

Ejercicio 1

Considera una serie de tiempo estacionaria e invertible X_t dada por

$$X_t = \sum_{j=1}^{\infty} \omega_j e_{t-j} + e_t, \quad \sum_{j=1}^{\infty} \omega_j^2 < \infty, \quad e_t \sim \text{no-correlacionadas}(0, \sigma^2),$$

$$X_t = \sum_{j=1}^{\infty} \pi_j X_{t-j} + e_t, \quad \sum_{j=1}^{\infty} |\pi_j| < \infty.$$

Muestre que $\mathbb{E} \left[\left(\hat{X}_{t,n} - \sum_{j=1}^{\infty} \pi_j X_{t-j} \right)^2 \right] \rightarrow 0$ cuando $n \rightarrow \infty$, donde $\hat{X}_{t,n}$ es el mejor estimador insesgado de X_t basado en X_{t-1}, \dots, X_{t-n} .

Solución:

Como $\hat{X}_{t,n}$ es un estimador insesgado

$$\mathbb{E} [\hat{X}_{t,n}] = X_t$$

entonces sumamos y restamos

$$\mathbb{E} \left[\left(\hat{X}_{t,n} - \sum_{j=1}^{\infty} \pi_j X_{t-j} \right)^2 \right] = \mathbb{E} \left[\left((\hat{X}_{t,n} - X_t) - \left(\sum_{j=1}^{\infty} \pi_j X_{t-j} - X_t \right) \right)^2 \right],$$

desarrollando el cuadrado

$$\mathbb{E} \left[(\hat{X}_{t,n} - X_t)^2 \right] - 2\mathbb{E} \left[(\hat{X}_{t,n} - X_t) \left(\sum_{j=1}^{\infty} \pi_j X_{t-j} - X_t \right) \right] + \mathbb{E} \left[\left(\sum_{j=1}^{\infty} \pi_j X_{t-j} - X_t \right)^2 \right] \quad (1)$$

Sabemos que

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[(\hat{X}_{t,n} - X_t)^2 \right] = \sigma^2 \quad (2)$$

y que

$$\mathbb{E} \left[\left(\sum_{j=1}^{\infty} \pi_j X_{t-j} - X_t \right)^2 \right] = \mathbb{E} [e_t^2] = \sigma^2 \quad (3)$$

donde para este último se usó la expresión para el AR(∞).

Para el término de enmedio,

$$\begin{aligned} -2\mathbb{E} \left[(\hat{X}_{t,n} - X_t) \left(\sum_{j=1}^{\infty} \pi_j X_{t-j} - X_t \right) \right] &= \\ &= -2\mathbb{E} [(\hat{X}_{t,n} - X_t)e_t], \end{aligned}$$

Pero ya hemos mostrado, Notas pág. 127 que $Y_t - \hat{Y}_{n,t} = e(n)$ converge en media cuadrática a una sucesión de v.a. no-correlacionadas $(0, \sigma^2)$ cuando n tiende a infinito si Y_t es una serie regular. Otra forma de decirlo es $e(n) \xrightarrow{MC} e_t$.

Primeramente, para que Y_t sea una serie regular pedimos que $\sigma \neq 0$. Luego,

$$-2 \lim_{n \rightarrow \infty} \mathbb{E} [(\hat{X}_{t,n} - X_t)e_t] = -2\mathbb{E} [e_t^2] = -2\sigma^2 \quad (4)$$

juntando (2) (3) (4) en (1) se obtiene que

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\left(\hat{X}_{t,n} - \sum_{j=1}^{\infty} \pi_j X_{t-j} \right)^2 \right] = \sigma^2 - 2\sigma^2 + \sigma^2 = 0$$

lo que concluye la prueba.

Ejercicio 2

Muestra que si una serie de tiempo estacionaria satisface la ecuación en diferencias

$$Y_t - Y_{t-1} = e_t,$$

entonces $\mathbb{E} [e_t^2] = 0$.

Solución:

Observemos que el modelo propuesto es un ARIMA(0,1,0). Veamos que por sustitución recursiva se tiene

$$\begin{aligned} Y_t &= e_t + Y_{t-1}, \\ &= e_t + e_{t-1} + Y_{t-2}, \\ &= e_t + e_{t-1} + e_{t-2} + Y_{t-3}, \\ &\vdots \\ &= \sum_{j=1}^t e_j + Y_0 \end{aligned} \quad (5)$$

que de hecho es un modelo de caminata aleatoria si $Y_0 = 0$. Por tanto, como la serie es estacionaria por hipótesis, entonces tanto la esperanza como la varianza no pueden depender del tiempo explícitamente.

Por tanto, de (5) se sigue que

$$\mathbb{E} [Y_t] = \sum_{j=1}^t \mathbb{E} [e_j] = 0,$$

y

$$\begin{aligned} \mathbb{V}ar[Y_t] &= \sum_{j=1}^t \mathbb{V}ar[e_j], \\ &= t\sigma^2, \end{aligned}$$

entonces para que la varianza sea independiente del tiempo ya que la serie es estacionaria se requiere que $\sigma^2 = 0$. Lo que es lo mismo a pedir que $\mathbb{V}ar[e_t] = \mathbb{E} [e_t^2] = 0$.

Ejercicio 3

Sea Y_t un proceso estacionario con media μ (desconocida) y función de autocovarianza $\gamma(h)$ conocida. Dado $(Y_1, \dots, Y_n)' = \tilde{Y}$ con $\mathbb{V}ar[\tilde{Y}] = V$ no-singular.

1. Muestra que el mejor predictor lineal insesgado de μ es $\hat{\mu} = (\mathbf{1}'V^{-1}\mathbf{1})^{-1} \mathbf{1}'V^{-1}\tilde{Y}$ con $\mathbf{1}$ vector columna de unos.
2. Sea $\tilde{c} = Cov(\tilde{Y}, Y_{n+1})$ y $b = V^{-1}\tilde{c}$. Sea $\tilde{Y}_{n+1} = \tilde{b}'\tilde{Y}$ (que es el mejor estimador lineal insesgado (BLUP) para \tilde{Y} cuando $\mu = 0$). Muestra que el BLUP para Y_{n+1} es $\hat{Y}_{n+1} = a_1\tilde{Y}_{n+1} + a_2\hat{\mu}$, donde $a_1 = 1$ y $a_2 = (1 - \mathbf{1}'V^{-1}\tilde{c})$.

