Last ML2 session

Time Series Forecasting - 5

→ Exogenous vars

→ ARIMA

→ CI

→ FB Prophet

ARIMA

→ AR : Auto Regressor

AR(p)

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

Regression

feature			target
y_{t-3}	y_{t-2}	y_{t-1}	y_t
$t-4$	$t-3$	y_{t-2}	y_{t-1}
$t-5$	$t-4$	$t-3$	y_{t-2}
X			y

past p values are feat

Past 3 value feat

$p+1$ params

AR(p)

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

Key : idea → future depends on past values

MA(q)

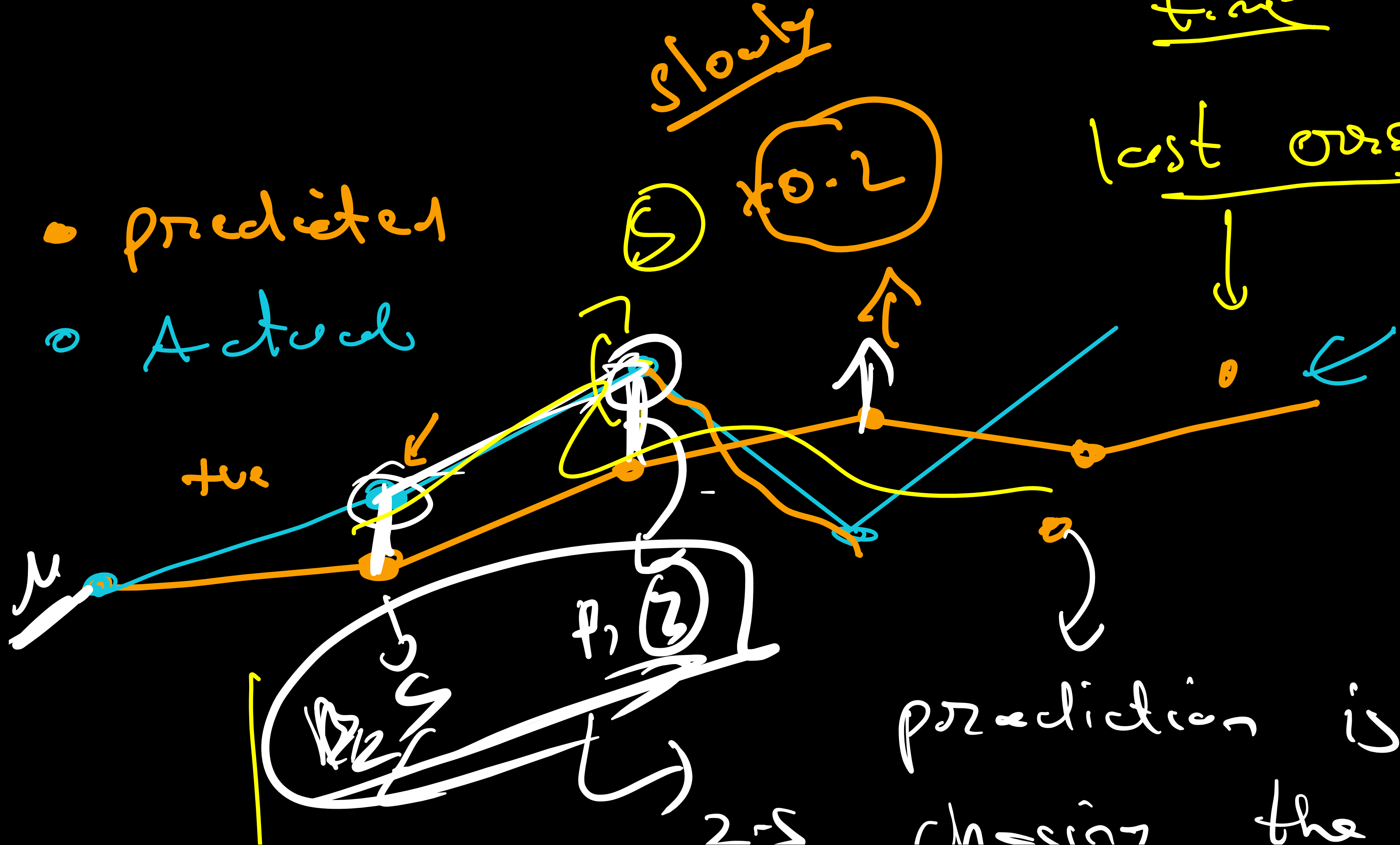
global mean

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

Key: future value depends on the past errors

Stationary TS : Mean does not change with time

last errors



prediction is chasing the

true value

$$\hat{y}_t = \mu + \beta_1 e_{t-1}$$

0.2

$$+ \beta_2 e_{t-2}$$

$$\dots$$

$$\beta_q \underline{e_{t-q}}$$

ARMA(p, q)

