STARS

University of Central Florida

STARS

Electronic Theses and Dissertations, 2004-2019

2015

Calibration of Option Pricing in Reproducing Kernel Hilbert Space

Lei Ge
University of Central Florida

Part of the Mathematics Commons

Find similar works at: https://stars.library.ucf.edu/etd

University of Central Florida Libraries http://library.ucf.edu

Showcase of Text, Archives, Research & Scholarship

CALIBRATION OF OPTION PRICING IN REPRODUCING KERNEL HILBERT SPACES

by

LEI GE
M.A. Harbin University of Sciences and Technology, 2010
M.A. University of Central Florida, 2012

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Mathematics
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2015

Major Professor: M. Zuhair Nashed

ABSTRACT

A parameter used in the Black-Scholes equation, volatility, is a measure for variation of the price of a financial instrument over time. Determining volatility is a fundamental issue in the valuation of financial instruments. This gives rise to an inverse problem known as the calibration problem for option pricing. This problem is shown to be ill-posed. We propose a regularization method and reformulate our calibration problem as a problem of finding the local volatility in a reproducing kernel Hilbert space. We defined a new volatility function which allows us to embrace both the financial and time factors of the options. We discuss the existence of the minimizer by using regularized reproducing kernel method and show that the regularizer resolves the numerical instability of the calibration problem. Finally, we apply our studied method to data sets of index options by simulation tests and discuss the empirical results obtained.

Keywords: RKHS, Tikhonov regularization, local volatility, kernel estimation, Ridge Regression

To my parents Hongan Ge and Shuxian Li, and my love, James Barr.

ACKNOWLEDGMENTS

I cannot have any of this without the help of my advisor, Dr. M. Zuhair Nashed. Also thanks to Dr. Qiyu Sun who patiently showed me the direction and helped me understand the models.

I am also grateful to the Mathematics Department at the University of Central Florida for supporting me financially through my doctoral program. Last but not least, thanks to my family and my boyfriend, James Barr, for giving me all the support and helping me proofread this dissertation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$C(S, t; T, K.\sigma, \delta)$ Black-Scholes Option pricing function, page 2

$\sigma, \sigma(S/K, T-t)$ volatility function, page 1, 10

$a(y, \tau)$, $a(x)$ transformed volatility function, page 18

$S, S(t)$ stock Price, page 1

T maturity time of a option, page 1

K strike price of a option, page 1

δ stock dividend, page 2

$H_{\mathbb{K}}$ Reproducing kernel Hilbert Space, RKHS, page 14

$\mathcal{K}$ kernel function, page 14

$R_{\alpha} : Y \to X$ Tikhonov regularization operator, page

F_t filtration, page 7

$u(a)$ option price as a function of volatility, page 20

$Q = R \times [0, T] \in R^2$ the set of accessible time, stock price and strike, page20

CHAPTER 1: INTRODUCTION

An option, by definition, is something that may or may not be chosen. An option in Finance is a financial instrument that gives sellers or buyers a right, but not the obligation, to buy or sell a certain stock at a promised price in a certain period or on a certain date. In 1900, Louis Bachelier, a French mathematician, first derived the option pricing formula based on the assumption the stock prices follow a Brownian motion with zero drift. This derivation has served as the foundation for option pricing research in the years since, and the option pricing problem has been an ongoing topic of interest in Financial Mathematics.

Thanks to the Itô formula in stochastic calculus, the Brownian motion model can be illustrated nicely as a parabolic partial differential equation(PDE). Volatility, the leading character in this dissertation, appears as a parameter in the Black-Scholes partial differential equation. The estimation or recovery of the volatility parameter is then brought up and is well-known as the calibration problem of option pricing, which is also known as the inverse problem of option pricing.

There are ambiguities in the definitions of inverse and direct problems. If two problems are inverse to one another, the formulation of one is fully replied on the other. It is obviously arbitrary which of the two problems can be called direct or inverse. We usually consider a problem as the direct problem when it is studied earlier and is understood in more detail when compared to the other. In the study of option pricing, how to establish option price by giving values of the parameters, for example volatility, is recognized as the direct problem. The other direction, how to find a(some) parameter value(s) by giving observed option of price, is called the inverse problem. In the case of option pricing, the inverse can be recognized as how to find a parameter used in the Black-Scholes PDE.

Mathematically, the identification of parameters in differential equations are never new topics. For

example, we see this kind of problem in various application research areas like image processing, geophysics, medical diagnostics, signal recovery, etc. In many cases it is not always easy to get the solutions as desired given so many questions are ill-posed. It is well-known by now that the calibration problem of option pricing is, not surprisingly, also ill-posed. An inverse problem is ill-posed in that the solution is not unique and stable. By saying stable, we mean the solution does not continuously depend on data. Obtaining a unique and stable solution is a fundamental goal in establishing a robust pricing model. Moreover, in practice, market data is often sparse with respect to the maturity time, so getting a nicely featured solution is even more challenging.

Open questions and various applications have brought the attention of numerous researchers in different areas to this volatility recognition and identification problem. For example, stochastic control methods are used in Jiang's [26], least squares approximations with Tikhonov regularization application is first introduced in this area by Lagnado and Osher[8], a linearized integral equation is applied to get the solution by Isakov[18], and more recently, Tikhonov regularization has been widely studied in the calibration problem [2,3,4,8].

There are various ways to investigate the ill-posed problems. Generally, most of the approaches used in inverse problems involve one or more of the following ideas: [9]

- A change of the concept of the solution;

- A change of the topologies or spaces;

- A change of the operator itself;

- Use of a regularization operator;

- Use of probabilistic and stochastic methods.

Following these directions, instead of solving the equation directly, we use approximation of the

real solution through a family of well-posed equations in a new topological space, the Reproducing Kernel Hilbert Space, with the Tikhonov regularization technique applied. This dissertation is aimed at addressing the theoretical and numerical aspects of the calibration problem by applying Tikhonov regularization with Reproducing Kernel Hilbert Spaces in real time option data.

European options, also known as vanilla options, are the simplest form of options. Unlike their American counterparts, which may be traded on any date within the maturity period, European options may only be traded on the maturity dates. European vanilla call and put options are both considered throughout this dissertation.

Reproducing kernel method is used to recover a smooth volatility surface function. It is a nonparametric method compare to other researchers' work. As we know that a parametric model assumes a specific function structure to model the unknown, the optimal fitting is obtained by tuning or adjusting the model parameters. For instance, linear regression model assume the linear structure between dependent variable and independent variable, the slope and the intercept control the model fitting. Nonparametric model on the contrary, allow more freedom of the model structure. Especially for some complex forms of the model, nonparametric model can conquer the limitation of the forced structural assumption from the parametric models. For this reason, we choose to use kernel method to recover the volatility function.

This dissertation is planned as following:
In Chapter 2 we will discuss the models used for option pricing problem and the direct and inverse problem for option pricing. First we review the Black-Scholes equation with a generalized model proposed by Dupire and Rubinstein. We reviewed the arbitrage and expectation prices and the geometric Brownian motion model. A new volatility function is defined. Compared to others' work, we focus on nonparametric methods for the volatility function structure, especially the use of Reproducing Kernel Hilbert Spaces for the volatility functions. The newly defined volatility

function embraces both the financial and time factors that are used to price options. We discuss the constraints of the volatility function and the related kernel function choices. Based on the Dupire's equation, we get our new model for the parabolic equation. We prove that the solution exists for the new model, which proves that the direct problem of option pricing is well-posed. Then we introduce the forward operator and research the properties. We prove the new results in the reproducing kernel Hilbert spaces. We then expand the existing results from Engl and Egger. Also we analyze some background of the direct problem and showed the well-posedness of it. After that we introduce the inverse problem, *given market price, how to get volatility*. We show that the inverse problem of option pricing is closely related to a parameter identification problem in parabolic equations. Then we prove the ill-posedness of the problem as the result of compactness of the forward operator in infinite dimensional spaces.

In Chapter 3 we use Tikhonov regularization for the calibration problem of the option pricing. We define the primary optimization function by using a regularization technique and we introduce the Tikhonov regularization function. We first review the classic Tikhonov regularization theory, then we explore the nonlinear case, which is mainly based on Engl and Egger [4]. After that we have a brief review of Reproducing Kernel Hilbert Spaces, then we apply it to regularization scheme and present the well-posedness of the new posed problem, the calibration problem for option pricing in the case for continuous setting assumption in Reproducing Kernel Hilbert Spaces. We analyze the uniqueness, stability and the convergence rate of the art of regularization method.

In Chapter 4, we apply the moment discretization of operator equations and implementation of the nonlinear model is given. We present the discrete models associated with discrete data. Numerical implementation by using the gradient descent method is presented. The cross validation method is used to find the optimal parameter in the regularization method. Monte Carlo simulations are used to produce the numerical results. We apply the control variate method to reduce the simulated variance by introducing the control variate, the true data from the market. The minimal variance

is obtained in the discrete model. We apply the studied method on both European call and put options with different kernel functions, especially, we used Gaussian kernel, polynomial kernel and Laplace kernel.

Chapter 5 represents the linear model by applying option Greeks in the Black-Scholes equation. The Greeks are essential tools in option trading. They are part of the option chain data in the market that are available to download in most of the option trading resources. Although they are numerically approximated by using the finite difference method, we can use them as handy tools to estimate the option price as a quick and fast approach. The great feature of application of the greeks is that we obtained a linear model between the volatility and option price. The solution of the new proposed model is calculated and by applying Tikhonov regularization, we obtain a stable solution. Apple. Inc real world trading data are used to study the theoretical method. Numerical results and conclusion of the work are given. We also do some error analysis by using R-square, which is a common method used in statistics. It is proved that our results are very promising. We compare different kernel functions and give the conclusion that Gaussian and Laplace kernel functions are more suitable for the model. Also based on the studied method, the prediction of the future market could be obtained. We believe this method could be of good practical use.

In chapter 6, conclusions are given. Illustration comparison between ordinary least squares method and regularized method are presented in Section 6.1. Also we showed that the volatility function will be rough and noisy without using kernel functions. Then we discuss the possible directions of future work. Jumps could be added in the ordinary Black-Scholes model as the consideration of extreme events happen in the real world. As we know that lognormal models assume very slow return rate, which in reality is very unpractical. Example of such is given in 6.3. Jumping models came to attention since 1980s. The inverse problem of the models with jumps would be of great interest and machine learning techniques could be used to explore this problem in the future. Bayesian theory also can be considered, especially for prediction and inference. By

assuming the volatility itself is a random process, also known as a prior, we can find the posterior likelihood. Then we can find the maximum likelihood of the posterior (MAP). Alternately, the mean of the posterior is another common used method for the approximation. It is interesting to find the relationship between the Bayesian theory and regularization techniques. Last but not least, we propose some other machine learning method, especially, the nonparametric methods that could be used in the future studies.

First and foremost, we would like to give a brief view of the main topic in this dissertation, the *Calibration Problem*.

CHAPTER 2: BACKGROUND OF VOLATILITY RECOVERY IN OPTION PRICING PROBLEM

2.1 A Brief View of Option Pricing

Consider all the trades in a complete financial market, where cash can be lent or borrowed at a given constant interest rate r, and for a risky asset or stock $S = S(t)$ which promises a dividend q. Options are financial derivatives that give you the opportunity to buy or sell but not the obligation.

We all know that buying low and selling high will make profit in stock market. If one expects the stock price to go up, it is wise to place a call option. Purchasing this option grants the holder the contractual authority to purchase a stock on a set future maturation date at a predetermined price, or strike, noted as K, $K > 0$. The time period to maturity is noted as T, $T > 0$. If the stock price goes up, the option holder may purchase the stock at the strike, which will now be lower than the current market price, and realize a profit. Should the stock price go down, the option holder may opt to not purchase the stock and simply be out the cost of the option. For example, if you are interested in a stock today that has price $S_0 = \$90$ and you bet the stock price will rise to \$120 in two months. You buy an option with strike $K = 100$, maturity $T = 60$ days. After two months, the stock price is \$118 and you may call or buy the stock at the strike price of \$100 by exercising the option. Your net payoff is 118-100 = 18.

The attractive part of option is that you don't need to pay \$90 now to hold the stock, which reduces the investment risk. Even if the price falls, you only lose the money you paid for option and you have the choice not to exercise your options. The option price, which is also called the option premium in Finance, is a key factor in this scenario and determining how to price an option is the focus of this dissertation.

In the famous Black-Scholes framework, a complete financial market is used to model the option prices. Before we talk about the Black-Scholes formulas, we want to have a look at the assumptions that were required,

- The stock prices are assumed to be lognormally distributed and followed a geometric Brownian motion model.

- The volatility in the model is assumed to be constant.

- Underlying asset dividends are given as a fixed dividend yield.

- The risk-free interest rate is known and considered to be a constant.

- Transaction cost and taxes are ignored.

- Short sell or borrow at risk-free rate.

For a European call option on a stock that pays continuous dividends, the price of the option $C(S, t; T, K, \sigma, q)$ follows,

$$C_t + \frac{1}{2}\sigma^2 S^2 C_{SS} + (r - q)SC_s - rC = 0 \tag{2.1}$$

The European call option is with maturity T, striking price K, and the boundary condition is given as

$$C(S, t = T) = (S - K)^+ \tag{2.2}$$

Notice that $(S_T - K)^+ = max(0, S_T - K)$. The variables used here are

- S is the stock price.

- K is the strike price of the option.

- σ is the option price volatility.

- r is the annual risk-free interest rate, continuously compounding is used.

- T is the time to maturity.

- q is the continuously compounded dividend yield.

- $C(S, t; T, K, \sigma, q)$ is the price of the option, a.k.a the option premium.

Equation (2.1) is called the Black-Scholes equation, we will use B-S in short in the following, which is a backward parabolic equation. (2.2) is the payoff at expiry if the underlying is known. In our context, we do not talk about the early exercise American options. Notice the payoff function is not linear. The option price formula based on payoff function for European call option can be given as the expectation of future payoff

$$C(S, 0) = E(e^{-\int_t^T r(\tau)d\tau}(S_T - K)^+ | F_t) \tag{2.3}$$

Notice, here $e^{-\int_t^T r(\tau)d\tau}$ is the compounding interest when interest rate is considered a function of time. But in this dissertation, we only consider constant interest, then the equation (2.3) can be simplified as

$$C(S, 0) = E(e^{-rT}(S_T - K)^+ | F_t) \tag{2.4}$$

We know that the stock price follows a random movement process. But in a well known saying in statistics, things are random in non-random ways. Kolmogorov's strong law of large numbers states that the average of outcomes from a large number of independent random trials tends to move to the probability expectation. Thus applying the strong law, we use the expectation as the theoretical fair price of the options.

Definition 2.1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{A}$ is a $\sigma - algebra$ of subsets of Ω, $\mathbb{P}$ is a non-negative measure on Ω such that $\mathbb{P}(\Omega) = 1$.*

Stock price follows a random process, in this dissertation, we consider the lognormal process as used in B-S. F_t used in the equation (2.3) and (2.4) is filtration. Recall that a filtration $F_t = (\mathcal{A}_t)_{t \geq 0}$ is an increasing family of $\sigma - algebras$ $\mathcal{A}_t$); i.e., for $t > \tau$, we have

$$\mathcal{A}_\tau \subset \mathcal{A}_t \subset \mathcal{A}$$

The $\sigma - algebra$ usually represents a past history available at time t. By solving equation (2.1) with boundary condition (2.3), B-S formula for call options are,

$$C(S, T - t, K, \sigma, q) = Se^{-q(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2) \tag{2.5}$$

where
$$d_1 = \frac{\ln(S/K) + (r - q + 0.5\sigma^2)(T - t)}{\sigma\sqrt{T - t}} \tag{2.6}$$

and
$$d_2 = \frac{\ln(S/K) + (r - q - 0.5\sigma^2)(T - t)}{\sigma\sqrt{T - t}} \tag{2.7}$$

$N(x)$ denote the cumulative normal distribution function.

In order to find the put option's formula, we need the well-known put-call parity,

$$C(S, T - t, K, \sigma, q) - P(S, T - t, K, \sigma, q) = Se^{-q(T-t)} - Ke^{-r(T-t)} \tag{2.8}$$

where $P(S, T - t, K, \sigma, q)$ represents the European vanilla put option premium function under the same asset S as call, same maturity T and strike K. Then one can solve for European put option's

formula,

$$P(S, T - t, K, \sigma, q) = Ke^{-r(T-t)}N(-d_2) - Se^{-q(T-t)}N(-d_1) \qquad (2.9)$$

Based on the expectation price and arbitrage free assumption, Black and Scholes discovered the B-S equation for option prices. With application of the Itô lemma, the B-S equation is a deterministic partial differential equation (2.1). Consider interest rate and dividend are constants, then the only parameter that is not freely observable from the market in (2.1) is the volatility σ. This parameter describes the diffusion properties of the underlying process. If the diffusion process for the underlying stock is given, then the option price can be determined from the Black-Scholes equation and the payoff condition. Therefore gaining information on σ is a fundamental challenge of the option pricing problem.

Given option pricing as a function of volatility, one is interested in the inverse of this relationship to get the market parameter or implied volatility. This is called the inverse problem of option pricing. Volatility is assumed here as a constant by Black and Scholes at first. However, in practice, the constant volatility shows *smile and skew* effects with respect to in the money or out of the money conditions. Now let us have a close look at what volatility means and why it is important.

> In finance, volatility is a measure for variation of price of a financial instrument over time. It can be understood as the nervousness or the agitation of the market.
> —-Wikipedia.com

In this sense, it is used as a noise scalar in the geometric Brownian motion model. There are different kinds of volatilities in this option pricing scenario, for example, historic volatility and implied volatility. Historic volatility is calculated from the available information of past market prices. It is the standard deviation of variations of the prices in the time series of real time trading. It is also used to postulate in the classical Black-Scholes model.

An implied volatility is usually derived from some given option pricing model, for example, the Black-Scholes model, with knowledge of market prices of a traded derivative (in particular an option). From all the practical experience and research we learned that volatility is never as simple as a fixed constant, it is a function of a couple of different indicators related to the derivatives on underlying stock, so we are aiming to get volatility by giving the market prices of options. With an accurately prescribed volatility, it would not only help us to price different financial instruments but also help us to analyze and minimize risk in the market.

