ARIMA with/without Drift.

ARIMA :

$$Y_t = m_t + X_t$$

m_t: trend — random or non random.

X_t: stationary series

If m_t is random walk.

$$\nabla Y_t = \nabla m_t + \nabla X_t$$

still stationary.

iid Normal errors $N(0, \sigma^2)$

model with ARMA(p, q) mean 0.

$\longrightarrow$ Y_t is ARIMA(p, 1, q) without drift.

If m_t is linear non-random trend

$$m_t = a + bt$$

$$\nabla m_t = b.$$

$$\nabla Y_t = \underbrace{b + \underbrace{\nabla X_t}_{\text{model with } \mathrm{ARMA}(p,q)}}_{\mathrm{ARMA}(p,q) \text{ with mean } b}$$

$\Rightarrow$ Y_t is $\mathrm{ARIMA}(p,1,q)$ with drift.

If m_t is RW with drift.

$$m_t = \sum_{i=1}^{t} e_i \qquad e_i \sim N(\underset{\substack{\uparrow \\ \text{drift}}}{\delta}, \sigma^2)$$

$$\nabla Y_t = \underbrace{\underset{\substack{\uparrow \\ N(\delta,\sigma^2)}}{\nabla m_t} + \underset{\substack{\uparrow \\ \text{stationary}}}{\nabla X_t}}_{\substack{\text{Model with} \\ \text{ARMA}(p,q) \text{ with} \\ \text{mean } \delta}} \quad \Rightarrow \quad \begin{array}{c} Y_t \text{ is} \\ \text{ARIMA}(p,1,q) \\ \text{with drift.} \end{array}$$

If ∇Y_t is ARMA(p, q) with mean μ,

1. use $\bar{X}$. determine if μ is significantly different from 0.

$$\bar{X} \sim N\left(\mu, \sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \gamma(h)\right)$$

(2) when estimating ϕ_i and θ_i, estimate μ as well.

```
arima( diff (X) )
or
Arima ( X , order = c(1,1,1), include.drift =T)
```

in forecast package.

Determine if $\hat{\mu}$ is significantly different from 0.

<u>Warning</u> :

If you use

arima(X, d= 1)

then it is assumed that drift = 0

after differencing.

Regression with Time-Series Errors

Regression

$$Y_t = X_{1t}\beta_1 + X_{2t}\beta_2 + X_{3t}\beta_3 + e_t$$

Covariates

$$e_t \sim IID(0, \sigma^2)$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} X_{11} & X_{21} & X_{31} \\ \vdots & \vdots & \vdots \\ X_{1n} & X_{2n} & X_{3n} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$$

Design Matrix

errors

parameters

$$\underline{Y} = \underline{X}\,\underline{\beta} + \underline{e}$$

OLS

$$\underline{Y} = \mathbb{X}\underline{\beta} + \underline{e}$$

$$\underline{\hat{e}} = \underline{Y} - \mathbb{X}\underline{\hat{\beta}}$$

residuals

OLS $\doteq$ minimum sum of squared residuals (errors).

$$SSE = \|\underline{\hat{e}}\|^2$$

$$= (\underline{Y} - \mathbb{X}\underline{\hat{\beta}})^T(\underline{Y} - \mathbb{X}\underline{\hat{\beta}})$$

$$(AB)^T = B^T A^T$$
$$(A+B)^T = A^T + B^T$$

$$\|\hat{\underline{e}}\|^2 = (\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}})^T(\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}})$$

$$= \underbrace{\underline{Y}^T\underline{Y}} - \underbrace{\hat{\underline{\beta}}^T \underline{\underline{X}}^T \underline{Y}}_{\text{scalar}} - \underbrace{\underline{Y}^T \underline{\underline{X}}\hat{\underline{\beta}}}_{\text{scalar}} + \underbrace{\hat{\underline{\beta}}^T \underline{\underline{X}}^T \underline{\underline{X}}\hat{\underline{\beta}}}$$

find $\hat{\underline{\beta}}$ that minimize $\|\hat{\underline{e}}\|^2$

$$\frac{d}{d\hat{\underline{\beta}}}\|\hat{\underline{e}}\|^2 = -2\underline{Y}^T\underline{\underline{X}} + 2\hat{\underline{\beta}}^T\underline{\underline{X}}^T\underline{\underline{X}} := 0.$$

$$\boxed{\hat{\underline{\beta}} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T\underline{Y}}$$ OLS

if $\underline{A}^T\underline{\beta}$ = scalor

$= (\underline{A}^T\underline{\beta})^T = \underline{\beta}^T\underline{A}$

$$\hat{\underline{\beta}} = (\mathbb{X}^T \mathbb{X})^{-1} \mathbb{X}^T \underline{Y}$$

$$= (\mathbb{X}^T \mathbb{X})^{-1} \mathbb{X}^T (\mathbb{X} \underline{\beta} + \underline{e})$$

