Variational and Topological Methods: Theory, Applications, Numerical Simulations, and Open Problems (2012). *Electronic Journal of Differential Equations*, Conference 21 (2014), pp. 183–195. ISSN: 1072-6691. http://ejde.math.txstate.edu, http://ejde.math.unt.edu ftp ejde.math.txstate.edu

THE INVERSE VOLATILITY PROBLEM FOR EUROPEAN OPTIONS

IAN KNOWLES, LI FENG, AJAY MAHATO

Abstract. The problem of determining equity volatility from a knowledge of European call option prices for a range of exercise (strike) prices and expirations is solved by minimization of a convex functional.

1. Introduction

The inner workings of financial markets, from a modeling perspective, are still not well understood, despite more than a century of effort dating back to the pioneering work of Bachelier [2]. Modern physics and its associated PDE modeling is supported by the laws of physics, which have withstood the test of time over centuries. Not so the "laws of finance", which appear quite flimsy in comparison. What we do know is that a market is a large collection of people acting individually and collectively, each with their own goals and economic reasons for participating. We also know that in transactions associated with future-oriented instruments, such as stock options and other financial derivatives, a huge amount of data is available buried inside of which is the market's best guess as to what the future holds. We are concerned here with the possibility of extracting information from this type of data with the aid of certain computational inverse algorithms.

It is common to model a financial asset (such as a stock or a commodity) via a stochastic differential equation

$$\frac{dS_t}{S_t} = m(S_t, t)dt + \sigma(S_t, t)dB_t, \tag{1.1}$$

where, for each time t, $S_t(\omega)$ is a random variable representing the price of the financial asset for the trial ω, m is the drift, which relates to the "trend" of the asset, σ is the volatility ("wobble"), and $B_t(\omega)$ is the Brownian motion stochastic process used to model the randomness. Financial derivatives are contracts that derive their value from such an underlying asset. In particular, a *European call option* on a stock is the right to buy one share of the stock at a specified price K (the strike, or exercise, price) at a specified future time T (expiration date). In their Nobel Prize winning paper [3] Black and Scholes showed that, under certain rather

2000 *Mathematics Subject Classification.* 34B24, 65L09, 45J40.
Key words and phrases. Inverse volatility; European option; Dupire equation; convex functional.

Published February 10, 2014.

severe restrictions, the arbitrage-free price of a European call contract, $v(S,t)$, satisfies a deterministic PDE of diffusion type in time t and the value S of the underlying asset:

$$\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 v}{\partial S^2} + \mu s \frac{\partial v}{\partial S} - rv = 0, \tag{1.2}$$

where σ is the (assumed constant) volatility, μ is the risk-neutral drift, and r is the short-term interest rate. For practical purposes the latter may be taken to be the interest rate on a 13-week US government treasury bill.

Many authors over the years [9, 17, 19, 20, 22, 23, 24] have noted that several of the assumptions laid down by Black and Scholes are basically incompatible with market data. Notable among these are objections to the constancy of σ. As recently as the early nineteen eighties, assertions such as "the Black-Scholes volatility is constant" seemed to hold true, at least while the market believed it so; but then, after the crash of 1987, volatility was anything but constant, and in fact it has recently become fashionable to speak of (and invest in) the volatility of the volatility! So it is common to regard the volatility as a function of S and t, $\sigma = \sigma(S,t)$, and we assume this in the sequel.

Now, financial data specifying the market price v of an option is readily available in quantity at various strike values K around the current price of the underlying asset (the "spot" price), and for values of the expiration T up to around six months into the future. Given that the computer projections currently used by most stock analysts are only valid for a week or so into the future, one is led quite naturally to the so-called *inverse volatility problem*: determine a market-inspired estimate of the future volatility function $\sigma(S,t)$ from a knowledge of current market prices v of options with different strikes and future expirations.

The solution of this problem generally goes as follows. The value v of an option contract also depends on the exercise (strike) price, K, and the expiration date, T, of the contract. In 1994 Bruno Dupire [8] noticed that the function $v(S,t;K,T)$ satisfies the "dual" Black-Scholes equation

$$\frac{\partial v}{\partial T} - \frac{1}{2}K^2\sigma(K,T)\frac{\partial^2 v}{\partial K^2} + \mu K \frac{\partial v}{\partial K} - (\mu - r)v = 0, \tag{1.3}$$

known also as the Dupire equation. If v is known for all strikes K and expirations T then, as was noted first in [8], the volatility is uniquely determined in principle from the equation (1.3). But such a formula for σ is of little use in practice, as the market data for v is not only noisy (which would make the estimation of these derivatives highly ill-posed), but even worse, the data is both discrete in T and somewhat sparse in K. A number of alternate approaches have been proposed subsequent to the appearance of [8], none of which has offered a definitive solution. Minimization methods using regularized least-squares fitting have been proposed in [1, 4, 18]; the possible presence of spurious local minima is always an issue here. An integral equation approach is presented in [5, 6], where convergence problems are possible given the underlying ill-posedness, and in [7, 11] linearization of the inherently non-linear inverse problem is discussed.

In this article we present a new variational algorithm for computing, via the Dupire equation (1.3), the volatility $\sigma(K,T)$ from a knowledge of European option prices at various strikes and expirations. The method used is an adaption of the variational approach involving the minimization of convex functionals (with the associated distinct advantage of having unique global minima and stationary points) used in [16] for numerical differentiation (formulated as an inverse problem), and in [14, 15] for solving the inverse groundwater modeling problem.

