

1 Introduction

We review the fundamentals of the theory of stochastic optimal control and its application to market making in limit order books. We assume some familiarity with Stochastic Calculus (e.g. Ito's lemma and martingales) and the dynamic programming formulation leading to the Hamilton-Jacobi-Bellman (HJB) equations.

The usual setup is to work with controlled diffusion processes of the form,

$$dX_s = b(X_s, \alpha_s)ds + \sigma(X_s, \alpha_s)dW_s$$

and formulate the problem as a dynamic programming equation. For clarity, we motivate the subject with a brief discussion of the classic Merton's Problem.

1.1 Merton's Problem

In the classic Merton's Problem, we have the following,

- A stock price S_t which follows : $dS_t = \mu S_t dt + \sigma S_t dW_t$.
- An agent trades in the stock and cash, holding H_t shares of the stock at time t .
- The cash in the portfolio is $X_t - H_t S_t$.
- The wealth process X_t has the dynamics,

$$\begin{aligned} dX_t &= H_t dS_t + r(X_t - H_t S_t) dt \\ &= H_t (\mu S_t dt + \sigma S_t dW_t) + r(X_t - H_t S_t) dt \\ &= rX_t dt + H_t [(\mu - r)S_t dt + \sigma S_t dW_t] \\ &= rX_t dt + H_t S_t [(\mu - r)dt + \sigma dW_t] \\ &= rX_t dt + \underbrace{H_t S_t}_{\pi_t} \sigma \left[\underbrace{\frac{(\mu - r)}{\sigma}}_{\lambda} dt + dW_t \right] \\ &= rX_t dt + \pi_t \sigma [\lambda dt + dW_t] \end{aligned}$$

where in the above, π_t is the control and λ is the market price of risk for the stock.

- The finite horizon problem is posed as having the agent maximise his expected utility of terminal wealth over the horizon $[0, T]$. The value function for the agent is,

$$u(x) := \sup_{\pi \in \mathcal{A}(x)} \mathbb{E}[U(X_T) | X_0 = x]$$

where x is the initial wealth, $U(\cdot)$ is some utility function and $\mathcal{A}(x)$ is the set of admissible controls.

1.2 The General Setup

In general, the value function is of the form,

$$u(x) := \sup_{\pi \in \mathcal{A}(x)} \mathbb{E} \left[\int_0^T f(s, X_s, \pi_s) ds + F(X_T) \middle| X_0 = x \right]$$

Why is the value function defined as such?

1. We can think of $F(\cdot)$ as quantifying the value of the terminal wealth X_T .
2. For $f(\cdot, \cdot, \cdot)$, we can think of it as quantifying the "instantaneous" utility at every time instance, and the "full reward" for carrying out our strategy is obtained by integrating $f(\cdot, \cdot, \cdot)$ from 0 to T .
3. Of course, since X_T is a stochastic quantity, naturally we would take expectations.
4. Finally, we take the supremum (over the set of admissible strategies) of the expectation.

For example, in Merton's Utility from Terminal Wealth and Intermediate Consumption Problem, the value function is,

$$u(x) := \sup_{(\pi, c) \in \mathcal{A}(x)} \mathbb{E} \left[\int_0^T e^{-\delta t} U_1(c_t) ds + U_2(X_T) \middle| X_0 = x \right]$$

where c_t is consumption at time t and U_1, U_2 are utility functions.

So what the above intuitively says is that, we want to chose optimal controls (π, c) to maximize the expected utility of,

1. Our sum of utility from consumption across time from 0 to t , given by $\int_0^T e^{-\delta t} U_1(c_t) ds$.
2. Our utility of terminal wealth X_T , given by $U_2(X_T)$.

2 Dynamic Programming and the Hamilton-Jacobi-Bellman Equation

We now explore how to go about solving control problems such as the above using the dynamic programming formulation and explore its link to HJB equations. As usual, we work with controlled diffusions of the form,

$$dX_s = b(X_s, \alpha_s) ds + \sigma(X_s, \alpha_s) dW_s$$

And we're interested in the **finite horizon problem**,

$$u(x) := \sup_{\alpha \in \mathcal{A}(x)} \mathbb{E} \left[\int_0^T f(s, X_s, \alpha_s) ds + F(X_T) \middle| X_0 = x \right]$$

where $T \in (0, \infty)$, $\mathcal{A}(x)$ is the set of admissible controls, given $X_0 = x$.

The dynamic programming formulation tackles the problem by considering a starting state (t, x) , and let $\mathcal{A}(t, x)$ be the set of admissible controls with this starting state.

