

High-Frequency Algorithmic Trading with Momentum and Order Imbalance

My goal is to establish and solve the stochastic optimal control problem that captures the momentum and order imbalance dynamics of the Limit Order Book (LOB). The solution will yield an optimal trading strategy that will permit statistical arbitrage of the underlying stock, which will then be backtested on historical data.

Progress Timeline

DATE	THESIS	STA4505
Dec 2014	• Complete CTMC calibration	
Dec 2014	• Backtest naive strategies based on CTMC	
Jan-May	• Study stochastic controls: ECE1639, STA4505	
Jun 5	• Establish models	Exam Study
Jun 12	• Establish performance criteria	Exam Study
Jun 15	• Derive DPP/DPE	EXAM
Jun 19	• Derive DPP/DPE	
Jun 26	• Derive continuous-time equations	
Jul 3	• Derive discrete-time equations	
Jul 10	• Set up MATLAB numerical integration	
Jul 17	• Integrate functions and plot dynamics	Integrate and analyze too!
Jul 24	• More dynamics, and calib/choose parameters	
Jul 31	• Backtest on historical data	Simulate results
Aug 7	• More backtesting, comparing with previous	
Aug 14	• Dissertation writeup / buffer	Project writeup
Aug 21	• Dissertation writeup / buffer	
Aug 28	• Dissertation writeup	Presentation

The Academic Week in Review

Reminder of our processes (a little bit of abuse of notation going on):

$\mathbf{z}_k = (\rho_k, \Delta_k)$ - 2-D time-homogenous Markov Chain with transition probabilities P_{ij} , where $\rho_k \in \Gamma$ and Γ represents the set of imbalance bins, and $\Delta_k = \text{sgn}(s_k - s_{k-1}) \in \{-1, 0, 1\}$.

$$\begin{aligned} \text{State } \vec{x}_k &= \begin{pmatrix} x_k \\ s_k \\ \mathbf{z}_k \\ q_k \end{pmatrix} && \begin{array}{l} \text{cash} \\ \text{stock price} \\ \text{Markov chain state, as above} \\ \text{inventory} \end{array} \\ \text{Control } \vec{u}_k &= \begin{pmatrix} \delta_k^+ \\ \delta_k^- \\ M_k^+ \\ M_k^- \end{pmatrix} && \begin{array}{l} \text{bid posting depth} \\ \text{ask posting depth} \\ \text{buy MO - binary control} \\ \text{sell MO - binary control} \end{array} \\ \text{Random } \vec{w}_k &= \begin{pmatrix} K_k^+ \\ K_k^- \\ \omega_k \end{pmatrix} && \begin{array}{l} \text{other agent buy MOs - binary} \\ \text{other agent sell MOs - binary} \\ \text{random variable uniformly distributed on } [0,1] \end{array} \end{aligned}$$

We'll write the evolution of the Markov chain as a function of the current state and a uniformly distributed random variable ω :¹

$$\mathbf{z}_{k+1} = T(\mathbf{z}_k, \omega_k) = \sum_{i=0}^{|\Gamma|} i \cdot \mathbb{1}_{(\sum_{j=0}^{i-1} P_{\mathbf{z}_k, j}, \sum_{j=0}^i P_{\mathbf{z}_k, j}]}(\omega_k) \quad (1)$$

Here $\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$, and hence \mathbf{z}_{k+1} is assigned to the value i for which ω_k is in the indicated interval of probabilities.

Our Markovian state evolution function f , given by $\vec{x}_{k+1} = f(\vec{x}_k, \vec{u}_k, \vec{w}_k)$, can be written explicitly as

$$\begin{pmatrix} x_{k+1} \\ s_{k+1} \\ \mathbf{z}_{k+1} \\ q_{k+1} \end{pmatrix} = \begin{pmatrix} x_k + (s_k + \pi + \delta_k^-)L_k^- - (s_k - \pi - \delta_k^+)L_k^+ + (s_k - \pi)M_k^- - (s_k + \pi)M_k^+ \\ s_k + \eta_{k, \mathbf{z}_k} T(\mathbf{z}_k, \omega_k)^{(2)} \\ T(\mathbf{z}_k, \omega_k) \\ q_k + L_k^+ - L_k^- + M_k^+ - M_k^- \end{pmatrix} \quad (2)$$

