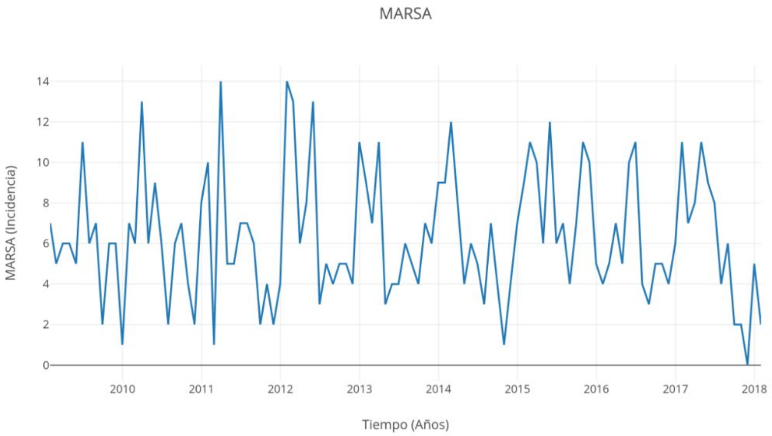
Series Temporales

Modelos de IA

Francisco Gallego Perona

¿Qué es una serie temporal?

- Un conjunto de puntos en el espacio ordenados en el tiempo.
- ¿Qué nos interesa de una serie temporal?
  - Estudiar elementos como la estacionariedad y la autocorrelación.
  - **Forecast** o predicción.



Procesos autorregresivos

- Un proceso autorregresivo es aquel que se puede expresar como una combinación lineal de sus propios valores pasados.

$$z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_k z_{t-k} + a_t$$

- z → Valor i de la serie.
- φ: Coeficiente autorregresivo.
- C: Constante relacionada con la media de la serie y los coeficientes autorregresivos.
- a: Perturbación aleatoria.

El proceso anterior es un AR(2).

Procesos de media móvil

Los procesos de media móvil (MA) siguen una estructura similar a los procesos autorregresivos.

$$z_t = C - \boxed{\theta_1 a_{t-1}} - \theta_2 a_{t-2} - \ldots - \theta_k a_{t-k} + a_t$$

a → perturbaciones aleatorias de instantes anteriores.

El proceso anterior sería un MA(2).

Procesos Mixtos

Los procesos mixtos tienen la siguiente forma:

$$z_t = C + \phi_1 z_{t-1} - \theta_1 a_{t-1} + a_t$$

El ejemplo anterior es un proceso con componente AR(1) y componente MA(1), y se expresa como ARMA(1, 1).

Procesos integrados

Un proceso integrado es aquel proceso que incluye el paso de "diferenciación". Debemos realizar la operación de integración para volver a la serie temporal original.

¿Por qué no estudiar la serie temporal original?

Porque pueden existir patrones (ruido) que no describen el proceso de la serie temporal, haciendo que, los patrones realmente importantes no puedan estudiarse correctamente (descubrimos estos patrones).

Estacionalidad

Los modelos ARIMA también pueden presentar estacionalidad. Un modelo con estacionalidad básico y muy común en la práctica es el $ARIMA(0,1,1)_s$, que tiene las mismas componentes no estacionales que el módelo no estacional $ARIMA(0,1,1)$, pero la componente MA(1) no se detecta en el lag 1 como sucedería al ser no estacional, sino que se presenta en el lag s. De igual manera ocurriría con un modelo que presente componente AR estacional.

Esto hace más complejo el estudio de las funciones de autocorrelación, al tener que disociar la parte estacional de la parte no estacional.

El ejemplo anterior quedaría de la siguiente manera:

$$z_t = z_{t-s} + C - \phi_s a_{t-s} + a_t$$

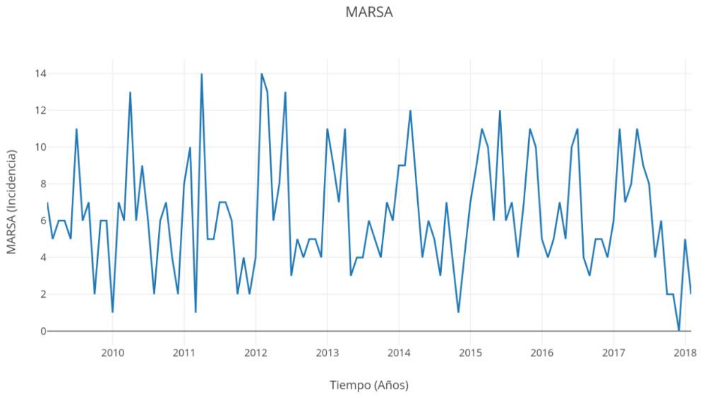
Donde z_{t-s} es el valor de la serie para el instante de tiempo $t-s$ y $\phi_s a_{t-s}$ es el coeficiente MA multiplicado por el valor de media móvil en el instante $t-s$.

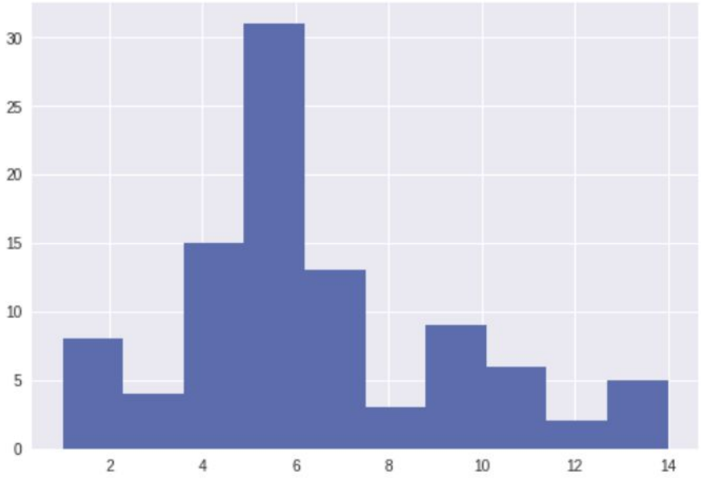
Estudio preliminar de una serie temporal

En primer lugar, realizaremos un estudio gráfico de la serie, viendo si hay patrones que se repitan a lo largo del tiempo, posibles tendencias o presencia de estacionalidad. También veremos si existen valores atípicos (outliers) en los datos. Estudiaremos si existen variaciones en la media y la varianza que hagan necesarias transformaciones matemáticas para estabilizar la serie de cara a la obtención de buenos modelos ARIMA.

