Research on Stock Selection and Stock Portfolio Plan and Volatility based on GARCH (1, 1)

Jieru Li

School of Statistics, Shanxi University of Finance and Economics, Taiyuan, Shanxi, 030000

Keywords: Topsis; inertia-reversal theory; EGRMA

Abstract: According to the characteristics of the stock market's return rate sequence such as peaks and thick tails, skewness, volatility clustering, and leverage effect, this paper constructs two forecasting models, GARCH (1, 1) and EGRMA (1,1) to explore and analyze return rates The inherent laws of the sequence use R language to estimate and predict the parameters of the two models established. The results show that the unconditional standard deviation of the EGARCH(1,1) model is 0.03454091 closer to the sample standard deviation of 0.03470566; the AIC(-4.020409) of the EGARCH(1,1) model is compared to the AIC characteristics of the GARCH(1,1) model , Improve forecast accuracy. At the same time, based on the model results, it provides investors with suggestions and strategies for stock selection.

1. Introduction

With the development of my country's capital market in recent years and the expansion of the scale of securities transactions, more and more capital has been devoted to the securities market, and the fluctuations in market transaction prices have also been very intense. Volatility is an essential attribute and characteristic of the securities market. Market volatility has a strong guiding significance for investor risk and return analysis, shareholder equity maximization and effective supervision at the regulatory level. Therefore, it is particularly important to study the regularity and causes of fluctuations in the securities market. It can provide a traceable basis for investors, listed companies and regulators.

This article uses attached stock index data and supplemented complete data to reasonably model the current index volatility and the index volatility in the next year. First of all, we use the daily closing price data of the stock "abc003" and take the logarithm to perform the first-order difference and record it as the rate of return. Then we performed descriptive statistics on the rate of return, and found that the series did not obey the normal distribution, showing the characteristics of peaks and thick tails. After the stationarity test and ARCH test, the results showed that the rate of return series was a stationary series and there was no serial correlation. But there is conditional heteroscedasticity, and asset volatility models can be constructed. Establish GARCH(1,1) and EGRAMA(1,1) models respectively, compare the results of the models, and find that EGRMA(1,1) model can better reflect the fluctuation of the return rate series. Based on the model estimation results and the corresponding stock volatility theory, provide investors with reasonable investment advice and strategies.

2. GARCH model

2.1. Model summary

(1) Volatility and its characteristics

Volatility is the inherent basic attribute and most essential feature of the stock market, and it is also the reason and result of investors trading stocks. Generally speaking, volatility cannot be directly observed, but it also has some characteristics worth studying. These characteristics include: 1) the volatility is clustered, that is, the volatility may be higher in some time periods and lower in other time periods; 2) the volatility changes in a continuous manner, and the volatility Jumping phenomena are rare; 3) Volatility is stable and does not diverge to infinity, but continuously changes over time within a certain range; 4) Volatility reacts differently to good news and bad news, namely There is a leverage effect.

(2)GARCH model

The steps of the GARCH model are:

1) Determine the mean value equation of the sequence and obtain the residual sequence;

2) ARCH effect (conditional heteroscedasticity) test on the residual sequence;

3) Use AIC or BIC to select a specific GARCH model.

(Bollerslev, 1986) proposed an important extension model of the ARCH model, called the GARCH model. For a logarithmic return series r_t ,as $a_t = r_t - \mu_t = r_t - E(r_t|F_{t-1})$ Is its innovation sequence, called$\{a_t\}$ is GARCH(m,s) model,if a_t:

$$a_t = \sigma_t \varepsilon_t, \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

Among them, $\{\varepsilon_t\}$ is an independent and identically distributed white noise column with zero mean unit variance. $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, 0 < \sum_{i=1}^{m} \alpha_i + \sum_{j=1}^{s} \beta_j < 1$,. This last condition is used to ensure that the unconditional variance of a_t that satisfies the model is limited and constant, while the conditional variance σ_t^2 can change with time t.

