Optimal Execution Problem with Random Market Impact

Kensuke Ishitani, Takashi Kato
Mitsubishi UFJ Trust Investment Technology Institute Co., Ltd. (MTEC)
2-6, Akasaka 4-Chome, Minato-ku, Tokyo 107-0052 Japan
E-mail: ishitani@mtec-institute.co.jp, kato@mtec-institute.co.jp

Abstract

In [4], we study mathematical formulation of an optimal execution problem in consideration of market impact and some behavior of the corresponding value functions. But there are few studies, including [4], which treat the noise of market impact. In this study we construct a model with random market impact as a generalization of [4]. We consider the case where the noise of market impact in a discrete-time model is given as i.i.d. random variables, and then we derive a continuous-time model as a limit in which the noise is described as a jump of a Lévy process.

1 Introduction

An optimal portfolio management problem is important and fundamental in mathematical finance theory. There are various papers dealing with such a problem and recently more realistic problems are focused, such as liquidity problems. In this paper we pay special attention to market impact (MI), which is the effect of the investment behavior of traders on security prices. MI plays an important role in portfolio theory and is also significant when we consider the case of an optimal execution problem as our basic model and define the corresponding value function. In Section 3, we give our main results. Similar to [4], we show that the discrete-time value functions converge to some function which can be regarded as a value function in the limit continuous-time model under some additional technical assumptions. Moreover we have the continuity property of the continuous-time value function like Theorem 2 of [4]. In Section 4 we treat some examples of our model and present the result of numerical experiments. Section 5 is the conclusion of this paper. Proofs are omitted in this paper but you can refer to [3].

2 The Model

In this section we present the details of the model. Let \((\Omega, \mathcal{F}, P)\) be a complete probability space and let \((B_t)_{0 \leq t \leq T}\) be a standard one-dimensional Brownian motion. \(T > 0\) means a time horizon and we assume \(T = 1\) for brevity. We suppose that the market consists of one risk-free asset (namely cash) and one risky asset (namely a security.) The price of cash is always equal
to 1, which means that a risk-free rate is equal to zero. The price of a security fluctuates according to a certain stochastic flow, and is influenced by sales of a trader.

First we consider the discrete-time model with the time interval $\frac{1}{n}$. We consider a single trader who has an endowment $\Phi_0 > 0$ shares of a security. This trader executes the shares $\Phi_0$ over a time interval $[0, 1]$ considering the effect of MI with noise. We assume that the trader executes only at time $0, \frac{1}{n}, \ldots, \frac{n-1}{n}$ for $n \in \{1, 2, 3, \ldots\}$.

For $l = 0, \ldots, n$, we denote $S^n_l$ the price of the security at time $\frac{l}{n}$ and $X^n_l = \log S^n_l$. Let $s_0 > 0$ be an initial price (i.e. $S^n_0 = s_0$) and $X^n_0 = \log s_0$. If the trader sells the amount $\psi^n_l$ at the time $\frac{l}{n}$, the log-price changes to $X^n_l - g^n_l(\psi^n_l)$ and this execution (selling) gives him/her the amount of cash $\psi^n_l S^n_l \exp(-g^n_l(\psi^n_l))$ as proceeds. Here the random function

$$g^n_l(\psi, \omega) = c^n_l(\omega) g_0(\psi), \quad \psi \in [0, \Phi_0], \omega \in \Omega$$

means MI with noise, which is given by the product of positive random variable $c^n_l(\omega)$ and a deterministic function $g_0 : [0, \Phi_0] \rightarrow [0, \infty)$. The random variable $c^n_l(\omega)$ means the noise of MI and is independent of trading volume $\psi$. The function $g_0(\psi)$ is assumed to be non-decreasing, continuous and differentiable and satisfying $g_0(0) = 0$. Moreover we assume that $(c^n_l)_k$ is i.i.d., so the noise of MI is time-homogeneous. We also assume that $(c^n_l)_k$ is independent of $(B_t)_t$. We remark that if $c^n_l$ is a constant (i.e. $c^n_l \equiv c$ for some $c > 0$) then this setting is the same as in [4].

