Topics Covered today

1. ACF/PACF Contd.
2. AR
3. MA
4. ARMA
5. ARIMA
6. SARIMA
7. Exogenous Variable
8. SARIMAX

⇒ ARIMA family of Models ← Agenda

Start at 9:05

① TS Decompisition → trend, seasonality, ~~~

Simple method →

↓

Smoothing method → SES, DES, TES

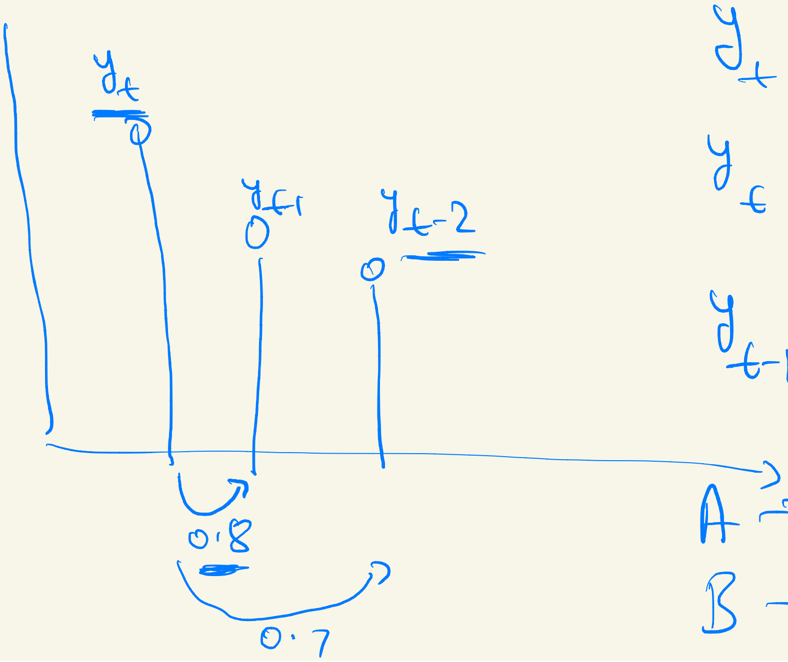
Holt's whitier's method

Stationarity & ACF →

⇒ Autocorrelation → Correlation b/w y_t & any other prev value.

y_t $\quad$ y_{t-1}

y_{t-2}



$y_t \quad y_{t-1} = 0.8$

$y_t \quad y_{t-2} = 0.7$

$y_{t-1} \quad y_{t-2}$

A → B

B → C.

$A \to C$ to certain extent

Partial Auto Correlation → Correlation b/w 2 time values after isolating them from impact of everything else.

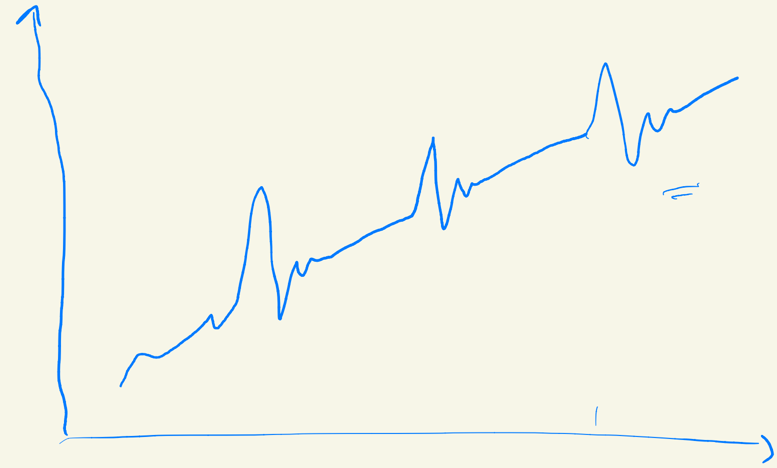
$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots$$

$$y_{t-1} = \alpha_1 y_{t-2} + \alpha_3 y_{t-3} + \dots$$

$$y_{t-2} = \frac{1}{\alpha_1} y_{t-1} - \frac{1}{\alpha_1}(\alpha_3 y_{t-3} + \dots)$$

$$y_t = \alpha_1 y_{t-1} + \frac{\alpha_2}{\alpha_1} y_{t-1} - (\quad) + (\quad)$$

