Problem Chosen	2022	Team Control Number
C	MCM/ICM Summary Sheet	2219846

I am the Wolf of Wall Street

Summary

In financial markets, market traders are always looking for better ways to invest in order to reach higher profits. Our task is to develop a mathematical model that helps investors to find the best trading strategy. In this article, we try to find **the optimal trading strategy** in a market consisting of Bitcoin, Gold and USD.

1. **Optimal Risk Asset Portfolio Model:** Investors are always looking for greater returns with less risk.We modify the **mean-variance model** by adding a **time** factor and a **cost** factor. This results in $\overline{SharpeRatio_T}$, which measures the return-to-risk ratio over a future time period when transaction costs are present. By finding the maximum $\overline{SharpeRatio_T}$, we can obtain the optimal investment decision.

2. **SVR+GARCH model:** In order to obtain the optimal investment strategy for the future time period. We need to forecast the future **return**. We first build a **multiple regression model** by performing **PCA** analysis on a variety of financial indicators and extracting the main components, however, we find significant differences between the predicted and true values of the model. We speculate that this is due to complex nonlinear factors. So we tried to build the **SVR model**. By performing residual analysis on the SVR model, we found the**ARCH** effect of the financial series. So we fit the residuals of the SVR model by **GARCH model** to build the SVR+GARCH model and successfully forecast the future return curve.

3. **Genetic Algorithm:** Next, we split the time series according to the ratio of the training and test sets of the SVR+GARCH model (7:3). Within each time period, the $\overline{SharpeRatio_T}$ is used as the **objective function** to find the optimal strategy using **Genetic Algorithm.**. By solving, the investor will have 111,383.15 USD on September 10, 2021!

4. **Sensitivity Analysis:** Finally, we performed a sensitivity analysis of the model and we found that our strategy is extremely sensitive to the commission price of bitcoin, but not to the commission price of gold.By analyzing the price curves and trading volatility, we found that they are caused by the huge difference in price volatility between bitcoin and gold and the increased difficulty of withdrawing funds due to higher trading costs, respectively. At the end of the paper, we also wrote a memo to market traders conveying our strategy, model and results.

Keywords: mean-variance model, PCA, GARCH, Genetic Algorithm, SVR

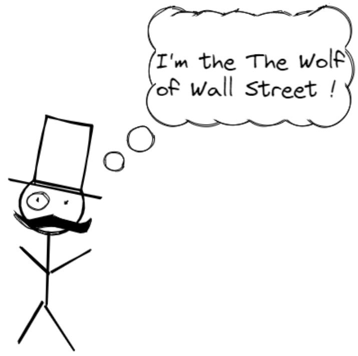
Contents

1 Introduction

1.1 Problem Background

Finance has been a constant topic since the dawn of human society. When people 200 years ago saw the financial markets of today, they would have been shocked. Because of the sheer volume of data!

Where once there could be just a blackboard in a Wall Street exchange, today, with the rise of the big data wave, quantitative trading has become a hot topic. We try to analyse data on Bitcoin and gold to establish the best strategy.



1.2 Restatement of the Problem

We have to develop a model that uses only the past stream of daily prices to date to determine each day if the trader should buy, hold, or sell their assets in their portfolio. We are asked to start with 1000 dollars on 9-11-2016 with the five-year trading period, from 9-11-2016 to 9-10-2021. On each trading day, the trader will have a portfolio consisting of cash, gold, and bitcoin [C, G, B] in U.S. dollars, troy ounces, and bitcoins, respectively. The commission for each transaction costs $\alpha\%$ of the amount traded. Assume the fare of gold is 1% and the fare of bitcoin is 2%. We should complete that:

- Develop a model that gives the best daily trading strategy based on the price data of the day only. And use your model and strategy to predict how much the initial investment of $1,000 will be worth on 9-10-2021.
- Present evidence that your model provides the best strategy.
- Analyze whether the strategy is sensitive to transaction costs.

1.3 Our work

1. We begin by setting out the black box model of quantitative trading to give direction to our modelling. It requires us to consider the three elements of return, risk and cost and then make a best strategy.

2. We then begin by considering the relationship between return and risk. Referring to the mean-variance model, we add a cost factor and a time factor to it. This allows us to use it to calculate the Sharpe ratio over a time period and thus address how we should measure risk and return to make optimal decisions.Therefore, the process of finding the optimal decision is the process of finding the maximum Sharpe ratio.

3. We then attempted to forecast returns versus risk.We first try to extract the factors related to returns and run a multiple regression on the return series. But the regression model did not work well. We believe this is due to the complex non-linearities in financial markets. So we tried to use the SVR model. Happily, the SVR model worked brilliantly. But there are still unexplained residuals between the SVR model's predicted results and the true values. We found that this could be due to ARCH effects. This is because the returns of Bitcoin and gold show a clear spike and thick tail.

We then ran the Mcleod-Li test on him and successfully verified our conjecture. We then fitted the residuals of the SVR model through the GARCH model and successfully obtained residuals that showed a normal distribution. We then completed our forecasting work with the SVR+GARCH model.

4. We then try to find our optimal decision. From the previous analysis we know that the process of finding the optimal decision is the process of finding the maximum Sharpe ratio. Also, because of our hypothesis: investors will look for the maximum Sharpe ratio over a time period where the forecast is good enough. So we split the time series by the ratio of the training and test sets of the SVR+GARCH model. Within each segmented time period, we find the optimal decision by Genetic Algorithm. Thus we get the best strategy.

5. Finally we perform a sensitivity analysis of our model.we found that our strategy is extremely sensitive to the commission price of bitcoin, but not to the commission price of gold.By analyzing the price curves and trading volatility, we found that they are caused by the huge difference in price volatility between bitcoin and gold and the increased difficulty of withdrawing funds due to higher trading costs, respectively.

