UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2009:18

Pairs Trading in Optimal Stopping Theory

Li Qiang

Examensarbete i matematik, 30 hp
Handledare och examinator: Erik Ekström
September 2009

Department of Mathematics
Uppsala University

Contents

List of Figures

Abstract

Pairs trading is an investment strategy used by many hedge funds. Consider two similar stocks which trade at some spread. If the spread widens, then short the high stock and buy the low stock. As the spread narrows again to some equilibrium value, a profit results. This thesis firstly provides an overview of such investment strategies. Secondly, a model of how to value a pairs trade is built by optimal stopping theory and the model is studied analytically in infinite horizon. In the third chapter, we solve the corresponding PDEs with finite horizon numerically by finite difference methods. All programming is done by Matlab, and the code is attached to the appendix.

Acknowledgements

First of all, I would like to express my warmest gratitude to my supervisor, Erik Ekström, for his patience and instructive suggestions on the writing of this thesis. Without his invaluable help and generous encouragement, I would not accomplish my thesis.

Besides, I wish to thank the department of mathematics of Uppsala University for giving me the opportunity to study here. I am also greatly indebted to the professors and teachers in Uppsala University, who have instructed and helped me a lot in the past two years.

The last but not the least, my thanks would go to my beloved family for their loving considerations and great confidence in me all through these years. I also owe my sincere gratitude to my girlfriend Yang Zeng who helped me with the Matlab programming part and my fellow classmates who gave me their help and time in listening to me and helping me work out my problems during the difficult course of the thesis.

Chapter 1

Introduction

1.1 Introduction to Pairs Trading

1.1.1 History

The background of pairs trading can be found in [1], we just give a brief summary here. Pairs trading practice is first introduced to Wall Street by quant Nunzio Tartaglia, who worked for Morgan Stanley in the mid-1980's. During then, he assembled a group of mathematicians, physicists, and computer scientists to uncover arbitrage opportunities in the stock markets. The strategies developed by the group involved using sophisticated statistical techniques to design a kind of trading programs. If the filter rules were detected, the programs were executed by automated trading systems. At that time, trading programs of this kind were considered to the cutting edge of technology.

Among those techniques, Tartaglia's group used a strategy called pairs trading in their portfolios. The strategy involved identifying pairs of securities whose prices tended to move together. Whenever an anomaly in the relationship was noticed, the pair would be traded with the idea that the anomaly would correct itself. Tartaglia and his group traded these pairs with great success in 1987 – a year when the group made a $50 million profit for the firm. However, because of bad performance in the following years, the group disbanded in 1989. Members of the group then found themselves in some other trading firms, and knowledge of the idea of pairs trading gradually spread. Pairs trading has since increased in popularity and become a common investment strategy used by individual and institutional traders as well as hedge funds.

1.1.2 Nuts and Bolts

Before getting started, we should present an important definition from [2].

Definition 1.1 (Market-neutral Strategy). *A trading strategy that derives its returns from the relationship between the performance of its long position and the performance of its short position, regardless of whether the market goes up or down, in good times or bad times.*

This definition makes apparent the derivation of the idea of market neutrality: portfolio performance is driven by relative performance rather than by the absolute performance manifest in a traditional long only or short only portfolio. In a market-neutral strategy, the return on the portfolio is a function of the return differential between the securities that are held long and those that are held short. As the market appreciates, both the long and short positions appreciate in value, so the overall portfolio value remains constant. Similarly, if the market declines, both the long and short positions will decline in value. If the change in value of the long positions equals that of the short positions, the value added from equity selection will be zero. If a manager is skillful enough, who can translate into the long securities having a higher return than the short securities, investors will enjoy a consistently positive return regardless of the overall market return.

Pairs trading is a market-neutral strategy in its most primitive form. The market-neutral portfolios are constructed using just two securities, which are a long position in one security and a short position in the other, in a predetermined ratio. At any given time, the portfolio is associated with a quantity called spread. This quantity is computed using the quoted prices of the two securities and forms a time series.

The general idea for investing in the marketplace from a valuation point of view is to sell overvalued securities and buy undervalued ones. However, it is possible to determine that a security is overvalued or undervalued only if we also know the true value of the security in absolute terms. However, it is very hard to do so. Pairs trading attempts to resolve this using the idea of relative pricing as we mentioned before. If two securities have similar characteristics, then the prices of both securities must be more or less the same. Note that the specific price of the security is not of importance. It is only important that the prices of the two securities be the same. If the prices happen to be different, it could be that one of the securities is overpriced, the other security is undervalued, or the mispricing is a combination of both. Pairs trading involves selling the higher-priced security and buying the lower-priced security with the idea that the mispricing will correct itself in the future. The mutual mispricing between the two securities is captured by the spread. The greater the spread is, the higher the magnitude of mispricing and the greater the profit potential.

1.1.3 Some Examples

The followings are some potential correlated pairs in real-life world.

- Dow Jones and S&P500
- Coca Cola and Pepsi
- Dell and HP
- Ford and General Motors

1.2 Introduction to Optimal Stopping Theory and Free-boundary Problems

1.2.1 Background

Optimal stopping theory is concerned with the problem of choosing a time to take a particular action based on sequentially observed random variables, in order to maximize an expected payoff or to minimize an expected cost. Optimal stopping problems can be found in many areas, such as statistics, economics, and financial mathematics (related to the pricing of American options).

According to [3], we know that the problem arose in the sequential analysis of statistical observations with Wald's theory of the sequential probability test in Wald (1945) and the subsequent books, *Sequential Analysis* (1947) and *Statistical Decision Functions* (1950). The Bayesian perspective on these problems was treated in the basic paper of Arrow, Blackwell and Girshick (1948). The generalization of sequential analysis to problems of pure stopping without statistical structure was made by Snell (1952). In the 1960's, papers of Chow and Robbins (1961) and (1963) gave impetus to a new interest and rapid growth of the subject. The book, *Great Expectations: The Theory of Optimal Stopping* by Chow, Robbins and Siegmund (1971), summarizes this development.

There are many ways to solve optimal stopping problems. Generally, we reduce them to free-boundary problems. Free-boundary problem is a sort of PDE with initial or final condition, boundary conditions as well as smooth fit condition. In most cases, we solve such PDEs numerically.

1.2.2 Markovian Approach in Optimal Stopping Problems

Now, we are moving to the mathematical part. The aim of this part is to exhibit some basic results of optimal stopping theory in continuous time. There are two methods, the Martingale approach and the Markovian approach. We focus on the latter one since this is the approach we are going to use in this thesis. To illustrate the approach, we follow the presentation of [4]. For more information about the Martingale approach, please refer to [4] as well.

Consider a strong Markov process $X = (X_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P}_x)$ with values in a measurable space $(E, \mathcal{B})$ where $E = \mathbb{R}^d$ for some $d \geq 1$ and $\mathcal{B} = \mathcal{B}(\mathbb{R}^d)$ is the Borel σ-algebra on

$\mathbb{R}^d$. It is assumed that the process X starts at x under P_x for $x \in E$ and is both right and left continuous over stopping times. It is also assumed that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous. In addition, it is assumed that the mapping $x \mapsto \mathsf{P}_x(F)$ is measurable for each $F \in \mathcal{F}$. Then given measurable function $G : E \mapsto \mathbb{R}$ satisfying the following condition (with $G(X_T) = 0$ if $T = \infty$):

$$\mathsf{E}_x\Big(\sup_{0\leq t\leq T} |G(X_t)|\Big) < \infty \tag{1.1}$$

for all $x \in E$, and the optimal stopping problem is given by

$$V(x) = \sup_{0\leq \tau\leq T} \mathsf{E}_x G(X_\tau) \tag{1.2}$$

where τ is a stopping time with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Note also that in (1.2) T can be ∞ as well.

Generally, V is called the *value function*, and G is called the *gain function.* To solve the optimal stopping problem means two things. Firstly, we need to find out the optimal stopping time τ^*, and then we need to compute the value $V(x)$ for $x \in E$ as explicitly as possible. Note that at each time t, we can optimally choose the process X either continue or stop. In this way the state space E can be divided into two areas: the *continuation set* C and the *stopping set* $D = E\backslash C$. It follows that as soon as the observed value $X_t(\omega)$ enters D, the observation should be stopped and an optimal stopping time is obtained.

In the sequel we will treat the *finite horizon* formulation ($T < \infty$) and the *infinite horizon* formulation ($T = \infty$) of the optimal stopping problem at the same time. In the former case ($T < \infty$) we need to replace X_t by $Z_t = (t, X_t)$ for $t \geq 0$ so that the problem becomes

$$V(t,x) = \sup_{0\leq \tau\leq T-t} \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}) \tag{1.3}$$

where the "rest of time" $T-t$ changes when the initial state $(t,x) \in [0,T]\times E$ changes in its first argument. It may be noted in (1.3) that at time T we have the "terminal" condition $V(T,x) = G(T,x)$ for all $x \in E$ so that the first entry time of Z to the stopping set D, denoted below by τ_D, will always be smaller than or equal to T and thus finite. In the latter case ($T = \infty$), problem (1.2) can be written as follows:

$$V(x) = \sup_{\tau} \mathsf{E}_x G(X_\tau) \tag{1.4}$$

the continuation set and stopping set are given by

$$C = \{x \in E : V(x) > G(x)\}$$
$$D = \{x \in E : V(x) = G(x)\}$$

respectively.

Definition 1.2. *A measurable function $F : E \mapsto \mathbb{R}$ is said to be superharmonic if*

$$\mathsf{E}_x F(X_\sigma) \leq F(x)$$

for all stopping times σ and all $x \in E$.

