MODELOS DE SERIES DE TIEMPO

I CONCEPTOS PRELIMINARES

1.1 Procesos Autorregresivos y de Promedio Móvil

Se dice que Y_t sigue un **proceso autorregresivo**:

$$Y_t = \mu + \gamma Y_{t-1} + \varepsilon_t \qquad \varepsilon_t \text{ es ruido blanco} \qquad (1)$$

porque su esperanza (condicional) depende de su valor en el período pasado:

$$E(Y_t \,|Y_{t-1}) = \mu + \gamma Y_{t-1}$$

En general, un proceso autorregresivo de orden p, AR(p), se define como:

$$Y_t = \mu + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \ldots + \gamma_p Y_{t-p} + \varepsilon_t \qquad (2)$$

donde la esperanza y varianza de Y_t, condicional en sus p primeros rezagos, vienen dadas por:

$$E(Y_t \,|Y_{t-1}, Y_{t-2}, ..., Y_{t-p}) = \mu + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \ldots + \gamma_p Y_{t-p}$$

$$Var(Y_t \,|Y_{t-1}, Y_{t-2}, \ldots Y_{t-p}) = \sigma_\varepsilon^2$$

Consideremos ahora el modelo:

$$Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1} \qquad (3)$$

La variable Y_t es una constante más un proceso de **promedio móvil** de orden 1, MA(1), $\varepsilon_t + \theta\varepsilon_{t-1}$, de modo tal que:

$$E(Y_t) = E(\mu + \varepsilon_t + \theta\varepsilon_{t-1}) = \mu$$

$$Var(Y_t) = Var(\varepsilon_t + \theta\varepsilon_{t-1}) = (1+\theta^2)\sigma_\varepsilon^2$$

$$\mathrm{Cov}(Y_t, Y_{t-s}) = \mathrm{Cov}(\varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_{t-s} + \theta\varepsilon_{t-s+1}) = \begin{cases} \theta\sigma_\varepsilon^2 & s = 1 \\ 0 & s > 1 \end{cases}$$

Si hacemos uso del operador de rezagos, L, podemos reescribir la ecuación (3) como sigue[1]:

$$Y_t = \mu + (1-\theta L)\varepsilon_t$$

$$\Leftrightarrow \qquad \frac{Y_t}{1-\theta L} = \frac{\mu}{1-\theta L} + \varepsilon_t$$

con $|\theta| < 1$, $\dfrac{1}{1-\theta L} = 1 + \theta L + \theta^2 L^2 + \ldots$

Entonces

$$\frac{\mu}{1-\theta L} = (1 + \theta L + \theta^2 L^2 + \ldots)\,\mu = \frac{\mu}{1-\theta}$$

$$\frac{Y_t}{1-\theta L} = (1 + \theta L + \theta^2 L^2 + \ldots)\, Y_t = Y_t + \sum_{i=1}^{\infty} \theta^i Y_{t-i}$$

Con ello la ecuación (3) pasa a ser:

$$Y_t + \sum_{i=1}^{\infty} \theta^i Y_{t-i} = \frac{\mu}{1-\theta} + \varepsilon_t$$

$$\Leftrightarrow \qquad \boxed{Y_t = \frac{\mu}{1-\theta} - \sum_{i=1}^{\infty} \theta^i Y_{t-i} + \varepsilon_t} \qquad (4)$$

Es decir, un modelo MA(1) puede representarse como un AR(∞).

En general, un modelo **autorregresivo-promedio móvil, ARMA(p, q)** se representa como:

[1] Se tiene que $LX_t = X_{t-1}$, $L^2X_t = X_{t-2}$, etcétera.

$$Y_t = \mu + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + ... + \gamma_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q} \quad (5)$$

Como se aprecia de la ecuación (5), existen 'p' términos autorregresivos (rezagos de la variable dependiente, Y_t) y 'q' rezagos de ε_t, correspondientes a la parte del promedio móvil. El término del error ε_t se denomina "**innovación**". Es decir, éste corresponde a toda aquella nueva información que se conoce cada período.

Consideremos el caso particular de un modelo AR(1):

$$Y_t = \mu + \gamma Y_{t-1} + \varepsilon_t$$

Si utilizamos el operador de rezagos, podemos reescribir la ecuación anterior como:

$$(1-\gamma L)Y_t = \mu + \varepsilon_t$$

con $|\gamma|<1$.

$$Y_t = \frac{\mu}{1-\gamma L} + \frac{\varepsilon_t}{1-\gamma L}$$

$$= \frac{\mu}{1-\gamma} + (1 + \gamma L + \gamma^2 L^2 + ...)\,\varepsilon_t = \frac{\mu}{1-\gamma} + \sum_{i=0}^{\infty} \gamma^i\, \varepsilon_{t-i} \quad (6)$$

De lo anterior, se deduce que un proceso AR(1) puede ser representado como un proceso MA(∞). Es decir, en este caso se tiene que Y_t es una agregación de la historia de **todas** las innovaciones. En contraste, un modelo MA(q) sólo incorpora la información de las innovaciones correspondiente a los **q períodos más recientes**:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q}$$

- Para resumir, los modelos de series de tiempo pueden ser representados como procesos autorregresivos, o bien como procesos de promedio móvil. En este último caso, la variable dependiente corresponde a la suma ponderada de las innovaciones pasadas.

