4-2011

Confidence Weighted Mean Reversion Strategy for On-Line Portfolio Selection

Bin LI
Nanyang Technological University

Steven C. H. HOI
Singapore Management University, chhoi@smu.edu.sg

Peilin ZHAO
Nanyang Technological University

Vivek Gopalkrishnan
Nanyang Technological University

Citation

LI, Bin; HOI, Steven C. H.; ZHAO, Peilin; and Gopalkrishnan, Vivek. Confidence Weighted Mean Reversion Strategy for On-Line Portfolio Selection. (2011). *JMLR Workshop and Conference Proceedings: 14th International Conference on Artificial Intelligence and Statistics (AISTATS) 2011, Fort Lauderdale, FL, April 11-13*. 15, 434-442.

Available at: https://ink.library.smu.edu.sg/sis_research/2292

Confidence Weighted Mean Reversion Strategy for On-Line Portfolio Selection

Bin Li, Steven C.H. Hoi, Peilin Zhao, Vivekanand Gopalkrishnan
School of Computer Engineering, Nanyang Technological University, Singapore
{s080061, chhoi, zhao0106, asvivek}@ntu.edu.sg

Abstract

This paper proposes a novel on-line portfolio selection strategy named "Confidence Weighted Mean Reversion" (CWMR). Inspired by the mean reversion principle and the confidence weighted on-line learning technique, CWMR models a portfolio vector as Gaussian distribution, and sequentially updates the distribution by following the mean reversion trading principle. The CWMR strategy is able to effectively exploit the power of mean reversion for on-line portfolio selection. Extensive experiments on various real markets demonstrate the effectiveness of our strategy in comparison with the state of the art.

1 Introduction

On-line Portfolio Selection (PS), also termed "sequential portfolio selection", aims to determine a practical strategy for investing wealth among a set of assets to achieve some financial objectives in the long run. The finance community has studied the problem by mainly concerning the objective of maximizing risk-adjusted returns [12, 26, 28, 29]. On the other hand, the *learning to trade* techniques, often aiming to maximize the logarithmic compound return or growth rate, have also been actively explored in the information theory [6, 7, 21, 22, 27, 31] and the machine learning community [1, 3, 4, 15–19, 23, 24, 30, 33].

Some state-of-the-art PS strategies [15, 16] assume that the current best performing stocks would also perform well the next trading day, but empirical evidence [20] indicates that such trends may be often violated, especially in the short term. This observation leads to the strategy of buying poor performing stocks and selling those with good performance. This trading principle, known as "mean reversion", is followed by some methods, including Constant Rebalanced Portfolios (CRP) [7] and Anticor [4], among others.

Appearing in Proceedings of the 14^{th} International Conference on Artificial Intelligence and Statistics (AISTATS) 2011, Fort Lauderdale, FL, USA. Volume 15 of JMLR: W&CP 15.

However, best CRP [7], which is theoretically grounded and passively reverses to the mean, performs significantly poorly in comparison with Anticor, which is heuristic and actively reverses to the mean via statistical correlation. This calls for the need of integrating a powerful learning method to actively exploit the mean reversion property. Besides, all existing learning to trade algorithms (c.f., Section 3 for a review) only exploit the first order information of the portfolio, while the change in the distribution of the portfolio is better reflected in its first order and second order information (mean and volatility).

To address these two drawbacks, we present a new on-line portfolio selection strategy named "Confidence Weighted Mean Reversion" (CWMR). In short, CWMR models the portfolio vector as a Gaussian distribution and sequentially updates the distribution according to the mean reversion trading idea. Thus, CWMR exploits the mean reversion property in the financial markets and both first and second order information of the portfolio vector by the powerful Confidence Weighted (CW) on-line learning [10, 11].

The salient features of the proposed CWMR strategy are:

1. It is the first learning to trade study that exploits the second order information of the *portfolio* (not the second order information of *price*);
2. Our novel algorithms effectively exploit the mean reversion property of the financial market by applying the powerful confidence weighted learning technique.

Through extensive numerical experiments on a variety of up-to-date real testbeds, we show that the proposed CWMR algorithms significantly surpass a number of state-of-the-art strategies in terms of long-term compound return. The experiments also show that CWMR is robust with respect to different settings of the parameters and it can withstand small transaction costs.

The rest of the paper is organized as follows. Section 2 formally defines the problem of on-line portfolio selection. Section 3 reviews related work and highlights their limitations. Section 4 presents our proposed CWMR algorithms, and Section 5 compares these approaches on historical stock markets. Finally, we conclude in Section 6 with directions for future work.

2 Problem Setting

Consider a financial market with m assets to be invested. The changes of asset prices for n trading periods are represented by a sequence of positive *price relative vectors* $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}_+^m$. We use $\mathbf{x}^n$ to denote such a sequence of vectors. The j^{th} component of the i^{th} vector x_{ij} denotes the ratio of the closing price to the last closing price of the j^{th} asset on the i^{th} trading period, thus an investment in asset j on the i^{th} period increases by a factor of x_{ij}.

An investment on the market is specified by a *portfolio vector*, denoted as $\mathbf{b} = (b_1, \ldots, b_m)$, where b_i represents the proportion of wealth invested on the i^{th} asset. Typically, we assume the portfolio is self-financed and no margin/shorting is allowed, which means $\mathbf{b} \in \triangle_m$, where $\triangle_m = \{\mathbf{b} : \mathbf{b} \in \mathbb{R}_+^m, \sum_{i=1}^m b_i = 1\}$. The investment procedure is represented by the *portfolio strategy*, i.e., a sequence of mappings $\mathbf{b}_i : \mathbb{R}_+^{m(i-1)} \rightarrow \triangle_m, i = 1, 2, \ldots$, where $\mathbf{b}_i = \mathbf{b}_i(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1})$ is the portfolio used on the i^{th} trading period given past market price relatives $\mathbf{x}^{i-1} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}\}$. Let us denote by $\mathbf{b}^n$ the portfolio strategy for the n consecutive trading periods.

For the i^{th} trading period, an investment defined by portfolio $\mathbf{b}_i$ produces a *portfolio period return* $\mathbf{s}_i$, i.e., the wealth increases by a factor of $\mathbf{s}_i = \mathbf{b}_i^\top \mathbf{x}_i = \sum_{j=1}^m b_{ij} x_{ij}$. Since we re-invest all the wealth, the investment results in multiplicative cumulative return. Thus, after *n* trading periods, the investment of a portfolio strategy $\mathbf{b}^n$ produces a *portfolio cumulative wealth* $\mathbf{S}_n$, which is increased by a factor of $\prod_{i=1}^n \mathbf{b}_i^\top \mathbf{x}_i$, i.e., $\mathbf{S}_n(\mathbf{b}^n, \mathbf{x}^n) = \mathbf{S}_0 \prod_{i=1}^n \mathbf{b}_i^\top \mathbf{x}_i$, where $\mathbf{S}_0$ denotes the initial wealth, which is set to $\$1$ in this paper.

Finally, we formulate the on-line PS problem as a sequential decision task. The portfolio manager aims to design a strategy to maximize the portfolio cumulative wealth. The portfolios are selected in a sequential fashion. On each trading period i, given the historical information, including all the previous sequences of price relative vectors $\mathbf{x}^{i-1} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}\}$, and the previous sequences of portfolio vectors $\mathbf{b}^{i-1} = \{\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}\}$, the manager learns to decide a new portfolio vector $\mathbf{b}_i$ for the coming price relative vector $\mathbf{x}_i$. The resulting portfolio is scored based on the portfolio period return. This procedure repeats until the end of the trading period. The portfolio strategy is scored according to the cumulative wealth achieved finally.

