Modelos Lineares Generalizados

Estimação por mínimos quadrados generalizados

Terezinha K. A. Ribeiro

terezinha.ribeiro@unb.br

Instituto de Ciências Exatas
Departamento de Estatística
Universidade de Brasília

Introdução

O modelo de regressão linear múltiplo normal considera que $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ com $\boldsymbol{\varepsilon} \sim \mathrm{N}_n(0, \sigma^2 I_n)$ em que I_n é matriz identidade de ordem n, e $r(X) = k + 1$.

Segue que $\boldsymbol{Y} \sim \mathrm{N}_n(X\boldsymbol{\beta}, \sigma^2 I_n)$.

O que acontece se trocarmos a matriz I_n por uma matriz V positiva definida cujo formato é conhecido?

Modelo linear com $\varepsilon \sim \mathrm{N}_n(0, \sigma^2 V)$, V conhecida

Considere o modelo linear $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ com $\boldsymbol{\varepsilon} \sim \mathrm{N}_n(0, \sigma^2 V)$ em que $V \neq I_n$ é matriz positiva definida, e $r(X) = k + 1$.

Segue que $\boldsymbol{Y} \sim \mathrm{N}_n(X\boldsymbol{\beta}, \sigma^2 V)$.

Note que

- se pelo menos dois elementos da diagonal principal de V são distintos, então o modelo linear é heteroscedástico;
- se V é uma matriz diagonal, então as respostas Y_i e Y_j são não correlacionadas (independentes);
- se V não é uma matriz diagonal, então as respostas Y_i e Y_j são correlacionadas (dependentes).

Modelo linear com $\varepsilon \sim \mathrm{N}_n(0, \sigma^2 V)$, V conhecida

Como estamos supondo que V é uma matriz positiva definida, pela decomposição de Cholesky, existe uma matriz $P_{n\times n}$ não singular tal que

$$V = PP^\top.$$

Considere a transformação $\boldsymbol{Z} = P^{-1}\boldsymbol{Y}$. Veja que

$$\begin{aligned}
\mathrm{E}(\boldsymbol{Z}) &= P^{-1}\,\mathrm{E}(\boldsymbol{Y}) = P^{-1}X\beta, \\
\\
\mathrm{Var}(\boldsymbol{Z}) &= P^{-1}\,\mathrm{Var}(\boldsymbol{Y})P^{-1\top} \\
&= P^{-1}\sigma^2 V P^{-1\top} \\
&= \sigma^2(P^{-1}P)(P^\top P^{-1\top}) \\
&= \sigma^2(P^{-1}P)(P^\top P^{\top -1}) \\
&= \sigma^2 I_n.
\end{aligned}$$

Modelo linear com $\varepsilon \sim \mathrm{N}_n(0, \sigma^2 V)$, V conhecida

Como P é de posto completo (pois P é não singular), segue que

$$\boldsymbol{Z} = P^{-1}\boldsymbol{Y} \sim \mathrm{N}_n(P^{-1}X\boldsymbol{\beta}, \sigma^2 I_n).$$

Assim, o modelo linear transformado é dado por

$$\boldsymbol{Z} = Q\boldsymbol{\beta} + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathrm{N}_n(\mathbf{0}, \sigma^2 I_n), \tag{1}$$

em que $Q_{n\times(k+1)} = P^{-1}X$ é a matriz do modelo de posto completo pois $r(Q) = r(P^{-1}X) = r(X) = k+1$, e $\boldsymbol{\eta} = P^{-1}\boldsymbol{\varepsilon}$ é o vetor de erros aleatórios.

O modelo linear definido em (1) é homoscedástico e satisfaz todas as condições da regressão linear múltipla normal.

Ainda, veja que os parâmetros de interesse $\boldsymbol{\beta}$ e σ^2 do modelo linear original são os mesmos do modelo transformado.

Modelo linear com $\varepsilon \sim \mathrm{N}_n(0, \sigma^2 V)$, V conhecida

Sendo assim, generalizaremos os resultados inferenciais obtidos sob a regressão linear normal.

