Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 314104, 7 pages
http://dx.doi.org/10.1155/2014/314104

Research Article

The Adjoint Method for the Inverse Problem of Option Pricing

Shou-Lei Wang,[1,3] Yu-Fei Yang,[1,2] and Yu-Hua Zeng[3]

[1] *College of Mathematics and Econometrics, Hunan University, 410082 Changsha, China*
[2] *Department of Information and Computing Science, Changsha University, 410003 Changsha, China*
[3] *Department of Mathematics, Hunan First Normal University, 410205 Changsha, China*

Correspondence should be addressed to Shou-Lei Wang; shouleiw@163.com

Received 25 November 2013; Accepted 17 February 2014; Published 26 March 2014

Academic Editor: Yang Tang

The estimation of implied volatility is a typical PDE inverse problem. In this paper, we propose the $TV - L^1$ model for identifying the implied volatility. The optimal volatility function is found by minimizing the cost functional measuring the discrepancy. The gradient is computed via the adjoint method which provides us with an exact value of the gradient needed for the minimization procedure. We use the limited memory quasi-Newton algorithm (L-BFGS) to find the optimal and numerical examples shows the effectiveness of the presented method.

1. Introduction

An option is classified as either a call option or a put option. A call (or put) option is a contract which gives its holder the right to buy (or sell) a prescribed asset, known as the underlying asset, by a certain date (expiration date) for predetermined price (commonly called the strike price or exercise price). The revolution in trading and pricing derivative securities began in the early 1970s. In 1973, Black and Scholes [1] published their seminal papers on the theory of option pricing and obtained the partial differential equation depicting the option prices:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - q) S \frac{\partial V}{\partial S} - rV = 0, \tag{1}$$

where $(S, t) \in (0, \infty) \times (0, T)$, and $V(S, t)$ is the value of option price. The asset price S is modeled to satisfy the Geometric Brownian motion, σ is the volatility, r is the riskless interest rate, and T is the maturity.

The payoff function at maturity and boundary conditions are given by

$$\begin{gathered} V(S, t)|_{t=T} = (S - K)^+ = \max(0, S - K), \quad \text{call option}, \\ V(0, t) = 0, \quad (S, t) \in (0, \infty) \times (0, T), \\ \lim_{S \to \infty} \frac{V(S, t)}{S} = 1, \end{gathered} \tag{2}$$

where K is the strike price. The analytical solution of the European call option is

$$V(S, t) = SN(d_1) - Ke^{-r(T-t)} N(d_2), \tag{3}$$

where

$$\begin{gathered} N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\omega^2/2} d\omega, \\ d_1 = \frac{\ln(S/K) + \left(r - q + \left(\sigma^2/2\right)\right)(T - t)}{\sigma\sqrt{(T - t)}}, \\ d_2 = d_1 - \sigma\sqrt{(T - t)}. \end{gathered} \tag{4}$$

The option prices $V(S,t)$ are functions of five parameters: S, K, r, T, and σ. Except for the volatility, the other four parameters S, K, r, and T are assumed or can be directly observed in the market. If the volatility is a constant, (1) becomes the classical Black-Scholes model. However, in the actual market volatility is changing [2, 3]. Volatility is a measure of the amount of flaction in the asset prices, that is, a measure of the randomness. It is necessary to measure it accurately in portfolio, asset pricing, risk management, and monetary policy. The estimation of volatility has been an important research topic of modern financial markets.

The volatility value implied by an observed option price is called the implied volatility. In market, empirical studies have revealed that no constant or merely time dependent local volatility function is consistent with most sets of market quotes; such phenomena are commonly called the volatility smile by market practitioners. In this paper, we are interested in the inverse problem of option pricing (IPOP). One possibility to explain the volatility smiles in Black-Scholes model is to use a deterministic function of underlying asset price S and time t; that is, $\sigma = \sigma(S,t)$. We only discuss the case of $\sigma(S,t) = \sigma(S)$ on European call options.

The inverse problem of option pricing was first considered by Dupire [4]. He obtained a local volatility formula for all strike prices and maturities; however it is instable. Bharadia et al. [5] derived a simple volatility formula that does not require the option to be exactly at-the-money. Quasi-iterative technique for computing the implied volatility was proposed by Chance [6]. Chambers and Nawalkha [7] restricted Chance's Taylor expansion to be only in volatility, improving its accuracy. Utilizing the third-order Taylor series expansion, Li [8] developed a new close formula of implied volatility. Ballestra and Cecere [9] proposed a highly efficient approach to compute the volatility of the Fractional Brownian Motion implied by American options. Research results concerning inverse problem of option pricing with Tikhonov regularization [10] strategies have been intensively published in recent years; see, for example, Chiarella et al. [11], Crépey [12], Deng et al. [13], Egger and Engl [14], Isakov [15], Jiang and Tao [16], Leland [17], Lagnado and Osher [18], Ngnepieba [19], and references therein. However, the classical Tikhonov regularization may oversmooth the solution of the origin problem. If the exact solution is nonsmooth or even has some singularities, the regularized solution cannot approximate effectively. These shortcomings will blur the edge of the restored image in image processing. Based on the advantage that the total variation (TV) regularization can preserve the edge of the image, Rudin et al. [20] proposed the $\mathrm{TV} - L^2$ model (also called the ROF model):

$$\min_{u\in\Omega} \frac{\lambda}{2}\|u - f\|^2_{L^2(\Omega)} + |\nabla u|_{L^1(\Omega)}. \tag{5}$$

Considering the jump, overnight, and weekend effect [21, 22] of volatility, the total variation regularization might be able to depict the properties of volatility better. So whether the TV regularization strategy could be applied to identify the implied volatility is a question worth pondering.

