日本オペレーションズ・リサーチ学会
2022年 春季研究発表会

Execution game in a Markovian environment

01007984 大阪大学 大学院経済学研究科 大西 匡光 OHNISHI Masamitsu
数理・データ科学教育研究センター
05001073 東京都立大学 大学院経営学研究科 *下清水 慎 SHIMOSHIMIZU Makoto

1. Introduction

In a discrete time framework $t \in \{1, \dots, T, T+1\}$ $(, T \in \mathbb{Z}_{++} := \{1, 2, \dots\})$, we assume that two large traders, denoted by $i \in \{1, 2\}$, purchase one risky asset in a trading market. Each large trader has a CARA vN-M (or negative exponential) utility function with the absolute risk aversion parameter $\gamma^i > 0$.

2. Model

Each large trader must purchase $\mathfrak{Q}^i (\in \mathbb{R})$ volume of one risky asset by the time $T+1$. $q_t^i (\in \mathbb{R})$ stands for large amount of orders of submitted by the large trader $i \in \{1, 2\}$ at time $t \in \{1, \dots, T\}$. We denote by $\overline{Q}_t^i$ the number of shares remained to purchase by the large trader at time $t \in \{1, \dots, T, T+1\}$. So

$$\overline{Q}_{t+1}^i = \overline{Q}_t^i - q_t^i, \tag{1}$$

with the initial and terminal conditions: $\overline{Q}_1^i = \mathfrak{Q}^i \in \mathbb{R}$; $\overline{Q}_{T+1}^i = 0 \in \mathbb{R}$, for each large trader $i \in \{1, 2\}$.

The market price of the risky asset at time $t \in \{1, \dots, T, T+1\}$ is P_t. Then, the execution price of the asset becomes $\widehat{P}_t$. We assume that submitting one unit of (large) order at time $t \in \{1, \dots, T\}$ causes the instantaneous price impact denoted as $\lambda_t (> 0)$.

We subsequently define the residual effect of past price impact caused by both large traders at time $t \in \{1, \dots, T\}$, represented by $R_t \in \mathbb{R}$. The dynamics of R_t for $t \in \{1, \dots, T\}$ is defined as:

$$R_{t+1} = \mathrm{e}^{-\rho} \left[R_t + \alpha_t \lambda_t \left(q_t^1 + q_t^2 \right) \right], \tag{2}$$

where $\alpha_t \in [0, 1]$ represents the linear price impact coefficients representing the temporary price impacts.

Furthermore, we define a sequence of independent random variables ϵ_t at time $t \in \{1, \dots, T\}$ as the effect of the public news/information about the economic situation between t and $t+1$. ϵ_t for $t \in \{1, \dots, T\}$ are assumed to follow a normal distribution with mean $\mu_t^\epsilon \in \mathbb{R}$ and variance $(\sigma_t^\epsilon)^2 \in \mathbb{R}_{++}$,

$$\epsilon_t \sim N\left(\mu_t^\epsilon, (\sigma_t^\epsilon)^2 \right), \quad t = 1, \dots, T. \tag{3}$$

A *Markovian environment*, denoted by $\mathcal{I}_t$, influences the fundamental price. The distribution of $\mathcal{I}_t$ is assumed to have a Markovian dependence:

$$\mathcal{I}_{t+1}|_{\mathcal{I}_t} \sim N\left(a_{t+1}^{\mathcal{I}} - b_{t+1}^{\mathcal{I}} \mathcal{I}_t, \left(\sigma_{t+1}^{\mathcal{I}} \right)^2 \right). \tag{4}$$

Assumption 2.1. $\mathcal{I}_t$ and ϵ_t are *correlated* for each $t \in \{1, \dots, T\}$. In addition, no other sequential dependence between two stochastic sequences exists.

By definition of ϵ_t, we define the dynamics of the fundamental price $P_t^f := P_t - R_t$ with a Markovian environment and the permanent price impact for time $t \in \{1, \dots, T\}$ as follows:

$$P_{t+1}^f := P_t^f + \beta_t \lambda_t \left(q_t^1 + q_t^2 \right) + \mathcal{I}_t + \epsilon_t. \tag{5}$$

According to Eq. (2) and (5), the dynamics of market price for $t \in \{1, \dots, T\}$ are described as

$$P_{t+1} = P_t - (1 - \mathrm{e}^{-\rho}) R_t + \left(\alpha_t \mathrm{e}^{-\rho} + \beta_t \right) \lambda_t \left(q_t^1 + q_t^2 \right) + \mathcal{I}_t + \epsilon_t. \tag{6}$$

