Local optimality in the multi-dimensional multi-period mean-variance portfolio problem.

Manfred Schäl

Abstract. A variant of the mean-variance problem is studied where the conditional variances of the gain in the last periods are minimized by a backward iteration.

1. Introduction and Summary

In discrete time $t = 0, 1, ..., T$, a frictionless financial market is studied which is free of arbitrage. The investor can invest in a bond and d stocks with unrestricted short selling. The bond is described by a price process $\{B_t,\ 0 \leq t \leq T\}$ and serves as a numeraire. The stocks are described by a d-dimensional price process $\{S_t,\ 0 \leq t \leq T\}$. An initial wealth x and a (self-financing dynamic) portfolio plan ξ define the value process $\{V_t^\xi(x),\ 0 \leq t \leq T\}$. The following mean-variance problem is classical (see [9] for the single-period case and [1], [4], [10], [13] for the multi-period case):

$$\min_\xi Var[V_T^\xi(x)/B_T] \quad \text{subject to} \quad E[V_T^\xi(x)/B_T] = v.$$

If the underlying probability measure P is not just a so-called martingale measure, then the constraint can be satisfied for all given numbers v (by some ξ). Of course, if B_T is deterministic then the problem is the same for the undiscounted and the discounted value process. But we don't make such an assumption and consider the discounted values. Then the problem is the same if we consider the gain process $\{G_t(\xi)\}$ in the discounted model defined by $x + G_t(\xi) := V_t^\xi(x)/B_t$ and study the problem:

$$\min_\xi Var[G_T(\xi)] \quad \text{subject to} \quad E[G_T(\xi)] = g. \tag{1.1}$$

It is known that the solution can be obtained by a solution of a utility maximization problem with a quadratic utility function $U(x) = -\{1 - x\}^2$. For the utility maximization problem one can use the classical criterion of maximizing the expected utility of the gain $G_T(\xi)$ which is considered here as a global optimization problem (see [16] – [21]). If ξ^* is a solution of

$$E[U(G_T(\xi^*))] = \max_\xi E[U(G_T(\xi))]\ , \tag{1.2}$$

then $c \cdot \xi^*$ is a solution of (1.1) with $c = g/(1 - \gamma_0)$ where $\gamma_0 := E[1 - G_T(\xi^*)]$. Obviously, $1 - G_T(\xi^*)$ is the orthogonal projection in $L^2(P)$ of 1 on the orthogonal complement of the linear space $\mathcal{G} = \{G_T(\xi),\ \xi \text{ any plan}\}$. Given the quadratic utility U, one can also define a locally optimal portfolio plan as a modification of Bellman's backward iteration which is different from the classical optimality when the utility is not linear. Let $G_{t,T}(\xi) := G_T(\xi) - G_t(\xi)$ describe the gain between t and T and $E_t[\cdot]$ the conditional expectation given the history of the market up to time t. Then we consider an investor who will not care about the past at any time t but will only consider the future gains. Thus at time $T-1$ the investor considers $\max_\xi E_{T-1}[U(G_{T-1,T}(\xi))] =: E_{T-1}[U(G_{T-1,T}(\xi^0))]$. At time

t the investor knows how he will decide at the times $t+1, ..., T-1$ and thus he knows $G_{t+1,T}(\xi^0)$. Then he wants to solve the following problems by backward iteration :

$$E_t[U(G_{t,T}(\xi^0))] = \max_{\xi} E_t[U(G_{t,T}(\xi))] \tag{1.3}$$

$$\text{subject to } G_{t+1,T}(\xi) = G_{t+1,T}(\xi^0) \ , \ 0 \le t < T.$$

The portfolio plan ξ^0 is obtained by a sequence of orthogonal projections for the one-period models concerning the periods $(t-1, t]$. We will use here the condition 3.2 that the so-called (extended) mean-variance tradeoff process (in the sense of [19]) is deterministic and we will present some sufficient conditions. Condition 3.2 is satisfied if the (vector valued) returns in different periods are independent. But 3.2 is also fulfilled in situations with a deterministic relative risk return process. Under this condition 3.2, it is known that the variance-optimal (signed) martingale measure and the minimal (signed) martingale measure coincide (see [20], [21]). Here, the following will be shown: The expected gains $E[G_T(\xi^*)]$ and $E[G_T(\xi^0)]$ coincide. Thus, given the constraint $E[G_T(\xi)] = 1-\gamma_0$, the investor can choose between ξ^* and ξ^0. Moreover, $E_t[G_{t,T}(\xi^0)] =: 1-\gamma_t$ is deterministic for all t and $c \cdot \xi^0$ solves the following iterative mean-variance problems for the conditional variances Var_t :

$$Var_t[G_{t,T}(c \cdot \xi^0)] = \min_{\xi} Var_t[G_{t,T}(\xi)] \tag{1.4}$$

$$\text{subject to } E_m[G_{m,T}(\xi)] = c \cdot (1-\gamma_m),\ t \le m < T,\ 0 \le t < T.$$

Both ξ^* and ξ^0 can be defined in terms of the adjustment process $\{\lambda_t\}$ (in the sense of [19] – [21]), i.e., the mutual fund principle holds with $\{\lambda_t\}$ as dynamic mutual fund. Here, under condition 3.2, the adjustment process is obtained by simple local (one-step) optimization problems. In (1.4) one can choose c such that $c \cdot \xi^0$ satisfies the desired constraint for $m = 0$. Then the other constraints are given through the numbers γ_m. The situation is particularly simple in the important (stationary) case where the (vector-valued) returns in different periods are independent and identically distributed. Then one has

$$\lambda_t^k = \lambda_0^k \cdot S_0^k \cdot B_t / S_t^k, \quad \gamma_t = (1-\kappa)^{T-t} \quad \text{for some} \quad 0 \le \kappa < 1.$$

