UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

Maestría en Estadística Aplicada

Técnicas de análisis de series de tiempo

Modelos de Box Jenkins

Mg. Jesús Eduardo Gamboa Unsihuay

21 de mayo de 2019

Estructura del capítulo IV

Introducción

- Describen y pronostican series de tiempo **estacionarias**.
- Por practicidad, vamos a asumir teóricamente modelos estacionarios en $\mu = 0$, o que las observaciones están centradas.
- Suelen ser parsimoniosos (respecto a los modelos estudiados previamente).
- Es recomendable tener experiencia y conocimiento de los datos.
- No puede ser modelado “manualmente”, requiere de *software*.

Operadores

Operador retardo

Considere la serie de tiempo $Y_1, ...Y_n$, entonces el operador retardo B (*back*), a veces también denominado L (*lag*), se define como:

$$BY_t = Y_{t-1}$$

de manera general:

$$B^m Y_t = Y_{t-m}$$

para $t = m, m+1, ..., n$

Operadores

Operador adelanto

Considere la serie de tiempo $Y_1, ...Y_n$, entonces el operador adelanto F (*future*) se define como:

$$FY_t = Y_{t+1}$$

de manera general:

$$F^m Y_t = Y_{t+m}$$

para $t = 1, ..., n - m$

Operadores

Operador diferencia

Considere la serie de tiempo $Y_1, ...Y_n$, entonces el operador diferencia $\triangle$ se define como:

$$\triangle Y_t = Y_t - Y_{t-1} = Y_t - BY_t = (1-B)Y_t$$

para $t = 2, ..., n$. Entonces:

$$\triangle = 1 - B$$

Además:

$$(\triangle)^2 Y_t = \triangle Y_t - \triangle Y_{t-1} = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$$

En general:

$$(\triangle)^n Y_t = (\triangle)^{n-1} Y_t - (\triangle)^{n-1} Y_{t-1}$$

La diferenciación permite eliminar la tendencia de una serie de tiempo.

Operadores

Integración

Considere la serie de tiempo $Y_1, ...Y_n$, se dice que es integrada de orden d, $I(d)$, si la d-ésima diferencia de Y_t es un ruido blanco:

$$(\triangle)^d y_t = \epsilon_t$$

$$(1 - B)^d y_t = \epsilon_t$$

Lo más común es emplear, como máximo, $d = 2$.

Función de autocorrelación

- Mide la relación lineal entre las observaciones de la serie temporal separadas por un desfase de k unidades de tiempo:

$$\rho_k = Corr(X_t, X_{t-k})$$

- Puede truncarse o puede extinguirse. Este comportamiento sirve para identificar un modelo tentativo.
- Se **trunca** cuando la última autocorrelación significativa es grande y las restantes son significativamente iguales a cero. Un truncamiento rápido es indicador de estacionariedad.
- Se **extingue** cuando decrece de forma permanente (lenta).

Función de autocorrelación

"Al seleccionar un modelo, recuerde que las distribuciones que se muestran son teóricas y que es muy improbable que las autocorrelaciones de datos reales sea idénticas exactamente a cualquiera de las distribuciones teóricas. Sin embargo, ustede debe, mediante prueba y error, poder ubicar en forma adecuada la mayoría de series de tiempo de datos, y al ganar experiencia, la tarea se hará más fácil (Hanke y Reitsch, 1996)"

Función de autocorrelación parcial

- Es la correlación resultante entre valores de la serie que están desfasados k periodos, que no es explicada por los retrasos de orden 1 a $k-1$ (aislando el efecto de dichos retrasos). Puede entenderse como una correlación condicional:

$$\rho_{k,k} = Corr(X_t, X_{t+k} | X_{t+1}, X_{t+2}, ..., X_{t+k-1})$$

  De ahí que $\rho_1 = \rho_{1,1}$
- Es análogo al coeficiente β es un modelo de regresión lineal.
- Coincide con el k-ésimo término de un modelo Autorregresivo de orden k, es decir AR(k).
- También puede truncarse o extinguirse. Este comportamiento sirve para identificar un modelo tentativo.

Pruebas de estacionariedad
Prueba KPSS

- Fue propuesta por Kwiatkowski, Phillips, Schmidt y Shin en 1992 a fin de probar la estacionariedad de una serie de tiempo.
- Prueba principalmente si la serie es estacionaria en media (H_0) o no lo es (H_1).
- También permite evaluar si una serie presenta tendencia estacionaria (determinística) o no.
- En R puede ser obtenida con la función `kpss.test` del paquete `tseries`.

