Prácticas de la asignatura Series Temporales

Cuarta Entrega

1 Modelos de Series Temporales no estacionarias

1.1 Procesos Integrados

Los procesos no estacionarios más comunes son los procesos integrados. Un proceso x_t es integrado de orden d, $I(d)$, si tras diferenciarlo d veces obtenemos una serie estacionaria. Por supuesto las series estacionarias son $I(0)$. Una manera muy sencilla de comprobar si una serie es estacionaria o no es mediante el correlograma. Las correlaciones muestrales de este tipo de procesos tienen un decrecimiento muy lineal. Este tipo de procesos una vez diferenciados las veces suficientes son series estacionarias por lo que a continuación podemos pasar a identificar el resto de su estructura. En general un proceso ARIMA(p, d, q) viene dado por la ecuación:

$$\left(1-\phi_1 B-\cdots-\phi_p B^p\right)(1-B)^d x_t = c+\left(1-\theta_1 B-\ldots-\theta_q B^q\right) a_t \qquad a_t \sim N\left(0,\sigma^2\right) \text{, independientes}$$

De esta manera si definimos el proceso $y_t = (1-B)^d x_t$, lo que obtenemos es que el proceso y_t sigue un modelo ARMA(p, q). El proceso integrado más simple es el paseo aleatorio pero ya lo vimos anteriormente

1.1.1 Proceso de alisado exponencial simple - IMA(1,1)

El proceso IMA(1) tiene por ecuación generadora:

$$(1-B)\, x_t = c + a_t - \theta a_{t-1}$$

La ecuación nos dice que el valor de la variable en tiempo t es el valor de la variable en el tiempo anterior más el valor de la innovación en tiempo t, a_t, menos θ veces el valor de la innovación en tiempo $t-1$, a_{t-1}, más el valor de la constante. Es fácil ver que este proceso es un proceso ARMA(1,1) donde el parámetro autorregresivo es 1. Suponiendo que tenemos $x_0 = a_0 = 0$, se puede demostrar que:

$$Var\left[x_t\right] = \sigma_a^2 \left(1+(t-1)(1-\theta)^2\right)$$

$$\rho\left[t, t+k\right] = \frac{(1-\theta)\left(1+(t-1)(1-\theta)\right)}{\sqrt{\left(1+(t-1)(1-\theta)^2\right)\left(1+(t+k-1)(1-\theta)^2\right)}} \simeq \frac{1}{\sqrt{1+\frac{k}{t}}}.$$

Vamos a generar varias series para diferentes parámetros θ. Para ello creamos un workfile de 400 datos de la manera habitual. En primer lugar generamos una serie donde $c = 0$, $\theta = -0.5$ y $\sigma_a^2 = 1$. Para ello debemos seguir los pasos:

1. File→New→Workfile.

2. Undated or Irregular→1 to 400.
3. Genr→a=nrnd. (Esto nos define la serie de ruido blanco).
4. Genr→x=0. (Esto nos define la serie como 0).
5. Genr→IMPORTANTE: Sample: 2:400 → x=x(-1)+a+0.5*a(-1). (Aqui la serie toma sus valores definitivos).

El proceso de generación asume que $x_1 = 0$. Vemos la serie:

View → Line Graph

La serie debe tomar valores que no se muevan alrededor de una media común, es decir, tiene periodos de crecimiento y decrecimiento pronunciados. Vemos el correlograma. Comprobamos que la función de autocorrelación estimada de la serie decae muy lentamente y además de manera aproximadamente lineal. Comprobar como la función de autocorrelación parcial tiene un primer retardo muy grande, cercano a 1. Por último, definimos el proceso primera diferencia de este, $y_t = x_t - x_{t-1}$. Para ello, en el menú del wokfile:

$$y = d(x, 1) \qquad Sample : 2 : 400$$

Aquí debemos comprobar como la serie y que obtenemos es una serie estacionaria. Si además comprobamos el correlograma (View→Correlogram), veremos como la fac y la facp corresponden a la de un proceso MA(1).

Podemos repetir el mismo ejercicio con un proceso donde la constante sea diferente de 0. Generamos una serie donde $c = 1$, $\theta = -0.8$ y $\sigma_a^2 = 1$. Para ello debemos seguir los pasos:

1. Genr→x=1. (Esto nos define la serie como 1).
2. Genr→IMPORTANTE: Sample: 2:400 → x=1+x(-1)+a+0.8*a(-1). (Aqui la serie toma sus valores definitivos).

El proceso de generación asume que $x_1 = 1$. Vemos la serie:

View → Line Graph

La serie debe tener una clara tendencia positiva de pendiente 1, es decir, el valor de la constante. Vemos el correlograma. Comprobamos que la función de autocorrelación estimada de la serie decae incluso más lentamente que en el caso anterior y además de manera aproximadamente lineal. Comprobar como la función de autocorrelación parcial tiene un primer retardo muy grande, cercano a 1 y el resto son prácticamente 0. Por último, definimos el proceso primera diferencia de este, $y_t = x_t - x_{t-1}$. Para ello, en el menú del wokfile:

$$y = d(x, 1) \qquad Sample : 2 : 400$$

Aquí debemos comprobar como la serie y que obtenemos es una serie estacionaria. Si además comprobamos el correlograma (View→Correlogram), veremos como la fac y la facp corresponden a la de un proceso MA(1).

1.1.2 Procesos I(2)

Veamos las características principales de estos procesos mediante la generación de un proceso muy sencillo, $(1-B)^2 x_t = a_t$. Para ello debemos seguir los pasos:

1. Genr→x=0. (Esto nos define la serie como 0).
2. Genr→IMPORTANTE: Sample: 3:400 → x=2*x(-1)-x(-2)+a. (Aqui la serie toma sus valores definitivos).