Under ξ^0 the situation is more stable in the sense that $E_t[G_{t,T}(\xi^0)] = 1-\gamma_t$ is already known from the beginning. Moreover, an investor using ξ^0 is more risk averse at the beginning than under ξ^* in the sense that he uses a portfolio which is smaller by the factor $0 < \gamma_1 \le 1$. Local optimization in the sense of (1.3) is similar to the concept of local risk minimization (see [7]). In the latter procedure one solves the following problems again by backward iteration for an arbitrary square-integrable contingent claim H (formulated here for self-financing plans):

$$Var_t[\{H - G_{t,T}(\xi^H)\}^2] = \min_{\xi} Var_t[\{H - G_{t,T}(\xi)\}^2]$$

$$\text{subject to } G_{t+1,T}(\xi) = G_{t+1,T}(\xi^H),\ 0 \le t < T.$$

There one has $Var_t[\{H - G_{t,T}(\xi)\}^2] = \min_x E_t[\{H - x - G_{t,T}(\xi)\}^2]$. Of course, in the case $H = 1$ this would lead to the trivial solution $\xi^H = 0$.

2. The financial market model

The evolution of the price of $1+d$ assets will be modeled by a 1-dimensional and a d-dimensional stochastic process $\{B_t,\ t=0,1,...,T\}$ and $\{S_t,\ t=0,1,...,T\}$ on some probability space (Ω, F, P) where T is the time horizon. Then B_t and the components S_t^k, $1 \le k \le d$, of S_t are positive. The asset with price process $\{B_t\}$ is called the **bond** (or bank account). We assume (w.l.o.g.) $B_0 = 1$. The **interest rates** of the bond are then defined through $B_t =: (1+r_1) \cdot ... \cdot (1+r_t)$. The other d assets are called **stocks**. In many cases B_t will be deterministic or predictable. But we don't need such an assumption. The **discounted stock price process** $\breve{S}_t = (\breve{S}_t^1, ..., \breve{S}_t^d)$ is defined by $\breve{S}_t^k := S_t^k / B_t$. It is assumed that $\breve{S}_t^k$ is square-integrable for all k and t. For any vector-valued process $\{Z_t\}$ we define the backward increment by $\Delta Z_t := Z_t - Z_{t-1}$. Further, we write $x^\top$ for the transposed vector and $x^\top \cdot y$ for the inner product of $x, y \in \mathbf{R}^d$. There is some filtration $\{F_t,\ t = 0,\ 1, ..., T\}$ on (Ω, F) such that B_t and S_t are F_t-measurable, *i.e.*, $\{B_t\}$ and $\{S_t\}$ are adapted to $\{F_t\}$. Here, F_t describes the events of the **history** up to time t, where (w.l.o.g.) $F_0 = \{\emptyset, \Omega\}$ and $F_T = F$. Then S_0 is deterministic. A **portfolio plan** is given by an $\mathbf{R}^d$-valued adapted stochastic process $\xi = \{\xi_t,\ 0 \le t < T\}$ such that $\xi_{t-1}^\top \cdot \Delta\breve{S}_t$ is square-integrable for $1 \le t \le T$. During $(t-1, t]$ the investor holds a portfolio $\xi_{t-1} = (\xi_{t-1}^1, ..., \xi_{t-1}^d)$ where ξ_{t-1}^k denotes the number of shares of the $k-th$ stock. We will consider self-financing portfolio plans. Given the initial wealth x, the number η_{t-1} of shares of the bond in $(t-1, t]$ is then specified for $1 \le t \le T$ by ξ according to the **budget equations** $\eta_0 + \xi_0^\top \cdot S_0 = x$, $\eta_t B_t + \xi_t^\top \cdot S_t = \eta_{t-1} B_t + \xi_{t-1}^\top \cdot S_t$. Then the **value process** $\{V_t^\xi(x)\}$ is given through $V_t^\xi(x) := \eta_t B_t + \xi_t^\top \cdot S_t$. As is well known, one can use a reduction to the discounted model where the interest rates of the bond are zero. This leads to

$$V_t^\xi(x)/B_t = x + G_t(\xi)\ ,\ \text{where}\ \ G_t(\xi) := \sum_{1 \le m \le t} \xi_{m-1}^\top \cdot \Delta\breve{S}_m\ ,\ 0 \le t \le T, \tag{2.1}$$

is the accumulated gain in the discounted model. As usual, we assume the well-known **no-arbitrage condition** (see [6], [11]) :

$$\text{for } 1 \le t \le T \text{ and any } \xi: \quad \xi_{t-1}^\top \cdot \Delta\breve{S}_t \ge 0 \ \ a.s. \Longrightarrow \xi_{t-1}^\top \cdot \Delta\breve{S}_t = 0 \ \ a.s.. \tag{NA}$$

3. The one-period models

Throughout the paper we will work with fixed regular conditional probabilities

$$P_t[B](\omega) := P[B|F_t](\omega)\ ,\ B \in F,\ \omega \in \Omega,\ 0 \le t < T. \tag{3.1}$$

There $P_0 = P$. We may assume (w.l.o.g.)

