The Adjoint Method Formulation for an Inverse Problem in the Generalized Black-Scholes Model

Pierre NGNEPIEBA
Department of Mathematics, Florida A&M University
Tallahassee, Florida 32307, USA

Abstract

A general framework is developed to treat optimal control problems for a generalized Black-Scholes model, which is used for option pricing. The volatility function is retrieved from a set of market observations.

The optimal volatility function is found by minimizing the cost functional measuring the discrepancy between the model solution (pricing) and the observed market price, via the unconstrained minimization algorithm of the quasi-Newton limited memory type. The gradient is computed via the adjoint method. The effectiveness of the method is demonstrated on an European call option.

Key words: Generalized Black-Scholes model, adjoint method, volatility function, gradient, options.

1 Introduction and Problem Statement

An option is a contract that gives the owner the right to buy or sell a specified amount of a particular underlying asset at a fixed price, called the exercise (strike) price on or before a specified date, called the maturity date. Options are generally either American or European a combination of both. American options can be exercised at any time up to expiry date, whereas the European options can be exercised only at the expiry date.

The present work focuses on the optimal control problem for data assimilation with the aim of ascertaining the optimal volatility in the generalized Black-Scholes equation using a set of market observation.

Consider the following generalized Black-

Scholes model:

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2(s)s^2\frac{\partial^2 V}{\partial s^2}+(r-q)\,s\frac{\partial V}{\partial s}-rV=0, \tag{1.1}$$

where $(s,t) \in I\!R^{+}\times(0,T)$, $V = V(s,t)$ is the value of the option price at asset price s, $\sigma = \sigma(s)$ is the volatility function of underlying asset s, r is the risk-free interest rate, q is the constant dividend yield, t is the current time, and T is the maturity date. When the volatility σ is a constant function, it becomes the famous model for the price of options, the classical Black-Scholes model [1].

Eq. (1.1) is a parabolic partial differential equation. To make it well-posed, we have to specify the initial or end condition, the payoff function at maturity and the boundary conditions at zero and at infinity. Here, we are concerned on the valuation of the vanilla European call options since the put options are almost identical mathematically, and no exact solution exists for the American options. Therefore, the payoff function at maturity and boundary conditions are given by:

$$\begin{cases} V(s,T)=(s-K)^{+}=\mathrm{Max}\,(0,s-K)\,,\ s\in I\!R^{+}\\ V(0,t)=0,\ t\in(0,T)\\ \lim\limits_{s\to\infty}V(s,t)=\text{payoff},\ t\in(0,T)\end{cases} \tag{1.2}$$

where K is the exercise or strike price. Eq. (1.1) is described in an infinite domain $I\!R^{+}\times(0,T)$, which makes it difficult in constructing numerical solutions. This motivates the consideration of the following model defined on a truncated domain $\mathcal{D} = (0,S_{max})\times(0,T)$, where S_{max} is the suitable chosen positive number representing the final value of the asset price. Thus, Eq. (1.1) and boundary conditions (1.2) become

$$\begin{cases} \dfrac{\partial V}{\partial t}+\dfrac{1}{2}\sigma^2(s)s^2\dfrac{\partial^2 V}{\partial s^2}+(r-q)\,s\dfrac{\partial V}{\partial s}-rV=0\\ \\ V(s,T)=(s-K)^{+},\ \ s\in[0,S_{max}]\\ V(0,t)=0,\ t\in[0,T],\ \ (s,t)\in\mathcal{D}\\ V(S_{max},t)=(S_{max}-K)^{+}_{;}\ \ t\in(0,T)\end{cases} \tag{1.3}$$

The existence and uniqueness of a solution of the Eq. (1.3) can be found in [2].

The generalized Black-Scholes model given by Eq. (1.3) consists of five parameters K, T, r, q, and $\sigma(s)$. The constant parameters T, K, q and r are assumed or can be directly observed in the market, whereas the volatility function $\sigma(s)$ cannot be directly obtained from the market.

