UNIVERSITAT DE BARCELONA

Universitat Oberta de Catalunya

Estimación del modelo lineal
Soluciones a los ejercicios opcionales

Francesc Carmona

9 de abril de 2020

Ejercicios del libro de Faraway

Ejercicio 2.3

In this question, we investigate the relative merits of methods for computing the coefficients. Generate some artificial data by:

```
> x <- 1:20
> y <- x+rnorm(20)
```

Fit a polynomial in `x` *for predicting* `y`*. Compute* $\hat{\beta}$ *in two ways — by* `lm()` *and by using the direct calculation described in the chapter. At what degree of polynomial does the direct calculation method fail? Note the need for the* `I()` *function in fitting the polynomial, that is,* `lm(y ~ x + I(x^2))`.

```
> g2 <- lm(y ~ x + I(x^2))
> coef(g2)

 (Intercept)            x       I(x^2)
-0.541944037  0.983851729  0.002956118

> mx <- model.matrix(g2)
> solve(t(mx) %*% mx) %*% t(mx) %*% y

                    [,1]
(Intercept) -0.541944037
x            0.983851729
I(x^2)       0.002956118

> g3 <- update(g2, . ~ . + I(x^3))
> mx <- model.matrix(g3)
> cbind(coef(g3),solve(t(mx) %*% mx) %*% t(mx) %*% y)

                     [,1]          [,2]
(Intercept) -0.8768172228 -0.8768172228
x            1.1547234042  1.1547234042
I(x^2)      -0.0168980239 -0.0168980239
I(x^3)       0.0006302902  0.0006302902

> g4 <- update(g3, . ~ . + I(x^4))
> mx <- model.matrix(g4)
> cbind(coef(g4),solve(t(mx) %*% mx) %*% t(mx) %*% y)
```

```
                     [,1]         [,2]
(Intercept) -1.744929973 -1.744929973
x            1.833830051  1.833830051
I(x^2)      -0.154319254 -0.154319254
I(x^3)       0.010638178  0.010638178
I(x^4)      -0.000238283 -0.000238283

> g5 <- update(g4, . ~ . + I(x^5))
> mx <- model.matrix(g5)
> cbind(coef(g5),solve(t(mx) %*% mx) %*% t(mx) %*% y)

                     [,1]          [,2]
(Intercept) -2.790923e+00 -2.790923e+00
x            2.948393e+00  2.948393e+00
I(x^2)      -4.907529e-01 -4.907529e-01
I(x^3)       5.170617e-02  5.170617e-02
I(x^4)      -2.408823e-03 -2.408823e-03
I(x^5)       4.134362e-05  4.134362e-05

> g6 <- update(g5, . ~ . + I(x^6))
> mx <- model.matrix(g6)
> cbind(coef(g6),solve(t(mx) %*% mx) %*% t(mx) %*% y)

Error in solve.default(t(mx) %*% mx): system is computationally singular: reciprocal condition
number = 3.54243e-18
```

Con un polinomio de grado 6 la matriz $\mathbf{X}'\mathbf{X}$ es quasisingular y en la práctica no podemos calcular su inversa.

Ejercicio 2.8

An experiment was conducted to examine factors that might affect the height of leaf springs in the suspension of trucks. The data may be found in `truck`*. The five factors in the experiment are set to* $-$ *and* $+$ *but it will be more convenient for us to use* -1 *and* $+1$*. This can be achieved for the first factor by:*

```
> truck$B <- sapply(truck$B, function(x) ifelse(x == "-" ,-1 ,1))
```

Repeat for the other four factors.

```
> library(faraway)
> data("truck")
> mf <- function(x) sapply(x, function(x) ifelse(x == "-" ,-1 ,1))
> truck[,1:5] <- apply(truck[,1:5],2,mf)
```

(a) Fit a linear model for the height in terms of the five factors. Report on the value of the regression coefficients.

```
> g <- lm(height ~ ., data=truck)
> coef(g)

(Intercept)           B           C           D           E           O
  7.6360417   0.1106250  -0.0881250  -0.0143750   0.0518750  -0.1297917
```

