $\mathrm{HC5}$ e $\mathrm{HC4m}$.
- Para detalhes consultar `?sandwich`. 😉
- https://cran.r-project.org/web/packages/sandwich/vignettes/sandwich.pdf

# Exercício (entregar próxima segunda-feira)

**Exercício 1**: Leia o Capítulo 6 de Hoffman (2016).

**Exercício 2**: Faça **todos** os exercícios do Capítulo 6 de Hoffman (2016).

**Exercício 3**: Reproduza o exemplo de Cribari-Neto (2004, Asymptotic inference under heteroskedasticity of unknown form, *Computational Statistics and Data Analysis*)

# Mínimos Quadrados Generalizados

- Anteriormente, apresentamos o procedimento a ser usado quando há heteroscedasticidade, mas admitimos que as covariâncias entre as fontes de variação de duas observações distintas eram todas iguais a zero, fazendo com que a matriz de variâncias e covariâncias do vetor de fontes de variação $\boldsymbol{e}$ fosse uma matriz diagonal.
- Vamos analisar o caso mais geral, em que se admite a possibilidade de que as fontes de variação de diferentes observações sejam correlacionadas.
- Para isto, consideremos a seguinte forma funcional do MRLM

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}, \tag{6}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{V})$, i.e., $\mathbb{E}[\boldsymbol{e}] = \boldsymbol{0}$ e $\mathbb{E}[\boldsymbol{e}\boldsymbol{e}^\top] = \sigma^2\boldsymbol{V}$, em que $\boldsymbol{V}$ é uma matriz positiva-definida não necessariamente diagonal.

# Mínimos Quadrados Generalizados

- Anteriormente, apresentamos o procedimento a ser usado quando há heteroscedasticidade, mas admitimos que as covariâncias entre as fontes de variação de duas observações distintas eram todas iguais a zero, fazendo com que a matriz de variâncias e covariâncias do vetor de fontes de variação $\boldsymbol{e}$ fosse uma matriz diagonal.
- Vamos analisar o caso mais geral, em que se admite a possibilidade de que as fontes de variação de diferentes observações sejam correlacionadas.
- Para isto, consideremos a seguinte forma funcional do MRLM

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}, \tag{6}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{V})$, i.e., $\mathbb{E}[\boldsymbol{e}] = \boldsymbol{0}$ e $\mathbb{E}[\boldsymbol{e}\boldsymbol{e}^\top] = \sigma^2\boldsymbol{V}$, em que $\boldsymbol{V}$ é uma matriz positiva-definida não necessariamente diagonal.

# Mínimos Quadrados Generalizados

- Anteriormente, apresentamos o procedimento a ser usado quando há heteroscedasticidade, mas admitimos que as covariâncias entre as fontes de variação de duas observações distintas eram todas iguais a zero, fazendo com que a matriz de variâncias e covariâncias do vetor de fontes de variação $\boldsymbol{e}$ fosse uma matriz diagonal.
- Vamos analisar o caso mais geral, em que se admite a possibilidade de que as fontes de variação de diferentes observações sejam correlacionadas.
- Para isto, consideremos a seguinte forma funcional do MRLM

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \tag{6}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{V})$, i.e., $\mathbb{E}[\boldsymbol{e}] = \boldsymbol{0}$ e $\mathbb{E}[\boldsymbol{e}\boldsymbol{e}^\top] = \sigma^2\boldsymbol{V}$, em que $\boldsymbol{V}$ é uma matriz positiva-definida não necessariamente diagonal.

# Mínimos Quadrados Generalizados

- Como $\boldsymbol{V}$ é uma matriz é uma matriz positiva-definida, e consequentemente não singular, então $\boldsymbol{V}^{-1}$ é bem-definida e também é positiva definida.

- Portanto, usando as decomposições usuais, Cholesky, SVD, espectral, p.e., é possível mostrar que existe uma matriz $\boldsymbol{\Lambda}$ tal que

$$\boldsymbol{V}^{-1} = \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda},$$

i.e., $\boldsymbol{\Lambda} = \boldsymbol{V}^{-1/2}$.

- Dada a obtenção de $\boldsymbol{\Lambda}$, então pré-multiplicamos todos os elementos de (6) por $\boldsymbol{\Lambda}$ de forma a obter

$$\begin{aligned} \boldsymbol{\Lambda y} &= \boldsymbol{\Lambda X}\boldsymbol{\beta} + \boldsymbol{\Lambda e}, \\ \boldsymbol{y}^* &= \boldsymbol{X}^*\boldsymbol{\beta} + \boldsymbol{e}^* \end{aligned} \tag{7}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{I})$, ou seja, agora estamos sob um MRLM homoscedástico com vetor de variável resposta $\boldsymbol{y}^*$, matriz de especificação $\boldsymbol{X}^*$ e vetor de fontes de variação $\boldsymbol{e}^*$.