Suppose we are given a set of $\tilde{m}\tilde{n}$ market observations values of the option, $\gamma_{ij\,1\le i\le\tilde{m},\,1\le j\le\tilde{n}}$, where γ_{ij} is the observed market price of the option with exercise price K_i at the exercise time T_j, we can estimate the volatility function $\sigma(s)$ using the generalized Black-Scholes model (1.3) and this set of the observations. We denote by $\tilde{\mathbf{V}}$ the observational vector constructed using all the options price

γ_{ij} for the maturity T_j and the exercise K_i, $i = 1, 2, \cdots, \tilde{m},\ j = 1, 2, \ldots, \tilde{n}$.

The cost function, $J(\sigma) : \mathcal{V}_{ad} \longrightarrow I\!R^{+}$, where $\mathcal{V}_{ad}$ is the set of admissible parameters $\sigma(s)$ which defines the discrepancy between the simulated values, $V(s,t)$, and the corresponding observations, $\tilde{\mathbf{V}}$, is defined as follows:

$$J(\sigma) = \frac{1}{2}\int_0^T \left\| V - \tilde{\mathbf{V}} \right\|^2 dt + \frac{\alpha}{2}\left\| \sigma - \sigma_{ex} \right\|^2 \quad (1.4)$$

where $\alpha \geq 0$ is the regularization coefficient and σ_{ex} is an *a priori* estimation of σ.

We then defined the optimal control problem as follows:

$$\begin{cases} \text{Find } V \text{ and } \sigma^{\star} \text{ such that} \\ \\ J(\sigma^{\star}) = \inf\limits_{\sigma \in \mathcal{V}_{ad}} J(\sigma). \end{cases} \quad (1.5)$$

If J has a minimum (assuming that the set of all admissible parameters $\mathcal{V}_{\mathrm{ad}}$ is the whole set) then the optimality condition $\nabla J(\sigma^{\star}) = 0$ holds. The gradient of the cost function J is obtained by using an adjoint equation (see Le Dimet and Talagrand [4], Lions [3]).

The key idea is that the adjoint method provides us with an exact value of the gradient of the cost function needed for the minimization procedure. The main difficulty in implementing the adjoint technique is the derivation of the adjoint equations. Once this exact value of the gradient of the cost function is obtained via the adjoint equation, the unconstrained minimization algorithm of the quasi-Newton limited memory type [5] is used to derive the optimal volatility function. The inverse model allows us to determine the optimal scale parameters and the model sensitivity.

2 Solution Technique

The gradient is the key element in the minimization procedure that requires computing the gradient of the cost function. It is obtained by the adjoint model, which we briefly describe in section 2.3 using the semi-discrete equation for the sake of completeness.

2.1 Discretization of the Black-Scholes equation

In the following, we consider the discretization of the Black-Scholes partial differential Eq. (1.3). We apply a uniform grid for the computational domain $[0, S_{max}] \times [0, T]$ which is formed with the space step $\Delta s = \dfrac{S_{max}}{N_s}$ and time step $\Delta T = \dfrac{T}{N_T}$. Moreover, we use the notation

$$V_i^n = V(s_i, t_n), \quad (2.6)$$

where $s_i = i\Delta s$ and $t_n = n\Delta T$; $i = 0, \cdots, N_s$, $n = 0, \cdots, N_T$, for the numerical approximation of the solution. The efficiency of numerical solution can be improved by coordinate transformation or using nonuniform grids [6].

2.2 Space Discretization

For the space discretization, a combination of first-order and second-order accurate finite differences

$$\begin{cases} \dfrac{\partial V}{\partial s}(s_i, t) \approx \dfrac{V_{i+1}(t) - V_i(t)}{\Delta s} \\ \\ \dfrac{\partial^2 V}{\partial s^2}(s_i, t) \approx \dfrac{V_{i+1}(t) - 2V_i(t) + V_{i-1}(t)}{\Delta s^2} \end{cases} \tag{2.7}$$

is used for the approximation of the space derivative in the Black-Scholes partial differential equation. This lead to the following semi-discrete equation

