Time Series Analysis

- Pearson's Correlation coeff: (PCC) (metric b/w 2 vars)

$$r = PCC(x,y) = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum (x_i-\bar{x})^2 \sum (y_i-\bar{y})^2}} = \frac{covar(x,y)}{\sigma_x \sigma_y}$$

(basically normalised covariance)

$r > 0 \Rightarrow$ directly $\propto$
$r < 0 \Rightarrow$ inversely $\propto$

to measure if there is a good rel^n b/w the 2 vars;

we check for the null hypothesis:

- Null Hypothesis: r does not significantly differ from 0.

$\Downarrow$

there is no linear rel^n ship.

- t-test:

$$t = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}}$$

(n is the sample size)

- r^2 = Coefficient of determination

⇒ % of variance in one variable that is accounted for by the variance in other variable.

(ie variability induction in one var on inducing variability in the first. hence greater r^2 ⇒ better relⁿ.

- A p-value is calculated from the test statistic t
- If pvalue is less than the significance level. (generally $\alpha = 0.05$)

reject Null hypothesis

∴ $p \leq 0.05$ ⇒ corr. exists.

ACF: Autocorrelation f^n:

gives pcc for a time series with a window size (k)

ex. there is a list of amt of rainfall each day in jan.

To find the correl^n of today's rainfall with day before yesterday's rainfall; we have to take values like

x:	dec 30	dec 31	jan 1	jan 2	...	jan 29
y:	jan 1	jan 2	jan 3	jan 4	...	jan 31

this is the x, y column. taking pcc; we get the auto-correlation (A f^n of window size k).

PACF: Partial autocorrelation f^n:

This only captures the direct affect of S_{t-2} (rainfall on day t-2) on S_t.

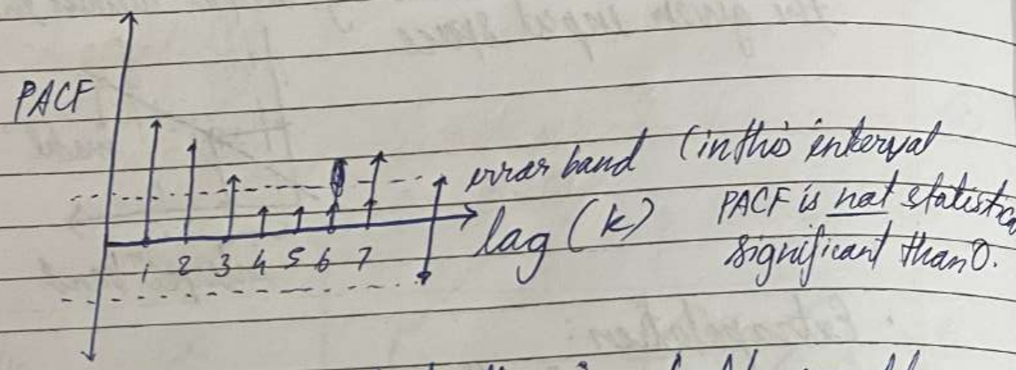
but S_t can also be indirectly affected by S_{t-1}

as $S_{t-2} \to S_{t-1} \to S_t$; so for independently capturing these direct & indirect effects, we use a regression approx^n.

ex $S_t = \phi_{21} S_{t-1} + \boxed{\phi_{22}} S_{t-2} + \underbrace{\varepsilon_t}_{\text{error}}$

PACF

now say we get the plot:



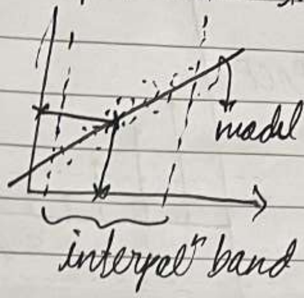
∴ a deduction can be that there is a feeble weekly periodicity in rainfall

• we will now consider $S_t = \phi_0 + \phi_1 S_{t-1} + \phi_2 S_{t-2} + \phi_3 S_{t-3} + \phi_4 S_{t-7}$

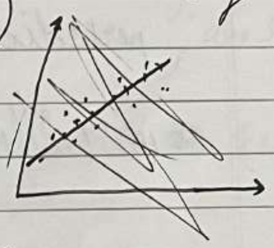
Auto regression Model

• In an autoregressive model we generally get a decaying ACF.

