22/01/14 Applied Forecasting

Kozem Dohyot

web page rozem doyt tcd

monday 1 at LB01

thursday 12 at Salm chem

friday 3pm LB04

Steps in Forecasting

- Define what you want to forecast
- Gather data (all inputs)
- Do exploratory analysis (correlation)
- Fitting models select best model
- Use model to forecast future

Quantitative forecasting

use past data to predict

- hypothesis - continuity (change or no change)

- predicting temperature

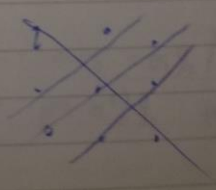
$$T_{t+1} = f\Big(\underset{\text{now}}{T_t},\ \underbrace{(\text{land use}_{T},\ P_t,\ T_{t-1\,\text{year}},\ \ldots)}_{\text{explanatory variables}}\Big) + \underset{\text{error}}{\varepsilon_t}$$

T = load(?)

response variable

Time Series $T_{t+1} = f(T_t, T_{t-1}, T_{t-2} \ldots T_{t-n})$

trend increasing, decreasing or seasonal



Forecastin

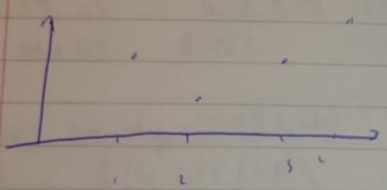
R

plot / Seasonal plot } tsdisplay() time series display
acf
pacf

lag max = 40 → do 40 lags.

Simple Exponential Smoothing SES.

$y_1 \quad y_2 \quad y_3 \quad \dots \quad y_n$ recorded data



$F = y_1$ F equals first value

α between 0 and 1

$F_2 = F_1 + \alpha(y_1 - F_1)$ error between real and forecasted
compensation for error

$F_{t+1} = F_t + \alpha(y_t - F_t)$

$F_3 = F_2 + \alpha(y_2 - F_2)$

$= y_1 + \alpha(y_2 - y_1)$

$F_3 = (1-\alpha)y_1 + \alpha y_2$

$F_4 = F_3 + \alpha(y_3 - F_3)$

$= (1-\alpha)y_1 + \alpha y_2 + \alpha[y_3 - (1-\alpha)y_1 - \alpha y_2]$

$= \alpha y_3 + [(1-\alpha) - \alpha(1-\alpha)]y_1 + (\alpha - \alpha^2)y_2$ weighted average of last 3

SES
Suitable for no trend or no seasonal component

Double Exponential Smoothing

- For data with a trend

b_t for slope of line l_t for level of line

29/09/14

Forecasting

Double Exponential Smoothing Algorithm
H(w) trend
no Seasonality
Trying to fit a line continuously to last point & predict next point

t	y_t	L_t	b_t	$F_t = L_{t-1} + b_{t-1}$	$y_t - F_t$
1	3	3	1	3	0
2	4	4	1	~~4~~	0
3	2	3.5.	0.85	5	3
4				4.35	x
5.				5.2	x
6				6.05	

1. $L_1 = y_1 = 3$ $\quad b_1 = y_2 - y_1 = 4-3 = 1$ $\quad F_1 = y_1 = 3$

$\alpha = 0.5$ $\quad \beta = 0.1$

2. $L_2 = \alpha y_2 + (1-\alpha)(L_1 + b_1)$ $\quad 0.5(4) + 0.5(3+1) = 4$

$b_2 = \beta(L_2 - L_1) + (1-\beta) b_1$ $\quad 0.1(4-3) + (1-0.1)1 = 1$

$F_2 = L_1 + b_1 = 3+1 = 4$

$y_t - F_1 \quad 4 - 4 = 0$

3. $F_3 = L_2 + b_2 = 4+1 = 5$

$L_3 = \alpha y_3 + (1-\alpha)(L_2 + b_2)$ $\quad 0.5(2) + (1-0.5)(4+1) = 3.5$

$b_3 = \beta(L_3 - L_2) + (1-\beta) b_2$ $\quad 0.1(3.5-4) + (1-0.1)(1) = 0.85$

$F_3 = y_t - F_t = 2 - 5 = -3$

4. $F_4 = L_3 + b_3 = 4.35$

$L_4 = \alpha y_4 + (1-\alpha)(L_3 + b_3)$ $\quad 0.5$

$F_5 = L_5 + 2b_2$

$F_{n+k} = L_n + k \cdot b_n$ $\quad F_5 = L_3 + 2b_3 = 3.5 + 2(0.85) = 5.2$

$F_6 = L_3 + 3b_3 = 3.5 + 3(0.85) = 6.05$

$SSE = 0^2 + 0^2 + (-3)^2 = 9 \quad (\alpha = 0.5 \quad \beta = 0.1)$

Pick α, β that minimise SSE
Does not account for relative error i.e. 0.1 vs 1 and 100 vs 101
big difference

Other criterion - mean absolute percent error MAPE

$$100 \frac{1}{n} \sum \left| \frac{y_t - F_t}{y_t} \right| = MAPE$$

SSE used in R

Root mean square error RMSE: $\sqrt{\frac{SSE}{N}}$

RMSE and SSE should give same α and β

SES (α) $\quad \alpha$ minimise $SSE(\alpha)$
DES (α, β) $\quad \alpha, \beta$ minimise $SSE(\alpha, \beta)$

SES: Time series with no trend & no seasonality
DES: Time series with trend & no seasonality

Two more algorithms to consider seasonality

1. Additive Holt Winters

Initialise $L_s = \frac{1}{s} \sum_{i=1}^{s} y_i$

beer data - s=12 for period in TS

$$b_s = \frac{1}{s} \left[\left(\frac{y_{s+1} - y_1}{s} \right) + \frac{y_{s+2} - y_2}{s} + + \dots \frac{y_{2s} - y_s}{s} \right) \right]$$

this year v time last year.