2. Reconstruction of volatility

For simplicity we assume that there is no dividend for the underlying asset. Thus the risk-neutral drift μ in (1.2) and (1.3) is equal to the interest rate r, and the Dupire equation (1.3) can be written as

$$\frac{\partial v}{\partial T} - \frac{1}{2}K^2\sigma^2(K,T)\frac{\partial^2 v}{\partial K^2} + rK\frac{\partial v}{\partial K} = 0. \tag{2.1}$$

Let $T_0 < T_1 < \dots < T_n$ be expiration times, and for each expiration T_i, $0 \le i \le n$, let $K_{i1}, \dots, K_{im_i}$ be the associated strike prices. We assume that the volatility is piecewise constant in time, so that, for $1 \le i \le n$, $\sigma = \sigma_i(K)$ over the i-th time sub-interval $[T_{i-1}, T_i]$. Fixing i, set

$$w_\lambda(K) = \int_{T_{i-1}}^{T_i} e^{-\lambda T} v(K,T)\,dT, \tag{2.2}$$

where $\lambda > 0$ is a parameter. For each such fixed i, $1 \le i \le n$, we now Laplace transform the Dupire equation over $[T_{i-1}, T_i]$ to obtain

$$\int_{T_{i-1}}^{T_i} e^{-\lambda T} v_T\,dT - \frac{1}{2}K^2\sigma_i^2 \underbrace{\int_{T_{i-1}}^{T_i} e^{-\lambda T} v_{KK}\,dT}_{w''_\lambda} + rK \underbrace{\int_{T_{i-1}}^{T_i} e^{-\lambda T} v_K\,dT}_{w'_\lambda} = 0,$$

where the primes indicate differentiation with respect to K. On integrating the first term by parts we get

$$[e^{-\lambda T} v]_{T_{i-1}}^{T_i} + \lambda \underbrace{\int_{T_{i-1}}^{T_i} e^{-\lambda T} v\,dT}_{w_\lambda} - \frac{1}{2}K^2\sigma_i^2 w''_\lambda + rKw'_\lambda = 0,$$

and rearranging terms gives,

$$-\frac{1}{2}K^2\sigma_i^2 w''_\lambda + rKw'_\lambda + \lambda w_\lambda = -v(K,T_i)e^{-\lambda T_i} + v(K,T_{i-1})e^{-\lambda T_{i-1}}.$$

Next, dividing by $\frac{1}{2}K^2\sigma_i^2$ throughout, we obtain

$$-(w''_\lambda - \frac{2r}{K\sigma_i^2}w'_\lambda) + \frac{\lambda}{\frac{1}{2}K^2\sigma_i^2}w_\lambda = \frac{-v(K,T_i)e^{-\lambda T_i} + v(K,T_{i-1})e^{-\lambda T_{i-1}}}{\frac{1}{2}K^2\sigma_i^2}.$$

Finally, on multiplying by the integrating factor

$$P(K) = e^{-2r\int^K \frac{dk}{k\sigma_i^2(k)}}, \tag{2.3}$$

we now have an equation in Sturm-Liouville form:

$$-(P(K)w'_\lambda)' + \lambda Q(K)w_\lambda = \beta(K,\lambda)Q(K), \tag{2.4}$$

where

$$Q(K) = (\frac{2}{K^2\sigma_i^2(K)})P(K), \tag{2.5}$$

$$\beta(K,\lambda) = -v(K,T_i)e^{-\lambda T_i} + v(K,T_{i-1})e^{-\lambda T_{i-1}}. \tag{2.6}$$

If we can recover the functions $P(K)$ and $Q(K)$ for each i, $1 \le i \le n$, we can find the volatility $\sigma_i(K)$ from the formula

$$\sigma_i(K) = \sqrt{\frac{2P(K)}{K^2 Q(K)}}. \tag{2.7}$$

We now focus attention on a variational approach to the recovery of one such pair of positive coefficient functions P, Q defined on an interval $a \le K \le b$. It is assumed that we are given the functions $w_\lambda(K)$ for K in $[a, b]$ and all $\lambda > 0$. For positive functions p and q also defined on $[a, b]$, let $c = (p, q)$. Define $w_{\lambda,c}(K)$ to be the solution to the boundary value problem

$$L_{p,\lambda q} w_{\lambda,c} = -(p(K) w'_{\lambda,c})' + \lambda q(K) w_{\lambda,c} = \beta(K,\lambda) q(K), \tag{2.8}$$

$$w_{\lambda,c}(a) = w_\lambda(a), \quad w_{\lambda,c}(b) = w_\lambda(b). \tag{2.9}$$

Let $\mathcal{D}$ be the set of all positive function pairs $c = (p, q)$ such that boundary value problem (2.8), (2.9) is *disconjugate* on [a,b], i.e. every non-trivial solution has at most one zero on [a,b]. It is known [10, Theorem 6.1, p. 351] that (2.8) is disconjugate if and only if the boundary value problem (2.8), (2.9) can always be solved uniquely. It is also known (c.f. [16, Proposition 2.1]) that this set is open and convex in $\mathcal{L}[a, b] \times \mathcal{L}[a, b]$ and $\mathcal{L}^2[a, b] \times \mathcal{L}^2[a, b]$. For each $\lambda > 0$ define the functional G_λ on the convex set $\mathcal{D}$ by

$$G_\lambda(c) = \int_a^b p(K)(w_\lambda'^2 - w_{\lambda,c}'^2) + \lambda q(K)(w_\lambda^2 - w_{\lambda,c}^2) - 2\beta q(K)(w_\lambda - w_{\lambda,c})\, dK. \tag{2.10}$$