⇒ TES → 3.6% Error



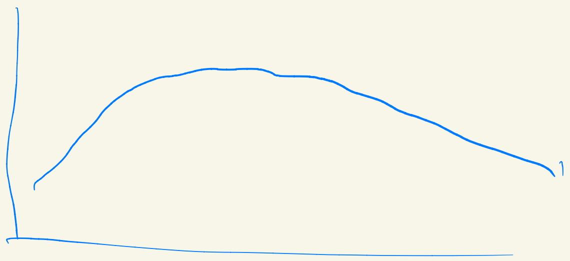
Sales

Smoothing method

↳ Trended

Seasonal

TES → 2-4%



→ • Stock market

• GDP rates

• Sensor data

$$y = b(t) + s(t) + r(t)$$

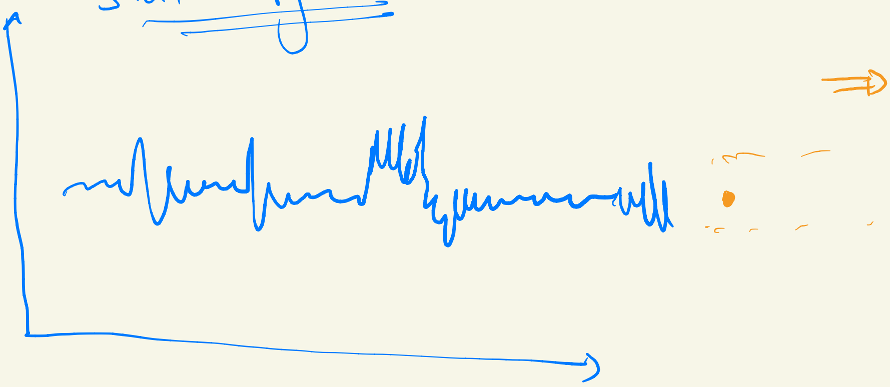
$z(t)$

↓

??

$e(t)$

Stationary TS



⇒ mean

⇒ ARIMA Models

1) AR
2) MA
3) ARMA
4) ARIMA
5) SARIMA
6) SARIMAX

---

① AR → Auto regression

on self

linear regression

$$\hat{y}_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_0$$

$$3 y_{t-1} + 2y_{t-2} + 1$$

$$\hat{y}_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots \alpha_{10} y_{t-\underbrace{10}_{p}} + \alpha_0$$

Smoothing weights → • sum to 1 • decline rapidly

p → Auto regressor Parameter

$AR(p) \to AR(3), \; AR(4), \; \dots$

$$\hat{y}_t = \boxed{AR(3) = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \alpha_0}$$

⇒ 10, 11, 12, 13, 14, 20, 25, 27, 28, 29 — TS Data

y_{t-3}	y_{t-2}	y_{t-1}	y_t
10	13	14	20
13	14	20	25
14	20	25	27
20	25	27	28
25	27	28	29

