Numerical Methods 46.950

Lecture 1

August 30, 2016

Overview of the Lecture:

- Overview and Examples
- Feynman-Kac Theorem and its applications
- Cauchy Problem formulation

Introduction

There are several common approaches to calculate the price of the instrument: Analytic formula, Binomial trees, Monte Carlo, Finite Differences.

Factors affecting Finite Difference and Monte Carlo methods:

- Number of dimensions;
- Functional form of coefficients;
- Boundary/final conditions;
- Decision features;
- Linear or non-linear;
- Additional factors;

Finite difference is good for low dimensions, method of choice for a contract with embedded decisions. They are excellent for nonlinear problems.

Why to use FDM

- Convergence rates are better compare to Monte Carlo;
- Easy to adapt for early exercise, discrete sampling, boundaries and barriers;
- FDM is extremely effective way of capturing early exercise;
- Also with jumps and path dependency;
- However, once the number of dimensions become greater than 4, it gets very expensive to use Finite Difference Methods;
- Greeks are easy to calculate;

European Option

We will be looking at the PDE that arise pricing different financial instruments, we list a few of them next.
Assume the stock dynamics:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

We want to price a contigent claim that depends on S_t. One way to approach is to solve the corresponding PDE.
Consider a European Call option, by constructing a replicating portfolio, one can derive the PDE the price function $V(t, S_t)$ must satisfy:

$$V_t + rSV_S + \frac{1}{2}\sigma^2 S^2 V_{SS} - rV = 0,$$

and terminal condition $V(T, S) = (S - K)^+$.

Heston Stochastic Volatility model

We can also consider the model for the volatility of the stock:

$$
\begin{aligned}
dS(t) &= rS(t)dt + \sqrt{v(t)}S(t)dW(t) \\
dv(t) &= \kappa(\eta - v(t))dt + \sigma(t)\sqrt{v(t)}dB(t) \\
V(T, S(T), v(T)) &= (S(T) - K)^+
\end{aligned}
$$

with $dW(t)dB(t) = \rho dt$, $0 < S < \infty$, $0 < v < \infty$, and $t \in [0, T)$. Then the PDE for a call option with a price $V(t, S_t, \sigma_t)$ satisfies the following PDE

$$
\begin{aligned}
V_t + \frac{vS^2}{2} V_{SS} &+ \rho\sigma vSV_{Sv} \\
&+ \frac{\sigma^2 v}{2} V_{vv} + rSV_S + \kappa(\eta - v)V_v - rV = 0
\end{aligned}
$$

and terminal condition $V(T, S, \sigma) = (S - K)^+$.

Asian Option

Path dependent options can be priced by solving the appropriate PDE as well. For example, Asian call option with the payoff depending on:

$$\frac{1}{t}\int_0^t S_u du$$

Current framework doesn't "support" certain types of path dependent payoffs! **Solution:** Extend the framework by adding a new state process Y_t:

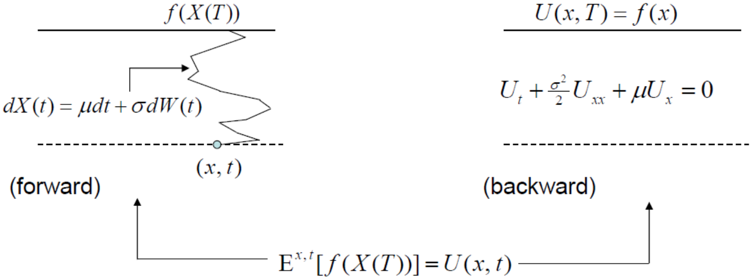
$$Y_t = \int_0^t S_u du$$

Now the function $V(t, S_t, Y_t)$ satisfies

$$V_t + rSV_S + \frac{1}{2}\sigma^2 S^2 V_{SS} + SV_Y - rV = 0,$$

and terminal condition $V(T, S, Y) = \left(\frac{Y}{T} - K\right)$.

Connection between SDE and PDE



Approaches to price a financial instrument

Recall from The Fundamental theorem of Asset Pricing: In the presence of the equivalent martingale measure(EEM) associated with a numeraire B the prices of the financial instruments can be defined as

$$V(t) = B(t)\mathbb{E}^B\left[\frac{V(T)}{B(T)}\right]$$

Approach 1: PDE approach: Derive and solve explicitly or numerically the PDE for $V(t)$;
Approach 2: Risk-neutral approach: Simulate price paths of the underlying under EMM to estimate the expectation of a discounted payoff.
Let's explore those approaches in detail and the theorem that connects the two.

Derivation of the pricing equation/replication

One can construct a replicating portfolio to derive the PDE! Consider the value of the option at time $t \in [0, T]$, $V(t, S(t))$ and the dynamics of underlying:

$$dS(t) = \mu S(t)dt + \sigma S(t)dW_t$$

Applying Ito's lemma to $V(t, S_t)$:

$$dV(t, S) = \left(V_t + \mu S(t)V_s + \frac{1}{2}\sigma^2 S(t)^2 V_{ss} \right) dt + \sigma S(t)V_s dW_t$$

Consider the hedging portfolio $P(t)$ that replicates the cash flows of the derivative with n units of the underlying, plus borrowing and lending with interest rate r. The change in the replicating portfolio value is

$$dP(t) = ndS + (P - nS(t))rdt$$
$$dP(t) = [Pr + nS(t)(\mu - r)]dt + n\sigma S(t)dW_t$$

Cont

Equating the dynamics for drift and volatility we find:

$$\begin{aligned} n &= V_s(t, S(t)), \\ V_t + \mu S(t) V_s + \frac{1}{2}\sigma^2 S(t)^2 V_{ss} &= rP + n(\mu - r)S(t). \end{aligned}$$

we require the hedging portfolio should replicate the option for any $t \in [0, T]$, i.e. $P(t) = V(t, S(t))$ and it is satisfied for all values of t and $S(t)$, if V is a function of two variables satisfying the partial differential equation:

$$V_t + rsV_s + \frac{1}{2}\sigma^2 s^2 V_{ss} - rV = 0,\ 0 \leq t < T,\ 0 < s < \infty$$

and appropriate terminal condition. It remains to show that we can indeed construct a replicating self-financing investment strategy for European options.

Theorem

Theorem

Let p be a locally integrable function, $S(t)$ follow Black-Scholes dynamics and r be the risk-free interest rate. Let V be the solution of the partial differential equation

$$V_t + rsV_s + \frac{1}{2}\sigma^2 s^2 V_{ss} - rV = 0,\ 0 \le t < T,\ 0 < s < \infty$$

satisfying the final condition

$$V(T, s) = p(s),\ 0 < s < \infty.$$

Assume that V is twice differentiable in the region $(0, \infty), [0, T)$. Then the price of the European option with the exercise date T and payoff $p(S(T))$ at any time $0 \le t \le T$ is $V(t, S(t))$ and the option can be replicated with a self-financing investment strategy with initial value $P(0) = V(0, S(0))$ and the stock holding $n = V_s(t, S(t))$.

Derivation of the pricing equation/The Markov Property

Suppose $X(t)$ solves the SDE:

$$dX(t) = \alpha(t, X(t))dt + \beta(t, X(t))dW(t),$$

provided there exist some constant C, s.t. for $0 \leq t \leq T$ and $x, y \in \mathbb{R}$

$$|\alpha(t,x) - \alpha(t,y)| + |\beta(t,x) - \beta(t,y)| \leq C|x-y|, \tag{1}$$
$$|\alpha(t,x) - \alpha(t,y)| \leq C(1 + |x|). \tag{2}$$

Let $0 \leq t < T$ be given, define

$$V(t,x) = \mathbb{E}[f(X(T))|X(t) = x]$$

then

$$\mathbb{E}[f(X(T))|X(t) = x] = V(t, X(t))$$

i.e. solutions to the SDEs are Markov processes.

Feynman-Kac theorem

Theorem

Consider the SDE

$$dX(t) = \alpha(t, X(t))dt + \beta(t, X(t))dW(t), \ t \geq 0.$$

Let $h(y)$ be a Borel measurable function. Fix $T > 0$, and let $t \in [0, T]$ be given. Define the function

$$V(t, x) = \mathbb{E}_{t,x}[f(X(T))].$$

Then $V(t, x)$ satisfies the partial differential equation

$$V_t(t, x) + \alpha(t, x)V_x(t, x) + \frac{1}{2}\beta^2(t, x)V_{xx}(t, x) = 0 \tag{3}$$

and the terminal condition $V(T, x) = f(x)$ for all x.

Proof

Applying Ito formula:

$$
\begin{aligned}
& d(V(t,X(t))) \\
&= V_t(t,X(t))dt + V_x(t,X(t))dX(t) + \frac{1}{2}V_{xx}(t,X(t))dX_t dX_t \\
&= \left(V_t(t,X(t)) + \alpha(t,X(t))V_x(t,X(t)) + \frac{1}{2}\beta^2(t,X(t))V_{xx}(t,X(t)) \right) dt \\
&+ \beta(t,X(t))V_x(t,X(t))dW(t)
\end{aligned}
$$

Since we know that $V(t,X(t))$ is a martingale we have to set the dt term to zero.