3. Properties of the Functional G_λ

The main properties of the functional G_λ are summarized in the following

Theorem 3.1. (a) *For any $c = (p, q)$ in $\mathcal{D}$,*

$$G_\lambda(c) = \int_a^b p(w'_\lambda - w'_{\lambda,c})^2 + \lambda q (w_\lambda - w_{\lambda,c})^2. \tag{3.1}$$

(b) *$G_\lambda(c) \ge 0$ for all $c = (p, q)$ in $\mathcal{D}$, and $G_\lambda(c)=0$ if and only if $w_\lambda = w_{\lambda,c}$.*
(c) *The first Gâteaux derivative of G_λ is given by*

$$G'_\lambda(p,q)[h_1,h_2] = \int_a^b \underbrace{(w_\lambda'^2 - w_{\lambda,c}'^2)}_{\text{L}^2 \text{ gradient in p}} h_1 + \underbrace{[\lambda(w_\lambda^2 - w_{\lambda,c}^2) - 2\beta(w_\lambda - w_{\lambda,c})]}_{\text{L}^2 \text{ gradient in q}} h_2. \tag{3.2}$$

(d) *The second Gâteaux derivative of G_λ is given by*

$$G''_\lambda(c)[h,k] = 2(L^{-1}_{p,\lambda q}(e(h)), e(k)), \tag{3.3}$$

where $h = (h_1, h_2)$, $k = (k_1, k_2)$,

$$e(h) = -(h_1 w'_{\lambda,c})' + \lambda h_2 w_{\lambda,c} - \beta h_2,$$

and $(\cdot\,,\cdot)$ denotes the usual inner product in $L^2[a,b]$.

Proof. (a) If $v \in W^{1,2}[a,b]$ and $\phi \in W_0^{1,2}[a,b]$ then by integration by parts we have

$$\int_a^b p(x) v' \phi'\, dx = \underbrace{p(x) v' \phi|_a^b}_{=0} - \int_a^b \phi (p(x) v')'\, dx = -\int_a^b \phi(p(x)v')'\, dx. \tag{3.4}$$

Consequently, from (3.4) using $\phi = w_\lambda - w_{\lambda,c} \in W_0^{1,2}[a,b]$,

$$\begin{aligned}
G_\lambda(c) &= \int_a^b p(w_\lambda'^2 - w_{\lambda,c}'^2) + \lambda q((w_\lambda^2 - w_{\lambda,c}^2) - 2\beta q(w_\lambda - w_{\lambda,c}) \\
&= \int_a^b p(w_\lambda' - w_{\lambda,c}')^2 + 2pw_{\lambda,c}'(w_\lambda' - w_{\lambda,c}') \\
&\quad + \lambda q((w_\lambda^2 - w_{\lambda,c}^2) - 2\beta q(w_\lambda - w_{\lambda,c}) \\
&= \int_a^b p(w_\lambda' - w_{\lambda,c}')^2 - 2(w_\lambda - w_{\lambda,c})(pw_{\lambda,c}')' \\
&\quad + \lambda q((w_\lambda^2 - w_{\lambda,c}^2) - 2\beta q(w_\lambda - w_{\lambda,c}), \\
&\quad \text{using (2.8) for } (pw_{\lambda,c}')' , \\
&= \int_a^b p(w_\lambda' - w_{\lambda,c}')^2 - 2(w_\lambda - w_{\lambda,c})(\lambda q w_{\lambda,c} - \beta q) \\
&\quad + \lambda q((w_\lambda^2 - w_{\lambda,c}^2) - 2\beta q(w_\lambda - w_{\lambda,c}) \\
&= \int_a^b p(w_\lambda' - w_{\lambda,c}')^2 + \lambda q(w_\lambda - w_{\lambda,c})^2,
\end{aligned}$$

after some rearrangement.

(b) As p and q are chosen to be positive and $\lambda > 0$, from (a) we get (b).

(c) The first Gâteaux derivative of the functional G_λ is given by

$$\begin{aligned}
&G_\lambda'(p,q)[h_1,h_2] \\
&= \lim_{\varepsilon\to 0} \frac{G_\lambda(c+\varepsilon h) - G_\lambda(c)}{\varepsilon} \\
&= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_a^b (p+\varepsilon h_1)(w_\lambda'^2 - w_{\lambda,c+\varepsilon h}'^2) + \lambda(q+\varepsilon h_2)(w_\lambda^2 - w_{\lambda,c+\varepsilon h}^2) \\
&\quad 2\beta(q+\varepsilon h_2)(w_\lambda - w_{\lambda,c+\varepsilon h}) - p(w_\lambda'^2 - w_{\lambda,c}'^2) \\
&\quad - \lambda q(w_\lambda^2 - w_{\lambda,c}^2) + 2\beta q(w_\lambda - w_{\lambda,c}) \\
&= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_a^b \varepsilon(w_\lambda'^2 - w_{\lambda,c+\varepsilon h}'^2)\, h_1 + \varepsilon\lambda(w_\lambda^2 - w_{\lambda,c+\varepsilon h}^2)\, h_2 \\
&\quad - 2\varepsilon\beta(w_\lambda - w_{\lambda,c})\, h_2 + p(w_\lambda'^2 - w_{\lambda,c+\varepsilon h}'^2) \\
&\quad + \lambda q(w_\lambda^2 - w_{\lambda,c+\varepsilon h}^2) - 2q\beta(w_\lambda - w_{\lambda,c+\varepsilon h}) \\
&= \lim_{\varepsilon\to 0} \int_a^b (w_\lambda'^2 - w_{\lambda,c+\varepsilon h}'^2)\, h_1 + [\lambda(w_\lambda^2 - w_{\lambda,c+\varepsilon h}^2) - 2\beta(w_\lambda - w_{\lambda,c})]\, h_2 \\
&\quad + \lim_{\varepsilon\to 0} \int_a^b \frac{1}{\varepsilon} p(w_{\lambda,c}'^2 - w_{\lambda,c+\varepsilon h}'^2) + \frac{1}{\varepsilon}\lambda q(w_{\lambda,c}^2 - w_{\lambda,c+\varepsilon h}^2) \\
&\quad - \frac{1}{\varepsilon} 2q\beta(w_{\lambda,c} - w_{\lambda,c+\varepsilon h})
\end{aligned}$$

