5

Contraste de hipótesis lineales

5.1. Hipótesis lineales contrastables

Consideremos el modelo lineal $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ y $\mathrm{var}(\mathbf{Y}) = \sigma^2\mathbf{I}$.

Una hipótesis lineal consiste en una o varias restricciones lineales planteadas sobre los parámetros $\boldsymbol{\beta}$. En un diseño de rango máximo $\mathrm{rg}\,\mathbf{X} = m$ vamos a ver que cualquier hipótesis lineal es contrastable (testable o *demostrable*), es decir, es posible encontrar un estadístico (el test F del teorema 5.3.1) mediante el cual podemos decidir si se rechaza o acepta la hipótesis. Si $\mathrm{rg}\,\mathbf{X} = r < m$, entonces pueden existir hipótesis estadísticamente no contrastables.

Definición 5.1.1

Una hipótesis lineal de rango q sobre los parámetros $\boldsymbol{\beta}$ es un conjunto de restricciones lineales

$$a_{i1}\beta_1 + \cdots + a_{im}\beta_m = 0 \qquad i = 1, \ldots, q$$

Si escribimos la matriz de la hipótesis como

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{q1} & \cdots & a_{qm} \end{pmatrix} \qquad rg\ \mathbf{A} = q$$

entonces las restricciones se resumen en

$$H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{0}$$

Una hipótesis se dice que es contrastable o demostrable *si el conjunto $\mathbf{A}\boldsymbol{\beta}$ es un sistema de funciones paramétricas estimables. Entonces, las filas de $\mathbf{A}$ son combinación lineal de las filas de la matriz de diseño $\mathbf{X}$, es decir, que existe una matriz $\mathbf{B}$ de tamaño $q \times n$ tal que*

$$\mathbf{A} = \mathbf{B}\mathbf{X}$$

También $\mathbf{B}$ puede ser $q \times k$ si consideramos la matriz de diseño reducida $\mathbf{X}_R$ $k \times m$.

Cuando $\mathbf{X}$ no es de rango máximo, un conjunto de restricciones $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$ donde las filas de $\mathbf{A}$ son linealmente independientes de las filas de $\mathbf{X}$ no forman una alternativa al modelo general, en el sentido de un modelo más sencillo. En realidad son restricciones que permiten identificar mejor las estimaciones indeterminadas que resultan de las ecuaciones normales. Por ello exigimos que las filas de $\mathbf{A}$ sean linealmente dependientes de las filas de $\mathbf{X}$ y que el rango de la matriz $\mathbf{A}$ $q \times m$ sea q. De hecho, cualquier ecuación $\mathbf{a}_i'\boldsymbol{\beta} = 0$ para la que $\mathbf{a}_i'$ sea linealmente independiente de las filas de $\mathbf{X}$ puede ignorarse y la hipótesis contrastable estará formada por el resto de las ecuaciones.

Una caracterización para saber si una hipótesis lineal es contrastable es

$$\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X} = \mathbf{A}$$

Este resultado es una generalización del que se ha demostrado en la página 46 para una función paramétrica estimable (ver ejercicio 5.3).

5.2. El modelo lineal de la hipótesis

El modelo lineal inicial $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, que se supone válido, constituye la hipótesis alternativa

$$H_1 : \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \qquad \text{rg}\, \mathbf{X} = r$$

Por otra parte, el modelo lineal junto con la restricción lineal contrastable forman la hipótesis nula

$$H_0 : \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad \mathbf{A}\boldsymbol{\beta} = \mathbf{0} \qquad \text{rg}\, \mathbf{A} = q$$

Pero esta restricción lineal transforma los parámetros $\boldsymbol{\beta}$ y la matriz de diseño $\mathbf{X}$ en un nuevo modelo llamado el modelo lineal de la hipótesis

$$H_0 : \mathbf{Y} = \widetilde{\mathbf{X}}\boldsymbol{\theta} + \boldsymbol{\epsilon} \qquad \text{rg}\, \widetilde{\mathbf{X}} = r - q > 0$$

que es otra forma de plantear la hipótesis nula.
Existen varios procedimientos para estimar $\boldsymbol{\beta}$ o $\boldsymbol{\theta}$ bajo la hipótesis nula y calcular la suma de cuadrados residual.

Método 1

Si la hipótesis es contrastable, las filas de $\mathbf{A}$ son combinación lineal de las filas de $\mathbf{X}$. El subespacio $\langle \mathbf{A}' \rangle$ generado por las filas de $\mathbf{A}$ está incluido en el subespacio $\langle \mathbf{X}' \rangle$ generado por las filas de $\mathbf{X}$. Existe entonces una base ortogonal

$$\mathbf{v}_1, \ldots, \mathbf{v}_q, \mathbf{v}_{q+1}, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1} \ldots, \mathbf{v}_m$$

tal que

$$\langle \mathbf{A}' \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_q \rangle \subset \langle \mathbf{v}_1, \ldots, \mathbf{v}_q, \mathbf{v}_{q+1}, \ldots, \mathbf{v}_r \rangle = \langle \mathbf{X}' \rangle \subset \mathbb{R}^m$$

Sea entonces $\mathbf{C}$ una matriz $m \times r'$, con $r' = r - q$, construida tomando los vectores columna $\mathbf{v}_{q+1}, \ldots, \mathbf{v}_r$

$$\mathbf{C} = (\mathbf{v}_{q+1}, \ldots, \mathbf{v}_r)$$

y definamos el vector paramétrico $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{r'})'$ tal que

$$\boldsymbol{\beta} = \mathbf{C}\boldsymbol{\theta}$$

Los parámetros $\boldsymbol{\theta}$ constituyen la reparametrización inducida por la hipótesis H_0, pues

$$\mathbf{A}\boldsymbol{\beta} = \mathbf{A}\mathbf{C}\boldsymbol{\theta} = \mathbf{0}\boldsymbol{\theta} = \mathbf{0}$$

El modelo $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ bajo la restricción $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$, se convierte en

$$E(\mathbf{Y}) = \mathbf{X}\mathbf{C}\boldsymbol{\theta} = \widetilde{\mathbf{X}}\boldsymbol{\theta}$$

y la matriz de diseño se transforma en

$$\widetilde{\mathbf{X}} = \mathbf{X}\mathbf{C}$$

relación también válida para la matriz de diseño reducida

$$\widetilde{\mathbf{X}}_R = \mathbf{X}_R\mathbf{C}$$

La estimación MC de los parámetros $\boldsymbol{\theta}$ es

$$\widehat{\boldsymbol{\theta}} = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\mathbf{Y}$$

La suma de cuadrados residual bajo la restricción $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$ es

$$\begin{aligned} \text{SCR}_H &= \min_{\mathbf{A}\boldsymbol{\beta}=\mathbf{0}} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = (\mathbf{Y} - \widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}})'(\mathbf{Y} - \widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}) \\ &= \mathbf{Y}'\mathbf{Y} - \widehat{\boldsymbol{\theta}}'\widetilde{\mathbf{X}}'\mathbf{Y} \end{aligned}$$

Método 2

Introduzcamos q multiplicadores de Lagrange

$$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_q)'$$

uno para cada restricción lineal. El mínimo restringido de $(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ se halla igualando a cero las derivadas respecto a cada β_i de

$$\sum_{i=1}^{n} (y_i - x_{i1}\beta_1 - \cdots - x_{im}\beta_m)^2 + \sum_{i=1}^{q} \lambda_i (a_{i1}\beta_1 + \cdots + a_{im}\beta_m)$$

En notación matricial, donde ahora $\mathbf{X}$ es la matriz ampliada, escribiremos

$$f(\boldsymbol{\beta}, \boldsymbol{\lambda}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) + (\boldsymbol{\beta}'\mathbf{A}')\boldsymbol{\lambda}$$

$$\partial f / \partial \boldsymbol{\beta} = -2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + \mathbf{A}'\boldsymbol{\lambda} = \mathbf{0}$$

$$\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{Y} - \frac{1}{2}\mathbf{A}'\boldsymbol{\lambda} \tag{5.1}$$

La solución es

$$\widehat{\boldsymbol{\beta}}_H = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y} - \frac{1}{2}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'\widehat{\boldsymbol{\lambda}}_H$$

$$= \widehat{\boldsymbol{\beta}} - \frac{1}{2}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'\widehat{\boldsymbol{\lambda}}_H$$

y como $\mathbf{A}\widehat{\boldsymbol{\beta}}_H = \mathbf{0}$, resulta

$$\mathbf{0} = \mathbf{A}\widehat{\boldsymbol{\beta}} - \frac{1}{2}\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'\widehat{\boldsymbol{\lambda}}_H$$

La matriz $\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'$ posee inversa, puesto que es de rango q, así

$$\frac{1}{2}\widehat{\boldsymbol{\lambda}}_H = (\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}})$$

y finalmente tenemos que la estimación MC restringida es

$$\widehat{\boldsymbol{\beta}}_H = \widehat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-}\mathbf{A}'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}')^{-1}\mathbf{A}\widehat{\boldsymbol{\beta}} \tag{5.2}$$

La suma de cuadrados residual es

$$\mathrm{SCR}_H = (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_H)'(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_H)$$

