Problem Chosen	2022	Team Control Number
C	MCM/ICM Summary Sheet	2202124

Optimal and Dynamic Trading Strategy: Research on Investment in Gold and Bitcoin Markets

Summary

In this paper, our team provide traders with an optimal and dynamic trading strategy every day by using the given data. The specific task is divided into three parts. We solve the problem from three aspects: model establishment, strategy determination and sensitivity analysis, and finally get reasonable results and strategy suggestions.

First, we set up the prediction model and the optimization model. After data preprocessing, we set up ARIMA model to predict the future price of bitcoin and gold. The stationarity of time series is tested by ADF test, and the order of model is determined by ACF and PACF, thus establishing ARIMA model. We use RMSE to evaluate the applicability of the model, and the results show that the RMSE of the two models is small, which reflects the good predictive ability of the model.

Secondly, we use nonlinear programming to establish the optimization model. Considering the rest of the gold market on non-trading days, we divide the model into two cases: both gold and bitcoin markets are on trading days and only bitcoin market is on trading days. We define volatility, risk and yield, and establish equations, so that we can calculate the optimal portfolio and the maximum return of each day. Here we use dynamic programming to solve nonlinear programming.

Further, we make the daily optimal strategy based on the above two models. For each day, firstly, the ARIMA model is established to forecast the price of gold and bitcoin in the next six days, and record the data from two days before the current day to the next three days respectively. The new time series is used to calculate the risk and yield, which is brought into the optimization model, and dynamic programming is used to solve the nonlinear programming by iteration, which can get the optimal portfolio and the maximum return of the day. We use the existing data to test the optimal strategy. By using Python to program, from the fourth day of data, the optimal strategy is brought into every day's data to form dynamic simulation, and finally the result can be obtained. It is clear that every day is regarded as a cycle. According to the results, on September 10^{th}, 2021, there are \$ 3346387.17 not in investment project, and the total assets is \$ 6327123.66. The investment portfolio is that the trader should put in gold of 0, bitcoin of \$ 2980736.4980757544. At the same time, it can be seen from the result chart that our strategy is optimal and suitable for long-term investment.

Finally, we analyze the sensitivity of the model. In order to get the relationship between the model and transaction costs of gold and bitcoin, we parameterize these two quantities in the model, and take dynamic values for different transaction costs. We use the final total assets to represent the influence results of transaction costs on the model, and use three-dimensional images to represent them. On the other hand, we use these two parameters to make a regression model for the final total assets, considering the least square regression model with and without interactive effects respectively. Through the corresponding tests, it is concluded that the model results are sensitive to the change of transaction costs of bitcoin , but insensitive to the change of transaction costs of gold.

Keywords: ARIMA model; dynamic programming; lagrange multiplier method; optimization algorithm; the least square regression

Contents

1 Introduction

1.1 Problem Background

Nowadays, market transactions have become a way for people to make money. As market traders, they often buy and sell volatile assets, so that they can get the maximum return. Two common investments are gold and bitcoin. The gold market is used to trade gold and exchange gold coins. It is an independent resource with high value. Investment in gold can help investors avoid possible problems in the economic environment. Bitcoin is a virtual currency, and bitcoin economy uses a distributed database composed of many nodes in the entire P2P network to confirm and record all transactions.

Although there are differences between them, they have some similar characteristics in market transactions. They can trade through a certain commission, and buy, hold or sell with the change of daily price, so as to get a return. However, different returns can also be determined by different investments because of difference commission, trading time and changing trend between gold market and bicoin market. Therefore, it leads to a question that how to find the best portfolio in the face of these two investments so that traders can get the maximum return.

1.2 Restatement of Problems

In this problem, we are given two time series spreadsheets, the data of which are the closing price of a troy ounce of gold in U.S. dollars and the price in U.S. dollars of a single bitcoin on the indicated date from November 9, 2016 to October 9, 2021.

In the question, the trader will have a portfolio consensus of cash, gold, and coins [C, G, B] in U.S. dolls, troyounces, and coins, respectively on each trading day. The initial state is [1000, 0, 0]. The commission for each transaction (purchase or sale) costs α% of the amount traded. Assume $\alpha_{\text{gold}} = 1\%$ and $\alpha_{\text{bitcoin}} = 2\%$. There is no cost to hold an asset. We should use these data to solve the following three problems under hypothetical conditions:

- Develop a model that gives the best daily trading strategy based only on price data up to that day. How much is the initial \$ 1000 investment worth on 9/10/2021 using your model and strategy?
- Present evidence that the model provides the best strategy.
- Determine how sensitive the strategy is to transaction costs. How do transaction costs affect the strategy and results?

1.3 Our Work

Our work mainly includes the following, as shown in the figure 1. We mainly use ARIMA model to predict the price and use dynamic programming method to solve nonlinear programming, establishing the daily optimal trading strategy. In every day, we establish ARIMA model to predict the price of gold and bitcoin in the next six days, calculate the volatility, risk and yield, and establish nonlinear programming. The optimal portfolio and the maximum return are solved by the dynamic programming method, and finally the optimal solution can be obtained.

Then, sensitivity analysis is carried out by changing transaction costs and establishing least square regression, and finally advantages and disadvantages are analyzed and suggestions are put forward.

Finally, we carry out the sensitivity analysis by changing transaction costs and establishing least square regression. And we analyze the strengths and weakness of the models, and put forward the reasonable suggestions.

