Unit 3

Method of Maximum Likelihood Estimation

Maximum Likelihood Estimation of the Parameters of ARMA Models

(p, q)

p + q + 1

Let $\boldsymbol{\theta} = (\phi_1, ..., \phi_p, \theta_1, ..., \theta_q, \sigma^2)'$ denote the vector of population parameters.

Suppose we have observed a sample of size T n

$$\mathbf{x} = (x_1, ..., x_T)$$

Maximum Likelihood Estimation of the Parameters of ARMA Models

Let the joint probability density function (p.d.f.) of these data be denoted

$$f(x_T, x_{T-1}, \ldots, x_1; \boldsymbol{\theta}) \longrightarrow (*)$$

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

$$L(\boldsymbol{\theta}|\mathbf{x}) = f(x_T, x_{T-1}, \ldots, x_1; \boldsymbol{\theta})$$

$$\frac{df(x)}{dx} = 0$$

$$\frac{d^2 f(x)}{dx^2} < 0$$

Maximum Likelihood Estimation of the Parameters of ARMA Models

The maximum likelihood estimator (MLE) is

$$\hat{\boldsymbol{\theta}}_{MLE} = \arg\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L(\boldsymbol{\theta}|\mathbf{x})$$

where $\boldsymbol{\Theta}$ is the parameter space.

Maximum Likelihood Estimation of the Parameters of ARMA Models

For simplifying calculations, it is customary to work with the natural logarithm of L, given by

Log

$$\log L(\boldsymbol{\theta}|\mathbf{x}) = l(\boldsymbol{\theta}|\mathbf{x}).$$

This function is commonly referred to as the log-likelihood.

Maximum Likelihood Estimation of the Parameters of ARMA Models

Since the logarithm is a monotone transformation the values that maximize $L(\boldsymbol{\theta}|\mathbf{x})$ are the same as those that maximize $l(\boldsymbol{\theta}|\mathbf{x})$, that is

$$\hat{\boldsymbol{\theta}}_{MLE} = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} L(\boldsymbol{\theta}|\mathbf{x}) = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} l(\boldsymbol{\theta}|\mathbf{x})$$

but the the log-likelihood is computationally more convenient.

Maximum Likelihood Estimation of the Parameters of ARMA Models

Now, we assume that the derivative of $l(\boldsymbol{\theta}|\mathbf{x})$ (w.r. $\boldsymbol{\theta}$) exists and is continuous for all $\boldsymbol{\theta}$.

The necessary condition for maximizing $l(\boldsymbol{\theta}|\mathbf{x})$ is

$$\frac{\delta l(\boldsymbol{\theta}|\mathbf{x})}{\delta \boldsymbol{\theta}} = \mathbf{0}$$

which is called likelihood equation.

Maximum Likelihood Estimation of the Parameters of ARMA Models

The maximum likelihood estimate, $\hat{\boldsymbol{\theta}}_{MLE}$, will be the solution of

$$\frac{\delta l(\boldsymbol{\theta}|\mathbf{x})}{\delta \boldsymbol{\theta}} = \mathbf{0}$$

Properties of Maximum Likelihood Estimators

Maximum Likelihood Estimators are most attractive because of their asymptotic properties.
Under regularity conditions, the Maximum Likelihood Estimator, $\hat{\boldsymbol{\theta}}_{MLE}$, will have the following asymptotic properties:

1. It is consistent
2. It is asymptotically normally distributed
3. It is asymptotically efficient

These three properties explain the prevalence of the maximum likelihood technique in time series analysis

① $E(\hat{\theta}) = \theta$

② $\lim_{n\to\infty} P\left[\,|\theta - \hat{\theta}| \geq \varepsilon\,\right] \to 0$

③ $Var[\hat{\theta}_1] < Var[\hat{\theta}_2]$

$\downarrow$

efficient

The exact Gaussian likelihood of an ARMA process

(Normal)

To write down the likelihood function for an ARMA process, one must assume a particular distribution for the white noise process u_t. Here, we assume that u_t is a Gaussian white noise:

or ε_t

ε_t

$$u_t \sim i.i.d.N(\underbrace{0}_{\text{mean}}, \underbrace{\sigma^2}_{\text{var}})$$

The exact Gaussian likelihood of an ARMA process

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}$$

This implies that the exact Gaussian likelihood of $\mathbf{x}=(x_1, x_2, ..., x_T)'$ is given by

$$L(\boldsymbol{\theta}|\mathbf{x}) = (2\pi)^{-T/2}\,|\Gamma(\boldsymbol{\theta})|^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{x}'\Gamma(\boldsymbol{\theta})^{-1}\mathbf{x}\right\}$$

where $\Gamma(\boldsymbol{\theta}) = E(\mathbf{x}\mathbf{x}')$ is the $T \times T$ covariance matrix of $\mathbf{x}$ depending on $\boldsymbol{\theta}$.

