Optimal Execution under Liquidity Constraints

Hongsik Kim

A dissertation submmited in partial fulfillment

of the requierments for the degree of

Doctor of Philosophy

Department of Mathematics

Courant Institute of Mathematical Sciences

New York University

May 2014

Marco Avellaneda

Contents

Abstract

This paper introduces a method of finding the optimal execution schedule to minimize risks under liquidity constraints. The liquidity of stocks is a significant factor for trading because a lack of liquidity leads to an inability to perform according to a preset execution schedule. The idea of liquidity constraints comes from the fact that the market impact of a small volume of trading is negligible when a large volume of trades exceeding the trading limit cannot be settled in practice. We formulate the problem as a minimization of objective functions, which measures the risk of minimum profit/loss(P/L) during an execution and the risk of cumulative market exposure. Minimum P/L during an execution is designed to measure the amount of money required to cover an execution schedule without going into bankruptcy. The optimal execution schedule is found by numerical approximation. We formulate the optimization problem as a quadratic programming problem under linear equality and inequality constraints. The objective function is positive-definite and can be solved using polynomial running time methods, such as the interior point method. Moreover, Karush-Kuhn-Tucker(KKT) conditions provides necessary and sufficient conditions for the optimal execution, proving that the optimal execution must be a piecewise-linear function with additional constraints. Finally, we empirically illustrate the construction of

an optimal execution. The sample research suggests that the optimal execution schedule can be uniquely constructed by pasting linear segments to reach hedging hyperplanes and minimize risks. Moreover, the risk of the optimal execution is proportional to $||Q||^{1.5}$ and $\frac{1}{||L||^{0.5}}$, where Q is the quantity of the initial portfolio and L represents liquidity constraints.

Introduction

The key to successful portfolio management begins with construction of a portfolio with a high expected return over risks and ends with the liquidation of a portfolio in the market without incurring additional trading costs. After Markowitz introduced mean-variance optimization in his Portfolio Selection (1952), the early stage of portfolio theory focused on developing the value function of a portfolio. Fama (1965) and Elton and Gruber (1974) suggests additional terms in the value function to measure the market in practice. Nevertheless, the essential frame of portfolio theory remains unchanged: optimizing the expected value(mean) of a portfolio over the risk(variance).

The optimal trading strategy is found using a similar frame. However, the limitation of the traditional mean-variance approach is clear. The traditional mean-variance approach yields the optimal portfolio in terms of single-period market exposure, whereas trading is a multi-period action. Analyzing the market using a single-period data is not sufficient to establish how investors optimally trade assets to develop that portfolio. Moreover, trading is not merely an evaluation using market data. Trading impacts markets and changes the price of assets for a short periods of time. Under the worst case scenario, market condition does not allow for a large volume of trading, which causes the trading strategy to fail.

An approach to address market impact is that market impact is explicitly included in the price function. Bertsimas and Lo (1998) introduces a linear price impact with serial market information. Price impact is the function of trading volume in addition to stated variables for the market. Almgren and Chriss (2000) divides market impact into two parts: a permanent market impact and a temporary market impact. Both impacts are designed as linear functions of trading volume. Hubermann and Stanzl (2004) suggests that the permanent market impact is a linear function of trading volume, and market impact decays exponentially over time.

Assuming that market impact is an explicit function provides several advantages in modelling. First, market impact is included in the price function and acts as a penalty function for large trading. When multiplied by trading volume, trading cost becomes a quadratic function of trading volume. The quadratic term prevents the optimal strategy from liquidating all assets in the first few seconds, which is not possible in practice. For a simple case, a linear price impact function is assumed to converts the optimization problem into the Hamilton-Jacobi equation, which suggests an explicit form of optimal execution.

Despite these advantages, the price impact function cannot be easily adapted in practice. Researchers have observed a bid/ask level impact. However, the market impact is difficult to establish when an asset is highly liquid, the trading volume is small or trading occurs at a low frequency level. In addition, market impact is not explicit because the coefficient is changing continuously.

In this paper, we assume that market impact exists, but that it is small and decays sufficiently fast to be negligible in price dynamics. We set up the lower and upper limit of daily trading volume in a discrete sense. At a continuous level, the same limits are applied to the speed of trading. These lower and upper boundaries work as threshold values such that the market impact in negligible.

The liquidity boundary provides inequality constraints and complicates the optimization problem. However, there are several advantages in using liquidity constraints. These boundaries can be obtained directly from the market. For example, the daily average trading volume is a popular indicator of asset liquidity. In this paper, we set the lower and upper boundaries as negative and positive 10% of daily average trading volume, respectively. Another advantage of using liquidity constraints is that it is possible to reflect the trading limit set by laws or a regulator. Sometimes regulators set trading limits for stocks in an unstable market to protect the market; for example, finance stocks are not allowed to trade for a few days after a financial crisis. To reflect those market conditions, we can simply change the boundary conditions. Liquidity constraints can be setup independently in discrete execution. Moreover, choosing different boundaries over time does not increase the complexity of numerical optimization.

This paper consists of five chapters, which are summarized below.

Chapter 1 introduces liquidity constraints. Liquidity constraints for an asset and a given period of time appear as the lower and upper boundaries. The lower boundary is a sell-side limit and a negative number, whereas the upper boundary is a buy-side limit and positive number. Without additional restrictions by regulators, these numbers are constants over time and will assume the same absolute value.

Chapter 2 introduces the profit and loss (P/L) random variable of an execution. P/L is the function of an execution for a fixed price scenario, and becomes a random variable. For a fixed execution, P/L tracks the sum of the realized P/L and unrealized market exposure for each time t. Asset prices are a stochastic process, and P/L at t becomes a random variable; the probability distribution is determined by the distribution of asset prices. We take the minimum P/L during an execution, and the value represents the money required to liquidate a portfolio without going into bankruptcy.

Chapter 3 measures the risk of the worst P/L using the 99% expected shortfall. Expected shortfall is the expectation of Value-at-Risk (VaR) over a given level of significance. An advantage of expected shortfall is that it is more sensitive to the tail event distribution and is more suitable for measuring the risk of tail events. We assume that stock prices are arithmetic Brownian motion without drift; then, we have an explicit formula for the expected shortfall of the worst P/L, which becomes the objective function.

Chapter 4 builds and solves an optimization problem. The objective function is the 99% expected shortfall of the worst P/L. Optimizing this function under liquidity constraints and initial and final portfolio position yields the optimal execution. Numerical approximation converges to continuous optimal execution, and solving Karush-Kuhn-Tucker(KKT) conditions yields the necessary conditions for the optimal execution; the optimal execution is piecewise-linear, and must trade at the liquidation limit or meet the specific hedging ratio. At a discrete level, optimization satisfies the conditions of quadratic programming with linear constraints.

Chapter 5 consists of two parts.

First, we prove that there is a relation between the initial position of portfolio and the expected shortfall of the worst P/L. For a single stock portfolio, the expected shortfall of the worst P/L is proportional to $||Q||^{1.5}$, when Q is the initial position of a portfolio. Moreover, the expected shortfall is proportional to $\frac{1}{||L||^{0.5}}$ when L represents liquidity constraints.

We demonstrate that under proper assumptions, the optimal execution of a portfolio under liquidity constraints appears as a piecewise-linear function. During the optimal execution, each stock must be liquidated as fast as possible or controlled to achieve a proper hedging ratio to reduce risks. These hedging ratios are determined by the covariance of stocks. When we follow the optimal execution, the worst P/L follows the 1.5 rule, which represents that the expected shortfall of the linear segment of the optimal execution is proportional to $||Q||^{1.5}$ and $\frac{1}{||L||^{0.5}}$, where Q is the initial position of a portfolio and L is liquidity constraints.

Second, empirical tests demonstrate that the numerical solution converges into the optimal execution and the continuous optimal execution can be constructed from the origin by pasting linear segments. The change of trading speed of the ith stock is possiple at t if and only if $\int_t^T [CX|_i(s)ds = 0$, when X is the optimal execution.

Chapter 1

Portfolio execution and liquidity constraints

Modern portfolio theory, first introduced by Harry Markowitz in 1952 uses mean-variance analysis to find the best composition of financial assets to maximize expected return relative to risk. MPT uses historical data to derive statistical measures of given portfolios and suggest the optimal portfolio that has the highest expected returns or lowest possible risk. Finding an efficient and accurate method to measure the expected return and risk of a portfolio is a major component of portfolio theory.

The task of generating or liquidating a portfolio in the market becomes another field of study. In theory, trading cost is disregarded by the assumption that a portfolio can be converted directly into an equivalent amount of cash instantly. In practice, it is impossible to trade a portfolio at the spot price. Rather, the liquidation of a portfolio appears as a continuous function of time. The price of stock changes over time and becomes a random variable. At the final time of liquidation, the realized value of a portfolio depends on how stock prices changed and the liquidation strategy of a portfolio. Modelling and measuring realized value consists of two parts: accurately modeling future stock prices and finding the optimal liquidation strategy from the price model.

In this paper, we do not discuss market prediction furthur. Instead, we focus on the interaction between the liquidation of a given portfolio and the impact of the corresponding market. In this chapter, we discuss previous market impact models and introduce liquidity constraints.

1.1 Portfolio executions and market impact functions

Assume that there is a portfolio of N stocks and that the initial position is Q_i shares for the i-th stock for $i = 1 \cdots N$. A portfolio can take either a long or short position for stocks, with positive values for long positions and negative values for short positions. This portfolio should be fully liquidated no later than the final trading day T. An execution X is a liquidation plan from 0 to T. $X(t)$ is a $1 \times N$ real vector such that $X_i(t)$ is the number of remaining shares of the ith stock at t. $X(0) = Q$ and $X(T) = 0$. Without restrictions, we can trade any amount of stock in the market. However, trading large amounts of stock in a short period of time is either impossible to achieve or incurs an extremely large trading cost.

One way to control the trading volume of X at the appropriate level is adding a penalty function as part of the trading cost. This penalty function is called the market impact function. Almgren and Chriss (2000) and Almgren, et al (2005) introduced market impact as a function of trading volume. They state that each trade causes temporary and permanent market impacts and assume that both impacts are linear functions of trading volume. Huberman and Stanzl (2004) and Gatheral (2008) suggest that permanent market impact is a linear function of trading volume and that the impact decays exponentially. The market impact function, multiplied by trading volume, adds a quadratic penalty function of the trading volume to the value of a portfolio. For a single stock portfolio, the quadratic penalty function causes the optimal execution to appear as an exponential decay function over time. Trading stocks in the market will change the price of stocks against the action for a short period of time. The market impact function can incorporate this philosophy into the model without any additional assumptions.

Despite the advantages of the market impact function, this function is not suitable for low frequency trading because the market impact decays in a sufficiently fast manner to be negligible in the market impact function in typical trading situations. The market impact function appears in high frequency trading at the bid/ask spread level and is not measurable into a closed form formula. There is not a sufficient amount of information to determine the function, even if the market impact can be measured at a specific time. The market impact function changes in such a rapid and unpredictable manner that there is no universal impact function.

