MAT-468: Modelos de regresión para datos agrupados

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Bibliografía básica

Davidian, M., and Giltinan, D. M. (1995).
Nonlinear Models for Repeated Measurement Data.
Chapman & Hall, London.

Pinheiro, J. C., and Bates, D. M. (2000).
Mixed-Effects Models in S and S-Plus.
Springer, New York.

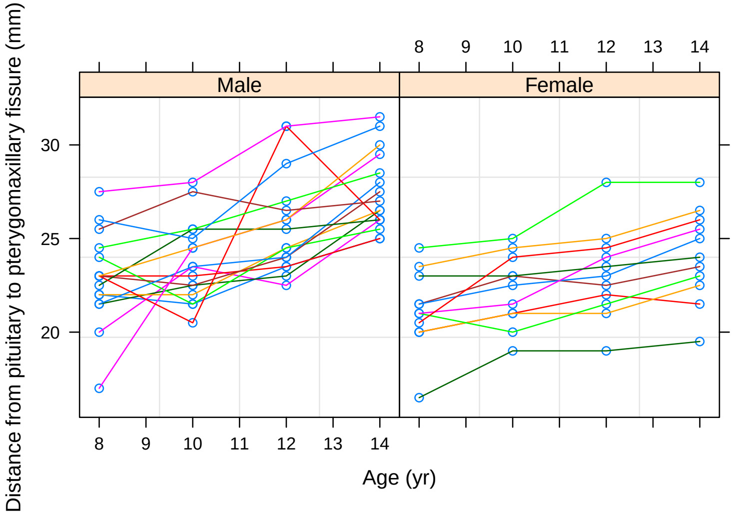
Vonesh, E. F., and Chinchilli, V. M. (1997).
Linear and Nonlinear Models for the Analysis of Repeated Measurements.
Marcel Dekker, New York.

Definición 1 (medidas repetidas):

Datos repetidos surgen cuando una respuesta es medida repetidamente en un conjunto de unidades.

- Unidades.
  - Individuos, pacientes,
  - Animales, plantas,
  - Familias, cuidades, industrias, ...
- Caso especial: **Datos longitudinales**.

Estudio de Ortodoncia (Potthoff y Roy, 1964)



Estudio de Ortodoncia (Potthoff y Roy, 1964)

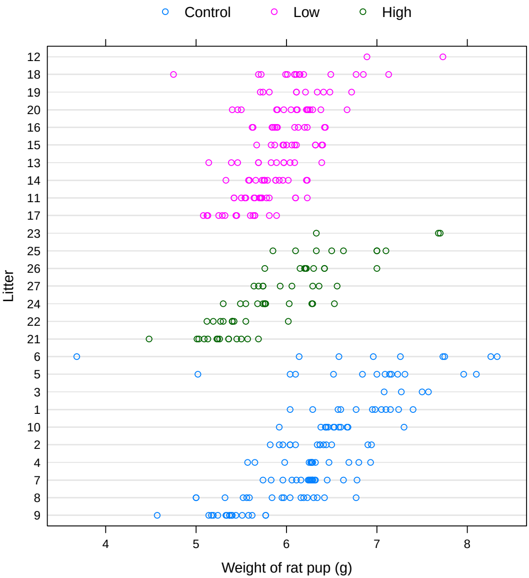
Objetivo:

- Investigar el patrón de crecimiento dental.
- Estudiar diferencias significativas entre ambos grupos.

Observaciones:

- Bastante variabilidad entre individuos con menor variabilidad dentro de individuos (especialmente en grupo femenino).
- Igual número de mediciones por individuo **(datos balanceados)**.
- Mediciones tomadas en instantes fijos.
- Algunos datos con crecimiento atípico.

Estudio reproductivo en roedores (Dempster et al., 1984)



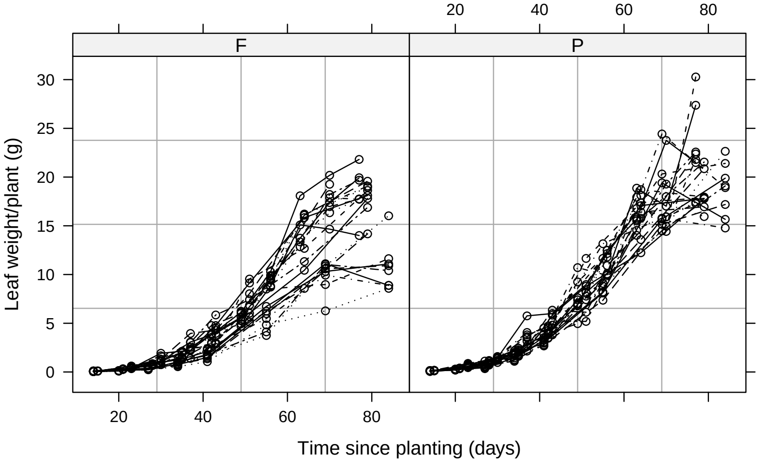
Estudio reproductivo en roedores (Dempster et al., 1984)

Observaciones:

- Datos altamente desbalanceados.
- Comportamiento similar por tratamiento.
- Gran variabilidad dentro de algunos grupos (nidadas).
- Algunas crías con peso "atípico" (Rosa, Padovani y Gianola, 2003).