1.2 Vectores ARMA

Los modelos vistos anteriormente pueden generalizarse a un conjunto de variables:

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Gamma}_1 \mathbf{y}_{t-1} + \boldsymbol{\Gamma}_2 \mathbf{y}_{t-2} + \ldots + \boldsymbol{\Gamma}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t + \boldsymbol{\Theta}_1 \boldsymbol{\varepsilon}_{t-1} + \ldots + \boldsymbol{\Theta}_q \boldsymbol{\varepsilon}_{t-q} \quad (7)$$

donde $\mathbf{y}_t$ y $\boldsymbol{\varepsilon}_t$ son vectores Mx1 de variables aleatorias, $\boldsymbol{\mu}$ es el vector de medias, Mx1, y $\boldsymbol{\Gamma}_1$, $\boldsymbol{\Gamma}_2$, …, $\boldsymbol{\Gamma}_p$, $\boldsymbol{\Theta}_1$, $\boldsymbol{\Theta}_2$, …, $\boldsymbol{\Theta}_q$ son matrices de parámetros MxM:

$$\mathbf{y}_t = \begin{pmatrix} Y_{1t} \\ Y_{2t} \\ \ldots \\ Y_{Mt} \end{pmatrix} \qquad \boldsymbol{\varepsilon}_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \ldots \\ \varepsilon_{Mt} \end{pmatrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \ldots \\ \mu_M \end{pmatrix}$$

$$\boldsymbol{\Gamma}_j = \begin{pmatrix} (\Gamma_j)_{11} & (\Gamma_j)_{12} & \ldots & (\Gamma_j)_{1M} \\ (\Gamma_j)_{21} & (\Gamma_j)_{22} & \ldots & (\Gamma_j)_{2M} \\ \ldots & \ldots & \ldots & \ldots \\ (\Gamma_j)_{M1} & (\Gamma_j)_{M2} & \ldots & (\Gamma_j)_{MM} \end{pmatrix} \qquad j=1, 2, \ldots, p$$

$$\boldsymbol{\Theta}_j = \begin{pmatrix} (\Theta_j)_{11} & (\Theta_j)_{12} & \ldots & (\Theta_j)_{1M} \\ (\Theta_j)_{21} & (\Theta_j)_{22} & \ldots & (\Theta_j)_{2M} \\ \ldots & \ldots & \ldots & \ldots \\ (\Theta_j)_{M1} & (\Theta_j)_{M2} & \ldots & (\Theta_j)_{MM} \end{pmatrix} \qquad j=1, 2, \ldots, p$$

En la práctica, la mayor parte del tiempo se utilizan los llamados **vectores autorregresivos (VAR):**

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Gamma}_1 \mathbf{y}_{t-1} + \boldsymbol{\Gamma}_2 \mathbf{y}_{t-2} + \ldots + \boldsymbol{\Gamma}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (8)$$

La ecuación i-ava del modelo (8) viene dada por:

$$Y_{it} = \mu_i + \sum_{j=1}^{p} (\Gamma_j)_{i1} Y_{1,t-j} + \sum_{j=1}^{p} (\Gamma_j)_{i2} Y_{2,t-j} + ... + \sum_{j=1}^{p} (\Gamma_j)_{iM} Y_{M,t-j} + \varepsilon_{it}$$

donde $(\Gamma_j)_{im}$ indica el elemento (i, m) de Γ_j.

Veremos más detalles de estos modelos cuando discutamos los temas de causalidad a la Granger y cointegración.

1.3 Estacionariedad e Invertibilidad

Definición: Se dice que un proceso Y_t es **débilmente estacionario**[2] o estacionario en términos de covarianza si se cumplen las siguientes condiciones:

a) $E(Y_t)$ es independiente de t
b) $Var(Y_t)$ es una constante, independiente de t
c) $Cov(Y_t, Y_s)$ es una función de $|t-s|$, pero no de t ó s.

La autocovarianza para el rezago k se define como:

$$\lambda_k = Cov(Y_t, Y_{t-k}) \qquad (9)$$

Bajo estacionariedad, $\lambda_{-k} = Cov(Y_t, Y_{t+k}) = \lambda_k$ porque las autocovarianzas son función de k, pero no de t.