In order to get the consistency between the model and the market prices, we need to calibrate the model of option pricing by adjusting our volatility, so that the model can be used confidently in the market, not only for a single option, but also for portfolios and more complex financial products.

2.2 Literature Review

This section gives a short survey of calibration problem. First and foremost, we start with the most famous model in quantitative Finance, the Black-Scholes model [19]. Inspired by Louis Bachelier, they applied that the risky stock followed the Brownian motion model,

$$dS = \mu S dt + \sigma S dW(t) \tag{2.10}$$

In this theory the drift is not considered, which lies at the core of the assumption of the risk-neutral measure. This, in other words, is the arbitrage free assumption mentioned previously in Section 2.1. The risk-neutral measure is for which the discounted process is a martingale. See details at [19, 59, 60]. They derived the Black-Scholes equation (2.1) and (2.2), which was a big hit and still plays a central role in Financial world. Although Black-Scholes model gives a straight forward formula to get option prices, it erroneously assumes that the volatility is a constant. Empirical

12

studies of the implied volatilities show a dependence of volatility on the strikes and the time to maturity. The shape of the implied volatility versus strike curve is a convex, parabolic shape which is known as the *volatility smile*.

Dupire, Kani and Derman[50], Rubinstein[49, 51], Lagnado and Osher[8] have expanded the Brownian motion model by using the underlying asset which the options are written on follows the stochastic differential equation(called Geometric Brownian motion model):

$$dS = \mu S dt + \sigma(S, t) S dW(t) \tag{2.11}$$

The volatility is no longer a constant, but a deterministic function of asset price and time. Dupire applied this model to derive another commonly used equation, *Dupire's Equation*, by considering the option prices as functions of strike and maturity, which is mathematically equivalent to the Black-Scholes equation(2.1) and (2.2), but better in dealing with the smile effect,

$$-C_T + \frac{1}{2}\sigma^2 K^2 C_{KK} + (\delta - r)K C_K - qV = 0 \tag{2.12}$$

with the initial value condition for present stock price S_0

$$C(S, t = 0) = (S_0 - K)^+ \tag{2.13}$$

This equation is also called the Dupire forward equation.

A good model should match the market price consistently and, as we mentioned above, σ is crucial to the option pricing model. Therefore, the inverse problem of option pricing, i.e. the calibration problem is brought up by Dupire for the first time. Naturally, one wants to inverse the Dupire's

equation above and solving for volatility, and the volatility can be solved as

$$\sigma(T, K) = \sqrt{\left(\frac{2}{K^2 \frac{\partial^2 C}{\partial K^2}}\right)\left(\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K}\right)} \qquad (2.14)$$

A formula of volatility is simply given, however, this method of getting volatility is very dangerous for several reasons. First, in the real time trading data, the accessible option prices are only for some discrete set of strike prices and maturities. Also notice that the formula demands the differentiability of second order of the option pricing function $C(K, T)$ with respect to strike and maturity, which it is not achievable with the available data structure, since it is not even dense enough to approximately continuous. with only sparse data available in practice, the differentiability requirement used in the formula leads to instability phenomenon. Also the data is in general corrupted by noise, so an accurate interpolation for missing data is impossible. Thus simply inverting the Dupire equation to solve for volatility is known to perform poorly in reality. Due to unknown information about the boundedness of the differential operator or the second derivative operator, the results do not depend continuously on the given data [5].

Binomial or trinomial trees are used by Rubinstein[49, 51], Dupire, Deman and Kani[50] independently as a discrete approximation to the risk-neutral process for the stock prices. But again the market alone does not provide enough information to make a unique determination of the implied process. Also the implied trees have limitations: the backward induction is limited to pricing options with a single maturity, and forward induction procedures rely on the ability to interpolate and extrapolate market prices without introducing arbitrage violations. Also, it cannot be extended to complex financial products [8]. More general models with a similar degree of variability were introduced. See [2, 5, 26, 37]. Local volatility models were developed as a class of one dimensional Markov models.[52, 4, 3]. Binomial tree methods introduced by Crepey to approximate the least squares solution for Tikhonov regularization in 2003, which features as an intuitive and fast

algorithm. However, for complex financial products, it is difficult to apply.

Meanwhile, researchers have different views about the what volatility as a function depends on. For example, volatility only depends on time $\sigma(t)$, in which the nonlinear forward operator is decomposed into the inner linear Volterra integral operator and the outer Nemytskii operator[18]. Also, volatility itself was considered a stochastic process, for example, see[61, 62, 63]. Dupire, Deman and Kani considered the volatility as the form[4, 18].

$$\sigma(K, T) = \sigma(K)\rho(T) \tag{2.15}$$

Lagnado and Osher in [5] consider the volatility function as

$$\sigma(S, t) = \frac{Constant}{S} \tag{2.16}$$

They assumed specific structures for volatility which can be very hard to reconcile with reality. Those approaches utilize parametric methods and make simplifying assumptions based on beliefs. For example, by assuming volatility as function of only strike and maturity, stock price information is ignored. Alternately, assuming volatility is as a function of stock price only misses the option maturity and strike features. Assuming volatility grows inversely as the price of stock gives a very strong structure of the function though in reality we see much more complicated picture.

Contrasting with those approaches, we propose the use of a nonparametric method and introduce the use of kernel functions to give a smooth volatility surface more considerate of reality. By tuning and adjusting the coefficients of the kernel functions, we can get unique and stable volatility function. This function can be used confidently to predict and generalize for option pricing.

The calibration problem of option pricing is a well known example of an ill-posed inverse problem in which the challenge is to find a unique and stable parameter used in a parabolic partial differ-

15

ential equation. It turns out that with reguarization techniques with some *prior* information for the solution, one can construct stable approximations to the solution with the desired level of accuracy.

The use of regularization techniques to solve Ill-posed problems has gained vast theoretical and practical value since 1940s. (See [9,10]). Classic Tikhonov regularization theory is mainly for linear bounded operators. In general, the calibration problem for option pricing is a nonlinear case which creates difficulties for applying the regularization strategy. However, with the help of option greeks, we find a linear model for this problem. Although a linear model is always more desirable, it can only be used for short term approximation due to the greeks being calculated by the approximations via the finite difference method.

In summary, we want to build a model that can deal with the ill-posedness by moderately adjusting the operator equation and also consider the volatility as a function of several different factors of stocks and options as well. By considering the log-ratio of stock and strike and time to maturity, we propose a new function for volatility, which we believe reflects the true nature of options. Also we need to consider the practical value of the model which has to cope with the sparsity and noise of the data, as well as its suitability for expansion for use in more complex financial products. A nonparametric model is proposed as compared to others' work that assumes strong structures of the volatility function. With the newly defined volatility function, we reformulate the calibration problem by using Tikhonov regularization in Reproducing Kernel Hilbert Spaces. Reproducing Kernel Hilbert Space is used for the first time in the calibration problem for option pricing.

The concept of reproducing kernel can be tracked back to the paper of Zaremba in 1908. [30] It was proposed for discussing the boundary value problems of the harmonic functions. Bergman gave the most important discussion of kernel methods in 1930's. Then Mercer discovered the positive definite property of continuous kernels. In 1950, N. Aronszajn summarized the work of the predecessors and gave out a systematic theory including using the Bergman kernel function[31].

16

Since then, mathematicians have used reproducing kernel theory to solve problems for numerous special fields such as: image processing, statistical machine learning, artificial intelligence, pattern recognition, etc. See [9,10].

By applying the regularization in Reproducing Kernel Hilbert Space, we have achieved a stable and unique solution and solved the ill-posed problem not only theoretically, but also the practically. The kernel method makes it easy to deal with the sparse and noisy data and can be expanded for use with more complex financial products.

We apply the Reproducing Kernel method for both linear and nonlinear cases. In the linear model, we use the option Greeks which are ready data from everyday trading. In the nonlinear model, we use the gradient descent to find the iterative solution. Also we apply both of the models to real trading data of Apple. Inc. options. Both put and call options are evaluated with several different kernel function applications. Numerical results and comparison are given. Finally, we discuss some future directions for research.

2.3 Volatility Assumption

As we mentioned in 2.1, volatility is a critical input for B-S equation for option pricing; however, it is not directly observable from the market. Further, it is shown from the market that volatility varies over time and both theoretical and practical models are trying to discover the behavior of volatility. In practical terms, it is an important issue for market-makers to hedge volatility to stay in profitability.

In this section, we discuss the specific techniques for measuring volatility and also demonstrate how volatility models can be influential to the B-S pricing framework. A new volatility function is defined. Then we transformed the Dupire equation into a parabolic partial differential equa-

tion. The solution of the pde will be our theoretical option price. A nonparametric method with reproducing kernels is used to estimate the volatility function. Then we give a quick review for reproducing kernel Hilbert space. As a preparation of the model setting up, we have to give some assumptions about volatility function, specifically, we assume volatility function is smooth and bounded.

2.3.1 The New Volatility Function Assumption

Neuberger(1994) used $\ln(S/S_0)$ to hedge and speculate the variance of forward contracts, which is called a *log contract*. From everyday market data, we can clearly see the collinearity between the strike price and stock price. We choose to use the logarithm ratio of the stock price over the strike price to be the moneyness factor, $ln(S/K)$, and the time to maturity to reflect the time factor. $T - t$. This assumption can solve the collinearity problem between the strike prices and stock prices. The Pearson correlation is calculated with real time market data in Chapter 3. One can't help questioning what factors in fact affect the option price. Only stock price and time? Or only strike and time? By definition of the option, as a contract on a underlying stock,intuitively, we believe that both the strike price and the stock price affect the option price. Then, naturally, the ratio idea comes out, not for the first time, the log-ratio between S and K is also used in the closed form Black-Scholes solution for option prices,

$$d_1 = \frac{ln(S/K) + ((r - q) + .5\sigma^2)(T - t)}{\sigma\sqrt{(T - t)}} \tag{2.17}$$

$$d_2 = d_1 - \sigma\sqrt{(T - t)} \tag{2.18}$$

18

Where $N(x)$ is the standard normal distribution,

$$N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} dt \tag{2.19}$$

As we can see, $ln(S/K)$ and $T - t$ definitely affect the option prices from B-S pricing framework. Volatility as a measure of price changes, will be affected accordingly. With a little generalization, we can get a totally new volatility function structure, instead of only using either pair of S, t or T, K. This assumption allows us fully apply the market data and better reflect the nature of the meaning of the options. The implementation on real data details can be found in chapter 3 and 4.

Definition 2.2. *Let S be underlying stock price for option derivatives, $S > 0$, K be the option price, $K > 0$, we define the volatility function is*

$$\sigma(\ln(S/K), T - t) \tag{2.20}$$

Which compare to (2.10) and (2.11) in Engl and Lagnado's work. Lagnado's volatility function is defined as (2.14) (see [8]) In Isakov's model, the volatility is a function of t only, [56]

$$\sigma := \sigma(t) \tag{2.21}$$

Engl and Egger chose a volatility function as

$$a = \sqrt{2}\sigma(y) = 0.15 + 0.05 \exp[-(y + 0.3)^2] \sin(2\pi y) + 0.05 erf(20y) \tag{2.22}$$

where $erf(\cdot)$ is the error function (also called the Gauss error Function),

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt \tag{2.23}$$

19

and in this definition

$$y = \ln(K/S) \tag{2.24}$$

In summary, one couldn't help to ask this question what volatility function should depend on? $\sigma(S, t)$ was used in Rambustain's work; $\sigma(T, K)$ in Durpire, Derman and Kani's research; $\sigma(S)$ in Lagnado and Osher's, then $\sigma(t)$ in Isakov's. Engl and Egger used $\sigma(\ln(K/S))$. Two trends can be seen, one trend relates the volatility to stock price, as the volatility used as a standard deviation of stock annual return; the other trend concerns the volatility embraced the option parameters, maturity T and strik K, as it is shown from the market, that volatility shows skew and smile effect without considering T and K.

Inspired from others's work, we defined a new volatility function for the options to both consider the financial and time factors.

Remark 2.3. *Logarithm ratio of S and K is considered, because the log-ratio changes slower compare to S/K only. Especially in the trading process, same stock price can lead to many different strike prices, the log ratio will give us denser data to work with. As it is known that the option chain data is sparse. Hence a denser data structure is desired. Details can be found in Chapter 3 and 4.*

Definition 2.4. *The new generalized B-S is,*

$$dS = \mu S dt + \sigma(\ln S/K, T - t) S dW(t) \tag{2.25}$$

2.3.2 Nonparametric Kernel Method for Volatility Estimation

Our research is concerned with using RKHS for all the accessible functions of volatility, which allows us to write volatility function in terms of linear combinations of kernel functions. This is

also called the nonparametric method to define a function in statistics. Compared to the parametric approach, this method gives the flexibility of the specific assumption of the structure of the unknown function.

First, to make our life a little easier, we need to do some transformation for Dupire's equation, [20]

$$-C_T + \frac{1}{2}\sigma^2 K^2 C_{KK} - (r-q)KC_K - rC = 0 \tag{2.26}$$

Apply the logarithm ratio assumption, substitute

$$y = \ln(K/S), \tau = T - t, v(y,\tau) = C(K,T) \tag{2.27}$$

Then we get the parabolic partial differential equation,

$$-v_\tau + a(-y,\tau)v_{yy} - (r-q)v_y - rv = 0 \tag{2.28}$$

With initial condition,

$$v(y,0) = S(1 - e^y)^+, \ y \in \mathbf{R} \tag{2.29}$$

Use vector $x \in Q = R \times [0,T]$, then

$$x := (ln(S/K), T - t) =: (-y,\tau), \ \ a(x) = \frac{1}{2}\sigma(x)^2. \tag{2.30}$$

Next, we want to illustrate the rationale to use Reproducing Kernel Hilbert Space for all accessible volatility function. First, the linear combination form in terms of kernels is a desired feature for estimation of the unknown structure of a function, which is a nonparametric method in statistics.

Definition 2.5. *We assume $a \in H_\mathcal{K}$ defined on set Q in the discrete case, then for given kernel*
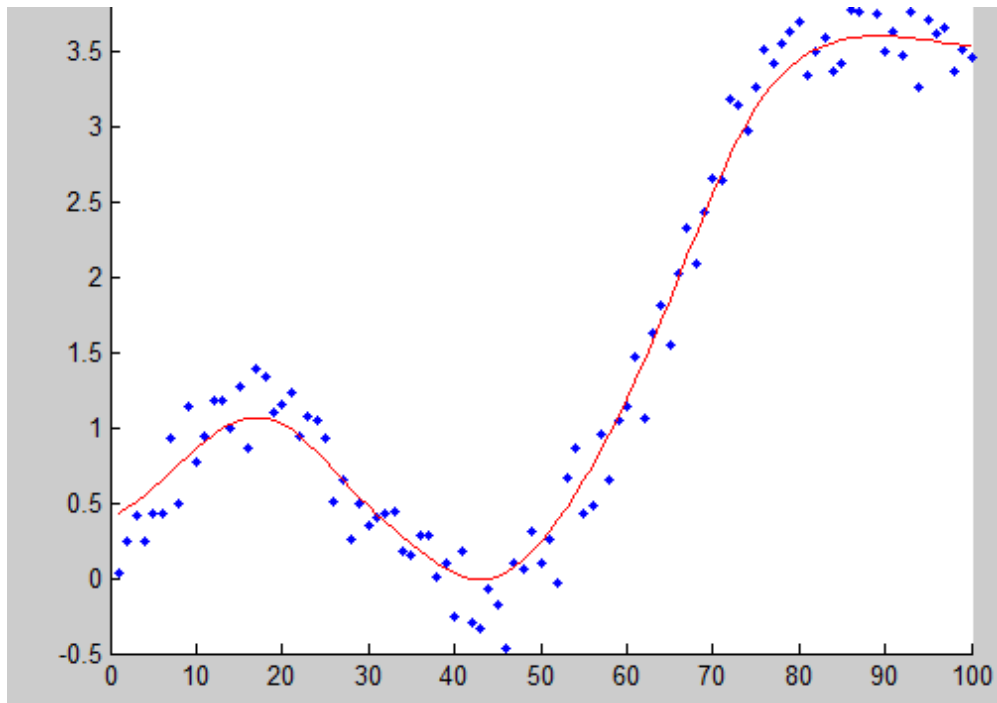
Figure 2.1: Example of kernel estimate

Source: Kernel Density Estimation
`http://www.wikipedia.com`

function $\mathcal{K}$, suitable coefficient β and $x \in Q$, $x_1 \cdots x_N \in Q$ are dense in Q,

$$a(x) = \sum_{i=1}^{N} \beta \mathcal{K}(x, x_i) \tag{2.31}$$

The kernel function smooth out the structure, the figure 1.1 can illustrate this. As we can see from the example, the nonparametric method concerns more about the natural trend of the data itself.

The nonparametric method is good in the way of following the natural structure of the data. But it could lead to jaggedness and jerks if kernel smoothing is not applied. The Figure 2.2 gives a quick illustration. Based on given data, the left side of the Figure (2.2) without use of kernel function, shows the noisy and inconsistent estimate compare to the right hand side graph with the use of

22
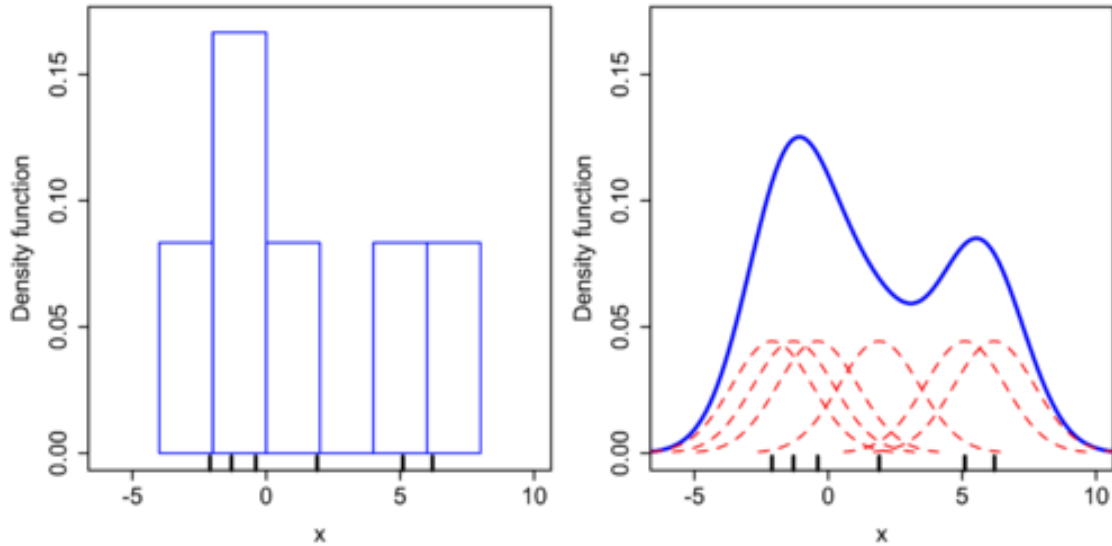
Figure 2.2: Comparison of kernel versus without kernel

Source: Kernel Density Estimation
http://www.wikipedia.com

kernel function. Especially for sparse data set, a smooth and stable result is unable to achieve.

