# A Maximum Wealth Asset Index and Mixture Strategies for Universal Portfolios on Subsets of Stocks

Wei Qiu And Andrew R. Barron

*Yale University*

**Abstract**

In this paper we first show how to compute a combination of assets producing an appropriate index of past performance. The desired index is equal to $S_T^{max} = \max_{\underline{b}} S_T(\underline{b})$ which is the maximum of $T$-period investment return $S_T(\underline{b}) = \prod_{t=1}^{T} \underline{b} \cdot \underline{x}_t$, where $\underline{x}_i$ is the vector of returns for the $t^{th}$ investment period, and $\underline{b}$ is the portfolio vector specifying the fraction of wealth allocated to each asset. We provide an iterative algorithm to approximate this index, where at step $k$ the algorithm produces a portfolio with at most $k$ assets selected among $M$ available assets. We show that the multi-period wealth factor $S_T(\underline{b}_k)$ converges to the maximum $S_T^{max}$ as $k$ increases. Furthermore, in the exponent the wealth factor is within $c^2/k$ of the maximum, where $c$ is determined by the empirical volatility of the stock returns, and we compare this computation to what is achieved by general procedures for convex optimization. This $S_T^{max}$ provides an index of historical asset performance which corresponds to the best constant rebalanced portfolio with hindsight. Surprisingly, we find empirically that a small handful of stocks among hundreds of candidate stocks are sufficient to have come close to $S_T^{max}$.

Universal portfolios are strategies for updating portfolios each period to achieve actual wealth with exponent provably close to what is provided by $S_T^{max}$. Not only do we show approximate computation of $S_T^{max}$ using subsets of stocks, we also derive new universal portfolio strategies based on such subsets. Under a volatility condition, the universal portfolio achieves a wealth exponent that drops from the maximum not more than order $\sqrt{\frac{\log M}{T}}$.

# 1 INTRODUCTION

In multi-period investment with a total of $M$ stocks it is important to decide which stocks are to be included in the fund and what fractions of resources are to be allocated to each of them. An investor may choose to rebalance the portfolio each trading period in accordance with a portfolio $\underline{b} = (b_1, b_2, \ldots, b_M)$ specifying the fraction of wealth to be invested in each of the stocks. Price fluctuation leads to new asset values at the end of each investment period, in which case the investor may trade portions of each asset to restore the specified fractions. For a succession of investment periods $t = 1, 2, \cdots, T$, let $x_{t,i}$ be the return, also called "wealth factor", for stock $i$ at time $t$, which means the ratio of the price plus dividend at the end of period $t$ to the price at start of period $t$. It provides a vector of returns $\underline{x}_t = (x_{t,1}, x_{t,2}, \cdots, x_{t,M})$ for period $t$. Then with portfolio $\underline{b}$ the investor achieves a return of $\underline{b} \cdot \underline{x}_t = b_1 x_{t,1} + b_2 x_{t,2} + \cdots + b_M x_{t,M}$ for that period. Let $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_T$ be such vectors of returns for a sequence of $T$ investment periods. Then, with rebalancing to portfolio $\underline{b}$ each period, the multi-period return (compounded wealth per dollar initially invested) is,

$$S_T(\underline{b}) = \prod_{t=1}^{T} \underline{b} \cdot \underline{x}_t. \tag{1}$$

The study of the wealth surface $S_T(\underline{b})$ is relevant for examining with hindsight what performance would have been achieved for various portfolios. In particular, there is an interest in the maximum wealth $S_T^{max} = \max_{\underline{b}} S_T(\underline{b})$ and in the portfolio $\underline{b}^{max} = (b_1^{max}, b_2^{max}, \cdots, b_M^{max})$ that would have achieved it. Characteristics of that maximum wealth portfolio reveal historically important stocks and the best fraction of wealth to have retained in each. Moreover, identification of such portfolios from past data may be useful for speculation as to which stocks to invest for subsequent trading periods. We will regard $S_T^{max}$ as an asset index, which refers to the collective performance of a given set of stocks over a given historical time period. Specifically, this index corresponds to the best constant rebalanced portfolio with hindsight.

A focus of attention in this paper is an iterative algorithm for the maximization of $S_T(\underline{b})$, which constructs the portfolio of historically optimal performance. The hindsight maximum wealth at the end of investment period $T$ is $S_T(\underline{b}^{max}) = \prod_{t=1}^{T} \underline{b}^{max} \cdot \underline{x}_t$. We provide an algorithm for this maximization, which chooses stocks from the pool of candidates in a greedy fashion. At the $k^{th}$ step the algorithm either introduces an additional stock to the portfolio or adjusts the weight given to a stock already in the portfolio so as to best balance with the weight of other stocks in the proceeding steps.

Thus the algorithm produces a sequence of portfolios $\underline{b}_k$ where at step $k$ we have included at most $k$ stocks. The multi-period wealth factor $S_T(\underline{b}_k)$, $k = 1, 2, \cdots$ achieved by this sequence of portfolios $\underline{b}_k$ is shown to converge to the maximum $S_T^{\max} = \max_{\underline{b}} S_T(\underline{b})$. In practice we see that it rarely requires more than a few stocks to come close to the maximum. Moreover, we provide theory which shows that an exponent characterizing the wealth at step $k$ is below the maximum by not more than $c^2/k$ for $k = 1, 2, \cdots$. Thus with $k$ stocks we reach approximately the same return as that of the optimal portfolio which has the freedom to have allocated wealth in all the stocks.

Writing $S_T(\underline{b}) = e^{Ty(\underline{b})}$ we find that the wealth exponent $y(\underline{b}) = y_T(\underline{b})$ is concave function of the portfolio $\underline{b}$. Thus we may regard the algorithm provided here as solving a concave optimization problem. We will contrast the method developed here with a general purpose algorithm for maximizing concave function subject to convex constraint sets (Nesterov and Nemirovski's interior point method [3]) for which there are also bounds on the number of computation steps required for specified accuracy.

In the practice of investment one requires a sequence of portfolios $\underline{b}_t$ updated each period $t$ based on what has been observed up to that time. A result of Cover [8] (refined further in Cover and Ordentlich [9] and Xie and Barron [36]) shows that $S_T^{max}$ is achievable by a universal portfolio updating strategy, in the sense that the actual wealth exponent drops from what $S_T^{max}$ achieves by not more than $\frac{M-1}{2T} \log \frac{T}{2\pi} + \frac{c_M}{T}$, uniformly over all possible stock return outcomes, where $c_M$ is a constant. The universal portfolios use at each time $t$ a weighted combination of portfolios $\underline{b}$ weighted by the wealth $S_t(\underline{b})$ up to that time. In the present paper, we give related mixture portfolios that we show achieve a wealth exponent that is within $c\sqrt{(\log M)/T}$ of the maximum where $c$ depends on an empirical relative volatility of the stocks. As we shall discuss, since the drop depends only logarithmically on the number of stock $M$, this new bound is preferable to Cover's bound when $M$ is large. Helmbold et al. [16] also showed a similar drop when using their portfolio updating rule with learning rate $\eta$. However, choosing such a $\eta$ requires the knowledge of both the number of trading periods $T$ and a lower bound of price return $x_{t,i}$ for all $t$ before starting to invest at time $t = 1$. We devise mixtures that do not require such knowledge in advance.

In the last section, we explore the use of our mixture portfolios imbedded in a strategy for updating our stock portfolios every investment period on actual stock return data. We also provide a strategy, which uses our wealth maximization algorithm, of selecting past optimal portfolios. In particular, one may use for each month a portfolio equal to the portfolio that made the most wealth with hindsight over a suitable number of preceding months. It shows impressive return compared to other investment strategies, such as the Standard and Poor 500 Index.

Constant rebalanced portfolios have played an important role in finance literature in both arbitrary sequence analysis and in stochastic models. As we have mentioned, their role in analysis of wealth for arbitrary return sequences as in Cover [8], Cover & Ordentlich [9] and Helmbold et al. [16] is to provide a target wealth $\max_{\underline{b}} S_T(\underline{b})$ approximately reached by practical update strategies. In stochastic analysis constant rebalanced portfolios are shown to be optimal when stock returns are modeled as independent across time for growth rate optimal portfolios as in Kelly [24], Brieman [4], Algoet & Cover [2], [3], and for certain utility functions as shown in von Neumann & Morgenstern [33],[34]. We revisit much of this literature further below.

The purpose of this paper is to provide a provably accurate algorithm for computation of $S_T^{max}$ and $\underline{b}^{max}$ and to provide a mixture strategy for updating portfolios which achieve wealth exponent provably close to $y_{max}$.

Our analysis uses the arbitrary sequence perspective. We show that the maximal wealth is nearly realized by our mixture strategy for all return sequences with a small drop in the exponent of wealth dependent upon volatility properties of the sequence. In discussing the wealth $S_T(\underline{b})$ and its maximum $S_T^{max}$ it is equivalent to work with the representation $S_T(\underline{b}) = e^{Ty(\underline{b})}$ and $S_T(\underline{b}) = (1 + r(\underline{b}))^T$ where $e^{Ty(\underline{b})}$ is the $T^{th}$ root of $S_T(\underline{b})$ (a geometric mean) and $r(\underline{b}) = e^{Ty(\underline{b})} - 1$ is the corresponding compounding rate of return. The arbitrary sequence analysis may be regarded as applying to any monotone increasing function (utility) of the multi-period wealth $S_T(\underline{b})$ as all such will share the same target of performance based on $S_T^{max}$ and the associated optimal $\underline{b}^{max}$. In contrast, *expected* utility analysis is quite a different matter. For instance, one can have a dramatically different portfolio maximizing the expectation of a power of $S_T(\underline{b})$ compared to maximizing the expected logarithm of $S_T(\underline{b})$.

There is much previous work in portfolio theory that has focused on the mean-variance criterion and associated efficient frontier, which was formulated by Markowitz [30] as an optimization problem with quadratic objective and linear constraints. It seeks the portfolio weights that minimize the variance for a given value of mean return or equivalently maximize the mean return for a given variance. In this setting, variance becomes a proxy for risk and the investor tries to maximize expected return for a given level of risk. This forms the basis of the Sharpe-Markowitz theory of investment. Sharp [35] gives an introduction on this topic. Goetzmann [15] discusses and gives empirical results for this mean-variance criterion using Standard & Poor 500 stocks, corporate and government bonds and other asset classes over the period 1970 through 1995. In Section 5 we compare the wealth achieved by our strategy to that given in Goetzmann. Latane [25], Hakansson [17] and Elton & Gruber [12] discuss maximization of expected geometric mean return. Bernstein and Wilkinson [6] modified the mean-variance formalism by maximizing geometric mean return with a variance constraint.