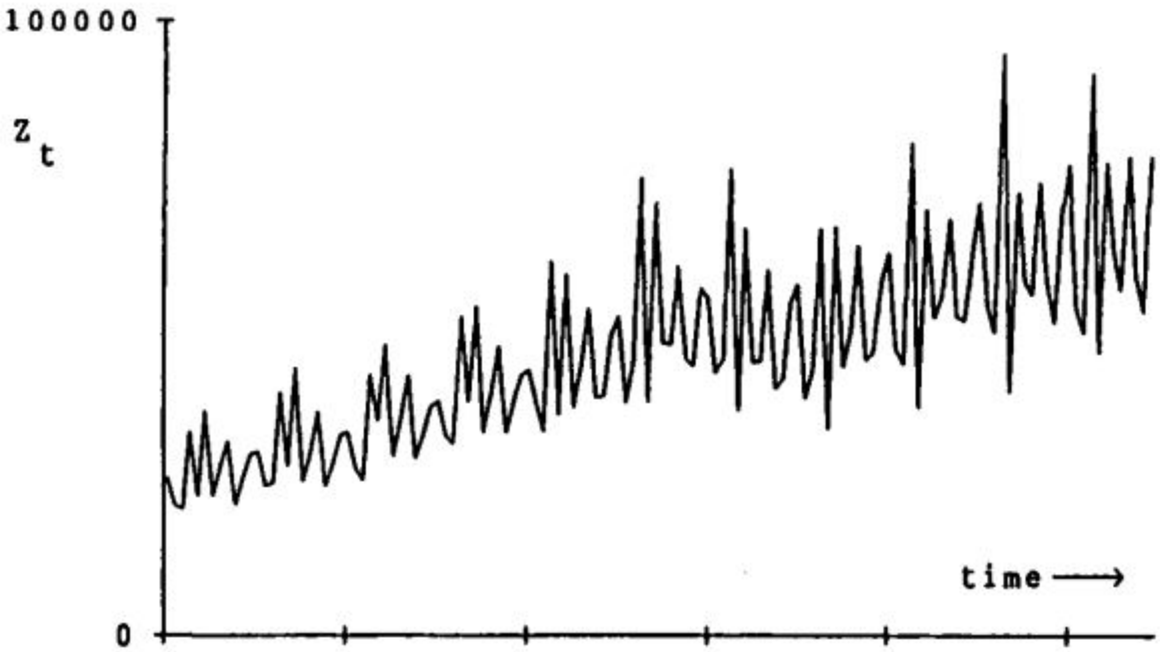
Aplicaremos una descomposición en las componentes principales y se representarán los residuos de la serie. Por último, haremos un estudio de las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) para ver posibles correlaciones presentes en los datos. Estas correlaciones, como veremos en posteriores secciones, podrán ser no estacionales o estacionales.

Estudio preliminar de una serie temporal





Estudio de estacionariedad de una serie



Media estacionaria

Para que una serie sea estacionaria (entendemos estacionariedad como estacionariedad débil) debe tener media, varianza y función de autocorrelación constante en el tiempo. Para ello haremos uso de la media móvil y de la desviación típica móvil, representando ambas para ver si es necesaria alguna transformación.

En caso de que la media no sea constante en el tiempo aplicaremos un tipo de transformación llamada *diferenciado regular*, que consiste en lo siguiente:

$$w_t = z_t - z_{t-1}$$

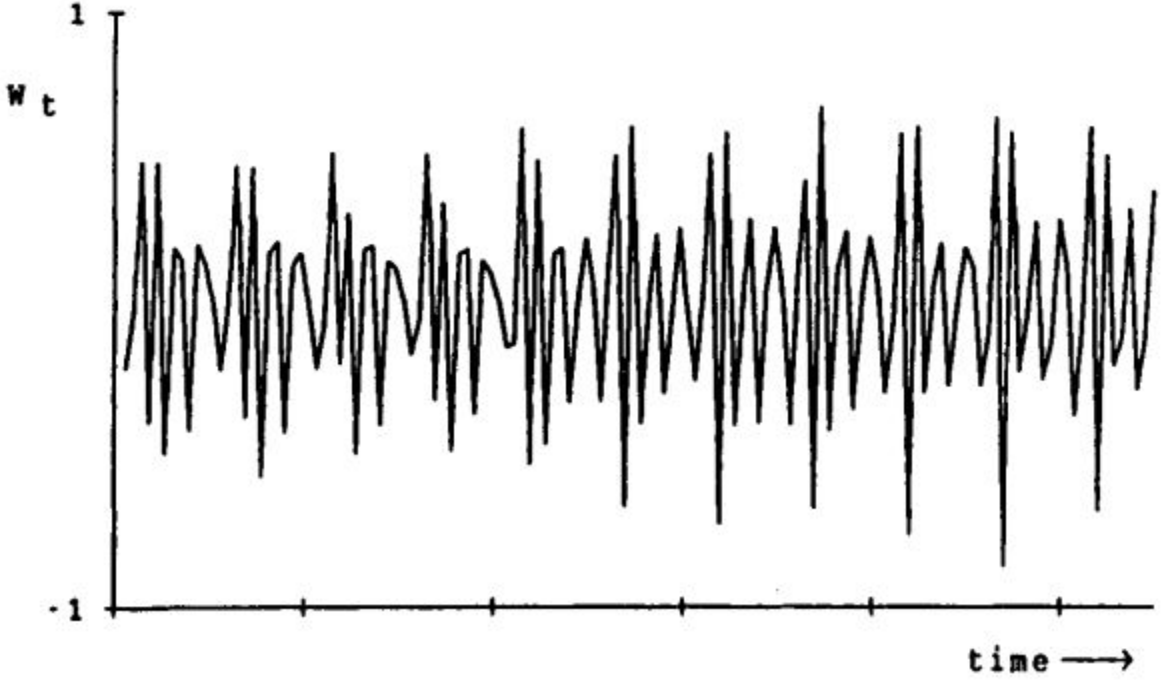
Media estacionaria

El nuevo valor en el instante de tiempo t se calcula como el valor original en dicho instante menos el valor anterior, por lo que si tenemos un aumento de la media a medida que avanzamos en el tiempo, esto hace que dicha subida se atenúe.
Podemos diferenciar tantas veces como necesitemos para obtener una media constante, pero no suele ser necesario diferenciar en más de dos niveles.