Nonetheless, we are aiming to find a general function form of the volatility by totally embracing the data structure, but not artificially forcing the data structure used in parametric method. The kernel method helps us to achieve this goal. By considering volatility functions in the RKHS, we can get a generalize theory for this matter. A continuous form is discussed in the following. Let's have a quick review of RKHS, some other selected properties of RKHS can be found in Appendix A.

2.3.3 The Review of RKHS

Reproducing Kernel Hilbert Space firstly is a Hilbert space with pointwise evaluation functionals bounded and continuous. RKHS appears in a wide range of research areas due to the additional

structure by kernel producing features which Hilbert spaces do not usually process. The review material is from [11, 31, 57, 65, 68].

Proposition 2.6. *If we have a Hilbert space H of functions defined on a set Ω for which the point evaluation functionals $I_y : f \to f(\Omega)$ are bounded for all $\omega \in \Omega$.*

$$|I_y f| = |f(\omega)| \leq \|f\| \tag{2.32}$$

By Riesz representation theorem (can be found in any Functional Analysis book), we can find a unique function $\mathcal{K}(z, \omega)$ defined on $\Omega \times \Omega$ with the following two properties:

Proposition 2.7. *Hilbert space H of functions defined on a set Ω,*

1. *For every $\omega \in \Omega$ the function $\mathcal{K}_\omega : z \to \mathcal{K}(z, w)$ belongs to H.*

2. *For every $\omega \in \Omega$ and $f \in H_\mathcal{K}$*
$$\langle f, K_\omega \rangle_H = f(\omega)$$

The function $\mathcal{K}(z, \omega)$ is uniquely defined and is called the reproducing kernel of the space. Note that we also can get,

$$\mathcal{K}(z, \omega) = \mathcal{K}_\omega(z) = \langle \mathcal{K}_\omega, \mathcal{K}_z \rangle$$

and

$$\|I_y\|^2 = \|\mathcal{K}_\omega\|^2 = \langle \mathcal{K}_\omega, \mathcal{K}_\omega \rangle = \mathcal{K}(\omega, \omega)$$

Let us have a look at the positive definite feature of kernel function and kernel matrix.

Proposition 2.8. *A symmetric function: $\mathcal{K} : H_\mathcal{K} \times H_\mathcal{K} \to R$ is called Kernel function if it is*

positive definite Positive definiteness for kernel means, for all n,

$$\sum_{i,j=1}^{n} c_i c_j \mathcal{K}(x_i, x_j) \geq 0, \forall c_i \in R, \forall\, x_i \in H_K \tag{2.33}$$

also the Gram matrix or kernel matrix defined by $K(i, j) = \mathcal{K}(x_i, x_j)$ is positive definite.

The positive definite property says that not all functions can be kernel functions.

Under some general circumstances, Kernel matrix K has an eigenvector eigenvalue decomposition that generalizes the eigenvector eigenvalue decomposition of a postitive definite matrix

$$K = \Gamma D \Gamma' \tag{2.34}$$

With Γ and D are orthogonal, and D diagonal. See[68, 10]

Proposition 2.9. *If there exists an orthonormal sequence of continuous eigenfunctions $\phi_1, \cdots, \phi_n$ in $L_2(Q)$ and eigenvalues $\lambda_1, \cdots, \lambda_n$ for kernel matrix decomposition, then*

$$\int_Q K(s,t)\phi_i(t)dt = \lambda_i \phi_i(s) \tag{2.35}$$

$$K(s,t) = \sum_{i=1}^{n} \lambda_i \phi_i(s)\phi(t) \tag{2.36}$$

$$\int_Q \int_Q K^2(s,t)dsdt = \sum_{i=1}^{n} \lambda_i^2 < \infty \tag{2.37}$$

Lemma 2.10. *If*

$$\int_Q \int_Q K^2(s,t)dsdt < \infty \ \text{ and } \ f_i = \int f(t)\phi_i(t)dt \tag{2.38}$$

Then $f \in H_K$ if and only if

$$\sum_i \frac{f_i^2}{\lambda_i} < \infty \qquad (2.39)$$

and

$$\|f\|_K^2 = \sum_i \frac{f_i^2}{\lambda_i} \qquad (2.40)$$

2.3.4 Assumption of Volatility Function in RKHS

First we assume volatility function $a(-y, \tau)$ is bounded, i.e.

$$\underline{a} \leq a, a_0 \leq \overline{a}$$

$a_0 \in H_{\mathcal{K}}$ is a prior. We are concerning the following set for the admissible class of parameter

$$\mathcal{D}(F) := \{a \in a_0 + H_{\mathcal{K}}(Q) : \underline{a} \leq a \leq \overline{a}\} \qquad (2.41)$$

$\mathcal{D}(F)$ is assumed to be a closed and convex subset in $H_{\mathcal{K}}$. Also we suppose that the consistent data is available for the theoretical analysis purpose.

Assumption for $a(-y, \tau)$:

Volatility function is assumed to be uniformly continuous (Hölder continuous, or Hölder condition) on $Q = R \times [0, T] \in R^2$, i.e.

$$\|a(x) - a(y)\|_{H_{\mathcal{K}}(Q)} \leq \mathcal{C}\|x - y\|_{R^2}^\alpha \quad \forall x, y \in Q \qquad (2.42)$$

where there are non-negative real constants $\mathcal{C}$, if $\alpha = 1$ then it satisfies a Lipschitz condition, and α is called Hölder coefficient.

Remark 2.11. *The Hölder condition assumption is essential for the existence and uniqueness of the solution in partial differential equation (2.28) and (2.29). We will discuss this in the following Section 2.5.*

2.4 The Direct Problem and Forward Operator

In this section, we discuss the direct problem of option pricing, a.k.a as how to find the option price with given volatility. Then we present the conditions under which the direct problem has a solution. We give some general assumptions for option price function and the volatility function as the parameter appeared in the partial differential equation. Also we discuss the features of the forward operator, which is also called the parameter-to-solution operator. The assumptions for option pricing function are given in 2.4.1. With the smooth assumption for volatility function, a smooth structure for the kernel function is as well needed. Some examples of kernel functions are given that are not suitable for this scenario. Then the direct problem is proved to be well-posed with some given corollaries in 2.4.2. The properties of forward operator are discussed in 2.4.3.

2.4.1 The Direct Problem

From the analysis above that we know that we need to solve the partial differential equations (PDE) (2.28) and (2.29) to find the option price. Those logarithmic transformed equations will serve us for the following research as the primary PDE for option pricing.

Assumption for $v(y, \tau)$**:**

$Q := [0, T] \times R \in R^2$ as the set defined above. $v(y, \tau)$ satisfies (2.17) and (2.18), assume $v(y, \tau) \in W_2^{2,1}(Q)$, where $W_2^{2,1}(Q)$ is a Sobolev space, which is the space of smooth functions

27

satisfying

$$\|u\|_{W_2^{2,1}(Q)} := \|u\|_{L_2(Q)} + \|u_\tau\|_{L_2(Q)} + \|u_y\|_{L_2(Q)} + \|u_{yy}\|_{L_2(Q)} < \infty, \quad \text{for } \forall u \in W_2^{2,1}(Q) \quad (2.43)$$

Definition 2.12. *The Parameter-to-solution operator is noted as $F : \mathcal{D}(F) \to W_p^{2,1}$, where $\mathcal{D}(F)$ is defined in (2.19), it is the all admissible class of parameter a. $W_2^{2,1}$ is a a Sobolev space on given set $Q = [0, T] \times R \in R^2$. The Parameter-to-solution operator is defined as for any $a \in \mathcal{D}(F)$, $\exists v(a) \in W_2^{2,1}(Q)$, such that*

$$F : a \to v(a) \tag{2.44}$$

F is also called forward operator.

2.4.2 Existence and Uniqueness of the Solution

In [3, 4], they assume the volatility function is in Sobolev space $H^1(\Omega)$. In the direct problem of finding price for options, our particular focus concerns the existence and uniqueness of the solution of the partial differential equation in the space $W_2^{1,2}(\Omega)$ by given the parameter in RKHS $H_{\mathcal{K}}(\Omega)$. Our conclusion is based on the well known result of [64] and [68]. The partial differential equation has a unique solution for smooth coefficients. Thus the Hölder condition has to be imposed for the volatility function in RKHS to satisfy the required smoothness for the existence and uniqueness of the solution.

Proposition 2.13. *Let $H_{\mathcal{K}}(\Omega)$ be RKHS defined on Ω with kernel function $\mathcal{K}$, for any $a \in H_{\mathcal{K}}(\Omega)$, there exists $\mathcal{K}_x = a(x) \in H_{\mathcal{K}}(\Omega)$ bounded and Hölder continuous, which means*

$$\|\mathcal{K}_x - \mathcal{K}_y\|_{L_2} \leq M\|x - y\|^\alpha, \quad \alpha > 0, \quad M > 0$$

The smoothness of the functional in RKHS dependents on the choice of kernel function. Thus in order to have unique solution, a smooth kernel function is required.

Example 2.14. *Uniform kernel*

$$\mathcal{K}(u) = \frac{1}{2}\boldsymbol{I}_{|u|\leq 1} \tag{2.45}$$

or triangular kernel

$$\mathcal{K}(u) = (1 - |u|)\boldsymbol{I}_{|u|\leq 1} \tag{2.46}$$

are not good candidates.

The following lemma is from [64, Theorem IV 9.2],

Lemma 2.15. *Let a be Hölder continuous with a is bounded, $b \in L^{\infty}(Q)$ and $f \in L^{P}(Q) \cap L^{2}(Q)$. Then*

$$-v_{\tau} + av_{yy} - bv_y - rv = 0 \tag{2.47}$$

With initial condition,

$$v(y, 0) = S(1 - e^y)^+, \ y \in \mathbf{R} \tag{2.48}$$

has a unique solution $v \in W_2^{1,2}(Q)$. Moreover,

$$\|v\|_{W_2^{1,2}(Q)} \leq L\|f\|_{L_2(Q} \tag{2.49}$$

where

$$L = \max(\bar{a}, r - q)$$

Now we let a be Hölder continuous in RKHS $H_{\mathcal{K}}(Q)$ and bounded. Then by the density embedding theorems for Sobolev spaces, we have

Theorem 2.16. *Let $a \in H_{\mathcal{K}}(Q)$ bounded and Hölder continuous, $b = r - q$ is considered to be*

constant, $f \in L^p(Q) \cap Ł^2(Q)$. Then for partial differential equation (2.47) and (2.48) has a unique solution $v \in W_2^{1,2}(Q)$. Moreover, (2.49) holds for RKHS assumption.

Sketch of Proof : Given a Hölder continuous coefficient a_n, Lemma 2.15 implies that there exists a unique solution $v^n \in W_2^{1,2}(Q)$ of

$$-v_\tau^n + a_n v_{yy}^n - b v_y^n = f \tag{2.50}$$

and the initial condition is,

$$v(0, y) = 0 \tag{2.51}$$

Let $\omega^n := v_y^n$, notice that $\omega^n \in W_2^{1,0}(Q)$, then by linearity of differential operators,

$$-\omega_\tau + (a_n \omega_y^n)_y - b\omega^n = f_y \tag{2.52}$$

$$\omega^n(0, y) = 0 \tag{2.53}$$

with the estimate

$$\|\omega^n\|_{W_2^{1,0}(Q)} \leq \mathcal{C}_1 \int_0^T \|f_y(t)\|_{L_2}^2 dt \leq \mathcal{C}_2 \|f\|_{L_2(Q)} \tag{2.54}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are constants depend on the limitation of the coefficients. From the definition of $W_2^{2,1}(Q)$ norm

$$\|v^n\|_{W_2^{2,1}(Q)} \leq \mathcal{C}\|f\|_{L_2(Q)}. \tag{2.55}$$

By the density of Hölder space on Hilbert space, there exists a sequence of Hólder coefficients a_n such that $a_n \to a$ in defined RKHS norm. The weak compactness of Hilbert space, remember $H_{\mathcal{K}}$

is a Hilbert space itself, implies that there exists a subsequence v^{n_k} of v^n such that

$$v^{n_k} \to v^* \in W_2^{1,2}(Q),$$

For some $\phi \in C_0^\infty(Q)$,

$$\int_Q (-v_\tau^* + a v_{yy}^* - b v_y^*) \phi d(y, \tau) = \lim_{k \to \infty} \int_Q (-v_\tau^{n_k} + a v_{yy}^{n_k} - b v_y^{n_k}) \phi d(y, \tau) = \int_Q f \phi d(y, \tau). \quad (2.56)$$

By the weak lower semi-continuity of $W_2^{2,1}(Q)$-norm v^* satisfies (2.47). Hence $v := v^*$ is the unique weak solution to (2.47) and (2.48).

Remark 2.17. *In [4], the density embedding theorems are applied for the proof for Sobolev space $H^1(Q)$. $H_\mathcal{K}(Q)$ with smooth kernel applied here which gives a smooth parameter. Detailed proof can be found in [4, 64, 68].*

Corollary 2.18. *From the classical existence theorem of parabolic differential equations, a fundamental solution $w(y, \tau)$ for (2.28) and (2.29) satisfies,*

$$w_\tau + a w_{yy} - b w_y = \delta(\tau)\delta(y) \quad (2.57)$$

$\delta(\tau), \delta(y)$ are Dirac delta functions.

Corollary 2.19. *The fundamental solution of (2.28) and (2.29) are the Green's function, such that*

$$v(y, \tau) = \int_{-\infty}^{y_0} \frac{1}{\sqrt{4\pi\tau}} exp\{-\frac{(y - \mu)^2}{4\tau}\}(S - e^\mu)d\mu \quad (2.58)$$

31

2.4.3 The Properties of the Forward Operator

In this subsection, we can see the volatility to price operator is compact and continuous, also Fréchet differentiable. This explains the ill-posedness of the volatility identification problem for infinite dimensional space assumption.

Proposition 2.20. *The volatility to price operator defined above as $F : \mathbf{D}(F) \subset H_{\mathcal{K}} \to W_2^{2,1}(Q)$ is continuous.*

Proof. Consider volatility function sequence a_n, and $a_n \to a$ in $H_{\mathcal{K}}$. Let v_n be the corresponding solutions of PDE (2.28) and (2.29), (substitute $b = r - q$ in (2.28)), $v^n \in W_2^{2,1}(Q)$. Define $\omega^n := v_n - v(a) \in W_2^{2,1}(Q)$, by linearity, which solves the following,

$$-\omega_\tau^n + a(\omega_{yy}^n - \omega_y^n) + b\omega_y = -(a_n - a)(v_{yy}^n - v_y^n) \tag{2.59}$$

with homogeneous boundary conditions.

Then by the norm defined in Sobolev space for ω^n,

$$\|\omega^n\|_{W_2^{2,1}(Q)} \leq M \|a^n - a\|_{L_2(Q)} \|v_y^n\|_{W_2^{2,1}(Q)} \tag{2.60}$$

For some $M > 0$, where $M = max(\|a\|_{L_\infty(Q)}, |b|)$. Then by RKHS, $H_{\mathcal{K}}$, is a complete subspace of L_2 we have $a^n \to a$ also in $L_2(Q)$, with the inner product defined in L_2 for bounded kernel $\mathcal{K}(\tau, y)$ on Q, with a upper bound $M^* > 0$, i.e. $|\mathcal{K}(\tau, y)| \leq M^*$, so

$$\|a_n - a\|_{H_{\mathcal{K}(Q)}}^2 \leq M^* \|a_n - a\|_{L_2(Q)}^2 \tag{2.61}$$

then, let $\tilde{M} = max\{M, M^*\}$,

$$\|\omega^n\|_{W_2^{2,1}(Q)} \le \tilde{M}\|a^n - a\|_{L_2(Q)}\|v_y^n\|_{W_2^{2,1}(Q)} \tag{2.62}$$

Thus $\|\omega^n\|_{W_2^{2,1}(Q)} \to 0$, by $\omega^n = v_n - v(a)$, $\|v_y^n\|_{W_2^{2,1}(Q)}$ is bounded. Thus $v^n \to v(a) \in W_2^{2,1}(Q)$. Hence the volatility to price operator F is continuous. $\qquad\square$

Definition 2.21. *An operator $F : X \to Y$ is compact, if for every bounded sequence $\{x_n\} \in X$, $\{F(x_n)\}$ has a convergent subsequence.*

Compact operator can also be viewed as operator that maps a bounded set into another compact(i.e. bounded and closed) set. Compactness is assumed here, since we can use the properties which allow us to resemble the properties of operators in infinite dimensional vector spaces with finite dimensional spaces.

Proposition 2.22. *The volatility to price operator defined above as $F : \mathbf{D}(F) \subset H_{\mathcal{K}} \to W_2^{2,1}(Q)$ is compact. Moreover, weak sequential closedness is weakly continuous and the domain $D(F)$ is weakly closed in RKHS.*

Proof. Let volatility function sequence a_n and option pricing transformed function v_n, $a_n \rightharpoonup a$ in $H_{\mathcal{K}}$ topology, and $v^n \rightharpoonup v$ in $W_2^{2,1}(Q)$ topology. Define $\omega^n := v_n - v(a) \in W_2^{2,1}(Q)$, by linearity, which solves the following

$$-\omega_\tau^n + a(\omega_{yy}^n - \omega_y^n) + b\omega_y = -(a_n - a)(v_{yy}^n - v_y^n) \tag{2.63}$$

The proof similar to the continuity one. $\qquad\square$

Remark 2.23. *Weakly sequentially closed in RKHS means if a sequence $\{x_n \in H_{\mathcal{K}}(Q)\}$ weakly converges to some $x^* \in H_{\mathcal{K}}(Q)$, i.e. $x_n \rightharpoonup x^*$, if present as inner product in $H_{\mathcal{K}}(Q)$, for any*

$\tilde{x} \in H_{\mathcal{K}}(Q)$,

$$\langle x_n - x, \tilde{x} \rangle_{H_{\mathcal{K}}(Q)} \to 0, \tag{2.64}$$

and if $F(x_n) \rightharpoonup y$ for some option price function y in $W_2^{2,1}(Q)$, then $x^* \in D(F) \subset H_{\mathcal{K}}$, and $F(x^*) = y$.

