

Making Money: a LR+ARIMA and Multi-Index Based Investment Model

Yuting Liu

Beijing Sport University, Beijing, China

Corresponding author:1228203005@qq.com

Abstract. In order to achieve the goal of maximizing returns, market traders tend to buy and sell assets with high volatility. Now, in this paper, traders will buy and sell only two assets, namely gold, and bitcoin, based on their requirements, using past daily prices to determine the trading prices to determine whether they should buy, continue to hold, or sell the assets in their portfolio on that date. Determine a portfolio consisting of cash, gold, and bitcoin that is based on an investment model for a five-year trading period starting at \$1,000 on September 11, 2016, and ending on September 10, 2021. During this period, they will follow a trading strategy when appropriate. (The gold market is only open on weekdays and the bitcoin market is open daily with commission fees for each trade).

Keywords: Linear Regression; ARIMA; Bollinger Bands; Rate of Change.

1. Introduction

There are certain risks in business investment. Generally, the higher the rate of return on a risky investment, the greater the risk. How to pursue maximum returns with minimum risk, has been a question for all investors[1]. With the development of the Internet, London Gold and Bitcoin were born. London Gold is not a name for gold, but rather a name for a way of trading gold[2].

The London gold market is not an actual trading place, but an invisible online market that connected many major gold dealers[3]. Many people on the planet have joined this gold trade so that the volume of gold traded is enormous. In 2016, it is measured in trillions of dollars per day. Therefore, no consortium or institution can artificially manipulate such a huge market, it is completely dependent on market self-regulation[4]. The gold market doesn't have bankers. There is no bull market or bear market in the gold market[5]. Whether the gold price is a big up or down, it is an opportunity for investors. The gold market is all about the market[6].

Unlike most currencies, Bitcoin does not rely on a specific monetary institution to issue it, it is based on a specific algorithm, which generated by a large number of calculations, and it is the first distributed virtual currency with an entire network of users and it does not have a central bank. Bitcoins can be managed from any computer with Internet access[7]. Anyone, regardless of location, can buy, sell, or receive bitcoins. Bitcoin has gone from less than \$0.01 apiece at its inception to \$20,000 apiece in 2017, then its price has plunged by 80%. In a few years, some people have become rich overnight, but some people become destitute[8]. Even so, there are still many investors who are willing to take huge risks to achieve huge benefits.

On one side is London Gold, which has stable fluctuations, and on the other side is Bitcoin, which has a lot of ups and downs, how to balance the two investments so that we can maximize the benefits with minimal risk is the subject of this article. Model assumptions and notation

2. Assumptions

The investor is a rational person who is willing to take a certain amount of risk in order to maximize total return.

In this model, there is no limit on the number of purchases per transaction, including a lower limit on purchases.

Ignore the phenomenon of dollar depreciation or appreciation.

The gold market is relatively stable for a short period of time, i.e., it can maintain a basic balance within three days

There is no trading process failure. In other words, there are no technical failures and other objective factors that make trading fail.

3. Model construction and solving

3.1 Cold Boot

Suppose we begin at t_0 and want to forecast $t_0 + 1$. Because there isn't any additional information at this time, we can only assume that $t_0 + 1$ remains unchanged, that is.

$$\hat{p}_{t_0+1}^{(i)} = p_{t_0}^{(i)} \quad (1)$$

It is difficult to do time series analysis due to too few data points known. Hence we use the linear regression model to forecast these t . More explicitly, the method is to fit.

$$(t, p_t^{(i)}) \quad t \in (t_0 + 1, \text{thre}^{(i)}) \quad (2)$$

into a straight line

$$y = b_0 + b_1 t \quad (3)$$

Satisfying

$$\min_{b_0, b_1} \left\| \begin{pmatrix} 1 & t_1 \\ \mathbf{M} & \mathbf{M} \\ 1 & t_n \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} - \begin{pmatrix} y_1 \\ \mathbf{M} \\ y_n \end{pmatrix} \right\|_2 = \min_b \|Ab - Y\|_2 \quad (4)$$

By solving above equations, get

$$b_1 = \frac{\sum_{i=1}^n t_i y_i - n \bar{t} \bar{y}}{\sum_{i=1}^n t_i^2 - n (\bar{t})^2} \quad b_0 = \bar{y} - b_1 \bar{t} \quad (5)$$