1.2 Liquidity constraints

In this paper, we assume that market impact from trading does not affect the proceeding trading, i.e., that the market impact is 0. Instead, we use another technique to control the speed of liquidation, namely, liquidity constraints. Liquidity constraints for the i-th stock at time t are the lower and upper boundaries $L_l(i,t)$ and $L_u(i,t)$ such that

$$\dot{X}_i(t) \in [L_l(i,t), L_u(i,t)] \tag{1.1}$$

when $\dot{X}_i(t)$ is the speed of trading at time t for the ith stock.

There are several advantages of using liquidity constraints in trading. First, values of the lower and upper boundaries can be easily obtained by taking a certain percentile of the daily average trading volume. The daily average trading volume is a commonly used indicator of the liquidity of a stock and is updated on a daily basis. This property enables us to set up liquidity constraints that are sensitive to the current market status.

We can also use liquidity constraints to represent the trading limit from regulators. In most cases, the lower and upper limits have the same absolute value with different signs $L_l(i,t) = -L_u(i,t)$ and there is no restriction on selling or buying stocks. After the 2008 financial crisis, South Korea banned short-selling finance stocks for 2008-2013. In that situation, any financial stock must start from the long position and be liquidated in one direction, which leads to liquidity constraints $[L_l(i,t), 0]$ to prevent reverse trading. The choice of the liquidation limit does not change the degree of complexity of the optimization problem. In this research, we use 10% of daily trading volume as both the short and long side liquidation limits.

Chapter 2

The worst profit and loss and the risk of execution

In the previous chapter, we introduced executions under liquidity constraints. The main objective of an execution is to present a trading schedule from the start of liquidation until the end. Liquidity constraints must also be established to ensure that an execution is safely liquidated in the market.

The risk measure of an execution starts from calculating the value of that execution under a given price scenario. If we measure the value of a static portfolio at t, it is sufficient to know the distribution of stock prices at t. When we calculate the value of an execution schedule X at t, we need the distribution of stochastic process $P(s)$ for $o \leq s \leq t$, where $P(s)$ is a vector of stock prices at time s because X trades stocks for each time s, and the market value of X depends on the path of the stock prices(i.e., the price scenario).

In this chapter, we introduce the profit and loss(PNL) function of an execution X, which becomes a random variable with a distribution that depends on the distribution of stock prices. The worst PNL of X is defined as the minimum PNL value during the execution time. This value is the amount of money required to prevent X from going into bankruptcy during an execution. The 99% expected shortfall of the worst PNL provides an average value of extreme tail losses.

2.1 Minimum cost of liquidation

If we let $P(s)$ be the price vector of stocks at time s, the initial market value of a portfolio Q becomes $Q \cdot P(0)$. Unlike a portfolio, the calculation of the value of an execution consists of two parts: realized and unrealized stocks. The value of unrealized assets or remaining assets at t uses the stock priice vector $P(t)$ to measure the market exposure at time t. Realized stocks have their own time of trading. For a continuous execution X, stocks are traded at the speed of $\dot{X}(s)$ for $0 \leq s \leq t$, and the prices are $P(s)$. For a multi-stock portfolio, X and P are both vectors.

The cost of liquidation, or PNL of an execution X at t can be defined as the difference between the value of X at time t and the initial market value. We first assume that X is discrete; trading volume is calculated once a day and the daily trading occurs at the closing price of the day. The cost of trading at day t, which is the sum of realized and unrealized PNL, is expressed as follows:

$$\left[\sum_{s=1}^{t}(X(s-1)-X(s))\cdot(P(s)-P(0))\right]+X(t)(P(t)-P(0)) \tag{2.1}$$

$$=\left[\sum_{s=0}^{t-1}X(s)\cdot(P(s+1)-P(s))\right] \tag{2.2}$$

$$\tag{2.3}$$

The first part of the equation is the sum of realized PNL from 0 to t, and the second part is the market exposure of the unrealized portfolio. The following approximation is obtained by tracking the cost of trading continuously:

$$U(X,t)=\lim_{\Delta s\to 0}\left[\sum_{s=0}^{t-\Delta s}X(s)\cdot(P(s+\Delta s)-P(s))\right] \tag{2.4}$$

$$=\int_0^t X(s)dP(s) \tag{2.5}$$

$U(X,t)$ is a random variable whith a distribution originating from P. When $U(X,t)$ takes a negative value, the same amount of money should be prepared to cover the loss. By taking X as our execution schedule, risk should be measured by the budget required to cover the maximum loss during an execution.

$$\tilde{U}(X)=\min_{0\le t\le T}U(X,t) \tag{2.6}$$

$\tilde{U}(X)$ takes the minimum value of $U(X,t)$ for all t. $U(X)$ combined with a price scenario in probability space provides the worst loss during an execution when the price follows a certain scenario. We refer to $\tilde{U}(X)$ as the minimum cost of liquidation.

2.2 Risk measure of minimum cost of liquidation

The risk of a portfolio has been measured in various ways over the years. Markowitz introduced mean-variance portfolio theory to compare portfolio PNL on the mean-variance plane and measure the portfolio efficiency. Value-at-Risk(VaR) has been widely used since the mid-1990s despite the criticism that VaR does not represent extreme tail events. Although VaR is still widely used, expected shortfall (also referred to as conditional VaR(cVaR)) also becomes widely used after the financial crisis in 2008.

In this paper, we use a 99% level of expected shortfall to measure the negative side of the risk of $\tilde{U}(X)$. Thus, we require the VaR to have a level above 99%, and from the definition of an expected shortfall, we obtain

$$ES_{0.99}(\tilde{U}(X)) = E[x|x < VaR_{0.99}(\tilde{U}(X))] \tag{2.7}$$

A VaR of 99% indicates that the probability of taking a number above VaR is less than 1%.

$$VaR_{0.99}(\tilde{U}(X)) = \sup_x \Big[P(\tilde{U}(X) < x) \leq (1 - 0.99) \Big] \tag{2.8}$$

VaR and expected shortfall take negative tail values to represent a loss. $\tilde{U}$ measures the minimum PNL, which is more sensitive to the tail event distribution. Thus, we use expected shortfall as our risk measure.

Assume that the increments of stock price dP have no serial correlation and follow an arithmetic Brownian motion with zero drift. Because X is deterministic, we can derive the connection between the probability distribution of $U(X, T)$ and $\tilde{U}(X)$.

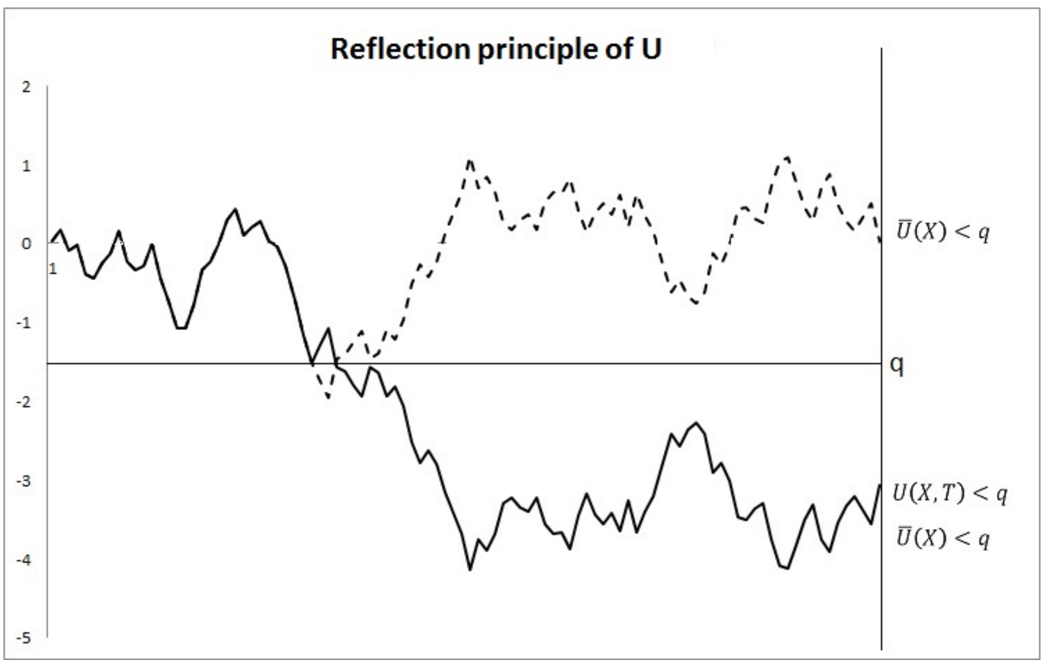

Figure 2.1: Reflection principle illustrating the relation of the probability distributions of $U(X, T)$ and $\tilde{U}(X)$.

$$P(\tilde{U}(X) < q) = 2P(U(X,T) < q) \tag{2.9}$$

The proof is the same as that for the reflection principle of Brownian motion, which states that an event from $U(X,T) < q$ can be reflected by the line $y = q$ at the first cross time to generate another event $\tilde{U}(X) < q$.

This principle provides a direct connection between the VaR values of $\tilde{U}(X)$ and $U(X,T)$. From the definition of VaR,

$$\begin{aligned} VaR_{0.99}(\tilde{U}(X)) &= \sup_x \left[P(\tilde{U}(X) < x) \leq 0.01 \right] && (2.10) \\ &= \sup_x \left[2P(U(X,T) < x) \leq 0.01 \right] && (2.11) \\ &= \sup_x \left[P(U(X,T) < -x) \leq 0.005 \right] && (2.12) \\ &= VaR_{0.995}(U(X,T)) && (2.13) \end{aligned}$$

99% VaR of $\tilde{U}(X)$ has the same value as 99.5% VaR of $U(X,T)$. Moreover, the VaR of a Gaussian process is always $\eta_\alpha \sigma$, where η_α is a constant determined from the confidence level and σ is the standard deviation of a Gaussian process. Therefore,

$$VaR_{0.99}(\tilde{U}(X)) = VaR_{0.995}(U(X,T)) = -2.58\sigma_{U(X,T)} \tag{2.14}$$

This relation demonstrate that the minimum cost of liquidation has thinner tail events than the cost of liquidation at time T. When we calculate the expected shortfall,

$$ES_{0.99}(\tilde{U}(X)) = ES_{0.995}(U(X,T)) \tag{2.15}$$

$$= \frac{1}{0.005}\int_{0.995}^{1} \eta_\theta \sigma_{U(X,T)} d\theta \tag{2.16}$$

$$= -2.94\sigma_{U(X,T)} \tag{2.17}$$

Chapter 3

Calculation of the expected shortfall of the worst PNL

In this chapter, we derive the closed form formula for the expected shortfall of the worst PNL of an execution from the assumption that the prices of stocks are arithmetic Brownian motions. Using this property, we demonstrated above that the expected shortfall of the worst PNL is a constant multiple of the volatility of the cost of trading at final time T. To address the high correlation between stocks, we decompose stock prices into the weighted sum of independent standard Brownian motion. Mathematically, this decomposition can be derived by using a covariance matrix of stock prices. Economically, stock prices are decomposed into mutually uncorrelated common risk factors from principal component analysis(PCA).

Using this decomposition and considering that there is no serial correlation for arithmetic Brownian motion, we have a closed form formula for the volatility of the total cost of trading, an explicit formula for our main objective function, and the 99% Expected shortfall of the worst PNL.