Weak closeness of operator F guarantees the existence of solution to operator equation $F(a) = u$, but it cannot guarantee the uniqueness, which is one of the criterion for well-posedness of this problem.

Proposition 2.24. *Let* $F : \mathbf{D}(F) \subset H_{\mathcal{K}} \to W_2^{2,1}(Q)$ *be the volatility to price operator, and* $F(a) = u$ *satisfies (2.17) and (2.18). Then F admits a one side derivative at $a \in \mathcal{D}(F)$ in the direction h and $a + h \in \mathcal{D}h$, the derivative $F'(a)$ satisfies*

$$\|F'(a)h\|_{W_2^{2,1}(Q)} \le c\|h\|_{H_{\mathcal{K}}(Q)}, \tag{2.65}$$

Moreover, F is Fréchet differentiable and satisfies Lipschitz condition, there exists $L > 0$ such that,

$$\|F'(a) - F'(a^\dagger)\| \le L\|a - a^\dagger\| \tag{2.66}$$

for all a and h such that $a, a + h \in \mathcal{D}(F)$.

Proof. The proof follows similar ideas in [4, Proposition 4.1]. Without loss of generality, we assume $b = 0$ in (2.17). By linearity of (2.17), the direction derivative $v'h$ in the direction h satisfies,

$$-(v'h)_\tau + a((v'h)_{yy} - (v'h)_y) = -h(v_{yy} - v_y) \tag{2.67}$$

with homogeneous initial conditions. Then the solution of this Cauchy problem is $v'h \in W_2^{1,2}$. Linearity of $v'h = F'(a)h$ in h with the continuity of a implies Fréchet differentiability. $\qquad \square$

First, by applying the Sobolev embedding theorem, we can expand the Proposition 2.18 to $L_2(Q)$.

Corollary 2.25. *The volatility to price operator defined above as $F : \mathbf{D}(F) \subset H_\mathcal{K} \rightarrow L_2(Q)$ is compact. Moreover, weak sequential closedness is weakly continuous and the domain $D(F)$ is weakly closed in RKHS.*

2.5 The Inverse Problem and the Ill-posedness

Inverse problems can be put as the converse of the direct problems. It may sound still ambiguous. One may think it is random to pick up which one is *direct* and which one is *inverse*. In general, two motivations may bring researchers attention to the inverse problems: first, the past states or the parameters of certain existing system need to be adjusted or tuned . Second, one may need to find out how the parameters or the past states influence the system and usually, it is critical to give good prescribed parameters and value for the past states. The detail discussion can be seen in [32].

Definition 2.26. *Let X and Y be vector spaces, an inverse problem can be understood as: if there is an operator $\mathcal{F} : X \rightarrow Y$, given $y \in Y$, we are aiming to find $x \in X$ such that*

$$\mathcal{F}x = y \tag{2.68}$$

In applications of real world scenario, it is unlikely to know precise value for y, so we may need to allow some degree of precision, which means if $y \in Y$, let vector space Y equipped with norm $\| \cdot \|$, such that for some $\delta > 0$

$$\|y - y^\delta\| \leq \delta \tag{2.69}$$

we will see the data is within δ precision. In this sense, we see in option pricing problem, given parameter to get the option prices would be the direct problem. Conversely, given option prices

35

to get parameter value will be the inverse problem. The good part of this pair is that the direct problem is well-posed, by saying well-posed, we mean precisely

Definition 2.27. *Hadmard's definition of well-posedness/ill-posedness:*

For inverse problem $\mathcal{F}x = y$ is said to be well-posed if

1. For each $y \in Y$, there exists a $x \in X$ such that $\mathcal{F}x = y$

2. For each $y \in Y$, there is a unique $x \in X$ such that $\mathcal{F}x = y$

3. The solution x depends continuously on y.

If it is not well-posed, then we say the operator equation is ill-posed.

Theorem 2.28. *The compact, continuous and weakly closed operator, a.k.a the parameter-to-solution operator $F : \mathcal{D}(F) \subset H_{\mathcal{K}}(Q) \rightarrow W_2^{2,1}(Q)$ implies local ill-posedness of the inverse problem of option pricing calibration, $\mathcal{D}(F)$ is considered to be infinite dimensional.*

Proof. If we choose a bounded sequence $\{a_n\} \in \mathcal{D}(F)$, a_n has no convergent subsequence, but it has a weakly convergent subsequence by the boundedness, the subsequence noted as a_{n_k}. By F is a weakly closed and compact operator, so the sequence $\{F(a_{n_k})\}$ converges. It implies that similar option prices may be associated with differing volatilities. Thus, we can see the inverse problem is ill-posed. The details of the explanation can be seen in [4, 54] □

CHAPTER 3: NONLINEAR MODEL WITH REGULARIZATION IN RKHS

The contribution of this chapter is that we propose Tikhonov regularization by using RKHS. To date, it has not yet been used in the calibration problem for option pricing. Sobolev space is used in [3, 4, 19]. Usually, a smooth volatility function is assumed. By using RKHS we help attain a smooth structure of volatility. The properties of RKHS can help us deal with the sparse and noisy data. This approach is commonly used in other areas, for example, signal processing, machine learning, neural network, artificial intelligence, etc, [10, 57, 58].

We choose to regularize the ill-posed problem motivated by the ideas of [3, 4, 19]. Tikhonov regularization is one of the most often used techniques in regularization schemes. In general, it can be viewed as a trade-off between precision and stability, wherein a small compromise in accuracy will be rewarded with an increased stability and uniqueness of the solution. Overall it is a popular approach to solve ill-posed inverse problems. Tikhonov regularization is well understood and applied for the linear ill-posed problems by now; however, the nonlinear cases are still not very well developed. Usually, some restrictive assumptions are needed for nonlinear equations.

In this chapter, we consider the continuous case for option pricing function $C(S, t; T, K, \sigma, q)$ for theoretical analysis. We analyze the convergence of the solution as the parameter of regularization and noisy level are controlled to approach zero. Also the existence and uniqueness are discussed while the stability of the solution is proven for continuous data accessible for option prices case.

This chapter is structured as:

In the first section, we have a short review of the general regularization theory, which is the foundation of regularization.

3.1 Tikhonov Regularization

Regularization is a popular way of controlling ill-posedness of operator equations in a flexible and intuitive manner. We have a quick review of regularization technique in Appendix B.

Definition 3.1. *In general, given $\mathcal{F} : X \to Y$, X and Y are any two topology spaces, a regularization strategy is a family of bounded operators*

$$R_\alpha : Y \to X, \quad \alpha > 0, \tag{3.1}$$

such that

$$\lim_{\alpha \to 0} R_\alpha \mathcal{F} x = x \quad for\ all\ x \in X. \tag{3.2}$$

The regularization operators R_α usually are a sequence of approximation of the inverse of $\mathcal{F}$. $R_\alpha \mathcal{F}$ is point-wisely convergent to identity.

3.2 Nonlinear Tikhonov Regularization for Calibration Problem

3.2.1. The Nonlinear Operator Equation

In more general cases, we have nonlinear operator equations, i. e. operator $\mathcal{F}$ is not linear. We can assume that

(i) $\mathcal{F}$ is a continuous operator from X to Y in the defined norm $\| \cdot \|$

(ii) $\mathcal{F}$ is also sequentially closed, which means for any sequence $\{x_n\} \subset D(F)$, $x_n \rightharpoonup x$, and $\mathcal{F}(x_n) \rightharpoonup y$ for $y \in Y$, then $x \in X$ and $\mathcal{F}(x) = y$.

Notice the above assumption is valid in varies application, especially, the calibration for option

pricing problems. [3,4,56,60] Sequentially closed is satisfied if $\mathcal{F}$ is continuous and compact with the domain of $\mathcal{F}$, $D(\mathcal{F})$ is weakly closed, i.e. closed and convex. The inverse problems for nonlinear compact operator, we also have the counterpart for the ill-posedness. Next, let us have a look at the difference between linear and nonlinear operator equations for inverse problem.

3.2.2. The Comparison with the Linear Case

We follow some standard results about the linear and nonlinear comparison, see [12,32,45]. It is necessary because we apply both of the scenario in our application. Consider the inverse problem

$$\mathcal{F}x = y, \ \ \mathcal{F} : X \to Y$$

In the linear case, $\mathcal{F}$ is a linear operator, on the contrary, $\mathcal{F}$ is nonlinear for the nonlinear case.

If the inverse problem is ill-posed, which essentially means the unboundedness of the inverse of $\mathcal{F}$ if it exists, or the generalized inverse in general cases $\mathcal{F}^{\dagger}$,

$$\mathcal{F}^{\dagger} \ \ unbounded \Leftrightarrow R(F) \ \ not \ closed \tag{3.3}$$

Here $R(F) \subset Y$ denotes the range of operator F. Also if F is compact,

$$\mathcal{F}^{\dagger} \ \ unbounded \ \ \Leftrightarrow dim R(F) = \infty \tag{3.4}$$

For linear operator regularization in RKHS, the solution exist from the famous Representer theorem, see [65]. Next we can see a sufficient condition for ill-posedness of the nonlinear inverse problem for compact operator $\mathcal{F}$.

3.3 The Convergence Analysis for Nonlinear Tikhonov Regularization for RKHS

If $\mathcal{F}$ is nonlinear, in consequence, the solution may not be unique in general.

Theorem 3.2. *The regularized solution depends continuously on given data in the nonlinear Tikhonov regularization for ill-posed operator equation $\mathcal{F}x = y$*

Proof. If a data sequence y_k in Y, $y_k \to y^\delta$ then minimizer sequence (x_k) has convergent subsequence also noted as x_k, and if $x_k \to \bar{x}$, then $\bar{x}$ is a minimizer, too. By continuous dependence on data, we mean for any small $\epsilon > 0$, there exists $\delta > 0$ by the continuity of $\mathcal{F}$ in X, such that as $\|y_k - y^\delta\| \le \delta$,

$$\|x_k - \bar{x}\| \le \epsilon$$

from assumption that (x_k) are the minimum of J_α so,

$$\|\mathcal{F}(x_k) - y_k\|^2 + \alpha\|x_k - x_0\|^2 \le \|\mathcal{F}(x) - y_k\|^2 + \alpha\|x - x_0\|^2, \quad x \in X$$

since $\|y_k - y^\delta\| \le \tau$, for any ϵ, we can find $k > N$, $N > 0$ a big number such that

$$\|\mathcal{F}(x_k) - c_k\|^2 + \alpha\|x_k - x_0\|^2 \le \|\mathcal{F}(\bar{x}) - c_\delta\|^2 + \alpha\|x_k - \bar{x}\|^2 + \alpha\|\bar{x} - x_0\|^2 \le M$$

$\square$

Compare to the results from Engl and Zou [32] and Engl and Egger [4], we can get the convergence rate of the regularization in RKHS $H_{\mathcal{K}(Q)}$,

Theorem 3.3. *If $\mathcal{D}(F) \in H_{\mathcal{K}}(Q)$ is Convex, let $a^\dagger$ be the true solution and $v = v(a)$. a_λ^δ is a minimizer of the regularization for a_0 prior, and if volatility to price operator satisfies the following*

conditions:

(i) F is Fréchet differentiable. F' denotes the derivative

(ii) there exists $L > 0$ with $\|F'(a^\dagger) - F'(a^*)\| \le L\|a^\dagger - a_0\|$, for all $a^* \in \mathcal{D}(F)$.

(iii) there exists $\omega \in W_2^{2,1}(Q)$, with $a^\dagger - a_0 = F'(x^\dagger)\omega$.

(iv) $L\|\omega\| < 1$.

Then, for $\alpha \sim \delta$,

$$\|a_\alpha^\delta - a^\dagger\| = O(\sqrt{\delta}). \tag{3.5}$$

Proof. First note that, in the RKHS case,

$$\|a^\delta - a^\dagger\|_{H_{\mathcal{K}(Q)}}^2 = \|a^\dagger - a_0\|_{H_{\mathcal{K}(Q)}}^2 + \|a^\dagger - a^\delta\|_{H_{\mathcal{K}(Q)}}^2 + 2\langle a^\dagger - a_0, a^\dagger - a^\delta\rangle_{H_{\mathcal{K}(Q)}} \tag{3.6}$$

the triangular inequality implies

$$\|v^\delta - v\|^2 + \alpha\|a^\delta - a^\dagger\|_{H_{\mathcal{K}(Q)}}^2 \le \delta^2 + 2\alpha\langle a^\dagger - a_0, a^\dagger - a^\delta\rangle_{H_{\mathcal{K}(Q)}} \tag{3.7}$$

and

$$\alpha\langle a^\dagger - a_0, a^\dagger - a^\delta\rangle_{H_{\mathcal{K}(Q)}} = \alpha\int_Q [(a^\dagger - a^\delta)(v_{yy} - v_y)]\phi d(y,\tau) \tag{3.8}$$

$$= \alpha\int_Q [-(v^\delta - v)_t + a^\delta(v^\delta - v)_{yy} - a^\delta(v^\delta - v)_y]\phi d(y,\tau) \tag{3.9}$$

$$= \alpha\int_Q (v^\delta - v)[\phi_t + (a^\delta\phi)_{yy} + (a^\delta\phi)_y]d(y,\tau) \tag{3.10}$$

$$\le \epsilon\|v - v^\delta\|_L^2(Q) + \frac{\alpha^2}{4\epsilon}(1 + 2\|a^\delta\|_{H_{\mathcal{K}(Q)}}) \tag{3.11}$$

When move the first term to the left, the proof is completed. □

41

CHAPTER 4: NONLINEAR MODEL WITH MONTE CARLO SIMULATIONS

In this Chapter, we discuss the discrete model of the nonlinear regularization model we discussed in Chapter 2. Gradient Descent is used to find the volatility surface. Monte Carlo Simulations are a very commonly used method in Finance. See [66]. We use them to generate a heuristic future stock price, then use the payoff formula to calculate the option price. With the simulated results, we apply the gradient descent to find the volatility function.

4.1 Moment Discretization

4.1.1 Regularization for Discrete Model

Tikhonov regularization is the most commonly used method of regularization of ill-posed problems. In this section, we are going to talk about the discrete model with this regularization technique. It is known as a technique that allows bias in-trade of a stable and well-posed solution.

Problem 3

The discrete case of Tikhonov Regularization for calibration problem for option pricing is

$$\min_{x \in X_m} \left\{ \|C_m(a) - y_m^\delta\|_{l^2(Q)}^2 + \lambda \|a - a_0\|_{H_\mathcal{K}(Q)}^2 \right\} \tag{4.1}$$

- $C_m(a)$ is the option model price,

- λ Tikhonov regularization parameter,

- y_m^δ is the observed market price with spread δ,

- a is volatility function as defined in (2.30), a_0 is a good guess, for example a good historical volatility,

- $\|f_i\|_{l^2(Q)} = \sum_i^{\infty} (f_i)^2, \ \ \forall f_i \in l^2(Q)$

4.1.2 Cross Validation for λ

Lambda as a parameter appeared in the regularization functional chosen by using cross validation method. Cross-validation, also known as Generalized cross validation (GCV), is a model validation technique for assessing how are the results of a statistical analysis for a given parameter. The process of Leave-1-out cross-validation is used in this paper. It involves using 1 observation as the validation for the parameter and the the other observations as the training set. In the regularization case, we need to find a suitable λ, which is the parameter used in the regularization.

The same technique can also be used to find the optimal parameter in kernel functions.

4.1.3 Gradient Descent Method

The gradient descent method, a.k.a the steepest descent, says multivariate differentiable functions decrease fastest if one goes in the direction of the negative gradient. The directional derivative along direction $[\eta]$ is followed the idea given by [4], the partial differential equation used here is slightly different from Engl and Egger [4],

$$J'(a)_{[\eta]} = 2(v(a) - v_0, v'(a)_{[\eta]})_v + 2\beta(a - a*, \eta)_a \tag{4.2}$$

43

If we let $V := v(a) - v^\delta$, $\omega := v'(a)_{[\eta]}$,

$$(v'(a)_{[\eta]}, v(a) - v^\delta) = \int \omega(y, T) r(y) dy = \int_0^T \frac{d}{dt} \int \omega(y, t) V(y, t) dy dt \qquad (4.3)$$

$$= \int_0^T \omega_t V - \omega V_t dy dt = \int_0^T \int [a(\omega_{yy} - \omega_y) - \eta(v_{yy} - v_y)]V + \omega V_t dy dt \qquad (4.4)$$

$$= \int_0^T \int [V_t + (aV)_{yy} + (bV)_y]\omega - \eta(v_{yy} - v_y)V dy dt = - \int_0^T \int \eta(v_{yy} - v_y)V dy dt \qquad (4.5)$$

V is the solution of the homogeneous pde of (2.28)

$$V_t + aV_{yy} + bV_y = 0 \qquad (4.6)$$

with condition $V(y, t) = 0$.

Then a gradient direction can be determined from

$$(g, \eta)_a = 2(v(a) - v^\delta, v'(a)_{[\eta]})_v + 2\lambda(a - a_0, \eta)_a = 2 \int_0^T \int (v_{yy} - v_y) V dy dt + 2\lambda(a - a_0, \eta)_a \qquad (4.7)$$

$(\cdot, \cdot)_a$ is the norm in H_K, (4.10) is then to solve,

$$-g_{yy} + g = -2(v_{yy} - v_y)V - 2\lambda(-da_{xx} + da) \qquad (4.8)$$

where $da = a - a_0$. The gradient direction can be derived from the numerical solutions of the parabolic partial differential equations and the solution of the variational problem (4.9).

