Statistical domain approaches in forecasting techniques

K Sreenivasulu and G Mokesh Rayalu

Abstract

Time series analysis has been created by four types of experts namely Engineers and Physical Scientists; Economists; Applied Statisticians and Économetrists; and Mathematical Statisticians and Probabilists. Time series analysis was divided into two approaches namely Frequency Domain and Time Domain approaches. For Engineers and Physical Scientists, it is natural to regard a time series as made up of oscillations at different frequencies. This gives time domain analysis. There is duality between the frequency domain and time domain approaches to the time series analysis.

Keywords: Statistical domain approaches, forecasting techniques

1. Introduction

In univariate time series regression models, the regressor variable is time. A linear trend relationship can be written as

\[ Y_T = \alpha + \beta T + \epsilon \]  

(1.1)

Where \( T \) indicates time. The \( T \) variable may be specified in many ways. When \( T \) has zero mean, the normal equations for fitting (1.1) will become

\[ a = \bar{Y} \text{ and } b = \frac{\sum Y}{\sum T} \]

One way of modeling such behavior is by means of autoregression. The simplest autoregressive model is

\[ Y_t = \alpha + \beta Y_{t-1} + \epsilon_t \]  

(1.2)

This is called a first order autoregressive model and is denoted by AR(1). The order indicates the maximum lag in the equation. From the equation (1.2), we make the following assumptions about \( \epsilon \) variable

\[ E(\epsilon_i) = 0 \quad \text{for all } i \]

\[ E(\epsilon_i^2) = \sigma^2 \quad \text{for all } i \]

\[ E(\epsilon_i \epsilon_j) = 0 \quad \text{for all } i \neq j \]  

(1.3)

These assumptions define a white noise series. Here the crucial question is, how does the \( Y \) series behave over time. Assuming that process stated a very long time ago, we take

\[ Y_t = \alpha (1 + \beta + \beta^2 + \ldots) + (\epsilon_t + \beta \epsilon_{t-1} + \beta^2 \epsilon_{t-2} + \ldots) \]  

(1.4)

\[ E(Y_t) = \alpha (1 + \beta + \beta^2 + \ldots) \]
This expectation exists only when the infinite geometric series on the right hand side has limit. The necessary and sufficient conditions is \(|\beta| < 1\).

The expectation is then
\[
E(Y_t) = \mu = \frac{\alpha}{1 - \beta}
\]  
(1.5)

Here variance \(Y\) will be
\[
\text{Var}(Y) = \sigma^2 - \frac{\sigma^2}{1 - \beta^2}
\]  
(1.6)

The \(Y\) series has a constant unconditional variance, independent of time.

The covariance of \(Y\) with a lagged value itself is known as autocovariance.

The lag auto covariance is defined as
\[
\gamma_s = \beta^s \sigma^2_y \quad s = 0, 1, 2, \ldots
\]  
(1.7)

So that first lag autocovariance is \(\gamma_1 = \beta \sigma^2_y\).

The autocovariances depend only on the lag length and are independent of \(t\) Dividing the covariances by the variance gives the autocorrelation coefficients, also known as series correlation coefficients. These will be defined as
\[
\rho_s = \frac{\gamma_s}{\sigma^2_y} . \quad s = 0, 1, 2, \ldots
\]  
(1.8)

Plotting the autocorrelation coefficients against the lag lengths gives the correlogram of series.

When \(|\beta| < 1\) the mean, variance, and covariances of \(Y\) series are constants independent of time. The \(Y\) series is then said to be weekly or covariance stationary.

When \(\beta = 1\), the AR(1) process is said to have unit root. The equation becomes.
\[
Y_t = \alpha + Y_{t-1} + \epsilon_t
\]  
(1.9)

Which is called a random walk with drift. The conditional expectation and conditional variance are
\[
E(Y_t/Y_0) = at + Y_0
\]

Which increases or decreases without limit as \(t\) increases. 

Which increases without limit. In this case the unconditional mean and variance do not exist. The \(Y\) series is then said to be nonstationary.