After the trading at the time $\frac{l}{n}$, $X^{n+1}_l$ and $S^{n+1}_l$ are given by

$$X^{n+1}_l = Y\left(\frac{l+1}{n} - \frac{l}{n} ; X^n_l - g^n_l(\psi^n_l), S^{n+1}_l = e^{X^{n+1}_l}, (1)\right)$$

where $Y(t; \rho, x)$ is the solution of the following stochastic differential equation (SDE) on the filtered space $(\Omega, F, (F^B_t), \mathbb{P})$.

$$\begin{cases} 
\frac{dY(t; \rho, x)}{dt} = \sigma(Y(t; \rho, x))dB_t + b(Y(t; \rho, x))dt, \quad t \geq r, \\
Y(r; \rho, x) = x.
\end{cases}$$

Here $(F^B_t)$ is the Brownian filtration and $b, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ are Borel functions. We assume that $b$ and $\sigma$ are bounded and Lipschitz continuous. Then for each $r \geq 0$ and $x \in \mathbb{R}$ there exists a unique solution.

At the end of the time interval $[0, 1]$, the trader has the amount of cash $W^n_n$ and the amount of the security $\varphi^n_n$, where

$$W^{n+1}_l = W^n_l + \psi^n_l S^n_l e^{-g^n_l(\psi^n_l)}, \quad \varphi^{n+1}_l = \varphi^n_l - \psi^n_l$$

for $l = 0, \ldots, n - 1$ and $W^n_0 = 0, \varphi^n_0 = \Phi_0$. We say that an execution strategy $(\psi^n_l)_{l=0}^{n-1}$ is admissible if $(\psi^n_l)_l \in A^n(\Phi_0)$ holds, where $A^n(\Phi)$ is the set of strategies $(\psi^n_k)_{k=1}^n$ such that $\psi^n_l$ is $\mathcal{F}_{\frac{l}{n}}$-measurable, $\psi^n_l \geq 0$ for each $l = 0, \ldots, k - 1$ and \( k-1 \sum_{i=0}^{k-1} \psi^n_i \leq \varphi$ almost surely.

Then the investor's problem is to choose an admissible strategy to maximize the expected utility $E[u(W^n_n, \varphi^n_n, S^n_n)]$, where $u \in C$ is his/her utility function and $C$ is the set of non-decreasing, non-negative and continuous functions on $D = \mathbb{R} \times [0, \Phi_0] \times [0, \infty)$ such that

$$u(w, \varphi, s) \leq C_u(1 + |w|^{m_u} + s^{m_u}), \quad (w, \varphi, s) \in D \quad (3)$$

for some constants $C_u > 0$ and $m_u > 0$.

For $k = 1, \ldots, n$, $(w, \varphi, s) \in D$ and $u \in C$, we define the (discrete-time) value function $V^n_k(w, \varphi, s; u)$ by

$$V^n_k(w, \varphi, s; u) = \sup_{(\varphi^n_l)_{l=0}^{n-1} \in A^n(\Phi)} E[u(W^n_n, \varphi^n_n, S^n_n)]$$

subject to (1) and (2) for $l = 0, \ldots, k - 1$ and $(W^n_l, \varphi^n_l, S^n_l) = (w, \varphi, s)$. (For $s = 0$, we set $S^n_l \equiv e^{X^n_l}$.) We denote $V^n_0(w, \varphi, s; u) = u(w, \varphi, s)$. Then our problem is the same as $V^n_n(0, \Phi_0, \Phi_0; u)$. We consider the limit of the value function $V^n_k(w, \varphi, s; u)$ as $n \rightarrow \infty$.

