Chapter 29
Time Series Analysis

1 Scope of the Chapter

This chapter provides procedures for the analysis of time series data.

2 Available Modules

Module 29.1: `nag_tsa_identify` — **Time series analysis**

This module contains procedures for calculating the sample autocorrelations of a univariate time series. Procedures are provided for

- calculating the autocorrelation function;
- calculating the partial autocorrelation function.

Module 29.2: `nag_tsa_kalman` — **Kalman filtering**

This module contains procedures for the Kalman filtering of time series data. The procedures are based on the square-root algorithm. Facilities are provided for

- an initial estimate for the state covariance matrix;
- a prediction step of the square-root covariance Kalman filter;
- a combined update-prediction step, using the time-varying square-root covariance Kalman filter;
- a combined update-prediction step, using the time-invariant square-root covariance Kalman filter.

3 Background

3.1 Introduction

Time series data, x_t $(t = 1, \ldots, n)$, generally consist of both *deterministic* and *stochastic* components. The deterministic component gives rise to trends, seasonal patterns and cycles, while the stochastic component causes statistical fluctuations which have a short term correlation structure.

A time series is *stationary* if the structure of the series depends only the relative position of the observations; that is, the joint distribution of x_t and x_{t+l} depends only on l not on t. If we define $E(x_t) = \mu_t$, $\text{var}(x_t) = E(x_t - \mu_t)^2$, and $\text{cov}(x_t, x_{t-l}) = \gamma(t, t-l) = E\left((x_t - \mu_t)(x_{t-l} - \mu_{t-l})\right)$, then a series is *second-order* stationary if $\mu_t = \mu$ for all t and $\gamma(t, t-l) = \gamma_l$ for all t and l.

If the variance of the observations in the series is not constant across the range of observations it may be useful to apply a variance stabilizing transformation to the series. A common situation is for the variance to increase with the magnitude of the observations, and in this case typical transformations used are the log or square root transformation.

3.2 ARMA Time Series Models

A stationary time series may often be modelled as an auto-regressive moving average (ARMA) process. For the univariate time series x_t $(t = 1, 2, \ldots, n)$, with mean μ, an ARMA(p, q) model is

$$w_t = \phi_1 w_{t-1} + \cdots + \phi_p w_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q},$$

where the ϕ are the p auto-regressive parameters, the θ are the q moving average parameters, $w_t = x_t - \mu$ and ϵ_t is an uncorrelated sequence of noise with zero mean and variance σ_ϵ^2.

If we define the back shift operator $B^j x_t = x_{t-j}$ then the above equation can be written more compactly as

$$\phi_p(B) w_t = \theta_q(B) \epsilon_t,$$

where the auto-regressive operator $\phi_p(B)$ is

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 \cdots - \phi_p B^p,$$

and the moving average operator $\theta_q(B)$ is

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 \cdots - \theta_q B^q.$$

For the process to be stationary the roots of the polynomial equation $\phi_p(z) = 0$ must lie outside the unit circle. A second important property is invertibility; this allows the process to be written as a pure autoregressive process and avoids model multiplicities. For the process to be invertible we require that the roots of the polynomial equation $\theta_q(z) = 0$ lie outside the unit circle.

The advantage of an ARMA model over a pure AR model is that a complex series can often be represented using an ARMA model with fewer parameters than would be needed by a suitable AR model.

A k-variate ARMA(p, q) can be represented using similar equations to those given above, except that w_t and ϵ_t are now vectors of length k and the ϕ and θ are now $k \times k$ matrices.

Seasonal, SARMA(p, q, P, Q, s), models also allow for correlation at lags which are multiples of seasonal period s. In this case the model can be written as

$$\Phi_P(B^s) \phi_p(B) w_t = \Theta_Q(B^s) \theta_q(B) \epsilon_t,$$

where the seasonal auto-regressive operator $\Phi_P(B)$ is

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} \cdots - \Phi_P B^{Ps},$$

and the seasonal moving average operator $\Theta_Q(B^s)$ is

$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} \cdots - \Theta_Q B^{Qs}.$$

For the SARMA process to be stationary the roots of the polynomial equation $\Phi_P(z)\phi_p(z) = 0$ must lie outside the unit circle, while for the process to be invertible the roots of the polynomial equation $\Theta_Q(z)\theta_q(B) = 0$ must lie outside the unit circle.

3.3 ARIMA Time Series Models

In many cases while a time series w_t may be non-stationary the differenced series $w_t - w_{t-1}$ may be stationary. Single differencing will remove linear trends, higher order and seasonal differencing may remove more complex trends and cycles.