Hemos visto (teorema 2.5.1) que la forma canónica de la suma de cuadrados residual bajo el modelo sin restricciones es

$$\mathrm{SCR} = z_{r+1}^2 + \cdots + z_n^2$$

La hipótesis $H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{0}$, que implica $\widetilde{\mathbf{X}} = \mathbf{X}\mathbf{C}$, significa que las columnas de $\widetilde{\mathbf{X}}$ son combinación lineal de las de $\mathbf{X}$. Luego los subespacios generados por dichas columnas verifican

$$\langle \widetilde{\mathbf{X}} \rangle \subset \langle \mathbf{X} \rangle \subset \mathbb{R}^n \tag{5.3}$$

Podemos entonces construir una base ortogonal

$$\mathbf{u}_1, \ldots, \mathbf{u}_{r'}, \mathbf{u}_{r'+1}, \ldots, \mathbf{u}_r, \mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$$

tal que

$$\langle \widetilde{\mathbf{X}} \rangle = \langle \mathbf{u}_1, \ldots, \mathbf{u}_{r'} \rangle \subset \langle \mathbf{X} \rangle = \langle \mathbf{u}_1, \ldots, \mathbf{u}_r \rangle$$

Entonces, si se cumple la hipótesis, por idéntico razonamiento al seguido en el teorema 2.5.1 tendremos que la forma canónica de la suma de cuadrados residual bajo el modelo H_0 es

$$\mathrm{SCR}_H = z_{r'+1}^2 + \cdots + z_n^2$$

Además, siempre se verificará que $\mathrm{SCR}_H > \mathrm{SCR}$ pues

$$\mathrm{SCR}_H - \mathrm{SCR} = \sum_{r'+1}^{r} z_i^2$$

Ejemplo 5.2.1

Consideremos el siguiente modelo lineal normal

$$\begin{aligned} y_1 &= \beta_1 + \beta_2 + \epsilon_1 \\ y_2 &= 2\beta_2 + \epsilon_2 \\ y_3 &= -\beta_1 + \beta_2 + \epsilon_3 \end{aligned}$$

y la hipótesis lineal

$$H_0 : \beta_1 = 2\beta_2$$

Las matrices de diseño y de la hipótesis son

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ -1 & 1 \end{pmatrix} \qquad \mathbf{A} = (1 \quad -2) \qquad \mathrm{rg}\,\mathbf{X} = 2 \qquad \mathrm{rg}\,\mathbf{A} = 1$$

Como $\mathbf{A}$ *es combinación lineal de las filas de* $\mathbf{X}$*,* H_0 *es una hipótesis contrastable. Además, en este caso particular el rango de la matriz de diseño es máximo, de modo que toda hipótesis lineal es contrastable.*

Con unos sencillos cálculos, tenemos:

Ecuaciones normales

$$2\beta_1 + 0\beta_2 = y_1 - y_3 \qquad 0\beta_1 + 6\beta_2 = y_1 + 2y_2 + y_3$$

Estimaciones MC

$$\widehat{\beta}_1 = (y_1 - y_3)/2 \qquad \widehat{\beta}_2 = (y_1 + 2y_2 + y_3)/6$$

Suma de cuadrados residual

$$\mathrm{SCR} = y_1^2 + y_2^2 + y_3^2 - 2\widehat{\beta}_1^2 - 6\widehat{\beta}_2^2$$

Si consideramos los vectores columna

$$\mathbf{v}_1 = (1, -2)' \qquad \mathbf{v}_2 = (2, 1)'$$

que constituyen una base ortogonal de $\mathbb{R}^2$*, se verifica*

$$\langle \mathbf{A}' \rangle = \langle \mathbf{v}_1 \rangle \subset \langle \mathbf{X}' \rangle = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$$

Podemos entonces tomar la matriz

$$\mathbf{C} = (2, 1)'$$

que verifica $\mathbf{AC} = \mathbf{0}$*. La reparametrización* $\boldsymbol{\beta} = \mathbf{C}\boldsymbol{\theta}$ *es*

$$\beta_1 = 2\theta \qquad \beta_2 = \theta$$

El modelo bajo la hipótesis es ahora

$$\begin{aligned} y_1 &= 3\theta + \epsilon_1 \\ y_2 &= 2\theta + \epsilon_2 \\ y_3 &= -\theta + \epsilon_3 \end{aligned}$$

Finalmente

$$\widehat{\theta} = (3y_1 + 2y_2 - y_3)/14$$
$$\mathrm{SCR}_H = y_1^2 + y_2^2 + y_3^2 - 14\widehat{\theta}^2$$

5.3. Teorema fundamental del Análisis de la Varianza

En esta sección vamos a deducir el test F que nos permite decidir sobre la aceptación de una hipótesis lineal contrastable.

Teorema 5.3.1

Sea $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ un modelo lineal normal, de manera que $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$. Consideremos una hipótesis lineal contrastable

$$H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{0} \qquad \text{rango } \mathbf{A} = q$$

entonces, los estadísticos

$$\begin{aligned} \mathrm{SCR} &= (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}) \\ \mathrm{SCR}_H &= (\mathbf{Y} - \widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}})'(\mathbf{Y} - \widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}) \end{aligned}$$

verifican:

(i) $\mathrm{SCR}/\sigma^2 \sim \chi^2_{n-r}$

(ii) Si H_0 es cierta

$$\begin{aligned} \mathrm{SCR}_H/\sigma^2 &\sim \chi^2_{n-r'} \qquad (r' = r - q) \\ (\mathrm{SCR}_H - \mathrm{SCR})/\sigma^2 &\sim \chi^2_q \end{aligned}$$

(iii) Si H_0 es cierta, los estadísticos $\mathrm{SCR}_H - \mathrm{SCR}$ y SCR son estocásticamente independientes.

(iv) Si H_0 es cierta, el estadístico

$$F = \frac{(\mathrm{SCR}_H - \mathrm{SCR})/q}{\mathrm{SCR}/(n-r)} \tag{5.4}$$

sigue la distribución F de Fisher-Snedecor con q y $n-r$ grados de libertad.

Demostración:

(i) Aunque este resultado ya se ha establecido en el teorema 3.4.2, nos interesa ahora su demostración explícita. En el teorema 2.5.1 se ha visto que

$$\mathrm{SCR} = z^2_{r+1} + \cdots + z^2_n$$

donde las z_i son normales, independientes y además $E(z_i) = 0$, $\mathrm{var}(z_i) = \sigma^2$. Luego SCR/σ^2 es suma de los cuadrados de $n-r$ variables $N(0,1)$ independientes.

(ii) La forma canónica de la suma de cuadrados residual bajo la restricción $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$ es

$$\mathrm{SCR}_H = z^2_{r'+1} + \cdots + z^2_n$$

luego análogamente tenemos que $\mathrm{SCR}_H/\sigma^2 \sim \chi^2_{n-r'}$, donde $r' = r - q$. Además

$$\mathrm{SCR}_H - \mathrm{SCR} = z^2_{r'+1} + \cdots + z^2_r$$

es también una suma de cuadrados en las mismas condiciones.

(iii) Las variables $z_{r'+1}, \ldots, z_n$ son normales e independientes. $\mathrm{SCR}_H - \mathrm{SCR}$ depende de las q primeras, mientras que SCR depende de las $n-r$ últimas y no hay términos comunes. Luego son estocásticamente independientes.

(iv) Es una consecuencia evidente de los apartados anteriores de este teorema. Si H_0 es cierta, el estadístico

$$F = \frac{[(\mathrm{SCR}_H - \mathrm{SCR})/\sigma^2]/q}{(\mathrm{SCR}/\sigma^2)/(n-r)} = \frac{(\mathrm{SCR}_H - \mathrm{SCR})/q}{\mathrm{SCR}/(n-r)}$$

sigue la distribución F de Fisher-Snedecor con q y $n-r$ grados de libertad.

■

Obsérvese que F no depende del parámetro desconocido σ^2 y se puede calcular exclusivamente en función de las observaciones $\mathbf{Y}$.
La expresión de SCR es

$$\mathrm{SCR} = \mathbf{Y}'(\mathbf{I}-\mathbf{P})\mathbf{Y} = \mathbf{Y}'\mathbf{Y} - \widehat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y}$$

Veamos que, del mismo modo, la expresión de SCR_H es

$$\mathrm{SCR}_H = \mathbf{Y}'\mathbf{Y} - \widehat{\boldsymbol{\beta}}_H'\mathbf{X}'\mathbf{Y}$$

donde $\widehat{\boldsymbol{\beta}}_H$ es la estimación MC de $\boldsymbol{\beta}$ restringida a $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$.
En efecto,

$$\mathrm{SCR}_H = (\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}_H)'(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}_H) = \mathbf{Y}'\mathbf{Y} - 2\mathbf{Y}'\mathbf{X}\widehat{\boldsymbol{\beta}}_H + \widehat{\boldsymbol{\beta}}_H'\mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}}_H$$

