

Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

TAREA 1

EJERCICIOS

- Sean Y y Z dos variables aleatorias conjuntamente definidas. Supongamos que $E[Z] = 0$. Demuéstrese que $E[YZ] = \text{Cov}(Y, Z)$.
- Demuéstrese que si (X, Y) son dos variables aleatorias conjuntamente definidas, $E(Y) = E_X[E_{Y|X}[Y|X]]$.
- Sea el proceso autorregresivo,

$$y_t - 0,3y_{t-1} - 0,2y_{t-2} + 0,3y_{t-3} = \epsilon_t$$

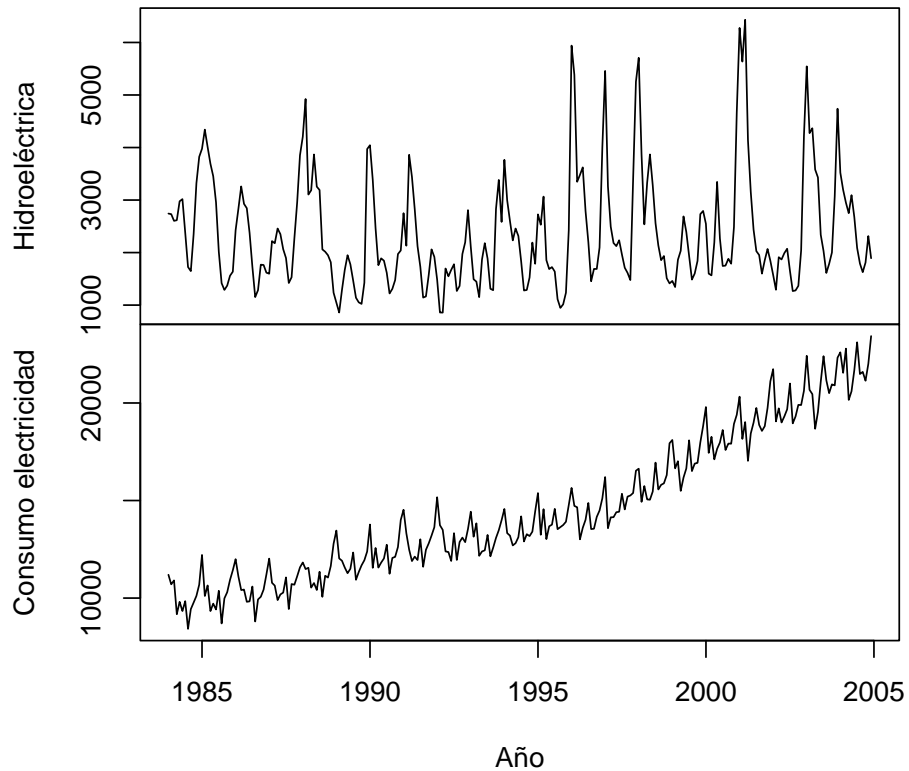
- Obtén las raíces de su polinomio característico. ¿Se trata de un proceso autorregresivo estacionario?
- Genera una muestra o realización del mismo de longitud $n = 500$, tras prescindir de un transitorio de longitud 50.
- Estima las funciones ACF y PACF y verifica que se adecúan a lo que la teoría sugiere.
- Estima los coeficientes del proceso y verifica que son aproximadamente los que has empleado en generar la muestra.
- Obtén los tres primeros términos de su representación en forma de media móvil,

$$y_t = \epsilon_t + \sum_{i=1}^{\infty} b_i \epsilon_{t-i}$$

si es que dicha representación existe.

- Obtén la varianza óptima del error de predicción uno, dos y tres periodos hacia adelante, en función de σ_ϵ^2 , varianza de la perturbación.
4. Considera los datos en las series temporales proporcionadas en el archivo `energia.dge` (lee con un `dget`). Se trata de datos mensuales de producción de energía hidroeléctrica y de consumo de energía eléctrica total en España. Sus gráficas aparecen a continuación.

**Producción hidroeléctrica y consumo eléctrico total
(en millones de Kw-h)**



- a) Haz un análisis de cada serie examinando su ACF y PACF. Justifica la necesidad o no de tomar diferencias, ordinarias o estacionales. También puede ser conveniente analizar las series en logaritmos (si es el caso, ¿por qué?).
- b) Sugiere un modelo (S)(AR)(I)(MA), estímalo y proporciona predicciones para los seis meses posteriores al fin de cada serie.
- Ajusta un modelo autorregresivo. Selecciona el orden con ayuda de un criterio como AIC. (Nota que el uso del criterio AIC sólo está realmente justificado cuando se emplea máxima verosimilitud como método de estimación).
 - Compara ambos modelos — el (S)(AR)(I)(MA) y el AR para cada serie. en términos de varianza residual y AIC.

AYUDAS, SUGERENCIAS Y COMPLEMENTOS

Objetivos:

Después de realizar esta tarea debes haber recordado métodos y conceptos de un curso previo de series temporales y estar en condiciones de llevar a término el análisis de una serie temporal univariante.

En el proceso de realización de la tarea habrás de familiarizarte, entre otras, con las siguientes funciones de R: `ts`, `filter`, `diff`, `ar`, `arima`, `predict`, `polyroot`, `Mod`.