EJEMPLOS

i) Proceso AR(1)

$Y_t = \mu + \gamma Y_{t-1} + \varepsilon_t$ $\quad$ ε_t es ruido blanco

$E(Y_t | Y_{t-1}) = \mu + \gamma Y_{t-1}$ $\quad$ $Var(Y_t | Y_{t-1}) = \sigma_\varepsilon^2$

Como vimos, si $|\gamma|<1$, entonces Y_t se puede representar como un proceso MA(∞):

[2] Se dice que un proceso estocástico $\{Y_t\}$ es **estrictamente estacionario** o **estacionario en un sentido fuerte** si, ∀ m-tupla $(t_1, t_2, \ldots, t_m)$ y ∀ k entero, la distribución conjunta de $(Y_{t_1}, Y_{t_2}, ..., Y_{t_m})$ es idéntica a la de $(Y_{t_1+k}, Y_{t_2+k}, ..., Y_{t_m+k})$.

$$Y_t = \frac{\mu}{1-\gamma} + \sum_{i=0}^{\infty} \gamma^i \, \varepsilon_{t-i}$$

De la relación anterior, podemos obtener los **momentos incondicionales** de Y_t. En particular,

$$\boxed{E(Y_t) = \frac{\mu}{1-\gamma} + \sum_{i=0}^{\infty} \gamma^i E(\varepsilon_{t-i}) = \frac{\mu}{1-\gamma}}$$

$$Cov(Y_t, Y_{t-k}) = Cov\left(\sum_{i=0}^{\infty} \gamma^i \varepsilon_{t-i}, \sum_{i=0}^{\infty} \gamma^i \varepsilon_{t-k-i} \right) = Cov\left(\sum_{i=k}^{\infty} \gamma^i \varepsilon_{t-i}, \sum_{i=0}^{\infty} \gamma^i \varepsilon_{t-k-i} \right)$$

$$= Cov(\gamma^k \varepsilon_{t-k}, \varepsilon_{t-k}) + Cov(\gamma^{k+1} \varepsilon_{t-k-1}, \gamma \varepsilon_{t-k-1}) + Cov(\gamma^{k+2} \varepsilon_{t-k-2}, \gamma^2 \varepsilon_{t-k-2})$$

$$+....$$

$$= \gamma^k (Cov(\varepsilon_{t-k}, \varepsilon_{t-k}) + \gamma^2 Cov(\varepsilon_{t-k-1}, \varepsilon_{t-k-1}) + \gamma^4 Cov(\varepsilon_{t-k-2}, \varepsilon_{t-k-2}) + ...)$$

$$= \gamma^k (Var(\varepsilon_{t-k}) + \gamma^2 Var(\varepsilon_{t-k-1}) + \gamma^4 Var(\varepsilon_{t-k-2}) + ...)$$

$$= \gamma^k (\sigma_\varepsilon^2 + \gamma^2 \sigma_\varepsilon^2 + \gamma^4 \sigma_\varepsilon^2 + ...) = \gamma^k \frac{\sigma_\varepsilon^2}{1-\gamma^2} \qquad k=0, 1, 2...$$

bajo estacionariedad débil.

De lo anterior, $\boxed{\lambda_{-k} = \lambda_k = \gamma^k \frac{\sigma_\varepsilon^2}{1-\gamma^2}}$ k=0, 1, 2, ...

Para k=0 se tiene que $Var(Y_t) = \frac{\sigma_\varepsilon^2}{1-\gamma^2}$.

ii) Proceso MA(1)

$$Y_t = \mu + \varepsilon_t + \theta\, \varepsilon_{t-1}$$

$$E(Y_t) = \mu + E(\varepsilon_t) + \theta\, E(\varepsilon_{t-1})$$

$$Var(Y_t) = Var(\varepsilon_t + \theta\varepsilon_{t-1}) = (1+\theta^2)\sigma_\varepsilon^2$$

$$\lambda_k = \begin{cases} \theta\sigma_\varepsilon^2 & k = 1 \\ 0 & k > 1 \end{cases}$$

En general, para una serie MA(q):

$$Y_t = \mu + \varepsilon_t + \theta_1\, \varepsilon_{t-1} + \theta_2\, \varepsilon_{t-2} + ... + \theta_q\, \varepsilon_{t-q}$$

$$E(Y_t) = \mu + E(\varepsilon_t) + \theta_1\, E(\varepsilon_{t-1}) + \theta_2\, E(\varepsilon_{t-2}) + ... + \theta_q\, E(\varepsilon_{t-q}) = \mu$$

$$Var(Y_t) = Var(\varepsilon_t) + \theta_1^2 Var(\varepsilon_{t-1}) + ... + \theta_q^2 Var(\varepsilon_{t-q}) = \sigma_\varepsilon^2 (1 + \sum_{i=1}^{q} \theta_i^2)$$

$$Cov(Y_t, Y_{t-1}) = (\varepsilon_t + \theta_1\varepsilon_{t-1} + ... + \theta_q\varepsilon_{t-q},\ \varepsilon_{t-1} + \theta_1\varepsilon_{t-2} + ... + \theta_q\varepsilon_{t-q-1})$$

$$= \theta_1\, Cov(\varepsilon_{t-1}, \varepsilon_{t-1}) + \theta_1\, \theta_2\, Cov(\varepsilon_{t-2}, \varepsilon_{t-2}) + \theta_2\, \theta_3\, Cov(\varepsilon_{t-3}, \varepsilon_{t-3})$$

$$+ ... + \theta_{q-1}\, \theta_q\, Cov(\varepsilon_{t-q-1}, \varepsilon_{t-q-1}) = \left(\theta_1 + \sum_{i=1}^{q-1} \theta_i \theta_{i+1} \right) \sigma_\varepsilon^2$$

