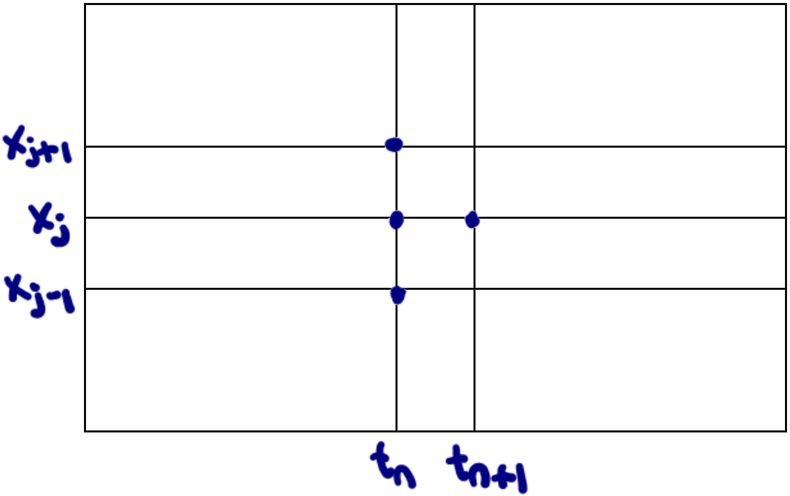
Financial Mathematics 32000

Lecture 4

Roger Lee

2025 April 14

Implicit FD scheme



Implicit FD scheme

Ignoring drift terms, the "model" problem is

$$C_t + \frac{1}{2}\sigma^2 C_{xx} = 0$$

The fully implicit FD scheme is defined by

$$\frac{C_{n+1}^j - C_n^j}{\Delta t} + \frac{1}{2}\sigma^2 \frac{C_{n}^{j+1} - 2C_{n}^j + C_{n}^{j-1}}{(\Delta x)^2} = 0$$

Given C_{n+1}, we need to solve for C_n (in both explicit and implicit scheme.

But in the explicit scheme, the blue subscripts would be $n+1$).

Implicit FD: calculating one column

The explicit scheme

$$C_n^j = q_u C_{n+1}^{j+1} + q_m C_{n+1}^j + q_d C_{n+1}^{j-1}$$

was easy to calculate. The implicit scheme (where $\alpha := \frac{1}{2}\sigma^2 \frac{\Delta t}{(\Delta x)^2}$)

$$-\alpha C_n^{j+1} + (1+2\alpha) C_n^j - \alpha C_n^{j-1} = C_{n+1}^j \qquad j = -J+1, \ldots, J-1$$

with C_n^{-J} and C_n^J given by boundary conditions, is harder, requiring solution of a system of $2J-1$ equations in $2J-1$ unknowns.

$$\begin{pmatrix} 1+2\alpha & -\alpha & & & \\ -\alpha & 1+2\alpha & -\alpha & & \\ & -\alpha & 1+2\alpha & -\alpha & \\ & & \cdots & \cdots & \cdots \\ & & & -\alpha & 1+2\alpha \end{pmatrix} \begin{pmatrix} C_n^{J-1} \\ C_n^{J-2} \\ C_n^{J-3} \\ \vdots \\ C_n^{-J+1} \end{pmatrix} + \begin{pmatrix} -\alpha C_n^J \\ 0 \\ \vdots \\ 0 \\ -\alpha C_n^{-J} \end{pmatrix} = \begin{pmatrix} C_{n+1}^{J-1} \\ C_{n+1}^{J-2} \\ C_{n+1}^{J-3} \\ \vdots \\ C_{n+1}^{-J+1} \end{pmatrix}$$

Implicit FD: error analysis

- Consistency: Taylor analysis shows accuracy $O(\Delta t) + O(\Delta x)^2$.
- Stability: Look for a solution of the form $C_n^j = \lambda_k^{N-n} e^{i\pi k j/J}$. Substitute into the implicit scheme and cancel C_n^j:

$$\lambda_k - 1 = \alpha\lambda_k(e^{i\pi k/J} - 2 + e^{-i\pi k/J}) = -4\alpha\lambda_k \sin^2 \frac{\pi k}{2J}$$

where α is as before. Solving for λ_k,

$$\lambda_k = \frac{1}{1 + 4\alpha \sin^2 \frac{\pi k}{2J}}$$

So $|\lambda_k| \leq 1$ without any restriction on $\Delta x, \Delta t$.

