Investment Strategy

Effanville

April 28, 2019

The point of this short note is to outline the desired manner without short selling by which money will be invested. The general structure takes inspiration from [1] and [2]. We introduce a mathematically rigorous grounding for the processes we go by in the investment strategy.

Recall that the notation for the moving average of the previous N values of a discrete function u_i is given by

$$A^N(u_i, t) = \frac{1}{N}\sum_{k=1}^{N} u_i(t-k). \tag{1}$$

1 Alpha Strategy

We have a collection of $K \in \mathbb{N}$ stocks, for example all those in the FTSE100. The purpose of the Alpha strategy is, for each stock, to give an investment decision. This is, for each $i = 1, \dots, K$, an outcome

$$\alpha \colon S_i(t) \mapsto \{\text{Buy}, \text{Sell}, \text{Hold}\} \tag{2}$$

and we desire this to be performed in a systematic manner, and using technical indicators. It is also sufficient to approximate the function

$$\text{Pred} \colon S_i(t) \times \text{TI} \mapsto S(t+j) \tag{3}$$

where $S_i(t)$ is the price history up to time t, TI is a space of technical indicators, and $S(t+j)$ is the price at time $t+j$. This function then gives a value of α, for example with

$$\alpha(P_i(t)) = \begin{cases} \text{Buy} & \text{Pred}(P_i(t), \cdot) > 1.05 \\ \text{Hold} & \text{Pred}(P_i(t), \cdot) \in [1, 1.05] \\ \text{Sell} & \text{Pred} < 1 \end{cases}$$

where these are relative values to the current price.

We assume we are able to harvest the following daily data:

- Opening price I_t
- Closing price C_t
- High H_t
- Low L_t
- The volume V_t of buying and selling,

We also assume we have complete data for this over some time period from $[t_1, P]$ where P is the present time. The remainder of this section is the process by which we approximate the function in equation (2).

1.1 Stock Statistics

In particular recall the following tests

Definition 1.1. *We define the* **Hodrick-Prescott Trend** *by* $\hat{\tau} = [I + \lambda K^T K]^{-1}x$ *where* I *is a* $k \times k$ *identity matrix and* K *is given by*

$$k_{ij} = \begin{cases} 1 & \textit{if } i = j \textit{ or } i = j+2 \\ -2 & \textit{if } i = j+1 \\ 0 & \textit{else.} \end{cases}$$

The standard parameters are $\lambda = 5000$, $n = 20$ and k is the length of the data for the stock.

Definition 1.2. *The* **Volume test** *is given by*

$$\mathrm{Vol}(S,t,n,m) = \begin{cases} \text{Strong} & A^n(V,t) > A^m(V,t) \\ \text{weak} & \textit{else} \end{cases}$$

where A^N is defined in equation (1)*, for some to be determined $n < m$. The defaults are $n = 2, m = 14$.*

Definition 1.3. *The* **ADX statistic** *is the following. Define*

$$\mathrm{DM}_t^+ = \begin{cases} \max\{H_t - H_{t-1}, 0\} & H_t - H_{t-1} \geq L_t - L_{t-1} \\ 0 & \textit{else} \end{cases}$$

$$\mathrm{DM}_t^- = \begin{cases} \max\{L_t - L_{t-1}, 0\} & H_t - H_{t-1} \leq L_t - L_{t-1} \\ 0 & \textit{else} \end{cases}$$

and then set $\mathrm{TR}_t = \max\{H_t, C_{t-1}\} - \min\{L_t - C_{t-1}\}$. Define $\mathrm{DI}_t^\pm = \mathrm{DM}_t^\pm / \mathrm{TR}_t$ and then we define

$$\mathrm{DX}_t(n) = \frac{A^n(\mathrm{DI}^+, t) - A^n(\mathrm{DI}_t^-, t)}{A^n(\mathrm{DI}^+, t) + A^n(\mathrm{DI}_t^-, t)}$$

The ADX statistic is then given by $\mathrm{ADX}_n = 100\, A^n(DX, t)$, where $n = 14$ is the usual value of this parameter.

