MLE for Gaussian ARMA(p, q)

$$X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p}$$
$$+ \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$\epsilon_t \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$$

$\underset{\sim}{\eta} = (c, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)'$ - model parameters

σ^2 - noise parameter

Conditional MLE formulation

Consider the likelihood, conditioned on p initial values of x and q initial values of ϵ

$$\underset{\sim}{x_0} = (x_0, x_{-1}, x_{-2}, \ldots, x_{-(p-1)})'$$

and $$\underset{\sim}{\epsilon_0} = (\epsilon_0, \epsilon_{-1}, \ldots, \epsilon_{-(q-1)})'$$

Given $\underset{\sim}{x_0}$ and $\underset{\sim}{\epsilon_0}$ and using $\epsilon_t = x_t - c - \sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$, we can write $\epsilon_1, \ldots, \epsilon_n$ in terms of x_ts, $c, \phi_1, \ldots \phi_p, \theta_1, \ldots, \theta_q$

Note that

$$X_1 \mid \underset{\sim}{x_0}, \underset{\sim}{\epsilon_0} \sim N\left(c + \sum_{i=1}^{p} \phi_i x_{1-i} + \sum_{i=1}^{q} \theta_i \epsilon_{1-i}, \sigma^2\right)$$

In general,

$\forall t \geq 2$

$$X_t \mid X_{t-1}, \ldots, X_1, \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0} \equiv X_t \mid X_{t-1}, \ldots, X_{t-p}, \epsilon_{t-1}, \ldots, \epsilon_{t-q}$$

$$\sim N\left(c + \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}, \sigma^2\right)$$

Conditional likelihood f^n is given by

$$L_c(\underset{\sim}{\eta}) = f_{X_n, \ldots, X_1 \mid \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}}(x_n, \ldots, x_1; \underset{\sim}{\eta} \mid \underset{\sim}{x_0}, \underset{\sim}{\epsilon_0})$$

$$= f_{X_n \mid X_{n-1}, \ldots X_1, \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}} \; f_{X_{n-1}, \ldots X_1 \mid \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}}$$

$$\vdots \qquad\qquad \vdots$$

$$L_c(\underset{\sim}{\eta}) = f_{X_1 \mid \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}} \prod_{t=2}^{n} f_{X_t \mid X_{t-1}, \ldots X_1, \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}}$$

Conditional log likelihood f^n is given by

$$l_c(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(x_t - c - \sum_{i=1}^{p}\phi_i x_{t-i} - \sum_{i=1}^{q}\theta_i \epsilon_{t-i}\right)^2$$

i.e. $$l_c(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\epsilon_t^2$$

$$\hat{\underset{\sim}{\eta}}_{CMLE} = \arg\max_{\underset{\sim}{\eta}} l_c(\underset{\sim}{\eta}).$$

Note: Choice of initial values in CMLE formulation

Option I : Set initial xs and ϵs equal to their expected values

i.e. $x_s = \frac{c}{1-\phi_1 - \cdots - \phi_p}$; $s = 0, -1, \ldots, -(p-1)$

& $\epsilon_s = 0$; $s = 0, -1, \ldots, -(q-1)$

and use $\epsilon_t = x_t - c - \sum \phi_i x_{t-i} - \sum \theta_i \epsilon_{t-i}$

for $t = 1, \ldots, n$ to write $\epsilon_1, \ldots, \epsilon_n$

in terms of xs and $c, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$

Option II : Set ϵs at their expected values & set xs at their actual values

i.e. start with $(x_1, \ldots, x_p)$ as initial set of xs

and set $\epsilon_p = \epsilon_{p-1} = \cdots = \epsilon_{p-(q-1)} = 0$

& $\epsilon_t = x_t - c - \sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$

$t = p+1, p+2, \ldots, n$

Note that under this option the conditional log likelihood changes to

$$l_c(\underset{\sim}{\eta}) = -\frac{n-p}{2} \log 2\pi - \frac{n-p}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=p+1}^{n} \epsilon_t^2$$

Option iii : Set ϵs at the expected values

$$\epsilon_0 = \epsilon_{-1} = \cdots = \epsilon_{-(q-1)} = 0$$

and set xs at their "backforecasted" values

Back forcasting is a techniques for forcasting in backward direction.