A new notation is also introduced where we will write $(X_s^{t,x})_{s \in (t, T)}$ to denote a path of the controlled diffusion starting at $X_t = x$.

To cater to different starting points, we define an objective functional,

$$J(t, x; \alpha) := \mathbb{E} \left[\int_t^T f(s, X_s^{t,x}, \alpha_s) ds + F(X_T^{t,x}) \middle| X_t = x \right]$$

The objective functional states that, if we start from time t , with initial value of $X_t = x$, and apply the control α (any arbitrary control), the expectation of the terminal reward $F(X_T^{t,x})$ and the ongoing sum of rewards $\int_t^T f(s, X_s^{t,x}, \alpha_s) ds$, is defined as $J(t, x; \alpha)$.

With that we can write the value function starting at time t as,

$$u(t, x) := \sup_{\alpha \in \mathcal{A}(t, x)} J(t, x; \alpha)$$

and using this notation, we can write the original value function starting at $t = 0$ as,

$$u(x) \equiv u(0, x)$$

So far, we have only been introducing additional notation to facilitate the formulation of the problem as a dynamic programming problem.

2.1 The Dynamic Programming Principle

In this section we introduce the dynamic programming principle, which will lead us to the HJB equation involving a PDE on $u(t, x)$. Solving the PDE and maximising it over the space of admissible controls would give us the optimal control.

The dynamic programming principle starts of with comparing two strategies,

- **I** : Using the optimal control $(\hat{\alpha}_s)_{s \in [t, T]}$ over the interval $[t, T]$, versus :
- **II** : Using an arbitrary control $(\alpha_s)_{s \in [t, t+h]}$ over the interval $[t, t+h]$ (where h is a small time interval), and then using the optimal control $(\hat{\alpha}_s)_{s \in [t+h, T]}$.

With the above formulation, it is immediately obvious that strategy **I** is at least as good (since $\hat{\alpha}_s$ is the optimal control over the whole period) as strategy **II**.

Given this, we can write,

$$u(t, x) \geq \mathbb{E} \left[\underbrace{\int_t^{t+h} f(s, X_s^{t,x}, \alpha_s) ds}_{\text{running reward from } t \text{ to } t+h} + \underbrace{u(t+h, X_{t+h}^{t,x})}_{\text{the optimal value from } t+h \text{ to } T} \right]$$

Note that in the above, the inequality results from the term $\mathbb{E} \left[\int_t^{t+h} f(s, X_s^{t,x}, \alpha_s) ds \right]$, due to the fact that it is using an **arbitrary control** α_s . To get equality, we need to maximize the RHS over the space of controls. This means we must write,

$$u(t, x) = \sup_{(\alpha_x)_{s \in [t, t+h]}} \mathbb{E} \left[\underbrace{\int_t^{t+h} f(s, X_s^{t,x}, \alpha_s) ds}_{\text{running reward from } t \text{ to } t+h} + \underbrace{u(t+h, X_{t+h}^{t,x})}_{\text{the optimal value from } t+h \text{ to } T} \right]$$

In the next step, we apply Ito's lemma on $u(t, x)$,

$$\begin{aligned} du &= u_t dt + u_x dX_t + \frac{1}{2} u_{xx} d\langle X \rangle_t \\ &= u_t dt + u_x [b(X_t, \alpha_t) dt + \sigma(X_t, \alpha_t) dW_t] + \frac{1}{2} u_{xx} \sigma^2(X_t, \alpha_t) dt \\ &= u_t dt + b(X_t, \alpha_t) u_x dt + \frac{1}{2} \sigma^2(X_t, \alpha_t) u_{xx} dt + \sigma(X_t, \alpha_t) u_x dW_t \end{aligned}$$

where in the above subscripts for u denotes partial differential, i.e., u_t means $\frac{\partial u}{\partial t}$.

In integral form, we can also write,

$$\begin{aligned} u(t+h, X_{t+h}^{t,x}) &= u(t, x) + \int_t^{t+h} \frac{\partial u}{\partial t}(s, X_s^{t,x}) + \underbrace{b(X_s, \alpha_s) \frac{\partial u}{\partial x}(s, X_s^{t,x}) + \frac{1}{2} \sigma^2(X_s, \alpha_s) \frac{\partial^2 u}{\partial x^2}(s, X_s^{t,x})}_{\text{this is termed } \mathcal{L}^\alpha(\cdot), \text{ the generator of diffusion}} ds \\ &\quad + \int_t^{t+h} \sigma(X_s, \alpha_s) \frac{\partial u}{\partial x}(X_s, \alpha_s) dW_s \\ &= u(t, x) + \int_t^{t+h} \left(\frac{\partial u}{\partial t} + \mathcal{L}^\alpha u \right) (s, X_s^{t,x}) ds + \int_t^{t+h} \sigma(X_s, \alpha_s) \frac{\partial u}{\partial x}(X_s, \alpha_s) dW_s \end{aligned}$$