Además (ver página 71), se verifica

$$\mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}}_H = \mathbf{X}'\mathbf{Y} - \frac{1}{2}\mathbf{A}'\widehat{\boldsymbol{\lambda}}_H$$

luego

$$\begin{aligned}\mathrm{SCR}_H &= \mathbf{Y}'\mathbf{Y} - 2\mathbf{Y}'\mathbf{X}\widehat{\boldsymbol{\beta}}_H + \widehat{\boldsymbol{\beta}}_H'(\mathbf{X}'\mathbf{Y} - \frac{1}{2}\mathbf{A}'\widehat{\boldsymbol{\lambda}}_H)\\ &= \mathbf{Y}'\mathbf{Y} - 2\mathbf{Y}'\mathbf{X}\widehat{\boldsymbol{\beta}}_H + \mathbf{Y}'\mathbf{X}\widehat{\boldsymbol{\beta}}_H - \frac{1}{2}\widehat{\boldsymbol{\beta}}_H'\mathbf{A}'\widehat{\boldsymbol{\lambda}}_H\end{aligned}$$

Pero como $\mathbf{A}\widehat{\boldsymbol{\beta}}_H = \mathbf{0}$, nos queda

$$\mathrm{SCR}_H = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\widehat{\boldsymbol{\beta}}_H$$

Calculemos ahora $\mathrm{SCR}_H - \mathrm{SCR}$. Considerando 5.2 tenemos

$$\widehat{\boldsymbol{\beta}}' - \widehat{\boldsymbol{\beta}}_H' = (\mathbf{A}\widehat{\boldsymbol{\beta}})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{A}')^{-1}\mathbf{A}(\mathbf{X}'\mathbf{X})^-$$

luego

$$\begin{aligned}\mathrm{SCR}_H - \mathrm{SCR} &= (\widehat{\boldsymbol{\beta}}' - \widehat{\boldsymbol{\beta}}_H')\mathbf{X}'\mathbf{Y}\\ &= (\mathbf{A}\widehat{\boldsymbol{\beta}})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{A}')^{-1}\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}\\ &= (\mathbf{A}\widehat{\boldsymbol{\beta}})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}})\end{aligned} \tag{5.5}$$

El estadístico F puede escribirse entonces

$$F = \frac{(\mathbf{A}\widehat{\boldsymbol{\beta}})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}})}{q\,\widehat{\sigma}^2} \tag{5.6}$$

donde $\widehat{\sigma}^2 = \mathrm{SCR}/(n-r)$.
Cuando $q > 2$ es mejor obtener SCR y SCR_H directamente por minimización de $\boldsymbol{\epsilon}'\boldsymbol{\epsilon}$ sin restricciones y con restricciones, respectivamente. Sin embargo, si $q \leq 2$ se puede utilizar la fórmula 5.6, ya que la matriz a invertir $\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{A}'$ es sólo de orden uno o dos.

Obsérvese que si $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$ es cierta, entonces $\widehat{\mathbf{A}\boldsymbol{\beta}} \approx \mathbf{0}$. Luego es probable que F no sea significativa. Cuando sea posible, también se puede utilizar la matriz de diseño reducida $\mathbf{X}_R$, junto con las matrices $\mathbf{D}$ y $\overline{\mathbf{Y}}$. Las expresiones son entonces

$$\text{SCR} = \mathbf{Y}'\mathbf{Y} - \overline{\mathbf{Y}}'\mathbf{D}\mathbf{X}_R(\mathbf{X}_R'\mathbf{D}\mathbf{X}_R)^{-}\mathbf{X}_R'\mathbf{D}\overline{\mathbf{Y}}$$
$$\text{SCR}_H - \text{SCR} = (\widehat{\mathbf{A}\boldsymbol{\beta}})'(\mathbf{A}(\mathbf{X}_R'\mathbf{D}\mathbf{X}_R)^{-}\mathbf{A}')^{-}(\widehat{\mathbf{A}\boldsymbol{\beta}})$$

El cálculo de ambas cantidades se suele expresar en forma de tabla general del análisis de la varianza (ver tabla 5.1).

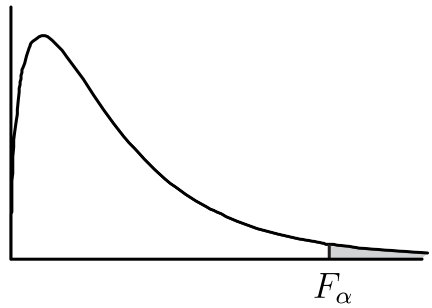
	grados de libertad	suma de cuadrados	cuadrados medios	cociente
Desviación hipótesis	q	$\text{SCR}_H - \text{SCR}$	$(\text{SCR}_H - \text{SCR})/q$	F
Residuo	$n - r$	SCR	$\text{SCR}/(n - r)$	

Tabla 5.1: Tabla general del análisis de la varianza

Criterio de decisión

Si $F > F_\alpha$ se rechaza H_0; si $F \leq F_\alpha$ se acepta H_0.

Donde, para un nivel de significación α, F_α se elige de forma que $P(F_{q,n-r} > F_\alpha) = \alpha$.



Del teorema 5.3.1 deducimos que, si H_0 es cierta, entonces

$$E[(\text{SCR}_H - \text{SCR})/q] = \sigma^2$$

Luego $(\text{SCR}_H - \text{SCR})/q$ y $\text{SCR}/(n - r)$ son dos estimaciones independientes de la varianza σ^2. El test F nos indica hasta que punto coinciden. Un valor grande de F indica que la primera estimación difiere demasiado de la varianza σ^2 y entonces H_0 debe ser rechazada. Se puede demostrar además (ver ejercicio 5.8) que en general

$$E(\text{SCR}_H - \text{SCR}) = q\sigma^2 + (\mathbf{A}\boldsymbol{\beta})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}')^{-}(\mathbf{A}\boldsymbol{\beta}) \tag{5.7}$$

Ejemplo 5.3.1

Para decidir sobre la hipótesis $H_0 : \beta_1 = 2\beta_2$ en el ejemplo 5.2.1 calcularemos

$$F = \frac{(\text{SCR}_H - \text{SCR})/1}{\text{SCR}/(3-2)} = \frac{-14\widehat{\theta}^2 + 2\widehat{\beta}_1^2 + 6\widehat{\beta}_2^2}{y_1^2 + y_2^2 + y_3^2 - 2\widehat{\beta}_1^2 - 6\widehat{\beta}_2^2}$$

Si utilizamos 5.6, se obtiene una expresión más sencilla

$$F = \frac{(\widehat{\beta}_1 - 2\widehat{\beta}_2)^2}{(\text{SCR}/1)(7/6)}$$

En cualquier caso, se decide por la significación en una distribución $F_{1,1}$ con 1 y 1 grados de libertad.

Ejemplo 5.3.2 ***Diseño "cross-over" simplificado***

Supongamos una experiencia clínica en la que se desean comparar dos fármacos **a** *y* **b***, para combatir una determinada enfermedad. El estado de los pacientes se valora mediante una cierta variable cuantitativa Y.*

En el diseño "cross-over" la experiencia se organiza asignando a $N_{\mathbf{a}}$ pacientes el tratamiento **a** *y a $N_{\mathbf{b}}$ pacientes el tratamiento* **b***, en un primer periodo. En un segundo periodo, los que tomaban* **a** *pasan a tomar* **b** *y recíprocamente. En este diseño los datos son de la forma:*

	Grupo 1				***media***	***varianza***
a *(primera vez)*	y_{11}	y_{12}	$\ldots$	$y_{1N_{\mathbf{a}}}$	$\bar{y}_{1\cdot}$	$s_1^2 = \frac{1}{N_{\mathbf{a}}}\sum_{i=1}^{N_{\mathbf{a}}}(y_{1i}-\bar{y}_{1\cdot})^2$
b *(después de* **a***)*	y_{21}	y_{22}	$\ldots$	$y_{2N_{\mathbf{a}}}$	$\bar{y}_{2\cdot}$	$s_2^2 = \frac{1}{N_{\mathbf{a}}}\sum_{i=1}^{N_{\mathbf{a}}}(y_{2i}-\bar{y}_{2\cdot})^2$
	Grupo 2					
b *(primera vez)*	y_{31}	y_{32}	$\ldots$	$y_{3N_{\mathbf{b}}}$	$\bar{y}_{3\cdot}$	$s_3^2 = \frac{1}{N_{\mathbf{b}}}\sum_{i=1}^{N_{\mathbf{b}}}(y_{3i}-\bar{y}_{3\cdot})^2$
a *(después de* **b***)*	y_{41}	y_{42}	$\ldots$	$y_{4N_{\mathbf{b}}}$	$\bar{y}_{4\cdot}$	$s_4^2 = \frac{1}{N_{\mathbf{b}}}\sum_{i=1}^{N_{\mathbf{b}}}(y_{4i}-\bar{y}_{4\cdot})^2$

Indicando

$$\begin{aligned} \mu &= \textit{media general} \\ \alpha &= \textit{efecto fármaco } \mathbf{a} \\ \beta &= \textit{efecto fármaco } \mathbf{b} \\ \gamma &= \textit{efecto recíproco entre } \mathbf{a} \textit{ y } \mathbf{b} \end{aligned}$$

se propone el siguiente modelo:

$$\begin{array}{lll} \mathbf{a}\ \textit{(primera vez)} & y_{1i} = \mu + \alpha + \epsilon_{1i} & i = 1,\ldots,N_{\mathbf{a}} \\ \mathbf{b}\ \textit{(después de } \mathbf{a}\textit{)} & y_{2i} = \mu + \beta + \gamma + \epsilon_{2i} & i = 1,\ldots,N_{\mathbf{a}} \\ \mathbf{b}\ \textit{(primera vez)} & y_{3i} = \mu + \beta + \epsilon_{3i} & i = 1,\ldots,N_{\mathbf{b}} \\ \mathbf{a}\ \textit{(después de } \mathbf{b}\textit{)} & y_{4i} = \mu + \alpha + \gamma + \epsilon_{4i} & i = 1,\ldots,N_{\mathbf{b}} \end{array}$$

Es decir, cuando sólo se ha tomado un fármaco actúa un solo efecto, pero cuando se ha tomado uno después del otro actúa entonces un efecto aditivo γ que recoge la mejoría del enfermo que ya ha tomado el primer medicamento.

