

CAPÍTULO 8

Modelos ARMA para la Componente Aleatoria

8.1. Introducción

En los modelos de descomposición $Y_t = T_t + S_t + \varepsilon_t$, $t = 1, 2, \dots$ se estima $\hat{\varepsilon}_t$ y se determina si es o no ruido blanco mediante las pruebas Ljung-Box y Durbin-Watson. En caso de encontrar que $\hat{\varepsilon}_t$ no es ruido blanco, el siguiente paso es modelar esta componente mediante cuatro posibles modelos.

1. Medias Móviles de orden q , $MA(q)$.
2. Autoregresivos de orden q , $AR(p)$.
3. Autoregresivos de Medias Móviles, $ARMA(p, q)$.
4. Autoregresivos de Medias Móviles Estacionales, $SARMA(p, q)(p_s, q_s)$.

Con relación a los tres primeros modelos Diebold [1999](ver pag. 129) anota lo siguiente : “Los tres modelos varían en su capacidad de capturar distintos tipos de comportamiento de autoregresión. Comenzaremos dando las características de las

funciones de autocorrelación y las cantidades relacionadas con cada modelo; éstas no tiene nada que ver con los datos ni con la estimación, pero son fundamentales para desarrollar una comprensión básica de las propiedades de los modelos, necesarias para llevar a cabo pronósticos inteligentes”.

Anotamos que Diebold [1999] no incluye el caso de Autoregresivos de Medias Móviles Estacionales, $SARMA(p, q)(p_s, q_s)$.

Definición 8.1.1. *El Operador Rezago se denota L (lag, en inglés) y se define por*

$$L(Y_t) = Y_{t-1}, \quad t \in \mathbb{Z}.$$

Es decir, L opera sobre una serie rezagándola un período hacia atrás. De igual manera $L(Y_{t-1}) = Y_{t-2}$, luego $L(L(Y_t)) = L^2(Y_t) = Y_{t-2}$ y en general $L^p(Y_t) = Y_{t-p}$, $p \geq 1$. Se define también $L^0 = I$, el operador identidad.

Un polinomio de grado p en el operador L se define como el operador formado por una combinación lineal de potencias de L . Por ejemplo, para constantes β_j , $j = 0, 1, \dots, p$ si se considera el polinomio en L

$$B_p(L) = \beta_0 I + \beta_1 L + \beta_2 L^2 + \dots + \beta_p L^p = \sum_{j=0}^p \beta_j L^j, \quad (8.1)$$

entonces, al aplicarlo a Y_t se obtiene

$$\begin{aligned} B_p(L)(Y_t) &= (\beta_0 I + \beta_1 L + \beta_2 L^2 + \dots + \beta_p L^p)Y_t, \\ &= \sum_{j=0}^p \beta_j L^j Y_t = \sum_{j=0}^p \beta_j Y_{t-j}, \\ &= \beta_0 Y_t + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p}. \end{aligned}$$

8.2. Procesos de Medias Móviles de orden q , $MA(q)$

Definición 8.2.1 (Proceso $MA(q)$). *Se dice que una serie $(Y_t, t \in \mathbb{Z})$ sigue un proceso $MA(q)$, $q = 1, 2, \dots$ de media móvil de orden q , si se cumple que*

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad t \in \mathbb{Z}, \quad (8.2)$$

para constantes $\theta_1, \dots, \theta_q$ y $\varepsilon_t \sim RB(0, \sigma^2)$. La expresión para (8.2) con el operador L es $Y_t = \Theta_q(L)(\varepsilon_t)$, donde

$$\Theta_q(z) = 1 + \theta_1 z + \dots + \theta_q z^q \quad (8.3)$$

se denomina el polinomio de media móvil.

Ejemplo 8.2.1. Defina $Y_t \sim MA(2)$ dado por

$$\begin{aligned} y_t &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \\ \theta_1 &= -0.4, \quad \theta_2 = 0.4, \\ \varepsilon_t &\sim iid N(0, 9), \quad t \in \mathbb{Z}. \end{aligned}$$

8.2.1. Propiedades de los procesos MA(q)

Suponga que $Y_t \sim MA(q)$ según (8.2). Entonces

1. $\mathbb{E}(Y_t) \equiv 0$. Se cumple inmediatamente pues $\mathbb{E}(Y_t) = \mathbb{E}(\varepsilon_t) + \theta_1 \mathbb{E}(\varepsilon_{t-1}) + \dots + \mathbb{E}(\theta_q \varepsilon_{t-q}) = 0$, ya que ε_t es ruido blanco. Se puede definir procesos MA(q) con media μ diferente de cero colocando $Y_t = \mu + \Theta_q(L)(\varepsilon_t)$.
2. $Var(Y_t) = (1 + \theta_1^2 + \dots + \theta_q^2)\sigma^2$. Debido a que ε_t es incorrelacionado se tiene que

$$\begin{aligned} Var(Y_t) &= Var(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}) \\ &= (1 + \theta_1^2 + \dots + \theta_q^2)\sigma^2. \end{aligned}$$

Nótese que $Var(Y_t) > Var(\varepsilon_t) = \sigma^2$.

3. $Cov(Y_t, Y_{t+k}) = R(k)$, donde

$$R(k) = \sigma^2 \sum_{j=0}^{q-k} \theta_j \theta_{j+k} = \sigma^2 (\theta_k + \theta_1 \theta_{1+k} + \dots + \theta_q), \quad 0 \leq k \leq q, \quad (8.4)$$

Y $R(k) = 0$, $k \geq q + 1$, con $\theta_0 = 1$. La ecuación (8.4) se puede interpretar como una evidencia de que un MA(q) es un proceso débilmente correlacionado, ya que su autocovarianza es cero a partir de un cierto rezago. Por esta razón se pueden ver los procesos MA(q) como alternativas al Ruido Blanco, que es completamente incorrelacionado.

Ejemplo 8.2.2. *Evaluar $R(k)$ con el proceso $MA(2)$ del Ejemplo 8.2.1*

El modelo es $Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$, $t \in \mathbb{Z}$, con $\varepsilon_t \sim iid N(0, 9)$, y los parámetros $\theta_1 = -0.4$, $\theta_2 = 0.4$, $\sigma_2 = 9$. Entonces hay que evaluar

$$R(k) = 9 \sum_{j=0}^{2-k} \theta_j \theta_{j+k}, \quad 0 \leq k \leq 2,$$

con $\theta_0 = 1$,

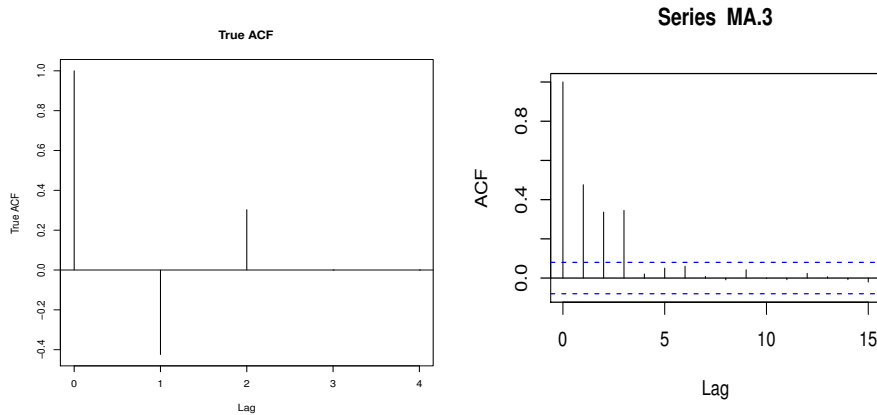
$$R(0) = 9(1 + 0.4^2 + 0.4^2) = 11.88$$

$$\begin{aligned} R(1) &= 9 \sum_{j=0}^{2-1} \theta_j \theta_{j+1} = 9(\theta_0 \theta_1 + \theta_1 \theta_2) \\ &= 9(-0.4 + (-0.4)(0.4)) = -5.04 \end{aligned}$$

$$R(2) = 9 \sum_{j=0}^{2-2} \theta_j \theta_{j+2} = 9(\theta_0 \theta_2) = 9(0.4) = 3.6.$$

Por tanto, la fac es

$$\begin{aligned} \rho(0) &= 1, \quad \rho(1) = -\frac{5.04}{11.88} = -0.42, \quad \rho(2) = \frac{3.6}{11.88} = 0.30 \\ \rho(3) &= \rho(4) = \dots = 0 \end{aligned}$$



(a) Función de Autocorrelación del Ejemplo 8.2.1

(b) fac muestral de un $MA(3)$

Figura 8.1: Fac teóricas de modelos $MA(q)$

Conclusión. De acuerdo con (8.4), si la fac muestral de una serie Y_t termina abruptamente puede tratarse de un $MA(q)$. Por ejemplo, en la Figura 8.1(b) sería factible un modelo $MA(3)$.

4. Un $MA(q)$ siempre es un proceso estacionario con fac, $\rho(k) = \frac{R(k)}{R(0)}$, $k \in \mathbb{Z}$.
5. Los $MA(q)$ son aproximadores universales de procesos estacionarios en covarianza (lineales). La posibilidad de encontrar un $MA(q)$ óptimo (aunque el orden q puede ser alto) está garantizada por un resultado teórico que dice que dado un proceso estacionario en covarianza que cumple cierta condición técnica (su densidad espectral deber ser continua en $[-\pi, \pi]$), entonces siempre existirá un $MA(q)$ que lo aproxime con un grado deseado de error (ver Brockwell and Davis [2002], pag. 132, Corollary 4.4.1).
6. La simulación de un proceso $MA(q)$ se puede hacer con la función `filter()`, descrita en la sección 6.1, pag.122, ya que el proceso es una media móvil unilateral de un ruido blanco. Si se denota por `et` el vector de ruido blanco con varianza σ^2 y se indica por `theta` un vector con los parámetros $\theta_1, \dots, \theta_q$, se simulan n valores de $Y_t \sim MA(q)$ ejecutando el comando

```
et = rnorm(n, 0, sigma)
y = filter(et, theta, "conv", sides=1, circular=TRUE) .
#Produce el mismo resultado que
y=arima.sim(list=(order(0, 0, q), ma=theta), n=n, sd=sigma) .
```

8.2.2. Condición de Invertibilidad del Proceso $MA(q)$

Los procesos $MA(q)$ pueden presentar un problema de falta de identificabilidad. Puede ocurrir que existan dos conjuntos de parámetros $\underline{\theta}^{(i)} = (\theta_1^{(i)}, \dots, \theta_q^{(i)})$, para $i = 1, 2$, con $\underline{\theta}^{(1)} \neq \underline{\theta}^{(2)}$ y sin embargo se cumple que las respectivas autocovarianzas (8.4) son iguales, $R(k, \underline{\theta}^{(1)}) \equiv R(k, \underline{\theta}^{(2)})$. Si se cumple la condición siguiente, “de invertibilidad”, no ocurre este problema.

Definición 8.2.2. Dado un proceso $MA(q)$, $Y_t = \theta_q(L)(\varepsilon_t)$ donde

$$\Theta_q(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q,$$

Si las q raíces de la ecuación

$$\Theta_q(z) = 0,$$

$z_1, z_2, \dots, z_q \in \mathbb{C}$, cumplen $|z_j| > 1$, se dice que el proceso Y_t es invertible.

Es decir, si $\Theta_q(z)$ cumple que “sus raíces están fuera del círculo unitario”, lo cual se expresa también como

$$\Theta_q(z) \neq 0, \forall z \in \mathbb{C}, |z| \leq 1. \quad (8.5)$$

Ejemplo 8.2.3. Consideremos el proceso $Y_t \sim MA(2)$ en el Ejemplo 8.2.1,

$$Y_t = \varepsilon_t - 0.4\varepsilon_{t-1} + 0.4\varepsilon_{t-2},$$

y veamos si es invertible. Hallamos las raíces del polinomio $\Theta_2(z)$

$$\begin{aligned} \Theta_2(z) &= 1 - 0.4z + 0.4z^2 = 0, \\ \therefore z &= \frac{0.4 \pm \sqrt{0.4^2 - 4(0.4)(1)}}{2(0.4)} = \frac{1}{2} \pm \frac{3}{2}i \end{aligned}$$

por tanto

$$|z| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{3}{2}\right)^2} = \sqrt{\frac{1}{4} + 9} > 1,$$

luego Y_t es invertible.

8.2.3. Función fac parcial de un Proceso MA(q) invertible

Definición 8.2.3 (Función de Autocorrelación Parcial (facp)). Suponga que $(Y_t, t \in \mathbb{Z})$ es serie estacionaria. La facp es una función de k , $\alpha(k)$, $k = 1, 2, \dots$ definida por

$$1. \alpha(1) = \rho(1)$$

$$2. \alpha(k) = \text{Corr}(U_1, U_k) \text{ donde}$$

$$U_1 = Y_1 - E(Y_1|Y_2, \dots, Y_k)$$

$$U_k = Y_{k+1} - E(Y_{k+1}|Y_2, \dots, Y_k), \quad k = 2, \dots$$

Lo que la facp hace es correlacionar Y_1 y Y_{k+1} pero eliminando primero el efecto de las variables intermedias Y_2, \dots, Y_k . La facp muestral se define por $\hat{\alpha}(k)$

1. $\hat{\alpha}(1) = \hat{\rho}(1)$
2. $\hat{\alpha}(2)$: se regresa Y_t sobre Y_{t-1} y Y_{t-2} tal que $Y_t = \phi_{21}Y_{t-1} + \phi_{22}Y_{t-2} + \varepsilon_t$, entonces $\hat{\alpha}(2) = \hat{\phi}_{22}$
3. $\hat{\alpha}(k)$: se regresa Y_t sobre Y_{t-1}, \dots, Y_{t-k} tal que $Y_t = \phi_{k1}Y_{t-1} + \dots + \phi_{kk}Y_{t-k} + \varepsilon_t$, entonces $\hat{\alpha}(k) = \hat{\phi}_{kk}$

La facp de un proceso $Y_t \sim MA(q)$ se puede encontrar sólo si se asume la condición de invertibilidad para un $MA(q)$. En este caso se tiene el siguiente resultado:

$$\lim_{k \rightarrow \infty} \alpha(k) = 0, \quad (8.6)$$

y se concluye que la facp de un $MA(q)$ siempre decrece a cero.