L^1 fidelity-based model has many desirable and unexpected consequences in applications, such as data-driven parameter selection and multiscale image decomposition. Since the TV regularization should be used in the second step in order to reconstruct jump discontinuities in inverse problem, the reasonable choice of fidelity in the first step is the L^1 fidelity. In this paper, we consider the minimization of TV regularization under L^1 fidelity. The adjoint method provides us with an exact value of the gradient needed for the minimization procedure.

This paper is organized as follows. In the next section, we put forward the $\mathrm{TV} - L^1$ model for determining the implied volatility. In Section 3, we deduce the semidiscrete form of the Black-Scholes equation and introduce the adjoint model. Time discretization and the L-BFGS algorithm [23] are given in Section 4. In Section 5, we present a selection of numerical examples. In the last section, we give some remarks to conclude the paper.

2. Total Variation Regularization Model

Let $\mathbb{X}$ and $\mathbb{Y}$ be Hilbert spaces. The standard form of an inverse problem is as follows. Give $y \in \mathbb{Y}$ and $F : D(F) \subset \mathbb{X} \to \mathbb{Y}$; find $a \in D(F)$ such that $F(a) = y$, where F is a nonlinear operator between $\mathbb{X}$ and $\mathbb{Y}$. Recall that an inverse problem is well posed if and only if the three conditions of Hadamard are satisfied: the existence, uniqueness, and continuous dependence of the solutions. Most inverse problems are ill posed.

We assume that only noisy data y^δ of the exact data y is available. To obtain a well-posed problem, the classical Tikhonov regularization strategy is minimizing

$$J^\delta_\alpha(a) = \left\|F(a) - y^\delta\right\|^2 + \alpha\left\|a - a_0\right\|^2, \quad a \in D(F). \tag{6}$$

In this section we consider the problem of inferring a local volatility function $\sigma(S)$ from the observed option prices (take call option for example). Equation (1) is described in an infinite domain $\mathbb{R}^+ \times (0,T)$ which makes it difficult to obtain numerical solutions. We replace the region $\mathbb{R}^+ \times (0,T)$ with the finite rectangle $\Omega := [0, S_{\max}] \times [0,T]$, where $S_{\max}$ is the suitable chosen positive number representing the final value of the asset price; then we have

$$\begin{gathered} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S)\,S^2\frac{\partial^2 V}{\partial S^2} + (r-q)\,S\frac{\partial V}{\partial S} - rV = 0, \quad (S,t) \in \Omega, \\ V(0,t) = 0, \quad t \in [0,T], \\ V(S,T) = (S-K)^+, \quad S \in [0, S_{\max}], \\ V(S_{\max}, t) = (S_{\max} - K)^+, \quad t \in [0,T]. \end{gathered} \tag{7}$$

In the current work, we assume that the market prices V_{ij} for a series of options are known, where V_{ij} is the observed market prices of the options with exercise dates T_i $(T_1, T_2, \dots T_N)$ and strike prices K_{ij} $(K_{i1}, K_{i2}, \dots K_{iM_i})$. We would like to estimate the volatility function $\sigma(S)$ that satisfies the Black-Scholes model (7) using this set of the observations.

In [18] Lagnado and Osher determined this inverse problem by using Tikhonov regularization strategy, that is, attempting to minimise

$$\widetilde{G}(\sigma) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M_i}\left(V\left(S_0, 0, K_{ij}, T_i, \sigma\right) - V_{ij}\right)^2 + \|\nabla\sigma\|_2^2, \quad (8)$$

where $\nabla = (\partial/\partial S, \partial/\partial t)$ denotes the gradient operator. This regularization strategy proposed by Lagnado and Osher was just for one fixed value of underlying asset S_0, at one fixed point in time $t = 0$. There is no guarantee that the value of σ calculated by this approach will be correct either for other underlying assets or at future times; there is also no guarantee that volatility will be positive everywhere.

Based on their work, Chiarella et al. [11] modified the objective functional as follows:

$$\widehat{G}(\sigma) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M_i}\int_0^{\infty}\int_0^{T_{\text{cur}}}\left[V\left(S, t, K_{ij}, T_i, \sigma\right) - V_{ij}\right]^2 dS\,dt + \|\nabla\sigma\|_2^2, \quad (9)$$

where T_{cur} is the current time.