EX UMBRA IN SOLEM

Crecimiento de plantas de Soya (Davidian y Giltinan, 1995)



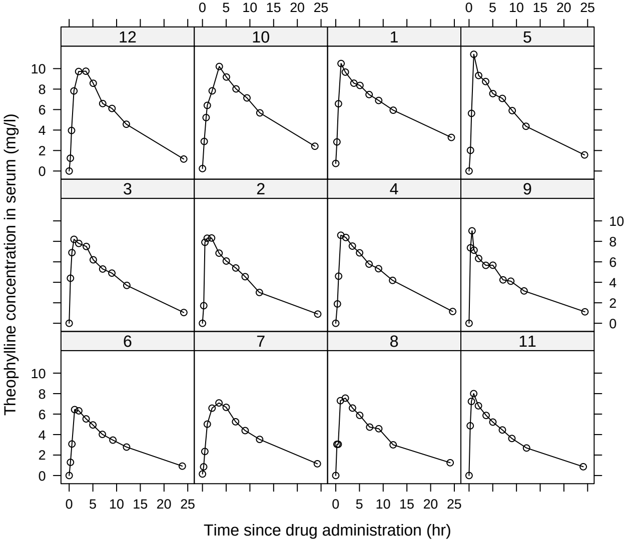
Crecimiento de plantas de Soya (Davidian y Giltinan, 1995)

Observaciones:

- Estudio típico en Agricultura, Genética con observaciones ligeramente desbalanceadas.
- Relación no lineal entre peso y tiempo.
- **Modelo logístico** usualmente utilizado para explicar el crecimiento

$$Y = \frac{\beta_1}{1 + \beta_2 \exp(\beta_3 x)}, \qquad \beta_1, \beta_2 > 0, \beta_3 < 0.$$

Cinética de Teofilina (Boeckmann, Sheiner y Beal, 1984)



Cinética de Teofilina (Boeckmann, Sheiner y Beal, 1984)

Observaciones:

- Estudio de laboratorio (farmacocinética) con pocos individuos y varias observaciones balanceadas.

- Modelo mecánico: modelo de primer orden de un compartimento.

$$C_t = \frac{\Delta K k_a}{Cl(k_a - K)}[\exp(-Kt) - \exp(-k_a t)].$$

donde Δ es la dosis inicial, k_a tasa de absorción, K tasa de eliminación, Cl clearance.

Definición 1 (Modelo GMANOVA)

Un modelo GMANOVA (curvas de crecimiento) está definido como

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z} + \boldsymbol{E},$$

donde $\boldsymbol{X} \in \mathbb{R}^{n\times m}$ y $\boldsymbol{Z} \in \mathbb{R}^{q\times p}$ son matrices de diseño con $\operatorname{rg}(\boldsymbol{X}) = m$ y $\operatorname{rg}(\boldsymbol{Z}) = q$, respectivamente. $\boldsymbol{B} \in \mathbb{R}^{q\times r}$ es matriz de coeficientes de regresión y

$$\boldsymbol{E} \sim \mathsf{N}_{n,p}(\boldsymbol{0}, \boldsymbol{I}_n, \boldsymbol{\Sigma}).$$

De ahí que

$$\boldsymbol{Y} \sim \mathsf{N}_{n,p}(\boldsymbol{X}\boldsymbol{B}\boldsymbol{Z}, \boldsymbol{I}_n, \boldsymbol{\Sigma}).$$

Considere

$$\boldsymbol{Q}(\boldsymbol{B}) = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})^\top(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z}).$$

El estimador LS en el modelo GMANOVA está definido como la solución del problema:

$$\min_{B} \operatorname{tr} \boldsymbol{Q}(\boldsymbol{B}) := \min_{B} \operatorname{tr}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})^\top(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z}).$$

En efecto, diferenciando con relación a $\boldsymbol{B}$, obtenemos

$$\begin{aligned}
\mathrm{d}_B \operatorname{tr} \boldsymbol{Q}(\boldsymbol{B}) &= -\operatorname{tr} \boldsymbol{Z}^\top(\mathrm{d}\boldsymbol{B})^\top\boldsymbol{X}^\top(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z}) - \operatorname{tr}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})^\top\boldsymbol{X}(\mathrm{d}\boldsymbol{B})\boldsymbol{Z} \\
&= -2\operatorname{tr} \boldsymbol{X}^\top(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})\boldsymbol{Z}^\top(\mathrm{d}\boldsymbol{B})^\top,
\end{aligned}$$

de ahí que la ecuación de estimación para $\boldsymbol{B}$ es dada por:

$$\boldsymbol{X}^\top(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})\boldsymbol{Z}^\top = \boldsymbol{0}.$$

Estimación LS

Desde la ecuación de estimación $\boldsymbol{X}^\top(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})\boldsymbol{Z}^\top=\boldsymbol{0}$, tenemos que

$$\boldsymbol{X}^\top\boldsymbol{X}\widehat{\boldsymbol{B}}\boldsymbol{Z}\boldsymbol{Z}^\top=\boldsymbol{X}^\top\boldsymbol{Y}\boldsymbol{Z}^\top,$$

es decir el estimador LS para $\boldsymbol{B}$ asume la forma:

$$\widehat{\boldsymbol{B}}=(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{Y}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}.$$

Por otro lado,

$$\begin{aligned}\mathsf{d}_B^2\operatorname{tr}\boldsymbol{Q}(\boldsymbol{B})&=2\operatorname{tr}\boldsymbol{X}^\top\boldsymbol{X}(\mathsf{d}\boldsymbol{B})\boldsymbol{Z}\boldsymbol{Z}^\top(\mathsf{d}\boldsymbol{B})^\top\\&=2(\mathsf{d}\operatorname{vec}\boldsymbol{B})^\top(\boldsymbol{Z}\boldsymbol{Z}^\top\otimes\boldsymbol{X}^\top\boldsymbol{X})\,\mathsf{d}\operatorname{vec}\boldsymbol{B},\end{aligned}$$

y como $\boldsymbol{Z}\boldsymbol{Z}^\top\otimes\boldsymbol{X}^\top\boldsymbol{X}$ es matriz positiva definida, sigue que $\widehat{\boldsymbol{B}}$ es mínimo (global).

