Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

MARC C. STEINBACH

Markowitz Revisited: Single-Period and Multi-Period Mean-Variance Models

Preprint SC 99-30 (August 1999)

Markowitz Revisited: Single-Period and Multi-Period Mean-Variance Models

Marc C. Steinbach

Abstract

Mean-variance portfolio analysis provided the first quantitative treatment of the trade-off between profit and risk. We investigate in detail the interplay between objective and constraints in a number of single-period variants, including semi-variance models. Particular emphasis is laid on avoiding the penalization of overperformance. The results are then used as building blocks in the development and theoretical analysis of multi-period models based on scenario trees. A key property is the possibility to remove surplus money in future decisions, yielding approximate downside risk minimization.

Key words. mean-variance analysis, downside risk, multi-period model

AMS subject classifications. 90A09, 90C15, 90C20

0 Introduction

The classical mean-variance approach for which Harry Markowitz received the 1990 Nobel Price in Economics offered the first systematic treatment of a dilemma that each investor faces: the conflicting objectives *high profit* versus *low risk*. In dealing with this fundamental issue Markowitz came up with a parametric optimization model that was both sufficiently general for a significant range of practical situations and simple enough for theoretical analysis and numerical solution. As the Swedish Academy of Sciences put it: his *primary contribution consisted of developing a rigorously formulated, operational theory for portfolio selection under uncertainty* [62].

Indeed, the subject is so complex that Markowitz' seminal work of the fifties [51, 52, 54] probably raised more questions than it answered, thus initiating a tremendous amount of related research. Before placing the present paper into perspective, the following paragraphs give a coarse overview of these issues. A substantial number of references is included, but the list is far from complete and cannot even contain all the relevant papers. We cite just a few references on each subject (in chronological order, and usually only the most recent from several contributions of the same author) to provide some starting points for the interested reader.

An important aspect of pareto-optimal (efficient) portfolios is that each determines a von Neumann–Morgenstern utility function [80] for which it maximizes the *expected utility* of the return on investment. This allowed Markowitz to interpret his approach by the theory of rational behavior under uncertainty [52], [54, Part IV]. Further, certain measures of *risk averseness* evolved as a basic concept in economic theory. These are derived from utility functions and justified by their relationship to the corresponding risk premiums, see Pratt [64], Arrow [2], Rubinstein [68], Duncan [17], Kihlstrom and Mirman [36], Ross [67], Li and Ziemba [49]. Applications of utility theory and risk averseness measures to portfolio selection are reported, e.g., by Mossin [59], Levy and Markowitz [48], Kroll et al. [46], Jewitt [31], King and Jensen [40], Kijima and Ohnishi [37].

A fundamental (and still debated) question is how risk should be measured properly. Markowitz discusses the pros and cons of replacing the variance by alternative risk measures in a more general mean-risk approach [54, Chap. XIII]. These considerations and the theory of stochastic dominance (see Bawa [5], Fishburn [20], Levy [47], Kijima and Ohnishi [38]) stimulated the research in asymmetric risk measures like expectation of loss and semi-variance, cf. Konno [41], King [39], Zenios and Kang [83], Uryasev and Rockafellar [79]. The properties of real return distributions also led to risk models involving higher moments, see Ziemba [84], Kraus and Litzenberger [45], Konno and Suzuki [43]. More recently the theoretical concept of *coherent* risk measures was introduced by Artzner et al. [3], while portfolio tracking (or replication) approaches became popular in practice, see King [39], Konno and Watanabe [44], Dembo and Rosen [15].

It is quite interesting that the mean-variance approach has received very little attention in the context of long-term investment planning. Although Markowitz does consider true multi-period models (where the portfolio may be readjusted several times during the planning horizon) [54, Chap. XIII], these considerations use a utility function based on the consumption of wealth over time rather than mean and variance of the final wealth, which places the problem in the realm of dynamic programming (Bellman [6]). Further long-term and simplified multi-period approaches are discussed, e.g., by Mossin [59], Samuelson [71], Hakansson [24], Merton and Samuelson [57], Konno et al. [42]. Much research has also been carried out in the closely related field of continuous-time models, see Merton [56], Harrison and Pliska [25], Heath et al. [27], Karatzas [34], Dohi and Osaki [16]. Over the past decade more detailed multi-period models have become tractable due to the progress in computing technology (both algorithms and hardware), see, e.g., Mulvey and Vladimirou [60], Dantzig and Infanger [14], Consigli and Dempster [13], Beltratti et al. [7].

With the exception of this group, most of the work cited above neglects details like asset liquidity or transaction costs. At least the second idealization causes serious errors when many transactions are performed, as in continuous-time models. Imperfect markets are briefly discussed by Markowitz [54, p. 297+]; later studies include Perold [63], He and Pearson [26], Karatzas et al. [35], Jacka [30], Shirakawa [73], Morton and Pliska [58], Atkinson et al. [4].

A final issue concerns the assumptions of the investor about the future, which is represented by probability distributions of the asset returns. Being based on assessments of financial analysts or estimated from historical data (or both), these distributions are never exact. (Markowitz calls them *probability beliefs*.) The question of the sensitivity of optimization results with respect to errors in the distribution is discussed, e.g., by Jobson [32], Broadie [11], Chopra and Ziemba [12], Best and Ding [8], MacLean and Weldon [50].

Additional material and references are found in a more recent book by Markowitz [55] and any standard text on mathematical finance, like Sharpe [72], Elton and Gruber [18], Ingersoll [29], Alexander and Sharpe [1], Zenios [82], Ziemba and Mulvey [85].

The present paper develops a fairly complete theoretical understanding of the multi-period mean-variance approach based on scenario trees. This is achieved by analyzing various portfolio optimization problems with gradually increasing complexity. Primal and dual solutions of these problems are derived, and dual variables are given an interpretation if possible. The most important aspect in our discussion is the precise interaction of objective (or risk measure) and constraints (or set of feasible wealth distributions), a subject that has not much been studied in the previous literature. It should be clear that arguing the properties of risk measures may be meaningless in an optimization context unless it is clear which distributions are possible. A specific goal in our analysis is to avoid a penalization of overperformance. In this context we discuss the role of cash and, in some detail, variance versus semi-variance. A key ingredient of our most complex multi-period model is an artificial arbitrage-like mechanism involving riskless though inefficient portfolios and representing a choice between immediate consumption or future profit.

Each of the problems considered tries to isolate a certain aspect, usually under the most general conditions even if practical situations typically exhibit more specific characteristica. However, we give higher priority to a clear presentation, and inessential generality will sometimes be sacrificed for technical simplicity. In particular, no inequality constraints are included unless necessary. (A separate section is devoted to the influence of such restrictions.) Neither do we attempt to model liquidity constraints or short-selling correctly, or to include transaction costs; we consider only idealized situations without further justification. The work here is based on a multi-period mean-variance model that was first proposed by Frauendorfer [21] and later refined by Frauendorfer and Siede [23]. A complete application model (including transaction costs and market restrictions) together with implications of the results developed here will be presented later in a joint paper.

Due to future uncertainty the portfolio optimization problems in this paper are all stochastic. More precisely, they are *deterministic equivalents* of convex stochastic programs, cf. Wets [81]. Except for the semi-variance problems they are also quadratic programs involving a second-order approximation of the return distribution in some sense, cf. Samuelson [70]. Based on earlier work in nonlinear optimal control [74, 75, 78], the author has developed structure-exploiting numerical algorithms for multi-stage convex stochastic programs like the ones discussed here [76, 77]. More general problem classes and duality are studied by Rockafellar and Wets [66], and Rockafellar [65]. For background material on stochastic programming we refer the reader to Ermoliev and Wets [19], Kall and Wallace [33], Birge [10], Birge and Louveaux [9], Ruszczyński [69].

The paper is organized as follows. Our analysis begins with single-period models in Section 1. Although many of the results are already known, the systematic discussion of subtle details adds insight that is essential in the multi-period case. To some extent this section has tutorial character; the problems may serve as examples in an introductory course on optimization. Next, multi-period mean-variance models are analyzed in Section 2, where the final goal consists in constructing an approximate downside risk minimization through appropriate constraints. This material is completely new; the research was motivated by practical experience with the application model mentioned above. Some concluding remarks are finally given in Section 3.

1 Single-period mean-variance analysis

Consider an investment in n assets over a certain period of time. Denote by x^ν the capital invested in asset ν, by $x \in \mathbb{R}^n$ the *portfolio* vector, and by $r \in \mathbb{R}^n$ the random vector of asset returns, yielding asset capitals $r^\nu x^\nu$ at the end of the investment period. Suppose that r is given by a joint probability distribution with expectation $\bar{r} := \mathbf{E}(r)$ and covariance matrix

$$\Sigma := \mathbf{E}[(r-\bar{r})(r-\bar{r})^*] = \mathbf{E}[rr^*] - \bar{r}\bar{r}^*.$$

(The existence of these two moments is assumed throughout the paper.) The choice of a specific portfolio determines a certain distribution of the associated total return (or final wealth) $w \equiv r^*x$. Mean-variance analysis aims at forming the most desirable return distribution through a suitable portfolio, where the investor's idea of desirability depends solely on the first two moments.

Definition 1 (Reward). The reward of a portfolio is the mean of its return,

$$\rho(x) := \mathbf{E}(r^*x) = \bar{r}^*x.$$

Definition 2 (Risk). The risk of a portfolio is the variance of the return,

$$R(x) := \sigma^2(r^*x) = \mathbf{E}[(r^*x - \mathbf{E}(r^*x))^2] = \mathbf{E}[x^*(r-\bar{r})(r-\bar{r})^*x] = x^*\Sigma x.$$

Various formulations of the mean-variance problem exist. From the very beginning [52, 54], Markowitz related his approach to the utility theory of von Neumann and Morgenstern [80]. As we shall see later, maximizing the expectation of a concave quadratic utility function leads to a formulation like

$$\begin{array}{ll} \max\limits_x & \mu\rho(x) - \frac{1}{2}R(x) \\ \text{s.t.} & e^*x = 1, \end{array} \tag{1}$$

where $e \in \mathbb{R}^n$ denotes the vector of all ones. The objective models the actual goal of the investor, a tradeoff between risk and reward,[1] while the *budget equation* $e^*x = 1$ simply specifies her initial wealth w_0 (normalized to $w_0 = 1$ w.l.o.g.). Our preferred formulation comes closer to the original

[1] Many authors attach a tradeoff parameter θ to the *risk* term and maximize $\rho(x) - \theta R(x)/2$, which is equivalent to (1) if $\mu \equiv \theta^{-1} > 0$. However, this problem becomes unbounded for $\theta \le 0$, whereas (1) remains solvable for $\mu \le 0$. This is important in our analysis.

one; it minimizes risk subject to the budget equation and subject to the condition that a certain desired reward ρ be obtained,

$$\begin{aligned} \min_x \quad & \frac{1}{2} R(x) \\ \text{s.t.} \quad & e^* x = 1, \\ & \rho(x) = \rho. \end{aligned} \tag{2}$$

Here the investor's goal is *split* between objective and reward condition.

In this section we study the precise relation of Problems (1) and (2), and a number of increasingly general single-period variants. We will include a cash account, then consider certain inequality constraints, utility functions, and finally downside risk. Many of the results are already known, but usually in a different form. Here we choose a presentation that facilitates the study of nuances in the optimization problems and that integrates seamlessly with the more general case of multi-period problems in Section 2.

1.1 Risky assets only

The simplest situation is given by portfolios consisting exclusively of risky assets. In this case we impose two conditions on the return distribution.

Basic assumptions.

A1) The covariance matrix is positive definite, $\Sigma > 0$.

A2) $\bar{r}$ is not a multiple of e.

Remarks. The first assumption means that all n assets (and any convex combination) are indeed risky; riskless assets like cash will be treated separately if present. The second assumption implies $n \geq 2$ and guarantees a non-degenerate situation: otherwise Problem (1) would always have the same optimal portfolio $x = \Sigma^{-1} e / (e^* \Sigma^{-1} e)$ regardless of the tradeoff parameter, and Problem (2) would have contradicting constraints except for one specific value of the desired reward: $\rho = \bar{r}^* e / n$. Notice that no formal restrictions are imposed on the value of $\bar{r}$, although $\bar{r} > 0$ (and even $\bar{r} > e$) will usually hold in practice.

Due to assumption A1 we can define the following constants that will be used throughout this section:

$$\alpha := e^* \Sigma^{-1} e, \quad \beta := e^* \Sigma^{-1} \bar{r}, \quad \gamma := \bar{r}^* \Sigma^{-1} \bar{r}, \qquad \delta := \alpha\gamma - \beta^2.$$

Lemma 1. *The constants α, γ, and δ are positive. More precisely,*

$$\alpha \in \left[\frac{n}{\lambda_{\max}(\Sigma)}, \frac{n}{\lambda_{\min}(\Sigma)} \right], \quad \gamma \in \left[\frac{\|\bar{r}\|_2^2}{\lambda_{\max}(\Sigma)}, \frac{\|\bar{r}\|_2^2}{\lambda_{\min}(\Sigma)} \right], \quad |\beta| < \frac{\sqrt{n}\, \|\bar{r}\|_2}{\lambda_{\min}(\Sigma)},$$

where $\lambda_{\min}, \lambda_{\max}$ denote the minimal and maximal eigenvalue, respectively.

Proof. Since $\Sigma > 0$ (by A1), we have

$$\alpha = e^* \Sigma^{-1} e \in \left[\|e\|_2^2 \, \lambda_{\min}(\Sigma^{-1}), \|e\|_2^2 \, \lambda_{\max}(\Sigma^{-1}) \right] = \left[\frac{n}{\lambda_{\max}(\Sigma)}, \frac{n}{\lambda_{\min}(\Sigma)} \right].$$

The inclusion for γ is analogous. Since $\bar{r}, e$ are linear independent and $\Sigma > 0$ (by A2 and A1), the (2,2)-matrix

$$\begin{pmatrix} e^* \\ \bar{r}^* \end{pmatrix} \Sigma^{-1} \begin{pmatrix} e & \bar{r} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} > 0$$

has positive determinant δ. Thus $|\beta| < \sqrt{\alpha\gamma} \leq \sqrt{n}\, \|\bar{r}\|_2 / \lambda_{\min}(\Sigma)$. $\square$

Remark. The inclusions for α and γ are sharp but not the bound on $|\beta|$, and neither $\alpha < \gamma$ nor $\beta > 0$ hold in general. Anyway, we need only $\alpha, \gamma, \delta > 0$.

Problem 1. Let us first consider the standard tradeoff formulation. To simplify the comparison with our preferred formulation, we minimize negative utility

$$\begin{aligned} \min_x \quad & \frac{1}{2} x^* \Sigma x - \mu \bar{r}^* x \\ \text{s.t.} \quad & e^* x = 1. \end{aligned}$$

The Lagrangian is

$$L(x, \lambda; \mu) = \frac{1}{2} x^* \Sigma x - \mu \bar{r}^* x - \lambda(e^* x - 1).$$

Theorem 1. *Problem 1 has the unique primal-dual solution*

$$x = \Sigma^{-1}(\lambda e + \mu \bar{r}), \qquad \lambda = (1 - \mu\beta)/\alpha$$

and associated reward

$$\rho = \lambda\beta + \mu\gamma = (\beta + \mu\delta)/\alpha.$$

Proof. From the Lagrangian one obtains the system of first order necessary conditions

$$\begin{pmatrix} \Sigma & e \\ e^* & \end{pmatrix} \begin{pmatrix} x \\ -\lambda \end{pmatrix} = \begin{pmatrix} \mu \bar{r} \\ 1 \end{pmatrix}.$$

Its first row (dual feasibility) yields the optimal portfolio x. The optimal multiplier λ and reward ρ are obtained by substituting x into the second row (primal feasibility) and the definition of ρ, respectively. Uniqueness of the solution follows from strong convexity of the objective and full rank of the constraint. □

Remark. Although the qualitative interpretation of the tradeoff function is clear, the precise value of the tradeoff parameter μ should also have an interpretation. In specific, the resulting reward is of interest. This is one reason why we prefer a different formulation of the mean-variance problem. Other important reasons are greater modeling flexibility, and sparsity in the multi-period formulation, see Section 2.

Problem 2. The mean-variance problem with prescribed reward reads

$$\begin{aligned} \min_x \quad & \frac{1}{2} x^* \Sigma x \\ \text{s.t.} \quad & e^* x = 1, \\ & \bar{r}^* x = \rho. \end{aligned}$$

Its Lagrangian is

$$L(x, \lambda, \mu; \rho) = \frac{1}{2} x^* \Sigma x - \lambda(e^* x - 1) - \mu(\bar{r}^* x - \rho).$$

We refer to the dual variables λ, μ as *budget multiplier* and *reward multiplier*, respectively. It will soon be shown that the optimal reward multiplier μ is precisely the tradeoff parameter of Problem 1.

Theorem 2. *Problem 2 has the unique primal-dual solution*

$$x = \Sigma^{-1}(\lambda e + \mu \bar{r}), \qquad \lambda = (\gamma - \beta\rho)/\delta, \quad \mu = (\alpha\rho - \beta)/\delta.$$

Proof. The system of first-order optimality conditions reads

$$\begin{pmatrix} \Sigma & e & \bar{r} \\ e^* & & \\ \bar{r}^* & & \end{pmatrix} \begin{pmatrix} x \\ -\lambda \\ -\mu \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \rho \end{pmatrix}.$$

As in Theorem 1, the optimal portfolio x is obtained from the first row. Substitution of x into rows two and three yields the optimal multipliers

$$\begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \left[\begin{pmatrix} e^* \\ \bar{r}^* \end{pmatrix} \Sigma^{-1} \begin{pmatrix} e & \bar{r} \end{pmatrix} \right]^{-1} \begin{pmatrix} 1 \\ \rho \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ \rho \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} \gamma - \beta\rho \\ \alpha\rho - \beta \end{pmatrix}.$$

Uniqueness of the solution follows as in Theorem 1. □

Theorem 3. *Problem 1 with parameter μ and Problem 2 with parameter ρ are equivalent if and only if μ equals the optimal reward multiplier of Problem 2 or, equivalently, ρ equals the optimal reward of Problem 1.*

Proof. The required conditions, $\mu = (\alpha\rho - \beta)/\delta$ and $\rho = (\beta + \mu\delta)/\alpha$, are clearly equivalent. It follows that the optimal budget multipliers of both problems are identical,

$$\frac{1 - \mu\beta}{\alpha} = \frac{\delta - \alpha\beta\rho + \beta^2}{\alpha\delta} = \frac{\alpha\gamma - \alpha\beta\rho}{\alpha\delta} = \frac{\gamma - \beta\rho}{\delta}.$$

Hence optimal portfolios also agree. The "only if" direction is trivial. □

Remarks. Apparently, the optimality conditions of Problem 2 include the optimality conditions of Problem 1, and additionally the reward condition. These $n+2$ equations define a one-dimensional affine subspace for the $n+3$ variables x, λ, μ, ρ, which is parameterized by μ in Problem 1 and by ρ in Problem 2. As an immediate consequence, the optimal risk is a quadratic function of ρ, denoted as $\sigma^2(\rho)$. Its graph is called the *efficient frontier.*[2]

Theorem 4. *In Problems 1 and 2, the optimal risk is*

$$\sigma^2(\rho) = (\alpha\rho^2 - 2\beta\rho + \gamma)/\delta = (\mu^2\delta + 1)/\alpha.$$

Its global minimum over all rewards is attained at $\hat{\rho} = \beta/\alpha$ and has the positive value $\sigma^2(\hat{\rho}) = 1/\alpha$. The associated solution is $\hat{x} = \Sigma^{-1}e/\alpha$, $\hat{\lambda} = 1/\alpha$, $\hat{\mu} = 0$.

Proof. By Definition 2 and Theorem 2,

$$\begin{aligned} \sigma^2(\rho) &= x^*\Sigma x = (\lambda e + \mu\bar{r})^*\Sigma^{-1}(\lambda e + \mu\bar{r}) \\ &= \lambda^2\alpha + 2\lambda\mu\beta + \mu^2\gamma = \lambda(\lambda\alpha + \mu\beta) + \mu(\lambda\beta + \mu\gamma). \end{aligned}$$

Using λ, ρ from Theorem 1 and λ, μ from Theorem 2 gives

$$\sigma^2(\rho) = \lambda + \mu\rho = (\alpha\rho^2 - 2\beta\rho + \gamma)/\delta = (\mu^2\delta + 1)/\alpha.$$

The remaining statements follow trivially. □

Discussion. The optimal portfolio is clearly a reward-dependent linear combination of the reward-independent portfolios $\Sigma^{-1}e$ and $\Sigma^{-1}\bar{r}$. Moreover, it is an affine function of ρ. The efficient frontier and optimal investments into two risky assets are depicted in Fig. 1. Here, since $n = 2$, the optimal portfolio is completely determined by the budget condition and the reward condition; it does not depend on Σ and is thus correlation-independent. Not so the risk: for negatively correlated assets,

[2]More generally, *efficient frontier* refers to the set of all pareto-optimal solutions in any multi-objective optimization problem. The solutions (portfolios) are also called efficient. Strictly speaking this applies only to the *upper branch* here, that is, $\rho \geq \hat{\rho}$ or, equivalently, $\mu \geq 0$ (see discussion below).