Sobre R:

- `energia` es un objeto de clase `mts` (*multiple time series* o serie temporal multivariante). Las columnas de `energia` son objetos de clase `ts`. R soporta la noción de *clases* de objetos, que permite que funciones del mismo nombre (*funciones genéricas*) hagan diferentes cosas dependiendo de la clase de su(s) argumento(s). Compara por ejemplo el resultado de:

```
print(energia[,1])
plot(energia[,1])
```

(en que `print` y `plot` actúan sobre un objeto de clase `ts`) y

```
scratch <- as.vector(energia[,1])
print(scratch)
plot(scratch)
```

(en que lo hacen sobre un simple vector de datos). En lo posible te interesa trabajar sobre objetos de la clase adecuada, porque el comportamiento “de oficio” de las funciones que emplees será probablemente el que desees.

- Para convertir un vector en una serie temporal de clase `ts` cuentas con las funciones `ts` y `as.ts`. Has de señalar también la fecha comienzo y la periodicidad, con lo que los gráficos serán adecuadamente rotulados. Para agrupar varias series temporales univariantes sobre el mismo lapso de tiempo en una serie multivariante tienes la función `ts.union`. Observa que las clases `ts` y `mts` permiten representar series regularmente espaciadas. Tienes en R alternativas para representar series irregularmente espaciadas (paquetes `its`, `tseries` y `zoo`, por ejemplo).
- Para ajustar modelos ARIMA cuentas en R con la función `arima`. Para ajustar modelos autorregresivos podrías hacerlo con dicha función, pero cuentas con funciones especializadas como `ar` y compañeras: `ar.mle`, `ar.burg`, `ar.yw`, `ar.ols`. Estas últimas difieren en cuestiones de detalle (que tienen que ver con la estimación de la matriz de momentos de segundo orden, a partir de la cual se realiza la estimación). Las diferencias sobre las estimaciones serán en la mayoría de los casos mínimas, pero la diferencia en tiempo de cómputo es sustancial. Las funciones `ar` y asociadas permiten hacer uso cómodamente del criterio AIC para seleccionar el orden de un modelo autorregresivo.
- Tienes en R funciones que permiten calcular cómodamente las raíces de un polinomio, incluso complejas: haz `help(polyroot)`. Tienes también funciones para manejar números complejos, como `Mod`, que calcula el módulo de un número complejo (haz `help(complex)`). Puedes también recurrir a calcular los valores propios de una cierta matriz (la de la ecuación de estado) para ver si un proceso autorregresivo es o no estacionario.

Sobre los datos:

- Proceden del Banco de España. Pueden obtenerse de su página Web, pero aún más fácil es hacerlo desde el programa GRETL corriendo sobre una máquina con conexión a Internet. Puedes acceder a las series del Banco de España directamente y exportar las series leídas a formato legible directamente en R. Sobre GRETL, que puedes haber empleado en otras asignaturas, tienes información y documentación en <http://gretl.sourceforge.net>.
- Observa la previsible estacionalidad en la serie de producción hidroeléctrica, debido a la diferente hidraulicidad de los distintos meses del año. Observa también periodos de sequía y la práctica ausencia de tendencia: en la fecha de comienzo de la serie, el potencial hidroeléctrico del país estaba instalado prácticamente al completo.

Lectura sugerida:

- Sobre el uso de funciones para manipulación de series temporales cuentas con la ayuda *on line*. Te será de ayuda también [13], especialmente su Capítulo 13.
- Sobre cosas que debes ya haber visto de series temporales univariantes puedes emplear los libros que hayas manejado y tus apuntes. Una monografía muy breve y clara es [6]. Sobre la metodología Box-Jenkins tienes el libro clásico de sus proponentes, [5]. Pero no hay escasez de buenas y comprensivas monografías sobre series temporales: [11], [7], [10], [8], [12] son sólo algunas.
- Mencionamos el criterio AIC en clase. Artículos legibles son [2], [3] y [1]. Hay también un volumen completo con trabajos sobre AIC: [4].

Referencias

- [1] H. Akaike. Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, 22:203–217, 1970.
- [2] H. Akaike. Use of an information theoretic quantity for statistical model identification. In *Proc. 5th. Hawaii Int. Conf. on System Sciences*, pages 249–250, 1972.
- [3] H. Akaike. Information theory and an extension of the maximum likelihood principle. In B.N. Petrov and F. Csaki, editors, *Second International Symposium on Information Theory*, pages 267–281, Budapest, 1974. Akademia Kiado. Reimpreso en Johnson-Kotz(1991), vol. 1, p. 610 y ss.
- [4] H. Akaike. Information theory and an extension of the maximum likelihood principle. In Johnson and Kotz [9], page 610 y ss.
- [5] G.E.P. Box and G.M. Jenkins. *Time series Analysis, Forecasting and Control*. Holden Day, second edition, 1970. Signatura: 519.246.8 BOX.
- [6] C. Chatfield. *The Analysis of Time Series: an introduction*. Chapman and Hall, 1980.
- [7] W.A. Fuller. *Introduction to Statistical Time Series*. Wiley, New York, second edition, 1996.
- [8] J.D. Hamilton. *Time Series Analysis*. Princeton, 1994. Signatura: 519.246.8 HAM.
- [9] Johnson and Kotz, editors. *Breakthroughs in Statistics*. Springer Verlag, 1991.
- [10] H. Lütkepohl. *Introduction to Multiple Time Series Analysis*. Springer-Verlag, 1993. Signatura: 519.246.8 LUT.
- [11] G.C. Reinsel. *Elements of Multivariate Time Series Analysis*. Springer, 1993. Signatura: 519.246.8 REI.
- [12] R.H. Shumway and D.S. Stoffer. *Time Series Analysis and Its Applications*. Springer Verlag, 2000. Signatura: 519.246.8 SHU.
- [13] W.N. Venables and B.D. Ripley. *Modern Applied Statistics with S-PLUS*. Springer-Verlag, New York, third edition, 1999.