3. Calcula $\mathbb{E} \left[(Y_{n+1} - \hat{Y}_{n+1})^2 \right]$.

4. Supón que $Y_t = \mu + e_t + e_{t-1}$, μ desconocida. Dado Y_1, Y_2, Y_3 , predice Y_4 .

Solución:

Ejercicio 4

Simula un modelo ARMA(1,1) de longitud $n = 100$, con coeficientes autorregresivo $\phi = 0.8$ y $\theta = 0.4$. Fija la semilla en 910. Para este ejercicio puedes usar R o Python.

1. Calcula y grafica la función de autocorrelación (ACF) teórica. Grafica suficientes lags (retraso) hasta que las correlaciones se desvanezcan. Calcula y grafica la ACF muestral de la muestra simulada. ¿Qué tan bien ajustan los valores y patrones de la ACFs teórica y muestral?
2. Lee sobre la función de autocorrelación extendida (EACF) en el libro “Time Series Analysis with Applications in R” de Jonathan D. Cryer y King-sik Chan (disponible en el Springer Link). Calcula e interpreta la EACF muestral para la serie simulada. ¿La EACF te permite identificar los órdenes correctos para el modelo?
3. Repite las partes (a) y (b) con una nueva simulación usando los mismos parámetros y tamaño de muestra.
4. Repite las partes (a) y (b) con una nueva simulación usando los mismos parámetros y un tamaño de muestra $n = 48$. Vuelve a repetir la simulación pero ahora con un tamaño de muestra $n = 200$. Comenta al respecto.

Solución:

1. Sea X_t un ARMA(1,1)

$$X_t - \phi X_{t-1} = e_t + \theta e_{t-1},$$

con $|\phi| < 1$ y $|\theta| < 1$ entonces de la pág. 72 de [3] la función de autocorrelación es

$$\rho(h) = \begin{cases} 1, & h = 0 \\ \frac{(1+\theta\phi)(\phi+\theta)}{1+\theta^2+2\theta\phi} \phi^{h-1}, & h = 1, 2, \dots \end{cases} \quad (6)$$

Así para los parámetros dados, tenemos que

$$\rho(h) = \begin{cases} 1, & h = 0 \\ 0.88(0.8)^h, & h = 1, 2, \dots \end{cases} \quad (7)$$

Para dar intuición, podemos considerar la expansión

$$\phi(B)X_t = \theta(B)e_t$$

con $\phi(z) = 1 - 0.8z$ y $\theta(z) = 1 + 0.4z$. Y por la representación del ARMA

$$X_t = \psi(B)e_t = \frac{\theta(B)}{\phi(B)}e_t$$

Por tanto, sabemos de [1] que la correlación para una serie de tiempo causal es

$$\rho(h) = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+h}}{\sum_{i=0}^{\infty} \psi_i^2} \quad (8)$$

donde ψ_i son tales que $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$. Por tanto, analizamos

$$\begin{aligned}\psi(z) &= \frac{1 + 0.4z}{1 - 0.8z}, \\ &= (1 + 0.4) \left(1 + 0.8z + (0.8^2) z^2 + (0.8)^3 z^3 + \dots \right), \\ &= 1 + (0.8 + 0.4)z + (0.8)^2(0.8 + 0.4)z^2 + (0.8)^3(0.8 + 0.4)z^3 + \dots\end{aligned}$$

por tanto, $\psi_0 = 1$ y $\psi_i = 1.2(0.8)^{i-1}$. Sustituyendo estos valores en (8) se obtiene (7).

Para la simulación de la un ARMA(1,1) consideramos la implementación en R con la librería nativa de R stats (no es necesario cargar). Primeramente ajustamos la semilla dada con el comando

```
|| set.seed(910)
```

Luego, simulamos el ARMA(1,1) y se gráfica la serie con el comando

```
|| x <- arima.sim(n=100, model=list(ar=0.8, ma=0.4))
|| plot.ts(x)
```

mostrada en la siguiente figura.

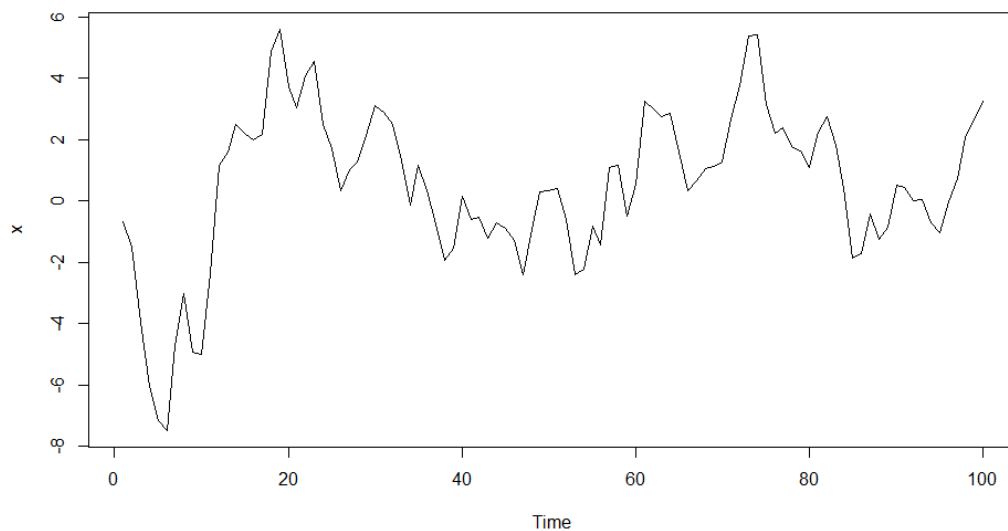


Figura 1: Simulación del ARMA(1,1) con $\phi = 0.8$ y $\theta = 0.4$

Notemos que la simulación tiene un valle en valores de t cercanos a 10, sin embargo no es problema porque parece que la simulación es homocedástica, por estacionaria, aunque buena parte de la serie se concentre por encima de 0. Además, se gráfica la función de autocorrelación muestral llamando $acf(x)$ junto con la función de autocorrelación teórica.

```
|| acf(x)
|| h <- 0:20
|| lines(h, 0.88*(0.8)**(h-1), col="red")
```