$$E_t[\|\Delta\breve{S}_{t+1}\|^2](\omega) < \infty \text{ for all } \ \omega \in \Omega,\ 0 \le t < T, \text{ where } \quad E_t[Z] := E[Z|F_t] \tag{3.2}$$

denotes the expectation of a random variable Z w.r.t. the fixed conditional probability. Thus (3.2) means that

$$\Delta\breve{S}_{t+1}^k \in L_t^2(\omega) := \{Z \in L^2(P);\ E_t[Z^2](\omega) < \infty\}\ ,\ 1 \le k \le d, \tag{3.3}$$

where $L_t^2(\omega)$ is endowed with the inner product $< Z, Z' >:= E_t[Z \cdot Z'](\omega)$. We may also assume (w.l.o.g.) that the no-arbitrage condition holds locally (see [2] Lemma 2.3, [11](3.22), [16]§2), i.e.,

$$P_t[\vartheta^\top \cdot \Delta\breve{S}_{t+1} \geq 0](\omega) = 1 \Longrightarrow P_t[\vartheta^\top \cdot \Delta\breve{S}_{t+1} = 0](\omega) = 1 \qquad (NA)^*$$
$$\forall \quad \omega \in \Omega, \quad 0 \leq t < T, \quad \vartheta \in \mathbf{R}^{\mathbf{d}}.$$

Let $\mathcal{L}_t(\omega)$ be the smallest linear space L in $\mathbf{R}^d$ such that $P_t[\Delta\breve{S}_{t+1} \in L](\omega) = 1$. Now we define a finite-dimensional linear subspace of $L_t^2(\omega)$ by

$$\mathcal{G}_t := \{\vartheta^\top \cdot \Delta\breve{S}_{t+1};\ \vartheta \in \mathbf{R}^d\}\ . \tag{3.4}$$

Further define $\Pi_t(\omega)Z$ as the orthogonal projection of $Z \in L_t^2(\omega)$ onto $\mathcal{G}_t$ and $\Pi_t^\perp(\omega)Z$ as the orthogonal projection onto the orthogonal complement $\mathcal{G}_t^\perp(\omega)$ of $\mathcal{G}_t$ where
$\mathcal{G}_t^\perp(\omega) = \{Z \in L_t^2(\omega);\ E_t[Z \cdot G](\omega) = 0 \forall G \in \mathcal{G}_t\} = \{Z \in L_t^2(\omega);\ E_t[Z \cdot \Delta\breve{S}_{t+1}^k](\omega) = 0 \forall k\}$.
There $\mathcal{G}_t^\perp(\omega)$ is a closed subset of $L_t^2(\omega)$. Since $\mathcal{G}_t$ is a finite-dimensional linear space, it is also closed and we have $\mathcal{G}_t^{\perp\perp}(\omega) = \mathcal{G}_t$. We define $\lambda_t(\omega) \in \mathcal{L}_t(\omega) \subset \mathbf{R}^d$ such that

$$\Pi_t 1 = \lambda_t^\top \cdot \Delta\breve{S}_{t+1},\ i.e.\ \min_{\vartheta \in \mathbf{R}^d} E_t[\{1 - \vartheta^\top \cdot \Delta\breve{S}_{t+1}\}^2] \text{ is attained at } \ \vartheta = \lambda_t, \tag{3.5}$$

where we know:

$$\min_{\vartheta \in \mathbf{R}^d} E_t[\{1 - \vartheta^\top \cdot \Delta\breve{S}_{t+1}\}^2] = \min_{\vartheta \in \mathcal{L}_t} E_t[\{1 - \vartheta^\top \cdot \Delta\breve{S}_{t+1}\}^2]. \tag{3.6}$$

Proposition 3.1. (a) λ_t *is unique.*
(b) $1 - \lambda_t^\top \cdot \Delta\breve{S}_{t+1} = \Pi_t 1$.
(c) $0 < E_t[\{1 - \lambda_t^\top \cdot \Delta\breve{S}_{t+1}\}^2] = E_t[1 - \lambda_t^\top \cdot \Delta\breve{S}_{t+1}] \leq 1$.

Proof. a) Assume that the minimum in (3.6) is attained at $\vartheta_t \in \mathcal{L}_t$. Then $P_t[\lambda_t^\top \cdot \Delta\breve{S}_{t+1} = \vartheta_t^\top \cdot \Delta\breve{S}_{t+1}] = 1$ and hence $P_t[(\lambda_t - \vartheta_t)^\top \cdot \Delta\breve{S}_{t+1} = 0] = 1$. This implies that $\lambda_t - \vartheta_t \perp \mathcal{L}_t$ where $\lambda_t - \vartheta_t \in \mathcal{L}_t$ and hence $\lambda_t - \vartheta_t = 0$. b) We have $1 = \Pi_t^\perp 1 + \Pi_t 1 \ \ a.s.$.
c) We have $E_t[\Pi_t^\perp 1] = E_t[\Pi_t^\perp 1 \cdot (\Pi_t 1 + \Pi_t^\perp 1)] = E_t[\{\Pi_t^\perp 1\}^2]$. Further we know that
$1 = E_t[\{\Pi_t^\perp 1\}^2] + E_t[\{\Pi_t 1\}^2]$. Now assume that $E_t[\{\Pi_t^\perp 1\}^2] = 0$; then
$P_t[\lambda_t^\top \cdot \Delta\breve{S}_{t+1} = 1] = 1$ which is a contradiction to (NA)*. []
The arguments for 3.1c are the same as in [21](1.5). We set

$$\kappa_t := E_t[\lambda_t^\top \cdot \Delta\breve{S}_{t+1}]. \tag{3.7}$$

Then by (3.5) and 3.1c we have

$$0 < 1 - \kappa_t = \min_{\vartheta \in \mathbf{R}^d} E_t[\{1 - \vartheta^\top \cdot \Delta\breve{S}_{t+1}\}^2] \leq 1. \tag{3.8}$$

There $1 - \kappa_t$ can be looked upon as the measure of a local risk. For the main results of the paper we will use the following condition:

Condition 3.2 . *The quantities κ_t are deterministic for* $0 \leq t < T$.