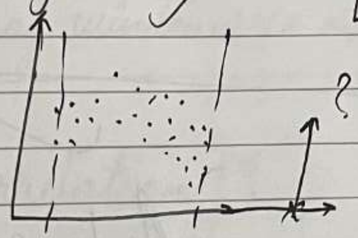
- Interpolation:
  Predicting output within the range of the given input space



- Extrapolation:
  Predicting output outside the band of the given input space. (forecasting)



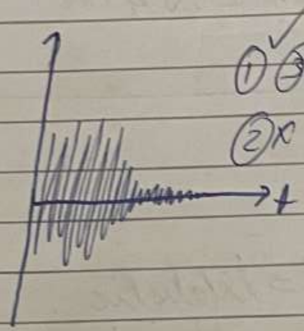
(time series probs. are generally extrapol^n)

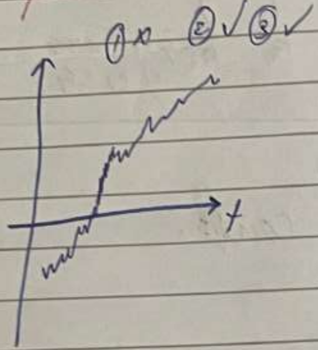


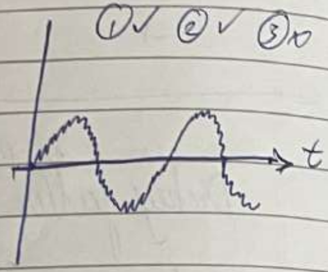
★ Always try to convert the time series in a form with ≈ stationary mean. (like how an audio signal looks like)

⇒ Stationarity

- μ is constant ①
- σ is constant ②
- there's no seasonality (predictable periodic behaviour) ③







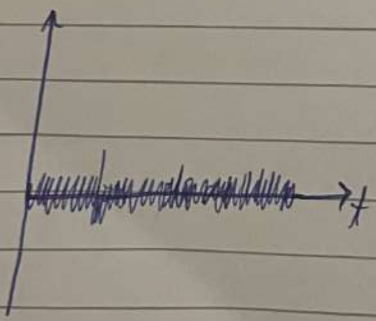
- white noise is stationary + $\mu = 0$.
- for checking statio--y we can do it

1. visually 2. check local vs global.

3. Augmented Dickey Fuller test

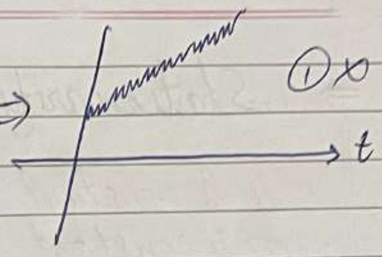
⇒ How to make a time series stationary.

- We want to convert to the following form



$$\approx f(t) = \underbrace{\beta_0}_{\text{mean}} + \overbrace{\varepsilon_t}^{\sigma \,=\, \text{fluctuation}}$$

ex. if $y_{(t)} = \beta_0 + \beta_1 t + \varepsilon_t$ $\Rightarrow$



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

$= \beta_1 + (\underbrace{\varepsilon_t}_{k_1^2} - \underbrace{\varepsilon_{t-1}}_{k_2^2})$ } variance of ε_t & ε_{t-1}

$\mu = \beta_1$

$\sigma = \sqrt{k_1^2 + k_2^2}$

Dickey fuller contd.

$$t_{\hat{\delta}} = \frac{\hat{\delta}}{SE(\hat{\delta})}$$

t ⇒ t statistic

SE ⇒ std. error.

$$\underbrace{y_t - y_{t-1}}_{\Delta y_t} = \mu + \underbrace{(\phi - 1)}_{\delta} y_{t-1} + \varepsilon_t$$

$H_0 : \delta = 0$

$H_1 : \delta < 0$

compare $t_{\hat{\delta}}$ to dickey fuller distn

$t_{\hat{\delta}} < DF_{critical}$ ⇒ reject H_0 ⇒ no unit root → stationary

Augmented DF test:

let the model be ARP:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t$$

$$\Delta y_t = \mu + \underset{(\phi_i - 1)}{\delta} y_{t-1} + \sum_{i=1}^{r} \beta_i \Delta y_{t-i} + \varepsilon_t$$

$$t_{\hat{\beta}_i} = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}$$

$H_0 : \delta = 0$

$H_1 : \delta < 1$

then compare with DF distributⁿ.

