

Sesión 14 (D) — La matriz de regresores en la práctica: verificación con datos simulados

Marcos Bujosa

Descripción de la práctica

Cerramos la sesión de regresión múltiple volviendo, por primera vez con Gretl, sobre la notación matricial de la lección 10: la matriz de datos $\mathbf{X}$, las ecuaciones normales $\mathbf{X}^\top \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{X}^\top \mathbf{y}$, y la condición de ortogonalidad $\mathbf{X}^\top \hat{\mathbf{e}} = \mathbf{0}$. Trabajamos con datos *simulados* porque ahora no queremos *interpretar* ningún fenómeno económico; queremos *verificar* que la maquinaria algebraica de la lección 10 funciona exactamente como se demostró: primero en un caso general, y después en el caso especial de regresores mutuamente ortogonales (“familia ortogonal”), donde cada coeficiente se puede calcular de forma aislada.

Objetivo

1. Construir explícitamente, con el comando `matrix` de Gretl, la matriz de datos $\mathbf{X}$ de un modelo con dos regresores no constantes.
2. Calcular $\mathbf{X}^\top \mathbf{X}$ y $\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$, comparando el resultado con el que da `ols` directamente.
3. Verificar $\mathbf{X}^\top \hat{\mathbf{e}} = \mathbf{0}$, en sus dos lecturas (matriz por vector y vector por matriz).
4. Construir un caso de *familia ortogonal* (regresores mutuamente ortogonales) y comprobar que cada $\hat{\beta}_j$ se puede calcular de forma aislada, sin resolver ningún sistema.
5. Verificar, en ese mismo caso, que $R^2 = \rho_{\mathbf{z}_2 \mathbf{y}}^2 + \rho_{\mathbf{z}_3 \mathbf{y}}^2$.
6. Ver qué hace Gretl ante la situación contraria, la *colinealidad exacta*: dos regresores que son la misma variable en dos unidades.

Actividad 1 - Datos simulados: un caso general, no ortogonal

Generamos $n = 40$ observaciones con dos regresores no constantes, x_2 y x_3 , deliberadamente *correlacionados entre sí* (para no caer, sin darnos cuenta, en el caso especial de la Actividad 4).

en línea de comandos:

```
nulldata 40 --preserve
set seed 97531

series x2 = uniform(0,10)
series x3 = 0.6*x2 + normal(0,2)
series u = normal(0,3)

scalar b1 = 5
scalar b2 = 2
scalar b3 = -1
series y = b1 + b2*x2 + b3*x3 + u

printf "Correlacion entre x2 y x3 = %.4f (no son ortogonales)\n", corr(x2,x3)
```

Licencia: Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Recuerde la advertencia de siempre sobre los vectores de diseño: $\mathbf{u}$ es el vector que nosotros usamos para generar los datos; no debe confundirse con el vector de residuos $\hat{\mathbf{e}}$ que calculará más adelante MCO a partir de los datos simulados.

Actividad 2 - La matriz de regresores y las ecuaciones normales

Construimos explícitamente la matriz de datos $\mathbf{X} = [\mathbf{1}; \mathbf{x}_2; \mathbf{x}_3]$ (lección 10) con el comando `matrix` de Gretl, a partir de una lista de series.

en línea de comandos:

```
list Xlist = const x2 x3
matrix X = { Xlist }
list ylist = y
matrix Y = { ylist }

matrix XtX = X'X
matrix XtY = X'Y

printf "Matriz X'X (3x3):\n"
print XtX
printf "\nVector X'y (3x1):\n"
print XtY
```

Matriz X'X (3x3):
XtX (3 x 3)

40,000	193,05	114,98
193,05	1290,6	715,42
114,98	715,42	612,46

Vector X'y (3x1):
XtY (3 x 1)