From [4], We can prove that V is the smallest superharmonic function which dominates G on E.

Remark 1.3. *Solving the optimal stopping problem* (1.2) *is equivalent to finding the smallest superharmonic function $\hat{V}$ which dominates G on E. Once $\hat{V}$ is found it follows that $V = \hat{V}$ and τ_D is the optimal stopping time.*

Generally, there are two traditional ways for finding $\hat{V}$:

1. Iterative procedure (constructive but non-explicit)

2. Free-boundary problem (explicit or non-explicit)

We are going to use the second approach, that is to say, reduce the optimal stopping problem to free-boundary problem.

1.2.3 Reduction to Free-boundary Problem

As we talked before, the optimal stopping problem should be reduced to the free-boundary problem. Then it follows that $\hat{V}$ and C should solve the free-boundary problem

$$\mathcal{L}_X \hat{V} \leq 0 \quad (\hat{V} \text{ minimal}), \tag{1.5}$$

$$\hat{V} \geq G \quad (\hat{V} > G \text{ on } C \quad \& \quad \hat{V} = G \text{ on } D)$$

where $\mathcal{L}_X$ is the infinitesimal generator of X. Let $V = \hat{V}$, then it follows that V admits the following representation:

$$V(x) = \mathsf{E}_x G(X_{\tau_D})$$

for $x \in E$ where τ_D is the first entry time of X into D. Thus V solves the following Dirichlet problem:

$$\mathcal{L}_X V = 0 \quad \text{in } C, \tag{1.6}$$

$$V|_D = G|_D. \tag{1.7}$$

If X starting at ∂C enters region D immediately then we have the smooth fit condition

$$\left.\frac{\partial V}{\partial x}\right|_{\partial C} = \left.\frac{\partial G}{\partial x}\right|_{\partial C} \quad \text{(smooth fit)} \tag{1.8}$$

where $d = 1$ for simplicity. Infinite horizon problems in dimension one are generally easier than finite horizon problems since the equation (1.6) can often be solved explicitly. However, for finite horizon problems, equation (1.6) contains a $\partial/\partial t$ term so the equation cannot be solved explicitly. To solve such problems, let us consider the optimal stopping problem (1.3), where the supremum is taken over all stopping times τ of X, and $X_t = x$ under $\mathsf{P}_{t,x}$ for $(t,x) \in [0,T] \times E$. When X is a diffusion and ∂C is sufficiently regular, we see that equation (1.6), (1.7) and (1.8) can then be written:

$$\begin{aligned} V_t + \mathcal{L}_X V &= 0 \quad \text{in } C, \\ V|_D &= G|_D \\ \left.\frac{\partial V}{\partial x}\right|_{\partial C} &= \left.\frac{\partial G}{\partial x}\right|_{\partial C} \quad \text{(smooth fit)} \end{aligned}$$

where $d = 1$ for simplicity. We cannot get an explicit solution to such problems, so we will solve the problems by numerical methods.

Chapter 2

Pairs Trading with Infinite Horizon

2.1 Ornstein-Uhlenbeck Process

Definition 2.1. *The Ornstein-Uhlenbeck process (named after Leonard Ornstein and George Eugene Uhlenbeck), also known as the mean-reverting process, is a stochastic process X_t given by the following stochastic differential equation:*

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t$$

where θ, μ and $\sigma > 0$ are parameters and W_t denotes a Wiener process.

The Ornstein-Uhlenbeck process is an example of a Gaussian process that has a bounded variance and a stationary probability distribution. The difference between an Ornstein-Uhlenbeck process and a Wiener process is in their "drift" term. For the Wiener process the drift term is constant, whereas for the Ornstein-Uhlenbeck process it is dependent on the current value of the process: if the value is less than the mean, the drift will be positive; if the value is greater than the mean, the drift will be negative. In other words, the mean acts as an equilibrium level for the process, that is why we call it mean-reverting process. By Itô's lemma and Itô isometry, we can show that

$$\begin{aligned}\mathsf{E}(X_t) &= x_0e^{-\theta t} + \mu(1-e^{-\theta t}) &(2.1)\\ \mathsf{Var}(X_t) &= \frac{\sigma^2}{2\theta}(1-e^{-2\theta t}) &(2.2)\end{aligned}$$

where x_0 is initial value, and while $t \to \infty$, $\mathsf{Var}(X_t) \to \frac{\sigma^2}{2\theta}$, which is called the stationary variance. To motivate the particular form above, we apply Itô's lemma to the function

$$f(X_t,t) = X_te^{\theta t}$$

to get

$$\begin{aligned}df(X_t,t) &= \theta X_te^{\theta t}dt + e^{\theta t}dX_t\\ &= e^{\theta t}\theta\mu dt + \sigma e^{\theta t}dW_t.\end{aligned}$$

Integrating from 0 to t we get

$$X_te^{\theta t} = X_0 + \int_0^t e^{\theta s}\theta\mu ds + \int_0^t \sigma e^{\theta s}dW_s,$$

hence we see

$$X_t = X_0e^{-\theta t} + \mu(1-e^{-\theta t}) + \int_0^t \sigma e^{\theta(s-t)}dW_s.$$

Then taking expectations on both sides gives (2.1). Denote $s\wedge t = \min(s,t)$ we can use Itô isometry to calculate the covariance function by

$$\begin{aligned}\mathsf{Cov}(X_s,X_t) &= \mathsf{E}[(X_s-\mathsf{E}[X_s])(X_t-\mathsf{E}[X_t])]\\ &= \mathsf{E}\left[\int_0^s \sigma e^{\theta(u-s)}dW_u \int_0^t \sigma e^{\theta(v-t)}dW_v\right]\\ &= \sigma^2e^{-\theta(s+t)}\mathsf{E}\left[\int_0^s e^{\theta u}dW_u \int_0^t e^{\theta v}dW_v\right]\\ &= \frac{\sigma^2}{2\theta}e^{-\theta(s+t)}(e^{2\theta(s\wedge t)}-1).\end{aligned}$$

Then let $s=t$, we can get (2.2). Figure 2.1 shows what Ornstein-Uhlenbeck processes look like when given different initial values. The picture is done by R and the code is attached to the appendix.

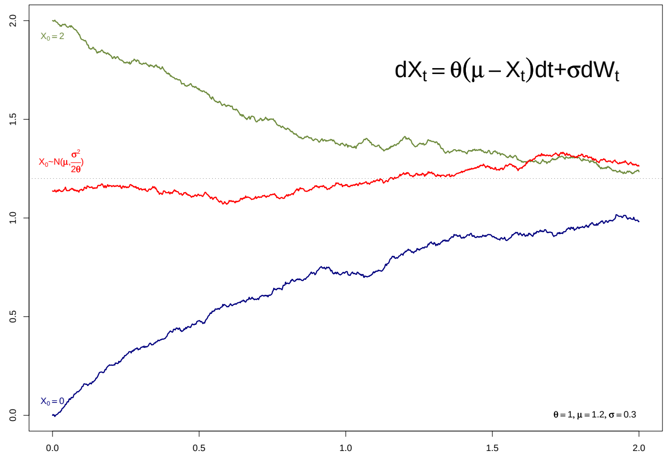

Figure 2.1: Ornstein-Uhlenbeck Process

2.2 Model Building

Recall from Chapter 1 that pairs trading is a kind of market-neutral strategy which involves holding a long position in one stock and a short position in the other, in a predetermined ratio. When the correlation broke down, i.e. one stock traded up while the other traded down, we would sell the outperforming stock and buy the underperforming one, betting that the "spread" between the two would eventually converge. According to the last subsection, we can consider the spread itself as a mean-reverting process. For simplicity, it is assumed that $\mu = 0$, which can be achieved by buying a certain number of shares of one stock and selling one share of the other stock. Theoretically, the longer you hold your portfolio, the more potential value your portfolio will have. However, it is not the case, you have to pay for the time. Thus, we need a discount factor $e^{-r\tau}$ in the gain function. Mathematically, the optimal stopping problem can be expressed as follows: let A and B be pairs, X_t be spread of the pairs. Then X_t is a mean-reverting process given by the following SDE

$$dX_t = -\theta X_t dt + \sigma dW_t \tag{2.3}$$

where $\theta > 0$, $\sigma > 0$. $V(x)$ is then given by

$$V(x) = \sup_{\tau} \mathsf{E}_x \, e^{-r\tau} X_\tau \tag{2.4}$$

where τ is stopping time, r is risk-free interest rate. In Chapter 1, we have seen that in order to solve such problems, we often reduce them to free-boundary problems. We present the model first, then explain it in detail. Mathematically, the free-boundary problem is given by

$$\begin{cases} \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} - \theta x \frac{\partial V}{\partial x} - rV = 0 \quad (-\infty < x < b) & (2.5) \\ V(b) = b \quad (x \geq b) & (2.6) \\ V'(b) = 1 \quad (x = b) & (2.7) \\ V(-\infty) = 0 & (2.8) \\ V(x) > x \quad (-\infty < x < b) & (2.9) \end{cases}$$

where b is its free-boundary.