Assim, o estimador de mínimos quadrados para β sob o modelo transformado (1) é

$$\widehat{\beta} = (Q^\top Q)^{-1} Q^\top \boldsymbol{Z}.$$

Voltando, ao modelo linear original temos que o estimador para β fica dado por

$$\begin{aligned}\widehat{\beta}_G &= [(P^{-1}X)^\top P^{-1}X]^{-1}(P^{-1}X)^\top P^{-1}\boldsymbol{Y} \\ &= [X^\top (P^{-1})^\top P^{-1}X]^{-1} X^\top (P^{-1})^\top P^{-1}\boldsymbol{Y}.\end{aligned}$$

De $V = PP^\top$ tem-se que $V^{-1} = P^{\top^{-1}} P^{-1} = P^{-1^\top} P^{-1}$.

Modelo linear com $\varepsilon \sim \mathrm{N}_n(0, \sigma^2 V)$, V conhecida

Assim, obtemos que

$$\widehat{\boldsymbol{\beta}}_G = (X^\top V^{-1} X)^{-1} X^\top V^{-1} \boldsymbol{Y}.$$

Note que $\widehat{\boldsymbol{\beta}}_G \neq \widehat{\boldsymbol{\beta}} = (X^\top X)^{-1} X^\top \boldsymbol{Y}$.

Sob o modelo linear original $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ com $\boldsymbol{\varepsilon} \sim \mathrm{N}_n(0, \sigma^2 V)$, o estimador $\widehat{\boldsymbol{\beta}}_G$ é chamado de estimador de mínimos quadrados generalizados de $\boldsymbol{\beta}$.

Sob o modelo linear transformado $\boldsymbol{Z} = Q\boldsymbol{\beta} + \boldsymbol{\eta}$ com $\boldsymbol{\eta} \sim \mathrm{N}_n(0, \sigma^2 I_n)$, o estimador $\widehat{\boldsymbol{\beta}}_G$ é usual estimador de mínimos quadrados de $\boldsymbol{\beta}$.

Modelo linear com $\varepsilon \sim \mathrm{N}_n(0, \sigma^2 V)$, V conhecida

A soma de quadrados dos desvios sob o modelo linear transformado é

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \eta_i^2 = \boldsymbol{\eta}^\top \boldsymbol{\eta} = (\boldsymbol{Z} - Q\boldsymbol{\beta})^\top (\boldsymbol{Z} - Q\boldsymbol{\beta}).$$

Voltando o modelo linear original, temos que

$$\begin{aligned} S_G(\boldsymbol{\beta}) &= (P^{-1}\boldsymbol{Y} - P^{-1}X\boldsymbol{\beta})^\top (P^{-1}\boldsymbol{Y} - P^{-1}X\boldsymbol{\beta}) \\ &= [P^{-1}(\boldsymbol{Y} - X\boldsymbol{\beta})]^\top [P^{-1}(\boldsymbol{Y} - X\boldsymbol{\beta})] \\ &= (\boldsymbol{Y} - X\boldsymbol{\beta})^\top {P^{-1}}^\top P^{-1}(\boldsymbol{Y} - X\boldsymbol{\beta}) \\ &= (\boldsymbol{Y} - X\boldsymbol{\beta})^\top V^{-1}(\boldsymbol{Y} - X\boldsymbol{\beta}) \\ &= \boldsymbol{\varepsilon}^\top V^{-1} \boldsymbol{\varepsilon}. \end{aligned}$$

Assim, sob o modelo linear original, $S_G(\boldsymbol{\beta})$ é a soma de quadrados dos desvios generalizados.

Modelo linear com $\varepsilon \sim \mathrm{N}_n(0, \sigma^2 V)$, V conhecida

O estimador de mínimos quadrados para σ^2 sob o modelo linear transformado é

$$\widehat{\sigma}^2 = \frac{1}{n-k-1}\sum_{i=1}^{n}\widehat{\eta}_i^2 = \frac{1}{n-k-1}\widehat{\boldsymbol{\eta}}^\top\widehat{\boldsymbol{\eta}}.$$

em que $\widehat{\boldsymbol{\eta}} = P^{-1}\widehat{\boldsymbol{\varepsilon}}$.

