Time Series Analysis

ARMA Models: Data Examples

Nicoleta Serban, Ph.D.
Professor
Stewart School of Industrial and Systems Engineering

IBM Stock Price: Forecasting

About This Lesson

IBM Stock Price

Stock Price:

- Perceived company's worth
- Multiplied by number of shares give the total company's worth
- affected by a number of things including volatility in the market, current economic conditions, and popularity of the company

Study Objective:

- Develop a model to predict IBM stock price given that no major events are to be released

Time Series Data:

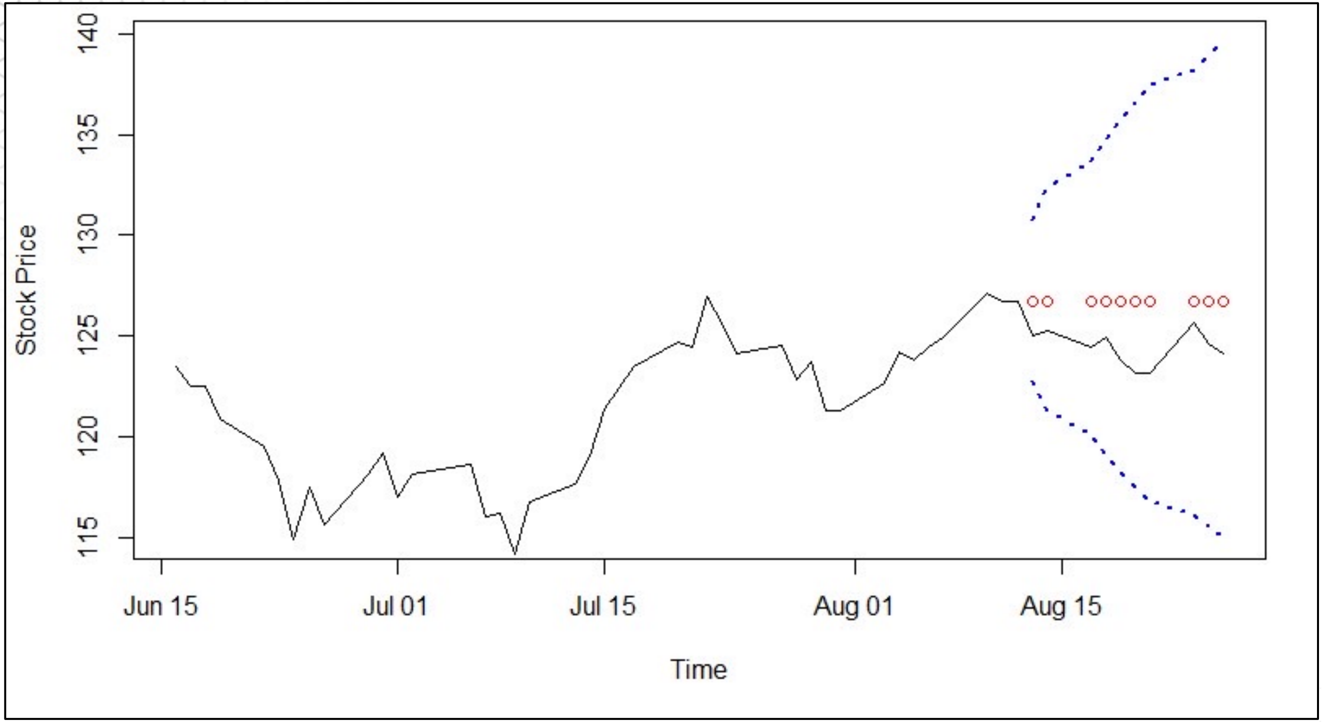
- Daily stock price from January 2nd 1960 until April 18th 2017
- High, Low, Close, Adj.Close

ARIMA Forecasting

```
## Forecasting with ARIMA: 10 Days Ahead
nfit = n-10
outprice = arima(ts.price[1:nfit], order = c(porder,1,qorder),method = "ML")
outpred = predict(outprice,n.ahead=10)
ubound = outpred$pred+1.96*outpred$se; ymax = max(ubound)
lbound = outpred$pred-1.96*outpred$se; ymin = min(lbound)

plot(ibm.date[(n-50):n],exp(ts.price[(n-50):n]),type="l", ylim=c(ymin,ymax),
xlab="Time", ylab="Stock Price")
points(ibm.date[(nfit+1):n],exp(outpred$pred),col="red")
lines(ibm.date[(nfit+1):n],exp(ubound),lty=3,lwd= 2, col="blue")
lines(ibm.date[(nfit+1):n],exp(lbound),lty=3,lwd= 2, col="blue")
```