2.2. Model preparation and data processing

(1) Data selection.

Through the stock index data in the annex and the supplemented data in Question 1, for these 10 sets of stock data, we now choose"abc003"this stock as the key case, the study of the yield form is the yield of the day number.

$$r_t = \ln p_t - \ln p_{t-1}$$

Where the day's closing price $p_t$for the "abc003"” stock is the previous day's p_{t-1}closing price. By estimating and predicting the volatility of the stock, this is a reasonable model of the current and future index fluctuations of 10 groups of stocks, and gives reasonable investment advice and strategies in the context of the models established. Model with R language software, code and corresponding charts can be found in the appendix.

(2) Descriptive statistics.

Table 1 provides a preliminary analysis of the yield data.

index.	r1.	index.	r1.
Nobs.	279.	Sum.	-0.4072.
Minimum.	-0.10565.	SE Mean.	0.002078.
Maximum.	0.09659.	Variance.	0.034706.
Mean.	-0.00146.	Stdev.	-0.04211.
Median.	-0.00202.	Skewness.	1.193698.

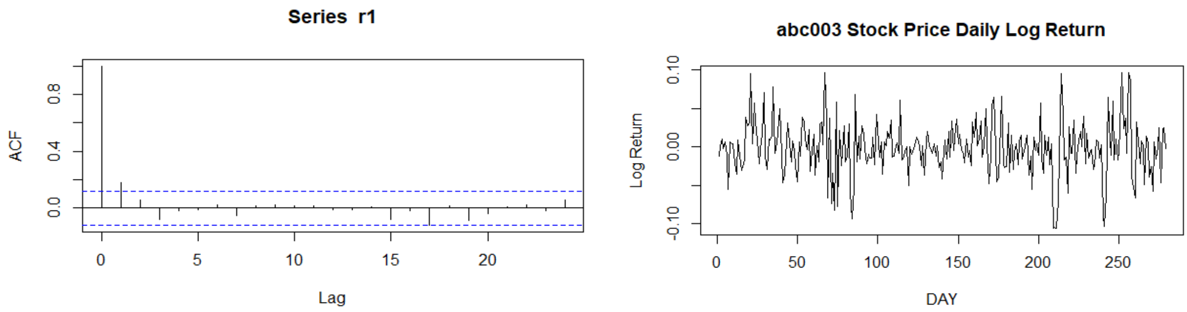
As can be seen from the data inabc001 Table 10, thestock's stock-day yield mean is very small, at -0.001460, which can beconsidered 0. The distribution of the rate of return has a negative bias (-0.042113), so the tail of the distribution drags slightly to the left, indicating that the probability of profit is less than the probability of loss. The excess peak is 1.193698, the maximum is 0.096590, and the minimum is -0.105646.

2.3. The establishment and solution of the model.

The mean equation mentioned in this article can simply be considered as the establishment of an ARMA or ARIMA model, GARCH which is similar to the Process of Building the ARCH model. Compared to the latter, theGARCH model is able to reduce the need for parameters.[9] In this paper, the GARCH (1,1) model is established using the given"abc003"stock day yield as an example. After the smoothness test of the selected data series, the ARMA (1,1) model of the mean equation is established, the volatility model is GARCH (1,1),the model is tested, and the future stock index volatility is predicted.

(1) The establishment of the mean equation.

First import the data into Rstudioto test the steadyness of the convection yield (calculated at the daily closing price) of the"abc003".Use the sequence timing graph and the ACF function in two ways to judge. The results show that in Figure 5a), therange of the day-to-day yield of abc0013 is between the range of the number of days , and its self-correlation function fluctuates in the confidence band range after the first period, and it can be assumed that the day-by-day yield is a weak and stable sequence.



a) abc003 logilate yield time series   b) abc003 log yieldACF function diagram.

Figure 1 Sequence stability test.

