Boosting Moving Average Reversion Strategy for Online Portfolio Selection: A Meta-Learning Approach

Lin Xiao[1], Zhang Min[2], Zhang Yongfeng[3], Gu Zhaoquan[4], Liu Yiqun[2], and Ma Shaoping[2]

[1] Institute of Interdisciplinary Information Sciences, Tsinghua University, Beijing, China,
jackielinxiao@gmail.com;
[2] Tsinghua National Laboratory for Information Science and Technology, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China,
z-m@tsinghua.edu.cn; yiqunliu@tsinghua.edu.cn; msp@tsinghua.edu.cn
[3] College of Information and Computer Science, University of Massachusetts Amherst, MA 01003, USA, yongfeng@cs.umass.edu
[4] Department of Computer Science, HongKong University, HongKong, China,
demin456@gmail.com

Abstract. In this paper, we study the online portfolio selection problem from the perspective of meta learning for mean reversion. The online portfolio selection problem aims to maximize the final accumulated wealth by rebalancing the portfolio at each time period based on the portfolio prices announced before. Mean Reversion is a typical principle in portfolio theory and strategies that utilize this principle achieve the superior empirical performances so far. However there are some important limits of existing Mean Reversion strategies: First, the mean reversion strategies have to set a fixed window size, where the optimal window size can only be chosen in hindsight. Second, most existing mean reversion techniques ignore the temporal heterogeneity of historical price relatives from different periods. Moreover, most mean reversion methods suffer from noises and outliers in the data, which greatly affects the performances. In order to tackle the limits of previous approaches, we exploit mean reversion principle from a meta learning perspective and propose a boosting method for price relative prediction. More specifically, we generate several experts where each expert follows a specific mean reversion policy and predict the final price relatives with meta learning techniques. The sampling of multiple experts involves mean reversion strategies with various window sizes; while the meta learning technique brings temporal heterogeneity and stronger robustness for prediction. We adopt online passive-aggressive learning for portfolio optimization with the predicted price relatives. Extensive experiments have been conducted on real-world datasets and our approach outperforms the state-of-the-art approaches significantly.

1 Introduction

Online portfolio selection problem aims to allocate the wealth among different assets at different time periods to maximize the long-term wealth. There are two models describing the problem: the Mean-Variance model[19] and Kelly's Capital Growth model [11]. The first model uses a weighted sum of expected return (mean) and risk (variance of the return) as a trade-off between the two objectives, and it is suitable for single-period portfolio selection; the second model sees the problem as a sequential decision problem

that aims to maximize the expected return at the end of multiple time periods. Kelly's Capital Growth model has a nature of online decision making, which is widely adopted by the studies from AI and Machine Learning researchers.

In online portfolio selection problem, each asset is associated with a price in each period. The ratio of prices between current period and last period is called price relative, which reflects the return of wealth invested on the assets after one period. The agent allocates the wealth among different assets based on their price relatives at different periods. Most portfolio selection strategies follow a two-phase scheme: price relative prediction phase and portfolio optimization phase. The first phase aims to predict the price relative at next period based on historical data; while the second phase aims to compute the optimal portfolio given the prediction of price relative.

One common methodology for this problem is Mean Reversion, which assumes the portfolios that perform poorly at current period will perform well next (and vice versa). The methods include PAMR (Passive Aggressive Mean Reversion)[17], CWMR (Confidence Weighting Mean Reversion) [15], OLMAR (OnLine Moving Average Reversion)[14] and RMR (Robust Mean Reversion)[9], which adopt the mean reversion idea in different ways. They achieve superior empirical results in experiments compared with other state-of-art methods. This proves the effectiveness of mean reversion policy.

Although PAMR, CWMR and OLMAR achieve good performances, they still face some difficulties. All existing mean reversion strategies do not fully consider the noisy data and outliers (RMR is proposed to alleviate the problem), which often leads to estimation error (see [20]). Furthermore, the assumption of single-period prediction [17][15] also leads to estimation error, which makes the performance poor. RMR (Robust Mean Reversion)[9] and OLMAR[14] uses multi-period prediction, but the algorithm sees each period equally, which ignores the temporal heterogeneity of historical price relatives and causes inaccuracy of predictions. We utilize meta learning to exploit the benefit of multi-period prediction and the periods are assigned with weights according to their performances. Moreover, this alleviates the impact of noisy data and outliers. The results show that our strategy outperforms RMR and OLMAR.

More specifically, in order to utilize multi-period historical data, we generate multiple experts for price relative prediction following typical MAR methods. Then we adopt the meta learning method for price prediction. Each expert is assigned a weight that is updated according to their performances and a weighted aggregation is used as the final prediction. Meanwhile, we choose the typical passive-aggressive learning method for portfolio optimization. This method captures the recent portfolio performances and the objective of enhancing the wealth return at each time period.

The contributions of our work are: first, to our knowledge, we are the first to exploit the Mean Reversion strategy with meta learning in online portfolio selection problems; second, we make a better use of multiple-period history, which is robust to the outliers and noises in historical data; third, we conduct extensive experiments on real-world datasets and achieve superior results compared with other state-of-the-art approaches.