(b) Fit a linear model using just factors B, C, D and E and report the coefficients. How do these compare to the previous question? Show how we could have anticipated this result by examining the X matrix.

```
> g2 <- lm(height ~ B + C + D + E, data=truck)
> coef(g2)

(Intercept)           B           C           D           E
   7.636042    0.110625   -0.088125   -0.014375    0.051875
```

Si calculamos $\mathbf{X}'\mathbf{X}$

```
> x <- model.matrix(g)
> t(x) %*% x

            (Intercept)  B  C  D  E  O
(Intercept)          48  0  0  0  0  0
B                     0 48  0  0  0  0
C                     0  0 48  0  0  0
D                     0  0  0 48  0  0
E                     0  0  0  0 48  0
O                     0  0  0  0  0 48
```

observamos que hay ortogonalidad, de manera que la estimación de los coeficientes es la misma independientemente de la presencia o no de la variable O.

(c) Construct a new predictor called A which is set to B+C+D+E. Fit a linear model with the predictors A, B, C, D, E and O. Do coefficients for all six predictors appear in the regression summary? Explain.

```
> attach(truck)
> A <- B + C + D + E
> g3 <- lm(height ~ A + B + C + D + E + O, data=truck)
> coef(g3)

(Intercept)           A           B           C           D           E
  7.6360417   0.0518750   0.0587500  -0.1400000  -0.0662500          NA
          O
 -0.1297917
```

En este último modelo, una de las variables regresoras (A) es combinación lineal de otras. Eso hace que el rango de la matriz de diseño no sea máximo y por tanto no podemos estimar todos los parámetros. La propuesta de **R** es desestimar la variable E.

(d) Extract the model matrix $\mathbf{X}$ *from the previous model. Attempt to compute* $\hat{\boldsymbol{\beta}}$ *from* $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. *What went wrong and why?*

```
> x <- model.matrix(g3)
> solve(t(x) %*% x)

Error in solve.default(t(x) %*% x): Lapack routine dgesv: system is exactly singular: U[6,6]
= 0
```

La matriz $\mathbf{X}'\mathbf{X}$ no es de rango máximo y, por tanto, no tiene inversa.

(e) Use the QR decomposition method as seen in Section 2.7 to compute $\hat{\boldsymbol{\beta}}$. *Are the results satisfactory?*

```
> qrx <- qr(x)
> f <- t(qr.Q(qrx)) %*% height
> backsolve(qr.R(qrx),f)

              [,1]
[1,]  7.636042e+00
[2,]  8.071747e+12
[3,] -8.071747e+12
[4,] -8.071747e+12
[5,] -8.071747e+12
[6,] -1.311865e-01
[7,] -8.071747e+12
```

Claramente el resultado no es satisfactorio. Los coeficientes tienen valores muy elevados que hacen sospechar que algo no ha funcionado.

Si miramos la matriz **R** de la descomposición

```
> qr.R(qrx)

  (Intercept)         A         B         C         D             O
1   -6.928203   0.00000  0.000000  0.000000  0.000000  0.000000e+00
2    0.000000 -13.85641 -3.464102 -3.464102 -3.464102  2.442491e-15
3    0.000000   0.00000  6.000000 -2.000000 -2.000000 -4.440892e-16
4    0.000000   0.00000  0.000000 -5.656854  2.828427  0.000000e+00
5    0.000000   0.00000  0.000000  0.000000 -4.898979 -1.110223e-16
6    0.000000   0.00000  0.000000  0.000000  0.000000  6.928203e+00
7    0.000000   0.00000  0.000000  0.000000  0.000000  0.000000e+00
              E
1  0.000000e+00
2 -3.464102e+00
3 -2.000000e+00
4  2.828427e+00
5  4.898979e+00
6 -1.197259e-15
7 -8.504822e-15
```

observamos que el último elemento de la diagonal tiene un valor muy pequeño cuando, en realidad, debería ser cero. Eso hace que la solución del sistema con esa matriz **R** sea totalmente errónea. Vemos también que la última columna de dicha matriz representa la variable E. Si eliminamos esa variable del sistema, la solución será correcta:

```
> backsolve(qr.R(qrx)[,-7],f[-7])

[1]  7.6360417  0.0518750  0.0587500 -0.1400000 -0.0662500 -0.1297917
```

(f) Use the function `qr.coef` *to correctly compute* $\hat{\beta}$.

