

Vector Autoregressive Weighting Reversion Strategy for Online Portfolio Selection

Xia Cai

School of Information Science and Technology, Fudan University, China
xcai17@fudan.edu.cn

Abstract

Aiming to improve the performance of existing reversion based online portfolio selection strategies, we propose a novel multi-period strategy named “Vector Autoregressive Weighting Reversion” (VAWR). Firstly, vector autoregressive moving-average algorithm used in time series prediction is transformed into exploring the dynamic relationships between different assets for more accurate price prediction. Secondly, we design the modified online passive aggressive technique and advance a scheme to weigh investment risk and cumulative experience to update the closed-form of portfolio. Theoretical analysis and experimental results confirm the effectiveness and robustness of our strategy. Compared with the state-of-the-art strategies, VAWR greatly increases cumulative wealth, and it obtains the highest annualized percentage yield and sharp ratio on various public datasets. These improvements and easy implementation support the practical applications of VAWR.

1 Introduction

The past decade has seen the rapid development of machine learning in business application. A hot research topic in this area is to develop machine-learning techniques for portfolio selection (PS). The goal of PS is to maximize some related performance measures, for example, cumulative return or risk-adjusted return, with the wealth invested in some financial markets in the long run [Li *et al.*, 2011]. One of the most important events in 1952 is that Markowitz first carried out a systematic analysis of PS problem [Markowitz, 1952]. Then in 1956, Kelly presented the “Capital Growth Theory” [Kelly, 1956], which is suitable for multi-period sequential PS. These are two major schools for investigating PS problem. Online learning algorithms devised to learn models for data in a sequential manner are in good agreement with the feature of data sequence in the financial field. So, it can be applied to portfolio research to deal with online PS problem. Related framework can be seen in [Ordentlich and Cover, 1996; Das *et al.*, 2014; Li and Hoi, 2014].

Recently, considerable literature has grown up around the theme of online PS, such as follow-the-winner, follow-the-loser, meta-learning algorithms, and so on [Li and Hoi, 2014]. [Li *et al.*, 2012; 2013] first proposed single-period reciprocal strategies, which adopt passive aggressive online learning algorithm [Crammer *et al.*, 2006] to produce the next portfolio. One major problem in this research concerned is that the information of only one period is inadequate, and it cannot satisfy the operation of real market. After that, some multi-period strategies [Li and Hoi, 2012; Huang *et al.*, 2016] have been advanced, but these are challenged by applying equal consideration to all considered periods. To overcome this drawback, [Ye *et al.*, 2017] presented the strategy of Gaussian weighting reversion, and related performance has been greatly improved. Such approaches, however, have failed to address the correlation between different assets. Meanwhile, these reversion based strategies just tend to focus on cumulative return, while investment risk as an essential aspect has not been particularly optimized. When investors face the risk adjustment in financial markets, their trading portfolios may be different. Therefore, how investors should trade in the presence of trading risks remains an open yet important question.

The key innovations of our paper lie in these aspects: on the one hand, considering that the previous reversion based strategies do not take the relationships between assets into account, and vector autoregressive moving-average (VARMA) algorithm as one of the most mainstream models of multivariate time series directly studies the correlation of varieties, we first associate these two points to apply VARMA to online PS problem for exploring the dynamic relationships of different assets. Parameter identification of the traditional VARMA model is done in a batch manner. Here, we combine it with online learning algorithm to realize dynamic parameter updating. So, the parameters of prediction model can be adaptively changed with the market, and this can effectively improve the accuracy of the next price prediction. On the other hand, concerning determining the portfolio, based on passive aggressive online technique, we propose a modified methodology through introducing the term of risk minimization to the objective function for a trade-off between cumulative experience and investment risk under a certain constrain of expected return. This new strategy is named “Vector Autore-

gressive Weighting Reversion” (VAWR). To validate the proposed strategy, we led extensive experiments on various real datasets. Our experimental results show that VAWR outperforms the existing strategies including state-of-the-art ones.

2 Related Work

Much of current literature on PS pays attention to the principle of Kelly’s investment [Kelly, 1956]. The most common benchmark is *Market*, in which one invests wealth among different assets with an initial portfolio and holds until the end. *Constantly Re-balanced Portfolios* (CRP) is another benchmark, and it keeps a fixed portfolio on each asset for all periods. *Best CRP* (BCRP) [Cover, 1991], over a whole market sequence, is an optimal strategy if market is i.i.d. *Successful CRP* (SCR) and *Weighted SCR* (WSCR) [Gaivoronski and Stella, 2000] implicitly predict the next price ratio by all historical data with a uniform probability.

Some algorithms do not focus on price prediction. The pioneering study of online PS was reported by [Ordentlich and Cover, 1996], and they put forward *Universal Portfolios* (UP). [Huang et al., 2015] pointed out that the effectiveness of avoiding re-balancing the portfolio when transaction costs outweigh trading benefit, whose strategy is called *Semi-universal Portfolio* (SUP). After that, [Li et al., 2018] proposed a novel framework named *Transaction Cost Optimization* (TCO) to further improve the strategy’s performance in the case of non-zero transaction costs. Considering portfolio risks, many papers focus on Markowitz’s mean-variance framework. However, when portfolio’s size increases, the estimation for the mean and variance of assets return becomes less reliable [Shen et al., 2019]. So related work devotes to eliminating the estimation risks caused by the uncertainty of parameters [Huang et al., 2011; Shen and Wang, 2017].