The remainder of the paper is organized as follows: the next section gives a brief introduction to the related work; Section 3 formally introduces the online portfolio selection problem while Section 4 introduces some preliminary works about mean reversion theory and some related concepts. Section 5 proposes BMAR strategy that utilizes meta

learning in online portfolio selection problem. Section 6 presents the results of experiments conducted on real-world datasets and a thorough comparison with the baselines. The conclusion and the future work are presented in Section 7.

2   Related Work

The study of online portfolio selection problem first concentrates on some benchmark algorithms, including Buy and Hold, Best Stock and Constant Rebalanced Portfolios. The Buy and Hold strategy means the agent invests wealth with an initial portfolio and holds it to the end without changing the portfolio. The Best Stock strategy means that one puts all the wealth on the stock whose performance is best in hindsight. Constant Rebalanced Portfolios is a strategy that rebalances the wealth to a fixed portfolio in all periods. The best CRP strategy which achieves highest accumulated wealth is called BCRP. BCRP is an optimal strategy if the market is i.i.d. [4]. Successive Constantly Rebalanced Portfolios (SCRP) [5] and Online Newton Step (ONS) [1] implicitly estimate next price relative via all historical price relatives with a uniform probability. However, both Best Stock and BCRP strategies have to be computed in hindsight.

There are two main categories of algorithms: follow-the-winner approach and follow-the-loser approach. The intuition behind first approach is to track the stock with best performance in history and raise the weights in the portfolio of these stocks. Most of the Follow-the-Winner approaches aim to imitate the BCRP strategy: including the universal portfolio selection (UP)[10], Exponential Gradient (EG)[8], follow-the-leader and follow-the-regularized-leader approaches. However, the prices of assets are unstable, even a good following of winner assets can not guarantee superior performances.

Follow-the-loser approach utilizes a typical assumption of mean reversion [18], which means that the good (poor)-performing assets will perform poor (good) in the following periods. The approaches in this category include Anti-Correlation (Anticor) [2], Passive-Aggressive Mean Reversion (PAMR)[17], Confidence-weighted Mean Reversion (CWMR)[15], OLine Moving Average Reversion (OLMAR)[14] and Robust Mean Reversion (RMR)[9]. CRP[4][5] implicitly envolves follow-the-loser approach since rebalancing the wealth means to transfer the wealth from winning stocks to losing stocks in some extent.

Another important category of the portfolio selection algorithms is pattern-matching, which estimates the portfolio price based on sampled similar historical patterns. Nonparametric kernel based moving window (BK) [7] measures the similarity by kernel method. Following the same framework, Nonparametric Nearest Neighbor (BNN) [12] locates the set of price relatives via nearest neighbor methods. [16] proposed Correlation-driven Nonparametric learning (CORN), which measures the similarity via correlation.

Since the mean-reversion technique is widely adopted in financial fields, it is useful in online learning algorithms as well. Passive Aggressive Mean Reversion (PAMR) [17] and Confidence Weighted Mean Reversion (CWMR) [15] estimate next price relative as the inverse of last price relative, which is in essence the mean reversion principle. Recently, [14] proposed On-Line Moving Average Reversion (OLMAR), which predicts the next price relative using moving averages and explores the multi-period mean reversion. Robust Mean Reversion is proposed to alleviate the impact of outliers and

noises existing in the data. The empirical experiments indicate that OLMAR and RMR outperforms the other state-of-the-art algorithms. However, most of the Mean Reversion algorithms have important limits: first, the mean reversion strategies require to select a fixed time window for prediction, which can not be easily determined; second, the strategies treat each time period equally in prediction, which ignores the temporal heterogeneity; third, the strategies do not have strong robustness against noises and outliers. We utilize mean reversion strategies for its good depiction of reality and further use meta learning approach to tackle the limits. A detailed comparison between our strategy and existing Mean Reversion strategies is presented in Section. 4.

3 Problem Setting

In this section, we formally introduce the Online Portfolio Selection problem. Assume that there exist m assets in market, and time is divided into T periods. Each asset i has a closing price $p_{t,i}$ at period t and p_t is denoted as the closing price vector (column vector for all vectors mentioned): $p_t = [p_{t,1}, p_{t,2}, ..., p_{t,m}]$. $x_{t,i}$ is the price relative that captures the ratio of closing prices between two consecutive periods: $x_{t,i} = \frac{p_{t,i}}{p_{t-1,i}}$ and $x_t = [x_{t,1}, x_{t,2}, ..., x_{t,m}]$ is the price relative vector.

