Applied Time Series Analysis Final Project
----Analysis and Prediction for 3M stock returns

Written by:

Yuxing Liu, Jiahui Zheng

April 30th, 2018

1. Data description and problem statement

The dataset used in this project contains 755 rows and 2 columns, which are date and the monthly simple returns for 3M stock. This dataset can be obtained from https://github.com/strengthening/AFTS/blob/master/chapter1/text/m-3m4608.txt.

2. Data Analysis

2.1 Data Transformation

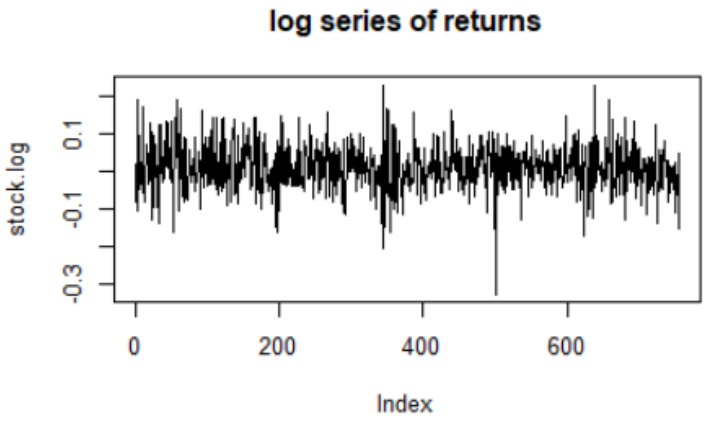

Figure 1: Time series of log returns

The first step was to transform the returns to log returns and then plotted the time series data. It looked like that there was no trend, so differencing was unnecessary.

2.2 Order Determination

Next step was to determine the order according to ACF, PACF and EACF.

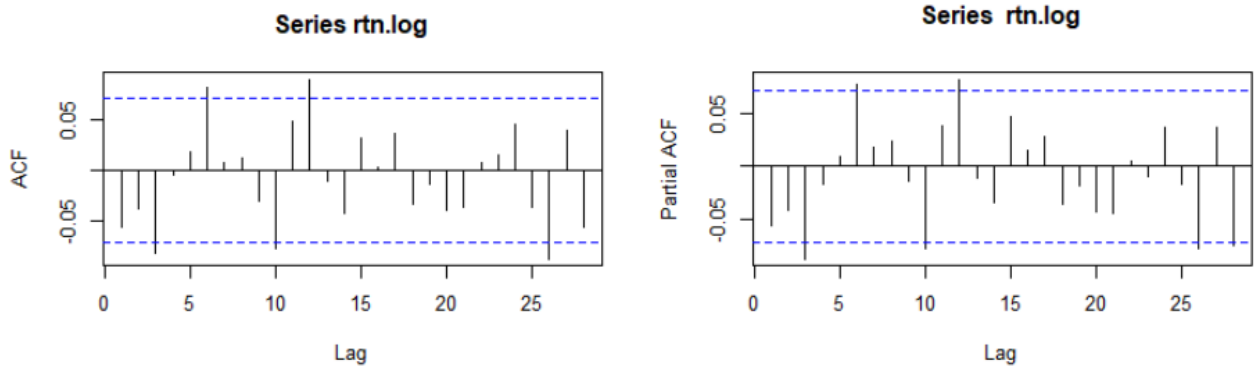

Figure 2: ACF and PACF of log returns

ACF and PACF shown above indicated that the orders should be p=3 and q=3.

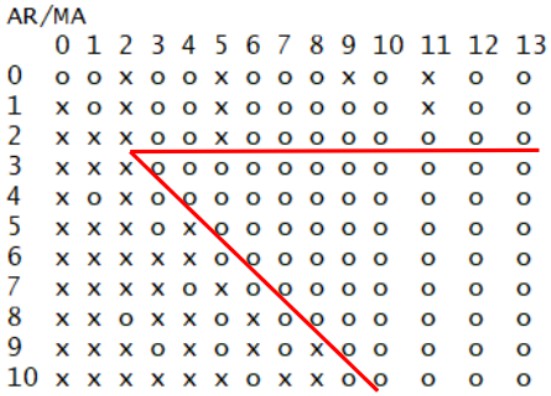

Figure 3: EACF of log returns

According to EACF above, it showed obviously that the model should be ARMA(3, 3). Therefore, the fitted model was ARIMA(3, 0, 3).