If we substitute $u(t+h, X_{t+h}^{t,x})$ into the inequality involving $u(t, x)$, the stochastic integral would disappear under the expectation operator, and we get,

$$\begin{aligned} u(t, x) &\geq \mathbb{E} \left[\int_t^{t+h} f(s, X_s^{t,x}, \alpha_s) ds + u(t, x) + \int_t^{t+h} \left(\frac{\partial u}{\partial t} + \mathcal{L}^\alpha u \right) (s, X_s^{t,x}) ds \right] \\ \Rightarrow \mathbb{E} \left[\int_t^{t+h} f(s, X_s^{t,x}, \alpha_s) ds + \left(\frac{\partial u}{\partial t} + \mathcal{L}^\alpha u \right) (s, X_s^{t,x}) ds \right] &\leq 0 \end{aligned}$$

Now we divide by h , move it within the expectation, and let h tends to zero. This gives us,

$$\begin{aligned} f(t, x, \alpha) + \left(\frac{\partial u}{\partial t} + \mathcal{L}^\alpha u \right) (t, x) &\leq 0 \\ \Rightarrow \frac{\partial u}{\partial t} + \mathcal{L}^\alpha u(t, x) + f(t, x, \alpha) &\leq 0 \end{aligned}$$

This inequality result is for any **arbitrary control** α . To get equality, we must have the optimal control $\hat{\alpha}$, with which we can write the **HJB Equation**,

$$\frac{\partial u}{\partial t} + \sup_{\alpha \in \mathcal{A}} [\mathcal{L}^\alpha u(t, x) + f(t, x, \alpha)] = 0$$

In other words, for the optimal $\alpha = \hat{\alpha}$, we have,

$$\frac{\partial u}{\partial t} + \mathcal{L}^{\hat{\alpha}} u(t, x) + f(t, x, \hat{\alpha}) = 0$$

3 Stochastic Control for Counting Process

So far we've worked with diffusions as the driving source of uncertainty in the control problem. In market-making, where-by participants post limit-orders and market-orders, it is also important to incorporate counting process for driving uncertainty in the model.

We will provide a quick introduction to Poisson Counting Processes and see how they fit in to the same dynamic programming framework as we have described above.

3.1 The Poisson Process

A Poisson process is the continuous time analogue of the Bernoulli process with the following properties,

- it is subject to jumps of fixed (or possibly random) size.

- its mean arrival rate over an interval dt is λ . Specifically, we define the probability of k arrivals in time dt as

$$P(k, dt) = \begin{cases} 1 - \lambda dt, & \text{if } k = 0 \\ \lambda dt, & \text{if } k = 1 \\ 0, & \text{if } k > 1 \end{cases}$$

and note that in the above, we mean the above holds in the limit as dt goes to zero. The "arrival rate" λ is the expected number of arrivals per unit time.

Using the above assumption, we consider the probability that there are k arrivals in time interval $t + \Delta t$, denoted by $P(k, t + dt)$. By the law of total probability, we must have the following relation,

$$\begin{aligned} P(k, t + dt) &= \underbrace{P(k, t)}_{k \text{ arrivals in time } t} \times \underbrace{P(0, dt)}_{0 \text{ arrivals in time } dt} + \underbrace{P(k-1, t)}_{k-1 \text{ arrivals in time } t} \times \underbrace{P(1, dt)}_{1 \text{ arrival in time } dt} \\ &= P(k, t) \times (1 - \lambda dt) + P(k-1, t) \times \lambda dt \end{aligned}$$

The next step is to manipulate the above expression into a differential equation.

