

MODELOS DE ANÁLISE DE VARIÂNCIA
(ANOVA)
MODELOS DE POSTO INCOMPLETO

CAPÍTULO 12. MODELOS DE ANÁLISE DE VARIÂNCIA

PROBLEMA: O pesquisador aplica diversos tratamentos a unidades experimentais e deseja comparar as médias dos tratamentos para alguma resposta y .

- Utilizamos modelos lineares para facilitar as comparações entre as médias dos tratamentos.

OBJETIVO: Estudar procedimentos para estimação de parâmetros e testes de hipóteses em modelos de posto incompleto, também chamados modelos de ANOVA.

Trataremos de modelos balanceados, cujos tratamentos têm números iguais de observações.

Os modelos desbalanceados serão tratados no Capítulo 15.

12.1 MODELOS DE POSTO INCOMPLETO

12.1.1 MODELO COM UM FATOR (*one-way model*)

Exemplo: Um pesquisador deseja comparar dois aditivos usados para melhorar o desempenho da gasolina. Suponhamos que:

Sem aditivos $\Rightarrow$ carro percorre μ quilômetros por litro

Com aditivo 1 $\Rightarrow$ acréscimo de τ_1 quilômetros por litro.

Com aditivo 2 $\Rightarrow$ acréscimo de τ_2 quilômetros por litro.

O modelo linear para estudar o problema pode ser expresso como:

$$y_1 = \mu + \tau_1 + \varepsilon_1$$

$$y_2 = \mu + \tau_2 + \varepsilon_2$$

onde y_1 é a quilometragem por litro para um tanque de gasolina contendo o aditivo 1 e ε_1 é um erro aleatório. As variáveis y_2 e ε_2 são definidas similarmente.

A partir da observação dos valores y_1 e y_2 o pesquisador deseja:

- Estimar os valores de μ , τ_1 e τ_2
- Testar a hipótese $H_0: \tau_1 = \tau_2$

Como resolver este problema?

Uma sugestão: Encher os tanques de seis carros idênticos com gasolina, colocar o aditivo 1 nos tanques de três carros escolhidos ao acaso e o aditivo 2 nos tanques de outros três carros e medir o desempenho (*km/litro*) de cada um dos carros.

O modelo final para as 6 observações pode ser escrito como:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2 \text{ e } j = 1, 2, 3 \quad (12.2)$$

onde

y_{ij} é a quilometragem/litro observada no j -ésimo carro que recebeu o i -ésimo aditivo em seu tanque.

ε_{ij} é um erro aleatório associado a y_{ij} .

Matricialmente as equações em (12.2) ficam:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{cases} \mu + \tau_1 + \varepsilon_{11} \\ \mu + \tau_1 + \varepsilon_{12} \\ \mu + \tau_1 + \varepsilon_{13} \\ \mu + \tau_2 + \varepsilon_{21} \\ \mu + \tau_2 + \varepsilon_{22} \\ \mu + \tau_2 + \varepsilon_{23} \end{cases} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \end{bmatrix} \quad (12.3)$$

Na sua forma matricial o modelo fica:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Em que

$$\mathbf{y} \text{ é } 6 \times 1, \mathbf{X} \text{ é } 6 \times 3, \boldsymbol{\beta} \text{ é } 3 \times 1 \text{ e } \boldsymbol{\varepsilon} \text{ é } 6 \times 1.$$

Problemas:

- $\mathbf{X}$ é 6×3 e $\text{posto}(\mathbf{X}) = 2$, ou seja, $\mathbf{X}$ é de posto incompleto $\Rightarrow$ Teoremas dos Capítulos 7 e 8 não podem ser usados diretamente para estimar $\boldsymbol{\beta}$. O modelo (12.2) é dito superparametrizado.
- Os parâmetros μ , τ_1 e τ_2 não são únicos!

Vamos entender o motivo: Admitindo, por exemplo, que $\mu = 15$, $\tau_1 = 1$ e $\tau_2 = 3$, o modelo (12.2) fica:

$$\begin{cases} y_{1j} = 16 + \varepsilon_{1j}, & j = 1, 2, 3 \\ y_{2j} = 18 + \varepsilon_{2j}, & j = 1, 2, 3 \end{cases} \quad (12.4)$$

Pergunta: Partindo das equações em (12.4) nós conseguimos deduzir que $\mu = 15$, $\tau_1 = 1$ e $\tau_2 = 3$?

Problema: O sistema $\begin{cases} \mu + \tau_1 = 16 \\ \mu + \tau_2 = 18 \end{cases}$ tem solução única, como na regressão linear múltipla?

Resposta: Não! O sistema $\begin{cases} \mu + \tau_1 = 16 \\ \mu + \tau_2 = 18 \end{cases}$ tem infinitas soluções!

Por exemplo: (1) $\mu = 10, \tau_1 = 6$ e $\tau_2 = 8$

(2) $\mu = 25, \tau_1 = -9$ e $\tau_2 = -7$

(3) $\mu = 17, \tau_1 = -1$ e $\tau_2 = 1$

Obtemos três soluções diferentes para o sistema $\Rightarrow$ os parâmetros μ, τ_1 e τ_2 não são únicos no modelo superparametrizado e não podem ser estimados como nos capítulos 7 e 8.

Algumas formas de remediar a falta de unicidade dos parâmetros no modelo superparametrizado:

- 1) Redefinir o modelo (12.2) usando dois novos parâmetros que sejam únicos (realizar uma reparametrização do modelo)

Sugestão: Utilizar o modelo $y_{ij} = \mu_i + \varepsilon_{ij}$, em que $\mu_i = \mu + \tau_i$ é a média de consumo dos carros depois de receberem o aditivo i .

Matricialmente tem-se:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \end{bmatrix} = \mathbf{W}\boldsymbol{\mu} + \boldsymbol{\varepsilon}$$

Como $\text{posto}(\mathbf{W}) = 2 \Rightarrow \hat{\boldsymbol{\mu}} = (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{y}$ é a solução única!

Esta proposta é chamada de reparametrização e o modelo final é chamado modelo de médias de caselas.

Note que a hipótese de interesse $H_0: \tau_1 = \tau_2$ a ser testada no modelo (12.2), também pode ser testada neste novo modelo pois sabendo que $\mu_1 = \mu + \tau_1$ e $\mu_2 = \mu + \tau_2$, fazendo $\mu_1 = \mu_2$ reproduzimos a hipótese: $H_0: \tau_1 = \tau_2$.

Isso quer dizer que testar $H_0: \tau_1 = \tau_2$ no modelo (12.2) é equivalente a testar $H_0: \mu_1 = \mu_2$ no modelo de médias de caselas.

2) Usar o modelo (12.2) e incluir algumas restrições nos parâmetros de modo a torná-los únicos.

Sugestão: Impor a condição a condição marginal $\tau_1^* + \tau_2^* = 0$ aos parâmetros do modelo (restrito) $y_{ij} = \mu^* + \tau_i^* + \varepsilon_{ij}$.

Neste novo contexto: μ^* é a nova quilometragem média depois dos aditivos químicos serem aplicados e τ_1^* e τ_2^* são desvios dessa média (ou efeitos dos aditivos).

Com esta restrição o sistema
$$\begin{cases} \mu^* + \tau_1^* = 16 \\ \mu^* + \tau_2^* = 18 \\ \tau_1^* + \tau_2^* = 0 \end{cases}$$
 passa a ter solução

única:

$$\mu^* = 17, \tau_1^* = -1 \text{ e } \tau_2^* = 1$$

A interpretação dos parâmetros sujeitos à condição $\sum \tau_i^* = 0$ é alterada para:

$\Rightarrow$ Depois do uso de aditivos, o desempenho médio dos carros é 17 km/l ; com o aditivo 1 o desempenho médio é 1 km/l abaixo desta média e com o aditivo 2, é 1 km/l acima desta média.

O modelo $y_{ij} = \mu^* + \tau_i^* + \varepsilon_{ij}$, sujeito à restrição $\tau_1^* + \tau_2^* = 0$, pode ser expresso no formato de um modelo de posto completo, substituindo $\tau_2^* = -\tau_1^*$ para obter:

$$y_{1j} = \mu^* + \tau_1^* + \varepsilon_{1j}$$

$$y_{2j} = \mu^* - \tau_1^* + \varepsilon_{2j}$$

Matricialmente temos:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \mu^* \\ \tau_1^* \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \end{bmatrix} = \mathbf{X}^* \boldsymbol{\beta}^* + \boldsymbol{\varepsilon}$$

Incorporando a restrição ao modelo a matriz $\mathbf{X}^*$ passa a ter posto completo e os parâmetros $\boldsymbol{\mu}^*$ e $\boldsymbol{\tau}_1^*$ podem ser estimados de forma única (capítulos 7 e 8). Depois estimamos $\boldsymbol{\tau}_2^* = -\boldsymbol{\tau}_1^*$.

3) Trabalhar somente com combinações lineares estimáveis (únicas) dos parâmetros.

Sugestão: Como $\text{posto}(\mathbf{X}) = 2$, vamos escolher somente duas combinações lineares dos parâmetros que sejam únicas, dentre as três combinações que já conhecemos:

$$\tau_1 - \tau_2, \quad \mu + \tau_1 \quad \text{e} \quad \mu + \tau_2$$

Note que os valores dessas combinações de parâmetros não se alteram quando usamos qualquer uma das soluções obtidas para o modelo (12.4).

Nota: Aumentar o número de repetições de cada tratamento não resolve o problema de estimação dos parâmetros do modelo (12.2).

12.1.2 MODELO COM DOIS FATORES (*two-way model*)

Exemplo: Pesquisador deseja avaliar o efeito de duas diferentes vitaminas e de dois diferentes métodos de administrar essas vitaminas sobre o ganho de peso de frangos $\Rightarrow$ temos um modelo com dois fatores (*two-way model*).

Seja α_i o efeito da i -ésima vitamina e β_j o efeito do j -ésimo método de administração. Se o pesquisador assume que esses efeitos são aditivos e com uma única repetição (sem perda de generalidade) para cada combinação vitamina-método, o modelo pode ser escrito:

$$y_{11} = \mu + \alpha_1 + \beta_1 + \varepsilon_{11}$$

$$y_{12} = \mu + \alpha_1 + \beta_2 + \varepsilon_{12}$$

$$y_{21} = \mu + \alpha_2 + \beta_1 + \varepsilon_{21}$$

$$y_{22} = \mu + \alpha_2 + \beta_2 + \varepsilon_{22}$$

ou

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \text{ para } i=1, 2 \text{ e } j = 1, 2 \quad (12.5)$$

onde y_{ij} é o ganho de peso do (ij) -ésimo frango e ε_{ij} é o erro aleatório associado a y_{ij}

Matricialmente temos:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{22} \end{bmatrix} \quad (12.6)$$

Problema: A matriz $\mathbf{X}$ é 4×5 , temos 5 parâmetros para estimar e $\text{posto}(\mathbf{X}) = 3 \Rightarrow$ os Teoremas dos Capítulos 7 e 8 também não são aplicáveis.