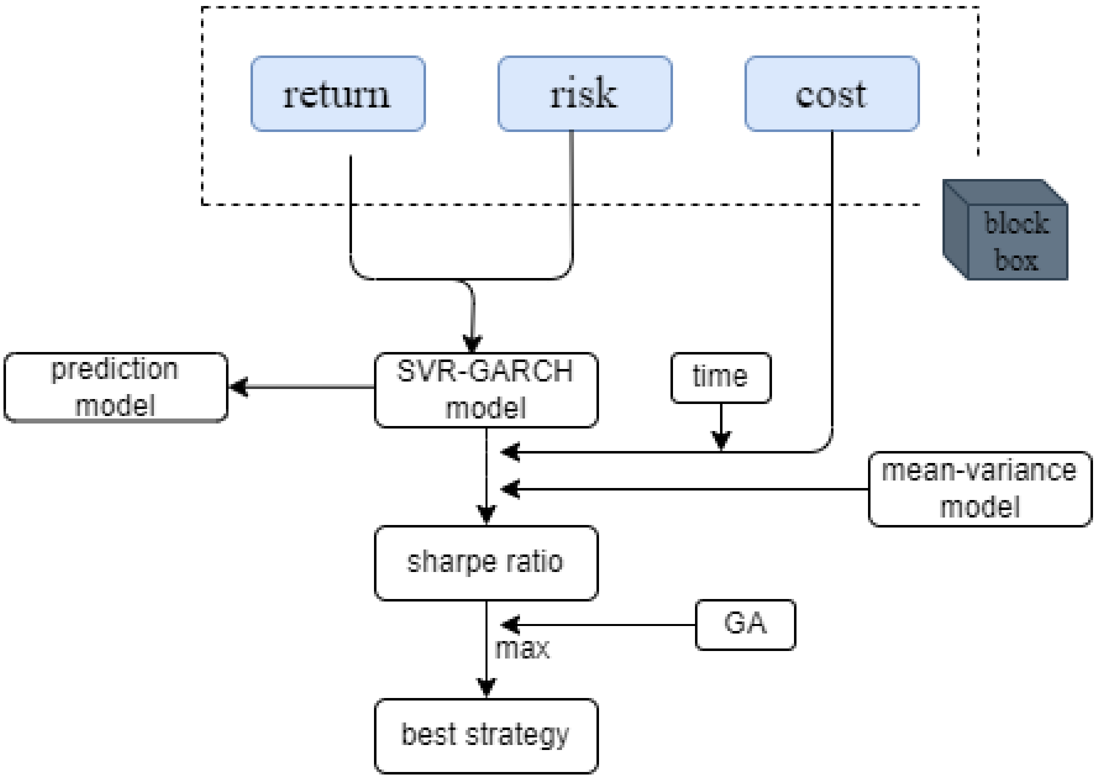


Figure 1: Flow Chart

2 Assumptions and Notations

2.1 Assumptions

Considering that practical problems always contain many complex factors, first of all, we need to make reasonable assumptions to simplify the model, and each hypothesis is closely followed by its corresponding explanation:

- **Investors consider each investment choice based on a forecast of future returns.**
  Explanation: It is a natural assumption that investors must buy assets when there is an excess return in the future cycle. When there is a risk of loss in the future cycle, investors must sell assets.
- **Investors measure the risk of an asset based on the variance of the asset's future returns.**
  Explanation: Judgments about risk are often made through variance, because the size of the variance affects the decisions that investors complete based on return expectations
- **Investors make decisions when the variance of the forecast results in the future cycle is more desirable.**
  Explanation: By fitting past data and predicting future data, investors will make decisions in the maximum period with better prediction results to achieve the optimal decision in the maximum period
- **For a given level of risk, investors expect the maximum return; correspondingly, for a given level of return, investors expect the minimum risk.**
  Explanation: It is also a natural assumption that, as an investor, one must carry an aversion to risk

2.2 Notations

Table 1: **Notations**

Symbol	Defifinition
R	Rate of return on risky assets
σ	Variance of risky asset prices
C	The value of cash in the hands of investors
G	The value of gold in the hands of investors
B	The value of bitcoin in the hands of investors
ρ	The correlation coefficient between bitcoin and gold
$cost$	Costs incurred by the transaction

3 Model Preparation

3.1 Definition of financial indicators

We have used a number of financial indicators in the modelling process. To make it easier to illustrate the model, we will not explain them in detail in subsequent articles.

3.2 Data Processing

There is missing data in the GOLD dataset, and we process it in two ways.

Table 2: **Indicators**

2MA	Average of closing prices within two days
10MA	Average of closing prices within ten days
30MA	Average of closing prices within Thirty days
bollup	Mid-track line + twice the standard deviation
bolldown	Mid-track line - twice the standard deviation
RSI	Percentage of volatility generated by price increases
ADX	Average direction of motion index
PPO	Price Oscillator
ROC10	Price ratio from 10 days ago

- First, in the prediction model of the future we fill the missing values of gold by **linear interpolation** to facilitate the completion of our prediction model. We put the t who's boolean value is 0 into T_0
- We then add constraints to our planning model by adding a set of **boolean values** to the dataset that represent whether the gold market is open or not.

4 Optimal Risk Asset Portfolio Model

4.1 Mean-variance model

We first introduce the mean-variance model proposed by Harry M. Markowitz in 1952.

In this model, there are one risk-free asset and two types of risky assets. We can call them: f, A and B.

Harry M. Markowitz note that the return and expectation of risky asset should be a curve as shown in the Fig.2, and r_f represents the rate of return on the risk-free asset. The different points on the curve represent the relationship between return and risk of different portfolios.

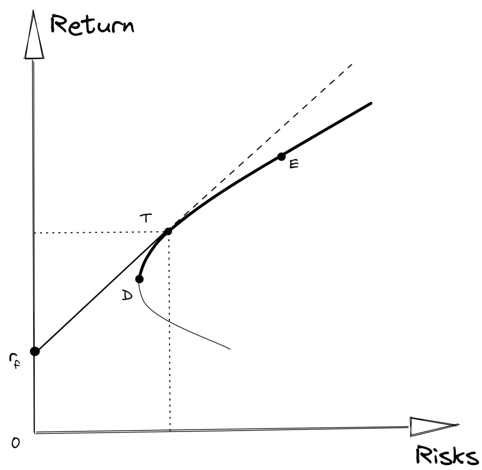


Figure 2: Mean-Var model

Obviously, we should focus our attention on the bolded part of this curve, because it represents a higher return for the same risk.

The straight line in the graph represents a portfolio relationship between risky and risk-free assets

To measure the mysteries on this graph in the future, let's first define the return and risk of asset portfolio:

$$E\left(R_N\right) = wE\left(R_1\right) + (1-w)\,E\left(R_2\right) \tag{1}$$

$$\sigma_N = \sqrt{w^2{\sigma_1}^2 + (1-w)^2{\sigma_2}^2 + 2\rho\sigma_1\sigma_2 w(1-w)} \tag{2}$$

where Eq.(1) represents the portfolio return of the assets, w is the weight, We let it represent the weight of A. Eq.(2) represents the portfolio risk of the assets, and like the risk of the individual assets, it is defined by the variance of the portfolio. Where ρ is their correlation coefficient.

And the measure of return versus risk is the SharpeRatio, which represents the magnitude of return per unit of risk:

$$SharpeRatio = \frac{E(R_N) - R_f}{\sigma_N} \tag{3}$$

Obviously, when the Sharpe ratio reaches its maximum, that is, when the line in the figure is tangent to the curve at point T. Mathematically, we can calculate the weight w represented by the point T.

$$w = \frac{\left(E\left(R_1\right) - R_f\right)\sigma_2{}^2 - \left(E\left(R_2\right) - R_f\right)\rho\sigma_1\sigma_2}{\left(E\left(R_1\right) - R_f\right)\sigma_2{}^2 + \left(E\left(R_2\right) - R_f\right)\sigma_1{}^2 - \left[E\left(R_1\right) - R_f + E\left(R_2\right) - R_f\right]\rho\sigma_1\sigma_2} \tag{4}$$

From Eq.(4) , we get the w which is the weight under the optimal combination.

4.2 Mean-variance model with time and costs

Of course, the mean-variance model is only the optimal combination sought without transaction costs and with a single transaction.

But we can draw inspiration from it. We propose Mean-variance model with time and costs based on our assumptions. Fig.3 demonstrates the process.

