

Mean-Variance Analysis in a Multiperiod Setting*

Karl Frauendorfer, Heiko Siede

July 15, 1997

Abstract

Similar to the classical Markowitz approach it is possible to apply a mean-variance criterion to a multiperiod setting to obtain efficient portfolios. To represent the stochastic dynamic characteristics necessary for modelling returns a process of asset returns is discretized with respect to time and space and summarized in a scenario tree. The resulting optimization problem is solved by means of stochastic multistage programming. The optimal solutions show equivalent structural properties as the classical approach, however, by taking rebalancing activities into consideration a different efficient frontier is obtained.

1 Introduction

One widely accepted and applied measure for risk in portfolio optimization is the variance of a chosen portfolio. Therefore, one has to estimate the variance-covariance-matrix and expected returns of the possible assets for a predefined time horizon. Starting with wealth one chooses the portfolio with the smallest variance for an expected return of the whole portfolio. By varying the expected return a set of minimum variances and corresponding portfolios are obtained. This set is called efficient frontier. In particular, the optimal portfolio is given by a vector valued function affine linear in the expected return, and the efficient frontier is convex quadratic. This technique was firstly introduced by Markowitz [?, ?] and provides the basis for the capital asset pricing model.

The goodness of the results obtained by this classical one period optimization model is significantly depending on the estimated input data, i.e. the variance-covariance-matrix and the expected returns of every single asset considered. However, it is observed, that these data vary during the chosen time horizon, and therefore, one has to restate the efficient frontier and rebalance the portfolio.

By modelling the price of an asset as a diffusion process one takes changes of the input data into consideration. However, this is a continuous time model and thus, to apply the classical Markowitz approach, the model has to be discretized to a one stage asset price model with matching returns. Of course, this discretization can not account for changing input data. To circumvent this

*Research of this report was supported by *Grundlagenforschungsfonds, University of St. Gallen*, Nr.: S-510

restriction a scenario tree can be modelled, which adequately represents the stochastic evolution of the underlying continuous time process. The number of stages is chosen corresponding to the number of times rebalancing is allowed. Obviously, on the one hand, the more stages considered, the better is the representation of the process. On the other hand, the larger the tree the higher the necessary computational effort of the optimization algorithm. Furthermore, the transaction costs of rebalancing, which are not modelled explicitly here, might be prohibitive high.

A scenario tree is a prerequisite for applying multistage stochastic optimization, and, therefore, is an ideal tool to develop a multiperiod mean-variance approach. Frauendorfer [?] formulated a problem to optimize a portfolio with the considered type of tree structure. The objective is to minimize the sum of expected variances over all considered time periods. Subsequently, he unveiled analytical characteristics of the efficient frontier, which are analogous to the classical Markowitz. Since the scenario tree may consist of different variance-covariance-matrices a varying volatility may directly influence the optimal portfolio and the necessary rebalancing activity.

In section 2 a multiperiod mean-variance approach is stated. The objective function remains as in Frauendorfer, though, the restrictions are slightly different. However, these changes leave the analytical characteristics unchanged. Since the proof yields an algorithm for a numerical implementation, the algorithm to solve this kind of problem is elaborated in detail. In section 3 a scenario tree is constructed and computational results are given for the one path case. Finally, section 4 summarizes the results and states further remarks.

2 The Extension of the Markowitz Approach

The major intention of the multistage stochastic programming methodology is to optimize decision u_t at any stage $t = 0, 1, \dots, T$, taking into consideration the information of the past up to stage $t - 1$, the potential scenarios of the future, and anticipating to implement the optimal corrections in future stages. Of particular importance for practitioners is to select the optimal decision u_0 , which has to be done at the beginning of the planning horizon, taking into account the stochastic evolution of the data and the associated optimal corrections.

In the context of portfolio optimization one desires to minimize the expected variance of the portfolio $u \in \mathbb{R}^n$ at every stage for a predefined expected return μ . Formally, for a T -stage formulation this corresponds to

$$\begin{aligned} \min_u \quad & \int_{\Omega} \sum_{t=0}^T \rho_t(Q_t, u_t) dP \\ \text{s.t.} \quad & f_t(u_{t-1}, u_t, r_{t-1}) = 0 \quad \forall t = 0, \dots, T \\ & g_T(u_T, r_T) = 0 \\ & u \text{ nonanticipative} \end{aligned}$$

with the discrete n -dim. processes r and $n \times n$ -dim. matrix process Q, Q_t positive definite, on a given probability space. The exact formulations and

interpretations of the functions ρ_t, f_t and g_t are discussed in the next subsection.

To solve the problem by applying stochastic optimization one has to integrate and minimize the implicitly given optimal value function. Unlike to many control type formulations, no analytical expressions can be expected within the stochastic multistage setting due to the subdifferentiability of the value functions. One way to overcome these difficulties numerically is to discretize the stochastic processes with respect to their outcomes. This is performed in the next subsection.

2.1 The Discrete Case

Let ω be a nonanticipative stochastic process on a probability space $(\Omega, \mathcal{A}, P)$ with time points $0, \dots, T+1$ and finite set of realisations. The setting $T+1$ for the final stage is due to modelling reasons. The initial risk factor ω_0 is a constant. The process ω represents the risk driving factors, which determine the processes of the $n \times n$ variance-covariance matrix Q and the n -dim. vector r of returns. By this way $Q_t = Q_t(\omega^t)$ and $r_t = r_t(\omega^{t+1})$ are nonanticipative with the notation $\omega^t = (\omega_0, \dots, \omega_t)$. Notice, Q_t is the variance-covariance-matrix of the returns between period t and $t+1$. The realised return r_t during this time can be observed in period $t+1$. For $t = 1, \dots, T+1$ let $\mathcal{A}^t$ be the projection of $\mathcal{A}$ on the horizon $[0, t]$ with associated probabilities $p_t(\omega^t) \neq 0$ of the realisations of ω^t . Of course, the projection $\mathcal{A}^{T+1}$ is $\mathcal{A}$. Furthermore, denote $\mathcal{A}_t(\omega^{t-1})$ as the set of finite many outcomes ω_t conditioned on ω^{t-1} .