$S_i = y_i - L_s \qquad i = 1 \dots s \qquad \gamma$

$0 \le \alpha \le 1 \qquad 0 \le \beta \le 0 \qquad 0 \le \gamma \le 1$

$L_t = \alpha (y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1}) \qquad S_t = \gamma (y_t - L_t) + (1-\gamma)(S_{t-s})$

$b_t = \beta (L_t - L_{t-1}) + (1-\beta) b_{t-1} \qquad F_{t+1} = L_t + b_t + S_{t+1-s}$

$$F_{n+k} = L_n + k\,b_n + S_{n+k-s}$$

$$F_{n+k} = (L_n + k \cdot b_n)\,S_{n+k-s}$$

holtwinter (beer) holtwinter additive form

(Beta=False gamma=FALSE for SES)

1. look at data ⇒ question mark file name ⇒ know what it is about. Nature of data.
2. plot data tsdisplay(beer) in R Trend? Seasonal component? noise? ACF PACF
3. Select algorithm from SES, DES SHW_+, SHW_x
4. look at SSE and pick alg with lowest SSE used to compute forecast

/14

Forecasting

Holt winters additive

$\alpha = 0.5$ $\beta = 0.3$ $\gamma = 0.9$

Month	Production bar.	L_t	b_t	S_t	F_t
1	164			5.75	-
2	148			-10.25	-
3	152			-6.25	-
4	144			-14.25	
5	155			-3.25	
6	125			-33.25	
7	153			~~18.25~~ -5.25	
8	143			-12.25	
9	139			-20.15	
10	190			31.75	
11	192			33.75	
12	192	158.25	-0.65	33.75	
13	147	149.425	-3.1025	-1.6075	
14	133				
15	163				

$L_{13} = 0.5(147 - 5.75) + (0.5)(158.25 - 0.65) = 149.425$

$b_{13} = 0.3(\cancel{158.25}\ 149.425 - 158.25) + (1 - 0.3)(-0.65) = -3.1025$

$S_{13} = 0.9(147 - 149.425) + (1 - 0.9)(5.75) = -1.6075$

$F_{13} = 158.25 + (-0.65) + 5.75 =$

$F_{14} = 149.425 + (-3.1025) - 10.25 =$

6/10/14 Forecasting

Seasonal Holt Winters

$$F_{t+1} = L_t + b_t + S_t$$
$\alpha \quad \beta \quad \gamma$

Fitted value at $t+1$
"prediction" for y_{t+1}

$$\varepsilon_{t+1} = y_{t+1} - F_{t+1}$$

Pick $\hat{\gamma}\ \hat{\beta}\ \hat{\alpha}$ = min SSE $\sum_{t=1}^{n} (y_t - F_t)^2$

Linear Regression

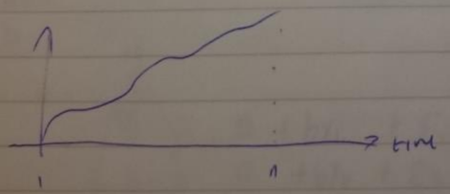
$$y_i = a + bx_i + \varepsilon_i \quad \forall i = 1 \dots n$$

response variable — explanatory var — error or noise.

ε_i's follow $\sim N(0, \sigma^2)$ (mean, variance)

$$p(\varepsilon) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp \frac{(-\varepsilon^2)}{2\sigma^2}$$

When $\sigma = 0$ dirac distribution



predict next value $\text{dow jones}_t = a + b\,\text{time}_t + \varepsilon_t$

$$\text{beer}_t = \beta_0 + \beta_1 \text{Jan}_t + \beta_2 \text{Feb}_t + \dots + \beta_{12} \text{December}_t + \varepsilon_t$$

$$\text{Jan}_t = \begin{cases} 1 & \text{if } t = \text{January} \\ 0 & \text{otherwise} \end{cases}$$

2

$t = \text{January} \Rightarrow beer_t = \beta_0 + \beta_1 + \varepsilon_t$

$t = \text{Dec} \quad beer = \beta_0 + \beta_{12} + \varepsilon_t$

$\{(x_i, y_i)\} \quad i \to 1..n \quad$ Best (a, b)?

Choose a, b that maximise the likelihood

$$p(\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n) = \prod_{i=1}^{n} p(\varepsilon_i) \qquad (y_i - a - bx_i)^2$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(\frac{-\varepsilon_i^2}{2\sigma^2}\right)$$

$$\hat{a}, \hat{b} = \arg\max \sum_{i=1}^{n} \log\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right) - \frac{\varepsilon_i^2}{2\sigma^2} = \arg\max \overset{\text{negative}}{-} \sum_{i=1}^{n} \varepsilon_i^2$$

$$\overset{\text{positive}}{=} \arg\min \sum_{i=1}^{n} \varepsilon_i^2$$

10/11 Forecasting

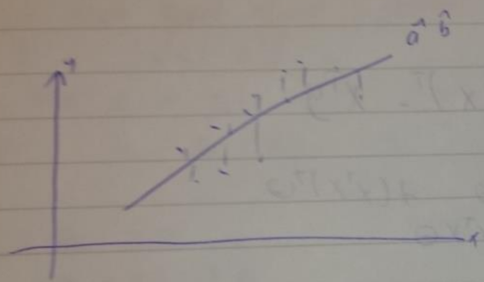
$y_i = a + bx_i + \varepsilon_i$

$\varepsilon_i \sim N(0, \sigma^2)$

$\varepsilon_i, \varepsilon_j$ independent $i \neq j$.

$(\hat{a}, \hat{b}) = \min (\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n)$

$= \max \prod_{i=1}^{n} p(\varepsilon_i) = \arg\min \sum_{i=1}^{n} \varepsilon_i^2$



SE squared distance from line

$\dfrac{d \sum \varepsilon_i^2}{da} = 0$

$\dfrac{d \sum \varepsilon_i^2}{db} = 0$

y

$y_1 = a + bx_1 + \varepsilon_1$

$y_2 = a + bx_2 + \varepsilon_2$

$\vdots$

$y_n = a + bx_n + \varepsilon_n$

$\vec{y} = X \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & x_n \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \vec{\varepsilon}$

$= \| \vec{\varepsilon} \|^2$

$= \vec{\varepsilon}^T \vec{\varepsilon} \quad = (\vec{y} - \vec{X}\theta)^T (\vec{y} - X\vec{\theta})$ $\binom{A}{B}$

$= \vec{y}^T \vec{y} - \vec{y}^T X\theta \quad - (X\theta)^T \vec{y} + (X\theta)^T (X\theta)$

$\theta^T X^T \vec{y}$

$\frac{d}{d\theta} \vec{y}^T y = 0$

$\frac{d}{d\theta} (y^T X)\theta = (y^T X)^T = X^T y$

$= X^T X \theta + (X^T X)^T \theta$

$= 2 X^T X \theta$

1. $y = a + bx + \varepsilon$
2. $y = a' + b'z + \varepsilon$
3. $y = a'' + b''x + c''z + \varepsilon$ least squar best model

$$\begin{cases} y_2 & a + by_1 + \varepsilon_2 \\ y_3 & a + by_2 + \varepsilon_3 \\ \\ y_n & a + by_{n-1} + \varepsilon_n \end{cases}$$ use past values

3/10/14 Forecasting

Auto regressive models AR 1

$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$ y_{t-1} as explanatory variable

AR 2

$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$ lag difference = 2 order = 2

Ar(p)