As we know, Tiknonov regularization may oversmooth the solution, so it cannot preserve the singularities of the solution well. In image processing this shortcoming will blur the edge of the restored image. To over this defect, Rudin et al. [20] proposed the total variation regularization strategy. Considering the jump, overnight, and weekend effect of the volatility, we introduce the following optimal control problem ($\text{TV} - L^1$ model):

$$\min_{\sigma} J(\sigma) = \frac{1}{2}\int_0^T \left\|V - V_{ij}\right\| dt + \alpha\int_0^{S_{\max}} |\nabla\sigma|\, dS, \quad (10)$$

where α denotes the regularization parameter, V_{ij} is the corresponding observations, V is the related vector of prices in the Black-Scholes model with volatility function $\sigma(S)$, and ∇ denotes the gradient, in this paper, $|\nabla\sigma| = |\sigma'(S)|$.

To avoid the case $|\nabla\sigma| \approx 0$ in the flat area, as is done in image processing, the problem (10) is usually approximated by using the problem

$$\min_{\sigma} J(\sigma) = \frac{1}{2}\int_0^T \sqrt{\left\|V - V_{ij}\right\|^2 + \varepsilon_1^2}dt + \alpha\int_0^{S_{\max}} \sqrt{\left|\sigma'(S)\right|^2 + \varepsilon_2^2}dS, \quad (11)$$

where ε_1 and ε_2 are two positive parameters which should usually be taken as a constant, for example, $\varepsilon_1 = \varepsilon_2 = 10^{-6}$.

3. Semidiscretization and Adjoint Model

The vega $\partial V/\partial\sigma$ will appear in the optimal necessary condition if we compute the gradient of cost function $\nabla_\sigma J(\sigma)$ directly. The vega (sometimes called kappa) of derivatives is the rate of change of its value with respect to the volatility of the underlying asset. Chiarella et al. [11] determined the vega by using the Black-Scholes formula as an approximation:

$$\begin{aligned}\frac{\partial V}{\partial\sigma} &= SN'\left(d_1\right)\frac{\partial d_1}{\partial\sigma} - Ke^{-r(T-t)}N'\left(d_2\right)\frac{\partial d_2}{\partial\sigma} \\ &= \frac{S\sqrt{T-t}e^{-d_1^2/2}}{\sqrt{2\pi}};\end{aligned} \quad (12)$$

however, it is not an exact value. In this paper, we introduce the adjoint method [19] which provides us with an exact value of the gradient needed for the minimization procedure.

We apply a uniform grid for the computational domain $[0, S_{\max}] \times [0, T]$; let

$$\Delta T = \frac{T}{N_T}, \qquad \Delta S = \frac{S_{\max}}{N_S}. \quad (13)$$

Moreover, we use the notation

$$V_i^n = V\left(S_i, t_n\right), \quad (14)$$

where

$$\begin{gathered} S_i = i\Delta S, \qquad t_n = n\Delta T, \\ i = 0, 1, 2, \ldots, N_S, \quad n = 0, 1, 2, \ldots, N_T. \end{gathered} \quad (15)$$

The first-order and second-order finite differences are used to approximate the space partial derivative $\partial V/\partial S$ and $\partial^2 V/\partial S^2$ in Black-Scholes equation:

$$\begin{aligned}\frac{\partial V}{\partial S}\left(S_i, t\right) &\approx \frac{V_{i+1}(t) - V_i(t)}{\Delta S}, \\ \frac{\partial^2 V}{\partial S^2}\left(S_i, t\right) &\approx \frac{V_{i+1}(t) - 2V_i(t) + V_{i-1}(t)}{\Delta S^2}.\end{aligned} \quad (16)$$

Then, we have (we replace $\sigma(S_i)$ by σ_i for convenience sake)

$$\begin{aligned}&\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_i^2(i\Delta S)^2\left(\frac{V_{i+1}(t) - 2V_i(t) + V_{i-1}(t)}{\Delta S^2}\right) \\ &\quad + (r-q)\,i\Delta S\frac{V_{i+1}(t) - V_i(t)}{\Delta S} - rV_i = 0;\end{aligned} \quad (17)$$

this leads to the following semidiscrete equation:

$$\begin{aligned}&\frac{\partial V}{\partial t} + \frac{1}{2}[\sigma_i i]^2 V_{i-1}(t) - \left[\sigma_i i + (r-q)\,i + r\right]V_i \\ &\quad + \left[\frac{1}{2}[\sigma_i i]^2 + (r-q)\,i\right]V_{i+1} = 0.\end{aligned} \quad (18)$$

$A^\sigma = [a_{ij}^\sigma]$ is a tridiagonal matrix with nonzero elements defined as follows:

$$\begin{aligned} a_{i,i-1}^\sigma &= -\frac{1}{2}(\sigma_i i)^2, \quad i = 1, 2, 3, \ldots, N_S, \\ a_{i,i}^\sigma &= (\sigma_i i)^2 + (r-q)\,i + r, \quad i = 1, 2, 3, \ldots, N_S, \\ a_{i,i+1}^\sigma &= -\left(\frac{1}{2}(\sigma_i i)^2 + (r-q)\,i\right), \quad i = 1, 2, 3, \ldots, N_S. \end{aligned} \quad (19)$$

