description of strategies used in the future paper

Martin Šmíd, Peter Hron, Gabriela Suchopárová, Karel Vrbenský

June 21, 2022

Market Maker

Consider a Market Maker, i.e. an agent who is obliged to quote a bid b_t and an ask a_t, $a_t, b_t \in \mathbb{N}$, $b_t < a_t$. The time interval between actions can be chosen, starting from some technical limit τ_0.

We assume that, at any time, the MM can consume certain amount of cash, preferring today's consumption, hence maximizing

$$\sum_{\tau=i}^{\infty} \gamma^{-\tau_i} c_\tau \tag{1}$$

where $\gamma < 1$ is a discount factor very close to one and $c_\tau \in \{0, \mu\}$. In addition to c_t, the actions at t include setting the quotes $b_t < a_t$ so that the amount of money and stocks including those blocked by the market to back limit order is non-negative.

The maximalistic DP

Denoting m_t and n_t the amount of money, stocks, respectively, and $a_t > \beta_t$, the other's best quotes, the action space is

$$A_t = A(m_t, n_t, \beta_t, \alpha_t) = \{(c, b, a, w, v, \Delta\tau) \in \mathbb{N}_0 \times \mathbb{N}_0 \times \mathbb{N}_0 \times \mathbb{N}_0 \times \mathbb{N}_\infty \times \mathbb{R}_+ : a_t > b_t, a_t > \beta_t, b_t < \alpha_t, m_t - w b_t - c_t \geq 0, n_t - v \geq 0, n_t = 0 \Leftrightarrow a_t = \infty, m_t = 0 \Leftrightarrow b_t = 0\} \tag{2}$$

where quoting $b = 0$ or $a = \infty$ means not quoting (yet the MM is obliged to quote, it may happen that he cannot quote b or a due to the lack of money, stocks, respectively). Here,

β is the best buy limit order of the other participants (if there is any) or own previous value of b (if some was set) or the initial price minus one

α is the best sell limit order of the other participants (if there is any) or own previous value of a (if some was set) or the initial price plus one

In addition to m_t and n_t, the state space includes the state of the market which we denote by R_t, i.e.

$$S = \mathbb{N}_0 \times \mathbb{N}_0 \times \mathcal{R} \tag{3}$$

where $\mathcal{R}$ is the state space containing the complete history. The transition function is

$$m_t = m_{t-1} - c_{t-1} - D_t b_{t-1} + C_t a_{t-1} \tag{4}$$

$$n_t = n_{t-1} + D_t - C_t \tag{5}$$

$$R_t = F(r_{t-1}, a_{t-1}, b_{t-1}) \tag{6}$$

where $D_t \in \mathbb{N}_0$,$C_t \in \mathbb{N}_0$ are random variables indicating the actual volume of a buy transaction, sell transaction, respectively, by the MM, and F is a very complex random function.

The Bellman equation is then

$$V(m,n,r) = \max_{(c,b,a,w,v,\Delta\tau)\in A(m,n)} [c+\gamma^{-\Delta\tau}\mathbb{E}_{C,D,R} V(m-c-Db+Ca, n+D-C, R)] \tag{7}$$

where A is given by (2).

Note that the distribution of C, D depends on the quotes a and b (quoting near or inside the spread clearly increases the probability of transaction) and their volumes w and v. Specially, given no quoting a or b, it is $C \equiv 0$, $D \equiv 0$, respectively.

First note that:

1. $V(0,0,r) = 0$ (no possibility to earn anything).

2. $V(m,n,r)$ is increasing in m with $\Delta V_m \geq 1$ (once a strategy u is optimal for $V(m_1,n,r)$ and $m_2 > m_1$, the strategy to consume $c := m_2 - m_1$ and use u brings value $c + V(m_1,n,r)$).

3. $V(m,n,r)$ is non-decreasing in n, provided that the number of stocks held is not part of r (once a strategy u is optimal for $V(m,n_1,r)$ and $n_2 > n_1$, the strategy u is feasible for the new problem as we just keep $n_2 - n_1$ stocks forever; hence $V(m,n_2,R) \geq V(m,n_2,r)$)

Clearly, the problem suffers from all the three curses of dimensionality.

ADP

As a first approach, assume that the time interval $\Delta\tau$ between the decision points is constant, without loss of generality equal to one.