$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$ order p

Ar(1): Least Squares estimate of parameters

$\hat{\theta} = (X^T X)^{-1} X^T \vec{y}$

Time Series: $y_1 = 14$, $y_2 = 3$, $y_3 = 6$, $y_4 = 7$ $y_5 = 12$

Model 1: $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$

$3 = \phi_0 + \phi_1 14 + \varepsilon_2$
$6 = \phi_0 + \phi_1 3 + \varepsilon_3$
$7 = \phi_0 + \phi_1 6 + \varepsilon_4$
$12 = \phi_0 + \phi_1 7 + \varepsilon_5$

$X = \begin{bmatrix} 1 & 14 \\ 1 & 3 \\ 1 & 6 \\ 1 & 7 \end{bmatrix}$ $\theta = \begin{bmatrix} \phi_0 \\ \phi_1 \end{bmatrix}$ $\vec{y} = \begin{bmatrix} 3 \\ 6 \\ 7 \\ 12 \end{bmatrix}$

$\hat{\sigma}^2 = \frac{1}{4-2}(\varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 + \varepsilon_5^2)$

n-2

AR(1) $\hat{\sigma}^2 = \frac{1}{(n-1)-2} \sum_{i=2}^{n} \varepsilon_i^2$

(n-1) eqns

AR(2) (n-2) eqns ϕ_0, ϕ_1, ϕ_2 3 parameters

$\hat{\sigma}^2 = \frac{1}{(n-2)-3} \sum_{i=3}^{n} \varepsilon_i^2$

AR(1) $y_t = \hat{\phi}_0 + \hat{\phi}_1 y_{t-1} + \varepsilon_t$ $\quad$ $\varepsilon_t \sim (0, \hat{\sigma}^2)$

$E(\varepsilon_n)$

$\{y_t\}$ $t = 1 \ldots n$ Forecast?

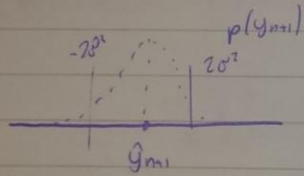
$$y_{n+1} = \boxed{\hat{\phi}_0 + \hat{\phi}_1 y_n} + \varepsilon_{n+1}$$

I only know this

Forecast $\hat{y}_{n+1}$

$$\begin{aligned} E(y_{n+1}) &= E(\hat{\phi}_0 + \hat{\phi}_1 y_n + \varepsilon_{n+1}) \\ &= E(\hat{\phi}_0 + \hat{\phi}_1 y_n) + E(\varepsilon_{n+1}) \\ &= \hat{\phi}_0 + \hat{\phi}_1 y_n + 0 \end{aligned}$$

95% prediction Interval for y_{n+1}?



$$y_{n+1} \in [\hat{y}_{n+1} - 2\sigma, \hat{y}_{n+1} + 2\sigma]$$

View of Situation
Information on other people - why they want it.
How will other departments rate

2 4 5 3 7 6 1 9 8 0

Ar(3).

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi y_{t-2} + \phi y_{t-3}$$

$$\begin{aligned}
3 &= \phi_0 + \phi_1 5 + \phi_2 4 + \phi_3 2 + \varepsilon_4 \\
7 &= \phi_0 + \phi_1 3 + \phi_2 5 + \phi_3 4 + \varepsilon_5 \\
6 &= \phi_0 + \phi_1 7 + \phi_2 3 + \phi_3 5 + \varepsilon_6 \\
1 &= \phi_0 + \phi_1 6 + \phi_2 7 + \phi_3 3 + \varepsilon_7 \\
9 &= \phi_0 + \phi_1 1 + \phi_2 6 + \phi_3 7 + \varepsilon_8 \\
8 &= \phi_0 + \phi_1 9 + \phi_2 1 + \phi_3 6 + \varepsilon_9 \\
0 &= \phi_1 + \phi_1 8 + \phi_2 9 + \phi_3 1 + \varepsilon_{10}
\end{aligned}$$

$$X = \begin{bmatrix} 1 & 5 & 4 & 2 \\ 1 & 3 & 5 & 4 \\ 1 & 7 & 3 & 5 \\ 1 & 6 & 7 & 3 \\ 1 & 1 & 6 & 7 \\ 1 & 9 & 1 & 6 \\ 1 & 8 & 9 & 1 \end{bmatrix} \quad d = \begin{bmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 3 \\ 7 \\ 6 \\ 1 \\ 9 \\ 8 \\ 0 \end{bmatrix}$$

$$\phi = (X^T X)^{-1} X^T \vec{y}$$

$$X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 3 & 7 & 6 & 1 & 9 & 8 \\ 4 & 5 & 3 & 7 & 6 & 1 & 4 \\ 2 & 4 & 5 & 3 & 7 & 6 & 1 \end{bmatrix}$$

$$X^T X = \begin{matrix} 46 & 44 & 58 & 65 & 44 & 62 & 79 \\ 44 & 51 & 57 & 66 & 62 & 57 & 74 \\ 58 & 57 & 84 & 79 & 61 & 97 & 89 \\ 65 & 66 & 79 & 95 & 70 & 80 & 115 \\ 44 & 62 & 61 & 70 & 87 & 58 & 70 \\ 62 & 57 & 97 & 80 & 58 & 119 & 88 \\ 79 & 74 & 89 & 115 & 70 & 88 & 147 \end{matrix}$$

/10/14 Forecasting

$y_{n+1} \in [\hat{y}_{n+1} - 2\sigma, \hat{y}_{n+1} + 2\sigma]$

$y_{n+2} = \hat{\phi}_0 + \hat{\phi}_1 y_{n+1} + \varepsilon_{n+2}$

$\hat{\phi}_0 + \hat{\phi}_1 y_n + \varepsilon_{n+1}$

$y_{n+2} = \boxed{\hat{\phi}_0 + \hat{\phi}_1\hat{\phi}_0 + \hat{\phi}_1^2 y_n} + \hat{\phi}_1\varepsilon_{n+1} + \varepsilon_{n+2}$ goes to zero

$\mathbb{E}(\text{error}) = 0$

$\mathbb{E}(\text{error}^2) = \sigma^2(1+\phi_1)^2$

$\mathbb{E}(\text{error}) = \mathbb{E}(\hat{\phi}_1 \varepsilon_{n+1} + \varepsilon_{n+2})$

$= \hat{\phi}_1 \mathbb{E}(\varepsilon_{n+1}) + \mathbb{E}(\varepsilon_{n+2})$

$\forall t \quad \varepsilon_t = N(0, \sigma^2)$

① $\mathbb{E}(\varepsilon_t) = 0 \qquad \mathbb{E}[(\varepsilon_t - \mathbb{E}(\varepsilon_t))^2]$

② $\sigma^2 = \mathbb{E}(\varepsilon_t^2)$

③ $\mathbb{E}(\varepsilon_{t_1}, \varepsilon_{t_2}) = 0$ not correlated

Variance of error term $\phi_1 \varepsilon_{n+1} + \varepsilon_{n+2}$?

