Calibration of Derivative Pricing Models

Christoph Reisinger

June 2018

Introduction

Motivation

What is calibration?

- Derivative pricing models depend on parameters: Black-Scholes σ, interest rate r, Heston reversion speed κ, etc.
- For a *given* parameter θ, one can compute model prices of derivatives, say, $P(\theta; S_0, t_0)$, conditional on the value of the underlying stock at time t_0, S_0.
- *But*: parameters are *a priori* unknown.
- $\longrightarrow$ Observe them indirectly via traded market prices $P^\star$.
- Calibration seeks to identify θ from the requirement

$$P(\theta; S_0, t_0) = P^\star,$$

where P and $P^\star$ can be vectors.

Motivation

Why calibrate?

- Calibration can be used to match the prices of liquidly traded contracts.
- It interpolates between prices, say for different strikes, maturities;
- to discover the true (?) model,
- to price more 'exotic' derivatives.
- Perhaps most importantly, it allows us to use the model to translate prices into hedge parameters.
- ...

Model-free density estimation

- Under no-arbitrage, option prices are discounted expectations under a risk-neutral measure.
- A key building block is the risk-neutral probability measure.
- The (risk-neutral) transition $p(S, t; S', T)$ can be viewed in two ways:
  1. If S and t are fixed (today's spot price and date), we can regard $p(S, t; S', T)$ as the probability density that at time $T > t$ the spot price will be S'.
  2. If S' and T are fixed (some given value of the spot price and date, say) then $p(S, t; S', T)$ is the probability density that at time $t < T$ the spot price *was* S.

  p is governed by the (forward and backward) Kolmogorov equations, which are model-specific (see later).

Breeden-Litzenberger formula

The price C of a European call option with strike K and expiration T can be written as

$$\begin{aligned} C(S,t;K,T) &= e^{-r(T-t)} \int_0^\infty p(S,t;S',T)\,C(S,T;K,T)\,\mathrm{d}S' \\ &= e^{-r(T-t)} \int_K^\infty p(S,t;S',T)(S'-K)\,\mathrm{d}S', \end{aligned}$$

and from this

$$\begin{aligned} \frac{\partial C}{\partial K} &= -e^{-r(T-t)} \int_K^\infty p\,\mathrm{d}S', \\ \frac{\partial^2 C}{\partial K^2} &= e^{-r(T-t)} p(S,t;K,T). \end{aligned}$$

Risk-neutral densities

- Matlab example: numerical risk-neutral densities from calls

Main models

- The Black-Scholes model assumes geometric Brownian motion

$$\frac{dS_t}{S_t} = \mu_t \, dt + \sigma \, dW_t$$

with a constant volatility parameter σ.

- The local volatility model allows a function

$$\frac{dS_t}{S_t} = \mu_t \, dt + \sigma(S_t, t) \, dW_t.$$

- The Heston stochastic volatility model with a volatility process

$$\begin{aligned} dS_t &= \mu_t S_t \, dt + \sqrt{v_t} S_t \, dW_t, \\ dv_t &= \kappa(\theta - v_t) \, dt + \xi \sqrt{v_t} \, dW_t^v. \end{aligned}$$

- The local-stochastic volatility (LSV) model combines both (see later).

Practicalities

Demands:

- Fast direct solver
  - will need to generate prices for many different strikes, maturities, as part of some sort of iteration;
  - ideally analytic or semi-analytic formulae should be available,
  - e.g., Fourier-based methods for *affine* models
- Stable parameter sensitivities
  - search directions of solvers usually based on gradients
  - Monte Carlo parameter 'bumping' can be noisy
  - use algorithmic differentiation, adjoint methods etc
- Good initial guess
  - solver may only converge locally
  - parameter ballpark may be unknown
  - use guess based on analytic (asymptotic) approximation

Issues

Potential, more fundamental problems:

- May not be able to fit all observed prices:

$$\{\theta : \; P(\theta; S_0, t_0) = P^\star\} = \emptyset$$

- May not have enough observed prices to determine all parameters:

$$|\{\theta : \; P(\theta; S_0, t_0) = P^\star\}| > 1$$

- Estimation may be unstable:

$$P^{-1}(P^\star; S_0, t_0) \;\text{ not continuous}$$

Trade-off between goodness of fit and stability of estimators, Greeks etc.