This setting allows the interpretation of a scenario tree with variance-covariance-matrix and a vector of returns at every node. The associated probability at the node is thus

$$p_t(\omega^t) = \prod_{s=1}^t p_s(\omega_s | \omega^{s-1})$$

Because of the constant start ω_0 the probability p_0 at the root of the tree is one.

The intention is to minimize risk for a predefined expected return μ achieved at the end of the planning horizon. In this setting the risk is defined as the aggregation of expected variances of the portfolio on all considered periods. Formally this is

$$\sigma_0^2(\mu) = \min_{u_0} \left\langle u_0, \frac{1}{2} Q_0 u_0 \right\rangle + \sum_{t=1}^T \sum_{\omega^t \in \mathcal{A}^t} \left\langle u_t(\omega^t), \frac{1}{2} Q_t(\omega^t) u_t(\omega^t) \right\rangle p_t(\omega^t), \quad (1)$$

with the n -dim. portfolio vector u_t depending on time and the risk factor ω^t . Of course, $\sigma_0^2(\mu)$ is not the variance of the portfolio over the entire planning horizon, since the process ω and thus Q consists of depending states. Rather more, it is a risk measure containing the variances of all states weighted by their probability.

The aim is to find an optimal policy under the restrictions

$$\langle \mathbf{1}, u_0 \rangle = 1 \quad (??a)$$

$$- \langle r_{t-1}(\omega^t), u_{t-1}(\omega^{t-1}) \rangle + \langle \mathbf{1}, u_t(\omega^t) \rangle = 0 \quad \forall \omega^t, \forall t \quad (??b)$$

$$\sum_{\omega^T \in \Omega^T} \langle E(r_T(\omega^{T+1})|\omega^T), u_T(\omega^T) \rangle p_T(\omega^T) = \mu, \quad (??c)$$

which describe the dynamic evolvement of wealth. By $\langle \mathbf{1}, u_0 \rangle = 1$ with $\mathbf{1} = (1 \dots 1)$, wealth is set to one at the beginning of the planning horizon. Equation (??b) describes the movement of wealth in time for all realisations. Taking a realisation ω then the wealth in period t is $\langle \mathbf{1}, u_t(\omega^t) \rangle$ and must equal the wealth of period $t - 1$ and the return between both periods. Since $r_t(\omega^{t-1})$ is the process for $(1+\text{return})$ equality is accomplished. Furthermore, the expected return, respectively wealth, over the whole planning horizon is set to μ and is achieved in period $T + 1$. Formally,

$$\begin{aligned} \mu &= \sum_{\omega^{T+1} \in \Omega^{T+1}} \langle r_T(\omega^{T+1}), u_T(\omega^T) \rangle p(\omega^{T+1}) \\ &= \sum_{\omega^T \in \Omega^T} \langle \sum_{\omega^{T+1} \in \mathcal{A}(\omega^T)} r_T(\omega_{T+1}, \omega^T) p(\omega_{T+1}|\omega^T), u_T(\omega^T) \rangle p(\omega^T) \end{aligned}$$

and this is equivalent to equation (??c).

Theorem 1. *Suppose the variance-covariance matrices $Q_t(\omega^t)$ are positive definite for all $\omega^t \in \mathcal{A}^t$ ($t = 0, \dots, T$) and $E(r_T(\omega^{T+1})|\omega^T)$ is linearly independent of $\mathbf{1}$ for all $\omega^T \in \mathcal{A}^T$, then the optimal portfolio is given by a vector valued function affine linear in μ and the efficient frontier is convex quadratic.*

Outline of the proof: Due to the inf-compactness of the convex quadratic objective function the dynamic representation of the convex multistage stochastic program (??), (??) may be applied (confer Rockafellar and Wets [?]). This amounts to considering the expected efficient frontier at the final stage T

$$\begin{aligned} \sigma_T^2(u_{T-1}, \mu) &= \min \sum_{\omega^T} \langle u_T(\omega^T), \frac{1}{2} Q_T(\omega^T) u_T(\omega^T) \rangle p_T(\omega^T) \\ \text{s.t.} \quad &\langle r_{T-1}(\omega^T), u_{T-1}(\omega^{T-1}) \rangle - \langle e, u_T(\omega^T) \rangle = 0 \quad \forall \omega^T \in \mathcal{A}^T \\ &\mu - \sum_{\omega^T} \langle E(r_T(\omega^{T+1})|\omega^T), u_T(\omega^T) \rangle p_T(\omega^T) = 0 \end{aligned} \quad (3)$$

Clearly, the expected efficient frontier at T depends on the decision of the previous stage u_{T-1} and on the prescribed portfolio return μ .

Defining backwards for $t = T - 1, \dots, 1$, one obtains the expected efficient frontier at t

$$\begin{aligned} \sigma_t^2(u_{t-1}, \mu) &= \min \sum_{\omega^t} \langle u_t(\omega^t), \frac{1}{2} Q_t(\omega^t) u_t(\omega^t) \rangle p_T(\omega^T) + \sigma_{t+1}^2(u_t, \mu) \\ \text{s.t.} \quad &\langle e, u_t(\omega^t) \rangle = \langle r_{t-1}(\omega^t), u_{t-1}(\omega^{t-1}) \rangle \quad \forall \omega^t \in \mathcal{A}^t, \end{aligned} \quad (4)$$

where $\sigma_t^2(u_{t-1}, \mu)$ is determined by minimizing both the expected risk one is exposed to at current stage t and the expected minimum risk one is exposed to beyond t .

