Time Series Analysis - Estimation of ARMA models

ECON 722

Elena Pesavento — Emory University

Fall 2024

MLE

For iid data with marginal pdf $f\left(y_t;\boldsymbol{\theta}\right)$, the joint density function for a sample $\mathbf{y}=\left(y_1,\ldots,y_T\right)$ is simply the product of the marginal densities for each observation

$$f(\mathbf{y};\boldsymbol{\theta})=f\left(y_1,\ldots,y_T;\boldsymbol{\theta}\right)=\prod_{t=1}^{T} f\left(y_t;\boldsymbol{\theta}\right)$$

The likelihood function is this joint density treated as a function of the parameters $\boldsymbol{\theta}$ given the data $\mathbf{y}$:

$$L(\boldsymbol{\theta}\mid\mathbf{y})=L\left(\boldsymbol{\theta}\mid y_1,\ldots,y_T\right)=\prod_{t=1}^{T} f\left(y_t;\boldsymbol{\theta}\right)$$

The log-likelihood then as the simple form

$$\ln L(\boldsymbol{\theta}\mid\mathbf{y})=\sum_{t=1}^{T}\ln f\left(y_t;\boldsymbol{\theta}\right)$$

MLE for Dependent Data

- For a sample from a covariance stationary time series $\{y_t\}$, the construction of the log-likelihood give above doesn't work because the random variables in the sample $(y_1, \ldots, y_T)$ are not iid.
- One solution is to try to determine the joint density $f(y_1, \ldots, y_T; \boldsymbol{\theta})$ directly, which requires, among other things, the $T \times T$ variancecovariance matrix $\operatorname{var}(\mathbf{y})$. Hamilton describes this approach in detail for Gaussian ARMA processes.
- An alternative approach relies on factorization of the joint density into a series of conditional densities and the density of a set of initial values.

Conditional MLE

Consider the joint density of two adjacent observations $f(y_2, y_1; \boldsymbol{\theta})$ from a covariance stationary time series. The joint density can always be factored as the product of the conditional density of y_2 given y_1 and the marginal density of y_1 :

$$f(y_2, y_1; \boldsymbol{\theta}) = f(y_2 \mid y_1; \boldsymbol{\theta}) \, f(y_1; \boldsymbol{\theta})$$

For three observations, the factorization becomes

$$f(y_3, y_2, y_1; \boldsymbol{\theta}) = f(y_3 \mid y_2, y_1; \boldsymbol{\theta}) \, f(y_2 \mid y_1; \boldsymbol{\theta}) \, f(y_1; \boldsymbol{\theta})$$

In general, the conditional marginal factorization has the form

$$f(y_T, \ldots, y_1; \boldsymbol{\theta}) = \left(\prod_{t=p+1}^{T} f(y_t \mid I_{t-1}, \boldsymbol{\theta}) \right) \cdot f(y_p, \ldots, y_1; \boldsymbol{\theta})$$

where $I_t = \{y_t, \ldots, y_1\}$ denotes the information available at time t, and $y_p, \ldots, y_1$ denotes the initial values.

Conditional MLE

The log-likelihood function may then be expressed as

$$\ln L(\boldsymbol{\theta} \mid \mathbf{y}) = \sum_{t=p+1}^{T} \ln f\left(y_t \mid I_{t-1}, \boldsymbol{\theta}\right) + \ln f\left(y_p, \ldots, y_1; \boldsymbol{\theta}\right)$$

The full log-likelihood function is called the exact log-likelihood. The first term is called the conditional log-likelihood, and the second term is called the marginal loglikelihood for the initial values.

Conditional MLE

$$\hat{\boldsymbol{\theta}}_{\text{CMLE}} = \arg\max_{\theta} \sum_{t=p+1}^{T} \ln f\left(y_t \mid I_{t-1}, \boldsymbol{\theta}\right)$$

Exact MLE

$$\hat{\boldsymbol{\theta}}_{MLE} = \arg\max_{\theta} \sum_{t=p+1}^{T} \ln f\left(y_t \mid I_{t-1}, \boldsymbol{\theta}\right) + \ln f\left(y_p, \ldots, y_1; \boldsymbol{\theta}\right)$$

Conditional MLE

- For stationary models, $\hat{\boldsymbol{\theta}}_{\text{CMLE}}$ and $\hat{\boldsymbol{\theta}}_{\text{MLE}}$ are consistent and have the same limiting normal distribution. Intuition?
- For non stationary data the initial observation matters (initial conditions).
- In finite samples, however, $\hat{\boldsymbol{\theta}}_{\text{cmle}}$ and $\hat{\boldsymbol{\theta}}_{\text{mle}}$ are generally not equal and my differ by a substantial amount if the data are close to being non-stationary or non-invertible.

