A Guided Tour of Chapter 8: Order Book Algorithms

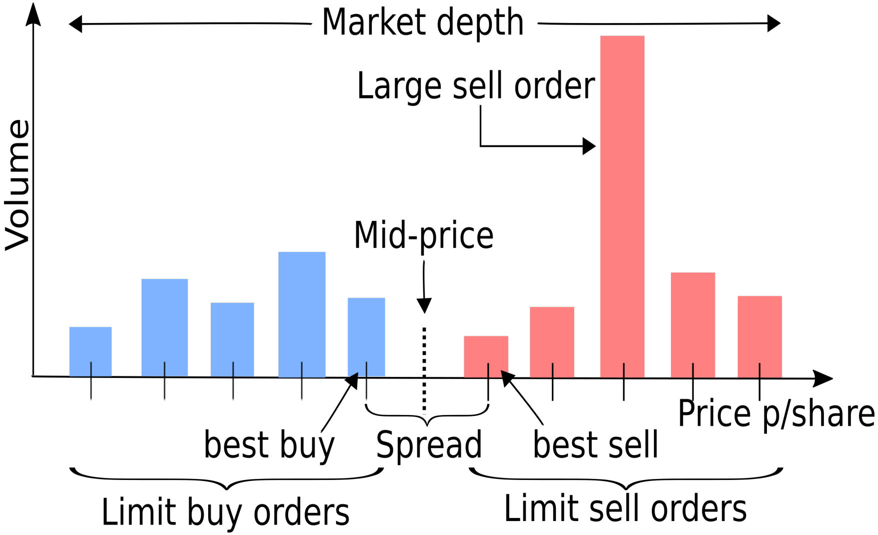
Ashwin Rao

ICME, Stanford University

February 12, 2021

Overview

Trading Order Book (abbrev. OB)



Basics of Order Book (OB)

- Buyers/Sellers express their intent to trade by submitting bids/asks
- These are Limit Orders (LO) with a price P and size N
- Buy LO (P, N) states willingness to buy N shares at a price $\leq P$
- Sell LO (P, N) states willingness to sell N shares at a price $\geq P$
- Order Book aggregates order sizes for each unique price
- So we can represent with two sorted lists of (Price, Size) pairs

$$\text{Bids: } [(P_i^{(b)}, N_i^{(b)}) \mid 0 \leq i < m], P_i^{(b)} > P_j^{(b)} \text{ for } i < j$$

$$\text{Asks: } [(P_i^{(a)}, N_i^{(a)}) \mid 0 \leq i < n], P_i^{(a)} < P_j^{(a)} \text{ for } i < j$$

- We call $P_0^{(b)}$ as simply *Bid*, $P_0^{(a)}$ as *Ask*, $\frac{P_0^{(a)} + P_0^{(b)}}{2}$ as *Mid*
- We call $P_0^{(a)} - P_0^{(b)}$ as *Spread*, $P_{n-1}^{(a)} - P_{m-1}^{(b)}$ as *Market Depth*
- A Market Order (MO) states intent to buy/sell N shares at the *best possible price(s)* available on the OB at the time of MO submission

The class OrderBook

```
@dataclass(frozen=True)
class DollarsAndShares:
    dollars: float
    shares: int

PriceSizePairs = Sequence[DollarsAndShares]

@dataclass(frozen=True)
class OrderBook:
    descending_bids: PriceSizePairs
    ascending_asks: PriceSizePairs
```

Order Book (OB) Activity

- A new Sell LO (P, N) potentially removes best bid prices on the OB

$$\text{Removal: } [(P_i^{(b)}, \min(N_i^{(b)}, \max(0, N - \sum_{j=0}^{i-1} N_j^{(b)}))) \mid (i : P_i^{(b)} \geq P)]$$

- After this removal, it adds the following to the asks side of the OB

$$(P, \max(0, N - \sum_{i:P_i^{(b)} \geq P} N_i^{(b)}))$$

- A new Buy MO operates analogously (on the other side of the OB)
- A Sell Market Order N will remove the best bid prices on the OB

$$\text{Removal: } [(P_i^{(b)}, \min(N_i^{(b)}, \max(0, N - \sum_{j=0}^{i-1} N_j^{(b)}))) \mid 0 \leq i < m]$$

- A Buy Market Order N will remove the best ask prices on the OB

$$\text{Removal: } [(P_i^{(a)}, \min(N_i^{(a)}, \max(0, N - \sum_{j=0}^{i-1} N_j^{(a)}))) \mid 0 \leq i < n]$$

OrderBook Activity methods

```
def eat_book(
    ps_pairs: PriceSizePairs,
    shares: int
) -> Tuple[DollarsAndShares, PriceSizePairs]:

def sell_limit_order(
    self,
    price: float,
    shares: int
) -> Tuple[DollarsAndShares, OrderBook]:

def sell_market_order(
    self,
    shares: int
) -> Tuple[DollarsAndShares, OrderBook]:
```

Price Impact and Order Book Dynamics

- We focus on how a Market order (MO) alters the OB
- A large-sized MO often results in a big *Spread* which could soon be replenished by new LOs, potentially from either side
- So a large-sized MO moves the Bid/Ask/Mid (*Price Impact* of MO)
- Subsequent Replenishment activity is part of *OB Dynamics*
- Models for OB Dynamics can be quite complex
- We will cover a few simple Models in this lecture
- Models based on how a Sell MO will move the OB *Bid Price*
- Models of Buy MO moving the OB *Ask Price* are analogous

Optimal Trade Order Execution Problem

- The task is to sell a large number N of shares
- We are allowed to trade in T discrete time steps
- We are only allowed to submit Market Orders
- We consider both *Temporary* and *Permanent* Price Impact
- For simplicity, we consider a model of just the *Bid Price* Dynamics
- Goal is to maximize Expected Total Utility of Sales Proceeds
- By breaking N into appropriate chunks (timed appropriately)
- If we sell too fast, we are likely to get poor prices
- If we sell too slow, we risk running out of time
- Selling slowly also leads to more uncertain proceeds (lower Utility)
- This is a Dynamic Optimization problem
- We can model this problem as a Markov Decision Process (MDP)