# Mínimos Quadrados Generalizados

- Como $\boldsymbol{V}$ é uma matriz é uma matriz positiva-definida, e consequentemente não singular, então $\boldsymbol{V}^{-1}$ é bem-definida e também é positiva definida.
- Portanto, usando as decomposições usuais, Cholesky, SVD, espectral, p.e., é possível mostrar que existe uma matriz $\boldsymbol{\Lambda}$ tal que

$$\boldsymbol{V}^{-1} = \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda},$$

i.e., $\boldsymbol{\Lambda} = \boldsymbol{V}^{-1/2}$.

- Dada a obtenção de $\boldsymbol{\Lambda}$, então pré-multiplicamos todos os elementos de (6) por $\boldsymbol{\Lambda}$ de forma a obter

$$\boldsymbol{\Lambda y} = \boldsymbol{\Lambda X}\boldsymbol{\beta} + \boldsymbol{\Lambda e},$$
$$\boldsymbol{y}^* = \boldsymbol{X}^*\boldsymbol{\beta} + \boldsymbol{e}^* \quad (7)$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2\boldsymbol{I})$, ou seja, agora estamos sob um MRLM homoscedástico com vetor de variável resposta $\boldsymbol{y}^*$, matriz de especificação $\boldsymbol{X}^*$ e vetor de fontes de variação $\boldsymbol{e}^*$.

# Mínimos Quadrados Generalizados

- Como $\boldsymbol{V}$ é uma matriz é uma matriz positiva-definida, e consequentemente não singular, então $\boldsymbol{V}^{-1}$ é bem-definida e também é positiva definida.
- Portanto, usando as decomposições usuais, Cholesky, SVD, espectral, p.e., é possível mostrar que existe uma matriz $\boldsymbol{\Lambda}$ tal que

$$\boldsymbol{V}^{-1} = \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda},$$

i.e., $\boldsymbol{\Lambda} = \boldsymbol{V}^{-1/2}$.

- Dada a obtenção de $\boldsymbol{\Lambda}$, então pré-multiplicamos todos os elementos de (6) por $\boldsymbol{\Lambda}$ de forma a obter

$$\begin{aligned} \boldsymbol{\Lambda y} &= \boldsymbol{\Lambda X}\beta + \boldsymbol{\Lambda e}, \\ \boldsymbol{y}^* &= \boldsymbol{X}^*\beta + \boldsymbol{e}^* \end{aligned} \tag{7}$$

em que $\boldsymbol{e} \sim (\boldsymbol{0}, \sigma^2 \boldsymbol{I})$, ou seja, agora estamos sob um MRLM homoscedástico com vetor de variável resposta $\boldsymbol{y}^*$, matriz de especificação $\boldsymbol{X}^*$ e vetor de fontes de variação $\boldsymbol{e}^*$.

# Mínimos Quadrados Generalizados

- Quando consideramos $\boldsymbol{V}$ conhecida o procedimento é similar ao procedimentro utilizado no método de MQP, com a diferença de que $\boldsymbol{V}$ pode ser uma matriz PD qualquer, tornando o processo de obtenção de $\boldsymbol{\Lambda} := \boldsymbol{V}^{-1/2}$ algebricamente mais complicado e como consequência, não conseguimos expressar a função objetivo e os respectivos estimadores em termos de somas de forma simples como acontecia no caso anterior.
- Dado que se tem um MRLM homoscedástico nas variáveis transformadas, podemos utilizar todos os resultados já vistos anteriormente. Por exemplo, o estimador de Mínimos Quadrados Generalizados (MQG) é dado por

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

com respectiva matriz de variâncias-covariâncias dada por

$$\sigma^2(\boldsymbol{X}^\top \boldsymbol{V}^{-1}\boldsymbol{X})^{-1}.$$

# Mínimos Quadrados Generalizados

- Quando consideramos $\boldsymbol{V}$ conhecida o procedimento é similar ao procedimentro utilizado no método de MQP, com a diferença de que $\boldsymbol{V}$ pode ser uma matriz PD qualquer, tornando o processo de obtenção de $\boldsymbol{\Lambda} := \boldsymbol{V}^{-1/2}$ algebricamente mais complicado e como consequência, não conseguimos expressar a função objetivo e os respectivos estimadores em termos de somas de forma simples como acontecia no caso anterior.
- Dado que se tem um MRLM homoscedástico nas variáveis transformadas, podemos utilizar todos os resultados já vistos anteriormente. Por exemplo, o estimador de Mínimos Quadrados Generalizados (MQG) é dado por

