

Capítulo 14

Modelos No lineales

Howell Tong (

Estadístico chino. Estudio en Inglaterra y es actualmente Professor de Estadística en la London School of Economics y en la Universidad de Hong-Kong. Inventor de los procesos autoregresivos por umbrales, ha hecho muchas contribuciones importantes a las series temporales no lineales.

14.1 Introducción

Tendremos un modelo no lineal si la relación entre una variable y su pasado es no lineal. Por ejemplo, la relación entre las ventas de hoy y las de ayer puede ser distinta cuando las ventas de ayer fueron altas que cuando las ventas fueron bajas. En este caso la función que relaciona la media en el instante t con el pasado será una función no lineal. Podríamos expresar esta situación mediante

$$z_t = m(\mathbf{Z}_{t-1}) + a_t$$

donde $\mathbf{Z}_{t-1} = (z_{t-1}, \dots, z_1)$ es el vector de valores pasados de la serie y a_t un proceso de ruido blanco. En este modelo la esperanza condicionada es

$$E(z_t | \mathbf{Z}_{t-1}) = m(\mathbf{Z}_{t-1}),$$

ya que la innovación tiene esperanza nula. La varianza condicionada es constante, ya que

$$\text{Var}(z_t | \mathbf{Z}_{t-1}) = \sigma^2$$

y a_t representa como en un modelo lineal el error de predicción a un paso. Por ejemplo, un modelo AR no lineal es

$$z_t = \alpha z_{t-1}^2 + \beta z_{t-1} z_{t-2} + a_t$$

Además de en la media una serie puede tener estructura no lineal en la varianza. Por ejemplo, con datos financieros y climatológicos se observa con frecuencia que la varianza de la serie en el instante t depende de la variabilidad de los datos anteriores, de manera que hay rachas de alta varianza seguidas de datos con poca varianza. Una forma general de representar procesos estacionarios no lineales es añadir esta dependencia mediante la formulación

$$z_t = m(\mathbf{Z}_{t-1}) + u_t v(\mathbf{Z}_{t-1}), \quad (14.1)$$

donde u_t es un proceso de ruido blanco de variables $N(0, 1)$ independiente de los valores \mathbf{Z}_{t-1} , y las funciones m y v son no lineales en los valores pasados de la serie. Si tomamos esperanzas en esta expresión condicionadas a \mathbf{Z}_{t-1} , tenemos que, como antes

$$E(z_t | \mathbf{Z}_{t-1}) = m(\mathbf{Z}_{t-1}),$$

ya que:

$$E(u_t v(\mathbf{Z}_{t-1}) | \mathbf{Z}_{t-1}) = E(u_t) v(\mathbf{Z}_{t-1}) = 0,$$

y si la función $m(\mathbf{Z}_{t-1})$ es no lineal, la función de medias condicionadas del proceso, $E(z_t|\mathbf{Z}_{t-1})$, también lo será. Por otro lado, la varianza condicionada de este proceso es

$$\text{Var}(z_t|\mathbf{Z}_{t-1}) = E[(z_t - m(\mathbf{Z}_{t-1}))^2|\mathbf{Z}_{t-1}] = v^2(\mathbf{Z}_{t-1})E(u_t^2) = v^2(\mathbf{Z}_{t-1})$$

y si la función $v^2(\mathbf{Z}_{t-1})$ es no lineal, la varianza condicional de este proceso será, en general, una función no lineal de los valores pasados. El modelo con varianza constante se obtiene suponiendo $v^2(\mathbf{Z}_{t-1}) = \sigma^2$.

En este capítulo nos centraremos en los procesos no lineales en la media y dejaremos para el siguiente los procesos no lineales en la varianza. Estos dos tipos de efectos pueden o no darse simultáneamente en una serie temporal.

La forma general de un proceso no lineal en la media se obtiene mediante la llamada representación de Volterra. Esta es una representación similar a la de los procesos lineales mediante una expansión MA(∞) :

$$\begin{aligned} z_t = & \mu + \sum_{i=-\infty}^{\infty} a_i \varepsilon_{t-i} + \sum_{i,j=-\infty}^{\infty} a_{ij} \varepsilon_{t-i} \varepsilon_{t-j} + \\ & \sum_{i,j,k=-\infty}^{\infty} a_{ijk} \varepsilon_{t-i} \varepsilon_{t-j} \varepsilon_{t-k} + \dots \end{aligned} \quad (14.2)$$

Esta expresión tiene un número infinito de coeficientes y, como la representación de Wold no es operativa. De la misma forma que los procesos ARIMA tratan de aproximar la representación de Wold se han desarrollado muchos modelos para aproximar la representación general anterior. En este capítulo veremos algunas aproximaciones simples a esta formulación.

