Review: empirical evidence

- The volatility $\sigma$ 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 strike and expiry.

Implied volatility for S&P 500 index call options.
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

\[
\frac{dS}{S} = \mu \, dt + \sigma(S, t) \, dW.
\]

The function \((S, t) \rightarrow \sigma(S, t)\) is called local volatility.
- The risk neutral transition density satisfies

\[
\frac{\partial p}{\partial T} = \frac{1}{2} \frac{\partial^2}{\partial S'^2} \left( \sigma(S', T)^2 S'^2 p \right) - \frac{\partial}{\partial S'} \left( rS' p \right)
\]

\[
- \frac{\partial p}{\partial t} = \frac{1}{2} \sigma(S, t)^2 S^2 \frac{\partial^2 p}{\partial S^2} + rS \frac{\partial p}{\partial S}
\]
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 $\sigma(\cdot, \cdot)$, there is a unique arbitrage-free price

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

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

$$\sigma(\cdot, \cdot) \rightarrow V(\cdot, \cdot).$$

(the dependence on $r$ is suppressed).
Finite differences

What is different to the Black-Scholes PDE, where coefficients only depended on $S$, is the time-dependence. We think of the $\theta$-scheme as an approximation of the PDE at time $t^m - \theta \Delta t$,

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

As earlier, 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 $\Delta t$ if $\theta = \frac{1}{2}$ and of first order otherwise.
For the term containing $\sigma$, 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

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

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

Re-arranging,

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

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 $\frac{\partial^2 V}{\partial K^2}$ tends to zero for $K \to \infty$

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

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

where, as usual,

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

\[ F(x) = \int_0^x f(z) \, dz \]

If \( f = 0 \) with some ‘noise’ \( f_\epsilon \) with \( |f_\epsilon(x)| < \epsilon \) for \( x \in [0,1] \), then for the perturbed integral \( F_\epsilon \)

\[ |F(x) - F_\epsilon(x)| \leq \left| \int_0^x f_\epsilon(z) \, dz \right| \leq \epsilon. \]

Differentiation:

\[ f(x) = \frac{dF}{dx}(x) \]

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 made arbitrarily large, even if we satisfy

\[ \max_{x \in [0,1]} |F_\epsilon(x)| \leq \epsilon. \]
Numerical differentiation

If we replace the derivative by a finite difference,

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

In the presence of noise

$$\frac{dF}{dx}(x) \approx \frac{F_\epsilon(x + h) - F_\epsilon(x)}{h} + c \cdot h + \ldots$$

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

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’ $\epsilon$ is known, the optimal choice is when $\epsilon/h \sim h$, i.e. $h \sim \sqrt{\epsilon}$. 
Regularisation

For an abstract ill-posed problem

\[ F(x) = y, \ x \in X, \ y \in Y, \] (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

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

with some \( \lambda > 0 \).
Example

Many examples found in the literature have the form

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

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 $\|\cdot\|$, and parametrisation of $\sigma$.

Jackson, Süli, Howison choose

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

with greater weight for liquid options (short dated, close to the money). $\sigma$ 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.
Egger uses the Dupire PDE

\[
\frac{\partial V}{\partial T} = \frac{1}{2} \sigma^2(K, T)K^2 \frac{\partial^2 V}{\partial K^2} - rK \frac{\partial V}{\partial K}
\]

\[V(S, t; K, t) = (K - S)_+\]

solved again by finite elements, and

\[\|\sigma\|^2 = \int_0^\infty \left( \frac{\partial \sigma}{\partial S} \right)^2 + \left( \frac{\partial^2 \sigma}{\partial S^2} \right)^2 \, dS\]

with \(\sigma\) 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
\[ x = \log \left( \frac{S}{K} \right) + r \tau, \]
\[ \tau = T - t, \] and
\[ \nu(x, \tau) = \exp(r \tau) C(S, T - \tau; K, T)/K, \] such that
\[
Lv = \nu_{\tau} - \frac{1}{2} \sigma^2(x, \tau)(\nu_{xx} - \nu_x) = 0
\]
\[ \nu(x, 0) = (e^x - 1)_+ \]

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

\[\nu(x, \tau) = u(x, \tau \phi^2(x, \tau)).\]

Some calculus gives

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

with

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

Therefore, \( \phi \) is the solution of the quasi-linear equation \( F = 0. \)
We can set $\tau = 0$ to ‘formally’ get the limiting equation

\[
\phi^2 = \sigma^2 (1 - x\phi_x/\phi)^2
\]

$\Leftrightarrow$ \[
\frac{1}{\sigma(x,0)} = \frac{\phi - x\phi_x}{\phi^2} = \frac{d}{dx} \left( \frac{x}{\phi} \right)
\]

$\Leftrightarrow$ \[
\int_0^x \frac{ds}{\sigma(s,0)} = \frac{x}{\phi}
\]

or

\[
\frac{1}{\phi(x,0)} = \int_0^1 \frac{ds}{\sigma(xs,0)}.
\]

