The worst-case risk of a portfolio

Miguel Sousa Lobo[1]　　Stephen Boyd[2]

Technical Report

September 2000

[1]mlobo@duke.edu
[2]boyd@stanford.edu

Abstract

We show how to compute in a numerically efficient way the maximum risk of a portfolio, given uncertainty in the means and covariances of asset returns. This is a semidefinite programming problem, and is readily solved by interior-point methods for convex optimization developed in recent years. While not as general, this approach is more accurate and much faster than Monte Carlo methods. The computational effort required grows gracefully, so that very large problems can be handled. The proposed approach is extended to portfolio selection, allowing for the design of portfolios which are robust with respect to model uncertainty.

- Contact: `mlobo@duke.edu`

- First draft: February 1999; Second draft: August 1999; This version: September 2000
- Research supported in part by NSF Grant ECS-9707111, by AFOSR Grant F49620-98-1-0147, by MURI Grant 49620-95-1-0525, and by the Portuguese Government under Praxis XXI

1 Introduction

Consider a portfolio of risky assets held over a single period, where the distribution of the asset returns is imprecisely known. That is, estimates exist for the expected value and covariance matrix of the asset returns, but these estimates are subject to estimation errors, and possibly to modeling errors.

Modeling errors arise, for example, because statistical procedures make the unrealistic assumption of distributions being stationary. Further, since the covariance matrix has $n\left(n+1\right)/2$ independent entries, if a large number of assets is to be considered then a very large number of samples is required in order to have good, uncorrelated estimates of all the entries in the covariance matrix. The usual fix consists of assuming a structure for the covariance matrix (typically diagonal plus low-rank), but it is not clear that this is statistically justified.

We use the following notation. There are n assets, with expected return $\mu \in \mathbf{R}^n$ and covariance $\Sigma \in \mathbf{R}^{n\times n}$. The portfolio under analysis is described by the weight vector $w \in \mathbf{R}^n$, where w_i represents the fraction of the total wealth held in asset i.

In a simple analysis of the risk associated with the portfolio, these estimates of asset statistics are assumed exact. The portfolio expected return and variance are then assumed to be $\mu^T w$ and $w^T \Sigma w$. Such an approach does not account for the imprecision in the estimates of the asset statistics, which may have a significant effect on the risk associated with the portfolio.

Repeating the analysis under a small number of different scenarios (*i.e.*, with different values for the expected returns and covariance matrix) is a simple way of dealing with inaccuracy in the estimates. It is, however, an inadequate approach for problems with a moderate to large number of assets. A large number of scenarios may be run in a Monte Carlo procedure, but, as the number of assets becomes large, obtaining accurate results by this method quickly becomes numerically too expensive. (Further, while Monte Carlo analysis can work well in analyzing the risk of moderately large portfolios, it is not easily incorporated in an optimization procedure with the purpose of finding a portfolio with desirable characteristics.)

The purpose of this article is to present a new approach for upper bounding the risk associated with a portfolio, for a given description of the uncertainty in the estimates of the first and second moments of the asset returns. We also show how to design portfolios that are robustly optimal, in the sense that they minimize this upper bound on risk. In fact, solving portfolio optimization problems with great precision, when the problem parameters (say μ and Σ) are not precisely known is not a reasonable proposition. A better approach is to explicitly account for such parameter uncertainty in the optimization, and to design a portfolio that performs reasonably for any set of parameters within the range of parameter uncertainty.

Consider an example of maximum risk analysis. Given an entry-wise description of uncertainty in the estimate of Σ, how large can the portfolio variance be? Assume we have the following information. We know an upper and lower bound on the variances of each asset, and an upper and lower bound on the covariances of each pair of assets. That is, for each entry Σ_{ij} of the covariance matrix we have an upper bound $\overline{\Sigma}_{ij}$ and a lower bound $\underline{\Sigma}_{ij}$. And, of course, we know that the covariance matrix must be positive semidefinite. The problem is then

$$\begin{array}{ll} \text{maximize} & w^T \Sigma\, w \\ \text{subject to} & \underline{\Sigma}_{ij} \leq \Sigma_{ij} \leq \overline{\Sigma}_{ij}, \quad i,j = 1,\ldots,n \\ & \Sigma \succeq 0, \end{array} \tag{1}$$

where $w \in \mathbf{R}^n$ is fixed, and $\Sigma \in \mathbf{R}^{n\times n}$ is the problem variable.

This is a convex program, with linear objective, and convex constraints. The positive semidefinite constraint is nonlinear and non-differentiable. The problem is, in fact, a *semi-definite program* (SDP), for which new optimization methods have been developed in recent years (Nesterov and Nemiroski's book [NN94] is a fundamental source, for an overview see Vandenberghe and Boyd [VB96], and software is now freely available [VB94, AHN$^+$97, Stu98].) With these methods, the global solution of many convex programs, including SDPs, can be efficiently computed. The computational effort is shown to grow polynomially with problem size, even for nonlinear, non-differentiable problems, making the methods suitable for large problems.