472,54
2881,9
1421,0

No pase de largo ante esa matriz: los apuntes de “Ecuaciones normales” afirman que cada elemento de $\mathbf{X}^\top \mathbf{X}$ es un producto escalar entre dos columnas de $\mathbf{X}$, y eso se puede comprobar aquí mismo, a ojo. El elemento (1, 1) es $\langle \mathbf{1}, \mathbf{1} \rangle_e$, es decir, n : debe valer 40, y vale 40. El elemento (1, 2) es $\langle \mathbf{1}, \mathbf{x}_2 \rangle_e = \sum_i x_{2i} = n\mu_{\mathbf{x}_2}$: divídalo entre 40 y obtendrá la media de $\mathbf{x}_2$. Lo mismo con el (1, 3) y la media de $\mathbf{x}_3$. La matriz no es una caja negra que Gretl produce; es la tabla de todos los productos escalares entre regresores, y la primera fila y la primera columna son, sencillamente, sumas.

Ahora resolvemos las ecuaciones normales *usando la inversa* —tal como aparece en la mayoría de los manuales (lección 10)— y comparamos el resultado con lo que da `ols` directamente.

en línea de comandos:

```
matrix beta_matricial = invpd(XtX)*XtY
printf "beta calculado como (X'X)^(-1) X'y:\n"
print beta_matricial

ols y const x2 x3 --quiet
matrix beta_ols = $coeff
printf "\nbeta calculado por 'ols' de Gretl:\n"
print beta_ols

matrix diferencia = beta_matricial - beta_ols
printf "\nDiferencia (debe ser practicamente cero):\n"
print diferencia
```

```
beta calculado como (X'X)^(-1) X'y:
beta_matricial (3 x 1)
```

```
4,4307
2,1139
-0,98078
```

```
beta calculado por 'ols' de Gretl:
beta_ols (3 x 1)
```

```
4,4307
2,1139
-0,98078
```

```
Diferencia (debe ser practicamente cero):
diferencia (3 x 1)
```

```
2,6645e-15
4,4409e-16
-1,3323e-15
```

Recuerde la advertencia de la lección 10: esta fórmula con inversa EXISTE, y es lo que Gretl calcula “por dentro” al ejecutar ols; pero el camino conceptual del curso nunca ha sido, ni será, invertir una matriz: es siempre la condición de ortogonalidad $\mathbf{X}^\top \hat{\mathbf{e}} = \mathbf{0}$, que verificamos en la siguiente actividad.

Actividad 3 - Las dos lecturas de la ortogonalidad: $\mathbf{X}^\top \hat{\mathbf{e}}$ y $\hat{\mathbf{e}}\mathbf{X}$

Recuperamos el vector de residuos de la última regresión y verificamos que es ortogonal a las tres columnas de $\mathbf{X}$ simultáneamente, en sus dos lecturas equivalentes (lección 10, “Transposición”).

en línea de comandos:

```
series ehat = $uhat
list elist = ehat
matrix E = { elist }

matrix XtE = X'E
printf "X'e (columna 3x1, debe ser (aprox.) el vector nulo):\n"
print XtE

matrix eX = E'X
printf "\ne*X (Gretl lo presenta como fila 1x3; mismos tres numeros):\n"
print eX
```

```
X'e (columna 3x1, debe ser (aprox.) el vector nulo):
XtE (3 x 1)
```

```
-1,5543e-13
5,7732e-13
-4,0945e-13
```

```
e*X (Gretl lo presenta como fila 1x3; mismos tres numeros):
eX (1 x 3)
```

```
-1,5543e-13  5,7732e-13  -4,0945e-13
```

Observe que $\mathbf{XtE}$ y $\mathbf{eX}$ contienen *los mismos tres números*: es exactamente la identidad $\hat{\mathbf{e}}\mathbf{X} = \mathbf{X}^\top \hat{\mathbf{e}}$ de la lección 10, comprobada aquí numéricamente.

