MSDM5058 Information Science
Computational Project II: Portfolio Management with Prediction Tools

Zhang Mingtao

1 Data Preprocessing

We choose stocks MTR(0066.HK) and Transport International (0062.HK), download their daily closing prices from Yahoo Finance which start from 2008-01-01 and end at 2024-05-03.

We choose MTR(0066.HK) to be the risker stock, denoted as S1, and Transport International (0062.HK) to be the safer stock as S2.

1.1 Update X(t) from the daily return

Two X(t) (daily-return) plots:

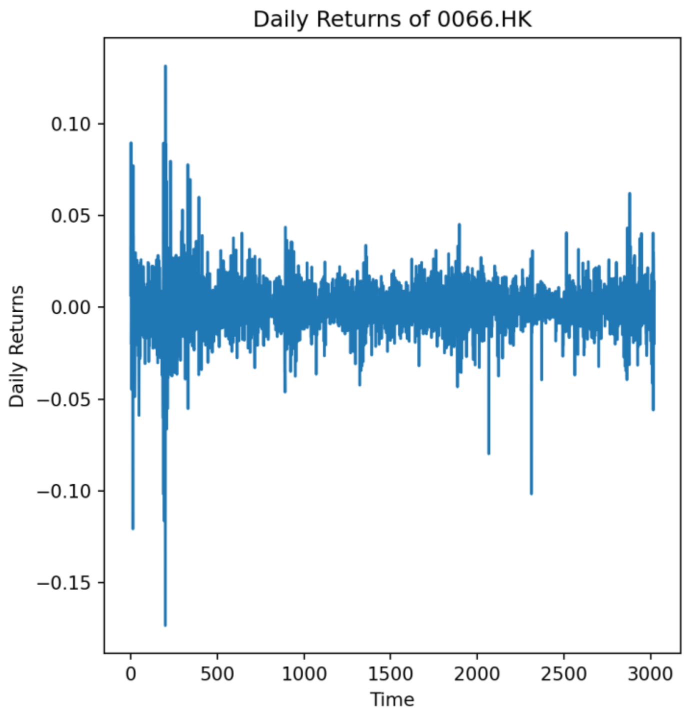

Fig. 1.1.1. Daily returns of 0066.HK

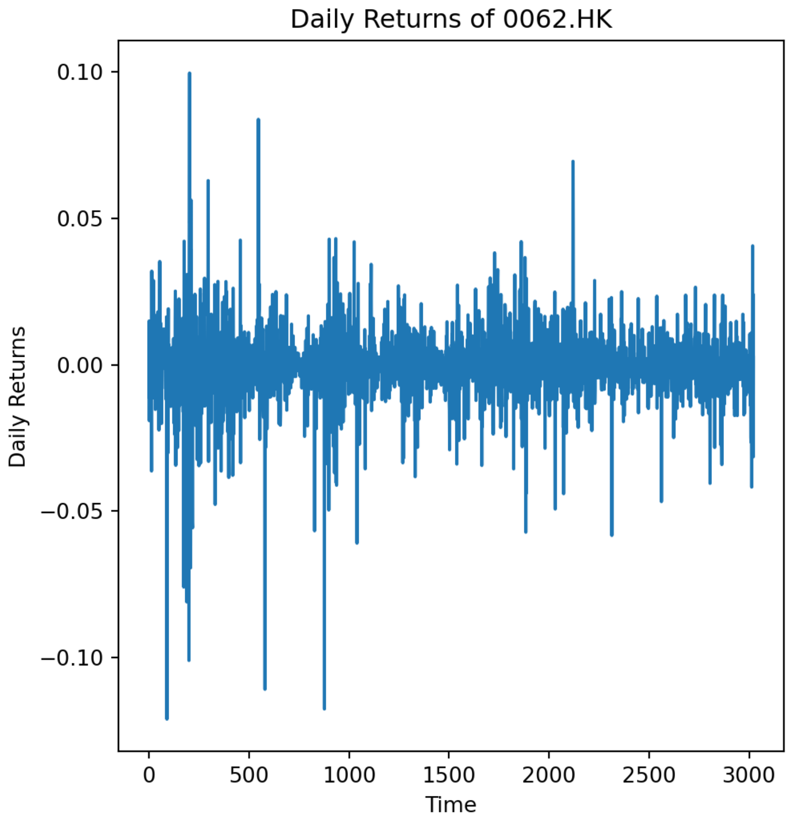

Fig. 1.1.2. Daily returns of 0062.HK

1.2 Plot S(t)

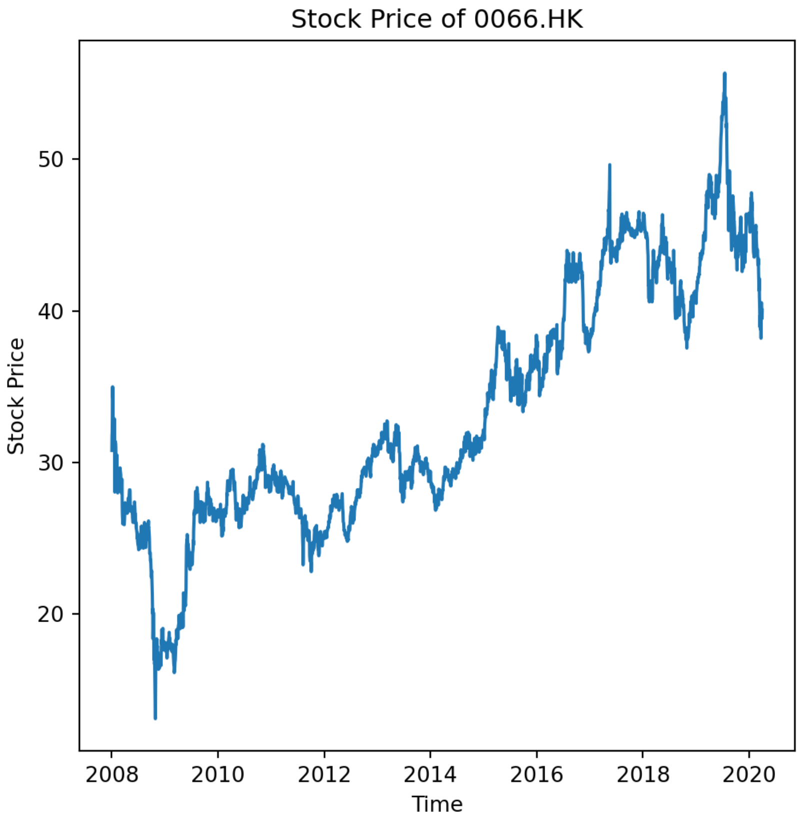

Fig. 1.2.1. Stock price of 0066.HK

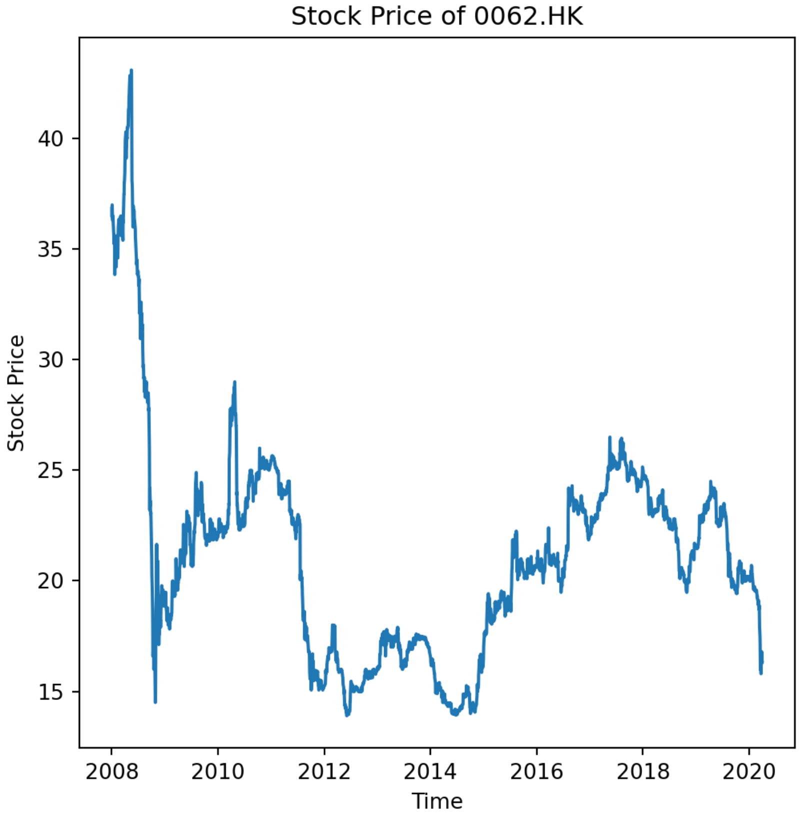

Fig. 1.2.2. Stock price of 0062.HK

2 Mean-Variance Analysis

Let us first construct the stocks' minimum-risk portfolio:

$$S_0 = p_0 S_1 + (1 - p_0) S_2$$

On one hand, we may infer $p_0(t)$ from all the data that we have observed so far, on the other hand, since the relevance of old data should decay, we may limit ourselves and infer it only from the data on the h most recent days.

