Adaptive Bayesian Optimisation for Online Portfolio Selection

Favour M. Nyikosa, Michael A. Osborne and Stephen J. Roberts
Department of Engineering Science
University of Oxford
{favour,mosb,sjrob}@robots.ox.ac.uk

Abstract

We present a Bayesian approach for online portfolio selection, a fundamental problem in computational finance. We pose the problem as the global optimisation of an expensive, time-varying, black-box function. As the optimum is itself dynamic, we use a model that allows us to capture time-dependent patterns of the function and to provide sequential decision processes that enable us to select optimal portfolios to invest in an online manner.

1 Introduction

Portfolio selection is the allocation of wealth across a set of assets with the goal of achieving some objectives. In this work, we are interested in the real-world portfolio selection task which involves the sequential allocation of wealth over multiple investment periods.

Our problem is choosing optimal allocations of wealth in a dynamic financial market. This is characterised by the feature that the relevant optimal portfolios evolve over time because the underlying market that affects our model of the problem is changing in time. This environment requires an agent to repeatedly select optimal portfolios in order to adapt to the market. This leads to the manifestation of a sequential decision problem where each decision in that chain is an optimal decision acclimatising to the environment's state at that point.

We tackle this problem by treating the function modelling this dynamic environment as a black-box. We fold the time information from our observations into this black-box model, which we will henceforth call the *objective function*. Gaussian process (GP) inference machinery is employed to give us the flexibility of folding in much prior domain knowledge into the objective function's model. Using Bayesian optimisation (BO) [1–5] we automatically balance the exploration and exploitation tradeoff to carefully select the varying maximisers over time. This model allows us to capture time-dependent patterns of the function. We take a non-Markov approach and assume that the changes in the function are caused by a temporally increasing data stream associated with the objective function. BO is used to find where the optimum will be a short time into the future. We are unaware of sequential BO or active-learning criteria that are specifically designed for the type of problem considered in this paper.

The rest of the paper is structured as follows. Sections 2 and 3 describe the setting of the problem and the related work. Section 4 delineates the prescribed framework to address it. Section 5 gives the empirical results and analysis of our suggested approach in online portfolio selection. Section 6 gives conclusions.

2 Problem setting

Our formalisation of the problem is based on the survey on online portfolio selection in [6] and is outlined in algorithm 1. We invest our wealth over m assets in the market for a sequence of n trading periods. The price changes (*returns*) are represented by a return vector $\boldsymbol{r}_t \in \mathbb{R}_+^m$, $t = 1, ..., n$, where the ith element of tth return vector, $r_{t,i}$, denotes the ratio of tth closing price to last closing price for the ith asset. Thus, an investment in asset i in period t increases by a factor of $r_{t,i}$. We are interested in a fixed time horizon so the market window $\boldsymbol{R}$ that starts from period 1 to n is $\boldsymbol{R}_1^n = \{\boldsymbol{r}_1, ..., \boldsymbol{r}_n\}$.

Algorithm 1 Online Portfolio Selection Framework

Input: $\boldsymbol{R}_1^n$ historical asset return data
Output: S_n, final cumulative wealth
Procedure:
Initialisation: $\boldsymbol{x} = \frac{1}{m}, S_0 = 1$
for $t = 1, 2, ...n$ **do**
 Agent learns portfolio $\boldsymbol{x}_t$
 Market reveals returns $\boldsymbol{r}_t$
 Portfolio incurs period return $s_t = \boldsymbol{x}^\top \boldsymbol{r}_t$
 Update cumulative return $S_t = S_{t-1} \times (\boldsymbol{x}_t^\top \boldsymbol{r}_t)$
 Agent updates portfolio selection model
end for

At the beginning of the tth period, an investment is specified by a portfolio vector $\boldsymbol{x}_t$, our decision variable. We assume a portfolio is self-financed and no short-selling is allowed. Therefore, every entry of $\boldsymbol{x}$ is non-negative and all sum up to one.