In the above model, we make several general assumptions:

1. Transaction cost: we assume no transaction costs (commissions, taxes, and slippage, etc.) in the model;
2. Market liquidity: we assume each asset is arbitrarily divisible, and one can buy and sell required quantities at the last closing price of any given trading period;
3. Impact cost: we assume the market behavior is not affected by the trading strategy in our evaluation.

The implications and effects of these assumptions are discussed in Section 5.3 and Section 5.5.

3 Related Work

Some common and well-known benchmarks for PS include the *Buy-And-Hold* (BAH) strategy and the *Constant Rebalanced Portfolios* (CRP) [7, 8]. In our study, we refer to the equal-weighted BAH strategy as the *market* strategy. Contrary to the static BAH strategy, CRP actively adjusts the portfolio by keeping a fixed fraction of the investor's total wealth on each asset. The best possible CRP strategy, known as *Best CRP* (BCRP), is a hindsight strategy.

One group of learning to trade research aims to approach the same daily wealth growth rate as the BCRP strategy. Cover [7] proposed *Universal Portfolio* (UP) strategy, which is based on the weighted average of the historical performance of all CRP experts. Helmbold et al. [19] proposed the *Exponential Gradient* (EG) strategy to maximize the expected logarithmic daily return. Agarwal et al. [1] proposed the *Online Newton Step* (ONS) strategy to maximize the expected logarithmic cumulative wealth. Hazan and Seshadhri [18] proposed an adaptive ONS approach.

Another promising research direction for new on-line PS strategies tries to approach the Oracle strategy. Such idea was adopted in Borodin et al. [4] where they proposed a non-universal portfolio strategy named *Anticor* to exploit statistical information from historical markets and to rebalance the portfolio according to the mean reversion trading idea. Györfi et al. [15] recently introduced a framework of *Nonparametric Kernel-based Moving Window* (B^{K}) strategy attempting to construct portfolios based on similar historical price relatives measured via Euclidean distance. Following the same framework, *Nonparametric Nearest Neighbor* (B^{NN}) [16] strategy locates the similar price relatives via nearest neighbor. Li et al. [24] further proposed *Correlation-driven Nonparametric learning* (CORN) strategy by locating the similar price relatives via correlation.

Aggregating algorithms [32] have also been used for Online PS. Singer [30] proposed *Switching Portfolio* (SP), which switches among the underlying strategies according to a prior distribution. Levina and Shafer [23] introduced *Gaussian Random Walk* (GRW), which applies the aggregating strategy and switches according to Gaussian distribution. Sequential prediction techniques can also be applied for tackle this task, for example, *Add-beta* [3] prediction strategy (*T0* & *M0* algorithm).

3.1 Limitation of existing work

Most existing learning to trade strategies (UP, EG, ONS, B^{K}, and B^{NN}) often adopt the trend following trading idea by assuming that the price relative for the next trading day follows the same trend as today's price relative, i.e., the winning stocks over others tend to win the following trading day. However, in the short-term, the stock price relatives may not follow the previous trends as empirically evidenced by Jegadeesh [20].

Besides the trend following idea, another trading principle, i.e., "mean reversion", assumes that if a stock performs worse than others, it tends to perform better in the next trading day. Thus, a mean reversion strategy tends to purchase the securities of poor performance and to sell the securities of good performance in the past. Some strategies that adopt this idea include CRP [7] and Anticor [4]. Empirically, the CRP strategy that passively reverses to the mean often performs worse than Anticor, which actively reverses to the mean and thus can better exploit the fluctuation of assets [4]. On the other hand, because Anticor heuristically transfers the proportion within the portfolio based on statistical correlations, it often produces sub-optimal results. A new strategy to exploit the mean reversion property with a powerful learning method is necessary.

Finally, all existing algorithms only consider the first order information of the portfolio vectors, while the second order information (volatility of the portfolio vector) could provide useful volatility information for the PS task.

4 Confidence Weighted Mean Reversion for On-Line Portfolio Selection

4.1 Motivation and Overview

Our proposed method is motivated by the best CRP strategy, which theoretically has a nice performance guarantee [7], and the Anticor strategy, which has a good empirical performance [4], with their underlying mean reversion trading idea. In the context of portfolio, or multiple stocks, it implies that better performing stocks tend to perform worse than others in the subsequent trading days, and the worse performing stocks are inclined to perform better. Thus if we want to maximize the portfolio return for the next trading day, we could minimize the expected portfolio return with respect to today's price relative since the next price relative tends to revert. This is a bit contra-intuitive, but according to Lo and MacKinlay [25], the effectiveness of mean reversion is due to the positive cross-autocovariances across securities.

The proposed method is also inspired by Confidence Weighted (CW) learning [10, 11], which was originally proposed for classification. The basic idea of CW is to maintain a Gaussian distribution for the classifier, and sequentially update the classifier distribution according to the Passive Aggressive (PA) learning [9]. CW takes advantage of both first and second order information of the solution.

To address the limitations described in Section 3.1, in this paper, we present a novel on-line PS method named "Confidence Weighted Mean Reversion", or CWMR for short. We model the portfolio vector as a Gaussian distribution and sequentially update the distribution according to the mean reversion trading idea. Different from CRP and Anticor, CWMR actively exploits the mean reversion property of the financial market with a powerful learning method. Traditional learning to trade algorithms, to the best of our knowledge, all focus on the first order information of portfolio vector, while the proposed CWMR algorithm considers both first and second order information where the additional second order information could benefit the PS task.

4.2 Formulation

Let us model the portfolio vector for the i^{th} trading day as a Gaussian distribution with mean $\mu\in\mathbb{R}^m$ and the diagonal covariance matrix $\mathbf{\Sigma}\in\mathbb{R}^{m\times m}$ with nonzero diagonal elements σ^2 and zero for off-diagonal elements. The value μ_j represents the knowledge of asset j in the portfolio. The diagonal covariance matrix term Σ_{jj} or σ_j^2 stands for the confidence we have in the portfolio mean value μ_j.

At the beginning of i^{th} trading day, we construct a portfolio $\mathbf{b}_i$ based on the distribution $\mathcal{N}(\mu, \mathbf{\Sigma})$, $\mathbf{b}_i\sim\mathcal{N}(\mu, \mathbf{\Sigma})$. Then after the price relative $\mathbf{x}_i$ is revealed, the portfolio increases its wealth by a factor of $\mathbf{b}_i\cdot\mathbf{x}_i$. It is straightforward that the portfolio daily return can be viewed as a random variable of a univariate Gaussian distribution, $\mathrm{D}\sim\mathcal{N}\left(\mu\cdot\mathbf{x}_i, \mathbf{x}_i^\top\mathbf{\Sigma}\mathbf{x}_i\right)$. The mean of portfolio daily return is the return of the mean portfolio vector and the variance is proportional to the length of the projection of $\mathbf{x}_i$ on $\mathbf{\Sigma}$.

