Conceptos básicos sobre Series de Tiempo

Jorge Hans Alayo Gamarra

27 de junio de 2014

En el presente reporte se presentan los conceptos básicos sobre series de tiempo.

1. Enfoque para modelar series de tiempo

- Graficar la serie y examinar las principales características del gráfico. Chequeando en particular si existe una tendencia, un componente periódico, un cambio en el comportamiento, outliers.
- Remover el componente periódico y la tendencia para obtener los residuos estacionarios. Para alcanzar este propósito a veces es necesario transformar los datos.
- Escoger un modelo que se ajuste a los residuos haciendo uso de varias muestras estadísticas incluyendo la función de autocorrelación muestral.
- Se alcanzará la estimación al estimar os residuos y luego invertir las transformaciones efectuadas para estimar la serie original.
- Un alternativa extremadamente útil es expresar la serie en términos de sus componentes de Fourier.

2. Conceptos básicos

Una serie de tiempo es una colección de observaciones indexadas a la fecha de cada observación (ver Figura 1), comenzando desde una fecha en particular ($t = 1$) hasta otra ($t = T$):

$$(y_1, y_2, y_3 \ldots, y_T) \tag{1}$$

La descomposición clásica de una serie tiene una tendencia, una componente periódica y una componente estacionaria:

$$y_t = m_t + s_t + x_t \tag{2}$$

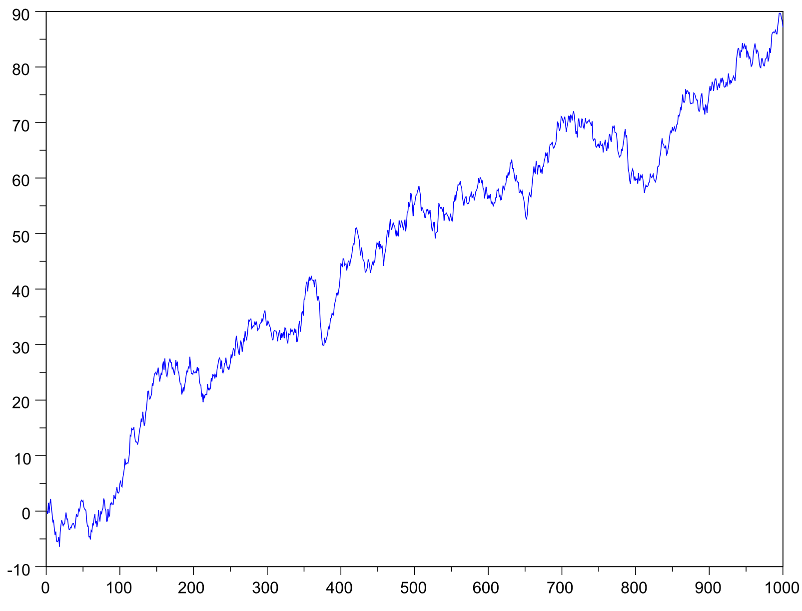

Figura 1: Ejemplo de una serie temporal

Los métodos se enfocan en analizar la componente estacionaria x_t; no obstante, estos métodos se pueden aplicar a todas las series luego de remover adecuadamente la tendencia y la componente periódica.

Para comenzar el estudio de series estacionarias, hay que definir formalmente una serie estacionaria. Para ello, definimos el esperado y la varianza de una serie de tiempo. Sea la variable $Y_t = y_1, y_2, \ldots, y_T$ una variable aleatoria de tal forma que la probabilidad de que una observación tome un valor dado no depende de las observaciones restantes, el esperado de la observación y_t se define por la siguiente expresión:

$$E(Y_t) = \int y_t \cdot f(y_t) dy_t = \mu_t \tag{3}$$

La varianza de la variable aleatoria se define por la siguiente expresión:

$$\gamma_{0t} = E(Y_t - \mu_t)^2 \tag{4}$$

Por ejemplo considérese que $Y_t = \xi_1, \xi_2, \ldots, \xi_T$,y están idénticamente distribuidas según $\xi_t \sim N(0, \sigma^2)$. El esperado y la varianza están dados por:

$$E(Y_t) = \int \xi_t \cdot \frac{1}{\sqrt{2\pi}\sigma} exp\left(\frac{-\xi_t^2}{2\sigma^2}\right) d\xi_t = 0 \tag{5}$$

$$\gamma_{0t} = E(\xi_t - 0)^2 = \sigma^2 \tag{6}$$

Luego se define la función de autocovarianzas por el siguiente esperado:

$$\gamma_{jt} = E(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j}) \tag{7}$$

Esta función describe como covaría la variable aleatoria respecto a su propio valor con rezago. Luego se dice que un proceso es estacionario si ni el esperado μ_t ni las covarianzas γ_{jt} dependen de la fecha t.

