



Análisis de Series de Tiempo

Carrera de Especialización en Inteligencia Artificial

Clase

Ing. Magdalena Bouza, Ing. Carlos German Carreño Romano

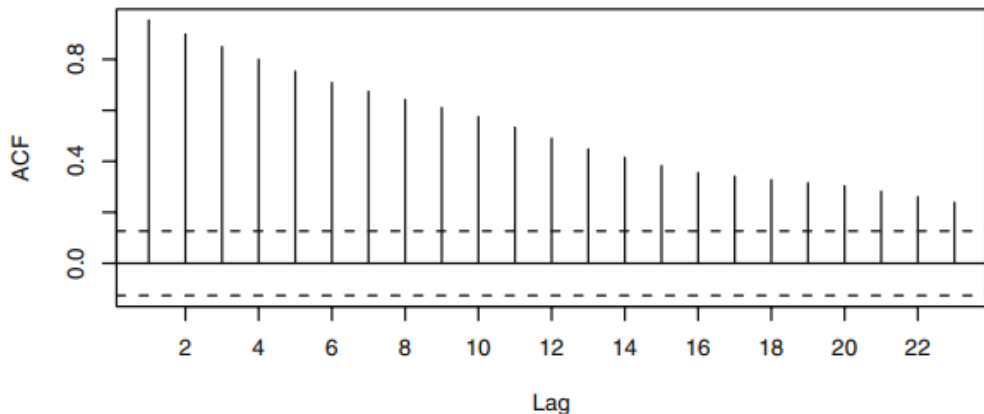
Agenda

1. Estacionariedad
2. Estacionalidad
3. Modelos ARIMA y SARIMA
4. Estimación del orden del modelo
5. Criterios de bondad de modelos

Estacionariedad

Autocorrelación

Un método relativamente fácil, aunque bastante a ojo, para verificar que una serie **no es estacionaria** es a través de su función de autocorrelación muestral.



Si la gráfica no alcanza valores nulos para lags grandes es porque posiblemente no sea estacionaria

Tests para determinar estacionariedad

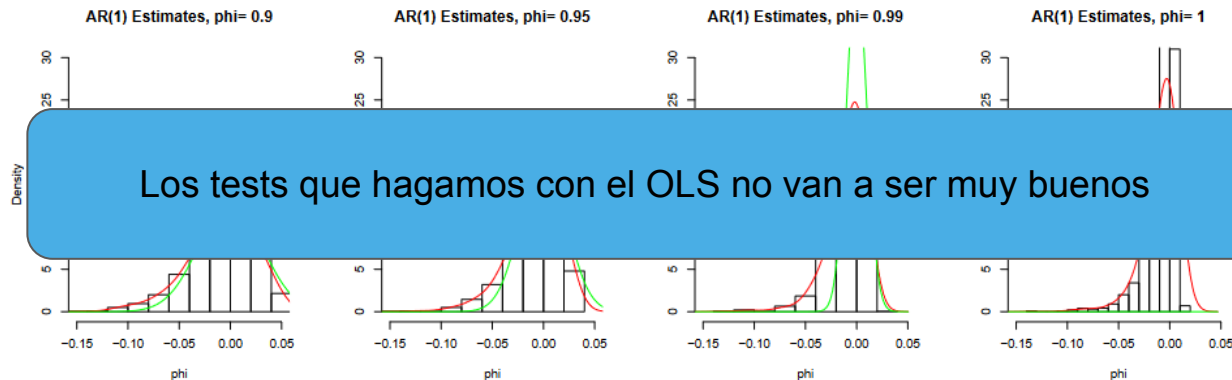
Tengo un modelo AR(1) de la forma $Y_t = a_1 Y_{t-1} + e_t$ y quiero saber si $a_1 = 1$.

Si desconocemos la estacionariedad del modelo, sólo podemos estimar el parámetro por OLS:

$$\hat{a}_1 = \frac{\sum_{t=1}^n Y_t Y_{t-1}}{\sum_{t=1}^n Y_{t-1}^2}$$

Se puede demostrar que $\sqrt{n}(\hat{a}_1 - a_1) \sim \mathcal{N}(0, 1 - a_1^2)$ si $|a_1| < 1$.

Cuando $a_1 \approx 1$, esta aproximación deja de ser válida.



Test de Dickey-Fuller

Propone el modelo $Y_t = aY_{t-1} + \boxed{X_t}$ ← Proceso estacionario

Si $a=1$ → El proceso es **no estacionario** (hay caminante aleatorio)

Si miramos la serie diferenciada una vez, tenemos que

$$Y_t - Y_{t-1} = aY_{t-1} + W_t - Y_{t-1} = (a - 1)Y_{t-1} + W_t$$

Dickey y Fuller proponen entonces el test

$$H_0 : (a - 1) = 0 \quad vs. \quad H_1 : (a - 1) \neq 0$$

Busco rechazar el test

Su gran aporte fue hallar la distribución asintótica de $n(\widehat{a} - 1)$ bajo H_0 .

Es un test para determinar si la serie (**a una diferenciación**), posee una componente de RW

Test de dickey-Fuller Aumentado

Incorpora al modelo un término de ruido dependiente (pero estacionario)

$$Y_t = aY_{t-1} + X_t \quad X_t = \sum_{j=1}^p \rho_j X_{t-j} + \boxed{w_t} \leftarrow \text{Ruido Blanco}$$

Si tomamos la primera diferencia, y observamos que bajo H_0

$$X_t = Y_t - Y_{t-1} \text{ tenemos que}$$

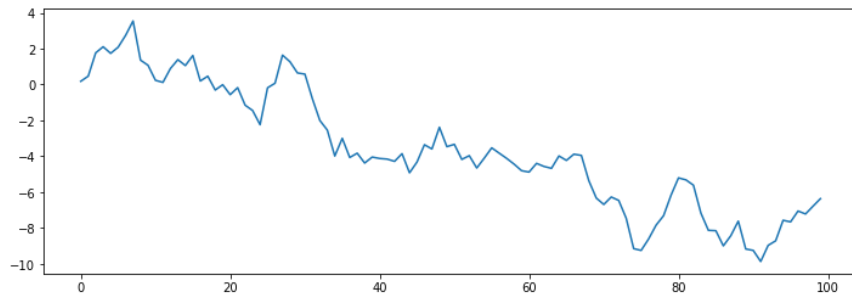
$$Y_t - Y_{t-1} = (a - 1)Y_{t-1} + \sum_{j=1}^p \rho_j (X_{t-j} - X_{t-j-1}) + W_t$$

Nuevamente se definen las hipótesis:

$$H_0 : (a - 1) = 0 \quad vs. \quad H_1 : (a - 1) \neq 0$$

p habría que estimarlo, pero podemos dejar que se encargue el software.