Another condition to be met by the model, the sequence is not related before or after, using the Box. Test test to test the self-correlation of the date yield of the abc003 stock, the p p-value of the test is 0.4148, gt;0.05,so the sequence is considered to exist before and after. So we'll build a model to eliminate this linear dependency.

Using the EACF method to establish the ARMA model, after determining the order, parameter estimation, model testing and other steps, we determine the ARMA (1,1) model. As can be seen from Figure 8, we established the end of the model by the EACF method, and then we estimated the parameters of the model, the estimate is shown in the schedule, and we examined the Ljung-Box statistics of the model residuals, p.093211,0.1, so we cannot reject the original assumption at the level of significance of 0.1, i.e. we think the ARMA (1,1) model is sufficient.

Using the EACF method to establish the ARMA model, after determining the order, parameter estimation, model testing and other steps, we have determined the ARMA (1,1) model [10]. As can be seen from Figure 6, we establishthed the end of the model by the EACF method, and then we estimated the parameters of the model, the estimate is shown in the schedule, and we examined the Ljung-Box statistics of the model residuals, the results are shown in Figure 6b),as can be seen from the figure,p.0.993211,0.1, so significant in 0.1.

The assumption that we believe the ARMA (1,1) model established is sufficientatthe sexual level.

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12
0	x	o	o	o	o	o	o	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o
3	x	o	x	o	o	o	o	o	o	o	o	o	o
4	x	o	o	x	o	o	o	o	o	o	o	o	o
5	x	o	o	x	o	o	o	o	o	o	o	o	o
6	x	o	x	o	o	x	o	o	o	o	o	o	o
7	x	x	x	x	x	x	o	o	o	o	o	o	o
8	x	x	x	x	x	o	o	o	o	o	o	o	o
9	x	x	o	x	x	o	o	x	o	o	o	o	o
10	x	x	o	x	o	o	x	o	o	x	o	o	o
11	x	x	x	o	o	o	x	x	o	x	o	o	o
12	x	x	o	o	o	o	x	o	o	x	o	o	o

```
> Box.test(zdr1$residuals,lag=12,type = 'Ljung')

        Box-Ljung test

data:  zdr1$residuals
X-squared = 4.814, df = 12, p-value = 0.9639
```

a) EACF method ARMA model order establish   b) ARMA (1,,1)model test results.

Figure 2 The establishment of the ARMA (1,,1)model.

Finally, we will build the ARMA (1,,1) model as the mean equation to be established. The equation is:

$$(1-0.2485B)(x_t+0.001968)=(1-0.0704B)\alpha_t$$

(2) ARCH effect test.

Based on the models we need to build, large past squared "disturbances" can create a large $\alpha_t$information condition heterodox, and thus a $|\alpha_t|$tendency to have a greater value. This means that under the ARCH framework, large "disturbances" tend to follow another large "disturbance", similar to what we call volatility aggregation. The so-called ARCH model test, that is, the correlation test of the conditional heterovariance sequence.

There are two main test methods: the Ljung-Box test and Engle's Lagrangian multiplier test, commonly referred to as the F-test. There is no big difference between the two methods in terms of results, so we choose the former to test the above-mentioned mean model.

Using the R language for testing, it is concluded that the p-value of the test p is less than 0.05,and the original assumption cannot be rejected, i.e. the sequence is not self-correlation through the ARCH test.

(3) Theestablishment of the GARCH model.

The GARCH (1,1) model is established with a normal condition distribution, as shown in Figure 7. The results of the white noise test of the standardized residuals and their squares have passed, and the normality test of the condition distribution is still passable. The model can be written as:

$$\begin{cases} \hat{r}_t = -0.00139\hat{r}_{t-1}+\varepsilon_t, \\ \sigma_t^2 = 0.00026+0.17454\varepsilon_{t-1}^2+0.59829\sigma_{t-1}^2 \end{cases}$$

σ_t^2The reliance on the past comes mainly from $\beta_{1=}$0.598.