Under this condition, the process $\{\lambda_t\}$ coincides with the *adjustment process* in the sense of [19] – [21]. Moreover, Schweizer calls the process $\tilde{K}_t := \sum_{m=0}^{t-1} \kappa_m$ the *extended mean-variance tradeoff*. Now we provide sufficient conditions for 3.2. Of course κ_0 is always deterministic. The **relative risk process** $\{R_t = (R_t^1, ..., R_t^d),\ 1 \le t \le T\}$ is defined by

$$1 + R_t^k := 1 + \Delta \breve{S}_t^k / \breve{S}_{t-1}^k = (1 + \Delta S_t^k /\ S_{t-1}^k)/(1 + r_t).$$

The condition of the next proposition is fulfilled in many market models.

Proposition 3.3. *For* $1 \le t < T$, κ_t *is deterministic if* R_{t+1} *and* F_t *are stochastically independent.*

The proof will be given for the more general situation of 3.4 . It shows that in the situation of 3.3:

$$\lambda_t^k = \delta_t^k / \breve{S}_t^k, 1 \le k \le d, \text{for some deterministic numbers } \delta_t^k. \tag{3.9}$$

Condition 3.4. *There exist some* $n \ge 1$, *a sequence* $\{\theta_t,\ 0 \le t < T\}$ *of* n*-dimensional vectors, an adapted* $\mathbf{R}^n$*-valued process* $\{W_t,\ 0 \le t \le T\}$ *and an adapted* $\mathbf{R}^{d \times n}$*-valued process* $\{\sigma_t,\ 0 \le t < T\}$, *such that*:

(1) $R_t^k = \sum_{j=1}^n \sigma_{t-1}^{k,j} \cdot (\theta_{t-1}^j + \Delta W_t^j)$, $1 \le k \le d$;

(2) ΔW_t *and* F_{t-1} *are stochastically independent for* $2 \le t \le T$;

(3) *the range* $\{\sigma_{t-1}^\top(\omega)\vartheta;\ \vartheta \in \mathbf{R}^d\}$ *of* $\sigma_{t-1}^\top(\omega)$ *does not depend on* $\omega \in \Omega$, $1 \le t < T$.

Obviously, condition 3.4 (3) is satisfied if $n = d$ and $\sigma_{t-1}(\omega)$ has full rank for all ω and t. The situation in 3.3 is included; we just have to choose $\Delta W_t = R_t$. Another interesting example is a discrete-time stochastic difference equation where $\{W_t\}$ is a discrete-time white noise, i.e., ΔW_t is independent of F_{t-1} and for each t the random variables $\Delta W_t^1, ..., \Delta W_t^d$ are uncorrelated and standardized (with mean zero and variance one). Then θ_t is called **relative risk vector** or **unit price for risk** which is deterministic under condition 3.4**.** Then we have for the conditional expectation and the conditional covariance matrix of R_t :

$$E_{t-1}[R_t] = \sigma_{t-1}\theta_{t-1} \ ,\ Cov_{t-1}[R_t] = \sigma_{t-1}\sigma_{t-1}^\top \ . \tag{3.10}$$

For $n = d = 1$, the latter situation can also be found in GARCH-M-models (see [5]).

Theorem 3.5. *Under condition* 3.4, κ_t *is deterministic for all* t, *i.e., Condition* 3.2 *is fulfilled.*

Proof. We set $I_{t-1} := \{\sigma_{t-1}^\top \vartheta;\ \vartheta \in \mathbf{R}^d\} \subset \mathbf{R}^n$ and $\theta_{t-1}^j + \Delta W_t^j =: U_t^j$. Then $U_t = (U_t^1, ..., U_t^d)^\top$ is independent of F_{t-1} and I_{t-1} is deterministic by assumption. According to (3.8) we have $1 - \kappa_t = \min_{\vartheta \in \mathbf{R}^d} E_t[\{1 - \vartheta^\top \cdot \Delta \breve{S}_{t+1}\}^2] = \min_{\delta \in \mathbf{R}^d} E_t[\{1 - \sum_{k=1}^d (\delta^k / \breve{S}_t^k) \cdot \Delta \breve{S}_{t+1}^k\}^2] = \min_\delta E_t[\{1 - \sum_{k=1}^d \delta^k \cdot \left(\sum_{j=1}^n \sigma_t^{k,j} \cdot U_{t+1}^j\right)\}^2] = \min_\delta E_t[\{1 - \delta^\top \sigma_t U_{t+1}\}^2] = \min_{\zeta \in I_t} E_t[\{1 - \zeta^\top \cdot U_{t+1}\}^2] = \min_{\zeta \in I_t} E[\{1 - \zeta^\top \cdot U_{t+1}\}^2]$.[]

Lemma 3.6. *For any* $Z \in L^2(P)$ *we have:* $\Pi_t(Z) = \Pi_t(E_{t+1}[Z])$, $0 \le t < T$.