Já sabemos que:

- Aumentar o número de repetições não altera o ***posto*(X)**.
- Se ***posto*(X) = 3** $\Rightarrow$ somente 3 parâmetros são únicos (ou estimáveis).
- É do interesse do pesquisador, se possível, testar as hipóteses:

$$H_0: \alpha_1 = \alpha_2 \text{ e } H_0: \beta_1 = \beta_2$$

SOLUÇÕES POSSÍVEIS:

- 1) Reparametrização: Consideremos os três novos parâmetros γ_1 , γ_2 e γ_3 , definidos como:

$$\gamma_1 = \mu + \alpha_1 + \beta_1, \quad \gamma_2 = \alpha_2 - \alpha_1 \quad \text{e} \quad \gamma_3 = \beta_2 - \beta_1$$

Note que γ_2 e γ_3 são parâmetros de interesse do pesquisador.

O novo modelo pode ser escrito em termos dos γ 's como:

$$y_{11} = (\mu + \alpha_1 + \beta_1) + \varepsilon_{11} = \gamma_1 + \varepsilon_{11}$$

$$y_{12} = (\mu + \alpha_1 + \beta_1) + (\beta_2 - \beta_1) + \varepsilon_{12} = \gamma_1 + \gamma_3 + \varepsilon_{12}$$

$$y_{21} = (\mu + \alpha_1 + \beta_1) + (\alpha_2 - \alpha_1) + \varepsilon_{21} = \gamma_1 + \gamma_2 + \varepsilon_{21}$$

$$\begin{aligned} y_{22} &= (\mu + \alpha_1 + \beta_1) + (\alpha_2 - \alpha_1) + (\beta_2 - \beta_1) + \varepsilon_{22} \\ &= \gamma_1 + \gamma_2 + \gamma_3 + \varepsilon_{22} \end{aligned}$$

Matricialmente, o novo modelo fica:

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad (12.7)$$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \end{bmatrix} = \begin{cases} \gamma_1 + \varepsilon_{11} \\ \gamma_1 + \gamma_3 + \varepsilon_{12} \\ \gamma_1 + \gamma_2 + \varepsilon_{21} \\ \gamma_1 + \gamma_2 + \gamma_3 + \varepsilon_{22} \end{cases} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{22} \end{bmatrix}$$

Como $\text{posto}(\mathbf{Z}) = 3 \Rightarrow \hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$ fornece as estimativas dos novos parâmetros $\boldsymbol{\gamma}_1 = \boldsymbol{\mu} + \boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1$, $\boldsymbol{\gamma}_2 = \boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1$ e $\boldsymbol{\gamma}_3 = \boldsymbol{\beta}_2 - \boldsymbol{\beta}_1$.

Testar $H_0: \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2$ e $H_0: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ no modelo (12.6)

é equivalente a

testar $H_0: \boldsymbol{\gamma}_2 = \mathbf{0}$ e $H_0: \boldsymbol{\gamma}_3 = \mathbf{0}$, respectivamente, no modelo (12.7) reparametrizado.

Na seção 12.2.2 mostraremos que funções lineares como

$$(\boldsymbol{\mu} + \boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1), \quad (\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1) \quad \text{e} \quad (\boldsymbol{\beta}_2 - \boldsymbol{\beta}_1)$$

são únicas e estimáveis, mesmo quando os parâmetros $\boldsymbol{\mu}$, $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\boldsymbol{\beta}_1$ e $\boldsymbol{\beta}_2$ não são únicos, nem estimáveis (como neste exemplo!)

2) Restrições sobre os parâmetros

O modelo (12.6) tem 5 parâmetros e $\text{posto}(\mathbf{X}) = 3$

⇒ A deficiência de *rank* da matriz $\mathbf{X}$ é $5 - 3 = 2$

⇒ Devemos impor 2 (duas) condições marginais *l. i.* para completar o *posto* da matriz $\mathbf{X}$.

Antes de escolher e impor as condições marginais vamos escrever o sistema de equações normais $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$:

$$\begin{bmatrix} 4 & 2 & 2 & 2 & 2 \\ 2 & 2 & 0 & 1 & 1 \\ 2 & 0 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 & 0 \\ 2 & 1 & 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{.1} \\ y_{.2} \end{bmatrix}$$

Ou

$$\begin{aligned} 4\hat{\mu} + 2(\hat{\alpha}_1 + \hat{\alpha}_2) + 2(\hat{\beta}_1 + \hat{\beta}_2) &= y_{..} \\ 2\hat{\mu} + 2\hat{\alpha}_1 + (\hat{\beta}_1 + \hat{\beta}_2) &= y_{1.} \\ 2\hat{\mu} + 2\hat{\alpha}_2 + (\hat{\beta}_1 + \hat{\beta}_2) &= y_{2.} \\ 2\hat{\mu} + (\hat{\alpha}_1 + \hat{\alpha}_2) + 2\hat{\beta}_1 &= y_{.1} \\ 2\hat{\mu} + (\hat{\alpha}_1 + \hat{\alpha}_2) + 2\hat{\beta}_2 &= y_{.2} \end{aligned}$$

Sugestão: Impor as restrições $\hat{\alpha}_1 + \hat{\alpha}_2 = 0$ e $\hat{\beta}_1 + \hat{\beta}_2 = 0$

Resolvendo o S.E.N. sob essas condições tem-se:

$$\hat{\mu} = (1/4)y_{..} = \bar{y}_{..}$$

$$\hat{\alpha}_1 = (1/2)y_{1.} - \hat{\mu} = \bar{y}_{1.} - \bar{y}_{..}$$

$$\hat{\alpha}_2 = \bar{y}_{2.} - \bar{y}_{..}$$

$$\hat{\beta}_1 = (1/2)\bar{y}_{.1} - \hat{\mu} = \bar{y}_{.1} - \bar{y}_{..}$$

$$\hat{\beta}_2 = \bar{y}_{.2} - \bar{y}_{..}$$

Com as condições marginais impostas aos parâmetros:

- $\hat{\alpha}_i$'s e $\hat{\beta}_j$'s passam a ser interpretados como desvios de médias, ou como os efeitos de vitamina e dos métodos de administração das vitaminas no peso dos frangos, respectivamente.
- Impondo $\hat{\alpha}_1 + \hat{\alpha}_2 = 0$ e $\hat{\beta}_1 + \hat{\beta}_2 = 0 \Rightarrow$ os novos parâmetros são únicos.

Para entender o significado de um modelo aditivo com dois fatores, vamos escrever a média da casela (i, j) , $E(y_{ij}) = \mu_{ij}$, como:

$$\begin{aligned}\mu_{ij} &= \bar{\mu}_{..} + (\bar{\mu}_{i.} - \bar{\mu}_{..}) + (\bar{\mu}_{.j} - \bar{\mu}_{..}) + (\mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..}) \\ &= \mu^* + \alpha_i^* + \beta_j^* + (\mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..})\end{aligned}$$

- Para que α_i^* e β_j^* sejam considerados aditivos, precisamos admitir que $\mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..} = 0$.
- Este termo, $\mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..}$, é chamado interação entre os níveis dos fatores vitamina e método de administração.
- Modelos com interação serão discutidos no Capítulo 14.

12.2. ESTIMAÇÃO

12.2.1 Estimabilidade de β

Consideremos o modelo linear geral $y = X\beta + \varepsilon$, em que $E(y) = X\beta$, $cov(y) = \sigma^2 I$ e X é $n \times p$, de posto $k < p \leq n$.

Utilizando o Método dos Quadrados Mínimos buscamos $\hat{\beta}$ que minimize:

$$\hat{\varepsilon}'\hat{\varepsilon} = (y - X\hat{\beta})'(y - X\hat{\beta}) = y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta} \quad (12.10)$$

Diferenciando (12.10) com respeito a $\hat{\beta}$ e igualando a 0 tem-se o Sistema de Equações Normais:

$$X'X\hat{\beta} = X'y \quad (12.11)$$

Teorema 12.2A. Se $\mathbf{X}$ é $n \times p$ de posto $k < p \leq n$, o sistema $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ é consistente. Uma solução do sistema é:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y} \quad (12.13)$$

onde $(\mathbf{X}'\mathbf{X})^{-}$ é *qualquer* inversa generalizada de $(\mathbf{X}'\mathbf{X})$.

Note que:

$$E(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'E(\mathbf{y}) = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \quad (12.14)$$

Ou seja:

- $\hat{\boldsymbol{\beta}}$ em (12.13) é um estimador **viesado** de $\boldsymbol{\beta}$.
- $\hat{\boldsymbol{\beta}}$ estima $(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$ sem viés, ou seja, $\hat{\boldsymbol{\beta}}$ estima sem viés **apenas** algumas combinações dos parâmetros em $\boldsymbol{\beta}$ (Quais?)

Problema: O produto $(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$ não é invariante a escolhas de diferentes inversas generalizadas $(\mathbf{X}'\mathbf{X})^{-}$.

Exemplo 12.2.1. Consideremos o modelo $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, $i = 1, 2$ e $j = 1, 2, 3$, em (12.2). A matriz $\mathbf{X}$ e o vetor $\boldsymbol{\beta}$ são dados como:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{e} \quad \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix}$$

Então:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix} \quad \text{e} \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{bmatrix}$$

onde $y_{..} = \sum_{i=1}^2 \sum_{j=1}^3 y_{ij}$ e $y_{i.} = \sum_{j=1}^3 y_{ij}$.

Pelo Teorema 2.2.C(i) e pelo Corolário 1 do Teorema 2.8B tem-se que **um** estimador de $\boldsymbol{\beta}$ é igual a:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{y}_{1.} \\ \bar{y}_{2.} \end{bmatrix}$$

onde $\bar{y}_{i.} = \frac{1}{3} y_{i.}$

Por outro lado:

$$\begin{aligned} E(\bar{y}_{i.}) &= E\left(\frac{1}{3}\sum_{j=1}^3 y_{ij}\right) = \frac{1}{3}\sum_{j=1}^3 E(y_{ij}) \\ &= \frac{1}{3}\sum_{j=1}^3 E(\mu + \tau_i + \varepsilon_{ij}) = \frac{1}{3}(3\mu + 3\tau_i) = \mu + \tau_i \end{aligned}$$

De (12.14) temos que

$$\begin{aligned}
 E(\hat{\boldsymbol{\beta}}) &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \mu + \tau_1 \\ \mu + \tau_2 \end{bmatrix}
 \end{aligned}$$

$\therefore \hat{\boldsymbol{\beta}} = \begin{bmatrix} 0 \\ \bar{y}_{1\bullet} \\ \bar{y}_{2\bullet} \end{bmatrix}$ estima **sem viés** as combinações dos parâmetros:

$$\mu + \tau_1 \quad \text{e} \quad \mu + \tau_2$$

12.2.2. Funções Estimáveis de β

- Não podemos estimar todos os parâmetros do vetor β de um modelo superparametrizado.
- Quais combinações lineares dos β 's, $\lambda'\beta$, poderão ser estimadas?