¿Cómo implementar en statsmodels?



```
from statsmodels.tsa.stattools import adfuller
adfuller(y)
```

✓ 0.9s

```
(-1.394757894883641, ← Estadístico
0.584786653359185, ← p-valor
4, ← # lags usados
95, ← # observaciones usadas
{'1%': -3.5011373281819504,
 '5%': -2.8924800524857854, ← umbrales
 '10%': -2.5832749307479226},
```

regression : {"c","ct","ctt","n"}

Constant and trend order to include in regression.

- "c" : constant only (default).
- "ct" : constant and trend.
- "ctt" : constant, and linear and quadratic trend.
- "n" : no constant, no trend.

Ejemplo con datos reales

```
inputfile = "../Datasets/TEC02.2000.2021.csv"
```

Scripts/Dickey-Fuller_dataset_example.ipynb

```
ts = pd.read_csv(inputfile, header=0, index_col=0, squeeze=True)
ts.fechaHora = pd.to_datetime(ts.fechaHora)
ts.fechaHora=pd.to_datetime(ts.fechaHora).dt.date
ts.fechaHora=pd.DatetimeIndex(ts.fechaHora)
ts=ts.sort_index(ascending=False)
```

```
adfuller(ts.ultimoPrecio)
```

```
(0.57890231841091,
 0.9870838971395222,
 32,
 4807,
 {'1%': -3.431711097447145,
  '5%': -2.862141452575749,
  '10%': -2.5670901548483767},
```



Ejemplo con datos reales

```
log_ultimoPrecio = np.log(ts.ultimoPrecio)
```

```
adfuller(log_ultimoPrecio, regression='ct')
```

```
(-3.037699968528972,  
 0.1218440464079108,  
 2,  
 4837,  
 {'1%': -3.9606428515465306,  
  '5%': -3.4113980566548285,  
  '10%': -3.127584714163724},
```

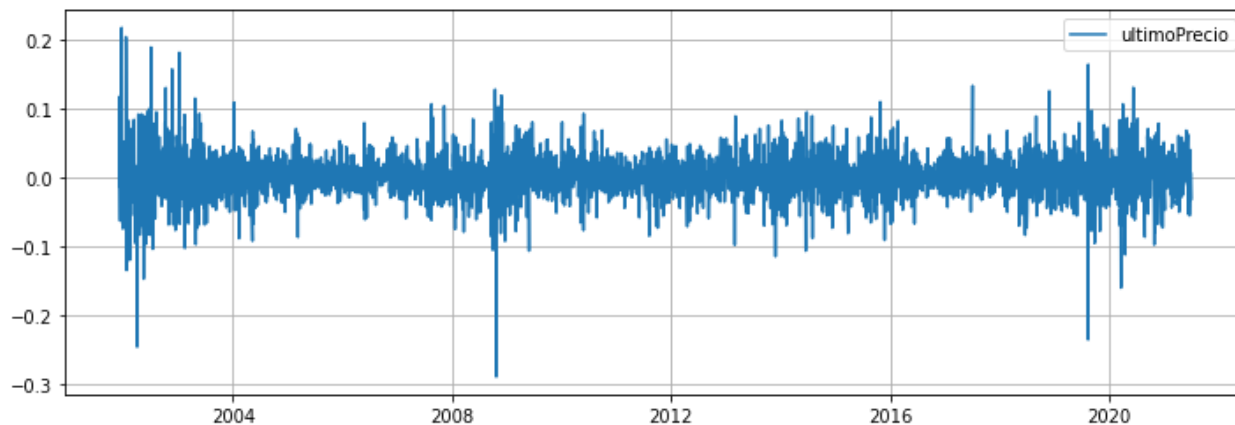


Ejemplo con datos reales

```
diff_log_ultimoPrecio = log_ultimoPrecio - log_ultimoPrecio.shift(1)
```

```
adfuller(diff_log_ultimoPrecio.dropna())
```

```
(-46.783753322231384,  
 0.0,  
 1,  
 4837,  
 {'1%': -3.4317026511738518,  
  '5%': -2.862137721117907,  
  '10%': -2.567088168437432},  
 -20357.65641243886)
```



Test de Kwiatkowski–Phillips–Schmidt–Shin (KPSS)

Es un test para determinar si la serie es **estacionaria** alrededor de **una tendencia determinística**.

Plantea que la serie se puede descomponer como

$$Y_t = \boxed{W_t} + X_t + \boxed{\mu_t}$$

Caminante aleatorio Tendencia determinística

Para luego plantear las hipótesis:

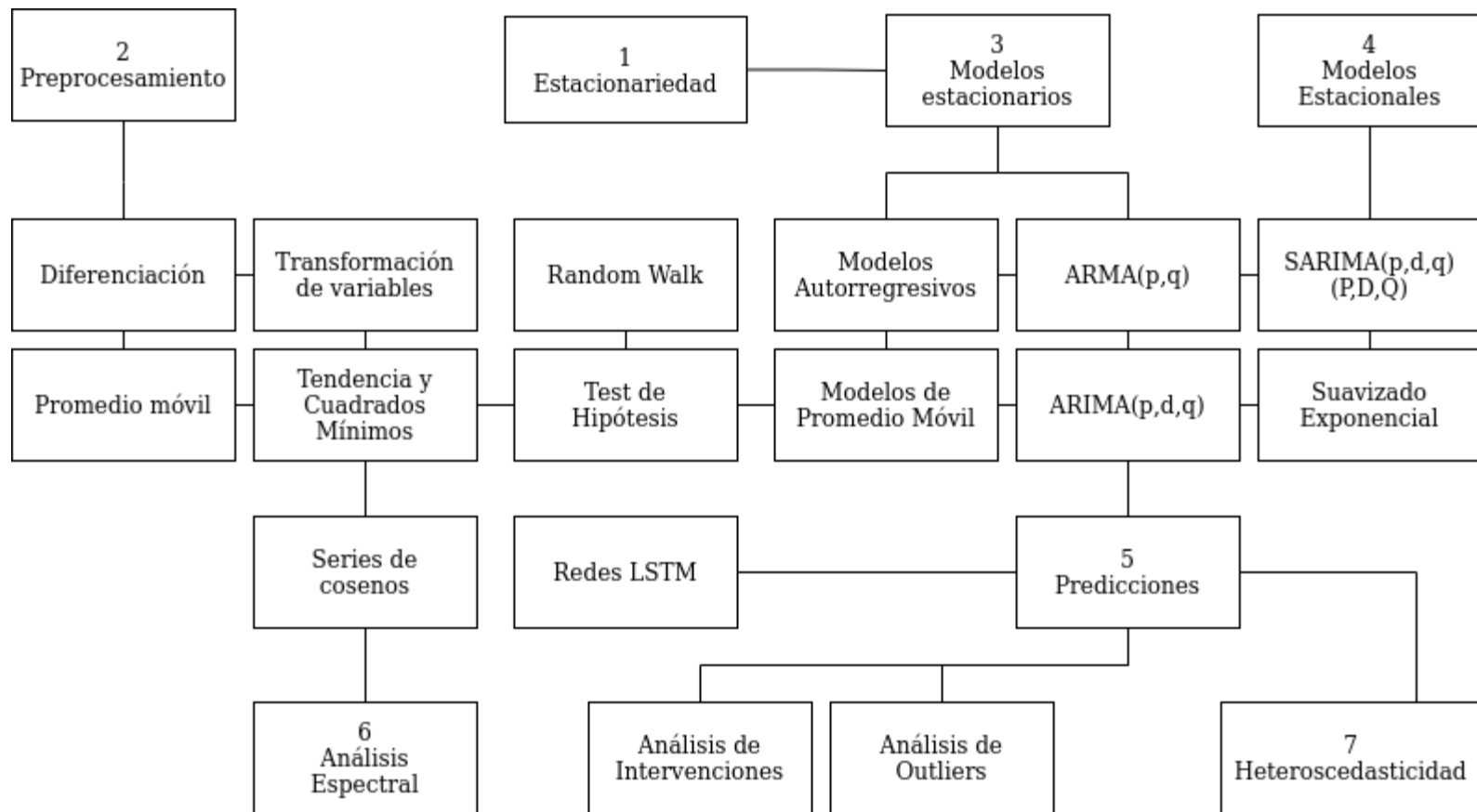
$$H_0 : var(W_t) = 0 \quad H_1 : var(W_t) > 0$$

Busco **no**
rechazar el test

Bibliografía extra para D-F y KPSS

- [Testing for unit roots](#)
- [Augmented Dickey-Fuller root tests](#)
- [Statsmodels: Stationarity and detrending](#)

ARIMA(p,d,q)



Operadores

- Operador de Backshift (B), Operador Nabla
- Polinomio característico
- AR, MA compactos
- SARIMA
- Ejemplo

$$B^K y_t = y_{t-K}$$

∇ : Nabla

$$\nabla y_t = y_t - y_{t-1}$$

$$\nabla^2 y_t = \nabla \nabla y_t$$

$$= \nabla (y_t - y_{t-1})$$

$$= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$$

$$= y_t - 2y_{t-1} + y_{t-2}$$

$$= \nabla (1-B) y_t$$

$$= y_t - B y_t$$

$$= y_t (1-B)$$

$$= (1-B)(1-B)y_t = (1-B)^2 y_t$$

$$[B^k Y_t = Y_{t-k}]$$

$$\nabla^2 Y_t = (1-B)^2 Y_t$$

\vdots

$$\rightarrow \nabla^k Y_t = (1-B)^k Y_t$$

AR(p): $Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_p Y_{t-p}$

MA(q): $Y_t = \epsilon_t - b_1 \epsilon_{t-1} - b_2 \epsilon_{t-2} - \dots - b_q \epsilon_{t-q}$

ARMA(p, q)

$$a_1 Y_{t-1} + \dots + a_p Y_{t-p} = \sum_{k=1}^p a_k Y_{t-k} = \sum_{k=1}^p a_k B^k Y_t$$

$$b_1 \epsilon_{t-1} + \dots + b_q \epsilon_{t-q} = \sum_{k=1}^q b_k \epsilon_{t-k} = \sum_{k=1}^q b_k B^k \epsilon_t$$

$$\nabla Y_t = Y_t - Y_{t-1}$$

$$\nabla_{\text{seasonal}} Y_t = Y_t - Y_{t-s}$$

$$\epsilon_t = Y_t + b_1 \epsilon_{t-1} + \dots$$

$$\text{ARMA}(p, q) = \text{AR}(p) + \text{MA}(q)$$

$$\rightarrow \underline{y_t} = \underbrace{a_1 \underline{y_{t-1}} + \dots + a_p \underline{y_{t-p}}}_{\text{AR}(p)} + \underline{e_t}$$

$$\rightarrow \underline{y_t} = \underline{e_t} - \underbrace{b_1 \underline{e_{t-1}} - \dots - b_q \underline{e_{t-q}}}_{\text{MA}(q)}$$

$$\rightarrow \underline{y_t} = a_1 \underline{y_{t-1}} + \dots + a_p \underline{y_{t-p}} + \underbrace{\underline{e_t} - b_1 \underline{e_{t-1}} - \dots - b_q \underline{e_{t-q}}}_{\text{ARMA}(p, q)}$$

