Research on Optimal Investment Strategy Combination Based on ARIMA Model and mean-variance analysis -- Taking Gold and Bitcoin assets as examples

Benchen Liu *

School of Informatics, Xiamen University, Xiamen, China, 361005

* Corresponding Author Email: liubenchen18@163.com

Abstract. Gold and Bitcoin are popular trading products in today's trading market. In order to build a trading portfolio that maximizes returns, the prices of two trading products need to be predicted first. This article utilizes ARIMA to deal with the non-stationarity and predict the future prices of gold and bitcoin. In this article, the choice of parameters is ARIMA (4, 1, 4) for both bitcoin and gold. To find the best timing to sell and buy the two assets, the article first rate them with well-designed rating system by three important factors: Changes in value, Moving averages, and Bias. Then based on these factors, the model further linearly composes the indicator for risk and trend. By utilizing the information, the model gets with the main factor to make trading decisions.

Keywords: ARIMA Model; Mean-variance Analysis; Investment Strategy Portfolio.

1. Introduction

Traders buy and sell volatile assets to maximize their income. Recently, with the growing interest of the public in cryptocurrencies, gold and bitcoin become a feasible kind of portfolio. But what comes with interest is a risk. How can we do our best to increase our income? Previous researchers have studied the optimal proportion of gold in household assets [1] and the impact of different proportions. It also studies whether bitcoin should be included in the entire asset allocation [2][3], and tries to use some factors to trade bitcoin [4]. One of our approaches is to predict the prices of bitcoin and gold and make decisions, the article used ARIMA (Autoregressive integrated moving average). To predict future prices with existing data is a difficult problem because too many factors may influence the prices. The international situation, national policies, and even social media can have a considerable impact on the prices. Moreover, the data shows non-stationarity. To take as many factors as possible and predict accurately according to their inner laws, this article adopts ARIMA to predict the prices, which is proved to give satisfying results. Besides we design a comprehensive rating system to make decisions on trading for making decisions on trading Based on various factors drawn out from price data, this article computes the main factor by composing the trend indicator and risk indicator. The article then sets thresholds based on both our observations and findings in data. Finally, this article decides the final strategies Markowitz portfolio theory.

2. Build an optimal portfolio of gold and Bitcoin

2.1. ARIMA Predictor

ARIMA is widely used for forecasting time series data, which is a generalization of the ARMA (Autoregressive Moving Average) model. ARMA requires the mean function of the data to be stationary because history data of stationary series can be used to predict the future. But time series are not always stationary, so ARIMA takes an initial differencing step to eliminate the non-stationarity of the mean function. In this case, the article finds the first-order difference is stationary, so the model takes first-order difference. Figure 1 shows the procedure this article takes to build the ARIMA model.

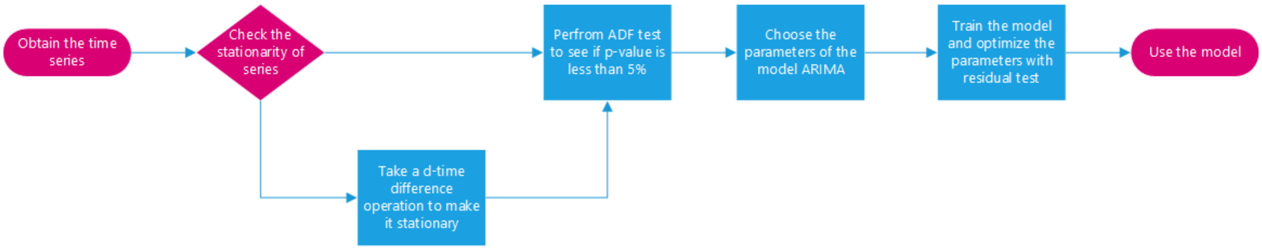


Figure 1. Procedure to Build ARIMA Model

To apply the model, this article has to test the stationarity firstly.