Besides, there is increasing literature on predicting the next price ratio by partial historical data. *Exponential Gradient* (EG) [Helmbold et al., 1998] uses the last price ratio as the prediction to next price ratio. *Passive Aggressive Mean Reversion* (PAMR) [Li et al., 2012] and *Confidence Weighted Mean Reversion* (CWMR) [Li et al., 2013] predict the next price ratio with the inverse of the last price ratio, which exploits in essence the mean reversion principle. *Online Moving Average Reversion* (OLMAR) was advanced by [Li and Hoi, 2012], and it is based on the moving average of historical data in a fixed window to predict the next price and explores average reversion. [Huang et al., 2016] proposed *Robust Median Reversion* (RMR) using L1-median of stock prices to explore the reversion mechanism. [Ye et al., 2017] presented *Gaussian Weighting Reversion* (GWR), which adopts a fixed Gaussian function to exploit the “time validity” of historical data and achieves Gaussian average reversion.

Recently, deep learning (DL) in portfolio management is popular [Selvin et al., 2017; Liang et al., 2018]. Nevertheless, DL is hugely dependent on the structure of neural network and needs large-scale datasets for training [Cai and Ye, 2019]. Many existing methods in this field are not tested on public datasets yet and lack theoretical analysis for performance guarantee [Liang et al., 2018]. It still has much room for further development in the application of DL in this field.

3 Problem Formulation

3.1 Problem Setting

The financial market with d assets for n trading periods to be invested is considered. At the end of the t^{th} period, the prices of assets are represented as a vector of closing price $\mathbf{p}_t \in R_+^d$, and each element p_t^i means the closing price of asset i . The price ratio vector $\mathbf{x}_t \in R_+^d$ represents the variation of each asset price, where each element x_t^i indicates the ratio of current closing price (corresponding to the t^{th} period) to the previous closing price (corresponding to the $(t-1)^{th}$ period) of asset i , i.e., $x_t^i = p_t^i / p_{t-1}^i$. Denote $\mathbf{x}_1^n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ as a sequence of price ratio from the first to the n^{th} period.

At the beginning of the t^{th} period, the capital is allocated in d assets through a portfolio vector \mathbf{b}_t , and each element of \mathbf{b}_t is used as the proportion of wealth invested in each asset. Generally, because we assume the portfolio is self-financed and no margin/short is inadmissibility, it implies each element of portfolio vector is non-negative and the sum of a portfolio is one, that is, $\mathbf{b}_t \in \Delta d$, where $\Delta d = \{\mathbf{b}_t : \mathbf{b}_t \in R_+^d, \sum_{i=1}^d b_t^i = 1\}$. The investment procedure is represented as a portfolio strategy, which means $\mathbf{b}_1 = \frac{1}{d} \mathbf{1}$ following a series of mappings $\mathbf{b}_t: (R_+^d)^{t-1} \rightarrow \Delta d, t = 2, 3, \dots$, where $\mathbf{b}_t = \mathbf{b}_t(\mathbf{x}_1^{t-1})$ is the portfolio adopted for the t^{th} trading period, given the past market sequence $\mathbf{x}_1^{t-1} = (\mathbf{x}_1, \dots, \mathbf{x}_{t-1})$. $\mathbf{b}_1^n = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ is the strategy from the first to n^{th} period.

When the t^{th} trading period finishes, the portfolio \mathbf{b}_t brings portfolio’s period return s_t , which shows the wealth grows by a factor of $s_t = \mathbf{b}_t \cdot \mathbf{x}_t = \sum_{i=1}^d b_t^i x_t^i$. As we reinvest, the wealth will get multiplicative increase. So, after n trading periods, the portfolio strategy \mathbf{b}_1^n generates the portfolio’s cumulative wealth S_n , which achieves $\Pi_{t=1}^n \mathbf{b}_t \cdot \mathbf{x}_t$ times the initial wealth, namely, $S_n(\mathbf{b}_1^n, \mathbf{x}_1^n) = S_0 \Pi_{t=1}^n (\mathbf{b}_t \cdot \mathbf{x}_t)$, where S_0 is initial wealth and set to one in this paper. The critical goal of a portfolio manager is to play a strategy \mathbf{b}_1^n to get the maximization of portfolio’s cumulative wealth S_n .

In the above model, we ideally presume no transaction costs, perfect market liquidity, and zero market impact. These are common in the existing work [Li et al., 2012; 2013; Huang et al., 2016; Ye et al., 2017; Cai and Ye, 2019], and we will analyze the effect of transaction costs in Section 5.5.

3.2 Analysis of Existing Work

As stated in Section 2, Kelly’s portfolio selection formula is the foundation of related strategies:

$$\mathbf{b}_{t+1} = \arg \max_{\mathbf{b} \in \Delta_d} \sum_{i=1}^k q_i \log(\mathbf{b} \cdot \hat{\mathbf{x}}_{t+1}^i). \quad (1)$$

Generally, a portfolio manager predicts $\hat{\mathbf{x}}_{t+1}$ with k possible values $\hat{\mathbf{x}}_{t+1}^1, \dots, \hat{\mathbf{x}}_{t+1}^k$, and their corresponding possibilities are q_1, \dots, q_k , respectively. Here, $\hat{\mathbf{x}}_{t+1}^i$ means one possible combination vector of the next price ratio prediction. Different methods adopted by portfolio manager will produce different $\hat{\mathbf{x}}_{t+1}^i$ and q_i , which will also cause different portfolios. Hence a precise and effective prediction method is essential to the success of a PS strategy. Also, some constraints can be

added to the process of generating portfolios with the consideration of the optimization of other indicators while maximizing cumulative return.

