Bayesian Time Series modelling and Analysis

6.12.2024

$\{y_t\}$ $\{y_t, t = 1, 2, \dots\}$ $y_{1:T}$

strong stationarity

$\{y_t\}$ any $n > 0$, any sequence $t_1, \dots, t_n$
and any $h > 0$

$(y_{t_1}, \dots, y_{t_n})'$ $(y_{t+h}, \dots, y_{t_n+h})$

weak stationarity (second order stationarity)

$\{y_t\}$ first and second moments of the two sequences exist and are identical.

$E[y_t] = \mu$, $var(y_t) = V$

$cov(y_t, y_s)$ depends on $|t - s|$

Strong stationarity implies weak stationarity

autocorrelation function (ACF)

$\{Y_t\}$

autocovariance

$$\gamma(t,s) = cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)]$$

$\mu_t = E[Y_t]$, $\mu_s = E[Y_s]$

stationarity $\Rightarrow$ $E[Y_t] = \mu$ for all t

and $\gamma(t,s) = \gamma(|t-s|)$

if $h > 0$ $\quad \gamma(h) = cov(Y_t, Y_{t-h})$

autocorrelation

$$\rho(t,s) = \frac{\gamma(t,s)}{\sqrt{\gamma(t,t)}\sqrt{\gamma(s,s)}}$$

stationarity $\Rightarrow$ $\rho(h) = \gamma(h)/\gamma(0)$

$\gamma(0) = var(Y_t)$

How to get estimates of γ and ρ ?

$y_{1:T}$

sample ACF: assuming stationarity

$$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=1}^{T-h} (y_{t+h} - \bar{y})(y_t - \bar{y})$$

with $\bar{y} = \frac{1}{T} \sum_{t=1}^{T} y_t$,

$$\hat{\rho}(h) = \hat{\gamma}(h) / \hat{\gamma}(0)$$

The AR(1)

$$y_t = \phi y_{t-1} + \varepsilon_t \quad , \quad \varepsilon_t \overset{iid}{\sim} N(0, v)$$

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

$$= \phi^2 y_{t-2} + \phi \varepsilon_{t-1} + \varepsilon_t$$

$$= \dots$$

$$= \phi^k y_{t-k} + \sum_{j=0}^{k-1} \phi^j \varepsilon_{t-j}$$

if $\phi \in [-1, 1] \Rightarrow$

$$y_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} \qquad (a)$$

$$E[Y_t] = 0 \quad , \quad var(Y_t) = var(a)$$
$$= \sum_{j=0}^{\infty} \phi^{2j} v$$
$$= \frac{v}{1-\phi^2}$$
does not dependont (stationary)

$$\gamma(h) = E[Y_t Y_{t-h}]$$
$$= E\left[\left(\sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}\right)\left(\sum_{k=0}^{\infty} \phi^k \varepsilon_{t-h-k}\right)\right]$$
$$= E\left[(\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \dots + \phi^h \varepsilon_{t-h})\right.$$
$$\left.(\varepsilon_{t-h} + \phi\varepsilon_{t-h-1} + \phi^2\varepsilon_{t-h-2} + \dots +)\right]$$

if time is not the same in the cross-producds the expectation is 0 because cov() and are iid.

$$= E[\phi^h \varepsilon_{t-h}^2 + \phi^{h+1}\phi\varepsilon_{t-h-1}^2 + \phi^{h+2}\phi^2\varepsilon_{t-h-2}$$
$$+ \dots]$$
$$= v \sum_{j=0}^{\infty} \phi^{h+j}\phi^j$$
$$= v\phi^h \sum_{j=0}^{\infty} \phi^{2j}$$
$$= \frac{v\phi^h}{1-\phi^2}$$

$$\rho(h) = \gamma(h) / \gamma(0) = \phi^h$$

In general, for any integer h, we have

$$\rho(h) = \phi^{|h|}$$

$$\gamma(h) = \frac{v\,\phi^{|h|}}{1-\phi^2}$$

The larger the value of ϕ, the decay will be slower.