Indeed $a_{i,j}^{\sigma} \leq 0$ for $i \neq j$; this property guarantees that the space discretization does not cause undesired oscillations into the numerical solution. Equation (18) can be written as

$$\frac{\partial V}{\partial t} = A^{\sigma} V. \tag{20}$$

The directional derivative of option price V, also called the sensitivity in financial theory context, is

$$\widehat{V}(\sigma, h) = \lim_{\beta \to 0} \frac{V(\sigma + \beta h) - V(\sigma)}{\beta}, \tag{21}$$

where h is the perturbation on σ. This combined with (20) implies

$$\begin{gathered} \frac{\partial \widehat{V}}{\partial t} = A^{\sigma} \cdot \widehat{V} + \left[\frac{\partial A^{\sigma}}{\partial \sigma}\right] \cdot Vh, \\ \widehat{V}|_{t=0} = 0. \end{gathered} \tag{22}$$

Similarly, the directional derivative of the cost function $J(\sigma)$ is

$$\begin{aligned} \widehat{J}(\sigma, h) &= \lim_{\beta \to 0} \frac{J(\sigma + \beta h) - J(\sigma)}{\beta} \\ &= \langle h, \nabla_{\sigma} J \rangle = \frac{1}{2} \int_0^T \left\langle \frac{V - V_{ij}}{\sqrt{\left\|V - V_{ij}\right\|^2 + \varepsilon_1^2}}, \widehat{V} \right\rangle dt \\ &\quad + \alpha \left\langle -\frac{\sigma''\left(\left(\sigma'\right)^2 + \varepsilon_2^2\right) - \left(\sigma'\right)^2}{\left(\left(\sigma'\right)^2 + \varepsilon_2^2\right)^{3/2}}, h \right\rangle. \end{aligned} \tag{23}$$

Introducing the adjoint variable P, we have

$$P \cdot \frac{\partial \widehat{V}}{\partial t} = P \cdot A^{\sigma} \cdot \widehat{V} + P \cdot \left[\frac{\partial A^{\sigma}}{\partial \sigma}\right] \cdot Vh; \tag{24}$$

by using integration by parts, the above equation is integrated between 0 and T:

$$\begin{aligned} &\int_0^T P \cdot d\widehat{V} = \int_0^T \left(P \cdot A^{\sigma} \cdot \widehat{V} + P \cdot \left[\frac{\partial A^{\sigma}}{\partial \sigma}\right] \cdot Vh\right) dt, \\ &\Longrightarrow \left\langle P(t), \widehat{V}(t) \right\rangle\Big|_0^T - \int_0^T \widehat{V} dP \\ &= \int_0^T \left[P \cdot A^{\sigma} \cdot \widehat{V} + P \cdot \left[\frac{\partial A^{\sigma}}{\partial \sigma}\right] \cdot Vh\right] dt, \\ &\Longrightarrow \left\langle P(t), \widehat{V}(t) \right\rangle\Big|_0^T - \int_0^T \left\langle \frac{dP}{dt}, \widehat{V} \right\rangle dt \\ &= \int_0^T \left\langle [A^{\sigma}]^T P, \widehat{V} \right\rangle dt \\ &\quad + \left\langle h, \int_0^T V^T \left[\frac{\partial A^{\sigma}}{\partial \sigma}\right]^T P dt \right\rangle, \\ &\Longrightarrow \left\langle P(T), \widehat{V}(T) \right\rangle - \int_0^T \left\langle \frac{dP}{dt}, \widehat{V} \right\rangle dt \\ &\quad - \int_0^T \left\langle [A^{\sigma}]^T P, \widehat{V} \right\rangle dt \\ &= \left\langle h, \int_0^T V^T \left[\frac{\partial A^{\sigma}}{\partial \sigma}\right]^T P dt \right\rangle, \\ &\Longrightarrow -\left\langle P(T), \widehat{V}(T) \right\rangle \\ &\quad + \int_0^T \left\langle \frac{dP}{dt} + [A^{\sigma}]^T P, \widehat{V} \right\rangle dt \\ &= \left\langle h, -\int_0^T V^T \left[\frac{\partial A^{\sigma}}{\partial \sigma}\right]^T P dt \right\rangle. \end{aligned} \tag{25}$$

If we define P, the adjoint variable is the solution of the equation

$$\begin{gathered} \frac{dP}{dt} + [A^{\sigma}]^T P = \frac{1}{2} \frac{V - V_{ij}}{\sqrt{\left\|V - V_{ij}\right\|^2 + \varepsilon_1^2}}; \\ P(T) = 0, \end{gathered} \tag{26}$$

then we have

$$\begin{aligned} &\frac{1}{2} \int_0^T \left\langle \frac{V - V_{ij}}{\sqrt{\left\|V - V_{ij}\right\|^2 + \varepsilon_1^2}}, \widehat{V} \right\rangle dt \\ &= \left\langle h, -\int_0^T V^T \left[\frac{\partial A^{\sigma}}{\partial \sigma}\right]^T P dt \right\rangle. \end{aligned} \tag{27}$$