Voltando ao modelo linear original, temos que

$$\begin{aligned}
\widehat{\sigma}^2_G &= \frac{1}{n-k-1}(P^{-1}\widehat{\boldsymbol{\varepsilon}})^\top P^{-1}\widehat{\boldsymbol{\varepsilon}} \\
&= \frac{1}{n-k-1}\widehat{\boldsymbol{\varepsilon}}^\top (P^{-1})^\top P^{-1}\widehat{\boldsymbol{\varepsilon}} \\
&= \frac{1}{n-k-1}\widehat{\boldsymbol{\varepsilon}}^\top V^{-1}\widehat{\boldsymbol{\varepsilon}}.
\end{aligned}$$

Assim, $\widehat{\sigma}^2_G$ é o estimador de mínimos quadrados generalizados para σ^2 sob o modelo linear original.

Caso 2: Propriedades de $\widehat{\beta}_G$

P_1. $\widehat{\boldsymbol{\beta}}_G = B\boldsymbol{Y}$ com $B = (X^\top V^{-1} X)^{-1} X^\top V^{-1}$. Assim, $\widehat{\boldsymbol{\beta}}_G$ é linear em $\boldsymbol{Y}$.

P_2. $\widehat{\boldsymbol{\beta}}_G$ é um estimador não viciado para $\boldsymbol{\beta}$.

De fato, veja que

$$\begin{aligned}
\mathrm{E}(\widehat{\boldsymbol{\beta}}_G) &= \mathrm{E}\left[(X^\top V^{-1} X)^{-1} X^\top V^{-1} \boldsymbol{Y}\right] \\
&= (X^\top V^{-1} X)^{-1} X^\top V^{-1}\, \mathrm{E}(\boldsymbol{Y}) \\
&= (X^\top V^{-1} X)^{-1} X^\top V^{-1} X \boldsymbol{\beta} \\
&= \boldsymbol{\beta}.
\end{aligned}$$

Caso 2: Propriedades de $\widehat{\beta}_G$

P_3. A variância de $\widehat{\boldsymbol{\beta}}_G$ possui forma fechada:

$$
\begin{aligned}
\mathrm{Var}(\widehat{\boldsymbol{\beta}}_G) &= \mathrm{Var}\left[(X^\top V^{-1} X)^{-1} X^\top V^{-1} \boldsymbol{Y}\right] \\
&= (X^\top V^{-1} X)^{-1} X^\top V^{-1}\, \mathrm{Var}(\boldsymbol{Y}) \left[(X^\top V^{-1} X)^{-1} X^\top V^{-1}\right]^\top \\
&= (X^\top V^{-1} X)^{-1} X^\top V^{-1}\, \sigma^2 V \; {V^{-1}}^\top X (X^\top V^{-1} X)^{-1} \\
&= \sigma^2 (X^\top V^{-1} X)^{-1} X^\top V^{-1} X (X^\top V^{-1} X)^{-1} \\
&= \sigma^2 (X^\top V^{-1} X)^{-1}.
\end{aligned}
$$

Lembre-se que V é uma matriz simétrica.

P_4. $\widehat{\boldsymbol{\beta}}_G \sim \mathrm{N}_{k+1}(\boldsymbol{\beta}, \sigma^2 (X^\top V^{-1} X)^{-1})$.

Caso 2: Propriedades de $\widehat{\boldsymbol{\beta}}_G$

P_5. Cada componente de $\widehat{\boldsymbol{\beta}}_G$ segue uma distribuição normal univariada. Mais especificamente, se $\widehat{\beta}^j_G$ é a $(j+1)$-ésima entrada de $\widehat{\boldsymbol{\beta}}_G$, segue que

$$\widehat{\beta}^j_G \sim N(\beta_j,\ \sigma^2\ b_{(j+1)(j+1)}\ ),$$

em que $b_{(j+1)(j+1)}$ é o elemento $(j+1, j+1)$ da matriz $(X^\top V^{-1} X)^{-1}$.

