Michał Motoczyński (Warszawa)

RISK MINIMIZATION IN
THE MODEL WITH TRANSACTION COSTS

Abstract. The problem of hedging a contingent claim with minimization of quadratic risk is studied. Existence of an optimal strategy for the model with proportional transaction cost and nondelayed observation is shown.

1. Introduction. Assume that on a given probability space $(\Omega, \mathcal{F}, P)$ equipped with a sequence $\mathcal{F}$ of increasing σ-fields $\mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_T$ there is defined a stock price process $S = (S_k)_{k=1,\ldots,T}$, that is, a sequence of $\mathcal{F}$-adapted, almost surely positive random variables such that

$$\mathrm{E}(S_k^2) = \|S_k\|_{L^2(P)} < \infty$$

for every $k = 1, \ldots, T$.

Assume that there are proportional transaction costs given by positive constants μ and λ. If an investor buys one unit of the asset S at time k, he pays $(1 + \lambda)S_k$, while when selling it he receives only $(1 - \mu)S_k$.

We consider so called trading strategies which are defined as follows.

DEFINITION 1. A *trading strategy* $\theta = (\theta_k)_{k=1,\ldots,T}$ is a pair of $\mathcal{F}$-adapted processes $(m_k, l_k)_{k=1,\ldots,T}$ which satisfy the following conditions:

(a1) $m_k, l_k \geq 0$,
(a2) $m_k S_k, l_k S_k \in L^2(P)$,

for all $k = 1, \ldots, T$. Denote by Θ the set of all trading strategies.

The financial interpretation of m_k and l_k is that they represent quantities of stocks that are sold or bought respectively at time k.

In what follows we assume that the investor has access to all available information about the price of asset S.

DEFINITION 2. Let $\theta = (\theta_k)_{k=1,\ldots,T}$, $\theta_k = (m_k, l_k)$, be a trading strategy.

(i) A process $X(\theta) = (X_k(\theta))_{k=1,\ldots,T}$ such that $X_0 = c \in \mathbb{R}$ is the initial endowment and

$$X_k(\theta) = (1 + r)X_{k-1}(\theta) + (1 - \mu)m_k S_k - (1 + \lambda)l_k S_k$$

is called the *process of capital invested in a riskless bond*.

(ii) A process $Y(\theta) = (Y_k(\theta))_{k=1,\ldots,T}$ such that

$$Y_k(\theta) = \bar{\xi}_k S_k,$$

where $\bar{\xi}_k = \sum_{i=1}^{k}(l_i - m_i)$, is called the *process of capital invested in the asset S*. Notice that $\bar{\xi}_k$ denotes the number of stocks held in the time interval $[k, k+1)$.

Without loss of generality one may assume that the interest rate r is equal to 0. Then

$$X_k(\theta) = X_0 + \sum_{i=1}^{k} \Delta X_i(\theta) = X_0 + \sum_{i=1}^{k}((1 - \mu)m_i S_i - (1 + \lambda)l_i S_i).$$

Notice that for each $k = 1, \ldots, T$, $X_k(\theta) \in L^2(P)$. However square integrability of the process $Y(\theta)$ is not assumed. To obtain it we impose an additional assumption (RB) on the price process S, introduced in [4].

DEFINITION 3. The price process S satisfies the *relative boundedness condition* (RB) if there exists a constant C such that

$$\mathrm{E}\left[\frac{S_{k+1}^2}{S_k^2} \,\middle|\, \mathcal{F}_k\right] \leq C$$

almost surely for each $k = 1, \ldots, T - 1$.

We recall a simple lemma which will be very useful later on (for a proof see [4]).

LEMMA 1. *Let Z be any $\mathcal{F}_i$-measurable random variable such that $\|ZS_j\|_{L^2} < K$ for some $j \in \{i, \ldots, T\}$. If the price process S satisfies the (RB) condition then*

$$\|ZS_k\|_{L^2} < KC^{k-j}$$

for any $k \in \{j, \ldots, T\}$.

