Project 1.C *Trade stocks to make as much money as possible.*

CEE 251L. Uncertainty, Design, and Optimization

Department of Civil & Environmental Engineering
Duke University
Henri P. Gavin and Steve Lattanzio, Duke CEE '07

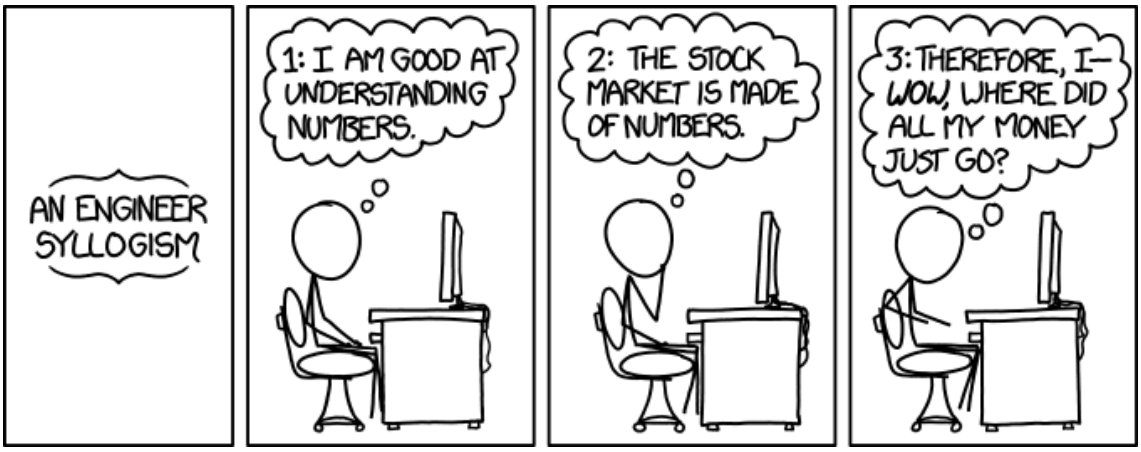
Due Monday, April 7, 2025

Think big, think positive, never show any sign of weakness. Always go for the throat. Buy low, sell high. Fear? That's the other guy's problem. Nothing you have ever experienced will prepare you for the absolute carnage you are about to witness. Super Bowl, World Series - they don't know what pressure is. In this building, it's either kill or be killed. You make no friends in the pits and you take no prisoners. One minute you're up half a million in soybeans and the next, boom, your kids don't go to college and they've repossessed your Bentley. Are you with me?

Dan Aykroyd as Louis Winthorpe III, to Eddie Murphy as Billy Ray Valentine, in Trading Places, 1983.

Thomas L. Friedman, Time to Reboot America, New York Times, December 24, 2008.

Greg Smith, Why I am Leaving Goldman Sachs, New York Times, March 14, 2012.



Randall Munroe, http://xkcd.com/1570/ Aug. 2015

---

1 Introduction

In this project, your group will design and optimize a methodology for selling and buying shares of stock. Naturally, your objective is to make as much money as possible! Groups will use 200 days of actual stock closing prices to design, test, and optimize an automatic stock trading method. Once groups are satisfied with the performance of their method, the method will be applied to the subsequent 200 days of stock closing prices. Ten points of the final grade will be apportioned according to how well your automatic stock trading algorithm works on the subsequent 200 days of stock closing prices.

The closing price of stock "s" on day "d" is denoted $p_{d,s}$.

Day-to-day changes in stock prices can be highly volatile; the change from yesterday to today can have very low correlation with the change from today to tomorrow. For this reason, it can be helpful to estimate trends over a recent period of time from a smoothed record of recent stock prices. One simple way to smooth the series of stock prices is to make a recursive (running) average of the closing prices of a particular stock. The day-to-day changes of the smoothed stock prices, and the rates of those changes, will have better day-to-day correlations, and might therefore serve as a better prediction for decision-making purposes.

Consider a sequence of numbers,

$$p_0 \ , \ p_1 \ , \ p_2 \ , ...$$

A running average

$$\bar{p}_0 \ , \ \bar{p}_1 \ , \ \bar{p}_2 \ , ...$$

of the data sequence can be defined as the weighted arithmetic average of the previous running average, $\bar{p}_k$, and the current data value, p_{k+1}.

$$\bar{p}_{k+1} = (1 - \phi) \ \bar{p}_k + \phi \ p_{k+1}$$

where ϕ is called the *forgetting factor* and $0 < \phi < 1$. If ϕ is close to one, $\bar{p}$ is close to p. If ϕ is close to zero, $\bar{p}$ varies more smoothly and with a longer time-lag.