4.2 Monte-Carlo Simulations for Options

Monte Carlo is a famous casino in Monaco and the name of this method comes from the idea of luck with the draw. Applications include: financial products pricing, integration approximation, Network reliability. etc.

Random sampling is an essential component of the Monte Carlo method and it applies well to the deterministic problem. This ability to provide solutions to both probabilistic and deterministic problems makes the Monte Carlo method more versatile than most other numerical approximation techniques.

4.2.1 The Generated Random Process

First, we need to use Monte-Carlo Simulation to generate the random variable of the lognormal process, which is assumed to be the future stock price movement path according to the B-S equation assumption. The steps are:

1. get standard norm distribution random number $z_i = N(0, 1)$,

2. generate an normal random number $n_i = N(\mu, v) = \mu + vz_i$

3. generate lognormal random number $x_i = e^{n_i}$

By using the generated random lognormal numbers, we can get the future stock price, $S_T = S_0 x_i$. Then apply for the payoff function for options for call option,

$$\text{payoff}_{call} = max(S_T - K, 0)$$
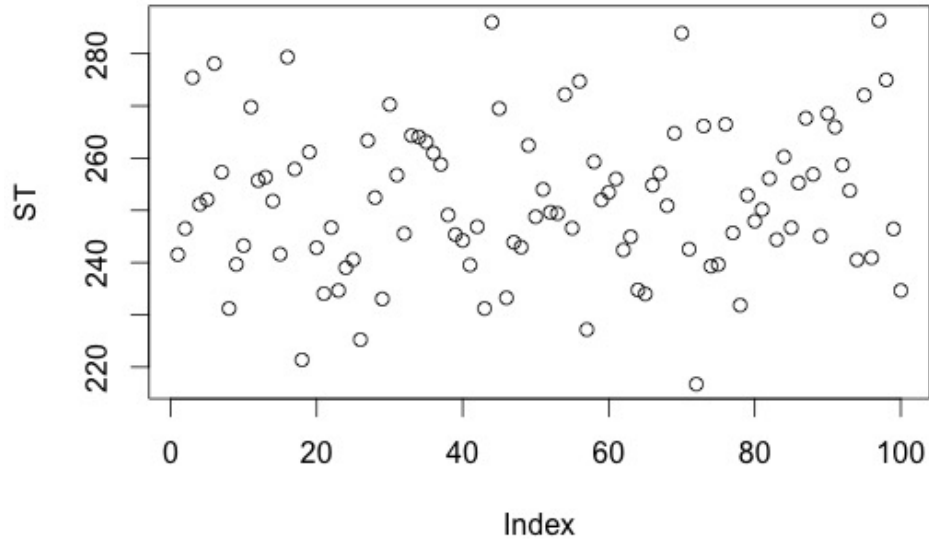
Figure 4.1: Plot of simulated Stock Prices

Then the option price at $t = 0$ is

$$C = e^{-rT}\text{payoff}_{call}$$

In this way, we can get heuristic option price data. The following table (3.1) is part of the generated result for $T - t = 0.024$, $K = 245$, $S_0 = 250$, $r = 0.03$, $q = 0.02$, we chose $\sigma = 0.012$ from the Apple. Inc Option Chain data as a starting point.

4.2.2 Control Variate and Minimum Variance

The control variate method is a common method used in Monte Carlo simulations in order to increase the accuracy of the simulation. The variance can be reduced by introducing a known result, the true value. It is especially useful in the sparse option data case. We use the market true
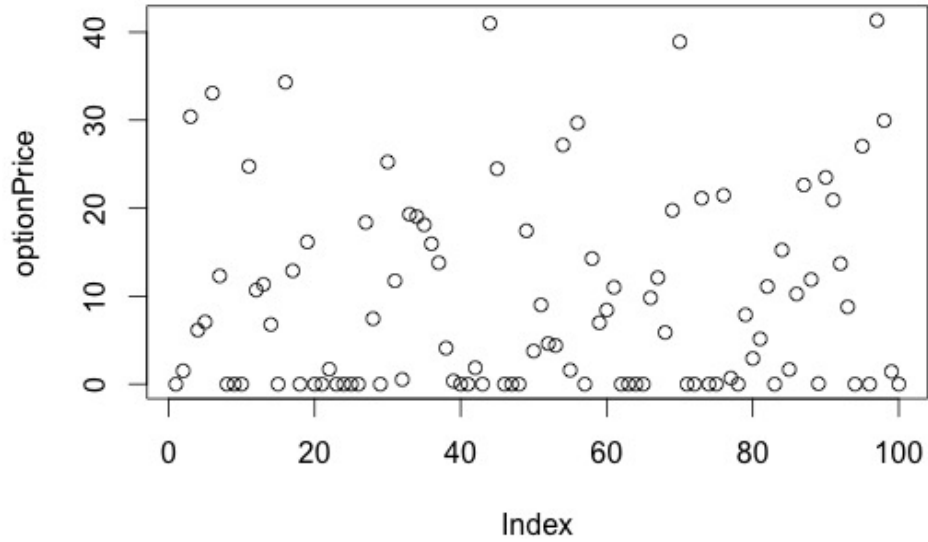
Figure 4.2: Plot of simulated Option Prices

data as a *control variate* to our simulated data. The control variate estimated is then defined as

$$C^* = C_S + \beta(V_t - V_S) \tag{4.9}$$

The reduction in variance is

$$Var(C_s) - Var(C^*) \tag{4.10}$$

Ideally, the control variate will be highly correlated with the option being priced. The following results followed from [64](Mcdonald-derivatives markets).

Theorem 4.1. *The variance* $Var(C^*)$ *is minimized when*

$$\beta = \frac{Cov(C_s, V_s)}{Var(V_s)} \tag{4.11}$$

47

Corollary 4.2. *Let $\{x_1, \cdots, x_n\}$ be the observations for the distribution of a random variable X, then the sample mean is*

$$\bar{X} = \frac{x_1 + \cdots + x_n}{n} \tag{4.12}$$

The sample variance is

$$Var(X) = E(X^2) - (E(X))^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n - 1} \tag{4.13}$$

If $\{y_1, \cdots, y_n\}$ is another set of observations of a random variable Y, then the covariance of X and Y is given by

$$Cov(X, Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{n - 1} \tag{4.14}$$

Thus, in control variate estimation $C^* = C_S + \beta(V_t - V_S)$, The variance $Var(C^*)$ is minimized when

$$\beta = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{X})^2} \tag{4.15}$$

Corollary 4.3. *The minimal variance is*

$$Var(C^*) = Var(C_S) - \frac{Cov^2(C_s, V_s)}{Var(V_s)} \tag{4.16}$$

The Variance reduction factor is

$$\frac{Var(C*)}{Var(C_s)} = 1 - \rho^2_{C_S, V_s} \tag{4.17}$$

where ρ_{C_S, V_s} is the correlation coefficient between C_s and V_s.

After applied Control variate by given true data, $N = 100$, V_t=observed prices. $\zeta = -0.0033636$
The minimal variance is $Var(C^*) = 2997.5436$ from 3031.6975

Table 4.1: Comparison of control variate and naive Monte Carlo

Controlled	Naive Monte Carlo
257.1046	257.4738
267.5082	267.8606
327.6275	327.9631
279.1841	277.5029

Remark 4.4. *A key question in error analysis for Monte Carlo valuations is one of determining the number of simulated stock price points required in order to realize the desired level of accuracy. The error reduces at the rate of $1/\sqrt{N}$. Details can be found in [8].*

4.2.3 Implementation with Gaussian Kernel for Call Options

The Gaussian radial basis function kernel is a popular kernel function used in various kernelized learning algorithms. The Gaussian kernel is given by

$$k(x_i, x_j) = exp(\frac{-\|x_i - x_j\|^2}{h^2}) \tag{4.18}$$

where h is the kernel bandwidth, $\|x_i - x_j\|$ is the Euclidean distance.

Algorithms

1. Use Naive Monte Carlo to generate simulated option prices,

2. Apply the AAPL true data to control variance, find the minimal variance coefficients, then calculate the new simulated prices,

3. Implement the regularization theory with $\lambda = 1$,

4. Use GCV to find the optimal parameter λ,

5. Find the directional gradient according to (4.10),

6. Implement Gradient descent method find the minimizer of the regularization functional,

7. Update the coefficients of the kernel function representation for volatility.

4.3 Implementation with Put Option and Other Kernel Choices

This section is the result of the implementation of European put options. Compared to call options, the payoff of put options with strike K for stock S at maturity T is

$$\text{payoff}_{put} = \max(K - S_T, 0) \tag{4.19}$$

4.3.1 European Put Option with Gaussian Kernel

First we use Monte-Carlo method to generate a new put option data. Table is omitted. Applying Gaussian kernel with parameter, the bandwith, $h = 1$.

The optimal bandwith in Gaussian kernel can be found by using cross validation, similar to the regularization.

4.3.2 Other kernels application

- Polynomial Kernel Implementation with degree 2 and 3

- Laplace kernel

Polynomial kernel of degree 2 is applied to the call and put data. The polynomial kernel is defined

as,

$$K(x, x') = (\langle x, x' \rangle + 1)^2 \tag{4.20}$$

With degree 3, the polynomial kernel is,

$$K(x, x') = (\langle x, x' \rangle + 1)^3 \tag{4.21}$$

Notice here the scale and offset of the polynomial kernel are both 1 by choice. Laplace is another general purpose kernel along with Gaussian kernel and polynomial kernel. It is usually called Laplace radial basis kernel in machine learning, the kernel function is defined as

$$K(x, x') = exp(-\sigma \|x - x'\|) \tag{4.22}$$

CHAPTER 5: LINEAR MODEL WITH GREEKS

In this chapter, we analyze the discrete case for operation equations. Comparing with the continuous case, we use the real time market data on Apple. Inc stock options. A linear model is obtained from the Black-Scholes equation in the Greeks form. The greeks are the estimations of changing rate of options and stocks, which are the derivatives in the B-S equation. They are imperative tools for analysis of option pricing. Also greeks are valuable information that are given in everyday trading data, thus we believe that a model embracing the greeks is desirable for the quick estimations.

5.1 The Option Greeks

5.1.1 Option Greeks Definition

"Trading options without an understanding of the Greeks - the essential risk measures and profit/loss guide in options strategies - is synonymous to flying a plane without the ability to read instruments. " –John Summer, Ph.D.

In finance, the Greeks used for options are some ratios defined to represent the sensitivities of prices of financial instrumentals to a change in price of the underlying asset. In practice, option values change on several different factors: variations in the stock price of the underlying asset, changing of time, movements in the interest rate or volatility or the dividend yield.

Greeks as it is used in other context, include delta, gamma, theta, rho and vega. Notice vega is not a normal greek letter, it is only used in the option scenario. We only use delta, gamma and theta according to the Black-Scholes equation, vega and rho are not considered in this dissertation due

to the assumption that r and q are constant in our model. By adopting the definition in [65], the greeks are defined as,

- Delta (Δ) is defined as the rate of change of the option price with respect to the underlying stock price, which is the first derivative of option price with respect to stock price

$$\frac{\partial C}{\partial S} = \Delta \tag{5.1}$$

- Gamma(Γ) is ratio of the change in delta and stock price, which is essentially the second derivative,

$$\frac{\partial^2 C}{\partial S^2} = \Gamma \tag{5.2}$$

- Theta(Θ) is the ratio of the change of option price with respect to time, which gives the rate of the change of the option price with respect to a decrease in the time to maturity of 1 day.

$$\frac{\partial C}{\partial t} = \Theta \tag{5.3}$$

5.1.2 Black-Scholes with Greeks

Let $C(S, t; K, T, \sigma, q)$ be B-S option price function, r is interest rate, we use it as a constant, $r = 0.03$, q is stock dividend, used as a constant, too. $q = 0.023$ according to the report of Apple Stock in November 2013. (dividend data from www. streetInsider.com). Then one can rewrite the B-S equation by using Greeks,

$$\frac{\partial C}{\partial t} + (r - q)S\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2} = rC \tag{5.4}$$

Use the greeks, $\frac{\partial C}{\partial t} = \Theta$, $\frac{\partial C}{\partial S} = \Delta$, $\frac{\partial^2 C}{\partial S^2} = \Gamma$, we have

$$\Theta + (r - q)S\Delta + \frac{1}{2}\sigma^2 S^2 \Gamma = rC \tag{5.5}$$

Proposition 5.1. *Delta is defined as $\frac{\partial C}{\partial S}$, then we have*

$$\Delta_{call} = e^{-q(T-t)}N(d_1) \tag{5.6}$$

$$\Delta_{put} = e^{-q(T-t)}N(-d_1) \tag{5.7}$$

$0 < \Delta_{call} < 1$ *and* $-1 < \Delta_{put} < 0$.

This can be derived from the B-S formula.

Proposition 5.2. *Gamma as the second derivative of option price with respect to stock price can be calculated as*

$$\Gamma_{call} = \frac{e^{-q(T-t)}}{S\sigma\sqrt{T-t}}N'(d_1) \tag{5.8}$$

and

$$\Gamma_{call} = \Gamma_{put} \tag{5.9}$$

where

$$N'(d_1) = \frac{1}{\sqrt{2\pi}}e^{-d_1^2/2} \tag{5.10}$$

Proposition 5.3. *We have*

$$\Theta_{call} = Sqe^{-q(T-t)}N(d_1) - rKe^{-r(T-t)}N(d_2) - \frac{Se^{-q(T-t)}\sigma}{2\sqrt{T-t}}N'(d_1) \tag{5.11}$$

By put-call parity, one can also get

$$\Theta_{put} = rKe^{-r(T-t)}N(-d_2) - Sqe^{-q(T-t)}N(-d_1) - \frac{Se^{-q(T-t)}\sigma}{2\sqrt{T-t}}N'(d_1) \tag{5.12}$$

5.1.3 The Greeks in Option data

In option chain data, option Greeks are estimated from data using finite difference method. They are estimated by the following formulas derived from binomial trees,

- Delta can be calculated as

$$\Delta = e^{-qh}\frac{C_u - C_d}{S_u - S_d} \tag{5.13}$$

- Gamma measures the change in delta, hence

$$\Gamma = \frac{\Delta_{uu} - \Delta_{ud}}{S_u - S_d} \tag{5.14}$$

Where

$$\Delta_{uu} = e^{-qh}\frac{C_{uu} - C_{ud}}{S_{uu} - S_{ud}} \tag{5.15}$$

$$\Delta_{ud} = e^{-qh}\frac{C_{ud} - C_{dd}}{S_{ud} - S_{dd}} \tag{5.16}$$

- Theta can be given as the rate of change of option price with respect to time h,

$$\Theta = \frac{C_{ud} - (C_0 + \Delta\epsilon + 0.5\Gamma\epsilon^2)}{2h} \tag{5.17}$$

where

$$\epsilon = S_{ud} - S_0, \quad S_0 \text{is the stock price at } t = 0 \tag{5.18}$$

55

	1	2	3	4	5	6	7	8	9	10	11	12
OpSym	Bid (pts)	Ask (pts)	Extrinsic Bid/Ask (pts)	IV Bid/Ask (%)	Delta Bid/Ask (%)	Gamma Bid/Ask (%)	Vega Bid/Ask (pts/% IV)	Theta Bid/Ask (pts/day)	Volume	Open Interest	Strike
IBM MAR10 110 C	16.25	16.70	0.00 0.37	19.77 35.15	99.16 92.06	0.27 1.15	0.007 0.053	0.0009 -0.0279	0	479	110.000
IBM MAR10 115 C	11.65	11.80	0.32 0.47	25.37 27.68	90.52 88.67	1.82 1.90	0.060 0.069	-0.0227 -0.0290	47	552	115.000
IBM MAR10 120 C	7.15	7.30	0.82 0.97	21.85 23.30	79.89 78.51	3.53 3.45	0.101 0.105	-0.0344 -0.0385	360	1179	120.000
IBM MAR10 125 C	3.40	3.50	2.07 2.17	19.04 19.75	58.20 57.98	5.65 5.46	0.141 0.141	-0.0431 -0.0448	1268	5782	125.000
Stock											126.33
IBM MAR10 130 C	1.10	1.14	1.10 1.14	17.41 17.73	28.66 29.04	5.40 5.33	0.123 0.124	-0.0349 -0.0358	1868	5947	130.000
IBM MAR10 135 C	0.23	0.25	0.23 0.25	16.73 17.08	8.45 8.91	2.56 2.61	0.056 0.058	-0.0154 -0.0164	666	6539	135.000
IBM MAR10 140 C	0.04	0.06	0.04 0.06	17.04 18.12	1.82 2.47	0.72 0.88	0.016 0.021	-0.0045 -0.0062	80	4284	140.000
IBM MAR10 145 C	0.00	0.03	0.00 0.03	0.00 21.03	0.00 1.17	0.00 0.40	0.000 0.011	0.0000 -0.0038	10	1747	145.000

Figure 5.1: Example of option data with greeks

Source: Option Chain Data
`http:www.inverstopedia.com`

By including the Greeks in Black-Scholes equation, we will use the `option chain data` downloaded from `historicaloptiondata.com`. The numerical result of this paper is from the downloaded real world data. The greeks are available in the data sheet of option chain, which can be used directly as a tool for our research. Also, greeks are available for almost all the option trading software packages, which makes greeks ready to use. (see example table Figure 4.1)

5. 2 The Linear Model and the Solution

In this section, we will show how the Greeks make the model linear. Following the linear model, we defined a new linear regularization problem.

56

5.2.1 The Linear Model

With the help of Greeks, option price appears as a linear relation with volatility.

If we let

$$\frac{1}{r}(\Theta + (r-q)S\Delta) = \gamma_0, \quad \frac{1}{r}S^2\Gamma = \gamma$$

Then the option price function $f(x_i)$ can be written as,

$$f(x_i) = \gamma_0 + \gamma a(x_i) \tag{5.19}$$

γ and γ_0 are calculated from the Greeks, which gives us a nice linear relationship between option price function and volatility function.

5.2.2 Tikhonov Regularization

Calibration from the real world option data is used to estimate the important parameter in the Black-Scholes equation, the volatility. Tikhonov regularization is also used to for the linear case to cure the ill-posedness of the inverse problem as discussed in chapter 1.

$$\mathcal{J}(a) = \|C(S, t, T, K, \sigma, r, q) - c^\delta\|^2 + \lambda\|a\|^2 \tag{5.20}$$

In this equation,

- $\mathcal{J}(a)$ is Tikhonov regularization functional for linear model.

