WORCESTER POLYTECHNIC INSTITUTE

PORTFOLIO VALUATION AND RISK MANAGEMENT

ARIMA/Margin Call Project Report

Author:
Yijiang XU
Weixi LIU

Instructor:
Professor Marcel Blais

November 12, 2017

Contents

List of Figures

Chapter 1

Introduction

In this project, in chapter 2 to 4, we use a stock data and a forex data to fit ARIMA model, by using AIC and BIC common critirions to pick a best model, then we use the model to predict future data. In chapter 5, we create a interactive Broker paper account to exercise margin call trades.

Chapter 2

Data Collection

Through bloomberg terminal, we download 2-years GBP_USD_FOREX (the ratio of GBP currency and USD currency), through Yahoo Finance[1], we download 2-years AAPL(AAPL Inc) historical data and created a program which extracted all data. Python Pandas package helps us to read excel file into data frame. For forex, we extracted PX_LAST value and transferred them into i-dimensional arrays, for stock, we extracted the closed price seperately and transferred them into 1-dimesional arrays. And they would be accessed in all problems of the project.

We exclude the last three data points as testing data, and regard remaining data as training data. We fit the training data to different ARIMA model and use testing data to compare with our prediction value calculated by our model.

[1] https://finance.yahoo.com/

Chapter 3

Methodology

3.1 Time series

Time series are sequences of data sampled over time, so much of the data from financial markets are time series.

3.1.1 Stationary

Def 1. *A process is said to be* ***stationary*** *if* $\forall m, n \in \mathcal{Z}$, *the distribution of* $y_1, y_2, \cdots, y_n$ *and* $y_{1+m}, y_{2+m}, \cdots, y_{n+m}$ *are the same (Ruppert, 2004, p. 102).*

Stationarity is a very strong assumption because it requires that "all aspects" of behavior be constant in time. In reality, most of time series are not **stationary**, but may be **Weakly staionary**, which we will define next. Although a time series is not weak stationary, it would always become weak stationary after be taken several differences. Generally, we can get by with assuming less, namely, weak stationarity. Here we need to first clearify what is weak stationary. According to what we learnt it has to satisfy 3 conditions as below (Ruppert, 2004, p. 102) :
Def 2. *A process is* ***Weakly staionary*** *if*

- $E[y_i] = \mu, \forall i,$
- $Var(y_i) = \sigma^2, \forall i,$
- $Corr(y_i, y_j) = \rho(|i - j|), \forall i, j$ & *some functions* ρ.

3.1.2 Weak white noise

Def 3. *The sequence* $\epsilon_1, \epsilon_2, \cdots$ *is a* ***weak white noise*** *(Ruppert, 2004, p. 103) process with mean* μ, *denoted WhiteNoise(0,* σ_ϵ^2*), if*

- $E(\epsilon_i) = \mu \; \forall i,$
- $Var(\epsilon_i) = \sigma_\epsilon^2 \; \forall i,$
- $Corr(\epsilon_i, \epsilon_j) = 0 \; \forall i \neq j.$

3.1.3 AR(p) processes

Def 4. *The stochastic process Y_t is an $AR(p)$ process (Ruppert, 2004, p. 115) if*

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2-\mu}) + \cdots + \phi_p(Y_{t-p} - \mu) + \epsilon_t,$$

where $\epsilon_1, \cdots, \epsilon_n$ is WhiteNoise(0, σ_ϵ^2).

3.1.4 MA(q) processes

Def 5. *The MA(q) process (Ruppert, 2004, p. 118) is*

$$Y_t = \mu + \epsilon_t - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q},$$

where $\epsilon_1, \cdots, \epsilon_n$ is WhiteNoise(0, σ_ϵ^2).

3.1.5 ARMA(p,q) processes

Def 6. *The ARMA(p,q) process (Ruppert, 2004, p. 120) is*

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \cdots = \phi_p(Y_{t-p} - \mu) + \epsilon_t - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q},$$

where $\epsilon_1, \cdots, \epsilon_n$ is WhiteNoise(0, σ_ϵ^2).

3.1.6 ARIMA(p,d,q) processes

Def 7. *A time series Y_t is said to be ARIMA(p,d,q) (Ruppert, 2004, p. 122) if $\Delta^d Y_t$ is ARMA(p,q), where*

$$\Delta Y_t = Y_t - Y_{t-1}.$$

3.2 Model selection

3.2.1 Maximum likelihood Estimation (MLE)

Let $\bar{y} = [y_1, y_2, \cdots, y_n]^T$ be a vector of data and $\bar{\theta} = [\theta_1, \cdots, \theta_p]^T$ be a vector of parameters for a distribution. Let $f(\bar{y}, \bar{\theta})$ be the density of $\bar{y}$ that depends on $\bar{\theta}$.