In each period, the market reveals the closing prices of assets and the investor has to assign the capital with a portfolio vector: $b_t = [b_{t,1}, b_{t,2}, ..., b_{t,m}]$ where $b_{t,i}$ represents the proportion of wealth assigned to asset i at time t. We follow the typical assumption that no margin/short sale is allowed, therefore $b_{t,i} \geq 0, \forall t, i$ and $\sum_{i=1}^{m} b_{t,i} = 1, \forall t$. An investment means to select a portfolio b_t from the simplex: $\{b_{t,i} \geq 0, \sum_{i=1}^{m} b_{t,i} = 1\}$ at period t. Usually we assume the portfolio is uniformly distributed in the beginning: $b_0 = [\frac{1}{m}, ..., \frac{1}{m}]$. The sequence of portfolios from t_1 to t_2 is denoted as $b_{t_1}^{t_2}$. Denote the wealth accumulated at t as W_t, w.l.g. the initial wealth is assumed to be 1: $W_0 = 1$. Therefore given the selected portfolio b_t at period t, the wealth becomes $W_t = W_{t-1} b_t^T x_t = \sum_{i=1}^{m} b_{t,i} x_{t,i} W_{t-1} = \prod_{\tau=1}^{t} b_\tau^T x_\tau$ (T is transpose here).

Given the notations and introduction above, the online portfolio selection problem refers to a sequential decision making problem with periods from $t = 1$ to $t = n$. In each period, the investor has to decide the portfolio based on historical closing prices of assets and the market reveals the newest closing price of assets, which leads to the change of wealth. The investor needs to strategically design portfolios b_1^n so that the accumulated wealth W_n at time $t = n$ is maximized.

We summarize the procedures from the introduction above and formulate the whole online portfolio selection process in Alg. 1 as [13].

4 Preliminary

In this section, we briefly introduce the mean reversion principle and how former works exploit this principle.

4.1 Mean Reversion

In each period, the algorithm tries to estimate the price relatives of assets in price prediction phase and compute the portfolio with Passive-Aggressive or Confidence Weighted

---

Algorithm 1 ONLINE PORTFOLIO SELECTION

Input: x_1^n : Historical market price relative sequence
Output: W_t : Final cumulative wealth
Procedure:
1: Initialize the portfolio $b_1 = \frac{1}{m}\mathbf{1}$, the wealth is initialized as $W_0 = 1$;
2: **for** $t = 1, 2, ..., n$ **do**
3:    Portfolio mannager learns the portfolio $\mathbf{b_t}$;
4:    Market reveals the market price relative $\mathbf{x_t}$;
5:    Portfolio incurs period return $w_t = \mathbf{b_t^T x_t}$ and updates cumulative return $W_t = W_{t-1} \times (\mathbf{b_t^T x_t})$;
6:    Portfolio manager updates the online portfolio selection rules;
7: **end for**;

Learning given the predicted price relatives. The mean reversion principle is reflected in the first phase by assuming that the poor-performing assets will have good performances in next periods (and vice versa). Denote the estimated price relative as $\tilde{x}_t$ and the estimated closing price as $\tilde{p}_t$. PAMR and CWMR assumes that the assets with high/low price relatives will have low/high price relatives in next period: $\tilde{x}_{t+1} = \frac{1}{x_t}, \forall t$, which means $\frac{\tilde{p}_{t+1}}{p_t} = \frac{p_{t-1}}{p_t}$. Therefore the principle assumes that $\tilde{p}_{t+1} = p_{t-1}$. Although the two methods work well, they can not perform consistently on some datasets. There are two reasons that lead to this inefficiency: first, the fluctuating prices may contain noises that affect the precision of mean reversion principle; second, the single period price reversion effect may not exist widely as expected.

4.2 Online Moving Average Reversion

The OLMAR (online moving average reversion) principle is proposed to model the mean reversion principle with multiple-period historical data. Denote the time window of OLMAR as w, the closing price p_t at period t is assumed to be: $\tilde{p}_t = \frac{1}{w}\sum_{\tau=t-w}^{t-1} p_\tau$. The price is therefore considered as the average of prices in a time window and the price relative becomes: $\tilde{x}_t = \frac{p_t}{p_{t-1}} = \frac{1}{w}\left(1 + \frac{1}{x_{t-1}} + ... + \frac{1}{\bigotimes_{\tau=t-w}^{t-1} x_\tau}\right)$. This Moving Average Reversion strategy with time window is denoted as SMAR.

Usually, it is assumed that the price at current period is closer to the price at recent periods due to the continuity of price changes. Therefore the price can be estimated with MAR by adding a decay factor α: $\tilde{p}_t = \alpha p_{t-1} + (1 - \alpha\tilde{p}_{t-1})$, which results with a price relative: $\tilde{x}_t = \alpha + (1 - \alpha)\frac{\tilde{x}_{t-1}}{x_{t-1}}$. Therefore the decay factor frees the algorithm from choosing a time window and utilizes the prices from the whole history. This Moving Average Reversion strategy with time window is denoted as EMAR.

Given the price relative predictions, OLMAR method further utilizes the online passive aggressive learning policy for portfolio optimization, which is also adopted by PAMR. Notice that the choice of time window size and decay factor determines the performance of this method and can not be pre-defined in hindsight.

Table 1. Comparison between our strategy (BMAR) with Existing Mean Reversion Approaches.