Among the related reversion strategies, PAMR implicitly assumes $\hat{\mathbf{x}}_{t+1} = \frac{1}{\mathbf{x}_t} \Rightarrow \frac{\hat{\mathbf{p}}_{t+1}}{\mathbf{p}_t} = \frac{\mathbf{p}_{t-1}}{\mathbf{p}_t} \Rightarrow \hat{\mathbf{p}}_{t+1} = \mathbf{p}_{t-1}$, that is, PAMR supposes that the next prices $\hat{\mathbf{p}}_{t+1}$ will revert to the last prices \mathbf{p}_{t-1} . It is a single-period strategy and does not make full use of market information. OLMAR improves PAMR by predicting the next prices with moving average,

i.e., $\hat{\mathbf{p}}_{t+1} = \text{avg}_t(L) = \frac{1}{L} \sum_{i=t-L+1}^t \mathbf{p}_i$, where L denotes the size of moving window. Then, L1-median is used by RMR as the estimator, which is the point with minimal sum of Euclidean distances to L given points of price data. In math,

$\hat{\mathbf{p}}_{t+1} = \arg \min_{\mathbf{p}} \sum_{i=0}^{L-1} \|\mathbf{p}_{t-i} - \mathbf{p}\|$, where $\|\cdot\|$ denotes the Euclidean norm. OLMAR and RMR use multi-period historical data, whereas attaching equal importance. In GWR, a Gaussian function is adopted to weight the historical data in a sliding window for all assets, the aim of which is paying more attention to recent data. The next predicted prices are expressed

as $\hat{\mathbf{p}}_{t+1} = \frac{\sum_{i=t-L+1}^t e^{-\frac{(t+1-i)^2}{2\tau^2}} \mathbf{p}_i}{\sum_{i=t-L+1}^t e^{-\frac{(t+1-i)^2}{2\tau^2}}}$, where τ is a parameter of Gaussian function and L is the size of sliding window. In addition, the method of double estimations is developed to mitigate the effect of noise and outliers. But the weights in GWR are fixed for all periods and assets. These estimators do not adequately take the statistical nature of the data and prediction model into account.

Next, we conduct the analysis of competing experimental results about the above multi-period strategies. The way of OLMAR using historical information leads to its disadvantages in return and risk-adjusted return. For GWR, it is meaningful to consider historical information in a non-equal weighting way, and its return has an order of magnitude growth compared with the previously state-of-the-art RMR. However, they all ignore the relationships of assets in prediction model. The information of this correlation is important for price prediction. We try to fully exploit the property of “time validity” and explore the relationships of price trend among different assets. This is considering the influence factors of asset prices from two dimensions. Moreover, although GWR performs well in terms of cumulative wealth, it is not as good as RMR in risk-related indicators. These strategies all do not take specific action to decrease trading risks during producing a portfolio. These drawbacks of existing strategies encourage us to design a new and effective online PS strategy.

4 The VAWR Strategy

4.1 Methodology

The proposed VAWR explores the dynamic relationships between different assets in different historical periods to obtain next price ratio $\hat{\mathbf{x}}_{t+1}$ via online time series prediction technology VARMA, which is different from the traditional way of updating model parameters in batch. Then VAWR adopts modified passive aggressive online learning to exploit the re-

versal property of financial market, which means to maximize the expected return $\mathbf{b} \cdot \hat{\mathbf{x}}_{t+1}$ while trying to keep the last portfolio information and minimize the portfolio risks.

In order to estimate the closing prices of the t^{th} period, the traditional VARMA(k, l) model is given [Yang *et al.*, 2018]:

$$\hat{\mathbf{p}}_t = \sum_{j=1}^k \Phi_j \mathbf{p}_{t-j} + \sum_{j=1}^l \Theta_j \epsilon_{t-j} + \epsilon_t. \quad (2)$$

Here, VARMA(k, l) model is characterized by two horizon terms k, l , and coefficient matrices $\Phi_j \in \mathbb{R}^{d \times d}$ and $\Theta_j \in \mathbb{R}^{d \times d}$. Find that $\hat{\mathbf{p}}_t$ is hard to predict, because the noise terms ϵ_{t-j} are unknown at any time, and coefficient matrices are required to be optimized. Therefore, for simplifying the problem, we use the below equation to follow the underlying VARMA model, which can be seen as a modified VARMA model with horizon term L :

$$\hat{\mathbf{p}}_t = \mathbf{W}^t \mathbf{P}_{t-1} = \sum_{j=1}^L \mathbf{W}_j^t \mathbf{p}_{t-j}, \quad (3)$$

where $\mathbf{W}^t = [\mathbf{W}_1^t, \mathbf{W}_2^t, \dots, \mathbf{W}_L^t]$, \mathbf{W}_j^t is weighting matrix, $\mathbf{W}_j^t \in \mathbb{R}^{d \times d}$, and $\mathbf{P}_{t-1} = [\mathbf{p}_{t-1}^T, \mathbf{p}_{t-2}^T, \dots, \mathbf{p}_{t-L}^T]^T$. The price prediction for asset i can be given:

$$\hat{p}_t^i = \sum_{j=1}^L \mathbf{W}_j^t(i, :) \mathbf{p}_{t-j}, \quad (4)$$

where $\mathbf{W}_j^t(i, :)$ is the i^{th} row of \mathbf{W}_j^t and represents the weighting values for asset i . Therefore, the problem of forecasting price is transformed into the problem of solving the modified VARMA model.

