Optimizing Execution Cost Using Stochastic Control

A Thesis Submitted
In Partial Fulfillment of the Requirements
for the Degree of
Master of Technology

by

Akshay Bansal

to the

Sampling and Official Statistics Unit
Indian Statistical Institute, Kolkata
June 2017

CERTIFICATE

It is certified that the work contained in the thesis titled ***Optimizing Execution Cost Using Stochastic Control*** by ***Akshay Bansal*** has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Dr. Diganta Mukherjee
Sampling and Official Statistics Unit
ISI Kolkata

June 2017

"I am tomorrow, or some future day, what I establish today. I am today what I established yesterday or some previous day."

James Joyce

INDIAN STATISTICAL INSTITUTE, KOLKATA

Abstract

Dr. Diganta Mukherjee
Sampling and Official Statistics Unit

Master of Technology

OPTIMIZING EXECUTION COST USING STOCHASTIC CONTROL

by Akshay Bansal

In this work, we devise an optimal allocation strategy for the execution of a predefined no. of stocks in a given time period using the technique of discrete-time Stochastic Control Theory for two different market models. The market model-I (MM-I) which allows an instant execution of *market* orders has been analyzed by assuming geometric Brownian motion of the stock prices for two different cost functions where the first function involves just the fiscal cost while the cost function of the second kind incorporates market risks along with fiscal costs. Subsequently, we improvise an investment strategy for the delayed stock execution (MM-II) and compare the performance of the resulting policies with some of the commonly used execution strategies.

Acknowledgements

I'd like to thank my supervisor **Dr. Diganta Mukherjee** for introducing me to the theory of *Stochastic Control* and developing my intuitive mindset to understand some of the more sophisticated concepts of mathematical finance. His constant motivation and support during this one year has left me with a lasting impression to keep improving upon the existing research.

It's difficult to imagine an academically fruitful passage of the last two years consisting of numerous crests and trenches without the help and support of my family members and hostel mates whose consortium brought instances of conducive learning along with some unforgettable moments to cherish and remember.

Akshay

Contents

List of Tables

Dedicated to my sister Kanika

Chapter 1

Introduction

The search for an 'optimal' action given the current state (or total information) in a discrete-time setup for processes involving some degree of stochasticity has led to the evolution of numerous data-driven as well as data-independent techniques with the precise definition of 'optimal' being highly subjective in general. The stochastic process we'll be looking forward to involves the time evolution of stock prices which an investor encounters while devising a useful strategy to buy a pre-defined block of shares in a given time duration. This stock execution problem is highly correlated with the fundamental difficulty of forecasting stock prices as a practical solution to any one of these two would bring some insight to solve the other. Investors and professional analysts frequently try to model stock prices with the help of the available information and certain *noise* factors whose distribution depends on various market aspects such as inflationary rates, financial status of the company and its competitive workforce. Therefore, any attempt to trade a given block of shares would result in a decision (policy) whose final investment cost attributes its dependence to this probabilistic behavior of stock prices with its precise mathematical distribution still lingering in uncertainty. Formally, the problem statement can be redefined as

Problem 1. *Given an investor with an obligation to buy K no. of shares (market order) in a time frame of T units, devise a trading strategy such that the investor's total monetary investment is minimized.*

Assuming that one proposes a strategy that fares better than the naive ones for most of the cases, then the policy can be executed in High Frequency Trading (HFT) markets where the role of an investor is replaced by a finite-time Turing Machine except for the fact that the no. of transactions made in the same duration has now increased in leaps as the trading can be electronically automated leading to an even smaller investment cost.

Mathematically, Problem 1 can be reformulated as:

Problem 2. *Determine a cost-efficient strategy (policy)*

$$\boldsymbol{\pi^*} = \{\mu_0^*(x_0, R_0), \mu_1^*(x_1, R_1) \dots \mu_N^*(x_N, R_N)\}$$

such that

$$\begin{aligned} x_{k+1} &= g(x_k, u_k, \epsilon_k)\, \forall k \in \mathbb{Z} \\ u_k &= \mu_k^*(x_k, R_k)\, \forall k \in \{0, 1, \dots, N\} \\ \sum_{r=0}^{N} u_r &= K \end{aligned}$$

where $g(x, R, \epsilon)$ is a known function which updates itself at each of the N equispaced time points in the time duration T, R_k is the stock position held at time point t_k and x_k is the stock price at time t_k.