Obtenemos

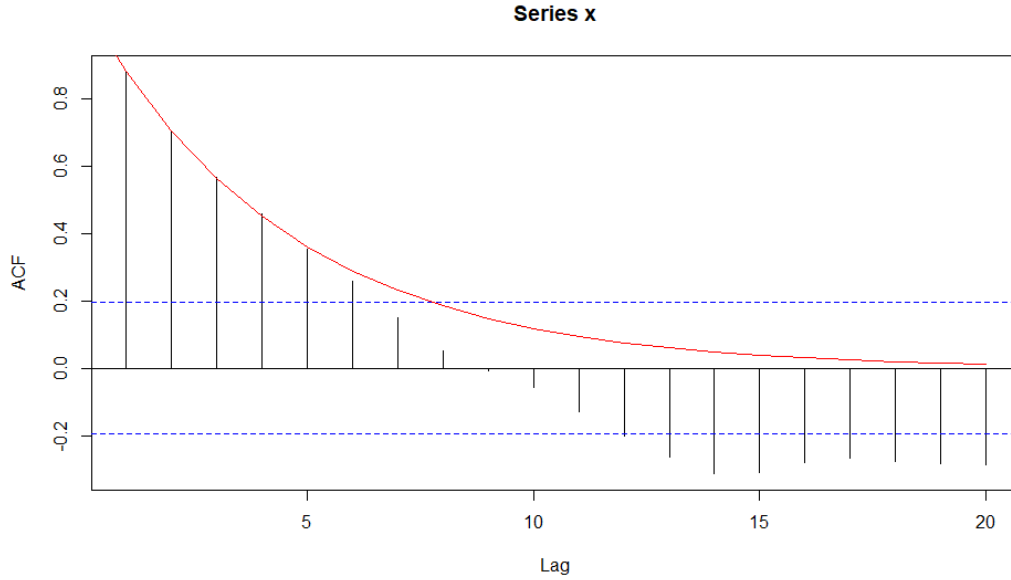


Figura 2: Función de autocorrelación muestral para la simulación del ARMA(1,1) y función de autocorrelación teórica.

Se gráfica hasta el lag 20 y notemos que tenemos un cambio de signo para la autocorrelación para lag cercano a 15. Pero para valores de lag chicos el ajuste muestral contra el teórico es plausible.

2. Las ACF y PACF son útiles para identificar parámetros de procesos AR(p) y MA(q) puros. Sin embargo, para modelos mixtos se tiene una cantidad numerable de ceros, lo que hace difícil identificar los modelos de sus respectivas funciones muestrales.

El método de la función de autocorrelación extendida usa el hecho de que si se conoce la parte autoregresiva del modelo ARMA, entonces podemos filtrarlo en un modelo MA puro. Los coeficientes del AR se estiman por una secuencia finita de regresiones.

Consideremos

$$Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1}$$

En este caso, una simple regresión lineal de Y_t en Y_{t-1} resulta en una inconsistencia en el estimador de ϕ . De hecho, de la ecuación (6) sabemos que $\rho(1) = (\phi - \theta) / (1 - 2\phi\theta + \theta^2)$. El coeficiente de Y_{t-1} en la segunda regresión, denotado por $\tilde{\phi}$ es un estimador consistente de ϕ . Definamos $W_t = Y_t - \tilde{\phi}Y_{t-1}$, que es un modelo MA(1) aproximadamente. Para un ARMA(2,1), la tercera regresión de Y_t en su primer lag con el residual del segundo lag de la primera regresión no da el coeficiente de Y_{t-1} como estimador consistente de ϕ . Similarmente, los coeficientes de un ARMA(p,q) pueden ser estimados consistentemente via una secuencia de q regresiones.

Como los ordenes del AR y MA son desconocidos, se requiere de un método iterativo. Sea

$$W_{t,k,j} = Y_t - \tilde{\phi}_1 Y_{t-1} - \cdots - \tilde{\phi}_k Y_{t-k}$$

los residuales del modelo autorregresivo con los coeficientes estimados iterativamente suponiendo que AR es de orden k y el orden del MA es j . Para $k = p$ y $j \geq q$, $\{W_{t,j,k}\}$ es aproximadamente un modelo MA(q), entonces su autocorrelación teórica de lag $q + 1$ o mayor son iguales a cero. Para $k > p$, ocurre una sobreparametrización del problema, y esto incrementa el orden del MA para el proceso W por el mínimo de $k - p$ y $j - p$. Tsay y Tiao (1984) sugieren resumir la información de la función de autocorrelación extendida muestral en una tabla con el elemento en la k -ésima fila y j -ésima columna igual al símbolo \times si el lag $j + 1$ de la correlación muestral de $W_{t,k,j}$ es significativamente diferente de 0 (es decir, con magnitud mayor que $1.96 / \sqrt{n - j - k}$), en caso contrario se anota un 0.

En dicha tabla, un proceso ARMA(p,q) tendrá un patrón teórico de triángulo de ceros, con vértice en la parte superior izquierda que corresponde a los ordenes del ARMA. [4]

Para el caso particular de la simulación, obtenemos que la tabla generada por la EACF es

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	o	o	o	o	o	o	x	x
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	o	o	o	o	o	o	o	o	o	o	o	o	o
3	x	o	x	o	o	o	o	o	o	o	o	o	o	o
4	o	x	o	o	o	o	o	o	o	o	o	o	o	o
5	x	o	o	x	o	o	o	o	o	o	o	o	o	o
6	x	o	x	o	o	o	o	o	o	o	o	o	o	o
7	o	x	x	x	o	o	o	o	o	o	o	o	o	o

Figura 3: Tabla de la función de autocorrelación extendida.