We begin by re-arranging terms,

$$\begin{aligned} P(k, t + dt) &= P(k, t) \times (1 - \lambda dt) + P(k-1, t) \times \lambda dt \\ P(k, t + dt) &= P(k, t) - P(k, t) \lambda dt + P(k-1, t) \lambda dt \\ P(k, t + dt) - P(k, t) &= \lambda [P(k-1, t) - P(k, t)] dt \end{aligned}$$

From the above, if we go to the limit as dt tends to zero,

$$\begin{aligned} \frac{\partial P(k, t)}{\partial t} &= \lim_{dt \rightarrow 0} \frac{P(k, t + dt) - P(k, t)}{dt} \\ &= \lambda [P(k-1, t) - P(k, t)] \end{aligned}$$

We now show that the following function satisfy the above differential equation,

$$P(k, t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

Differentiating wrt t gives,

$$\begin{aligned} \frac{\partial P(k, t)}{\partial t} &= e^{-\lambda t} \frac{\lambda k (\lambda t)^{k-1}}{k!} - \lambda e^{-\lambda t} \frac{(\lambda t)^k}{k!} \\ &= \lambda \left[e^{-\lambda t} \frac{k (\lambda t)^{k-1}}{k!} - e^{-\lambda t} \frac{(\lambda t)^k}{k!} \right] \\ &= \lambda \left[e^{-\lambda t} \frac{(\lambda t)^{k-1}}{(k-1)!} - e^{-\lambda t} \frac{(\lambda t)^k}{k!} \right] \\ &= \lambda [P(k-1, t) - P(k, t)] \end{aligned}$$

which is as desired.

3.2 The Poisson Counting Process

The simplest Poisson Process is that with a jump size of 1, which is termed the Poisson Counting Process. It is defined as $(N_t)_{t \geq 0}$ with,

- $N_t \geq 0$
- N_t is an integer

- if $s \leq t$ then $N_s \leq N_t$
- $\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$, i.e. the probability that N_t is equal to n is given by the Poisson pdf we derived earlier.
- $\mathbb{E}[N_t] = \lambda t$ and $\text{Var}[N_t] = \lambda t$ (this is standard result from Poisson distribution).
- Importantly, it also has **stationary independent increments**, i.e. for $0 \leq s \leq t$, the following holds,

$$\mathbb{P}(N_{t+h} - N_{s+h} = k) = \mathbb{P}(N_t - N_s = k) = e^{-\lambda(t-s)} \frac{(\lambda(t-s))^k}{k!}$$
 for all $h > 0$.
- Another useful result is that $N_t - \lambda t$ is a martingale.

This is very important because remember in the derivation of the HJB equation, we need to make the stochastic integral wrt to the Wiener Process disappear under the expectation operator. We would need to same for the derivation later on involving the Counting Process.

Since we often need to work with SDEs of stock price involving jump diffusions like the below,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t - \delta dN_t$$

we need to now define what is meant by dN_t .

Recall that, roughly speaking, the probability of one count in an interval of amplitude dt is of order dt (given by λdt earlier), while the probability of more than one count is of higher order. We can translate this into the rule,

$$dN_t = \begin{cases} 1, & \text{if there is a jump in the time interval } (t, t + dt], \\ 0, & \text{otherwise,} \end{cases}$$

$$d\langle N \rangle_t = dN_t, \quad dN_t dt = 0$$

3.3 Ito's Lemma for Jumps

We also need Ito's lemma to work with jumps, and it is but a straight forward extension of the usual Ito's lemma with the addition work of adding back the all the jumps which occurred in between the limits of integration.

Recall that for a continuous process $dX_s = b(X_s, \alpha_s)ds + \sigma(X_s, \alpha_s)dW_s$, Ito's lemma is as follows,

$$\begin{aligned} df(X_t) &= f_x dX_t + \frac{1}{2} f_{xx} d\langle X \rangle_t \\ &= f_x [b(\cdot)dt + \sigma(\cdot)dW_t] + \frac{1}{2} f_{xx} \sigma^2(\cdot)dt \end{aligned}$$

In integral form, we have,

$$X_t = X_0 + \int_0^t f_x b(\cdot)ds + \int_0^t f_x \sigma(\cdot)dW_s + \frac{1}{2} \int_0^t f_{xx} \sigma^2(\cdot)ds$$

Since conceptually, the integral is a sum, and if we add jumps at finite number of points to the process X_t , what we need to extend Ito's lemma is simply to just add back the jumps. Assume the following notation,

$$N_s = N_{s-} + 1 \quad \text{if} \quad dN_s = 1$$

Then the extended Ito's lemma is,

$$\begin{aligned}
X_t &= X_0 + \int_0^t f_x b(\cdot) ds + \int_0^t f_x \sigma(\cdot) dW_s + \frac{1}{2} \int_0^t f_{xx} \sigma^2(\cdot) ds + \sum_{0 \leq s \leq t} [f(X_s) - f(X_{s-})] \\
&= X_0 + \int_0^t f_x b(\cdot) ds + \int_0^t f_x \sigma(\cdot) dW_s + \frac{1}{2} \int_0^t f_{xx} \sigma^2(\cdot) ds + \int_0^t f(X_s) - f(X_{s-}) dN_s
\end{aligned}$$