We are, in effect, departing from a Bayesian approach, in which the uncertainty about the distribution would be incorporated by defining a new distribution on the returns (with a larger variance.) An approach similar to the one we discuss here has, however, been used with great success in the field of robust control.

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 independently 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 Salomon Brothers [LBK96] is one of many sources for the downside-risk approach, which has been increasingly used in recent years (see also [Roy52]).

The RiskMetrics technical document [MR96], by J. P. Morgan and Reuters, provides an introductory overview of portfolio risk analysis. With the increased use of downside risk approaches, more attention is being paid to the effects of uncertainty in the covariance matrix estimates. A recent article by Ju and Pearson [JP99] in the *Journal of Risk* is titled "Using value-at-risk to control risk taking: how wrong can you be?" This article provides the tools to answer this question, and proposes robust optimization as a methology to deal with it.

In §2 we describe three worst-case portfolio analysis problems that can be solved via numerically efficient convex optimization methods. In §3 and §4 we look at other descriptions of the uncertainty set for the mean and covariance matrix that are convex and can be efficiently handled. In §5 we look at solution methods for the analysis problem. In additions to the general approach via interior-point methods, specialized projection methods can also be used. In §6 we describe the corresponding design problem, that is, finding a portfolio that has good worst-case performance. While software that can be used to solve the analysis problem is now widely available, the design problem requires specialized methods. In §7 we describe how the design problem can be efficiently solved using cutting plane methods.

2 The analysis problem

Assume the following data: a) the vector $w \in \mathbf{R}^n$ specifies the weights of the portfolio under analysis; b) the sets $\mathcal{M} \subset \mathbf{R}^n$ and $\mathcal{S} \subset \mathbf{R}^{n \times n}$ define the range of uncertainty for the estimates of the mean and covariance matrix of asset returns, that is, they define the values that μ and Σ may take. Consider the following problems.

- The *variance problem*: compute the worst-case variance of portfolio w, *i.e.*

$$\sup_{\Sigma \in \mathcal{S}} \quad w^T \Sigma w. \tag{2}$$

- The *downside risk problem*: compute the worst-case value at risk in portfolio w, for a confidence level of η, *i.e.*

$$\sup_{\mu \in \mathcal{M}, \Sigma \in \mathcal{S}} \quad \gamma \, (w^T \Sigma w)^{1/2} - \mu^T w - 1, \tag{3}$$

  with $\eta = \Phi(\gamma)$, where $\Phi(\cdot)$ is the c.d.f. of a zero norm, unit variance Gaussian random variable.

- The *tracking error problem*: compute the worst-case expected square tracking error of portfolio w relative to a reference portfolio with weights v, *i.e.*

$$\sup_{\mu \in \mathcal{M}, \Sigma \in \mathcal{S}} \quad (w - v)^T \left(\Sigma + \mu \mu^T \right) (w - v). \tag{4}$$

Note that we have assumed the uncertainty description for Σ and μ to be separable. Although we won't discuss it in this thesis, this doesn't have to be the case, and some joint descriptions can also be handled. As a second side note, the interpretation of problem (3) as the *value at risk* for a confidence level of η requires the assumption of jointly Gaussian distribution of returns. With proper care and adequate interpretation, the downside risk problem can still be used without this assumption. The derivations for problems 3 and 4 can be found, *e.g.*, in Lobo, Fazel and Boyd [LFB99].

The sets $\mathcal{M}$ and $\mathcal{S}$ can be obtained from statistic considerations (*i.e.*, from confidence intervals), and from assumptions about modeling errors. The process by which $\mathcal{M}$ and $\mathcal{S}$ are obtained, however, is not discussed in this thesis. Our focus is on identifying descriptions of $\mathcal{S}$ and $\mathcal{M}$ that can be handled with new efficient convex optimization methods. In particular, one critical element in any description of the set $\mathcal{S}$ is that all its elements must be positive semidefinite, since they must be valid covariance matrices. In this article we show how to handle problems with such a positivity constraint on Σ, as well as with other types of constraints, including box constraints on Σ and μ and ellipsoidal constraints on μ.

3 Uncertainty sets for the expected returns vector

The expected asset returns appear as a linear term in the downside risk problem, and as a quadratic term in the tracking error problem. We now show how to obtain solutions for the worst-case analysis of these terms, for two different types of uncertainty sets. In several cases an analytical solution exists.