2.4. The test of the model.

For a reasonably established GARCH model, standardized residuals: , which constitute a sequence of random variables that are independently and distributed. Therefore, we can examine the adequacy of the fitted GARCH model by examining the sequence. In Figures 8 and 9,we made a chart of the chart of thestock's fitted volatility, which shows that volatility peaks around 205-215 days, with the highest volatility. The green horizontal line in the figure is the sample standard deviation, and we can calculate its value of 0.03470566. At the same time, we learned that the model-based variance yield unconditional standard deviation is 0.03397701, which is very similar to the sample standard deviation, you can initially judge the adequacy of the model.

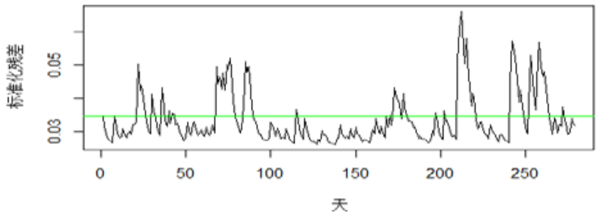


Figure 3 abc003 stock build simulation combined volatility.

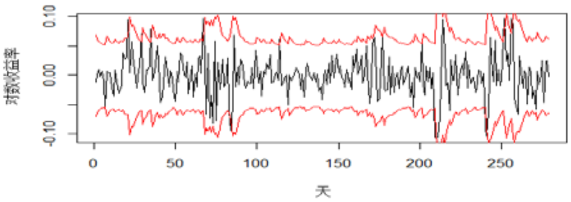


Figure 4 Stock logtologic yield snout 95% forecasting.

3. EGARCH model

3.1. A summary of the model

Nelson's index GARCH model, proposed in 1991, allows positive and negative returns on equity to have an asymmetrical effect on volatility. Consider the following transformation:

$$g(\varepsilon_t) = \theta\varepsilon_t + \gamma[|\varepsilon_t| - E|\varepsilon_t|]$$

Among them and the real $\theta\gamma$constant, $\{\varepsilon_t\}and\{|\varepsilon_t|$-E$|\varepsilon_t|\}$are the zero mean independent and distributed white noise, the distribution is continuous distribution. Easy to see. $Eg(\varepsilon_t) = 0$

$$g(\varepsilon_t) = \theta\varepsilon_t + \gamma[|\varepsilon_t| - E|\varepsilon_t|]$$

Take EGARCH (1, 1) as an example. The model is:.

$$a_t = \sigma_t\varepsilon_t, (1 - \alpha B)\, ln\sigma_t^2 = (1 - \alpha)\alpha_0 + g(\varepsilon_{t-1})$$

where $\varepsilon_t \sim N(0,1), \alpha = \alpha_1$. Such a model makes conditional variance dependent on perturated symbols in a nonlinear way, which is derived from segmentation functions and exponential transformations. The volatility equation can be written as $\sigma_t^2 a_{t-1}$: .

$$\sigma_t^2 = \sigma_{t-1}^2 e^{\alpha^*} \begin{cases} \exp[(\gamma+\theta)\dfrac{a_{t-1}}{\sigma_{t-1}}], & a_{t-1} \geq 0 \\ \exp[(\gamma+\theta)\dfrac{|a_{t-1}|}{\sigma_{t-1}}, & a_{t-1} < 0 \end{cases}$$

The EGARCH model shows that the rate of return reacts to positive and negative information [11] inadesic. In the model, radon is the key parameter for measuring information asymmetry. If the positive interference causes the condition variance change is greater than the negative interference, if the negative interference causes the condition variance change is greater than the positive interference, and the positive interference, indicates that there is no information asymmetry effect.

3.2. Establish and solve.

(1) Data processing.

