Financial Time Series – Forecasting of ARMA processes

Andreas Petersson

TMS088/MSA410 – May 2020

UNIVERSITY OF GOTHENBURG

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

Recall: 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

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

Recall: Innovations algorithm

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

Method (Innovations algorithm)

Compute the coefficients $\theta_{n1}, \ldots, \theta_{nn}$ recursively from the equations

$$v_0 := \kappa(1,1)$$

and

$$\theta_{n(n-k)} := v_k^{-1} \left(\kappa(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k(k-j)} \theta_{n(n-j)} v_j \right)$$

for $0 \leq k < n$ and

$$v_n := \kappa(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n(n-j)}^2 v_j.$$

Innovations algorithm for the transformation W

Idea: Apply algorithm to $W = (W_t, t \in \mathbb{N})$, where, with $m = \max\{p, q\}$,

$$W_t := \begin{cases} \sigma^{-1} X_t & t = 1, \ldots, m, \\ \sigma^{-1} \phi(B) X_t & t > m. \end{cases}$$

Innovations algorithm for the transformation W

Idea: Apply algorithm to $W = (W_t, t \in \mathbb{N})$, where, with $m = \max\{p, q\}$,

$$W_t := \begin{cases} \sigma^{-1} X_t & t = 1, \dots, m, \\ \sigma^{-1} \phi(B) X_t & t > m. \end{cases}$$

The values $\kappa(i, j) = \mathbb{E}(W_i W_j)$ are given by

$$\kappa(i,j) = \begin{cases} \sigma^{-2} \gamma_X(i-j) & \max(i,j) \le m, \\ \sigma^{-2} \left(\gamma_X(i-j) - \sum_{r=1}^{p} \phi_r \gamma_X(r - |i-j|) \right) & \min\{i,j\} \le m < \max\{i,j\} \le 2m, \\ \sum_{r=0}^{q} \theta_r \theta_{r+|i-j|} & \min\{i,j\} > m, \\ 0 & \text{otherwise.} \end{cases}$$

Innovations algorithm for the transformation W

Idea: Apply algorithm to $W = (W_t, t \in \mathbb{N})$, where, with $m = \max\{p, q\}$,

$$W_t := \begin{cases} \sigma^{-1} X_t & t = 1, \ldots, m, \\ \sigma^{-1} \phi(B) X_t & t > m. \end{cases}$$

The values $\kappa(i, j) = \mathbb{E}(W_i W_j)$ are given by

$$\kappa(i,j) = \begin{cases} \sigma^{-2} \gamma_X(i-j) & \max(i,j) \le m, \\ \sigma^{-2} \left(\gamma_X(i-j) - \sum_{r=1}^{p} \phi_r \gamma_X(r - |i-j|) \right) & \min\{i,j\} \le m < \max\{i,j\} \le 2m, \\ \sum_{r=0}^{q} \theta_r \theta_{r+|i-j|} & \min\{i,j\} > m, \\ 0 & \text{otherwise.} \end{cases}$$

$$\hat{W}_{n+1} = \begin{cases} \sum_{j=1}^{n} \theta_{nj} (W_{n+1-j} - \hat{W}_{n+1-j}) & 1 \le n < m, \\ \sum_{j=1}^{q} \theta_{nj} (W_{n+1-j} - \hat{W}_{n+1-j}) & n \ge m, \end{cases}$$

Forecasting X using W

Compute forecasts for the (causal) $\mathrm{ARMA}(p, q)$ process

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

Forecasting X using W

Compute forecasts for the (causal) $\mathrm{ARMA}(p, q)$ process

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

Note that

$$\hat{W}_{n+1} = b^l_{n+1}(W^n)$$

with $W^n := (W_1, \ldots, W_n)$.