Restriction of A

To restrict dimensionality, we will allow to consume only zero or $\omega \in \mathbb{N}$ and to quote only zero or $\eta \in \mathbb{N}$. Clearly, quoting far from the spread leads to a negligible probability of transaction and the cross-quoting leads to the same results regardless of the quoted price, so we can restrict the action space to

$$
\begin{aligned}
& A(m,n,\beta,\alpha) \\
&= \{(c,w,v,b,a) \in \{0,\omega\}\times\{0,\eta\}^2\times\{\alpha-\Delta,\ldots,\alpha+\delta\}\cup\{0\}\times\{\beta-\delta,\ldots\beta+\Delta\}\cup\{\infty\} : \\
&\qquad b<a, a>\beta, b<\alpha, m-bw-c\geq 0, n-v\geq 0, \\
&\qquad\qquad m=0 \Leftrightarrow b=0, n=0 \Leftrightarrow a=\infty\}.
\end{aligned}
$$

Here $\delta,\Delta \in \mathbb{N}$ are given constants. It is however suitable to reformulate

$$
\begin{aligned}
& A(m,n,\beta,\alpha) = \{(c,w,v,\Delta b,\Delta a) \\
&\qquad \in \{0,\omega\}\times\{0,\eta\}^2\times\{-\Delta,\ldots,\delta\}\cup\{\square\}\times\{-\delta,\ldots\Delta\}\cup\{\square\} : \\
&\beta+\Delta b<a, a>\beta, b<\alpha, m-bw-c\geq 0, n-v\geq 0, m=0 \Leftrightarrow b=0, n=0 \Leftrightarrow a=\infty\}.
\end{aligned}
$$

where $\Delta a = a-\alpha$, $\Delta b = b-\beta$, $\Delta b = \square \Leftrightarrow b=0$, $\Delta a = \square \Leftrightarrow a=\infty$.

Restriction of S

As it is clear from the definition, the MM's objective depends directly only on process $c_t, b_t, w_t, D_t, a_t, v_t, C_t$. However, C and D depend on the state of the market, which is very complex in reality. In our opinion, a minimal reasonable market state space $\mathcal{R}$ should contain the the bid β and ask α save own quotes, i.e. $\mathcal{R} = \{\alpha \in \mathbb{N}_\infty, \beta \in \mathbb{N}_0 : \alpha > \beta\}$

Formulation of the problem

The Bellman equation is

$$
V(m,n,\beta,\alpha) = \max_{(c,\Delta b,\Delta a)\in A(m,n,\beta,\alpha)} [c+\gamma\mathbb{E}_{D,C,B,A}V(m-c-Db+Ca, n+D-C, B, A)]
$$

Here, B and A are new values of β, α, respectively.

Approximation of V

The future earnings of the MM do not depend much on the actual situation on the market, rather they depend on his resources held (if $m=n=0$ for instance, then all future earnings are zero). Thus, we approximate

$$
V(m,n,\beta,\alpha) \doteq W(m,n) \tag{8}
$$

As such an approximation does not depend on B and A, we need only the distribution of C and D for our computations so this distribution depends only on $\Delta a := a-\alpha$ and $\Delta b := b-\beta$. We further assume that the entire quote may be either fulfilled all or not at all (i.e. $C, D \in \{0,\eta\}$), so $C = \eta C'$ and $D = \eta D'$ for some zero-one variables C' and D'. Thus, we have to estimate/learn only $(\Delta+\delta+1)^2$ probabilities.

Note that $W(0,0) = 0$. As V is monotonic both in m and n (with more resources, there are more options, hence higher income), it is reasonable that its approximation W is monotonic too. We may use [AN APPROXIMATE DYNAMIC PROGRAMMING ALGORITHM FOR MONOTONE VALUE FUNCTIONS by Jiang and Powell][1] [2]to achieve the monotonicity and use value approximation:

1. Initialize W: $W(m,n) \leftarrow \frac{1-\gamma^{m+np}}{1-\gamma}$ for each m,n where p is the current price of the asset (corresponds to the situation when all is consumed as quickly as possible)