Bertsimas & Lo[1] devised one such policy by partitioning the entire time frame into N intervals of equal length and performing the transaction of buying K/N shares at the start of each interval. In order to analyze the expected investment cost of such policy, Bertsimas utilized the discrete form of Arithmetic Brownian Motion (ABM) ($x_t = x_{t-1} + h(u_t) + \eta\epsilon_t$) to periodically update the stock price. The major drawback with ABM model for stock price updation is that the non-negative behavior of stock price prevails only for shorter time frames T and the resultant optimal action (no. of shares bought out of the remaining stock pool) at each transaction point remained independent of any current/previous state information. Almgren & Chriss[2] extended the Bertismas' model for *limit order* markets by incorporating the variance associated with the execution shortfall in the objective function. More recently, application of some of data-driven statistical techniques based on Reinforcement Learning[3] by Kakade et al.[4] and Nevmyvaka et al.[5] have resulted in significant improvement over simpler execution strategies such as *submit and leave*. In 2014, Cont & Kukanov[6] developed a more generalized mathematical framework for optimal order execution in *limit order* markets by incorporating targeted execution size due to bounded execution capacity of limit orders

In this work, we've introduced the preliminary concepts of discrete-time control theory (with perfect state information) for two possible reductions in the form of Infinite and Finite-horizon problems (Chapter 2) followed by its application to two different market scenarios (MM-I and MM-II) where we've analyzed the resultant action using fiscal cost and its combination with market risks (Chapter 3 & 4). Conclusively, we've compared the performance of our output with some commonly used strategies and provided a few suggestions to improve upon the existing results.

Chapter 2

Optimal Policy Formulation Using Stochastic Control Theory

The uncertainty factor (ϵ) involved in the state-updation function leads us to one such pathway of determining a cost-efficient policy (satisfying the conditions of (2)) by minimizing the expected future cost leading to the application of well-established theory of Stochastic Control. Mathematically, the exact optimization problem reduces to determining optimal policy $\boldsymbol{\pi^*} = \{\mu_0^*(x_0, R_0), \mu_1^*(x_1, R_1) \ldots \mu_N^*(x_N, R_N)\}$ for the objective

$$\min_{\{\pi\}} \mathbb{E}_0[\sum_{r=0}^{N} u_r x_r] \tag{2.1}$$

Subject to the conditions:

$$\begin{aligned} u_k &= \mu_k(x_k, R_k) \,\forall k \in \{0, 1, \ldots, N\} \\ R_{k+1} &= R_k - u_k \\ x_{k+1} &= g(x_k, u_k, \epsilon_k) \forall k \\ \sum_{r=0}^{N} u_r &= K \\ u_k &\geq 0 \;\forall k \end{aligned} \tag{2.2}$$

where x_k is the stock price at time point t_k, R_k is the stock position held at time t_k and u_k is the appropriate action (investment strategy).

The two different discrete-time formulations for determining optimal action using Stochastic Control in order to minimize the expected future cost (monetary) with respect to the execution of *market orders* are prescribed in the following sections:

2.1 Reduction to Infinite horizon problem

The approximation of the finite discrete-time objective (2.1) to an Infinite horizon problem implies the following deduction provided $\beta = 1 - \epsilon$ where $\epsilon \to 0^+$

$$\min_{\{\mu\}} \mathbb{E}_0[\sum_{r=0}^{\infty} \beta^r u_r x_r] \implies \min_{\{\pi\}} \mathbb{E}_0[\sum_{r=0}^{N} u_r x_r] \tag{2.3}$$

On application of Bellman's principle of optimality for discrete-case Infinite horizon problem[7], the optimal policy of objective function (*LHS* of (2.3)) can be determined by solving the functional equation (2.4) using some of the simplified class of functions for $\mu(x, R)$ such as variable separability/lower order polynomial forms.

$$V(x, R) = x.\mu(x, R) + \beta\mathbb{E}[V(g(x, \mu(x, R), \epsilon), R - \mu(x, R))] \tag{2.4}$$

where $V(x, R)$ is the optimal value of $\mathbb{E}_0[\sum_{r=0}^{\infty} \beta^r u_r x_r]$ corresponding to the optimal policy $\mu(x, R)$.

2.1.1 Pitfalls of reduction to Infinite horizon case

1. Solving the multivariate functional equation (2.4) even with some of the simplified assumptions for the functional form of $\mu(x, R)$ is quite difficult.
2. The sufficient conditions for the convergence of infinite series given by *LHS* of (2.3) are not known in general.
3. The optimal actions predicted by the resultant policy $\mu(x, R)$ would in general be non-integers (α-approximate policy).

