Report: Trading in a volatile time

Chang Deng U71725017 changdeng@bu.edu

Yanbin Liu U14312397 ybliu17@bu.edu

Data

For data used in our project, we acquire a time series data of SP500 index from Jan. 30, 2020 until Mar. 30, 2020 from yahoo.finance website.

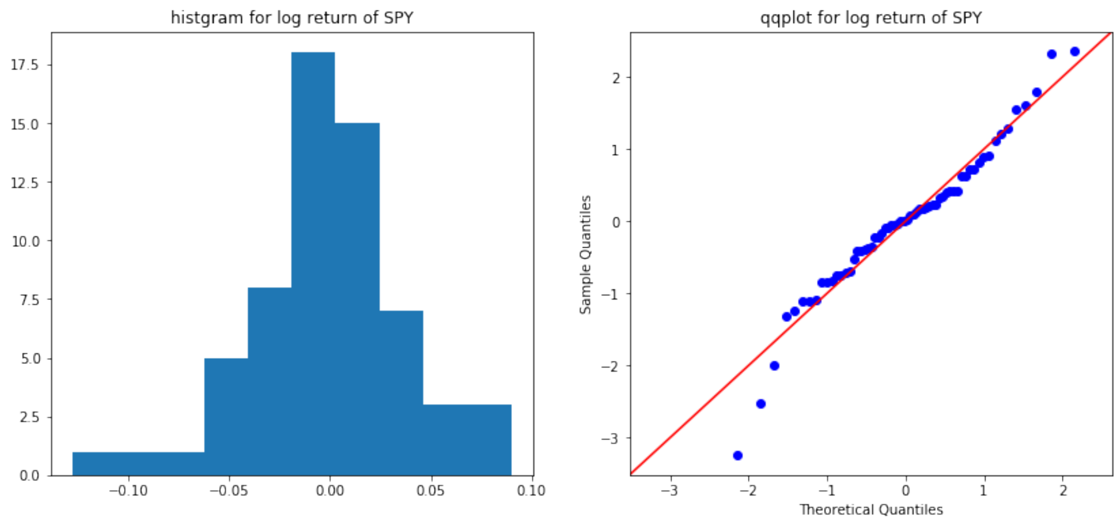
Summary

Using the data, the project aims to discuss how to trade in a volatile time. Intuitively, when exposed to price data, we are inspired to check how the return is distributed. By assumption, the log-return may exhibit normal distribution. So we examine it and find the exact distribution of log-return has thick tail. And this effect of the thick tail is extremely important in our data, since the market experienced slump in that period. After that, to be somewhat simple, we try to track the fluctuation of our price data after removal of the time trend. If we find some pattern then we can use the deviation from the time trend to construct daily signal for buy and sell with the belief the deviation is mean-reverting. Then we are encourage to construct a AR(1) model for the difference of that deviation, from which we may extract the mean-reverting model of the deviation. However, in our data, we find the AR(1) is not significant. We think this result may be owed to the relatively short gap of our data. So with the belief that the deviation is mean-reverting, we deduct the mean, volatility, speed of mean-reverting and use it to construct ad-hoc bands strategy. After trying different ad-hoc bands, we acquire the best strategy which seems to be reasonable for it recognize a reversion signal. Later we go through optimal problems to find the best exit and entry point given specific time. Applied the optimal signal, the strategy still is profitable. Finally, since in the real market sometimes we may not trade whenever we want,

we further discuss the delay execution of the ad-hoc bands strategy. Not to our surprise, the performance of the strategy is not as good as before, which is also reasonable. For more detailed of our project, please see the following part.

1. Data preparation and illustration

Calculate the log return of SP500 index using the "adj close" price. Draw the histogram of the log return, fit to a normal distribution, compare the tail of the fitted distribution with the empirical distribution.



As seen from the figures, the distribution of the log return is close to the fitted distribution of the empirical distribution (normal distribution) except that the tail of in the distribution of the log return is relatively thicker, which implies higher probability of extreme situation, especially extreme slump, as the left tail deviate more from the empirical distribution.