$$\hat{y}_t = c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots \alpha_p y_{t-p}$$

$$\beta_1 e_{t-1} + \beta_2 e_{t-2} \dots \beta_q e_{t-p}$$

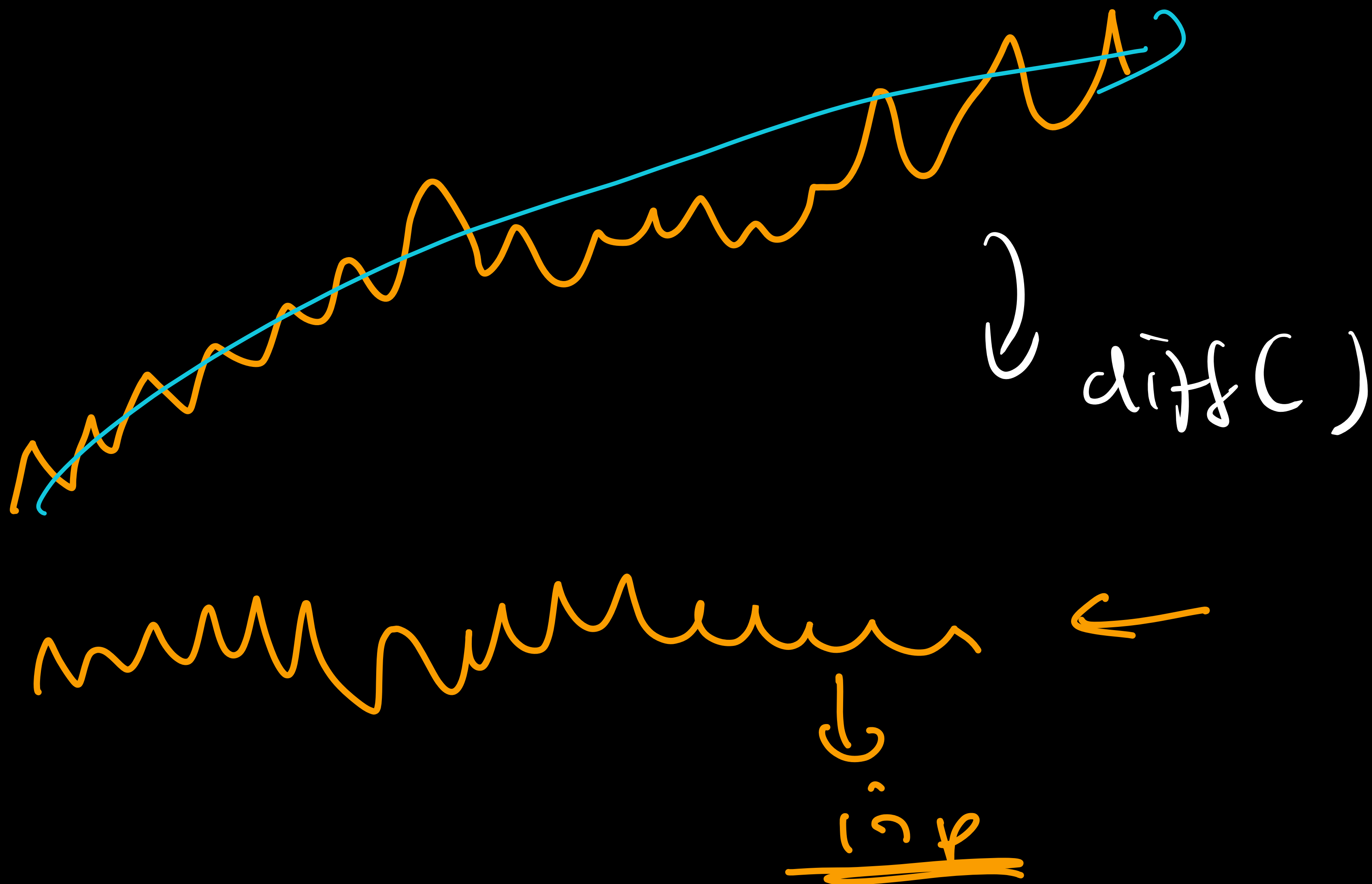
these models assume stationarity

inp = mobile_sales.diff()

pred = SARIMAX( )

cumsum() + sales[-1]

oup = pred.