We use the dynamic function prior to model the objective function $f\colon (\boldsymbol{x}, t) \mapsto \mathbb{R}$ that maps $\boldsymbol{x}$ at time t to some gained wealth. Then BO is used with some heuristics to find the optimal $\boldsymbol{x}_t$. The portfolio $\boldsymbol{x}_t$ is scored using the portfolio period return $\boldsymbol{x}_t^\top \boldsymbol{r}_t$. This procedure is repeated until period n and the strategy is finally scored according to the portfolio cumulative wealth S_n.

We also make the following domain-specific assumptions: (i) Transaction cost: no transaction costs or taxes in the model; (ii) Market liquidity: one can buy and sell any quantity of any asset in its closing prices; (iii) Impact cost: market behaviour is not affected by any portfolio selection strategy.

3 Related work

Table 1 shows the general classification of the state-of-the-art online portfolio selection algorithms and their corresponding representative references. For this work, we use Bayesian optimisation (see [1–3, 5, 19]) with the multi-points EI (qEI) (see [20, 21]).

4 Adaptive Bayesian optimisation (ABO)

4.1 Dynamic function prior model for online portfolio selection

Let us consider an unknown function $y = f(\boldsymbol{x}, t; \mathcal{D}_t)$ that we wish to optimise for a decision process where $y, t \in \mathbb{R}$ and $\boldsymbol{x} \in \boldsymbol{X} \subset \mathbb{R}^N$. The value of the function f associated with the decision variable $\boldsymbol{x}$ at time t is also dependent on a data stream $\mathcal{D}_t = \{\boldsymbol{d}_1, ..., \boldsymbol{d}_t\}$. In portfolio selection, this corresponds to the asset returns per investment period. For every time step, $\mathcal{D}_t$ is updated $\mathcal{D}_t = \mathcal{D}_{t-1} \cup \boldsymbol{d}_t$. However, we often do the optimisation of f for a particular future time $t + \nu$. If the present time is t, then we have only observed $\mathcal{D}_t$. Thus, to optimise f, we need to make estimations of $\mathcal{D}_{t+\nu}$. Since every evaluation of f is dependent on $\mathcal{D}$, the temporal evolution of the data stream affects the function by having its values depend on time. Therefore, the function f is changing in time.

The function f is expensive to evaluate where, at time t, we would only observe y as a response to a portfolio $\boldsymbol{x}$. Consequently, our sequence of decisions would usually be incrementally collated into a single data set, $\big[\{(\boldsymbol{x}_1, 1), y_1\}, ..., \{(\boldsymbol{x}_t, t), y_t\}\big]$.

For every time t of interest, our goal is to find an $\boldsymbol{x}_{\text{best}}$ that gives us the best response y_{best} after maximising $f(\boldsymbol{x}, t; \mathcal{D}_t)$. In addition to function input and output observations, we also collect the times t that the observations are made, and the data stream $\mathcal{D}$ that affects the response. Therefore, if we place a GP prior on f and use an appropriate covariance function, we can effectively capture the how the latent function varies in time as we move away from the observed data. We can achieve this by including an extra input dimension to the data to infuse the time-related information. This adds an extra term to the covariance that folds in the time evolution aspect of the problem. For this paper, we use the separable covariance function

$$k(\{\boldsymbol{x}_i, t_i\}, \{\boldsymbol{x}_j, t_j\}) = k_{\text{Gabor}}(t_i, t_j) \times k_{\text{Linear}}(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{1}$$

where

$$k_{\text{Gabor}}(t_i, t_j) = \exp\Big(-\frac{(t_i - t_j)^2}{2l^2}\Big)\cos\Big(\frac{2\pi(t_i - t_j)}{p}\Big) \text{ and } k_{\text{Linear}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^\top \boldsymbol{P}^{-1}\boldsymbol{x}_j$$

where hyperparameter l is the lengthscale in time, p is the period and $\boldsymbol{P}$ is an $N \times N$ diagonal matrix with lengthscale hyperparameters of the portfolio dimensions along its diagonal. The separable kernel we consider is of the form $k_{SpectralMixture} \cdot k_{Linear}$. We used the linear kernel over the portfolios because of the linear way that the investment rewards are realised when no transaction costs are assumed. The spectral mixture kernel is used in order to capture time-dependent patterns and was inspired by work in [22] where a kernel for pattern discovery is proposed. In the context of online portfolio selection, these patterns may include mean reversion, anti-correlation and periodicity, among others. The kernel proposed in [22] is dense in the space of all stationary kernels [23]. It is a weighted mixture of Gabor kernels and in this paper we consider the simplest case.