Ejemplo 8.2.4. En las Figuras siguientes 8.2 se observa la fac y la facp de un $MA(3)$ invertible.

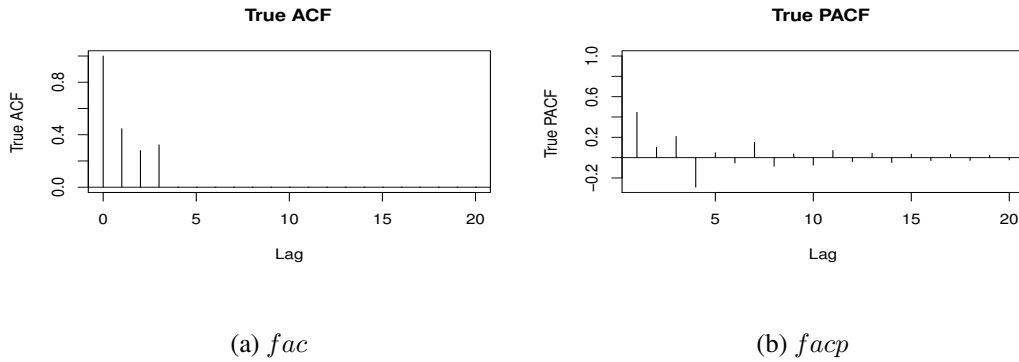


Figura 8.2: fac y facp de un $MA(3)$ invertible.

La prueba de (8.6) es como sigue.

Demostración. Suponga que se cumple la condición de invertibilidad (8.5) para el polinomio

$$\Theta_q(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q$$

Entonces la función $\Theta_q(z)^{-1} = 1/\Theta_q(z)$ se puede desarrollar en una serie de Taylor alrededor de $z = 0$, de la forma

$$\Theta_q(z)^{-1} = \sum_{j=0}^{\infty} \psi_j z^j, \quad (8.7)$$

para $|z| < 1$ y tal que $\sum_{j=0}^{\infty} |\psi_j| < \infty$, $\psi_0 = 1$.

Si se reemplaza z por L se tiene un operador en el rezago L de la forma $\Theta_q(L)^{-1} = \sum_{j=0}^{\infty} \psi_j L^j$, el cual, aplicado a Y_t produce

$$\begin{aligned} \Theta_q(L)^{-1} Y_t &= \sum_{j=0}^{\infty} \psi_j L^j (Y_t) = \sum_{j=0}^{\infty} \psi_j Y_{t-j} \\ &= Y_t + \psi_1 Y_{t-1} + \psi_2 Y_{t-2} + \dots \end{aligned}$$

Ahora, multiplicando ambos miembros de por $\Theta_q(L)^{-1}$ se obtiene

$$\begin{aligned} Y_t &= \Theta_q(L) \varepsilon_t \\ \Leftrightarrow \Theta_q(L)^{-1} Y_t &= \Theta_q(L)^{-1} \Theta_q(L) \varepsilon_t = \varepsilon_t \\ \therefore \Theta_q(L)^{-1} Y_t &= \varepsilon_t \\ \Leftrightarrow Y_t + \psi_1 Y_{t-1} + \psi_2 Y_{t-2} + \dots &= \varepsilon_t \\ \therefore Y_t &= -\psi_1 Y_{t-1} - \psi_2 Y_{t-2} - \dots + \varepsilon_t, \end{aligned}$$

Por tanto, si se hace la regresión de Y_t sobre los primeros k rezagos Y_{t-j} , $j = 1, \dots, k$,

$$Y_t = -\psi_1 Y_{t-1} - \psi_2 Y_{t-2} - \dots - \psi_k Y_{t-k} + \varepsilon_t \quad (8.8)$$

entonces el k -ésimo coeficiente es $\alpha(k) = \psi_k$. Pero la serie $\sum_{k=0}^{\infty} \psi_k$ es absolutamente convergente, entonces $\psi_k \rightarrow 0$. Por tanto, $\alpha(k) \rightarrow 0$ cuando $k \rightarrow \infty$ y se concluye que la facp de un MA(q) siempre decrece a cero. \square

8.2.4. Estimación de procesos MA(q)

La estimación de procesos $Y_t \sim MA(q)$ no es directa. Una de las razones es que en $Y_t = \theta_q(L) \varepsilon_t$ los valores del ruido blanco ε_t no son observables. Supongamos una muestra de tamaño T de la serie, Y_1, \dots, Y_T .

En el caso de estimadores de mínimos cuadrados ordinarios ver (2.7), pag.30, el estimador de $\underline{\theta} = (\theta_1, \dots, \theta_q)'$ se define como el vector $\hat{\underline{\theta}}$ que minimiza la función

$$G(\underline{\theta}) = \sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T \left(Y_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j} \right)^2, \quad (8.9)$$

Pero los ε_{t-j} no son observables. Se puede recurrir a la invertibilidad, y utilizar la serie en (8.8) pero truncada hasta un cierto n . Permitiendo seguir con el signo igual, se escribe

$$Y_t = -\psi_1 Y_{t-1} - \psi_2 Y_{t-2} - \dots - \psi_n Y_{t-n} + \varepsilon_t. \quad (8.10)$$

Considerando la identidad

$$\Theta_q(z)^{-1} \Theta_q(z) = (1 + \psi_1 z + \psi_2 z^2 + \dots)(1 + \theta_1 z + \dots + \theta_q z^q) = 1$$

se pueden despejar los ψ_j en función del vector $\underline{\theta} = (\theta_1, \dots, \theta_q)'$, escribiendo $\psi_j = \psi_j(\underline{\theta})$. La minimización se hace con base en (8.10) con la función objetivo

$$G(\underline{\theta}) = \sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T \left(Y_t + \sum_{j=1}^n \psi_j(\underline{\theta}) Y_{t-j} \right)^2. \quad (8.11)$$

El correspondiente estimador $\hat{\underline{\theta}}$ se denomina estimador de mínimos cuadrados condicionales. Es asintóticamente eficiente y tiene las mismas propiedades asintóticas del estimador de máxima verosimilitud.

En R hay varias funciones para estimar procesos MA(q). La función `arma` de la librería `tseries` y la función `arima`. La primera calcula los valores p de los coeficientes, y la segunda permite calcular pronósticos. Ambas utilizan estimadores de mínimos cuadrados condicionales. El siguiente código muestra cómo estimar y pronosticar un proceso MA(q).

Ejemplo 8.2.5. *Retomando el Ejemplo (7.3.2) en la pag. 141, para modelación de la serie de producción de cemento Portland, trimestral se ajustó un modelo de la forma*

$$Y_t = \beta_0 + \beta_1 t + \sum_{j=1}^3 \delta_j I_t(t) + \varepsilon_t.$$

Al estimar los residuos estructurales $\hat{\varepsilon}_t$ se observó que posiblemente muestran una dinámica autocorrelacionada. Se confirmó este hecho con las pruebas Ljung-Box y Durbin-Watson. En este Ejemplo se investiga si los residuos podrían seguir una dinámica MA(q).

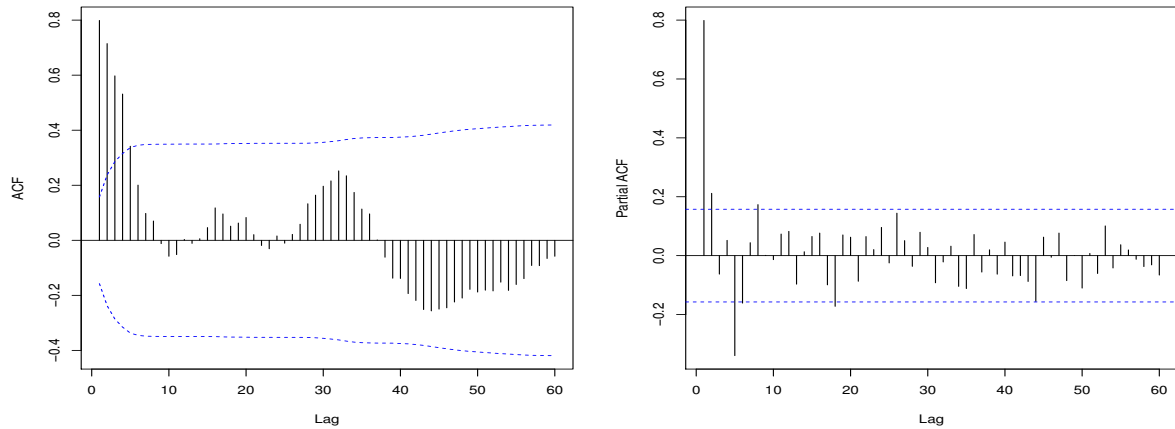
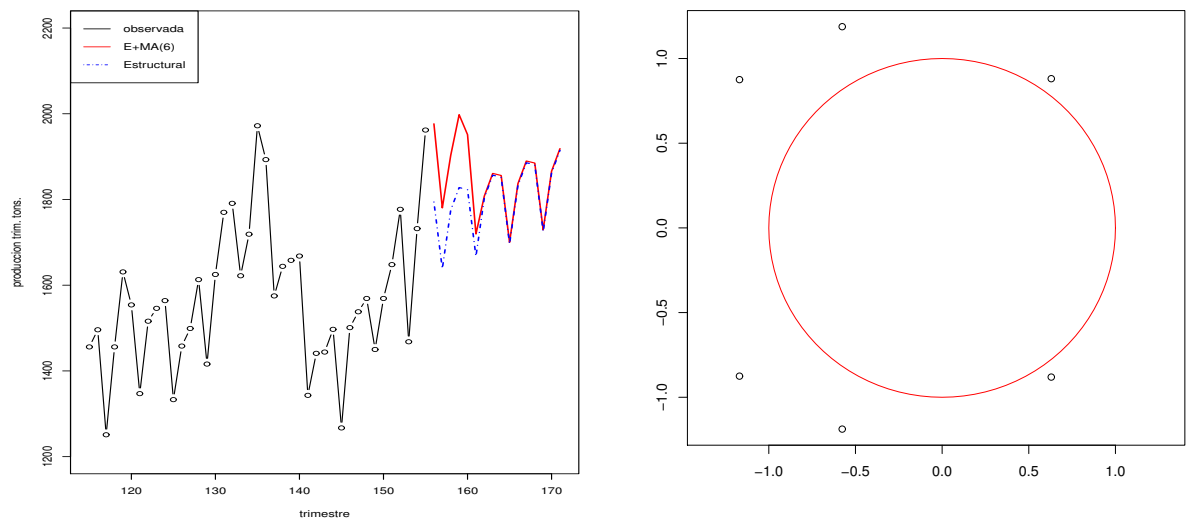


Figura 8.3: Residuos estructurales de la serie de producción de cemento



(a) Pronósticos

(b) Raíces círculo unitario

Figura 8.4: (a) Pronosticos Estructurales versus Estructurales + MA(6), (b) Raíces y círculo unitario

En la Figura 8.3 se muestran la fac y la fac parcial de los residuos estructurales, que muestran que un MA(q) es factible, con $q \geq 4$. El código R siguiente muestra que un

MA(6) puede presentar mejor ajuste. En la Figura 8.4, panel izquierdo, se muestran pronósticos con el modelo estructural y el modelo MA(6) para los residuos. Las raíces del polinomio del modelo MA(6) se grafican en la Figura 8.4, panel derecho. El código R siguiente muestra la estimación y análisis del MA(6).

Estima MA(6) para los residuos con dos funciones de estimación: arma y arima.

```
#-----
require(tseries)
mod4.1 = arma(r,order = c(0,6))
summary(mod4.1)
mod4.2 = arima(r,order = c(0,0,6))
summary(mod4.2)
```

Examinar residuos del MA(6)

```
#-----
r4.2 = residuals(mod4.2)

par(mfrow=c(2,2))
plot(t,r4.2,type='o',ylab='residuo')
abline(h=0,lty=2)
plot(density(r4.2),xlab='x',main= '')
acf(r4.2,60,ci.type="ma",drop.lag.0=TRUE,main="")
pacf(r4.2,60,main="")
cpgram(r4.2)
#----- prueba Ljung-Box
Box.test(r4.2, lag = 12, type = "Ljung-Box")
X-squared = 9.5687, df = 12, p-value = 0.6537
```

Calcular 16 pronósticos estructurales añadiendo 16 pronósticos MA(6).

```
m = 16
tp = seq(T+1,T+m)
ltp = seasonaldummyf(y,m)
# el objeto mod1 es el modelo estructural
```

```

y.p = predict(mod1,data.frame(t=tp,It=I(Itp)))
r.p = predict(mod4.2,n.ahead=m)$pred
#----- pronosticos conjuntos
yr.p = y.p+r.p

#-----
require(polynom)
theta = coef(mod4.2); theta = c(1,theta[-7]);
(Mod(polyroot(theta)))
require(signal)
zplane(filt=rev(theta),a=c(1))

#----- graficar
m0 = 40; N = seq(T-m0,T);
par(mfrow=c(1,1))
plot(seq(T-m0,T),y[N],type='b',
ylim=c(1200,2200),xlim=c(T-m0,T+m),
ylab='produccion trim. tons.',xlab='trimestre')
lines(seq(T+1,T+m),yr.p,lty=7,lwd=2,col='red')
lines(seq(T+1,T+m),y.p,lty=4,lwd=2,col='blue')
legend("topleft",
legend=c("observada","E+MA(6)","Estructural"),
col=c('black','red','blue'),lty=c(1,7,4))

```

8.3. Procesos Autorregresivos de Orden p , AR(p)

Definición 8.3.1 (Proceso AR(p)). *Un proceso $(Y_t, t \in \mathbb{Z})$ sigue un modelo AR(p) de media cero si es solución de la ecuación recursiva siguiente.*

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t, \quad (8.12)$$

donde $\varepsilon_t \sim RB(0, \sigma^2)$, $\varphi_j, j = 1, 2, \dots, p$ son los coeficientes (parámetros) y $p = 1, 2, \dots$ es el orden. El caso $p = 0$ se interpreta como $Y_t = \varepsilon_t$.