Approaches	Mean Reversion	Multi-Period	Robustness	Temporal Heterogenity
PAMR	$\checkmark$	$\backslash$	$\backslash$	$\backslash$
CWMR	$\checkmark$	$\backslash$	$\backslash$	$\backslash$
OLMAR	$\checkmark$	$\checkmark$	$\backslash$	$\backslash$
RMR	$\checkmark$	$\checkmark$	$\checkmark$	$\backslash$
BMAR	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$

4.3 Robust Mean Reversion

RMR uses L1 estimator to estimate the closing prices of assets so that the resulting price has a better robustness compared to other methods: with a window size of historical periods w, RMR estimates the closing at t by minimizing this objective: $\sum_{\tau=t-1}^{t-w} \|\tilde{p}_t - p_\tau\|_2$, and the price relative is estimated as $\frac{\tilde{p}_t}{p_{t-1}}$. This estimation is named as L1 estimator and is relatively more robust to noises.

4.4 Temporal Heterogeneity

The price relatives of different periods are correlated in different extents. Usually, it is assumed that the price at current period is closer to the price at recent periods due to the continuity of price changes. Meanwhile, the changes of price relatives in some time windows are more similar than others (which is the foundation of CORN (CORrelation-driven Nonparametric learning) [16]). Therefore, when predicting the price relatives in the future, the algorithm should consider how to make use of the historical time windows differently.

Given these existing algorithms exploiting Mean Reversion principles, we make a comparison between our approach (Boosting Moving Average Reversion: BMAR) and these algorithms in Table. 1. The multi-period column shows whether the approaches use multiple-period historical data for prediction; the Robustness shows whether the approaches are robust to noises and outliers; the Temporal Heterogenity column shows whether the approaches can utilize the data from different periods Heterogeneously. The table shows that our strategy (BMAR) preserves all the good properties, which shows the superiority.

5 Proposed Strategy: Boosting Moving Average Reversion

Like most methods, we solve the problem with two phases: price relative prediction and portfolio optimization. We first generate a set of experts and each expert is a predictor of the price relative following a mean reversion policy. In each period, each expert first makes its predictions on the price relatives in next period; then we compute the cumulated losses induced by each expert from their historical predictions and the true past price relatives. With the cumulated losses, we compute the weights assigned to each expert following the boosting methods introduced later and make a final prediction. Then we use Online Passive-Aggressive learning method to compute an optimized portfolio
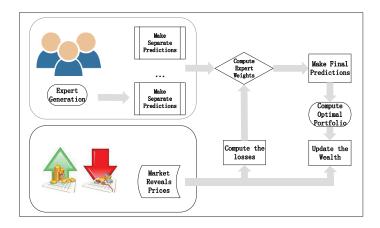
Fig. 1. BMAR Strategy for Online Portfolio Selection

with the final prediction of price relatives in next period. When the true price relatives are revealed, we can update the cumulated losses of each predictor. The process of BMAR strategy for online portfolio selection is illustrated in Fig. 1.

5.1 Boosting Moving Average Reversion for Price Relative Prediction

We generate the experts of price relative prediction with different parameters from OL-MAR. Denote the set of experts as E and the SMAR expert with time window w is denoted as E_w, the EMAR expert with Decaying factor α as E_α.

Uniform Sampling: We generate the experts by sampling the parameters uniformly from the range: $w \sim U(w_{min}, w_{max})$ and $\alpha \sim U(0, 1)$. We generate $M = w_{max} - w_{min} + 1$ experts from MAR ($w = w_{min}, w_{min} + 1, ..., w_{max}$) and N experts from EMAR ($\alpha = \{0.1, 0.2, ..., 0.9\}$ when $N = 9$). Based on the different ways of generating experts, we denote the strategy of generating experts with time window as BMAR-1 and the strategy of generating experts by sampling α as BMAR-2. With the generated experts, we use weighted aggregation of their decisions to predict the price relatives at different periods. Each expert represents an approximation of the price relative in next period with a certain time window. By utilizing the predictions of these experts with a weighted scheme, we can induce temporal heterogeneity into our approach and the details will be introduced later.

As shown in Theorem 1 later, the regret of our strategy is closely related to the number of experts we generate. We will present the influence of expert numbers and sampling methods on the performances in the experiment.

We assume a weighted sum of the experts as the estimator of the price relative:

$$BMAR-1 : \tilde{x}(t) = \frac{\sum_{i=1}^{i=N} \theta_{i,t-1} \tilde{x}(t, w_i)}{\sum_{i=1}^{i=N} \theta_{i,t-1}}; BMAR-2 : \tilde{x}(t) = \frac{\sum_{j=1}^{j=M} \theta_{j,t-1} \tilde{x}(t, \alpha_j)}{\sum_{j=1}^{j=M} \theta_{j,t-1}} \quad (1)$$

where $\tilde{x}(t, w)$ and $\tilde{x}(t, \alpha)$ are the predicted price relatives of expert E_w and E_α; $\theta_{w,t-1}$ and $\theta_{\alpha,t-1}$ are the weights assigned to these experts given their performances until period $t - 1$.