3.1 Box constraint

Consider a set of entry-wise upper and lower bounds on μ (or box constraint, or ℓ_∞ constraint), *i.e.*

$$\underline{\mu}_i \leq \mu_i \leq \overline{\mu}_i, \quad i = 1, \dots, n.$$

With this constraint specifying the uncertainty in the expected asset returns, analytical solutions are easily derived for the terms in μ that appear in the worst-case problems. For the downside risk problem, which is linear in μ, the solution is:

$$\sup_{\mu \in \mathcal{M}} -\mu^T w \ = \ \overline{\mu}^T (w)_- - \underline{\mu}^T (w)_+,$$

where $(w)_+$ is a vector with entries $\max\{w_i, 0\}$, and $(w)_-$ is a vector with entries $\max\{-w_i, 0\}$. For the tracking error problem, which is quadratic in μ, the solution is:

$$\begin{aligned} &\sup_{\mu \in \mathcal{M}} (\mu^T (w - v))^2 \ = \\ &\ = \ \max \left\{ \left(\overline{\mu}^T (w - v)_+ - \underline{\mu}^T (w - v)_- \right)^2, \left(\underline{\mu}^T (w - v)_+ - \overline{\mu}^T (w - v)_- \right)^2 \right\}. \end{aligned}$$

3.2 Ellipsoidal constraint

A more interesting case is when $\mathcal{M}$ is an ellipsoid, *i.e.*,

$$\mathcal{M} = \{\mu : (\mu - \bar{\mu})^T S^{-1} (\mu - \bar{\mu}) \leq 1\}.$$

Note that for many statistical procedures M will indeed be an ellipsoid, with S collinear with Σ (*i.e.*, a scaled version). This is not the case if μ is inferred by some other process (such as some pricing model, or from information from analysts' reports). We now show how to incorporate worst-case ellipsoidal uncertainty in μ in a practical optimization program. We do not require an assumption of collinearity with Σ.

For the downside risk problem, which is linear in μ, we can easily compute the worst-case value analytically.

$$\lambda \;=\; \sup_{\mu\in\mathcal{M}} -\mu^T w \;=\; -\inf_{\mu\in\mathcal{M}} \mu^T w \;=\; -\inf_{\|S^{-1/2}\tilde{\mu}\|\leq 1} (\bar{\mu}+\tilde{\mu})^T w \;=\; -\bar{\mu}^T w - \inf_{\|z\|\leq 1} z^T S^{1/2} w,$$

where $S^{1/2}$ is the (symmetric, positive-semidefinite) matrix square root of S. The z that achieves the infimum is easily seen to be

$$z^* = -\frac{S^{1/2}w}{\|S^{1/2}w\|},$$

so that the value of the supremum is

$$\lambda \;=\; -\bar{\mu}^T w + \frac{w^T S w}{\sqrt{w^T S w}} \;=\; -\bar{\mu}^T w + \sqrt{w^T S w} \;=\; -\bar{\mu}^T w + \|S^{1/2} w\|.$$

Therefore, the solution to the downside risk problem (3) with ellipsoidal uncertainty in μ is obtained by adding the constant

$$-\bar{\mu}^T w + \|S^{1/2} w\| - 1$$

to the solution of the worst-case variance problem (2).

For the tracking error problem, which is quadratic in μ, we can also obtain a numerically tractable program. The derivation is bit more involved. To obtain the worst-case error we need to evaluate

$$\lambda = \sup_{\mu\in\mathcal{M}} (w-v)^T \mu\mu^T (w-v).$$

It is easily seen that λ is the smallest number that satisfies

$$\mu^T (w-v)(w-v)^T \mu \leq \lambda \quad \text{for all } \mu \text{ such that} \quad (\mu-\bar{\mu})^T S^{-1} (\mu-\bar{\mu}) \leq 1. \quad (5)$$

By using the $\mathcal{S}$-procedure [BEFB94, p.23], λ satisfies (5) if and only if there is a τ such that

$$\begin{bmatrix} -(w-v)(w-v)^T & 0 \\ 0 & \lambda \end{bmatrix} - \tau \begin{bmatrix} -S^{-1} & S^{-1}\bar{\mu} \\ \bar{\mu}^T S^{-1} & 1-\bar{\mu}^T S^{-1}\bar{\mu} \end{bmatrix} \succeq 0, \quad \tau \geq 0.$$

This can be rewritten as

$$\begin{bmatrix} \tau S^{-1} & -\tau S^{-1}\bar{\mu} \\ -\tau\bar{\mu}^T S^{-1} & \tau(\bar{\mu}^T S^{-1}\bar{\mu}-1)+\lambda \end{bmatrix} - \begin{bmatrix} w-v \\ 0 \end{bmatrix} \begin{bmatrix} (w-v)^T & 0 \end{bmatrix} \succeq 0, \quad \tau \geq 0,$$

and, with the Schur complement [BEFB94, p.7], we get the equivalent inequality

$$\begin{bmatrix} \tau S^{-1} & -\tau S^{-1}\bar{\mu} & w-v \\ -\tau\bar{\mu}^T S^{-1} & \tau(\bar{\mu}^T S^{-1}\bar{\mu}-1)+\lambda & 0 \\ (w-v)^T & 0 & 1 \end{bmatrix} \succeq 0, \quad \tau \geq 0.$$

This last formulation involves a positive semidefinite constraint, or linear matrix inequality, and can be efficiently and globally handled [VB96].