From Lemma 1 we immediately obtain the following corollary.

COROLLARY 1. *If S satisfies (RB) then the process $[X(\theta), Y(\theta)]'$ is square integrable, i.e. $[X(\theta), Y(\theta)]' \in L^2(P) \times L^2(P)$ for any trading strategy $\theta \in \Theta$.*

DEFINITION 4. The *gain functional* is the transformation $G_T : \Theta \to L^2(P) \times L^2(P)$ of the form

$$G_T(\theta) = \begin{bmatrix} X_T(\theta) - X_0 \\ Y_T \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{T}((1-\mu)m_k S_k - (1+\lambda)l_k S_k) \\ (\sum_{k=1}^{T}(l_k - m_k))S_T \end{bmatrix}.$$

The image $G_T(\Theta) := \{G_T(\theta) : \theta \in \Theta\}$ is called the *set of attainable gains*.

Using the notation introduced in [4] we formulate the following optimization problem (we write x' for the transposition of the vector x): *For a given $\widehat{c} = [c,0]' \in \mathbb{R}^2$ and any contingent claim $H \in L^2(P) \times L^2(P)$ find*

(P) $$\inf_{g \in G_T(\Theta)} \mathrm{E}[(H - (\widehat{c} + g))'(H - (\widehat{c} + g))].$$

Here we consider a very general model, which describes a so called incomplete market. One of the well known optimization criteria proposed in the literature is quadratic risk minimization, which was considered in general as minimization of global or local risk (see [6] for details). In a discrete time model, the problem was studied e.g. in [7]. There are many papers devoted to this problem, but without taking into account any transaction costs. Historically, the first attempt to incorporate transaction costs was made by Lamberton, Pham and Schweizer [3], but they only considered the local risk minimization. In [4], the existence and uniqueness of an optimal strategy minimizing global risk was proved. One of the technical assumptions sufficient for that result is predictability of the strategy, which practically means that the investor does not know the current price (he knows the previous price) when taking an investment decision. The reason for this assumption was the necessity to move the controlled, predictable part outside the conditional expectation, under fairly weak conditions (MND), (RS) and (RB) imposed on S.

In this paper we relax the predictability condition on θ, making the trading strategy more realistic. Consequently, we are forced to use different arguments to prove the closedness of $G_T(\Theta)$. In order to achieve this objective we require that the price process S satisfies slightly stronger conditions, which we introduce in the next section.

Finally in Section 3 we prove the main theorem of the paper.

The paper is part of a Ph.D. thesis written under the guidance of Prof. Ł. Stettner to whom the author wishes to express his thanks.

2. Additional assumptions. In order to relax the assumption on predictability of a strategy θ, we impose some additional assumptions that ensure closedness of $G_T(\Theta)$. The nondegeneracy condition (MND) (considered in [4]) is now replaced by the following.

DEFINITION 5. The process $S \in L^2(P)$ satisfies the condition:

• (A1) if there exists a constant $M > 0$ such that

$$0 < \frac{1}{M} \leq \frac{\mathrm{E}[\Delta S_k^2 \mathbf{1}(\Delta S_k > 0) \,|\, \mathcal{F}_{k-1}]}{\mathrm{E}[\Delta S_k^2 \mathbf{1}(\Delta S_k < 0) \,|\, \mathcal{F}_{k-1}]} \leq M;$$

- (A2) if there exists a constant $L > 0$ such that

$$0 < \frac{1}{L} \le \frac{\mathrm{E}[\mathbf{1}(\Delta S_k > 0) \,|\, \mathcal{F}_{k-1}]}{\mathrm{E}[\mathbf{1}(\Delta S_k < 0) \,|\, \mathcal{F}_{k-1}]} \le L,$$

for every $k = 2, \ldots, T$.