Note that the running average at the k-th time step, $\bar{p}_k$, involves the entire data sequence, from p_0 and up to and including p_k, but older data contributes less and less to the running average. As a rule of thumb, the most recent $(5/\phi)$ points of data contribute significantly to the running average. For example, for $\phi = 0.2$, the most recent 25 points of data contribute significantly to the running average, and for $\phi = 0.8$, only the last 6 points contribute significantly to the running average.

For example, running averages of the sequence of 11 points: 20, 21, ... , 30 for various values of the forgetting factor ϕ are:

k	0	1	2	3	4	5	6	7	8	9	10
p_k	20	21	22	23	24	25	26	27	28	29	30
$\phi = 0.1$	20.000	20.100	20.290	20.561	20.905	21.314	21.783	22.305	22.874	23.487	24.138
$\phi = 0.1312$	20.000	20.131	20.376	20.720	21.151	21.656	22.225	22.852	23.527	24.245	25.000
$\phi = 0.5$	20.000	20.500	21.250	22.125	23.062	24.031	25.016	26.008	27.004	28.002	29.001
$\phi = 0.9$	20.000	20.900	21.890	22.889	23.889	24.889	25.889	26.889	27.889	28.889	29.889

2 Financial Data Analysis ... quantifying quality

The decision to buy and sell stock will be based on a stock statistic which we shall call "quality" $\mathcal{Q}$. Stocks with higher quality are better candidates for purchase. Stocks with lower quality are better candidates for sale. A measure of quality of stock s on day d ($\mathcal{Q}_{d,s}$) could be a weighted sum of the fractional change in a stock price, the rate of the fractional change of a stock price, and the volatility of a stock price over a recent period of time.

$$\mathcal{Q}_{d,s} = q_1 \ \dot{\bar{p}}_{d,s} + q_2 \ \ddot{\bar{p}}_{d,s} + q_3 \ \bar{v}_{d,s} \ , \tag{1}$$

Or, more terms could be added ...

$$\mathcal{Q}_{d,s} = q_1 \ \dot{\bar{p}}_{d,s} + q_2 \ \ddot{\bar{p}}_{d,s} + q_3 \ \bar{v}_{d,s} + q_4 \ \bar{v}_{d,s}\dot{\bar{p}}_{d,s} + q_5 \ \bar{v}_{d,s}\ddot{\bar{p}}_{d,s} \ , \tag{2}$$

where

- q_1, ... , q_5 are weighting coefficients with values to be found via optimization.

- $\bar{p}_{d,s}$ is a running average of the price of stock s on day d,

$$\bar{p}_{(d+1),s} = (1-\phi) \ \bar{p}_{d,s} + \phi \ p_{(d+1),s} \tag{3}$$

- $\dot{\bar{p}}_{d,s}$ is the fractional change of a running average of the price of stock s on day d. It is dimensionless.

$$\dot{\bar{p}}_{(d+1),s} = (\bar{p}_{(d+1),s} - \bar{p}_{(d-1),s}) \ / \ (\bar{p}_{(d+1),s} + \bar{p}_{(d-1),s}) \tag{4}$$

- $\ddot{\bar{p}}_{d,s}$ is the rate of the fractional change of the running average of stock s on day d. It is also dimensionless.

$$\ddot{\bar{p}}_{(d+1),s} = (\dot{\bar{p}}_{(d+1),s} - \dot{\bar{p}}_{(d-1),s}) \ / \ 2 \tag{5}$$

- $\bar{v}_{d,s}$ is the running average of the volatility of the fractional change of stock s on day d

$$\bar{v}^2_{(d+1),s} = (1-\phi) \ \bar{v}^2_{d,s} + \phi \ \left(\dot{\bar{p}}_{(d+1),s}\right)^2 \ . \tag{6}$$

  $\bar{v}^2_{d,s}$ is a running variance of $\dot{\bar{p}}$.