White Noise

- A time series with $\Rightarrow \mu = 0$
  $\Rightarrow$ st dev $\Rightarrow$ const with time
  $\Rightarrow$ correlation b/w lags is 0.

- White noise by defn is not predictable.

$$y_t = \underbrace{\text{signal}}_{\text{predictable}} + \overbrace{\text{noise}}^{\text{not pred.}}$$

* fit model, get y_t - signal;
  check if it satisfies white noise criteria

(visual test, global vs local checks, check ACF; check correlogram).

if white noise → good model.

AR ⇒ auto regressive
MA ⇒ moving avg

⇒ Backshift Operator :- (Lag operator)

ARMA (3,3) model (AR3 & MA3)

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \varepsilon_t$$

$$\underbrace{y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \phi_3 y_{t-3}}_{\text{AR terms}} = \underbrace{\theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \varepsilon_t}_{\text{MA terms}}$$

Lag operator / Back shift operator.

$$\hat{L} y_t = y_{t-1} = \hat{B} y_t$$

$$\hat{L}^2 y_t = y_{t-2} \text{ --}$$

$$\therefore \underbrace{(1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3)}_{\Phi(L)} y_t = \underbrace{(1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3)}_{\Theta(L)} \varepsilon_t$$

$$\therefore \Phi(L) y_t = \Theta(L) \varepsilon_t$$

- If we predict 3 yrs in future, we have progressively increasing variability ⇒ error, hence

what we do is:

Train on months 1, 2 - - $k-3$ → predict month $k-2$
1 - - $k-3, k-2$ - - - $k-1$
1 - - $k-2, k-1$ - - - k

the averaging out will give a good pic.

Moving average model (MA)

predicted no. coeff.

$$\hat{f}_t = \underset{mean}{\mu} + \phi_1 \varepsilon_{t-1}$$ } MA(1) model

ε_{t-1} → error from the prev. run

$\varepsilon_t \Rightarrow$ error in μ, $\varepsilon_t \Rightarrow$ Normal $(\overset{0}{\mu_\varepsilon}, \overset{1}{\sigma_\varepsilon})$

Say every day Prof tells me to bring μ no. of cupcakes, & I bring $\mu \pm \varepsilon_t$ every day;

now tomorrow I'll bring $10 + \phi_1 \varepsilon_{t-1}$ to adjust for ~~yester~~ today's error.

MA(2) : $\hat{f}_t = \mu + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2}$

$$f_t = \mu + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \varepsilon_t$$

ε_t → error (~~today~~ b/w pred & actual today)

* We use ACF to tell us about the MA part of an ARMA model

- Predicting over longer periods will converge to the mean.

($\star$ $corr(a,b) = ACF(a,b)$

$$= \frac{\mathbb{E}(ab) - \mathbb{E}(a)\mathbb{E}(b)}{\sqrt{\sigma^2(a)}\sqrt{\sigma^2(b)}}$$

$\hat{x}_t$

Let $x_t = \mu + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} \cdots \phi_q \varepsilon_{t-q} + \varepsilon_t$ } $MA(q)$

$corr(x_t, x_{t-k}) = [\mathbb{E}(x_t x_{t-k}) - \mathbb{E}(x_t)\mathbb{E}(x_{t-k})] / \sqrt{\sigma^2(x_t)}\sqrt{\sigma^2(x_{t-k})}$