Definição: Uma função linear dos parâmetros, $\lambda'\beta$, é dita ser estimável se existe uma combinação linear das observações ($\mathbf{a}'\mathbf{y}$) que tenha um valor esperado igual a $\lambda'\beta$. Isto é,

$\lambda'\beta$ é estimável se existe um vetor $\mathbf{a}$ tal que $E(\mathbf{a}'\mathbf{y}) = \mathbf{a}'\mathbf{X}\beta = \lambda'\beta$.

⇒ Encontrar este vetor $\mathbf{a}$ pode ser uma tarefa bem difícil...

⇒ **Dica:** O vetor $\mathbf{a}$ deve indicar uma combinação das linhas de $\mathbf{X}\beta$, que reproduza a função linear $\lambda'\beta$.

Teorema 12.2B. No modelo $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, onde $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$, $\mathbf{X}$ é $n \times p$ de posto $k < p \leq n$, a função linear $\boldsymbol{\lambda}'\boldsymbol{\beta}$ é dita **estimável** se e somente se qualquer uma das seguintes condições é satisfeita:

- i) $\boldsymbol{\lambda}'$ é uma combinação linear das linhas de $\mathbf{X}$, isto é, existe um vetor $\mathbf{a}$ tal que:

$$\mathbf{a}'\mathbf{X} = \boldsymbol{\lambda}' \quad (12.15)$$

Isto acontece se $\text{posto}(\mathbf{X}' : \boldsymbol{\lambda}) = \text{posto}(\mathbf{X}') = \text{posto}(\mathbf{X})$.

- ii) $\boldsymbol{\lambda}'$ é uma combinação linear das linhas (ou das colunas) de $\mathbf{X}'\mathbf{X}$, isto é, existe um vetor $\mathbf{r}$ tal que:

$$\mathbf{r}'\mathbf{X}'\mathbf{X} = \boldsymbol{\lambda}' \quad \text{ou} \quad \mathbf{X}'\mathbf{X} \mathbf{r} = \boldsymbol{\lambda} \quad (12.16)$$

Isto acontece se $\text{posto}(\mathbf{X}'\mathbf{X} : \boldsymbol{\lambda}) = \text{posto}(\mathbf{X}'\mathbf{X}) = \text{posto}(\mathbf{X})$.

...continua

iii) λ (ou λ') é tal que

$$\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\lambda = \lambda \quad \text{ou} \quad \lambda'(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X} = \lambda' \quad (12.17)$$

onde $(\mathbf{X}'\mathbf{X})^{-}$ é uma inversa generalizada (simétrica) de $\mathbf{X}'\mathbf{X}$.

Exemplo 12.2.2(a). Será que a função $\lambda'\boldsymbol{\beta} = [0, 1, -1]\boldsymbol{\beta} = \tau_1 - \tau_2$ é estimável no modelo $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ do Exemplo 12.2.1?

Como exercício, vamos verificar a estimabilidade da função $\tau_1 - \tau_2$ usando cada uma das condições impostas no Teorema 12.2B.

Problema: Encontrar uma combinação das linhas de $\mathbf{X}$ que reproduza $\lambda' = [0, 1, -1]$.

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \text{ e } \boldsymbol{\lambda}' = [0, 1, -1]$$

i) Note que 3ª linha – 4ª linha de $\mathbf{X}$ é igual a $\boldsymbol{\lambda}' \Rightarrow$ Considerando $\boldsymbol{a}' = [0, 0, 1, -1, 0, 0]$ tem-se:

$$\boldsymbol{a}'\mathbf{X} = [0, 0, 1, -1, 0, 0] \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = [0, 1, -1] = \boldsymbol{\lambda}'$$

Conclusão: A função $\boldsymbol{\lambda}'\boldsymbol{\beta} = [0, 1, -1]\boldsymbol{\beta} = \tau_1 - \tau_2$ é estimável.

Existem outras escolhas para o vetor $\mathbf{a}$ que satisfazem $\mathbf{a}'\mathbf{X} = \lambda'$, como por exemplo:

$$\mathbf{a}' = [1, 0, 0, 0, 0, -1] \text{ ou } \mathbf{a}' = [2, -1, 0, 0, 1, -2].$$

Para mostrar que $\lambda'\boldsymbol{\beta} = [0, 1, -1]\boldsymbol{\beta}$ é estimável é mais fácil verificar que:

$$\text{posto}(\mathbf{X}': \lambda) = \text{posto}(\mathbf{X}') = \text{posto}(\mathbf{X}) = 2$$

Outro problema: Encontrar uma combinação das colunas de $\mathbf{X}'\mathbf{X}$ que reproduza λ .

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix} \text{ e } \lambda = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

ii) $(1/3)[2^{\text{a}} \text{ coluna} - 3^{\text{a}} \text{ coluna de } \mathbf{X}'\mathbf{X}]$ é igual a $\lambda \Rightarrow$ Escolhendo $\mathbf{r} = [0, 1/3, -1/3]'$ tem-se:

$$\mathbf{X}'\mathbf{X} \mathbf{r} = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1/3 \\ -1/3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \lambda$$

$\therefore \lambda' \boldsymbol{\beta} = \tau_1 - \tau_2$ é estimável.

Certamente, existem outros valores possíveis de $\mathbf{r}$, tais como $\mathbf{r} = [-1/3, 2/3, 0]'$, que servem para verificar $\mathbf{X}'\mathbf{X} \mathbf{r} = \lambda$.

Alternativa mais simples: $\lambda' \boldsymbol{\beta} = [0, 1, -1] \boldsymbol{\beta}$ é estimável porque

$$\text{posto}(\mathbf{X}'\mathbf{X} : \lambda) = \text{posto}(\mathbf{X}'\mathbf{X}) = \text{posto}(\mathbf{X}) = 2$$

iii) Utilizando a inversa generalizada $(\mathbf{X}'\mathbf{X})^- = \text{diag}[0, 1/3, 1/3]$, obtida no Exemplo 12.2.1, temos:

$$\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^- = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^- \boldsymbol{\lambda} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \boldsymbol{\lambda}$$

Concluimos de diversas formas que

$$\boldsymbol{\lambda}'\boldsymbol{\beta} = [0, 1, -1]\boldsymbol{\beta} = \tau_1 - \tau_2$$

é uma função estimável no modelo 12.2.1.

Definição: Um conjunto de funções lineares e estimáveis $\{\lambda_1'\boldsymbol{\beta}, \lambda_2'\boldsymbol{\beta}, \dots, \lambda_m'\boldsymbol{\beta}\}$ é dito linearmente independente (*l. i.*) se todos os vetores $\lambda_1, \lambda_2, \dots, \lambda_m$ forem linearmente independentes.

Teorema 12.2C. No modelo de posto incompleto $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, o número de funções estimáveis e *l. i.* de $\boldsymbol{\beta}$ é igual ao $\text{posto}(\mathbf{X}) = k$.

(*Prova:* ver Graybill, 1976, pp. 485-486)

Importante:

- Toda linha de $\mathbf{X}\boldsymbol{\beta}$ é uma função estimável $\Rightarrow \mathbf{X}\boldsymbol{\beta}$ é estimável.
- Toda linha de $\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$ é uma função estimável $\Rightarrow \mathbf{X}'\mathbf{X}\boldsymbol{\beta}$ é estimável.

Teorema 12.2D. No modelo $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, onde $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ e $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$, qualquer função estimável $\boldsymbol{\lambda}'\boldsymbol{\beta}$ pode ser obtida como uma combinação linear das linhas (elementos) de $\mathbf{X}\boldsymbol{\beta}$ ou das linhas de $\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$.

Exemplo 12.2.2(b). Consideremos o modelo (12.6) com dois fatores da Seção 12.1.2 com

$$\mathbf{X}\boldsymbol{\beta} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \mu + \alpha_1 + \beta_1 \\ \mu + \alpha_1 + \beta_2 \\ \mu + \alpha_2 + \beta_1 \\ \mu + \alpha_2 + \beta_2 \end{bmatrix}$$

- Todas as linhas de $\mathbf{X}\boldsymbol{\beta}$ são estimáveis, como por exemplo a 1ª linha: $\boldsymbol{\lambda}'_1\boldsymbol{\beta} = [1 \ 1 \ 0 \ 1 \ 0]\boldsymbol{\beta} = \mu + \alpha_1 + \beta_1$.

- Combinações lineares das linhas de $\mathbf{X}\boldsymbol{\beta}$ também são estimáveis!

Exemplos:

2ª linha - 1ª linha: $\boldsymbol{\lambda}'_2\boldsymbol{\beta} = \beta_2 - \beta_1$ é estimável

3ª linha - 1ª linha: $\boldsymbol{\lambda}'_3\boldsymbol{\beta} = \alpha_2 - \alpha_1$ é estimável

Observe que:

- As funções lineares $\boldsymbol{\lambda}'_1\boldsymbol{\beta}$, $\boldsymbol{\lambda}'_2\boldsymbol{\beta}$ e $\boldsymbol{\lambda}'_3\boldsymbol{\beta}$ são estimáveis e os vetores

$$\boldsymbol{\lambda}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{\lambda}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \text{ e } \boldsymbol{\lambda}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ são l.i.}$$

Note que essas funções são idênticas às funções γ_1 , γ_2 e γ_3 que foram usadas na Seção 12.1.2 para **reparametrizar** o modelo de posto incompleto para um modelo de posto completo.

Conclusão: Funções dos parâmetros que sejam *l. i.* e estimáveis podem ser usadas em reparametrizações.

Note que os coeficientes das funções estimáveis $\lambda'_2 \boldsymbol{\beta} = \beta_2 - \beta_1$ e $\lambda'_3 \boldsymbol{\beta} = \alpha_2 - \alpha_1$ somam zero.

- Toda combinação linear desse tipo, em que a soma dos seus coeficientes é nula é chamada de contraste e todo contraste é estimável.

12.3. ESTIMADORES

12.3.1 Estimadores de $\lambda'\beta$

Teorema 12.3A. Seja $\lambda'\beta$ uma **função estimável** de β no modelo $y = X\beta + \varepsilon$, onde $E(y) = X\beta$ e X é $n \times p$, de posto $k < p \leq n$. Seja $\hat{\beta}$ qualquer solução do sistema de equações $X'X\hat{\beta} = X'y$ e seja r qualquer solução para $X'Xr = \lambda$. Então os dois estimadores $\lambda'\hat{\beta}$ e $r'X'y$ têm as seguintes propriedades:

- i) $E(\lambda'\hat{\beta}) = E(r'X'y) = \lambda'\beta$, ou seja, $\lambda'\hat{\beta}$ é um estimador justo ou não viesado de $\lambda'\beta$.
- ii) $\lambda'\hat{\beta} = r'X'y$ para qualquer $\hat{\beta}$ e qualquer r .
- iii) Os estimadores $\lambda'\hat{\beta}$ e $r'X'y$ são invariantes para escolhas de $\hat{\beta}$ ou de r .