The exact Gaussian likelihood of an ARMA process

The exact Gaussian log-likelihood is then given by

$$l(\boldsymbol{\theta}|\mathbf{x}) = -\frac{1}{2}\left[T\log(2\pi) + \log|\Gamma(\boldsymbol{\theta})| + \mathbf{x}'\Gamma(\boldsymbol{\theta})^{-1}\mathbf{x}\right]$$

$\log m^n = n \log m$

$\log mn = \log m + \log n$

The exact Gaussian likelihood of an AR(1) process

A Gaussian AR(1) process takes the form

$$x_t = \phi_1 x_{t-1} + u_t \quad \text{or } \varepsilon_t$$

with

$$u_t \sim i.i.d.N(0, \sigma^2)$$

For this case, the vector of popolation parameters to be estimated consists of $\boldsymbol{\theta} = (\phi_1, \sigma^2)'$.

The exact Gaussian likelihood of an AR(1) process

The exact Gaussian likelihood of $\mathbf{x}=(x_1, x_2, ..., x_T)'$ is given by

$$L(\boldsymbol{\theta}|\mathbf{x}) = (2\pi)^{-T/2} |\Gamma(\boldsymbol{\theta})|^{-1/2} \exp\left\{-\frac{1}{2}\mathbf{x}'\Gamma(\boldsymbol{\theta})^{-1}\mathbf{x}\right\}$$

where

$$\Gamma(\boldsymbol{\theta}) = \frac{\sigma^2}{1-\phi_1^2}\begin{bmatrix} 1 & \phi_1 & \phi_1^2 & \cdots & \phi_1^{T-1} \\ \phi_1 & 1 & \phi_1 & \cdots & \phi_1^{T-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \phi_1^{T-1} & \phi_1^{T-2} & \phi_1^{T-3} & \cdots & 1 \end{bmatrix}$$

In fact we recall that the j-th autovariance for an $AR(1)$ process is given by

$$E(x_t x_{t-j}) = \frac{\sigma^2 \phi_1^j}{1-\phi_1^2}$$

The exact Gaussian likelihood of an MA(1) process

The exact Gaussian likelihood of $\mathbf{x}=(x_1, x_2, ..., x_T)'$ is given by

$$L(\boldsymbol{\theta}|\mathbf{x}) = (2\pi)^{-T/2}\,|\Gamma(\boldsymbol{\theta})|^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{x}'\Gamma(\boldsymbol{\theta})^{-1}\mathbf{x}\right\}$$

where

$$\Gamma(\boldsymbol{\theta}) = \sigma^2\begin{bmatrix} (1+\theta_1) & \theta_1 & 0 & \cdots & 0 \\ \theta_1 & (1+\theta_1) & \theta_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & (1+\theta_1) \end{bmatrix}$$

Non-zero mean μ

Consider an ARMA process $\{x_t; t \in \mathbb{Z}\}$ with mean $\mu \neq 0$, defined by the equation

y_t or

$$\cancel{x_t} - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p} = c + \overset{\varepsilon_t}{\cancel{u_t}} + \theta u_{t-1} + \dots + \theta u_{t-p}$$

$$\varepsilon_t \text{ or } \cancel{u_t} \sim WN(0, \sigma^2)$$

where $\phi^{-1}(1)c = \mu$. The unknown parameters in this model are

- $\phi = (\phi_1, \dots, \phi_p)'$
- $\theta = (\theta_1, \dots, \theta_q)'$
- σ^2
- c

Non-zero mean μ

The equation

y_t

$$x_t - \phi_1 x_{t-1} - \ldots - \phi_p x_{t-p} = c + \underset{\varepsilon_t}{u_t} + \theta \underset{\varepsilon_{t-1}}{u_{t-1}} + \ldots + \theta u_{t-p}$$

can be rewritten as

$$(x_t - \mu) - \phi_1(x_{t-1} - \mu) - \ldots - \phi_p(x_{t-p} - \mu) = u_t + \theta u_{t-1} + \ldots + \theta u_{t-p}$$

Non-zero mean μ

We estimate μ by

$$\bar{x}_T = \sum_{t=1}^{T} x_t$$

and proceed to analyze the demeaned series

$$\{(x_t - \bar{x}_T); t = 1, ..., T\}$$