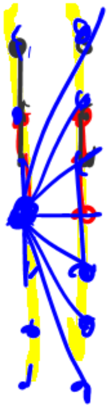
So the fully implicit FD scheme is unconditionally stable.

This is an advantage vs. the explicit scheme.

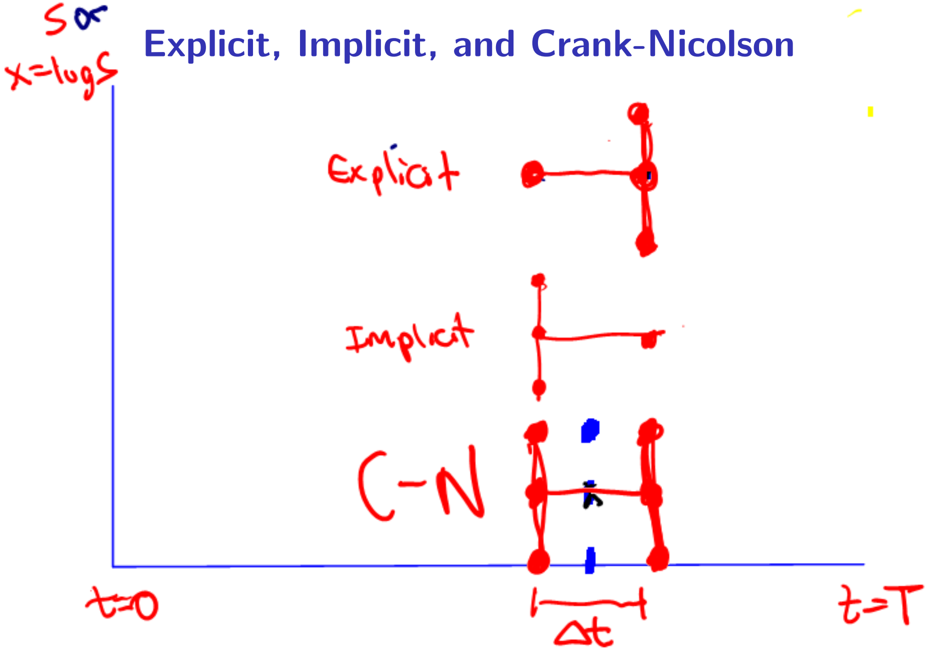
Implicit FD: stability intuition

Intuition for the stability of the implicit scheme:

- Recall the explicit scheme was like a trinomial tree.
  If $\sigma^2 \Delta t$ is too big relative to $(\Delta x)^2$, then the tree tries to squeeze too much variance into three branches, causing the up and down weights to sum to > 1.
- In the implicit scheme, each node at time slice n is linked to *every node* at time slice $n + 1$.
  It's like a tree in which the branches go to *all* levels.
  No problem with squeezing too much variance.



Explicit, Implicit, and Crank-Nicolson



Crank and Nicolson

John Crank (1916-2006) and Phyllis Nicolson (1917-1968)

Crank-Nicolson

$dS = rSdt + \sigma S dW$

$dX = (r - \frac{1}{2}\sigma^2)dt + \sigma dW$

We want to solve for $C(x,t)$ the full PDE (not just model problem)

$$\frac{\partial C}{\partial t} + f(x,t)\frac{\partial^2 C}{\partial x^2} + g(x,t)\frac{\partial C}{\partial x} + h(x,t)C = 0$$

with given f, g, h, and given terminal conditions. Example:

Black-Scholes model for no-dividend stock S. Let $X = \log S$.