The proof follows from

$$E_t[\{Z - \vartheta^\top \cdot \Delta \breve{S}_{t+1}\}^2] = E_t[\{Z - E_{t+1}[Z]\}^2] + E_t[\{E_{t+1}[Z] - \vartheta^\top \cdot \Delta \breve{S}_{t+1}\}^2] \ . \tag{3.11}$$

4. Global optimality

Definition 4.1. *A portfolio plan ξ^* is called globally optimal for $c \in \mathbf{R}$ if:*

$$E[\{c - G_T(\xi^*)\}^2] \leq E[\{c - G_T(\xi)\}^2] \text{ for all } \xi. \tag{4.1}$$

By the following relation we can restrict attention to the case $c = 1$. There we set $c \cdot \xi := (c \cdot \xi_t)$.

Lemma 4.2. *If ξ^* is globally optimal for 1,then $c \cdot \xi^*$ is globally optimal for $c \in \mathbf{R}$.*

For a proof of the following result see [20] or the proof of Theorem 5.4 below.

Proposition 4.3. *If ξ^* globally optimal for 1 , then*

$$Var[G_T(c \cdot \xi^*)] \leq Var[G_T(\xi)] \quad \text{for all } \xi \text{ with } E[G_T(\xi)] = c \cdot E[G_T(\xi^*)].$$

We will compute $E[G_T(\xi^*)]$ in 4.4 below. In order to solve (4.1) for $c = 1$, one can write $1 - G_T(\xi) = 1 + G_T(-\xi)$. There $1 + G_T(-\xi)$ is the discounted value process $of -\xi$ starting with an initial value 1 and we want to study

$$\min_{\xi} E[\{1 - G_T(\xi)\}^2] = \min_{\xi} E[\{1 + G_T(\xi)\}^2]. \tag{4.2}$$

It is convenient to describe a portfolio plan $\pm\xi$ by a *proportional portfolio plan* $\pm\varphi = (\pm\varphi_t, 0 \leq t < T)$. As a portfolio plan, a proportional portfolio plan φ is an $\mathbf{R}^d$-valued adapted process where $\pm\varphi_t^k$ represents the proportion $\pm\xi_t^k/(1 \pm G_t(\xi))$ for $1 \pm G_t(\xi) \neq 0$ and $\varphi_t^k = 0$ for $1 \pm G_t(\xi) = 0$. Then

$$1 \pm G_t(\xi) = (1 \pm \varphi_0^\top \cdot \Delta\breve{S}_1) \cdot \ldots \cdot (1 \pm \varphi_{t-1}^\top \cdot \Delta\breve{S}_t) \ . \tag{4.3}$$

Note that the property "$1 \pm G_{t-1}(\xi) = 0 \implies 1 \pm G_t(\xi) = 0$", $1 \leq t < T$, is both necessary and sufficient to describe a portfolio plan by a proportional portfolio plan. But obviously, when studying (4.2) we can restrict attention to such plans.

Theorem 4.4. *Under condition 3.2 we have for $\xi_t^* := (1 - G_t(\xi^*))\lambda_t$, $0 \leq t < T$:*
ξ^* is a portfolio plan which is globally optimal for 1 and

$$1 - G_t(\xi^*) = (1 - \lambda_0^\top \cdot \Delta\breve{S}_1) \cdot \ldots \cdot (1 - \lambda_{t-1}^\top \cdot \Delta\breve{S}_t),$$

$$E[1 - G_t(\xi^*)] = E[\{1 - G_t(\xi^*)\}^2] = \Pi_{m=0}^{t-1}(1 - \kappa_m) \ , \ 1 \leq t \leq T.$$

Proof. In view of (3.5) and (3.8) we know that $E_t[\{1 + \varphi_{t-1}^\top \cdot \Delta\breve{S}_t\}^2] \geq E_t[\{1 - \lambda_{t-1}^\top \cdot \Delta\breve{S}_t\}^2] = 1 - \kappa_t$. Thus $E[\{\Pi_{t=1}^T(1 + \varphi_{t-1}^\top \cdot \Delta\breve{S}_t)\}^2] \geq E[\{(1 - \lambda_{T-1}^\top \cdot \Delta\breve{S}_T) \cdot \Pi_{t=1}^{T-1}(1 + \varphi_{t-1}^\top \cdot \Delta\breve{S}_t)\}^2] = (1 - \kappa_{T-1}) \cdot E[\{\Pi_{t=1}^{T-1}(1 + \varphi_{t-1}^\top \cdot \Delta\breve{S}_t)\}^2]$. By induction we obtain: $E[\{\Pi_{t=1}^T(1 + \varphi_{t-1}^\top \cdot \Delta\breve{S}_t)\}^2] \geq E[\{\Pi_{t=1}^T(1 - \lambda_{t-1}^\top \cdot \Delta\breve{S}_t)\}^2]$ and $E[\{\Pi_{m=1}^t(1 - \lambda_{m-1}^\top \cdot \Delta\breve{S}_m)\}^i] = \Pi_{m=0}^{t-1}(1 - \kappa_m)$ for $i = 1, 2$. Applying well-known selection theorems, it is easy to see that λ_t can be chosen as F_t-measurable random variable (see [12] Proposition 2.2, [16] Proposition 4.7).