3.2 ARIMA

According to our hypothesis, the transaction price has a high degree of autocorrelation and stable noise. Therefore, under the condition of sufficient data, adopt the ARMA(p, q) model in the following form:

$$p_t = \phi_0 + \sum_{n=1}^p \phi_n p_{t-n} + \varepsilon_t - \sum_{n=1}^q \theta_n \varepsilon_{t-n} \quad (6)$$

Where

$$E(\varepsilon_t) = 0, \text{var}(\varepsilon_t) = \sigma_\varepsilon^2, \text{cov}(\varepsilon_t, \varepsilon_s) = 0 (s \neq t), \text{cov}(p_s, \varepsilon_t) = 0 \quad (\forall s < t) \quad (7)$$

Besides, to stationarize the series, substitute $W_t = \nabla^d p_t$ into p_t in the above formula:

$$W_t = \phi_0 + \sum_{n=1}^p \phi_n W_{t-n} + \varepsilon_t - \sum_{n=1}^q \theta_n \varepsilon_{t-n} \quad (8)$$

To determine the best p, q , we enumerate all possible (p, q) in $([1, 8] \times [1, 8]) \cap (Z \times Z)$, and calculate the corresponding AIC of each:

$$\text{AIC} = 2k + n \ln \left(\frac{\text{RSS}}{n} \right) \quad (9)$$

where n is the observation number and RSS is the sum of residual squares. (p, q) with the smallest AIC is taken as the final parameter.

To determine the best order d , we conducted ADF tests at $d = 0, 1, 2$ in order to select the d which inducts to most stable sequence.

3.3 MIBDM Decision-making Model

In this section, propose the MIBDM (Multi-index Based Decision-making) model to evaluate the buying/selling intention of bitcoin or gold at a certain day[9]. This model comprehensively examines the Bollinger Band index, ROC index, bitcoin/gold linkage effect, trading frequency, and future forecast, and so on, so as to get the best trading strategy.

3.3.1 Indicators Chosen

3.3.1.1 Confidence

Through the LA+ARIMA model, we can make predictions of the future trend, and the prediction in the next three days is quite accurate. Therefore, the forecast value in the next three days can be used as an important reference for “buy low and sell high”

Let

$$\hat{\delta}^{(i)} = (\hat{p}_{t_0+1}, \hat{p}_{t_0+2}, \hat{p}_{t_0+3}) \quad \delta^{(i)} = (p_{t_0}, p_{t_0}, p_{t_0}) \quad (10)$$

Hence

$$c^{(i)} = \frac{1}{\pi} \tan^{-1} \left(\left\| \hat{\delta}^{(i)} - \delta^{(i)} \right\|_2 \right) \quad (11)$$

represents our confidence of price rise or decline in the near future.

Also, let

$$\delta_{\pm} = \text{sgn} \left(E_3 \left(\hat{\delta}^{(i)} - \delta^{(i)} \right)^T \right) \quad (12)$$

Where E_3 denotes the 3×3 unit matrix and $\text{sgn}(\cdot)$ the sign function. It can represent the overall trend in the near future. $\delta_{\pm} = 1, 0, -1$ represents the price rising, staying flat and falling respectively.

3.3.1.2 Bollinger Bands Factor

Bollinger Bands are a type of price envelope developed by John Bollinger. They are envelopes drawn at a standard deviation level above and below a simple moving average of the price.

The simple moving average of the price here is.

$$B_t^{(n)} = \frac{1}{n} \sum_{i=0}^{n-1} p_{t-i} \quad (13)$$

and the upper band is

$$\bar{B}_t^0 = B_t^{(n)} + 1.5\sigma_{n,t}^{(i)} \quad (14)$$

the lower band is

$$\underline{B}_t^{(n)} = B_t^{(n)} - 1.5\sigma_{n,t}^{(i)} \quad (15)$$

$$\sigma_t^{(n)} = \left[\frac{1}{n-1} \sum_{i=0}^{n-t} (p_{t-i} - \bar{p}_t^{(n)})^2 \right]^{\frac{1}{2}} \quad (16)$$

Prices have a tendency to bounce within the bands' envelope, touching one band then moving to the other band. These swings can be used to predict the trend. If a price bounces off the lower band and then crosses above the moving average, there may be a rise in prices. If the opposite happens, there may be a decline in prices[10]. To summary, we define Bollinger bands factor by

$$b_t^{(i)} = \begin{cases} 1, & \text{uptrend} \\ 0, & \text{balance} \\ -1, & \text{downtrend} \end{cases} \quad (17)$$