①

$$\underline{y_t - a_1 y_{t-1} - \dots - a_p y_{t-p}} = e_t - b_1 e_{t-1} - \dots - b_q e_{t-q}$$

ARMA

$$y_t (1 - a_1 B - a_2 B^2 - \dots - a_p B^p) = e_t (1 - b_1 B - \dots - b_q B^q)$$

②

$$y_t \left(1 - \sum_{k=1}^p a_k \cdot B^k \right) = e_t \left(1 - \sum_{k=1}^q b_k B^k \right)$$

3

$$\gamma_t (1 - \sum_{k=1}^p a_k B^k) = \epsilon_t (1 - \sum_{k=1}^q b_k B^k)$$

polinomio característico: ω

$$a(x) = 1 - a_1 x - a_2 x^2 - \dots - a_n x^n$$

$$b(x) = 1 - b_1 x -$$

$$\phi(x) =$$

$$\theta(x) =$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

3

$$\left. \begin{aligned} \gamma_t (1 - \sum_{k=1}^p a_k B^k) &= \phi(B) \gamma_t \\ \epsilon_t (1 - \sum_{k=1}^q b_k B^k) &= \theta(B) \epsilon_t \end{aligned} \right\}$$

$$\phi(B) \gamma_t = \theta(B) \epsilon_t$$

ARMA(p, q)

Backshift

$$BY_t = Y_{t-1}$$

$$B^2Y_t = Y_{t-2}$$

...

$$B^kY_t = Y_{t-k}$$

$$\nabla Y_t = Y_t - Y_{t-1}$$

$$\begin{aligned}\nabla^2 Y_t &= \nabla(\nabla Y_t) \\ &= \nabla(Y_t - Y_{t-1})\end{aligned}$$

$$B^k Y_t = B_k Y_t = Y_{t-k}$$

$$\begin{aligned}B_s^k Y_t &= B^{ks} Y_t = Y_{t-ks} \\ &= Y_{t-ks}\end{aligned}$$

$$B_{12}^k Y_t = Y_{t-k \cdot 12}$$

AR, MA + Operador de Backshift (B)

MA(q):

$$\begin{aligned}Y_t &= e_t - b_1 e_{t-1} - \dots - b_q e_{t-q} \\&= e_t - b_1 B e_t - \dots b_q B^q e_t \\&= (1 - b_1 B - \dots - b_q B^q) e_t\end{aligned}$$

- **Proceso Lineal General:** es útil para encontrar propiedades generales, por ejemplo para condicionar los coeficientes del modelo

AR(p)

$$\begin{aligned}Y_t &= e_t + a_1 Y_{t-1} + \dots + a_p Y_{t-p} \\Y_t - a_1 Y_{t-1} - \dots - a_p Y_{t-p} &= e_t \\Y_t - a_1 B Y_t - \dots - a_p B^p Y_t &= e_t \\(1 - a_1 B - \dots - a_p B^p) Y_t &= e_t\end{aligned}$$

- **Polinomio característico:** es útil para expresar relaciones analíticas entre las raíces del polinomio y las expresiones de autocorrelación y autocovarianza del modelo.

AR, MA usando Operadores

AR(p)

$$e_t = Y_t - \sum_{k=1}^p a_k Y_{t-k}$$

①

$$e_t = (1 - \sum_{k=1}^p a_k B^k) Y_t$$

②

$$e_t = \phi(B) Y_t$$

③

MA(q)

$$Y_t = e_t - \sum_{k=1}^p b_k e_{t-k}$$

$$Y_t = (1 - \sum_{k=1}^p b_k B^k) e_t$$

$$Y_t = \theta(B) e_t$$

Modelo ARIMA

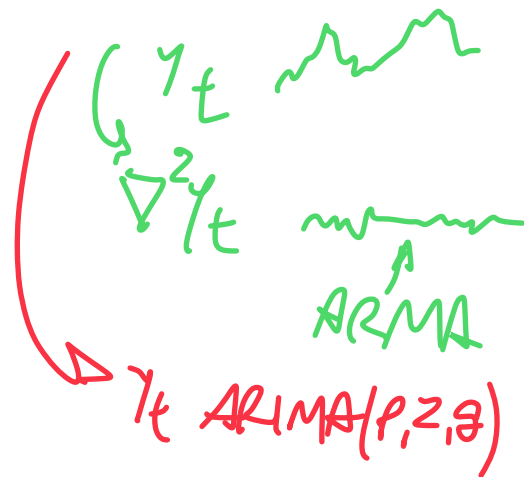
Diremos que $\{Y_t\}$ sigue un modelo ARIMA si $W_t = \nabla^d Y_t$ es un proceso ARMA(p,q) estacionario. Diremos en este caso que Y_t sigue un proceso ARIMA(p,d,q).

Operador de *backshift*: el operador de backshift (B) opera sobre el índice temporal de la serie de tiempo y la desplaza en una unidad de tiempo:

$$BY_t = Y_{t-1}$$

El operador B es lineal: $B(aY_t + bX_t + c) = aY_{t-1} + bX_{t-1} + c$

En general escribimos que $B^d Y_t = Y_{t-d}$



ARMA

$$\phi(B) Y_t = \theta(B) e_t$$

ARIMA: $\nabla^d Y_t = W_t$

or ARIMA

$$(1-B)^d Y_t$$

→ ARMA

$$\phi(B) \nabla^d Y_t = \theta(B) e_t$$

$$\phi(B) (1-B)^d Y_t = \theta(B) e_t$$

ARIMA(p,d,q)

ARMA

$$\begin{pmatrix} 1 & -2 & \dots & -q_1 \\ 1 & -a_1 \end{pmatrix}$$

$$W_t = a_1 W_{t-1} + \dots + e_t - b_1 e_{t-1} - \dots$$

$$(1-B)^d$$

$$\nabla^d$$

$$\nabla^d y_t = w_t$$

$$y_t - y_{t-1} = w_t$$

$$\text{ARMA } y_t = a_1 y_{t-1} + \dots + a_p y_{t-p} + e_t - b_1 e_{t-1} - \dots$$

$$y_t - y_{t-1} =$$

$$y_t - y_{t-1} = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + e_t - b_1 e_{t-1} - \dots$$

$$y_t = \frac{1}{(a_1 - 1)} y_{t-1} + \dots + a_p y_{t-p} + e_t = b_1 e_{t-1} - \dots - b_q e_{t-q} - \frac{(1 - b_1)}{(1 - 2b_1 + b_1^2)} e_t$$

