Modelos univariantes de series temporales

Alfonso Novales
Departamento Economía Cuantitativa
Universidad Complutense

Procesos estocásticos

Conceptos y estadísticos

Proceso estocástico

- Un proceso estocástico es una sucesión de variables aleatorias.
- Consideramos procesos estocásticos indexados por la variable tiempo.
- Cada una de estas variables aleatorias tiene su propia distribución de probabilidad
- Una serie temporal de datos se interpreta como una realización de un proceso estocástico. Cada dato es una realización (una muestra de tamaño uno) de la distribución de probabilidad de la variable aleatoria correspondiente a dicho momento.

Procesos estocásticos fundamentales

- Un *ruido blanco* es una sucesión de variables aleatorias independientes en el tiempo, con esperanza matemática cero e igual varianza $\{\varepsilon_t\}$.

- Un *paseo o camino aleatorio* es un proceso estocástico $\{y_t\}$, cuyas primeras diferencias constituyen un ruido blanco,

$$y_t = y_{t-1} + \varepsilon_t \Rightarrow y_t - y_{t-1} = \varepsilon_t, \quad t = \ldots,-1,-2,0,1,2,\ldots$$

- *Ejercicios de simulación*

Estacionariedad

- Un proceso estocástico es *estacionario* si la distribución de probabilidad de cualquier subconjunto de variables del mismo (no necesariamente consecutivas) es constante en el tiempo.

- En particular, todas las variables aleatorias que configuran el proceso han de tener la misma distribución de probabilidad.

- Esto es muy exigente. Nos conformamos con el concepto de *estacionariedad débil o de segundo orden*: las variables individuales, así como un par de variables separado por *s* periodos, tiene distribución de probabilidad invariante en el tiempo.

- El ruido blanco es un proceso estocástico estacionario de segundo orden.

El camino aleatorio no es estacionario

- El *camino aleatorio* no es un proceso estacionario. A partir de un valor inicial, el valor en *t* puede escribirse,

$$y_t = y_{t-1} + \varepsilon_t = y_{t-2} + \varepsilon_{t-1} + \varepsilon_t = \ldots = y_0 + \sum_{s=1}^{t} \varepsilon_s$$

por lo que su esperanza es,

$$Ey_t = y_0 + E\left(\sum_{s=1}^{t} \varepsilon_s\right) = y_0,\ t = 1, 2, 3, \ldots$$

- Pero su varianza,

$$Var(y_t) = Var\left(\sum_{s=1}^{t} \varepsilon_s\right) = t\sigma_\varepsilon^2,\ t = 1, 2, 3, \ldots$$

crece (cambia) en el tiempo sin límite.

Funciones de autocovarianza y de autocorrelación simple

- La *función de autocovarianza* de un proceso estocástico es,

$$\gamma_t^k = Cov(y_t, y_{t-k})$$

que será una función únicamente de *k* en un proceso estacionario de segundo orden.

- La *función de autocorrelación simple* es,

$$\rho_t^k = Corr(y_t, y_{t-k}) = \frac{\gamma_t^k}{\sqrt{Var(y_t)}\sqrt{Var(y_{t-k})}}$$

- En un proceso estacionario de segundo orden,

$$\rho_t^k = Corr(y_t, y_{t-k}) = \frac{\gamma_t^k}{\sigma_y^2}$$

Función de autocorrelación parcial

- La *función de autocorrelación parcial* de un proceso estocástico es una función que para cada instante t y cada entero k toma un valor numérico igual a la correlación entre y_t e y_{t-k} corregida del efecto común de los retardos intermedios $y_{t-1}, y_{t-2},..., y_{t-k+1}$.

- Se estima mediante una serie de de regresiones, explicando y_t , en desviaciones respecto a su media muestral, por k retardos suyos, k =1,2,3,..., y tomando en cada caso el coeficiente estimado en el último retardo.