To figure out the order of an MA model use ACF

$\mathbb{E}(x_t) = \mu + \phi_1(0) + \cdots + \phi_{t-q}(0) + 0 = \mu$

similarly $\mathbb{E}(x_{t-k}) = \mu$

$(x_{t-k} = \mu + \phi_1' \varepsilon_{t-k-1} + \phi_2' \varepsilon_{t-k-2} \cdots \phi_q \varepsilon_{t-q-k} + \varepsilon_{t-k})$

$\mathbb{E}(x_t x_{t-k}) = ?$

$\mathbb{E}(\varepsilon_t \cdot \varepsilon_{t-x}) = 0$ ($\because$ errors are indep. on lagging

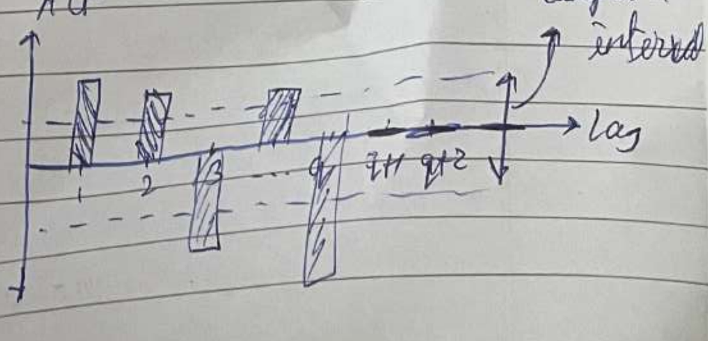
but $\mathbb{E}(\varepsilon_{t-3}^2) \neq 0$

$Var(\varepsilon_t) = \mathbb{E}(\varepsilon_t^2) - (\cancel{\mathbb{E}(\varepsilon_t)})^2$ (0)

overlapping terms iff $t-q \leq t-k \Rightarrow k \leq q$

$\therefore$ Autocorrel^n of a time series with (k) lagged version of itself

is Not 0 iff $k \leq q$;



We can figure out if a time series is well modelled by an MA(q) model if the ACF drops to 0 ~~from~~ after some lag q. The ~~val~~ Lag

★ The lag after which ACF of a time series drops to 0 & stays 0 ⇒ indicates an MA model of the same order

⇒ Invertability:

$MA(1) \Leftrightarrow AR(\infty)$

say $MA(1)$: $c_t = -\phi \varepsilon_{t-1} + \varepsilon_t = (1-\phi L)\varepsilon_t$

$$\varepsilon_t = \frac{c_t}{1-\phi L}$$

or $c_t(1+\phi L + \phi^2 L^2 + \cdots \infty) = \varepsilon_t$ $(|\phi|<1)$

or $c_t = -\phi c_{t-1} - \phi^2 c_{t-2} + \cdots + \varepsilon_t$ } $AR(\infty)$

(we cannot have ∞ lag terms practically but $\because$ $|\phi|<1$; $\therefore$ ϕ^k progressively falls, so we can cut down @ some k, w/o losing

Intuition: $\varepsilon_{t-k-1} = \dfrac{c_{t-k} - \varepsilon_{t-k}}{\phi}$ } plug this in c_t expression & recursively open it, to get all c_{t-i} terms

$\therefore$ $MA(1) \Leftrightarrow AR(\infty)$ if $|\phi|<1$

A time series is <u>invertible</u> if $|\phi| < 1$

How is $AR(1) \Leftrightarrow MA(\infty)$

$$C_t = \phi C_{t-1} + \cancel{\phi}\varepsilon_t$$

$$C_t = \phi L C_t + \varepsilon_t \Rightarrow \varepsilon_t = C_t(1-\phi L)$$

$$C_t = \frac{\varepsilon_t}{1-\phi L} = \varepsilon_t(1 + \phi L + \phi^2 L^2 + \dots \infty) \; ; \; |\phi| < 1$$

$$\text{or } C_t = \underbrace{\phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \dots}_{\infty \text{ lags}} + \varepsilon_t \; \Big\} \; MA(\infty)$$

($\because |\phi|<1 \therefore$ terms after some practical lag will be insignificant.

$$\begin{array}{ccc} C_t & \leftarrow & \varepsilon_t \\ \uparrow & & \\ C_{t-1} & \leftarrow & \varepsilon_{t-1} \\ \uparrow & & \\ C_{t-2} & \leftarrow & \varepsilon_{t-2} \\ \vdots & & \\ \vdots & - & \\ \uparrow & & \end{array}$$

