Local Volatility Model

QF607 Numerical Methods

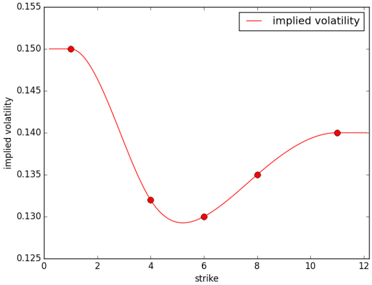
Zhenke Guan
zhenkeguan@smu.edu.sg

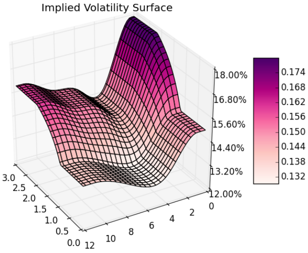
Outline

- Black-Scholes Term Structure Model I , II
- Local Volatility Model
  - Derivation using Fokker-Planck Equation
  - Local Volatility Model - Strike Arbitrage/Calendar Arbitrage

Black-Schole Model Is Inadequate

- Black-Scholes formula gives options prices as a function of implied volatility
- The Black-Scholes model is not consistent with the market such that for options with different strike and maturity have different implied volatility





Black-Scholes Term Structure Model I

- First attempt to extend the Black-Scholes model is to make it term structured:

$$\frac{dS}{S} = \mu(t)dt + \sigma(t)dW \tag{1}$$

  where $\mu(t) = r_d(t) - r_f(t)$ is the drift and $\sigma(t)$ is the instantaneous volatility that is time dependent
- Note the difference between $\sigma(t)$ and the implied volatility $\sigma_{iv}(t)$
- Instantaneous volatility is the parameter of the model, implied volatility is the parameter for Black-Scholes formula, they are equal only when the model is Black-Scholes, not for Black-Scholes term structure model
- Linking them by the total variance $v(T)$ we have

$$v(T) = \sigma_{iv}(T)^2 T = \int_0^T \sigma(t)^2 dt \tag{2}$$

Black-Scholes Term Structure Model II

- Assume there is no volatility smile in the market, but ATM implied volatility for different maturity differs
- That is, the market quotes the implied ATMVOL for several pillars: $\sigma_{iv}(T_i)$ for $i \in [1, \ldots, N]$
- The instantaneous volatility in between each pillar dates can then be assumed constant
- For $t \in [T_i, T_{i+1}]$, we have

$$\sigma(t) = \sqrt{\frac{\sigma_{iv}^2(T_{i+1})T_{i+1} - \sigma_{iv}^2(T_i)T_i}{T_{i+1} - T_i}} \tag{3}$$

- (3) is also consistent with our temporal interpolation method: linear in variance — constant $\sigma(t)$ results in linear growth of variance with respect to time
- Now we have a model and its diffusion process generates us consistent price for ATM options of different maturities, including the interpolated maturities

Dependency On Strikes

- The dependence of implied volatility on the strike is trickier
- But the overall goal is the same — build a diffusion process such that
  - It is compatible with the whole implied volatility surface
  - It is self consistent
- Some models resort to an extra source of randomness, e.g., stochastic volatility (Heston, SABR), or jump process, to achieve the smile effect.
  - Additional risk factors, not friendly to some of our numerical integrator (e.g., Tree or PDE)
  - The additional diffusion process is normally not hedge-able through market instruments
  - Limited degree of freedom, not necessarily recover exactly the vanilla prices, i.e., the calibration error could be larger

Local Volatility Model

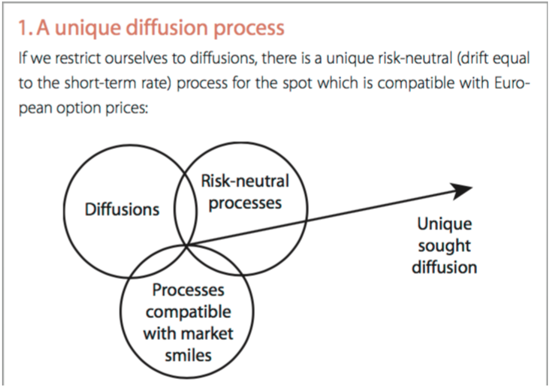
If we want to restrict ourselves to the simple one-factor form, the natural extension to the Black-Scholes term structure model would be:

$$\frac{dS}{S} = \mu(t)dt + \sigma(S,t)dW \tag{4}$$

- The instantaneous volatility itself is a surface dependent on spot and time — **local volatility surface**
- There is abundant degree of freedom to match the implied volatility surface: so $\sigma_{iv}(K, T)$
- The remaining questions are
  - **How do we calculate** $\sigma(S,t)$?
  - **Is it unique?**
- The questions were answered by Bruno Dupire [3], so the model is also called **Dupire model**.