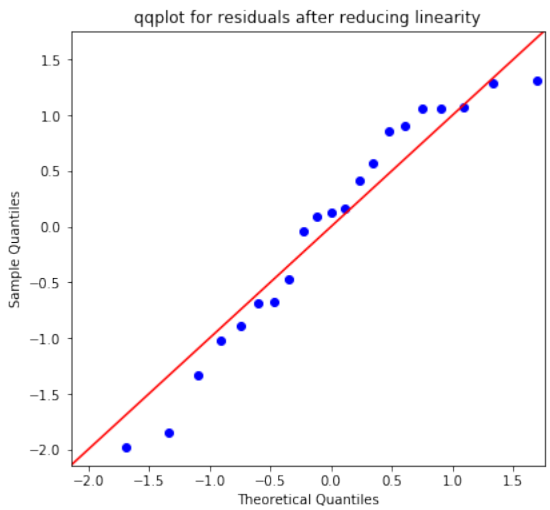
2. AR(1) model

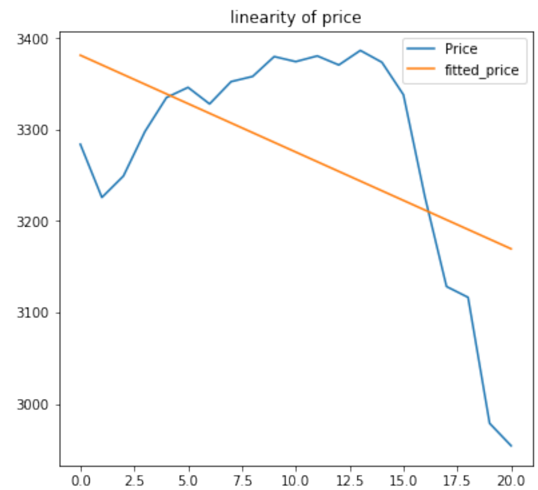
a) Split the data into a portfolio formation period and a portfolio testing period. You can download more data during your project.

Formation_period	Formation_price	Testing_period	Testing_price
2020-1-30	3283.659912	2020-3-2	3090.229980
2020-1-31	3225.520020	2020-3-3	3003.370117
2020-2-3	3248.919922	2020-3-4	3130.120117
2020-2-4	3297.590088	2020-3-5	3023.939941
2020-2-5	3334.689941	2020-3-6	2972.370117
2020-2-6	3345.780029	2020-3-9	2746.560059
2020-2-7	3327.709961	2020-3-10	2882.229980
2020-2-10	3352.090088	2020-3-11	2741.379883
2020-2-11	3357.750000	2020-3-12	2480.639893
2020-2-12	3379.449951	2020-3-13	2711.020020
2020-2-13	3373.939941	2020-3-16	2386.129883
2020-2-14	3380.159912	2020-3-17	2529.189941
2020-2-18	3370.290039	2020-3-18	2398.100098
2020-2-19	3386.149902	2020-3-19	2409.389893
2020-2-20	3373.229980	2020-3-20	2304.919922
2020-2-21	3337.750000	2020-3-23	2237.399902
2020-2-24	3225.889893	2020-3-24	2447.330078
2020-2-25	3128.209961	2020-3-25	2475.560059

2020-2-26	3116.389893	2020-3-26	2630.070068
2020-2-27	2978.760010	2020-3-27	2541.469971
2020-2-28	2954.219971	2020-3-30	2626.649902

b) In the portfolio formation period, fit the SP500 price as a linear function of time, call this linear function $S^{ave}(t)$ and think it as the linear trend of SP500 in the portfolio formation period.





Generally speaking, there exists a liner relationship between time variable and the SP500 price. However, since our data is in a relatively short period, so the relationship is not much obvious. But the linear trend does exist, in this case, **S(t)** = -2.64416932 **t** + 3380.92

c) In the portfolio formation period, calculate Y(t) = S(t) - $S^{ave}(t)$, where S(t) is the SP500 price at time t. You can think Y(t) is the deviation from the SP500 trend. Fit a AR(1) model on ΔY(t) = Y(t)-Y(t-1) and see if there is a significant negative auto-correlation of lag 1.

	value	p-value
alpha	-1.597845	0.886110
beta	0.384673	0.096906

From the result of our AR(1) model, the p-value of each coefficient shows there is not a significant negative auto-correlation of lag 1.

d) Extract a continuous time mean reverting model for Y using the AR(1) model for ΔY(t). Use the sample volatility of Y(t) in the portfolio formation period as the volatility parameter in your model.

From AR(1) model for ΔY(t) we have

$$\Delta Y_t = \alpha + \beta \Delta Y_{t-1}$$

for a continuous time mean reverting model for Y

$$dY_t = k(\theta - Y_{t-1})dt + \sigma dW_t$$

$$dY_{t-1} = k(\theta - Y_{t-2})dt + \sigma dW_{t-1}$$

then we have

$$Y_t - Y_{t-2} = k(Y_{t-2} - Y_{t-1})dt + \sigma(dW_t - dW_{t-1})$$

$$Y_t - Y_{t-1} = (1 - kdt)(Y_{t-1} - Y_{t-2}) + \epsilon$$

$$\epsilon = dW_t - dW_{t-1}$$

$$dY_t = (1 - kdt)dY_{t-1} + \epsilon$$

So by approximation,

$$(1 - kdt) = \beta$$

$$\alpha \approx 0$$

$$\theta \approx \frac{1}{T}\sum_{t=1}^{T}[S_t - S_{t-1} + kS_{t-1}\Delta t]\frac{1}{kdt}$$