14.2 Los procesos bilineales

Los modelos habituales de procesos no lineales pueden considerarse como aproximaciones a la expansión general 14.2. Un primer enfoque es suponer que podemos expresar la dependencia en función de un número finito de retardos y escribir el modelo ARMA no lineal:

$$z_t = f(z_{t-1}, \dots, z_{t-p}; a_{t-1}, \dots, a_{t-q}) + a_t \quad (14.3)$$

donde f es arbitraria y la innovación a_t es una secuencia de variables independientes e idénticamente distribuidas (iid) independientes de z_{t-k} para cualquier $k > 0$. Si $q = 0$ tenemos un AR no lineal y si $p = 0$ un MA no lineal. Suponiendo que f es una función suave con derivadas continuas, podemos desarrollarla en serie y, despreciando términos de orden superior, aproximarla por:

$$z_t = c + \sum_{i=1}^p \phi_i z_{t-i} + \sum_{j=1}^q \theta_j a_{t-j} + \sum_{i=1}^m \sum_{j=1}^h \gamma_{ij} z_{t-i} a_{t-j} + a_t \quad (14.4)$$

que conduce a la clase de procesos *bilineales*. Estos modelos han sido estudiados por Granger y Anderson (1978). Puede demostrarse que estos procesos pueden aproximar la expansión de Volterra en un intervalo finito cualquier precisión determinada. Estos procesos se representan como $BL(p, q, m, h)$, indicando el grado mayor de cada uno de los términos.

El proceso bilineal más simple es el BL(1,0,1,1) dado por:

$$z_t = \phi z_{t-1} + \gamma z_{t-1} a_{t-1} + a_t \quad (14.5)$$

Según esta ecuación a una estructura AR(1) de dependencia estacionaria se añade un término $\gamma z_{t-1} a_{t-1}$ que depende del último valor observado y del último error de predicción. Supongamos que $\gamma \geq 0$. Entonces, cuando a_{t-1} es positivo el siguiente valor se genera con un coeficiente AR $\phi + \gamma a_{t-1}$, que puede ser mayor que uno cuando $a_{t-1} \geq (1 - \phi)/\gamma$. El comportamiento entonces será como un proceso explosivo. Sin embargo, cuando a_{t-1} sea negativo, el siguiente valor se genera con un coeficiente AR $\phi - \gamma |a_{t-1}|$, que puede ser cero o negativo. La estructura de dependencia cambia aparentemente de un punto a otro y si el valor de γ no es

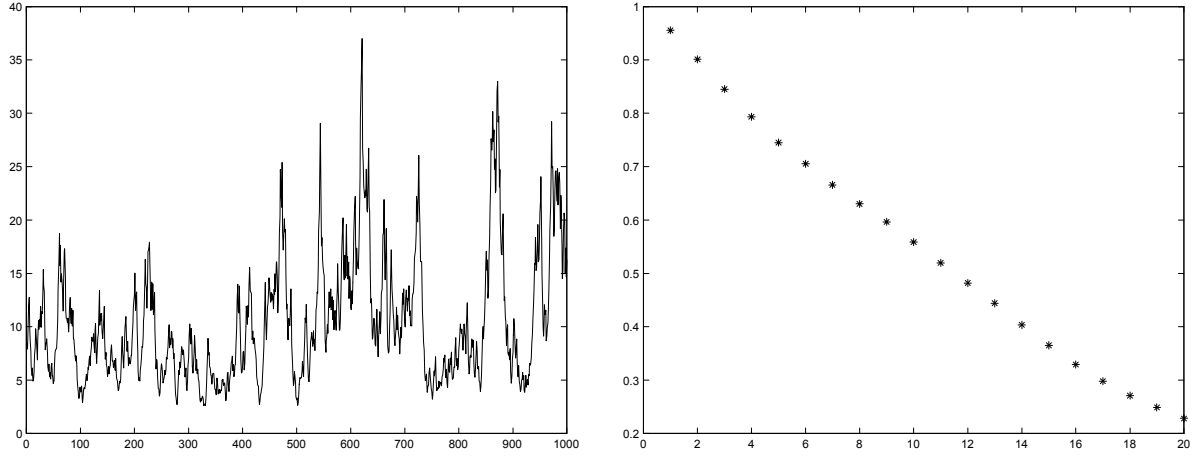


Tabla 14.1: Simulación del proceso bilineal $z(t) = .7z(t-1) + .5z(t-1)a(t-1) + a(t)$. Grafico del proceso y función de autocorrelación simple

pequeño la serie puede reflejar un comportamiento no estacionario. Este proceso puede también escribirse como:

$$\begin{aligned} z_t &= \phi_t z_{t-1} + a_t \\ \phi_t &= \phi + \gamma a_{t-1} \end{aligned}$$

que indica que se trata de un proceso AR(1) con parámetros cambiantes en cada instante y el cambio en los parámetros es proporcional al último error de predicción cometidos.