To motivate the particular form of (2.5), we argue as follows. To distinguish the two functions, assume $V(x)$ is the solution of (2.5), we need to prove that $V(x) = V^*(x)$, where $V^*(x)$ is given by (2.4) above. Let $Y_t = e^{-rt}V(X_t)$, then Itô's formula can be applied to Y_t in its standard form. This gives (see more details in Appendix A)

$$dY_t = e^{-rt}\sigma V'(X_t)dW_t + e^{-rt}(\mathcal{L}V - rV)(X_t)I_{\{X_t > b\}}dt \tag{2.10}$$

Integrate from 0 to t to get

$$\begin{aligned} Y_t = e^{-rt}V(X_t) \quad &= \quad V(x) + \int_0^t e^{-rs}(\mathcal{L}V - rV)(X_s)I_{\{X_s > b\}}ds \\ &\quad + \int_0^t e^{-rs}\sigma V'(X_s)dW_s. \end{aligned} \tag{2.11}$$

We can see that Y_t is a martingale, when $X_s \leq b$; a supermartingale, when $-\infty < X_s < \infty$. Then by the definition of supermartingale and for every stopping time τ of X, we have

$$\mathsf{E}_x \, e^{-r\tau}V(X_\tau) = \mathsf{E}_x Y_\tau \leq Y_0 = V(X_0) = V(x). \tag{2.12}$$

Put together (2.9) and (2.12), then take the supremum over all stopping times τ of X, we can conclude that

$$\sup_{\tau} \mathsf{E}_x \, e^{-r\tau} X_\tau \leq V(x), \tag{2.13}$$

which is $V^*(x) \leq V(x)$. To prove the reverse inequality (equality), let

$$\tau_b = \inf\{t \geq 0 : X_t \geq b\}, \tag{2.14}$$

that is τ_b is the first time that X reaches b. Then we have

$$V(X_{\tau_b}) = V(b) = b = X_{\tau_b}.$$

For τ_b, all inequalities are equalities, and we observe from (2.11) upon using (2.5) that

$$\mathsf{E}_x \, e^{-r\tau_b} V(X_{\tau_b}) = V(x) = \mathsf{E}_x \, e^{-r\tau_b} X_{\tau_b}. \tag{2.15}$$

This shows that τ_b is optimal in (2.4). Then take supremum over τ

$$\sup_{\tau} \mathsf{E}_x \, e^{-r\tau} X_\tau \geq V(x) \tag{2.16}$$

Thus, (2.13) together with (2.16) give $V(x) = V^*(x)$, and the proof is complete.

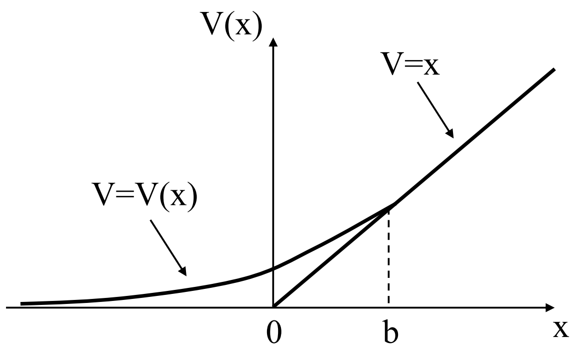

Figure 2.2: Pairs Value with Infinite Horizon

Figure 2.2 shows the shape of pairs value with infinite horizon. I would like to explain it intuitively. Generally, we enter a pairs trading when $x < 0$, that means the spread deviates from the mean and exit when x reaches b, our profit will then be b. It is not optimal to exit when x just equals 0, because the drift in our model is 0, it is better to wait till x hits the free-boundary b while it fluctuates around the origin.

Now, we are moving to the solution part. The general solution to $\frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial x^2} - \theta x\frac{\partial V}{\partial x} - rV = 0$ has the form

$$V(x) = CF(x) + DG(x)$$

where

$$\begin{aligned} F(x) &= \int_0^\infty u^{\beta-1}e^{kxu-u^2/2}du \\ G(x) &= \int_0^\infty u^{\beta-1}e^{-kxu-u^2/2}du \end{aligned}$$

C and D, k and β are suitably chosen constants. Take the first and second derivative of $F(x)$ with respect to x

$$\begin{aligned} F'(x) &= k\int_0^\infty u^{\beta}e^{kxu-u^2/2}du \\ F''(x) &= k^2\int_0^\infty u^{\beta+1}e^{kxu-u^2/2}du \\ &= k^2\int_0^\infty u^{\beta}(u-kx)e^{kxu-u^2/2}du + k^3x\int_0^\infty u^{\beta}e^{kxu-u^2/2}du \\ &= -k^2u^{\beta}e^{kxu-u^2/2}\Big|_0^\infty + k^2\int_0^\infty \beta u^{\beta-1}e^{kxu-u^2/2}du \\ &\quad +k^3x\int_0^\infty u^{\beta}e^{kxu-u^2/2}du \\ &= \beta k^2F(x) + k^2xF'(x) \end{aligned}$$

Then compare the two equations

$$\begin{cases} V''(x) - \frac{2\theta}{\sigma^2}XV'(x) - \frac{2r}{\sigma^2}V(x) = 0 \\ F''(x) - k^2XF'(x) - \beta k^2F(x) = 0, \end{cases}$$

we have

$$\begin{cases} k = \frac{\sqrt{2\theta}}{\sigma} \\ \beta = \frac{r}{\theta}. \end{cases}$$

The boundary condition (2.8) gives us $D = 0$, since $G(-\infty) = \infty$, which is contradictive with (2.8). The solution then is given by

$$V(x) = CF(x). \tag{2.17}$$

Plug (2.6) and (2.7) into (2.17), we have

$$\begin{cases} CF(b) = b & (2.18) \\ CF'(b) = 1. & (2.19) \end{cases}$$

(2.18) gives us

$$C = \frac{b}{\int_0^\infty u^{\frac{r}{\theta}-1} e^{\frac{\sqrt{2\theta}}{\sigma}bu - \frac{u^2}{2}} du}. \tag{2.20}$$

(2.18) divided by (2.19) gives us

$$\frac{b\sqrt{2\theta}}{\sigma} \int_0^\infty u^{\frac{r}{\theta}} e^{\frac{\sqrt{2\theta}}{\sigma}bu - \frac{u^2}{2}} du = \int_0^\infty u^{\frac{r}{\theta}-1} e^{\frac{\sqrt{2\theta}}{\sigma}bu - \frac{u^2}{2}} du. \tag{2.21}$$

From (2.21), we can see that b is difficult to be solved explicitly, but we can solve it numerically. Assume $r = 0.1$, $\sigma = 0.5$, $\theta = 0.5$, then (2.21) becomes

$$2b \int_0^\infty u^{0.2} e^{2bu - \frac{u^2}{2}} du = \int_0^\infty u^{-0.8} e^{2bu - \frac{u^2}{2}} du.$$

Let $f := 2b \int_0^\infty u^{0.2} e^{2bu - \frac{u^2}{2}} du - \int_0^\infty u^{-0.8} e^{2bu - \frac{u^2}{2}} du$, using Maple to plot it gives figure 2.3. When $f(b) = 0$, $b = 0.6982557296$.

Inserting (2.20) into (2.17) yields

$$V(x) = \frac{b}{\int_0^\infty u^{\frac{r}{\theta}-1} e^{\frac{\sqrt{2\theta}}{\sigma}bu - \frac{u^2}{2}} du} \int_0^\infty u^{\frac{r}{\theta}-1} e^{\frac{\sqrt{2\theta}}{\sigma}xu - \frac{u^2}{2}} du \quad (\text{for } x < b) \tag{2.22}$$

Plug $b = 0.6982557296$ into (2.22)

$$V(x) = \frac{0.6982557296}{\int_0^\infty u^{\frac{r}{\theta}-1} e^{0.6982557296\frac{\sqrt{2\theta}}{\sigma}u - \frac{u^2}{2}} du} \int_0^\infty u^{\frac{r}{\theta}-1} e^{\frac{\sqrt{2\theta}}{\sigma}xu - \frac{u^2}{2}} du, \tag{2.23}$$

which is the solution of (2.5).

2.3 Model Analysis

The solution to equation (2.5) looks complicated, it contains three parameters: r, σ and θ. Recall that when we assume $r = 0.1$, $\sigma = 0.5$, $\theta = 0.5$, b

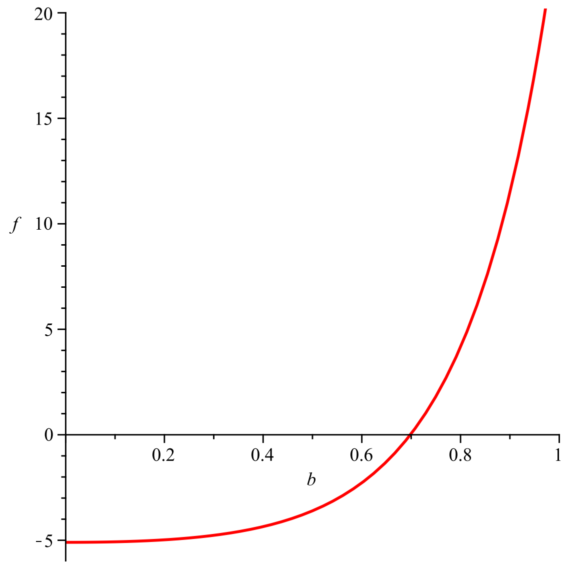

Figure 2.3: The Figure of $f(b)$

can be derived numerically. In the former case, $b = 0.6982557296$. So we can conclude that the change in each parameter will lead to the change in b and even in $V(x)$. This section is going to study what happens to b and $V(x)$ if we alter one of the parameters above.