2.1 Plot $p_0(t, h)$ for h = 30, 100, 300, and ∞

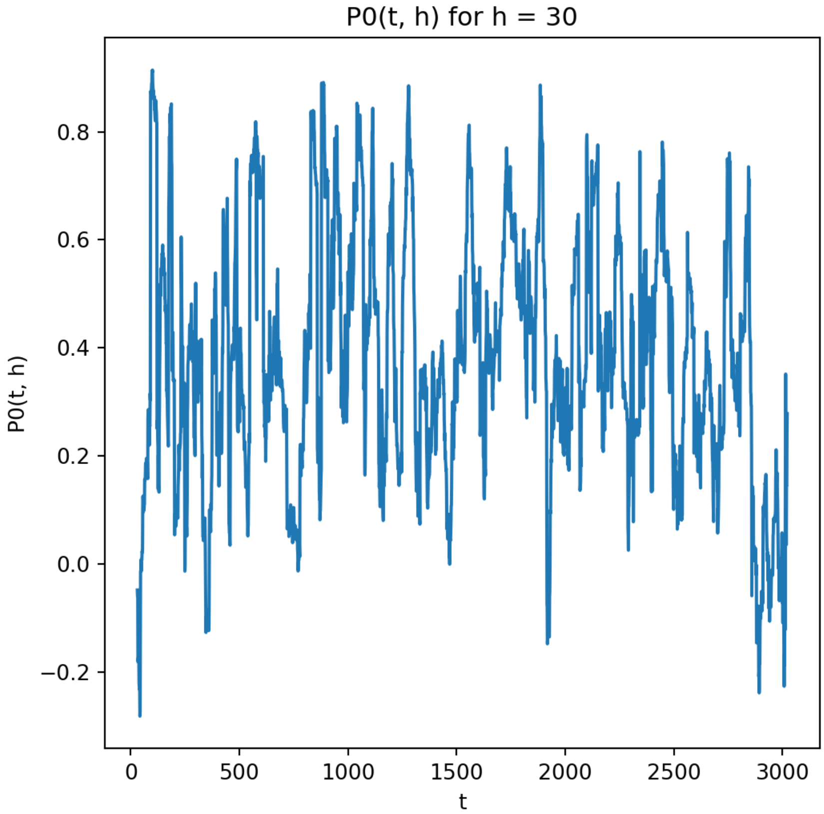

Fig. 2.1.1. $p_0(t, h)$ for h = 30

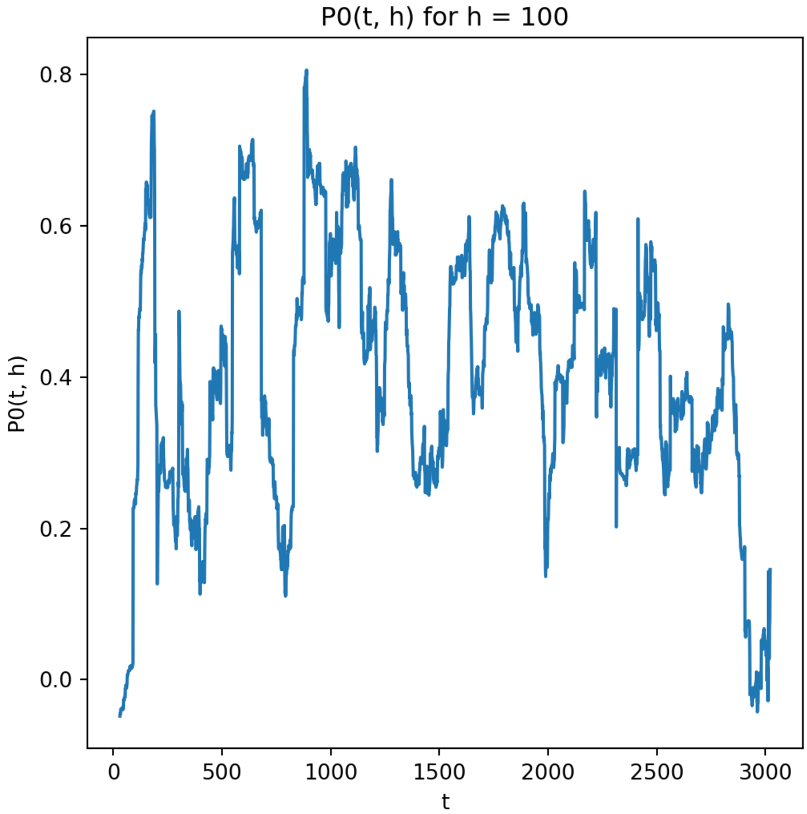

Fig. 2.1.2. $p_0(t,h)$ for h = 100

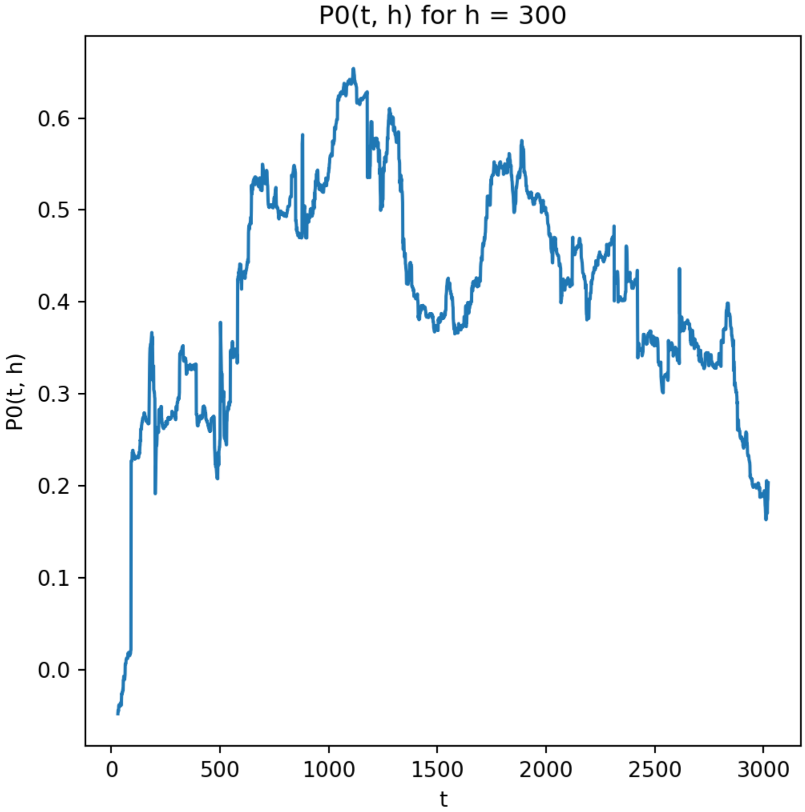

Fig. 2.1.3. $p_0(t,h)$ for h = 300

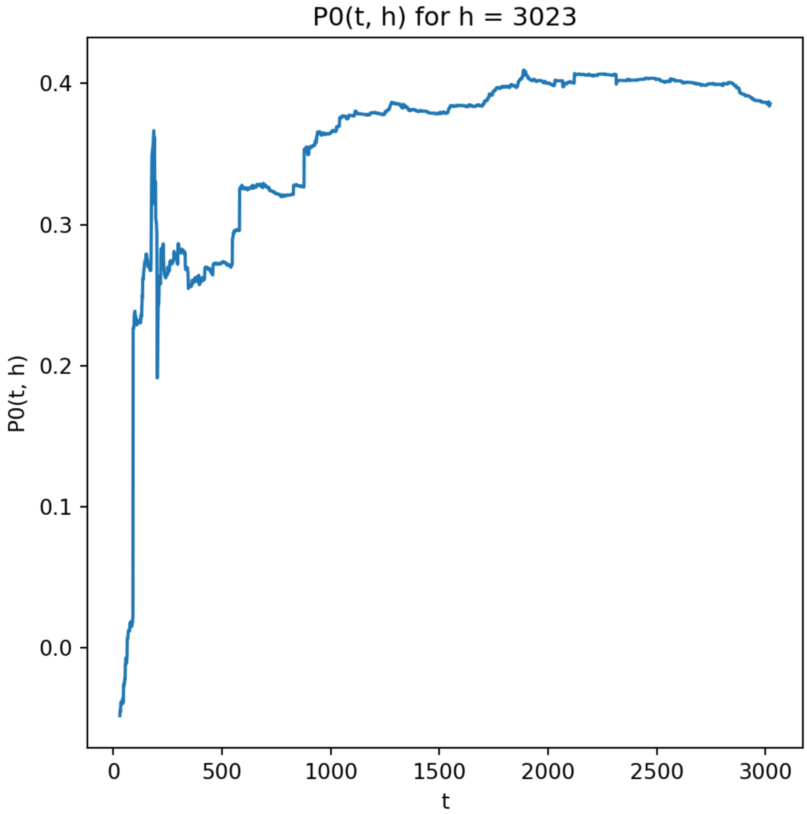

Fig. 2.1.4. $p_0(t, h)$ for h = ∞

2.2 Plot $S_0(t,h)$ and $\gamma_0(t,h)$ for h = 30, 100, 300, and ∞

$\gamma_0(t,h)$ is the Sharpe ratio of stock price.