Therefore, the differential form of Ito's lemma is,

$$df(X_t) = f_x dX_t + \frac{1}{2} f_{xx} d\langle X \rangle_t + [f(X_s) - f(X_{s-})] dN_s$$

Note: The reason for all the above hard work on Ito's lemma for Counting Process is so that we can approach the problem of market making using HJB equations with jumps. For example, one of the simplest way to model the arrival of market orders is to use counting processes associated with the number of market orders up to time t . We will get to that in the next few sections.

4 HJB Formulation for Counting Processes

For the counting process we defined earlier, the assumption was that the intensity, or “arrival rate”, is a constant λ . In actual fact, we can make λ a function of time as well, using $\lambda(t)$,

$$P(k, dt) = \begin{cases} 1 - \lambda(t)dt, & \text{if } k = 0 \\ \lambda(t)dt, & \text{if } k = 1 \\ 0, & \text{if } k > 1 \end{cases}$$

and it can be shown using similar approach as previous that we must have the differential equation,

$$\frac{\partial P(k, t)}{\partial t} = \lambda(t) [P(k-1, t) - P(k, t)]$$

and that the solution is,

$$P(k, t) = \exp\left(-\int_0^t \lambda(u) du\right) \frac{\left(-\int_0^t \lambda(u) du\right)^k}{k!}$$

If N_t is the counting process with the time-dependent intensity $\lambda(t)$, then it is standard result that $M_t = N_t - \int_0^t \lambda(u) du$ is a martingale.

The above result also generalizes to,

- the case when the intensity is a stochastic process λ_t as well as,
- the case when it is a controlled stochastic process $\lambda_t^u = \lambda(t, N_t^u, u)$, where u is the control.

To begin the HJB formulation, we consider again the familiar setup with,

- $(N_t^u)_{0 \leq t \leq T}$ a controlled doubly stochastic Poisson process starting at $N_{0-} = n$.
- N_t^u has intensity $\lambda_t^u = \lambda(t, N_t^u, u_t)$ where u_t is the control.

The value function is,

$$H(n) := \sup_{u \in \mathcal{A}} \mathbb{E} \left[\int_0^T F(s, N_s^u, u_s) ds + G(N_T^u) \right]$$

If u is an **arbitrary control**, the **performance criteria** is,

$$H^u(n) := \mathbb{E} \left[\int_0^T F(s, N_s^u, u_s) ds + G(N_T^u) \right]$$

and the agent seeks to maximise this performance criteria, i.e.,

$$H(n) = \sup_{u \in \mathcal{A}} H^u(n)$$

Next we again introduce the time-indexed notation to facilitate the resolution of the problem using the dynamic programming approach.

We define,

$$H(t, n) = \sup_{u \in \mathcal{A}} H^u(t, n)$$

with

$$H^u(t, n) = \mathbb{E}_{t,n} \left[\int_t^T F(s, N_s^u, u_s) ds + G(N_T^u) \right]$$

where $\mathbb{E}_{t,n}[\cdot]$ represents expectation conditional on $N_t = n$, and u is an arbitrary control when we refer to $H^u(t, n)$.

4.1 The Martingale Optimality Principle

We take a slightly easier route to the derivation of the HJB equation this time round by using the martingale optimality principle.

Consider the process,

$$\left(\underbrace{\int_0^t F(s, N_s^u, u_s) ds}_{\text{using arbitrary control } u \text{ up to time } t} + \underbrace{H(t, N_t)}_{\text{using the optimal control from time } t \text{ onwards}} \right)_{t \in [0, T]}$$

The martingale optimality principle says that the above process is a **super-martingale** for an **arbitrary control** u , and it is a **martingale** for the optimal control \hat{u} .

More formally, let $(R_t)_{t \in [0, t]}$ be defined as,

$$R_t = \int_0^t F(s, N_s^u, u_s) ds + H(t, N_t)$$

By definition, $R_0 = H(0, n)$, which is our original objective $H(0, n) = \sup_{u \in \mathcal{A}} H^u(0, n)$ (see definitions above). Now let us examine R_T , which is,

$$R_T = \int_0^T F(s, N_s^u, u_s) ds + G(N_T^u)$$

We can then write the following inequality,

$$\underbrace{R_0 = H(0, n)}_{\text{the optimal value function using optimal control } \hat{u}} \geq \underbrace{\mathbb{E}[R_T] = \mathbb{E} \left[\int_0^T F(s, N_s^u, u_s) ds + G(N_T^u) \right]}_{\text{the value function using arbitrary control } u}$$

What the above says is that $R_0 \geq \mathbb{E}[R_T]$, meaning R_t is a **super-martingale**.