P_6. $\text{Cov}(\widehat{\beta}^j_G,\ \widehat{\beta}^l_G) = \sigma^2\ b_{(j+1)(l+1)}$, com $j, l = 0, 1, 2, \ldots, k$.

Da propriedade P_4, para $j \neq l$, segue que $\widehat{\beta}^j_G$ e $\widehat{\beta}^l_G$ são independentes se $b_{(j+1)(l+1)} = 0$.

P_7. O estimador $\widehat{\boldsymbol{\beta}}_G$ obtido por mínimos quadrados generalizados coincide com o estimador $\widehat{\boldsymbol{\beta}}_{MV}$ obtido por máxima verossimilhança.

Estimadores ótimos para funções de β e σ^2

Resultado. o estimador $\widehat{\boldsymbol{\beta}}$ é um estimador não viciado para $\boldsymbol{\beta}$ sob o modelo $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, com $\boldsymbol{\varepsilon} \sim \mathrm{N}_n(\mathbf{0}, \sigma^2 V)$.

Entretanto, $\widehat{\boldsymbol{\beta}}$ **não é** um estimador ótimo para $\boldsymbol{\beta}$ (não é função de uma estatística suficiente e completa para $\boldsymbol{\beta}$). Logo, $\widehat{\boldsymbol{\beta}}$ não é o ENVVUM de $\boldsymbol{\beta}$.

Resultado. Sejam $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, com $\boldsymbol{\varepsilon} \sim \mathrm{N}_n(\mathbf{0}, \sigma^2 V)$, e $t(\boldsymbol{\theta})$ uma função qualquer de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top \ \sigma^2)^\top$. Se $q(\widehat{\boldsymbol{\theta}}_G)$, com $\widehat{\boldsymbol{\theta}}_G = (\widehat{\boldsymbol{\beta}}_G^\top \ \widehat{\sigma}_G^2)^\top$, é um estimador não viciado para $t(\boldsymbol{\theta})$, então $q(\widehat{\boldsymbol{\theta}}_G)$ é o ENVVUM para $t(\boldsymbol{\theta})$.

Estimadores ótimos para funções de β e σ^2

Assim, temos os seguintes resultados:

(i) $\widehat{\boldsymbol{\beta}}_G = (X^\top V^{-1} X)^{-1} X^\top V^{-1} \boldsymbol{Y}$ é o ENVVUM para $\boldsymbol{\beta}$;

(ii) $\widehat{\beta}^j_G$ é o ENVVUM para β_j com $j = 0, 1, \ldots, k$;

(iii) $\boldsymbol{a}^\top \widehat{\boldsymbol{\beta}}_G = \boldsymbol{a}^\top (X^\top V^{-1} X)^{-1} X^\top V^{-1} \boldsymbol{Y}$ é o ENVVUM para $\boldsymbol{a}^\top \boldsymbol{\beta}$ com $\boldsymbol{a}^\top = (a_0 \ a_1 \ \ldots \ a_k) \in \mathbb{R}^{k+1}$ um vetor de constantes;

(iv) $\widehat{\mathrm{E}(Y_i)} = \boldsymbol{X}_i^\top \widehat{\boldsymbol{\beta}}_G$ é o ENVVUM para $\mathrm{E}(Y_i) = \boldsymbol{X}_i^\top \boldsymbol{\beta}$.

(v) $\widehat{\sigma}^2_G$ é o ENVVUM para σ^2

Testes de hipóteses via TRV

Queremos testar $H_0\colon C\boldsymbol{\beta} = \boldsymbol{m}$ contra $H_1\colon C\boldsymbol{\beta} \neq \boldsymbol{m}$, com $C_{p\times(k+1)}$ de constantes tal que $r(C) = p \leq k+1$ (posto completo), $\boldsymbol{m}$ vetor de constantes $p \times 1$.