$$\begin{aligned} \frac{\partial V}{\partial t} &+ \left[\frac{1}{2}(\sigma_i\, i)^2\right] V_{i-1} - \left[(\sigma_i\, i)^2 + (r-q)\, i + r\right] V_i \\ &+ \left[\frac{1}{2}(\sigma_i\, i)^2 + (r-q)\, i\right] V_{i+1} = 0 \end{aligned} \tag{2.8}$$

where $i = 1, 2, \cdots, N_s$, and $\sigma_i = \sigma(s_i)$. This equation can be written as follows

$$\frac{dV}{dt} = \mathbf{A}^\sigma V, \tag{2.9}$$

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

$$\begin{cases} a_{i,i-1}^\sigma = -\dfrac{1}{2}(\sigma_i\, i)^2, \ i = 1, \cdots, N_s \\ a_{i,i}^\sigma = \left((\sigma_i\, i)^2 + (r-q)\, i + r\right), \ i = 1, \cdots, N_s \\ a_{i,i+1}^\sigma = -\left(\dfrac{1}{2}(\sigma_i\, i)^2 + (r-q)\, i\right), \ i = 1, \cdots, N_s \end{cases}$$

With this discretization, $\mathbf{A}^\sigma$ has an M-matrix property. Indeed, $a_{i,j}^\sigma \leq 0$ for $i \neq j$. This property guarantees that the space discretization doest not cause undesired oscillations into the numerical solution. (see [6]). The time discretization of the semi-discrete Eq. (2.9) is discussed in the section 2.4.

2.3 Adjoint Model

In this section, we derive the adjoint equation to the semi-discretized equation (2.9).

Using the definition of the directional derivative (also called *sensitivity* in financial theory context),

$$\widehat{V}(\sigma, h) = \lim_{\beta \to 0} \frac{V(\sigma + \beta h) - V(\sigma)}{\alpha} \tag{2.10}$$

(where h is the perturbation on σ) to Eq. (2.9) gives rise to the tangent linear system,

$$\begin{cases} \dfrac{d\widehat{V}}{dt} &= \left[\dfrac{\partial \mathbf{A}^\sigma}{\partial V}\right] . \widehat{V} + \left[\dfrac{\partial \mathbf{A}^\sigma}{\partial \sigma}\right] . h \\ \\ \widehat{V}(0) &= 0, \end{cases} \tag{2.11}$$

which is used to find the adjoint model. Similarly, the directional derivative of the cost function is

$$\begin{aligned} \widehat{J}(\sigma, h) &= \lim_{\alpha \to 0} \frac{J(\sigma + \beta h) - J(\sigma)}{\beta} = \langle h, \nabla_\sigma J \rangle \\ &= \int_0^T \left\langle V - \tilde{\mathbf{V}}, \widehat{V} \right\rangle dt + \\ &\quad \alpha \langle \sigma - \sigma_{ex}, h \rangle \end{aligned} \tag{2.12}$$

Introducing an adjoint variable P, the scalar product of P and the tangent linear system (2.11) is integrated between 0 and T to yield (using integration by parts) the adjoint model,

$$\begin{cases} \left\langle P(T), \widehat{V}(T) \right\rangle - \left\langle P(0), \widehat{V}(0) \right\rangle \\ -\displaystyle\int_0^T \left\langle \frac{dP}{dt}, \widehat{V} \right\rangle dt = \int_0^T \left\langle \left[\frac{\partial \mathbf{A}^\sigma}{\partial \sigma}\right]^T P, \widehat{V} \right\rangle dt \\ \quad + \left\langle h, \displaystyle\int_0^T \left[\frac{\partial \mathbf{A}^\sigma}{\partial \sigma}\right]^T P dt \right\rangle \end{cases} \tag{2.13}$$

Eq. (2.13) can be rewritten

$$-\left\langle P\left(T\right),\widehat{V}\left(T\right)\right\rangle+\int_0^T\left\langle\frac{dP}{dt}+\left[\frac{\partial\mathbf{A}^\sigma}{\partial V}\right]^T P,\widehat{V}\right\rangle dt$$
$$=\left\langle h,-\int_0^T\left[\frac{\partial\mathbf{A}^\sigma}{\partial\sigma}\right]^T Pdt\right\rangle$$