La figura 14.1 presenta una realización de este proceso y su función de autocorrelación simple. Se observa que el proceso produce valores extremos que se identificarían como atípicos en un ajuste lineal, y que tiene una estructura distinta de la lineal. Por otro lado la función de autocorrelación decae más lentamente que la de un proceso estacionario lineal. Si hacemos un histograma de los datos la distribución es muy asimétrica.

Si en lugar de modificar la estructura AR de primer orden mediante la innovación introducimos estructura de tipo AR(2) sobre una estructura lineal de primer orden, tendremos el proceso

$$z_t = \phi z_{t-1} + \beta a_{t-1} z_{t-2} \phi + a_t$$

que puede interpretarse como un proceso AR(2) con el parámetro de primer orden constante y el segundo variable en el tiempo, $\phi_2 = \beta a_{t-1}$. Como z_{t-2} y a_{t-1} están incorrelados y $E(a_t) = 0$, tomando esperanzas tenemos que $\mu = \phi\mu$, de donde concluimos que $E(z_t) = 0$. Las covarianzas serán, para $k \geq 2$

$$E[z_t z_{t-k}] = \phi E[z_{t-1} z_{t-k}] \quad (14.6)$$

y la varianza

$$\sigma_z^2 = \phi^2 \sigma_z^2 + \beta^2 \sigma_z^2 \sigma_a^2 + \sigma_a^2$$

de donde deducimos

$$\sigma_z^2 = \frac{\sigma_a^2}{(1 - \phi^2 - \beta^2 \sigma_a^2)}$$

que será positiva si

$$\phi^2 + \beta^2 \sigma_a^2 < 1$$

que es una condición para que el proceso sea estacionario.

Por la condición (14.6), este proceso puede confundirse con un AR(1) lineal. Sin embargo, si ajustamos un AR(1) lineal y calculamos sus residuos, estos vendrán dados por

$$e_t = (\phi - \hat{\phi}) z_{t-1} + \beta z_{t-2} a_{t-1} + a_t$$

y si $(\phi - \hat{\phi})$ es pequeño estos residuos estarán aproximadamente incorrelados. Sin embargo, es fácil comprobar que los cuadrados de los residuos si estarán correlados, lo que no ocurrirá si el proceso fuese lineal.

14.3 Los procesos autorregresivos por umbrales

Otra forma de introducir no linealidad es suponer que podemos aproximar la función no lineal que relaciona los valores presentes con los pasados por funciones lineales por tramos. Por ejemplo en el modelo AR(1) estacionario no lineal general :

$$z_t = f(z_{t-1}) + a_t$$

podemos aproximar la función $f(z_{t-1})$ por tramos lineales. Si hacemos una aproximación con dos términos un posible modelo es

$$\begin{aligned} z_t &= c^{(1)} + \phi^{(1)} z_{t-1} + a_t^{(1)}, & z_{t-1} > d \\ z_t &= c^{(2)} + \phi^{(2)} z_{t-1} + a_t^{(2)}, & z_{t-1} \leq d \end{aligned} \quad (14.7)$$

donde $a_t^{(i)}$ para $i = 1, 2$ es ruido blanco con distribución $N(0, \sigma_i^2)$. Esta formulación equivale a suponer que un modelo lineal distinto opera para distintos valores de la variable z_{t-1} . Estos modelos son debidos a Tong (1983) y se denominan modelos autorregresivos por umbrales o Threshold AutoRegresive (TAR) models. La variable z_{t-1} se denomina variable de umbral y el valor d es el umbral del modelo.

Las dos ecuaciones del modelo pueden escribirse conjuntamente introduciendo una variable auxiliar, $\alpha_t = F(z_{t-1} - d)$ que toma el valor 1 cuando el argumento es positivo y cero en otro caso. Entonces el modelo (14.7) puede escribirse:

$$z_t = [\alpha_t c^{(1)} + (1 - \alpha_t) c^{(2)}] + [\alpha_t \phi^{(1)} + (1 - \alpha_t) \phi^{(2)}] z_{t-1} + [\alpha_t \sigma_1 + (1 - \alpha_t) \sigma_1] \epsilon_t$$

donde ϵ_t es $N(0, 1)$.