In summary, we place a GP prior over the portfolio-time space in order to obtain a posterior over the function's time scale to capture the time evolution.

Table 1: General classification for the state-of-the-art online portfolio selection algorithms.

Classification	**Algorithm**	**Reference**
Benchmarks	Equally weighted	
	Best-stock	
	Constant Rebalanced Portfolios (CRP)	[6–8]
Follow Best	Universal Portfolios (UP)	[7]
	Exponential Gradient (EG)	[9]
	Follow Regularised Leader (ONS)	[10]
Follow Worst	Anti Correlation (AntiCor)	[11, 12]
	Passive Aggressive Mean Reversion (PAMR)	[13]
	Confidence Weighted Mean Reversion (CWMR)	[14]
	Online Moving Average Reversion (OLMAR)	[15]
Pattern Matching	Nonparametric Kernel (B^K)	[16]
	Nonparametric Nearest-Neighbour (B^{NN})	[17]
	Correlation-driven Nonparametric Learning (CORN)	[18]

4.2 The ABO framework

There are several aspects of this problem that require illumination.

Firstly, we are solving multiple related problems rather than one global problem.

Secondly, each of these sub-problems is a global optimisation problem that requires us to find the best portfolio constrained at particular times where we assume there is a smoothly varying objective function at each of these times. We are also constrained to make evaluations at particular time steps during the optimisation. We cannot choose evaluations of our function that occur at times before the evaluations we have already gathered. Unlike standard BO where we have several steps to return the proposed minimum for a fixed objective function, we only have one step to return this best point for any particular time step for the time-varying objective function.

Since we only have one evaluation to return an optimal portfolio for the next investment period, we resort to using epistemic explorative evaluations of our function based on our GP. To aid with this, we use the qEI heuristic [20, 21] to generate a batch of q proposals for that time step. Because the actual evaluation and optimisation of qEI is expensive, we use an approximation similar to the Kriging believer (KB) in [20] that produces these q points sequentially. Our approach differs with KB in this way: after generating the q points from the EI optimisations, we rank these proposals based on the GP posterior means they correspond to rather than the EI values they give. We do this because some points with high EI values may correspond to exploratory steps that may be riskier and we wish avoid them. This is an application-specific heuristic meant to avoid high drawdowns, which is pivotal in financial investment applications. This stems from the need to reduce the risk of losing money when making investment decisions. Therefore, from the q proposals, we chose a single point that gives us the best return based on our epistemic beliefs of our objective function and discard the others.

Adding time to the covariance function introduces uncertainty about the past data we have observed. To deal with this, we have to recalculate the GP prior hyperparameters at every time step to update our priors with information of how the function has changed in time.

In summary, at every step t, we perform BO on our objective function constrained at a future time $t+\nu$ that we would like to find the best portfolio for. We gather our batch of q multi-points based on qEI, return the best portfolio and discard the others. Algorithm 2 shows an outline of the approach.

5 Experiments

In this section we present experiments to empirically analyse adaptive Bayesian optimisation for online portfolio selection.

Algorithm 2 Adaptive Bayesian optimisation (ABO)