2.1.1 Augmented Dickey-Fuller Test

An augmented Dickey–Fuller test (ADF) is a kind of hypothesis test. The null hypothesis is H_0: A unit root is present in a time series sample. Subject to the χ^2 distribution, the article accepts it when the observed value of the waviness of the series, P, also known as the p-value, satisfies

$$P \geq 0.05 \tag{1}$$

It can test whether the data is stationary or trend-stationary. If the unit root is present in the series, the time series is non-stationary. In the raw data, we found:

$$P_{0,bitcoin} = 0.8420802907198858 > 0.05 \tag{2}$$

$$P_{0,gold} = 0.9042384812941653 > 0.05 \tag{3}$$

So, this article accepts H_0. The raw time series is non-stationary. And in the first-order difference, we found:

$$P_{1,bitcoin} = 9.276161468079189e-13 < 0.05 \tag{4}$$

$$P_{1,gold} = 9.269711421536524e-13 < 0.05 \tag{5}$$

The first-order difference is stationary, So, we plot them in Figure 2.

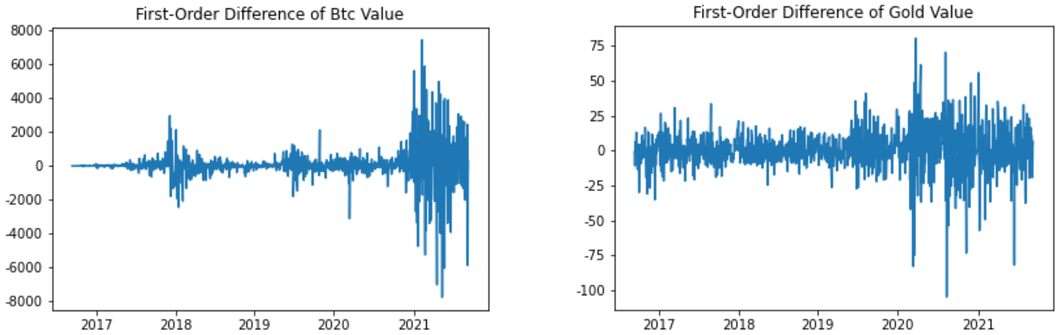


Figure 2. The 1st Order Difference of Gold and Bitcoin

This article also can see from the plot that it is stationary. Therefore, this article uses it for our ARIMA model. The parameter d of ARIMA (p, d, q) is d = 1.

2.1.2 Finite Difference

The difference of a function y is defined as:

$$\Delta^k y_t = \Delta(\Delta^{k-1} y_t) = \Delta^{k-1} y_{t+1} - \Delta^{k-1} y_t = \Sigma_{i=0}^{k} (-1)^i C_k^i y_{t+k-1} \ (k = 1,2,3, \dots) \tag{6}$$

Specially, the first-order difference, which is commonly used in ARIMA model, is defined as

$$\Delta y_t = y_{t+1} - y_t \tag{7}$$

Finite difference is useful in processing non-stationary data.

2.1.3 Choice of Parameters

To select the best parameters of the ARIMA model [5], this article first relies on observation on PACF and ACF plots [6]. And then, if it does not fall in a range that makes sense, this article searches for possible parameters by grid searching the feasible domain.