La función `qr.coef` de **R** hace justamente lo que se ha dicho en el apartado anterior y además conserva el orden de las variables regresoras.

```
> qr.coef(qrx, height)

(Intercept)           A           B           C           D           E
  7.6360417   0.0518750   0.0587500  -0.1400000  -0.0662500          NA
          O
 -0.1297917
```

```
> detach(truck)
```

Ejercicios del libro de Carmona

Ejercicio 3.2

En un modelo lineal, la matriz de diseño es

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Hallar la expresión general de las funciones paramétricas estimables.

La matriz de diseño del modelo lineal propuesto tiene 5 columnas, es decir, que el modelo lineal tiene 5 parámetros. El rango de esa matriz es

```
> matriz <- matrix(c(1,1,1,1,1,
+                    1,0,1,0,0,
+                    1,1,1,0,0,
+                    1,0,1,1,1), byrow=T, ncol=5)
> qr(matriz)$rank

[1] 3
```

De forma que en realidad solo tenemos 3 parámetros efectivos.

Sabemos que una función paramétrica es estimable si es combinación lineal de las filas de la matriz de diseño. Como el rango es 3, tomamos 3 filas linealmente independientes de la matriz de diseño, por ejemplo las tres primeras:

```
> qr(matriz[1:3,])$rank

[1] 3
```

y vemos como son las combinaciones lineales de esas tres filas:

$$(a_1, a_2, a_3, a_4, a_5) = \lambda_1(1,1,1,1,1) + \lambda_2(1,0,1,0,0) + \lambda_3(1,1,1,0,0)$$

Ahora planteamos el sistema de ecuaciones:

$$\begin{aligned} a_1 &= \lambda_1 + \lambda_2 + \lambda_3 \\ a_2 &= \lambda_1 + \quad + \lambda_3 \\ a_3 &= \lambda_1 + \lambda_2 + \lambda_3 \\ a_4 &= \lambda_1 \\ a_5 &= \lambda_1 \end{aligned}$$

La solución de este sistema en función de las a_i (eliminar las lambdas) en este caso es muy sencilla y tendrá $5 - 3 = 2$ ecuaciones:

$$a_1 = a_3, \quad a_4 = a_5$$

Así pues, cualquier función paramétrica del tipo $a_1\beta_1 + a_2\beta_2 + a_3\beta_3 + a_4\beta_4 + a_5\beta_5$ será estimable si los coeficientes a_i verifican estas dos restricciones.

Ejercicio 3.7

Consideremos el modelo lineal

$$\begin{aligned} y_1 &= \beta_1 + \beta_2 + \epsilon_1 \\ y_2 &= \beta_1 + \beta_3 + \epsilon_2 \\ y_3 &= \beta_1 + \beta_2 + \epsilon_3 \end{aligned}$$

Se pide:

1) ¿Es la función paramétrica $\psi = \beta_1 + \beta_2 + \beta_3$ estimable?

Este modelo tiene 3 parámetros y su rango es 2 ya que la tercera fila de la matriz de diseño es igual a la primera.
Para ver si la función paramétrica propuesta es estimable debemos asegurarnos de que es combinación lineal de las filas de la matriz de diseño del modelo. Para ello, podemos añadir los coeficientes $(1, 1, 1)$ a la matriz de diseño y ver si conserva el rango.

```
> matriz <- matrix(c(1,1,0,
+                    1,0,1,
+                    1,1,0,
+                    1,1,1), byrow=T, ncol=3)
> qr(matriz)$rank

[1] 3
```

Como el rango ahora es 3, la fila añadida no es combinación lineal de las otras y la función paramétrica no será estimable.

2) Probar que toda función paramétrica

$$\psi = a_1\beta_1 + a_2\beta_2 + a_3\beta_3$$

es estimable si y sólo si $a_1 = a_2 + a_3$.