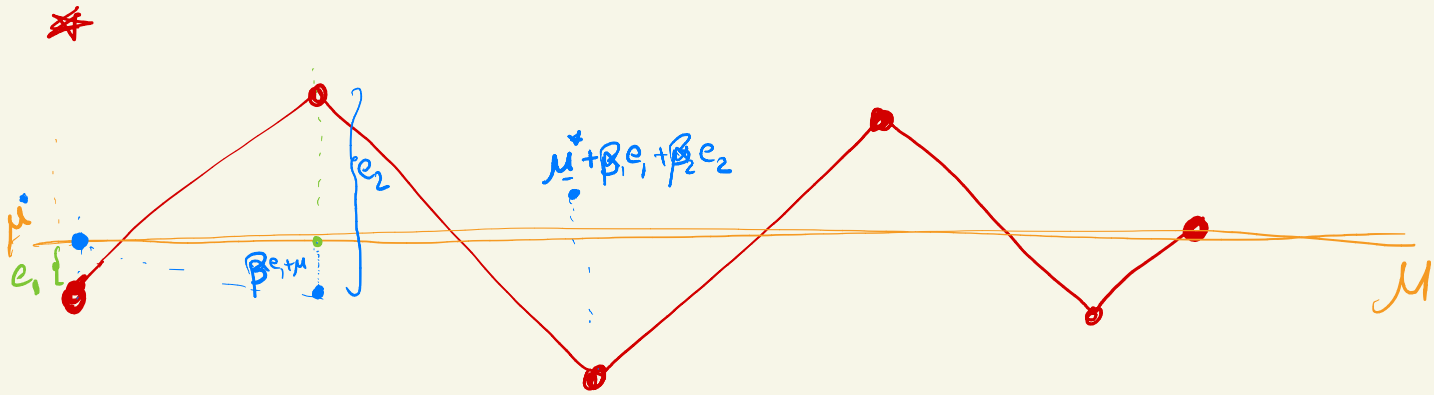
$\underbrace{\qquad\qquad}_{X}$ $\underbrace{\quad}_{Y}$

normal LR problem

Auto Regression

② MA(q) → # Moving Avg method (Simple method)

rolling avg



MA key Idea ⇒ I should learn from each error.

$$\boxed{\hat{y}_t = \mu + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots \dots \beta_q e_{t-q}}$$

⇒ μ_{t-1} → data till $t-1$

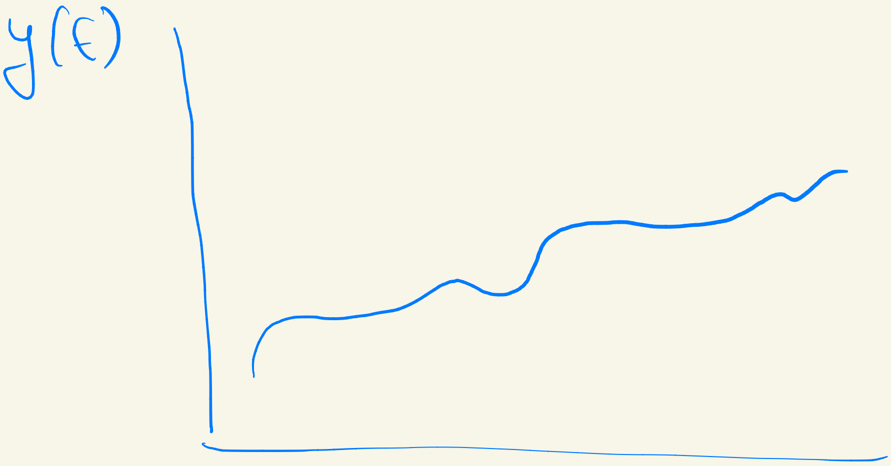
to predict t → $\hat{y}_t - y_t$

e

$\mu_t + \beta e_1 \rightarrow \hat{y}_{t+1}$

③ ARMA → AR + MA

$$\hat{y}_t = c_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots \alpha_p y_{t-p}$$
$$+ \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots \alpha_q e_{t-q}$$