If we can show the second term is zero we get (3.2). Let the integral in the second term be denoted by I. Now,

$$-(pw_{\lambda,c}')' + \lambda q w_{\lambda,c} = \beta q, \tag{3.5}$$

$$-((p+\varepsilon h_1)w_{\lambda,c+\varepsilon h}')' + \lambda(q+\varepsilon h_2)w_{\lambda,c+\varepsilon h} = \beta(q+\varepsilon h_2). \tag{3.6}$$

The first term in the integral I can be expanded as

$$\begin{aligned}
&\varepsilon^{-1}\int_a^b p(w'^2_{\lambda,c} - w'^2_{\lambda,c+\varepsilon h})\\
&= \varepsilon^{-1}\int_a^b p(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h})(w'_{\lambda,c} - w'_{\lambda,c+\varepsilon h}),\\
&\quad \text{from (3.4) using } \phi = w_{\lambda,c} - w_{\lambda,c+\varepsilon h}\ ,\\
&= \varepsilon^{-1}\int_a^b (w_{\lambda,c+\varepsilon h} - w_{\lambda,c})(p(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h}))'\\
&= \varepsilon^{-1}\int_a^b (w_{\lambda,c+\varepsilon h} - w_{\lambda,c})[(pw'_{\lambda,c})' + (pw'_{\lambda,c+\varepsilon h})'],\\
&\quad \text{using (3.5) and (3.6)},\\
&= \varepsilon^{-1}\int_a^b (w_{\lambda,c+\varepsilon h} - w_{\lambda,c})[\lambda q w_{\lambda,c} - \beta q + \lambda(q+\varepsilon h_2)w_{\lambda,c+\varepsilon h}\\
&\quad - \beta(q+\varepsilon h_2) - \varepsilon(h_1 w'_{\lambda,c+\varepsilon h})']\\
&= \int_a^b (w_{\lambda,c+\varepsilon h} - w_{\lambda,c})[\lambda h_2 w_{\lambda,c+\varepsilon h} - \beta h_2 - (h_1 w'_{\lambda,c+\varepsilon h})']\\
&\quad + \varepsilon^{-1}\int_a^b (w_{\lambda,c+\varepsilon h} - w_{\lambda,c})[\lambda q(w_{\lambda,c} + w_{\lambda,c+\varepsilon h}) - 2\beta q]\\
&= \int_a^b (w_{\lambda,c+\varepsilon h} - w_{\lambda,c})[\lambda h_2 w_{\lambda,c+\varepsilon h} - \beta h_2 - (h_1 w'_{\lambda,c+\varepsilon h})']\\
&\quad + \varepsilon^{-1}\int_a^b \lambda q(w^2_{\lambda,c+\varepsilon h} - w^2_{\lambda,c})(-2\beta q(w_{\lambda,c+\varepsilon h} - w_{\lambda,c})).
\end{aligned}$$

Substituting the above for $\varepsilon^{-1}\int_a^b p(w'^2_{\lambda,c} - w'^2_{\lambda,c+\varepsilon h})$ in I we obtain

$$I = \int_a^b (w_{\lambda,c+\varepsilon h} - w_{\lambda,c})[\lambda h_2 w_{\lambda,c+\varepsilon h} - \beta h_2 - (h_1 w'_{\lambda,c+\varepsilon h})'].$$

It follows that $I \to 0$ as $\varepsilon \to 0$.

(d) To find the second Gâteaux derivative of the functional G_λ we will need the following result:

$$\begin{aligned}
L_{p,\lambda q}(w_{\lambda,c+\varepsilon h} - w_{\lambda,c}) &= -(p(w_{\lambda,c+\varepsilon h} - w_{\lambda,c})')' + \lambda q(w_{\lambda,c+\varepsilon h} - w_{\lambda,c})\\
&= -(pw'_{\lambda,c+\varepsilon h})' + \lambda q w_{\lambda,c+\varepsilon h} - [-(pw'_{\lambda,c})' + \lambda q w_{\lambda,c}],\\
&\quad \text{using (3.5) and (3.6)},\\
&= \varepsilon[(h_1 w'_{\lambda,c+\varepsilon h})' - \lambda h_2 w_{\lambda,c+\varepsilon h} + \beta h_2]
\end{aligned} \tag{3.7}$$