2.1. Procesos *Moving Average (MA)*

Sea ξ_t un proceso de ruido blanco tal que $\xi_t \sim N(0, \sigma^2)$ y considérese el siguiente proceso:

$$Y_t = \mu + \xi_t + \theta\xi_{t-1} \tag{8}$$

El proceso es conocido como *moving average* de orden uno MA(1) (Ver Figura 3), debido a que depende de un rezago de la componente aleatoria, se puede generalizar a un proceso MA(q) para q rezagos. El esperado está dado por:

$$E(Y_t) = E(\mu + \xi_t + \theta\xi_{t-1}) = \mu \tag{9}$$

La varianza está dada por:

$$E(Y_t - \mu)^2 = E(\mu + \xi_t + \theta\xi_{t-1} - \mu)^2 = E(\xi_t + \theta\xi_{t-1})^2 \tag{10}$$

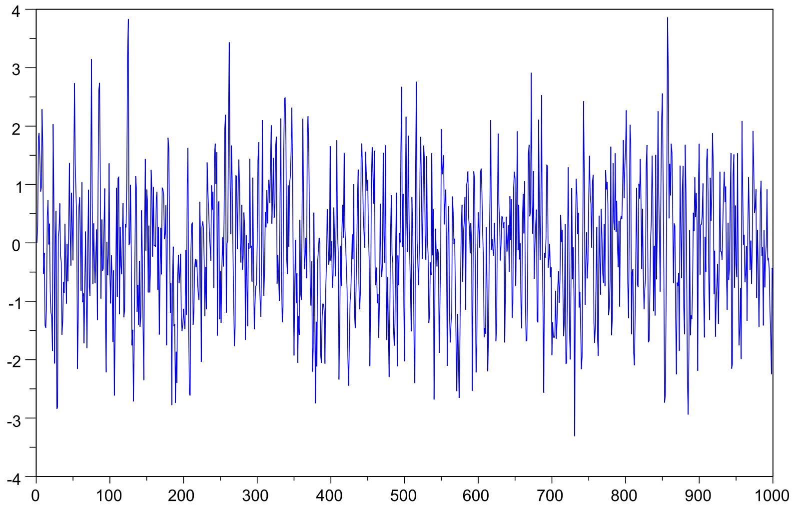

Figura 2: Proceso MA(1)

$$E(Y_t - \mu)^2 = E(\xi_t^2 + \theta^2 \xi_{t-1}^2 + 2\xi_t \xi_{t-1}) = \sigma^2(1 + \theta^2) \tag{11}$$

La primera autocovarianza está dada por:

$$E(Y_t - \mu)(Y_{t-1} - \mu) = E(\xi_t \xi_{t-1} + \theta \xi_{t-1}^2 + \theta \xi_t \xi_{t-2} + \theta^2 \xi_{t-1} \xi_{t-2}) \tag{12}$$

$$E(Y_t - \mu)(Y_{t-1} - \mu) = \theta \sigma^2 \tag{13}$$

Para mayores rezagos la función de autocovarianzas es cero.

2.2. Procesos Autoregresivos (AR)

Sea ξ_t un proceso de ruido blanco tal que $\xi_t \sim N(0, \sigma^2)$ y considérese el siguiente proceso:

$$Y_t = c + \phi Y_{t-1} + \xi_t \tag{14}$$

El proceso es conocido como autoregresivo de orden uno AR(1) (ver Figura 2), debido a que depende de un rezago de la misma variable, se puede generalizar a un proceso AR(p) para p rezagos. Luego por recursión se tiene lo siguiente:

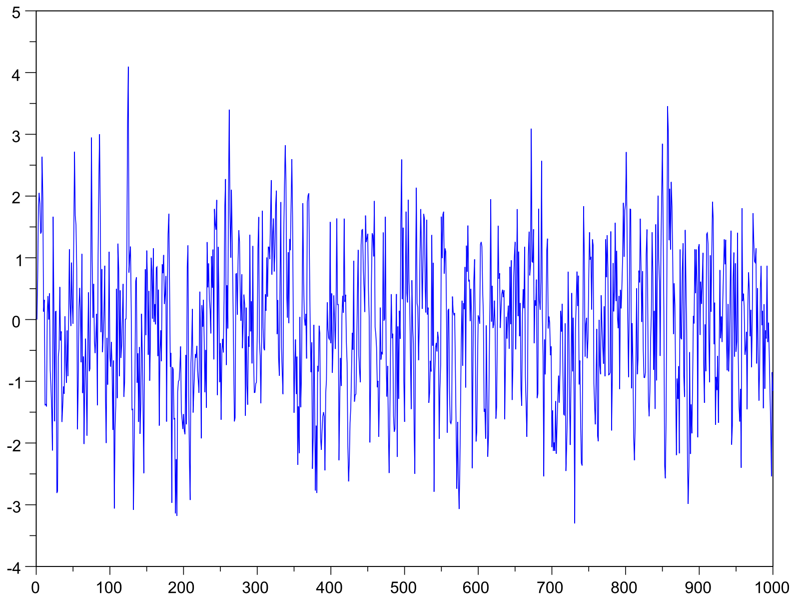

Figura 3: Proceso AR(1)

$$
\begin{aligned}
Y_1 &= c + \phi Y_0 + \xi_1 \\
Y_2 &= c + \phi Y_1 + \xi_2 = c + c\phi + \phi^2 Y_0 + \phi\xi_1 + \xi_2 \\
Y_3 &= c + \phi Y_2 + \xi_3 = c + c\phi + c\phi^2 + \phi^3 Y_0 + \phi^2\xi_1 + \phi\xi_2 + \xi_3 \\
Y_t &= c\sum_{k=0}^{t-1}\phi^k + \phi^t Y_0 + \sum_{k=0}^{t-1}\phi^k\xi_{t-k}
\end{aligned}
\tag{15}
$$

El proceso es estacionario para valores de $\phi < 1$. Luego si el proceso es estacionario, entonces el esperado está dado por:

$$E(Y_t) = c + E(\phi Y_{t-1}) + E(\xi_t) \tag{16}$$

$$E(Y_t) - E(\phi Y_t) = c \tag{17}$$

$$E(Y_t) = \frac{c}{1-\phi} \tag{18}$$

2.3. Procesos ARMA

Viene a ser la combinación de los procesos autoregresivos AR(p) y los procesos moving average MA(q) para dar lugar al proceso ARMA(p,q). Considérese que el operador L sobre una serie rezaga un periodo el valor de la variable:

$$Lx_t = x_{t-1} \tag{19}$$

Luego un proceso ARMA queda representado por la siguiente ecuación:

$$A(L)x_t = B(L)e_t \tag{20}$$

Donde $A(L)$ y $B(L)$ son polinomios del el operador L de grados p y q respectivamente.

2.4. La autocorrelación muestral

La función de autocorrelación está definida por la siguiente expresión:

$$\rho = \frac{\gamma_j}{\gamma_0} \tag{21}$$

Donde:

$$\gamma_j = E(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j}) \tag{22}$$

Entonces el estimado natural de una población está dado por:

$$\hat{\rho} = \frac{\hat{\gamma}_j}{\hat{\gamma}_0} \tag{23}$$

Donde:

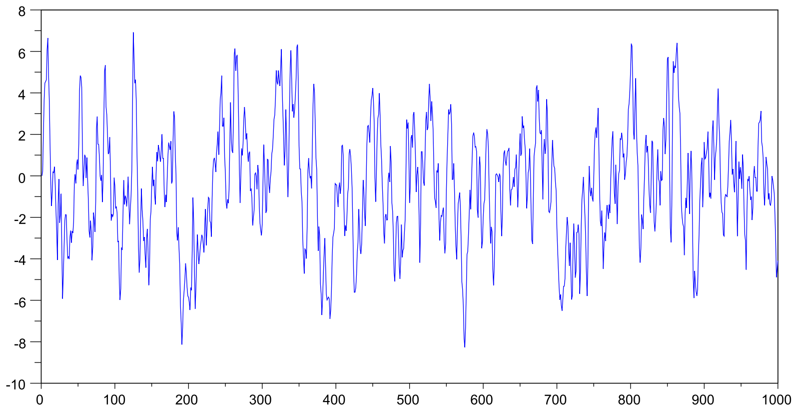

Figura 4: Proceso ARMA(1,1)