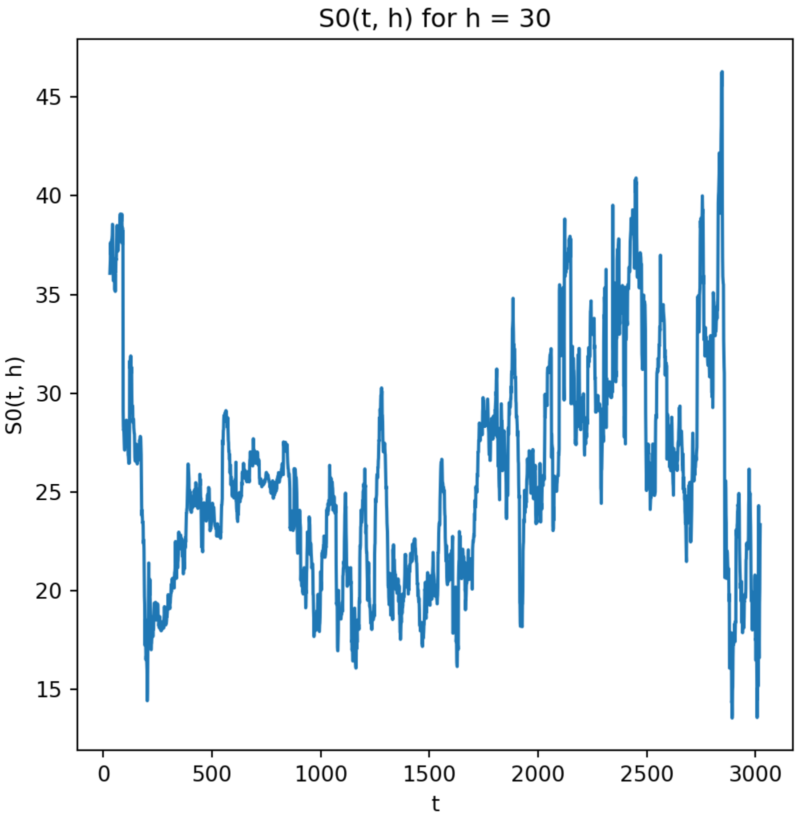

Fig. 2.2.1. $S_0(t,h)$ for h = 30

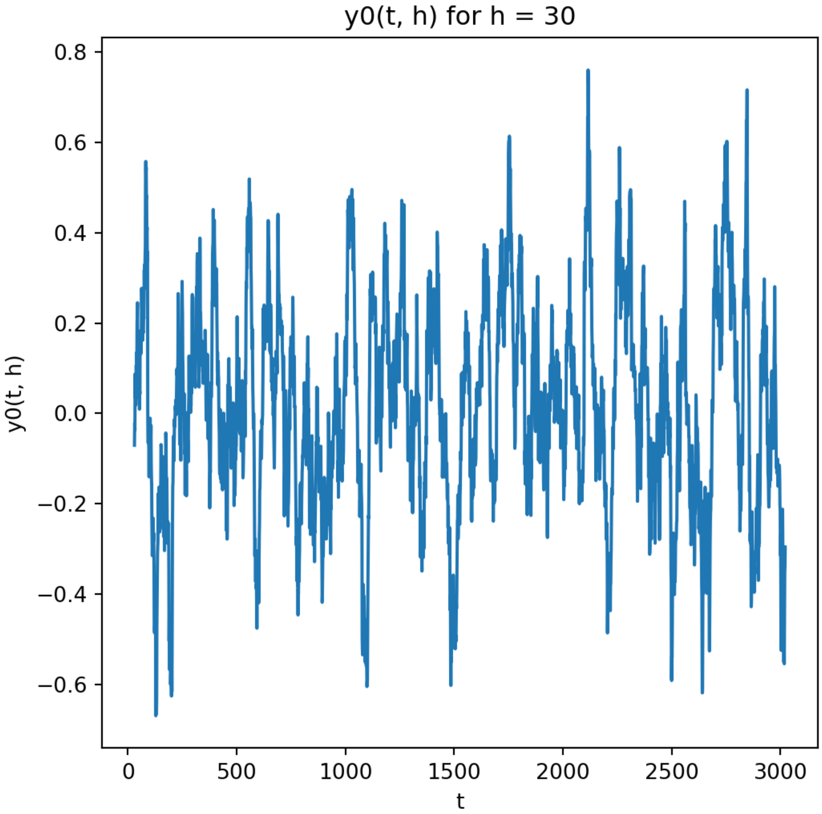

Fig. 2.2.2. $\gamma_0(t,h)$ for h = 30

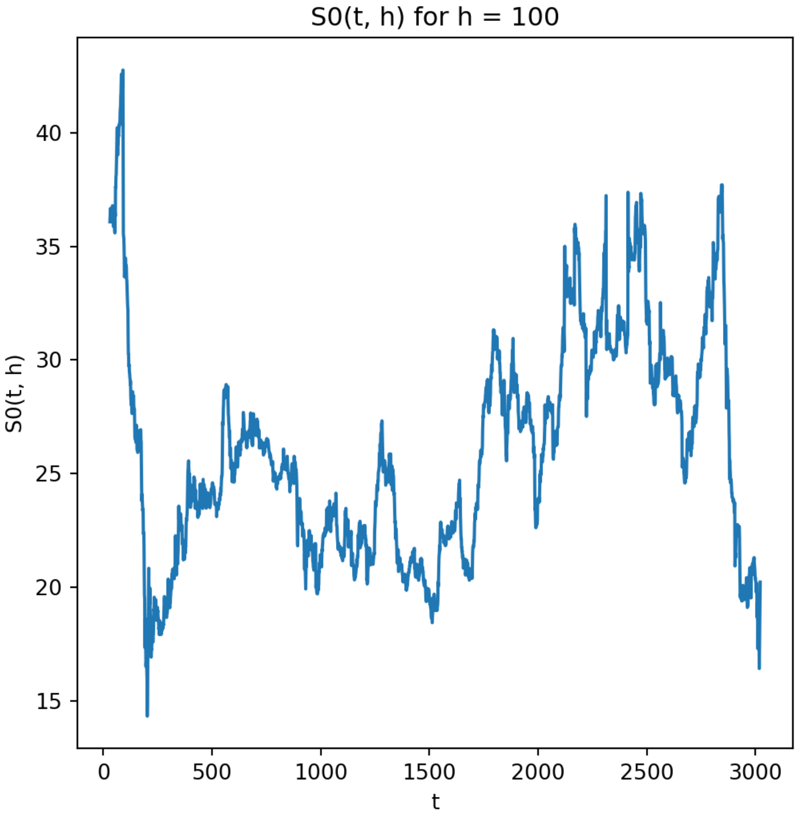

Fig. 2.2.3. $S_0(t, h)$ for h = 100

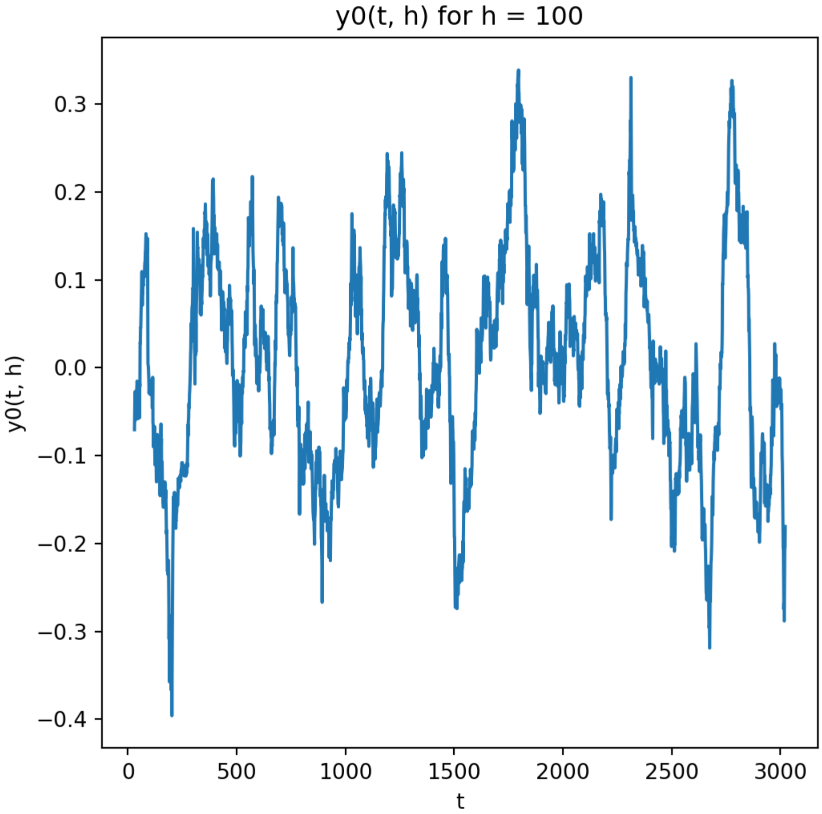

Fig. 2.2.4. $\gamma_0(t, h)$ for h = 100

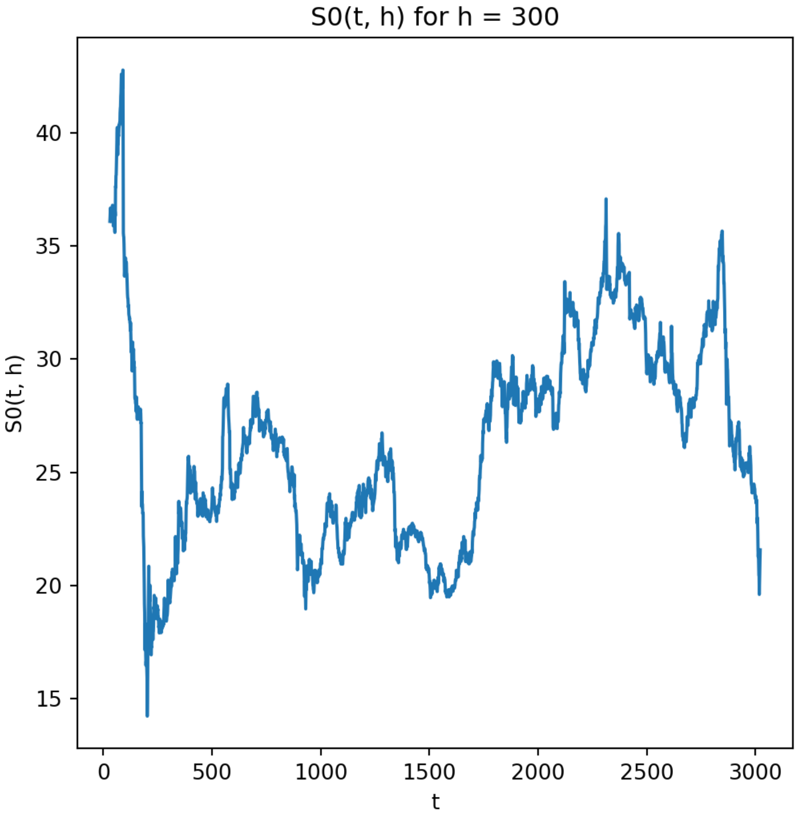

Fig. 2.2.5. $S_0(t,h)$ for h = 300

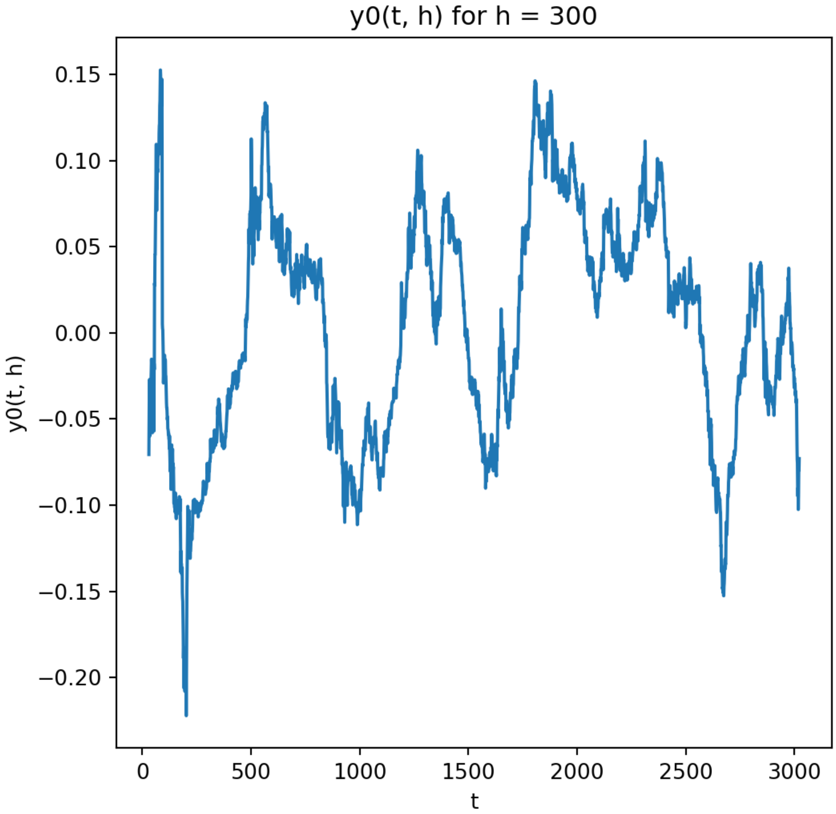

Fig. 2.2.6. $\gamma_0(t,h)$ for h = 300

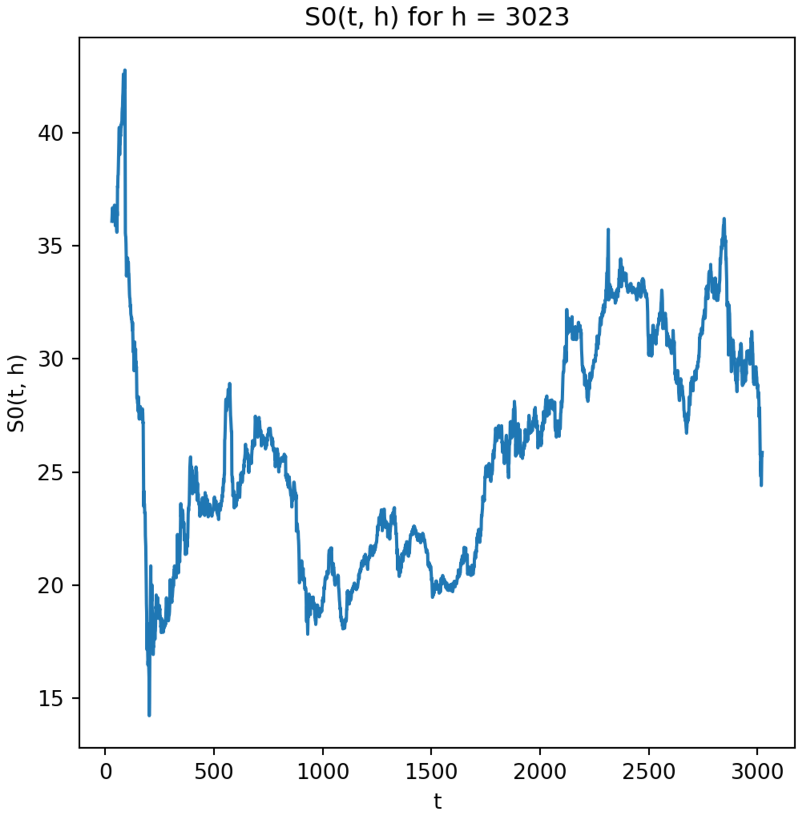

Fig. 2.2.7. $S_0(t,h)$ for h = ∞

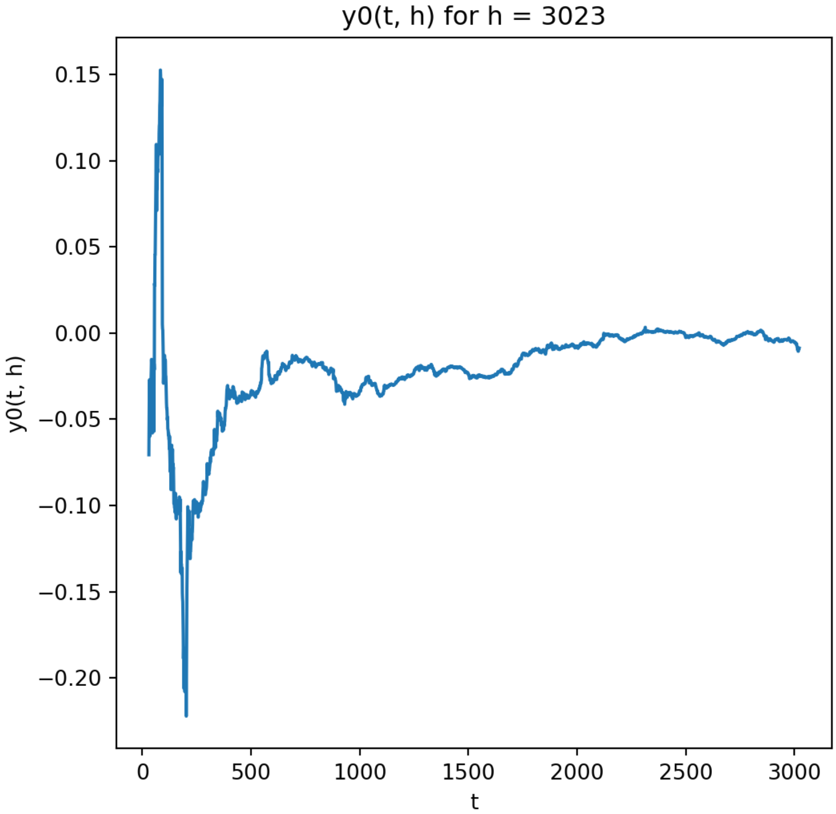

Fig. 2.2.8. $\gamma_0(t,h)$ for h = ∞

We can that the larger h is, the smoother all plots are.