La figura 14.2 presenta la relación entre z_t y z_{t-1} en el modelo con $c^{(1)} = c^{(2)} = d = 10$, $\phi^{(2)} = .8$. Los valores de $\phi^{(1)}$, σ_1, σ_2 en los cuatro casos son $(.2, 1, 3)$, $(.2, 1, 1)$, $(-.8, 1, 3)$, $(-.8, 1, 1)$. En el caso (a) la relación es siempre positiva pero presenta menos varianza para valores altos de z_{t-1} (mayores de 10) que para valores bajos. La pendiente también es menor para valores altos. Estas propiedades son claramente visibles en la figura (a). En el caso (b) la varianza es la misma en ambos regímenes y lo que cambia es la pendiente de la relación. En los casos (c) y (d) la relación es positiva para valores bajos pero negativa para valores más altos, y en el (c) la varianza es distinta en ambos regímenes mientras que en el (d) es la misma.

Los gráficos de dispersión entre z_t y sus retardos son muy útiles para identificar la posible presencia de estos modelos.

Los modelos TAR en el caso general se escriben como:

$$z_t = c^{(r)} + \sum_{i=1}^p \phi_i^{(r)} z_{t-i} + a_t^{(r)}, \quad \mathbf{x}_{t-1} \in A_r, \quad r = 1, \dots, R$$

donde $\mathbf{x}_{t-1} = (z_{t-1}, \dots, z_{t-p})'$, es el vector de estado de la variable de umbral, y A_r la partición de los valores del vector de estado correspondiente a los R regímenes. Para cada posible estado r , tenemos distintos parámetros, tanto para los coeficientes del AR como para la varianza de las innovaciones, σ_r^2 . La identificación de estos modelos cuando la variable que determina los umbrales es vectorial es difícil y en la práctica se supone que el vector de estado es escalar, de manera que hay un único retardo que determina los umbrales y los conjuntos A_r son una partición de los valores de la variable de umbral. Tong(1983) ha propuesto un procedimiento iterativo para ajustar estos modelos basado en el criterio AIC como sigue:

Paso 1. A partir de los gráfico de dispersión bivariantes entre la variable y sus retardos, se selecciona la variable z_{t-j} más prometedora para ser la variable de umbral. Esta variable será aquella donde la relación entre z_t y z_{t-j} sea más no lineal. Una vez decidida esta variable se selecciona el valor del umbral o punto de corte, d , que separa el gráfico en dos ecuaciones lineales. Con esta información se divide la muestra en dos submuestras según el valor de la variable umbral y se estiman modelos AR en cada tramo para distintos valores de p . Se selecciona el mejor modelo en cada uno de los regímenes mediante el AIC. Sea $AIC(\hat{p}_1)$ el valor de este criterio para un modelo AR(p_1). Se define el AIC del modelo total como la suma de los AIC correspondientes a cada régimen:

$$AIC(j, d) = AIC(\hat{p}_1) + AIC(\hat{p}_2)$$

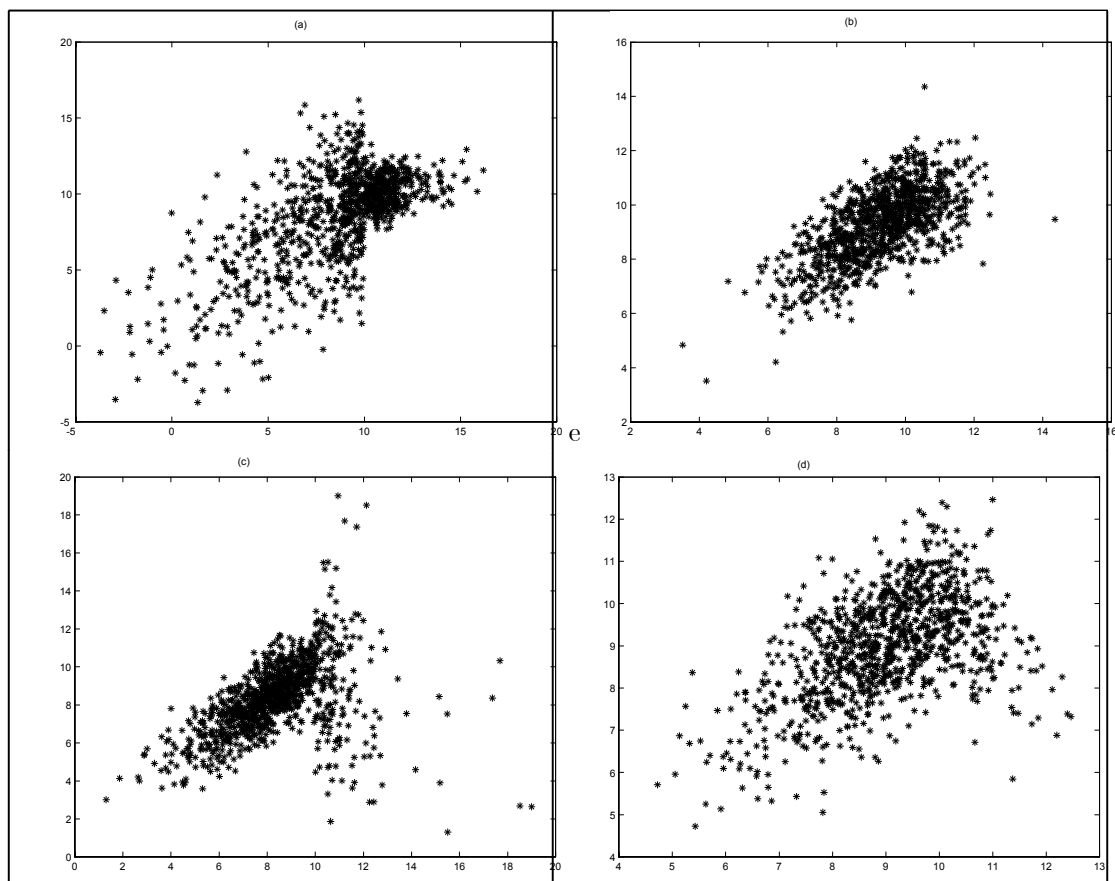


Tabla 14.2: Relacion entre la variable y su retardo para cuatro modelos TAR. En los cuatro casos $d=10$ y el primer coeficiente AR es $.8$. Los modelos difieren en las varianzas y en el segundo coeficiente AR que es $.2$ arriba y $-.8$ en las de abajo

donde \hat{p}_1 y \hat{p}_2 son los órdenes de los procesos AR ajustados a cada una de los regímenes.

Paso 2. Se considera ahora un conjunto posible de valores de d , $-\infty \leq d_1 \leq \dots \leq d_k \leq \infty$, dividiendo el rango de variación de z_{t-j} , la variable umbral, en tramos. Se repite el Paso 1 para cada valor de d_j , $j = 1, \dots, k$, y se selecciona con el criterio AIC el mejor modelos. Sea $AIC(\hat{d})$ el valor del criterio correspondiente al mínimo de los modelos estimados.

Paso 3. Se repiten los dos pasos anteriores probando con otra posible variable de umbral, z_{t-h} , para $h = 1, \dots, p$. Se comparan los AIC obtenidos para distintos valores de la variable umbral y se elige como modelo final el que tenga mínimo AIC.

Ejemplo 14.1

La figura presenta los residuos de un MA(2) ajustados a los datos del PIB de EEUU. Los datos se encuentran en el fichero pibusa.dat. El modelo ajustado al periodo 4/83-2/03 es

$$\nabla \log y_t = (1 + 0.28B + .32B^2)a_t$$

La figura muestra los residuos de este modelo con relación a su retardo. Aparece claramente un comportamiento por umbrales, ya que la relación es distinta por debajo y por encima de cero. Además, por debajo de cero se observan claramente dos regímenes distintos, que Tiao y Tsay (1994) identificaron como separados por el valor de y_{t-2} . De esta manera se obtiene un modelo TAR con cuatro umbrales según que las variables y_{t-1} y y_{t-2} sean positivas o negativas. Por debajo del cero el comportamiento es distinto que por encima. Véase Tiao y Tsay (1994) para un análisis de esta serie mediante un modelo por umbrales, y Peña y Rodriguez (2004) para un análisis de la linealidad de esta serie.