$\dots + \phi_q \varepsilon_{t-q} + \varepsilon_t$

ARMA Model

Auto regressive moving average model.

ex ARMA (1, 1)

$\Rightarrow l_t = \beta_0 + \beta_1 l_{t-1} + \phi_1 \epsilon_{t-1} + \epsilon_t$

$l_t \Rightarrow$ no. of light bulbs manuf.——d this month

$\hat{l}_t = \beta_0 + \beta_1 l_{t-1} + \phi_1 \epsilon_{t-1}$

ARMA(p, q):

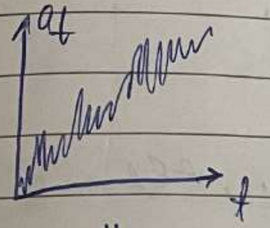
$y_t = \mu_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_p y_{t-p} + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q} + \epsilon_t$

$\Rightarrow$ ACF helps us get the MA order
PACF helps us get the AR order

integrated
↑
ARIMA Model

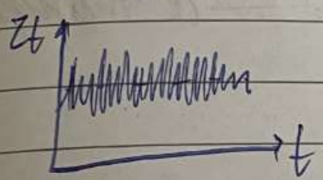
ex: you are an anchor seller

$a_t \Rightarrow$ no. of anchors sold/month



moving mean

⇓



$z_t = a_{t+1} - a_t$ (const mean) } stationary

ARIMA model is used when timeseries has a variable mean (moving average) &

so in ARIMA; instead of predicting time series itself; we predict differences of time series @ diff time stamps.

ARIMA (__ , __ , __)
p d q

ARIMA(1,1,1) $\Rightarrow$ $z_t = \phi_1 z_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$

ex ARIMA(2,2,3) $\Rightarrow$ $a_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \theta_3 \epsilon_{t-3} + \epsilon_t$

where $w_t = z_{t+1} - z_t$; where $z_t = a_{t+1} - a_t$
↓
second difference.

$d = k \Rightarrow z_t = a_t - a_{t-k}$

time

how to recover a_t

$$z_t = a_{t+1} - a_t ; \text{or} \quad a_{t+1} = a_t + z_t$$

$$\text{or } a_k = z_{k-1} + a_{k-1} = z_{k-1} + z_{k-2} + a_{k-2} = z_{k-1} + z_{k-2} + z_{k-3} + a_{k-3}$$

$$\text{or } a_k = \sum_{i=1}^{k-l} z_{k-i} + a_l$$

last a value for which we have the data.

Seasonality:

- periodicity

remaining seasonality;

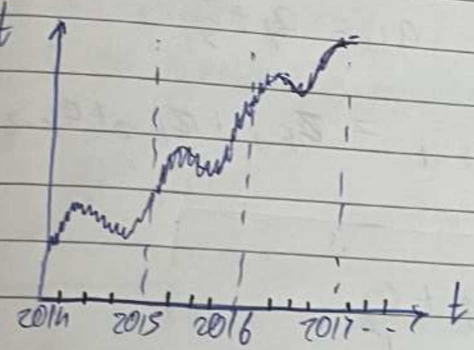
say y_t is sales of icecream perday $\Rightarrow$ ~~season~~ periodicity or seasonality of one year

so take $\quad z_t = y_{t+365} - y_t$

Seasonality vs. Cycles:

- A cycle is a ~~repeating~~ trend occuring over the course of ~~a year~~ a few years
- Cycles are not very predictable.