Clearly, the cash process at a subsequent is equal to the cash at the previous step, plus the costs of profits of executing market or limit orders. There are two noteworthy observations regarding this formulation of the evolution function. First, note that the price paid/received for limit orders depends on the stock price at time k . This implies that at k , if the agent posts a sell limit order, and the binary random variable L_k^- (which depends on the binary random variable M_k^+) is equal to 1, then the agent's order is filled "between timesteps" k and $k+1$, but using the price at time k . Second, since the second dimension $T(\mathbf{z}_k, \omega_k)^{(2)} = \Delta_{k+1} = \text{sgn}(s_{k+1} - s_k)$ determines the directionality of the price jump between times k and $k+1$, multiplying it by the random variable η_{k, \mathbf{z}_k} determines the size of the price change.

¹Borrowed from ECE1639 notes.

Case 1: Max Terminal Wealth (Discrete)

Following traditional dynamic programming, we introduce the value function V_k^u . In this first case, our objective is to maximize the value function given by

$$V_k^u(x, s, \mathbf{z}, q) = \mathbb{E}[W_T^u] = \mathbb{E}_{k,x,s,\mathbf{z},q}[X_T^u + Q_T^u(S_T - \text{sgn}(Q_T^u)\pi) - \alpha(Q_T^u)^2] \quad (3)$$

where, as before, the notation $\mathbb{E}_{k,x,s,\mathbf{z},q}[\cdot]$ represents the conditional expectation

$$\mathbb{E}[\cdot \mid X_k = x, S_k = s, \mathbf{Z}_k = \mathbf{z}, Q_k = q]$$

In this case, our dynamic programming equations (DPEs) are given by

$$V_k(x, s, \mathbf{z}, q) = \sup_{\mathbf{u}} \{ \mathbb{E}_{\mathbf{w}} [V_{k+1}(f((x, s, \mathbf{z}, q), \mathbf{u}, \mathbf{w}_k))] \} \quad (4)$$

$$V_T(x, s, \mathbf{z}, q) = \sup_{\mathbf{u}} \{ \mathbb{E} [x + q(s - \text{sgn}(q)\pi) - \alpha q^2] \} \quad (5)$$

where expectation is with respect to the random vector \mathbf{w}_k .

To simplify the DPEs, we introduce a now familiar ansatz:

$$V_k = x + q(s - \text{sgn}(q)\pi) + h_k(\mathbf{z}, q) \quad (6)$$

with boundary condition $h_T(\mathbf{z}, q) = -\alpha q^2$. Substituting this into the DPE, we obtain

$$\begin{aligned} 0 &= \sup_{\mathbf{u}} \{ \mathbb{E}_{\mathbf{w}} [V_{k+1}(f((x, s, \mathbf{z}, q), \mathbf{u}, \mathbf{w}_k))] \} - V_k(x, s, \mathbf{z}, q) \\ &= \sup_{\mathbf{u}} \left\{ \mathbb{E}_{\mathbf{w}} \left[(s + \pi + \delta^-) L_k^- - (s - \pi - \delta^+) L_k^+ + (s - \pi) M_k^- - (s + \pi) M_k^+ \right. \right. \\ &\quad \left. \left. + (L_k^+ - L_k^- + M_k^+ - M_k^-) \right. \right. \\ &\quad \left. \left. \times (s + \eta_{0,\mathbf{z}} T(\mathbf{z}, \omega)^{(2)} - \text{sgn}(q + L_k^+ - L_k^- + M_k^+ - M_k^-) \pi) \right. \right. \\ &\quad \left. \left. + q (\eta_{0,\mathbf{z}} T(\mathbf{z}, \omega)^{(2)} - (\text{sgn}(q + L_k^+ - L_k^- + M_k^+ - M_k^-) - \text{sgn}(q)) \pi) \right. \right. \\ &\quad \left. \left. + h_{k+1}(T(\mathbf{z}, \omega), q + L_k^+ - L_k^- + M_k^+ - M_k^-) - h_k(\mathbf{z}, q) \right] \right\} \quad (7) \end{aligned}$$