Se define

$$\Phi_p(z) = 1 - \sum_{j=1}^p \varphi_j z^j \quad (8.13)$$

el polinomio autorregresivo. Usando el operador de rezago L se puede escribir (8.12) como

$$\Phi_p(L)(Y_t) = \varepsilon_t.$$

Un proceso $Y_t \sim AR(p)$ con media diferente de cero, $\mathbb{E}(Y_t) = \mu \neq 0$, se define mediante la ecuación

$$\Phi_p(L)(Y_t - \mu) = \varepsilon_t. \quad (8.14)$$

Obsérvese que

$$\begin{aligned} \Phi_p(L)(Y_t - \mu) &= \Phi_p(L)(Y_t) - \Phi_p(L)(\mu) \\ &= \Phi_p(L)(Y_t) - \mu(1 - \sum_{j=1}^p \varphi_j) = \varepsilon_t. \end{aligned}$$

Luego,

$$\begin{aligned} \Phi_p(L)(Y_t) &= \varphi_0 + \varepsilon_t \\ \Leftrightarrow Y_t &= \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t, \end{aligned}$$

con $\varphi_0 = \mu(1 - \sum_{j=1}^p \varphi_j)$. La constante φ_0 se denomina intercepto.

8.3.1. Condición Suficiente para que un $AR(p)$ sea Estacionario en Covarianza

Una condición suficiente para que $Y_t \sim AR(p)$ sea estacionario en covarianza es que las p raíces de la ecuación $\Phi_p(z) = 0$, z_i , $i = 1, 2, \dots, p$, cumplan

$$|z_i| > 1. \quad (8.15)$$

Es decir, si $\Phi_p(z)$ cumple que “sus raíces están fuera del círculo unitario”, expresado como

$$\forall z \in \mathbb{C}, |z| \leq 1, \Phi_p(z) \neq 0. \quad (8.16)$$

En la Figura 8.5 se observa el círculo unitario junto con una raíz z_j y su conjugado \bar{z}_j , ambas fuera del círculo.

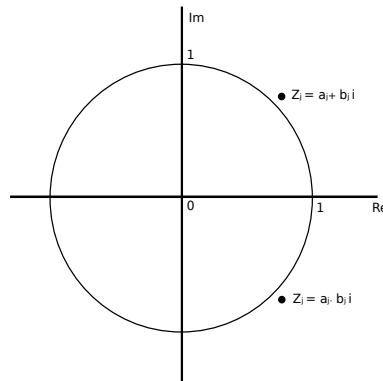


Figura 8.5: Círculo Unitario

Sin embargo, la condición (8.16) no es necesaria. Puede suceder que algunas de las raíces de $\Phi_p(z) = 0$ estén adentro y otras afuera del círculo unitario, pero si se cumple que $\Phi_p(z) \neq 0, \forall z \in \mathbb{C}, |z| = 1$, en este caso Y_t es estacionario en covarianza.

Ejemplo 8.3.1. El proceso $Y_t \sim AR(2)$,

$$Y_t = 1.5Y_{t-1} - 0.9Y_{t-2} + \varepsilon_t, \quad (8.17)$$

cumple que es estacionario en covarianza ya que las raíces características de $\Phi_2(z) = 1 - 1.5z + 0.9z^2 = 0$, que son $\frac{5}{6} \pm \frac{\sqrt{15}i}{6}$, están fuera del círculo unitario. Se puede hacer la comprobación con las funciones `InvertibleQ()` de la librería `FitAR`, y `armaRoots()` de la librería `fArma`, pero que se puede cargar como programa independiente, como sigue.

```
phi = c(1.5, -0.9)
require(FitAR)
InvertibleQ(phi)
[1] TRUE
source("armaRoots.r")
armaRoots(phi)
      re      im      dist
```

```

1 0.8333 0.6455 1.0541
2 0.8333 -0.6455 1.0541

```

Para un proceso $Y_t \sim AR(p)$ de media cero, que cumple la condición para ser estacionario en covarianza (8.16), se cumple:

$$i) \quad \mathbb{E}(Y_t) = 0, \quad (8.18)$$

$$ii) \quad \sum_{j=1}^p \varphi_j < 1. \quad (8.19)$$

Demostración. i). Para comprobar (8.18), si Y_t es estacionario en covarianza entonces $\mathbb{E}(Y_t) \equiv \mu$. Además,

$$\mathbb{E}(Y_t) = \varphi_1 \mathbb{E}(Y_{t-1}) + \varphi_2 \mathbb{E}(Y_{t-2}) + \cdots + \varphi_p \mathbb{E}(Y_{t-p}) + 0,$$

pero todas las esperanzas son iguales a μ , luego

$$\mu = \varphi_1 \mu + \varphi_2 \mu + \cdots + \varphi_p \mu.$$

Si $\mu \neq 0$ entonces $1 = \varphi_1 + \cdots + \varphi_p$, y por tanto el polinomio autorregresivo evaluado en $z = 1$ se anula, $\Phi_p(1) = 0$, lo cual es una contradicción ya que $\forall z \in \mathbb{C}, |z| \leq 1$ $\Phi_p(z) \neq 0$. Luego debe tenerse que $\mu = 0$, es decir, el proceso definido en (8.12) es de media cero. \square

Demostración. ii). Usando la técnica de reducción al absurdo, para comprobar (8.19), se asume lo contrario, es decir $\sum_{j=1}^p \varphi_j \geq 1$. Si ocurre el caso $\sum_{j=1}^p \varphi_j = 1$, equivale a $\Phi_p(1) = 0$, lo cual es una contradicción con la condición (8.16). Si se asume $\sum_{j=1}^p \varphi_j > 1$ entonces $\Phi_p(1) < 0$. Pero $\Phi_p(0) = 1$. Como la función $\Phi_p(z)$ es continua en $[0, 1]$ debe existir un punto $0 < z < 1$ en el cual $\Phi_p(z) = 0$. Pero eso es nuevamente una contradicción con la condición (8.16), porque este z sería una raíz de $\Phi_p(z) = 0$ que cumple $|z| < 1$. \square

8.3.2. La Función de Autocovarianza de un Proceso AR(p)

La función de autocovarianza $R(k) = Cov(Y_t, Y_{t+k})$, de un proceso $Y_t \sim AR(p)$, se puede calcular resolviendo una ecuación recursiva lineal denominada la ecuación

de Yule–Walker, dada por

$$R(k) = \sum_{j=1}^p \varphi_j R(k-j), \quad k = 1, 2, \dots, \quad (8.20)$$

donde el valor inicial está dado por $R(0) = \sum_{j=1}^p \varphi_j R(j) + \sigma^2$.

A partir de (8.20) es inmediato que la fac $\rho(k) = R(k)/R(0)$ cumple una ecuación similar, al dividir estas ecuaciones por $R(0)$.

$$\rho(k) = \sum_{j=1}^p \varphi_j \rho(k-j), \quad k = 1, 2, \dots \quad (8.21)$$

La ecuación (8.21) es una ecuación recursiva lineal de orden p , homogénea. Requiere p valores iniciales

$$\rho(0) = 1, \rho(1) = q_1, \dots, \rho(p-1) = q_{p-1}.$$

Se puede resolver por fórmula ó utilizando un software de cálculo simbólico como por ejemplo, Maple ó Maxima ⁽¹⁾.

Demostración. Con $\mu = \mathbb{E}(Y_t)$, la ecuación $Y_t = \varphi_0 + \sum_{j=1}^p \varphi_j Y_{t-j} + \varepsilon_t$ equivale a $Y_t - \mu = \sum_{j=1}^p \varphi_j (Y_{t-j} - \mu) + \varepsilon_t$. Multiplicando ambos miembros por $Y_{t-k} - \mu$, con $1 \leq k \leq t$, y tomando valor esperado $\mathbb{E}(\cdot)$ se obtiene

$$\begin{aligned} R(k) &= \mathbb{E}((Y_t - \mu)(Y_{t-k} - \mu)) \\ &= \sum_{j=1}^p \varphi_j \mathbb{E}((Y_{t-j} - \mu)(Y_{t-k} - \mu)) + \mathbb{E}(\varepsilon_t(Y_{t-k} - \mu)) \\ &= \sum_{j=1}^p \varphi_j R(k-j). \end{aligned}$$

Ya que se cumple $\mathbb{E}(\varepsilon_t(Y_{t-k} - \mu)) = 0$ porque, a partir de la definición del proceso Y_t en (8.12), Y_{t-k} depende de ε_s con $s \leq t-k$, que son variables incorrelacionadas con ε_t . El caso $k = 0$ es similar, con $\mathbb{E}(\varepsilon_t(Y_t - \mu)) = \text{Var}(\varepsilon) = \sigma^2$. \square

Ejemplo 8.3.2. Retomando el Ejemplo 8.3.1, para el proceso $Y_t \sim AR(2)$, $Y_t = 1.5Y_{t-1} - 0.9Y_{t-2} + \varepsilon_t$, estacionario en covarianza, con $\varepsilon_t \sim iidN(0, \sigma^2)$, $\sigma^2 = 9$. Encontrar su autocorrelación.

¹<http://maxima.sourceforge.net/es/index.html>

La ecuación de Yule-Walker para $\rho(k)$

$$\rho(k) = 1.5\rho(k-1) - 0.9\rho(k-2), \quad k \geq 2. \quad (8.22)$$

Hay que encontrar las condiciones iniciales, $\rho(0), \rho(1)$. Primero se encuentra la varianza del proceso, $R(0)$.

$$\begin{aligned} R(0) &= \sum_{j=1}^2 \varphi_j R(j) + \sigma^2 \\ \Leftrightarrow 1 &= 1.5\rho(1) - 0.9\rho(2) + \sigma^2/R(0) \\ \Leftrightarrow R(0) &= \frac{\sigma^2}{1 - 1.5\rho(1) + 0.9\rho(2)} \end{aligned}$$

Reemplazando $k = 1$ en (8.22) y con $\rho(0) = 1$ se obtiene

$$\begin{aligned} \rho(1) &= 1.5\rho(0) - 0.9\rho(1) \\ \therefore \rho(1) &= \frac{15}{19}. \end{aligned}$$

Se puede encontrar la solución de (8.22) usando una teoría básica de ecuaciones lineales recurrentes homogéneas, ver ⁽²⁾. Con Maxima se obtiene

$$\rho(n) = \frac{(2\sqrt{10})^n \sqrt{15} 3^{n-1} \sin\left(\operatorname{atan}\left(\frac{\sqrt{15}}{5}\right) n\right)}{19 \times 20^n} + \frac{(2\sqrt{10})^n 3^n \cos\left(\operatorname{atan}\left(\frac{\sqrt{15}}{5}\right) n\right)}{20^n}$$

Nótese que en este caso la fac tiene una semi-periodicidad, de período

$$\frac{2\pi}{\arctan\left(\frac{\sqrt{15}}{5}\right)} = 9.5,$$

y además es amortiguada exponencialmente, tendiendo a cero.

La función de autocovarianza, $R_X(k)$, se pueden calcular en el lenguaje R con la función `ARMAacf()` como sigue. La gráfica está en la Figura 8.6,

```
phi = c(1.5, -0.9)
acf.ar2 = ARMAacf(ar=phi, ma=numeric(0), lag.max=30)
str(acf.ar2)
Named num [1:31] 1 0.789 0.284 -0.284 -0.682 ...
- attr(*, "names")= chr [1:31] "0" "1" "2" "3" ...
```

²https://en.wikipedia.org/wiki/Recurrence_relation#Roots_of_the_characteristic_polynomial

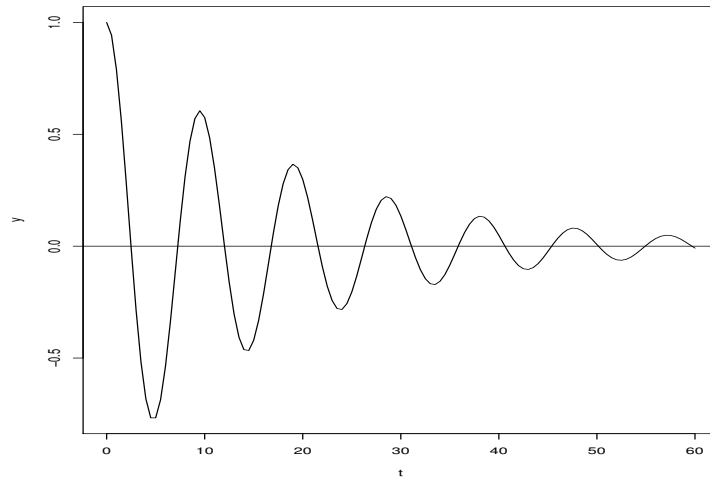


Figura 8.6: fac de ar(2)

8.3.3. Función fac parcial de un Proceso AR(p)

El razonamiento para determinar la facp de un AR(p) es directo. La facp de un proceso $AR(p)$ es $\alpha(k)$ tal que $\hat{\alpha}(k)$ es el coeficiente $\hat{\beta}_{k,k}$ en la regresión

$$Y_t = \beta_0 + \beta_{k,1}Y_{t-1} + \cdots + \beta_{k,k}Y_{t-k} + \eta_t, \quad k = 2 \quad (8.23)$$

pero como $\beta_{k,k} = 0$ si $k \geq p + 1$ entonces $\hat{\alpha}(k) = 0$ si $k \geq p + 1$. Se concluye que la fac parcial de un AR(p) es cero para rezagos $k \geq p + 1$.

Ejemplo 8.3.3. Retomando el Ejemplo 8.3.1, para el proceso $Y_t \sim AR(2)$,

$$Y_t = 1.5Y_{t-1} - 0.9Y_{t-2} + \varepsilon_t,$$

estacionario en covarianza, la función de autocovarianza parcial, $\alpha_X(k)$, se pueden calcular en R con la función siguiente

```
phi = c(1.5, -0.9)
PAcf = ARMAacf(ar=phi,
ma=numeric(0), lag.max=30, pacf=TRUE)
str(PAcf)
num [1:30] 7.89e-01 -9.00e-01 3.10e-15 -8.67e-16 -9.40e-16 ...
```

8.3.4. Estimación de procesos AR(p)

La estimación de un proceso $Y_t \sim AR(p)$ es directa y existen varios métodos para realizarla. Si se tiene una muestra del proceso, Y_1, \dots, Y_T , se pueden utilizar:

1. Los estimadores Yule-Walker con base en la fac muestral. Se basan en la siguiente identidad. La fac de un $AR(p)$, $\rho(k) = \text{Corr}(Y_t, Y_{t+k})$, $k = 1, 2, \dots, p, p+1, \dots$, cumple lo siguiente.

