Financial Time Series – Estimation of ARMA parameters

Andreas Petersson

TMS088/MSA410 – April 2020

Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, Sweden

Parameter estimation

- Goal: Assume we know p, q and estimate $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \sigma^2$ in

$$X_t - \sum_{j=1}^{p} \phi_j X_{t-j} = Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j}$$

  based on observations $(x_1, \ldots, x_n)$ of $(X_1, \ldots, X_n)$
- Main tool: Maximum likelihood
- Yule–Walker estimation: preliminary estimation of $AR(p)$
- Hannan–Rissanen estimation: preliminary estimation of $ARMA(p, q)$

Yule–Walker motivation: Calculating the ACVF

Method

We have for *causal* $\mathrm{ARMA}(p, q)$:

$$\gamma(k) - \sum_{j=1}^{p} \phi_j \gamma(k-j) = \sigma^2 \sum_{j=0}^{q-k} \theta_{k+j} \psi_j \tag{1}$$

for $0 \leq k \leq q$ and

$$\gamma(k) - \sum_{j=1}^{p} \phi_j \gamma(k-j) = 0 \tag{2}$$

for $k > q$, where $\psi_j := 0$ for $j < 0$, $\theta_0 := 1$, and $\theta_j := 0$ for $j \notin \{0, \ldots, q\}$.

Yule–Walker estimation

Rewritten as

$$\sum_{j=1}^{p} \phi_j \gamma(k-j) = \begin{cases} \gamma(k) & k \in \{1, \ldots, p\}, \\ \gamma(0) - \sigma^2 & k = 0. \end{cases}$$

Yule–Walker estimation

Rewritten as

$$\sum_{j=1}^{p} \phi_j \gamma(k-j) = \begin{cases} \gamma(k) & k \in \{1, \ldots, p\}, \\ \gamma(0) - \sigma^2 & k = 0. \end{cases}$$

Leads to the linear system

$$(\gamma(i-j))_{i,j=1}^{p} (\phi_1, \ldots, \phi_p)' = (\gamma(1), \ldots, \gamma(p))'$$

and to

$$(\phi_1, \ldots, \phi_p) \cdot (\gamma(1), \ldots, \gamma(p))' = \gamma(0) - \sigma^2.$$

Yule–Walker estimation

Method (Yule–Walker estimation)

Compute estimators $\hat{\sigma}^2$ and $(\hat{\phi}_j, j = 1, \ldots, p)$ from the equations

$$(\hat{\phi}_1, \ldots, \hat{\phi}_p)' = \hat{R}_p^{-1}(\hat{\rho}(1), \ldots, \hat{\rho}(p))',$$

and

$$\hat{\sigma}^2 = \hat{\gamma}(0)\left(1 - (\hat{\rho}(1), \ldots, \hat{\rho}(p))\hat{R}_p^{-1}(\hat{\rho}(1), \ldots, \hat{\rho}(p))'\right),$$

where $\hat{R}_p$ denotes the sample autocorrelation matrix.

For large n, $(\hat{\phi}_1, \ldots, \hat{\phi}_p) \sim \mathcal{N}((\phi_1, \ldots, \phi_p)', n^{-1}\sigma^2\Gamma_p^{-1})$ approximately, where $\Gamma_p := (\gamma(i-j))_{i,j=1}^p$

Hannan–Rissanen estimation

1. Fit a high-order $\mathrm{AR}(m)$ model (with $m > \max\{p, q\}$) to the data using the Yule–Walker estimates from Method 5. For estimated coefficients $(\hat{\phi}_{m1}, \ldots, \hat{\phi}_{mm})$, compute the estimated residuals $\hat{Z}_t$ from the equations

$$\hat{Z}_t = X_t - \sum_{j=1}^{m} \hat{\phi}_{mj} X_{t-j}$$

for $t = m + 1, \ldots, n$.