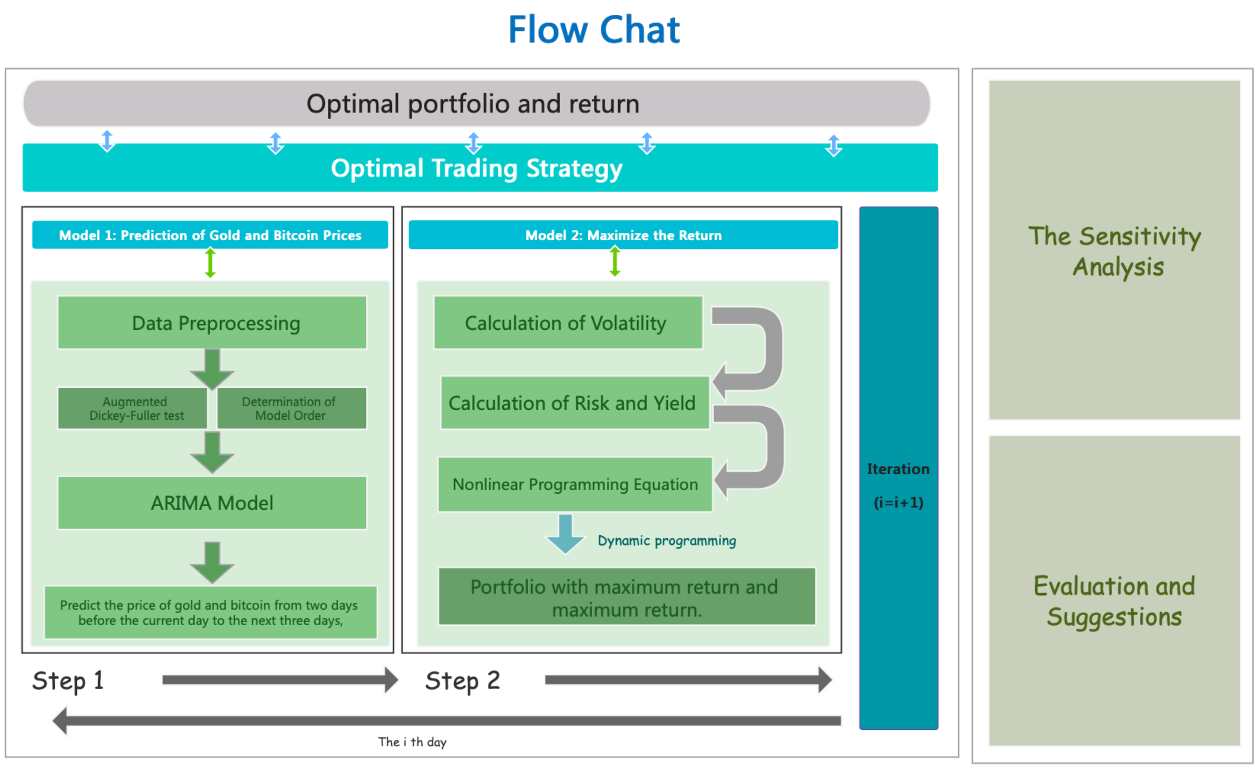


Figure 1: The Flow Chat in this paper

2 Assumptions and Nomenclature

2.1 Assumptions

We make several assumptions in our model:

- Investors are risk-averse.
- The handling fee for gold and bitcoin must be paid in US dollars.
- There is no minimum unit for the transaction.

- Gold and bitcoin can only be traded once on the same day, regardless of the transaction order.

2.2 Detailed Definitions

We give some detailed definitions below in Table 1:

Table 1: Detailed Definitions

Symbol	Definition
α_{gold}	Transaction costs of gold
α_{bitcoin}	Transaction costs of bitcoin
LG	Daily gold price after data preprocessing
LB	Daily bicoin price after data preprocessing
SG_i	Market price of gold on the i^{th} day
SB_i	Market price of bitcoin on the i^{th} day
fSG_i	Predicted price of gold on the i^{th} day
fSB_i	Predicted price of bitcoin on the i^{th} day
D_i	Total assets on the i^{th} day
G_i	Total assets of gold on the i^{th} day
B_i	Total assets of bitcoin on the i^{th} day
P_{gi}	Volatility of gold price on the i^{th} day
P_{bi}	Volatility of bitcoin price on the i^{th} day
q_{gi}	Risk rate of gold price on the i^{th} day
q_{bi}	Risk rate of bitcoin price on the i^{th} day
R_{gi}	Yield of gold price on the i^{th} day
R_{bi}	Yield of bitcoin price on the i^{th} day
r_{gi}	Real yield of gold price on the i^{th} day
r_{bi}	Real yield of bitcoin price on the i^{th} day
y_{qi}	Risk of gold price on the i^{th} day
x_{qi}	Risk of bitcoin price on the i^{th} day
y_i	Assets transferred from the gold market on the i^{th} day
x_i	Assets transferred from the bitcoin market on the i^{th} day

Note: Negative numbers represent withdrawal.

3 Model 1 – Prediction of Gold and Bitcoin Prices

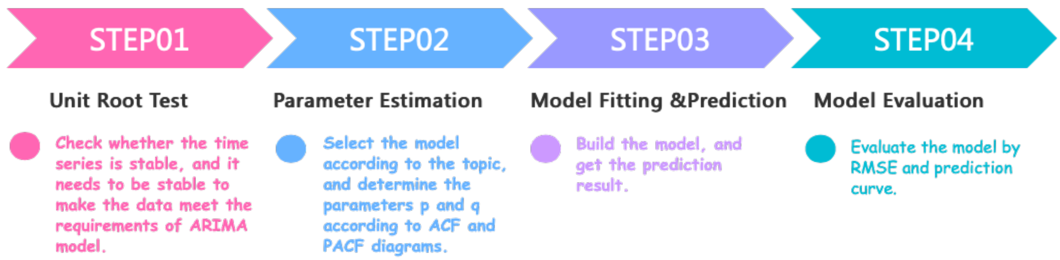
3.1 Data Preprocessing

Firstly, we use Python to preprocess the data because the availability of data must be guaranteed. All the data are time series, so in order to facilitate the model we build in later, we first standardize and normalize the data, which can eliminate the dimension. Then we take logarithm processing to make the time series more stable, which can eliminate the heteroscedasticity problem.

Secondly, the data containing gold prices had missing values on non-trading days. So in order to ensure the continuity of time series, we filled them by backward filling method, and finally got 1,825 items of daily gold prices and 1,826 items of daily bitcoin prices.

3.2 ARIMA Model

In order to find the optimal investment strategy, we consider establishing a model to predict the bitcoin and gold prices in the future, and then carry out the next step of optimization strategy through the trend. For price forecast, based on time series, we consider using ARIMA model to forecast. The steps to build ARIMA model are as below:



ARIMA model is a model for forecasting time series. It mainly has five forms: AR(p), MA (q),AR(p), MA (q), ARMA(p, q), ARIMA(p, d, q) and ARIMA $(\mathrm{p}, \mathrm{d}, \mathrm{q}) \times (\mathrm{P}, \mathrm{D}, \mathrm{Q})_{\mathrm{s}}$. In this paper, the model of ARIMA(p,d,q) is used to solve the problem. ARIMA(p, d, q) model is an extension of ARMA(p, q) model, which can be expressed as:

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)(1-L)^d X_t=\left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t \tag{1}$$

L is the lag operator, $d \in \mathbb{Z}$, d>0 .