- PDE for $C(S,t)$: and $\frac{\partial C}{\partial S}$ and $\frac{\partial^2 C}{\partial S^2}$

  $f(S,t) = \frac{1}{2}\sigma^2 S^2, \quad g(S,t) = rS, \quad h(S,t) = -r$

- PDE for $C(x.t)$: and $\frac{\partial C}{\partial x}$ and $\frac{\partial^2 C}{\partial x^2}$

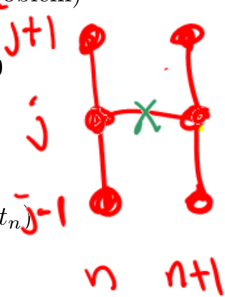
  $f(x,t) = \frac{1}{2}\sigma^2, \quad g(x,t) = r - \frac{1}{2}\sigma^2, \quad h(S,t) = -r$

Crank-Nicolson

Δt

We want to solve for $C(x,t)$ the full PDE (not just model problem)

$$\frac{\partial C}{\partial t} + f(x,t)\frac{\partial^2 C}{\partial x^2} + g(x,t)\frac{\partial C}{\partial x} + h(x,t)C = 0$$



with given f, g, h, and given terminal conditions. Let

$$f_n^j := f(x_j, t_n), \quad g_n^j := g(x_j, t_n), \quad h_n^j := h(x_j, t_n)$$

Discretize the PDE.

$$\frac{C_{n+1}^j - C_n^j}{\Delta t} + \frac{1}{2}\left[f_n^j \frac{C_n^{j+1} - 2C_n^j + C_n^{j-1}}{(\Delta x)^2} + g_n^j \frac{C_n^{j+1} - C_n^{j-1}}{2\Delta x} + h_n^j C_n^j\right]$$
$$+ \frac{1}{2}\left[f_{n+1}^j \frac{C_{n+1}^{j+1} - 2C_{n+1}^j + C_{n+1}^{j-1}}{(\Delta x)^2} + g_{n+1}^j \frac{C_{n+1}^{j+1} - C_{n+1}^{j-1}}{2\Delta x} + h_{n+1}^j C_{n+1}^j\right] = 0.$$

Crank-Nicolson

H

Then for $j = -J+1, \ldots, J-1$,

$$-F_n^j C_n^{j+1} + (1+G_n^j) C_n^j - H_n^j C_n^{j-1} = F_{n+1}^j C_{n+1}^{j+1} + (1-G_{n+1}^j) C_{n+1}^j + H_{n+1}^j C_{n+1}^{j-1}$$

where

$$\begin{aligned} F_n^j &:= \frac{1}{2} \frac{\Delta t}{(\Delta x)^2} f_n^j + \frac{1}{4} \frac{\Delta t}{\Delta x} g_n^j \\ G_n^j &:= \frac{\Delta t}{(\Delta x)^2} f_n^j - \frac{\Delta t}{2} h_n^j \\ H_n^j &:= \frac{1}{2} \frac{\Delta t}{(\Delta x)^2} f_n^j - \frac{1}{4} \frac{\Delta t}{\Delta x} g_n^j. \end{aligned}$$

Assume that the boundary values C_n^J and C_n^{-J} for $n = 0, \ldots, N$ are given.

Crank-Nicolson

Rewriting (for $j = -J+1, \dots, J-1$)

$$-F_n^j C_n^{j+1} + (1+G_n^j) C_n^j - H_n^j C_n^{j-1} = F_{n+1}^j C_{n+1}^{j+1} + (1-G_{n+1}^j) C_{n+1}^j + H_{n+1}^j C_{n+1}^{j-1}$$

in matrix form,

$$\begin{pmatrix} 1+G_n^{J-1} & -H_n^{J-1} & & & \\ -F_n^{J-2} & 1+G_n^{J-2} & -H_n^{J-2} & & \\ & -F_n^{J-3} & 1+G_n^{J-3} & -H_n^{J-3} & \\ & & \cdots & \cdots & \cdots \\ & & & -F_n^{-J+1} & 1+G_n^{-J+1} \end{pmatrix} \begin{pmatrix} C_n^{J-1} \\ C_n^{J-2} \\ \vdots \\ \vdots \\ C_n^{-J+1} \end{pmatrix} + \begin{pmatrix} -F_n^{J-1} C_n^J \\ 0 \\ \vdots \\ 0 \\ -H_n^{-J+1} C_n^{-J} \end{pmatrix}$$