If the time series w_t $(t = 1, 2, \ldots, n)$ is non-stationary and the differenced series follows an ARMA(p, q) then the series is termed an ARIMA(p, d, q) and can be written as

$$\phi_p(B)(1 - B)^d w_t = \theta_q(B) \epsilon_t,$$

where $(1 - B)^d$ is the dth-order differencing operator, i.e.,

$$(1 - B) w_t = w_t - w_{t-1}, \quad (1 - B)^2 w_t = w_t - 2w_{t-1} + w_{t-2}, \text{ etc.}$$

In the case of a seasonal series it can be reduced to a SARMA(p, q, P, Q, s) by using seasonal differencing, with the model being written as follows:

$$\Phi_P(B^s) \phi_p(B)(1 - B^s)^D (1 - B)^d w_t = \Theta_Q(B^s) \theta_q(B) \epsilon_t,$$

where $(1 - B^s)^D$ is the Dth-order seasonal differencing operator, i.e.,

$$(1 - B^s) w_t = w_t - w_{t-s}, \quad (1 - B^s)^2 w_t = w_t - 2w_{t-s} + w_{t-2s}, \text{ etc.}$$

3.4 Structural Models

Structural models represent a time series directly in terms of trend, seasonal and irregular components and thus offer an alternative to ARIMA models.

The basic structural model for the univariate time series y_t, $t = 1, \ldots, T$ has the form

$$y_t = \mu_t + \gamma_t + \epsilon_t$$

where μ_t, γ_t and ϵ_t are respectively the trend, seasonal and irregular components.
The equations for the process generating the trend are

$$\text{level:} \quad \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad t = 1, \ldots, T$$

and

$$\text{slope:} \quad \beta_t = \beta_{t-1} + \zeta_t, \quad t = 1, \ldots, T$$

where η_t and ζ_t are normally distributed white noise processes with zero means and variances σ_η^2 and σ_ζ^2, respectively. The effect of η_t is to allow the level of the trend to shift up and down, while ζ_t allows the slope to change.

If the seasonal components were deterministic then, at time t, they would have to satisfy

$$\sum_{j=0}^{s-1} \gamma_{t-j} = 0$$

where s is the number of 'seasons' in the 'year'. The introduction of a random component, $\omega_t = N(0, \sigma_\omega^2)$ on the right-hand side,

$$\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t,$$

allows the seasonal effect to change over time, but still ensures that the sum of any s consecutive seasonal components has an expected value of zero. A recursive formula for the seasonal components is

$$\gamma_t = \sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t.$$

The above is called the dummy variable form of seasonality.

An alternative way of modelling seasonality is by a set of trigonometric terms at the seasonal frequencies, $\lambda_j = 2\pi j/s$, $j = 1, \ldots, s/2$. The lowest frequency, $2\pi/s$, is known as the fundamental frequency while the remaining frequencies are harmonics. If, as above, a random component is added to allow the season effects to change over time but still have an expected value of zero, the following recursive formula for the seasonal effect at time t is obtained:

$$\gamma_t = \sum_{j=1}^{s-1} \gamma_{j,t}$$

where

$$\left.\begin{array}{rcl} \gamma_{j,t} & = & \gamma_{j,t-1}\cos\lambda_j + \gamma^*_{j,t-1}\sin\lambda_j + \omega_{j,t} \\ \gamma^*_{j,t} & = & -\gamma_{j,t-1}\sin\lambda_j + \gamma^*_{j,t-1}\cos\lambda_j + \omega^*_{j,t} \end{array}\right\} \quad j = 1, \ldots, s/2,$$

where $\omega_{j,t}$ and $\omega^*_{j,t}$ are zero mean white noise processes which are uncorrelated and have a common variance σ_j^2. Assigning different variances to each harmonic allows them to evolve at varying rates. However, it is usually desirable to let $\text{var}(\omega^*_{j,t}) = \text{var}(\omega_{j,t}) = \sigma_j^2 = \sigma_\omega^2, j = 1, \ldots, s/2$. When s is even and $j = s/2$, the sine term vanishes giving $\gamma_{j,t} = \gamma_{j,t-1}\cos\lambda_j + \omega_{j,t}$. For even s the number of parameters, γ_j and γ^*_j, is therefore always $s-1$, which is the same as the number of coefficients in the seasonal dummy form. Provided that the full set of trigonometric terms is included, the trigonometric and dummy variable approach should give identical results. However, the advantage of the trigonometric approach is that it allows unimportant frequency components to be dropped and thus enables more compact models to be constructed.