However, before establishing the model, it needs to be clear that we need stationary time series to build the model, but it is not difficult to find that the time series in this problem is not stationary, so the data should be transformed into stationarity first.

Using the moving average method to calculate rolling mean and standard deviation of gold and bitcoin. The result is shown in figure 2 and 3:

It can be seen from figures that this method does not play a significant role in the stability of time series, so we take logarithms to make time series more stable, which can also eliminate the heteroscedasticity problem as well. In order to further determine whether the time series is stable, the unit root test is carried out next.

3.2.1 Augmented Dickey-Fuller Test

We use ADF test method to verify the stationarity. If it is a stationary sequence, the model can be built directly. If it is a non-stationary sequence, it needs to be differentiated to transform the

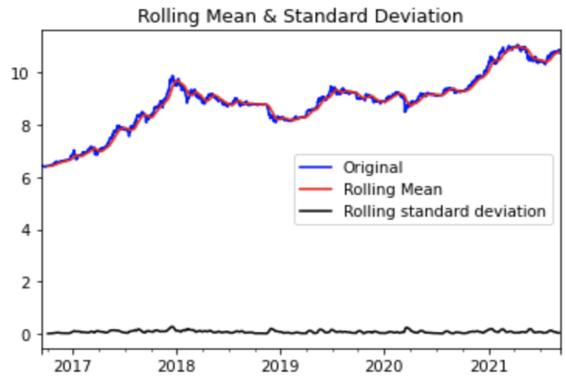


Figure 2: Mean and standard deviation of gold

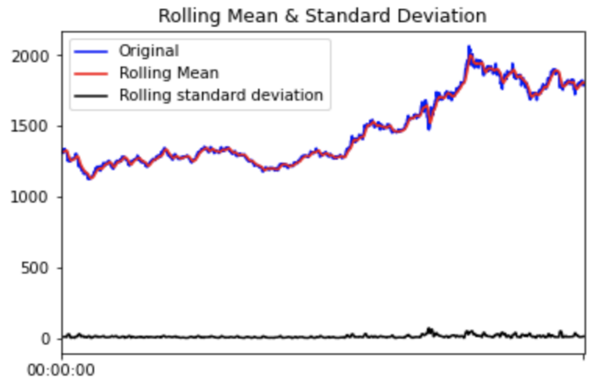


Figure 3: Mean and standard deviation of Bitcoin

non-stationary sequence into stationary sequence to ensure the more accurate analysis results. We still use Python to program the data, and the results are as follows:

Table 2: The result chat of ADF test

Variable	DF test statistic	5% Critical value	P-value	Result
LG	-0.430483	-2.863134	0.904883	Nonstationary
LB	-1.521970	-2.863134	0.522616	Nonstationary
D(LG)	-1.733088e+01	-2.863144	5.410944e-30	Stationary**
D(LB)	-1.224698e+01	-2.863134	9.733162e-23	Stationary**

Note: * * means that the original hypothesis is rejected at the significance level of 5 %, and D(X) means the first-order difference of variables.

It can be concluded from Table 2 that LG and LB are both non-stationary sequences, but after the difference, they both reject the original hypothesis at the significance level of 5 %.

3.2.2 Determination of Model Order

After the first-order difference, the data is stable, and then it is necessary to determine the order of model, which are p and q. We use the images of Autocorrelation function(ACF) and Partial Autocorrelation function(PACF) to solve this problem. ACF is a complete autocorrelation function, which describes the degree of correlation between the current value of the sequence and its past value. PACF is a partial autocorrelation function or partial autocorrelation function, which is used to find the correlation between the residual error and the next lag value. We can determine the order of the model by using ACF or PACF, which means to find a certain time when ACF or PACF tends to 0. ACF determines q and PACF determines p. We draw two time series images of ACF and PACF, as shown in Figures 4, 5, 6 and 7, in which figures 4 and 5 represent bitcoin price of ACF and PACF, and figures 6 and 7 represent gold price of ACF and PACF:

In figures, it can be concluded that two data are basically show the same results, so let's take bitcoin data as an example. Figure 4 and 5 show that both ACF and PACF have tailing characteristics, and both of them have obvious first-order correlation. Meanwhile, ACF shows that after the first time

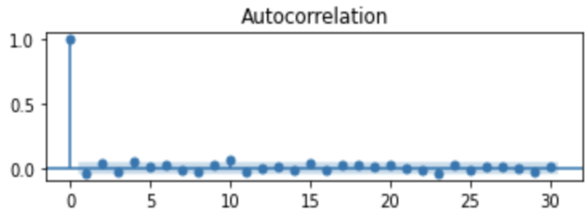


Figure 4: Bitcoin: Image of ACF

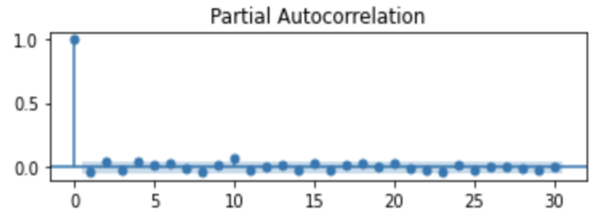


Figure 5: Bitcoin: Image of PACF

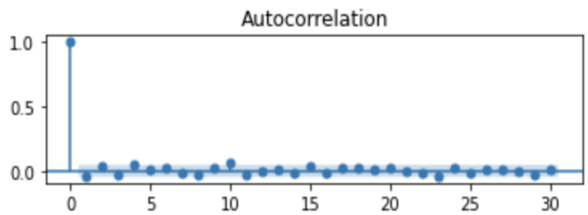


Figure 6: Gold: Image of ACF

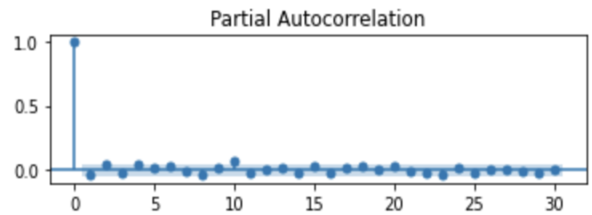


Figure 7: Gold: Image of PACF

lag, it gradually tends to 0, and PACF shows that the correlation value does not exceed the effective boundary (0.5) after the first time lag. Therefore, the parameters can be determined according to AIC minimum information criterion and relevant model fitting index, which are d = 1, p = 1 and q = 1. So the model we build is ARIMA (1,1,1).