Funciones de autocorrelación: significación estadística

- Si el verdadero modelo es *AR*(*p*), la distribución de los valores estimados de la ***fap*** es, asintóticamente N(0,1/T) (aprox.) para *j*>*p*, por lo que examinamos si dichos valores caen en el intervalo (-2/T,2/T). Si a partir de un cierto orden *p* dichos valores caen dentro del intervalo, no rechazaremos la hipótesis de que el proceso es *AR* de orden igual o inferior a *p*.

- El mismo procedimiento se aplica a la estimación de la ***fas*** para contrastar si el proceso es *MA* de orden igual o inferior a *q*.

Funciones de autocorrelación: significación estadística

- Si el proceso no es ruido blanco, el cociente 1/T puede no ser una buena aproximación a la varianza de las estimaciones de las ***fas, fap***.

- Si ρ_k es distinto de cero para $k \leq q$ e igual a cero para $k > q$, entonces la varianza de r_k es aproximadamente [Bartlett],

$$Var(r_k) = \frac{1 + 2\left(\rho_1^{\ 2} + \rho_2^{\ 2} + \ldots + \rho_k^{\ 2}\right)}{T}$$

lo que explica la aproximación 1/T cuando se quiere contrastar la estructura de ruido blanco

Funciones de autocorrelación: significación estadística

- La fórmula de Bartlett tiende a sobreestimar la varianza de los valores de las ***fas, fap***, por lo que sus intervalos de confianza tienen tendencia a ser "demasiado" amplios. El sesgo es hacia no detectar estructura en algunos casos en que ésta existe.

- Lo que sugiere contrastar significación a niveles de 80%, 85%, es decir, utilizar umbrales críticos de 1,4 (aprox.).

- En general, 1/T es una cota superior para la varianza de ***fas, fap.***

Hipótesis nula:
“los residuos son ruido blanco”

- Box-Pierce, $$Q_{BP} = T(r_1^2 + r_2^2 + ... + r_k^2)$$

  se distribuye *chi-cuadrado* con *k-p-q* grados de libertad

- Ljung-Box, $$Q_{LB} = T(T+2)\sum_{j=1}^{k} \frac{r_j^2}{T-j}$$

  se distribuye *chi-cuadrado* con *k* grados de libertad

Características de un proceso estacionario

- Tiene varianza finita
- Sus funciones de autocorrelación simple y parcial tienden a cero rápidamente.
- Su gráfico temporal cruza su valor medio muestral frecuentemente.
- Una perturbación transitoria sobre su nivel tiene efectos puramente transitorios.
- Su previsión converge a su esperanza matemática al alargarse el horizonte de predicción.

Características de un proceso no estacionario

- Su varianza generalmente aumenta al transcurrir el tiempo, tendiendo a infinito.
- Su función de autocorrelación simple no tiende a cero rápidamente.
- El número esperado de períodos entre dos cruces consecutivos con su nivel medio es infinito.Su gráfico *deambula* durante períodos largos de tiempo.
- Una perturbación transitoria sobre su nivel tiene efectos permanentes
- Su previsión no converge a su esperanza matemática al alargarse el horizonte de predicción.

Modelos ARMA

Especificación, predicción, estimación y diagnosis

Modelos ARMA

Especificación

Modelos ARMA: especificación

- En un modelo univariante, utilizamos una única ecuación para representar la evolución dinámica de una variable económica.

- En cada instante de tiempo, una variable se ve afectada por su innovación. El valor numérico de dicha innovación es imprevisible en base a información pasada.

- Los factores explicativos del valor de la variables son: a) sus propios valores previos, b) los valores actual y pasados de su innovación

El proceso autorregresivo de orden 1, AR(1)

Proceso AR(1): $y_t = \delta + \phi y_{t-1} + \varepsilon_t$

donde ϕ, δ son constantes, y ε_t es un ruido blanco.
Si un proceso AR(1) es estacionario, su esperanza y varianza son,

$$Ey_t = \frac{\delta}{1-\phi}, \quad Var(y_t) = \frac{\sigma_\varepsilon^2}{1-\phi^2}$$

Un proceso AR(1)

- Puede escribirse en función de los valores de la innovación actual e innovaciones pasadas si el parámetro autoregresivo ϕ tiene valor inferior a la unidad

- Si el parámetro autoregresivo ϕ tiene valor superior a la unidad, admite una representación con varianza finita en función de las innovaciones futuras. En Economía, esta representación no es útil.