Equality in the above is achieved when we use the optimal control \hat{u} , in which case R_t is a martingale, i.e., $R_0 = \mathbb{E}[R_T]$. The dynamics of R_t would provide a clue on when R_t would be a martingale.

By Ito,

$$dR_t = F(t, N_t^u, u_t)dt + dH(t, N_t)$$

Using Ito's lemma on $H(t, n)$, we have,

$$dH(t, n) = H_t dt + [H(t, n+1) - H(t, n)] dN_t$$

Continuing, we have,

$$dR_t = F(t, N_t^u, u_t)dt + H_t dt + [H(t, n+1) - H(t, n)] dN_t$$

Note that dN_t in the above is not a martingale, which is problematic. But recall the standard result that $M_t = N_t - \int_0^t \lambda(u)du$ is a martingale. This implies that,

$$dM_t = dN_t - \lambda(t)dt$$

So again continuing,

$$\begin{aligned} dR_t &= F(t, N_t^u, u_t)dt + H_t dt + [H(t, n+1) - H(t, n)] [dM_t + \lambda(t)dt] \\ &= F(t, N_t^u, u_t)dt + H_t dt + \lambda(t) [H(t, n+1) - H(t, n)] dt + \underbrace{[H(t, n+1) - H(t, n)] dM_t}_{\text{local martingale}} \end{aligned}$$

This implies that if we take expectations, the local martingale term disappears and we get,

$$\begin{aligned} \mathbb{E}[dR_t] &= F(t, N_t^u, u_t)dt + H_t dt + \lambda(t) [H(t, n+1) - H(t, n)] dt \\ &= [F(t, N_t^u, u_t) + H_t + \lambda(t) [H(t, n+1) - H(t, n)]] dt \end{aligned}$$

For R_t to be a martingale, we need the entire expression in the brackets to disappear, and we must do this by optimizing over the space of controls $u \in \mathcal{A}$.

This leads us to the HJB equation,

$$\frac{\partial H}{\partial t}(t, n) + \sup_{u \in \mathcal{A}} [F(t, N_t^u, u_t) + \lambda(t) [H(t, n+1) - H(t, n)]] = 0$$

5 Application to Market Making

With all the basics in place, we are now ready to formulate the market making problem in terms of an optimal control problem. We model how a market maker (MM) maximises terminal wealth by trading in and out of positions using limit orders (LOs). The MM provides liquidity to the limit order book (LOB) by posting buy and sell LOs and the control variable is the depth, which is measured from the midprice, at which these LOs are posted. To formalise the problem, we list the relevant variables that we use throughout this section:

- $(S_t)_{0 \leq t \leq T}$ is the midprice, with dynamics $S_t = S_0 + \sigma W_t$, where W_t is a standard brownian motion.
- $(\delta_t^\pm)_{0 \leq t \leq T}$ denote the depth at which the agent posts LOs. Sell LOs are posted at $S_t + \delta_t^+$ and buy LOs at $S_t - \delta_t^-$. For each round-trip of buying and selling, we make the spread $\delta_t^+ + \delta_t^-$. Our PnL depends on making as many of such round trips as possible, while at the same time managing the risk of inventory.

- $(M_t^\pm)_{0 \leq t \leq T}$ denote counting processes corresponding to the arrival of buy (+) and sell (-) market orders (MOs), which arrive at Poisson times with intensity λ^\pm .
- $(N_t^{\delta, \pm})_{0 \leq t \leq T}$ denote ****controlled counting processes**** (not Poisson according to Cartea book, but why??) for the agent's filled sell (+) and buy (-) LOs.
- condition on the arrival of a market order (MO), the posted LOs are filled with probability $e^{-\kappa^\pm \delta_t^\pm}$, with $\kappa^\pm > 0$. This means that the fill rate of LOs is given by $\Lambda_t^{\delta, \pm} = \lambda^\pm e^{-\kappa^\pm \delta_t^\pm}$. Intuitively, λ^\pm is the arrival rate, and also the probability of a jump. This is modulated by the probability of being filled when posting at depth δ_t^\pm , $e^{-\kappa^\pm \delta_t^\pm}$, to give $\Lambda_t^{\delta, \pm}$.
- $(X_t^\delta)_{0 \leq t \leq T}$ denotes the market maker's cash process, and satisfies the SDE,

$$dX_t^\delta = \underbrace{(S_t^- + \delta_t^+) dN_t^{\delta, +}}_{\text{selling high}} - \underbrace{(S_t^- - \delta_t^-) dN_t^{\delta, -}}_{\text{buying low}}$$

- $(Q_t^\delta)_{0 \leq t \leq T}$ denote the agent's inventory process, with $Q_t^\delta = \underbrace{N_t^{\delta, -}}_{\text{qty bought}} - \underbrace{N_t^{\delta, +}}_{\text{qty sold}}$.

