

ARIMA

The ARIMA procedure computes the parameter estimates for a given seasonal or non-seasonal univariate ARIMA model. It also computes the fitted values, forecasting values, and other related variables for the model.

Notation

The following notation is used throughout this chapter unless otherwise stated:

$y_t (t=1, 2, \dots, N)$	Univariate time series under investigation
N	Total number of observations
$a_t (t = 1, 2, \dots, N)$	White noise series normally distributed with mean zero and variance σ_a^2
p	Order of the non-seasonal autoregressive part of the model
q	Order of the non-seasonal moving average part of the model
d	Order of the non-seasonal differencing
P	Order of the seasonal autoregressive part of the model
Q	Order of the seasonal moving-average part of the model
D	Order of the seasonal differencing
s	Seasonality or period of the model
$\phi_p(B)$	AR polynomial of B of order p , $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$
$\theta_q(B)$	MA polynomial of B of order q , $\theta_q(B) = 1 - \vartheta_1 B - \vartheta_2 B^2 - \dots - \vartheta_q B^q$
$\Phi_P(B)$	SAR polynomial of B of order P , $\Phi_P(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_P B^P$
$\Theta_Q(B)$	SMA polynomial of B of order Q , $\Theta_Q(B) = 1 - \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_Q B^Q$
∇	Non-seasonal differencing operator $\nabla = 1 - B$
∇_s	Seasonal differencing operator with seasonality s , $\nabla_s = 1 - B^s$
B	Backward shift operator with $By_t = y_{t-1}$ and $Ba_t = a_{t-1}$

Models

A seasonal univariate ARIMA(p,d,q)(P,D,Q)_s model is given by

$$\phi_p(B)\Phi_P(B)[\nabla^d \nabla_s^D y_t - \mu] = \theta_q(B)\Theta_Q(B)a_t \quad t = 1, \dots, N \quad (1)$$

where μ is an optional model constant. It is also called the stationary series mean, assuming that, after differencing, the series is stationary. When NOCONSTANT is specified, μ is assumed to be zero.

When $P = Q = D = 0$, the model is reduced to a (non-seasonal) ARIMA(p,d,q) model:

$$\phi_p(B)[\nabla^d y_t - \mu] = \theta_q(B)a_t \quad t = 1, \dots, N \quad (2)$$

An optional log scale transformation can be applied to y_t before the model is fitted. In this chapter, the same symbol, y_t , is used to denote the series either before or after log scale transformation.

Independent variables x_1, x_2, \dots, x_m can also be included in the model. The model with independent variables is given by

$$\phi_p(B)\Phi_P(B)[\nabla^d \nabla_s^D (y_t - \sum_{i=1}^m c_i x_{it}) - \mu] = \theta_q(B)\Theta_Q(B)a_t$$

$$\text{or } \Phi(B)[\nabla(B)(y_t - \sum_{i=1}^m c_i x_{it}) - \mu] = \Theta(B)a_t \quad (3)$$

where

$$\Phi(B) = \phi_p(B)\Phi_P(B)$$

$$\nabla(B) = \nabla^d \nabla_s^D$$

$$\Theta(B) = \theta_q(B)\Theta_Q(B)$$

and $c_i, i = 1, 2, \dots, m$, are the regression coefficients for the independent variables.

Estimation

Basically, two different estimation algorithms are used to compute maximum likelihood (ML) estimates for the parameters in an ARIMA model:

- **Melard's algorithm** is used for the estimation when there is no missing data in the time series. The algorithm computes the maximum likelihood estimates of the model parameters. The details of the algorithm are described in Melard (1984), Pearlman (1980), and Morf, Sidhu, and Kailath (1974).
- A **Kalman filtering algorithm** is used for the estimation when some observations in the time series are missing. The algorithm efficiently computes the marginal likelihood of an ARIMA model with missing observations. The details of the algorithm are described in the following literature: Kohn and Ansley (1986) and Kohn and Ansley (1985).