2.2 Finite horizon problem for integral states

Let the uniform partition $\Pi(T) = \{t_0, t_1, \ldots, t_N\}$ be given with X being the finite set of all possible stock prices and $P = \{r \in \mathbb{Z}^+ \mid r \leq K\}$ the set of all possible stock positions. Then at any given time point $t \in \Pi(T)$, the state vector $(x, R) \in X \times P$. If the function $f(x, u, R)$ computes the *instantaneous* cost for the current state (x, R) and action u, the optimal policy ($\boldsymbol{\pi^*} = \{\mu_0^*(x_0, R_0), \mu_1^*(x_1, R_1) \ldots \mu_N^*(x_N, R_N)\}$) for the objective function (2.1) can be computed dynamically for each discrete time point using Bellman's principle of optimality[8]. Precisely, to determine the time t_k policy function $\mu_k^*(x_k, R_k)$, optimal action u_k^{opt} is tabulated as a function of all $(x_k, R_k) \in X \times P$ using the *adaptive* cost objective

$$J_k(x_k, R_k) = \min_{\{u_k\}} \sum_{i}^{\infty} \Pr(\epsilon_k^i | \mathcal{F}_k)[f(x_k, u_k, R_k) + J_{k+1}(g(x_k, u_k, \epsilon_k^i), R_k - u_k)] \tag{2.5}$$

where $\mathcal{F}_k$ is the t_k-filtration (information contained till time t_k).

At the final time point t_N, the optimal action would be to buy all the remaining R_N. Thus $J_N(x_N, R_N)$ simply reduces to $f(x_N, R_N, R_N)$.

If the uncertainty parameter (ϵ_k) is independent of information $\mathcal{F}_k$, then (2.5) further simplifies to

$$J_k(x_k, R_k) = \min_{\{u_k\}} \sum_{i}^{\infty} \Pr(\epsilon_k^i)[f(x_k, u_k, R_k) + J_{k+1}(g(x_k, u_k, \epsilon_k^i), R_k - u_k)] \tag{2.6}$$

At any time point $t_k \in \Pi(T)$, the optimal action u_k^{opt} and $J_k(x_k, R_k)$ can be dynamically computed using (2.6) for each of the state element $(x_k, R_k) \in X \times P$.

2.2.1 Pitfalls of reduction to integral finite horizon case

1. The numerical algorithm for its implementation mandates the construction of a three-dimensional matrix where each two-dimensional sub-matrix corresponds to a unique time point. Therefore its space complexity is of the order $\Omega(x_{max} K N)$.

2. The method imposes an additional restraint of the finiteness and countability of the set of all possible states $(X \times P)$.

3. The numerical search for the optimal integral solution can at best be accomplished using branch and bound algorithm[9] whose worst case complexity is still K (initial stock position). Thus the eventual time complexity for this algorithm is $\Omega(x_{max}K^2N)$.

Chapter 3

Optimal Investment Strategy for Market Model-I

In this chapter, we'll develop an investment strategy based on the idea of Stochastic Control Theory discussed briefly in Chapter 2 for the market structure of first kind which sanctions the investor to buy any number of stocks on an instant basis at the current *market price* (mid-point of bid-ask spread). Unlike Almgren & Chriss [see 2], we've modelled stock prices realistically using Geometric Brownian Motion (discrete form of Ito's differential) as the Bachelier's model ($x_t = x_{t-1} + h(u_t) + \eta\epsilon_t$) would eventually return negative stock prices with non-zero probability in the limit of longer time duration. The discrete time stock price model we intend to use in our analysis is given by:

$$x_{k+1} = x_k(1 + \beta u_k + \epsilon_k) \tag{3.1}$$

here x_{t+1} is the stock price at time t_{k+1}, ϵ_k is a random noise with $\mathbb{E}[\epsilon_t] = 0$ and βu_k is the drift in stock price due to the buying action of u_k no. of stocks with β being some kind of *prominence factor* which varies according to one's influence in the stock market. For our case, we'll assume β belonging to the range $[10^{-5}, 10^{-4}]$. In the next two sections, we'll establish the general nature of some of the investment strategies for different kinds of *instantaneous* cost functions and compare their performance with some well-established policies.

3.1 Allocation Policy for Fiscal Cost Function

For this particular case, the *instantaneous* cost function is exclusively monetary i.e $f(x_k, u_k, R_k)$ (Section 2.2) is simply given by

$$f(x_k, u_k, R_k) = x_k u_k \tag{3.2}$$

where $u_k \leq R_k$.

Accordingly, the expression for optimal expected cost (2.5) modifies to

$$J_k(x_k, R_k) = \min_{\{u_k\}} [x_k u_k + \sum_{i}^{\infty} Pr(\epsilon_k^i) J_{k+1}(g(x_k, u_k, \epsilon_k^i), R_k - u_k)] \tag{3.3}$$

On rewriting the above expression for penultimate time point ($t = t_{N-1}$) by modelling the stock price using GBM, the objective simplifies to

$$\begin{aligned} J_{N-1}(x_{N-1}, R_{N-1}) = \min_{\{u_{N-1}\}} & [x_{N-1}u_{N-1} + x_{N-1}(R_{N-1} - u_{N-1})(1 \\ & + \beta u_{N-1})] (\because \mathbb{E}[\epsilon_{N-1}] = 0, J_N(x_N, R_N) = x_N R_N) \end{aligned} \tag{3.4}$$

leading to the following deduction.