- Un *camino aleatorio* es una caso particular de proceso AR(1), no estacionario.

Proceso AR(1): función autocorrelación simple

$$\gamma_t^0 = Var(y_t) = \frac{\sigma_\varepsilon^2}{1-\phi^2}$$

$$\gamma_t^k = \phi\gamma_t^{k-1} \Rightarrow \gamma_t^k = \phi^k\gamma_t^0 \Rightarrow \rho^k = \phi^k$$

que converge hacia cero según las potencias de ϕ, si $-1 < \phi < 1$. La convergencia será monótona o alternando en signo, según que ϕ sea positivo o negativo.

Proceso AR(1): función autocorrelación parcial

- Formada por los sucesivos coeficientes de los retardos más largos en regresiones del tipo,

$$\tilde{y}_t = \phi_{k1}\tilde{y}_{t-1} + \phi_{k2}\tilde{y}_{t-2} + ... + \phi_{kk}\tilde{y}_{t-k}$$

donde $\tilde{y}_t$ denota la variable en diferencias respecto a su media muestral,

- Luego la f.a.p. de un proceso AR(1) tiene un primer valor igual a ϕ (positivo o negativo), y todos los demás iguales a cero.

Proceso autorregresivo de orden 2: AR(2)

Proceso AR(2): $y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$

con esperanza matemática y varianza:

$$Ey_t = \frac{\delta}{1-\phi_1-\phi_2},$$

$$Var(y_t) = \gamma_0 = \frac{(1-\phi_2)\sigma_\varepsilon^2}{(1+\phi_2)\left[(1-\phi_2)^2-\phi_1^2\right]}$$

- puede generar ciclos si su parámetro ϕ_2 es negativo. Cuando lo hace, el período *T* de dicho ciclo viene dado por:

$$\cos\frac{2\pi}{T} = \frac{\phi_1}{2\sqrt{-\phi_2}}$$

  y factor de amortiguamiento, $d = \sqrt{-\phi_2}$

Proceso AR(2): funciones de autocorrelación

- Los sucesivos valores de la ***fas*** satisfacen,

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

  que se conoce como sistema de ecuaciones de Yule-Walker. Todos los modelos AR(p) satisfacen un sistema análogo.

- Por lo que pueden tener un comportamiento oscilatorio (aunque no necesariamente lo tienen)

- La función de autocorrelación parcial tiene únicamente sus dos primeros valores no nulos, siendo iguales a los coeficientes del modelo AR(2).

El modelo de medias móviles de orden 1: MA(1)

- El valor contemporáneo de la variable depende de la innovación actual y la del período previo, $y_t = \delta + \varepsilon_t - \theta \varepsilon_{t-1}$, $t \geq 1$, a partir de un ε_0 dado, y siendo ε_t un proceso de *ruido blanco*.

- Su esperanza y varianza son, $E(y_t) = \delta; \quad Var(y_t) = \sigma_\varepsilon^2 \left(1 + \theta^2\right);$

- La función de autocorrelación simple del proceso MA(1) tiene tan sólo su primer valor no nulo, debido a la presencia común en y_t y en y_{t-1} del término ε_{t-1}:

$$\rho_1 = -\frac{\theta}{1 + \theta^2}$$

  que no excede de 0,50 en valor absoluto.
- Sin embargo, entre y_t e y_{t-2} no hay ningún término común, por lo que el segundo valor de la f.a.s. es igual a cero, así como todos los sucesivos valores de dicha función.

Invertibilidad de modelos MA

- Mediante sustitución reiterada, un modelo MA(1) puede expresarse como un proceso AR de orden infinito, en el que los coeficientes no son sino las potencias del parámetro del modelo MA(1).

- Tal transformación sólo tiene sentido si dicho parámetro es inferior a 1 en valor absoluto. Decimos, en tal caso, que el modelo MA(1) es *invertible*.