AR, MA + operador Backshift

MA(q):

$$\begin{aligned}Y_t &= e_t - b_1 e_{t-1} - \dots - b_q e_{t-q} \\&= e_t - b_1 B e_t - \dots b_q B^q e_t \Rightarrow Y_t = b(B) e_t \\&= (1 - b_1 B - \dots - b_q B^q) e_t\end{aligned}$$

ARMA(p,q)

$$a(B)Y_t = b(B)e_t$$

AR(p)

$$\begin{aligned}Y_t &= e_t + a_1 Y_{t-1} + \dots + a_p Y_{t-p} \\Y_t - a_1 Y_{t-1} - \dots - a_p Y_{t-p} &= e_t \\Y_t - a_1 B Y_t - \dots - a_p B^p Y_t &= e_t \Rightarrow a(B) Y_t = e_t \\(1 - a_1 B - \dots - a_p B^p) Y_t &= e_t\end{aligned}$$

ARIMA(p,d,q)

$$a(B)(1 - B)^d Y_t = b(B)e_t$$

ARMA, ARIMA + operator Backshift

ARMA(p,q)

$$Y_t = \theta(B)e_t$$

$$e_t = \phi(B)Y_t$$

$$\phi(B)Y_t = \theta(B)e_t$$

$$a(B)Y_t = b(B)e_t$$

↙

}

ARIMA(p,d,q)

$$\rightarrow \nabla^d Y_t = W_t$$

$$\uparrow W_t : ARMA(p, q)$$

$$a(B)(1 - B)^d Y_t = b(B)e_t$$

Ejemplos

https://github.com/charlieromano/TimeSeries/blob/main/Scripts/predicciones_AR_MA.py

Statsmodels

[https://www.statsmodels.org/devel/generated/
statsmodels.tsa.arima.model.ARIMA.html](https://www.statsmodels.org/devel/generated/statsmodels.tsa.arima.model.ARIMA.html)

Estacionalidad

$$\nabla^d y_t = w_t$$

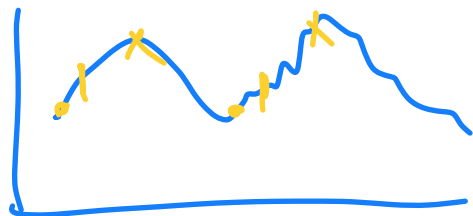
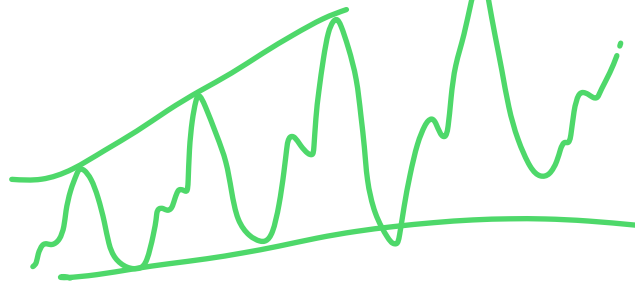
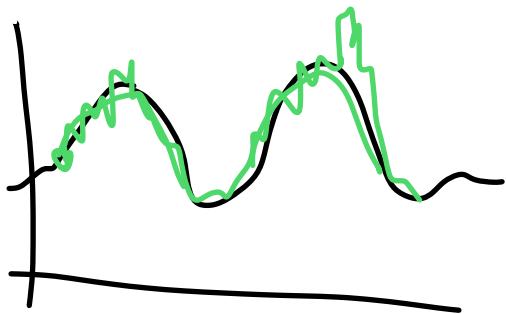
$$y_t - y_{t-1}$$

$$\phi(B) w_t = \theta(B) e_t \quad \text{ARMA}$$

$$w_t = a_1 w_{t-1} + \dots + a_p w_{t-p} + e_t - b_1 e_{t-1} - \dots$$

$$\phi(B) (w_t) = \theta(B) e_t \quad (\text{if } w_t \text{ is ARMA})$$

$$\nabla^d y_t = (1-B)^d y_t$$



$$S_t = 0.2 S_{t-12} + \epsilon_t$$

$\uparrow 0.2$

$$Y_t = \underbrace{\mu_t}_{\text{trend}} + \underbrace{X_t}_{\text{seasonality}}$$

AR(1)

$$Y_t = 0.3 Y_{t-1} + \epsilon_t$$

$$Y_t = S_t + \epsilon_t$$

$\epsilon \neq \epsilon$

$$S_t = S_{t-12} + \epsilon_t$$

$s=12$

$$(1 - B_{12})(1 - B)Y_t = \theta \varepsilon_t \quad \checkmark$$

$$(1 - B_{12} - B + B_{13})Y_t = \underbrace{Y_t - Y_{t-1}}_{AR(p=1)} - \underbrace{Y_{t-12} + Y_{t-13}}_{S=12}$$

$$\rightarrow \nabla \nabla_s Y_t = Y_t - Y_{t-s}$$

$$Y_t = S_t + \varepsilon_t$$

mode de une
saison
cambié espace

$$S_t = S_{t-1} + \varepsilon_t$$

SARIMA : y_t es multiplicativo seasonal
 parameters no seasonal (p, d, q)
 y orderes (P, D, Q)
 y período estacional "s"

$$\nabla^d \nabla_s^D y_t = w_t$$

w_t tiene que ser $\boxed{\text{ARMA}(p, q) \text{ (P, Q)}}$

SARMA.

$$\nabla^d y_t = w_t$$

ARMA
 (p, d, q)