Now let us update the distribution. According to the mean reversion trading idea, the probability of a profitable portfolio for the next trading day $\mathbf{b}$ with respect to a mean reversion threshold ϵ is defined as,

$$\mathrm{Pr}_{\mathbf{b}\sim\mathcal{N}(\mu,\mathbf{\Sigma})}\left[\mathrm{D}\leq\epsilon\right] = \mathrm{Pr}_{\mathbf{b}\sim\mathcal{N}(\mu,\mathbf{\Sigma})}\left[\mathbf{b}\cdot\mathbf{x}_i\leq\epsilon\right].$$

For simplicity, we write $\mathrm{Pr}\left[\mathbf{b}\cdot\mathbf{x}_i\leq\epsilon\right]$ instead. The manager adjusts the distribution to ensure the probability of a profitable portfolio is higher than a confidence level $\theta\in[0, 1]$,

$$\mathrm{Pr}\left[\mathbf{b}\cdot\mathbf{x}_i\leq\epsilon\right]\geq\theta.$$

If the expected return using the i^{th} price relative is less than a threshold with high probability, the actual return for the $i{+}1^{th}$ trading day tends to be high with correspondingly high probability, since the price relative tends to reverse. Then, following the intuition underlying PA algorithms [9], our algorithm chooses the distribution closest (in the KL divergence sense) to the current distribution $\mathcal{N}(\mu_i, \mathbf{\Sigma}_i)$. As a result, on the $i{+}1^{th}$ trading day, the algorithm sets the parameters of the distribution by solving the following optimization problem:

Original Optimization Problem of CWMR:

$$\begin{aligned}(\mu_{i+1}, \mathbf{\Sigma}_{i+1}) = \arg\min_{\mu,\mathbf{\Sigma}} \quad & \mathrm{D}_{\mathrm{KL}}\left(\mathcal{N}(\mu, \mathbf{\Sigma})\,\|\mathcal{N}(\mu_i, \mathbf{\Sigma}_i)\right) \\ \text{s.t.} \quad & \mathrm{Pr}\left[\mu\cdot\mathbf{x}_i\leq\epsilon\right]\geq\theta \\ & \mu\in\triangle_m.\end{aligned} \tag{1}$$

Under the distribution of $\mathcal{N}(\mu, \mathbf{\Sigma})$, the return for the i^{th} trading day has a Gaussian distribution with mean $\mu_D{=}\mu\cdot\mathbf{x}_i$ and covariance $\Sigma_D{=}\mathbf{x}_i^\top\mathbf{\Sigma}\mathbf{x}_i$ of diagonal elements σ_D^2. Thus, the probability of a profitable portfolio,

$\Pr[D \le \epsilon] = \Pr\left[\frac{D-\mu_D}{\sigma_D} \le \frac{\epsilon-\mu_D}{\sigma_D}\right]$. In this formula, $\frac{D-\mu_D}{\sigma_D}$ is a normally distributed random variable, the above probability equals $\Phi\left(\frac{\epsilon-\mu_D}{\sigma_D}\right)$, where Φ is the cumulative distribution function of the Gaussian distribution. As a result, we can rewrite the constraint as, $\frac{\epsilon-\mu_D}{\sigma_D} \ge \Phi^{-1}(\theta)$. Substituting μ_D and σ_D by their definitions and rearranging terms, we obtain the constraint, $\epsilon - \mu \cdot \mathbf{x}_i \ge \phi\sqrt{\mathbf{x}_i^\top \boldsymbol{\Sigma}\mathbf{x}_i}$, where $\phi = \Phi^{-1}(\theta)$.

To make our research more realistic and consistent with previous studies, we replace the portfolio return term $\mu \cdot \mathbf{x}_i$ by its logarithmic $\log(\mu \cdot \mathbf{x}_i)$ in order to reflect the risk aversion preference of the investors. Moreover, using logarithmic utility function and holding other variables constant, imply loosening the constraint. However, since both ϵ and ϕ are adjustable, by choosing appropriate values, we can weaken the loose effect. Thus, in our formulation we modify the constraint using the logarithmic return function as, $\epsilon - \log(\mu \cdot \mathbf{x}_i) \ge \phi\sqrt{\mathbf{x}_i^\top \boldsymbol{\Sigma}\mathbf{x}_i}$.

To this end, we rewrite the above optimization problem as:

Revised Optimization Problem of CWMR:

$$
\begin{aligned}
(\mu_{i+1}, \boldsymbol{\Sigma}_{i+1}) = \arg\min_{\mu, \boldsymbol{\Sigma}} \quad & \frac{1}{2}\left(\log\left(\frac{\det\boldsymbol{\Sigma}_i}{\det\boldsymbol{\Sigma}}\right) + \mathrm{Tr}\left(\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Sigma}\right)\right) \\
& + \frac{1}{2}\left((\mu_i - \mu)^\top \boldsymbol{\Sigma}_i^{-1}(\mu_i - \mu)\right) \\
\text{s.t.} \quad & \epsilon - \log(\mu \cdot \mathbf{x}_i) \ge \phi\sqrt{\mathbf{x}_i^\top \boldsymbol{\Sigma}\mathbf{x}_i} \\
& \mu \cdot \mathbf{1} = 1, \ \mu \succeq 0.
\end{aligned}
$$

For the above revised optimization problem, the constraint is not convex in $\boldsymbol{\Sigma}$. We suggest two ways to handle it. The first way is to linearize it by omitting the square root [11], i.e., $\epsilon - \log(\mu \cdot \mathbf{x}_i) \ge \phi \mathbf{x}_i^\top \boldsymbol{\Sigma}\mathbf{x}_i$. As a result, we have the first final optimization problem named **CWMR-Var**, whose solution is an approximate solution to the original optimization problem (1).

Final Optimization Problem 1 (CWMR-Var):

$$
\begin{aligned}
(\mu_{i+1}, \boldsymbol{\Sigma}_{i+1}) = \arg\min_{\mu, \boldsymbol{\Sigma}} \quad & \frac{1}{2}\left(\log\left(\frac{\det\boldsymbol{\Sigma}_i}{\det\boldsymbol{\Sigma}}\right) + \mathrm{Tr}\left(\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Sigma}\right)\right) \\
& + \frac{1}{2}\left((\mu_i - \mu)^\top \boldsymbol{\Sigma}_i^{-1}(\mu_i - \mu)\right) \\
\text{s.t.} \quad & \epsilon - \log(\mu \cdot \mathbf{x}_i) \ge \phi \mathbf{x}_i^\top \boldsymbol{\Sigma}\mathbf{x}_i \\
& \mu \cdot \mathbf{1} = 1, \ \mu \succeq 0.
\end{aligned} \tag{2}
$$

Following Crammer et al. [10], the second reformulation is to decompose $\boldsymbol{\Sigma}$ since it is positive semidefinite (PSD), i.e., $\boldsymbol{\Sigma} = \boldsymbol{\Upsilon}^2$ with $\boldsymbol{\Upsilon} = \mathrm{Q}\mathrm{diag}\left(\lambda_1^{1/2}, \ldots, \lambda_m^{1/2}\right)\mathrm{Q}^\top$, where Q is orthonormal and $\lambda_1, \ldots, \lambda_m$ are the eigenvalues of $\boldsymbol{\Sigma}$ and thus $\boldsymbol{\Upsilon}$ is also PSD. This reformulation yields the second final optimization problem named **CWMR-Stdev**, whose solution is the exact solution of the original problem (1).

Final Optimization Problem 2 (CWMR-Stdev):

$$
\begin{aligned}
(\mu_{i+1}, \boldsymbol{\Upsilon}_{i+1}) = \arg\min_{\mu, \boldsymbol{\Upsilon}} \quad & \frac{1}{2}\left(\log\left(\frac{\det\boldsymbol{\Upsilon}_i^2}{\det\boldsymbol{\Upsilon}^2}\right) + \mathrm{Tr}\left(\boldsymbol{\Upsilon}_i^{-2}\boldsymbol{\Upsilon}^2\right)\right) \\
& + \frac{1}{2}\left((\mu_i - \mu)^\top \boldsymbol{\Upsilon}_i^{-2}(\mu_i - \mu)\right) \\
\text{s.t.} \quad & \epsilon - \log(\mu \cdot \mathbf{x}_i) \ge \phi\left\|\boldsymbol{\Upsilon}\mathbf{x}_i\right\|, \boldsymbol{\Upsilon} \text{ is PSD} \\
& \mu \cdot \mathbf{1} = 1, \ \mu \succeq 0.
\end{aligned} \tag{3}
$$

4.3 Algorithms

Now let us develop the proposed algorithms based on their solutions using the typical techniques from convex optimization [5]. The solutions to the optimizations are shown in Proposition 1 & Proposition 2, with their corresponding proofs in Appendix A & B, respectively.