Dadas la matriz A y los vectores $\underline{\varphi}, \underline{\rho}$

$$A = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(p-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(p-3) \\ \vdots & \vdots & & & \vdots \\ \rho(p-1) & \rho(p-2) & \cdots & & 1 \end{bmatrix}, \quad \underline{\varphi} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix}, \quad \underline{\rho} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{bmatrix},$$

entonces es válido el sistema lineal $p \times p$

$$A\underline{\varphi} = \underline{\rho}. \quad (8.24)$$

Luego, dados los estimadores de la fac $\hat{\rho}(1), \dots, \hat{\rho}(p)$ se pueden reemplazar en la expresiones para A y $\underline{\rho}$, y resolver (8.24) para obtener

$$\hat{\underline{\varphi}} = \hat{A}^{-1}\hat{\underline{\rho}}. \quad (8.25)$$

Estos estadísticos $\hat{\underline{\varphi}}$ así calculados se denominan los estimadores de Yule-Walker de $\underline{\varphi}$.

2. Estimadores por mínimos cuadrados ordinarios. En el caso de estimadores de mínimos cuadrados ordinarios, ver (2.7), pag.30, el estimador de $\underline{\varphi} = (\varphi_1, \dots, \varphi_p)'$ se define como el vector $\hat{\underline{\varphi}}$ que minimiza la función

$$G(\underline{\varphi}) = \sum_{t=1+p}^T \varepsilon_t^2 = \sum_{t=1+p}^T \left(Y_t - \sum_{j=1}^p \varphi_j Y_{t-j} \right)^2. \quad (8.26)$$

Un problema que surge con estos estimadores es que el vector estimado $\hat{\underline{\varphi}}$ puede no cumplir la condición (8.16), para que la solución Y_t sea estacionaria en covarianza. El método de Burg permite corregir esta posibilidad.

3. estimadores de máxima verosimilitud, asumiendo que los errores cumple $\varepsilon_t \sim iidN(0, \sigma^2)$,

Los métodos de máxima verosimilitud y Burg se pueden consultar en Brockwell and Davis [2002] y Grandell [2014]

Estimación de un proceso AR(p) en R

La estimación de un proceso AR(p) en R se puede hacer mediante varias funciones en R. La función `ar()` de la librería `stats` provee la estimación por los métodos: mínimos cuadrados, Yule-Walker, máxima verosimilitud y Burg.

```
ar(x, aic = TRUE, order.max = NULL,
    method = c("yule-walker", "burg", "ols", "mle", "yw")
#-----
aic : Logical flag. If TRUE then the Akaike Information
Criterion is used to choose the order of the
autoregressive model. If FALSE, the model of order
order.max is fitted.
```

Ejemplo 8.3.4. Retomando el Ejemplo 8.3.1, para el proceso $Y_t \sim AR(2)$,

$$Y_t = 1.5Y_{t-1} - 0.9Y_{t-2} + \varepsilon_t,$$

estacionario en covarianza, los resultados de una estimación con base en una muestra simulada, utilizando `ar()` con las opciones “ols”, “burg”, “yw” y “mle” se muestran a continuación.

	ols	burg	yw	mle
1	1.48	1.48	1.48	1.50
2	-0.88	-0.88	-0.88	-0.90

Tabla 8.1: Estimadores OLS, Burg, Y-W, MaxVer en un AR(2)

La función `FitAR()` de la librería `FitAR` utiliza un algoritmo para máxima verosimilitud. Por ejemplo, las instrucciones siguientes determinan un orden p óptimo mediante el mínimo AIC, considerando todas las posibilidades hasta un máximo, en el ejemplo de 30.

```
p = SelectModel(y, lag.max=32, Best=1)
m2 = FitAR(y, p)
coef(m2)
```

La función `arima()` de `stats` utiliza máxima verosimilitud. Un inconveniente con esta función es que reporta la media con el nombre de intercepto. Así, podría generar confusión ya que intercepto se refiere a φ_0 .

8.3.5. Simulación de procesos AR(p)

La simulación de un proceso AR(p) se puede hacer mediante la función `filter()`, escribiendo `y=filter(et, phi, "recursive", init=y0)`, donde `phi=c(f1, ..., fp)` es el vector de coeficientes φ_j en la ecuación $Y_t = \sum_{j=1}^p \varphi_j Y_{t-j} + \varepsilon_t$, y $(\varepsilon_t) \sim RB(0, \sigma^2)$. El vector `y0=c(y1, ..., yp)` consiste de los primeros p valores necesarios para iniciar la recursión AR(p).

Ejemplo 8.3.5. *En el caso del Ejemplo 8.3.1 la instrucción es*

```
et = rnorm(1000, 0, 2)
y=filter(et, c(1.5, -0.9), "recursive", init=y0)
```

donde `y0=c(0, 0)`, si se toman los valores iniciales como la media del proceso, que es cero.

La simulación de un proceso AR(p) de media μ diferente de cero, se realiza colocando `y=mu+filter(et, phi, "recursive", init=y0)`

8.3.6. Análisis de un caso de aplicación de AR(p)

La serie de aportes hidrológicos diarios en Colombia muestra la cantidad de agua convertida en energía, en todas las represas para generación de electricidad en el país. La unidad es Gw/hora. Para los datos entre Julio 1 de 1998 y 31 de Mayo de 2014, la gráfica de los aportes, Y_t está en la Figura 8.7. Inicialmente se propone un

modelo de componentes lineal y trigonométricas, de la forma

$$Y_t = b_0 + b_1 t + \sum_{j=1}^k \alpha_j \cos(2\pi j t / 365) + \beta_j \sin(2\pi j t / 365) + \varepsilon_t, \quad (8.27)$$

donde (ε_t) es una serie con autocorrelacionada, que posiblemente puede ser AR(p). Los análisis en el programa R siguiente muestran que este supuesto es correcto.

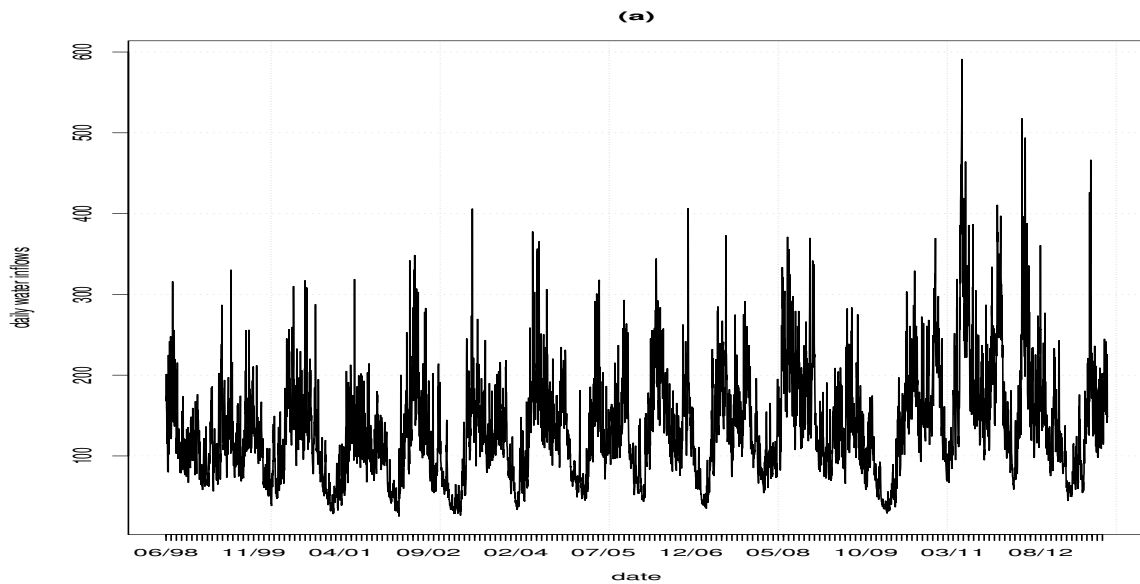


Figura 8.7: Aportes hidrológicos en Gw/h, 01/07/98 - 31/05/14

```
# Ejemplo de analisis AR(p)
# Serie de aportes hidrológicos diarios
# en Gw/h
#-----modela estacionalidad con sen, cos
#----- serie con frecuencia diaria
Ap = ts(Ap,frequency = 365)
require(forecast)
It.trig = fourier(Ap,4)
t = seq(1,length(Ap))/100
mod2 = lm(Ap ~ t + It.trig)
summary(mod2)
```

Los resultados son aceptables, aunque el R^2 está bajo.

```
Call: lm(formula = Ap ~ t + It.trig)
```

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	104.80070	1.34523	77.906	< 2e-16	***
t	1.00769	0.04182	24.094	< 2e-16	***
It.trigS1-365	32.45044	0.95059	34.137	< 2e-16	***
It.trigC1-365	24.35062	0.95159	25.589	< 2e-16	***
It.trigS2-365	-11.17645	0.95123	-11.749	< 2e-16	***
It.trigC2-365	31.80005	0.95038	33.460	< 2e-16	***
It.trigS3-365	0.43385	0.95037	0.457	0.648	
It.trigC3-365	-3.70841	0.95124	-3.899	9.79e-05	***
It.trigS4-365	-7.41607	0.95115	-7.797	7.52e-15	***
It.trigC4-365	4.54505	0.95022	4.783	1.77e-06	***

```
---
```

```
Multiple R-squared: 0.4021, Adjusted R-squared: 0.4012
```

```
F-statistic: 415.6 on 9 and 5561 DF, p-value: < 2.2e-16
```

Identificar los residuos con fac y fac parcial en la Figura 8.8.

```
#-----examen fac y fac parcial residuos
r = residuals(mod2)
require(TSA)
par(mfrow=c(1,2))
TSA::acf(r,60, drop.lag.0 = TRUE, ci.type="ma")
pacf(r,60)
```

Se puede concluir que los residuos estructurales pueden ser modelados mediante un AR(p), con $p \geq 9$. Para identificar p se puede usar la función `SelectModel`

```
#-----identifica AR(p)
#install.packages("FitAR")
require(FitAR)
n = length(Ap)
```

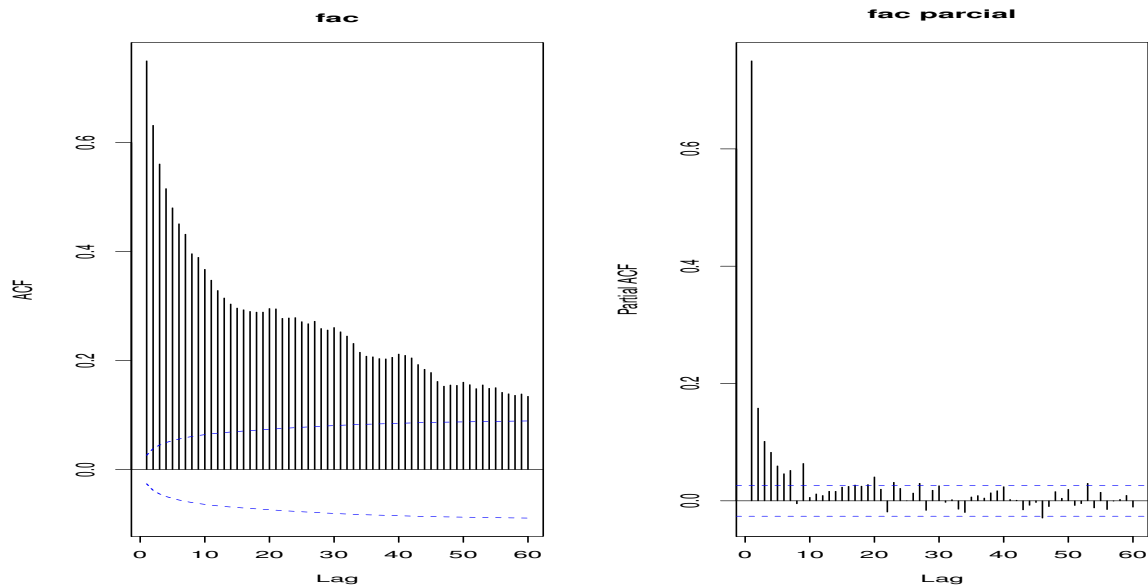


Figura 8.8: fac y facp de Aportes hidrológicos

```
pvec = SelectModel(r, ARModel="AR",
Criterion="BIC",
lag.max=floor(10*log10(n)), Best=1)
(p=pvec)
9
```

ó también la función ar

```
# aic Logical flag. If TRUE then the Akaike Information
# Criterion is used
mod1 = ar(r,aic = TRUE, method=c("burg"))
print(mod1)
Order selected 30
```

Los residuos del modelo AR(30) deben ser ruido blanco. La Figura 8.9 muestra que la fac está dentro de las bandas de Bartlett. Finalmente, la prueba Ljung-Box muestra que no se rechaza la nula de ruido blanco.

```
Box.test(et,50,type= "Ljung-Box")
```

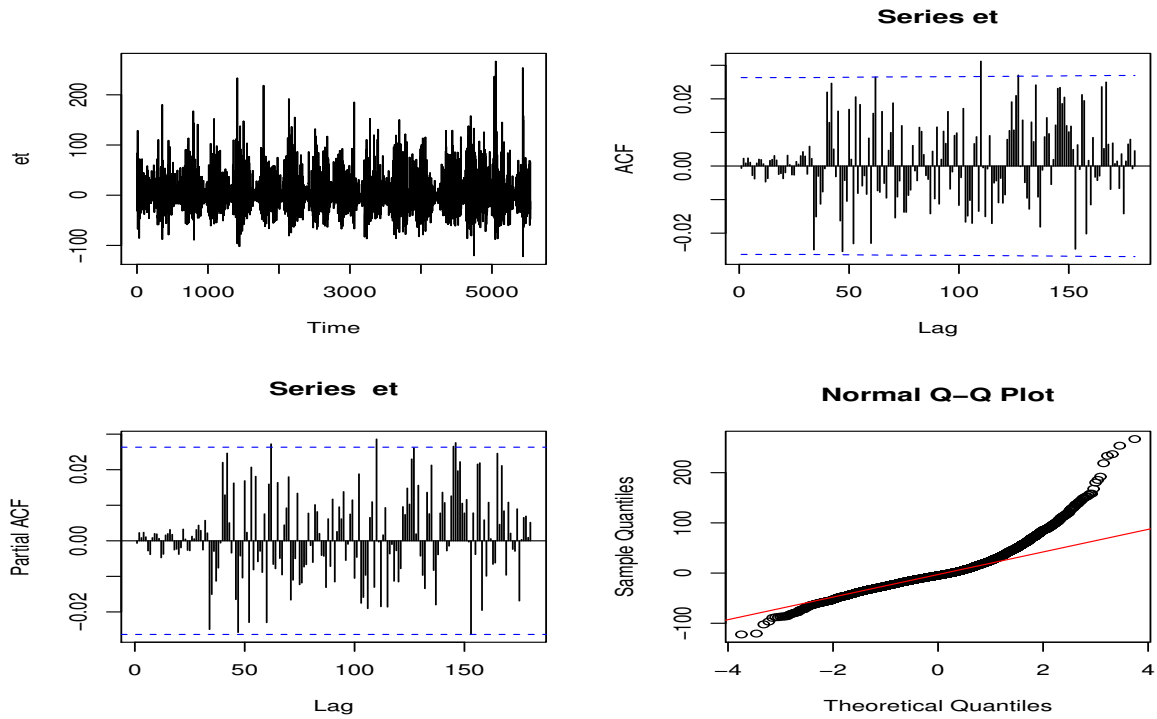



Figura 8.9: fac y facp de residuos de AR(30)