Modelo SARMA multiplicativo

Cuando el proceso es estacionario podemos definir un modelo SARMA multiplicativo **ARMA(p,q)x(P,Q)s**, con período estacional **s** como un modelo AR con polinomio característico $a(x)\alpha(x)$ más un MA con polinomio característico $b(x)\beta(x)$. Con

$$a(x) = 1 - a_1x - a_2x^2 - \dots - a_px^p$$

$$\alpha(x) = 1 - \alpha_1x^s - \alpha_2x^{2s} - \dots - \alpha_Px^{Ps},$$

$$b(x) = 1 - b_1x - b_2x^2 - \dots - b_qx^q$$

$$\beta(x) = 1 - \beta_1x^s - \beta_2x^{2s} - \dots - \beta_Px^{Ps},$$

Modelo SARIMA multiplicativo

In general, then, we define a **multiplicative seasonal ARMA(p, q) \times (P, Q) $_s$ model with seasonal period s** as a model with AR characteristic polynomial $\phi(x)\Phi(x)$ and MA characteristic polynomial $\theta(x)\Theta(x)$, where

$$\left. \begin{aligned} \phi(x) &= 1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p \\ \Phi(x) &= 1 - \Phi_1 x^s - \Phi_2 x^{2s} - \dots - \Phi_P x^{Ps} \end{aligned} \right\}$$

and

$$\left. \begin{aligned} \theta(x) &= 1 - \theta_1 x - \theta_2 x^2 - \dots - \theta_q x^q \\ \Theta(x) &= 1 - \Theta_1 x^s - \Theta_2 x^{2s} - \dots - \Theta_Q x^{Qs} \end{aligned} \right\}$$

Notación equivalente

$$a(B) \equiv \phi(B)$$

$$b(B) \equiv \theta(B)$$

Usamos ϕ y θ para
evitar notar $B(B)$.

The model may also contain a constant term θ_0 . Note once more that we have just a special ARMA model with AR order $p + Ps$ and MA order $q + Qs$, but the coefficients are not completely general, being determined by only $p + P + q + Q$ coefficients. If $s = 12$, $p + P + q + Q$ will be considerably smaller than $p + Ps + q + Qs$ and will allow a much more parsimonious model.

Modelo SARIMA multiplicativo

Una herramienta importante es el análisis de procesos estacionales **no estacionarios** es la diferenciación estacional de período s para la serie $\{Y_t\}$, denotada

$$\nabla_s Y_t = Y_t - Y_{t-s}$$

Se dice que una serie estacional no estacionaria sigue un modelo SARIMA(p,d,q)x(P,D,Q)_s de período s si la serie diferenciada

sigue un proceso SARMA(p,q)x(P,Q)_s.

$$W_t = \nabla^d \nabla_s^D Y_t$$

$$\nabla_s^d Y_t = (1 - B_s)^d Y_t$$

Resumen de modelos

Resumen de los modelos

ARMA(p,q)

$$\phi(B)Y_t = \theta(B)e_t$$

ARIMA(p,d,q)

$$\phi(B)(1 - B)^d Y_t = \theta(B)e_t$$

SARMA(p,q)(P,Q)_s

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)e_t$$

Recordar que $\nabla_s^d Y_t = (1 - B_s)^d Y_t$

SARIMA(p,d,q)(P,D,Q)_s

$$\phi(B)\Phi(B^s)((1 - B)^d(1 - B^s)^D Y_t) = \theta(B)\Theta(B^s)e_t$$

Modelo SARIMA multiplicativo

SAR(p)(P)s como AR(p) multiplicativo con AR(P)s del polinomio de la componente estacional

$$\phi(B)\Phi(B^s)Y_t = e_t$$

SMA(q)(Q)s como MA(q) multiplicativo con MA(Q)s del polinomio de la componente estacional

$$\theta(B)\Theta(B^s)e_t = Y_t$$

ARIMA como ARMA diferenciado

Si

$$W_t = \nabla^d Y_t$$

es ARMA, entonces Y_t es ARIMA.

SARIMA como SARMA diferenciado

Si

$$W_t = \nabla^d \nabla_s^D Y_t$$

es SARMA, entonces Y_t es SARIMA

Estimación del orden

Función de autocorrelación parcial

Define el efecto que tiene $Y(t-k)$ sobre $Y(t)$, descontando el efecto de las variables $Y_{t-1}, \dots, Y_{t-(k-1)}$,

$$\alpha(1) = \text{corr}(Y_{t+1}, Y_t), \text{ para } k = 1$$

$$\alpha(k) = \text{corr}(Y_t - (\beta_1 Y_{t-1} + \dots + \beta_{k-1} Y_{t-(k+1)}), \\ Y_{t-k} - (\beta_1 Y_{t-(k-1)} + \dots + \beta_{k-1} Y_{t-1})) \text{ para } k \geq 2$$



$$\alpha(k) = \text{corr}((Y_t - P_{t,k}(Y_t), Y_{t-k} - P_{t,k}(Y_{t-K})), \text{ para } k \geq 2$$

$P_{t,k}(x)$ ← proyección ortogonal sobre las funciones lineales generadas por $Y_{t-1}, \dots, Y_{t-(k-1)}$

¿Cómo estimar el orden del modelo?

Model	PACF
White noise	The partial autocorrelation is 0 for all lags.
Autoregressive model	The partial autocorrelation for an $AR(p)$ model is nonzero for lags less than or equal to p and 0 for lags greater than p .
Moving-average model	If $\phi_{1,1} > 0$, the partial autocorrelation <i>oscillates</i> to 0.
	If $\phi_{1,1} < 0$, the partial autocorrelation <i>geometrically</i> decays to 0.
Autoregressive–moving-average model	An $ARMA(p, q)$ model's partial autocorrelation geometrically decays to 0 but only after lags greater than p .

¿Cómo estimar el orden del modelo?

Usando la función de autocorrelación y la función de autocorrelación parcial.