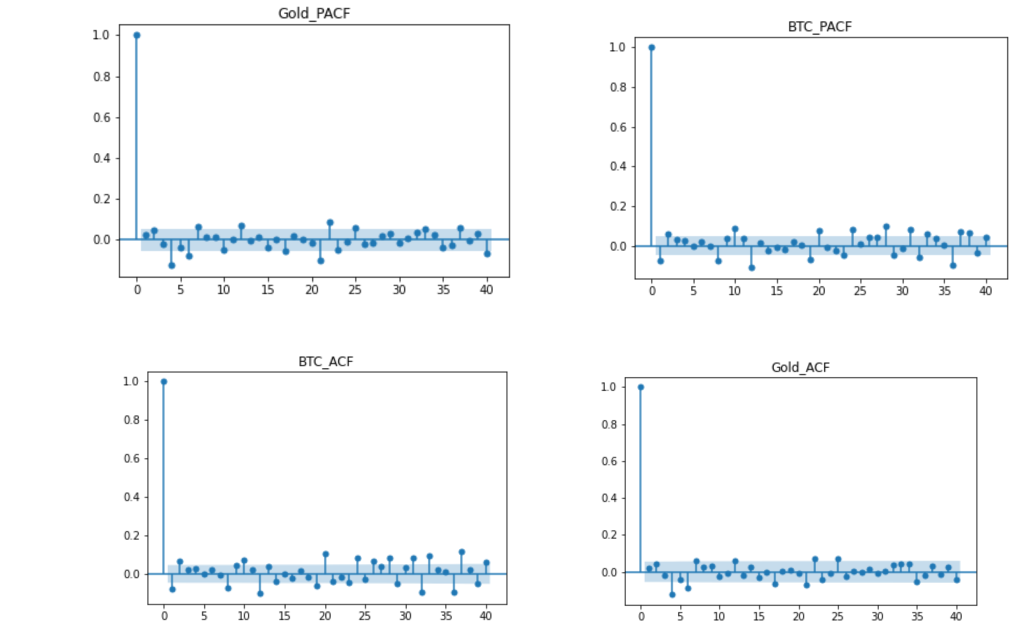


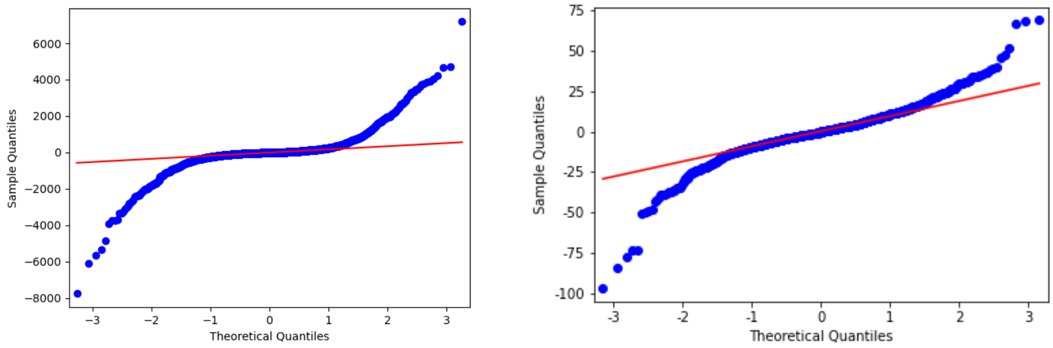
Figure 3. ACF and PACF Plot for Gold and Bitcoin

The parameter p of ARIMA (p, d, q) is the number of lag observations in the model, which is also known as the lag order. To assign p, the ACF plot in Figure 3 is inspected. This article finds that p=4 is the best choice for both bitcoin and gold.

The parameter q is the size of the moving average window, which is also called the order of moving window. To assign q, the PACF plot in Figure 3 is inspected. This article finds that q=4 is the best choice for both bitcoin and gold [7].

2.1.4 Residual Test and Durbin-Watson Test

To test whether the model has extracted all information in the data, this article inspects the residuals and test if they look like white noise. White noise residuals reveal that all information to predict has been extracted so that we can utilize them to predict the future. And what is left is only random perturbation and is of no use for our ARIMA predictor. To carry out the test, this article first plots the residual Q-Q plot and then carry out a Durbin-Watson (DW) test. The test results are as follows.



(a) Bitcoin Residual Q-Q Plot (b) Gold Residual Q-Q Plot

Figure 4. Residual Q-Q Plot for Gold and Bitcoin

Results are in Figure 4 that the scatters appear to be around a straight line, which indicates the parameters are good enough to predict the future.

