ME720 - Modelos Lineares Generalizados

Parte 7 - Parametrização, identificabilidade, estimabilidade e testabilidade em modelos normais lineares

Profa. **Larissa Avila Matos**

Exemplo 2 (dados de absorbância)

Voltemos ao exemplo dos solventes (um único fator com vários níveis).

$$Y_{ij} = \mu + \alpha_i + \xi_{ij},\ i = 1, 2, \cdots, k,\ j = 1, \cdots, n_i.$$

- Erros (parte aleatória): $\xi_{ij} \overset{i.i.d}{\sim} N(0, \sigma^2)$;
- μ, α_i e σ^2 parâmetros desconhecidos;
- μ : média geral.
- α_i : efeito relacionada ao i-ésimo fator.

Implicação das suposições acima: $Y_{ij} \overset{ind.}{\sim} N(\mu + \alpha_i, \sigma^2)$.

Exemplo 2

Matricialmente: $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\xi}$, onde

$$\boldsymbol{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{15} \\ Y_{21} \\ \vdots \\ Y_{51} \\ \vdots \\ Y_{55} \end{bmatrix}, \; \boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \; \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix}, \; \boldsymbol{\xi} = \begin{bmatrix} \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{21} \\ \vdots \\ \xi_{51} \\ \vdots \\ \xi_{55} \end{bmatrix}$$

Identificabilidade

Seja $L(\boldsymbol{\theta})$ a verossimilhança associada à um determinado modelo (probabilístico, estatístico, regressão).

O modelo é dito estar (globalmente) identificado (sob a ótica frequentista) se $\forall \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2, \boldsymbol{\theta}_i, i = 1, 2 \in \boldsymbol{\Theta}$, tivermos

$$L(\boldsymbol{\theta}_1) \neq L(\boldsymbol{\theta}_2).$$

Se o modelo está identificado, tem-se que, para um determinado método de estimação (em geral), teremos um único conjunto de estimativas, para uma dada amostra.

Se o modelo não está identificado, podemos ter mais de um conjunto de estimativas.

Porque (re)parametrizar o modelo?

Interpretações mais apropriadas dos parâmetros (os modelos são mais facilmente interpretáveis).

Testes de hipótese mais simples (em termos dos parâmetros) principalmente para modelos mais complexos.

Para garantir a identificabilidade do modelo.

Mais facilidade na identificação de efeitos significativos.

Estimadores com propriedades mais interessantes (ortogonalidade).

Parametrizações para o modelo em questão

Modelo de médias. Modelo identificado.

Desvios com restrição: $\sum_{i=1}^{k} \alpha_i = 0$. Modelo identificado.

Casela (cela de referência): igualar um único α_i à 0, por exemplo α_1. Modelo identificado.

Desvios sem restrição: não se coloca nenhuma restrição. Modelo não identificado. Neste caso, somente as funções estimáveis podem ser "estimadas".

Para interpretação dos parâmetros basta lembrar que $\mu_i = \mu + \alpha_i$

Parametrização de médias

Modelo de médias:

$$Y_{ij} = \mu_i + \xi_{ij},\ i = 1, 2, \cdots, k,\ j = 1, \cdots, n_i.$$

- Erros $\xi_{ij} \overset{i.i.d}{\sim} N(0, \sigma^2)$.
- μ_i não aleatório.
- $\mathbb{E}(Y_{ij}) = \mu_i,\ \text{Var}(Y_{ij}) = \sigma^2$.
- μ_i : média populacional relacionada ao i-ésimo fator.

Parametrização de médias

Neste caso, temos que

$$\boldsymbol{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{15} \\ Y_{21} \\ \vdots \\ Y_{51} \\ \vdots \\ Y_{55} \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{21} \\ \vdots \\ \xi_{51} \\ \vdots \\ \xi_{55} \end{bmatrix}$$

Desvios com restrição

Desvios com restrição: $\sum_{i=1}^{k} \alpha_i = 0$.

- μ : média das médias dos grupos $\mu = \overline{\mu} = \frac{1}{k} \sum_{i=1}^{k} \mu_i$.
- α_i : incremento (positivo ou negativo) da média do grupo i com relação à média das médias, $\alpha_i = \mu_i - \overline{\mu}$.

Em geral, considera-se que $\alpha_k = -\sum_{i=1}^{k-1} \alpha_i$.