Para el ejemplo de datos dentales tenemos que:

$$\boldsymbol{X} = \begin{pmatrix} \mathbf{1}_{16} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{11} \end{pmatrix}, \qquad \boldsymbol{Z} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 8 & 10 & 12 & 14 \end{pmatrix},$$

mientras que $\boldsymbol{Y} \in \mathbb{R}^{27\times 4}$. De este modo, tenemos que $\boldsymbol{B} \in \mathbb{R}^{2\times 2}$ y $\boldsymbol{\Sigma}$ es matriz definida positiva 4×4.

Comandos en R

```
library(nlme) # Biblioteca nlme contiene los datos 'dentales'
data(Orthodont)
names(Orthodont)
[1] "distance" "age"       "Subject"  "Sex"

# matriz de respuestas
y <- Orthodont$distance
y <- matrix(y, ncol = 4, byrow = TRUE)

y
      [,1] [,2] [,3] [,4]
 [1,] 26.0 25.0 29.0 31.0
 [2,] 21.5 22.5 23.0 26.5
 [3,] 23.0 22.5 24.0 27.5
 [4,] 25.5 27.5 26.5 27.0

...

[24,] 23.0 23.0 23.5 24.0
[25,] 20.0 21.0 22.0 21.5
[26,] 16.5 19.0 19.0 19.5
[27,] 24.5 25.0 28.0 28.0
```

```r
# matrices de diseno
x <- cbind(c(rep(1,16), rep(0,11)), c(rep(0,16), rep(1,11)))
z <- rbind(rep(1,4), c(8,10,12,14))

# contruye estimador para B
xx <- crossprod(x)
zz <- crossprod(t(z))
xyz <- crossprod(x, y %*% t(z))
B <- solve(xx, xyz %*% solve(zz))

# Salida
B
        Intercept       age
Male     16.34063 0.7843750
Female   17.37273 0.4795455
```

Comandos en R

```
# construye estimador para Sigma
res <- y - x %*% B %*% z
n <- nrow(y)
Sigma <- crossprod(res) / n

# Salida
Sigma
         [,1]     [,2]     [,3]     [,4]
[1,] 5.054480 2.457757 3.615701 2.531994
[2,] 2.457757 3.958162 2.717032 3.039186
[3,] 3.615701 2.717032 5.978775 3.821699
[4,] 2.531994 3.039186 3.821699 4.629217

# Calcula covarianza (estimada) del estimador de B
kronecker(solve(xx), solve(zz, z %*% Sigma %*% t(z)) %*% solve(zz))
            [,1]         [,2]       [,3]        [,4]
[1,]  0.96056230 -0.071385371  0.0000000  0.00000000
[2,] -0.07138537  0.006848071  0.0000000  0.00000000
[3,]  0.00000000  0.000000000  1.3971815 -0.10383327
[4,]  0.00000000  0.000000000 -0.1038333  0.00996083
```

Estimación ML GMANOVA

- Primeramente, asumiremos que la matriz $\boldsymbol{\Sigma}$ es no estructurada, es decir corresponde a una matriz simétrica y definida positiva.

- Luego consideraremos estructuras lineales del tipo:

  $$\boldsymbol{\Sigma} = \boldsymbol{Z}^\top \boldsymbol{\Gamma} \boldsymbol{Z} + \boldsymbol{G}^\top \boldsymbol{\Phi} \boldsymbol{G},$$

  donde $\boldsymbol{G} \in \mathcal{Q}$ tal que

  $$\mathcal{Q} = \{\boldsymbol{G} : \boldsymbol{G} \in \mathbb{R}^{p \times (p-m)}, \boldsymbol{G}\boldsymbol{Z}^\top = \boldsymbol{0}\}.$$

  Clase que es conocida como estructura de covarianza (simple) de Rao.

Estimación ML en GMANOVA: caso no estructurado

Sea $\boldsymbol{\Theta} = (\boldsymbol{B}, \boldsymbol{\Sigma})$, entonces la función de log-verosimilitud adopta la forma

$$\begin{aligned}\ell(\boldsymbol{\Theta}) &= \log \Big\{(2\pi)^{-np/2}|\boldsymbol{\Sigma}|^{-n/2} \exp\big\{ -\tfrac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{Q}(\boldsymbol{B})\big\}\Big\} \\ &= -\frac{np}{2}\log 2\pi - \frac{n}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{Q}(\boldsymbol{B}),\end{aligned}$$

donde

$$\boldsymbol{Q}(\boldsymbol{B}) = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})^\top(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z}),$$

corresponde a la matriz de suma de productos cruzados (de errores).