The two conditions above also play a role of a nondegeneracy property, but are stronger than the (ND) condition (see e.g. [7] or [5]) and stronger than the well known condition (see [1]) that implies the existence of a martingale measure in the model. Therefore we have

COROLLARY 2. *If S satisfies* (A2) *then there exists an equivalent martingale measure for S.*

REMARK 1. Notice that the assumptions allow us to consider S as an element of a quite general class of discrete time processes. In particular if S is modeled by the equation

$$S_k = S_{k-1}(1 + \varrho_k)$$

where the random variables ϱ_k are i.i.d. and moreover almost surely separated from -1 and bounded, then both (A1) and (A2) are satisfied.

Notice that under (RB) and (A1), (A2) we do not need the (RS) condition required in [4].

In what follows we denote by ξ_k the number of stocks at time $k - 1$, i.e. $\xi_k = \sum_{i=1}^{k-1}(l_i - m_i)$. Notice that the process ξ is predictable and $\xi_1 = 0$.

3. Closedness of $G_T(\Theta)$. Before we formulate the main theorem of the paper we introduce an auxiliary lemma.

LEMAT 2. *If S satisfies conditions* (A1), (A2) *and*

$$\left\| \left(\sum_{k=1}^{T} \xi_k \Delta S_k \right)^{-} \right\|_{L^2} \le C_1$$

for a constant $C_1 > 0$, then there exists a constant C_2 such that

$$\left\| \sum_{k=1}^{T} \xi_k \Delta S_k \right\|_{L^2} \le C_2.$$

Proof. We argue by induction on T. Since we put $\xi_1 = 0$, we start the induction at $T = 2$. By condition (A1) for $T = 2$, we have

$$E[\xi_2^2 \Delta S_2^2 \,|\, \mathcal{F}_1] \le (M + 1)\xi_2^2 E[\Delta S_2^2 \mathbf{1}(\Delta S_2 < 0) \,|\, \mathcal{F}_1],$$
$$E[\xi_2^2 \Delta S_2^2 \,|\, \mathcal{F}_1] \le (M + 1)\xi_2^2 E[\Delta S_2^2 \mathbf{1}(\Delta S_2 > 0) \,|\, \mathcal{F}_1].$$

Consequently,

$$E[\xi_2^2 \mathbf{1}(\xi_2 > 0)\Delta S_2^2 \,|\, \mathcal{F}_1] \le (M + 1)E[\xi_2^2 \mathbf{1}(\xi_2 > 0)\Delta S_2^2 \mathbf{1}(\Delta S_2 < 0) \,|\, \mathcal{F}_1]$$

and analogously

$$E[\xi_2^2 \mathbf{1}(\xi_2 < 0)\Delta S_2^2 \mid \mathcal{F}_1] \le (M+1)E[\xi_2^2 \mathbf{1}(\xi_2 < 0)\Delta S_2^2 \mathbf{1}(\Delta S_2 > 0) \mid \mathcal{F}_1].$$

This implies that

$$\|\xi_2 \Delta S_2\|_{L^2}^2 \le (M+1)\|(\xi_2 \Delta S_2)^-\|_{L^2}^2 \le (M+1)C_1^2,$$

which yields the assertion for $T = 2$.

Assume that the assertion holds for $T - 1$. Let $f_{T-1} := \sum_{k=1}^{T-1} \xi_k \Delta S_k$. Then

(1) $\qquad \|(f_{T-1} + \xi_T \Delta S_T)\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T < 0)\mathbf{1}(\Delta S_T > 0)\|_{L^2} \le C_1,$

(2) $\qquad \|(f_{T-1} + \xi_T \Delta S_T)\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T > 0)\mathbf{1}(\Delta S_T < 0)\|_{L^2} \le C_1.$

Since f_{T-1} and $\xi_T \Delta S_T$ are negative on the given sets, we have

(3) $\qquad \|f_{T-1}\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T < 0)\mathbf{1}(\Delta S_T > 0)\|_{L^2} \le C_1,$

(4) $\qquad \|f_{T-1}\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T > 0)\mathbf{1}(\Delta S_T < 0)\|_{L^2} \le C_1,$

(5) $\qquad \|f_{T-1}\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\Delta S_T = 0)\|_{L^2} \le C_1.$