Discounted Feynman-Kac theorem

Theorem

Consider the SDE

$$dX(t) = \alpha(t, X(t))dt + \beta(t, X(t))dW(t).\ t \geq 0.$$

Let $f(y)$ *be a Borel measurable function. Fix* $T > 0$*, and let* $t \in [0, T]$ *be given. Define the function*

$$V(t, x) = \mathbb{E}_{t,x}[\exp^{-r(T-t)} f(X(T))].$$

Then $V(t, x)$ *satisfies the partial differential equation*

$$V_t(t, x) + \alpha(t, x)V_x(t, x) + \frac{1}{2}\beta(t, x)^2 V_{xx}(t, x) - rV(t, x) = 0 \quad (4)$$

and the terminal condition $V(T, x) = f(x)$ *for all* x*.*

Proof

Applying Ito formula:

$$d\left(e^{-rt}V(t,X(t))\right) = -re^{-rt}V(t,X(t))dt + e^{-rt}dV(t,X(t))$$

$$= e^{-rt}\left[\, V_t(t,X(t)) + \alpha(t,X(t))V_x(t,X(t)) + \frac{1}{2}\beta^2(t,X(t))V_{xx}(t,X(t))\right.$$

$$\left. - rV(t,X(t))\,\right]dt + e^{-rt}\beta(t,X(t))V_x(t,X(t))dW(t)$$

Similarly, we know that $e^{-rt}V(t,X(t))$ is a martingale, thus we have to set the *dt* term to zero.

Recap of the proofs:

- Write all the dynamics under the chosen measure;
- Find a martingale;
- Take the differential of it and set the *dt* term to zero;

Generalizations

- Non constant interest rates?
- What about change of measure?
- Running credit/debit for the derivative?
- Is there a multi-dimensional version?

Application/Fixed Income

Consider an interest rate process $r(t)$ that satisfies

$$dr(t) = \beta(t, r(t))dt + \gamma(t, r(t))d\widetilde{W}(t),$$

under the risk-neutral measure.
For $t \geq 0$, define the discount process $D(t)$ by

$$D(t) = e^{-\int_0^t r(u)du}$$

From this definition,

$$dD(t) = -D(t)r(t)dt.$$

Pricing a Bond

Consider a zero-coupon bond that pays \$1 at time T. The value of the bond at time $t, 0 \leq t \leq T$, is given by

$$B(t, T) = \widetilde{\mathbb{E}}\left[e^{-\int_t^T r(u)du} | \mathcal{F}(t)\right] \tag{5}$$

$$= \frac{1}{D(t)} \tilde{\mathbb{E}}\left[D(T) | \mathcal{F}(t)\right]. \tag{6}$$

Since $r(t)$ is given by a stochastic differential equation, it is Markov process, i.e.

$$B(t, T) = V(t, r(t)).$$

Cont

Consider the discounted bond price process where

$$d\Big(D(t)V(t,r(t))\Big) = D(t)[(-r(t)V(t,r(t)) + V_t(t,r(t))$$
$$+ \beta(t,r(t))V_r(t,r(t)) + \frac{1}{2}\gamma^2(t,r(t))V_{rr}(t,r(t)))dt$$
$$\gamma(t,r(t))V_r(t,r(t))d\widetilde{W}(t)\Big).$$

Since the discounted bond price is a martingale under risk neutral measure, $V(t,r)$ satisfies

$$V_t(t,r) + \beta(t,r)V_r(t,r) + \frac{1}{2}\gamma^2(t,r)V_{rr}(t,r) - rV(t,r) = 0$$

with terminal condition

$$V(T,r) = 1$$

Consider the Hull-White interest rate model where

$$dr(t) = b(t)(a(t) - r(t))dt + \sigma(t)d\widetilde{W}(t), t \geq 0.$$

Let $V(t, r)$ be the time t value of a zero-coupon bond that pays \$1 at time T. Then

$$V_t(t, r) + b(t)(a(t) - r)V_r(t, r) + \frac{1}{2}\sigma^2(t)V_{rr}(t, r) - rV(t, r) = 0$$

with terminal condition

$$V(T, r) = 1$$

Application/Bond Option with Hull-White model

Consider a call option on a zero-coupon bond that pays \$1 at time T_2 where the call has a strike K and maturity at time T_1, $T_1 < T_2$. The value of the call at time $t, 0 \le t \le T_1$:

$$C(t, r(t)) = \frac{1}{D(t)} \widetilde{\mathbb{E}} \left(D(T_1)(B(T_1, T_2) - K)^+ | \mathcal{F}_t \right).$$

Then by Feynman-Kac theorem the value of the call satisfies

$$C_t(t, r) + b(t)(a(t) - r)C_r(t, r) + \frac{1}{2}\sigma^2(t) C_{rr}(t, r) - rC(t, r) = 0$$

with terminal condition

$$C(T_1, r) = (B(T_1, T_2) - K)^+$$

Cauchy problem

Initially, we will consider the numerical solution of the general one-dimensional terminal value problem

$$\frac{\partial V}{\partial t} + \mathcal{L}V = 0,$$

where $\mathcal{L}$ is the operator

$$\mathcal{L} = \mu(t,x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma(t,x)^2\frac{\partial^2}{\partial x^2} - r(t,x)$$

and where $V = V(t,x)$ satisfies a terminal condition $V(T,x) = f(x)$. This terminal value problem is called **Cauchy problem** to be solved for $V(t,x)$.

Well-posed problems

What are the properties of the solution to

$$\frac{\partial V}{\partial t} + \mathcal{L}V = 0$$

we would like to have?

- Existence: there is a solution to look for;
- Uniquness: we don't end up having two distinct prices for a derivative;
- Stability: able to bound a solution of a PDE in terms of its data.

Recommended reading

- [S]: Section 4.5, Chapter 6.
- [TR]: pages 23-29.

Numerical Methods 46.950

Lecture 2

September 6, 2016

Overview of the Lecture:

- Well-Posed problems
- Heat equation and the properties of the solution
- Conversion of BS PDE to the heat equation
- Application to the down-and-out call option
- Boundary conditions types
- Degenerate Equations and the Fichera test
- Far boundary conditions
- Finite difference approximations of the partial derivatives

Why PDE and not Expectations?

- Feynman-Kac is a link between the solution of the PDE and the expectations;
- But there are more methods available for solving PDE(exactly and approximately) than obtaining expectations.

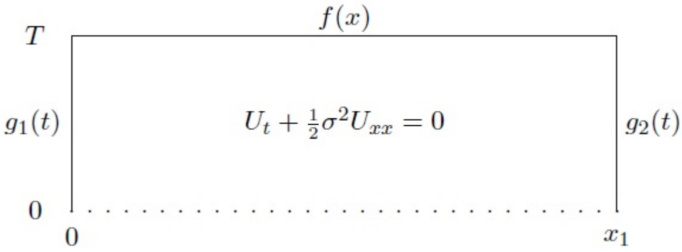
However, before we start looking at any numerical methods to solve the PDE we must ensure the problem is *well-posed*, i.e. we discuss

- Existence
- Uniqueness
- Stability

Backward heat Equation

Let's start with a simple heat equation:

$$U_t + \frac{1}{2}\sigma^2 U_{xx} = 0$$



with a given terminal condition $U(T, x) = f(x)$. Consider some region $[0, x_1]$ and boundary conditions $g_1(t)$ and $g_2(t)$ for $x = 0$ and $x = x_1$ respectively. And assume $g_1 = 0$ and $g_2 = 0$.

Solution for BVP

If we take $[0, x_1] = [0, \pi]$, then it can be shown that

$$U(t, x) = \sum_{k=1}^{\infty} b_k e^{-\frac{1}{2}\sigma^2 k^2 (T-t)} \sin(kx)$$

and b_k are chosen to be Fourier coefficients to match the terminal condition:

$$b_k = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(kx) dx, \quad k = 0, 1, \ldots$$

- If $\sigma > 0$, the amplitude of the solution mode decays in time. The larger the wave number k, the faster it decays. Typical for *parabolic* PDE.
- If $\sigma = 0$, then the amplitude of the solution remains the same for all k. Typical for *hyperbolic* PDE.

Diffusion smoothing

Consider an example of the terminal value as $f(x) = 1$, then the sine expansion is

$$f(x) = \sum_{k>0 \ odd} \frac{4\sin(kx)}{\pi k}$$

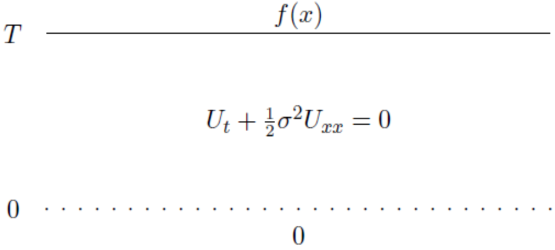
which converges very slowly. The reason is discontinuity at $x = 0, \pi$. However, the solution near those "corners" inside the region looks like:

$$U(t,x) = \sum_{k>0 \ odd} e^{-\frac{1}{2}\sigma^2 k^2 (T-t)} \frac{4\sin(kx)}{\pi k}$$

Thus, discontinuity between terminal/boundary condition doesn't propagate into the solution region but causes issues close to the boundaries.

Initial value problem for the heat equation(on $\mathbb{R}$)

Extend the boundaries to infinity in both directions and consider the new problem



$$
\begin{aligned}
U_t + \frac{1}{2}\sigma^2 U_{xx} &= 0, \\
U(T, x) &= f(x).
\end{aligned}
$$

Cont.

We can solve it probabilistically!

$$dX(t) = \sigma dW(t)$$

then by Feynman-Kac $U(t,x) = \mathbb{E}[f(X(T))|X(t) = x]$ solves the equation.
We know that conditioned on $X_t = x$, $X(T)$ is distributed $N(0, \sigma^2(T-t))$ and therefore it follows that:

$$U(t,x) = \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \int_{-\infty}^{\infty} e^{-\frac{(y-x)^2}{2\sigma^2(T-t)}} f(y)dy$$

or, in a different notation:

$$U(t,x) = \int_{-\infty}^{\infty} K(\sigma^2(T-t), y-x) f(y)dy$$

which can be derived using Fourier transforms.
Be mindful of **non-unique** solution when there is no boundary conditions!

Applicable payoffs

To know it is the *only* solution, we must prove its uniqueness! It can be shown that the solution is in fact unique if the boundary conditions are imposed.