Como en el ejercicio anterior, planteamos la combinación lineal de 2 filas independientes (rango 2) de la matriz de diseño:

$$(a_1, a_2, a_3) = \lambda_1(1, 1, 0) + \lambda_2(1, 0, 1)$$

El sistema de ecuaciones es:

$$\begin{aligned} a_1 &= \lambda_1 + \lambda_2 \\ a_2 &= \lambda_1 \\ a_3 &= \quad + \lambda_2 \end{aligned}$$

De modo que la solución es sencilla: $a_1 = a_2 + a_3$.

Ejercicio 3.8

Consideremos el modelo lineal

$$\begin{aligned} y_1 &= \mu + \alpha_1 + \beta_1 + \epsilon_1 \\ y_2 &= \mu + \alpha_1 + \beta_2 + \epsilon_2 \\ y_3 &= \mu + \alpha_2 + \beta_1 + \epsilon_3 \\ y_4 &= \mu + \alpha_2 + \beta_2 + \epsilon_4 \\ y_5 &= \mu + \alpha_3 + \beta_1 + \epsilon_5 \\ y_6 &= \mu + \alpha_3 + \beta_2 + \epsilon_6 \end{aligned}$$

(a) ¿Cuando es $a_0\mu + a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3 + a_4\beta_1 + a_5\beta_2$ estimable?

Se trata de un típico diseño de dos factores sin interacción. Si consideramos los parámetros $\mu, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$, la matriz de diseño asociada es:

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

cuyo rango es

```
> matriz <- matrix(c(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), byrow=T, ncol=6)
> qr(matriz)$rank

[1] 4
```

También podemos observar que la suma de la segunda, la tercera y la cuarta columnas es la primera y que la suma de quinta y la sexta también es la primera, de modo que efectivamente el rango es 4. De los 6 parámetros, únicamente 4 son efectivos.
Veamos si las 4 primeras filas son linealmente independientes:

```
> qr(matriz[1:4,])$rank

[1] 3
```

No lo son. Probamos con la primera, la segunda, la tercera y la quinta:

```
> qr(matriz[c(1:3,5),])$rank

[1] 4
```

Ahora sí que podemos escribir la combinación lineal:

$$(a_0, a_1, a_2, a_3, a_4, a_5) = \lambda_1(1,1,0,0,1,0) + \lambda_2(1,1,0,0,0,1) + \lambda_3(1,0,1,0,1,0) + \lambda_4(1,0,0,1,1,0)$$

y el sistema de ecuaciones asociado:

$$\begin{aligned} a_0 &= \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \\ a_1 &= \lambda_1 + \lambda_2 \\ a_2 &= \lambda_3 \\ a_3 &= \lambda_4 \\ a_4 &= \lambda_1 + \lambda_3 + \lambda_4 \\ a_5 &= \lambda_2 \end{aligned}$$

cuya solución, en forma de $6 - 4 = 2$ ecuaciones, es evidente:

$$a_0 = a_1 + a_2 + a_3, \quad a_0 = a_4 + a_5$$

(b) ¿Es $\alpha_1 + \alpha_2$ estimable?

Los coeficientes $(a_0, a_1, a_2, a_3, a_4, a_5)$ para esta función paramétrica son $(0, 1, 1, 0, 0, 0)$ de manera que

$$0 \neq 1 + 1 + 0, \quad 0 = 0 + 0$$

Como una de las restricciones no se verifica, $\alpha_1 + \alpha_2$ no es estimable.

(c) ¿Es $\beta_1 - \beta_2$ estimable?

En este caso los coeficientes son $(0, 0, 0, 0, 1, -1)$ de manera que

$$0 = 0 + 0 + 0, \quad 0 = 1 - 1$$

Se verifican las dos restricciones y $\beta_1 - \beta_2$ es estimable.

(d) ¿Es $\mu + \alpha_1$ estimable?

En este caso los coeficientes son $(1, 1, 0, 0, 0, 0)$ de manera que

$$1 = 1 + 0 + 0, \quad 1 \neq 0 + 0$$

No se verifican las dos restricciones y $\mu + \alpha_1$ no es estimable.

(e) ¿Es $6\mu + 2\alpha_1 + 2\alpha_2 + 2\alpha_3 + 3\beta_1 + 3\beta_2$ estimable?