The second Gâteaux derivative of the functional G_λ is given by

$$\begin{aligned}
&G''_\lambda(c)[h,k]\\
&= \lim_{\varepsilon\to 0}\frac{G'(c+\varepsilon h)[k] - G'(c)[k]}{\varepsilon}\\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_a^b (w'^2_\lambda - w'^2_{\lambda,c+\varepsilon h})k_1 + [\lambda(w^2_\lambda - w^2_{\lambda,c+\varepsilon h}) - 2\beta(w_\lambda - w_{\lambda,c+\varepsilon h})]k_2\\
&\quad - (w'^2_\lambda - w'^2_{\lambda,c})k_1 - [\lambda(w^2_\lambda - w^2_{\lambda,c}) - 2\beta(w_\lambda - w_{\lambda,c})]k_2
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_a^b (w'^2_{\lambda,c} - w'^2_{\lambda,c+\varepsilon h})k_1 + [\lambda(w^2_{\lambda,c} - w^2_{\lambda,c+\varepsilon h}) - 2\beta(w_{\lambda,c} - w_{\lambda,c+\varepsilon h})]k_2 \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_a^b k_1(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h})(w'_{\lambda,c} - w'_{\lambda,c+\varepsilon h}) \\
&\quad + [\lambda(w^2_{\lambda,c} - w^2_{\lambda,c+\varepsilon h}) - 2\beta(w_{\lambda,c} - w_{\lambda,c+\varepsilon h})]k_2, \\
&\quad \text{from (3.4) using } \phi = w_{\lambda,c} - w_{\lambda,c+\varepsilon h}\ , \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_a^b (w_{\lambda,c} - w_{\lambda,c+\varepsilon h})(-k_1(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h}))' \\
&\quad + [\lambda(w^2_{\lambda,c} - w^2_{\lambda,c+\varepsilon h}) - 2\beta(w_{\lambda,c} - w_{\lambda,c+\varepsilon h})]k_2, \\
&\quad \text{factoring } (w_{\lambda,c} - w_{\lambda,c+\varepsilon h})\ , \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_a^b (w_{\lambda,c} - w_{\lambda,c+\varepsilon h})\,[(-k_1(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h}))' \\
&\quad + (\lambda(w_{\lambda,c} + w_{\lambda,c+\varepsilon h}) - 2\beta)k_2], \\
&\quad \text{using (3.7)}, \\
&= \lim_{\varepsilon\to 0}\int_a^b L^{-1}_{p,\lambda q}[-(h_1 w'_{\lambda,c+\varepsilon h})' + \lambda h_2 w_{\lambda,c+\varepsilon h} - \beta h_2] \\
&\quad \times [(-k_1(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h}))' + (\lambda(w_{\lambda,c} + w_{\lambda,c+\varepsilon h}) - 2\beta)k_2] \\
&= \lim_{\varepsilon\to 0}\int_a^b L^{-1}_{p,\lambda q}[-(h_1(w'_{\lambda,c+\varepsilon h} - w'_{\lambda,c}))' + \lambda h_2(w_{\lambda,c+\varepsilon h} - w_{\lambda,c})] \\
&\quad \times [(-k_1(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h}))' + (\lambda(w_{\lambda,c} + w_{\lambda,c+\varepsilon h}) - 2\beta)k_2] \\
&\quad + \lim_{\varepsilon\to 0}\int_a^b L^{-1}_{p,\lambda q}[-(h_1 w'_{\lambda,c})' + \lambda h_2 w_{\lambda,c}) - \beta h_2] \\
&\quad \times [(-k_1(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h}))' + (\lambda(w_{\lambda,c} + w_{\lambda,c+\varepsilon h}) - 2\beta)k_2], \\
&\quad \text{expanding the second integral}, \\
&= \lim_{\varepsilon\to 0}\int_a^b L^{-1}_{p,\lambda q}[-(h_1(w'_{\lambda,c+\varepsilon h} - w'_{\lambda,c}))' + \lambda h_2(w_{\lambda,c+\varepsilon h} - w_{\lambda,c})] \\
&\quad \times [(-k_1(w'_{\lambda,c} + w'_{\lambda,c+\varepsilon h}))' + (\lambda(w_{\lambda,c} + w_{\lambda,c+\varepsilon h}) - 2\beta)k_2] \\
&\quad + \lim_{\varepsilon\to 0}\int_a^b L^{-1}_{p,\lambda q}[-(h_1 w'_{\lambda,c})' + \lambda h_2 w_{\lambda,c}) - \beta h_2] \\
&\quad \times [(-k_1(w'_{\lambda,c+\varepsilon h} - w'_{\lambda,c}))' + \lambda(w_{\lambda,c+\varepsilon h} - w_{\lambda,c})k_2] \\
&\quad + 2\int_a^b L^{-1}_{p,\lambda q}[-(h_1 w'_{\lambda,c})' + \lambda h_2 w_{\lambda,c} - \beta h_2] \\
&\quad \times [-(k_1 w'_{\lambda,c})' + \lambda k_2 w_{\lambda,c} - \beta k_2].
\end{aligned}
$$

The first and second terms equal zero. Thus we obtain

$$
G''_\lambda(c)[h,k] = 2(L^{-1}_{p,\lambda q}(e(h)), e(k)), \tag{3.8}
$$

where

$$
\begin{aligned}
e(h) &= -(h_1 w'_{\lambda,c})' + \lambda h_2 w_{\lambda,c} - \beta h_2, \\
e(k) &= -(k_1 w'_{\lambda,c})' + \lambda k_2 w_{\lambda,c} - \beta k_2.
\end{aligned}
$$