Problem Notation

- Time steps indexed by $t = 0, 1, \ldots, T$
- P_t denotes Bid Price at start of time step t
- N_t denotes number of shares sold in time step t
- $R_t = N - \sum_{i=0}^{t-1} N_i$ = shares remaining to be sold at start of time step t
- Note that $R_0 = N, R_{t+1} = R_t - N_t$ for all $t < T, N_{T-1} = R_{T-1} \Rightarrow R_T = 0$
- Price Dynamics given by:

$$P_{t+1} = f_t(P_t, N_t, \epsilon_t)$$

  where $f_t(\cdot)$ is an arbitrary function incorporating:
  - Permanent Price Impact of selling N_t shares
  - Impact-independent market-movement of Bid Price over time step t
  - ϵ_t denotes source of randomness in Bid Price market-movement
- Sales Proceeds in time step t defined as:

$$N_t \cdot Q_t = N_t \cdot (P_t - g_t(P_t, N_t))$$

  where $g_t(\cdot)$ is an arbitrary func representing Temporary Price Impact
- Utility of Sales Proceeds function denoted as $U(\cdot)$

Markov Decision Process (MDP) Formulation

- This is a discrete-time, finite-horizon MDP
- MDP Horizon is time T, meaning all states at time T are terminal
- Order of MDP activity in each time step $0 \leq t < T$:
  - Observe *State* $s_t := (P_t, R_t) \in \mathcal{S}_t$
  - Perform *Action* $a_t := N_t \in \mathcal{A}_t$
  - Receive *Reward* $r_{t+1} := U(N_t \cdot Q_t) = U(N_t \cdot (P_t - g_t(P_t, N_t)))$
  - Experience Price Dynamics $P_{t+1} = f_t(P_t, N_t, \epsilon_t)$
- Goal is to find a Policy $\pi_t^*((P_t, R_t)) = N_t^*$ that maximizes:

$$\mathbb{E}[\sum_{t=0}^{T-1} \gamma^t \cdot U(N_t \cdot Q_t)] \text{ where } \gamma \text{ is MDP discount factor}$$

A Simple Linear Impact Model with No Risk-Aversion

- We consider a simple model with Linear Price Impact
- N, N_t, P_t are all continuous-valued ($\in \mathbb{R}$)
- Price Dynamics: $P_{t+1} = P_t - \alpha N_t + \epsilon_t$ where $\alpha \in \mathbb{R}$
- ϵ_t is i.i.d. with $\mathbb{E}[\epsilon_t | N_t, P_t] = 0$
- So, Permanent Price Impact is $\alpha \cdot N_t$
- Temporary Price Impact given by $\beta \cdot N_t$, so $Q_t = P_t - \beta \cdot N_t$ ($\beta \in \mathbb{R}_{\geq 0}$)
- Utility function $U(\cdot)$ is the identity function, i.e., no Risk-Aversion
- MDP Discount factor $\gamma = 1$
- This is an unrealistic model, but solving this gives plenty of intuition
- Approach: Define Optimal Value Function & invoke Bellman Equation

Optimal Value Function and Bellman Equation

- Denote Value Function for policy π as:

$$V_t^\pi((P_t, R_t)) = \mathbb{E}_\pi[\sum_{i=t}^{T} N_i(P_i - \beta \cdot N_i)|(P_t, R_t)]$$

- Denote Optimal Value Function as $V_t^*((P_t, R_t)) = max_\pi V_t^\pi((P_t, R_t))$
- Optimal Value Function satisfies the Bellman Eqn ($\forall 0 \leq t < T - 1$):

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

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

- From the above, we can infer $V_{T-2}^*((P_{T-2}, R_{T-2}))$ as:

$$\max_{N_{T-2}}\{N_{T-2}(P_{T-2} - \beta N_{T-2}) + \mathbb{E}[R_{T-1}(P_{T-1} - \beta R_{T-1})]\}$$

$$= \max_{N_{T-2}}\{N_{T-2}(P_{T-2} - \beta N_{T-2}) + \mathbb{E}[(R_{T-2} - N_{T-2})(P_{T-1} - \beta(R_{T-2} - N_{T-2}))]\}$$

Optimal Policy & Optimal Value Function for case $\alpha \geq 2\beta$

$$= \max_{N_{T-2}} \{R_{T-2}P_{T-2} - \beta R_{T-2}^2 + (\alpha - 2\beta)(N_{T-2}^2 - N_{T-2}R_{T-2})\}$$

- For the case $\alpha \geq 2\beta$, we have the trivial solution: $N_{T-1}^* = 0$ or R_{T-1}
- Substitute N_{T-2}^* in the expression for $V_{T-2}^*((P_{T-2}, R_{T-2}))$:

$$V_{T-2}^*((P_{T-2}, R_{T-2}) = R_{T-2}(P_{T-2} - \beta R_{T-2})$$

- Continuing backwards in time in this manner gives:

$$N_t^* = 0 \text{ or } R_t$$

$$V_t^*((P_t, R_t)) = R_t(P_t - \beta R_t)$$

- So the solution for the case $\alpha \geq 2\beta$ is to sell all N shares at any one of the time steps $t = 0, \ldots, T-1$ (and none in the other time steps) and the Optimal Expected Total Sale Proceeds = $N(P_0 - \beta N)$

Optimal Policy & Optimal Value Function for case $\alpha < 2\beta$

- For the case $\alpha < 2\beta$, differentiating w.r.t. N_{T-2} and setting to 0 gives:

$$(\alpha - 2\beta)(2N^*_{T-2} - R_{T-2}) = 0 \Rightarrow N^*_{T-2} = \frac{R_{T-2}}{2}$$

- Substitute N^*_{T-2} in the expression for $V^*_{T-2}((P_{T-2}, R_{T-2})$:

$$V^*_{T-2}((P_{T-2}, R_{T-2})) = R_{T-2}P_{T-2} - R^2_{T-2}(\frac{\alpha + 2\beta}{4})$$

- Continuing backwards in time in this manner gives:

$$N^*_t = \frac{R_t}{T-t}$$

$$V^*_t((P_t, R_t)) = R_t P_t - \frac{R_t^2}{2}(\frac{2\beta + \alpha(T-t-1)}{T-t})$$