$$= \begin{pmatrix} 1-G_{n+1}^{J-1} & H_{n+1}^{J-1} & & & \\ F_{n+1}^{J-2} & 1-G_{n+1}^{J-2} & H_{n+1}^{J-2} & & \\ & F_{n+1}^{J-3} & 1-G_{n+1}^{J-3} & H_{n+1}^{J-3} & \\ & & \cdots & \cdots & \cdots \\ & & & F_{n+1}^{-J+1} & 1-G_{n+1}^{-J+1} \end{pmatrix} \begin{pmatrix} C_{n+1}^{J-1} \\ C_{n+1}^{J-2} \\ \vdots \\ \vdots \\ C_{n+1}^{-J+1} \end{pmatrix} + \begin{pmatrix} F_{n+1}^{J-1} C_{n+1}^J \\ 0 \\ \vdots \\ 0 \\ H_{n+1}^{-J+1} C_{n+1}^{-J} \end{pmatrix}.$$

If given terminal conditions, then we know C_{n+1} and we solve for C_n.

Comparison: Explicit, Implicit, C-N

Consistency: C-N has accuracy $O(\Delta t)^2 + O(\Delta x)^2$, by Taylor analysis

Stability: C-N is unconditionally stable, by Fourier analysis.

Summary:

	Accuracy	Unconditionally stable	Requires solving linear systems
Explicit	$O(\Delta t) + O(\Delta x)^2$	No	No
Implicit	$O(\Delta t) + O(\Delta x)^2$	Yes	Yes
C-N	$O(\Delta t)^2 + O(\Delta x)^2$	Yes	Yes

c^n c^{n+1}

Solving a linear system

For implicit and C-N schemes, obtain nth column $\mathbf{c}^n$ by solving

$$\mathbf{M}\mathbf{c}^n = \mathbf{b}, \qquad \text{where } \mathbf{b} \text{ is a vector that depends on } \mathbf{c}^{n+1}$$

where $\mathbf{M}$ is an $m \times m$ matrix ($m = 2J - 1$).

- Could calculate $\mathbf{M}^{-1}$; then $\mathbf{c}^n = \mathbf{M}^{-1}\mathbf{b}$.

  But this is too much work: $O(m^2)$ for a tri-diagonal matrix.

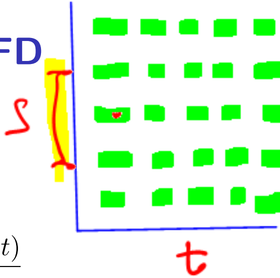
  (Tri-diagonal means that the matrix has a zero for every element not on the main diagonal and the two adjacent diagonals.)
- Better: Gaussian elimination.

  Build augmented matrix $[\mathbf{M}|\mathbf{b}]$. Then row-reduce/back-substitute.

  In the case of a tri-diagonal matrix, the amount of work is $O(m)$.

  Python: `numpy.linalg.solve(M,b)`

Computing Greeks using FD



At (S, t),

$$\Delta := \frac{\partial C}{\partial S} \approx \frac{C(S+\Delta S, t) - C(S-\Delta S, t)}{2\Delta S}$$

$$\Gamma := \frac{\partial^2 C}{\partial S^2} \approx \frac{C(S+\Delta S, t) - 2C(S,t) + C(S-\Delta S, t)}{(\Delta S)^2}$$

$$\Theta := \frac{\partial C}{\partial t} \approx \frac{C(S, t+\Delta t) - C(S, t)}{\Delta t} \text{ or } \frac{C(S, t+\Delta t) - C(S, t-\Delta t)}{2\Delta t}$$

If the grid does not have constant ΔS then

$$\frac{\partial C}{\partial S}(S_j, t_n) \approx \frac{C_{j+1}^n - C_{j-1}^n}{S_{j+1} - S_{j-1}}$$

or, if grid has constant Δx where $x = \log S$ then

$$\frac{\partial C}{\partial S} = \frac{\partial C}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}S} = \frac{\partial C}{\partial x}\frac{1}{S}$$