Forecasting X using W

Compute forecasts for the (causal) $\mathrm{ARMA}(p,q)$ process

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

Note that

$$\hat{W}_{n+1} = b^l_{n+1}(W^n)$$

with $W^n := (W_1, \ldots, W_n)$.

$$\hat{W}_t = \begin{cases} \sigma^{-1}\hat{X}_t & t = 1, \ldots, m, \\ \sigma^{-1}\left(\hat{X}_t - \sum_{j=1}^{p} \phi_j X_{t-j}\right) & t > m \end{cases}$$

Forecasting X using W

Compute forecasts for the (causal) $\mathrm{ARMA}(p, q)$ process

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

Note that

$$\hat{W}_{n+1} = b^l_{n+1}(W^n)$$

with $W^n := (W_1, \ldots, W_n)$.

$$\hat{W}_t = \begin{cases} \sigma^{-1}\hat{X}_t & t = 1, \ldots, m, \\ \sigma^{-1}\left(\hat{X}_t - \sum_{j=1}^p \phi_j X_{t-j}\right) & t > m \end{cases}$$

$$\hat{W}_t - W_t = \sigma^{-1}(\hat{X}_t - X_t) \text{ for } t \geq 1$$

Forecasting X using W

Compute forecasts for the (causal) $\mathrm{ARMA}(p,q)$ process

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

Note that

$$\hat{W}_{n+1} = b^l_{n+1}(W^n)$$

with $W^n := (W_1, \ldots, W_n)$.

$$\hat{W}_t = \begin{cases} \sigma^{-1}\hat{X}_t & t = 1, \ldots, m, \\ \sigma^{-1}\left(\hat{X}_t - \sum_{j=1}^p \phi_j X_{t-j}\right) & t > m \end{cases}$$

$$\hat{W}_t - W_t = \sigma^{-1}(\hat{X}_t - X_t) \text{ for } t \geq 1$$

$$\hat{X}_{n+1} = \begin{cases} \sum_{j=1}^n \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}) & 1 \leq n < m, \\ \sum_{j=1}^p \phi_j X_{n+1-j} + \sum_{j=1}^q \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}) & n \geq m, \end{cases}$$

Forecasting X using W

Compute forecasts for the (causal) $\mathrm{ARMA}(p,q)$ process

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

Note that

$$\hat{W}_{n+1} = b^l_{n+1}(W^n)$$

with $W^n := (W_1, \dots, W_n)$.

$$\hat{W}_t = \begin{cases} \sigma^{-1}\hat{X}_t & t = 1, \dots, m, \\ \sigma^{-1}\left(\hat{X}_t - \sum_{j=1}^p \phi_j X_{t-j}\right) & t > m \end{cases}$$

$$\hat{W}_t - W_t = \sigma^{-1}(\hat{X}_t - X_t) \text{ for } t \geq 1$$

$$\hat{X}_{n+1} = \begin{cases} \sum_{j=1}^n \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}) & 1 \leq n < m, \\ \sum_{j=1}^p \phi_j X_{n+1-j} + \sum_{j=1}^q \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}) & n \geq m, \end{cases}$$

$$\mathrm{MSE}(\hat{X}_{n+1}, X_{n+1}) = \sigma^2 \,\mathrm{MSE}(\hat{W}_{n+1}, W_{n+1}) = \sigma^2 v_n,$$

Forecasting for $h > 1$

$b^l_{n+h}(W^n)$ is obtained by dropping the first $h-1$ terms of

$$\hat{W}_{n+h} = b^l_{n+h}(W^{n+h-1}) = \sum_{j=1}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j}).$$

Why?