2. For each Δa, $\Delta b \neq \Box$ initialize

$$N(C' = 1, D' = 1|\Delta b, \Delta a) = 0,$$

$$N(C' = 1, D = 0|\Delta b, \Delta a) = K \times \begin{cases} 1/3 & \Delta a < 0 \\ 1/4 & \Delta a = 0 \\ 0 & \Delta a > 0 \end{cases}$$

$$N(C' = 0, D = 1|\Delta b, \Delta a) = K \times \begin{cases} 1/3 & \Delta b > 0 \\ 1/4 & \Delta b = 0 \\ 0 & \Delta b < 0 \end{cases}$$

$$N(C' = 0, D' = 0|\Delta b, \Delta a) = K - N(C' = 1, D' = 0|\Delta b, \Delta a) - N(C' = 0, D' = 1|\Delta b, \Delta a)$$

   where K is a number divisible by 3 (e.g. 300), and put $N(C' = 0, D' = 0|\Delta b, \Delta a) = 1$ whenever $\Delta a = \Box$ or $\Delta b = \Box$.

3. Initialize $m = m_0$, $n = n_0$, $a = p + 1, b = p - 1$

4. Repeat until the end of the experiment

   (a) Get (C, D, β, α) from the simulator

   (b) Update $N(C = \bullet, D = \bullet|\Delta b_{last}, \Delta a_{last})$

   (c) Find $v \leftarrow \max_{(c,b,a)\in A(m,n,\beta.\alpha)}[c + \gamma\mathbb{E}(W(m - c - Db + Ca, n + D - C)|\Delta b, \Delta a)]$ where the expectation is computed by means of $p(\bullet|\Delta b, \Delta a)$ which is got from N by division by K.

   (d) $w \leftarrow \epsilon v + (1 - \epsilon)W(m, n)$

   (e) $W \leftarrow \Pi_M((m, n), w, W)$ (4.1 of Jjang and Powell)

[1] https://castlelab.princeton.edu/html/Papers/Jiang-MonotoneADP_arxiv_V4_May252015.pdf

[2] Other options could be nearest neighbor. We can also approximate V by means of a piece-wise linear function during the ADP algorithm, namely to compute its values by interpolation of values (approximations) in existing points. Efficient implementation of this approximation is a non-trivial task, library https://www.sintef.no/globalassets/upload/ikt/9011/geometri/ttl/ttl_1.1.0_doc/index.html will hopefully help.

Parametric (UNDER CONSTRUCTION!)

In order to apply a parameterized approximation of V, we conjecture

1. $V(m,n,r)$ is non-decreasing in β (once the prices are higher, the same inventory is more worth).
2. $\Delta_m V(m,n,r) = 1$ for $m > h_{m,R}$ where $h_{m,R}$ is some constant (once there is h_m of cash available, leaving additional cash unit in business does not help the business and the best thing to do would be to consume it)
3. $\Delta_n V(m,n,r) = \gamma^{\Delta\tau}\overline{\alpha}$, $\overline{\alpha} = \min(a,\alpha)$ for $n > h_n$ where h_n is some constant (once there is h_n stocks available, leaving additional stocks in business does not help the business and the best thing to do would be to try to sell it in the next step)
4. V is concave in both m and n (first, additional funds and stocks holdings bring high additional value for the business but then their contribution falls below one, see above)
5. The higher prices, the more money has to be present in the business ($h_{m,r}$ is increasing in $\overline{\beta}$)
6. It pays off to quote once there is enough money and stocks in the business: $m > h_{m,f}, n > h_n \Rightarrow V(m+\overline{\alpha}, n-1) > V(m,n), V(m-\overline{\beta}, n+1) > V(m,n)$ (i.e. selling or buying a stock pays off).

Additive approximation

To simplify things, we assume that V is additive in the sense that the contribution of m and n add up (there is no super-additive effect) , i.e. V is parametrized/approximated as

$$V(m,n,r) = V_m(m/\overline{\beta})\overline{\beta} + \gamma^{\Delta\tau}\overline{\alpha}V_n(n), \qquad \overline{\beta} = \max(\beta, b)$$

where

$$V_\bullet(x) = x + \sum_{i=1}^{k_\bullet} \phi_i^\bullet s_{ih_\bullet}(x), \qquad s_t(x) = \min(tx, t)$$

where h_m, h_n, k_m, k_n are given integers, ϕ's are unknown non-negative parameters. Note that, as s's are kink concave functions, constant from t, $V_\bullet$ is concave with unit derivative starting from $k_\bullet h_\bullet$. so V fulfills 1. - 6. By adjusting the ϕ's and possibly refining h's and k's

The action space is given by (2) and the transition equation by (4) and (5).