Generally, the traditional view of finance has been that an investor shall choose a portfolio by optimizing an expected utility function. Fishburn [14] and Kreps [23] provide an introduction. In this literature, a utility function is regarded as resonable if it is both increasing, because more money is better, and concave, because investors are risk adverse. Quadratic utility and exponential utility are among the most commonly used utility functions. Many other types of utility functions have been studied, such as the von Neumann-Morgenstern [33], [34] class of utility functions, which includes the power utility $U(s) = (s^\alpha - 1)/\alpha$ for $\alpha < 1$ and the logarithmic utility $U(s) = \log s$. These utilities are distinguished by the property that for $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_T$ i.i.d. the sequence of portfolio actions best for $EU(S_T)$ is the same as the choice best for $EU(\underline{X}_t \cdot \underline{b})$ each period. The logarithmic utility is shown to produce the highest growth rate of wealth in probability in Algoet and Cover (1988), where if the $\underline{X}_t$ are not i.i.d. the optimal action is to maximize the conditional expected logarithm given the past. In the i.i.d. case one simply observes that the exponent $y_T(\underline{b})$ converge to $E_P \log \underline{X} \cdot \underline{b}$ in probability, so any $\underline{b}$ other than $\underline{b}^* = \arg\max E_P \log \underline{X} \cdot \underline{b}$ will have exponentially smaller growth in probability. Nevertheless, we emphasize that we do not need any stochastic assumptions for the main conclusions of this thesis. We use as the standard of comparison $S_T^{max} = \max_{\underline{b}}(\prod_{t=1}^{T} \underline{x}_t \cdot \underline{b})$. We show how to compute it for a given sequence of returns and we give strategies for updating investment portfolios which achieve an exponent that matches what $S_T^{max}$ achieves with a drop from the maximum explicitly controlled.

# 2 SUMMARY OF METHODS AND RESULTS

We present our algorithm and theory for maximum wealth portfolio computation and for mixture portfolios strategies in this section.

## 2.1 WEALTH MAXIMIZATION THEORY

We first introduce a tool for constructing an asset index, namely the computation of $S_T^{max} = \max_{\underline{b}} S_T(\underline{b})$. It is an algorithm, which, when given a series of returns in $T$ periods for $M$ stocks, determines the rebalancing portfolio that would have made the maximal wealth for these stocks in that time frame. We will show that the total computations needed to achieve the targeted accuracy $\varepsilon$ by our algorithm is $N_{new}(\varepsilon) = cMT/\varepsilon$. Here $c$ will depend on the sequence of returns $\underline{x}_1 \cdots \underline{x}_T$ and is not a universal constant. Nevertheless, we argue that for moderate accuracies $\varepsilon$ the typical computation time is such that the $cMT/\varepsilon$ is much smaller than the computation time $TM^{4.5}\log(M/\varepsilon)$ guaranteed by an interior point method.

Our algorithm is a multi-step stock selection procedure during which at each step we select one stock from all $M$ stocks. We let $S_T^k$ denote its multiperiod wealth after $k$ steps. The stock selected at step $k$ may be either a stock already selected or a previously unselected stock. For $k = 1$ we put full weight $\alpha_1 = 1$ on the best single stock. Then for $k \geq 2$ the incremental contribution to the portfolio weight of the selected stock is $\alpha_k = 2/(k+2)$. Correspondingly, the portfolio weight of previously selected stocks is downweighted by the factor $1 - \alpha_k = k/(k+2)$. This yields a portfolio $\underline{b}_k$ and portfolio returns $Z_{t,k} = \underline{x}_t \cdot \underline{b}_k$ for $t = 1, \ldots, T$ with contribution to the wealth only from the selected stocks. The compounded wealth with this portfolio is $S_T^k = S_T(\underline{b}_k)$. The algorithm is greedy in that at step $k$ the stock $i = i_k$ selected is the one that (given the $k-1$ previous choices) yields the best such multiperiod portfolio return $\prod_{t=1}^T [(1-\alpha_k)Z_{t,k-1} + \alpha_k x_{t,i}]$ that balances the previous portfolio with the newly chosen stock. Here the portfolio return is updated by $Z_{t,k} = (1-\alpha_k)Z_{t,k-1} + \alpha_k x_{t,i_k}$ and its product $S_T^k = \prod_{t=1}^T Z_{t,k}$ is its multiperiod wealth factor. Instead of the prespecified $\alpha_k = /(k+2)$ we are free alternatively to optimize over choices $0 \le \alpha_k \le 1$ for each $k \ge 2$. As we shall see the bound we develop holds in either case. Now we give our wealth maximization theorem.

**Theorem 1** *Let a sequence of return vectors $\underline{x}_1, \ldots, \underline{x}_T$ be given and let $S_T^{max}$ and $S_T^k$ be defined as above. Our $k$ step algorithm provides a portfolio $\underline{b}_k$ for which*

$$0 \le \frac{1}{T} \log \frac{S_T^{max}}{S_T^k} \le \frac{c^2}{k+3} \tag{2}$$

*or equivalently,*

$$S_T^{max} \ge S_T^k \ge S_T^{max} e^{-T\frac{c^2}{k+3}}, \tag{3}$$

*where $c^2 = 4I\log(2v\sqrt{e})$. Here $I = \frac{1}{T}\sum_{t=1}^T \sum_{i=1}^M b_i^{max}(\frac{x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t})^2$ and $v = \max_{1\le t\le T, 1\le i,j\le M}\{x_{t,i}/x_{t,j}\}$ are empirical measures of volatility which depend on the sequence of returns $\underline{x}_1, \ldots, \underline{x}_T$. They are constants in the sense that they do not depend on the number of iterations $k$.*

Concerning the quantities $I$ and $v$ that arise in the definition of $c$ one may think of $I$ an average squared empirical relative volatility of only the stocks that arise in the optimal $\underline{b}^{max}$. Likewise $v = \max_{t,i,j} x_{t,i}/x_{t,j}$ is a worst case relative volatility over all candidate stocks. The appearance of this $v$ in the bound is somewhat bothersome but we are pleased that at least it appears only through a logarithm.

To summarize the conclusion of Theorem 1, the wealth that would have been achieved at $\underline{b}_k$ has a drop from the maximal wealth exponent by not more than $\frac{c^2}{k+3}$. Furthermore, $S_T^k$ converges to $S_T^{max}$ as $k \to +\infty$. We emphasize that $S_T^{max}$ and its approximation $S_T^k$ are indices based on historically given return sequences. It can also be conceptualized as a target level of (possibly unachievable) performance for future periods.

Next we develop portfolios updated each time period and relate their exponential growth to that of $S_T^{max}$. Further details of the algorithm and proof of Theorem 1 are given in Section 3.

## 2.2 UNIVERSAL PORTFOLIOS

Cover's universal portfolio update strategy [8] gives (initially equal) weight to all portfolios that use all $M$ stocks and each time period updates the weights given to portfolios by the wealth achieved thus far. For Cover's universal portfolio strategy to achieve a nearly maximal wealth exponent, the number of periods $T$ needs to be large compared to $M$. Moreover, computation of the full mixture is a challenge. Now we will develop a universal portfolio strategy, in which we form a mixture of portfolios involving subsets of the stocks, with weights determined by the wealth achieved by these subset portfolios. As we shall see the wealth achieved by certain subsets is nearly maximal. As a consequence our algorithm is shown to achieve wealth near the maximum. This is possible when $T$ is large compared to $\log M$ even if $M$ is large compared to $T$.

Our strategy is to build a mixture of portfolios involving subsets of all stocks, with weights determined by wealth achieved by these subset portfolios. Let $i_1, \ldots, i_k$ be the indices of a subcollection of the $M$ stocks in which repeats are allowed. There are $M^k$ such ordered subcollection and our strategy distributes wealth (initially equally) across all of these subcollection. For each ordered subcollection we provide portfolios weights to which these assets are rebalanced. The resulting wealth $S_{T,k}^{mix}$ after $T$ investment periods is obtained by adding up the multi-period contribution from each subcollection. Further details on this construction are in Section 4.1.

One may think of there being a portfolio manager for each of the subcollection of stocks, each of whom is contracted to follow a prospectus specifying particular portfolio weights to which the stocks are to be rebalanced each period. Our wealth then is the sum of the wealths achieved by each of these managers, weighted by the (equal) fraction of our money initially placed in these funds. The mixture produces at the start of each period a portfolio depending only on the returns up to that time. An alternative implementation is to compute that portfolio update each period and buy and sell as needed to achieve it.

We emphasize the distinction between the wealth $S_{T,k}^{mix}$ which is achievable (either as a mixture of funds or as an updating rule depending only on available return history each period) and the unachievable wealths $S_{T,k}$ and $S_T^{max}$ which we use as target wealths. The following theorem gives a sense in which $S_{T,k}^{mix}$ is near $S_T^{max}$ for every reasonable return sequence provided the number of time periods $T$ is large compared to the logarithm of the number of candidate stocks.