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

com respectiva matriz de variâncias-covariâncias dada por

$$\sigma^2(\boldsymbol{X}^\top \boldsymbol{V}^{-1}\boldsymbol{X})^{-1}.$$

# Comentários

- De forma similar ao MQP, temos por exemplo que o MINQUE de $\sigma^2$ é dado por:

$$\begin{aligned}\hat{\sigma}^2 = \text{QMRes} &= \frac{\boldsymbol{y}^\top \boldsymbol{\Lambda}^\top [\boldsymbol{I} - \boldsymbol{\Lambda}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{\Lambda}^\top]\boldsymbol{\Lambda}\boldsymbol{y}}{n-p} \\ &= \frac{\boldsymbol{y}^\top [\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{V}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{V}^{-1}]\boldsymbol{y}}{n-p}.\end{aligned}$$

- Assim como no caso anterior, o EMQO continua não viesado e consistente, todavia ele não é mais o BLUE, perdendo assim sua eficiência.
- Note que o método de mínimos quadrados ponderados, estudado anteriormente, é um caso particular do método de mínimos quadrados generalizados, em que a matriz $\boldsymbol{V}$ é diagonal e quando $\boldsymbol{V} = \boldsymbol{I}_n$, obtemos o caso associado ao método de mínimos quadrados ordinários.
- No software `R` podemos usar a função `'gls'` dentro do pacote `nlme` para utilizar o método de mínimos quadrados generalizados.

# Comentários

- De forma similar ao MQP, temos por exemplo que o MINQUE de $\sigma^2$ é dado por:

$$\hat{\sigma}^2 = \mathrm{QMRes} = \frac{\boldsymbol{y}^\top \boldsymbol{\Lambda}^\top [\boldsymbol{I} - \boldsymbol{\Lambda}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{\Lambda}^\top]\boldsymbol{\Lambda}\boldsymbol{y}}{n-p} = \frac{\boldsymbol{y}^\top [\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{V}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{V}^{-1}]\boldsymbol{y}}{n-p}.$$

- Assim como no caso anterior, o EMQO continua não viesado e consistente, todavia ele não é mais o BLUE, perdendo assim sua eficiência.

- Note que o método de mínimos quadrados ponderados, estudado anteriormente, é um caso particular do método de mínimos quadrados generalizados, em que a matriz $\boldsymbol{V}$ é diagonal e quando $\boldsymbol{V} = \boldsymbol{I}_n$, obtemos o caso associado ao método de mínimos quadrados ordinários.

- No software `R` podemos usar a função `'gls'` dentro do pacote `nlme` para utilizar o método de mínimos quadrados generalizados.

# Comentários

- De forma similar ao MQP, temos por exemplo que o MINQUE de $\sigma^2$ é dado por:

$$\hat{\sigma}^2 = \mathrm{QMRes} = \frac{\boldsymbol{y}^\top \boldsymbol{\Lambda}^\top [\boldsymbol{I} - \boldsymbol{\Lambda}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{\Lambda}^\top]\boldsymbol{\Lambda}\boldsymbol{y}}{n-p} = \frac{\boldsymbol{y}^\top [\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{V}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{V}^{-1}]\boldsymbol{y}}{n-p}.$$

- Assim como no caso anterior, o EMQO continua não viesado e consistente, todavia ele não é mais o BLUE, perdendo assim sua eficiência.
- Note que o método de mínimos quadrados ponderados, estudado anteriormente, é um caso particular do método de mínimos quadrados generalizados, em que a matriz $\boldsymbol{V}$ é diagonal e quando $\boldsymbol{V} = \boldsymbol{I}_n$, obtemos o caso associado ao método de mínimos quadrados ordinários.
- No software `R` podemos usar a função `'gls'` dentro do pacote `nlme` para utilizar o método de mínimos quadrados generalizados.

# Comentários

- De forma similar ao MQP, temos por exemplo que o MINQUE de $\sigma^2$ é dado por:

$$\hat{\sigma}^2 = \text{QMRes} = \frac{\boldsymbol{y}^\top \boldsymbol{\Lambda}^\top [\boldsymbol{I} - \boldsymbol{\Lambda X}(\boldsymbol{X}^\top \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda X})^{-1})\boldsymbol{X}^\top \boldsymbol{\Lambda}^\top]\boldsymbol{\Lambda y}}{n-p} = \frac{\boldsymbol{y}^\top [\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{V}^{-1}\boldsymbol{X})^{-1})\boldsymbol{X}^\top \boldsymbol{V}^{-1}]\boldsymbol{y}}{n-p}.$$

- Assim como no caso anterior, o EMQO continua não viesado e consistente, todavia ele não é mais o BLUE, perdendo assim sua eficiência.
- Note que o método de mínimos quadrados ponderados, estudado anteriormente, é um caso particular do método de mínimos quadrados generalizados, em que a matriz $\boldsymbol{V}$ é diagonal e quando $\boldsymbol{V} = \boldsymbol{I}_n$, obtemos o caso associado ao método de mínimos quadrados ordinários.
- No software `R` podemos usar a função `'gls'` dentro do pacote `nlme` para utilizar o método de mínimos quadrados generalizados.