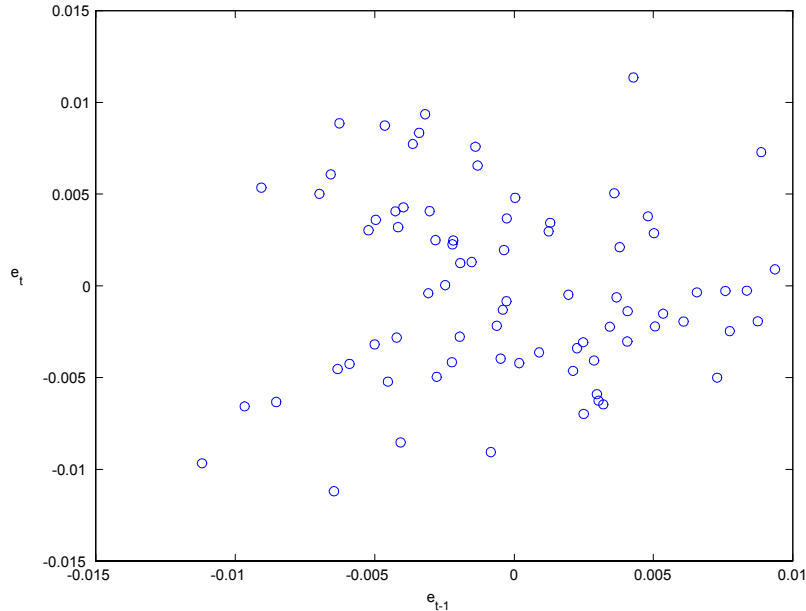


Figura 14.1: Grafico de los residuos de un ajuste lineal a la serie del pib de EEUU. Se observa que la dependencia por debajo del cero es distinta que por encima. Por debajo de cero se ven además dos regímenes distintos

14.4 Otros modelos no lineales

Los modelos TAR pueden generalizarse permitiendo una transición suave entre los regímenes. Se obtienen entonces los modelos STAR, o Smooth Threshold AutoRegresive. En el caso más simple de un AR(1) su representación es

$$z_t = F(z_{t-1})(c^{(1)} + \phi^{(1)}z_{t-1} + a_t^{(1)}) + (1 - F(z_{t-1}))(c^{(2)} + \phi^{(2)}z_{t-1} + a_t^{(2)})$$

donde $F(z_{t-1})$ es una función suave que toma valores en $[0,1]$. Por ejemplo una función que se utiliza mucho es la logística

$$F(z_{t-1}) = \frac{1}{1 + e^{-\alpha(z_{t-1}-d)}}$$

donde el parámetro α es positivo. Si $z_{t-1} - d$ tiende a menos infinito la función tiende a cero mientras que si $z_{t-1} - d$ tiende a más infinito la función tiende a uno. La transición entre estos dos valores depende del valor de α . Si este parámetro es muy grande se produce un cambio muy rápido al pasar $z_{t-1} - d$ de negativo a positivo y tenemos un caso muy similar al TAR. Por el contrario si α es pequeño la transición será suave. Los modelos TAR son casos particulares de estos modelos con

$$F(z_{t-1}) = \begin{cases} 0, & z_{t-1} - d \leq 0 \\ 1 & z_{t-1} - d > 0 \end{cases}$$

Finalmente, en lugar de utilizar modelos paramétricos se ha explorado la utilización de métodos no paramétricos para explicar estructuras de dependencia no lineales. Véase Peña, Tiao and Tsay (2001) para ejemplos de estos modelos.

14.5 Contrastes de no linealidad

14.5.1 Tipos de contrastes

La identificación de la no linealidad se realiza habitualmente sobre los residuos de un modelo lineal. Los tests existentes pueden clasificarse en tres grupos. El primer grupo parte de que aunque los residuos estén incorrelados por construcción con los valores retardados de la variable no son independientes de estos y tendrán una relación de dependencia general del tipo:

$$e_t = m(\mathbf{Z}_{t-1}), \quad (14.8)$$

donde $\mathbf{Z}_{t-1} = (z_{t-1}, \dots, z_1)$ es el vector de valores pasados de la serie. Podemos aproximar esta función mediante un desarrollo de Taylor y escribir

$$e_t = \sum \alpha_i z_{t-i}^2 + \sum \beta_{ij} z_{t-i} z_{t-j} + u_t$$

ya que en esta relación los términos lineales no estarán por la incorrelación entre los residuos y los retardos de la variable. Por tanto si realizamos esta regresión y encontramos coeficientes significativos esto sugiere no linealidad.