We set $\Delta\tau$ so that the expected traded volume between this interval is 2λ where λ is a pre-chosen constant. We further assume that the total volume of sell/buy market orders Q_S/Q_B in the interval $\Delta\tau$is both Poisson with intensity λ. Then the number of fulfilled buy orders given that the MM posted bid b with volume v and there is π_b orders preferred, is $D = v \wedge [Q_s - \pi_b]_+$ distribution of which is analytically tractable.

Speculator

Further consider a speculator having no obligations but wanting to earn as much as possible. Assuming unit order volumes again, we let her to put at most one buy and one sell limit orders b_t and a_t, respectively. Contrary to MM, we do not require the speculator to quote (i.e. she may avoid putting orders) and we let her put also crossing limit orders (which are in fact equivalent to market orders). The criterion she maximizes is (1).

Maximalistic

Keeping the notation of MM, the action set of the speculator is

$$
\begin{aligned}
&A(m, n, \beta, \alpha) \\
&= \{(c, w, v, b, a) \in \{0, \omega\} \times \{0, \eta\}^2 \times \{\alpha - \Delta, \ldots, \alpha + \delta\} \cup \{0\} \times \{\beta - \delta, \ldots \beta + \Delta\} \cup \{\infty\} : \\
&\qquad b < a, m - bw - c \geq 0, n - v \geq 0, \\
&\qquad\qquad m = 0 \Rightarrow b = 0, n = 0 \Rightarrow a = \infty\}.
\end{aligned}
$$

(note the implications and $\geq' s$ having replaced equivalences and $>$'s). The state space is given by (3) and the transition function by (4), (5) and (6). The Bellman equation is as in (7).

Very minimalistic

Here we assume the speculator to put only market orders with unit volume:

$$A(m, n, r) = \{(c, d) \in \{0, \omega\} \times \{-1, 0, 1\}, m - d\alpha \geq 0, n + d \geq 0\}.$$

with the Bellman equation

$$V(m, n, r) = \max_{(c,d) \in A(m,n,r)} [c + \gamma \mathbb{E}_R V(m - d^+\alpha + d^-\beta - c, n + d, R)]$$

(we neglect the possibiliy that that our trade does not succeed). Note also that, similarly as above, $V(m, n, r) \geq 1 + V(m - 1, n, r)$ and that V is non-decreasing in n. It is a question which quantities to include into r: the absolute minimum is the midpoint price p (and to use the fact that $d^+\alpha - d^-\beta = dp - |d|s/2$ where s is the spread, which we can approximate by a constant); however, it is desirable to experiment with adding another values, like past increases etc.

Optimal buyer/seller

Finally consider an agent who only intents to buy as much as possible of the asset, having m_t units of cash at his disposal. We assume time preference, so the reward function is given by

$$\mathbb{E}(\sum_{\tau=0}^{\infty} \gamma^{-\tau} D_\tau)$$

where D_t is the numer of stocks bought at t.

The maximalistic DP of a buyer

The buyer is supposed to put only buy market orders. His action space is

$$A_t \quad = \quad A(m_t, R_t) \quad = \quad \{v_t \quad \in \quad \mathbb{N}_0 \quad : \quad \pi(v_t,\, R_t) \quad \leq \quad m_t\}$$

where v_t stands for volume ordered (i.e. the volume of a market buy order he puts) at t, $\pi(v, r)$ is the price paid for the market order with value v given the state of the market r. The transition equations are

$$m_{t+1} = m_t - C_{t,R,v}$$

where $C_{t,R,v}$ are random variables determining the true cost, respectively, as a result of putting the market order with volume v to the market with state R.

The Bellman equation is

$$V(m, r) = \max_{h \in A(m,r)} [\mathbb{E}D_{r,v} + \gamma \mathbb{E}_{C,R} V(m - C_{r,v}, F)] \tag{9}$$

where $D_{r,v}$ is a random variable determining how much stocks will be actually bought by a market order v given state of market r.