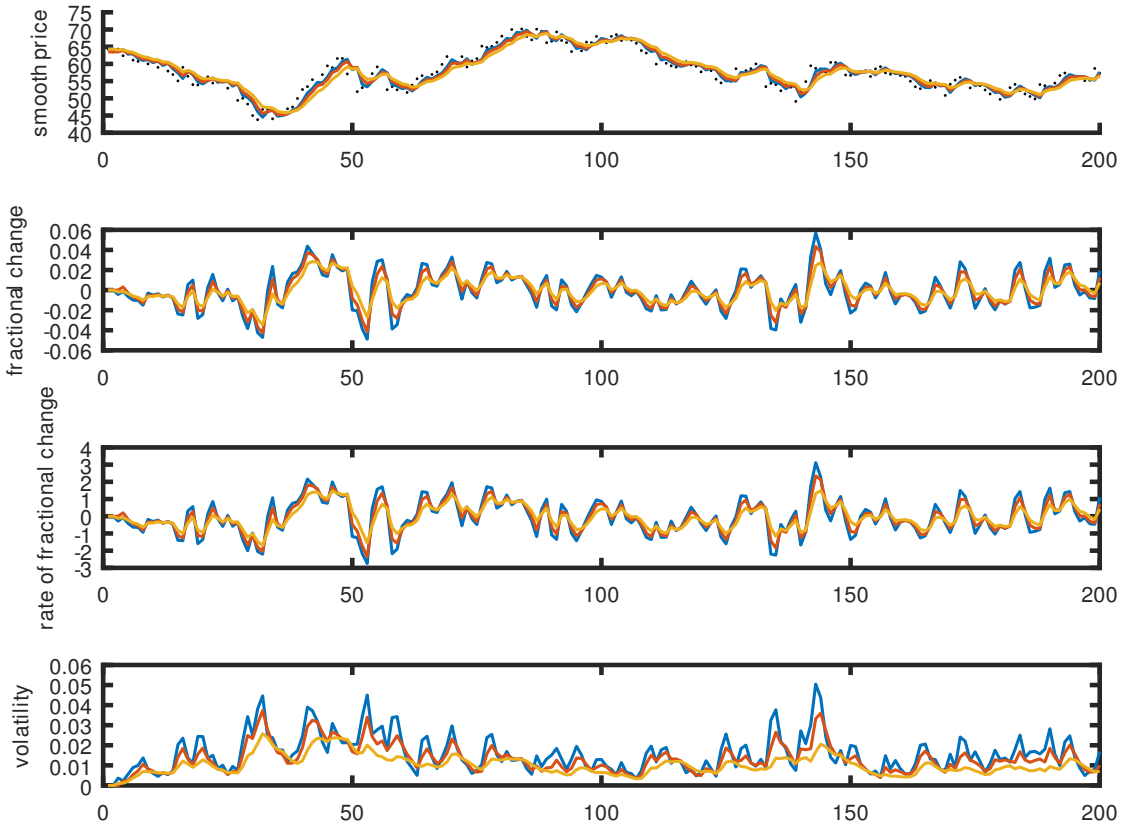


Figure 1. Effect of the forgetting factor ϕ on the running average of stock prices, the fractional change of smoothed price, and the rate of fractional change of smoothed price, for stock #8 and $\phi = 0.7$ (blue), $\phi = 0.5$ (red), and $\phi = 0.3$ (orange).