2. Autoregressive Model with Order \(p\): AR(\(p\))

A common approach for modeling univariate time series is the autoregressive (AR) model. The AR\((p)\) model is defined as

\[
X_t = \delta + \phi X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \epsilon_t
\]  
(2.1)

where \(X_t\) is the time series, \(\delta = \left(1 - \sum_{i=1}^{p} \phi_i \right) \mu, \) where \(\mu\) is the process mean
\(\epsilon_t\) is the white noise
\(\phi_1, \phi_2, \ldots, \phi_p\) are the parameters of the model
\(p\) is the order of the AR model

Some constraints are necessary on the values of the parameters of this model so that model remains wide-sense stationary. For an AR\((P)\) model to be wide-sense stationary, the roots of the polynomial \(Z^p - \sum_{i=1}^{p} \phi_i Z^{p-i}\) must lie within the unit circle, i.e., each root \(Z_i\) must satisfy \(|Z_i| < 1\).

3. Moving Average Model With Order \(q\): MA(\(q\))

In time series analysis, the moving average (MA) model is common approach for modeling univariate time series models. The moving average model with order qMA(\(q\)) is

\[
X_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}
\]  
(3.1)
Where, $\mu$ is the mean of the series, 
$\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the model, 
$\varepsilon_t, \varepsilon_{t-1}, \ldots$ are white noise error terms.
$q$ is the order of the moving average model.

### 4. Autoregressive moving average (arma) model

It is convenient to use the notation ARMA $(p, q)$, where $p$ is the order of the autoregressive part and $q$ is the order of the moving average part.

The general AR $(p)$ model was represented as

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \varepsilon_t$$  \hspace{1cm} (4.1)

Multiplying both sides by $X_{t+k}$ yields

$$X_{t+k}X_t = \phi_1 X_{t+k}X_{t-1} + \phi_2 X_{t+k}X_{t-2} + \ldots + \phi_p X_{t+k}X_{t-p} + X_{t-k} \varepsilon_t$$

Taking expected value both sides and assuming stationarity gives

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \ldots + \phi_p \gamma_{k-p}$$

Where $\gamma_k$ is the covariance between $X_t$ and $X_{t-k}$

The MA $(q)$ model is written as

$$X_k = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}$$

Multiplying both sides by $X_{t-k}$ yields

$$X_{t-k}X_t = (\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}) \times (\varepsilon_{t-k} - \theta_1 \varepsilon_{t-k-1} - \theta_2 \varepsilon_{t-k-2} - \ldots - \theta_q \varepsilon_{t-k-q})$$  \hspace{1cm} (4.2)

The expected value of the above equation will depend upon the value of $k$.

If $k=0$, and all other terms of the equation on (4.2) dropout because

$$E(\varepsilon_i \varepsilon_{i+j}) = 0 \text{ for } i \neq 0 \text{ and } E(\varepsilon_i \varepsilon_{i+j}) = \sigma_\varepsilon^2 \text{ for } i = 0$$

Thus (4.1) becomes, 

$$\gamma_0 = \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 + \theta_2^2 \sigma_\varepsilon^2 + \ldots + \theta_q^2 \sigma_\varepsilon^2$$

To obtain the initial estimates for ARMA models, combine AR and MA models:

$$\gamma_k = \phi_1 E(X_t X_{t-k}) + \ldots + \phi_p E(X_{t-p} X_{t-k}) + E(\varepsilon_t X_{t-k}) - \theta_1 E(\varepsilon_{t-1} X_{t-k}) - \ldots - \theta_q E(\varepsilon_{t-q} X_{t-k})$$  \hspace{1cm} (4.3)

If $k>q$, the terms $E(\varepsilon_t X_{t-k}) = 0$ which leaves

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \ldots + \phi_p \gamma_{k-p}$$

When $k<q$, the past errors and the $X_{t-k}$ will be correlated and the autocovariances will be affected by the moving average part of the process, requiring that it will be included. The variance and autocovariances of an ARMA $(1, 1)$ process are therefore obtained as follows:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Multiplying both sides by $X_{t-k}$ gives

$$X_{t-k}X_t = \phi_1 X_{t-k} X_{t-1} + X_{t-k} \varepsilon_t - \theta_1 X_{t-k} \varepsilon_{t-1}$$