Hannan–Rissanen estimation

2. Estimate the vector of parameters $\beta := (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$ by a least squares linear regression of X_t onto $(X_{t-1}, \ldots, X_{t-p}, \hat{Z}_{t-1}, \ldots, \hat{Z}_{t-q})$, $t = m+1+q, \ldots, n$, i.e., minimize the sum of squares

$$S(\beta) = \sum_{t=m+1+q}^{n} \left(X_t - \sum_{j=1}^{p} \phi_j X_{t-j} - \sum_{i=1}^{q} \theta_i \hat{Z}_{t-i} \right)^2$$

with respect to β. This gives the *Hannan–Rissanen estimator*

$$\hat{\beta} = (Z'Z)^{-1} Z'(X_{m+1+q}, \ldots, X_n)',$$

where

$$Z = \begin{pmatrix} X_{m+q} & X_{m+q-1} & \cdots & X_{m+q+1-p} & \hat{Z}_{m+q} & \hat{Z}_{m+q-1} & \cdots & \hat{Z}_{m+1} \\ X_{m+q+1} & X_{m+q} & \cdots & X_{m+q+2-p} & \hat{Z}_{m+q+1} & \hat{Z}_{m+q} & \cdots & \hat{Z}_{m+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{n-1} & X_{n-2} & \cdots & X_{n-p} & \hat{Z}_{n-1} & \hat{Z}_{n-2} & \cdots & \hat{Z}_{n-q}. \end{pmatrix}$$

3. Set $\hat{\sigma}^{(\mathrm{HR})2} := \frac{S(\hat{\beta})}{n-m-q}$.

Maximum likelihood estimation

- Assume that $X \sim \mathrm{ARMA}(p, q)$ is Gaussian
- $\mathbf{X}_n := (X_1, \ldots, X_n)'$
- $\phi := (\phi_1, \ldots, \phi_p)$ and $\theta := (\theta_1, \ldots, \theta_q)$
- $\Gamma_n = \Gamma_n(\phi, \theta, \sigma^2)$
- $L(\Gamma_n) = f(\mathbf{X}_n | \Gamma_n) = (2\pi)^{-n/2} (\det \Gamma_n)^{-1/2} \exp\left(-2^{-1} \mathbf{X}_n' \Gamma_n^{-1} \mathbf{X}_n\right)$
- Idea: replace $X_1, \ldots, X_n$ with $X_1 - \hat{X}_1, \ldots, X_n - \hat{X}_n$

Maximum likelihood estimation

- Assume that $X \sim \mathrm{ARMA}(p,q)$ is Gaussian
- $\mathbf{X}_n := (X_1, \ldots, X_n)'$
- $\phi := (\phi_1, \ldots, \phi_p)$ and $\theta := (\theta_1, \ldots, \theta_q)$
- $\Gamma_n = \Gamma_n(\phi, \theta, \sigma^2)$
- $L(\Gamma_n) = f(\mathbf{X}_n | \Gamma_n) = (2\pi)^{-n/2} (\det \Gamma_n)^{-1/2} \exp\left(-2^{-1} \mathbf{X}_n' \Gamma_n^{-1} \mathbf{X}_n\right)$
- Idea: replace $X_1, \ldots, X_n$ with $X_1 - \hat{X}_1, \ldots, X_n - \hat{X}_n$

Lemma

There exist unique coefficients $(\theta_{ij},\ 1 \le j \le i \le n)$ *such that the best linear predictors satisfy*

$$\hat{X}_{n+1} = \begin{cases} 0 & \textit{for } n = 0, \\ \sum_{j=1}^{n} \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}) & \textit{for } n \ge 1. \end{cases}$$

Maximum likelihood estimation

$$\mathbf{X}_n = C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)$$

Maximum likelihood estimation

$$\mathbf{X}_n = C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)$$

$$\Gamma_n = \mathbb{E}(\mathbf{X}_n \mathbf{X}_n') = \mathbb{E}(C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)(\mathbf{X}_n - \hat{\mathbf{X}}_n)' C_n') = C_n D_n C_n'.$$

Maximum likelihood estimation

$$\mathbf{X}_n = C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)$$

$$\Gamma_n = \mathbb{E}(\mathbf{X}_n \mathbf{X}_n') = \mathbb{E}(C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)(\mathbf{X}_n - \hat{\mathbf{X}}_n)' C_n') = C_n D_n C_n'.$$

$$\mathbf{X}_n' \Gamma_n^{-1} \mathbf{X}_n = \sum_{j=1}^{n} (X_j - \hat{X}_j)^2 / v_{j-1}$$

