A semi-smooth Newton method for an inverse problem in option pricing

Bertram Düring*

Institut für Analysis und Scientific Computing, Technische Universität Wien, 1040 Wien, Austria.

We present an optimal control approach using a Lagrangian framework to identify local volatility functions from given option prices. We employ a globalized sequential quadratic programming (SQP) algorithm and implement a line search strategy. The linear-quadratic optimal control problems in each iteration are solved by a primal-dual active set strategy which leads to a semi-smooth Newton method. We present first- and second-order analysis as well as numerical results.

1 Introduction

In an idealized financial market the price of a European option can be obtained as the solution of the celebrated Black-Scholes equation. This equation has been derived under several assumptions, in particular that the volatility of the underlying asset is constant. However, if one computes the *implied volatility* from market option prices by inverting the closed-form solution to the Black-Scholes equation, it is typically not constant, but rather shows a *smile* or *skew* pattern. These observations lead to a natural generalization of the Black-Scholes model replacing the constant volatility by a *local volatility function* $\sigma(T, E)$, where T denotes the time to maturity and E the exercise price. Then, the option price $V(T, E)$ as a function of the exercise time T and the exercise price E satisfies the (forward) differential equation

$$V_T(T,E) - \tfrac{1}{2}\sigma^2(T,E)E^2V_{EE}(T,E) + rEV_E(T,E) = 0, \quad T > 0,\ E > 0, \tag{1}$$

with initial condition $V(0, E) = \max(S_0 - E, 0)$ and boundary conditions $V(T, 0) = S_0$, $\lim_{E\to\infty} V(T, E) = 0$, where S_0 denotes the current price of the underlying [1]. Our goal is to identify from market option prices the volatility function in (1).

2 The optimal control problem

To streamline the presentation we restrict ourselves to the case of zero interest rate ($r = 0$) in the analytical part of the paper. For $R > E > M > 0$ and $T > 0$ let $\Omega = (M, R)$ and $Q = (0, T) \times \Omega$. Let $V = \{\varphi \in H^1(\Omega) : \varphi(R) = 0\}$, $W(0, T) = \{\varphi \in L^2(0, T; V) : \varphi_t \in L^2(0, T; V')\}$, and $H^{2,1}(Q) = H^1(0, T; L^2(\Omega)) \cap L^2(0, T; H^2(\Omega))$. We define the two Hilbert spaces $X = H^{2,1}(Q) \times W(0, T)$ and $Y = L^2(0, T; H^1_0(\Omega)) \times L^2(0, T) \times L^2(\Omega)$ and introduce the bilinear operator $e = (e_1, e_2, e_3) : X \to Y'$ by

$$e_1(\omega) = u_t - qu_{xx}, \quad e_2(\omega) = u(\cdot\,, M) - u_D, \quad e_3(\omega) = u(0) - u_0, \tag{2}$$

where $\omega = (q, u)$. Our goal is to identify the coefficient $q(t, x) = \frac{1}{2}E^2\sigma^2(T, E)$. We use the cost functional $J : X \to [0, \infty)$,

$$J(\omega) = \frac{1}{2}\int_\Omega |u(T) - u_T|^2\,\mathrm{d}x + \frac{\beta}{2}\,\|q - q_d\|^2_{H^{2,1}(Q)} \quad \text{for } \omega = (q, u) \in X,$$

where u_T is a given observed option price at the end-time T, q_d is an *a priori* guess and $\beta > 0$ is a regularization parameter.

The parameter identification problem is given by a constrained optimal control problem in the following form

$$\min J(\omega) \quad \text{s.t.} \quad \omega \in K_{\mathsf{ad}} \quad \text{and} \quad e(\omega) = 0, \tag{3}$$

where $K_{\mathsf{ad}} = \mathcal{Q}_{\mathsf{ad}} \times W(0, T)$ and $\mathcal{Q}_{\mathsf{ad}} = \big\{q \in H^{2,1}(Q) : q_{\mathsf{min}} \le q \le q_{\mathsf{max}} \text{ in } Q \text{ a.e.}\big\}$ is the set of admissible coefficient functions. Note that both the state variable u and the coefficient q are considered as independent variables while the realization of (2) is an explicit constraint. The existence of at least one (global) solution to (3) was proved in [2].

The bilinear operator $e : X \to Y'$ is twice continuously Fréchet differentiable and the mapping $\omega \mapsto e''(\omega)$ is Lipschitz continuous on X. Moreover, its linearization $e'(\omega) : X \to Y'$ at any point $\omega = (q, u) \in K_{\mathsf{ad}}$ is surjective. This guarantees a constraint qualification, so that there exists a (unique) Lagrange multiplier λ^* satisfying the first-order necessary optimality condition, i.e. stationarity of the Lagrange functional $L(\omega, \lambda) = J(\omega) + \langle e(\omega), \lambda\rangle$ associated with problem (3). Using an error estimate for the Lagrange multiplier λ, one can ensure that a second order sufficient optimality condition holds if the residual $\|u^*(T) - u_T\|_{L^2(\Omega)}$ is sufficiently small. Furthermore, there exists a unique Lagrange multiplier associated with the inequality constraints for the optimal coefficient q^* (for details, see [2]).

