HAL open science

Control-theoretic framework for a quasi-Newton local volatility surface inversion

Gabriel Turinici

To cite this version:

Gabriel Turinici. Control-theoretic framework for a quasi-Newton local volatility surface inversion. ICCMSE, Sep 2008, Crete, Greece. pp.254-257. hal-00298960

HAL Id: hal-00298960

https://hal.science/hal-00298960v1

Submitted on 16 Jul 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Control-theoretic framework for a quasi-Newton local volatility surface inversion

Gabriel Turinici

***Abstract*— We investigate in this paper the recovery of the local volatility surface in a parametric framework similar to that of Coleman, Li and Verma [4]. The quality of a surface is assessed through a functional which is optimized; the specificity of the approach is to separate the optimization on the parametric space (performed with any suitable optimization algorithm) from the computation of the functional where we use an adjoint formulation similar to that of the optimal control; the procedure can thus incorporate information from any derivative contract compatible with the adjoint approach. The procedure was implemented and was shown to perform satisfactory on real-world data.**

I. Introduction: the local volatility surface

Let us consider a security (e.g. a stock) whose price follows the stochastic differential equation

$$dS_t/S_t = \mu dt + \sigma dW_t \tag{1}$$

where μ is the return rate, σ is the volatility (we will make explicit its dependence latter) and W_t a Brownian motion. Under the risk-neutral [1] measure, the evolution becomes

$$dS_t/S_t = r(t)dt + \sigma dW_t \tag{2}$$

with $r(t)$ being the time dependent risk-free rate.

If the security S_t distributes dividends at a known proportional rate $q(t)$, under the Black-Scholes local volatility model, the price $C(S,t)$ of a derivative contract on S_t with pay-off $h(S)$ at maturity $t = T$, will satisfy the (Black-Scholes) equation [2] for all $S \geq 0$ and $t \in [0,T]$:

$$\partial_t C + (r-q)\partial_S C + \frac{\sigma^2 S^2}{2}\partial_{SS} C - rC = 0 \tag{3}$$

$$C(S, t = T) = h(S) \tag{4}$$

G. Turinici is with CEREMADE, Université Paris Dauphine, Place du Marechal de Lattre de Tassigny, 75016 Paris, France gabriel.turinici@dauphine.fr

The price at $t = 0$ of the contract is $C(S_{t=0}, t = 0)$; recall that the pay-off of an European call is $h(S) = (S-K)_+$ where we have introduced the strike $K \geq 0$ and the notation $x_+ = \max(x, 0)$. Note the retrograde nature of the equation (3).

The equation (3) gives the price of a derivative contract $C(S,t)$ as function of the risk-free rate $r(t)$, dividend $q(t)$, maturity T, pay-off h and volatility σ. Among those factors, σ alone is not explicitly known or quoted in the markets. The corresponding inverse problem, called "calibration of the local volatility surface" is to recover σ given some known prices $C^{market}_{K_l,T_l}$ of derivative contracts (usually European calls) with (given) maturities T_l and strikes K_l, $l = 1, 2, ..., L$.

When the number of known contracts prices $C^{market}_{K_l,T_l}$ is large enough such that K_l, T_l cover well the range of S and t the local volatility can be expressed using the Dupire formula [2], [3]. However, when only a few prices are known the Dupire formula is less effective and other methods have to be used.

II. Cost functional and the gradient

We take the choice here to formulate the problem in a control setting: the goal is to find the control $\sigma(S,t)$ which ensures that the known prices $C^{market}_{K_l,T_l}$ are hit, i.e. $C_l(S_0, 0) = C^{market}_{K_l,T_l}$ where C_l is the solution of (3) with final data $C_l(S, t = T_l) = (S - K_l)_+$.