Una advertencia sobre lo que ve en pantalla, para que no deshaga lo aprendido en la lección 10. Gretl presenta uno de los dos resultados “en columna” y el otro “en fila”, y el guion ha tenido que escribir $\mathbf{E}'\mathbf{X}$

—transponiendo lo que el curso llama un vector— porque el lenguaje matricial de Gretl no ofrece otra cosa: *para Gretl* un vector es una matriz de una sola fila o de una sola columna, y por tanto tiene orientación. Para nosotros no la tiene: un vector es una lista de coordenadas, y solo las matrices se transponen (lección 10, “Transposición”; y la pregunta 2 del repaso de esa lección, que quizá acaba de responder). De hecho, el sentido de aquella transparencia es precisamente este: $\hat{e}X$ *no exige* transponer $\hat{e}$; la transposición recae sobre la matriz, no sobre el vector. Lo que ve en pantalla es una convención de Gretl, no un ingrediente del curso.

Actividad 4 - Familia ortogonal: el mismo experimento, cambiando una sola cosa

Vamos a construir ahora el caso especial de la transparencia “Familia ortogonal”. Y vamos a construirlo con un cuidado que conviene explicitar, porque de él depende que la comparación signifique algo.

Podríamos inventarnos un mundo nuevo: otro tamaño muestral, otros coeficientes, otro nivel de ruido. Sería más fácil de escribir, pero *no probaría nada*: si cambiamos cinco cosas a la vez y el resultado cambia, no sabemos a cuál de las cinco atribuirlo. Así que haremos lo contrario: dejamos **todo** exactamente como en la Actividad 1 —las mismas $n = 40$ observaciones, los mismos $\beta_1 = 5$, $\beta_2 = 2$, $\beta_3 = -1$, y *la misma perturbación* u , sin volver a sortearla— y cambiamos **una sola cosa**: la correlación entre los dos regresores, que pasará de 0,50 a cero exacto.

Para conseguirlo no hace falta ningún truco nuevo, solo la proyección ortogonal de siempre: centramos x_2 para que tenga media nula, y a x_3 le quitamos su proyección sobre la constante y sobre z_2 —es decir, nos quedamos con sus /residuos/. Por construcción (ecuaciones normales, lección 10), lo que queda es exactamente ortogonal a 1 y a z_2 .

en línea de comandos:

```
series z2 = x2 - mean(x2)          # centrado: media nula
ols x3 const z2 --quiet
series z3 = $uhat                  # residuos: ortogonales a 1 y a z2 por construccion

series yo = b1 + b2*z2 + b3*z3 + u # MISMO beta, MISMA perturbacion u

printf "media(z2)    = %.10f    (debe ser 0)\n", mean(z2)
printf "media(z3)    = %.10f    (debe ser 0)\n", mean(z3)
printf "suma(z2*z3)  = %.10f    (debe ser 0: ortogonalidad mutua)\n", sum(z2*z3)
printf "corr(z2,z3)  = %.10f    (era 0,5046 entre x2 y x3)\n", corr(z2,z3)
```

```
media(z2)    = 0,0000000000    (debe ser 0)
media(z3)    = -0,0000000000    (debe ser 0)
suma(z2*z3)  = 0,0000000000    (debe ser 0: ortogonalidad mutua)
corr(z2,z3)  = 0,0000000000    (era 0,5046 entre x2 y x3)
```

Observe que la ortogonalidad no es aproximada sino EXACTA: las cifras salen a cero hasta el décimo decimal. No es casualidad ni buena suerte con la semilla, sino consecuencia directa de las ecuaciones normales: el residuo de una regresión es siempre, por construcción, ortogonal a todos los regresores de esa regresión. Otra forma habitual de conseguir ortogonalidad exacta es un diseño factorial con valores ± 1 —lo reencontraremos, con otro propósito, al estudiar variables dummy—, pero exigiría cambiar los datos, y aquí queremos justamente conservarlos.

Actividad 5 - Cada coeficiente por separado: la familia ortogonal en acción

Recordando la transparencia “Familia ortogonal” de la lección 10: si los regresores son mutuamente ortogonales, cada $\hat{\beta}_j$ puede calcularse *de forma aislada*, sin resolver ningún sistema, con la misma fórmula de Cauchy–Schwarz de la lección 3; y además R^2 se descompone como suma de correlaciones al cuadrado.