Local Volatility Model

- The uniqueness of the local volatility surface is claimed in the Dupire's paper to be due to the risk neutrality [3]



1. A unique diffusion process

If we restrict ourselves to diffusions, there is a unique risk-neutral (drift equal to the short-term rate) process for the spot which is compatible with European option prices:

- The derivation of the local volatility requires the celebrated **Fokker-Planck** equation

Fokker Planck Equation

Given a general form of stochastic process X_t:

$$dX_t = a(X_t, t)dt + b(X_t, t)dW. \tag{5}$$

Suppose $g(x)$ is an arbitrary function satisfying $\lim_{x\to\pm\infty} g(x) = 0$. Denote the process $G_t = g(X_t)$. Since G_t is only a function of X_t, applying Itô lemma we have

$$dG_t = \frac{\partial g}{\partial X} dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial X^2} dX_t^2 \tag{6}$$

yielding

$$dG_t = \left(\frac{\partial g}{\partial X} a(X_t, t) + \frac{1}{2} b^2(X_t, t) \frac{\partial^2 g}{\partial X^2} \right) dt + \frac{\partial g}{\partial X} b(X_t, t) dW \tag{7}$$

Fokker Planck Equation

Taking the expectation on both side, and note that the diffusion part has 0 expectation:

$$dE(G_t) = E\left[\frac{\partial g}{\partial X}a(X_t, t) + \frac{1}{2}b^2(X_t, t)\frac{\partial^2 g}{\partial X^2}\right]dt \tag{8}$$

$$\frac{dE(G_t)}{dt} = \int p(x, t)\left(a(x, t)\frac{\partial g}{\partial x} + \frac{1}{2}b^2(x, t)\frac{\partial^2 g}{\partial x^2}\right)dx \tag{9}$$

where $p(x, t)$ is the probability density function of X_t.

Fokker Planck Equation

Extending the left hand side we have

$$E[G_t] = \int p(x,t)g(x)dx \quad \Rightarrow \quad \frac{dE[G_t]}{dt} = \int \frac{\partial p(x,t)}{\partial t} g(x)dx \tag{10}$$

Integrating by parts the two components of the right hand side and recall that $\lim_{x\to\pm\infty} g(x) = 0$, we can derive

$$\frac{dE(G_t)}{dt} = -\int \frac{\partial[p(x,t)a(x,t)]}{\partial x} g(x)dx + \frac{1}{2}\int \frac{\partial^2[b^2(x,t)p(x,t)]}{\partial x^2} g(x)dx. \tag{11}$$

Integration by parts

$$\int u dv = uv - \int v du$$

Fokker Planck Equation

Equalizing (10) and (11):

$$\int \frac{\partial p(x,t)}{\partial t} g(x) dx = -\int \frac{\partial [p(x,t)a(x,t)]}{\partial x} g(x) dx + \frac{1}{2} \int \frac{\partial^2 [b^2(x,t)p(x,t)]}{\partial x^2} g(x) dx \qquad (12)$$

and the arbitrariness of $g(x)$ yields the **Fokker-Planck equation**

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial [p(x,t)a(x,t)]}{\partial x} + \frac{1}{2} \frac{\partial^2 [p(x,t)b^2(x,t)]}{\partial x^2}. \qquad (13)$$

It is basically an PDE that governs the forward transition probability density function $p(x, t; x_0, t_0)$ with the initial value a Dirac delta function at x_0 (because there is no uncertainty at (t_0, x_0)):

$$p(x, t_0) = \delta_{x_0}(x).$$

Fokker Planck Equation With Local Volatility Model

Under a local volatility model

$$\frac{dS}{S} = \mu(t)dt + \sigma(S,t)dW_t \tag{14}$$

we have $a(S,t) = S\mu(t)$ and $b(S,t) = S\sigma(S,t)$ in the Fokker-Planck equation (13):

$$\frac{\partial p(s,t)}{\partial t} = -\frac{\partial[s\mu(t)p(s,t)]}{\partial s} + \frac{1}{2}\frac{\partial^2[s^2\sigma^2(s,t)p(s,t)]}{\partial s^2} \tag{15}$$