For what type of functions $f(x)$ can we write the solution as

$$U(t,x) = \int_{-\infty}^{\infty} K(y-x, \sigma^2(T-t)) f(y) dy$$

without specifying any boundary condition? We must require modest restriction on the growth as x approaches infinity:

$$f(x) \leq Me^{c|x|^2}$$

Initial boundary value problem for heat equation on half space (on $\mathbb{R}^+$)

Consider now the case of $\mathbb{R}^+$. That is only one boundary condition $g_1(0)$ is given. Take $g_1 = 0$ on that boundary. And extend the terminal payoff function to the whole real line as

$$F(x) = \begin{cases} f(x), & x > 0 \\ -f(-x), & x < 0 \end{cases}$$

then we can write the solution as

$$\begin{aligned} U(t,x) = &\int_0^\infty K(y-x, \sigma^2(T-t)) f(y) dy \\ &+ \int_0^\infty K(y+x, \sigma^2(T-t)) f(y) dy, \quad (x > 0) \end{aligned}$$

and it will be a unique solution to the half space problem. (provided the payoff growth restriction)

Compatible data

If solution is to be continuous(up to the boundary, and up to the terminal condition), then the boundary and terminal condition must be *compatible*, i.e. they should satisfy

$$f(x_0) = g_1(T)$$

When the data are incompatible, the solution is bound to be singular near the corners even if f and g_1 are individually smooth.
Relevance to finance: down-and-out put if the strike is above the barrier. These options are difficult to hedge numerically if, near time of maturity the stock price wanders near the barrier.

Stability

We defined

$$U(t,x) = \mathbb{E}^{t,x}[f(X(T))]$$

and by the properties of expectation

$$|U(t,x)| \leq \max_{x \in I} |f(x)|$$

Why stability matters:

- We don't evaluate things exactly in practice but because of stability we are not in danger of those errors;
- If the relative error is small compare to absolute error, then it translates to the relative error in the solution being small compare to absolute error;
- Often used in proofs of uniqueness;

Change of variables in parabolic PDE

The starting point is the BS equation for a standard call option:

$$
\begin{aligned}
&V_t + rsV_s + \frac{1}{2}\sigma^2 s^2 V_{ss} - rV = 0, \\
&V(T, s) = (s - K)^+, \quad s \in \mathbb{R}^+.
\end{aligned}
$$

Note there is a multiplier of S of the same order as the partial derivative. Use change of variables $x = \ln s$ to remove the multiplier and define $V(t, s) = u(t, x)$:

$$
\begin{aligned}
&u_t + \left(r - \frac{1}{2}\sigma^2\right) u_x + \frac{1}{2}\sigma^2 u_{xx} - ru = 0, \\
&u(T, x) = f(x) = (e^x - K)^+, \quad x \in \mathbb{R}.
\end{aligned}
$$

Cont.

Now consider a transformation to remove lower terms:

$$U(t,x) = e^{\alpha x + \beta(T-t)} u(t,x)$$

Then $U(t,x)$ solves the heat equation

$$\begin{aligned} U_t + \frac{1}{2}\sigma^2 U_{xx} &= 0, \\ U(T,x) &= e^{\alpha x} f(x). \end{aligned}$$

Thus, Black-Scholes PDE can also be transformed to the heat equation! The solution can be written as

$$V(t,s) = e^{\alpha \ln s + \beta(T-t)} U(t, \ln s)$$

And any payoff that grows no faster than $M \exp\left(c(\ln s)^2\right)$ will produce a unique solution. Most of the payoffs are such.

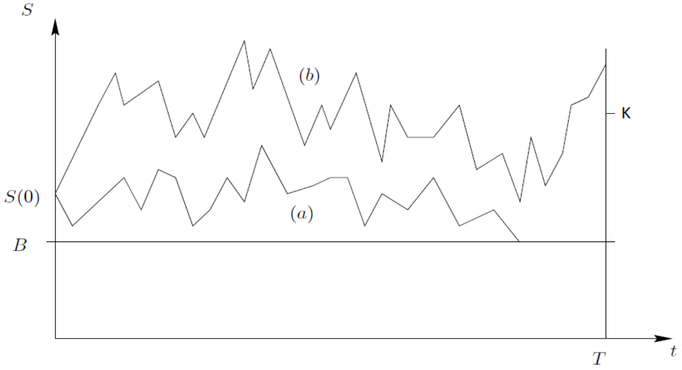
Applications

All we saw so far provides the main ideas for pricing any European barrier options:

- up-and-out: the option expires worthless if the barrier $S = B$ is reached from below;
- down-and-out: the option expires worthless if the barrier $S = B$ is reached from above;
- up-and-in: the option expires worthless unless the barrier $S = B$ is reached from below;
- down-and-in: the option expires worthless unless the barrier $S = B$ is reached from above;

Down-and-out call

Assume $B < K$, where B is the barrier for the option, i.e. the payoff vanishes at $S = B$. Then we have compatible data.



Value of the down-and-out call option

It can be shown that the value of this option is

$$V(t,S) = C(t,S) - \left(\frac{S}{B}\right)^{(1-k)} C(t, \frac{B^2}{S})$$

where $k = r/\left(\frac{\sigma^2}{2}\right)$. Assume a change of variables:

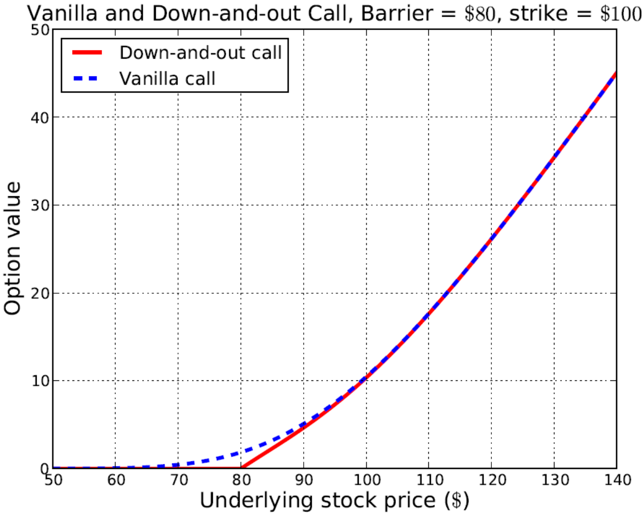
$$S = Be^x, \quad U(t,x) = Be^{\alpha x + \beta(T-t)} u(t,x)$$

to convert the problem to the heat equation:

$$U_t + \frac{1}{2}\sigma^2 U_{xx} = 0$$

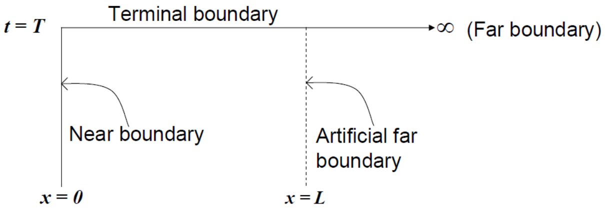
for $x \in \mathbb{R}$ and $U(T,x) = Be^{-\alpha x}(e^x - K)^+$
Trick: Reflect the problem about in an odd way and use the solution to the heat problem on the entire real line!

Cont


Vanilla and Down-and-out Call, Barrier = $80, strike = $100
Down-and-out call
Vanilla call
Option value
Underlying stock price ($)
0
10
20
30
40
50
50
60
70
80
90
100
110
120
130
140

Boundary conditions

Numerically we need to truncate the domain and set the boundary conditions on those boundaries.



Thus there are two boundaries to consider:

- Near boundary, i.e. S_{min}, often $S_{min} = 0$;
- Far boundary, i.e. S_{max};

We discuss those boundaries in details next.

Types of BC considered

The more general approach is to express the boundary condition at $x = 0$ as

$$p(t)U_x(t,0) + q(t)U(t,0) = g(t)$$

where p, q and g are given functions of t.

- If $p = 0, q = 1$, the BC is a **Dirichlet** boundary condition;
- If $p = 1, q = 0$, the BC is a **Neumann** boundary condition;

Degenerate PDE

Consider the following case of $r = 0$ in Black-Scholes equation, then the PDE becomes:

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} = 0,$$

which **degenerates** on $S = 0$, i.e. the higher order terms disappear on the boundary $S = 0$.
Is explicit(exogenous) boundary condition needed at $S = 0$? In this particular case, **no**, but it could be needed in the presence of lower order terms.

Fichera test

Consider the general case of a parabolic PDE:

$$U_t + a(t,x)U_{xx} + b(t,x)U_x + c(t,x)U + d(x,t) = 0, \quad x \geq x_0 \qquad (1)$$

where $a(t, x_0) = 0$. How do we know if the boundary condition is built in so we can omit the boundary condition and still have a well-posed problem?

Fichera test: we **don't need** to specify any boundary condition if

$$(b(x_0) - a_x(x_0))n \leq 0$$

at $x = x_0$ where n is the outer normal vector to the boundary!

Example: Black-Scholes

One application comes when looking at the BS PDE at $S = 0$:

$$U_t + \frac{1}{2}\sigma^2 S^2 U_{SS} + rSU_S - rU = 0$$

$$(b(0) - a_S(0))(-1) = -rS|_{S=0} + \sigma^2 S|_{S=0} = 0$$

No need for the boundary condition at $S = 0$! Use the equation itself: plug $S = 0$ into the equation to get:

$$U_t - rU = 0$$

which has an explicit solution $U(t, 0) = f(0)e^{-r(T-t)}$.
Intuition: The paths of a stochastic process are forced back when approaching the degenerate boundary and by placing boundary condition one can introduce extra(possibly conflicting) information.