Box-Ljung test

X-squared = 23.374, df = 50, p-value = 0.9995

8.4. Procesos Autoregresivos y de Medias Móviles AR-MA(p,q)

Una manera de introducir los procesos ARMA(p,q) es la siguiente. Si en un proceso $Y_t \sim AR(p)$

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t, \quad t \in \mathbb{Z}$$

donde $\varepsilon_t \sim RB(0, \sigma^2)$, se reemplaza ε_t por un proceso $MA(q)$,

$$Z_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

entonces se obtiene un proceso ARMA(p,q),

$$\begin{aligned} Y_t &= \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + Z_t \\ &= \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}. \end{aligned}$$

Por tanto, un proceso ARMA(p,q) es un modelo que combina las propiedades de memoria larga de los AR(p) con las propiedades de ruido débilmente autocorrelacionado en los MA(q), y que tiene suficiente flexibilidad y parsimonia para representar una variedad grande de procesos estacionarios en covarianza.

Usando la notación del operador de rezago L se puede escribir un proceso ARMA(p, q) como

$$\begin{aligned} Y_t &= \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \\ \Leftrightarrow Y_t - \varphi_1 Y_{t-1} - \varphi_2 Y_{t-2} - \cdots - \varphi_p Y_{t-p} &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \\ \Leftrightarrow \Phi_p(L)(Y_t) &= \Theta_q(L)(\varepsilon_t), t \in \mathbb{Z}, \end{aligned} \quad (8.28)$$

donde

$$\Phi_p(z) = 1 - \sum_{j=1}^p \varphi_j z^j, \quad p = 1, 2, \dots \quad (8.29)$$

$$\Theta_q(z) = 1 + \sum_{j=1}^q \theta_j z^j, \quad q = 1, 2, \dots \quad (8.30)$$

son los polinomios autoregresivo y de media móvil, respectivamente.

Un proceso $Y_t \sim ARMA(p, q)$ de media cero, se define como una solución de la ecuación recursiva (8.28), dados los coeficientes φ_j , θ_j , y dado el ruido blanco $(\varepsilon_t, t \in \mathbb{Z})$. En la práctica, sin embargo, se trata es de un problema inverso: dada la serie (Y_1, \dots, Y_T) , encontrar los coeficientes (parámetros) y estimar el ruido blanco.

Las condiciones de estacionariedad de la parte AR(p) y de invertibilidad de la parte MA(q) se asumen en el modelo ARMA(p,q), (8.28). Es decir, se asume que las raíces de las ecuaciones $\Phi_p(z) = 0$ y $\Theta_q(z) = 0$ están fuera del círculo unitario. Además se asume que estos polinomios no tienen raíces en común. Si se cumplen estas condiciones el proceso $Y_t \sim ARMA(p, q)$ es estacionario e identificable. Un ARMA(p,q) de media diferente de cero se define por la ecuación

$$\Phi_p(L)(Y_t) = \varphi_0 + \Theta_q(L)(\varepsilon_t)$$

donde $\varphi_0 = \mu(1 - \sum_{j=1}^p \varphi_j)$.

Ejemplo 8.4.1. Suponga un proceso $Y_t \sim ARMA(1, 1)$ de media cero, dado por

$$\Phi_1(L)(Y_t) = \Theta_1(L)(\varepsilon_t) \quad (8.31)$$

donde $\Phi_1(L) = 1 - \varphi L$ y $\Theta_1(L) = 1 + \theta L$. Es decir $Y_t = \varphi Y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$. Si $|\varphi| < 1$ y $|\theta| < 1$ es estacionario e invertible. Por ejemplo

$$Y_t = 0.9Y_{t-1} + \varepsilon_t - 0.4\varepsilon_{t-1},$$

con $\varepsilon_t \sim RB(0, \sigma^2)$. Si se asume $E(Y_t) = \mu = 2$ entonces el proceso queda

$$\begin{aligned} Y_t &= (1 - \varphi)\mu + \varphi Y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \\ &= 2(1 - 0.9) + 0.9Y_{t-1} + \varepsilon_t - 0.4\varepsilon_{t-1}. \end{aligned}$$

Ejemplo 8.4.2. Considere el proceso $Y_t \sim ARMA(2, 1)$ dado por

$$Y_t = 2(1 - 1.5 + 0.9) + 1.5Y_{t-1} - 0.9Y_{t-2} + \varepsilon_t - 0.4\varepsilon_{t-1}$$

que equivale a

$$\begin{aligned} Y_t - 1.5Y_{t-1} + 0.9Y_{t-2} - 2(1 - 1.5 + 0.9) &= \varepsilon_t - 0.4\varepsilon_{t-1} \\ (1 - 1.5L + 0.9L^2)(Y_t - 2) &= (1 - 0.4L)\varepsilon_t, \end{aligned} \quad (8.32)$$

con ecuación característica $1 - 1.5z + 0.9z^2 = 0$, y sus raíces dadas por

$$z = \frac{1.5 \pm \sqrt{1.5^2 - 4(0.9)}}{2(0.9)} = 0.83 \pm 0.645i$$

con módulo $|z| = 1.05 > 1$. Por tanto Y_t es un proceso estacionario en covarianza. También es invertible. En este caso, es válido dividir ambos miembros de la ecuación (8.32) por $1 - 1.5L + 0.9L^2$, para obtener una expresión equivalente

$$Y_t = 2 + \frac{1 - 0.4L}{1 - 1.5L + 0.9L^2} \varepsilon_t.$$

8.4.1. Procesos ARMA y la representación de Wold

Suponga que $Y_t \sim ARMA(p, q)$ y que las raíces de las ecuaciones $\Phi_p(z) = 0$ y $\Theta_q(z) = 0$ están fuera del círculo unitario.

La capacidad de los procesos $ARMA(p, q)$ de representar una amplia variedad de procesos estacionarios en covarianza se basa el siguiente resultado conocido como Teorema de Wold.

Proposición 8.4.1. *Todo proceso estacionario en covarianza X_t , sin componentes determinísticos (ver 7.2.2, pag. 134), satisface una identidad de la forma*

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots, \quad (8.33)$$

con $\psi_0 = 1$ y $\sum_{j=0}^{\infty} |\psi_j| < \infty$, y $\varepsilon_t \sim RB(0, \sigma^2)$. La identidad (8.33) se denomina representación $MA(\infty)$ ó de Wold, de X_t .

La relación entre la representación $MA(\infty)$ y los procesos ARMA(p,q) consiste en una aproximación. Si se define la función de variable $z \in \mathbb{C}$

$$\Psi(z) := \sum_{j=0}^{\infty} \psi_j z^j, |z| < 1$$

se puede aplicar un resultado que se conoce como aproximación de Padé, ver ⁽³⁾.

Aproximación de Padé: siempre es posible encontrar una función racional, es decir, un cociente de polinomios, de la forma $\Theta_q(z)/\Phi_p(z)$, que aproxime bien a $\Psi(z)$, en el sentido de que existe una constante $C > 0$ y enteros $p, q \geq 0$, para los cuales se cumple

$$\forall z \in \mathbb{C}, \quad \left| \Psi(z) - \frac{\Theta_q(z)}{\Phi_p(z)} \right| \leq C|z|^{p+q+1}. \quad (8.34)$$

La aplicación es: si se toma un Y_t , estacionario en covarianza, entonces aplica la representación $MA(\infty)$ y debe existir la función $\Psi(z)$. Pero por la aproximación de Padé se pueden encontrar p, q tales que

$$\Psi(z) \approx \frac{\Theta_q(z)}{\Phi_p(z)} \quad (8.35)$$

Entonces Y_t puede aproximarse con un modelo ARMA(p,q), tal que $\Phi_p(L)Y_t = \Theta_q(L)\varepsilon_t$.

Fué la idea base de George E. P. Box y Gwilym Jenkins, implementada en su libro Box et al. [1976], en donde se introdujo la metodología Box-Jenkins, que resultó muy exitosa y se difundió extensamente, ver ⁽⁴⁾

En la práctica no se conoce la función $\Psi(z)$, pero se espera que, al encontrar un posible modelo ARMA(p,q), se cumpla la aproximación (8.34). Entonces el

³https://es.wikipedia.org/wiki/Aproximaci3n_de_Pad3

⁴https://en.wikipedia.org/wiki/Box-Jenkins_method

problema básico inicial es encontrar los órdenes p y q adecuados, que se conoce como el problema de identificación de un modelo ARMA(p, q).

Dados los coeficientes $a = (\varphi_1, \dots, \varphi_p)$ y $b = (\theta_1, \dots, \theta_q)$ se pueden calcular los coeficientes $(\psi_j, j = 0, 1, \dots)$ en (8.33), mediante la función `ARMAtoMA(a, b, m)`.

Ejemplo 8.4.3. Considere el proceso ARMA(2,1) dado por $(1 - L + \frac{1}{4}L^2)Y_t = (1 + L)\varepsilon_t$. La representación $MA(\infty)$ está dada por: (el primero es ψ_1)

```
ARMAtoMA(c(1.0, -0.25), c(1.0), 10)
[1] 2.00000000 1.75000000 1.25000000 0.81250000 0.50000000
[6] 0.29687500 0.17187500 0.09765625 0.05468750 0.03027344
```

El plan de análisis con procesos ARMA consiste en los siguientes pasos, con algunas librerías R para implementarlos.

1. Identificación
2. Estimación
3. Chequeo del ajuste.
4. Pronósticos con el modelo. Utilizar la función de R `predict` para pronósticos.

8.4.2. Simulación de modelos ARMA

La simulación de un proceso ARMA(p, q) se puede hacer mediante la función `arima.sim`. El ejemplo siguiente muestra los polinomios autoregresivo y de media móvil de un ARMA(3,2)

```
#----- Ejemplo la funcion arima.sim
# con los coeficientes a y b.
Phi3 = 1 - 0.1466381*x - 0.1430615*x^2 - 0.3576538*x^3
Theta2 = 1 + 1.420118*x + 0.591716*x^2
a = c(1, - 0.1466, - 0.1430, - 0.3576)
b = c(1, 1.4201, 0.5917)
```

```
y = arima.sim(list(order=c(3,0,2), ar=-a[2:4],
ma=b[2:3]), n = 300, sd=sqrt(0.3))
```

Utilizando la función `filter()`. Requiere dos aplicaciones de esta función. En la primera calcula la componente MA(q), y en la segunda calcular la componente AR(p) con errores MA(q). La instrucción es

```
#----- usar la funcion filter.
et = rnorm(n,0,sqrt(0.3))
x = stats::filter(et,b,"convolution",sides=1,circular=TRUE)
y = stats::filter(x,-a[2:4],"recursive",init=rep(0,3))
```

donde se toman los 3 valores iniciales como la media del proceso, que es cero.

La simulación de un proceso AR(p) de media μ diferente de cero, se realiza añadiendo esta cantidad en la segunda ecuación de la simulación anterior.

8.4.3. La función de autocovarianza de un proceso ARMA

La función de autocovarianza de un proceso $Y_t \sim \text{ARMA}(p,q)$ estacionario de media cero se indica por $R(k) = \text{Cov}(Y_t, Y_{t+k})$, para $k = 0, 1, \dots$. Multiplicando ambos miembros por Y_{t-k} y tomando esperanza $\mathbb{E}(\cdot)$ en la representación $MA(\infty)$ en (8.33) se obtienen las ecuaciones recursivas siguientes, similares a las ecuaciones Yule-Walker para AR(p), en (8.20). Defina $n = \max(p, q + 1)$,

$$R(k) = \varphi_1 R(k-1) + \dots + \varphi_p R(k-p) + \sigma^2 \sum_{j=k}^q \theta_j \psi_{j-k}, \quad (8.36)$$

$$k = 0, 1, \dots, n-1,$$

$$R(k) = \varphi_1 R(k-1) + \dots + \varphi_p R(k-p), \quad k = n, n+1, \dots \quad (8.37)$$

Ejemplo 8.4.4. (tomado de Brockwell and Davis [2002], pag. 93). Considere el proceso ARMA(2,1) dado por $(1 - L + \frac{1}{4}L^2)Y_t = (1 + L)\varepsilon_t$. Entonces $n = \max(p, q + 1) = 2$, por tanto, para $k = 0, 1$ se tiene el sistema lineal

$$R(0) - R(1) + \frac{1}{4}R(2) = \sigma^2(\psi_0\theta_0 + \psi_1\theta_1) = \sigma^2(1 + \psi_1),$$

$$R(1) - R(0) + \frac{1}{4}R(1) = \sigma^2(\theta_1\psi_0) = \sigma^2\psi_0 = \sigma^2. \quad (8.38)$$

Se tiene $\psi_1 = 2$ del Ejemplo anterior. Usando la segunda ecuación (8.37), se obtiene $R(2) = R(1) - \frac{1}{4}R(0)$, luego, reemplazando en el sistema (8.38), y resolviendo, se obtienen $R(0) = 32\sigma^2/3$, $R(1) = 28\sigma^2/3$. Se requiere especificar σ^2 . Utilizando la segunda ecuación (8.37),

$$R(k) = R(k-1) - \frac{1}{4}R(k-2), \quad k = 2, 3, \dots$$

se puede calcular la autocovarianza recursivamente.