Exemplo 12.3.1. No Exemplo 12.2.2(a) mostramos que a função $\lambda'\beta = \tau_1 - \tau_2$ é estimável. Para estimar $\tau_1 - \tau_2$ com $\mathbf{r}'\mathbf{X}'\mathbf{y}$, usamos $\mathbf{r}' = [0 \ 1/3 \ -1/3]$ do Exemplo 12.2.2(a) para obter:

$$\begin{aligned} \mathbf{r}'\mathbf{X}'\mathbf{y} &= [0 \ 1/3 \ -1/3] \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} \\ &= [0 \ 1/3 \ -1/3] \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{bmatrix} = \bar{y}_{1.} - \bar{y}_{2.} \end{aligned}$$

Para obter o mesmo resultado usando $\lambda'\hat{\beta}$, precisamos encontrar uma solução do sistema $\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}$:

$$\begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{bmatrix}$$

Ou

$$\begin{cases} 6\hat{\mu} + 3\hat{\tau}_1 + 3\hat{\tau}_2 = y_{..} \\ 3\hat{\mu} + 3\hat{\tau}_1 = y_{1.} \\ 3\hat{\mu} + 3\hat{\tau}_2 = y_{2.} \end{cases}$$

Como a primeira equação é redundante (é igual à soma da segunda e terceira equações), podemos tomar $\hat{\mu}$ como uma constante arbitrária e daí obter:

$$\hat{\tau}_1 = \bar{y}_{1.} - \hat{\mu} \text{ e } \hat{\tau}_2 = \bar{y}_{2.} - \hat{\mu}$$

Assim

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{y}_{1\bullet} \\ \bar{y}_{2\bullet} \end{bmatrix} + \hat{\mu} \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

Para estimar $\tau_1 - \tau_2 = [0 \ 1 \ -1]\boldsymbol{\beta} = \boldsymbol{\lambda}'\boldsymbol{\beta}$, podemos tomar $\hat{\mu} = 0$ em $\hat{\boldsymbol{\beta}}$ para obter:

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} 0 \\ \bar{y}_{1\bullet} \\ \bar{y}_{2\bullet} \end{bmatrix} \Rightarrow \text{estimador de } \boldsymbol{\lambda}'\boldsymbol{\beta} = \tau_1 - \tau_2 \text{ é } \boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} = \bar{y}_{1\bullet} - \bar{y}_{2\bullet}.$$

Mesmo mantendo $\hat{\mu}$ arbitrário, obtemos o mesmo estimador:

$$\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} = [0 \ 1 \ -1] \begin{bmatrix} \hat{\mu} \\ \bar{y}_{1\bullet} - \hat{\mu} \\ \bar{y}_{2\bullet} - \hat{\mu} \end{bmatrix} = (\bar{y}_{1\bullet} - \hat{\mu}) - (\bar{y}_{2\bullet} - \hat{\mu}) = \bar{y}_{1\bullet} - \bar{y}_{2\bullet}.$$

PROBLEMA: Como $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y}$ não é único no modelo de posto incompleto $\mathbf{y} = \mathbf{X}\beta + \varepsilon$ onde $cov(\mathbf{y}) = \sigma^2\mathbf{I}$, $cov(\hat{\beta})$ também não é única.

Para uma particular inversa generalizada (simétrica) de $(\mathbf{X}'\mathbf{X})$ nós podemos usar o Teorema 3.6(D)i para obter:

$$cov(\hat{\beta}) = cov[(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y}] = \sigma^2(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-} \quad (12.18)$$

que **não é invariante** para diferentes escolhas de $(\mathbf{X}'\mathbf{X})^{-}$.

Mas $var(\lambda'\hat{\beta}) = \sigma^2\lambda'(\mathbf{X}'\mathbf{X})^{-}\lambda$ é invariante a diferentes escolhas de $(\mathbf{X}'\mathbf{X})^{-}$ se $\lambda'\beta$ é uma função estimável.

Teorema 12.3B. Seja $\lambda'\beta$ uma função estimável no modelo linear $y = X\beta + \varepsilon$, onde X é $n \times p$ de posto $k < p \leq n$ e $\text{cov}(y) = \sigma^2 I$. Seja r qualquer solução para $X'Xr = \lambda$ e seja $\hat{\beta}$ qualquer solução de $X'X\hat{\beta} = X'y$. Então, a variância de $\lambda'\hat{\beta}$ (ou de $r'X'y$) tem as seguintes propriedades:

$$i) \text{var}(r'X'y) = \sigma^2 r'(X'X)^- r = \sigma^2 r' \lambda$$

$$ii) \text{var}(\lambda'\hat{\beta}) = \lambda' \text{cov}(\hat{\beta}) \lambda = \sigma^2 \lambda'(X'X)^- \underbrace{X'X(X'X)^-}_{\lambda} \lambda$$

$$= \sigma^2 \lambda'(X'X)^- \lambda, \text{ porque } \lambda'\beta \text{ é estimável.}$$

iii) $\text{var}(\lambda'\hat{\beta})$ é única, ou seja, é invariante para as escolhas de r ou de $(X'X)^-$.

Teorema 12.3C. Se $\lambda'_1\beta$ e $\lambda'_2\beta$ são duas funções estimáveis no modelo $\mathbf{y} = \mathbf{X}\beta + \varepsilon$ onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$ e $\text{cov}(\mathbf{y}) = \sigma^2\mathbf{I}$, a covariância entre seus estimadores é:

$$\text{cov}(\lambda'_1\hat{\beta}, \lambda'_2\hat{\beta}) = \sigma^2 \mathbf{r}'_1 \lambda_2 = \sigma^2 \lambda'_1 \mathbf{r}_2 = \sigma^2 \lambda'_1 (\mathbf{X}'\mathbf{X})^{-} \lambda_2$$

onde $\mathbf{X}'\mathbf{X}\mathbf{r}_1 = \lambda_1$ e $\mathbf{X}'\mathbf{X}\mathbf{r}_2 = \lambda_2$.

Teorema 12.3D. Se $\lambda'\beta$ é uma função estimável no modelo $\mathbf{y} = \mathbf{X}\beta + \varepsilon$ onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$, então os estimadores $\lambda'\hat{\beta}$ e $\mathbf{r}'\mathbf{X}'\mathbf{y}$ são os melhores estimadores não viesados (BLUE) de $\lambda'\beta$.

12.3.2. Um estimador da variância (σ^2)

Por analogia com (7.23) nós definimos

$$SQResiduo = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (12.19)$$

onde $\hat{\boldsymbol{\beta}}$ é qualquer solução do sistema de equações normais $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$.

Duas expressões alternativas para $SQRes$ são:

$$SQResiduo = \mathbf{y}'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \quad (12.20)$$

$$SQResiduo = \mathbf{y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y} \quad (12.21)$$

Para um estimador de σ^2 , nós definimos:

$$s^2 = \frac{SQResiduo}{n - k} = QMResiduo \quad (12.22)$$

onde n é o número de linhas de $\mathbf{X}$ e $k = posto(\mathbf{X})$.

Teorema 12.3E. Para s^2 definido em (12.22) para um modelo $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ de posto incompleto, com $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ e $\text{cov}(\mathbf{y}) = \sigma^2 \mathbf{I}$, temos as seguintes propriedades:

- i) $E(s^2) = \sigma^2$ (provar!)
- ii) s^2 é invariante para escolhas de $\hat{\boldsymbol{\beta}}$ ou escolhas de inversas generalizadas $(\mathbf{X}'\mathbf{X})^-$

12.3.3. Modelo Normal

Para o modelo de posto incompleto $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, assumiremos que $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ ou que $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$. Com isso nós podemos obter estimadores de máxima verossimilhança.

Teorema 12.3F. Se $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$, então os estimadores de máxima verossimilhança de $\boldsymbol{\beta}$ e σ^2 são dados por:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (12.23)$$

$$\hat{\sigma}^2 = \frac{1}{n}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (12.24)$$

- O estimador de máxima verossimilhança $\hat{\boldsymbol{\beta}}$ em (12.23) é idêntico ao estimador de mínimos quadrados em (12.13).

- O estimador $\hat{\sigma}^2$ em (12.24) é viesado $\Rightarrow$ Vamos substituir $\hat{\sigma}^2$ pelo estimador não viesado, s^2 , dado em (12.22).

Teorema 12.3G. Se $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$, então os estimadores de máxima verossimilhança $\hat{\boldsymbol{\beta}}$ e s^2 (corrigido para o viés) têm as seguintes propriedades:

- i) $\hat{\boldsymbol{\beta}} \sim N_p[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]$
- ii) $(n-k)s^2/\sigma^2 \sim \chi^2_{(n-k)}$
- iii) $\hat{\boldsymbol{\beta}}$ e s^2 são independentes.

Teorema 12.3H. Se $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$, e se $\boldsymbol{\lambda}'\boldsymbol{\beta}$ é uma função estimável, então $\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}}$ tem variância mínima dentre todos os estimadores não viesados de $\boldsymbol{\lambda}'\boldsymbol{\beta}$.

12.4. REPARAMETRIZAÇÃO

Na reparametrização, transformamos um modelo de posto incompleto:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \text{ onde } \mathbf{X} \text{ é } n \times p, \text{ de posto } k < p \leq n$$

em um modelo de posto completo

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \text{ onde } \mathbf{Z} \text{ é } n \times k \text{ de posto } k$$

em que

$$\boldsymbol{\gamma} = \mathbf{U}\boldsymbol{\beta}, \text{ onde } \mathbf{U} \text{ é } k \times p \text{ de posto } k$$

é um conjunto de k funções estimáveis e linearmente independentes de $\boldsymbol{\beta}$.

Problema: Se conhecemos a reparametrização $\boldsymbol{\gamma} = \mathbf{U}\boldsymbol{\beta}$, mas precisamos encontrar a matriz $\mathbf{Z}$ do modelo reparametrizado $\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$.

Como $\mathbf{Z}\boldsymbol{\gamma} = \mathbf{X}\boldsymbol{\beta}$ e $\boldsymbol{\gamma} = \mathbf{U}\boldsymbol{\beta}$ podemos escrever:

$$\mathbf{Z}\boldsymbol{\gamma} = \mathbf{Z}\mathbf{U}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta} \Rightarrow \mathbf{X} = \mathbf{Z}\mathbf{U} \quad (12.30)$$

Pós-multiplicando ambos os lados de (12.30) por $\mathbf{U}'$, tem-se:

$$\mathbf{X}\mathbf{U}' = \mathbf{Z}\mathbf{U}\mathbf{U}'$$

Como a matriz $\mathbf{U}\mathbf{U}'$ é não singular, pós multiplicando ambos os lados por $(\mathbf{U}\mathbf{U}')^{-1}$ temos:

$$\mathbf{X}\mathbf{U}'(\mathbf{U}\mathbf{U}')^{-1} = \mathbf{Z}\mathbf{U}\mathbf{U}'(\mathbf{U}\mathbf{U}')^{-1} = \mathbf{Z}$$

ou seja:

$$\mathbf{Z} = \mathbf{X}\mathbf{U}'(\mathbf{U}\mathbf{U}')^{-1} \quad (12.31)$$

que é uma matriz $n \times k$ de posto coluna completo, porque

$$\text{posto}(\mathbf{Z}) \geq \text{posto}(\mathbf{Z}\mathbf{U}) = \text{posto}(\mathbf{X}) = k$$

⇒ Conhecendo as matrizes $\mathbf{X}$ e $\mathbf{U}$ que definem os novos parâmetros $\boldsymbol{\gamma} = \mathbf{U}\boldsymbol{\beta}$ obtemos a nova matriz $\mathbf{Z}$ utilizando (12.31).