Vamos a comprobarlo, pero haciendo *las dos cuentas en los dos escenarios*: el general de la Actividad 1 y el ortogonal que acabamos de construir. Porque comprobar que la fórmula funciona donde debe funcionar solo es la mitad del experimento; la otra mitad es ver qué ocurre donde no debe.

en línea de comandos:

```
printf "===== CASO GENERAL (Actividad 1) =====\n"
printf "correlacion entre los regresores = %.4f\n", corr(x2,x3)
ols y const x2 x3 --quiet
printf "beta2: conjunto = %10.6f    aislado (cov/var) = %10.6f\n", $coeff(x2), cov(x2,y)/var(x2)
printf "beta3: conjunto = %10.6f    aislado (cov/var) = %10.6f\n", $coeff(x3), cov(x3,y)/var(x3)
printf "R2 : conjunto = %10.6f    suma de rho^2 = %10.6f\n", $rsq, corr(x2,y)^2 + corr(x3,y)^2

printf "\n===== CASO ORTOGONAL (Actividad 4) =====\n"
printf "correlacion entre los regresores = %.4f\n", corr(z2,z3)
ols yo const z2 z3 --quiet
printf "beta2: conjunto = %10.6f    aislado (cov/var) = %10.6f\n", $coeff(z2), cov(z2,yo)/var(z2)
printf "beta3: conjunto = %10.6f    aislado (cov/var) = %10.6f\n", $coeff(z3), cov(z3,yo)/var(z3)
printf "R2 : conjunto = %10.6f    suma de rho^2 = %10.6f\n", $rsq, corr(z2,yo)^2 + corr(z3,yo)^2
```

```
===== CASO GENERAL (Actividad 1) =====
correlacion entre los regresores = 0,5046

beta2: conjunto = 2,113860    aislado (cov/var) = 1,675269
beta3: conjunto = -0,980779    aislado (cov/var) = 0,222597
R2 : conjunto = 0,771932    suma de rho^2 = 0,651820

===== CASO ORTOGONAL (Actividad 4) =====
correlacion entre los regresores = 0,0000

beta2: conjunto = 2,122456    aislado (cov/var) = 2,122456
beta3: conjunto = -0,980779    aislado (cov/var) = -0,980779
R2 : conjunto = 0,835808    suma de rho^2 = 0,835808
```

Lo que debe observar

La salida tiene dos bloques, y hay que leerlos como las dos mitades de un mismo experimento. Recuerde que *lo único* que los separa es la correlación entre los regresores: mismo n , mismos β , misma perturbación.

En el bloque de abajo (*ortogonal*), las dos columnas coinciden en las tres filas. Cada coeficiente se puede calcular aislado, ignorando por completo al otro regresor, y sale el mismo número que da el modelo conjunto. Y R^2 coincide con $\rho_{z_2y}^2 + \rho_{z_3y}^2$. Es lo que predice la lección 10.

En el bloque de arriba (*general*), ninguna de las tres filas coincide, y conviene mirarlás una a una:

- $\hat{\beta}_2$: el modelo conjunto da 2,11; el cálculo aislado, 1,68. Un error de en torno al 20 %.
- $\hat{\beta}_3$: el modelo conjunto da $-0,98$ —muy cerca del valor verdadero, $\beta_3 = -1$ —, mientras que el cálculo aislado da $+0,22$. **No es que se equivoque en la magnitud: se equivoca en el signo.** Quien calculase así concluiría que x_3 influye positivamente sobre y , cuando en el mundo que hemos fabricado influye negativamente.
- R^2 : el modelo conjunto da 0,772; la suma de correlaciones al cuadrado, 0,652.

Este segundo bloque es el que da sentido al primero. La transparencia “Familia ortogonal” no describe una simplificación cómoda que uno pueda usar siempre con un poco de cuidado: describe un caso *excepcional*, y fuera de él la fórmula no se degrada suavemente, sino que puede llegar a invertir una conclusión económica. Que en la Actividad 1 los regresores estuvieran correlacionados sólo a 0,50 —una correlación nada extrema para datos reales— basta para producir un cambio de signo.