Aunque es difícil detectar el patrón triangular, intrínsecamente sabemos que debería ser un triángulo con punta en 1,1 por la simulación. Vemos que se puede ver una separación diagonal, sin embargo hay muchos ceros debajo de ella. Por tanto, me parece que la estimación en este caso según el EACF es (1,1)

3. Con una nueva semilla, volvemos a simular a 100 pasos para el mismo ARMA(1,1). Esta vez se obtiene la gráfica de la función de autocorrelación siguiente

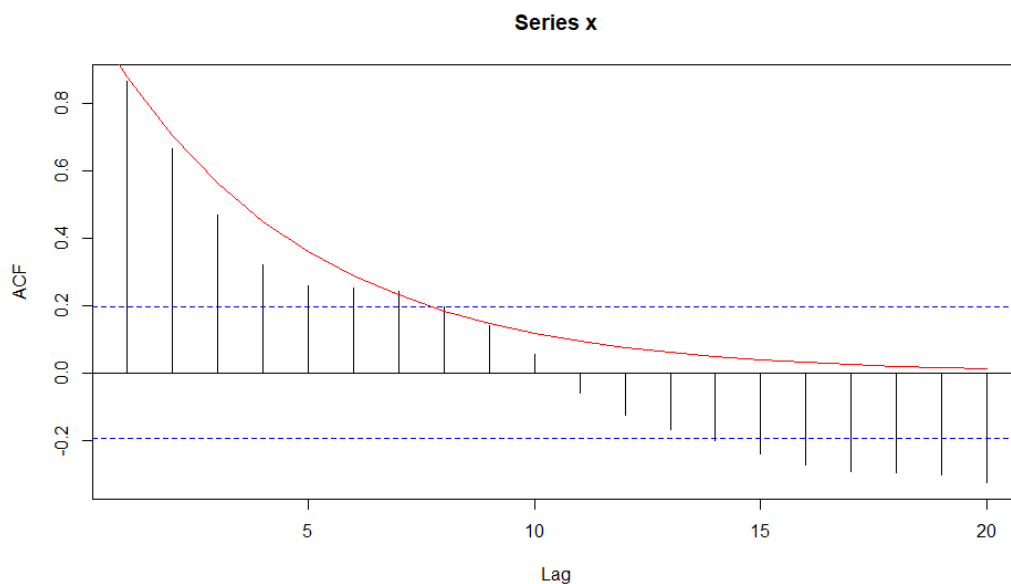


Figura 4: Función de autocorrelación muestral y teórica para otra simulación del ARMA(1,1).

La tabla generado por el EACF es

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	0	0	0	0	0	0	0
1	x	x	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	x	0	0	0	x	0	0	0
3	x	0	0	0	0	0	0	0	0	0	x	0	0	0
4	x	x	x	0	0	0	0	0	0	0	x	0	0	0
5	0	x	x	0	0	0	0	0	0	0	0	0	0	0
6	0	0	x	0	0	0	0	0	0	0	x	0	0	0
7	x	0	x	0	0	0	0	0	0	x	x	0	0	0

Figura 5: Tabla de la función de autocorrelación extendida.

En este caso, no es claro la división triangular observamos que definitivamente por la parte superior el p se estima en 1 sin embargo para q no es claro si se toma 0 o 1.

4. Para la simulación con $n = 48$ se obtiene la siguiente función de autocorrelación

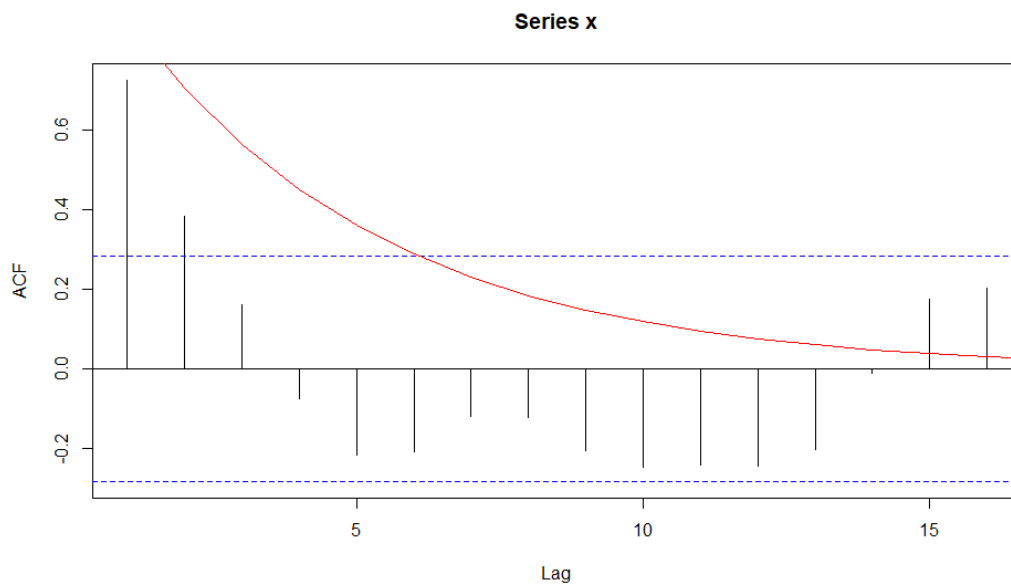


Figura 6: Función de autocorrelación muestral y teórica para otra simulación del ARMA(1,1).

Notamos un ajuste bastante pobre de la AFC muestral respecto al teórico.