$\mathbb{E}[(\hat{\phi}_1 \varepsilon_{n+1} + \varepsilon_{n+2})^2] = \mathbb{E}(\hat{\phi}_1^2 \varepsilon_{n+1}^2 + 2\hat{\phi}_1 \varepsilon_{n+1}\varepsilon_{n+2} + \varepsilon_{n+2}^2)$

$= \hat{\phi}_1^2 \underbrace{\mathbb{E}(\varepsilon_{n+1}^2)}_{\sigma^2} + 2\hat{\phi}_1 \mathbb{E}(\underbrace{\varepsilon_{n+1} + \varepsilon_{n+2}}_{0}) + \underbrace{\mathbb{E}(\varepsilon_{n+2}^2)}_{\sigma^2}$

$\hat{\phi}_1^2 \sigma^2 + \sigma^2$

$y_{n+2} \in [\hat{y}_{n+2} - 2\sqrt{\hat{\phi}_1^2+1}\,\sigma, \hat{y}_{n+2} + 2\sigma\sqrt{\hat{\phi}_1^2+1}$

$y_{n+3} \in [\hat{y}_{n+3} - 2\sqrt{\hat{\phi}_1^4 + \hat{\phi}_1^2 + 1}\,\sigma, \ldots]$

then 6, 8, 10 etc

$S = 1 + \phi_1^2 + \phi_1^4 + \phi_1^6 \ldots + \phi_1^{2n}$

Geometric Series if $|\phi_1| < 1$

AR(1) with $|\phi| < 1 \Rightarrow$ time series, no trend, no seasonality

$y = f(t) = a + bt \qquad f'(t) = b$

transform by differencing and apply the model to LS resids separately

General model like $y_{t+1} = y_t + \varepsilon_{t+1}$

/14

Forecasting

Auto regressive models - AR(p)

$$y_t - \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

Moving Average MA(1)

$$y_t = \underbrace{\phi_0 + \phi_1}_{\text{estimate}} \underbrace{\varepsilon_{t-1}}_{\text{explanation}} - \varepsilon_t.$$

$y_1 \; y_2 \; \dots \; y_n$ $\qquad \varepsilon_0 = 0$

$$y_1 = \phi_0 + \phi \varepsilon_0 + \varepsilon_1 = \phi_0 + \varepsilon_1$$

$$y_2 = \phi_0 + \phi_1 \varepsilon_1 + \varepsilon_2 = \phi_0 + \phi_1(y_1 - \phi_0) + \varepsilon_2$$

$$y_3 = \phi_0 + \phi_1 \varepsilon_2 + \varepsilon_3 = \phi_0 + \phi_1(y_2 - \phi_0 - \phi_1 y_1 + \phi_0 \phi_1) + \varepsilon_3$$

$$y_3 = \cancel{\phi_0 + \phi_1 y_2 - \phi_1 y_1 - \phi_1^2 y_2}$$

Non linear parameters ϕ_0, ϕ_1

uses all the past observations (MA) $\qquad$ T.S with no trend $|\phi_1| < 1$

less weight on first obs more on second etc because of $\phi_1, \phi^2, \phi^{\dots}$

Moving Average (q)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_1 y_{t-2} + \dots + \phi_p \varepsilon_{t-q} + \varepsilon_t$$

Exercises:

1. $y_t = \phi_0 + \phi_{12} y_{t-12} + \varepsilon_t \Rightarrow AR(12)$ $\qquad$ lag series. 12 steps behind $y_t \; y_{t-12}$
2. $y_t = w_0 + w_{12} \varepsilon_{t-12} + \varepsilon_t \Rightarrow MA(12)$ $\quad p \; q$
3. $y_t = c + d_{12} y_{t-12} + w_{12} \varepsilon_{t-12} + \varepsilon_t \Rightarrow ARMA(12,12)$
4. $y_t = \phi_0 + \phi_1 \varepsilon_{t-1} + \phi_{12} \varepsilon_{t-12} + \varepsilon_t \Rightarrow MA(12)$

$$y_t = \phi_0 + \underbrace{\phi_1 y_{t-1}}_{AR(1)} + \underbrace{\phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2}}_{MA(2)} + \varepsilon_t$$

$ARMA(\overset{p}{1}, \overset{q}{2})$ $\qquad p = 1 \; q = 2$

$$y_{t+1} = \phi_0 + \phi_1 y_{t-1} + \phi_3 \varepsilon_{t-3} + \varepsilon_{t+1} \qquad ARMA(\overset{p}{2}, \overset{q}{4})$$

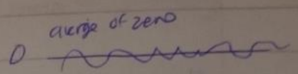
2

2 MA(1) model, what is expectation $E[y_t]$? Is it stationary mean or changing over time t.

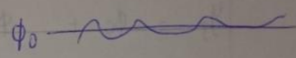
assume $\phi_0 = 0$

MA(1) $y_t = \phi_0 + \phi_1 \varepsilon_{t-1} + \varepsilon_t$

$$E[y_t] = E[\phi_1 \varepsilon_{t-1} + \varepsilon_t]$$
$$= \phi_1 \underbrace{E[\varepsilon_{t-1}]}_{0} + \underbrace{E[\varepsilon_t]}_{0}$$
$$= 0$$



$$E[y_t] = E[\phi_0 + \phi_1 \varepsilon_{t-1} + \varepsilon_t]$$
$$= \phi_0 + \phi_1 \underset{\uparrow 0}{E[\varepsilon_{t-1}]} + \underset{\uparrow 0}{E[\varepsilon_t]}$$
$$= \phi_0$$



Stationary in mean → expection of y_t will not change over time
if t.s has trend ⇒ no stationary mean
t.s with seasonality ⇒ not stationary in mean.

3 MA(2) model, what is expectation $E[y_t]$? Is it stationary in mean?

$$E[y_t] = E[\phi_0 + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \varepsilon_t]$$
$$= \phi_0 + \phi_1 \underset{\uparrow 0}{E[\varepsilon_{t-1}]} + \phi_2 \underset{\uparrow 0}{E[\varepsilon_{t-2}]} + \underset{\uparrow 0}{E[\varepsilon_t]}$$
$$= \phi_0$$

31/10/14 Forecasting

Backshifter operator - compute lag of timeseries

$B y_t = y_{t-1}$

$B.B\, y_t = y_{t-2} \quad B^2 y_t = y_{t-2}$

AR(1) $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$

backshift $\Rightarrow$ $y_t = \phi_0 + \phi_1 B y_t + \varepsilon_t$

$(1 - \phi_1 B) y_t = \phi_0 + \varepsilon_t$

MA(1) $y_t = \phi_0 - \phi_1 \underset{B\varepsilon_t}{\underset{\shortparallel}{\varepsilon_{t-1}}} + \varepsilon_t$