3 Moving Average

Now we define $S = S_1$ and $X = X_1$, We are going to examine the effectiveness of moving average, which some traders heavily rely on and regard as accurate indicators.

3.1 EMA analysis

First we plot the w-day exponential moving average (EMA) of S for w = 30, 100, and 300.

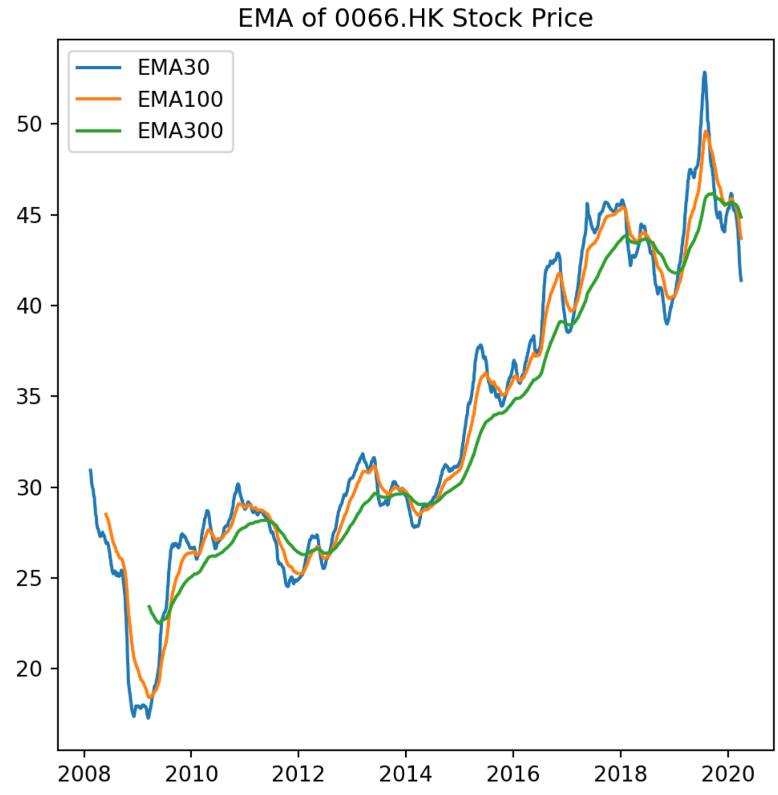

Fig. 3.1. EMA of 0066.HK

We can see that the larger w is, the smoother EMA line is.

3.2 SMA analysis

Then we plot the w-day simple moving average (SMA) of S for the same values of w, and do comparison with EMA.

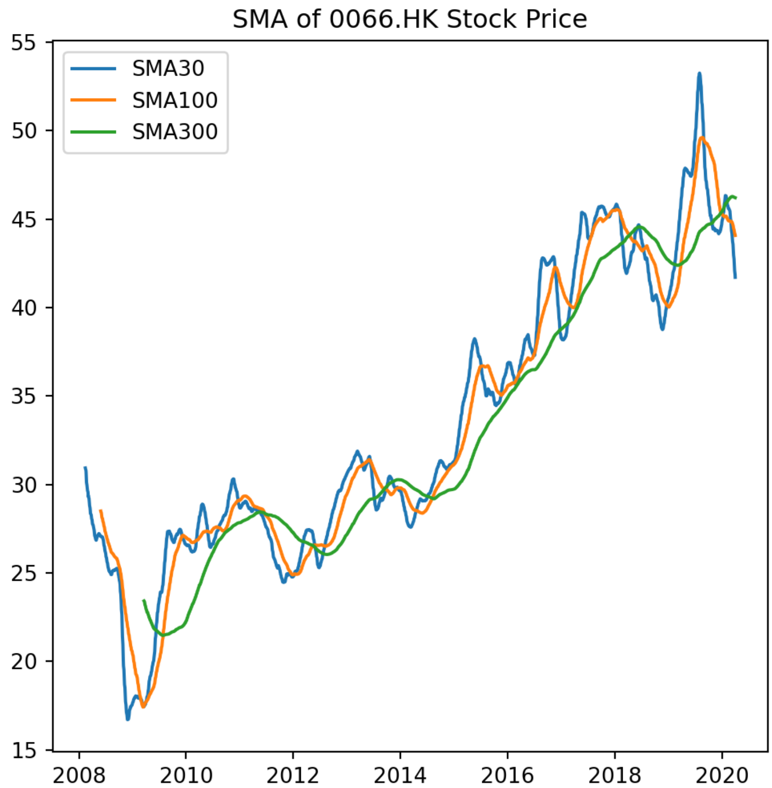

Fig. 3.2.1. SMA of 0066.HK

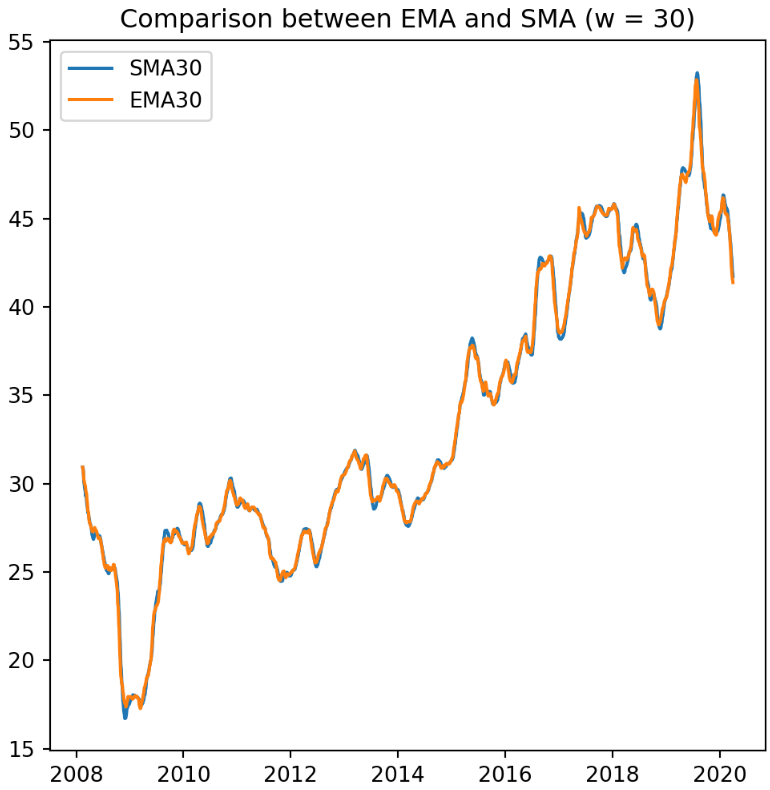

Fig. 3.2.2. Comparison of EMA and SMA with w = 30

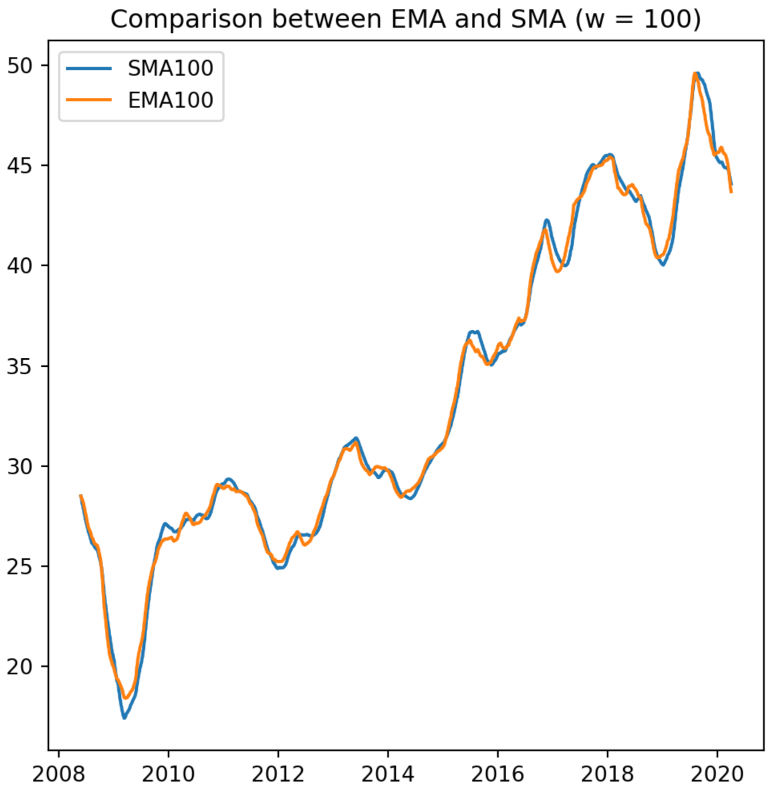

Fig. 3.2.2. Comparison of EMA and SMA with w = 100

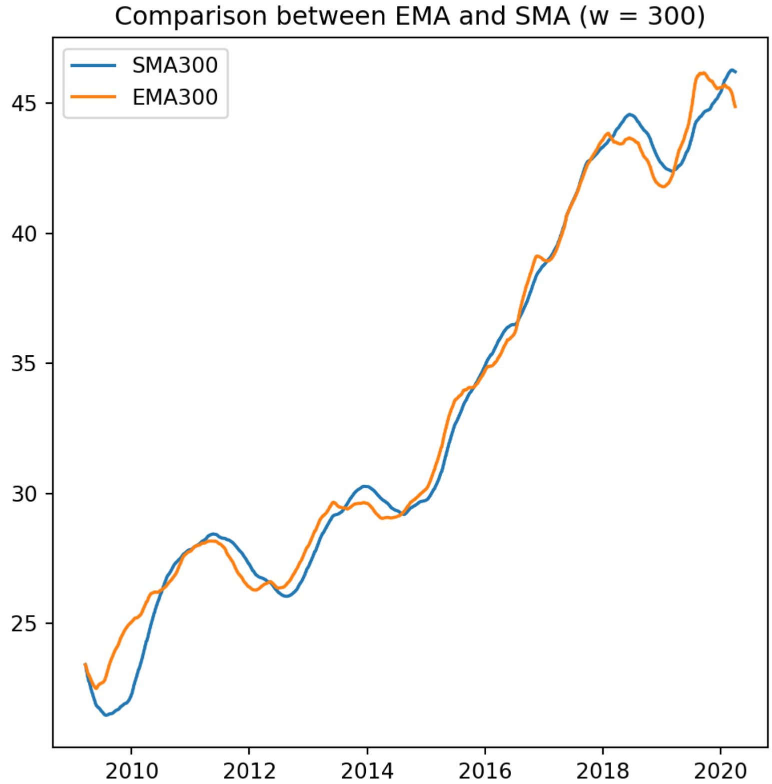

Fig. 3.2.2. Comparison of EMA and SMA with w = 300

We can see that they are quite similar.

3.3 MACD analysis

Next compute the stock' s moving-average-convergence-divergence (MACD) line and signal line.