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

$$\begin{cases}\dfrac{dP}{dt}+\left[\dfrac{\partial\mathbf{A}^\sigma}{\partial V}\right]^T P=\left(V-\tilde{\mathbf{V}}\right)\\[2ex] P\left(T\right)=0\end{cases}\tag{2.14}$$

Then, the directional derivative of the cost function defined in (2.12) can be written as

$$\begin{aligned}\tilde{J}\left(h,\sigma\right)=\left\langle h,\nabla_\sigma J\right\rangle &= \left\langle h,-\int_0^T\left[\frac{\partial\mathbf{A}^\sigma}{\partial\sigma}\right]^T Pdt\right\rangle\\ &+ \alpha\left\langle\sigma-\sigma_{ex},h\right\rangle\end{aligned}$$

wherefrom we obtain

$$\nabla_\sigma J=\alpha\left(\sigma-\sigma_{ex}\right)-\int_0^T\left[\frac{\partial\mathbf{A}^\sigma}{\partial\sigma}\right]^T Pdt\tag{2.15}$$

Thus, to obtain the gradient of the cost function with respect to the control variable, σ, the optimality system (2.9, 2.14, 2.15) is solved simultaneously.

The solution of the minimization problem defined in section 1 may be found by using the Newton's method:

$$\sigma_{n+1}=\sigma_n-\left[\nabla_\sigma^2 J\left(\sigma_n\right)\right]^{-1}.\nabla_\sigma J\left(\sigma_n\right)\tag{2.16}$$

where σ_n is the current estimate.

To solve the minimization problem, the limited-memory quasi-Newton minimization algorithm is used [5]. It is based on the BFGS (Broyden-Fletcher-Goldfarb-Shanno) update formula for the inverse Hessian $\left[\nabla_\sigma^2 J\left(\sigma_n\right)\right]^{-1}$.

2.4 Time Discretization

In the option pricing problems, the stability of the time discretization scheme is an important issue because of the nonsmooth of initial data. The stability of time discretization schemes for the parabolic partial differential equation is considered in [7].

For the time discretization of the semi-discrete Eq. (2.9), we consider the Crank-Nicolson scheme. The Crank-Nicolson time discretization scheme can be interpreted as the average of the explicit and implicit Euler schemes. The Scheme

$$\frac{V^{n+1}-V^n}{\Delta T} = \mathbf{A}^\sigma\left(\frac{V^{n+1}+V^n}{2}\right)\tag{2.17}$$

is second-order accurate and unconditionally stable.

Eq. (2.17) lead to the matrix solution

$$\left(\mathbf{I}+\frac{\Delta T}{2}\mathbf{A}^\sigma\right)V^n=\left(\mathbf{I}-\frac{\Delta T}{2}\mathbf{A}^\sigma\right)V^{n+1}\tag{2.18}$$

Discrete Version of the Adjoint

The complete discrete adjoint equation associated to this time scheme discretization with the boundary condition $P_0^n = 0$ and $P_{N_s}^n = 0$, $n = 0, \cdots, N_T$ is defined by solving backward the following linear system

$$\begin{cases} \left(\mathbf{I} - \mathbf{B}^\sigma\right) P^{n-1} = \left(\mathbf{I} + \mathbf{B}^\sigma\right) P^n - \left(V^n - \tilde{\mathbf{V}}\right) \\ \\ P^{N_T+1} = 0, \quad n = N_T, \cdots, 1 \end{cases} \tag{2.19}$$

where $\mathbf{B}^\sigma = \left[b_{ij}^\sigma\right]$ is a tridiagonal matrix with nonzero elements

$$\begin{cases} b_{i,i-1}^\sigma = -\Delta T \left(\dfrac{1}{2}(\sigma_{i-1}\,(i-1))^2 + (r-q)\,(i-1)\right)/2 \\ b_{i,i}^\sigma = \Delta T \left((\sigma_i\, i)^2 + (r-q)\, i + r\right)/2 \\ b_{i,i+1}^\sigma = -\Delta T \dfrac{1}{2}(\sigma_{i+1}\,(i+1))^2/2 \\ \qquad i = 1, \cdots, N_s - 1 \end{cases}$$