3.1 Risk factors for the stock price

In the previous chapter, we demonstrate that the calculation of the expected shortfall of the worst PNL is equivalent to calculating the variance of the PNL function at the final time T under the assumption that the price of stocks follows an arithmetic Brownian motion. Increments of stock prices are not independent due to the correlation of stocks. Rather, the increment of the i-th stock dP_i will be decomposed into the weighted sum of standard Brownian increments dW_j. These increments have an important property; namely, each dW_j becomes a common risk factor derived from PCA.

$$dP_i = \sum_{j=1}^{N} \beta_{ij} dW_j \tag{3.1}$$

PCA can be used to generate mutually uncorrelated common risk factors from historical data. It is also designed to cover the maximum amount of variance for the first few factors. Thus, each j-th factor has a unique expression given by the linear sum of the dataset.

$$dP = O \cdot [O^{\star} dP] \tag{3.2}$$

where $C = O\Lambda O^\star$, the orthonormal eigenvalue decomposition of the covariance matrix of dP. The j-th element of $O^\star dP$ is the j-th common risk factor, and its variance is λ_j, the j-th eigenvalue. When we define $d\tilde{W}_j = [O^\star dP]_j$, $d\tilde{W}_j$ becomes the linear sum of the arithmetic Brownian motion, which is also Brownian motion. The factors are mutually uncorrelated, and the variance of $d\tilde{W}_j$ is given by $[d\tilde{W}_j, d\tilde{W}_j] = \lambda_j dt$.

We call $d\tilde{W}_j = \sqrt{\lambda_j} dW_j$,

$$dP_i = \sum_{j=1}^{N} \sqrt{\lambda_j} O_{ij} dW_j \tag{3.3}$$

where dW_j are i.i.d standard Brownian motion.

3.2 Evaluation of the expected shortfall

The expected shortfall of a Gaussian process is a constant multiplied by the volatility of the process. Recall that the expected shortfall of $\tilde{U}(X)$, the worst PNL of an execution, can be derived from the expected shortfall of $U(X,T)$, which is the PNL of X at T.

$$ES_{0.99}(\tilde{U}(X)) = ES_{0.995}(U(X,T)) = \eta_{0.995} \sigma_{U(X,T)} \tag{3.4}$$

We calculate the variance of $U(X,T)$, $\sigma^2_{U(X,T)}$.

$$\sigma^2_{U(X,T)} = E\left[\left(\int_0^T X(s)dP(s)\right)^2\right] = \int_0^T E\left[(X(s)dP(s))^2\right] \tag{3.5}$$

$$\tag{3.6}$$

When we apply the risk factor decomposition $dP_i = \sum_{j=1}^N \sqrt{\lambda_j} O_{ij} dW_j$,

$$\begin{aligned}
E\left[(X(s)dP(s))^2\right] &= E\left[(\sum_{i=1}^N X_i(s)dP_i(s))^2\right] && (3.7)\\
&= E\left[\left(\sum_{i=1}^N \sum_{j=1}^N X_i(s)O_{ij}\sqrt{\lambda_j}dW_j\right)^2\right] && (3.8)\\
&= E\left[\left(\sum_{j=1}^N \left[\sum_{i=1}^N X_i(s)O_{ij}\sqrt{\lambda_j}\right] dW_j\right)^2\right] && (3.9)\\
&= \sum_{j=1}^N E\left[\left(\left[\sum_{i=1}^N X_i(s)O_{ij}\sqrt{\lambda_j}\right] dW_j\right)^2\right] && (3.10)\\
&= \sum_{j=1}^N \left[\sum_{i=1}^N X_i(s)O_{ij}\right]^2 \lambda_j dt && (3.11)\\
&= X^T C X dt && (3.12)
\end{aligned}$$

C is the covariance matrix of dP and $C = O\Lambda O^\star$, $\lambda_i = \Lambda_{ii}$. O is a unitary matrix, and $O = O^\star$. This calculation leads to

$$
\begin{aligned}
ES_{0.99}(\tilde{U}(X)) &= \eta_{0.995}\sigma_{U(X,T)} && (3.13)\\
&= \eta_{0.995}\sqrt{\left[\int_0^T XCXdt\right]} && (3.14)
\end{aligned}
$$

Chapter 4

Optimal execution under liquidity constraints

The expected shortfall of the worst PNL appears as a negative number. For two given executions X_1 and X_2, $ES(X_1) < ES(X_2) < 0$ indicates that X_1 is exposed to more risk than X_2.

In this chapter, we optimize the utility functions to find the optimal execution schedule. The optimal execution can be obtained from the numerical approximation of the discrete optimal execution, which is the solution of Karush-Kuhn-Tucker(KKT) conditions. KKT conditions imply that the optimal execution must be a piecewise-linear functions and also provides additional sufficient conditions.

4.1 Setting up the optimization problem

The expected shortfall of the worst PNL is the function of an execution schedule X, and represents the risk of following X. When we compare various execution candidates, choosing the X that maximizes the expected shortfall will minimize the risk of trading. $ES(\tilde{U}(X))$ takes a negative value, and maximizing this function gives the minimum absolute value.

Recall that when we assume that the prices of stocks follow arithmetic Brownian motion, we have a closed form formula for the expected shortfall.

$$ES_{0.99}(\tilde{U}(X)) = \eta_{0.995}\sqrt{\left[\int_0^T XCXdt\right]} \tag{4.1}$$

We also have boundary conditions and liquidity constraints for X.

$$X(0) = Q, X(T) = 0 \tag{4.2}$$

$$L^l \leq \dot{X}(t) \leq L^u \tag{4.3}$$

To find the optimal execution among all possible Xs, we must maximize $ES_{0.99}(\tilde{U}(X))$ for all Xs that satisfy the constraints. $\eta_{0.995}$ is negative, and thus, maximizing $ES_{0.99}(\tilde{U}(X))$ is equivalent to minimizing $\int_0^T XCXdt$. In summary, our problem is reduced to the following optimization problem.

$$\min_X \int_0^T XCX dt \tag{4.4}$$

$$X(0) = Q, X(T) = 0 \tag{4.5}$$

$$L^l \leq \dot{X}(t) \leq L^u \tag{4.6}$$

$\int_0^T XCXdt$ is the convex function of X with linear constraints and it has a unique solution if and only if there exists at least one feasible X.

4.2 Numerical approximation

For an execution X and a positive integer M, we define a $MNT \times 1$ vector Y_X such that

$$Y_X(MT(i-1)+j) = X_i(\frac{j}{M}) \tag{4.7}$$

$$j = 1, 2, \cdots, MT \tag{4.8}$$

Y_X is uniquely well-defined for every X and is the discretized execution of $X(0)$ with mesh size $\frac{1}{M}$. For Y_X, the expected shortfall of Y_X becomes a matrix multiplication $Y_X\hat{C}Y_X$, where each elements of $\hat{C}$ corresponds to the covariance matrix C.

Continuous optimization is changed into the following quadratic programming with linear constraints(QPLC),

$$\min_{Y_X} f(Y_X) \tag{4.9}$$

subect to

$$\begin{aligned} g_+(i,j)(Y_X) \leq& 0 && (4.10)\\ g_-(i,j)(Y_X) \leq& 0 && (4.11)\\ h_j(Y_X) =& 0 && (4.12)\\ i =& 1, 2, \cdots, MT && (4.13)\\ j =& 1, 2, \cdots, N && (4.14) \end{aligned}$$

when

$$\begin{aligned} f(Y_X) =& Y_X^T \hat{C} Y_X && (4.15)\\ g_+(i,j)(X_M) =& \frac{X_M(i,j) - X_M(i-1,j)}{M} - L(j) && (4.16)\\ g_-(i,j)(X_M) =& \frac{-X_M(j,i) + X_M(j-1,i)}{M} - L(j) && (4.17)\\ h_j(Y_X) =& Y_X(MT(N-1)+j) && (4.18) \end{aligned}$$

f is the objective function, g demonstrates liquidity constraints and h_i represents the final condition. The initial condition is removed from the fact that it also gives direct values of an execution.

$\hat{C}$ is a positive definite matrix which provides that f is a convex quadratic function. Quadratic programming with linear constraints(QPLC) combined with positive definite matrix can be solved by stable numerical methods in polynomial time.

4.3 Karush-Kuhn-Tucker conditions

Karush-Kuhn-Tucker(KKT) conditions provide necessary conditions for a local minimum point. From the fact that objective function is convex, local minimum point becomes our desired optimal execution.

Suppose Y is the optimal execution of the numerical optimization defined as previous section. KKT conditions claim that Y must satisfy the following conditions :

$$\nabla f(Y) = \sum_i \mu_i \nabla g_i(Y) + \sum_j \lambda_j \nabla h_j(Y) \tag{4.19}$$

$$g_i(Y) \leq 0 \tag{4.20}$$

$$h_j(Y) = 0 \tag{4.21}$$

$$\mu_i \geq 0 \tag{4.22}$$

$$\mu_i g_i(Y) = 0 \tag{4.23}$$

For each $Y(q)$ for $q = MT(i-1) + j$ and $j \neq M$, $Y(q) = X_i(\frac{j}{M})$ and

$$\frac{\partial f}{\partial Y(q)} = \sum_i \mu_i \frac{\partial g_i}{\partial Y(q)} + \sum_j \lambda_j \frac{\partial h_j}{\partial Y(q)} \tag{4.24}$$

Direct calculation gives that $\frac{\partial f}{\partial Y(q)} = [CX(\frac{j}{M})]_i$ and $\frac{\partial h_j}{\partial Y(q)} = 0$. Note that each element of Y appears at most 4 times in inequality constraints :

$$\bar{g}_1(Y) = \frac{Y(q) - Y(q+1)}{M} - L_u(i, \frac{j}{M}) \leq 0 \tag{4.25}$$
$$\bar{g}_2(Y) = \frac{-Y(q) + Y(q+1)}{M} + L_l(i, \frac{j}{M}) \leq 0 \tag{4.26}$$
$$\bar{g}_3(Y) = \frac{Y(q-1) - Y(q)}{M} - L_u(i, \frac{j-1}{M}) \leq 0 \tag{4.27}$$
$$\bar{g}_4(Y) = \frac{-Y(q-1) + Y(q)}{M} + L_u(i, \frac{j-1}{M}) \leq 0 \tag{4.28}$$

If none of $\bar{g}_i$ is zero, from the fact that $\mu_i g_i(Y) = 0$ we derive $\bar{\mu}_i = 0$, and it leads the necessary conditions that

$$\left[CX(\frac{j}{M})\right]_i = 0 \tag{4.29}$$

When we send $M \to \infty$, the converging continuout optimal execution must satisfy at least one of the following property for each i and t,

$$\dot{Y}_{i+} = L_l(i, t) \tag{4.30}$$
$$\dot{Y}_{i+} = L_u(i, t) \tag{4.31}$$
$$\dot{Y}_{i-} = L_l(i, t) \tag{4.32}$$
$$\dot{Y}_{i-} = L_u(i, t) \tag{4.33}$$
$$[CY(t)]_i = 0 \tag{4.34}$$

These also leads that Y is a piecewise-linear function of t.