Como o modelo $\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$ é de posto completo, o sistema de equações normais, $\mathbf{Z}'\mathbf{Z}\hat{\boldsymbol{\gamma}} = \mathbf{Z}'\mathbf{y}$, tem solução única, dada por:

$$\hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$$

O estimador não viesado de σ^2 no modelo reparametrizado é dado por:

$$s^2 = \frac{1}{n-k} (\mathbf{y} - \mathbf{Z}\hat{\boldsymbol{\gamma}})'(\mathbf{y} - \mathbf{Z}\hat{\boldsymbol{\gamma}}) = \frac{SQResiduo}{n-k} \quad (12.32)$$

Observe que:

- Desde que $\mathbf{Z}\boldsymbol{\gamma} = \mathbf{X}\boldsymbol{\beta} \Rightarrow \mathbf{Z}\hat{\boldsymbol{\gamma}} = \mathbf{X}\hat{\boldsymbol{\beta}} \Rightarrow SQResiduo$ definida em (12.19) e em (12.32) também são iguais, ou seja:

$$(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = (\mathbf{y} - \mathbf{Z}\hat{\boldsymbol{\gamma}})'(\mathbf{y} - \mathbf{Z}\hat{\boldsymbol{\gamma}}) \quad (12.33)$$

- O conjunto de novos parâmetros, $\boldsymbol{\gamma} = \mathbf{U}\boldsymbol{\beta}$, é somente um conjunto de funções estimáveis e l. i.
- Podemos usar outros conjuntos de k funções estimáveis e l. i. de $\boldsymbol{\beta}$ em outras reparametrizações e a igualdade (12.33) continua válida.

Exemplo 12.4. Nós ilustraremos uma reparametrização para o modelo $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, $i = 1, 2$ e $j = 1, 2$. Na sua forma matricial este modelo pode ser escrito como $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, em que:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \end{bmatrix}$$

Como $\text{posto}(\mathbf{X}) = 2$, precisamos de duas funções estimáveis e linearmente independentes para realizar a reparametrização, como $\gamma_1 = \mu + \tau_1$ e $\gamma_2 = \mu + \tau_2$. Assim:

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \mu + \tau_1 \\ \mu + \tau_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix} = \mathbf{U}\boldsymbol{\beta} \Rightarrow \mathbf{U} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Apesar de ser fácil (neste caso!) obter a matriz $\mathbf{Z}$ do modelo reparametrizado, vamos obtê-la usando $\mathbf{Z} = \mathbf{XU}'(\mathbf{UU}')^{-1}$.

$$\mathbf{UU}' = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow (\mathbf{UU}')^{-1} = \left(\frac{1}{3}\right) \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\mathbf{U}'(\mathbf{UU}')^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \left(\frac{1}{3}\right) \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \left(\frac{1}{3}\right) \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\mathbf{Z} = \mathbf{XU}'(\mathbf{UU}')^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \left(\frac{1}{3}\right) \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

de tal modo que:

$$\mathbf{Z}\boldsymbol{\gamma} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_1 \\ \gamma_1 \\ \gamma_2 \\ \gamma_2 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \mu + \tau_1 \\ \mu + \tau_1 \\ \mu + \tau_1 \\ \mu + \tau_2 \\ \mu + \tau_2 \\ \mu + \tau_2 \end{bmatrix} = \mathbf{X}\boldsymbol{\beta}$$

Obs: Verifique que $\mathbf{ZU} = \mathbf{X}$.

Exercício: Escreva o modelo reparametrizado para os novos parâmetros $\gamma_1 = 2\mu + \tau_1 + \tau_2$ e $\gamma_2 = \tau_1 - \tau_2$

IMPORTANTE: Ver material adicional sobre reparametrização em modelos de posto incompleto, disponível no Moodle.

12.5. CONDIÇÕES MARGINAIS

Condições marginais fornecem restrições (lineares) que tornam únicos e individualmente estimáveis os parâmetros de um modelo linear.

Outro uso para as condições marginais: impor restrições sobre as estimativas de β somente para simplificar a resolução do sistema de equações normais, $\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}$.

Situação: $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n \Rightarrow \mathbf{X}'\mathbf{X}\beta$ representa um conjunto de k funções estimáveis de β .

Já sabemos: As condições marginais devem ser funções não estimáveis de β .

Justificativa: Se uma condição marginal for uma função estimável de β , ela pode ser expressa como uma combinação linear das linhas de $\mathbf{X}'\mathbf{X}\beta$ e não contribuirá para anular a deficiência de posto de $\mathbf{X}$, nem para obter um vetor solução $\hat{\beta}$ do sistema $\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}$.

Se a deficiência no posto de $\mathbf{X}$ é $(p - k)$, devemos impor $(p - k)$ condições marginais que completem a deficiência de posto de $\mathbf{X}$.

Definiremos as condições marginais como:

$$\mathbf{T}\beta = \mathbf{0} \text{ (nos parâmetros do modelo)}$$

ou

$$\mathbf{T}\hat{\beta} = \mathbf{0} \text{ (nas estimativas } \hat{\beta}, \text{ que são soluções do S.E.N.)}$$

onde $\mathbf{T}$ é uma matriz $(p - k) \times p$ de posto $p - k$, tal que $\mathbf{T}\beta$ é um conjunto de funções não estimáveis e *l. i.* dos parâmetros em β .

Teorema 12.5A. Se $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$ e se $\mathbf{T}$ é uma matriz $(p-k) \times p$, de posto $p-k$, tal que $\mathbf{T}\boldsymbol{\beta}$ é um conjunto de funções não estimáveis de $\boldsymbol{\beta}$, então existe um vetor único $\hat{\boldsymbol{\beta}}$ que satisfaz, simultaneamente, $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ e $\mathbf{T}\hat{\boldsymbol{\beta}} = \mathbf{0}$.

Prova: O modelo $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ e o conjunto de condições marginais $\mathbf{0} = \mathbf{T}\boldsymbol{\beta} + \mathbf{0}$ podem ser combinados como:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{X}\boldsymbol{\beta} \\ \mathbf{T}\boldsymbol{\beta} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{X} \\ \mathbf{T} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{\varepsilon} \\ \mathbf{0} \end{bmatrix} \quad (12.35)$$

Como as linhas de $\mathbf{T}$ são *l. i.* e não são combinações das linhas de $\mathbf{X}$, a matriz $\begin{bmatrix} \mathbf{X} \\ \mathbf{T} \end{bmatrix}$ é $(n+p-k) \times p$ e tem posto p . O novo sistema de equações normais fica:

$$\begin{bmatrix} \mathbf{X}' \\ \mathbf{T}' \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{T} \end{bmatrix} \hat{\boldsymbol{\beta}} = \begin{bmatrix} \mathbf{X}' \\ \mathbf{T}' \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} \quad (12.36)$$

Ou

$$(\mathbf{X}'\mathbf{X} + \mathbf{T}'\mathbf{T})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} + \mathbf{T}'\mathbf{0} = \mathbf{X}'\mathbf{y}$$

e tem solução única:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X} + \mathbf{T}'\mathbf{T})^{-1}\mathbf{X}'\mathbf{y} \quad (12.37)$$

Note que $[\mathbf{X}' \ \mathbf{T}'] \begin{bmatrix} \mathbf{X} \\ \mathbf{T} \end{bmatrix} = \mathbf{X}'\mathbf{X} + \mathbf{T}'\mathbf{T}$ é uma matriz quadrada $p \times p$ e de posto completo.

A solução $\hat{\boldsymbol{\beta}}$ em (12.37) também satisfaz o sistema de equações normais $(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$, desde que, por (12.36):

$$(\mathbf{X}'\mathbf{X} + \mathbf{T}'\mathbf{T})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} + \mathbf{T}'\mathbf{0} \Rightarrow \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{T}'\mathbf{T}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} \quad (12.38)$$

Como assumimos $\mathbf{T}\hat{\boldsymbol{\beta}} = \mathbf{0} \Rightarrow$ (12.38) reduz-se a $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$.

Exemplo 12.5. Consideremos o modelo $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, $i = 1, 2, j = 1, 2$, como no Exemplo 12.4. No Problema 12.5(b) pode ser mostrado que $\tau_1 + \tau_2$ é uma função não é estimável.

Impondo a condição marginal $\mathbf{T}\boldsymbol{\beta} = [0 \ 1 \ 1]\boldsymbol{\beta} = \tau_1 + \tau_2 = 0$, em que $\mathbf{T} = [0 \ 1 \ 1]$, temos:

$$\mathbf{X}'\mathbf{X} + \mathbf{T}'\mathbf{T} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} [0 \ 1 \ 1] = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

Então

$$(\mathbf{X}'\mathbf{X} + \mathbf{T}'\mathbf{T})^{-1} = \frac{1}{4} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

Como $\mathbf{X}'\mathbf{y} = [y_{..} \ y_{1.} \ y_{2.}]'$ e utilizando (12.37), obtemos:

$$\begin{aligned}
 \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X} + \mathbf{T}'\mathbf{T})^{-1}\mathbf{X}'\mathbf{y} = \frac{1}{4} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{bmatrix} \\
 &= \frac{1}{4} \begin{bmatrix} 2y_{..} - y_{1.} - y_{2.} \\ 2y_{1.} - y_{..} \\ 2y_{2.} - y_{..} \end{bmatrix} = \begin{bmatrix} \bar{y}_{..} \\ \bar{y}_{1.} - \bar{y}_{..} \\ \bar{y}_{2.} - \bar{y}_{..} \end{bmatrix} \quad (12.39)
 \end{aligned}$$

Porque: $y_{1.} + y_{2.} = y_{..}$, $\bar{y}_{i.} = y_{i.}/2$ e $\bar{y}_{..} = y_{..}/4$.