Ahora, para el EACF

AR/MA		0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	0	0	0	0	0	0	0	0	0	0	0	0	0
1	x	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	x	x	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	x	0	0	0	0	x	0	0	0	0	0	0	0	0	0

Figura 7: Tabla de la función de autocorrelación extendida.

notamos que, aunque hay bastantes ceros, se alcanza a distinguir el pico en el (1,1) que es la estimación correcta.

Para $n = 200$ con la misma semilla, la función de autocorrelación es

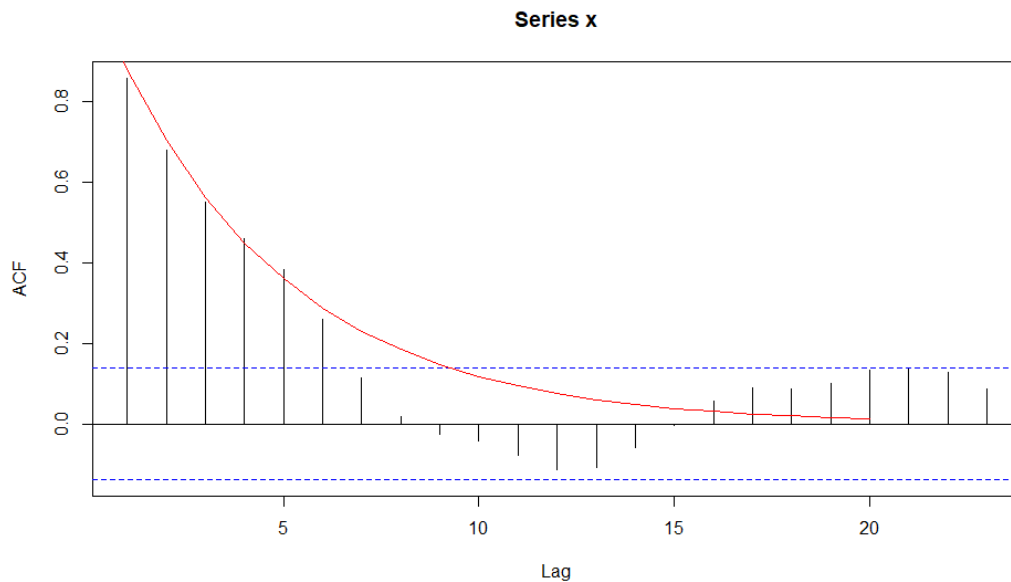


Figura 8: Función de autocorrelación muestral y teórica para otra simulación del ARMA(1,1).

que se observa un mejor ajuste entre el muestral y el teórico. Luego, en la tabla asociada a el EACF es

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	0	0	0	0	0	0	0	0
1	x	0	0	0	x	x	0	0	0	0	0	0	0	0
2	x	x	0	0	x	x	0	0	0	0	0	0	0	0
3	x	x	0	0	0	x	x	0	0	0	0	0	0	0
4	x	0	0	0	0	x	0	0	0	0	0	0	0	0
5	0	0	x	x	0	x	x	x	0	0	0	0	0	0
6	x	x	x	x	x	0	0	x	0	0	0	0	0	0
7	x	0	0	0	0	x	0	x	0	0	0	0	0	0

Figura 9: Tabla de la función de autocorrelación extendida.

Nuevamente, notamos que la fila superior esta bien definida, la diagonal se alcanza a distinguir, luego el estimador queda nuevamente en el (1,1) que es la estimación correcta.

Ejercicio 6

Sea $\{X_t : t = 0, \pm 1, \pm 2, \dots\}$ definida como

$$X_t + \alpha X_{t-1} + \alpha_2 X_{t-2} = e_t \quad e_t \sim \text{no-correlacionados}(0, \sigma^2).$$

Sean r_1, r_2 las raíces (reales) con $|r_i| < 1, i = 1, 2$. Muestra que

$$X_t - r_1 X_{t-1} = Y_t$$

es un AR(1) con parámetro r_2 .

Reescribamos la serie de tiempo de la siguiente manera

$$\begin{aligned} e_t &= X_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} \\ &= (1 + \alpha_1 B + \alpha_2 B^2) X_t \end{aligned}$$

donde B es el operador *backshift*. Tenemos el polinomio característico $\phi(z) = 1 + \alpha_1 z + \alpha_2 z^2$ con raíces reales r_1, r_2 . Entonces, de algebra elemental podemos escribir el polinomio característico como

$$\begin{aligned} \phi(z) &= 1 - (r_1 + r_2)z + r_1 r_2 z^2 \\ &= (1 - r_1 z)(1 - r_2 z) \end{aligned}$$

Como $|r_1| < 1$ y $|r_2| < 1$ por hipótesis, entonces

$$\begin{aligned} 0 &= 1 - r_1 z_1 \Leftrightarrow z_1 = \frac{1}{r_1} \Leftrightarrow |z_1| > 1 \\ 0 &= 1 - r_2 z_2 \Leftrightarrow z_2 = \frac{1}{r_2} \Leftrightarrow |z_2| > 1 \end{aligned}$$

Recordemos en [1] el siguiente:

Teorema 3.2

Un proceso AR(p) es causal si las raíces del polinomio característico $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ todas están fuera del círculo unitario.

Por tanto, X_t es un proceso causal. Recordemos además, que un proceso se dice causal si existen ψ_j con $\sum_{j=0}^{\infty} |\psi_j| < \infty$ tal que

$$Y_t = \sum_{j=0}^{\infty} \psi_j e_{t-j}$$

Particularmente, hemos visto que la expansión para $Y_t = \phi Y_{t-1} + e_t$ un AR(1) causal de parámetro ϕ ($|\phi| < 1$) tiene como expansión

$$Y_t = \sum_{j=0}^{\infty} \phi^j e_{t-j} \quad (9)$$

es decir, $\psi_j = \phi^j$, pag. 30 [1].

