Portfolio optimization with linear and fixed transaction costs

Miguel Sousa Lobo[1]   Maryam Fazel[2]   Stephen Boyd[3]

December 25, 2003

[1]Duke University, Fuqua School of Business, `mlobo@duke.edu`

[2]Contact author. California Institute of Technology, Control and Dynamical Systems Department, `maryam@cds.caltech.edu`

[3]Stanford University, Information Systems Lab., `boyd@stanford.edu`

Portfolio optimization with linear and fixed transaction costs

Abstract

We consider the problem of portfolio selection, with transaction costs and constraints on exposure to risk. Linear transaction costs, bounds on the variance of the return, and bounds on different shortfall probabilities are efficiently handled by convex optimization methods. For such problems, the globally optimal portfolio can be computed very rapidly. Portfolio optimization problems with transaction costs that include a fixed fee, or discount breakpoints, cannot be directly solved by convex optimization. We describe a relaxation method which yields an easily computable upper bound via convex optimization. We also describe a heuristic method for finding a suboptimal portfolio, which is based on solving a small number of convex optimization problems (and hence can be done efficiently). Thus, we produce a suboptimal solution, and also an upper bound on the optimal solution. Numerical experiments suggest that for practical problems the gap between the two is small, even for large problems involving hundreds of assets. The same approach can be used for related problems, such as that of tracking an index with a portfolio consisting of a small number of assets.

1   Introduction

This paper deals with the problem of single-period portfolio optimization. We consider the maximization of expected return, taking transaction costs into account, and subject to different types of constraints on the feasible portfolios.

Our approach is based on the fact that *convex* optimization problems, even if nonlinear or of large-scale, can be numerically solved with great efficiency, using recently developed algorithms. We show that a number of portfolio optimization problems can be cast as convex optimization problems, and hence globally, and efficiently, solved. This class of convex portfolio optimization problems includes those with linear transactions costs, margin and diversification constraints, and limits on variance and on shortfall risk.

We also consider problems with fixed transaction costs (possibly in addition to linear transaction costs). These nonconvex portfolio optimization problems cannot be solved directly via convex optimization, but we describe two approaches that are based on convex optimization. These problems can be solved exactly (*i.e.*, globally) by solving a number of convex problems which, unfortunately, grows exponentially with the number of assets. This method, as well as other more sophisticated methods of global optimization, is practical only for portfolios with about fifteen or fewer assets.

Our main contribution is to describe a method for solving *approximately* much larger nonconvex portfolio optimization problems, by solving a small number of convex optimization problems. The method yields a possibly suboptimal portfolio, as well as a guaranteed upper bound on the global optimum. While there is no guarantee that the gap between the performance of the suboptimal portfolio and the upper bound will always be small, we find that in practice it is. Our method therefore gives an effective practical solution to nonconvex portfolio optimization problems, even with portfolio constraints and hundreds of assets. (If higher guaranteed accuracy is needed, our method can be embedded in a branch and bound algorithm.)

1.1   Related work

Broadly speaking, our approach falls in the Markowitz framework, where a tradeoff between return mean and variance is present. The genesis of the field has been attributed to Markowitz [Mar52, Mar59] and Roy [Roy52]. Implications for the valuation of assets arose with the capital asset pricing model (CAPM) of Sharpe [Sha64] and Lintner [Lin65]. Recent general references are, *e.g.*, Rudolf [Rud94], and Luenberger [Lue98]. The book from Leibowitz et al. [LBK96] is one of many sources for the downside-risk approach, which has been increasingly used in recent years (although already described in Roy's 1952 paper.)

For fixed transaction costs, solutions have been found for specific structures of the covariance matrix. Blog *et al.* [BHKT83] describe a solution for a single factor model, *i.e.*, a diagonal plus rank-one covariance matrix. Patel and Subrahmanyam [PS82]) assume an even more specific structure, namely that there

is an identical correlation coefficient between all assets and the single factor. In contrast, we make no assumptions about the correlation matrix, and moreover, allow the addition of any other (convex) cost terms and constraints. Bertsimas, Darnell and Soucy [BDS99] use generic mixed-integer programming methods to deal with fixed costs and other integer constraints in several practical cases.

Many treatments have been presented for problems with linear costs. Most methods described in the literature are modifications of the simplex method for quadratic programming, which can handle a quadratic objective but not quadratic constraints. The variance is included in the program objective, weighted by a parameter λ, and the solutions on the efficient frontier are found by varying the parameter λ. See, *e.g.*, Perold [Per84], where a method for efficiently ranging over such a parameterization of the efficient frontier is proposed.

The iterative heuristic we propose for finding a suboptimal solution was developed simultaneously and independently by Jason Schattman, whose numerical results confirm the good performance of the heuristic (see [Sch00], section 3.5). It is also related to the one given by Delaney and Bresler [DB98], in the context of image reconstruction. Meyer [Mey74] establishes the convergence of a large class of algorithms that includes the heuristic discussed in this paper.

For branch and bound methods and integer programming see, for instance, Lawler and Wood [LW66], and Schrijver [Sch86].

1.2 Convex optimization

Our approach is made feasible by relatively recent advances in interior-point methods for nonlinear convex optimization. While these methods can be traced back to the late 1960s (see, *e.g.*, [FM68]), the modern era was initiated by Karmarkar's interior-point method for linear programming [Kar84], which was shown to be more efficient than the simplex method in terms of worst-case complexity analysis and also in practice. More recently, Nesterov and Nemirovsky [NN94] observed that interior-point methods for linear programming can be extended to handle a wide variety of *nonlinear*, nondifferentiable convex optimization problems. Interior-point methods for nonlinear convex optimization problems have been found to have many of the same characteristics of the methods for LP. They have polynomial-time worst case complexity, and they are extremely efficient in practice, requiring a number of iterations that hardly varies with problem size, and is typically between 5 and 50. Each iteration requires the solution (or approximate solution) of a linear system of equations. Current algorithms can solve problems with hundreds of variables and constraints, in times measured in seconds or at most a few minutes, on a personal computer. Larger problems can be handled if problem structure, such as sparsity, is exploited in the solution of the linear system. The current state of the art is described, for example, in the books by Wright [Wri97] and Ye [Ye97]. The forthcoming book by Boyd and Vandenberghe [BV] give an introduction to the field and describe a large number of applications.

Some specific types of nonlinear convex optimization problems have recently been the focus of much research, both in terms of algorithms and applications.

These include semi-definite programming (SDP) [VB96] and second-order cone programming (SOCP) [LVBL98]. In this paper, we will make use of second-order cone programs, with the form

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & \|A_i x + b_i\| \leq c_i^T x + d_i, \quad i = 1, \ldots, L, \\
& Fx = g,
\end{array} \tag{1}
$$

where $\|\cdot\|$ denotes the Euclidean norm, $i.e.$, $\|z\| = \sqrt{z^T z}$. The numerical examples in this paper were solved using the convex optimization software SOCP, by Lobo $et\ al.$ [LVB97]. Other software packages that handle this class of problems are now available, such as MOSEK by Andersen and Andersen [And99] (which currently seems to be the fastest), SEDUMI by Sturm [Stu98], or SDPPACK by Alizadeh $et\ al.$ [AHN$^+$97].

1.3 Overview

The single-period portfolio selection problem is stated in section 2. Transaction costs functions and portfolio constraints are described in section 2.1 and section 2.2. An example of a convex problem with linear transaction costs is presented in section 2.6. Fixed costs are included in section 3, where it is shown how to compute a global bound on performance and how to obtain an approximate solution. Numerical examples are given in section 4. Related problems, such as index tracking, are briefly discussed in section 5. Our conclusions and final comments are given in section 6.

2 The portfolio selection problem

Consider an investment portfolio that consists of holdings in some or all of n assets. This portfolio is to be adjusted by performing a number of transactions, after which the portfolio will be held over a fixed time period. The investor's goal is to maximize the expected wealth at the end of period, while satisfying a set of constraints on the portfolio. These constraints typically include limits on exposure to risk, and bounds on the amount held in each asset. The problem of an investor averse to risk in terms of "mean-variance" preferences can be treated in a similar fashion.