Next, the modified VARMA model is solved by online learning. When the t^{th} period finishes, the real closing price of each asset is revealed, and we receive the loss of each asset. According to equation (4), the formula definition of the loss for asset i is given as follows:

$$l_t(\mathbf{W}^t(i, :)) = l_t(p_t^i, \hat{p}_t^i) = (p_t^i - \sum_{j=1}^L \mathbf{W}_j^t(i, :) \mathbf{p}_{t-j})^2, \quad (5)$$

where $\mathbf{W}^t(i, :)$ is the i^{th} row of \mathbf{W}^t and represents the weighting matrix related to asset i in the sliding window after the $(t-1)^{th}$ iteration updates, and $\mathbf{W}^t(i, :) \in \mathbb{R}^{1 \times dL}$. The above formula shows the difference between predicted and real value of asset i . For online convex optimization, loss function should satisfy α -exp-concavity property, that is, there is an exp-concavity parameter $\alpha > 0$ such that $\exp(-\alpha l_t)$ is concave for all t . Our goal is to minimize the loss of each asset:

$$\mathbf{W}^{t+1}(i, :) = \arg \min l_t(\mathbf{W}^t(i, :)). \quad (6)$$

Through experiencing losses, the weighting matrix of modified VARMA model is updated, and the next prediction $\hat{\mathbf{p}}_{t+1}$ can be made. During this process, we apply a first-order online convex optimization solver, Online Gradient Descent [Zinkevich, 2003]. Every time updates the weighting matrix, it takes steps which are proportional to the negative of

Algorithm 1 PriceRatioPrediction($\mathbf{P}_t, \mathbf{W}^t, L, \eta$)

Input: historical data \mathbf{P}_t , weighting matrix \mathbf{W}^t , parameter L , learning rate η , and the number of assets d .

Output: predicted $\hat{\mathbf{x}}_{t+1}$

```

1: Procedure:
2: for  $i = 1 \rightarrow d$  do
3:   Predict  $\hat{p}_t^i = \sum_{j=1}^L \mathbf{W}_j^t(i, :) \mathbf{p}_{t-j}$ .
4:   Release  $p_t^i$  and suffer a loss  $l_t(\mathbf{W}^t(i, :))$ .
5:    $\nabla_t = \nabla l_t(\mathbf{W}^t(i, :))$ 
6:    $\mathbf{W}^{t+1}(i, :) \leftarrow \prod_{\kappa_i}(\mathbf{W}^t(i, :) - \eta \nabla_t)$ 
7: end for
8:  $\hat{\mathbf{p}}_{t+1} = \mathbf{W}^{t+1} \mathbf{P}_t$ , and  $\hat{\mathbf{x}}_{t+1} = \frac{\hat{\mathbf{p}}_{t+1}}{\mathbf{P}_t}$ 
9: return  $\hat{\mathbf{x}}_{t+1}$ 

```

the gradient of loss function at the current point. Algorithm 1 shows the price prediction for general convex loss function. The notation \prod_{κ_i} refers to the Euclidean projection onto the nearest point in κ_i , i.e., $\prod_{\kappa_i}(\mathbf{y}) = \arg \min_{\mathbf{x} \in \kappa_i} \|\mathbf{x} - \mathbf{y}\|_2$, and κ_i is defined as:

$$\kappa_i = \left\{ \mathbf{W}(i, :) \in \mathbb{R}^{1 \times dL}, \sum_{k=1}^{dL} \mathbf{W}(i, k) = 1, \mathbf{W}(i, k) \geq 0 \right\}. \quad (7)$$

Based on the prediction $\hat{\mathbf{p}}_{t+1}$, the portfolio \mathbf{b}_{t+1} is figured out according to the modified passive aggressive online learning algorithm, which can exploit the reversal phenomenon in financial markets and minimize the portfolio risks. Passive aggressive online learning algorithm [Crammer *et al.*, 2006] is generally used for classification. When the classification is correct, this algorithm passively keeps unchanged and aggressively obtains a new solution when it is wrong.

Optimize the following problem to form next portfolio:

$$\begin{aligned} \mathbf{b}_{t+1} = \arg \min_{\mathbf{b} \in \Delta^d} \|\mathbf{b} - \mathbf{b}_t\|^2 + \lambda \mathbf{b}^T \Sigma \mathbf{b} \\ \text{s.t. } \mathbf{b} \cdot \hat{\mathbf{x}}_{t+1} \geq \delta, \quad \mathbf{b}^T \vec{\mathbf{1}} = 1 \end{aligned} \quad (8)$$

Here, $\|\cdot\|$ means the Euclidean norm, Σ is co-variance matrix calculated by historical closing prices of assets, and λ is a trade-off factor. $\hat{\mathbf{x}}_{t+1}$ is worked out by Algorithm 1, δ is the given threshold of expected return, and $\vec{\mathbf{1}}$ is a vector with dimension d , that is, $\vec{\mathbf{1}} \in \mathbb{R}^d$.

The optimization objective (8) consists of two parts: the term about cumulative experience $\|\mathbf{b} - \mathbf{b}_t\|^2$ and risk term $\mathbf{b}^T \Sigma \mathbf{b}$. The goal of the above problem is to find an optimal portfolio by minimizing the deviation from the last portfolio \mathbf{b}_t as well as trading risks under the designed constraint $\mathbf{b} \cdot \hat{\mathbf{x}}_{t+1} \geq \delta$. Specifically, if this constraint is satisfied, namely, the expected return is higher than the threshold δ and the last portfolio can minimize current risks, the resulting portfolio passively keeps the same as the last. If not, a new portfolio will be aggressively calculated to promise that the expected return is higher than δ , and new portfolio is as close as possible to the previous. Such formulation explains the idea of market reversion. Actually, $\hat{\mathbf{x}}_{t+1}$ is the predicted price ratio via online modified VARMA, while the constraint $\mathbf{b} \cdot \hat{\mathbf{x}}_{t+1} \geq \delta$ implies that the next prices will reverse to our prediction.