El proceso de generación asume que $x_1 = 0$. Vemos la serie:

View → Line Graph

La serie debe tener altos y bajos demostrandose que la serie es no estacionaria. Vemos el correlograma. Comprobamos que la función de autocorrelación estimada de la serie decae incluso muy lentamente además de manera lineal. Comprobar como la función de autocorrelación parcial tiene un primer retardo muy grande, cercano a 1 y el resto son prácticamente 0. ¿Cómo distinguir este proceso de un $I(1)$? Definimos el proceso primera diferencia de este, $y_t = x_t - x_{t-1}$. Para ello, en el menú del wokfile:

$$y = d(x,1) \qquad Sample : 2 : 400$$

Aquí debemos comprobar como la serie y que obtenemos es una serie no estacionaria de nuevo y parecida a la que obtuvimos al principio de la práctica. Si además comprobamos el correlograma (View→Correlogram), veremos como la fac y la facp corresponden a la de un proceso no estacionario de nuevo.

1.2 Procesos ARIMA estacionales

Una serie es estacional si toma valores diferentes en diferentes periodos de tiempo que se repiten. Es decir, para datos mensuales, si los valores correspondientes a enero son diferentes a los valores que toma la serie en agosto. Por ejemplo, una serie que mida la temperatura máxima en Madrid en cada uno de los meses. Es evidente que la temperatura máxima en enero será muy diferente a la temperatura máxima en agosto y esto se repite en todos los años. Por lo tanto, este tipo de modelos son no estacionarios puesto que su valor medio dependerá del mes en que se tome el dato. El modelo ARIMA$(p,d,q)\times(P,D,Q)_s$ estacional completo es complicado:

$$\Phi_P(B^s)\,\phi_p(B)\,(1-B)^d\,(1-B^s)^D\,x_t = \theta_q(B)\,\Theta_Q(B^s)\,a_t$$

por lo que veremos casos un poco más sencillos.

1.2.1 Modelo de líneas aéreas

El modelo de líneas aéreas fue empleado por Box y Jenkins para ajustar datos del número de pasajeros en líneas aéreas entre enero de 1949 y diciembre de 1960, de donde ha tomado el nombre. El modelo en términos ARIMA es el siguiente:

$$(1-B)\left(1-B^{12}\right)x_t = (1-\theta B)\left(1-\Theta B^{12}\right)a_t$$

es decir, un ARIMA$(0,1,1)\times(0,1,1)_{12}$ o un doble IMA para la parte regular y la estacional. Para generar una serie a partir de este modelo tomamos por ejemplo, $\theta = -0.7$ y $\Theta = -0.8$. Entonces podemos escribir:

$$x_t = x_{t-1} + x_{t-12} - x_{t-13} + a_t + 0.7a_{t-1} + 0.8a_{t-12} + 0.56a_{t-13}$$

y procedemos como sigue:

1. Genr→x=0. (Esto nos define la serie como 0).
2. Genr→IMPORTANTE: Sample: 14:400 → x=x(-1)+x(-12)-x(-13)+a+0.7*a(-1)+0.8*a(-12)+0.56*a(-13). (Aqui la serie toma sus valores definitivos).

El proceso de generación asume que $x_1 = 0$. Vemos la serie:

View → Line Graph

La serie debe tener altos y bajos además de un fuerte componente estacional. Vemos el correlograma. Comprobamos que la función de autocorrelación estimada de la serie decae muy lentamente además de manera lineal. Comprobar como la función de autocorrelación parcial tiene un primer retardo muy grande, cercano a 1 y el resto son prácticamente 0 salvo en la altura del retardo 12, 13, 24, 25 y alrededores. Parece que es un I(1) en cuanto su correlograma. Definimos el proceso primera diferencia de este, $y_t = x_t - x_{t-1}$. Para ello, en el menú del wokfile:

$$dx = d(x, 1) \qquad Sample : 1 : 400$$

Aquí debemos comprobar como la serie dx que obtenemos es una serie claramente estacional por lo que es no estacionaria de nuevo. Si además comprobamos el correlograma (View→Correlogram), veremos como la fac y la facp tienen muy marcadas las autocorrelaciones de orden 12, 24 etc... Pasamos a tomar una diferencia estacional. Para ello, tomamos:

$$ddx = d(x, 1, 12) \qquad Sample : 1 : 400$$

Vemos que ahora la serie ya es estacionaria. Si vamos a ver el correlograma, veremos que tenemos picos significativos al principio de los valores y a la altura de los retardos 12, 13 etc... Si vemos los primeros de estos, debemos comprobar que tenemos un MA(1) en la parte regular, mientras que en la parte estacional es mucho más complicado de determinar. Cuando pasemos a ver estimación volveremos a este tipo de modelos y a determinar más claramente como proceder con ellos para poder identificarlos de manera sencilla.

Ejemplo: Vemos el ejemplo de los datos del libro de Box y Jenkins. Para ello creamos un workfile de datos mensuales de 1949:1 a 1960:12. Una vez importados los datos, comprobamos como los datos no tienen varianza constante, si no creciente. Tomamos el logaritmo de dichos datos y comprobamos que ahora los datos no aparentan un cambio de varianza. Vemos el correlograma y comprobamos que el correlograma se parece al que vimos para la serie generada. Tomamos una diferencia de la serie y como antes vemos que el gráfico muestra estacionalidad y el correlograma se parece mucho al de la serie generada para el modelo de líneas aéreas. Lo mismo se puede decir de la serie despues de diferenciar una raíz estacional.