SARIMA Model



(Seasonal auto regressive Integrated moving average model)

A SARIMA model is given by 7 whole parameters | seasonal eqv's of (p d q)

$$ARIMA\ (p,d,q)(P,D,Q)_m \rightarrow \text{seasonal factor}$$

ex. $ARIMA(1,1,1)(1,1,1)_4 \Rightarrow$ | No. of periods within a year for the seasonality to repeat.

$$(1-L\phi_1)(1-L^4\Phi_1)(1-L)(1-L^4)y_t = (1+L\theta_1)(1+L^4\Theta_1)\varepsilon_t$$

LHS : AR & I terms
RHS : MA terms

ex. $ARIMA(1,0,0)(0,1,1)_4$:

$$(1-L\phi_1)(1-L^4)y_t = (1+L^4\Theta_1)\varepsilon_t$$

$$\Rightarrow (1-L\phi_1-L^4-L^5\phi_1)y_t = \varepsilon_t + \Theta_1\varepsilon_{t-4}$$

$$\Rightarrow \quad y_t = \phi_1 y_{t-1} + y_{t-4} + \phi_1 y_{t-5} + \Theta\varepsilon_{t-4} + \varepsilon_t$$

let $z_t = y_t - y_{t-4}$

or $\underbrace{z_t}_{I_4\ (D=1)} = \underbrace{\phi_1 z_{t-1}}_{AR_1\ (p=1)} + \underbrace{\Theta\varepsilon_{t-4}}_{MA_4\ (Q=1)} + \varepsilon_t$

ARCH Model

ARCH(1):

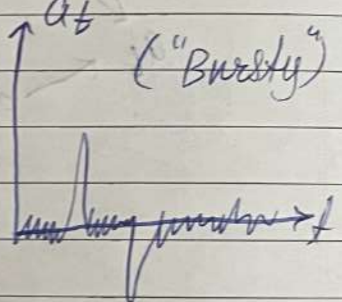
$$Var(\varepsilon_t) = \sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2$$

(When t.s' have variable volatility.

ARCH = Auto regressive conditional $\overset{\text{volatility}}{\text{Heteroskedasticity}}$

$$\varepsilon_t = \overset{\text{whitenoise}}{\omega_t} \sqrt{\alpha_0+\alpha_1\varepsilon_{t-1}^2} \; ; \; a_t = \varepsilon_t\sqrt{\alpha_0+\alpha_1 a_{t-1}^2} = \varepsilon_t\sigma_t$$

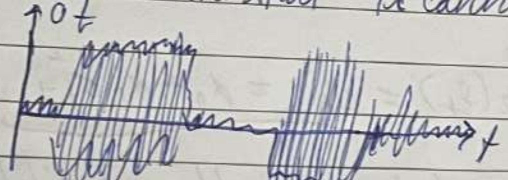
$$\underbrace{\varepsilon_t^2 = \omega_t^2\alpha_0 + \omega_t^2\alpha_1\varepsilon_{t-1}^2}_{\text{AR part}}$$



Use Correlogram (b/w ε_t & ε_{t+n}) to get which lags to consider generalised

GARCH Model

ARCH is Bursty, ~~we dont want that~~ it cannot model σ_t's like



here value stays at same val

GARCH(1,1)

$$a_t = \epsilon_t \overbrace{\sqrt{\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2}}^{\sigma_t} = \epsilon_t \sigma_t$$

$$\begin{tikzcd}
 & \epsilon_t \arrow[dl] & & a_{t-1} \arrow[dl] \\
a_t & & \sigma_t \arrow[ll] & \\
 & & & \sigma_{t-1} \arrow[ul]
\end{tikzcd}$$

} Causal diagram

VAR

Vector Autoregression

$$a_t = c_{11} a_{t-1} + c_{12} b_{t-1} + \varepsilon_{a,t}$$

$$b_t = c_{21} a_{t-1} + c_{22} b_{t-1} + \varepsilon_{b,t}$$

interdependent AR Model (2 vars) (VAR(1))

or $\begin{bmatrix} a_t \\ b_t \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} a_{t-1} \\ b_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{at} \\ \varepsilon_{bt} \end{bmatrix}$

or $\vec{f}_t = [C]\vec{f}_{t-1} + \vec{\varepsilon}_t$

Granger Causality

$$P_t = \phi_1 P_{t-1} + \phi_3 P_{t-3} + \varepsilon_t$$

$\cancel{\phi_3 P_{t-3}}$ $P_t = \phi_1 P_{t-1} + \phi_3 P_{t-3} + \psi x_{t-3} + \psi_{t-5} + \varepsilon_t$.