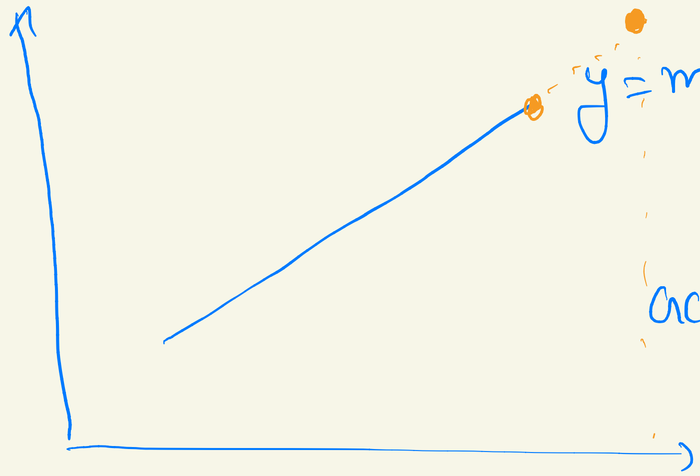


Actual 100, 120, 110, 115, 98, 127

Diff ($y'(t)$) +20, −10, +5, −17, +29,

$\hat{y}'(t)$ 19 −7 +4 −13 +20

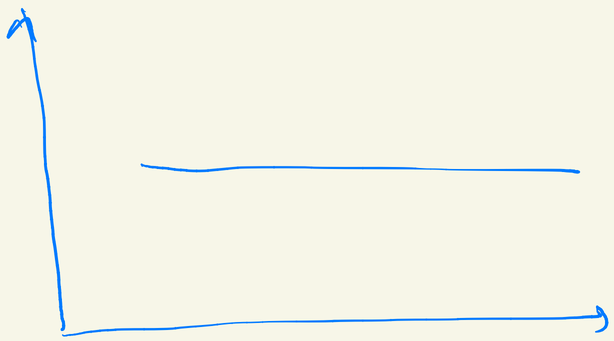
$\Rightarrow$ Differentiate $\rightarrow$ Integrate.



$y = mx + c$

actual $y(t)$

$\Rightarrow$ last y value $+ \; m(t_{t+1} - t)$



$y' = m$

$y'(t)$

m

$$y_t = m x_t + c$$

$$y_{t+1} = m x_{t+1} + c$$

$$\boxed{x_{t+1} = x_t + 1}$$

$\hat{y} \Rightarrow$ = Last Value + cumsum of Prediction of y'