Implicit FD scheme

Crank-Nicolson scheme

Solving a linear system

Greeks

Americans

Taking *strike* derivatives

American options using FD

- LCP approach: use PSOR (projected successive over-relaxation) to solve discrete Linear Complementarity Problem.
- Simple approach:

  Given column-$(n+1)$ American option prices C^{n+1}, compute the "continuation value" H in column n, using the FD scheme Q.

$$H^n = Q(C^{n+1}).$$

  Take max of continuation value and exercise value, to produce American option price C^n. For example, in case of a put,

$$C_j^n = \max(H_j^n, (K - S_j^n)^+)$$

Smoothing to improve convergence and Greeks

- Prices as functions of (N, S) can be smoothed, by using FD not in column $N-1$, but rather starting at $N-2$ (then $N-3, N-4, \ldots$)
- In column $N-1$, don't use FD. Instead, compute the continuation value in each row of the grid using Black-Scholes formula. Then take max(continuation value, exercise value) in column $N-1$.
- Then compute continuation values in other columns using FD (or tree)
- Plotting the resulting price vs. N, smoother convergence to exact price allows easier detection of convergence, and extrapolation if desired
- Plotting the resulting price vs. S, smoother profile improves accuracy of Greeks

Let p(s,T) be the risk-neutral probability density of S_T

Model-independent fact: $-\frac{\partial C}{\partial K}$ = binary call price

Let C(K,T) be time-0 price of K-strike T-expiry European call on S

Note that

$$C(K,T) = e^{-rT}\int_0^\infty p(s,T)(s-K)^+ \mathrm{d}s$$

Differentiate once wrt K

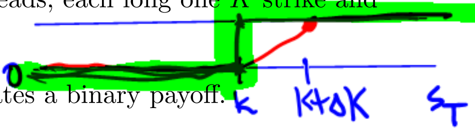
$$\frac{\partial C}{\partial K}(K,T) = -e^{-rT}\int_K^\infty p(s,T)\mathbf{1}_{s>K}\mathrm{d}s = -(\text{binary call price})$$

hence $-\partial C/\partial K$ is the price of the binary call.

- Intuition: Consider $1/\Delta K$ call spreads, each long one K strike and short one $K+\Delta K$ strike call

  On one hand, its payoff approximates a binary payoff.

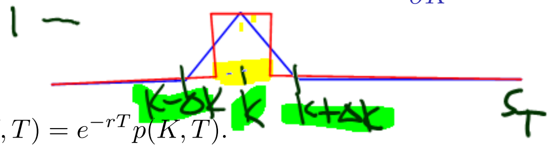
  On the other hand, its price $(C(K) - C(K+\Delta K))/(\Delta K)$ is a finite difference approximation to $-\partial C/\partial K$.



Model-independent fact: S_T density $= e^{rT}\frac{\partial^2 C}{\partial K^2}$

Differentiate again wrt K

$$\frac{\partial^2 C}{\partial K^2}(K,T) = e^{-rT} p(K,T).$$



This is known as the **Breeden-Litzenberger** formula.

- Intuition: Consider $1/\Delta K$ butterflies (1x2x1): long one $K-\Delta K$ strike, short two K strike, long one $K+\Delta K$ strike calls. On one hand, its price $\approx e^{-rT}\mathbb{P}(S_T$ lies within $\Delta K/2$ of $K) \approx e^{-rT}p(K,T)\Delta K$, so $1/(\Delta K)^2$ butterflies (1x2x1) has price $\approx e^{-rT}p(K,T)$.
  On the other hand, the price of $1/(\Delta K)^2$ butterflies (1x2x1) is also $(C(K-\Delta K) - 2C(K) + C(K+\Delta K))/(\Delta K)^2$, a finite difference approximation to $\partial^2 C/\partial K^2$.

Breeden-Litzenberger

Douglas Breeden and Robert Litzenberger

Morgan Stanley

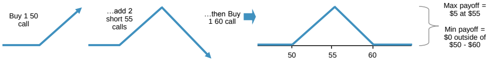
Morgan Stanley

How Options Implied Probabilities Are Calculated

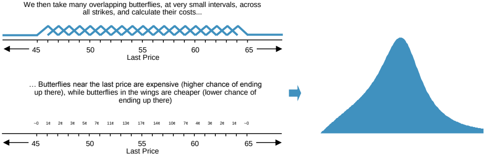
The implied probability distribution is an approximate risk-neutral distribution derived from traded option prices using an interpolated volatility surface. In a risk-neutral world (i.e., where we are not more adverse to losing money than eager to gain it), the fair price for exposure to a given event is the payoff if that event occurs, times the probability of it occurring. Worked in reverse, the probability of an outcome is the cost of exposure to the outcome divided by its payoff.