By (3) and (A2), we have

$$\|f_{T-1}\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T < 0)\mathbf{1}(\Delta S_T < 0)\|_{L^2}$$
$$\le \sqrt{L}\,\|f_{T-1}\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T < 0)\mathbf{1}(\Delta S_T > 0)\|_{L^2} \le \sqrt{L}\,C_1.$$

Therefore using (3) and (5) we obtain

$$\|f_{T-1}\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T < 0)\|_{L^2}^2 \le C_1^2(2 + L^2).$$

In an analogous way we prove that

$$\|f_{T-1}\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T > 0)\|_{L^2}^2 \le C_1^2(2 + 1/L^2).$$

Since by the assumption $\|f_{T-1}\mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T = 0)\|_{L^2} \le C_1$, there is a constant C_3 such that

$$\|f_{T-1}\mathbf{1}(f_{T-1} < 0)\|_{L^2} \le C_3.$$

Consequently, by the induction assumption we obtain $\|f_{T-1}\|_{L^2} \le C_4$. Now by (1),

$$\|\xi_T \Delta S_T \mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T < 0)\mathbf{1}(\Delta S_T > 0)\|_{L^2} \le C_1,$$

and by (2),

$$\|\xi_T \Delta S_T \mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T > 0)\mathbf{1}(\Delta S_T < 0)\|_{L^2} \le C_1,$$

Using (A1) we have

$$\|\xi_T \Delta S_T \mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T < 0)\|_{L^2} \le \frac{C_5}{1+M},$$

$$\|\xi_T \Delta S_T \mathbf{1}(f_{T-1} < 0)\mathbf{1}(\xi_T > 0)\|_{L^2} \le \frac{MC_6}{1+M}.$$

Therefore

$$\|\xi_T \Delta S_T \mathbf{1}(f_{T-1} < 0)\|_{L^2}^2 \leq \frac{2(C_5^2 + M^2 C_6^2)}{(1+M)^2}.$$

To complete the proof, it suffices to show that

$$\|\xi_T \Delta S_T \mathbf{1}(f_{T-1} \geq 0)\|_{L^2}^2 \leq C_7$$

for a constant C_7. Observe that

$$\|(f_{T-1} + \xi_T \Delta S_T)\mathbf{1}(f_{T-1} \geq 0)\mathbf{1}(\xi_T \Delta S_T < 0)\|_{L^2}^2$$
$$= \|(f_{T-1} + \xi_T \Delta S_T)\mathbf{1}(f_{T-1} \geq 0)\mathbf{1}(\xi_T \Delta S_T < 0)\mathbf{1}(f_{T-1} + \xi_T \Delta S_T > 0)\|_{L^2}^2$$
$$\quad + \|(f_{T-1} + \xi_T \Delta S_T)\mathbf{1}(f_{T-1} \geq 0)\mathbf{1}(\xi_T \Delta S_T < 0)\mathbf{1}(f_{T-1} + \xi_T \Delta S_T < 0)\|_{L^2}^2$$
$$\leq C_2^2 + C_1^2.$$

Since in the first term f_{T-1} is positive and $\xi_T \Delta S_T < 0$, we have $f_{T-1} + \xi_T \Delta S_T \leq f_{T-1}^+$, while in the second term both f_{T-1} and $\xi_T \Delta S_T$ are negative, and by the assumption $\|(f_{T-1} + \xi_T \Delta S_T)^-\|_{L^2} \leq C_1$.