Como en los apartados anteriores, comprobamos si

$$6 = 2 + 2 + 2, \quad 6 = 3 + 3$$

de manera que es estimable.

(f) ¿Es $\alpha_1 - 2\alpha_2 + \alpha_3$ estimable?

Como en los apartados anteriores, comprobamos si

$$0 = 1 - 2 + 1, \quad 0 = 0 + 0$$

de manera que es estimable.

(g) Hallar la covarianza entre los estimadores lineales MC de las funciones paramétricas $\beta_1 - \beta_2$ y $\alpha_1 - \alpha_2$

En primer lugar comprobamos que sean estimables. La primera ya lo hemos visto y la segunda

$$0 = 1 - 1 + 0, \quad 0 = 0 + 0$$

también. Las dos son estimables.
Sabemos que la matriz de covarianzas de los parámetros estimados es

$$\boldsymbol{\Sigma} = \text{cov}(\boldsymbol{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^-$$

de modo que la covarianza entre las dos f.p.e. es

$$\text{cov}(\beta_1 - \beta_2, \alpha_1 - \alpha_2) = \text{cov}(\beta_1, \alpha_1) - \text{cov}(\beta_1, \alpha_2) - \text{cov}(\beta_2, \alpha_1) + \text{cov}(\beta_2, \alpha_2) = \sigma_{52} - \sigma_{53} - \sigma_{62} + \sigma_{63}$$

que podemos concretar más si calculamos la matriz $(\mathbf{X}'\mathbf{X})^-$

```
> library(MASS)
> ginvXtX <- ginv(t(matriz) %*% matriz)
```

así pues

```
> ginvXtX[5,2] - ginvXtX[5,3] - ginvXtX[6,2] + ginvXtX[6,3]

[1] 3.989864e-17
```

y la covarianza entre las dos f.p.e. es $\sigma^2(-2.94903 \cdot 10^{-17})$, prácticamente cero.

El mismo cálculo se puede hacer de forma matricial. Si $\mathbf{a}'\boldsymbol{\beta}$ y $\mathbf{b}'\boldsymbol{\beta}$ son las dos f.p.e., la covarianza viene dada por

$$\text{cov}(\mathbf{a}'\boldsymbol{\beta}, \mathbf{b}'\boldsymbol{\beta}) = \sigma^2 \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-}\mathbf{b}$$

En nuestro caso:

$$\mathbf{a}' = (0,0,0,0,1,-1), \quad \mathbf{b}' = (0,1,-1,0,0,0)$$

y su covarianza es

```
> t(c(0,0,0,0,1,-1)) %*% ginvXtX %*% c(0,1,-1,0,0,0)

             [,1]
[1,] 3.989864e-17
```

multiplicado por σ^2.

(h) Hallar la dimensión del espacio paramétrico.

La dimensión del espacio paramétrico es el número de parámetros realmente efectivos.
Sabemos que el vector de observaciones está en un espacio de dimensión $n =$ número de observaciones. Dicho espacio se descompone en dos subespacios ortogonales $\mathbb{R}^n = \boldsymbol{\Omega} + \boldsymbol{\Omega}^{\perp}$. El espacio de estimaciones es $\boldsymbol{\Omega} = \langle \mathbf{X} \rangle$ y su dimensión será el rango de la matriz de diseño. En nuestro caso $r = 4$.

(i) Obtener una expresión del espacio de los errores.

El espacio de los errores es $\boldsymbol{\Omega}^{\perp}$ que tiene por dimensión $n - r$. Para hallar una base de ese subespacio bastará con buscar una base del espacio $\boldsymbol{\Omega} = \langle \mathbf{X} \rangle$ (en nuestro caso 4 vectores columna linealmente independientes de la matriz de diseño) y luego hallar $n - r$ vectores de una base ortogonal.
La construcción del complemento ortogonal de un subespacio consiste en plantear y resolver un sistema de ecuaciones para buscar los vectores (una base de ellos) que son ortogonales a todos los de la base del subespacio original.