- Por lo que la función de autocorrelación parcial de un modelo MA(1) invertible converge exponencialmente hacia cero, alternando en signo si θ es negativo.

El modelo de medias móviles de orden 2: MA(2)

- El valor contemporáneo de la variable depende de la innovación actual y la del período previo,

$$y_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}, t \geq 1$$

a partir de $\varepsilon_{0,}$, ε_{-1} dados, y siendo ε_t un proceso de *ruido blanco*.

- Su esperanza y varianza son,

$$E(y_t) = \delta; \quad Var(y_t) = \sigma_\varepsilon^2 \left(1 + \theta_1^{\ 2} + \theta_2^{\ 2}\right);$$

Modelos ARMA

- Se componen de una estructura AR junto con una estructura MA.
- Heredan las propiedades de una y otra. Esto implica que sus funciones de autocorrelación tienen ambas unas características relativamente complejas, dificultando la identificación de estos procesos.
- Por ejemplo, en un modelo ARMA(1,1), compuesto de una estructura AR(1) y otra MA(1), ambas funciones decaen exponencialmente hacia cero, pudiendo hacerlo alternando en signo.
- En muchos casos se llega a estos modelos en varios pasos, tras comprobar que los residuos de un modelo AR(2) tienen estructura MA(1); por ejemplo.

Estacionariedad mediante transformaciones funcionales

- La transformación logarítmica pretende reducir la cantidad de heteroscedasticidad. En muchos casos, reduce la evidencia de exceso de curtosis. Por eso es que suponemos que la variable original es log-normal.

- Una transformación más general es la familia de funciones Box-Cox, de la cual la logarítmica es un caso particular. El parámetro λ que caracteriza cuál de las transformaciones de la familia es la más adecuada, puede estimarse por procedimientos habituales.

Estacionariedad mediante diferencias

- En muchas ocasiones, las diferencias de orden 1 ó 2 de una variable no estacionaria es una variable estacionaria.
- Recordemos que la primera diferencia del logaritmo de una variable es su tasa de variación logarítmica, aproximadamente igual a la tasa de variación porcentual (para tasas pequeñas).
- En Finanzas, la primera diferencia del logaritmo de un precio es una rentabilidad. Los precios son, generalmente no estacionarios; las rentabilidades asociadas pueden ser estacionarias. Los tipos de interés, no son estacionarios
- Por eso se busca un modelo ARMA en dicha(s) diferencia(s), constituyendo los denominados modelos ARIMA (Integrated Autorregressive Moving Average models)

Predicción

Predicción estática y predicción dinámica

- *Predicción dinámica* es la efectuada en un determinado instante de tiempo, sobre cada uno de los períodos del horizonte de predicción.
  - Para el primer período, nos basamos en los últimos datos observados.
  - Para el segundo y sucesivos períodos, habremos de utilizar las previsiones de períodos intermedios, previamente calculadas.

- *Predicción estática* es la formada por un conjunto de predicciones, cada una de ellas efectuada un período hacia delante. Para obtenerla hay que utilizar un conjunto de información distinto en cada período, para lo que haya que disponer de la necesaria información muestral.

- La predicción intra-muestral puede ser dinámica o estática. La predicción extra-muestral ha de ser estática.

- La predicción estática proporciona siempre mejores resultados, pero no son predicciones comparables.

Proceso AR(1): Predicción

$$E_T y_{T+1} = \delta + \phi y_T$$

$$E_T y_{T+2} = \delta + \phi E_T y_{T+1} = \delta\left(1+\phi\right) + \phi^2 y_T$$

.....

$$E_T y_{T+k} = \phi^k y_T + \delta\left(1+\phi+\phi^2+\ldots+\right)\phi^k = \phi^k y_T + (1-\phi^k)Ey$$

que converge a la media del proceso según aumenta el horizonte de predicción

Proceso AR(1): Errores de predicción

$$E_T y_{T+1} = \delta + \phi y_T \Rightarrow \zeta_{T+1} = y_{T+1} - E_T y_{T+1} = \varepsilon_{T+1}$$

$$E_T y_{T+2} = \delta + \phi E_T y_{T+1} \Rightarrow \zeta_{T+2} = y_{T+2} - E_T y_{T+2} = \varepsilon_{T+2} + \phi\varepsilon_{T+1}$$