Deduction 1. *When the nature of the instantaneous cost function is completely fiscal i.e.* $f(x, u, R) = xu$ *and the stock price is modeled using 3.1, the optimal investment policy due to stochastic control (Problem 2.1) simply converges to the purchase of the entire stock block of size* K *at time* $t = t_N$*. In general, the result holds for any stock price updation function of the form* $x_{t+1} = x_t(1 + h(u_t) + \epsilon_t)$ *where* $h(u_t)$ *is a non-decreasing drift with* $h(0) = 0$

Proof. On rearranging the terms of penultimate time objective for the drift $h(u)$, (3.4) modifies to

$$J_{N-1}(x_{N-1}, R_{N-1}) = \min_{\{u_{N-1}\}} [x_{N-1}R_{N-1} + x_{N-1}(R_{N-1} - u_{N-1})h(u_{N-1})] \tag{3.5}$$

As $(R_{N-1} - u_{N-1})h(u_{N-1}) \geq 0$, the optimal action (u^{opt}_{N-1}) results in zero with $J_{N-1}(x_{N-1}, R_{N-1}) = x_{N-1}R_{N-1}$. By recursively calculating u^{opt}_k and $J_k(x_k, R_k)$ using the functional form of $J_{k+1}(x_{k+1}, R_{k+1})$ (3.3), it's trivial to observe the identical nature of the objective function for all $0 \leq k \leq N-1$. Hence the above deduction follows. □

3.1.1 Resultant policy and its comparison with Bertsimas' model

Deduction 1 can be further generalized by observing the degenerate nature of the objective function at the penultimate time point i.e. both 0 and R_{N-1} are the optimal solutions to the objective 3.5. Henceforth, the optimal allocation policy modifies to the total investment for the entire stock block (K) at any one of the time point $t \in \{t_0, t_1, \ldots, t_N\}$.

Tabulated below is the total expenditure resulting from Bertsimas' policy and one-time investment at the midpoint $T/2$.[1]

Stock	Investment Cost(B)	Investment Cost(OT)	Ratio(OT:B)
GOOG	$719770.69	$738000	1.02532
AAPL	$97670.42	$106636.21	1.09179
QCOM	$48983.12	$48808.40	0.99643
NVDA	$36247.86	$35704.41	0.98500
LXS.DE	€39972.39	€40986.76	1.02537

TABLE 3.1: Comparison of total expenditure between Bertsimas'(B) and One-Time(OT) policy based on their daily opening price spanning a total of 100 working days (Feb'16 - Jun'16)

As evident from the data above, the one-time investment policy may frequently fail to perform better than the distributed investment policy (due to Bertsimas).

3.2 Allocation Policy for Constrained Cost Function

Due to the possibility of positive accumulation of random noise (ϵ_t) over large no. of discrete time steps, the allocation policy devised in the last section has a tendency of resulting in a greater investment cost compared to the policy of distributed trading

[1] As per the stock data obtained from Yahoo Finance

over the same no. of time steps. Thus, we've made an attempt to modify the *instantaneous* cost $f(x, u, R)$ by incorporating non-negative penalty in addition to the fiscal cost if the current action (u_k) violates certain market specific bounds. Specifically, a pre-determined set of bounds - an upper bound (UB) and a lower bound (LB) restricts the fractional consumption (u_k/R_k) at every time point t_k. The effect of penalty imposed for the case when the fractional consumption goes below the lower bound (LB) is less pronounced at initial time points compared to the later ones as the opportunistic time window to minimize the total expenditure decreases gradually with the passage of another transaction opportunity. The non-existence of such a restriction would eventually result in the investor holding a large fraction of his initial stock position at later time points with fewer opportunities to improve his total investment cost. Similarly, by restricting the investor to buy a large fraction of his current stock position (exceeding the upper bound (UB)) at the earlier time points of the transaction window, one instructs the investor to employ a distributed investment strategy till the near end of the transaction window where this constraint is liberalized. Mathematically, these two kind of restrictions can be summarized by modifying *instantaneous* cost $(f(x, u, R))$ using the logarithmic barrier[10] resulting in the functional form:

$$\begin{aligned} f(x_k, u_k, R_k) = x_k u_k - x_k C_l \Big(\frac{t_k}{t_N}\Big)^{\gamma} \log\Big(1 - \max(0, LB - \frac{u_k}{R_k})\Big) \\ - x_k C_u \Big(\frac{t_N}{t_k}\Big)^{\gamma} \log\Big(1 - \max(0, \frac{u_k}{R_k} - UB)\Big) \forall k \in \{0, 1, 2, \ldots, N-1\} \end{aligned} \tag{3.6}$$

Here C_l, C_u and γ are positive market specific constants with $C_l \gg C_u$.