Let us start with the parameter r. r is risk-free rate in our model. When r is smaller, we say, $r = 0.01$ instead of $r = 0.1$ (it happens in financial crisis) and keep the other parameters unchanged, we can calculate b by Maple. So $b_r = 1.110813194$, which is bigger than 0.6982557296. In the same way, while σ and θ become smaller, assume $\sigma = 0.3$ and $\theta = 0.3$, the corresponding $b_\sigma = 0.4189534378$ and $b_\theta = 0.7821082007$ can be derived respectively as well. Thus, in infinite horizon, if we fix each parameter, the optimal exercise boundary is also fixed, that is to say, the optimal exercise boundary does not depend on time. In the next chapter, we will see that the optimal exercise boundary with finite horizon is not constant anymore.

Figure 2.4 illustrates $V(x)$ with different parameters. From the picture, we see that when r and θ decrease in value, the optimal exercise boundary tends to be increased in value; when σ is decreased in value, the optimal exercise boundary tends to be decreased in value. In the real world, we can get the same conclusion. When interest rate is decreasing, people prefer to put money in the stock market rather than put money in the bank, then the price of pairs value tends to be increased, so the optimal exercise boundary will be increased as well. When σ is decreasing, the spread price is more stable, so the optimal exercise boundary tends to be decreased in value. When θ is decreasing, it takes more time for the spread price to revert to mean, so the spread price has more chance to hit the optimal exercise boundary, the optimal exercise boundary will be increased in value.

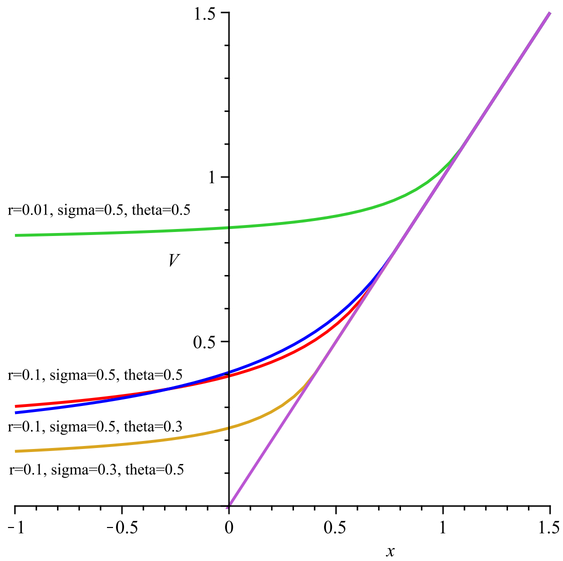

Figure 2.4: $V(x)$ with Different Parameters

Chapter 3

Pairs Trading with Finite Horizon

Now, we are moving to finite horizon problems. Those problems are difficult, and typically impossible to solve analytically, so we should solve them numerically. From [5], we know that there are many ways to solve such problems.

- Monte Carlo Methods, which involve valuing and analyzing portfolios by simulating the various sources of uncertainty affecting their value, and then determining their average value over the range of resultant outcomes.

- Finite Element Methods, which involve finding approximate solutions of PDEs as well as of integral equations. The solution approach is based either on eliminating the differential equation completely, or rendering the PDEs into an approximating system of ODEs, which are then numerically integrated using standard techniques such as Euler's Method, Runge-Kutta, etc.

- Finite Difference Methods, which involve approximating the solution to differential equations using finite difference equations to approximate derivatives, which can be solved either directly or iteratively.

In this section, we will focus on finite difference methods and then apply it to our optimal stopping problem.

3.1 Finite Difference Methods

Finite difference methods value a derivative by solving the differential equation that the derivatives satisfies. The differential equation is converted into a set of difference equations, and the difference equations are solved iteratively.

To illustrate the approach, we follow [6] and consider how it might be used to value the pairs trading strategy in our case. Recall that in Chapter 1 the optimal stopping problem with finite horizon is given by

$$V(t,x) = \sup_{0 \le \tau \le T-t} \mathsf{E}_{t,x}(e^{-r\tau} X_{t+\tau}) \tag{3.1}$$

where τ is a stopping time and T is the final time. As we argued above, to solve such a problem we often reduce it to a free-boundary problem, which is given by

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}\sigma^2 \dfrac{\partial^2 V}{\partial x^2} - \theta x \dfrac{\partial V}{\partial x} - rV = 0 \quad (-\infty < x < b(t)) & (3.2) \\ V(b(t),t) = b(t) \quad (x \ge b(t)) & (3.3) \\ V_x(b(t),t) = 1 & (3.4) \\ V(-\infty, t) = 0 & (3.5) \\ V(x,T) = \max(x,0) & (3.6) \end{cases}$$

where $b(t)$ is free-boundary. Intuitively, we can see that $b(t)$ is decreasing as t approaches to T. The reason for this is that when we have less time left, we have less profit expectation, which leads to the decreasing optimal exercise boundaries. We divide T into N equally spaced intervals of length $\Delta t = T/N$. A total of $N+1$ times are therefore considered

$$0,\ \Delta t,\ 2\Delta t,\ \ldots,\ T$$

Suppose that $x_{\max}$ is the spread price sufficiently high and $x_{\min}$ is the spread price sufficiently low that, when it is reached, the pairs has virtually no value.

We define $\Delta x = (x_{\max} - x_{\min})/M$ and consider a total of $M + 1$ equally spaced spread prices:

$$x_{\min},\ \ldots,\ -2\Delta x,\ -\Delta x,\ 0,\ \Delta x,\ 2\Delta x,\ \ldots,\ x_{\max}$$

The time points and spread price points define a grid consisting of a total of $(M + 1)(N + 1)$ points. The (i, j) point on the grid is the point that corresponds to time $i\Delta t$ and spread price $j\Delta x$. We will use the variable $V_{i,j}$ to denote the value of the pairs at the (i, j) point.

3.1.1 Implicit Finite Difference Method

For an interior point(i, j) on the grid, $\partial V_{i,j}/\partial x$ can be approximated as

$$\frac{\partial V_{i,j}}{\partial x} = \frac{V_{i,j+1} - V_{i,j}}{\Delta x} \tag{3.7}$$

or as

$$\frac{\partial V_{i,j}}{\partial x} = \frac{V_{i,j} - V_{i,j-1}}{\Delta x} \tag{3.8}$$

Equation (3.7) is known as the *forward difference approximation*; equation (3.8) is known as the *backward difference approximation*. We use a more symmetrical approximation by averaging the two:

$$\frac{\partial V_{i,j}}{\partial x} = \frac{V_{i,j+1} - V_{i,j-1}}{2\Delta x} \tag{3.9}$$

For $\partial V_{i,j}/\partial t$, we will use a forward difference approximation so that the value at time $i\Delta t$ is related to the value at time $(i + 1)\Delta t$:

$$\frac{\partial V_{i,j}}{\partial t} = \frac{V_{i+1,j} - V_{i,j}}{\Delta t} \tag{3.10}$$

Then a finite difference approximation for $\partial^2 V_{i,j}/\partial x^2$ at the (i, j) point is

$$\frac{\partial^2 V_{i,j}}{\partial x^2} = \left(\frac{V_{i,j+1} - V_{i,j}}{\Delta x} - \frac{V_{i,j} - V_{i,j-1}}{\Delta x}\right) \Big/ \Delta x$$

or

$$\frac{\partial^2 V_{i,j}}{\partial x^2} = \frac{V_{i,j+1} + V_{i,j-1} - 2V_{i,j}}{(\Delta x)^2} \tag{3.11}$$

substituting equations (3.9), (3.10) and (3.11) into the differential equation (3.2) and noting that $x = j\Delta x$ gives

$$\frac{V_{i+1,j} - V_{i,j}}{\Delta t} + \frac{1}{2}\sigma^2 \frac{V_{i,j+1} + V_{i,j-1} - 2V_{i,j}}{(\Delta x)^2} - \theta j \Delta x \frac{V_{i,j+1} - V_{i,j-1}}{2\Delta x} - rV_{i,j} = 0$$

for $j = 1, 2, \ldots, M-1$ and $i = 0, 1, \ldots, N-1$. Rearranging terms, we obtain

$$a_j V_{i,j-1} + b_j V_{i,j} + c_j V_{i,j+1} = V_{i+1,j} \tag{3.12}$$

where

$$\begin{aligned} a_j &= -\frac{1}{2}\theta j \Delta t - \frac{1}{2}\frac{\sigma^2}{(\Delta x)^2}\Delta t \\ b_j &= 1 + \frac{\sigma^2}{(\Delta x)^2}\Delta t + r\Delta t \\ c_j &= \frac{1}{2}\theta j \Delta t - \frac{1}{2}\frac{\sigma^2}{(\Delta x)^2}\Delta t \end{aligned}$$

Then we need to find final condition and two boundary conditions. The value of the pairs at time T is $\max(x_T, 0)$, where x_T is the spread price at time T. Hence,

$$V_{N,j} = \max(j\Delta x, 0), \quad j = 0, 1, \ldots, M \tag{3.13}$$

The value of the pairs when the spread price is $x_{\min}$ is zero. Hence,

$$V_{i,0} = 0, \quad i = 0, 1, \ldots, N \tag{3.14}$$

The difficult part is to find the value of the pairs when the spread price is $x_{\max}$. Recall that while $r = 0.1$, $\sigma = 0.5$, $\theta = 0.5$ we have $b = 0.6982557296$ from last chapter, which is the optimal exercise boundary when t tends to be ∞. In finite horizon, $b(t)$ is of course smaller than 0.6982557296. Thus, we can simply choose $x_{\max} > b$, that is $x_{\max} = \hat{b}$. So that,

$$V_{i,M} = \hat{b}, \quad i = 0, 1, \ldots, N \tag{3.15}$$

Equations (3.13), (3.14) and (3.15) define the value of the pairs along the three edges of the grid. First, the points corresponding to time $T - \Delta t$ are tackled. Equation (3.12) with $i = N-1$ gives

$$a_j V_{N-1,j-1} + b_j V_{N-1,j} + c_j V_{N-1,j+1} = V_{N,j} \tag{3.16}$$

for $j = 1, 2, \ldots, M-1$. The right-hand sides of these equations are known from equation (3.13), Furthermore, form equations (3.14) and (3.15),

$$V_{N-1,0} = 0 \tag{3.17}$$

$$V_{N-1,M} = \hat{b} \tag{3.18}$$

Equations (3.16) are therefore $M-1$ simultaneous equations that can be solved for the $M-1$ unknowns: $V_{N-1,1}$, $V_{N-1,2}$, $\ldots$, $V_{N-1,M-1}$. After this has been done, each value of $V_{N-1,j}$ is compared with $j\Delta x$. If $V_{N-1,j} < j\Delta x$, early exercise at time $T-\Delta t$ is optimal and $V_{N-1,j}$ is set equal to $j\Delta x$. The nodes corresponding to time $T-2\Delta t$ are handled in a similar way, and so on. Eventually, $V_{0,1}$, $V_{0,2}$, $\ldots$, $V_{0,M-1}$ are obtained.