Outline

- Parametric models
  - Black-Scholes and implied volatility
  - Heston
  - Best-fit and iterative solvers
- Local volatilities
  - Breeden-Litzenberger and the probability density function
  - Dupire local volatility and the adjoint problem
  - Inverse and ill-posed problems, regularisation
  - Local-stochastic volatility
- Techniques
  - Monte Carlo and particle methods
  - Finite differences
  - Application

Parametric models

Black-Scholes model and PDE

The Black-Scholes model assumes geometric Brownian motion

$$\frac{\mathrm{d}S}{S} = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W$$

with a constant volatility parameter σ. The PDE for the no-arbitrage price of a European option in this model is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

In the simplest case of a vanilla call, there is a terminal condition

$$V(S,T) = \max(S-K,0).$$

Black-Scholes formula

This problem can be solved *analytically* and the solution is

$$V(S,t) = S\,N(d_1) - e^{-r(T-t)} K\,N(d_2),$$

where

$$d_1 = \frac{\log(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2 = d_1 - \sigma\sqrt{T-t}.$$

This gives us a mapping

$$V = V(S,t;K,T;\sigma,r)$$

from the asset price S and the parameters to the option price.

Implied volatility

- For an externally given price $V^\star$, the (Black-Scholes) *implied volatility* is the (constant) volatility parameter $\sigma^\star$, which, if inserted in the Black-Scholes model, gives this price.
- This definition only makes sense if such a parameter value exists, and is unique. From

$$\frac{\partial \mathcal{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \mathcal{V}}{\partial S^2} + rS\frac{\partial \mathcal{V}}{\partial S} - r\mathcal{V} = -\sigma S^2 \Gamma$$

  with $\mathcal{V}(S, T) = 0$ it can be shown that for convex payoffs the implied volatility is unique.
- In the case of European calls this means inverting the Black-Scholes formula with respect to σ.
- Moreover, the implied volatility is the same for calls and puts (put-call-parity).

Nonlinear iteration

Review iterative solvers:

- Fixed-point iteration:
- (linear) convergence for contractions.
- Newton-Raphson method:
- local, quadratic convergence.

Empirical evidence

- The volatility σ is a parameter of the model for the stock (the Black-Scholes model), and not of the option contract.
- If we believe in the model, we should expect to get the same implied volatility independent of what strike and expiry the option has that we use to identify σ.

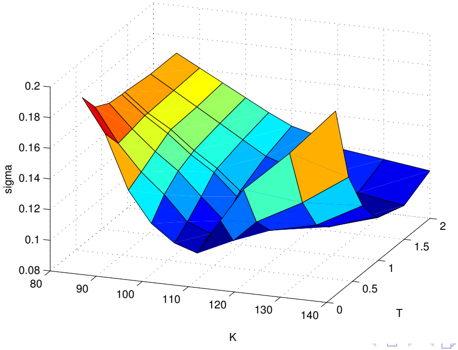
Implied volatility for S&P 500 index call options.

T/K	85	90	95	100	105	110	115	120	130
0.175	0.190	0.168	0.133	0.113	0.102	0.097	0.120	0.142	0.169
0.425	0.177	0.155	0.138	0.125	0.109	0.103	0.100	0.114	0.130
0.695	0.172	0.157	0.144	0.133	0.118	0.104	0.100	0.101	0.108
0.940	0.171	0.159	0.149	0.137	0.127	0.113	0.106	0.103	0.100
1.000	0.171	0.159	0.150	0.138	0.128	0.115	0.107	0.103	0.099
1.500	0.169	0.160	0.151	0.142	0.133	0.124	0.119	0.113	0.107
2.000	0.169	0.161	0.153	0.145	0.137	0.130	0.126	0.119	0.115

S&P 500 volatility surface

It is evident that the *volatility surface* is not flat as predicted by the Black-Scholes model. We often see both a

- dependence on K, termed *smile* or *skew*, and a
- dependence on T, referred to as *term structure*.



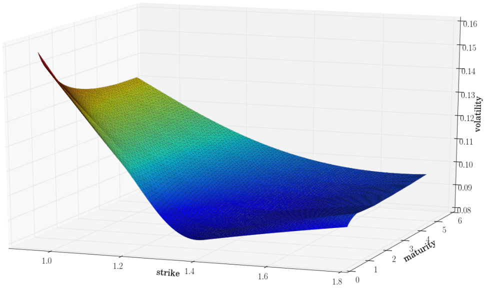
FX volatility surface

- The EURUSD spot value on 28/03/2013 is $S_0 = 1.2837$.
- Quoted are for each maturity 5 volatilities on a delta scale,

$$10\text{D-Put}, 25\text{D-Put}, 50\text{D}, 25\text{D-Call}, 10\text{D-Call},$$