5.1 Problem Setup

In this section we assume that the MM seeks the strategy $(\delta_t^\pm)_{0 \leq t \leq T}$ that maximises cash at the terminal date T . We also assume that at time T the MM liquidates her terminal inventory Q_T using an MO at a price which is worse than the midprice to account for liquidity taking fees as well as the MO walking the LOB. Finally, the MM caps her inventory so that it is bounded above by $\bar{q} > 0$ and below by $\underline{q} < 0$, both finite, and also includes a running inventory penalty so that the **performance criterion** is,

$$H^\delta(t, x, S, q) = \mathbb{E}_{t, x, q, S} \left[\underbrace{X_T}_{\text{terminal cash}} + \underbrace{Q_T^\delta (S_T^\delta - \alpha Q_T^\delta)}_{\text{liquidate inventory with penalty at time } T} - \underbrace{\phi \int_t^T (Q_u)^2 du}_{\text{running penalty term for holding inventory}} \right]$$

In the above where $\alpha \geq 0$ represents the fees for taking liquidity (i.e. using an MO) as well as the impact of the MO walking the LOB, and $\phi \geq 0$ is the running inventory penalty parameter. The MM's value function is,

$$H(t, x, S, q) = \sup_{\delta^\pm \in \mathcal{A}} H^\delta(t, x, S, q)$$

As done previously, we use the martingale optimality principle, but this time round, we take addition care to break up the process R_t into its continuous part and the jump part. The process R_t is,

$$R_t = -\phi \int_0^t Q_u^2 du + H(t, X_t, S_t, Q_t)$$

which we break up into the continuous part $R_{t,C}$ (assuming x and q fixed, while t and S vary),

$$R_{t,C} = -\phi \int_0^t Q_u^2 du + H_C(t, \cdot, S_t, \cdot)$$

and the jump part $R_{t,J}$ (assuming S fixed),

$$R_{t,J} = -\phi \int_0^t Q_u^2 du + H_J(t, X_t, \cdot, Q_t)$$

Using Ito on $R_{t,C}$,

$$\begin{aligned} dR_{t,C} &= H_t dt + H_S dS_t + \frac{1}{2} H_{SS} d\langle S \rangle_t \\ &= H_t dt + H_S \sigma dW_t + \frac{1}{2} H_{SS} \sigma^2 dt \\ \mathbb{E}[dR_{t,C}] &= \left[\underbrace{H_t + \frac{1}{2} H_{SS} \sigma^2}_{\text{we need this to be zero to make } R_t \text{ a martingale}} \right] dt \end{aligned}$$

Next we go on to the jump part. The variables x and q are related as follows,

- if a MO hits the ask-quote of the market maker, $q \mapsto q - 1$ and $x \mapsto x + (S + \delta^+)$.
- if a MO hits the bid-quote of the market maker, $q \mapsto q + 1$ and $x \mapsto x - (S - \delta^-)$.

That is to say that Q_t^δ jumps due to two different scenarios, $dN_t^{\delta,+}$ jumping or $dN_t^{\delta,-}$ jumping.

Using Ito on $R_{t,J}$,

$$\begin{aligned} dR_{t,J} &= -\phi q^2 dt + H_t dt + [H(t, x + (S + \delta^+), S, q - 1) - H(t, x, S, q)] dN_t^{\delta,+} \\ &\quad + [H(t, x - (S - \delta^-), S, q + 1) - H(t, x, S, q)] dN_t^{\delta,-} \\ &= -\phi q^2 dt + H_t dt + [H(t, x + (S + \delta^+), S, q - 1) - H(t, x, S, q)] [dM_t^{\delta,+} + \Lambda_t^{\delta,+} dt] \\ &\quad + [H(t, x - (S - \delta^-), S, q + 1) - H(t, x, S, q)] [dM_t^{\delta,-} + \Lambda_t^{\delta,-} dt] \\ \mathbb{E}[dR_{t,J}] &= [-\phi q^2 + H_t + \lambda^+ e^{-\kappa^+ \delta_t^+} [H(t, x + (S + \delta^+), S, q - 1) - H(t, x, S, q)] \\ &\quad + \lambda^- e^{-\kappa^- \delta_t^-} [H(t, x - (S - \delta^-), S, q + 1) - H(t, x, S, q)]] dt \end{aligned}$$