Applying the classical Lagrangian methodology, it immediately becomes apparent, that the expected efficient frontiers $\sigma_t^2(u_{t-1}, \mu)$ at stages $t = 1, \dots, T$ are convex quadratic value functions. The set of efficient portfolios turn out to be linearly affine in (u_{t-1}, μ) . As a consequence, at $t = 0$, a simple convex quadratic single-period program of the form

$$\begin{aligned} \sigma_0^2(\mu) &= \min < u_0, Q_0 u_0 > + \sigma_1^2(u_0, \mu) \\ \text{s.t.} \quad &< e, u_0 > = 1 \end{aligned} \quad (5)$$

remains to be solved. The current efficient frontier $\sigma_0^2(\mu)$, which is convex quadratic, is of particular interest. The program, aggregating backwards the variance-covariances and returns, yields the *aggregation principle of stochastic multistage mean-variance analysis*. The outlined algorithm is elaborated in detail in the remaining section.

2.2 The Expected Efficient Frontier at T

The Lagrangian function to problem (??) is

$$\begin{aligned} L_T(u_T, \lambda_T, \gamma_T) &= \sum_{\omega^T} < u_T(\omega^T), \frac{1}{2} Q_T(\omega^T) u_T(\omega^T) > p_T(\omega^T) \\ &\quad + \lambda_T(\omega^T) [< r_{T-1}(\omega^T), u_{T-1} > - < e, u_T(\omega^T) >] \\ &\quad + \gamma_T [\mu - \sum_{\omega^T} p_T(\omega^T) < E(r_T(\omega^{T+1}) | \omega^T), u_T(\omega^T) >], \end{aligned} \quad (6)$$

which yields after differentiation with respect to $u_T(\omega^T)$,

$$\frac{\partial L_T(u_T, \lambda_T, \gamma_T)}{\partial u_T(\omega^T)} = p_T(\omega^T) Q_T(\omega^T) u_T(\omega^T) - \lambda_T(\omega^T) e - \gamma_T p_T(\omega^T) r_T(\omega^T),$$

the optimal solution $u_T^* := (\dots, u_T^*(\omega^T), \dots)$ with

$$u_T^*(\omega^T) = \frac{\lambda_T(\omega^T)}{p_T(\omega^T)} Q_T^{-1}(\omega^T) e + \gamma_T Q_T^{-1}(\omega^T) E(r_T(\omega^{T+1}) | \omega^T) \quad \forall \omega^T. \quad (7)$$

Setting

$$\begin{aligned} A_T(\omega^T) &= < e, Q_T^{-1}(\omega^T) e > \\ B_T(\omega^T) &= < e, Q_T^{-1}(\omega^T) E(r_T(\omega^{T+1}) | \omega^T) > \\ C_T(\omega^T) &= < E(r_T(\omega^{T+1}) | \omega^T), Q_T^{-1}(\omega^T) E(r_T(\omega^{T+1}) | \omega^T) > \\ \alpha_T &= \sum_{\omega^T} p_T(\omega^T) C_T(\omega^T) \end{aligned}$$

and using (??) and the first restriction of problem (??), one may write for all ω^T

$$\begin{aligned} \langle e, u_T^*(\omega^T) \rangle &= \lambda_T(\omega^T) \frac{A_T(\omega^T)}{p_T(\omega^T)} + \gamma_T B_T(\omega^T) \\ &= \langle r_{T-1}(\omega^T), u_{T-1}(\omega^{T-1}) \rangle \end{aligned} \quad (8)$$

and because of the second restriction one gets

$$\begin{aligned} \sum_{\omega^T} p_T(\omega^T) \langle r_{T-1}(\omega^T), u_{T-1}(\omega^{T-1}) \rangle &= \sum_{\omega^T} B_T(\omega^T) \lambda_T(\omega^T) + \gamma_T \sum_{\omega^T} p_T(\omega^T) C_T(\omega^T) \\ &= \sum_{\omega^T} B_T(\omega^T) \lambda_T(\omega^T) + \alpha_T \gamma_T = \mu. \end{aligned} \quad (9)$$

The linear system (??)-(??) may be rewritten in matrix form

$$\begin{pmatrix} \ddots & 0 & 0 & \vdots \\ 0 & \frac{A_T(\omega^T)}{p_T(\omega^T)} & 0 & B_T(\omega^T) \\ 0 & 0 & \ddots & \vdots \\ \dots & B_T(\omega^T) & \dots & \alpha_T \end{pmatrix} \begin{pmatrix} \vdots \\ \lambda_T(\omega^T) \\ \vdots \\ \gamma_T \end{pmatrix} = \begin{pmatrix} \vdots \\ \langle r_{T-1}(\omega^T), u_{T-1}(\omega^{T-1}) \rangle \\ \vdots \\ \mu \end{pmatrix} \quad (10)$$

The matrix of the linear system is symmetric and regular due to the Cauchy-Schwartz inequality, which implies $B_T^2(\omega^T) < A_T(\omega^T)C_T(\omega^T)$ for exactly those ω^T for which $E(r_T(\omega^{T+1})|\omega^T)$ and $\mathbf{1}$ are linearly independent. Setting

$$\beta_T := \left(\sum_{\omega^T} p_T(\omega^T) [C_T(\omega^T) - \frac{B_T^2(\omega^T)}{A_T(\omega^T)}] \right)^{-1}$$

the multipliers are given by

$$\begin{aligned} \lambda_T^*(\omega^T) &:= \sum_{\hat{\omega}^T \neq \omega^T} \frac{p_T(\omega^T) B_T(\omega^T) p_T(\hat{\omega}^T) B_T(\hat{\omega}^T)}{A_T(\omega^T) A_T(\hat{\omega}^T)} \beta_T \langle r_{T-1}(\hat{\omega}^T), u_{T-1}(\hat{\omega}^{T-1}) \rangle \\ &+ \left(\frac{p_T(\omega^T)}{A_T(\omega^T)} + \frac{p_T^2(\omega^T) B_T^2(\omega^T) \beta_T}{A_T^2(\omega^T)} \right) \langle r_{T-1}(\omega^T), u_{T-1}(\omega^{T-1}) \rangle \\ &- \frac{p_T(\omega^T) B_T(\omega^T)}{A_T(\omega^T)} \beta_T \mu \end{aligned}$$

and

$$\gamma_T^* := - \sum_{\omega^T} \frac{p_T(\omega^T) B_T(\omega^T)}{A_T(\omega^T)} \beta_T \langle r_{T-1}(\omega^T), u_{T-1}(\omega^{T-1}) \rangle + \beta_T \mu.$$