Interpreting the solution for the case $\alpha < 2\beta$

- Rolling forward in time, we see that $N_t^* = \frac{N}{T}$, i.e., uniformly split
- Hence, Optimal Policy is a constant (independent of *State*)
- Uniform split makes intuitive sense because Price Impact and Market Movement are both linear and additive, and don't interact
- Essentially equivalent to minimizing $\sum_{t=1}^T N_t^2$ with $\sum_{t=1}^T N_t = N$
- Optimal Expected Total Sale Proceeds = $NP_0 - \frac{N^2}{2}(\alpha + \frac{2\beta-\alpha}{T})$
- So, *Implementation Shortfall* from Price Impact is $\frac{N^2}{2}(\alpha + \frac{2\beta-\alpha}{T})$
- Note that Implementation Shortfall is non-zero even if one had infinite time available ($T \rightarrow \infty$) for the case of $\alpha > 0$
- If Price Impact were purely temporary ($\alpha = 0$, i.e., Price fully snapped back), Implementation Shortfall is zero with infinite time available

Models in Bertsimas-Lo paper

- Bertsimas-Lo was the first paper on Optimal Trade Order Execution
- They assumed no risk-aversion, i.e. identity Utility function
- The first model in their paper is a special case of our simple Linear Impact model, with fully Permanent Impact (i.e., $\alpha = \beta$)
- Next, Betsimas-Lo extended the Linear Permanent Impact model
- To include dependence on Serially-Correlated Variable X_t

$$P_{t+1} = P_t - (\beta N_t + \theta X_t) + \epsilon_t, X_{t+1} = \rho X_t + \eta_t, Q_t = P_t - (\beta N_t + \theta X_t)$$

- ϵ_t and η_t are i.i.d. (and mutually independent) with mean zero
- X_t can be thought of as market factor affecting P_t linearly
- Bellman Equation on Optimal VF and same approach as before yields:

$$N_t^* = \frac{R_t}{T - t} + h(t, \beta, \theta, \rho) X_t$$

$$V_t^*((P_t, R_t, X_t)) = R_t P_t - (\text{quadratic in } (R_t, X_t) + \text{ constant})$$

- Seral-correlation predictability ($\rho \neq 0$) alters uniform-split strategy

A more Realistic Model: LPT Price Impact

- Next, Bertsimas-Lo present a more realistic model called "LPT"
- *Linear-Percentage Temporary* Price Impact model features:
  - Geometric random walk: consistent with real data, & avoids prices ≤ 0
  - % Price Impact $\frac{g_t(P_t, N_t)}{P_t}$ doesn't depend on P_t (validated by real data)
  - Purely Temporary Price Impact

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

- Z_t is a random variable with mean μ_Z and variance σ_Z^2
- With the same derivation as before, we get the solution:

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

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

Incorporating Risk-Aversion/Utility of Proceeds

- For analytical tractability, Bertsimas-Lo ignored Risk-Aversion
- But one is typically wary of *Risk of Uncertain Proceeds*
- We'd trade some (Expected) Proceeds for lower Variance of Proceeds
- Almgren-Chriss work in this Risk-Aversion framework
- They consider our simple linear model maximizing $E[Y] - \lambda Var[Y]$
- Where Y is the total (uncertain) proceeds $\sum_{t=0}^{T-1} N_t Q_t$
- λ controls the degree of risk-aversion and hence, the trajectory of N_t^*
- $\lambda = 0$ leads to uniform split strategy $N_t^* = \frac{N}{T}$
- The other extreme is to minimize $Var[Y]$ which yields $N_0^* = N$
- Almgren-Chriss derive *Efficient Frontier* and solutions for specific $U(\cdot)$
- Much like classical Portfolio Optimization problems

Real-world Optimal Trade Order Execution (& Extensions)

- Arbitrary Price Dynamics $f_t(\cdot)$ and Temporary Price Impact $g_t(\cdot)$
- Non-stationarity/non-linear dynamics/impact require (Numerical) DP
- Frictions: Discrete Prices/Sizes, Constraints on Prices/Sizes, Fees
- Incorporating various markets factors in the State bloats State Space
- We could also represent the entire OB within the State
- Practical route is to develop a simulator capturing all of the above
- Simulator is a *Market-Data-learnt Sampling Model* of OB Dynamics
- In practice, we'd need to also capture *Cross-Asset Market Impact*
- Using this simulator and neural-networks func approx, we can do RL
- References: Nevmyvaka, Feng, Kearns; 2006 and Vyetrenko, Xu; 2019
- Exciting area for Future Research as well as Engineering Design

OB Dynamics and Market-Making

- Modeling OB Dynamics involves predicting arrival of MOs and LOs
- Market-makers are liquidity providers (providers of Buy and Sell LOs)
- Other market participants are typically liquidity takers (MOs)
- But there are also other market participants that trade with LOs
- Complex interplay between market-makers & other mkt participants
- Hence, OB Dynamics tend to be quite complex
- We view the OB from the perspective of a single market-maker who aims to gain with Buy/Sell LOs of appropriate width/size
- By anticipating OB Dynamics & dynamically adjusting Buy/Sell LOs
- Goal is to maximize *Utility of Gains* at the end of a suitable horizon
- If Buy/Sell LOs are too narrow, more frequent but small gains
- If Buy/Sell LOs are too wide, less frequent but large gains
- Market-maker also needs to manage potential unfavorable inventory (long or short) buildup and consequent unfavorable liquidation

Notation for Optimal Market-Making Problem

- We simplify the setting for ease of exposition
- Assume finite time steps indexed by $t = 0, 1, \ldots, T$
- Denote $W_t \in \mathbb{R}$ as Market-maker's trading PnL at time t
- Denote $I_t \in \mathbb{Z}$ as Market-maker's inventory of shares at time t ($I_0 = 0$)
- $S_t \in \mathbb{R}^+$ is the OB Mid Price at time t (assume stochastic process)
- $P_t^{(b)} \in \mathbb{R}^+, N_t^{(b)} \in \mathbb{Z}^+$ are market maker's Bid Price, Bid Size at time t
- $P_t^{(a)} \in \mathbb{R}^+, N_t^{(a)} \in \mathbb{Z}^+$ are market-maker's Ask Price, Ask Size at time t
- Assume market-maker can add or remove bids/asks costlessly
- Denote $\delta_t^{(b)} = S_t - P_t^{(b)}$ as Bid Spread, $\delta_t^{(a)} = P_t^{(a)} - S_t$ as Ask Spread
- Random var $X_t^{(b)} \in \mathbb{Z}_{\geq 0}$ denotes bid-shares "hit" *up to* time t
- Random var $X_t^{(a)} \in \mathbb{Z}_{\geq 0}$ denotes ask-shares "lifted" *up to* time t

$$W_{t+1} = W_t + P_t^{(a)} \cdot (X_{t+1}^{(a)} - X_t^{(a)}) - P_t^{(b)} \cdot (X_{t+1}^{(b)} - X_t^{(b)}) \text{ , } I_t = X_t^{(b)} - X_t^{(a)}$$