Actividad 6 - El caso contrario: colinealidad exacta

La lección 10 puso un ejemplo de regresores *linealmente dependientes*: la temperatura de un mismo día en grados Celsius y en Fahrenheit, $F = 32 + 1,8C$. Cada una varía; juntas no aportan dos direcciones, sino

una. Fabriquemos ese caso con los datos de la Actividad 1, tomando x_2 como si fuera una temperatura en Celsius.

en línea de comandos:

```
series C = x2          # "temperatura" en Celsius
series F = 32 + 1.8*C  # la misma temperatura, en Fahrenheit
printf "corr(C, F) = %.10f\n", corr(C, F)

list Xcol = const C F
matrix Xc = { Xcol }
printf "determinante de X'X con C y F = %g  (una matriz invertible tiene determinante distinto de cero)\n", det(Xc'Xc)

ols y const C F
```

```
corr(C, F) = 1,0000000000
determinante de X'X con C y F = 0  (una matriz invertible tiene determinante distinto de cero)
```

Modelo 1: MCO, usando las observaciones 1-40

Variable dependiente: y

Omitidas debido a colinealidad exacta: F

	coeficiente	Desv. típica	Estadístico t	valor p
const	3,72826	1,15050	3,241	0,0025 ***
C	1,67527	0,202541	8,271	5,04e-10 ***

Media de la vble. dep.	11,81341	D.T. de la vble. dep.	6,338798
Suma de cuad. residuos	559,5827	D.T. de la regresión	3,837429
R-cuadrado	0,642903	R-cuadrado corregido	0,633506
F(1, 38)	68,41377	Valor p (de F)	5,04e-10
Log-verosimilitud	-109,5238	Criterio de Akaike	223,0476
Criterio de Schwarz	226,4253	Crit. de Hannan-Quinn	224,2688

Lo que debe observar

La correlación entre C y F es exactamente 1: son el mismo vector en desviaciones, salvo escala. El determinante de $\mathbf{X}^T \mathbf{X}$ es nulo (o del orden del error de redondeo): la matriz no es invertible, y las ecuaciones normales tienen infinitas soluciones, como advertía la lección 10. Gretl no se detiene: detecta la dependencia, *elimina* uno de los dos regresores y lo dice en la salida. Es la respuesta razonable —el ajuste $\hat{y}$ es único aunque los coeficientes no lo sean—, pero conviene saber que ha ocurrido: el coeficiente de C que queda ya no es “el efecto de C manteniendo F fija”, porque fijar F es fijar C . En la lección 14 veremos el caso menos limpio y más frecuente: regresores *casi* alineados, que Gretl no elimina y que hacen inestables los coeficientes.

Actividad 7 - Síntesis

Hoy hemos verificado con Gretl la maquinaria matricial completa de la lección 10: (i) que $\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ coincide con lo que calcula `ols`, y que $\mathbf{X}^T \hat{\mathbf{e}} = \mathbf{0}$ (que es lo mismo que $\hat{\mathbf{e}} \mathbf{X} = \mathbf{0}$); (ii) que con regresores mutuamente ortogonales cada coeficiente puede aislarse sin resolver ningún sistema, y R^2 se descompone como suma de correlaciones al cuadrado.

Y lo hemos verificado de una manera concreta que conviene retener, porque vale para cualquier experimento y no solo para este: comparando dos escenarios que se diferencian en *una sola cosa*. Si el caso ortogonal hubiera traído consigo otro tamaño muestral, otros coeficientes y otro nivel de ruido, habríamos visto cambiar los números sin poder atribuir el cambio a la ortogonalidad. Al dejar fijo todo lo demás —incluida la perturbación—, la única explicación posible de la diferencia entre los dos bloques de la Actividad 5 es la que queríamos estudiar.