Input: $\mathcal{D}_0$, prior $f \sim \mathcal{GP}(\mu, k) \to \mathcal{I}_0$
Input: Budget of BO steps n, iteration label i
Input: Number of batch-points q
Input: Time label t, Time window ν
Output: Optima $\{\boldsymbol{x}_{t_0+\nu}, \boldsymbol{x}_{t_0+2\nu}, ..., \boldsymbol{x}_{t_0+N\nu}\}$
Output: $\{\boldsymbol{y}_{t_0+\nu}, \boldsymbol{y}_{t_0+2\nu}, ..., \boldsymbol{y}_{t_0+N\nu}\}$
for $\boldsymbol{i} = 1, 2, ...n$ **do**
  $I_{\text{temp}} = I_{i-1}$
  $(\boldsymbol{x}_1^*, ..., \boldsymbol{x}_q^*) = \underset{x|t=t_0+i\nu}{\text{arg max}}\ \text{qEI}(\boldsymbol{x}_1, ..., \boldsymbol{x}_q)$
  Set $\boldsymbol{x}_{t_0+i\nu} = \boldsymbol{x}_j^*$ where $\max\{\mu_{posterior}(\boldsymbol{x}_j^*), j = 1, ..., q\}$
  $y_{t_0+i\nu} = f(\boldsymbol{x}_{t_0+i\nu}, i)$, take action & observe true response
  $\mathcal{D}_i = \mathcal{D}_{i-1} \cup \boldsymbol{d}_i$, observe update in datastream
  $I_i = I_{i-1} \cup \{\boldsymbol{x}_{t_0+i\nu}, y_{t_0+i\nu}\}$
  Update GP model
end for

The data used in the experiments is shown in table 2. For more details on the datasets see [13, 24]. The implementation was in MATLAB with qEI optimisation done with the DIRECT search solver for derivative-free optimisation in where the portfolio and optimisation constraints were set. The GP prior had a zero mean function and the separable covariance function in equation 1 with MAP estimation of hyperparameters. We used a maximum moving data window of size 5 times the input dimensions for all experiments in order to deal with computational complexity of the GP when the number of data points increased.

Table 2: Summary of 4 real datasets in used in experiments.

Dataset	# Assets	# Days	Period
SP500	25	1276 (5 yrs)	Jan/2/1998 – Jan/31/2003
DJIA	30	507 (2 yrs)	Jan/14/2001 – Jan/14/2013
TSE	88	1258 (5 yrs)	Jan/4/1994 – Dec/31/1998
MSCI	24	1043 (4 yrs)	Apr/1/2006 – Mar/31/2010

The averaged results performed on 5 runs of the experiment on each dataset are shown in table 3. The time window (ν) used was one day. The results for competitor methods are based on configurations described in [25]. The results show that the ABO strategy with qEI achieves better performance than the competitors and significantly outperforms the state-of-the-art, Online Moving Average Reversion [15, 25]. BO benefits from the specification of priors and subsequent adaptation to new trends in the observed signals. An examination of the portfolios generated by BO showed that it invested in a small subset of the assets in the portfolio at every timestep, mostly following trends most consistent with the posterior mean and avoiding riskier positions with higher uncertainty. This was due to the selection criteria of the winning portfolio from the q generated proposals and also from the constraints imposed when optimising qEI.

6 Conclusions

We presented an adaptive Bayesian optimisation approach for online portfolio selection by modelling the objective function at every step as an expensive, time-varying, black-box function. We described an appropriate dynamic function prior model suited to this problem. Furthermore, we

Table 3: Cumulative wealth achieved by various trading strategies on 4 datasets.

Methods	SP500	DJIA	TSE	MSCI
Equally weighted	1.34	0.76	1.61	0.91
Best-stock	3.78	1.19	6.28	1.50
BCRP	4.04	1.24	6.78	1.51
UP	1.64	0.81	1.59	0.92
EG	1.63	0.81	1.60	0.92
ONS	3.34	1.53	1.62	0.86
B^K	2.24	0.68	1.62	2.64
B^{NN}	3.07	0.88	2.27	13.47
CORN	6.35	0.84	3.56	26.10
AntiCor	5.89	2.29	39.36	3.22
PAMR	5.09	0.68	264.86	15.23
CWMR	5.90	0.68	332.62	17.28
OLMAR-S	8.63	2.12	424.80	16.39
OLMAR-E	8.63	1.20	678.44	21.29
ABO	13.01	107.39	1682	342.83

empirically demonstrated that the resulting algorithm skilfully navigates the portfolio-time space seeking out the best portfolios to hold.

References

[1] Eric Brochu, Vlad M. Cora, and Nando de Freitas. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *CoRR*, 2010.

[2] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical Bayesian optimization of machine learning algorithms. In *NIPS*, 2012.

