Homework #9

Junwu Zhang
CME 241: Reinforcement Learning for Finance

February 04, 2020

Problem 1.
Write out the full derivation (in LaTeX) of the Avellaneda-Stoikov result we derived in class

Solution.
In this problem, we are still focusing on the optimal market-making problem. One way to solve it can be the method discussed in the paper by Avellanda and Stoikov in 2006.

The question is formulated as a Markov Decision Process (MDP). In discrete-time notation, we can write the MDP as:

States are (t, S_t, W_t, I_t) where t represents the time step, S_t represents the Trading Order Book (TOB) mid price at time t, W_t is the Market-maker's trading profit-and-loss (PnL) at time t, and I_t is the Market-maker's inventory of shares at time t.

Actions is $(P_t^{(b)}, N_t^{(b)}, P_t^{(a)}, N_t^{(a)})$

TOB Price Dynamics is as follows:

- random bid-shares hit $= X_{t+1}^{(b)} - X_t^{(b)}$ and ask-shares lifted $= X_{t+1}^{(a)} - X_t^{(a)}$
- $W_t \to W_{t+1},\ I_t \to I_{t+1}$
- Stochastic update of S_t to S_{t+1}

Reward at time-step $t+1$ is:

$$R_{t+1} := \begin{cases} 0 & \text{for } 1 \leq t+1 \leq T-1 \\ U\left(W_{t+1} + l_{t+1} \cdot S_{t+1}\right) & \text{for } t+1 = T \end{cases} \tag{1}$$

Therefore, the objective is to find optimal policy $\pi^*(t, S_t, W_t, I_t) = (P_t^{(b)}, N_t^{(b)}, P_t^{(a)}, N_t^{(a)})$ that maximizes $\mathbb{E}[\sum_{t=1}^{T} R_t]$ with discount factor $\gamma = 1$.

Once we have this formulated MDP, the discrete-time notation can be converted to the continuous-time setting mentioned in the paper.

■