In our case, by using the sample volatility of Y(t) in the portfolio formation period as the volatility parameter σ in the mean reverting model,

$$kdt = 1 - \beta = 0.62$$

$$\theta = 15.98$$

$$\sigma = 106.32$$

$$dY_t = 0.62(15.98 - Y_t) + 106.32dW_t$$

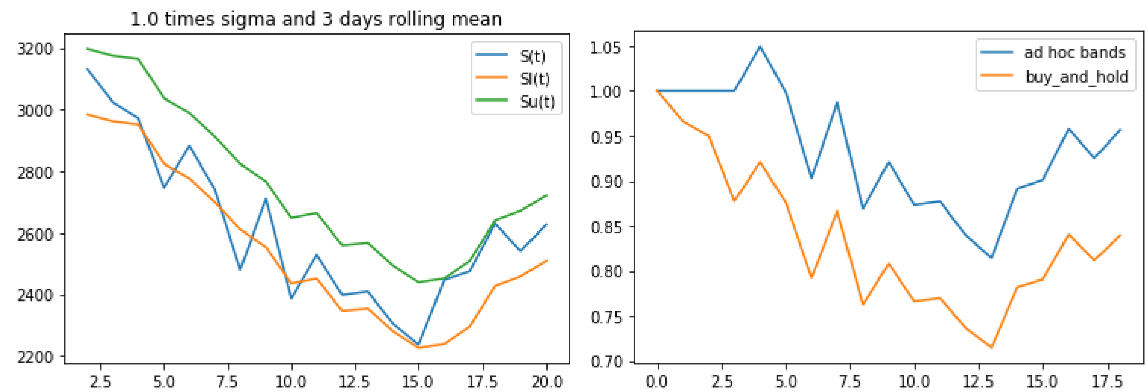
3. Trading strategy using ad-hoc bands

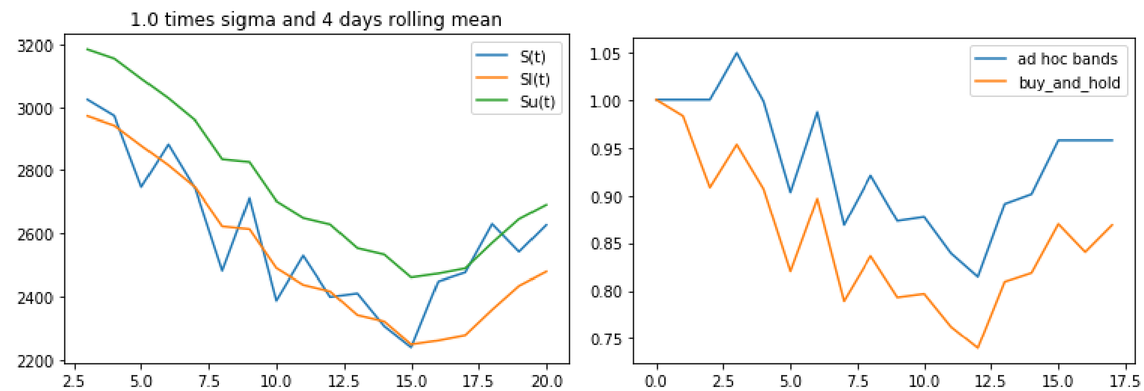
Design trading strategies by choose several ad-hoc bands using the volatility of Y(t) in the portfolio formation period. Call the lower bound Yl and the upper bound Yu. Test these trading strategies using data in the portfolio testing period. In the testing period, you can calculate $S^{ave}(t)$ in a rolling window (for example, average of SP500 prices in the last three days). When S(t) reaches $S^{ave}(t)$ + Yl, buy SP500; when S^ave(t) + Yu is reached, sell SP500.