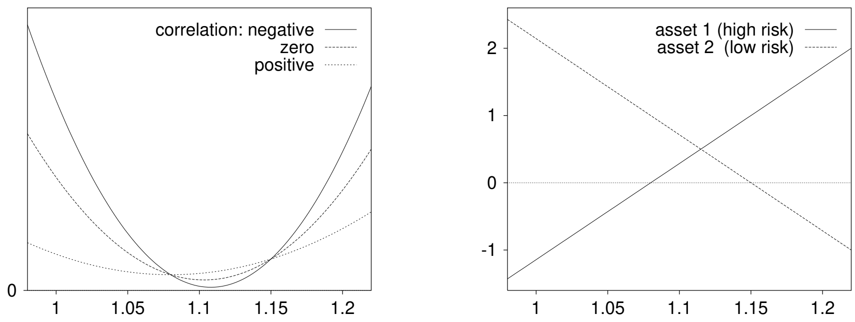

Figure 1: Portfolio with two risky assets; $\bar{r}^1 = 1.15$, $\bar{r}^2 = 1.08$. Left: efficient frontier for negatively correlated, uncorrelated, and positively correlated assets. Right: optimal portfolio vs. reward.

it has a pronounced minimum at a fairly large reward $\hat{\rho}$. As the correlation increases, the lowest possible risk also increases and is attained at a smaller reward. (These statements do not simply generalize to the case $n > 2$.)

A serious drawback of the model (in this form) is the fact that positive deviations from the prescribed reward are penalized, and hence the "risk" increases when ρ is reduced below $\hat{\rho}$. Indeed, the penalization cannot be avoided, indicating that the model is somehow incomplete. We will see, however, that unneccessary positive deviations from ρ do not occur if the model is extended appropriately. For the moment let us accept that only the upper branch is relevant in practice.

1.2 Risky assets and riskless cash

Now consider n risky assets and an additional cash account x^c with deterministic return $r^c \equiv \bar{r}^c$. The portfolio is (x, x^c), and $x, r, \bar{r}, \Sigma$ refer only to its risky part.

Basic assumptions. Assumption A2 is replaced by a similar condition on the extended portfolio, which may now consist of just one risky asset and cash.

A1) $\Sigma > 0$.

A3) $\bar{r} \neq r^c e$.

Remarks. Again, the second assumption excludes degenerate situations, and no restrictions are imposed to ensure realistic returns. In practice one can typically assume $\bar{r} > r^c e > 0$ (or even $r^c e \geq e$), which satisfies A3. The constants α, β, γ are defined as before; they are related to the risky part of the portfolio only. (Replacing A2 makes $\delta = 0$ possible, but δ plays no role here.)

Problem 3. Any covariance associated with cash vanishes, so that the risk and reward are $R(x, x^c) = x^* \Sigma x$ and $\rho(x, x^c) = \bar{r}^* x + r^c x^c$, respectively, and the optimization problem reads

$$\begin{aligned} \min_{x, x^c} \quad & \frac{1}{2} \begin{pmatrix} x \\ x^c \end{pmatrix}^* \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ x^c \end{pmatrix} = \frac{1}{2} x^* \Sigma x \\ \text{s.t.} \quad & e^* x + x^c = 1, \\ & \bar{r}^* x + r^c x^c = \rho. \end{aligned}$$

Theorem 5. *Problem 3 has the unique primal-dual solution*

$$\begin{aligned} x = \Sigma^{-1}(\lambda e + \mu \bar{r}) = \mu \Sigma^{-1}(\bar{r} - r^c e), \qquad & \lambda = -r^c \mu, \\ x^c = 1 - \mu(\beta - r^c \alpha), \qquad & \mu = (\rho - r^c)/\delta^c, \end{aligned}$$

where $\delta^c := (r^c)^2 \alpha - 2 r^c \beta + \gamma > 0$. *The resulting optimal risk is*

$$\sigma^2(\rho) = (\rho - r^c)^2 / \delta^c.$$

Its global minimum over all rewards is attained at $\hat\rho = r^c$ and has value zero. The associated solution has 100 % cash: $(\hat x, \hat x^c) = (0,1)$, $\hat\lambda = \hat\mu = 0$.

Proof. The system of optimality conditions is

$$\begin{pmatrix} \Sigma & 0 & e & \bar r \\ 0 & 0 & 1 & r^c \\ e^* & 1 & & \\ \bar r^* & r^c & & \end{pmatrix} \begin{pmatrix} x \\ x^c \\ -\lambda \\ -\mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \rho \end{pmatrix}.$$

The optimal budget multiplier λ is obtained from row two. Substitution into row one yields the expression for x, and substituting x into row three yields x^c. Substitution of x and x^c into row four gives

$$\rho = \bar r^* x + r^c x^c = \mu(\gamma - r^c\beta) + r^c - \mu(r^c\beta - (r^c)^2\alpha) = r^c + \mu\delta^c,$$

yielding μ. The positivity of δ^c follows (with A3) from

$$\delta^c = (\bar r - r^c e)^* \Sigma^{-1} (\bar r - r^c e).$$

Finally, the second formula for x yields

$$\sigma^2(\rho) = \mu^2 (\bar r - r^c e)^* \Sigma^{-1} (\bar r - r^c e) = \mu^2 \delta^c = (\rho - r^c)^2/\delta^c.$$

The remaining statements ($\hat\rho = r^c$ etc.) follow trivially. □

Problem 4. Problem 3 also has a tradeoff version:

$$\begin{aligned} \min_{x,x^c} \quad & \frac{1}{2} x^* \Sigma x - \mu(\bar r^* x + r^c x^c) \\ \text{s.t.} \quad & e^* x + x^c = 1. \end{aligned}$$

Theorem 6. *Problem 3 with parameter ρ and Problem 4 with parameter μ are equivalent if and only if $\rho = r^c + \mu\delta^c$.*

Proof. The proof is analogous to the proof of Theorem 3 and therefore omitted. □

Discussion. Basically the situation is quite similar to Problem 2, the only qualitative difference being the existence of one zero risk portfolio: for $\rho = r^c$, the capital is completely invested in cash and the risk vanishes. Otherwise a fraction of $e^* x = \mu(\beta - r^c\alpha)$ is invested in risky assets and the risk is positive, see Fig. 2. The optimal portfolio is now a mix of the (reward-independent) risky portfolio $(\Sigma^{-1}(\bar r - r^c e), 0)$ and cash (0,1). The following comparison shows how precisely the cash account reduces risk when added to a set of (two or more) risky assets.

Theorem 7. *The risk in Problem 3 is almost always lower than in Problem 2: If $\beta \neq r^c\alpha$, then the efficient frontiers touch in the single point*

$$\rho = r^c + \frac{\delta^c}{\beta - r^c\alpha} = \frac{\gamma - r^c\beta}{\beta - r^c\alpha}, \qquad \sigma^2(\rho) = \frac{\delta^c}{(\beta - r^c\alpha)^2},$$

(see Fig. 2), where the solutions of both problems are "identical": x or $(x,0)$. If $\beta = r^c\alpha$, then $x^c \equiv 1$ and $e^* x \equiv 0$, and risks differ by the constant $1/\alpha$,

$$\frac{(\rho - r^c)^2}{\delta^c} + \frac{1}{\alpha} = \frac{\alpha\rho^2 - 2\beta\rho + \gamma}{\delta}.$$

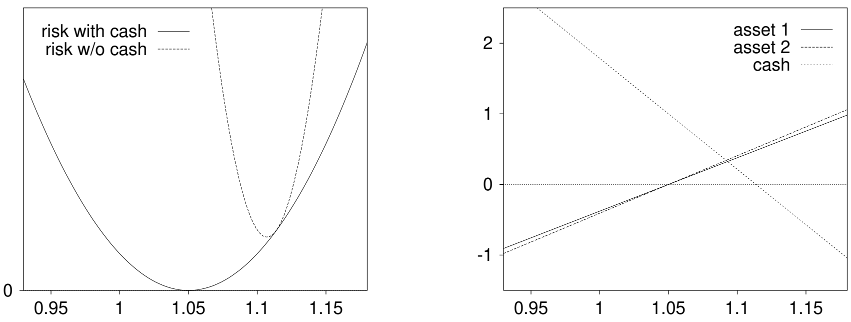

Figure 2: Portfolio with two risky assets and cash; $\bar{r}^1, \bar{r}^2$ as before, $r^c = 1.05$. Left: efficient frontiers with/without cash. Right: optimal portfolio vs. reward.

Proof. If $\beta \neq r^c\alpha$, then Problem 3 has a unique zero-cash solution, $x^c = 0$, with

$$\mu = \frac{1}{\beta - r^c\alpha}, \qquad \lambda = -\frac{r^c}{\beta - r^c\alpha}.$$

This gives the stated values of ρ and $\sigma^2(\rho)$ by Theorem 5. Substituting ρ into the formulae for λ, μ in Theorem 2 yields identical values in both problems. Hence the portfolios agree, too. The curvatures of the efficient frontiers, $d^2\sigma^2(\rho)/d\rho^2$, are $2\alpha/\delta$ and $2/\delta^c$, respectively. Now,

$$\alpha\delta^c - \delta = (r^c)^2\alpha^2 - 2r^c\alpha\beta + \alpha\gamma - \alpha\gamma + \beta^2 = (r^c\alpha - \beta)^2 > 0.$$

Thus $2\alpha/\delta > 2/\delta^c > 0$, implying that Problem 3 has lower risk if $x^c \neq 0$. The case $\beta = r^c\alpha$ is trivial: both efficient frontiers have $\hat{\rho} = r^c$ and identical curvatures. $\square$

To conclude this section, we show that it does not make sense to consider portfolios with more than one riskless asset (and no further restrictions).

Lemma 2 (Arbitrage). *Any portfolio having at least two riskless assets x^c, x^d with different returns r^c, r^d can realize any desired reward at zero risk.*

Proof. Choose $x^c = (\rho - r^d)/(r^c - r^d)$, $x^d = 1 - x^c$, and invest nothing in other assets. $\square$

1.3 Risky assets, cash, and guaranteed total loss

Let us now consider a portfolio with $n \geq 1$ risky assets, a riskless cash account as in Problem 3, and in addition an "asset" x^l with guaranteed total loss, i.e., $r^l \equiv \bar{r}^l = 0$. (Notice that x^l is not "risky" in the sense of an uncertain future.) At first glance this situation seems strange, but it will turn out to be useful.[3]

Basic assumptions. In addition to the conditions of the previous section we now require positive cash return ($r^c \leq r^l$ does not make sense).

A1) $\Sigma > 0$.

A3) $\bar{r} \neq r^c e$.

A4) $r^c > 0$.

[3]The suggestive (but actually misleading) notion of an "asset with guaranteed total loss" is due to Infanger [28] who constructed an example problem similar to the one considered here.

Problem 5. All covariances associated with x^c or x^l vanish, so that the risk and reward are $R(x, x^c, x^l) = x^*\Sigma x$ and $\rho(x, x^c, x^l) = \bar{r}^* x + r^c x^c$, respectively, and the optimization problem reads

$$\begin{aligned} \min_{x,x^c,x^l} \quad & \frac{1}{2}\begin{pmatrix} x \\ x^c \\ x^l \end{pmatrix}^* \begin{pmatrix} \Sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ x^c \\ x^l \end{pmatrix} = \frac{1}{2} x^* \Sigma x \\ \text{s.t.} \quad & e^* x + x^c + x^l = 1, \quad x^l \geq 0, \\ & \bar{r}^* x + r^c x^c = \rho. \end{aligned}$$

Note that the no-arbitrage condition $x^l \geq 0$ must be imposed; otherwise one could borrow arbitrary amounts of money without having to repay. However, Lemma 2 still works for *sufficiently small* ρ. This is precisely our intention.

Theorem 8. *Problem 5 has unique primal and dual solutions x, x^c, x^l, λ, μ, η, where η is the multiplier of the nonnegativity constraint $x^l \geq 0$. For $\rho > r^c$, the optimal solution has $x^l = 0$ and $\eta = -\lambda > 0$, and is otherwise identical to the solution of Problem 3. Any reward $\rho \leq r^c$ is obtained at zero risk by investing in a linear combination of the two riskless assets, with primal-dual solution*

$$x = 0, \quad x^c = \frac{\rho}{r^c}, \quad x^l = 1 - \frac{\rho}{r^c}, \qquad \lambda = \mu = \eta = 0.$$

Proof. The system of necessary conditions can be written

$$\begin{pmatrix} \Sigma & 0 & 0 & e & \bar{r} \\ 0 & 0 & 0 & 1 & r^c \\ 0 & 0 & 0 & 1 & 0 \\ e^* & 1 & 1 & & \\ \bar{r}^* & r^c & 0 & & \end{pmatrix} \begin{pmatrix} x \\ x^c \\ x^l \\ -\lambda \\ -\mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \eta \\ 1 \\ \rho \end{pmatrix}, \qquad x^l \geq 0, \quad \eta \geq 0, \quad x^l \eta = 0.$$

As in Theorem 5, the first two rows yield $\lambda = -r^c \mu$ and $x = \mu \Sigma^{-1}(\bar{r} - r^c e)$. The third row yields $\eta = -\lambda = r^c \mu$. Hence, by complementarity of x^l and η, rows four and five yield either x^c and μ as in Problem 3 (if $x^l = 0$ and $\eta \geq 0$; case 1), or $x^c + x^l = 1$ and $r^c x^c = \rho$ (if $x^l \geq 0$ and $\eta = 0$; case 2). Due to the nonnegativity of x^l and η, case 1 can only hold for $\rho \geq r^c$ and case 2 only for $\rho \leq r^c$. (Indeed, for $\rho = r^c$ both cases coincide so that all variables are continuous with respect to the parameter ρ.) ☐

Problem 6. The tradeoff version of Problem 5 reads

$$\begin{aligned} \min_{x,x^c,x^l} \quad & \frac{1}{2} x^* \Sigma x - \mu(\bar{r}^* x + r^c x^c) \\ \text{s.t.} \quad & e^* x + x^c + x^l = 1, \quad x^l \geq 0. \end{aligned}$$

Theorem 9. *Problem 6 with $\mu > 0$ is equivalent to Problem 5 (with $\rho > r^c$) iff $\rho = r^c + \mu \delta^c$.* Every *solution of Problem 5 with $\rho \leq r^c$ is optimal for Problem 6 with $\mu = 0$. Problem 6 is unbounded for $\mu < 0$: no solution exists.*

Proof. The necessary conditions of both problems are identical, except that in the tradeoff problem μ is given and the reward condition is missing. The condition $\eta = r^c \mu$ together with nonnegativity and complementarity of x^l, η leads immediately to the three given cases. ☐

Discussion. Apparently, at the price of slightly increased complexity, Problem 5 correctly captures the case of an overly pessimistic investor. It minimizes something that qualitatively resembles a quadratic *downside risk* (or *shortfall risk*): the risk to obtain less than the desired amount, see Fig. 3. In that sense the model is now more realistic. (In contrast, its tradeoff version becomes degenerate for $\mu = 0$ and does not extend to $\mu < 0$.) But what does it mean to "invest" knowingly in an asset with guaranteed total loss? Does it not imply that one could as well burn the money?

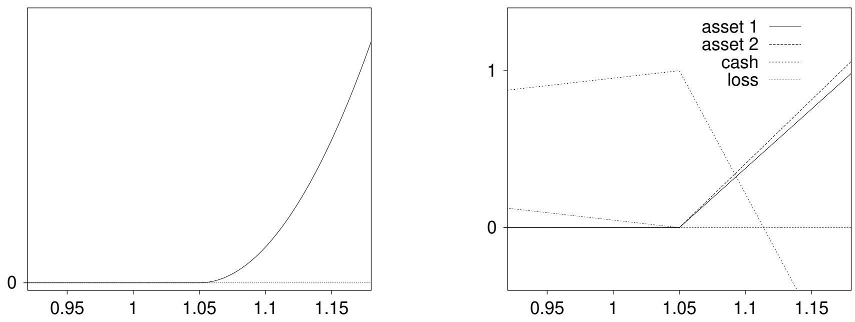

Figure 3: Portfolio with two risky assets, cash, and asset with guaranteed loss; $\bar{r}^1, \bar{r}^2, r^c$ as before. Left: efficient frontier. Right: optimal portfolio vs. reward.

Let us first give the provocative answer *"Yes, why not?"*. From the point of view of the model, the investor's goal is minimizing the "risk" of earning less *or* more than the specified reward. Therefore it makes sense to get rid of money whenever this reduces the variance—which it indeed does for $\rho < \hat{\rho}$. The model cannot know and consequently does not care how the investor will interpret that, and it will use *any* possible means to take out capital if appropriate.

Of course, we can also offer a better interpretation. The fraction invested in x^l is simply surplus capital: the desired reward ρ is achieved at zero risk without that amount, so it should not be invested in the first place—at least not into the portfolio under consideration. The investor may enjoy a free lunch instead or support her favorite artist, if she prefers that to burning the money. Or she may reconsider and decide to pursue a more ambitious goal: the model does not suggest *how* to spend the surplus money. This interpretation of the new riskless (but inefficient) solutions becomes obvious after the following observation.

Lemma 3. *Problem 5 is equivalent to the modification of Problem 3 where the budget equation $e^*x + x^c = 1$ is replaced by the inequality $e^*x + x^c \leq 1$, i.e., less than 100 % investment is allowed.*

Proof. With a slack variable $s \geq 0$, the modified condition is equivalent to $e^*x + x^c + s = 1$, and the modified Problem 3 becomes *identical* to Problem 5: the ominous loss asset is simply a slack variable, $x^l \equiv s$. □

1.4 Utility functions

Let us start a brief excursion into utility-based portfolio optimization by considering Problem 1, the tradeoff formulation of the mean-variance model for n risky assets. In utility theory, the portfolio is chosen so that some function $U(w)$, the investor's subjective *utility* of final wealth $w = r^*x$, has maximal expectation for the given return distribution. The connection is apparent: minimizing the tradeoff function with parameter μ_ρ is equivalent to maximizing the expectation $\mathbf{E}[U_\rho(r^*x)]$ if we define the family of concave quadratic utility functions

$$U_\rho(w) := \mu_\rho w - \frac{1}{2}(w-\rho)^2, \qquad \mu_\rho \equiv \frac{\alpha\rho - \beta}{\delta}.$$

(μ_ρ is the optimal budget multiplier of the desired reward ρ.) If $\rho + \mu_\rho > 0$, then this equivalence remains valid for the normalized utility functions

$$\bar{U}_\rho(w) := \frac{1}{(\rho+\mu_\rho)^2}\left[U_\rho(w) + \frac{1}{2}\rho^2\right] = \frac{w}{\rho+\mu_\rho} - \frac{w^2}{2(\rho+\mu_\rho)^2},$$

satisfying $\bar{U}_\rho(0) = 0$ and $\max_{w\in\mathbb{R}} \bar{U}_\rho(w) = \bar{U}_\rho(\rho+\mu_\rho) = \frac{1}{2}$. For a portfolio with two positively correlated risky assets, Fig. 4 shows the normalized utility functions associated with several desired rewards, and the resulting optimal wealth distribution functions given normally distributed

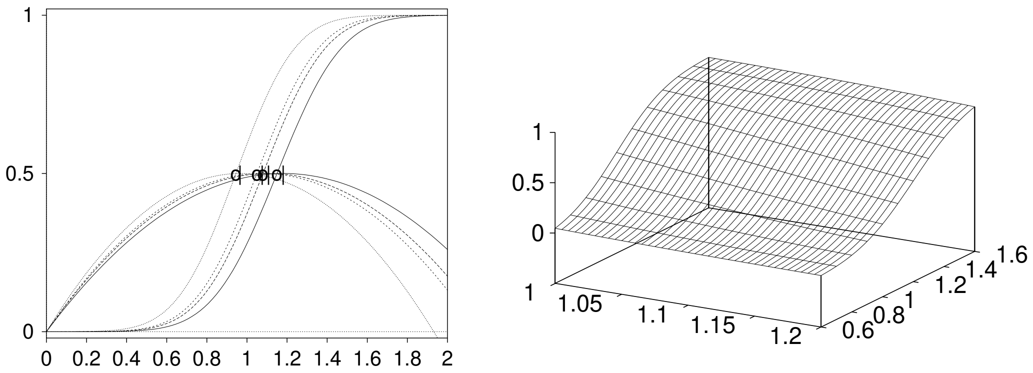

Figure 4: Utility-based portfolio optimization; two positively correlated assets. Left: Normalized utility functions $\bar{U}_\rho$ and optimal wealth distribution functions Φ_ρ for $\rho \in \{0.9r^c, r^c, \bar{r}^1, \bar{r}^2\}$. Symmetry center of Φ_ρ curves at $(\rho, \frac{1}{2})$ marked by '$\circ$'; top of $\bar{U}_\rho$ parabolas at $(\rho + \mu_\rho, \frac{1}{2})$ marked by '$|$'. Right: Family of optimal wealth distribution functions Φ_ρ over the range $\rho \in [1, 1.2]$.