Def 8. ***The likelihood function*** *L($\bar{\theta}$) defined as*

$$L(\bar{\theta}) = f(\bar{y}, \bar{\theta}).$$

The maximum likelihooh estimation (MLE)[1] is the value of $\bar{\theta}$ that maximize L($\bar{\theta}$).

[1] https://en.wikipedia.org/wiki/Maximum_likelihood_estimation

3.2.2 How to choose d

We look at sample auto correlation function (SACF) (Ruppert, 2004, p. 103) of

$$Y_t, \Delta Y_t, \Delta^2 Y_t,$$

and check these for evidence of week stationarity. Since lack of auto-correlation is evidence of staionarity, that means SACF deays to 0 quickly. Then we choose the smallest value of d such that $\Delta^d Y_t$ appears staionary.

3.2.3 How to choose p & q

There are two common criteria to use to choose p & q (Ruppert, 2004, p. 124):

- Akaikes information Criterion (AIC):

  $$2 \log L + 2(p + q),$$

  where L is likelihood function evaluated at the Maximum likelihood Estimation(MLE) for $\bar{\theta}$.

- Baytesian information Criterion (BIC):

  $$2 \log L + (\log n)(p + q),$$

  where n is the number of data points in time series.

3.3 Prediction

We use two ways to forecast future value, which are manual prediction and model prediction.

3.3.1 Manual prediction

Since forecasting (Ruppert, 2004, p. 128) AR(p) or MA(p), $p > 1$, processes is very similar with AR(1) or MA(1) processes, we just take AR(1) and MA(1) process as examples.

- Forecasting AR(1) processes.

  Consider the AR(1) process, $Y_t - \mu = \epsilon_t - \theta\epsilon_{t-1}$. The next observation will be

  $$Y_{n+1} = \mu + \epsilon_{n+1} - \theta\epsilon_n. \tag{3.1}$$

  We replace μ and θ by estimated and ϵ_n by the residual $\hat{\epsilon_n}$ in the right-hand side of equation (3.1). And since ϵ_{n+1} is independent of the observed data, we can replace it by 0. Then the forecast is

  $$\hat{Y}_{n+1} = \hat{\mu} - \hat{\theta}\hat{\epsilon_n}.$$

Moreover, since ϵ_{n+1} and ϵ_{n+2} are independent of the observed data, the two-step ahead forecast of $Y_{n+2} = \mu + \epsilon_{n+2} - \theta\epsilon_{n+1}$ is simply $\hat{Y}_{n+2} = \hat{\mu}$. Hence, we can conclude that $\hat{Y}_{n+k} = \hat{\mu}$ for all $k > 2$.

- Forecasting MA(1) process.

  Consider the MA(1) process, $Y_t - \mu = \epsilon_t - \theta\epsilon_{t-1}$. Then the next observation will be

  $$Y_{n+1} = \mu + \epsilon_{n+1} - \theta\epsilon_n.$$

  We replace μ and θ by estimates and ϵ_n by the residual $\hat{\epsilon_n}$, the forecast becomes

  $$\hat{Y}_{n+1} = \hat{\mu} - \hat{\theta}\hat{\epsilon}_n.$$

Forecasting ARMA and ARIMA processes is similar to forecasting AR and MA processes.

3.3.2 Model prediction

Using Python inner function to predict future value directly.

Chapter 4

Implementation

4.1 Using AAPL stock daily log returns as time series

4.1.1 Select staionary time series

We plot 3 sample autocorrelation functions of $y_t(D = 0), \Delta y_t(D = 1), \Delta^2 y_t(D = 2)$ as shown in figure 4.1.

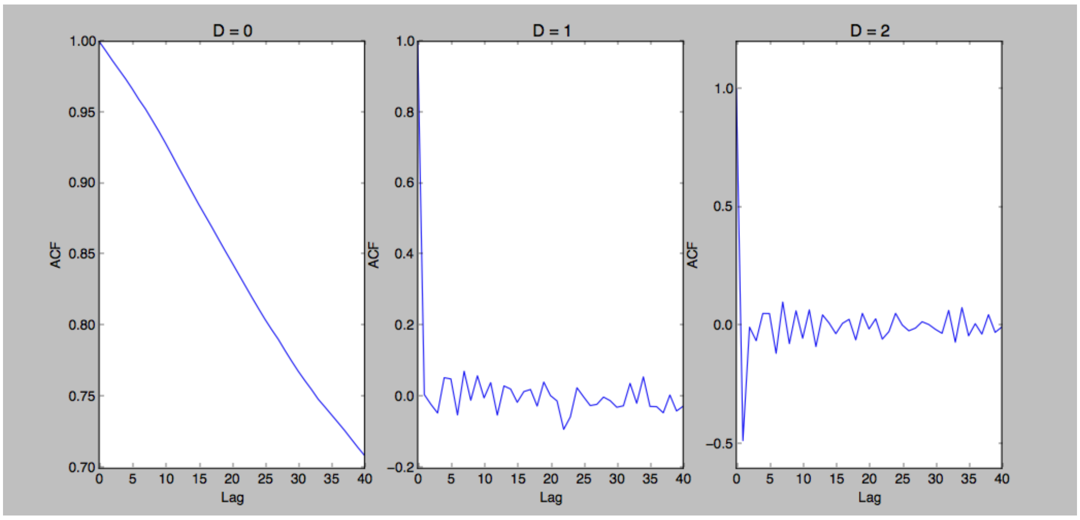

FIGURE 4.1: AAPL Sample Autocorrelation function.