The minimalistic DP of a buyer

Both the distributions of C and D may depend non-trivially on R (clearly C will be different for $a = 10$ from that given $a = 100$). We approximate them as follows: we assume that all the volume will be always bought, so we take

$$D_{r,v} \doteq v.$$

Further, reflecting the fact that, with increasing volume, the price becomes higher, we assume

$$C_{r,v} \doteq C_{a,v} = v(a + I(v))$$

where I is a random function (price impact). It would be reasonable to take I affine on average, i.e.

$$I(x) = \mathrm{round}(\phi(x - 1)) + E$$

and estimate ϕ by regression; here, E is a centered discrete variable.

As we cannot evaluate π with the knowledge of only a,we have to approximate the action set as well:

$$A(m, r) \doteq A'(m, a) = \{v \in \mathbb{N}_0 : \mathrm{round}(\phi(v - 1)) + \mathrm{stdev}(E))) \leq m_t\}$$

Consequently, our (approximated) problem is

$$\begin{aligned} V(m,r) &\doteq W(m,a) \\ &= \max_{v \in A'(m,r)} [v + \gamma \mathbb{E}_{E,F} W(\max(0, m - v(a + \mathrm{round}(\phi(v-1)) + E)), \max(0, a+F))] \end{aligned} \tag{10}$$

where F is a (centered) random variable, independent of E, estimating the shift of the ask. Note that $W(m, \bullet) = 0$ for $m = 0$.

As for the distribution of E and F, we assume $E = \mathrm{round}(N(0, \sigma_E))$,$F = \mathrm{round}(N(0, \sigma_F))$.

The minimalistic DP of a buyer with repeated orders

If we assume that new money (requests to buy stocks), which only changes (10) to

$$V(m,r) \doteq \max_{v \in A'(m,r)} [v + \gamma \mathbb{E}_{E,F,M} W(\max(0, m + M - v(a + \mathrm{round}(\phi(v-1)) + E)), \max(0, a+F))]$$

where E, F are as above and M is zero with a relatively mall probability p_M and $\kappa_{\mathrm{m}}\mathrm{Poisson}(\lambda_M)$ for some κ_m and λ_m (say $p_M = 0.01$, $\kappa_M = 100$ and $\lambda_M = 3$)

Optimal seller

TBD

The maximalistic DP of a seller

UNDER CONSTRUCTION: The action space of a seller is

$$A_t \quad = \quad A(h_t) \quad = \quad \{v_t \quad \in \quad \mathbb{N}_0 \quad : \quad v_t \quad \leq \quad h_t\}$$

where v_t stands for volume offered (i.e. volume of market sell order put) at t. The revenue c_{t+1} from such an order is equal to

$$c_{t+1} = C'_{R_t, v_t}, \qquad h_{t+1} = D'_{R_t, v_t}$$

where $C'_{R,v}$ and $D'_{R,v}$ are random variables determining the total revenue, number bought, respectively, as a result of putting the market order with volume v to the market with state R.

The approximation is analogous.

Apokryfy

A realistic DP

It may be beneficial to take more than α and β into account. Denote r the additional information (this may be for instance trend) and assume this information is discrete taking finitely many values (hence we can assume $r \in \mathbb{N}$, $r \leq r_{max}$).

As for the solution, we again approximate

$$V(m, n, \beta, \alpha, r) \doteq W(m, n).$$

Denoting R the (random) next value of r, we modify the algorithm

1. same as minimalistic
2. For each $1 \leq r \leq r_{max}$. For each Δa, Δb and r, initialize

   $$N(C = 1, D = 1 | \Delta b, \Delta a, r) = 0,$$

   etc.
3. same
4. Repeat until the end of the experiment
   (a) Get (C, D, β, α, r) from the simulator
   (b) Update $N(C = \bullet, D = \bullet | \Delta b_{last}, \Delta a_{last}, r_{last})$
   (c) Find $v \leftarrow \max_{(c,b,a) \in A(m,n,\beta.\alpha)}[c + \gamma \mathbb{E}(W(m - c - Db + Ca, n + D - C) | \Delta b, \Delta a, r)]$ where the expectation is computed by means of $p(\bullet | \Delta b, \Delta a, r)$ which is got from N.
   (d) same
   (e) same