The Bellman's criteria for optimality (2.6) can now be applied for the *instantaneous* cost $f(x, u, R)$ given by eq. 3.6 resulting in another useful deduction.

Deduction 2. *Let X be the set of all possible stock prices and the instantaneous cost $f(x_k, u_k, R_k)$ be taken of the form given by (3.6). Then the adaptive cost objective $(J_k(x_k, R_k))$ given by (2.6) is linearly dependent on x_k $(\forall x_k \in X)$*

Proof. Let $P(n)$ be the proposition that the cost $J_k(x_k, R_k)$ is linearly dependent on x_k $\forall k \geq n$

Base Case: The objective function at penultimate time point (t_{N-1}) is given by

$$\begin{aligned} J_{N-1}(x_{N-1}, R_{N-1}) = x_{N-1}u_{N-1} - x_{N-1}C_l\Big(\frac{t_{N-1}}{t_N}\Big)^{\gamma} \log \Big(1 \\ - \max(0, LB - \frac{u_{N-1}}{R_{N-1}})\Big) - x_{N-1}C_u\Big(\frac{t_N}{t_{N-1}}\Big)^{\gamma} \log \Big(1 \\ - \max(0, \frac{u_{N-1}}{R_{N-1}} - UB)\Big) + x_{N-1}(1 + \beta u_{N-1})(R_{N-1} - u_{N-1}) \end{aligned}$$

which is evidently linearly dependent on x_{N-1}. Thus $P(N-1)$ holds true.

Inductive Step: Let $P(k+1)$ holds true for some $k \leq N-1$. Then

$$\begin{aligned} J_k(x_k, R_k) = x_k u_k - x_k C_l\Big(\frac{t_k}{t_N}\Big)^{\gamma} \log \Big(1 - \max(0, LB - \frac{u_k}{R_k})\Big) - x_k C_u\Big(\frac{t_N}{t_k}\Big)^{\gamma} \log \Big(1 \\ - \max(0, \frac{u_k}{R_k} - UB)\Big) + \mathbb{E}[J_{k+1}(x_k(1 + u_k + \epsilon_k), R_k - u_k)] \end{aligned}$$

From the induction hypothesis, $J_{k+1}(x_k(1 + u_k + \epsilon_k), R_k - u_k)$ is linearly dependent on $x_k(1 + u_k + \epsilon_k)$ thus $J_k(x_k, R_k)$ is linearly dependent on x_k. Hence $P(n)$ holds $\forall 0 \leq n \leq N-1$ □

Corollary 3.1. *Let X be the set of all possible stock prices and the instantaneous cost be taken of the form given by (3.6). Then the optimal action u_k for the objective (2.6) is independent of x_k $\forall k \in \{0, 1, \ldots, N-1\}$*

This computationally useful corollary follows trivially from the previous deduction.

3.2.1 Numerical Algorithm for Policy Evaluation

The Deduction 2 (and thus Corollary 3.1) is extremely advantageous to develop an efficient algorithm for determining the policy as the optimal action resulting from the theory of stochastic control is independent of stock price x. Hence all future computations can be performed by assuming stock price to be *unity*.

The time and space complexity of this algorithm is $\mathcal{O}(K^2N)$ and $\mathcal{O}(KN)$ respectively compared to the previous estimate of $\Omega(x_{max}K^2N)$ and $\Omega(x_{max}KN)$.

The C code for the algorithm 1 can be downloaded from this link.

Algorithm 1 An efficient algorithm to compute optimal policy for constrained cost

procedure OPTIMAL˙ALLOCATION(J, U) ▷ 2-D arrays to store optimal cost and action
  while $r \in \{0, 1, \ldots, InitialSize\}$ **do**
    $J[N][r] \leftarrow r$ ▷ Optimal Cost at time t_N for $x = 1$
    $U[N][r] \leftarrow r$ ▷ Optimal action at t_N (ind. of x from last corollary)
  while $i \in \{N-1, N-2, \ldots, 0\}$ **do** ▷ To evaluate u_{opt} and $J_i(1, r)$ at each time point t_i
    $J[i][0] \leftarrow 0$ ▷ Optimal Cost when stock position is null
    $U[i][0] \leftarrow 0$ ▷ Optimal Action when stock position is null
    while $r \in \{1, 2, \ldots, InitialSize\}$ **do** ▷ To determine the optimal action for each possible stock position
      $u_{opt} \leftarrow 0$
      $val_{opt} \leftarrow f(1, u_{opt}, r) + (1 + \beta u_{opt}) J[i+1][R - u_{opt}]$
      while $u \in \{1, 2, \ldots, r\}$ **do** ▷ Brute force search to determine optimal action dynamically
        $val_u \leftarrow f(1, u, r) + (1 + \beta u) J[i+1][R - u]$
        if $val_u \leq val_{opt}$ **then**
          $u_{opt} \leftarrow u$
          $val_{opt} \leftarrow val_u$
      $J[i][r] \leftarrow val_{opt}$
      $U[i][r] \leftarrow u_{opt}$