Para mostrar que $\hat{\boldsymbol{\beta}}$ em (12.39) também é solução do sistema de equações normais $(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$, basta verificar que:

$$\begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} \bar{y}_{..} \\ \bar{y}_{1.} - \bar{y}_{..} \\ \bar{y}_{2.} - \bar{y}_{..} \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{bmatrix}$$

Exercício: Considere o modelo $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, $i = 1, 2, j = 1, 2, 3$, como no Exemplo 12.4 e os dados:

Aditivo 1	Aditivo 2
14	18
16	19
15	17

Obtenha estimativas de $\boldsymbol{\beta}$ impondo as seguintes condições marginais:

- a) $\tau_1 = 0$ (Condição imposta pelo R)
- b) $\tau_2 = 0$ (Condição imposta pelo SAS)
- c) $\tau_1 + \tau_2 = 0$ (Condição + comum da Estatística Experimental)

12.6. TESTANDO HIPÓTESES

12.6.1. Hipóteses testáveis

Definição: A hipótese $H_0: \beta_1 = \beta_2 = \dots = \beta_q$ é testável se existe um conjunto de funções estimáveis linearmente independentes $\{\lambda'_1\beta, \lambda'_2\beta, \dots, \lambda'_{q-1}\beta\}$ tal que H_0 é verdadeira se e somente se

$$\lambda'_1\beta = \lambda'_2\beta = \dots = \lambda'_{q-1}\beta = 0$$

Ou

$H_0: \beta_1 = \beta_2 = \dots = \beta_q$ é testável se encontrarmos um conjunto de funções estimáveis e *l.i.* $\{\lambda'_1\beta, \lambda'_2\beta, \dots, \lambda'_{q-1}\beta\}$ que quando iguadas a zero, reproduzem a hipótese H_0 .

Exemplo: Para testar $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4$ precisamos encontrar um conjunto de $4 - 1 = 3$ funções estimáveis e *l. i.* tais que $\lambda'_1 \boldsymbol{\beta} = \lambda'_2 \boldsymbol{\beta} = \lambda'_3 \boldsymbol{\beta} = 0$ reproduzam H_0 .

Sugestão: Usar contrastes (funções estimáveis e *l. i.*) do tipo:

$$\lambda'_1 \boldsymbol{\beta} = 3\beta_1 - (\beta_2 + \beta_3 + \beta_4)$$

$$\lambda'_2 \boldsymbol{\beta} = 2\beta_2 - (\beta_3 + \beta_4)$$

$$\lambda'_3 \boldsymbol{\beta} = \beta_3 - \beta_4$$

Igualando $\lambda'_3 \boldsymbol{\beta} = \lambda'_2 \boldsymbol{\beta} = \lambda'_1 \boldsymbol{\beta} = 0$ reproduzimos H_0 (verifique!)

Caso geral: Para testar $H_0: \beta_1 = \beta_2 = \cdots = \beta_q$ vamos usar $q-1$ contrastes que formam um conjunto de funções estimáveis e *l. i.*, tais que:

$$\begin{bmatrix} \lambda'_1 \boldsymbol{\beta} \\ \lambda'_2 \boldsymbol{\beta} \\ \vdots \\ \lambda'_{q-1} \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} q-1 & -1 & -1 & \cdots & -1 & -1 \\ 0 & q-2 & -1 & \cdots & -1 & -1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_{q-1} \\ \beta_q \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

se e somente se $\beta_1 = \beta_2 = \cdots = \beta_q$.

Para ilustrar hipóteses testáveis, seja o modelo aditivo com dois fatores

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \text{ para } i = 1, 2, 3 \text{ e } j = 1, 2, 3$$

e a hipótese de interesse $H_0: \alpha_1 = \alpha_2 = \alpha_3$. Já sabemos que todas as linhas de $\mathbf{X}\boldsymbol{\beta}$ são estimáveis:

Linha	Função estimável
1	$\mu + \alpha_1 + \beta_1$
2	$\mu + \alpha_2 + \beta_1$
3	$\mu + \alpha_3 + \beta_1$
4	$\mu + \alpha_1 + \beta_2$
5	$\mu + \alpha_2 + \beta_2$
6	$\mu + \alpha_3 + \beta_2$
7	$\mu + \alpha_1 + \beta_3$
8	$\mu + \alpha_2 + \beta_3$
9	$\mu + \alpha_3 + \beta_3$

- Combinações lineares dessas linhas também são estimáveis:

(1) Linha 1 - Linha 2 = $\alpha_1 - \alpha_2$

(2) Linha 1 + Linha 2 - 2*Linha3 = $\alpha_1 + \alpha_2 - 2\alpha_3$

Igualando as duas funções estimáveis e *l. i.* a zero temos:

$$(1) \alpha_1 - \alpha_2 = 0 \Rightarrow \alpha_1 = \alpha_2$$

$$(2) \alpha_1 + \alpha_2 - 2\alpha_3 = 0 \Rightarrow \alpha_2 + \alpha_2 - 2\alpha_3 = 0 \Rightarrow 2\alpha_2 = 2\alpha_3$$

$$\Rightarrow \alpha_2 = \alpha_3 \Rightarrow \alpha_1 = \alpha_2 = \alpha_3 \text{ e reproduzimos } H_0$$

$\therefore H_0: \alpha_1 = \alpha_2 = \alpha_3$ é uma hipótese testável e testar essa hipótese é equivalente a testar:

$$H_0: \begin{bmatrix} \alpha_1 - \alpha_2 \\ \alpha_1 + \alpha_2 - 2\alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (12.40)$$

Dúvida: Como testar $H_0: \alpha_1 = \alpha_2 = \alpha_3$?

Usar a abordagem do Modelo completo *versus* Modelo reduzido?

Usar a abordagem da hipótese linear geral?

12.6.2. Modelo Completo e Modelo Reduzido

Problema: Testar a hipótese $H_0: \beta_1 = \beta_2 = \dots = \beta_q$ no modelo de posto incompleto $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$ e $\boldsymbol{\beta}$ é $p \times 1$.

Argumento: Se H_0 é testável nós podemos encontrar um conjunto de t funções estimáveis e *l. i.* $\{\lambda'_1\boldsymbol{\beta}, \lambda'_2\boldsymbol{\beta}, \dots, \lambda'_t\boldsymbol{\beta}\}$ tais que $H_0: \beta_1 = \beta_2 = \dots = \beta_q$ seja equivalente a:

$$H_0: \boldsymbol{\gamma}_1 = \begin{bmatrix} \lambda'_1\boldsymbol{\beta} \\ \lambda'_2\boldsymbol{\beta} \\ \vdots \\ \lambda'_t\boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Também é possível encontrar outro conjunto de $(k - t)$ funções:

$$\boldsymbol{\gamma}_2 = \begin{bmatrix} \lambda'_{t+1} \boldsymbol{\beta} \\ \lambda'_{t+2} \boldsymbol{\beta} \\ \vdots \\ \lambda'_k \boldsymbol{\beta} \end{bmatrix}$$

de tal forma que todas as k funções:

$$\underbrace{\lambda'_1 \boldsymbol{\beta}, \lambda'_2 \boldsymbol{\beta}, \dots, \lambda'_t \boldsymbol{\beta}} \quad \underbrace{\lambda'_{t+1} \boldsymbol{\beta}, \dots, \lambda'_k \boldsymbol{\beta}}$$

sejam estimáveis e *l. i.*, onde $k = \text{posto}(\mathbf{X})$.

Seja $\boldsymbol{\gamma} = \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix}$ então podemos reparametrizar o modelo de posto incompleto $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, obtendo o modelo de posto completo:

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} = \mathbf{Z}_1\boldsymbol{\gamma}_1 + \mathbf{Z}_2\boldsymbol{\gamma}_2 + \boldsymbol{\varepsilon}$$

em que $\mathbf{Z} = [\mathbf{Z}_1 \ \mathbf{Z}_2]$ é particionada conforme o número de elementos em $\boldsymbol{\gamma}_1$ e $\boldsymbol{\gamma}_2$.

Na abordagem modelo completo *versus* modelo reduzido, usamos $\mathbf{y} = \mathbf{Z}_1\boldsymbol{\gamma}_1 + \mathbf{Z}_2\boldsymbol{\gamma}_2 + \boldsymbol{\varepsilon}$ como o modelo completo.

Para testar $H_0: \boldsymbol{\gamma}_1 = \mathbf{0}$ usamos o modelo reduzido $\mathbf{y} = \mathbf{Z}_2\boldsymbol{\gamma}_2^* + \boldsymbol{\varepsilon}^*$

Pelo Teorema 7.10A, a estimativa de $\boldsymbol{\gamma}_2^*$ no modelo reduzido é a mesma estimativa de $\boldsymbol{\gamma}_2$ no modelo completo desde que as colunas de $\mathbf{Z}_2$ sejam ortogonais às colunas de $\mathbf{Z}_1$, isto é, se $\mathbf{Z}_2'\mathbf{Z}_1 = \mathbf{0}$.

Como para modelos balanceados esta ortogonalidade geralmente se verifica, usaremos $\boldsymbol{\gamma}_2$ e $\hat{\boldsymbol{\gamma}}_2$, ao invés de $\boldsymbol{\gamma}_2^*$ e $\hat{\boldsymbol{\gamma}}_2^*$.

- Desde que $\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$ é um modelo de posto completo, a hipótese $H_0: \boldsymbol{\gamma}_1 = \mathbf{0}$ pode ser testada como na Seção 8.2 (Regressão linear múltipla)
- O número de graus de liberdade (t) associado a $SQ(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)$ é igual ao número de funções estimáveis e *l. i.* necessárias para expressar a hipótese H_0 .

Tabela 12.2 Análise de variância para testar $H_0: \boldsymbol{\gamma}_1 = \mathbf{0}$ no modelo reparametrizado balanceado

F. Variação	g.l.	Somas de quadrados
Devida a $\boldsymbol{\gamma}_1 \boldsymbol{\gamma}_2$	t	$SQ(\boldsymbol{\gamma}_1 \boldsymbol{\gamma}_2) = \hat{\boldsymbol{\gamma}}'\mathbf{Z}'\mathbf{y} - \hat{\boldsymbol{\gamma}}'_2\mathbf{Z}'_2\mathbf{y}$
Resíduo	$n-k$	$SQResiduo = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\gamma}}'\mathbf{Z}'\mathbf{y}$
Total	$n-1$	$SQT = \mathbf{y}'\mathbf{y} - n\bar{y}^2$

- Os resultados apresentados na Tabela 12.2 são análogos aos da Tabela 8.3 (modelo de regressão linear múltipla)

A soma de quadrados $\hat{\mathbf{y}}'\mathbf{Z}'\mathbf{y}$ é obtida no modelo completo $\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$ e a soma de quadrados $\hat{\mathbf{y}}'_2\mathbf{Z}'_2\mathbf{y}$, no modelo reduzido $\mathbf{y} = \mathbf{Z}_2\boldsymbol{\gamma}_2^* + \boldsymbol{\varepsilon}^*$, que assume que a hipótese $H_0: \boldsymbol{\gamma}_1 = \mathbf{0}$ é verdadeira.

Este procedimento envolvendo a reparametrização é interessante, mas na prática a obtenção da matriz $\mathbf{Z}$ pode consumir muito tempo! Veremos que este passo pode ser evitado.

Perceba que de (12.20) e (12.33), nós temos:

$$\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} = \mathbf{y}'\mathbf{y} - \hat{\mathbf{y}}'\mathbf{Z}'\mathbf{y}$$

de onde obtemos:

$$\hat{\beta}' \mathbf{X}' \mathbf{y} = \hat{\gamma}' \mathbf{Z}' \mathbf{y} \quad (12.41)$$

onde $\hat{\beta}$ representa **qualquer** solução do sistema de equações normais $(\mathbf{X}'\mathbf{X})\hat{\beta} = \mathbf{X}'\mathbf{y}$.