Luego, volviendo a la serie de tiempo de interés.

$$X_t = \phi(B)^{-1} e_t = \frac{1}{(1 - r_1 B)(1 - r_2 B)} e_t.$$

Por fracciones parciales

$$\begin{aligned} \frac{1}{(1 - r_1 z)(1 - r_2 z)} &= \frac{A}{1 - r_1 z} + \frac{B}{1 - r_2 z}, \\ &= \frac{A(1 - r_2 z) + B(1 - r_1 z)}{(1 - r_1 z)(1 - r_2 z)} \end{aligned}$$

se sigue

$$1 = A(1 - r_2 z) + B(1 - r_1 z)$$

Resolviendo como es usual, propongamos $z = 1/r_2$ entonces

$$1 = B \left(1 - \frac{r_1}{r_2} \right) \Leftrightarrow B = \frac{r_2}{r_2 - r_1}$$

proponiendo $z = 1/r_1$

$$1 = A \left(1 - \frac{r_2}{r_1} \right) \Leftrightarrow A = -\frac{r_1}{r_2 - r_1}$$

entonces

$$\frac{1}{(1 - r_1 z)(1 - r_2 z)} = \frac{1}{r_2 - r_1} \left(\frac{r_2}{1 - r_2 z} - \frac{r_1}{1 - r_1 z} \right)$$

La serie de tiempo queda entonces, expresada como

$$\begin{aligned} X_t &= \frac{1}{(1 - r_1 B)(1 - r_2 B)} e_t \\ &= \frac{1}{r_2 - r_1} \left(\frac{r_2}{1 - r_2 B} - \frac{r_1}{1 - r_1 B} \right) e_t \\ &= \frac{1}{r_2 - r_1} \left(r_2 \sum_{j=0}^{\infty} r_2^j B^j - r_1 \sum_{j=0}^{\infty} r_1^j B^j \right) e_t \\ &= \frac{1}{r_2 - r_1} \sum_{j=0}^{\infty} (r_2^{j+1} - r_1^{j+1}) e_{t-j} \end{aligned} \quad (10)$$

donde podemos notar que $\psi_j = (r_2^{j+1} - r_1^{j+1}) / (r_2 - r_1)$ para la expansión dada por la definición de causalidad.

Ahora, analizamos el siguiente proceso

$$Y_t = X_t - r_1 X_{t-1}$$

de la expansión en (10)

$$\begin{aligned} Y_t &= \frac{1}{r_2 - r_1} \sum_{j=0}^{\infty} (r_2^{j+1} - r_1^{j+1}) e_{t-j} - r_1 \frac{1}{r_2 - r_1} \sum_{j=0}^{\infty} (r_2^{j+1} - r_1^{j+1}) e_{t-1-j} \\ &= \frac{1}{r_2 - r_1} \left(\sum_{j=0}^{\infty} (r_2^{j+1} - r_1^{j+1}) e_{t-j} - \sum_{j=0}^{\infty} (r_1 r_2^{j+1} - r_1^{j+2}) e_{t-(j+1)} \right) \\ &= \frac{1}{r_2 - r_1} \left(\sum_{j=0}^{\infty} (r_2^{j+1} - r_1^{j+1}) e_{t-j} - \sum_{j=1}^{\infty} (r_1 r_2^j - r_1^{j+1}) e_{t-j} \right) \end{aligned}$$

donde se cambio la indentación a la segunda suma. Abriendo a el primer término para la primer suma y cancelando con la segunda

$$\begin{aligned} Y_t &= \frac{1}{r_2 - r_1} \left((r_2 - r_1) e_t + \sum_{j=1}^{\infty} (r_2^{j+1} - r_1^{j+1}) e_{t-j} - \sum_{j=1}^{\infty} (r_1 r_2^j - r_1^{j+1}) e_{t-j} \right) \\ &= \frac{1}{r_2 - r_1} \left((r_2 - r_1) e_t + \sum_{j=1}^{\infty} r_2^{j+1} e_{t-j} - \sum_{j=1}^{\infty} r_1 r_2^j e_{t-j} \right) \\ &= \frac{1}{r_2 - r_1} \left((r_2 - r_1) e_t + \sum_{j=1}^{\infty} (r_2 - r_1) r_2^j e_{t-j} \right) \\ &= e_t + \sum_{j=1}^{\infty} r_2^j e_{t-j} \\ &= \sum_{j=0}^{\infty} r_2^j e_{t-j} \end{aligned}$$

Por tanto, tenemos la expansión de un proceso causal. De la relación (9) podemos notar que Y_t es un AR(1) de parámetro $\phi = r_2$ y

$$Y_t = r_2 Y_{t-1} + e_t$$

lo que concluye la demostración.

Ejercicio 7

Define el proceso $\{X_t\}$ como $X_t = \sum_{j=0}^{\infty} \rho^j e_{t-j} (= \rho X_{t-1} + e_t)$, donde $e_t \sim$ no-correlacionadas $(0, \sigma_e^2)$ y $|\rho| < 1$. Sea u_t una sucesión de v.a. no-correlacionadas $(0, \sigma_u^2)$ independientes de $\{e_t\}$. Sea $Y_t = X_t + u_t$ (i.e. Y_t es X_t medida con un error u_t).

1. Encuentra la función de autocovarianza de Y_t .
2. Encuentra un proceso ARMA(1,1) $Z_t = \alpha_1 Z_{t-1} + \beta_1 e_{t-1} + e_t$, donde $e_t \sim$ no-correlacionadas $(0, \sigma_e^2)$ tal que $\gamma_Z(h) = \gamma_Y(h)$ para toda h .
3. Si X_t fuera un ARMA(p,q) estacionario e invertible, ¿qué puedes decir de $Y_t = X_t + u_t$? donde u_t es WN $(0, \sigma_u^2)$ independiente de X_t .