Diferenciando con relación a $\boldsymbol{B}$, obtenemos:

$$\mathsf{d}_B\,\ell(\boldsymbol{\Theta}) = -\tfrac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\,\mathsf{d}_B\,\boldsymbol{Q}(\boldsymbol{B}),$$

por otro lado,

$$\mathsf{d}_B\,\boldsymbol{Q}(\boldsymbol{\Theta}) = -\operatorname{tr}\Big\{(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})^\top\boldsymbol{X}(\mathsf{d}\boldsymbol{B})\boldsymbol{Z} + \boldsymbol{Z}^\top(\mathsf{d}\boldsymbol{B})^\top\boldsymbol{X}^\top(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})\Big\},$$

De este modo, recordando que $\operatorname{tr}\boldsymbol{A} = \operatorname{tr}\boldsymbol{A}^\top$, obtenemos

$$\mathsf{d}_B\,\ell(\boldsymbol{\Theta}) = \operatorname{tr}\boldsymbol{Z}\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})^\top\boldsymbol{X}\,\mathsf{d}\boldsymbol{B}.$$

Por tanto, la ecuación de estimación para $\boldsymbol{B}$ (obtenida desde $\mathsf{d}_B\,\ell(\boldsymbol{\Theta}) = 0$) asume la forma:

$$\boldsymbol{X}^\top(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}^\top = \boldsymbol{0},$$

Ahora, diferenciando con relación a $\boldsymbol{\Sigma}$, obtenemos:

$$\begin{aligned} \mathsf{d}_{\Sigma}\,\ell(\boldsymbol{\Theta}) &= -\frac{n}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\,\mathsf{d}\boldsymbol{\Sigma} + \frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}(\mathsf{d}\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}\boldsymbol{Q}(\boldsymbol{B}) \\ &= -\frac{n}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\Big(\boldsymbol{\Sigma} - \frac{1}{n}\boldsymbol{Q}(\boldsymbol{B})\Big)\boldsymbol{\Sigma}^{-1}\,\mathsf{d}\boldsymbol{\Sigma}. \end{aligned}$$

De este modo, la ecuación de estimación para $\boldsymbol{\Sigma}$ es dada por:

$$n\boldsymbol{\Sigma} - \boldsymbol{Q}(\boldsymbol{B}) = \mathbf{0}.$$

Por tanto, los MLEs de $\boldsymbol{B}$ y $\boldsymbol{\Sigma}$ deben satisfacer el siguiente sistema de ecuaciones:

$$\begin{aligned} &\boldsymbol{X}^\top(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{Z}^\top = \boldsymbol{0} \\ &n\widehat{\boldsymbol{\Sigma}} - \boldsymbol{Q}(\widehat{\boldsymbol{B}}) = \boldsymbol{0}. \end{aligned} \tag{1}$$

El siguiente resultado, presenta la solución $(\widehat{\boldsymbol{B}}, \widehat{\boldsymbol{\Sigma}})$ para las ecuaciones de verosimilitud anteriores.

Resultado 1 (MLE-UN en GMANOVA)

Para el modelo GMANOVA con matrices de diseño $\boldsymbol{X}$, $\boldsymbol{Z}$ de rango completo. Se tiene que la solución de la ecuación de verosimilitud en (1) es única y es dada por

$$\widehat{\boldsymbol{B}} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{Y}\boldsymbol{S}^{-1}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{S}^{-1}\boldsymbol{Z}^\top)^{-1},$$

y

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n}(\boldsymbol{Y}-\boldsymbol{X}\widehat{\boldsymbol{B}}\boldsymbol{Z})^\top(\boldsymbol{Y}-\boldsymbol{X}\widehat{\boldsymbol{B}}\boldsymbol{Z})$$

donde $\boldsymbol{S} = \boldsymbol{Y}^\top(\boldsymbol{I}-\boldsymbol{H}_X)\boldsymbol{Y}$ y $\boldsymbol{H}_X = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}$.

Comandos en R

```
## Supondremos que los datos 'x', 'y', 'z' estan cargados en
## nuestro directorio de trabajo.

# calculos en el modelo de regresion multivariado
xy <- crossprod(x, y)
r  <- y - x %*% solve(xx, xy)
S  <- crossprod(r)

# Salida
S
          [,1]      [,2]      [,3]      [,4]
[1,] 135.38636  67.92045  97.75568  67.75568
[2,]  67.92045 104.61932  73.17898  82.92898
[3,]  97.75568  73.17898 161.39347 103.26847
[4,]  67.75568  82.92898 103.26847 124.64347
```

Comandos en R

```
# construye estimador para B
zsz <- z %*% solve(S, t(z))
rhs <- solve(S, t(z) %*% solve(zsz))
B <- solve(xx, xy) %*% rhs

# Salida
B
        Intercept       age
Male     15.84229 0.8268033
Female   17.42537 0.4763647

# construye estimador para Sigma
res <- y - x %*% B %*% z
n <- nrow(y)
Sigma <- crossprod(res) / n

# Salida
Sigma
         [,1]     [,2]     [,3]     [,4]
[1,] 5.119199 2.440902 3.610510 2.522243
[2,] 2.440902 3.927948 2.717514 3.062349
[3,] 3.610510 2.717514 5.979798 3.823461
[4,] 2.522243 3.062349 3.823461 4.617984
```

¿Existe alguna condición en la que el estimador ML bajo el modelo GMANOVA

$$\widehat{\boldsymbol{B}}_S = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{Y}\boldsymbol{S}^{-1}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{S}^{-1}\boldsymbol{Z}^\top)^{-1},$$

coincida con el estimador LS?

$$\widehat{\boldsymbol{B}} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{Y}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1},$$

Observación:

La respuesta a la pregunta anterior puede ser resuelto mediante modelar la estructura de covarianza.