Forecasting for $h > 1$

$b^l_{n+h}(W^n)$ is obtained by dropping the first $h-1$ terms of

$$\hat{W}_{n+h} = b^l_{n+h}(W^{n+h-1}) = \sum_{j=1}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j}).$$

Why?

$$\mathbb{E}\left(\left(W_{n+h} - \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j})\right) W_k\right)$$

Forecasting for $h > 1$

$b^l_{n+h}(W^n)$ is obtained by dropping the first $h-1$ terms of

$$\hat{W}_{n+h} = b^l_{n+h}(W^{n+h-1}) = \sum_{j=1}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j}).$$

Why?

$$\mathbb{E}\left(\left(W_{n+h} - \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j})\right) W_k\right)$$

$$= \mathbb{E}\left(\left(W_{n+h} - \hat{W}_{n+h} + \sum_{j=1}^{h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j})\right) W_k\right)$$

Forecasting for $h > 1$

$b^l_{n+h}(W^n)$ is obtained by dropping the first $h-1$ terms of

$$\hat{W}_{n+h} = b^l_{n+h}(W^{n+h-1}) = \sum_{j=1}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j}).$$

Why?

$$\mathbb{E}\left(\left(W_{n+h} - \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j})\right) W_k\right)$$

$$= \mathbb{E}\left(\left(W_{n+h} - \hat{W}_{n+h} + \sum_{j=1}^{h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j})\right) W_k\right)$$

$$= \mathbb{E}\left((W_{n+h} - \hat{W}_{n+h})W_k\right) + \sum_{j=1}^{h-1} \theta_{(n+h-1)j}\, \mathbb{E}\left((W_{n+h-j} - \hat{W}_{n+h-j})W_k\right)$$

Forecasting for $h > 1$

$b^l_{n+h}(W^n)$ is obtained by dropping the first $h-1$ terms of

$$\hat{W}_{n+h} = b^l_{n+h}(W^{n+h-1}) = \sum_{j=1}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j}).$$

Why?

$$\mathbb{E}\left(\left(W_{n+h} - \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j})\right) W_k\right)$$

$$= \mathbb{E}\left(\left(W_{n+h} - \hat{W}_{n+h} + \sum_{j=1}^{h-1} \theta_{(n+h-1)j}(W_{n+h-j} - \hat{W}_{n+h-j})\right) W_k\right)$$

$$= \mathbb{E}\left((W_{n+h} - \hat{W}_{n+h})W_k\right) + \sum_{j=1}^{h-1} \theta_{(n+h-1)j}\, \mathbb{E}\left((W_{n+h-j} - \hat{W}_{n+h-j})W_k\right)$$

Therefore

$$b^l_{n+h}(W^n) = \sigma^{-2} \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j}(X_{n+h-j} - \hat{X}_{n+h-j}).$$

Forecasting for $h > 1$

From

$$\hat{W}_t = \begin{cases} \sigma^{-1} \hat{X}_t & t = 1, \ldots, m, \\ \sigma^{-1} \left(\hat{X}_t - \sum_{j=1}^{p} \phi_j X_{t-j} \right) & t > m \end{cases}$$

one obtains

$$b^l_{n+h}(X^n) = \begin{cases} \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j} (X_{n+h-j} - \hat{X}_{n+h-j}) \\ \sum_{i=1}^{p} \phi_i b^l_{n+h-i}(X^n) + \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j} (X_{n+h-j} - \hat{X}_{n+h-j}) \end{cases}$$

Forecasting for $h > 1$

From

$$\hat{W}_t = \begin{cases} \sigma^{-1} \hat{X}_t & t = 1, \ldots, m, \\ \sigma^{-1} \left(\hat{X}_t - \sum_{j=1}^{p} \phi_j X_{t-j} \right) & t > m \end{cases}$$

one obtains

$$b^l_{n+h}(X^n) = \begin{cases} \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j} (X_{n+h-j} - \hat{X}_{n+h-j}) \\ \sum_{i=1}^{p} \phi_i b^l_{n+h-i}(X^n) + \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j} (X_{n+h-j} - \hat{X}_{n+h-j}) \end{cases}$$

When $n > m := \max\{p, q\}$, then for all $h \geq 1$

$$b^l_{n+h}(X^n) = \sum_{i=1}^{p} \phi_i \, b^l_{n+h-i}(X^n) + \sum_{j=h}^{q-1} \theta_{(n+h-1)j} (X_{n+h-j} - \hat{X}_{n+h-j}).$$