DW test is commonly used for testing whether autocorrelation presents in the residuals from a regression analysis. The DW statistics has a null hypothesis.

$$H_0: \rho = 0 \tag{8}$$

And the test statistic is:

$$d = \frac{\Sigma_{t=2}^{T}(e_t - e_{t-1})^2}{\Sigma_{t=1}^{T} e_t^2} \tag{9}$$

This article calculates d in equation (9), which is the approximation of $2(1-\hat{\rho})$ where $\hat{\rho}$ is the sample autocorrelation of the residuals. So, d = 2 indicates no autocorrelation. In our case, we calculate the values in by equation (9), and find

$$d_{gold} = 2.0179203733750732 \tag{10}$$

$$d_{bitcoin} = 2.000089654788288 \tag{11}$$

which are very close to 2, indicating very low autocorrelation.

2.2. Trade Decision

2.2.1 Position Management

According to Markowitz portfolio theory [8], the final return of using a multi-asset portfolio investment can reach the weighted average of each asset's return, but the portfolio can significantly reduce systematic risk. The aim is to flatten the return curve and reduce retracements.

Markowitz utilizes mean-variance analysis in his model to measure the expected return and risk of each asset, and in the following, we will also use this model to quantify the expected return and risk of this portfolio of bitcoin and gold.

To measure the risk of the portfolio, this article calculates the standard deviation. The standard deviation of the portfolio is defined as

$$\sigma = \sqrt{w_T \cdot \Sigma \cdot w} \tag{12}$$

Where σ is the standard deviation of the portfolio, Σ is the covariance matrix of the portfolio, and vector w is the weight of the portfolio. So, to solve the given problem, the vector w is defined as

$$w = (w_{cash}, w_{bitcoin}, w_{gold}) \tag{13}$$

Where w_{cash}, $w_{bitcoin}$, and w_{gold} is the weight in the portfolio for cash, bitcoin and gold, respectively. The covariance matrix and the correlation matrix are shown in (14) and (15) respectively.

$$\begin{pmatrix} 0.675269 & 0.007781 \\ 0.007781 & 0.007153 \end{pmatrix} \tag{14}$$

$$\begin{pmatrix} 1.000000 & 0.111957 \\ 0.111957 & 1.000000 \end{pmatrix} \tag{15}$$

According to Markowitz portfolio theory and mean-variance model, the correlation matrix only represents the correlation between asset returns but does not reveal the volatility of the assets, while the covariance matrix can reveal this information. Therefore, this article will use the covariance matrix to measure the portfolio weights among assets in our portfolio risk assessment.

To measure the return of the portfolio, this article uses the annualized return of the portfolio [9].

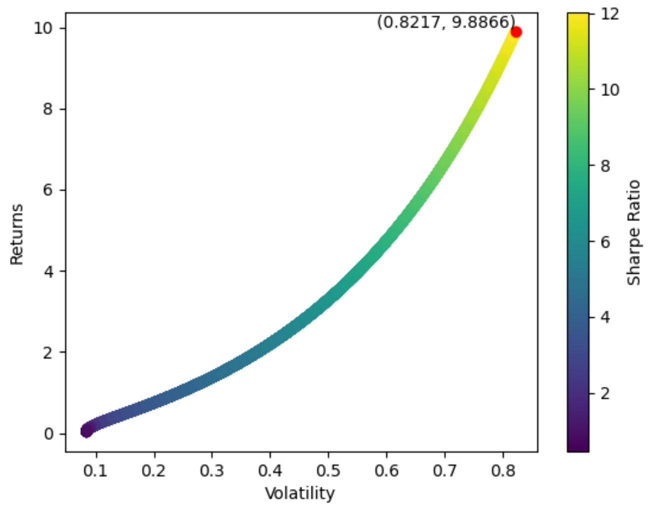


Figure 5. Simulation based on Monte-Carlo Method and Markowitz Theory

Using the mean-variance analysis, this article tries to find a portfolio that minimizes the risk and maximizes the return. In this model, we use the Monte-Carlo method to simulate the positions of bitcoin and gold. The result of the simulation is plotted in Figure 5.

