

TAREA DE ESTIMACIÓN

A continuación se encuentran 5 ejercicios de los cuales el estudiante sólo debe entregar 3 de ellos. El primero es obligatorio para todos. Los otros dos ejercicios son asignado de la siguiente manera: **Adriangela: 2 y 4. Angel: 2 y 5. Ivonne: 3 y 5. Rossana: 3 y 4.**

1. Crea en R la siguiente matriz \mathbf{X} de orden 5×3 .

$$\begin{pmatrix} 3 & 5 & 1 \\ 4 & 2 & 2 \\ 3 & 1 & 3 \\ 2 & 6 & 3 \\ 0 & 1 & 4 \end{pmatrix}$$

- a) Forma la matriz $\mathbf{X}'\mathbf{X}$. ¿Es de rango completo la matriz $\mathbf{X}'\mathbf{X}$? ¿Qué te permite llegar a esta conclusión? ¿Existe la matriz $(\mathbf{X}'\mathbf{X})^{-1}$? Si existe, calcúlela
 - b) Calcula $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. ¿Qué orden tiene? ¿Es simétrica? ¿Idempotente?. Justifica.
 - c) Calcula $I - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. ¿Qué orden tiene? ¿Es simétrica? ¿Idempotente?. Justifica.
 - d) ¿Existe $[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']^{-1}$? ¿Qué te permite llegar a tal conclusión? ¿Existe la matriz $[I - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']^{-1}$?
2. El fichero de datos choco.dat, contiene el precio en euros en el 2006 por cada 100 gramos de distintas marcas de chocolate. Se recoge también su contenido porcentual en cacao. Con los datos disponibles, y prescindiendo de un posible efecto "marca", estima la influencia de las variables CacaoPC y Peso sobre el precio, usando el método de

Cholesky, operaciones matriciales con R y la función `lm()`. La data esta disponible en la página.

3. Un geyser es una fuente de agua caliente que, de tanto en tanto, experimenta variaciones de régimen y produce una erupción de vapor o agua sobre calentada. El mecanismo físico subyacente es complejo y no perfectamente predecible. La experiencia muestra, sin embargo, que aunque las erupciones no se producen a intervalos regulares, hay cierta relación entre la duración de una erupción y el tiempo que transcurre hasta la siguiente. Uno de los geyser más famosos está situado Yellowstone National Park, y se conoce como Old Faithful². Entra en erupción a intervalos de entre 40 y 100 minutos, y cada erupción dura entre 1 y 6 minutos. Los guardas han observado que, cuanto más larga es una erupción, más tarda en presentarse la siguiente, y emplean la fórmula empírica $T = 30 + 10d$, donde T es el tiempo aproximado hasta la siguiente erupción, y d es la duración de la erupción previa. ¿Puedes tú hacerlo mejor? En el fichero `geyser.dat` tienes datos correspondientes a 272 erupciones consecutivas, acaecidas en Octubre de 1.980. Las variables recogidas son duración e intervalo, ambas en minutos. Para su respuesta use el método de Cholesky, operaciones matriciales con R y la función `lm()`. La data esta disponible en la página.
4. El fichero de datos `gala`, contiene el numero de especies de tortugas en las distintas Islas Galápagos. Hayo 30 islas y siete variables en el fichero. Las variables son `Species` (Numero de especies de tortugas halladas en la isla), `Endemics` (el número de especies endémicas), `Area` (el área de la isla, km^2), `Elevation` (la elevación mas alta de la isla, m), `Nearest` (La distancia a la isla mas cercana, km), `Scruz` (la distancia desde la isla de santa Cruz, km) y `Adjacent` (el área de la isla adyacente, km). La data esta disponible en el paquete `faraway` de R, por lo tanto se debe cargar la librería usando el siguiente comando `library(faraway)`. Ajuste un modelo usando operaciones matriciales con R y la función `lm()`.
5. El fichero de datos `prostate`, tiene los datos de una muestra de los salarios semanales para los trabajadores masculinos de Estados Unidos de la Encuesta de Población en 1988. Entre las variables disponibles se encuentran `wage` (Salario semanal en dólares), `educ` (años de educación) y `exper` (años de experiencia). Ajuste un modelo usando operaciones matriciales con R y la función `lm()`. La data esta disponible en el paquete `faraway` de R, por lo tanto se debe cargar la librería usando el siguiente comando `library(faraway)`.