4 Uncertainty sets for the covariance matrix

We now turn to the specification of $\mathcal{S}$, that is, to the description of our knowledge of Σ (or, more accurately, the description of the uncertainty in our knowledge of Σ).

The variance and tracking error problems (2) and (4) are linear in Σ. The downside risk problem (3) can be cast as a program where Σ appears linearly:

$$\begin{array}{ll} \text{minimize} & t - \mu^T w \\ \text{subject to} & t^2 \leq \gamma^2 \, w^T \Sigma w \\ & \mu \in \mathcal{M}, \; \Sigma \in \mathcal{S}, \end{array}$$

where $t \in \mathbf{R}$ is an extra variable. The extra inequality is convex quadratic in t, and linear in Σ.

For a practical optimization method to be effective we want, if at all possible, a description of the uncertainty set $\mathcal{S}$ that leads to a convex program. All the problems in consideration are (or can be made) linear in Σ, and they will be convex if the set $\mathcal{S}$ is convex.

We have already introduced a positivity constraint and box constraints, both of which are convex. We represent the constraint that Σ be positive semidefinite by

$$\Sigma \succeq 0.$$

This constraint is required to ensure that all the Σ in the uncertainty set $\mathcal{S}$ are valid covariance matrices. It may, however, be omitted: a) if the other constraints are shown to define a subset of the positive semidefinite cone; or b) if Σ is parameterized in such a way that it is guaranteed to always be positive semidefinite (as will be discussed for factor models in §4.4). The most straightforward description of an uncertainty set for Σ is by adding to the positivity constraint a set of box constraints, on each entry Σ_{ij}:

$$\underline{\Sigma}_{ij} \leq \Sigma_{ij} \leq \overline{\Sigma}_{ij}$$

for $i = 1, \ldots, n$, $j = 1, \ldots, i$.

4.1 Second-moment of the Wishart distribution

If the returns are normally distributed, the estimate of covariance matrix has a Wishart distribution. The second-moment of the Wishart distribution can be used to define a region of confidence for the estimate. The moments of the Wishart distribution are given, *e.g.*, in Muirhead's book [Mui82, p.90]. This leads to an ellipsoidal constraint on the entries of the covariance matrix, *i.e.*,

$$(s - s_0)^T Q (s - s_0) \leq 1$$

where $s \in \mathbf{R}^{n(n+1)/2}$ is a vector representation of the (upper triangular) entries of Σ. This can also be written as a second-order cone constraint, using a square-root of Q. The size of this constraint can be quite large; Q has about $n^4/4$ entries. Note, however, that the matrix Q has a great deal of structure (again, see [Mui82]). Further work is required to determine how to exploit this structure for fast algorithms.

The region of confidence can be scaled by a factor that specifies the "conservativeness" (large factor) or "aggressiveness" (small factor) of the analysis. Under a probabilistic interpretation, the size of the ellispsoid corresponds to the confidence level, *i.e.*, to the probability of the covariance matrix being in the uncertainty set.

4.2 Constraints on the correlation coefficients

A more natural approach may be to write constraints not on the covariances, but on the correlation coefficents. The constraint for coefficient ij,

$$\underline{\rho}_{ij} \leq \rho_{ij} \leq \overline{\rho}_{ij},$$

is equivalent to

$$\underline{\rho}_{ij}\sqrt{\Sigma_{ii}\Sigma_{jj}} \leq \Sigma_{ij} \leq \overline{\rho}_{ij}\sqrt{\Sigma_{ii}\Sigma_{jj}}.$$

If Σ_{ii} and Σ_{jj} are constant (*i.e.*, if the variance of each asset is known exactly), this constraint is linear in Σ. In this case, it is equivalent to a box constraint.

In the general case, however, the constraint on the correlation coefficient is convex only if $\underline{\rho}_{ij} \leq 0$ and $\overline{\rho}_{ij} \geq 0$. In particular, it is then a convex hyperbolic constraint. By adding the auxiliary variable $t \in \mathbf{R}$, it can be written as one *second-order cone* (SOC) and two linear constraints:

$$\left\| \begin{bmatrix} 2\,t \\ \Sigma_{ii} - \Sigma_{jj} \end{bmatrix} \right\| \leq \Sigma_{ii} + \Sigma_{jj}, \quad \underline{\rho}_{ij} t \geq \Sigma_{ij} \geq \overline{\rho}_{ij} t.$$

4.3 Constraints on the variance of specific portfolios

Box constraints are a particular case of linear constraints on the covariance matrix. As an example of the use of other linear constraints, suppose we have better statistical information about the return of a given portfolio u (*e.g.*, the market portfolio). That is, we have a confidence interval for the variance of the portfolio return. We can then define the linear constraint

$$\underline{s} \leq u^T \Sigma\, u \leq \overline{s},$$

where $\underline{s}$ and $\overline{s}$ are the lower and upper bounds on the variance of portfolio u. This is only useful if it results in a tighter specification for Σ than, say, the confidence intervals for its individual entries. Of course, constraints such as this can be included simultaneously for any number of portfolios.