Probability Density Function

The probability density function $p(s,t)$ is known from the implied volatility surface:

$$C(K,T) = E[max(S-K,0)] = \int_K^{\infty} p(s,T)(s-K)ds \tag{16}$$

where $C(K,T)$ is the undiscounted call option price struck at K with maturity T. And (see appendix page 5)

$$\frac{\partial^2 C(K,T)}{\partial K^2} = p(K,T) \tag{17}$$

Local Volatility Derivation — Drift-less Case

For simplicity, we consider only the drift-less case ($\mu(t) = 0$), where

$$\frac{1}{2}\frac{\partial^2[s^2\sigma^2(s,t)p(s,t)]}{\partial s^2} = \frac{\partial p(s,t)}{\partial t} \quad \text{— Fokker Planck equation} \tag{18}$$

Change the variables from (s, t) to (K, T), and replace $p(K, T)$ at the right hand side by $\frac{\partial^2 C(K,T)}{\partial K^2}$ we have

$$\frac{1}{2}\frac{\partial^2[K^2\sigma^2(K,T)p(K,T)]}{\partial K^2} = \frac{\partial}{\partial T}\left[\frac{\partial^2 C(K,T)}{\partial K^2}\right] = \frac{\partial^2}{\partial K^2}\left[\frac{\partial C(K,T)}{\partial T}\right]$$

Integrating twice with respect to K and eliminate the first order K term and the constants due to the asymptotic behaviour of European call prices, we have

$$\frac{1}{2}K^2\sigma^2(K,T)\frac{\partial^2 C(K,T)}{\partial K^2} = \frac{\partial C(K,T)}{\partial T}$$

The formula of local volatility is thus

$$\sigma(K,T) = \sqrt{\left[2\frac{\partial C(K,T)}{\partial T}\right] \Big/ \left[K^2\frac{\partial^2 C(K,T)}{\partial K^2}\right]} \tag{19}$$

Local Volatility And Arbitrage

The denominator of the local volatility formula:

$$\left[K^2 \frac{\partial^2 C(K,T)}{\partial K^2}\right]$$

- K^2 times the second derivative of the call option price with respect to the strike, or the density function of the marginal distribution
- If the call option price as a function of strike is non-convex, meaning the density is negative, there is a **strike arbitrage**, also called negative butterfly in practice
- Local volatility is ill defined at the non-convex point so requires the denominator to be floored and effectively the local volatility being capped

Local Volatility And Arbitrage

The numerator of the local volatility formula:

$$\left[2\frac{\partial C(K,T)}{\partial T}\right]$$

- Twice the dual theta
- If the denominator becomes negative, it means given other attributes the same, the option with longer maturity is cheaper — **calendar arbitrage** under the drift-less assumption
- When there is calendar arbitrage, the denominator needs to be floored and so is the local volatility

When the local volatility is floored or capped, the calibration error will be larger than normal — implied volatility surface is arbitrage-able so violating the model assumption

Local Volatility — General Case

- The local volatility formula for non-zero drift:

$$\sigma(K, T) = \sqrt{2\left[\frac{\partial C(K,T)}{\partial T} + (r_d - r_f)K\frac{\partial C(K,T)}{\partial K} + r_f C\right] / \left[K^2 \frac{\partial^2 C(K,T)}{\partial K^2}\right]} \tag{20}$$

- Only the numerator changes — and so is the precise definition of calendar arbitrage

- The derivation can be found in [2].

Local Volatility Formula In Terms Of Implied Volatility

In practice, since the FX markets are quoted by implied volatilities, the local volatility formula is normally translated to be based on implied volatility (see [1] and Appendix A):

$$\sigma(K,T) = \sqrt{\frac{\sigma_{iv}^2 + 2T\sigma_{iv}\frac{\partial \sigma_{iv}}{\partial T} + 2(r_d - r_f)KT\sigma_{iv}\frac{\partial \sigma_{iv}}{\partial K}}{(1 + Kd_1\sqrt{T}\frac{\partial \sigma_{iv}}{\partial K})^2 + K^2T\sigma_{iv}(\frac{\partial^2 \sigma_{iv}}{\partial K^2} - d_1\sqrt{T}(\frac{\partial \sigma_{iv}}{\partial K})^2)}} \quad (21)$$

where

$$d_1 = \frac{\ln\frac{F}{K} + \frac{1}{2}\sigma_{iv}^2 T}{\sigma_{iv}\sqrt{T}} \quad (22)$$

Implementation

In our implementation of implied volatility, we simply need to provide the functions to calculate $\frac{\partial \sigma_{iv}}{\partial T}$, $\frac{\partial \sigma_{iv}}{\partial K}$, and $\frac{\partial^2 \sigma_{iv}}{\partial K^2}$.