Large Sample asymptotic distⁿ of MLE

Let $\{X_t\}$ be a Causal and invertible ARMA(p, q)

$$\phi(B) X_t = \theta(B) \epsilon_t$$

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$$

$$\underset{\sim}{\beta} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)'$$

Asymptotic distⁿ result:

$$\sqrt{n}\left(\hat{\underset{\sim}{\beta}}_{MLE} - \underset{\sim}{\beta}\right) \xrightarrow{L} N_{p+q}\left(\underset{\sim}{0}, V(\underset{\sim}{\beta})\right)$$

$$V(\underset{\sim}{\beta}) = \sigma^2 \begin{bmatrix} E\, \underset{\sim}{U}_t \underset{\sim}{U}_t' & E\, \underset{\sim}{U}_t \underset{\sim}{V}_t' \\ E\, \underset{\sim}{V}_t \underset{\sim}{U}_t' & E\, \underset{\sim}{V}_t \underset{\sim}{V}_t' \end{bmatrix}^{-1}$$

where, $\underset{\sim}{U}_t = (U_t, U_{t-1}, \ldots, U_{t-(p-1)})'$

$\underset{\sim}{V}_t = (V_t, V_{t-1}, \ldots, V_{t-(q-1)})'$

$\{U_t\}$ & $\{V_t\}$ are AR processes (stationary) given by

$$\phi(B) U_t = \epsilon_t$$

$$\theta(B) V_t = \epsilon_t$$

Note: If $\underbrace{p = 0}_{MA}$, then $V(\underset{\sim}{\beta}) = \sigma^2 \left(E\, \underset{\sim}{V}_t \underset{\sim}{V}_t'\right)^{-1}$.

If $\underbrace{q = 0}_{AR}$, then $V(\underset{\sim}{\beta}) = \sigma^2 \left(E\, \underset{\sim}{U}_t \underset{\sim}{U}_t'\right)^{-1}$

Example : AR(p)

$$\phi(B) X_t = \epsilon_t$$

$$\underset{\sim}{\phi} = (\phi_1, \ldots, \phi_p)'$$

$$V(\underset{\sim}{\phi}) = \sigma^2 \left(E\, \underline{U}_t \underline{U}_t' \right)^{-1}$$

U_t is $\ni$ $\phi(B) U_t = \epsilon_t$

$$E\, \underline{U}_t \underline{U}_t' = \Gamma_p = Cov(\underset{\sim}{U}_t) \; ; \; \underline{U}_t = \begin{pmatrix} U_t \\ U_{t-1} \\ \vdots \\ U_{t-(p-1)} \end{pmatrix}.$$

$$= \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ & \ddots & & \vdots \\ & & \ddots & \gamma_1 \\ & & & \gamma_0 \end{pmatrix}$$

$$V(\underset{\sim}{\phi}) = \sigma^2 \Gamma_p^{-1}$$

$$\sqrt{n}\left(\hat{\underset{\sim}{\phi}}_{MLE} - \underset{\sim}{\phi} \right) \xrightarrow{\mathcal{L}} N_p\left(\underset{\sim}{0}, \sigma^2 \Gamma_p^{-1}\right)$$

i.e. $\hat{\underset{\sim}{\phi}}_{MLE} \sim N_p\left(\underset{\sim}{\phi}, \frac{1}{n}\sigma^2 \Gamma_p^{-1}\right)$ for large n

Use Galbraith's formula to write elements of Γ_p^{-1}.

AR(1): $\sqrt{n}\left(\hat{\phi}_{MLE} - \phi \right) \xrightarrow{\mathcal{L}} N\left(0, \sigma^2 \left(\frac{\sigma^2}{1-\phi^2}\right)^{-1}\right)$

i.e. $N(0, 1-\phi^2)$

Random sampling from stationary time series

Let $X_1, \ldots, X_n$ be a sample of size n from a stationary time series with

(i) $E X_t = \mu \quad \forall t$

(ii) $\gamma_h = \mathrm{Cov}(X_t, X_{t+h}) = E(X_t - \mu)(X_{t+h} - \mu) \quad \forall t$

& (iii) $\sum_h |\gamma_h| < \infty$

Estimation of μ

$\bar{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t$ is an unbiased estimator for μ

$$E \bar{X}_n = \mu$$

$$V \bar{X}_n = \frac{1}{n} \sum_{|h| \leq n} \left(1 - \frac{|h|}{n}\right) \gamma_h$$

Some important asymptotic results!

Result 1: $E(\bar{X}_n - \mu)^2 \longrightarrow 0$ as $n \to \infty$

i.e. $\bar{X}_n \xrightarrow{m.s.} \mu$