$y_t = \phi_0 + (1 - \phi_1 B)\varepsilon_t$

ARMA(1,1) $(1 - \phi B) y_t = \underset{\text{combined constant}}{\underset{\uparrow}{c}} + (1 - \phi B)\varepsilon_t$

MA(2) $y^t = \phi_0 + (1 - \phi_1 B - \phi_2 B^2)\varepsilon_t$

To differentiate y_t

$y_t = y_t - y_{t-1}$ difference/slope

$= (1 - B) y_t$

d=2 differentiating of order 2

$y''_t = y'_t - y'_{t-1}$

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

$y'_{t-1} = y_{t-1} - y_{t-2} = (1-B) y_{t-1}$

$(1-B) B y_t$

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

$= (1-B)^2 y_t$

$y_t \text{ d } 1 = (1-B)^d y^t$

ARIMA

i - integrated differentiation

ARIMA $\rightarrow (p, d, q)$

$(1-B)^d y_t$ – trend removal

constant from ϕ_0

$$\underbrace{(1-\phi_1 B-\phi_2 B^2-\dots-\phi_p B^p)}_{AR(p)}\underbrace{(1-B)^d}_{\text{trend removal}} y_t = c + (1-\theta_1 B-\theta_2 B^2 \dots \theta_q B^q)\varepsilon_t$$

/11/14 Forecasting

, integrated difference order k

ARIMA

auto regressive order p — moving average order q

ACF: Auto correlation function

ACF(k) : correlation between y_t and y_{t-k}

$$ACF(k) = \frac{\mathbb{E}\left[(y_t - \mathbb{E}(y_t))(y_{t-k} - \mathbb{E}(y_{t-k}))\right]}{\sqrt{\mathbb{E}\left[(y_t - \mathbb{E}(y_t))^2\right]\mathbb{E}\left[(y_{t-k} - \mathbb{E}(y_{t-k}))^2\right]}}$$

square root

For MA(1) what is ACF(k)?

Model: $y_t = \underbrace{\phi_0}_{=0} + \phi_1 \varepsilon_{t-1} + \varepsilon_t$ — noise on response variable

response variable, for simplicity, explanatory variable

$$\mathbb{E}(y_t) = \mathbb{E}(\phi_1 \varepsilon_{t-1} + \varepsilon_t)$$

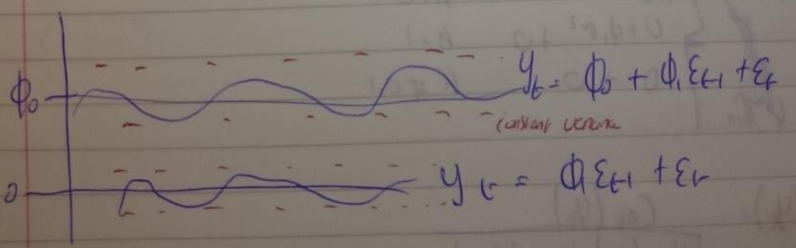
$$= \phi_1 \underbrace{\mathbb{E}(\varepsilon_{t-1})}_{=0} + \underbrace{\mathbb{E}(\varepsilon_t)}_{=0} \quad \mathbb{E} \text{ is a linear operator.}$$

$$= 0$$

Hyp for ARIMA $\mathbb{E}(\varepsilon_t) = 0 \ \forall t$

$Var(\varepsilon_t) = \sigma^2 \ \forall t$

ε_i independent for $i \neq i_2$



$$Var(y_t) = \mathbb{E}\{(y_t - \underbrace{\mathbb{E}(y_t)}_{=0})^2\} \quad \text{from first result}$$

$$= \mathbb{E}(y_t^2)$$

$$= \mathbb{E}\{(\phi \varepsilon_{t-1} + \varepsilon_t)^2\} \quad MA(1)$$

2

$= \mathbb{E}\left[\phi_1^2 \varepsilon_{t-1}^2 + 2\phi_1 \varepsilon_{t-1}\varepsilon_t + \varepsilon_t^2\right]$

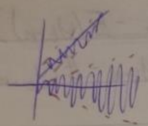
$= \phi_1^2 \underbrace{\mathbb{E}(\varepsilon_{t-1}^2)}_{=\sigma^2} + 2\phi_1 \underbrace{\mathbb{E}[\varepsilon_{t-1}\varepsilon_t]}_{=0 \text{ as independent}} + \underbrace{\mathbb{E}(\varepsilon_t^2)}_{=\sigma^2}$

Hypothesis for ARMA

$Var(\varepsilon_t) = \sigma^2$ $\quad = \phi_1^2\sigma^2 + \sigma^2$

$\mathbb{E}(\varepsilon_t - \underbrace{\mathbb{E}(\varepsilon_t)}_{=0})^2 = \sigma^2$ $\quad = Var(y_t) = \phi_1^2\sigma^2 + \sigma^2 \quad \forall t$

$\mathbb{E}(\varepsilon_t^2) = \sigma^2$

NOT:  not constant variance

$Cov(y_t) = \mathbb{E}\left[(y_t - \mathbb{E}(y_t))(y_{t-k} - \mathbb{E}(y_{t-k}))\right]$

$= \mathbb{E}[y_t - y_{t-k}]$

$= \mathbb{E}\left[(\phi_1\varepsilon_{t-1} + \varepsilon_t)(\phi_1 \varepsilon_{t-k-1} + \varepsilon_{t-k})\right]$

$= \mathbb{E}\left[\phi_1^2 \varepsilon_{t-1}\varepsilon_{t-k-1} + \phi_1 \varepsilon_{t-1}\varepsilon_{t-k} + \phi_1 \varepsilon_t \varepsilon_{t-k-1} + \varepsilon_t \varepsilon_{t-k}\right]$

$= \phi_1^2 \mathbb{E}[\varepsilon_{t-1}\varepsilon_{t-k-1}] + \phi_1 \mathbb{E}[\varepsilon_{t-1}\varepsilon_{t-k}] + \phi_1 \mathbb{E}[\varepsilon_t \varepsilon_{t-k-1}] + \mathbb{E}[\varepsilon_t\varepsilon_{t-k}]$

First term, when $k=0$.