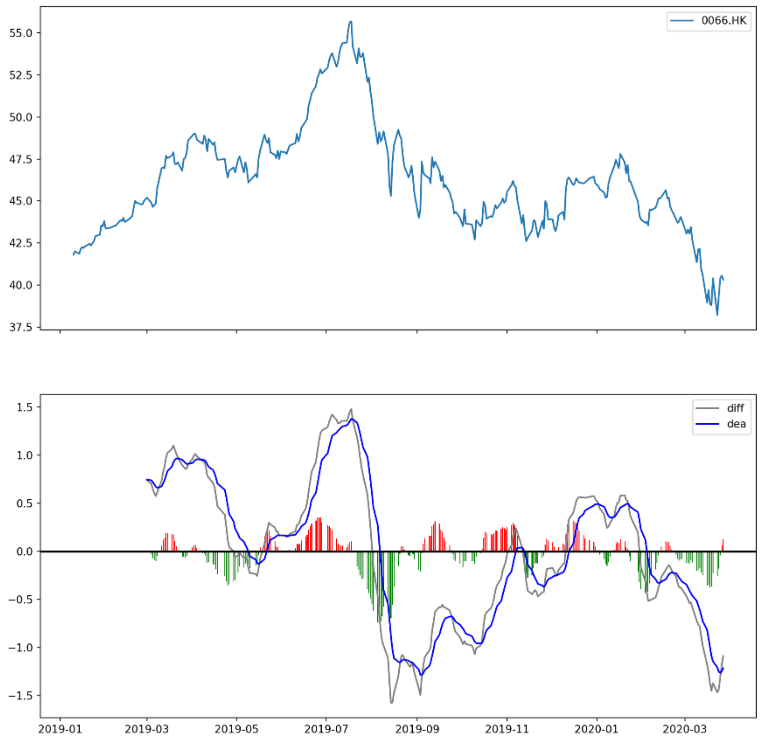

Fig. 3.3. MACD of 0066.HK

When the MACD line crosses above the signal line, it is considered a bullish signal. Conversely, when the MACD line crosses below the signal line, it is viewed as a bearish signal. When the MACD line crosses the zero line, it signifies a change in the direction of the trend. The observation match the empirical rules.

4 Probability Density Function

We may assume X is normally distributed and estimate $f(x)$ with a normal distribution or logistic distribution. From our calculation, we can fit the function parameters:

In $g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$, μ = 8.8989e-05, σ^2 = 0.000207.

In $L(x) = \frac{1}{1+exp[-b(x-x^*)]}$, x^* = 0.000224 , b = 143.971.

Then let' s plot our fits $G(x)$ ≡ ∫ $g(x)\mathrm{d}x$ and $L(x)$ atop $F(x)$, to check if they are reasonable.

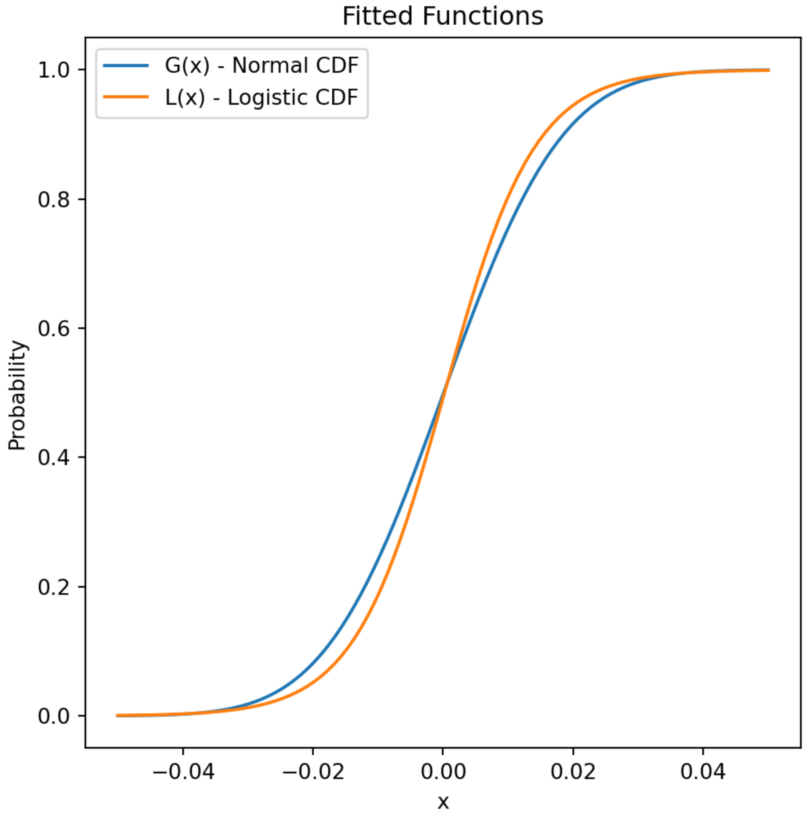

Fig. 4. $G(x)$ and $L(x)$ as CDF

4.1 Digitization and Conditional Probability

Given the return $X(t)$ on a day, we would like to predict value $X(t+1)$ on the next day. It is more realistic to predict the trend qualitatively, so first we digitize $X(t)$ as $Y\ (t)$ with three alphabets, viz. D for "down", U for "up", and H for "hold". We choose the threshold ε to be 0.04 to fit the 0062.HK's prices.

First let’s plot $F(x|y) \equiv \mathrm{CDF}[X(t) = x | Y(t+1) = y]$. Compare with CDF’s definition with raw data, we can see that Normal assumption is more reasonable.

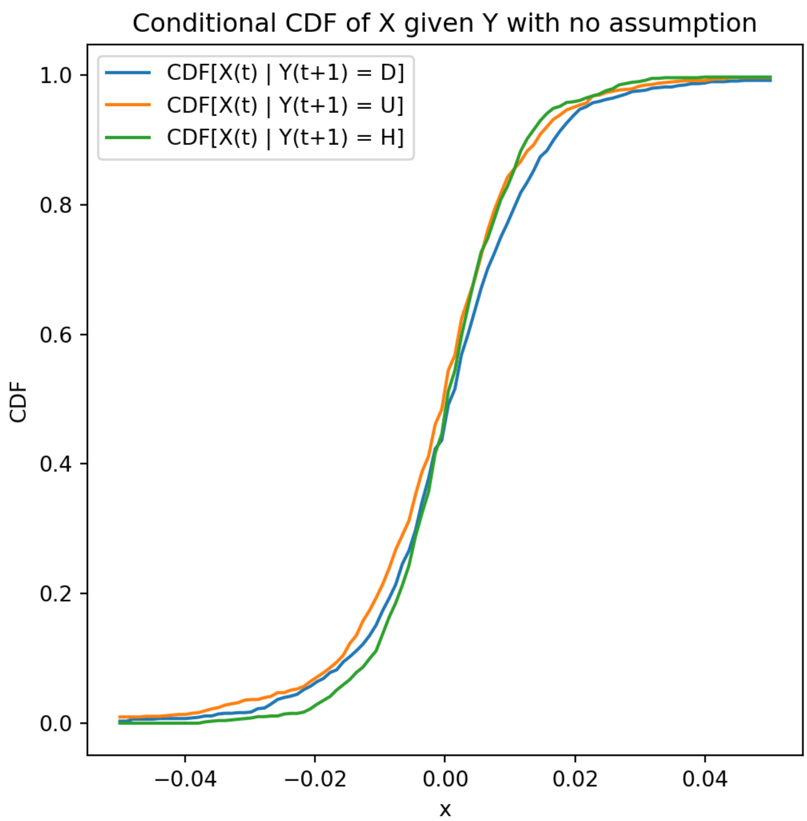

Fig. 4.1.1. CDF of definition

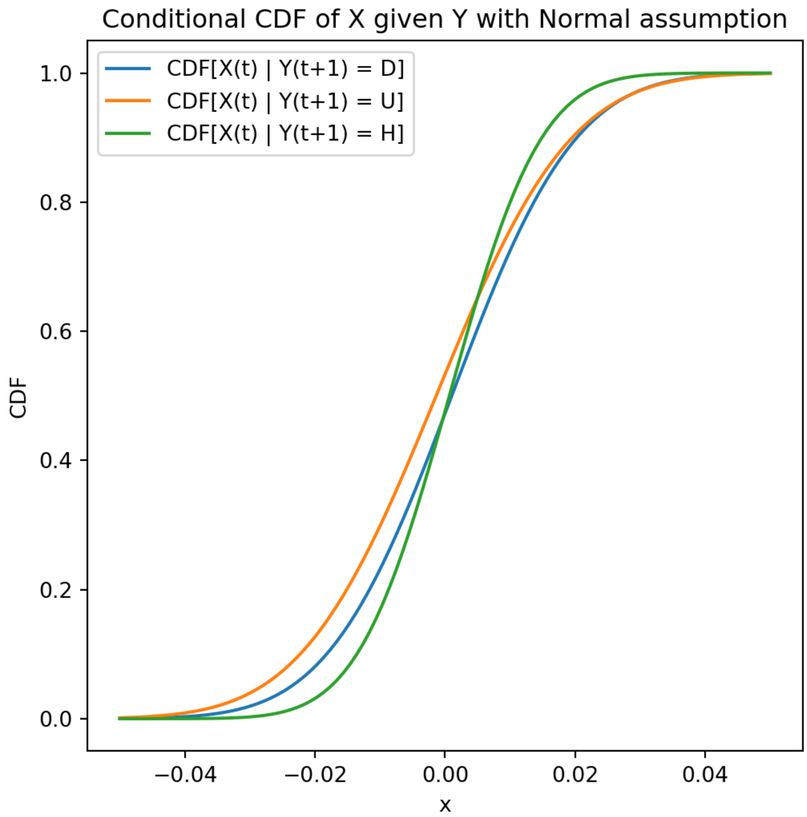

Fig. 4.1.2. CDF of Normal assumption

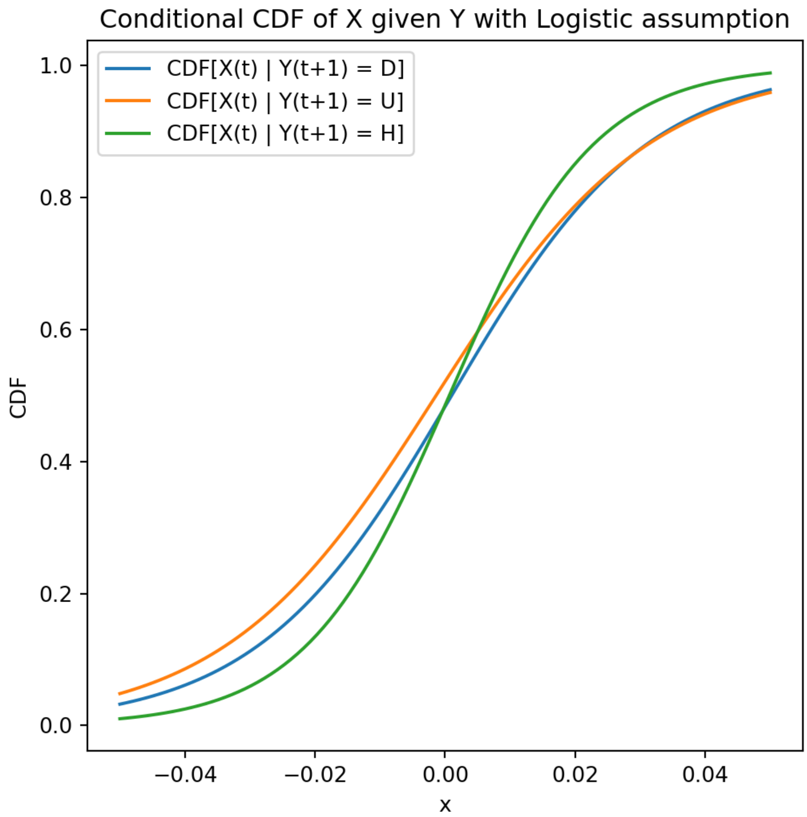

Fig. 4.1.3. CDF of Logistic assumption

When coming with PDF, we follow the same methods to extract the three corresponding conditional PDFs $f_y(x)$, then plot them on the same graph.

For convenience, we choose Normal assumption for later steps.