   The current holdings in each asset are $w = (w_1, \ldots, w_n)$. The total current wealth is then $\mathbf{1}^T w$, where $\mathbf{1}$ is a vector with all entries equal to one. The dollar amount transacted in each asset is specified by $x = (x_1, \ldots, x_n)$, with $x_i > 0$ for buying, $x_i < 0$ for selling. After transactions, the adjusted portfolio is $w + x$. Representing the sum of all transaction costs associated with x by $\phi(x)$, the budget constraint is

$$
\mathbf{1}^T x + \phi(x) = 0. \tag{2}
$$

The adjusted portfolio $w + x$ is then held for a fixed period of time. At the end of that period, the return on asset i is the random variable a_i. All random

variables are on a given probability space, for which $\mathbf{E}$ denotes expectation. We assume knowledge of the first and second moments of the joint distribution of $a = (a_1, \ldots, a_n)$,

$$\mathbf{E}\, a = \bar{a}, \qquad \mathbf{E}(a - \bar{a})(a - \bar{a})^T = \Sigma.$$

A riskless asset can be included, in which case the corresponding $\bar{a}_i$ is equal to its (certain) return, and the ith row and column of Σ are zero.

The end of period wealth is a random variable, $W = a^T(w+x)$, with expected value and variance given by

$$\mathbf{E}\, W = \bar{a}^T(w + x), \qquad \mathbf{E}(W - \mathbf{E}\, W)^2 = (w + x)^T \Sigma\,(w + x). \qquad (3)$$

The budget constraint (2) can also be written as an inequality,

$$\mathbf{1}^T x + \phi(x) \leq 0. \qquad (4)$$

With some obvious assumptions ($\bar{a}_i > 0$, $\phi \geq 0$), solving an expected wealth maximization problem with either form of the budget constraint yields the same result. The inequality form is more appropriate for use with numerical optimization methods. (For example, if ϕ is convex, the inequality constraint (4) defines a convex set, while the equality constraint (2) does not.)

We summarize the portfolio selection problem as

$$
\begin{aligned}
\text{maximize} \quad & \bar{a}^T(w + x) \\
\text{subject to} \quad & \mathbf{1}^T x + \phi(x) \leq 0 \\
& w + x \in \mathcal{S}
\end{aligned}
\qquad (5)
$$

where

$\bar{a} \in \mathbf{R}^n$	is the vector of expected returns on each asset,
$w \in \mathbf{R}^n$	is the vector of current holdings in each asset,
$x \in \mathbf{R}^n$	is the vector of amounts transacted in each asset,
$\phi : \mathbf{R}^n \to \mathbf{R}$	is the transaction costs function,
$\mathcal{S} \subseteq \mathbf{R}^n$	is the set of feasible portfolios.

A related problem is that of minimizing the total transaction costs subject to portfolio constraints. Among all possible transactions that result in portfolios achieving a given expected return and meeting the other portfolio constraints, we would like to perform those transactions that incur the smallest total cost. This problem is written as

$$
\begin{aligned}
\text{minimize} \quad & \phi(x) \\
\text{subject to} \quad & \bar{a}^T(w + x) \geq r_{\min} \\
& w + x \in \mathcal{S},
\end{aligned}
\qquad (6)
$$

where $r_{\min}$ is the desired lower bound on the expected return. In this paper we focus mostly on problem (5), but we will also consider problem (6) in section 3.

In the next two sections we describe a variety of transaction costs functions ϕ and portfolio constraint sets $\mathcal{S}$.

2.1 Transaction costs

Transaction costs can be used to model a number of costs, such as brokerage fees, bid-ask spreads, taxes, or even fund loads. In this paper, we assume the transaction costs to be separable, *i.e.*, the sum of the transaction costs associated with each trade

$$\phi(x) = \sum_{i=1}^{n} \phi_i(x_i),$$

where ϕ_i is the transaction cost function for asset i.

The simplest model for transaction costs is that there are none, *i.e.*, $\phi(x) = 0$. In this case the original portfolio is irrelevant, except for its total value. We can make whatever transactions are necessary to arrive at the optimal portfolio. A better model of realistic transactions costs is a linear one, with the costs for each transaction proportional to the amount traded:

$$\phi_i(x_i) = \begin{cases} \alpha_i^+ x_i, & x_i \geq 0 \\ -\alpha_i^- x_i, & x_i \leq 0. \end{cases} \tag{7}$$

Here α_i^+ and α_i^- are the cost rates associated with buying and selling asset i. Linear transaction costs can be used, for example, to model the gap between bid and ask prices. Since the linear transaction costs functions ϕ_i are convex, the budget constraint (4) can be handled by convex optimization. Specifically, linear costs can be handled by introducing new variables $x^+, x^- \in \mathbf{R}^n$, expressing the total transaction as

$$x_i = x_i^+ - x_i^-,$$

with the constraints $x_i^+ \geq 0$, $x_i^- \geq 0$. The transaction costs function ϕ_i is then represented as

$$\phi_i = \alpha_i^+ x_i^+ + \alpha_i^- x_i^-.$$

Any piecewise linear convex transaction costs function can be handled in a similar way.

In practice, transaction costs are not convex functions of the amount traded. Indeed, the costs for either buying or selling are likely to be *concave*. For example, a fixed charge for any nonzero trade is common, and there may be one or more breakpoints above which the transaction costs per share decrease. We will consider a simple model that includes fixed plus linear costs, but our method is readily extended to handle more complex transaction costs functions. Let β_i^+ and β_i^- be the fixed costs associated with buying and selling asset i. The fixed-plus-linear transaction costs function is given by

$$\phi_i(x_i) = \begin{cases} 0, & x_i = 0 \\ \beta_i^+ + \alpha_i^+ x_i, & x_i > 0 \\ \beta_i^- - \alpha_i^- x_i, & x_i < 0. \end{cases} \tag{8}$$

which is illustrated in Figure 1. Evidently this function is not convex, unless the fixed costs are zero. Therefore, the budget constraint (4) cannot be handled by convex optimization.
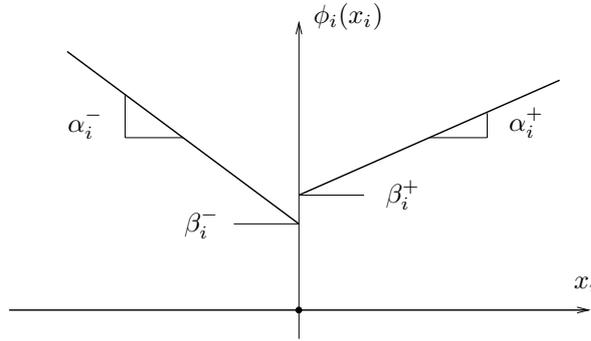
Figure 1: Fixed plus linear transaction costs $\phi_i(x_i)$ as a function of transaction amount x_i. There is no cost for no transaction, *i.e.*, $\phi_i(0) = 0$.

2.2 Diversification constraints

Constraints on portfolio diversification can be expressed in terms of linear inequalities, and therefore are readily handled by convex optimization. Individual diversification constraints limit the amount invested in each asset i to a maximum of p_i,

$$w_i + x_i \le p_i, \qquad i = 1, \ldots, n. \tag{9}$$

Alternatively, we can limit the fraction of the total (post transaction) wealth held in each asset:

$$w_i + x_i \le \gamma_i \mathbf{1}^T (w + x), \qquad i = 1, \ldots, n.$$

These are linear, and therefore convex, inequality constraints on x.