3.2.2 Resultant policy for constrained objective

The optimal allocation vector (in row-major form) with initial stock position of 1000 shares for $\beta = 5 \times 10^{-5}$, $C_l = 1000$, $C_u = 10$, $\gamma = 2$, $LB = 0.2$, $UB = 0.6$ and different no. of time points is depicted as under

- $N = 10$

$$\vec{u}^{opt} = \begin{bmatrix} 600 & 240 & 96 & 38 & 15 & 6 & 1 & 2 & 1 & 1 \end{bmatrix}$$

- $N = 30$

$$\vec{u}^{opt} = \begin{bmatrix} 0 & 600 & 110 & 58 & 47 & 37 & 30 & 24 & 19 & 15 \\ 12 & 10 & 8 & 6 & 5 & 4 & 3 & 3 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- $N = 50$

$$\vec{u}^{opt} = \begin{bmatrix} 0 & 0 & 0 & 600 & 39 & 72 & 58 & 46 & 37 & 30 \\ 24 & 19 & 15 & 12 & 10 & 8 & 6 & 5 & 4 & 3 \\ 3 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- $N = 100$

$$\vec{u}^{opt} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 119 \\ 176 & 141 & 113 & 90 & 72 & 58 & 46 & 37 & 30 & 24 \\ 19 & 15 & 12 & 10 & 8 & 6 & 5 & 4 & 3 & 3 \\ 2 & 2 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

With a steady increment in the no. of available time points N for transaction in the fixed interval T, the resultant allocation policy follows a strategy of smaller stock acquisition towards the beginning and end of the interval T whereas bigger transactions are made towards the middle. Intuitively, this kind of allocation behaviour can be explained the observing the effect of the drift βu which has a tendency to increase the stock price resulting in a larger investment cost. Therefore, it is advantageous to make small transactions towards the beginning in such a way that the stock prices have a little tendency to drift upwards and at the same time a noticeable fraction of the initial stock position is also fulfilled followed by a major acquisition towards the middle. The resultant hefty drift would eventually have a little effect on the total investment cost as the remaining stocks constitute a small fraction of the initial stock position K.

Chapter 4

Optimal Investment Strategy for Market Model-II

In this chapter, we've considered a market model (MM-II) of the second type where an investor makes a decision to trade u_k no. of shares using the information contained till time t_{k-1} and the order is executed at time t_k at the market price x_k. Accordingly, the mathematical formulation of this control problem for the uniform time partition $\Pi(T) = \{ t_1, t_2, \ldots, t_N \}$ modifies to

$$\min_{\{\pi\}} \mathbb{E}_0[\sum_{r=1}^{N} u_r x_r] \tag{4.1}$$

Subject to the conditions:

$$\begin{aligned} u_k &= \mu_k(x_{k-1}, R_k) \, \forall k \in \{1, 2, \ldots, N\} \\ R_k &= R_{k-1} - u_{k-1} \\ x_k &= g(x_{k-1}, u_k, \epsilon_k) \forall k \\ \sum_{r=1}^{N} u_r &= K \\ u_k &\geq 0 \ \forall k \end{aligned} \tag{4.2}$$

The slightly modified stock updation function $g(x_{k-1}, u_k, \epsilon_k)$ follows Geometric Brownian Motion of the form:

$$x_k = x_{k-1}(1 + \beta u_k + \epsilon_k) \tag{4.3}$$

4.1 Allocation Policy for Fiscal Cost Function (MM-II)

Using a similar analysis as discussed previously for market model of type-I, the Bellman's criteria for optimality to determine an optimal policy for problem 4.1 redefines the set of objectives to

$$V_k(x_{k-1}, R_k) = \min_{\{u_k\}} \mathbb{E}[x_k u_k + V_{k+1}(x_k, R_k - u_k)] \tag{4.4}$$

The proof of the following deduction is analogous to the one described in (2).