Pruebas de estacionariedad
Prueba Dickey - Fuller de raíces unitarias

- Fue propuesta por David Dickey y Wayne Fuller (1979). Si establecemos el modelo:
$$Y_t = \rho Y_{t-1} + \epsilon_t$$
- Se dice que tiene una raíz unitaria si $\rho = 1$. De ser ese el caso, la serie no sería estacionaria. Entonces si una serie presenta raíz unitaria, no es estacionaria.
- La prueba de Dickey Fuller contrasta las hipótesis:
  H_0 : Existe raíz unitaria en la serie (la serie no es estacionaria)
  H_1 : No existe raíz unitaria en la serie (la serie es estacionaria)
- En R se puede usar las funciones `adf.test` (en el paquete `tseries`) y `ur.df` (en el paquete `urca`)

Pruebas de estacionariedad

Prueba Dickey - Fuller Aumentada de raíces unitarias

Ajustando un modelo autorregresivo de desfase k, esta prueba examina la hipótesis nula de un $ARIMA(p, 1, 0)$ versus el proceso estacionario $ARIMA(p + 1, 0, 0)$, para valores de $p \leq k - 1$. Se usan las mismas funciones que Dickey Fuller, incrementando el argumento `lag`.

Pruebas de estacionariedad

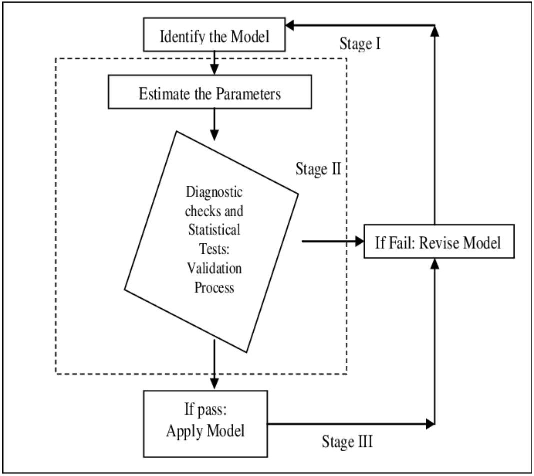
Prueba Phillips-Perron de raíces unitarias

Propuesta por Peter Phillips y Pierre Perron en 1988. Trabaja de manera similar a la prueba de Dickey Fuller. Estas pruebas serán muy importantes en la fase de identificación del modelo

Fases del modelamiento

1. Identificación
   - Determinar el modelo tentativo
   - Determinar el orden tentativo
2. Estimación
   - Uso de *software*. Mínimos cuadrados no lineal (método iterativo)
3. Diagnóstico
   - Agregar más rezagos: ¿son significativos?
   - Análisis de residuos: media cero, no autocorrelación, evaluación de indicadores (MSE, MAPE, ...)
   - Validación cruzada
   - De ser necesario, se puede volver al paso de identificación.
4. Predicción
   - Aplicar el modelo elegido para realizar pronósticos de horizonte h a partir del tiempo n.

Fases del modelamiento



Modelo AR(p)

- El modelo autorregresivo de orden p se define como:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t$$

$$Y_t = \phi_1 B Y_t + \phi_2 B^2 Y_t + \ldots + \phi_p B^p Y_t + \epsilon_t$$

$$\epsilon_t = Y_t - \phi_1 B Y_t - \phi_2 B^2 Y_t - \ldots - \phi_p B^p Y_t$$

$$\epsilon_t = Y_t(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$$

$$\epsilon_t = \phi(B) Y_t$$

  donde $\phi(B)$ es conocido como polinomio autorregresivo y $\epsilon_t \sim N(0, \sigma_e^2)$
- Un modelo de suavización exponencial simple es un AR(∞) con $\phi_i = \alpha(1-\alpha)^i$ para $i = 1, 2, \ldots.$
- Un paseo aleatorio es un AR(1) con $\phi_1 = 1$.

Modelo AR(1)

- El modelo autorregresivo de orden 1 se define como:

$$Y_t = \phi_1 Y_{t-1} + \epsilon_t$$

$$\epsilon_t = \phi(B) Y_t$$

donde $\phi(B) = 1 - \phi B$ es el polinomio autorregresivo. Este polinomio sirve en la evaluación de la estacionariedad del modelo.