Las funciones `fac` y `facp` teóricas se pueden calcular con la función `ARMAacf`, como se muestra en el código siguiente.

```
#----- fac facp teórica
Phi3 = 1 - 0.1466381*x - 0.1430615*x^2 - 0.3576538*x^3
Theta2 = 1 + 1.420118*x + 0.591716*x^2
a = c(1, - 0.1466, - 0.1430, - 0.3576)
b = c(1, 1.4201, 0.5917)
fac.k = ARMAacf(ar = -a[2:4], ma=b[2:3], lag.max = 20)
facp.k = ARMAacf(ar = -a[2:4], ma=b[2:3], lag.max = 20, pacf=TRUE)
```

8.4.4. Identificación de modelos ARMA

No es inmediato cómo identificar los modelos $ARMA(p, q)$ mediante la `fac` y `facp`. Si $q \geq p \geq 1$ entonces para $k \geq q+1$ se cumple (8.36) y para $1 \leq k \leq p-1$, $\rho(k)$ no tiene un patrón general, luego la `fac` muestral presenta un patrón definido solamente para $k \geq q+1$.

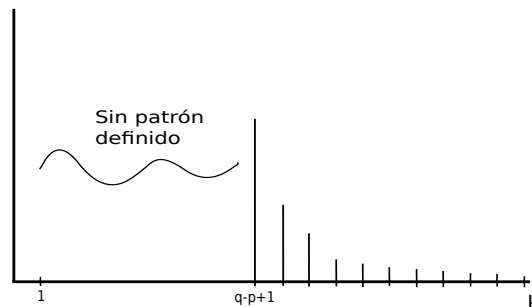


Figura 8.10: Fac Muestral de $ARMA(p, q)$.

Un método directo (“de barrido”) consiste en la identificación de p y q con el AIC. Se basa en comparar varios modelos $ARMA(p, q)$ para valores de, por ejemplo, $p = 0, 1, 2, 3$ y $q = 0, 1, 2, 3$, buscando el que produzca el menor AIC , definido por

$$AIC = -2L(\hat{\underline{\beta}}) + 2k, \quad k = p + q \quad (8.39)$$

donde $\hat{\underline{\beta}}$ está dada más adelante en (8.43) y $L()$ es la función de log-verosimilitud asumiendo distribución Normal para el ruido blanco ε_t . Se identifica un posible modelo parsimonioso con este mínimo AIC . El código siguiente en R muestra una programación básica con base en la función `arima()` para estimación por máxima verosimilitud.

Código R 8.4.1.

```
#---metodo de barrido, para escoger un posible
#   modelo ARMA(p,q), con p,q=0,1,2,3
nll = matrix(nrow=4,ncol=4,
dimnames=list(paste("p=", 0:3, sep=""),
paste("q=", 0:3, sep="")))
aic = nll
lb = nll
for (p in (0:3)) {
  for (q in (0:3)) {
    ARMAlabel = sprintf("ARMA(%d,%d)",p,q)
    armamodel = arima(y,order=c(p, 0, q),include.mean = TRUE)
    lbtest = Box.test(na.omit(armamodel$resid), lag = 12,type="Ljung")
    nll[p+1,q+1] = armamodel$loglik;
    aic[p+1,q+1] = armamodel$aic;
    lb[p+1,q+1] = lbtest$p.value}}
cat("LogLik:\n")
print(nll)
cat("AIC:\n")
print(aic)
cat("Box-Ljung-Test:\n")
print(lb)
min(aic)
```


En el lenguaje R existen varias funciones para identificar un posible modelo ARMA.

1. La función `auto.arima()` de la librería `forecast`. Para ampliar ver el sitio ⁽⁵⁾. De este sitio se toma esta nota inicial:

La función `auto.arima()` en R usa una variación del algoritmo de Hyndman-Khandakar (Hyndman & Khandakar, 2008), que combina pruebas de raíz unitaria, minimización del AICc y MLE para obtener un modelo ARIMA. Los argumentos de `auto.arima()` proporcionan muchas variaciones en el algoritmo.

2. la función `autoarmafit()` de la librería `timsac`; el programa asume que el modelo $\text{ARMA}(p,q)$ se define con $\Theta_q(z) = 1 - \sum_{j=1}^q \theta_j z^j$, es decir, los coeficientes θ_j cambian de signo en esta librería con respecto a la convención utilizada en R y en estas Notas. Sin embargo, es una función muy inestable y genera en ocasiones desbordamiento de memoria.
3. La función `armasubsets()` con la librería `library(TSA)`.

8.4.5. Estimación de modelos ARMA

La estimación de procesos $Y_t \sim \text{ARMA}(p,q)$ se basa en un supuesto de que el vector $\underline{Y} = (Y_1, \dots, Y_n)'$ se distribuye Normal multivariado con media $\underline{\mu}$, y matriz de covarianzas $\Sigma = [\text{Cov}(Y_i, Y_j)]_{n \times n}$. Como el proceso se asume estacionario, se cumple $\text{Cov}(Y_i, Y_j) = R(j-i)$, donde $R(k)$ es la función de autocovarianza de Y_t . La forma de Σ es la de una matriz tipo Toeplitz: simétrica con las sub-diagonales idénticas:

$$\Sigma = \begin{bmatrix} R(0) & R(1) & R(2) & \cdots & R(n-1) \\ R(1) & R(0) & R(1) & \cdots & R(n-2) \\ R(2) & R(1) & R(0) & \cdots & R(n-3) \\ \vdots & \vdots & & \ddots & \\ R(n-1) & R(n-2) & R(n-3) & \cdots & R(0) \end{bmatrix}. \quad (8.40)$$

$R(k)$ se calcula mediante las ecuaciones Yule-Walker, (8.36) y (8.37). Por tanto, colocando $\underline{\beta} = (\mu, \sigma^2, \varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q)'$, la matriz Σ depende del vector $\underline{\beta}$,

⁵<https://otexts.com/fpp2/arima-r.html>

y se escribe $\Sigma(\underline{\beta})$. Este supuesto permite implementar la estimación por máxima verosimilitud. Se escribe la densidad Normal Multivariada como

$$f(\underline{y}, \underline{\beta}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(\Sigma(\underline{\beta}))}} \exp \left(-\frac{1}{2} (\underline{y} - \underline{\mu})' \Sigma(\underline{\beta})^{-1} (\underline{y} - \underline{\mu}) \right) \quad (8.41)$$

donde $\underline{\mu} = (\mu, \dots, \mu)' \in \mathbb{R}^n$. La función de log-verosimilitud se define a partir del logaritmo de la densidad (8.41), $\log(f(\underline{y}, \underline{\beta}))$, y está dada por

$$\begin{aligned} L(\underline{\beta}) &:= \sum_{j=1}^n \log f(\underline{y}_j, \underline{\beta}) \\ &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log \det(\Sigma(\underline{\beta})) - \frac{1}{2} (\underline{y} - \underline{\mu})' \Sigma(\underline{\beta})^{-1} (\underline{y} - \underline{\mu}). \end{aligned} \quad (8.42)$$

El estimador ML de máxima verosimilitud se define como

La función de verosimilitud $L(\underline{\beta})$ se maximiza con respecto a los parámetros del modelo (coeficientes AR y MA) en $\underline{\beta}$.

$$\hat{\underline{\beta}} = \underset{\underline{\beta}}{\operatorname{argmin}} (-L(\underline{\beta})). \quad (8.43)$$

Para hacerlo, se requiere un algoritmo de optimización y el filtro de Kalman ⁽⁶⁾. El filtro de Kalman se usa para calcular el valor de la función de verosimilitud para un conjunto dado de valores de parámetros. No es un tema directo, requiere elaboración, ver un ejemplo en ⁽⁷⁾.

O se puede usar el algoritmo de Durbin-Levinson, en la función `DLLoglikelihood` de la librería `ltsa`.

Este valor se pasa como argumento a un algoritmo de optimización, por ejemplo, la función `R optim`.

La librería `ltsa`: Linear time series analysis, de McLeod, Yu y Krougly, tiene funciones relacionadas con la estimación. De la descripción: “Methods of developing linear time series modelling. Methods are given for loglikelihood computation, forecasting and simulation”.

Existen varias funciones del lenguaje R para estimación de modelos ARMA. Por ejemplo,

⁶https://es.wikipedia.org/wiki/Filtro_de_Kalman

⁷<https://cran.r-project.org/web/packages/FKF/vignettes/FKF.html>

1. La función `arima()` estima primero por mínimos cuadrados condicionales y luego por máxima verosimilitud.
2. La función `coeftest()` de la librería `lmtest` calcula los valores p de los coeficientes con base en el modelo estimado con `arima`. Además, con `arima()` se provee el AIC.
3. La función `arma()`, de `tseries`, estima mediante mínimos cuadrados condicionales.
4. La función `FitARMA()`, de `FitARMA` para estimación de $ARIMA(p,d,q)$ con el algoritmo Fast Maximum Likelihood de McLeod y Zhang(2007).
5. La función `armafit()`, de `timsac`, para ajuste de modelos ARMA con el algoritmo Davidon, ⁽⁸⁾. La librería es un producto del IMS, ver ⁽⁹⁾

Ejemplo 8.4.5. *Se simula un $ARMA(3,2)$ tal que las raíces de $\Phi_3(z) = 0$ y $\Theta_2(z) = 0$, tienen módulos fuera del círculo unitario. Las tres raíces de $\Phi_3(z) = 0$ se indican por $z1[1]$, $z1[2]$, $z1[3]$ y los coeficientes del polinomio autorregresivo de grado 3, mónico, $\Phi_3(z)$,*

$$\begin{aligned}\Phi_3(z) &= (1 - z/z1[1])(1 - z/z1[2])(1 - z/z1[3]) \\ &= 1 - \varphi_1 z - \varphi_2 z^2 - \varphi_3 z^3 \\ &= 1 - 0.1466381z - 0.1430615z^2 - 0.3576538z^3\end{aligned}$$

Estos coeficientes se calculan en el código siguiente en el vector `a`. El código incluye la identificación mediante la función `auto.arima`, la estimación con `arima`.

Código R 8.4.2.

```
# simulación, identificación, estimación y
# predicción de un ARMA(3,2)
#-----raices autoregresivas 3
library(polynom)
z1 = complex(3)
z1[1] = -0.8 - 1.3i
```

⁸https://en.wikipedia.org/wiki/Davidon-Fletcher-Powell_formula

⁹<https://www.ism.ac.jp/editsec/csm/index.html>

```

z1[2] = Conj(z1[1])
z1[3] = 1.2
#-----polinomio AR
a = poly.calc(z1)
a = a/a[1]
#-----raices media movil  2
z2 = complex(2)
z2[1] = -1.2 -0.5i
z2[2] = Conj(z2[1])
#-----polinomio MA
b = poly.calc(z2)
(b = b/b[1])
#----- usar la funcion arima.sim
#           con los coeficientes a y b.
n = 300
y = arima.sim(list(order=c(3,0,2), ar=-a[2:4],
ma=b[2:3]), n=n,sd=sqrt(0.3))
#-----graficar
plot.ts(y)
#----- identificar el modelo
y = ts(y,frequency=250)
auto.arima(y)
#----- para su estimacion se usa la instruccion
mod1 = arima(y1, c(3,0,2))

```

Ejemplo 8.4.6. En el artículo Nowicka-Zagrajek and Weron [2002], los autores presentan un modelo de series de tiempo ARMA(p,q) con ruido hiperbólico para la demanda diaria de electricidad en el mercado eléctrico de California. Primero eliminan la componente estacional semanal y anual.

“Los estimadores de ML utilizados aquí se basan en la suposición de ruido blanco Gaussiano. Sin embargo, esto no excluye modelos con ruido no Gaussiano dado que la distribución muestral asintótica de los estimadores es la misma para $Z_t \text{ iid}(0, \sigma^2)$, independientemente de si Z_t es o no Gaussiano, ver Brockwell and Davis [1987, Sección 10.8]”.

El modelo identificado con el menor AIC modificado es

$$X_t = 0.332776X_{t-1} + Z_t - 0.383245Z_{t-1} \\ - 0.12908Z_{t-2} - 0.149307Z_{t-3} - 0.0531862Z_{t-6}$$

donde $Z_t \sim iidHyp(\mu, \alpha, \beta, \delta)$, es una distribución Hiperbólica. La distribución Hiperbólica es una distribución de 4 parámetros, α , β (asimetría), $\delta > 0$ (escala), $\mu \in \mathbb{R}$ (localización), con $0 \leq |\beta| < \alpha$, y $\gamma = \sqrt{\alpha^2 - \beta^2}$,

$$f(x) = \frac{\gamma}{2\alpha\delta K_1(\delta\gamma)} e^{\beta(x-\mu) - \alpha\sqrt{\delta^2 + (x-\mu)^2}},$$

donde $K_\lambda(x)$ es la función de Bessel modificada de segunda clase. En el artículo se encuentra que

$$\hat{\alpha} = 1.671304, \hat{\beta} = -0.098790, \hat{\delta} = 0.298285, \hat{\mu} = 0.076975.$$

Ejemplo 8.4.7. La serie de los precios anuales del cobre desde 1800 hasta 1997 que se puede descargar en R a través de la librería `tsdl`, tomada de Time Series Data Library. El modelo identificado con el menor AIC modificado es

$$X_t = 0.332776X_{t-1} + Z_t - 0.383245Z_{t-1} \\ - 0.12908Z_{t-2} - 0.149307Z_{t-3} - 0.0531862Z_{t-6}$$

donde $Z_t \sim iidHyp(\mu, \alpha, \beta, \delta)$, es una distribución Hiperbólica. Mediante la librería de R `GeneralizedHyperbolic` se pueden estimar los parámetros y además, se puede realizar una prueba de ajuste tipo Cramer-vonMises, con la función `hyperbCvMTest`.

8.5. Modelos ARMA estacionales, SARMA

Los modelos ARMA estacionales estacionarios se introducen con el fin de modelar estacionalidades remanentes en los residuos estructurales. Por tanto, dependen del parámetro s definido como el período de la serie. Pero también son modelos útiles por sí solos.