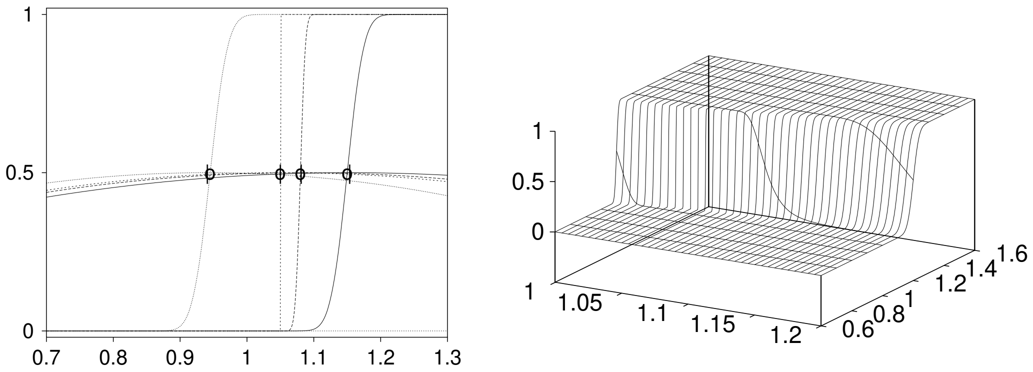

Figure 5: The same situation as in Fig. 4 with cash included.

returns, $r \sim \mathcal{N}(\bar{r}, \Sigma)$. The optimal wealth distributions have the explicit form

$$\Phi_\rho(w) := \frac{1}{\sqrt{2\pi}\,\sigma_\rho} \int_{-\infty}^{w} \exp\left(-\frac{(t-\rho)^2}{2\sigma_\rho^2}\right) dt = \Phi\left(\frac{w-\rho}{\sqrt{2}\,\sigma_\rho}\right)$$

where $\sigma_\rho^2 := \sigma^2(\rho) = (\mu_\rho^2 \delta + 1)/\alpha$, and Φ is the standard error integral,

$$\Phi(w) := \frac{1}{\sqrt{\pi}} \int_{-\infty}^{w} \exp(-t^2)\, dt.$$

When cash is included in the portfolio (Problem 4), utility functions $U_\rho, \bar{U}_\rho$ and distribution functions Φ_ρ have precisely the same form, except that the final wealth becomes $w = r^* x + r^c x^c$, and that $\mu_\rho := (\rho - r^c)/\delta^c$ and $\sigma_\rho^2 := \mu_\rho^2 \delta^c$. The properties of Φ_ρ give another indication of the risk reduction mechanism described in Theorem 7: Fig. 5 shows that the slope of Φ_ρ is rather steep and becomes a jump in the zero-risk case, $\rho = r^c$.

When a loss asset is also added to the portfolio in Problem 6, the utility functions $U_\rho, \bar{U}_\rho$ are exactly identical to the previous case, and even their optimal wealth distributions for $\rho \geq r^c$ coincide, see Fig. 6. For $\rho \leq r^c$, however, the wealth distributions $\Phi_\rho := \chi_{[\rho,\infty)}$ become indicator functions rather than normal distributions: they all have $\mu_\rho = \sigma_\rho = 0$ and a jump discontinuity at ρ, producing zero risk. (In Problem 4 this happens only for $\rho = r^c$.)

In this paper we do not wish to pursue the subject further. The interested reader should refer to the original considerations of Markowitz [54, Part IV], the literature cited in the introduction (especially [31, 37, 39, 40, 46, 48, 59]), and the references therein.

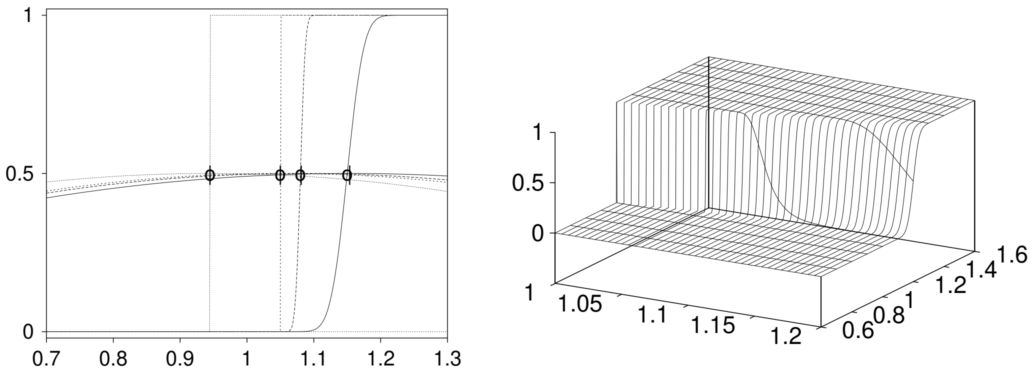

Figure 6: The same situation as in Fig. 4 with cash and loss included.

1.5 Influence of inequalities

Except for the no-arbitrage condition $x^l \geq 0$, all the problems considered so far were purely equality-constrained. Now we study problems with inequalities. Let us first view ρ as a lower bound (not an exact value) for the desired reward.

Theorem 10. *Consider the following modifications of Problems 2, 3, and 5: the reward equation is replaced by the inequality $\bar{r}^*x \geq \rho$ in Problem 2 and by $\bar{r}^*x + r^c x^c \geq \rho$ in Problems 3 and 5. Then:*

1. *The solution of each original problem for $\rho \geq \hat{\rho}$ is also the unique solution of the corresponding modified problem (upper branch).*
2. *The solution of Problem 2 or 3 with reward $\hat{\rho}$ is also the unique solution of the corresponding modified problem for any $\rho \leq \hat{\rho}$ (lower branch).*
3. *Any solution of Problem 5 with $\rho(x, x^c, x^l) \in [\rho, r^c]$ is a riskless solution of the modified problem with $\rho < r^c$. That is, any portfolio $(0, x^c, 1 - x^c)$ with $x^c \in [\rho/r^c, 1]$ is optimal.*

Proof. Obvious. □

Discussion. Specifying the desired reward as a lower bound rather than an exact value leads to reasonable behavior on the lower branch of the efficient frontier. The minimal-risk solution is simply extended to all sufficiently small rewards, yielding again a quadratic downside-like risk in each case. In Problem 3 this provides an alternative to introducing a loss asset. (By statement 3, the combination yields non-unique solutions but no further advantages.) The lower bound formulation—as originally introduced by Markowitz—might appear more natural than the loss asset, but mathematically both are equivalent: optimal investments in risky assets and the resulting risk are identical, only the surplus money is put in x^l (removed immediately) in the latter case and invested in cash x^c in the former case—to be removed afterwards. This difference could be interpreted as reflecting certain attitudes of the investor toward surplus money: $e^*x + x^c \leq 1$ and $\rho(x) \geq \rho$ would model respective preferences for immediate consumption or for future profit, whereas the combination would express indecision. Instead, we simply interpret the loss model as giving the investor a *choice* between consumption and profit in some situations. Leaving the choice open seems preferable in view of the multi-period case.

In practice, nonnegativity constraints $x \geq 0$, $x^c \geq 0$ will usually be included to prohibit short-selling assets or borrowing cash.[4] The budget equation then implies that only *convex combinations* of the assets are permitted. (That is, strictly speaking, convex combinations of single-asset portfolios). Of course, the unconstrained solution remains valid if and only if it is nonnegative anyway.

[4] Even *if* borrowing is allowed, it should also be modeled as a (separate) nonnegative asset since the interest rate differs from the one for investing.

Otherwise some constraints become tight, excluding the corresponding assets from the portfolio and increasing risk. More precisely, the following simple facts hold.

Theorem 11. *Include nonnegativity constraints $x \geq 0$, $x^c \geq 0$ in Problems 2, 3, and 5, and assume for simplicity that $\bar{r} > r^c e > 0$. Denote by $\bar{r}^{\min}, \bar{r}^{\max}$ the minimal and maximal expected return in the portfolio, and choose corresponding assets $x^{\min}, x^{\max}$.*

1. *Problems 3 and 5 have $x^{\min} = x^c$ and $x^{\min} = x^l$, respectively.*
2. *In each problem, an optimal solution exists iff $\rho \in [\bar{r}^{\min}, \bar{r}^{\max}]$.*
3. *The efficient frontier is convex and piecewise quadratic (or linear).*

Proof. Statement 1 is trivial. Since each problem is convex, an optimal solution exists if and only if the feasible set is nonempty. For $\rho \in [\bar{r}^{\min}, \bar{r}^{\max}]$, the feasible set clearly contains a (unique) convex combination of $x^{\min}$ and $x^{\max}$. Conversely, every convex combination of assets yields as reward the same convex combination of individual expected returns, which lies in the range $[\bar{r}^{\min}, \bar{r}^{\max}]$. This proves statement 2. To prove statement 3 consider Problem 2 first. Let $\rho_0, \rho_1 \in [\bar{r}^{\min}, \bar{r}^{\max}]$ with respective solutions x_0, x_1. Then $x_t := (1-t)x_0 + tx_1$ is feasible for $\rho_t := (1-t)\rho_0 + t\rho_1$, $t \in [0,1]$, and convexity of the efficient frontier follows from convexity of R,

$$\sigma^2(\rho_t) \leq R(x_t) \leq (1-t)R(x_0) + tR(x_1) = (1-t)\sigma^2(\rho_0) + t\sigma^2(\rho_1).$$

At $\bar{r}^{\min}$ and $\bar{r}^{\max}$ all the money is invested in one single asset: $x^{\min}$ or $x^{\max}$. Each $\rho \in (\bar{r}^{\min}, \bar{r}^{\max})$ determines a subset of two or more nonnegative assets whose efficient frontier gives the optimal risk in that point. Since strict positivity is a generic property, each of these sub-portfolios is optimal either in a single point or on an entire interval. Thus, the efficient frontier (in Problem 2) is composed of finitely many quadratic pieces. Precisely the same arguments hold for Problem 3 since $R(x, x^c) \equiv R(x)$. In Problem 5, the efficient frontier consists of the segment $\sigma^2(\rho) \equiv 0$ on $[0, r^c]$ and the pieces of Problem 3 on $[r^c, \bar{r}^{\max}]$. ☐

Discussion. The theorem gives a simple characterization of the influence of standard nonnegativity constraints. In a portfolio with three risky assets, the respective efficient frontiers of subportfolios that contribute to the optimal solution in Problems 2, 3, and 5 might look as in Fig. 7. Other inequalities, like upper bounds on the assets or limits on arbitrary asset combinations, will further restrict the range of feasible rewards and increase the risk in a similar manner. This case was already considered by Markowitz: he handles general linear inequalities by *dummy assets* (slacks) and constraints $Ax = b$, $x \geq 0$, where $A = e^*$ (with $x \geq 0$ and $\rho(x) \geq \rho$) is called the *standard case* [54, p. 171]. Moreover, Markowitz devised an algorithm to trace the *critical lines*, that is, the segments of the efficient frontier [53], [54, Chap. VIII]. The number of assets in an optimal portfolio is ivestigated, e.g., by Nakasato and Furukawa [61].

1.6 Downside risk

In the discussion of Section 1.3 we made the remark that Problem 5 resembles a downside risk. This will now be investigated in detail. We have to work with the distribution of returns, but the entire analysis can be given in geometric terms using its *support*. Consider a probability space $(\mathbb{R}^n, \mathcal{B}, P)$ and let $\Xi \in \mathcal{B}$ denote the support of P, i.e., the smallest closed Borel set with measure one,

$$\Xi = \operatorname{supp}(P) := \bigcap_{P(S)=1} \overline{S}.$$

(Of course, if P has a density ϕ, then $\Xi = \operatorname{supp}(\phi) = \overline{\{x \in \mathbb{R}^n : \phi(x) > 0\}}$.) In the following we will actually use the convex hull of the support in most cases, denoted by $C := \operatorname{conv}(\Xi)$.

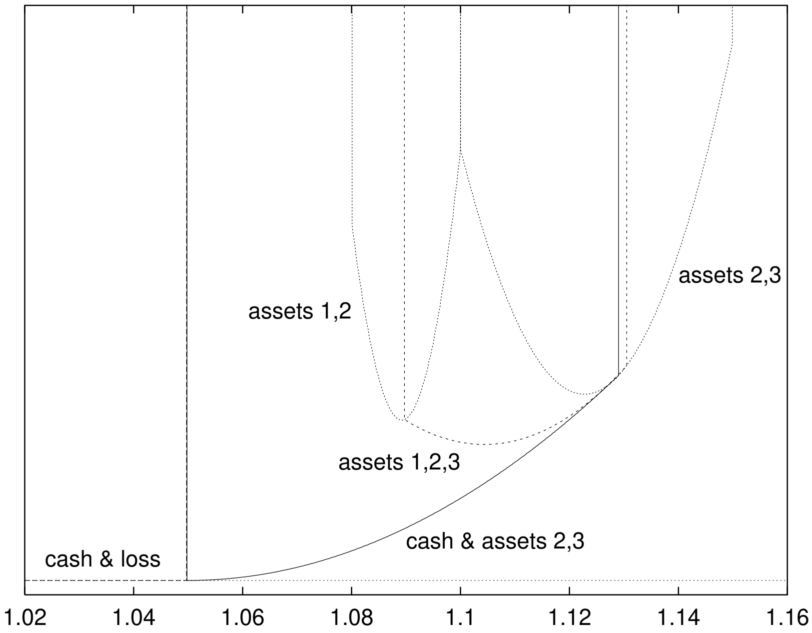

Figure 7: Efficient frontiers for Problems 2, 3, and 5 with nonnegative assets.

Definition 3 (Downside risk). For a function w of the random vector r with distribution P, the downside risk of order $q > 0$ with target $\tau \in \mathbb{R}$ is

$$R^q_\tau(w) := \mathbf{E}\big[|\min(w(r) - \tau, 0)|^q\big] = \int_{\mathbb{R}^n} |\min(w(r) - \tau, 0)|^q \, dP.$$

Remarks. Without the risk context such expectations are neutrally called *lower partial moments*, with *downside expected value* or *semi-deviation* (order 1) and *downside variance* or *semi-variance* (order 2) as special cases. In [54, Chap. XIII] Markowitz gives a qualitative discussion of the linear case (*expected value of loss*, $q = 1$), the quadratic case (*semi-variance*, $q = 2$), and some other measures of risk, by examining the associated utility functions. Expectation of loss has recently gained interest as a *coherent* replacement for the popular *Value-at-Risk (VaR),* often under alternative names like *mean shortfall, tail VaR,* or *conditional VaR*, cf. [3, 79].

In the following we are only interested in quadratic downside risk of portfolio returns like $w_{x,x^c}(r, r^c) = r^*x + r^c x^c$. Moreover, we always use the desired reward as a natural choice for the shortfall target, $\tau = \rho$, and write simply $R_\rho(x, x^c)$ instead of $R^2_\rho(w_{x,x^c})$. The problems considered in this section are downside risk versions of Problems 3 and 5, and of the modification of Problem 3 with $\rho(x, x^c) \geq \rho$. In each case only the objective is changed: standard risk R is replaced by downside risk R_ρ. Before considering these problems we need some technical preparations.

For $x \neq 0$ and $c \in \mathbb{R}$ let us introduce open and closed half-spaces

$$H(x, c) := \{r \in \mathbb{R}^n : r^*x < c\}, \qquad \bar{H}(x, c) := \{r \in \mathbb{R}^n : r^*x \leq c\},$$

and portfolio-dependent semi-variance matrices

$$\Sigma(x) := \int_{\bar{r}+H(x,0)} (r - \bar{r})(r - \bar{r})^* \, dP, \quad x \neq 0.$$

For $x = 0$ let $\Sigma(0) := \frac{1}{2}\Sigma$ where Σ is of course the usual covariance matrix,

$$\Sigma := \int_{\mathbb{R}^n} (r - \bar{r})(r - \bar{r})^* \, dP.$$

Lemma 4. *Denote by* $\uplus$ *a disjoint union. Then, for* $x \neq 0$ *and* $a > 0$,

1. $H(ax, c) = H(x, a^{-1}c)$, $H(x, ac) = H(a^{-1}x, c)$, $H(ax, ac) = H(x, c)$.
2. $H(-x, -c) = \mathbb{R}^n \setminus \bar{H}(x, c)$.
3. $\bar{H}(x, c) = H(x, c) \uplus \partial H(x, c)$, $\bar{H}(x, 0) = H(x, 0) \uplus \{x\}^\perp$.

Statements 1 and 2 remain valid when H *and* $\bar{H}$ *are exchanged everywhere.*

Proof. Trivial. □

Lemma 5. *For* $x \in \mathbb{R}^n$ *and* $a > 0$,

1. $\Sigma(ax) = \Sigma(x)$.
2. $0 \leq \Sigma(x) \leq \Sigma$. *(In particular, each* $\Sigma(x)$ *is positive semidefinite.)*
3. $x^*\Sigma(x)x = \mathbf{E}[\min((r - \bar{r})^*x, 0)^2]$.
4. $x^*[\Sigma(x) + \Sigma(-x)]x = x^*\Sigma x$.

Proof. Statements 1, 2, 3 are obvious from the definitions and the first identity in Lemma 4. The expressions in statement 4 are identical for $x = 0$; otherwise they differ by the integral of $((r - \bar{r})^*x)^2$ over $\bar{r} + \{x\}^\perp$, which is clearly zero. □

Lemma 6. *For any random vector* r *the following holds.*

1. *The expectation lies in the convex hull of the support:* $\bar{r} \in C$.
2. *The covariance matrix and all semi-variance matrices are positive definite if and only if* Ξ *has full dimension in the sense that its convex hull has nonempty interior:*
   $\operatorname{int}(C) \neq \emptyset \iff \Sigma > 0 \iff \Sigma(x) > 0 \ \forall x \in \mathbb{R}^n$.
3. *If* r *is discrete with* $\Sigma > 0$, *then it has at least* $n + 1$ *realizations.*

Proof. Assume $\bar{r} \notin C$. Then $\bar{r}$ has positive distance to C, and a vector $x \neq 0$ exists so that $(r - \bar{r})^*x > 0$ for all $r \in C$. Since expectation is the integral over $\Xi \subseteq C$, this yields the contradiction $0 < \mathbf{E}[(r - \bar{r})^*x] = 0$, proving statement 1. Now assume $\operatorname{int}(C) = \emptyset$. Then C is contained in some hyperplane $\bar{r} + \{x\}^\perp$ with $x \neq 0$, implying

$$x^*\Sigma x = \mathbf{E}[((r - \bar{r})^*x)^2] = 0.$$

Hence Σ is only positive semidefinite. Conversely, assume $\operatorname{int}(C) \neq \emptyset$ and $x \neq 0$. Then $(r - \bar{r})^*x < 0$ for all $r \in \bar{r} + H(x, 0)$. By Lemma 7 below, $\bar{r} + H(x, 0)$ has positive measure. Therefore

$$x^*\Sigma(x)x = \int_{\bar{r}+H(x,0)} ((r - \bar{r})^*x)^2 \, dP > 0,$$

showing that $\Sigma(x) > 0$. The proof of statement 2 is complete since $\Sigma \geq \Sigma(x)$ for all x. Now statement 3 is an immediate consequence. □

Lemma 7. *Let* $\operatorname{int}(C) \neq \emptyset$ *and* $x \neq 0$. *Then* $\bar{r} + H(x, 0)$ *has positive measure.*

Proof. The inner product $s(x) := (r - \bar{r})^*x$ is negative, zero, and positive on the respective sets $\bar{r} + H(x, 0)$, $\bar{r} + \{x\}^\perp$, and $\bar{r} + H(-x, 0)$. Furthermore,

$$\int_{\bar{r}+H(x,0)} s(x) \, dP + \int_{\bar{r}+\{x\}^\perp} s(x) \, dP + \int_{\bar{r}+H(-x,0)} s(x) \, dP = \mathbf{E}[s(x)] = 0.$$

Therefore $\bar{r} + H(x, 0)$ and $\bar{r} + H(-x, 0)$ have either both positive measure or both measure zero. The second case implies $\Xi \subseteq \bar{r} + \{x\}^\perp$, which leads to the contradiction $\operatorname{int}(C) = \emptyset$. □

Let us now study the downside risk versions of Problems 3 and 5 under the same assumptions as before (A1 and A3 resp. A1, A3, and A4). It will be seen that in these cases the qualitative behavior does not change significantly. This is mainly because the constraints are linear and $\Sigma(x)$ depends only on the direction and not on the magnitude of x (cf. Lemma 5).

Problem 7. We minimize downside risk $R_\rho(x, x^c)$ for risky assets and cash, with fixed desired reward $\rho(x, x^c) = \rho$,

$$\begin{aligned} \min_{x,x^c} \quad & \frac{1}{2}\int_{\mathbb{R}^n} \min(r^* x + r^c x^c - \rho, 0)^2 \, dP \\ \text{s.t.} \quad & e^* x + x^c = 1, \\ & \bar{r}^* x + r^c x^c = \rho. \end{aligned}$$

Problem 8. Now minimize downside risk $R_\rho(x, x^c, x^l)$ for risky assets, cash, and loss, with fixed desired reward $\rho(x, x^c, x^l) = \rho$,

$$\begin{aligned} \min_{x,x^c,x^l} \quad & \frac{1}{2}\int_{\mathbb{R}^n} \min(r^* x + r^c x^c - \rho, 0)^2 \, dP \\ \text{s.t.} \quad & e^* x + x^c + x^l = 1, \quad x^l \geq 0, \\ & \bar{r}^* x + r^c x^c = \rho. \end{aligned}$$

Lemma 8. *With $x^c \equiv 1 - e^* x - x^l$ and $\theta \equiv r^c x^l$, Problem 8 is equivalent to*

$$\begin{aligned} \min_{x,\theta} \quad & \frac{1}{2} x^* \Sigma(x) x \\ \text{s.t.} \quad & (\bar{r} - r^c e)^* x = \rho + \theta - r^c, \quad \theta \geq 0. \end{aligned}$$

When fixing $\theta = 0$, the resulting problem is equivalent to Problem 7.

Proof. The modified reward condition is immediately obtained by the identity $x^c = 1 - e^* x - x^l$. Using $r^* x + r^c x^c - \rho = (r - \bar{r})^* x$ gives the downside risk

$$\int_{\mathbb{R}^n} \min((r - \bar{r})^* x, 0)^2 \, dP = \int_{\bar{r}+H(x,0)} ((r - \bar{r})^* x)^2 \, dP = x^* \Sigma(x) x.$$

(The special case $x = 0$ is easily verified.) Clearly, $\theta = 0$ means $x^l = 0$, yielding Problem 7. □

We are now ready to analyze Problems 7 and 8. In general, closed-form solutions cannot be found due to the nonlinearity of downside risk with respect to the risky assets. However, we can derive some important properties of the solutions and give a qualitative comparison to Problems 3 and 5.