Para un modelo **MA(q)**

- Para $k > q$ la **autocorrelación** es nula
- **Autocorrelación parcial** decae a 0

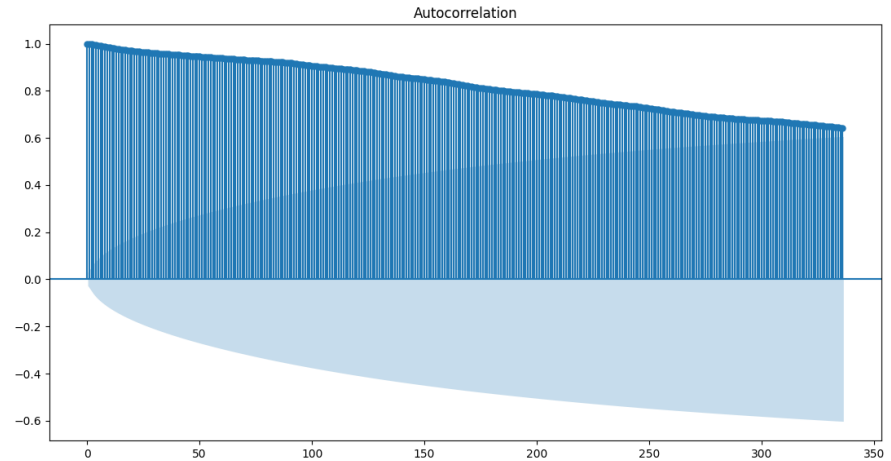
Para un modelo **AR(p)**

- Para $k > p$, la **autocorrelación parcial** es nula
- La **autocorrelación** decae a 0

Para un modelo **ARMA(p,q)**

- Para $k > p$ la **autocorrelación parcial** decae a cero
- La **autocorrelación** decae a 0 después de q lags

Función de autocorrelación

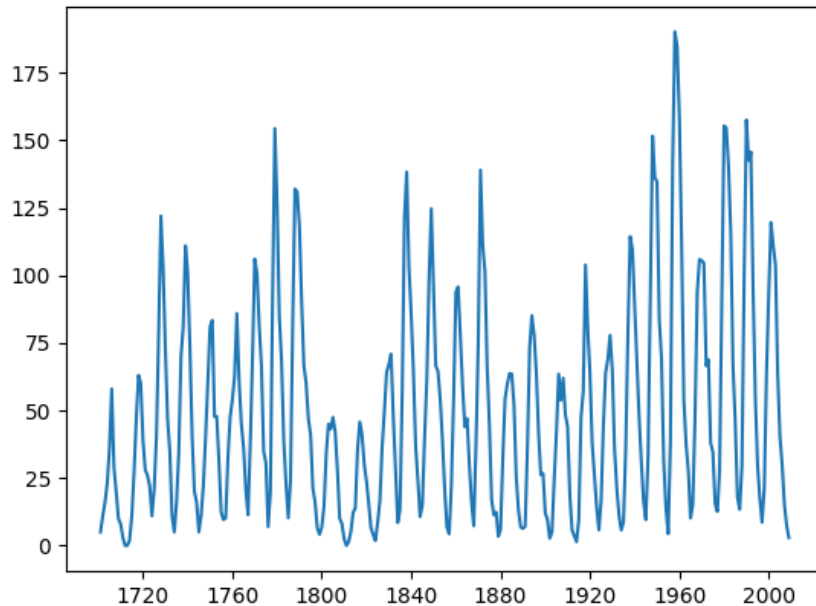


→ Indicio de que la serie es NO estacionaria

Función de autocorrelación

Ejemplo con dataset público

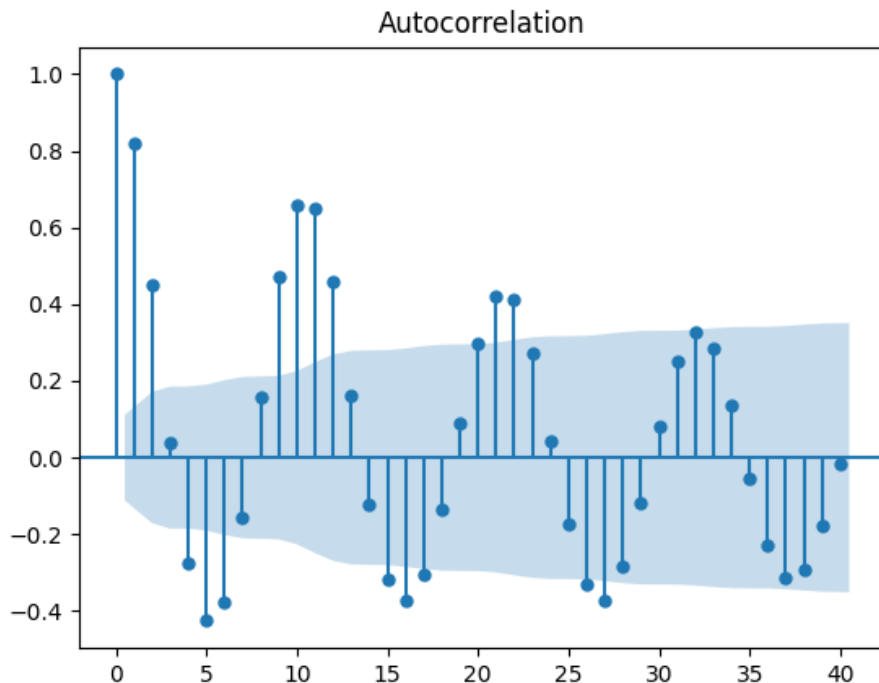
```
import statsmodels.api  
as sm  
  
dta =  
sm.datasets.sunspots.load_pandas().data
```