.....

$$E_T y_{T+k} = \phi^k y_T + (1-\phi^k) Ey \Rightarrow$$

$$\Rightarrow \quad \zeta_{T+k} = y_{T+k} - E_T y_{T+k} = \varepsilon_{T+k} + \phi\varepsilon_{T+k-1} + \phi\ \varepsilon_{T+k-2} + \ldots + \phi^{k-1}\varepsilon_{T+1}$$

por lo que la varianza del error de predicción y, con ella, la amplitud de las bandas de confianza de la predicción, va aumentando con el horizonte de predicción, pero permanece finita, convergiendo a la propia varianza del proceso.

Proceso AR(2): Predicción

$$E_T y_{T+1} = \delta + \phi_1 y_T + \phi_2 y_{T-1}$$

$$E_T y_{T+2} = \delta + \phi_1 E_T y_{T+1} + \phi_2 y_T =$$

$$= \delta\left(1+\phi_1\right) + \left(\phi_1^{\,2} + \phi_2\right) y_T + \phi_1 \phi_2 y_{T-1}$$

$$E_T y_{T+3} = \delta + \phi_1 E_T y_{T+2} + \phi_2 E_T y_{T+1} =$$

$$= \delta\left(1+\phi_1+\phi_1^{\,2}+\phi_2\right) + \phi_1\left(\phi_1^{\,2}+2\phi_2\right) y_T + \left(\phi_1^{\,2}\phi_2 + \phi_2^{\,2}\right) y_{T-1}$$

La varianza del error de predicción de un proceso AR(2) tiene las mismas propiedades que las de un proceso AR(1), siempre que sea estacionario, es decir, que las dos raíces de su ecuación característica estén fuera del círculo unidad. En ese caso, el tamaño del error de predicción aumenta, pero converge a la varianza del proceso AR(2).

Proceso MA(1): Predicción

$$y_t = \varepsilon_t - \theta\varepsilon_{t-1} \Rightarrow y_{T+1} = \varepsilon_{T+1} - \theta\varepsilon_T$$

$$E_T y_{T+1} = E_T \varepsilon_{T+1} - \theta\varepsilon_T = -\theta\varepsilon_T$$

$$E_T y_{T+2} = E_T \varepsilon_{T+2} - \theta E_T \varepsilon_{T+1} = 0$$

....

$$k \geq 2 \Rightarrow E_T y_{T+k} = E_T \varepsilon_{T+k} - \theta E_T \varepsilon_{T+k-1} = 0$$

por lo que la varianza del error de predicción es:

$$Var(\xi_T^1) = \sigma_\varepsilon^{\ 2}; \quad Var(\xi_T^k) = \sigma_\varepsilon^{\ 2}\left(1+\theta^2\right) \; para \; k \geq 2$$

Predicción en modelo ARMA(1,1)

- Este proceso comparte las propiedades de sus componentes, que se superponen unas a otras. La varianza del error de predicción en el modelo ARMA(1,1) aumenta con el horizonte de predicción pero, al alargarse dicho horizonte, converge a un límite finito, la propia varianza del proceso.

Predicción en modelos ARMA: características generales

- El error de predicción un período hacia delante es, en todos los modelos, la propia innovación del proceso, ε_t. Ello no significa que el error de predicción sea independiente del modelo utilizado: sólo uno de ellos (el verdadero modelo) tiene la innovación como error de predicción un período hacia delante.

- Un intervalo de confianza alrededor del valor numérico de la predicción obtenida a un determinado horizonte puede obtenerse bajo el supuesto de Normalidad de la innovación. En tal caso, se lleva a la izquierda y derecha de dicha predicción puntual la desviación típica de la misma, que es función del horizonte de predicción.

Predicción de una variable en diferencias

- En muchas ocasiones se modeliza una transformación de la variable original, mediante una o varias diferencias. Cuando se obtienen predicciones para dicha variable diferenciada, hay que recuperar las predicciones correspondientes para la variable original.