Estructura de covarianza simple de Rao

Definición 2 (Estructura de covarianza simple de Rao)

La estructura de covarianza simple de Rao (SCS) es dada por

$$\boldsymbol{\Sigma} = \boldsymbol{Z}^\top \boldsymbol{\Gamma} \boldsymbol{Z} + \boldsymbol{G}^\top \boldsymbol{\Phi} \boldsymbol{G},$$

donde $\boldsymbol{Z} \in \mathbb{R}^{q \times p}$ $(\mathrm{rg}(\boldsymbol{Z}) = q)$ y $\boldsymbol{G} \in \mathcal{Q}$, con

$$\mathcal{G} = \{\boldsymbol{G} : \boldsymbol{G} \in \mathbb{R}^{p \times (p-q)}, \boldsymbol{G}\boldsymbol{Z}^\top = \boldsymbol{0} = \boldsymbol{Z}\boldsymbol{G}^\top\}.$$

Resultado 2 (MLE-SCS en GMANOVA)

Para el modelo GMANOVA con estrutura de covarianza simple de Rao, los estimadores ML pueden ser expresados como:

$$\widehat{\boldsymbol{B}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y} \boldsymbol{Z}^\top (\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}$$

$$\widehat{\boldsymbol{\Gamma}} = \frac{1}{n} (\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1} \boldsymbol{Z}\boldsymbol{S}\boldsymbol{Z}^\top (\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}$$

$$\widehat{\boldsymbol{\Phi}} = \frac{1}{n} (\boldsymbol{G}\boldsymbol{G}^\top)^{-1} \boldsymbol{G}\boldsymbol{Y}^\top \boldsymbol{Y}\boldsymbol{G}^\top (\boldsymbol{G}\boldsymbol{G}^\top)^{-1}.$$

Sea

$$\widetilde{\boldsymbol{Z}} = \begin{pmatrix} \boldsymbol{Z} \\ \boldsymbol{G} \end{pmatrix}.$$

Note que $\boldsymbol{\Sigma} = |\boldsymbol{Z}^\top \boldsymbol{\Gamma} \boldsymbol{Z} + \boldsymbol{G}^\top \boldsymbol{\Phi} \boldsymbol{G}|$, puede ser escrito como

$$\begin{aligned} |\boldsymbol{\Sigma}| &= \left| (\boldsymbol{Z}^\top, \boldsymbol{G}^\top) \begin{pmatrix} \boldsymbol{\Gamma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Phi} \end{pmatrix} \begin{pmatrix} \boldsymbol{Z} \\ \boldsymbol{G} \end{pmatrix} \right| = \left| \begin{pmatrix} \boldsymbol{Z} \\ \boldsymbol{G} \end{pmatrix} (\boldsymbol{Z}^\top, \boldsymbol{G}^\top) \begin{pmatrix} \boldsymbol{\Gamma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Phi} \end{pmatrix} \right| \\ &= \begin{vmatrix} \boldsymbol{Z}\boldsymbol{Z}^\top & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}\boldsymbol{G}^\top \end{vmatrix} \begin{vmatrix} \boldsymbol{\Gamma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Phi} \end{vmatrix} = |\boldsymbol{Z}\boldsymbol{Z}^\top||\boldsymbol{G}\boldsymbol{G}^\top||\boldsymbol{\Gamma}||\boldsymbol{\Phi}|. \end{aligned}$$

Por otro lado,

$$\begin{aligned} \boldsymbol{\Sigma}^{-1} &= \begin{pmatrix} \boldsymbol{Z} \\ \boldsymbol{G} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\Gamma}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Phi}^{-1} \end{pmatrix} (\boldsymbol{Z}^\top, \boldsymbol{G}^\top)^{-1} \\ &= (\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}, \boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}) \begin{pmatrix} \boldsymbol{\Gamma}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Phi}^{-1} \end{pmatrix} \begin{pmatrix} (\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}\boldsymbol{Z} \\ (\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{G} \end{pmatrix} \\ &= \boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}\boldsymbol{\Gamma}^{-1}(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}\boldsymbol{Z} + \boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{\Phi}^{-1}(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{G} \end{aligned}$$

Considere $\boldsymbol{\Theta} = (\boldsymbol{B}, \boldsymbol{\Gamma}, \boldsymbol{\Phi})$, entonces la función de log-verosimilitud es dada por

$$\begin{aligned}\ell(\boldsymbol{\Theta}) &= \log\Big\{(2\pi)^{-np/2}|\boldsymbol{\Sigma}|^{-n/2}\exp\big\{-\tfrac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{Q}(\boldsymbol{B})\big\}\Big\}\\ &= -\frac{np}{2}\log 2\pi - \frac{n}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{Q}(\boldsymbol{B}),\end{aligned}$$

donde

$$\boldsymbol{Q}(\boldsymbol{B}) = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z})^\top(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}\boldsymbol{Z}).$$

Note que

$$\log|\boldsymbol{\Sigma}| = \log|\boldsymbol{Z}\boldsymbol{Z}^\top| + \log|\boldsymbol{G}\boldsymbol{G}^\top| + \log|\boldsymbol{\Gamma}| + \log|\boldsymbol{\Phi}|.$$

Por otro lado,

$$\begin{aligned}
\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{Q}(\boldsymbol{B}) &= \operatorname{tr}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}\boldsymbol{\Gamma}^{-1}(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}\boldsymbol{Z}\boldsymbol{Q}(\boldsymbol{B}) \\
&\quad + \operatorname{tr}\boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{\Phi}^{-1}(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{G}\boldsymbol{Q}(\boldsymbol{B}) \\
&= \operatorname{tr}\boldsymbol{\Gamma}^{-1}(\boldsymbol{Y}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1} - \boldsymbol{X}\boldsymbol{B})^\top(\boldsymbol{Y}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1} - \boldsymbol{X}\boldsymbol{B}) \\
&\quad + \operatorname{tr}\boldsymbol{\Phi}^{-1}(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{G}\boldsymbol{Y}^\top\boldsymbol{Y}\boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}
\end{aligned}$$