Discrete version of the Gradient

The gradient of the cost function with respect to the control parameter, $\sigma_k = \sigma(s_k)$, $k = 1, 2, \cdots, N_s$; is given by

$$\begin{aligned} \nabla_{\sigma_k} J \;&= \Delta t \sum_{n=1}^{N_T} \sum_{i=1}^{N_s} \left[V_{i-1}^n + V_{i-1}^{n+1} - 2\left(V_i^n + V_i^{n+1}\right)\right. \\ &+ \left. V_{i+1}^n + V_{i+1}^{n+1}\right] \sigma_i i^2 p_i^n/2 \\ &+ \alpha\left(\sigma_k - (\sigma_{ex})_k\right) \end{aligned} \tag{2.20}$$

Remark

If we consider the volatity as a function of time and the asset price, $\sigma = \sigma(s,t)$, the problem of estimating the volatility surface $\sigma(s,t)$ becomes more complex. However, using a wavelets analysis, the additional time dimension can be easily mitigated [8].

The basic idea behind the wavelet analysis is to decompose a time dependent function into a number of component, each one of which can be associated with a particular scale at a particular time.

In short, a wavelet $\varphi \in L^2(I\!R)$ is function whose binary dilations and dyadic tralations generated a Riesz basis on $L^2(I\!R)$. Any $f \in L^2(I\!R)$ can be expanded into a wavelet series,

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} w_{jk}\varphi_{jk}(t)$$

where $\varphi_{jk} \in L^2(I\!R)$ denotes the dilated and translated wavelet, defined by

$$\varphi_{jk} = 2^{j/2}\varphi\left(2^j t - k\right)$$

$j \in \mathcal{Z}$ is the scale of the wavelet, corresponding to a dilation by 2^j and $k \in \mathcal{Z}$ is the position (translation). In the present setting we can use this expansion to represent the volatility coefficient $\sigma(s,t)$ in a Wavelets basis.

3 Numerical Experiments

In this section, we present numerical experiments to illustrate the theory presented in the previous sections. All numerical computations were executed on a HP PC (AMD, 1,8 GHZ, 512 MB).

In these numerical experiments, the European call option problem (1.3) is described by the parameter values $S_{max} = 100$, $T = 5$, $K = 50$, $N_s = 100$, $N_T = 50$, and $\alpha = 0.5$. The payoff function is given by $V_i^{N_T} = (i\Delta s - K, 0)^+$, $i = 1, 2, \cdots, N_s$. The present methodology is applied to two examples.

Example 1

First, we assume that the true volatility parameter, $\sigma_{ex}(s)$, is defined as $\sigma_{ex}(s) = 0.01e^{-(0.01s)}$. This exact volatility is used to solve generalized Black-Scholes equation (1.3) with $r = 0.35$ and $q = 0.3$. The solution obtained serves as the observed market prices $\tilde{\mathbf{V}}$. Figure 1 displays $\tilde{\mathbf{V}}$.

The unconstrained minimization algorithm of the quasi-Newton limited memory type [5] with the convergence criterion either on the number of iterations or the gradient norm of the cost function is used to determine the optimal volatility.

Figure 2 shows the evolution of the (relative) gradient norm of the cost function. The optimal volatility is recovered in 12 iterations.