2.3 Model Fitting

	Ar1	Ar2	Ar3	Ma1	Ma2	Ma3	Intercept
	0.2701	-0.2628	-0.6866	-0.3179	0.2778	0.6135	0.0103
s.e.	0.3198	0.3208	0.3093	0.3409	0.3464	0.3236	0.0021
σ^2 estimated as 0.003955			Log likelihood=1017.16			AIC=-2020.31	

Table 1: Coefficients output of ARIMA(3, 0, 3)

From Table 1, we could write the ARIMA model as follows:

$$r_t = 0.0103 + 0.2701r_{t-1} - 0.2628r_{t-2} - 0.6866r_{t-3} - 0.3179a_{t-1} + 0.2778a_{t-2} + 0.6135a_{t-3}$$

2.4 ARCH Effects Testing

In order to test that if there was any ARCH effects in the log returns, Ljung–Box statistics with 6 and 12 lags of autocorrelations were applied here. The significance level was 5%.

	Chi-square	df	p-value
lag=6	29.781	6	4.325e-05
lag=12	38.971	12	0.0001064

Table 2: Box-test output

Null hypothesis was that first m lags of a_t^2 were zero. From the results, p-values were both less than 0.05, which suggested that null hypothesis should be rejected. So, there was ARCH effects in the log returns.

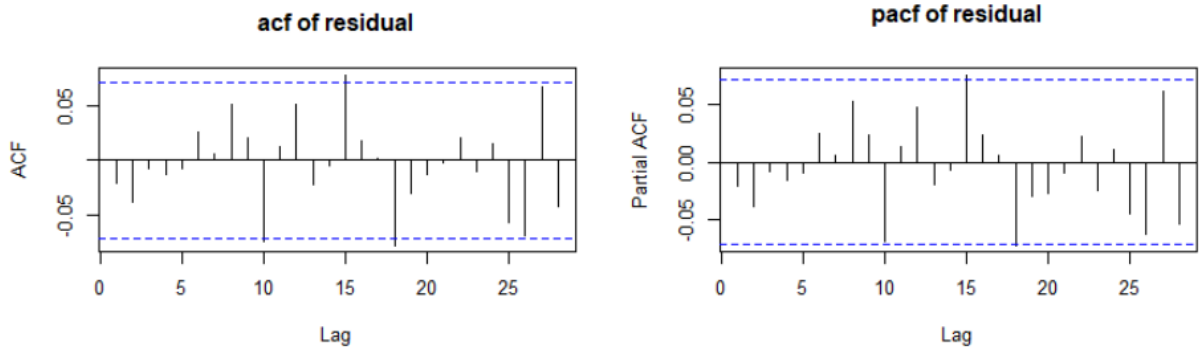

Figure 4: ACF and PACF of residuals

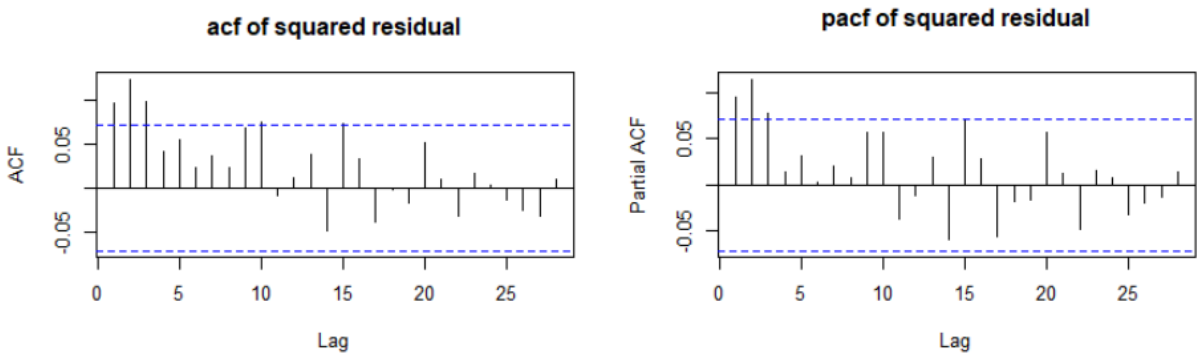

Figure 5: ACF and PACF of squared residuals

Also, the same conclusion could be drawn from ACF and PACF above. If there were no ARCH effects in the residuals, the ACF and PACF should be zero at all lags. However, here in the ACF lags 1,2,3,10,15 were out of the band; and in the PACF first three lags were out of band. Therefore, we could conclude that there was ARCH effects in the log returns.