3.2.3 Results of ARIMA Model

Through the analysis of ARIMA model, we get the fitting result, and the predicted value is reduced by inverse correlation transformation. We use root mean square error (RMSE) to evaluate the model, and eliminate the influence of "non-predicted" data in advance. The final prediction results are shown in figure 8 and figure 9:

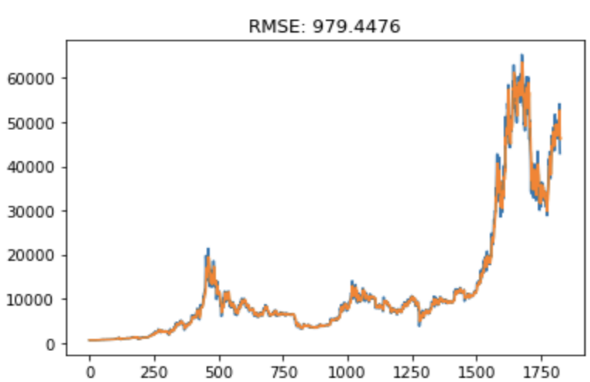


Figure 8: The forecast result of bitcoin

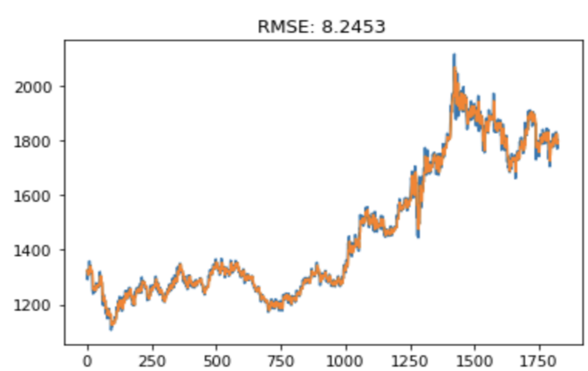


Figure 9: The forecast result of gold

From the figures, the RMSE of bitcoin and gold prediction results are 979.4476 and 8.2453 respectively. Through fitting images and RMSE, it can be concluded that the fitting effect of the model is good.

4 Model 2 ¯ Maximize the return

Because our goal is to find the best trading strategy based only on price data up to that day, we need to consider finding the best investment portfolio between gold and bitcoin to maximize the return. In addition to the constraint condition of the problem itself, the two most important indicators that need to be considered are risks and benefits. Therefore, based on the given risk index calculation method, we calculate the risk, establish the variance of nonlinear programming with the maximum total return as the decision objective. Then, using dynamic programming method, we get the optimization model of portfolio. The specific steps are as follows:

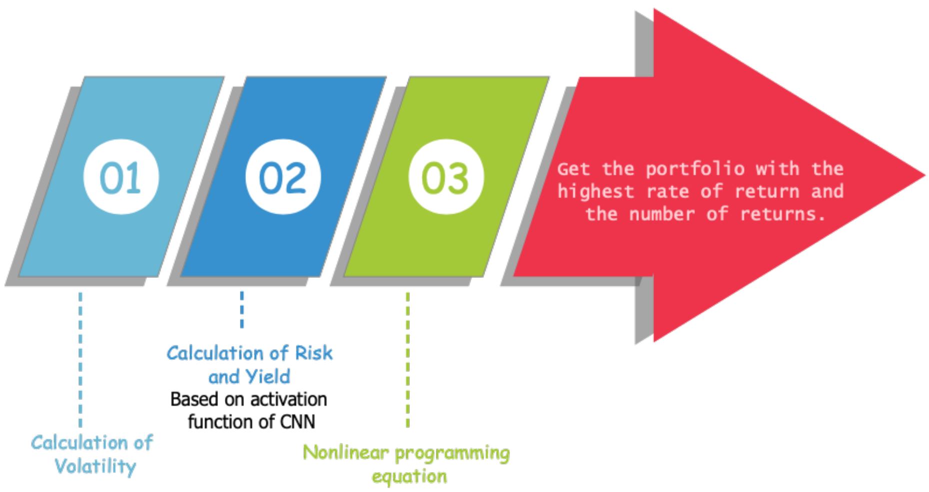


Figure 10: The Structure of Model 2

4.1 Calculation of Volatility

Volatility is the fluctuation degree of financial asset prices, which is used to reflect the risk level of bitcoin and gold prices. The higher the volatility, the stronger the uncertainty of the yield.

In order to measure the risk, we first give the formula for calculating the volatility. The mathematical formula is as follows:

$$P_g = \frac{s_g}{SG_i} \tag{2}$$

$$P_b = \frac{s_b}{SB_i} \tag{3}$$

$$s_g = \sqrt{\frac{\sum_{k=i-2}^{i+3}\left(SG_k - \bar{SG}\right)^2}{6}} \tag{4}$$

$$s_b = \sqrt{\frac{\sum_{k=i-2}^{i+3}\left(SB_k - \bar{SB}\right)^2}{6}} \tag{5}$$

P_{gi} is the volatility of gold price on the i^{th} day, P_{bi} is the volatility of bitcoin price on the i^{th} day, and s_g, s_b are the price standard variance of gold and bitcoin from two days before the current day to the third day after the current day respectively.

4.2 Calculation of Risk and Yield

※ Calculation of Risk

First, we define the risk rate. Let's put forward the concept of hyperbolic tangent function. Hyperbolic tangent function is a kind of hyperbolic function, which is defined by exponential function and can be used as the activation function of neural network. The expression is:

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{6}$$

We use the hyperbolic tangent function to calculate the risk rate through the volatility, in order to project the risk into the range of $0 - 1$, which is convenient for calculation. On this basis, we define the calculation formula of the risk rate of gold and bitcoin on the i^{th} day as follows:

$$q_{gi} = 1 - \tanh\left(3P_{gi}\right) = 1 - \frac{e^{3P_{gi}} - e^{-3P_{gi}}}{e^{3P_{gi}} + e^{-3P_{gi}}} \tag{7}$$

$$q_{bi} = 1 - \tanh\left(3P_{bi}\right) = 1 - \frac{e^{3P_{bi}} - e^{-3P_{bi}}}{e^{3P_{bi}} + e^{-3P_{bi}}} \tag{8}$$

Then, we define the assets after considering risks as:

$$y_{qi} = q_{gi} y_i \tag{9}$$

$$x_{qi} = q_{bi} x_i \tag{10}$$

※ Calculation of Yield

According to historical documents, we define the yields of gold and bitcoin as follows:

$$R_{gi} = \frac{f SG_{i+1}}{SG_i} \tag{11}$$

In which fSG_i refers to the predicted gold price on the ith day, and fSB_i refers to the predicted gold price on the Ith day.

$$R_{bi} = \frac{fSB_{i+1}}{fSB_i} \tag{12}$$

After defining risk and yield, we set up nonlinear programming.