Algorithm 2 The VAWR algorithm

Input: market sequence \mathbf{x}_1^n , reversion threshold δ , learning rate η , window size $L \geq 2$, and parameter λ .

Output: cumulative wealth after the n^{th} period S_n .

```

1: Procedure:
2: Initialization:  $\mathbf{b}_1 = \frac{1}{d} \vec{\mathbf{1}}$ ,  $S_0 = 1$ . Set  $\mathbf{W}^1(i, :) \in \kappa_i, i = 1, 2, \dots, d$  arbitrarily.
3: for  $t = 1 \rightarrow n$  do
4:    $\mathbf{p}_{t+1} = \mathbf{p}_t \cdot \mathbf{x}_{t+1}$ 
5: end for
6: for  $t = 1 \rightarrow n$  do
7:   Receive stock price ratio:  $\mathbf{x}_t$ 
8:   Update cumulative return:
9:      $s_t = \mathbf{b}_t \cdot \mathbf{x}_t$ , and  $S_t = S_{t-1} * s_t$ 
10:  Predict the vector of next price ratio:
11:     $\hat{\mathbf{x}}_{t+1} = \text{PriceRatioPrediction}(\mathbf{P}_t, \mathbf{W}^t, L, \eta)$ 
12:  Update the portfolio  $\mathbf{b}_{t+1}$ :
13:     $\mathbf{b}_{t+1} = \arg \min_{\mathbf{b} \in \Delta^d} \|\mathbf{b} - \mathbf{b}_t\|^2 + \lambda \mathbf{b}^T \Sigma \mathbf{b}$ 
14:    s.t.  $\mathbf{b} \cdot \hat{\mathbf{x}}_{t+1} \geq \delta, \quad \mathbf{b}^T \vec{\mathbf{1}} = 1$ 
15: end for

```

Problem (8) is a standard quadratic programming (QP) problem, which has the optimum solution because covariance matrix is positive semi-definite. The algorithm used for solving QP problem can be primal-dual path-following method based on the Nesterov-Todd scaling. For more background, history, and analysis of related algorithms, refer to [Andersen *et al.*, 2003]. The portfolio solved by the above methods may go out of the simplex domain. Therefore, to ensure that the portfolio is non-negative, the resulting portfolio is finally projected to the simplex domain [Duchi *et al.*, 2008].

Therefore, the whole solving process of VAWR divides into two steps: the first is predicting the next closing price ratio based on online VARMA; the second is determining the portfolio by modified passive aggressive online learning. And the complete process of online PS following the problem setting in Section 3.1 for VAWR is outlined in Algorithm 2.

4.2 Theoretical Analysis

Before analysis, related definitions should be given. We define the regret of asset i as:

$$R_T^i = \sum_{t=1}^T l_t(\mathbf{W}^t(i, :)) - \min_{\mathbf{W}(i, :)} \sum_{t=1}^T l_t(\mathbf{W}(i, :)), \quad (9)$$

where $\mathbf{W}(i, :)$ is fixed coefficient parameter from hindsight and it can produce the smallest total loss. We set

$$D = \max_{\mathbf{W}(i_1, :), \mathbf{W}(i_2, :) \in \kappa_i} \|\mathbf{W}(i_1, :) - \mathbf{W}(i_2, :)\|_2, \quad (10)$$

where D identifies the diameter of κ_i . The upper bound of $\|\nabla l_t(\mathbf{W}^t(i, :))\|$ for all t is denoted by G , and $\mathbf{W}^t(i, :) \in \kappa_i$. We use η_t as learning rate at time point t . According to these definitions, we can prove the following theorem about bounding the regret of Algorithm 1.

Theorem 1 For any market data $\{\mathbf{p}_t\}_{t=1}^T$, Algorithm 1 generates a sequence of online weighting matrix $\{\mathbf{W}^t(i, :)\}_{t=1}^T$,

which guarantees

$$R_T^i = \sum_{t=1}^T l_t(\mathbf{W}^t(i, :)) - \min_{\mathbf{W}} \sum_{t=1}^T l_t(\mathbf{W}(i, :)) < \frac{D^2}{2\eta_T} + \frac{G^2}{2} \sum_{t=1}^T \eta_t.$$

We can get the above from the result of Online Gradient Descent in [Zinkevich, 2003], relying on the fact that our loss function is α -exp-concave. Then according to the work [Yang *et al.*, 2018] about the regret bound of VARMA model used in multivariate time series, we can bound the regret for all asset,

$$R_T = \sum_{t=1}^T l_t(\mathbf{W}^t) - \min_{\mathbf{W}} \sum_{t=1}^T l_t(\mathbf{W}) < d(\frac{D^2}{2\eta_T} + \frac{G^2}{2} \sum_{t=1}^T \eta_t).$$

Besides, if we set $\eta_t = \frac{1}{\sqrt{t}}$, then

$$\sum_{t=1}^T \eta_t = \sum_{t=1}^T \frac{1}{\sqrt{t}} \leq 1 + \int_{t=1}^T \frac{dt}{\sqrt{t}} \leq 1 + [2\sqrt{t}]_1^T \leq 2\sqrt{T} - 1.$$

So, we can get the regret bound of all assets

$$R_T < d(\frac{D^2\sqrt{T}}{2} + G^2(\sqrt{T} - \frac{1}{2})) < O(\sqrt{T}).$$

5 Performance Evaluation

5.1 Datasets

The datasets involved in the experiments are four real public datasets, which have been widely used in the test of online PS strategies [Li *et al.*, 2012; 2013; Huang *et al.*, 2016; Ye *et al.*, 2017; Cai and Ye, 2019], and are summarized in Table 1.