Solución:

1. Como X_t es un AR(1) causal ya hemos calculado su función de covarianza, esta es

$$\gamma_X(h) = \sigma_e^2 \frac{\rho^h}{1 - \rho^2}$$

Luego, calculando la función de covarianza para Y_t

$$\begin{aligned}\gamma_Y(h) &= \text{Cov}(Y_t, Y_{t+h}), \\ &= \text{Cov}(X_t + u_t, X_{t+h} + u_{t+h}), \\ &= \text{Cov}(X_t, X_{t+h}) + \text{Cov}(X_t, u_{t+h}) + \text{Cov}(u_t, X_{t+h}) + \text{Cov}(u_t, u_{t+h})\end{aligned}$$

como X_t es independiente de u_t

$$\begin{aligned}\gamma_Y(h) &= \gamma_X(h) + \text{Cov}(u_t, u_{t+h}), \\ &= \sigma_e^2 \frac{\rho^h}{1 - \rho^2} + \sigma_u^2 \delta(h).\end{aligned}\tag{11}$$

2. Para este caso, de pag. 72 en [3] tenemos que la función de autocovarianza para el ARMA(1,1) invertible y causal

$$Z_t = -\alpha Z_{t-1} + \varepsilon_t + \beta \varepsilon_{t-1}$$

es

$$\gamma_Z(h) = \begin{cases} \frac{1+\beta^2+2\beta\alpha}{1-\alpha^2} \sigma_\varepsilon^2, & h = 0 \\ \frac{(1+\beta\alpha)(\alpha+\beta)}{1-\alpha^2} \alpha^{h-1} \sigma_\varepsilon^2, & h = 1, 2, \dots \end{cases}\tag{12}$$

por lo que el proceso será tal que resuelva el sistema de ecuaciones comparando (11) y (12).

3. Como X_t es un ARMA(p,q) entonces existen polinomios $\phi(B), \theta(B)$ de orden p, q , respectivamente con raíces fuera del círculo unitario tal que

$$\phi(B)X_t = \theta(B)e_t.$$

Entonces

$$\begin{aligned}\phi(B)Y_t &= \phi(B)(X_t + u_t), \\ &= \phi(B)X_t + \phi(B)u_t, \\ &= \theta(B)e_t + \phi(B)u_t,\end{aligned}$$

por tanto, del lado izquierdo mantenemos un AR(p) mientras que en el lado derecho tenemos un MA(r) donde $r = \max\{p, q\}$ ya que tenemos la suma de dos MA independientes.

Ejercicio 8

El conjunto de datos deere3 contiene 57 mediciones consecutivas registradas en una máquina de la compañía Deere and Co. Los valores están dados en desviaciones de un valor ideal en diezmillonésimas de una pulgada. El proceso emplea un mecanismo de control que reinicia algunos parámetros de la máquina dependiendo de la magnitud de la desviación del valor ideal del último ítem producido. Los datos están disponibles en el paquete TSA de R.

1. Grafica la ST y comenta al respecto. ¿Un modelo estacionario es apropiado?
2. Grafica la ACF y PACF para esta serie y selecciona los órdenes tentativos de un modelo ARMA para la serie.
3. Selecciona el orden empleando un criterio distinto o alguno que complemente el inciso anterior.

Solución:

La grafica de la serie de tiempo es

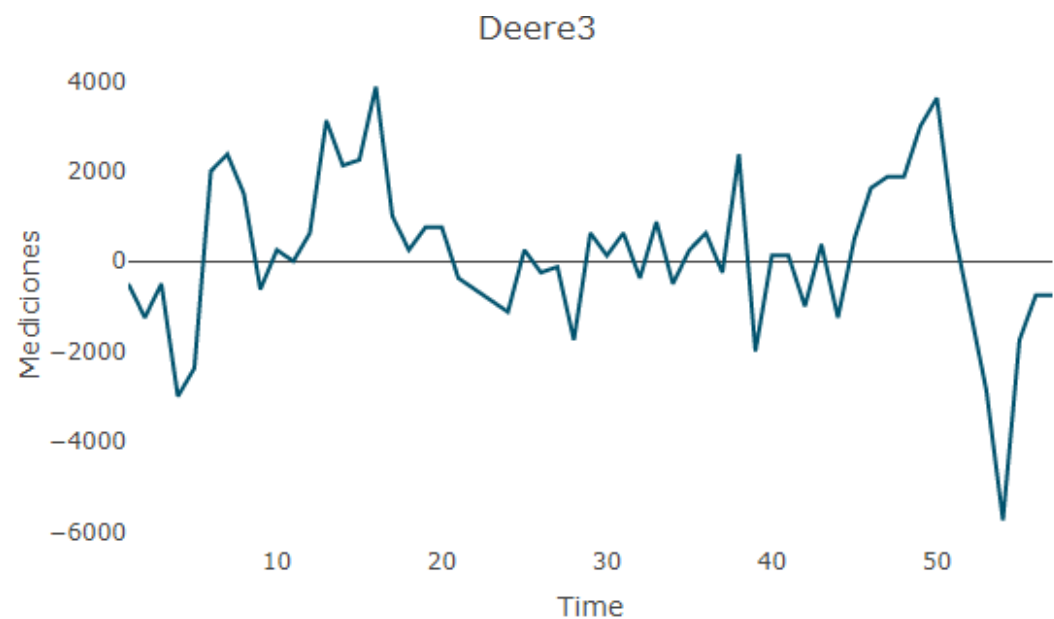


Figura 10: Serie temporal para las mediciones consecutivas de una máquina de Deere and co.

podemos ver que parece tener una esperanza centrada en cero, por lo que pudiéramos pensar que es estacionaria. Sin embargo, notemos que parece ser un proceso heterocedástico debido a que la varianza por la medición 50 es mayor que en el resto. Sin embargo, podemos pensar que dicha observación se debe a que el ruido blanco detrás puede ser un ruido de cola ligera y está es una observación atípica. Entonces, aún podemos dar el supuesto de estacionaridad y ajustar un modelo ARMA. Para ello veamos el ACF y PACF