Since our buy/sell market order controls are binary, the supremum over the control vector \mathbf{u} can be treated as a simultaneous supremum over δ^\pm and maximum over the four possible values for M^\pm . Notably, however, a quick substitution shows that the case where $M^+ = M^- = 1$ is not possible as it is always strictly 2π less in value than the case of only limit orders, where $M^+ = M^- = 0$. This should be evident, as buying and selling with market orders in a single timestep yields a guaranteed loss as the agent is forced to cross the spread. Thus, our DPE takes

the form:

$$\begin{aligned}
0 = \max \Bigg\{ & \sup_{\delta^\pm} \left\{ \mathbb{E}_{\mathbf{w}} \left[(s + \pi + \delta^-) L_k^- - (s - \pi - \delta^+) L_k^+ \right. \right. \\
& + (L_k^+ - L_k^-) (s + \eta_{0,z} T(\mathbf{z}, \omega)^{(2)} - \text{sgn}(q + L_k^+ - L_k^-) \pi) \\
& + q (\eta_{0,z} T(\mathbf{z}, \omega)^{(2)} - (\text{sgn}(q + L_k^+ - L_k^-) - \text{sgn}(q)) \pi) \\
& \left. \left. + h_{k+1}(T(\mathbf{z}, \omega), q + L_k^+ - L_k^-) - h_k(\mathbf{z}, q) \right] \right\} ; \\
& \sup_{\delta^\pm} \left\{ \mathbb{E}_{\mathbf{w}} \left[(s + \pi + \delta^-) L_k^- - (s - \pi - \delta^+) L_k^+ - (s + \pi) \right. \right. \\
& + (L_k^+ - L_k^- + 1) (s + \eta_{0,z} T(\mathbf{z}, \omega)^{(2)} - \text{sgn}(q + L_k^+ - L_k^- + 1) \pi) \\
& + q (\eta_{0,z} T(\mathbf{z}, \omega)^{(2)} - (\text{sgn}(q + L_k^+ - L_k^- + 1) - \text{sgn}(q)) \pi) \\
& \left. \left. + h_{k+1}(T(\mathbf{z}, \omega), q + L_k^+ - L_k^- + 1) - h_k(\mathbf{z}, q) \right] \right\} ; \\
& \sup_{\delta^\pm} \left\{ \mathbb{E}_{\mathbf{w}} \left[(s + \pi + \delta^-) L_k^- - (s - \pi - \delta^+) L_k^+ + (s - \pi) \right. \right. \\
& + (L_k^+ - L_k^- - 1) (s + \eta_{0,z} T(\mathbf{z}, \omega)^{(2)} - \text{sgn}(q + L_k^+ - L_k^- - 1) \pi) \\
& + q (\eta_{0,z} T(\mathbf{z}, \omega)^{(2)} - (\text{sgn}(q + L_k^+ - L_k^- - 1) - \text{sgn}(q)) \pi) \\
& \left. \left. + h_{k+1}(T(\mathbf{z}, \omega), q + L_k^+ - L_k^- - 1) - h_k(\mathbf{z}, q) \right] \right\} \Bigg\}
\end{aligned} \tag{8}$$

We'll begin by concentrating on the first case where $M^+ = M^- = 0$. Thus, we want to solve

$$\begin{aligned}
\sup_{\delta^\pm} \Bigg\{ & \mathbb{E}_{\mathbf{w}} \left[(s + \pi + \delta^-) L_k^- - (s - \pi - \delta^+) L_k^+ \right. \\
& + (L_k^+ - L_k^-) (s + \eta_{0,z} T(\mathbf{z}, \omega)^{(2)} - \text{sgn}(q + L_k^+ - L_k^-) \pi) \\
& + q (\eta_{0,z} T(\mathbf{z}, \omega)^{(2)} - (\text{sgn}(q + L_k^+ - L_k^-) - \text{sgn}(q)) \pi) \\
& \left. \left. + h_{k+1}(T(\mathbf{z}, \omega), q + L_k^+ - L_k^-) - h_k(\mathbf{z}, q) \right] \right\}
\end{aligned} \tag{9}$$