Función de autocorrelación

Ejemplo con dataset público

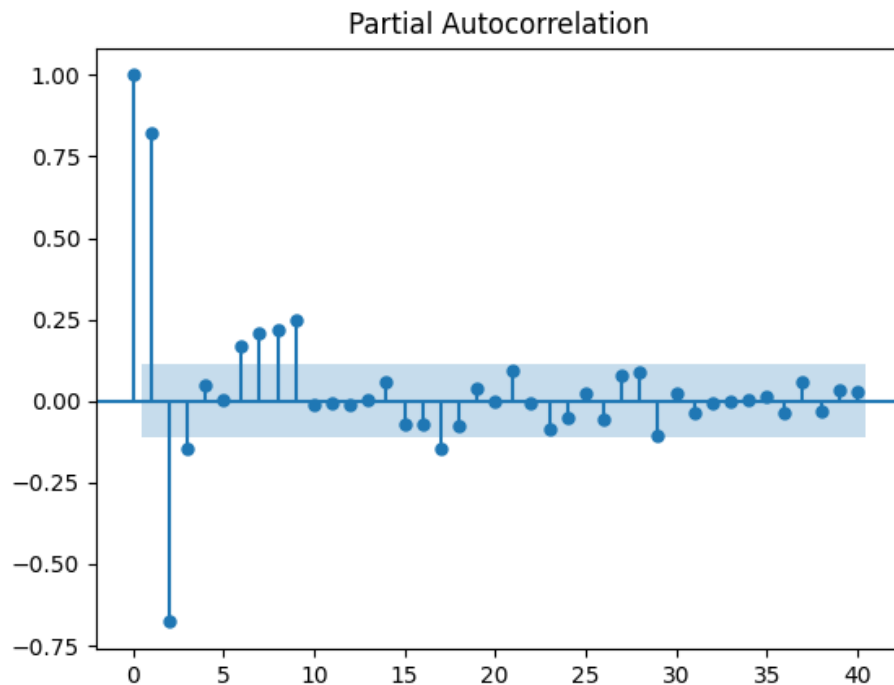
```
sm.graphics.tsa.plot_acf  
(dta.values.squeeze(),  
lags=40)
```



Función de autocorrelación parcial

Ejemplo con dataset público

```
sm.graphics.tsa.plot_pacf  
(dta.values.squeeze(),  
lags=40, method="ywm")
```



Sobre la estimación del orden de integración

¡Cuidado con no diferenciar de más!

La diferencia de cualquier serie estacionaria sigue siendo estacionaria, **pero**:

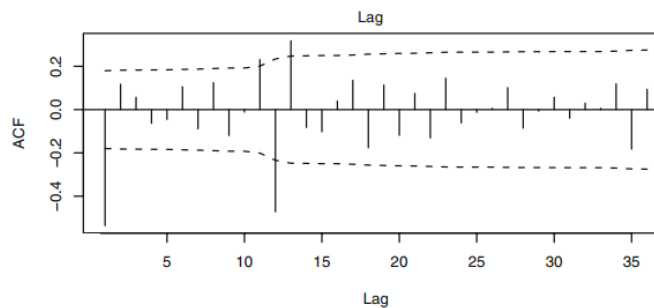
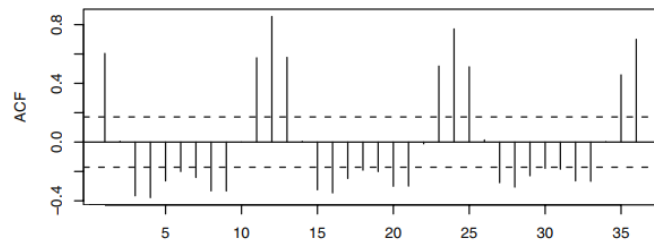
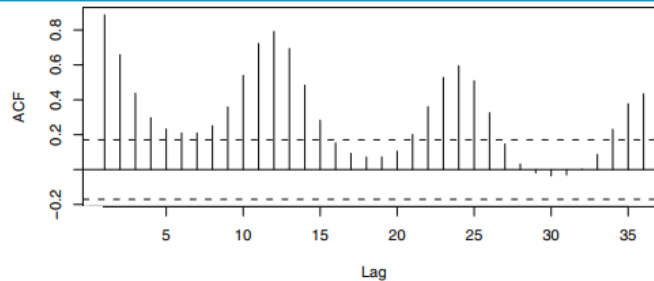
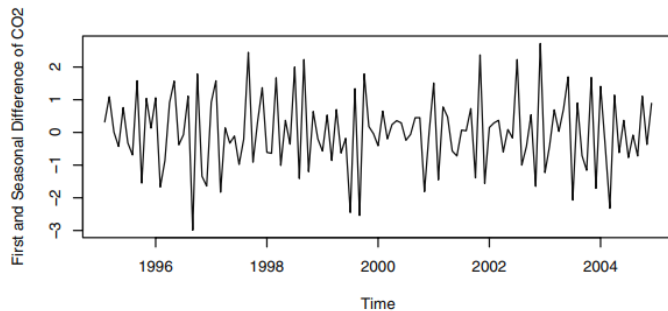
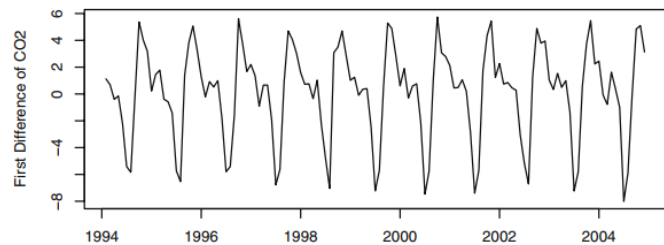
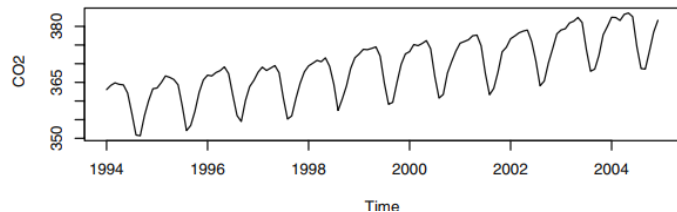
- Diferenciar más de lo necesario introduce correlaciones innecesarias en los datos, complicando el modelado.
- Sobre-diferenciar lleva a modelos no invertibles (el mismo modelo se puede representar con distintas combinaciones de parámetros AR y MA)

Especificación del modelo - modelos estacionales

Se usan las mismas ideas introducidas para modelos ARMA. y ARIMA.

- 1) Inspeccionar la serie de tiempo y su función de autocorrelación muestral.
¿Observo alguna tendencia y/o algún comportamiento estacional?
- 2) Proponer transformaciones (ej. diferenciar) y volver a analizar la serie resultante

Especificación del modelo - ejemplo



¿preguntas?