For modeling the stock data index fluctuations for the coming year, first we use the AMRM (1,1) model involved in the GARCH model in Model 1 的 to forecast the daily yield of the "abc003" stock in the coming year;

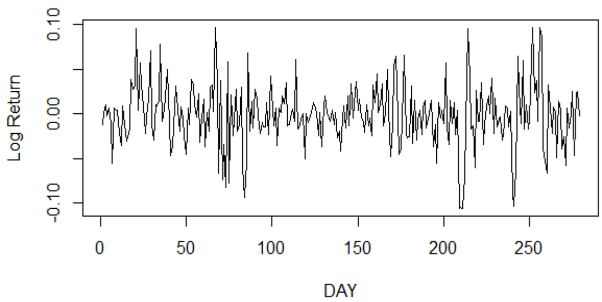


Figure 5 predicts a time series of logbirth slower yields for the next year.

(2) Model-building.

Since the sequence is predicted by the model's a dating rate of one by ARMA (1,1), we directly skip the sequence stability test, self-correlation and AECH effect test to directly build the model. Consider IBM stock monthly variance yield modeling, 280 observations from January 8, 2019 to March 26, 20020. To build an EGARCH model, taking into account some of the conclusions obtained in Model 1, to establish the following model form, the model is:

$$r_t = \phi_0 + a_t, ln\sigma_t^2 = \alpha_0 + \alpha_1(|\varepsilon_{t-1}| + \gamma_1\varepsilon_{t-1}) + \beta_1\, ln\sigma_{t-1}^2$$

Using the R language, the newly generated data series are basically processed, such as the outlier test. The EGARCH (1,1) model is then established using its unique fGrach package, and the model is tested and predicted, similar to the test prediction in Model 1.

Estimation results of EGARCH(1,1) model:

	phi0	alpha0	alpha1	gamma1	beta1
Estimate	0.006722463	-0.5980087	0.21763575	-0.4237355	0.92021763
SE	0.002876230	0.2347611	0.05916033	0.1681051	0.03883014
t-ratio	2.337248075	-2.5473075	3.67874502	-2.5206581	23.69854001

Figure 6 Screenshot of the results of the eGARCH (1,1) model section.

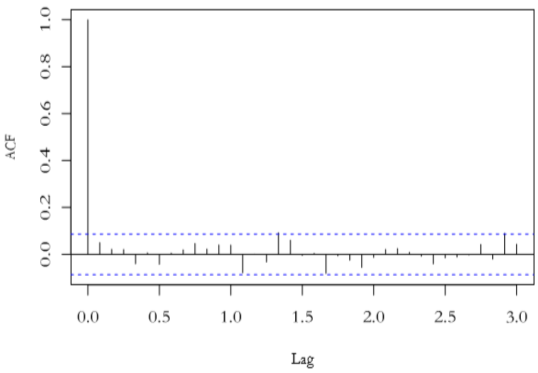
Based on the results obtained in Figure 9, you can determine the estimated model:

$$r_t = 0.0067 + a_t, a_t = \sigma_t \varepsilon_t, \varepsilon_t \sim N(0,1)$$

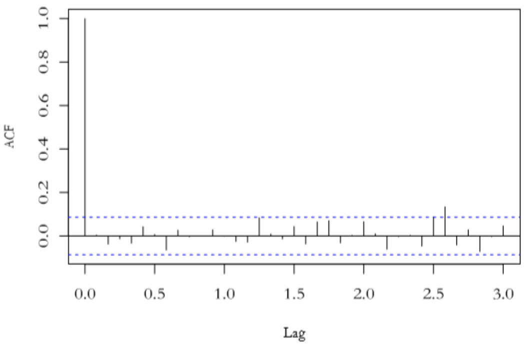
$$ln\sigma_t^2 = -\,0.5980 + 0.2176(|\varepsilon_t| - 0.4237\varepsilon_t) + 0.9202\, ln\sigma_{t-1}^2$$

3.3. Model test

Similar to a model test method, first we look at the ACF graph of standardized residuals, while the"Ljung-Box"white noise test for standardized residual, standardized residualsquares. The result is Figure 13. As can be seen from the figure, the ACF chart of both standardized residuals and standardized residual squares shows that the model is well established, and thatthe P-value of theBox test is greater than 0.05,which alsoillustrates the adequacy of the model.



a) abc003 stock yield model standardized residual ACF



b) abc003 stock yield model standardizes the residual square of the ACF.