First, some preliminary results:

$$\begin{aligned}
\mathbb{P}[L_k^- = 1 | K_k^+ = 1] &= e^{-\kappa \delta^-} \\
\mathbb{P}[L_k^- = 0 | K_k^+ = 1] &= 1 - e^{-\kappa \delta^-} \\
\mathbb{P}[L_k^- = 0 | K_k^+ = n] &= (1 - e^{-\kappa \delta^-})^n \\
\mathbb{P}[L_k^- = 1 | K_k^+ = n] &= 1 - (1 - e^{-\kappa \delta^-})^n \\
\mathbb{P}[K_k^+ = n] &= \frac{e^{-\lambda \Delta t} (\lambda \Delta t)^n}{n!}
\end{aligned}$$

Theorem 1. $\mathbb{E}[L_k^-] = \mathbb{P}[L_k^- = 1] = 1 - e^{-e^{-\kappa\delta^-} \lambda \Delta t}$

Proof.

$$\begin{aligned}
\mathbb{E}[L_k^-] &= \sum_{n=0}^{\infty} \mathbb{P}[L_k^- = 1 | K_k^+ = n] \cdot \mathbb{P}[K_k^+ = n] \\
&= \sum_{n=0}^{\infty} [1 - (1 - e^{-\kappa\delta^-})^n] \frac{e^{-\lambda\Delta t} (\lambda\Delta t)^n}{n!} \\
&= 1 - e^{-\lambda\Delta t} \sum_{n=0}^{\infty} \frac{(1 - e^{-\kappa\delta^-})^n (\lambda\Delta t)^n}{n!} \\
&= 1 - e^{-\lambda\Delta t} \sum_{n=0}^{\infty} \frac{(\lambda\Delta t - e^{-\kappa\delta^-} \lambda\Delta t)^n}{n!} \\
&= 1 - e^{-\lambda\Delta t} e^{\lambda\Delta t - e^{-\kappa\delta^-} \lambda\Delta t} \\
&= 1 - e^{-e^{-\kappa\delta^-} \lambda\Delta t}
\end{aligned}$$

□

For ease of notation, we'll write the probability of the $L_k^- = 0$ event as

$$p(\delta^-) = e^{-e^{-\kappa\delta^-} \lambda\Delta t}$$

and its derivative as

$$d(\delta^-) = \partial_{\delta^-} p(\delta^-) = \kappa \lambda \Delta t e^{-e^{-\kappa\delta^-} \lambda\Delta t - \kappa\delta^-}$$

This gives us the results:

$$\begin{aligned}
\mathbb{P}[L_k^- = 1] &= 1 - p(\delta^-) = \mathbb{E}[L_k^-] \\
\mathbb{P}[L_k^- = 0] &= p(\delta^-) \\
\partial_{\delta^-} \mathbb{P}[L_k^- = 1] &= -d(\delta^-) \\
\partial_{\delta^-} \mathbb{P}[L_k^- = 0] &= d(\delta^-)
\end{aligned}$$

Let's pre-compute some of the terms that we'll encounter in the supremum, namely the expectations of the random variables.

$$\begin{aligned}
\mathbb{E}[\text{sgn}(q + L_k^+ - L_k^-)] &= \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 1] \cdot \text{sgn}(q) \\
&\quad + \mathbb{P}[L_k^- = 1] \cdot \mathbb{P}[L_k^+ = 0] \cdot \text{sgn}(q - 1) \\
&\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 1] \cdot \text{sgn}(q + 1) \\
&\quad + \mathbb{P}[L_k^- = 0] \cdot \mathbb{P}[L_k^+ = 0] \cdot \text{sgn}(q) \\
&= (1 - p(\delta^-))(1 - p(\delta^+)) \text{sgn}(q) \\
&\quad + (1 - p(\delta^-))p(\delta^+) \text{sgn}(q - 1) \\
&\quad + p(\delta^-)(1 - p(\delta^+)) \text{sgn}(q + 1) \\
&\quad + p(\delta^-)p(\delta^+) \text{sgn}(q) \\
&= \text{sgn}(q) [1 - p(\delta^+) - p(\delta^-) + 2p(\delta^+)p(\delta^-)] \\
&\quad + \text{sgn}(q - 1) [p(\delta^+) - p(\delta^+)p(\delta^-)] \\
&\quad + \text{sgn}(q + 1) [p(\delta^-) - p(\delta^+)p(\delta^-)] \\
&= \begin{cases} 1 & q \geq 2 \\ 1 - p(\delta^+)(1 - p(\delta^-)) & q = 1 \\ p(\delta^-) - p(\delta^+) & q = 0 \\ -[1 - p(\delta^-)(1 - p(\delta^+))] & q = -1 \\ -1 & q \leq -2 \end{cases} \tag{10}
\end{aligned}$$