Sob o modelo linear transformado $\boldsymbol{Z} = Q\boldsymbol{\beta} + \boldsymbol{\eta}$ com $\boldsymbol{\eta} \sim \mathrm{N}_n(0, \sigma^2 I_n)$, a estatística do TRV é

$$W_0 = \frac{(C\widehat{\boldsymbol{\beta}} - \boldsymbol{m})^\top [C(Q^\top Q)^{-1} C^\top]^{-1} (C\widehat{\boldsymbol{\beta}} - \boldsymbol{m})}{p\widehat{\sigma}^2}.$$

Testes de hipóteses via TRV

Sob o modelo linear original $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ com $\boldsymbol{\varepsilon} \sim \mathrm{N}_n(0, \sigma^2 V)$, a estatística do TRV é

$$
\begin{aligned}
W_0 &= \frac{(C\widehat{\beta}_G - \boldsymbol{m})^\top \{C[(P^{-1}X)^\top P^{-1}X]^{-1}C^\top\}^{-1}(C\widehat{\beta}_G - \boldsymbol{m})}{p\widehat{\sigma}_G^2} \\
&= \frac{(C\widehat{\beta}_G - \boldsymbol{m})^\top \{C[X^\top {P^{-1}}^\top P^{-1}X]^{-1}C^\top\}^{-1}(C\widehat{\beta}_G - \boldsymbol{m})}{p\widehat{\sigma}_G^2} \\
&= \frac{(C\widehat{\beta}_G - \boldsymbol{m})^\top \{C[X^\top V^{-1}X]^{-1}C^\top\}^{-1}(C\widehat{\beta}_G - \boldsymbol{m})}{p\widehat{\sigma}_G^2}.
\end{aligned}
$$

Rejeita-se H_0: $C\boldsymbol{\beta} = m$ ao nível de significância $\alpha \cdot 100\%$ se $W_0 \geq w_0$ com w_0 obtido tal que

$$\mathrm{P}(F_{p,n-k-1} \geq w_0) = \alpha.$$

Exemplo

Antes de definirmos as somas de quadrados, considere o exemplo a seguir.

Considere o modelo linear

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \ldots, n$$

em que $\varepsilon_i \overset{\text{ind}}{\sim} \mathrm{N}(0, \sigma^2 w_i)$ com $w_i > 0$.

Daí, temos que $\boldsymbol{\varepsilon} \sim \mathrm{N}_n(\mathbf{0}, \sigma^2 V)$ com $V = \mathrm{diag}\{w_1, w_2, \ldots, w_n\}$.

Pela decomposição de Cholesky, sabemos que $V = PP^\top$, e portanto,

$$P = V^{1/2} = \mathrm{diag}\{\sqrt{w_1}, \sqrt{w_2}, \ldots, \sqrt{w_n}\}.$$

Exemplo

Segue que

$$P^{-1} = V^{-1/2} = \mathrm{diag}\left\{\frac{1}{\sqrt{w_1}}, \frac{1}{\sqrt{w_2}}, \ldots, \frac{1}{\sqrt{w_n}}\right\}.$$

Então, o modelo linear transformado é $\boldsymbol{Z} = Q\boldsymbol{\beta} + \boldsymbol{\eta}$ com

$$\boldsymbol{Z} = P^{-1}\boldsymbol{Y} = \begin{bmatrix} \frac{Y_1}{\sqrt{w_1}} \\ \frac{Y_2}{\sqrt{w_2}} \\ \vdots \\ \frac{Y_n}{\sqrt{w_n}} \end{bmatrix}, \quad Q = P^{-1}X = \begin{bmatrix} \frac{1}{\sqrt{w_1}} & \frac{x_1}{\sqrt{w_1}} \\ \frac{1}{\sqrt{w_2}} & \frac{x_2}{\sqrt{w_2}} \\ \vdots & \vdots \\ \frac{1}{\sqrt{w_n}} & \frac{x_n}{\sqrt{w_n}} \end{bmatrix}, \quad \boldsymbol{\eta} = P^{-1}\boldsymbol{\varepsilon} = \begin{bmatrix} \frac{\varepsilon_1}{\sqrt{w_1}} \\ \frac{\varepsilon_2}{\sqrt{w_2}} \\ \vdots \\ \frac{\varepsilon_n}{\sqrt{w_n}} \end{bmatrix}$$