Moreover, we have $E[\{\xi_t^{*\top} \cdot \Delta\breve{S}_{t+1}\}^2] = E[\{(\Pi_{m=1}^t(1-\lambda_{m-1}^\top \cdot \Delta\breve{S}_t) \cdot (\lambda_t^\top \cdot \Delta\breve{S}_{t+1})\}^2] = \Pi_{m=0}^{t-1}(1-\kappa_m) \cdot \kappa_t \leq 1$. []

5. Local optimality

In addition to (2.1) let us define :

$$G_{t,T}(\xi) := \sum_{t<m\leq T} \xi_{m-1}^\top \cdot \Delta\breve{S}_m \ , \ W_{t,T}(\xi) := E_t[G_{t,T}(\xi)] \ . \tag{5.1}$$

Definition 5.1. *A portfolio plan ξ^0 is called locally optimal for $c \in \mathbf{R}$ if $\forall 0 \leq t < T$:*

$$E_t[\{c - G_{t,T}(\xi^0)\}^2] \leq E_t[\{c - G_{t,T}(\xi)\}^2] \quad \text{for all } \ \xi \text{ with } \ G_{t+1,T}(\xi) = G_{t+1,T}(\xi^0).$$

For a plan ξ as in 5.1 we have $G_{t,T}(\xi) = \xi_t^\top \cdot \Delta\breve{S}_{t+1} + G_{t+1,T}(\xi^0)$. The following lemma shows that local optimality can be based on the conditional future gains $W_{t,T}(\xi)$. Similarly, the concept of recursive utility is based on the conditional utilitiy of the future gains (see [8]).

Lemma 5.2. *A portfolio plan ξ^0 is locally optimal for c if and only if $\forall 0 \leq t < T$, $\forall \vartheta \in \mathbf{R}^d$:*

$$E_t[\{c - \xi_t^{0\top} \cdot \Delta\breve{S}_{t+1} - W_{t+1,T}(\xi^0)\}^2] \leq E_t[\{c - \vartheta^\top \cdot \Delta\breve{S}_{t+1} - W_{t+1,T}(\xi^0)\}^2] \ .$$

Proof. In view of 3.6, ξ^0 is locally optimal if and only if
$\xi_t^{0\top} \cdot \Delta\breve{S}_{t+1} = \Pi_t(c - G_{t+1,T}(\xi^o)) \forall 0 \leq t < T \iff \xi_t^{0\top} \cdot \Delta\breve{S}_{t+1} = \Pi_t(E_{t+1}[c - G_{t+1,T}(\xi^0)])$ $= \Pi_t(c - W_{t+1,T}(\xi^0)) \quad \forall 0 \leq t < T$ which is equivalent to the condition of the lemma. []
For many considerations it is again sufficient to consider the case $c = 1$. Indeed we have:

Lemma 5.3. *If ξ^0 is locally optimal for 1, then $c \cdot \xi^0$ is locally optimal fiir c.*

Now we can give a characterization of local optimality in a larger set of plans and the relation to the minimization of the conditional variances given the conditional means where we set

$$Var_t[Z] := E_t[Z^2] - E_t[Z]^2.$$

Theorem 5.4. *If ξ^0 is locally optimal for* 1, *then* for $0 \leq t < T$:
(a) $E_t[\{1 - G_{t,T}(\xi^0)\}^2] \leq E_t[\{1 - G_{t,T}(\xi)\}^2] \quad \forall\xi$ *with* $W_{m,T}(\xi) = W_{m,T}(\xi^0)$, $t < m < T$;
(b)$Var_t[G_{t,T}(\xi^0)] \leq Var_t[G_{t,T}(\xi)] \quad \forall\xi$ *with* $W_{m,T}(\xi) = W_{m,T}(\xi^0)$, $t \leq m < T$.

The quantities $W_{m,T}(\xi^0)$ will be computed in 5.6 below.
Proof. First consider the case $t = T-1$. Then (a) is clear and (b) follows as below. Now assume that (b) holds for $t+1$ and $W_{m,T}(\xi) = W_{m,T}(\xi^0)$, $t < m < T$. Then from (3.11) and Lemma 5.2 we obtain $E_t[\{1 - G_{t,T}(\xi)\}^2] = E_t[\{1 - \xi_t^\top \cdot \Delta\breve{S}_{t+1} - W_{t+1,T}(\xi)\}^2] + E_t[\{G_{t+1,T}(\xi) - W_{t+1,T}(\xi)\}^2] = E_t[\{1 - \xi_t^\top \cdot \Delta\breve{S}_{t+1} - W_{t+1,T}(\xi^0)\}^2] + E_t[Var_{t+1}[G_{t+1,T}(\xi)]]$ $\geq E_t[\{1 - \xi_t^{0\top} \cdot \Delta\breve{S}_{t+1} - W_{t+1,T}(\xi^0)\}^2] + E_t[Var_{t+1}[G_{t+1,T}(\xi^0)]] = E_t[\{1 - G_{t,T}(\xi^0)\}^2]$.