Denote the loss of expert E_w by period t as $l(w, t)$ and the loss of weighted expert (BMAR) by period t as $l(t)$. The cumulated losses of expert E_w and weighted expert by period T are denoted as $L(w, T) = \sum_{t=1}^{t=T} l(w, t)$ and $L(T) = \sum_{t=1}^{t=T} l(t)$ respectively. The difference between the two losses is seen as the regret of weighted expert with respect to expert E_w: $R(w, T) = L(T) - L(w, T)$.

We introduce the weights assigned to the expert, i.e. exponential weights:

$$\theta_{i,t-1} = \frac{e^{-\eta R_{i,t-1}}}{\sum_{j=1}^{N} e^{-\eta R_{j,t-1}}} \tag{2}$$

where η is a nonnegative parameter. Notice that $R(w, T) = L(T) - L(w, T)$, the exponential weights make the predictions simpler:

$$\theta_{i,t-1} = \frac{e^{-\eta L_{i,t-1}}}{\sum_{j=1}^{N} e^{-\eta L_{j,t-1}}} \tag{3}$$

It has been proved that this expert learning procedure guarantees a proper upper bound of regret in prediction, as shown in the theorems from [3]:

Theorem 1. *Assume that the loss function is convex in its first argument and takes values from $[0, 1]$, then the regret of exponentially weighted average predictor satisfies (N is the number of experts, n is the number of periods and η is the parameter in exponential weights):*

$$\hat{L}_n - \min_{i=1,\dots,N} L_{i,n} \leq \frac{lnN}{\eta} + \frac{n\eta}{2} \tag{4}$$

The details of the proofs can be found in [3] and we omit the details here. Given these theorems, we can design loss functions that satisfy the requirements:

$$l(\hat{p}, y) = \frac{1}{N\varepsilon} \times \|\hat{p} - y\|_2^2 \tag{5}$$

where ε is the constant that rescales $l(\hat{p}, y)$ into $[0, 1]$. And it is easy to verify that the function is convex, which satisfies the requirement of the theorems. Notice that ε actually works as coefficients of $l(\hat{p}, y)$ with η, we can simply tune the value of η to adjust the performances, therefore when using this loss function, we do not explicitly set the value of ε.

Remarks on Robustness: Notice that we do not explicitly model robustness in our prediction, however the utilization of multiple experts involves robustness: if the outliers and noises causes degradation of the experts' prediction accuracy, the weights assigned to these affected experts are lowered, which prevents the final predictions suffering from the noises and outliers.

5.2 Portfolio Optimization

Given the predicted price relatives shown in former section, we utilize the passive aggressive learning procedure to solve an optimal portfolio. The basic idea of passive

aggressive learning is to keep the portfolio the same if the predefined requirement is satisfied, otherwise the portfolio is computed to satisfy the requirement with a minimal change. More specifically, we formulate the optimization problem as follows:

$$
\begin{aligned}
min. \qquad & \|b_t - b_{t-1}\|^2 \\
s.t. \quad & b_t \tilde{x}_t \geq \epsilon, \text{ and } b_t \succeq \mathbf{0}
\end{aligned}
\tag{6}
$$

where ϵ is the threshold for the return at each period. Usually ϵ is a constant greater than 1 to ensure the return under predicted price relative is increasing. The optimum to this problem is the portfolio assigned for period t. Notice that if we keep the portfolio same with that in last period and the return under predicted price relatives still exceeds the required value, we will keep the portfolios unchanged; otherwise, we will try to minimize the change between current portfolio and that in last period as long as the return can exceed the requirement. Since this optimization problem is convex, we can derive the portfolio in a closed form. The solution without considering the nonnegativity constraint is presented in the following proposition:

$$
b_{t+1} = b_t - \alpha_{t+1}(\hat{x}_{t+1} - \bar{x}_{t+1} \cdot \mathbf{1})
\tag{7}
$$

where $\bar{x}_{t+1} = \frac{1}{d}(\mathbf{1}\hat{x}_{t+1})$ denotes the average predicted price relative and α_{t+1} is the Lagrangian multiplier calculated as,

$$
\alpha_{t+1} = \min\{0, \frac{\hat{x}_{t+1}b_t - \epsilon}{\|\hat{x}_{t+1} - \bar{x}_{t+1} \cdot \mathbf{1}\|^2}\}
\tag{8}
$$

In order to ensure that the portfolio is non-negative, we project the above portfolio into the simplex domain as [14].

5.3 Transaction Costs

In this section, we will introduce the transaction cost, which is an important factor in practical scenarios. In practice, each transaction of wealth from one asset to another is charged with transaction fees. The transaction cost is imposed by markets, and a portfolios behavior cannot change the properties of transaction costs, such as commission rates or tax rates. Usually we assume the transaction fee follows a proportional model, which means rebalancing a portfolio incurs transaction costs on every buy and sell operation, based upon a transaction cost rate of $\gamma \in (0, 1)$. Therefore the transaction cost for a rebalancing from $\hat{b}_{t-1}$ to b_t is computed as:

$$
\frac{\gamma}{2} \times \sum_{i=1}^{m} |b_{t,i} - \hat{b}_{t-1,i}|
\tag{9}
$$

Therefore the cumulated wealth after n periods becomes:

$$
W_n^{\gamma} = W_0 \prod_{t=1}^{n} [(b_t \cdot x_t) \times (1 - \frac{\gamma}{2} \times \sum_{i=1}^{m} |b_{t,i} - \hat{b}_{t-1,i}|)]
\tag{10}
$$

Notice that the main intuition of Passive-Aggressive portfolio optimization is to keep the portfolio unchanged unless the requirement can not be satisfied. This avoids unnecessary rebalancing of wealth among assets and saves the transaction costs induced.