Estimación ML en GMANOVA: estructura simple de Rao

Finalmente, la función de log-verosimilitud para $\mathbf{\Theta} = (\boldsymbol{B}, \mathbf{\Gamma}, \mathbf{\Phi})$ asume la forma

$$\begin{aligned}\ell(\mathbf{\Theta}) = & -\frac{np}{2}\log 2\pi - \frac{n}{2}\log|\boldsymbol{Z}\boldsymbol{Z}^\top| - \frac{n}{2}\log|\boldsymbol{G}\boldsymbol{G}^\top| \\ & - \frac{n}{2}\log|\mathbf{\Gamma}| - \frac{1}{2}\operatorname{tr}\mathbf{\Gamma}^{-1}(\boldsymbol{Y}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1} - \boldsymbol{X}\boldsymbol{B})^\top(\boldsymbol{Y}\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1} - \boldsymbol{X}\boldsymbol{B}) \\ & - \frac{n}{2}\log|\mathbf{\Phi}| - \frac{1}{2}\operatorname{tr}\mathbf{\Phi}^{-1}(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{G}\boldsymbol{Y}^\top\boldsymbol{Y}\boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}.\end{aligned}$$

De este modo, diferenciando con relación a $\mathbf{\Gamma}$, obtenemos:

$$
\begin{aligned}
\mathsf{d}_{\Gamma}\,\ell(\mathbf{\Theta}) &= -\frac{n}{2}\operatorname{tr}\mathbf{\Gamma}^{-1}\,\mathsf{d}\mathbf{\Gamma} \\
&\quad + \frac{1}{2}\operatorname{tr}\mathbf{\Gamma}^{-1}(\mathsf{d}\mathbf{\Gamma})\mathbf{\Gamma}^{-1}(\boldsymbol{Y}\boldsymbol{Z}^{\top}(\boldsymbol{Z}\boldsymbol{Z}^{\top})^{-1}-\boldsymbol{X}\boldsymbol{B})^{\top}(\boldsymbol{Y}\boldsymbol{Z}^{\top}(\boldsymbol{Z}\boldsymbol{Z}^{\top})^{-1}-\boldsymbol{X}\boldsymbol{B}) \\
&= -\frac{n}{2}\operatorname{tr}\mathbf{\Gamma}^{-1}\Big\{\mathbf{\Gamma}-\frac{1}{n}(\boldsymbol{Y}\boldsymbol{Z}^{\top}(\boldsymbol{Z}\boldsymbol{Z}^{\top})^{-1}-\boldsymbol{X}\boldsymbol{B})^{\top}(\boldsymbol{Y}\boldsymbol{Z}^{\top}(\boldsymbol{Z}\boldsymbol{Z}^{\top})^{-1}-\boldsymbol{X}\boldsymbol{B})\Big\}\mathbf{\Gamma}^{-1}\,\mathsf{d}\mathbf{\Gamma}.
\end{aligned}
$$

Por tanto el estimador ML para $\mathbf{\Gamma}$ es dado por:

$$
\begin{aligned}
\widehat{\mathbf{\Gamma}} &= \frac{1}{n}(\boldsymbol{Y}\boldsymbol{Z}^{\top}(\boldsymbol{Z}\boldsymbol{Z}^{\top})^{-1}-\boldsymbol{X}\widehat{\boldsymbol{B}})^{\top}(\boldsymbol{Y}\boldsymbol{Z}^{\top}(\boldsymbol{Z}\boldsymbol{Z}^{\top})^{-1}-\boldsymbol{X}\widehat{\boldsymbol{B}}) \\
&= \frac{1}{n}(\boldsymbol{Z}\boldsymbol{Z}^{\top})^{-1}\boldsymbol{Z}\boldsymbol{S}\boldsymbol{Z}^{\top}(\boldsymbol{Z}\boldsymbol{Z}^{\top})^{-1}
\end{aligned}
$$

De este modo, diferenciando con relación a $\mathbf{\Phi}$, obtenemos:

$$\begin{aligned}\mathsf{d}_{\Phi}\,\ell(\mathbf{\Theta}) &= -\frac{n}{2}\operatorname{tr}\mathbf{\Phi}^{-1}\,\mathsf{d}\mathbf{\Phi} + \frac{1}{2}\operatorname{tr}\mathbf{\Phi}^{-1}(\mathsf{d}\mathbf{\Phi})\mathbf{\Phi}^{-1}(\boldsymbol{G}\boldsymbol{G}^{\top})^{-1}\boldsymbol{G}\boldsymbol{Y}^{\top}\boldsymbol{Y}\boldsymbol{G}^{\top}(\boldsymbol{G}\boldsymbol{G}^{\top})^{-1} \\ &= -\frac{n}{2}\operatorname{tr}\mathbf{\Phi}^{-1}\Big\{\mathbf{\Phi} - \frac{1}{n}(\boldsymbol{G}\boldsymbol{G}^{\top})^{-1}\boldsymbol{G}\boldsymbol{Y}^{\top}\boldsymbol{Y}\boldsymbol{G}^{\top}(\boldsymbol{G}\boldsymbol{G}^{\top})^{-1}\Big\}\mathbf{\Phi}^{-1}\,\mathsf{d}\mathbf{\Phi}.\end{aligned}$$

Así, el estimador ML para $\mathbf{\Phi}$ resulta:

$$\widehat{\mathbf{\Phi}} = \frac{1}{n}(\boldsymbol{G}\boldsymbol{G}^{\top})^{-1}\boldsymbol{G}\boldsymbol{Y}^{\top}\boldsymbol{Y}\boldsymbol{G}^{\top}(\boldsymbol{G}\boldsymbol{G}^{\top})^{-1}.$$