**Theorem 2** *For our mixture strategy, there are choices of $k$ of order $\sqrt{T/\log M}$ such that at time $T$ we achieve a return $S_{T,k}^{mix}$ which has a wealth exponent that drops from $S_T^{max}$ by not more than order $\sqrt{\frac{\log M}{T}}$. Specifically,*

$$S_{T,k}^{mix} \geq S_T^{max} e^{-T(a\sqrt{\frac{\log M}{T}})} \tag{4}$$

*where $a = 2c$ and $c$ is the function of stock return relative volatility specified in Theorem 1.*

If we have prior knowledge of the value of $c$ determined by the stock return relative volatility, we could set $k = c\sqrt{T/\log M}$ which would optimize our bound on the drop to be $2c\sqrt{(\log M)/T}$ where $c$ is as given in Theorem 1. Prior knowledge of the value of $c$ is generally not available. Thus we may use $k = \sqrt{T/\log M}$, which also leads to the same order bound (albeit with a dependence on volatility with $c^2$ in place of $c$). Alternatively we may adapt to what is achieved by the best $k$, by distributing our initial wealth according to a prior $q(k)$ on the subcollection size $k = 1, \ldots, M$. That is $S_T^{mix} = \sum_{k=1}^M q(k) S_{T,k}^{mix}$. For example when $q(k)$ equals $1/M$, we distribute initial wealth evenly across all $k$. Both the fixed $k$ and this adaptive strategy are shown to provide the bound in the proof of Theorem 2 which is given in Section 3. Strengthening of the conclusions is also given there. In particular we see that in fact $S_{T,k}^{mix} \geq M^2 S_T^{max} e^{-T(a\sqrt{\frac{\log M}{T}})}$ which is larger by the factor $M^2$

The Cover and Ordentlich [9] universal portfolio strategy achieves a wealth exponent that is within

$$\frac{M-1}{2T} \log \frac{T}{2\pi} + \frac{c_M}{T} \tag{5}$$

of the maximum wealth exponent $y^{max}$. Compared to our mixture, their universal portfolio achieves an exponent closer to $y^{max}$ for $T$ large compared to $M$. However, an often more realistic setting has $M$ large compared to $T$, but $T$ large compared to $\log M$. In this case, our mixture strategy drop $\sqrt{\frac{\log M}{T}}$ is smaller. A slight refinement of expression (5) is minimax optimal (as shown in Xie and Barron [36]) where in the minimax formulations the maximum is taken over all possible return vectors. Indeed we emphasize that the Cover and Ordentlich (and the Xie and Barron) portfolio strategies achieve the bound on the drop in wealth exponent relative to the maximum *uniformly* over all return sequences. Our improvement (in which the $M$ is replaced by a $\log M$) is not uniform over all return vectors but rather it depends on the observed volatility. Our analysis is related in that we also use a mixture based portfolio based on an initial distribution $\pi$ on the set of partition $\underline{b}$ in the simplex $\{\underline{b} : b_i \geq 0, \sum_{i=1}^{M} b_i = 1\}$. The difference is that our $\pi$ is discrete with points on faces of the simplex determined by subsets of stocks.

Helmbold et al. [16] showed a similar exponent drop bound when using the following portfolio updating rule at time $t$ with learning rate $\eta = 2c'\sqrt{2(\log M)/T}$

$$b_{t+1,i} = \frac{b_{t,i} \exp(\eta x_{t,i}/\underline{b}_t \cdot \underline{x}_t)}{\sum_{j=1}^{M} b_{t,j} \exp(\eta x_{t,j}/\underline{b}_t \cdot \underline{x}_t)} \tag{6}$$

where $c' = \min x_{t,i}$ for all $t \geq 1$ and $i \geq 1$. The choice of proper $\eta$ requires the knowledge of both the price relative volatility bound $c'$ and the number of trading periods $T$.

Having outlined above our conclusions for wealth maximization and for mixture portfolios, we turn in the next two sections, respectively, to develop these two theories in further detail. A key feature is that one rarely needs more than a few stocks. We will show experiments in Section 5 based on our algorithm with real stock market data.

# 3 WEALTH MAXIMIZATION ANALYSIS

This section focuses on the computational task of wealth maximization, that is, the computation of $S_T^{max} = \max_{\underline{b}} S_T(\underline{b})$ and the determination of the rebalancing portfolio $\underline{b}^{max}$ which would achieve this maximum. When given a series of return in $T$ periods for $M$ stocks, this maximization determines the portfolio that would have achieved the maximal wealth for these stocks in that time frame. We relate the computation task to concave optimization, we give further details of our algorithm, and we prove Theorem 1. The total computations needed to achieve the targeted accuracy $\varepsilon$ by our algorithm is $N_{new} = cTM/\varepsilon$ which is typically much smaller than the $TM^{4.5}\log(M/\varepsilon)$ guaranteed by the interior point method.

## 3.1 CONCAVE OPTIMIZATION AND THE INTERIOR POINT METHOD

For any given sequence of returns $\underline{x}_t$, our interest is to determine the portfolio which maximizes $S_T(\underline{b})$, which is equivalent to maximizing the log-wealth function, given by $y(\underline{b}) = \frac{1}{T}\log S_T(\underline{b}) = \frac{1}{T}\sum_{t=1}^{T}\log(\underline{b}\cdot\underline{x}_t)$. We will develop an optimization of $S_T(\underline{b})$, taking advantage of the fact that $y(\underline{b})$ is a concave function of $\underline{b}$ constrained to the $(M-1)$-dimensioned simplex of values where $b_i \geq 0$ and $1 - \sum_{i=1}^{M-1} b_i \geq 0$.

One approach to concave optimization is by existing general purpose algorithms. Consider optimization problems of the following form: $\hat{b} = argmax\ y(b)$ where the $n$-dimensional parameter $b$ is constrained to a convex set. In particular we may have an optimization problem of the form:

$$\begin{array}{ll} \text{maximize} & y(b) \\ \text{subject to} & y_i(b) \geq 0, i = 1, \cdots, M, \end{array} \tag{7}$$

where the functions $y, y_1, \cdots, y_M : R^n \to R$ are concave.

The interior-point method achieves optimization by going through the middle of the solid defined by the problem rather than around its surface. General polynomial algorithm for concave maximization subject to convex constraints have existed since 1976 by Nemirovski and Yudin [37],[38] (who developed the ellipsoidal method). Subsequently, Karmarkar [22] announced a fast polynomial-time interior method for linear programming which is related to classical barrier methods. Later, Nesterov and Nemirovski [32] extended interior-point theory to cover general nonlinear convex optimization problems. The method of solution involves Newton algorithm steps applied to the objective function with a logarithmic barrier penalty with a particular schedule of values of Lagrange multipliers. In [38] they show that for solving the problem (7) with a specified accuracy $\varepsilon$, the total number of operations $N(\varepsilon)$ satisfies

$$N(\varepsilon) \leq CM^{1/2}(Mn^2 + n^3)\ln(\frac{2MB}{\varepsilon}). \tag{8}$$

where $C$ and $B$ are constants. In their analysis, each call to a subroutine to evaluate a function $y(\underline{b})$ or $y_i(\underline{b})$ is regarded as one operation. In our stock setting each evaluation of $y(\underline{b})$ requires $T$ times $M$ elementary operations, where $T$ is the number of time periods and $M$ is the number of stocks. In this case the dimension $n$ and the number of constraints are both of order $M$, the number of stocks. So the total computation time bound for the interior point method is of order

$$N(\varepsilon) = CTM^{4.5}\log(M/\varepsilon). \tag{9}$$

We will contrast the computation time with what is achieved by our algorithm in seeking $S_T^{max}$.

## 3.2 DETAILS OF THE ALGORITHM

As discussed in the introduction, our algorithm is an iterative procedure to select stocks into the new portfolio. We select only one stock at each step, where the stock selected may be among the previously selected stocks (but assigned new weight) or it may be a stock not previously selected by the algorithm.

Let $S_T^k$ be the wealth of the newly constructed portfolio at the end of $k^{th}$ step. We know that,

$$S_T^{max} = S_T(\underline{b}^{max}) = e^{Ty^{max}} \tag{10}$$

where $y^{max} = y(\underline{b}^{max})$.

We show that for each sequence of stock wealth factor $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_T$, there is a $c = c(\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_T)$, such that,

$$S_T^k \geq e^{T(y^{max} - c^2/(k+3))} \tag{11}$$

By inequality (11), we see that after $k$ steps, we are assured a wealth exponent within $c^2/(k+3)$ of the maximum.

Portfolios with one stock correspond to vectors $\underline{b}$ which are non-zero in only one of the $M$-coordinates, where the weight assigned is trivially $\alpha_1 = 1$. The wealth exponent $y_1$ in the single stock case is equal to $\frac{1}{T}\sum_{t=1}^{T} \log x_{t,i}$. The first step picks the $i = i_1$ among $\{1, \ldots, M\}$ at which this is largest. The initial portfolio returns (with $k = 1$) are $Z_{t,1} = x_{t,i_1}$ for $t = 1, 2, \ldots, T$.

Likewise in the second step, given $i_1$, we select stock $i_2$ among $\{1, \ldots, M\}$ together with a weight $\alpha_2 \in [0, 1]$ to maximize the resulting $y(\underline{b})$ which now takes the form

$$\frac{1}{T}\sum_{t=1}^{T} \log\left[(1-\alpha_2)x_{t,i_1} + \alpha_2 x_{t,i_2}\right]. \tag{12}$$

The current portfolio $\underline{b}$ is now non-zero in at most two coordinates. The portfolio returns (with $k = 2$) are now $Z_{t,2} = (1-\alpha_2)x_{t,i_1} + \alpha_2 x_{t,i_2}$ for $t = 1, 2, \ldots, T$. The corresponding stocks in our portfolio after step 2 are $i_1$ and $i_2$ (the two could be the same stock).