Lemma 9. *Optimal solutions always exist in Problems 7 and 8. The resulting downside risk is nonnegative and not greater than the optimal risk in Problem 3 or 5, respectively. Moreover, the riskless solutions of Problems 7 and 3 (8 and 5) are identical. (In general the solutions are* not *unique.)*

Proof. Convexity of $\min(w, 0)$ implies convexity of downside risk $x^* \Sigma(x) x$ and thus of Problems 7 and 8. The existence of solutions and the stated inclusion follow since $0 \leq \Sigma(x) \leq \Sigma$ by Lemma 4. By assumption A1 and Lemma 6, zero risk requires $x = 0$, which holds under the same conditions as in the standard risk case. □

Theorem 12. *In Problem 7, choose respective optimal portfolios $(x_\pm, x^c_\pm)$ for $\rho_\pm := r^c \pm 1$. Then $(a x_\pm, a x^c_\pm - a + 1)$ is optimal for $\rho = r^c \pm a$ if $a \geq 0$. Moreover, $x_\pm \neq 0$ and $x_+ \neq x_-$.*

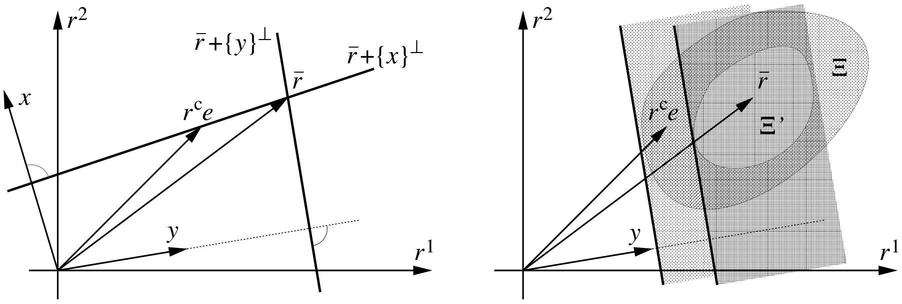

Figure 8: Hyperplanes (left) and half-spaces (right) in downside risk ($n = 2$).

Proof. The transformation of x_+^c follows from $e^*x + x^c = 1$ if ax_+ is optimal. Assume ax_+ is *not* optimal for $\rho = r^c + a > r^c$. By Lemma 8, $x \neq ax_+$ exists so that $(\bar r - r^c e)^* x = \rho - r^c$ and

$$x^*\Sigma(x)x < ax_+^*\Sigma(ax_+)ax_+ = a^2 x_+^*\Sigma(x_+)x_+,$$

where the last equality holds by Lemma 5. Hence,

$$(\bar r - r^c e)^* a^{-1} x = a^{-1}(\rho - r^c) = 1 = \rho_+ - r^c$$

and

$$a^{-1}x^*\Sigma(a^{-1}x)a^{-1}x = a^{-2}x^*\Sigma(x)x < x_+^*\Sigma(x_+)x_+.$$

Thus x_+ cannot be optimal for ρ_+: a contradiction. The case $\rho < r^c$ is analogous, and $\rho = r^c$ is trivial. Finally observe that $(\bar r - r^c e)^* x_- < 0 < (\bar r - r^c e)^* x_+$. $\square$

Theorem 13. *Constants $c_\pm \in (0, 1)$ exist so that the optimal risk in Problem 7 is c_+ (c_-) times the optimal risk of Problem 3 on the upper (lower) branch.*

Proof. The existence of $c_\pm \in (0, 1]$ with the stated properties follows from Lemma 9 and Theorem 12. Statement 4 of Lemma 5 implies $c_\pm < 1$. $\square$

Theorem 14. *The same statements as in Theorems 12 and 13 hold on the upper branch in Problem 8. On the lower branch one has the (unique) riskless solution $(x, x^c, x^l) = (0, \rho/r^c, 1 - \rho/r^c)$.*

Proof. This is a simple case distinction. $\square$

Remarks. When assumption A1 ($\Sigma > 0$) is dropped, the following can be shown using Lemmas 4–8. A feasible portfolio with $x \neq 0$ has zero risk in Problem 7 or 8 iff Ξ is contained in the hyperplane $\bar r + \{x\}^\perp$. In Problem 7 such a portfolio exists for $\rho = r^c$ iff Ξ lies in a hyperplane containing both $\bar r$ and $r^c e$ (see x in Fig. 8), and for $\rho \neq r^c$ iff Ξ lies in a hyperplane containing $\bar r$ but not $r^c e$ (see y in Fig. 8). Similarly, in Problem 8 such a portfolio exists for $\rho \le r^c$ iff Ξ lies in *any* hyperplane containing $\bar r$ (see x, y in Fig. 8), and for $\rho > r^c$ iff Ξ lies in a hyperplane containing $\bar r$ but not $r^c e$ (see y in Fig. 8). In both cases, an arbitrage is thus possible if Ξ lies in a hyperplane containing $\bar r$ but not $r^c e$.

Discussion. Up to now downside risk behaves qualitatively similar to standard risk: The efficient frontier is still piecewise quadratic, and optimal portfolios are always combinations of reward-independent portfolios. Only uniqueness is not guaranteed any more, and the curvatures of upper and lower branches of the efficient frontier may differ. (Optimal portfolios will usually also differ from their standard risk counterparts, of course). The similarity is caused by fixing the reward: this places the mean $\bar r$ on the boundary of semi-variance half-spaces so that the properties of $\Sigma(x)$ come into play. The two risk measures become identical if the return distribution is symmetric with respect to rotations about $\bar r$. In that case $\Sigma(x) = \frac{1}{2}\Sigma$ for all x, and $c_+ = c_- = \frac{1}{2}$.

The last problem considered in this section is the downside risk version of the modification of Problem 3. Again the riskless solutions are of interest.

Problem 9. We minimize downside risk $R_\rho(x, x^c)$ for risky assets and cash, with desired minimal reward $\rho(x, x^c) \geq \rho$,

$$\begin{aligned} \min_{x,x^c,\theta} \quad & \frac{1}{2}\int_{\mathbb{R}^n} \min(r^*x + r^c x^c - \rho, 0)^2 \, dP \\ \text{s.t.} \quad & e^*x + x^c = 1, \\ & \bar{r}^*x + r^c x^c = \rho + \theta, \quad \theta \geq 0. \end{aligned}$$

Remark. Note that downside risk is still calculated with respect to the *desired* reward ρ whereas the *actual* reward is now $\rho + \theta$. Otherwise the problem would be equivalent to Problem 8.

Lemma 10. *$(x, x^c) = (0, 1)$ is feasible for Problem 9 iff $\rho \leq r^c$. Otherwise, with $x^c \equiv 1 - e^*x$, Problem 9 is equivalent to*

$$\begin{aligned} \min_{x,\theta} \quad & \frac{1}{2}\int_{\bar{r}+H(x,-\theta)} (\theta + (r - \bar{r})^*x)^2 \, dP \\ \text{s.t.} \quad & (\bar{r} - r^c e)^*x = \rho + \theta - r^c, \quad \theta \geq 0. \end{aligned}$$

Proof. The first part is trivial. The second part is proved as Lemma 8, the only difference being that the reward condition now yields $r^*x + r^c x^c - \rho = (r - \bar{r})^*x + \theta$ instead of $(r - \bar{r})^*x$. □

Theorem 15. *If $r^c e \in \operatorname{int}(C)$, then the following holds in Problem 9.*

1. *For every $\rho \leq r^c$, $(x, x^c) = (0, 1)$ is a riskless solution.*
2. *A portfolio with $x \neq 0$ has zero risk iff $\Xi \subseteq r^c e + \bar{H}(-x, r^c - \rho)$.*
3. *For $\rho < r^c$ such a portfolio exists iff Ξ is contained in any closed half-space.*
4. *For $\rho \geq r^c$ such a portfolio does not exist.*

Proof. Statement 1 is trivial. If $x \neq 0$, then downside risk clearly vanishes iff Ξ does not intersect $\bar{r} + H(x, -\theta)$, that is, iff $(r - \bar{r})^*x + \theta \geq 0$ for $r \in \Xi$. Substituting $\theta = (\bar{r} - r^c e)^*x - \rho + r^c$ from the reward equation yields the equivalent condition $(r - r^c e)^*x \geq \rho - r^c$ for $r \in \Xi$, which proves statement 2. (This condition implies feasibility of x since $\bar{r} \in C$.) Now, if Ξ is contained in some closed half-space, then $y \neq 0$ exists so that $(r - r^c e)^*y \geq -1$ for $r \in \Xi$, see Fig. 8. For $\rho < r^c$ let $x := (r^c - \rho)y$ to satisfy the zero-risk condition. The "only if" direction of statement 3 is trivial. Finally observe that C contains an open ball centered at $r^c e \in \operatorname{int}(C)$. On that ball the inner product $(r - r^c e)^*x$ takes positive and negative values for any $x \neq 0$, showing that for $\rho \geq r^c$ the zero-risk condition $(r - r^c e)^*x \geq \rho - r^c \geq 0$ cannot be satisfied. □

Remarks. Similar arguments show that for $\rho > r^c$ ($\rho = r^c$) a zero-risk portfolio with $x \neq 0$ exists iff Ξ lies in a closed half-space not containing $r^c e$ (not containing $r^c e$ in its interior), see Ξ' in Fig. 8. This is why we need an additional no-arbitrage condition. Although $r^c e \in C$ would suffice, we choose the stronger condition $r^c e \in \operatorname{int}(C)$ to ensure a unique riskless solution (100 % cash) for $\rho = r^c$. Thus $x \neq 0$ will produce positive risk for any $\rho \geq r^c$.

Discussion. Risk vanishes on the lower branch in Problem 9, but for the upper branch we know only that it is convex; even the optimal portfolios for different $\rho > r^c$ may be unrelated. This is because the actual reward may exceed the shortfall target, resulting in semi-variance half-spaces far from $\bar{r}$ and producing asymmetric (or decentral) risk integrals instead of $\Sigma(x)$. One might expect such truly nonlinear behavior from any downside risk measure, but it occurs only if the actual reward may differ from the shortfall target. One might also associate zero risk on the lower branch with downside risk, but this property occurs for standard risk as well and has nothing to do with the objective; it is caused by the inequality $\rho(x) \geq \rho$ or $e^*x + x^c \leq 1$.

1.7 Summary

We have discussed various formulations of the classical mean-variance approach to obtain single-period models that give a qualitatively correct description of risk, particularly for unreasonably small desired rewards. Positivity constraints and other inequalities have been studied, and downside risk models have been analyzed in detail. Thus we have clarified the effects and interaction of all components in the portfolio optimization problems. In the following we use the results of this section in developing multi-period models. The goal is to achieve an approximate minimization of downside risk.

Tradeoff formulations or utility functions will not be considered any more since extra constraints provide higher modeling flexibility and facilitate a better understanding of subtle details.

2 Multi-period mean-variance analysis

For the multi-period mean-variance models we consider a planning horizon of $T+1$ periods (not necessarily equidistant) in discrete time $t = 0, \dots, T+1$. The portfolio is allocated at $t = 0$ and thereafter restructured at $t = 1, \dots, T$, before the investor obtains her reward after the final period, at time $T+1$. The portfolios and return vectors are $x_t, r_{t+1} \in \mathbb{R}^n$, $t = 0, \dots, T$, yielding asset capitals $r_t^\nu x_{t-1}^\nu$ just before the decision at time t. Cash, its return, and loss assets (if present) are denoted by x_t^c, r_t^c, and x_t^l; the wealth is $w_t = r_t x_{t-1}$ (or $r_t x_{t-1} + r_t^c x_{t-1}^c$). Cash returns r_t^c are assumed to be known a priori, whereas the evolvement of asset returns is of course random. The decision at time t is made after observing the realizations of $r_1, \dots, r_t$ but prior to observations of $r_{t+1}, \dots, r_{T+1}$, leading to a *nonanticipative* policy $x = (x_0, \dots, x_T)$.

Suppose that the distribution of returns until T is given by a *scenario tree*: each r_t has finitely many realizations r_j with probabilities $p_j > 0$, $j \in L_t$, so that L_t forms a level set in the tree. The set of all nodes is $V := \bigcup_{t=0}^T L_t$, and the set of leaves, each representing a scenario, is $L := L_T$. We denote by $0 \in L_0$ the root, by $j \in L_t$ the current node (a partial scenario), by $i \equiv \pi(j) \in L_{t-1}$ its parent node, and by $S(j) \subseteq L_{t+1}$ the set of child nodes (successors). The return in the final period may be given by continuous distributions in each leaf. Thus r_t, x_t are random vectors on a discrete-continuous probability space that possesses a filtration generated by the tree. The conditional expectation $\bar{r}_T := \mathbf{E}(r_{T+1} | L_T)$ and its covariance matrix

$$\Sigma_T := \mathbf{E}[(r_{T+1} - \bar{r}_T)(r_{T+1} - \bar{r}_T)^* | L_T] = \mathbf{E}(r_{T+1} r_{T+1}^* | L_T) - \bar{r}_T \bar{r}_T^*$$

define random variables on the same space, with realizations $\bar{r}_j, \Sigma_j$ on L_T.

The discrete decision vector is denoted $x = (x_j)_{j \in V}$. As before, reward and risk are defined as mean and variance of the final wealth, w_{T+1}. In the absence of cash these definitions read

$$\rho(x) = \mathbf{E}(r_{T+1}^* x_T) = \mathbf{E}(\bar{r}_T^* x_T) = \sum_{j \in L} p_j \bar{r}_j^* x_j$$

and

$$R(x) = \sigma^2(r_{T+1}^* x_T) = \mathbf{E}[(r_{T+1}^* x_T - \rho(x))^2].$$

Lemma 11 (Siede [23]). *The risk is given by*

$$R(x) = \mathbf{E}[x_T^*(\Sigma_T + \bar{r}_T \bar{r}_T^*) x_T] - \rho(x)^2 = \sum_{j \in L} p_j x_j^* (\Sigma_j + \bar{r}_j \bar{r}_j^*) x_j - \rho(x)^2.$$

Proof. By definition,

$$\begin{aligned} R(x) &= \mathbf{E}[(r_{T+1}^* x_T - \rho(x))^2] = \mathbf{E}(x_T^* r_{T+1} r_{T+1}^* x_T) - \rho(x)^2 \\ &= \mathbf{E}[\mathbf{E}(x_T^* r_{T+1} r_{T+1}^* x_T | L_T)] - \rho(x)^2 \\ &= \mathbf{E}[x_T^* \mathbf{E}(r_{T+1} r_{T+1}^* | L_T) x_T] - \rho(x)^2 = \mathbf{E}[x_T^* (\Sigma_T + \bar{r}_T \bar{r}_T^*) x_T] - \rho(x)^2. \end{aligned}$$

The discrete representation follows immediately. □

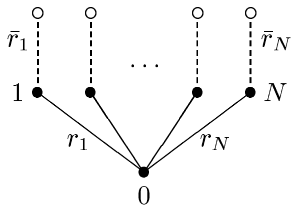

Figure 9: Scenario tree for the two-period mean-variance model.

Remarks. Notice that this representation yields a block-diagonal risk matrix because of the separate term $\rho(x)^2$. If the Hessian of the latter were included in the risk matrix, it would add a completely dense rank-1 term: the dyadic product containing all the covariances $-p_j p_k \bar r_j \bar r_k^*$ between different leaves $j, k \in L$. Since $\rho(x) = \rho$ is fixed in the optimization problems below, we can neglect the term ρ^2 except when considering the reward-dependence of risk.

Corollary 1. *Denote by $\rho_T(x_T) := \bar r_T^* x_T$ and $R_T(x_T) := x_T^* \Sigma_T x_T$ the conditional reward and risk of the final period, respectively, with realizations $\rho_j(x_j) = \bar r_j^* x_j$ and $R_j(x_j) = x_j^* \Sigma_j x_j$ on L_T. Then $R(x) = R_c(x) + R_d(x)$ with*

$$R_c(x) := \mathbf{E}[R_T(x_T)] = \sum_{j \in L} p_j x_j^* \Sigma_j x_j$$

and

$$R_d(x) := \mathbf{E}[\rho_T(x_T)^2] - \rho(x)^2 = \sum_{j \in L} p_j \rho_j(x_j)^2 - \rho(x)^2.$$

Proof. Obvious from Lemma 11. □

Remark. We call R_c the *continuous* part and R_d the *discrete* part of the risk. This distinction will be useful in the subsequent analysis: R_c is the expectation of the conditional variance (of w_{T+1}), measuring the average final-period risk, whereas R_d is the variance of the conditional expectation, measuring how well the individual scenario returns are balanced.

Apparently, period $T+1$ with its continuous distribution (of which only the conditional mean and variance enter the problem) corresponds to the single period in the classical model. Indeed, for $T = 0$ the multi-period model reduces precisely to the one-period case, where the "scenario tree" consists of the root only, and $x \equiv x_T$, $r \equiv r_{T+1}$, $\bar r \equiv \bar r_T$, $\Sigma \equiv \Sigma_T$, $w \equiv w_{T+1}$. Assuming idealized transactions with no loss of capital in the whole multi-period situation, the single budget equation $e^* x = 1$ is supplemented by the set $\{e^* x_t = r_t^* x_{t-1}\}_{t=1}^T$ with discrete representation $\{e^* x_j = r_j^* x_{\pi(j)}\}_{j \in V^*}$ where $V^* := V \setminus \{0\}$.

The following analysis requires some lengthy technical proofs; these are moved into the appendix. Moreover, for technical simplicity, increasing amounts of the presentation will be specialized for the two-period problem, $T = 1$. In that case it is convenient to number the leaves as $j = 1, \dots, N$, so that $V^* = S(0) = L = \{1, \dots, N\}$, see Fig. 9. We would like to point out, however, that the two-period models exhibit almost all the properties of the general multi-period case. Comments will be given at the end of each section.

2.1 Risky assets only

Let us begin the investigation of multi-period problems with the case of purely risky assets. The model discussed here was originally proposed by Frauendorfer [21] (with slightly different objective) and later refined by Frauendorfer and Siede [23].

In this section we impose regularity conditions similar to A1 and A2 in all nodes and in at least one node, respectively. The conditions for L_{t-1} require certain definitions on L_t which in turn depend on the conditions for L_t. To avoid a nested presentation, the conditions will be stated after the definitions.

Let $\tilde{r}_j := p_j\bar{r}_j$ and $\tilde{\Sigma}_j := p_j(\Sigma_j + \bar{r}_j\bar{r}_j^*)$ in the leaves, $j \in L$. By assumption A5 below, $\tilde{\Sigma}_j > 0$; therefore we can define

$$\tilde{\alpha}_j := e^*\tilde{\Sigma}_j^{-1}e, \quad \tilde{\beta}_j := e^*\tilde{\Sigma}_j^{-1}\tilde{r}_j, \quad \tilde{\gamma}_j := \tilde{r}_j^*\tilde{\Sigma}_j^{-1}\tilde{r}_j, \qquad \tilde{\delta}_j := \tilde{\alpha}_j\tilde{\gamma}_j - \tilde{\beta}_j^2.$$

Recursively for $t = T, \dots, 1$ and $i \in L_{t-1}$ let

$$\tilde{r}_i := \sum_{j\in S(i)} \frac{\tilde{\beta}_j}{\tilde{\alpha}_j} r_j, \qquad \tilde{\Sigma}_i := \sum_{j\in S(i)} \frac{1}{\tilde{\alpha}_j} r_j r_j^*,$$

and employ A5 again to define $\tilde{\alpha}_i, \dots, \tilde{\delta}_i$ in analogy to $\tilde{\alpha}_j, \dots, \tilde{\delta}_j$. In the subsequent analysis (and in the solution algorithm) these quantities will play a similar role as their counterparts in the leaves, but they do not have the same meaning. In particular, $\bar{r}_i := \tilde{r}_i/p_i$ and $\Sigma_i := \tilde{\Sigma}_i/p_i - \bar{r}_i\bar{r}_i^*$ are usually not the expectation and covariance matrix of the discrete distribution $\{r_j\}_{j\in S(i)}$.

Basic assumptions.

A5) $\forall j \in V$: $\tilde{\Sigma}_j > 0$.

A6) $\exists j \in V$: $\tilde{r}_j$ is not a multiple of e.

Remarks. The role of these conditions is analogous to the single-period case: they ensure strict convexity and avoid degenerate constraints. Assumption A5 also implies $N \geq n$ as a technical requirement on the return discretization in each period. In practice one will usually have $N > n$, otherwise the covariance matrices are only positive *semi*definite by Lemma 6. Suitable multi-period discretizations can be generated, e.g., by barycentric approximations [22].