Jan

	J	F	M	A	May
$y =$	100	120	110	115	95
$y' =$		20	-10	+5	-20

diff

i.i.d

$\downarrow$

AR / MA / ARMA

Jun Jul Aug . .

$\hat{y}' =$ +10 -5 +15

inter

$y[-1] = 95$

$\hat{y}_t = y_0[-1] + \hat{y}'_{t+1}$

	Jun	Jul	Aug
y	105	$\boxed{100}$	$\boxed{115}$

$95 + \hat{y}'.\text{cumsum}()$

+10	−5	+15
+10	+5	+20

+10

105 100 115

d=1

3

0

2

3

ARIMA(p, d, q)

Integration

AR(p)

MA(q)

I(d)

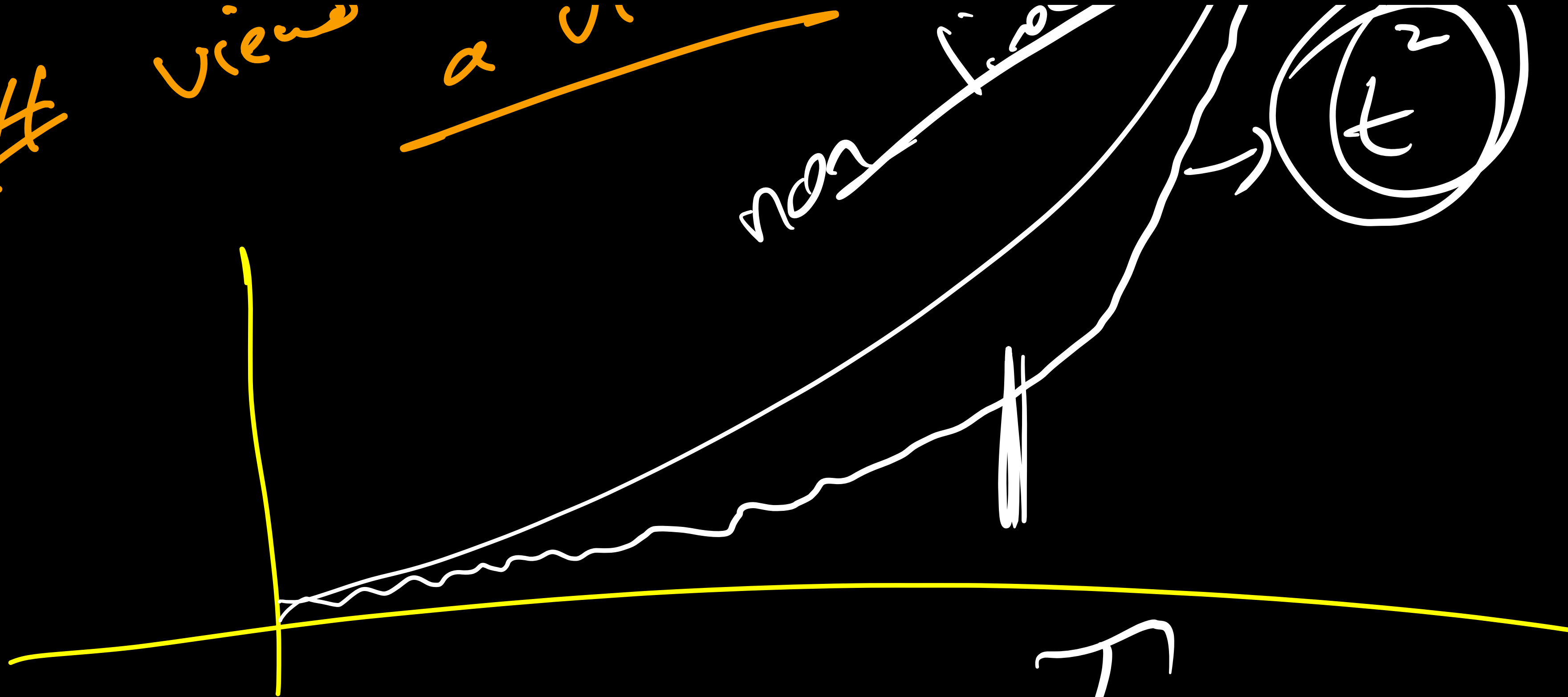
$$y = [\, y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \,]$$