Example: CIR model

Another application comes from Interest Rate models. Consider a dynamics for the short rate model

$$dr(t) = \alpha(\beta - r(t))dt + \sigma\sqrt{r_t}dW_t$$

then the price function of any derivative that depends on $r(t)$ satisfies

$$U_t + \frac{1}{2}\sigma^2 r U_{rr} + \alpha(\beta - r)U_r - rU = 0 \tag{2}$$

on $r = 0$ the function equation becomes first order, i.e. degenerates, then we check the Fichera condition

$$(b(0) - a_r(0))(-1) = -\alpha\beta + \frac{1}{2}\sigma^2$$

Cont

- Thus, if $-\alpha\beta + \frac{1}{2}\sigma^2 \leq 0$, then no need for the boundary condition!
- Note, this is the very well known condition for the model not to produce negative rates! That is, if the initial interest rate is strictly positive, the interest rate is always strictly positive.
- In the absence of this condition, the interest rate can become zero, but is we assume that α and β are positive, the interest rate doesn't become negative. When it hits zero, it bounces off and becomes positive again.

Consequences of not following the Fichera rule

What if you impose boundary condition when Fichera test says no boundary condition is needed?

- You add an extra information to the equation and may end up with a solution to another problem!

What if you use the equation itself when Fichera test says exogenous boundary condition is required?

- Possible to get non unique solution!

The domain/Choice of the far boundary position

How far should we place the far boundary? In a particular case of a Black Scholes PDE, in a case of a call option, consider a trasformation $X_t = \ln S_t$, then the PDE becomes with constant coefficients:

$$u_t + \left(r - \frac{1}{2}\sigma^2\right) u_x + \frac{1}{2}\sigma^2 u_{xx} - ru = 0,$$

with a terminal condition $u(T, x) = (e^x - K)^+$. We also know that from $S(T) = S(t)e^{(r-\frac{1}{2}\sigma^2)(T-t)+\sigma(W(T)-W(t))}$ and therefore,

$$X(T) = X(0) + \left(r - \frac{1}{2}\sigma^2\right) T + \sigma(W(T) - W(0)),$$

which is a Gaussian random variable with mean $\bar{x} = X(0) + \left(r - \frac{1}{2}\sigma^2\right) T$ and variance $\sigma^2 T$.

Probabilistic approach

Suitable truncation of the domain can often be done probabilistically, based on the confidence interval for $X(T)$. Replace the unbounded domain with

$$[\bar{x} - \alpha\sigma\sqrt{T}, \bar{x} + \alpha\sigma\sqrt{T}].$$

- The likelihood of $X(T)$ falling outside of this interval is $2N(-\alpha)$
- For $\alpha = 4$, $2N(-\alpha) = 6.3 * 10^{-5}$, which is small for most applications.
- Good choice of $\alpha \in [3, 5]$ as this is optimizes the trade off between the error estimates and the numerical complexity.

Taylor expansion

Recall that a function $f(x)$ that is infinitely differentiable at the point $x = x_0$ can be represented as

$$f(x) = f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\frac{1}{3!}f''(x-x_0)^3+\frac{1}{(n+1)!}f^{(n+1}$$

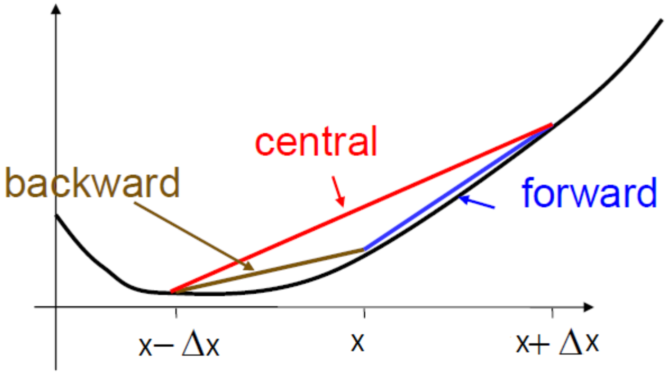
where ξ is in open interval (x_0, x). Note that if we truncate after the first term

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + O((x - x_0)^2)$$

and we can express

$$f'(x_0) = \frac{f(x) - f(x_0)}{x - x_0} + O(x - x_0)$$

Cont



Finite Difference Formulas

Thus, using Taylor series we can approximate the first derivative on the grid:

- **Forward difference:** $D_x^+ V(t, x_j) = \frac{V(x_{j+1}) - V(x_j)}{x_{j+1} - x_j}$
- **Backward difference:** $D_x^- V(t, x_j) = \frac{V(x_j) - V(x_{j-1})}{x_j - x_{j-1}}$
- **Central difference:** $D_x V(t, x_j) = \frac{V(x_{j+1}) - V(x_{j-1})}{x_{j+1} - x_{j-1}}$

And we can also approximate the second derivative on the grid:

- $D_{xx} V(t, x_j) = \frac{V(t, x_{j+1}) - 2V(t, x_j) + V(t, x_{j-1})}{(x_{j+1} - x_j)(x_j - x_{j-1})}$

Recommended reading

- [SM] p.6-9
- [S] Example 6.2.3, Section 7.3
- [TR] p.61-63, 67-70
- Uploaded to Blackboard document on Fichera function in finance

Numerical Methods 46.950

Lecture 3

September 13, 2016

Overview of the Lecture:

- Semi-discretization in space
- Time-discretization
- Explicit, Implicit, Crank-Nicolson methods
- Stability results

Mathematical formulation

Initially we consider a general parabolic terminal value problem

$$\frac{\partial V}{\partial t} + \mathcal{L}V = 0,$$

where $\mathcal{L}$ is the operator

$$\mathcal{L} = \mu(t,x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2}{\partial x^2} - r(t,x),$$

and where $V(t,x)$ satisfies a terminal condition $V(T,x) = f(x)$.
Underneath the PDE lies a physical model where a state variable process $X(t)$ follows an SDE of the form

$$dX(t) = \mu(t,X(t))dt + \sigma(t,X(t))dW(t)$$

in the risk-neutral probability measure $\mathbb{Q}$.

Semi-discretization

Let's treat the time variable t as special among the independent variables. Consider discretizing the spatial domain and derivatives first. This gives a *semi-discretization*, which leads a set of initial value ODEs.

The advantages are:

- Use of ODE methods: method of lines;
- Different approaches for the spatial discretization may be introduced;
- Analysis can often be simpler for the semi-discrete formulation.

The disadvantages(pitfalls) are:

- Not all useful full discretization methods can be obtained in this stepwise way;
- Number of ODEs is infinite as mesh size goes to zero.

Finite Difference Formulas

Formulas for thee first derivative on the grid:

- **Forward difference:** $D_x^+ V(t, x_j) = \frac{V(x_{j+1}) - V(x_j)}{x_{j+1} - x_j}$
- **Backward difference:** $D_x^- V(t, x_j) = \frac{V(x_j) - V(x_{j-1})}{x_j - x_{j-1}}$
- **Central difference:** $D_x V(t, x_j) = \frac{V(x_{j+1}) - V(x_{j-1})}{x_{j+1} - x_{j-1}}$

The formula for the second derivative on the grid:

- $D_{xx} V(t, x_j) = \frac{V(t, x_{j+1}) - 2V(t, x_j) + V(t, x_{j-1})}{(x_{j+1} - x_j)(x_j - x_{j-1})}$

Why this choice for the second derivative approximation? How do we find approximations for any partial derivatives? We use *method of undetermined coefficients*.

LTE for First derivative with central difference

Consider a uniform grid where $x_{j+1} - x_j = x_j - x_{j-1} = \delta x$. Then,

$$
\begin{aligned}
D_x V(t, x_j) &= \frac{V(x_{j+1}) - V(x_{j-1})}{x_{j+1} - x_{j-1}} \\
&= \frac{V(x_j + \delta x) - V(x_j - \delta x)}{2\delta x} \\
&= \frac{1}{2\delta x} \Bigg[V(x_j) + \delta x V_x(x_j) + \frac{\delta x^2}{2} V_{xx}(x_j) + \frac{\delta x^3}{6} V_{xxx}(\xi_1) \\
&- V(x_j) + \delta x V_x(x_j) - \frac{\delta x^2}{2} V_{xx}(x_j) + \frac{\delta x^3}{6} V_{xxx}(\xi_2) \Bigg] \\
&= V_x(x_j) + \frac{\delta x^2}{6} \frac{V_{xxx}(\xi_1) + V_{xxx}(\xi_2)}{2}.
\end{aligned}
$$

if V_{xxx} is continuous,

$$
V_x(x_j) - \frac{V(x_{j+1}) - V(x_{j-1})}{2\delta x} = -\frac{\delta x^2}{6} V_{xxx}(\xi)
$$

for some $\xi \in [x_{j-1}, x_{j+1}]$

LTE for Second derivative with central difference

For the second derivative approximation,

$$
\begin{aligned}
D_{xx}V(t,x_j) &= \frac{V(x_j+\delta x) - 2V(x_j) + V(x_j-\delta x)}{\delta x^2} \\
&= \frac{1}{\delta x^2}\Bigg[V(x_j) + \delta x V_x(x_j) + \frac{\delta x^2}{2}V_{xx}(x_j) + \frac{\delta x^3}{6}V_{3x}(x_j) + \frac{\delta x^4}{24}V_{4x}(\xi_1) \\
&\quad - 2V(x_j) + V(x_j) - \delta x V_x(x_j) + \frac{\delta x^2}{2}V_{xx}(x_j) - \frac{\delta x^3}{6}V_{3x}(x_j) + \frac{\delta x^4}{24}V_{4x}(\xi_2)\Bigg] \\
&= V_{xx}(x_j) + \frac{\delta x^2}{24}\left(V_{4x}(\xi_1) + V_{4x}(\xi_2)\right).
\end{aligned}
$$

if V_{4x} is continuous,

$$V_{xx}(x_j) - \frac{V(x_{j+1}) - 2V(x_j) + V(x_{j-1})}{\delta x^2} = -\frac{\delta x^2}{12}V_{4x}(\xi)$$

for some $\xi \in [x_{j-1}, x_{j+1}]$.

