Series de Tiempo

Licenciatura en Ciencia de Datos

IIMAS-UNAM 2

Ramsés H. Mena

Clase 2: ACF y transformaciones a estacionariedad

6. Función de autocorrelación (ACF)

En la Definición 4 se introdujo la definición de la función de autocovarianza, que proporciona una medida de la dependencia subyacente a un modelo dado. Por otro lado, la Definición 5, de un proceso débilmente estacionario, utiliza dicha función.

Ejemplo 1. *Consideren la serie de la Figura 2 como las mediciones en el tiempo de un fenómeno de interés.*

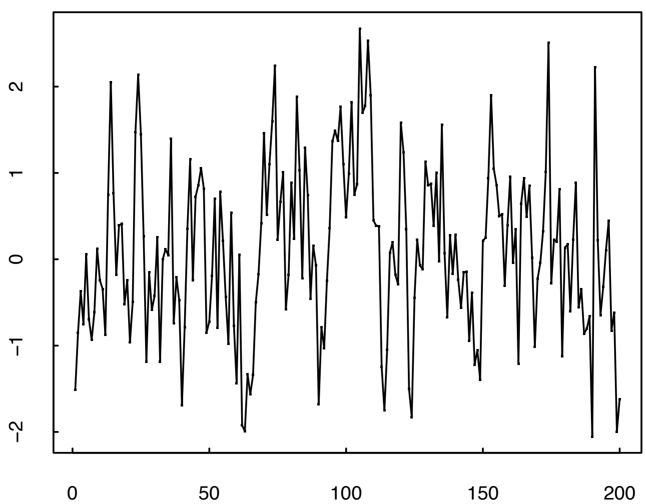

Figura 2: Serie simulada.

Para describir a a dicha serie se propone el siguiente modelo

$$X_t = \phi\, X_{t-1} + \sqrt{1-\phi^2}\,\varepsilon_t, \tag{11}$$

donde $0 < \phi < 1$ *y* $\varepsilon_t \sim N(0,1)$.

Una forma de ver como el modelo captura la dinámica de los datos es a través de la función de autocorrelación. Primero, observemos la ACF muestral en la Figura 3, donde se puede apreciar una que la autocorrelación subyacente a los datos tiende a desaparecer conforme el tiempo pasa.

Por otro lado, se puede ver que la ACF correspondiente al modelo (11), está dada por Corr$(X_t, X_{t+h}) = \phi^h$. *Si estimamos el valor de* ϕ *(e.g. usando máxima verosimilitud), se puede ver que* $\hat{\phi}$ *=0.49. Así pues, se podría decir que el modelo no es " tan malo ", ya que al menos la ACF*(1) *se ajusta bien.*

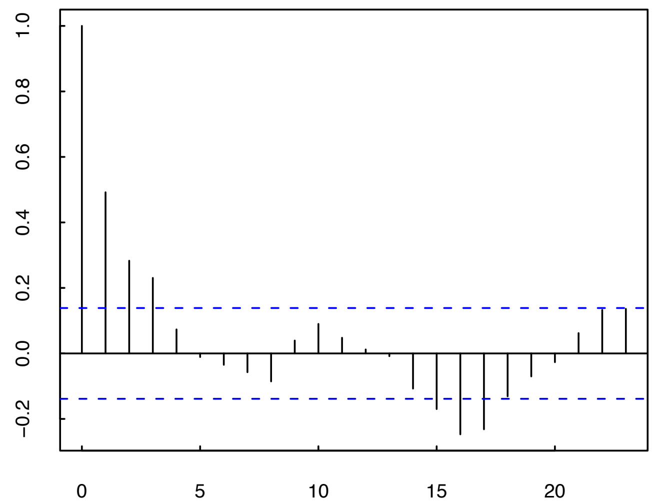

Figura 3: ACF de la serie en Figura 2. Los primeros cuatro valores son: 1, 0.492, 0.283, 0.231 y 0.073 respectivamente.

Tarea: Calcular la ACF del modelo (11).

7. Transformación del modelo clásico a estacionaridad.

El Ejemplo 1 muestra una serie " fácil " de analizar, sin embargo las situaciones que se observan de series de datos reales no son tan sencillas.

En la Figura 4 se observa la serie correspondientes al número de pasajeros en en aerolíneas internacionales. Se puede ver claramente que dicha serie exhibe cierta ***tendencia*** y cierta ***estacionalidad*** o ***ciclicidad***.

Estos datos se podrían ver como la realización de un modelo de la forma

$$X_t = m_t + s_t + Y_t, \tag{12}$$

donde m_t y s_t son funciones " deterministas " que representan cierta tendencia y ciclicidad respectivamente y Y_t denota un " ruido aleatorio " dado a través de un proceso débilmente estacionario (ver Definición 5). De manera ideal el componente de tendencia m_t, se debe de comportar de una manera " suave " y sin cambios muy " abruptos " mientras que la función de ciclicidad deberá ser una función con periodo conocido d.