Propiedades del modelo AR(1)

- El modelo AR(1) estacionario si $|\phi| < 1$:
  La raíz de B debe ser mayor que la unidad (caer fuera del círculo unitario), es decir $|B| > 1$, lo que equivale a decir $\left|\frac{1}{\phi}\right| > 1$, en consecuencia $|\phi| < 1$
- El modelo AR(1) siempre es invertible (puede expresarse como un modelo MA):

$$Y_t = \phi Y_{t-1} + \epsilon_t$$

$$Y_t = \phi(\phi Y_{t-2} + \epsilon_{t-1}) + \epsilon_t$$

$$...$$

$$Y_t = \epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + ...$$

Parámetros del modelo AR(1)

Funciones media y varianza

- Función media: dado que el proceso es estacionario:

$$E(Y_t) = E(\phi Y_{t-1} + \epsilon_t)$$

$$E(Y_t) = \phi E(Y_t)$$

$$(1-\phi)E(Y_t) = 0 \rightarrow E(Y_t) = 0$$

con la restricción de que $(1-\phi) \neq 0$, es decir $\phi \neq 1$

- Función varianza: puede demostrarse que :

$$V(Y_t) = \gamma_0 = \frac{\sigma_e^2}{1-\phi^2}$$

donde $|\phi| < 1$.

Parámetros del modelo AR(1)

Funciones de autocovarianza y autocorrelación

- Función de autocovarianza:

$$\gamma_k = \begin{cases} \frac{\sigma_e^2}{1-\phi^2} & \text{k=0} \\ \phi\gamma_{k-1} & \text{k=1,2,3,...} \end{cases}$$

- Función de autocorrelación: se extingue de manera exponencial y amortiguada .

$$\rho_k = \phi^k$$

para $k = 0, 1, 2, 3, ...$

- Función de autocorrelación parcial: se trunca luego del orden $k = 1$.

$$\rho_{k,k} \begin{cases} \neq 0 & \text{k=1} \\ = 0 & \text{k=2,3,...} \end{cases}$$

Parámetros del modelo AR(1)

Autocovarianza y autocorrelación

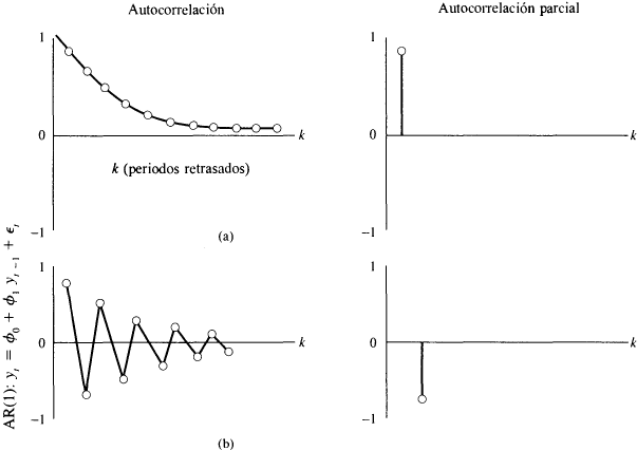


Figura: Fuente: Hanke y Reitsch

Simulación de un modelo AR(1)

Código en R

Para simular un AR(1):

```
arima.sim(model=list(ar=c(phi)), n = muestra, sd = sigma)
arima.sim(model=list(order=c(1,0,0),ar=c(phi)), n = muestra, sd = sigma)
```

donde:

- `phi` es el parámetro del modelo AR(1).
- `muestra` es el tamaño de muestra a ser simulado.
- `sigma` es la desviación del ruido blanco, es decir σ_e.

Modelo AR(2)

- El modelo autorregresivo de orden 2 se define como:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

$$\epsilon_t = \phi(B) Y_t$$

donde $\phi(B) = 1 - \phi_1 B - \phi_2 B^2$ es el polinomio autorregresivo.

Propiedades del modelo AR(2)

- El modelo AR(2) estacionario si se cumplen al mismo tiempo las 3 siguientes condiciones: $\phi_1 + \phi_2 < 1$, $-\phi_1 + \phi_2 < 1$ y $|\phi_2| < 1$:
- El modelo AR(2) siempre es invertible (puede expresarse como un modelo MA).