Lemma 12. *Under assumptions A5 and A6, the constants $\tilde{\alpha}_j, \tilde{\gamma}_j$ are all positive, the $\tilde{\delta}_j$ are all nonnegative, and at least one $\tilde{\delta}_j$ is positive.*

Proof. Positivity and nonnegativity are proved as in Lemma 1, where $\tilde{\delta}_j = 0$ iff $\tilde{r}_j, e$ are linearly dependent. □

Problem 10. The multi-period mean-variance problem (using $i \equiv \pi(j)$) reads

$$\begin{aligned} \min_x \quad & \sum_{j\in L} \frac{1}{2} x_j^* \tilde{\Sigma}_j x_j - \frac{1}{2}\rho^2 \\ \text{s.t.} \quad & e^* x_0 = 1, \\ & e^* x_j = r_j^* x_i \quad \forall j \in V^*, \\ & \sum_{j\in L} \tilde{r}_j^* x_j = \rho, \end{aligned}$$

Its Lagrangian is

$$L(x, \lambda, \mu; \rho) = \sum_{j\in L} \frac{1}{2} x_j^* \tilde{\Sigma}_j x_j - \frac{1}{2}\rho^2 - \lambda_0(e^* x_0 - 1) - \sum_{j\in V^*} \lambda_j (e^* x_j - r_j^* x_i) - \mu\Bigg(\sum_{j\in L} \tilde{r}_j^* x_j - \rho\Bigg).$$

Theorem 16. *Problem 10 has the unique primal-dual solution*

$$x_j = \tilde{\Sigma}_j^{-1}(\lambda_j e + \mu\tilde{r}_j), \qquad \lambda_j = \frac{w_j - \mu\tilde{\beta}_j}{\tilde{\alpha}_j}, \qquad \mu = \left(\rho - \frac{\tilde{\beta}_0}{\tilde{\alpha}_0}\right) \Big/ \sum_{j\in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j},$$

where $w_0 = 1$ and $w_j = r_j^* x_i$ for $j \in V^*$. The associated optimal risk is

$$R(x) \equiv \sigma^2(\rho) = \frac{1}{\tilde{\alpha}_0} + \left(\rho - \frac{\tilde{\beta}_0}{\tilde{\alpha}_0}\right)^2 \Big/ \sum_{j\in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} - \rho^2.$$

Its global minimizer $\hat\rho$ and minimal risk, respectively, are

$$\hat\rho = \frac{\tilde\beta_0}{\tilde\alpha_0} \Bigg/ \Bigg(1 - \sum_{j\in V} \frac{\tilde\delta_j}{\tilde\alpha_j}\Bigg), \qquad \sigma^2(\hat\rho) = \frac{1}{\tilde\alpha_0} - \Bigg(\frac{\tilde\beta_0}{\tilde\alpha_0}\Bigg)^2 \Bigg/ \Bigg(1 - \sum_{j\in V} \frac{\tilde\delta_j}{\tilde\alpha_j}\Bigg).$$

Proof. See appendix (for $T = 1$). □

Discussion. The proof of Theorem 16 is given for the two-period case only, where expressions for the solution variables are derived first in the leaves and then in the root. In the multi-period case this generalizes readily to a recursive procedure which is actually a highly efficient algorithm for practical computations. We call that recursion the *tree-sparse Schur complement method* [76, 77].

The specialization to a single period gives similar results as Theorems 1, 2, and 4, but under slightly weaker conditions: only $\Sigma + \bar r \bar r^* > 0$ rather than $\Sigma > 0$ is now required, so that riskless portfolios may exist. (We use the weaker condition anyway since $\tilde\Sigma = \Sigma + \bar r \bar r^*$ appears naturally in the problem.) The two following lemmas establish the precise relationship.

Lemma 13. *If $\tilde\Sigma > 0$ but not $\Sigma > 0$, then the nullspace is $N(\Sigma) = \operatorname{span}(\tilde\Sigma^{-1}\bar r)$.*

Proof. If $\Sigma x = 0$, then $\tilde\Sigma x = \bar r \bar r^* x$ and $x = (\bar r^* x)\tilde\Sigma^{-1}\bar r \in \operatorname{span}(\tilde\Sigma^{-1}\bar r)$. □

Lemma 14. *Consider the single-period case of Problem 10 under assumptions A1 and A2, i.e., $\Sigma \equiv \Sigma_0 > 0$. Then*

$$\tilde\alpha_0 = \frac{\alpha+\delta}{1+\gamma}, \qquad \tilde\beta_0 = \frac{\beta}{1+\gamma}, \qquad \tilde\gamma_0 = \frac{\gamma}{1+\gamma}, \qquad \tilde\delta_0 = \frac{\delta}{1+\gamma},$$

and

$$\tilde x = x, \qquad \tilde\lambda = \lambda, \qquad \tilde\mu = \mu + \rho,$$

where quantities with and without tilde refer to problems 10 and 2, respectively.

Proof. The Sherman–Morrison formula applied to $\tilde\Sigma = \Sigma + \bar r \bar r^*$ yields

$$\tilde\Sigma^{-1} = \Sigma^{-1} - \frac{\Sigma^{-1}\bar r \bar r^* \Sigma^{-1}}{1 + \bar r^* \Sigma^{-1} \bar r} = \Sigma^{-1} - \frac{\Sigma^{-1}\bar r \bar r^* \Sigma^{-1}}{1+\gamma}.$$

Using this in the definitions of $\tilde\alpha_0, \tilde\beta_0, \tilde\gamma_0, \tilde\delta_0$ gives the first set of identities after a few elementary calculations. The relation of the solutions is similarly obtained by simple but more lengthy calculations (first $\tilde\mu$, then $\tilde\lambda$, then $\tilde x$). □

2.2 Risky assets and cash

Here we define local constants for the two-period case only but state the general optimization problem. Denote by r_1^c, r_2^c the deterministic (and thus scenario-independent) cash returns in periods one and two, respectively, and by $r^c := r_2^c r_1^c$ their combined return. As before, let $\tilde r_j = p_j \bar r_j$, $\tilde\Sigma_j = p_j(\Sigma_j + \bar r_j \bar r_j^*)$ for $j \in L$, and define in addition $\tilde r_j^c := p_j r_2^c$. With assumption A7 below let

$$\alpha_j := e^* \Sigma_j^{-1} e, \qquad \beta_j := e^* \Sigma_j^{-1} \bar r_j, \qquad \gamma_j := \bar r_j^* \Sigma_j^{-1} \bar r_j,$$

and

$$\delta_j^c := (r_2^c)^2 \alpha_j - 2 r_2^c \beta_j + \gamma_j = (\bar r_j - r_2^c e)^* \Sigma_j^{-1} (\bar r_j - r_2^c e).$$

In the root define

$$\tilde p_0 := \sum_{j\in S(0)} \frac{p_j}{\delta_j^c + 1}, \qquad \tilde r_0 := \sum_{j\in S(0)} \frac{p_j}{\delta_j^c + 1} r_j, \qquad \tilde\Sigma_0 := \sum_{j\in S(0)} \frac{p_j}{\delta_j^c + 1} r_j r_j^*,$$

and furthermore $\bar r_0 = \tilde r_0 / \tilde p_0$, $\Sigma_0 = \tilde\Sigma_0/\tilde p_0 - \bar r_0 \bar r_0^*$, and $\tilde r_1^c := \tilde p_0 r^c$ (not $\tilde p_0 r_1^c$). Using A7 again, $\alpha_0, \beta_0, \gamma_0$ are then defined in analogy to $\alpha_j, \beta_j, \gamma_j$. Finally let $\delta_0^c := (r_1^c)^2 \alpha_0 - 2 r_1^c \beta_0 + \gamma_0$ and recall that t is the current time in $j \in V$.

Basic assumptions. For the general multi-period case assume

A7) $\forall j \in V$: $\Sigma_j > 0$.

A8) $\exists j \in V$: $\bar{r}_j \neq r_t^c e$.

A9) $r^c \neq 0$.

Remark. The conditions here are similar to A5 and A6 (or A1 and A3), but in addition we require nonzero cash returns r_t^c, $t = 1, \dots, T+1$. The opposite case would unnecessarily complicate the analysis and is not considered.

Lemma 15. *Under assumptions A7–A9, the constants α_j, γ_j are all positive, the δ_j^c are nonnegative, and at least one δ_j^c is positive. Moreover, $\tilde{p}_0 \in (0, 1]$.*

Proof. Positivity and nonnegativity are obvious, where $\delta_j^c = 0$ iff $\bar{r}_j = r_t^c e$. Now $\delta_j^c \geq 0$ implies $p_j/(\delta_j^c + 1) \in (0, p_j]$ and hence $\tilde{p}_0 \in (0, 1]$. □

In the following we use two different formulations for both the reward and the risk, involving again the conditional final-period risk and return. The latter is now $\rho_T(x_T, x_T^c) := \bar{r}_T^* x_T + r_{T+1}^c x_T^c$ with realizations $\rho_j(x_j, x_j^c) := \bar{r}_j^* x_j + r_{T+1}^c x_j^c$, and the discrete decision vector is $x = (x_j, x_j^c)_{j \in V}$.

Lemma 16. *The reward in the presence of cash can be written*

$$\rho(x) = \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = \sum_{j \in L} p_j \rho_j(x_j, x_j^c).$$

The risk has the two representations

$$\begin{aligned} R(x) &= \sum_{j \in L} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \rho(x)^2 \\ &= \sum_{j \in L} p_j x_j^* \Sigma_j x_j + \sum_{j \in L} p_j \rho_j(x_j, x_j^c)^2 - \rho(x)^2 =: R_c(x) + R_d(x). \end{aligned}$$

Proof. By definition, the continuous representation of the reward reads

$$\rho(x) = \mathbf{E}[r_{T+1}^* x_T + r_{T+1}^c x_T^c] = \mathbf{E}[\bar{r}_T^* x_T + r_{T+1}^c x_T^c] = \mathbf{E}[\rho_T(x_t, x_T^c)].$$

Similarly, using Lemma 11,

$$\begin{aligned} R(x) &= \mathbf{E}\left\{ \begin{pmatrix} x_T \\ x_T^c \end{pmatrix}^* \left[\begin{pmatrix} \Sigma_T & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} r_{T+1} \\ r_{T+1}^c \end{pmatrix} \begin{pmatrix} r_{T+1} \\ r_{T+1}^c \end{pmatrix}^* \right] \begin{pmatrix} x_T \\ x_T^c \end{pmatrix} \right\} - \rho(x)^2 \\ &= \mathbf{E}(x_T^* \Sigma_T x_T) + \mathbf{E}[(\bar{r}_T^* x_T + r_{T+1}^c x_T^c)^2] - \rho(x)^2. \end{aligned}$$

In both cases, the stated discrete formulae are readily obtained. □

Problem 11. Using the first formulation for reward and risk in Lemma 16, the multi-period mean-variance problem with cash reads

$$\begin{aligned} \min_x \quad & \sum_{j \in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2} \rho^2 \\ \text{s.t.} \quad & e^* x_0 + x_0^c = 1, \\ & e^* x_j + x_j^c = r_j^* x_i + r_t^c x_i^c \quad \forall j \in V^*, \\ & \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = \rho. \end{aligned}$$

Its Lagrangian is

$$
\begin{aligned}
L(x,\lambda,\mu;\rho) = {} & \sum_{j\in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2}\rho^2 \\
& - \lambda_0 (e^* x_0 + x_0^c - 1) - \sum_{j\in V^*} \lambda_j (e^* x_j + x_j^c - r_j^* x_i - r_t^c x_i^c) \\
& - \mu \Big(\sum_{j\in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c - \rho \Big).
\end{aligned}
$$

Theorem 17. *Problem 11 with $T = 1$ has the unique primal-dual solution*

$$
\begin{aligned}
x_j &= -\frac{\lambda_j}{\tilde{r}_j^c} \Sigma_j^{-1} (\bar{r}_j - r_2^c e), & x_0 &= -\frac{\lambda_0}{r_2^c \tilde{r}_1^c} \Sigma_0^{-1} (\bar{r}_0 - r_1^c e), \\
x_j^c &= \frac{1}{r_2^c} \left[\frac{\lambda_j}{\tilde{r}_j^c} (\gamma_j - r_2^c \beta_j + 1) + \mu \right], & x_0^c &= \frac{1}{r^c} \left[\frac{\lambda_0}{\tilde{r}_1^c} (\gamma_0 - r_1^c \beta_0 + 1) + \mu \right], \\
\lambda_j &= \frac{\tilde{r}_j^c}{\delta_j^c + 1} [r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu], & \lambda_0 &= \frac{\tilde{r}_1^c}{\delta_0^c + 1} (r^c - \mu) \equiv \tilde{\rho} (r^c - \mu), \\
\mu &= r^c \frac{\rho - \tilde{\rho}}{r^c - \tilde{\rho}}, & \tilde{\rho} &:= \frac{\tilde{r}_1^c}{\delta_0^c + 1} \in (0, r^c).
\end{aligned}
$$

The associated risk is

$$
R(x) \equiv \sigma^2(\rho) = \tilde{\rho} \frac{(r^c - \rho)^2}{r^c - \tilde{\rho}}.
$$

Its global minimum is attained at $\hat{\rho} = r^c$ and has value zero. The associated solution has 100 % cash: $(\hat{x}_0, \hat{x}_0^c) = (0, 1)$, $(\hat{x}_j, \hat{x}_j^c) = (0, r_1^c)$, $\hat{\lambda} = 0$, $\hat{\mu} = r^c$.

Proof. See appendix. □

Remark. Problem 11 also has a unique solution if $r_2^c \neq 0$ and $r_1^c = 0$, and it has multiple solutions for $r_2^c = 0$ (regardless of r_1^c). These situations are qualitatively different and quite unrealistic, however, and therefore not of interest here.

Discussion. Zero risk is now possible (with $(x_j, x_j^c) = (0, r_1^c)$ for $j \in L$) since balanced scenario returns do not require investments in risky assets. In fact, they require 100 % cash in both periods so that the riskless solution is unique.

It can be seen that the whole situation is actually covered by the results of the previous section if one replaces in each node x_j by (x_j, x_j^c), etc. This way the analysis extends again to the general multi-period case. The details are obtained precisely as in the previous section when proper replacements are carried out everywhere. However, the splitting into cash and risky assets makes the definitions of intermediate quantities and the specialization to the single-period case somewhat more involved.

To conclude this section we show that specifying ρ as a desired *minimal* reward has precisely the same effect as in the single-period case.

Problem 12. We modify Problem 11 by requiring ρ to be a lower bound on the reward (with associated slack $\theta \geq 0$ and dual slack $\eta \geq 0$),

$$
\begin{aligned}
\min_{x,\theta} \quad & \sum_{j\in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2}(\rho + \theta)^2 \\
\text{s.t.} \quad & e^* x_0 + x_0^c = 1, \\
& e^* x_j + x_j^c = r_j^* x_i + r_t^c x_i^c \quad \forall j \in V^*, \\
& \sum_{j\in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = \rho + \theta, \quad \theta \geq 0.
\end{aligned}
$$

The Lagrangian is now

$$L(x,\theta,\lambda,\mu,\eta;\rho) = \sum_{j\in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2}(\rho+\theta)^2$$
$$- \lambda_0(e^* x_0 + x_0^c - 1) - \sum_{j\in V^*} \lambda_j (e^* x_j + x_j^c - r_j^* x_i - r_t^c x_i^c)$$
$$- \mu\Bigg(\sum_{j\in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c - \rho - \theta\Bigg) - \eta\theta.$$

Theorem 18. *Problem 12 has a unique primal-dual solution. For $\rho \geq r^c$ one obtains $\theta = 0$, $\eta = \mu - \rho \geq 0$, and otherwise the same solution as in Problem 11. For $\rho \leq r^c$ one obtains $\eta = 0$ and $\theta = \rho - r^c \geq 0$, giving reward $\rho + \theta = r^c$ and the associated riskless solution of Problem 11. (At $\rho = r^c$ both cases coincide.)*

Proof. See appendix. □

2.3 Risky assets, cash, and loss

We are now entering the main section. Although it will still be a simplification, Problem 13 below covers all the essential aspects of the multi-period application model mentioned in the introduction. The notation and constants of the previous section remain valid, and the discrete decision vector is $x = (x_j, x_j^c, x_j^l)_{j\in V}$. We keep the general multi-period notation only in the problem statement and Lagrangian; the remaining analysis now concentrates on the two-period case.

Basic assumptions. In addition to the assumptions of the previous section we require positive cash returns and (as in Theorem 15) a no-arbitrage condition on the discrete part of the return distribution, involving $C_0 := \operatorname{conv}(\{r_j\}_{j\in L})$.

A7) $\forall j \in V$: $\Sigma_j > 0$.

A8) $\exists j \in V$: $\bar{r}_j \neq r_t^c e$.

A10) $r_1^c > 0$, $r_2^c > 0$.

A11) $r_1^c e \in \operatorname{int}(C_0)$.

Problem 13. The multi-period mean-variance problem with cash and loss reads

$$\begin{aligned} \min_x \quad & \sum_{j\in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2}\rho^2 \\ \text{s.t.} \quad & e^* x_0 + x_0^c + x_0^l = 1, \quad x_0^l \geq 0, \\ & e^* x_j + x_j^c + x_j^l = r_j^* x_i + r_t^c x_i^c, \quad x_j^l \geq 0 \quad \forall j \in V^*, \\ & \sum_{j\in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = \rho. \end{aligned}$$

Its Lagrangian is

$$L(x,\lambda,\eta,\mu;\rho) = \sum_{j\in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2}\rho^2$$
$$- \lambda_0(e^* x_0 + x_0^c + x_0^l - 1) - \sum_{j\in V^*} \lambda_j (e^* x_j + x_j^c + x_j^l - r_j^* x_i - r_t^c x_i^c)$$
$$- \sum_{j\in V} \eta_j x_j^l - \mu\Bigg(\sum_{j\in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c - \rho\Bigg).$$

Theorem 20 in the appendix characterizes optimal solutions in a way similar to the previous theorems. However, the loss variables x_j^l still appear in the formulae, and the case distinctions are more involved than in the single-period case (cf. Theorem 8) or in Problem 12. Later we will discuss this in part; for the time being, the following results provide more insight.

Lemma 17 (Arbitrage in Problem 13). *If assumption A11 is strictly violated, $r_1^c e \notin C_0$, then Problem 13 has a riskless solution for arbitrary ρ.*

Proof. Since C_0 is convex, $y \in \mathbb{R}^n$ exists so that $(r - r_1^c e)^* y \geq 1$ for $r \in C_0$. Given $\rho \in \mathbb{R}$, let $x_0 := (\rho/r_2^c - r_1^c)y$, $x_0^c := 1 - e^* x_0$, and $x_0^l := 0$ to obtain $w_j = (r_j - r_1^c e)^* x_0 + r_1^c \geq \rho/r_2^c$ for all $j \in L$. Now let $x_j := 0$, $x_j^c := \rho/r_2^c$, and $x_j^l := w_j - \rho/r_2^c \geq 0$. $\square$

For the analysis of actual riskless solutions consider in $\mathbb{R}^n$ the family of closed convex polyhedra

$$Z_0(\rho) := \left\{x \colon \left(\frac{r_j}{r_1^c} - e\right)^* x \geq \frac{\rho}{r^c} - 1 \;\forall j \in L\right\} = \bigcap_{j \in L} \bar{H}\left(e - \frac{r_j}{r_1^c}, 1 - \frac{\rho}{r^c}\right).$$

Lemma 18. *$Z_0(\rho)$ is nonempty (containing the origin) iff $\rho \leq r^c$. For $\rho < r^c$, $0 \in \mathrm{int}(Z_0(\rho))$ and*

$$Z_0(\rho) = \left(1 - \frac{\rho}{r^c}\right) Z_0(0).$$

Each $Z_0(\rho)$ is bounded and hence compact. In particular, $Z_0(r^c) = \{0\}$.

Remark. More generally one can show that the following four conditions are equivalent under the weaker no-arbitrage condition $r_1^c e \in C_0$.

1. $r_1^c e \in \mathrm{int}(C_0)$.
2. $Z_0(r^c) = \{0\}$.
3. $\exists \rho \in (-\infty, r^c]$: $Z_0(\rho)$ is bounded.
4. $\forall \rho \in (-\infty, r^c]$: $Z_0(\rho)$ is bounded.

As in Theorem 15 we choose the stronger condition $r_1^c e \in \mathrm{int}(C_0)$ to avoid unnecessary technical complications and ambiguities.

Proof. By A11 there exists a convex combination $\sum_{j \in L} \xi_j (r_j - r_1^c e) = 0$. Every $x \in Z_0(\rho)$ then satisfies

$$0 = \sum_{j \in L} \xi_j \left(\frac{r_j}{r_1^c} - e\right)^* x \geq \frac{\rho}{r^c} - 1.$$

Hence $Z_0(\rho) = \emptyset$ if $\rho > r^c$ and $0 \in Z_0(\rho)$ if $\rho \leq r^c$. Now let $\rho < r^c$. For $c > 0$,

$$\left(\frac{r_j}{r_1^c} - e\right)^* x \geq -1 \iff \left(\frac{r_j}{r_1^c} - e\right)^* cx \geq -c,$$

showing that $x \in Z_0(0)$ iff $(1 - \rho/r^c)x \in Z_0(\rho)$. To verify $0 \in \mathrm{int}(Z_0(\rho))$, observe that $Z_0(\rho)$ contains the ball around the origin with radius

$$\left(1 - \frac{\rho}{r^c}\right) \Big/ \max_{j \in L} \left\| \frac{r_j}{r_1^c} - e \right\|_2 > 0.$$

cf. Fig. 10. Condition 1) in the remark holds by A11. We prove 2), 3), and 4) in natural order. Assume first that 2) does not hold and let $y \in Z_0(r^c) \setminus \{0\}$. Then $(r_j/r_1^c - e)^* y \geq 0$ for all $j \in L$ and hence $(r/r_1^c - e)^* y \geq 0$ for $r \in C_0$, i.e., $C_0 \subseteq r_1^c e + \bar{H}(-y, 0)$. This yields the contradiction $r_1^c e \notin \mathrm{int}(C_0)$ and proves 2), which obviously implies 3). Assume now that 4) does not hold, that is, $Z_0(\rho)$ is unbounded for some $\rho \leq r^c$. Then, since $Z_0(\rho)$ is convex and $0 \in Z_0(\rho)$, $y \neq 0$ exists so that $cy \in Z_0(\rho)$ for all $c > 0$, that is, $(r_j/r_1^c - e)^* cy \geq \rho/r^c - 1$ for $j \in L$. This implies $(r_j/r_1^c - e)^* cy \geq 0$ and hence $(r_j/r_1^c - e)^* cy \geq \rho/r^c - 1$ for $c > 0$, $j \in L$, and $\rho \leq r^c$. Thus 3) implies 4): either none of the polyhedra is bounded or all of them. $\square$

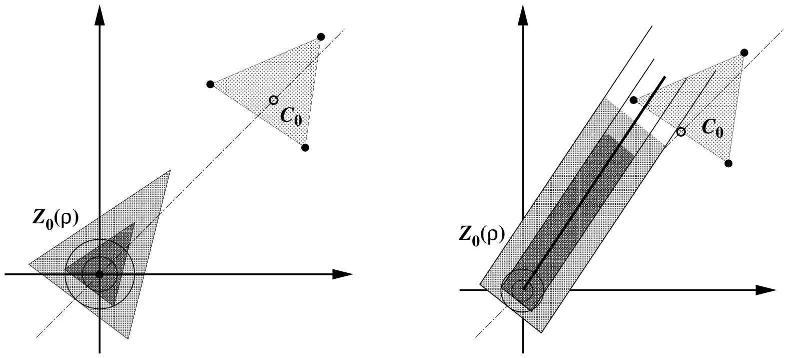

Figure 10: First-period returns $\{r_j\}, r_1^c e \in \mathbb{R}^2$ and associated zero-risk polyhedra $Z_0(\rho)$ with enclosed balls for three values of ρ (black $= r^c$, dark grey, medium grey). Left: compact case, $r_1^c e \in \operatorname{int}(C_0)$. Right: unbounded case, $r_1^c e \in \partial C_0$.