To begin, let us view the price

$$C(t = 0; S = S_0) =< \delta_{t=0,S=S_0}, C(S,t) > \tag{5}$$

as a functional in the optimal control setting to be optimized with respect to σ. To compute the gradient $\frac{\delta C}{\delta(\sigma^2)}$ of C with respect to σ^2 one has to introduce an adjoint state χ solution of:

$$\partial_t \chi + \partial_S((r-q)S\chi) + \partial_{SS}(\frac{\sigma^2 S^2}{2}\chi) + r\chi = 0 \quad (6)$$

$$\chi(S, t=0) = \delta_{t=0, S=S_0} \quad (7)$$

and obtain

$$\frac{\delta C}{\delta(\sigma^2)} = \frac{S^2}{2}(\partial_{SS} C)\chi. \quad (8)$$

Of course we also have

$$\frac{\delta C}{\delta \sigma} = 2\sigma \frac{S^2}{2}(\partial_{SS} C)\chi. \quad (9)$$

Both problems (3) and (6) are implemented through a Crank-Nicholson finite-difference scheme; we took care to use for (6) the numerical adjoint of (3). To illustrate the nature of this gradient we display an example in Figure 1 where we note two singularities appearing in $(t = 0, S = S_0)$ (from eqn (7)) and $(t = 1, S = K)$ (from $\partial_{SS}(S-K)_+$).

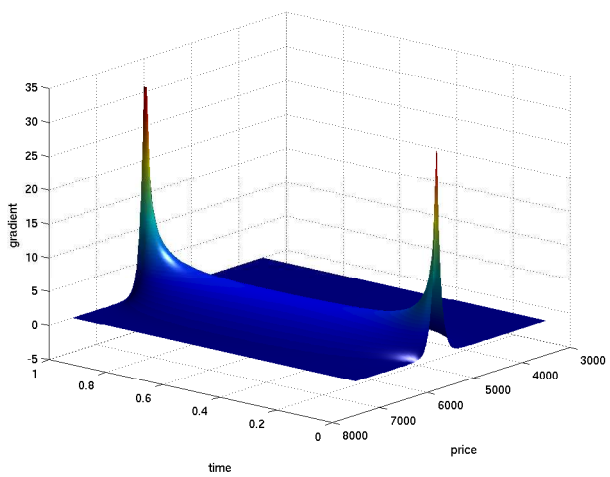

Fig. 1. Gradient (cf. eqn. (8)) of the price of the derivative contract with respect to the volatility surface (squared). Here $r = 4.2\%$, $q = r$, $K = 6000$, $S_0 = 5800$, $\sigma = 20\%$(constant), $T = 1$.

When several option prices are available, we can build an extended cost functional

$$J_e(\sigma) = \sum_{l=1}^{L} (C_l(0; S_0) - C^{market}_{K_l, T_l})^2. \quad (10)$$

Other forms are also possible, for instance the sum of relative errors

$$J_e(\sigma) = \sum_{l=1}^{L} \left(\frac{C_l(0; S_0)}{C^{market}_{K_l, T_l}} - 1 \right)^2 \quad (11)$$

or other weighted variants.

Another setting appears when bid/ask quotes are available i.e. $C_l(0; S_0) \in [C^{bid}_{K_l, T_l}, C^{ask}_{K_l, T_l}]$

$$J_e(\sigma) = \sum_{l=1}^{L} \left[(C_l(0; S_0) - C^{bid}_{K_l, T_l})_+ \right]^2 \quad (12)$$

$$+ \left[(C^{ask}_{K_l, T_l} - C_l(0; S_0))_+ \right]^2. \quad (13)$$

In fact, the final functional to be optimized will contain J_e and also some terms that ensure smoothness; see Section III for details.

The formulas (6)-(9) serve as building bricks for the optimization of this aggregated cost functional. Let us suppose that a simple fixed step ($\rho > 0$) gradient algorithm is used:

$$\sigma^2_{n+1} = \sigma^2_n - \rho \frac{\delta J_e}{\delta(\sigma^2)}(\sigma^2_{n+1}). \quad (14)$$

We see thus immediately that the singularities will propagate into the solution which will have a full list of singularities at $(0, S_0)$ and (T_l, K_l), $l = 1, ..., L$. Such properties are not natural for the local volatility surface $\sigma(t, S)$ and the inversion procedure has to address them. Note that obtaining a smoother local surface is possible because of its underdertermination : in the extreme situation $L = 1$ only one price $C^{market}_{K_l, T_l}$ is available which brings a limited information on the volatility surface that will not be unique; in this case the most natural volatility surface will a constant, equal to the Black-Scholes implied volatility.