$\mathbb{E}(\varepsilon_{t-1}^2) = \sigma^2$

$k \neq 0$ $\mathbb{E}[\varepsilon_{t-1}\varepsilon_{t-k-1}] = 0$

$$Cov(y_t) = \begin{cases} \phi_1^2\sigma^2 + \sigma^2 & k=0 \\ 0 + \phi_1\sigma^2 + 0 & k=1 \\ 0 + 0 & k \neq 0,1 \end{cases}$$

$$Corr(y_t) = \frac{Cov(y_t)}{\sqrt{Var(y_t)\, Var(y_{t-k})}}$$

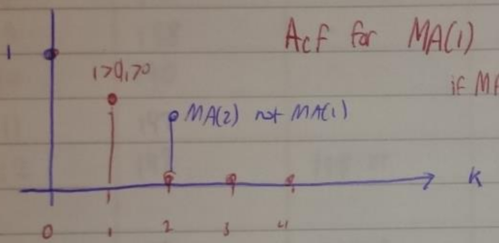
$(\phi_1^2\sigma^2 + \sigma^2) \quad (\phi_1^2\sigma^2 + \sigma^2$

/11/14 Forecasting

1. $\text{Corr}(y_t, y_{t-k}) = \dfrac{\phi_1^2\sigma^2 + \sigma^2}{\phi_1^2\sigma^2 + \sigma^2} = 1$ $\quad k=0 \quad$ (y_t and y_t)

2. $\text{Corr}(y_t, y_{t-k}) = \dfrac{\phi_1 \sigma^2}{\phi_1^2\sigma^2 + \sigma^2} = \dfrac{\phi_1}{\phi_1^2+1}$ $\quad k=1$

3. $\text{Corr}[y_t, y_{t-k}] = 0$ $\quad k \neq 0, k \neq 1.$



if MA(q) .. ACF(k) ≠ 0 all non zero then

MA(2) $y_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \varepsilon_t$ will have value for $k=2$ but 0 after.
Acf (MA(2)) 1, value, value, 0, 0 ... 0 for every value

AR

AR(1) = $y_t = \phi_1 y_{t-1} + \varepsilon_t$ Acf?

DAVID WETBRENT
12300644

Month No.	Production	Level L_t	Trend b_t	Seasonal S_t	Forecast F_t
1	164	-	-	1.036	-
2	148	-	-	0.935	-
3	152	-	-	0.961	-
4	144	-	-	0.910	-
5	155	-	-	0.979	-
6	125	-	-	0.790	-
7	153	-	-	0.967	-
8	148	-	-	0.935	-
9	138	-	-	0.872	-
10	190	-	-	1.201	-
11	192	-	-	1.213	-
12	192	158.25	-0.65	1.213	-
13	147	149.7	-3.01	0.987	163.33
14	133	144.5	-3.69	0.92	137.2
15	163	155.24	0.65	1.04	135.21
16					141.85

DAVID WEITBRECHT
12300644

Initialise: $L_s = \frac{1}{s} \sum_{i=1}^{s} y_i$ $\quad a = 0.5,\ \beta = 0.3\ \ \gamma = 0.9$

$b_s = \frac{1}{s} \left| \frac{y_{s+1}}{s} + \frac{y_{s+2} - y_2}{s} + \dots + \frac{y_{2s} - y_s}{s} \right|$

$S_i = y_t / L_s \quad i = 1 .. s.$

Compute for $t > s$

Level $L_t = \alpha\, y_t / S_{t-s} + (1-\alpha)(L_{t-1} + b_{t-1})$

Trend $b_t = \beta(L_t - L_{t-1}) + (1-\beta) b_{t-1}$

Seasonal, $S_t = \gamma\, y_t / L_t + (1-\gamma) S_{t-s}$

Forecast : $F_{t+1} = (L_t + b_t) S_{t+1-s}$

$S_1 = y_1 / L_{12} = 164 / 158.25 = 1.036$

$S_2 = y_2 / L_{12} = 148 / 158.25 = 0.935$

$S_3 = y_3 / L_{12} = 152 / 158.25 = 0.961$

$S_4 = y_5 / L_{12} = 144 / 158.21 = 0.910$

$S_5 = y_5 / L_{12} = 155 / 158.25 = 0.979$

$S_6 = y_6 / L_{12} = 125 / 158.25 = 0.790$

$S_7 = y_7 / 12 = 153 / 158.25 = 0.967$

$S_8 = y_8 / 12 = 148 / 158.25 = 0.935$

$S_9 = y_9 / L_{12} = 138 / 158.25 = 0.872$

$S_{10} = y_{10} / L_{12} = 190 / 158.25 = 1.201$

$S_{11} = y_{11} / L_{12} = 192 / 158.25 = 1.213$

$S_{12} = y_{12} / L_{12} = 192 / 158.25 = 1.213$

$F_{13} = (L_{12} + b_{12}) S_1 = (158.25 + (-0.65))(1.036) = 163.3264$

$L_{13} = \alpha\, y_{13} / S_{13-12} + (1 - 0.5)(L_{12} + b_{12})$

$0.5 (147 / 1.036) + (1 - 0.5)(158.25 + (-0.65)) = 149.72$

$b_t = \beta (L_{13} - L_{12}) + (1 - \beta) b_{12}$

$0.3 (149.7 - 158.25) + (0.7)(-0.65) = -3.01$

$S_{13} = \gamma \frac{y_{13}}{L_{13}} + (1-\gamma) S_1 = 0.9 \left(\frac{147}{149.7} \right) + 0.1 (1.036) = 0.987$

$$F_{14} = (L_{13} + b_{13}) S_{14-12} = (149.7 + (-3.01))\, 0.935 = 137.2$$

$$L_{14} = \alpha \frac{y_{14}}{S_2} + (1-\alpha)(L_{13} + b_{13}) = 0.5\,(133/0.935) + 0.5(149.7 + (-3.01)) = 144.5$$

$$b_{14} = \beta(L_{14} - L_{13}) + (1-\beta) b_{13} = 0.3(144.5 - 149.7) + (0.7)(-3.01) = -3.69$$

$$S_{14} = \gamma \, y_{14}/L_{14} + (1-\gamma) S_2 = 0.9(133/144.5) + 0.1(0.935) = 0.92$$

$$F_{15} = (L_{14} + b_{14}) S_3 = (144.5 + (-3.69))(0.961) = 135.21$$

$$L_{15} = \alpha \frac{y_{15}}{S_3} + (1-\alpha)(L_{14} + b_{14}) = 0.5(163/0.961) + 0.5(144.5 + (-3.69)) = 155.24$$

$$b_{15} = \beta(L_{15} - L_{14}) + (1-\beta) b_{14} = 0.3(155.24 - 144.5) + 0.7(-3.69) = 0.65$$

$$S_{15} = \gamma \, y_{15}/L_{15} + (1-\gamma) S_3 = 0.9(163/155.24) + 0.1(0.961) = 1.04$$

$$F_{16} = (L_{15} + b_{15}) S_4 = (155.24 + 0.65)\, 0.910 = 141.85$$

14/11/14

Forecasting

MA(1)
$$y_t = \phi_1 \varepsilon_{t-1} + \varepsilon_t$$
$$y_{t-1} = \phi_1 \varepsilon_{t-2} + \varepsilon_{t-1}$$

$$y_t = \phi_1 (y_{t-1} - \phi \varepsilon_{t-2}) + \varepsilon_t$$

$$y_{t+2} = \phi_1 \varepsilon_{t-3} + \varepsilon_{t-2}$$

$$y_t = \phi_1 y_{t-1} - \phi_1^2 (y_{t-2} - \phi \varepsilon_{t-3}) + \varepsilon_t$$
$$= \phi_1 y_{t-1} - \phi_1^2 y_{t-2} - \phi_1^3 \varepsilon_{t-3} + \varepsilon_t$$