También podemos aplicar un diferenciado estacional. Como veremos en las pruebas realizadas, muchos de los modelos obtenidos requieren este tipo de operación, por lo que podremos saber, una vez acabado el proceso de selección, si necesitamos aplicar un diferenciado estacional para generar buenos modelos. El diferenciado estacional indica que algunos de los valores de la serie están variando de forma estacional o periódica.

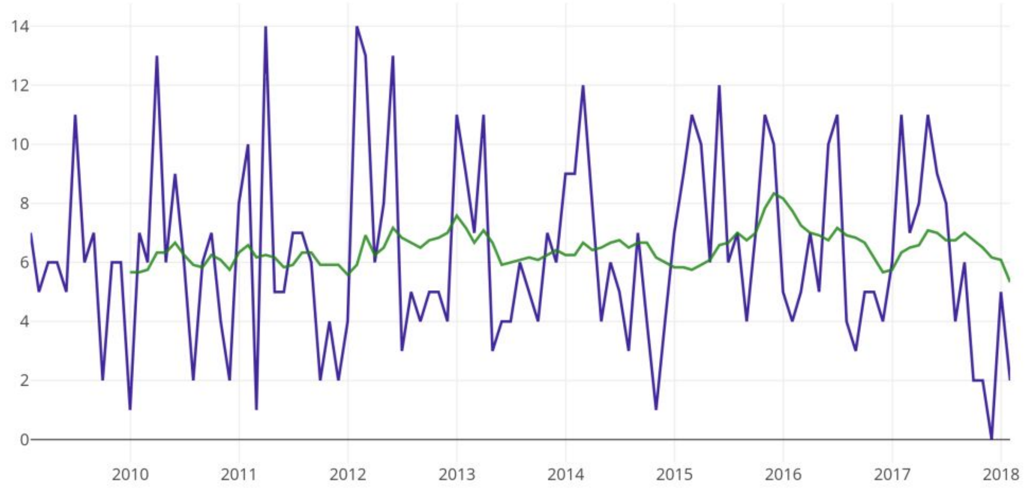
$$w_t = z_t - z_{t-s}$$

Media estacionaria

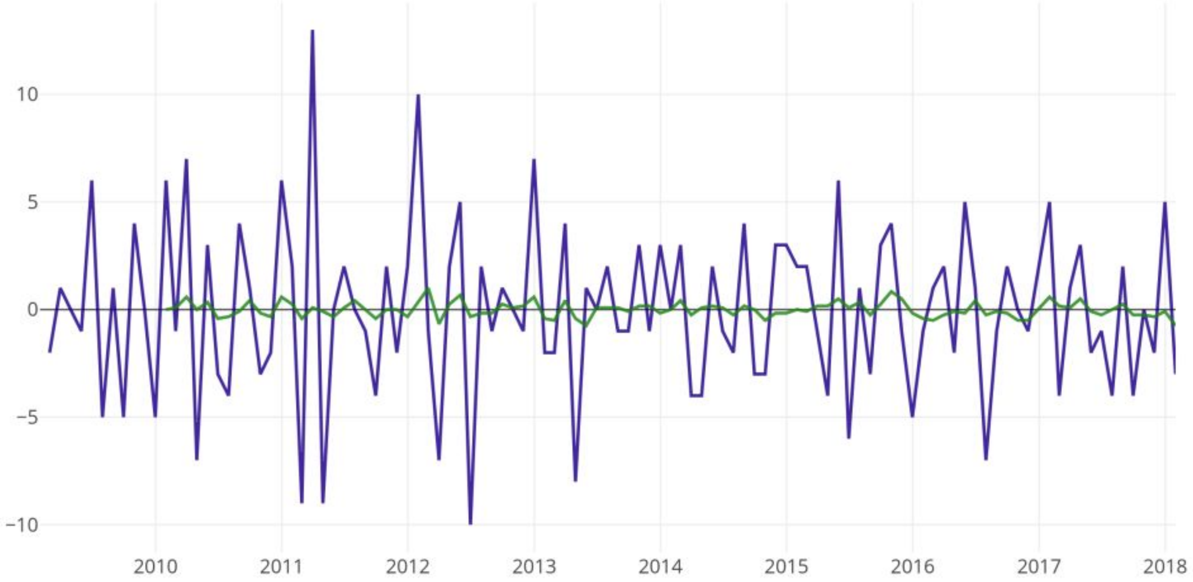


Cálculo de media móvil

Sirve para comprobar si la media de una serie es estacionaria. Para ello debemos indicar una ventana, es decir, un conjunto de datos previos que se van a utilizar para calcular el valor de la media en un punto concreto del tiempo.



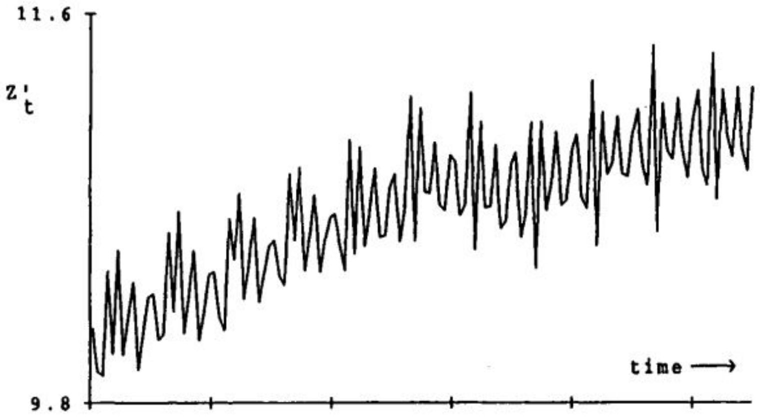
Cálculo de media móvil diferenciada d = 1