- Goal to maximize $\mathbb{E}[U(W_T + I_T \cdot S_T)]$ for appropriate concave $U(\cdot)$

Markov Decision Process (MDP) Formulation

- Order of MDP activity in each time step $0 \leq t \leq T-1$:
  - Observe *State* $:= (S_t, W_t, I_t) \in \mathcal{S}_t$
  - Perform *Action* $:= (P_t^{(b)}, N_t^{(b)}, P_t^{(a)}, N_t^{(a)}) \in \mathcal{A}_t$
  - Experience OB Dynamics resulting in:
    - random bid-shares hit $= X_{t+1}^{(b)} - X_t^{(b)}$ and ask-shares lifted $= X_{t+1}^{(a)} - X_t^{(a)}$
    - update of W_t to W_{t+1}, update of I_t to I_{t+1}
    - stochastic evolution of S_t to S_{t+1}
  - Receive next-step $(t+1)$ *Reward* R_{t+1}

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

- Goal is to find an *Optimal Policy* $\pi^* = (\pi_0^*, \pi_1^*, \ldots, \pi_{T-1}^*)$:

$$\pi_t^*((S_t, W_t, I_t)) = (P_t^{(b)}, N_t^{(b)}, P_t^{(a)}, N_t^{(a)}) \text{ that maximizes } \mathbb{E}[R_T]$$

- Note: Discount Factor when aggregating Rewards in the MDP is 1

Avellaneda-Stoikov Continuous Time Formulation

- We go over the landmark paper by Avellaneda and Stoikov in 2006
- They derive a simple, clean and intuitive solution
- We adapt our discrete-time notation to their continuous-time setting
- $X_t^{(b)}, X_t^{(a)}$ are *Poisson processes* with *hit/lift-rate* means $\lambda_t^{(b)}, \lambda_t^{(a)}$

$$dX_t^{(b)} \sim Poisson(\lambda_t^{(b)} \cdot dt) \text{ , } dX_t^{(a)} \sim Poisson(\lambda_t^{(a)} \cdot dt)$$

$$\lambda_t^{(b)} = f^{(b)}(\delta_t^{(b)}) \text{ , } \lambda_t^{(a)} = f^{(a)}(\delta_t^{(a)}) \text{ for decreasing functions } f^{(b)}, f^{(a)}$$

$$dW_t = P_t^{(a)} \cdot dX_t^{(a)} - P_t^{(b)} \cdot dX_t^{(b)} \text{ , } I_t = X_t^{(b)} - X_t^{(a)} \text{ (note: } I_0 = 0)$$

- Since infinitesimal Poisson random variables $dX_t^{(b)}$ (shares hit in time dt) and $dX_t^{(a)}$ (shares lifted in time dt) are Bernoulli (shares hit/lifted in time dt are 0 or 1), $N_t^{(b)}$ and $N_t^{(a)}$ can be assumed to be 1
- This simplifies the *Action* at time t to be just the pair: $(\delta_t^{(b)}, \delta_t^{(a)})$
- OB Mid Price Dynamics: $dS_t = \sigma \cdot dz_t$ (scaled brownian motion)
- Utility function $U(x) = -e^{-\gamma x}$ where $\gamma > 0$ is coeff. of risk-aversion

Hamilton-Jacobi-Bellman (HJB) Equation

- We denote the Optimal Value function as $V^*(t, S_t, W_t, I_t)$

$$V^*(t, S_t, W_t, I_t) = \max_{\delta_t^{(b)}, \delta_t^{(a)}} \mathbb{E}[-e^{-\gamma \cdot (W_T + I_T \cdot S_T)}]$$

- $V^*(t, S_t, W_t, I_t)$ satisfies a recursive formulation for $0 \leq t < t_1 < T$:

$$V^*(t, S_t, W_t, I_t) = \max_{\delta_t^{(b)}, \delta_t^{(a)}} \mathbb{E}[V^*(t_1, S_{t_1}, W_{t_1}, I_{t_1})]$$

- Rewriting in stochastic differential form, we have the HJB Equation

$$\max_{\delta_t^{(b)}, \delta_t^{(a)}} \mathbb{E}[dV^*(t, S_t, W_t, I_t)] = 0 \text{ for } t < T$$

$$V^*(T, S_T, W_T, I_T) = -e^{-\gamma \cdot (W_T + I_T \cdot S_T)}$$

Converting HJB to a Partial Differential Equation

- Change to $V^*(t, S_t, W_t, I_t)$ is comprised of 3 components:
  - Due to pure movement in time t
  - Due to randomness in OB Mid-Price S_t
  - Due to randomness in hitting/lifting the Bid/Ask
- With this, we can expand $dV^*(t, S_t, W_t, I_t)$ and rewrite HJB as:

$$
\begin{aligned}
\max_{\delta_t^{(b)}, \delta_t^{(a)}} \{ & \frac{\partial V^*}{\partial t} dt + \mathbb{E}[\sigma \frac{\partial V^*}{\partial S_t} dz_t + \frac{\sigma^2}{2} \frac{\partial^2 V^*}{\partial S_t^2} (dz_t)^2] \\
& + \lambda_t^{(b)} \cdot dt \cdot V^*(t, S_t, W_t - S_t + \delta_t^{(b)}, I_t + 1) \\
& + \lambda_t^{(a)} \cdot dt \cdot V^*(t, S_t, W_t + S_t + \delta_t^{(a)}, I_t - 1) \\
& + (1 - \lambda_t^{(b)} \cdot dt - \lambda_t^{(a)} \cdot dt) \cdot V^*(t, S_t, W_t, I_t) \\
& - V^*(t, S_t, W_t, I_t) \} = 0
\end{aligned}
$$

Converting HJB to a Partial Differential Equation

We can simplify this equation with a few observations:

- $\mathbb{E}[dz_t] = 0$
- $\mathbb{E}[(dz_t)^2] = dt$
- Organize the terms involving $\lambda_t^{(b)}$ and $\lambda_t^{(a)}$ better with some algebra
- Divide throughout by dt

$$\max_{\delta_t^{(b)},\delta_t^{(a)}} \{\frac{\partial V^*}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 V^*}{\partial S_t^2}$$
$$+ \lambda_t^{(b)} \cdot (V^*(t, S_t, W_t - S_t + \delta_t^{(b)}, I_t + 1) - V^*(t, S_t, W_t, I_t))$$
$$+ \lambda_t^{(a)} \cdot (V^*(t, S_t, W_t + S_t + \delta_t^{(a)}, I_t - 1) - V^*(t, S_t, W_t, I_t))\} = 0$$

