MC3-P3

Chapman Siu

Introduction

The purpose of this report is to look at reinforcement learning and how it can be used to develop trading strategies. This report is in two parts:

1. Developing the model for reinforcement learning
2. Results and comparison of the reinforcement learning approach

Discussion

Part 1 - Building the Reinforcement Learning Model

As this question was fairly open ended, the definition of the states I used and the reward values are highly important.

States

The states I used were comprised of three components which were then discretised:

1. Bollinger Bands
2. RSI
3. Current position (grouped to be one of: short, no action, long)

The first and third option were discussed in the lecture and will not be repeated here.

Bollinger Bands were grouped by taking the value of “current value - SMA / 2 * rolling_std”:

- < -1; assigned a value of -1
- > 1; assigned a value of 1
- Everything else; assigned a value of 0

RSI is a momentum indicator which I used in my previous assignments (e.g. MC2P2) which intuitively measures the ratio of higher closes to lower closes as a measure of momentum. The rationale is when prices fall rapidly, it suggests that at some point it was oversold, conversely if it rises rapidly it suggests that it was at some point undervalued.

The equation for RSI is as follows[1]:

```
RSI = 100 - 100/RS
```

Where

```
RS = Average of x days' up closes / Average of x days' down closes
```

The number "x" was chosen to be 14 as suggested by J. Wilder[2]. Wilder also suggested that if RSI was over 70, then the stock was overbought, and if it was under 30 then it was oversold. Thus the groupings for RSI were:

- < 30; assigned a value of -1
- > 70; assigned a value of 1
- Everything else; assigned a value of 0

Discretizing the States

To discretize the states, it was done differently to the lectures. From above, I have placed each chosen feature to be in 3 groups. I then added the two features together to create a new feature which had 5 categories, -2, -1, 0, 1, 2. These would represent the "agreement" level between the two factors. This could be trivially extended if we were to add more factors.

With these 5 categories and 3 actions I have then created 15 possible states. I experimented with increasing the number of these states and adding additional indicators, unfortunately the performance turned out to be worse, than the result shown below.

Rewards

The reward value was the same as discussed in the lecture; the daily return on the portfolio.

Running the Model

To run the model with these states and actions, I made the following assumptions:

- The time period is 2007-12-31 to 2009-12-31
- all positions only held 100 stock
- Only IBM stock considered

Using the notation from the assignment page, the parameters used were:

- rar = 0.98
- radr = 0.9999

---

[1] The equation is based on the version as described on investopedia: http://www.investopedia.com/terms/r/rsi.asp

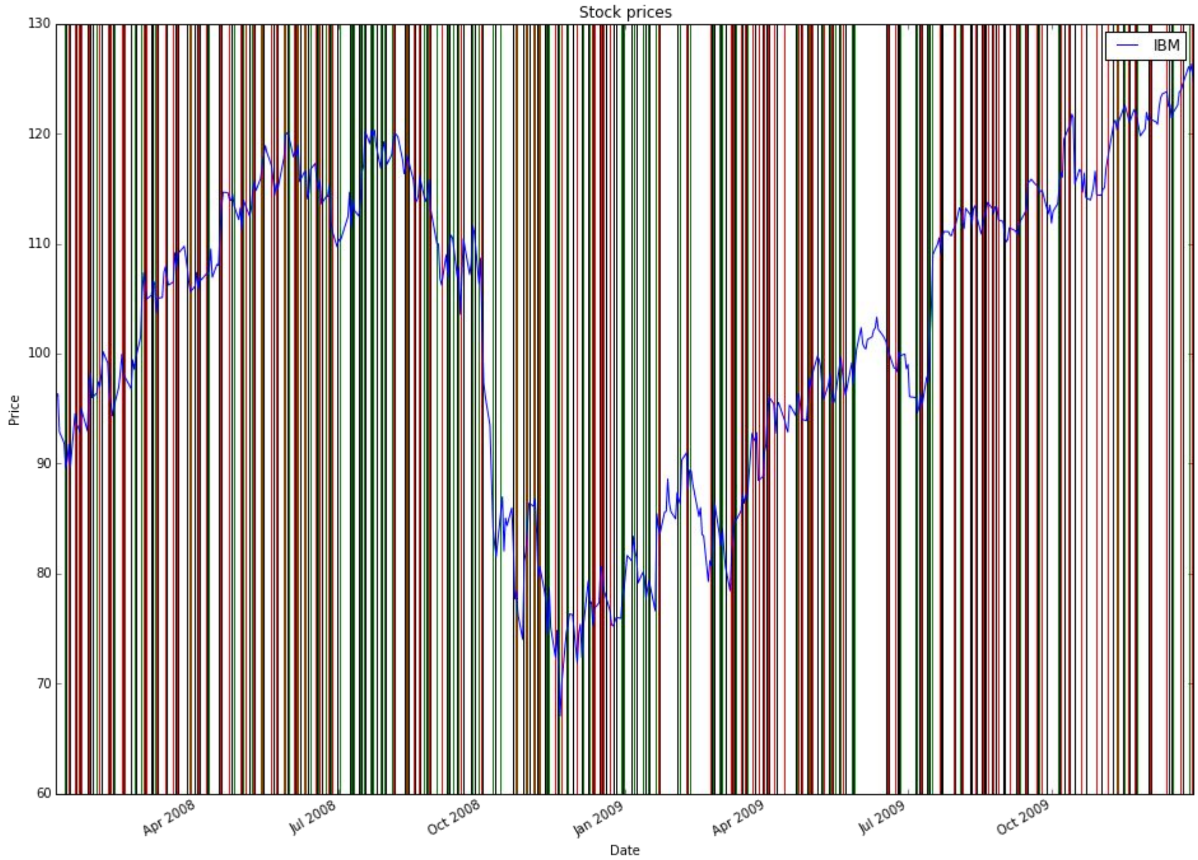
[2] This attribution isn't totally clear based on my research. Though it is expressed on this website: http://www.worden.com/TeleChartHelp/Content/Indicators/Wilders_RSI.htm