AR(1)
PACF (k=1) $\quad y_t = \phi_0 + \boxed{\phi_1} y_{t-1} + \varepsilon_t$

pacf at k=1

AR(2) k=2 $\quad y_t = \phi_0 + \phi_1 y_{t-1} + \boxed{\phi_2} y_{t-1} + \varepsilon_t$

lag 2 of pacf at k=2

ARIMA (p=0, d=0, q=0) $\quad y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$

ARIMA (p=0, d=1, q=0) $\quad y_t = 0 + y_{t-1} + \varepsilon_t$

Remove trend by diffing
d < 3 dont go beyond 3

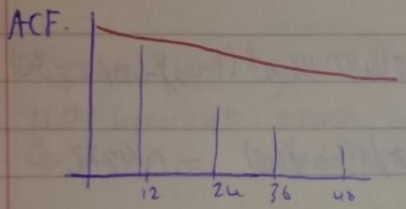
8/11/14 Forecasting

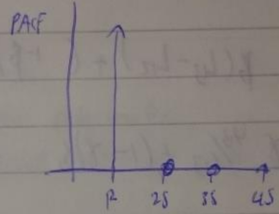
1. ARMA $(0,0,0)(1,0,0)_{12}$

2. ARIMA$(0,0,0)(0,1,0)_{12}$

What is differen.

1. $y_t = c + \Phi_1 y_{t-12} + \varepsilon_t$ Difference is we have no coefficient
2. $y_t = c + y_{t-12} + \varepsilon_t$





In autoregressive exponentially decrease in Acf of AR model

have spike because of y_{t-12}
height of this is proportional to Φ_1
in D case it should be close to one

Red line is D model.
same story between Φ=1 and p=1

Difference between D and P is Acf if exponential or general decrease

2

Q3 Assignment Q3 correction

$S_1 = \frac{y_1}{L_{12}} = \frac{64}{158.25} = 1.0363$

$S_2 = y_2/L_{12} = 148/158.25 = 0.93$

$S_{12} = y_{12}/L_{12} = 192/158.25 = 1.2133$

$L_{13} = \alpha \, y_{13}/S_1 + (1-\alpha)(L_{12}+b_{12}) = 0.5 \cdot 147/1.0365 + (1-0.5)(158.25-0.65) = 149.72$

$b_{13} = \beta(L_{13}-L_{12}) + (1-\beta)b_{12} = 0.3(149.72-158.25) + (1-0.3)(-0.65) = -3.01$

$S_{13} = \gamma \, y_{13}/L_{13} + (1-\gamma)S_1 = 0.9(147/149.72) + (1-0.9)(1.03) = 0.9872$

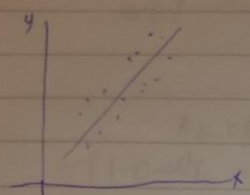
$F_{13} = (L_{12}+b_{12})S_1 = (158.25-0.65)(1.0363) = 163.52$

$F_{14} = (L_{13}+b_{13})S_2 = (149.72 + -3.01)(0.93) = 137.20$

Compare forecast values to 12 months before for a rough check ecasting

number of significant values on pacf is telling you order p of ar(p)

Overfitting



$y = b + ax + \varepsilon$ ← if true model (c=0 in other model and d=0)

can propose another model:

$y = b + ax + cx^2 + \varepsilon$

$y = b + ax + cx^2 + dx^3 + \varepsilon$

In practice c,d greater than 0 say 0.01

Is this significant enough to incorporate into model?

Don't want to over fit and explain general randomness, want to explain fit not noise

ACF(1) for MA(1)

Criteria to select best ARIMA model

AIC = $-2\,\text{Log}(L) + 2M$

BIC = $-2\log(L) + m\log n$

$y_1 \dots y_n$

residuals $\varepsilon_1 \dots \varepsilon_n$

Likelihood is $p(\varepsilon_1 \dots \varepsilon_n) = \prod p(\varepsilon_i)$ (independence ε_i and ε_j $i \neq j$)

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\varepsilon^2}{2\sigma^2}\right)$$

(normally distributed with mean 0 and fixed variance σ^2)

$$-2\,\text{Log}(\text{likelihood}) = -2\,\text{Log}\left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\varepsilon_i^2}{2\sigma^2}\right)\right)$$

$$= -2\sum_{i=1}^{n} \log\left(\frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\varepsilon_i^2}{2\sigma}\right)\right)$$

$$= -2\sum_{i=1}^{n} -\log(\sqrt{2\pi}\sigma) - \frac{\varepsilon_i^2}{2\sigma^2}$$

$$= 2 \left(n\log(\sqrt{2\pi}\sigma) + \frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^2 \right)$$

$$2n\log(\sqrt{2\pi}\sigma) + \frac{1}{\sigma^2}\sum_{i=1}^{n}\varepsilon_i^2$$

~~AIC 2nlog(√2π σ~~

$$AIC = n\log(2\pi\sigma^2) + \frac{\sum_{i=1}^{n}\varepsilon_i^2}{\sigma^2} + 2m$$

Find p,d,q AIC or BIC are minimized. Stick to same method whole time

pdq ↑ (big) will reduce ↓ $\Sigma\varepsilon_i^2$ but (2m) will increase and (mlogn) ↑

Show: $-2\log(L) = n(1+\log(2\pi)) + n\log(s^2)$

$$s^2 \approx \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2$$

$$ns^2 \simeq \sum_{i=1}^{n}\varepsilon_i^2$$

$$\sum_{i=1}^{n}\varepsilon_i^2 \simeq ns^2$$

$$s^2 \simeq \sigma^2$$

$$-2\log(L) = n\log(2\pi\sigma^2) + \frac{\overbrace{\sum_{i=1}^{n}\varepsilon_i^2}^{ns^2}}{\underbrace{\sigma^2}_{s^2}}$$

$$= n\log(2\pi) + n\log(\sigma^2) + n$$

$$= n(1+\log(2\pi)) + n\log(\sigma^2)$$

$$= n(1+\log(2\pi)) + n\log(s^2) \quad \text{Q.E.D.}$$

21/11/14 Forecasting

ARIMA(p d q)

ARIMA(1,1,1)

$$\underbrace{(1-0.4B)}_{ar(1)}\underbrace{(1-B)}_{\text{differencing } d=1}y_t = 0.1 + \underbrace{(1-0.9B)}_{MA(1)}\varepsilon_t$$

