Applications

Kuttykrishnan A.P “Laplace autoregressive time series models”, Department of Statistics, University of Calicut, 2006
Chapter-VII
Applications

7.1. Introduction

Several models have been used for modeling time series observations. A mathematical model representing the set of observations consists of a certain mathematical form and a set of parameters. Hence fitting an appropriate model to an observed series involves two interrelated problems, namely determining the order and estimating the parameters of the model. A systematic approach to modeling time series observations involves following phases.

(i) Select the type of model among the various linear time series models.
(ii) Determine the order of the model.
(iii) Use proper estimation method to estimate the parameters of the model.
(iv) Testing the goodness of fit of the model.

The overall time series modeling is an iterative process with the feedback and interaction between the above-referred phases.

The selection of the type of the linear model generally depends on the statistical characteristics like autocorrelation and partial autocorrelation functions of the given time series. Let \( \{x_0, x_1, \ldots, x_n\} \) be the given time series values, then the sample autocorrelation function \( \hat{\rho}(h) \) of order \( h \) is given by
\[
\hat{\rho}(h) = \frac{\sum_{j=0}^{n-h} (x_{j+h} - \bar{x})(x_j - \bar{x})}{\sum_{j=0}^{n} (x_j - \bar{x})^2},
\]

where \( \bar{x} = \frac{1}{n+1} \sum_{j=0}^{n} x_j \) and \( h < n \).

The partial autocorrelation function at lag \( h \) denoted by \( \alpha(h) = \phi_{hh}, h \geq 1 \) of the stationary process \( \{X_n\} \) is uniquely determined by the equation

\[
\begin{bmatrix}
\rho(0) & \rho(1) & \ldots & \rho(h-1) \\
\rho(1) & \rho(0) & \ldots & \rho(h-2) \\
\vdots & \vdots & \ddots & \vdots \\
\rho(h-1) & \rho(h-2) & \ldots & \rho(0)
\end{bmatrix}
\begin{bmatrix}
\phi_{h1} \\
\phi_{h2} \\
\vdots \\
\phi_{hh}
\end{bmatrix}
= \begin{bmatrix}
\rho(1) \\
\rho(2) \\
\vdots \\
\rho(h)
\end{bmatrix}.
\]

Hence the sample partial autocorrelation \( \hat{\alpha}(h) \) at lag \( h \) of \( \{x_0, x_1, \ldots, x_n\} \), provided \( x_i \neq x_j \) for some \( i \) and \( j \), is given by \( \hat{\alpha}(h) = \hat{\rho}_{hh}, 1 \leq h < n \), where \( \hat{\phi}_{hh} \) is uniquely determined by (7.1.2) with each \( \rho(h) \) replaced by the corresponding sample autocorrelation \( \hat{\rho}(h) \) given by (7.1.1).

Box and Jenkins (1970) discussed the model identification procedure using autocorrelation and partial autocorrelation functions. If the data exhibits stationarity character and have rapidly decreasing autocorrelation function we choose an autoregressive process with order \( p \) to model the data. To determining the order of
the process we choose the partial autocorrelation function. It can be seen that for an AR (p) process the partial autocorrelations $\phi_{mm}$ are equal to zero when $m > p$. In particular the estimate of $p$ is obtained by finding the smallest value of $r$ such that the sample partial autocorrelation function $\hat{\phi}_{mm}$ satisfies $|\hat{\phi}_{mm}| < \frac{1.96}{\sqrt{n}}$ for $m > r$, where $n$ is the number of observations. If the sample autocorrelation and partial autocorrelation functions of observed series is in tune with the theoretical pattern of the known model, we use such model to model the given series of observations.

For an AR (1) process the first order autocorrelation function decays exponentially, while the partial autocorrelation function cuts off after the first lag. If the autocorrelation function and partial autocorrelation function of the sample are consistent with that of an AR (1) model we can identify the model as AR (1). To select the proper marginal distribution to fit the observed data, statistical tests like chi-square and Kolmogrov- Smirnov tests are used to test the goodness of fit. The same problem can also be addressed using the methods percentage-percentage (P-P) probability plots / quantile- quantile (Q-Q) probability plots (for more discussions see Sim (1994)).