ARIMA Forecasting



Prediction Accuracy

Prediction Error Measures

- Compare observed response Y_i to the predicted $Y_i{}^*$
- Mean squared prediction error (MSPE) = $\frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i*)^2$
- Mean absolute prediction errors (MAE) = $\frac{1}{n}\sum_{i=1}^{n}|Y_i - Y_i*|$
- Mean absolute percentage error (MAPE) = $\frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - Y_i*|}{Y_i}$
- Precision error (PM) = $\frac{\sum_{i=1}^{n}(Y_i - Y_i*)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2}$
- Confidence Interval error (CIM) = $\frac{1}{n}\sum_{i=1}^{n} I(Y_i* \notin CI)$

Prediction Error Measure Insights

Mean squared prediction error (MSPE)

- Appropriate for models estimated by minimizing square prediction errors. It depends on scale and it is sensitive to outliers

Mean absolute prediction errors (MAE)

- Appropriate for models estimated by minimizing absolute prediction errors. It depends on scale but robust to outliers.

Mean absolute percentage error (MAPE)

- Appropriate for models estimated by minimizing absolute prediction errors. It does not depend on scale and it is robust to outliers.

Precision error (PM)

- Appropriate for models estimated by minimizing square prediction errors. It does not depend on scale.

Confidence Interval error (CIM)

Prediction Error Measure Insights

Mean squared prediction error (MSPE)
- Appropriate for models estimated by minimizing square prediction errors. It depends on scale and it is sensitive to outliers

Mean absolute prediction errors (MAE)
- Appropriate for models estimated by minimizing absolute prediction errors. It depends on scale but robust to outliers.

Mean absolute percentage error (MAPE)
- Appropriate for models estimated by minimizing absolute prediction errors. It does not depend on scale and it is robust to outliers.

Precision error (PM)
- Appropriate for models estimated by minimizing square prediction errors. It does not depend on scale.

Confidence Interval error (CIM)

- **Choose measures that do not depend on scale**
- **Use appropriate measure depending on the estimation approach, e.g. normal estimation uses square prediction errors whereas t-distribution estimation uses absolute prediction error.**
- **Many time series model estimation is similar to regression modeling, which are estimated by minimizing sum of least squares hence use precision measure**

Prediction Accuracy: 10 Days Ahead

```
## Compute Accuracy Measures
obsprice = exp(ts.price[(nfit+1):n])
predprice = exp(outpred$pred)
### Mean Squared Prediction Error (MSPE)
mean((predprice-obsprice)^2)
### Mean Absolute Prediction Error (MAE)
mean(abs(predprice-obsprice))
### Mean Absolute Percentage Error (MAPE)
mean(abs(predprice-obsprice)/obsprice)
### Precision Measure (PM)
sum((predprice-obsprice)^2)/sum((obsprice-
mean(obsprice))^2)
### Does the observed data fall in the prediction
intervals?
sum(obsprice<exp(lbound))+sum(obsprice>exp(ubound))
```

Accuracy Measures

MSPE = $\frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i^*)^2$

MAE = $\frac{1}{n}\sum_{i=1}^{n}|Y_i - Y_i^*|$

MAPE = $\frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - Y_i^*|}{Y_i}$

PM = $\frac{\sum_{i=1}^{n}(Y_i - Y_i^*)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$