* Corresponding author: e-mail: bduering@anum.tuwien.ac.at

Strike E	95	97.5	100	102.5	105
True value	0.1500	0.1500	0.1500	0.1500	0.1500
Good guess	0.1454	0.1500	0.1517	0.1506	0.1470
Good guess & noise	0.1457	0.1500	0.1517	0.1506	0.1470
Bad guess	0.1458	0.1500	0.1517	0.1506	0.1472
Bad guess & noise	0.1460	0.1500	0.1517	0.1506	0.1472
Good guess, fine grid	0.1488	0.1500	0.1509	0.1503	0.1494

Table 1 Reconstructed volatilities for different strikes E.

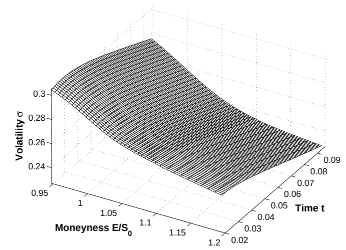

Fig. 1 Local volatility function $\sigma(E, T)$.

3 The optimization method

The method is based on an iterative algorithm proposed in [2]. We initialize our method and choose updates in such a way, that the constraints $e_2(\omega^n) = 0$ and $e_3(\omega^n) = 0$ are fulfilled by construction for all iterates ω^n. We are left with only one constraint, $e_1(\omega) = 0$. In an *outer loop*, we minimize at each iteration a quadratic approximation of the Lagrangian over an affine subspace of solutions, i.e. we solve for the update $\delta\omega^n = (\delta q^n, \delta u^n)$ the linear-quadratic optimal control problem

$$\min \mathsf{L}(\omega^n, \lambda^n) + \mathsf{L}'(\omega^n, \lambda^n)\delta\omega + \tfrac{1}{2}\, L''(\omega^n, \lambda^n)(\delta\omega^n, \delta\omega^n) \quad \text{s.t.} \quad e_1'(\omega^n)\delta\omega^n + e_1(\omega^n) = 0,\ q_{\mathsf{min}} \le q^n + \delta q^n \le q_{\mathsf{max}}.$$

This subproblem involves linear inequality constraints. For the solution of the subproblems in an *inner loop* we use a primal-dual active set method [3]. In contrast to [2], here we do not iterate until convergence, but only perform one step of the primal-dual active set method in each SQP iteration. This leads to a semi-smooth Newton method [5], that converges locally superlinearly. The relationship between the primal dual active set strategy and semi-smooth Newton methods has been investigated in [4]. We also implement a globalization strategy that is realized by a modification of the Hessian matrix and by a line search strategy to ensure that every SQP step is a descent direction. Since we do not iterate the *inner loop* until convergence, the inequality constraints are not necessarily fulfilled in each SQP iteration. Therefore, we need to modify the line search strategy presented in [2]. To determine the step size parameter α^n we here use a first-order variation of the merit function

$$\varphi^n(\alpha^n) = J(\omega^n + \alpha^n\delta\omega^n) + \mu\, \|e_1(\omega^n + \alpha^n\delta\omega^n)\|_{Y_1'} + \zeta\, \|\max\{0, q^n - q_{\mathsf{max}}\}\|_{L^2(Q)} + \eta\, \|\max\{0, q_{\mathsf{min}} - q^n\}\|_{L^2(Q)}.$$

For appropriately chosen penalty parameters, our line search strategy is then based on the well-known Armijo rule.

4 Numerical experiments

For the discretization we use linear finite elements on a non-uniform spatial grid with 140 nodes locally refined around $x = S_0$. In time, we employ a fixed, non-equidistant grid consisting of 35 points with small time steps close to $t = 0$. The linear systems are solved by a preconditioned GMRES method. We define a decreasing sequence of regularization parameters β and start our method with the highest value. Then we subsequently decrease the regularization parameter, restarting the method at the minimizers obtained in the previous step. As a first example we apply our method to an artificial data set of Black-Scholes prices with $S_0 = 100$, $r = 0$, one month to maturity and constant volatility $\sigma = 0.15$. We consider four different cases with *a priori* guess $q_d = \frac{1}{2}\sigma_d^2 x^2$. We use a 'good' *a priori* guess $\sigma_d = 0.16$ and a 'bad' *a priori* guess $\sigma_d = 0.1$, and compare these results to those from runs where we added 0.1% uniformly distributed noise. Table 1 shows the results. The volatilities are well identified, with small differences remaining due to discretization errors, which can be reduced by using a finer grid as seen in a fifth run on a grid with halved mesh width in space and time. Overall, the method shows only a very small dependence on the chosen *a priori* guess and it is robust regarding to additional data noise. In a second example we use market data of FTSE index call options. The resulting local volatility function is shown in Figure 1. It is skewed and decreasing as time approaches maturity. This is consistent with empirically observed patterns in equity index options.

Acknowledgements The author is supported by the Deutsche Forschungsgemeinschaft, grant JU 359/6 (Forschergruppe 518).

References

[1] B. Dupire. Pricing with a smile. *Risk* **7**, 18-20 (1994).
[2] B. Düring, A. Jüngel, S. Volkwein. A SQP method for volatility estimation in option pricing. *J. Optim. Theory Appl.*, to appear (2007).
[3] M. Hintermüller. A primal-dual active set algorithm for bilaterally control constrained optimal control problems. *Quarterly Appl. Math.* **61**, 131-161 (2003).
[4] M. Hintermüller, K. Ito, K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM J. Optim.* **13**, 865-888 (2003).
[5] M. Ulbrich. Semismooth Newton methods for operator equations in function spaces. *SIAM J. Optim.* **13**, 805-842 (2003).