Desvios com restrição

Neste caso, temos que

$$\boldsymbol{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{15} \\ Y_{21} \\ \vdots \\ Y_{51} \\ \vdots \\ Y_{55} \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & -1 & -1 & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & -1 & -1 & -1 & -1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{21} \\ \vdots \\ \xi_{51} \\ \vdots \\ \xi_{55} \end{bmatrix}$$

Casela de referências ($\alpha_1 = 0$)

Casela de referência: $\alpha_1 = 0$.

- μ : média do grupo 1 (grupo de referência), $\mu = \mu_1$.
- α_i : incremento (positivo ou negativo) da média do grupo i com relação à média do grupo 1 (grupo de referência), $\alpha_i = \mu_i - \mu_1$.

Em geral, esta será a parametrização utilizada no curso.

Casela de referências ($\alpha_1 = 0$)

Neste caso, temos que

$$\boldsymbol{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{15} \\ Y_{21} \\ \vdots \\ Y_{51} \\ \vdots \\ Y_{55} \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{21} \\ \vdots \\ \xi_{51} \\ \vdots \\ \xi_{55} \end{bmatrix}$$

Desvios sem restrição

Desvios sem restrição: k médias e (k+1) parâmetros.

Os parâmetros, por si só, não tem interpretações mas, combinações lineares deles, podem ter:

- $\mu_i = \mu + \alpha_i$ (média do i-ésimo grupo).
- $\mu_i - \mu_j = \alpha_i - \alpha_j$ (diferença entre as médias do grupo i e do grupo j).

As combinações lineares acima pertencem à classe das **funções estimáveis.**

Desvios sem restrição

Neste caso, temos que

$$\boldsymbol{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{15} \\ Y_{21} \\ \vdots \\ Y_{51} \\ \vdots \\ Y_{55} \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{21} \\ \vdots \\ \xi_{51} \\ \vdots \\ \xi_{55} \end{bmatrix}$$

Estimação por MQO

O estimador de MQO é dado pela solução da seguinte equação:

$$(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{Y}$$

Nos casos das parametrizações: de médias, casela de referência e desvios com restrição, a matriz $\boldsymbol{X}$ é de posto coluna completo e, portanto

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$$

Contudo, para a parametrização desvios sem restrição, a matriz $\boldsymbol{X}$ é de posto coluna incompleto e, portanto

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{G}\boldsymbol{X}'\boldsymbol{Y},$$

em que $\boldsymbol{G}$ é uma (entre várias) inversa generalizada (ig) de $(\boldsymbol{X}'\boldsymbol{X})$. Ou seja, $(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{G}(\boldsymbol{X}'\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{X}$.

Desse modo, para um determinado conjunto de dados, podemos ter infinitos conjuntos de estimativas.

Que implicações (limitações) ocorrem neste caso?

Primeiramente, note que $\widehat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{G}')$.

Resultados importantes

Se $\boldsymbol{X}'\boldsymbol{X} = \mathbf{0} \rightarrow \boldsymbol{X} = \mathbf{0}$.

Se $\boldsymbol{P}\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{X}'\boldsymbol{X} \rightarrow \boldsymbol{P}\boldsymbol{X}' = \boldsymbol{Q}\boldsymbol{X}'$.

Se $\boldsymbol{G}$ é uma ig de de $\boldsymbol{X}'\boldsymbol{X}$ (usando-se os dois itens acima), tem-se que

1. $\boldsymbol{G}'$ é uma ig de $\boldsymbol{X}'\boldsymbol{X}$.
2. $\boldsymbol{G}\boldsymbol{X}'$ é uma ig de $\boldsymbol{X}$, ou seja $\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}$.
3. $\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'$ é invariante relativamente à escolha de $\boldsymbol{G}$.
4. $\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'$ é simétrica, mesmo quando $\boldsymbol{G}$ não o é.

Somas de Quadrados

Seja $\widehat{\boldsymbol{\beta}} = \boldsymbol{G}\boldsymbol{X}'\boldsymbol{Y}$ (note que temos um conjunto de soluções).