# Comentários

- Na prática, não conhecemos a matriz $\boldsymbol{V}$ e como consequência, precisamos estimá-la. É comum utilizar um processo em dois estágios.
- Primeiramente, estimamos $\beta$ considerando $\boldsymbol{V}$ conhecida e posteriormente, estima-se $\boldsymbol{V}$ e considera-se o estimador

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

tal estimador não é mais linear e consequentemente não é o BLUE. Todavia, sob fracas condições a respeito do estimador de $\boldsymbol{V}$, o estimador acima (EBLUE) continua não viesado (Demidenko 2013, Mixed Models, Lemma 16, pag. 138).

# Comentários

- Na prática, não conhecemos a matriz $\boldsymbol{V}$ e como consequência, precisamos estimá-la. É comum utilizar um processo em dois estágios.

- Primeiramente, estimamos $\beta$ considerando $\boldsymbol{V}$ conhecida e posteriormente, estima-se $\boldsymbol{V}$ e considera-se o estimador

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

  tal estimador não é mais linear e consequentemente não é o BLUE. Todavia, sob fracas condições a respeito do estimador de $\boldsymbol{V}$, o estimador acima (EBLUE) continua não viesado (Demidenko 2013, Mixed Models, Lemma 16, pag. 138).

# Comentários

- A matriz simétrica $\boldsymbol{V}$ pode ter até $n(n+1)/2$ elementos possivelmente distintos que precisamos conhecer, para poder aplicar os resultados anteriores. Quando a matriz possui exatamente $n(n+1)/2$ elementos distintos, diz-se que ela é não estruturada.
- Na prática é comum parametrizar esta matriz em função de uma quantidade menor de parâmetros, gerando estruturas paramétricas que facilitam o processo de estimação. Vamos considerar aqui uma estrutura que é bastante utilizada quando se tem um processo estocástico particular, gerando fontes de variação autocorrelacionadas.
- Na função `'gls'` é possível utilizar diversas estruturas fora a que será apresentada.

# Comentários

- A matriz simétrica $\boldsymbol{V}$ pode ter até $n(n+1)/2$ elementos possivelmente distintos que precisamos conhecer, para poder aplicar os resultados anteriores. Quando a matriz possui exatamente $n(n+1)/2$ elementos distintos, diz-se que ela é não estruturada.
- Na prática é comum parametrizar esta matriz em função de uma quantidade menor de parâmetros, gerando estruturas paramétricas que facilitam o processo de estimação. Vamos considerar aqui uma estrutura que é bastante utilizada quando se tem um processo estocástico particular, gerando fontes de variação autocorrelacionadas.
- Na função `'gls'` é possível utilizar diversas estruturas fora a que será apresentada.

# Comentários

- A matriz simétrica $\boldsymbol{V}$ pode ter até $n(n+1)/2$ elementos possivelmente distintos que precisamos conhecer, para poder aplicar os resultados anteriores. Quando a matriz possui exatamente $n(n+1)/2$ elementos distintos, diz-se que ela é não estruturada.
- Na prática é comum parametrizar esta matriz em função de uma quantidade menor de parâmetros, gerando estruturas paramétricas que facilitam o processo de estimação. Vamos considerar aqui uma estrutura que é bastante utilizada quando se tem um processo estocástico particular, gerando fontes de variação autocorrelacionadas.
- Na função `'gls'` é possível utilizar diversas estruturas fora a que será apresentada.

# Autocorrelação

- Para ilustrar a situação de autocorrelação, consideremos o modelo (6) em que

$$e_t = \rho e_{t-1} + u_t, t \geq 1, \tag{8}$$

em que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, $e_0 := 0$ e $|\rho| < 1$.

- Este processo é denominado processo auto-regressivo de ordem 1 de parâmetro $\rho$ e é denotado por $\mathrm{AR1}(\rho)$. A condição de que $|\rho| < 1$ é para garantir que o processo estocástico associado seja estacionário.
- Utilizamos a letra $t$ para indicar o índice associado às diferentes observações, pois a autocorrelação surge, geralmente, quando estamos trabalhando com séries temporais/ dados longitudinais, em que cada observação corresponde a um certo período em uma escala ordenada (ano, mês, dia, profundidade, etc).
- Por (8) tem-se que que a fonte de variação da observação relativa a um período está correlacionado com a fonte de variação da observação anterior, especificamente

$$\mathrm{Cor}(e_t, e_{t-1}) = \rho.$$

# Autocorrelação

- Para ilustrar a situação de autocorrelação, consideremos o modelo (6) em que

$$e_t = \rho e_{t-1} + u_t, t \geq 1, \tag{8}$$

em que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, $e_0 := 0$ e $|\rho| < 1$.