- and the following maturities

$$3\text{M}, 6\text{M}, 1\text{Y}, 2\text{Y}, 3\text{Y}, 4\text{Y}, 5\text{Y}.$$



Heston model

- As a *parametric* example, we study the Heston stochastic volatility model

$$
\begin{aligned}
dS_t &= rS_t\,dt + \sqrt{v_t}S_t\,dW_t, \\
dv_t &= \kappa(\theta - v_t)\,dt + \xi\sqrt{v_t}\,dW_t^v,
\end{aligned}
$$

  where Y_t is the *variance process* of the asset. $\xi > 0$ governs the volatility of variance, and $\kappa, \theta > 0$ the mean reversion.
- The Wiener processes W_t, W_t^v have correlation ρ.
- The variance process is non-negative by construction.
- Under the *Feller condition* $2\kappa\theta > \xi^2$, the variance process stays strictly positive.
- The distribution at time t is non-central ξ^2.

Heston model properties

- The log-process is *affine*.
- A hedging argument gives the pricing PDE

$$\frac{\partial C}{\partial t}+\frac{1}{2}\left[S^2 v\frac{\partial^2 C}{\partial S^2}+2\rho\gamma vS\frac{\partial^2 C}{\partial S\partial v}+\xi^2 v\frac{\partial^2 C}{\partial v^2}\right]+rS\frac{\partial C}{\partial S}+\kappa(\theta-v)\frac{\partial C}{\partial v}-rC=0$$

Heston model properties

- The *characteristic function*

$$\phi(u,t) = \mathbb{E}\left[\exp\left(iu\log(S_t)\right)\right]$$

is analytically known as

$$\begin{aligned}
\phi(u,t) &= \exp(iu(\log(S_0)+(r-q)t)) \times \exp\left(\theta\kappa\xi^{-2}((\kappa-\rho\xi iu+d)t) - 2\log((1-ge^{dt})/(1-g))\right) \\
&\quad \times \exp\left(v_0\xi^{-2}(\kappa-\rho\xi iu+d)(1-ge^{dt})/(1-ge^{dt})\right) \\
d &= \sqrt{(\rho\xi ui-\kappa)^2+\xi^2(iu+u^2)} \\
g &= (\kappa-\rho\xi iu+d)/(\kappa-\rho\xi iu-d)
\end{aligned}$$

- This allows for a semi-closed-from solution for calls as

$$\begin{aligned}
C(K,T) &= \frac{\exp(-\alpha\log K)}{\pi}\int_0^\infty \exp(-iv\log K)\varrho(v)\,\mathrm{d}v \\
\varrho(v) &= \frac{\exp(-rT)\phi(v-(\alpha+1)i,T)}{\alpha^2+\alpha-v^2+i(2\alpha+1)v}
\end{aligned}$$

(α arbitrary but such that $\alpha+1$st moment exists.)

Fitted model

- Matlab example: Heston model implied vol and market fit

Local volatility

Local volatility

- One interpretation of the observed data is that the volatility is not constant but depends on the value of the stock.
- A model that accounts for this dependence on state and time is
$$\frac{\mathrm{d}S}{S} = \mu\,\mathrm{d}t + \sigma(S,t)\,\mathrm{d}W.$$
The function $(S,t) \to \sigma(S,t)$ is called *local volatility*.
- The risk neutral transition density satisfies

$$\begin{aligned} \frac{\partial p}{\partial T} &= \tfrac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma(S',T)^2 S'^2 p\right) - \frac{\partial}{\partial S'}\left(rS'p\right) \\ -\frac{\partial p}{\partial t} &= \tfrac{1}{2}\sigma(S,t)^2 S^2 \frac{\partial^2 p}{\partial S^2} + rS\frac{\partial p}{\partial S} \end{aligned}$$

Local volatility

The model still describes a complete market, where a standard hedging argument determines the price of a European option as the solution of the PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2} S^2 \sigma^2(S,t) \frac{\partial^2 V}{\partial S^2} + r\left(S\frac{\partial V}{\partial S} - V\right) = 0.$$

For any choice of the local volatility function $\sigma(.,.)$, there is a unique arbitrage-free price

$$V = V(S,t;K,T;\sigma(.,.),r).$$

In general, numerical solution is necessary. This defines a *parameter-to-solution* map

$$\Psi:\ \sigma(.,.) \to V(.,.), \quad \text{or} \quad \Phi:\ \sigma(.,.) \to \widehat{\sigma}(.,.).$$

(the dependence on r is suppressed).