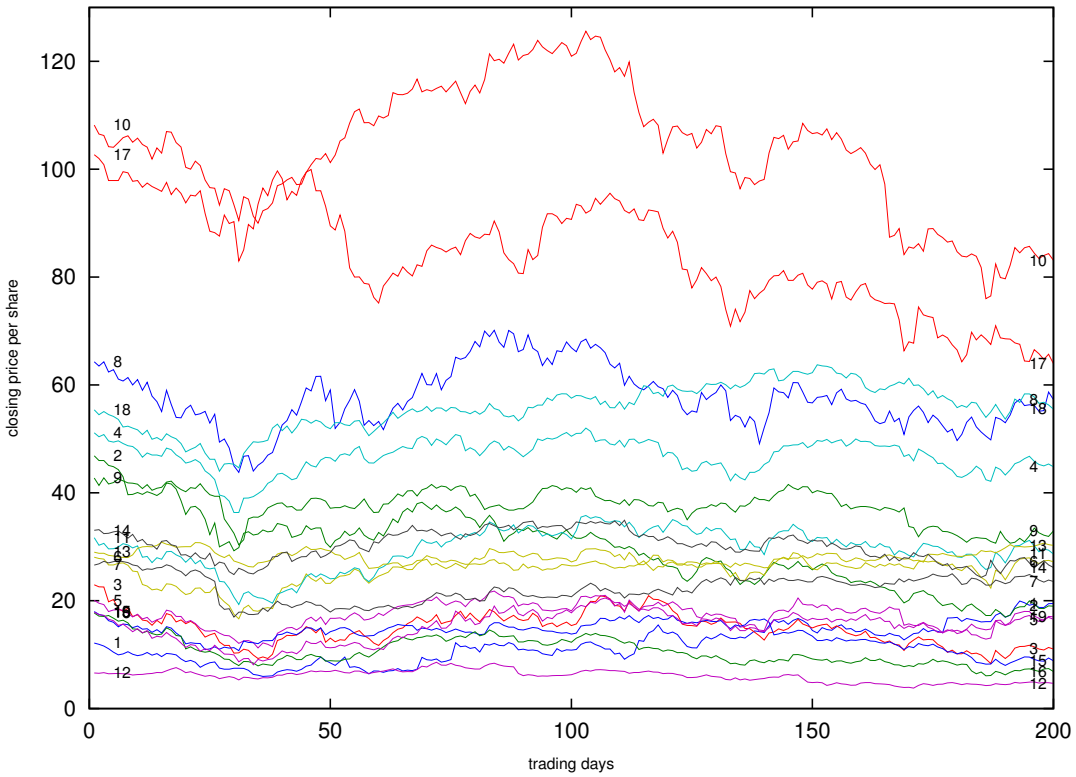


Figure 2. 200 days of stock closing prices for nineteen stocks

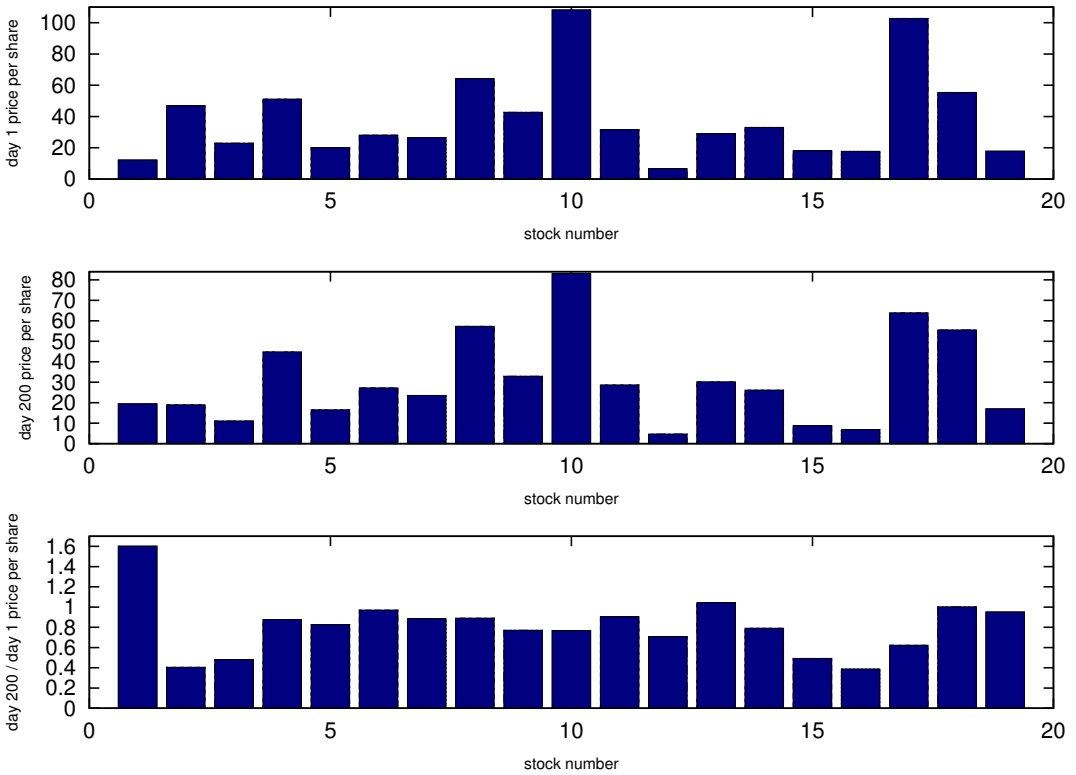


Figure 3. Initial prices, final price, and ratio of final-to-initial price

3 Stock Trading Decisions and Rules

If $\mathcal{Q}_{d,s} < S_d$ on day d then stock s is potentially *sellable*. If $\mathcal{Q}_{d,s} > B_d$ on day d then stock s is potentially *buyable*. The sell- and buy-thresholds S_d and B_d are variables that can change over time. The sell threshold S_d shall not be greater than the buy threshold B_d, so a stock can not be both potentially buyable and potentially sellable. The set of stocks to buy and sell are governed by the following rules:

- All shares of stock s are sold on day d if:
  - shares of stock s are owned (no shorting);
  - $\mathcal{Q}_{d,s} < S_d$ (Stock s is potentially sellable);
  - (optionally) $\mathcal{Q}_{d,\cdot} > B_d$, for another stock, not stock s,
    Shares of one stock may be sold if a different stock is potentially buyable.
- Shares of stock s are bought on day d if:
  - $\mathcal{Q}_{d,s} > B_d$, for any stock not sold on day d.
- Each time M_s shares of stock s are sold, the cash in hand, C, to purchase stocks is increased by the current value of the shares. A transaction cost of two dollars is paid for every stock sold or bought.

$$C_{d+1} = C_d + \sum_{s \in \mathsf{sold}} p_{s,d} M_s - 2.00 \left(\sum_{s \in \mathsf{sold}} 1 \right) - 2.00 \left(\sum_{s \in \mathsf{bought}} 1 \right) \tag{7}$$

- A fraction of the cash in hand ($0 < f_c < 1$) is then invested in equal parts in all of the stocks to be purchased.
- Each day the buy threshold B_d and the sell threshold S_d are updated to the maximum and minimum quality metric of all stocks owned on that day. Doing so attempts to ensure that future decision will only improve the investments. But this updating of the buy and sell thresholds does not guarantee that the value of the investments will grow with time.

4 Example

As an example, consider the optimization of a stock trading method that is calibrated for two-hundred days of closing prices for nineteen stocks, as shown in Figure 2. Figure 3 shows the "day 1" prices, the "day 200" prices and the ratio of the "day 200" prices to the "day 1" prices for each of the nineteen stocks.

Applying the methodology shown above with the following parameter values: $q_1 = -10$, $q_2 = 2$, $q_3 = -2$, $f_c = 0.9$, $\phi = 0.5$, $B_1 = 0.1$ and $S_1 = -0.1$ gives the results shown in Figures 4, 5, and 6.