Varianza estacionaria

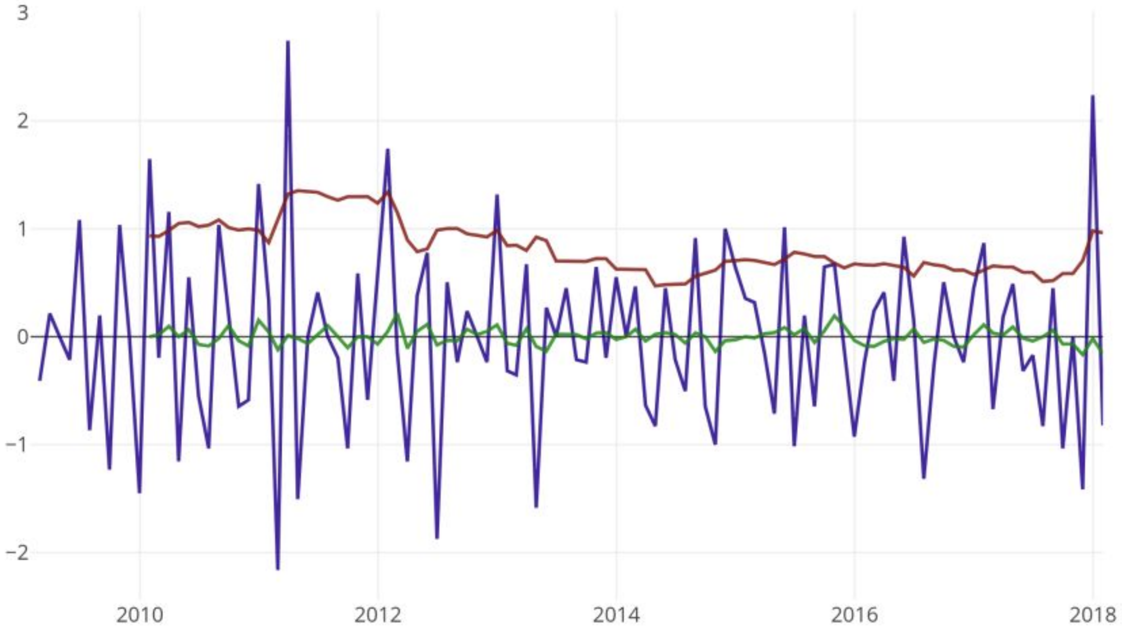
Para transformar la serie y obtener una varianza estacionaria debemos aplicar transformaciones de las llamadas “power transformations”. Algunas de estas son: log y raíz cuadrada.

$$z_t' = \frac{z_t^{\lambda} - 1}{\lambda}$$



Varianza estacionaria

Realizamos una transformación Box-Cox para estabilizar la varianza de la serie.



Test avanzado de Dickey-Fuller para estudio de la estacionariedad

Con lo anteriormente visto, podríamos necesitar hacer la primera diferencia de la serie y estabilizar la varianza aplicando una transformación Box-Cox. Con el Test de Dickey-Fuller se puede estudiar si existe la presencia de raíces unitarias, viendo si la serie es estacionaria de orden 0.

Consideramos como hipótesis nula H0: la serie no es estacionaria.

Aplicando adfuller, el p-value obtenido para dicha serie es de p = $3{,}30e^{-06}$, lo que indica que la serie es estacionaria al ser mucho menor que el valor crítico de 0.05.

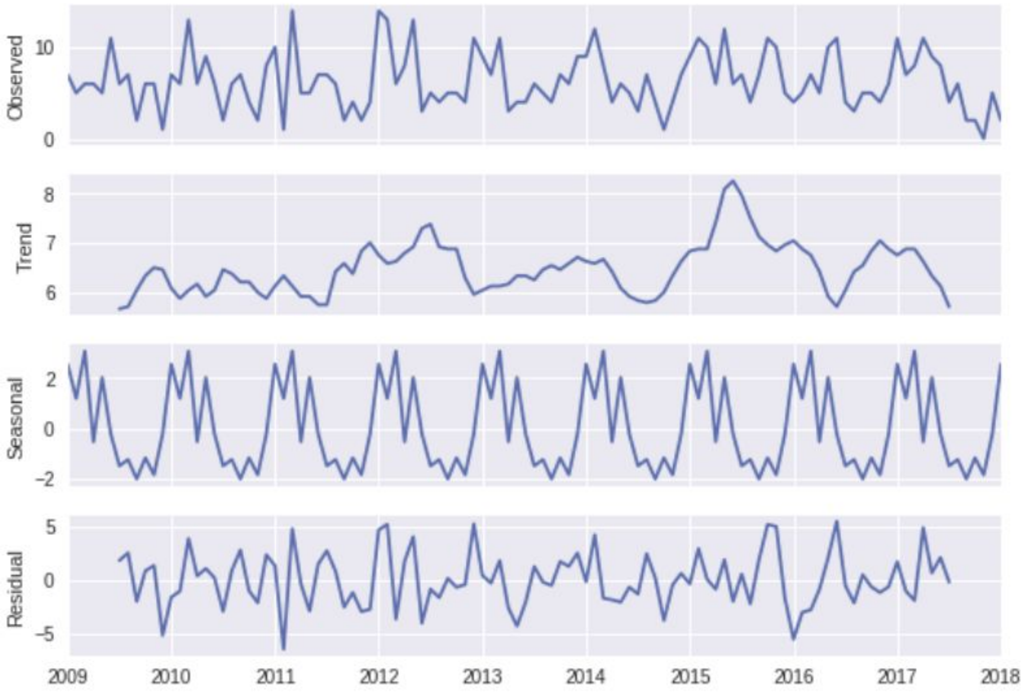
El test de Dickey-Fuller puede no detectar bien la estacionariedad en series cortas.

Para aplicar el test de Dickey-Fuller:

```
from statsmodels.tsa.stattools import adfuller

result = adfuller(series, autolag='AIC')
```

Descomposición de una serie



Suavizado exponencial

Un modelo de suavizado exponencial calcula los valores de la serie como una media de los valores anteriores, aplicando un peso a cada una de las observaciones.

En este tipo de modelos, el peso de las observaciones va descendiendo a medida que se va alejando en el tiempo.

Los modelos de suavizado exponencial son una alternativa a los modelos ARIMA.