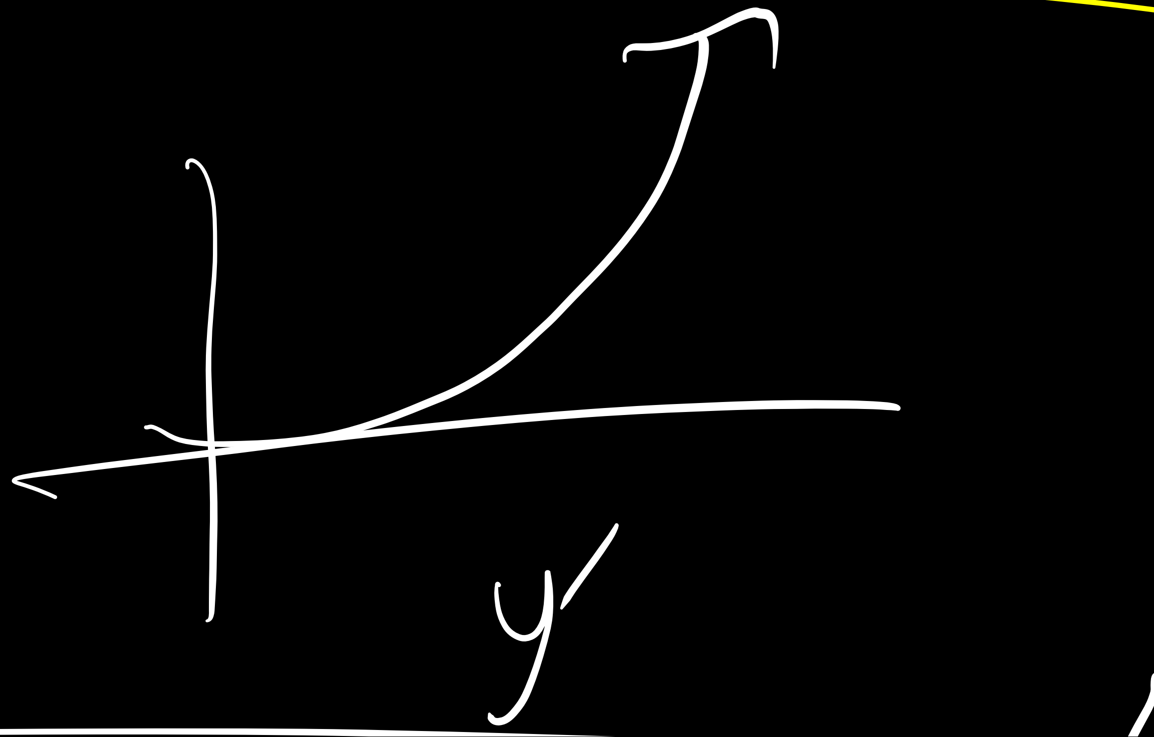
$$y' = [\, y'_1 \quad y'_2 \quad y'_3 \quad y'_4 \,]$$

$$y'' \; [\, y''_1 \quad y''_2 \quad y''_3 \,]$$

$$\frac{d}{dt}\left(\frac{d(y)}{dt}\right) \rightarrow \frac{\text{diff}}{d(t)}$$

views a





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

$$
\begin{aligned}
\hat{y}_t = c + {} & \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_p y_{t-p} \\
& \beta_1 e_{t-1} + \beta_2 e_{t-2} + \beta_q e_{t-q} \\
& \gamma_1 y_{t-s} + \gamma_2 y_{t-2s} + \gamma_P y_{t-Ps} \\
& \delta_1 e_{t-s} + \delta_2 e_{t-2s} + \delta_Q e_{t-Qs}
\end{aligned}
$$