Tenemos $k = 4$ condiciones experimentales, que en el "cross-over" simplificado se consideran independientes, y $N_1 = N_2 = N_{\mathbf{a}}$, $N_3 = N_4 = N_{\mathbf{b}}$. El vector de observaciones $\mathbf{Y}$ *y la matriz de diseño reducida* $\mathbf{X}_R$ *son*

$$\mathbf{Y} = (y_{11},\ldots,y_{1N_{\mathbf{a}}},y_{21},\ldots,y_{2N_{\mathbf{a}}},y_{31},\ldots,y_{3N_{\mathbf{b}}},y_{41},\ldots,y_{4N_{\mathbf{b}}})'$$

$$\mathbf{X}_R = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \qquad \text{rg}\ \mathbf{X}_R = 3$$

La hipótesis nula de mayor interés es

$$H_0 : \alpha = \beta \qquad \mathbf{a} \textit{ y } \mathbf{b} \textit{ tienen la misma efectividad}$$

que expresada en forma de hipótesis lineal es

$$H_0 : \begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha \\ \beta \\ \gamma \end{pmatrix} = 0$$

Como el vector $\begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix}$ *es combinación lineal de las filas de* $\mathbf{X}_R$*, se trata de una hipótesis contrastable. Para reparametrizar el diseño bajo* H_0 *tomaremos como matriz ortogonal a* $\mathbf{A}$

$$\mathbf{C} = \begin{pmatrix} 2/3 & 0 \\ 1/3 & 0 \\ 1/3 & 0 \\ 0 & 1 \end{pmatrix}$$

Obsérvese que las columnas de $\mathbf{C}$ *son también combinación lineal de las filas de* $\mathbf{X}_R$.
Al establecer la relación $\boldsymbol{\beta} = \mathbf{C}\boldsymbol{\theta}$ *tendremos*

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$$

siendo $\theta_1 = \mu + \alpha = \mu + \beta$ *y* $\theta_2 = \gamma$.
Es decir, bajo H_0 *el diseño reparametrizado depende de dos parámetros:*

θ_1 : *efecto debido a la medicación (común a* **a** *y* **b** *bajo* H_0*)*
θ_2 : *efecto recíproco entre* **a** *y* **b**

y la nueva matriz de diseño es

$$\widetilde{\mathbf{X}}_R = \mathbf{X}_R\mathbf{C} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$$

siendo $\operatorname{rg} \widetilde{\mathbf{X}}_R = r - t = 3 - 1 = 2$.
Si el diseño es balanceado ($N_\mathbf{a} = N_\mathbf{b}$)*, entonces* $N = 4N_\mathbf{a} = 4N_\mathbf{b}$ *y se puede calcular que*

$$\text{SCR} = \frac{N_\mathbf{a}}{4}(y_{1\cdot} + y_{2\cdot} - y_{3\cdot} - y_{4\cdot})^2 + N_\mathbf{a}\left(\sum_{i=1}^{4} s_i^2\right)$$

con $N - 3$ *grados de libertad*

$$\text{SCR}_H = \frac{N_\mathbf{a}}{4}[(y_{1\cdot} + y_{2\cdot} - y_{3\cdot} - y_{4\cdot})^2 + (y_{1\cdot} - y_{2\cdot} - y_{3\cdot} + y_{4\cdot})^2] + N_\mathbf{a}\left(\sum_{i=1}^{4} s_i^2\right)$$

con $N - 2$ *grados de libertad.*
Luego, si H_0 *es cierta, bajo el modelo lineal normal, el estadístico*

$$F = \frac{(y_{1\cdot} - y_{2\cdot} - y_{3\cdot} + y_{4\cdot})^2}{4\,\text{SCR}} N_\mathbf{a}(4N_\mathbf{a} - 3)$$

sigue la distribución F con 1 y $N - 3$ *g.l..*
La tabla 5.2 contiene los datos de dos grupos de 10 y 10 enfermos reumáticos a los que se valoró la variación del dolor respecto del estado inicial, mediante una escala convencional, con el deseo de comparar dos fármacos antirreumáticos **a** *y* **b***, administrados a lo largo de dos meses.*

Se incluye además la tabla del análisis de la varianza para contrastar H_0.
Con estos datos se han detectado diferencias significativas entre los dos fármacos **a** *y* **b***. Para estimar la eficacia de cada fármaco, pasaremos a considerar las funciones paramétricas*

$$\psi_\mathbf{a} = \mu + \alpha \qquad \psi_\mathbf{b} = \mu + \beta$$

que son ambas estimables.

Grupo 1		Grupo 2	
a (mes 1)	b (mes 2)	b (mes 1)	a (mes 2)
17	17	21	10
34	41	20	24
26	26	11	32
10	3	26	26
19	-6	42	52
17	-4	28	28
8	11	3	27
16	16	3	28
13	16	16	21
11	4	-10	42

Tabla 5.2: Datos de los enfermos reumáticos

	g.l.	suma de cuadrados	cuadrados medios	F
Entre fármacos	1	783.2	783.2	4.71 ($p < 0.05$)
Residuo	37	6147.9	166.2	

Tabla 5.3: Análisis de la varianza para el contraste $H_0 : \alpha = \beta$

Para estimar $\psi_\mathbf{a}, \psi_\mathbf{b}$ hallaremos primeramente "una" estimación MC de los parámetros:

$$\widehat{\mu} = 0 \qquad \widehat{\alpha} = 20.975 \qquad \widehat{\beta} = 12.125$$

Aplicando el teorema de Gauss-Markov, las estimaciones óptimas de $\psi_\mathbf{a}, \psi_\mathbf{b}$ se obtienen sustituyendo parámetros por estimaciones MC, es decir

$$\widehat{\psi_\mathbf{a}} = \widehat{\mu} + \widehat{\alpha} = 20.975 \qquad \widehat{\psi_\mathbf{b}} = \widehat{\mu} + \widehat{\beta} = 12.125$$

Por otra parte, las expresiones en función de las medias y las varianzas mínimas correspondientes son:

$$\begin{aligned} \widehat{\psi_\mathbf{a}} &= 3/4\bar{y}_1 - 1/4\bar{y}_2 + 1/4\bar{y}_3 + 1/4\bar{y}_4 \qquad & \operatorname{var}(\widehat{\psi_\mathbf{a}}) &= 0.075\sigma^2 \\ \widehat{\psi_\mathbf{b}} &= 1/4\bar{y}_1 + 1/4\bar{y}_2 + 3/4\bar{y}_3 - 1/4\bar{y}_4 \qquad & \operatorname{var}(\widehat{\psi_\mathbf{b}}) &= 0.075\sigma^2 \end{aligned}$$

5.3.1. Un contraste más general

Consideremos la hipótesis nula

$$H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{c} \qquad \mathbf{A} \text{ es } q \times m, \text{ rg } \mathbf{A} = q$$

donde $\mathbf{c}$ es un vector columna que lógicamente debe ser combinación lineal de las columnas de $\mathbf{A}$. También suponemos que las filas de $\mathbf{A}$ son combinación lineal de las filas de $\mathbf{X}$, de manera que $\mathbf{A}\boldsymbol{\beta}$ es un conjunto de funciones paramétricas estimables.

Sea $\boldsymbol{\beta}_0$ tal que $\mathbf{A}\boldsymbol{\beta}_0 = \mathbf{c}$ y consideremos $\boldsymbol{\gamma} = \boldsymbol{\beta} - \boldsymbol{\beta}_0$. Entonces, si en el modelo lineal

$$\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0 = \mathbf{X}(\boldsymbol{\beta} - \boldsymbol{\beta}_0) + \boldsymbol{\epsilon}$$

ponemos $\widetilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0$, obtenemos el modelo transformado

$$\widetilde{\mathbf{Y}} = \mathbf{X}\boldsymbol{\gamma} + \boldsymbol{\epsilon} \tag{5.8}$$

y en este modelo la hipótesis planteada adopta la expresión

$$H_0 : \mathbf{A}\boldsymbol{\gamma} = \mathbf{0}$$

La estimación MC del conjunto de funciones paramétricas estimables $\mathbf{A}\boldsymbol{\gamma}$ en este modelo transformado es

$$\begin{aligned}\mathbf{A}\widehat{\boldsymbol{\gamma}} &= \mathbf{BX}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\widetilde{\mathbf{Y}} \\ &= \mathbf{BP}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}_0) = \mathbf{BX}\widehat{\boldsymbol{\beta}} - \mathbf{BX}\boldsymbol{\beta}_0 \\ &= \mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{A}\boldsymbol{\beta}_0 = \mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{c}\end{aligned}$$