Hasta el momento no se ha introducido un procedimiento para identificar y estimar el período en una serie dada. Es útil contar con algunos. Se pueden mencionar, por ejemplo

1. La función `periodicity` de la librería `xts`, ver Ryan and Ulrich [2017]. De la página de ayuda: “Estimate the periodicity of a time-series-like object by calculating the median time between observations in days”.

```
et = resid(modelo)
et = ts(et, frequency=1)
periodicity(et)
```

2. Un programa de aplicación por R. Hyndman ⁽¹⁰⁾

Si realmente no tiene idea de cuál es la periodicidad, probablemente el mejor enfoque sea encontrar la frecuencia correspondiente al máximo de la densidad espectral. Sin embargo, el espectro a bajas frecuencias se verá afectado por la tendencia, por lo que primero debe eliminar la tendencia de la serie. La función R siguiente debería hacer el trabajo para la mayoría de las series ... devuelve 1 para los datos que no tienen una periodicidad fuerte y, de lo contrario, la duración del período

```
find.freq <- function(x)
{
  n <- length(x)
  spec <- spec.ar(c(x), plot=FALSE)
  if(max(spec$spec)>10)
# Arbitrary threshold chosen by trial and error.
  {
    period <- round(1/spec$freq[which.max(spec$spec)])
    if(period==Inf) # Find next local maximum
    {
      j <- which(diff(spec$spec)>0)
      if(length(j)>0)
      {
        nextmax <- j[1] + which.max(spec$spec[j[1]:500])
        period <- round(1/spec$freq[nextmax])
      }
    }
  }
}
```

¹⁰<https://stats.stackexchange.com/questions/1207/period-detection-of-a-generic-time-series>

```

else
    period <- 1
}
}
else
    period <- 1
return(period) }

```

Ejemplo 8.5.1. Con los datos de demanda de energía eléctrica domiciliaria a la hora 6:00 pm, diaria, muestran una periodicidad semanal, período igual a 7.

Los modelos ARMA estacionales, S-ARMA modelan la estacionalidad incorporando, a un modelo ARMA, correlaciones entre valores de la serie separadas por un período s .

Definición 8.5.1. Un proceso $(Y_t, t \in \mathbb{Z})$ sigue un modelo ARMA Estacional de período $s > 1$, ó S-ARMA, denotado por $Y_t \sim SARMA(p, q)(p_s, q_s)[s]$, donde p, q, p_s, q_s son enteros no negativos, si para $\epsilon_t \sim RB(0, \sigma^2)$, cumple

$$\Phi_{p_s}(L^s)\Phi_p(L)Y_t = \Theta_{q_s}(L^s)\Theta_q(L)\epsilon_t, \quad (8.44)$$

donde

$$\Phi_p(L) = 1 - \sum_{j=1}^p \varphi_j L^j, \quad (8.45a)$$

$$\Theta_q(L) = 1 + \sum_{j=1}^q \theta_j L^j, \quad (8.45b)$$

$$\Phi_{p_s}(L^s) = 1 - \sum_{j=1}^{p_s} \varphi_{s,j} L^{js}, \quad (8.45c)$$

$$\Theta_{q_s}(L^s) = 1 + \sum_{j=1}^{q_s} \theta_{s,j} L^{js}. \quad (8.45d)$$

En (8.44) p, q son los órdenes AR y MA del modelo ARMA ordinario y p_s, q_s los órdenes AR y MA estacionales. Nótese que en (8.45c) y (8.45d) las potencias de L son de la forma L^s, L^{2s}, \dots

Se asume que las raíces de los cuatro polinomios anteriores están fuera del círculo unitario y además, no hay raíces comunes. De este supuesto se puede escribir el modelo de la forma siguiente

$$Y_t = \frac{\Theta_{q_s}(L^s)}{\Phi_{p_s}(L^s)} \frac{\Theta_q(L)}{\Phi_p(L)} e_t,$$

y se define $X_t = \frac{\Theta_q(L)}{\Phi_p(L)} e_t$, entonces X_t tiene autocorrelaciones tipo ARMA(p,q), y queda

$$Y_t = \frac{\Theta_{q_s}(L^s)}{\Phi_{p_s}(L^s)} X_t,$$

y por Teorema de Wold se tiene la identidad:

$$Y_t = X_t + \psi_1^{(s)} X_{t-s} + \psi_2^{(2s)} X_{t-2s} + \dots, \quad (8.46)$$

Una interpretación de (8.46) es que Y_t tiene efectos adicionales cada s períodos, incorporados a la autocorrelación entre períodos consecutivos, que se asemeja a una dinámica con autocorrelación ordinaria más estacionalidad de período s .

Ejemplo 8.5.2. El modelo $Y_t \sim ARMA(1, 1)(1, 2)[4]$ corresponde a

$$(1 - \varphi_1 L)(1 - \varphi_{4,1} L^4) Y_t = (1 + \theta_1 L)(1 + \theta_{4,1} L^4 + \theta_{4,2} L^8) \epsilon_t, \quad (8.47)$$

Y el modelo $Y_t \sim ARMA(1, 2)(1, 1)[12]$ corresponde a

$$(1 - \varphi_1 L)(1 - \varphi_{12,1} L^{12}) Y_t = (1 + \theta_1 L + \theta_2 L^2)(1 + \theta_{12,1} L^{12}) \epsilon_t, \quad (8.48)$$

Nótese que en un S-ARMA, por ejemplo el (8.48), al multiplicar los polinomios autorregresivos y de media móvil se obtiene

$$(1 - \varphi_1 L)(1 - \varphi_{12,1} L^{12}) = 1 - \varphi_1 L - \varphi_{12,1} L^{12} + \varphi_1 \varphi_{12,1} L^{13} \quad (8.49)$$

$$(1 + \theta_1 L + \theta_2 L^2)(1 + \theta_{12,1} L^{12}) = 1 + \theta_1 L + \theta_2 L^2 + \theta_{12,1} L^{12} + \theta_1 \theta_{12,1} L^{13} + \theta_2 \theta_{12,1} L^{14}. \quad (8.50)$$

luego (8.48) es un ARMA en el cual hay un coeficiente autorregresivo $\varphi_1 \varphi_{12}$ en el rezago $1 + 12 = 13$, y también hay un coeficiente de media móvil $\theta_1 \theta_{12}$ en el rezago $1 + 12 = 13$ y otro $\theta_2 \theta_{12,1}$ en el rezago $2 + 12 = 14$. Estos podrían verse en la fac y en la fac parcial, además del rezago 12 (en ocasiones se denominan rezagos satélites).

8.5.1. La simulación de los modelos SARMA

Una manera de realizar la simulación de un modelo S-ARMA es generar los coeficientes auto-regresivos y de media móvil de las partes ordinaria y estacional y colocarlos en vectores, como se muestra en (8.49) y (8.50), de tal forma que se puedan ingresar a la función `arima.sim`.

La librería `CombMSC` tiene la función `sarima.Sim` que implementa la simulación de estos modelos.

Ejemplo 8.5.3. *Un ejemplo con base en el modelo en (8.49) y (8.50)*

```
library(polynom)
# Ejemplo coeficientes s-arma(1,2)(1,1)[12]
#-----
ar = polynomial(c(1,-0.3))
ma = polynomial(c(1,0.072,- 0.629))
ars = polynomial(c(1,-0.12))
mas = polynomial(c(1,- 0.021))
#----genera coeficientes ar y ma completos
period = 12
bs = polynomial(c(rep(0,period),1))
ars = predict(ars,bs)
mas = predict(mas,bs)

fullarpoly <- ar*ars
fullmapoly <- ma*mas

(Mod(polyroot(fullarpoly)))
(Mod(polyroot(fullmapoly)))

mo <- list()
mo$ar <- coef(fullarpoly)
mo$ma <- coef(fullmapoly)

require(signal)
par(mfrow=c(1,1))
```

```

zplane(filt=rev(mo$ma), a=rev(mo$ar))
#-----
n = 300
y = arima.sim(list(order=c(13,0,14), ar=-mo$ar[-1],
  ma=mo$ma[-1]), n=n, sd=sqrt(0.3))

```

8.5.2. La identificación de los modelos SARMA

Esta observación es útil con referencia al siguiente procedimiento de identificación gráfico, en la librería `library(TSA)`, con la función `armasubsets()`. Calcula un panel para identificación para modelos ARMA y SARMA, que detecta estos rezagos. Sin embargo, tiene limitaciones gráficas para rezagos de órdenes altos. A continuación se muestra una aplicación e interpretación de los resultados de la función.

Para identificar un posible modelo SARMA en una serie con período s , el procedimiento de identificación gráfico con `armasubsets()` permite escoger los órdenes p, q, p_s, q_s con base en un panel que muestra cuáles son los modelos más opcionados con respecto al BIC.

En el panel la información sobre los rezagos a incluir en el modelo está organizada en dos sectores. A la izquierda los autorregresivos, identificada con “r-lag” y a la derecha los de media móvil, “error-lag”. Los modelos más opcionados ó factibles son los que incluyen los rezagos con menor BIC, es decir, los cuadros oscuros de la primera línea. La instrucción que grafica el panel es:

```

mod2 = lm(yi ~ t + It)
summary(mod2)
r = mod2$residuals
library(TSA)
res=armasubsets(y=r, nar=14, nma=14, y.name='r',
ar.method='ols')
plot(res)

```

En la Figura 8.11 aparecen cuatro paneles de ejemplo, generados con las instrucciones anteriores. Corresponden a un AR en la celda (1,1). En las celdas (1,2) y (2,1)

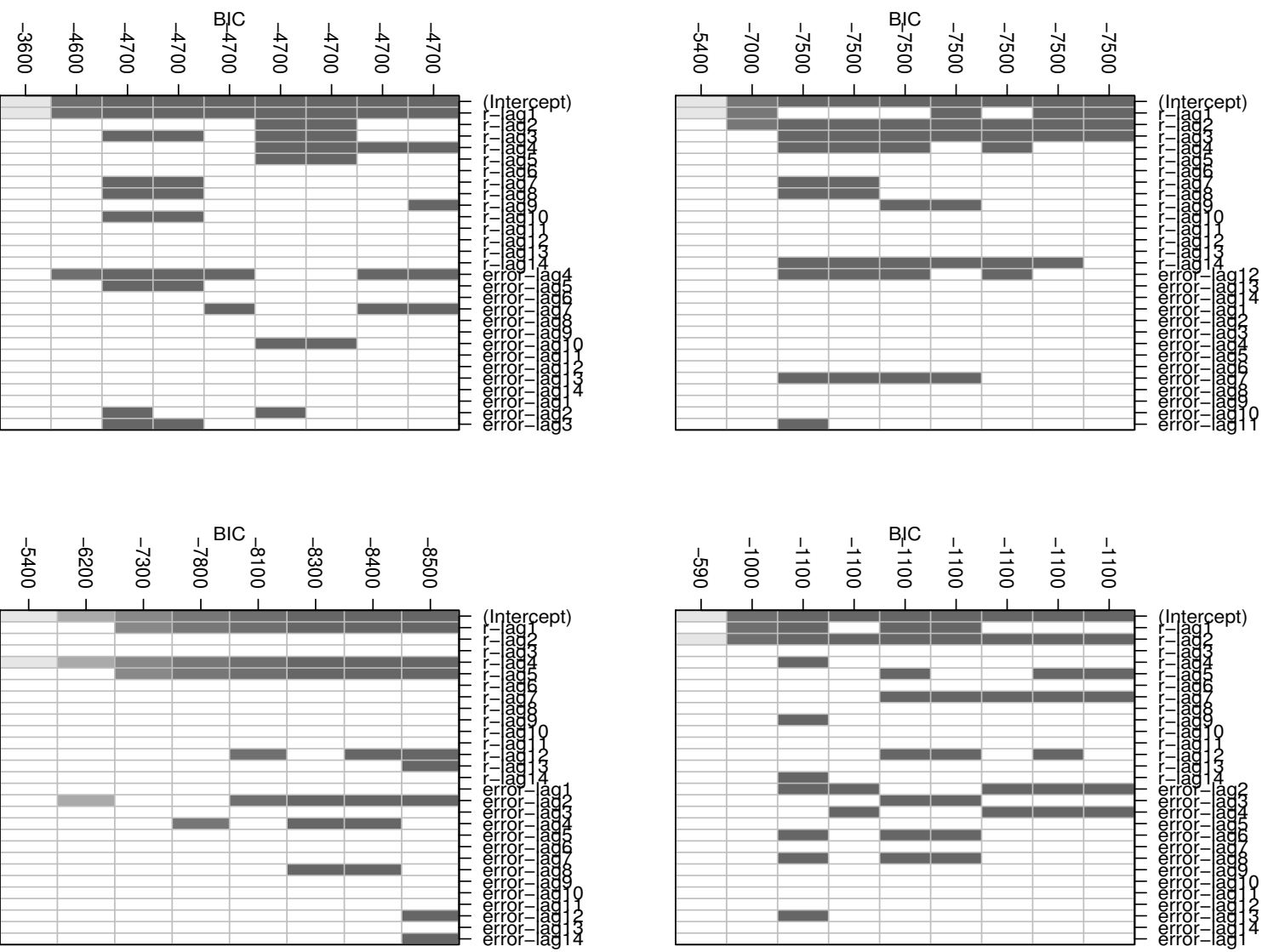


Figura 8.11: Posibles submodelos (1,1)=AR, (1,2),(2,1)=ARMA, (2,2)=SARMA

modelos ARMA. Y en la (2,2) un SARMA.

En la gráfica (1,1) aparecen muy evidentes los rezagos autorregresivos hasta $p=3$. Entonces se puede identificar un AR(3).

En la celda (1,2), arriba-derecha, se observan rezagos autorregresivos $p = 5,7$, y de media móvil $q = 2,4$. Entonces un modelo factible es un ARMA(4,4). Nótese que con `auto.arima()` se identifica un ARMA(4,4).

En la celda (2,1), abajo-izquierda, se observan rezagos autorregresivos $p = 4$. Y se observan rezagos de media móvil $q = 4,7$. Un posible modelo es un ARMA(4,4). Con `auto.arima()` se identifica un ARMA(4,5).