- Para ello, basta añadir a los últimos datos de la variable original, las predicciones de su versión diferenciada. El modo exacto de hacerlo depende de las diferencias que se hayan aplicado a la variable original.

Predicción de una variable a partir de su logaritmo

- En muchos casos, habremos tomado el logaritmo de la variable original, posiblemente antes de tomar una o más diferencias de la misma. La predicción de la variable original no se obtiene elevando el número **e** (base del logaritmo nepperiano) a la predicción de la variable transformada.

- Es preciso hacer una corrección por la varianza de la serie: a la predicción del logaritmo se suma 0,5 veces la varianza del error de predicción. Esto se hace para cada horizonte de predicción.

- Esta práctica se basa en el supuesto de que la variable original obedece a una distribución de probabilidad log-normal, por lo que su transformada logaritmica es Normal.

Modelos ARIMA

Estimación

Estimación de modelos autorregresivos

- Es un modelo de regresión con variables explicativas aleatorias, aunque predeterminadas. El estimador MCO tiene buenas propiedades (es consistente) siempre que las variables explicativas y_{t-i} estén incorrelacionadas con el término de error. Esto es equivalente a que el término de error carezca de autocorrelación.

- La razón es que un proceso estacionario admite una representación con varianza finita en función de las innovaciones actual y pasadas.

- La presencia de autocorrelación en el residuo es el principal indicio de mala especificación en un modelo ARIMA.

- La consistencia no precisa Normalidad del término de error.

Estimación de modelos de medias móviles

- Se lleva a cabo por Máxima Verosimilitud, por lo que descansa en un supuesto acerca del tipo de distribución del término de error.

- Bajo Normalidad, maximizar la verosimilitud equivale a minimizar la Suma de Cuadrados de los Residuos (SCR), y la estimación de la varianza del error se obtiene dividiendo la SCR resultante por el número de *gdl*: *T-p-q*.

- En caso de un reducido número de parámetros, puede llevarse a cabo un procedimiento de búsqueda, pero es bastante ineficiente

Estimación de modelos de medias móviles

- Como alternativa, se escribe el modelo para el término de error contemporáneo, como función de a) la variable en *t* y períodos anteriores, b) valores previos del término de error,

- Por ejemplo, en el caso de una proceso MA(2),

$$\varepsilon_t = y_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

Pueden utilizarse pre-estimaciones de los 2 parámetros para generar una serie temporal de ε_t partiendo de condiciones iniciales iguales a cero, y se aproxima linealmente en serie de Taylor alrededor de dichas pre-estimaciones, obteniendo:

Estimación de MA(2)

$$\varepsilon_t = \varepsilon_t^{\,0} + (\theta_1 - \theta_1^{\,0})\left(\frac{\partial \varepsilon_t}{\partial \theta_1}\right)_{\theta_1=\theta_1^{\,0}} + (\theta_2 - \theta_2^{\,0})\left(\frac{\partial \varepsilon_t}{\partial \theta_2}\right)_{\theta_2=\theta_2^{\,0}}$$

$$\Rightarrow \quad \varepsilon_t^{\,0} - \theta_1^{\,0}\varepsilon_{t-1}^{\,0} - \theta_2^{\,0}\varepsilon_{t-2}^{\,0} = -\theta_1\varepsilon_{t-1}^{\,0} - \theta_2\varepsilon_{t-2}^{\,0} + \varepsilon_t$$

$$\Rightarrow$$

$$w_t = \theta_1 x_{1t} + \theta_2 x_{2t} + \varepsilon_t$$

$$w_t = \varepsilon_t^{\,0} - \theta_1^{\,0}\varepsilon_{t-1}^{\,0} - \theta_2^{\,0}\varepsilon_{t-2}^{\,0};\; x_{1t} = -\varepsilon_{t-1}^{\,0};\; x_{2t} = -\varepsilon_{t-2}^{\,0};$$

Lo que conduce a un procedimiento iterativo que utiliza la estimación por MCO de una regresión auxiliar

Estimación de modelos MA(2)

- En cada iteración, dados los valores en ese instante de los dos parámetros, podemos generar una series de *residuos,*

$$\varepsilon_t = y_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

y estimar, utilizando esta serie y los datos disponibles, la aproximación lineal del modelo MA(2). Ello nos permite pasar de las actuales estimaciones, a unas nuevas.