Parte del la tarea inicial de la teoría de series de tiempo es la estimación y extracción de dichos componentes, m_t

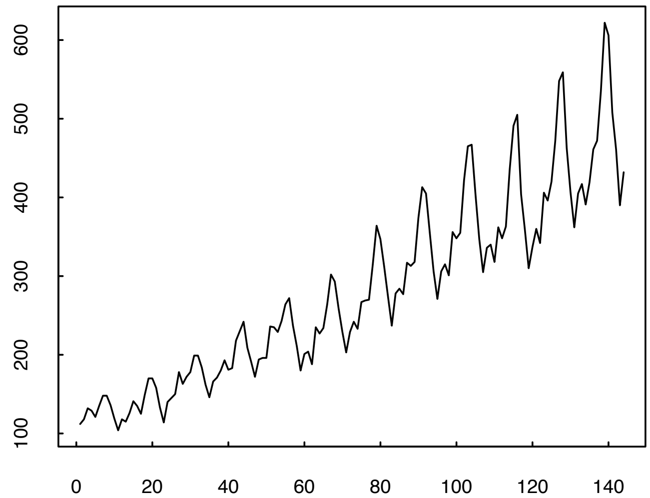

Figura 4: Número de pasajeros en aerolíneas internacionales: totales mensuales de enero de 1949 a Diciembre de 1960. Las cifras reportadas están en miles. Fuente: Box and Jenkins (1970).

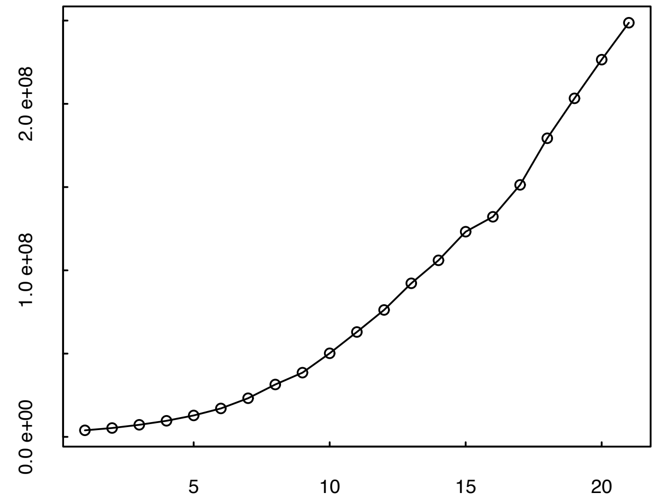

Figura 5: Población (en millones) de EUA durante los años 1790–1990. Fuente: Brockwell and Davis (2002).

y s_t, con la esperanza de que el componente de ***residuo*** o componente de ***ruido***, Y_t, resulte en una serie de tiempo estacionaria, y así modelar este último mediante alguno de los modelos básicos de series de tiempo.

8. Estimación y eliminación de tendencia en ausencia de estacionalidad

En esta sección exploraremos algunas técnicas para remover la tendencia de un modelo sin componente cíclico, es decir de la forma

$$X_t = m_t + Y_t, \tag{13}$$

donde $\mathbb{E}[Y_t] = 0$[1].

8.1. Estimación de m_t por mínimos cuadrados

Este método consiste en modelar, "ajustar", una familia de funciones paramétricas, por ejemplo

$$m_t = a_0 + a_1\, t + a_2\, t^2, \tag{14}$$

a los datos, escogiendo los parámetros (a_0, a_1 y a_2 en el ejemplo arriba) de tal forma que se minimize

$$\sum_t (x_t - m_t)^2. \tag{15}$$

Ejemplo 2. *Considérese la serie proveniente de la población en EUA durante el periodo de 1970–1980. Ver Figura 5*

Si suponemos un modelo de la forma (13) y modelamos la tendencia con una función cuadrática de la forma (14), los estimadores resultantes del procedimiento de mínimos cuadrados son $\hat{a}_0 = 2.097911 \times 10^{10}$ $\hat{a}_1 = -2.334962 \times 10^7$ *y* $\hat{a}_2 = 6.498591 \times 10^3$.

Como consecuencia la idea sería analizar los residuales $Y_t = X_t - \hat{m}_t$*, claro que no para cualquier serie dicha cantidad resultará en un proceso débilmente estacionario.*

8.2. Método de diferenciación

Este método consiste en transformar los datos mediante el ***operador de diferencias***,∇, definido por

$$\nabla X_t = X_t - X_{t-1}, \tag{16}$$

con la finalidad de que el modelo resultante sea débilmente estacionario. Por ejemplo veamos que sucedería si aplicamos este operador a el modelo con tendencia lineal dado por:

$$X_t = a_0 + a_1\, t + Y_t,$$

donde una ves más suponemos que Y_t es débilmente estacionario y $\mathbb{E}[Y_t] = 0$. Tenemos

$$\nabla\, X_t = a_1 + \nabla\, Y_t, \tag{17}$$

en donde se verifica que la tendencia ha sido removida.

El operador de diferencias se puede aplicar varias veces, por ejemplo $\nabla^2 X_t = \nabla(\nabla X_t) = X_t - 2\, X_{t-1} + X_{t-2}$.