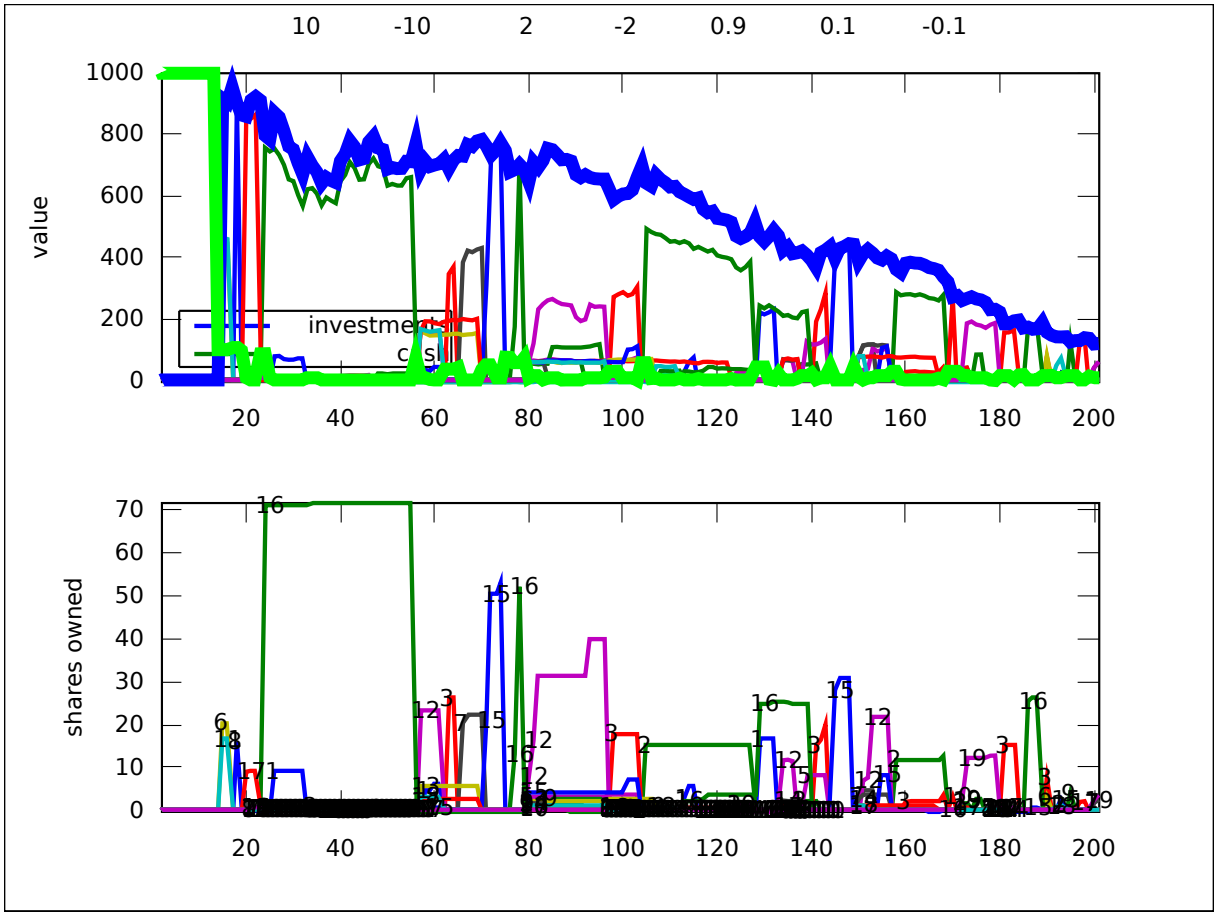


Figure 4. (a) Value of investments and value of cash on hand. (b) Number of shares owned over time.

The change of the value of the investments and the cash on hand is shown in the top part of Figure 4. Initially the cash on hand is \$1000.00. The value of cash and investments after 200 trading days is \$123.43[1]. The change in the shares owned is shown in the bottom part of Figure 4. A number of metrics for a select set of stocks is shown in Figure 5. The evolution of the buy threshold B_d and the sell threshold S_d is shown in Figure 6.

[1] no bonus for me this year

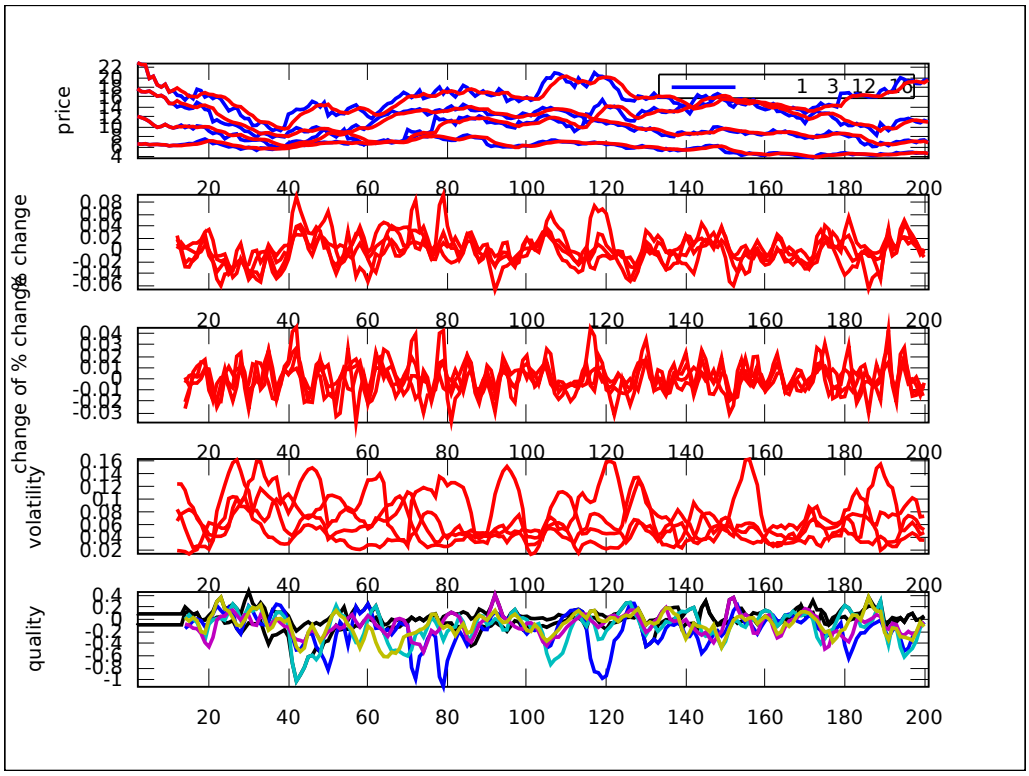


Figure 5. Price, fractional change, rate of fractional change, volatility, and quality of stocks 1, 3, 12, and 16