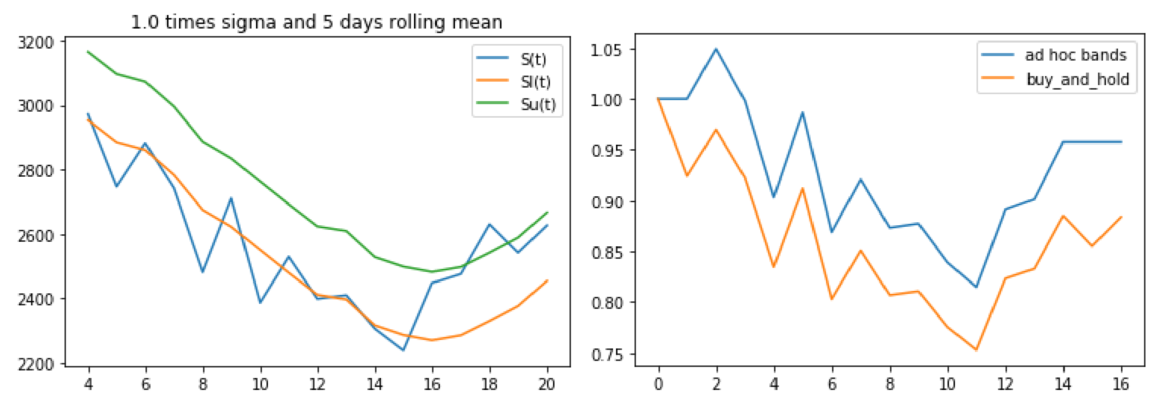
Please see the attachment for the detail of the code of the ad-hoc bands strategies.

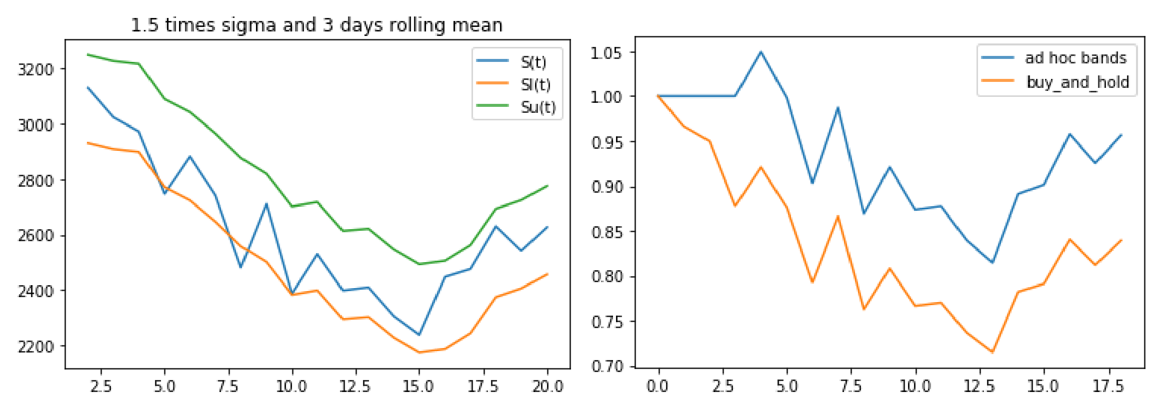
The initial capital is assumed to be $1,000. For different choice of ad-hoc bands, Calculate the cumulative P&L of your strategy at the end of the portfolio testing period. When the end of the portfolio testing period is reached, you are forced to close your position. Compare the performance of your strategies and also against the buy and hold strategy in the same period of time using mean, variance and Sharpe ratio of P&Ls.

The left figure is the daily track of price, the right figure is the cumulative P&L of all strategies.

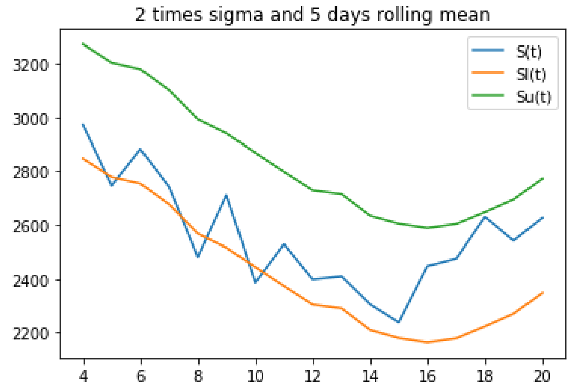

1.0 times sigma and 3 days rolling mean
S(t)
Sl(t)
Su(t)
ad hoc bands
buy_and_hold

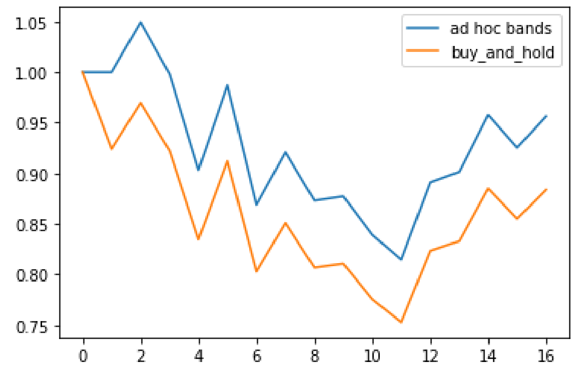

1.0 times sigma and 4 days rolling mean
S(t)
Sl(t)
Su(t)
ad hoc bands
buy_and_hold