"507"
Taking expected values both sides results in
\[
E[X_{t-k}X_t] = \phi E[X_{t-k}X_{t-1}] + E[X_{t-k}e_t] - \theta E[X_{t-k}e_{t-1}]
\]
If \( k = 0 \),
\[
\gamma_0 = \theta_1 \gamma_1 + E[(\phi_1 X_{t-1} + e_t - \theta_1 e_{t-1})e_t] - \theta_1 E[(\phi_1 X_{t-1} + e_t - \theta_1 e_{t-1})e_{t-1}]
\]  
(4.4)

Since \( x_t = \phi_1 k_{t-1} + e_t - \theta_1 e_{t-1} \)
\[
\gamma_0 = \phi_1 \gamma_1 + \sigma_e^2 - \theta_1 (\phi_1 - \theta_1) \sigma_e^2
\]

Similarly if \( k = 1 \),
\[
\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_e^2
\]  
(4.5)

Solving the equations (4.2) ad (4.3) for \( \gamma_0 \) and \( \gamma_1 \) we get
\[
\gamma_0 = \frac{1 + \theta_1^2 - 2 \phi_1 \theta_1}{1 - \theta_1^2}
\]  
(4.6)
\[
\gamma_1 = \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 - \theta_1^2}
\]  
(4.7)

Dividing (4.5) by (4.4) gives
\[
\rho_t = \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2 \phi_1 \theta_1}
\]  
(4.8)

5. Autoregressive integrated moving average (aima) model
To identify the appropriate ARIMA model for a time series, one has to begin by identifying the order (r) of differencing needing to stationarize the series and remove the gross features of seasonality, in conjunction with a variance–stabilizing transformation such as lagging or deflating.

The equation for the simplest case ARIMA (1, 1, 1) is as follows
\[
(1 - B)(1 - \phi_1 B)X_t = \mu^t + (1 - \phi_1 B)e_t
\]

The terms can be multiplied out and rearranged as follows.
\[
\left[1 - B \left(1 + \phi_1 \right) + \phi_1 B^2 \right]X_t = \mu^t + e_t - \theta_1 e_{t-1}
\]
\[X_t = (1 + \phi_1)X_{t-1} - \phi_2 X_{t-2} + \mu^t + e_t - \theta_1 e_{t-1}
\]  
(5.1)

In this form, the ARIMA model looks more like a conventional regression equation, except that there is more than one error term on the right hand side. ARIMA models are quite flexible and can represent a wide range of characteristics of time series that occur in practice.

5.1 ARIMA Model : \( p \)th order autoregressive model
\[
Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t
\]  
(5.2)

5.2 ARIMA Model: \( q \)th order moving average model
\[
Y_t = \mu + \epsilon_t - W_1 \epsilon_{t-1} - W_2 \epsilon_{t-2} - \ldots - W_q \epsilon_{t-q}
\]  
(5.3)

5.3 ARIMA Model: ARMA (p, q) model
\[
Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t - W_1 \epsilon_{t-1} - W_2 \epsilon_{t-2} - \ldots - W_q \epsilon_{t-q}
\]  
(5.4)
5.4 ARIMA (0, 1, 1) Model

\[ Y_t - Y_{t-1} = \epsilon_t - w_{t-1} \]  
(5.6)

6. Autoregressive distributed lag models (ARDL)

In case where the variables in the long-run relation of interest are trend stationary, the general practice has been to de-trend the series and to model the detrended series as stationary distributed lag or autoregressive distributed lag (ARDL) models. Estimation and inference concerning the long-run properties of the model are then carried out using standard asymptotic normal theory. The analysis becomes more complicated when the variables are difference–stationary, or integrated of order 1. The autoregressive distributed lag model with \( p \) lags of dependent variable \( Y_t \) and \( q \) lags of additional regressor \( X_t \), ADL \((p, q)\) is defined as:

\[
\begin{align*}
ADL(p,q) & = Y_t = \beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p} + \delta_1 X_{t-1} + \ldots + \delta_q X_{t-q} + \mu_t, \\
\beta(L) Y_t & = \beta_0 + \delta(L) X_{t-1} \quad \text{with lag-polynomials defined by} \\
\beta(L) & = 1 - \beta_1 L - \ldots - \beta_p L_p, \\
\delta(L) & = \delta_1 + \delta_2 L + \ldots + \delta_q L_{q-1} \\
k \text{additional predictors} & = ADL(p,q) \quad (6.1)
\end{align*}
\]