We can see when $D = 1$ and $D = 2$ the two time series decay to 0 very quickly which mean they all appear weak staionary, and we choose the time series with the smallest value of D, which is Δy_t, for studying.

4.1.2 Select Model

We fit the time series by using ARIMA(p,1,q) model with different p, q values, including all cases for p and q in the set $\{0, 1, 2\}$. Then we select a model if this model's AIC or BIC value is minimum. The result is shown as below:

Model	AIC	BIC
ARIMA(1,1,0)	-2858.7173067704593	-2846.07948848
ARIMA(2,1,0)	-2856.848233527661	-2839.99780914
ARIMA(0,1,1)	-2858.722321967311	-2846.08450368
ARIMA(1,1,1)	-2856.7402530582935	-2839.88982868
ARIMA(2,1,1)	-2855.2162743687027	-2834.15324389
ARIMA(0,1,2)	-2856.8210890117302	-2839.97066463
ARIMA(1,1,2)	-2854.942556540382	-2833.87952606
ARIMA(2,1,2)	-2857.408061028908	-2832.13242445

FIGURE 4.2: Values of AIC and BIC of different models that fitted with AAPL time series.

We can see ARIMA(0,1,1) is the best fitted model no matter what using AIC criteria or BIC criteria.

In practice the best AIC model is usually close to the best BIC model and often they are the same model. But, we will see, in our GBP_USD_FOREX part, the best AIC model is different from the best BIC model.

4.1.3 Prediction

As we mentioned before, we exclude the last three points in our data as validation set, and using remaining data as training set to build model. In this way, one can use the validation set to compare with the values calculated by model. We plot validation value, manual predictions and model predictions in one graph

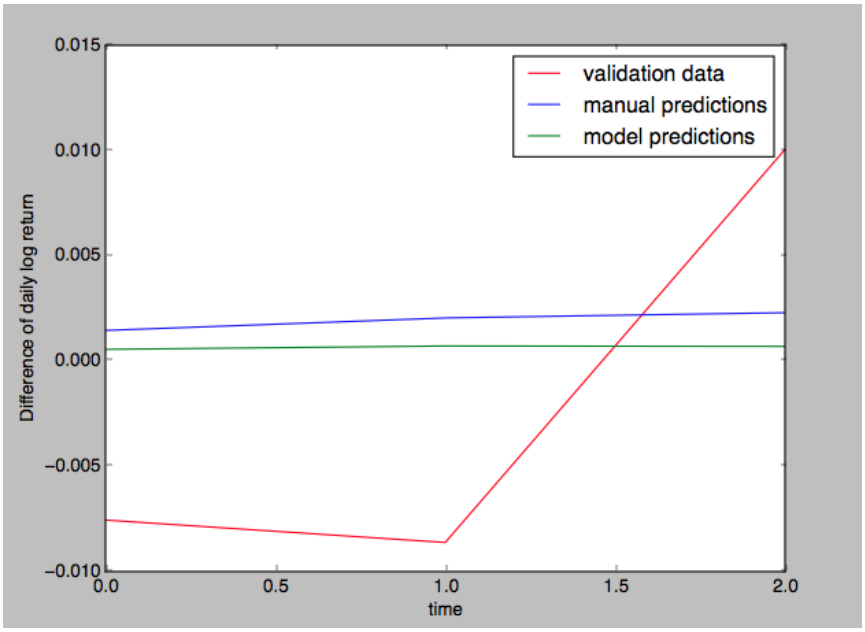

FIGURE 4.3: Validation and prediction curve.

and present the MSE

```
Manual MSE: 8.5045606255e-05
Model MSE: 8.0169821435e-05
```

FIGURE 4.4: Mean Squared Error(MSE)

Since absolute value of the validation data is around 0.00203, the root of MSE is around 0.009 which is much larger than 0.002, and the graph shows the two prediction curve are not fit to validation curve, thus the model doesn't fit the AAPL data well.

4.2 Using GBP_USD_FOREX ratios as time series

4.2.1 Select staionary time series

We plot 3 sample autocorrelation functions of $y_t(D = 0), \Delta y_t(D = 1), \Delta^2 y_t(D = 2)$ as shown in figure 4.1.