Proposition 1. *The solution to the final optimization problem* (2) ***(CWMR-Var)*** *is expressed as:*

$$\mu_{i+1} = \mu_i - \lambda_{i+1}\boldsymbol{\Sigma}_i\left(\frac{\mathbf{x}_i - \bar{\mathbf{x}}_i\mathbf{1}}{\mu_i \cdot \mathbf{x}_i}\right), \ \boldsymbol{\Sigma}_{i+1}^{-1} = \boldsymbol{\Sigma}_i^{-1} + 2\lambda_{i+1}\phi\mathbf{x}_i\mathbf{x}_i^\top,$$

where λ_{i+1} corresponds to the Lagrangian multiplier calculated by Eq. (5) *and $\bar{\mathbf{x}}_i = \frac{\mathbf{1}^\top\boldsymbol{\Sigma}_i\mathbf{x}_i}{\mathbf{1}^\top\boldsymbol{\Sigma}_i\mathbf{1}}$ denotes the confidence weighted price relative average.*

Proposition 2. *The solution to the final optimization problem* (3) ***(CWMR-Stdev)*** *is expressed as:*

$$\mu_{i+1} = \mu_i - \lambda_{i+1}\boldsymbol{\Sigma}_i\frac{\mathbf{x}_i - \bar{\mathbf{x}}_i\mathbf{1}}{\mu_i \cdot \mathbf{x}_i}, \ \boldsymbol{\Sigma}_{i+1}^{-1} = \boldsymbol{\Sigma}_i^{-1} + \lambda_{i+1}\phi\frac{\mathbf{x}_i\mathbf{x}_i^\top}{\sqrt{U_i}},$$

where $V_i = \mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{x}_i$ and $\sqrt{U_i} = \frac{-\lambda_{i+1}V_i\phi + \sqrt{\lambda_{i+1}^2V_i^2\phi^2 + 4V_i}}{2}$ denote the variance of the portfolio return for the i^{th} and $i{+}1^{th}$ trading day, and λ_{i+1} denotes the Lagrangian multiplier calculated by Eq. (7)*, and $\bar{\mathbf{x}}_i = \frac{\mathbf{1}^\top\boldsymbol{\Sigma}_i\mathbf{x}_i}{\mathbf{1}^\top\boldsymbol{\Sigma}_i\mathbf{1}}$ represents the confidence weighted average of the i^{th} price relative.*

Initially, with no information available for the on-line PS task, we simply initialize the portfolio mean μ_1 to uniform portfolio and the portfolio covariance matrix $\boldsymbol{\Sigma}_1$ to equally standard deviation $\frac{1}{m}$, or equivalent variance $\frac{1}{m^2}$. One remaining issue is that the resulting μ can be negative since we do not consider the non-negativity constraint in the solution. To solve this issue we simply project the resulting μ to the simplex domain to ensure the simplex constraint. Another remaining issue is that although the covariance matrix is non-singular in theory, in real computation, the covariance matrix $\boldsymbol{\Sigma}$ sometimes may be singular due to the computer precision. To avoid this problem and be consistent with the projection of the μ, we try to rescale $\boldsymbol{\Sigma}$ by normalizing its maximum value to $\frac{1}{m^2}$. The final CWMR algorithm is presented in Figure 1.

4.4 Discussion

The CWMR algorithm is motivated by the Confidence Weight learning (CW) [10, 11], thus its formulation and subsequent derivations are similar. However, they address

Algorithm 1 The CWMR Algorithms for On-Line PS

INPUT: $\phi=\Phi^{-1}(\theta)$: Confidence parameter; $\epsilon<0$: Mean reversion parameter

INITIALIZE: $\mu_1 = \frac{1}{m}\mathbf{1},\ \mathbf{\Sigma}_1 = \frac{1}{m^2}\mathbf{I},\ \mathbf{S}_0 = 1$

For $t = 1, 2, \ldots$

1: Draw a portfolio $\mathbf{b}_t$ from $\mathcal{N}(\mu_t, \mathbf{\Sigma}_t)$
2: Receive stock price relatives: $\mathbf{x}_t = (x_{t1}, \ldots, x_{tm})$
3: Calculate the daily return and cumulative return: $\mathbf{S}_t = \mathbf{S}_{t-1} \times (\mathbf{b}_t \cdot \mathbf{x}_t)$
4: Calculate the following variables:

$$M_t = \mu_t \cdot \mathbf{x}_t, \quad V_t = \mathbf{x}_t^\top \mathbf{\Sigma}_t \mathbf{x}_t, \quad \bar{\mathbf{x}}_t = \frac{\mathbf{1}^\top \mathbf{\Sigma}_t \mathbf{x}_t}{\mathbf{1}^\top \mathbf{\Sigma}_t \mathbf{1}}$$

5: Update the portfolio distribution:

$$\text{CWMR-Var} \begin{cases} \lambda_{t+1} \text{ as in Eq. (5)} \\ \mu_{t+1} = \mu_t - \lambda_{t+1}\mathbf{\Sigma}_t \frac{\mathbf{x}_t - \bar{\mathbf{x}}_t \mathbf{1}}{M_t} \\ \mathbf{\Sigma}_{t+1} = \left(\mathbf{\Sigma}_t^{-1} + 2\lambda_{t+1}\phi \text{diag}^2(\mathbf{x}_t)\right)^{-1} \end{cases}$$

$$\text{CWMR-Stdev} \begin{cases} \lambda_{t+1} \text{ as in Eq. (7)} \\ \sqrt{U}_t = \frac{-\lambda_{t+1}\phi V_t + \sqrt{\lambda_{t+1}^2 \phi^2 V_t^2 + 4V_t}}{2} \\ \mu_{t+1} = \mu_t - \lambda_{t+1}\mathbf{\Sigma}_t \frac{\mathbf{x}_t - \bar{\mathbf{x}}_t \mathbf{1}}{M_t} \\ \mathbf{\Sigma}_{t+1} = \left(\mathbf{\Sigma}_t^{-1} + \phi \frac{\lambda_{t+1}}{\sqrt{U_t}} \text{diag}^2(\mathbf{x}_t)\right)^{-1} \end{cases}$$

6: Normalize μ_{t+1} and $\mathbf{\Sigma}_{t+1}$:

$$\mu_{t+1} = \arg\min_{\mu \in \triangle_m} \|\mu - \mu_{t+1}\|^2, \ \mathbf{\Sigma}_{t+1} = \frac{\mathbf{\Sigma}_{t+1}}{m^2 \text{Tr}(\mathbf{\Sigma}_{t+1})}$$

Figure 1: The proposed Confidence Weighted Mean Reversion (CWMR) algorithms.

different problems since CWMR aims to handle on-line portfolio selection while CW focuses on classification. Although both objectives adopt KL divergence to measure the closeness between two distributions, their constraints reflect that they are different problems oriented. To be specific, CW's constraint is the probability of a correct prediction, while CWMR's constraint is the probability of a mean reversion profitable portfolio plus a simplex constraint. The formulations' differences result in different derivations.