Diagnostic Statistics

The following definitions are used in the statistics below:

N_p	Number of parameters
	$N_p = \begin{cases} p + q + P + Q + m & \text{without model constant} \\ p + q + P + Q + m + 1 & \text{with model constant} \end{cases}$
SSQ	Residual sum of squares
	$SSQ = \mathbf{e}'\mathbf{e}$, where \mathbf{e} is the residual vector
$\hat{\sigma}_a^2$	Estimated residual variance
	$\hat{\sigma}_a^2 = \frac{SSQ}{df}$, where $df = N - N_p$
SSQ'	Adjusted residual sum of squares
	$SSQ' = (SSQ) \Omega ^{1/N}$, where Ω is the theoretical covariance matrix of the observation vector computed at MLE

4 ARIMA

Log-Likelihood

$$L = -N \ln(\hat{\sigma}_a) - \frac{SSQ'}{2\hat{\sigma}_a^2} - \frac{N \ln(2\pi)}{2}$$

Akaike Information Criterion (AIC)

$$AIC = -2L + 2N_p$$

Schwartz Bayesian Criterion (SBC)

$$SBC = -2L + \ln(N)N_p$$

Generated Variables

Predicted Values

Forecasting Method: Conditional Least Squares (CLS or AUTOINT)

In general, the model used for fitting and forecasting (after estimation, if involved) can be described by Equation (3), which can be written as

$$y_t - D(B)y_t = \Phi(B)\mu + \Theta(B)a_t + \sum_{i=1}^m c_i \Phi(B)\nabla(B)x_{it}$$

where

$$D(B) = \Phi(B)\nabla(B) - 1$$

$$\Phi(B)\mu = \Phi(1)\mu$$

Thus, the predicted values (FIT)_{*t*} are computed as follows:

$$(FIT)_t = \hat{y}_t = D(B)\hat{y}_t + \Phi(B)\mu + \Theta(B)\hat{a}_t + \sum_{i=1}^m c_i \Phi(B)\nabla(B)x_{it} \quad (4)$$

where $\hat{a}_t = y_t - \hat{y}_t \quad 1 \leq t \leq n$

Starting Values for Computing Fitted Series. To start the computation for fitted values using Equation (4), all unavailable beginning residuals are set to zero and unavailable beginning values of the fitted series are set according to the selected method:

- **CLS.** The computation starts at the $(d+sD)$ -th period. After a specified log scale transformation, if any, the original series is differenced and/or seasonally differenced according to the model specification. Fitted values for the differenced series are computed first. All unavailable beginning fitted values in the computation are replaced by the stationary series mean, which is equal to the model constant in the model specification. The fitted values are then aggregated to the original series and properly transformed back to the original scale. The first $d+sD$ fitted values are set to missing (SYSMIS).
- **AUTOINIT.** The computation starts at the $[d+p+s(D+P)]$ -th period. After any specified log scale transformation, the actual $d+p+s(D+P)$ beginning observations in the series are used as beginning fitted values in the computation. The first $d+p+s(D+P)$ fitted values are set to missing. The fitted values are then transformed back to the original scale, if a log transformation is specified.

Forecasting Method: Unconditional Least Squares (EXACT)

As with the CLS method, the computations start at the $(d+sD)$ -th period. First, the original series (or the log-transformed series if a transformation is specified) is differenced and/or seasonally differenced according to the model specification. Then the fitted values for the differenced series are computed. The fitted values are one-step-ahead, least-squares predictors calculated using the theoretical autocorrelation function of the stationary autoregressive moving average (ARMA) process corresponding to the differenced series. The autocorrelation function is computed by treating the estimated parameters as the true parameters. The fitted values are then aggregated to the original series and properly transformed back to the original scale. The first $d+sD$ fitted values are set to missing (SYSMIS). The details of the least-squares prediction algorithm for the ARMA models can be found in Brockwell and Davis (1991).

Residuals

Residual series are always computed in the transformed log scale, if a transformation is specified.

$$(ERR)_t = y_t - (FIT)_t \quad t = 1, 2, \dots, N$$

Standard Errors of the Predicted Values

Standard errors of the predicted values are first computed in the transformed log scale, if a transformation is specified.

Forecasting Method: Conditional Least Squares (CLS or AUTOINIT)

$$(SEP)_t = \hat{\sigma}_a \quad t = 1, 2, \dots, N$$

Forecasting Method: Unconditional Least Squares (EXACT)

In the EXACT method, unlike the CLS method, there is no simple expression for the standard errors of the predicted values. The standard errors of the predicted values will, however, be given by the least-squares prediction algorithm as a byproduct.