So the directional derivative of the cost function can be written as follows:

$$\hat{J}(\sigma,h) = \langle h, \nabla_\sigma J\rangle = \left\langle h, -\int_0^T V^T\left[\frac{\partial A^\sigma}{\partial\sigma}\right]^T P dt\right\rangle + \alpha\left\langle -\frac{\sigma''\left((\sigma')^2+\varepsilon_2^2\right)-(\sigma')^2}{\left((\sigma')^2+\varepsilon_2^2\right)^{3/2}}, h\right\rangle; \quad (28)$$

thus, the gradient of the cost function $J(\sigma)$ with respect to the control variable σ is

$$\nabla_\sigma J(\sigma) = -\alpha\frac{\sigma''\left((\sigma')^2+\varepsilon_2^2\right)-(\sigma')^2}{\left((\sigma')^2+\varepsilon_2^2\right)^{3/2}} - \int_0^T V^T\left[\frac{\partial A^\sigma}{\partial\sigma}\right]^T P dt. \quad (29)$$

4. Time Discretization and Algorithm

Considering the stability and high accuracy of the Crank-Nicolson time discretization scheme which can be interpreted as the average of the explicit and implicit Euler schemes, the time discretization of the semidiscrete equation (20) can be written as

$$\frac{V^{n+1}-V^n}{\Delta T} = A^\sigma\left(\frac{V^{n+1}+V^n}{2}\right); \quad (30)$$

then we have

$$\left(I+\frac{\Delta T}{2}A^\sigma\right)V^n = \left(I-\frac{\Delta T}{2}\right)V^{n+1}. \quad (31)$$

The above discrete scheme is second-order accurate and unconditionally stable. Let the boundary condition $P_0^n = 0$, $P_{N_S}^n = 0$, $n = 0,1,\ldots,N_T$; we also use this scheme for the discrete of adjoint equation (26):

$$\frac{P^n - p^{n-1}}{dt} + \left[A^\sigma\right]^T\frac{P^n+P^{n-1}}{2} = \frac{1}{2}\frac{V^n - V_{ij}}{\sqrt{\left\|V^n - V_{ij}\right\|^2+\varepsilon_1^2}}, \quad P^{N_T+1} = 0; \quad (32)$$

thus

$$\left(I+\frac{\Delta T}{2}\left[A^\sigma\right]^T\right)P^n - \Delta T\frac{1}{2}\frac{V^n - V_{ij}}{\sqrt{\left\|V^n - V_{ij}\right\|^2+\varepsilon_1^2}} = \left(I-\frac{\Delta T}{2}\left[A^\sigma\right]^T\right)P^{n-1}; \quad P^{N_T+1} = 0, \quad (33)$$

Let $B^\sigma = (\Delta T/2)[A^\sigma]^T$; (33) can be written as

$$\left(I-B^\sigma\right)P^{n-1} = \left(I+B^\sigma\right)P^n - \frac{\Delta T}{2}\frac{V^n - V_{ij}}{\sqrt{\left\|V^n - V_{ij}\right\|^2+\varepsilon_1^2}}, \quad P^{N_T+1} = 0; \quad (34)$$

$B^\sigma = [b_{ij}^\sigma]$ is a tridiagonal matrix with nonzero elements:

$$b_{i,i-1}^\sigma = -\frac{\Delta T}{2}\left(\frac{1}{2}\sigma_{i-1}^2(i-1)^2 + (r-q)(i-1)\right),$$
$$b_{i,i}^\sigma = \frac{\Delta T}{2}\left[(\sigma_i i)^2 + (r-q)i + r\right], \quad (35)$$
$$b_{i,i+1}^\sigma = -\frac{\Delta T}{2}\left[\frac{1}{2}\sigma_{i+1}^2(i+1)^2\right], \quad i = 1,2,3,\ldots,N_S.$$

The discrete form of the gradient $\nabla_\sigma J(\sigma)$ is given by

$$\begin{aligned}\nabla_{\sigma_k} J(\sigma) &= -\alpha\frac{\sigma''\left((\sigma')^2+\varepsilon_2^2\right)-(\sigma')^2}{\left((\sigma')^2+\varepsilon_2^2\right)^{3/2}} - \int_0^T V^T\left[\frac{\partial A^\sigma}{\partial\sigma}\right]^T P dt \\ &= \Delta T\sum_{n=1}^{N_T}\sum_{i=1}^{N_S}\left[V_{i-1}^n + V_{i-1}^{n+1} - 2\left(V_i^n + V_i^{n+1}\right) + V_{i+1}^n + V_{i+1}^{n+1}\right]\sigma_i i^2 P_i^n \times (2)^{-1} \\ &\quad - \alpha\left(\frac{\sigma_{k+1}-2\sigma_k+\sigma_{k-1}}{\Delta S^2}\left[\left(\frac{\sigma_{k+1}-\sigma_k}{\Delta S}\right)^2+\varepsilon_2^2\right] - \left(\frac{\sigma_{k+1}-\sigma_k}{\Delta S}\right)^2\right) \\ &\quad \times\left(\left(\left(\frac{\sigma_{k+1}-\sigma_k}{\Delta S}\right)^2+\varepsilon_2^2\right)^{3/2}\right)^{-1},\end{aligned} \quad (36)$$