4.3 Nonlinear Programming Equation

Lagrange Multiplier Method

In mathematical optimization problems, the Lagrange multiplier method (named after mathematician Joseph Louis Lagrange) is a method of finding the extremum of a multivariate function whose variables are bounded by one or more conditions. This method transforms an optimization problem with n variables and k constraints into an extremum problem of a system of equations with n + k variables, whose variables are not subject to any constraints. This approach introduces a new scalar unknown, the Lagrange multiplier: the coefficient of each vector in the linear combination of the gradients of the constraint equation. This method involves multi-directional knowledge such as partial differential, total differential or chain method (not explained in depth here), so as to find the value of the unknown that can make the differential of the implicit function set to zero.

According to the above calculation method of risk and yield, a nonlinear programming can be established. However, since gold is closed every week on non-working days, it is not necessary to consider the impact of the gold market on non-trading days. Based on this, we will discuss two situations in the problem. The first is when both gold and bitcoin are on the trading day, and the second is when only bitcoin is on the trading day. The details are as follows.

- **Both Gold and Bitcoin Are on the Trading Day**

In this part, the nonlinear programming is set as follows:

$$\begin{aligned} \max \quad & D_i - x_i - y_i - 0.02\,|x_i| - 0.01|y_i| + (B_i + x_i)\,R_{bi} + (G_i + y_i)\,R_{gi} \\ s.t. \quad & \begin{cases} D_{i+1} = D_i - x_i - y_i - 0.02|x_i| - 0.01|y_i| \geqslant 0 \\ B_{i+1} = B_i + x_i \geqslant 0 \\ G_{i+1} = G_i + y_i \geqslant 0 \end{cases} \end{aligned} \tag{13}$$

- **Only Bitcoin Is on the Trading Day**

In this part, it is only need to delete the information about gold. The specific nonlinear programming is as follows:

$$\begin{aligned} \max \quad & D_i - x_i - 0.02\,|x_i| + (B_i + x_i)\,R_{bi} \\ s.t. \quad & \begin{cases} D_{i+1} = D_i - x_i - 0.02|x_i| \geqslant 0 \\ B_{i+1} = B_i + x_i \geqslant 0 \end{cases} \end{aligned} \tag{14}$$

Through dynamic programming, we can get the estimated portfolio with the maximum return and x_i,yi without the consideration of risk.

5 Optimal Trading Strategy

5.1 Description of the Strategy

Based on Model 1 and Model 2, we conclude a strategy that can get the optimal portfolio and the maximum return every day through the data only up to that day, and we give the following optimal trading strategy:

Firstly, we use Model 1 to forecast the trading price of gold and bitcoin in the next six days, and record the data from two days before the current day to the next three days, which can get two new time series, each of which has only six data. Using the newly obtained data, the volatility, risk, yield and real yield are calculated, which means we get the values needed in the next nonlinear programming of Model 2.

Secondly, we use the dynamic programming method to get the optimal portfolio and the maximum return of the day. Considering the rest of the gold market on non-trading days, the corresponding nonlinear programming is brought in for iteration to calculate the optimal allocation portfolio and the maximum return of the day.

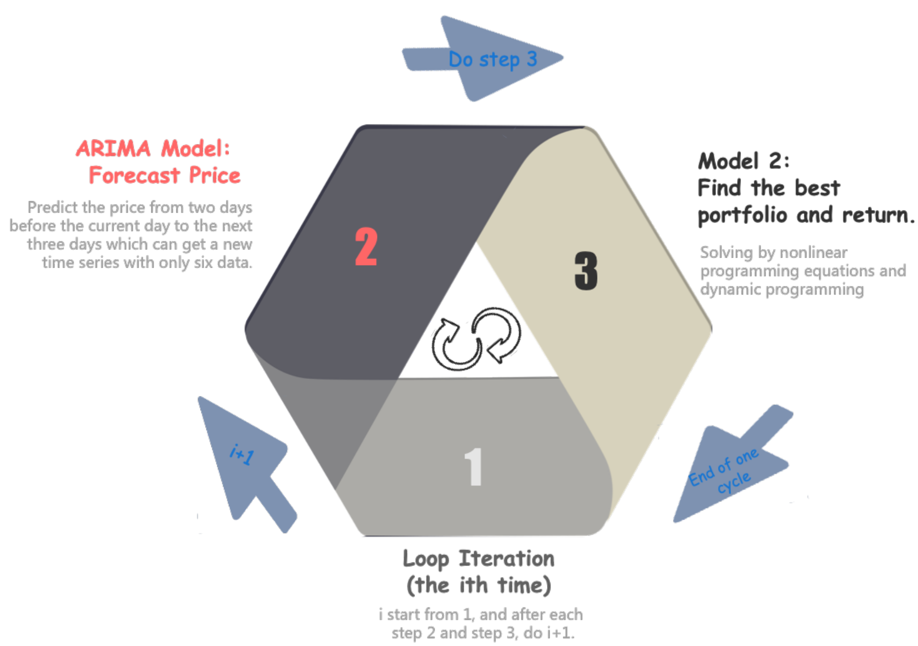


Figure 11: The Trading Strategy

We use Model 1 to forecast the trading price of gold and bitcoin in the next three days, and record the data from two days before the current day to the next three days, which can get two new time series, each of which has only six data. Using loop iteration and dynamic programming to solve the optimal portfolio and maximum return, which is also the core of our optimal strategy.

Specifically, Model 1 and Model 2 are brought into the loop. From the third day, September 14^{th}, 2016, we do the following operation for each cycle. ARIMA model is established based on price data up to that day to predict the prices of gold and bitcoin in the next three days, so as to obtain the new time series mentioned in the previous step. Then, new data is brought into Model 2, and the optimal portfolio and return of each day can be obtained by dynamic programming . This is the optimal trading strategy model established in this paper.