Model Assumptions

1. \( \mathbb{E}\left[ u_t / Y_{t-i}, X_{t-i}, \ldots, X_{t-k-i} \right] = 0 \)  
(6.4)

2. (a). \( \left( Y_t, X_{t-i}, \ldots, X_{t-k-i} \right) \) are (strictly) stationary  
(b). \( \left( Y_t, X_{t-i}, \ldots, X_{t-k-i} \right) \) are ergodic  
(6.5)

3. \( Y_t, X_{t-i}, \ldots, X_{t-k-i} \) have nonzero, finite fourth moments

4. no perfect multicollinearity

7. Non stationary time series regression models

A time series \( X_t \) is said to be stationary if its expected value and population variance are independent of time and if the population covariance between its values at time \( t \) and \( t+s \) depends on \( s \) but not on time.

An example of a stationary time series is an AR (1) process

\[ X_t = \beta_2 X_{t-1} + \epsilon_t \]  
(7.1)

Here \(-1 < \beta_2 < 1\) and \( \epsilon_t \) is white noise

If equation (7.1) is valid for time period \( t \), it is also valid for time period \( t-1 \) i.e. the series is stationary.

\[ X_{t-1} = \beta_2 X_{t-2} + \epsilon_{t-1} \]

Substituting for \( X_{t-1} \) in equation (7.1), we get

\[ X_t = \beta_2^2 X_{t-2} + \beta_2 \epsilon_{t-1} + \epsilon_t \]

Continuing this process of lagging and substituting, we get

\[ X_t = \beta_2^2 X_0 + \beta_2^{t-1} \epsilon_1 + \ldots + \beta_2 \epsilon_{t-1} + \epsilon_t \]  
(7.2)

In the previous examples, if \( \beta_2 \) is equal to 1, the original series becomes

\[ X_t = X_{t-1} + \epsilon_t \]  
(7.3)
This is an example of a nonstationary process which is known as a random walk.
If it starts at \( X_0 \) at time 0, its value at time \( t \) is given by

\[
X_t = X_0 + \varepsilon_t + \ldots + \varepsilon_t
\]  
(7.4)

The key difference between this process and the corresponding A.R(1) process is that the contribution of each innovation is permanently built into the time series. In the more general version of the autoregressive process with the constant \( \beta_1 \), the process becomes a random walk with drift, if \( \beta^2 \) equals 1.

\[
X_t = \beta_t + X_{t-1} + \varepsilon_t
\]  
(7.5)

If the series starts at \( X_0 \) at time 0, \( X_t \) is given by

\[
X_t = X_0 + \beta_{t} + \varepsilon_t + \ldots + \varepsilon_t
\]

The expectation of \( X_t \) at time ‘0’, \( E(X_0 + \beta_t) \) is a function of ‘t’. Another common example of a nonstationary time series is one possessing a time trend:

\[
X_t = \beta_1 + \beta_2 t + \varepsilon_t
\]  
(7.6)

This type of trend is sometimes described as a deterministic trend. The expected value of \( X_t \) at time ‘t’ \( E(\beta_1 + \beta_2 t) \) is not independent of \( t \) and so \( X_t \) is nonstationary.

8. Conclusions
Time series analysis has advance from univariate modeling based on a single variable to multivariate models that employ the interrelationships between several such variables. Constructing such models requires the performing of tests to determine and to discover the interactions that exist between a given time series variables and one or more other variables. The given variable can be influenced not only by certain exogenous events occurring at particular points in time but also by contemporaneous, lagged and leading values of another variable or additional variables.

9. References