3.5 State Space Models

A general class of time series models is the state space model. This comprises two equations, the state equation that models the unobserved stochastic state vector, X_t,

$$X_{t+1} = A_t X_t + B_t W_t$$

and the measurement equation,

$$Y_t = C_t X_t + V_t$$

for the observed Y_t. W_t is the state noise, V_t is the measurement noise, A_t is the state transition matrix, B_t is the noise coefficient matrix, and C_t is the measurement coefficient matrix. The state noise and the measurement noise are assumed to be uncorrelated and have zero mean, and the covariance matrices are

$$\mathrm{cov}(W_t) = Q_t \text{ and } \mathrm{cov}(V_t) = R_t.$$

With state space models the primary objective of the analysis is to estimate the value of the state vector, X_t, given the observed $Y_1 \ldots Y_t$ and to forecast X_{t+1}. This is achieved as recursive operation using the *Kalman Filter*. The current estimate of X_t given observations $Y_1 \ldots Y_{t-1}$, denoted by $\hat{X}_{t|t-1}$, is updated by the Y_t observation to give the estimate $\hat{X}_{t|t}$. This operation can be seen as being within a Bayesian framework.

Both the ARMA and the structural models described above can be written as state space models. The matrices A_t and B_t are then functions of the parameters of the model.

3.6 Model Identification

Model identification is the stage in the modelling process in which the form of the model that is thought to be suitable is selected. Identification can either be conceptual, based on theoretical knowledge of the system, or empirical, making use of data observed on the system. Because of the generality of state space models any direct model identification would have to be, in the first instance, conceptual. For ARIMA models empirical identification is possible. The basic tools for ARIMA model identification are data plot, the *autocorrelation function (acf)* and the *partial autocorrelation function (pacf)*. A simple plot of the series will indicate if the data has to be differenced or transformed in order to produce a stationary series. Given a stationary series the autocorrelation function is the correlation of observations l apart for $l = 1, 2, \ldots$, that is an estimate of $\rho_l = \gamma_l/\gamma_0$. For pure MA processes of order q the autocorrelations should be zero for $l > q$. The partial autocorrelation is a measure of the correlation between observations l apart having removed the effect of the correlation due to the intervening observations. For a pure AR process of order p the partial autocorrelations should be zero for $l > p$. Thus examination of plots of the acf and pacf should indicate if a MA or AR model is suitable or if a more complex ARMA model is needed.

3.7 Model Fitting

Given a specified form of a model in terms of unknown parameters, θ_i, ϕ_i, etc., the model can be fitted by estimating the parameters using maximum likelihood. If the model is written as a state space model the likelihood can be computed by making use of a Kalman filter. The Kalman filter computes an estimate of the state vector X_t and its covariance matrix P_t given observations $Y_1 \ldots Y_t$ and values of the model parameters. As a by-product one-step-ahead residuals, r_t and their covariance matrix, H_t are also computed; from these the log-likelihood for the unknown parameters can be computed as

$$\kappa - \frac{1}{2}\sum_{i=1}^{t} \ln(\det(H_i)) - \frac{1}{2}\sum_{i=1}^{t} r_i^T H_i^{-1} r_i,$$

where κ is a constant. This likelihood can then be maximized using a non-linear optimizer.

3.8 Model Checking

In addition to testing the significance of the estimated parameters by using likelihood ratio tests or z-test, the fit of the model can be checked by examining the one-step-ahead residuals. These residuals should be independently and identically distributed Normal variates. The independence can be checked by computing the acf of the residuals, and the Normality can be checked using graphical methods.

3.9 Forecasting

Given a fitted state-space model, forecasts can be computed by using

$$\hat{X}_{t+i|t} = A_{t+i-1}\hat{X}_{t+i-1|i} \qquad i = 1, 2, \ldots,$$

where $\hat{X}_{t+i|t}$ is the estimate of the state vector X_{t+i} given observation up to time t. The covariance matrix, $P_{t+i|t}$, of the forecasts can be computed from

$$P_{t+i|t} = A_{t+i-1}P_{t+i-1|t}A^T_{t+i-1|t} + B_iQ_iB^T_i,$$

where $P_{t+1|t}$ is the state covariance matrix at $t+1$ for observations up to time t and is computed by the Kalman filter along with $\hat{X}_{t+1|t}$.