Finite Difference discretization in x-Direction

Consider replacing the partial derivatives with difference operators:

$$D_x V(t, x_j) = \frac{V(x_{j+1}) - V(x_{j-1})}{2\delta x},$$

$$D_{xx} V(t, x_j) = \frac{V(t, x_{j+1}) - 2V(t, x_j) + V(t, x_{j-1}))}{\delta x^2}$$

These operators are accurate to second order. i.e.

$$D_x V(t, x_j) = \frac{\partial V(t, x_j)}{\partial x} + O(\delta x^2),$$

$$D_{xx} V(t, x_j) = \frac{\partial^2 V(t, x_j)}{\partial x^2} + O(\delta x^2),$$

Finite Difference discretization in x-Direction

We introduce the discrete operator

$$\widehat{\mathcal{L}} = \mu(t,x)D_x + \frac{1}{2}\sigma^2(t,x)D_{xx} - r(t,x),$$

and for all $\{x_j\}_{j=1}^{m}$, we have

$$\mathcal{L}V(t,x) = \widehat{\mathcal{L}}V(t,x) + O(\delta x^2).$$

With the Dirichlet boundary condition at x_0 and x_m

$$V(t,x_0) = g_1(t,x_0), V(t,x_m) = g_2(t,x_m)$$

$$\widehat{\mathcal{L}}\mathbf{V}(t) = \mathbf{A}(t)\mathbf{V}(t) + \mathbf{\Sigma}(t)$$

where $\mathbf{\Sigma}(t)$ is a vector containing boundary values.

Tridiagonal matrix

we can view $A(t)$ as a matrix:

$$\mathbf{A}(t) = \begin{pmatrix} c_1(t) & u_1(t) & 0 & 0 & 0 & \cdots & 0 \\ l_2(t) & c_2(t) & u_2(t) & 0 & 0 & \cdots & 0 \\ 0 & l_3(t) & c_3(t) & u_3(t) & 0 & \cdots & 0 \\ 0 & 0 & l_4(t) & c_4(t) & u_4(t) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 & \\ 0 & 0 & 0 & 0 & l_{m-1}(t) & c_{m-2}(t) & u_{m-2}(t) \\ 0 & 0 & 0 & 0 & 0 & l_{m-1}(t) & c_{m-1}(t) \end{pmatrix}$$

Elements of the $\mathbf{A}(t)$ matrix and $\mathbf{\Sigma}(t)$

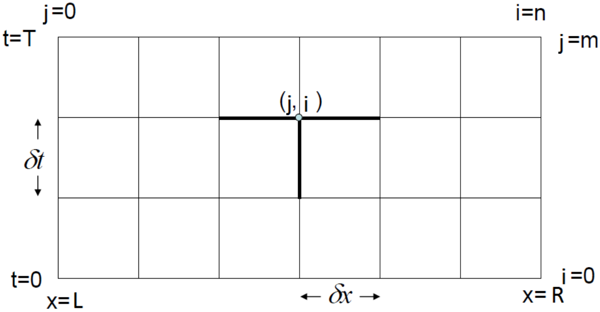
$$
\begin{aligned}
c_j(t) &= -\sigma(t, x_j)^2 \frac{1}{\delta x^2} - r(t, x_j), \\
u_j(t) &= \frac{1}{2}\mu(t, x_j)\frac{1}{\delta x} + \frac{1}{2}\sigma(t, x_j)^2 \frac{1}{\delta x^2}, \\
l_j(t) &= -\frac{1}{2}\mu(t, x_j)\frac{1}{\delta x} + \frac{1}{2}\sigma(t, x_j)^2 \frac{1}{\delta x^2},
\end{aligned}
$$

and

$$
\mathbf{\Sigma}(t) = \begin{pmatrix} l_1(t) g_1(t, x_0) \\ 0 \\ \vdots \\ 0 \\ u_{m-1}(t) g_2(t, x_m) \end{pmatrix}.
$$

Grid notation

Consider a bounded domain $[0, T]$ in time and $[L, R]$ in space. And consider the equidistant grids $\{t^i\}_{i=0}^n$ and $\{x_i\}_{j=0}^m$ where $t^i = iT/n = i\delta t$ and $x_j = L + j(R-L)/m = L + j\delta x$. The terminal value $V(T, x) = f(x)$ is imposed at $t^n = T$.



Time discretization

Assume $g_1 = g_2 = 0$ and thus $\mathbf{\Sigma}(\mathbf{t}) = 0$, then the original PDE can be written as

$$\frac{\partial \mathbf{V}(t)}{\partial t} = -\mathbf{A}(t)\mathbf{V}(t) + O(\delta x^2)$$

which defines a system of coupled ordinary differential equations! Thus, on a particular interval $[t^i, t^{i+1}]$, the choice for the approximation is:

$$\frac{\partial \mathbf{V}}{\partial t} \approx \frac{\mathbf{V}(t^{i+1}) - \mathbf{V}(t^i)}{\delta t}.$$

But to which time in the interval $[t^i, t^{i+1}]$ we should associate this derivative?

Consider $t_i^{i+1}(\theta) \in [t^i, t^{i+1}]$ given by

$$t_i^{i+1}(\theta) = (1-\theta)t^{i+1} + \theta t^i,$$

for $\theta \in [0, 1]$ Then we write

$$\frac{\partial \mathbf{V}\left(t_i^{i+1}(\theta)\right)}{\partial t} \approx \frac{\mathbf{V}(t^{i+1}) - \mathbf{V}(t^i)}{\delta t}.$$

Theta scheme

One can show by Taylor expansion,

$$\frac{\partial \mathbf{V}\left(t_i^{i+1}(\theta)\right)}{\partial t} = \frac{\mathbf{V}(t^{i+1}) - \mathbf{V}(t^i)}{\delta t} + 1_{\{\theta \neq \frac{1}{2}\}} O(\delta t) + O(\delta t^2)$$

This time-discretization technique is known as *theta scheme*. The special cases are:

- $\theta = 1$: fully implicit scheme:

$$(\mathbf{I} - \delta t \mathbf{A}(t^i))\mathbf{V}(t^i) = \mathbf{V}(t^{i+1})$$

- $\theta = 0$: fully explicit scheme

$$\mathbf{V}(t^i) = (\mathbf{I} + \delta t \mathbf{A}(t^{i+1}))\mathbf{V}(t^{i+1})$$

- $\theta = \frac{1}{2}$: Crank-Nicolson scheme

$$(\mathbf{I} - \frac{1}{2}\delta t \mathbf{A}(t^i))\mathbf{V}(t^i) = (\mathbf{I} + \frac{1}{2}\delta t \mathbf{A}(t^{i+1}))\mathbf{V}(t^{i+1})$$

Local truncation errors

Combining the error from the space and time discretizations one can calculate the locat truncation errors:

- **Explicit:** $O(\delta t) + O(\delta x^2)$
- **Implicit:** $O(\delta t) + O(\delta x^2)$
- **Crank-Nicolson:** $O(\delta t^2) + O(\delta x^2)$

Explicit scheme for heat equation

Consider the backward heat equation

$$U_t + U_{xx} = d(t, x)$$

First step is to discretize in x direction:

$$U_t + \frac{U(t, x_{j-1}) - 2U(t, x_j) + U(t, x_{j+1})}{\delta x^2} = d(t, x_j)$$

Then, consider approximation of time derivative at t^i, x_j, i.e. $\theta = 0$:

$$U_t \approx \frac{U(t^{i+1}, x_j) - U(t^i, x_j)}{\delta t}$$

denoting $U(t^i, x_j)$ as U_j^i the FD scheme can be written as

$$\frac{U_j^{i+1} - U_j^i}{\delta t} + \frac{U_{j-1}^{i+1} - 2U_j^{i+1} + U_{j+1}^{i+1}}{\delta x^2} = d_j^{i+1}$$

Implicit scheme for heat equation

Similarly, we can derive the implicit finite difference scheme approximating the time derivative at (t^i) according to $\theta = 1$:

$$\frac{U_j^{i+1} - U_j^i}{\delta t} + \frac{U_{j-1}^i - 2U_j^i + U_{j+1}^i}{\delta x^2} = d_j^i$$

Crank-Nicolson for heat equation

Crank-Nicolson Scheme corresponds to approximating time derivative at $t = \frac{1}{2}(t^i + t^{i+1})$, therefore the space approximation is a combination of the approximation at t^i, x_j and t^{i+1}, x_j:

$$\begin{aligned}\frac{U_j^{i+1} - U_j^i}{\delta t} &+ \frac{1}{2}\frac{U_{j-1}^i - 2U_j^i + U_{j+1}^i}{\delta x^2} \\ &+ \frac{1}{2}\frac{U_{j-1}^{i+1} - 2U_j^{i+1} + U_{j+1}^{i+1}}{\delta x^2} = d(t^{i+\frac{1}{2}\delta t}, x_j)\end{aligned}$$

which is basically a blend between explicit and implicit FDMs.

Gauss elimination without pivoting

For the implicit finite difference methods we end up with a tridiagonal system of the form

$$Mx = r,$$

which can be solved efficiently by Gaussian Elimination without pivoting method, also known as **Thomas algorithm**. It is based on LU decomposition and where the above matrix equation can be written as

$$LUx = r,$$

and solved in 2 steps: setting $Ux = \rho$, and solving first for ρ

$$L\rho = r,$$

and then solving for x

$$Ux = \rho.$$

Stability for the explicit FDM

Explicit Finite Difference Method can be written as:

$$U_j^i = \rho U_{j-1}^{i+1} + (1 - 2\rho) U_j^{i+1} + \rho U_{j+1}^{i+1}$$

where $U_j^i = U(t^i, x_j)$ and $\rho = \frac{\delta t}{\delta x^2}$.
Notice the difference in the results if we take $\rho = 1$ or $\rho = \frac{1}{2}$. Thus, the stability of the scheme depends on the value of ρ.