Furthermore recall the timing indicators

Definition 1.4. *The* **stochastic statistic** *is the function D^n given by*

$$D^n(t) = A^3(A^3(K_n, \cdot), t) \qquad , K_n(t) = 100\frac{C_t - L_n}{H_n - L_n},$$

for $H_n = \max_{i=t-n,\dots,t-1} H(i)$ the highest price of the previous n days, and $L_n = \min_{i=t-n,\dots,t-1} L(i)$ is the lowest price of the previous n days. Usually choose $n = 14$.

1.2 Decision system

The natural manner in which to approximate the function specified in (3) is via a regression based approach. We set this up in the following manner. We assume that the prediction $\mathrm{Pred}(P(t))$ can be written as a linear combination of the prices and volumes of the previous 10 days, and of various features such as moving averages of these.

We denote by β the vector of coefficients, X the matrix of all possible "normalised" data, and y the vector of the open price after $(t + j)$ days. We then look for the values of β that minimise

$$\left\| y_i - X^i \beta_i \right\|_2^2 + \lambda \left\| \beta_i \right\|_1$$

for $i = 1, \dots, N$. Here, the second term penalises too many large β terms and so reduces the chances of over-fitting, dependent upon some the size of $\lambda \geq 0$. Remark that $\lambda = 0$ is least squares regression.

The normalisation of the data we take is non-standard. All prices used are divided by I_t, and all volumes used are divided by V_t, thereby providing an analysis of the relative change for each stock compared to the last opening data.

Observe that λ is a parameter to be specified by the model. We use a process of cross validation to chose the optimal parameter value.

To account for the massive number of potential variables one has, we use the following 12 features at time t:

1. H_t/I_t
2. L_t/I_t
3. C_t/I_t
4. HP
5. $DX_t(14)$
6. $A^5(I,t)/I_t$
7. $A^{20}(I,t)/I_t$
8. $A^{50}(I,t)/I_t$
9. $D^{14}(t)$
10. $D^{28}(t)$
11. $A^2(V,t)/V_t$
12. $A^{14}(V,t)/V_t$.

We then use the Lasso regression process above to produce the optimal coefficients of these values to predict the value after j days.

Remark 1.5. *We have many choices as to what features to include the model. To identify possible dependencies we could first perform correlation analysis on the data set. This however increases the risk of over-fitting.*

2 Portfolio Building and Trading

The Alpha strategy provides a function which generates a buy/sell decision for a stock. We then from this must decide daily how to modify the current portfolio.

We suppose that we currently have invested in portfolio $\mathcal{P}$, and that we have FC currently not invested.

1. First we sell those stocks i that we hold for which $\alpha(i) = \text{Sell}$
2. We order the stocks for which $\alpha(i) = \text{Buy}$ based on the ADX_n value, as given in Definition 1.3.
3. We invest an amount
$$\min\left\{f\,\text{val}(\mathcal{P}),\, FC,\, 20000\right\}$$
in each stock in the list to buy.
4. While $FC > 0$ sequentially invest in these stocks in the order given above.

3 Simulation

The algorithm provides trading decisions. To test, we need to specify the timing of when these decisions are carried out, how much money is exchanged in each trade, and which stocks we require decisions upon.

We specify that we purchase each stock first thing in the morning, and that the price we pay is given by the following algorithm:

1. Prior to test run specify a parameter $\gamma\,(= 0.03)$,
2. For each transaction, generate a random number $s \in \{-1, 1\}$,
3. On day t, Buy/Sell at the price given by $I_t\,(1 + \gamma\, s)$

This simulates reality in that one does not know the transaction price until the transaction, and adds uncertainty due to the randomness of the price.

We then obtain the data as said at the start of section 1 for the stocks in the FTSE100 from 2013 to the end of 2018, and split this into two sections, the first from 2005 to 2008, and the second from 2008 until the end of 2018. The former is considered training data, and the latter the test simulation.

We use as a benchmark an equal weighting of all the stocks in the list.

3.1 Quality Statistics

The above specifies a two parameter family of simulations, denoted (f, j), with parameters j for prediction day ahead of current time, and f the fraction of the total book value to invest. We thus require statistics to determine which parameters are the best.