More sophisticated diversification constraints limit the amount of the total wealth that can be concentrated in any small group of assets. Suppose, for example, that we require that no more than a fraction γ of the total wealth be invested in fewer than r assets. Letting $f_{[i]}$ denote the ith largest component of the vector f, this constraint can be expressed as

$$\sum_{i=1}^{r} (w + x)_{[i]} \le \gamma \mathbf{1}^T (w + x). \tag{10}$$

The left-hand side gives maximum wealth held in any subset of r assets. The right-hand side is a factor γ times the total (post transaction) wealth. To see that the constraint (10) is convex, we can express it as a set of $\binom{n}{r}$ linear inequalities, one for each possible combination of r assets chosen from the n assets. This representation is clearly impractical, however, as this number of

linear inequalities can be extremely large. The diversification constraint (10) can be far more efficiently represented by $1 + 2n$ linear inequalities,

$$
\begin{aligned}
\gamma \mathbf{1}^T (w + x) &\geq rt + \mathbf{1}^T y \\
t + y_i &\geq w_i + x_i, \quad i = 1, \ldots, n \\
y_i &\geq 0, \qquad\qquad i = 1, \ldots, n,
\end{aligned}
\tag{11}
$$

where $y \in \mathbf{R}^n$ and $t \in \mathbf{R}$ are new variables.

Several extensions of this type of diversification constraint are possible. For example, we can divide our n assets into N classes of assets, and require that no more than a fraction γ of the total wealth be invested in fewer than R of these classes.

2.3 Shortselling constraints

Shortselling constraints also lead to linear inequalities. Individual bounds s_i on the maximum amount of shortselling allowed on asset i are

$$
w_i + x_i \geq -s_i, \qquad i = 1, \ldots, n.
\tag{12}
$$

(In the case of a riskless asset, s_i is a credit line.) If shortselling is not permitted, the s_i are set to zero. A bound S on total shortselling is

$$
\sum_{i=1}^n (w_i + x_i)_- \ \leq\ S,
$$

where $(\xi)_- = \max\{-\xi, 0\}$. This can be rewritten as a set of linear constraints by introducing an auxiliary variable $t \in \mathbf{R}^n$,

$$
\begin{aligned}
t_i &\geq -(w_i + x_i), \qquad t_i \geq 0, \qquad i = 1, \ldots, n \\
\mathbf{1}^T t &\leq S.
\end{aligned}
\tag{13}
$$

Other variations can be handled in a similar fashion, among which we mention, for its practical interest, the collateralization requirement

$$
\sum_{i=1}^n (w_i + x_i)_- \ \leq\ \gamma \sum_{i=1}^n (w_i + x_i)_+ ,
$$

which limits the total of short positions to a fraction γ of the total of long positions.

2.4 Variance

The standard deviation of the end of period wealth W is constrained to be less than $\sigma_{\max}$ by the (convex) quadratic inequality

$$
(w + x)^T \Sigma (w + x) \leq \sigma_{\max}^2,
$$

7

which is readily handled by convex optimization. We can also express this constraint as

$$\|\Sigma^{1/2}(w+x)\| \le \sigma_{\max}, \tag{14}$$

where $\|\cdot\|$ is the Euclidean or ℓ_2 norm, and $\Sigma^{1/2}$ is the (symmetric) matrix square root of Σ. The constraint (14), which is convex, is a *second-order cone constraint*, and is efficiently handled by recently developed interior-point methods [LVBL98].

Equivalently, a maximum $\sigma_{R,\max}$ can be imposed on the standard deviation of the return R, defined as the ratio of end of period wealth to current wealth, *i.e.*, $R = W/(\mathbf{1}^T w)$. This constraint can be expressed as

$$\|\Sigma^{1/2}(w+x)\| \le \sigma_{R,\max}\mathbf{1}^T w,$$

which is also a second-order cone constraint.

2.5 Shortfall risk constraints

In this section we assume that the returns, the random vector a, have a jointly Gaussian distribution, $a \sim \mathcal{N}(\bar{a}, \Sigma)$. We impose the requirement that the end of period wealth W be larger than some undesired level W^{low} with a probability (or confidence level) exceeding η, where $\eta \ge 0.5$:

$$\mathbf{Prob}\left(W \ge W^{\text{low}}\right) \ge \eta. \tag{15}$$

(For $W^{\text{low}} < \mathbf{1}^T w$, this corresponds to a *value at risk* per dollar invested of $\text{VaR} = 1 - W^{\text{low}}/\mathbf{1}^T w$, for a confidence level of η.) We will show that this probability constraint can be expressed as a second-order cone constraint.

The end of period wealth is a Gaussian random variable, $W = a^T(w+x) \sim \mathcal{N}(\mu, \sigma^2)$. The constraint (15) can be written as

$$\mathbf{Prob}\left(\frac{W - \mu}{\sigma} \le \frac{W^{\text{low}} - \mu}{\sigma}\right) \le 1 - \eta.$$

Since $(W - \mu)/\sigma$ is a zero mean, unit variance Gaussian variable, the probability above is simply $\Phi((W^{\text{low}} - \mu)/\sigma)$, where

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} \, dt$$

is the cumulative distribution function of a zero mean, unit variance Gaussian random variable. Thus, the probability constraint (15) can be expressed as

$$\frac{W^{\text{low}} - \mu}{\sigma} \le \Phi^{-1}(1 - \eta),$$

or, using $\Phi^{-1}(1 - \eta) = -\Phi^{-1}(\eta)$,

$$\mu - W^{\text{low}} \ge \Phi^{-1}(\eta)\sigma.$$

From $\mu = \bar{a}^T(w + x)$ and $\sigma^2 = (w + x)^T \Sigma(w + x)$, we obtain

$$\Phi^{-1}(\eta)\|\Sigma^{1/2}(w + x)\| \leq \bar{a}^T(w + x) - W^{\text{low}}.$$

Now, provided $\eta \geq 0.5$ (and therefore $\Phi^{-1}(\eta) \geq 0$), this constraint is a second-order cone constraint. (If $\eta < 0.5$, the shortfall risk constraint becomes concave in x.) It follows that, under the Gaussian assumption, we can impose one or more shortfall risk constraints, and preserve convexity of the problem. We might impose, for example, a constraint on a merely bad return, with some modest confidence, as well as a constraint on a truly disastrous return, with much greater confidence.

Constraints on loss probability have a simple rectangular shape on a figure showing the cumulative distribution of the return, as illustrated in Figure 2. In the usual expected return versus standard deviation plane, loss probability constraints define half-planes of feasible points, as illustrated in Figure 3.

We can optimize for any convex objective, say the expected return, under shortfall probability constraints of the type described here. Alternatively we can include one shortfall probability constraint and maximize W^{low}. This is also a convex problem: the objective is linear and the constraint is a second-order cone. (In another variation, one shortfall probability constraint is included but the objective is to maximize the confidence level η. This is not a convex problem. It is, however, a quasiconvex problem, and can be efficiently and globally solved by bisection on η.) These three approaches are sometimes called the Telser, Kataoka, and Roy criteria, respectively (see Rudolf [Rud94]).

While in the rest of this paper we only assume knowledge of the first and second moments of the joint distribution of asset returns, this treatment of shortfall risk constraints requires the assumption of a jointly Gaussian distribution. In practice, the observed returns are seldom Gaussian. They often are skewed, or have "fat tails", *i.e.*, resemble a Gaussian distribution in the central area but have higher probability mass for high deviations.

Nevertheless, a shortfall probability approach can be effective if used in an informed manner. Alternatively, we can again assume no knowledge of the distribution except for the first and second moments, and use the Chebyshev bound to limit the shortfall probability (as suggested by Roy [Roy52]). In this case, the factor $\Phi^{-1}(\eta)$ in the formulas above is replaced by $(1 - \eta)^{-1/2}$. Other references for downside risk approaches include Telser [Tel55], Rudolf [Rud94], Leibowitz *et al.* [LBK96], and Lucas and Klaasen [LK98].

2.6 Convex portfolio optimization problems

If any number of convex transaction costs and convex constraints are combined, the resulting problem is convex. Linear transaction costs, as well as all the portfolio constraints described above, are convex, indeed, second-order cone programs. Such problems can be globally solved with great efficiency, even for problems with a large number of assets and constraints.