Explicit method results

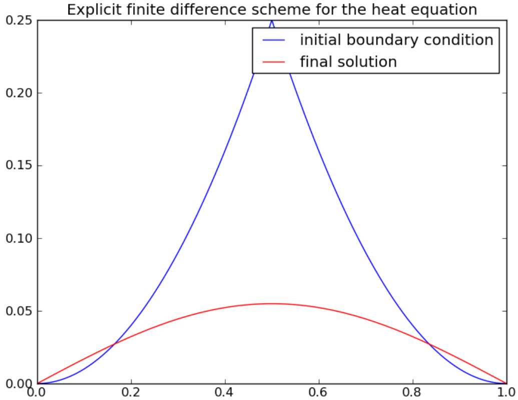


Figure: Solution when we choose $\delta t \leq 1/2\delta x^2$

Cont

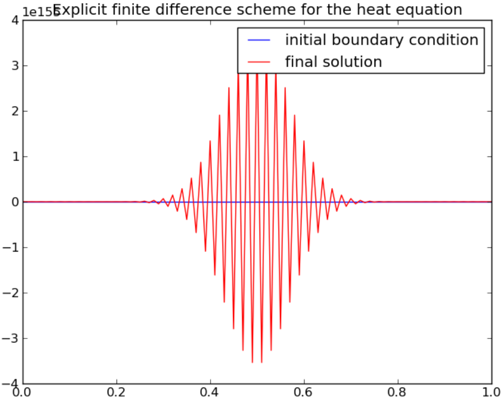


Figure: Solution when we choose $\delta t > 1/2\delta x^2$

Sufficient stability condition for explicit method

Consider the explicit finite difference method which can be written as

$$U_j^i = aU_{j-1}^{i+1} + bU_j^{i+1} + cU_{j+1}^{i+1}$$

where a, b, c are constants and $j = 1, \ldots, m-1$ are all interior domain points the equation to be solved for.
Assume $a, b, c \geq 0$ and $a + b + c \leq 1$, then

$$\begin{aligned} |U_j^i| &\leq a|U_{j-1}^{i+1}| + b|U_j^{i+1}| + c|U_{j+1}^{i+1}| \\ &\leq (a+b+c)\max_j |U_j^{i+1}| \leq \max_j |U_j^{i+1}| \end{aligned}$$

Taking maximum over all j on the left we deduce

$$\max_j |U_j^i| \leq \max_j |U_j^{i+1}|$$

Iterating over all i, we obtain $\max_j |U_j^0| \leq \max_j |U_j^n| = \max_j |f(x_j)|$, therefore the solution is bounded by the data and the equation is stable in that sense.

Stability for the implicit FDM

Any $\theta \geq \frac{1}{2}$ produces an implicit finite difference scheme. In particular, fully implicit Finite Difference Method can be expressed as:

$$-\rho U_{j-1}^{i} + (1+2\rho)U_j^i - \rho U_{j+1}^i = U_j^{i+1}$$

We rewrite it as

$$aU_{j-1}^{i} + bU_j^i + cU_{j+1}^i = U_j^{i+1}$$

Assume $a, c \leq 0, b \geq 0$, and $b + a + c \geq 1$, then

$$\begin{aligned} bU_j^i &= -aU_{j-1}^i - cU_{j+1}^i + U_j^{i+1}, \\ b|U_j^i| &= -a|U_{j-1}^i| - c|U_{j+1}^i| + U_j^{i+1} \\ &\leq (-a-c)M_i + \max_j |U_j^{i+1}| \end{aligned}$$

where $M_i = \max_j |U_j^i|$.

Cont

Taking j on the LHS such that $M_i = U_j^i$

$$bM_i \leq (-a-c)M_i + \max_j |U_j^{i+1}|,$$
$$(a+b+c)M_i \leq \max_j |U_j^{i+1}|$$

and therefore

$$M_i \leq \frac{1}{(a+b+c)} \max_j |U_j^{i+1}| \leq \max_j |U_j^{i+1}|$$

Iterating over all i we get an estimate

$$\max_j |U_j^0| \leq \max_j |U_j^n| = \max_j |f(x_j)|$$

Cont

Therefore we showed

$$\max_j |U_j^i| \leq \max_j |U_j^{i+1}|$$

Repeating iteratively,

$$\max_j |U_j^0| \leq^n \max_j |U_j^n| =^n \max_j |f(x_j)|$$

Lax Equivalence Theorem

Theorem: For a well-posed linear terminal value PDE, a consistent finite difference scheme is convergent if and only if it is stable.

- Consistency means local approximation of derivatives
- Stability means boundedness by the initial(or terminal payoff) data

Recommended reading

- [TR]: Chapter 3
- [SM]: p12-29, p38-59, p71-75
- Uploaded to Blackboard document on Thomas algorithm

Numerical Methods 46.950

Lecture 4

September 20, 2016

Overview of the Lecture:

- Different ways to impose and implement boundary conditions;
- Matrix and Fourier stability approaches;
- Connection of the Finite Difference Methods with Binomial Trees;
- Final Error estimates;
- Common problems with Crank-Nicolson and remedies to handle them;
- Barrier options;

Boundary Conditions

One has to be careful how and what type of boundary conditions(BC) to impose. Bad BC can spoil the solution or its properties. Several options for the choice to consider:

- Asymptotic behavior;
- Known boundary values;
- Use equation itself;
- Use modified pricing equation;

Implementation of Dirichlet BC

Approach 1: Embed the BC in discretization.
Consider the system

$$A_L \mathbf{V}^{i-1} = A_R \mathbf{V}^i$$

that arises when PDE is discretized. The unknowns are $V_0, V_1, \ldots, V_m$, but the discrete equations are written only at grid points $S_1, S_2, \ldots, S_{m-1}$. Then we can embed BC in discretization and known boundary values drop out of the equation.

- Makes A_L a square matrix, however A_R is preserved and rectangular;
- Need to create an additional vector and work with a combination of square and rectangular matrices;

Approach 2: Expand the system by adding equations for the boundary unknowns.

- All matrices are of the same size;
- However, the approximations can ruin the structure of the matrix A_L(see Homework)

Ghost unknowns For Dirichlet type BC

The approximation formula of the BC must at least match the order of accuracy of the interior formula, or else a deterioration in the global error may result, at least locally. Consider approximating Neumann BC $U'(x_0) = g_0(t)$ for heat equation $U_t + U_{xx} = 0$ at the left boundary:

$$\begin{aligned}\frac{dU_0}{dt} + \frac{U_{-1} - 2U_0 + U_1}{\delta x^2} &= 0,\\ \frac{U_1 - U_{-1}}{2\delta x} &= g_0(t)\end{aligned}$$

We call U_{-1} a *ghost unknown*. And thus, U_{-1} can be eliminated

$$U_{-1} = U_1 - 2\delta x g_0(t)$$

hence

$$\frac{dU_0}{dt} + \frac{2U_1 - 2U_0 - 2\delta x g_0(t)}{\delta x^2} = 0.$$

Note that the system is still tridiagonal.

Liner BC

In a large number of option structures, the option value is nearly linear with respect to spot price. Which translates to the simple boundary condition

$$\frac{\partial^2 V}{\partial S^2}(t, S_{max}) = 0$$

and if we use a downward discretization of the second derivative.

Another example of BC

If the PDE discretized in the logarithm of some asset, it may be natural to assume that $V(t,x) \approx e^x$ at the boundaries, or, equivalently,

$$\frac{\partial V}{\partial x} = \frac{\partial^2 V}{\partial x^2}$$

then on the grid, it is approximated as

$$\frac{V(t,x_m) - V(t,x_{m-1})}{\delta x} = \frac{V(t,x_m) - 2V(t,x_{m-1}) + V(t,x_{m-2})}{\delta x^2}$$

from where the boundary condition for $V(t,x_m)$ and $V(t,x_0)$

$$V(t,x_m) = V(t,x_{m-2})\frac{1}{\delta x - 1} + V(t,x_{m-1})\frac{\delta x - 2}{\delta x - 1},$$

$$V(t,x_0) = V(t,x_1)\frac{2+\delta x}{1+\delta x} - V(t,x_2)\frac{1}{\delta x + 1}.$$

Stability/Matrix methods

Ignoring the contributions from boundary conditions, the finite difference scheme can be written as

$$\mathbf{V}(t_i) = B_i^{i+1}\mathbf{V}(t_{i+1}),$$

where

$$B_i^{i+1} = \left(I - \theta\delta t\mathbf{A}(t_i^{i+1}(\theta))\right)^{-1}\left(I + (1-\theta)\delta t\mathbf{A}(t_i^{i+1}(\theta))\right).$$

That is, for any $0 \leq k < n$,

$$\mathbf{V}(t_k) = B_k^n\mathbf{V}(t_k), \quad B_k^n = B_k^{k+1}B_{k+1}^{k+2}\dots B_{n-1}^n.$$

We say that the scheme is *stable* if $||\mathbf{V}(t_k)||$ is bounded for all $0 \leq k < n$. Thus, a necessary and and sufficient condition for stability is

$$||B_k^n|| \leq K$$

for any matrix norm.