Since portfolio mean is our main concern for the on-line PS problem, in this section, we mainly provide a preliminary analysis of update schemes of portfolio mean μ to reflect its underlying mean reversion trading idea. Both CWMR-Var and CWMR-Stdev have the same update equation for the portfolio mean, i.e., $\mu_{t+1}=\mu_t-\lambda_{t+1}\mathbf{\Sigma}_t\frac{\mathbf{x}_t-\bar{\mathbf{x}}_t\mathbf{1}}{\mu_t\cdot\mathbf{x}_t}$. It is obvious that λ_{t+1} is non-negative and $\mathbf{\Sigma}_t$ is PSD. The denominator term $\mathbf{x}_t-\bar{\mathbf{x}}_t\mathbf{1}$ can be viewed as excess return vector of asset pool for previous trading day, where $\bar{\mathbf{x}}_t$ represents the confidence weighted mean return. Holding other terms constant, the portfolio mean tends to move towards previous one while the magnitude is negatively related to the previous excess return, which is in effect the mean reversion trading idea. At the same time, the negative movements are dynamically adjusted by optimal λ_{t+1}, previous portfolio confidence $\mathbf{\Sigma}_t$ and mean return $\mu_t\cdot\mathbf{x}_t$, which catch both first and second order information. To the best of our knowledge, none of previous learning to trade algorithms has explicitly exploit the second order information of portfolio vector, however, the second order information may contribute to the success of the proposed algorithms.

5 Numerical Experiments

We now examine the efficacy of our proposed approach by performing extensive experiments on publicly available, real and diverse data from stock markets.

Details of the experimental datasets[1] are summarized in Table 1. Two of these datasets have been used by in previous work (NYSE (O) [1, 4, 7, 15, 16, 19] and TSE [4]), while the rest datasets are collected by us.

Dataset	Region	Time frame	# days	# Assets
NYSE (O)	US	July 3$^{\text{rd}}$ 1962 - Dec 31$^{\text{st}}$ 1984	5651	36
NYSE (N)	US	Jan 1$^{\text{st}}$ 1985 - Jun 30$^{\text{th}}$ 2010	6431	23
TSE	CA	Jan 4$^{\text{th}}$ 1994 - Dec 31$^{\text{st}}$ 1998	1259	88
STI	SG	Jan 1$^{\text{th}}$ 2005 - Jun 30$^{\text{th}}$ 2010	1404	22
MSCI	Global	Oct 17$^{\text{th}}$ 2005 - Oct 15$^{\text{th}}$ 2010	1304	4

Table 1: Summary of 5 real datasets from various markets.

In this paper, we measure investment performance using the most common metric, *cumulative wealth*. Other detailed experiments, including those on risk adjusted return, are presented in the long version.

For our approach, we provide deterministic and stochastic versions. For the former (CWMR-Var and CWMR-Stdev), we eliminate the randomness of the portfolio (in reality, no investors like random portfolio) and stabilize experiment performance, by deterministically drawing a portfolio from portfolio Gaussian distribution, i.e., directly set the portfolio $\mathbf{b}=\mu$. For the latter (CWMR-Var-s and CWMR-Stdev-s), we repeat it for 50 times and provide their average value. Since the stochastic portfolio may be negative, the projection to the simplex domain becomes necessary. We set the parameters empirically without tuning, i.e., confidence parameter $\phi=2$ or equivalently confidence level $\theta=95\%$, and mean reversion parameter $\epsilon=-0.5$. Section 5.2 will examine the parameter sensitivity.

We compare the proposed strategy with several existing strategies (c.f., Section 3), whose parameters were set according to the suggestions from their respective studies.

5.1 Cumulative Wealth

The first experiment evaluates the cumulative wealth at the end of the trading period. From the results illustrated in Table 2(a), we find that CWMR (both deterministic CWMR-Var/CWMR-Stdev and stochastic CWMR-Var-s/CWMR-Stdev-s) significantly outperform all competitors, including Anticor and B^{NN}, which are the state-of-the-art techniques. As widely done in the fund management industry [14], we

[1] All datasets and their compositions can be downloaded from `http://www.cais.ntu.edu.sg/~libin/portfolios`.

also performed statistical tests to examine if the claimed excess return can be generated by simple luck. As the results (Table 2(b)) show, the possibility for achieving the excess return due to simple luck is only 0.01% on the TSE dataset and almost 0% on other datasets. Finally, we also plot the wealth curve of the cumulative wealth in Figure 3. The figures show that the proposed CWMR algorithm performs consistently over the entire trading period. These results show that the CWMR approach is a promising and reliable PS technique to achieve high return with high confidence.

5.2 Parameter Sensitivity

We now evaluate how different choices of the parameters affect the performance of CWMR. Since confidence parameter ϕ generally does not have a decisive influence on the final performance, we evaluate the scalability of the proposed approach with respect to the negative mean reversion sensitivity $-\epsilon$. Figure 4 depicts the results, plus the final cumulative wealth achieved by Market and BCRP strategy for comparison. The figures show clearly that final cumulative wealth increases as the negative sensitivity grows, and becomes stable as the negative sensitivity exceeds certain critical values, indicating that the power of mean reversion has been thoroughly exploited by our strategy. Needless to say, even though our parameter setting, i.e., $-\epsilon{=}0.5$, is not the *best* setting, the proposed CWMR still significantly surpasses existing approaches.

5.3 Practical Issues: Transaction Cost

To evaluate the performance when the market model is not friction-less, we conduct empirical experiments on the proposed CWMR strategy with proportional transaction costs [2, 4]. Figure 5 shows the results on the five datasets with varying transaction costs from 0% to 1% (we extend x-axis on the STI dataset since the break-even level exceeds 1%), plus the cumulative wealth achieved by Market, BCRP and the state of the arts (Anticor and B^{NN}). As we can observe, the performance with transaction costs is market dependent, in most cases, especially with small rates, CWMR outperforms the state of the arts. In other cases, though both powered by mean reversion, CWMR underperforms Anticor, showing that aggressiveness results in more transaction costs. Nevertheless, the results compared with the benchmarks clearly demonstrate that on most datasets (except NYSE (N)), the performance is considerably robust with respect to the transaction costs, where the break-even rates are always above 0.6% (around 0.2% on NYSE (N)). Thus, CWMR can withstand small transaction costs even though we do not explicitly tackle it in our study.

5.4 Computational Time

Other than the promising cumulative wealth performance, CWMR also runs quite efficiently. Table 2(c) shows the computational time of the CWMR-Stdev and three performance-competing strategies (Anticor, B^{K} and B^{NN}) on all datasets with the same platform. From the table, we can observe that CWMR costs much less computational time than the three performance-competing strategies, which validates its computational efficiency.

5.5 Discussion and Thread of Validity

Any PS strategy claiming excess returns should be carefully scrutinized, including CWMR. To recall, we had made several assumptions in Section 2 regarding transaction costs, market liquidity, and impact cost, which would affect the practical deployment of the proposed strategy. Ignoring transaction costs can reduce the problem complexity, which is common in existing studies. In Section 5.3, we had examined the effect of varying transaction costs with results showing that CWMR can withstand moderate transaction costs. The second assumption is that the market is liquid and one can trade any quantity at quoted price. In the experiments, we have tried to minimize the effect of the market liquidity by arbitrary choosing stocks from the market index, which usually have large capitalization and thus have a high market liquidity. The last assumption is that portfolio has no impact to the market. However, as we observed, the portfolio increases astronomically and would inevitably impact the market. In reality, we can reduce the market impact by controlling the size of the portfolio, as typically done by some quantitative funds. Finally, we note again, even in such theoretically “perfect market” typically adopted in previous studies, none has ever claimed such eye-catching performance on the benchmark testbeds.

Back tests in the historical market may suffer from “data-snooping bias” issue. In particular, following previous works, in our datasets the composition stocks never delisted from the markets, i.e., survived over the entire trading period. Another possible “data-snooping bias” is the dataset selection. In fact, we developed CWMR approaches based solely on the widely adopted NYSE (O) dataset, and collected other three datasets (NYSE (N), STI, and MSCI) after the algorithm was fully developed.