Dupire formula

Using the integral formula for call prices and Kolmogorov forward equation, the value function

$$V = V(S, t; K, T),$$

satisfies a PDE in T and K (!),

$$\frac{\partial V}{\partial T} + rK\frac{\partial V}{\partial K} = \tfrac{1}{2}K^2\sigma^2(K,T)\frac{\partial^2 V}{\partial K^2}.$$

Re-arranging,

$$\sigma(S,t;K,T)^2 = \frac{(\partial V/\partial T)(S,t;K,T) + rK(\partial V/\partial K)(S,t;K,T)}{\tfrac{1}{2}K^2(\partial^2 V/\partial K^2)(S,t;K,T)}$$

Recall that, in practice, when we compute $\sigma(S,t;K,T)$, today's spot price S and date t are fixed. We can vary only the strike K and the maturity T.

Discussion

Practical problems with this approach:

- requires continuum of strikes and maturities (interpolation, extrapolation)
- numerical differentiation is ill conditioned
- the nominator $\partial^2 V/\partial K^2$ tends to zero for $K \to \infty$

The last problem can be circumvented to some extent by switching from quoted prices to implied vols

$$\sigma(K,T)^2 = \frac{\hat\sigma^2 + 2\hat\sigma\,(T-t)\frac{\partial\hat\sigma}{\partial T} + 2r\,\hat\sigma\,K(T-t)\frac{\partial\hat\sigma}{\partial K}}{(1+K\,d_1\sqrt{T-t}\,\frac{\partial\hat\sigma}{\partial K})^2 + \hat\sigma\,(T-t)\,K^2\left(\frac{\partial^2\hat\sigma}{\partial K^2} - d_1\left(\frac{\partial\hat\sigma}{\partial K}\right)^2\sqrt{(T-t)}\right)}$$

where, as usual,

$$d_1 = \frac{\log(S/K) + (r+\frac{1}{2}\hat\sigma^2)(T-t)}{\hat\sigma\sqrt{T-t}}.$$

Iterative solution (A. Rehai, Risk Magazine, April, 2006)

Can use this for an iterative procedure to solve

$$\Phi(\sigma) = \widehat{\sigma}^*,$$

the market implied volatility.

- Set $\sigma^0 = \widehat{\sigma}^*$.
- Then define, for $k \geq 0$,

$$\sigma^{k+1} = \sigma^k \cdot \frac{\widehat{\sigma}^*}{\Phi(\sigma^k)} = \phi(\sigma^k).$$

- Any fixed-point σ^* solves $\Phi(\sigma^*) = \widehat{\sigma}^*$.
- The map ϕ is found to be a contraction.
- In practice, only few iterations needed.

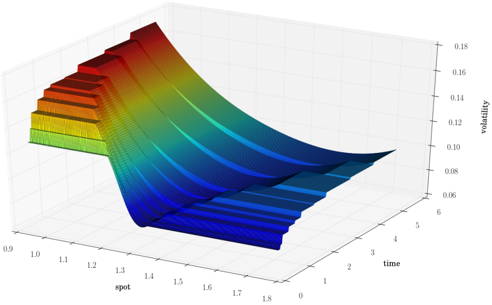
Calibrated FX local volatility surface

- The EURUSD spot value on 28/03/2013 is $S_0 = 1.2837$.
- Quoted are for each maturity 5 volatilities on a delta scale,

$$10\text{D-Put}, 25\text{D-Put}, 50\text{D},\ 25\text{D-Call},\ 10\text{D-Call},$$

- and the following maturities

$$3\text{M},\ 6\text{M},\ 1\text{Y},\ 2\text{Y},\ 3\text{Y},\ 4\text{Y},\ 5\text{Y}.$$



Noise magnification

- Matlab example: local vol without/with small data noise.

An ill-posed problem

Parameter identification problems are often inherently ill-posed.

A problem is called *well-posed* (in the sense of Hadamard), if

1. a solution exists,
2. the solution is unique, and
3. the solution depends continuously on the data.

A problem that is not well-posed is called *ill-posed*.

When speaking of 'inverse' problems, a natural reflex is to ask "inverse to what?".

A simple instructive example of a direct/inverse problem pair are differentiation and integration.