In the general step $k$, we select stock $i = i_k$ among $\{1, \ldots, M\}$ with a weight $\alpha_k \in [0, 1]$ to optimize the wealth

$$y_k = \frac{1}{T}\sum_{t=1}^{T} \log\left[(1-\alpha_k)Z_{t,k-1} + \alpha_k x_{t,i_k}\right] \tag{13}$$

where $Z_{t,k-1}$ is the portfolio return for period $t$ at step $k-1$. Similarly,

$$Z_{t,k} = \left[(1-\alpha_k)Z_{t,k-1} + \alpha_k x_{t,i_k}\right]. \tag{14}$$

After $k$ steps the contribution from a previous steps $j$ to the the portfolio weight for stock $i_j$ is $\alpha_j \cdot \prod_{m=j+1}^{k}(1-\alpha_m)$ for $j = 1, \ldots, k-1$. Therefore, our $k$-step portfolio $\underline{\alpha}_k$ for the $k$ selected stocks is

$$\underline{\alpha}_k = \left(\alpha_1 \cdot \prod_{m=2}^{k}(1-\alpha_m), \ldots, \alpha_{k-1}(1-\alpha_k), \alpha_k\right). \tag{15}$$

Recognizing that our procedure permits a stock to be revisited as the selection, we see that the total weight for stock $i$ in the resulting portfolios $\underline{b}_k$ is

$$b_{i,k} = \sum_{j=1}^{k} 1_{\{i_j = i\}} \cdot \alpha_j \prod_{m=j+1}^{k} (1 - \alpha_m), \tag{16}$$

where at $j = k$ the empty product is interpreted as equal to 1. The associated return is $Z_{t,k} = \underline{b}_k \cdot \underline{x}_t = \sum_{i=1}^{k} b_{i,k} x_{t,i}$ and $y_k = y(\underline{b}_k)$. Consequently, our multiperiod portfolio return achieved after step $k$ is given by $S_T^k = e^{Ty(\underline{b}_k)}$, where

$$y(\underline{b}_k) = \frac{1}{T} \sum_{t=1}^{T} \log \underline{b}_k \cdot \underline{x}_t = \frac{1}{T} \sum_{t=1}^{T} \log Z_{t,k}. \tag{17}$$

We can see that during each step of the iterative procedure, we only consider two components, one of which is the combination of stocks which has already been selected (with their previously determined relative weights), the other one is the selected stock from $\{1, \ldots, M\}$. The newly selected stock may be either new or the one which has been already selected. In the latter case, the optimization steps serves to adjust the weight of the selected stocks relative to the others. The wealth factor $S_T^k$ after $k$ steps is close to the maximum $S_T^{max}$ in the sense that it has an exponential drop of order $c^2/(k+3)$ as shown in (2). Consequently with $k$ increasing, our $S_T^k$ converges to the maximum $S_T^{max}$.

As we have now our optimization procedure, the next thing is to know how close the wealth achieved after $k$ steps is to $S_T^{max}$.

## 3.3 PROOF OF THE WEALTH MAXIMIZATION BOUND

We define $D_k$ as the average logarithm ratio of $\underline{b}^{max} \cdot \underline{x}_t$ and $\underline{b}_k \cdot \underline{x}_t$ for $t = 1, \ldots, T$. That is

$$D_k = \frac{1}{T} \sum_{t=1}^{T} \log \frac{\underline{b}^{max} \cdot \underline{x}_t}{\underline{b}_k \cdot \underline{x}_t} = \frac{1}{T} \log \frac{S_T^{max}}{S_T(\underline{b}_k)}. \tag{18}$$

Theorem 1 states a bound on $D_k$ of $c^2/(k+3)$. We prove this theorem through the following lemmas.

**Lemma 1** *Suppose a sequence of nonnegative numbers $D_k$, with $k \geq 1$, satisfies*

$$D_k \leq (1-\alpha)D_{k-1} + \alpha^2 c^2/4 \tag{19}$$

*for all $\alpha \in (0,1)$ and $k \geq 2$, for some $c$ which is independent of $k$. Also suppose $D_1 \leq c^2/4$, then we have for all $k \geq 1$*

$$D_k \leq \frac{c^2}{k+3}. \tag{20}$$

**Proof**: We proceed by induction. First, the bound holds by assumption when $k=1$. Now suppose $D_{k-1} \leq \frac{c^2}{k+2}$ for $k \geq 2$. Then invoking (18) with $\alpha = \frac{2}{k+2}$

$$\begin{aligned} D_k &\leq (1-\frac{2}{k+2})\frac{c^2}{k+2} + \frac{c^2}{(k+2)^2} \\ &= c^2 \cdot \left[\frac{k}{(k+2)^2} + \frac{1}{(k+2)^2}\right] \\ &= \frac{c^2}{k+3} \cdot \frac{k^2+4k+3}{(k+2)^2} \\ &\leq \frac{c^2}{k+3}. \end{aligned}$$

Though the statement of Lemma 1 requires the inequality (20) to hold for all $\alpha \in (0,1)$, we see from the proof that having (20) hold for $\alpha \leq 1/2$ and indeed for the particular choice $\alpha_k = 2/(k+2)$ with $k = 2, 3, \ldots$ is sufficient for the validity of the claim.

To show Theorem 1 from Lemma 1, we prove that $D_k$ defined as in expression (18) indeed satisfies the requirement of inequality (19). Demonstration of this property of $D_k$ is the focus of our remaining efforts in this section.

Here we need some useful inequalities for pairs of nonnegative real numbers.

**Lemma 2** *For all numbers $r, r_0$ with $0 < r_0 \leq r$, we have the inequality*

$$-\log r \leq -(r-1) + [\frac{-\log r_0 + r_0 - 1}{(r_0-1)^2}](r-1)^2. \tag{21}$$

**Lemma 3** *For all $r > 0$, the following inequality holds*

$$2[\frac{-\log r + r - 1}{r-1}] \leq \log r. \tag{22}$$

**Lemma 4** *For all $r > 0$, we have*

$$\frac{-\log r + r - 1}{(r-1)^2} \leq 1/2 + \max(0, -\log r). \tag{23}$$

Here, in accordance with extension by continuity, the expression $\frac{-\log r+(r-1)}{(r-1)^2}$ is taken to be $1/2$ at $r = 1$ and likewise $\frac{-\log r+(r-1)}{r-1}$ is taken to be 0 at $r = 1$.

**Remark** Li and Barron [25] use the same inequalities as above but different setting in their work on mixture density estimation. Li and Barron [25] originated the technique we use here. They showed that for mixture density estimation, a $k$-component mixture estimated by maximum likelihood achieves log-likelihood within order $1/k$ of the log likelihood achievable by any convex combination.

For our analysis, we define $r_t$ as the ratio of our portfolio return to the optimal portfolio return at time $t$ at step $k$ when stock $i$ is introduced, that is,

$$r_t = \frac{(1-\alpha)Z_{t,k-1} + \alpha x_{t,i}}{\underline{b}^{max} \cdot \underline{x}_t}. \tag{24}$$

Also let $r_{0,t} = \frac{(1-\alpha)Z_{t,k}}{\underline{b}^{max} \cdot \underline{x}_t}$ where $0 < r_{0,t} \leq r_t$ and $0 \leq \alpha \leq 1$ and where $Z_{t,k-1}$ and $\underline{b}^{max}$ are as defined in the previous subsection. Now we can start to show our main result. Plug $r_t$ and $r_{0,t}$ into (21) and use (22) to obtain

$$\begin{aligned}
-\log r_t &\leq -(r_t - 1) + [\frac{-\log(r_{0,t}) + r_{0,t} - 1}{(r_{0,t}-1)^2}](r_{0,t} - 1 + \frac{\alpha x_{t,i}}{\underline{b}^{max} \cdot \underline{x}_t})^2 \\
&= -(r_{0,t} - 1 + \frac{\alpha x_{t,i}}{\underline{b}^{max} \cdot \underline{x}_t}) + (-\log r_{0,t} + r_{0,t} - 1) \\
&\quad + (\frac{\alpha x_{t,i}}{\underline{b}^{max} \cdot \underline{x}_t})^2 (\frac{-\log r_{0,t} + r_{0,t} - 1}{(r_{0,} - 1)^2}) + \frac{2\alpha x_{t,i}}{\underline{b}^{max} \cdot \underline{x}_t}(\frac{-\log r_{0,t} + r_{0,t} - 1}{r_{0,t} - 1}) \\
&\leq -\log r_{0,t} - \frac{\alpha x_{t,i}}{\underline{b}^{max} \cdot \underline{x}_t} + (\frac{\alpha x_{t,i}}{\underline{b}^{max} \cdot \underline{x}_t})^2 (\frac{-\log r_{0,t} + r_{0,t} - 1}{(r_{0,t} - 1)^2}) \\
&\quad + \frac{\alpha x_{t,i}}{\underline{b}^{max} \cdot \underline{x}_t} \log r_{0,t}
\end{aligned} \tag{25}$$

This $-\log r_t$ appears in our update rule for $D_k$. Indeed, by the definition of $D_k$ (equation (18)) with $i_k$ is chosen to maximize expression (13). We have that

$$D_k \le \min_i D_{k,i} \tag{26}$$

where

$$D_{k,i} = \frac{1}{T}\sum_{t=1}^{T}\left[-\log\frac{(1-\alpha)Z_{t,k-1}+\alpha x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\right]. \tag{27}$$

This minimum is not less than the weighted average of $D_{k,i}$ for any weights that add to 1 for $i=1,2,\ldots,M$. In particular, the minimum is smaller than the average using $\underline{b}^{max}$. Hence a sequence of inequalities can be given as follows,

$$\begin{aligned}
D_k &\le \sum_{i=1}^{M} b_i^{max} D_{k,i} \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{M} b_i^{max}\Bigg[-\log r_{0,t} - \frac{\alpha x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t} \\
&\quad +\Big(\frac{\alpha x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\Big)^2\Big(\frac{-\log r_{0,t}+r_{0,t}-1}{(r_{0,t}-1)^2}\Big) + \frac{\alpha x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\log r_{0,t}\Bigg] \\
&= -\log(1-\alpha) - \alpha + \alpha\log(1-\alpha) + (1-\alpha)D_{k-1} \\
&\quad +\frac{\alpha^2}{T}\sum_{t=1}^{T}\sum_{i=1}^{M} b_i^{max}\Big(\frac{x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\Big)^2\Big(\frac{-\log r_{0,t}+r_{0,t}-1}{(r_{0,t}-1)^2}\Big) \\
&\le (1-\alpha)D_{k-1} + \frac{\alpha^2}{T}\sum_{t=1}^{T}\sum_{i=1}^{M} b_i^{max}\Big(\frac{x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\Big)^2\Big(\frac{-\log r_{0,t}+r_{0,t}-1}{(r_{0,t}-1)^2}\Big)
\end{aligned}$$

where the last inequality is established by noting that $(\alpha-1)\log(1-\alpha)-\alpha \le 0$ for $\alpha \in [0,1]$.