En consecuencia, de la ecuación 5.5 se deduce

$$\begin{aligned}\mathrm{SCR}_H - \mathrm{SCR} &= (\mathbf{A}\widehat{\boldsymbol{\gamma}})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\gamma}}) \\ &= (\mathbf{A}\widehat{\boldsymbol{\beta}}-\mathbf{c})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}}-\mathbf{c})\end{aligned}$$

donde $\widehat{\boldsymbol{\beta}}$ es tal que $\mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$. Se verifica también

$$E(\mathrm{SCR}_H - \mathrm{SCR}) = q\,\sigma^2 + (\mathbf{A}\boldsymbol{\beta}-\mathbf{c})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}')^{-1}(\mathbf{A}\boldsymbol{\beta}-\mathbf{c})$$

Finalmente, a partir de la fórmula 5.6 el test para contrastar la hipótesis es

$$F = \frac{(\mathbf{A}\widehat{\boldsymbol{\beta}}-\mathbf{c})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}}-\mathbf{c})/q}{\mathrm{SCR}/(n-r)} \tag{5.9}$$

donde, si es cierta la hipótesis nula, el estadístico F sigue una distribución $F_{q,n-r}$.
En el caso particular $q = 1$, donde la hipótesis es $H_0 : \mathbf{a}'\boldsymbol{\beta} = c$, el test F se puede simplificar en un test t con

$$t = \frac{\mathbf{a}'\widehat{\boldsymbol{\beta}} - c}{(\widehat{\sigma}^2(\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-}\mathbf{a}))^{1/2}} \tag{5.10}$$

que sigue una distribución t_{n-r}, si H_0 es cierta.

Ejemplo 5.3.3

Contraste de medias en poblaciones normales con igual varianza

Sean $u_1, u_2, \ldots, u_{n_1}$ y $v_1, v_2, \ldots, v_{n_2}$ dos muestras aleatorias simples de dos poblaciones normales $N(\mu_1, \sigma^2)$ y $N(\mu_2, \sigma^2)$, respectivamente.
Vamos a contrastar la hipótesis lineal $H_0 : \mu_1 - \mu_2 = d$ con la ayuda de la teoría de los modelos lineales.
Podemos pensar que las observaciones son de la forma

$$\begin{aligned} u_i &= \mu_1 + \epsilon_i & i &= 1,\ldots,n_1 \\ v_j &= \mu_2 + \epsilon_{n_1+j} & j &= 1,\ldots,n_2 \end{aligned}$$

o en notación matricial

$$\begin{pmatrix} u_1 \\ \vdots \\ u_{n_1} \\ v_1 \\ \vdots \\ v_{n_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n_1} \\ \epsilon_{n_1+1} \\ \vdots \\ \epsilon_n \end{pmatrix}$$

donde $n = n_1 + n_2$. Observemos que, gracias a la igualdad de varianzas en las dos poblaciones, se trata de un modelo lineal y se verifican las condiciones de Gauss-Markov.
En este modelo, la matriz de diseño reducida es 2×2 de rango máximo

$$\mathbf{X}_R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad y \qquad \mathbf{D} = \begin{pmatrix} n_1 & 0 \\ 0 & n_2 \end{pmatrix}$$

Así pues, la hipótesis nula es lineal y contrastable

$$H_0 : \mu_1 - \mu_2 = d \qquad \Leftrightarrow \qquad H_0 : \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = d \qquad q = 1$$

Con unos sencillos cálculos se obtiene

$$\widehat{\boldsymbol{\beta}} = (\hat{\mu}_1, \hat{\mu}_2)' = (\mathbf{X}_R'\mathbf{D}\mathbf{X}_R)^{-1}\mathbf{X}_R'\mathbf{D}\overline{\mathbf{Y}} = \overline{\mathbf{Y}} = (\bar{u}, \bar{v})'$$

$$\mathbf{A}\widehat{\boldsymbol{\beta}} = \hat{\mu}_1 - \hat{\mu}_2 = \bar{u} - \bar{v}$$

$$\begin{aligned} \text{SCR} &= \mathbf{Y}'\mathbf{Y} - \overline{\mathbf{Y}}'\mathbf{D}\mathbf{X}_R(\mathbf{X}_R'\mathbf{D}\mathbf{X}_R)^{-}\mathbf{X}_R'\mathbf{D}\overline{\mathbf{Y}} \\ &= \sum_i u_i^2 + \sum_j v_j^2 - n_1\bar{u}^2 - n_2\bar{v}^2 \\ &= \sum_i (u_i - \bar{u})^2 + \sum_j (v_j - \bar{v})^2 \end{aligned}$$

$$\mathbf{A}(\mathbf{X}_R'\mathbf{D}\mathbf{X}_R)^{-1}\mathbf{A}' = \frac{1}{n_1} + \frac{1}{n_2}$$

de modo que

$$F = \frac{(\mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{c})'(\mathbf{A}(\mathbf{X}_R'\mathbf{D}\mathbf{X}_R)^{-1}\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{c})}{q\,\widehat{\sigma}^2} = \frac{(\bar{u} - \bar{v} - d)^2}{\widehat{\sigma}^2(1/n_1 + 1/n_2)}$$

donde $\widehat{\sigma}^2 = \text{SCR}/(n_1 + n_2 - 2)$ *y cuya distribución, bajo* H_0*, es una* F_{1,n_1+n_2-2}.
Pero cuando $q = 1$*, tenemos que* $F_{1,n_1+n_2-2} \equiv t^2_{n_1+n_2-2}$ *y se deduce que el contraste es equivalente al test t usual, en especial el caso* $d = 0$*.*

5.3.2. Test de la razón de verosimilitud

Para simplificar, consideremos un modelo de rango máximo. Bajo la hipótesis de normalidad de las observaciones, ya sabemos (ver pág. 36) que las estimaciones de máxima verosimilitud de los parámetros son

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \qquad \widehat{\sigma}^2_{MV} = \text{SCR}/n$$

y el valor máximo de la función de verosimilitud es

$$L(\widehat{\boldsymbol{\beta}}, \widehat{\sigma}^2_{MV}) = (2\pi\widehat{\sigma}^2_{MV})^{-n/2}e^{-n/2}$$

Del mismo modo, los estimadores de máxima verosimilitud de los parámetros con las restricciones $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ son

$$\widehat{\boldsymbol{\beta}}_H \qquad \widehat{\sigma}^2_H = \text{SCR}_H/n$$

y el valor máximo de la función de verosimilitud, bajo la hipótesis nula, es

$$L(\widehat{\boldsymbol{\beta}}_H, \widehat{\sigma}^2_H) = (2\pi\widehat{\sigma}^2_H)^{-n/2}e^{-n/2}$$

De modo que el estadístico de la razón de verosimilitud es

$$\Lambda = \frac{L(\widehat{\boldsymbol{\beta}}_H, \widehat{\sigma}^2_H)}{L(\widehat{\boldsymbol{\beta}}, \widehat{\sigma}^2_{MV})} = \left[\frac{\widehat{\sigma}^2_{MV}}{\widehat{\sigma}^2_H}\right]^{n/2}$$

Es fácil ver que

$$F = \frac{n - m}{q}(\Lambda^{-2/n} - 1)$$

luego son contrastes equivalentes.

5.4. Cuando el test es significativo

Si el estadístico F para $H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ es significativo, podemos investigar la causa del rechazo de dicha hipótesis. Una posibilidad consiste en contrastar cada una de las restricciones $\mathbf{a}_i'\boldsymbol{\beta} = c_i$, $i = 1, \ldots, q$ por separado, utilizando un test t para ver cual es la responsable.

Hemos visto de varias formas que, bajo la hipótesis lineal $H_i : \mathbf{a}_i'\boldsymbol{\beta} = c_i$, el estadístico t_i verifica

$$t_i = \frac{\mathbf{a}_i'\widehat{\boldsymbol{\beta}} - c_i}{[\widehat{\sigma}^2\mathbf{a}_i'(\mathbf{X}'\mathbf{X})^-\mathbf{a}_i]^{1/2}} \sim t_{n-r}$$

de modo que podemos rechazar $H_i : \mathbf{a}_i'\boldsymbol{\beta} = c_i$ con un nivel de significación α si

$$|t_i| \geq t_{n-r}(\alpha)$$

donde $t_{n-r}(\alpha)$ es el valor de la tabla tal que $P(|t_{n-r}| \geq t_{n-r}(\alpha)) = \alpha$.

También podemos construir intervalos de confianza para cada $\mathbf{a}_i'\boldsymbol{\beta}$

$$\mathbf{a}_i'\widehat{\boldsymbol{\beta}} \pm t_{n-r}(\alpha) \cdot \widehat{\sigma}(\mathbf{a}_i'(\mathbf{X}'\mathbf{X})^-\mathbf{a}_i)^{1/2}$$

Este procedimiento en dos etapas para el contraste de $H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{c}$, es decir, un contraste global F seguido de una serie de test t cuando F es significativo, se conoce con el nombre de MDS[1] o *mínima diferencia significativa*. El valor significativo mínimo es $t_{n-r}(\alpha)$ y la palabra "diferencia" se refiere a que este método se utiliza con frecuencia para comparar parámetros tales como *medias* dos a dos.

Este método es simple y versátil, sin embargo tiene sus debilidades: es posible rechazar H_0 y no rechazar ninguna de las H_i. Este problema, otras dificultades y, en general, otros métodos de inferencia simultánea se estudian de forma más completa en lo que se llama *Métodos de comparación múltiple*.