6 Experiment

We conduct extensive experiments on several real-world datasets to evaluate the performances of our strategy and make comparisons with state-of-the-art approaches.

6.1 Experiment Setting

In our experiment, we use the real-world datasets that are frequently used in related works. There are four datasets that contain price relatives of assets from US and Global markets. The time frames of these datasets range from decades to years, which reflect the performances of both long-term and short-term portfolio selections. The details of the datasets are listed in Table. 2. In the experiment, we use the metrics that are adopted in the literatures for evaluation: i.e. the total wealth achieved at the final period.

Table 2. Statistics of the real-world datasets for experiment.

Dataset	Region	Time Frame	# Periods	# Assets
NYSE(O)	US	Jul.3rd 1962-Dec. 31st 1984	5651	36
NYSE(N)	US	Jan.1st 1985-Jun 30th 2010	6431	23
SP500	US	Jan.2nd 1998-Jan. 31st 2003	1276	25
MSCI	Global	Apr. 1st 2006-Mar.31st 2010	1043	24

6.2 Comparison Approaches

We select the state-of-art algorithms (most of them have been introduced in the related works) for comparison, including those Benchmark algorithms (Market, Best-Stock and BCRP), follow-the-winner algorithms (UP, EG, ONS), pattern-matching algorithms (B^k, B^{NN}, CORN, Anticor) and all the variants of mean reversion algorithms (PAMR, CWMR, OLMAR, RMR). For all the algorithms above, we choose the parameters with best performances as reported in related works. Notice that we select some algorithms that use information in hindsight for comparison (which are strong baseline algorithms). For the default setting of BMAR, we set $W = 8$ and $N = 9$ for BMAR-1 and BMAR-2. The other parameters are chosen as: $\eta = 1, \varepsilon = 5$ for all datasets.

6.3 Performance Evaluation

We present the cumulative wealth of our strategy and the comparative approaches in table. 3. As shown in the table, our strategy achieves the best performance on all datasets and outperforms other comparative algorithms significantly on long-term portfolio selection problems, i.e. on NYSE(O) and NYSE(N). Notice that the parameters fit for each dataset can be different, we also list the best performances of algorithms for comparison. The results show that the (including best or conventional) performances achieved by BMAR are better than other comparison algorithms. We also conduct significance test (following [6]) on the performances and the results are listed in Table. 4. The significance tests shown above indicate that the performance of our strategy is significantly better on all datasets, which is not the consequence of luck. Notice that our

Table 3. Cumulative Wealth on four datasets.

Categories	Approaches	NYSE(O)	NYSE(N)	SP500	MSCI
	Market	14.5	18.06	1.34	0.91
Baselines	Best Stock	54.14	83.51	3.78	1.50
	BCRP	250.60	120.32	4.07	1.51
Follow	UP	26.68	31.49	1.62	0.92
the winner	EG	27.09	31.00	1.63	0.93
	ONS	109.19	21.59	3.34	0.86
	B^K	1.08E+09	4.64E+03	2.24	2.64
Pattern	B^{NN}	3.35E+11	6.80E+04	3.07	13.47
Matching	$CORN$	1.48E+13	5.37E+05	6.35	26.10
	Anticor	2.41E+08	6.21E+06	5.89	3.22
	PAMR	5.14E+15	1.25E+06	5.09	15.23
	CWMR	6.49E+15	1.41E+06	5.90	17.28
	OLMAR-1	3.68E+16	2.54E+08	5.83	16.39
Mean	OLMAR-1(max)	1.62E+17	3.95E+08	20.91	25.49
Reversion	OLMAR-2	1.09E+18	5.10E+08	8.63	21.21
	OLMAR-2(max)	2.19E+18	2.84E+09	14.63	27.05
	RMR	1.64E+17	3.25E+08	8.28	16.76
	RMR(max)	2.81E+17	4.73E+08	17.05	19.07
	BMAR-1	**2.02E+18**	**8.95E+08**	**9.99**	23.01
Our	BMAR-1(max)	**4.59E+18**	**5.04E+09**	**16.02**	**30.98**
Strategy	BMAR-2	**8.11E+18**	**1.16E+09**	**10.11**	22.93
	BMAR-2(max)	**8.33E+18**	**6.64E+09**	**26.45**	**30.04**

Table 4. Significance Test of the Performances on experimental datasets. MER means Mean Excess Return; All the statistics are preferred to be higher; p-value is expected to be low.