- $C(S, t, T, K, \sigma, r, q)$ is the B-S option price.

- c^δ is the observed market price for options.

- $a(\ln(S/K), T - t) = \frac{1}{2}\sigma^2$ is the volatility function defined as before

We are aiming to minimize the total cost by applying $a(\ln(S/K), T - t)$ in the reproducing kernel Hilbert Space $H_{\mathcal{K}}$

$$a(x, x_i) = \sum_{i=1}^{N} \beta_i K(x, x_i) \tag{5.21}$$

With $x_i = (ln(S_i/K_i); T_i - t_i)$ is the input, $K(\cdot, \cdot)$ is the kernel function in $H_{\mathcal{K}}$.

Remark 5.4. *Notice, if we solve C in equation (4.5), and we have*

$$f(x_i) := C = \frac{1}{r}(\Theta_i + (r - q)S_i\Delta_i + a(x_i)S_i^2\Gamma_i) \tag{5.22}$$

Problem 2 The linear model for option pricing calibration by using greeks in RKHS $H_{\mathcal{K}}$ is defined as

$$argmin_a \sum_{i=1}^{N} (y_i - f(x_i))^2 + \lambda\beta^T K\beta \tag{5.23}$$

where $y_i := c_\delta$ is the observed option prices.

5.2.3 The Representer Theorem

In statistics, the well-known Representer Theorem introduced by Kimeldorf and Wahba is applied for minimizing functional [10]

$$J(f) = \frac{1}{n}\sum_{i=1}^{n}(f(x_i) - y_i)^2 + \lambda\|f\|_{H_{\mathbb{K}}}^2 \tag{5.24}$$

the minimizer can be written as

$$f = \sum_{i=1}^{n} C_i f(x_i) \tag{5.25}$$

Applying the Representer Theorem, we have justify our usage of $a(x)$ in $H_\mathcal{K}$.

Theorem 5.5. *In **Problem 2**, the minimizer $a(x)$ can be written as*

$$a(x, x_i) = \sum_{i=1}^{N} \beta_i K(x, x_i) \tag{5.26}$$

With $x_i = (ln(S_i/K_i); T_i - t_i)$ is the input, $K(\cdot, \cdot)$ is the kernel function in $H_\mathcal{K}$.

5.2.4 The Solution of the Linear Model

Theorem 5.6. *The solution to **Problem 2** is calculated as*

$$\beta = (\gamma^2 K + \lambda I)^{-1}(\gamma(Y - \gamma_0)) \tag{5.27}$$

Proof. In **Problem 2**, we are trying to minimize

$$\mathcal{J}(a) = \sum_{i=1}^{N}(y_i - f(x_i))^2 + \lambda \beta^T K \beta \tag{5.28}$$

$$=: (Y - f(x))^T(Y - f(x)) + \lambda \beta^T K \beta \tag{5.29}$$

Where $Y = (y_1, \cdots, y_N)$. Differentiate $\mathcal{J}$ with respect to β, we can get the coefficient of the kernel function as desired. Apply $f(x) = \frac{1}{r}(\Theta + (r - q)S\Delta + a(x_i)S^2\Gamma)$ in (4.22), then the first part in (4.29) is

$$(Y - \frac{1}{r}(\Theta + (r - q)S\Delta + K\beta S^2\Gamma))^T(Y - \frac{1}{r}(\Theta + (r - q)S\Delta + K\beta S^2\Gamma)) \tag{5.30}$$

Let $\frac{1}{r}(\Theta + S\Delta^T(r - q)) = \gamma_0$ and $\frac{1}{r}S^2\Gamma = \gamma$.

59

Table 5.1: Example of AAPL option chain data

Stock	T-t	Strike	Option	Delta	Gamma	Theta
520.03	0.024	360	159.95	0.8152	0.0004	-12.7165
520.03	0.024	365	154.95	0.8061	0.0005	-11.9357
520.03	0.024	370	149.95	0.798	0.0006	-11.0972
520.03	0.024	375	144.95	0.7912	0.0006	-10.2167

Then

$$(Y - \gamma_0 - \gamma K\beta)^T (Y - \gamma_0 - K\beta) + \lambda \beta^T K\beta \qquad (5.31)$$

Take derivatives with respect to β and set it equal 0

$$2(Y - \gamma_0 - \gamma K\beta)^T (-\gamma K) + 2\lambda \beta^T K = 0 \qquad (5.32)$$

Solve for β, the coefficient for volatility with kernel $\mathcal{K}$,

$$\beta = (\gamma^2 K + \lambda I)^{-1}(\gamma(Y - \gamma_0)) \qquad (5.33)$$

$\square$

5.3 Numerical Implementation

Apple stock data on November 1st, 2013 is downloaded in historicaloptiondata.com. The index symbol for apple stock is AAPL. The interest rate used in our implementation is 3%, apple stock in November 2013 had dividend rate as 2.3% according to streetinsider.com. Option has bid and ask prices, for calculation purpose we use the average of the two to be the option price, which is a common way to use the data from other researchers.

Table 5.2: Result of linear model for call option with Gaussian Kernel

Stock	T-t	Strike	Option	beta	volatility
520.03	0.024	360	159.95	-2.5127	0.01439
520.03	0.024	365	154.95	-2.40539	0.01209
520.03	0.024	370	149.95	-2.28382	0.01005
520.03	0.024	375	144.95	-2.14995	0.00823
520.03	0.024	380	139.95	-2.0050	0.00663

5.3.1 Numerical Implementation with Gaussian kernel

In order to calculate for our purpose, cross sections data are used to serve different variable values used in Black-Scholes equation. The selected data table is 1168 by 8, with 1168 entries trading information for European Vanilla options. Call options and put options are tested separately. The comparison results are given in the following context. Gaussian kernel is given in (3.3). R programming language is used for the numerical implementation. Here are the steps:

1. Prepare Input data $x_i = (ln(S_i/K_i), T - t)$, $y = \gamma_0 + \gamma a(x)$

2. Get the kernel matrix, use Gaussian kernel.

3. Calculate the inflated matrix $M = (\gamma^2 K + \lambda I)^{-1}$.

4. Output β and volatility surface.

5.3.2 The Result of the Gaussian Kernel Implementation

In the Table 4.2, only partial results are displayed.

Figure 5.2 shows the a stable and unique solution of volatility function.
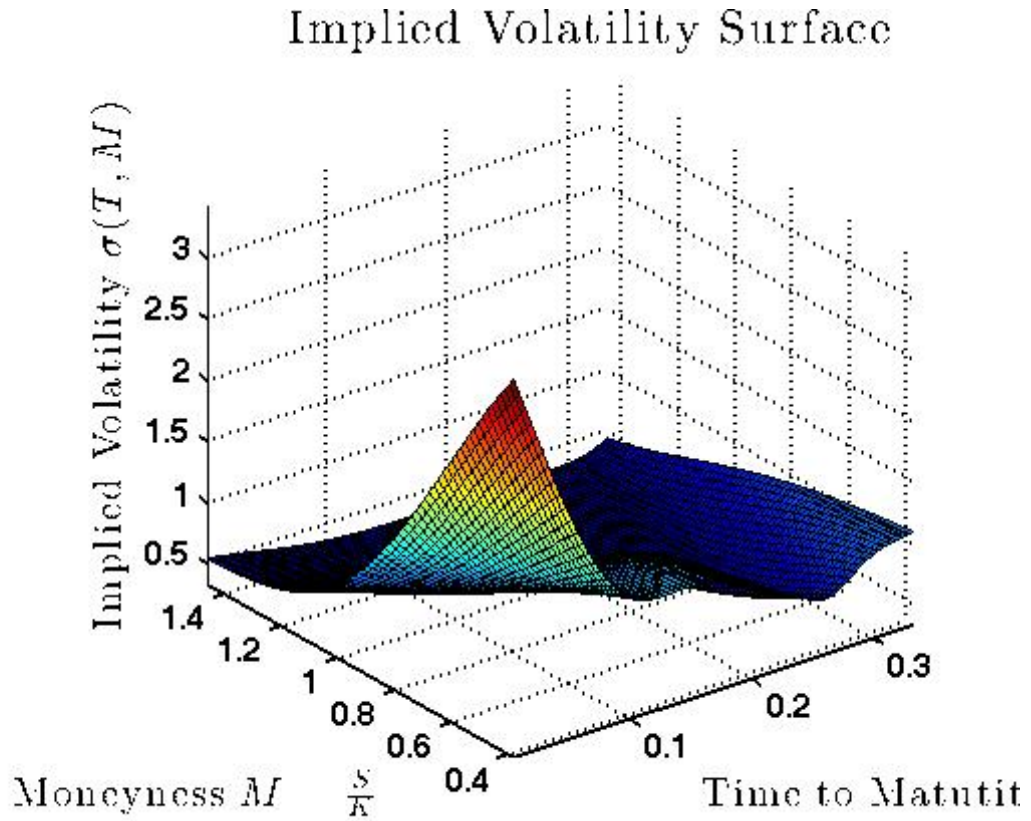
61

Figure 5.2: Plot of volatility surface with Gaussian kernel

5.3.3 Error Analysis

In statistics, r squared is a number used to calculate the fit of the regression model.

After get the coefficients for kernel functions, the R squared for both call and put are calculated to see the fit of the model. The results are promising. If $\bar{y}$ is the mean of the observed data y_i, $\hat{y}_i$ is the estimate, $i = 1, 2, \ldots$,

$$\bar{y} = \frac{1}{n} \sum y$$

Table 5.3: Result of linear model of put option with Gaussian

Stock	T-t	Strike	Option	beta	volatility
520.03	0.024	360	0.02	1.48354	-0.01642
520.03	0.024	365	0.025	1.33887	-0.0142
520.03	0.024	370	0.025	1.04024	-0.001219
520.03	0.024	375	0.025	0.889048	-0.0104
520.03	0.024	380	0.025	0.73829	-0.0879

The total sum of the squares SS_{tot} is defined as

$$SS_{tot} = \sum (y_i - \bar{y})^2$$

From the idea of regression, we care about the residues of the fitting, the residue sum of the squares is defined as

$$SS_{res} = \sum (y_i - \hat{y}_i)^2$$

R squared calculate the residues of all the data fitting, the formulas is given as

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

$R^2 = 1$ is the perfect fit. $R^2 = 0$ means no fitting at all. From the testing result in appl data set, R squared for 1168 entries of European put data is 0.9695575. European call option R squared is 0.813300828.

5.4 European Put option implementation

Applying on European vanilla put options on the same option chain data set, the implementation result is given in the table 4.4. As we see the coefficients beta for put options are mostly negative.

5.5 Application of Other Kernels

By changing the kernel functions,we can expect different result.

5.5.1 Polynomial Kernel Implementation

Polynomial kernels are considered to be applied.

Polynomial kernel of degree 2 is applied to the call and put data. The polynomial kernel is defined as,

$$K(x, x') = (\langle x, x' \rangle + 1)^2$$

Notice here the scale and offset of the polynomial kernel are both 1 by choice. The result for call is not as good as Gaussian kernel, the R squared is 0.4573. Either the result for put data, the R squared is around 0.4494.

5.5.2 Laplace Kernel Implementation

The Laplace kernel is another general purpose kernel along with Gaussian kernel and polynomial kernel. It is usually called Laplace radial basis kernel in machine learning, Laplace kernel function is defined as

$$K(x, x') = exp(-\sigma \|x - x'\|)$$

The distance used here is Euclidean distance. The result for call option is even better, the R-squared is 0.999865.

CHAPTER 6: CONCLUSIONS AND FUTURE WORK

This chapter discusses the results and conclusions with some future directions of work opened by applying this method. The rationale for using the kernel method and regularization technique are given. We use Apple option chain data to the nonparametric method fitting without using kernel smoothing and regularization. The graph shows an unstable and noisy result. We also show that it is necessary by applying the regularization compared to ordinary least squares. Figure 6.2 gives an illustration of the error analysis for the method without using regularization techniques. The comparison of our approach and other methods are given in section 6.1.3. The conclusions of our contribution are given in 6.2. Some directions of future work are proposed.

6.1 Comparison with Other Works

6.1.1 Reproducing Kernel Use

In the nonparametric method, the kernel functions are assumed to be smooth. Without the smoothing of the kernel functions, we tend to get a jagged and noisy result. The following graph gives the illustration of the volatility function estimation without using the kernel smoothing. The following graph is an illustration of the use of reproducing kernel method by applying to real data, AAPL option chain data on November 13, 2013.

By comparing to the example we used in Figure 2.2, we can draw the conclusion that one can only get a rough and jagged graph without applying the kernels in nonparametric methods. Hence the reproducing kernel Hilbert space is the way to go for a generally smooth volatility function.
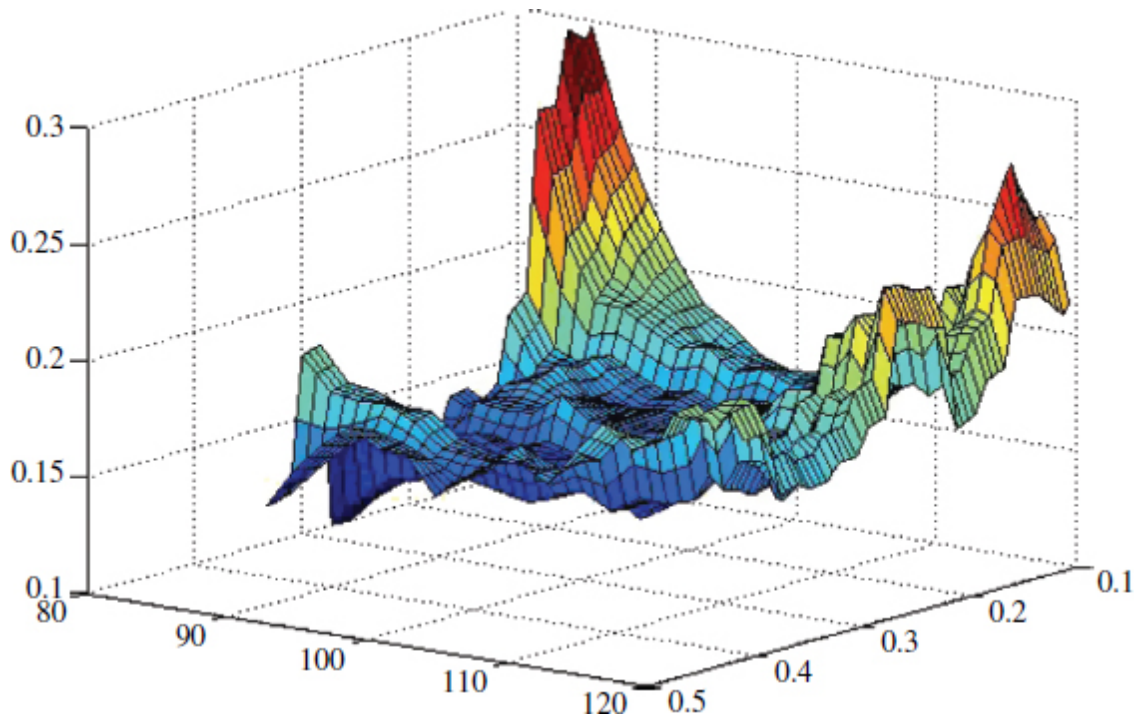
Figure 6.1: Plot of volatility without Kernel smoothing

6.1.2 Regularization Use

In this subsection we want to show the difference of using Tikhonov regularization and without the regularization in the linear model, also known as, the ordinary least squares, i.e.

$$x = argmin\|f(x_i) - y^\delta\|^2 \tag{6.1}$$

where $f(x)$ is defined in (5.19). y^δ is the observed option prices. Figure 6.2 gives the error analysis for ordinary least squares without using regularization. The residual data of the ordinary least square model gives the difference between the observed data of the dependent variable y_i and

66

the fitted values $\hat{y} = f(\hat{x})$, it is defined as

$$\text{Residual} = y - \hat{y} \tag{6.2}$$

Figure 6.2 is a plot of the residual of the simple linear regression model of the data set against the independent variable. The residue curve decreases first as number of samples increases; however, it increases dramatically near 90 samples. Also the Scale-Location plot gives the similar result as it shows the standard deviation increases dramatically as the sample number increases to approximately 90. Normal Quantile-Quantile (Q-Q) plots, are a useful tool for assessing how well data sets fit distributions. An ideal regression model looks like a straight curve with minimal variation. As it is shown that it is impossible to get a stable solution as we get in Figure 5.2. Cook's distance is the total distance of the variation between the model prediction and real data. OLS seeks to minimize the vertical distances between the data and the model line, when the data points that are further out towards the extremes will push / pull harder on the lever (i.e., the regression line) like with more leverage.