3.1.2 Explicit Finite Difference Method

The implicit finite difference method has the advantage of being very robust. It always converges to the solution of the differential equation as Δx and Δt approach zero. One of the disadvantages of the implicit finite difference method is that $M-1$ simultaneous equations have to be solved in order to calculate the $V_{i,j}$ from the $V_{i+1,j}$. The method can be simplified if the values of $\partial V_{i,j}/\partial x$ and $\partial^2 V_{i,j}/\partial x^2$ at point (i,j) on the grid are assumed to be the same as at point $(i+1,j)$. Equations (3.10), (3.9) and (3.11) then become

$$\begin{aligned}
\frac{\partial V_{i+1,j}}{\partial t} &= \frac{V_{i+1,j} - V_{i,j}}{\Delta t} \\
\frac{\partial V_{i+1,j}}{\partial x} &= \frac{V_{i+1,j+1} - V_{i+1,j-1}}{2\Delta x} \\
\frac{\partial^2 V_{i+1,j}}{\partial x^2} &= \frac{V_{i+1,j+1} + V_{i+1,j-1} - 2V_{i+1,j}}{(\Delta x)^2}
\end{aligned}$$

The difference equation is

$$\begin{aligned}
\frac{V_{i+1,j} - V_{i,j}}{\Delta t} + \frac{1}{2}\sigma^2 \frac{V_{i+1,j+1} + V_{i+1,j-1} - 2V_{i+1,j}}{(\Delta x)^2} \\
-\theta j \Delta x \frac{V_{i+1,j+1} - V_{i+1,j-1}}{2\Delta x} - rV_{i,j} = 0
\end{aligned}$$

or

$$V_{i,j} = a_j^* V_{i+1,j-1} + b_j^* V_{i+1,j} + c_j^* V_{i+1,j+1} \tag{3.19}$$

where

$$\begin{aligned}
a_j^* &= \frac{1}{1+r\Delta t}\left(\frac{1}{2}\frac{\sigma^2}{(\Delta x)^2}\Delta t + \frac{1}{2}\theta j \Delta t\right) \\
b_j^* &= \frac{1}{1+r\Delta t}\left(1 - \frac{\sigma^2}{(\Delta x)^2}\Delta t\right) \\
c_j^* &= \frac{1}{1+r\Delta t}\left(\frac{1}{2}\frac{\sigma^2}{(\Delta x)^2}\Delta t - \frac{1}{2}\theta j \Delta t\right)
\end{aligned}$$

The final condition and boundary conditions are exactly the same as those of the implicit method.

3.1.3 The Crank-Nicolson Method

The Crank-Nicolson method is an average of the explicit and implicit methods. For the implicit method, equation (3.12) gives

$$V_{i,j} = a_j V_{i-1,j-1} + b_j V_{i-1,j} + c_j V_{i-1,j+1}$$

For the explicit method, equation (3.19) gives

$$V_{i-1,j} = a_j^* V_{i,j-1} + b_j^* V_{i,j} + c_j^* V_{i,j+1}$$

The Crank-Nicolson method averages these two equations to obtian

$$V_{i,j} + V_{i-1,j} = a_j V_{i-1,j-1} + b_j V_{i-1,j} + c_j V_{i-1,j+1} + a_j^* V_{i,j-1} + b_j^* V_{i,j} + c_j^* V_{i,j+1}$$

Putting

$$W_{i,j} = V_{i,j} - a_j^* V_{i,j-1} - b_j^* V_{i,j} - c_j^* V_{i,j+1}$$

we obtain

$$W_{i,j} = a_j V_{i-1,j-1} + b_j V_{i-1,j} + c_j V_{i-1,j+1} - V_{i-1,j}$$

This shows that implementing the Crank-Nicolson method is similar to implementing the implicit finite difference method. However, it is more numerically intensive than implicit method as it requires solving a more complex system of linear equations. We just use implicit and explicit finite difference method to analyze our model, this approach is not going to be used in this thesis. For more information about this method, please see [7].

Remark 3.1. *When we use explicit method, we have to consider its the stability and convergence. [7] tells us that explicit method is only conditionally stable for some special time-step, Δt; while implicit method and Crank-Nicolson method are unconditionally stable. The advantage of the Crank-Nicolson method is that it has faster convergence than both the explicit and the implicit method.*

3.2 Numerical Solution–the θ-method

The implicit and explicit methods require specific form, such as equations (3.12) and (3.19), and you have to compute them respectively. Furthermore, the equations are hard to solve, because you have to consider x and t at the same time in a big matrix. Hence, in order to get rid of such inconveniences, we introduce the θ^{*}[1]-method.

We consider equation (3.2), rearranging terms, we obtain

$$\frac{\partial V}{\partial t} = \theta x \frac{\partial V}{\partial x} - \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} + rV$$

Using central difference on space and define $U(t) = \frac{\partial V}{\partial t}$ yield

$$\begin{aligned} U(t) &= \theta x \frac{V(x+\Delta x) - V(x-\Delta x)}{2\Delta x} \\ &\quad - \frac{1}{2}\sigma^2 \frac{V(x+\Delta x) - 2V(x) + V(x-\Delta x)}{(\Delta x)^2} + rV(x) \\ &= \left(- \frac{\theta x}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) V(x-\Delta x) + \left(\frac{\sigma^2}{(\Delta x)^2} + r \right) V(x) \\ &\quad + \left(\frac{\theta x}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) V(x+\Delta x) \end{aligned}$$

When $j = 1$,

$$\begin{aligned} U_1(t) &= \left(- \frac{\theta(x_{\min} + \Delta x)}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) V_0 + \left(\frac{\sigma^2}{(\Delta x)^2} + r \right) V_1 \\ &\quad + \left(\frac{\theta(x_{\min} + \Delta x)}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) V_2 \\ &= b_1 V_1 + c_1 V_2 \end{aligned}$$

[1] In order to distinguish the θ in equation (2.5), we use θ^* instead of θ here.

When $j = 2, \ldots, M-2$,

$$\begin{aligned}
U_j(t) &= \left(-\frac{\theta(x_{\min} + j\Delta x)}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) V_{j-1} + \left(\frac{\sigma^2}{(\Delta x)^2} + r \right) V_j \\
&\quad + \left(\frac{\theta(x_{\min} + j\Delta x)}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) V_{j+1} \\
&= a_j V_{j-1} + b_j V_j + c_j V_{j+1}
\end{aligned}$$

When $j = M-1$,

$$\begin{aligned}
U_{M-1}(t) &= \left(-\frac{\theta(x_{\max} - \Delta x)}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) V_{M-2} + \left(\frac{\sigma^2}{(\Delta x)^2} + r \right) V_{M-1} \\
&\quad + \left(\frac{\theta(x_{\max} - \Delta x)}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) V_M \\
&= a_{M-1} V_{M-2} + b_{M-1} V_{M-1} + \left(\frac{\theta(x_{\max} - \Delta x)}{2\Delta x} - \frac{\sigma^2}{2(\Delta x)^2} \right) x_{\max} \\
&= a_{M-1} V_{M-2} + b_{M-1} V_{M-1} + \beta x_{\max}
\end{aligned}$$

So the linear system is given by

$$\mathbf{V}'(t) = \mathbf{A}\mathbf{V}(t) + \mathbf{c}(t) \tag{3.20}$$

where

$$\mathbf{V}(t) = \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_{M-2} \\ V_{M-1} \end{pmatrix}, \ \mathbf{A} = \begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{M-2} & b_{M-2} & c_{M-2} \\ & & & a_{M-1} & b_{M-1} \end{pmatrix}, \ \mathbf{c}(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \beta x_{\max} \end{pmatrix}$$