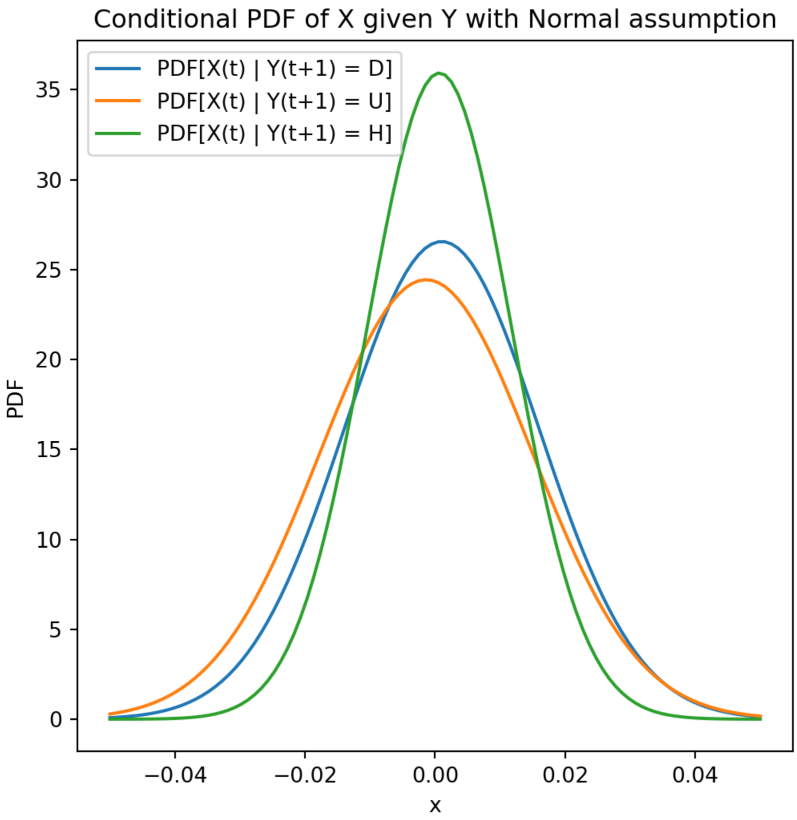

Fig. 4.1.4. PDF of $f_y(x)$, y = D,U or H

5 Bayes Detector

We will now construct a Bayes detector with the obtained PDFs to predict $Y(t+1)$ after observing $X(t) = x$.

We should first compute the probabilities $P[Y\ (t\ +\ 1) = y]$ for y = D, U, and H, used as the prior probabilities $q(y)$ of the hypotheses —— $H_D\!: Y(t+1) = D,\ H_U\!: Y(t+1) = U,\ H_H\!: Y(t+1) = H.$

We can next predict with the maximum posterior probability. We will predict the $y^* = arg\,max_y P\big[Y(t+1) = y|X_{(t)} = x\big]$ = $arg\,max_y\ q_y f_y\ (x)$. Compute $y^*(x)$ and make a graph.

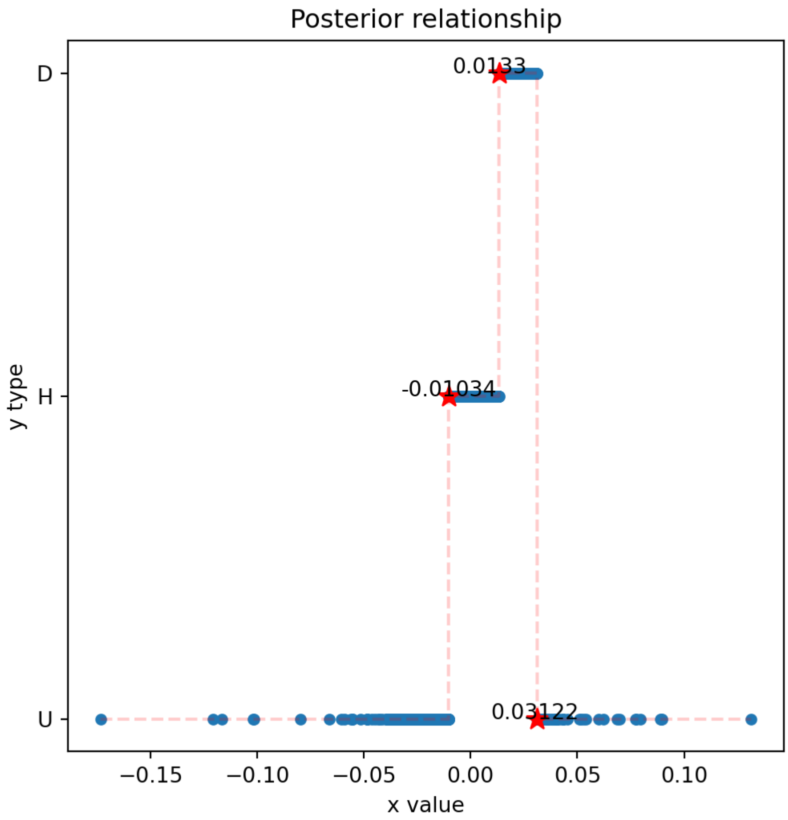

Fig. 5. Bayesian predictor

We can see that, with x-value increasing at x-value -0.01036, y will change from U to H, at x-value 0.01354, y will change from H to D, at x-value 0.03122, y will change from D to H.

6 Association Rules

Next we can consider the association rules. When $k = 5$, we have 3 729 possible association rules in the form of :

$\{Y\ (t-4), Y\ (t-3), Y\ (t-2), Y\ (t-1), Y\ (t)\} \rightarrow Y\ (t+1) = y$.

We analyze the 10 rules with the highest support and the 10 rules with the highest confidence.

```
##                                    antecedents consequents   support
## 75986  [Y_1_H, Y_2_H, Y_3_H, Y_4_H, Y_5_H]       (Y_6_H)  0.004642
## 76110  [Y_1_H, Y_2_H, Y_3_H, Y_4_H, Y_5_U]       (Y_6_H)  0.004310
## 69973  [Y_1_D, Y_2_U, Y_3_D, Y_4_U, Y_5_U]       (Y_6_D)  0.003979
## 73444  [Y_1_H, Y_2_D, Y_3_H, Y_4_H, Y_5_H]       (Y_6_H)  0.003979
## 83240  [Y_1_U, Y_2_D, Y_3_U, Y_4_U, Y_5_U]       (Y_6_H)  0.003979
## 69106  [Y_1_D, Y_2_U, Y_3_D, Y_4_D, Y_5_U]       (Y_6_D)  0.003979
## 74870  [Y_1_H, Y_2_H, Y_3_D, Y_4_H, Y_5_H]       (Y_6_H)  0.003647
## 65387  [Y_1_D, Y_2_D, Y_3_U, Y_4_D, Y_5_D]       (Y_6_U)  0.003647
## 80702  [Y_1_U, Y_2_D, Y_3_D, Y_4_U, Y_5_D]       (Y_6_U)  0.003647
## 64641  [Y_1_D, Y_2_D, Y_3_D, Y_4_U, Y_5_U]       (Y_6_D)  0.003647
##                                    antecedents consequents  confidence
## 76238  [Y_1_H, Y_2_H, Y_3_H, Y_4_U, Y_5_D]       (Y_6_U)    1.000000
## 65818  [Y_1_D, Y_2_D, Y_3_U, Y_4_H, Y_5_H]       (Y_6_H)    0.833333
## 81320  [Y_1_U, Y_2_D, Y_3_H, Y_4_H, Y_5_D]       (Y_6_U)    0.777778
## 64767  [Y_1_D, Y_2_D, Y_3_H, Y_4_D, Y_5_D]       (Y_6_U)    0.750000
## 69539  [Y_1_D, Y_2_U, Y_3_D, Y_4_H, Y_5_U]       (Y_6_D)    0.750000
## 87084  [Y_1_U, Y_2_U, Y_3_D, Y_4_U, Y_5_H]       (Y_6_H)    0.727273
## 76610  [Y_1_H, Y_2_H, Y_3_U, Y_4_D, Y_5_H]       (Y_6_D)    0.714286
## 85970  [Y_1_U, Y_2_H, Y_3_U, Y_4_U, Y_5_D]       (Y_6_D)    0.714286
## 82435  [Y_1_U, Y_2_D, Y_3_U, Y_4_H, Y_5_D]       (Y_6_D)    0.700000
## 81442  [Y_1_U, Y_2_D, Y_3_H, Y_4_H, Y_5_U]       (Y_6_H)    0.700000
```

Fig. 6. 10 rules with highest support or confidence

6.1 Usefulness of Association Rules

The usefulness u of a rule should be proportional to its support s and its confidence c because a useful rule is both frequent and accurate.

We may assume $u \sim s^{1-\lambda} c^{\lambda}$ for some tuning parameter $\lambda \in [0,1]$. As λ rises, the emphasis of this quantity smoothly slides from support to confidence.

$\lambda = \frac{1}{2}$, u is the geometric mean of support and confidence.

$\lambda = \frac{2}{3}$, $u = \sqrt[3]{sc^2} = \sqrt[3]{r}$, where r is a rule’ s rule power factor (RPF).

We analyze the 10 rules with the highest geometric mean and the 10 rules with the highest RPF.

```
##                                      antecedents consequents  geometric_mean
## 76110  [Y_1_H, Y_2_H, Y_3_H, Y_4_H, Y_5_U]     (Y_6_H)        0.054306
## 74870  [Y_1_H, Y_2_H, Y_3_D, Y_4_H, Y_5_H]     (Y_6_H)        0.048579
## 64641  [Y_1_D, Y_2_D, Y_3_D, Y_4_U, Y_5_U]     (Y_6_D)        0.048579
## 83240  [Y_1_U, Y_2_D, Y_3_U, Y_4_U, Y_5_U]     (Y_6_H)        0.047682
## 69973  [Y_1_D, Y_2_U, Y_3_D, Y_4_U, Y_5_U]     (Y_6_D)        0.047682
## 75986  [Y_1_H, Y_2_H, Y_3_H, Y_4_H, Y_5_H]     (Y_6_H)        0.047338
## 69539  [Y_1_D, Y_2_U, Y_3_D, Y_4_H, Y_5_U]     (Y_6_D)        0.047308
## 83056  [Y_1_U, Y_2_D, Y_3_U, Y_4_U, Y_5_H]     (Y_6_D)        0.045952
## 73695  [Y_1_H, Y_2_D, Y_3_H, Y_4_U, Y_5_H]     (Y_6_U)        0.045522
## 87084  [Y_1_U, Y_2_U, Y_3_D, Y_4_U, Y_5_H]     (Y_6_H)        0.043922
##                                       antecedents consequents      RPF
## 76110  [Y_1_H, Y_2_H, Y_3_H, Y_4_H, Y_5_U]      (Y_6_H)  0.002018
## 69539  [Y_1_D, Y_2_U, Y_3_D, Y_4_H, Y_5_U]      (Y_6_D)  0.001679
## 74870  [Y_1_H, Y_2_H, Y_3_D, Y_4_H, Y_5_H]      (Y_6_H)  0.001527
## 64641  [Y_1_D, Y_2_D, Y_3_D, Y_4_U, Y_5_U]      (Y_6_D)  0.001527
## 81320  [Y_1_U, Y_2_D, Y_3_H, Y_4_H, Y_5_D]      (Y_6_U)  0.001404
## 87084  [Y_1_U, Y_2_U, Y_3_D, Y_4_U, Y_5_H]      (Y_6_H)  0.001403
## 76238  [Y_1_H, Y_2_H, Y_3_H, Y_4_U, Y_5_D]      (Y_6_U)  0.001326
## 83240  [Y_1_U, Y_2_D, Y_3_U, Y_4_U, Y_5_U]      (Y_6_H)  0.001299
## 69973  [Y_1_D, Y_2_U, Y_3_D, Y_4_U, Y_5_U]      (Y_6_D)  0.001299
## 73695  [Y_1_H, Y_2_D, Y_3_H, Y_4_U, Y_5_H]      (Y_6_U)  0.001295
```

Fig. 6.1.1. 10 rules with highest support or confidence

We think the geometric mean and RPF are enough for estimation, we tried other λ but the results are almost the same as RPF or geometric mean.

Finally we can take the intersection results of geometric mean and RPF, and take the most credible 8 rules.