1.0 times sigma and 5 days rolling mean
S(t)
Sl(t)
Su(t)
ad hoc bands
buy_and_hold


1.5 times sigma and 3 days rolling mean
S(t)
Sl(t)
Su(t)
ad hoc bands
buy_and_hold


1.5 times sigma and 4 days rolling mean
S(t)
Sl(t)
Su(t)
ad hoc bands
buy_and_hold
1.5 times sigma and 5 days rolling mean
S(t)
Sl(t)
Su(t)
ad hoc bands
buy_and_hold
2 times sigma and 3 days rolling mean
S(t)
Sl(t)
Su(t)
ad hoc bands
buy_and_hold
2 times sigma and 4 days rolling mean
S(t)
Sl(t)
Su(t)
ad hoc bands
buy_and_hold

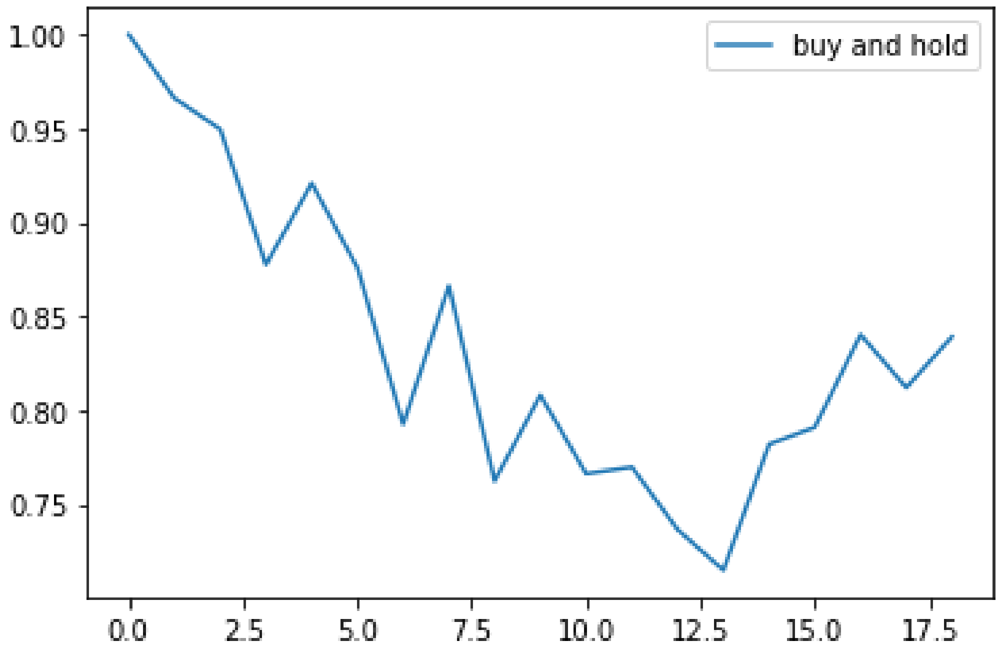




The cumulative P&L of strategies and the mean, variance and Sharpe ratio of P&Ls are shown in the following.

sigma	win	P&L	mean	std	SR
1.0	3.0	-0.043658	-0.182634	0.923993	-0.197657
1.0	4.0	-0.042413	-0.190382	0.933010	-0.204051
1.0	5.0	-0.042413	-0.201581	0.961719	-0.209604
1.5	3.0	-0.043658	-0.182634	0.923993	-0.197657
1.5	4.0	-0.042413	-0.190382	0.933010	-0.204051
1.5	5.0	-0.042413	-0.201581	0.961719	-0.209604
2.0	3.0	0.058860	1.071859	0.809227	1.324548
2.0	4.0	-0.043658	-0.192781	0.950777	-0.202761
2.0	5.0	-0.043658	-0.204121	0.980034	-0.208279

Besides, The cumulative P&L of buy and hold strategy in the same period of time and the mean, variance and Sharpe ratio of P&L are the following,