Existen tres niveles de suavizado exponencial:

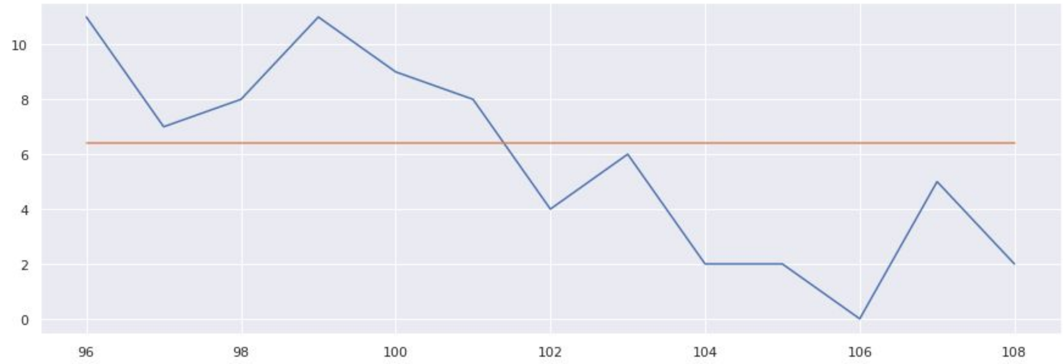
- **Simple**: Cuando tenemos una serie sin tendencia ni estacionalidad.
- **Doble**: Cuando contamos con series con tendencia (aditiva o multiplicativa).
- **Triple**: Contamos con tendencia(aditiva o multiplicativa) y estacionalidad (aditiva o multiplicativa).

Suavizado exponencial: Simple

Sigue la siguiente función:

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots,$$

Donde α es un parámetro de suavizado.



Suavizado exponencial: Simple

Hay dos casos extremos en el parámetro α:

- Si es == 0: Se predicen los valores futuros como la **media** de la serie.
- Si es == 1: Se predicen los valores futuros como el **último valor** obtenido en la serie.

Suavizado exponencial: Doble

Método de **Holt**. Mediante este método podemos predecir series temporales con tendencia.

$$
\begin{aligned}
\text{Forecast equation} \qquad & \hat{y}_{t+h|t} = \ell_t + hb_t \\
\text{Level equation} \qquad & \ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
\text{Trend equation} \qquad & b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1},
\end{aligned}
$$

Suavizado exponencial: Holt-Winters

Método utilizado para realizar forecasting en series temporales con tendencia y estacionalidad. Hay dos variaciones del método:

- **Método aditivo**:

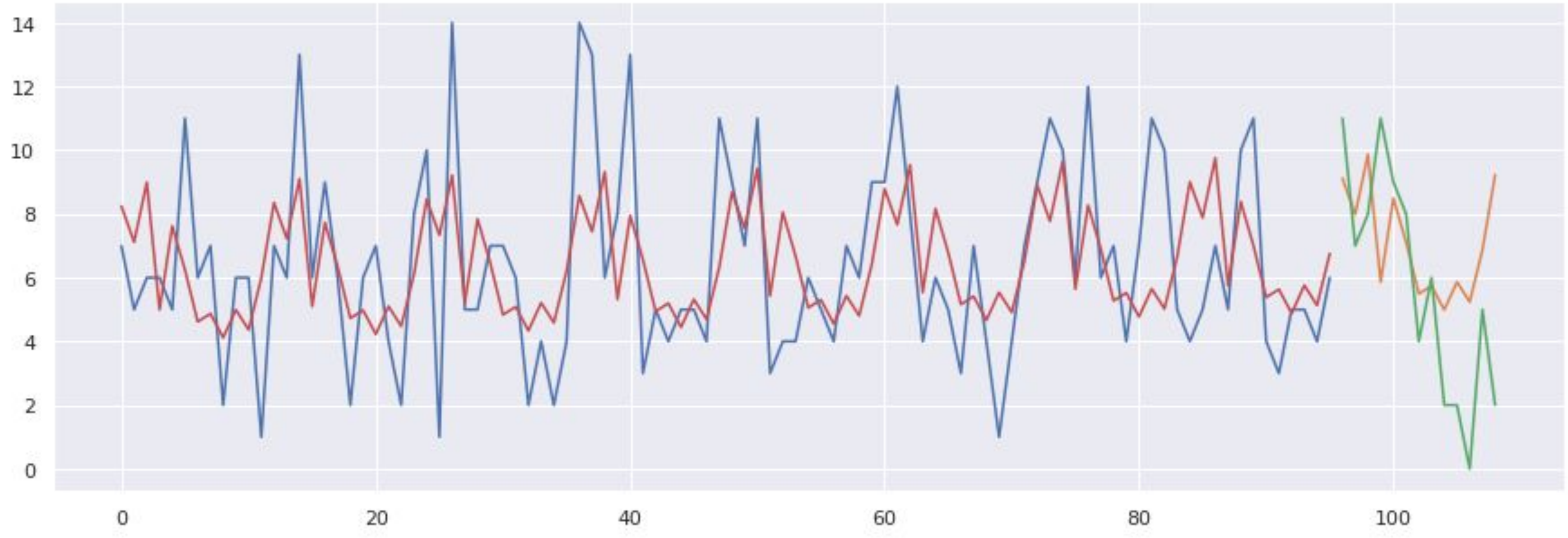
  y(t) = Nivel + Tendencia + Estacionalidad + Ruido

- **Método multiplicativo**:

  *y(t) = Nivel * Tendencia * Estacionalidad * Ruido*

En el método aditivo las variaciones debidas a la componente de estacionalidad son prácticamente constantes, mientras que, en el método multiplicativo cambian proporcionalmente al nivel de la serie.