- Este processo é denominado processo auto-regressivo de ordem 1 de parâmetro $\rho$ e é denotado por $\mathrm{AR1}(\rho)$. A condição de que $|\rho| < 1$ é para garantir que o processo estocástico associado seja estacionário.

- Utilizamos a letra $t$ para indicar o índice associado às diferentes observações, pois a autocorrelação surge, geralmente, quando estamos trabalhando com séries temporais/ dados longitudinais, em que cada observação corresponde a um certo período em uma escala ordenada (ano, mês, dia, profundidade, etc).

- Por (8) tem-se que que a fonte de variação da observação relativa a um período está correlacionado com a fonte de variação da observação anterior, especificamente

$$\mathrm{Cor}(e_t, e_{t-1}) = \rho.$$

# Autocorrelação

- Para ilustrar a situação de autocorrelação, consideremos o modelo (6) em que

$$e_t = \rho e_{t-1} + u_t, t \geq 1, \tag{8}$$

em que $u_t \stackrel{\text{ind}}{\sim} (0, \sigma^2)$, $e_0 := 0$ e $|\rho| < 1$.

- Este processo é denominado processo auto-regressivo de ordem 1 de parâmetro $\rho$ e é denotado por $\mathrm{AR1}(\rho)$. A condição de que $|\rho| < 1$ é para garantir que o processo estocástico associado seja estacionário.

- Utilizamos a letra $t$ para indicar o índice associado às diferentes observações, pois a autocorrelação surge, geralmente, quando estamos trabalhando com séries temporais/ dados longitudinais, em que cada observação corresponde a um certo período em uma escala ordenada (ano, mês, dia, profundidade, etc).

- Por (8) tem-se que que a fonte de variação da observação relativa a um período está correlacionado com a fonte de variação da observação anterior, especificamente

$$\mathrm{Cor}(e_t, e_{t-1}) = \rho.$$

# Autocorrelação

- Para ilustrar a situação de autocorrelação, consideremos o modelo (6) em que

$$e_t = \rho e_{t-1} + u_t, t \geq 1, \tag{8}$$

em que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, $e_0 := 0$ e $|\rho| < 1$.

- Este processo é denominado processo auto-regressivo de ordem 1 de parâmetro $\rho$ e é denotado por $\mathrm{AR1}(\rho)$. A condição de que $|\rho| < 1$ é para garantir que o processo estocástico associado seja estacionário.
- Utilizamos a letra $t$ para indicar o índice associado às diferentes observações, pois a autocorrelação surge, geralmente, quando estamos trabalhando com séries temporais/ dados longitudinais, em que cada observação corresponde a um certo período em uma escala ordenada (ano, mês, dia, profundidade, etc).
- Por (8) tem-se que que a fonte de variação da observação relativa a um período está correlacionado com a fonte de variação da observação anterior, especificamente

$$\mathrm{Cor}(e_t, e_{t-1}) = \rho.$$

# Autocorrelação

- Se $\rho > 0$ dizemos que as fontes de variação são positivamente autocorrelacionados e se $\rho < 0$ dizemos que há autocorrelação negativa.
- Se $\rho = 0$ teremos, obviamente, o MRLM homoscedástico usual.
- Se $\rho = \pm 1$ teremos essencialmente um passeio aleatório.
- No caso em que $\rho = 1$, tem-se

$$e_t = e_{t-1} + u_t \tag{9}$$

implicando que $e_t - e_{t-1} = u_t$.

# Autocorrelação

- Se $\rho > 0$ dizemos que as fontes de variação são positivamente autocorrelacionados e se $\rho < 0$ dizemos que há autocorrelação negativa.
- Se $\rho = 0$ teremos, obviamente, o MRLM homoscedástico usual.
- Se $\rho = \pm 1$ teremos essencialmente um passeio aleatório.
- No caso em que $\rho = 1$, tem-se

$$e_t = e_{t-1} + u_t \tag{9}$$

implicando que $e_t - e_{t-1} = u_t$.

# Autocorrelação

- Se $\rho > 0$ dizemos que as fontes de variação são positivamente autocorrelacionados e se $\rho < 0$ dizemos que há autocorrelação negativa.
- Se $\rho = 0$ teremos, obviamente, o MRLM homoscedástico usual.
- Se $\rho = \pm 1$ teremos essencialmente um passeio aleatório.
- No caso em que $\rho = 1$, tem-se

$$e_t = e_{t-1} + u_t \tag{9}$$

implicando que $e_t - e_{t-1} = u_t$.