Obviously, $\lambda_T^*(\omega^T)$, γ_T^* and, hence, the optimal solution $u_T^* := (\dots, u_T^*(\omega^T), \dots)$ are linear in (u_{T-1}, μ) . The expected efficient frontier at T depends on u_{T-1} and may be written after simple transformations

$$\begin{aligned} \sigma_T^2(u_{T-1}, \mu) = & \frac{1}{2} \beta_T \left[\sum_{\omega^T} \frac{B_T(\omega^T) p_T(\omega^T)}{A_T(\omega^T)} < r_{T-1}(\omega^{T-1}), u_{T-1}(\omega^{T-1}) > - \mu \right]^2 \\ & + \frac{1}{2} \sum_{\omega^T} \frac{p_T(\omega^T)}{A_T(\omega^T)} < r_{T-1}(\omega^{T-1}), u_{T-1}(\omega^{T-1}) >^2. \end{aligned} \quad (11)$$

Observe that $\beta_T > 0$ guarantees the expected efficient frontier to be quadratic convex in both arguments (u_{T-1}, μ) .

Equation (??) can be rewritten in a more convenient form. Therefore, suppose $|\mathcal{A}(\omega_i^{t-1})| = l_{i,t}$. Define for all $t = 1, \dots, T-1$

$$\begin{aligned} R_t &= \begin{pmatrix} \ddots & & & 0 \\ & r(\omega_{t,1}, \omega_i^{t-1}) & \dots & r(\omega_{t,l_{i,t}}, \omega_i^{t-1}) \\ 0 & & & \ddots \end{pmatrix} \in \mathbb{R}^{n|\mathcal{A}^{t-1}| \times |\mathcal{A}^t|} \\ X &= \begin{pmatrix} \ddots & & 0 \\ & \frac{B_T(\omega^T) p_T(\omega^T)}{A_T(\omega^T)} & \\ 0 & & \ddots \end{pmatrix} \in \mathbb{R}^{|\mathcal{A}^T| \times |\mathcal{A}^T|} \\ Q_a &= R_T X \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} X' R_T' \\ q_a &= \mathbf{1}' X' R_T' \\ Q_b &= \begin{pmatrix} \ddots & & 0 \\ & \sum_{\omega^T \in \mathcal{A}(\omega_i^{T-1})} \frac{p_T(\omega^T)}{A_T(\omega^T)} r(\omega^T) r'(\omega^T) & \\ 0 & & \ddots \end{pmatrix} \in \mathbb{R}^{n|\mathcal{A}^{T-1}| \times n|\mathcal{A}^{T-1}|}, \end{aligned}$$

then

$$\begin{aligned} \sigma_T^2(u_{T-1}, \mu) = & \frac{1}{2} \beta_T (< u_{T-1}, Q_a u_{T-1} > - 2\mu < q_a, u_{T-1} > + \mu^2) + \frac{1}{2} < u_{T-1}, Q_b u_{T-1} > \end{aligned} \quad (12)$$

with $u_t = (\dots, u_t'(\omega^t), \dots)'$ for all t .

2.3 The Aggregation Principle

The aggregation starts with the representation of σ_{T-1}^2 . For calculational purposes it is necessary to find Q_{T-1} , q_{T-1} and β_{T-1} , which transform the objective

function of program (??) with $t = T - 1$ to

$$\begin{aligned} \sum_{\omega^{T-1}} &< u_{T-1}(\omega^{T-1}), \frac{1}{2}Q_{T-1}(\omega^{T-1})u_{T-1}(\omega^{T-1}) > p_{T-1}(\omega^{T-1}) + \sigma(u_{T-1}, \mu) \\ &= < u_{T-1}, \frac{1}{2}Q_{T-1}u_{T-1} > - \mu < q_{T-1}, u_{T-1} > + \beta_{T-1}\mu^2. \end{aligned} \quad (13)$$

Let $Q_{c,t}$ be the blockdiagonalmatrix with the elements $Q_t(\omega^t)p_t(\omega^t)$, i.e. for $t = T - 1$

$$Q_{c,T-1} = \begin{pmatrix} \ddots & & 0 \\ & Q_{T-1}(\omega^{T-1})p_{T-1}(\omega^{T-1}) & \\ 0 & & \ddots \end{pmatrix} \in \mathbb{R}^{n|\mathcal{A}^{T-1}| \times n|\mathcal{A}^{T-1}|}. \quad (14)$$

Combining equations (??) with (??) yields

$$\begin{aligned} Q_{T-1} &= \beta_T Q_a + Q_b + Q_{c,T-1}, \\ q_{T-1} &= \beta_T q_a, \\ \beta_{T-1} &= \frac{1}{2}\beta_T. \end{aligned}$$

Starting with the objective function (??) of period $T - 1$, in the following it is shown, that a problem of the form

$$\begin{aligned} \sigma_t^2(u_{t-1}, \mu) &= \min < u_t, \frac{1}{2}Q_t u_t > - \mu < q_t, u_t > + \beta_{t+1}\mu^2 \\ \text{s.t.} \quad &< \mathbf{1}, u_t(\omega^t) > = < r_{t-1}(\omega^t), u_{t-1}(\omega^{t-1}) > \quad \forall \omega^t \end{aligned} \quad (15)$$

in period t yields an analogous problem in period $t - 1$.