It can be concluded from the figure that if the trader is looking for a higher Sharpe Ratio, we should look for the point with the largest value of the vertical coordinate over the horizontal coordinate, or on the top-left area of the figure.

Extracting the portfolio data behind this point, this article finds that most portfolios with a Sharpe Ratio of 10 or higher have a portfolio share of 95% or more in bitcoin.

Therefore, this article concludes to invest bitcoin as much as possible in the whole portfolio. When both gold and bitcoin are on an upward trend, traders are supposed to buy more bitcoin for their portfolio to achieve a high expected return.

2.2.2 Rating Gold and Bitcoin

Based on the closing price every day, this article first calculates the required normalized factors using including the change today, and the average change of 15 days and the bias, for both bitcoin and gold. Then the model predicts the price the next day with ARIMA model.

Then, the model calculates the indicators with the formula below for bitcoin and gold.

$$\text{Trend indicator} = \text{TI} = 0.7 \times 30\ \text{day's averagechange} + 0.3 \times 20\ \text{day's average bias} \quad (16)$$

$$\text{Risk indicator} = \text{RI} = 0.4 \times \text{Normalized 15 day's average change} + 0.6 \times \text{bias} \quad (17)$$

Then this article composes TI and RI linearly with ratios chosen according to Markowitz portfolio theory and then normalize it to get our main factor (MF) for bitcoin and gold.

This article sets following thresholds for selling or buying bitcoin and gold in Table 1.

Table 1. Trading Thresholds

	Bitcoin	Gold
Buy (BS)	0.775	0.57
Sell (SS)	0.55	0.3

Then, the model trade bitcoin and gold subject to the following rules which are shown in Table 2.

Table 2. Basic Trade Rules

	$MF_{Bitcoin} \geq BS_{Bitcoin}$	Others	$MF_{Bitcoin} \leq SS_{Bitcoin}$
$MF_{gold} > BS_{gold}$	**Further Judgement Required**	Buy Gold	Sell Bitcoin and Buy Gold
Other	Bug Bitcoin	Do Nothing	Sell Bitcoin
$MF_{gold} < SS_{gold}$	Sell Gold and Buy Bitcoin	Sell Gold	Sell both

When further judgment is required, we come up with the following factor F whose formula is shown in equation (18).

$$F = 2.5 \times (MF_{bitcoin} - BS_{bitcoin}) - (MF_{gold} - BS_{gold}) \quad (18)$$

Then, the model recommends buying gold when F<0 and buying bitcoin when F > 0. In the above formula, the ratio 2.5 comes from the weights that reach the maximum in return that we find in the previous sub-section Position Management.

2.3. Result

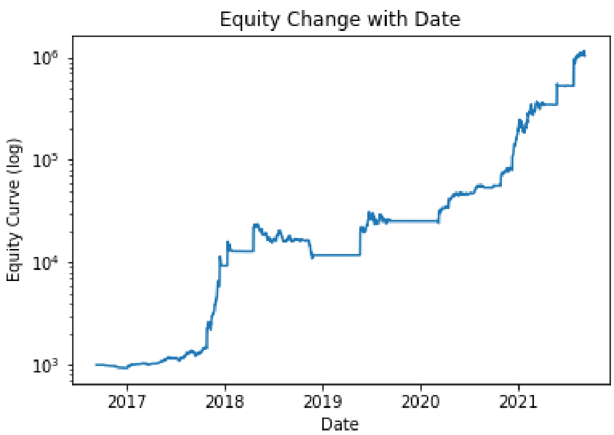


Figure 6. Equity Change Under Our Strategy

As is plotted in Figure 6, our equity change rises a lot with time going by. From the plot we can know the following facts.