4.4 Factor models

Factor models, unfortunately, are not easily handled. Consider a one factor model, that is, a diagonal plus rank one covariance matrix:

$$\Sigma = \mathbf{diag}(d) + bb^T.$$

In general, if the uncertainty is specified by a set which is convex in the d_i and b_i, $i = 1, \ldots, n$ (*e.g.*, a box constraint), a non-convex quadratic problem results. (Note that, in this case, we can parameterize Σ in the variables d_i, b_i, $i = 1, \ldots, n$, and dispense with the positive-semidefinite constraint.)

5 Solving the analysis problem

If one of the objectives (2), (3), or (4) given in §2 is coupled with any of the convex uncertainty sets described in §3 and §4, a convex program results. While nonlinear and non-differentiable, these programs have a global optimum, and can be efficiently solved. Thus, we can treat problems like: find the worst case downside risk given box uncertainty in Σ; or, find the worst case tracking error given box uncertainty in Σ and ellipsoidal uncertainty in μ. We will discuss two solution techniques: semidefinite programming, and projection methods.

5.1 Solution by semidefinite programming

Convex programming methods developed in recent years can efficiently solve all the convex problems previously described, even for large problem sizes.

These methods are discussed, *e.g.*, in Vandenberghe and Boyd [VB96]. Currently available semidefinite programming software packages can handle problems of size over $n = 100$, on an inexpensive personal computer.

The computational complexity of these methods is provably polynomial in problem size. On the other hand, computing resources show no sign of departing from exponential growth and, as a consequence, within a few years semidefinite programs of size well into the thousands will be readily handled.

5.2 Solution by projection methods

There are alternative methods for the solution of the analysis problem that further exploit specific structure in the problem. These methods are based on the fact that it is easy to compute the projection on the sets defined by some of the constraints discussed above. Unlike in §5.1, this approach is not general, and depends on the particular problem under consideration. Its effectiveness relies on the projections on the objective level and constraint sets being computationally inexpensive.

For conciseness, we consider the example (1) given in the introduction, that is,

$$\begin{array}{ll} \text{maximize} & w^T \Sigma\, w \\ \text{subject to} & \underline{\Sigma}_{ij} \leq \Sigma_{ij} \leq \overline{\Sigma}_{ij}, \quad i,j = 1,\ldots,n \\ & \Sigma \succeq 0, \end{array} \tag{6}$$

where the variable is the symmetric matrix $\Sigma \in \mathbf{R}^{n \times n}$. This is a problem for which a projection method can be used, since all the required projections are readily computed.

Several optimization algorithms have been developed that are based on iterated projections (see, *e.g.*, Bertsekas [Ber95]). In practice, we have found a variation of the projection arc algorithm to be effective – but, as noted before, with the caveat that our numerical experience is not extensive.

We next show how to compute projections onto objective function level sets, and onto the sets defined by the two constraints in the above problem: the box constraint and the positivity constraint. All projections are given in the Frobenius norm (consistency of norms is required to ensure the convergence of a projection based method.)

We also describe the problem dual to (6) which, for some problems, is more readily solved.

5.2.1 Projection on the objective level set

The gradient of the variance objective function (6) is a rank-one matrix,

$$\nabla_\Sigma(w^T \Sigma\, w) = ww^T.$$

Hence, the Euclidean projection (*i.e.*, for matrix 2-norm) on the objective level set $\mathcal{V} = \{\Sigma : w^T \Sigma\, w \geq \sigma^2\}$ is given by

$$P_{\mathrm{obj}}(\Sigma) = \Sigma + \gamma w w^T, \qquad \text{where} \quad \gamma = \frac{\sigma^2 - w^T \Sigma w}{(w^T w)^2}.$$

It can be shown that this also yields the Frobenius metric projection.

5.2.2 Projection on the box constraint set

The Frobenius norm projection on the set defined by the box constraint (6) is easily seen to be

$$[P_{\text{box}}(\Sigma)]_{ij} = \begin{cases} \underline{\Sigma}_{ij}, & \Sigma_{ij} < \underline{\Sigma}_{ij} \\ \overline{\Sigma}_{ij}, & \Sigma_{ij} < \overline{\Sigma}_{ij} \\ \Sigma_{ij}, & \text{otherwise.} \end{cases}$$

5.2.3 Projection on the positive semidefinite cone

The projection of Σ on the positive semidefinite cone (6) is also easily computed. This projection is obtained by computing the eigenvalue decomposition of Σ and taking only the dyads with positive eigenvalues, *i.e.*,

$$P_{\text{psd}}(\Sigma) = U \operatorname{diag}\Big((\lambda_i(\Sigma))_+ \Big)\, U^T,$$

where $\Sigma = U \operatorname{diag}(\, \lambda_i(X)\,)\, U^T$ is the decomposition of Σ, and $(x)_+ = \max\{x, 0\}$. This projection is valid in both Euclidean and Frobenius metrics. (For proof, consider the problem of minimizing $\|X - \Sigma\|_{\text{F}}$ subject to $X \succeq 0$, and use the decomposition of Σ plus the invariance of the norm with respect to unitary transformations.)