Suavizado exponencial: Holt-Winters

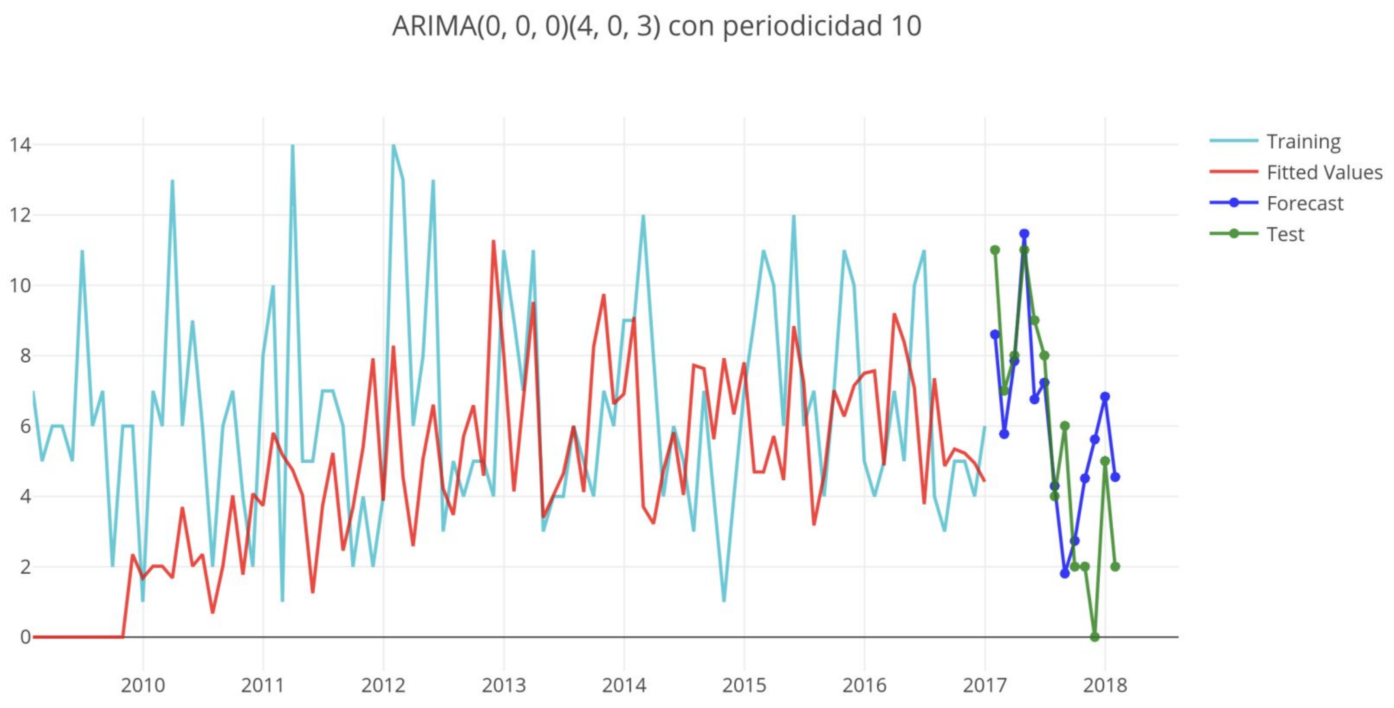


Modelos ARIMA

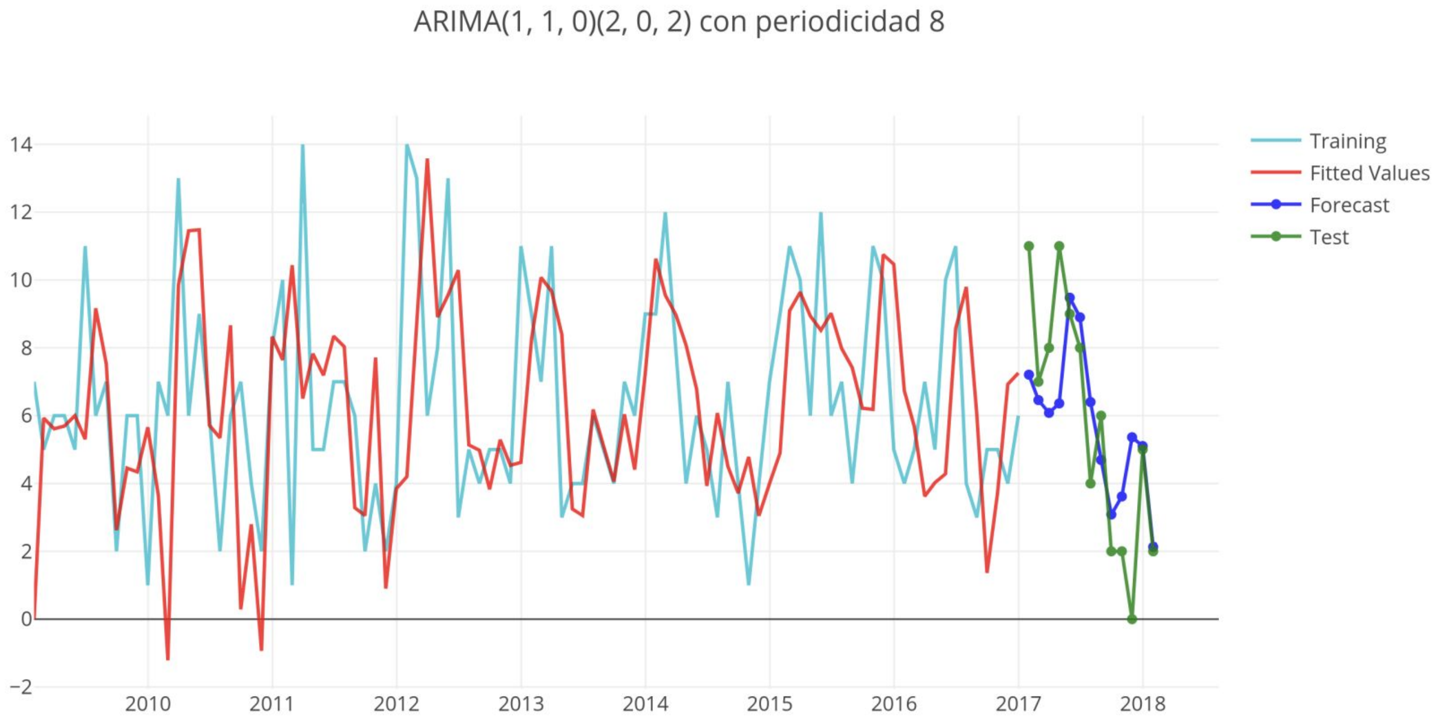
En un proceso **ARIMA o ARMA integrado**, se modela la dinámica de la serie diferenciada.