Using (??) respectively (??) one obtains the associated Lagrangian

$$\begin{aligned} L_t(u_t, \lambda_t) &= < u_t, \frac{1}{2}Q_t u_t > - \mu < q_t, u_t > + \beta_{t+1}\mu^2 \\ &+ \sum_{\omega^t} \lambda_t(\omega^t) [< r_{t-1}(\omega^t), u_{t-1}(\omega^{t-1}) > - < \mathbf{1}, u_t(\omega^t) >]. \end{aligned}$$

For ease of exposition write

$$\sum_{\omega^t} \lambda_t(\omega^t) [< r_{t-1}(\omega^t), u_{t-1}(\omega^{t-1}) > - < e, u_t(\omega^t) >]$$

in vector form as

$$(R_t' u_{t-1} - E_t' u_t) \lambda_t,$$

with

$$E_t = \begin{pmatrix} 1 & & 0 \\ \vdots & & \\ 1 & & \\ & \ddots & \\ & & 1 \\ & & \vdots \\ 0 & & 1 \end{pmatrix} \in \mathbb{R}^{n|\mathcal{A}^t| \times |\mathcal{A}^t|}.$$

Differentiating the Lagrangian with respect to u_t ,

$$\frac{\partial L_t(u_t, \lambda_t)}{\partial u_t} = Q_t u_t - \mu q_t - E_t \lambda_t$$

yields

$$u_t^* = \mu Q_t^{-1} q_t + Q_t^{-1} E_t \lambda_t. \quad (16)$$

Using

$$E_t' u_t^* = \mu E_t' Q_t^{-1} q_t + E_t' Q_t^{-1} E_t \lambda_t = R_t' u_{t-1},$$

the multipliers associated with (u_{t-1}, μ) are given by the existence of $(E_t' Q_t^{-1} E_t)^{-1}$ with

$$\lambda_t^* = (E_t' Q_t^{-1} E_t)^{-1} R_t' u_{t-1} - \mu (E_t' Q_t^{-1} E_t)^{-1} E_t' Q_t^{-1} q_t.$$

Further, by (??)

$$u_t^* = \underbrace{Q_t^{-1} E_t (E_t' Q_t^{-1} E_t)^{-1} R_t'}_{D_t \in \mathbb{R}^{n k_t \times n k_{t-1}}} u_{t-1} - \underbrace{\mu (Q_t^{-1} E_t (E_t' Q_t^{-1} E_t)^{-1} E_t' Q_t^{-1} q_t - Q_t^{-1} q_t)}_{d_t \in \mathbb{R}^{n k_t}}. \quad (17)$$

$\lambda_t^*(u_{t-1}, \mu)$, $u_t^*(u_{t-1}, \mu)$ are linearly affine in (u_{t-1}, μ) . Clearly, by substituting (??) into (??), the expected efficient frontier $\sigma_t^2(u_{t-1}, \mu)$ is convex quadratic in (u_{t-1}, μ) . With the optimal u^* the associated value function is

$$\begin{aligned} \sigma_t^2(u_{t-1}, \mu) = & \langle u_{t-1}, D_t' \frac{1}{2} Q_t D_t u_{t-1} \rangle \\ & - \mu \langle D_t' Q_t d_t + D_t' q_t, u_{t-1} \rangle \\ & + \underbrace{\mu^2 \left(\langle d_t, \frac{1}{2} Q_t d_t \rangle + \langle q_t, d_t \rangle + \beta_t \right)}_{\beta_{t-1}} \end{aligned}$$

which yields

$$\sigma_t^2(u_{t-1}, \mu) = \langle u_{t-1}, \frac{1}{2} R_t (E_t' Q_t^{-1} E_t)^{-1} R_t' u_{t-1} \rangle - \mu \underbrace{\langle D_t' q_t, u_{t-1} \rangle}_{q_{t-1}} + \mu^2 \beta_{t-1}. \quad (18)$$

With q_{t-1}, β_{t-1} and setting

$$Q_{t-1} = Q_{c,t-1} + R_t (E_t' Q_t^{-1} E_t)^{-1} R_t' \quad (19)$$

the problem in period $t-1$ is analogous to period t . Aggregating successively up to period 0 yields the deterministic problem

$$\begin{aligned} \sigma_0^2(\mu) = \min & \quad \langle u_0, \frac{1}{2} Q u_0 \rangle - \mu \langle q, u_0 \rangle + \beta_1 \mu^2 \\ \text{s.t.} & \quad \langle \mathbf{1}, u_0 \rangle = 1. \end{aligned} \quad (20)$$

We emphasize that Q is again regular of rank n . Both Q and q are determined by a proper aggregation of the variance-covariance matrices and returns on $\mathcal{A}^T$ and of the quadratic and bilinear terms inherent in the expected efficient frontier achieved at and beyond stage $t = 1$. Differentiating the associated Lagrangian yields

$$\frac{\partial L_0(u_0, \lambda_0)}{\partial u_0} = Qu_0 - \mu q - \mathbf{1}\lambda_0$$

and

$$u_0^* = \mu Q^{-1}q + Q^{-1}\mathbf{1}\lambda_0.$$

With

$$\langle \mathbf{1}, u_0^* \rangle = \mu \langle \mathbf{1}, Q^{-1}q \rangle + \langle \mathbf{1}, Q^{-1}\mathbf{1} \rangle \lambda_0 = 1$$

the multiplier is

$$\lambda_0^* = \langle \mathbf{1}, Q^{-1}\mathbf{1} \rangle^{-1} (1 - \mu \langle \mathbf{1}, Q^{-1}q \rangle).$$

The set of efficient portfolios at $t = 0$ is given by

$$u_0^* = \mu Q^{-1}q + Q^{-1}\mathbf{1}(1 - \mu \langle \mathbf{1}, Q^{-1}q \rangle) \langle \mathbf{1}, Q^{-1}\mathbf{1} \rangle^{-1} \quad (21)$$

and represents a linearly affine function in μ . Furthermore, substituting (??) into (??) reveals that the current efficient frontier is convex quadratic in μ , which completes the proof of the theorem.