5.5. Contraste de hipótesis sobre funciones paramétricas estimables

Sea $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_q)' = \mathbf{A}\boldsymbol{\beta}$ un sistema de funciones paramétricas estimables, de modo que las filas de la matriz $\mathbf{A}$ sean linealmente independientes. La distribución F que sigue la expresión 3.3 permite construir diferentes contrastes de hipótesis bajo el modelo lineal normal.

Sea $\mathbf{c} = (c_1, \ldots, c_q)'$ un vector de constantes, con la condición de que $\mathbf{c}$ sea combinación lineal de las columnas de $\mathbf{A}$. Planteamos la hipótesis nula

$$H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{c} \tag{5.11}$$

Para decidir la aceptación de H_0, como una consecuencia de 3.3, podemos utilizar el estadístico

$$F = \frac{(\mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{c})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{c})/q}{\mathrm{SCR}/(n-r)} \tag{5.12}$$

con distribución $F_{q,n-r}$. Pero es evidente que 5.11 es una hipótesis lineal contrastable, de modo que podemos utilizar el test F que resulta ser idéntico al anterior. Es otra forma de demostrar 5.9 y también que

$$\mathrm{SCR}_H - \mathrm{SCR} = (\mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{c})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{A}')^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{c})$$

Además, podemos plantear otras hipótesis sobre las funciones paramétricas estimables $\boldsymbol{\psi}$, siempre que sean lineales. Por ejemplo, consideremos ahora la hipótesis lineal planteada sobre las q funciones linealmente independientes

$$H_0 : \psi_1 = \psi_2 = \cdots = \psi_q \tag{5.13}$$

1. en inglés: LSD o *least significant difference*

es decir, bajo H_0 las q funciones son iguales. Si consideramos las nuevas funciones

$$\widetilde{\psi}_i = \psi_1 - \psi_{i+1} \qquad i = 1, \ldots, q-1$$

entonces 5.13 se reduce a 5.11 tomando $\widetilde{\boldsymbol{\psi}} = (\widetilde{\psi}_1, \ldots, \widetilde{\psi}_{q-1})'$, $\mathbf{c} = \mathbf{0}$ y sustituyendo q por $q-1$. Dicho de otra manera, sea la matriz

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{q1} & a_{q2} & \ldots & a_{qm} \end{pmatrix}$$

Entonces 5.13 es equivalente a la hipótesis lineal

$$H_0 : \mathbf{A}^* \boldsymbol{\beta} = \mathbf{0}$$

tomando como matriz de hipótesis

$$\mathbf{A}^* = \begin{pmatrix} a_{11} - a_{21} & a_{12} - a_{22} & \ldots & a_{1m} - a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{11} - a_{q1} & a_{12} - a_{q2} & \ldots & a_{1m} - a_{qm} \end{pmatrix}$$

Luego podemos utilizar el estadístico F de 5.6, con $\mathbf{A}^*$ y $q-1$, que bajo H_0 tiene distribución $F_{q-1,n-r}$, para decidir si 5.13 debe ser aceptada.

5.6. Elección entre dos modelos lineales

5.6.1. Sobre los modelos

Para la estimación en el modelo lineal

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \qquad E(\boldsymbol{\epsilon}) = \mathbf{0}, \operatorname{var}(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}$$

hemos establecido (ver pág. 32) que el punto crucial es la utilización de la matriz $\mathbf{P}$, proyección ortogonal sobre el espacio de las estimaciones $\Omega = \langle \mathbf{X} \rangle$. Así, dos modelos son iguales si tienen el mismo espacio de las estimaciones. Dos de estos modelos darán las mismas predicciones y el mismo estimador de σ^2.

Sean $\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\epsilon}_1$ y $\mathbf{Y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}_2$ dos modelos lineales tales que $\langle \mathbf{X}_1 \rangle = \langle \mathbf{X}_2 \rangle$. La matriz proyección no depende de $\mathbf{X}_1$ o $\mathbf{X}_2$ sino sólo de $\Omega (= \langle \mathbf{X}_1 \rangle = \langle \mathbf{X}_2 \rangle)$. La estimación de σ^2 es la misma $\widehat{\sigma}^2 = \mathrm{SCR}/(n-r)$ y las predicciones también

$$\widehat{\mathbf{Y}} = \mathbf{P}\mathbf{Y} = \mathbf{X}_1\widehat{\boldsymbol{\beta}}_1 = \mathbf{X}_2\widehat{\boldsymbol{\beta}}_2$$

En cuanto a las funciones paramétricas estimables, hemos visto que la estimabilidad se restringe a las combinaciones lineales de las filas $\mathbf{X}_1$, es decir, $\mathbf{a}_1'\boldsymbol{\beta}_1$ es estimable si se escribe como $\mathbf{b}'\mathbf{X}_1\boldsymbol{\beta}_1$. Pero $\mathbf{X}_1\boldsymbol{\beta}_1$ pertenece a Ω de forma que $\mathbf{X}_1\boldsymbol{\beta}_1 = \mathbf{X}_2\boldsymbol{\beta}_2$ para algún $\boldsymbol{\beta}_2$ y así

$$\mathbf{a}_1'\boldsymbol{\beta}_1 = \mathbf{b}'\mathbf{X}_1\boldsymbol{\beta}_1 = \mathbf{b}'\mathbf{X}_2\boldsymbol{\beta}_2 = \mathbf{a}_2'\boldsymbol{\beta}_2$$

Las funciones paramétricas estimables son las mismas pero están escritas con diferentes parámetros. Su estimador $\mathbf{b}'\mathbf{P}\mathbf{Y}$ también es único.

Ejemplo 5.6.1

El ANOVA de un factor se puede escribir de dos formas:

$$\begin{array}{ll} y_{ij} = \mu + \alpha_i + \epsilon_{ij} & i = 1,\ldots,I,\ j = 1,\ldots,n_i \\ y_{ij} = \mu_i + \epsilon_{ij} & i = 1,\ldots,I,\ j = 1,\ldots,n_i \end{array}$$

pero son equivalentes puesto que $\langle \mathbf{X}_1 \rangle = \langle \mathbf{X}_2 \rangle$.
En este modelo las relaciones entre los dos conjuntos de parámetros son sencillas

$$\mu_i = \mu + \alpha_i \qquad \mu_1 - \mu_2 = \alpha_1 - \alpha_2 \qquad etc.$$

Ejemplo 5.6.2

La regresión lineal simple admite dos modelos:

$$\begin{array}{ll} y_i = \beta_0 + \beta_1 x_i + \epsilon_i & i = 1,\ldots,n \\ y_i = \gamma_0 + \gamma_1 (x_i - \bar{x}) + \epsilon_i & i = 1,\ldots,n \end{array}$$

pero son equivalentes ya que

$$\begin{aligned} \gamma_0 &= \beta_0 + \beta_1 \bar{x} \\ \gamma_1 &= \beta_1 \end{aligned}$$

En resumen, en un modelo lineal $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ la esencia es el subespacio $\Omega = \langle \mathbf{X} \rangle$. Si conservamos Ω, podemos cambiar $\mathbf{X}$ a nuestra conveniencia.

5.6.2. Contraste de modelos

El contraste de hipótesis en modelos lineales se reduce esencialmente a restringir el espacio de las estimaciones.
Si partimos de un modelo que sabemos o suponemos válido

$$\text{Modelo inicial:} \qquad \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \qquad \text{rg}\, \mathbf{X} = r$$

debemos intentar reducir este modelo, es decir, ver si algún modelo más simple se ajusta aceptablemente a los datos, como

$$\text{Modelo restringido:} \qquad \mathbf{Y} = \widetilde{\mathbf{X}}\boldsymbol{\theta} + \boldsymbol{\epsilon} \qquad \text{rg}\, \widetilde{\mathbf{X}} = \widetilde{r}$$

Dado que la esencia de un modelo está en el subespacio generado por las columnas de la matriz de diseño o espacio de las estimaciones, es absolutamente necesario que el modelo restringido verifique

$$\Omega_0 = \langle \widetilde{\mathbf{X}} \rangle \subset \langle \mathbf{X} \rangle = \Omega$$

Sólo en este caso se puede plantear la elección entre dos modelos alternativos como un contraste de hipótesis

$$\begin{array}{l} H_0 : \mathbf{Y} = \widetilde{\mathbf{X}}\boldsymbol{\theta} + \boldsymbol{\epsilon} \\ H_1 : \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \end{array} \Leftrightarrow \begin{array}{l} H_0 : E(\mathbf{Y}) \in \Omega_0 = \langle \widetilde{\mathbf{X}} \rangle \\ H_1 : E(\mathbf{Y}) \in \Omega = \langle \mathbf{X} \rangle \end{array} \tag{5.14}$$

donde $E(\mathbf{Y}) = \widetilde{\mathbf{X}}\boldsymbol{\theta}$ y $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$, respectivamente.

Sean $\mathbf{P}_\Omega$ y $\mathbf{P}_{\Omega_0}$ las proyecciones ortogonales sobre $\Omega = \langle \mathbf{X} \rangle$ y $\Omega_0 = \langle \widetilde{\mathbf{X}} \rangle$ respectivamente. Bajo el modelo inicial el estimador de $E(\mathbf{Y})$ es $\mathbf{P}_\Omega \mathbf{Y}$, mientras que bajo el modelo restringido el estimador es $\mathbf{P}_{\Omega_0}\mathbf{Y}$. Si la hipótesis H_0 es cierta, ambas estimaciones deben estar próximas.