Dataset	Region	Time frame	#Days	#Assets
DJIA	US	1/1/2001-14/1/2003	507	30
NYSE_O	US	3/7/1962-31/12/1984	5651	36
NYSE_N	US	1/1/1985-30/6/2010	6431	23
TSE	CA	4/1/1994-31/12/1998	1259	88

Table 1: The summary of four public datasets

The first one ‘‘DJIA’’ consists of Dow Jones 30 composite shares. The second is from New York Stock Exchange, one benchmark dataset pioneered by [Cover, 1991], and denote it by ‘‘NYSE_O’’. The third is the extension of ‘‘NYSE_O’’ and collected by [Li *et al.*, 2011]. It is set as ‘‘NYSE_N’’. The final dataset ‘‘TSE’’ with ‘‘DJIA’’ was both collected by [Borodin *et al.*, 2004], and is from Toronto Stock Exchange.

5.2 Cumulative Wealth

Table 2 provides the experimental results of cumulative wealth achieved by different strategies. The concrete parameter settings of these strategies keep the same with the original papers. With the goal of consistency in the comparison, we empirically set the size of sliding window $L=5$, $\lambda=0.1$, and $\delta=50$ in the experiments for VAWR. Considering that competitive algorithms start from different initial periods in the original papers, to be fair, the wealth of all algorithms in Table 2 is accumulated from the first period. If the number of past periods is smaller than the size of sliding window, the last closing prices are used as the prediction to next period.

In Table 2, VAWR beats all existing strategies on DJIA, NYSE_O and NYSE_N, and is ranked the 2nd on TSE

Strategies	DJIA	NYSE_O	NYSE_N	TSE
Market	0.79	14.30	18.32	1.57
BCRP	1.20	248.75	119.32	6.71
CRP	0.80	22.57	26.09	1.55
SCRIP	0.75	18.39	19.92	1.47
UP	0.80	26.44	27.41	1.58
SUP	0.76	18.04	19.03	1.56
PAMR	0.70	4.87E+15	1.32E+06	257.78
OLMAR	2.51	7.42E+16	4.42E+08	60.15
RMR	2.69	1.59E+17	3.27E+08	179.59
GWR	2.50	5.45E+17	7.18E+08	460.60
VAWR	2.99	4.14E+18	1.10E+09	346.31

Table 2: The contrast of cumulative wealth.

Stat.	DJIA	NYSE_O	NYSE_N	TSE
Size	507	5651	6431	1259
MER(VAWR)	0.0025	0.0081	0.0038	0.0062
MER(Market)	-0.0005	0.0004	0.0004	0.0002
α	0.0033	0.0077	0.0034	0.0058
β	1.2569	1.0885	1.0734	1.4873
t-statistics	2.7634	16.9448	7.7557	3.4910
p-value	0.0030	0.0000	0.0000	0.0002

Table 3: Statistical Test of VAWR.

(only inferior to GWR). Meanwhile, there is a significant and remarkable growth of cumulative wealth on NYSE_O and NYSE_N. Compared with ever state-of-the-art strategies RMR and GWR, the cumulative wealth of VAWR on DJIA, NYSE_O, and NYSE_N is increased by 11.23%, 658.90%, and 52.45%, respectively, and completely defeats RMR on four datasets. So from the perspective of cumulative wealth, VAWR performs more outstanding than other strategies, which shows the price prediction model in VAWR is efficient.

In addition, the widely recognized t-test statistics are used to analyze the relationships between the achieved wealth and simple luck, which is a common way in the fund management industry [Li *et al.*, 2012]. The results of p-value, as listed in Table 3, indicate that the obtained wealth whether is caused by simple luck or not. The corresponding chance for getting excess return by luck is at most 0.0030 on DJIA. This reveals that our proposed VAWR is reliable, which can be superior to competitors with high confidence. Therefore, the effectiveness and practicability of our strategy are further illustrated.

5.3 APY and SR

Annualized Percentage Yield (APY) is a criterion to measure return with compound interest effect. This interest effect indi-

Criteria	Strategy	DJIA	NYSE_O	NYSE_N	TSE
APY	RMR	0.6387	5.0522	1.1256	1.8239
	GWR	0.5820	5.4005	1.1909	2.4093
	VAWR	0.7282	6.0182	1.2268	2.2203
SR	RMR	1.1698	8.8740	1.9232	1.9341
	GWR	1.0499	9.3376	2.0165	2.4771
	VAWR	1.3203	10.5940	2.0282	2.2752

Table 4: Comparison of APY and SR.

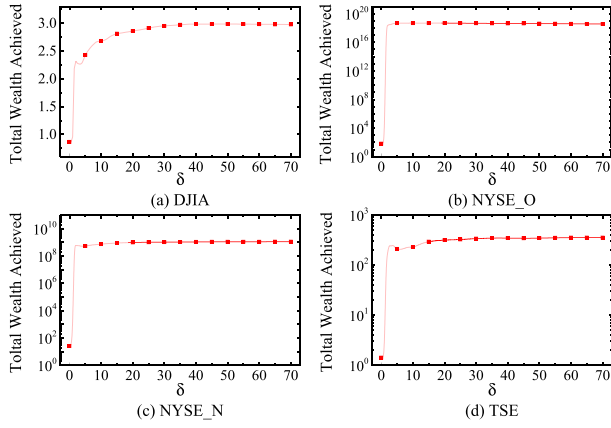


Figure 1: Sensitivity tests of the parameter δ .

icates that a strategy can achieve average wealth growth within one year and $APY = \sqrt[y]{S_n} - 1$, where y is the number of years corresponding to n trading periods. *Sharpe Ratio* (SR) is a main indicator to measure wealth and risks synthetically. It is known as “reward-to-variability ratio” and concludes the mean and variance with a simple measure of risk-adjusted return. Here, we use annualized SR calculated by $SR = \frac{APY - R_f}{\sigma_p}$, where R_f is the risk-free return (typically the return of Treasury bills, fixed at 4% [Li *et al.*, 2012]), and σ_p is annualized standard deviation of period return.