Example

We consider the specific problem

$$
\begin{aligned}
\text{maximize} \quad & \bar{a}^T(w + x^+ - x^-) \\
\text{subject to} \quad & \mathbf{1}^T(x^+ - x^-) + \sum_{i=1}^{n}(\alpha_i^+ x_i^+ + \alpha_i^- x_i^-) \leq 0 \\
& x_i^+ \geq 0, \quad x_i^- \geq 0, \quad i = 1, \ldots, n \\
& w_i + x_i^+ - x_i^- \geq s_i, \quad i = 1, \ldots, n \\
& \Phi^{-1}(\eta_j)\|\Sigma^{1/2}(w + x^+ - x^-)\| \leq \bar{a}^T(w + x^+ - x^-) - W_j^{\text{low}}, \\
& \hspace{8cm} j = 1, \ldots, 2.
\end{aligned}
$$

with 100 risky and one riskless assets (so $n = 101$). This specifies linear transaction costs, a limit on shortselling of s_i per asset, and two shortfall risk constraints.

The mean and covariance of the risky assets was estimated from one year of daily closing prices of S&P 500 stocks (the first 100, alphabetically by ticker, with a full year of data from January 9, 1998 to January 8, 1999). The distribution was scaled for a portfolio holding period of 20 days. The riskless asset was assumed to have unit return. Far more sophisticated methods could have been used to estimate the return mean and covariance; our only goal here is to demonstrate the optimization method.

For transactions cost and constraint parameters, we (arbitrarily) selected the values

$$
\begin{array}{ll}
w_1, \ldots, w_{100} = 1/101, & w_{101} = 1/101 \\
\alpha_1^+, \ldots, \alpha_{100}^+ = 1\%, & \alpha_{101}^+ = 0 \\
\alpha_1^-, \ldots, \alpha_{100}^- = 1\%, & \alpha_{101}^- = 0 \\
s_1, \ldots, s_{100} = 0.005, & s_{101} = 0.5
\end{array}
$$

(where index 101 is the riskless asset). For the shortfall constraints, we chose

$$
\eta_1 = 80\%, \quad W_1^{\text{low}} = 0.9; \quad \text{and} \quad \eta_2 = 97\%, \quad W_2^{\text{low}} = 0.7,
$$

which correspond to a limit on probabilities of a bad and of a disastrous return, respectively.

This problem is a second-order cone program (SOCP), with 202 variables and 306 constraints. The optimal portfolio was obtained in approximately three minutes on a personal computer, using the general purpose software `socp`, which does not take any advantage of the (substantial) sparsity in the problem data.

Figure 2 plots the cumulative distribution of the return for the optimal portfolio. The 50% probability level corresponds, on the horizontal axis, to the expected return. The loss probability constraints are also drawn in the figure. Note that the 0.7 return (0.3 value at risk) for a 97% confidence level is the active constraint.

Figure 3 plots the tradeoff curve of expected return versus standard deviation of return, which is the efficient frontier for the problem (ignoring the shortfall probability constraints). On this plot, the shortfall probability constraints correspond to half-planes, whose boundaries are shown as dashed lines
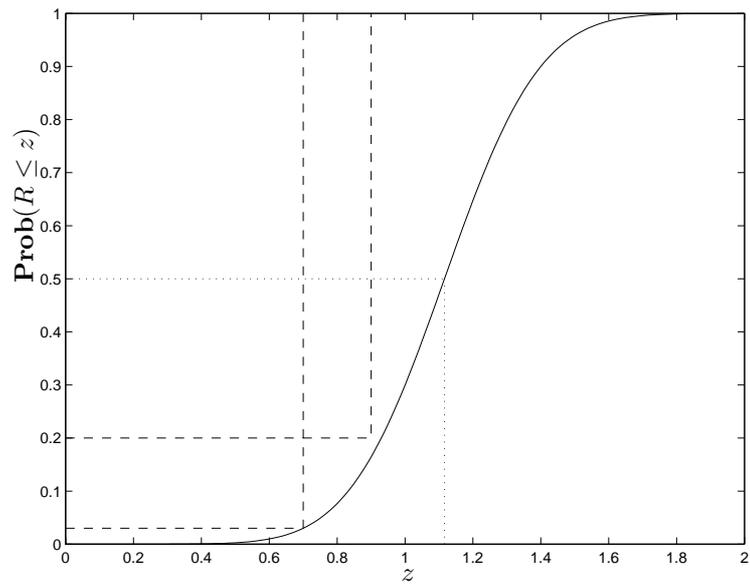
10

Figure 2: Cumulative distribution function of the return, for the optimal portfolio in the example with 100 stocks plus a riskless asset. The expected return, which is also the median return since the distribution is assumed Gaussian, is shown with the dotted line. The two limits on shortfall probability are shown as dashed lines. The limit on the right, and higher, limits the probability of a return below 0.9 (*i.e.*, a bad return) to no more than 20%; the limit on the left, and lower, limits the probability of a return below 0.7 (*i.e.*, a disastrous return) to no more than 3%.
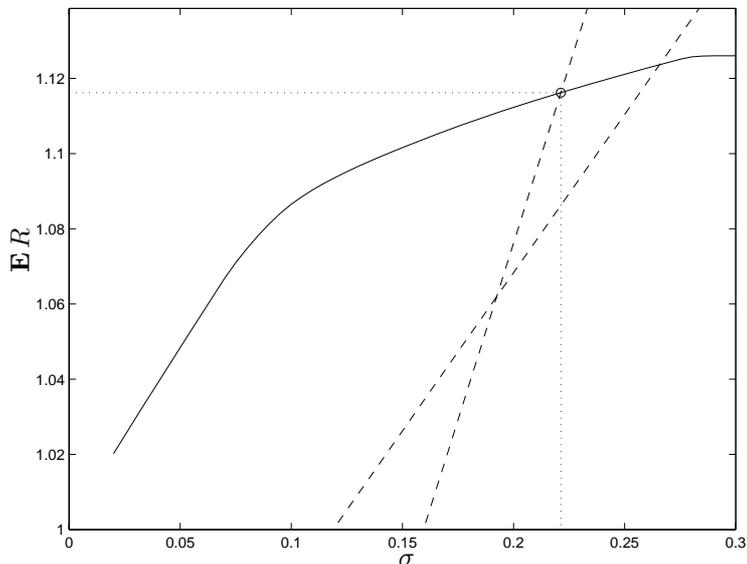
Figure 3: Efficient frontier of return mean versus return variance for example problem with 100 stocks plus riskless asset, ignoring the shortfall probability constraints. The sloped dashed lines show the limits imposed by the shortfall probability constraints. The optimal solution of the problem with the shortfall constraints is shown as the small circle.

in the figure. The dashed line with the smaller slope corresponds to the limit on the probability of the bad return; the line with steeper slope corresponds to the limit on the probability of the disastrous loss. The optimal solution is marked with a small circle.

3   Fixed transaction costs

To simplify notation, we assume from now on equal costs for buying and selling, the extension for nonsymmetric costs being straightforward. The transaction costs function is then

$$\phi(x) = \sum_{i=1}^{n} \phi_i(x_i),$$

with

$$\phi_i(x_i) = \begin{cases} 0, & x_i = 0 \\ \beta_i + \alpha_i |x_i|, & x_i \neq 0. \end{cases} \tag{16}$$

In the general case, costs of this form lead to a hard combinatorial problem.

The simplest way to obtain an approximate solution is to ignore the fixed costs, and solve for $\phi_i(x_i) = \alpha_i |x_i|$. If the β_i are very small, this may lead to

an acceptable approximation. In general, however, it will generate inefficient solutions with too many transactions. (Note that if this approach is taken and the solution is computed disregarding the fixed costs, some margin must be added to the budget constraint to allow for the payment of the fixed costs.)

On the other hand, by considering the fixed costs, we discourage trading small amounts of a large number of assets. Thus, we obtain a *sparse* vector of trades; *i.e.*, one that has many zero entries. This means most of the trading will be concentrated in a few assets, which is a desirable property.

3.1 Finding the global optimum

To compute the exact solution, an exhaustive combinatorial search can be performed. Each x_i can either be set to zero (with $\phi_i(x_i) = 0$), or assumed to be nonzero (with $\phi_i(x_i) = \alpha_i|x_i| + \beta_i$). These two alternatives for each of the n assets yields 2^n combinations. For each of these combinations a convex optimization problem results, since the ϕ_i are convex (either zero or linear plus a constant). If we solve each of these 2^n convex problems, the global optimum is found by choosing, out of the 2^n solutions, the one with the highest expected end of period wealth. Because of computational requirements, this approach becomes difficult for n larger than 10, and certainly unrealistic for n larger than 20.