Standard errors of the predicted values are then transformed back to the original scale for each predicted value, if a transformation is specified.

Confidence Limits of the Predicted Values

Confidence limits of the predicted values are first computed in the transformed log scale, if a transformation is specified:

$$(LCL)_t = (FIT)_t - t_{1-\alpha/2, df} (SEP)_t \quad t = 1, 2, \dots, N$$

$$(UCL)_t = (FIT)_t + t_{1-\alpha/2, df} (SEP)_t \quad t = 1, 2, \dots, N$$

where $t_{1-\alpha/2, df}$ is the $(1 - \alpha / 2)$ -th percentile of a t distribution with df degrees of freedom and α is the specified confidence level (by default $\alpha = 0.05$).

Confidence limits of the predicted values are then transformed back to the original scale for each predicted value, if a transformation is specified.

Forecasting

Forecasting Values

Forecasting Method: Conditional Least Squares (CLS or AUTOINIT)

The following forecasting equation can be derived from Equation (3):

$$\hat{y}_t(l) = D(B)\hat{y}_{t+l} + \Phi(B)\mu + \Theta(B)\hat{a}_{t+l} + \sum_{i=1}^m c_i \Phi(B)\nabla(B)x_{i,t+l} \quad (5)$$

where

$$D(B) = \Phi(B)\nabla(B) - 1, \quad \Phi(B)\mu = \Phi(1)\mu$$

$\hat{y}_t(l)$ denotes the l -step-ahead forecast of y_{t+l} at the time t .

$$\hat{y}_{t+l-i} = \begin{cases} y_{t+l-i} & \text{if } l \leq i \\ \hat{y}_t(l-i) & \text{if } l > i \end{cases}$$

$$\hat{a}_{t+l-j} = \begin{cases} y_{t+l-i} - \hat{y}_{t+l-i-1}(1) & \text{if } l \leq i \\ 0 & \text{if } l > i \end{cases}$$

Note that $\hat{y}_{t'}(1)$ is the *one-step-ahead forecast* of $y_{t'+1}$ at time t' , which is exactly the predicted value $\hat{y}_{t'+1}$ as given in Equation (4).

Forecasting Method: Unconditional Least Squares (EXACT)

The forecasts with this option are finite memory, least-squares forecasts computed using the theoretical autocorrelation function of the series. The details of the least-squares forecasting algorithm for the ARIMA models can be found in Brockwell and Davis (1991).

Standard Errors of the Forecasting Values

Forecasting Method: Conditional Least Squares (CLS or AUTOINIT)

For the purpose of computing standard errors of the forecasting values, Equation (1) can be written in the format of ψ weights (ignoring the model constant):

$$y_t = \frac{\vartheta_q(B)\Theta_Q(B)}{\phi_p(B)\Phi_P(B)} a_t = \psi(B)a_t = \sum_{i=0}^{\infty} \psi_i B^i a_{t-i}, \quad \psi_0 = 1 \quad (6)$$

Let $\hat{y}_t(l)$ denote the l -step-ahead forecast of y_{t+l} at time t . Then

$$\text{se}[\hat{y}_t(l)] = \{1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{l-1}^2\}^{1/2} \hat{\sigma}_a$$

Note that, for the predicted value, $l = 1$. Hence, $(SEP)_t = \hat{\sigma}_a$ at any time t .

Computation of ψ Weights. ψ weights can be computed by expanding both sides of the following equation and solving the linear equation system established by equating the corresponding coefficients on both sides of the expansion:

$$\phi_p(B)\Phi_P(B)\nabla^d \nabla_s^D \psi(B) = \theta_q(B)\Theta_Q(B)$$

An explicit expression of ψ weights can be found in Box and Jenkins (1976).

Forecasting Method: Unconditional Least Squares (EXACT)

As with the standard errors of the predicted values, the standard errors of the forecasting values are a byproduct during the least-squares forecasting computation. The details can be found in Brockwell and Davis (1991).

References

- Box, G. E. P., and Jenkins, G. M. 1976. *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.
- Brockwell, P. J., and Davis, R. A. 1991. *Time series: Theory and methods*, 2nd ed. New York: Springer-Verlag.