Análogamente,

$$Cov(Y_t, Y_{t-2}) = \left(\theta_2 + \sum_{i=1}^{q-2} \theta_i \theta_{i+2} \right) \sigma_\varepsilon^2,\ ...,\ Cov(Y_t, Y_{t-q}) = \theta_q \sigma_\varepsilon^2,$$

$Cov(Y_t, Y_{t-s}) = 0,\ \forall s > q.$

Esto implica que Y_t es estacionario.

iii) Proceso AR(2)

Supongamos el siguiente proceso autorregresivo de segundo orden:

$$Y_t = \mu + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \varepsilon_t$$

La expresión anterior se puede reexpresar en términos del operador de rezagos:

$$Y_t (1-\gamma_1 L - \gamma_2 L^2) = \mu + \varepsilon_t$$

Sea $C(L) = 1-\gamma_1 L - \gamma_2 L^2$ un polinomio de segundo orden en el operador de rezagos. ¿Bajo qué condiciones Y_t se puede expresar como un MA(∞)?:

$$Y_t = C^{-1}(L)(\mu + \varepsilon_t)$$

❖ Ello dependerá de los valores de γ_1, γ_2. Concretamente, si C(L) se puede expresar como:

$$1-\gamma_1 L - \gamma_2 L^2 = (1-\lambda_1 L)(1-\lambda_2 L)$$

donde $|\lambda_1|<1$, $|\lambda_2|<1$, entonces:

$$Y_t = (1-\lambda_1 L)^{-1}(1-\lambda_2 L)^{-1}(\mu + \varepsilon_t) = \left(\sum_{i=0}^{\infty}(\lambda_1 L)^i\right)\left(\sum_{i=0}^{\infty}(\lambda_2 L)^i\right)(\mu + \varepsilon_t)$$

$$\equiv \sum_{i=0}^{\infty}\pi_i(\mu + \varepsilon_{t-i})$$

dado que $\frac{1}{1-\lambda_1 L} = \sum_{i=0}^{\infty}(\lambda_1 L)^i$, $\frac{1}{1-\lambda_2 L} = \sum_{i=0}^{\infty}(\lambda_2 L)^i$.

- **Ejemplo**: Supongamos $Y_t = 0.6Y_{t-1} - 0.08\, Y_{t-2} + u_t$. Queremos descomponer $1-0.6L+0.08\, L^2$ en el producto $(1-\lambda_1 L)(1-\lambda_2 L)$. ¿Existen λ_1 y λ_2, tal que $|\lambda_1|<1$, $|\lambda_2|<1$?

$$1 - 0.6L + 0.08\, L^2 = (1-\lambda_1 L)(1-\lambda_2 L) = 1-(\lambda_1+\lambda_2)L+\lambda_1\lambda_2 L^2$$

Por lo tanto, $\lambda_1 + \lambda_2 = 0.6$, $\lambda_1\lambda_2 = 0.08$. La solución es $\lambda_1 = 0.4$, $\lambda_2 = 0.2$. De ello,

$$Y_t = (1-0.4L)^{-1}(1-0.2L)^{-1}\,\varepsilon_t = \left(\sum_{i=0}^{\infty}(0.4L)^i\right)\left(\sum_{i=0}^{\infty}(0.2L)^i\right)\varepsilon_t$$

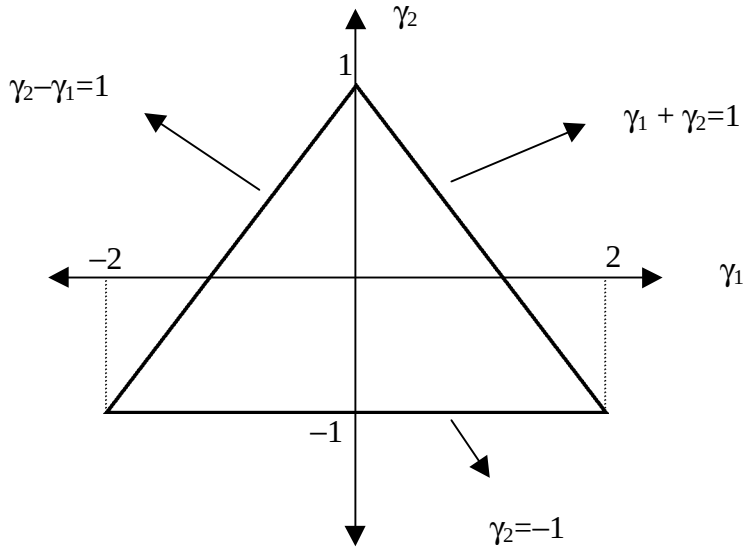
$$= (1 + 0.4L + 0.4^2L^2 + 0.4^3L^3 + ...)(1 + 0.2L + 0.2^2L^2 + 0.2^3L^3 + ...)\,\varepsilon_t$$

$$= \{1 + (0.2+0.4)L + (0.2^2 + 0.2x0.4 + 0.4^2)L^2 + (0.2^3 + 0.2^2x0.4 + 0.2x0.4^2 + 0.4^3)L^3 + ...\}\varepsilon_t$$

$$= \varepsilon_t + 0.6\varepsilon_{t-1} + 0.28\varepsilon_{t-2} + 0.12\varepsilon_{t-3} + ... \blacklozenge$$