Converting HJB to a Partial Differential Equation

Next, note that $\lambda_t^{(b)} = f^{(b)}(\delta_t^{(b)})$ and $\lambda_t^{(a)} = f^{(a)}(\delta_t^{(a)})$, and apply the max only on the relevant terms

$$
\begin{aligned}
&\frac{\partial V^*}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 V^*}{\partial S_t^2} \\
&+ \max_{\delta_t^{(b)}} \{ f^{(b)}(\delta_t^{(b)}) \cdot (V^*(t, S_t, W_t - S_t + \delta_t^{(b)}, I_t + 1) - V^*(t, S_t, W_t, I_t)) \} \\
&+ \max_{\delta_t^{(a)}} \{ f^{(a)}(\delta_t^{(a)}) \cdot (V^*(t, S_t, W_t + S_t + \delta_t^{(a)}, I_t - 1) - V^*(t, S_t, W_t, I_t)) \} = 0
\end{aligned}
$$

This combines with the boundary condition:

$$V^*(T, S_T, W_T, I_T) = -e^{-\gamma \cdot (W_T + I_T \cdot S_T)}$$

Converting HJB to a Partial Differential Equation

- We make an "educated guess" for the structure of $V^*(t, S_t, W_t, I_t)$:

$$V^*(t, S_t, W_t, I_t) = -e^{-\gamma(W_t + \theta(t, S_t, I_t))} \tag{1}$$

  and reduce the problem to a PDE in terms of $\theta(t, S_t, I_t)$
- Substituting this into the above PDE for $V^*(t, S_t, W_t, I_t)$ gives:

$$\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2}\left(\frac{\partial^2 \theta}{\partial S_t^2} - \gamma\left(\frac{\partial \theta}{\partial S_t}\right)^2\right)$$
$$+ \max_{\delta_t^{(b)}}\left\{\frac{f^{(b)}(\delta_t^{(b)})}{\gamma} \cdot \left(1 - e^{-\gamma(\delta_t^{(b)} - S_t + \theta(t, S_t, I_t + 1) - \theta(t, S_t, I_t))}\right)\right\}$$
$$+ \max_{\delta_t^{(a)}}\left\{\frac{f^{(a)}(\delta_t^{(a)})}{\gamma} \cdot \left(1 - e^{-\gamma(\delta_t^{(a)} + S_t + \theta(t, S_t, I_t - 1) - \theta(t, S_t, I_t))}\right)\right\} = 0$$

- The boundary condition is:

$$\theta(T, S_T, I_T) = I_T \cdot S_T$$

Indifference Bid/Ask Price

- It turns out that $\theta(t, S_t, I_t + 1) - \theta(t, S_t, I_t)$ and $\theta(t, S_t, I_t) - \theta(t, S_t, I_t - 1)$ are equal to financially meaningful quantities known as *Indifference Bid and Ask Prices*
- Indifference Bid Price $Q^{(b)}(t, S_t, I_t)$ is defined as:

$$V^*(t, S_t, W_t - Q^{(b)}(t, S_t, I_t), I_t + 1) = V^*(t, S_t, W_t, I_t) \tag{2}$$

- $Q^{(b)}(t, S_t, I_t)$ is the price to buy a share with *guarantee of immediate purchase* that results in Optimum Expected Utility being unchanged
- Likewise, Indifference Ask Price $Q^{(a)}(t, S_t, I_t)$ is defined as:

$$V^*(t, S_t, W_t + Q^{(a)}(t, S_t, I_t), I_t - 1) = V^*(t, S_t, W_t, I_t) \tag{3}$$

- $Q^{(a)}(t, S_t, I_t)$ is the price to sell a share with *guarantee of immediate sale* that results in Optimum Expected Utility being unchanged
- We abbreviate $Q^{(b)}(t, S_t, I_t)$ as $Q_t^{(b)}$ and $Q^{(a)}(t, S_t, I_t)$ as $Q_t^{(a)}$

Indifference Bid/Ask Price in the PDE for θ

- Express $V^*(t, S_t, W_t - Q_t^{(b)}, I_t + 1) = V^*(t, S_t, W_t, I_t)$ in terms of θ:

$$-e^{-\gamma(W_t - Q_t^{(b)} + \theta(t, S_t, I_t+1))} = -e^{-\gamma(W_t + \theta(t, S_t, I_t))}$$

$$\Rightarrow Q_t^{(b)} = \theta(t, S_t, I_t + 1) - \theta(t, S_t, I_t) \tag{4}$$

- Likewise for $Q_t^{(a)}$, we get:

$$Q_t^{(a)} = \theta(t, S_t, I_t) - \theta(t, S_t, I_t - 1) \tag{5}$$

- Using equations (4) and (5), bring $Q_t^{(b)}$ and $Q_t^{(a)}$ in the PDE for θ

$$\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2}\left(\frac{\partial^2 \theta}{\partial S_t^2} - \gamma\left(\frac{\partial \theta}{\partial S_t}\right)^2\right) + \max_{\delta_t^{(b)}} g(\delta_t^{(b)}) + \max_{\delta_t^{(a)}} h(\delta_t^{(b)}) = 0$$

$$\text{where } g(\delta_t^{(b)}) = \frac{f^{(b)}(\delta_t^{(b)})}{\gamma} \cdot (1 - e^{-\gamma(\delta_t^{(b)} - S_t + Q_t^{(b)})})$$

$$\text{and } h(\delta_t^{(a)}) = \frac{f^{(a)}(\delta_t^{(a)})}{\gamma} \cdot (1 - e^{-\gamma(\delta_t^{(a)} + S_t - Q_t^{(a)})})$$