Combining the continuous and jump parts, and recall that R_t is a martingale for the optimal control, this means we have the HJB equation,

$$\begin{aligned} 0 &= H_t + \frac{1}{2} \sigma^2 H_{SS} - \phi q^2 \\ &\quad + \lambda^+ \sup_{\delta^+} \left\{ e^{-\kappa^+ \delta_t^+} [H(t, x + (S + \delta^+), S, q - 1) - H(t, x, S, q)] \right\} \\ &\quad + \lambda^- \sup_{\delta^-} \left\{ e^{-\kappa^- \delta_t^-} [H(t, x - (S - \delta^-), S, q + 1) - H(t, x, S, q)] \right\} \end{aligned}$$

The terminal condition (when $t = T$) is $H(T, x, S, q) = x + q(S - \alpha q)$, since the penalty term vanish.

5.2 Solving the HJB Equation

Since the terminal condition for H is $H(T, x, S, q) = x + q(S - \alpha q) = x + qS - \alpha q^2$, we use the *ansatz* $H(t, x, S, q) = x + qS + g(t, q)$, with the requirement that $g(T, q) = -\alpha q^2$ so that we recover the terminal condition. Next we substitute our *ansatz* into the HJB equation,

$$\begin{aligned} 0 &= g_t - \phi q^2 \\ &\quad + \lambda^+ \sup_{\delta^+} \left\{ e^{-\kappa^+ \delta_t^+} [\delta^+ + g(t, q - 1) - g(t, q)] \right\} \\ &\quad + \lambda^- \sup_{\delta^-} \left\{ e^{-\kappa^- \delta_t^-} [\delta^- + g(t, q + 1) - g(t, q)] \right\} \\ g(T, q) &= -\alpha q^2 \end{aligned}$$

To find the supremum, we take first order conditions wrt δ^+ and δ^- and set them to zero,

$$\begin{aligned}\frac{\partial}{\partial \delta^+} \left[e^{-\kappa^+ \delta_t^+} [\delta^+ + g(t, q-1) - g(t, q)] \right] &= 0 \\ e^{-\kappa^+ \delta_t^+} - \kappa^+ e^{-\kappa^+ \delta_t^+} [\delta^+ + g(t, q-1) - g(t, q)] &= 0 \\ e^{-\kappa^+ \delta_t^+} [1 - \kappa^+ [\delta^+ + g(t, q-1) - g(t, q)]] &= 0 \\ \delta^+ &= \frac{1}{\kappa^+} - [g(t, q-1) - g(t, q)]\end{aligned}$$

The expression for δ^- is similar, and we can summarize by, saying that the optimal control $\delta^{\pm,*}$ is given by,

$$\delta^{\pm,*} = \frac{1}{\kappa^{\pm}} - [g(t, q \mp 1) - g(t, q)]$$

Substituting the expression for the optimal control back into the HJB equation gives,

$$\begin{aligned}0 &= g_t - \phi q^2 \\ &+ \lambda^+ \left\{ e^{-\kappa^+ (\frac{1}{\kappa^+} - [g(t, q-1) - g(t, q)])} \left[\frac{1}{\kappa^+} - [g(t, q-1) - g(t, q)] + g(t, q-1) - g(t, q) \right] \right\} \\ &+ \lambda^- \left\{ e^{-\kappa^- (\frac{1}{\kappa^-} - [g(t, q+1) - g(t, q)])} \left[\frac{1}{\kappa^-} - [g(t, q+1) - g(t, q)] + g(t, q+1) - g(t, q) \right] \right\} \\ &= g_t - \phi q^2 + \frac{\lambda^+}{e\kappa^+} e^{\kappa^+ [g(t, q-1) - g(t, q)]} + \frac{\lambda^-}{e\kappa^-} e^{\kappa^- [g(t, q+1) - g(t, q)]}\end{aligned}$$

We need to solve for $g(t, q)$ to find the final form of the expression for $\delta^{\pm,*}$.

References

- [1] Cartea, Ivaro; Jaimungal, Sebastian; Penalva, Jos. *Algorithmic and High-Frequency Trading*. Cambridge University Press.