5.2 Test the Model and Strategy

We use the preprocessed data to test the daily optimal trading strategy. Here we make the optimal trading strategy dynamic, letting the data of each day is based on the data of the previous day. Before that, we first define a variable the real rate of return of the gold and bitcoin markets:

$$r_{gi} = \frac{SG_i}{SG_{i-1}} \tag{15}$$

$$r_{bi} = \frac{SB_i}{SB_{i-1}} \tag{16}$$

On the basis of the results of the current day, the next day's total assets, assets in the gold market and assets in bitcoin are iteratively updated. The calculation formulas of the next day's total assets, assets in the gold market and assets in bitcoin are as follows:

$$\begin{cases} D_{i+1} = D_i - x_{qi} - y_{qi} - 0.01|x_i| - 0.02|y_i| \\ B_{i+1} = (B + x_{qi}) \cdot r_{b(i+1)} \\ G_{i+1} = (G + y_{gi}) \cdot r_{g(i+1)} \end{cases} \tag{17}$$

Through the above algorithm, combined with the daily optimal trading strategy, we use Python to test the strategy and models start from the fourth day of data, September 14^{th}, 2016. We use the SciPy library in Python to realize the calculation results of dynamic programming. For trading day and non-trading day problems, we classify and discuss, and finally merge into a unified result. We put the results of each day's calculations (including cash, money in gold, and money in bitcoin) into a vector, which ultimately shows the changes in the flow of money as a graph. The change images of total assets, bitcoin market assets and gold market assets are as follows:

According to the relationship chart, we find that the total fund rises slowly in the first one thousand days, with the change range not exceeding one-seventh of the total. As time goes on, the total fund grows rapidly, and there is a period of rapid growth after 1250 days and 1500 days, which is exactly

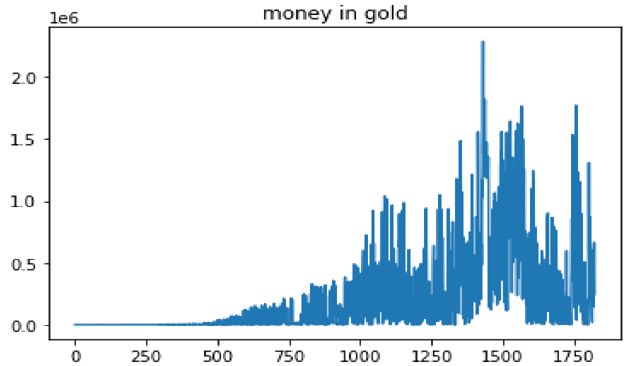


Figure 12: The change of total assets

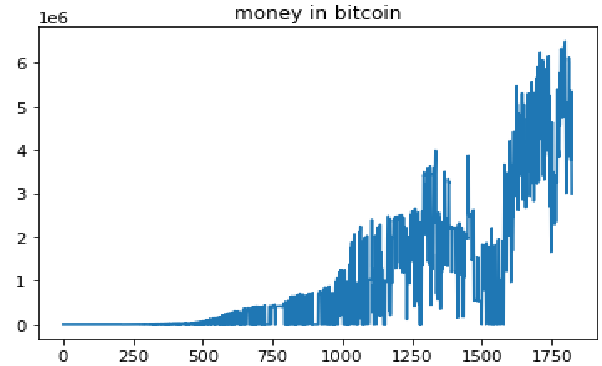


Figure 13: The change of cash assets

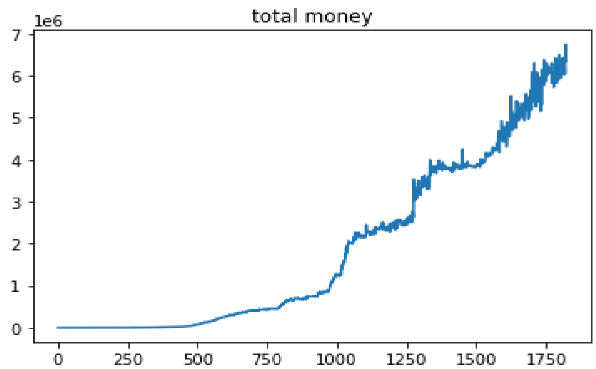


Figure 14: The change of gold

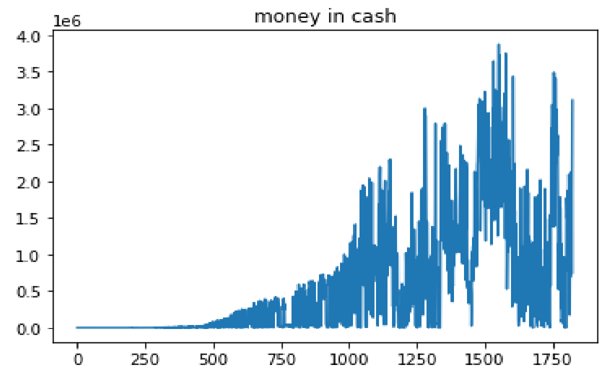


Figure 15: The change of bitcoin

consistent with the rapid growth of bitcoin price in the later period. And then when we look at the second chart, it's basically unchanged for the first 500 days, and then it starts to grow with great fluctuation. When it is in a declining or relatively small period of time, it means that most of the cash on hand is used as assets to invest in gold or bitcoin, while when the amount of cash is very high, it means that we have predicted the price decline of gold and bitcoin, and took out the assets to avoid losses, which is consistent with the fluctuation of funds in the last two figures.

At the same time, we get the result: in September 11^{th},2021, we should invest 0 in gold, \$ 2980736.4980757544 in bitcoin, and have \$ 3346387.17 not in investment project, in total assets \$ 6327123.66.

6 Sensitivity Analysis

6.1 Determine the Sensitivity of Our Strategy to Transaction Costs

According to the topic assumption, $\alpha_{\text{gold}} = 1\%$ and $\alpha_{\text{bitcoin}} = 2\%$, our strategy presents a good performance.

