Stochastic Control for Dynamic Market Making

Adnan Contractor

1 Introduction

Market making is designed to provide liquidity and facilitate efficient price discovery. Market makers capture the bid-ask spread as revenue while exposing themselves to risks such as inventory imbalances, adverse price movements, and transaction costs. Traditional approaches to market making rely on stationary models, where key variables such as mid-price dynamics, order flow, and market volatility are assumed to exhibit time-invariant statistical properties. While these models are analytically tractable, they fall short in capturing the complexities of modern markets which are driven by non-stationary order flows and abrupt liquidity shocks.

For example, inventory imbalances expose market makers to significant risk if asset prices move adversely. Rebalancing inventories through active trading incurs transaction costs and introduces additional price impacts, particularly in fragmented and high-frequency trading markets. The inherent trade-offs between maximizing short-term profitability and managing long-term risk require market makers to employ dynamic strategies that respond to changing market conditions in real time. This paper develops a stochastic control framework tailored to the realities of non-stationary markets.

I formalize the market-making problem as a stochastic control problem, introducing state variables for mid-price and inventory that evolve according to non-stationary dynamics. I then derive optimal strategies for order placement and inventory management using the HJB equation, addressing both profit maximization and risk mitigation.

2 Framework

The market-making problem is inherently dynamic, requiring the formulation of strategies that respond to evolving market conditions. To model this rigorously, I employ a stochastic control framework where the mid-price of the asset and the inventory held by the market maker are treated as state variables governed by stochastic dynamics. The objective is to maximize expected terminal wealth while managing the risks associated with inventory imbalances and adverse price movements.

Let S_t denote the mid-price of the traded asset at time t, modeled as a stochastic process. The evolution of S_t is governed by the stochastic differential equation:

$$dS_t = \mu_t \, dt + \sigma_t \, dW_t,$$

where μ_t is the drift term capturing the directional trend of the price, σ_t is the volatility term reflecting the uncertainty in price movements, and W_t is a standard Brownian motion. Importantly, the parameters μ_t and σ_t are time-varying and can depend on the state variables S_t and X_t, reflecting the non-stationary nature of the market.

The market maker's inventory X_t evolves based on their trading activity. Let q_t represent the rate at which limit orders are submitted, increasing inventory, and let ξ_t denote the intensity of market orders, which reduce inventory. The dynamics of X_t are then given by:

$$dX_t = q_t \, dt - \xi_t \, dt.$$

The market maker incurs costs associated with inventory risk, price impacts, and transaction fees. Let $c(q_t, \xi_t)$ denote the total cost function, which includes both a convex component representing price impact costs and a linear component representing fixed transaction fees. The market maker's objective is to maximize their expected terminal wealth W_T, which evolves as:

$$dW_t = X_t \, dS_t - c(q_t, \xi_t) \, dt.$$

The value function $V(t, S_t, X_t)$ represents the maximum expected utility of terminal wealth achievable from time t onward, given the state variables S_t and X_t. Formally, this can be expressed as:

$$V(t, S_t, X_t) = \sup_{\{q_u, \xi_u\}_{u \in [t,T]}} \mathbb{E} \left[U(W_T) \mid S_t, X_t \right],$$

where $U(\cdot)$ is the utility function, typically assumed to be concave to reflect risk aversion.

The optimization problem is subject to the dynamics of S_t and X_t, as well as constraints on the control variables q_t and ξ_t. To derive the optimal control policies, we employ the Hamilton-Jacobi-Bellman (HJB) equation, which characterizes the value function $V(t, S_t, X_t)$ as:

$$\frac{\partial V}{\partial t} + \sup_{q_t, \xi_t} \left\{ -c(q_t, \xi_t) + X_t \frac{\partial V}{\partial S_t} \mu_t + \frac{1}{2} \sigma_t^2 \frac{\partial^2 V}{\partial S_t^2} + q_t \frac{\partial V}{\partial X_t} - \xi_t \frac{\partial V}{\partial X_t} \right\} = 0.$$

The first-order conditions with respect to q_t and ξ_t yield the optimal strategies:

$$\frac{\partial c(q_t^*, \xi_t)}{\partial q_t} = \frac{\partial V}{\partial X_t}, \quad \frac{\partial c(q_t, \xi_t^*)}{\partial \xi_t} = -\frac{\partial V}{\partial X_t}.$$

These conditions balance the marginal cost of order placement against the sensitivity of the value function to inventory changes, ensuring that the chosen controls q_t^* and ξ_t^* maximize the expected utility of terminal wealth.

In non-stationary environments, the parameters μ_t and σ_t evolve dynamically, introducing additional complexity into the optimization problem. For example, a sudden increase in volatility σ_t amplifies the risk associated with holding inventory, prompting the market maker to adjust their trading intensity ξ_t accordingly. Similarly, changes in the drift term μ_t influence the expected trajectory of the mid-price, affecting the market maker's inventory management decisions.