Ejercicio 3.10

Un transportista realiza diversos trayectos entre tres poblaciones A, B y C. En cuatro días consecutivos ha hecho los recorridos que muestra la siguiente tabla:

trayecto	km
$A \to B \to A \to C$	533
$C \to A \to C \to B$	583
$B \to C \to A \to C \to A \to B \to A$	1111
$A \to B \to A \to C \to A \to B \to A$	1069

donde el kilometraje es, por diversas razones, aproximado.

(a) Proponer un modelo lineal, con la matriz de diseño y las hipótesis necesarias, para estimar las distancias kilométricas entre las tres poblaciones.
Con los datos proporcionados, ¿es posible estimar las distancias entre las tres poblaciones? ¿Cuales son las distancias o funciones paramétricas estimables (fpe) en este modelo?

Los parámetros a estimar en este problema son las distancias entre las tres ciudades. Podemos hacer la suposición razonable que las distancias entre dos ciudades son iguales, es decir, $d(A,B) = d(B,A)$.

Entonces, los parámetros son

$$\begin{aligned} \alpha &= d(A,B) \\ \beta &= d(A,C) \\ \gamma &= d(B,C) \end{aligned}$$

El modelo lineal será

$$\begin{aligned} 533 &= 2\alpha + \beta + \epsilon_1 \\ 583 &= 2\beta + \gamma + \epsilon_2 \\ 1111 &= 2\alpha + 3\beta + \gamma + \epsilon_3 \\ 1069 &= 4\alpha + 2\beta + \epsilon_4 \end{aligned}$$

con las hipótesis de Gauss-Markov.

En este modelo es evidente que el rango de la matriz de diseño es $r = 2$, ya que la tercera fila es la suma de las dos primeras y la cuarta es el doble de la primera. Además, $n = 4$.

Ahora vamos a ver como son las f.p.e. Para ello planteamos la combinación lineal siguiente:

$$(a_1, a_2, a_3) = \lambda_1(2, 1, 0) + \lambda_2(0, 2, 1)$$

y el sistema asociado:

$$\begin{aligned} a_1 &= 2\lambda_1 \\ a_2 &= \lambda_1 + 2\lambda_2 \\ a_3 &= \lambda_2 \end{aligned}$$

de aquí que $a_2 = a_1/2 + 2a_3$ o también $2a_2 = a_1 + 4a_3$.

Veamos si es posible estimar las distancias entre ciudades: $d(A, B) = \alpha \to (1, 0, 0)$ de modo que $2 \cdot 0 \neq 1 + 4 \cdot 0$ y no es estimable. Lo mismo ocurre con las otras dos distancias.

(b) ¿Se puede estimar el kilometraje del trayecto $M_{BC} \to B \to A \to C \to M_{AC}$, donde M_{IJ} es el punto medio entre dos poblaciones? ¿Es una buena estimación? ¿Cual es el error de esta estimación?

Vamos a traducir el trayecto a una suma con los parámetros del modelo:

$$\frac{1}{2}\gamma + \alpha + \beta + \frac{1}{2}\beta = \alpha + \frac{3}{2}\beta + \frac{1}{2}\gamma$$

y en este caso $2\frac{3}{2} = 1 + 4\frac{1}{2}$ es cierto y el trayecto es estimable.

La estimación la obtenemos así:

```
> y <- c(533,583,1111,1069)
> X <- matrix(c(2,1,0,
+                    0,2,1,
+                    2,3,1,
+                    4,2,0), byrow=T, ncol=3)
> betas <- ginv(t(X) %*% X) %*% t(X) %*% y
> a <- c(1,3/2,1/2)
> t(a) %*% betas

         [,1]
[1,] 556.9091
```

Esta estimación se obtiene por el método de los mínimos cuadrados, de forma que es la mejor estimación (BLUE) según el teorema de Gauss-Markov.

El error de la estimación se calcula con la fórmula $\text{var}(\mathbf{a}'\boldsymbol{\beta}) = \hat{\sigma}^2\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-}\mathbf{a}$:

```
> n <- length(y)
> r <- 2
> residuos <- y - X %*% betas
> sigma2 <- sum(residuos^2)/(n-r)
> var.error <- sigma2 * t(a) %*% ginv(t(X) %*% X) %*% a
> error <- sqrt(var.error)
> error                                # de la función paramétrica

        [,1]
[1,] 1.14632
```