Preguntas de interpretación para la clase

1. En la Actividad 2, ¿por qué la diferencia entre `beta_matricial` y `beta_ols` no es cero, sino un número muy pequeño?
2. ¿Por qué necesitamos que `x2` y `x3` estén correlacionados en la Actividad 1 para que el ejercicio tenga sentido pedagógico? ¿Qué habría pasado si, por descuido, hubieran resultado casi ortogonales?
3. En la Actividad 4, ¿por qué es importante que $\mathbf{z}_2$ y $\mathbf{z}_3$ tengan media *exactamente* cero, y no solo estén incorrelados?
4. En la Actividad 5, el cálculo “aislado” $\hat{\beta}_2 (\text{cov}(\mathbf{x}_2, \mathbf{y}) / \text{var}(\mathbf{x}_2))$ sobre los datos de la Actividad 1 (`x2`, `x3`, correlacionados) no coincide con el coeficiente de `x2` en el modelo conjunto `ols` y `const x2 x3`, y el de `x3` ni siquiera tiene el mismo signo. ¿Por qué?
5. ¿Diría que el uso de la inversa $(\mathbf{X}^\top \mathbf{X})^{-1}$ en esta práctica contradice la advertencia de la lección 10 de que el curso “nunca la usa como vía de razonamiento”? ¿Qué diferencia hay entre usarla para *verificar* una fórmula ya conocida y usarla como *método* para resolver MCO?
6. En la Actividad 6, ¿por qué el ajuste $\hat{\mathbf{y}}$ es único aunque los coeficientes no lo sean? ¿Qué subespacio generan $\mathbf{1}$, $\mathbf{C}$ y $\mathbf{F}$?

Para profundizar:

- Véase la lección 10 ([lección 10](#)) para la notación matricial completa, el operador selector y la transparencia “Familia ortogonal”.
- Cottrell, A. y Lucchetti, R. (2023). *Gretl User's Guide*. Sección sobre construcción y operaciones con matrices a partir de listas de series.