DUDUU→D, HDHUH→U, DUDHU→D, DDDUU→D, UUDUH→H, HHDHH→H, HHHHU→H, UDUUU→H.

```
RPF_result = []
for row in range(10):
    antecedents = rules_RPF_top10.iloc[row, 0]
    hdu = ''
    for i in antecedents:
        hdu += i.split('_')[2]
    hdu += next(iter(rules_RPF_top10.iloc[row, 1])).split('_')[2]
    RPF_result.append(hdu)
print(RPF_result)
```

```
## ['HHHHUH', 'DUDHUD', 'HHDHHH', 'DDDUUD', 'UDHHDU', 'UUDUHH', 'HHHUDU', 'UDUUUH', 'DUDUUD', 'HDHUHU']
```

```
gm_result = []
for row in range(10):
    antecedents = rules_geometric_mean_top10.iloc[row, 0]
    hdu = ''
    for i in antecedents:
        hdu += i.split('_')[2]
    hdu += next(iter(rules_geometric_mean_top10.iloc[row, 1])).split('_')[2]
    gm_result.append(hdu)
print(gm_result)
```

```
## ['HHHHUH', 'HHDHHH', 'DDDUUD', 'UDUUUH', 'DUDUUD', 'HHHHHH', 'DUDHUD', 'UDUUHD', 'HDHUHU', 'UUDUHH']
```

```
set(RPF_result).intersection(set(gm_result)) # 8 rules
```

```
## {'DUDUUD', 'HDHUHU', 'DUDHUD', 'DDDUUD', 'UUDUHH', 'HHDHHH', 'HHHHUH', 'UDUUUH'}
```

Fig. 6.1.2. 8 rules based on results intersection

7 A Portfolio with One Stock and Money

Let $M(t)$ be the amount of money and $N(t)$ the number of shares at the end of day t, so the portfolio' s value $V(t)$ is defined as $V = M + NS$.

We are initially given the minimum-risk portfolio that has V (0) = 100,000 dollars —— M(0) = 100,000 dollars at the beginning. Then we start trading at t = 1 according to the prediction by the tools.

If we decide to trade on day t,

- we may spend m = $gM(t$ − 1) dollars buying $m/S(t)$ shares, or
- we may sell n = $gN(t$ − 1) shares to get $nS(t)$ dollars.

The parameter g ∈ (0,1) quantifies the "greed". The greedier we choose to be, the more we want to earn and thus the more we trade per transaction.

Meanwhile, M is compounded with a daily interest rate r at the beginning of each day, so the minimum-risk portfolio's value amounts to :
$V_{0(t)} = M(0) \times (1 + r)^t$. We set r = 0.001%.

We choose the action to buy, hold or sell depending on the strategy in the picture, we use the tools of MACD results and Bayesian predictor.

Today's State	Tomorrow's State	Today's MACD Value	Tomorrow's MACD Value - Today's	Action
U	U	>0	>0	Buy
			<0	Hold
		<0	>0	Hold
			<0	Sell
U	H/D			Sell
H	U			Buy
H	H			Hold
H	D			Sell
D	U/H			Buy
D	D	>0	>0	Hold
			<0	Hold
		<0	>0	Sell
			<0	Sell

Fig. 7.1 Portfolio Strategy (1 stock with M)

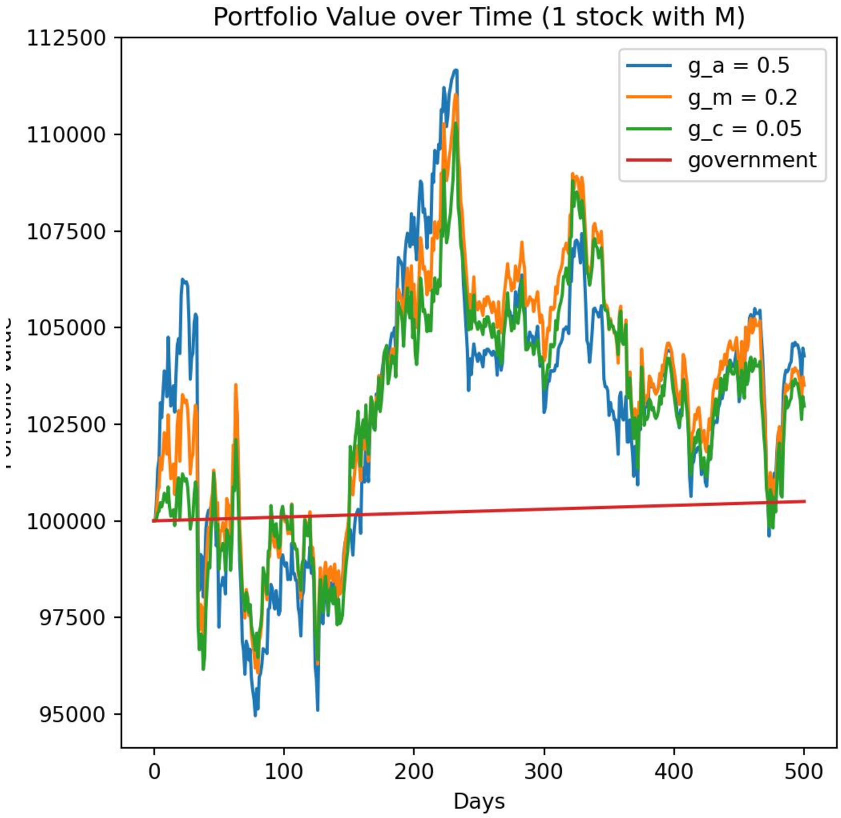

Fig. 7.2 Portfolio Value over Time (1 stock with M)

We can see that all greedy portfolios perform better than the minimum-risk portfolio significantly.

Actually, the middle greedy $g_M = 0.2$ performs better than either $g_A = 0.5$ or $g_C = 0.05$. We can just choose the middle point of g_A and g_C on this model's background.

The highst value can come to about 104000 at the end time.

8 A Portfolio with Two Stocks

Now consider a market model in which we are going to trade the riskier stock(S1) using the safer stock(S2) instead of money.

Let $N_1(t)$ and $N_2(t)$ be the number of shares in the riskier stock and the safer stock, so our portfolio' s value is $V = N_1 S_1 + N_2 S_2$.

We are initially given the minimum-risk portfolio that has V (0) = 100,000 dollars, then we can trade once per day. We can sell $n_i = gN_i(t-1)$ shares in either stock to buy $n_i S_i(t)/S_{3-i}(t)$ shares in the other stock on day t.

To do comparison, we must buy or sell the riskier stock on the days when we bought or sold it in section 7.

We choose the money division of p_0 depending on the minimum-risk portfolio result in section 2, which is 0.38.

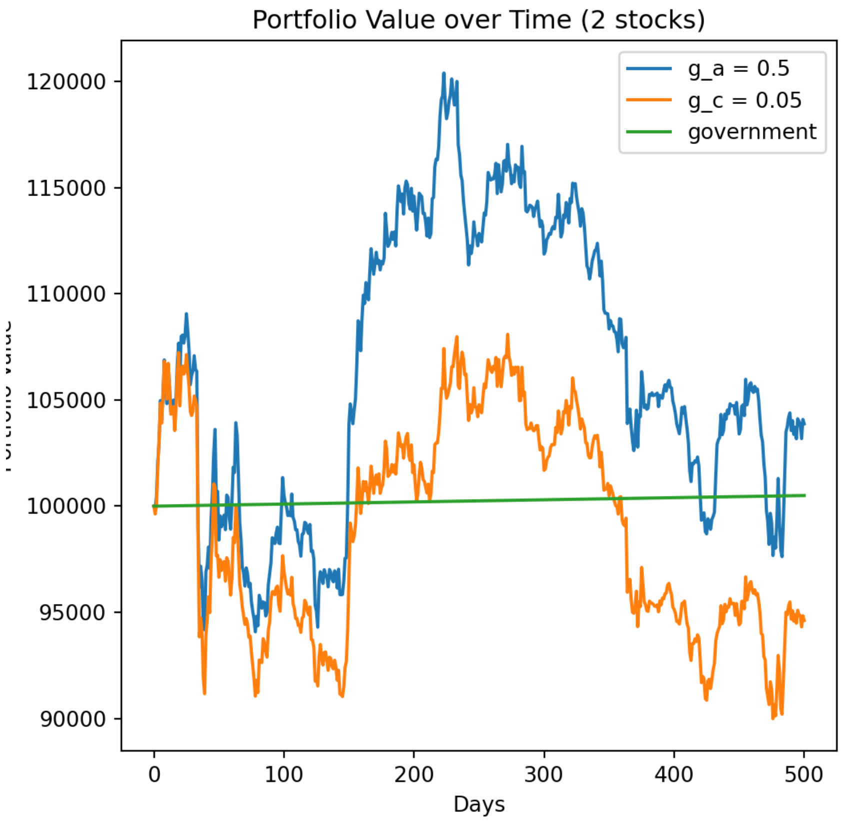

Fig. 8. Portfolio Value over Time (2 stocks)

We can see that it is worth being greedy this time, because we' ll lose money at the end if we choose to be conservative. Maybe it' s caused by stock2' s continuously decline.

The highst value can just come to about 102000 this time.

8.1 Efficient Frontier

The discussed scheme of investment is probably suboptimal since we may buy too many shares in the safer stock, resulting in a portfolio that yields a return even lower than the minimum-risk portfolio. We can prevent this situation by taking the stocks' efficient frontier into account.

Since the portfolio's risk $\sigma(t)$ is always between the minimum-risk portfolio's risk σ_0 and the riskier stock's risk σ_1.

This time we should set n_i in a way so that after each trade,
– the portfolio becomes safer and yields $\sigma(t) = \sigma(t-1) - g[\sigma(t-1) - \sigma_0]$, or
– the portfolio becomes riskier and yields $\sigma(t) = \sigma(t-1) + g[\sigma_1 - \sigma(t-1)]$.

Hence, the portfolio's return is always at least equal to that of the minimum risk portfolio.

Let's first plot $p_{(t)} = N_1(t)S_1(t) \,/\, V_1(t)$ for the original scheme of investment, then compare it to the minimum-risk fraction $p_0(t)$.

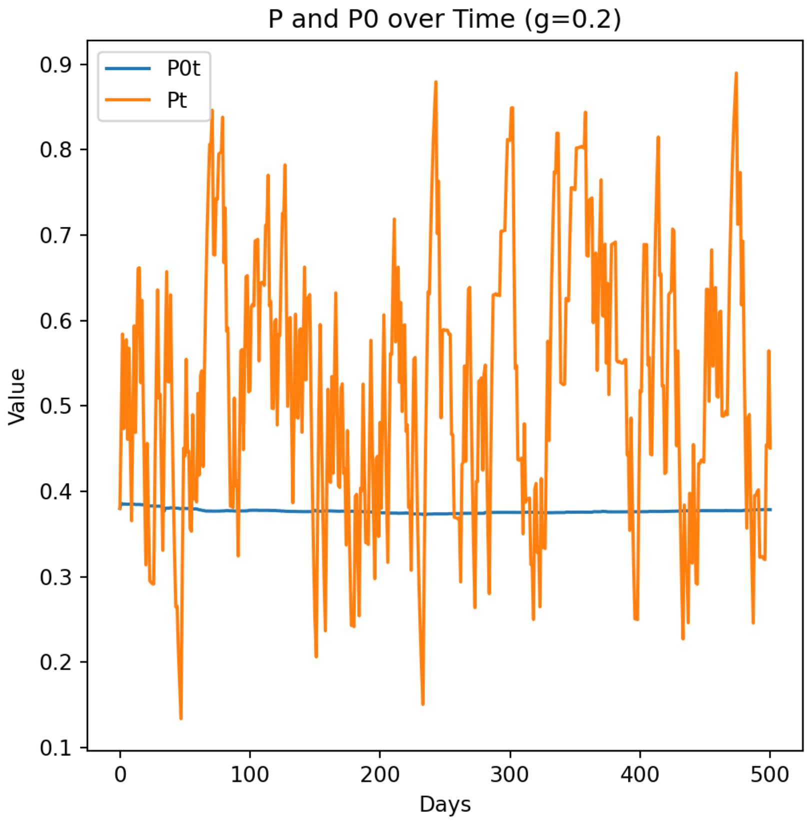

Fig. 8.1.1. P and P0 over Time (g=0.2)

We can see that there indeed exist some undesirable moments at which $p(t) < p_0(t)$.