Example: MLE for AR(1)

Consider the stationary AR(1) model

$$y_t = c + \phi y_{t-1} + \varepsilon_t, \varepsilon_t \sim iidN\left(0, \sigma^2\right), t = 1, \ldots, T$$
$$\boldsymbol{\theta} = \left(c, \phi, \sigma^2\right)', |\phi| < 1$$

Conditional on I_{t-1}

$$y_t \mid I_{t-1} \sim N\left(c + \phi y_{t-1}, \sigma^2\right), t = 2, \ldots, T$$

which only depends on y_{t-1}. The conditional density $f\left(y_t \mid I_{t-1}, \theta\right)$ is then

$$f\left(y_t \mid y_{t-1}, \boldsymbol{\theta}\right) = \left(2\pi\sigma^2\right)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\left(y_t - c - \phi y_{t-1}\right)^2\right), t = 2, \ldots, T$$

Example: MLE for AR(1)

To determine the marginal density for the initial value y_1, recall that for a stationary $\mathrm{AR}(1)$ process

$$E\left[y_1\right] = \mu = \frac{c}{1-\phi}$$
$$\operatorname{var}\left(y_1\right) = \frac{\sigma^2}{1-\phi^2}$$

It follows that

$$y_1 \sim N\left(\frac{c}{1-\phi}, \frac{\sigma^2}{1-\phi^2}\right)$$
$$f\left(y_1;\boldsymbol{\theta}\right) = \left(2\pi\frac{\sigma^2}{1-\phi^2}\right)^{-1/2} \exp\left(-\frac{1-\phi^2}{2\sigma^2}\left(y_1 - \frac{c}{1-\phi}\right)^2\right)$$

Example: MLE for AR(1)

The conditional log-likelihood function is

$$\sum_{t=2}^{T} \ln f\left(y_t \mid y_{t-1}, \boldsymbol{\theta}\right) = \frac{-(T-1)}{2} \ln(2\pi) - \frac{(T-1)}{2} \ln\left(\sigma^2\right) - \frac{1}{2\sigma^2} \sum_{t=2}^{T} \left(y_t - c - \phi y_{t-1}\right)^2$$

Notice that the conditional log-likelihood function has the form of the log-likelihood function for a linear regression model with normal errors. It follows that the conditional mles for c and ϕ are identical to the least squares estimates from the regression

$$y_t = c + \phi y_{t-1} + \varepsilon_t, t = 2, \ldots, T$$

and the conditional MLE for σ^2 is

$$\hat{\sigma}^2_{\text{CMLE}} = (T-1)^{-1} \sum_{t=2}^{T} \left(y_t - \hat{c}_{\text{CMLE}} - \hat{\phi}_{\text{CMLE}}\, y_{t-1}\right)^2$$

Example: MLE for AR(1)

The marginal log-likelihood for the initial value y_1 is

$$\ln f\left(y_1; \boldsymbol{\theta}\right) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\frac{\sigma^2}{1-\phi^2}\right) - \frac{1-\phi^2}{2\sigma^2}\left(y_1 - \frac{c}{1-\phi}\right)^2$$

The exact log-likelihood function is then

$$\begin{aligned}\ln L(\boldsymbol{\theta} \mid \mathbf{y}) = & -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\frac{\sigma^2}{1-\phi^2}\right) - \frac{1-\phi^2}{2\sigma^2}\left(y_1 - \frac{c}{1-\phi}\right)^2 \\ & - \frac{(T-1)}{2}\ln\left(\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{t=2}^{T}\left(y_t - c - \phi y_{t-1}\right)^2\end{aligned}$$

- The exact log-likelihood function is a non-linear function of the parameters θ, and so there is no closed form solution for the exact mles.
- The exact mles must be determined by numerically maximizing the exact log-likelihood function.

CMLE for MA and ARMA models

- For MA and ARMA models, even the conditional Likelihood is a non linear functions of the parameters and has to be numerically maximized.
- See Hamilton Chapter 5, for derivations. Useful review.
- Most time sereis packages will do this for you.