A traditional choice to avoid singularities and address the non-uniqueness is to parametrize the surface [3], [4] with a limited number of degrees of freedom; the result will be the optimal surface in the class. The parametrization also helps to keep natural properties of the local volatility surface e.g., finding the implicit volatility should only one option be given in the input.

We propose in this work a variant of the Coleman, Li & Verma [4] procedure with the distinction that here the optimization of the volatility surface, while still performed in the parametric space, is accelerated by the use of an first (and partially second) order approximation of the functional through the use of the adjoint (6).

III. Surface space and the quasi-Newton problem

A. Surface space

Continuing the arguments of the previous section, we set the space of available surface shapes to be interpolations of cubic spine type that have zero values except at some point of the grid (t_i, S_j), cf. Fig. 2; let us call $f_{ij}(S,t)$ such an element; each surface is a linear combination of these shapes $f_{ij}(S,t)$:

$$\sigma(S,t) = \sum \alpha_{ij} f_{ij}(S,t). \tag{15}$$

A possible procedure would be to optimize J_e (e.g. from (10)) expressed as a function of the coefficients α_{ij} of σ in (15). But this dependence is highly nonlinear and the resulting optimization will have many unwanted local extrema.

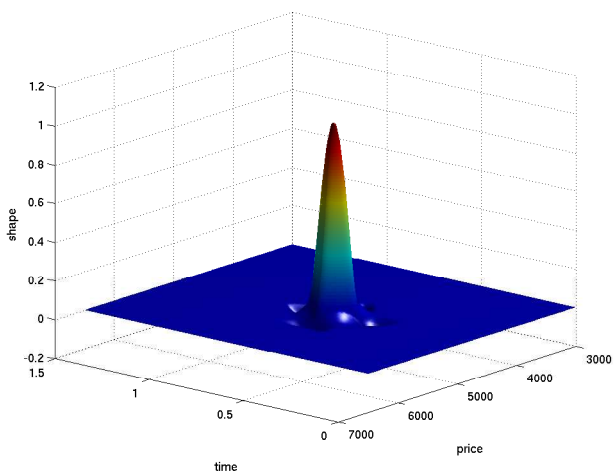

Fig. 2. Typical shape that is used to reconstruct the local volatility surface.

B. Approximation formula

For this discrete space, we can compute the gradient of any derivative contract $C(S,t)$ with respect to variations of the local surface σ inside the admissible surface space. This is in fact just a matter of projecting the exact gradient (9) onto each shape f_{ij}. We obtain an approximation formula around the current local volatility σ:

$$C(\sigma + \sum_{ij} \alpha_{ij} f_{ij}(S,t)) \simeq C(\sigma) \tag{16}$$

$$+ \sum_{ij} < \frac{\partial C}{\partial \sigma}, f_{ij} >_{L^2_{S,t}} \alpha_{ij}. \tag{17}$$

Note that this formula already provides second order information for J_e; we will therefore use a sequentially quasi-Newton iteration [5]; in order to ensure smoothness we add a quadratic term

$$< \alpha, Q\alpha > \tag{18}$$

where the matrix Q will represent the gradient of shapes f_{ij} with respect to t and S.

C. Optimization procedure

The algorithm will be as follows:

0/ choose L^∞ bounds for the local volatility surface :

$$\sigma_{min} \leq \sigma(t,S) \leq \sigma_{max} \tag{19}$$

1/ initialize σ_0

2/ compute the gradient of any derivative price with respect to variations of σ in the admissible space i.e. formula (16);

3/ construct a (quadratic) optimization problem

$$\begin{array}{c} \min \\ A\alpha \leq b \\ \sigma_{min} \leq \alpha \leq \sigma_{max} \end{array} \quad \frac{1}{2}\alpha^T H \alpha + w^t \alpha \tag{20}$$

with H containing the second order part of J_e expressed by using the approximation formula (16) and also the smoothness term (18); should a bid/ask functional (e.g. as in (12)) be used then the problem will not be quadratic any more but (16) is still used; the constraints arise from the requirement that α_k be in $[\sigma_{min}, \sigma_{max}]$; additional contraints, in a "trust-region" style, can be put to remain in a region where the approximation (16) holds. This quadratic problem can be solved by any suitable algorithm; for instance Matlab uses by default a subspace trust-region method based on the interior-reflective Newton method described in [6] (using a preconditioned conjugate gradient).