The comparison results of related criteria are included in Table 4. Compared with GWR, VAWR gets the largest APY and the highest SR in DJIA, NYSE_O and NYSE_N datasets, and beats RMR in all datasets. These results indicate that it is meaningful to add the minimization item of risk to passive aggressive online learning technology. And this consideration makes our strategy capable of achieving a good trade-off between return and risk.

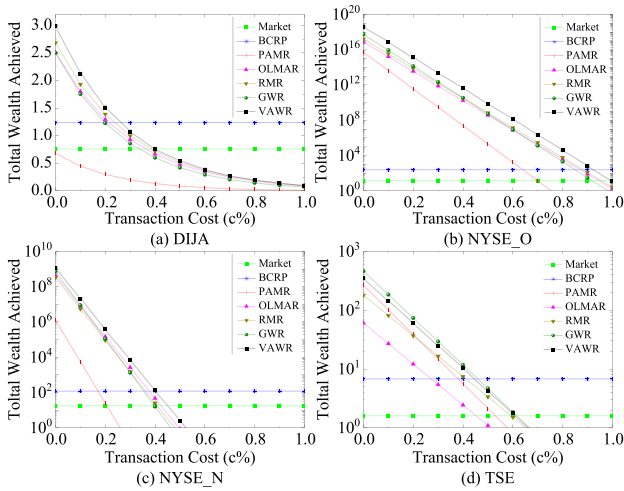


Figure 2: The effect of transaction costs on the total wealth.

5.4 Parameter Sensitivity

The sensitivity of parameter δ in VAWR is examined on four datasets, which shows how various selections of δ give effect on the final wealth achieved. The results are presented in Figure 1. It reflects that the cumulative wealth first sharply grows with the increase of δ , and then it turns steady when δ reaches a certain threshold, which is related to the specific dataset. Therefore, VAWR is no longer sensitive when δ reaches a certain threshold. We finally set $\delta = 50$ in other experiments.

5.5 The Effect of Transaction Costs

Actually, transaction cost is another important and non-negligible problem, which contains commission fees and taxes demanded by brokers and governments [Cover and Thomas, 1990]. It should be pointed out that portfolio cannot affect transaction costs. We follow the proportional transaction cost model proposed by [Blum and Kalai, 1999; Borodin *et al.*, 2004], which has wide application in the literature [Li *et al.*, 2012; Huang *et al.*, 2016; Ye *et al.*, 2017]. Specifically, at the beginning of the t^{th} trading day, the portfolio manager adjusts his/her wealth from the last unregulated portfolio $\hat{\mathbf{b}}_{t-1}$ to a new portfolio \mathbf{b}_t , and this action will produce transaction costs: $Cost(t) = \frac{c}{2} \sum_{j=1}^d |b_t^j - \hat{b}_{t-1}^j|$,

where $\hat{b}_{t-1}^j = \frac{b_{t-1}^j x_{t-1}^j}{\mathbf{b}_{t-1} \cdot \mathbf{x}_{t-1}}$ and c is transaction cost rate changed from 0% to 1%. The total wealth is calculated: $S_n^c = S_0 \prod_{t=1}^n [(\mathbf{b}_t \cdot \mathbf{x}_t) \times (1 - Cost(t))]$.

All experimental results of transaction costs are reported in Figure 2. When transaction cost rate increases, the cumulative wealth gained by all competing strategies gradually decreases. However, the cumulative wealth curve of VAWR is always above others and only inferior to GWR on TSE, which shows that our VAWR gains more wealth under the same rate and can withstand higher transaction costs. The break-even rate of VAWR with respect to Market ranges from 0.4% to 1.0%. So, even when the trade is not frictionless in practical application, VAWR can still perform satisfactorily.

6 Conclusion

This paper proposes a novel multi-period online portfolio selection strategy VAWR. Firstly, VAWR learns the dynamic relationships between different assets in different trading periods based on online VARMA model, which can help take full advantage of historical data. Then, considering the minimization of risks while keeping cumulative experience as much as possible for producing a portfolio, we propose a modified passive aggressive online learning technique. The experimental results show that VAWR is better than GWR and RMR in cumulative return, APY and SR. Furthermore, VAWR passes the statistical test and is robust to different parameter settings. Also it can withstand higher transaction costs. These results consistently present the feasibility and validity of VAWR. However, there are still some problems: the performance on TSE is not as good as GWR, which is likely to be caused by using a definite size of weighted window. A better solution is to make the data-driven choice. Also, it needs to explore more advanced models for further optimizing regret bound.