Generally, from [8], we know that the θ^*-method for solving (3.20) is defined by the difference equations

$$\xi(t+\Delta t) = \xi(t) + \Delta t\{\theta^*[\mathbf{A}\xi(t) + \mathbf{c}(t)] + (1-\theta^*)[\mathbf{A}\xi(t+\Delta t) + \mathbf{c}(t+\Delta t)]\} \tag{3.21}$$

where θ^* is a scalar and $\xi(0) = \mathbf{V_0}$. Mostly θ^* is a value in the interval $[0, 1]$. The θ^*-method has other names for special values of θ^*. Thus $\theta^* = 1$ yields the familiar forward method

$$\xi(t + \Delta t) = \xi(t) + \Delta t[\mathbf{A}\xi(t) + \mathbf{c}(t)],$$

$\theta^* = 0$ yields the backward method

$$\xi(t + \Delta t) = \xi(t) + \Delta t[\mathbf{A}\xi(t + \Delta t) + \mathbf{c}(t + \Delta t)],$$

while $\theta^* = \frac{1}{2}$ determines the Crank-Nicolson method

$$\xi(t + \Delta t) = \xi(t) + \frac{\Delta t}{2}\{\mathbf{A}[\xi(t) + \xi(t + \Delta t)] + \mathbf{c}(t) + \mathbf{c}(t + \Delta t)\}$$

Rewriting (3.21) as

$$[\mathbf{I} - (1-\theta^*)\Delta t\mathbf{A}]\xi(t+\Delta t) = [\mathbf{I} + \theta^*\Delta t\mathbf{A}]\xi(t) + \Delta t[(1-\theta^*)\mathbf{c}(t+\Delta t) + \theta^*\mathbf{c}(t)]$$

where $\mathbf{I}$ is identity matrix. We see that the θ^*-method is explicit only for $\theta^* = 1$, for all other values of θ^* the method is implicit requiring the solution of a linear algebraic system of equations at each step. Rewriting equation (3.21) as the form of follows

$$\xi(t + \Delta t) = \mathbf{B}\xi(t) + \mathbf{d}(t) \tag{3.22}$$

where

$$\begin{aligned} \mathbf{B} &= [\mathbf{I} - (1-\theta^*)\Delta t\mathbf{A}]^{-1}[\mathbf{I} + \theta^*\Delta t\mathbf{A}], \\ \mathbf{d}(t) &= \Delta t[\mathbf{I} - (1-\theta^*)\Delta t\mathbf{A}]^{-1}[(1-\theta^*)\mathbf{c}(t+\Delta t) + \theta^*\mathbf{c}(t)]. \end{aligned}$$

[8] gives us the following theorem:

Theorem 3.2. *Assume that equation* (3.20) *is asymptotically stable and that* $\mathbf{c}(t)$ *is bounded, all* $t \geq 0$. *Then*

- *Equation* (3.22), *with* $0 \leq \theta^* \leq \frac{1}{2}$, *is asymptotically stable* $\forall \Delta t > 0$. *(Unconditional stability)*

- *Equation* (3.22), *with* $\frac{1}{2} < \theta^* \leq 1$, *is asymptotically stable if and only if*

$$\Delta t < -\frac{2Re(\lambda_j)}{(2\theta^* - 1)|\lambda_j|^2}, \quad j = 1, 2, \ldots, \quad (\textit{Conditional stability})$$

where λ_j, $j = 1, 2, \ldots$, *denote the eigensolutions of* $\mathbf{A}$.

In our case, we should make a change. Because we use the final condition instead of initial condition, the order of calculation should be other way, and the value of θ^* is opposite as well. For example, now we can see that the θ^*-method is explicit only for $\theta^* = 0$, not for $\theta^* = 1$.

We can use Matlab to present the solutions. For explicit method, we let $\theta^* = 0$. In order to satisfy the conditional stability, we let Δt be small enough, say, $\Delta t = 0.0001$. Figure 3.1 shows the solutions to equation (3.2)–(3.6), which is stable as we expected. Because the time interval is fairly small, we only present the curves of $t = 0$ and $t = T$. Figure 3.2 shows the optimal exercise boundaries, which are consistent with the fact that $b(t)$ is decreasing as t increases. Note that the optimal exercise boundary is not unique anymore.

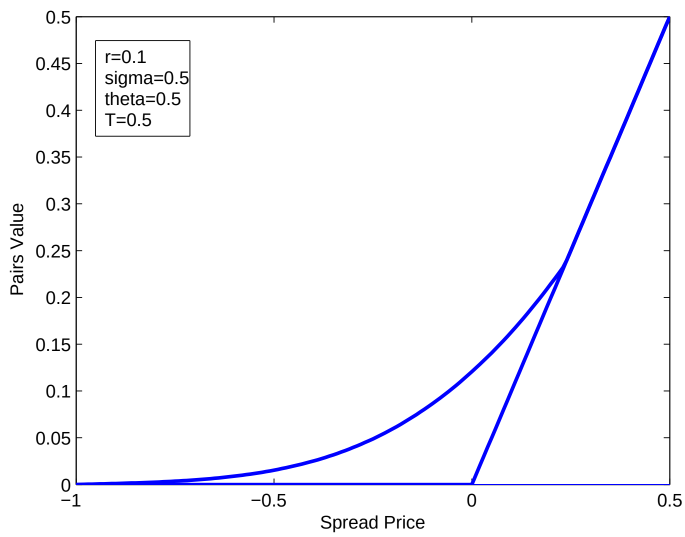

Figure 3.1: Explicit Method

For implicit method, we also let $r = 0.1$, $\sigma = 0.5$, $\theta = 0.5, T = 0.5$ but choose $\theta^* = 1$ instead. Figure 3.3 illustrates the pairs value of this method, which looks like the one we get from explicit method.

Recall that when we study optimal stopping with infinite horizon, we are also interested in what happens if we change parameters. We do the same thing in finite horizon. In order to observe clearly and draw curves with different parameters in one figure, we pick out the curves when $t = 0$ and $t = T$. Figure 3.4 shows the pairs value if we let $r = 0.01$, $\sigma = 0.3$, $\theta = 0.3$, respectively, and keep other parameters unchange. Figure 3.5 presents optimal exercise boundaries with different parameters. No matter what the parameters are, the fact always holds that $b(t)$ is decreasing as t approaches to T.

3.3 Adding A Barrier to the Spread Price

The model we have already discussed above is not consistent with what really happens in the real world, so we may worry the model is not correct. For example, in the real world, when x tends to $-\infty$, the pairs value tends to $-\infty$ as well. That means your prediction for the future movement is totally wrong, the spread price goes to the opposite way, and you will end up losing everything (of course, a skillful manager will not let it happen). However, the model tells that the pairs value is going to be zero instead of $-\infty$, so we have a contradiction here. Hence, this section is going to add an extra feature to the model.

Recall that this thesis is about optimal stopping, that is to say, we should find out the best time to exit. If the spread price never hits the optimal exercise boundary, we should just wait till time T. From figure 3.3, we see that when spread price is negative, it will by no means reach the optimal exercise boundary. However, it is different if we add a barrier to the spread price since the optimal exercise boundary will be altered. A barrier is like a stop losing threshold. When the barrier is reached, we should quit immediately or we would face the risk of losing even more. Figure 3.6 shows what pairs value with a barrier looks like. We have set the barrier $X_{\min} = -1.5$ in Figure 3.6, which means that when the barrier is hit, you quit the pairs trading, and the pairs value will become -1.5. You of course lose money in this situation, but it is acceptable. The question is how to set the barrier. If you are a risk-averse trader, you can set the barrier b' close to the entering point, say, when the spread price is $a(a < 0)$, you enter a pairs trade. If you are a risk-love trader, you can set the barrier b'' far away from the entering point, here we have $b'' < b' < a < 0$. We can also draw a picture of optimal exercise boundaries with different parameters. In order to make the picture as clear as possible, we use $r = 0.3$ instead of $r = 0.01$ and $\sigma = 0.6$ instead of $\sigma = 0.3$ here. From figure 3.7, we conclude that when we add a barrier to the spread price, $b(t)$ in most cases is still decreasing as t increases. We should also note that $b(t)$ should always be equal to or bigger than zero, because when the spread price is negative, the drift is positive, we prefer waiting to exiting since the positive drift will pull us back to the safe side.