5.2.4 Solution by Iterated projection

For a given value of the objective, σ^2, the problem can be determined to be feasible or infeasible by iterating the 3 projections, P_{obj}, P_{box}, and P_{psd}. The problem is feasible if the objective level set $\mathcal{V} = \{\Sigma : w^T \Sigma\, w \geq \sigma^2\}$

and the two sets defined by each of the problem constraints have non-empty intersection.

If the three sets do indeed have non-empty intersection (*i.e.*, if the optimal objective is higher than σ^2) the convexity of the sets and the fact that all projection are in the same (Frobenius) norm guarantees that the cycling of the projections will converge to a point in the intersection of the sets. Further, if the intersection of the sets is empty (*i.e.*, if the optimal objective is lower than σ^2), the cycling of the projections will converge to a fixed cycle.

This procedure can be repeated with bissection of σ^2, to find the optimal value of (6), which lies at the threshold between feasibility and infeasibility. We have then the following, simple algorithm:

1. Pick a value σ^2 for the objective, and cycle through the 3 projections: $S_1 = P_{\text{obj}}(\Sigma),\ S_2 = P_{\text{box}}(S_1),\ \Sigma = P_{\text{psd}}(S_2)$.

2. If σ^2 is less than the optimal value, step 1 will converge to a fixed point (Σ, S_1, S_2 equal to each other).
   If σ^2 is greater than the optimal value, step 1 will converge to a fixed cycle (Σ, S_1, S_2 repeating from cycle to cycle).

3. Repeat from step 1, with bisection on σ^2, to find the optimal value to the desired precision.

In practice, this is not an effective method. The convergence for the case when the objective value is feasible can be quite slow, which makes it hard to reliably detect infeasibility (since this detection relies on non-convergence to a point). Hence, what should be a lower bound for the bisection is easily mistaken for an upper bound, leading to erroneous results. A more effective method is presented next.

5.2.5 Solution by search along the arc of iterated projections

A more effective optimization method can be obtained by performing a line search along the gradient projection arc (see, *e.g.*, Bertsekas [Ber95, §2.3]). The idea is, from a given starting point, to scale the gradient of the objective and project it on the constraint set. Bisection is then performed on the gradient scaling, to find the projected point with the best objective value. This is then used as a new starting point for the next iteration. The procedure is repeated until convergence.

However, in our case the projection arc cannot be found easily. We easily project on each of two convex constraints (box and positive semidefinite), but no easy way to compute the projection on the intersection of these two constraints. We can, nevertheless, use the two projections iteratively to find an arc of feasible points (which we may call a "pseudo-projection arc"). With this modification, the method can still be shown to converge. We have then the following algorithm:

1. Given the current point Σ and gradient scaling κ, let $S := \Sigma + \kappa \nabla_\Sigma (w^T \Sigma w) = \Sigma + \kappa w w^T$.

2. Repeat $S := P_{\text{psd}}(P_{\text{box}}(S))$ until S is feasible.

3. Repeat steps 1 and 2 as required to perform bisection search on κ to maximize $w^T S w$.

4. Update the current point to be the result of the bissection search, $\Sigma := S$, and repeat from step 1 (until convergence of Σ).

5.2.6 The dual problem

In some instances it may be more effective to solve the dual problem. This should be the case, in particular, if the dual problem includes fewer constraints and the projections are therefore easier to apply. The dual of (6) is

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(\overline{\Sigma}\,\overline{\Lambda} - \underline{\Sigma}\,\underline{\Lambda}) \\ \text{subject to} & Z + \overline{\Lambda} - \underline{\Lambda} = -ww^T \\ & \overline{\Lambda}_{ij} \geq 0, \quad \underline{\Lambda}_{ij} \geq 0, \quad Z \succeq 0. \end{array} \tag{7}$$

If (6) is feasible and bounded (*i.e.*, has a finite optimal objective), then (7) is also finite and bounded, and the optimal objectives of the two problems are identical [VB96].

Note that this can be equivalently written without the variable Z, using the constraint $\overline{\Lambda} - \underline{\Lambda} + ww^T \succeq 0$. Alternatively, by using a translation to make the box constraints symmetric and using $\Lambda = \overline{\Lambda} - \underline{\Lambda}$, the number of dual variables can be reduced. The dual problem (7) is then equivalent to

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\left(\frac{\overline{\Sigma}+\underline{\Sigma}}{2}\,Z - \frac{\overline{\Sigma}-\underline{\Sigma}}{2}\,\Lambda\right) \\ \text{subject to} & \Lambda_{ij} \geq Z_{ij} + w_i w_j, \quad \Lambda_{ij} \geq -Z_{ij} - w_i w_j \\ & Z \succeq 0. \end{array} \tag{8}$$

Obvious extensions apply. For $\underline{\Sigma}_{ij} = \overline{\Sigma}_{ij}$ (*i.e.*, Σ_{ij} precisely known), the corresponding Λ_{ij} and the constraints associated with it are omitted. For Σ_{ij} unconstrained (*i.e.*, no knowledge about Σ_{ij}), we have that $\Lambda_{ij} = 0$, and the constraint $Z_{ij} = -w_i w_j$ is used.