Stability/Von Neumann Analysis

The basis for the Von Neumann analysis is the fact that a real function sampled on a finite number of points is uniquely defined by a complex Fourier series:

$$V(t_k, x_j) = \sum_l H_l(t_k) e^{i\omega_l j\delta x},\ l = 0, \dots, m$$

where $H_l(t_k)$ is the amplification factor and ω_l is the wave number for l-th mode. A key fact for our PDE problem is that:

$$H_l(t_{k-1}) = H_l(t_k)\xi_l$$

where ξ_l is the *amplification factor* independent of time. To determine how a solution is propagated back through the finite difference grid, it thus suffices to consider a test function of the form

$$V(t_k, x_j) = \xi(\omega)^{n-k} e^{i\omega j\delta x}.$$

According to the Von-Neumann criterion, stability requires that:

$$|\xi(\omega)| \leq 1$$

Application: stability proof for the explicit scheme

Consider the forward time-centered space approximation for the heat equation:

$$\frac{V(t_{k-1},x_j)-V(t_k,x_j)}{\delta t}+\frac{V(t_k,x_{j-1})-2V(t_k,x_j)+V(t_k,x_{j+1})}{\delta x^2}=0.$$

where we used index k for time not to confuse with complex i. Solving for $V(t_k,x_j)$ gives:

$$\begin{aligned}V(t_{k-1},x_j)&=\frac{\delta t}{\delta x^2}V(t_k,x_{j-1})\\&\quad+(1-2\frac{\delta t}{\delta x^2})V(t_k,x_j)+\frac{\delta t}{\delta x^2}V(t_k,x_{j+1})\end{aligned}$$

Application: stability proof for the explicit scheme

Substitute $\xi(\omega)^{n-k} e^{i\omega j\delta x}$ for $v(t_k, x_j)$:

$$\xi(\omega)^{n-k+1} e^{i\omega j\delta x} = \frac{\delta t}{\delta x^2}\xi(\omega)^{n-k} e^{i\omega(j-1)\delta x} + (1 - 2\frac{\delta t}{\delta x^2})\xi(\omega)^{n-k} e^{i\omega j\delta x} + \frac{\delta t}{\delta x^2}\xi(\omega)^{n-k} e^{i\omega(j+1)\delta x}$$

Substituting trigonometric functions for the complex exponential leads to

$$\xi(\omega) = 1 - 2\frac{\delta t}{\delta x^2} + \frac{\delta t}{\delta x^2} e^{i\omega\delta x} + \frac{\delta t}{\delta x^2} e^{-i\omega\delta x} = 1 - 2\frac{\delta t}{\delta x^2}(1 - \cos\omega\delta x)$$

To prevent amplification of the ω wave mode we must have

$$|\xi(\omega)| \leq 1$$

meaning that $-1 \leq 1 - 2\frac{\delta t}{\delta x^2}(1 - \cos\omega\delta x) \leq 1$ from which

$$0 \leq \frac{\delta t}{\delta x^2}(1 - \cos\omega\delta x) \leq 1.$$

Thus we require $\frac{\delta t}{\delta x^2} \leq \frac{1}{2}$ as $(1 - \cos\omega\delta x)$ can be 2 at most.

BC and stability

- Do we need to worry about stability when imposing boundary conditions?
- It turns out that for parabolic PDEs, sensible consistent BC approximations turn out to be stable.
- In practice, if the terminal condition and boundary condition do not agree at their points of meeting, it makes sense then to interpret the boundary conditions as holding only for $t < T$, to avoid ambiguity.

Connection between binomial trees and FDMs

- There is a connection between binomial trees seen previously in earlier courses and finite difference methods! In fact, trees are particular forms of finite difference schemes!
- A good finite difference solver is faster, more robust and more flexible compare to binomial/trinomial trees;

Binomial tree setup

We consider the Black-Scholes setup with the dynamics under risk-neutral measure:

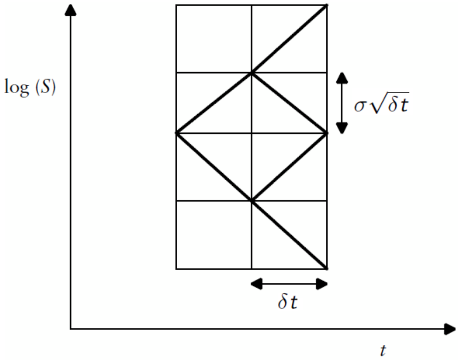
$$dS(t) = rS(t)dt + \sigma S(t)d\widetilde{W}(t),$$

and the moves up and down factors on the tree are

$$u := e^{\sigma\sqrt{\delta t}}, \quad d := e^{-\sigma\sqrt{\delta t}}$$

where δt is the spacing of the tree in time. Note that $\sigma\sqrt{\delta t}$ is the spacing of nodes in the logarithm of the price process $S(t)$. Then if r is a continuously compounded annual interest rate, then \$1 invested over a time interval of length δt grows to \$$e^{r\delta t}$. Therefore, for a period of length δt, the per period simple interest rate is $e^{r\delta t} - 1$.

Cont



Cont

Recall from Multi-Period Asset Pricing for the risk neutral probability of an up move is:

$$\tilde{p} = \frac{e^{r\delta t} - e^{-\sigma\sqrt{\delta t}}}{e^{\sigma\sqrt{\delta t}} - e^{-\sigma\sqrt{\delta t}}}$$

and Taylor series gives

$$\tilde{p} = \frac{1}{2} + \frac{1}{2\sigma}\left(r - \frac{\sigma^2}{2}\right)\sqrt{\delta t} + O(\delta t^{3/2})$$

and this converges to $\frac{1}{2}$ as δt approaches 0.

Taylor expansion for

Expand numerator and denominator separately. The numerator is easy:

$$e^{r\delta t} - e^{-\sigma\sqrt{\delta t}} = (1 + r\delta t) - (1 - \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t - \frac{1}{6}\sigma^3\delta t^{3/2}) + O(\delta t^2)$$

The denominator can be shown to be

$$\begin{aligned} e^{\sigma\sqrt{\delta t}} - e^{-\sigma\sqrt{\delta t}} =& (1 + \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t + \frac{1}{6}\sigma^3\delta t^{3/2}) \\ & - (1 - \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t - \frac{1}{6}\sigma^3\delta t^{3/2}) + O(\delta t^2) \\ &= 2\sigma\sqrt{\delta t} + \frac{1}{3}\sigma^3\delta t^{3/2} + O(\delta t^2) \\ &= 2\sigma\sqrt{\delta t}(1 + \frac{1}{6}\sigma^2\delta t + O(\delta t^{3/2})) \\ &= 2\sigma\sqrt{\delta t}(1 + x) \end{aligned}$$

where

$$x = \frac{1}{6}\sigma^2\delta t + O(\delta t^{3/2}).$$

Cont

The denominator can be then expended

$$\frac{1}{2\sigma\sqrt{\delta t}(1+x)} \approx \frac{1}{2\sigma\sqrt{\delta t}}(1-x)$$
$$= \frac{1}{2\sigma\sqrt{\delta t}} - \frac{1}{12}\sigma\sqrt{\delta t} + O(\delta t)$$

Combining numerator and denominator

$$\tilde{p} = \frac{e^{r\delta t} - e^{-\sigma\sqrt{\delta t}}}{e^{\sigma\sqrt{\delta t}} - e^{-\sigma\sqrt{\delta t}}} = \left[\sigma\sqrt{\delta t} + \left(r - \frac{1}{2}\sigma^2\right)\delta t + \frac{1}{6}\sigma^3\delta t^{3/2} + O(\delta t^2)\right]$$
$$\left[\frac{1}{2\sigma\sqrt{\delta t}} - \frac{1}{12}\sigma\sqrt{\delta t} + O(\delta t)\right] = \frac{1}{2}$$
$$+ \frac{1}{2\sigma}\left(r - \frac{1}{2}\sigma^2\right)\sqrt{\delta t} + \frac{1}{12}\sigma^2\delta t - \frac{1}{12}\sigma^2\delta t + O(\delta t^{3/2})$$
$$= \frac{1}{2} + \frac{1}{2\sigma}\left(r - \frac{1}{2}\sigma^2\right)\sqrt{\delta t} + O(\delta t^{3/2})$$

Finite difference method setup

Define $X(t) = \ln S(t)$, then the Black Scholes PDE becomes:

$$u_t + \left(r - \frac{1}{2}\sigma^2\right) u_x + \frac{1}{2}\sigma^2 u_{xx} - ru = 0$$

and we can apply explicit finite difference scheme to solve it:

$$\begin{aligned}
u_j^i &= Au_{j-1}^{i+1} + Bu_j^{i+1} + Cu_{j+1}^{i+1} \\
&= \left(\frac{1}{2}\sigma^2 \frac{\delta t}{\delta x^2} - \left(r - \frac{1}{2}\sigma^2\right)\frac{\delta t}{2\delta x}\right) u_{j-1}^{i+1} \\
&+ \left(1 - \sigma^2 \frac{\delta t}{\delta x^2} - r\delta t\right) u_j^{i+1} \\
&+ \left(\frac{1}{2}\sigma^2 \frac{\delta t}{\delta x^2} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\delta t}{2\delta x}\right) u_{j+1}^{i+1}
\end{aligned}$$

Cont

Setting the term $\left(1-\sigma^2\frac{\delta t}{\delta x^2}-r\delta t\right)$ multiplying u_j^{i+1} to zero, we get

$$\delta x=\sigma\sqrt{\delta t}+H.O.T.$$

Substitute this into the term $\left(\frac{1}{2}\left(r-\frac{1}{2}\sigma^2\right)\frac{\delta t}{\delta x}+\frac{1}{2}\sigma^2\frac{\delta t}{\delta x^2}\right)$ multiplying u_{j+1}^{i+1} and get

$$C=\frac{1}{2}+\frac{1}{2\sigma}\left(r-\frac{\sigma^2}{2}\right)\sqrt{\delta t}+H.O.T$$

which is precisely the p from the binomial tree model. Thus, there is a connection between the two and therefore trees inherit all the problems of the explicit finite difference schemes(stability). However, handling boundaries is much easier using FDM compare to trees.

Truncation error estimates

Basic principle of FDM is to replace partial derivatives by algebraic approximations. This process called *discretization*. This gives rise to local truncation errors for different methods:

- **Explicit:** $O(\delta t) + O(\delta x^2)$
- **Implicit:** $O(\delta t) + O(\delta x^2)$
- **Crank-Nicolson:** $O(\delta t^2) + O(\delta x^2)$

Truncation error is referred to as *accuracy of the discrete approximation*. TE is not the actual error! But what is the actual error between the numerical solution and the actual value at any given point x_j at time $t = 0$ for example?