$x_t = (1-B)y_t = y_t - y_{t-1}$

$(1-0.4B)x_t = 0.1 + (1-0.9B)\varepsilon_t$

Time	y_t	x_t	$\hat{x}_t$	ε_t	$\hat{y}_t$
1	9.5	$x_1 = y_1 - y_0$	–		
2	13.7	$y_2 - y_1 = 4.2$	0	$x_t - \hat{x}_t = 4.2$	$\hat{y}_2 = 9.5$
3	8.7	$x_3 = y_3 - y_2 = -5$	−2	$-5 - -2 = -3$	11.7
4	16.1	7.4	0.8	6.6	9.5

$x_t - 0.4x_{t-1} = 0.1 + \varepsilon_t - 0.9\varepsilon_{t-1}$

$x_t = \underbrace{\boxed{0.1 + 0.4x_{t-1} - 0.9\varepsilon_{t-1}}}_{\hat{x}_t} + \varepsilon_t$

$\hat{y}_t = \hat{x}_t + y_{t-1} \quad \hat{y}_2 \; 0 + 9.5$

$\hat{x}_3 = 0.1 + 0.4\overset{x_2}{(4.2)} - 0.9\overset{\varepsilon_2}{(4.2)} = -2$

$\varepsilon_3 = -5 - (-2) = -3$

$\hat{y}_3 = \hat{x}_3 + y_2 = -2 + 13.7 = 11.7$

$x_4 = y_4 - y_3 = 16.1 - 8.7 = 7.4$

$\hat{x}_4 = 0.1 + 0.4(-5) - 0.9(-3) \quad 0.1 - 2 + 2.7 = 0.8$

$\varepsilon_4 = x_4 - \hat{x}_4 = 7.4 - 0.8 = 6.6$

$\hat{y}_4 = \hat{x}_4 + y_3 = 0.8 + 8.7 = 9.5$

Beer data, arima mod

building model some form of βy_{t-12} $(y_{t-12}$ or y_{t-ks} (12 months prior).

in similar fashion can use ε_{t-12}, ε_{t-24} etc.

$$y_t = a + \beta y_{t-12} + \varepsilon_t$$

$\searrow AR(p=12)$

Season AR, SAR(P=1) capital P=1 only using one y_{t-x} not all 12 y_t's

$$SAR_{12}(P=1) \quad y_t = \alpha + \beta \overset{\text{backshift}}{B^{12}} y_t + \varepsilon_t$$

$$(1-\beta B^{12})y_t = \alpha + \varepsilon_t$$

$$SAR_{12}(P=2) \quad y_t = \alpha + \beta_1 y_{t-12} + \beta_2 y_{t-24} + \varepsilon_t$$

$$(1-\beta_1 B^{12} - \beta_2 B^{24})y_t = \alpha + \varepsilon_t$$

24/11/14 Forecasting

ARIMA (p, d, q) complexity $m = p+q$

SARIMA $(p, d, q)(P, D, Q)_s$ $m = p+q+P+Q$

$$\text{SARIMA}(p=0, d=0, q=0)(P=1, D=0, Q=0)_{s=4}$$

$$(1)(1-\beta_1 B^{4(1)})(1)(1)y_t = c + (1)(1)\varepsilon_t$$

$$(1-\beta_1 B^4)y_t = c + \varepsilon_t$$

$$y_t = \beta_1 y_{t-4} + c + \varepsilon_t$$

$\beta_1 = 1$ here

$$\text{SARIMA}(p=0, d=0, q=0)(P=0, D=1, Q=0)_{s=4}$$

$$(1)(1)(1)(1-B^4)^1 y_t = c + (1)(1)\varepsilon_t$$

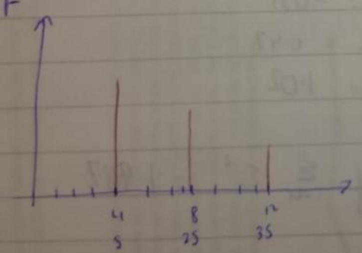
$$y_t = c + y_{t-4} + \varepsilon_t$$

AR $y_t = c + \beta_1 y_{t-4} + \varepsilon_t$

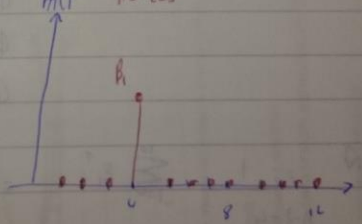
$$y_t = c + 0y_{t-1} + 0y_{t-2} + 0y_{t-3} + \beta_1 y_{t-4} + 0y_{t-5} \dots + \varepsilon_t$$

PACF coefs

ACF



PACF



Major spike S, 2S = 0, 3S = 0, 4S = 0 ...

Moving Average

Spike at S on ACF : small values otherwise

Every S values on PACF "

P D Q ≤ 3

3

Assignment Correction

Q4 A $y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t$ AR(p=3)

B $y_4 = c + \phi_1 \overset{5}{y_3} + \phi_2 \overset{4}{y_2} + \phi_3 \overset{7}{y_1} + \varepsilon_4$

$y_5 = c + \phi_1 y_4 + \phi_2 y_3 + \phi_3 y_2 + \varepsilon_5$

$y_6 = c + \phi_1 y_5 + \phi_2 y_4 + \phi_3 y_3 + \varepsilon_6$

$\vdots$

$y_{10} = c + \phi_1 y_9 + \phi_2 y_8 + \phi_3 y_7 + \varepsilon_{10}$

$$\underset{\vec{y}}{\begin{bmatrix} 3 \\ 7 \\ 6 \\ 1 \\ 9 \\ 8 \\ 0 \end{bmatrix}} = \underset{X}{\begin{bmatrix} 1 & 5 & 4 & 2 \\ 1 & 3 & 5 & 4 \\ 1 & 7 & 3 & 5 \\ 1 & 6 & 7 & 3 \\ 1 & 1 & 6 & 7 \\ 1 & 9 & 1 & 6 \\ 1 & 8 & 9 & 1 \end{bmatrix}} \underset{\theta}{\begin{bmatrix} c \\ \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}} + \underset{\varepsilon}{\begin{bmatrix} \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \end{bmatrix}}$$

$\hat{\theta} = (X^T X)^{-1} X^T \vec{y}$

$\hat{\theta}$: $\hat{c} = 5.14$

$\phi_1 = -0.35$

$\phi_2 = -0.47$

$\phi_3 = 1.02$

C $\hat{\sigma}^2 = \dfrac{1}{n-3-4} \sum_{i=4}^{10} \varepsilon_i^2 = \dfrac{1}{3} \sum_{i=4}^{10} \varepsilon_i^2 = 1.9937$