Direct/inverse problems

- Integration:

$$F(x) = \int_0^x f(z)\,\mathrm{d}z$$

If $f = 0$ with some 'noise' f_ϵ with $|f_\epsilon(x)| < \epsilon$ for $x \in [0,1]$, then for the perturbed integral F_ϵ

$$|F(x) - F_\epsilon(x)| \leq \left| \int_0^x f_\epsilon(z)\,\mathrm{d}z \right| \leq \epsilon.$$

- Differentiation:

$$f(x) = \frac{\mathrm{d}F}{\mathrm{d}x}(x)$$

For $F = 0$ and noise of the form $F_\epsilon(x) = \epsilon \sin(kx)$

$$|f(x) - f_\epsilon(x)| \approx \epsilon k,$$

which can be arbitrarily large, even if $\max_{x\in[0,1]} |F_\epsilon(x)| \leq \epsilon$.

Numerical differentiation

If we replace the derivative by a finite difference,

$$\frac{\mathrm{d}F}{\mathrm{d}x}(x) \approx \frac{F(x+h)-F(x)}{h} + c \cdot h + \ldots$$

In the presence of noise

$$\begin{aligned}\frac{\mathrm{d}F}{\mathrm{d}x}(x) &\approx \frac{F_\epsilon(x+h)-F_\epsilon(x)}{h} + c \cdot h + \ldots \\ &\approx \frac{F(x+h)-F(x)}{h} + \frac{\epsilon}{h} + c \cdot h + \ldots\end{aligned}$$

The finite stepsize h regularises the ill-posed problem, it acts as regularisation parameter. There is a tradeoff between approximation and stability. If the 'noise level' ϵ is known, the optimal choice is when $\epsilon/h \sim h$, i.e. $h \sim \sqrt{\epsilon}$.

Regularisation

For an abstract ill-posed problem

$$F(x) = y, x \in X, y \in Y, \tag{1}$$

remedies often fall in the following classes.

1. Non-existence: find the best fit, i.e. replace (1) by an optimisation problem

$$\|F(x) - y\|_Y \to \min_{x \in X}$$

2. Non-uniqueness: choose a particular solution, e.g.

$$\|x - x_0\|_X \to \min_{x \in X}$$

   for an initial guess x_0.
3. Continuous dependence on data: combine the above two to

$$\|F(x) - y\|_Y^2 + \lambda \|x - x_0\|_X^2 \to \min_{x \in X}$$

Example

Many examples found in the literature have the form

$$\sum_{i=1}^{N} w_i |V(S_0, 0; K_i, T_i; \sigma) - V_i|^2 + \lambda \|\sigma\|^2 \to \min$$

with some positive weights w_i and $\lambda > 0$, where V_i are quoted prices for various strike/maturity pairs (K_i, T_i). They differ in their choice of $\|.\|$, and parametrisation of σ.

Jackson, Süli, Howison choose

$$\|\sigma\|^2 = \int_0^T \int_0^\infty \left(\frac{\partial\sigma}{\partial S}\right)^2 + \left(\frac{\partial\sigma}{\partial t}\right)^2 \, \mathrm{d}S \, \mathrm{d}t$$

with greater weight for liquid options (short dated, close to the money). σ is represented by cubic splines in S and piecewise linear in t, with more points around the money, constant extrapolation in the far range. The solution of the direct problem is computed by adaptive finite elements.

Dual approach

Egger uses the Dupire PDE

$$\begin{aligned} \frac{\partial V}{\partial T} &= \frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 V}{\partial K^2} - rK\frac{\partial V}{\partial K} \\ V(S,t;K,t) &= (K-S)_+ \end{aligned}$$

solved again by finite elements, and

$$\|\sigma\|^2 = \int_0^\infty \left(\frac{\partial \sigma}{\partial S}\right)^2 + \left(\frac{\partial^2 \sigma}{\partial S^2}\right)^2 \, \mathrm{d}S$$

with σ from a class of cubic splines. A proof of stability and convergence rates are given. The optimisation problem is solved by a BSGF quasi-Newton method.

Transformation

Following H. Berestycki, J. Busca, and I. Florent: Asymptotics and calibration of local volatility models, Quantitative Finance (2002), we now address the problematic regions $t \to T$, $S \to 0$ and $S \to \infty$.

It is convenient to consider the transformations $x = \log(S/K) + r\tau$, $\tau = T - t$, and $v(x,\tau) = \exp(r\tau)C(S, T-\tau; K, T)/K$, such that

$$Lv = v_\tau - \frac{1}{2}\sigma^2(x,\tau)(v_{xx} - v_x) = 0$$
$$v(x,0) = (\mathrm{e}^x - 1)_+$$

where $\sigma(x,t)$ is actually $\sigma(K\exp(x - r(T-t)), T-t)$.