The exhaustive search can be improved by using branch and bound methods, which can greatly reduce the required computational effort. These methods require a procedure for computing upper and lower bounds. We will next propose a heuristic, which can be used to find approximate solutions (and therefore lower bounds). Our experience indicates that this heuristic performs consistently well, producing high-quality suboptimal solutions. We also show how to compute a global upper bound on the achievable end of period wealth, which makes it possible to embed the heuristic in branch and bound methods.

3.2 Convex relaxation and global bound

We assume that lower and upper bounds on the x_i are known, *i.e.*, l_i and u_i for which x_i must satisfy

$$-l_i \leq x_i \leq u_i.$$

(We will later describe how to obtain such bounds from the portfolio constraints.) The *convex envelope* of ϕ_i, which is the largest convex function which is lower or equal to ϕ_i in the interval $[-l_i, u_i]$, is given by

$$\phi_i^{\text{c.e.}}(x_i) = \begin{cases} \left(\dfrac{\beta_i}{u_i} + \alpha_i\right) x_i, & x_i \geq 0 \\ -\left(\dfrac{\beta_i}{l_i} + \alpha_i\right) x_i, & x_i \leq 0. \end{cases} \tag{17}$$

This is shown in Figure 4. The extension for the case when one or both bounds are not available, *i.e.*, $l_i = -\infty$ or $u_i = +\infty$, is trivial (*e.g.*, $\phi_i^{\text{c.e.}}(x_i) = \alpha_i x_i$ for $x_i \geq 0$ if $u_i = +\infty$.)

13

Using $\phi_i^{\text{c.e.}}$ for ϕ_i relaxes the budget constraint, in the sense that it enlarges the search set. Consider the portfolio selection problem (5), with $\phi_i^{\text{c.e.}}$ replaced for ϕ_i,

$$
\begin{aligned}
\text{maximize} \quad & \bar{a}^T(w + x) \\
\text{subject to} \quad & \mathbf{1}^T x + \sum_{i=1}^n \phi_i^{\text{c.e.}}(x_i) \leq 0 \\
& w + x \in \mathcal{S}.
\end{aligned}
\tag{18}
$$

This corresponds to optimizing the same objective (the expected end of period wealth), subject to the same portfolio constraints, but with a looser budget constraint. Therefore the optimal value of (18) is an upper bound on the optimal value of the unmodified problem (5). Since the problem (18) is convex, we can compute its optimal solution, and hence the upper bound on the optimal value of the original problem (5), very efficiently.

Note that in most cases the optimal transactions vector for the relaxed problem (18) will not be feasible for the original problem (5). The unmodified budget constraint will not be satisfied by the solution of the modified program, except in the very special case when each transaction amount x_i is either l_i, u_i, or 0. (These are the three values for which the convex envelope and the true transaction costs function agree.)

This relaxation can also be used in problem (6), where the goal is to minimize transaction costs. This results in the relaxed problem

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^n \phi_i^{\text{c.e.}}(x_i) \\
\text{subject to} \quad & \bar{a}^T(w + x) \geq r_{\min} \\
& w + x \in \mathcal{S}.
\end{aligned}
\tag{19}
$$

Here, compared to the original problem, the relaxed problem has the same feasible set, but a different objective function.

3.3 Bounds on the x_i

Upper and lower bounds on the x_i are required to perform the convex relaxation and obtain a global upper bound on the expected end of period wealth. Such bounds are easily derived from the diversification, shortselling and budget constraints. However, it is highly desirable to find tighter bounds on the x_i (because the tightness of the global upper bound will in turn depend on the tightness of these bounds). The computation of such bounds depends on the particular problem being addressed. If the problem includes a constraint on the variance of portfolio return, an upper bound on x_i is given by the solution of

$$
\begin{aligned}
\text{maximize} \quad & f(x) = x_i \\
\text{subject to} \quad & g(x) = (w + x)^T \Sigma (w + x) - \sigma^2 \leq 0.
\end{aligned}
$$

The gradients are

$$
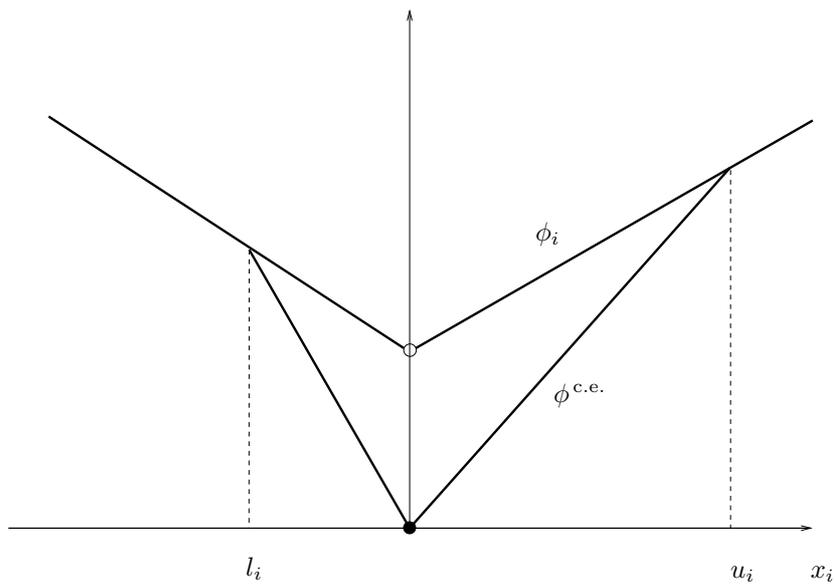\nabla_x f = \lambda e_i, \quad \nabla_x g = 2\Sigma(w + x).
$$

Figure 4: The convex envelope of ϕ_i over the interval $[l_i, u_i]$, is the largest convex function smaller than ϕ_i over the interval. For fixed plus linear costs, as shown here, the convex envelope is a linear transaction costs function.

Forming the Lagrangian and minimizing, a simple calculation leads to

$$x_i = \sigma \sqrt{(\Sigma^{-1})_{ii}} - w_i,$$

where $(\,\cdot\,)_{ii}$ denotes the i, i element of the matrix. The same can be done for a lower bound,

$$x_i = -\sigma \sqrt{(\Sigma^{-1})_{ii}} - w_i,$$

although it is likely that, in this case, the bounds derived from the shortselling constraints will be tighter.

3.4   Iterative heuristic

We now describe a heuristic for finding a feasible, suboptimal portfolio, which is based on the same method used to find an upper bound. At the end of that section, we noted that the transaction x that is optimal for the relaxed problem (18) would be feasible for the original problem (5) if, by chance, the optimal x for the relaxation satisfied $\phi^{\text{c.e.}}(x) = \phi(x)$. This only happens when each transaction is either zero, or at one of its bounds l_i or u_i.

   The iterative procedure uses a modified transaction costs function which, like the relaxed cost function, is convex. Unlike the relaxed cost function, however, we do not require the convex cost function to be a lower bound on the true transaction costs function.

   An iterated reweighting of this convex cost function is used, in such a way that most small transactions in the solution are driven to zero. We use δ (some nonnegative and small value) as a threshold for deciding when a transaction is considered to be zero. Since each of these reweighted modified functions is convex, each iteration consists in solving a convex program. The feasibility of the portfolio is obtained by ensuring that the modified transaction costs function $\widehat{\phi}_i^k$ agrees with the true ϕ_i at the solution transactions x_i^*. These ideas will become clearer with the ensuing discussion.

   Consider the following procedure.

1. $k := 0$.
   Solve the convex relaxed problem (18).
   Let x^0 be the solution to this problem.

2. $k := k + 1$.
   Given the solution to the previous problem x^{k-1}, define $\widehat{\phi}_i^k$ as

$$\widehat{\phi}_i^k(x_i) = \left(\frac{\beta_i}{|x_i^{k-1}| + \delta} + \alpha_i \right) |x_i|.$$

   Solve the modified (convex) portfolio selection problem

$$\begin{aligned}
\text{maximize} \quad & \bar{a}^T(w + x) \\
\text{subject to} \quad & \mathbf{1}^T x + \sum_{i=1}^n \widehat{\phi}_i^k(x_i) \leq 0 \\
& w + x \in \mathcal{S}.
\end{aligned} \tag{20}$$

16

Let x^k be the solution to this problem.

