description of strategies used in the future paper

Martin Šmíd, Peter Hron, Gabriela Suchopárová, Karel Vrbenský

January 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, $a_t, b_t \in \mathbb{N}$, $b_t < a_t$. As a first approach, assume that the quoted volumes are always unit and that the time interval between the decision points is constant, without loss of generality equal to one.

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

$$\sum_{\tau=0}^{\infty} \gamma^{-\tau} c_\tau \tag{1}$$

where $\gamma < 1$ is a discount factor very close to one. 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 reservations is non-negative.

The maximalistic DP

Denoting m_t and n_t the amount of money, stocks, respectively, the action space is

$$\begin{gathered} A_t = A(m_t, n_t, \beta_t, \alpha_t) = \{(c, b, a) \in \{0,1\} \times \mathbb{N}_0 \times \mathbb{N}_\infty : a_t > b_t, \\ a_t > \beta_t, b_t < \alpha_t, m_t - b_t - c_t \geq 0, n_t \geq 0, n_t = 0 \Leftrightarrow a_t = \infty, m_t = 0 \Leftrightarrow b_t = 0\} \end{gathered} \tag{2}$$

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 $D_t \in \{0,1\}$,$C_t \in \{0,1\}$ are indicators of a buy transaction, sell transaction, respectively, by the MM, and F is a complex random function.

The Bellman equation is then

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

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). Specially, given no quoting a or b, it is $C \equiv 0$, $D \equiv 0$, respectively.

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

A minimalistic DP

Restriction of A

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

$$A(m,n,\beta,\alpha) = \{(c,b,a) \in \{0,1\} \times \{\alpha - \Delta, \dots, \alpha + \delta\} \cup \{0\} \times \{\beta - \delta, \dots \beta + \Delta\} \cup \{\infty\} : \\ b < a, a > \beta, b < \alpha, m - b - c \geq 0, n \geq 0, m = 0 \Leftrightarrow b = 0, n = 0 \Leftrightarrow a = \infty\}.$$

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

$$A(m,n,\beta,\alpha) = \{(c,\Delta b,\Delta a) \in \{0,1\} \times \{-\Delta, \dots, \delta\} \cup \{\Box\} \times \{-\delta, \dots \Delta\} \cup \{\Box\} : \\ \beta + \Delta b < a, a > \beta, b < \alpha, m - b - c \geq 0, n \geq 0, m = 0 \Leftrightarrow b = 0, n = 0 \Leftrightarrow a = \infty\}.$$

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

Restriction of S

As it is clear from the definition, the MM's objective depends directly only on process $c_t, b_t, D_t, a_t C_t$. 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. We assume that this distribution depends only on $\Delta a := a - \alpha$ and $\Delta b := b - \beta$; 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, $\Delta b \neq \square$ initialize

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

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

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

[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.

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

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)

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 $r\in\mathbb{N}$, $r\le r_{max}$).

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 $1\le r\le r_{max}$. 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

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

$$A_t = A(m_t, n_t, \beta_t, \alpha_t) = \{(c, b, a) \in \{0,1\} \times \mathbb{N}_0 \times \mathbb{N}_\infty : a_t > b_t, a_t \geq \beta_t, b_t \leq \alpha_t,$$
$$m_t - b_t - c_t \geq 0, n_t \geq 0, n_t = 0 \Rightarrow a_t = \infty, m_t = 0 \Rightarrow b_t = 0\}$$

(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).

Minimalistic and realistic

Here we assume

$$A(m, n, \beta, \alpha) = \{(c, b, a) \in \{0,1\} \times \{\alpha - \Delta, \dots, \alpha + \delta\} \cup \{0\} \times \{\beta - \delta, \dots \beta + \Delta\} \cup \{\infty\} :$$
$$b < a, a \geq \beta, b \leq \alpha, m - b - c \geq 0, n \geq 0, m = 0 \Rightarrow b = 0, n = 0 \Rightarrow a = \infty\}.$$

with the Bellman equation identical to the MM's minimalistic problem. Computation is analogous, differing only in the action space.

Optimal buyer/seller

Finally consider an agent who only intents to buy/sell a predetermined a amount h_∞, h_0, respectively of the asset. As he has no reason to "keep money in the business", we assume him to pay as late as possible (buyer) immediately consume the money torn (seller). We assume time preference, so the reward function of both is given by (1) where c is negative/positive, being the cost of the purchase/revenue from selling.

The maximalistic DP of a buyer

The action space of a buyer is

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

where v_t stands for volume ordered (i.e. volume of market buy order put) at t (we assume unlimited money resources knowing that the agent will minimize their usage). The negative cost c_{t+1} of 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 cost, number bought, respectively, as a result of putting the market order with volume v to the market with state R.

The Bellman equation is

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

The minimalistic DP of a buyer

While the simple action space needs not be further simplified (except for some upper bound of v), and the distribution of $D_{R,v}$ may, with certain license, be viewed as not depending on R, the variable $C_{R,v}$ does depend on R (clearly C will be different for $a = 10$ from that given $a = 100$). Consequently, any approximation of V has to take at least a into account. Reflecting the fact that, with increasing volume, the price becomes higher, we can assume C to be convex in v, we can approximate

$$C_{R,v} = C_{a,v} = a + f(vE)$$

Further, we approximate

$$D_{R,v} \doteq vE.$$

Here, E is random, reflecting the ratio of orders fulfilled, taking values perhaps in $\{0, \frac{1}{2}, 1\}$ (we can assume $\mathbb{P}[E = 1] \doteq 1$ on liquid markets). It would be reasonable to estimate f by regression. Consequently we approximate

$$V(h,c,r) \doteq W(h,c,a) = \max_{v \in A(h)} [c + \gamma \mathbb{E}_{A,E} W(h - vE, -a - f(vE), a + F)]$$

where F is a random variable, independent of E, estimating the shift of the ask.

The maximalistic DP of a seller

The action space of a seller is

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

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.