Código completo de la práctica

```
# -----
# Copyright (C) 2026 Marcos Bujosa
# Licencia: GNU General Public License v3.0 o posterior
# Este código es software libre y puede ser redistribuido y/o modificado bajo los términos de la GPL.
# Ver el archivo LICENSE del repositorio para más detalles.
# -----

# Los dos primeros comandos son necesarios para que Gretl guarde los resultados de la práctica en el
↳ directorio de trabajo
# al ejecutar lo siguiente desde un terminal (use los nombres y ruta que correspondan)
#
#   DIRECTORIO="Nombre_Directorio_trabajo" gretlcli -b ruta/nombre_fichero_de_la_practica.inp
#
# Si esto no le funciona en su sistema, comente las siguientes dos líneas y sitúese en el directorio de
↳ trabajo de gretl
# que corresponda (configure dicho directorio de trabajo desde la ventana principal de Gretl).

string directory = getenv("DIRECTORIO")
set workdir "@directory"

nulldata 40 --preserve
set seed 97531

series x2 = uniform(0,10)
series x3 = 0.6*x2 + normal(0,2)
series u = normal(0,3)

scalar b1 = 5
scalar b2 = 2
scalar b3 = -1
series y = b1 + b2*x2 + b3*x3 + u

printf "Correlacion entre x2 y x3 = %.4f (no son ortogonales)\n", corr(x2,x3)

outfile --quiet matrizXyEcuacionesNormales.txt
  list Xlist = const x2 x3
  matrix X = { Xlist }
  list ylist = y
  matrix Y = { ylist }

  matrix XtX = X'X
  matrix XtY = X'Y

  printf "Matriz X'X (3x3):\n"
  print XtX
  printf "\nVector X'y (3x1):\n"
  print XtY
end outfile

outfile --quiet betaMatricialVsOls.txt
  matrix beta_matricial = invpd(XtX)*XtY
  printf "beta calculado como (X'X)^(-1) X'y:\n"
  print beta_matricial

  ols y const x2 x3 --quiet
  matrix beta_ols = $coeff
  printf "\nbeta calculado por 'ols' de Gretl:\n"
  print beta_ols

  matrix diferencia = beta_matricial - beta_ols
  printf "\nDiferencia (debe ser practicamente cero):\n"
  print diferencia
end outfile
```

```

outfile --quiet verificacionXteYeX.txt
series ehat = $uhat
list elist = ehat
matrix E = { elist }

matrix XtE = X'E
printf "X'e (columna 3x1, debe ser (aprox.) el vector nulo):\n"
print XtE

matrix eX = E'X
printf "\ne*X (Gretl lo presenta como fila 1x3; mismos tres numeros):\n"
print eX
end outfile

outfile --quiet datosFamiliaOrtogonal.txt
series z2 = x2 - mean(x2)           # centrado: media nula
ols x3 const z2 --quiet
series z3 = $uhat                  # residuos: ortogonales a 1 y a z2 por construccion

series yo = b1 + b2*z2 + b3*z3 + u  # MISMOS beta, MISMA perturbacion u

printf "media(z2)    = %.10f    (debe ser 0)\n", mean(z2)
printf "media(z3)    = %.10f    (debe ser 0)\n", mean(z3)
printf "suma(z2*z3)  = %.10f    (debe ser 0: ortogonalidad mutua)\n", sum(z2*z3)
printf "corr(z2,z3)  = %.10f    (era 0,5046 entre x2 y x3)\n", corr(z2,z3)
end outfile

outfile --quiet familiaOrtogonalComparacion.txt
printf "===== CASO GENERAL (Actividad 1) =====\n"
printf "correlacion entre los regresores = %.4f\n\n", corr(x2,x3)
ols y const x2 x3 --quiet
printf "beta2: conjunto = %.10f    aislado (cov/var) = %.10f\n", $coeff(x2), cov(x2,y)/var(x2)
printf "beta3: conjunto = %.10f    aislado (cov/var) = %.10f\n", $coeff(x3), cov(x3,y)/var(x3)
printf "R2      : conjunto = %.10f    suma de rho^2      = %.10f\n", $rsq, corr(x2,y)^2 + corr(x3,y)^2

printf "\n===== CASO ORTOGONAL (Actividad 4) =====\n"
printf "correlacion entre los regresores = %.4f\n\n", corr(z2,z3)
ols yo const z2 z3 --quiet
printf "beta2: conjunto = %.10f    aislado (cov/var) = %.10f\n", $coeff(z2), cov(z2,yo)/var(z2)
printf "beta3: conjunto = %.10f    aislado (cov/var) = %.10f\n", $coeff(z3), cov(z3,yo)/var(z3)
printf "R2      : conjunto = %.10f    suma de rho^2      = %.10f\n", $rsq, corr(z2,yo)^2 +
↪ corr(z3,yo)^2
end outfile

outfile --quiet colinealidadExacta.txt
series C = x2                      # "temperatura" en Celsius
series F = 32 + 1.8*C              # la misma temperatura, en Fahrenheit
printf "corr(C, F) = %.10f\n", corr(C, F)

list Xcol = const C F
matrix Xc = { Xcol }
printf "determinante de X'X con C y F = %g    (una matriz invertible tiene determinante distinto de
↪ cero)\n", det(Xc'Xc)

ols y const C F
end outfile

```

Enlace al guión: [S14-Prct-D-simulados-matricial.inp](#)