6 Conclusions

This paper proposed a novel on-line portfolio selection strategy named “Confidence Weighted Mean Reversion” (CWMR), which successfully applied machine learning techniques for on-line portfolio selection by exploiting the mean reversion property of the financial markets. Unlike the existing techniques using only the first order information, CWMR exploits both the first and second order information of the portfolio vectors. Empirically CWMR surpassed all the competing existing techniques on various up-to-date testbeds from real markets. Future work will study theoretical bounds of the logarithmic wealth achieved by CWMR and its performance with high transaction costs.

Acknowledgments

This work was supported by Singapore MOE ARC tier-1 research grant (RG67/07) and tier-2 research grant (T208B2203).

Methods	NYSE(O)	NYSE(N)	TSE	STI	MSCI
Market	14.50	18.06	1.61	1.84	1.03
Best-stock	54.14	83.51	6.28	3.67	1.05
BCRP	250.60	120.32	6.78	9.01	1.11
UP	26.68	31.49	1.60	2.85	1.09
EG	27.09	31.00	1.59	2.80	1.09
ONS	109.19	21.59	1.62	7.95	1.26
SP	27.08	31.55	1.60	2.83	1.09
GRW	27.73	30.45	1.61	2.84	1.10
M0	113.50	40.94	1.26	2.96	1.08
Anticor	1.71E+07	2.10E+05	28.77	628.89	3.10
B^{K}	1.08E+09	4.64E+03	1.62	22.59	2.84
B^{NN}	3.35E+11	6.80E+04	2.27	431.09	95.29
CWMR-Var	**6.40E+15**	**1.42E+06**	**324.65**	**6.79E+07**	**155.76**
CWMR-Stdev	**6.20E+15**	**1.28E+06**	**322.52**	**6.57E+07**	**155.82**
CWMR-Var-s	4.31E+15	1.23E+06	318.58	4.54E+07	90.23
CWMR-Stdev-s	4.32E+15	1.11E+06	318.70	4.52E+07	90.04

(a) Cumulative Wealth

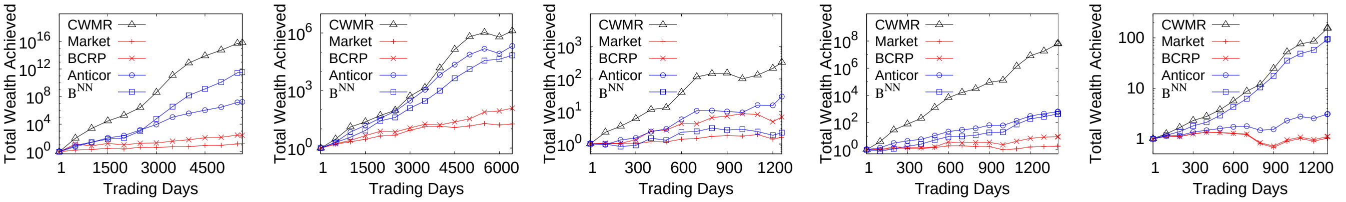
Stat. Attr.	NYSE (O)	NYSE (N)	TSE	STI	MSCI
Size	5651	6431	1259	1404	1304
MER (CWMR)	0.0070	0.0027	0.0057	0.0137	0.0040
MER (Market)	0.0005	0.0005	0.0004	0.0005	0.0001
Winning Ratio	0.5636	0.5197	0.5616	0.6731	0.6511
α	0.0064	0.0021	0.0051	0.0129	0.0039
β	1.2132	1.1377	1.5182	1.5640	1.0358
t-statistics	15.9256	5.9278	3.8944	14.1282	13.8449
p-value	0.0000	0.0000	0.0001	0.0000	0.0000

(b) Statistical Test of CWMR-Stdev

Methods	NYSE (O)	NYSE (N)	TSE	STI	MSCI
Anticor	1645	751	2118	284	8
B^{K}	7.89E+04	5.78E+04	6.35E+03	4.38E+03	1.36E+03
B^{NN}	4.93E+04	3.39E+04	1.32E+04	5.50E+03	1.46E+03
CWMR	123	68	162	14	2

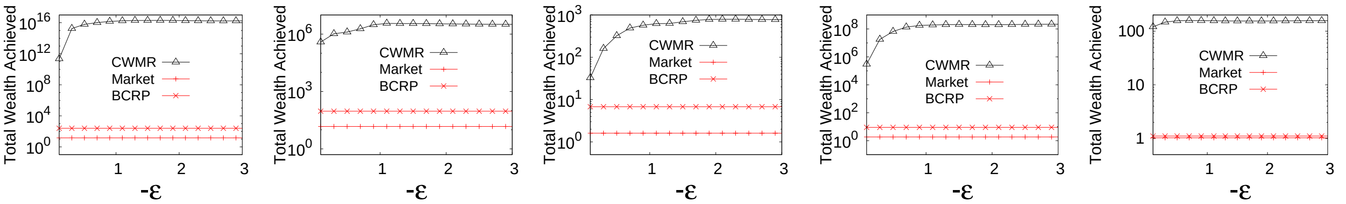
(c) Computational Time (seconds)

Figure 2: Performance evaluation: (a). Cumulative wealth achieved by various trading strategies on the five datasets. The top two best results in each dataset are highlighted in bold font. (b). Statistical t-test of the performance of the CWMR-Stdev on the stock datasets. MER denotes the Mean Excess Return. (c). Computational time costs (seconds) on the five datasets achieved by performance comparable state-of-the-art trading strategies.



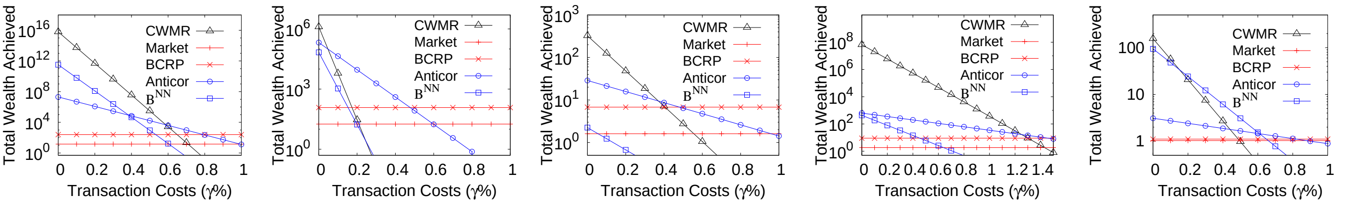
(a) NYSE (O) (b) NYSE (N) (c) TSE (d) STI (e) MSCI

Figure 3: Trend of cumulative wealth achieved by proposed CWMR-Stdev and various strategies during the entire trading period on the stock datasets.



(a) NYSE (O) (b) NYSE (N) (c) TSE (d) STI (e) MSCI

Figure 4: Parameter Sensitivity of the total wealth achieved by CWMR-Stdev with respect to sensitivity parameter $-\epsilon$.



(a) NYSE (O) (b) NYSE (N) (c) TSE (d) STI (e) MSCI

Figure 5: Scalability of the total wealth achieved by CWMR-Stdev with respect to transaction cost rate $\gamma\%$.

Appendix A: Proof of Proposition 1

Proof. First let us replace the log return function using its first order Taylor expansion at μ_i, i.e., $\log(\mu\cdot\mathbf{x}_i)\approx\log(\mu_i\cdot\mathbf{x}_i)+\frac{\mathbf{x}_i\cdot(\mu-\mu_i)}{\mu_i\cdot\mathbf{x}_i}$. Moreover, since considering the non-negativity constraint introduces too much complexity, at this stage we solve the problem without considering it, and instead later we will project the solution to simplex domain to obtain the required portfolio.