But in many cases, Our strategies and models need to be presented when transaction costs are unknown. Therefore, the robustness of the model to transaction costs is needed. In this problem, we use the change of final assets to represent the impact of transaction costs under different ratios. Since the final total asset value is large, we take the logarithm of it and put it into the data input to make it smoother. Combined with market transaction costs, we consider the transaction costs of gold and bitcoin ranging from 0 to 0.2 with the interval of 0.01, thus obtaining the three-dimensional figure as follows:

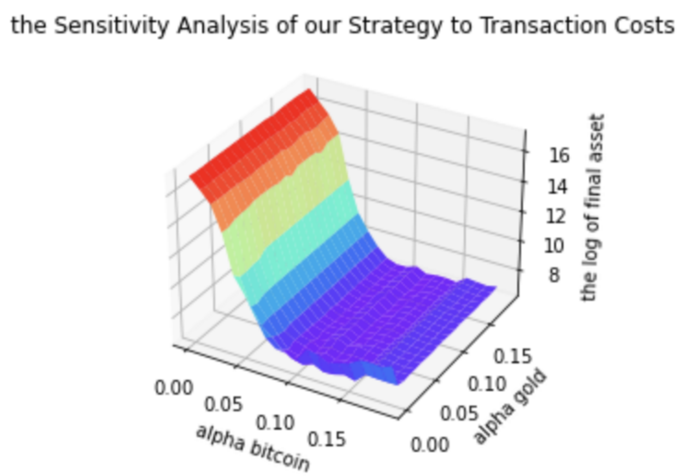


Figure 16: Sensitivity analysis on transaction costs

6.2 Impact of Transaction Cost on Results

We established a regression model to give the impact of the change of transaction costs on the final result. Here, we firstly adopted the least square regression model with interactive effect, and obtained the following results through programming calculation:

Table 3: OLS Regression Results

Dep. Variable:	final_asset_log	R-squared:	0.643
Model:	OLS	Adj. R-squared:	0.641
Method:	Least Squares	F-statistic:	238.0
Prob (F-statistic):	3.04e-88	Log-Likelihood:	-881.81
No. Observations:	400	AIC:	1772.
Df Residuals:	396	BIC:	1788.
Df Model:	3		
Covariance Type:	nonrobust		

Obtained by the data in the table, the transaction cost of bitcoin coefficient is about - 49, variance of about 3.7, and it is extremely significant, while the transaction cost of gold and the interaction effect between gold and bitcoin is highly insignificant , with their P values all greater than 0.5. Meanwhile

Table 4: Coefficient Info(including interaction effect)

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	13.7995	0.409	33.702	0.000	12.995	14.605
alpha_gold	0.4160	3.684	0.113	0.910	-6.828	7.660
alpha_bitcoin	-49.3714	3.684	-13.400	0.000	-56.615	-42.128
alpha_gold:alpha_bitcoin	-17.7201	33.154	-0.534	0.593	-82.901	47.460

the variance of the interaction effect reached nearly a staggering 33, further illustrate the interaction effect is not significant. Therefore, we re-considered the least square regression model without the interaction effect of the two. The results we got this time are not very different from the above. Similarly, the transaction commission of Bitcoin is significant extremely, with a coefficient of about -51 and a variance of about 1.9, while the transaction cost of gold is still not significant.

Table 5: OLS Regression Results

Dep. Variable:	final_asset_log	R-squared:	0.643
Model:	OLS	Adj. R-squared:	0.641
Method:	Least Squares	F-statistic:	357.5
Prob (F-statistic):	1.64e-89	Log-Likelihood:	-881.96
No. Observations:	400	AIC:	1770.
Df Residuals:	397	BIC:	1782.
Df Model:	2		
Covariance Type:	nonrobust		

Table 6: Coefficient Info(including interaction effect)

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	13.9595	0.279	49.988	0.000	13.410	14.508
alpha_gold	-1.2674	1.910	-0.664	0.507	-5.023	2.488
alpha_bitcoin	-51.0549	1.910	-26.730	0.000	-54.810	-47.300

All in all, our model does not have strong robustness for transaction cost of bitcoin, which reflects that different transaction costs have a great impact on total assets, especially changes in the transaction cost of bitcoin, which also confirms that our assets are mainly earned by bitcoin. Meanwhile, in terms of the transaction cost of gold, our model has a strong rubostness, which is attributed to the fact that the effect of gold's transaction cost is not significant. This is also well demonstrated in our three-dimensional figure.

7 Assessment and Promotion of Model

7.1 Strength and Weakness

7.1.1 Strengths

✓ Our solution process is robust, and the established model is intuitive and simple.

✓ The ARIMA model is used to forecast the price, and the model has good properties, which provides a basis for subsequent optimization.

✓ The dynamic programming algorithm is used to optimize and solve the nonlinear programming problem, considering both the gold trading day and the non-trading day. This algorithm can ensure that we can get the maximum return under the constraint conditions and ensure the feasibility of the optimal trading strategy.

✓ The research problems are closely combined with reality, which can effectively provide more accurate and optimized trading strategies for traders who invest in both gold and bitcoin markets from a statistical point of view, and is of great significance.

7.1.2 Weakness

⋄ We only considered the situation of trading once a day, without considering the transaction sequence of multiple transactions and various investments, ignoring some detailed conditions.

⋄ We are only based on two portfolios of bitcoin and gold, and the model does not consider more than two portfolios.

7.2 Promotion and Improvement

Considering only the markets of gold and bitcoin, and in terms of days, our model is very robust. Therefore, our model is applicable when we don't consider the number and order of daily transactions and the portfolio of other markets.

However, in real market investment, traders can't be limited to the two markets of gold and bitcoin. Therefore, our strategy can be slightly modified to add more investment markets to cope with the optimal trading strategy of more than two portfolios. For such problems, we can still improve on the basis of this model, introduce more variables on the basis of the prediction model, and still use the dynamic programming method to optimize multi-portfolios. In this way, we can get the optimal trading strategy of various portfolios, and promote the model more widely.

In addition, we can also improve the model itself. In the prediction model, we can consider comparing with other models and improving them to make the model have the best prediction ability. Short-term data can consider grey prediction model, and some machine learning algorithms can be used to solve prediction problems too, such as random forest and LSTH neural network. The prediction ability can also be optimized by adjusting the order of ARIMA model. In nonlinear programming, we can consider more indicators and better risk measurement indicators to optimize the equation.

8 Conclusion

In the current economic situation, investing in a single product carries huge risks and low returns. Combination investment of two products can effectively avoid risks and increase the level of returns.