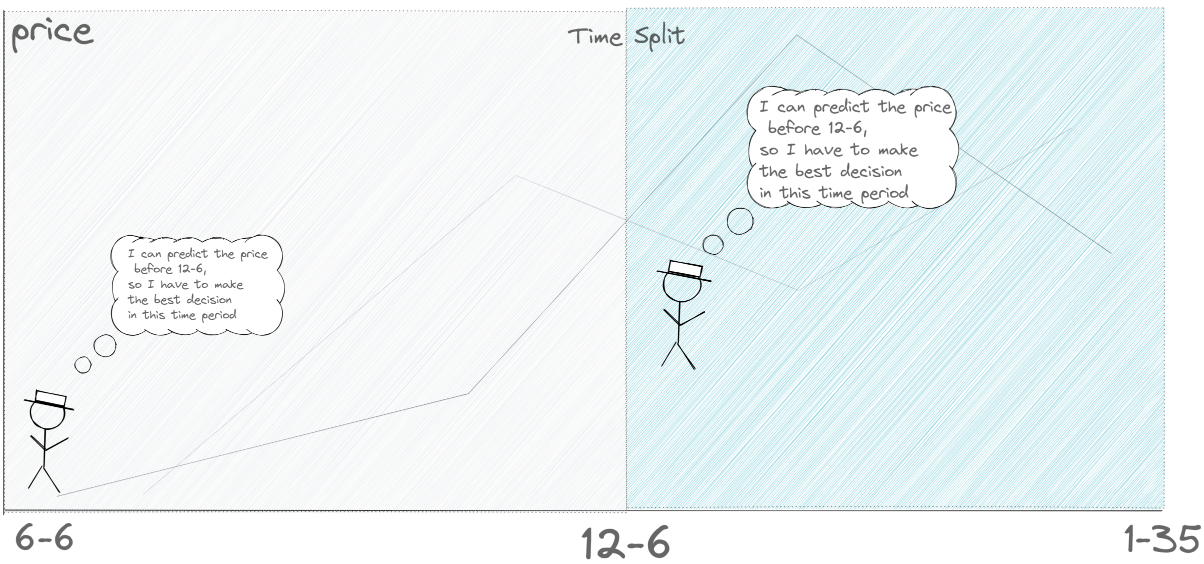


Figure 3: Decision-making process

This is because in the model assumptions we talk about: when the forecast model can achieve better results in a future time period, investors must seek the optimal portfolio in that time period based on the forecast. Therefore, all we have to do is to allocate assets in this time period to achieve the optimal Sharpe ratio.

We first modify Eq.(1) and Eq.(2) by adding a time factor and a cost factor:

$$E\left(R_{N_t}\right) = w_t E\left(R_{1_t}\right) + \left(1 - w_t\right) E\left(R_{2_t}\right) - cost \tag{5}$$

$$\sigma_{N_t} = \sqrt{{w_t}^2{\sigma_{1_t}}^2 + (1-w_t)^2{\sigma_{2_t}}^2 + 2\rho\sigma_{1_t}\sigma_{2_t}w_t(1-w_t)} \tag{6}$$

Eq.(5) represents the rate of return on t. And cost is given by the transaction cost:

$$cost = 1\% \left| w_t - w_{t-1} \right| A_t + 2\% \left| 1 - w_t - 1 + w_{t-1} \right| B_t$$

A_t and B_t represent the value of asset A and asset B owned by the investor at time t. Then we have the Sharpe ratio corresponding to time t:

$$SharpeRatio_t = \frac{E(R_{N_t}) - R_f}{\sigma_{T_t}}$$

$$\overline{SharpeRatio_T} = mean\left[SharpeRatio_t\right], t \in T \tag{7}$$

The optimal strategy we seek is the one that corresponds to when the mean value of the Sharpe ratio during time periods T is maximized.

Of course, this is only the optimal combination between two risky assets. It represents our optimal choice of risky assets. But we should note that the combination between risk-free and risky assets is linear, which means that we cannot hold any risk-free assets when we want to get the maximum risky return, because they are both linearly combined. Generally speaking, investors are happy to adjust the allocation between risk-free and risky assets according to the Sharpe ratio. We assume that investors will put all their money into risky markets when the Sharpe ratio is positive and pull back 75% of their money when the Sharpe ratio is negative.

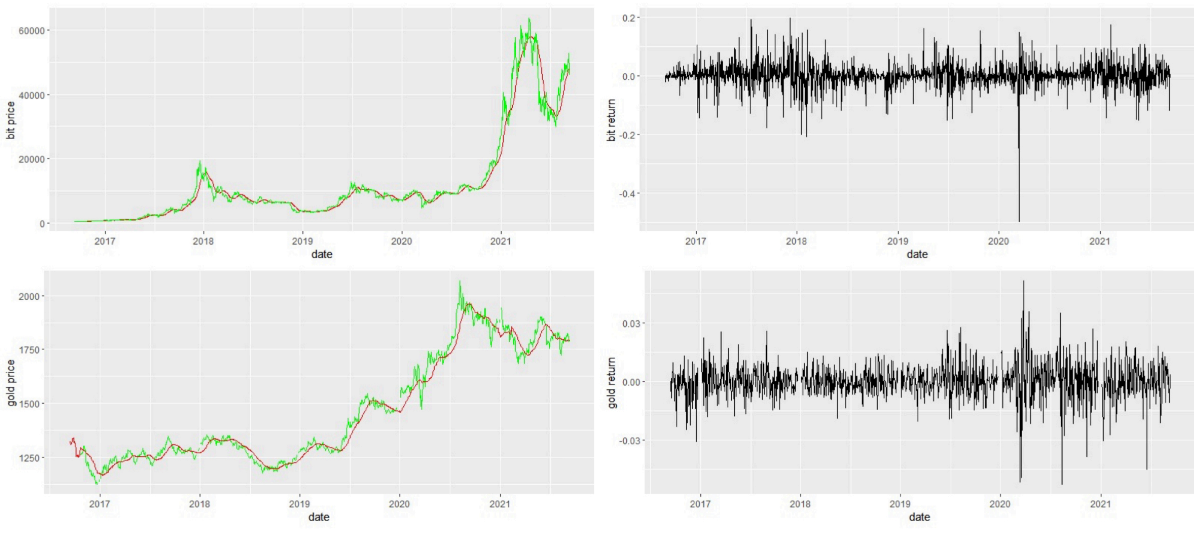


Figure 4: price and return

5 Prediction models based on SVR and GARCH

In financial markets, people are more interested in the return on a risky asset than in the price. Financiers often define return in this way

$$Return = diff\left(\log\left(price\right)\right) \tag{8}$$

In Eq.(8), $diff$ stands for making one difference and $price$ stands for the price series.

Based on this, we draw the price and return charts for bitcoin and gold respectively as a display

From the Fig.(4) we see that the return of bitcoin is almost ten times that of gold. Of course, with that comes a huge increase in risk. For example, on 2020-3-13, Bitcoin fell 20% in one day. Obviously, if our forecasting model does not have an excellent risk warning at this point, it could be a devastating disaster for our investment strategy!