3 Generation of the Tree-Structure

So far, we considered an optimization problem for a given probability structure, i.e. for a scenario tree with nodes containing a variance-covariance-matrix, a vector of returns, which are achieved in the previous period, and the probability for reaching the node. The task of this section is to model the stochastic process of returns which generates the input factors of the scenario tree.

Furthermore, the model has to allow different number of periods for a pre-defined horizon. Here, the horizon is set to length 1 and the lengths of the periods is $\frac{1}{k}$ for a k stage setting, $k = 1, 2, \dots$. In the case of $k = 1$ the classical Markowitz corresponds to the multiperiod model and the output of this approach is used as a benchmark.

3.1 A Specific Process of Asset Returns

As in the section above we consider a set of n different assets. The value of the n -dim. vector of stock prices S is supposed to follow the process

$$dS_t = S_t(\mu_t dt + \sqrt{\Sigma_t} dz_t), \quad (22)$$

where z_t is a n -dim. vector of independent Brownian motions. The drift μ_t is a n -dim. function of t representing the expected *rate of return* at time t . The term $\sqrt{\Sigma_t}$ is meant to be the Cholesky decomposition of the positive definite matrix Σ_t , whose entries are functions of t . Usually, the diagonal entries of this matrix are referred to as the *stock price volatilities*.

In the proceeding we restrict our attention to the case of $\mu_t = \mu \in \mathbb{R}^n$ and $\Sigma_t = \Sigma \in \mathbb{R}^{n \times n}$, i.e the stock prices S_t follows a n -dim. geometric Brownian motion. This process is widely assumed in the field of financial modelling and application, e.g. the Black-Scholes option pricing formula [?]. Thus, changes in the level of volatility, which is responsible for the well known ‘smile’-effect, or a changing correlation-matrix is ignored. This deficit of constant variances can be overcome by enlarging the scenario tree and assuming that the variances follow specific processes. A common way is to assume that the log-variances are governed by an arithmetic Ornstein-Uhlenbeck process with mean reverting. This is strongly connected to the subject of ARCH modelling treated by Bollerslev, Engle and Nelson [?].

Using Itô’s Lemma one may derive from (??) the logarithm of the stock price. This leads to the generalized Wiener process

$$d \ln S = \left(\mu - \frac{1}{2} \begin{pmatrix} \sigma_1^2 \\ \vdots \\ \sigma_n^2 \end{pmatrix} \right) dt + \sqrt{\Sigma} dz_t. \quad (23)$$

For generating the scenario tree one considers discrete time points rather than a continuous process. The time interval between two adjoining times is determined by the number of stages of the aspired optimization problem. The easiest way is to consider time intervals of equal length $\frac{1}{k}$. Still having the process (??) in mind, the change in $\ln S$ between time t and $t + \frac{1}{k}$ has the multivariate normal distribution

$$\ln S_{t+\frac{1}{k}} - \ln S_t \sim N \left[\left(\mu - \frac{1}{2} \begin{pmatrix} \sigma_1^2 \\ \vdots \\ \sigma_n^2 \end{pmatrix} \right) \frac{1}{k}, \Sigma \frac{1}{k} \right],$$

where $N(m, s)$ denotes a multivariate normal distribution with mean m and variance-covariance-matrix s . Thus, the return

$$r_{\frac{1}{k}} = \frac{S_{t+\frac{1}{k}}}{S_t} = e^{\ln S_{t+\frac{1}{k}} - \ln S_t}$$

of the stock price between two adjacent points of time t and $t + \frac{1}{k}$ is multidimensional lognormally distributed with moments

$$\mathbb{E}(r_{\frac{1}{k},i}) = \exp\left(\mu_i \frac{1}{k}\right) \quad (??a)$$

$$\text{Cov}(r_{\frac{1}{k},i}, r_{\frac{1}{k},j}) = \exp\left(\mu_i \frac{1}{k} + \mu_j \frac{1}{k}\right) (\exp(\sigma_{i,j} \frac{1}{k}) - 1) \quad (??b)$$

as reviewed in Jones and Miller [?].

Obviously one uses the variance-covariance resulting from (??b) as the input matrix for every node in the tree. For the return one can apply any discretization of the multivariate lognormal distribution. This may be the expected return for the following time interval. Thus, instead of considering a tree one

can use a single path, i.e with probability 1, with constant (expected) returns as calculated in (??a).

Notice, that by underlying the process (??), the expected return and the variance-covariance of the stock price on the entire planning horizon is the same regardless of the number of considered time intervalls, i.e. stages of the optimization problem. In particular, this remains valid for the classical one period Markowitz approach.

In the following, we restrain our attention to the case of a single path containing the expected returns. By nature, this is a very crude discretization for P , which shows no variance and may therefore be in contrast to the assumption of a given variance-covariance-matrix.

3.2 Properties of a Single Constant Path

By assuming a scenario tree with a single path containing constant input data $Q_t = Q, r_t = r$ and $p_t = 1$ for all $t = 1, \dots, T$ properties can be derived to lower the calculational effort of the optimization and increase the accuracy of the results.