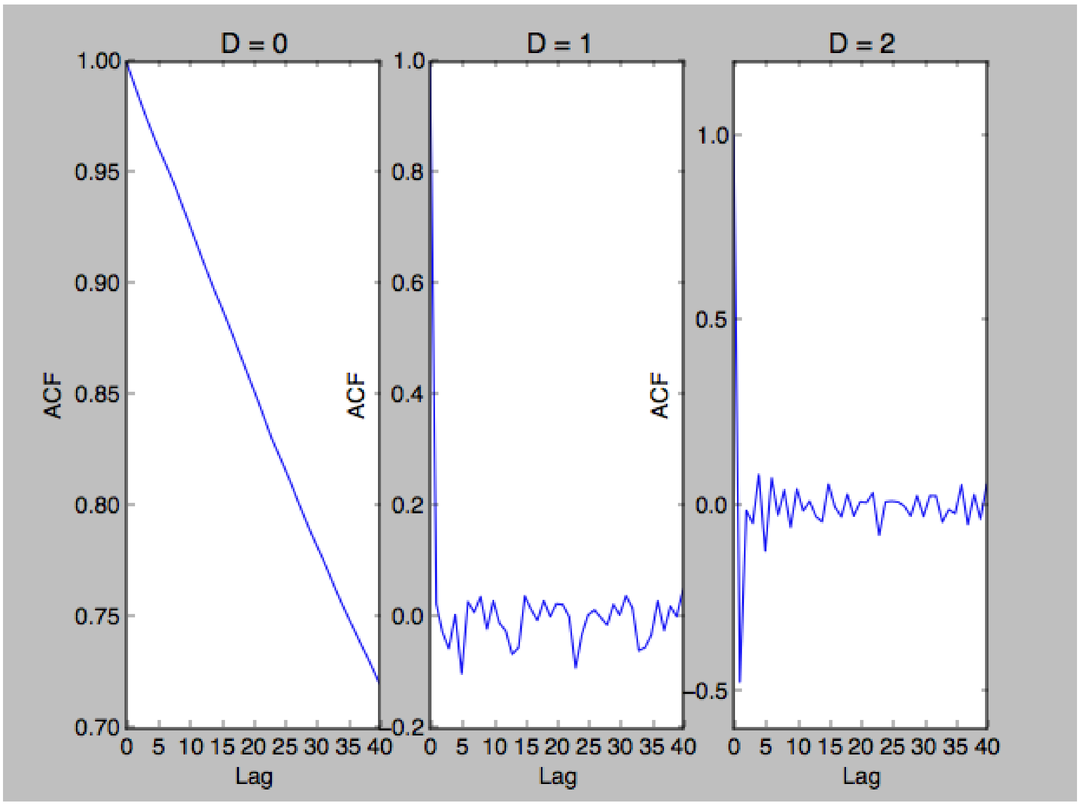

FIGURE 4.5: Forex Sample Autocorrelation function.

We can see when $D = 1$ and $D = 2$ the two time series decay to 0 very quickly which mean they all appear weak staionary, and we choose the time series with the smallest value of D, which is Δy_t, for studying.

4.2.2 Select Model

We fit the time series by using ARIMA(p,1,q) model with different p, q values, including all cases for p and q in the set $\{0, 1, 2\}$. Then we select a model if this model's AIC or BIC value is minimum. The result is shown as below:

Model	AIC	BIC
ARIMA(1,1,0)	-3318.6035713888223	-3305.83632126
ARIMA(2,1,0)	-3316.9836610954726	-3299.96066093
ARIMA(0,1,1)	-3318.6201386448556	-3305.85288852
ARIMA(1,1,1)	-3316.670425265801	-3299.6474251
ARIMA(2,1,1)	-3316.7860212790292	-3295.50727107
ARIMA(0,1,2)	-3316.906220311894	-3299.88322014
ARIMA(1,1,2)	-3316.729327101937	-3295.45057689
ARIMA(2,1,2)	-3322.755918893691	-3297.22141864

FIGURE 4.6: Values of AIC and BIC of different models that fitted with Forex time series.

We can see ARIMA(0,1,1) is the best fitted model by checking BIC criteria, and ARIMA(2,1,2) is the best fitted model by checking AIC criteria.

4.2.3 Prediction

As we mentioned before, we exclude the last three points in our data as validation set, and using remaining data as training set to build model. In this way, one can use the validation set to compare with the values calculated by model.