5.1 The work we did on forecast model

The time series model is an estimate of the future moment based on the historical observations $y_n, y_{n-1}, \ldots$ of the time series $\{y_i\}$ and other observable information $u_n, u_{n-1}, \ldots$ That is, we can use Eq.(9) to describe that:

$$y_{n+k} = f\left(y_n, y_{n-1}, \ldots, y_{n-N+1}; u_n, u_{n-1}, \ldots, u_{n-M+1}\right) \tag{9}$$

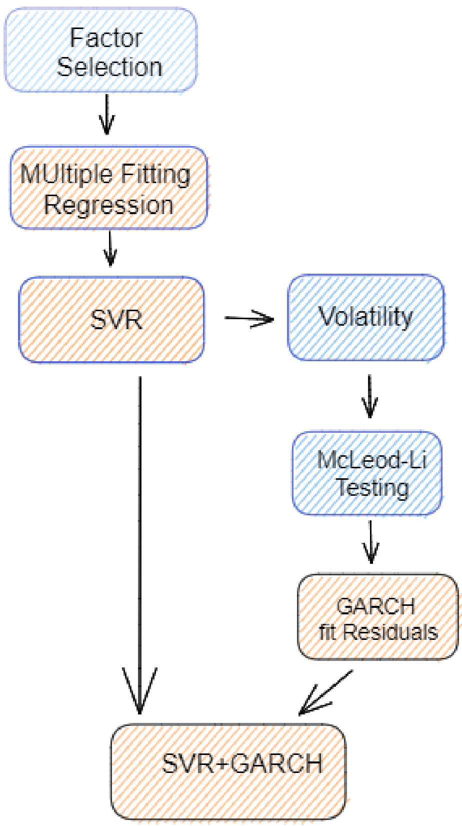


Figure 5: Mean-Var model

We first did multiple regressions on bitcoin and gold returns by trying to choose

some factors commonly used in financial markets, and found that the part that cannot be explained by linear fitting is evident in the return curve. We then used SVR models to predict the fluctuations of the return curves and achieved relatively good results. However, when analyzing the prediction of return, we found a distinctive fat-tail distribution, which we conjecture is due to the significant volatility clustering of the yield fluctuations. We used the McLeod-Li test and found that the guess was correct, so we performed a GARCH fit to the residuals of the SVR prediction result residuals.

5.2 Factor selection and multiple regression

We list the factors in the model preparation section, which are commonly used in the financial industry. In order to reduce the dimensionality and distinguish the importance of the factors, we first perform a PCA analysis on them.

We first proceed to calculate KMO values and perform KMO tests.

The table shows that the KMO value is greater than 0.6, which indicates that there is correlation between the variables, and the p-value of the bartlett test is <0.01, which allows for principal component analysis. According to the scree plot in Fig.6 we can see that 3 principal components should be extracted. The final result shows that when we extract the three principal components represent 89.898% of the variance.

We perform a multiple linear regression of the return series using the three extracted principal components. The following Fig.(7) shows the regression results for the first 48 days.

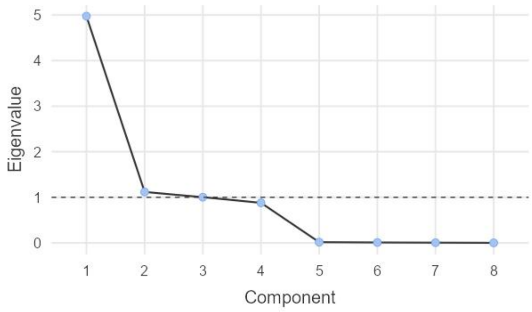


Figure 6: scree plot

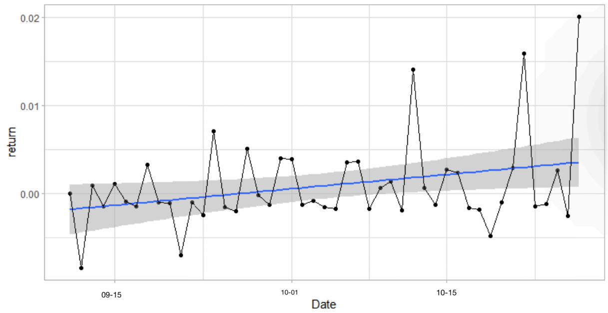


Figure 7: multiple linear regression

We can see from the F chart that the return curve is not well fitted. Obviously, this is due to the large nonlinear component of the data in the financial markets. Therefore, we use the SVR model to make regression forecasts of the return curve.

5.3 SVR Model

Let's first explain SVR.

Through the previous analysis, we found a strong nonlinear relationship between the return curve and the factors. Then let's see how the SVR model solves this problem.

First of all, the SVR model is based on the principle of kernel function, which makes the data show a linear relationship in a higher dimension by increasing the number of dimensions of the data. The kernel function we use here is the Radial Basis Function (RBF).

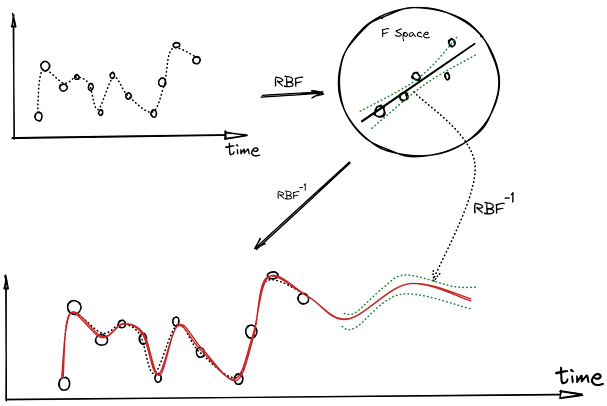


Figure 8: SVR

And then, when we map the data to a high-dimensional null by RBF function and in this space these data are linearly related to each other. Then we construct the linear function Eq.(10) in $F-space$ to fit it:

$$f\left(x\right) = w^T RBF\left(x\right) + b, w\epsilon F \tag{10}$$

Then by the RBF^{-1} function, we get the Eq.(8) in the time dimension.

We determined, through the literature, a ratio of 7:3 for the division of the data set, separated the time after the 50th by 7:3 approximation into time intervals $\{T_i\}$, and predicted stock prices on the first day of each T_i for the entire T_i time period.

We choose the 600th day point in time to show the effect of forecasting 300 days backwards. Fig.(9) and Fig.(10) are the forecast for bitcoin and the forecast for gold price respectively. Red represents the predicted value and blue represents the true value. It can be noticed that using the SVR model makes the prediction effect much stronger.