$$\hat{\gamma}_j = \frac{1}{T} \sum_{t=j+1}^{T} (y_t - \bar{y})(y_{t-j} - \bar{y}) \tag{24}$$

$$\bar{y} = \frac{1}{T} \sum_{t=1}^{T} y_t \tag{25}$$

Dicha autocorrelación muestral es una herramienta útil para analizar un conjunto de datos, ya que permite observar la dependencia de la variable con sus rezagos. por ejemplo para un proceso AR(1) se tiene la siguiente gráfica:

2.5. Raíces unitarias

Cuando se discutió los procesos AR(1), se indicó que el proceso es estacionario si el parámetro $|\phi| < 1$.

$$(1 - \phi L)y_t = \mu + e_t \tag{26}$$

Si $\phi = 1$ se dice que el proceso tiene una raíz unitaria y no es estacionario; ello puede causar problemas de inferencia estadística en los modelos de series de tiempo si no es detectada debidamente. La principal característica de este proceso es que los shock tienen un efecto permanente sobre la tendencia de la serie a diferencia de una serie que tiene una tendencia natural y que los shocks tienen efectos transitorios y no modifican la tendencia.

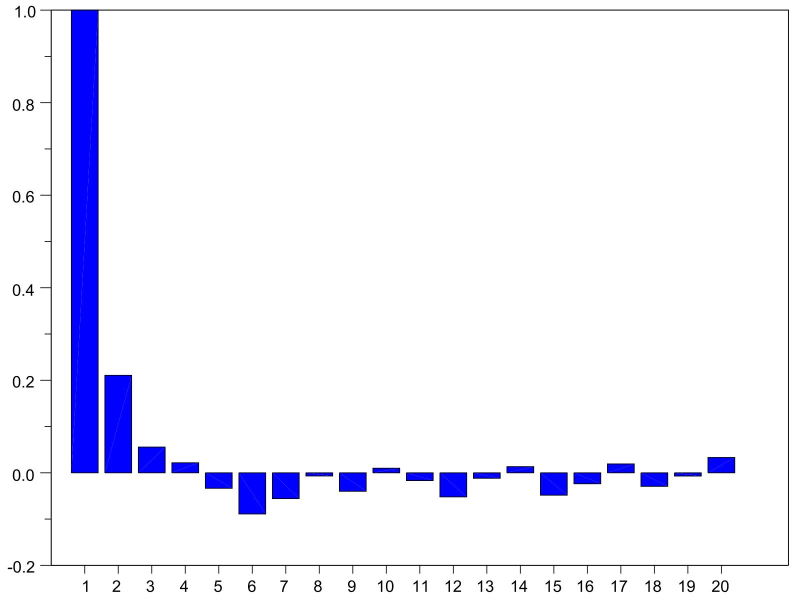

Figura 5: Autocorrelación muestral

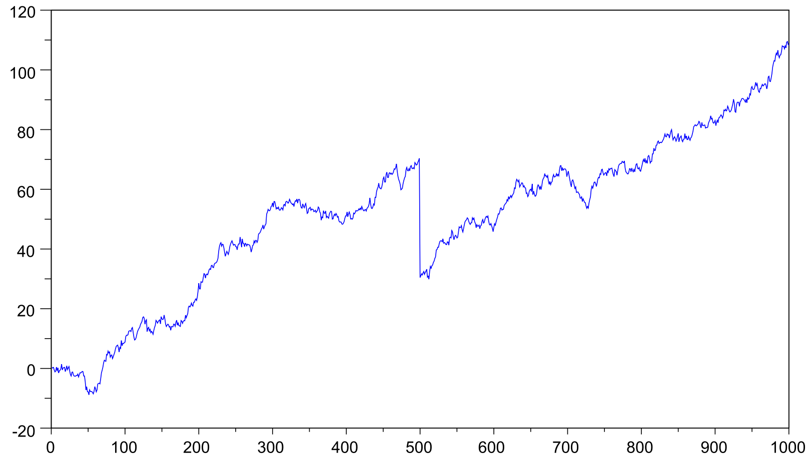

Figura 6: Respuesta de un proceso con raíz unitaria ante un shock

3. Técnicas para remover la tendencia y estacionalidad

3.1. Estimar y remover la tendencia sin estacionalidad

Existen dos técnicas para eliminar la tendencia de una serie, el primer método consiste en ajustar una función $f(t)$ a los datos,por ejemplo:

$$m_t = a_0 + a_1 t + a_2 t^2 \tag{27}$$

Luego, se estiman los parámetros por mínimos cuadrados ordinarios:

Los residuos de la serie es la parte estacionaria

El segundo método consiste en transformar la serie, denote a Δ al operador diferencia:

$$\Delta y_t = (1 - L) y_t = y_t - y_{t-1} \tag{28}$$

Luego el método consiste en aplicar el operador Δ de forma consecutiva:

$$\Delta^d y_t = (1 - L)^d y_t = (1 - L)^d m_t + (1 - L)^d x_t \tag{29}$$

Se debe seleccionar d de forma que $(1 - L)^d m_t$ sea constante, usualmente $d = 1, 2$. El problema entre ambos métodos es el problema de la raíz unitaria.

3.2. Estimar y remover la tendencia con estacionalidad

Cuando se tiene estacionalidad primero se estima la tendencia de los datos transformados para eliminar la estacionalidad. Para ello se utiliza un moving average elegido especialmente para remover la estacionalidad, suponerse que se tienen $x_0, x_1, \ldots, x_t$ observaciones. Si el periodo d es par, por ejemplo $d = 2q$ entonces usamos lo siguiente transformación:

$$\hat{m}_t = (0{,}5 x_{t-q} + x_{t-q+1} + \cdots + x_{t+q-1} + 0{,}5 x_{t+q})/d \tag{30}$$

Donde $t \in < q, n - q >$Si el periodo es impar se utiliza el siguiente moving average:

$$\hat{m}_t = \sum_{j=-q}^{q} X_{t-j}/(2q + 1) \tag{31}$$

Donde $t \in < q + 1, n - q >$, el siguiente paso es estimar la componente estacional, para ello se calcula para $k = 1, 2, \ldots, d$ la desviación promedio $w_k = x_{k+jd} - \hat{m}_{k+jd}$, donde $k + jd \in < q, n - q >$, luego

$$\hat{s}_k = w_k - \sum_{i=1}^{d} w_i/d, \qquad k = 1 \ldots d \tag{32}$$

Finalmente se resta el componente estacional y se aplica una de las técnicas mencionadas en la sección anterior para eliminar la tendencia.

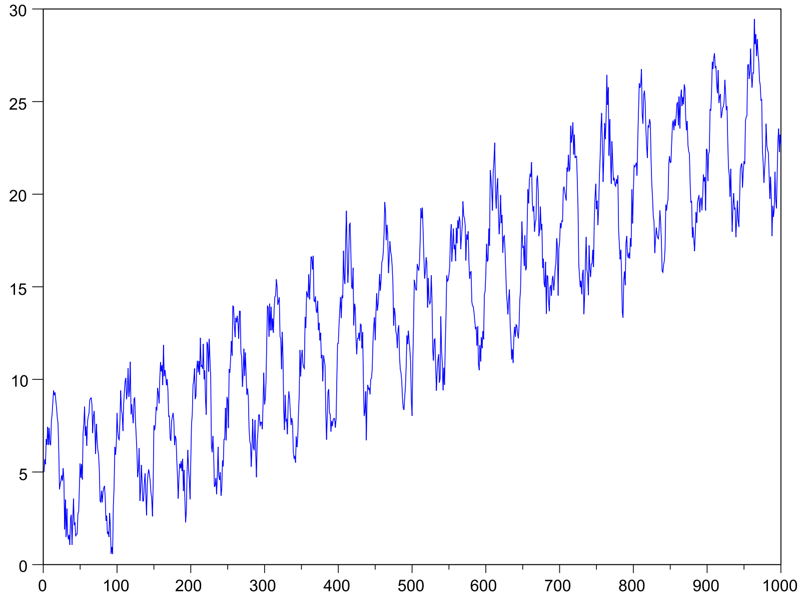

Figura 7: Ejemplo de una serie temporal con tendencia y estacionalidad

Otro método alternativo es aplicar el operador $\nabla_d = 1 - L^d$, con ello solo quedaría la serie de tiempo con la tendencia y luego se puede aplicar el operador ∇ para conseguir una serie estacionaria.