Using the notation of theorem 1, firstly, one may state $R_t = r$ for all t , $X = \frac{B_T}{A_T}$, $Q_b = \frac{1}{A_T}rr'$ and $Q_a = \frac{B_T}{A_T}$. This yields

$$Q_{T-1} = \gamma_{T-1}rr' + Q, \quad (25)$$

with the parameter $\gamma_T = \beta \frac{B_T^2}{A_T^2} + \frac{1}{A_T}$, and, furthermore, one obtains

$$q_{T-1} = \beta \frac{B_T}{A_T} r. \quad (26)$$

In this special case, E_t is equal to $\mathbf{1}$. Setting $A_t = \mathbf{1}'Q_t^{-1}\mathbf{1}$ and $B_t = \mathbf{1}'Q_t^{-1}q_t$ equation (??) yields

$$Q_{t-1} = \gamma_t rr' + Q \quad (27)$$

with $\gamma_t = \frac{1}{A_t}$, and due to $D_t' = r\mathbf{1}Q_t^{-1}$ equation (??) leads to

$$q_{t-1} = \frac{B_t}{A_t} r. \quad (28)$$

Proposition 1. *If no change of the variance-covariance-matrix and returns occurs in time, then the vectors $Q_t^{-1}e$ and $Q_t^{-1}q_t$ are linearly depending from $Q^{-1}e$ and $Q^{-1}r$ for all $t = 0, \dots, T$.*

Proof. Firstly, one examines the case $t = T - 1$. Suppose

$$Q_{T-1}^{-1}q_{T-1} = \alpha_{T-1}^1 Q^{-1}r.$$

By left multiplying Q_{T-1} , confer (??), one obtains

$$\beta \frac{B_T}{A_T} = \alpha_{T-1}^1 (\gamma_T rr'^T + Q) Q^{-1} r.$$

Knowing that $rr'Q^{-1}r = C_T r$ the above equation yields

$$\beta \frac{B_T}{A_T} = \alpha_{T-1}^1 (\gamma_T C_T + 1) r.$$

Thus

$$\alpha_{T-1}^1 = \frac{\beta B_T}{A_T (\gamma_T C_T + 1)}.$$

In order to show the linear dependence of $Q_{T-1}^{-1}e$ on $Q^{-1}e$ and $Q^{-1}r$ one supposes

$$Q_{T-1}^{-1}e = \alpha_{T-1}^2 Q^{-1}e + \alpha_{T-1}^3 Q^{-1}r$$

and multiplies again with Q_{T-1} . Continuing analogously as for α_{T-1}^1 one obtains the parameters

$$\begin{aligned} \alpha_{T-1}^2 &= 1, \\ \alpha_{T-1}^3 &= -\frac{\gamma_T B_T}{\gamma_T C_T + 1}. \end{aligned}$$

For the cases $t = 0, \dots, T-2$ one proceeds exactly the same way. From the assumptions $Q_{t-1}^{-1}q_{t-1} = \alpha_{t-1}^1 Q^{-1}r$ and $Q_{t-1}^{-1}e = \alpha_{t-1}^2 Q^{-1}e + \alpha_{t-1}^3 Q^{-1}r$ one extracts the parameters

$$\begin{aligned} \alpha_{t-1}^1 &= \frac{B_t}{C_T + A_t}, \\ \alpha_{t-1}^2 &= 1, \\ \alpha_{t-1}^3 &= -\frac{B_T}{C_T + A_t}. \end{aligned}$$

□

For a given expected return μ the optimal portfolio in period 0, see equation (??), is only depending on $A_0, B_0, Q_0^{-1}e$ and $Q_0^{-1}q_0$. However, those variables can be calculated by using the above proposition. Multiplying by $\mathbf{1}$ leads to

$$\begin{aligned} A_{T-1} &= A_T - \frac{B_T^2 C_T}{-B_T^2 + A_T C_T + C^2}, \\ B_{T-1} &= \frac{B_T^2}{-B_T^2 + A_T C_T + C^2}, \\ A_{t-1} &= A_T - \frac{B_T^2}{C_T + A_t} \quad \forall t = 1, \dots, T-1, \\ B_{t-1} &= \frac{B_t B_T}{A_t + C_T} \quad \forall t = 1, \dots, T-1. \end{aligned}$$

Thus, A_0 and B_0 can be calculated successively by A_t and B_t and by this way the optimal portfolio u_0 is determined by a single matrix inversion to compute Q^{-1} .

3.3 Computational Results

For simplicity one assumes that the drift μ_i of one return process consists of a riskless return R and a risk premium $\lambda\sigma_i^2$ with $R, \lambda \in \mathbb{R}$. In vector form this yields

$$\mu = R\mathbf{1} + \lambda \begin{pmatrix} \sigma_1^2 \\ \vdots \\ \sigma_n^2 \end{pmatrix}.$$

Thus, it suffices to estimate the Σ, R and λ to completely determine a scenario tree, single path respectively.

In the following example we consider a portfolio with 10 assets. Again, we chose a single path, the planning horizon is 1 and the length of the subintervals is 1 divided by the stages of the optimization problem. The parameters

$$\begin{aligned} \lambda &= 0.65, \\ R &= 0.06 \end{aligned}$$

are used. The level of volatility σ^2 is between 0.144 and 0.034 and corresponds to drifts between 1.1536 and 1.0821. Furthermore, the correlations consists exclusively of positive values where the majority lie between 0.3 and 0.5.

In the classical one period theory the market portfolio is defined as the constant portfolio, which gives the best possible mix with a riskless asset for any prediscrined level of return. Graphically, the market portfolio is obtained by thr tangent at the efficient frontier cutting the ordinate at the return of the riskless asset. It is illustrated in figure ?? . For the efficient frontier in the multiperiod case one may use this graphical method to calculate the market portfolio.

Figure 1: Graphical way for determining the market portfolio

Table ?? shows the market portfolio as percentage of the initial wealth and corresponding μ and σ for the given input parameters. For the computation exploited the properties of a single constant path.

