Análisis de Serie de Tiempo

Camilo Enrique Argoty Pulido y Verónica Estela Pastor

Especialización en Inteligencia Artificial

30 de agosto de 2024

Series temporales - Procesos Estocásticos

Proceso Estocástico

Un proceso estocástico es una secuencia de valores aleatorios $\{Y_t : t = 0, \pm 1, \pm 2, ...\}$ y sirve de modelo para una serie temporal observada.

Definiciones relacionadas a un proceso estocástico:

- La **función media** como:

$$\mu_t = E(Y_t), \forall t \in \mathbb{Z}.$$

- La **función de autocovarianza**, $\gamma_{t,s}$, se define como

$$\gamma_{t,s} = Cov(Y_t, Y_s), \forall t, s \in \mathbb{Z}$$

  donde $Cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E(Y_t, Y_s) - \mu_t \mu_s$.

- La **función de autocorrelación**, $\rho_{t,s}$, viene dada por

$$\rho_{t,s} = Corr(Y_t, Y_s), \forall t, s \in \mathbb{Z}$$

  donde $Corr(Y_t, Y_s) = \dfrac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} = \dfrac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}$.

Funciones de autocorrelación y autocorrelación parcial

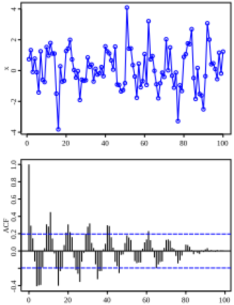
Se define la función $\mathcal{R}$, como la función que, dado un rezago τ:

$$\mathcal{R}(\tau) = E(x_t - \mu)(x_{t-\tau} - \mu)$$

Al calcular de nuevo la correlación entre la serie y_t y la serie $y_{t-\tau}$, se obtiene:

$$r(\tau) = \frac{\mathcal{R}(\tau)}{\mathcal{R}(0)}$$

La función $r(\tau)$ se conoce como función de autocorrelación, y su gráfica se conoce como correlograma.



La función $r_p(\tau)$ de autocorrelación parcial es la misma $r(\tau)$, pero restándole la parte que se debe a la interacción entre las series con rezagos intermedios (x_{t-s}) con $s \neq 1, \tau$.

Ejemplo: media móvil

Sean $e_1, e_2, ...$ una secuencia de variables aleatorias *iid*, con media cero y varianza σ^2. Supongamos que $\{Y_t\}$ se construye como

$$Y_t = \frac{e_t + e_{t-1}}{2}$$

Podemos mostrar que $\mu_t = 0, \forall t$ y también que $\gamma_{t,s} = 0{,}5\ var(e)$.

La función de autocorrelación, está dada por:

$$\rho_{t,s} = \begin{cases} 1 & si\ |t-s| = 0 \\ 0{,}5 & si\ |t-s| = 1 \\ 0 & si\ |t-s| > 1 \end{cases}$$

ya que $0{,}25\sigma_e^2 / 0{,}5\sigma_e^2 = 0,5$.

Notemos que: $\rho_{2,1} = \rho_{3,2} = 0{,}5$, $\rho_{3,1} = \rho_{4,2} = 1$ y, en general, $\rho_{t,t-k}$ es la misma para todos los valores de t. Esto nos lleva al concepto de estacionariedad.

Estacionaridad

Decimos que un proceso estocástico es estacionario si sus propiedades no cambian con el tiempo.

Un proceso Y_t se considera estrictamente estacionario si la distribución conjunta de las variables aleatorias es idéntica para cualquier elección de puntos temporales $t_1, t_2, ..., t_n$ y para todas las selecciones de desfase temporal k.

Así, resulta que $\mu_t = \mu_{t-k}$ y $\sigma_t^2 = \sigma_{t-k}^2 \; \forall t, \; k$. También se tiene que $\gamma_{t_1,k} = \gamma_{t_2,k}$. $\forall t_1, t_2, \; k$.

Estás son consecuencias de la estacionaridad, pero no son suficientes para la misma. Sin embargo, dado que la estacionaridad estricta es muy rara en la naturaleza, se suele utilizar las anteriores propiedades para definir una forma débil de estacionaridad.