For $\sigma = 1$ constant,

$$u_\tau = \frac{1}{2}(u_{xx} - u_x).$$

Implied volatility

Note that if $u(x,\tau)$ is a solution to

$$u_\tau = \frac{1}{2}(u_{xx} - u_x),$$

then $u(x,\tau\mu)$ with $\mu > 0$ is a solution to the problem with volatility $\sqrt{\mu}$.

Therefore, the implied volatility $\phi(x,\tau)$ is implicitly given via

$$v(x,\tau) = u(x, \tau\phi^2(x,\tau)).$$

Some calculus gives

$$Lv = u_\tau(x,\tau\phi^2) F(x,\tau,\phi,\phi_x,\phi_{xx})$$

with

$$F(x,\tau,\phi,\phi_x,\phi_{xx}) = \phi^2 - \sigma^2\left(1 - x\phi_x/\phi\right)^2 + 2\tau\phi\phi_\tau - \sigma^2\tau\phi^2\phi_{xx} + \sigma^2\tau^2\phi^2\phi_x^2/4$$

Therefore, ϕ is the solution of the quasi-linear equation $F = 0$.

Asymptotics

We can set $\tau = 0$ to 'formally' get the limiting equation

$$\begin{array}{rrcl} & \phi^2 & = & \sigma^2 \left(1 - x\phi_x/\phi\right)^2 \\ \Leftrightarrow & \frac{1}{\sigma(x,0)} & = & \frac{\phi - x\phi_x}{\phi^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x}{\phi}\right) \\ \Leftrightarrow & \int_0^x \frac{\mathrm{d}s}{\sigma(s,0)} & = & \frac{x}{\phi} \end{array}$$

or

$$\frac{1}{\phi(x,0)} = \int_0^1 \frac{\mathrm{d}s}{\sigma(xs,0)}.$$

Based on this asymptotic result, define the calibration functional as

$$\sum_{i=1}^{N} \left(\phi(x_i,\tau_i)^{-1} - \phi_i^{-1}\right)^2 + \lambda \, \|\nabla(\sigma^{-1})\|_2^2,$$

where ϕ_i are observed implied volatilities, and $\phi(x_i,\tau_i)$ are implied from the model for these strike/maturity pairs.

Markovian projections

Gyöngy's result

- Consider a model of the form

$$\frac{dS_t}{S_t} = r\,dt + \eta_t\,dW_t,$$

where η is a fairly arbitrary volatility process subject only to some technical conditions (mostly, $\int_0^t \eta_u^2 S_u^2\,du$ of bounded expectation).

- Then the law (density) of S_t agrees with that of the local volatility model provided

$$\mathbb{E}[\eta_t^2 | S_t = S] = \sigma^2(S,t).$$

Discussion

- If the density of the processes agrees, the prices of European options (e.g., calls and puts) agree.
- Hence for any continuous semi-martingale model there is a local volatility model indistinguishable from vanilla options.
- We can use the local volatility as a "code book", an equivalent way to quote vanilla prices, in the same way as we use the (Black-Scholes) implied volatility.

Local-stochastic volatility (LSV)

- Local volatility can fit all prices exactly.
- The dynamics of spot and vol is incorrect.
- Stochastic volatility gives an imperfect fit.
- It gives a realistic spot-vol dynamics.
- The Local-stochastic volatility model seeks to combine both advantages.

Local-stochastic volatility (LSV)

- Combining stochastic and local volatility,

$$
\begin{aligned}
\frac{\mathrm{d}S_t}{S_t} &= r\,\mathrm{d}t + \sigma(S_t, t)\sqrt{v_t}\,\mathrm{d}W_t, \\
\mathrm{d}v_t &= \kappa(\theta - v_t)\,\mathrm{d}t + \xi\sqrt{v_t}\,\mathrm{d}W_t^v,
\end{aligned}
$$

  with local and stochastic volatility.

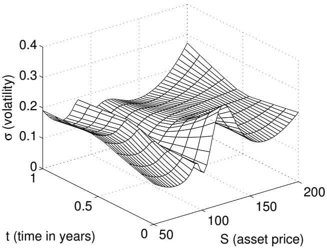
- To be consistent with European call/put option prices, it has to hold that

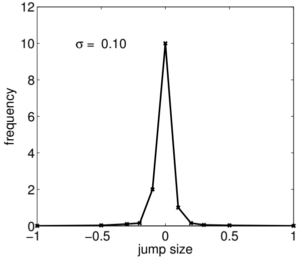
$$
\sigma_{LV}^2(K, T) = \sigma(K, T)^2\,\mathbb{E}\left[v_T \middle| S_T = K\right],
$$

  where σ_{LV} is the Dupire local volatility function calibrated to the calls and puts.