Figure 3 shows the comparison between the true volatility (continuous line) and the estimated volatility, $\sigma^\star(s)$, (dashed line) obtained by solving (1.5). It can be seen that the agree-

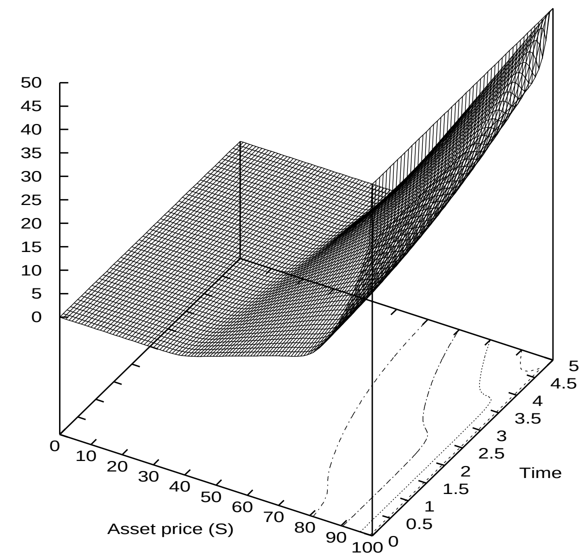

Figure 1: option values $\tilde{\mathbf{V}}$ for $\sigma_{ex}(s) = 0.01e^{-(0.01s)}$

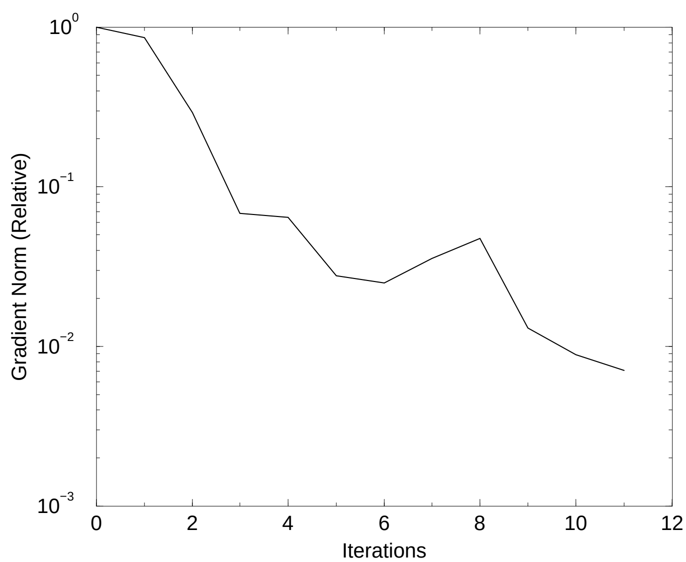

Figure 2: Relative gradient norm of the cost function: $\dfrac{\|\nabla J_k\|}{\|\nabla J_0\|}$ (the gradient at each iteration is divided by the gradient at the first iteration)

ment is excellent. We observed the discrepancies near the maturity time, $T = 100$. In other words, the mismatch occures for large values of the underling asset.

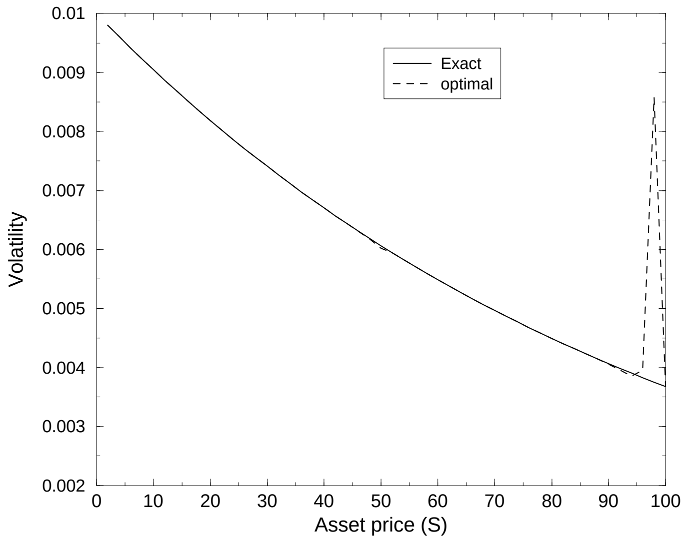

Figure 3: Volatility estimation: example 1

recovered. Once more the discrepancies occur at the vicinity of the maturity date.