Write

$$I = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{M} b_i^{max}\Big(\frac{x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\Big)^2 \tag{28}$$

and

$$v = \max_{1\le t\le T, 1\le i,j\le M}\{x_{t,i}/x_{t,j}\}. \tag{29}$$

Let $x_{t,max} = \max_i\{x_{t,i}\}$ and $x_{t,min} = \min_i\{x_{t,i}\}$. Since $\sum_{i=1}^{M} b_i^{max} x_{t,i} \leq x_{t,max}$ and likewise $Z_{t,k-1} \geq x_{t,min}$, we have by Lemma 4 that

$$\begin{aligned} \frac{-\log r_{0,t} + r_{0,t} - 1}{(r_{0,t}-1)^2} &\leq 1/2 + \log^-(r_{0,t}) \\ &= 1/2 + \log^+ \frac{\sum_{i=1}^{M} b_i^{max} x_{t,i}}{(1-\alpha)\sum_{i=1}^{M} b_{i,k-1} x_{t,i}} \\ &\leq 1/2 + \log 2v \\ &= \log 2v\sqrt{e} \end{aligned}$$

Thus the inductive inequality $D_k \leq (1-\alpha)D_{k-1} + \alpha^2 c^2/4$ is obtained, where,

$$c^2 = 4I \log 2v\sqrt{e}. \tag{30}$$

As mentioned before, $D_k$ is the normalized log-wealth ratio, so we conclude that

$$D_k = \frac{1}{T}\log\frac{S_T^{max}}{S_T^k} \leq \frac{c^2}{k+3}. \tag{31}$$

That is,

$$S_T^k \geq S_T^{max} e^{-T\frac{c^2}{k+3}}. \tag{32}$$

So we showed the proof of Theorem 1, which says that $S_T^k$ approximates $S_T^{max}$ with an exponent that is less than the maximum by not more than $c^2/(k+3)$.

Regarding the quantities $I$ and $v$ we note that

$$\begin{aligned} I &= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{M} b_i^{max}\left(\frac{x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\right)^2 \\ &= \frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{i=1}^{M} b_i^{max} x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\left(\frac{x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\right) \\ &\leq \frac{1}{T}\sum_{t=1}^{T}\max_{1\leq i\leq M}\left(\frac{x_{t,i}}{\underline{b}^{max}\cdot\underline{x}_t}\right) \end{aligned} \tag{33}$$

which in turn is not more than $v$. Thus $I$ is the mildest of these volatility expressions depending only on relative return of stocks in the portfolio

$\underline{b}^{max}$ relative to the portfolio return $\underline{b}^{max} \cdot \underline{x}_t$. In contrast the bound (33) depends on the maximum relative return over all stocks (though still relative to $\underline{b}^{max} \cdot \underline{x}_t$) and the coarsest measure of volatility $v$ depends on the worst case ratio overall.

We complete our discussion of the wealth maximization algorithm by noting its computation time. The aim is to find a $\underline{b}$ that achieves a value for $y(\underline{b}) = \frac{1}{T}\sum_{t=1}^{T} \log \underline{b} \cdot \underline{x}_t$ that is within $\varepsilon$ of the maximum. Our algorithm after $k$ steps achieves $|y(\underline{b_k}) - y(\underline{b}^{max})| \leq c^2/k$, that is, $k \geq c^2/\varepsilon$ suffices for the stated aim. During each step, the number of computations equals the number of periods $T$ times the number of candidate stocks $M$. Therefore, the total number of computations needed to reach the accuracy of $\varepsilon$ is in the form of $N_{new}(\varepsilon)$, where

$$N_{new}(\varepsilon) = \frac{c^2 MT}{\varepsilon} \tag{34}$$

When compared with the time bound (9) for the interior-point method, we see that while (9) has the advantage of logarithmic dependence on $1/\varepsilon$. Our algorithm is superior for large $M$. Also our algorithm time depends specifically on the volatility quantity $c$. Our procedure is better when the number of stock $M$ is large.

# 4 MIXTURE PORTFOLIO ANALYSIS

We have seen that Cover's universal portfolio analysis requires the number of stocks $M$ to be small compared to investment periods $T$ in order to insure that the wealth exponent achieved is close to $y^{max}$. However, when dealing with thousands of stocks and hundreds of time periods, $M$ is large compared to $T$, but $\log M$ is reasonably small compared to $T$. Then our mixture portfolio strategy provides a better bound.

The Cover [8] and Cover & Ordentlich [9] strategies use a sequences of Bayes-like portfolio update rules.

$$\hat{\underline{b}}_t = \frac{\int \underline{b} S_{t-1}(\underline{b})\pi(\underline{b})d\underline{b}}{\int S_{t-1(}\underline{b})\pi(\underline{b})d\underline{b}} \tag{35}$$

where $\pi(\underline{b})d\underline{b}$ is a density function on the $(M-1)$-dimensional simplex of all portfolio $\underline{b}$.

A remarkable property they show for these sequences of portfolios is that the associated actual wealth $S_T^{univ} = \prod_{t=1}^T \hat{\underline{b}}_t \cdot \underline{x}_t$ also equals

$$\int S_T(\underline{b})\pi(\underline{b})d\underline{b} \tag{36}$$

Equivalently, one may think of $\pi$ with $\int \pi(\underline{b})d\underline{b} = 1$ as providing an initial distribution of wealth across portfolios for which we subsequently accumulate the total wealth $\int S_T(\underline{b})\pi(\underline{b})d\underline{b}$.

Now $S_T(\underline{b})$ is a highly peaked function of $\underline{b}$. Laplace-type approximation is then used by them in [8] and [9] (with refinement in [36]) to show that to first order this integral is determined by $S_T^{max}$ and to quantify the drop from that maximum associated with the weighted average. The refinements in [9] and [36] show that for suitable $\pi$ (close to Dirichlet$(1/2, \ldots, 1/2)$) the universal portfolio achieves

$$S_T^{univ} \geq e^{T[y^{max} - \frac{M-1}{2T}\log(c_M T)]}. \tag{37}$$

Our analysis is related in that we also use a mixture based portfolio. The difference is that our $\pi$ is discrete with points on faces of the simplex determined by subsets of stocks. In this section, we bound our mixture strategy performance.

As we have written in Section 2, we allocate our wealth across all subgroups of size $k$, which are constructed by selecting $k$ stocks $(j_1, \ldots, j_k)$. There will be a total of $M^k$ subgroups with size $k$ when repeating is allowed. Let $S_T^{(j_1,\ldots,j_k)}$ be the wealth associated with the choice $(j_1, \ldots, j_k)$ where the portfolio is determined by the following weight vector

$$\underline{\alpha}_k = \left(\frac{6}{(k+1)(k+2)}, \frac{6}{(k+1)(k+2)}, \frac{8}{(k+1)(k+2)}, \ldots, \frac{2}{k+2}\right). \tag{38}$$

Here the coordinates follow the pattern $\alpha_{i,k} = \frac{2(i+1)}{(k+1)(k+2)}$ for $i = 2, \ldots, k-1$. These are the weights that arise by initializing $\alpha_{1,1} = 1$ and then for $k > 1$ obtaining $\underline{\alpha}_k$ by multiplying $\underline{\alpha}_{k-1}$ by $\frac{k}{k+2}$ and setting the new coordinate to $\frac{2}{k+2}$. For example for $k = 5$ this $\underline{\alpha}_k$ is $(\frac{3}{21}, \frac{3}{21}, \frac{4}{21}, \frac{5}{21}, \frac{6}{21})$. These weights give more attention to recent iterations than earlier ones. [In contrast if $\underline{\alpha}_k$ were formed by multiplying $\underline{\alpha}_{k-1}$ by $\frac{k-1}{k}$ and setting the new coordinates to $\frac{1}{k}$ then all $k$ coordinates of $\underline{\alpha}_k$ would be equal.]

We create two types of mixtures, one in which $k$ is prespecified and another in which we mix across $k$. For the prespecified $k$, we give equal initial allocation $1/M^k$ on each subgroup. Thus after $T$ investment periods, we will have the wealth $S_T^{mix,k}$, where

$$S_T^{mix,k} = \frac{1}{M^k} \sum_{(j_1,\ldots,j_k)\in\{1,\ldots,M\}} S_T^{(j_1,\ldots,j_k)}. \tag{39}$$

To explain the idea of this mixture we contrast it with $S_T^k = S_T^{(j_1^*,\cdots,j_k^*)}$ where $(j_1^*,\cdots,j_k^*)$ are the indices of the stocks selected the first $k$ steps of the greedy algorithm. That choice of $j_1^*,\cdots,j_k^*$ depends on the entire return sequence $\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_T$. The mixture overcomes the lack of advanced knowledge of which choice will perform well by giving some weight $(1/M^k)$ to every $(j_1,\ldots,j_k)$. Accordingly we find that

$$S_T^{mix,k} \geq \frac{1}{M^k} S_T^k. \tag{40}$$

**Proof of Theorem** 2: First we are to show that there is a $k$ such that

$$\begin{aligned} S_T^{mix,k} &\geq S_T^{max} e^{-T2c\sqrt{\frac{\log M}{T}}} \\ &= e^{T[y^{max}-2c\sqrt{\frac{\log M}{T}}]}. \end{aligned}$$

For any $k \geq 1$, among the subgroups we mix across it will happen that one of them will be the particular one $(j_1^*,\ldots,j_k^*)$ that arises by our greedy algorithm in Theorem 1. Then invoking our bound on its wealth we have

$$\begin{aligned} S_T^{mix,k} &= \frac{1}{M^k} \sum_{(j_1,\ldots,j_k)\in\{1,\ldots,M\}} S_T^{(j_1,\ldots,j_k)} \\ &\geq \frac{1}{M^k} \max_{(j_1,\ldots,j_k)} S_T^{(j_1,\ldots,j_k)} \\ &\geq S_T^{(j_1^*,\ldots,j_k^*)} \frac{1}{M^k} \\ &\geq S_T^{max} e^{-T\left[\frac{c^2}{k+3}+\frac{k\log M}{T}\right]}. \end{aligned} \tag{41}$$

Now among real-valued $k$ the best such bound would be at $k^* = c\sqrt{\frac{T}{\log M}} - 3$ which would yield

$$S_T^{mix,k} \geq M^3 S_T^{max} e^{-T2c\sqrt{\frac{\log M}{T}}}. \tag{42}$$

This is even better than claimed by the factor $M^3$. When that $k^*$ is not an integer we take the integer $k = \tilde{k}^*$ between $c\sqrt{\frac{T}{\log M}} - 3$ and $c\sqrt{\frac{T}{\log M}} - 2$ and achieve

$$S_T^{mix,k} \geq M^2 S_T^{max} e^{-T2c\sqrt{\frac{\log M}{T}}}. \tag{43}$$

When prior knowledge of $c$ is not available, we may use $k = \sqrt{T/\log M} - 3$ (rounding up to an integer if need be), which also results the same order bound. That is,

$$\begin{aligned} S_T^{mix,k} &\geq M^2 S_T^{max} e^{-T(c^2+1)\sqrt{\frac{\log M}{T}}} \\ &= M^2 e^{T\left[y^{max} - (c^2+1)\sqrt{\frac{\log M}{T}}\right]} \end{aligned}$$