Es decir, se modela la diferencia de los dos últimos valores de la serie en función de las diferencias anteriores. El orden de un modelo ARIMA se expresa como una terna (p, d, q), donde p y q son los órdenes de los procesos AR y MA, respectivamente, y d es el orden de diferenciación.

Modelos ARIMA: Ejemplo gráfico



Modelos ARIMA: Ejemplo gráfico II



Estacionalidad en ARIMA

Los modelos ARIMA también pueden presentar **estacionalidad**.

Un modelo con estacionalidad básico y muy común en la práctica es el ARIMA(0, 1, 1)s , que tiene las mismas componentes no estacionales que el módelo no estacional ARIMA(0, 1, 1), pero la componente MA(1) no se detecta en el lag 1 como sucedería al ser no estacional, sino que se presenta en el lag s.

De igual manera ocurriría con un modelo que presente componente AR estacional.

Autocorrelación: ACF y PACF

Un paso necesario en la metodología Box-Jenkins es la representación de la función de autocorrelación (**ACF**) y la función de autocorrelación parcial (**PACF**).

La **autocorrelación** nos indica cómo se relaciona el actual valor de una serie z con sus valores pasados (z-1, z-2, z-3…).

Indica una medida de la **relación** (positiva o negativa) y de la **fuerza** de la relación entre observaciones de una serie temporal.

Para calcular la correlación existente entre dos instantes de tiempo concretos de la serie utilizamos el coeficiente de autocorrelación poblacional (**population autocorrelation coefficient**) para varios lags K. Donde K indica cada cuántos instantes estamos estableciendo dicha correlación.

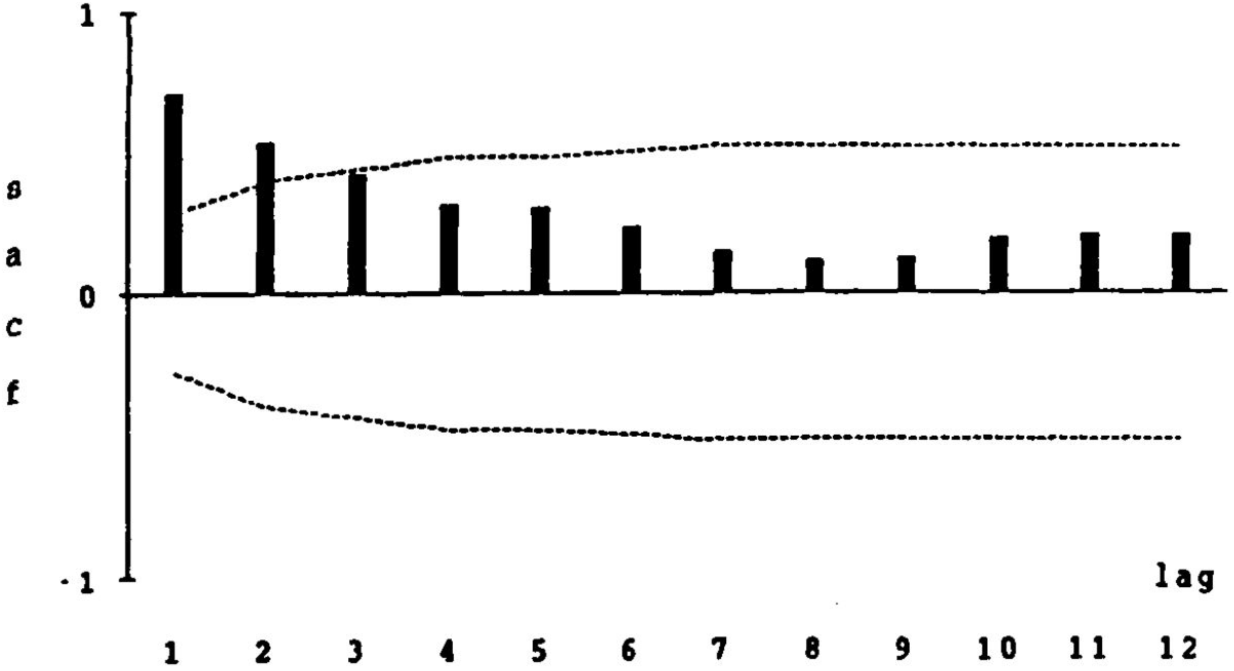
$$\rho_k = \mathrm{cov}(z_t, z_{t+k}) / \sigma_z^2$$

Autocorrelación: ACF y PACF

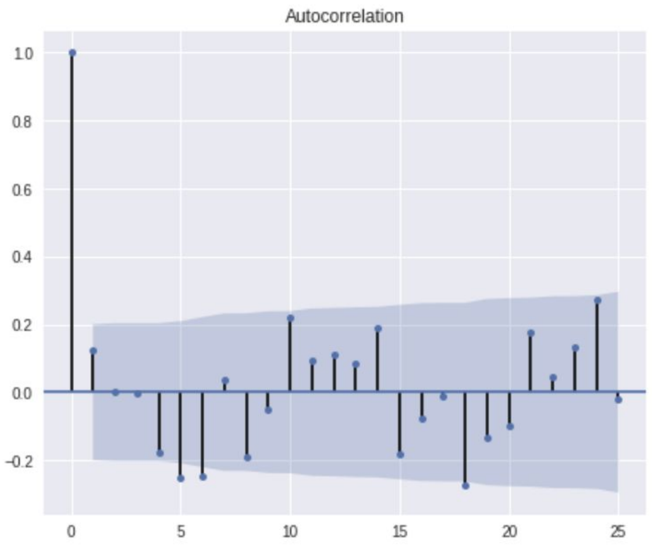
$$\rho_k = \mathrm{cov}(z_t, z_{t+k})/\sigma_z^2$$