$$= \Phi(q, \delta^+, \delta^-) \tag{11}$$

Hence, we can also compute the partial derivatives of this expectation:

$$\partial_{\delta^-} \Phi(q, \delta^+, \delta^-) = \begin{cases} 0 & q \geq 2 \\ d(\delta^-)p(\delta^+) & q = 1 \\ d(\delta^-) & q = 0 \\ d(\delta^-)(1 - p(\delta^+)) & q = -1 \\ 0 & q \leq -2 \end{cases} \tag{12}$$

$$\partial_{\delta^+} \Phi(q, \delta^+, \delta^-) = \begin{cases} 0 & q \geq 2 \\ -d(\delta^+)(1 - p(\delta^-)) & q = 1 \\ -d(\delta^+) & q = 0 \\ -d(\delta^+)p(\delta^-) & q = -1 \\ 0 & q \leq -2 \end{cases} \tag{13}$$

As it'll be required for later, we'll additionally introduce the shorthand notation

$$\frac{\partial_{\delta^-} \Phi(q, \delta^+, \delta^-)}{d(\delta^-)} = \Psi(q, \delta^+) \tag{14}$$

$$\frac{\partial_{\delta^+} \Phi(q, \delta^+, \delta^-)}{d(\delta^+)} = \Upsilon(q, \delta^-) \tag{15}$$

Recalling that we have \mathbf{P} the transition matrix for the Markov Chain \mathbf{Z} , with $\mathbf{P}_{\mathbf{z}, \mathbf{j}} = \mathbb{P}[\mathbf{Z}_{k+1} = \mathbf{j} | \mathbf{Z}_k = \mathbf{z}]$, then we can also write:

$$\begin{aligned} \mathbb{E}[h_{k+1}(T(\mathbf{z}, \omega), q + L_k^+ - L_k^-)] &= \sum_{\mathbf{j}} \mathbf{P}_{\mathbf{z}, \mathbf{j}} \left[h_{k+1}(\mathbf{j}, q) [1 - p(\delta^+) - p(\delta^-) + 2p(\delta^+)p(\delta^-)] \right. \\ &\quad + h_{k+1}(\mathbf{j}, q - 1) [p(\delta^+) - p(\delta^+)p(\delta^-)] \\ &\quad \left. + h_{k+1}(\mathbf{j}, q + 1) [p(\delta^-) - p(\delta^+)p(\delta^-)] \right] \end{aligned} \quad (16)$$

and its partial derivatives as

$$\begin{aligned} \partial_{\delta^-} \mathbb{E}[h_{k+1}(T(\mathbf{z}, \omega), q + L_k^+ - L_k^-)] &= \sum_{\mathbf{j}} \mathbf{P}_{\mathbf{z}, \mathbf{j}} \left[h_{k+1}(\mathbf{j}, q) [-d(\delta^-) + 2p(\delta^+)d(\delta^-)] \right. \\ &\quad + h_{k+1}(\mathbf{j}, q - 1) [-p(\delta^+)d(\delta^-)] \\ &\quad \left. + h_{k+1}(\mathbf{j}, q + 1) [d(\delta^-) - p(\delta^+)d(\delta^-)] \right] \end{aligned} \quad (17)$$

$$\begin{aligned} &= d(\delta^-) \sum_{\mathbf{j}} \mathbf{P}_{\mathbf{z}, \mathbf{j}} \left[h_{k+1}(\mathbf{j}, q) [-1 + 2p(\delta^+)] \right. \\ &\quad + h_{k+1}(\mathbf{j}, q - 1) [-p(\delta^+)] \\ &\quad \left. + h_{k+1}(\mathbf{j}, q + 1) [1 - p(\delta^+)] \right] \end{aligned} \quad (18)$$