Exemplo

Então, o modelo linear transformado é dado por

$$\frac{Y_i}{\sqrt{w_i}} = \beta_0 \frac{1}{\sqrt{w_i}} + \beta_1 \frac{x_i}{\sqrt{w_i}} + \frac{\varepsilon_i}{\sqrt{w_i}}, \quad i = 1, 2, \ldots, n.$$

Este modelo linear não possui intercepto. Perceba que β_0 é coeficiente da covariável $1/\sqrt{w_i}$.

Muitas vezes a transformação $\boldsymbol{Z} = P^{-1}\boldsymbol{Y}$ resulta em um modelo linear sem intercepto, mesmo que o modelo linear original tenha intercepto.

Sendo assim, devemos ter o cuidado de identificar se o modelo linear transformado possui intercepto para que possamos definir as somas de quadrados do modelo linear original.

Somas de quadrados com intercepto

Se o modelo linear transformado possuir intercepto, usamos as somas de quadrados SQT, SQReg, e SQRes definidas, respectivamente, por

$$\mathrm{SQT} = \sum_{i=1}^{n} (Z_i - \overline{Z})^2;$$

$$\mathrm{SQReg} = \sum_{i=1}^{n} (\widehat{Z}_i - \overline{Z})^2;$$

$$\mathrm{SQRes} = \sum_{i=1}^{n} \widehat{\eta}_i^2 = \sum_{i=1}^{n} (Z_i - \widehat{Z}_i)^2.$$

Somas de quadrados com intercepto

Agora, usando resultado do modelo de regressão linear normal,

$$
\begin{aligned}
\mathrm{SQT} &= \sum_{i=1}^{n}(Z_i - \overline{Z})^2 \\
&= \boldsymbol{Z}^\top \left[I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^\top \right] \boldsymbol{Z} \\
&= [P^{-1}\boldsymbol{Y}]^\top \left[I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^\top \right] P^{-1}\boldsymbol{Y} \\
&= \boldsymbol{Y}^\top \left[{P^{-1}}^\top P^{-1} - \frac{1}{n}{P^{-1}}^\top \mathbf{1}\mathbf{1}^\top P^{-1} \right] \boldsymbol{Y} \\
&= \boldsymbol{Y}^\top \left[V^{-1} - \frac{1}{n}{P^{-1}}^\top \mathbf{1}\mathbf{1}^\top P^{-1} \right] \boldsymbol{Y}.
\end{aligned}
$$

Somas de quadrados com intercepto

Ainda, temos que

$$
\begin{aligned}
\text{SQReg} &= \sum_{i=1}^{n}(\widehat{Z}_i - \overline{Z})^2 \\
&= \boldsymbol{Z}^\top \left[Q(Q^\top Q)^{-1}Q^\top - \frac{1}{n}\mathbf{1}\mathbf{1}^\top \right] \boldsymbol{Z} \\
&= [P^{-1}\boldsymbol{Y}]^\top \left\{ P^{-1}X[(P^{-1}X)^\top P^{-1}X]^{-1}(P^{-1}X)^\top - \frac{1}{n}\mathbf{1}\mathbf{1}^\top \right\} P^{-1}\boldsymbol{Y} \\
&= \boldsymbol{Y}^\top {P^{-1}}^\top \left\{ P^{-1}X[X^\top {P^{-1}}^\top P^{-1}X]^{-1}X^\top {P^{-1}}^\top - \frac{1}{n}\mathbf{1}\mathbf{1}^\top \right\} P^{-1}\boldsymbol{Y} \\
&= \boldsymbol{Y}^\top \left\{ V^{-1}X[X^\top V^{-1}X]^{-1}X^\top V^{-1} - \frac{1}{n}{P^{-1}}^\top \mathbf{1}\mathbf{1}^\top P^{-1} \right\} \boldsymbol{Y}.
\end{aligned}
$$