# Autocorrelação

- Se $\rho > 0$ dizemos que as fontes de variação são positivamente autocorrelacionados e se $\rho < 0$ dizemos que há autocorrelação negativa.
- Se $\rho = 0$ teremos, obviamente, o MRLM homoscedástico usual.
- Se $\rho = \pm 1$ teremos essencialmente um passeio aleatório.
- No caso em que $\rho = 1$, tem-se

$$e_t = e_{t-1} + u_t \tag{9}$$

implicando que $e_t - e_{t-1} = u_t$.

# Autocorrelação

- Logo, considerando o modelo (6) com a estrutura (9) tem-se que

$$\begin{aligned} y_t - y_{t-1} &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + (e_t - e_{t-1}) \\ &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + u_t, \forall t \geq 2. \end{aligned}$$

- Lembrando que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, logo podemos fazer uso de todos os resultados associados ao MRLM homoscesdástico usual.
- Note que no modelo acima a variável dependente é dada por $\Delta y_t := y_t - y_{t-1}$ e as variáveis explicativas são $\Delta x_{it} := x_{it} - x_{i(t-1)}$, $i = 1, \ldots, p$, de forma que o número de observações se reduz a $n - 1$.
- Perceba que o modelo transformado não possui intercepto.
- Para o caso em que $\rho = -1$, obtemos de forma similar o seguinte modelo transformado

$$y_t + y_{t-1} = \sum_{j=1}^{p} \beta_j (x_{jt} + x_{j(t-1)}) + u_t, \forall t \geq 2.$$

# Autocorrelação

- Logo, considerando o modelo (6) com a estrutura (9) tem-se que

$$\begin{aligned} y_t - y_{t-1} &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + (e_t - e_{t-1}) \\ &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + u_t, \forall t \geq 2. \end{aligned}$$

- Lembrando que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, logo podemos fazer uso de todos os resultados associados ao MRLM homoscesdástico usual.

- Note que no modelo acima a variável dependente é dada por $\Delta y_t := y_t - y_{t-1}$ e as variáveis explicativas são $\Delta x_{it} := x_{it} - x_{i(t-1)}$, $i = 1, \ldots, p$, de forma que o número de observações se reduz a $n - 1$.

- Perceba que o modelo transformado não possui intercepto.

- Para o caso em que $\rho = -1$, obtemos de forma similar o seguinte modelo transformado

$$y_t + y_{t-1} = \sum_{j=1}^{p} \beta_j (x_{jt} + x_{j(t-1)}) + u_t, \forall t \geq 2.$$

# Autocorrelação

- Logo, considerando o modelo (6) com a estrutura (9) tem-se que

$$\begin{aligned} y_t - y_{t-1} &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + (e_t - e_{t-1}) \\ &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + u_t, \forall t \geq 2. \end{aligned}$$

- Lembrando que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, logo podemos fazer uso de todos os resultados associados ao MRLM homoscesdástico usual.
- Note que no modelo acima a variável dependente é dada por $\Delta y_t := y_t - y_{t-1}$ e as variáveis explicativas são $\Delta x_{it} := x_{it} - x_{i(t-1)}$, $i = 1, \ldots, p$, de forma que o número de observações se reduz a $n - 1$.
- Perceba que o modelo transformado não possui intercepto.
- Para o caso em que $\rho = -1$, obtemos de forma similar o seguinte modelo transformado

$$y_t + y_{t-1} = \sum_{j=1}^{p} \beta_j (x_{jt} + x_{j(t-1)}) + u_t, \forall t \geq 2.$$

# Autocorrelação

- Logo, considerando o modelo (6) com a estrutura (9) tem-se que

$$\begin{aligned} y_t - y_{t-1} &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + (e_t - e_{t-1}) \\ &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + u_t, \forall t \geq 2. \end{aligned}$$

- Lembrando que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, logo podemos fazer uso de todos os resultados associados ao MRLM homoscesdástico usual.
- Note que no modelo acima a variável dependente é dada por $\Delta y_t := y_t - y_{t-1}$ e as variáveis explicativas são $\Delta x_{it} := x_{it} - x_{i(t-1)}$, $i = 1, \ldots, p$, de forma que o número de observações se reduz a $n - 1$.
- Perceba que o modelo transformado não possui intercepto.
- Para o caso em que $\rho = -1$, obtemos de forma similar o seguinte modelo transformado

$$y_t + y_{t-1} = \sum_{j=1}^{p} \beta_j (x_{jt} + x_{j(t-1)}) + u_t, \forall t \geq 2.$$

# Autocorrelação

- Logo, considerando o modelo (6) com a estrutura (9) tem-se que

$$\begin{aligned} y_t - y_{t-1} &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + (e_t - e_{t-1}) \\ &= \sum_{j=1}^{p} \beta_j (x_{jt} - x_{j(t-1)}) + u_t, \forall t \geq 2. \end{aligned}$$