Así, un proceso estocástico se dice débilmente estacionario si las funciones media y varianza son constantes en el tiempo y $\gamma_{t_1,k} = \gamma_{t_2,k}$. $\forall t_1, t_2, \; k$. Cuando la distribución de la serie cambia en el tiempo, se dice que la serie es no estacionaria.

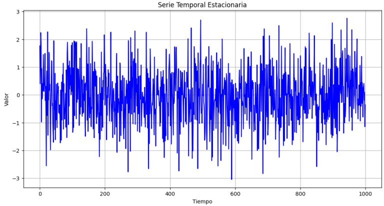
Promedios móviles

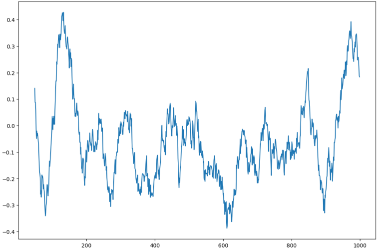
- Para una serie de tiempo y_n, el promedio móvil de $(2k+1)$ términos está dado por:

$$T_n = \frac{1}{2k+1} \sum_{j=-k}^{k} y_{n+j}$$

- Si modificamos la definición intercambiando la media por la mediana obtenemos la mediana móvil, definida como:

$$T_n = mediana\{y_{-k}, ..., y_k\}$$



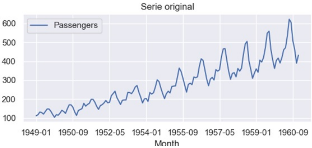


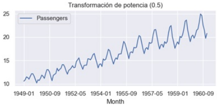
En general, la mediana móvil puede capturar cambios en la tendencia más rápido que el promedio móvil.

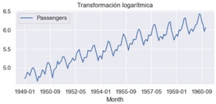
Transformaciones de Variables

Se conoce como Transformaciones de Box-Cox a la familia de funciones de potencial λ. Estas transformaciones combinan el objetivo de encontrar una relación simple, con homogeneidad de varianzas, mejorando la normalidad.

Potencia	Transformación	Descripción
$\lambda = 2$	$Y' = Y^2$	Cuadrado
$\lambda = 1$	$Y' = Y$	Datos sin Transformar
$\lambda = 0{,}5$	$Y' = \sqrt{Y}$	Raíz Cuadrada
$\lambda = 0{,}333$	$Y' = \sqrt[3]{Y}$	Raíz Cúbica
$\lambda = 0$	$Y' = ln(Y)$	Logaritmo
$\lambda = -0{,}5$	$Y' = \frac{1}{\sqrt{Y}}$	Raíz Cuadrada Inversa
$\lambda = -1$	$Y' = \frac{1}{Y}$	Recíproco







Diferenciación

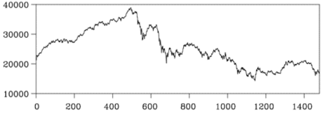
Si la serie presenta un tendencia se puede analizar en su lugar la serie diferenciada. Dada una serie de tiempo $y_n,\ n = 0, 1, \ldots$ definimos

$$z_n = \Delta y_n = y_n - y_{n-1}$$

La idea es que si tenemos $y_n = an + b$, al diferenciar se obtiene una constante.

Cuando se observa un ciclo anual en las series temporales, podemos utilizar la diferencia entre la serie temporal en el momento presente y un ciclo anterior definida por

$$\Delta_p y_n = y_n - y_{n-p}$$



(g) Stock price index data

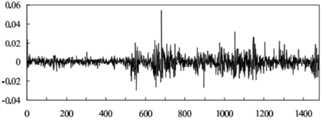


Figure 1.2: *Difference of the logarithm of the* Nikkei 225 *data.*

Ruido blanco

Es un proceso estacionario que se define como una secuencia de variables aleatorias independientes e idénticamente distribuidas e_t. El hecho de que e_t sea estrictamente estacionaria es fácil de ver,