In this project, we use the ARIMA model to predict future market changes through data analysis of the previous gold market and Bitcoin market. Combining the predicted data with the previous market performance to evaluate the risk, and under the premise of risk management and control, with the help of lagrange multiplier method, we propose our dynamic investment model to maximize the benefits of the gold and bitcoin portfolio investment. This method can be used not only in the combined investment of the gold market and the bitcoin market, but can also be extended to invest in a variety of other products.

Memorandum

To: The traders who are ready to invest in the market
Topic: Provide and introduce an optimal trading strategy for investing in gold and bitcoin.

Dear users:

We lend you a helping hand, because we can provide you with a reasonable daily optimal trading strategy, which can help you avoid risks and get the maximum return. Moreover, this strategy will no longer be limited to one kind of investment, but it is aimed at two kinds of common investments, gold and bitcoin. We can provide you with the best daily portfolio and return information only based only on price data up to that day. We appreciate the opportunity to introduce our strategy to you. Let me tell you about our solution in detail.

First of all, let me state our **strategy** and the **model** built in it. Our optimal strategy is mainly based on two models − ARIMA prediction model and nonlinear programming equations optimization model. The main idea is to forecast the market price first and then optimize the trading choice.

- **Model**

  The ARIMA model can be used to predict the future price of gold and bitcoin. The nonlinear programming can be used to optimize the portfolio and return. Here, the optimization equation is mainly based on the two indicators of risk and yield, which can truly reflect the change of return in the real market.

- **Strategy**

  Based on the above two models, we establish the daily optimal strategy. The optimal strategy is divided into two steps.

  In the first step, the ARIMA model is used to forecast the trading prices of gold and bitcoin three days after the current day, and the data from two days before the current day to six days after the third day are recorded to obtain two new time series.

  In the second step, using the newly obtained data, the volatility, risk, yield and real yield are calculated, which means we get the values needed in the next nonlinear programming of Model 2. Considering that gold is not open on non-working days every week, the problem is divided into two parts, both gold and bitcoin on the trading day and only bitcoin on the trading day. Using dynamic programming, the corresponding nonlinear programming is brought in for iteration to calculate the optimal allocation portfolio and the maximum return of the day.

- **Result**

  Secondly, we will show the results of the established model and strategy. We use Python to program, if the initial money is \$ 1000 on September 12^{th}, 2019, the maximum return will be obtained on September 10^{th}, 2021, in which the investment portfolio is the trader should put in gold of 0, bitcoin of $ 2980736.4980757544. In finally, there are $ 3346387.17 not in investment project, and the total assets is $ 6327123.66.

Finally, we do sensitive analysis, through the corresponding tests, it is concluded that the model results are sensitive to the change of transaction costs of bitcoin , but insensitive to the change of transaction costs of gold.

The above is the optimal trading strategy that we have suggested. We would like to help you make reasonable and diversified investments and get the maximum return. If you have any questions about our solution, please feel free to contact us. Our team are willing to solve any questions you have.

Yours sincerely,
Team: 2202124

References

[1] Xia Hansong. Research on the similarity measurement method of time series based on dynamic time regularity [D]. Chongqing University of Posts and Telecommunications, 2021.

[2] Anna M. Costello, The value of collateral in trade finance, Journal of Financial Economics, Volume 134, Issue 1, 2019, Pages 70-90.

[3] David Easley, Maureen O'Hara, Soumya Basu, From mining to markets: The evolution of bitcoin transaction fees, Journal of Financial Economics, Volume 134, Issue 1, 2019, Pages 91-109.

[4] Zhang Dingyuan. Establishment, evaluation and improvement of spot gold trading strategy [D]. xinjiang university of finance, 2015.

[5] Chen Ningjuan. Research on quantitative timing trading strategy based on neural network [D]. Shandong University, 2021.

[6] Zhou Baihui. Reinsurance portfolio problem and American option pricing problem under Lévy process [D]. Northeast Normal University, 2021.

[7] Huiying Wang, Hao Yongtao. Prediction algorithm of stock price trend based on technical index and random forest [J]. Modern Computer, 2021,27(27):43-47+52.

[8] Dynamic programming method of portfolio investment model [J]. journal of weifang university, 2012,12(06):20-21+47.

[9] Marten Hillebrand, Jan Wenzelburger, The impact of multiperiod planning horizons on portfolios and asset prices in a dynamic CAPM, Journal of Mathematical Economics, Volume 42, Issues 4–5, 2006, Pages 565-593.

[10] Chen Guohua, Liao Xiaolian, Wu Na. Dynamic programming solution of portfolio model [J]. statistics and decision, 2012(05):85-87.

[11] Han Yucai. Discussion on the calculation method of financial internal rate of return of investment projects [J]. Project cost management, 2017(01):67-71.

[12] Zhou Meiyan, Huang Ying, Yan Yulong, Zhou Jie, Bo Huang. Application of exponential smoothing method and ARIMA model in students' tuberculosis prediction [J]. Practical Preventive Medicine, 2022,29(01):18-22.

[13] Liu Zhongdian, Li Yanning. Prediction of epidemic development trend in novel coronavirus, Guangxi based on ARIMA model [J]. Journal of Guangxi Medical University, 2021,38 (12): 2367-2371.

Appendices

Appendix1 Code of three-dimensional graph

```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
plt.figure(dpi=300,figsize=(10,5))

df = pd.read_csv('fty.csv')
df.iloc[5,5] = (df.iloc[5,4]+df.iloc[5,6])/2

fig = plt.figure()
plt.axis('tight')
ax1 = plt.axes(projection='3d')

def get_value(x,y):
    zero = [0 for i in range(20)]
    Z = np.array([zero]*20)
    for i in range(20):
        for j in range(20):
            Z[i,j] = df.iloc[int(x[0,i]*100),int(y[j,0]*100)+1]
    return Z

xx = np.arange(0,0.2,0.01)
yy = np.arange(0,0.2,0.01)
X, Y = np.meshgrid(xx, yy)
Z = np.log(get_value(X,Y))
ax1.plot_surface(Y,X,Z,rstride = 1,cstride = 1,cmap = 'rainbow')
ax1.set_title("the Sensitivity Analysis of our Strategy to Transaction Costs")
ax1.set_xlabel("alpha bitcoin")
ax1.set_ylabel("alpha gold")
ax1.set_zlabel("the log of final asset")
plt.show()
```