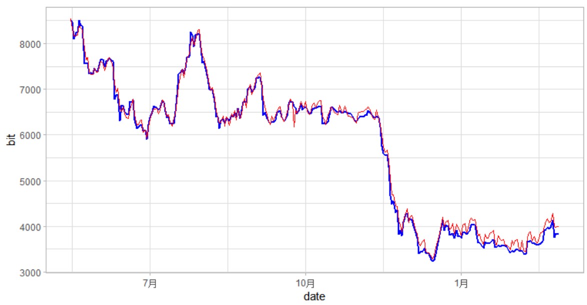


Figure 9: bit predictions

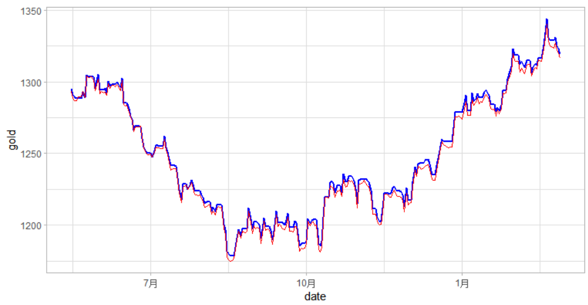


Figure 10: gold predictions

Of course, we are more interested in the forecast of the return as Fig.(11) and Fig.(12) show. And red represents the predicted value and blue represents the true value.Obviously, there is a clear error.

Fig.(14) and Fig.(15) show the return distribution graphs for bitcoin and gold, and clearly they show a clear spike thick tail situation. We then give Q-Q plots of the residuals between the predicted and true values of bitcoin and gold returns. Clearly, their residuals are not normally distributed.

We will then further decompose the residuals of the SVR model predictions. In the financial industry, people often work with time series through ARIMA model. When the residuals of the ARIMA model appear as described above (usually using the McLeod-Li test), we can fit the residuals by means of a GARCH model. Here we try to combine the SVR model and the GARCH model.

5.4 Fitting residuals with GRACH

We first use McLeod-Li test on the time series. It is build for the square of the residuals $\hat{\omega}_t^2$:

$$\hat{\omega}_t^2 = \alpha_0 + \alpha_i \hat{\omega}_{t-1}^2 + ... + \alpha_q \hat{\omega}_{t-q}^2$$

where q is the lag order, if there is no ARCH insignificance, the regression equation coefficient estimates should be 0. Fig.(16) gives the test results of bit and it can be found that almost all p-values are less than 0.05. ARCH is significant and GRACH model should be used.This is also true for gold.

Then, we decompose the residuals of the SVR model with GRACH(m,s):

$$\begin{cases} Return_t = SVR + \omega_t \\ {\sigma_t}^2 = \omega + \sum_{i=1}^{m} \alpha_i {\omega_{t-i}}^2 + \sum_{j=1}^{s} \beta_j {\sigma_{t-j}}^2 \end{cases} \tag{11}$$

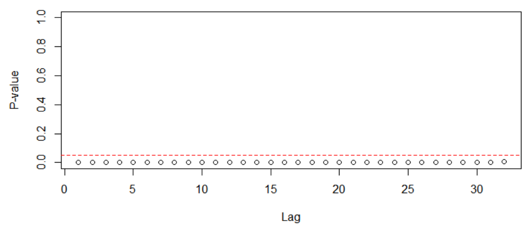


Figure 17: McLeod-Li test of bit

In Eq.(11), ω_t is the residual phase, σ_t is its variance, and m and s are its parameters, which he needs to establish through the ARMA model. We use GRACH(1,1).

Fig.(18) shows a plot of the residuals from the results of predicting bitcoin returns using