$$P(e_{t_1} \leqslant x_1,\ e_{t_2} \leqslant x_2, ...,\ e_{t_n} \leqslant x_n) \underbrace{=}_{indep.} P(e_{t_1} \leqslant x_1)(e_{t_2} \leqslant x_2)...(e_{t_n} \leqslant x_n) =$$

y al ser estacionaria, tiene idéntica distribución,

$$= P(e_{t_1-k} \leqslant x_1)(e_{t_2-k} \leqslant x_2)...(e_{t_n-k} \leqslant x_n) = P(e_{t_1-k} \leqslant x_1, ..., e_{t_n-k} \leqslant x_n).$$

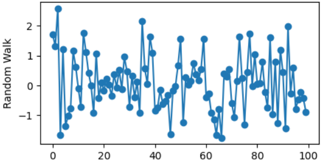
El término ruido blanco surge del hecho de que un análisis de frecuencia del modelo muestra que, por analogía con la luz blanca, todas las frecuencias entran por igual. El proceso de ruido blanco tiene media cero y $Var(e_t) = \sigma_e^2$.

Caminante aleatorio

Sean $e_1, e_2, ...$ una secuencia de variables aleatorias *iid*, cada una con media cero y varianza σ^2.

La serie temporal observada, $\{Y_t : t = 1, 2, ...\}$, se construye como:

$$\begin{cases} Y_1 = e_1 \\ Y_2 = e_1 + e_2 \\ \vdots \\ Y_t = e_1 + ... + e_t \end{cases}$$



Reescribiéndolo de manera recurrente: $Y_t = Y_{t-1} + e_t$, con c.i. $Y_1 = e_1$.

Las $e's$ representan el tamaño de los pasos dados (hacia delante o hacia atrás) a lo largo de una recta numérica, Y_t es la posición del caminante aleatorio en el momento t.

Podemos mostrar que $\mu_t = 0, \forall t$ y también que $\gamma_{t,s} = t\ var(e)$.

El caminante aleatorio se construye a partir de ruido blanco, pero no es estacionario.

Coseno aleatorio

Es un proceso aleatorio, que se define como:

$$Y_t = cos[2\pi(t/12 + \Phi)],\ \forall t \in \mathbb{Z}$$

donde Φ se selecciona (una vez) de una distribución uniforme en el [0,1].

Una muestra de un proceso de este tipo parecerá altamente determinista, ya que Y_t se repetirá idénticamente cada 12 unidades de tiempo y se parecerá a una curva coseno perfecta (en tiempo discreto).

Sin embargo su máximo no se producirá en $t = 0$ sino que vendrá determinado por la fase aleatoria Φ. La fase Φ puede interpretarse como la fracción de un ciclo completo completada en el tiempo $t = 0$.

Grafiquemos series temporales con Camilo...

Descomposición de series de tiempo

Modelo Aditivo - Modelos Multiplicativo

Dada una serie temporal denotada por y_t, se supone que puede descomponerse de modo:

$$\begin{aligned} \text{aditivo} \quad & y_t = T_t + S_t + I_t \\ \text{multiplicativo} \quad & y_t = T_t \times S_t \times I_t \end{aligned}$$

donde es la T tendencia, S es el componente estacional e I las fluctuaciones irregulares.

- La tendencia, dijimos, es la dirección general de la variable en el período de observación, debe recoger el movimiento a largo plazo de la serie, independientemente de otros componentes irregulares.
- La estacionalidad o fluctuaciones periódicas o sistemáticas de la variable, debe recoger las oscilaciones que se producen con un período inferior o igual al año. Es decir, son oscilaciones a corto plazo que se repiten en años sucesivos.
- Mientras que las fluctuaciones irregulares deben recoger, obviamente la parte aleatoria e imprevisible, pero también lo no previsible y que, luego, se pueda justificar (por ejemplo, es algún tipo crisis).

Algunos autores, incluyen un término de **ciclo**, C, que desde el punto de vista estadístico, incluye cualquier característica que no sea tendencia, estacionalidad y ruido.