Rather than using a fixed size $k$, when $c$ is unknown, it is preferrable to mix across choices of $k$. We may distribute our initial wealth according to a prior $q(k)$ on the subcollection size $k$ for $k = 1, \ldots, M$. This mixing procedure leads to wealth $S_T^{mix}$ after $T$ investment periods, where

$$S_T^{mix} = \sum_{k=1}^{M} q(k) S_{T,k}^{mix}. \tag{44}$$

We can take a uniform prior $q(k) = \frac{1}{M}$ yielding

$$S_T^{mix} = \sum_{k=1}^{M} \frac{1}{M^{k+1}} \sum_{j_1,\ldots,j_k \in \{1,\ldots,M\}} S_T^{(j_1,\ldots,j_k)} \tag{45}$$

As shown in Section 2.1, we have the inequality $\frac{1}{T}\log\frac{S_T^{max}}{S_T^k} \leq \frac{c^2}{k+3}$. Thus

$$\frac{1}{T}\log\frac{S_T^{max}}{S_T^k} \leq \frac{c^2}{k+3}. \tag{46}$$

Equivalently, $S_T^k \geq S_T^{max} e^{-T\frac{c^2}{k+3}}$. Hence using the same argument as above we have,

$$S_T^{mix} \geq \sum_{k=1}^{M} \frac{1}{M^{k+1}} S_T^{max} e^{-T\frac{c^2}{k+3}}$$

$$
\begin{aligned}
&= \sum_{k=1}^{M} \frac{S_T^{max}}{M} e^{-T[\frac{c^2}{k+3} + \frac{k \log M}{T}]} \\
&\geq \frac{S_T^{max}}{M} e^{-T[\frac{c^2}{\tilde{k}^*+3} + \frac{\tilde{k}^* \log M}{T}]} \\
&\geq M S_T^{max} e^{-T \cdot 2c\sqrt{\frac{\log M}{T}}} \\
&= M e^{T\left[y^{max} - 2c\sqrt{\frac{\log M}{T}}\right]}.
\end{aligned} \tag{47}
$$

This complete our proof of Theorem 2.

In the next section, we apply both our wealth maximization algorithm and the mixture strategy to real stock data from Wharton Research Data Services.

# 5 EXPERIMENTS WITH REAL MARKET DATA

In this section, we conduct several experiments with real stock data to examine historical stock performance with our wealth maximizing algorithm and to test the performance of our mixture portfolio strategy and some other practical strategies.

## 5.1 MIXTURE STRATEGY EXPERIMENTS

First for illustrative purpose, we examine investment in Standard and Poor 500 stocks. The stock price information is from the Wharton WRDS online data base. First we examine the consequence of following our mixture strategy over 10 year period from January 1996 through December 2005. From Figure 1, we can see that our mixture portfolio multiplies wealth by the factor $S_T^{mix} = 6.10$ (e.g. a \$1000 initial investment would have become \$6,100). Here we are rebalancing monthly the portfolios of subsets of these stocks and aggregates them together at the end. The mixture we form (with $k = 3$) totals the wealth from all $500^3$ subsets with repeats allowed. [Thus each of the subsets is given initial weight $\frac{1}{500^3}$]. For this mixture strategy he wealth factor 6.10 for these ten years coincides with an compounding annual return of about 19.82%. In contrast the S&P 500 index

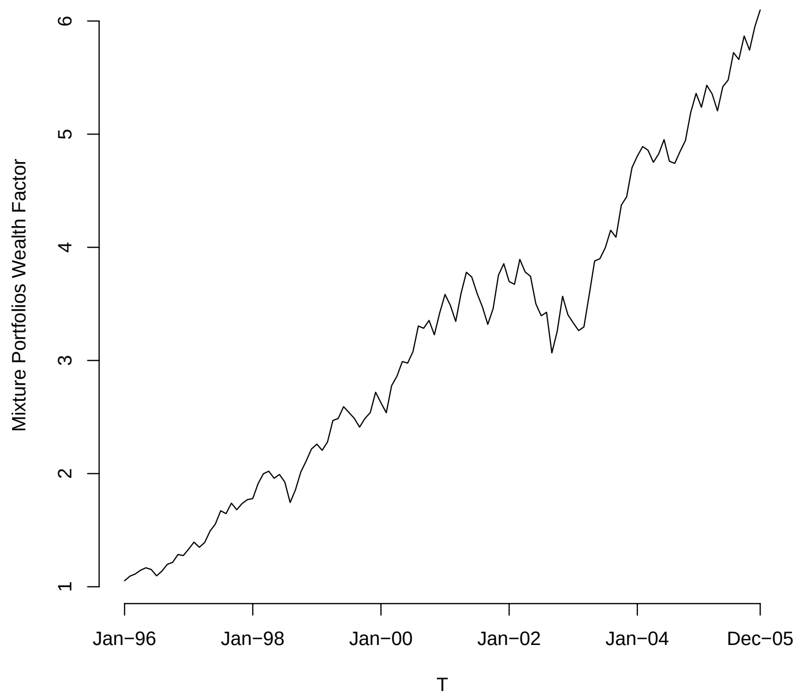


Figure 1: Performance of The Mixture Portfolio Strategy $S_T^{mix}$ on S&P500 for $T = 1$ to 120 months

increased by a factor of 1.96 (equivalent to compounding annual return of 6.9%). Using the weights specified by our mixture strategy the best subset stocks among all are *Dell Inc.*, *Jabil Circuit Inc.* and *Qlogic Corp.* which with weights $(0.3, 0.3, 0.4)$ had wealth factor $S_T = 125.1$. Now a wealth factor of 125 though large is swamped by the division by $500^3$, so in this case the mixture strategy is making its reasonable wealth factor 6.10 by aggregating the wealth from many good triplets of stocks.

So that the mixture strategy represents an achievable wealth we needed to incorporate a rule for handing stocks (among the S&P 500) that are acquired or otherwise replaced from the S&P (but still traded). Our rule was to continue to follow the 500 stocks that comprised the S&P 500 in January

1996. When a stock $i$ is delisted at month $t_0$, we convert it to cash using the price after delisting reported in the WRDS database (and thereafter use a cash wealth factor of $x_{t,i} = 1$ for $t > t_0$). When a stock $i$ is acquired at some month $t_0$ we convert it to stocks in the acquiring company (and thereafter use the return of the acquiring company in place of stock $i$, where if the acquiring company was already on of the initial 500, we give it an index $i'$, then its return are used for both $x_{i,t}$ and $x_{i',t}$ for $t > t_0$).

Over the same 10 year time period for stocks listed in *NASDAQ-100 Index* (We are using all stocks that have been included in the index. For those delisted, we use delisting price and replace them with new ones), the mixture strategy achieves a wealth factor of 9.9 corresponding to a compounding annual return of 25%.

Next we examine investment in Standard and Poor 500 stocks over the 26 years period from January 1970 through December 1995 with rebalancing monthly among subsets of these stocks. This time frame is chosen for comparison with result of Goetzmann (1996). Our mixture portfolio multiplies wealth by a factor of 34.4, which is equivalent to an annual compounding rate of return of 14.7%. Goetzmann chose hisportfolio according to a mean-variance criterion. He reports a 12% annual return, which is the average return for these periods instead of the actual compounding rate. As we know, the arithmetic mean is always greater than the geometric mean. That is,

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} \underline{b} \cdot \underline{x}_t &\geq (\prod_{t=1}^{T} \underline{b} \cdot \underline{x}_t)^{\frac{1}{T}} \\ &= e^{y(\underline{b})} \\ &= 1 + r(\underline{b}). \end{aligned}$$

Thus the actual compounding rate of return is less than the empirical average return. In particular, for Goetzmann's strategy the actual compounding rate of return which is not reported must be not more than 12% per year whereas our rate of 14.7% is higher. In our work we use only those stocks listed in S&P 500 for the entire period from January 1970 through December 1995. It is not clear how Goetzmann handles those stocks delisted from S&P 500 during this period.

A critic might complain that maximization of $y(\underline{b}) = \frac{1}{T}\sum_{t=1}^{T}\log(\underline{b}\cdot\underline{x}_t)$ does not have a variance constraint. Nevertheless we hasten to point out that the average logarithm (as arise in maximization of $S_T(\underline{b})$) is a more risk adverse criterion especially for $\underline{b}\cdot\underline{x}_t$ near 0 than quadratic utilities. Such risk adversion in $y(\underline{b})$ is necessary for identifying the highest attainable rate of growth for constant rebalanced portfolios.

Also we point out that the 14.7% return is the return attained by the mixture strategy which can be regarded as updated each period based only on preceding performance. In contrast the 12% average return given by Goetzmann is based on hindsight for the whole 26 year period. We turn attention next to what would be the best growth rate with hindsight.

## 5.2 MAXIMUM WEALTH INDEX CALCULATION

Now we compute the maximum constant rebalanced portfolio wealth factor $S_T^{max} = \max_{\underline{b}} S_T(\underline{b})$ for the ten year period from January 1996 through December 2005, rebalancing monthly (that is for $T = 120$ months). We first use as the pool of stocks those that have been included as Standard and Poor 500 stocks with their monthly return as reported by the Wharton data base. We find that with 3 or 4 stocks the greedy algorithm comes reasonably close to the maximum.