The Lagrangian for optimization problem (2) is,

$$\mathcal{L}=\frac{1}{2}\left(\log\left(\frac{\det\boldsymbol{\Sigma}_i}{\det\boldsymbol{\Sigma}}\right)+\mathrm{Tr}\left(\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Sigma}\right)+(\mu_i-\mu)^\top\boldsymbol{\Sigma}_i^{-1}(\mu_i-\mu)\right)$$
$$+\lambda\left(\phi\mathbf{x}_i^\top\boldsymbol{\Sigma}\mathbf{x}_i+\log(\mu_i\cdot\mathbf{x}_i)+\frac{\mathbf{x}_i\cdot(\mu-\mu_i)}{\mu_i\cdot\mathbf{x}_i}-\epsilon\right)+\eta(\mu\cdot\mathbf{1}-1).$$

Taking the gradient of the Lagrangian with respect to μ and setting it to zero, we can get the update of μ_{i+1} : $\mu_{i+1}=\mu_i-\boldsymbol{\Sigma}_i\left(\lambda\frac{\mathbf{x}_i}{\mu_i\cdot\mathbf{x}_i}+\eta\mathbf{1}\right)$, where $\boldsymbol{\Sigma}_i$ is assumed to be non-singular. Multiplying both sides of the update with $\mathbf{1}^\top$, we can get η, i.e., $1=1-\mathbf{1}^\top\boldsymbol{\Sigma}_i\left(\lambda\frac{\mathbf{x}_i}{\mu_i\cdot\mathbf{x}_i}+\eta\mathbf{1}\right)\Rightarrow\eta=-\lambda\frac{\bar{x}_i}{\mu_i\cdot\mathbf{x}_i}$, where $\bar{x}_i=\frac{\mathbf{1}^\top\boldsymbol{\Sigma}_i\mathbf{x}_i}{\mathbf{1}^\top\boldsymbol{\Sigma}_i\mathbf{1}}$ denotes the confidence weighted average of the i^{th} price relative. Plugging η to the update of μ_{t+1}, we can get, $\mu_{t+1}=\mu_i-\lambda\boldsymbol{\Sigma}_i\left(\frac{\mathbf{x}_i-\bar{x}_i\mathbf{1}}{\mu_i\cdot\mathbf{x}_i}\right)$. Moreover, calculating the derivative with respect to $\boldsymbol{\Sigma}$ and setting it to zero, we can also have the update of $\boldsymbol{\Sigma}_{i+1}$, i.e., $\boldsymbol{\Sigma}_{i+1}^{-1}=\boldsymbol{\Sigma}_i^{-1}+2\phi\lambda\mathbf{x}_i\mathbf{x}_i^\top$. Thus, the updates for μ_{i+1} and $\boldsymbol{\Sigma}_{i+1}$ are represented as:

$$\mu_{i+1}=\mu_i-\lambda\boldsymbol{\Sigma}_i\left(\frac{\mathbf{x}_i-\bar{x}_i\mathbf{1}}{\mu_i\cdot\mathbf{x}_i}\right),\ \boldsymbol{\Sigma}_{i+1}^{-1}=\boldsymbol{\Sigma}_i^{-1}+2\lambda\phi\mathbf{x}_i\mathbf{x}_i^\top. \quad (4)$$

Now let us solve the Lagrange multiplier λ_{i+1} using the KKT conditions. The inverse of the $\boldsymbol{\Sigma}_{i+1}$ can also be calculated using Woodbury equation [13], i.e., $\boldsymbol{\Sigma}_{i+1}=\boldsymbol{\Sigma}_i-\boldsymbol{\Sigma}_i\mathbf{x}_i\frac{2\lambda\phi}{1+2\lambda\phi\mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{x}_i}\mathbf{x}_i^\top\boldsymbol{\Sigma}_i$. The KKT conditions imply that either λ=0, and no update is needed, or the constraint in optimization (2) is an equality after the update. Taking Eq. (4) and Woodbury equation to the equality version of the constraint and rearranging the terms, we have:

$$a\lambda^2+b\lambda+c=0, \quad (5)$$

with $a=\frac{2\phi V_i^2-2\phi V_i\bar{x}_i\mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{1}}{M_i^2}$, $b=\frac{V_i-\bar{x}_i\mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{1}}{M_i^2}+2\phi V_i(\epsilon-\log M_i)$, $c=\epsilon-\log M_i-\phi V_i$, and $M_i=\mu_i\cdot\mathbf{x}_i$ is the return mean and V_i=$\mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{x}_i$ denotes the return variance of the i^{th} trading day. Above Eq. (5) is clearly a quadratic equation in λ. We can calculate its roots γ_{i1} and γ_{i2} as follows, $\gamma_{i1}=\frac{-b+\sqrt{b^2-4ac}}{2a}$, $\gamma_{i2}=\frac{-b-\sqrt{b^2-4ac}}{2a}$. To ensure the non-negativity of the Lagrangian multiplier, we can project the value to $[0,+\infty)$, $\lambda_{i+1}=\max\{\gamma_{i1},\ \gamma_{i2},\ 0\}$.

In practice, since we only adopt the diagonal elements of the covariance matrix, it is equivalent to computing λ_{i+1} as Eq. (5) but updating the covariance matrix with the following rule instead, $\boldsymbol{\Sigma}_{i+1}^{-1}=\boldsymbol{\Sigma}_i^{-1}+2\lambda_{i+1}\phi\mathrm{diag}^2(\mathbf{x}_i)$, where $\mathrm{diag}(\mathbf{x}_i)$ denotes the diagonal matrix with the elements of $\mathbf{x}_i$ on its main diagonal. □

Appendix B: Proof of Proposition 2

Proof. Following the same procedure as the proof of Proposition 1, we adopt the Taylor expansion of the log function and ignore the non-negativity of the portfolio vector first.

The Lagrangian for the optimization (3) is,

$$\mathcal{L}=\frac{1}{2}\left(\log\left(\frac{\det\boldsymbol{\Upsilon}_i^2}{\det\boldsymbol{\Upsilon}^2}\right)+\mathrm{Tr}\left(\boldsymbol{\Upsilon}_i^{-2}\boldsymbol{\Upsilon}^2\right)+(\mu_i-\mu)^\top\boldsymbol{\Upsilon}_i^{-2}(\mu_i-\mu)\right)$$
$$+\lambda\left(\phi\left\|\boldsymbol{\Upsilon}\mathbf{x}_i\right\|+\log(\mu_i\cdot\mathbf{x}_i)+\frac{\mathbf{x}_i\cdot(\mu-\mu_i)}{\mu_i\cdot\mathbf{x}_i}-\epsilon\right)+\eta(\mu\cdot\mathbf{1}-1).$$

Taking the gradient of the Lagrangian with respect to μ and setting it to zero, we get the update of μ_{i+1}, $\mu_{i+1}=\mu_i-\boldsymbol{\Upsilon}_i^2\left(\lambda\frac{\mathbf{x}_i}{\mu_i\cdot\mathbf{x}_i}+\eta\mathbf{1}\right)$, where $\boldsymbol{\Upsilon}_i$ is non-singular. Multiplying both sides by $\mathbf{1}^\top$, we get, $1=1-\mathbf{1}^\top\boldsymbol{\Upsilon}_i^2\left(\lambda\frac{\mathbf{x}_i}{\mu_i\cdot\mathbf{x}_i}+\eta\mathbf{1}\right)\Rightarrow\eta=-\frac{\lambda\bar{x}_i}{\mu_i\cdot\mathbf{x}_i}$, where $\bar{x}_i=\frac{\mathbf{1}^\top\boldsymbol{\Upsilon}_i^2\mathbf{x}_i}{\mathbf{1}^\top\boldsymbol{\Upsilon}_i^2\mathbf{1}}$ is the confidence weighted average of i^{th} price relative. Plugging it to the update of μ_{i+1}, we get, $\mu_{i+1}=\mu_i-\lambda\boldsymbol{\Upsilon}_i^2\frac{\mathbf{x}_i-\bar{x}_i\mathbf{1}}{\mu_i\cdot\mathbf{x}_i}$. Moreover, calculating the derivative of $\boldsymbol{\Upsilon}$ and set it to zero, we also have the update of $\boldsymbol{\Upsilon}_{i+1}^2$, $\boldsymbol{\Upsilon}_{i+1}^{-2}=\boldsymbol{\Upsilon}_i^{-2}+\lambda\phi\frac{\mathbf{x}_i\mathbf{x}_i^\top}{\sqrt{\mathbf{x}_i^\top\boldsymbol{\Upsilon}_{i+1}^2\mathbf{x}_i}}$. The two updates can be expressed in terms of the covariance matrix as follows,