4. Outliers

Un outlier es un punto en los datos que es significativamente diferente al resto de los puntos.

Asimismo existen diversos tipos de outliers:

- Outliers aditivos.
- Outliers innovativos.
- Outliers de cambio de nivel.
- Outliers de cambio transitorio.

A continuación se revisa la definición de cada uno.

4.1. Outliers aditivos

Es el tipo de outlier que solo afecta a una sola observación. Después de este disturbio, la serie retorna a su senda regular. El efecto de un outlier aditivo en el tiempo $t = T$ con magnitud ω está dado por:

$$Z_t = X_t + \omega I_t^T = \frac{B(L)}{A(L)} a_t + \omega I_t^T \tag{33}$$

Donde $I_t^T = 1, t = T$ y $I_t^T = 0, t \neq T$

4.2. Outliers innovativos

Es el tipo de outlier que afecta a las siguientes observaciones comenzando desde un shock inicial y que se propaga a las siguientes observaciones. El efecto de un outlier innovativo está dado por:

$$Z_t = X_t + \omega I_t^T = \frac{B(L)}{A(L)} (a_t + \omega I_t^T) \tag{34}$$

Donde $I_t^T = 1, t = T$ y $I_t^T = 0, t \neq T$

4.3. Outliers de cambio de nivel

El cambio de nivel es como la función escalón. Para un proceso estacionario, un cambio de nivel implica un cambio en la media del proceso después de un punto y consecuentemente el proceso es transformado en uno no estacionario.

4.4. Outliers de cambio transitorio

Es un pico que afecta a unos pocos periodos y desaparece exponencialmente; el impacto no es permanente.

4.5. Detección de outliers

Existen dos problemáticas asociadas a los outliers:

- Identificar que puntos de una serie de tiempo son outliers.
- Que hacer con los puntos que han sido identificados como outliers.

A continuación se presentan las metodologías revisadas para la detección y procesamiento de outliers:

4.5.1. Two-sided median method

El método está basado en Basu y Meckesheimer. Se utiliza el valor de la mediana en una vecindad de los datos para determinar si un punto en particular es un outlier. Sea una serie de tiempo $y_1, y_2, \ldots, y_N$, se define una vecindad de puntos $\eta_t^k = \{y_{t-k}, \ldots y_{t-1}, y_{t+1}, \ldots, y_{t+k}\}$ donde $2k$ es el tamaño de la ventana.

Para limpiar los datos se calcula la mediana en la vecindad, m_t^k y se compara con y_t, luego se calcula la desviación y se compara con un umbral $|y_t - m_t^k| \leq \tau$. Si no se supera el umbral se conserva el dato, de lo contrario se reemplaza el dato por m_t^k y se etiqueta a la observación y_t como un outlier.

4.5.2. Detección con parámetros de la serie conocidos

Sea:

$$\pi(L) = \frac{A(L)}{B(L)} \tag{35}$$

Los residuos de una serie de tiempo serían:

$$e_t = \pi(L) Z_t \tag{36}$$

Entonces para un outlier aditivo:

$$e_t = \omega \pi(L) I_t^T + a_t \tag{37}$$

Mientras que para un outlier innovativo:

$$e_t = \omega \pi(L) I_t^T + a_t \qquad var(e_t) = var(\hat{\omega}) = \delta_a^2 \tag{38}$$

a_t es ruido blanco, luego se puede estimar ω mediante una estimación de mínimos cuadrados ordinarios. Se recomienda un apropiado test estadístico para el proceso de detección.