El procedimiento termina cuando las variaciones en los valores estimados de los parámetros son despreciables.

Estimación de modelos ARMA

- Un procedimiento similar puede utilizarse para modelos que combinan estructura autorregresiva con estructura de medias móviles.
- En este caso, puede procederse por bloques, estimando primero los parámetros AR y, condicionales en ellos, los parámetros MA, iterando sucesivamente. Sin embargo, la eficiencia estadística requiere estimación simultánea.
- Pueden tomarse estimaciones iniciales de los parámetros de la estructura de media móvil a partir de los valores estimados de la función de autocorrelación simple del proceso.

Modelos ARMA

Diagnóstico del modelo

Modelos ARMA: diagnóstico

- Los parámetros estimados deben satisfacer las condiciones de estacionariedad e invertibilidad

- Obtener las raíces de los polinomios de retardos autoregresivos y de medias móviles. Comprobar la posible existencia de ciclos y su período.

- Una buena especificación requiere que los residuos del modelo no contengan evidencia de autocorrelación (son ruido blanco)

- Las estructuras ARMA de autocorrelación son aditivas: una estructura MA(1) en los residuos de un modelo AR(2) sugiere que una mejor especificación es un modelo ARMA(2,1). Una estructura AR(1) en un modelo AR(1) sugiere que un modelo AR(2) es preferible

Sobreparametrización y sobrediferenciación

- Es conveniente factorizar los polinomios de retardos estimados:
  - un factor *MA* (1-.95L) en un modelo en diferencias $\Delta y_t = y_t - y_{t-1} = \delta + \varepsilon_t - \theta\varepsilon_{t-1}$ puede sugerir una situación de sobrediferenciación
  - un factor *AR* (1-.95L) puede sugerir la necesidad de una diferencia adicional: $y_t - \phi y_{t-1} = \delta + \varepsilon_t$
  - En ocasiones puede aparecer un factor (aproximadamente) común en las estructuras *AR* y *MA:* $y_t - \phi y_{t-1} = \delta + \varepsilon_t - \phi\varepsilon_{t-1}$
- ¿Es mejor quedarse corto de diferencias, o tender a sobrediferenciar?

Sobreparametrización y sobrediferenciación

- ¿Es mejor quedarse corto de diferencias, o tender a sobrediferenciar?
  - La diferenciación puede eliminar información contenida en las fluctuaciones de corto plazo
  - Pero garantizar la estacionariedad de los residuos es crucial
- Estrategia: tomar una diferencia adicional e incorporar un factor *MA* compensatorio adicional. Si resulta claramente no invertible, retornar al modelo con una diferencia menos.
- Esta estrategia suele mejorar los resultados predictivos

Estacionalidad

Estacionalidad

- En datos económicos trimestrales o mensuales, es muy frecuente que las observaciones correspondientes a un mismo período (mes o trimestre) de cada año, estén correlacionados entre sí.
- De hecho, lo mismo puede observarse al trabajar con datos de cierre diario de mercados financieros.
- En ocasiones, incluso con datos financieros horarios o más frecuentes.
- Para variables consideradas relevantes (Contabilidad nacional, IPC, etc...) los organismos estatales construyen versiones desestacionalizadas.

Estacionalidad

- Las citadas autocorrelaciones de carácter estacional pueden representarse mediante los mismos modelos ARMA que hemos visto para datos consecutivos.

- Así, puede hablarse de una estructura AR(1) en las frecuencias mensuales, que relacionará el valor actual de una variable al retardo de orden 12. El modelo AR(12) añadiría el retardo 24.