where $\nabla_\sigma J(\sigma) = (\nabla_{\sigma_1}J(\sigma), \nabla_{\sigma_2}J(\sigma), \ldots, \nabla_{\sigma_{N_S}}J(\sigma))$.

The solution of the minimization problem (11) could be computed by Newton's method:

$$\sigma^{(k+1)} = \sigma^{(k)} - \left[\nabla_\sigma^2 J\left(\sigma^{(k)}\right)\right]^{-1}\cdot\nabla_\sigma J\left(\sigma^{(k)}\right), \quad (37)$$

where the inverse Hessian $[\nabla_\sigma^2 J(\sigma^{(n)})]^{-1}$ is approximated by L-BFGS formula.

We first need to introduce some notations. The iterates will be denoted by σ_k and we define $s_k = \sigma_{k+1} - \sigma_k$, $g_k = \nabla_{\sigma_k}J(\sigma)$, $y_k = g_{k+1} - g_k$. The method uses the inverse BFGS formula in the form

$$H_{k+1} = v_k^T H_k v_k + \rho_k s_k s_k^T, \quad (38)$$

where $\rho_k = 1/y_k^T s_k$ and $v_k = 1 - \rho_k y_k s_k^T$.

Algorithm 1 (TV − L^1 model for solving the implied volatility).

Step 1. Choose a function $\sigma_0(S)$. This will be the initial approximation to the true volatility $\sigma_{ex}(S)$.

Step 2. Give the initialization value $\alpha, T, N_S, N_T, K, S_{max}, \varepsilon_1 = \varepsilon_2 = 10^{-6}$, $0 < \beta' < 1/2 < \beta'' < 1$, and $H_0 = I$.

Step 3. Determine $V_i^{N_T} = (i\Delta S - K)^+$ and V_{ij} by the Black-Scholes formula using $\sigma = \sigma_{ex}(S)$:

$$V(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2). \quad (39)$$

Step 4. Let $\sigma = \sigma_k$; P is the solution of the following linear equation:

$$(I - B^{\sigma}) P^{n-1} = (I + B^{\sigma}) P^{n} - \frac{\Delta T}{2} \frac{V^n - V_{ij}}{\sqrt{\left\| V^n - V_{ij} \right\|^2 + \varepsilon_1^2}}, \quad (40)$$

$$P^{N_T+1} = 0,$$

and $V^n (n = N_T, N_T - 1, \ldots, 2, 1)$ is computed by

$$\left(I + \frac{\Delta T}{2} A^{\sigma}\right) V^n = \left(I - \frac{\Delta T}{2}\right) V^{n+1}, \quad (41)$$

where A^{σ} and B^{σ} are defined in (19) and (35).

Step 5. Compute $g_k = \nabla_{\sigma} J(\sigma_k)$, $d_k = -H_k g_k$; then

$$\sigma_{k+1} = \sigma_k + \alpha_k d_k, \quad (42)$$

where α_k satisfies the Wolfe condition:

$$J(\sigma_k + \alpha_k d_k) \leq J(\sigma_k) + \beta' \alpha_k g_k^T d_k, \quad (43)$$

$$g(\sigma_k + \alpha_k d_k) \geq \beta g_k^T d_k. \quad (44)$$

We always try the step length $\alpha_k = 1$ first.

Step 6. If $\|\sigma_{k+1} - \sigma_{\text{ex}}\|_{\infty} \leq \tau$, end; else go to next step.

Step 7. Let $\widehat{m} = \min\{k, m-1\}$; update H_0 $\widehat{m} + 1$ times using the pairs $\{y_j, s_j\}_{j=k-\widehat{m}}^{k}$; that is, let

$$\begin{aligned} H_{k+1} &= \left(V_k^T \cdots V_{k-\widehat{m}}^T\right) H_0 \left(V_{k-\widehat{m}} \cdots V_k\right) \\ &\quad + \rho_{k-\widehat{m}} \left(V_k^T \cdots V_{k-\widehat{m}+1}^T\right) s_{k-\widehat{m}} \cdot s_{k-\widehat{m}}^T \left(V_{k-\widehat{m}+1} \cdots V_k\right) \\ &\quad + \rho_{k-\widehat{m}+1} \left(V_k^T \cdots V_{k-\widehat{m}+2}^T\right) s_{k-\widehat{m}+1} \\ &\quad \cdot s_{k-\widehat{m}+1}^T \left(V_{k-\widehat{m}+2} \cdots V_k\right) \\ &\quad \vdots \\ &\quad + \rho_k s_k s_k^T. \end{aligned} \quad (45)$$

Step 8. Set $k = k + 1$ and go to Step 4.