Somas de quadrados com intercepto

Por fim,

$$
\begin{aligned}
\mathrm{SQRes} &= \sum_{i=1}^{n} \widehat{\eta}_i^2 \\
&= \boldsymbol{Z}^\top \left[I_n - Q(Q^\top Q)^{-1} Q^\top \right] \boldsymbol{Z} \\
&= [P^{-1}\boldsymbol{Y}]^\top \left\{ I_n - P^{-1}X[X^\top {P^{-1}}^\top P^{-1}X]^{-1}X^\top {P^{-1}}^\top \right\} P^{-1}\boldsymbol{Y} \\
&= \boldsymbol{Y}^\top \left\{ V^{-1} - V^{-1}X[X^\top V^{-1}X]^{-1}X^\top V^{-1} \right\} \boldsymbol{Y}.
\end{aligned}
$$

Ainda, $\mathrm{SQRes} = (n-k-1)\widehat{\sigma}_G^2$.

Somas de quadrados com intercepto

Sob o modelo linear transformado, sabemos que

$$\frac{\text{SQRes}}{\sigma^2} \sim \chi^2_{n-k-1};$$

$$\frac{\text{SQReg}}{\sigma^2} \sim \chi^2_{k,\lambda},$$

com

$$\lambda = \frac{1}{2\sigma^2} \boldsymbol{\beta}^{*\top} Q_1^\top \left[I_n - \frac{1}{n} \mathbf{1}\mathbf{1}^\top \right] Q_1 \boldsymbol{\beta}^*,$$

em que Q_1 é obtida de $Q = [\mathbf{1}\ Q_1]$.

Sob H_0: $\boldsymbol{\beta}^* = \mathbf{0}$, temos que

$$\frac{\text{SQReg}}{\sigma^2} \sim \chi^2_k.$$

A partir disto, pode-se montar a tabela de análise de variância.

Somas de quadrados sem intercepto

Se o modelo linear transformado **NÃO** possuir intercepto, usamos as somas de quadrados SQT_{nc}, SQReg_{nc}, e SQRes definidas, respectivamente, por

$$\mathrm{SQT}_{nc} = \sum_{i=1}^{n} Z_i^2;$$

$$\mathrm{SQReg}_{nc} = \sum_{i=1}^{n} \widehat{Z}_i^2;$$

$$\mathrm{SQRes} = \sum_{i=1}^{n} \widehat{\eta}_i^2 = \sum_{i=1}^{n} (Z_i - \widehat{Z}_i)^2.$$

Note que SQRes é definida da mesma forma.

Somas de quadrados sem intercepto

Veja que

$$\begin{aligned} \mathrm{SQT}_{nc} &= \sum_{i=1}^{n} Z_i^2 \\ &= \boldsymbol{Z}^\top \boldsymbol{Z} \\ &= [P^{-1}\boldsymbol{Y}]^\top P^{-1}\boldsymbol{Y} \\ &= \boldsymbol{Y}^\top V^{-1}\boldsymbol{Y}; \end{aligned}$$

$$\begin{aligned} \mathrm{SQReg}_{nc} &= \sum_{i=1}^{n} \widehat{Z}_i^2 \\ &= \boldsymbol{Z}^\top Q(Q^\top Q)^{-1}Q^\top \boldsymbol{Z} \\ &= \boldsymbol{Y}^\top V^{-1}X[X^\top V^{-1}X]^{-1}X^\top V^{-1}\boldsymbol{Y}. \end{aligned}$$