$$\mu_{i+1}=\mu_i-\lambda\boldsymbol{\Sigma}_i\frac{\mathbf{x}_i-\bar{x}_i\mathbf{1}}{\mu_i\cdot\mathbf{x}_i},\ \boldsymbol{\Sigma}_{i+1}^{-1}=\boldsymbol{\Sigma}_i^{-1}+\lambda\phi\frac{\mathbf{x}_i\mathbf{x}_i^\top}{\sqrt{\mathbf{x}_i^\top\boldsymbol{\Sigma}_{i+1}\mathbf{x}_i}}. \quad (6)$$

Here, $\boldsymbol{\Sigma}_{i+1}$ is PSD and non-singular.

Now let us solve the Lagrangian multiplier using its KKT condition. We compute the inverse using Woodbury equation [13], $\boldsymbol{\Sigma}_{i+1}=\boldsymbol{\Sigma}_i-\boldsymbol{\Sigma}_i\mathbf{x}_i\left(\frac{\lambda\phi}{\sqrt{\mathbf{x}_i^\top\boldsymbol{\Sigma}_{i+1}\mathbf{x}_i}+\lambda\phi\mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{x}_i}\right)\mathbf{x}_i^\top\boldsymbol{\Sigma}_i$. Then, let $M_i=\mu_i\cdot\mathbf{x}_i$, $V_i=\mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{x}_i$, and $U_i=\mathbf{x}_i^\top\boldsymbol{\Sigma}_{i+1}\mathbf{x}_i$, and multiplying the update of μ_{i+1} by $\mathbf{x}_i^\top$ (left) and $\mathbf{x}_i$ (right), we get U_i=$V_i-V_i\left(\frac{\lambda\phi}{\sqrt{U_i}+V_i\lambda\phi}\right)V_i$ which can be solved for U_i to obtain $\sqrt{U_i}=\frac{-\lambda V_i\phi+\sqrt{\lambda^2V_i^2\phi^2+4V_i}}{2}$. The KKT condition implies that either λ=0, and no update is needed, or the constraint in the optimization problem Eq. (3) is an equality after the update. Substitute Eq. (6) and Woodbury equation to the equality version of the constraint, after rearranging in terms of λ, we obtain:

$$a\lambda^2+b\lambda+c=0, \quad (7)$$

with $a=\left(\frac{V_i-\bar{x}_i\mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{1}}{M_i^2}+\frac{V_i\phi^2}{2}\right)^2-\frac{V_i^2\phi^4}{4}$, $b=2(\epsilon-\log M_i)\left(\frac{V_i-\bar{x}_i\mathbf{x}_i^\top\boldsymbol{\Sigma}_i\mathbf{1}}{M_i^2}+\frac{V_i\phi^2}{2}\right)$, and $c=(\epsilon-\log M_i)^2-V_i\phi^2$. Let γ_{i1} and γ_{i2} be its roots, thus $\gamma_{i1}=\frac{-b+\sqrt{b^2-4ac}}{2a}$, $\gamma_{i2}=\frac{-b-\sqrt{b^2-4ac}}{2a}$. To ensure the non-negativity of the Lagrangian multiplier, we project the value to $[0,+\infty)$, $\lambda_{i+1}=\max\{\gamma_{i1},\gamma_{i2},0\}$.

Similar to Proposition 1, we can update the diagonal covariance matrix as, $\boldsymbol{\Sigma}_{i+1}^{-1}=\boldsymbol{\Sigma}_i^{-1}+\phi\frac{\lambda_{i+1}}{\sqrt{U_i}}\mathrm{diag}^2(\mathbf{x}_i)$. □

References

[1] A. Agarwal, E. Hazan, S. Kale, and R. E. Schapire. Algorithms for portfolio management based on the newton method. In *ICML*, pages 9–16, 2006.

[2] A. Blum and A. Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3): 193–205, 1999.

[3] A. Borodin, R. El-Yaniv, and V. Gogan. On the competitive theory and practice of portfolio selection (extended abstract). In *LATIN*, pages 173–196, 2000.

[4] 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.

[5] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[6] L. Breiman. Optimal gambling systems for favorable games. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, 1:65–78, 1961.

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

[8] T. M. Cover and D. H. Gluss. Empirical bayes stock market portfolios. *Advances in applied mathematics*, 7(2):170–181, 1986.

[9] K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer. Online passive-aggressive algorithms. *Journal of Machine Learning Research*, 7:551–585, 2006.

[10] K. Crammer, M. Dredze, and F. Pereira. Exact convex confidence-weighted learning. In *NIPS*, pages 345–352, 2008.

[11] M. Dredze, K. Crammer, and F. Pereira. Confidence-weighted linear classification. In *ICML*, pages 264–271, 2008.

[12] E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann. *Modern Portfolio Theory and Investment Analysis*. Wiley, 2006.

[13] G. H. Golub and C. F. Van Loan. *Matrix Computations, 3rd ed*. Baltimore, MD: Johns Hopkins, 1996.

[14] R. Grinold and R. Kahn. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. McGraw-Hill, 1999.

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

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

[17] E. Hazan and S. Kale. On stochastic and worst-case models for investing. In *NIPS*, pages 709–717, 2009.

[18] E. Hazan and C. Seshadhri. Efficient learning algorithms for changing environments. In *ICML*, pages 393–400, 2009.

[19] D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth. On-line portfolio selection using multiplicative updates. *Mathematical Finance*, 8(4):325–347, 1998.

[20] N. Jegadeesh. Evidence of predictable behavior of security returns. *Journal of Finance*, 45(3):881–98, 1990.

[21] J. Kelly, J. A new interpretation of information rate. *Bell Systems Technical Journal*, 35:917–926, 1956.

[22] H. A. Latané. Criteria for choice among risky ventures. *The Journal of Political Economy*, 67(2):144–155, 1959.

[23] T. Levina and G. Shafer. Portfolio selection and online learning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 16(4):437–473, 2008.

[24] B. Li, S. C. Hoi, and V. Gopalkrishnan. Corn : Correlation-driven nonparametric learning approach for portfolio selection. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 2011.

[25] A. W. Lo and A. C. MacKinlay. When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3(2):175–205, 1990.

[26] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[27] E. Ordentlich and T. M. Cover. On-line portfolio selection. In *COLT*, pages 310–313, 1996.

[28] W. F. Sharpe. A simplified model for portfolio analysis. *Management Science*, 9:277–293, 1963.

[29] W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.

[30] Y. Singer. Switching portfolios. *International Journal of Neural Systems*, 8(4):488–495, 1997.

[31] E. Thorp. Portfolio choice and the kelly criterion. In *Business and Economics Section of the American Statistical Association*, pages 215–224, 1971.

[32] V. G. Vovk. Aggregating strategies. In *COLT*, pages 371–386, 1990.

[33] V. G. Vovk and C. Watkins. Universal portfolio selection. In *COLT*, pages 12–23, 1998.