- Lembrando que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, logo podemos fazer uso de todos os resultados associados ao MRLM homoscesdástico usual.
- Note que no modelo acima a variável dependente é dada por $\Delta y_t := y_t - y_{t-1}$ e as variáveis explicativas são $\Delta x_{it} := x_{it} - x_{i(t-1)}$, $i = 1, \ldots, p$, de forma que o número de observações se reduz a $n - 1$.
- Perceba que o modelo transformado não possui intercepto.
- Para o caso em que $\rho = -1$, obtemos de forma similar o seguinte modelo transformado

$$y_t + y_{t-1} = \sum_{j=1}^{p} \beta_j (x_{jt} + x_{j(t-1)}) + u_t, \forall t \geq 2.$$

# Autocorrelação

- Se $|\rho| < 1$, usando (8) de forma recursiva, tem-se a seguinte representação estocástica

$$e_t \overset{\mathcal{D}}{=} \sum_{r=0}^{\infty} \rho^r u_{t-r}. \tag{10}$$

- Por (10) e lembrando que $u_t \overset{\text{ind}}{\sim} (0, \sigma^2)$, tem-se que $\forall t, h \in \mathbb{N}$ e

$$\mathbb{E}[e_t] = 0, \tag{11}$$

$$\text{Var}[e_t] = \mathbb{E}[e_t^2] = \sum_{r=0}^{\infty} \rho^{2r}\sigma^2 = \sigma^2 \sum_{r=0}^{\infty} \rho^{2r} = \frac{\sigma^2}{1-\rho^2}, \tag{12}$$

$$\begin{aligned}\text{Cov}(e_t, e_{t-h}) &= \text{Cov}\left(\sum_{r_1=0}^{\infty} \rho^{r_1} u_{t-r_1}, \sum_{r_2=0}^{\infty} \rho^{r_2} u_{t-h-r_2}\right) \\ &= \sum_{r=0}^{\infty} \rho^{h+2r}\text{Var}(U_{t-h-r}) = \rho^h\sigma^2 \sum_{r=0}^{\infty} \rho^{2r} = \frac{\rho^h\sigma^2}{1-\rho^2}.\end{aligned} \tag{13}$$

# Autocorrelação

- Se $|\rho| < 1$, usando (8) de forma recursiva, tem-se a seguinte representação estocástica

$$e_t \stackrel{\mathcal{D}}{=} \sum_{r=0}^{\infty} \rho^r u_{t-r}. \tag{10}$$

- Por (10) e lembrando que $u_t \stackrel{\text{ind}}{\sim} (0, \sigma^2)$, tem-se que $\forall t, h \in \mathbb{N}$ e

$$\begin{aligned}
\mathbb{E}[e_t] &= 0, &(11)\\
\mathrm{Var}[e_t] = \mathbb{E}[e_t^2] &= \sum_{r=0}^{\infty} \rho^{2r}\sigma^2 = \sigma^2 \sum_{r=0}^{\infty} \rho^{2r} = \frac{\sigma^2}{1-\rho^2}, &(12)\\
\mathrm{Cov}(e_t, e_{t-h}) &= \mathrm{Cov}\left(\sum_{r_1=0}^{\infty} \rho^{r_1} u_{t-r_1}, \sum_{r_2=0}^{\infty} \rho^{r_2} u_{t-h-r_2}\right)\\
&= \sum_{r=0}^{\infty} \rho^{h+2r}\mathrm{Var}(U_{t-h-r}) = \rho^h\sigma^2 \sum_{r=0}^{\infty} \rho^{2r} = \frac{\rho^h\sigma^2}{1-\rho^2}. &(13)
\end{aligned}$$

# Autocorrelação

Portanto, por (11), (12) e (13) tem-se $\boldsymbol{e} \sim (\mathbf{0}, \sigma^2 \boldsymbol{V})$, em que

$$\boldsymbol{V} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix}.$$

# Autocorrelação

É possível mostrar que

$$\boldsymbol{V}^{-1} = \begin{pmatrix} 1 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{pmatrix},$$

e que $\boldsymbol{V}^{-1} = \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}$, tal que

$$\boldsymbol{\Lambda} = \begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{pmatrix}.$$