Let' s use the new model to do research and plot the result again.

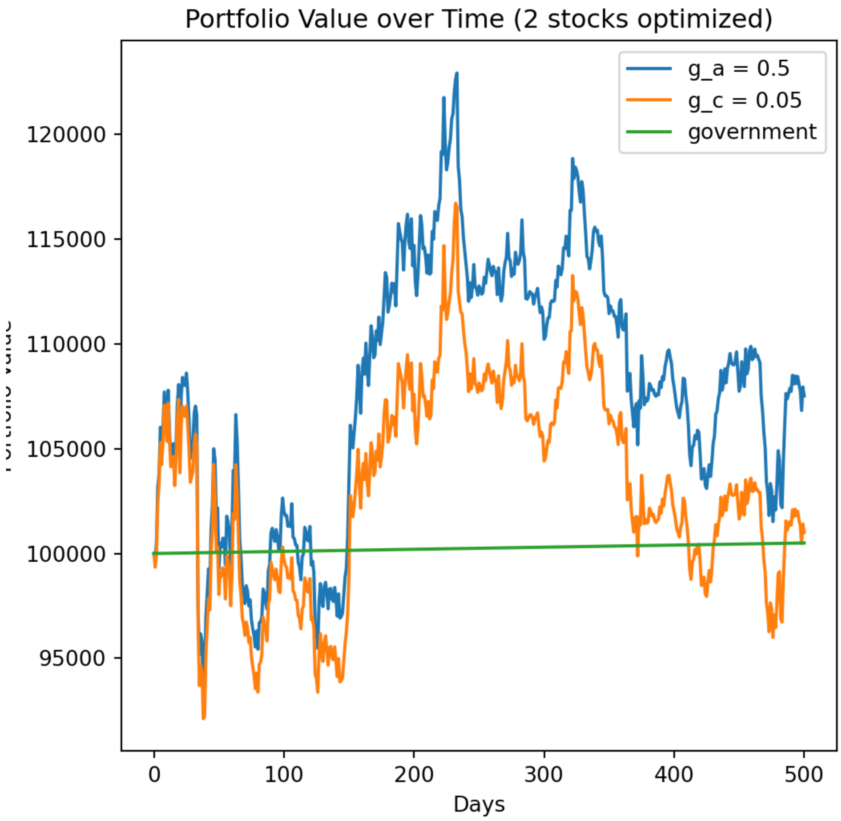

Fig. 8.1.2. Portfolio Value over Time (2 stocks optimized)

We can see this time they perform better than before, the highst value can come to about 106000 at the end time, even higher that of 1 stock portfolio value in section 7.

9 A Portfolio with Two Stocks and Money

Finally, we will work on a more realistic market model: we can now predict both stocks' prices and trade them with money independently, so we are going to manage a portfolio that comprises three assets.
We are initially given $M(0)$ = 100,000 dollars and $N_1(0) = N_2(0) = 0$ shares in the stocks, then we can trade each stock at most once per day.

We first update the action in sections 7 and 8 by both stocks' states predicted, to choose which to buy , which to sell and their orders.

S1 Action	S2 Action	Action	If Need to Balance pt
Buy	Buy	Buy both with half of gM	TRUE
Buy	Hold	Buy S1 with gM	FALSE
Buy	Sell	First sell gN2, then buy S1 with g(M+Sold_value)	FALSE
Hold	Buy	Buy S2 with gM	TRUE
Hold	Hold	Hold	FALSE
Hold	Sell	Sell S2 with gN2	FALSE
Sell	Buy	First sell gN1, then buy S2 with g(M+Sold_value)	TRUE
Sell	Hold	Sell S1 with gN1	TRUE
Sell	Sell	Sell both with half of gNi	TRUE

Fig. 9.1.1. Portfolio Strategy (2 stock with M)

Let's use the model to do research.

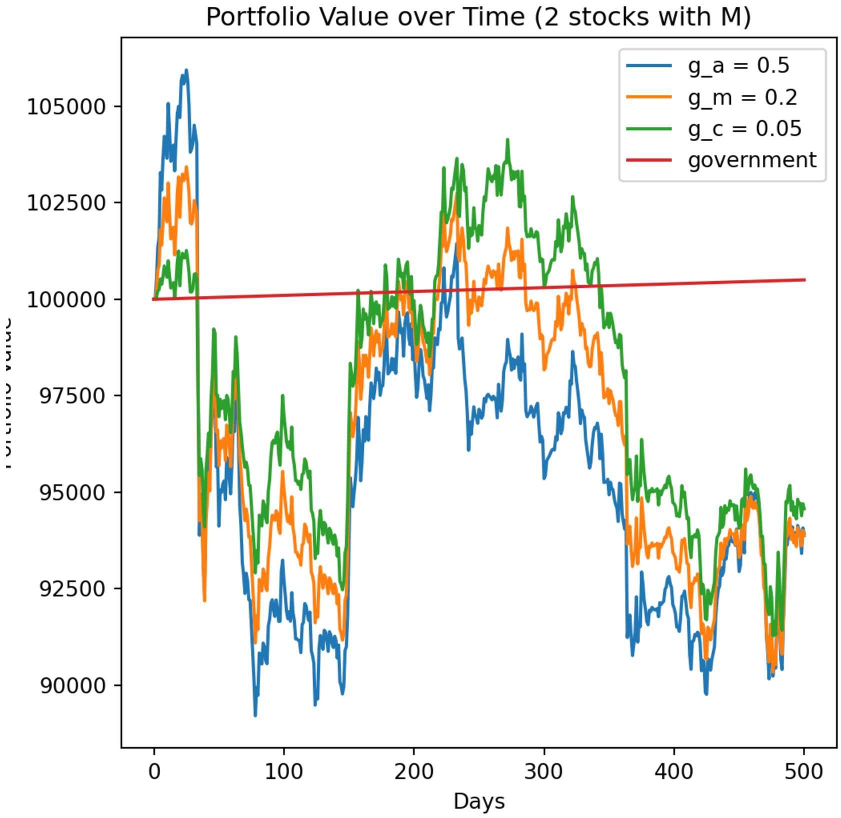

Fig. 9.1.2. Portfolio Value over Time (2 stocks with M)

However, the result looks bad, we need to find other methods. We decide to take RSI (Relative strength index) indicator into consideration.

The relative strength index (RSI) is a technical indicator used in the analysis of financial markets. It is intended to chart the current and historical strength or weakness of a stock or market based on the closing prices of a recent trading period. The indicator should not be confused with relative strength.

When n=14, the index is most critical. When the RSI of a security exceeds 70, it means that the security has been overbought and investors should consider selling the security. On the contrary, when the RSI of a security falls to 30, it means that the security is oversold and investors should buy the security.

After updating the action function, we first consider the two stocks as independent with each other, just add the 2 portfolios the same as that in section 7, to check the RSI' s effect.

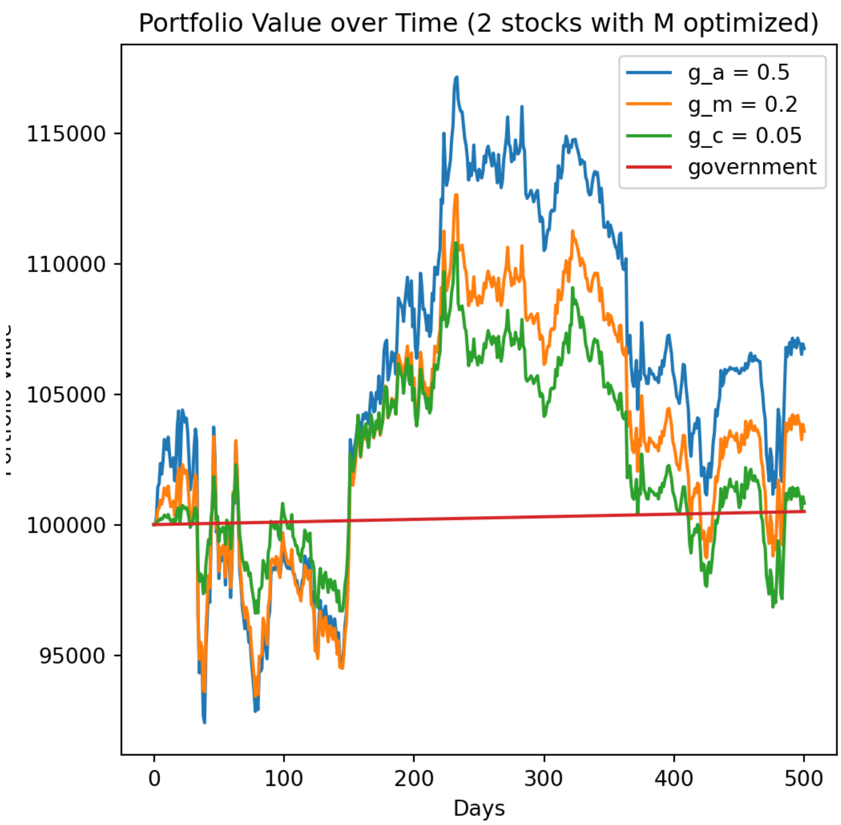

Fig. 9.1.3. Portfolio Value over Time (2 stocks with M optimized)

We can see the RSI really works, we can earn about 108000 value at the end.

Then we add the RSI method into the model with interaction effect.

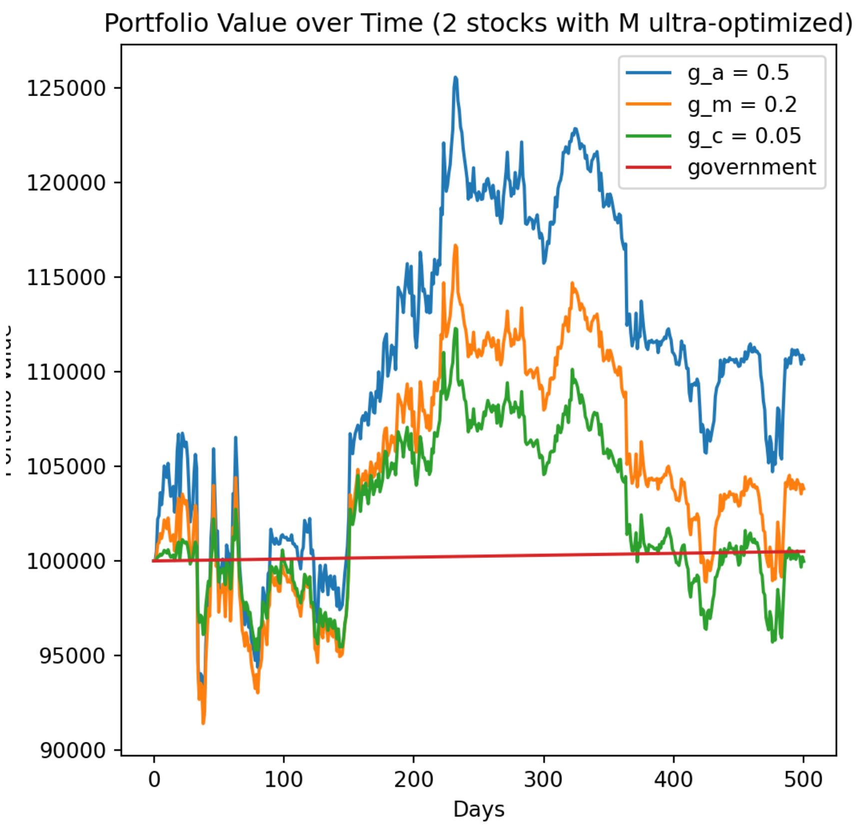

Fig. 9.1.4. Portfolio Value over Time (2 stocks with M optimized)

We can see this portfolio and strategy performs best, we can even get a final value of about 110000 with the aggressive greed parameter of 0.5.

Appendix

Special thanks for Chen Longyin’ s contribution to this project.

Codes:
https://github.com/ksye6/ksye6/tree/main/MSDM5058%E4%BF%A1%E6%81%AF%E7%A7%91%E5%AD%A6/%E4%BD%9C%E4%B8%9A/project2