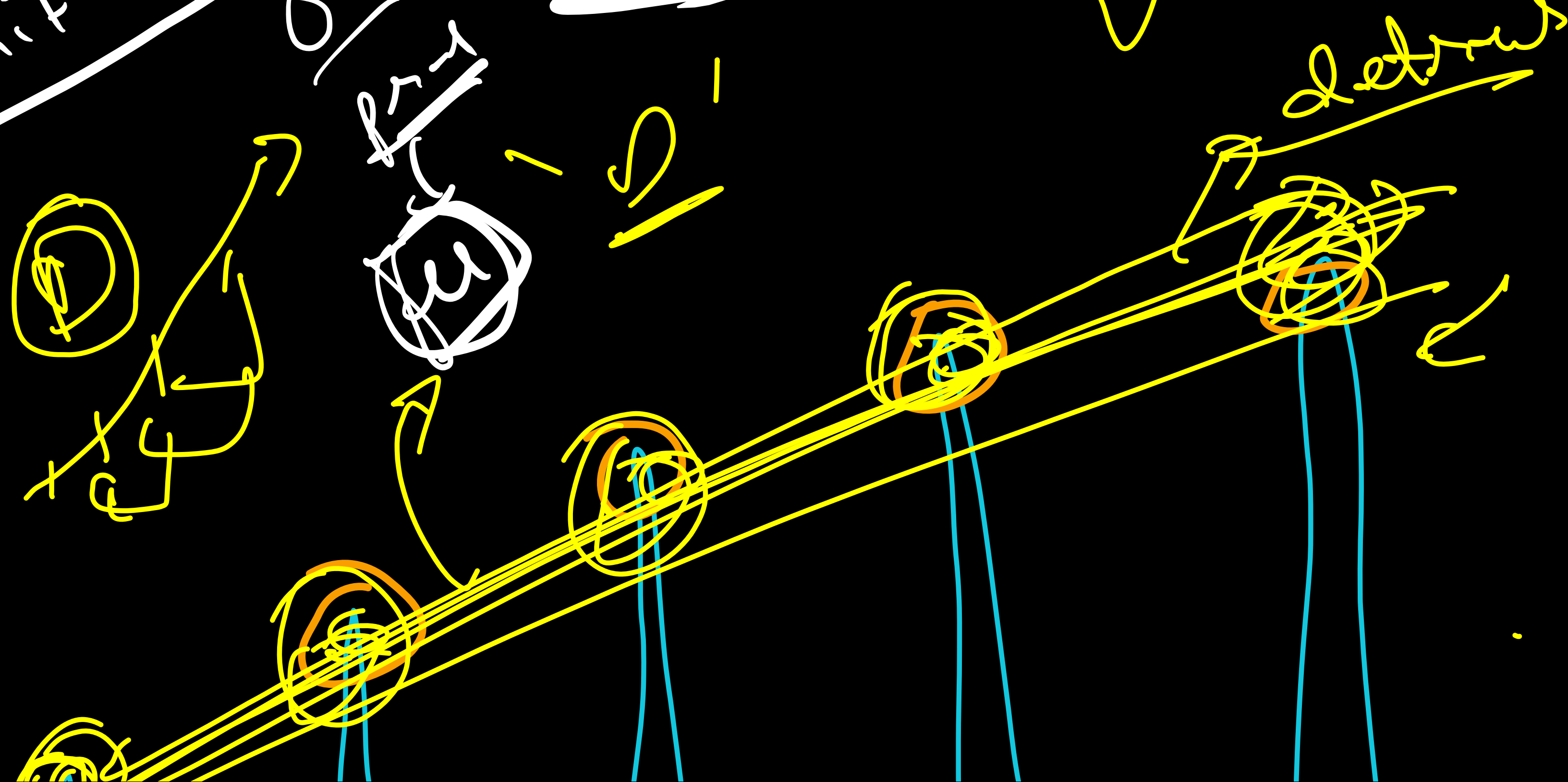
SARIMA (p, d, q, P, D, Q, s)

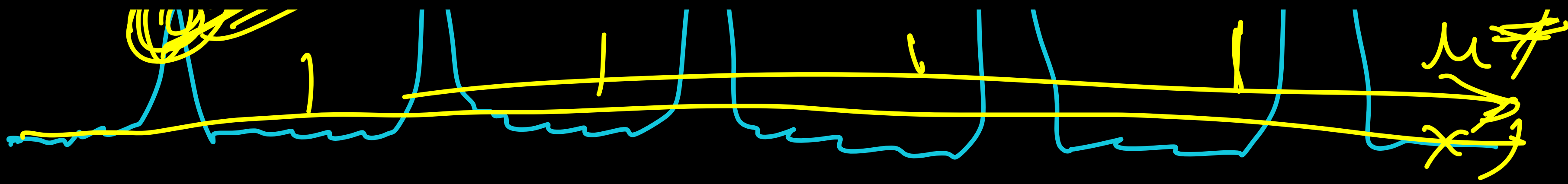
Init

0th

diff (D)

D





FlPKT → Big Billion Sale

1Bn

1.5B

2.5B

SARIMAX

$\hat{y}_t$ → endogenous vars

itself

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

$$+ \beta_1 e_{t-1} + \beta_2 e_{t-2} + \beta_q e_{t-q}$$

$$\gamma_1 y_{t-s} + \gamma_2 y_{t-2s} + \gamma_p y_{t-p}$$

$$\delta_1 e_{t-s} + \delta_2 e_{t-2s} + \delta_Q e_{t-}$$

$$+ \xi_1 (\text{holiday}) + \xi_2 (\text{GDP})$$

$$+ \xi_3 (\text{4G penetration})$$

SARIMAX

Which model to use !!

i) complete random

↳ Sample methods

mean
naive
ARMA
SES

↳ trend

DES
ARIMA

↳ Seasonity
TES

SARIMA

Exog

custom / manual

SARIMAX