Parámetros del modelo AR(2)

Funciones media y varianza

- Función media:

$$E(Y_t) = E(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t)$$

  Bajo las condiciones de estacionariedad que serán impuestas más adelante:

$$E(Y_t) = 0$$

- Función varianza: puede demostrarse que :

$$V(Y_t) = \gamma_0 = \frac{\sigma_e^2}{1 - \phi_1 - \phi_2}$$

  donde ϕ_1 y ϕ_2 siguen las condiciones de estacionariedad que serán impuestas más adelante.

Parámetros del modelo AR(2)

Funciones de autocovarianza y autocorrelación

- Función de autocovarianza, para $k > 0$:

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2}$$

- Función de autocorrelación: para $t > 0$: .

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$$

- Función de autocorrelación parcial: se trunca luego del orden $k = 2$.

$$\rho_{kk} \begin{cases} \neq 0 & si\text{k=1, 2} \\ = 0 & si\text{k=3, 4, ...} \end{cases}$$

Simulación de un modelo AR(2)

Código en R

Para simular un AR(2):

```
arima.sim(model=list(ar=c(phi1,phi2)), n = muestra, sd = sigma)
```

```
arima.sim(model=list(order=c(2,0,0),ar=c(phi1,phi2)), n = muestra, sd = sigma)
```

donde:

- `phi1` y `phi2` son los parámetros del modelo AR(2).
- `muestra` es el tamaño de muestra a ser simulado.
- `sigma` es la desviación del ruido blanco, es decir σ_e.

Modelo AR(p)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t$$

- Función media

$$\mu = 0$$

- Función varianza

$$\gamma_0 = \phi_1 \gamma_1 + \ldots + \phi_p \gamma_p + \sigma_e^2$$

- Función de autocovarianza

$$\gamma_k = \phi_1 \gamma_{k-1} + \ldots + \phi_p \gamma_{k-p} \qquad k > 0$$

- Función de autocorrelación

$$\rho_k = \phi_1 \rho_{k-1} + \ldots + \phi_p \rho_{k-p} \qquad k > 0$$

- Función de autocorrelación parcial:

$$\rho_{kk} = 0 \quad k > p$$

Modelo MA(q)

- El modelo de medias móviles de orden q se define como:

$$Y_t = \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - ... - \theta_q\epsilon_{t-q}$$
$$Y_t = \epsilon_t - \theta_1 B\epsilon_t t - \theta_2 B^2\epsilon_t - ... - \theta_q B^q\epsilon_t$$
$$Y_t = (1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q)\epsilon_t$$
$$Y_t = \theta(B)\epsilon_t$$

  donde $\theta(B)$ es conocido como polinomio de medias móviles y $\epsilon_t \sim N(0, \sigma_e^2)$
- Media móvil ponderada de ruidos blancos.
- No es tan intuitivo como un modelo AR(p) pero es útil.
- Importante: En R el modelo MA(q) es especificado como:

  X[t] = e[t] + b[1]e[t-1] + . . . + b[q]e[t-q]

  Es decir:

$$\theta_j^* = -\theta_j$$

  donde θ_j^* =b[j]

Modelo MA(1)

- El modelo de medias móviles de orden 1 se define como:

$$Y_t = \epsilon_t - \theta\epsilon_{t-1}$$

$$Y_t = \theta(B)\epsilon_t$$

donde $\theta(B) = 1 - \theta B$ es el polinomio de medias móviles.

Parámetros del modelo MA(1)

Media, varianza y autocovarianza

- Función media:

$$E(Y_t) = E(\epsilon_t - \theta\epsilon_{t-1}) = 0$$

- Función varianza: puede demostrarse que:

$$V(Y_t) = \gamma_0 = (1 + \theta^2)\sigma_e^2$$

- Función de autocovarianza: distinta de cero para $k = 1$:

$$\gamma_k = \begin{cases} -\theta\sigma_e^2 & si \text{k=1} \\ 0 & si \text{k=2,3,...} \end{cases}$$

Parámetros del modelo MA(1)

Autocorrelación y autocovarianza

- Función de autocorrelación:

$$\rho_k = \begin{cases} \frac{-\theta}{1+\theta^2} & si k=1 \\ 0 & si k=2,3,... \end{cases}$$

- Función de autocorrelación parcial: ρ_{kk} se extingue de manera exponencial amortiguada.