When a satisfactory model is found, the next stage is extrapolating past behavior into the future and hence forecast future values of the observed series. Given the information set $\{x_0, x_1, \ldots, x_n\}$, the variable $X_{n+m}, m = 1, 2, \ldots$ can be predicted. The classical statistical theory tells us that the minimum mean square error point forecast of $X_{n+m}, m = 1, 2, \ldots$, given the information $\{x_0, x_1, \ldots, x_n\}$ is
E(X_{n+m} / x_0, x_1, ..., x_n). Hence for the Markov process the point forecasts are obtained by finding E(X_{n+m} / x_n) where m = 1, 2, ...

7.2. An Application

The class of asymmetric Laplace distributions is well suited for modeling phenomena where the variable of interest results from a large random number of independent observations, while the empirical distribution appears to be asymmetric, with steep peak and tails heavier than those allowed by normal distribution. Usually the empirical distributions of data sets related to currency exchange rate, interest rate, stock price changes, industrial production rate etc. often do not exhibit the character of normal law but with asymmetric, steep peak and heavy tailed character. Kozubowski and Podgórski (1999, 2000) used asymmetric Laplace distributions for modeling interest rate and currency exchange rate. Although the theory and applications of asymmetric Laplace distributions is well developed in recent years, applications in time series modeling is not much developed. Here we are discussing an application of asymmetric Laplace distribution in the field of time series modeling.

As an application, ALAR (1) model is fitted to the monthly industrial production index of the USA. The historical data consists of 1020 values of index of industrial production of the USA (data is taken from the web site www.economagic.com) from January 1921 to December 2005. Time series plot of
the data \( \{Y_n\} \) is given in Figure 7.2.1. From the figure we can notice an upward trend in the index of industrial production.

![Figure 7.2.1](image-url)

**Figure 7.2.1**

**Time series plot of index of industrial production**

The given time series \( \{Y_n\} \) is non-stationary. Now we generate a new series by the transformation \( \ln Y_n \). Time series plot of \( \{\ln Y_n\} \) is given in the Figure 7.2.2.

![Figure 7.2.2](image-url)

**Figure 7.2.2**

**Time series plot of logarithm of index of industrial production**

The non-stationary historical series \( \{\ln Y_n\} \) is made stationary by taking the first order difference of the historical data. Let us define

\[
X_i = \ln Y_i - \ln Y_{i-1}, \quad i = 1, 2, \ldots
\]

Time series plot of series \( \{X_n\} \) is given in
Figure 7.2.3. Now it is seen that the resulting time series \( \{X_n\} \) can be modeled using a stationary process. From the time series plot of the stationary series \( \{X_n\} \) we can notice large fluctuations at the initial stage and as time increases fluctuations become small. Also a general asymmetry and sharp peaks in the fluctuations is exhibited.

![Time series plot of first order difference of logarithm of index of industrial production](image)

Figure 7.2.3

Time series plot of first order difference of logarithm of index of industrial production

The sample autocorrelation function (ACF) and partial autocorrelation (PACF) values of the same corresponding to different lags of the stationary series are given in the Table 7.2.1.

<table>
<thead>
<tr>
<th>Lag</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>0.508</td>
<td>0.213</td>
<td>0.084</td>
<td>-0.003</td>
<td>-0.02</td>
<td>-0.03</td>
<td>0.08</td>
<td>0.115</td>
</tr>
<tr>
<td>PACF</td>
<td>0.508</td>
<td>-0.061</td>
<td>-0.001</td>
<td>-0.049</td>
<td>0.007</td>
<td>-0.028</td>
<td>0.164</td>
<td>0.110</td>
</tr>
</tbody>
</table>

Table 7.2.1

Sample autocorrelation and partial autocorrelation function
From the above table we can note that, sample autocorrelation values are exponentially decreasing and lag-2 sample partial autocorrelation function 
\[ \hat{\phi}_{22} = -0.061 \] falls within the limits \[ \pm \frac{1.96}{\sqrt{n}} \]. Hence the autocorrelation and partial autocorrelation of the sample data are consistent with that of a first order autoregressive process. So we prefer to model the data using a first order autoregressive model.

Next we shall examine whether the data \( \{x_n\} \) can be modeled using an asymmetric Laplace distribution. The minimum value of the data is -0.11 and maximum value is 0.153. A frequency table is prepared by taking equal class intervals starting from -0.11 and ending at 0.153 and the histogram of first order difference of the logarithm of index of industrial production is presented in Figure 7.2.4.

![Histogram of first order difference of logarithm of index of industrial production data](image)

**Figure 7.2.4**

Histogram of first order difference of logarithm of index of industrial production data

199
The histogram of the series demonstrates that the historical data corresponding to the first order difference of logarithm of index of industrial production has peaked with asymmetric character. Now we estimate the model and distribution parameters using the method described in Section 2.2.2 of Chapter-II.

