ARIMA

Machine learning

Data → Model

- Linear reg
- log reg
- SVM
- DT

Time Series Data → Model

- ARIMA
- SARIMA, SARIMAX
- DL, RNN, Attention, transformer

AR
(Autoregression)

⇓

p

[0, 1, 2, 3, ... n]

lag value

PACF (Corrlogram)

I
(Integration)

⇓

d

- [0, 1, 2, 3 ... n]

lag value

Differencing

Non Stationary | Stationary

MA
(moving average model)

⇓

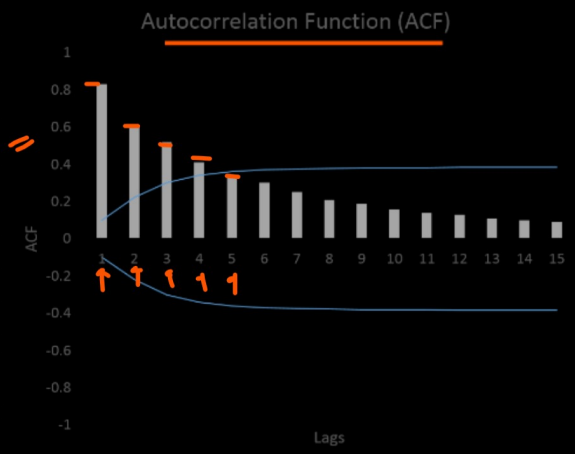
q

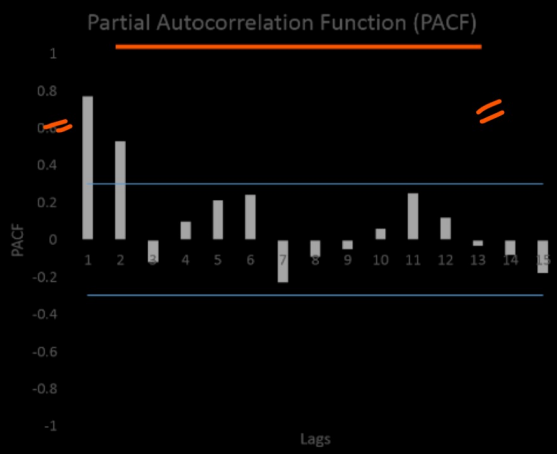
[0, 1, 2, 3 ... n]

lag value

ACF (Corrlogram)

- Y_t / Y_{t-1}
- Y_t / Y_{t-2}
- Y_t / Y_{t-3}





$\equiv$ $Y_{t-2} \longleftrightarrow Y_t$

AR (Autoregression)

$$Y_t = \psi_1 Y_{t-1} + \psi_2 Y_{t-2} + \psi_3 Y_{t-3} \cdots \psi_n Y_{t-n} + C$$

AR(1) $\Rightarrow$ $Y_t = \psi Y_{t-1} + C$

		lag1 TS	lag2
$\rightarrow D_1$	10	NA	NA
D_2	20	$D_1 = 10$	NA
D_3	30	$D_2 = 20$	$D_1 = 10$
D_4	40	$D_3 = 30$	$D_2 = 20$
D_5	45	$D_4 = 40$	$D_3 = 30$

lag = 1
AR = 1
P = 1

20	10
30	20
40	30
45	40
?	45

AR(2) or P = 2 or lag 2

$\equiv$ $Y_t = \psi_1 Y_{t-1} + \psi_2 Y_{t-2} + C$

$\Rightarrow$ $Y_t = \psi Y_{t-1} + C$

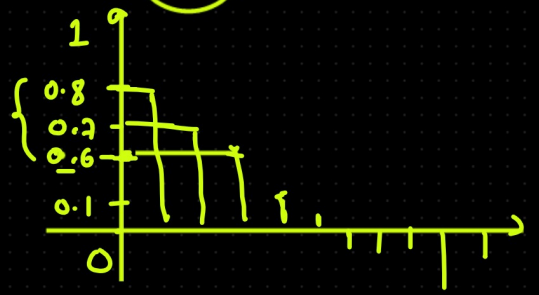
Y_t	Y_{t-1}	Y_{t-2}
30	20	10
40	30	20
45	40	30
?	45	40

?? ?