Maximum likelihood estimation

$$\mathbf{X}_n = C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)$$

$$\Gamma_n = \mathbb{E}(\mathbf{X}_n \mathbf{X}_n') = \mathbb{E}(C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)(\mathbf{X}_n - \hat{\mathbf{X}}_n)' C_n') = C_n D_n C_n'.$$

$$\mathbf{X}_n' \Gamma_n^{-1} \mathbf{X}_n = \sum_{j=1}^{n} (X_j - \hat{X}_j)^2 / v_{j-1}$$

$$\det(\Gamma_n) = v_0 v_1 \cdots v_{n-1}.$$

Maximum likelihood estimation

$$\mathbf{X}_n = C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)$$

$$\Gamma_n = \mathbb{E}(\mathbf{X}_n \mathbf{X}_n') = \mathbb{E}(C_n(\mathbf{X}_n - \hat{\mathbf{X}}_n)(\mathbf{X}_n - \hat{\mathbf{X}}_n)' C_n') = C_n D_n C_n'.$$

$$\mathbf{X}_n' \Gamma_n^{-1} \mathbf{X}_n = \sum_{j=1}^{n} (X_j - \hat{X}_j)^2 / v_{j-1}$$

$$\det(\Gamma_n) = v_0 v_1 \cdots v_{n-1}.$$

$$L(\phi, \theta, \sigma^2) = (2\pi\sigma^2)^{-n/2} \Big(\prod_{j=1}^{n} r_{j-1} \Big)^{-1/2} \exp\Big(-(2\sigma^2)^{-1} \sum_{j=1}^{n} r_{j-1}^{-1} (X_j - \hat{X}_j)^2 \Big)$$

Maximum likelihood estimation

Method (Maximum likelihood estimators)

The MLE of σ^2, ϕ, and θ are determined from the expressions

$$\hat{\sigma}^2 = n^{-1} S(\hat{\phi}, \hat{\theta}),$$

and

$$(\hat{\phi}, \hat{\theta}) = \arg\min_{(\phi,\theta)} \ell(\phi, \theta).$$

Here

$$S(\phi, \theta) = \sum_{j=1}^{n} r_{j-1}^{-1} (X_j - \hat{X}_j)^2,$$

where $\hat{X}_j$ and r_{j-1}^{-1} are computed using the parameters ϕ and θ and

$$\ell(\phi, \theta) = \ln\left(n^{-1} S(\phi, \theta)\right) + n^{-1} \sum_{j=1}^{n} \ln r_{j-1}.$$

Maximum likelihood estimation

- For large n,

$$(\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\theta}_1, \ldots, \hat{\theta}_q)' \sim \mathcal{N}((\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)', 2H^{-1}(\beta)/n)$$

approximately. Here $H(\beta) = (\delta^2 \ell(\beta)/\delta\beta_i \delta\beta_j)_{i,j=1}^{p+q}$.

Maximum likelihood estimation

- For large n,

$$(\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\theta}_1, \ldots, \hat{\theta}_q)' \sim \mathcal{N}((\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)', 2H^{-1}(\beta)/n)$$

  approximately. Here $H(\beta) = (\delta^2 \ell(\beta)/\delta\beta_i \delta\beta_j)_{i,j=1}^{p+q}$.
- Holds also for $Z \sim IID(0, \sigma^2)$ without Gaussian assumption

Maximum likelihood estimation

- For large n,

$$(\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\theta}_1, \ldots, \hat{\theta}_q)' \sim \mathcal{N}((\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)', 2H^{-1}(\beta)/n)$$

  approximately. Here $H(\beta) = (\delta^2 \ell(\beta)/\delta\beta_i \delta\beta_j)_{i,j=1}^{p+q}$.
- Holds also for $Z \sim IID(0, \sigma^2)$ without Gaussian assumption
- Common alternative: $Z \sim IID(0, \sigma^2)$ and that Z_t follows a Student t-distribution for all times $t \implies$ *conditional MLE*

Maximum likelihood estimation

- For large n,

$$(\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\theta}_1, \ldots, \hat{\theta}_q)' \sim \mathcal{N}((\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)', 2H^{-1}(\beta)/n)$$

  approximately. Here $H(\beta) = (\delta^2 \ell(\beta)/\delta\beta_i \delta\beta_j)_{i,j=1}^{p+q}$.
- Holds also for $Z \sim IID(0, \sigma^2)$ without Gaussian assumption
- Common alternative: $Z \sim IID(0, \sigma^2)$ and that Z_t follows a Student t-distribution for all times $t \implies$ *conditional MLE*
- The *rescaled residuals*

$$\hat{R}_t := (X_t - \hat{X}_t)/v_{t-1}^{1/2} = (X_t - \hat{X}_t(\hat{\phi}, \hat{\theta}))/v_{t-1}^{1/2}(\hat{\phi}, \hat{\theta}),$$

  should behave like Z_t/σ
- $\mathbb{E}((\hat{R}_t - Z_t/\sigma)^2) \to 0$