⇒ Take TS

↓

→ Differentiate → decompose
→ diff | detrend ⇒ ✱

↓

Predict using AR, MA, ARMA

↓

→ Integrate → (cumsum) ✱

↓

Final Answers

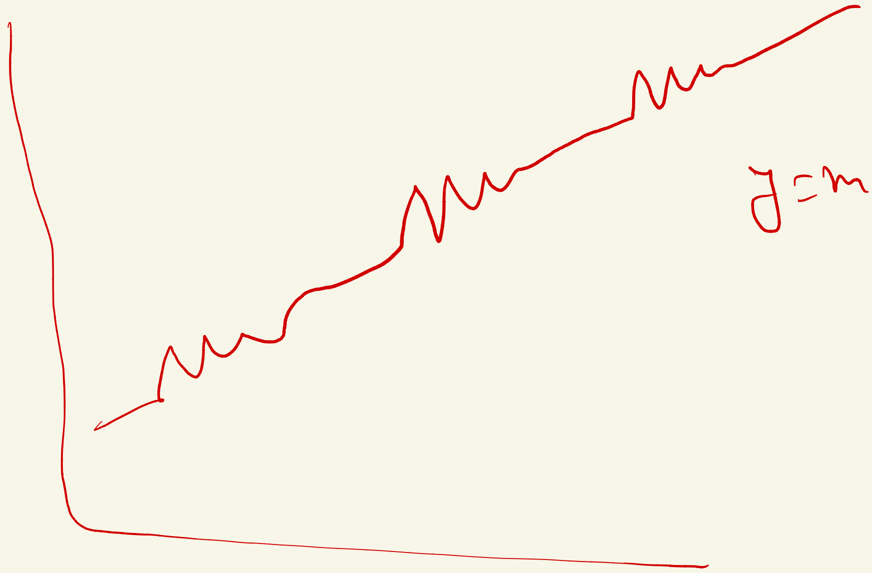
ARIMA

↓

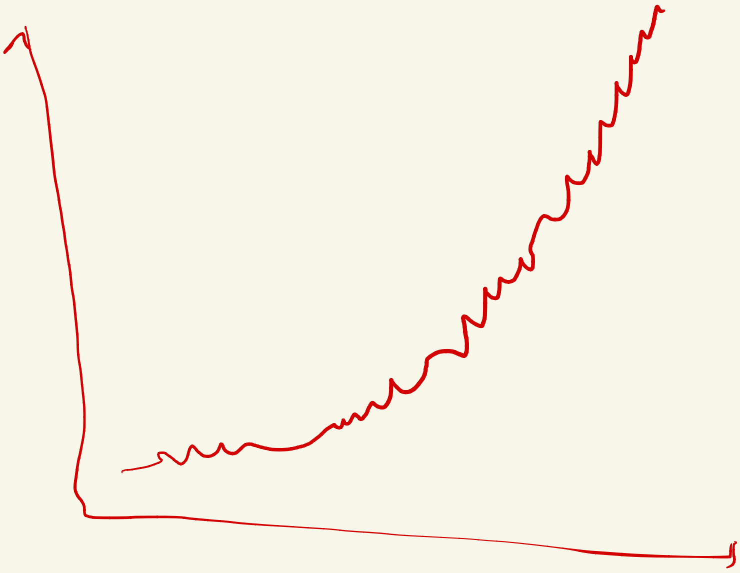
Integration → Auto

Trend $\rightarrow$ $y = mx + c$

'diff



$y = mt + c$



$y = bt^2 + ct + d$

Single differentiation

2 differentiation

$y \Rightarrow$	100	120	115	117	98	123
y'		20	-5	$+2$	-19	$+25$
$y'' \Rightarrow$			25	-7	21	-44

Order = (p, d, q)

- p → AR terms
- d → no. of differentiation to do
- q → Count of MA terms

⇒ SARIMA ← (p, d, q, P, D, Q, S)

S → Seasonal → → Impact of very old / Seasonal

Jan Sale → Seasonality →

$$AR(p) = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots \alpha_p y_{t-p} + \alpha_0$$

$$SAR(P) = \gamma_1 y_{t-S} + \gamma_2 y_{t-2S} + \dots \gamma_P y_{t-PS} + \gamma_0$$

$$SMA(Q) = \delta_1 e_{t-s} + \delta_2 e_{t-2s} + \dots \delta_Q e_{t-Qs} + \delta_0$$

$$\begin{aligned}\text{SARIMA} = C &+ \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots \alpha_p y_{t-p} \\ &+ \beta_1 e_{t-1} + \dots \beta_q e_{t-q} \\ &+ \gamma_1 y_{t-s} + \dots \gamma_P y_{t-Ps} \\ &+ \delta_1 e_{t-s} + \dots \delta_Q y_{t-Qs}\end{aligned}$$