Metrics	NYSE(O)	NYSE(N)	SP500	MSCI
Size	5651	6431	1276	1043
MER (BMAR-1)	0.0081	0.0038	0.0024	0.0033
MER (Market)	0.0005	0.0005	0.0003	0.0000
Winning Ratio	0.5735	0.5400	0.5306	0.5916
α	0.0075	0.0032	0.0020	0.0033
β	1.2884	1.1166	1.2852	1.2085
t-statistics	16.7284	8.1011	2.5402	6.3251
p-value	0.0000	0.0000	0.0056	0.0000

approach outperforms other baselines significantly, especially on datasets NYSE(O)and NYSE(N). The reason is that the NYSE datasets contain relatively long time periods, and the wealth accumulated has a "Matthew Effect" (the accumulated wealth will be increased with time passing by, the longer it goes, the more wealth will be accumulated).

6.4 Parameter Sensitivity

Notice that our strategy has several parameters: W (which is the maximum window size of experts generated from moving mean reversion) and η for strategy BMAR-1; α and η for strategy BMAR-2. The threshold for portfolio optimization ε is also a key parameter. We conduct experiments on all datasets with different values of the parameters.
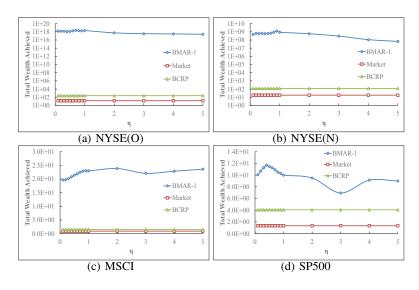
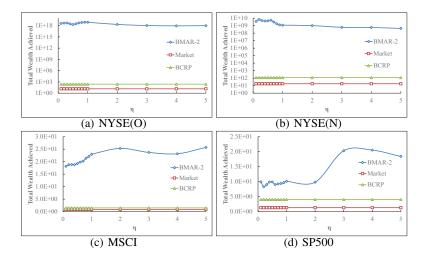**Fig. 2.** Parameter Sensitivity of η with $\varepsilon = 5, W = 8$ in BMAR-1

Fig. 3. Parameter Sensitivity of η with $\varepsilon = 5, N = 9$ in BMAR-2

Impact of η Notice that we use P_2 norm of the difference between prediction and true price relative as loss function, the value of η therefore has two effects: first, it scales the loss function into $[0, 1]$; second, it evaluates the weights assigned to each expert considering their performances. Therefore we conduct experiments with different values of η on the datasets to show their impact in Fig. 2 and Fig. 3. Notice that the two strategies are different according to their ways of estimating the price relatives at each period. The impact of η also varies with the two strategies. Generally, choosing $\eta = 1$ guarantees relative good performances on all the datasets, which is adopted in the experiments shown in Table.3.
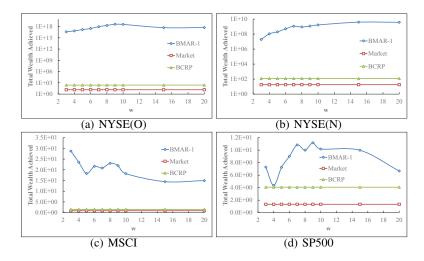
(a) NYSE(O)   (b) NYSE(N)

(c) MSCI   (d) SP500

Fig. 4. Parameter Sensitivity of W with $\varepsilon = 5, \eta = 1$ in BMAR-1

Impact of W BMAR-1 generates experts with different window sizes $w \in [2, W]$, each expert estimates the asset price as the mean of prices in most recent w periods. We choose different values of W to generate experts, where each W means $W - 1$ experts with window sizes are generated. The results are shown in Fig. 4. Notice that the impacts of W are different on the datasets, this is due to the fact that the optimal window sizes for the experts to work on different datasets are also different: the optimal window size for MSCI can be relatively low compared with other datasets. We find that the strategy can achieve consistently good performances when $W \in [8, 10]$ and set it as a conventional value.

Impact of ε The passive-aggressive portfolio optimization technique is applied in our scheme, which tends to keep the portfolios same unless they fail to reach the requirement of return from each period. Therefore, we conduct experiments with different values of ε to show the impact. The impact of ε on our strategies are similar. As shown in Fig. 5, both BMAR-1 and BMAR-2 achieve good performances on all datasets when $\varepsilon \in [5, 10]$. Similarly, we choose $\varepsilon = 5$ as a conventional setting.

6.5 Performance Under Transaction Costs

We also conduct experiments with different transaction cost ratio since it is an unavoidable issue in practice. We alter the transaction cost ratio from 0% to 1% and compute the cumulative wealth of different strategies. The results are presented in Fig. 6.