Maximum likelihood Estimation for AR(1)

$$Y_t = \phi Y_{t-1} + \varepsilon_t \,, \quad \varepsilon_t \overset{iid}{\sim} N(0, v)$$

(we work will the conditional likelihood)

$$\phi \in [-1, 1] \Rightarrow \{Y_t\} \text{ is stationary}$$

$$Y_1 \sim N(0, v/1-\phi^2)$$

$$Y_t | Y_{t-1} \sim N(\phi Y_{t-1}, v)$$

$$Y_{1:T}$$

$$p(y_{1:T} \mid \phi, v) = p(y_1 \mid \phi, v) \prod_{t=2}^{T} p(y_t \mid y_{t-1}, \phi, v)$$

$$= \frac{1 \cdot (1-\phi^2)^{1/2}}{(2\pi v)^{1/2}} \exp\left\{ - \frac{y_1^2 (1-\phi^2)}{2v} \right\} \cdot$$

$$\underbrace{\frac{1}{(2\pi v)^{(T-1)1/2}} \exp\left\{ - \frac{\sum_{t=2}^{T} (y_t - \phi y_{t-1})^2}{2v} \right\}}_{Q^*(\phi)}$$

$$Q^*(\phi) = y_1^2 (1-\phi^2) + \underbrace{\sum_{t=2}^{T} (y_t - \phi y_{t-1})^2}_{Q(\phi)}$$

$$= \frac{(1-\phi)^{1/2}}{(2\pi v)^{T/2}} \exp\left\{ - \frac{Q^*(\phi)}{2v} \right\}$$

This is the full likelihood.

<u>Conditional likelihood</u> (conditioned on the first obs)

$$p(y_{2:T} \mid y_1, \phi, v) = \prod_{t=2}^{T} \frac{1}{(2\pi v)^{1/2}} \exp\left\{ - \frac{(y_t - \phi y_{t-1})^2}{2v} \right\}$$

$$= \frac{1}{(2\pi v)^{(T-1)/2}} \exp\left\{ - \frac{Q(\phi)}{2v} \right\}$$

$$\begin{pmatrix} y_2 \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_{T-1} \end{pmatrix} \phi + \begin{pmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_t \end{pmatrix}$$

$\Rightarrow$

$$\underline{\tilde{y}} = \underline{X}\,\phi + \underline{\varepsilon}, \quad \underline{\varepsilon} \sim N(0, v\underline{I})$$

if $\underline{X}$ is full rank, then

$$\hat{\phi} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{\tilde{y}}$$

$$\hat{v} = s^2 = \frac{(\underline{\tilde{y}} - \underline{X}\hat{\phi})^T (\underline{\tilde{y}} - \underline{X}\hat{\phi})}{|\dim(\underline{\tilde{y}}) - \dim(\hat{\phi})|}$$

$$\boxed{\begin{cases} \hat{\phi}_{mle} = \left(\sum_{t=2}^{T} y_t y_{t-1}\right) \Big/ \left(\sum_{t=2}^{T} y_{t-1}^2\right) \\ s^2 = \sum_{t=2}^{T} (y_t - \hat{\phi}_{mle} y_{t-1})^2 / (T-2) \quad \text{unbiased} \end{cases}}$$

We have to resort to numerical maximization for the AR coefficient if we work with the full likelihood.

$$Y_t = \phi Y_{t-1} + \varepsilon_t \quad , \quad \varepsilon_t \overset{iid}{\sim} N(0,1)$$

$$Y_1 \mid \phi \sim N\left(0, \frac{1}{(1-\phi^2)}\right)$$

$$Y_t \mid Y_{t-1}, \phi \sim N(\phi Y_{t-1}, 1)$$

<u>full likelihood</u>

$$p(Y_{1:T} \mid \phi) = p(Y_1 \mid \phi) \cdot \prod_{t=2}^{T} p(Y_t \mid Y_{t-1}\, \phi) =$$

$$= \frac{(1-\phi^2)^{1/2}}{(2\pi)^{1/2}} \exp\left\{-\frac{Y_t (1-\phi^2)}{2}\right\} \cdot$$

$$\frac{1}{(2\pi)^{(T-1)/2}} \exp\left\{-\sum_{t=2}^{T} \frac{(Y_t - \phi Y_{t-1})^2}{2}\right\}$$

$$= \frac{(1-\phi^2)^{1/2}}{(2\pi)^{T/2}} \exp\left\{-\frac{1}{2} Q^*(\phi)\right\}$$

$$\Rightarrow \log p(Y_{1:T} \mid \phi) = \frac{1}{2} \log(1-\phi^2) - \frac{1}{2} Q^*(\phi) + K$$

$\Rightarrow$ numerical optimization

e.g. Newton - Raphson.