Indeed from Table 1, we see that the algorithm only needs $k = 4$ with three stocks to achieve a wealth factor $S_T^k$ of 1606.96. Further optimization, for instance to $k = 16$ steps reaches a factor $S_T^k$ of 1618.39 with only five stocks. These stocks are *Biogen Idec Inc. (BIIB), Citrix Systems Inc. (CTXS), Network Appliance Inc. (NTAP), Apoppo Group Inc. (APOL)* and *Dell Inc. (DELL)*. In this implementation for $k > 1$ we allowed $\alpha_k$ to be freely adjusted between 0 and 1/2 (here we used a fine grid of spacing 1/10000) rather than fixed at $\alpha_k = \frac{2}{k+2}$ (either way is permitted by our theory). Each step tries every stock for possible new inclusion or tunes an existing stock weight (relative to the others), whichever is best. For these data, the algorithm found no advantage after step 15 for including any additional stocks beyond the indicated five. Thus confirming that we were already very close to the maximum with a handful of steps.

| k | Stocks in Portfolio at Each Algorithm Step | Wealth Factor $S_T^k$ |
|---|---|---|
| 1 | BIIB | 155.97 |
| 2 | BIIB, CTXS | 942.90 |
| 3 | BIIB, CTXS, NTAP | 1585.51 |
| 4 | BIIB, CTXS, NTAP | 1606.96 |
| 5 | BIIB, CTXS, NTAP | 1613.13 |
| 6 | BIIB, CTXS, NTAP | 1615.14 |
| 7 | BIIB, CTXS, NTAP, APOL | 1616.54 |
| 8 | BIIB, CTXS, NTAP, APOL | 1617.17 |
| .<br>.<br>. | ...<br>...<br>... | .<br>.<br>. |
| 14 | BIIB, CTXS, NTAP, APOL | 1618.31 |
| 15 | BIIB, CTXS, NTAP, APOL, DELL | 1618.38 |
| 16 | BIIB, CTXS, NTAP, APOL, DELL | 1618.39 |
| .<br>.<br>. | ...<br>...<br>... | .<br>.<br>. |
| 1000 | BIIB, CTXS, NTAP, APOL, DELL | 1618.41 |

Table 1: Wealth Maximization Algorithm

The volatility quantities $I$ and $\log(v)$ in this This wealth growth of $S_T^{max} = 1618.4$ over the $T = 120$ months reflects a monthly wealth exponent of $y^{max} = \frac{1}{120} \log 1618.4 = 0.0615$ or equivalently an annualized return of $(1618.4)^{\frac{1}{10}} = 2.09$, that is, 109% growth per year. Though surprisingly high, we must emphasize that the tradeoffs required for achievable wealth are sobering, even in this example. Indeed consider the mixture wealth bound from Theorem 2. With $M = 500$ stock, $\sqrt{(\log M)/T} = \sqrt{0.05} = 0.22$ so even if $c^2$ were near $I = 1.04$ the bound on the drop in exponent swallows the otherwise spectacular gain of $y^{max}$. Over the same 10 year time period when maximizing on stocks listed in *NASDAQ-100 Index* the best constant rebalanced portfolio achieved a wealth factor of 2342.8 (with $I = 1.11$ and $\log(v) = 7.58$). This corresponds to a compounding annual return of 217%.

We also computed the maximum constant rebalanced portfolio wealth with monthly rebalancing for the twenty year period from January 1986 through December 2005 for Standard and Poor 500 stocks. At step $k = 4$ it

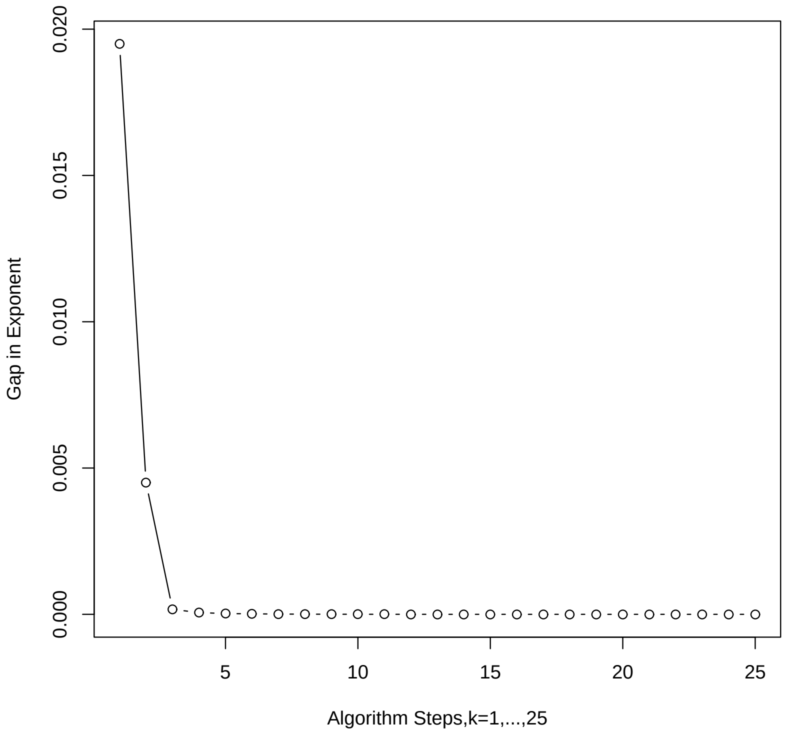


Figure 2: Gap in exponent between $S_T^{max}$ and $S_T^k$ for $k = 1, \ldots, 25$

uses the fours stocks, *Apple Inc.*(*AAPL*), *Countrywide Financial Corp. (C-FC)*, *Roclwell Automation.* (*ROK*) and *RadioShack Corp. (RSH)*, to achieve a wealth factor $S_T^k$ of 1667.09. There are five more stocks, *MBNA Corp. (KR-B)*, *Safeway Inc. (SWY)*, *Synovus Financial Corp. (SNV)*, *American Bankers Ins Group Inc. (ABI)*, and *Keyspan Energy Corp.(KSE)*, when further optimized to 21 steps and the wealth factor after 1000 steps is 1829.52 corresponding to a compounding annual return of 46%. The corresponding volatility quantities $I$ and $\log(v)$ are 1.008 and 2.18 respectively.

Finally, we consider the period from January 1970 through December 1995 and Standard and Poor 500 stocks as in Goetzmann (1996). With monthly rebalancing the best constant rebalanced portfolios with hindsight achieved a multiperiod compounding wealth factor of 561.8 running the algorithm up to

500 steps. The corresponding annual compounding rate of return is 27.5%. Here we also exclude those stocks that were delisted during this period. At step $k = 7$, the portfolio has a wealth factor of 561.08 and uses the five stocks, *Mylan Labs Inc. (MYLN)*, *Southwest Airlines Co. (LUV)*, *St Jude Medical Inc. (STJM)*, *Home Depot Inc. (HD)* and *Circuit City Stores Inc. (CC)*.

## 5.3 MOVING-WINDOW GREEDY UPDATING VERSUS THE MIXTURE STRATEGY

Here we report a strategy of greedy portfolio selection using what may be called moving window information. In this strategy at each trading time (e.g. at the end of each month), the portfolio we set for the next period (e.g. the next month) is the portfolio which, on the previous $T_w$ time periods, would have made the most wealth as computed by our greedy algorithm. We shift this time window (for training the next portfolio) each period so that it reflects the same window length $T_w$ of past return. Our theory gives no guarantee that the next period behavior is predicted best by the preceding time window, nonetheless, it is of interest to see how such a greedy strategy would have performed. Our wealth maximization algorithm is the key ingredient for computation of this moving-window algorithm.

Here we again use the Standard and Poor 500 Index stocks. For each month from January 1996 through December 2005, we consider a moving training window of preceding returns with which we determine the portfolio to use for the current month. For the length of the training window, $12, 18, 24$ and 30 months were tried. The best results as reported here were based on a 24 month window. There is also the issue of the rapidity of rebalancing for the portfolio wealth function that we maximize over the preceding years. For instance, it is unclear whether it is better to use the stock fractions that are optimal with daily rebalancing or with monthly rebalancing. For ease of computation we report results in which we tried monthly rebalancing on each training window here.

Thus at the start of each month we get the portfolio which would (with monthly rebalancing) have made the most over the preceding two years and set that to be our portfolio for the start of that month. It is then updated

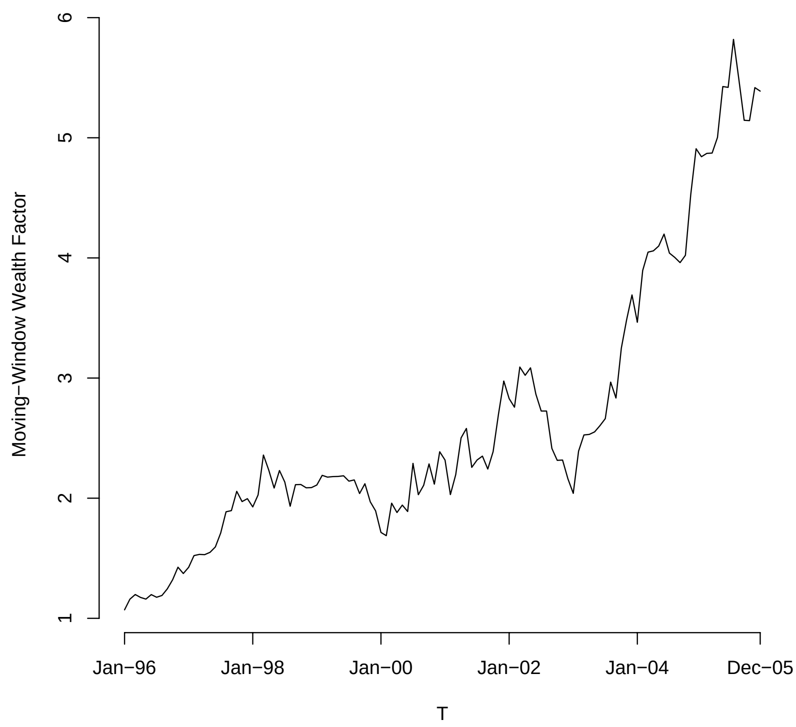


Figure 3: Wealth factor for moving-window portfolio updates

(with a moving two years window) at the start of the following month. Over the January 1996 to December 2005 time frame, the Standard and Poor 500 index had a wealth factor of 2.58 (or an annual return of 9%). During these 10 years our portfolio has a wealth factor over 5.39 (or an annual return of 18%) as shown in Figure 3. Each month it used a small handful of stocks that evolved across time.

Another result shows a twenty-year result in which our moving-window portfolio achieves a wealth factor of 23.43 (or an annual return of 17%) during this period from January 1986 through December 2005 with a training window of 24 months for Standard and Poor 500 stocks while the S&P 500 index gained a factor of 5.89 (or a annual return of 9%) during that period.