4/ update the local volatility σ; if the replication error J_e is too high return in 2/ otherwise exit.

IV. Results and discussions

Note that instead of a unique optimization in the parametric space we perform one optimization aroung each current point; this reduces the number of computations of the PDE (3). But, equally importantly, the separation between the optimization and the approximation of the functional provides flexibility in the information that can be fitted, e.g.

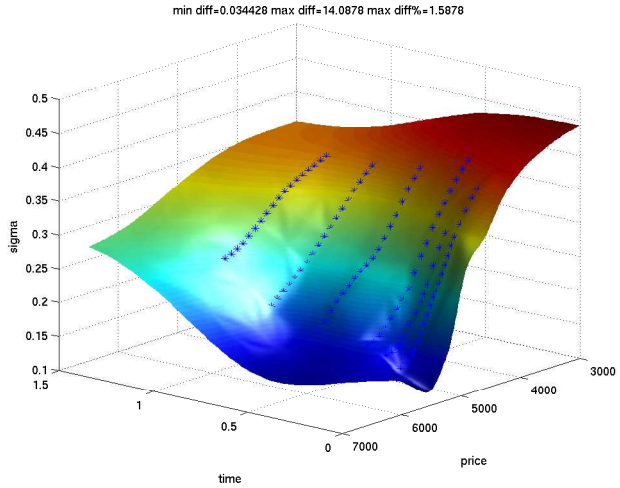

Fig. 3. Local volatility surface of the CAC40 index for 19/02/08 as recovered from the published end-of-data european call option prices. The marks on the surface indicate the option prices that were used to invert: the K_l and T_l, $l = 1, ..., 106$. To account for the dividends we adjusted the prices to refer to the futures although the options are written on the index.

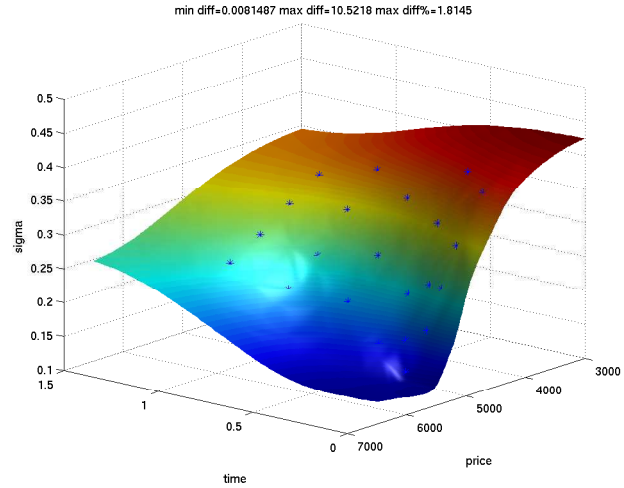

Fig. 4. Local volatility surface of the CAC40 index for 19/02/08 as recovered from the published end-of-data european call option prices. The marks on the surface indicate the option prices that were used to invert: the K_l and T_l.Five times less points with respect to Fig. 3 are used ($L = 22$).

we can readily accomodate any derivative contract (as soon as an gradient formula like (8) exists for it; when it does not one can use Malliavin calculus) such as options on futures, strategies, structured products etc. This is to be put into contrast with Dupire's approach [2] that needs a uniform set of data (e.g. vanilla options) to perform the inversion.

The use of the gradient not in an optimization procedure but to obtain an approximation of the functional around the current point is a acknowledgement of the fact that the main difficulty is not finding a solution but choosing one among all compatible surfaces (i.e. ill-posedness). In practice two conditions ensure a relevant selection of the local volatility surface

- the L^∞ constraints (19); the use of a separate optimization procedure is instrumental in satisfying the constraints

- the smoothness term in eqn. (18); we depart here from the view in [4] to use a parametrization dependent on the number L of available prices; the term (18) is used to avoid tunning the set of shapes f_{ij} and to automatize the process.