3 Optimization and Stochastic Control

The market-making problem is fundamentally a stochastic control problem, where the objective is to dynamically adjust trading strategies to maximize profits while mitigating risks such as inventory imbalances and adverse price movements. The optimization framework relies on the value function $V(t, S_t, X_t)$, which represents the maximum expected utility of terminal wealth achievable from time t onward, given the state variables S_t (the mid-price) and X_t (the inventory level).

The value function is formally defined as:

$$V(t, S_t, X_t) = \sup_{\{q_u, \xi_u\}_{u \in [t,T]}} \mathbb{E} \left[U(W_T) \,\middle|\, S_t, X_t \right],$$

where W_T denotes the terminal wealth, q_u is the rate of limit order placement, ξ_u is the intensity of market orders, and $U(W_T)$ is a concave utility function, reflecting the risk aversion of the market maker.

The dynamics of the state variables S_t and X_t are given by:

$$dS_t = \mu_t \, dt + \sigma_t \, dW_t,$$

$$dX_t = q_t \, dt - \xi_t \, dt,$$

where μ_t and σ_t are the drift and volatility of the mid-price S_t, and q_t and ξ_t control the market maker's inventory dynamics. The wealth process W_t evolves according to:

$$dW_t = X_t\,dS_t - c(q_t, \xi_t)\,dt,$$

where $c(q_t, \xi_t)$ represents the cost function, which includes price impact costs and fixed transaction costs.

The value function satisfies the Hamilton-Jacobi-Bellman (HJB) equation:

$$\frac{\partial V}{\partial t} + \sup_{q_t,\xi_t} \left\{ -c(q_t, \xi_t) + X_t \frac{\partial V}{\partial S_t} \mu_t + \frac{1}{2}\sigma_t^2 \frac{\partial^2 V}{\partial S_t^2} + q_t \frac{\partial V}{\partial X_t} - \xi_t \frac{\partial V}{\partial X_t} \right\} = 0.$$

The optimal controls q_t^* and ξ_t^* are determined by solving the first-order conditions derived from the HJB equation. For q_t, the condition is:

$$\frac{\partial c(q_t, \xi_t)}{\partial q_t} = \frac{\partial V}{\partial X_t},$$

which equates the marginal cost of placing limit orders to the marginal benefit, as measured by the sensitivity of the value function to inventory changes. Similarly, for ξ_t, the condition is:

$$\frac{\partial c(q_t, \xi_t)}{\partial \xi_t} = -\frac{\partial V}{\partial X_t},$$

indicating that the marginal cost of executing market orders must balance the negative sensitivity of the value function to inventory.

Substituting these optimal controls back into the HJB equation yields a fully specified partial differential equation for $V(t, S_t, X_t)$. Solving this equation provides the value function and optimal strategies for the market maker.

In non-stationary environments, the drift μ_t and volatility σ_t evolve dynamically, introducing additional challenges to the optimization problem. For example, consider a scenario where σ_t increases sharply due to market turbulence. This amplifies the second derivative term $\frac{\partial^2 V}{\partial S_t^2}$, increasing the risk of holding inventory and prompting the market maker to reduce X_t through more aggressive market orders ξ_t^*. Conversely, a positive drift $\mu_t > 0$ incentivizes the market maker to increase inventory, exploiting upward price momentum.

The interaction between μ_t and σ_t creates a dynamic trade-off between profit maximization and risk control. Consider a simplified quadratic cost model:

$$c(q_t, \xi_t) = \alpha q_t^2 + \beta \xi_t^2 + \gamma q_t \xi_t,$$

where $\alpha, \beta > 0$ and γ captures cross-effects between limit and market orders. Substituting this cost function into the first-order conditions, we obtain:

$$q_t^* = \frac{\frac{\partial V}{\partial X_t} - \gamma \xi_t}{2\alpha}, \quad \xi_t^* = \frac{-\frac{\partial V}{\partial X_t} - \gamma q_t}{2\beta}.$$

These expressions illustrate how the optimal controls depend on the sensitivity of the value function to inventory and the interaction between limit and market orders.