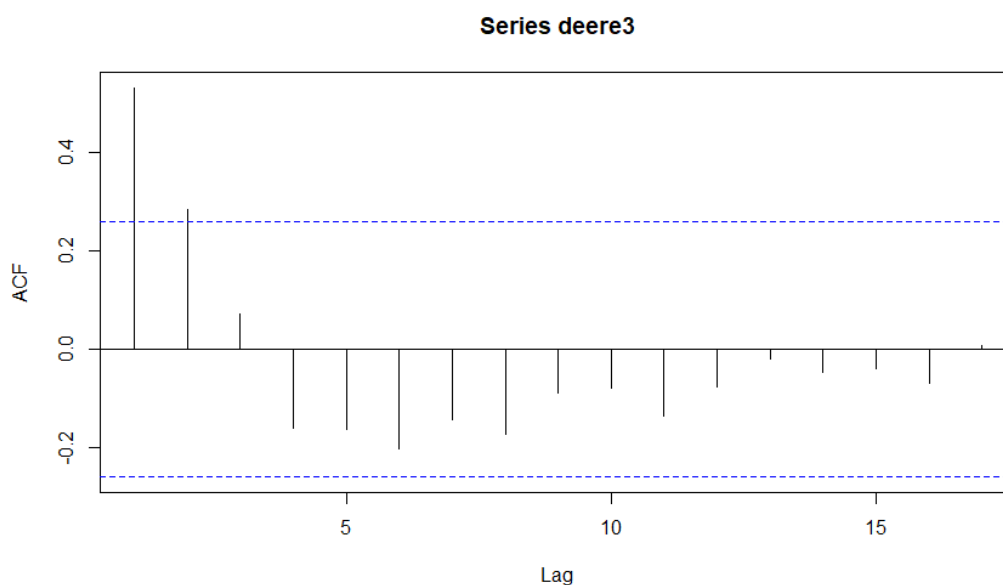


Figura 11: Función de autocorrelación para la los datos de Deere.

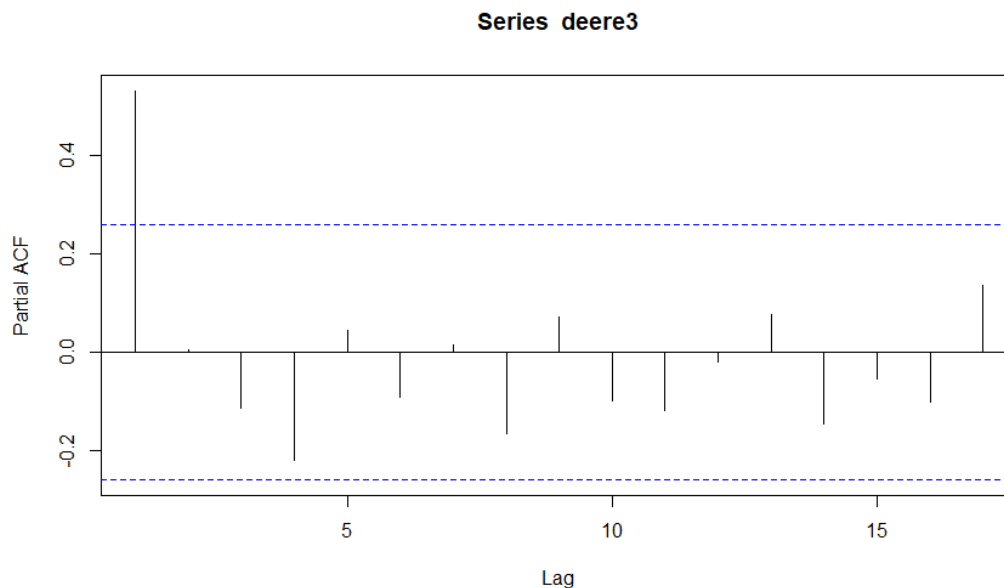


Figura 12: Función de autocorrelación parcial para la serie de Deere.

Posteriormente, para calcular el orden de selección, usamos los ordenes que minimicen el AIC. Con el siguiente ciclo, calculamos el AIC del modelo para las combinaciones posibles de modelos con p, q de cero a cuatro

```
for (i in 0:4) for (j in 0:4) {
  current.aic <- AIC(arima(x, order=c(i, 0, j)))
  if (current.aic < final.aic) {
    final.aic <- current.aic
    final.order <- c(i, 0, j)
    final.arma <- arima(x, order=final.order)
  }
}
```

lo que nos da que el mejor orden es $p = 3, q = 1$. Finalmente, con la prueba de Ljung-Box testeamos que se tenga un buen ajuste. Con el comando

```
Box.test(resid(final.arma), lag=20, type="Ljung-Box")
```

lo que nos da

```
Box-Ljung test

data:  resid(final.arma)
X-squared = 6.4666, df = 20, p-value = 0.9981
```

que tiene un p -valor mayor al nivel de significancia de 0.05. Por tanto, aceptamos los ordenes del modelo. Por otro lado, para buscar ordenes de otra forma empleamos el EACF cuya tabla se muestra a continuación

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	o	o	o	o	o	o	o	o	o	o	o	o
1	o	o	o	o	o	o	o	o	o	o	o	o	o	o
2	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3	x	o	o	o	o	o	o	o	o	o	o	o	o	o
4	o	x	o	o	o	o	o	o	o	o	o	o	o	o
5	o	x	o	x	o	o	o	o	o	o	o	o	o	o
6	o	x	o	x	o	o	o	o	o	o	o	o	o	o
7	o	x	x	o	o	o	o	o	o	o	o	o	o	o

Figura 13: Tabla para la estimación de ordenes por medio de la función de autocorrelación extendida.

Notamos que el estimador no es categórico, ya que se puede ver una línea difusa, entonces podemos estimar que $p = 2$, $q = 0$.

Referencias

- [1] Chan, N. H. (2011). Time series: Applications to finance with R and S-Plus (Vol. 837). John Wiley and Sons.
- [2] Brockwell, P. J., and Davis, R. A. (Eds.). (2002). Introduction to time series and forecasting. New York, NY: Springer New York.
- [3] Fuller, W. A. (2009). Introduction to statistical time series. John Wiley and Sons.
- [4] Chan, K. S., and Cryer, J. D. (2008). Time series analysis with applications in R. springer publication.