Parámetros del modelo MA(1)

Autocovarianza y autocorrelación

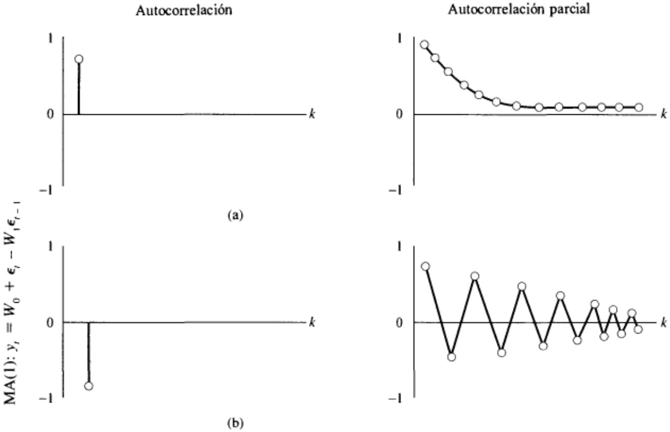


Figura: Fuente: Hanke y Reitsch

Parámetros del modelo MA(1)

Autocovarianza y autocorrelación

- Teóricamente no existen en simultáneo:

$$\gamma_k \begin{cases} \neq 0 & si \text{k=1,2,...,p} \\ = 0 & si \text{k=p+1, p+2, ...} \end{cases}$$

$$\rho_k \begin{cases} \neq 0 & si \text{k=1,2,...,p} \\ = 0 & si \text{k=p+1, p+2, ...} \end{cases}$$

- Sin embargo pueden darse muestralmente.
- Evaluar entre AR(1) y MA(1):
  - ¿cuál se corta de manera más abrupta?
  - Contribución significativa
  - Indicadores (RMSE, MAPE, etc)
  - AIC

Simulación de un modelo MA(1)

Código en R

Para simular un MA(1):

```
arima.sim(model=list(ma=c(theta)), n = muestra, sd = sigma)
```

```
arima.sim(model=list(order=c(0,0,1),ma=c(theta)), n = muestra, sd = sigma)
```

donde:

- `theta` es el parámetro del modelo MA(1).
- `muestra` es el tamaño de muestra a ser simulado.
- `sigma` es la desviación del ruido blanco, es decir σ_e.

Simulación de un modelo MA(1)

Actividad:

1. Simule un modelo MA(1) con $\theta = 0.9$ y $\sigma_\epsilon^2 = 4$
2. Compare la función media teórica y muestral.
3. Compare la función de varianza teórica y muestral.
4. Compare la función de autocorrelación teórica y muestral.
5. Compare la función de autocorrelación parcial teórica y muestral.
6. Repita este procedimiento para otros valores de ϕ y σ_e

Propiedades del modelo MA(1)

- Un modelo MA(1) siempre es estacionario.
- Un modelo MA(1) es invertible si $|\theta| < 1$ (que el modelo no sea explosivo).

Modelo MA(2)

- El modelo de medias móviles de orden 2 se define como:

$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}$$

$$Y_t = \theta(B)\epsilon_t$$

donde $\theta(B) = 1 - \theta_1 B - \theta_2 B^2$ es el polinomio de medias móviles.

Parámetros del modelo MA(2)

Media, varianza y autocovarianza

- Función media:

$$E(Y_t) = E(\epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2}) = 0$$

- Función varianza: puede demostrarse que:

$$V(Y_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma_e^2$$

- Función de autocovarianza: distinta de cero para $k = 1, 2$:

$$\gamma_k = \begin{cases} (-\theta_1 + \theta_1\theta_2)\sigma_e^2 & si \text{k=1} \\ -\theta_2\sigma_e^2 & si \text{k=2} \\ 0 & si \text{k=3,...} \end{cases}$$

Parámetros del modelo MA(2)

Autocorrelación y autocovarianza

- Función de autocorrelación: se corta luego del periodo q:

$$\rho_k = \begin{cases} \frac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2} & si k=1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & si k=2 \\ 0 & si k=3,... \end{cases}$$

- Función de autocorrelación parcial: ρ_{kk} se extingue de manera exponencial amortiguada.

Simulación de un modelo MA(2)

Código en R

Para simular un MA(2):

```
arima.sim(model=list(ma=c(theta1,theta2)), n = muestra, sd = sigma)
```

```
arima.sim(model=list(order=c(0,0,2),ma=c(theta1,theta2)),
n = muestra, sd = sigma)
```

donde:

- `theta1` y `theta2` son los parámetros del modelo MA(2).
- `muestra` es el tamaño de muestra a ser simulado.
- `sigma` es la desviación del ruido blanco, es decir σ_e.