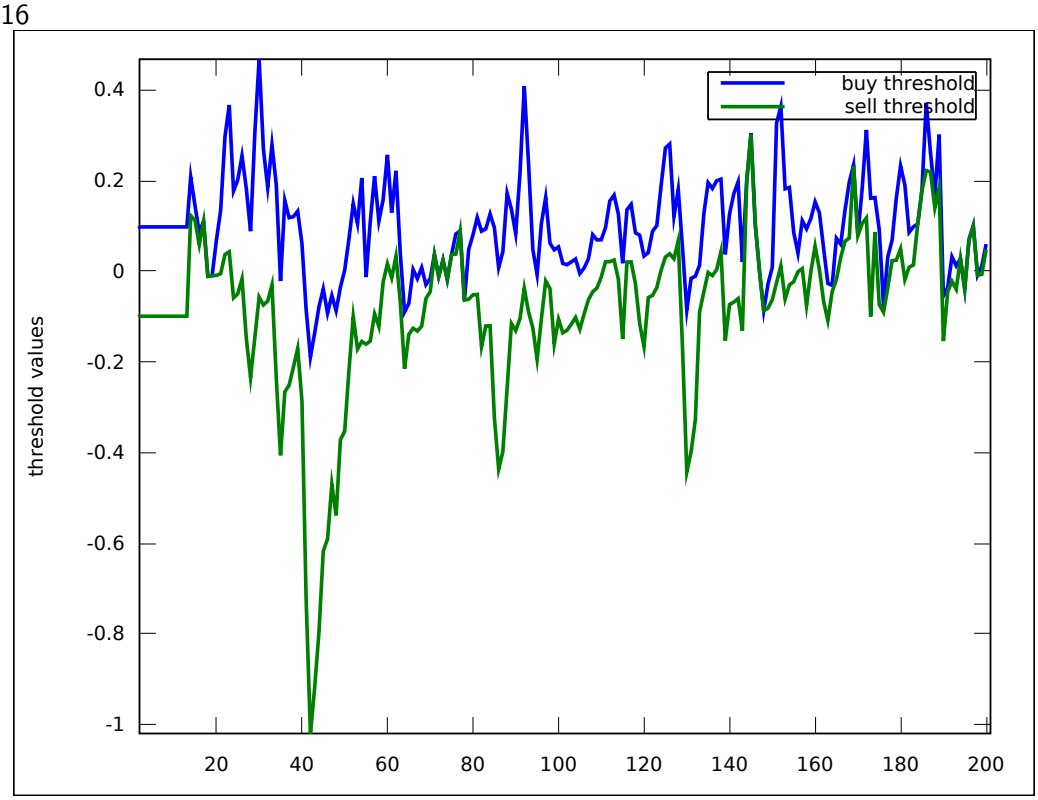


Figure 6. buy and sell thresholds, B_d and S_d.

5 Project Assignment

Study and adjust the method described in this document to make as much money as possible with the same data as used in the example. Download the data file of stock prices, stock_prices1.csv, and an example stock trading analysis program, exchange_analysis.m. If the stock price data file opens into a new browser window, you can right click in the window to save the data file.

```
1  function [ cost , constraint ] = exchange_analysis(param, constants )
2  % cost = exchange_analysis(param)
3  % analyze a policy for buying and selling stocks
4  %
5  % INPUT VARIABLES
6  % p(1)  = q(1)     .... quality veloc coeff
7  % p(2)  = q(2)     .... quality accel coeff
8  % p(3)  = q(3)     .... quality volatility coeff
9  % p(4)  = fc       .... fraction of cash to invest (must be between 0 and 1)
10 % p(5)  = phi      .... forgetting factor
11 % p(6)  = B        .... buy  threshold
12 % p(7)  = S        .... sell threshold
13 %
14 % OUTPUT VARIABLES
15 % cost             .... the negative of the value of cash and investments
16 %                       after 200 days of trading.
```

The stock trading methodology that your group develops will be parameterized by at least the following design variables: ϕ, q_1, q_2, f_c, B_1, and S_1, where B_1 is the buy threshold on "day 1" and $\hat{S}_1$ is the sell threshold on "day 1".