$p \to$ AR $\quad P \to$ SAR

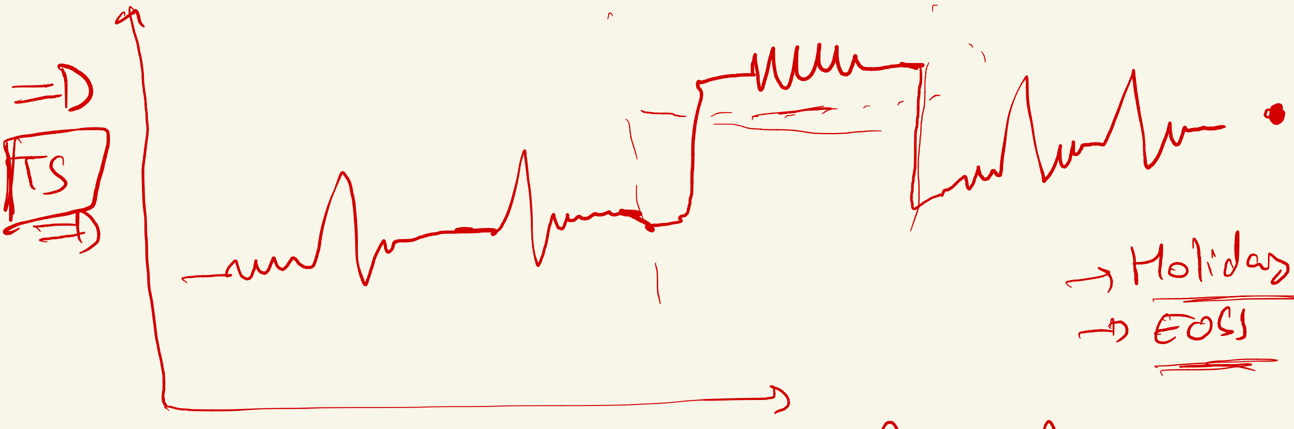
$q \to$ MA $\quad Q \to$ SMA

$d \to$ ARIMA $\quad D \to$ SARIMA

$S \to$ seasonality

SARIMAX

↳ Exogeneous Variables → ??

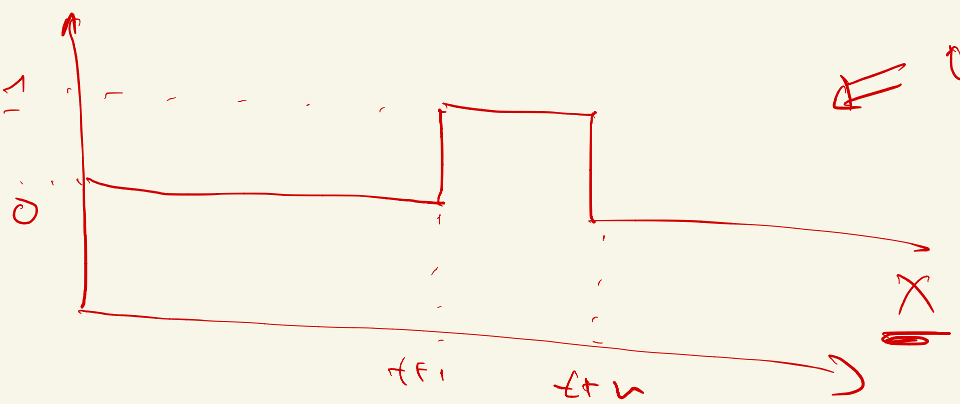


→ SES, Simple, ARIMA

⇒ Seasonality → Winters

① Holiday EOSS →

② we give extra information to the model.



⇐ were marketing campaign running

⇒ SARIMAX → α_1

β_1

γ_1

δ_1

weightage

extra variable → $\omega_1 x_1 + \omega_2 x_2 + \dots + \omega_z x_z$

Practical Approach

Data → EDA → missing values + outliers

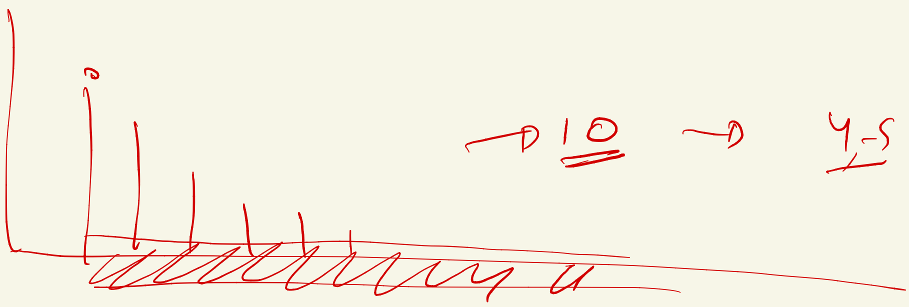
→ TS decompose

↳ TES → ( ) 2-4%

→ PACF Plot

↳ p, q, P, Q, S ←

↳ SARIMA model → 2-4%



→ 10 → 4-5