International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 8
August 2017

Stock Price Prediction Using ARIMA

Deepthi Praveenlal Kuttichira
Amrita Vishwa Vidyapeetham

ABSTRACT

Prediction of stock prices is interesting and challenging. Since all the variables affecting a stock price are not known, modeling based on independent variables is difficult.

We made predictions for stock price for two companies using Autoregressive Integrated Moving Average (ARIMA). ARIMA is a univariate time series model. The companies considered were two pharma companies listed in National Stock Exchange. The granularity of time was minute. Estimations done using the forecasting package in the tool R.

Results evaluated using the percentage error between the predicted price and actual price. It was encouraging. The lags of the model was determined by using autocorrelation and partial correlation plots.

The benefits and limitations discussed.

Keywords

Autoregression, Moving Average, Autoregressive Moving Average, Autoregressive Integrated Moving Average

INTRODUCTION

Stock price prediction is a challenging problem as the market is quite unpredictable. The prices fluctuate rapidly, making predictions difficult. There are many who advocate that stock prices are stochastic and cannot be predicted [1]. The recent studies challenge this argument [2]. These studies showed that current prices are affected by past prices. This provides scope for analyzing trends of stock market and making predictions based on it.

Time series analysis can be used to analyze the trend of a stock market. The stock prices vary in time, forming a time series. Regression is a method commonly used for prediction. It has a dependent variable that depends on one or more independent variables. Stock price can be taken as a dependent variable. The factors affecting it, like current demand, government policies, buyers' assessments etc, are independent variables. In the case of stock market all the factors affecting it are not known. In such situations a time series analysis is possible. In this method the current price is considered as dependent variable and previous prices as independent variables. Dependency of current price on previous price is the basis of time series analysis.

There are many time series analysis methods like Moving Average (MA), Auto regression(AR), Autoregressive Moving Average(ARMA) and Autoregressive Integrated Moving Average (ARIMA)[3][4]. MA tries to model the current price as a function of the white noise. AR tries to model current price as the function of previous prices. ARMA combines both and models the current price as a function of both the white noise and previous prices. The time series data of a stock price is non- stationary. This has to be made stationary, before modeling is done. ARIMA does this by differentiating the series until the series become stationary. Seasonal Autoregressive Integrated Moving Average (SARIMA) considers seasonal patterns in addition to white noise and previous price, to model the current price.

We used ARIMA to model and to make predictions. The time series we considered was minute- wise, hence seasonal effects were not seen in the data. Since the series was non-stationary, it had to be made into stationary first.

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 8
August 2017

METHODOLOGY

ARIMA is a Box-Jenkins method. It is a univariate model, that it does modeling taking in account only one time series. There are three steps in a box Jenkins method.

Model Selection- In this step the variables considered are ensured to be stationary. For this autocorrelation and partial autocorrelation plots of the dependent time-series are plotted. Based on these plots, what components of MA and AR has to be used is determined.

Parameter Estimation- In this step, the coefficients of the independent variables are estimated. Best fit by least square method is usually used. Log-likelihood method can also be used for estimation.

Model checking- This step checks if the model fitted for the data is correct. This is checked using residuals. Residual is the difference between actual data and the data obtained using the fitted model. A model is said to be suitable for the data only if the residuals form a Gaussian white noise.

If the residual obtained is not white noise, then the steps has to be repeated until white noise residual is obtained. Akaike values help in model selection, as they determine the quality of the model selected.

An ARIMA model is of the form,

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \ldots + \theta_q w_{t-q} \quad (1)$$

where x_t is stationary, and θ , ϕ are coefficients. Prices are labelled as x at times t, t-1, t-2 etc. We assumed that w at time tis a Gaussian white noise series with mean zero and variance 2. An ARIMA model is defined as ARIMA(p,d,q)[5]. p, d, q are the parameters. p is the lag of AR - the number of previous data points on which the current data point depends. q is the lag of MA - the number of previous residual values on which the current data point depends. d is the number of times the series was differenced inorder to make it stationary. We used the tool R to fit ARIMA model to our data. The data considered was the minute-wise stock data of CIPLA and ABBOT INDIA, two Pharmaceutical companies, for the month August 2014. The model was fit using 4000 data points.

RESULT

The results obtained when auto.arima () function in R was applied is shown in the tables 1 and 2. It shows the actual price, predicted price by the fitted ARIMA model and the error percent between the actual and predicted price.

Table 1. Prediction for ABBOTINDIA

SL:No	Actual	Arima (0,1,0)	Arima error percent
1	2109	2094	0.711
2	2109	2094	0.711
3	2109	2094	0.711
4	2109	2094	0.711
5	2109	2094	0.711

Table 2. Prediction for CIPLA

SL:No	Actual	Arima (3,1,2)	Arima error percent
1	455	459.45	0.978
2	454.05	460.35	1.387
3	455.45	461.25	1.273
4	456	462.15	1.348
5	456.9	463.05	1.346

DISCUSSION

ARIMA uses a linear fit. It can be observed from the table that the predictions are fairly accurate. As the number of predictions increased, the error percentage also increased.

The ARIMA model that fit for the company ABBOTINDIA was ARIMA(0, 1,0). It means that the current price of this company is not dependent on previous price. There was a trend that existed, which was removed by differencing the series once. This is a random walk model. For the company CIPLA, the model that fit was ARIMA(3,1,2). This means that the current price of this company depends on the past three prices and the past two residuals. It also exhibits a trend which can be removed by differencing the series once. The model that fit for these two time series are different.

While one exhibits a random walk tendency, other clearly exhibits dependency on past prices.

As the accuracy steadily decreases with the number of predictions made, it cannot be used for long term predictions in stock market. Also fitting a model is time consuming process. ARIMA needs a time series of considerable length, say 1000 data point, to fit a model. Model selection is a tedious task. But manual model selection is next to impossible, as there are so many options for it in Box-Jenkins method. Use of tool is inevitable for this.

CONCLUSION

ARIMA can effectively analyze the trend of a stock time series. But it is not efficient in making long term predictions.

REFERENCES

[1] B. G. Malkiel, “The efficient market hypothesis and its critics,” TheJournal of Economic Perspectives, vol. 17, no. 1, pp. 59–82, 2003.Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

[2] H. Bessembinder and K. Chan, “The profitability of technical trading rules in the asian stock markets,” Pacific-Basin Finance Journal, vol. 3, no. 2, pp. 257–284, 1995.Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[3] A. A. Ariyo, A. O. Adewumi, and C. K. Ayo, “Stock price predictionusing the arima model,” pp. 106–112, 2014.

[4] A. Gunasekarage and D. M. Power, “The profitability of moving averagetrading rules in south asian stock markets,” vol. 2, no. 1. Elsevier, 2001, pp. 17–33.

[5] Robert H Schumway, David S Stoffe “Time Series Analysis and Its Applications”.,2000