4.4 Convergence of numerical approximation

Suppose X is an execution of Q under liquidity constraints L. We assume that $\bar{L} = |L|_\infty < \infty$. For an integer M, we define a discrete execution Y_M of X as the following

$$\tilde{X}_M(i) = X(\frac{i}{M}) \tag{4.35}$$

We construct back a continuous-time function X_M from $\tilde{X}_M$,

$$X_M(t) = \tilde{X}_M(i) \tag{4.36}$$
$$\frac{i}{M} \leq t < \frac{i+1}{M} \tag{4.37}$$

Note that X_M is a step function, not continuous function.

Conversely, for each feasible $\tilde{X}_M$, there exists at least one continuous execution X which inducex $\tilde{X}_M$ when we set $X(t) = \tilde{X}_M(i+1) + M(\frac{i+1}{M} - t)(\tilde{X}_M(i) - \tilde{X}_M(i+1))$ for $\frac{i}{M} \leq t \leq \frac{i+1}{M}$.

For $\tilde{X}_M$,

$$\tilde{X}_M \hat{C} \tilde{X}_M = \int_0^T X_M C X_M dt \tag{4.38}$$

If we define $\gamma(t) = X_M(t) - X(t)$, $X_M(t) = X(t) + \gamma(t)$ and

$$\int_0^T X_M C X_M dt = \int_0^T (X+\gamma)C(X+\gamma)dt \tag{4.39}$$
$$= \int_0^T \left[XCX + \gamma C\gamma + 2\gamma CX\right] dt \tag{4.40}$$

From liquidity constraints, $|\gamma(t)| \leq \frac{\bar{L}}{M}$. X is also bounded by $|Q|_\infty + T\bar{L}$, which provides

$$|\gamma C\gamma| \leq ||\gamma|| \cdot ||C\gamma|| \leq ||\gamma||^2 ||C||_2 \leq N \cdot \frac{\bar{L}^2}{M^2} ||C||_2 \tag{4.41}$$
$$|\gamma CX| \leq ||\gamma|| \cdot ||CX|| \leq ||\gamma|| ||X|| ||C||_2 \leq N \cdot \frac{\bar{L}(|Q|_\infty + T\bar{L})}{M} ||C||_2 \tag{4.42}$$

Therefore,

$$\left|\int_0^T X_M C X_M dt - \int_0^T XCX\right| = \left|\int_0^T \gamma C\gamma + 2\gamma CX dt\right| \tag{4.43}$$
$$\leq \int_0^T |\gamma C\gamma| dt + 2\int_0^T |\gamma CX| dt \tag{4.44}$$
$$\leq \frac{TN\bar{L}||C||_2}{M} \cdot (\frac{\bar{L}}{M} + (|Q|_\infty + T\bar{L}) \tag{4.45}$$

when T, N, $||C||_2$, $|Q|_\infty$ and $\bar{L}$ are constant over M. If X_M is the projection of X, $\int_0^T X_M C X_M dt \to \int_0^T XCX dt$.

Now suppose X is the continuous optimal execution, $\tilde{Y}_M$ is the discrete optimal execution for mesh size $\frac{1}{M}$. Since C is positive definite matrix, it is clear that

$$< X_1, X_2 >_C = \int_0^T X_1 C X_2 dt \tag{4.46}$$

gives the inner product, and it induces a Hilbert space norm. For Y_M,

$$\int_0^T Y_M C Y_M dt \leq \tilde{Y}_M \hat{C} \tilde{Y}_M + \frac{C}{M} \tag{4.47}$$

$$\leq \tilde{X}_M \hat{C} \tilde{X}_M + \frac{C}{M} \leq \int_0^T XCXdt + \frac{2C}{M} \tag{4.48}$$

$\{Y_M\}$ is a bounded sequence, and there exists a subsequence $\{Y'_M\}$ and $\{X'\}$ such that Y'_M converges weakly to X'. Moreover, the domain of minimization $\{X|X(0) = Q, X(T) = 0, |X(s) - X(t)|_i \leq L_i|s - t|\}$ is a convex closed set, so X' is also in the domain and becomes an execution.

$$\int_0^T X'CX'dt \leq \liminf_{M\to\infty} \int_0^T Y'_M C Y'_M dt \tag{4.49}$$

$$\leq \liminf_{M\to\infty} \int_0^T X'_M C X'_M dt \leq \int_0^T XCXdt + \liminf_{M\to\infty} \frac{C}{M} \tag{4.50}$$

$$= \int_0^T XCXdt \tag{4.51}$$

This proves that $X = X'$. Now, if Y_M does not converge to X, there exist an infinite subsequence of Y_M that none of its subsequence converges to X. However this set is still bounded and have the same property, leading contradiction. Therefore, Y_M converges to X, or discrete optimal executions converge to continuous optimal execution.

Chapter 5

Analysis

In this chapter, we analyze the behavior of optimal execution of a portfolio. There is a relationship between the expected shortfall of the optimal execution and the quantity of the portfolio. The risk of the worst PNL is proportional to $||Q||^{1.5}$ and also proportional to $\frac{1}{||L||^{0.5}}$, where Q is the initial position of a portfolio.

Every optimal execution can be obtained by pasting linear segments from the origin. KKT condtions suggest necessary and sufficient conditions for possible backtracking directions for each point and each execution. This demonstrates complete method of constructing the optimal execution from the origin.

5.1 Risk of the worst PNL and the volume of the initial portfolio

In this section, we find the relationship between the expected shortfall of the worst PNL and the initial quantity of a portfolio Q. For a single stock portfolio, the optimal execution can be obtained explicitly as

$$X(t) = Q - lt \tag{5.1}$$

for $0 \leq t \leq \frac{Q}{l}$. In this case, the expected shortfall of X is

$$\begin{aligned} ES_{0.99}(\bar{U}(X)) &= \eta_{0.995}\sqrt{\int_0^{\frac{Q}{l}} (Q - ls)^2 ds} && (5.2) \\ &= \eta_{0.995}\sqrt{\frac{Q^3}{3l}} = \frac{\eta_{0.995}}{\sqrt{3l}} \cdot Q^{1.5} && (5.3) \end{aligned}$$

This equation provides a direct relationship between the expected shortfall of the worst PNL and the initial quantity Q as well as liquidity constraints l. For a multi-asset portfolio, it is required to find the relationship between the optimal execution of portfolios.

Suppose that X is the optimal execution of a multi-asset portfolio Q. Now, we define an execution schedule of a portfolio αQ ($\alpha > 0$) under the same liquidity constraints as the following:

$$X_\alpha(t) = \alpha X(\frac{t}{\alpha}) \tag{5.4}$$

We claim that $X_\alpha(t)$ is the optimal execution of αQ with the final trading day αT. It is trivial that for any execution $\bar{X}_\alpha$ of αQ, there exists a unique execution $\bar{X}$ of Q defined by $\bar{X}_\alpha(t) = \alpha \bar{X}(\frac{t}{\alpha})$. Moreover,

$$\begin{align}
ES^2_{0.99}(\tilde{U}(X_\alpha)) &= \eta^2_{0.995} \int_0^{|\alpha|T} X_\alpha(t) C X_\alpha(t) dt \tag{5.5}\\
&= \eta^2_{0.995} \int_0^{|\alpha|T} \left[\alpha X(\frac{t}{|\alpha|})\right] C \left[\alpha X(\frac{t}{|\alpha|})\right] dt \tag{5.6}\\
&= \eta^2_{0.995} \alpha^2 \int_0^{|\alpha|T} X(\frac{t}{|\alpha|}) C X(\frac{t}{|\alpha|}) dt \tag{5.7}\\
&= \eta^2_{0.995} \alpha^2 \int_0^{T} X(s) C X(s) \, |\alpha| \, ds \tag{5.8}\\
&= |\alpha|^3 \cdot \left[\eta^2_{0.995} \int_0^T X(s) C X(s) ds\right] \tag{5.9}\\
&= |\alpha|^3 \, ES^2_{0.99}(\tilde{U}(X)) \tag{5.10}
\end{align}$$

This shows that the optimal execution X of Q implies the optimal execution X_α for αQ for all $\alpha > 0$ under the same liquidity constraints and extended final trading date from T to αT.

This proves that there is a direct relation between the expected shortfall of the optimal execution schedules of Q and that of αQ.

$$\frac{ES_{0.99}(\tilde{U}(X_\alpha))}{ES_{0.99}(\tilde{U}(X))} = |\alpha|^{1.5} \tag{5.11}$$

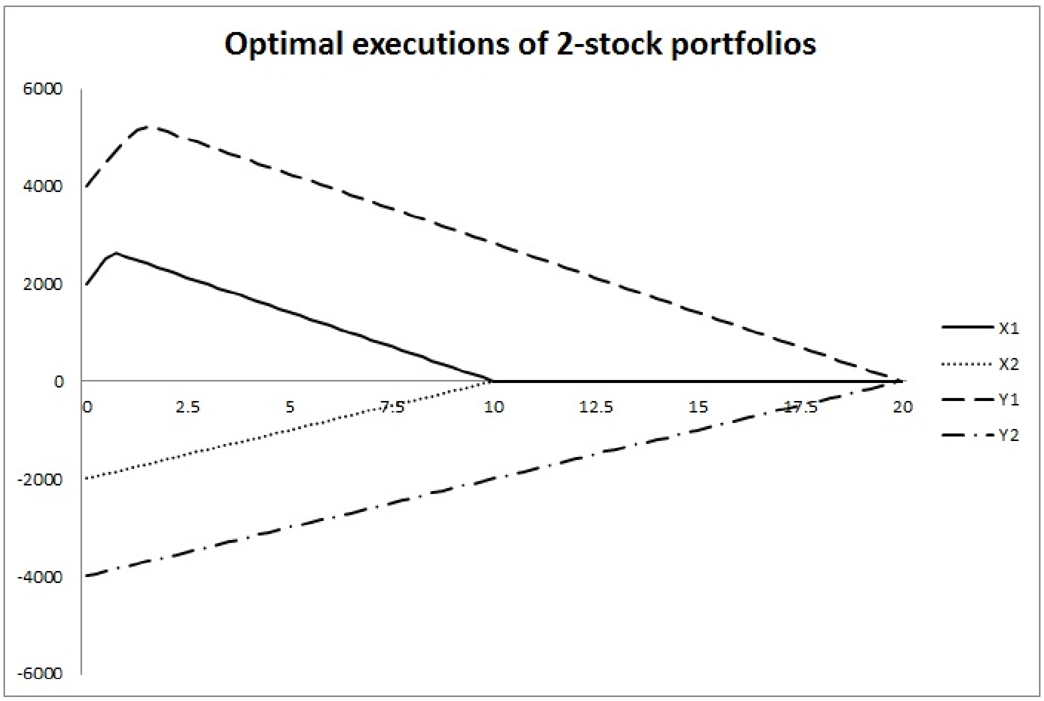

Figure 5.1: (X1,X2) represents the optimal execution of 2 stock portfolio with initial position [2000, -2000]. On the other hand, (Y1,Y2) is the optimal execution of [4000, -4000] with same liquidity constraints.

5.2 Risk of the worst PNL and liquidity constraints

Suppose there is a portfolio Q under liquidity constraints L, and call X as the optimal execution of Q under L. L contains all the information of lower and upper boundary of each stock of Q.