A warning – model uncertainty

- A local volatility model, jump diffusion model, and (Heston) stochastic volatility model calibrated to 60 observed European calls for different strike/maturity pairs within 3 basis points.





Parameter	rate of reversion	long run variance	volatility of volatility	correlation	initial variance
Value	0.0745	0.1415	0.1038	-0.2127	0.0167

- The value of an up-and-out barrier call with strike 90% and barrier 110% of the spot varies by 177 basis points.

Techniques

Path simulation – local volatility

- For Black-Scholes, we have exactly

$$S_{n+1} = S_n \exp\left((r - \sigma^2/2)\Delta t + \sigma \Delta W_n\right),$$

where $\Delta t \equiv t_{n+1} - t_n$ and $\Delta W_n \equiv W_{t_{n+1}} - W_{t_n} \sim N(0, \Delta t)$.

- The Euler-Maruyama scheme is

$$S_{n+1} = S_n \left(1 + r\Delta t + \sigma \Delta W_n\right).$$

- For local volatility, use either Euler-Maruyama or log-Euler

$$S_{n+1} = S_n \exp\left((r - \sigma^2(S_{t_n}, t)/2)\Delta t + \sigma(S_{t_n}, t)\Delta W_n\right),$$

which preserves positivity and is exact for constant σ.

Path simulation – Heston and LSV

- First, approximate the variance process, e.g. by the full-truncation Euler scheme

$$v_{n+1} = v_n + \kappa \left(\theta - v_n^+\right) \Delta t + \xi \sqrt{v_n^+} \Delta W_n^v,$$

where $x^+ = \max(x, 0)$.

- Note the standard Euler-Maruyama scheme generates negative variances.
- Then approximate S, e.g. by log-Euler

$$S_{n+1} = S_n \exp\left((r - v_n^+/2)\Delta t + \sqrt{v_n^+} \Delta W_n\right),$$

which is again positive. Similarly, for SLV

$$S_{n+1} = S_n \exp\left((r - v_n^+ \sigma^2(S_{t_n}, t)/2)\Delta t + \sqrt{v_n^+} \sigma(S_{t_n}, t) \Delta W_n\right).$$

Monte Carlo expectations

Want to compute an expected European payoff

$$\mathbb{E}\left[P(S_T)\right].$$

- Simulate $\widehat{S}_N$ for $T = N\Delta t$, to generate M independent samples $\widehat{S}_N^{(i)}$.
- Then define the estimator

$$\widehat{P} = \frac{1}{M}\sum_{i=1}^{M} P(\widehat{S}_N^{(i)}).$$

- The error consists of bias (neglect here) and sampling noise.
- Convergence for $M \to \infty$ by LLN and error bounds by CLT.

Monte Carlo conditional expectations – particle method

- From the model, the conditional expectation can be estimated by Markovian projection.
- A particle method estimator is

$$\widehat{\sigma}(S,t) = \frac{\sum_{i=1}^{N} \eta_t^{(i)} \delta_\epsilon(S_t^{(i)} - S)}{\sum_{i=1}^{N} \delta_\epsilon(S_t^{(i)} - S)},$$

where $(S_t^{(i)}, \eta_t^{(i)})$ is a sample of (S_t, η_t).

- The kernel function δ has small bandwidth ϵ, e.g.

$$\delta_\epsilon(S) = \frac{1}{\sqrt{2\pi\epsilon}} \exp\left(-\frac{S^2}{2\epsilon^2}\right).$$

- Samples are obtained by Monte Carlo simulation.

Finite differences – local volatility

We think of the θ-scheme as an approximation of the PDE at time $t^m - \theta\Delta t$,

$$\left[\frac{\partial V}{\partial t}\right]_{S=S_j}^{t=t_m-\theta\Delta t} + \left[\frac{1}{2}S^2\sigma^2(S,t)\frac{\partial^2 V}{\partial S^2}\right]_{S=S_j}^{t=t_m-\theta\Delta t} + \left[rS\frac{\partial V}{\partial S}\right]_{S=S_j}^{t=t_m-\theta\Delta t} - [rV]_{S=S_j}^{t=t_m-\theta\Delta t} = 0$$

Approximate the PDE by setting e.g.

$$[rV]_{S=S_j}^{t=t_m-\theta\Delta t} = \theta rV(S_j, t_{m-1}) + (1-\theta)rV(S_j, t_m)$$

and similarly for the first S-derivative.
If

$$\left[\frac{\partial V}{\partial t}\right]_{S=S_j}^{t=t_m-\theta\Delta t} = \frac{V(S_j, t_m) - V(S_j, t_{m-1})}{\Delta t},$$

the scheme is of second order accurate in Δt if $\theta = \frac{1}{2}$ and of first order otherwise.