Let $h : [0, \infty) \rightarrow [0, \infty)$ be a non-decreasing continuous function. We introduce the following condition for $g_0(\psi)$.

$$[A] \lim_{n \rightarrow \infty} \sup_{\psi \in [0, \Phi_0]} \frac{d}{d\psi} g_0(\psi) - h(np) = 0.$$

Moreover we assume the following conditions for $(c^n_k)$.\[[B1]\]

For any $n \in \mathbb{N}$ and $x \geq 0$ it holds that $\gamma_n > 0$ and

$$h(x/\gamma_n) \rightarrow 0, \quad n \rightarrow \infty,$$

where $\gamma_n = \text{essinf} c^n_k$.\[[B2]\]

Let $\mu_n$ be the distribution of $\frac{c^n_0 + \ldots + c^n_{n-1}}{n}$ Then $\mu_n$ has a weak limit $\mu$ as $n \rightarrow \infty$.\[[B3]\]

There is a sequence of infinitely divisible distributions $(p_n)_n$ on $\mathbb{R}$ such that $\mu_n = \mu \ast p_n$ and either

- $[B3a]$ \( \int_{\mathbb{R}} x^2 p_n(dx) = O(1/n^3) \) as $n \rightarrow \infty$
- $[B3b]$ There is a sequence $(K_n)_n \subset (0, \infty)$ such that $K_n = O(1/n)$, $p_n((\infty, \infty) - K_n) = 0$ (or $p_n((K_n, \infty)) = 0$) and

$$\int_{\mathbb{R}} x^2 p_n(dx) = O(1/n) \quad \text{as} \quad n \rightarrow \infty,$$

where $O$ denotes order notation (Landau's symbol.)

Let us give some remarks for condition $[B1]$. Since $c^n_k, k = 0, 1, 2, \ldots$, are identically distributed, $\gamma_n$ is independent of $k$. Moreover, if $h(\infty) < \infty$, then (4) is always fulfilled. If $h(\infty) = \infty$, we have the following example:

$$h(\zeta) = \alpha \zeta^p, \quad \gamma_n = \frac{1}{n^{1/p - \delta}} \quad (p, \delta > 0, \delta < 1/p).$$

Since $\mu$ is an infinitely divisible distribution, there is some Lévy process $(L_t)_{0 \leq t \leq 1}$ on a certain probability
space such that \( L_t \) is distributed as \( \mu \). Without loss of generality, we may assume that \((L_t), (B_t)\) are defined on the same filtered space and are independent. Let \( \nu \) be a Lévy measure of \((L_t)\). We assume the following moment condition for \( \nu \).

\[
\mathbb{E}[z + z^2 \nu(dz)] < \infty.
\]

Now we define the function which gives the limit of the discrete-time value functions. For \( t \in [0, 1] \) and \( \varphi \in \mathbb{R}^d \) we denote by \( A_t(\varphi) \) the set of \((F_r)_{0 \leq r \leq t}\)-adapted and càglàd process (i.e. left continuous and has a right limit at each point) \((\zeta(t))_{0 \leq t \leq 1}\) such that \( \zeta_t \geq 0 \) for each \( t \in [0, t] \), \( \int_0^t \zeta_r \, dr \leq \varphi \) almost surely and \( \sup \zeta_r(w) < \infty \). For \( t \in [0, 1], (w, \varphi, s) \in D \) and \( u \in \mathcal{C} \), we define \( V_t(w, \varphi, s; u) \) by

\[
V_t(w, \varphi, s; u) = \sup_{(\sigma, b) \in A_t(\varphi)} \mathbb{E}[u(W_t, \varphi_t, S_t)]
\]

subject to

\[
\begin{align*}
dW_r &= \zeta_r S_r \, dr - \zeta_r \, dW_r, \\
\frac{dS_r}{S_r} &= \delta(r) \, dr + b(S_r) \, dW_r - g(\zeta_r) \, S_r \, dL_r
\end{align*}
\]

and \((W_0, \varphi_0, S_0) = (w, \varphi, s)\), where

\[
\begin{align*}
\delta(s) &= s \sigma(\log s), \\
b(s) &= s \left( \frac{1}{2} \sigma^2(\log s)^2 \right)
\end{align*}
\]

for \( s > 0 \) \((\delta(0) = b(0) = 0)\) and

\[
g(\zeta) = \int_0^\infty h(\zeta') \, d\zeta'.
\]