If $\mathbb{X}$ is non random,

$$E(\hat{\underline{\beta}}) = (\mathbb{X}^T \mathbb{X})^{-1} \mathbb{X}^T \mathbb{X} \Big(\underline{\beta} + \underset{\substack{\uparrow \\ 0\,.}}{E(\underline{e})}\Big) = \underline{\beta}$$

$$V(\hat{\underline{\beta}}) = E\left[(\hat{\underline{\beta}} - \underline{\beta})(\hat{\underline{\beta}} - \underline{\beta})^T\right]$$

$$= E\left\{ \left[(\mathbb{X}^T \mathbb{X})^{-1} \mathbb{X}^T \underline{e} \right] \left[(\mathbb{X}^T \mathbb{X})^{-1} \mathbb{X}^T \underline{e} \right]^T \right\}$$

$$= (\mathbb{X}^T \mathbb{X})^{-1} \mathbb{X}^T \underbrace{E(\underline{e}\, \underline{e}^T)}_{\text{Cov matrix of } e_t .} \mathbb{X} \underbrace{(\mathbb{X}^T \mathbb{X})^{-1}}_{\text{symmetric}}$$

If $\Sigma = E(\underline{e}^{*} \underline{e}^{T}) = E \begin{bmatrix} e_1 e_1 & e_2 \cdot e_1 & e_3 \cdot e_1 & \cdots & e_n \cdot e_1 \\ \vdots & e_2 \cdot e_2 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ e_1 e_n & \cdots & \cdots & & e_n \cdot e_n \end{bmatrix}$

Cov matrix of errors

$= \mathbb{I} \sigma^2$ (i.e. if $e_t \sim iid(0, \sigma^2)$)

Then

$$Var(\hat{\beta}_{OLS}) = \sigma^2 (\mathbb{X}^T \mathbb{X})^{-1}$$

Regression with correlated errors

$$\underline{Y} = \mathbb{X}\underline{\beta} + \underline{\varepsilon}. \qquad \varepsilon_t \sim ARMA(p,q)$$

If you use OLS.

$$\underline{\hat{\beta}} = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\underline{Y}$$

then,

$$E(\underline{\hat{\beta}}_{OLS}) = \underline{\beta} \qquad \text{unbiased}$$

$$V(\underline{\hat{\beta}}_{OLS}) = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}\underbrace{E(\underline{\varepsilon}\,\underline{\varepsilon}^T)}_{\Gamma_n \text{ cov matrix of ARMA}}\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}.$$

GLS

Generalized Least Squares.

Instead of minimizing $\|\hat{\underline{e}}\|^2 = (\underline{Y} - \mathbb{X}\underline{\beta})^T(\underline{Y} - \mathbb{X}\underline{\beta})$,

minimize $(\underline{Y} - \mathbb{X}\underline{\beta})^T \boldsymbol{\Pi}_n^{-1} (\underline{Y} - \mathbb{X}\hat{\underline{\beta}})$.

$$E(\hat{\beta}_{GLS}) = \underline{\beta} \quad \text{if } \mathbb{X} \text{ non-random.}$$

$$V(\hat{\beta}_{GLS}) = (\mathbb{X}^T \boldsymbol{\Pi}_n^{-1} \mathbb{X})^{-1}$$

$$\hat{\underline{\beta}}_{GLS} = (\mathbb{X}^T \boldsymbol{\Pi}_n^{-1} \mathbb{X})^{-1} \mathbb{X}^T \boldsymbol{\Pi}_n^{-1} \underline{Y}$$

It can be shown that GLS is Best Linear Unbiased Estimator (BLUE) of β.

GLS : need Π_u $\rightarrow$ need $\phi_1 \ldots \phi_p$ of ARMA .
$\theta_1 \ldots \theta_q$

<u>Recursive Method :</u>

use OLS . get residuals . fit ARMA . estimate ϕ's θ's .

→ use GLS with $\hat{\phi}$, $\hat{\theta}$. get residuals . fit ARMA .

→ repeat until convergence .