Teorema 5.6.1

La condición necesaria y suficiente para que 5.14 sea contrastable es que se verifique

$$\Omega_0 = \langle \widetilde{\mathbf{X}} \rangle \subset \langle \mathbf{X} \rangle = \Omega \tag{5.15}$$

El test F se basa entonces en el estadístico

$$F = \frac{(\mathrm{SCR}_H - \mathrm{SCR})/(r - \widetilde{r})}{\mathrm{SCR}/(n-r)}$$

cuya distribución, bajo H_0, es $F_{r-\widetilde{r},n-r}$ y donde

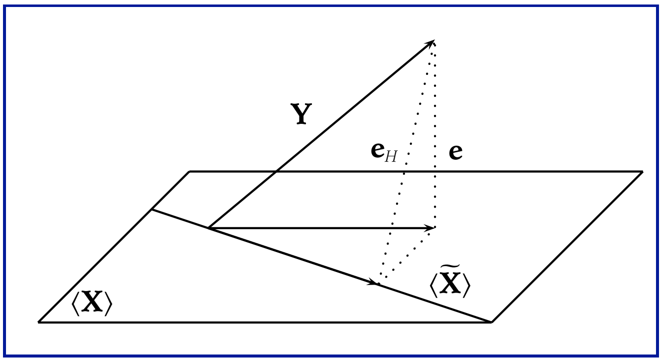
$$\mathrm{SCR}_H = \mathbf{Y}'(\mathbf{I} - \mathbf{P}_{\Omega_0})\mathbf{Y} \qquad \mathrm{SCR} = \mathbf{Y}'(\mathbf{I} - \mathbf{P}_{\Omega})\mathbf{Y}$$

Demostración:

La expresión 5.15 implica la relación $\widetilde{\mathbf{X}} = \mathbf{XC}$ para una cierta matriz $\mathbf{C}$. Entonces H_0 significa formular una hipótesis lineal contrastable al modelo $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$, que lo reduce a $E(\mathbf{Y}) = \widetilde{\mathbf{X}}\boldsymbol{\theta}$. El resto es consecuencia del *Método 1* explicado en la sección 5.2 y el teorema 5.3.1. ■

Observemos que si $\Omega_0 \not\subseteq \Omega$, entonces estamos ante modelos de naturaleza diferente. No podemos decidir entre ambos modelos mediante ningún criterio estadístico conocido. Si se verifica $\Omega_0 = \Omega$, entonces tenemos dos versiones paramétricas del mismo modelo, pudiendo pasar del uno al otro por una reparametrización. Un modelo $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ determina el espacio $\Omega = \langle \mathbf{X} \rangle$, y recíprocamente el espacio Ω determina el modelo (salvo reparametrizaciones que no disminuyan el rango).

Como ya hemos visto, la interpretación geométrica de la solución al modelo lineal $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ es considerar la proyección del vector $\mathbf{Y}$ sobre el subespacio $\Omega = \langle \mathbf{X} \rangle$ de $\mathbb{R}^n$. La relación 5.15 indica que las columnas de $\widetilde{\mathbf{X}}$ generan un subespacio de $\langle \mathbf{X} \rangle$. Entonces SCR y SCR_H son distancias de la observación $\mathbf{Y}$ a los subespacios $\langle \mathbf{X} \rangle$ y $\langle \widetilde{\mathbf{X}} \rangle$, respectivamente. El test F nos dice hasta que punto la diferencia $\mathrm{SCR}_H - \mathrm{SCR}$ es pequeña (comparada con SCR) para poder afirmar que el modelo se ajusta al subespacio $\langle \widetilde{\mathbf{X}} \rangle$ en lugar de $\langle \mathbf{X} \rangle$ (ver figura).



La longitud al cuadrado de la diferencia $\mathbf{P}_\Omega\mathbf{Y} - \mathbf{P}_{\Omega_0}\mathbf{Y}$ es

$$((\mathbf{P}_\Omega - \mathbf{P}_{\Omega_0})\mathbf{Y})'((\mathbf{P}_\Omega - \mathbf{P}_{\Omega_0})\mathbf{Y}) = \mathbf{Y}'(\mathbf{P}_\Omega - \mathbf{P}_{\Omega_0})\mathbf{Y}$$

ya que $\mathbf{P}_\Omega - \mathbf{P}_{\Omega_0} = \mathbf{P}_{\Omega_0^\perp \cap \Omega}$ es una matriz proyección (ver Apéndice). Pero además

$$\mathbf{Y}'(\mathbf{P}_\Omega - \mathbf{P}_{\Omega_0})\mathbf{Y} = \mathbf{Y}'(\mathbf{I} - \mathbf{P}_{\Omega_0})\mathbf{Y} - \mathbf{Y}'(\mathbf{I} - \mathbf{P}_\Omega)\mathbf{Y} = \mathrm{SCR}_H - \mathrm{SCR}$$

Cuando la hipótesis nula se plantea en términos de un grupo de funciones paramétricas estimables del tipo $H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{0}$, sabemos que existe una matriz $\mathbf{B} = \mathbf{A}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ tal que $\mathbf{A} = \mathbf{BX}$. De modo que

$$\mathbf{0} = \mathbf{A}\boldsymbol{\beta} = \mathbf{BX}\boldsymbol{\beta} = \mathbf{B}E(\mathbf{Y}) \Leftrightarrow E(\mathbf{Y}) \in \ker(\mathbf{B})$$

y el subespacio que define la hipótesis nula es $\Omega_0 = \ker(\mathbf{B}) \cap \Omega$. En este caso se puede demostrar (ver Apéndice) que $\Omega_0^\perp \cap \Omega = \langle \mathbf{P}_\Omega\mathbf{B}' \rangle$ y reencontrar así el test 5.6.

Ejemplo 5.6.3

Consideremos de nuevo el diseño cross-over explicado en el ejemplo 5.3.2. Supongamos ahora que la influencia γ de un fármaco sobre el otro no es recíproca. El efecto aditivo no es necesariamente el mismo cuando se administra **a** *después de* **b**, *que cuando se administra* **b** *después de* **a**. *Entonces debemos introducir los parámetros*

$$\begin{aligned} \gamma_1 &: \text{ influencia de } \mathbf{a} \text{ sobre } \mathbf{b} \\ \gamma_2 &: \text{ influencia de } \mathbf{b} \text{ sobre } \mathbf{a} \end{aligned}$$

y admitir que la matriz de diseño reducida, para los parámetros $\mu, \alpha, \beta, \gamma_1, \gamma_2$ es

$$\mathbf{X}_R = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} \qquad \text{rg } \mathbf{X}_R = 4$$

que representa una alternativa a la propuesta inicialmente para los parámetros $\mu, \alpha, \beta, \gamma$

$$\widetilde{\mathbf{X}}_R = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \qquad \text{rg } \widetilde{\mathbf{X}}_R = 3$$

Es fácil ver que se verifica 5.15. El análisis de la varianza para decidir entre $\widetilde{\mathbf{X}}_R$ y $\mathbf{X}_R$, sobre los datos de la tabla 5.2, se encuentra en la tabla 5.4. Como F no es significativo se admite como válido el modelo más simple representado por $\widetilde{\mathbf{X}}_R$.

	grados de libertad	suma de cuadrados	cuadrados medios	*F*
Desviación hipótesis	1	600.6	600.6	3.898
Residuo	36	5547.3	154.1	

Tabla 5.4: Análisis de la varianza para contrastar dos modelos de cross-over

5.7. Ejemplos con R

En esta sección vamos a ver como se contrastan las hipótesis que hemos planteado en el ejemplo 5.3.2 sobre el diseño cross-over simplificado.
En primer lugar procedemos a introducir los datos en el vector de observaciones.

```
> y<-c(17,34,26,10,19,17,8,16,13,11,
+ 17,41,26,3,-6,-4,11,16,16,4,
+ 21,20,11,26,42,28,3,3,16,-10,
+ 10,24,32,26,52,28,27,28,21,42)
```

A continuación construimos las columnas de la matriz de diseño que corresponden a los parámetros α, β, γ con las funciones de repetición.

```
> alpha<-c(rep(1,10),rep(0,10),rep(0,10),rep(1,10))
> beta<-c(rep(0,10),rep(1,10),rep(1,10),rep(0,10))
> gamma<-c(rep(0,10),rep(1,10),rep(0,10),rep(1,10))
```

Los modelos lineales se definen en R con la función `lm`. Así, el modelo general y el modelo bajo la hipótesis nula se definen como

```
> crossover.lm<-lm(y~alpha+beta+gamma)
> crossover.lm0<-lm(y~gamma)
```

La columna de unos que corresponde al parámetro μ no es necesario escribirla, ya que por defecto está incluida en cualquier modelo lineal de R así definido. Observemos además que bajo la hipótesis nula $H_0 : \alpha = \beta$, el modelo a considerar sólo tiene dos parámetros μ, γ. En este caso, el efecto del fármaco (común) se puede incluir en la media general.
La tabla del análisis de la varianza para el contraste de la hipótesis nula considerada se realiza mediante la función `anova`(*modelo* H_0,*modelo general*).

```
> anova(crossover.lm0,crossover.lm)
Analysis of Variance Table

Model 1: y ~ gamma
Model 2: y ~ alpha + beta + gamma
  Res.Df    RSS Df Sum of Sq      F  Pr(>F)
1     38 6931.1
2     37 6147.9  1     783.2 4.7137 0.03641 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Del mismo modo también se puede realizar el contraste de modelos propuesto en el ejemplo 5.6.3. En este caso, el modelo más general necesita las columnas correspondientes a los parámetros γ_1, γ_2.