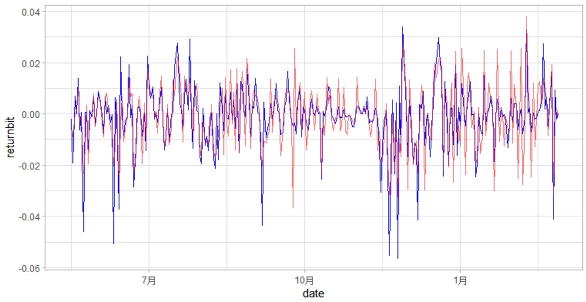


Figure 11: Bitcoin Return Forecast

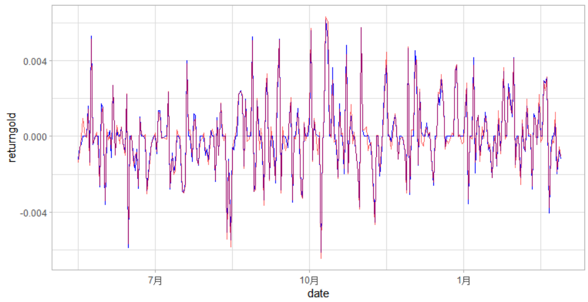


Figure 12: Gold Return Forecast

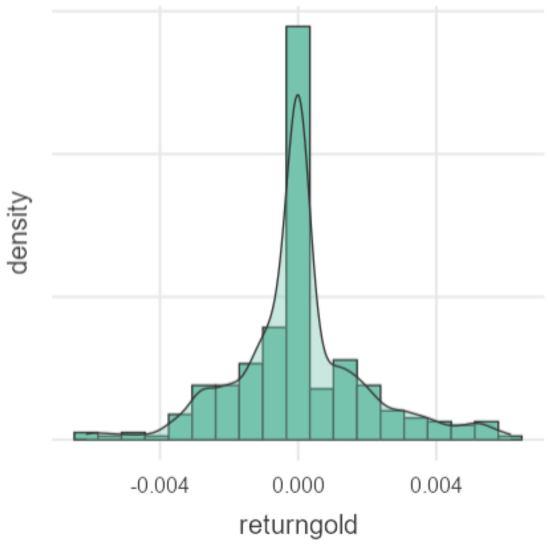


Figure 13: Return Distribution of bit

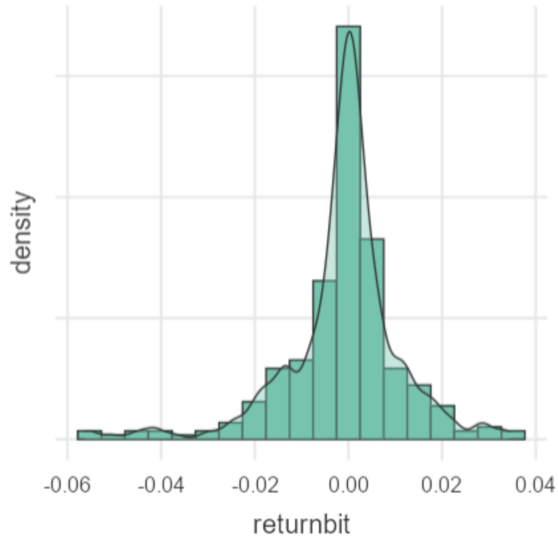


Figure 14: Return Distribution of gold

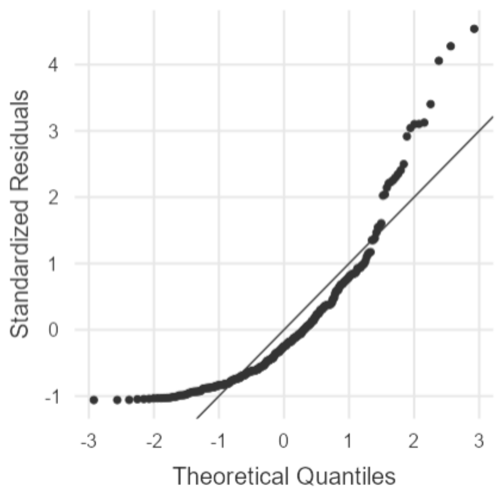


Figure 15: Q-Q of Bit Predictions' Residuals

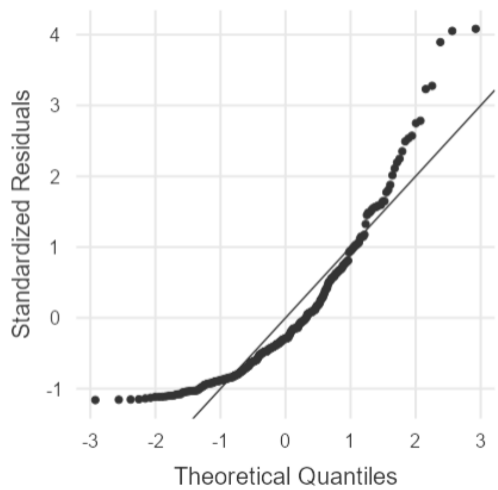


Figure 16: Q-Q of gold Predictions' Residuals

the SVR+GRACH model, and you can see that the normality of the residuals has improved nicely.

This means that the SVR+GARCH model extracts the vast majority of information from the data and is able to predict the vast majority of trends in the time series. Only the white noise xu'lie that cannot be physically measured remains.

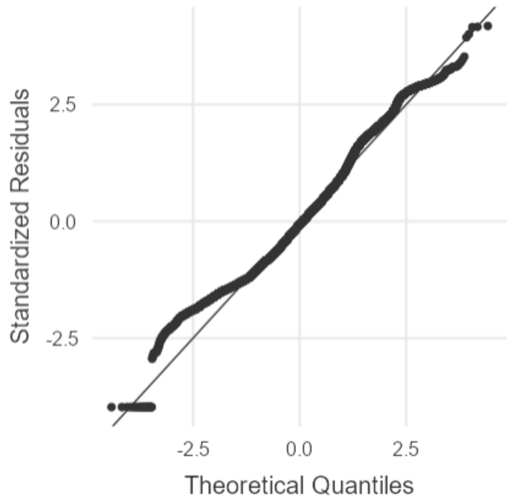


Figure 18: Q-Q of Residuals

6 The best Strategy Model

6.1 programming

After the forecasting model and the optimal portfolio model of risky assets were built, we started to try to build optimization models for our investment transactions. Before we start, we would like to emphasize that our investor is happy to maximize the return on his investment for the same risk. Therefore, in a time period with a positive Sharpe ratio, he will seek to maximize the return of the optimal risky asset.

$$MAX:\ \overline{SharpeRatio_T} = mean\left[SharpeRatio_t\right], t\ \epsilon\ T \tag{12}$$

Eq.(12) is the equation we want to find the optimum, which represents the optimal value of the Sharpe ratio in time period T. It is worth noting that he also has some constraints. We first note that the variables in this expression are the asset allocation of the investors for each day. Assuming that there are N days in time period T, then these variables should be a matrix with N columns in three dimensions:

$$\begin{bmatrix} w_{c1} & \cdots & w_{cN} \\ w_{g2} & \cdots & w_{gN} \\ w_{b3} & \cdots & w_{bN} \end{bmatrix} \tag{13}$$

The w_c,w_g,w_b in this equation represents the weight of cash, gold and bitcoin respectively. Don't be afraid of it because it's a matrix, by Eq.(4) we know that there is an optimal ration between w_g and w_b. And the sum of them is one. So this matrix variable can actually be turned into a one-dimensional vector!

Of course, do not forget our set T_0, who provides us with the constraints:

$$w_{gt} = w_{gt-1}, t\ \epsilon T_0$$

Then we solve this optimization problem by a genetic algorithm.

6.2 Algoritmo genetico

Genetic algorithm is an algorithm that simulates the law of biological evolution by computer to find the best based on the natural and genetic properties of things. It corresponds to the genetic space, encodes each possible value in binary, just like chromosomes, forming a string equivalent to genes, and then evaluates each set of codes according to the expected result and selects the most suitable one. It introduces the biological evolutionary principle of "survival of the fittest" into the coded string population formed by the optimized parameters, and selects individuals according to the selected fitness function and through selection, crossover and mutation in genetics, so that individuals with good fitness values are retained and those with poor fitness are eliminated, and the new population inherits the information of the previous generation and is better than the previous generation. The new population inherits the information of the previous generation and is better than the previous generation. This cycle is repeated until the conditions are met. It is worth noting that GA is not a simple repetition, but a spiral search for superiority.

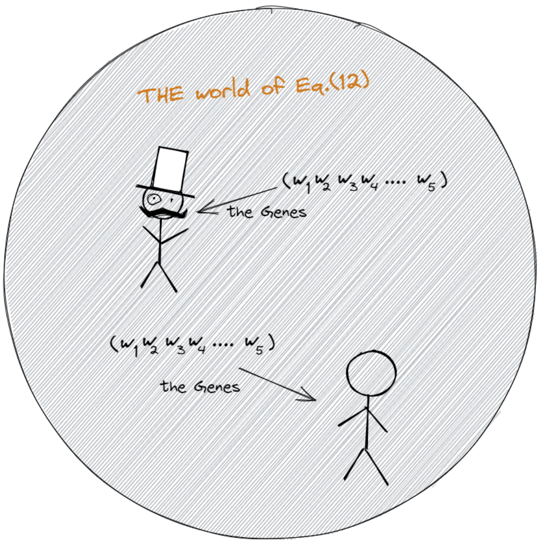


Figure 19: Schematic diagram of the principle of Genetic Algorithm

Let's take an example, in the Fig.(19), the $(w_1...w_n)$ constitutes the gene of two people in the diagram. And this world is the world of Eq.(12). This means that the gene that is closer to the optimal solution of Eq.(12) will fit the world better. Then his gene will be kept and crossed with other good genes into the gene closer to the optimal solution. Then the optimal solution sequence $(w_1...w_n)$ is generated!

We will perform this process in the computer in this way. First we encode the sequence w using floating point encoding. and generate these sequences in a population fashion. Then we simulate the evolution of nature by performing replication, swapping , and random changes to the encoding between populations until our solution converges to within the ideal error.

We demonstrate this process using pseudo-code.

The process of genetic algorithm

Generating $(w_1...w_2)$ genes using floating point encoding
Randomly generate n initial populations
Perform an adaptation assessment test, **if** ture print out
repeat:
Make a copy of $(w_1...w_2)$
Exchange the elements between sequences with probability p_c
Change the elements in the sequence with probability p_m
Perform an adaptation assessment test, **if** ture print out
until Stop Condition is met

6.3 Results Show

With the genetic algorithm, we obtain the optimal strategy for the time interval. Also we can calculate the return of our investment strategy. We show the return profile with the capital allocation in Fig.(20).