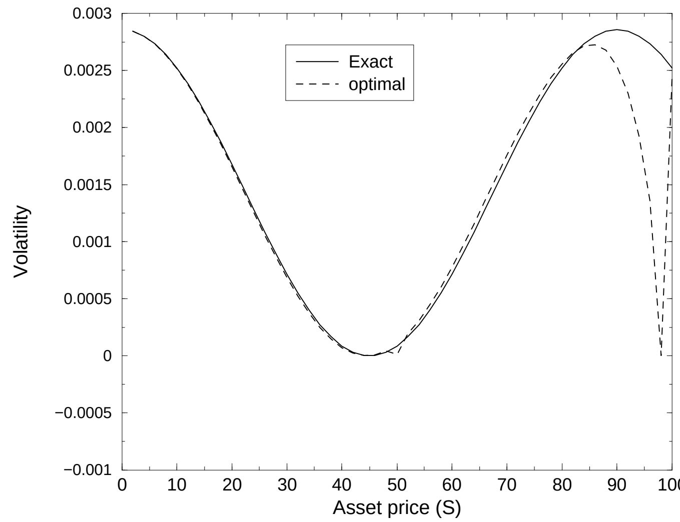

Figure 4: Volatility estimation: Example 2

Example 2

In this example, we assume that the true volatility function, σ_{ex}, is given by

$$\sigma_{ex}(s) = \frac{10^{-2}\left(1 + \cos\left(\frac{\pi s}{45}\right)\right)}{7}$$

The observation is the solution of the generalized Black-Scholes model (1.3) using this true volatility with $r = 0.01$ and $q = 0$.

Figure 4 shows the comparison between the true volatility (solid line) and the optimal one (dashed line). As in the example 1, using the true volatility with more structures, we observed that the optimal volatility $\sigma^{\star}(s)$ is well recovered. Once more the discrepancies occur at the vicinity of the maturity date.

4 Conclusions

A generalized Black-Scholes equation is considered as a mathematical model for the evaluation of the European options. We developed a scheme based on the adjoint method to calibrate the volatility parameter as the function of the asset price. The resulting ill-posed inverse problem is regularized by penalizing the cost functional. The Numerical experiments carried out support the theoretical result. For the future work, It is remained to calibrate the volatility surface as the function of the asset price and time to maturity; and to extend the study to the American call/put options.

References

[1] F. Black and M. Scholes, “The pricing of options and corporate liabilities,” **Journal of Political Economy**, Vol 81, 1973, pp.637-654.

[2] Q. A. Ladyzenskaja, V. A. Solonnikov and N. N. Uralceva, “Linear and Quasilinear equations of Parabolic Type, ” **AMS**, Providence, 1968.

[3] J. L. Lions, **Optimal control of systems governed by partial differential equations**, Berlin, New York, Springer-Verlag, 1971 (Translated by S. K. Mitter).

[4] F. X. Le Dimet and O. Talagrand, “Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects,” **Tellus**, Vol 38A, 1986, pp.97-110.

[5] J. C. Gilbert and C. Lemarechal, “Some numerical experiments with variable-storage quasi-Newton algorithms,” **Math. Program.**, 45, 1989, pp. 407-435.

[6] N. Clarke and K. Parrott, “Multigrid for American pricing with a stochastic volatility,” **Applied Mathematical Finance**, 6, 1999, pp 175-195.

[7] J. R. Cash, “The New finite difference schemes for parabolic equation,” *Journal of Numerical Analysis*, 21, 1984, pp 433-446.

[8] G. Erlebacher, MY Hussaini, and Leland Jamenson, **Wavelets -Theory and applications**, Oxford University Press, 1996.

[9] J. C. Hull, **Option Futures and other Derivatives** Precentice-Hall, Englewood Cliffs, NJ, 1997.

[10] I. Bouchouev and V. Isakov, “The inverse problem of option pricing,” **Inverse Problems.**, 13, 1997, pp. L11-L17.

[11] S. Crepey, “Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization,” **SIAM J. Math. Anal.**, 34, 2003, pp. 1183-1206.

[12] A. Tarantola, **Inverse Problem Theory: Methods for Data Fitting and Model Parameters Estimation**, Elsevier, 1987.