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 $\phi_i$ are observed implied volatilities, and $\phi(x_i, \tau_i)$ are implied from the model for these strike/maturity pairs.
Real-world FX calibration example

- The spot value is denominated as the price of one EUR in USD units
  \[
  \text{Spot} = 1.1271
  \]

- Let the discounted call option price under an arbitrage-free model for a notional of one unit of EUR be
  \[
  C(K, T) = \mathbb{E}^{Q^d} \left[ D_d^T (S_T - K)^+ \right].
  \]

- The implied volatility is retrieved for the Deltas
  10D-Put, 25D-Put, 50D, 25D-Call, 10D-Call

- And the maturity pillars
  3W, 1M, 2M, 3M, 6M, 1Y, 1Y6M, 2Y, 3Y, 5Y

- The domestic risk-neutral measure \( Q^d \) is the one associated with the USD money market account. The foreign risk-neutral measure \( Q^f \) is the one associated to the EUR money market account.
Market volatility surface for the EUR-USD (18/03/2016).
We recall the Dupire PDE, forward in maturity and with the strike as space variable,

\[
\begin{align*}
\frac{\partial C_{LV}(K,T)}{\partial T} &+ (r^d(T) - r^f(T)) K \frac{\partial C_{LV}(K,T)}{\partial K} + r^f(T) C_{LV}(K,T) \\
&- \frac{1}{2} K^2 \frac{\partial^2 C_{LV}(K,T)}{\partial K^2} \sigma^2_{LV}(K,T) = 0, \\
C_{LV}(K,0) &=(S_0 - K)^+, \quad C_{LV}(0,T)=S_0, \quad C_{LV}(S_{max},T)=0.
\end{align*}
\]

(2)

where \(r^d(t) = -\frac{\partial \ln P^d(0,t)}{\partial t}\) and \(r^f(t) = -\frac{\partial \ln P^f(0,t)}{\partial t}\), with \(P^d/f(0,T)\) the market zero coupon bond prices for the domestic and foreign money market accounts, respectively.
Calibration by fixed-point algorithm

\[
\text{for } ( i = 1 ; i \leq N_{\text{Maturities}} ; i ++ )
\]

\[\textbf{\texttt{// while error > tol}}\]

\[\textbf{\texttt{// solve Dupire PDE on } [T_{i-1}, T_i]}\]

\[\textbf{\texttt{// compute model implied vol } \Sigma_{\text{Model}} \text{ for maturity } T_i \text{ from the computed call prices}}\]

\[\textbf{\texttt{// compute}}\]

\[\text{error} = \sum_{m=1}^{M_i} (\Sigma_{\text{Model}} (K_{T_i,m}, T_i) - \Sigma_{\text{Target}} (K_{T_i,m}, T_i))^2\]

\[\textbf{\texttt{// for } ( j = 1 ; j \leq M_i ; j ++ )}\]

\[\textbf{\texttt{// update local volatility guess from state } n \text{ to } n + 1}\]

\[\sigma^{n+1}_{\text{LV}} (K_{T_i,j}, T_i) = \sigma^n_{\text{LV}} (K_{T_i,j}, T_i) \frac{\Sigma_{\text{Target}} (K_{T_i,j}, T_i)}{\Sigma_{\text{Model}} (K_{T_i,j}, T_i)}\]

\[\textbf{\texttt{// endfor}}\]

\[\textbf{\texttt{// endwhile}}\]

\[\textbf{\texttt{// endfor}}\]
As stated in [Regai, 2006], the map $f (\{\sigma_{LV}\}) \rightarrow \{\sigma_{LV}\} \ast \left(\frac{\Sigma_{Target}}{\Sigma_{Model}}\right)$ is contracting.
The calibrated risk-neutral density
A probabilistic angle

▶ Given a stochastic process $X = (X_t)_{t \geq 0}$, and parameter(s) $\theta \in \Theta$.

▶ Observe a price process $S = (S_t)_{t \geq 0}$ such that $S_t = S(t, (X_u)_{u \leq t}, \theta)$.

▶ Let the time $t$ value of an option with payoff $P$, expiry $T$, be

$$V_t(\theta) = \mathbb{E}[B(t, T)P(S_T) | S_t]$$

▶ Suppose at time $t \in [0, T]$ we observe a set of such option prices $\{V_t^{(i)}(\theta) : i \in I_t\}$ possibly with noise $\{e_t^{(i)} : i \in I_t\}$,

$$Y_t^{(i)} = V_t^{(i)}(\theta) + e_t^{(i)} \quad \text{for} \quad i \in I_t.$$

▶ The calibration problem is to find $\theta$ that best reproduces observed prices $\{Y_t^{(i)} : i \in I_t, t \in \{t_1, \ldots, t_n\}\}$. 