The resulting stochastic differential equation for X_t becomes:

$$dX_t = \frac{\frac{\partial V}{\partial X_t} - \gamma \xi_t}{2\alpha}\,dt - \frac{-\frac{\partial V}{\partial X_t} - \gamma q_t}{2\beta}\,dt.$$

4 Analytical Solutions

Consider a simplified setting where the cost function $c(q_t, \xi_t)$ is quadratic:

$$c(q_t, \xi_t) = \alpha q_t^2 + \beta \xi_t^2,$$

with no cross-effects between limit and market orders. The mid-price dynamics are modeled as:

$$dS_t = \mu\, dt + \sigma\, dW_t,$$

where μ and σ are constant, and the inventory dynamics are:

$$dX_t = q_t\, dt - \xi_t\, dt.$$

The wealth process evolves as:

$$dW_t = X_t\, dS_t - (\alpha q_t^2 + \beta \xi_t^2)\, dt.$$

The value function $V(t, S_t, X_t)$ satisfies the Hamilton-Jacobi-Bellman (HJB) equation:

$$\frac{\partial V}{\partial t} + \sup_{q_t, \xi_t} \left\{ X_t \mu \frac{\partial V}{\partial S_t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S_t^2} + q_t \frac{\partial V}{\partial X_t} - \xi_t \frac{\partial V}{\partial X_t} - (\alpha q_t^2 + \beta \xi_t^2) \right\} = 0.$$

The first-order conditions for q_t and ξ_t are:

$$\frac{\partial}{\partial q_t} \left(q_t \frac{\partial V}{\partial X_t} - \alpha q_t^2 \right) = 0, \quad \frac{\partial}{\partial \xi_t} \left(-\xi_t \frac{\partial V}{\partial X_t} - \beta \xi_t^2 \right) = 0.$$

Solving these conditions yields the optimal controls:

$$q_t^* = \frac{\frac{\partial V}{\partial X_t}}{2\alpha}, \quad \xi_t^* = \frac{-\frac{\partial V}{\partial X_t}}{2\beta}.$$

Substituting q_t^* and ξ_t^* into the HJB equation, we obtain:

$$\frac{\partial V}{\partial t} + X_t \mu \frac{\partial V}{\partial S_t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S_t^2} + \frac{1}{4\alpha} \left(\frac{\partial V}{\partial X_t} \right)^2 + \frac{1}{4\beta} \left(\frac{\partial V}{\partial X_t} \right)^2 = 0.$$

To simplify further, assume that the value function is separable in its dependence on S_t and X_t:

$$V(t, S_t, X_t) = V_1(t, S_t) + V_2(t, X_t).$$

The term $V_1(t, S_t)$ captures the mid-price dynamics and satisfies:

$$\frac{\partial V_1}{\partial t} + \mu X_t \frac{\partial V_1}{\partial S_t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V_1}{\partial S_t^2} = 0.$$

The term $V_2(t, X_t)$ reflects inventory dynamics and satisfies:

$$\frac{\partial V_2}{\partial t} + \frac{1}{4\alpha} \left(\frac{\partial V_2}{\partial X_t} \right)^2 + \frac{1}{4\beta} \left(\frac{\partial V_2}{\partial X_t} \right)^2 = 0.$$

When the utility of terminal wealth $U(W_T)$ is quadratic, i.e., $U(W_T) = -\gamma W_T^2$ for some risk-aversion parameter $\gamma > 0$, the separable structure allows explicit solutions. For $V_1(t, S_t)$, the solution is well-known from the Black-Scholes framework:

$$V_1(t, S_t) = A(t) + B(t)S_t + C(t)S_t^2,$$

where $A(t), B(t), C(t)$ are functions of time that solve a system of ordinary differential equations (ODEs).

For $V_2(t, X_t)$, we transform the HJB equation using $p = \frac{\partial V_2}{\partial X_t}$, yielding:

$$\frac{\partial p}{\partial t} = -\frac{1}{2\alpha}p^2 - \frac{1}{2\beta}p^2.$$

The solution for p determines the inventory sensitivity, allowing us to recover $V_2(t, X_t)$ by integration:

$$V_2(t, X_t) = \int p(t, X_t)\, dX_t.$$

These solutions highlight the dependence of optimal strategies on the market maker's risk tolerance, price dynamics, and the cost structure. For instance, higher values of β (market order costs) incentivize greater reliance on limit orders, while higher volatility σ reduces the desirability of large inventories.

What are the implications for market making strategies?

1. The quadratic dependence of inventory risk on X_t implies that market makers should maintain inventories close to zero in high-volatility environments to avoid excessive risk exposure.
2. A positive drift $\mu > 0$ motivates market makers to build inventory positions, capitalizing on expected price appreciation, while negative drift $\mu < 0$ necessitates rapid inventory reduction.
3. The balance between q_t^* and ξ_t^* is governed by the relative costs α and β, with limit orders preferred when α is small and market orders dominating when immediate execution is critical.

5 Extensions

A candidate *multi-asset market-making problem* introduces additional state variables, representing the prices and inventories of multiple correlated assets. Let $S_t = (S_t^{(1)}, S_t^{(2)}, \dots, S_t^{(n)})^\top$ denote the vector of mid-prices for n assets, and $X_t = (X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(n)})^\top$ the corresponding inventory vector. The price dynamics are modeled as a multivariate stochastic process:

$$dS_t = \mu_t\, dt + \Sigma_t\, dW_t,$$

where $\mu_t \in \mathbb{R}^n$ is the vector of drifts, $\Sigma_t \in \mathbb{R}^{n\times n}$ is the time-varying volatility matrix, and $W_t \in \mathbb{R}^n$ is an n-dimensional standard Brownian motion. The correlation structure of the assets is embedded in Σ_t, with off-diagonal elements reflecting interdependencies.

