Stanford CME 241 (Winter 2021) - Assignment 9

Assignments:

1. We'd like to build a simple simulator of Order Book Dynamics as a MarkovProcess using the code in rl/chapter9/order_book.py. An object of type OrderBook constitutes the *State*. Your task is to come up with a simple model for random arrivals of Market Orders and Limit Orders based on the current contents of the OrderBook. This model of random arrivals of Marker Orders and Limit Orders defines the probabilistic transitions from the current state (OrderBook object) to the next state (OrderBook object). Implement the probabilistic transitions as a MarkovProcess and use it's simulate method to complete your implementation of a simple simulator of Order Book Dynamics.

   Experiment with different models for random arrivals of Market Orders and Limit Orders.

2. Derive the expressions for the Optimal Value Function and Optimal Policy for the *Linear-Percentage Temporary* (LPT) Price Impact Model formulated by Bertsimas and Lo. The LPT model is described below for all $t = 0, 1, \ldots T - 1$:

$$P_{t+1} = P_t \cdot e^{Z_t}$$
$$X_{t+1} = \rho \cdot X_t + \eta_t$$
$$Q_t = P_t \cdot (1 - \beta \cdot N_t - \theta \cdot X_t)$$

   where Z_t are independent and identically distributed random variables with mean μ_Z and variance σ_Z^2 for all $t = 0, 1, \ldots, T - 1$, η_t are independent and identically distributed random variables with mean 0 for all $t = 0, 1, \ldots, T - 1$, Z_t and η_t are independent of each other for all $t = 0, 1, \ldots, T - 1$, and ρ, β, θ are given constants. The model assumes no risk-aversion (Utility function is the identity function) and so, the objective is to maximize the Expected Total Sales Proceeds over the finite-horizon up to time T (discount factor is 1). In your derivation, use the same methodology as we followed for the *Simple Linear Price Impact Model with no Risk-Aversion*.

   Implement this LPT model by customizing the class OptimalOrderExecution in rl/chapter9/optimal_order_exe
   Compare the obtained Optimal Value Function and Optimal Policy against the closed-form solution you derived above.

2. States: $S_t = (P_t, R_t)$

Actions: N_t

Reward $r_{t+1} = U(N_t \cdot Q_t) = U(N_t(P_t(1-\beta N_t - \theta X_t)))$

Price dynamics $P_{t+1} = f_t(P_t, z_t) = P_t e^{z_t}$

Want to find $\pi^*((P_t, R_t)) = N_t^*$ that

maximizes $E[\sum_{t=0}^{T-1} \gamma^t U(N_t Q_t)]$

$$V_t^*((P_t, R_t)) = \max_{N_t} \{ N_t(P_t(1-\beta N_t - \theta X_t)) + E[V_{t+1}^*((P_{t+1}, R_{t+1}))] \}$$

$$V_{T-1}^*((P_{T-1}, R_{T-1})) = N_{T-1}(P_{T-1}(1-\beta N_{T-1} - \theta X_{T-1})) = R_{T-1}(P_{T-1}(1-\beta N_{T-1} - \theta X_{T-1}))$$

$$V_{T-2}^*((P_{T-2}, R_{T-2})) = \max_{N_{T-2}} \{ N_{T-2}(P_{T-2}(1-\beta N_{T-2} - \theta X_{T-2})) + E[R_{T-1}(P_{T-1}(1-\beta N_{T-1} - \theta X_{T-1}))] \}$$

$$= \max_{N_{T-2}} \{ N_{T-2}(P_{T-2}(1-\beta N_{T-2} - \theta X_{T-2})) + (R_{T-2} - N_{T-2})(P_{T-1}(1-\beta(R_{T-2} - N_{T-2}) - \theta X_{T-1})) \}$$

$$= \max_{N_{T-2}} \{ N_{T-2}(P_{T-2}(1-\beta N_{T-2} - \theta X_{T-2})) + (R_{T-2} - N_{T-2})(P_{T-2} e^{z_{T-2}}(1-\beta(R_{T-2} - N_{T-2}) - \theta(\rho X_{T-2} + \eta_{T-2})) \}$$

differentiate w.r.t. N_{T-2}:

$$P_{T-2}(1-\beta N_{T-2} - \theta X_{T-2}) - P_{T-2} e^{z_{T-2}}(1-\beta(R_{T-2} - N_{T-2}) - \theta(\rho X_{T-2} + \eta_{T-2}))$$
$$+ (R_{T-2} - N_{T-2})(P_{T-2} e^{z_{T-2}} \beta) = 0$$

$\Rightarrow$ Substitute N_{T-2}^* in $V_{T-2}^*((P_{T-2}, R_{T-2}))$

$$N_t^* = c_t^{(1)} + c_t^{(2)} R_t + c_t^{(3)} X_t$$

$$V_t^*((P_t, R_t, X_t)) = e^{\mu_Z + \frac{\sigma_Z^2}{2}} \cdot P_t \cdot (c_t^{(4)} + c_t^{(5)} R_t + c_t^{(6)} X_t + c_t^{(7)} R_t^2 + c_t^{(8)} X_t^2 + c_t^{(9)} R_t X_t)$$