Simulación y propiedades del modelo MA(2)

Actividad:

1. Investigue acerca de los patrones de las funciones de autocorrelación simple y parcial en un modelo MA(2)
2. Simule modelos MA(2) que permitan identificar los patrones investigados. Recuerde la inversión del signo en R
3. ¿Bajo qué condiciones un modelo MA(2) es estacionario?
4. ¿Bajo qué condiciones un modelo MA(2) es invertible?

Modelo MA(q)

$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q}$$

- Función media

$$\mu = 0$$

- Función varianza

$$\gamma_0 = (1 + \theta_1^2 + \ldots + \theta_q^2)\sigma_e^2$$

- Función de autocovarianza

$$\gamma_k = (-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \ldots + \theta_{q-k}\theta_k)\sigma_e^2 \qquad k = 1, \ldots, q$$

$$\gamma_k = 0 \qquad k > q$$

- Función de autocorrelación

$$\rho_k = \frac{-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \ldots + \theta_{q-k}\theta_k}{1 + \theta_1^2 + \ldots + \theta_q^2} \qquad k = 1, \ldots, q$$

$$\rho_k = 0 \qquad k > q$$

- Función de autocorrelación parcial: se extingue

Modelo ARMA(p,q)

- El modelo de autorregresivo de media móvil de orden (p, q) se define como:

$$Y_t = \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - ... - \theta_q \epsilon_{t-q}$$

- Un modelo ARMA(p,0) es equivalente a un AR(p)
- Un modelo ARMA(0,q) es equivalente a un MA(q)
- El modelo ARMA suele ser más parsimonioso que AR o MA.

Modelo ARMA(1,1)

- El modelo de autorregresivo de media móvil de orden $(1, 1)$ se define como:

$$Y_t = \phi Y_{t-1} + \epsilon_t - \theta \epsilon_{t-1}$$

Parámetros del modelo ARMA(1,1)

Media, varianza y función de autocovarianza

- Función media

$$E(Y_t) = 0$$

- Función varianza:

$$V(Y_t) = \gamma_0 = \frac{1 + \theta^2 - 2\theta\phi}{1 - \phi^2}\sigma_e^2$$

- Función de autocovarianza:

$$\gamma_k = \begin{cases} \phi\gamma_0 - \theta\sigma_e^2 & si \text{ k=1} \\ \theta^{k-1}\gamma_1 & si \text{ k=2,3,...} \end{cases}$$

Parámetros del modelo ARMA(1,1)

Función de autocorrelación y autocorrelación parcial

- Función de autocorrelación:

$$\rho_k = \begin{cases} 1 & si \text{k=0} \\ \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta} & si \text{k=1} \\ \phi^{k-1}\rho_1 & si \text{k=2,3,4,...} \end{cases}$$

- Función de autocorrelación parcial, va decreciendo exponencialmente.

Parámetros del modelo ARMA(1,1)

Propiedades del modelo

- El proceso es estacionario si $|\theta| < 1$
- El proceso es invertible si $|\phi| < 1$

Simulación de un ARMA(1,1)

Código en R

Para simular un ARMA(1,1):

```
arima.sim(model=list(ar=c(phi),ma=c(theta)), n = muestra, sd = sigma)

arima.sim(model=list(order=c(1,0,1), ar=c(phi),ma=c(theta)),
n = muestra, sd = sigma)
```

donde:

- `phi` es el parámetro del componente AR(1).
- `theta` es el parámetro del componente MA(1).
- `muestra` es el tamaño de muestra a ser simulado.
- `sigma` es la desviación del ruido blanco, es decir σ_e

Modelo ARIMA

Definición

Una serie de tiempo $Y_1, ..., Y_n$ sigue un proceso ARIMA(p,d,q) si su d-ésima diferencia es un proceso ARMA(p,q). Se puede definir la relación:

$$Z_t = (1 - B)^d Y_t$$

donde Z_t sigue un ARMA(p,q) y Y_t un ARIMA(p,d,q). Entonces, la relación de ARMA

$$\phi_p(B) Z_t = \theta_q(B) \epsilon_t$$

es llevada a ARIMA para la serie de tiempo original (Y):

$$\phi_p(B)(1 - B)^d Y_t = \theta_q(B) \epsilon_t$$

Simulación de un ARIMA(1,d,1)