Since we are aiming to maximise return, we choose to maximise the compound annual rate of the investment. If we have invested for N years, the CAR is

$$CAR = \left(\frac{\text{Val}(N)}{\text{Val}(0)}\right)^{1/N} - 1.$$

To analyse the consistency of returns, we also look at the Sharpe ratio. This ratio is, for prices $Y_1, \ldots, Y_n$ over the n years of the simulation,

$$\frac{\text{Mean}\,(Y_i)}{\text{STD}\,(Y_i)}.$$

This measures return but aims to compare it to how much risk was present when one obtained that return. It should be noted that a higher Sharpe ratio is better, as this means either the mean of the returns was higher, i.e. made more money, or the variance of the returns was lower, i.e. made money in a less risky manner.

3.2 Test Run

As can be seen in Appendix A, the predictions for further into the future were better, both in terms of having a higher CAR, and with having a higher Sharpe ratio, although this is not the case for a 10 day in the future prediction.

To highlight the typical behaviour, we analyse the $(0.5, 9)$ model, see figure 1. One should here observe the tendency to have a sudden increase, followed by a period of stagnation, which suggests increases coming from relatively few investments. Furthermore, one shoud note the significant fall in portfolio value around middle to late 2015, which is at the end of a three year fall in portfolio value. To add to this, since the benchmark increases in this period, this fall is not due to a global fall, but due to poor decisions of the algorithm. This is clearly not desirable behaviour.

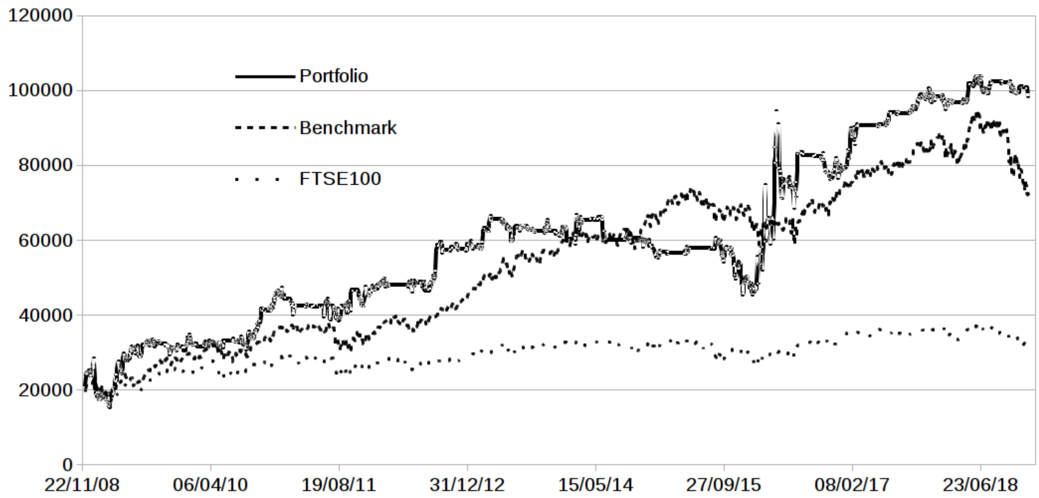

Figure 1: Portfolio value for the $(0.5, 8)$-parameter model

Furthermore, a hybrid model, where one averages the β values for a number of future day predictions was created. Averaging the predictions over days 5 to 9 into the future gives a return of 14% and a Sharpe ratio of 0.4. For example, the run with this hybrid prediction and investing 0.3 is shown in figure 2. As can be seen, the strategy outperforms the FTSE100 significantly, and also outperforms an equal weighting benchmark. However, most of the returns are made in the early stage of the simulation, with a massive increase after a few months, and relatively little increase thereafter. This is not so desirable.