Forecasting for $h > 1$

From

$$\hat{W}_t = \begin{cases} \sigma^{-1} \hat{X}_t & t = 1, \ldots, m, \\ \sigma^{-1} \left(\hat{X}_t - \sum_{j=1}^{p} \phi_j X_{t-j} \right) & t > m \end{cases}$$

one obtains

$$b^l_{n+h}(X^n) = \begin{cases} \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j} (X_{n+h-j} - \hat{X}_{n+h-j}) \\ \sum_{i=1}^{p} \phi_i b^l_{n+h-i}(X^n) + \sum_{j=h}^{n+h-1} \theta_{(n+h-1)j} (X_{n+h-j} - \hat{X}_{n+h-j}) \end{cases}$$

When $n > m := \max\{p, q\}$, then for all $h \geq 1$

$$b^l_{n+h}(X^n) = \sum_{i=1}^{p} \phi_i \, b^l_{n+h-i}(X^n) + \sum_{j=h}^{q-1} \theta_{(n+h-1)j} (X_{n+h-j} - \hat{X}_{n+h-j}).$$

$$\mathrm{MSE}(b^l_{n+h}(X^n), X_{n+h}) = \sum_{j=0}^{h-1} \left(\sum_{r=0}^{j} \chi_r \theta_{(n+h-r-1)(j-r)} \right)^2 v_{n+h-j-1},$$

From

$$\hat{W}_t = \begin{cases} \sigma^{-1}\hat{X}_t & t = 1, \dots, m, \\ \sigma^{-1}\left(\hat{X}_t - \sum_{j=1}^{p}\phi_j X_{t-j}\right) & t > m \end{cases}$$

one obtains

$$b^l_{n+h}(X^n) = \begin{cases} \sum_{j=h}^{n+h-1}\theta_{(n+h-1)j}(X_{n+h-j} - \hat{X}_{n+h-j}) \\ \sum_{i=1}^{p}\phi_i b^l_{n+h-i}(X^n) + \sum_{j=h}^{n+h-1}\theta_{(n+h-1)j}(X_{n+h-j} - \hat{X}_{n+h-j}) \end{cases}$$

When $n > m := \max\{p, q\}$, then for all $h \geq 1$

$$b^l_{n+h}(X^n) = \sum_{i=1}^{p}\phi_i\, b^l_{n+h-i}(X^n) + \sum_{j=h}^{q-1}\theta_{(n+h-1)j}(X_{n+h-j} - \hat{X}_{n+h-j}).$$

$$\mathrm{MSE}(b^l_{n+h}(X^n), X_{n+h}) = \sum_{j=0}^{h-1}\left(\sum_{r=0}^{j}\chi_r\theta_{(n+h-r-1)(j-r)}\right)^2 v_{n+h-j-1},$$

The coefficients $(\chi_j, j = 0, 1, \dots)$ are computed recursively by $\chi_0 := 1$ and $\chi_j = \sum_{k=1}^{\min\{p,j\}}\phi_k\chi_{j-k}$.

Prediction bounds

- If the $\mathrm{ARMA}(p, q)$ process is Gaussian then for each $h \geq 1$,

$$b^l_{n+h}(X^n) - X_{n+h} \sim \mathcal{N}\left(0, \mathrm{MSE}(b^l_{n+h}(X^n), X_{n+h})\right)$$

- Allows for the computation of confidence intervals $\leftarrow$ *prediction bounds* of $b^l_{n+h}(X^n)$
- If X is not necessarily Gaussian but invertible, then

$$\mathbb{E}((X_t - \hat{X}_t - Z_t)^2) \to 0$$

  as $t \to \infty$
- $\implies$ a confidence interval for Z_t is an approximate prediction bound for $\hat{X}_t$