We also applied this strategy to the NASDAQ 100 stocks over the January 1997 through December 2005 time frame with daily rebalacing on the training window. We tried $6, 9, 12, 15, 18, 21$ and 24 months as the length of training window. The best one is with 12 months window. It shows an impressive result that during these 9 yeas our portfolio has a wealth factor of 58.5 while the NASDAQ 100 index gained a factor of 2.14 in the same period. For a 24 months training window the wealth factor for these 9 years is about 32.5. It is still better than the result from Standard and Poor stocks. The NASDAQ companies are relative younger and faster growing companies with smaller capitalization compared with Standard and Poor companies, which might be one of the reasons for this phenomenon.

Thus the moving window greedy algorithm has performed quite well on recent historical data and should be given serious attention as an investment strategy. However, we caution that the greedy approach can be fooled. Sudden changes in best portfolios can lead to a situation in which a portfolio trained to be best on the past is miserable in the future (compared to appropriate targets).

As our Theorem 2 shows, the mixture strategy advocated in this paper does provides a performance guarantee. It is provably close in exponent to the best constant rebalanced portfolio (however high or low that might be) provided $c\sqrt{(\log M)/T}$ is small compared to that best exponent.

## 5.4 WEALTH TARGETS WITH PAST DEPENDENT PORTFOLIOS

In this paper we have presented the algorithm in the context of optimization of constant rebalanced portfolios, where we find $\underline{b}$ such that $\Pi_{t=1}^{T} \underline{b} \cdot \underline{x}_t$ is maximized where each $x_{t,i}$ is the return of an available asset. However, it is possible to incorporate interesting types of past dependence in this framework. The idea is to allow parameterized dependence of the portfolio on past returns so as to capture the possibility of putting higher or lower attention on stocks that went up the previous period. Cross & Barron [11] introduce such past dependent portfolio in a universal portfolio setting. As in [11], if the dependence of the portfolios on past returns is linear in portfolio weights then our theory readily adapts to this setting.

In particular consider vectors such as

$$\underline{\tilde{x}}_t^+ = (\frac{(x_{t,1}-1)_+}{s_t^+}, \ldots, \frac{(x_{t,M}-1)_+}{s_t^+}) \tag{48}$$

and

$$\underline{\tilde{x}}_t^- = (\frac{(x_{t,1}-1)_-}{s_t^-}, \ldots, \frac{(x_{t,M}-1)_-}{s_t^-}) \tag{49}$$

which are non-negative, sum to 1, and depend on $\underline{x}_t$. Here $s_t^+ = \sum_{i=1}^M (x_{t,i}-1)_+$ and $s_t^- = \sum_{i=1}^M (x_{t,i}-1)_-$. Then for period $t$, the above vectors $\underline{\tilde{x}}_{t-1}^+$ or $\underline{\tilde{x}}_{t-1}$ computed from the preceding period may be thought of as providing portfolios for new auxiliary assets $x_{t,M+1} = \underline{\tilde{x}}_{t-1}^+ \cdot \underline{x}_t$ and $x_{t,M+2} = \underline{\tilde{x}}_{t-1}^- \cdot \underline{x}_t$.

Now a portfolio $\underline{b} \cdot \underline{x}_t = \sum_{i=1}^{M+2} b_i x_{t,i}$ in all the assets (including the two newly created) may be regarded as investing in each stock $i$ a fraction $b_i + b_{M+1}\tilde{x}_{t-1,i}^+ + b_{M+2}\tilde{x}_{t-1,i}^-$ which depends on the past and is indeed linear in the weight $\underline{b}$. Including this freedom for past dependence (captured through the auxiliary assets) the wealth target $\max_{\underline{b}} \Pi_{t=1}^T \underline{b} \cdot \underline{x}_t$ is now higher than before. It is still available for computation by our wealth maximization algorithm and for construction of mixture-based or moving window portfolio updates.

For example, if each month we put weight 1/2 on $\underline{\tilde{x}}_{t-1}^-$ and weight 1/2 on the moving-window greedy portfolio updates described before, then the nine-year *NASDAQ 100* wealth factor ending December 2005 increases from 58.5 to 116.3.

# References


[1] Algoet, P. and Cover, T.M. (1988): Asymptotic Optimality and Asymptotic Equipartition Property of Log-optimal Investment, *Annals of Prob.*, 16(2), 876 − 898.

[2] Bell, R. and Cover, T.M. (1980): Competitive Optimality of Logarithmic Investment, *Mathematics of Operations Research*, 5(2), 161 − 166.

[3] BELL, R. and COVER, T.M. (1988): Game-theoretic Optimal Portfolios, *Management Science*, 34(6), $724-733$.

[4] BREIMAN, L. (1961): Optimal Gambling Systems For Favorable Games, *Fourth Berkeley Symposium on Prob. and Stat.,* 1, $65-78$.

[5] BARRON, A.R. and COVER, T.M. (1988): A Bound on the Financial Value of Information, *IEEE Trans. Info. Theory* 34(5), $1097-1100$.

[6] BERNSTEIN, W.J. and WILKINSON, D. (1997): Diversification, Rebalancing and the Geometric Mean Frontier, *Online Journal of Efficient Frontier.*

[7] BOYD, S. and L. VANDENBERGHE (2004): *Convex Optimization*, Cambridge University Press.

[8] COVER, T.M. (1991): Univsersal Portfolios, *Math. Finance* 1(1),$1-29$.

[9] COVER, T.M. and E. ORDENTLICH (1996): Universal Portfolio wilth Side Information, *IEEE Trans. Info. Theory*, 42(2), $348-363$.

[10] CROSS, J. and A.R. BARRON (2003): Efficient Universal Portfolios For Past-dependent Target Classes, *Math. Finance*, 13(2), $245-276$.

[11] CHAN, L., N. JEGADEESH and J. LAKONISHOK (1996): Momentum Strategies, *Journal of Finance* 51(5), $1681-1713$.

[12] ELTON, E.J and MARTIN, J.G. (1974), On the Optimality of Some Multiperiod Selection Criteria, *Journal of Business*, 47(2), $231-243$.

[13] FIACCO, A.V. and MCCORMICK, G.P. (1968), *Nonlinear Programming: Sequential Unconstraned Minimization Techniques*, Wiley and Sons.

[14] FISHBURN, P.C. (1970), *Utility for Decision Making*, Krieger Publishing Co.

[15] GOETZMANN, W.N. (Online Book): *An Introduction to Investment Theory*, Yale School of Management.

[16] HELMBOLD, D.P., R.E. SCHAPIRE, Y. SINGER and M.K. WARMUTH (1998): On-line Portfolio Selection Uing Multiplicative Updates, *Math. Finance*, 8(4), $325-347$.

[17] HAKANSSON, N.H. (1971), Capital Growth and the Mean-Variance Approach to Portfolio Selection, *Journal of Financial and Quantitative Analysis*, 6(1), $517 - 558$.

[18] HAKANSSON, N.H. (1979), A Characterization of Optimal Multiperiod Portfolio Policies, *TIMS Studies in Management Sciences*, 11, $169 - 177$.

[19] ISHIJIMA H. (2001): Nummerical Methods for Universal Portfolios, *Quantitative Methods in Finance Conference*.

[20] JEGADEESH, N. (1990): Evidence of Predictable Behavior of security Returns, *Journal of Finance* 45(3), $881 - 898$.

[21] JEDADEESH, N. and S. TITMAN (1993): Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance* 48(1), $65 - 91$.

[22] KARMARKAR, N. (1984): *A New Polynomial-Time Algorithm for Linear Programming*, Combinatorica 4, $373 - 395$.

[23] KREPS, D.M. (1988), *Notes on Theory of Choice*, Westview Press.

[24] KELLY, J.L. (1956), A New Interpretaion of Information Rate, *IEEE Trans. Info. Theory*, 2(3), $185 - 189$.

[25] LI J. and A.R. BARRON (2000): Mixture Density Estimation, *Advances in Neural Information Processing Systems*, 12, $279 - 285$, MIT Press.

[26] LEHMANN, B. (1990): Fads, Martingales, and Market Efficiency, *Quarterly Journal of Economics* 60(1), $1 - 28$.

[27] LUENBERGER, D.G. (1997): *Investment Science*, Oxford University Press.

[28] LATANE, H.A. (1959), Criteria for Choice Among Risky Ventures, *Journal of Political Economy*, 67(2), $144 - 155$.

[29] LATANE, H.A. and TUTTLE, D.L. (1967), Criteria for Portfolio Building, *Journal of Finance*, 22(3), $359 - 373$

[30] MARKOWITZ, H.M. (1952): Portfolio Selection, *Journal of Finance*, 7(1), $77 - 91$.

[31] Macbeth, J.D. (1995), What's the Long-term Expected Return to Your Portfolio?, *Financial Analyst Journal*, 51(5), $6-8$.

[32] Nestrov, Y. and A. Nemirovsky (1993): Interior-Point Polinomial Algorithms in Convesx Programming, *SIAM Studies in Applied Mathematics.*

[33] Neumann, J.V. and Morgenstern, O. (1944), *Theory of Games and Economic Behavior*, Princeton University Press.

[34] Neumann, J.V. and Morgenstern, O. (1947), *Theory of Games and Economic Behavior 2nd edition*, Princeton University Press.

[35] Sharpe, W.F. (1985): *Investment*, Prentice-Hall, Englewood Cliffs.

[36] Xie, Q. and A.R. Barron (2000): Asymptotic Minimax Regret for Data Compression, Gambling and Prediction, *IEEE Trans. Info. Theory*, 46(2) $431-445$.

[37] Yudin, D. and A. Nemirovsky (1976): Evaluation of the Information Complexity of Mathematical Programming Problems, (in Russian), *Ekonomikai Matematicheskie Metody* 12, $128-142$, (English translation: *Matekon* 13(2), $3-45$).

[38] Yudin, D. and A. Nemirovsky (1983): *Problem Complxity and Method Efficiency in Optimization*, New York: Willey.

[39] Yu, J. (2003), *Maximal Compounded Wealth for Portfolios of Stocks and Options*, Ph.D. Dissertation Prospectus, Department of Statistics, Yale University.