El segundo tipo de contraste parte de que los residuos de una serie no lineal a la que ajustamos un modelo lineal aunque estén incorrelados no serán independientes. Esto es consecuencia de la relación (14.8), que indica que e_t y e_{t-k} comparten información común y no son independientes. Si suponemos una relación general de dependencia general

$$e_t = g(\mathbf{E}_{t-1})$$

donde $\mathbf{E}_{t-1} = (e_{t-1}, \dots, e_1)$, podemos escribir

$$e_t = \sum \alpha_i e_{t-i}^2 + \sum \beta_{ij} e_{t-i} e_{t-j} + v_t$$

y ajustar esta regresión en busca de términos significativos. Alternativamente podemos investigar directamente la presencia de esta función viendo si cuando tenemos dos secuencias con valores similares de las variables retardadas los valores siguientes de las variables también son similares.

El tercer tipo de contraste supone que si existe una relación no lineal entre las variables existirá una transformación $h(e_t)$ donde esta relación aparezca más claramente y propone encontrar una transformación h para maximizar la dependencia

$$h(e_t) = \sum \alpha_i h(e_{t-i}) + \sum \beta_{ij} h(e_{t-i}) h(e_{t-j}) + \epsilon_t$$

La función h puede elegirse dentro de la familia de Box-Cox. Alternativamente podemos comenzar con una transformación simple, como los cuadrados, y comprobar si encontramos dependencia lineal entre los cuadrados de los residuos. En las secciones siguientes desarrollamos estos contrastes.

14.5.2 Contrastes de regresión entre los residuos y funciones de la variable

Un contraste de este tipo ha sido propuesto por Tsay (1986). La idea del contraste es construir ciertas variables explicativas \mathbf{x}_t que recojan los términos no lineales de z_t y sus retardos y relacionarlas con los residuos e_t . Si encontramos dependencia aceptamos la presencia de estructura no lineal.

(1) Estimar un modelo autorregresivo con un retardo M largo. Sean

$$\hat{z}_t = \hat{\alpha}_0 + \sum_{i=1}^M \hat{\alpha}_i z_{t-i} \quad (14.9)$$

las predicciones generadas por el modelo estimado. Los residuos e_t de este ajuste lineal se calculan de la forma habitual con:

$$e_t = z_t - \hat{z}_t = z_t - \hat{\alpha}_0 - \sum_{i=1}^M \hat{\alpha}_i z_{t-i} \quad (14.10)$$

(2) Hacer una regresión entre los residuos e_t y las variables $(z_{t-i}e_{t-i}, e_{t-i}e_{t-i-1})$ para $i = 1, \dots, M$. De esta manera, como según (14.10)

$$z_{t-i}e_{t-i} = z_{t-i}^2 - \hat{\alpha}_0 z_{t-i} - \sum_{j=1}^M \hat{\alpha}_j z_{t-i-j} z_{t-i}$$

y

$$e_{t-i}e_{t-i-1} = \hat{\alpha}_0^2 + z_{t-i}z_{t-i-1} - \hat{\alpha}_0 z_{t-i} - \hat{\alpha}_0 z_{t-i-1} + \text{dobles}$$

donde *dobles* incluye los dobles productos, incluimos los cuadrados de las variables y los productos cruzados como regresores. Como las variable introducidas para explicar los residuos incluyen términos lineales con los retardos de la variable y los residuos deben estar incorrelados con ellos, para contrarestar este efecto se introducen en la regresión entre los residuos y estas variables tambien los regresores $(1, z_{t-1}, \dots, z_{t-M})$. A continuación, se realiza el contraste F de regresión habitual y se rechaza la linealidad si este contraste es significativo. Véase Peña, Tiao and Tsay (2000) y Peña y Rodriguez (2004) para ejemplos y estudios de la potencia de este contraste.

14.5.3 Contrastes sobre la dependencia de su pasado de los residuos. Contraste BDS.

Un contraste que se basa en la relación no lineal de los residuos del ajuste lineal y su pasado es debido a Brock et al. (1991, 1996), y se conoce como contraste BDS. La idea del contraste es muy general: Si existe relación entre los residuos y su pasado, si tenemos dos bloques de datos consecutivos $(e_{t+1}, \dots, e_{t+k})$ y $(e_{t+1+h}, \dots, e_{t+k+h})$, y los dos bloques están próximos, esperamos que los valores futuros de ambas secuencias $(e_{t+k+1}, \dots, e_{t+k+g})$ y $(e_{t+k+h+1}, \dots, e_{t+k+h+g})$ también estén próximos. Para implementar esta idea necesitamos una medida de proximidad entre bloques de observaciones y en este test se toma la máxima distancia euclídea entre observaciones de ambos bloques que ocupan la misma posición. Es decir, dados dos bloques de observaciones de tamaño k ,

$$(e_{t+1}, \dots, e_{t+k}), \quad (e_{t+1+h}, \dots, e_{t+k+h})$$

se forman las distancias entre observaciones que ocupan la misma posición en el bloque, $d_j = |e_{t+j} - e_{t+h+j}|$ para $j = 1, \dots, k$. La distancia entre bloques se define como el máximo de estas distancias, y los bloques se consideran próximos si $\max(d_j) \leq c$.