4.5.3. Metodología de Perrón - Rodriguez

En el contexto de procesos con raíz unitaria, una forma iterativa de detectar outliers es utilizar las primeras diferencias de los datos. La metodología en palabras simples consiste en estimar por mínimos cuadrados el tamaño δ del outlier y aplicar la prueba estadística τ para verificar si $\delta = 0$ Si el proceso generador de datos está dado por la siguiente expresión:

$$y_t = d_t + \sum_{i=1}^{m} \delta_i D(T_{ao,i})_t + u_t$$
$$(1-L)u_t = v_t \tag{39}$$

Donde $D(T_{ao,i}) = 1$ si $t = T_{ao,i}$ y 0 en otro caso, $d_t = \mu$, y $v_t = \sum_{i=0}^{oo} \varphi_i L_i e_t$. Entonces la primera diferencia está dada por:

$$\Delta y_t = \delta[D(T_{ao,i})_t - D(T_{ao,i})_{t-1}] + v_t \tag{40}$$

Luego se puede estimar el valor de $\hat{\delta}$ por mínimos cuadrados ordinarios. Luego si se estima δ para cada t entonces se tiene el siguiente estadístico:

$$t_{\hat{\delta}}(T_{ao}) = \frac{\hat{\delta}}{2(\hat{R}(0) - \hat{R}(1))} \tag{41}$$

Luego se utiliza el valor de $\tau_d = sup_{T_{ao}} |t_{\hat{\delta}}(T_{ao})|$ para rechazar o no un outlier. utilizando los valores críticos encontrados por Perron y Rodriguez donde $\hat{R}(j) = T^{-1} \sum_{t=1}^{T-j} \hat{v}_t \cdot \hat{v}_{t-j}$.

5. Estimación por mínimos cuadrados ordinarios

La formulación general del modelo re regresión lineal clásico es la siguiente:

$$y = X\beta + e \tag{42}$$

En donde y es la variable dependiente y x es el vector de variables independientes. Los supuestos básicos de este modelo son los siguiente:

- A1: y es lineal en los parámetros β.
- A2: No existe una relación lineal entre las variables independientes.
- A3: $E(e_i|X) = 0$ el valor esperado del error no es una función de las observaciones.
- A4: $E(ee'|X) = \sigma^2 I$
- A5: $e|X \sim N(0, \sigma I)$

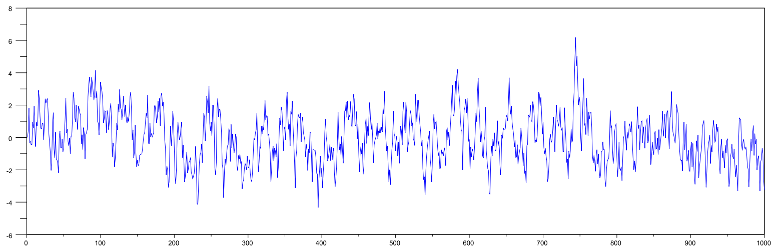

Figura 8: Proceso AR(1), $\phi = 0{,}8$

Luego se debe minimizar $e'e$

$$\text{Minimizar } e'e = (y - X\beta)'(y - \beta) \tag{43}$$

$$\text{Minimizar } e'e = (y' - \beta'X')(y - X\beta) \tag{44}$$

$$\text{Minimizar } e'e = (y'y - \beta'X'y - y'X\beta + \beta'X'X\beta) \tag{45}$$

$$\text{Minimizar } e'e = (y'y - 2y'X\beta + \beta'X'X\beta) \tag{46}$$

$$\frac{\partial e'e}{\partial \beta} = 0 \implies -2X'y + 2X'X\beta = 0 \tag{47}$$

$$\beta = (X'X)^{-1}X'y \tag{48}$$

5.1. Ejemplo:

Considérese el proceso AR(1) $y_t = \phi \cdot y_{t-1} + e_t$, luego $X = y_{t-1}$, $\beta = \phi$. Si se generan los datos con un computador de tal forma que $e_t \sim N(0,1)$ y $\phi = 0{,}8$ (ver Figura 5) es posible estimar a partir de los datos generados el parámetro ϕ a partir de la ecuación (41):

$$\phi = (y'_{t-1}y_{t-1})^{-1}y'_{t-1}y_t \tag{49}$$

$$\phi = \frac{y'_{t-1}y_t}{(y'_{t-1}y_{t-1})} \tag{50}$$

$$\phi = \frac{\sum_{t=1}^{T} y_{t-1}y_t}{\sum_{t=1}^{T} y_{t-1}^2} \tag{51}$$

Utilizando los datos y la formula obtenida se estimo $\phi = 0{,}82$