Somas de quadrados sem intercepto

Sob o modelo linear transformado, sabemos que

$$\frac{\text{SQRes}}{\sigma^2} \sim \chi^2_{n-k};$$

$$\frac{\text{SQReg}_{nc}}{\sigma^2} \sim \chi^2_{k,\lambda},$$

com

$$\lambda = \frac{1}{2\sigma^2}\boldsymbol{\beta}^\top Q^\top Q\boldsymbol{\beta}.$$

Sob H_0: $\boldsymbol{\beta} = \mathbf{0}$, temos que

$$\frac{\text{SQReg}_{nc}}{\sigma^2} \sim \chi^2_{k}.$$

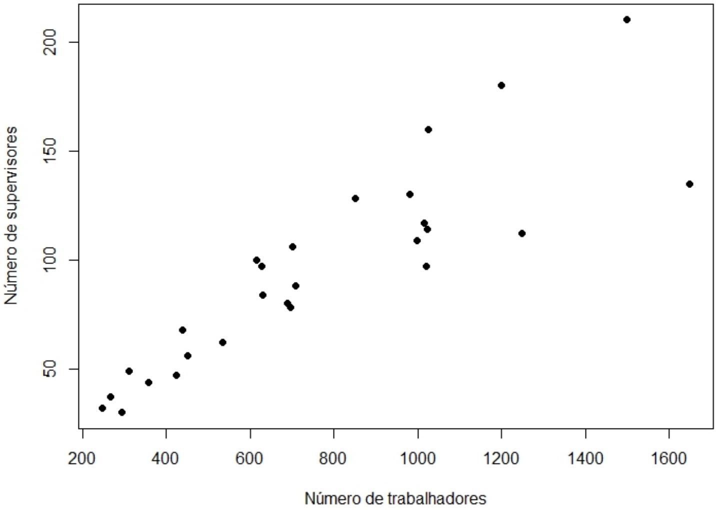
A partir disto, pode-se montar a tabela de análise de variância.

Aplicação prática

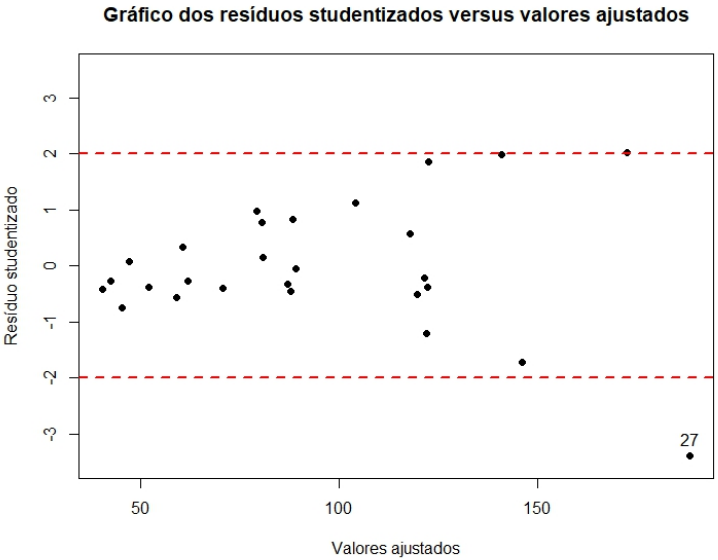
No arquivo **supervisores.txt** estão apresentados dados referentes a 27 estabelecimentos industriais. Neste estudo, o objetivo é explicar o número de supervisores em função do número de trabalhadores destes estabelecimentos.

Inicialmente foi ajustado um modelo de regressão linear normal simples.

Aplicação prática



Aplicação prática



Aplicação prática

Em seguida, ajustou-se um modelo de regressão linear sob mínimos quadrados ponderados supondo que $\mathrm{Var}(Y) = \sigma^2 V$ com

$$V = \mathrm{diag}\{x_1^2, x_2^2, \ldots, x_n^2\}.$$

Dada a matriz V, obtemos que os pesos do procedimento de estimação por mínimos quadrados ponderados são dados por $w_i = 1/x_i^2$.

Aplicação prática