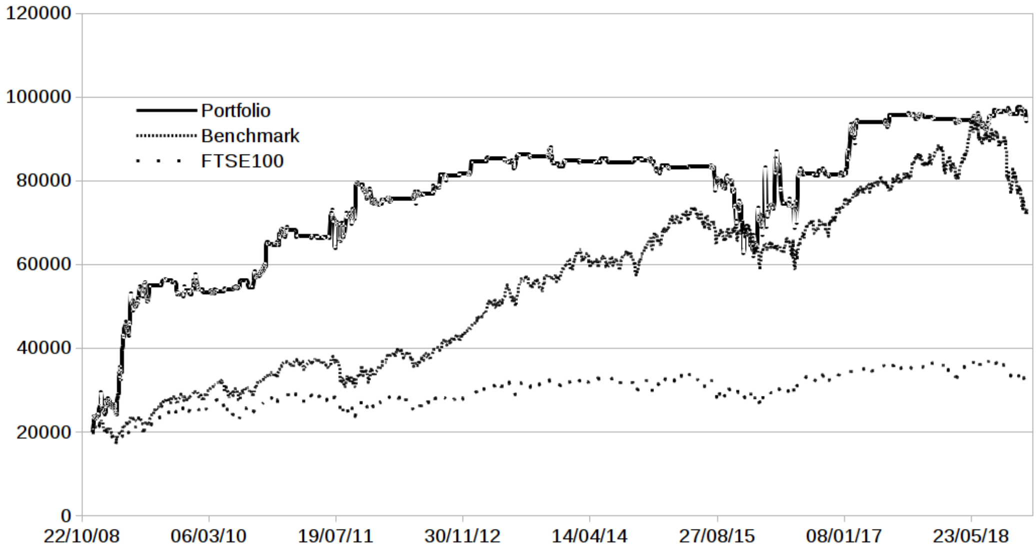

Figure 2: Portfolio value for the 0.3-hybrid model

4 Outlook

We will analyse in more depth the decisions and why they are not so suitable, with a view to providing a better prediction system, and approximation to Pred.

References

[1] J I Larsen. Predicting stock prices using technical analysis and machine learning. Master's thesis, Norwegian University of Science and Technology, 2010.

[2] RK Narang. *Inside the Black Box: A Simple Guide to Quantitative and High Frequency Trading*, volume 883. John Wiley & Sons, 2013.

A Training Data Results

	0	1	2	3	4	5	6	7	8	9	10
0.1	4.59	4.22	7.04	5.00	7.62	9.04	7.79	10.90	10.74	10.9	8.51
0.2	6.41	6.49	6.66	6.70	8.77	12.25	11.78	16.46	19.06	12.8	14.4
0.3	7.09	6.09	6.34	4.55	9.84	11.76	13.56	19.76	18.1	16.7	19.4
0.4	5.13	6.54	2.23	4.25	9.36	13.14	13.99	19.75	18.58	18	20.4
0.5	10.98	12.13	0.25	7.98	8.12	14.37	15.43	19.93	17.23	15.4	21.1
0.6	9.71	2.85	-0.7	5.97	11.10	15.60	14.45	21.26	18.06	17.6	22.4
0.7	6.09	3.14	-1.7	8.63	12.55	13.71	13.45	21.27	17.88	17.28	19.9
0.8	9.34	4.71	-2.7	6.38	9.62	15.98	14.96	21.79	19.2	12.7	22.4
0.9	8.05	8.67	-0.9	8.04	9.15	8.65	11.85	19.73	21.05	18.1	19.9
1	7.62	7.28	-2.8	4.20	11.87	12.23	15.28	20.95	20.10	17.8	22.4

Table 1: CAR results

	0	1	2	3	4	5	6	7	8	9	10
0.1	0.14	0.23	0.17	0.15	0.20	0.20	0.20	0.26	0.31	0.22	0.15
0.2	0.16	0.27	0.18	0.15	0.22	0.25	0.23	0.40	0.31	0.36	0.23
0.3	0.19	0.38	0.16	0.13	0.23	0.31	0.30	0.30	0.37	0.38	0.20
0.4	0.27	0.39	0.12	0.11	0.25	0.29	0.38	0.34	0.40	0.34	0.17
0.5	0.18	0.22	0.00	0.10	0.32	0.27	0.30	0.33	**0.53**	0.45	0.16
0.6	0.23	0.15	-0.13	0.12	0.22	0.24	0.41	0.26	0.47	0.39	0.17
0.7	0.29	0.14	-0.28	0.11	0.18	0.29	0.48	0.26	0.47	0.41	0.14
0.8	0.24	0.24	-0.35	0.14	0.24	0.22	0.35	0.24	0.38	0.45	0.11
0.9	0.27	0.35	-0.19	0.12	0.27	0.46	0.45	0.32	0.31	0.35	0.21
1	0.29	0.38	-0.36	0.097	0.21	0.36	0.31	0.27	0.30	0.38	0.11

Table 2: Sharpe Ratios