In this paper, we only discuss the estimation of implied volatility on European call options. The TV $-$ L^1 model and adjoint method are still valid in the case of put options.

5. Numerical Experiments

In this section, we present numerical experiments to illustrate the TV $-$ L^1 model and adjoint method presented in the previous sections. First, we assume that the true volatility function $\sigma_{\text{ex}}(S)$ is defined as

$$\sigma_{\text{ex}}(S) = \begin{cases} \sigma = 0.1e^{-0.01S} + 0.02, & S \in [0, 50]; \\ \sigma = 0.1e^{-0.01S} - 0.02, & S \in (50, 100]. \end{cases} \quad (46)$$

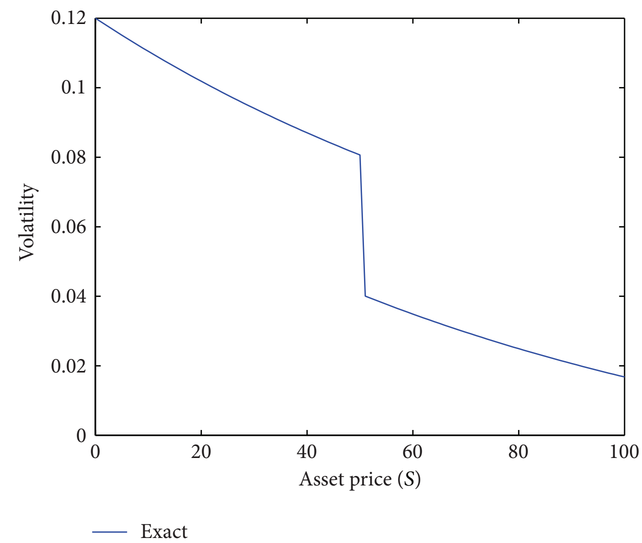

FIGURE 1: Volatility function $\sigma_{\text{ex}}(S, t)$.

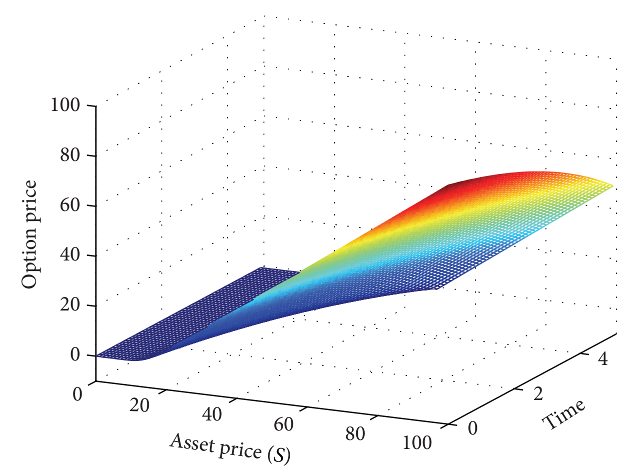

FIGURE 2: The observed market prices V_{ij}.

In numerical experiments, the interest rate $r = 0.25$, $S_{\max} = 100$; we consider only one time to option maturity $T = 5$. We take $\Delta S = 1, \Delta T = 0.01, N_S = 100, N_T = 500$, and $K = 50$. Figure 1 displays the true volatility function.

The observed market prices V_{ij} are obtained by solving the Black-Scholes equation with the true volatility. Figure 2 displays V_{ij}.

We solve the optimal volatility by Algorithm 1; Figure 3 shows the comparison between the true volatility $\sigma_{\text{ex}}(S)$ and the optimal estimated $\sigma(S)$,.

Our total variation regularization strategy has three advantages: the first one is it contains no terms involving the Dirac delta function [24] compared with Lagnado and Osher's model [18]; the second is that the total variation regularization can maintain the singularities of the solution better ($S = 50$); the third is that the gradient is computed via the adjoint method which provides us with an exact value of the gradient needed for the minimization.

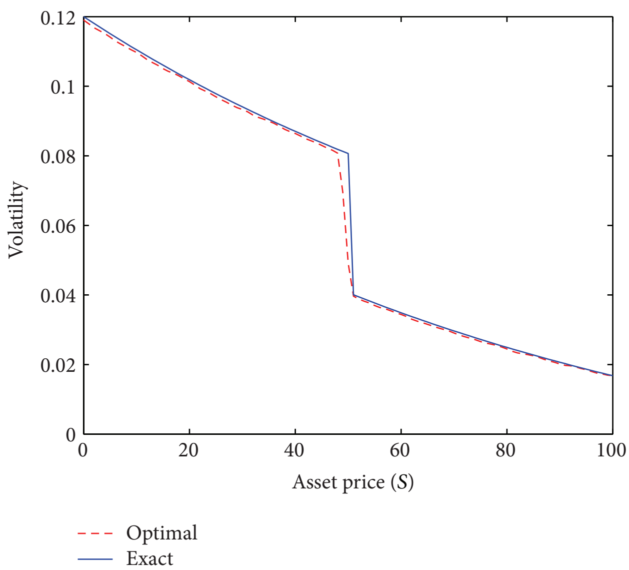

Figure 3: Volatility estimation.