The maximum asset achieved by using our strategy is 111383.15 USD.

where the black curve represents our total assets. The red curve is a graph of Bitcoin's price volatility, and the blue is the volatility curve for gold. We can see that gold fluctuates very little compared to bitcoin. It can almost be treated as a risk-free asset. This feature will be mentioned again in our sensitivity analysis.

Table 3: **SHAPERADIO**

time	50-75	75-100	100-150	150-200	200-300	300-400	400-600	600-900	900-1200	1200-1826
SR	1%	2%	1.23%	4.37%	4.37%	5.20%	6.07%	5.21%	8.32%	8.01%

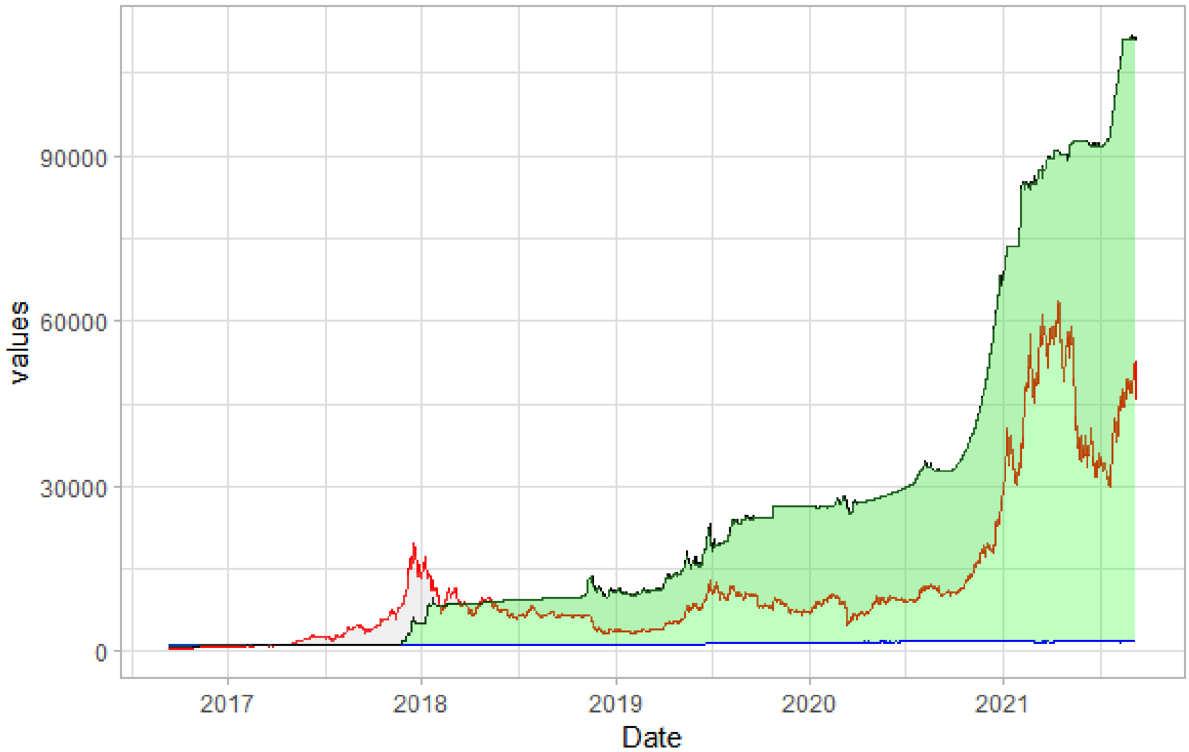
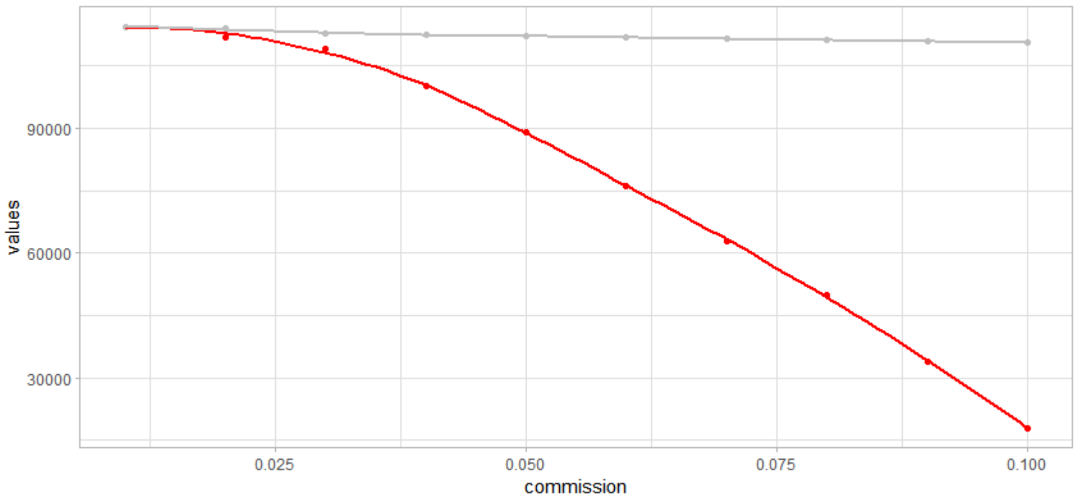


Figure 20: Total Asset Appreciation Curve under Optimal Strategy

We found that the vast majority of our gains came from Bitcoin, but at the same time we managed to sell Bitcoin assets at almost all the big risk points. This is mainly because for a risky asset mix of gold and bitcoin, bitcoin is far more volatile than gold. So as we maximize our Sharpe Ratio, when the price of Bitcoin drops sharply, the entire asset is quickly withdrawn into risk-free and low-risk assets. This mechanism works now, but when the situation changes, it can produce some less than pleasant results.We will discuss this situation in detail in the sensitivity analysis.

7 Sensitivity analysis on commissions

We must note that there is a very important COST factor in the EQ mentioned in Section 4. Although 1% and 2% may seem small. But we find that bitcoin's daily



H

Figure 21: Max asset under different commissions

returns rarely exceed 2% either. This means that when we need to pull money out of risky assets to hedge against risk, the cost of trading can be a significant deterrent.

Fig.(21)shows us the maximum assets that can be achieved with our strategy for different transaction costs.

The red curve represents what happens when bitcoin commissions change, and the gray represents what happens when gold commissions change. We can see that when bitcoin commissions change, our strategy is greatly affected.

To explain this phenomenon, we first need to clarify the relationship between risk-free assets and risky assets. First, in the formulation of this topic, gold is a risky asset. Making it true that gold is a risky asset compared to the US dollar. But when compared to the volatility of Bitcoin, the volatility of gold is insignificant.

So, we need to understand that most of our gains are brought about by bitcoin appreciation. A risky asset like gold can be considered a risk-free asset in comparison to bitcoin. So when money needs to be pulled back from a risky asset into a risk-free asset. Investors could very well choose the US dollar over gold which has a very high commission.