In the options market, we can buy exposure to a specific range of stock price outcomes with a strategy know as a butterfly spread (long 1 low strike call, short 2 higher strikes calls, and long 1 call at an even higher strike). The probability of the stock ending in that range is then the cost of the butterfly, divided by the payout if the stock is in the range.

Building a Butterfly:



To find a smooth distribution, we price a series of theoretical call options expiring on a single date at various strikes using an implied volatility surface interpolated from traded option prices, and with these calls price a series of very tight overlapping butterfly spreads. Dividing the costs of these trades by their payoffs, and adjusting for the time value of money, yields the future probability distribution of the stock as priced by the options market.



Note, adequate trading volume and liquidity is required to produce a volatility surface and price theoretical options. Therefore, options implied probabilities will not be available for all equities.

Interactive Brokers “Probability Lab”

as of 2023 April 12

The Dupire equation

skip

Given local volatility dynamics under risk-neutral probabilities

$$\mathrm{d}S_t = rS_t\mathrm{d}t + \sigma(S_t, t)S_t\mathrm{d}W_t$$

We want to derive the Dupire PDE. Compare and contrast:

- The (extended) B-S PDE fixes (K, T), and finds call prices for all (S, t).

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2 C}{\partial S^2} = rC$$

$$C(S,T) = (S-K)^+$$

- The Dupire PDE fixes (S, t) and finds call prices for all (K, T).

$$\frac{\partial C}{\partial T} - \frac{1}{2}K^2\sigma^2(K,T)\frac{\partial^2 C}{\partial K^2} + rK\frac{\partial C}{\partial K} = 0$$

$$C(K,t) = (S-K)^+$$

Dupire equation

skip

Application of Dupire equation

- Given the dynamics, the Dupire equation lets you solve *one* PDE to find the prices of *all* call options at all (K, T).
- Given the prices of all call options, the Dupire PDE lets you infer the dynamics. In other words, it lets you *calibrate* σ to options prices.

Proof of the Dupire equation

skip

We have already motivated this in the trinomial tree setting.

To prove this in continuous time, we will combine two facts about the time-0 probability density of S_T

$$p(s,T)\mathrm{d}s := \mathbb{P}(S_T \in (s - \mathrm{d}s/2,\ s + \mathrm{d}s/2))$$

- Breeden-Litzenberger formula: The risk-neutral probability density of S_T can be extracted from the second strike-derivative of European call prices, model-independently via:

$$p(s,T) = e^{rT}\frac{\partial^2 C}{\partial K^2}(s,T)$$

- Probability density satisfies Fokker-Planck (forward Kolmogorov) PDE.

Densities satisfy Fokker-Planck PDE

skip

Suppose that

$$\mathrm{d}S_t = a(S_t)\mathrm{d}t + b(S_t)\mathrm{d}W_t$$

Then the time-0 probability density $p(s,T)$ of S_T satisfies the *Fokker-Planck* or *forward Kolmogorov* PDE

$$\frac{\partial p}{\partial T}(s,T) + \frac{\partial}{\partial s}[a(s)p(s,T)] - \frac{1}{2}\frac{\partial^2}{\partial s^2}[b^2(s)p(s,T)] = 0$$

Contrast: Feynman-Kač says that $F(x,t) := \mathbb{E}_t[f(S_T)|\ S_t = x]$ satisfies the *backward Kolmogorov* PDE

$$\frac{\partial F}{\partial t}(x,t) + a(x)\frac{\partial F}{\partial x}(x,t) + \frac{1}{2}b^2(x)\frac{\partial^2 F}{\partial x^2}(x,t) = 0$$