When \( s > 0 \), we obviously see that the process of the log-price of the security \( X_r = \log S_r \) satisfies

\[
\begin{align*}
dX_r &= \sigma(X_r) \, dB_r + b(X_r) \, dW_r - g(\zeta_r) \, dL_r.
\end{align*}
\]

## 3 Main Results

In this section we present main results. First we give the convergence theorem for value functions.

**Theorem 1** ([3]) For each \((w, \varphi, s) \in D\), \( t \in [0, 1] \) and \( u \in \mathcal{C} \) it holds that

\[
\lim_{n \to \infty} V_{n}^u(w, \varphi, s; u) = V_t(w, \varphi, s; u),
\]

where \([nt]\) is the greatest integer less than / equal to \( nt \).

By this theorem, we see that the function \( V_t(w, \varphi, s; u) \) corresponds to the continuous-time model of an optimal execution problem with random MI. This result is almost the same as in [4], but we notice that the term of MI is given as an increment \( g(\zeta_r) \, dL_r \).

\[
L_t = \gamma t + \int_0^t \int_{(0, \infty)} z N(dr, dz) \quad \text{be the Lévy decomposition of} \quad (L_t)_t.
\]

Then \( g(\zeta_r) \, dL_r \) can be divided into the following two terms:

\[
g(\zeta_r) \, dL_r = \gamma g(\zeta_r) \, dL_r + g(\zeta_r) \int_0^t \int_{(0, \infty)} z N(dr, dz).
\]

The last term in the right-hand side refers to the effect of the noise of MI. This means that the noise of MI appears as a jump of Lévy process.

As for continuity of the continuous-time value function, we have the following.

**Theorem 2** ([3]) Let \( u \in \mathcal{C} \).

(i) If \( h(\infty) = \infty \), then \( V_t(w, \varphi, s; u) \) is continuous in \((t, w, \varphi, s) \in [0, 1] \times D\).

(ii) If \( h(\infty) < \infty \), then \( V_t(w, \varphi, s; u) \) is continuous in \((t, w, \varphi, s) \in [0, 1] \times D\) and \( V_t(w, \varphi, s; u) \) converges to \( J_t(w, \varphi, s; u) \) uniformly on any compact subset of \( D \) as \( t \to 0 \), where \( J_t(w, \varphi, s) \) is given as

\[
\sup_{u \in \mathcal{U}} u(w + s \varphi - \psi, \varphi - \psi, \psi) \quad \text{as} \quad \psi \to 0.
\]

This is also quite similar to [4]. The continuity in \( t \) at the origin is according to the state of the function \( h \) at the infinity point. When \( h(\infty) < \infty \), the value function is not always continuous at \( t = 0 \) and has the right limit \( J_t(w, \varphi, s) \). The function \( J_t(w, \varphi, s) \) indicates the utility of the profit of the execution of a trader who sells a part of the shares of a security \( \varphi \) by dividing infinitely in infinitely short time (enough to neglect the fluctuation of the price of a security) and makes the amount \( \varphi - \psi \) remain.

We pay attention to the fact that the noise part \( g(\zeta_r) \int_0^t \int_{(0, \infty)} z N(dr, dz) \) makes no change for the result.

We should notice that if \( \gamma = 0 \) and \( h(\infty) < \infty \), then the effect of MI disappears in \( J_t(w, \varphi, s) \).