De modo similar, correspondente ao modelo reduzido $\mathbf{y} = \mathbf{Z}_2 \boldsymbol{\gamma}_2^* + \boldsymbol{\varepsilon}^*$, temos o modelo reduzido $\mathbf{y} = \mathbf{X}_2 \boldsymbol{\beta}_2^* + \boldsymbol{\varepsilon}^*$, obtido quando assumimos que $\beta_1 = \beta_2 = \dots = \beta_q$.

Então,

$$\hat{\beta}_2^{*'} \mathbf{X}_2' \mathbf{y} = \hat{\gamma}_2^{*'} \mathbf{Z}_2' \mathbf{y} \quad (12.42)$$

onde $\hat{\beta}_2^*$ é qualquer solução do sistema $(\mathbf{X}_2' \mathbf{X}_2) \hat{\beta}_2^* = \mathbf{X}_2' \mathbf{y}$.

Note que: Para encontrar $\hat{\beta}$ e $\hat{\beta}_2^*$ em modelos balanceados também podemos usar condições marginais (ou restrições nos parâmetros).

Teorema 12.6A. Considere o modelo particionado $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$, onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$. Se $\mathbf{X}_2'\mathbf{X}_1 = \mathbf{0}$ (Seção 12.7.3), a estimativa de $\boldsymbol{\beta}_2^*$ no modelo reduzido $\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2^* + \boldsymbol{\varepsilon}^*$ é a mesma estimativa de $\boldsymbol{\beta}_2$ no modelo completo $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$.

Obs. Como no modelo balanceado de posto incompleto, a ortogonalidade de $\mathbf{X}_1$ e $\mathbf{X}_2$ geralmente se verifica $\Rightarrow$ usaremos $\boldsymbol{\beta}_2$ e $\hat{\boldsymbol{\beta}}_2$, ao invés de $\boldsymbol{\beta}_2^*$ e $\hat{\boldsymbol{\beta}}_2^*$.

Na Tabela 12.3, $\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$ e $\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{y}$ são as SQ 's obtidas nos modelos completo ($\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$) e reduzido ($\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$) pela hipótese $H_0: \beta_1 = \beta_2 = \cdots = \beta_q$.

O número de graus de liberdade, t , para $SQ(\boldsymbol{\beta}_1|\boldsymbol{\beta}_2)$ é o mesmo de $SQ(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)$ da Tabela 12.2.

Tabela 12.3 Análise de variância para testar $H_0: \beta_1 = \beta_2 = \dots = \beta_q$ no modelo balanceado de posto incompleto.

F. Variação	g.l.	Somas de quadrados
Devida a $\beta_1 \beta_2$	t	$SQ(\beta_1 \beta_2) = \hat{\beta}' \mathbf{X}' \mathbf{y} - \hat{\beta}'_2 \mathbf{X}'_2 \mathbf{y}$
Resíduo	$n-k$	$SQ_{Residuo} = \mathbf{y}' \mathbf{y} - \hat{\beta}' \mathbf{X}' \mathbf{y}$
Total	$n-1$	$SQT = \mathbf{y}' \mathbf{y} - n\bar{y}^2$

Neste exemplo, t corresponde ao número de funções estimáveis e linearmente independentes usadas para expressar H_0 .

Para testar $H_0: \beta_1 = \beta_2 = \dots = \beta_q$ são necessárias $t = q - 1$ funções estimáveis e *l. i.*

12.6.3 Hipótese Linear Geral $H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$

Teorema 12.6B. Se $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$, se $\mathbf{C}$ é $m \times p$ de posto $m \leq k$ tal que $\mathbf{C}\boldsymbol{\beta}$ é um conjunto de m funções estimáveis e *l.i.* e se $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, então:

- i) $\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'$ é não singular e invariante a escolhas de $(\mathbf{X}'\mathbf{X})^{-1}$
- ii) $\mathbf{C}\hat{\boldsymbol{\beta}} \sim N_m(\mathbf{C}\boldsymbol{\beta}, \sigma^2\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')$
- iii) $SQHip/\sigma^2 = (\mathbf{C}\hat{\boldsymbol{\beta}})' [\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}']^{-1} (\mathbf{C}\hat{\boldsymbol{\beta}})/\sigma^2 \sim \chi^2(m, \lambda)$,
 Onde $\lambda = (\mathbf{C}\boldsymbol{\beta})' [\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}']^{-1} \mathbf{C}\boldsymbol{\beta} / 2\sigma^2$
- iv) $SQResiduo/\sigma^2 = \mathbf{y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y}/\sigma^2 \sim \chi^2(n - k)$.
- v) $SQHip$ e $SQResiduo$ são independentes.

Teorema 12.6C Seja $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, onde $\mathbf{X}$ é $n \times p$, de posto $k < p \leq n$, e sejam $\mathbf{C}$, $\mathbf{C}\boldsymbol{\beta}$ e $\widehat{\boldsymbol{\beta}}$ como definidos no Teorema 12.6B. Então se $H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ é verdadeira, a estatística

$$F = \frac{SQHip/m}{SQResiduo/(n-k)} = \frac{(\mathbf{C}\widehat{\boldsymbol{\beta}})' [\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}']^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}})/m}{SQResiduo/(n-k)} \quad (12.46)$$

Em que $SQHip = (\mathbf{C}\widehat{\boldsymbol{\beta}})' [\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}']^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}})$ e F é distribuída como $F(m, n-k)$.

12.7. UMA ILUSTRAÇÃO DE ESTIMAÇÃO E TESTE DE HIPÓTESE

Suponha o modelo *aditivo* (sem interação) com dois fatores:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \text{ para } i = 1, 2, 3 \text{ e } j = 1, 2$$

As seis observações podem ser escritas na forma $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ como

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{31} \\ \varepsilon_{32} \end{bmatrix} \quad (12.47)$$

A matriz $\mathbf{X}'\mathbf{X}$ é dada por: $\mathbf{X}'\mathbf{X} = \begin{bmatrix} 6 & 2 & 2 & 2 & 3 & 3 \\ 2 & 2 & 0 & 0 & 1 & 1 \\ 2 & 0 & 2 & 0 & 1 & 1 \\ 2 & 0 & 0 & 2 & 1 & 1 \\ 3 & 1 & 1 & 1 & 3 & 0 \\ 3 & 1 & 1 & 1 & 0 & 3 \end{bmatrix}$

$\mathbf{X}'\mathbf{X}$ é 6×6 e $\text{posto}(\mathbf{X}'\mathbf{X}) = \text{posto}(\mathbf{X}) = 4$.

12.7.1 Funções Estimáveis

A hipótese $H_{01}: \alpha_1 = \alpha_2 = \alpha_3$ é equivalente a

$$H_{01}: \alpha_1 - \alpha_2 = \alpha_1 - \alpha_3 = 0 \text{ ou } H_{01}: \begin{bmatrix} \alpha_1 - \alpha_2 \\ \alpha_1 - \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Assim H_{01} é testável porque $\alpha_1 - \alpha_2$ e $\alpha_1 - \alpha_3$ são funções estimáveis (contrastes) *l. i.* e quando igualadas a zero, reproduzem a hipótese $H_{01}: \alpha_1 = \alpha_2 = \alpha_3$.

12.7.2 Testando uma hipótese

Como são necessárias duas funções estimáveis e *l. i.* dos α 's para expressar $H_{01}: \alpha_1 = \alpha_2 = \alpha_3$, a SQH_{01} tem dois graus de liberdade.

De modo similar, $H_{02}: \beta_1 = \beta_2$ também é testável e SQH_{02} tem um grau de liberdade.

O sistema de equações normais $(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ é dado por:

$$\begin{bmatrix} 6 & 2 & 2 & 2 & 3 & 3 \\ 2 & 2 & 0 & 0 & 1 & 1 \\ 2 & 0 & 2 & 0 & 1 & 1 \\ 2 & 0 & 0 & 2 & 1 & 1 \\ 3 & 1 & 1 & 1 & 3 & 0 \\ 3 & 1 & 1 & 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \\ y_{.1} \\ y_{.2} \end{bmatrix} \quad (12.48)$$

Impondo duas condições marginais $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 = 0$ e $\hat{\beta}_1 + \hat{\beta}_2 = 0$, nós obtemos as seguintes soluções:

$$\hat{\mu} = \bar{y}_{..} \quad \hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} \quad \hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..} \quad (12.49)$$

onde $\bar{y}_{..} = \sum_{ij} y_{ij}/6$, $\bar{y}_{i.} = \sum_j y_{ij}/2$ e $\bar{y}_{.j} = \sum_i y_{ij}/3$, para $i = 1, 2, 3$ e $j = 1, 2$.

Importante:

- Se impusermos as condições sobre os parâmetros e sobre as estimativas as equações (12.49) fornecem estimativas únicas de parâmetros com significado único.

Por exemplo: α_1 passa a ser $\alpha_1^* = \bar{\mu}_{1\bullet} - \bar{\mu}_{\bullet\bullet}$ (desvio esperado da média devido ao tratamento 1) e $\bar{y}_{1\bullet} - \bar{y}_{\bullet\bullet}$ é uma boa estimativa de α_1 .

- Se as condições forem usadas somente para obter estimativas e não forem impostas sobre os parâmetros, então α_i não é único e $\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}$ não estima um parâmetro.

Neste caso $\hat{\alpha}_i = \bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}$ somente pode ser usado com outros elementos em $\hat{\beta}$ para obter estimativas $\lambda' \hat{\beta}$ de funções estimáveis $\lambda' \beta$.

Para testar $H_0: \alpha_1 = \alpha_2 = \alpha_3$ precisamos calcular a SQ do modelo completo:

$$SQ(\mu, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2) = SQ(\mu, \alpha, \beta) = \hat{\beta}' \mathbf{X}' \mathbf{y}$$

$$SQ(\mu, \alpha, \beta) = \begin{bmatrix} \hat{\mu} & \hat{\alpha}_1 & \hat{\alpha}_2 & \hat{\alpha}_3 & \hat{\beta}_1 & \hat{\beta}_2 \end{bmatrix} \begin{bmatrix} y_{\bullet\bullet} \\ y_{1\bullet} \\ y_{2\bullet} \\ y_{3\bullet} \\ y_{\bullet 1} \\ y_{\bullet 2} \end{bmatrix}$$

$$SQ(\mu, \alpha, \beta) = \frac{y_{\bullet\bullet}^2}{6} + \left(\sum_{i=1}^3 \frac{y_{i\bullet}^2}{2} - \frac{y_{\bullet\bullet}^2}{6} \right) + \left(\sum_{j=1}^2 \frac{y_{\bullet j}^2}{3} - \frac{y_{\bullet\bullet}^2}{6} \right) \quad (12.50)$$

$$\begin{aligned} SQRes &= \mathbf{y}' \mathbf{y} - \hat{\beta}' \mathbf{X}' \mathbf{y} \\ &= \left(\sum_{ij} y_{ij}^2 - \frac{y_{\bullet\bullet}^2}{6} \right) - \left(\sum_{i=1}^3 \frac{y_{i\bullet}^2}{2} - \frac{y_{\bullet\bullet}^2}{6} \right) - \left(\sum_{j=1}^2 \frac{y_{\bullet j}^2}{3} - \frac{y_{\bullet\bullet}^2}{6} \right) \end{aligned}$$

Para obter a SQ do modelo reduzido, $\hat{\beta}'_2 \mathbf{X}'_2 \mathbf{y}$, da Tabela 12.3, nós usamos o modelo:

$$y_{ij} = (\mu + \alpha) + \beta_j + \varepsilon_{ij} = \mu + \beta_j + \varepsilon_{ij}$$

em que $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$ e a constante $(\mu + \alpha)$ é substituída por μ (sem perda de generalidade).