2.5 ARMA+GARCH Model

ARMA(3, 3) + GARCH(1, 1) was then fitted to the residuals with Gaussian assumption and skewed t distribution assumption. From the output, we found that:

	AIC	BIC	Log likelihood
Gaussian Assumption	-2.731023	-2.663614	1041.961
Skewed t Distribution Assumption	-2.729068	-2.655531	1042.223

Table 3: Partial output from two assumptions

It looked like that the skewed t distribution assumption outperformed normal assumption, so the latter should be applied. And from the output of the skewed t assumption, the fitted ARMA(3, 3) + GARCH(1, 1) model was written as:

$$r_t = 0.0222 - 0.009r_{t-1} - 0.3146r_{t-2} - 0.7114r_{t-3} - 0.0044a_{t-1} + 0.3469a_{t-2} + 0.655a_{t-3}$$

$$a_t = \sigma_t \ , \ \sigma_t^2 = 0.00053 + 0.08452a_{t-1}^2 + 0.78050\sigma_{t-1}^2$$

2.6 Prediction

We then used this fitted model to do a two steps prediction and the results showed that the one-step prediction was 0.0217 and the two-step prediction was 0.0085.

	Mean Forecast	Mean Error	SD	Lower Interval	Upper Interval
1	0.02167569	0.07380258	0.07380258	-0.1225251	0.1715735
2	0.00845601	0.07241294	0.07241214	-0.1331278	0.1555314

Table 4: Two steps prediction

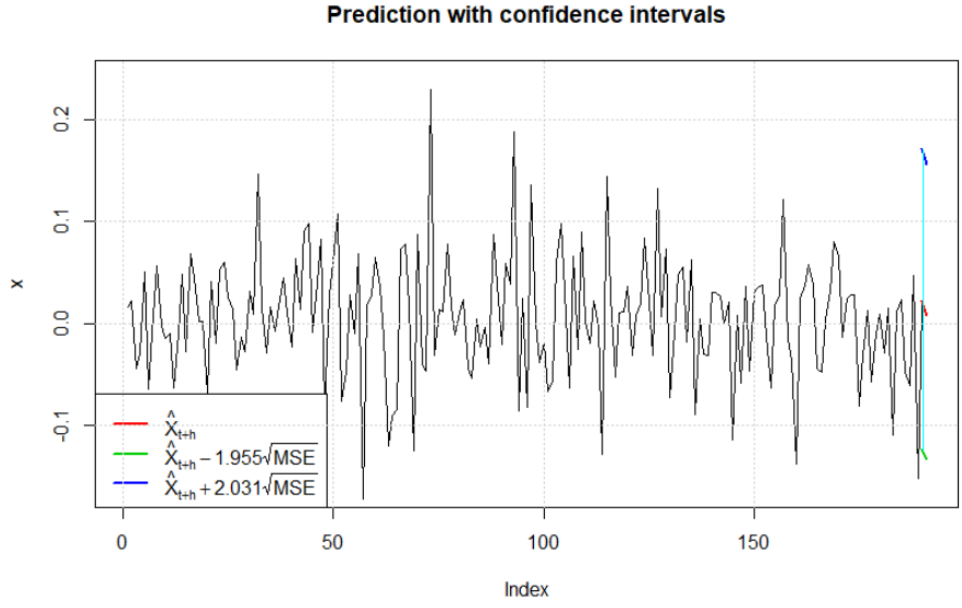

Figure 6: Plot of two steps prediction

3. Conclusion

To analyze this dataset, it was transformed to log returns first. Then according to ACF, PACF and EACF table, the fitted model was ARIMA(3, 0, 3). Two methods were used to test ARCH effects: Box-test and ACF, PACF of squared residuals. The results indicated that there was ARCH effects in log returns. And we compared two models with Gaussian assumption and skewed t distribution assumption, and finally ARMA(3, 3)+GARCH(1, 1) with skewed t distribution assumption was chosen. Then we used this fitted model to do two-step forecast.