3. If the portfolios x^k and x^{k-1} found in the two previous iterations are (approximately) equal, return $x^* := x^k$ and exit.

   Otherwise, go to step 2.

A rough interpretation of the algorithm is that, in each iteration, we amortize the fixed costs evenly over the transaction amount in the previous iteration.

If this iterative procedure exits, which occurs if two successive iterates are close to each other, the solution x^* will be nearly feasible for the original problem (see Figure 5). This is seen by noting that, for $x_i^* \gg \delta$,

$$\widehat{\phi}_i(x_i^*) \; = \; \left(\frac{\beta_i}{|x_i^*| + \delta} + \alpha_i \right) |x_i^*| \; \approx \; \beta_i + \alpha_i |x_i^*| \; = \; \phi_i(x_i^*),$$

and for $x_i^* = 0$,

$$\widehat{\phi}_i(x_i^*) \; = \; 0 \; = \; \phi_i(x_i^*).$$

In a sense, δ defines a *soft threshold* for deciding whether a given x_i^* is considered zero, *i.e.*, whether the corresponding transaction should be performed or not. In a practical implementation of the portfolio trades, a hard threshold is needed, and the x_i^* on the order of δ or smaller should be taken as zero. Note that while $\widehat{\phi}_i(x_i^*) \le \phi_i(x_i^*)$ for all x_i^*, this inequality is tight except for x_i^* in the order of δ. Such terms may lead to feasibility problems, with a nonnegligible violation of the budget constraint. In practice this is not an issue since terms in the order of δ will seldom appear in the solution. On each iteration, for the x_i^k that become small, the modified $\widehat{\phi}_i^{k+1}(x_i)$ has an increased derivative. This eventually pushes the small x_i to zero, leading to sparse solutions.

This also provides the motivation for the method, and an intuition to justify the quality of the approximate solutions that have been found in numerical experiments.

The same method is applicable to problem (6), where we seek to minimize the transaction costs. The iterative heuristic is identical, except that in each iteration instead of problem (20), the following problem is solved

$$
\begin{aligned}
& \text{minimize} && \sum_{i=1}^n \widehat{\phi}_i^k(x_i) \\
& \text{subject to} && \bar{a}^T(w + x) \ge r_{\min} \\
& && w + x \in \mathcal{S}.
\end{aligned}
\tag{21}
$$

A proof of convergence for this heuristic (for the successive relaxation of the objective function) is given in Appendix A. We note in the appendix that this heuristic is equivalent to finding a local minimum of a log-like, concave function.

Our numerical experiments (performed on problem (5)) indicate that convergence occurs in about 4 iterations or less for a wide range of problems.

Upper and lower bounds on the global optimum for the expected end of period wealth are given by $\bar{a}^T x^0$ and $\bar{a}^T x^*$. As a final step, an extra iteration can be included with the sparsity pattern fixed, and with the transaction costs
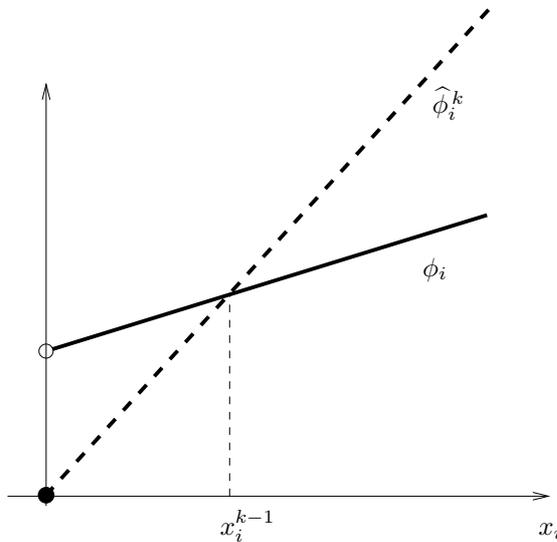
17

Figure 5: One iteration of the algorithm. Each of the nonconvex transaction costs (plotted as a solid line) is replaced by a convex one (plotted as a dashed line) that agrees with the nonconvex one at the current iterate. If two successive iterates are the same, then the iterates are feasible for the original nonconvex problem.

exact for that pattern. That is, the small x_i are set to zero, with $\phi_i(x_i) = 0$, and the others are assumed nonzero, with $\phi_i(x_i) = \beta_i + \alpha_i|x_i|$. (This is equivalent to one particular combination, out of the possible 2^n in the exhaustive search.)

4   Examples with fixed costs

For numerical examples, we use the same stock data as in the previous example. We specify fixed plus linear transaction costs, and constraints on shortselling and on variance. We first describe an example with 10 stocks, plus a riskless asset. The parameters used (again, arbitrarily) are

$$
\begin{aligned}
&w_1, \ldots, w_{11} = 1/11 \\
&\alpha_1, \ldots, \alpha_{10} = 1\%, \qquad \alpha_{11} = 0 \\
&\beta_1, \ldots, \beta_{10} = 0.01, \qquad \beta_{11} = 0 \\
&s_1, \ldots, s_{10} = 0.05, \qquad s_{11} = 0.5.
\end{aligned}
$$

The small size of this problem allows us to compute the exact solution, that is the global optimum, by combinatorial search. Figure 6 displays the resulting tradeoff curve, with expected return plotted against standard deviation of return. Four curves are shown: the upper bound; the exact solution; the heuristic solution; and the solution computed without incorporating the fixed cost.
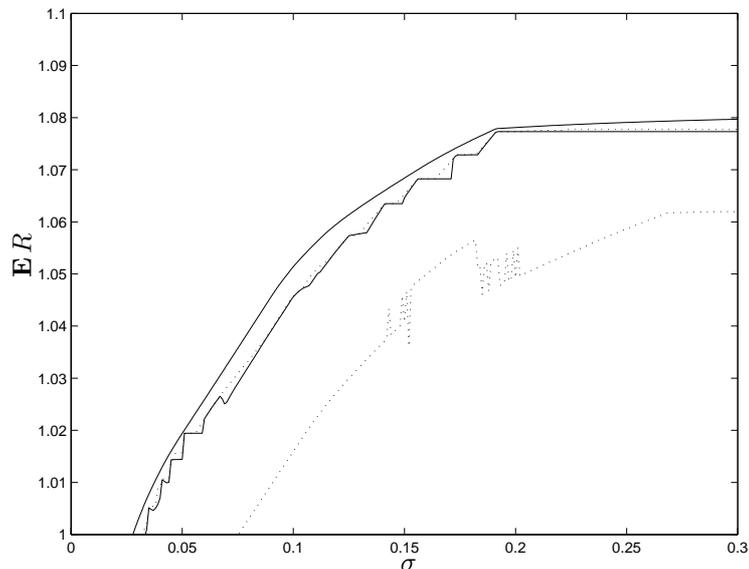
Figure 6: Example with 10 stocks plus riskless asset, plot of expected return as a function of standard deviation. Curves from top to bottom are: 1. global upper bound (*solid*), 2. true optimum by exhaustive search (*dotted*), 3. heuristic solution (*solid*), and 4. solution computed without regard for fixed costs (*dotted*). Note that curves 2 and 3 are nearly coincidental.

Note that the upper bound is close to the heuristic solution. Note also that the heuristic solution nearly coincides with the exact solution, even though the heuristic required only about one-thousandth the computational effort. For the heuristic, we used $\delta = 10^{-3}$ (and did not include a final iteration with fixed sparsity pattern).