Respuestas

1. ¿Por qué la diferencia no es cero? Por errores de redondeo en aritmética de coma flotante, la misma razón por la que $\sum \hat{e}_i$ no era cero en las prácticas anteriores. Invertir una matriz 3×3 y multiplicarla por otro vector implica varias operaciones aritméticas encadenadas, cada una con su pequeñísimo error de precisión; el resultado final coincide con el de `ols` hasta un margen del orden de 10^{-10} o menor, que es indistinguible de cero a efectos prácticos.
2. ¿Por qué necesitamos correlación entre `x2` y `x3`? Porque el objetivo pedagógico de la Actividad 1 es mostrar el caso *general* de la regresión múltiple, en el que resolver el sistema completo (o, equivalentemente, usar la matriz inversa) es imprescindible. Si `x2` y `x3` hubieran resultado casi ortogonales por accidente, el ejercicio se habría parecido demasiado al caso especial de la Actividad 4, y no habría quedado clara la diferencia entre “hay que resolver el sistema” (caso general) y “cada coeficiente se calcula aislado” (familia ortogonal): precisamente el contraste que se explota, con toda intención, entre las Actividades 1-3 y las Actividades 4-5.
3. ¿Por qué la media exactamente nula, no solo la incorrelación? Porque la fórmula de la transparencia “Familia ortogonal” (lección 10) exige que $\mathbf{1}, \mathbf{z}_2, \mathbf{z}_3$ sean *mutuamente* ortogonales: no basta con que $\mathbf{z}_2 \perp \mathbf{z}_3$, hace falta además que cada uno de ellos sea ortogonal a la constante $\mathbf{1}$, es decir, que tenga media cero. Sin esa condición, $\hat{\beta}_1$ (el coeficiente de la constante) no sería, en general, la media de $\mathbf{y}$, y las fórmulas aisladas de $\hat{\beta}_2, \hat{\beta}_3$ como covarianza dividida por varianza dejarían de coincidir con los coeficientes del modelo conjunto.
4. ¿Coincide el cálculo aislado con los datos de la Actividad 1? No, y ya no hace falta creerlo: la Actividad 5 lo calcula. La fórmula aislada $\hat{\beta}_2 = \sigma_{\mathbf{x}_2 \mathbf{y}} / \sigma_{\mathbf{x}_2}^2$ solo reproduce el coeficiente del modelo conjunto cuando los regresores son mutuamente ortogonales. Con `x2` y `x3` correlacionados, el cálculo aislado ignora la información compartida entre ambos y “atribuye” a uno parte de lo que corresponde al otro: $\hat{\beta}_2$ pasa de 2,11 a 1,68, y $\hat{\beta}_3$ —esto es lo llamativo— pasa de $-0,98$ a $+0,22$, cambiando de signo. Es el mismo fenómeno que se observó en la práctica de Ramanathan con `sqft` y `bedrms`, aquí con un mundo del que conocemos los coeficientes verdaderos y podemos por tanto ver quién acierta: el modelo conjunto ($-0,98$ frente a $\beta_3 = -1$).
5. ¿Contradice esta práctica la advertencia de la lección 10 sobre la inversa? No, y la distinción es importante. La advertencia de la lección 10 es que el curso nunca usa la inversa como *vía de razonamiento* para justificar por qué MCO funciona, ni como herramienta habitual de cálculo: el argumento conceptual sigue siendo, siempre, la condición de ortogonalidad $\mathbf{X}^\top \hat{\mathbf{e}} = \mathbf{0}$. Aquí hemos usado la inversa una única vez, con un propósito muy distinto: *verificar numéricamente* que una fórmula ya demostrada por otro camino (las ecuaciones normales, obtenidas de la ortogonalidad) da, en efecto, el mismo resultado que `ols`. Usarla como comprobación puntual de una fórmula ya justificada no es lo mismo que apoyarse en ella para *explicar* por qué MCO funciona.
6. *Ajuste único, coeficientes no.* $\mathbf{F} = 32\mathbf{1} + 1,8\mathbf{C}$ es combinación lineal de $\mathbf{1}$ y $\mathbf{C}$, así que $\mathcal{L}(\mathbf{1}, \mathbf{C}, \mathbf{F}) = \mathcal{L}(\mathbf{1}, \mathbf{C})$: un plano, no un subespacio de dimensión tres. La proyección de $\mathbf{y}$ sobre ese plano es única (es el punto más cercano), pero se puede escribir de infinitas maneras como combinación de tres generadores que solo aportan dos direcciones. Gretl elige una de esas maneras eliminando $\mathbf{F}$; cualquier otra daría el mismo $\hat{\mathbf{y}}$, los mismos residuos y el mismo R^2 .