Theorem 19. *Consider the* feasible *solutions of Problem 13. Then:*

1. *Risk vanishes iff $x_j = 0$ and $x_j^c = \rho/r_2^c$ in all scenarios; it is positive else.*
2. *Leaf variables of riskless solutions depend uniquely on the root variables.*
3. *For $\rho < r^c$, risk vanishes on an $(n+1)$-dimensional cone over $Z_0(\rho)$.*
4. *For $\rho = r^c$, the unique zero-risk portfolio has 100 % cash.*
5. *For $\rho > r^c$, risk is strictly positive.*

Remarks. Strictly speaking, the riskless solutions for $\rho < r^c$ form an $(n+1)$-dimensional cone whose *projection* on the (x_0, x_0^l)-space has $Z_0(\rho)$ as base, see Fig. 11. Similarly, the projection for $\rho = r^c$ is $Z_0(r^c) \times \{0\}$, and in both cases the remaining variables are uniquely determined by feasibility. (If $r_1^c e \in \partial C_0$, then $x_0 \in Z_0(r^c)$, x_0^c, and x_j^l, $j \in L$, would not be unique for $\rho = r^c$.)

Proof. The condition in 1 is obviously sufficient for zero risk (cf. Lemma 17). Let $p := (p_j)_{j\in L}$ and $q := (\rho_j(x_j, x_j^c))_{j\in L}$. Then

$$R_d(x) = \sum_{j\in L} p_j q_j^2 - \Big(\sum_{j\in L} p_j q_j\Big)^2 = q^*[\operatorname{Diag}(p) - pp^*]q.$$

By Lemma 19, this quadratic form vanishes if $q = \rho e$ and is positive else. Moreover, $R_c(x) \ge 0$ for all x, and $R_c(x) = 0$ iff $x_j = 0$ for all $j \in L$. This proves statement 1. The zero-risk condition requires $w_j \equiv r_j^* x_0 + r_1^c x_0^c \ge \rho/r_2^c$ in all scenarios. If this implied restriction on the root variables holds, then the leaf variables are uniquely determined by statement 1 and $x_j^l = w_j - x_j^c \ge 0$, proving statement 2. Since $x_0^c = 1 - e^* x_0 - x_0^l$, the inequality $w_j \ge \rho/r_2^c$ is equivalent to $x_0 \in Z_0(\rho + r^c x_0^l)$. By Lemma 18 we conclude that x is optimal for $\rho < r^c$ iff $x_0^l \in [0, 1 - \rho/r^c]$ and $x_0 \in Z_0(\rho + r^c x_0^l)$, see Fig. 11. This proves statement 3. Statements 4 and 5 follow similarly: for $\rho = r^c$ the zero-risk cone degenerates (only its vertex remains), and for $\rho > r^c$ it becomes empty. ☐

Lemma 19. *Let $p \in \mathbb{R}^N$, $p > 0$, $e^* p = 1$, and $f(q) := \frac{1}{2} q^*[\operatorname{Diag}(p) - pp^*]q$. Then*

$$\min_q f(q) \quad s.t. \quad p^* q = \rho$$

has the unique minimizer $q = \rho e$, with optimal value $f(\rho e) = 0$.

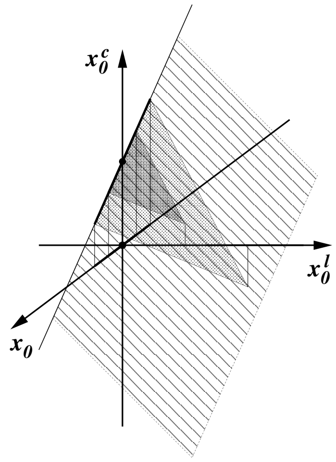

Figure 11: Zero-risk cones over $Z_0(\rho)$ for three values of ρ (black = r^c, dark grey, light grey); compact case. (Here $Z_0(\rho)$ is a segment of the x_0 axis.)

Proof. The Lagrangian of the minimization problem is

$$L(q,\eta) = \frac{1}{2}q^*[\mathrm{Diag}(p) - pp^*]q - \eta(p^*q - \rho),$$

yielding the optimality condition

$$0 = \frac{\partial L}{\partial q} = \mathrm{Diag}(p)q - pp^*q - \eta p = \mathrm{Diag}(q)p - \rho p - \eta p.$$

Since $p > 0$, this gives $q_j - \rho - \eta = 0$ for $j = 1, \ldots, N$, and hence $q = (\rho + \eta)e$. Now $p^*q = \rho$ implies $\eta = 0$, as required. Clearly, $f(\rho e) = 0$. □

Statement 3 of Theorem 19 says that $x_0^l > 0$ ($x_j^l > 0$) is possible when the total cash return $r^c = r_2^c r_1^c w_0$ (the second-period cash return $r_2^c w_j$) exceeds the desired reward. The following result shows that surplus money is actually *necessary* for positive loss variables: if there is no surplus money, then positive amounts x_j^l are better invested in cash to reduce risk.

Lemma 20. *Let x be an optimal solution of Problem 13 for $\rho > r^c$.*

1. *If $x_j \neq 0$ in a scenario $j \in L$, then $x_j^l = 0$.*
2. *$x_0^l = 0$.*

Remark. Statement 1 of the lemma is also obtained from Theorem 20, and moreover the plausible fact that $x_j^l = 0$ in scenario j if $\rho_j < \rho$. However, the proof of that theorem is not constructive.

Proof. Assume $x_j^l > 0$. We modify the local variables in scenario j to construct a better feasible solution. Let $a := (\bar{r}_j - r_2^c e)^* x_j$, define $\epsilon \in (0, 1]$ as

$$\epsilon := \begin{cases} \min(1, r_2^c x_j^l / a) & \text{if } a > 0, \\ 1 & \text{else,} \end{cases}$$

and replace (x_j, x_j^c, x_j^l) by

$$(\hat{x}_j, \hat{x}_j^c, \hat{x}_j^l) := (x_j, x_j^c, x_j^l) - \epsilon(x_j, -\bar{r}_j^* x_j / r_2^c, a / r_2^c).$$

Then $\hat{x}_j^l \geq 0$, $\hat{w}_j = w_j$, $\hat{\rho}_j = \rho_j$, and the risk is reduced by the positive amount $[1 - (1-\epsilon)^2] x_j^* \Sigma_j x_j$. This proves statement 1. For $x_0^l > 0$ we modify x_0^c and all x_j^l. Let $(\hat{x}_0, \hat{x}_0^c, \hat{x}_0^l) := (x_0, x_0^c + x_0^l, 0)$ and $(\hat{x}_j, \hat{x}_j^c, \hat{x}_j^l) := (x_j, x_j^c, x_j^l + r_1^c x_0^l)$ for $j \in L$. Clearly, $\hat{x}$ is feasible, $R(\hat{x}) = R(x)$, and $\hat{x}_j^l \geq r_1^c x_0^l > 0$ in all scenarios. Thus, by statement 1, $\hat{x}$ and consequently x are not optimal. □

Discussion. As in the one-period case, the 100 % cash solution plays a key role: it has the largest reward among all riskless solutions. But now the solutions become degenerate for small rewards $\rho < r^c$, even if loss is allowed only in the second period, i.e., if $x_0^l = 0$ is fixed. This does not happen in Problem 12 (the modification of Problem 11 with bounded reward $\rho(x) \geq \rho$), which behaves precisely as the corresponding single-period problem. Of course, the degeneracy occurs only for practically irrelevant rewards, and even then it can easily be avoided. (One may choose the vertex of the zero-risk cone, i.e., $x_0^l = 1 - \rho/r^c$. This removes any surplus money immediately and gives a unique solution.)

The only case of practical interest is $\rho > r^c$, when the solution of Problem 11 remains optimal in Problem 12. Why do we prefer the loss formulation, Problem 13? Obviously the risk cannot be higher than in Problem 11 since every optimal solution of the latter remains feasible in the former problem. Actually it turns out that the loss formulation gives *strictly lower* risk in most cases, i.e., it allows better solutions than Problem 12. To develop a geometric understanding for this observation, we compare Problems 11 and 13 in a simplified situation. A reformulation eliminates all the budget equations and most of the portfolio variables in favor of the individual scenario returns. The two risk terms R_c, R_d are then used to explain in which cases (and how) an optimal solution of Problem 11 can be modified to give a better feasible solution of Problem 13.

Problem 14. Consider as example a portfolio consisting of just one risky asset and cash, using the second formulation of reward and risk in Lemma 16. Include loss assets in the leaves but not in the root, and write the problem with scenario returns ρ_j as additional variables,

$$\begin{aligned}
\min_{x,\{\rho_j\}} \quad & \sum_{j\in L} \frac{1}{2} p_j [\Sigma_j x_j^2 + \rho_j^2] - \frac{1}{2}\rho^2 \\
\text{s.t.} \quad & \rho_j = \bar{r}_j x_j + r_2^c x_j^c \quad \forall j \in L, \\
& x_0 + x_0^c = 1, \\
& x_j + x_j^c + x_j^l = r_j x_0 + r_1^c x_0^c, \quad x_j^l \geq 0 \quad \forall j \in L, \\
& \sum_{j\in L} p_j \rho_j = \rho.
\end{aligned}$$

This specialization of Problem 13 is only considered for $\rho \geq r^c$, but for arbitrary $\rho \in \mathbb{R}$ as a specialization of Problem 11, i.e., when all the loss variables $x_j^l = 0$ are fixed. (In these cases the solution is unique by Theorems 17 and 20.)

Lemma 21. *For simplicity assume $\bar{r}_j \neq r_2^c$ for all $j \in L$. (A8 guarantees this for just one $j \in L$.) Let $\phi_j := \Sigma_j/(\bar{r}_j - r_2^c)^2 > 0$, $\psi_j := r_2^c(r_j - r_1^c)$, and $\theta_j \equiv r_2^c x_j^l$. Then Problem 14 is equivalent to*

$$\begin{aligned}
\min_{x_0,\{\rho_j,\theta_j\}} \quad & \sum_{j\in L} \frac{1}{2} p_j [\phi_j(\rho_j - r^c - \psi_j x_0 + \theta_j)^2 + \rho_j^2] - \frac{1}{2}\rho^2 \\
\textit{s.t.} \quad & \sum_{j\in L} p_j \rho_j = \rho, \quad \theta_j \geq 0 \quad \forall j \in L.
\end{aligned} \tag{3}$$

The optimal scenario returns are

$$\rho_j = \frac{\phi_j(r^c + \psi_j x_0 - \theta_j) + \mu}{\phi_j + 1}, \quad j \in L. \tag{4}$$

Moreover, the optimal reward multiplier μ has the same value in both problems.

Proof. Eliminate

$$x_0^c = 1 - x_0, \qquad x_j^c = (r_j - r_1^c)x_0 + r_1^c - x_j - x_j^l.$$

Then substitute x_j^c into the ρ_j equation and use ψ_j and θ_j to obtain

$$\rho_j = (\bar{r}_j - r_2^c)x_j + r^c + \psi_j x_0 - \theta_j.$$

Solving for x_j, inserting it into the objective and using ϕ_j yields Problem (3). Differentiating the Lagrangian with respect to the returns ρ_j gives optimality conditions

$$\phi_j(\rho_j - r^c - \psi_j x_0 + \theta_j) + \rho_j = \mu,$$

from which one obtains the expression (4). This derivation holds for the case without loss, too: one just has to set $x^l_j = \theta_j = 0$ everywhere. Finally, when the problem transformations above are applied to the full primal-dual system of optimality conditions, it is observed that μ has the same value in both problems. □

Discussion. Consider the case $\theta_j = 0$ first (Problem 11). We have to choose optimal values for x_0 and for the scenario returns ρ_j so that their mean equals ρ. Defining $d_j := r^c + \psi_j x_0$ gives the continuous risk part

$$R_c = \sum_{j \in L} p_j \phi_j (\rho_j - d_j)^2 \geq 0.$$

This is a weighted average of scenario risks $\phi_j(\rho_j - d_j)^2$, each of which defines a parabola characterized by its offset d_j and curvature ϕ_j. Both magnitude and distance of the offsets are influenced by the common "spread factor" x_0: they all coincide with r^c if $x_0 = 0$ (100 % cash), whereas $x_0 = 1$ (no cash) yields the discrete distribution $d_j = r^c_2 r_j$ with mean $d := r^c_2 \mathbf{E}(\{r_j\})$. Clearly, the continuous risk R_c is small when all the scenario returns are close to their respective offsets, while the discrete risk R_d is small when they are close to each other. Thus, loosely speaking, x_0 has the job to balance the scenarios by adjusting d (close to ρ) without spreading the offsets too much.

Only one detail changes when $\theta_j > 0$ is allowed (Problem 13): each offset d_j is replaced by $d_j - \theta_j$, that is, the parabolas may be shifted to the left *separately* in each scenario.

We can now explain the risk reduction mechanism. Consider an optimal solution of Problem 14 *without* loss. Typically there will be "good" scenarios (fortunate for the investor, with large offsets $d_j > \rho$) and "bad" scenarios (unfortunate for the investor, with small offsets $d_j < \rho$). Moreover, one expects that some of the scenario returns will lie on their local upper branch ($\rho_j > d_j$) and some on their local lower branch ($\rho_j < d_j$). If in a good scenario the optimal return lies on the lower branch, $\rho < \rho_j < d_j$, then its contribution to the risk is canceled by shifting the parabola to the left, $\theta_j := d_j - \rho_j > 0$. Clearly, nothing else changes, so that this gives a suboptimal feasible solution of Problem 13 which is better than the optimal solution of Problem 11. The mechanism here is precisely the same as in a single period (Problem 5) except that it now occurs *locally* in individual scenarios.

Of course, the *optimal* solution of Problem 13 will readjust all the variables *globally.* Now x_0 still has the job to balance scenarios by adjusting d close to ρ, but only without spreading the *small* offsets $d_j < \rho$ too much. The large offsets $d_j > \rho$ produce surplus money in *all* sufficiently good scenarios. These scenarios do not contribute to R_c, and all their returns are equal (and slightly larger than ρ to balance the bad ones). This is proved in Theorem 20: surplus money $x^l_j > 0$ implies $r^c_2 w_j > \mu$ and $\rho_j = \mu \geq \rho$. It means that a jump discontinuity is produced in the distribution of final wealth, to which each of the riskless scenarios contributes a fraction. It also means that the bad scenarios dominate the resulting risk: again we have an approximate minimization of downside risk.

When looking for an instance of Problem 11 with $\rho < \rho_j < d_j$ one might try Problem 3 with only two scenarios. However, the following results show that the effect cannot occur if $n = 1$ and $N = 2$: in that case (with $\rho > r^c$) the single degree of freedom in x_0 is sufficient to balance the scenarios well. After proving this we give an example of the risk reduction with $n = 1$ and $N = 3$. A slight modification finally shows that risk reduction can occur even in bad scenarios, that is, optimal solutions of Problem 11 may have $\rho_j < d_j < \rho$.

For the comparison of signs we define the equivalence relation

$$a \sim b \quad :\Longleftrightarrow \quad \operatorname{sign}(a) = \operatorname{sign}(b) \in \{-1, 0, 1\}$$

and use the fact that $\{-1, 0, 1\}$ is a multiplicative subgroup in $\mathbb{R}$.

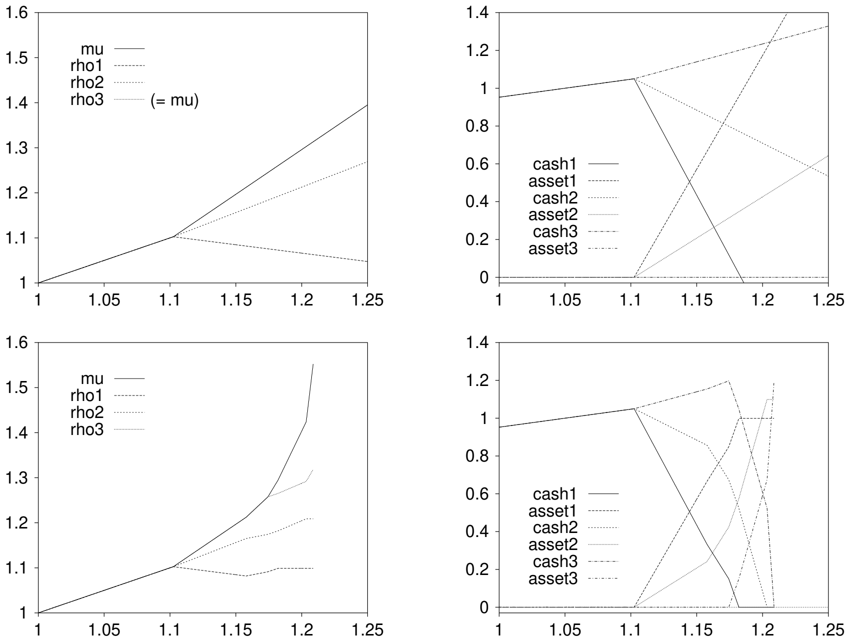

Figure 12: Optimal scenario returns (left) and investments (right) for the small example problem. Asset values unconstrained (top) and nonnegative (bottom; $\bar{r}^{\max} = 1.21$).

Lemma 22. *In Problem 14* without *loss assets,* $\rho_j - d_j \sim s_j(\rho - r^c)$ *where*

$$s_j := \sum_{k \in L \setminus \{j\}} \frac{p_k}{\delta_k^c + 1}(r_k - r_j)(r_k - r_1^c).$$

Proof. See appendix. □

Remark. Notice that s_j depends only on the stochastic data (the return distribution), and in particular *not* on ρ.

Corollary 2. *In Problem 14 with two scenarios and without loss assets, both ρ_1 and ρ_2 lie on their respective upper branches if $\rho > r^c$.*

Proof. Assumption A11 yields $r_1 < r_1^c < r_2$, and the sums s_j consist of one single term each since $N = 2$. This implies that s_1, s_2 are both positive. □

Remark. A weak version of this result holds for $r_1^c \in C_0$. Then $r_1 \le r_1^c \le r_2$, and s_1, s_2 are both nonnegative. Even that weaker version would prevent any beneficial shift $\theta_j > 0$.

The example with $N = 3$ is simple: let $(p_1, p_2, p_3) = (0.2, 0.6, 0.2)$, $(r_1, r_2, r_3) = (1.0, 1.1, 1.2)$, $r_1^c = 1.05$, $\Sigma_1 = \Sigma_2 > 0$, and $j = 3$. Then $\delta_1^c = 0.05^2/\Sigma_1 = \delta_2^c$, and hence

$$s_3 = \sum_{k=1}^{2} \frac{p_k}{\delta_k^c + 1}(r_k - r_3)(r_k - r_1^c) = \frac{0.2(-0.2)(-0.05) + 0.6(-0.1)(0.05)}{\delta_1^c + 1} < 0.$$

The corollary proves that ρ_3 lies on the lower branch even though $r_1^c \in \operatorname{int}(C_0)$; thus introducing the loss asset x_3^l reduces the optimal risk. (The values of Σ_3, r_2^c, and $\bar{r}_1, \bar{r}_2, \bar{r}_3$ do not matter here as long as A7, A8, A10, and A11 are satisfied. The reader may check that this holds for $\Sigma_j = 1$, $r_2^c = 1.05$, and $\bar{r}_j = 1.1$, $j \in \{1, 2, 3\}$.) Setting instead $r_1^c := 1.15$ and $\Sigma_2 = \Sigma_3 > 0$ yields $s_1 < 0$. This shows that risk reduction can also occur in bad scenarios—at least for unreasonably large r_t^c. Results for this small example problem (with loss assets included) are displayed in Fig. 12.

2.4 The general multi-period case

The general situation with loss assets becomes quite complicated and we do not attempt a thorough analysis. Instead, we state some plausible extensions of the two-period results that have been confirmed in practical computations.