In Figure 7, still for the same 11 assets example, $\sigma_{\max}$ was kept constant at 0.15, and the problem was solved for a range of fixed costs β. The optimal expected return is plotted as a function of fixed costs, with the four curves obtained by the same procedure as in the previous figure. Again we can see that the difference between our heuristic and the optimal is very small. In this figure we can also see the cost of ignoring the transaction costs, which, naturally, increases with increasing fixed transaction costs.
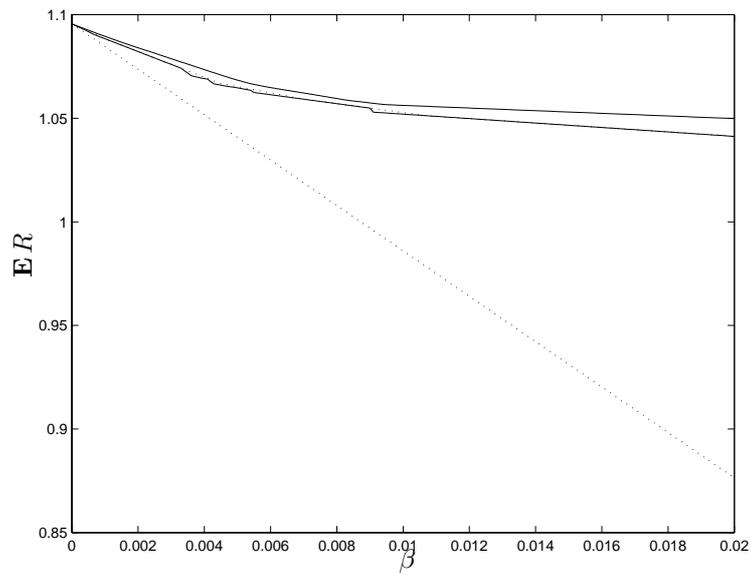
Figure 7: Example with 10 stocks plus riskless asset, plot of expected return as a function of fixed transaction costs. Curves from top to bottom are: 1. global upper bound (*solid*), 2. true optimum by exhaustive search (*dotted*), 3. heuristic solution (*solid*), and 4. solution computed without regard for fixed costs (*dotted*). Note that curves 2 and 3 are nearly coincidental.

As a second and larger example, we considered 100 stocks, plus a riskless asset, using the parameters

$$
\begin{aligned}
&w_1, \ldots, w_{101} = 1/101 \\
&\alpha_1, \ldots, \alpha_{100} = 1\%, \qquad \alpha_{101} = 0 \\
&\beta_1, \ldots, \beta_{100} = 0.001, \qquad \beta_{101} = 0 \\
&s_1, \ldots, s_{100} = 0.005, \qquad s_{101} = 0.5.
\end{aligned}
$$

Figure 8 displays the resulting tradeoff curve. The curves shown are the upper bound, the heuristic solution, and the solution computed without regard for fixed costs. The exact tradeoff curve is not shown since, in this case, it would require a prohibitive effort to compute. However, the fact that the upper bound and the heuristic are close to each other establishes that computing the exact, globally optimal solution would only yield a small improvement over the heuristic solution. Note that the heuristic takes only a few minutes to compute, while the time required to perform the combinatorial search with the same computational resources would be of the order of 2^{100}, or 10^{30} minutes. This is a rather long time for what is guaranteed to be only a marginal improvement (for reference, the age of the universe is approximately 10^{16} minutes.)

Again, we note that the upper bound and the heuristic can be embedded in a branch and bound algorithm to find the global optimum. Since the upper bound and the heuristic are often close, it can be expected that such a branch and bound search would exhibit good convergence properties.
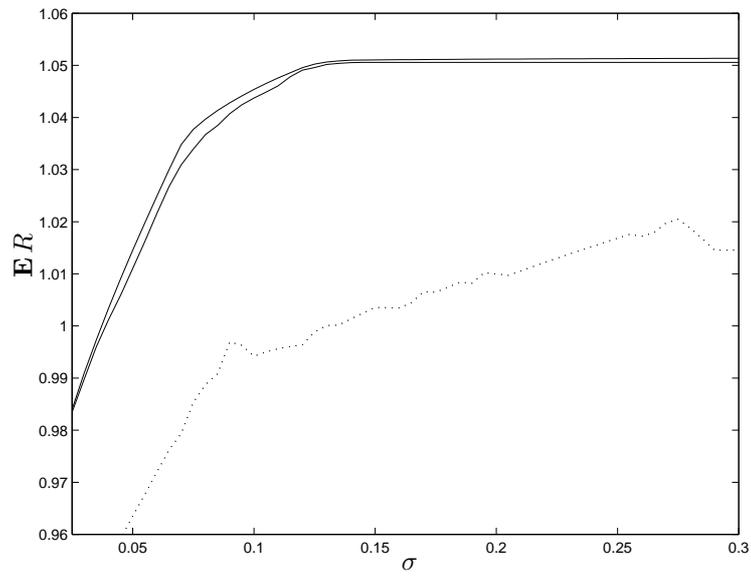
Figure 8: Example with 100 stocks plus riskless asset, plot of expected return as a function of standard deviation. Curves from top to bottom are: 1. global upper bound (*solid*), 2. heuristic solution (*solid*), and 3. solution computed without regard for fixed costs (*dotted*).

5 Related problems

A related problem is that of tracking a portfolio over a single period, with high accuracy and low cost. This arises, for instance, in the problem of reproducing an index formed by a large number of stocks, where trading in all the relevant stocks would incur excessive costs. In this case, the desired solution is a portfolio consisting of a relatively small subset of the stocks present in the index that, with high probability, behaves like the index. (An equivalent problem is that of portfolio insurance, where the goal is to reduce the downside risk as much as possible by buying put options on a small number of the assets.)

A measure of closeness between the adjusted portfolio $w + x$ and a reference portfolio v is given by the expected square difference in returns (v may be the index to be tracked, or the portfolio to be insured). The expected square difference in the returns of the two portfolios is equal to

$$\mathbf{E}\,(a^T(w + x) - a^T v)^2 = (w + x - v)^T(\Sigma + \bar{a}\bar{a}^T)(w + x - v),$$

which we refer to as the tracking error, for short. It has a number of interesting properties, but the one of main concern here is that it is convex in x.

The problem can be seen as having two conflicting objectives. In addition to low transaction costs $\phi(x)$, we want a portfolio $w + x$ with small tracking error. We can address this problem in any of the alternative formulations: minimize the costs subject to a bound on the tracking error; minimize the tracking error subject to a bound on the costs; or minimize a weighted combination of costs and tracking error. (Other constraints, such as on budget and shortselling, will also be included, of course.)

For linear transaction costs, each of these problem formulations is convex and easy to solve (a quadratic program or a quadratically constrained quadratic program). When fixed transaction costs are present, a difficult combinatorial program results. This can be addressed by the same method as in section 3, which will produce a global bound on achievable performance, and a suboptimal solution. Numerical simulations for the tracking problem have produced suboptimal solutions of consistently high quality. Nevertheless, if so desired, the global bound and the heuristic can be embedded in a branch and bound method to guarantee a high accuracy solution.

Of many other possible extensions, most worthy of mention are those to a multi-period setting or to continuous time. These are significantly more complex problems due to the stochastic dynamics. The desirability of a trade in a given stock must then take into account the alternative of delaying the trade. The challenge is to develop effective numerical methods for the (approximate) solution of the resulting stochastic programming (or optimal stopping) problems.

6   Conclusions

We have described a number of portfolio optimization problems that are convex, and therefore efficiently solved. If fixed transaction costs are included, the resulting problem is not convex. In this case, we showed how to compute a global upper bound from a convex relaxation, and proposed a heuristic for computing an approximate solution (which yields, of course, a lower bound). Computational experiments suggest that the gap between our heuristic suboptimal solution and our guaranteed upper bound is very often small. If further accuracy is needed, the upper bound and the heuristic method can be incorporated in a branch and bound method.

The unifying idea in this paper is to exploit new efficient interior-point methods for convex optimization. While such methods are of polynomial complexity (in simple implementations, cubic) in problem dimension, the availability of computing resources over time shows no signs of departing from geometric growth. As a consequence, interior-point methods will be able to handle very large problems, in very short run-times in the near future. Currently, run-times are in the order of a minute for problems with a few hundred variables, on an inexpensive personal computer, using generic software that is not optimized for portfolio problems.