- ARIMA(0,1,1): We plot validation value, manual predictions and model predictions in one graph

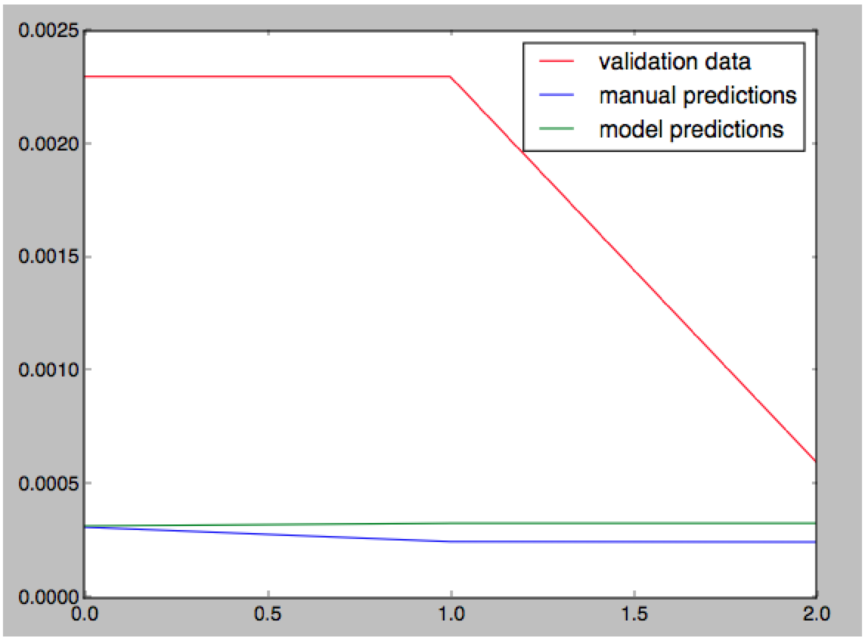

FIGURE 4.7: Validation and prediction curve with first difference data.

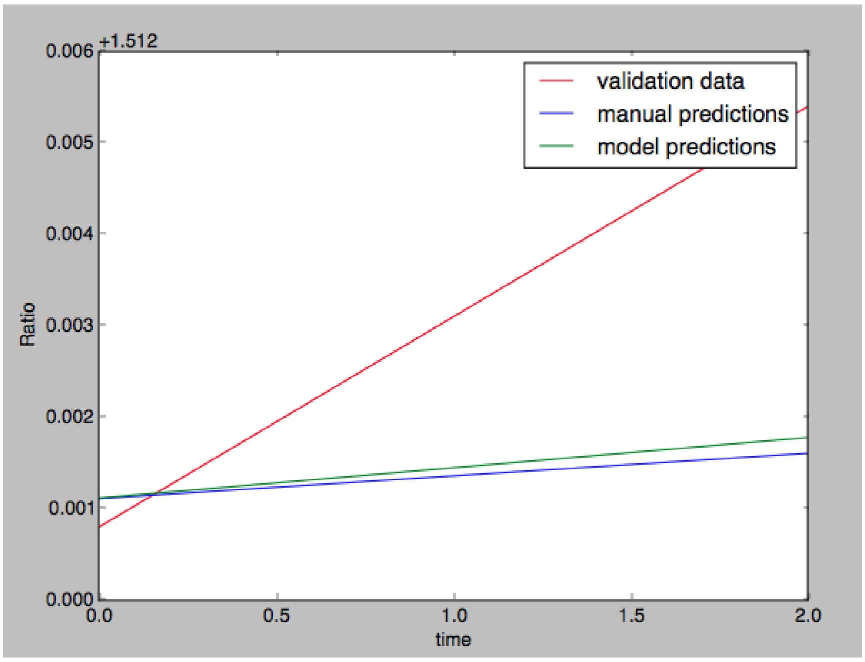

FIGURE 4.8: Validation and prediction curve with original data.

and present the MSE

```
Manual MSE: 2.76061991731e-06
Model MSE: 2.62459772399e-06
```

FIGURE 4.9: Mean Squared Error(MSE)

Since absolute value of the validation data is around 0.00173, and the root of MSE is around 0.00162, and the graph shows the two prediction curve are not fit to validation curve, so the model doesn't fit the AAPL data well.

- ARIMA(2,1,2): We plot validation value, manual predictions and model predictions in one graph

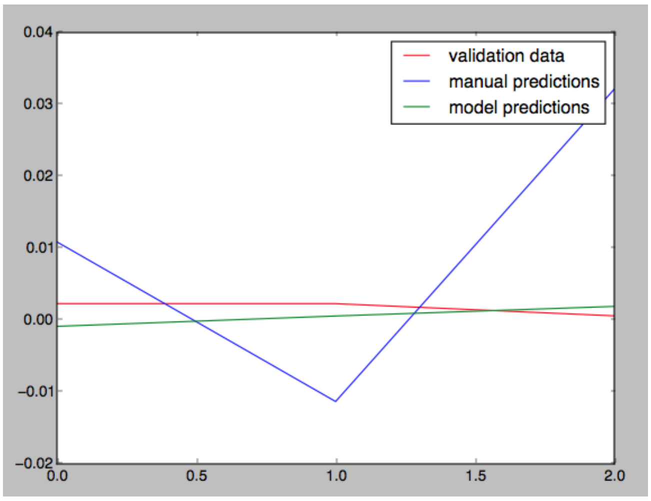

FIGURE 4.10: Validation and prediction curve with first difference data.