En general, un proceso AR(2) podrá ser representado como un MA(∞) si se satisfacen las siguientes condiciones:

$$|\gamma_2| < 1 \qquad \gamma_1 + \gamma_2 < 1 \qquad \gamma_1 - \gamma_2 < 1$$



El gráfico de arriba muestra todas aquellos pares (γ_1, γ_2) que satisfacen las condiciones anteriores. En efecto, se trata de todos aquellos puntos al **interior** del triángulo con vértices (−2, −1), (2, −1), (0, 1) [3].

[3] Para una discusión, véase James Hamilton, *Time Series Analysis*, capítulo 1.

❖ ¿Cómo obtener los momentos de Y_t?

Bajo estacionariedad débil, $E(Y_t)=E(Y_{t-1})=E(Y_{t-2})\equiv\mu_Y$:

$$\mu_Y = \mu + \gamma_1 \mu_Y + \gamma_2 \mu_Y$$

$$\Rightarrow \boxed{\mu_Y = \frac{\mu}{1-\gamma_1-\gamma_2}}$$

De ello,

$$Y_t = \mu_Y(1-\gamma_1-\gamma_2) + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \varepsilon_t$$

$$\Leftrightarrow \quad Y_t - \mu_Y = \gamma_1(Y_{t-1} - \mu_Y) + \gamma_2(Y_{t-2} - \mu_Y) + \varepsilon_t$$

Si multiplicamos ambos lados de la ecuación por $Y_{t-k}-\mu_Y$ y tomamos esperanza obtenemos:

$$E(Y_t-\mu_Y)(Y_{t-k}-\mu_Y) = \gamma_1 E(Y_{t-1}-\mu_Y)(Y_{t-k}-\mu_Y)$$

$$+ \gamma_2 E(Y_{t-2}-\mu_Y)(Y_{t-k}-\mu_Y) + E(Y_{t-k}-\mu_Y)\varepsilon_t$$

donde $\lambda_k = Cov(Y_t, Y_{t-k}) = Cov(Y_t, Y_{t+k}) = \lambda_{-k}$, bajo estacionariedad.

De la ecuación anterior, para k=0, 1, 2, se tiene, respectivamente:

$$\lambda_0 = \gamma_1\lambda_1 + \gamma_2\lambda_2 + \sigma_\varepsilon^2$$
$$\lambda_1 = \gamma_1\lambda_0 + \gamma_2\lambda_1$$
$$\lambda_2 = \gamma_1\lambda_1 + \gamma_2\lambda_0$$

Ecuaciones de Yule-Walker

De las expresiones anteriores para las autocovarianzas (las llamadas **ecuaciones de Yule-Walker**), se puede obtener $Var(Y_t)=\lambda_0$:

$$\boxed{Var(Y_t) = \frac{(1-\gamma_2)\sigma_\varepsilon^2}{(1+\gamma_2)(1-\gamma_2-\gamma_1)(1-\gamma_2+\gamma_1)}}$$

- ✓ Claramente, las condiciones $|\gamma_2|<1$, $\gamma_1+\gamma_2<1$, $\gamma_1-\gamma_2<1$ aseguran que la $\text{Var}(Y_t)$ sea positiva y finita.

En general, supongamos el proceso autorregresivo:

$$Y_t = \mu + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \ldots + \gamma_p Y_{t-p} + \varepsilon_t$$

Este será **estacionario** si las raíces de la **ecuación característica**:

$$C(z)=1-\gamma_1 z-\gamma_2 z^2-....-\gamma_p z^p = 0 \qquad (10)$$

tienen un módulo mayor a 1. Esto es, si éstas yacen fuera del círculo unitario (conjunto de valores a, b tal que $a^2+b^2=1$, incluyendo $a\pm bi$, $i=\sqrt{-1}$). En dicho caso, el proceso AR(p) puede ser re-escrito como un proceso MA(∞), cuya varianza es finita[4]:

$$\boxed{Y_t = C(L)^{-1}(\mu+\varepsilon_t) \equiv \sum_{i=0}^{\infty}\pi_i(\mu+\varepsilon_{t-i})} \qquad (11)$$

donde $\sum_{i=0}^{\infty}|\pi_i| < \infty$, $\pi_0=1$

$C(L)^{-1}=(1-\gamma_1 L-\gamma_2 L^2-....-\gamma_p L^p)^{-1}=(1-\lambda_1 L)^{-1}(1-\lambda_2 L)^{-1}...(1-\lambda_p L)^{-1}$.