This completes the proof of the theorem. □

With some additional work one can show that the first and second Gâteaux derivatives of G_λ are also Fréchet derivatives. As $L_{p,\lambda q}$ is a positive operator on $W_0^1[a,b]$, we have from Theorem 3.1(d) that $G_\lambda''(c) \geq 0$ for all c in the convex set $\mathcal{D}$. By [25, Corollary 42.8] the functional G_λ is therefore convex on $\mathcal{D}$. We know from Theorem 3.1(b) that G_λ has a global minimum (zero) at $c = (p,q)$ if and only if $w_\lambda = w_{\lambda,c}$. Choose $N \geq 3$ positive distinct real numbers λ_j, $1 \leq j \leq N$, so that

$$0 < \lambda_j T < 2, \quad T \in [T_{i-1}, T_i].$$

Define a convex functional G on the domain $\mathcal{D}$ (defined above) by

$$G(c) = \sum_{j=1}^{N} G_{\lambda_j}(c). \tag{3.9}$$

From the uniqueness theorem [12, Theorem 3.5] we know that, under certain (computer-verifiable) conditions on the nature of the flows of certain associated vector fields (which amount here to an admissibility restriction on the data $v(K,T)$), the condition $w_\lambda = w_{\lambda,c}$ for at least three distinct values of λ implies that $c = (p,q) = (P,Q)$. By [25, Proposition 42.6(1)] we know that if the convex functional G has a stationary point at (p,q) then it must have a global minimum there, and from the foregoing (assuming admissible data) that stationary point must uniquely occur at (P,Q). So, the desired function pair (P,Q) now appears as the unique global minimum of a convex functional with a unique stationary point. In practical numerics this is an important consideration, as many (if not most) least-square type minimization methods suffer greatly from the minimization process getting stuck in spurious local minima. That this cannot happen here is one of the significant advantages of our approach.

4. The Algorithm

$G(c)$ is a nonnegative convex functional since it is the sum of nonnegative convex functionals, and it also has a unique stationary point at $c = (P,Q)$. The idea here is that by using G rather than just one of the G_λ, in addition to gaining favourable uniqueness properties, we are blending additional time-based data into the inverse problem, and this is intended to improve the well-posedness of the problem. We note in passing from [13] that this inverse recovery is conditionally well-posed in the weak-L^2 sense, so from a theoretical standpoint, the recoveries are expected to be quite stable, which indeed is the case.

We minimize this functional for $N = 20$ using the steepest descent method to recover the coefficients $P(K)$ and $Q(K)$. The L^2-direction of steepest descent for G at $c_0 = (p_0, q_0)$ with respect to p is

$$-\nabla_{L^2,p} G(c_0) = \sum_{j=1}^{N} (w_{\lambda_j}'^2 - w_{\lambda_j,c_0}'^2),$$

and the L^2-direction of steepest descent for G at (p_0, q_0) with respect to the variable q is given by

$$-\nabla_{L^2,q} G(c_0) = \sum_{j=1}^{N} [\lambda_j (w_{\lambda_j}^2 - w_{\lambda_j,c_0}^2) - 2\beta(w_{\lambda_j} - w_{\lambda_j,c_0})].$$

Instead of using these L^2-gradients we use the corresponding Neuberger-gradients (see [21]) as the L^2-gradient has numerical problems that are extensively discussed in [16]. In particular, the L^2-gradient with respect to q is zero on the boundary of [a,b] given that w_λ and $w_{\lambda,c}$ are equal there, and thus the algorithm is unable to properly recover Q. The Neuberger-gradient smooths the L^2-gradient and preserves boundary data during the descent, an important property not shared by other descent techniques. Our Neuberger-gradient $g = \nabla_{H^1} G$ can be found from an L^2-gradient $\nabla_{L^2} G$ by solving the boundary value problem

$$\begin{aligned} -g'' + g &= \nabla_{L^2} G, \\ g(a) = g(b) &= 0. \end{aligned} \tag{4.1}$$

Below is the steepest descent algorithm used to get one descent step in p:

(1) Initialize $p(K)$ and $q(K)$ with $c_0 = (p_0, q_0)$.
(2) Find w_{λ,c_0} and w'_{λ,c_0} by solving (2.8),(2.9).
(3) Find the L^2 gradient of G in p, $\nabla_{L^2,p} G(c_0)$.
(4) Find the Neuberger gradient in p, $\nabla_{H^1,p} G(c_0)$.
(5) Evaluate $p_{new}(K) = p_0(K) - \alpha \nabla_{H^1,p} G(c_0)$.
(6) Find $G(p, q_0)$ using $p_{new}(K)$ for $p(K)$.
(7) Find α that gives the lowest value of $G(p, q_0)$.
(8) Set $p(K) = p_{new}(K)$.

The descent in q is similar to that of descent in p. Here we find corresponding gradients in q. The $q_{new}(K)$ is given by

$$q_{new}(K) = q(K) - \alpha \nabla_{H^1,q} G(c_0)$$

The specific order of descent is somewhat problem dependent, and different combinations of descents in p and q were tried to get the best minimization. Typically one needs more p-descent steps relative to q-descent steps as the descent progresses.

5. Results

One of the most popular European options traded on US exchanges is the option on the Standard & Poors 500 (SPX) index. Call option prices on the SPX index were taken from the official website of the Chicago Board Options Exchange (CBOE), for two consecutive maturities on the 22nd of February, 2012. The data includes only the near-the-money options as they are the most heavily traded. To recover the coefficient functions $P(K)$ and $Q(K)$ in (2.4) a computer code code was written in the programming language C. The volatility recovered was compared to the "implied volatility" obtained directly from the standard formula of Black and Scholes by substituting the known option price and solving for the implied volatility σ as an unknown.

We have option prices for discrete sets of strikes and expirations. We generated the function $v(K,T)$ by linearly interpolating the option price in both strike and expiration. The function $v(K,T)$ was mollified (c.f. [14, §6]) so that it could be differentiated, and the derivative $v_K(K,T)$ was found using central differences. For 20 fixed values of λ the functions $v(K,T)$ and $v'_K(K,T)$ were Laplace transformed using (2.2) to $w_\lambda(K)$ and $w'_\lambda(K)$ respectively. The functions $p(K)$ and $q(K)$ were initialized using (2.3) and (2.5) with the initial σ_i chosen to be the implied volatility. We performed a series of descents in p using the aforementioned Neuberger steepest descent algorithm such that the functional could not be minimized any further.