Therefore, we obtain

$$\|\xi_T \Delta S_T \mathbf{1}(f_{T-1} \geq 0)\mathbf{1}(\xi_T \Delta S_T < 0)\|_{L^2}^2$$
$$\leq 2\|(f_{T-1} + \xi_T \Delta S_T)\mathbf{1}(f_{T-1} \geq 0)\mathbf{1}(\xi_T \Delta S_T < 0)\|_{L^2}^2$$
$$\quad + 2\|f_{T-1}\mathbf{1}(f_{T-1} \geq 0)\mathbf{1}(\xi_T \Delta S_T < 0)\|_{L^2}^2$$
$$\leq 2(C_1^2 + C_2^2) + 2C_2^2.$$

It can be easily seen now that by considerations analogous to the proof for $T = 2$, using (A1) we obtain the existence of constant bounds for the norm of $\xi_T \Delta S_T \mathbf{1}(f_{T-1} \geq 0)\mathbf{1}(\xi_T > 0)$ and $\xi_T \Delta S_T \mathbf{1}(f_{T-1} \geq 0)\mathbf{1}(\xi_T < 0)$. Consequently,

$$\|\xi_T \Delta S_T \mathbf{1}(f_{T-1} \geq 0)\|_{L^2} \leq C_7$$

for a constant C_7 and the proof is complete. ∎

A similar result, but in the $L^1(P)$ norm instead of $L^2(P)$, was achieved in [2, Theorem 2]. The authors imposed slightly weaker assumptions and used different techniques in the proof.

Before we prove the main theorem of the paper we need an auxiliary lemma which is in some sense a generalization of [4, Lemma 4].

LEMMA 5. *Let $(\eta_k^n)_{n=1,2,\dots}$ be a sequence of $\mathcal{F}_k$-measurable random variables such that $\eta_k^n S_k \in L^2(P)$ for $n = 1, 2, \dots$, and let η_k be an $\mathcal{F}_k$-measurable random variable such that $\eta_k S_k \in L^2(P)$ and $\eta_k^n S_k$ is convergent to $\eta_k S_k$ in the weak topology of $L^2(P)$. If the process S satisfies* (RB) *then*

$$\eta_k^n S_{k+1} \to \eta_k S_{k+1}$$

in the weak topology of $L^2(P)$.

Proof. Observe that for any $\mathcal{F}_{k+1}$-measurable $f \in L^2(P)$,

$$\lim_{n \to \infty} \mathrm{E}[(\eta_k^n - \eta_k)S_{k+1}f] = \lim_{n \to \infty} \mathrm{E}\left[(\eta_k^n - \eta_k)S_k \frac{\mathrm{E}[S_{k+1}f \mid \mathcal{F}_k]}{S_k}\right] = 0$$

because by the Cauchy–Schwarz inequality and (RB) condition we have

$$\mathrm{E}\left[\frac{(\mathrm{E}[S_{k+1}f \mid \mathcal{F}_k])^2}{S_k^2}\right] \leq \mathrm{E}\left[\frac{\mathrm{E}[f^2 \mid \mathcal{F}_k]\mathrm{E}[S_{k+1}^2 \mid \mathcal{F}_k]}{S_k^2}\right] \leq C\|f\|_{\mathcal{L}^2(P)}^2$$

and hence by weak convergence of $\eta_k^n S_k$ to $\eta_k S_k$ we obtain weak convergence of $\eta_k^n S_{k+1}$ to $\eta_k S_{k+1}$. ∎

We are in a position to formulate the main result of the paper.

THEOREM 1. *If S satisfies* (RB), (A1) *and* (A2) *then the cone* $G_T(\Theta)$ *is closed.*

Proof. Let θ^n be a sequence of strategies such that $X(\theta^n)$ and $Y(\theta^n)$ converge in $L^2(P)$, i.e.

$$X(\theta^n) = \sum_{k=1}^{T}((1-\mu)m_k^n S_k - (1+\lambda)l_k^n S_k) \xrightarrow{L^2} z_1,$$

$$Y(\theta^n) = \sum_{k=1}^{T}(l_k^n - m_k^n)S_T = \sum_{k=1}^{T}(\xi_k^n \Delta S_k + (l_k^n - m_k^n)S_k) \xrightarrow{L^2} z_2.$$

Then

$$X(\theta^n) + (1-\mu)Y(\theta^n) = (1-\mu)\sum_{k=1}^{T}\xi_k^n \Delta S_k - (\mu+\lambda)\sum_{k=1}^{T}l_k^n S_k,$$

$$X(\theta^n) + (1+\lambda)Y(\theta^n) = (1+\lambda)\sum_{k=1}^{T}\xi_k^n \Delta S_k - (\mu+\lambda)\sum_{k=1}^{T}m_k^n S_k.$$