Fokker-Planck PDE: Idea of proof

skip

For any smooth $\varphi : \mathbb{R} \to \mathbb{R}$ vanishing outside some $[y, z] \subset \mathbb{R}$,

$$\mathbb{E}\varphi(S_T) = \varphi(S_0) + \mathbb{E}\int_0^T a(S_t)\frac{\partial \varphi}{\partial s}(S_t) + \frac{1}{2}b^2(S_t)\frac{\partial^2 \varphi}{\partial s^2}(S_t)\mathrm{d}t$$

$$\Rightarrow \frac{\partial}{\partial T}\mathbb{E}\varphi(S_T) = \mathbb{E}\left[a(S_T)\frac{\partial \varphi}{\partial s}(S_T) + \frac{1}{2}b^2(S_T)\frac{\partial^2 \varphi}{\partial s^2}(S_T)\right]$$

$$\Rightarrow \frac{\partial}{\partial T}\int_{\mathbb{R}}\varphi(s)p(s,T)\mathrm{d}s = \textstyle\int_{\mathbb{R}} a(s)\frac{\partial \varphi}{\partial s}(s)p(s,T)\mathrm{d}s + \frac{1}{2}\int_{\mathbb{R}} b^2(s)\frac{\partial^2 \varphi}{\partial s^2}(s)p(s,T)\mathrm{d}s$$

$$\Rightarrow \int_{\mathbb{R}}\varphi(s)\frac{\partial p}{\partial T}(s,T)\mathrm{d}s = -\int_{\mathbb{R}}\varphi(s)\frac{\partial}{\partial s}[a(s)p(s,T)]\mathrm{d}s + \frac{1}{2}\int_{\mathbb{R}}\varphi(s)\frac{\partial^2}{\partial s^2}[b^2(s)p(s,T)]\mathrm{d}s$$

$$\Rightarrow \int_{\mathbb{R}}\varphi(s)\left(\frac{\partial p}{\partial T}(s,T)+\frac{\partial}{\partial s}[a(s)p(s,T)] - \frac{1}{2}\frac{\partial^2}{\partial s^2}[b^2(s)p(s,T)]\right)\mathrm{d}s = 0$$

Since φ is arbitrary, the other factor must vanish for all s.

Combine the two facts to prove the Dupire equation

skip

Therefore, $e^{rT}(\partial^2 C/\partial K^2)$ satisfies the Fokker-Planck PDE,

so $G := \partial^2 C/\partial K^2$ satisfies the same PDE, but with an extra term rG.

$$\frac{\partial G}{\partial T} - \frac{\partial^2}{\partial K^2}\left(\frac{1}{2}\sigma^2(K,T)K^2 G\right) + r\frac{\partial}{\partial K}(KG) + rG = 0.$$

Integrating twice wrt K and applying the appropriate boundary conditions,

we have the Dupire equation

$$\frac{\partial C}{\partial T} - \frac{1}{2}K^2\sigma^2(K,T)\frac{\partial^2 C}{\partial K^2} + rK\frac{\partial C}{\partial K} = 0$$

$$C(K,t) = (S-K)^+$$

Calibration of local volatility

skip

- Observe available call prices C. Interpolate/extrapolate/fit a smooth function $C(K,T)$. (Carefully – we don't want the interpolated C values to allow arbitrage).
- Then
$$\sigma(K,T) = \left(\frac{\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}} \right)^{1/2}$$
We saw the discrete version of this formula for trinomial trees.
- Alternatively, this formula can be rewritten to calibrate $\sigma(K,T)$ from a smooth function fitted to implied volatilities, instead of option prices.

Calibration of local volatility

skip

Some other approaches:

Parametrically (for example, $\sigma(S,t) = \max(0, \alpha_0 + \alpha_1 S + \alpha_2 t)$ for constant parameters α) or non-parametrically (for example, allow σ to take at each (S,t) any positive value), solve for the σ which minimizes:

Error of model prices relative to observed prices, plus a penalty term.

- Error of model prices might be defined to depend on

  $\sum |$ model prices - observed prices $|^2$
- The penalty might be defined to depend on $\|\nabla\sigma\|$, or on $\|\sigma - \sigma_{prior}\|$.