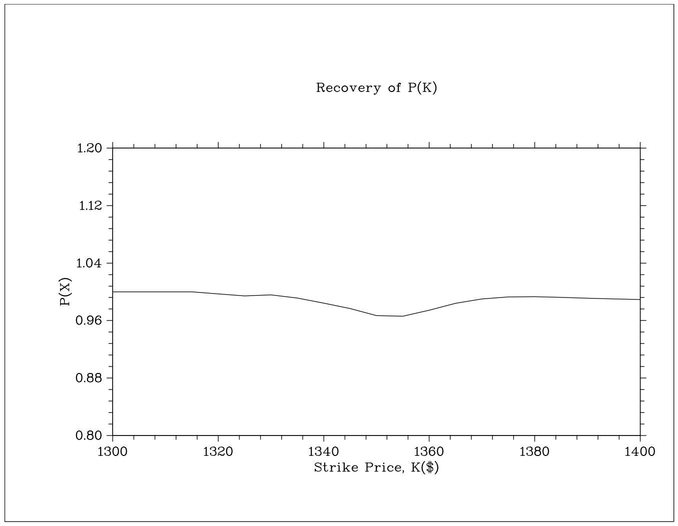
Spot Price(S_0)		\$ 1357.66
Maturity Time (T_1)		2 days
Maturity Time (T_2)		22 days
Strike Price(K)	$v(K,T_1)$	$v(K,T_2)$
1300	59.40	63.00
1305	54.40	58.60
1310	49.20	54.20
1315	44.60	50.00
1320	39.40	45.80
1325	34.80	41.00
1330	29.80	37.90
1335	25.10	34.10
1340	20.60	30.50
1345	16.40	27.00
1350	12.5	23.7
1355	7.9	20.6
1360	5.1	17.7
1365	2.85	14.5
1370	1.6	12.7
1375	0.9	10
1380	0.55	8.7
1385	0.3	7
1390	0.25	5.6
1395	0.2	4.5
1400	0.2	3.6

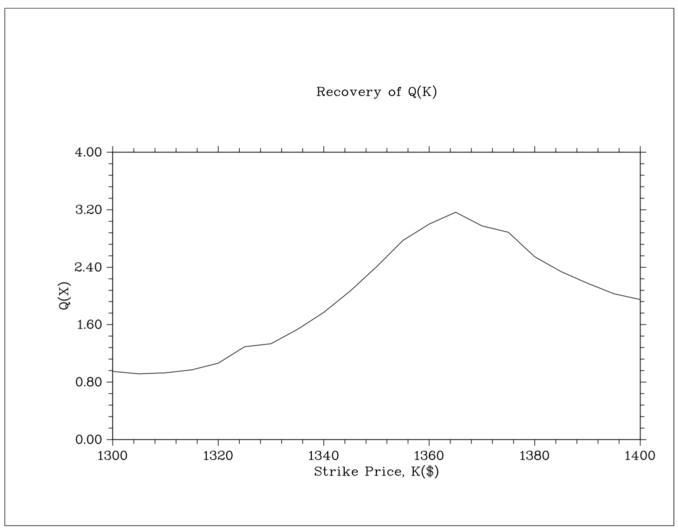
Then a series of descents in q were performed to the point where functional likewise could not be lowered any further. We repeated this sequence of descents in p and q. The minimization of $G(c)$ in α was done using the well known Brent minimization technique, by adapting the one-variable code in the Numerical Recipes in C function `brent()`. To avoid possible catastrophic cancellation in the Simpson rule formula used in the calculation of the integrals in the formula (2.10) for the functional G_λ, we used the alternate formula (3.1) instead. After running the code we recovered the functions $P(K)$ and $Q(K)$ graphed below.

From (2.7) we calculated the volatility and compared it to the implied volatility of the option at first and second expirations, as shown in the graph below. On taking subsequent maturity intervals a volatility surface can in principle be plotted.

Finally, from the recovered volatility we calculated the option price in MATLAB using the Binomial method and compared it to the actual price, as shown in the figure below.

Conlusion. We have shown that volatility can be recovered from published option prices using a steepest descent minimization technique. This provides a "market view" of future volatility which in principle can used to trade options more efficiently. The results obtained look promising. The analogous work on recovering volatility for the much more ubiquitous American options is in progress. It would be interesting to consider interest rate r (also known in this context as the risk-neutral





drift) as function of time and asset price, instead of treating it as a constant, and recover it in similar fashion from the option price.

References

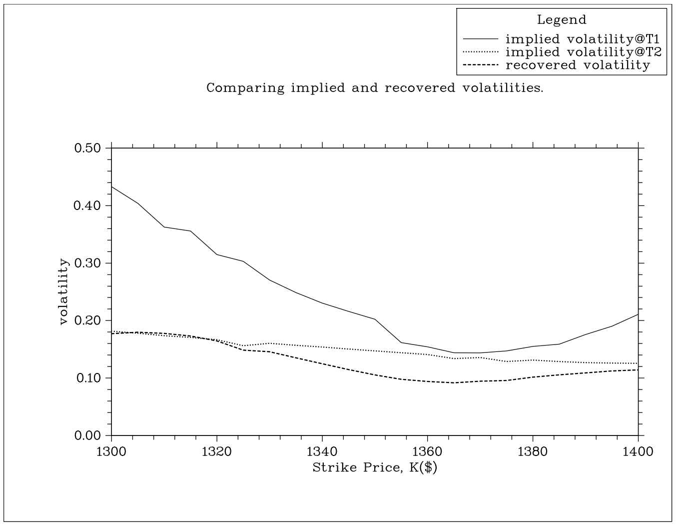
[1] M. Avellaneda, C. Friedman, L. Holmes, and L. Sampieri; Calibrating volatility surfaces via relative entropy minimization. *Appl. Math. Finance*, 4:37–64, 1997.