Suppose now that liquidity constraints are changed from L to αL for $\alpha > 0$. We claim that $Y(t) = X(\alpha t)$ is the optimal execution of Q under liquidity constraints αL. The proof is the same for the previous section, that there is a one-to-one correspondance between an execution of Q under L and under αL. The relation is given by $X_\alpha(t) = X(\alpha t)$. Moreover,

$$
\begin{aligned}
ES^2_{0.99}(\tilde{U}(Y)) &= \eta^2_{0.995} \int_0^{\frac{T}{\alpha}} Y(t)CY(t)dt && (5.12)\\
&= \eta^2_{0.995} \int_0^{\frac{T}{\alpha}} X(\alpha t)CX(\alpha t)dt && (5.13)\\
&= \eta^2_{0.995} \int_0^{T} X(s)CX(s)\frac{ds}{\alpha} && (5.14)\\
&= \frac{1}{\alpha} \cdot \left[\eta^2_{0.995} \int_0^{T} X(s)CX(s)ds\right] && (5.15)\\
&= \frac{1}{\alpha} ES^2_{0.99}(\tilde{U}(X)) && (5.16)
\end{aligned}
$$

The above equation also proves that the optimal execution of Q under αL is $X(\alpha t)$, when X is the optimal execution of Q under L. Combining results will give the formula of the expected shortfall of the optimal execution.

$$ES_{0.99}(\tilde{U}(X)) = c(Q, L) \cdot \eta_{0.995} \frac{||Q||^{1.5}}{||L||^{0.5}} \tag{5.17}$$

when $c(Q, L) = c(\alpha Q, \beta L)$ for any real number α and β.

5.3 The behavior of optimal execution in 2 stock case

In this section we assume that liquidity constraints are constant over time and $L_l(i) = -L_u(i)$. There are positive correlations between assets, which comprise more than 60% of stocks in the U.S. market. This high correlation originates from the fact that common risk factors are positively correlated to stock prices. The optimal execution of a multi-asset portfolio is designed to reduce correlation risks using a hedging hyperplane $[CX]_i$.

The following 2-asset case illustrates how the optimal execution minimizes the risk by trading at the hedging ratio. There are two sample stocks with correlation $\rho = 0.7$, volatility $\sigma_1 = 1\%$ and $\sigma_2 = 2\%$. The initial portfolio has $Q = [2000, -2000]$, liquidity constraints $L = [500, 200]$ for both long and short trading and the maximum time to execution is 10 days. All schedules are derived from numerical quadratic programming.

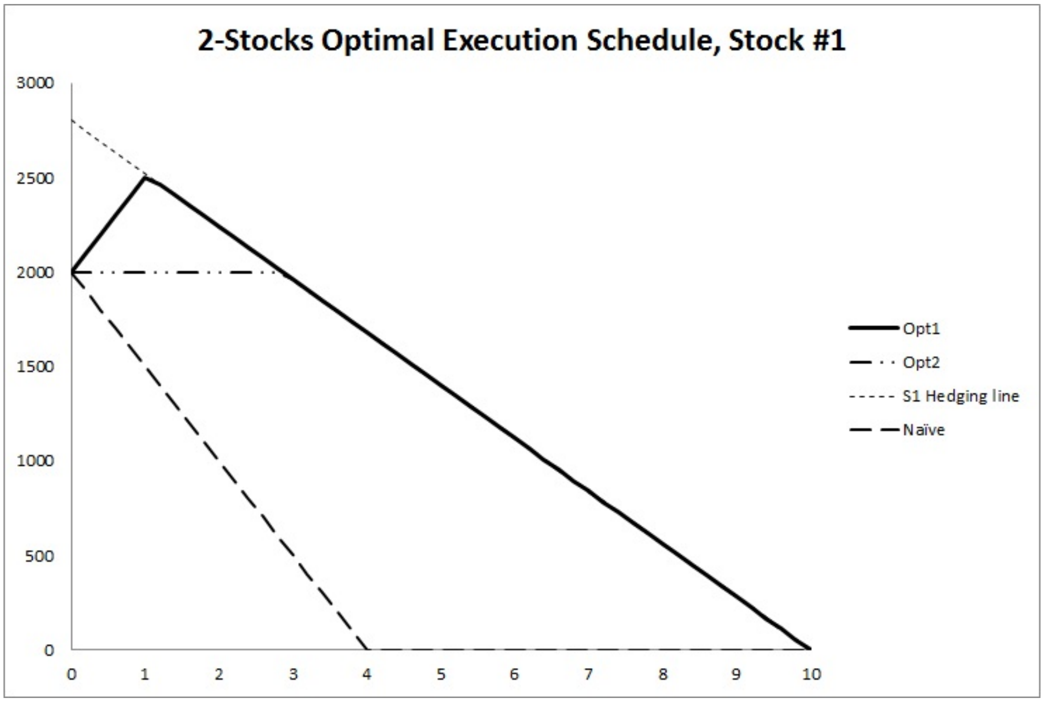

Figure 5.2: Optimal execution schedules for the sample portfolio. The naive strategy liquidates stocks at the maximum speed. Opt1 is the optimal execution when reverse trading is allowed, and Opt2 is the optimal execution when reverse trading is not allowed. Opt1 and Opt2 take the fastest path to arrive at the hedging ratio $[CX]_1 = 0$.

The second stock with short position must be liquidated as fast as possible to achieve the minimum risk. Opt1 is the optimal execution when reverse trading is allowed. On the other hand, 'Opt 2' is obtained under the assumption that reverse trading is not allowed. It is clear that Naive schedule liquidates stocks as fast as possible, but optimal schedules liquidate stocks at the trading limit for the first few days until it reaches a hedging line $[CX]_1 = 0$. When they reach the line, they follow the hedging line until the end of liquidation.

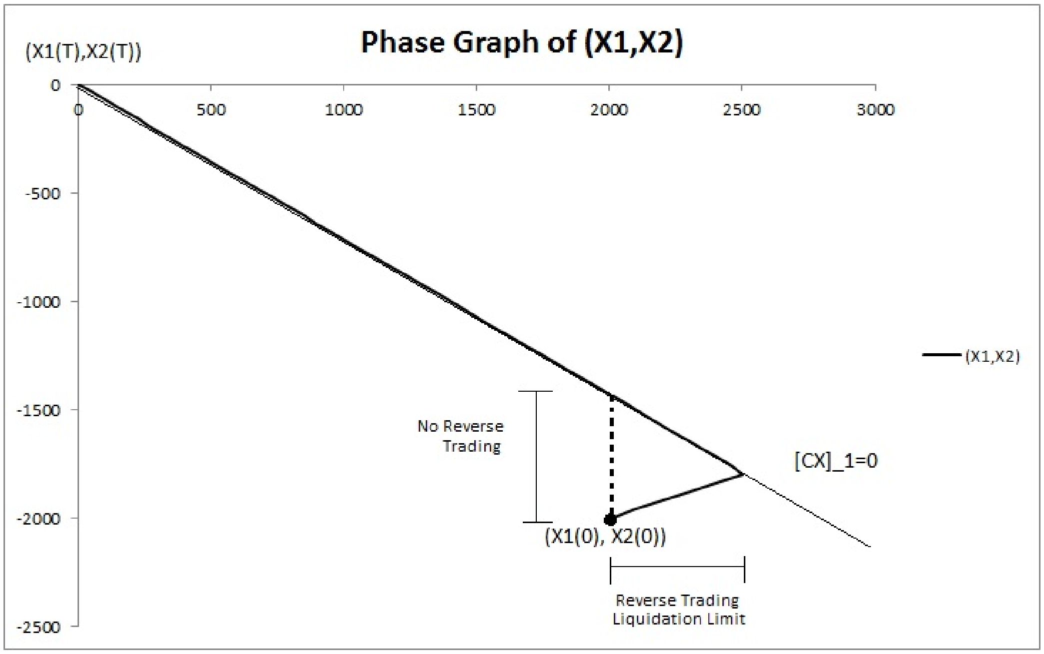

Figure 5.3: Phase graph of two optimal execution schedules. Reverse trading enables a schedule to reach the hedging line faster than without reverse trading. In a 2-dimensional plane, the optimal path is the fastest path to reach $[CX]_i$.

This becomes clear when we illustrate the movement of the optimal execution on the phase plane (X_1, X_2). When a phase at time t is not on the hedging line $[CX]_1 = 0$, optimal schedule seeks the 'fastest path' to the hedging line. The fastest path is different from the shortest path on R^2, because liquidity constraints work as speed limits. When there is a high correlation between stocks, allowing reverse trading gives us a shortcut to the hedging line.

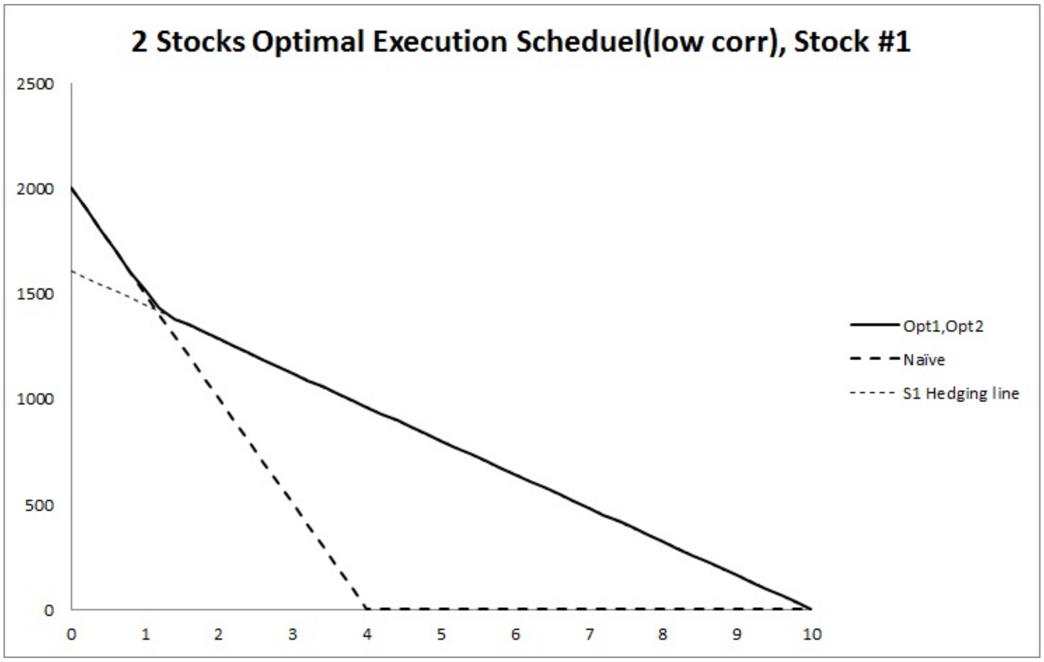

Figure 5.4: Optimal execution schedules of sample portfolio. Naive strategy liquidates stocks at the maximum speed. Opt1 is the optimal execution when reverse trading is allowed, where Opt2 is not. Opt1 and Opt2 takes the fastest way to arrive hedging ratio $[CX]_1 = 0$. When stock correlation is low($\rho = 0.4$), Hedging ratio $[CX]_1 = 0$ becomes smaller than the ratio for high correlation. Both Opt1 and Opt2 liquidates as fast as possible to reach the ratio.