Surplus money can now appear in any node $j \in L_t$, $t \in \{0, \dots, T\}$, if the (partial) scenario is sufficiently good up to that point, that is, if

$$w_j > w_t^{\min} := \frac{\mu}{r_{T+1}^c \cdots r_{t+1}^c}.$$

(As before the root can be excluded by considering only "large" desired rewards $\rho > r^c$, with $r^c := r_{T+1}^c \cdots r_1^c$.) When there is surplus money, the whole scenario subtree rooted in j does not contribute to the risk provided that sufficient node capitals w_k are maintained. Implicitly this condition defines a zero-risk polyhedron whose geometry is determined by the subtree's discrete return distribution. Of course, if $j \in L_{T-1}$, then one gets a cone similar to the one considered before, but depending on w_j. These observations imply that *the generic optimal solution is highly degenerate*: any reasonable return discretization will include good and bad partional scenarios on each level, and surplus money will almost always appear somewhere in the tree. However, even in the zero-risk subtrees all leaf variables are again uniquely determined by the lower-level variables.

Obviously $w_{t+s}^{\min}$ is the sufficient capital for a node $k \in L_{t+s}$ in the zero-risk subtree, and the easiest way to maintain that amount is to invest precisely $w_t^{\min}$ in cash and remove the rest ("invest" in x_j^l). Thus all surplus money is taken out immediately in the root of the subtree and each remaining node has 100 % cash.

3 Conclusions

We have seen that multi-period mean-variance problems behave similar to their single-period counterparts in many respects. In specific, it is possible to avoid overperformance by allowing to remove capital. Small desired rewards $\rho \le r^c$ are met exactly at zero risk. In that case all the capital is either invested in cash or removed; thus minimizing the variance is trivially equivalent to minimizing the semi-variance (or any other downside risk measure) *without* removing capital but allowing to exceed the desired reward. That is, with $x = (x_j, x_j^c)_{j\in V}$ and in abbreviated notation, the problem

$$\min_x R(x) \qquad \text{s.t.} \qquad \rho(x) = \rho, \quad e^* x_j + x_j^c \le w_j \quad \forall j \in V,$$

is equivalent to the downside risk problem

$$\min_x R_\rho^2(x) \qquad \text{s.t.} \qquad \rho(x) \ge \rho, \quad e^* x_j + x_j^c = w_j \quad \forall j \in V.$$

(Of course, the solutions of the second problem differ insofar as surplus money is invested in cash instead of being removed.) For moderate values $\rho > r^c$ one cannot avoid overperformance completely, but in effect the first problem still tends to minimize the semi-variance. More precisely, the discrete part R_d approximates its downside version due to the existence of zero-risk subtrees. The quality of that approximation decreases as ρ increases so that for large values the risk measure becomes a blending of variance and semi-variance. Note that there is no such gradual process in the single-period case, but there is a close similarity of single-period downside risk (Theorem 15) and multi-period zero-risk polyhedra (Lemma 18). We may conclude that Problem 13 is a reasonable multi-period model for an investor who wishes to minimize the semi-variance rather than the variance of final wealth.

The previous comparison also gives some hints how an optimal policy of Problem 13 should be interpreted. Again, positive values of x_j^l do not suggest to burn that amount. They indicate the presence of surplus money which the investor may spend immediately without risking to miss her goal, or which she may invest in cash to obtain a riskless extra profit. Of course, she may also consume part of the surplus and invest the rest. Thus, if the investor implements the optimal policy over the full planning horizon, she will approximately minimize the risk of ending up with less than the desired amount, regardless of her choice. Interestingly, the second alternative (investing) amounts to a *single-period* strategy with *predetermined intermediate decisions,* which may be useful when the investor cannot react to the market until the end of her planning horizon, or for some reason does not wish to do so.

However, it should be noted that the problem under consideration is not *time-invariant* due to the reward condition. That condition involves an expectation over all scenarios, that is, over the potential futures at $t = 0$. But at $t = 1$ most of these potential futures become impossible, whichever scenario realizes. The terminal condition $\rho(x) = \rho$ or $\rho(x) \geq \rho$ can usually not be satisfied when the restricted expectation over the subtree is taken—unless it happens to be a zero-risk subtree. Therefore only the immediate decision will be of interest for the typical investor. Rather than following the original future policy, she will adjust the reward and solve the problem anew for each decision. Of course, the investor may also build an extended model after each period in pursuing a moving horizon technique.

In any case it seems appropriate to consider all riskless strategies (in addition to the efficient ones) as reasonable choices in multi-period decision models. This does no harm since it includes all the standard alternatives, but it opens up new possibilities like the trick described above.

Let us finally point out two issues that might be interesting subjects of future research. First, the model presented here does not include any preferences of consumption, although one can easily specify hard constraints (exact, minimal, or maximal consumption) through a cash flow. However, it is not clear how one should incorporate (soft) preferences and how the result would be related to long-term models based on utility of consumption. Second, the multi-period setting enables the investor to control higher moments of her distribution of final wealth—at least to some extent. How would risk measures involving skewness, e.g., behave in the context of our model?

Appendix

The appendix contains some proofs and a theorem that would disrupt the line of thought in the main body of the multi-period section.

Proof of Theorem 16. The system of optimality conditions (for two periods) can be written

$$\left(\begin{array}{cccc|ccccc} 0 & & & & e & -r_1 & \dots & -r_N & 0 \\ & \tilde\Sigma_1 & & & & e & & & \tilde r_1 \\ & & \ddots & & & & \ddots & & \vdots \\ & & & \tilde\Sigma_N & & & & e & \tilde r_N \\ \hline e^* & & & & & & & & \\ -r_1^* & e^* & & & & & & & \\ \vdots & & \ddots & & & & & & \\ -r_N^* & & & e^* & & & & & \\ 0 & \tilde r_1^* & \dots & \tilde r_N^* & & & & & \end{array}\right) \left(\begin{array}{c} x_0 \\ x_1 \\ \vdots \\ x_N \\ \hline -\lambda_0 \\ -\lambda_1 \\ \vdots \\ -\lambda_N \\ -\mu \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ \hline 1 \\ 0 \\ \vdots \\ 0 \\ \rho \end{array}\right).$$

For $j \in L$, $x_j = \tilde\Sigma_j^{-1}(\lambda_j e + \mu \tilde r_j)$ is immediately obtained from the j-th dual feasibility condition. Substitution into the budget equation for x_j and into dual feasibility condition 0, respectively, yields $\lambda_j = (r_j^* x_0 - \mu\tilde\beta_j)/\tilde\alpha_j$ and

$$0 = -\lambda_0 e + \sum_{j\in L} \lambda_j r_j = -\lambda_0 e + \sum_{j\in L} \frac{r_j^* x_0 - \mu\tilde\beta_j}{\tilde\alpha_j} r_j = -\lambda_0 e + \tilde\Sigma_0 x_0 - \mu \tilde r_0,$$

giving $x_0 = \tilde\Sigma_0^{-1}(\lambda_0 e + \mu\tilde r_0)$. The budget equation for x_0 now reads

$$1 = e^* x_0 = e^* \tilde\Sigma_0^{-1}(\lambda_0 e + \mu \tilde r_0) = \lambda_0 \tilde\alpha_0 + \mu\tilde\beta_0,$$

yielding $\lambda_0 = (1-\mu\tilde\beta_0)/\tilde\alpha_0$. From the reward equation one finally obtains

$$\begin{aligned}
\rho &= \sum_{j\in L} \tilde r_j^* x_j = \sum_{j\in L} (\lambda_j \tilde\beta_j + \mu\tilde\gamma_j) = \sum_{j\in L} \left(\frac{\tilde\beta_j}{\tilde\alpha_j} r_j^* x_0 - \mu \frac{\tilde\beta_j^2}{\tilde\alpha_j} + \mu\tilde\gamma_j \right) \\
&= \tilde r_0^* x_0 + \mu \sum_{j\in L} \frac{\tilde\delta_j}{\tilde\alpha_j} = \lambda_0\tilde\beta_0 + \mu\tilde\gamma_0 + \mu \sum_{j\in L} \frac{\tilde\delta_j}{\tilde\alpha_j} \\
&= \frac{\tilde\beta_0}{\tilde\alpha_0} - \mu\frac{\tilde\beta_0^2}{\tilde\alpha_0} + \mu\tilde\gamma_0 + \mu \sum_{j\in L} \frac{\tilde\delta_j}{\tilde\alpha_j} = \frac{\tilde\beta_0}{\tilde\alpha_0} + \mu \sum_{j\in V} \frac{\tilde\delta_j}{\tilde\alpha_j}.
\end{aligned}$$

This gives the reward multiplier using Lemma 12,

$$\mu = \left(\rho - \frac{\tilde\beta_0}{\tilde\alpha_0} \right) \Big/ \sum_{j\in V} \frac{\tilde\delta_j}{\tilde\alpha_j}.$$

(Here we need assumption A6: the denominator vanishes if $\tilde\delta_j = 0$ for all $j \in V$.) As to the global minimum, we have

$$\begin{aligned}
x_j^* \tilde\Sigma_j x_j &= \lambda_j(\lambda_j\tilde\alpha_j + \mu\tilde\beta_j) + \mu(\lambda_j\tilde\beta_j + \mu\tilde\gamma_j) \\
&= \lambda_j r_j^* x_0 + \mu \frac{\tilde\beta_j r_j^* x_0 + \mu\tilde\delta_j}{\tilde\alpha_j} = \frac{(r_j^* x_0)^2}{\tilde\alpha_j} + \mu^2 \frac{\tilde\delta_j}{\tilde\alpha_j}
\end{aligned}$$

and similarly

$$x_0^* \tilde\Sigma_0 x_0 = \lambda_0 + \mu \frac{\tilde\beta_0 + \mu\tilde\delta_0}{\tilde\alpha_0} = \frac{1}{\tilde\alpha_0} + \mu^2 \frac{\tilde\delta_0}{\tilde\alpha_0}.$$

Therefore,

$$\begin{aligned}
\sum_{j\in L} x_j^* \tilde\Sigma_j x_j &= \sum_{j\in L} \left(\frac{(r_j^* x_0)^2}{\tilde\alpha_j} + \mu^2 \frac{\tilde\delta_j}{\tilde\alpha_j} \right) = x_0^* \tilde\Sigma_0 x_0 + \mu^2 \sum_{j\in L} \frac{\tilde\delta_j}{\tilde\alpha_j} \\
&= \frac{1}{\tilde\alpha_0} + \mu^2 \sum_{j\in V} \frac{\tilde\delta_j}{\tilde\alpha_j} = \frac{1}{\tilde\alpha_0} + \left(\rho - \frac{\tilde\beta_0}{\tilde\alpha_0} \right)^2 \Big/ \sum_{j\in V} \frac{\tilde\delta_j}{\tilde\alpha_j}.
\end{aligned}$$

Subtracting ρ^2 yields a risk expression of the form $\sigma^2(\rho) = s + (\rho - c)^2/d - \rho^2$. Since the optimal portfolio x is an affine function of ρ and R is convex quadratic, the efficient frontier is either strictly convex (iff $d < 1$), or $\sigma^2(\rho) \equiv 0$ (iff $d = 1$ and $c = s = 0$). But $s = \tilde\alpha_0^{-1} > 0$ by Lemma 12; therefore $\sigma^2(\rho)$ has the global minimum $s + c^2/(1-d)$ at $\hat\rho = c/(1-d)$, as stated. □

Proof of Theorem 17. The system of optimality conditions can be written

$$\begin{aligned}
\partial L/\partial x_0 = 0: &\qquad -\lambda_0 e + \textstyle\sum_{j\in L} \lambda_j r_j = 0, \\
\partial L/\partial x_0^c = 0: &\qquad -\lambda_0 + \textstyle\sum_{j\in L} \lambda_j r_1^c = 0, \\
\partial L/\partial x_j = 0: &\qquad \tilde\Sigma_j x_j + r_2^c \tilde r_j x_j^c - \lambda_j e - \mu \tilde r_j = 0 \quad \forall j \in L, \\
\partial L/\partial x_j^c = 0: &\qquad r_2^c \tilde r_j^* x_j + r_2^c \tilde r_j^c x_j^c - \lambda_j - \mu \tilde r_j^c = 0 \quad \forall j \in L, \\
\partial L/\partial \lambda_0 = 0: &\qquad e^* x_0 + x_0^c = 1, \\
\partial L/\partial \lambda_j = 0: &\qquad e^* x_j + x_j^c - r_j^* x_0 - r_1^c x_0^c = 0 \quad \forall j \in L, \\
\partial L/\partial \mu = 0: &\qquad \textstyle\sum_{j\in L} \tilde r_j^* x_j + \tilde r_j^c x_j^c = \rho.
\end{aligned}$$

The dual feasibility condition $\partial L/\partial x_j^c = 0$ gives

$$x_j^c = \frac{1}{r_2^c}\Big[-\bar r_j^* x_j + \frac{\lambda_j}{\tilde r_j^c} + \mu\Big]. \tag{5}$$

Substitution into $\partial L/\partial x_j = 0$ yields

$$0 = \tilde\Sigma_j x_j + \tilde r_j\Big[-\bar r_j^* x_j + \frac{\lambda_j}{\tilde r_j^c} + \mu\Big] - \lambda_j e - \mu\tilde r_j = p_j\Sigma_j x_j + \frac{\lambda_j}{r_2^c}(\bar r_j - r_2^c e),$$

which gives x_j and, upon substitution into (5), x_j^c:

$$x_j = -\frac{\lambda_j}{\tilde r_j^c}\Sigma_j^{-1}(\bar r_j - r_2^c e), \qquad x_j^c = \frac{\lambda_j}{r_2^c\tilde r_j^c}(\gamma_j - r_2^c\beta_j + 1).$$

Therefore the budget equation $\partial L/\partial\lambda_j = 0$ reads

$$0 = e^* x_j + x_j^c = -\frac{\lambda_j}{\tilde r_j^c}(\beta_j - r_2^c\alpha_j) + \frac{\lambda_j}{r_2^c\tilde r_j^c}(\gamma_j - r_2^c\beta_j + 1) + \frac{\mu}{r_2^c} = \frac{\lambda_j}{r_2^c}\frac{\delta_j^c + 1}{\tilde r_j^c} + \frac{\mu}{r_2^c},$$

from which one obtains

$$\lambda_j = \frac{\tilde r_j^c}{\delta_j^c + 1}[r_2^c(r_j^* x_0 + r_1^c x_0^c) - \mu].$$

Now we can proceed with the root variables. The condition $\partial L/\partial x_0^c = 0$ reads

$$0 = -\lambda_0 + \sum_{j\in L} r_1^c \frac{\tilde r_j^c}{\delta_j^c + 1}[r_2^c(r_j^* x_0 + r_1^c x_0^c) - \mu] = -\lambda_0 + r^c r_2^c \tilde r_0^* x_0 + r^c \tilde r_1^c x_0^c - \mu\tilde r_1^c,$$

giving

$$x_0^c = \frac{1}{r^c}\Big[-r_2^c\bar r_0^* x_0 + \frac{\lambda_0}{\tilde r_1^c} + \mu\Big]. \tag{6}$$

Similarly, after inserting λ_j and then x_0^c, the condition $\partial L/\partial x_0 = 0$ reads

$$\begin{aligned} 0 &= -\lambda_0 e + \sum_{j\in L} r_j \frac{\tilde r_j^c}{\delta_j^c + 1}[r_2^c(r_j^* x_0 + r_1^c x_0^c) - \mu] \\ &= -\lambda_0 e + (r_2^c)^2\tilde\Sigma_0 x_0 + r^c r_2^c \tilde r_0 x_0^c - \mu r_2^c \tilde r_0 = (r_2^c)^2 \tilde p_0 \Sigma_0 x_0 + \frac{\lambda_0}{r_1^c}(\bar r_0 - r_1^c e). \end{aligned}$$

This yields x_0 and, by substitution into (6), x_0^c:

$$x_0 = -\frac{\lambda_0}{r_2^c\tilde r_1^c}\Sigma_0^{-1}(\bar r_0 - r_1^c e), \qquad x_0^c = \frac{\lambda_0}{r^c\tilde r_1^c}(\gamma_0 - r_1^c\beta_0 + 1) + \frac{\mu}{r^c}.$$

Thus the root budget equation is

$$0 = e^* x_0 + x_0^c = -\frac{\lambda_0}{r_2^c\tilde r_1^c}(\beta_0 - r_1^c\alpha_0) - \frac{\lambda_0}{r^c\tilde r_1^c}(\gamma_0 - r_1^c\beta_0 + 1) + \frac{\mu}{r^c} = \frac{\lambda_0}{r^c}\frac{\delta_0^c + 1}{\tilde r_1^c} + \frac{\mu}{r^c},$$

which gives

$$\lambda_0 = \frac{\tilde r_1^c}{\delta_0^c + 1}(r^c - \mu) = \tilde\rho(r^c - \mu).$$

By Lemma 15 we have $\tilde{p}_0 \in (0,1]$ and $\delta_0^c \geq 0$. Moreover, if $\tilde{p}_0 = 1$ then $\delta_j^c = 0$ for $j \in L$ and thus $\delta_0^c > 0$. This proves the inclusion $\tilde{\rho} \equiv r^c \tilde{p}_0/(\delta_0^c + 1) \in (0, r^c)$. Altogether, the previous results give

$$\tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = -\frac{\lambda_j}{r_2^c}(\gamma_j - r_2^c \beta_j) + \frac{\lambda_j}{r_2^c}(\gamma_j - r_2^c \beta_j + 1) + p_j \mu = \frac{\lambda_j}{r_2^c} + p_j \mu \tag{7}$$

and similarly

$$r_2^c \tilde{r}_0^* x_0 + \tilde{r}_1^c x_0^c = -\frac{\lambda_0}{r^c}(\gamma_0 - r_1^c \beta_0) + \frac{\lambda_0}{r^c}(\gamma_0 - r_1^c \beta_0 + 1) + \tilde{p}_0 \mu = \frac{\lambda_0}{r^c} + \tilde{p}_0 \mu. \tag{8}$$

Upon insertion of λ_j and then λ_0 the reward condition reads

$$\begin{aligned} \rho = \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c &= \sum_{j \in L} \frac{p_j}{\delta_j^c + 1}[r_2^c(r_j^* x_0 + r_1^c x_0^c) - \mu] + \mu \\ &= r_2^c \tilde{r}_0^* x_0 + \tilde{r}_1^c x_0^c - \tilde{p}_0 \mu + \mu = \frac{\lambda_0}{r^c} + \mu = \frac{\tilde{\rho}}{r^c}(r^c - \mu) + \mu. \end{aligned}$$

This gives $\rho = \tilde{\rho} + \mu(r^c - \tilde{\rho})/r^c$ and thus $\mu = r^c(\rho - \tilde{\rho})/(r^c - \tilde{\rho})$. To calculate the risk, let $\bar{R}_j(x_j, x_j^c) := x_j^* \Sigma_j x_j + \rho_j(x_j, x_j^c)^2$ and use the last expression from (7) for $p_j \rho_j(x_j, x_j^c)$. Then

$$\begin{aligned} \bar{R}_j(x_j, x_j^c) &= \left(\frac{\lambda_j}{\tilde{r}_j^c}\right)^2 \delta_j^c + \left(\frac{\lambda_j}{\tilde{r}_j^c} + \mu\right)^2 = \left(\frac{\lambda_j}{\tilde{r}_j^c}\right)^2 (\delta_j^c + 1) + 2\frac{\lambda_j}{\tilde{r}_j^c}\mu + \mu^2 \\ &= \frac{[r_2^c(r_j^* x_0 + r_1^c x_0^c) - \mu]^2}{\delta_j^c + 1} + 2\frac{[r_2^c(r_j^* x_0 + r_1^c x_0^c) - \mu]}{\delta_j^c + 1}\mu + \mu^2 \\ &= \frac{[r_2^c(r_j^* x_0 + r_1^c x_0^c)]^2}{\delta_j^c + 1} + \mu^2 \frac{\delta_j^c}{\delta_j^c + 1}. \end{aligned}$$

Since $R(x) = \sum_{j \in L} p_j \bar{R}_j(x_j, x_j^c) - \rho^2$, we get

$$\begin{aligned} R(x) + \rho^2 &= (r_2^c)^2 \sum_{j \in L} \frac{p_j}{\delta_j^c + 1}\left[(r_j^* x_0)^2 + 2 r_1^c x_0^c r_j^* x_0 + (r_1^c x_0^c)^2\right] + \mu^2 \sum_{j \in L} \frac{p_j \delta_j^c}{\delta_j^c + 1} \\ &= (r_2^c)^2 \left[x_0^* \tilde{\Sigma}_0 x_0 + 2 r_1^c x_0^c \tilde{r}_0^* x_0 + \tilde{p}_0 (r_1^c x_0^c)^2\right] + \mu^2(1 - \tilde{p}_0). \end{aligned}$$

Using (8), the first term equals

$$r_2^c \tilde{r}_j^c \left[x_0^* \Sigma_0 x_0 + (\bar{r}_0^* x_0 + r_1^c x_0^c)^2\right] = r_2^c \tilde{r}_j^c \left[\left(\frac{\lambda_0}{r_2^c \tilde{r}_1^c}\right)^2 \delta_0^c + \left(\frac{\lambda_0}{r_2^c \tilde{r}_1^c} + \frac{\mu}{r_2^c}\right)^2\right],$$

which is further simplified to

$$\tilde{p}_0 \left[\left(\frac{\lambda_0}{\tilde{r}_1^c}\right)^2 (\delta_0^c + 1) + 2\frac{\lambda_0}{\tilde{r}_1^c}\mu + \mu^2\right] = \tilde{p}_0 \left[\frac{(r^c - \mu)^2}{\delta_0^c + 1} + 2\frac{r^c - \mu}{\delta_0^c + 1}\mu + \mu^2\right].$$