6. Conclusion

A lot of research works have been made to determine the implied volatility by regularization strategies. Based on the advantages and great success of the total variation regularization strategy in image processing, in this paper, we propose the $TV - L^1$ model for solving the implied volatility under the framework of the Black-Scholes model. We estimate implied volatility by solving an optimal control problem and the gradient is computed via the adjoint method. We use the limited memory quasi-Newton algorithm (L-BFGS) to find the optimal solution. Furthermore, the results of numerical experiments are presented.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the NNSF of China (nos. 60872129 and 11271117) and the Science and Technology Project of Changsha City of China (no. K1207023-31).

References

[1] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, pp. 637–654, 1973.

[2] J. R. Franks and E. S. Schwartz, "The stochastic behavior of market variance implied in the price of index options," *The Economics Journal*, vol. 101, pp. 1460–1475, 1991.

[3] R. Heynen, "An empirical investigation of observed smile patterns," *Review Futures Markets*, vol. 13, pp. 317–353, 1994.

[4] B. Dupire, "Pricing with a smile," *Risk*, vol. 7, pp. 18–20, 1994.

[5] M. A. Bharadia, N. Christofides, and G. R. Salkin, "Computing the Black-Scholes implied volatility," *Advances in Futures and Options Research*, 8, pp. 15–29, 1996.

[6] D. M. Chance, "A generalized simple formula to compute the implied volatility," *The Financial Review*, vol. 31, pp. 859–867, 1996.

[7] D. R. Chambers and S. K. Nawalkha, "An improved approach to computing implied volatility," *The Financial Review*, vol. 38, pp. 89–100, 2001.

[8] S. Li, "A new formula for computing implied volatility," *Applied Mathematics and Computation*, vol. 170, no. 1, pp. 611–625, 2005.

[9] L. V. Ballestra and L. Cecere, "A numerical method to compute the volatility of the fractional Brownian motion implied by American options," *International Journal of Applied Mathematics*, vol. 26, no. 2, pp. 203–220, 2013.

[10] A. N. Tikhonov, A. S. Leonov, and A. G. Yagola, *Nonlinear Ill-Posed Problems*, Chapman & Hall, London, UK, 1998.

[11] C. Chiarella, M. Craddock, and N. El-Hassan, "The calibration of stock option pricing models using inverse problem methodology," QFRQ Research Papers, UTS, Sydney, Australia, 2000.

[12] S. Crépey, "Calibration of the local volatility in a trinomial tree using Tikhonov regularization," *Inverse Problems*, vol. 19, no. 1, pp. 91–127, 2003.

[13] Z. C. Deng, J. N. Yu, and L. Yang, "An inverse problem of determining the implied volatility in option pricing," *Journal of Mathematical Analysis and Applications*, vol. 340, no. 1, pp. 16–31, 2008.

[14] H. Egger and H. W. Engl, "Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates," *Inverse Problems*, vol. 21, no. 3, pp. 1027–1045, 2005.

[15] V. Isakov, "The inverse problem of option pricing," in *Recent Development in Theories and Numerics*, pp. 47–55, World Scientific Publishing, Singapore, 2003.

[16] L. S. Jiang and Y. S. Tao, "Identifying the volatility of underlying assets from option prices," *Inverse Problems*, vol. 17, no. 1, pp. 137–155, 2001.

[17] H. E. Leland, "Option pricing and replication with transaction costs," *The Journal of Finance*, vol. 40, pp. 1283–1301, 1985.

[18] R. Lagnado and S. Osher, "A technique for calibrating derivative security pricing models: numerical solution of an inverse problem," *Journal of Computational Finance*, vol. 1, pp. 13–25, 1997.

[19] P. Ngnepieba, "The adjoint method formulation for an inverse problem in the generalized Black-Scholes model," *Journal of Systemics Cybernetics and Informatics*, vol. 4, pp. 69–77, 2006.

[20] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.

[21] M. J. Boes, F. C. Drost, and B. J. M. Werker, "The impact of overnight periods on option pricing," *Journal of Financial and Quantitative Analysis*, vol. 42, no. 2, pp. 517–534, 2007.

[22] I. Ishida, M. McAleer, and K. Oya, "Estimating the leverage parameter of continuous-time stochastic volatility models using high frequency S&P500 and VIX," *Managerial Finance*, vol. 37, pp. 1048–1067, 2011.

[23] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, vol. 45, no. 3, pp. 503–528, 1989.

[24] G. E. Andrews, R. Askey, and R. Roy, *Special Functions*, vol. 71, Cambridge University Press, Cambridge, UK, 1999.