The inventory dynamics generalize to:

$$dX_t^{(i)} = q_t^{(i)}\, dt - \xi_t^{(i)}\, dt, \quad i = 1, 2, \dots, n,$$

where $q_t^{(i)}$ and $\xi_t^{(i)}$ represent the limit order rate and market order intensity for asset i, respectively. The wealth process evolves as:

$$dW_t = \sum_{i=1}^{n} X_t^{(i)}\, dS_t^{(i)} - \sum_{i=1}^{n} c^{(i)}(q_t^{(i)}, \xi_t^{(i)})\, dt,$$

where $c^{(i)}(q_t^{(i)}, \xi_t^{(i)})$ is the cost function for asset i.

The value function $V(t, S_t, X_t)$ now represents the maximum expected utility of terminal wealth across all assets:

$$V(t, S_t, X_t) = \sup_{\{q_u, \xi_u\}_{u\in[t,T]}} \mathbb{E}\left[U(W_T) \mid S_t, X_t\right].$$

The corresponding Hamilton-Jacobi-Bellman (HJB) equation is:

$$\frac{\partial V}{\partial t} + \sup_{q_t, \xi_t} \left\{ \sum_{i=1}^{n} \left[X_t^{(i)} \mu_t^{(i)} \frac{\partial V}{\partial S_t^{(i)}} + \frac{1}{2} \sum_{j=1}^{n} \Sigma_t^{(i,j)} \Sigma_t^{(i,j)} \frac{\partial^2 V}{\partial S_t^{(i)} \partial S_t^{(j)}} + q_t^{(i)} \frac{\partial V}{\partial X_t^{(i)}} - \xi_t^{(i)} \frac{\partial V}{\partial X_t^{(i)}} - c^{(i)}(q_t^{(i)}, \xi_t^{(i)}) \right] \right\} = 0.$$

In multi-asset markets, volatility shocks often propagate across assets due to shared risk factors. To capture this, the volatility matrix Σ_t can be modeled using a stochastic factor process:

$$d\Sigma_t = f(S_t, X_t, \Theta_t)\, dt + g(S_t, X_t, \Theta_t)\, dZ_t,$$

where Θ_t represents external risk factors, and Z_t is an independent Brownian motion. The interaction between Σ_t and inventory X_t induces cross-asset rebalancing strategies, where excess inventory in one asset prompts offsetting trades in correlated assets.

Optimal strategies q_t^* and ξ_t^* now depend not only on individual asset dynamics but also on their correlations. For example, the first-order conditions for $q_t^{(i)}$ and $\xi_t^{(i)}$ are:

$$\frac{\partial c^{(i)}(q_t^{(i)}, \xi_t^{(i)})}{\partial q_t^{(i)}} = \frac{\partial V}{\partial X_t^{(i)}}, \quad \frac{\partial c^{(i)}(q_t^{(i)}, \xi_t^{(i)})}{\partial \xi_t^{(i)}} = -\frac{\partial V}{\partial X_t^{(i)}}.$$

One may also consider a venue-specific cost function $c_v(q_t^{(i)}, \xi_t^{(i)})$, indexed by venue v, with constraints:

$$\sum_v q_{t,v}^{(i)} \leq Q_t^{(i)}, \quad \sum_v \xi_{t,v}^{(i)} \leq \Xi_t^{(i)},$$

where $Q_t^{(i)}$ and $\Xi_t^{(i)}$ are the total allowable limit and market orders for asset i across all venues. The optimization problem extends to:

$$\sup_{q_t, \xi_t} \mathbb{E}\left[\int_0^T \sum_{i=1}^n \left(X_t^{(i)} \mu_t^{(i)} - \sum_v c_v(q_{t,v}^{(i)}, \xi_{t,v}^{(i)}) \right) dt \right].$$

In high-frequency trading, intraday patterns such as periodic liquidity surges or market close effects significantly impact optimal strategies. Let $h(t)$ represent a time-varying weight capturing these effects. The objective function becomes:

$$\sup_{q_t, \xi_t} \mathbb{E}\left[\int_0^T h(t) \sum_{i=1}^n \left(X_t^{(i)} \mu_t^{(i)} - c^{(i)}(q_t^{(i)}, \xi_t^{(i)}) \right) dt \right].$$

The optimal controls incorporate these temporal factors, with higher activity during periods of increased liquidity and reduced trading near illiquid intervals. This requires solving the HJB equation with time-dependent coefficients, which is hopefully the subject of future work!