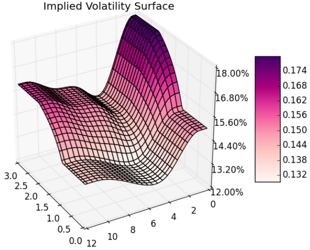
```
class ImpliedVol:
    ...
    def dVoldK(self, t, k):
        return (self.Vol(t, k+0.01) - self.Vol(t, k-0.01)) / 0.02
    def dVoldT(self, t, k):
        return (self.Vol(t+0.005, k) - self.Vol(t, k)) / 0.005
    def dVol2dK2(self, t, k):
        return (self.Vol(t, k+0.01) + self.Vol(t, k-0.01) - 2*self.Vol(t, k))
    / 0.0001

```

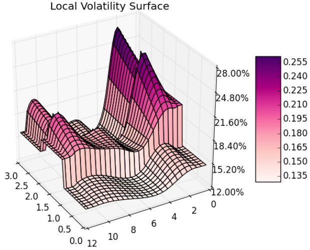
Finite difference or using the analytic derivatives of cubic spline

Local Volatility Implementation

```python
class LocalVol:
    def __init__(self, iv, S0, rd, rf):
        self.iv = iv
        self.S0 = S0
        self.rd = rd
        self.rf = rf
    def LV(self, t, s):
        if t < 1e-6:
            return self.iv.Vol(t, s)
        imp = self.iv.Vol(t, s)
        dvdk = self.iv.dVoldK(t, s)
        dvdt = self.iv.dVoldT(t, s)
        d2vdk2 = self.iv.dVol2dK2(t, s)
        d1 = (math.log(self.S0/s) + (self.rd-self.rf)*t + imp * imp * t / 2) / imp / math.sqrt(t)
        numerator = imp*imp + 2*t*imp*dvdt + 2*(self.rd-self.rf)*s*t*imp*dvdk
        denominator = (1+s*d1*math.sqrt(t)*dvdk)**2 + s*s*t*imp*(d2vdk2 - d1 * math.sqrt(t) * dvdk * dvdk)
        localvar = min(max(numerator / denominator, 1e-8), 1.0)
        if numerator < 0: # floor local volatility
            localvar = 1e-8
        if denominator < 0: # cap local volatility
            localvar = 1.0
        return math.sqrt(localvar)
```

What Does Local Volatility Surface Look Like





- The jumps at pillars are mainly due to discontinuities of $\frac{\partial \sigma_{iv}}{\partial T}$
- The local volatility along the strike axis is smooth because our cubic spline interpolator is C^2

References

I. Clark. *Foreign Exchange Option Pricing: A Practitioners Guide*. Wiley, 2010.

E. Derman and I. Kani. *Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility*. International Journal of Theoretical and Applied Finance, 1998.

B. Dupire. *Pricing with a smile. Risk*, (7):18-20, 1994.

Appendix A

Building blocks to derive (21) from (20)

$$\frac{\partial^2 C(T)}{\partial K^2} = \frac{\partial^2 [Se^{-r_f T} N(d_1) - Ke^{-r_d T} N(d_2)]}{\partial K^2}$$

$$d_1 = \frac{\ln \frac{F}{K} + \frac{1}{2}\sigma_{iv}^2 T}{\sigma_{iv}\sqrt{T}}, \quad d_2 = \frac{\ln \frac{F}{K} - \frac{1}{2}\sigma_{iv}^2 T}{\sigma_{iv}\sqrt{T}}$$

$$\frac{\partial d_1}{\partial K} = \frac{\ln F}{-\sigma^2\sqrt{T}}\frac{\partial \sigma}{\partial K} - (\frac{1}{K\sigma\sqrt{T}} - \frac{\ln K}{\sigma^2\sqrt{T}}\frac{\partial \sigma}{\partial K}) + \frac{1}{2}\sqrt{T}\frac{\partial \sigma}{\partial K} = \frac{\partial \sigma}{\partial K}\frac{-d_2}{\sigma} - \frac{1}{K\sigma\sqrt{T}}$$