Tasks

1. Read and study the provided stock exchange program `exchange_analysis.m`.
2. To help study the method of running averages used in this project, reproduce Figure 1 of this document for stock number 5 and describe the effect of ϕ on $\bar{p}$, $\dot{\bar{p}}$ and $\ddot{\bar{p}}$ in a few sentences.
3. (Optional) You may change the stock trading algorithm by modifying `exchange_analysis.m` in five places:
   - The definitions of quality $\mathcal{Q}$ (lines 72-78);
   - The decision to sell stocks and hold cash (lines 80-84);
   - The distribution of purchased shares (lines 115-118);
   - The method used to update the buy and sell thresholds, B_d and S_d (lines 129-137); and
   - The smallest allowable difference between the thresholds, B_d and S_d (lines 142-143);
4. Using the code as provided (or as modified by your group in the task 3, above), optimize the values of $q_1, q_2, ..., f_c, \phi, B_1$, and S_1 and/or other parameters your group may have introduced in order to maximize value.

(a) Write a MATLAB script `exchange_opt.m` that will use the MATLAB function `exchange_analysis.m` along with optimization routines, such as `ORSopt.m` and/or `NMAopt.m`, to optimize the design variables. The first two lines of code in `exchange_opt.m` must load in the stock price data file as follows:

```
stock_prices = csvread('stock_prices1.csv'); % load stock price data
constants = { FigNo , stock_prices };        % FigNo is a figure number
```

(b) Try different initial guesses for the design variables, list your initial guesses, the resulting optimal values, and the resulting earnings. Are your optimal values and optimal earnings very sensitive to the initial guesses?

(c) Run a sensitivity analysis. Run separate analyses with each of your parameters perturbed by +5% individually. That is, for analysis "i" the "i-th" parameter value is 1.05 times its optimal value, and all the other parameters are at their optimal value. What are your earnings with each of the perturbed parameter sets? Did your earnings increase or decrease significantly? Which design variables have the greatest effect on your earnings?

(d) If an optimal solution shows $q_1 < 0$, $q_2 > 0$, and $q_3 > 0$, what inference can be drawn regarding investment policy?

5. After completing questions 1 through 4:

(a) Specify your recommended set of parameters: $q_1, q_2, ..., f_c, \phi, B_1$, and S_1. and any other parameters you may have introduced.

(b) Produce versions of Figure 4 and Figure 6 associated with your recommended set of parameters. You can do this by simply running
`my_cost = exchange_analysis( my_params , constants )`

(c) If your group changed the methodology in any way in question 3, describe your changes in words and show the lines of MATLAB your group wrote to replace the corresponding sections of `exchange_analysis.m`.

(d) What is the greatest amount of money your group was able to make?

I will apply your group's method to the set of closing stock prices for the 200 trading days subsequent to the 200 days that you used to optimize your system.

I will distribute versions of Figure 4 for each group's results.

Have fun and good luck!

6 Report contents

Each group of two or three people should prepare a report including:

- Your written responses to questions 2, 3, 4, and 5.
- Plots from questions 2 and 5.
- Print-outs of the MATLAB code you wrote.
  If your group changed `exchange_analysis.m` send your revised version to me in an e-mail attachment.
- A table of your final optimized design variable values:

q_1	q_2	...	f_c	ϕ	B_1	S_1	day 1-200 value	day 251-400 value

. (Leave the last column blank.)

7 Grading

- 10 points: The final value of cash and investments at the end of the first two-hundred days of stock prices,

$$\text{points} \;=\; \text{day 200 value} \;/\; 200$$

- 10 points: The final value of cash and investments at the end of the subsequent two-hundred days of stock prices,

$$\text{points} \;=\; \text{day 400 value} \;/\; 100$$

- 20 points: Quality of the report:
  (typed, 11pt or 12 pt, 1.5 space, 1 inch margins, page numbers, figure numbers, figure captions, etc.)
- 30 points: Completeness and correctness of descriptions and discussions
- 30 points: Answers to questions 2-5. (Question 3 is optional.)