El test se realiza entonces como sigue: formamos secuencias de k elementos $(e_{t+1}, \dots, e_{t+k})$ para $t = 0, \dots, n - k$, y para cada una de ellas contamos cuantas de las restantes secuencias están próximas a la considerada. Para resumir el resultado de la comparación entre dos secuencias de tamaño k , una comenzando en t y otra en s , se define una variable ficticia, $C_{t,s}$, que toma el valor uno si ambas secuencias están próximas y cero en otro caso. La proporción de secuencias de tamaño k que están proximas, se calcula mediante

$$C_{k,T} = \frac{2}{(T-k)(T-k-1)} \sum_{t=1}^{T-k} \sum_{s=t+1}^{T-k-1} C_{t,s}.$$

y el estadístico para el contraste BDS es el valor estandarizado de la proporción de secuencias próximas, dado por :

$$w_{k,T} = \sqrt{T-k-1} \frac{(C_{k,T} - C_{1,T-k+1}^k)}{\sigma_{k,T-k+1}}$$

Se demuestra que, bajo la hipótesis de independencia de los residuos, este estadístico sigue una distribución normal. Para utilizar este test es necesario definir k , el tamaño de la secuencia, y el nivel crítico para las distancias, c .

14.5.4 Contrastes sobre los cuadrados de los residuos

Si tenemos un modelo no lineal, los residuos del mejor ajuste lineal estarán incorrelados pero no serán independientes, ya que la información no modelizada quedará en los residuos. Si los residuos son independientes las autocorrelaciones de sus cuadrados deben ser similares a los cuadrados de las autocorrelaciones originales. En efecto, Maravall (1983) demostró que si las a_t son normales incorreladas y por tanto independientes:

$$\rho_k(a_t^2) = [\rho_k(a_t)]^2$$

En consecuencia, como $|\rho_k(a_t)| \leq 1$ si elevamos los residuos al cuadrado y calculamos sus autocorrelaciones, éstas –en la hipótesis de normalidad– no deberán aumentar con relación a las de la serie original.

McLeod and Li (1983) han propuesto un contraste de esta propiedad. Suponiendo que los residuos estimados son normales, las autocorrelaciones de sus cuadrados definidas por:

$$r_k(e_t^2) = \frac{\sum_{t=k+1}^T (\hat{a}_t^2 - \hat{\sigma}^2) (\hat{a}_{t-k}^2 - \hat{\sigma}^2)}{\sum_{t=1}^n (\hat{a}_t^2 - \hat{\sigma}^2)^2}, \quad (k = 1, 2, \dots, m),$$

donde $\hat{\sigma}^2 = \sum \hat{a}_t^2 / T$ serán asintóticamente variables normales incorreladas de media cero y varianza $1/T$. Por tanto el estadístico

$$Q_{LB}(\hat{a}_t^2) = T(T+2) \sum_{k=1}^m (T-k)^{-1} r_k^2(\hat{a}_t^2),$$

seguirá una distribución χ^2 con m grados de libertad. Observamos que en el contraste de los cuadrados de los residuos no hay que restar el número de parámetros estimados.

Un contraste de no linealidad más potente es has sido propuesto por Peña y Rodriguez (2003) que utiliza que el estadístico D_m dado por

$$D_m = -\frac{T}{m+1} \log |\mathbf{R}_m|,$$

donde \mathbf{R}_m es la matriz de autocorrelación de los residuos al cuadrado, dada por

$$\mathbf{R}_m = \begin{bmatrix} 1 & r_1 & \cdots & r_m \\ r_1 & 1 & \cdots & r_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_m & r_{m-1} & \cdots & 1 \end{bmatrix}. \quad (14.11)$$

En la hipótesis de linealidad este estadístico sigue una distribución gamma $\mathcal{G}(\alpha, \beta)$ con $\alpha = 3m(m+1)/4(2m+1)$ y $\beta = 3m/2(2m+1)$.

14.6 Referencias adicionales

Los libros de Priestly (1988) y Tong (1990) son referencias básicas para estos modelos. Granger y Anderson (1978) estudian los modelos bilineales y Tong (1983) los autorregresivos por umbrales.