Our strategies work well to maximize income. As is seen in the plot, the equity rises in the order of magnitude from 103 to 106, which is a considerable achievement.

Our strategy has a distinct feature that it has very low retracement. As is seen in the plot, only there are only a few retracements, all of which have small value changes. It reveals that we have a good performance in dealing with risks.

Table 3. Detailed Performance Data

Performance Indicator	Measurement
Accumulated Equity	1.03743×10^6
Annualized Return	1469.58%
Maximum Retracement	-53.90%
Annualized Return/Retracement Ratio	27.77
Sharpe Ratio [10]	1.1358

The detailed performance data is shown in Table 3. From the table we can have a quantitative view on the brilliant performance of the model.

3. Conclusions

First, the article plotted the price data and found that they show non-stationarity in the sense of mean. This article, therefore, utilizes ARIMA (Autoregressive Integrated Moving Average) to deal with the non-stationarity and predict the future prices of gold and bitcoin. Our choice of parameters is ARIMA (4,1,4) for both bitcoin and gold. The ARIMA (4,1,4)-based predictor shows the good capability to accurately predict the price of the future.

To find the best timing to sell and buy the two assets, this article first rates them with our well-designed rating system by three important factors: Changes in value, Moving averages, and Bias. Then from the factors the model further linearly composes our indicator for risk and trend. By linearly composing these two indicators, our article gets our main factor, which can give us enough information to make trading decisions. By utilizing the information, the model gets with our main factor to make trading decisions. And this article furthermore set thresholds based on the observations and experiments in the data.

References

[1] Lucey, B.M. et al. (2019) "What is the optimal weight for gold in a portfolio?," Annals of Operations Research, 297(1-2), pp. 277–291. Available at: https://doi.org/10.1007/s10479-019-03496-5.

[2] Platanakis, E. and Urquhart, A. (2020) "Should investors include Bitcoin in their portfolios? A portfolio theory approach," The British Accounting Review, 52(4), p. 100837. Available at: https://doi.org/10.1016/j.bar.2019.100837.

[3] Kang, S.H. et al. (2019) "Bitcoin as hedge or safe haven: Evidence from stock, currency, Bond and derivatives markets," Computational Economics, 56(2), pp. 529–545. Available at: https://doi.org/10.1007/s10614-019-09935-6.

[4] Fang, F. et al. (2022) "Cryptocurrency trading: A comprehensive survey," Financial Innovation, 8(1). Available at: https://doi.org/10.1186/s40854-021-00321-6.

[5] Pierre Rostan, Alexandra Rostan, Mohammad Nurunnabi. Options trading strategy based on ARIMA forecasting[J]. PSU Research Review,2020.

[6] Vo, N. and Ślepaczuk, R. (2022) "Applying hybrid arima-SGARCH in algorithmic investment strategies on S&P500 index," Entropy, 24(2), p. 158. Available at: https://doi.org/10.3390/e24020158.

[7] Xiao, D. and Su, J. (2022) "Research on stock price time series prediction based on Deep Learning and autoregressive integrated moving average," Scientific Programming, 2022, pp. 1–12. Available at: https://doi.org/10.1155/2022/4758698.

[8] Sun, Lipo. An empirical study of Makowitz's portfolio theory based on Python[J]. TimesFinance,2020(25):46-47+50.

[9] Naccarato, A., Pierini, A. and Ferraro, G. (2019) "Markowitz portfolio optimization through pairs trading cointegrated strategy in long-term investment," Annals of Operations Research, 299(1-2), pp. 81–99. Available at: https://doi.org/10.1007/s10479-019-03225-y.

[10] Lin Hongmei, Du Jinyan, Zhang Shaodong. Sharpe Ratio: Estimation Method, Ap- plicability and Empirical Analysis [J]. Journal of Statistics, 2021, 2(06): 73-88. DOI: 10.19820/j.cnki.issn2096-7411.2021. 06.006.