Soma de quadrados dos resíduos

$$
\begin{aligned}
SQR &= (\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}) = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{Y})'(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{Y}) \\
&= \boldsymbol{Y}'(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}')'(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}')\boldsymbol{Y} \\
&= \boldsymbol{Y}'(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}' - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}' + \underbrace{\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}}_{\boldsymbol{X}}\boldsymbol{G}\boldsymbol{X}')\boldsymbol{Y} \\
&= \boldsymbol{Y}'(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}')\boldsymbol{Y}
\end{aligned}
$$

Assim, tem-se que $\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'$ é simétrica e idempotente. Além disso $r(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}') = n - r(\boldsymbol{X})$.

Pode-se provar, então, que $SQR/\sigma^2 \sim \chi^2_{(n-r(\boldsymbol{X}))}$.

Defina agora $SQM = \boldsymbol{Y}'(\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}' - n^{-1}\boldsymbol{J})\boldsymbol{Y} = \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$.

Pode-se provar que:

- $n^{-1}\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{J} = n^{-1}\boldsymbol{J}$.
- $\boldsymbol{A}$ é simétrica e idempotente, onde $r(\boldsymbol{A}) = r(\boldsymbol{X}) - 1$.
- $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{0}$
- Sob H_0, $SQM/\sigma^2 \sim \chi^2_{(r(\boldsymbol{X})-1)}$.
- Sob H_1, $SQM/\sigma^2 \sim \chi^2_{(r(\boldsymbol{X})-1,\delta)}$, em que

  $$\delta = \frac{1}{\sigma^2}\left(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - n^{-1}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{J}\boldsymbol{X}\boldsymbol{\beta}\right) = \frac{1}{\sigma^2}\left(\boldsymbol{\beta}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - n^{-1}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{J}\boldsymbol{X}\boldsymbol{\beta}\right)$$

Estatística do teste ANOVA

Assim, sob H_0,

$$F = \frac{SQM/(r(\boldsymbol{X}) - 1)}{SQR/n - r(\boldsymbol{X})} \sim F_{(r(\boldsymbol{X})-1, n-r(\boldsymbol{X}))}.$$

Assim, sob H_1,

$$F = \frac{SQM/(r(\boldsymbol{X}) - 1)}{SQR/n - r(\boldsymbol{X})} \sim F_{(r(\boldsymbol{X})-1, n-r(\boldsymbol{X}), \delta)},$$

em que

$$\delta = \frac{1}{\sigma^2}\left(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - n^{-1}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{J}\boldsymbol{X}\boldsymbol{\beta}\right) = \frac{1}{\sigma^2}\left(\boldsymbol{\beta}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - n^{-1}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{J}\boldsymbol{X}\boldsymbol{\beta}\right)$$

Os resultados apresentados valem, independentemente da escolha de $\boldsymbol{G}$.

Funções estimáveis de $\boldsymbol{\beta}$

Tendo estabelecido que não podemos estimar $\boldsymbol{\beta}$, a próxima dúvida é saber se podemos estimar alguma combinação linear dos β's, digamos $\boldsymbol{\lambda}'\boldsymbol{\beta}$, onde $\boldsymbol{\lambda}$ é um vetor $p \times 1$.

Considere o modelo normal linear usual e $\boldsymbol{X}$ geral (de posto coluna completo ou incompleto) e $\widehat{\boldsymbol{\beta}} = \boldsymbol{G}\boldsymbol{X}'\boldsymbol{Y}$.

Definição: Uma função linear dos parâmetros $\boldsymbol{\lambda}'\boldsymbol{\beta}$ é dita ser estimável se existe uma combinação linear das observações com um valor esperado igual a $\boldsymbol{\lambda}'\boldsymbol{\beta}$, ou seja, dizemos que $\theta = \boldsymbol{\lambda}'\boldsymbol{\beta}$ é uma função estimável se $\exists$ um vetor $\boldsymbol{a}_{(p\times 1)}$ tal que

$$\mathbb{E}(\boldsymbol{a}'\boldsymbol{Y}) = \boldsymbol{\lambda}'\boldsymbol{\beta} = \theta.$$

Funções estimáveis de $\boldsymbol{\beta}$

Teorema: No modelo linear $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\xi}$, onde $\mathbb{E}(\boldsymbol{Y}) = \boldsymbol{X\beta}$ e $\boldsymbol{X}$ é uma matriz $n \times p$ de posto $k < p \leq n$, a função linear $\boldsymbol{\lambda}'\boldsymbol{\beta}$ é estimável se e somente se qualquer uma das seguintes condições é satisfeita:

1. $\boldsymbol{\lambda}'$ é uma combinação linear das linhas de $\boldsymbol{X}$, isto é, existe um vetor $\boldsymbol{a}$ tal que $\boldsymbol{a}'\boldsymbol{X} = \boldsymbol{\lambda}'$.
2. $\boldsymbol{\lambda}'$ é uma combinação linear das linhas de $\boldsymbol{X}'\boldsymbol{X}$ ou $\boldsymbol{\lambda}$ é uma combinação linear das colunas de $\boldsymbol{X}'\boldsymbol{X}$, isto é, existe um vetor $\boldsymbol{r}$ tal que $\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{\lambda}'$ ou $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{r} = \boldsymbol{\lambda}$.
3. $\boldsymbol{\lambda}$ ou $\boldsymbol{\lambda}'$ é tal que $\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{\lambda} = \boldsymbol{\lambda}$ ou $\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{\lambda}'$, onde $(\boldsymbol{X}'\boldsymbol{X})^{-}$ é qualquer inversa generalizada (simétrica) de $\boldsymbol{X}'\boldsymbol{X}$.

Funções estimáveis de $\boldsymbol{\beta}$ - Exemplo

Voltando ao exemplo anterior, temos

1. $\boldsymbol{\lambda}' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, então $\theta = \mu$.
2. $\boldsymbol{\lambda}' = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix}$, então $\theta = \alpha_1 - \alpha_2$.
3. $\boldsymbol{\lambda}' = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$, então $\theta = \mu + \alpha_1$.

Para o segundo caso, se tomarmos

$$\boldsymbol{a}' = \begin{bmatrix} \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} & -\frac{1}{5} & -\frac{1}{5} & -\frac{1}{5} & -\frac{1}{5} & 0 & \dots & 0 \end{bmatrix},$$

temos que $\boldsymbol{a}'\boldsymbol{Y} = \overline{Y}_1 - \overline{Y}_2$, logo $\boldsymbol{a}'\mathbb{E}(\boldsymbol{Y}) = \mathbb{E}(\overline{Y}_1 - \overline{Y}_2) = \alpha_1 - \alpha_2$.

Para o terceiro caso, se tomarmos
$\boldsymbol{a}' = \begin{bmatrix} 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 & \dots & 0 \end{bmatrix}$, temos que $\boldsymbol{a}'\boldsymbol{Y} = \overline{Y}_1$, logo $\boldsymbol{a}'\mathbb{E}(\boldsymbol{Y}) = \mathbb{E}(\overline{Y}_1) = \mu + \alpha_1$.

Para o primeiro caso, não é possível encontrar $\boldsymbol{a}$ que satisfaça à restrição em questão.

Observação, se $\boldsymbol{a}'\mathbb{E}(\boldsymbol{Y}) = \boldsymbol{\lambda}'\boldsymbol{\beta} \rightarrow \boldsymbol{a}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{a}'\boldsymbol{\beta} \rightarrow \boldsymbol{a}'\boldsymbol{X} = \boldsymbol{\lambda}'$.

Teorema: No modelo de posto incompleto ($\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\xi}$), o número de funções estimáveis linearmente independentes de $\boldsymbol{\beta}$ é igual ao posto de $\boldsymbol{X}$.

Estimadores de $\boldsymbol{\lambda}'\beta$

O valor esperado de cada observação é estimável.

Combinações lineares de funções estimáveis são funções estimáveis.

Dos slides acima temos os estimadores $\boldsymbol{a}'\boldsymbol{Y}$ e $\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{Y}$ para $\boldsymbol{\lambda}'\boldsymbol{\beta}$, onde $\boldsymbol{a}'$ e $\boldsymbol{r}'$ satisfazem $\boldsymbol{\lambda}' = \boldsymbol{a}'\boldsymbol{X}$ e $\boldsymbol{\lambda} = \boldsymbol{r}'\boldsymbol{X}'\boldsymbol{X}$, respectivamente. Um terceiro estimador de $\boldsymbol{\lambda}'\boldsymbol{\beta}$ é $\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}}$, onde $\widehat{\boldsymbol{\beta}}$ é uma solução de $\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{Y}$.

Nós iremos focar nos estimadores $\boldsymbol{r}\boldsymbol{X}'\boldsymbol{Y}$ e $\lambda'\widehat{\boldsymbol{\beta}}$.