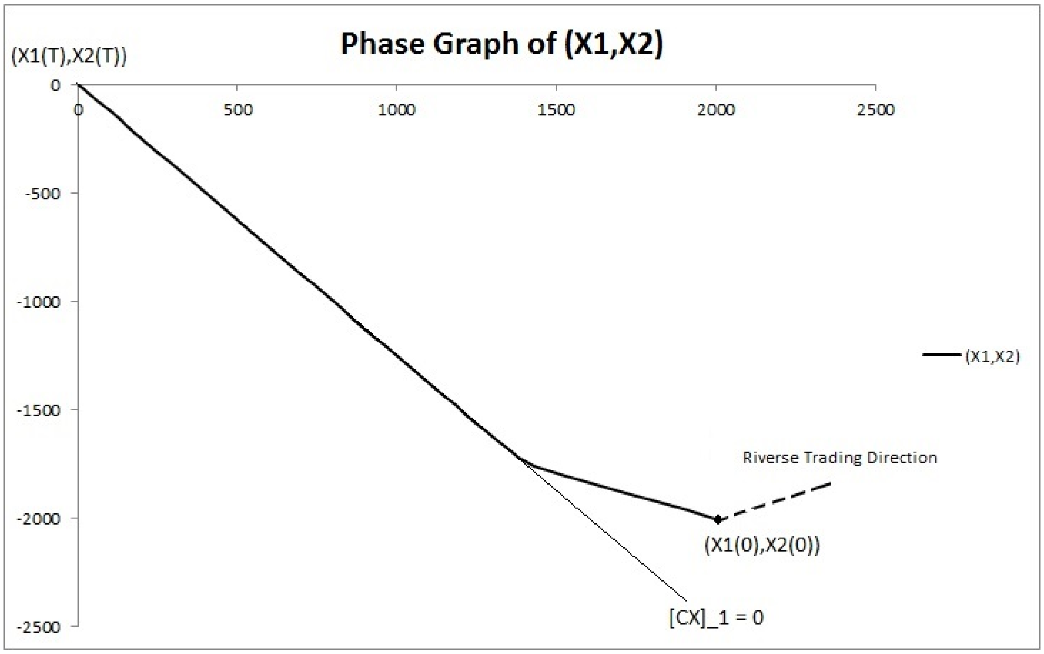

Figure 5.5: Optimal execution schedules of sample portfolio. Naive strategy liquidates stocks at the maximum speed. Opt1 allows reverse trading, Opt2 does not. Opt1 and Opt2 take the fastest way to arrive at the hedging ratio $[CX]_1 = 0$.

For a portfolio with low correlation $\rho = 0.4$, the hedging ratio also becomes lower. In this case, allowing reverse trading does not give any advantage on improving risk reductions. The reason is trivial on the phase graph; the hedging line moves left side of the initial position $X(0)$, and any reverse trading moves the portfolio far from the line.

For any Q, we show that the optimal execution of Q is piecewise-linear. Now for X, we call a vector v is a landing direction of X if and only if $\exists 0 \leq a < b \leq T$ such that $X(s) = 0\ \forall s \in [b, T]$ and $X(s) = v \cdot (b - s)$ $\forall s \in [a, b]$. Now, from KKT conditions, v must satisfy one of the following conditions for each i.

$$\dot{X}_i(s) = -v_i = L(i) \tag{5.18}$$

$$\dot{X}_i(s) = -v_i = -L(i) \tag{5.19}$$

$$[CX(s)]_i = (b-s)[Cv]_i = 0 \tag{5.20}$$

gives necessary conditions for possible landing directions.

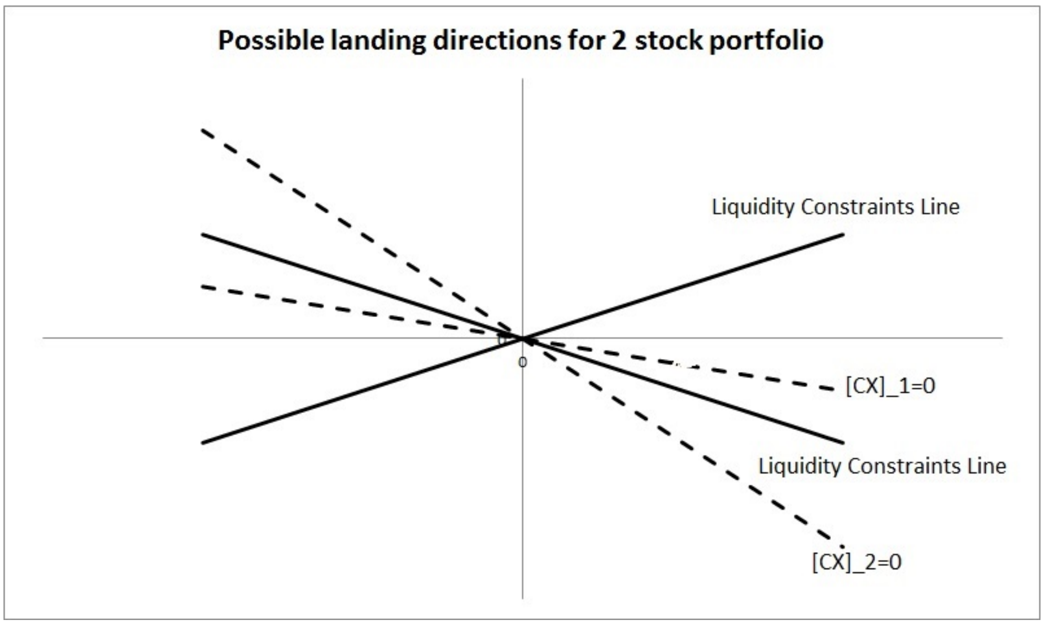

Figure 5.6: Possible landing directions for a 2-stock portfolio.

Suppose $[Cv_1]_1 = 0$. From the KKT condition, $v_1(2) = \pm L(2)$ and it leads $v_1(1) = \mp\frac{\rho\sigma_2}{\sigma_1}L(2)$. From the fact that $|v_1(1)| \leq L(1)$, $[CX]_1 = 0$ is a landing direction only if $\frac{\rho\sigma_2}{\sigma_1}L(2) \leq L(1)$, or $\rho\sigma_2 L(2) \leq \sigma_1 L(1)$ when ρ is the correlation of 2 stocks.

We claim that this condition is also sufficient condition for $[CX]_1 = 0$ to become a landing line. It is well known that KKT condition for quadratic programming with linear constraints is also a sufficient condition. Now, define an execution of a 2-stock portfolio such that $X(t) = (b - t) \cdot v_1$ for $0 \leq t \leq b$, and $X(t) = 0$ for $b \leq t \leq T$. In this case we set $b < T$ in order to make sure this solution does not depend on final time change.

To prove that X is the optimal execution of the initial position bv_1, it is sufficient to show that any numerical approximation given by $X_M(j) = X(\frac{j}{M})$ satisfies KKT condition. Suppose $v_1(2) = L(2) > 0$, then we have $X_M(i,2) - X_M(i-1,2) = -\frac{L(2)}{M}$ for $i \leq Mb$.

$$f(X_M) = X_M^T \hat{C} X_M \tag{5.21}$$

$$g_+(i,j)(X_M) = [X_M(i,j) - X_M(i-1,j)] - \frac{L(j)}{M} \tag{5.22}$$

$$g_-(i,j)(X_M) = [-X_M(j,i) + X_M(j-1,i)] - \frac{L(j)}{M} \tag{5.23}$$

$$h(j)(X_M) = X_M(MT,j) - 0 \tag{5.24}$$

for $j = 1, 2, \cdots, MT$ and $X_M(0) = X(0)$. KKT condition claims that X_M must satisfy

$$\nabla f + \sum_{i,j} [\mu_+(i,j)\nabla g_+(i,j) + \mu_-(i,j)\nabla g_-(i,j)] + \sum_j \lambda_j \nabla h(j) = 0 \tag{5.25}$$

$$\mu_+(i,j) \geq 0 \tag{5.26}$$

$$\mu_-(i,j) \geq 0 \tag{5.27}$$

When we put f, g and h to the equation, we have the following equations

$$\frac{\partial f}{\partial x(k,l)}(X_M) = [CX_M(k)]_l \tag{5.28}$$

$$\frac{\partial g_+(i,j)}{\partial x(k,l)}(X_M) = \begin{cases} 1, & i=k, j=l \\ -1, & i=k+1, j=l \\ 0, & otherwise \end{cases} \tag{5.29}$$

$$\frac{\partial g_-(i,j)}{\partial x(k,l)}(X_M) = \begin{cases} -1, & i=k, j=l \\ 1, & i=k+1, j=l \\ 0, & otherwise \end{cases} \tag{5.30}$$

$$\frac{\partial h(j)}{\partial x(k,l)}(X_M) = \begin{cases} 1, & k=MT, j=l \\ 0, & otherwise \end{cases} \tag{5.31}$$

Now, we let

$$\mu_+(i,j) = 0 \tag{5.32}$$

$$\mu_-(i,1) = 0 \tag{5.33}$$

$$\mu_-(i,2) = \begin{cases} 0, & i \geq bM \\ [Cv]_2 \left[\sum_{p=i}^{bM} p\right], & i < bM \end{cases} \tag{5.34}$$

$$\lambda_j = 0 \tag{5.35}$$

When we substitute μ and λ, it is clear that μ and λ satisfies gradient equation for X_M. Now, X_M is the optimal execution if and only if $[Cv]_2 \geq 0$ and X_M is feasible. $[Cv]_2 \geq 0$ comes from the fact C is positive difinite and $v^T Cv = v(2) \cdot [Cv]_2 \geq 0$. Therefore, X_M is the optimal execution if X_M is a feasible execution, which is equivalent to $\rho\sigma_2 L(2) \leq \sigma_1 L(1)$. When $V(2) < 0$, $[Cv]_2 < 0$ but in this case $\mu_+ = 0$ and μ_- works as positive KKT factor, which gives the same solution.

This proof can be generalized to find the final landing direction in n-dimensional case. However, the property $v(i)[Cv]_i \geq 0$ is not directly induced from positive-difinity, so we have to add another condition. When I is a set of index for a vector v_I such that

$$[Cv_I]_i = 0 \text{ for } i \in I \tag{5.36}$$

$$v_I(i) = \pm L(i) \text{ otherwise} \tag{5.37}$$

then v_I is a landing direction if and only if

$$v_I(i) \cdot [Cv_I]_i \geq 0 \tag{5.38}$$

$$X(t) = (T - t)v_I \text{ is feasible} \tag{5.39}$$

5.4 The behavior of optimal execution

When there are more than 2 stocks, the equation $[CX]_i = 0$ gives a hyperplane instead of a line equation.