6.1.3 Comparison with Other Works

Crepey's tree model is known to be hard to apply to complex financial products. By using RKHS, we can confidently use it for more complicated financial system. As we mentioned in the introduction, Engl and Egger, Laganado, Isako used a parametric method by assuming some fixed structures of the volatility functions. Instead of following their example, we use the nonparametric method with the help of reproducing kernels. This method follows the natural trend of the available data, which gives us very good results from the error analysis. The kernel functions can be used to smooth out the jaggedness from the sparse and noisy data. The results of the implementations with different kernels are very promising. Laplace kernel fitting's R-squared is around 0.99, Gaussian
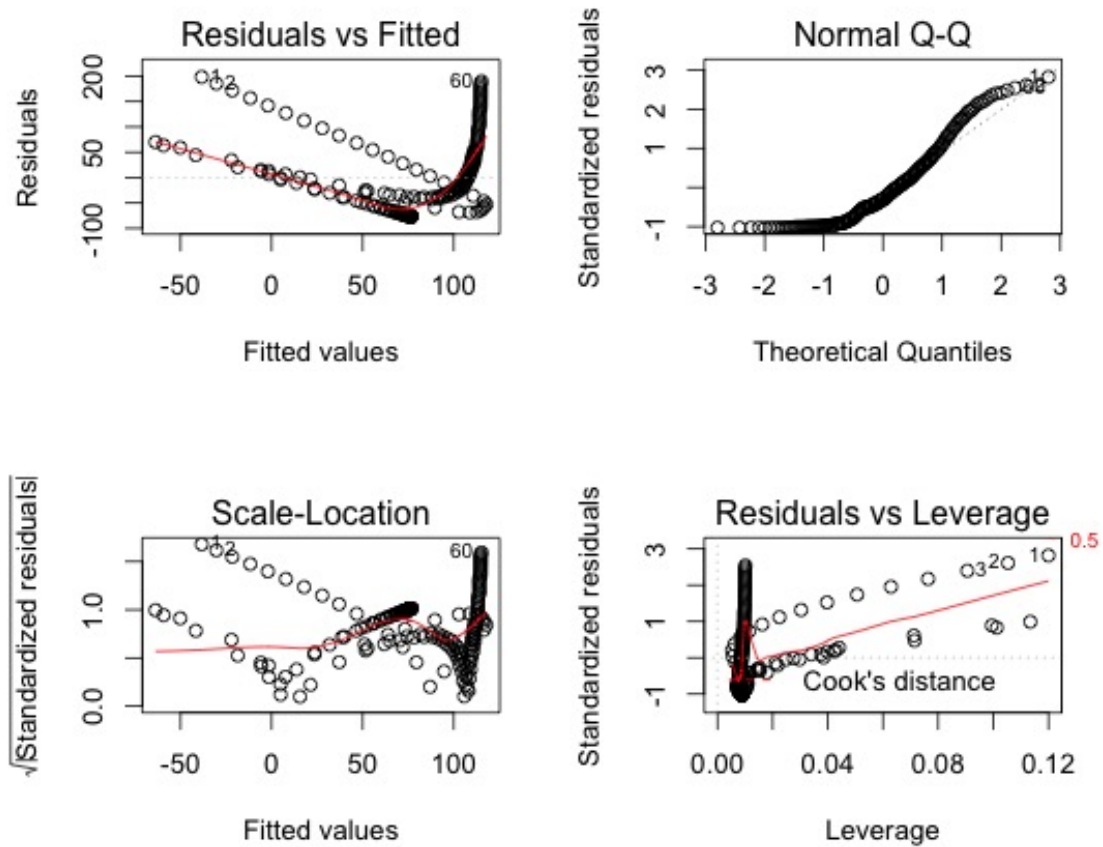
Figure 6.2: Plot of ordinary least squares without regularization

kernel implementation has 0.97 R-squared as well. With such close fitting result, the volatility function can be used for model prediction for hedging and speculations. Figure 6.3 is the plot of B-S implied volatility without using regularization. The deep valley showed in the graph is the so called volatility smile.
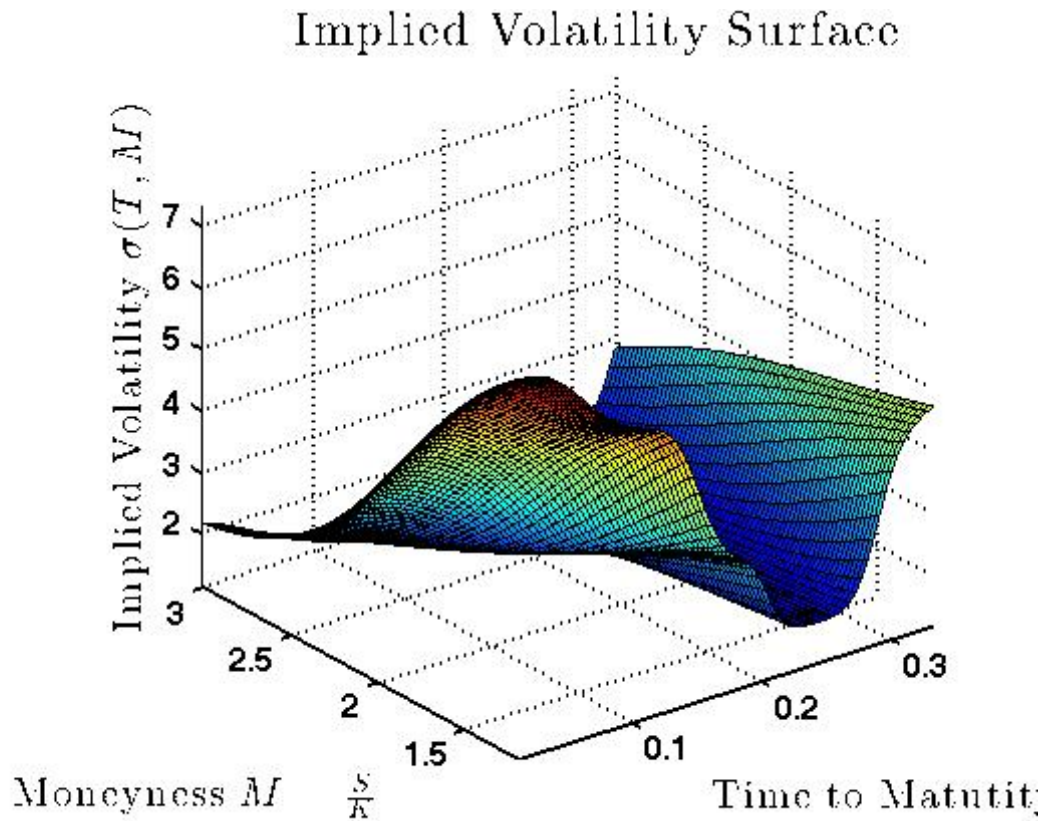
Figure 6.3: Plot of implied volatility with AAPL data

6.2 Conclusions

First we defined the new volatility function $\sigma(\ln(S/K), T-t)$. The new definition considers both financial and time factors. By using the new volatility function, we have a new model with new partial differential equation. We proved that the PDE has solution under the assumption of Hölder continuous of the volatility, which demands a smooth kernel function as well.

Then we estimated the volatility via a nonparametric method in RKHS which gave us the flexibility to fit the model to real data structures. We used the reproducing kernel functions to represent the volatility with a smooth structure, which was essential for a stable and generalized purpose. An

69

example of using kernel versus without using kernel is given in Figure 2.2. The implementation with real data without the using kernel approximation is given in Figure 6.1.

By tuning and adjusting the coefficients of the kernel function, we could represent the true structure of the data. Different kernel functions gave different results. Gaussian and Laplace kernels proved to be more efficient than polynomial kernels and we used the R-squared to show the fitness is very promising.

Both linear and nonlinear models with regularization were studied. The nonlinear model with regularization technique was studied in RKHS.

The linear model with greeks gave us a quick and ready way to estimate volatility in short period of time using easily available option chain data resources. With greeks in the Black-Scholes equation, a linear model was achieved. The regularized solution was calculated for the new model. We applied the studied method for AAPL option chain data and performed error analysis using R-squared. Results are given in Table 5.2 and Table 5.3. Both European put and call options were studied.

Monte Carlo simulations with control variate methods were used to deal with the sparse data. We especially needed consistent data for the nonlinear model, hence the simulation with control variate method used for the numerical analysis. The simulation was controlled by real world data to reduce the variance of the simulated result. Also the simulation solves the sparsity problem of real world data. We find the minimum variance with the control parameter's value. The comparison between controlled and uncontrolled variance is given in Table 4.1.

6.3 Future Work

In this section, we talk about some of the future directions for applying this method:

- Exotic options application, for example, American options, Asian options, Gap options, etc.

- Jump model for stock price

- Machine learning methods applications

    - Bayesian Inference

    - Neural Network

    - Deep learning

    - non-supervised learning

- Other Greeks application for option prices

6.3.1 Exotic Options and Other Derivatives Instruments

Other options can be applied the same technique, for example American Options, Asian Options, Gap options etc. Also, B-S equations are also to be used in other derivatives, for example, future contracts and currency exchanges. The studied method in this dissertation can be applied to them in the future work.

For American options, the early exercise is possible. Under risk neutral probability, an American option with payoff Q^0 is

$$Q_t = \sup_{\tau \in [t,T]} E(e^{-\epsilon_t^\tau} Q^0(S_\tau)|F_t) \tag{6.3}$$

Similarly, for an American put, from put-call parity, we have,

$$P_t = \sup_{\tau \in [t,T]} E(e^{-\in_t^\tau}(K - Se^{r - \frac{\sigma^2}{2}(\tau - t) + \sigma(W_\tau - W_t)})^+) \tag{6.4}$$

It is proven that for nondividend paying stocks, American options and European options have the same prices.

Asian Options are path-dependent options. One can use average of stock price as the strike, or average stock price as mean stock price. Also averages can be calculated as arithmetic way or geometric averages. For example, the arithmetic averaging,

$$I_T = \int_0^T S_\tau d\tau, \ \ A_T = \frac{1}{T} I_T, \tag{6.5}$$

or geometric averaging

$$A_T = \exp(\frac{1}{T} \int_0^T \log(S_\tau) d\tau). \tag{6.6}$$

Depending on how the average is calculated, and the mean as strike or stock price, there are many different kinds of Asian options.

6.3.2 Jump Models

A lognormal distribution is known to assign a considerably low probability for large stock price movements. Assuming jumps in moves can help deal with this issue. Then we want to know the number of jumps and also the magnitude of each jump.

A Poisson distribution in statistics is a discrete probability distribution that counts the number of the events, in this case large stock price moves, that occur over a period of time. The probability

that the event occurs exactly m times is given by Poisson distribution,

$$\frac{\lambda^k}{k!}e^{-\lambda}$$

For example, Merton's jump model which gives out the modified Black- Scholes formula for option prices. See [66]. Considering jumps,the B-S PDE becomes,

$$C_t + \frac{1}{2}C_{SS}\sigma^2 S^2 + C_S(r - \delta)S + \lambda E_Y[C(SY,t) - C(S,t)] = rC$$

when jumps are lognormal, the price of European call, with slightly modification of some parameters in the B-S formula, can be given as

$$\sum_{i=0}^{\infty}\frac{e^{-\lambda'T}(\lambda'T)^i}{i!}BSCall(S, K, \sqrt{\sigma^2 + i\sigma_J/T}, r - \lambda k + i\alpha_J/T, T, \delta)$$

6.3.3 Bayesian Inverse Theory

In Bayesian inverse theory, the unknown parameters are assumed to be random variables themselves and followed certain distributions which are specified according to our beliefs. In this way, Bayesian approach allows us to incorporate the various information into the estimation. Compared to other available and commonly used estimation techniques, for example, Maximum likelihood estimation, the Bayesian approach is easier to solve, especially with the help of computer softwares.

Let $p(x, y)$ be joint probability density for x and y in any given probability system, then the condi-

73

tional probability density by

$$p(x|y) = \frac{p(x,y)}{p(y)} \tag{6.7}$$

Where the marginal probability density $p(y)$ is

$$p(y) = \int p(x,y)dx \tag{6.8}$$

Bayesian theory is then built on the fundamental formula, the Bayes' Theorem:

$$p(x|y)p(y) = p(y|x)p(x) \tag{6.9}$$

Recently, there has been a significant increase in the use of Bayesian statistics as the Bayesian theory has become more widely accepted and software implementations have proliferated.

Next if we have $X_1, \cdots, X_n$ are i.i.d random variables, let $\pi(\theta)$ be the prior pdf that pre-assigned to the parameter by beliefs. If the joint pdf of $X_1, \cdots, X_n$ is

$$f(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \tag{6.10}$$

By the fact of independence, then the posterior pdf is

$$p(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta} \propto f(x|\theta)\pi(\theta) \tag{6.11}$$

In the calibration problem of option pricing, the Bayesian approach concerns the process of fitting a mathematical model M_θ to a set of observed market data V^* and recording the result as a probability distribution on the parameter θ of M_θ. The analysis can then be extended to find probability distributions for other quantities of interest relating to the model type M. Bayesian theory examines what extra information we can infer about an unknown quantity given observations of a

related quantity. The following idea can be found in [66].

We are looking for a volatility, θ_{cal} minimises the difference between the calculated prices $V(S(M, \theta))$ and market prices V^*, i.e.

$$\theta_{cal} = argmin \, \|V(S(M, \theta)) - V^*\|. \tag{6.12}$$

In [66], Gaussian distribution is used for the prior distribution, the Bayesian likelihood is given

$$P(V^*|\sigma) = \mathbf{1}_{G(\sigma) \leq \sigma^2} \exp[-\frac{1}{2}\lambda_l G(\sigma)] \tag{6.13}$$

Then the posterior is

$$P(\sigma|V^*) \propto P(V^*|\sigma)\pi(\sigma) = \mathbf{1}_{G(\sigma) \leq \sigma^2} \exp[-\frac{1}{2}\lambda_l G(\sigma) + \lambda\|\sigma - \sigma_0\|^2] \tag{6.14}$$

Where

- $G(\sigma)$ represents the difference between the model price and market price

- δ the given spread error

- λ_l is a scaling constant chosen before hand to ensure that the density is not too concentrated on one value of σ.

Then by maximizing the posterior (6.12), it is equivalent to the Tikhonov regularization in Problem 1 as in our research. The connection between Bayesian inverse theory and Tikhonov has been an interesting topic and been recognized by many researchers.

APPENDIX A: HILBERT SPACE AND REPRODUCING KERNEL

HILBERT SPACE

A.1 Reproducing Kernel Hilbert Space Basics

Definition A.1. *Inner product in Hilbert Space $\mathcal{H}$ is a pair satisfies:*

1. Symmetry

$$\langle \mu, v \rangle = \langle v, \mu \rangle \tag{A.1}$$

2. Bilinearity

$$\langle \alpha\mu + \beta v, \omega \rangle = \alpha \langle \mu, \omega \rangle + \eta \langle v, \omega \rangle \tag{A.2}$$

3. Positive definiteness

$$\langle \mu, \mu \rangle \geq 0, \quad \forall \mu \in \mathcal{H} \tag{A.3}$$

$$\langle \mu, \mu \rangle = 0 \Leftrightarrow \mu = 0 \tag{A.4}$$

Definition A.2. *A* Hilbert Space *is a complete vector space with the norm defined by the inner product.*

The space of all square integrable functions on given set Q is a Hilbert space, commonly denoted as L_2. Square integrable means $\int_Q f(s) < \infty, \forall f \in L_2, s \in Q$. The inner product of L_2 is defined as,

$$\langle f, g \rangle = \int_Q f(s)g(s)ds, \quad f, g \in L_2, s \in Q \tag{A.5}$$

Definition A.3. *In a Hilbert space of functions,$\mathcal{H}$, defined on a given nonempty set Q, for any $f \in \mathcal{H}$, $x \in Q$, $K(\cdot, \cdot)$ is called a reproducing kernel if*

$$f(x) = \langle K(x, \cdot), f(\cdot) \rangle \tag{A.6}$$

Theorem A.4. *A Reproducing Kernel Hilbert Space $\mathcal{H}_K$is a Hilbert space with a reproducing*

kernel $K(\cdot, \cdot)$whose span is dense in $\mathcal{H}$

A.2 The Representer Theorem

Theorem A.5. Riesz Representation Theorem *In a Hilbert space $\mathcal{H}$, all continuous linear functionals from $\mathcal{H}$ in to the field R or C can be represented uniquely as,*

$$f(x) = \langle f, h_x \rangle, \ \ \forall f \in \mathcal{H} \tag{A.7}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product defined in $\mathcal{H}$, h_x is unique in $\mathcal{H}^*$, the adjoint space of $\mathcal{H}$, which is the space of all linear bounded functionals from $\mathcal{H}$ in to the field R or C.

In statistical machine learning area, the well-known representer theorem introduced by Kimeldorf and Wahba is applied for minimizing functional

$$J(f) = \frac{1}{n} \sum_{i=1}^{n} (f(x_i) - y_i)^2 + \lambda ||f||^2 \tag{A.8}$$

We give the proof for a different empirical risk function for our option pricing calibration problem

$$J(\sigma) = \sum_{i=1}^{n} \sum_{j=1}^{m} [C(T_i, K_{ij}, \sigma) - \overline{C_{ij}}]^2 + \alpha ||\sigma||_{H_k}^{\ 2} \tag{A.9}$$

Let $(T, K) \in Q = R^+ \times (0, S)$, $\sigma(T, K) \in K(a^*)$, $\overline{C_{ij}}$ the average of bid and ask prices for asset S on day $t \in R^+$ for given maturity and strikes $\{T_i, K_{ij}\}$, $i \in (1, \ldots, n), j \in (1, \ldots, m)$ The minimizer over the Reproducing Kernel Hilbert Space $H_K := K(a^*)$ of the regularized functional

$$\sum_{i=1}^{n} \sum_{j=1}^{m} [C(T_i, K_{ij}, \sigma) - \overline{C_{ij}}]^2 + \alpha ||\sigma||_{H_k}^{\ 2} \tag{A.10}$$

Can be represented as

$$\sigma(T, K) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_i K(T, K; T_i, K_{ij}) \tag{A.11}$$

For some n-tuple $(c_1, \cdots c_n) \in R^n$.

Proof. Define kernel mapping $\phi : K(a^*) \to R$, $\phi(x) = K(\cdot, x)$

$$\phi(x)(x') = K(x', x) = < \phi(x'), \phi(x) > \tag{A.12}$$

where $< \cdot, \cdot >$ is the inner product on H_k. Then by the orthogonal projection decomposition of $\sigma \in H_k$

$$\sigma = \sum_{i=1}^{n} c_i \phi(x_i) + v_i \tag{A.13}$$

for some $v_i \in H_{\mathcal{K}}$, we can see the minimizer obtained if and only if $v_i \equiv 0$. Notice

$$||\sigma||^2 = || \sum_{i=1}^{n} c_i \phi(x_i)||^2 + ||v||^2 \geq || \sum_{i=1}^{n} c_i \phi(x_i)||^2 \tag{A.14}$$

Hence the minimum for two dimensional case can be represented as

$$\sigma(x) = \sum_{i=1}^{n} \beta K(x, x_i) \tag{A.15}$$

In this proof we consider the reproducing kernel Hilbert space $W_2^1(Q)$ as a special case for $H_{\mathcal{K}}$

$$W_2^1(Q) = \{u(x) | u(x) \text{ is absolutely continuous}, u'(x) \in L^2(Q)\} \tag{A.16}$$

With inner product defined as

$$< u(x), v(x) >_{W_2^1(Q)} = \int_{-\infty}^{\infty} (u(x)v(x) + u'(x)v'(x))dx \qquad \text{(A.17)}$$

$\square$

APPENDIX B: REGULARIZATION BACKGROUND MATERIAL

B. 1 Pseudo-inverse Operator and Condition Number

In this section, we use the results from [46]. The pseudo-inverse operator or the generalized inverse operator is defined as

Definition B.1. $F : X \rightarrow Y$, X, Y *are given two vector spaces with norms both defined as* $\| \cdot \|$, *the inverse problem of* $Fx = y$ *is ill-posed, then operator* $F^\dagger : Y \rightarrow X$ *is called the pseudo-inverse operator for* F *if,*

$$F^\dagger y = x^\dagger$$

$x^\dagger$ *is the pseudo solution to the inverse problem of* $Fx = y$.