Resultados sobre funções estimáveis

Seja $\boldsymbol{\lambda}'\boldsymbol{\beta}$ uma função estimável de $\boldsymbol{\beta}$ no modelo $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\xi}$, onde $E(\boldsymbol{Y}) = \boldsymbol{X}\boldsymbol{\beta}$ e $\boldsymbol{X}$ é $n \times p$ de posto $k < p \leq n$. Seja $\widehat{\boldsymbol{\beta}}$ qualquer solução do sistema de equações normais $\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{Y}$, e seja $\boldsymbol{r}$ qualquer solução para $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{r} = \boldsymbol{\lambda}$. Então os dois estimadores $\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}}$ e $\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{Y}$ têm as seguintes propriedades:

1. $E(\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}}) = E(\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{Y}) = \boldsymbol{\lambda}'\boldsymbol{\beta}$.
2. $\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}} = \boldsymbol{r}'\boldsymbol{X}'\boldsymbol{Y}$ para qualquer $\widehat{\boldsymbol{\beta}}$ e qualquer $\boldsymbol{r}$.
3. $\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}}$ e $\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{Y}$ são invariantes para escolhas de $\widehat{\boldsymbol{\beta}}$ ou $\boldsymbol{r}$.

Prova: ??

Resultados sobre funções estimáveis

A variância de $\boldsymbol{\lambda}'\boldsymbol{\beta}$ ou de $\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{Y}$ tem as seguintes propriedades:

1. $Var(\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{Y}) = \sigma^2\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{r} = \sigma^2\boldsymbol{r}'\boldsymbol{\lambda}$.
2. $Var(\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}}) = \sigma^2\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{\lambda}$.
3. $Var(\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}})$ é única, ou seja, é invariante para as escolhas de $\boldsymbol{r}$ ou de $(\boldsymbol{X}'\boldsymbol{X})^-$.

Se $\boldsymbol{\lambda}_1'\boldsymbol{\beta}$ e $\boldsymbol{\lambda}_2'\boldsymbol{\beta}$ são duas funções estimáveis do modelo, então

$$Cov(\boldsymbol{\lambda}_1'\widehat{\boldsymbol{\beta}}, \boldsymbol{\lambda}_2'\widehat{\boldsymbol{\beta}}) = \sigma^2\boldsymbol{r}_1'\boldsymbol{\lambda}_2 = \sigma^2\boldsymbol{\lambda}1'\boldsymbol{r}_2 = \sigma^2\boldsymbol{\lambda}1'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{\lambda}_2,$$

onde $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{r}_1 = \boldsymbol{\lambda}_1$ e $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{r}_2 = \boldsymbol{\lambda}_2$.

Prova: ??

Resultados sobre funções estimáveis

Então, temos o seguinte Teorema:

[TEO.] Se $\boldsymbol{\lambda}'\boldsymbol{\beta}$ é uma função estimável no modelo $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\xi}$, onde $\boldsymbol{X}$ é $n \times p$ de posto $k < p \leq n$, então os estimadores $\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}}$ e $\boldsymbol{r}'\boldsymbol{X}'\boldsymbol{Y}$ são os melhores estimadores não viesados (BLUE) de $\boldsymbol{\lambda}'\boldsymbol{\beta}$.

Prova: ??

Estimador de σ^2

Assim como nos modelos com posto completo, minimizar a soma de quadrados $\boldsymbol{\xi}\boldsymbol{\xi}'$ não fornece um estimador para σ^2. No entanto, um estimador não viaciado de σ^2 baseado nas estimativas de mínimos quadrados e dado por

$$\widehat{\sigma}^2 = \frac{1}{n-k}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)'\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)$$

O qual é não-viciado. Além disso, pode-se provar que $\widehat{\boldsymbol{\beta}}\perp\widehat{\sigma}^2$ e $\frac{(n-p)\widehat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-k)}$. $\widehat{\sigma}^2$ é invariante a escolha de $\widehat{\boldsymbol{\beta}}$ ou a escolha da inversa generalizada $(\boldsymbol{X}\boldsymbol{X})^{-}$.

Estimação por máxima verossimilhança

Seja $\boldsymbol{Y} \sim N_n(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I})$, onde $\boldsymbol{X}$ é $n \times p$ de posto $k < p \leq n$.