Figure 7 EGARCH (1, 1) model inspection.

4. Model results and investment advice.

4.1. The model result.

In Model 1, the GARCH (1,1) model that fits the number of yields on the closing day of the"abc003"” stock has the following characteristics: : .

(1) The largebeta coefficient in the model and the passing of the significance test show that exponential fluctuations have a "long-term memory", ” “ i.e. past price fluctuations are related to the size of their infinite long-term price fluctuations.

(2) The Close to $\alpha_1 + \beta_1 1$ in the GARCH equation indicates that the conditional variance

function has unit root and monocompleteity, that is, conditional variance fluctuations have a persistent memory, indicating that the reaction of the securities market to external shocks is decreasing at a relatively slow rate, and it is difficult to eliminate the large fluctuations in the stock market in a short period of time.

(3) In the GARCH $\alpha_1 + \beta_1 < 1$equation, it is shown that the sequence of variances of the yield conditions is stable and the model is predictable.

The EGARCH (1,1) model shows that the effect of the dating yield information is asymmetric and has a significant leverage effect, i.e. the negative impact of the stock market is more volatile than the positive impact. The information impact graph can also reflect the same conclusion, which is steep when the information shock is less than 0, that is, when it represents a negative shock, and is relatively flat at positive shock, which indicates that the negative shock makes the volatility change more.

In terms of the AIC values of the two evaluation models, the EGARCH (1,1) model has an AIC of -4.020409,, which is more advantageous than the AIC value of the GARCH (1,1) model -0.3975137, and the forecast results are better. At the same time, we can also look at the model calculated the variance yield unconditional standard deviation to compare. The standard deviation of the sample we calculated was 0.03470566, while the unconditional standard deviation calculated using the GARCH (1,1) and EGARCH (1,1) models were 0.033977010.03397701,0.03454091. From this result, it is also demonstrated that EGARCH (1,1) forecasts are better.

4.2. Investment advice and strategies

For the two models of stock volatility established in question three, it should be emphasized that it is only applicable to the volatility of the "abc003" stock onsp selected herein, but also to the other 9 stock data, the GARCH (1,1), EGARCH (1,1) model can also be established to estimate and predict the volatility of their respective stocks. It has proved to be well-applied. Based on the estimates of the two models, and combined with the actual situation of the current stock market, the following investment suggestions and strategies are put forward after reference to a large number of literatures.

1) When selecting a stock, pay attention to its volatility, depending on the size of the volatility of the stock to decide whether to invest.

2) To observe the band range of the stock, it is often that everyone confuses the volatility and band ranges.

3) In practice, we will find that volatility has a non-pairforming characteristic, i.e. positive "disturbance" and negative "disturbance" have different effects on volatility, which is what we often call leverage effect.

References

[1] Wang Jingliang. Research review of inertia and reversal effects[J]. Shang, 2014(26): 159.

[2] Li Haoyue. Empirical research on stock pricing factor model based on reversal effect [D]. Soochow University, 2016.

[3] Liu Yaguang. The selection of the best investment portfolio in China's stock market: based on the mean-variance model [J]. Postgraduate Journal of Central China Normal University, 2015, 22(03): 129-134.

[4] Xia Jing. Stock investment portfolio based on linear programming [J]. China Management Information Technology, 2019, 22(21): 142-143.

[5] Lian Xiangbin. Research on the relationship between market volatility, skewness and kurtosis and stock returns [D]. Dongbei University of Finance and Economics, 2016.

[6] Chen C T. Extensions of the TOPSIS for group decision-making under fuzzy environment[J]. Fuzzy Sets & Systems, 2000, 114