The estimates values of $\rho$, $\mu$ and $\sigma$ are $\hat{\rho} = 0.509$, $\hat{\mu} = 0.003$ and $\hat{\sigma} = 0.013$ respectively.

Hence the estimate $\hat{\kappa}$ of $\kappa$ is given by $\hat{\kappa} = \frac{2}{\frac{\hat{\mu}}{\hat{\sigma}} + \sqrt{4 + \left(\frac{\hat{\mu}}{\hat{\sigma}}\right)^2}} = 0.853$.

Using these estimate values simulate a sequence $\{X_n\}$ of ALAR (1) process of the form (2.2.3).

![Figure 7.2.5](image)

**Figure 7.2.5**

*Time series plot of simulated ALAR (1) sequence*

In order to assess the adequacy of fit, we use the technique of graphical method. The histogram of the sequence $\{X_n\}$ with the best-fitted sequence of ALAR (1) process super imposed on it is given in the Figure 7.2.6a. The P-P plot of the same is given in the Figure 7.2.6 b. Using the historical data $\{X_n\}$ we draw the
cumulative frequency curve with the cumulative frequency curve of simulated ALAR (1) process embedded on it. The same is given in the Figure 7.2.6 c. The plots of data quantile versus asymmetric Laplace quantile corresponding to the estimated parameters are given the Figure 7.2.6 d.

![Figure 7.2.6a](cumulative-frequency-curve.png)  ![Figure 7.2.6b](Q-Q-plot.png)

**Figure 7.2.6a**  **Figure 7.2.6b**

Histogram and Q-Q plot of ALAR (1) sequence

![Figure 7.2.6c](cumulative-frequency-curve.png)  ![Figure 7.2.6d](P-P-plot.png)

**Figure 7.2.6c**  **Figure 7.2.6d**

Cumulative frequency curve and P-P plot of ALAR (1) sequence

The histogram of the given historical data overlaid with the corresponding estimated ALAR (1) sequence shows that the data are in close agreement with the ALAR (1) model. Also the P-P plot and Q-Q plot are very close to a straight line.
Hence we conclude that the fitted model, ALAR (1) model, is adequate for modeling the given time series observations \( \{X_n\} \).

Next we shall verify whether the fitted model generates the historical sequence. The historical and simulated sequence of observations is presented in Figure 7.2.7 (the solid-line represents simulated sequence of observations and dot-line represents historical data). From this figure we can see that the historical data bear a close resemblance to the simulated data. It is observed that the historical and the simulated series agree in all the 100 repetitions we have made. Thus we can conclude that ALAR (1) process is appropriate for modeling the given set of time series data.

![Figure 7.2.7](image)

**Figure 7.2.7**

**Time series plot of simulated (solid line) and historical data (dot line) of ALAR (1) sequence**

For the ALAR (1) process we know \( E(X_n | X_{n-1} = x) = \rho x + (1 - \rho) \mu \). Now we forecast observations using the same technique adopted in simulation and the forecast is done by taking \( X_0 \) as the last observation in the historical data. The

202
generated 12 observations of $X_i$ are calculated and hence we can forecast index of industrial production in the USA for the year 2006. The results are given in the following table.

<table>
<thead>
<tr>
<th>t</th>
<th>$Y_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2006</td>
<td>112.098</td>
</tr>
<tr>
<td>February 2006</td>
<td>112.099</td>
</tr>
<tr>
<td>March 2006</td>
<td>112.102</td>
</tr>
<tr>
<td>April 2006</td>
<td>112.104</td>
</tr>
<tr>
<td>May 2006</td>
<td>112.107</td>
</tr>
<tr>
<td>June 2006</td>
<td>112.110</td>
</tr>
<tr>
<td>July 2006</td>
<td>112.113</td>
</tr>
<tr>
<td>August 2006</td>
<td>112.117</td>
</tr>
<tr>
<td>September 2006</td>
<td>112.120</td>
</tr>
<tr>
<td>October 2006</td>
<td>112.123</td>
</tr>
<tr>
<td>November 2006</td>
<td>112.126</td>
</tr>
<tr>
<td>December 2006</td>
<td>112.129</td>
</tr>
</tbody>
</table>

Table 7.2.1

Forecast of index of industrial production in USA for the year 2006