Finalmente, note que

$$\boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{G} = \boldsymbol{I}_q - \boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}\boldsymbol{Z}.$$

Esto permite construir una matriz asociada a $\boldsymbol{G}$ tal que $\boldsymbol{G}\boldsymbol{Z}^\top = \boldsymbol{0}$ y que además

$$\widehat{\boldsymbol{\Sigma}} = \boldsymbol{Z}^\top\widehat{\boldsymbol{\Gamma}}\boldsymbol{Z} + \boldsymbol{G}^\top\widehat{\boldsymbol{\Phi}}\boldsymbol{G}.$$

Basta apreciar que

$$\begin{aligned}\boldsymbol{G}^\top\widehat{\boldsymbol{\Phi}}\boldsymbol{G} &= \frac{1}{n}\boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{G}\boldsymbol{Y}^\top\boldsymbol{Y}\boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{G}^\top)^{-1}\boldsymbol{G} \\ &= \frac{1}{n}(\boldsymbol{I}_q - \boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}\boldsymbol{Z})\boldsymbol{Y}^\top\boldsymbol{Y}(\boldsymbol{I}_q - \boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{Z}^\top)^{-1}\boldsymbol{Z}).\end{aligned}$$

```
## Supondremos que los datos 'x', 'y', 'z' estan cargados en
## nuestro directorio de trabajo.

# calculos en el modelo de regresion multivariado
xy <- crossprod(x, y)
r  <- y - x %*% solve(xx, xy)
S  <- crossprod(r)

# Salida
S
          [,1]      [,2]      [,3]      [,4]
[1,] 135.38636  67.92045  97.75568  67.75568
[2,]  67.92045 104.61932  73.17898  82.92898
[3,]  97.75568  73.17898 161.39347 103.26847
[4,]  67.75568  82.92898 103.26847 124.64347
```

Comandos en R

```
# contruye estimador para B
xx <- crossprod(x)
zz <- crossprod(t(z))
xyz <- crossprod(x, y %*% t(z))
B <- solve(xx, xyz %*% solve(zz))

# Salida
B
       Intercept       age
Male    16.34063 0.7843750
Female  17.37273 0.4795455

# construye estimador para Gamma
zsz <- z %*% S %*% t(z)
Gamma <- solve(zz, zsz %*% solve(zz)) / n

# Salida
Gamma
          [,1]       [,2]
[1,] 15.368997 -1.1421659
[2,] -1.142166  0.1095691
```

Comandos en R

```
p <- ncol(z)
res <- diag(p) - t(z) %*% solve(zz, z)
res
     [,1] [,2] [,3] [,4]
[1,]  0.3 -0.4 -0.1  0.2
[2,] -0.4  0.7 -0.2 -0.1
[3,] -0.1 -0.2  0.7 -0.4
[4,]  0.2 -0.1 -0.4  0.3

yy <- crossprod(y)
gg <- res %*% yy %*% res /n
gg
           [,1]        [,2]       [,3]        [,4]
[1,]  0.4084259 -0.66972222  0.1141667  0.14712963
[2,] -0.6697222  1.40500000 -0.8008333  0.06555556
[3,]  0.1141667 -0.80083333  1.2591667 -0.57250000
[4,]  0.1471296  0.06555556 -0.5725000  0.35981481

# construye estimador para Sigma
Sigma <- crossprod(z, Gamma %*% z) + gg

# Salida
Sigma
         [,1]     [,2]     [,3]     [,4]
[1,] 4.515192 2.905818 3.158481 2.660218
[2,] 2.905818 4.887591 2.588808 3.362248
[3,] 3.158481 2.588808 4.994136 3.507796
[4,] 2.660218 3.362248 3.507796 5.223715
```

Observaciones:

- Soluciones explícitas (no iterativas) para estimadores ML y GLS.
- Tal resultado no es válido para estructuras de covarianza más generales.
- GMANOVA es útil solamente para datos balanceados

Modelo ANOVA

Considere el modelo

$$Y_{ij} = \mu + b_i + u_{ij}, \quad i = 1, \dots, p; j = 1, \dots, n,$$

donde $\{b_i\}$ son v.a independientes con $\mathsf{E}(b_i) = \theta_i$, $\mathsf{var}(b_i) = \psi^2$ y $\{u_{ij}\}$ son v.a. iid con media cero y varianza ϕ^2.

Sea $\widetilde{b}_i = b_i - \theta_i$ y $\epsilon_{ij} = b_i + u_{ij}$, luego

$$Y_{ij} = \mu + \theta_i + \epsilon_{ij}, \quad i = 1, \dots, p; j = 1, \dots, n.$$

EX UMBRA IN SOLEM

El modelo anterior puede escribirse como

$$\boldsymbol{Y}_j = \boldsymbol{X}_j \boldsymbol{\beta} + \boldsymbol{\epsilon}_j, \qquad j = 1, \dots, n,$$

donde $\boldsymbol{Y}_j = (Y_{1j}, \dots, Y_{pj})^T$, $\boldsymbol{\epsilon}_j = \widetilde{b}_i \mathbf{1} + \boldsymbol{u}_j$, de este modo

$$\mathsf{E}(\boldsymbol{\epsilon}_j) = \mathbf{0}, \qquad \mathsf{Cov}(\boldsymbol{\epsilon}_j) = \psi^2 \mathbf{1}\mathbf{1}^T + \phi^2 \boldsymbol{I}_p.$$

Sea $\boldsymbol{Y}_i = (Y_{i1}, \dots, Y_{im_i})^T$ vector aleatorio m_i-dimensional. Considere

$$\boldsymbol{Y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i, \qquad i = 1, \dots, n,$$

donde $\boldsymbol{\beta} \in \mathbb{R}^p$ es común para todos los indivíduos.