One recognizes, that the more frequently rebalancing is allowed, i.e. the higher the number of stages considered, the greater is the expected return of the market portfolio by a reduction of the riskmeasure σ . There is a substantial change in the market portfolio between the classical Markowitz and the 2-stage optimization up to 5.62% in asset 3. Between the 2- and 3-stage case there is still a change of up to 2.11% to note, but subsequently, changes are much smaller.

Overall, it suffices to examine the 10-stage case to take all rebalancing activities into consideration. Especially, in face of transaction costs, a greater amount of rebalancing activities might be prohibitive expensive. In table ?? the rebalancing of a 5-stage optimization is documented. One can find the resulting changes of the portfolio in table ?? and they turn out to be very stable.

Periods	μ	σ	Assets 1	2	3	4	5	6	7	8	9	10
1	1.1562	0.0467	-4.80	13.29	20.90	-0.74	24.33	-4.81	22.86	0.66	8.96	19.34
2	1.1587	0.0450	-7.97	10.79	26.52	-3.56	26.52	-5.95	24.72	-1.20	9.97	20.18
3	1.1595	0.0444	-9.11	9.86	28.61	-4.59	27.27	-6.38	25.38	-1.86	10.35	20.47
5	1.1602	0.0439	-10.06	9.08	30.37	-5.44	27.89	-6.73	25.93	-2.40	10.67	20.70
10	1.1608	0.0436	-10.79	8.48	31.74	-6.09	28.36	-7.01	26.34	-2.82	10.91	20.88
100	1.1612	0.0433	-11.46	7.92	33.01	-6.70	28.79	-7.27	26.73	-3.20	11.14	21.04
500	1.1615	0.0435	-11.67	7.75	33.38	-6.88	28.88	-7.32	26.84	-3.32	11.23	21.10

Table 1: Market portfolio for different length of subintervalls, respectively periods

Periods	μ	σ	Assets 1	2	3	4	5	6	7	8	9	10
0	1.1602	0.0439	-10.07	9.08	30.37	-5.44	27.89	-6.74	25.93	-2.40	10.66	20.71
1	1.1294	0.0355	-8.85	10.54	28.77	-4.34	28.14	-6.73	25.93	-1.57	10.30	20.91
2	1.0979	0.0269	-7.64	12.00	27.19	-3.25	28.40	-6.73	25.94	-0.73	9.94	21.13
3	1.0659	0.0182	-6.44	13.46	25.62	-2.16	28.67	-6.74	25.97	0.11	9.58	21.35
4	1.0332	0.0092	-5.23	14.94	24.06	-1.06	28.96	-6.74	26.01	0.94	9.23	21.59

Table 2: Effect of rebalancing for a 5-stage optimization problem

Period	Assets 1	2	3	4	5	6	7	8	9	10
1	1.22	1.46	-1.60	1.10	0.24	0.01	0.00	0.84	-0.37	0.20
2	1.21	1.46	-1.58	1.09	0.26	0.00	0.01	0.84	-0.36	0.21
3	1.21	1.47	-1.57	1.09	0.28	0.00	0.03	0.83	-0.35	0.23
4	1.21	1.48	-1.56	1.09	0.29	-0.01	0.04	0.84	-0.35	0.24

Table 3: Changes of portfolio for a 5-stage optimization problem

4 Final Remarks

So far the objective function is interpreted unsatisfactorily. The increment of wealth between t and $t + 1$ is $\langle r_t(\omega^{t+1}), u_t(\omega^t) \rangle - \langle \mathbf{1}, u_t(\omega^t) \rangle$ and has the variance $\langle u_t(\omega^t), Q_t(\omega^t) u_t(\omega^t) \rangle$. The increment depends firstly on the level of wealth $\langle \mathbf{1}, u_t(\omega^t) \rangle$ and secondly on the portfolio policy. The portfolio strategy is obtained by optimization, but the level of wealth is determined in the preceding periods. Thus the wealth process has depending increments and an addition of the variances does not lead to the variance of wealth of the whole planning horizon.

Taking the expectation firstly leads to a different interpretation. The expected variance of the increment of wealth between t and $t + 1$ is $E(\langle u_t(\omega^t) Q_t(\omega^t) u_t(\omega^t) \rangle)$ and is by nature of the expectation operator solely depending on t and not on the past of $u_t(\omega^t)$. Interpreting and handling the expected variances as normal variances one can add them up to a variance in mean for the portfolio on $[0, 1]$.

A nice feature of the given results is the seeming convergence of the policy if continuous rebalancing is allowed, i.e. $k \rightarrow \infty$. However, one has to emphasize, that these results are in respect to a very crude discretization. More sophisticated discretizations are due to the exponential growth of scenarios, also known as ‘curse of dimensionality’, by now means treatable. Nevertheless, following this way leads to a continuous model of portfolio optimization.

The results in the previous section propose, that one can get along with very few stages. But even in this case the computational effort of enhanced discretizations gets too onerous quickly. Therefore, the algorithm needs to be improved. Because of the quadratic form of the objective function the deterministic equivalent problem shows a specific block-sparse structure for which a solution algorithm is introduced by Steinbach [?]. In particular, this algorithm may optimize portfolios with additional restrictions (e.g. no short sales, transaction costs).

Finally, we want to mention the usual way of discrete time intertemporal portfolio selection. The restrictions of the mean-variance approach are similar to classical models, but differ from the utility based objective function (see, e.g. Ingersoll [?]).

We summarize that an extension of the classical Markowitz approach leads to a multiperiod model with the objective of minimizing a risk measure similar to a variance. The characteristic features of the classical efficient frontier are preserved and one obtains a level of risk and a portfolio selection u_0 for every predefined return. Furthermore, the optimization provides a rebalancing strategy, which takes the dynamics of the underlying stochastic return process into account. The numerical results are obtained by discretizing the log-normal distribution with respect to expectations. In the documented example the market portfolio differs essentially from the classical case. Furthermore, the rebalancing activities are substantial and occur in a stable way.