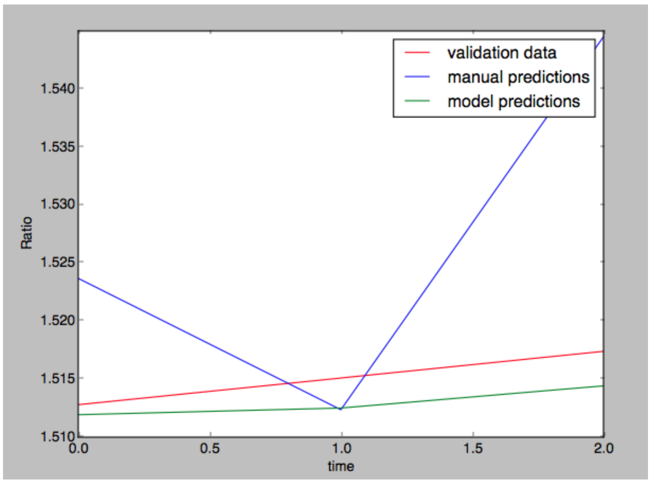

FIGURE 4.11: Validation and prediction curve with original data.

and present the MSE

```
Manual MSE: 0.000417918433359
Model MSE: 4.90002805137e-06
```

FIGURE 4.12: Mean Squared Error(MSE)

The absolute value of the validation data is around -0.00173, the root of manual MSE is around 0.0204, and the root of manual MSE is around 0.0022. We can see the difference between the two MSE is remarkable, the reason is the model hypothesis ϵ_n are i.i.d white noise, we use two residuals $\hat{\epsilon_{n-1}}, \hat{\epsilon_{n-2}}$ generated by model prediction with validation set to get residual $\hat{\epsilon_n}$. It results a remarkable error.

4.2.4 Conclusion

In conclusion, when we fit this forex data to ARIMA model, the performance of both AAPL and forex are not very good, i.e. these model cannot predict future value in three days perfectly. But when comparing forex model to AAPL, these models fitted by forex data have higher accuracy than model fitted by AAPL data.

Notice that, when use data for assets that trade on stock exchange or other financial data to fit model by following the procedures as presented in this report, we cannot predict much more future data points. Otherwise, when we exclude last 100 data as testing data, and use remaining data as training data. We will get such results as below:

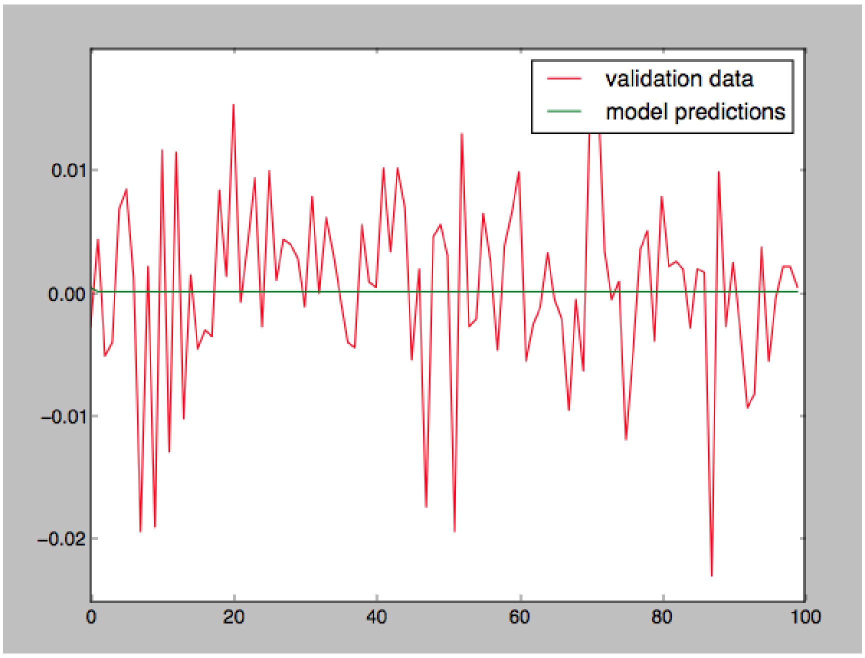

FIGURE 4.13: Predict 100 future data points using model prediction of the first difference forex data

The model totally cannot predict the future data.