Código en R

Para simular un ARIMA(1,d,1):

```
arima.sim(model=list(order=c(1,d,1), ar=c(phi),ma=c(theta)),
n = muestra, sd = sigma)
```

donde:

- `d` es el número de diferenciaciones que debe realizarse para conseguir estacionariedad en la serie.
- `phi` es el parámetro del componente AR(1).
- `theta` es el parámetro del componente MA(1).
- `muestra` es el tamaño de muestra a ser simulado.
- `sigma` es la desviación del ruido blanco, es decir σ_e

Modelo ARIMA

Modelamiento

1. La serie debe diferenciarse d veces, hasta que se consiga estacionariedad. Generalmente $d = 0$, $d = 1$ o $d = 2$.
2. Analizar las funciones de autocorrelación y autocorrelación parcial a fin de dar valores tentativos para p (AR) y q (MA).
3. Estimar los parámetros del o de los modelos tentativos.
4. En caso sea más de 1 (es lo usual), comparar los modelos.
5. Pronosticar para el horizonte h deseado.

Estacionalidad

- La estacionalidad ocasiona fuerte asociación entre retardos de orden s
- Se definen los operadores de diferencia y retardo estacional $\triangle^s$ y B^s, respectivamente:

$$\triangle^s y_t = y_t - y_{t-s}$$

$$B^s y_t = y_{t-s} \qquad (B^s)^n y_t = y_{t-sn}$$

$$\triangle^s y_t = (1 - B^s) y_t$$

- Recordar que:

$$(\triangle)^s \neq \triangle^s$$

$$(\triangle)^s = (1 - B)^s \qquad \triangle^s = 1 - B^s$$

- Tal como la diferencia simple, la estacional también puede llevarse a cabo D veces:

$$(\triangle^s)^D y_t = (1 - B^s)^D y_t$$

Estacionalidad

			Estacionalidad	
		No	Aditiva (Estacionaria)	Multiplicativa (No estacionaria)
Tendencia	No	AR(p)	ARIMA(p,0,q)×(P,0,Q)s	ARIMA(p,0,q)×(P,D,Q)s
		MA(q)		
		ARMA(p,q)		
	Sí	ARI(p,d)	ARIMA(p,d,q)×(P,0,Q)s	ARIMA(p,d,q)×(P,D,Q)s
		IMA(d,q)		
		ARIMA(p,d,q)		

Modelo SARIMA

Una serie de tiempo $Y_1, ..., Y_n$ sigue un proceso ARIMA(p,d,q)×(P,D,Q)$_s$ si su d-ésima diferencia simple y D-ésima diferencia estacional es un proceso ARMA(p,q). Se puede definir la relación:

$$Z_t = (1 - B)^d (1 - B^s)^D Y_t$$

donde Z_t sigue un ARMA(p,1) y Y_t un SARIMA(p,d,q)×(P,D,Q)$_s$. Entonces:

$$\phi_p(B)\Phi_P(B^s)(1 - B)^d (1 - B^s)^D Y_t = \theta_q(B)\Theta_Q(B^s)\epsilon_t$$

por ejemplo, un modelo ARIMA con valores $s = 12$, $p = d = Q = 0$ y $P = q = D = 1$:

$$\Phi_1 \left(B^{12}\right)\left(1 - B^{12}\right) Y_t = \theta_1(B)\epsilon_t$$

$$\left(1 - \Phi_1 B^{12}\right)\left(1 - B^{12}\right) Y_t = (1 - \theta_1 B)\, \epsilon_t$$

al ser expresado por extenso:

$$Y_t = (1 + \Phi)Y_{t-12} - \Phi Y_{t-24} + \epsilon_t - \theta\epsilon_{t-1}$$

Bibliografía

- Bowerman, B; O'Connell, R; Koehler, A. Pronósticos, series de tiempo y regresión, un enfoque aplicado. Cengage Learning, 4ta edición. México, 2006.
- Cryer, J; Chan, K. Time Series Analysis With Applications in R.Springer, 2da edición. EEUU. 2008.
- Court, E; William, E. Estadísticas y Econometría Financiera. Cengage Learning. 1ra edición. Argentina, 2011.
- Cowpertwait, P; Metcalfe, A. Introductory Time Series with R. Springer. EEUU, 2009
- Hanke, J; Reitsch, A. Pronósticos en los negocios. Prentice Hall, 5ta edición. México. 1996.
- Morettin, P; Toloi, C. Análise de séries temporais. Editora Blucher, 2da edición. Brasil, 2006.