[3] Jones Donald R. The taxonomy of global optimisation methods based on response surfaces. *The Journal of Global Optimization*, 21:345–383, 2001.

[4] Lizotte Daniel James. *Practical Bayesian Optimization*. PhD thesis, University of Alberta, Edmonton, Alberta, Fall 2008.

[5] Michael A. Osborne, Roman Garnett, and Stephen J. Roberts. Gaussian processes for global optimization. *International Conference on Learning and Intelligent Optimization (lion 2009)*, 2009.

[6] Bin Li and Steven C. H. Hoi. Online portfolio selection: A survey. *ACM Comput. Surv.*, 46(3):35:1–35:36, January 2014.

[7] Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

[8] J. J. Kelly. A new interpretation of information rate. *Bell Systems Technical Journal*, 35(917-926), 1956.

[9] David P Helmbold, Robert E Schapire, Yoram Singer, and Manfred K Warmuth. On-line portfolio selection using multiplicative updates. *Mathematical Finance*, 8(4):325–347, 1998.

[10] Amit Agarwal, Elad Hazan, Satyen Kale, and Robert E. Schapire. Algorithms for portfolio management based on the newton method. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, pages 9–16, New York, NY, USA, 2006. ACM.

[11] A Borodin, R El-Yaniv, and V Gogan. Can we learn to beat the best stock. In *Neural Information Processing Systems*, 2003.

[12] A Borodin, R El-Yaniv, and V Gogan. Can we learn to beat the best stock. *Journal of Artificial Intelligence Research*, 21(579-594), 2004.

[13] Bin Li, Peilin Zhao, Steven Hoi, and Vivekanand Gopalkrishnan. PAMR: Passive aggressive mean reversion strategy for portfolio selection. *Machine Learning*, 87(2):221–258, 2012.

[14] Bin Li, Steven C.H. Hoi, Peilin Zhao, and Viveknand Gopalkrishnan. Confidence weighted mean reversion strategy for on-line portfolio selection. In *ACM Transactions on Knowledge Discovery from Data*, 2013.

[15] Bin Li and Steven C. H. Hoi. On-line portfolio selection with moving average reversion. In *Proceedings of the International Conference on Machine Learning*, 2012.

[16] L Gyorfi, G Lugosi, and F Udina. Nonparametric kernel-based sequential investment strategies. *Mathematical Finance*, 16(2):337–357, 2006.

[17] L Gyorfi, F Udina, and H Walk. Nonparametric nearest neighbor based empirical portfolio selection strategies. *Statistics and Decisions*, 26(2):145–157, 2008.

[18] Bin Li, Steven C.H. Hoi, and Viveknand Gopalkrishnan. Corn: Correlation-driven nonparametric learning approach for portfolio selection. *ACM Transactions on Intelligent Systems and Technology*, 2(3):21:1–21:29, 2011.

[19] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the human out of the loop: A review of bayesian optimization. Technical report, Universities of Harvard, Oxford, Toronto, and Google DeepMind, 2015.

[20] David Ginsbourger, Rodolphe Le Riche, and Laurent Carraro. Kriging is well-suited to parallelize optimization. In *Computational Intelligence in Expensive Optimization Problems*, pages 131–162. Springer, 2010.

[21] Clément Chevalier and David Ginsbourger. Fast computation of the multi-points expected improvement with applications in batch selection. In *Learning and Intelligent Optimization*, pages 59–69. Springer, 2013.

[22] Andrew Gordon Wilson and Ryan Prescott Adams. Gaussian process kernels for pattern discovery and extrapolation. *arXiv preprint arXiv:1302.4245*, 2013.

[23] Yves-Laurent Kom Samo and Stephen Roberts. Generalized spectral kernels. *arXiv preprint arXiv:1506.02236*, 2015.

[24] DBOLPS: A database for on-line portfolio selection. `http://www.mysmu.edu.sg/faculty/chhoi/olps/datasets.html`. Accessed: 2015-02-15.

[25] Bin Li, Steven C.H. Hoi, Doyen Sahoo, and Zhi-Yong Liu. Moving average reversion strategy for on-line portfolio selection. *Artificial Intelligence*, 222:104 – 123, 2015.