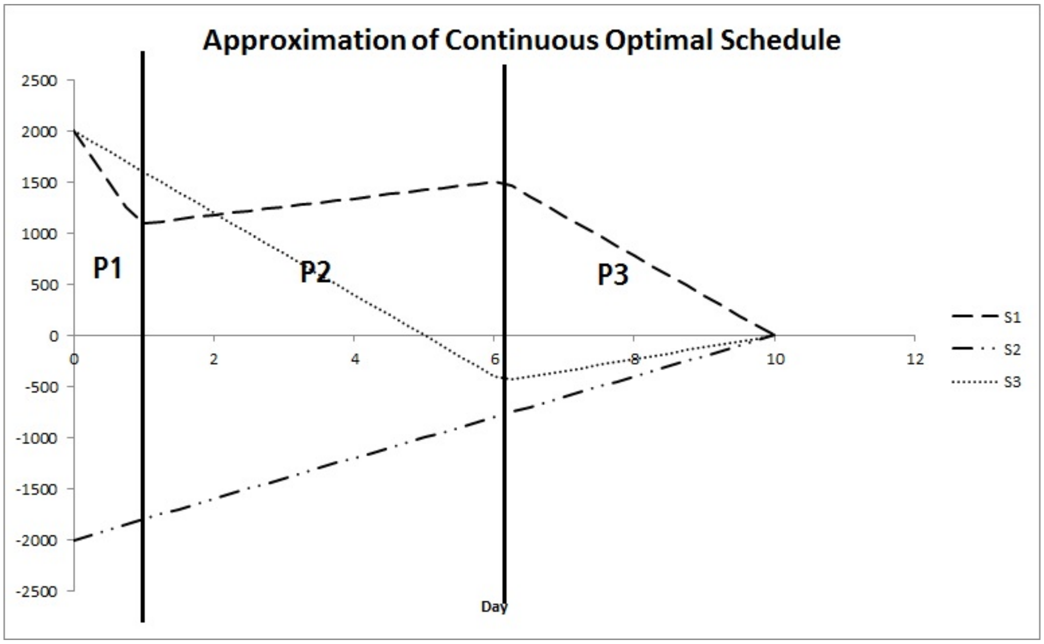

Figure 5.7: Optimal execution schedule for 3 stocks. Numerically approximated. All functions are piecewise-linear, as well as for all three phases of stocks trading at liquidation limit or $[CX]_i = 0$.

Phase	x_1''	x_2''	x_3''
P1	$(x_1')^2 - l_1^2$	$(x_2')^2 - l_2^2$	$(x_3')^2 - l_3^2$
P2	$[Cx]_1$	$(x_2')^2 - l_2^2$	$(x_3')^2 - l_3^2$
P3	$[Cx]_1$	$(x_2')^2 - l_2^2$	$[Cx]_3$

Table 5.1: The term which becomes 0 for each phase in 3-stock example. We can see that for each phase there are different combination of terms to make second derivative 0, which make the solution piecewise-linear.

This 3 stock example also shows that numerical approximations meet conditions from ODE analysis. Moreover, we can find that there are, at most, one 'state change' for each stock. All stocks start by liquidating at maximum speed to reach one or more hedging hyperplanes and stay on the plane until the end.

In this case, all stocks start liquidating up to trading limit until $(X_1(t), X_2(t), X_3(t))$ reaches to a hyperplane $[CX]_1 = 0$ under the condition $X_2(t) = 2000 - l_2 t$. After, point moves toward the intersection of two hyperplanes $[CX]_1 = 0$ and $[CX]_3 =$. The intersection of two hyperplanes in R^3 becomes a straight line from the origin. Once a schedule arrives at the line, it liquidates the rest of the assets following that line.

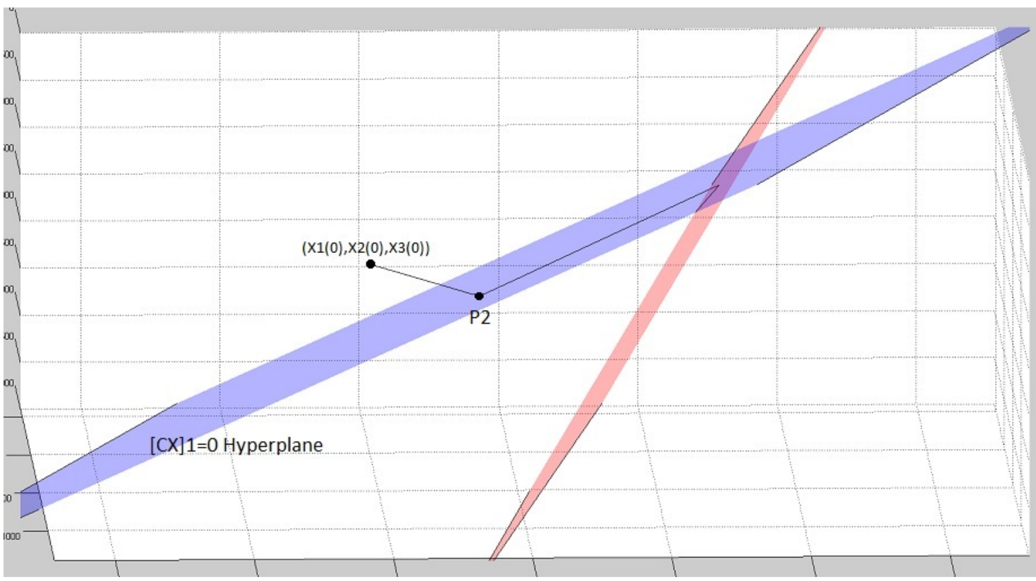

Figure 5.8: Phase graph of the optimal execution schedule for 3 stocks. It is clear that the optimal execution seeks the fastest way to the nearest hyperplane.

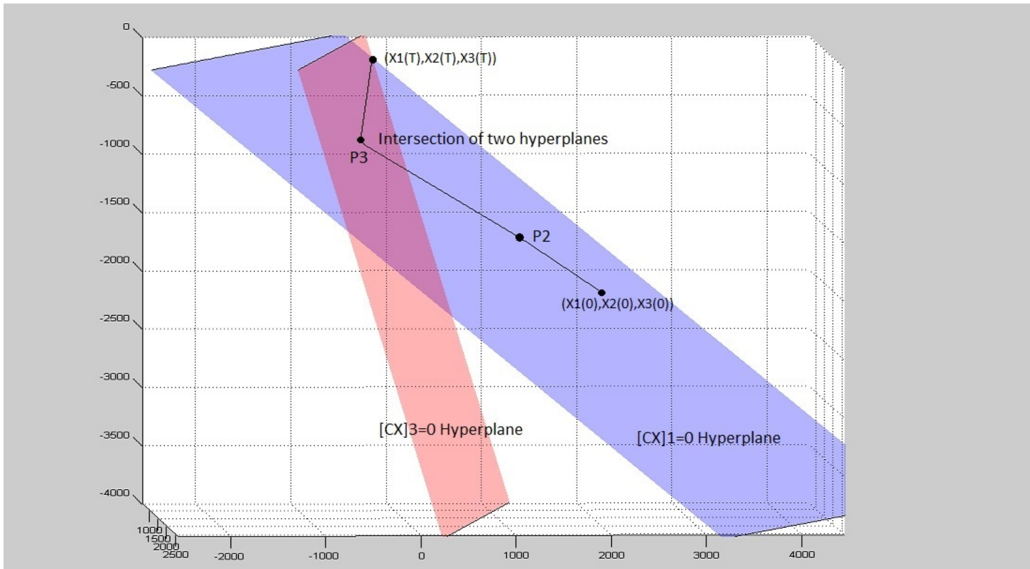

Figure 5.9: Phase graph of the optimal execution schedule for 3 stocks. It is clear that the optimal execution seeks the fasest way to the nearest hyperplane.

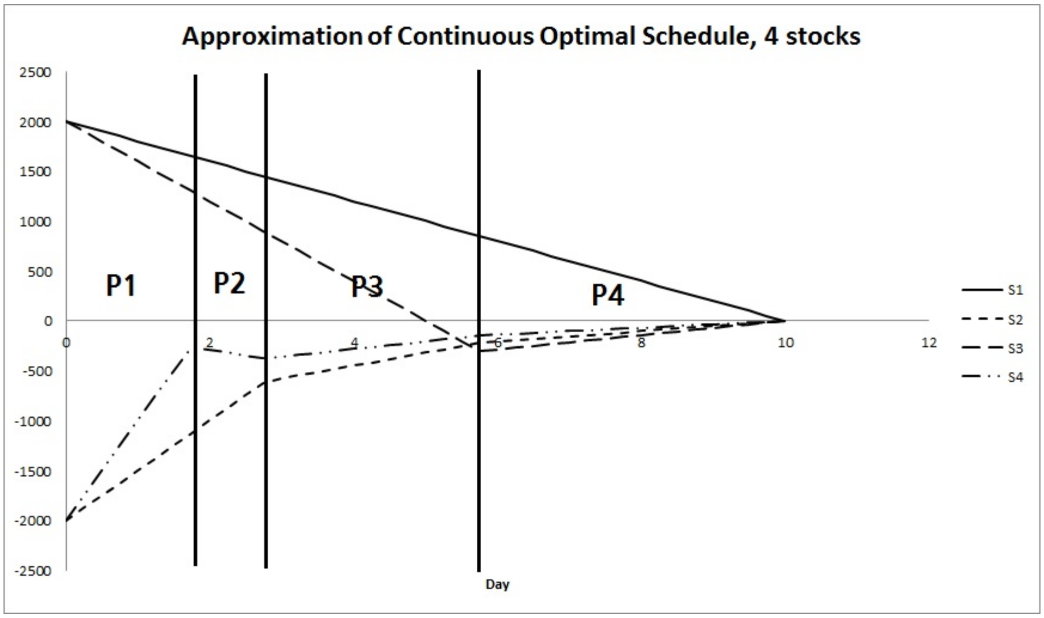

Figure 5.10: Continuous Optimal execution schedule for 4 stocks, approximated. All functions are piecewise-linear, as well as for all four phases, stocks trading at liquidation limit or $[CX]_i = 0$.

Phase	x_1''	x_2''	x_3''	x_4''
P1	$(x_1')^2 - l_1^2$	$(x_2')^2 - l_2^2$	$(x_3')^2 - l_3^2$	$(x_4')^2 - l_4^2$
P2	$(x_1')^2 - l_1^2$	$(x_2')^2 - l_2^2$	$(x_3')^2 - l_3^2$	$[Cx]_4$
P3	$(x_1')^2 - l_1^2$	$[Cx]_2$	$(x_3')^2 - l_3^2$	$[Cx]_4$
P4	$(x_1')^2 - l_1^2$	$[Cx]_2$	$[Cx]_3$	$[Cx]_4$

Table 5.2: The term which becomes 0 for each phase in 4-stock example. We can see that, for each phase, there are different combinations of terms to make second derivative 0, which make the solution piecewise-linear.

A sample test for a 4-stock portfolio suggests the general construction of the optimal execution of a given portfolio.

First, take the minimum trading day such that all assets can be liquidated without violating liquidity constraints. This forces at least one asset to be liquidated as fast as possible until the end of liquidation. Next, starting from the initial portfolio, seek the fastest way to one of the possible hyperplanes $[CX]_i = 0$. Until the execution reaches to a hyperplane, it has to liquidate assets as fast as possible. Once the execution reaches the nearest hyperplane, find the next nearest hyperplane and travel to it as fast as possible. The optimal execution must stay on the hyperplanes once reached. Repeat this process until the end of liquidation.

For example, the optimal execution of a 4-stock portfolio takes the following path:

$$X(0) \rightarrow [CX]_4 \rightarrow [CX]_4 \cap [CX]_2 \tag{5.40}$$

$$\rightarrow [CX]_4 \cap [CX]_2 \cap [CX]_3 \rightarrow 0 \tag{5.41}$$

5.5 Backward construction of the optimal execution from the origin

Phase plane enables ut to explain an execution in R^N space, which is a path from Q to the origin. In this section, we show that we can construct the optimal execution from the origin in R^N space by pasting linear segment.