6 The design problem

We now incorporate the previous analysis into portfolio optimization. The discussion so far has addressed the analysis problem, where the portfolio is fixed (and known). The corresponding design problem consists in selecting a portfolio using the worst-case risk as a criterion for the desirability of a particular portfolio. In the simplest form of the problem, the goal is to find the portfolio that is optimal in the sense of the previous analysis, that is, the portfolio that minimizes the worst-case risk. The minimization is subject to constraints, of course, such as an upper bound on the budget, a lower bound on the expected return, or others.

For simplicity, we consider worst-case design for the classical mean-variance tradeoff problem, with uncertainty in the covariance matrix Σ. The same approach can be used for the other problems introduced in §2.

Consider then the problem of selecting the portfolio with lowest risk (defined as the worst-case portfolio variance), subject to a lower bound on expected return, and subject to budget and shorting constraints:

$$\begin{array}{ll} \text{minimize} & \max_{\Sigma \in \mathcal{S}} (w^T \Sigma\, w) \\ \text{subject to} & \mathbf{1}^T w = 1, \quad w_i \geq w_{\min}, \quad \mu^T w \geq R_{\min}, \end{array} \tag{9}$$

where $\mathcal{S} = \{\, \Sigma \in \mathbf{R}^{n \times n} \mid \Sigma \succeq 0,\ \underline{\Sigma}_{ij} \leq \Sigma_{ij} \leq \overline{\Sigma}_{ij} \}$ is a convex compact set, so that the max is well defined. In practice, we will want to find the tradeoff curve between risk and return, which is obtained by solving this program ranging over different values of $R_{\min}$.

The uncertainty set can be scaled around some nominal point to make the design more "cautious" or more "aggressive". In the limit where the uncertainty set is reduced to a point (*i.e.*, to the nominal value of the mean and covariance of returns) the worst-case design reduces to the standard portfolio selection problem.

6.1 Min-max and max-min

Consider problem (9) in the form

$$\min_{w \in \mathcal{W}} \quad \max_{\Sigma \in \mathcal{S}} \quad w^T \Sigma\, w, \tag{10}$$

where $\mathcal{W} = \{\, w \in \mathbf{R}^n \mid \mathbf{1}^T w = 1,\ w_i \geq w_{\min},\ \mu^T w \geq R_{\min}\}$ is a convex compact set, so that the min is well defined. This is equivalent to the problem

$$\min_{C \succeq ww^T,\ w \in \mathcal{W}} \quad \max_{\Sigma \in \mathcal{S}} \quad \mathbf{Tr}(C\Sigma),$$

which has an inner product as objective function. We can now use the min-max theorem of game theory to switch the order of the min and max operators (see, *e.g.*, Bertsekas [Ber95, §5.4.3], or Luenberger [Lue69, §7.13]). The sets associated with both the min and max operators are convex. In particular, the newly added constraint is convex. Using the Schur complement, it is seen to be equivalent to

$$\begin{bmatrix} C & w \\ w^T & 1 \end{bmatrix} \succeq 0.$$

In the form stated the set associated with the min operator is not compact, as required by the theorem. However, since $\Sigma \succeq 0$ it is always possible to upper bound C without changing the solution of the problem, which makes the set compact (with the upper bound applied to, say, the spectral norm of C). Therefore, problem (10) is equivalent to

$$\begin{array}{ccc} \max & \min & w^T \Sigma\, w. \\ \Sigma \in \mathcal{S} & w \in \mathcal{W} & \end{array} \tag{11}$$

The design problem is convex overall, and we can equivalently solve it in either min-max or max-min form. In practice, the numerically most convenient form should be used.

7 Solving the design problem

In this section, we briefly indicate how the robust portfolio design problem can be effectively solved by analytic center cutting plane methods. For simplicity, the discussion here will focus on the variance problem. This approach is generalizable to the other robust portfolio problems. At the end of the section we briefly indicate alternative solution methods.

Define the function $\phi : \mathbf{R}^n \to \mathbf{R}$,

$$\phi(w) = \max_{\Sigma \in \mathcal{S}} \; w^T \Sigma\, w,$$

and denote the corresponding optimal variable by

$$\Sigma^*(w) = \underset{\Sigma \in \mathcal{S}}{\operatorname{argmax}} \; w^T \Sigma\, w.$$

For a convex $\mathcal{S}$, it is easily shown that the function

$$g(w) = 2\, \Sigma^*(w)\, w$$

is a subgradient of $\phi(w)$. Hence, once $\phi(w)$ has been computed (*i.e.*, the associated convex program has been solved) the subgradient is obtained at essentially no extra computational cost. The subgradient defines a supporting hyperplane for the sublevel set $\{v : \phi(v) \leq \phi(w)\}$ (the plane orthogonal to $g(w)$ that passes through w).