Georgia Tech

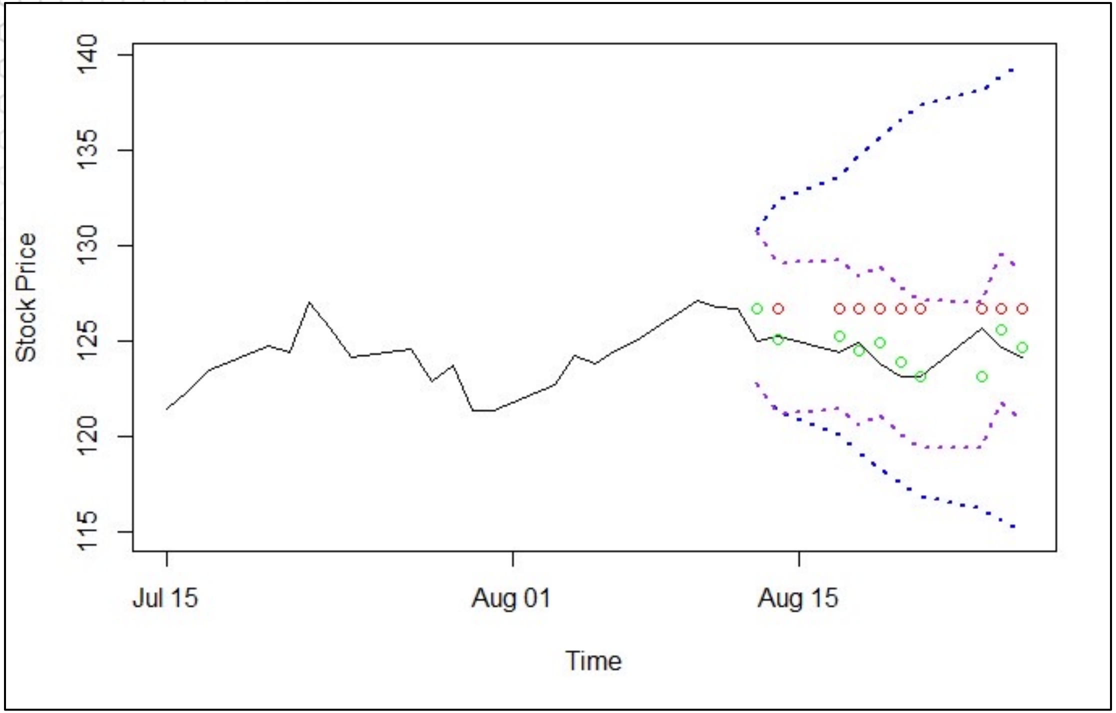
Prediction Accuracy: 10 Days Ahead

```
> ### Mean Squared Prediction Error (MSPE)
> mean((predprice-obsprice)^2)
[1] 5.816396
> ### Mean Absolute Prediction Error (MAE)
> mean(abs(predprice-obsprice))
[1] 2.271224
> ### Mean Absolute Percentage Error (MAPE)
> mean(abs(predprice-obsprice)/obsprice)
[1] 0.01829638
> ### Precision Measure (PM)
> sum((predprice-obsprice)^2)/sum((obsprice-mean(obsprice))^2)
[1] 8.840339
> ### Does the observed data fall outside the prediction intervals?
> sum(obsprice<exp(lbound))+sum(obsprice>exp(ubound))
[1] 0
```

ARIMA Forecasting: 1 rolling day vs 10 days

```
## Daily Prediction over a period of 10 days
outpred.10 = NULL
ubound.10 = NULL
lbound.10 = NULL
n = length(ts.price)
for(i in 1:10){
   nfit = n-(10-i+1)
   outprice = arima(ts.price[1:nfit], order = c(porder,1,qorder),method = "ML")
   pred.1 = predict(outprice,n.ahead=1)
   outpred.10 = c(outpred.10,pred.1$pred)
   ubound.10 = c(ubound.10, pred.1$pred+1.96*pred.1$se)
   lbound.10 = c(lbound.10, pred.1$pred-1.96*pred.1$se)
  }
predprice.10 = exp(outpred.10)
```

ARIMA Forecasting: 1 rolling day vs 10 days



Prediction Accuracy: 1-Day Ahead

```
> ## Compute Accuracy Measures
> predprice.10 = exp(outpred.10)
> ### Mean Squared Prediction Error (MSPE)
> mean((predprice.10-obsprice)^2)
[1] 1.294451
> ### Mean Absolute Prediction Error (MAE)
> mean(abs(predprice.10-obsprice))
[1] 0.8947639
> ### Mean Absolute Percentage Error (MAPE)
> mean(abs(predprice.10-obsprice)/obsprice)
[1] 0.007171295
> ### Precision Measure (PM)
> sum((predprice.10-obsprice)^2)/sum((obsprice-mean(obsprice))^2)
[1] 1.967435
> ### Does the observed data fall outside the prediction intervals?
> sum(obsprice<exp(lbound.10))+sum(obsprice>exp(ubound.10))
[1] 0
```

Summary