Optimal Bid Spread and Optimal Ask Spread

- To maximize $g(\delta_t^{(b)})$, differentiate g with respect to $\delta_t^{(b)}$ and set to 0

$$e^{-\gamma(\delta_t^{(b)^*} - S_t + Q_t^{(b)})} \cdot (\gamma \cdot f^{(b)}(\delta_t^{(b)^*}) - \frac{\partial f^{(b)}}{\partial \delta_t^{(b)}}(\delta_t^{(b)^*})) + \frac{\partial f^{(b)}}{\partial \delta_t^{(b)}}(\delta_t^{(b)^*}) = 0$$

$$\Rightarrow \delta_t^{(b)^*} = S_t - Q_t^{(b)} + \frac{1}{\gamma} \cdot \ln{(1 - \gamma \cdot \frac{f^{(b)}(\delta_t^{(b)^*})}{\frac{\partial f^{(b)}}{\partial \delta_t^{(b)}}(\delta_t^{(b)^*})})} \tag{6}$$

- To maximize $g(\delta_t^{(a)})$, differentiate h with respect to $\delta_t^{(a)}$ and set to 0

$$e^{-\gamma(\delta_t^{(a)^*} + S_t - Q_t^{(a)})} \cdot (\gamma \cdot f^{(a)}(\delta_t^{(a)^*}) - \frac{\partial f^{(a)}}{\partial \delta_t^{(a)}}(\delta_t^{(a)^*})) + \frac{\partial f^{(a)}}{\partial \delta_t^{(a)}}(\delta_t^{(a)^*}) = 0$$

$$\Rightarrow \delta_t^{(a)^*} = Q_t^{(a)} - S_t + \frac{1}{\gamma} \cdot \ln{(1 - \gamma \cdot \frac{f^{(a)}(\delta_t^{(a)^*})}{\frac{\partial f^{(a)}}{\partial \delta_t^{(a)}}(\delta_t^{(a)^*})})} \tag{7}$$

- (6) and (7) are implicit equations for $\delta_t^{(b)^*}$ and $\delta_t^{(a)^*}$ respectively

Solving for θ and for Optimal Bid/Ask Spreads

- Let us write the PDE in terms of the Optimal Bid and Ask Spreads

$$\begin{aligned}
&\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2}(\frac{\partial^2 \theta}{\partial S_t^2} - \gamma(\frac{\partial \theta}{\partial S_t})^2) \\
&+ \frac{f^{(b)}(\delta_t^{(b)*})}{\gamma} \cdot (1 - e^{-\gamma(\delta_t^{(b)*} - S_t + \theta(t,S_t,I_t+1) - \theta(t,S_t,I_t))}) \\
&+ \frac{f^{(a)}(\delta_t^{(a)*})}{\gamma} \cdot (1 - e^{-\gamma(\delta_t^{(a)*} + S_t + \theta(t,S_t,I_t-1) - \theta(t,S_t,I_t))}) = 0
\end{aligned} \tag{8}$$

with boundary condition $\theta(T, S_T, I_T) = I_T \cdot S_T$

- First we solve PDE (8) for θ in terms of $\delta_t^{(b)*}$ and $\delta_t^{(a)*}$
- In general, this would be a numerical PDE solution
- Using (4) and (5), we have $Q_t^{(b)}$ and $Q_t^{(a)}$ in terms of $\delta_t^{(b)*}$ and $\delta_t^{(a)*}$
- Substitute above-obtained $Q_t^{(b)}$ and $Q_t^{(a)}$ in equations (6) and (7)
- Solve implicit equations for $\delta_t^{(b)*}$ and $\delta_t^{(a)*}$ (in general, numerically)

Building Intuition

- Define *Indifference Mid Price* $Q_t^{(m)} = \frac{Q_t^{(b)} + Q_t^{(a)}}{2}$
- To develop intuition for Indifference Prices, consider a simple case where the market-maker doesn't supply any bids or asks

$$V^*(t, S_t, W_t, I_t) = \mathbb{E}[-e^{-\gamma(W_t + I_t \cdot S_T)}]$$

- Combining this with the diffusion $dS_t = \sigma \cdot dz_t$, we get:

$$V^*(t, S_t, W_t, I_t) = -e^{-\gamma(W_t + I_t \cdot S_t - \frac{\gamma \cdot I_t^2 \cdot \sigma^2 (T-t)}{2})}$$

- Combining this with equations (2) and (3), we get:

$$Q_t^{(b)} = S_t - (2I_t + 1)\frac{\gamma\sigma^2(T-t)}{2} \text{ , } Q_t^{(a)} = S_t - (2I_t - 1)\frac{\gamma\sigma^2(T-t)}{2}$$

$$Q_t^{(m)} = S_t - I_t\gamma\sigma^2(T-t) \text{ , } Q_t^{(a)} - Q_t^{(b)} = \gamma\sigma^2(T-t)$$

- These results for the simple case of no-market-making serve as approximations for our problem of optimal market-making

Building Intuition

- Think of $Q_t^{(m)}$ as *inventory-risk-adjusted* mid-price (adjustment to S_t)
- If market-maker is long inventory ($I_t > 0$), $Q_t^{(m)} < S_t$ indicating inclination to sell than buy, and if market-maker is short inventory, $Q_t^{(m)} > S_t$ indicating inclination to buy than sell
- Armed with this intuition, we come back to optimal market-making, observing from eqns (6) and (7): ${P_t^{(b)}}^* < Q_t^{(b)} < Q_t^{(m)} < Q_t^{(a)} < {P_t^{(a)}}^*$
- Think of $[{P_t^{(b)}}^*, {P_t^{(a)}}^*]$ as "centered" at $Q_t^{(m)}$ (rather than at S_t), i.e., $[{P_t^{(b)}}^*, {P_t^{(a)}}^*]$ will (together) move up/down in tandem with $Q_t^{(m)}$ moving up/down (as a function of inventory position I_t)

$$Q_t^{(m)} - {P_t^{(b)}}^* = \frac{Q_t^{(a)} - Q_t^{(b)}}{2} + \frac{1}{\gamma} \cdot \ln\left(1 - \gamma \cdot \frac{f^{(b)}({\delta_t^{(b)}}^*)}{\frac{\partial f^{(b)}}{\partial \delta_t^{(b)}}({\delta_t^{(b)}}^*)}\right) \tag{9}$$

$$ {P_t^{(a)}}^* - Q_t^{(m)} = \frac{Q_t^{(a)} - Q_t^{(b)}}{2} + \frac{1}{\gamma} \cdot \ln\left(1 - \gamma \cdot \frac{f^{(a)}({\delta_t^{(a)}}^*)}{\frac{\partial f^{(a)}}{\partial \delta_t^{(a)}}({\delta_t^{(a)}}^*)}\right) \tag{10}$$