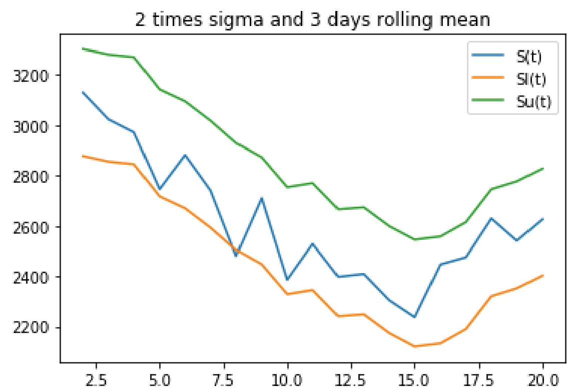


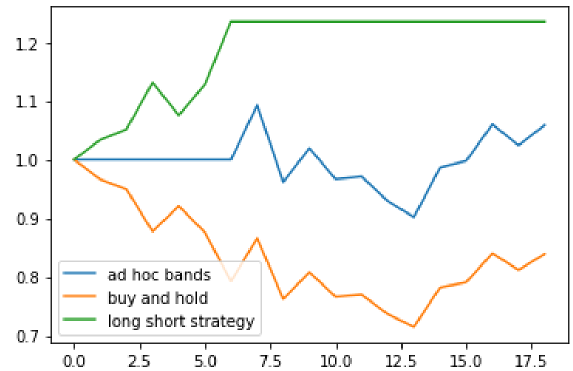
P&L	mean	std	SR
-0.160847	-1.866334	0.969688	-1.924675

In sum, the performance of our strategies depend on the choice of power of volatility of Y(t) and the rolling window size. In some cases, it will beat the buy-and-hold strategy, while in other cases it will not. Finally, we identify 2 times of volatility of Y(t) and 3 days window for rolling average in S^ave (t) be the best ad-hoc bands which has better mean, variance and Sharpe ration of P&L than the buy-and-hold strategy.

For the best ad-hoc bands that you identify, short $1000 value of SP500 (say X units of SP500) at the beginning of the portfolio testing period and use the $1000 to construct your strategy, buy back X units of SP500 at the end of the portfolio testing period. What is the performance of your long-short strategy?

From previous result, we identify 2 times volatility in Y(t) and 3 days window for rolling average in $S^{ave}(t)$ be the best ad-hoc bands. Then the performance of our long-short strategy are shown below





The left figure is the daily track of price, the right figure is the cumulative P&L of all three strategies. And the following table shows their mean, variance and Sharpe ratio.

	P&L	mean	std	SR
ad_bands	0.058860	1.071859	0.809227	1.324548
buy_hold	-0.160847	-1.866334	0.969688	-1.924675
long_short	0.219707	2.938193	0.508983	5.772675

To conclude, the performance of log-short strategy is relative good, for it manage to keep short position in the decreasing trend and identify a long signal, or say, market reversion signal, and close out its short position. Although the identified signal is not the best one, but it works.

4. Trading strategy using optimal bands

a) In the portfolio formation period, consider S(t) = Y(t) + $S^{ave}(t)$ as a model for the SP500 price. Here $S^{ave}(t)$ comes from 2 (b) and Y(t) is obtained from 2 (d). Take the end of the portfolio formation period as the investment horizon. Assume the discounting factor is zero. Write down the free boundary equations for the exiting and entering problems. Both these free boundary equations for functions with respect to t and S(t).

The HJB equation for optimal exit problem in finite horizon is

$$max\{(\partial_t + \mathcal{A})H(S_t, t), H(S_t, t) - (S_t - c)\} = 0$$

where the infinitesimal generator in this case is defined as

$$\mathcal{A} = (\beta + k(\theta - St - \alpha - \beta t))\partial_s + \frac{1}{2}\sigma^2\partial_{ss}^2$$

with terminal condition

$$H(S, T) = S - c$$

since the discounting factor is zero, so the free boundary equations for the optimal exit problem in finite horizon are

$$(\partial_t + \mathcal{A})H(S_t, t) = 0, \; S(t) < S^*(t)$$

$$H(S_t, t) = S(t) - c,\ S(t) \geq S^*(t)$$

$$\partial_s H(S_t, t) = 1,\ S(t) = S^*(t)$$

Similarly, the HJB equation for optimal entry problem in finite horizon is

$$max\{(\partial_t + A)G(S_t, t), G(S_t, t) - S_t - c + H(S_t, t)\} = 0$$

with terminal condition

$$G(S, T) = -2c$$

the free boundary equations for the optimal exit problem in finite horizon are

$$(\partial_t + \mathcal{A})G(S_t, t) = 0,\ S(t) < S^*(t)$$

$$G(S_t, t) = H(S_t, t) - S(t) - c,\ S(t) \geq S^*(t)$$

$$\partial_s G(S_t, t) = 1,\ S(t) = S^*(t)$$

b) Solve these two free boundary PDEs using finite difference methods (this is similar to American option pricing problem on a finite horizon).