If in addition $W_{t,T}(\xi) = W_{t,T}(\xi^0)$, then $Var_t[G_{t,T}(\xi)] = Var_t[1 - G_{t,T}(\xi)] = E_t[\{1 - G_{t,T}(\xi)\}^2] - \{1 - W_{t,T}(\xi)\}^2 \geq E_t[\{1 - G_{t,T}(\xi^0)\}^2] - \{1 - W_{t,T}(\xi^0)\}^2 = Var_t[G_{t,T}(\xi^0)]$. []

Proposition 5.5. *A plan ξ^0 is locally optimal for* 1 *if and only if*
$\xi^0_{t-1} \cdot \Delta\breve{S}_t = \Pi_{t-1}(1 - G_{t,T}(\xi^0)) = \Pi_{t-1}(E_t[1 - G_{t,T}(\xi^0)])$. *Then*
(a) $1 - G_{t,T}(\xi^0) = \Pi^{\perp}_t(1 - G_{t+1,T}(\xi^0)) = \Pi^{\perp}_t(\Pi^{\perp}_{t+1}(...\Pi^{\perp}_{T-1}(1)))$;
(b) $E_t[1 - G_{t,T}(\xi^0)] = E_t[\Pi^{\perp}_t(E_{t+1}[1 - G_{t+1,T}(\xi^0)])]$
$= E_t[\Pi^{\perp}_t(E_{t+1}[\Pi^{\perp}_{t+1}(...E_{T-1}[\Pi^{\perp}_{T-1}(1)])])])]$.

Proof. First we consider the case $t = T - 1$. Then we have: $\xi^{0\top}_{T-1} \cdot \Delta\breve{S}_T = \Pi_{T-1}(1)$ and $1 - G_{T-1,T}(\xi^0) = 1 - \Pi_{T-1}(1) = \Pi^{\perp}_{T-1}(1)$, hence $E_{T-1}[1 - G_{T-1,T}(\xi^0)] = E_{T-1}[\Pi^{\perp}_{T-1}(1)]$. According to 5.2 and 3.6 we know: $\xi^{0\top}_{t-1} \cdot \Delta\breve{S}_t = \Pi_{t-1}(1 - G_{t,T}(\xi^0)) = \Pi_{t-1}(E_t[1 - G_{t,T}(\xi^0)])$. Thus we obtain $1 - G_{t-1,T}(\xi^0) = 1 - G_{t,T}(\xi^0) - \xi^{0\top}_{t-1} \cdot \Delta\breve{S}_t = 1 - G_{t,T}(\xi^0) - \Pi_{t-1}(1 - G_{t,T}(\xi^o))$ hence $1 - G_{t-1,T}(\xi^0) = \Pi^{\perp}_{t-1}(1 - G_{t,T}(\xi^0))$. Similarly we get $E_t[1 - G_{t-1,T}(\xi^0)] = E_t[1 - G_{t,T}(\xi^0)] - \xi^{0\top}_{t-1} \cdot \Delta\breve{S}_t = E_t[1 - G_{t,T}(\xi^0)] - \Pi_{t-1}(E_t[1 - G_{t,T}(\xi^0)])$, hence $E_t[1 - G_{t-1,T}(\xi^0)] = \Pi^{\perp}_t(E_t[1 - G_{t,T}(\xi^0)])$. Now we can use backward induction. []
We remark that in 5.1, 5.2, 5.3, and 5.5 one can replace the function 1 by an arbitrary random variable $X \in L^2$. The result in 5.4 can be compared with the following relation for the plan ξ^* being globally optimal for 1: $1 - G_T(\xi^*) = \Pi^{\perp}(1)$ where $\Pi^{\perp}$ is the projection in $L^2(P)$ on $\mathcal{G}^{\perp}$, and $\mathcal{G}$ is the linear space of gains $G_T(\xi)$ under all plans ξ (see [21]). Now we will use our Condition 3.2 and we set

$$0 < \gamma_t := \Pi^{T-1}_{m=t}(1 - \kappa_m) \leq 1,\ 0 \leq t < T,\ \gamma_T := 1. \tag{5.5}$$

Theorem 5.6. *Under Condition* 3.2 *we have for $\xi^0_t = \gamma_{t+1}\lambda_t$, $0 \leq t < T$: ξ^0 is locally optimal for* 1. *Then*: $1 - W_{t,T}(\xi^0) = \gamma_t$, $0 \leq t < T$, *in particular* $E[1 - G_T(\xi^0)] = \gamma_0$.

Proof. According to 5.5 we have $\xi^{0\top}_{t-1} \cdot \Delta\breve{S}_t = \Pi_{t-1}(\gamma_t) = \gamma_t \cdot \Pi_{t-1}(1) = \gamma_t\lambda^{\top}_{t-1} \cdot \Delta\breve{S}_t$. From 5.8 below we conclude that ξ^0 is a plan. Then we know that $E_{T-1}[1 - G_{T-1,T}(\xi^0)] = E_{T-1}[\Pi^{\perp}_{T-1}(1)] = 1 - \kappa_{T-1} = \gamma_{T-1}$. Now assume that $1 - W_{t+1,T}(\xi^0) = \gamma_{t+1}$, then we obtain by use of 5.5b: $E_t[1 - G_{t,T}(\xi^0)] = E_t[\Pi^{\perp}_t(E_{t+1}[1 - G_{t+1,T}(\xi^0)])] = E_t[\Pi^{\perp}_t(\gamma_{t+1})] = \gamma_{t+1} \cdot E_t[\Pi^{\perp}_t(1)] = \gamma_{t+1} \cdot (1 - \kappa_t) = \gamma_t$. []
From 4.4 and 5.6 we obtain

$$\xi^0_t = \gamma_{t+1} \cdot \lambda_t,$$
$$\xi^*_t = (1 - G_t(\xi^*)) \cdot \lambda_t, \quad 0 \leq t < T.$$

Corollary 5.7. *Assume Condition* 3.2. *If ξ^0 and ξ^* are the locally and the globally optimal plan for* 1, *respectively, then*: $E[G_T(\xi^*)] = E[G_T(\xi^0)] = 1 - \gamma_0$.
Moreover, $E_t[\xi^{0\top}_t \cdot \Delta\breve{S}_{t+1}] = \kappa_t \cdot \gamma_{t+1}$ *is deterministic whereas*
$E_t[\xi^{*\top}_t \cdot \Delta\breve{S}_{t+1}] = (1 - G_t(\xi^*) \cdot \kappa_t = (1 - \lambda^{\top}_0 \cdot \Delta\breve{S}_1) \cdot ... \cdot (1 - \lambda^{\top}_{t-1} \cdot \Delta\breve{S}_t) \cdot \kappa_t$.