The wealth process for each large trader $i \in \{1, 2\}$, W_t^i $(\in \mathbb{R})$, for $t \in \{1, \dots, T\}$ evolves as follows:

$$W_{t+1}^i = W_t^i - \widehat{P}_t q_t^i = W_t^i - \left\{ P_t + \lambda_t \left(q_t^1 + q_t^2 \right) \right\} q_t^i. \tag{7}$$

3. Formulation as a Markov game

We define the state of the decision process at time $t \in \{1, \dots, T, T+1\}$ as 7–tuple and denote it as

$$\boldsymbol{s}_t = \left(W_t^1, W_t^2, P_t, \overline{Q}_t^1, \overline{Q}_t^2, R_t, \mathcal{I}_{t-1} \right) \in \mathbb{R}^7 =: S. \tag{8}$$

For $t \in \{1, \dots, T\}$, an allowable action chosen at state $\boldsymbol{s}_t$ is an execution volume $q_t^i \in \mathbb{R} =: A^i$ so that the set A^i of admissible actions is independent of the current state $\boldsymbol{s}_t$. When an action q_t^i is chosen in a state $\boldsymbol{s}_t$ at time $t \in \{1, \dots, T\}$, a transition to a next state: $\boldsymbol{s}_{t+1} = (W_{t+1}^1, W_{t+1}^2, P_{t+1}, \overline{Q}_{t+1}^1, \overline{Q}_{t+1}^2, R_{t+1}, \mathcal{I}_t) \in S$ occurs according to the law of motion which we have precisely described in the previous subsection.

A utility payoff (or reward) arises only in a terminal state s_{T+1} at the end of horizon $T+1$ as

$$g^i_{T+1}(\boldsymbol{s}_{T+1}) := \begin{cases} -\exp\{-\gamma^i W^i_{T+1}\} & \text{if } \overline{Q}^i_{T+1} = 0; \\ -\infty & \text{if } \overline{Q}^i_{T+1} \neq 0, \end{cases} \quad (9)$$

where $\gamma^i > 0$ represents the risk aversion parameter for the large trader $i \in \{1, 2\}$. The term $-\infty$ means a hard constraint enforcing the large trader to execute all of the remaining volume $\overline{Q}^i_T$ at the maturity T. The types of large traders could be defined by $(W^i, \mathfrak{Q}^i, \gamma^i),\ i \in \{1, 2\}$, and these are assumed to be their common knowledge.

If we define a (history–independent) one–stage decision rule f_t at time $t \in \{1, \ldots, T\}$ by a Borel measurable map from a state $\boldsymbol{s}_t \in S = \mathbb{R}^7$ to an action

$$q^i_t = f^i_t(\boldsymbol{s}_t) \in A = \mathbb{R}, \quad (10)$$

then a Markov execution strategy π is defined as a sequence of one–stage decision rules

$$\pi^i := (f^i_1, \ldots, f^i_t, \ldots, f^i_T). \quad (11)$$

We denote the set of all Markov execution strategies as Π_{M}. Further, for $t \in \{1, \ldots, T\}$, we define the sub–execution strategy after time t of a Markov execution strategy $\pi \in \Pi_{\mathrm{M}}$ as

$$\pi^i_t := (f^i_t, \ldots, f^i_T), \quad (12)$$

and the entire set of π^i_t as $\Pi^i_{\mathrm{M},t}$.

By definition (9), the value function under an execution strategy profile $(\pi^1.\pi^2)$ becomes an expected utility payoff arising from the terminal wealth W_{T+1}:

$$V^i_1(\pi^1, \pi^2)\big[\boldsymbol{s}_1\big] = \mathbb{E}^{(\pi^1,\pi^2)}_1\Big[g^i_{T+1}(\boldsymbol{s}_{T+1})\Big|\boldsymbol{s}_1\Big], \quad (13)$$

where, for $t \in \{1, \ldots, T\}$, $\mathbb{E}^{(\pi^1,\pi^2)}_t$ is a conditional expectation given a condition at time t under the strategy profile (π^{1*}, π^{2*}).

For $t \in \{1, \ldots, T, T+1\}$ and $\boldsymbol{s}_t \in S$, we let

$$V^i_t(\pi^1_t, \pi^2_t)\big[\boldsymbol{s}_t\big] = \mathbb{E}^{(\pi^1_t,\pi^2_t)}_t\Big[g^i_{T+1}(\boldsymbol{s}_{T+1})\Big|\boldsymbol{s}_t\Big], \quad (14)$$

be the expected utility payoff at time t.