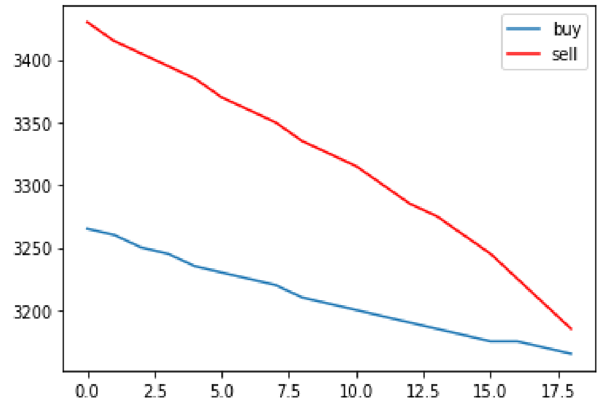
$$\begin{aligned} H\big(s_i, t_{j+1}\big) = \quad & \left(1 + \sigma^2 \frac{h_t}{h_s^2}\right) H\big(s_i, t_j\big) \\ & -\frac{1}{2}\left(\sigma^2 \frac{h_t}{h_s^2} - \frac{\beta + k(\theta - s_i - \alpha - \beta t_i)h_t}{h_s}\right) H\big(s_{i-1}, t_j\big) \\ & -\frac{1}{2}\left(\sigma^2 \frac{h_t}{h_s^2} + \frac{\beta + k(\theta - s_i - \alpha - \beta t_i)h_t}{h_s}\right) H\big(s_{i+1}, t_j\big) \end{aligned}$$

$$\begin{aligned} G\big(s_i, t_{j+1}\big) = \quad & \left(1 + \sigma^2 \frac{h_t}{h_s^2}\right) G\big(s_i, t_j\big) \\ & -\frac{1}{2}\left(\sigma^2 \frac{h_t}{h_s^2} - \frac{\beta + k(\theta - s_i - \alpha - \beta t_i)h_t}{h_s}\right) G\big(s_{i-1}, t_j\big) \\ & -\frac{1}{2}\left(\sigma^2 \frac{h_t}{h_s^2} + \frac{\beta + k(\theta - s_i - \alpha - \beta t_i)h_t}{h_s}\right) G\big(s_{i+1}, t_j\big) \end{aligned}$$

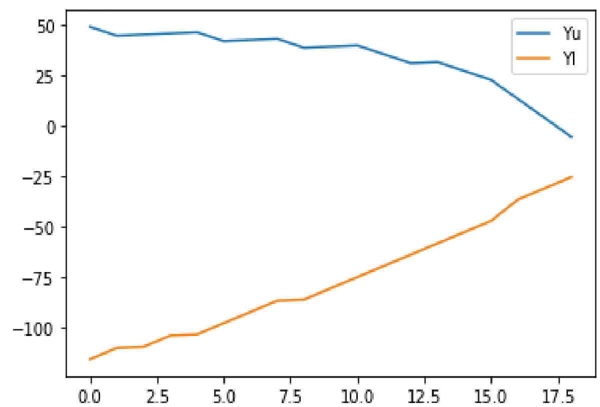
Using finite difference method with implicit scheme, we solve the previous two free bounday PDEs, please see the attachment for detailed code. The result are shown in following questions.

c) Find the optimal entering boundary Sl(t) and the optimal exiting boundary Su(t). Since we work with a finite horizon problem, both boundaries are time dependent. Using these two boundaries, we can calculate boundaries for Y(t) via Yl(t) = Sl(t) - $\boldsymbol{S^{ave}(t)}$ and Yu(t) = Su(t) -$\boldsymbol{S^{ave}(t)}$.

The optimal entering boundary Sl(t) and the optimal exiting boundary Su(t) are shown below,

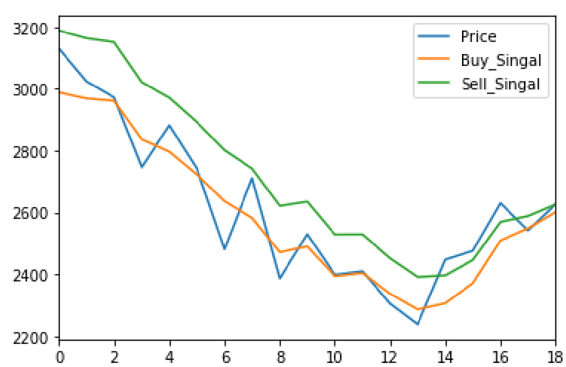


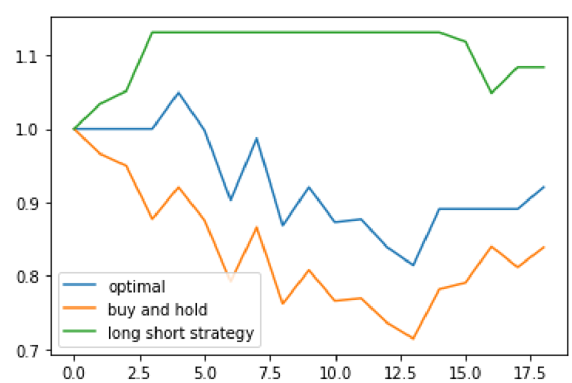
by subtraction S^ave(t), we calculate boundaries for Y(t) as following,