- En una serie trimestral, relacionaríamos el valor contemporáneo a los valores de hace 4, 8,... períodos

Estacionalidad

- Asimismo, pueden examinarse los valores en los retardos múltiplos de 4 (con datos trimestrales) o 12 (con datos mensuales) para evaluar la posible ausencia de estacionariedad estacional.

- Cuando existe apariencia de no estacionariedad, es aconsejable tomar diferencias de orden estacional, restando del valor contemporáneo de la variable, el valor de 12 meses antes (o 4 trimestres antes).

- Tras tomar las diferencias precisas, tanto de orden regular como de orden estacional, los términos AR regular y estacional se aplican multiplicativamente. Lo mismo para los términos MA de ambos tipos.

Análisis de intervención

Valores influyentes y anomalías: análisis de intervención

- En la serie transformada mediante diferencias, valores anómalos son los que tienen una magnitud muy superior a la desviación típica de la variable transformada

- Bajo supuestos de Normalidad, no sería razonable tener muchos más de un 5% de valores superiores a 2 veces la desviación típica, un 10% de ellos superiores a 1,65% la desviación típica, etc. Pero tampoco sería razonable observar valores alejados de la media en más de 8 desviaciones típicas.

- Valores anómalos pueden sugerir (de manera espúrea) la necesidad de diferenciar la serie.

- Pueden resultar influyentes en la estimación de los parámetros del modelo.

- O sesgar las predicciones, si están próximos al final de la serie temporal de datos.

Valores influyentes y anomalías: análisis de intervención

- Pueden tratarse indistintamente del resto de los valores de la serie
- O puede tratarse de estimar su influencia, mediante el análisis de intervención
- Este último debe hacerse sólo si el impacto numérico de tales datos es significativo
- O si se dispone de información acerca de su naturaleza: condiciones meteorológicas, intervenciones de política monetaria, efectos S.Santa, entrada en la CEE, etc...
- El *análisis de intervención* se lleva a cabo mediante variables impulso (efectos transitorios), y variables escalón (efectos permanentes)
- En ambos casos, pueden venir afectados de un polinomio de retardos

Valores influyentes y anomalías: análisis de intervención

- Una variable tipo *impulso* toma un valor igual a uno en un determinado período, siendo igual a cero en el resto.

- Una variable tipo *escalón* toma valor igual a uno en un determinado intervalo de tiempo, siendo igual a cero el resto.

- El intervalo de tiempo que define un escalón puede ser caracterizado por: “antes de determinado instante”, “después de”, o “entre dos instantes de tiempo”

- Una variable impulso puede estar definida para un conjunto de periodos aislados de tiempo: “los meses de julio”, “los meses en que cae la S.Santa”, “meses con alta/normal/baja precipitación pluviosa”, etc..

- En ocasiones, en presencia de un elevado número de intervenciones tipo impulso, se impone la igualdad de coeficientes entre un determinado grupo, para reducir el número de parámetros estimados.

Valores influyentes y anomalías: análisis de intervención

- Un polinomio de retardos asociado a una intervención tipo <u>impulso</u> genera una respuesta que se extiende a lo largo de varios períodos de tiempo:

$$\beta \frac{\xi_t^i}{1 - \lambda B}$$

genera una respuesta que cae en el tiempo a una tasa λ, comenzando a partir del coeficiente estimado

$$\beta \left(1 - \lambda_1 B - \lambda_2 B^2\right) \xi_t^i$$

genera una respuesta igual al coeficiente estimado β, seguido de respuestas $\beta\lambda_1$, $\beta\lambda_2$

Valores influyentes y anomalías: análisis de intervención

- Un polinomio de retardos asociado a una intervención tipo *escalón* genera una respuesta que se extiende a lo largo de varios períodos de tiempo:

$$\beta\left(1-\lambda_1 B-\lambda_2 B^2\right)\xi_t^s$$

genera un salto inicial en la media de magnitud igual al coeficiente estimado β, seguido de saltos en la media igual a $\beta\lambda_1$, $\beta\lambda_2$. En realidad, son impulsos seguidos de escalón

$$\beta\frac{\xi_t^s}{1-\lambda B}$$

genera una sucesión de saltos en media, no muy fácilmente justificable