Definition 3.1 (Nash Equilibrium). $(\pi^{1*}, \pi^{2*}) \in \Pi^1_{\mathrm{M}} \times \Pi^2_{\mathrm{M}}$ is a Nash equilibrium starting from a fixed initial state $\boldsymbol{s}_1$ if and only if

$$V^1_1(\pi^{1*}, \pi^{2*})\big[\boldsymbol{s}_1\big] \geq V^1_1(\pi^1, \pi^{2*})\big[\boldsymbol{s}_1\big], \quad \forall \pi^1 \in \Pi^1_{\mathrm{M}}; \quad (15)$$

$$V^2_1(\pi^{1*}, \pi^{2*})\big[\boldsymbol{s}_1\big] \geq V^2_1(\pi^{1*}, \pi^2)\big[\boldsymbol{s}_1\big], \quad \forall \pi^2 \in \Pi^2_{\mathrm{M}}. \quad (16)$$

Definition 3.2 (Markov Perfect Equilibrium). $(\pi^{1*}, \pi^{2*}) \in \Pi^1_{\mathrm{M}} \times \Pi^2_{\mathrm{M}}$ is a Markov perfect equilibrium if and only if for all $t \in \{1, \ldots, T\}$ and for all $\boldsymbol{s}_t \in S$,

$$V^1_t(\pi^{1*}_t, \pi^{2*}_t)\big[\boldsymbol{s}_t\big] \geq V^1_t(\pi^1_t, \pi^{2*}_t)\big[\boldsymbol{s}_t\big], \quad \forall \pi^1_t \in \Pi^1_{\mathrm{M},t}, \quad (17)$$

$$V^2_t(\pi^{1*}_t, \pi^{2*}_t)\big[\boldsymbol{s}_t\big] \geq V^1_t(\pi^{1*}_t, \pi^2_t)\big[\boldsymbol{s}_t\big], \quad \forall \pi^2_t \in \Pi^2_{\mathrm{M},t}. \quad (18)$$

Based on the One Stage [Step, Shot] Deviation Principle, we obtain an equilibrium execution strategy at a Markov perfect equilibrium by backward induction procedure of dynamic programming.

4. Equilibrium execution strategy

Theorem 4.1. There exists a Markov perfect equilibrium at which the following properties hold for each large trader $i \in \{1, 2\}$:

1. The execution volume at the Markov perfect equilibrium for the large trader $i \in \{1, 2\}$ at time $t \in \{1, \ldots, T\}$, denoted as q^{i*}_t, becomes an *affine function* of $(\overline{Q}^i_t, \overline{Q}^j_t, \mathcal{I}_{t-1})$:

$$q^{i*}_t = a^i_t + b^i_t\overline{Q}^i_t + c^i_t\overline{Q}^j_t + d^i_t R_t + e^i_t\mathcal{I}_{t-1}. \quad (19)$$

2. The value function $V^i_t(\pi^1, \pi^2)\big[\boldsymbol{s}_t\big]$ at time $t \in \{1, \ldots, T, T+1\}$ for each large trader $i \in \{1, 2\}$ is represented as a functional form:

$$\begin{aligned} V^i_t(\pi^1_t, \pi^2_t)\big[\boldsymbol{s}_t\big] = &-\exp\Big\{-\gamma\Big[W^i_t - P^\top_t\overline{Q}^i_t + G^{1i}_t\left(\overline{Q}^i_t\right)^2 \\ &+ G^{2i}_t\overline{Q}^i_t + H^{1i}_t\overline{Q}_t R_t + H^{2i}_t R^2_t + H^{3i}_t R_t + I^{1i}_t\overline{Q}^i_t\overline{Q}^j_t \\ &+ I^{2i}_t\overline{Q}^j_t R_t + I^{3i}_T\left(\overline{Q}^j_t\right)^2 + I^{4i}_t\overline{Q}^j_t + J^{1i}_t\overline{Q}^i_t\mathcal{I}_{t-1} + J^{2i}_t R_t\mathcal{I}_{t-1} \\ &+ J^{3i}_t\overline{Q}^j_t\mathcal{I}_{t-1} + J^{4i}_t\mathcal{I}^2_{t-1} + J^{5i}_t\mathcal{I}_{t-1} + Z^i_t\Big]\Big\}. \end{aligned} \quad (20)$$

The dynamics of $a^i_t, b^i_t, c^i_t, d^i_t, e^i_t;\ G^{1i}_t, G^{2i}_t, H^{1i}_t, H^{2i}_t, H^{3i}_t, I^{1i}_t, I^{2i}_t, I^{3i}_t, I^{4i}_t, J^{1i}_t, J^{2i}_t, J^{3i}_t, J^{4i}_t, J^{5i}_t, Z^i_t$ for $t \in \{1, \ldots, T, T+1\}$ are deterministic functions of time t which are dependent on the problem parameters, and can be computed backwardly in time t from maturity T.

参考文献

[1] Lehalle, C. A., & Neuman, E. (2019). Incorporating signals into optimal trading. *Finance and Stochastics*, **23**(2), 275–311.