Time Series Analysis

ARMA Models

Nicoleta Serban, Ph.D.
Professor
Stewart School of Industrial and Systems Engineering

Parameter Estimation: Maximum Likelihood Estimation

Georgia Tech

About This Lesson

ARMA Model

$\{X_t, t \in \mathbb{Z}\}$ an ARM$A(p, q)$ process,

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

where $Z_t \sim$ WN$(0, \sigma^2)$.

- Model Parameters: ϕ_1, .., ϕ_p (AR coefficients), $\theta_1, \ldots, \theta_q$ (MA coefficients), μ (mean) and σ^2 (variance) unknown
- Commonly, de-mean the process X_t (hence μ =0)
- Orders: p, q set fixed but they need to be selected

Gaussian Time Series

$\{X_t\}$ a *Gaussian time series* if all of its joint distributions are multivariate normal, i.e. if for any collection of integers $i_1, \ldots, i_n$, the vector $\left(X_{i_1}, \ldots, X_{i_n}\right)^T$ has a multivariate normal distribution.

Property: If $\{Z_t\} \sim \mathrm{N}(0, \sigma^2)$ iid, then any causal invertible ARMA process

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

is a Gaussian time series.

Maximum Likelihood Estimation

Let $\bar{X}_n = (X_1, \dots, X_n)^T$ and $\hat{\bar{X}}_n = \left(\hat{X}_1, \dots, \hat{X}_n\right)^T$, where $\hat{X}_n = P_{n-1}X_n$.

- Since $\bar{X}_n$ is MVN, its likelihood is

$$L(\Gamma_n) = \frac{1}{(2\pi)^{n/2}} (\det \Gamma_n)^{-\frac{1}{2}} e^{-\frac{1}{2}\bar{X}_n^T \Gamma_n^{-1} \bar{X}_n}$$

where $\Gamma_n = [\gamma_X(i-j)]_{i,j=1,\dots,n}$.

- The innovations algorithm allows the likelihood to be computed without directly computing $\det \Gamma_n$ or Γ_n^{-1}.

Maximum Likelihood Estimation (cont'd)

The likelihood of a stationary zero-mean Gaussian time series is given by

$$L(\Gamma_n) = \frac{1}{(2\pi)^{n/2}} (v_0 v_1 \ldots v_{n-1})^{-1/2} e^{-\frac{1}{2}\sum_{j=1}^{n}\left(X_j - \hat{X}_j\right)^2 / v_{j-1}},$$

where $v_0, \ldots, v_{n-1}$ *are the mean squared errors*

$$v_i = E\left[\left(X_{i+1} - \hat{X}_{i+1}\right)^2\right]$$

obtained using the innovations algorithm.

➡ The terms $v_1, \ldots, v_{n-1}$ and $\{(X_j - \hat{X}_j), j = 1, \ldots, n\}$ can be obtained recursively using the innovations algorithm.

Example: AR Model via Linear Regression

Let $\{X_t\}$ be an AR(p) process satisfying

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t, \quad Z_t \sim \text{IID}(0, \sigma^2)$$

- Predicting variables: X_{t-1},.., X_{t-p}
- Regression coefficients: ϕ_1,.., ϕ_p
- Model selection: AR order
- Assumptions: $Z_t \sim \text{N}(0, \sigma^2)$ iid

MLE versus Other Approaches

Advantages of MLE:

- Efficient (low variance estimator).
- Often the Gaussian assumption is reasonable.
- Even if $\{X_t\}$ is not Gaussian, the asymptotic distribution of the estimators is the same as the Gaussian case.

Disadvantages of MLE:

- Difficult optimization problem – minimization is done numerically.
- Need to choose a good starting point (often use other estimators for this).

MLE: Preliminary Parameter Estimates

AR(p): Yule-Walker

MA(q): Innovations algorithm f

ARMA(p,q): Hannan-Rissanen algorithm

1. Estimate high-order AR.
2. Use to 'estimate' (unobserved) noise Z_t: $\hat{Z}_t$
3. Regress X_t onto $X_{t-1}, \ldots, X_{t-p}, \hat{Z}_{t-1}, \ldots, \hat{Z}_{t-q}$
4. Regress again with improved estimates of Z_t

Georgia Tech

MLE: Statistical Property

$\{X_t\}$ be an ARMA(p, q): the MLE satisfies

$$\begin{pmatrix} \hat{\phi} \\ \hat{\theta} \end{pmatrix} - \begin{pmatrix} \phi \\ \theta \end{pmatrix} \sim \text{AN}\left(0, \frac{\sigma^2}{n} \begin{pmatrix} \Gamma_{\phi\phi} & \Gamma_{\phi\theta} \\ \Gamma_{\theta\phi} & \Gamma_{\theta\theta} \end{pmatrix}^{-1}\right)$$

where $\begin{pmatrix} \Gamma_{\phi\phi} & \Gamma_{\phi\theta} \\ \Gamma_{\theta\phi} & \Gamma_{\theta\theta} \end{pmatrix} = \text{Cov}\big((X, Y), (X, Y)\big)$

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

$$\theta(B)Y_t = Z_t$$

Summary