### 4 Examples

In this section we show two examples of our model, both of which are generalization of the ones in [4].

Let \( b(x) \equiv -\mu \) and \( \sigma(x) \equiv \sigma \) for some constants \( \mu, \sigma \geq 0 \) and suppose \( \bar{\mu} = \mu - \sigma^2/2 > 0 \). We assume that a trader has a risk-neutral utility function \( u(w, \varphi, s) = w \).

For the noise part of MI, we consider the Gamma distribution.

\[
P(c_t^\gamma - \gamma \in dx) = \text{Gamma}(\alpha_1/n, n\beta_1)(dx)
\]

\[
= \frac{1}{\Gamma(\alpha_1/n)(n\beta_1)^{\alpha_1/n+1}} e^{-x/(n\beta_1)} 1_{(0, \infty)}(x) dx,
\]
where $\Gamma(x)$ is the Gamma function. Here $\alpha_1, \beta_1, \gamma > 0$ are constants.

For the deterministic part of MI, we consider two patterns: (log-)linear case and (log-)quadratic case i.e. $g_n(\psi) = \alpha_0 \psi$ or $g_n(\psi) = n\alpha_0 \psi^2$ for $\alpha_0 > 0$. In each case the assumption [A] is satisfied, and so are [B1] - [B3] and [C].

4.1 Log-Linear Impact & Gamma Distribution

In this case we have the following theorem.

Theorem 3 ([3]) It holds that

$$V_t(w, \varphi, s; u) = w + \frac{1 - e^{-\gamma \alpha_0 \varphi}}{\gamma \alpha_0} s \tag{5}$$

for each $t \in (0, 1]$ and $(w, \varphi, s) \in D$.

An implication of this result is the same as in [4]: the right-hand side of (5) is equal to $J_t(w, \varphi, s)$ and converges to $w + \varphi s$ as $\alpha_0 \downarrow 0$ or $\gamma \downarrow 0$, which is the profit gained by choosing the execution strategy of so-called block liquidation such that a trader sells all shares $\varphi$ at $t = 0$ when there is no MI. So the optimal strategy in this case is to execute all shares dividing infinitely in infinitely short time at $t = 0$ (we call such a strategy an almost block liquidation at the initial time.) We see that the effect of (pure) noise part of MI does not influence the result. That is, the value of $V_t(w, \varphi, s; u)$ and the corresponding optimal strategy do not vary with respect to $\alpha_1$ and $\beta_1$.

4.2 Log-Quadratic Impact & Gamma Distribution

In [4], we partially get an analytical solution of the problem: when $\varphi$ is small enough or large enough, we have an explicit form of optimal strategies. However, the noise of MI makes the problem complicated and hard to solve, so deriving explicit solutions is more difficult. Thus we rely on numerical experiments. Thanks to the assumption that a trader is risk-neutral, we see that an optimal strategy is deterministic. We assume the following additional condition.

$$[D] \quad \gamma \geq \frac{\alpha_1 \beta_1}{8}.$$ 

Then we can replace our optimization problem with the deterministic control problem

$$f(t, \varphi) = \sup_{\zeta(t)} \int_0^t \exp \left( - \int_0^r q(\zeta_s)ds \right) \zeta \, dr,\]$$

$$q(\zeta) = \mu + \gamma \alpha_0 \zeta^2 + \alpha_1 \log(\alpha_0 \beta_1 \zeta^2 + 1)$$

for deterministic process $(\zeta_r)$. Indeed we have the following theorem.

Theorem 4 ([3]) $V_t(w, \varphi, s; u) = w + s f(t, \varphi)$ under [D].

So we can solve this example numerically by considering the deterministic control problem in the discrete-time model for large enough $n$. We set each parameters as follows: $\alpha_0 = 0.01$, $t = 1$, $\mu = 0.05$, $\beta_1 = 2$, $\gamma = 1$, $w = 0$, $s = 1$ and $n = 100$. For value of $\alpha_1$ and $\varphi$, we examine the following patterns: $\alpha_1 = 0, 1, 3$ and $\varphi = 1, 10, 100$. This is because forms of optimal strategies vary according to these values.