Similarly to the well-posed problems, we have the condition number for the ill-posed case,

Definition B.2. *Let* F *be an operator,* $F^\dagger$, *the generalized inverse of* F, *then the condition number of operator* F *is defined as*

$$cond(F) = \|F\| \|F^\dagger\|$$

A large or infinity condition number implies the problem is ill-posed, which is in the sense that small perturbation in the data results in a large change in the solution. By Theorem 3.2, $\|R_\alpha\| \rightarrow \infty$ as $\alpha \rightarrow 0$. The second part in (3.5) denotes the approximation error $\|R_\alpha Fx - x\|$ for $Fx = y$. By regularization idea, as α tends to zero, this term will go to zero accordingly. Hence the strategy is to keep the total error as small as possible for the choice of $\alpha(\delta)$ which depends on the noise level δ. In this section we have a look at the regularization method can improve the ill-posedness for inverse problem $Fx = y$ in a RKHS.

Theorem B.3. *If a kernel function is defined as* $K : Q \times Q \rightarrow R$, $\omega = (x, y)$, $X = (x_1, ...x_m) \in R^m$, $Y = (y_1, ...y_m) \in R^m$ *and* $\alpha > 0$, *then*

(i) Unique solution exists, and x_α in $H_K(Q)$.

(ii) When K is positive definite, the condition number of K is given by

$$cond(K_\alpha) := cond(K[x] + \alpha mI) = 1 + \frac{(condK[x] - 1)\lambda_{min}}{\lambda_{min} + \alpha m} \tag{B.1}$$

where λ_{min} is the smallest singular value of $K[x]$

From this theorem, we can see for $\alpha m \to \infty$, then $1 + \frac{(condK[x]-1)\lambda_{min}}{\lambda_{min}+\alpha m} \to 1$, which shows the ill-posedness of the inverse problem in RKHS. By applying regularizer with parameter α, we can control the ill-posedness so that the condition number can be reduced accordingly.

Theorem B.4. *Let $K : Q \times Q \to R$ be a kernel, m, the size of data set, and $\mathcal{K}(X)$the kernel matrix. Then given any $\alpha > 0$,*

(i) $cond(K_\alpha) \leq 1 + \frac{1}{\alpha m}$.

(ii) If $m \geq 2$ and $\alpha \geq \frac{1}{m}$, then $cond(K_\alpha) \leq 2$.

(iii) $cond(K[x] + \alpha mI) < cond(K[x])$.

Proof. By using the symmetry of the kernel matrix,

$$cond(K_\alpha) = 1 + \frac{(condK[x] - 1)\lambda_{min}}{\lambda_{min} + \alpha m} = 1 + \frac{(1/\lambda_{min} - 1)\lambda_{min}}{\lambda_{min} + \alpha m} = 1 + \frac{1 - \lambda_{min}}{\lambda_{min} + \alpha m} = \frac{1 + \alpha m}{\lambda_{min} + \alpha m} \tag{B.2}$$

the parameter $\alpha > 0$,

$$\frac{1 + \alpha m}{\lambda_{min} + \alpha m} < 1 + \frac{1}{\alpha m} \tag{B.3}$$

Hence, for any $\lambda_{min} > 0$,

$$\alpha m < \lambda_{min} + \alpha m \tag{B.4}$$

(ii) can be easily seen from the definition of $cond(K_\alpha)$.

(iii) Since $\frac{\lambda_{min}}{\lambda_{min}+\alpha m} < 1$, multiply $cond(K[x]) - 1$ on both sides of the inequality,

$$1 + \frac{(cond(K[x]) - 1)\lambda_{min}}{\lambda_{min} + \alpha m} < cond(K[x]) \tag{B.5}$$

Therefore,

$$cond(K[x] + \alpha m I) < cond(K[x]). \tag{B.6}$$

There are plenty of regularization techniques, one of the most popular methods is Tikhonov Regularization.

B.2 Classic Tikhonov Regularization

Lemma B.5. *Let us consider a linear bounded operator $F : X \to Y$, X, Y are given two vector spaces with norms both defined as $\| \cdot \|$, the inverse problem $Fx = y$, $y \in Y$, the pseudo-solution, $x^{\dagger} \in X$, exists if and only if*

$$F^* F x^{\dagger} = F^* y$$

$F^* : Y \to X$ *is the adjoint operator of F.*

Sketch of proof:

A linear bounded operator $F : X \to Y$, X, Y are given two vector spaces with norms both defined as $\| \cdot \|$, the inverse problem $Fx = y$, $y \in Y$, $x_0 \in X$ is a priori, $\alpha > 0$, the regularization parameter, is also a prior, the $Tikhonov\,functional$ is defined as

$$J_{\alpha}(x) := \|Fx - y\|^2 + \lambda\|x - x_0\|^2 \quad for\ x \in X. \tag{B.7}$$

where the regularization parameter λ increases smoothness at the cost of some additional error.

It is often a custom to choose a stronger norm for the regularizer. Instead of (3.6), we choose to minimize

$$J_\alpha(x) := \|Fx - y\|^2 + \alpha \|x - x_0\|_1^2 \quad for\ x \in X_1 \tag{B.8}$$

where $\|\cdot\|_1$ is a stronger norm on a subspace $X_1 \subset X$. This was firstly introduced by Phillips and Tikhonov. The stronger norm gives the smooth feature to the solution in the regularization framework. Inspired by this idea, we choose the reproducing kernel $\|\cdot\|_{H_K}$ in reproducing kernel Hilbert space H_K to establish the regularization scheme.

Theorem B.6. *For inverse problem $Fx = y$ with $F : X \to Y$ a compact operator and $dimX = \infty$, then we have*

(i) The regularization operators R_α are not bounded, which is equivalent to, for some sequence (α_k), $\|R_\alpha Fx\| \to \infty$ as $k \to \infty$.

(ii) Sequence $(R_\alpha F)$ does not converge uniformly on any bounded set of X, which means there is no convergence of $(R_\alpha F)$, such that $R_\alpha Fx = x$.

Proof: If it is not true, by contradiction, we assume that R_α is bounded, i.e. there exists $M > 0$, such that

$$\|R_\alpha\| \leq M, \quad for\ any\ given\ \alpha > 0.$$

Thus the inverse operator of F is bounded, which implies the identity operator $I = F^{-1}F : X \to X$ is compact, which contradicts with the infinity dimension of X. The conclusion can be seen that if F is compact, then the operators R_α are not uniformly bounded and the regularization $R_\alpha F$ can not converge uniformly to identity. [33] As we mentioned in section 2.4, the inverse problem $Fx = y$ with the data y observed with error, y^δ is denoted as the noisy data, δ is the noise level, i.e. $\|y - y^\delta\| \leq \delta$ so we will apply the regularization, the error of approximation solution $x_{\alpha,\delta}$ is,

with α the regularization parameter,

$$\|x_{\alpha,\delta} - x\| \leq \|R_\alpha y^\delta\| + \|R_\alpha y - x\| \tag{B.9}$$

$$\leq \|R_\alpha\|\|y^\delta - y\| + \|R_\alpha F x - x\| \tag{B.10}$$

therefore, the error is

$$\|x_{\alpha,\delta} - x\| \leq \delta\|R_\alpha\| + \|R_\alpha K x - x\| \tag{B.11}$$

From this inequality, we can see the error of the approximation solution depends not only on the penalty from the regularization but also on the error level of the perturbed data. The regularization strategy is a minimization of the data term and the regularization term. From the right hand of the inequality, we see there are two parts, one part is the error in the data multiplied by $\|R_\alpha\|$, which is the "condition number", which is defined as $cond(F) = \|F\|\|F^{-1}\|$ for a well-posed problem. Inverse problems are often ill-posed. When a solution does not exist, then we are looking for the best approximation solution, the *pseudo solution*. Then we can find a operator, called *pseudo-inverse operator*.

Next we will see Tikhonov regularization is a stable way to deal with the ill-posed inverse problem for linear operator equations. Details see [46].

Theorem B.7. *Given a linear bounded operator $F : X \to Y$, X, Y are given two vector spaces with norms both defined as $\|\cdot\|$, the inverse problem $Fx = y$, $y \in Y$. Then the Tikhonov functional J_α has a unique solution $x^\alpha \in X$, which can be represented by,*

$$\alpha x^\alpha + F^* F x^\alpha = F^* y \tag{B.12}$$

Here the operator $F^* : Y \to X$ is the adjoint operator of F. i.e. the operator such that

$$(Fx, y)_Y = (x, F^*y)_X \tag{B.13}$$

for all $x \in X$, $y \in Y$, $(\cdot, \cdot)_X$ and $(\cdot, \cdot)_Y$ are the inner products for the Hilbert spaces X and Y respectively.

Theorem B.8. *let $F : X \to Y$ be a linear and compact operator, X, Y are given two vector spaces with norms both defined as $\|\cdot\|$, the inverse problem $Fx = y$, $y \in Y$, the regularization parameter $\alpha > 0$, then operator $\alpha I + F^*F$ is bounded and invertible for $\alpha(\delta) \to 0$ with $\delta^2/\alpha(\delta) \to 0$ as $\delta \to 0$ If we define the regularization operator*

$$R_\alpha := (\alpha I + F^*F)^{-1}F^* : Y \to X$$

then $\|R_\alpha\| \leq 1/(2\sqrt{\alpha})$.

LIST OF REFERENCES

[1] F. Black and M. Scholes, *The pricing of options and corporate liabilities*. The Journal of Political Economy, 1973.

[2] M. Craddock, C. Chiarella and N. El-Hassan, *The calibration of stock option pricing models using inverse problem methodology*.

[3] S. Crepey, *Calibration of the local volatility in a generalized black-Scholes model using Tikhonov regularization*. SIAM Journal of Mathematical Analysis, 2003.

[4] H. Egger and H. W Engl, *Tikhonov regularization applied to the inverse problem of option pricing: Convergence analysis and rates*.

[5] T. Hein, *Some analysis of Tikhonov regularization for the inverse problem of option pricing in the price-dependent case*. Journal for Analysis and its Applications, 2005.

[6] L. S. Jiang and B. J. Bian, *Identifying the principal coefficient of parabolic equations with non-divergent form*. Journal of Physics: Conference Series, 2005.

[7] Y. Achdou, *Computational Methods for Option Pricing*, SIAM, 2006.

[8] R. Lagnado and S. Osher, *Reconciling differences*. Risk, 1997.

[9] M. Z. Nashed, *Approximate regularized solutions to improperly posed linear integral and operator equations*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1974.

[10] M. Z. Nashed and Grace Wahba, *Generalized inverses in reproducing kernel spaces: An approach to regularization of linear operator equations*. SIAM J. Math. Anal., 1974.

[11] M.G. Cui, *Nonlinear Numerical Analysis in the Reproducing Kernel Space*, Nova, 2008.

[12] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer, 1996.

[13] J. Locker and P. M. Prenter, *Regularization with differential operators*, Journal of Mathematical analysis and applications, 1980.

[14] M. Baxter and A. Rennie, *Financial Calculus an introduction to derivative pricing* Cambridge, 2001.

[15] A. Berlinet and C. Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Kluwer Academic Publishers, 2004.

[16] B.Oksendal, *Stochastic Differential Equation*, Springer-Verlag, 2000.

[17] H. W. Engl, A Binder and a Schatz. *Advanced numerical techniques for financial engineering*. Springer, 2000.

[18] I. Bouchouev and V. Isakov, *The inverse problem of option pricing*. Inverse Problems, 1997.

[19] F. Black and M. Scholes, *The pricing of options and corporate liabilities*. The Journal of Political Economy, 1973.

[20] B. Dupire, *Pricing with a smile*. Risk, 1994.

[21] H. Egger and H. W Engl, *Tikhonov regularization applied to the inverse problem of option pricing: Convergence analysis and rates*.?

[22] H. W. Engl, *Calibration problems - an inverse problems view*. WILMOTT magazine, ?.

[23] P. Friz and J. Gatheral, *Valuation of volatility derivatives as an inverse problem*.?

[24] K. H. and F. Klebaner, *On one inverse problem in financial mathematics*. Journal of Uncertain Systems, 2007.

[25] C. Jin and X. Ni, *A new regularized algorithm to calibrate implied volatility in option pricing models*. IEEE, 2010.

[26] L. Jiang and Y. Tao, *Identifying the volatility of the underlying assets from option prices*. Inverse Problems, 2001.

[27] Y. L. Keung and J. Zou, *Numerical identifications of parameters in parabolic systems*. Inverse Problems, 1997.

[28] L. Wang, L. Jiang, Q. Chen and J. E. Zhang, *A new well-posed Algorithm to recover implied local volatility*. Quantitative Finance, 2003.

[29] R. Lagnado and S. Osher, *Reconciling differences*. Risk, 1997.

[30] Zaremba S. *Sur le calcul numerique des founctions demandess dans le problems de direchlet et le problems hydrodynamique*. Bulletin International de I Academie des Sciences de Cracovie, 1908.

[31] Aronzaj N. *Theory of reproducing kernels*. Trans. Amer. Math. Soc, 1950.

[32] H. W. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems* Kluwer Academic Publishers, 1996.

[33] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*. Springer, Berlin, 1996.

[34] V. G. Romanov, S.I. Kabanikhin, *Ill-posed and Inverse Problem*. VSP, Boston, 2002.

[35] Pierre, *The adjoint method formulation for an inverse problem in the generalized Black-Scholes model*. [Sen] A. Sen, An optimization approach to computing the implied volatility of American options.

[36] A. Sever, *A stability estimate of an inverse problem in financial prospection—*. Applied Mathematics and Computation, 2004.

[37] A. Sever, *A linearization-based solution to an inverse problem in financial markets*. Applied Mathematics and Computation, 2007.

[38] S. Kotani, S. Kaji, *Financial inverse problem and reconstruction of infinitely divisible distributions with Gaussian component*.

[39] D. A. Mendes, S. R. Bentes, R. Menezes, *Stock market volatility: An approach based on Tsallis entropy*.

[40] M. O. Souza and J. P. Zubelli, *Real options under fast mean reversion stochastic volatility*.

[41] M. O. Souza and J. P. Zubelli, *Inverse problems and regularization techniques in option pricing*. Proceedings in Applied Mathematics and Mechanics, 2007.

[42] C. Tsallis and C. Anteneodo, *Nonextensive statistical mechanics and economics*.

[43] Y. Li, T. F. Coleman and A. Verma, *Reconstructing the unknown local volatility function*. Journal of Computational Finance, 1999.

[44] L. Yang, Zui-Cha Deng, Jian-Ning Yu, *An inverse problem of determining the implied volatility in option pricing*. Mathematical Analysis and Applications, 2008.

[45] M. Z. Nahsed, ed., *Generalized Inverses and Applications*. Academic Press, 1976.

[46] Katerina Hlavackova-Schindler, *Tikhonov Regularization Parameter in Reproducing Kernel Hilbert Spaces with Respect to the Sensitivity of the Solution* Springer-Verlag Berlin Heidelberg, 2008.

[47] M. Hintermuller, M. Tber *An Inverse Problem in American Options as a Matematical Program with Equilibrium Constraints: C-Sattionarity and An Active-Set-Newton Solver*. SIAM, 2010.

[48] R. Temam, *Numerical Analysis*. D. Reidel Publishing Company, 1973.

[49] M. Rubinstein, *Implied Binomial Trees* Journal of Finance, 1994.

[50] E. Derman and I. Kani, *Riding on a Smile* Risk, 1994.

[51] J. Cox and M. Rubinstein *Option Pricing: a Simplified Approach* Journal of Finance, 1979.

[52] P. Wilmott, *Quantitative Finance* Wiley, 2000.

[53] R. Adams and J. F. Fournier, *Sobolev Spaces*, Elsevier, 2003.

[54] B. Hofmann and O. Scherzer, *Local Ill-posedness and Source Conditions of Operator Equations in Hilbert Spaces*. Inverse Problems, 1998.

[55] O.A. Landyzenskaya, V. A. Solonikov and N. N. Urealceva, *Linear and Quasilinear Equations of Parabolic Type*. AMS, 1968.

[56] V. Isakov, *Inverse Problems for Partial Differential Equations*. Springer, New York, 2006.

[57] M. Z. Nashed, Q. Sun *Sampling and Reconstruction of Signals in a Reproducing Kernel Subspace of $L_p(R_d)$*, Journal of Functional Analysis, 2010.

[58] V. M. Cornelis, Z. M. Nashed, S. Seatzu, *Sampling Expansions and Interpolation in a Unitarily Translation Invariant Reproducing Kernel Spaces*, Kluwer Academic Publishers, 2003.

[59] I. Karatzas and S. E. Shreve, *Methods of Mathematics Finance*. Springer-Verlag, New York, 1998.

[60] R. Korn and E Korn, *Option Pricing and Portfolio Optimization*. AMS, 2001.

[61] A. R. Gallant, D. Hsieh and G. Tauchen, *Estimation of the Stochastic Volatility Models with Diagnostics*. Elsevier, Journal of Econometrics, 1997.

[62] S. L. Heston, *A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*. Society of Financial Studies, 1993.

[63] T. G. Anderson, B. E. Sorensen, *GMM estimation of a Stochastic Volatility Model*. Journal of Business and Economic, 1996.

[64] O.A. Landyzenskaya, Linear and Quasilinear equations of Parabolic type, 1968.

[65] Kimeldorf, George S., Wahba, Grace, *A correspondence between Bayesian estimation on stochastic processes and smoothing by splines*. The Annals of Mathematical Statistics, 1990.

[66] R.L.McDonald, *Derivatives Markets*, Pearson, 2006.

[67] W. Shaw, *Modeling Financial Derivatives with Mathematica*, Soviet Math, 1999.

[68] G. Wahba, *Spline Models for Observational Data*, SIAM, 1990.