Deduction 3. *Let X be the set of all possible stock prices then the adaptive cost objective ($V_k(x_{k-1}, R_k)$) given by (4.4) is linearly dependent on x_k ($\forall x_k \in X$). Thus the optimal action u_k for the objective (4.1) is independent of x_k $\forall k \in \{0, 1, \ldots, N-1\}$*

The above deduction can be used to simplify the objective (4.4) in the following way

$$\begin{aligned} V_k(x_{k-1}, R_k) = \min_{\{u_k\}} &[x_{k-1}(1 + \beta u_k)u_k + x_{k-1}(1 + \beta u_k)V_{k+1}(1, R_k - u_k)] \forall k \\ &\in \{1, 2, \ldots, N-1\} \quad (\because \mathbb{E}[\epsilon_k] = 0) \end{aligned} \tag{4.5}$$

$$V_N(x_{N-1}, R_N) = x_{N-1}(1 + \beta R_N)R_N \tag{4.6}$$

4.1.1 Numerical Algorithm for Policy Evaluation

The simplified set of equations (4.5) and Deduction 3 leads to the following algorithm to compute the optimal policy function.

The time and space complexity of this algorithm is $\mathcal{O}(K^2N)$ and $\mathcal{O}(KN)$ respectively.

The C code for the Algorithm 2 can be downloaded from this link.

Algorithm 2 An efficient algorithm to compute optimal policy for fiscal cost (MM-II)

procedure OPTIMAL_ALLOCATION(V, U) ▷ 2-D arrays to store optimal cost and action
  while $r \in \{0, 1, \ldots, InitialSize\}$ **do**
    $V[N][r] \leftarrow (1 + \beta r)r$ ▷ Optimal Cost at time t_N for $x = 1$
    $U[N][r] \leftarrow r$ ▷ Optimal action at t_N (ind. of x from last corollary)
  while $i \in \{N-1, N-2, \ldots, 1\}$ **do** ▷ To evaluate the optimal action and cost $V_i(1, r)$ at each time point t_i
    while $r \in \{0, 1, \ldots, InitialSize\}$ **do** ▷ To determine the optimal action for each possible stock position
      $u_{opt} \leftarrow 0$
      $val_{opt} \leftarrow (1 + \beta u_{opt})u_{opt} + (1 + \beta u_{opt})V[i+1][R - u_{opt}]$
      while $u \in \{1, 2, \ldots, r\}$ **do** ▷ Brute force search to determine optimal action dynamically
        $val_u \leftarrow (1 + \beta u)u + (1 + \beta u)V[i+1][R - u]$
        if $val_u \leq val_{opt}$ **then**
          $u_{opt} \leftarrow u$
          $val_{opt} \leftarrow val_u$
      $V[i][r] \leftarrow val_{opt}$
      $U[i][r] \leftarrow u_{opt}$

4.1.2 Resultant policy for fiscal cost objective (MM-II)

The resultant allocation vector (optimal action) in row-major form for $\beta = 5 \times 10^{-5}$ is demonstrated as below:

- $N = 10$

$$\vec{u}^{opt} = \begin{bmatrix} 103 & 103 & 102 & 101 & 101 & 99 & 99 & 98 & 97 & 97 \end{bmatrix}$$

- $N = 30$

$$\vec{u}^{opt} = \begin{bmatrix} 35 & 34 & 34 & 34 & 34 & 34 & 34 & 34 & 34 & 34 \\ 34 & 33 & 34 & 33 & 33 & 34 & 33 & 33 & 33 & 33 \\ 33 & 33 & 33 & 33 & 33 & 33 & 32 & 32 & 32 & 32 \end{bmatrix}$$

- $N = 50$

$$\vec{u}^{opt} = \begin{bmatrix} 21 & 21 & 21 & 21 & 21 & 20 & 21 & 21 & 20 & 21 \\ 21 & 21 & 20 & 20 & 20 & 20 & 20 & 21 & 20 & 20 \\ 20 & 20 & 19 & 20 & 20 & 20 & 20 & 19 & 19 & 20 \\ 20 & 20 & 20 & 20 & 20 & 20 & 20 & 20 & 19 & 20 \\ 20 & 19 & 19 & 19 & 19 & 20 & 20 & 19 & 19 & 19 \end{bmatrix}$$

- $N = 100$

$$\vec{u}^{opt} = \begin{bmatrix} 11 & 10 & 10 & 10 & 10 & 9 & 10 & 10 & 11 & 10 \\ 11 & 10 & 10 & 11 & 9 & 10 & 9 & 10 & 10 & 10 \\ 11 & 10 & 10 & 10 & 11 & 9 & 10 & 10 & 10 & 11 \\ 10 & 11 & 10 & 10 & 10 & 10 & 10 & 11 & 11 & 10 \\ 10 & 10 & 9 & 10 & 9 & 10 & 9 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \\ 9 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 9 & 9 \end{bmatrix}$$

The comparison of the total expenditure of the above policy with Bertsimas' is tabulated as under:[1]