A Convergence of the heuristic

We outline the proof of convergence of the following successive relaxation of a nonconvex objective function. Define the transformation $A : \mathbf{R}^n \to \mathbf{R}^n$,

$$A(y) = \underset{x \in \mathcal{S}}{\operatorname{arginf}} \sum_{i=1}^{n} \frac{x_i}{y_i + \delta},$$

with $\mathcal{S} \subset \mathbf{R}^n$ a convex, compact set, and $\delta > 0$. We show that the sequence x^k, such that

$$x^0 \in \mathcal{S} \quad \text{and} \quad x^{k+1} = A(x^k),$$

satisfies $x_i^{k+1} - x_i^k \to 0$, for $i = 1, \ldots, n$. For simplicity, we assume $x_i \geq 0$. This sequence corresponds to the heuristic applied to problem (6), where we successively relax the objective function. As far as we can tell, the proof is significantly harder for the successive relaxation of a constraint.

To prove this result, we first define the function $L : \mathbf{R}^n \to \mathbf{R}_+$, ($\delta > 0$)

$$L(x) = \prod_{i=1}^{n} (x_i + \delta),$$

and show that the sequence $L(x^k)$ is monotonically nonincreasing (accordingly, we refer to L as the *descent function*). Since $x^{k+1} = A(x^k)$ yields the infimum over $\mathcal{S}$, and $x^k \in \mathcal{S}$, we have that

$$\sum_{i=1}^{n} \frac{x_i^{k+1} + \delta}{x_i^k + \delta} \leq \sum_{i=1}^{n} \frac{x_i^k + \delta}{x_i^k + \delta} = n.$$

Using the inequality between the arithmetic and geometric means for nonnegative terms, we conclude

$$\prod_{i=1}^{n} \frac{x_i^{k+1} + \delta}{x_i^k + \delta} \leq 1,$$

which implies $L(x^{k+1}) \leq L(x^k)$. Since L is bounded below by δ^n, the sequence $L(x^k)$ converges.

Now, convergence of $L(x^k)$ to a nonzero limit implies that

$$\frac{L(x^{k+1})}{L(x^k)} = \prod_{i=1}^{n} \frac{x_i^{k+1} + \delta}{x_i^k + \delta} \to 1.$$

Define y^{k+1} to be

$$y_i^{k+1} = \frac{x_i^{k+1} + \delta}{x_i^k + \delta},$$

and write $y_1^k = 1 + \epsilon$, it follows that

$$\prod_{i=1}^{n} y_i^k = (1 + \epsilon) \prod_{i=2}^{n} y_i^k \leq (1 + \epsilon) \left(1 - \frac{\epsilon}{n-1}\right)^{n-1} = f(\epsilon),$$

25

where we used $\sum_{i=1}^{n} y_i^{k+1} \leq n$, and the inequality between the arithmetic and geometric means. The function f is continuous in ϵ and, with some algebra, it is easily checked that $f(0) = 1$, $f'(0) = 0$, and $f''(\epsilon) < 0$, for $|\epsilon| < 1$. Therefore, $f(\epsilon) < 1$ for $\epsilon \neq 0$, $|\epsilon| < 1$.

We conclude that $\prod_{i=1}^{n} y_i^k \to 1$ implies $f(\epsilon) \to 1$, and that this in turn implies $\epsilon \to 0$. Hence $y_1^k \to 1$, and likewise for all y_i^k.

Using $x_i^k + \delta \leq M < \infty$ for all k (since the set $\mathcal{S}$ is bounded), we obtain the desired result:

$$x_i^k - x_i^{k+1} \to 0.$$

Note that, upon convergence, the partial derivative with respect to x_i of the function minimized in the last iteration is given by

$$\frac{1}{x_i^* + \delta},$$

which is equal to the derivative of the function

$$\sum_{i=1}^{n} \log(x_i + \delta),$$

at $x_i = x_i^*$. From the equality of the first-order conditions for optimality, it is easy to see that the iterative procedure finds a local minimum in $\mathcal{S}$ of this logarithmic function.

References

[AHN⁺97] F. Alizadeh, J. P. Haeberly, M. V. Nayakkankuppam, M. L. Overton, and S. Schmieta. SDPPACK *User's Guide, Version 0.9 Beta*. NYU, June 1997.

[And99] E. Andersen. *MOSEK v1.0b User's manual*, 1999. Available from the url `http://www.mosek.com`.

[BDS99] D. Bertsimas, C. Darnell, and R. Soucy. Portfolio construction through mixed-integer programming at Grantham, Mayo, Van Otterloo and Company. *Interfaces*, 29(1):49–66, January-February 1999.

[BHKT83] B. Blog, G. V. D. Hoek, A. H. G. Rinnooy Kan, and G. T. Timmer. The optimal selection of small portfolios. *Management Science*, 29(7):792–798, July 1983.

[BV] S. P. Boyd and L. Vandenberghe. Convex optimization. Available at www.stanford.edu/~boyd. To be published in 2003.

[DB98] A. H. Delaney and Y. Bresler. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Transactions on Signal Processing*, 7(2):204–221, February 1998.

[FM68] A. Fiacco and G. McCormick. *Nonlinear programming: sequential unconstrained minimization techniques*. Wiley, 1968. Reprinted 1990 in the SIAM Classics in Applied Mathematics series.

[Kar84] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.

[LBK96] M. L. Leibowitz, L. N. Bader, and S Kogelman. *Return targets and shortfall risks: studies in strategic asset allocation*. Salomon Brothers, Irwin publishing, Chicago, 1996.

[Lin65] John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1):13–37, February 1965.

[LK98] A. Lucas and P. Klaasen. Extreme returns, downside risk, and optimal asset allocation. *The Journal of Portfolio Management*, pages 71–79, Fall 1998.

[Lue98] D. G. Luenberger. *Investment Science*. Oxford University Press, New York, 1998.

[LVB97] M. S. Lobo, L. Vandenberghe, and S. Boyd. SOCP*: Software for Second-Order Cone Programming*. Information Systems Laboratory, Stanford University, 1997.

[LVBL98] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and Applications*, 284(1-3):193–228, November 1998.

[LW66] E. L. Lawler and D. E. Wood. Branch-and-bound methods: A survey. *Operations Research*, 14:699–719, 1966.

[Mar52] H. M. Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.

[Mar59] H. M. Markowitz. *Portfolio Selection*. J. Wiley & Sons, New York, 1959.

[Mey74] R. R. Meyer. Sufficient conditions for the convergence of monotonic mathematical programming algorithms. *Journal of computer and system sciences*, 12:108–121, 1974.

[NN94]   Yu. Nesterov and A. Nemirovsky. *Interior-point polynomial methods in convex programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.

[Per84]   A. F. Perold. Large-scale portfolio optimization. *Management Science*, 30(10):1143–1160, October 1984.

[PS82]   N. R. Patel and M. G. Subrahmanyam. A simple algorithm for optimal portfolio selection with fixed transaction costs. *Management Science*, 28(3):303–314, 1982.

[Roy52]   A. D. Roy. Safety first and the holding of assets. *Econometrica*, 20:413–449, 1952.

[Rud94]   M. Rudolph. *Algorithms for portfolio optimization and portfolio insurance.* Haupt, Switzerland, 1994.

[Sch86]   A. Schrijver. *Theory of Linear and Integer Programming.* Wiley-Interscience series in discrete mathematics. John Wiley & Sons, 1986.

[Sch00]   Jason B. Schattman. *Portfolio selection under nonconvex transaction costs and capital gain taxes.* PhD thesis, Rutgers University, May 2000.

[Sha64]   William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, September 1964.

[Stu98]   J. Sturm. SeDuMi *Version 1.02.* McMaster University, 1998.

[Tel55]   L. G. Telser. Safety first and hedging. *Review of Economics and Statistics*, 23:1–16, 1955.

[VB96]   L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, March 1996.

[Wri97]   S. J. Wright. *Primal-Dual Interior-Point Methods.* SIAM, Philadelphia, 1997.

[Ye97]   Y. Ye. *Interior Point Algorithms: Theory and Analysis.* Discrete Mathematics and Optimization. Wiley, New York, 1997.