4.2.1 The case of $\varphi = 1$

If a trader has a few amount of the security holdings, the result is given as in Figure 1. In fact, if MI function has no noise, that is $\alpha_1 = 0$, then the optimal strategy is "to sell up all the amount at the same speed." (We remark that the roundness at the corner in the top graph of Figure 1 expresses the discritization error and this is no essential.) The same tendency is found in the case of $\alpha_1 = 10$, but in this case the execution time is longer than the case of $\alpha_1 = 0$. When we consider the case of $\alpha_1 = 3$, the situation undergoes a complete change. In this case the optimal strategy is to increase the execution speed as the time horizon comes near.

4.2.2 The case of $\varphi = 10$

When the amount of the security holdings is little larger than the case of Section 4.2.1, then we have the result of Figure 2. In this case a trader's optimal strategy is to increase the execution speed as the trading time comes to end like the case of Section 4.2.1 with $\alpha_1 = 3$. We see that the larger the value of $\alpha_1$, the higher the speed of execution near the time horizon.
Fig. 2: The result with \( \varphi = 10 \). The top graph: An optimal strategy \( \xi_r \). The bottom graph: an amount of the security holdings \( \varphi_r \). The horizontal axis means time \( r \).

We add that a trader cannot finish the liquidation when \( \alpha_1 = 3 \).

4.2.3 The case of \( \varphi = 100 \)

When the amount of the security holdings is too large, a trader cannot finish the liquidation regardless of the value of \( \alpha_1 \), as Figure 3 implies. Remaining is similar to the case of Section 4.2.2 with \( \alpha_1 = 3 \). The rest shares of the security at the time horizon is larger as the noise of MI becomes larger.

5 Concluding Remarks

In this paper we generalize the framework of [4] and consider an optimal execution problem with random MI. We define the (one-shot) MI function as a product of i.i.d. positive random variable and a deterministic function in discrete-time model. We derive the continuous-time model of an optimization problem as the limit of discrete-time models, and find out that the noise of MI in continuous-time model is described as a jump of a Lévy process.

It is interesting that our main results discussed in Section 3 are almost the same as in [4] and our numerical experiments suggest that the bigger the noise of MI, the longer we need liquidation time.

For getting a deeper insight, we should investigate the structure of the MI function punctually. In Theorem 2 (ii) and Theorem 3, the important parameter is \( \gamma \), which is the infimum of \( L_1 \) and is smaller than (or equal to) \( \mathbb{E}[L_1] \). We can give the interpretation that the (almost) block liquidation annihilates the effect of positive jumps of \( (L_t)_t \). We have another decomposition of \( L_t \) such that

\[
L_t = \bar{\gamma} t + \int_0^t \int_{(0, \infty)} z \tilde{N}(dr, dz),
\]

where \( \bar{\gamma} = \gamma + \int_{(0, \infty)} z \nu(dz) \) and \( \tilde{N}(dr, dz) = N(dr, dz) - \int_{(0, \infty)} z \nu(dz) dr \). This representation is essential from the view of martingale theory and the value of \( \bar{\gamma} \) is also important. Considering the meaning of this value may present significant findings.

In our settings, MI function is stationary in time, but in the real market, characteristics of MI is found to change according to the time zone. So it is meaningful to study the case where MI function is inhomogeneous in time. It is one of our further developments.

A characterization of the continuous-time value function as a viscosity solution of the corresponding Hamilton-Jacobi-Bellman equation like [4] is also a remaining task and significant in mathematical finance. It may be difficult to consider viscosity solutions corresponding to a stochastic control problem with jumps, but we surmise that it is possible if we give a suitable technical assumptions.

References