$$\begin{aligned} \partial_{\delta^+} \mathbb{E}[h_{k+1}(T(\mathbf{z}, \omega), q + L_k^+ - L_k^-)] &= d(\delta^+) \sum_{\mathbf{j}} \mathbf{P}_{\mathbf{z}, \mathbf{j}} \left[h_{k+1}(\mathbf{j}, q) [-1 + 2p(\delta^-)] \right. \\ &\quad + h_{k+1}(\mathbf{j}, q - 1) [1 - p(\delta^-)] \\ &\quad \left. + h_{k+1}(\mathbf{j}, q + 1) [-p(\delta^-)] \right] \end{aligned} \quad (19)$$

Now we tackle solving the supremum in equation 9. First we consider the first-order condition on δ^- , namely that the partial derivative with respect to it must be equal to zero.

$$\begin{aligned} 0 &= (1 - p(\delta^-)) - d(\delta^-)(s + \pi + \delta^-) + d(\delta^-)(s + \mathbb{E}[\eta_{0, \mathbf{z}} T(\mathbf{z}, \omega)^{(2)}] - \Phi(q, \delta^+, \delta^-)\pi) \\ &\quad - \partial_{\delta^-} \Phi(q, \delta^+, \delta^-)\pi(p(\delta^-) - p(\delta^+)) - q\pi \partial_{\delta^-} \Phi(q, \delta^+, \delta^-) + d(\delta^-) \sum_{\mathbf{j}} \mathbf{P}_{\mathbf{z}, \mathbf{j}} \\ &\quad \times \left[h_{k+1}(\mathbf{j}, q) [-1 + 2p(\delta^+)] + h_{k+1}(\mathbf{j}, q - 1) [-p(\delta^+)] + h_{k+1}(\mathbf{j}, q + 1) [1 - p(\delta^+)] \right] \end{aligned} \quad (20)$$

and diving through by $d(\delta^-)$, which is nonzero, we get

$$\begin{aligned}
&= \frac{1 - p(\delta^-)}{d(\delta^-)} - (s + \pi + \delta^-) + (s + \mathbb{E}[\eta_{0,z}T(\mathbf{z}, \omega)^{(2)}] - \Phi(q, \delta^+, \delta^-)\pi) \\
&\quad - \pi(p(\delta^-) - p(\delta^+))\Psi(q, \delta^+) - q\pi\Psi(q, \delta^+) \\
&\quad + \sum_{\mathbf{j}} \mathbf{P}_{z,\mathbf{j}} \left[h_{k+1}(\mathbf{j}, q) [-1 + 2p(\delta^+)] + h_{k+1}(\mathbf{j}, q - 1) [-p(\delta^+)] + h_{k+1}(\mathbf{j}, q + 1) [1 - p(\delta^+)] \right]
\end{aligned} \tag{21}$$

Re-arranging and collecting terms in δ^- on the left-hand side:

$$\begin{aligned}
&\frac{1 - p(\delta^-)}{d(\delta^-)} + \delta^- + \pi\Phi(q, \delta^+, \delta^-) + \pi p(\delta^-)\Psi(q, \delta^+) \\
&= \mathbb{E}[\eta_{0,z}T(\mathbf{z}, \omega)^{(2)}] + \pi(-1 + p(\delta^+)\Psi(q, \delta^+) - q\Psi(q, \delta^+)) \\
&\quad + \sum_{\mathbf{j}} \mathbf{P}_{z,\mathbf{j}} \left[h_{k+1}(\mathbf{j}, q) [-1 + 2p(\delta^+)] + h_{k+1}(\mathbf{j}, q - 1) [-p(\delta^+)] + h_{k+1}(\mathbf{j}, q + 1) [1 - p(\delta^+)] \right]
\end{aligned} \tag{22}$$

$$\frac{1 - p(\delta^-)}{d(\delta^-)} = \frac{e^{\kappa\delta^-} \left(e^{\lambda\Delta t e^{-\kappa\delta^-}} \right)}{\kappa\lambda\Delta t}$$