Simple Functional Form for Hitting/Lifting Rate Means

- The PDE for θ and the implicit equations for $\delta_t^{(b)*}, \delta_t^{(a)*}$ are messy
- We make some assumptions, simplify, derive analytical approximations
- First we assume a fairly standard functional form for $f^{(b)}$ and $f^{(a)}$

$$f^{(b)}(\delta) = f^{(a)}(\delta) = c \cdot e^{-k \cdot \delta}$$

- This reduces equations (6) and (7) to:

$$\delta_t^{(b)*} = S_t - Q_t^{(b)} + \frac{1}{\gamma} \ln{(1 + \frac{\gamma}{k})} \tag{11}$$

$$\delta_t^{(a)*} = Q_t^{(a)} - S_t + \frac{1}{\gamma} \ln{(1 + \frac{\gamma}{k})} \tag{12}$$

$\Rightarrow P_t^{(b)*}$ and $P_t^{(a)*}$ are equidistant from $Q_t^{(m)}$

- Substituting these simplified $\delta_t^{(b)*}, \delta_t^{(a)*}$ in (8) reduces the PDE to:

$$\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2} (\frac{\partial^2 \theta}{\partial S_t^2} - \gamma (\frac{\partial \theta}{\partial S_t})^2) + \frac{c}{k + \gamma} (e^{-k \cdot \delta_t^{(b)*}} + e^{-k \cdot \delta_t^{(a)*}}) = 0 \tag{13}$$

with boundary condition $\theta(T, S_T, I_T) = I_T \cdot S_T$

Simplifying the PDE with Approximations

- Note that this PDE (13) involves $\delta_t^{(b)^*}$ and $\delta_t^{(a)^*}$
- However, equations (11), (12), (4), (5) enable expressing $\delta_t^{(b)^*}$ and $\delta_t^{(a)^*}$ in terms of $\theta(t, S_t, I_t - 1), \theta(t, S_t, I_t), \theta(t, S_t, I_t + 1)$
- This would give us a PDE just in terms of θ
- Solving that PDE for θ would not only give us $V^*(t, S_t, W_t, I_t)$ but also $\delta_t^{(b)^*}$ and $\delta_t^{(a)^*}$ (using equations (11), (12), (4), (5))
- To solve the PDE, we need to make a couple of approximations
- First we make a linear approx for $e^{-k\cdot\delta_t^{(b)^*}}$ and $e^{-k\cdot\delta_t^{(a)^*}}$ in PDE (13):

$$\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2}\left(\frac{\partial^2 \theta}{\partial S_t^2} - \gamma\left(\frac{\partial \theta}{\partial S_t}\right)^2\right) + \frac{c}{k+\gamma}\left(1 - k \cdot \delta_t^{(b)^*} + 1 - k \cdot \delta_t^{(a)^*}\right) = 0 \quad (14)$$

- Equations (11), (12), (4), (5) tell us that:

$$\delta_t^{(b)^*} + \delta_t^{(a)^*} = \frac{2}{\gamma} \ln\left(1 + \frac{\gamma}{k}\right) + 2\theta(t, S_t, I_t) - \theta(t, S_t, I_t + 1) - \theta(t, S_t, I_t - 1)$$

Asymptotic Expansion of θ in I_t

- With this expression for $\delta_t^{(b)^*} + \delta_t^{(a)^*}$, PDE (14) takes the form:

$$\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2}(\frac{\partial^2 \theta}{\partial S_t^2} - \gamma(\frac{\partial \theta}{\partial S_t})^2) + \frac{c}{k+\gamma}(2 - \frac{2k}{\gamma} \ln(1 + \frac{\gamma}{k}) \\ - k(2\theta(t, S_t, I_t) - \theta(t, S_t, I_t + 1) - \theta(t, S_t, I_t - 1))) = 0 \tag{15}$$

- To solve PDE (15), we consider this asymptotic expansion of θ in I_t:

$$\theta(t, S_t, I_t) = \sum_{n=0}^{\infty} \frac{I_t^n}{n!} \cdot \theta^{(n)}(t, S_t)$$

- So we need to determine the functions $\theta^{(n)}(t, S_t)$ for all $n = 0, 1, 2, \ldots$
- For tractability, we approximate this expansion to the first 3 terms:

$$\theta(t, S_t, I_t) \approx \theta^{(0)}(t, S_t) + I_t \cdot \theta^{(1)}(t, S_t) + \frac{I_t^2}{2} \cdot \theta^{(2)}(t, S_t)$$

Approximation of the Expansion of θ in I_t

- We note that the Optimal Value Function V^* can depend on S_t only through the current *Value of the Inventory* (i.e., through $I_t \cdot S_t$), i.e., it cannot depend on S_t in any other way
- This means $V^*(t, S_t, W_t, 0) = -e^{-\gamma(W_t + \theta^{(0)}(t, S_t))}$ is independent of S_t
- This means $\theta^{(0)}(t, S_t)$ is independent of S_t
- So, we can write it as simply $\theta^{(0)}(t)$, meaning $\frac{\partial \theta^{(0)}}{\partial S_t}$ and $\frac{\partial^2 \theta^{(0)}}{\partial S_t^2}$ are 0
- Therefore, we can write the approximate expansion for $\theta(t, S_t, I_t)$ as:

$$\theta(t, S_t, I_t) = \theta^{(0)}(t) + I_t \cdot \theta^{(1)}(t, S_t) + \frac{I_t^2}{2} \cdot \theta^{(2)}(t, S_t) \qquad (16)$$

Solving the PDE

- Substitute this approximation (16) for $\theta(t, S_t, I_t)$ in PDE (15)

$$\frac{\partial \theta^{(0)}}{\partial t} + I_t \frac{\partial \theta^{(1)}}{\partial t} + \frac{I_t^2}{2} \frac{\partial \theta^{(2)}}{\partial t} + \frac{\sigma^2}{2} (I_t \frac{\partial^2 \theta^{(1)}}{\partial S_t^2} + \frac{I_t^2}{2} \frac{\partial^2 \theta^{(2)}}{\partial S_t^2})$$
$$- \frac{\gamma \sigma^2}{2} (I_t \frac{\partial \theta^{(1)}}{\partial S_t} + \frac{I_t^2}{2} \frac{\partial \theta^{(2)}}{\partial S_t})^2 + \frac{c}{k + \gamma} (2 - \frac{2k}{\gamma} \ln{(1 + \frac{\gamma}{k})} + k \cdot \theta^{(2)}) = 0$$

with boundary condition:

$$\theta^{(0)}(T) + I_T \cdot \theta^{(1)}(T, S_T) + \frac{I_T^2}{2} \cdot \theta^{(2)}(T, S_T) = I_T \cdot S_T \tag{17}$$