$$\frac{\partial d_2}{\partial K} = \frac{\ln F}{-\sigma^2\sqrt{T}}\frac{\partial \sigma}{\partial K} - (\frac{1}{K\sigma\sqrt{T}} - \frac{\ln K}{\sigma^2\sqrt{T}}\frac{\partial \sigma}{\partial K}) - \frac{1}{2}\sqrt{T}\frac{\partial \sigma}{\partial K} = \frac{\partial \sigma}{\partial K}\frac{-d_1}{\sigma} - \frac{1}{K\sigma\sqrt{T}}$$

$$\frac{\partial N(d_1)}{\partial K} = \phi(d_1)\frac{\partial d_1}{\partial K}, \quad \frac{\partial KN(d_2)}{\partial K} = N(d_2) + K\phi(d_2)\frac{\partial d_2}{\partial K}$$

$$\frac{\partial^2 N(d_1)}{\partial K^2} = \phi(d_1)\frac{\partial^2 d_1}{\partial K^2} + \frac{\partial \phi(d_1)}{\partial d_1}\left(\frac{\partial d_1}{\partial K}\right)^2 = \phi(d_1)\frac{\partial^2 d_1}{\partial K^2} - \phi(d_1)d_1\left(\frac{\partial d_1}{\partial K}\right)^2$$

$$\frac{\partial^2 KN(d_2)}{\partial K^2} = \phi(d_2)\frac{\partial d_2}{\partial K} + K\phi(d_2)\frac{\partial^2 d_2}{\partial K^2} + \frac{\partial d_2}{\partial K}\left(\phi(d_2) - Kd_2\phi(d_2)\frac{\partial d_2}{\partial K}\right)$$

Appendix B: Smile Arbitrage

Two conditions to ensure there is no smile arbitrage:

1. European call prices are monotonically decreasing with respect to the strike:

$$C(S_0, K_1, T, \sigma(K_1), r, q) \geq C(S_0, K_2, T, \sigma(K_2), r, q) \quad \text{if} \quad K_1 < K_2 \tag{23}$$

2. The European call price as a function of strike has to be convex every where: for any three points $K_1 < K_2 < K_3$

$$\frac{C(K_2) - C(K_1)}{K_2 - K_1} < \frac{C(K_3) - C(K_2)}{K_3 - K_2} \tag{24}$$

or

$$C(K_2) < C(K_3)\frac{K_2 - K_1}{K_3 - K_1} + C(K_1)\frac{K_3 - K_2}{K_3 - K_1} \tag{25}$$

This is also equivalent to saying "butterfly price has to be non-negative"

When Could Smile Arbitrage Happen?

- Recall that call price is the expectation of payoff under risk neutral measure

$$C(K) = e^{-rT}\mathbb{E}_{\mathbb{Q}}[\max(S-K,0)] \tag{26}$$

- And expectation is an integral over the probability density function $p(S)$

$$C(K) = e^{-rT}\int_{K}^{+\infty}(S-K)p(S)dS \tag{27}$$

- The first non-arbitrage condition translates to

$$\begin{aligned}&C(K_1) - C(K_2) && (28)\\ =&e^{-rT}\left[\int_{K_1}^{K_2}(S-K_1)p(S)dS + \int_{K_2}^{+\infty}(K_2-K_1)p(S)dS\right] && (29)\end{aligned}$$

which is positive by definition

- Similarly, the second non-arbitrage condition translates to

$$C(K_3)\frac{K_2 - K_1}{K_3 - K_1} + C(K_1)\frac{K_3 - K_2}{K_3 - K_1} - C(K_2) \tag{30}$$

$$= \frac{K_3 - K_2}{K_3 - K_1}\int_{K_1}^{K_2}(S - K_1)p(S)dS + \frac{K_2 - K_1}{K_3 - K_1}\int_{K_2}^{K_3}(K_3 - S)p(S)dS \tag{31}$$

  which is also positive by definition

- So, when could smile arbitrage happen? — when the probability density does not exist.

- If we can start with valid probability density function $p(S)$, arbitrage-freeness is guaranteed by construction

- Note that the second derivative of call price is the probability density

$$
\begin{aligned}
\frac{dC}{dK} &= d\frac{\int_K^\infty Sp(S)dS}{dK} - d\frac{K\int_K^\infty p(S)dS}{dK} \\
&= -Kp(K) - \left(\int_K^\infty p(S)dS - Kp(K)\right) \\
&= -\int_K^\infty p(S)dS
\end{aligned}
$$

So

$$
\frac{d^2C}{dK^2} = p(K) \tag{32}
$$

- These go back to the requirement that call prices have to be monotonically decreasing, and convex.