3.3.1.3 ROC Factor

The rate of change (ROC) is the speed at which the price changes over a specific period. It's defined by.

$$ROC = \frac{p_t - p_{t-n}}{p_{t-n}} \quad (18)$$

A rising ROC above zero typically confirms an uptrend while a falling ROC below zero indicates a downtrend[2]. Thus, the ROC factor is defined by:

$$r_t^{(i)} = \begin{cases} 1, & \text{uptrend} \\ 0, & \text{balance} \\ -1, & \text{downtrend} \end{cases} \quad (19)$$

3.3.1.4 Correlation Factor Between Bitcoin and Gold Price

Dr. He[3] suggested that the similarities and differences between bitcoin and gold would promote the generation of asset overflow effect and substitution effect, and then make them show correlation. Moreover, through analysis, he thought that the substitution effect of gold on bitcoin was more obvious, and the fluctuating overflow effect of bitcoin on gold was more obvious. The asset overflow effect caused by similarity shows a positive linkage correlation; The asset substitution effect caused by differences is a negative linkage correlation.

To take this into account, define the correlation factor between bitcoin and gold price by:

$$l_t^{(i)} = \begin{cases} 1.1, & \text{positive linkage} \\ 0.9, & \text{negative linkage} \end{cases} \quad (20)$$

Generation mechanism of overflow effect and substitution effect (Source: Translated from Time-varying relationship between bitcoin and gold: the perspective of asset allocation) is shown in Figure 1.

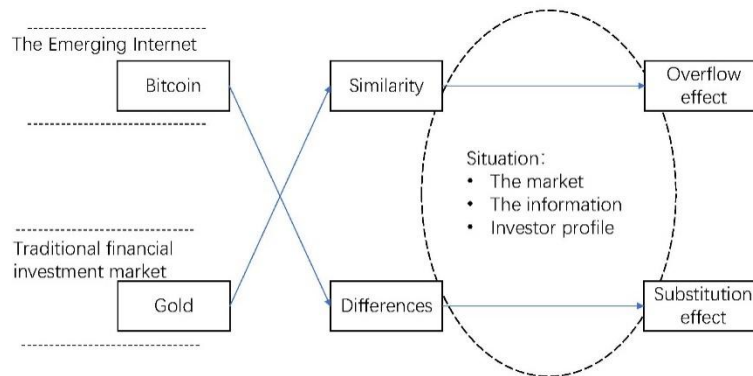


Figure 1 Generation mechanism of overflow effect and substitution effect (Source: Translated from Time-varying relationship between bitcoin and gold: the perspective of asset allocation)

3.3.1.5 Transaction Frequency Factor

Because transactions require transaction costs, high-frequency transactions may cause losses when the profits are low, so the frequency of transactions must be limited. On the other hand, it is necessary to trade more during the arbitrage period. Therefore, we use the peak value of the account in the past 30 days to determine whether the arbitrage period has entered. If the ratio of the consumed transaction costs to the peak value is higher than a certain percentage $\gamma^{(i)}$ in the past 30 days, it is considered that the transaction cost paid is too much and should be limited.

Therefore, in order to control the frequency of transactions, the transaction frequency coefficient is defined as follows:

$$f_t^{(i)} = \begin{cases} 1, & \alpha_t^{(i)} / \max_{0 \leq i \leq 30} \{w_{t-i}\} < \gamma^{(i)} \\ 0.9, & \alpha_t^{(i)} / \max_{0 \leq i \leq 30} \{w_{t-i}\} \geq \gamma^{(i)} \end{cases} \quad (21)$$

3.3.2 Strategies

3.3.2.1 Sell or Buy?

After knowing the future trend, Bollinger bands and the ROC factor, the trader can calculate her intention upon these three indicators by

$$will^{(i)} = \delta_{\pm}^{(i)} + b^{(i)} + r^{(i)} \quad (22)$$

3.3.2.2 Weights

When this is the case, it shows that both gold and bitcoin need to buy, but due to the limited available funds, we must weigh their weights. Our basic idea is to comprehensively evaluate the transaction weights by calculating the expected returns of both of them and combining the previous indicators.

Reference[4] puts forward that the distribution of the simple moving average can be regarded as a normal distribution with known parameters. Therefore, $P(\xi_{upper}^{(i)}) = P(\xi_{bottom}^{(i)}) = \Phi(1.5) - \Phi(0)$.