Matthieu Mariapragassam and Christoph Reisinger
Calibration Lecture 3: Local Volatility
Bayesian theory

- Assume we have some prior information for $\theta$ (for example that it is positive, or is a smooth surface).
- Summarise this by a prior density $p(\theta)$ for $\theta$.
- We observe some noisy data

$$Y_t = V_t(\theta) + e_t$$

for $t \in \{t_1, \ldots, t_n\}$.
- $p(Y|\theta)$ is the probability of observing the data $Y$ given $\theta$, called the likelihood function.
- Bayes rule gives the posterior density of $\theta$,

$$p(\theta | Y) = \frac{p(Y|\theta) \, p(\theta)}{p(Y)}$$

- $p(Y)$ is given by $p(Y) = \int p(Y|\theta) \, p(\theta) \, d\theta$. 
Estimation and uncertainty

- Parameter estimates can be found from the posterior distribution.
- Example: The mean value $\mathbb{E}[\theta|Y]$ with respect to the posterior.
- Example: The maximum a posteriori (MAP) estimator $\text{argmax}\{p(\theta|Y)\}$.
- The posterior also contains information about the parameter uncertainty, which can be used to construct measures of model risk.
Combining ‘smoothness’ prior and likelihood functions, we get the posterior explicitly as

\[ p(\theta|V) \propto \exp \left\{ -\frac{1}{2\delta^2} \left[ \lambda \|\theta - \theta_0\|^2_\kappa + G(\theta) \right] \right\}, \]

\[ G(\theta) = \frac{10^8}{S_0^2} \sum_{i \in I} w_i \left| f_0^{(i)}(\theta) - V_0^{(i)} \right|^2. \]

Can construct estimators from \( p(\theta|V) \), eg, the MAP estimator

\[ \theta_{MAP}(V) = \arg \max_{\theta \in \Theta}\{ p(\theta|V) \}. \]

The Bayesian approach reformats Tikhonov regularisation methods into a unified framework, as is already noted by Fitzpatrick (1991).
We price 66 European call options on the local volatility surface given in Jackson, Suli and Howison (1999), with 11 strikes and 6 maturities.

To each of the prices we add Gaussian noise with mean zero and standard deviation 0.1% of the original price, and treat these as the observed prices, similar to the approach in Coleman, Li and Verma (2001) and Hamida and Cont (2005).

We take the calibration error acceptance level to be $\delta = 3$ basis points following Jackson, Suli and Howison (1999).
We discretise $\sigma(S, t)$ on a grid:

$S_{\text{min}} = s_1 < \ldots < s_j < \ldots < s_J = S_{\text{max}}$ and 
$0 = t_1 < \ldots < t_l < \ldots < t_L = t_{\text{max}}$.

This defines a parameter vector

$$\theta = (\log \sigma_1, \ldots, \log \sigma_m, \ldots, \log \sigma_M),$$

and a spline interpolant $\Theta(\cdot, \cdot)$ of $\theta$.

We take nodes positioned on the grid given by

$s = 2500, 4000, 4500, 4750, 5000, 5250, 5500, 7000, 10000,$

$t = 0.0, 0.5, 1.0,$

so $J = 9$, $L = 3$ with a total of $M = J \times L = 27$ parameters.
MCMC simulation

- Select a starting point $\theta_0$ for which $p(\theta_0|V) > 0$.

- For $r = 1, \ldots, n$, sample a proposal $\theta^\#$ from a symmetric jumping distribution $J(\theta^\#|\theta_{r-1})$ and set

$$
\theta_r = \begin{cases} 
\theta^\# & \text{with probability } \min \left\{ \frac{p(\theta^\#|V)}{p(\theta_{r-1}|V)}, 1 \right\}, \\
\theta_{r-1} & \text{otherwise}.
\end{cases}
$$

- Run $m$ parallel chains, starting from different points $\theta_0^{(j)}$.

- Discard the first $b$ iterations of the run (known as burn-in).

- Keep every $k$th draw from the remaining iterations (thinning).
Take the jumps of a random walk whose transition kernel is associated with the prior, Beskos & Stuart (2007).

Let $A^{-1}$ be the inverse non-singular covariance matrix. By Cholesky decomposition $A = BB^T$, the jump is

$$\theta' = \theta + \sqrt{2du} B \xi,$$

where $\xi \sim N(0, I_M)$ and $du$ is a step size.

The value of $du$ is chosen so that the acceptance rate is close to the optimum value of 23% found by Gelman et al (2004).
Figure: Using MCMC, 479 surfaces from the posterior distribution were sampled (from 16 chains of 10000 surfaces) and are plotted with the same degree of transparency. The true surface is plotted in opaque black.
The Bayesian model average price of a new contract is given by

$$\frac{1}{N} \sum_{i=1}^{N} f(\theta_i) \approx \int_{\theta} f(\theta) p(\theta|V) \, d\theta.$$ 

**Figure:** Prices for up-and-out barrier call option strike 5000 ($S_0 = 5000$), barrier 5500 and maturity 3 months. Included are the true price with an assumed bid-ask spread of 6 basis points, the MAP/Tikhonov price, and the Bayes price with its associated posterior pdf.
Robustness
Knots versus options

Figure: Prices for an up-and-out barrier call option with strike 5000 ($S_0 = 5000$), barrier 5500 and maturity 3 months for different number of calibration options and different number of spline knots.