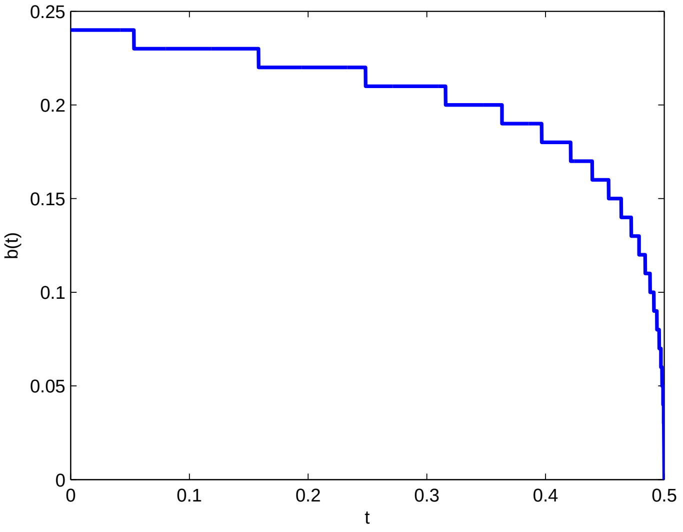

Figure 3.2: Optimal Exercise Boundaries with Explicit Method

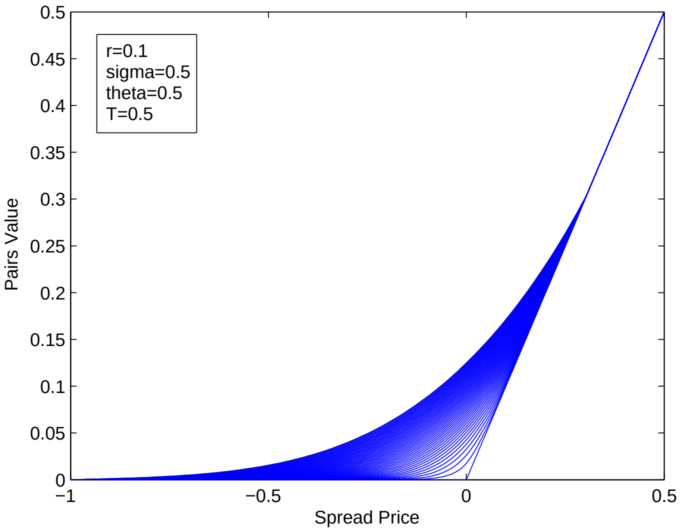

Figure 3.3: Implicit Method

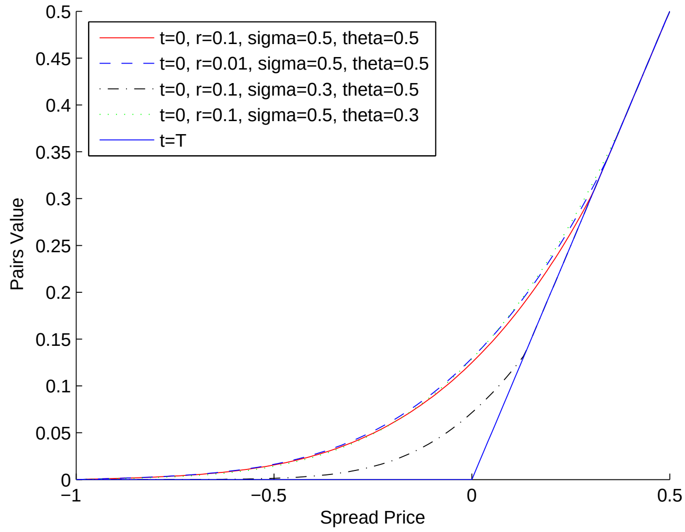

Figure 3.4: Pairs Value with Different Parameters When $t = 0$ and $t = T$

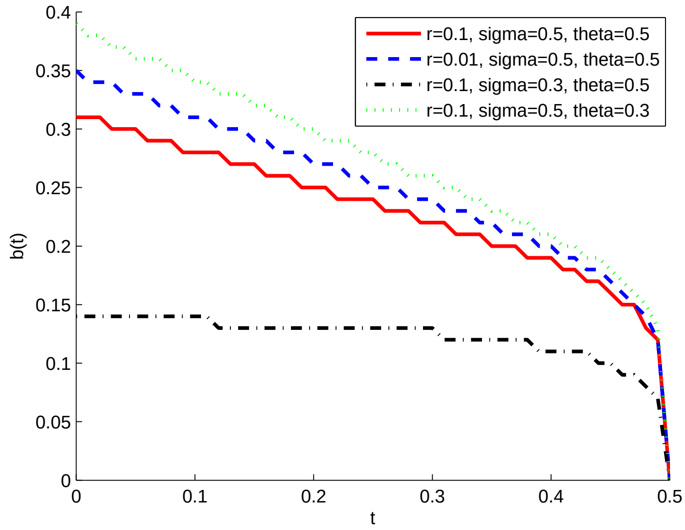

Figure 3.5: Optimal Exercise Boundaries with Different Parameters

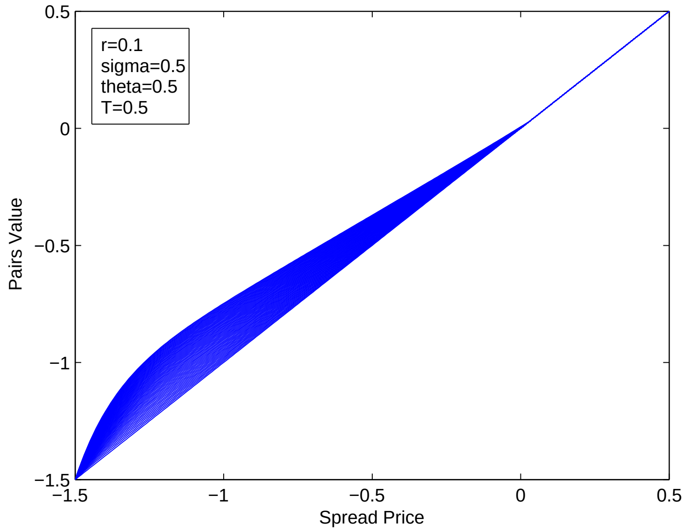

Figure 3.6: Pairs Value with A Barrier

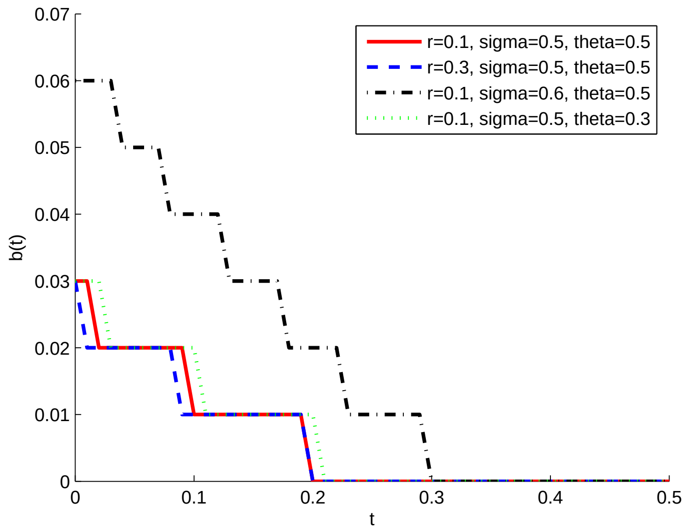

Figure 3.7: Optimal Exercise Boundaries with A Barrier

Chapter 4

Conclusion

This thesis is mainly about using optimal stopping theory to model a pairs trading, then solve the corresponding free-boundary problems both analytically and numerically.

In the first chapter, we introduce the background of pairs trading, such as the history, the trading details and even some examples in the real world. Then we discuss optimal stopping theory and reduce the optimal stopping problems to free-boundary problems. We also present some approaches to solve such problems.

The second chapter is the basic part of our thesis. In this chapter, we set up a model for valuing pairs through Ornstein-Uhlenbeck process. Then, for simplicity, we let time t be ∞, and use Maple to solve the free-boundary problem and get an optimal exercise boundary. In the third part of this chapter, we study the dependence of the boundary on the different parameters of the model.

Chapter 3 is the core part. We use finite t instead of infinite t in this part. Pairs value with finite horizon cannot be solved analytically, so we introduce a numerical method–finite difference methods. Generally, finite difference methods include two approaches, explicit method and implicit method. The θ-method is a weighted average between the two methods, so we use the θ-method to solve the PDEs both explicitly and implicitly only by changing θ. Section 3 of this part is an addition to our model. We add a barrier to the spread price to make the model more close to the real world.

Because this thesis is only focused on optimal stopping, we talk nothing about optimal entering, that is to say, finding out the best time to enter a pairs trade. We will leave such problems for further studies.

Appendix A

Recall that $dX_t = -\theta X_t dt + \sigma dW_t$, according to Itô's formula, we have

$$\begin{aligned}
dV(X_t) &= \frac{\partial V}{\partial x} dX_t + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2}(dX_t)^2 \\
&= \underbrace{(-\theta X_t \frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2})}_{\mathcal{L}V} dt + \sigma \frac{\partial V}{\partial x} dW_t \\
&= \mathcal{L}V dt + \sigma \frac{\partial V}{\partial x} dW_t
\end{aligned}$$

So consider

$$\begin{aligned}
dY_t = d\left(e^{-rt}V(X_t)\right) &= V(X_t)de^{-rt} + e^{-rt}dV(X_t) \\
&= \mathcal{L}V(X_t)e^{-rt}dt + \sigma e^{-rt}V'(X_t)dW_t - re^{-rt}V(X_t)dt \\
&= e^{-rt}(\mathcal{L}V - rV)(X_t)dt + e^{-rt}\sigma V'(X_t)dW_t
\end{aligned}$$

Now assume that $V(x)$ solves

$$(\mathcal{L}V - rV)(X_t) = 0 \quad (-\infty < x < b)$$

Hence, the equations above become

$$dY_t = e^{-rt}\sigma V'(X_t)dW_t + e^{-rt}(\mathcal{L}V - rV)(X_t)I_{\{X_t > b\}}dt$$

which is exactly what we need.

