Lecture Twenty-One

MA(q) models

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q},$$
$$t = \ldots, -3, -2, -1, 0 \ldots$$

$\{\varepsilon_t\} \overset{iid}{\sim} N(0, \sigma^2)$

$q=1$: $y_t = \boxed{\mu} + \underset{\sim}{\varepsilon_t} + \theta_1 \underset{\sim}{\varepsilon_{t-1}}$ ← Slutzky 1937 Summation of Random Causes

$\theta_1 = 0 \rightarrow y_t \overset{iid}{\sim} N(\mu, \sigma^2)$

$\theta_1 > 0 \rightarrow$ positive correlation between y_t & y_{t+1}

$\theta_1 < 0 \rightarrow$ negative correlation between y_t & y_{t-1}

$$\mathrm{Corr}(y_t, y_{t+1}) = \frac{\theta_1}{1+\theta_1^2}$$

$$\mathrm{Corr}(y_t, y_{t+h}) = 0 \quad \text{for } h \geq 2$$

Sample ACF

$y_1 \ldots y_n$

sample ACF(h)

= sample correlation between y_t & y_{t+h}

$= \begin{matrix} (y_1, y_{1+h}) \\ (y_2, y_{2+h}) \\ \vdots \\ (y_{n-h}, y_n) \end{matrix}$

Theoretical ACF for a Stationary Time Series Model

$\rho(h) = \mathrm{Corr}(y_t, y_{t+h})$

→ cannot depend on t

$$(a_1, b_1) \; \ldots \; (a_m, b_m)$$

$$\text{Correlation} = \frac{\sum_{i=1}^{m} (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^{m} (a_i - \bar{a})^2 \sum_{i=1}^{m} (b_i - \bar{b})^2}}$$

$m = n - h$

$a_i = y_i$

$b_i = y_{i+h}$

$$\bar{a} = \frac{1}{n-h} \sum_{i=1}^{n-h} y_i \approx \bar{y}, \quad \bar{b} = \sum_{i=1}^{n-h} y_{i+h} \approx \bar{y}$$

$$\frac{\sum_{t=1}^{n-h} (y_t - \bar{y})(y_{t+h} - \bar{y})}{\sqrt{\underbrace{\sum_{t=1}^{n-h} (y_t - \bar{y})^2}\ \underbrace{\sum_{t=1}^{n-h} (y_{t+h} - \bar{y})^2}}}$$

Replace by $\sum_{t=1}^{n} (y_t - \bar{y})^2$

$$\text{Sample ACF}(h) = \frac{\sum_{t=1}^{n-h} (y_t - \bar{y})(y_{t+h} - \bar{y})}{\sqrt{\sum_{t=1}^{n} (y_t - \bar{y})^2 \sum_{t=1}^{n} (y_t - \bar{y})^2}}$$

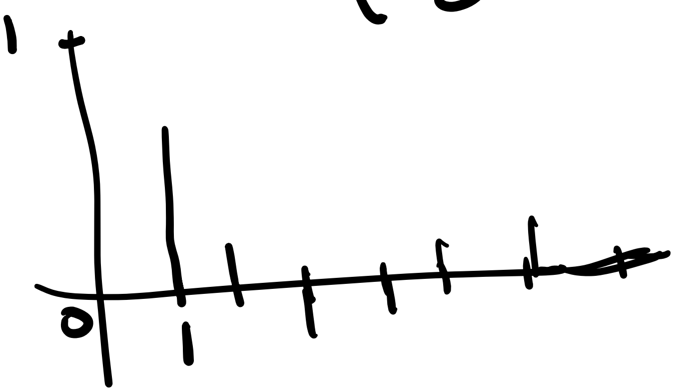
$$= \frac{\sum_{t=1}^{n-h} (y_t - \bar{y})(y_{t+h} - \bar{y})}{\sum_{t=1}^{n} (y_t - \bar{y})^2} \longrightarrow \text{Calculated from Data}$$

$\hat{\rho}(h)$

Stationary Model : $\rho(h)$: ACF of the model

MA(1) model: $y_t = \mu + \varepsilon_t + \theta\, \varepsilon_{t-1}$

$$\rho(h) = \begin{cases} 1 & \text{if } h = 0 \\ \dfrac{\theta}{1+\theta^2} & \text{if } |h| = 1 \\ 0 & \text{if } |h| \geq 2 \end{cases}$$



→ looks like MA(1)

MA(q) model: $y_t = \underbrace{\mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}}$

ALWAYS CAUSAL & STATIONARY

$$\rho(h) = \begin{cases} \text{something} & \text{if } |h| \leq q \\ 0 & \text{if } |h| > q \end{cases}$$

MA(1) vs AR(1)

$$\underbrace{y_t = \mu + \varepsilon_t + \theta\, \varepsilon_{t-1}}_{\text{ALWAYS STATIONARY}}$$

$$\underbrace{y_t - \phi_1 y_{t-1}}_{\text{NOT NECESSARILY CAUSAL STATIONARY}} = \phi_0 + \varepsilon_t$$

Parameter Estimation:

ϕ_0, ϕ_1, σ

$y = (y_t)$

$X = [1, y_{t-1}]$

sm.OLS(y, X).fit()

Likelihood:

$$(y_1 \, \cdots \, y_n) \sim N\left(\begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}, \begin{bmatrix} \ddots \end{bmatrix} \right)$$

$$\mathrm{var}(y_t) = \mathrm{var}(\mu + \varepsilon_t + \theta \varepsilon_{t-1})$$
$$= \sigma^2 (1+\theta^2)$$

$$\mathrm{Cov}(y_t, y_{t+1}) = \theta \sigma^2$$

$$\Sigma = \begin{bmatrix} \sigma^2(1+\theta^2) & \theta\sigma^2 & 0 & \cdots & & 0 \\ \theta\sigma^2 & \sigma^2(1+\theta^2) & \theta\sigma^2 & 0 & \cdots & 0 \\ 0 & \theta\sigma^2 & \sigma^2(1+\theta^2) & \theta\sigma^2 & 0 \cdots & 0 \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ 0 & \cdots & & 0 & \theta\sigma^2 & \sigma^2(1+\theta^2) \end{bmatrix}$$ ✓

$$\frac{1}{(\sqrt{2\pi})^n \sqrt{\det \Sigma}} \exp\left[-\frac{1}{2} (y-\mu)^T \Sigma^{-1} (y-\mu) \right]$$

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \mu = \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}$$

ARIMA → function in statsmodels.

AR model: $y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p} = \phi_0 + \varepsilon_t$

MA : $y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$

ARMA(p, q)

$y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p} = \phi_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$

$\boxed{\text{ARIMA}(p, d, q)}$

differencing

take data y_t, difference d times & then use ARMA(p, q)

arima(data, order = (p, d, q))

MA(q): arima(data, order = (0, 0, q))

MA(q) to $y_t - y_{t-1}$:

arima(data, order = (0, 1, q))

arima(differenced data, order = (0, 0, q))

AR(p): arima(data, order = (p, 0, 0))

AR(p) AutoReg

$\phi(B) y_t = \phi_0 + \varepsilon_t$

$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p$

$\longrightarrow$ AR(p) polynomial

Roots of $\phi(z)$:

Moduli of every root $> 1 \rightarrow$ Causal Stationary Regime

$$y_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots$$

$\underline{AR(1)}$: $|\phi_1| < 1$

$$y_t = \frac{\phi_0}{1-\phi_1} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \phi_1^3 \varepsilon_{t-3} + \cdots$$

$\underbrace{\text{ARMA}}$ 2 $\underbrace{\text{MA}}$