With $\mathcal{S} = \{\, \Sigma \in \mathbf{R}^{n\times n} \mid \Sigma \succeq 0,\ \underline{\Sigma}_{ij} \leq \Sigma_{ij} \leq \overline{\Sigma}_{ij}\}$ as before, problem (9) can be written

$$\begin{array}{ll} \text{minimize} & \phi(w) \\ \text{subject to} & w \in \mathcal{W}. \end{array}$$

Solving the mathematical program that computes $\phi(w)$ (and $\Sigma^*(w)$, $g(w)$) for a given w, provides an "oracle" for an analytic center cutting plane method, for which polynomial complexity has been established. For more on cutting plane methods see, *e.g.*, [Kel60, Nes95, GV99] (and the references therein.)

A method for solving this problem is as follows. Starting from any $w \in \mathcal{W}$:

1. Solve the inner problem: compute $\phi(w)$, a subgradient, and the corresponding supporting hyperplane.

2. Find the analytic center of the intersection of $\mathcal{W}$ with the half-planes defined by the previously computed supporting hyperplanes. Let w be the analytic center.

3. Repeat from 1, until w converges.

The overall complexity of this algorithm is polynomial in problem size. The cutting plane method is of polynomial complexity, and the same is true for a number of methods that solve the semidefinite program from which the cutting plane is obtained. Finding the analytic center requires the minimization of the logarithmic barrier function. In practice, the analytic centering problem does not have to be very precisely solved, which can save a significant amount of computational effort.

An alternative solution method exploits duality. The optimal value of the inner problem can equivalently be obtained by solving the associated dual problem. The min-max problem is then reduced to a minimization over all program variables. It is an open question whether this always results in a program that is easily handled by existing methods. In the case of (9), this approach results in an SDP, which is readily solved (Laurent El Ghaoui, personal communication.)

Another approach consists in developing self-concordant barrier functions for the min-max problems, allowing for the direct application of interior-point methods (Reha Tutuncu, personal communication.)

8 Conclusions

For many cases of interest, computing the maximum risk of a portfolio given uncertainty in the means and covariances of asset returns is a semidefinite programming problem. Its global solution can be efficiently computed, even for large problems. While not as general, this approach is more accurate and much faster than Monte Carlo methods. The computational effort required grows gracefully, which allows very large problems to be handled.

Also, solving portfolio optimization problems with great precision, when the problem parameters are inherently uncertain, is not a reasonable proposition. By using cutting plane methods, the worst-case risk analysis can be incorporated into portfolio selection, which allows "robust" portfolios to be designed.

Acknowledgements

Byunggyoo Kim implemented tests of the iterated projection algorithm. The authors would also like to acknowledge comments and sugestions from Maryam Fazel, Laurent El Ghaoui, Marc Nunes, Darrell Duffie, Bill Sharpe, John Watson, Michael Saunders, David Donoho, Stephen Wright, Lieven Vandenberghe, and Haitham Hindi.

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.

[BEFB94] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, June 1994.

[Ber95] D. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, Massachusetts, 1995.

[GV99] J.-L. Goffin and J.-P. Vial. Shallow, deep and very deep cuts in the analytic center cutting plane method. *Mathematical Programming*, 84(1):89–103, January 1999.

[JP99] X. Ju and N. D. Pearson. Using value-at-risk to control risk taking: how wrong can you be? *Journal of Risk*, 1(2):5–36, Winter 1999.

[Kel60] J. E. Kelley. The cutting-plane method for solving convex programs. *J. Soc. Indust. Appl. Math*, 8(4):703–712, December 1960.

[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.

[LFB99] M. S. Lobo, M. Fazel, and S. Boyd. Portfolio optimization with linear and fixed transaction costs and bounds on risk. Working paper, `http://www.stanford.edu/~boyd`, March 1999.

[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.

[Lue69] D. G. Luenberger. *Optimization By Vector Space Methods*. John Wiley & Sons, New York, N. Y., 1969.

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

[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.

[MR96] J. P. Morgan and Reuters. Riskmetrics – technical document. Technical report, December 1996. Fourth Edition.

[Mui82] Robb J. Muirhead. *Aspects of Multivariate Statistical Theory*. Wiley, 1982.

[Nes95] Y. Nesterov. Complexity estimates of some cutting plane methods based on the analytic barrier. *Mathematical Programming*, 69(1):149–176, July 1995.

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

[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.

[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.

[VB94] L. Vandenberghe and S. Boyd. sp*: Software for Semidefinite Programming. User's Guide, Beta Version.* Stanford University, October 1994. Available at `www.stanford.edu/~boyd`.

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