Appendix B

R Code

```
ou<-function(theta=1,mu=1.2,sigma=0.3,
init=c(mu,sigma^2/(2*theta)),from=0,to=2,steps=500*(to-from))
{
  t<-seq(from,to,length=steps)
  oup<-rnorm(n=1,mean=init[1],sd=sqrt(init[2]))
  for (i in (1:(steps-1)))
  {
    oup[i+1]=oup[i]+theta*(mu-oup[i])*(to-from)/steps+
    sigma*rnorm(n=1,mean=0,sd=sqrt(sigma^2*(to-from)/steps))
  }
  return(oup)
}

cols=c("navy","darkolivegreen4","red")
t<-seq(0,2,length=1000)
set.seed(311051)

png(filename="OrnsteinUhlenbeck3.png",width=1200,height=900,
pointsize=12)
matplot(t,ou(init=c(0,0)),type="l",xlab="",
```

```
ylab="",col=cols[1],lwd=2,ylim=c(0,2))
lines(t,ou(init=c(2,0)),col=cols[2],lwd=2)
lines(t,ou(),col=cols[3],lwd=2)
text(x=1.85,y=0,expression(list(theta==1,mu==1.2,sigma==0.3)))
text(x=1.55,y=1.75,expression(paste("d",X[t]==theta(mu-X[t]),
"dt+",sigma,"d",W[t])),cex=2.5)
text(x=0,y=0.07,expression(X[0]==0),col=cols[1])
text(x=0,y=1.92,expression(X[0]==2),col=cols[2])
text(x=0.03,y=1.29,expression(paste(X[0], "~N(",
mu,",",frac(sigma^2,2*theta), ")") ),col=cols[3])
abline(h=1.2,lty=3,col="grey25")
```

Appendix C

Matlab Code

C.1 Pairs value in finite horizon

```
% Define function
  function [x,t,c,U]=myFDM(r,sigma,theta)

% Parameter initialization
  gamma=1;        % gamma=0 implies explicit method
  h=0.01;         % spread step
  k=0.01;         % time step, k=0.0001 for explicit method
  Xmax=0.5;
  Xmin=-1;
  T=0.5;
  M=(Xmax-Xmin)/h-1;
  N=T/k+1;

% Matrix initialization
  I=eye(M);
  x=linspace(Xmin,Xmax,M+2);
  t=linspace(0,T,N);
  Ah=zeros(M,M);
```

```
    v=zeros(M,N);
    u=zeros(M+2,N);
    b=zeros(M,1);

% Define matrix Ah
    Ah(1,1)=(sigma^2)/(h^2)+r;
    Ah(1,2)=theta*(Xmin+h)/(2*h)-0.5*(sigma^2)/(h^2);
    b(1,1)=0;
    for j=2:(M-1)
        Ah(j,j)=(sigma^2)/(h^2)+r;
        Ah(j,j-1)=-theta*(Xmin+j*h)/(2*h)-0.5*(sigma^2)/(h^2);
        Ah(j,j+1)=theta*(Xmin+j*h)/(2*h)-0.5*(sigma^2)/(h^2);
    end
    Ah(M,M-1)=-theta*(Xmax-h)/(2*h)-0.5*(sigma^2)/(h^2);
    Ah(M,M)=(sigma^2/h^2)+r;
    b(M,1)=(theta*(Xmax-h)/(2*h)-0.5*(sigma^2)/(h^2))*Xmax;

% When t=T, use the final condition u(x,T)=max(x,0)
    for j=1:M
        v(j,N)=max(x(j+1),0);
    end

% When t<T, use the gamma-method
    for i=(N-1):-1:1
        v(:,i)=inv(I+gamma*k.*Ah)*...
         ((I-(1-gamma)*k.*Ah)*v(:,i+1)-k.*b);
    end

% Use u(x,t) instead of v(x,t)
    for j=1:M
        u(j+1,:)=v(j,:);
    end
```

```
u(M+2,:)=Xmax;

% Find U=max(x,u) and optimal exercise boundaries
U=zeros(M+2,N);
c=zeros(1,N);
for i=1:N
     U(:,i)=max(x(:),u(:,i));
     nodes=find((x(:))==U(:,i));
     sub=min(nodes);
     c(:,i)=x(sub);
end

% Implement with different parameters
% The first choice: r=0.1, sigma=0.5, theta=0.5
r=0.1;
sigma=0.5;
theta=0.5;
[x,t,c1,U1]=myFDM(r,sigma,theta);
N=length(t);
figure
for i=1:N
   plot(x,U1(:,i))
   hold on
end
hold off

% The second choice: r=0.01, sigma=0.5, theta=0.5
r=0.01;
sigma=0.5;
theta=0.5;
[x,t,c2,U2]=myFDM(r,sigma,theta);
figure
```

```
  hold on
  plot(x,U1(:,1),'-r')
  plot(x,U2(:,1),'--b')

% The third choice: r=0.1, sigma=0.3, theta=0.5
  r=0.1;
  sigma=0.3;
  theta=0.5;
  [x,t,c3,U3]=myFDM(r,sigma,theta);
  plot(x,U3(:,1),'-.k')

% The fourth choice: r=0.1, sigma=0.5, theta=0.3
  r=0.1;
  sigma=0.5;
  theta=0.3;
  [x,t,c4,U4]=myFDM(r,sigma,theta);
  plot(x,U4(:,1),':g')
  plot(x,U4(:,N))

% Display all results of b in one figure
  figure
  hold on
  plot(t,c1,'-r')
  plot(t,c2,'--b')
  plot(t,c3,'-.k')
  plot(t,c4,':g')
  hold off
```

C.2 Pairs value with a barrier

```
% Define function
```

```
function [x,t,c,U]=mybarrier(r,sigma,theta)

% Parameter initialization
  gamma=1;
  h=0.01;          % spread step
  k=0.01;          % time step
  Xmax=0.5;
  Xmin=-1.5;
  T=0.5;
  M=(Xmax-Xmin)/h-1;
  N=T/k+1;

% Matrix initialization
  I=eye(M);
  x=linspace(Xmin,Xmax,M+2);
  t=linspace(0,T,N);
  Ah=zeros(M,M);
  v=zeros(M,N);
  u=zeros(M+2,N);
  b=zeros(M,1);

% Define matrix Ah
  Ah(1,1)=(sigma^2)/(h^2)+r;
  Ah(1,2)=theta*(Xmin+h)/(2*h)-0.5*(sigma^2)/(h^2);
  b(1,1)=(-theta*(Xmin+h)/(2*h)-0.5*(sigma^2)/(h^2))*Xmin;
  for j=2:(M-1)
      Ah(j,j)=(sigma^2)/(h^2)+r;
      Ah(j,j-1)=-theta*(Xmin+j*h)/(2*h)-0.5*(sigma^2)/(h^2);
      Ah(j,j+1)=theta*(Xmin+j*h)/(2*h)-0.5*(sigma^2)/(h^2);
  end
  Ah(M,M-1)=-theta*(Xmax-h)/(2*h)-0.5*(sigma^2)/(h^2);
  Ah(M,M)=(sigma^2/h^2)+r;
```

```
    b(M,1)=(theta*(Xmax-h)/(2*h)-0.5*(sigma^2)/(h^2))*Xmax;

% When t=T, use the final condition u(x,T)=x
    for j=1:M
        v(j,N)=x(j+1);
    end

% When t<T, use the gamma-method
    for i=(N-1):-1:1
        v(:,i)=inv(I+gamma*k.*Ah)*...
        ((I-(1-gamma)*k.*Ah)*v(:,i+1)-k.*b);
    end

% Use u(x,t) instead of v(x,t)
    for j=1:M
        u(j+1,:)=v(j,:);
    end
    u(1,:)=Xmin;
    u(M+2,:)=Xmax;

% Find U=max(x,u) and optimal exercise boundaries
    U=zeros(M+2,N);
    c=zeros(1,N);
    for i=1:N
        U(:,i)=max(x(:),u(:,i));
        nodes=find((x(2:M+2))==U(2:M+2,i));
        sub=min(nodes);
        c(1,i)=max(x(sub),0);
    end
    c(1,N)=0;
% Implement with different parameters
% The first choice: r=0.1, sigma=0.5, theta=0.5
```

```
 r=0.1;
 sigma=0.5;
 theta=0.5;
 [x,t,c1,U1]=mybarrier(r,sigma,theta);
 N=length(t);
 figure
 for i=1:N
    plot(x,U1(:,i))
    hold on
 end
 hold off

% The second choice: r=0.3, sigma=0.5, theta=0.5
 r=0.3;
 sigma=0.5;
 theta=0.5;
 [x,t,c2,U2]=mybarrier(r,sigma,theta);
 figure
 hold on
 plot(x,U1(:,1),'-r')
 plot(x,U2(:,1),'--b')

% The third choice: r=0.1, sigma=0.6, theta=0.5
 r=0.1;
 sigma=0.6;
 theta=0.5;
 [x,t,c3,U3]=mybarrier(r,sigma,theta);
 plot(x,U3(:,1),'-.k')

% The fourth choice: r=0.1, sigma=0.5, theta=0.3
 r=0.1;
 sigma=0.5;
```

```
theta=0.3;
[x,t,c4,U4]=mybarrier(r,sigma,theta);
plot(x,U4(:,1),':g')
plot(x,U4(:,N))

% Display all results of b in one figure
figure
hold on
plot(t,c1,'-r')
plot(t,c2,'--b')
plot(t,c3,'-.k')
plot(t,c4,':g')
hold off
```

Bibliography

[1] Ganapathy Vidyamurthy *Pairs Trading – Quantitative Methods and Analysis.* John Wiley & Sons, Inc.

[2] Douglas S. Ehrman *The Hand Book of Paris Trading – Strageties Using Equities, Options and Futures.* John Wiley & Sons, Inc., 2006.

[3] Thomas S. Ferguson *Optimal Stopping and Applications.* http://www.math.ucla.edu/ tom/Stopping/Contents.html.

[4] Peskir G. and Shiryaev A. *Optimal Stopping and Free-boundary Problems.* Birkhäuser Verlag, Basel· Boston· Berlin, 2006.

[5] Jeffery J. Leader *Numerical Analysis and Scientific Computation* Addison Wesley, 2004.

[6] John C. Hull *Options, Futures and Other Derivatives.* Pearson Education, Inc.,2006.

[7] Daniel J. Duffy *Finite Difference Methods in Financial Engineering* Wiley, 2006.

[8] Owe Axelsson *The Finite Differences Method* 2000.