Judging from the results, the transaction costs has a significant impact on the wealth return. When the transaction cost ratio is greater than 0.005, the wealth achieved on most datasets is close to 0. Since our algorithms can still outperform the baselines, they have good scalability for transaction costs. Notice that the real transaction cost ratio is usually below 0.005, our algorithm can work well in practice.
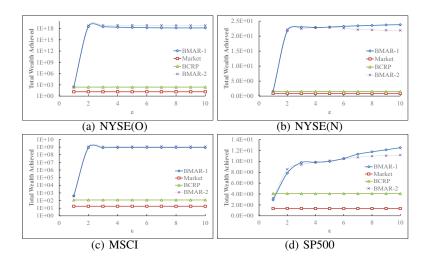
Fig. 5. Parameter Sensitivity of ε with $\eta = 1, w = 8$ in BMAR-1 and $\eta = 1, N = 9$ in BMAR-2
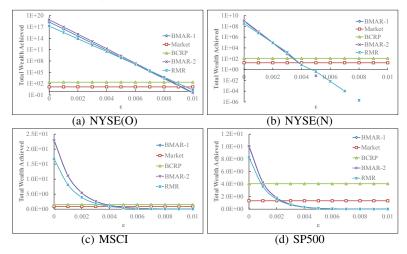

Fig. 6. Performance with different transaction cost ratios

7 Conclusion

In this paper, we consider the online portfolio selection problem from the perspective of mean reversion and meta learning. So far, mean reversion strategies have achieved best empirical results, however they face limits of unknown window size and ignores the temporal heterogeneity of different periods. Meanwhile they are easily affected by outliers and noises in the data. We utilize meta learning to tackle the limits and propose Boosting Moving Average Reversion (BMAR) strategies. The experiments on real-world datasets show that BMAR outperforms state-of-the-art strategies. We believe more accurate prediction of price relatives can further improve the performances and we will consider this as future works.

8 Acknowledgement

References

1. Agarwal, A., Hazan, E., Kale, S., Schapire, R.E.: Algorithms for portfolio management based on the newton method. pp. 9–16. ICML '06, ACM, New York, NY, USA (2006)
2. Borodin, A., Elyaniv, R., Gogan, V.: Can we learn to beat the best stock. Journal of Artificial Intelligence Research 21(1), 579–594 (2004)
3. Cesa-Bianchi, N., Lugosi, G.: Prediction, Learning, and Games. Cambridge University Press, New York, NY, USA (2006)
4. Cover, T.M., Thomas, J.A.: Elements of information theory. John Wiley & Sons (2012)
5. Gaivoronski, A.A., Stella, F.: Stochastic nonstationary optimization for finding universal portfolios. Annals of Operations Research 100(1), 165–188 (2000)
6. Grinold, R., Kahn, R.: Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk. McGraw-Hill Education (1999)
7. Gyorfi, L., Lugosi, G., Udina, F.: Nonparametric kernel-based sequential investment strategies. Mathematical Finance 16(2), 337–357 (2006)
8. Helmbold, D.P., Schapire, R.E., Singer, Y., Warmuth, M.K.: On line portfolio selection using multiplicative updates. Mathematical Finance 8(4), 325–347 (1998)
9. Huang, D.J., Zhou, J., Li, B., Hoi, S., Zhou, S.: Robust median reversion strategy for online portfolio selection. IEEE Transactions on Knowledge and Data Engineering 28(9), 2480–2493 (2016)
10. Kalai, A., Vempala, S.: Efficient algorithms for universal portfolios. Journal of Machine Learning Research 3(3), 423–440 (2003)
11. Kelly, J.L.: A new interpretation of information rate. Bell System Technical Journal 35(4), 917–926 (1956)
12. Laszlo, G., Frederic, U., Harro, W.: Nonparametric nearest neighbor based empirical portfolio selection strategies. Statistics & Decisions 26(2), 145–157 (2008)
13. Li, B., Hoi, S.C.H.: Online portfolio selection: A survey. ACM Comput. Surv. 46(3), 1–36 (Jan 2014)
14. Li, B., Hoi, S.C.H., Sahoo, D., Liu, Z.: Moving average reversion strategy for on-line portfolio selection. Artificial Intelligence 222, 104–123 (2015)
15. Li, B., Hoi, S.C.H., Zhao, P., Gopalkrishnan, V.: Confidence weighted mean reversion strategy for online portfolio selection. ACM Trans. Knowl. Discov. Data 7(1), 1–38 (Mar 2013)
16. Li, B., Hoi, S.C., Gopalkrishnan, V.: Corn: Correlation-driven nonparametric learning approach for portfolio selection. ACM Trans. Intell. Syst. Technol. 2(3), 1–29 (May 2011)
17. Li, B., Zhao, P., Hoi, S.C.H., Gopalkrishnan, V.: Pamr: Passive aggressive mean reversion strategy for portfolio selection. Machine Learning 87(2), 221–258 (2012)
18. Lo, A.W., Mackinlay, A.C.: When are contrarian profits due to stock market overreaction. Review of Financial Studies 3(2), 175–205 (1989)
19. MARKOWITZ, H.: Portfolio selection. The Journal of Finance 7, 1, 77C91 (1952)
20. Merton, R.C.: On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8(4), 323–361 (1980)
21. Suykens, J.A.: Advances in learning theory: methods, models, and applications, vol. 190. IOS Press (2003)