Então, os estimadores de máxima verossimilhança $\widehat{\boldsymbol{\beta}}$ e $\widehat{\sigma}^2$ (corrigindo o viés) são dados por

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{Y} \;\text{ e }\; \widehat{\sigma}^2 = \frac{n}{n-k}(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}),$$

e têm as seguintes propriedades:

1. $\widehat{\boldsymbol{\beta}} \sim N_p((\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-})$;
2. $\frac{(n-k)\widehat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-k)}$;
3. $\widehat{\boldsymbol{\beta}}$ e $\widehat{\sigma}^2$ são independentes.

Se $\boldsymbol{\lambda}'\boldsymbol{\beta}$ é uma função estimável no modelo, então $\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}}$ tem variância mínima dentre todos os estimadores não viesados.

Hipóteses testáveis

Uma hipótese tal como $H_0 : \beta_1 = \beta_2 = \cdots = \beta_q$ é dita ser testável se existe um conjunto de funções estimáveis linearmente independentes $\boldsymbol{\lambda}_1'\boldsymbol{\beta}, \boldsymbol{\lambda}_2'\boldsymbol{\beta}, \cdots, \boldsymbol{\lambda}_t'\boldsymbol{\beta}$, tal que H_0 é verdadeira se e somente se $\boldsymbol{\lambda}_1'\boldsymbol{\beta} = \boldsymbol{\lambda}_2'\boldsymbol{\beta} = \cdots = \boldsymbol{\lambda}_t'\boldsymbol{\beta} = 0$.

Hipóteses testáveis

Sejam as hipóteses

$$H_0 : \boldsymbol{C}_{(q\times p)}\boldsymbol{\beta}_{(p\times 1)} = \boldsymbol{M}_{(q\times 1)} \ vs \ H_1 : \boldsymbol{C}_{(q\times p)}\boldsymbol{\beta}_{(p\times 1)} \neq \boldsymbol{M}_{(p\times 1)}.$$

A estatística (usual) do teste dependerá de $\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}$.

Se $\boldsymbol{C} = \left[\begin{array}{cccc} \boldsymbol{\lambda}_1' & \boldsymbol{\lambda}_2' & \ldots & \boldsymbol{\lambda}_q \end{array} \right]$, com $\boldsymbol{\lambda}_i\boldsymbol{\beta}$ sendo uma função estimável, $i = 1, 2, ..., q$, então $\boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{M}$ é uma hipótese testável.

Antes de discutirmos sobre a situação em que $\boldsymbol{X}$ tem posto coluna incompleto, voltemos ao caso anterior.

Construção da Estatística do Teste (posto coluna completo)

Vimos que:

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M} \sim N_q(\boldsymbol{C}\boldsymbol{\beta} - \boldsymbol{M}, \sigma^2 \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}').$$

Como $\widehat{\boldsymbol{\beta}} \perp \widehat{\sigma}^2$, então $\widehat{\boldsymbol{\theta}} \perp \widehat{\sigma}^2$, em que

$$\widehat{\sigma}^2 = \frac{1}{n-p}(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}) = \frac{1}{n-p}\boldsymbol{Y}'(\boldsymbol{I} - \boldsymbol{H})\boldsymbol{Y} = \frac{SQR}{n-p} = QMR$$

Portanto, sob $H_0(\boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{M})$ e usando alguns resultados de distribuições de formas quadráticas, temos que

$$Q^* = \frac{1}{\sigma^2}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right)'\left(\boldsymbol{C}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{C}'\right)^{-1}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right) \sim \chi^2_{(q)}$$

Além disso, vimos que $(n-p)\widehat{\sigma}^2/\sigma^2 \sim \chi^2_{(n-p)}$.

Portanto, pelos resultados acima, temos, sob H_0, que:

$$F = \frac{Q^*/q}{\widehat{\sigma}^2/\sigma^2} = \frac{1}{q\widehat{\sigma}^2}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right)'\left(\boldsymbol{C}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{C}'\right)^{-1}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right) \sim F_{(q,n-p)}$$

$p - valor = P(F > f|H_0)$, em que f é o valor calculado da estatística definida acima, e $F \sim F_{(q,n-p)}$.