So next let's analyze what happens when the bitcoin commission changes. We find a very interesting phenomenon. That is as bitcoin commissions increase. We can see the size of fluctuations in amount in the whole trading cycle is significantly lower in Fig.(22)

This is obviously because rising transaction costs make it difficult for investors to withdraw money at high risk. If it is not possible to withdraw funds at high risk, then investors will definitely consider reducing their asset holdings at high risk. This is because investors will use the Sharpe ratio to decide how to hold risky assets. At random this time he will choose gold over the high yielding bitcoin. Thus leading to

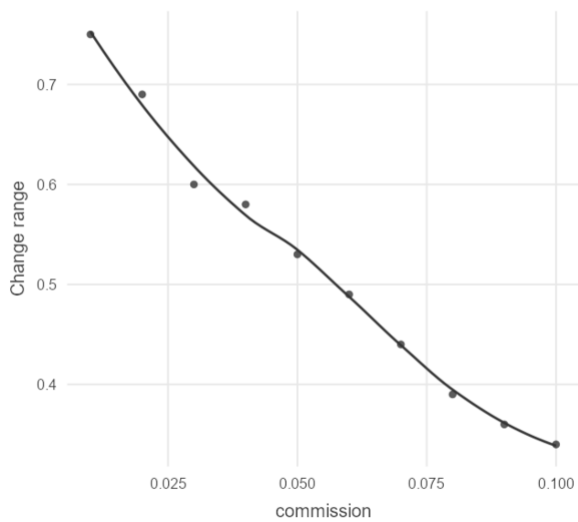


Figure 22: Change range

reduced returns.

8 Strengths and Weaknesses

8.1 Strengths

- Our SVR+GRACH model model can predict the yield curve in the financial industry very well. Especially, it has a good fitting and prediction effect on curves that are distributed into sharp peaks and thick tails
- Optimal Risk Asset Portfolio Model takes into account the combined effect of several key factors and the factors considered are more comprehensive, making the model a better predictor.

8.2 Weaknesses

- The programming implementation of genetic algorithm is complicated, firstly, the problem needs to be coded, and after finding the optimal solution, the problem needs to be decoded,
- Also, the implementation of the other three operators of the genetic algorithm has many parameters, such as crossover rate and variance, and the choice of these parameters seriously affects the quality of the solution, while the current choice of these parameters mostly relies on experience.

References

[1] Rishi K.Narang. Inside the Black Box: A Simple Guide to Quantitative and High Frequency Trading[M]. Beijing: China Machine Press, 2016.

[2] Research on the application of multi-factor model in digital currency market[D]. Chunlin Su. University of Electronic Science and Technology 2019.

[3] A log-optimal asset portfolio model based on VaR risk control [D]. Feng Guoyong. Lanzhou University 2010

[4] Mathematical modeling algorithms and applications / edited by Shoukui Si,Zhaoliang Sun. -2 edition. -Beijing:National Defense Industry Press,2015.4

[5] Research on nonlinear time series prediction methods based on support vector machines [D]. Wang Hongbo. Jiangnan University 2008

[6] SVR Modeling and Parameter Optimization for Financial Time Series Forecasting[D]. Jiao Shuai. Donghua University of Technology 2012

[7] Dong Changhong. MATLAB neural networks and applications [M]. Beijing: National Defense Industry Press, 2007: 96

Appendices

```
MATLAB CODE
tic
%data=csvread('Gdate.csv');
%data=csvread('goldfactor.csv');
data=csvread('bitfactor.csv');
[M,N]=size(data);
input=data(:,1:N-1);
output=data(:,N);


TestNum=M-1226;
input_train=input(1:TestNum,:)';
output_train=output(1:TestNum,:)';
input_test=input(TestNum+1:M-926,:)';
output_test=output(TestNum+1:M-926,:)';

[inputn,inputps]=mapminmax(input_train);
[outputn,outputps]=mapminmax(output_train);

Mdl=fitrsvm(inputn',outputn');

inputn_test=mapminmax('apply',input_test,inputps);

an=predict(Mdl,inputn_test');
SVMoutput=mapminmax('reverse',an,outputps);

plot(SVMoutput,'r');
hold on;
plot(output_test,'b')
toc
output_test=output_test';
b=SVMoutput;
d=output_test;

rmse=sqrt(sum((b-d).^2)/length(b));
mape=sum(abs(b-d)./d)/length(b);
mae=sum(abs(b-d))/length(b);

\textbf(R CODE)
t<-ts(df$Value, frequency =1, start =as.Date("2016-09-11"))
s<- as.Date("2016-09-11")
dates<-seq(from=s, by=1, length.out=1826)
t<-c(0,diff(log(t)))
bit<-data.frame(dates,t)
head(bit)
#
ggplot(bit,aes(x = dates, y = t))+
geom_line(colour="green")
```

```
df<-BCHAIN_MKPRU
bit<-df
ggplot(df,aes(x = Date, y = Value))+
geom_line(colour="green")
#bit garch
acf(bit$return)
pacf((bit$return))
acf((bit$return)^2)
pacf((bit$return)^2)
#mcleod.li
McLeod.Li.test(y=bit$return)
fit1=garchFit(~garch(1,0),bit$t)
summary(fit1)

#gold garch
df2<-lbma_gold1
ggplot(df2,aes(x = Date, y = Value))+
geom_line(colour="green")

df2$return<-c(0,diff(log(df2$Value)))
ggplot(df2,aes(x = Date, y = return))+
geom_line(colour="green")



df2$return[is.na(df2$return)]<-0.000162

acf(df2$return)
pacf((df2$return))
acf((df2$return)^2)
pacf((df2$return)^2)
McLeod.Li.test(y=df2$return)
fit1=garchFit(~garch(2,0),df2$return)
summary(fit1)

adf.test(df2$return,alt="stationary")

armamodel=auto.arima(df2$return)
armamodel
plot(residuals(armamodel))
```

Memo

To： MCM/ICM trader

From: Team 2219846

Date: February 21,2022

Subject: A memo about trading models, strategies and results

Finance is the human eternal topic.Market traders buy and sell volatile assets frequently, with a goal to maximize their total return.Our mission is to analyze the two kinds of assets, gold and bitcoins, to provide the best strategy to traders.

Model:

1.Optimal Risk Asset Portfolio Model, which makes us find greater returns with less risk. We consider both cost and time into the mean-variance model, which allowed us to make the best decision by calculating the Sharpe ratio. By finding maximum Sharpe ratio, we can obtain the optimal investment strategy.

2.Prediction model based on SVR+GARCH aims to get the optimal trading strategy for some time to come.predict risk and return. Since there is a strong nonlinear relationship between the return curve and various factors, we use the SVR +GARCH model to make regression prediction and successfully forecast the future return.

Strategy:

We take three elements of return, risk and cost into account.

And we split the time series according to the ratio of the training sets of the SVR+GARCH model(7:3), within each time period, the Sharpe ratio is used as the objective function to find the best strategy using Genetic Algorithm.

Results:

Based on the modeling and analysis of the data, we get the following results.

1.The initial $1000 investment worth $111383 on 9/10/2021.

2. Cost has great influence on trading strategies. The reasons are as follows.

   (1) the huge difference in price volatility between bitcoin and gold.

   (2) the increased difficulty of withdrawing funds due to higher costs.

3. It is unrealistic to try to make a greater profit by trading multiple times.