Finite differences – local volatility

For the term containing σ, we could either use

$$\left[\frac{1}{2}S^2\sigma^2(S,t)\frac{\partial^2 V}{\partial S^2}\right]_{S=S_j}^{t=t_m-\theta\Delta t} = \theta\,\frac{1}{2}S_j^2\sigma^2(S_j,t_{m-1})\left[\frac{\partial^2 V}{\partial S^2}\right]_{S=S_j}^{t=t_{m-1}} + (1-\theta)\,\frac{1}{2}S_j^2\sigma^2(S_j,t_m)\left[\frac{\partial^2 V}{\partial S^2}\right]_{S=S_j}^{t=t_m}$$

with

$$\left[\frac{\partial^2 V}{\partial S^2}\right]_{S=S_j}^{t=t_m} = \frac{V(S_j+\Delta S,t_m) - 2V(S_j,t_m) + V(S_j-\Delta S,t_m)}{\Delta S^2},$$

or

$$\left[\frac{1}{2}S^2\sigma^2(S,t)\frac{\partial^2 V}{\partial S^2}\right]_{S=S_j}^{t=t_m-\theta\Delta t} = \frac{1}{2}S_j^2\sigma^2(S_j,t_m-\theta\Delta t)\left\{\theta\left[\frac{\partial^2 V}{\partial S^2}\right]_{S=S_j}^{t=t_{m-1}} + (1-\theta)\left[\frac{\partial^2 V}{\partial S^2}\right]_{S=S_j}^{t=t_m}\right\}$$

Finite differences – forward equations

- For the computation of (conditional) expectations, the probability density is useful.
- This requires a numerical solution of the Kolmogorov forward equation.
- Can approximate

$$\left[\frac{1}{2}\frac{\partial^2}{\partial S^2}S^2\sigma^2(S,t)V\right]_{S=S_j}^{t=t_m-\theta\Delta t} = \theta\,\frac{1}{2}\left[\frac{\partial^2}{\partial S^2}S^2\sigma^2(S,t)V\right]_{S=S_j}^{t=t_{m-1}} + (1-\theta)\,\frac{1}{2}\left[\frac{\partial^2}{\partial S^2}S^2\sigma^2(S,t)V\right]_{S=S_j}^{t=t_m},$$

where

$$\left[\frac{\partial^2}{\partial S^2}S^2\sigma^2(S,t)V(s,t)\right]_{S=S_j}^{t=t_m} = \frac{(S_j+\Delta S)^2\sigma^2(S_j+\Delta S,t_m)V(S_j+\Delta S,t_m) - 2S_j^2\sigma^2(S_j,t)V(S_j,t_m) + \dots}{\Delta S^2}.$$

- The Dirac initial datum needs care (e.g., Rannacher start-up).

Computing gradients

- In the optimisation, need to compute derivatives

$$\nabla_\theta \sum_{i=1}^{n} |V(S,t;\theta) - V_i^*|^2 = 2\nabla_\theta V(S,t;\theta) \sum_{i=1}^{n} (V(S,t;\theta) - V_i^*)$$

- The pricing method must be able to return stable parameter sensitivities.
- Can differentiate PDE with respect to parameters and solve extra PDEs.
- Monte Carlo:
  - Pathwise sensitivities
  - Likelihood ratios

Generic tools

- 'Bumping'
  - Evaluate algorithm for parameters θ and $\theta + \delta\theta$
  - approximate derivative by $\frac{V(\theta+\delta\theta)-V(\theta)}{\delta\theta}$
  - potentially large error magnification; trade off with bias
  - inefficient for large number of parameters
- Algorithmic differentiation
  - break code down into elementary instructions ($+$,$\times$,...)
  - differentiate by a giant chain rule
  - automatic software tools available
- AAD – Adjoint algorithmic differentiation
  - evaluate the chain rule in reverse direction
  - ultimately, backward product of transposed matrices
  - efficient when many inputs, i.e. high-dimensional gradient
  - complexity always within 3-4 times that of function evaluation

Example: Local-stochastic volatility

- Whiteboard example LSV: first fix the Heston parameters, then fine-tune by 'boot-strapping' the function σ; see Ren, Madan and Qian (2007).