O sistema de equações normais fica:

$$\mathbf{X}'_2 \mathbf{X}_2 \hat{\beta}_2 = \mathbf{X}'_2 \mathbf{y} \quad \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{.1} \\ y_{.2} \end{bmatrix} \quad (12.51)$$

Usando a condição $\hat{\beta}_1 + \hat{\beta}_2 = 0$, a solução (única) para o sistema é obtida como:

$$\hat{\mu} = \bar{y}_{..} \quad \hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..} \quad (12.52)$$

Observe que essa solução é a mesma obtida no modelo completo em (12.49).

A soma de quadrados do modelo reduzido é calculada por

$$SQ(\mu, \beta) = \hat{\beta}'_2 \mathbf{X}'_2 \mathbf{y} = \frac{y_{..}^2}{6} + \left(\sum_{j=1}^2 \frac{y_{.j}^2}{3} - \frac{y_{..}^2}{6} \right) \quad (12.53)$$

A soma de quadrados associada à hipótese $H_{01}: \alpha_1 = \alpha_2 = \alpha_3$ é definida como:

$$SQ(\alpha|\mu, \beta) = \hat{\beta}' \mathbf{X}' \mathbf{y} - \hat{\beta}'_2 \mathbf{X}'_2 \mathbf{y} = \sum_{i=1}^3 \frac{y_{i.}^2}{2} - \frac{y_{..}^2}{6} \quad (12.54)$$

Esses resultados estão resumidos na Tabela 12.4, observando que não foi incluída a soma de quadrados $SQ(\beta|\mu, \alpha)$.

Tabela 12.4 Análise de variância para testar $H_{01}: \alpha_1 = \alpha_2 = \alpha_3$

Fonte de Variação	g.l.	Somas de quadrados
Devida a α ajust. μ, β	2	$SQ(\alpha \mu, \beta) = \sum_{i=1}^3 \frac{y_{i\cdot}^2}{2} - \frac{y_{\cdot\cdot}^2}{6}$
Resíduo	2	$SQRes = \sum_{ij} y_{ij}^2 - \hat{\beta}'\mathbf{X}'\mathbf{y}$
Total	5	$SQT = \sum_{ij} y_{ij}^2 - \frac{y_{\cdot\cdot}^2}{6}$

Note que $SQ(\alpha|\mu, \beta)$ é a mesma fórmula algébrica usada na Estatística Experimental para calcular $SQ(\alpha)$.

12.7.3 Ortogonalidade das Colunas de $\mathbf{X}$

- As estimativas de μ , β_1 e β_2 dadas em (12.52) no modelo reduzido são as mesmas de μ , β_1 e β_2 em (12.49) no modelo completo.
- A soma de quadrados $\hat{\boldsymbol{\beta}}_2' \mathbf{X}_2' \mathbf{y}$ em (12.53) é claramente uma parte de $\hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{y}$ em (12.50). De fato,

$$SQ(\alpha|\mu, \beta) = SQ(\alpha)$$

$$SQ(\mu, \alpha, \beta) = SQ(\mu) + SQ(\alpha) + SQ(\beta)$$

Esses resultados simplificados são devidos à ortogonalidade dos blocos de colunas na matriz $\mathbf{X}$ em (12.47), como comentado no Teorema 12.6A.

As colunas de $\mathbf{X}$ em (12.47) não são ortogonais a todas as outras colunas, mas nós podemos identificar três grupos de colunas ortogonais se cada coluna depois da primeira for centrada na média da coluna.

$$\mathbf{X}_c = \left(\mathbf{I} - \frac{1}{n} \mathbf{J} \right) \mathbf{X}_1 = \begin{bmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \dots & x_{1k} - \bar{x}_k \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \dots & x_{2k} - \bar{x}_k \\ \vdots & \vdots & & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \dots & x_{nk} - \bar{x}_k \end{bmatrix}$$

onde $\mathbf{X}_1$ é a matriz formada pelas colunas de $\mathbf{X}$, com exceção da primeira.

No exemplo em questão (modelo aditivo com dois fatores) temos

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Matriz centrada:

$$[\mathbf{j} \ \mathbf{X}_c] = [\mathbf{j} \ \mathbf{X}_\alpha \ \mathbf{X}_\beta] = \begin{bmatrix} 1 : & 2/3 & -1/3 & -1/3 : & 1/2 & -1/2 \\ 1 : & 2/3 & -1/3 & -1/3 : & -1/2 & 1/2 \\ 1 : & -1/3 & 2/3 & -1/3 : & 1/2 & -1/2 \\ 1 : & -1/3 & 2/3 & -1/3 : & -1/2 & 1/2 \\ 1 : & -1/3 & -1/3 & 2/3 : & 1/2 & -1/2 \\ 1 : & -1/3 & -1/3 & 2/3 : & -1/2 & 1/2 \end{bmatrix}$$

(12.55)

Os três grupos de colunas $[\mathbf{j} \ \mathbf{X}_\alpha \ \mathbf{X}_\beta]$ em (12.55) são ortogonais. Por exemplo: cada uma das colunas 2, 3 e 4 é ortogonal a cada uma das colunas 5 e 6, mas as colunas 2, 3 e 4 não são ortogonais entre si.

Vale lembrar que $\text{posto}(\mathbf{X}) = 4$ e que os estimadores dos parâmetros $\boldsymbol{\beta}_1$ no modelo centrado $\mathbf{y} = [\mathbf{j} \ \mathbf{X}_c] \begin{bmatrix} \alpha \\ \boldsymbol{\beta}_1 \end{bmatrix} + \boldsymbol{\varepsilon}$ são os mesmos estimadores de $\boldsymbol{\beta}_1$ do modelo não centrado, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$.

Ilustraremos o uso de condições marginais para obter uma ortogonalização que é de posto completo. Para tanto, considere o modelo com dois fatores e com interação:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad (12.56)$$

para $i = 1, 2, j = 1, 2$ e $k = 1, 2$ (duas repetições). Matricialmente o modelo fica:

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{21} \\ \gamma_{22} \end{bmatrix} + \begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{221} \\ \varepsilon_{222} \end{bmatrix} \quad (12.57)$$

Condições marginais úteis ficam evidentes no sistema de equações normais (12.58):

$$8\hat{\mu} + 4(\hat{\alpha}_1 + \hat{\alpha}_2) + 4(\hat{\beta}_1 + \hat{\beta}_2) + 2(\hat{\gamma}_{11} + \hat{\gamma}_{12} + \hat{\gamma}_{21} + \hat{\gamma}_{22}) = y_{\bullet\bullet}$$

$$4\hat{\mu} + 4\hat{\alpha}_i + 2(\hat{\beta}_1 + \hat{\beta}_2) + 2(\hat{\gamma}_{i1} + \hat{\gamma}_{i2}) = y_{i\bullet} \quad i = 1, 2$$

$$4\hat{\mu} + 2(\hat{\alpha}_1 + \hat{\alpha}_2) + 4\hat{\beta}_j + 2(\hat{\gamma}_{1j} + \hat{\gamma}_{2j}) = y_{\bullet j} \quad j = 1, 2$$

$$2\hat{\mu} + 2\hat{\alpha}_i + 2\hat{\beta}_j + 2\hat{\gamma}_{ij} = y_{ij} \quad i = 1, 2, \quad j = 1, 2$$

A solução das equações em (12.58) será simplificada impondo as seis condições marginais:

$$\begin{aligned} \hat{\alpha}_1 + \hat{\alpha}_2 &= 0, \quad \hat{\beta}_1 + \hat{\beta}_2 = 0, \\ \hat{\gamma}_{i1} + \hat{\gamma}_{i2} &= 0 \quad \text{para } i = 1, 2 \\ \hat{\gamma}_{1j} + \hat{\gamma}_{2j} &= 0 \quad \text{para } j = 1, 2 \end{aligned} \tag{12.59}$$

- A matriz $\mathbf{X}$ é 8×9 e $\text{posto}(\mathbf{X}) = 4$, porque as cinco primeiras colunas podem ser expressas como combinações lineares das quatro últimas colunas, que são *l. i.*
- Assim, $\mathbf{X}'\mathbf{X}$ é 9×9 e tem uma deficiência de $\text{posto } 9 - 4 = 5$. Entretanto, existem seis condições marginais em (12.59).
- Esta discrepância é resolvida notando-se que entre as quatro últimas equações em (12.59) existem somente três restrições *l. i.* (uma das quatro equações é redundante).

Podemos obter uma ortogonalização de posto completo impondo as condições marginais apresentadas em (12.59) sobre os parâmetros do modelo e usando essas relações para expressar os parâmetros redundantes em termos dos quatro parâmetros μ , α_1 , β_1 e γ_{11} , pois:

$$\alpha_2 = -\alpha_1, \quad \beta_2 = -\beta_1, \quad (12.60)$$

$$\gamma_{12} = -\gamma_{11}, \gamma_{21} = -\gamma_{11} \quad \text{e} \quad \gamma_{22} = \gamma_{11}$$

Usando (12.60), nós podemos expressar as oito observações y_{ijk} em (12.56) em termos de μ, α_1, β_1 e γ_{11} .

Para $k = \{1, 2\}$ tem-se:

$$y_{11k} = \mu + \alpha_1 + \beta_1 + \gamma_{11} + \varepsilon_{11k}$$

$$y_{12k} = \mu + \alpha_1 + \beta_2 + \gamma_{12} + \varepsilon_{12k} = \mu + \alpha_1 - \beta_1 - \gamma_{11} + \varepsilon_{12k}$$

$$y_{21k} = \mu + \alpha_2 + \beta_1 + \gamma_{21} + \varepsilon_{21k} = \mu - \alpha_1 + \beta_1 - \gamma_{11} + \varepsilon_{21k}$$

$$y_{22k} = \mu + \alpha_2 + \beta_2 + \gamma_{22} + \varepsilon_{22k} = \mu - \alpha_1 - \beta_1 + \gamma_{11} + \varepsilon_{21k}$$

A “nova” matriz $\mathbf{X}$ fica:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Que é uma matriz de posto completo com colunas ortogonais $\Rightarrow$ os métodos apresentados nos Capítulos 7 e 8 podem ser usados na estimação dos parâmetros e em testes de hipóteses.

Exercícios: ver pág. 336-338 do livro do Rencher.