- We will separately collect terms involving specific powers of I_t, each yielding a separate PDE:
  - Terms devoid of I_t (i.e., I_t^0)
  - Terms involving I_t (i.e., I_t^1)
  - Terms involving I_t^2

Solving the PDE

- We start by collecting terms involving I_t

$$\frac{\partial \theta^{(1)}}{\partial t} + \frac{\sigma^2}{2} \cdot \frac{\partial^2 \theta^{(1)}}{\partial S_t^2} = 0 \text{ with boundary condition } \theta^{(1)}(T, S_T) = S_T$$

- The solution to this PDE is:

$$\theta^{(1)}(t, S_t) = S_t \tag{18}$$

- Next, we collect terms involving I_t^2

$$\frac{\partial \theta^{(2)}}{\partial t} + \frac{\sigma^2}{2} \cdot \frac{\partial^2 \theta^{(2)}}{\partial S_t^2} - \gamma\sigma^2 \cdot (\frac{\partial \theta^{(1)}}{\partial S_t})^2 = 0 \text{ with boundary } \theta^{(2)}(T, S_T) = 0$$

- Noting that $\theta^{(1)}(t, S_t) = S_t$, we solve this PDE as:

$$\theta^{(2)}(t, S_t) = -\gamma\sigma^2(T - t) \tag{19}$$

Solving the PDE

- Finally, we collect the terms devoid of I_t

$$\frac{\partial \theta^{(0)}}{\partial t} + \frac{c}{k+\gamma}(2 - \frac{2k}{\gamma}\ln{(1+\frac{\gamma}{k})} + k \cdot \theta^{(2)}) = 0 \text{ with boundary } \theta^{(0)}(T) = 0$$

- Noting that $\theta^{(2)}(t, S_t) = -\gamma\sigma^2(T-t)$, we solve as:

$$\theta^{(0)}(t) = \frac{c}{k+\gamma}((2 - \frac{2k}{\gamma}\ln{(1+\frac{\gamma}{k})})(T-t) - \frac{k\gamma\sigma^2}{2}(T-t)^2) \quad (20)$$

- This completes the PDE solution for $\theta(t, S_t, I_t)$ and hence, for $V^*(t, S_t, W_t, I_t)$
- Lastly, we derive formulas for $Q_t^{(b)}, Q_t^{(a)}, Q_t^{(m)}, {\delta_t^{(b)}}^*, {\delta_t^{(a)}}^*$

Formulas for Prices and Spreads

- Using equations (4) and (5), we get:

$$Q_t^{(b)} = \theta^{(1)}(t, S_t) + (2I_t + 1) \cdot \theta^{(2)}(t, S_t) = S_t - (2I_t + 1)\frac{\gamma\sigma^2(T-t)}{2} \quad (21)$$

$$Q_t^{(a)} = \theta^{(1)}(t, S_t) + (2I_t - 1) \cdot \theta^{(2)}(t, S_t) = S_t - (2I_t - 1)\frac{\gamma\sigma^2(T-t)}{2} \quad (22)$$

- Using equations (11) and (12), we get:

$$\delta_t^{(b)*} = \frac{(2I_t + 1)\gamma\sigma^2(T-t)}{2} + \frac{1}{\gamma}\ln\left(1 + \frac{\gamma}{k}\right) \quad (23)$$

$$\delta_t^{(a)*} = \frac{(1 - 2I_t)\gamma\sigma^2(T-t)}{2} + \frac{1}{\gamma}\ln\left(1 + \frac{\gamma}{k}\right) \quad (24)$$

$$\text{Optimal Bid-Ask Spread } \delta_t^{(b)*} + \delta_t^{(a)*} = \gamma\sigma^2(T-t) + \frac{2}{\gamma}\ln\left(1 + \frac{\gamma}{k}\right) \quad (25)$$

$$\text{Optimal ``Mid'' } Q_t^{(m)} = \frac{Q_t^{(b)} + Q_t^{(a)}}{2} = \frac{P_t^{(b)*} + P_t^{(a)*}}{2} = S_t - I_t\gamma\sigma^2(T-t) \quad (26)$$

Back to Intuition

- Think of $Q_t^{(m)}$ as *inventory-risk-adjusted* mid-price (adjustment to S_t)
- If market-maker is long inventory ($I_t > 0$), $Q_t^{(m)} < S_t$ indicating inclination to sell than buy, and if market-maker is short inventory, $Q_t^{(m)} > S_t$ indicating inclination to buy than sell
- Think of $[P_t^{(b)^*}, P_t^{(a)^*}]$ as "centered" at $Q_t^{(m)}$ (rather than at S_t), i.e., $[P_t^{(b)^*}, P_t^{(a)^*}]$ will (together) move up/down in tandem with $Q_t^{(m)}$ moving up/down (as a function of inventory position I_t)
- Note from equation (25) that the Optimal Bid-Ask Spread $P_t^{(a)^*} - P_t^{(b)^*}$ is independent of inventory I_t
- Useful view: $P_t^{(b)^*} < Q_t^{(b)} < Q_t^{(m)} < Q_t^{(a)} < P_t^{(a)^*}$, with these spreads:

$$\text{Outer Spreads } P_t^{(a)^*} - Q_t^{(a)} = Q_t^{(b)} - P_t^{(b)^*} = \frac{1}{\gamma} \ln\left(1 + \frac{\gamma}{k}\right)$$

$$\text{Inner Spreads } Q_t^{(a)} - Q_t^{(m)} = Q_t^{(m)} - Q_t^{(b)} = \frac{\gamma \sigma^2 (T - t)}{2}$$

Real-world Market-Making and Reinforcement Learning

- Real-world OB dynamics are non-stationary, non-linear, complex
- Frictions: Discrete Prices/Sizes, Constraints on Prices/Sizes, Fees
- Need to capture various market factors in the *State* & OB Dynamics
- This leads to Curse of Dimensionality and Curse of Modeling
- The practical route is to develop a simulator capturing all of the above
- Simulator is a *Market-Data-learnt Sampling Model* of OB Dynamics
- Using this simulator and neural-networks func approx, we can do RL
- References: 2018 Paper from University of Liverpool and 2019 Paper from JP Morgan Research
- Exciting area for Future Research as well as Engineering Design