Stock	**Investment Cost(B)**	**Investment Cost(MM2)**	**Ratio(MM2:B)**
GOOG	\$719770.69	\$719837.14	1.00009
AAPL	\$97670.42	\$97668.46	0.99997
QCOM	\$48983.12	\$48964.32	0.99961
NVDA	\$36247.86	\$36195.93	0.99856
LXS.DE	€39972.39	€39954.41	0.99955

Table 4.1: Comparison of total expenditure between Bertsimas'(B) and Market Model-II(MM2) based on their daily opening price spanning a total of 100 working days (Feb'16 - Jun'16)

Thus from the above data we can safely infer that the optimal action resulting from the Geometric Brownian Motion of the stock prices performs equally well when compared to their corresponding Arithmetic Brownian Motion for the market model of second type.

[1]As per the stock data obtained from Yahoo Finance

Chapter 5

Conclusion

The policy resulting from the analysis performed in Chapter 3 by incorporating several risk-factors has shown considerable improvement over the Bertsimas' policy with its total expenditure tabulated as under:[1]

Stock	**Investment Cost(B)**	**Investment Cost(WR)**	**Ratio(WR:B)**
GOOG	$719770.69	$699576.13	0.97194
AAPL	$97670.42	$94117.02	0.96361
QCOM	$48983.12	$45666.70	0.93229
NVDA	$36247.86	$28670.91	0.79096
LXS.DE	€39972.39	€35319.80	0.88360

TABLE 5.1: Comparison of total expenditure between Bertsimas'(B) and Cost with Risks(WR) based on their daily opening price spanning a total of 100 working days (Feb'16 - Jun'16)

In summary, the non-performance of one-time investment policy (Table 3.1) and significant improvement of the policy resulting from the modified cost function (Table 5.1) by incorporating market risks can be safely established for the average case analysis of market model-I keeping in mind the existence of a non-zero probability of the occurrence of a case scenario where the above deduction fails to hold. The subsequent analysis conducted for market structure-II (Table 4.1) that recognizes the aspect of delayed trading for the geometric Brownian motion of stock prices yielded results with a performance similar to those proposed in [1].

[1] As per the stock data obtained from Yahoo Finance

The *instantaneous* cost objective 3.6 could be improved further by factoring constraints in a rational manner such that the penalty levied upon their violation does not undermine or overestimate the effective fiscal cost. Another way to improve the cost objective is by estimating the effect of current stock price before converging to any possible action. For instance, if the bounds on the possible stock prices and its probability distribution throughout the entire time duration T is already known, then one can possibly make use of this information by tuning the penalty functions appropriately as a significantly lower stock price and higher probability density would result in a net reduced risk for the case when one intends to invest in a large fraction even at the earlier time points. Similarly, a higher price (close to upper bound) would levy a high penalty even when one is within the bounds of the imposed constraints. These kind of formulations would bring in the dependence of the stock price resulting in improved policies but with a slight trade-off of an increased time and space complexity.

Another possible way to improve the performance of the resulting control action is by utilizing a more general form of the stock price updation function based on the theory of Linear Price Impact with Information as suggested in [1] i.e. the stock price at each successive time point can now be modeled as:

$$\begin{aligned} x_{t+1} &= f(x_t, u_t, Z_t, \epsilon_t) \\ Z_t &= g(Z_{t-1}, \eta_t) \end{aligned} \tag{5.1}$$

Finally, in future we plan to work upon the above generalizations by simultaneously improving upon the structural form of the information acquired and the algorithmic approach to reduce the average case time complexity of the final control problem.

Bibliography

[1] Dimitris Bertsimas and Andrew W Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50, 1998.

[2] Robert Almgren and Neil Chriss. Value under liquidation. *Risk*, 12(12):61–63, 1999.

[3] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[4] Eyal Even-Dar, Sham Kakade, and Yishay Mansour. Reinforcement learning in POMDPs without resets. 2005.

[5] Yuriy Nevmyvaka, Yi Feng, and Michael Kearns. Reinforcement learning for optimized trade execution. In *Proceedings of the 23rd international conference on Machine learning*, pages 673–680. ACM, 2006.

[6] Rama Cont and Arseniy Kukanov. Optimal order placement in limit order markets. 2013.

[7] Lars Ljungqvist and Thomas J Sargent. *Recursive macroeconomic theory.* MIT press, 2012.

[8] Dimitri P. Bertsekas. *Dynamic Programming and Stochastic Control.* Academic Press New York, 1976.

[9] Ailsa H Land and Alison G Doig. An automatic method of solving discrete programming problems. *Econometrica: Journal of the Econometric Society*, pages 497–520, 1960.

[10] Dimitri P Bertsekas. *Nonlinear programming.* Athena scientific Belmont, 1999.