We first checked that when $L = 1$ the problem recovers the implicit volatility; it did so with only one cycle 2/-4/. We next moved to real-world examples to analyze the local volatility surface of options on the CAC40 index (cf www.euronext.com). The convergence is attained after, typically, 10 to 20 iterations. A first test in Fig. 3 indicates that fine structures on the volatility surface are recovered if enough prices ($L = 106$) are available for the inversion; but if only a few prices are to be used ($L = 22$ in Fig. 4 and $L = 10$ in Fig. 5) the overall shape remains very stable. For comparison we also draw in Fig. 6 the implicit volatility surface.

REFERENCES

[1] M. Musiela and M. Rutkowski, *Martingale methods in financial modelling*, 2nd ed., ser. Stochastic Modelling and Applied Probability. Berlin: Springer-Verlag, 2005, vol. 36.

[2] J. Hull, *Options, Futures, and Other Derivatives*, 6th ed. Prentice Hall, 2006.

[3] Y. Achdou and O. Pironneau, *Computational methods for option pricing*, ser. Frontiers in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2005, vol. 30.

[4] T. F. Coleman, Y. Li, and A. Verma, "Reconstructing the unknown local volatility function [J. Comput. Finance **2** (1999), no. 3, 77–100]," in *Quantitative analysis in financial markets*. World Sci. Publ., River Edge, NJ, 2001, pp. 192–215.

[5] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, *Numerical optimization*, 2nd ed., ser. Universitext. Berlin: Springer-Verlag, 2006, theoretical and practical aspects.

[6] T. F. Coleman and Y. Li, "A reflective Newton method for minimizing a quadratic function subject to bounds on some of the variables," *SIAM J. Optim.*, vol. 6, no. 4, pp. 1040–1058, 1996.

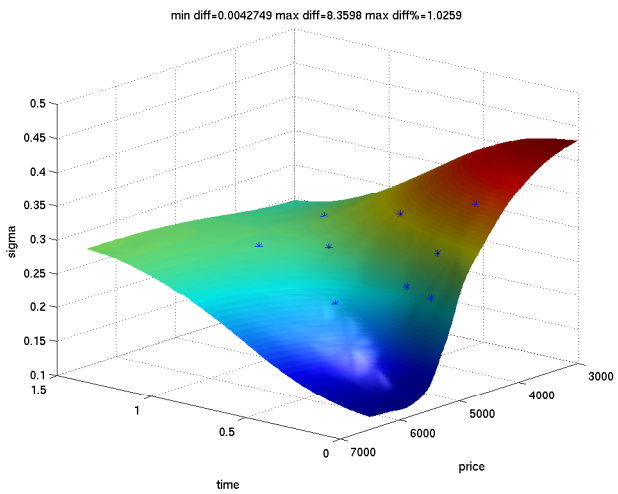

Fig. 5. Local volatility surface of the CAC40 index for 19/02/08 as recovered from the published end-of-data european call option prices. The marks on the surface indicate the option prices that were used to invert: the K_l and T_l. Ten times less points with respect to Fig. 3 are used ($L = 10$).

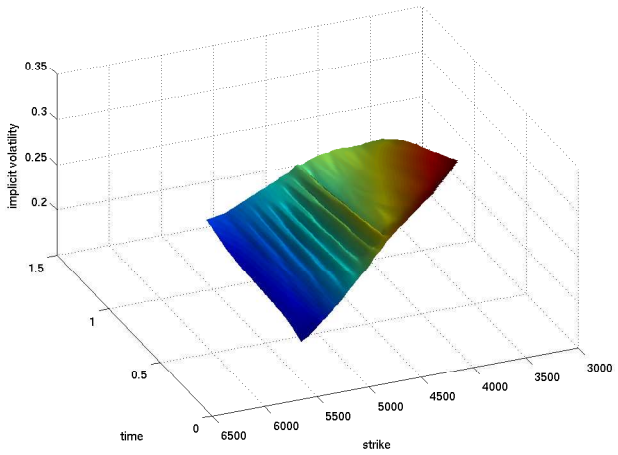

Fig. 6. Implicit volatility surface for the example in Fig. 3-5.