Since the sequences on the left side are bounded in the $L^2(P)$ norm, by nonnegativity of the strategies m_k^n, l_k^n we have

$$\left\|\left(\sum_{k=1}^{T}\xi_k^n \Delta S_k\right)^{-}\right\|_{L^2} \leq C_1$$

for $n \in \mathbb{N}$ and by Lemma 2,

$$\left\|\sum_{k=1}^{T}\xi_k^n \Delta S_k\right\|_{L^2} \leq C_2.$$

Therefore, there exists a constant K such that

$$\left\|\sum_{k=1}^{T}l_k^n S_k\right\|_{L^2} \leq K, \quad \left\|\sum_{k=1}^{T}m_k^n S_k\right\|_{L^2} \leq K.$$

Since every bounded set in Hilbert space is weakly compact, there is a subsequence n_j for which $m_k^{n_j} S_k \xrightarrow{w} \phi_k$ and $l_k^{n_j} S_k \xrightarrow{w} \psi_k$. Letting $\widehat{m}_k := \phi_k/S_k$ and $\widehat{l}_k := \psi_k/S_k$ we see that $m_k^{n_j} S_k \xrightarrow{w} \widehat{m}_k S_k$ and $l_k^{n_j} S_k \xrightarrow{w} \widehat{l}_k S_k$. Moreover, by the (RB) condition and Lemma 3, we obtain the weak convergence $m_k^{n_j} S_l \xrightarrow{w} \widehat{m}_k S_l$ and $l_k^{n_j} S_l \xrightarrow{w} \widehat{l}_k S_l$ for any $l \geq k$. Then $\xi_l^{n_j} \Delta S_l \xrightarrow{w} \widehat{\xi}_l \Delta S_l$, where $\widehat{\xi}_l = \sum_{k=1}^{l} (\widehat{l}_k - \widehat{m}_k)$, which indicates that the limits z_1 and z_2 may be represented in terms of weak limits and therefore they belong to $G_T(\Theta)$. ∎

REMARK 2. Analyzing the proof we observe that analogous arguments also show the closedness of the cone $G_T(\Theta)$ in the case when it is a subset of $L^2(P)$ instead of $L^2(P) \times L^2(P)$, e.g. $G_T(\Theta) = X_T(\Theta) + Y_T(\Theta)$ or $G_T(\Theta) = X_T(\Theta) + (1-\mu)Y_T(\Theta)$, which can be interpreted as a book value or mark-to-market value. The "one-dimensional" version of the set $G_T(\Theta)$ was considered in [3].

References

[1] R. C. Dalang, A. Morton and W. Willinger, *Equivalent martingale measures and no-arbitrage in stochastic securities market models*, Stochastics Stochastics Rep. 29 (1990), 185–201.
[2] J. Jacod and A. N. Shiryaev, *Local martingales and the fundamental asset pricing theorems in the discrete-time case*, Finance Stoch. 2 (1998), 259–273.
[3] D. Lamberton, H. Pham and M. Schweizer, *Local risk-minimization under transaction costs*, Math. Oper. Res. 23 (1998), 585–612.
[4] M. Motoczyński, *Existence of variance-optimal hedging strategy in discrete time model with transaction costs*, Bull. Polish Acad. Sci. Math. 47 (1999), 191–207.
[5] M. Schäl, *On quadratic cost criteria for option hedging*, Math. Oper. Res. 19 (1994), 121–131.
[6] M. Schweizer, *Hedging of options in a general semimartingale model*, Diss., ETH Zurich, 8615 (1988).
[7] —, *Variance-optimal hedging in discrete time*, Math. Oper. Res. 20 (1995), 1–32.

Risk Management Department
Kredyt Bank S.A.
Kasprzaka 2/8
01-211 Warszawa, Poland
E-mail: mmotoczynski@kredytbank.pl