En la celda (2,2), abajo-derecha, se observa algo que no aparece en las anteriores. Aparecen rezagos autorregresivos 1,12,13. Lo que da lugar a suponer una estructura $(1 - \varphi L)(1 - \varphi_{12}L^{12})$. Porque así se justificaría el rezago 13 como el producto $L \times L^{12} = L^{13}$. Entonces una primera decisión es una estructura SARMA $(1, q)(1, q_{12})[12]$. En los rezagos de media móvil se observan rezagos de media móvil 2,12,14. Con lo cual una opción podría ser colocar $(1, 2)(1, 1)[12]$ porque así se justificaría el rezago 14 que proviene de $2+12 = 14$. Es el modelo siguiente.

$$(1 - \varphi_1 L)(1 - \varphi_{12,1} L^{12}) Y_t = (1 + \sum_{j=1}^2 \theta_j L^j)(1 + \theta_{12,1} L^{12}) \epsilon_t, \quad (8.51)$$

8.5.3. La estimación de los modelos SARMA

Los modelos SARMA se estiman con la función `arima()`. Por ejemplo, la estimación del modelo (8.51), $Y_t \sim ARMA(1, 2)(1, 1)[12]$, se hace con los comandos siguientes.

```
mod4 = arima(r, order=c(1, 0, 2),
seasonal = list(order = c(1, 0, 1), period = 12))
```

Este modelo tiene residuos dados por `r4 = mod4$residuals`. El análisis de inco-rrelación de éstos se hace por ejemplo con el examen de la *fac* y la prueba Ljung-Box.

Ejemplo 8.5.4. *La serie de llamadas diarias a un Call Center de cierta entidad bancaria muestra una componente estacional de período 7 días y una tendencia*

lineal, en la Figura 8.12. Se procedió a estimar estas componentes mediante un modelo de componentes con una función lineal y funciones trigonométricas. En la Figura 8.13 se pueden observar las características de los residuos del modelo estructural. Los comandos R siguientes implementan los pasos para identificar, estimar y pronosticar un modelo S-ARMA para estos residuos.

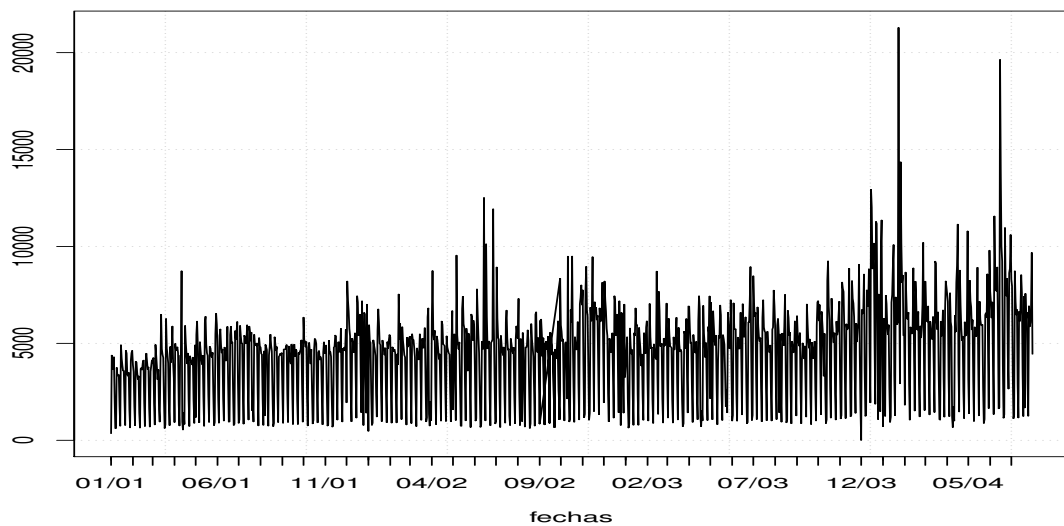


Figura 8.12: Serie de llamadas diarias a un Call Center.

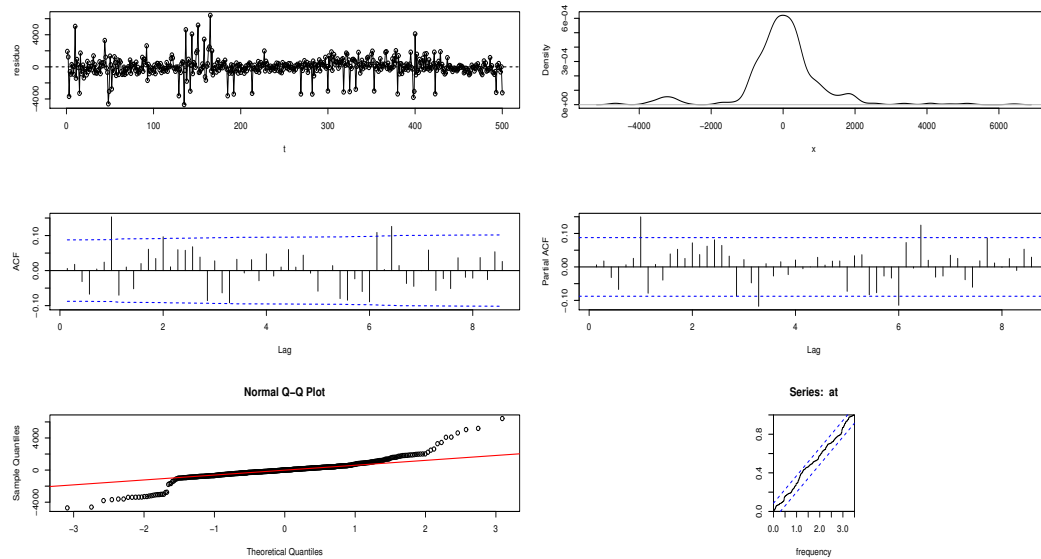


Figura 8.13: Residuos estructurales para la serie de llamadas diarias a un Call Center.
Posible modelo S-ARMA

```
#-----modelo de componentes serie llamadas call center
require(forecast)
It = fourier(y,2)
t = seq(1,length(y))
m.l = lm(y ~ t + It)
summary(m.l)
#-----residuos estructurales
yr = resid(m.l)
yr = ts(yr,frequency=7)

par(mfrow=c(3,2))
require(TSA)
plot(t,yr,type='o',ylab='residuo')
abline(h=0,lty=2)
plot(density(yr),xlab='x',main='')
acf(yr,60,drop.lag.0 = TRUE,main="")
```

```

pacf(yr, 60, drop.lag.0 = TRUE, main="")
qqnorm(yr)
qqline(yr, col=2)
#-----identificador armasubsets
require(TSA)
res=armasubsets(y=yr,
nar=14,nma=14,y.name='y',ar.method='ols')
par(mfrow=c(1,1))
plot(res)
#-----identificador auto.arima
auto.arima(yr)
ARIMA(3,0,0)(2,0,0)[7] with zero mean

```

Observamos que el modelo identificado es S-ARMA, luego el residuo estructural contiene todavía una componente estacional. La estimación con la función `arima` muestra los siguientes resultados. Nótese que se aumentó en una unidad el orden autorregresivo estacional, para obtener residuos ruido blanco.

```

m.2 = arima(yr, order=c(3,0,0), seasonal=
list(order=c(3,0,0), period=7))
require(lmtest)
coeftest(m.2)
z test of coefficients:

```

ar1	0.227720	0.028435	8.0084	1.162e-15	***
ar2	-0.060973	0.028743	-2.1213	0.03389	*
ar3	0.064545	0.027813	2.3207	0.02030	*
sar1	0.245015	0.027231	8.9976	< 2.2e-16	***
sar2	0.158905	0.028121	5.6508	1.597e-08	***
sar3	0.198965	0.028062	7.0903	1.338e-12	***
intercept	-11.869552	144.912236	-0.0819	0.93472	

Los residuos de este modelo S-ARMA(5,3)(2,0)[7] deberían ser incorrelacionados. La prueba Ljung-Box arroja el siguiente resultado

```

at = resid(m.2)
Box.test(x = at, lag = 14, type="Ljung-Box")

```

$X\text{-squared} = 19.926$, $df = 14$, $p\text{-value} = 0.1325$

Por lo que no rechaza la nula de incorrelación. Una conclusión es que la serie de llamadas tiene una estacionalidad de período 7 tanto en la componente estructural como en los residuos.

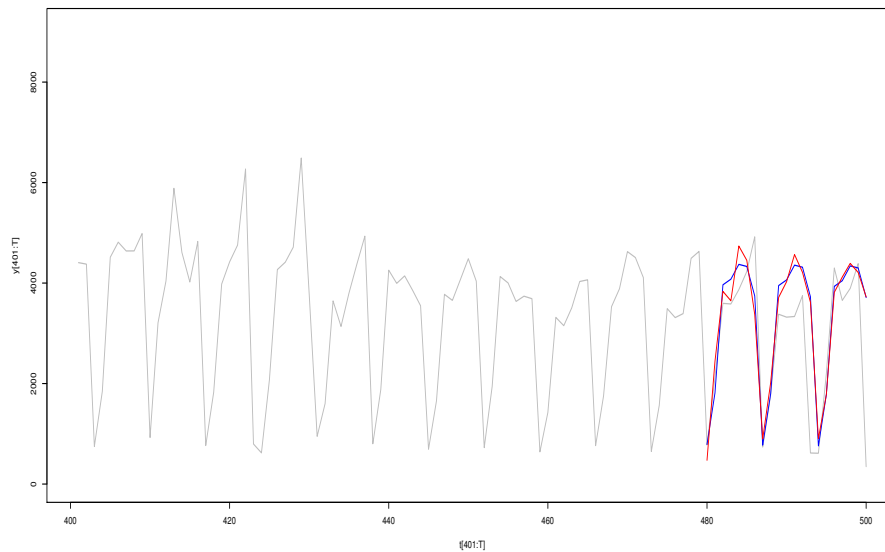


Figura 8.14: Pronósticos estructurales (azul) versus pronósticos estructurales incorporando pronósticos con S-ARMA(5,3)(2,0)[7], para la serie de llamadas diarias a un Call Center.

Los pronósticos estructurales a 21 días y los que se obtienen añadiéndoles los pronósticos del modelo S-ARMA(5,3)(2,0)[7] estimado están en la Figura 8.14

```
#-----calcular pronosticos a 21 dias
T = length(y)
tt = seq((T-21+1),T,1); Itt = fourier(yr,3,21);
py.l = predict(m.l,data.frame(t=tt,It=I(Itt)))
p.arma = predict(m.2,n.ahead=21)$pred
py.tot = py.l+p.arma
```

Ejemplo 8.5.5. Retomando el Ejemplo (7.3.2) en la pag. 141, para modelación de la serie de producción de cemento Portland, trimestral se ajustó un modelo de la

forma

$$Y_t = \beta_0 + \beta_1 t + \sum_{j=1}^3 \delta_j I_t(t) + \varepsilon_t.$$

Un esquema de análisis es el siguiente. Tomando en cuenta que anteriormente en el Ejemplo 8.2.5, en la pag. 169, se ajustó un MA(6) a los residuos estimados $\hat{\varepsilon}_t$. Ahora interesa examinar si se puede tener también $\hat{\varepsilon}_t \sim S\text{-ARMA}$, y cuál de éstos modelos produce mejores pronósticos.

Un modelo posible es SARMA(1,0,2)(1,0,2)[4], con media cero. En la Tabla 8.2 se pueden ver los resultados de la estimación, utilizando la función `arima()`. El resultado de la prueba Ljung-Box para incorrelación de los residuos es $X\text{-squared} = 23.9558$, $df = 26$, $p\text{-value} = 0.5785$.

En la Tabla 8.3 se pueden comparar las medidas de calidad de pronósticos para los modelos: Estructural (3.19), Estructural con residuos MA(6) y Estructural con residuos SARMA(1,0,2)(1,0,2)[4]. Este último modelo tiene los pronósticos de menor MAPE.

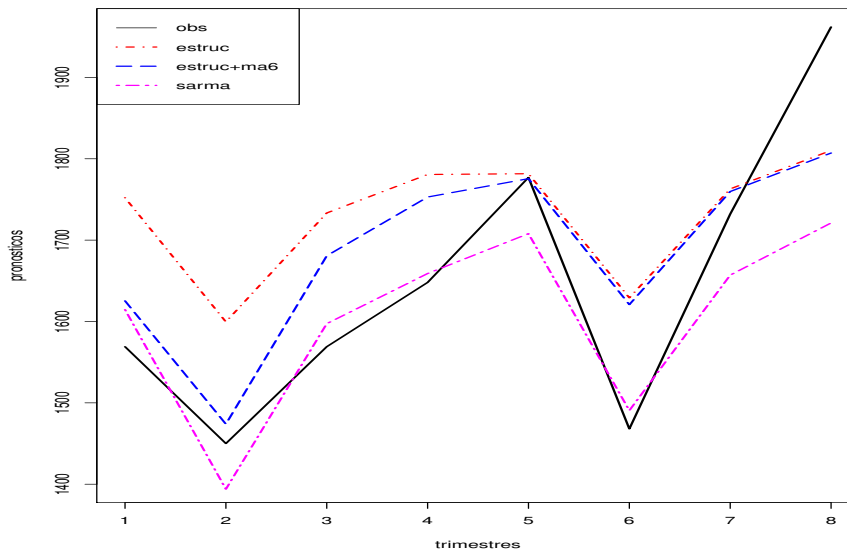


Figura 8.15: Pronósticos con tres modelos para producción de cemento

Tabla 8.2: Coeficientes estimados del modelo SARMA(1,0,2)(3,0,1)[4]

	Estimate	Std. Error	z value	Pr(> z)
ar1	0.6852	0.0991	6.9143	0.0000
ma1	0.0234	0.1144	0.2047	0.8378
ma2	0.4114	0.1097	3.7503	0.0002
sar1	1.1032	0.1511	7.3023	0.0000
sar2	-0.4470	0.1400	-3.1927	0.0014
sar3	0.2762	0.0896	3.0833	0.0020
sma1	-0.5645	0.1461	-3.8638	0.0001
intercept	-44.0386	93.4644	-0.4712	0.6375

Tabla 8.3: Comparación de medidas de calidad de pronósticos

	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
estructural	-84.3254	137.1416	122.2615	-5.7245	7.6581	0.1503	0.6695
E+ma6	-40.1263	97.0538	79.2159	-2.8445	4.8387	-0.0503	0.5103
E+sarma	41.6997	96.8725	68.5528	2.1812	3.9090	0.0445	0.5023