Bayesian Inference in the AR(1)

model based on the conditional likelihood and a reference prior.

$$y_t = \phi y_{t-1} + \varepsilon_t \, , \quad \varepsilon_t \overset{iid}{\sim} N(0, v)$$

we can write as,

$$\underset{\sim}{\tilde{y}} = \underset{\sim}{X} \phi + \underset{\sim}{\varepsilon}$$

ϕ is a scalar or β

$\underset{\sim}{\varepsilon} \sim N(0, v\underline{I})$

simplify posterior

$$\begin{pmatrix} y_2 \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_{T-1} \end{pmatrix} \underset{(\beta)}{\phi} \begin{pmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix}$$

conditional likelihood:

$$p(y_{2:T} \mid \phi, v, y_1) = \frac{1}{(2\pi v)^{(T-1)/2}} \exp\left\{ - \frac{(\underline{\tilde{y}} - \underline{X}\phi)^T (\underline{\tilde{y}} - \underline{X}\phi)}{2v} \right\}$$

reference prior:

$$p(\phi, v) \propto \frac{1}{v}$$

posterior

$$p(\phi, v \mid y_{1:T}) \propto p(\phi, v) \, p(y_{2:T} \mid y_1, \phi, v)$$

$$p(\phi, v) \propto \frac{1}{v} \cdot (\underline{\beta} \mid v, \underline{X}, \underline{\tilde{y}}) \sim N(\underline{\hat{\beta}}_{mle}, v(\underline{X}^T X)^{-1})$$

$$\cdot (v \mid \underline{X}, \underline{\tilde{y}}) \sim IG\left(\frac{(T-2)}{2}, \frac{Q(\underline{\hat{\beta}}_{mle})}{2}\right)$$

$$\underline{\hat{\beta}}_{mle} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{\tilde{y}}$$

$$\underline{\hat{\beta}}_{mle} = \underline{\hat{\phi}}_{mle} = \left(\sum_{t=2}^{T} y_t y_{t-1}\right) \Big/ \sum_{t=2}^{T} y_{t-1}^2$$

$$Q(\underline{\hat{\phi}}_{mle}) = \sum_{t=2}^{T} (y_t - \underline{\hat{\phi}}_{mle} y_{t-1})^2$$

We will obtain samples for $\underline{\hat{\beta}}_{mle}$ and $\underline{\hat{\phi}}_{mle}$.

Definition of the state-space representation

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t ,$$

$$\varepsilon_t \overset{iid}{\sim} N(0, v)$$

general AR(p) setting

The AR characteristic polynomial is

$$\Phi(u) = 1 - \phi_1 u - \ldots - \phi_p u^p$$

u is any complex value.

<u>stable</u> : if $\Phi(u) = 0$ only if $|u| > 1$

(outside of the unit circle)

stable $\Rightarrow$ stationarity.

Then it has a moving average representation.

$$Y_t = \Psi(B)\,\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

with $\psi = 1$

and

$$\sum_{j=0}^{\infty} |\psi_j| < \infty$$

$B^j Y_t = Y_{t-j}$: B : Backshift operator

$$\Psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$$

we can also write :

$$\tilde{\Phi}(u) = \prod_{j=1}^{p} (1 - \alpha_j u) \quad (\alpha_1, \alpha_2, \dots, \alpha_p) \text{ are characteristic roots}$$

State - space (or dynamic linear model) representation:

$$Y_t = F^T \underline{X}_t$$

$$\underline{X}_t = G \underline{X}_{t-1} + \underline{w}_t$$

For AR(p), we have

$$\underline{X}_t = (Y_t, Y_{t-1}, \dots, Y_{t-p+1})^T$$

$$F = (1, 0, \dots, 0)^T$$

$$\underline{w}_t = (\varepsilon_t, 0, 0, \dots, 0)$$

$$G = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & & & 0 & 0 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & & \\ 0 & \cdots & & 1 & 0 \end{pmatrix}$$

if we do the matrix computation, we get back the general formula for AR(p).

$$f_t(h) = E[Y_{t+h} \mid Y_{1:t}]$$

for forecasting h steps ahead

$$= F^T E[\underline{X}_{t+h} \mid Y_{1:t}]$$

$$= F^T G \, E[\underline{X}_{t+h-1} \mid Y_{1:t}]$$

$$= F^T G^h E[\underline{X}_t \mid Y_{1:t}]$$

$$= F^T G^h \underline{X}_t$$

Eigenstructure of G? The eigenvalues corresponds to the characteristic roots of the process.

Assume we have p distinct reciprocal roots

$$G = E \Lambda E^{-1}$$

E: matrix of eigenvectors

$$\Lambda = \text{diagonal}(\alpha_1, \dots, \alpha_p)$$

$$G^h = E \Lambda^h E^{-1}$$

the power h (number of steps ahead) are the α at the power of h.

$$f_t(h) = \sum_{j=1}^{p} (c+j) \alpha_j^h$$

If the process is stable, the reciprocal roots are gonna decay exponentially. It provides a way to interpret the AR process.