d) In the portfolio testing period, calculate $S^{ave}(t)$ in a rolling window as in #3. When S(t) reaches Yl(t) + $S^{ave}(t)$ buy SP500 and when S(t) reaches Yu(t) + $S^{ave}(t)$ sell SP500. Form a long-short strategy like #3, what is the performance of this strategy? P&L is calculated by the end of the portfolio testing period.

The performance, P&L, mean, variance and and Sharpe ratio of P&L are shown in the following.





	P&L	mean	std	SR
ad_bands	0.058860	1.071859	0.809227	1.324548

buy_hold	-0.160847	-1.866334	0.969688	-1.924675
long_short	0.219707	2.938193	0.508983	5.772675

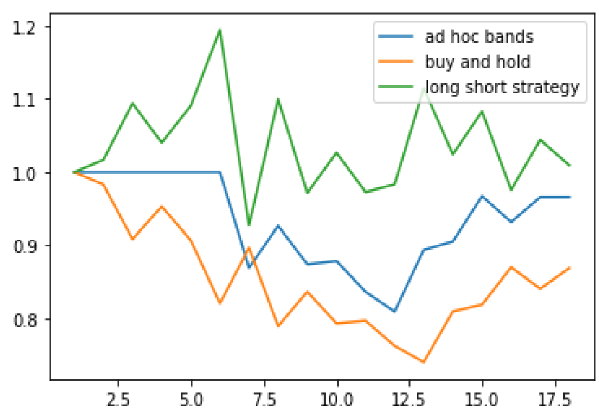
The performance of this strategy is also profitable, as seen from the mean, variance and Sharpe ratio. However, since the constructed optimal bands involving rolling average, so it is somewhat biased. As a result, our long-short strategy close out short position earlier and re-enter short position earlier. So its profit mainly comes from the initial period when the market is slumping but the short position is still kept.

5. Delay in execution

For the strategies constructed in #3 and #4, if execution (buy and sell) is delayed by 1 day (typical case when trading mutual funds), what is the performance of these long-short strategies?

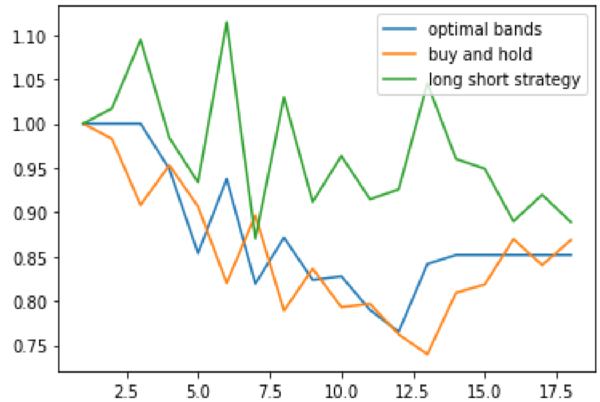
Either the strategy in #3 or the strategy in #4 fails to maintain its original relatively good performance. This is reasonable since we use rolling average when constructing the strategy. The rolling average may hide some information, for example, if the stock price decreases for the previous two days but reverts in the next day. The rolling average of these three days may all exhibit decreasing trend. But if we delay our execution from the second day to the third day to buy, then we miss the lowest price and buy assets with higher cost.

Delay in execution of strategy in #3



	P&L	mean	std	SR
ad_bands	-0.034011	-0.151622	0.832449	-0.182140
buy_hold	-0.131382	-1.495110	0.992272	-1.506754
long_short	0.097370	1.343488	1.577054	0.851897

Delay in execution of strategy in #4



	P&L	mean	std	SR
optimal_bands	-0.148065	-1.826190	0.923517	-1.977429
buy_hold	-0.131382	-1.495110	0.992272	-1.506754
long_short	-0.016684	-0.331079	1.662826	-0.199107

Teamwork

This paragraph is to confirm that each person in the team contribute equally for this project. The exact tasks for each team member are allocated in the following way. To maintain accuracy of the project, for problems #1 and #2, each member works separately (write their own code and formulation) and verifies the result with each other. For problem #3 and #4, each team member is asked to pick one problem and write down his code first. Then team member send the code to each other to go through cross-validation and further modification. Problem #5 is just simply repeat the procedure of problem #3, so is initially done by the team member who is responsible for problem #3. All team member has written part of code and feel fair to the teamwork.