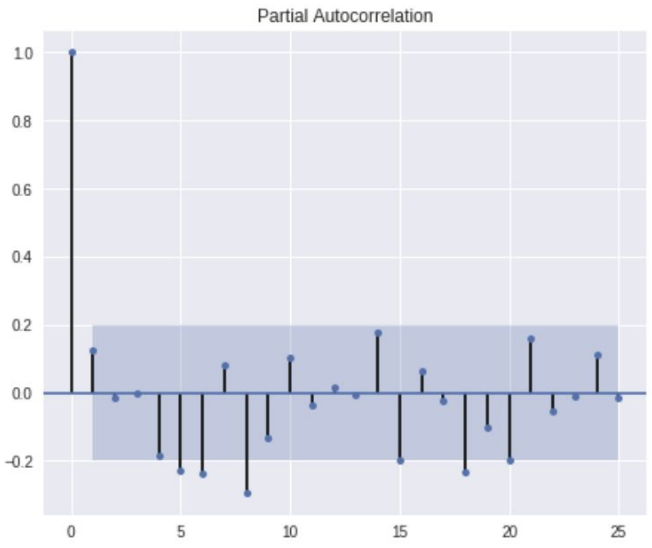
$$r_k = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^{n} (z_t - \bar{z})^2}$$

Autocorrelación: ACF y PACF

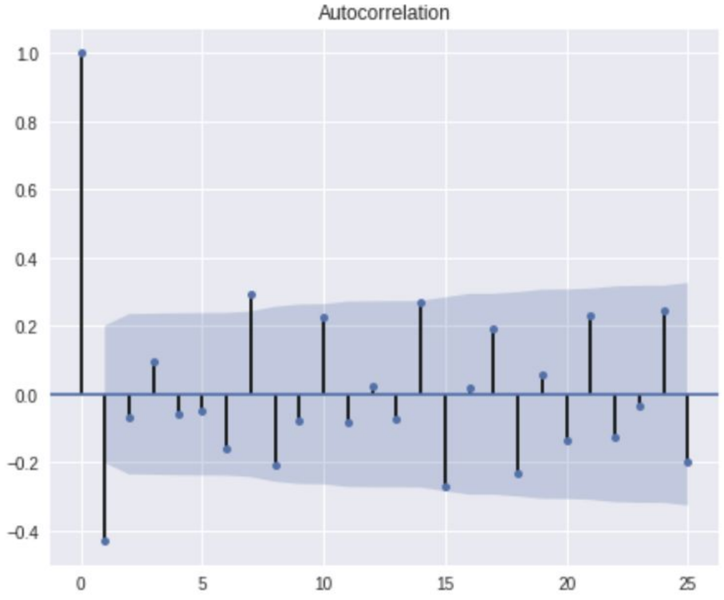


Ejemplos prácticos de ACF y PACF





Ejemplos prácticos de ACF y PACF



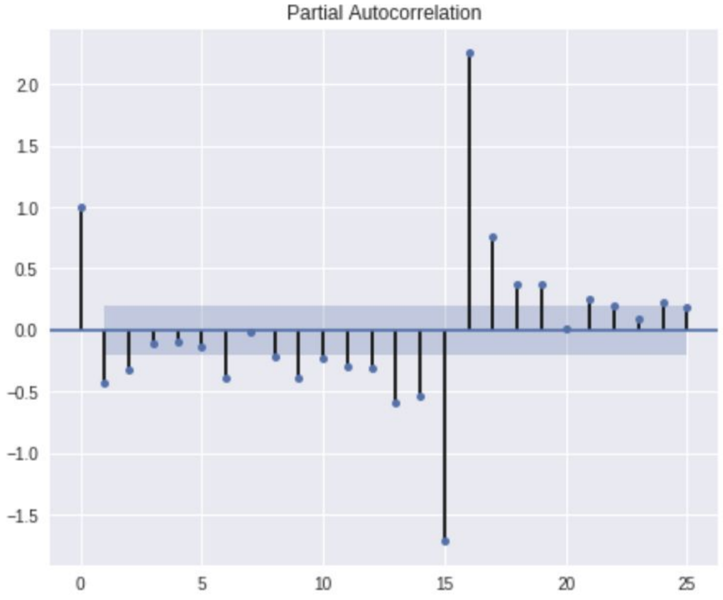


Tabla de ACF y PACF teóricos

	AR(p)	MA(q)	ARMA(p,q)
ACF	varios puntos con coef>0 decayendo	0 excepto los q primeros	varios puntos con coef>0 decayendo
PACF	0 excepto los p primeros	varios puntos con coef>0	varios puntos con coef>0

¿Cómo comprobamos la bondad de nuestro modelo?

RSME: Error cuadrático medio, tanto de la fase de entrenamiento como de la de test.

AIC: Criterio de información de Akaike.

Parsimonia: elegir modelos más generales.

AIC: Criterio de información de Akaike

Es una medida de la calidad **relativa** del modelo. Se obtiene a partir de la suma de los cuadrados de los residuos y de un factor que penaliza la inclusión de parámetros.

$$AIC = 2k – 2ln(L)$$

$$AIC = Nlog(\frac{SSE}{N}) + 2(k+2)$$

Donde N es el número de observaciones, SSE es la suma de los errores al cuadrado, y k es el número de predictores utilizados.

Principio de Parsimonia

Un criterio importante a la hora de elegir buenos modelos es que siga el principio de parsimonia.

Queremos elegir **modelos que expliquen bien los datos**, es decir, que tengan un buen ajuste de los datos de entrenamiento y a su vez que sean generales.

Preferimos estos **modelos más generales** ya que funcionan mejor en un rango más amplio de situaciones.

Principio de Parsimonia

Establecemos una fórmula muy sencilla para calcular la parsimonia de un modelo, en la que incluimos dos secciones, la sección sin estacionalidad y la sección con estacionalidad.

$$Parsimonia = p + d + q + P + D + Q$$

En la ecuación anterior, los valores de p, d y q son:

- p: factor de AR.
- d: diferenciación → Aplicación del proceso de diferenciación a la serie.
- q: factor de MA.
- P: factor de AR estacional.
- D: factor de diferenciación estacional.
- Q: factor de MA estacional.

Automatización del modelado ARIMA

- Criterio de capacidad de predicción usando el RMSE (Root Mean Square Error o Error Cuádrático Medio).
- Criterio de complejidad con los valores de AIC (Criterio de información de Akaike) de los modelos.
- Criterio basado en la autocorrelación del modelo usando las gráficas de ACF (Función de autocorrelación) y PACF (Función de autocorrelación parcial).
- Criterio de invertibilidad y estacionariedad.