Now we want to compare the conditional variances of ξ^0 and ξ^*. From 3.1c we conclude:

Lemma 5.8. $E_t[\{\lambda^{\top}_t \cdot \Delta\breve{S}_{t+1}\}^2] = \kappa_t$, $Var_t[\lambda^{\top}_t \cdot \Delta\breve{S}_{t+1}] = (1 - \kappa_t) \cdot \kappa_t$.

Now we get:

Corollary 5.9. *Assume Condition* 3.2, *then we obtain for* ξ^* *and* ξ^0 *as in* 5.7 :
$Var_t[\xi_t^{0\top} \cdot \Delta\check{S}_{t+1}] = (1-\kappa_t) \cdot \kappa_t \cdot \gamma_{t+1}^2$; $Var_t[\xi_t^{*\top} \cdot \Delta\check{S}_{t+1}] = (1-G_t(\xi^*))^2 \cdot (1-\kappa_t) \cdot \kappa_t$.

Comparing ξ^0 and ξ^* for $t=0$ and $t=T-1$ we see that:

$$\xi_0^0 = \gamma_1\lambda_0, \quad \xi_0^* = \lambda_0, \quad \xi_{T-1}^0 = \lambda_{T-1}, \quad \xi_{T-1}^* = (1-G_{T-1}(\xi^*))\lambda_{T-1},$$

where $0 < \gamma_1 \leq 1$, $0 < E[1-G_{T-1}(\xi^*)] \leq 1$.
Thus, at the beginning the investor is less risk averse under ξ^0 than under ξ^*. At the end, the opposite property holds but "only in the mean".

References.

[1] E.J. Elton and M.J. Gruber, *Finance as a Dynamic Process*, Englewood cliffs: Prentice Hall, 1975.

[2] R.C. Dalang, A. Morton and W. Willinger, *Equivalent martingale measures and no-arbitrage in stochastic securities market models*, Stochastics and Stochastic Reports **29** (1990), 185 — 201, .

[3] P. Grandits, *The p-optimal martingale measure and its asymptotic relation with the minimal-entropy martingale measure*", Bernoulli **5** (1999), 225-248.

[4] R.R. Grauer and N.H. Hakansson, *On the use of mean-variance and quadratic approximations in implementing dynamic investment policies*, Manag. sci., **39** (1993), 856-871.

[5] W. Härdle and C.M. Hafner, *Discrete time option pricing with flexible volatility estimation*, Finance Stochast. **4** (2000), 189-208.

[6] J. Jacod and A. N. Shiryaev, *Local martingales and the fundamental asset pricing theorems in the discrete-time case*, Finance Stochast. **3** (1998), 259-273.

[7] D. Lamberton, H. Pham and M. Schweizer, *Local risk-minimization under transaction costs*, *Math. Oper. Res.*, **23** (1998), 585-612.

[8] C. Ma, *A discrete-time intertemporal asset pricing model*, Mathematical Finance **8** (1998), 249-275.

[9] H.M. Markowitz , *Portfolio selection*, J. of Fin., **7** (1952), 77-91.

[10] J. Mossin, *Optimal multiperiod portfolio policies*, J. of Bus., **41**, (1068), 215-229.

[11] S.R. Pliska, *Introduction to Mathematical Finance*, Blackwell publisher, Malden, USA, Oxford, UK, 1997.

[12] L.C.G. Rogers, *Equivalent martingale measures and no-arbitrage*, Stochastics and Stochastic Reports **51** (1994), 41-49.

[13] P.A. Samuelson, *Lifetime portfolio selection by dynamic stochastic programming*, Rev. Econ. Stat., **51** (1969), 239-246.

[14] M. Schäl, *On quadratic cost criteria for option hedging*, Math. Oper. Res. **19** (1994), 121-131.

[15] M. Schäl, *Martingale measures and hedging for discrete-time financial markets*, Math. Oper. Res. **24** (1999), 509-528.

[16] M. Schäl, *Portfolio optimization and martingale measures*, Mathematical Finance **10** (2000), 289-304.

[17] M. Schäl, *Price systems constructed by optimal dynamic portfolios*, Math. Meth. Oper. Res. **51** (2000), 375-397.

[18] M. Schäl, *Markov decision theory in finance and dynamic options.* To be published in Shwartz A, Feinberg E (eds.) Markov decision processes, (2000) Kluwer.
[19] M. Schweizer, *Approximating random variables by stochastic integrals*, Ann. Probab. **22** (1994), 1536-1575.
[20] M. Schweizer, *Variance-optimal hedging in discrete time*, Math. Oper. Res. **20** (1995), 1-32.
[21] M. Schweizer, *Approximation pricing and the variance-optimal martingale measure,* Ann. Probability **24** (1996), 206-236.

Institut für Angewandte Mathematik,
Universität Bonn,
Wegelerstr. 6,
$D-53115$ Bonn, Germany
E-mail address: schael@uni-bonn.de