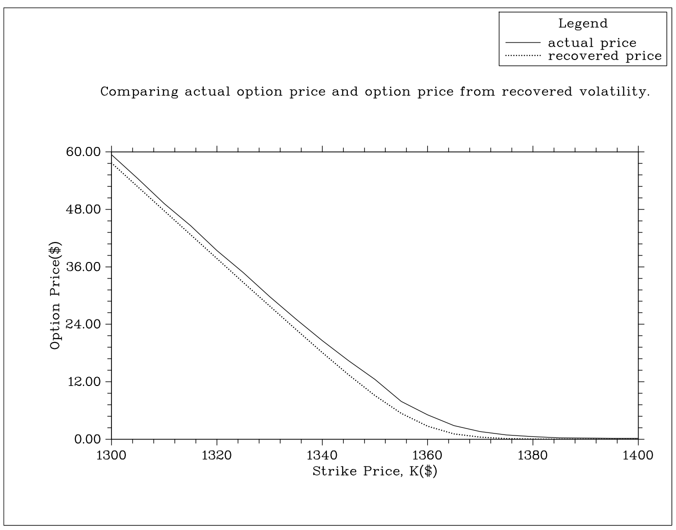
[2] Louis Bachelier; *Théorie de la Spéculation*. PhD thesis, L'École Normale Supériere, 1900. Translation: Cootner, 1964.

[3] Fischer Black and Myron Scholes; The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[4] J. N. Bodurtha and M. Jermakyan; Non-parameric estimation of an implied volatility surface. *J. Computational Finance*, 2:29–61, 1999.

[5] Ilia Bouchouev and Victor Isakov; The inverse problem of option pricing. *Inverse Problems*, 13:L11–L17, 1999.





[6] Ilia Bouchouev and Victor Isakov; Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. *Inverse Problems*, 15:R95–R116, 1999.

[7] Ilia Bouchouev and Victor Isakov; Recovery of volatility coefficient by linearization. *Quant. Finance*, 2:257–263, 2002.

[8] B. Dupire; Pricing with a smile. *RISK*, 7:18–20, 1994.

[9] Eugene F. Fama; The behavior of stock-market prices. *Journal of Business*, 38:34–105, 1965.

[10] Philip Hartman; *Ordinary differential equations*. S. M. Hartman, Baltimore, Md., 1973. Corrected reprint.

[11] Victor Isakov; The inverse problem of option pricing. Preprint, 2004.

[12] Ian Knowles; Uniqueness for an elliptic inverse problem. *SIAM J. Appl. Math.*, 59(4):1356–1370, 1999.

[13] Ian Knowles, Mary A. LaRussa; Conditioinal well-posedness for an elliptic inverse problem. *SIAM J. Appl. Math.*, 71:952–971, 2011. Available online at http://www.math.uab.edu/knowles/pubs.html.

[14] Ian Knowles, Tuan A. Le, Aimin Yan; On the recovery of multiple flow parameters from transient head data. *J. Comp. Appl. Math.*, 169:1–15, 2004.

[15] Ian Knowles, Michael Teubner, Aimin Yan, Paul Rasser, Jong Wook Lee; Inverse groundwater modelling in the Willunga Basin, South Australia. *Hydrogeology Journal*, 15:1107–1118, 2007.

[16] Ian Knowles, Robert Wallace; A variational method for numerical differentiation. *Numerische Mathematik*, 70:91–110, 1995.

[17] S. J. Kon; Models of stock returns - a comparison. *Journal of Finance*, 39(1):147–165, 1984.

[18] R. Lagnado and S. Osher; A technique for calibrating derivation of the security pricing models: numerical solution of the inverse problem. *J. Computational Finance*, 1:13–25, 1997.

[19] D. B. Madan and E. Seneta; The variance gamma model for share market returns. *Journal of Business*, 63(4):511–524, 1990.

[20] Benoit B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36:394–41, 1963.

[21] J. W. Neuberger; *Sobolev gradients and differential equations*, volume 1670 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2010.

[22] R. R. Officer; The distribution of stock returns. *Journal of the American Statistical Association*, 67(340):807–812, 1972.

[23] P. D. Praetz. The distribution of share price changes. *Journal of Business*, 45(1):49–55, 1972.

[24] S. J. Press; A compound events model for security prices. *Journal of Business*, 40(July):317–335, 1967.

[25] Eberhard Zeidler; *Nonlinear functional analysis and its applications. III.* Springer-Verlag, New York, 1985. Variational methods and optimization, Translated from the German by Leo F. Boron.

Ian Knowles
Department of Mathematics, University of Alabama at Birmingham, Birmingham AL 35294, USA
E-mail address: iknowles@uab.edu

Li Feng
Department of Mathematics, University of Alabama at Birmingham, Birmingham AL 35294, USA
E-mail address: lifeng@uab.edu

Ajay Mahato
Department of Mathematics, University of Alabama at Birmingham, Birmingham AL 35294, USA
E-mail address: amahato7@gmail.com