```
> gamma1<-c(rep(0,10),rep(1,10),rep(0,10),rep(0,10))
> gamma2<-c(rep(0,10),rep(0,10),rep(0,10),rep(1,10))
> crossover.lm1<-lm(y~alpha+beta+gamma1+gamma2)
> anova(crossover.lm,crossover.lm1)
Analysis of Variance Table

Model 1: y ~ alpha + beta + gamma
Model 2: y ~ alpha + beta + gamma1 + gamma2
  Res.Df    RSS Df Sum of Sq      F  Pr(>F)
1     37 6147.9
2     36 5547.3  1     600.6 3.8978 0.05606 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5.8. Ejercicios

Ejercicio 5.1

Sean $X \sim N(\mu_1, \sigma)$, $Y \sim N(\mu_2, \sigma)$ variables independientes. En muestras de extensión n_1 de X, n_2 de Y, plantear la hipótesis nula

$$H_0 : \mu_1 = \mu_2$$

mediante el concepto de hipótesis lineal contrastable y deducir el test t de Student de comparación de medias como una consecuencia del test F.

Ejercicio 5.2

Una variable Y depende de otra x (variable control no aleatoria) que toma los valores $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4$ de acuerdo con el modelo lineal normal

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$$

Encontrar la expresión del estadístico F para la hipótesis

$$H_0 : \beta_2 = 0$$

Ejercicio 5.3

Probar que una hipótesis lineal de matriz $\mathbf{A}$ es contrastable si y sólo si

$$\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X} = \mathbf{A}$$

Ejercicio 5.4

Con el modelo del ejercicio 3.11:

(a) ¿Podemos contrastar la hipótesis $H_0 : \theta_1 + \theta_8 = 0$?

(b) Contrastar la hipótesis $H_0 : \theta_1 = \theta_2$.

Ejercicio 5.5

Dado el siguiente modelo lineal normal

$$\begin{aligned} \beta_1 + \beta_2 &= 6.6 \\ 2\beta_1 + \beta_2 &= 7.8 \\ -\beta_1 + \beta_2 &= 2.1 \\ 2\beta_1 - \beta_2 &= 0.4 \end{aligned}$$

estudiar si se puede aceptar la hipótesis $H_0 : \beta_2 = 2\beta_1$.

Ejercicio 5.6

Continuación del ejercicio 3.10:

El transportista discute con un amigo que afirma que el doble de la distancia entre A y B es equivalente a la distancia del trayecto $A \to C \to B$. ¿Podemos aclarar en términos estadísticos su discusión?

Ejercicio 5.7

Consideremos el modelo lineal normal $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. Probar que para la hipótesis lineal

$$H_0 : \mathbf{X}\boldsymbol{\beta} = \mathbf{0}$$

se verifica $\mathrm{SCR}_H - \mathrm{SCR} = \widehat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y}$. Hallar el estadístico F correspondiente.

Ejercicio 5.8

Demostrar que para una hipótesis lineal contrastable se verifica

$$E(\mathrm{SCR}_H - \mathrm{SCR}) = q\sigma^2 + (\mathbf{A}\boldsymbol{\beta})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-}\mathbf{A}')^{-}(\mathbf{A}\boldsymbol{\beta})$$

Indicación: Utilizar la propiedad 2 del Apéndice de Estadística Multivariante con la expresión 5.5.

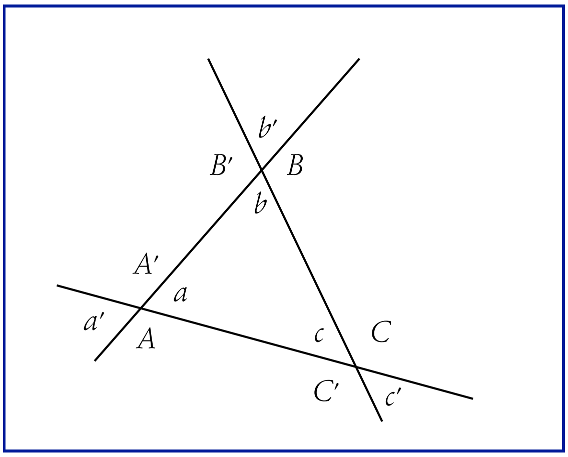
Ejercicio 5.9

Demostrar que para una hipótesis lineal contrastable se verifica la siguiente descomposición en suma de cuadrados

$$||\mathbf{Y} - \widehat{\mathbf{Y}}_H||^2 = ||\mathbf{Y} - \widehat{\mathbf{Y}}||^2 + ||\widehat{\mathbf{Y}} - \widehat{\mathbf{Y}}_H||^2$$

Ejercicio 5.10

Supongamos que cada uno de los valores $x_1, x_2, \ldots, x_{12}$ son las observaciones de los ángulos $a, a', A, A', b, b', B, B', c, c', C, C'$ del triángulo del gráfico adjunto. Los errores de las observaciones $\epsilon_1, \ldots, \epsilon_{12}$ se asume que son independientes y con distribución $N(0, \sigma)$.



Antes de escribir el modelo asociado a estos datos observemos que, aunque aparentemente hay 12 parámetros $a, a', \ldots,$ éstos están ligados por las conocidas propiedades de un triángulo, es decir

$$a = a' \qquad A = A' \qquad a + A = 180 \qquad a + b + c = 180$$

y de forma similar para b, b', B, B' y c, c', C, C'. El conjunto de estas relaciones nos conduce a que, realmente, sólo hay dos parámetros independientes, les llamaremos α y β. Si trasladamos a la izquierda las cantidades 180 y con estos parámetros, el modelo es

$$\begin{array}{llll} y_1 = \alpha + \epsilon_1 & y_2 = \alpha + \epsilon_2 & y_3 = -\alpha + \epsilon_3 & y_4 = -\alpha + \epsilon_4 \\ y_5 = \beta + \epsilon_5 & y_6 = \beta + \epsilon_6 & y_7 = -\beta + \epsilon_7 & y_8 = -\beta + \epsilon_8 \\ y_9 = -\alpha - \beta + \epsilon_9 & y_{10} = -\alpha - \beta + \epsilon_{10} & y_{11} = \alpha + \beta + \epsilon_{11} & y_{12} = \alpha + \beta + \epsilon_{12} \end{array}$$

donde

$$\begin{array}{llll} y_1 = x_1 & y_2 = x_2 & y_3 = x_3 - 180 & y_4 = x_4 - 180 \\ y_5 = x_5 & y_6 = x_6 & y_7 = x_7 - 180 & y_8 = x_8 - 180 \\ y_9 = x_9 - 180 & y_{10} = x_{10} - 180 & y_{11} = x_{11} & y_{12} = x_{12} \end{array}$$

Deseamos contrastar la hipótesis de que el triángulo es equilátero, es decir, que $a = b = c = 60$. Pero si $a = 60, b = 60$, c es automáticamente 60, luego la hipótesis es

$$H_0 : \alpha = \beta = 60$$

con 2 grados de libertad, no 3. Resolver el contraste.

Ejercicio 5.11

Con el modelo cross-over expuesto en el ejemplo 5.3.2 calcular los siguientes elementos:

(a) Una estimación de los parámetros mediante la fórmula $(\mathbf{X}_R'\mathbf{D}\mathbf{X}_R)^{-}\mathbf{X}_R'\mathbf{D}\overline{\mathbf{Y}}$.

(b) La suma de cuadrados residual

$$\begin{aligned} \mathrm{SCR} = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{P}\mathbf{Y} &= \sum y_{ij}^2 - \mathbf{Y}'\mathbf{P}\mathbf{Y} \\ &= N_{\mathbf{a}}\left(\sum_{i=1}^{4} \bar{y}_{i\cdot}^2 + \sum_{i=1}^{4} s_i^2\right) - \mathbf{Y}'\mathbf{P}\mathbf{Y} \\ &= N_{\mathbf{a}}\left(\sum_{i=1}^{4} \bar{y}_{i\cdot}^2 + \sum_{i=1}^{4} s_i^2\right) - \overline{\mathbf{Y}}'\mathbf{D}\mathbf{X}_R(\mathbf{X}_R'\mathbf{D}\mathbf{X}_R)^{-}\mathbf{X}_R'\mathbf{D}\overline{\mathbf{Y}} \end{aligned}$$

(c) La estimación de la función paramétrica $\alpha - \beta$ y su varianza.

(d) El estadístico con distribución t de Student para contrastar la hipótesis $H_0 : \alpha = \beta$

$$t = \frac{\hat{\alpha} - \hat{\beta}}{\mathrm{ee}(\hat{\alpha} - \hat{\beta})}$$

cuyo cuadrado coincide con el estadístico F del ejemplo.