Ejemplo

Supongamos un proceso AR(1), $Y_t=\mu + \gamma Y_{t-1}+ \varepsilon_t$. Entonces la raíz de la ecuación característica:

$$C(z) = 1-\gamma z=0 \qquad \Leftrightarrow z = 1/\gamma$$

yace fuera del círculo unitario si $|\gamma|<1$. En dicho caso, el proceso AR(1) se puede representar como un MA(∞), tal como vimos anteriormente♦

[4] La condición $\sum_{i=0}^{\infty}|\pi_i| < \infty$ implica $\sum_{i=0}^{\infty}\pi_i^2 < \infty$.

Supongamos ahora un **proceso MA(q)**:

$$Y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + ... + \theta_q\varepsilon_{t-q}$$

Sea $D(z) = 1 - \theta_1 z - \theta_2 z^2 - ... - \theta_q z^q$. El promedio será **invertible** si todas las raíces de la ecuación característica D(z)=0 yacen fuera del círculo unitario. En dicho caso, el promedio móvil de orden q tiene una representación de AR(∞):

$$Y_t (D(L)^{-1}) = \mu(D(L))^{-1} + \varepsilon_t \qquad (12)$$

donde

$$D(L)^{-1} = (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q)^{-1} = (1 - \varphi_1 L)^{-1}(1 - \varphi_2 L)^{-1}...(1 - \varphi_q L)^{-1}$$

Ejemplo

Supongamos un proceso MA(1), $Y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1}$. Este se puede reescribir como:

$$Y_t = \mu + \varepsilon_t (1 + \theta_1 L)$$

Entonces, $D(z) = 1 + \theta_1 z = 0 \quad \Rightarrow$ z= $-1/\theta$, la cual estará fuera del círculo unitario si $|\theta|<1$ ♦

En general, si todas las raíces de la ecuación característica de un proceso de promedio móvil, D(z), están fuera del círculo unitario, la serie se dice **invertible**. Esto es, que da origen a una representación autorregresiva.

IMPORTANTE: En un proceso ARMA, la **estacionariedad** depende sólo de la parte **autorregresiva**. Ello, porque todos los procesos de promedio móvil con coeficientes finitos son estacionarios.

1.4 Autocovarianza, Autocorrelación Simple y Parcial

Recordemos la definición de autocovarianza de un proceso estocástico $\{Y_t\}$:

$$\lambda_k = Cov(Y_t, Y_{t-k})$$

La función de **autocorrelación simple** (**ACF**) se define como:

$$\rho_k = \frac{\lambda_k}{\lambda_0} \qquad -1 \leq \rho_k \leq 1 \qquad (13)$$

donde $\lambda_0 = Var(Y_t)$.

Para un proceso estacionario, ACF sólo depende de k y de los parámetros del proceso estocástico.

La **función de autocorrelación parcial** entre Y_t e Y_{t-k} **(PACF**) es la correlación simple entre Y_t e Y_{t-k}, una vez que se toma en cuenta la influencia de los k−1 primeros rezagos:

$$\rho_k^* = Corr(Y_t - E^*(Y_t \mid Y_{t-1}, Y_{t-2,}..., Y_{t-k+1}), Y_{t-k}) \qquad (14)$$

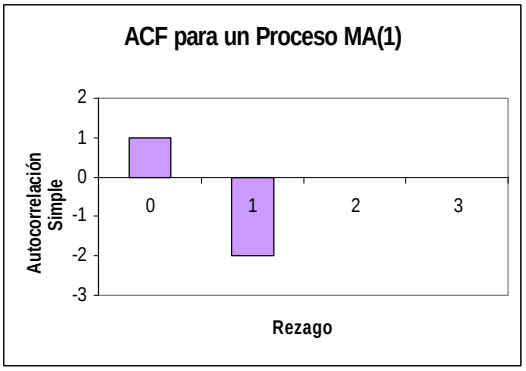
donde $E^*(Y_t \mid Y_{t-1}, Y_{t-2,}..., Y_{t-k+1}) = \mu + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_{k-1} Y_{t-k+1}$ es el mejor predictor **lineal** de Y_t en base a Y_{t-1} $Y_{t-2,}$, ..., Y_{t-k-1}.

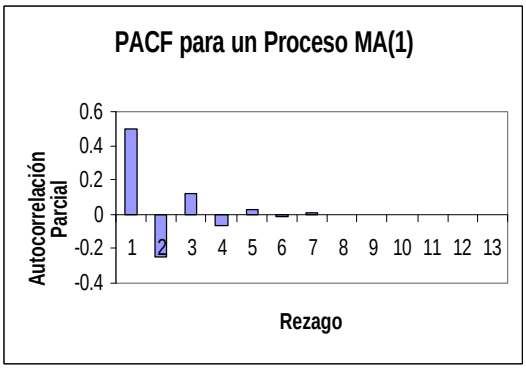
En términos prácticos, ρ_k^* es simplemente el coeficiente β_k en el modelo:

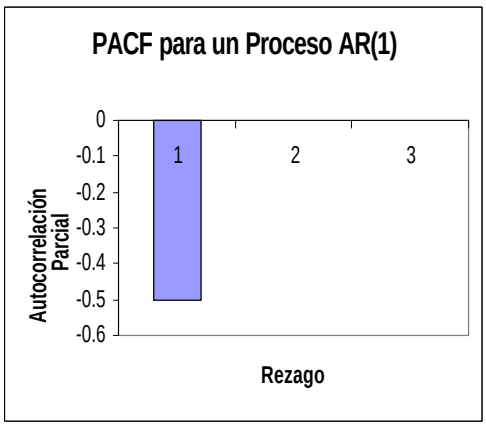
$$Y_t = \mu + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_{k-1} Y_{t-k+1} + \beta_k Y_{t-k} + \varepsilon_t$$

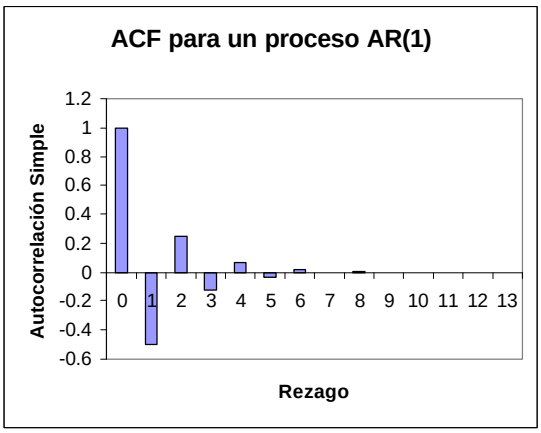
Veamos algunos ejemplos

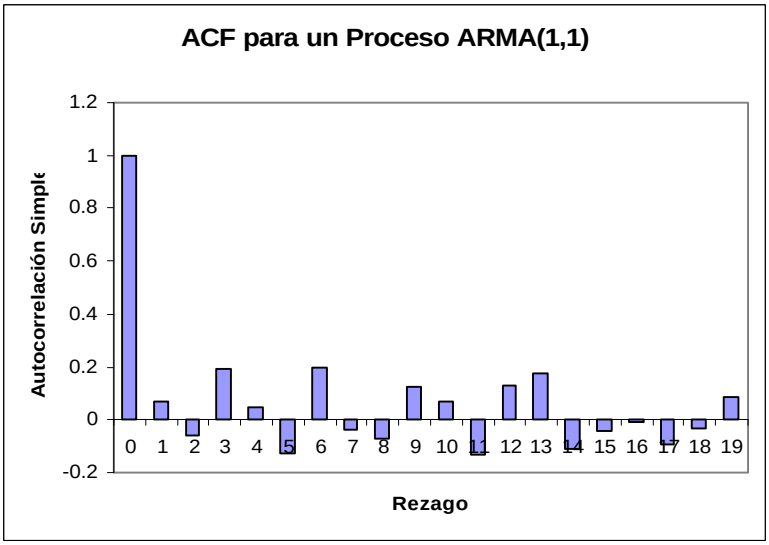
Gráfico 1 ACF y PACF para Procesos Varios

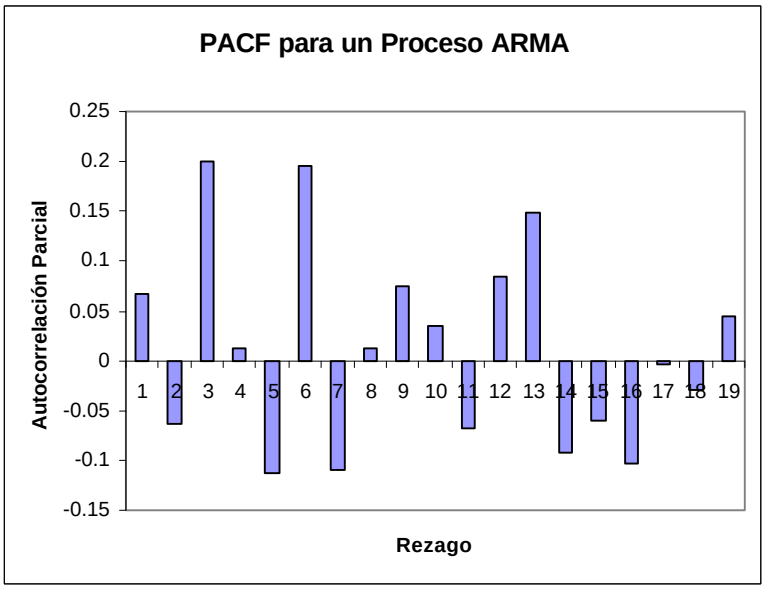












Una característica de un proceso estacionario es que su función de autocorrelación simple eventualmente converge a cero. Por ejemplo, para un proceso AR(1), $\rho_k=\gamma^k$, la cual es una serie geométrica que declina monotónicamente al aumentar k (alternando signos si $\gamma<0$).

Para un proceso AR(p) existen "p" coeficientes de autocorrelación parciales distintos de cero. En el caso de un MA(1), por ejemplo, sabemos, por la ecuación (4), que:

$$Y_t = \frac{\mu}{1-\theta} - \sum_{i=1}^{\infty} \theta^i Y_{t-i} + \varepsilon_t$$

Por lo tanto, su PACF converge a cero exponencialmente, alternando signos si $\theta<0$.