Supuesto distribucional:

$$\boldsymbol{Y}_i \overset{\text{ind}}{\sim} \mathsf{N}_{m_i}(\boldsymbol{X}_i \boldsymbol{\beta}, \boldsymbol{\Sigma}_i(\boldsymbol{\tau})), \qquad i = 1, \dots, n,$$

con $\boldsymbol{\tau} = (\tau_1, \dots, \tau_K)^T$.

Función de log-verosimilitud

Considere

$$\boldsymbol{\Sigma}_i = \sigma^2 \boldsymbol{R}_i, \qquad \boldsymbol{R}_i = \boldsymbol{R}_i(\boldsymbol{\lambda}).$$

La log-verosimilitud para $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma^2, \boldsymbol{\lambda}^T)^T$ asume la forma

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= -\frac{N}{2}\log 2\pi - \frac{N}{2}\log\sigma^2 \\ &\quad - \frac{1}{2}\sum_{i=1}^n \left\{ \log|\boldsymbol{R}_i| + (\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})^T \boldsymbol{R}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})/\sigma^2 \right\},\end{aligned}$$

donde $N = \sum_{i=1}^n m_i$.

Estimadores ML

Para $\boldsymbol{\lambda}$ fijo, los estimadores ML de $\boldsymbol{\beta}$ y σ^2 corresponden a GLS, esto es

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda}) = \Big(\sum_{i=1}^n \boldsymbol{X}_i^T \boldsymbol{R}_i^{-1} \boldsymbol{X}_i\Big)^{-1} \sum_{i=1}^n \boldsymbol{X}_i^T \boldsymbol{R}_i^{-1} \boldsymbol{Y}_i,$$

$$\widehat{\sigma}^2(\boldsymbol{\lambda}) = \frac{1}{N} \sum_{i=1}^n (\boldsymbol{Y}_i - \boldsymbol{X}_i \boldsymbol{\beta})^T \boldsymbol{R}_i^{-1} (\boldsymbol{Y}_i - \boldsymbol{X}_i \boldsymbol{\beta}).$$

$\widehat{\boldsymbol{\lambda}}$ es obtenido maximizando la log-verosimilitud perfilada

$$\ell_*(\boldsymbol{\lambda}) = c - \frac{1}{2} \sum_{i=1}^n \log |\boldsymbol{R}_i| - N \log \sum_{i=1}^n \|\boldsymbol{R}_i^{-1/2} (\boldsymbol{Y}_i - \boldsymbol{X}_i \widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda}))\|^2,$$

donde c representa una constante.

El cálculo de $\widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda})$, $\sigma^2(\boldsymbol{\lambda})$ y $\widehat{\boldsymbol{\lambda}}$ puede ser amenizado si consideramos una descomposición raíz cuadrada (Cholesky, LDL) de $\boldsymbol{R}_i$,

$$\boldsymbol{R}_i = \boldsymbol{L}_i \boldsymbol{L}_i^T, \qquad \boldsymbol{L}_i \text{ triangular inferior},$$

y considere

$$\tilde{\boldsymbol{Y}}_i = \boldsymbol{L}_i^{-1} \boldsymbol{Y}_i, \qquad \tilde{\boldsymbol{X}}_i = \boldsymbol{L}_i^{-1} \boldsymbol{X}_i,$$

En cuyo caso $\widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda})$ y $\widehat{\sigma}^2(\boldsymbol{\lambda})$ están dados por

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda}) = \Big(\sum_{i=1}^n \tilde{\boldsymbol{X}}_i^T \tilde{\boldsymbol{X}}_i \Big)^{-1} \sum_{i=1}^n \tilde{\boldsymbol{X}}_i^T \tilde{\boldsymbol{Y}}_i, \quad \widehat{\sigma}^2(\boldsymbol{\lambda}) = \frac{1}{N} \sum_{i=1}^n \|\tilde{\boldsymbol{Y}}_i - \tilde{\boldsymbol{X}}_i \widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda})\|^2.$$

Estimadores ML

La función de log-verosimilitud perfilada puede ser evaluada como

$$\ell_*(\boldsymbol{\lambda}) = c - \sum_{i=1}^n \log|\boldsymbol{L}_i| - N\log\sum_{i=1}^n \|\tilde{\boldsymbol{Y}}_i - \tilde{\boldsymbol{X}}_i\widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda})\|^2.$$

Note que,

$$\sum_{i=1}^n \|\tilde{\boldsymbol{Y}}_i - \tilde{\boldsymbol{X}}_i\widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda})\|^2 = \|\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{X}}\widehat{\boldsymbol{\beta}}(\boldsymbol{\lambda})\|^2$$

donde

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_n \end{pmatrix}, \qquad \boldsymbol{Y} = \begin{pmatrix} \boldsymbol{Y}_1 \\ \vdots \\ \boldsymbol{Y}_n \end{pmatrix}.$$

Modelo:

$$Y_{ij} = \beta_0 + \beta_1 \mathsf{Sexo} + \beta_2(\mathsf{edad}_j - 11) + \beta_3(\mathsf{edad}_j - 11)\mathsf{Sexo} + \epsilon_{ij},$$

donde

$$\boldsymbol{\epsilon}_i \overset{\mathsf{ind}}{\sim} \mathsf{N}_4(\mathbf{0}, \sigma^2 \boldsymbol{R}_i), \qquad i = 1, \dots, 27,$$

con $\boldsymbol{R}_i$ matriz no estructurada.

Gráfico con rectas ajustadas