We define a set of semi-landing direction LD such that

$$LD = \{v \in R^N |\, |v(j)| \leq L(j), \text{ and}(v(i) = \pm L(i) \text{ or } [Cv]_i = 0, \text{ i=1, 2, } \cdots N)\} \tag{5.42}$$

We assume that for $v \in LD$, $v(i) \neq \pm L(i)$ if $[Cv]_i = 0$. For a given initial portfolio Q and for a mesh size $\frac{1}{M}$, let X_M be the optimal execution of Q. We call $|I_{X_M}|$ is the degree of X_M when $I_{X_M} = \{i | [CX_M(t)]_i = 0\ \forall t\}$. When the degree of X_M is n, it means that X_M stays on the intersection of n hyperplanes during its execution. For N stock case, $|I_X| = N$ indicates that $X = 0$ for all t.

We also call X_M is the universal optimal execution when $X_M(s) = 0$ for some $s < MT$. It is clear that if X_M is the universal optimal execution for final time T and $\bar{X}_M$ is the universal optimal execution for $\bar{T}$ when $T < \bar{T}$, $X_M(t) = \bar{X}_M(t)$ for all $0 \leq t \leq T$ and $\bar{X}_M(t) = 0$ for $T \leq t \leq \bar{T}$.

X_M must satisfy KKT conditions for the set of constraints g_+, g_- and h with KKT multiplier set μ_+, μ_- and λ. If X_M is universal, KKT multiplier set is uniquely constructed from $X_M(MT)$ backwardly and $\lambda = 0$.

For $v \in LD$, construct Y_v as an execution of $X_M(0) - \frac{v}{M}$ such that

$$Y_v(i) = \begin{cases} X_M(0) - \frac{v}{M}, & i = 0 \\ X_M(i-1), & 1 \leq i \leq 1 + MT \end{cases} \tag{5.43}$$

Y_v is the universal optimal execution of $X_M(0) - \frac{v}{M}$ if and only if Y_v meets

KKT condition. It is clear that Y_v is universal, and KKT multiplier set $\bar{\mu}_{\pm}$ is uniquely constructed as the following :

$$\bar{\mu}_{\pm}(i,j) = \mu_{\pm}(i-1,j) \tag{5.44}$$

for $i = 2, 3, \cdots, 1 + MT$. The value of $\bar{\mu}_{\pm}(1,j)$ depends on the sign of $[CY_v(0)]_j - (\mu_+(1,j) - \mu_-(1,j))$. If $[Cv]_j \neq 0$, without loss of generality we assume that $v(j) = L(j)$. In this case, $\bar{\mu}_-(1,j) = 0$ and $\bar{\mu}_+(1,j) = [CY_v(0)]_j - (\mu_+(1,j) - \mu_-(1,j))$.

If $[Cv]_j = 0$, by the assumption $v(j) \neq \pm L(j)$ and it leads that $[CY_v(0)]_j = 0$ if Y_v is the universal optimal execution. Moreover $[CX_M(0)]_j = [CY_v(0)]_j = 0$ and it leads the following property :

Y_v is the universal optimal execution of $X_M(0) + \frac{v}{M}$ if and only if for each j, $v(j)$ takes one of the following value according to the rule :

$$\sum_{t=0}^{MT} [CX_M(t)]_j + [CY_v(0)]_j \begin{cases} > 0 & \to v(j) = L(j) \\ < 0 & \to v(j) = -L(j) \\ = 0 & \to v(j) = \pm L(j) \end{cases} \tag{5.45}$$

$$[Cv]_j = 0 \to [CX_M(0)]_j = 0, \ \sum_{t=0}^{MT} [CX_M(t)]_j = 0 \tag{5.46}$$

When $[CX_M(t)]_j = 0$ for all t, it is always free to exit the plane for the direction of v only if $v(j)[Cv]_j \geq 0$ because $\sum_{t=0}^{MT}[CX_M(t)]_j = 0$. Recall that $v(j)[Cv]_j \geq 0$ is a necessary condition for v become a landing direction.

Conversely, If we assume that $[CX_M(0)]_j = 0$ and $\sum_{t=0}^{MT}[CX_M(t)]_j = 0$ implies $[CX_M(t)]_j = 0$ for $0 \leq t \leq T$, then it is trivial that if a universal optimal execution X enters the plane $[CX]_j = 0$ and stays for a short period of time, X stays on the plane forever. Even in the case that $\sum_{t=0}^{MT}[CX_M(t)]_j = 0$ leads $[CX_M(t)]_j = 0$, it is still possible for X_M to 'penetrate' planes. For each j, the change of speed of the j-th stock is possible only when $\sum_{s=t}^{MT}[CX_M(t)]_j = 0$.

For the continuous optimal execution, X is the optimal execution only if

$$\dot{X}_i^-(t) \int_t^T [CX(s)]_i ds \leq 0 \tag{5.47}$$

for all $0 \leq t \leq T$ and $i = 1, 2, \cdots, N$, when $\dot{X}_i^-(t)$ is the first derivative of X_i obtained from t_-.

Chapter 6

Summary

In this paper, we discuss a new approach for measuring risk and obtaining optimal execution schedule. The new approach replaces the market impact function with liquidation constraints of the trading volume. Our objective is to minimize the risk of the worst PNL during an execution. Under the assumption that the prices of stocks are arithmetic Brownian motions without a drift, we can derive the explicit formula for the expected shortfall of the worst PNL of an execution X.

$$ES_{0.99}(\tilde{U}(X)) = \eta_{0.995}\sqrt{\left[\int_0^T XCX dt\right]} \tag{6.1}$$

The optimal execution of a portfolio can be obtained by minimizing $ES_{0.99}(\tilde{U}(X))$ overall all X. X must satisfy boundary conditions $X(0) = Q$, $X(T) = 0$ as well as liquidity constraints.

We find the optimal execution by numerical approximation. Optimizing $ES_{0.99}(\tilde{U}(X))$ over discrete execution becomes a quadratic programming with linear constraints(QPLC) and KKT conditions implies necessary and sufficient conditions. KKT analysis suggests that optimal execution must satisfy one of the following conditions.

$$\dot{X}_i(t) = L_i^u \tag{6.2}$$

$$\dot{X}_i(t) = L_i^l \tag{6.3}$$

$$[CX(t)]_i = 0 \tag{6.4}$$

Analysis reveals the relation between the risk and quantity of the initial portfolio Q and liquidity constraints L. The risk increases as the initial quantity increases, and the risk decreases as the liquidity increases. The explicit relation is given as

$$ES_{0.99}(\tilde{U}(X)) = c(Q, L) \cdot \eta_{0.995} \frac{||Q||^{1.5}}{||L||^{0.5}} \tag{6.5}$$

when $c(Q, L) = c(\alpha Q, \beta L)$.

KKT conditions gives sufficient conditions for X to be the optimal execution, and it enables us to construct optimal executions from the origin. Optimal executions must be made by pasting linear segment, which directions must be confirmed by KKT conditions. If we assume that $[CX(t)]_i = 0$

and $\int_t^T [CX(s)]_i ds = 0$ implies $[CX(s)]_i = 0$ for $t \leq s \leq T$, the optimal execution stays on the hedging plane once it arrives unless it penetrates it due to liquidity unbalance.

In general, continuous optimal execution must satisfy the following equation for all t and i

$$\dot{X}_i^-(t) \int_t^T [CX(s)]_i ds \leq 0 \tag{6.6}$$

Bibliography

[1] Acerbi, Carlo, and Giacomo Scandolo. *"Liquidity Risk Theory and Coherent Measures of Risk."* Quantitative Finance 8.7 (2008):682-92. Print.

[2] Almgren, Robert, and Neil Chriss. *"Optimal Execution of Portfolio Transactions."* N.p., 8 Apr. 1999. Web. `http://www.courant.nyu.edu/~almgren/papers/optliq.pdf`

[3] Almgren, Robert, Chee Thum et al. *"Direct Estimation of Equity Market Impacts."*. N.p., 10 May. 2005. Web. `http://www.courant.nyu.edu/~almgren/papers/costestim.pdf`

[4] Avellaneda, Marco, and Jeong-Hyun Lee. *"Statistical Arbitrage in the U.S. Equities Market."*. N.p., 11 July. 2008. Web. `http://www.math.nyu.edu/faculty/avellane/AvellanedaLeeStatArb071108.pdf`

[5] Avellaneda, Marco, and Rama Cont. *"Close-Out Risk Evaluation(CORE): A new risk-management Approach for Central Counterparties."*. N.p., Web. `http://www.math.nyu.edu/faculty/avellane/LiquidityMarketRisk.pdf`

[6] Bertsimas, Dimitris, and Andrew W. Lo. *"Optimal control of execution costs."*. Journal of Financial Markets 1.1 (1998):1-50. Print.

[7] Chen, Nai-Fu, Richard Roll, and Stephan A. Ross. *"Economic Forces and the Stock Market."* The Journal of Business 59.3 (1986):383. Print.

[8] Duffie, Darrell, and Alexandre Ziegler. *"Liquidation Risk"* Financial Analysts Journal 59.3 (2003): 42-51. Print.

[9] Elton, Edwin J., and Martin J. Gruber. *"Modern Portfolio Theory, 1950 to date."* Journal of Banking and Finance 21.11-12 (1997): 1743-759. Print.

[10] Elton, Edwin J., and Martin J. Gruber, *"Finance as a dynamic process."* Englewood Cliff., NJ: Prentice-Hall, 1975. Print.

[11] Fama, Eugene F. *"The Behaviour of Stock-Market Prices."* The Journal of Business 38.1 (1965): 34-105. Print.

[12] Fama, Eugene F., and Kenneth R. French. *"Common risk factors in the returns on stocks and bonds."* Journal of Financial Economics 33.1 (1993): 3-56. Print.

[13] Gatheral, Jim. *"No-dynamic-arbitrage and market impact."*. Quantitative Finance 10.7 (2010): 74959. Print.

[14] Huberman, Gur, and Wener Stanzl. *"Price manipulation and Quasi-Arbitrage"*. Econometrica 72.4 (2004): 1247275. Print.

[15] Jorion, Philippe. *"Value at Risk, The new benchmark of evaluating financial risk, 2nd edition."* New York: McGraw-Hill, 2001. Print.

[16] Kozlov, M.k., S.p. Tarasov, and L.g. Khachiyan. *"The Polynomial Solvability of Convex Quadratic Programming."* USSR Computational Mathematics and Mathemarical Physics 20.5 (1980): 223-28. Print.

[17] Lillo, Fabrizio, J. Doyne Farmer, and Rosario N. Mantegna. *"Econophysics: Master curve for price-impact function."* Nature 421.6919 (2003): 129-30. Print.

[18] Ross, Stephen A. *"The Arbitrage Theory of Capital Asset Pricing."* Journal of Economic Theory 13.3 (1976): 341-60. Print.

[19] Roy, A. D., and Harry M. Markowitz. *"Portfolio Selection."* Econometrica 29.1 (1961): 99. Print.