[1]Si $\mathbb{E}[Y_t] \neq 0$ entonces podemos reemplazar a m_t por $m_t + \mathbb{E}[Y_t]$ y a Y_t por $Y_t - \mathbb{E}[Y_t]$.

Tarea: Aplicar el operador de diferencias dos veces al modelo (13) con tendencia dada por (14).

En general una tendencia polinómica de orden k se puede remover mediante la aplicación del operador de diferencias k veces.

8.3. Suavizamiento con filtros de promedios móviles finito

Sea q un entero no-negativo y consideremos el ***promedio móvil*** dado por

$$W_t = \frac{1}{2q+1} \sum_{j=-q}^{q} X_{t-j} \tag{18}$$

correspondiente al proceso $\{X_t\}$ definido mediante 13. Entonces para $q+1 \leq t \leq n-q$,

$$W_t = \frac{1}{2q+1} \sum_{j=-q}^{q} m_{t-j} + \frac{1}{2q+1} \sum_{j=-q}^{q} Y_{t-j} \approx m_t. \tag{19}$$

La approximation de arriba se vale si asumimos que m_t es aproximadamente lineal sobre el intervalo $[t-q, t+q]$ y que el promedio de los errores sobre este intervalo tiende a cero. Entonces los promedios móviles proporcionan el siguiente estimador

$$\hat{m}_t = \frac{1}{2q+1} \sum_{j=-q}^{q} X_{t-j}, \quad q+1 \leq t \leq n-q. \tag{20}$$

Dado que la serie no se observa para $t < 0$ o $t > n$, entonces no se puede usar (20) para $t \leq q$ o $t > n-q$. Una manera de manipular esto en práctica se puede hacer mediante $X_t := X_1$ para $t < 1$ y $X_t := X_n$ para $t > n$.

9. Eliminación de tendencia y ciclicidad

Ahora consideremos un modelo dado de la forma (12), es decir

$$X_t = m_t + s_t + Y_t, \tag{21}$$

con $E[Y_t] = 0$, $s_{t+d} = s_t$ y $\sum_{j=1}^{d} s_j = 0$.

9.1. Estimación de la tendencia y ciclicidad

Supongamos que observamos la serie $\{x_t\}_{t=1}^{n}$. Un método para remover ambas, tendencia y ciclicidad, se da mediante la aplicación de un filtro de promedios móviles construido de manera especial para eliminar el componente cíclico. Si el periodo, denotado por d, es par $d = 2q$, entonces se usa

$$\hat{m}_t = (0.5\, x_{t-q} + x_{t-q+1} + \cdots + 0.5\, x_{t+q})/d, \tag{22}$$

para $q < t < n-q$. Por otro lado si el periodo es impar $d = 2q+1$, entonces se utiliza (18).

El segundo paso es estimar el componente de ciclicidad. Para cada $k = 1, \ldots, d$ se calcula el promedio, w_k, de las desviaciones

$$\{(x_{k+jd} - \hat{m}_{k+jd}), \quad q < k + jd \leq n-q\}.$$

Así pues, el componente de ciclicidad se estima mediante

$$\hat{s}_k = w_k - \frac{1}{d} \sum_{i=1}^{d} w_i, \quad k = 1, \ldots, d, \tag{23}$$

y $\hat{s}_k = \hat{s}_{k-d}$, $k > d$.

Entonces los datos sin ciclicidad se pueden estimar mediante

$$d_t = x_t - \hat{s}_t, \quad t = 1, \ldots, n. \tag{24}$$

Finalmente se re-estima la tendencia de los datos sin ciclicidad, $\{d_t\}$ mediante el uso de alguno de los métodos descritos en la clase anterior. Así pues, los residuales se pueden estimar mediante

$$\hat{Y}_t = x_t - \hat{m}_t - \hat{s}_t, \quad t = 1, \ldots, n. \tag{25}$$

Hacer ejemplo con los datos "deaths" usando ITSM: `Transform >Classical`.

9.2. Método de diferencias

Al igual que en el modelo de sin ciclicidad (13), el método de diferencias se puede adaptar para eliminar o "difuminar" cierta ciclicidad subyacente a los datos. Considérese el ***operador de diferencias con rezago-d***, definido como

$$\nabla_d X_t := X_t - X_{t-d}. \tag{26}$$

Una aplicación de este operador al modelo dado por (21, obtenemos

$$\nabla_d X_t = m_t - m_{t-d} + Y_t - Y_{t-d}, \tag{27}$$

es decir se descompone en un componente de tendencia $(m_t - m_{t-d})$ y un componente de ruido $(Y_t - Y_{t-d})$. De igual forma, la tendencia se puede eliminar con alguno de los métodos descritos en la Sección 8.

Hacer ejemplo con los datos "deaths" usando ITSM: `Transform >Difference`, *con* `lag` *12. Para eliminar la tendencia restante se puede aplicar el operador de diferencias una vez más, es decir* $\nabla\,(\nabla_{12}\, x_t)$.