En el caso de los procesos ARMA(p,q), las funciones de autocorrelación simple y parcial son una mezcla de aquéllas de los procesos AR(p) y MA(q).

Debido a que las ACF y PACF dependen de los parámetros del proceso en cuestión, los cuales son generalmente desconocidos, trabajamos con sus análogos muestrales.

La ACF muestral o **correlograma** viene dada por:

$$\hat{\rho}_k = \frac{\frac{1}{T-(k+1)}\sum_{t=k+1}^{T}(Y_t - \overline{Y})(Y_{t-k} - \overline{Y})}{\frac{1}{T-1}\sum_{t=1}^{T}(Y_t - \overline{Y})} \qquad (15)$$

Como vimos anteriormente, los estadígrafos de Box-Pierce (Q) y Ljung-Box (Q^*) nos permiten discernir la **significancia conjunta** de los coeficientes de autocorrelación de los primeros p rezagos:

$$Q = T\sum_{k=1}^{p}\hat{\rho}_k^2 \xrightarrow{d} \chi^2(p) \qquad Q^* = T(T+2)\sum_{k=1}^{p}\frac{\hat{\rho}_k^2}{T-k} \xrightarrow{d} \chi^2(p) \qquad (16)$$

En el caso de la **función de autocorrelación parcial muestral** de Y_t, $\hat{\rho}_k^*$ corresponde al coeficiente k-avo en una regresión de Y_t en una constante y k rezagos de Y_t, $Y_t = \hat{\mu} + \hat{\beta}_1 Y_{t-1} + \hat{\beta}_2 Y_{t-2} + ... + \hat{\beta}_{k-1} Y_{t-k+1} + \hat{\beta}_k Y_{t-k} + \hat{\varepsilon}_t$:

$$\hat{\rho}_k^* = \hat{\beta}_k \qquad (17)$$

II PROCESOS INTEGRADOS Y DIFERENCIACION

2.1 Paseo aleatorio (*random walk*) con un parámetro de tendencia (*drift*)

Consideremos el proceso:

$$Y_t = \mu + Y_{t-1} + \varepsilon_t \qquad (18)$$

Podemos reescribirlo como:

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

Si hacemos este ejercicio repetidamente, tenemos:

$$Y_t = \mu T + Y_{t-T} + \sum_{i=0}^{T-1}\varepsilon_{t-i} \qquad T=1, 2, ... \qquad (19)$$

Entonces,

$$E(Y_t | Y_{t-T}) = \mu T + Y_{t-T}$$

$$Var(Y_t | Y_{t-T}) = \sigma_\varepsilon^2 T$$

con $E(\varepsilon_t)=0$, $Var(\varepsilon_t) = \sigma_\varepsilon^2 \quad \forall t$

❖ Claramente la distribución de Y_t, **condicional** en Y_{t-T}, es **NO** estacionaria: tanto su media como su varianza dependen de T.

Si reemplazamos repetidamente la expresión correspondiente a Y_{t-T} en la ecuación (19), llegamos a:

$$Y_t = \sum_{i=0}^{\infty} (\mu + \varepsilon_{t-i}) \qquad (20)$$

la cual es una suma infinita. Tanto la esperanza como la varianza **incondicional** de Y_t divergirán:

$$E(Y_t) = \sum_{i=0}^{\infty} \mu \qquad Var(Y_t) = \sum_{i=0}^{\infty} \sigma_\varepsilon^2$$

Sin embargo, la primera diferencia de Y_t es (débilmente) **estacionaria**:

$$Z_t \equiv Y_t - Y_{t-1} = \mu + \varepsilon_t \qquad (21)$$

donde $E(Z_t)=\mu$, $Var(Z_t)=\sigma_\varepsilon^2$, $Cov(Z_t, Z_{t-s})=Cov(Z_{t+s}, Z_t)=0$, $\forall s \geq 1$, NO dependen de t.

En este caso se dice que Y_t es **integrado de orden 1**, I(1), porque al tomar la primera diferencia de Y_t se tiene un proceso estacionario.

En general, se dice que una serie es **integrada de orden d**, I(d), si se vuelve (débilmente) estacionaria después de diferenciarla d veces:

$$Z_t = (1-L)^d Y_t = \Delta^d Y_t$$

donde L es el operador de rezagos.

Por ejemplo, Y_t es I(2) si $(1-L)^2 Y_t = (1-2L+L^2)Y_t = Y_t - 2Y_{t-1} + Y_{t-2}$ es estacionaria.

Una generalización de los modelos ARMA son los modelos **ARIMA** (proceso autorregresivo **integrado** con promedio móvil). Un modelo ARIMA establece que la estructura ARMA no se aplica a Y_t, sino a la variable diferenciada $\Delta^d Y_t$. En particular, un modelo ARIMA(p, d, q) viene dado por:

$$\Delta^d Y_t = \mu + \gamma_1 \Delta^d Y_{t-1} + ... + \gamma_p \Delta^d Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q} \quad (22)$$