Sob H_1, $F \sim F_{\left[q,n-p,\delta=\frac{1}{\sigma^2}\left((\boldsymbol{C}\boldsymbol{\beta}-\boldsymbol{M})'\left(\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}'\right)^{-1}(\boldsymbol{C}\boldsymbol{\beta}-\boldsymbol{M})\right)\right]}$.

Construção da Estatística do Teste (posto incompleto)

Posto coluna incompleto: Os desenvolvimnto são semelhantes ao caso em que $\boldsymbol{X}$ tem posto coluna completo.

Sabemos que:

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{G}\boldsymbol{X}'\boldsymbol{Y} \sim N_{col(\boldsymbol{X})}(\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{G}'\sigma^2).$$

Portanto, segue que:

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M} \sim N_q(\boldsymbol{C}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{M}, \boldsymbol{C}\boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{G}'\boldsymbol{C}'\sigma^2).$$

Defina:

$$\widehat{\sigma}^2 = \frac{1}{n - r(\boldsymbol{X})}\boldsymbol{Y}'\left(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}'\right)\boldsymbol{Y} = \frac{SQR}{n - r(\boldsymbol{X})}$$

Temos que $(n - r(\boldsymbol{X}))\widehat{\sigma}^2/\sigma^2 \sim \chi^2_{(n-r(\boldsymbol{X}))}$.

Portanto, sob $H_0(\boldsymbol{C\beta} = \boldsymbol{M})$ e usando alguns resultados de distribuições de formas quadráticas e algebrismos, temos que

$$Q^{**} = \frac{1}{\sigma^2}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right)'\left(\boldsymbol{CGC}'\right)^{-1}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right) \sim \chi^2_{(q)}$$

Para provar o resultado acima, usa-se a seguinte igualdade

$$Q^{**} = \frac{1}{\sigma^2}(\boldsymbol{Y}-\boldsymbol{XC}'(\boldsymbol{CC}')^{-}\boldsymbol{M})'(\boldsymbol{XG}'\boldsymbol{C}'(\boldsymbol{CGC}')^{-1}\boldsymbol{CGX}')(\boldsymbol{Y}-\boldsymbol{XC}'(\boldsymbol{CC}')^{-}\boldsymbol{M}) \quad (1)$$

Através da expressão (1), pode-se demonstrar ainda que $\widehat{\boldsymbol{\theta}} \perp \widehat{\sigma}^2$, verificando que

$$(\boldsymbol{X}\boldsymbol{G}'\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{G}\boldsymbol{C}')^{-1}\boldsymbol{C}\boldsymbol{G}\boldsymbol{X}')(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{G}\boldsymbol{X}') = \mathbf{0}$$

Finalmente, com os desenvolvimentos acima, concluí-se que, sob H_0

$$F^* = \frac{Q^{**}/q}{\widehat{\sigma}^{*2}/\sigma^2} = \frac{1}{q\widehat{\sigma}^{*2}}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right)'\left(\boldsymbol{C}\boldsymbol{G}\boldsymbol{C}'\right)^{-1}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right) \sim F_{(q,n-r(\boldsymbol{X}))}$$

e, sob H_1 $F^* \sim F_{(q,n-r(\boldsymbol{X}),\delta)}$

$$\delta = \frac{1}{\sigma^2}\left[(\boldsymbol{C}\boldsymbol{\beta} - \boldsymbol{M})'(\boldsymbol{C}\boldsymbol{G}\boldsymbol{C}')^{-1}(\boldsymbol{C}\boldsymbol{\beta} - \boldsymbol{M}))\right]$$

Exemplo 2 (dados de absorbância)

Utilizando o modelo desvios sem restrição, obtemos:

$\widehat{\boldsymbol{\beta}} = (0,393; 0,145; 0,173; 0,056; 0,214; -0,196)$.

As médias dos grupos estimadas pelo modelo são

$\widehat{\boldsymbol{\mu}} = (0,539; 0,566; 0,449; 0,607; 0,197)$.

as quais correspondem às médias amostrais (como esperado).

As diferenças entre a média do grupo (E50) e os demais grupos, estimadas pelo modelo foram

$\widehat{\alpha} = (0,028; -0,090; 0,069; -0,343)$.

as quais correspondem às estimativas obtidas anteriormente sob a parametrização casela de referência.

Referência

- Notas de aula do Prof. Caio Azevedo.
- Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. Wiley series in probability and statistics.