But we can modify our method to improve the accuracy, that is, suppose the training data size is n, using this training data to fit model for predicting the first point in testing data, and add the first testing data into training data to fit a new model for predicting the second testing data, and so on. In this way, we will have a better result than original method.

Chapter 5

Margin Call

5.1 Plots

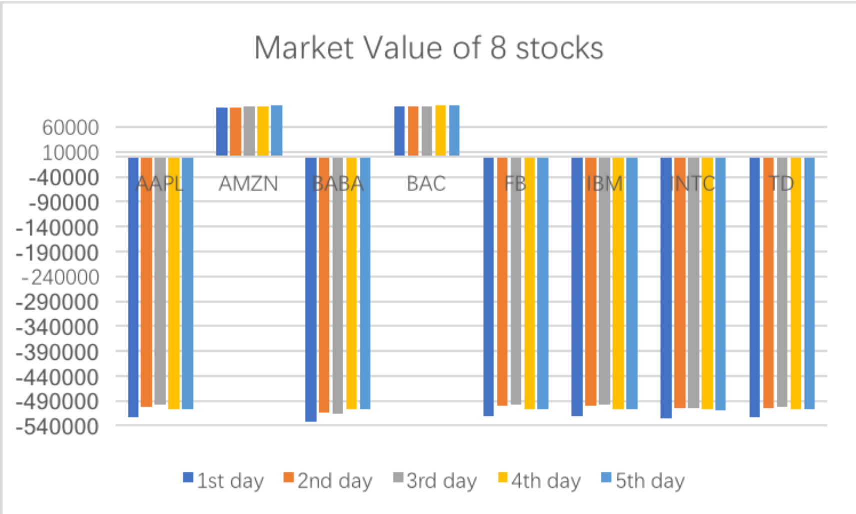

FIGURE 5.1: Market Value of 8 stocks

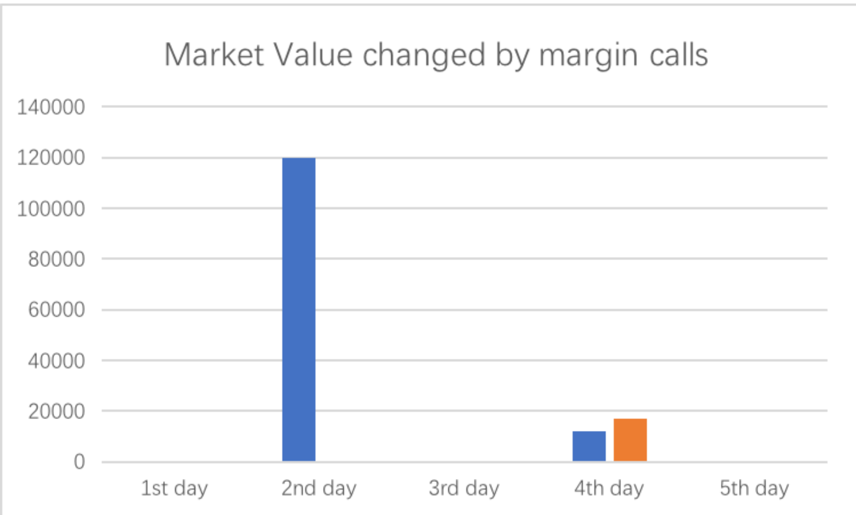

FIGURE 5.2: Market value changed by Margin calls

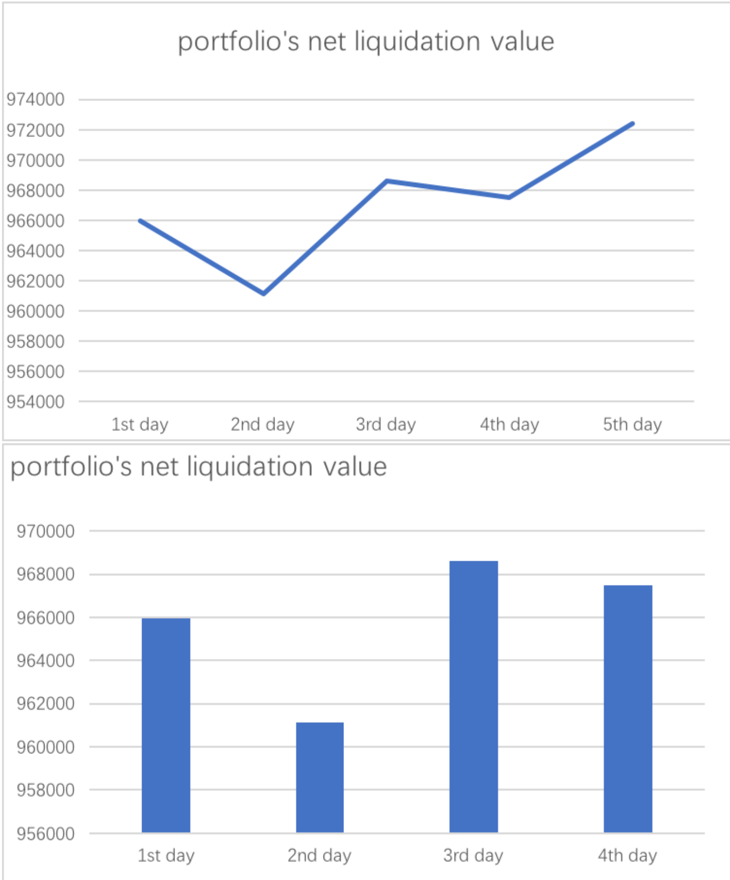

FIGURE 5.4: Portfolio's new liquidation value

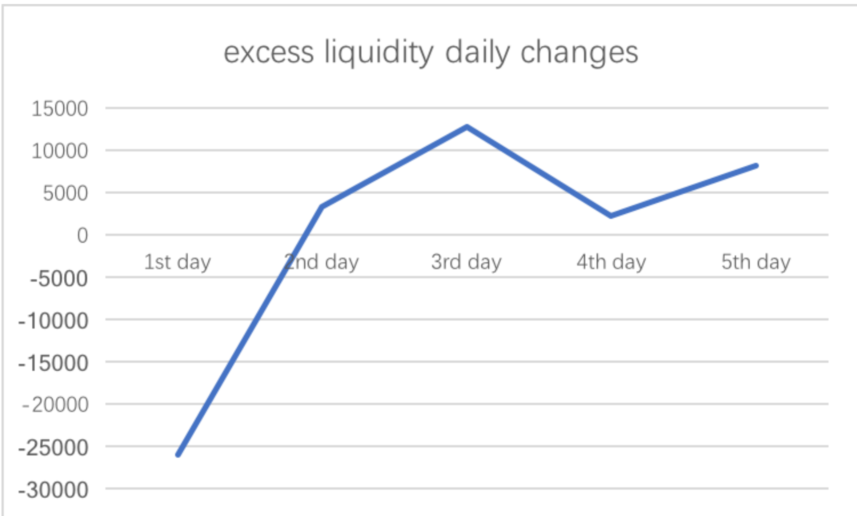

FIGURE 5.3: Excess liquidity daily changes

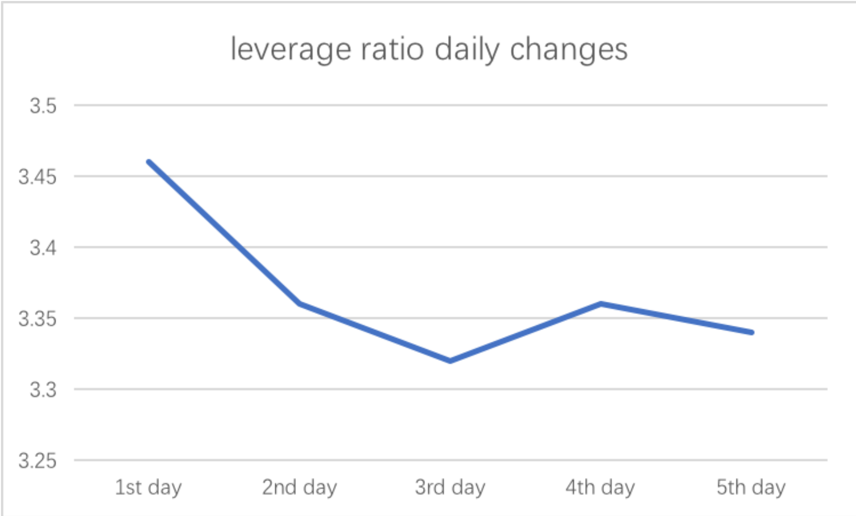

FIGURE 5.5: Leverage ratio daily changes

5.2 Conclusion

Explain the reason why we have a negative value of excess liquidity at end of the first day: Our last observation of the first day is a bit too late and because we are paper accounts so we did not receive the margin call during the day. We did not have enough time to change the excess liquidity by closing the short positions. But we immediately change the excess liquidity to positive the next day.

We did 3 margin calls trades over five days. Because we did not do trades the first day so the market value we changed during the first margin call trades is huge. At the end of the third day in order to do another margin call trade practice, we controlled our excess liquidity around 3000. And then we received two margin calls in the fourth day so we do two trades.

From the plots, our leverage fluctuates and the trend is going down. Maybe because it is more risker doing a lot short sales. And our net liquidity varies a lot daily. Generally our net liquidation value is going up. At first we setted the market value of each short positons equally, if we did not do margin call trades over the days they will vary much from each other.

Bibliography

Ruppert, D. (2004). *Statistics and Finance: An Introduction*. Springer Texts in Statistics. Springer. ISBN: 9780387202709.