ACF

AR(3)

lag 3



D_1	10	
D_2	20	
D_3	30	
D_4	40	20
D_5	40	
D_6	60	20

lag 2 → $Y_t - Y_{t-2}$

$Y_t = Y_{t-1}\Psi_1 + Y_{t-2}\Psi_2 + \cdots Y_{t-n}\Psi_n + C$

$Y_t = Y_{t-1}\Psi + \Psi_{t-2}\Psi_2 + \Psi_{t-1c}\,Y_{t-1c} + C$

I ⇒ Integration ⇒ Difference

	D=0	D=1	D=2
Y_{t-4}	D_1		
Y_{t-3}	D_2	$(D_2 - D_1)$	
Y_{t-2}	D_3	$(D_3 - D_2)$	
Y_{t-1}	D_4	$D_4 - D_3$	
Y_t	D_5	$(D_5 - D_4)$	

$$D_1 = (Y_t - Y_{t-1})$$

$$D_2 = (Y_{t-1} - Y_{t-2})$$

$$D_2 = (Y_t - Y_{t-1} - Y_{t-2})$$

Moving average (error) ⇒

$$y_t = \varepsilon_{t-1}\psi + c$$

$$y_t = \varepsilon_{t-1}\psi_{t-1} + \varepsilon_{t-2}\psi_{t-2} + \varepsilon_{t-3}\psi_{t-3} -- \varepsilon_{t-n}\psi_{t-n} + c$$

ARIMA ⇒

$$(y_t) = (y_{t-1}\psi_{t-1} + y_{t-2}\psi_{t-2} + -- y_{t-n}\psi_{t-n} + c) + (y_{t-1} - y_t - y_{t-2} -- y_{t-n}) +$$

$$(\varepsilon_{t-1}\psi_{t-1} + \varepsilon_{t-2}\psi_{t-2} + \varepsilon_{t-3}\psi_{t-3} - ---- - \varepsilon_{t-n}\psi_{t-n} + c)$$

AR(1) ⇒ P=1
I (2) ⇒ d=2
MA(1) ⇒ q=1

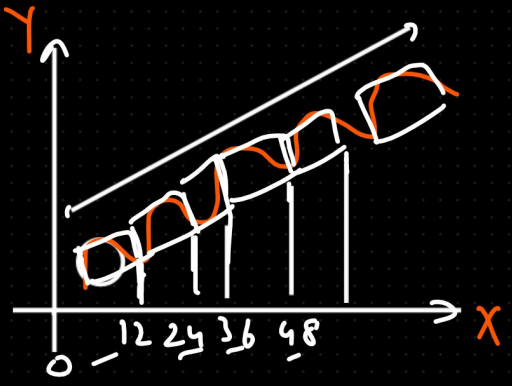
$$y_t = (y_{t-1}\psi_{t-1} + c) + (y_t - y_{t-1} - y_{t-2}) + (\varepsilon_{t-1}\psi_{t-1} + c)$$

SARIMA

↳ Seasonal

p, d, q

(PDQ)



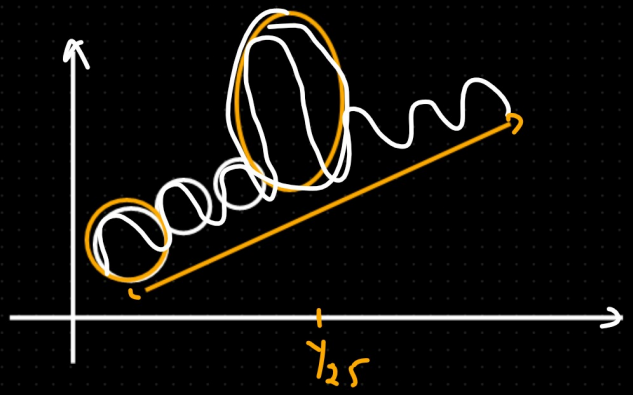
$(p,d,q)(P,D,Q)_s$

$\Rightarrow (1,1,2)(1,0,2)_{12}$

SARIMx

↳ exgeoug

(outlier) (entraeffect)



$(p,d,q)(P,D,Q)_s$ X

$\Rightarrow (0,1,2)(1,1,2)_{12}$ Y_{25}