Part 2 - Results

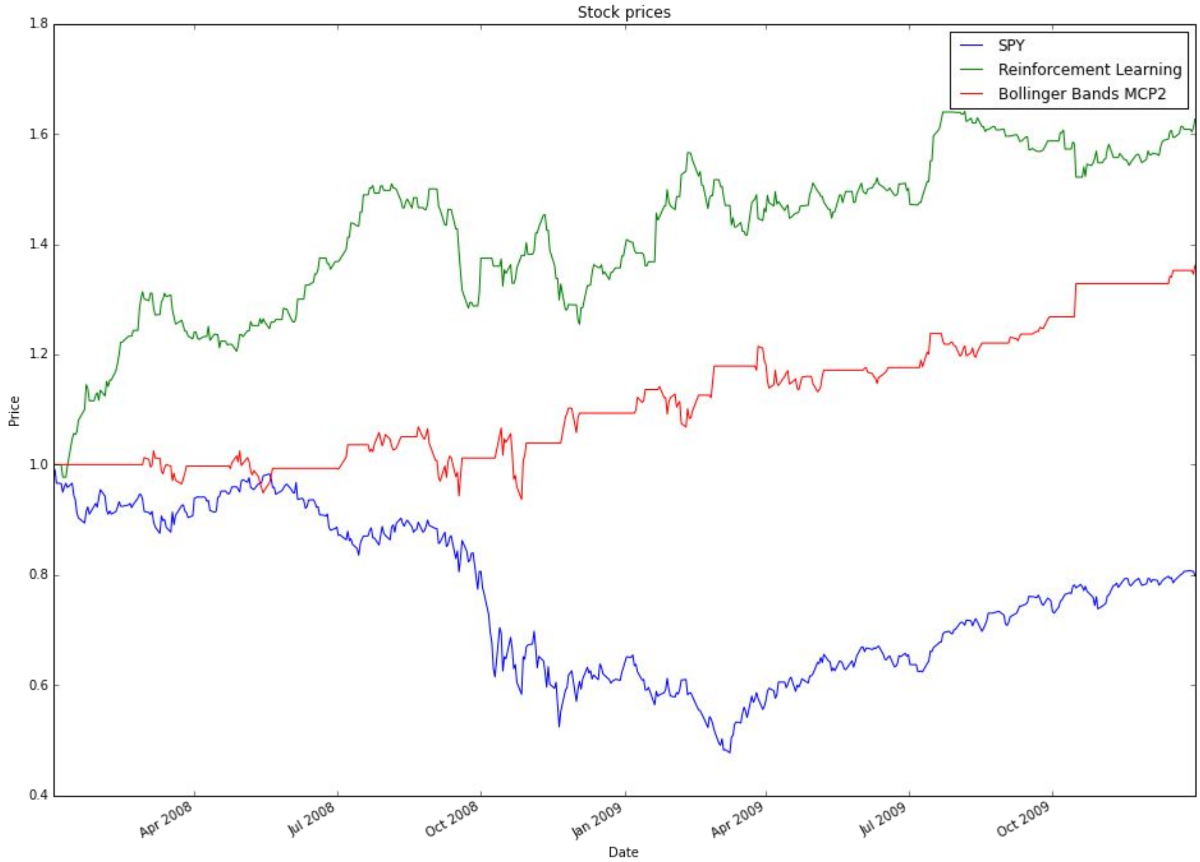
It would appear that the reinforcement learning strategy can at least beat the Bollinger Band Strategy which was described in MC2P2 assignment

```
Reinforcement Learning Strategy

Sharpe Ratio of Fund: 1.45406598596
Cumulative Return of Fund: 0.6269
Standard Deviation of Fund: 0.0112312516098
Average Daily Return of Fund: 0.00102875510113
```

```
Bollinger Band Strategy from MC2P2

Sharpe Ratio of Fund: 0.978427130372
Cumulative Return of Fund: 0.3614
Standard Deviation of Fund: 0.010891061211
Average Daily Return of Fund: 0.000671271818757
```



Above is the strategy chart where green is the long position, red is the short positions and black are the exit positions



Above is the backtest of the results.