Therefore we have

$$\begin{aligned} R(x) + \rho^2 &= \tilde{p}_0 \frac{(r^c)^2 + \delta_0^c \mu^2}{\delta_0^c + 1} + \mu^2(1 - \tilde{p}_0) = r^c \tilde{\rho} + \mu^2 \left(1 - \frac{\tilde{p}_0}{\delta_0^c + 1}\right) \\ &= r^c \tilde{\rho} + \mu^2 \left(1 - \frac{\tilde{\rho}}{r^c}\right) = r^c \tilde{\rho} + r^c \frac{(\rho - \tilde{\rho})^2}{r^c - \tilde{\rho}}. \end{aligned}$$

Subtracting ρ^2 gives the stated risk formula whose minimum over all ρ is easily determined. The remaining statements follow trivially. $\square$

Proof of Theorem 18. The optimality conditions include the ones of Problem 11 except that ρ must be replaced by $\rho + \theta$ (in $\partial L/\partial\mu = 0$). Additional conditions are $\mu - (\rho + \theta) = \eta$ (from $\partial L/\partial\theta = 0$), $\theta \geq 0$, $\eta \geq 0$, and complementarity $\theta\eta = 0$. The proof now proceeds precisely as the proof of Theorem 17, with ρ replaced by $\rho + \theta$ in all instances. Finally nonnegativity and complementarity of θ and η lead to the case distinction: either $\theta = 0$ and $\mu \geq \rho$, or $\theta > 0$ and $\mu = \rho + \theta$. The formula $\mu = r^c(\rho + \theta - \tilde{\rho})/(r^c - \tilde{\rho})$ gives $\rho \geq r^c$ in the first case and $\rho + \theta = r^c$ in the second case. □

Theorem 20. *In Problem 13 let $\theta_j \equiv r_2^c x_j^l$ for $j \in L$, $\theta_0 \equiv r^c x_0^l$, and*

$$\bar{\theta}_0 := \frac{1}{\tilde{p}_0} \sum_{j \in L} \frac{p_j}{\delta_j^c + 1} \theta_j, \qquad s_0 := \frac{1}{\tilde{p}_0} \sum_{j \in L} \frac{p_j}{\delta_j^c + 1} \theta_j r_j.$$

Then every primal-dual solution satisfies

$$\begin{aligned}
x_j &= -\frac{\lambda_j}{\tilde{r}_j^c} \Sigma_j^{-1}(\bar{r}_j - r_2^c e), \qquad x_j^c = \frac{1}{r_2^c}\left[\frac{\lambda_j}{\tilde{r}_j^c}(\gamma_j - r_2^c \beta_j + 1) + \mu\right], \\
x_0 &= -\frac{\lambda_0}{r_2^c \tilde{r}_1^c} \Sigma_0^{-1}(\bar{r}_0 - r_1^c e) - \frac{1}{r_2^c} \Sigma_0^{-1}(\bar{\theta}_0 \bar{r}_0 - s_0), \\
x_0^c &= \frac{1}{r^c}\left[\frac{\lambda_0}{\tilde{r}_1^c}(\gamma_0 - r_1^c \beta_0 + 1) + \mu + (\gamma_0 + 1)\bar{\theta}_0 - \bar{r}_0^* \Sigma_0^{-1} s_0\right], \\
\lambda_j &= \frac{\tilde{r}_j^c}{\delta_j^c + 1}[r_2^c(r_j^* x_0 + r_1^c x_0^c) - \mu - \theta_j] = -\eta_j \leq 0, \\
\lambda_0 &= \frac{\tilde{r}_1^c}{\delta_0^c + 1}[r^c - \mu - \theta_0 - (\gamma_0 - r_1^c \beta_0 + 1)\bar{\theta}_0 + (\bar{r}_0 - r_1^c e)^* \Sigma_0^{-1} s_0] = -\eta_0 \leq 0, \\
\mu &= r^c \frac{\rho - \tilde{\rho}}{r^c - \tilde{\rho}} + \frac{\tilde{\rho}}{r^c - \tilde{\rho}}[(\gamma_0 - r_1^c \beta_0 + 1)\bar{\theta}_0 + \theta_0 - (\bar{r}_0 - r_1^c e)^* \Sigma_0^{-1} s_0].
\end{aligned}$$

In particular, the following case distinction can be made in each scenario. If $\eta_j > 0$, then $x_j^l = 0$, $\rho_j(x_j, x_j^c) < \mu$, and $r_2^c w_j < \mu$. Conversely, if $\eta_j = 0$, then $x_j^l \geq 0$, $\rho_j(x_j, x_j^c) = \mu \geq \rho$, and $r_2^c w_j \geq \mu$. In this case the leaf variables are

$$x_j = 0, \qquad x_j^c = \mu/r_2^c, \qquad x_j^l = 1 - \mu/r_2^c,$$

giving $x_j^l = 0$ (100 % cash) if $r_2^c w_j = \mu$ and $x_j^l > 0$ else.

The following case distinction holds in the root. If $\eta_0 > 0$, then $x_0^l = 0$ and $\rho < \mu$. Otherwise, if $\eta_0 = 0$, then $x_0^l \geq 0$ and $\rho = \mu$.

Proof. The optimality conditions of Theorem 17 remain valid except that x_j^l now appears in all the budget conditions $\partial L/\partial\lambda_j = 0$, $j \in V$. Additional optimality conditions in each node are $\lambda_j = -\eta_j$ (from $\partial L/\partial x_j^l = 0$), $x_j^l \geq 0$, $\eta_j \geq 0$, and complementarity $x_j^l \eta_j = 0$. The expressions above are obtained precisely as in the proof of Theorem 17 when slack variables x_j^l and derived quantities $\theta_j, \bar{\theta}_0, s_0$ are included. This derivation also yields intermediate results

$$\rho_j(x_j, x_j^c) = \mu + \frac{\lambda_j}{\tilde{r}_j^c} \leq \mu, \quad j \in L, \qquad \rho = \mu + \frac{\lambda_0}{r^c} \leq \mu,$$

cf. (7) and the reward condition after (8). The additional optimality conditions above together with these identities lead to the stated case distinctions when it is observed that λ_j has the same sign as $r_2^c w_j - \mu - \theta_j$ for $j \in L$. □

Remark. Note that all the multipliers now have a natural interpretation. The reward multiplier μ is the maximal scenario return and a *threshold value for surplus money:* there is surplus money in scenario j iff $\rho_j = \mu$ and $r_2^c w_j > \mu$. The budget multipliers λ_j (up to a scaling factor) measure the difference of desired return or scenario returns to the threshold.

Proof of Lemma 22. Formula (4) for ρ_j (with $\theta_j = 0$) gives

$$(\phi_j + 1)(\rho_j - d_j) = \phi_j d_j + \mu - (\phi_j + 1)d_j = \mu - d_j$$

and hence $\rho_j - d_j \sim \mu - d_j \equiv \mu - r^c - \psi_j x_0$. By Theorem 17,

$$x_0 = -\frac{\lambda_0}{r_2^c \tilde{r}_1^c} \frac{\bar{r}_0 - r_1^c}{\Sigma_0} = -\tilde{\rho} \frac{r^c - \mu}{r_2^c \tilde{r}_1^c} \frac{\bar{r}_0 - r_1^c}{\Sigma_0}.$$

Therefore we have

$$\begin{aligned}
\rho_j - d_j &\sim \mu - r^c + r_2^c(r_j - r_1^c)\tilde{\rho}\frac{r^c - \mu}{r_2^c \tilde{r}_1^c} \frac{\bar{r}_0 - r_1^c}{\Sigma_0} \\
&= (\mu - r^c)\left(1 - \frac{r_j - r_1^c}{\delta_0^c + 1} \frac{\bar{r}_0 - r_1^c}{\Sigma_0}\right).
\end{aligned}$$

By Theorem 17, $\mu - r^c = r^c(\rho - r^c)/(r^c - \tilde{\rho}) \sim \rho - r^c$. The second factor equals

$$\begin{aligned}
1 - \frac{(r_j - r_1^c)(\bar{r}_0 - r_1^c)}{[(\bar{r}_0 - r_1^c)\Sigma_0^{-1}(\bar{r}_0 - r_1^c) + 1]\Sigma_0} &\sim (\bar{r}_0 - r_1^c)^2 + \Sigma_0 - (r_j - r_1^c)(\bar{r}_0 - r_1^c) \\
&= \Sigma_0 + \bar{r}_0^2 - (r_j + r_1^c)\bar{r}_0 + r_j r_1^c \\
&\sim \tilde{\Sigma}_0 - (r_j + r_1^c)\tilde{r}_0 + \tilde{p}_0 r_j r_1^c \\
&= \sum_{k \in L} \frac{p_k}{\delta_k^c + 1}[r_k^2 - (r_j + r_1^c)r_k + r_1^c r_j] \\
&= \sum_{k \neq j} \frac{p_k}{\delta_k^c + 1}(r_k - r_j)(r_k - r_1^c) = s_j.
\end{aligned}$$

This completes the proof. □

Acknowledgement

This work would have been impossible without many intensive discussions in the Operations Research group at the University of St. Gallen. My sincerest thanks especially to K. Frauendorfer, H. Siede, and D. Steiner.

References

[1] A. J. Alexander and W. F. Sharpe, *Fundamentals of investment*, Prentice-Hall, 1989.

[2] K. J. Arrow, *Essays in the theory of risk-bearing*, Markham, Chicago, 1971.

[3] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath, *Thinking coherently*, Risk, 10 (1997), pp. 68–71.

[4] C. Atkinson, S. R. Pliska, and P. Wilmott, *Portfolio management with transaction costs.*, Proc. R. Soc. Lond., Ser. A, 453 (1997), pp. 551–562.

[5] V. S. Bawa, *Stochastic dominance: A research bibliography*, Manag. Sci., 28 (1982), pp. 698–712.

[6] R. E. Bellman, *Dynamic Programming*, Princeton University Press, 1957.

[7] A. Beltratti, A. Consiglio, and S. A. Zenios, *Scenario modeling for the management of international bond portfolios.*, Ann. Oper. Res., 85 (1999), pp. 227–247.

[8] M. Best and B. Ding, *On the continuity of the minimum in parametric quadratic programs.*, J. Optimization Theory Appl., 86 (1995), pp. 245–250.

[9] J. R. Birge, *Stochastic programming computations and applications*, INFORMS J. Comput., 9 (1997), pp. 111–133.

[10] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer-Verlag, 1997.

[11] M. Broadie, *Computing efficient frontiers using estimated parameters*, Ann. Oper. Res., 45 (1993), pp. 21–58.

[12] V. K. Chopra and W. T. Ziemba, *The effects of errors in means, variances, and covariances on optimal portfolio choice*, J. Portfolio Manag., (1993), pp. 6–11.

[13] G. Consigli and M. A. H. Dempster, *Solving dynamic portfolio problems using stochastic programming*, Z. Angew. Math. Mech., 77 (1997), pp. S535–S536.

[14] G. B. Dantzig and G. Infanger, *Multi-stage stochastic linear programs for portfolio optimization*, Ann. Oper. Res., 45 (1993), pp. 59–76.

[15] R. S. Dembo and D. Rosen, *The practice of portfolio replication. A practical overview of forward and inverse problems.*, Ann. Oper. Res., 85 (1999), pp. 267–284.

[16] T. Dohi and S. Osaki, *A note on portfolio optimization with path-dependent utility*, Ann. Oper. Res., 45 (1993), pp. 77–90.

[17] G. T. Duncan, *A matrix measure of multivariate local risk aversion*, Econometrica, 45 (1977), pp. 895–903.

[18] E. J. Elton and M. J. Gruber, *Modern portfolio theory and investment analysis*, John Wiley, New York, 3rd ed., 1987.

[19] Y. Ermoliev and R. J.-B. Wets, eds., *Numerical techniques for stochastic optimization*, Springer-Verlag, 1988.

[20] P. C. Fishburn, *Mean-risk analysis with risk associated with below-target returns*, Amer. Econ. Rev., 67 (1977), pp. 116–125.

[21] K. Frauendorfer, *The stochastic programming extension of the Markowitz approach*, Int. J. Mass-Parallel Comput. Inform. Syst., 5 (1995), pp. 449–460.

[22] ———, *Barycentric scenario trees in convex multistage stochastic programming*, Math. Programming, 75 (1996), pp. 277–293.

[23] K. Frauendorfer and H. Siede, *Portfolio selection using multi-stage stochastic programming*, tech. report, University of St. Gallen, 1998.

[24] N. H. Hakansson, *Multi-period mean-variance analysis: Toward a general theory of portfolio choice*, J. Fin., 26 (1971), pp. 857–884.

[25] J. M. Harrison and S. R. Pliska, *A stochastic calculus model of continuous trading: Complete markets*, Stochastic Processes Appl., 15 (1983), pp. 313–316.

[26] H. He and N. D. Pearson, *Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case*, J. Econ. Theory, 54 (1991), pp. 259–304.

[27] D. C. Heath, S. Orey, V. C. Pestien, and W. D. Sudderth, *Minimizing or maximizing the expected time to reach zero*, SIAM J. Control Optimization, 25 (1987), pp. 195–205.

[28] G. Infanger, 1994. Private Communication (via K. Frauendorfer, 1998).

[29] J. E. Ingersoll, Jr., *Theory of financial decision making*, Rowman & Littlefield, 1987.

[30] S. D. Jacka, *A martingale representation result and an application to incomplete financial markets*, Math. Finance, 2 (1992), pp. 239–250.

[31] I. Jewitt, *Choosing between risky prospects: The characterization of comparative statistics results, and location independent risk*, Manag. Sci., 35 (1989), pp. 60–70.

[32] J. D. Jobson, *Confidence regions for the mean-variance efficient set: An alternative approach to estimation risk*, Rev. Quant. Fin. Accounting, 1 (1991), p. 293.

[33] P. Kall and S. W. Wallace, *Stochastic Programming*, John Wiley, New York, 1994.

[34] I. Karatzas, *Optimization problems in the theory of continuous trading*, SIAM J. Control Optimization, 27 (1989), pp. 1221–1258.

[35] I. Karatzas, J. P. Lehoczky, S. E. Shreve, and G. L. Xu, *Martingale and duality methods for utility maximization in an incomplete market*, SIAM J. Control Optimization, 29 (1991), pp. 702–730.

[36] R. E. Kihlstrom and L. J. Mirman, *Constant, increasing, and decreasing risk aversion with many commodities*, Rev. Econ. Studies, 48 (1981), pp. 271–280.

[37] M. Kijima and M. Ohnishi, *Mean-risk analysis of risk aversion and wealth effects on optimal portfolios with multiple investment opportunities*, Ann. Oper. Res., 45 (1993), pp. 147–163.

[38] ———, *Portfolio selection problems via the bivariate characterization of stochastic dominance relations.*, Math. Finance, 6 (1996), pp. 237–277.

[39] A. J. King, *Asymmetric risk measures and tracking models for portfolio optimization under uncertainty*, Ann. Oper. Res., 45 (1993), pp. 165–177.

[40] A. J. King and D. L. Jensen, *Linear-quadratic efficient frontiers for portfolio optimization.*, Appl. Stochastic Models Data Anal., 8 (1992), pp. 195–207.

[41] H. Konno, *Piecewise linear risk function and portfolio optimization*, J. Oper. Res. Soc. Japan, 33 (1990), pp. 139–156.

[42] H. Konno, S. R. Pliska, and K.-I. Suzuki, *Optimal portfolios with asymptotic criteria*, Ann. Oper. Res., 45 (1993), pp. 187–204.

[43] H. Konno and K.-I. Suzuki, *A mean-variance-skewness portfolio optimization model.*, J. Oper. Res. Soc. Japan, 38 (1995), pp. 173–187.

[44] H. Konno and H. Watanabe, *Bond portfolio optimization problems and their applications to index tracking: A partial optimization approach.*, J. Oper. Res. Soc. Japan, 39 (1996), pp. 295–306.

[45] A. Kraus and R. Litzenberger, *Skewness preference and the valuation of risky assets*, J. Fin., 21 (1976), pp. 1085–1094.

[46] Y. Kroll, H. Levy, and H. M. Markowitz, *Mean-variance versus direct utility maximization*, J. Fin., 39 (1984), pp. 47–62.

[47] H. Levy, *Stochastic dominance and expected utility: survey and analysis*, Manag. Sci., 38 (1992), pp. 555–593.

[48] H. Levy and H. M. Markowitz, *Approximating expected utility by a function of mean and variance*, Amer. Econ. Rev., 69 (1979), pp. 308–317.

[49] Y. Li and W. T. Ziemba, *Univariate and multivariate measures of risk aversion and risk premiums*, Ann. Oper. Res., 45 (1993), pp. 265–296.

[50] L. C. MacLean and K. Weldon, *Estimating multivariate random effects without replication.*, Comm. Statist. Theory Methods, 25 (1996), pp. 1447–1469.

[51] H. M. Markowitz, *Portfolio selection*, J. Fin., 7 (1952), pp. 77–91.

[52] ———, *The utility of wealth*, J. Political Econ., 60 (1952), pp. 151–158.

[53] ———, *The optimization of a quadratic function subject to linear constraints*, Nav. Res. Logist. Quarterly, 3 (1956), pp. 111–133.

[54] ———, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley, New York, 1959.

[55] ———, *Mean-Variance analysis in portfolio choice and capital markets*, Basil Blackwell, 1987.

[56] R. C. Merton, *Optimal consumption and portfolio rules in a continuous-time model*, J. Econ. Theory, 3 (1971), pp. 373–413.

[57] R. C. Merton and P. A. Samuelson, *Fallacy of the log-normal approximation to optimal portfolio decision-making over many periods*, J. Fin. Econ., 1 (1974), pp. 67–94.

[58] A. J. Morton and S. R. Pliska, *Optimal portfolio management with fixed transaction costs.*, Math. Finance, 5 (1995), pp. 337–356.

[59] J. Mossin, *Optimal multiperiod portfolio policies*, J. Bus., 41 (1968), pp. 215–229.

[60] J. M. Mulvey and H. Vladimirou, *Stochastic network optimization models for investment planning*, Ann. Oper. Res., 20 (1989), pp. 187–217.

[61] M. Nakasato and K. Furukawa, *On the number of securities which constitute an efficient portfolio*, Ann. Oper. Res., 45 (1993), pp. 333–347.

[62] *The Sveriges Riksbank (Bank of Sweden) Prize in Economic Sciences in Memory of Alfred Nobel.* Press Release of The Royal Swedish Academy of Sciences, Oct. 16, 1990. URL `http://www.nobel.se/laureates/economy-1990-press.html`.

[63] A. F. Perold, *Large-scale portfolio optimization*, Manag. Sci., 30 (1984), pp. 1143–1160.

[64] J. W. Pratt, *Risk aversion in the small and in the large*, Econometrica, 32 (1964), pp. 122–136.

[65] R. T. Rockafellar, *Duality and optimality in multistage stochastic programming*, Ann. Oper. Res., 85 (1999), pp. 1–19.

[66] R. T. Rockafellar and R. J.-B. Wets, *Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete time*, SIAM J. Control Optimization, 28 (1990), pp. 810–822.

[67] S. Ross, *Some stronger measures of risk aversion in the small and in the large*, Econometrica, 49 (1981), pp. 621–638.

[68] M. E. Rubinstein, *A comparative statistics analysis of risk premiums*, J. Bus., 12 (1973), pp. 605–615.

[69] A. Ruszczyński, *Decomposition methods in stochastic programming*, Math. Programming, 79 (1997), pp. 333–353.

[70] P. A. Samuelson, *The fundamental approximation theorem of portfolio analysis in terms of means, variances, and higher moments*, Rev. Econ. Studies, 25 (1958), pp. 65–86.

[71] ———, *Lifetime portfolio selection by dynamic stochastic programming*, Rev. Econ. Studies, 51 (1969), pp. 239–246.

[72] W. F. Sharpe, *Investment*, Prentice-Hall, 1978.

[73] H. Shirakawa, *Optimal consumption and portfolio selection with incomplete markets and upper and lower bound constraints.*, Math. Finance, 4 (1994), pp. 1–24.

[74] M. C. Steinbach, *Fast Recursive SQP Methods for Large-Scale Optimal Control Problems*, Ph. D. dissertation, University of Heidelberg, 1995.

[75] ———, *Structured interior point SQP methods in optimal control*, Z. Angew. Math. Mech., 76 (1996), pp. 59–62.

[76] ———, *Recursive direct optimization and successive refinement in multistage stochastic programs*, Preprint SC-98-27, ZIB, 1998.

[77] ———, *Recursive direct algorithms for multistage stochastic programs in financial engineering*, in Operations Research Proceedings 1998, P. Kall and H.-J. Lüthi, eds., Springer-Verlag, 1999, pp. 241–250.

[78] M. C. Steinbach, H. G. Bock, G. V. Kostin, and R. W. Longman, *Mathematical optimization in robotics: Towards automated high speed motion planning*, Surv. Math. Ind., 7 (1998), pp. 303–340.

[79] S. Uryasev and R. T. Rockafellar, *Optimization of Conditional Value-at-Risk*, Research Report 99-4, University of Florida, 1999.

[80] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 3rd ed., 1953.

[81] R. J.-B. Wets, *Programming under uncertainty: The equivalent convex program*, SIAM J. Appl. Math., 14 (1984), pp. 89–105.

[82] S. A. Zenios, *Financial optimization*, Cambridge University Press, Cambridge, UK, 1992.

[83] S. A. Zenios and P. Kang, *Mean-absolute deviation portfolio optimization for mortgage-backed securities*, Ann. Oper. Res., 45 (1993), pp. 433–450.

[84] W. T. Ziemba, *Choosing investments when the returns have stable distributions*, in Mathematical programming in theory and practice, P. L. Hammer and G. Zoutendijk, eds., North-Holland, Amsterdam, 1974, pp. 443–482.

[85] W. T. Ziemba and J. M. Mulvey, eds., *Worldwide asset and liability modeling*, Cambridge University Press, Cambridge, UK, 1998.