References

- [Andersen *et al.*, 2003] Erling D Andersen, Cornelis Roos, and Tamas Terlaky. On implementing a primal-dual interior-point method for conic quadratic optimization. *Mathematical Programming*, 95(2):249–277, 2003.
- [Blum and Kalai, 1999] Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999.
- [Borodin *et al.*, 2004] Allan Borodin, Ran El-Yaniv, and Vincent Gogan. Can we learn to beat the best stock. *Journal of Artificial Intelligence Research*, 21:579–594, 2004.
- [Cai and Ye, 2019] Xia Cai and Zekun Ye. Gaussian weighting reversion strategy for accurate online portfolio selection. *IEEE Transactions on Signal Processing*, 67(21):5558–5570, 2019.
- [Cover and Thomas, 1990] Thomas M Cover and Joy A Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, 1990.
- [Cover, 1991] Thomas M Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.
- [Crammer *et al.*, 2006] Koby Crammer, Ofer Dekel, Joseph Keshet, Shai Shalev-Shwartz, and Yoram Singer. Online passive-aggressive algorithms. *Journal of Machine Learning Research*, 7:551–585, 2006.
- [Das *et al.*, 2014] Puja Das, Nicholas Johnson, and Arindam Banerjee. Online portfolio selection with group sparsity. In *28th AAAI Conference on Artificial Intelligence*, 2014.
- [Duchi *et al.*, 2008] John Duchi, Shai Shalev-Shwartz, Yoram Singer, and Tushar Chandra. Efficient projections onto the l_1 -ball for learning in high dimensions. In *25th International Conference on Machine Learning*, pages 272–279, 2008.
- [Gaivoronski and Stella, 2000] Alexei A Gaivoronski and Fabio Stella. Stochastic nonstationary optimization for finding universal portfolios. *Annals of Operations Research*, 100(1):165–188, 2000.
- [Helmbold *et al.*, 1998] 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.
- [Huang *et al.*, 2011] Junzhou Huang, Tong Zhang, and Dimitris Metaxas. Learning with structured sparsity. *Journal of Machine Learning Research*, 12(11):3371–3412, 2011.
- [Huang *et al.*, 2015] Dingjiang Huang, Yan Zhu, Bin Li, Shuigeng Zhou, and Steven C Hoi. Semi-universal portfolios with transaction costs. In *24th International Conference on Artificial Intelligence*, pages 178–184, 2015.
- [Huang *et al.*, 2016] Dingjiang Huang, Junlong Zhou, Bin Li, Steven C Hoi, and Shuigeng Zhou. Robust median reversion strategy for online portfolio selection. *IEEE Transactions on Knowledge and Data Engineering*, 28(9):2480–2493, 2016.
- [Kelly, 1956] Jr Kelly. A new interpretation of information rate. *Bell System Technical Journal*, 35(4):917–926, 1956.
- [Li and Hoi, 2012] Bin Li and Steven C Hoi. On-line portfolio selection with moving average reversion. In *29th ICML*, pages 563–570, 2012.
- [Li and Hoi, 2014] Bin Li and Steven C Hoi. Online portfolio selection: A survey. *ACM Computing Surveys*, 46(3):35:1–35:36, 2014.
- [Li *et al.*, 2011] Bin Li, Steven C. H. Hoi, and Vivekanand Gopalkrishnan. Corn: Correlation-driven nonparametric learning approach for portfolio selection. *ACM Transactions on Intelligent Systems and Technology*, 2(3):21:1–21:29, 2011.
- [Li *et al.*, 2012] Bin Li, Peilin Zhao, Steven C Hoi, and Vivekanand Gopalkrishnan. Passive aggressive mean reversion strategy for portfolio selection. *Machine Learning*, 87(2):221–258, 2012.
- [Li *et al.*, 2013] Bin Li, Steven C Hoi, Peilin Zhao, and Vivekanand Gopalkrishnan. Confidence weighted mean reversion strategy for online portfolio selection. *ACM Transactions on Knowledge Discovery from Data*, 7(1):4:1–4:38, 2013.
- [Li *et al.*, 2018] Bin Li, Jialei Wang, Dingjiang Huang, and Steven C Hoi. Transaction cost optimization for online portfolio selection. *Quantitative Finance*, 18(8):1411–1424, 2018.
- [Liang *et al.*, 2018] Zhipeng Liang, Kangkang Jiang, Hao Chen, Junhao Zhu, and Yanran Li. Adversarial deep reinforcement learning in portfolio management. *arXiv preprint arXiv:1808.09940v3*, 2018.
- [Markowitz, 1952] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.
- [Ordentlich and Cover, 1996] Erik Ordentlich and Thomas Cover. On-line portfolio selection. In *9th Annual Conference on Computational Learning Theory*, 1996.
- [Selvin *et al.*, 2017] Sreelekshmy Selvin, R Vinayakumar, EA Gopalakrishnan, Vijay Krishna Menon, and KP Soman. Stock price prediction using lstm, rnn and cnn-sliding window model. In *2017 International conference on advances in computing, communications and informatics*, pages 1643–1647, 2017.
- [Shen and Wang, 2017] Weiwei Shen and Jun Wang. Portfolio selection via subset resampling. In *31st AAAI*, 2017.
- [Shen *et al.*, 2019] Weiwei Shen, Bin Wang, Jian Pu, and Jun Wang. The kelly growth optimal portfolio with ensemble learning. In *AAAI*, volume 33, pages 1134–1141, 2019.
- [Yang *et al.*, 2018] Haimin Yang, Zhisong Pan, Qing Tao, and Junyang Qiu. Online learning for vector autoregressive moving-average time series prediction. *Neurocomputing*, 315:9–17, 2018.
- [Ye *et al.*, 2017] Zekun Ye, Kai Huang, Shuigeng Zhou, and Jihong Guan. Gaussian weighting reversion strategy for accurate on-line portfolio selection. In *IEEE 29th ICTAL*, pages 929–936, 2017.
- [Zinkevich, 2003] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *20th ICML*, pages 928–936, 2003.