Case 1:

Non-risk level

$$L = 1 - \frac{P_t^{(i)} - \xi_{bottom}^{(i)}}{4\sigma} \quad (23)$$

Expected loss value (risk)

$$\int_{\xi_{bonaum}^{(0)}}^{P_t^{(i)}} F(x)dx \quad (24)$$

Expected profit value

$$\int_{P_t^{(i)}}^{\xi_{sipper}^{(i)}} F(x)dx \quad (25)$$

Overall profit value

$$W_t^{(i)} = \int_{P_t^{(i)}}^{\xi_{inper}^{(i)}} F(x)dx - \int_{\xi_{bataum}^{(i)}}^{P_t^{(i)}} F(x)dx \quad (26)$$

Therefore, from the perspective of expected return, the best weight is

$$\frac{\omega^{(1)}}{\omega^{(2)}} = \frac{L^{(1)} \cdot W^{(1)}}{L^{(2)} \cdot W^{(2)}} \quad (27)$$

Case 2:

In this case, there is only expected profit value

$$W^{(i)} = \int_{\xi_{batram}^{(i)}}^{\xi_{suppor}^{(in)}} F(x)dx \quad (28)$$

The best weight is

$$\frac{\omega^{(1)}}{\omega^{(2)}} = \frac{W^{(1)}}{W^{(2)}} \quad (29)$$

Case 3:

Similar to case 1, the optimal weight ratio is still

$$\frac{\omega^{(1)}}{\omega^{(2)}} = \frac{L^{(1)} \cdot W^{(1)}}{L^{(2)} \cdot W^{(2)}} \quad (30)$$

Combining other useful indicators, get:

$$\frac{w^{(1)}}{w^{(2)}} = \frac{\omega^{(1)} l^{(1)} f^{(1)} c^{(1)}}{\omega^{(2)} l^{(12)} f^{(2)} c^{(2)}} \quad (31)$$

$will^{(i)} > 0$ for an i and $will^{(i)} < 0$ for another

When this is the case, it shows that one product will show a downward trend in the future, while another product will show a good trend in the future. Therefore, in order to better arbitrage and stop loss in time, we will sell all the falling products and buy as many rising products as possible.

$will^{(1)} < 0$ and $will^{(2)} < 0$

When this is the case, it shows that both two product will show a downward trend in the future, we need to sell a certain amount of goods in order to stop losses in time and improve the trading

freedom. However, in order to make profits in the future, we can't sell all the products. The proportion of products sold is defined as follows

$$\text{sell}_{\%}^{(i)} = \begin{cases} 100\%, & |will^{(i)}| = 3 \\ 80\%, & |will^{(i)}| = 2 \\ 50\%, & |will^{(i)}| = 1 \end{cases} \quad (32)$$

3.3.3 A Remark on Gold Trading

One important hypothesis in our model is that we assume the gold price is stable over a short period. This hypothesis significantly simplifies our model, at least in the following two ways:

1. In the data washing stage, when encounters blank days, we just use the data of the day before to fill it. Since the gold price is stable over a short period, it will not significantly influence the forecast results.

2. Though gold is only traded on days the market is open, we need still not distinguish the days that the gold market is closed from normal days. If ever "made a transaction" on days which the gold market was closed, it simply means would make it actually happen on the next normal day. Because of the stability of the gold prices, it will not significantly influence the results.

3.4 Results

After experimenting, it take these parameters: $n = 20$, $\gamma = 0.5$

Then implement our model and get a final income of \$1237034.252456963

Below are some of our transaction records as shown in Figure 2.