# Autocorrelação

- Usando $\boldsymbol{\Lambda}$ obtido anteriormente, obtemos o MRLM homoscedástico

$$\boldsymbol{\Lambda y} = \boldsymbol{\Lambda X \beta} + \boldsymbol{\Lambda e},$$

no qual consegue ser reescrito da seguinte forma para $t = 1$

$$\begin{aligned}\sqrt{1-\rho^2}y_1 &= \sum_{j=0}^{p-1} \beta_j \sqrt{1-\rho^2}x_{j1} + \sqrt{1-\rho^2}e_1 \\ &= \sum_{j=0}^{p-1} \beta_j \sqrt{1-\rho^2}x_{j1} + u_1\end{aligned}$$

e para, lembrando que $e_t = \rho e_{t-1} + u_t$, $t = 2, \ldots, n$

$$\begin{aligned}y_t - \rho y_{t-1} &= \sum_{j=0}^{p-1} \beta_j (x_{jt} - \rho x_{j(t-1)}) + e_t - \rho e_{t-1} \\ &= \sum_{j=0}^{p-1} \beta_j (x_{jt} - \rho x_{j(t-1)}) + u_t.\end{aligned}$$

# Autocorrelação

- Todavia, na prática $\rho$ é desconhecido. Em geral, é obtido um estimador de $\rho$ e usa-se o EBLUE de $\boldsymbol{\beta}$.
- Um estimador consistente de $\rho$ é obtido em função dos resíduos ordinários de um ajuste do MRLM através do MQO, e o mesmo é dado por

$$\widehat{\rho} = \frac{\frac{1}{n-1}\sum_{t=2}^{n} \hat{e}_t \hat{e}_{t-1}}{\frac{1}{n-p}\sum_{t=1}^{n} \hat{e}_t^2}.$$

- Na literatura, existem alguns testes para verificar a existência de autocorrelação, dentre eles um dos mais famosos é o teste de Durbin-Watson que pode ser visto em Hoffman (2016) ou Singer, Nobre e Rocha (2019), por exemplo.

# Autocorrelação

- Todavia, na prática $\rho$ é desconhecido. Em geral, é obtido um estimador de $\rho$ e usa-se o EBLUE de $\boldsymbol{\beta}$.

- Um estimador consistente de $\rho$ é obtido em função dos resíduos ordinários de um ajuste do MRLM através do MQO, e o mesmo é dado por

$$\widehat{\rho} = \frac{\frac{1}{n-1}\sum_{t=2}^{n} \hat{e}_t \hat{e}_{t-1}}{\frac{1}{n-p}\sum_{t=1}^{n} \hat{e}_t^2}.$$

- Na literatura, existem alguns testes para verificar a existência de autocorrelação, dentre eles um dos mais famosos é o teste de Durbin-Watson que pode ser visto em Hoffman (2016) ou Singer, Nobre e Rocha (2019), por exemplo.

# Autocorrelação

- Todavia, na prática $\rho$ é desconhecido. Em geral, é obtido um estimador de $\rho$ e usa-se o EBLUE de $\boldsymbol{\beta}$.

- Um estimador consistente de $\rho$ é obtido em função dos resíduos ordinários de um ajuste do MRLM através do MQO, e o mesmo é dado por

$$\widehat{\rho} = \frac{\frac{1}{n-1}\sum_{t=2}^{n} \hat{e}_t \hat{e}_{t-1}}{\frac{1}{n-p}\sum_{t=1}^{n} \hat{e}_t^2}.$$

- Na literatura, existem alguns testes para verificar a existência de autocorrelação, dentre eles um dos mais famosos é o teste de Durbin-Watson que pode ser visto em Hoffman (2016) ou Singer, Nobre e Rocha (2019), por exemplo.

# Exercícios

**Exercício 1**: Faça uma breve redação sobre modelos com variáveis dummy e sobre os métodos de mínimos quadrados ordinários, ponderados e generalizados.

**Exercício 2**: Apresente o teste de Durbin-Watson e os comandos no R para realizar o teste em questão.

**Exercício 3**: Faça todas as questões do Capítulo 7 de Hoffman (2016).

# Reflexão

*"Wir müssen wissen,*

*Wir werden wissen"*.

David Hilbert

*"Cabeça chata sim,*

*quadrada não!"*.

Autor desconhecido

# Agradecimentos

- Vocês estão prontas crianças?
- Estamos capitão!
- Vocês estão prontos para ganhar o mundo (principalmente depois de estudar modelos de regressão) !!!

# Agradecimentos

- Vocês estão prontas crianças?
- Estamos capitão!
- Vocês estão prontos para ganhar o mundo (principalmente depois de estudar modelos de regressão) !!!

# Agradecimentos

- Vocês estão prontas crianças?
- Estamos capitão!
- Vocês estão prontos para ganhar o mundo (principalmente depois de estudar modelos de regressão) !!!

# Agradecimentos

- Obrigado pelo **Excelente semestre**. ☺
- Em breve estaremos juntos novamente! ☺

# Agradecimentos

- Obrigado pelo **Excelente semestre**. 🙂
- Em breve estaremos juntos novamente! 😉