09/21/16 Buy 1.6370427412215218 Bitcoin at a price of 598.88. Money: 0
09/24/16 Sell 0.8185213706107609 Bitcoin at a price of 603.88. Money: 484.4029115787378
09/25/16 Sell 0.40926068530538046 Bitcoin at a price of 601.74. Money: 725.7460658588843
09/26/16 Sell 0.20463034265269023 Bitcoin at a price of 598.98. Money: 845.8641588481505
09/27/16 Sell 0.20463034265269023 Bitcoin at a price of 605.96. Money: 967.3820052332982
09/27/16 Buy 1.443562872008324 gold at a price of 1327.0. Money: 0.0
10/12/16 Sell 0.360890718002081 gold at a price of 1256.5. Money: 448.92459529791864
...
08/23/21 Sell 16.400852263356835 Bitcoin at a price of 49335.68. Money: 792964.2550124035
08/24/21 Sell 8.200426131678418 Bitcoin at a price of 49523.5. Money: 1190955.7824739362
08/25/21 Buy 12.227604770088883 Bitcoin at a price of 47744.58. Money: 595477.8912369681
08/25/21 Buy 329.6148434794273 gold at a price of 1788.7. Money: 297738.94561848405
08/26/21 Sell 20.4280309017673 Bitcoin at a price of 48972.09. Money: 1278134.2461057308
08/26/21 Buy 1581.4417568900915 gold at a price of 1786.6. Money: 0.0
08/27/21 Sell 436.5622946574513 gold at a price of 1798.5. Money: 777305.7140720119
08/28/21 Buy 1292.397288144336 gold at a price of 1798.5. Money: 0.0
08/31/21 Sell 864.4797914008936 gold at a price of 1814.85. Money: 1553212.1379296726
08/31/21 Buy 32.34762480985605 Bitcoin at a price of 47074.77. Money: 0.0
09/06/21 Sell 16.173812404928025 Bitcoin at a price of 51769.06. Money: 820557.0035230739
09/07/21 Sell 8.086906202464013 Bitcoin at a price of 52677.4. Money: 1238034.252456958

Figure 2 Transaction records (parts)

3.5 Evaluation on Forecast Model



Figure 3 Forecast series and actually series

In order to test the accuracy of our LR+ARIMA model, we compared it with the linear regression model. Since the curve of bitcoin is more prone to fluctuation, we take bitcoin trend prediction as an example. It can be clearly seen from the figure that LR model has a great prediction error, and the LR+ARIMA model we use can not only reduce the error, but also predict the ups and downs of big events. Forecast series and actually series is shown in Figure 3.

Moreover, we calculated RMSE and MAE of the two models, and these two indexes of LR+ARIMA model are significantly lower than those of LR model, which shows high accuracy.

3.6 Evaluation on Decision Mode

In order to compare the effect of decision-making, we compare MIBDM model with other common investment strategies. The strategies used for comparison are as follows:

Strategy 1 Buy \$1000 bitcoin on 9/12/2016 and sell all bitcoin on 9/11/2021

Strategy 2 Buy \$1000 gold on 9/12/2016 and sell all gold on 9/11/2021

Strategy 3 A fixed investment of \$200 a year in bitcoin (On January 1st every year)

The result is as follows: It can be seen that the trading strategy derived from MIBDM model is the best.

4. Conclusion

Based on the LR+ARIMA prediction model and the MIBDM decision model, and restarting the model at the appropriate time, this paper draw conclusions as follows after modeling analysis:

- According to our investment strategy, the return is 123,703.43% after 5 years. Investing scientifically can bring us unexpected returns, but it is also accompanied by certain risks.
- High returns mean high risk, and low risk can bring limited returns, which can be seen visually from the price changes of bitcoin and gold.
- If want a higher rate of return, our financial distribution has to be more focused on investing in bitcoin, which also means taking a higher risk.

References

- [1] Pan Qingbin. Research on programmed trading strategy based on improving BOLL index and fund management[D]. Fujian: Huaqiao University, 2018
- [2] Jin Zining. Legal regulation of uncertainty in risk assessment[J]. China Law Review. 2022(02)
- [3] Yu Dongyue. Procedural trading strategy based on technical index combination[J]. Jingchu University of Technology Review. 2019(06)
- [4] Lin Yong. Financial time series analysis and prediction based on deep learning and network big data[D]. Chengdu: Sichuan Normal University, 2022
- [5] Liang Ying. Application of functional time series analysis method in high frequency stock price prediction[D]. Xinjiang University, 2021
- [6] Chen Yuan. Analysis on the value evaluation of technical intangible assets based on income method[J]. Communication Of Finance And Accounting. 2019(35)
- [7] Wang Peidong. Analysis and forecast of stock price based on multiple linear regression[J]. Science & Technology Economy Market. 2020(01)
- [8] Fan Yuhan. An empirical study on the influencing factors of stock returns based on multiple linear regression model[J]. China Circulation Economy. 2019(06)
- [9] Yu Hanjun, Liu Cheng. Using multiple linear regression model to analyze macro factors affecting stock price[J]. Times Finance. 2017(23)
- [10] He Fei. Time-varying Relationship Between Bitcoin And Gold[D]. 2017