Error estimate for the explicit FDM

Stability and error estimation for the difference approximations are closely connected topics. The connection comes about through the **error equation**. For example in the case of a heat equation $U_t + U_{xx} = 0$ the explicit FDM can be written

$$\frac{\hat{U}_j^{i+1} - \hat{U}_j^i}{\delta t} + \frac{\hat{U}_{j-1}^{i+1} - 2\hat{U}_j^{i+1} + \hat{U}_{j+1}^{i+1}}{\delta x^2} = 0$$

while the equation for the actual function U is

$$\begin{aligned}\frac{U_j^{i+1} - U_j^i}{\delta t} + \frac{U_{j-1}^{i+1} - 2U_j^{i+1} + U_{j+1}^{i+1}}{\delta x^2} \\ &= U_t + U_{xx} + O(\delta t) + O(\delta x^2) \\ &= U_t + U_{xx} + \tau_j^i\end{aligned}$$

Cont

Define $\epsilon_j^i := \hat{U}_j^i - U_j^i$, then the equation for the error becomes

$$\frac{\epsilon_j^{i+1} - \epsilon_j^i}{\delta t} + \frac{\epsilon_{j-1}^{i+1} - 2\epsilon_j^{i+1} + \epsilon_{j+1}^{i+1}}{\delta x^2} = -\tau_j^i$$

or in other notation

$$\epsilon_j^i = \rho\epsilon_{j-1}^{i+1} + (1-2\rho)\epsilon_j^{i+1} + \rho\epsilon_{j+1}^{i+1} + \delta t \tau_j^i$$

Note that if $i+1 = n$, all $\epsilon_j^n = 0$, therefore, if $\rho \leq \frac{1}{2}$ then taking max over j:

$$\max_j \epsilon_j^i \leq \max_j \epsilon_j^{i+1} + \delta t \tau \leq \max_j \epsilon_j^{i+2} + 2\delta t \tau \leq \ldots \leq \max_j \epsilon_j^n + n\delta t \tau$$

where τ denotes the maximal LTE on the grid. Thus, one can show that for a stable finite difference method the solution error is bounded proportional to the truncation error:

$$|\epsilon| \leq K|\tau|$$

Spurious oscillation for non-smooth data

- In general, CN is well-suited for approximating parabolic problems.
- However it may suffer from spurious oscillations which do not decay quickly. In fact, it can be renewed at every monitoring date.

In the case of application of CN scheme to the heat equation

$$U_t + U_{xx} = 0,\ x \in [0, \pi]$$

one can derive the amplification factor to be:

$$\xi(\omega) = \frac{1 - \frac{\delta t}{\delta x^2} sin^2\left(\frac{\omega \delta x}{2}\right)}{1 + \frac{\delta t}{\delta x^2} sin^2\left(\frac{\omega \delta x}{2}\right)}$$

The amplification factor $\xi(\omega)$ has to satisfy $|\xi(\omega)| \leq 1$. However, when $\omega = m, m-1, \ldots,$ we use the fact that $\delta x = \frac{\pi}{m}$ and therefore $\xi(\omega) \approx -1$. Meaning those components that correspond to large ω don't get damped and oscillations can occur as high wave Fourier components change the sign.

Cont

One can show theoretically that the oscillations can be controlled if the following(approximate) condition holds:

$$\delta t < \frac{\delta x^2}{2}$$

Which is a serious restriction on a time step.
There are several approaches how to avoid the strict constraint on a time step:

- Perform one or more fully implicit time steps before applying Crank-Nicolson;
- Smooth the payoff function;

For best results, apply both!

Implicit time stepping

One way to control the oscillations is to perform one or two implicit time steps before applying Crank-Nicolson scheme. It is also called *Rannacher stepping*.

As an example, consider a digital call having knock out barriers at $x = 0, 100$. parameters are: $\sigma = 10\%$, $T = 1$, $r = 12\%$. Maturity is one year and the derivative pays \$1 if the terminal spot value is greater than the strike $K = 50$. Results for Crank-Nicolson and Rannacher time stepping shown next.

Digital call results

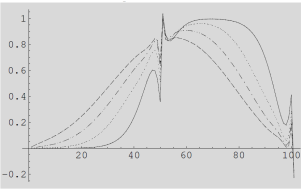


Figure: Crank-Nicolson

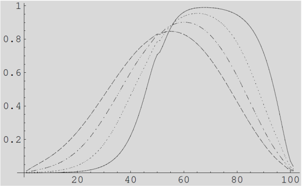


Figure: first implicit time step and then CN

Non-smooth payoff/Quantization error

Another source of error observed in finite difference solutions is *quantization error*. By setting option expiration value on the finite difference grid to the payoff:

$$V(T, S_j) = f_j = f(S_j)$$

can lead to omitting the irregularities between grid points if the payoff function $f(S)$ or its derivative is not smooth.

To reduce this error, and smooth the payoff function, we must ensure that the calculated present value of an option depends as little as possible on the location of the strike relative to the grid.

One can use the following *continuity correction*: the value of a function on a grid represents the average over the surrounding grid cell instead of the value sampled at each grid point:

$$f_j \approx \frac{1}{\delta S} \int_{S_j - \delta S/2}^{S_j + \delta S/2} f(s) ds \neq f(S_j)$$

Non-smooth payoff/Grid shifting

Another way to decrease the quantization error is to shift the grid. Consider a digital call option with

$$f(x) = \mathbb{1}_{\{x>H\}}$$

where H is located between x_k and x_{k+1}. The the continuity correction only affects $V(x_k)$ and $V(x_{k+1})$ and sets it is value between zero and one. Then the solution is to arrange the spatial grid such that x-values where the payoff(or its derivatives) is discontinuous are exactly midway between the grid nodes. Aligning the grid in this way will, in a loose sense, make the payoff smooth.

More options examples

So far we have been looking at options with some payoff function f. In reality, many options are more complicated than this and may involve early exercise decision, pre-maturity cash flows, path dependancy and more. Let's look at the following types of options:

- Continuous barrier options;
- Discrete barrier options;

Continuous barrier options

- Many of these options involve "up" and "down" type of barriers, some of them give a payout(rebate) at the time the barrier is hit, thus we impose
$$V(t, x_0) = g_1(t)$$
and
$$V(t, x_1) = g_2(t)$$
- Some options involve time dependent barriers, e.g. *step-up* and *step-down* options;

Up-and-out single barrier option

- Consider as an example zero-rebate up-and-out single-barrier option, s.t. the barrier moves from value H^* to H at some future time $T^* < T$.
- Extension to of the FD algorithm is straightforward to cover this case: We start with domain $[x_{min}, H]$ and after time T^* the PDE applies only to $[x_{min}, H^*]$.
- Make sure that there is a grid point at exactly H^*.

Discrete barrier options

Consider a double barrier option: knock-out option, i.e. a call option that expires worthless if one of the two barriers has been hit at a monitoring date. Define a set of monitoring dates $T_0, T_1, \ldots, T_n$, lower barrier S_d and upper barrier S_u that are monitored only at times $T_i, i = 0, 1, \ldots, n$. The price $V(t, S)$ satisfies the Black-Scholes PDE with initial and boundary conditions:

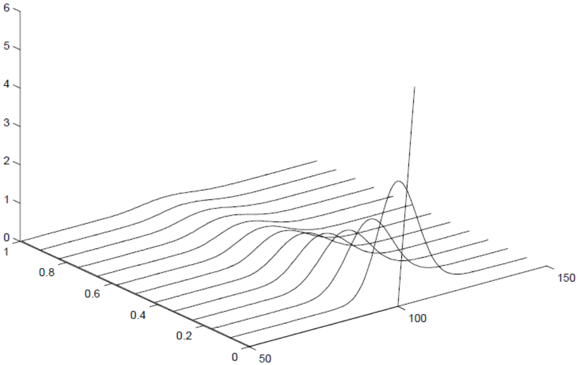
$$V_t + rSV_S + \frac{1}{2}\sigma^2 S^2 V_{SS} - rV = 0,$$
$$V(T, S) = (S - K)^+ \mathbb{1}_{\{S_d, S_u\}}(S),$$
$$V(0, T) = 0, V(S_{max}) = 0.$$

and the discrete monitoring introduces an updating of the solution at the monitoring times:

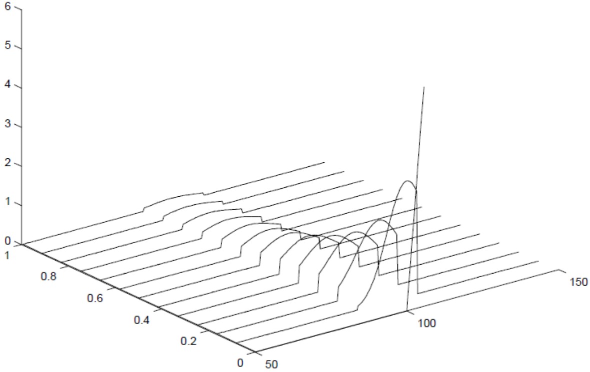
$$V(T_i, S) = V(T_i, S)\mathbb{1}_{\{S_d, S_u\}}(S)$$

where $\mathbb{1}_{[S_d, S_u]}(S) = 1$ if $S \in [S_d, S_u]$ and 0 otherwise.

Barrier prices before the monitoring times



Barrier prices after the monitoring times



Recommended reading

- [TR]: Relevant parts of Chapter 4 and 5;
- [SM]: pages 45-124;
- Paper on convergence remedies by Pooley, Vetzal and Forsyth uploaded to Blackboard.

Numerical Methods 46.950

Lecture 5

September 27, 2016

Overview of the Lecture:

- Comparison of far Boundary condition for Black Scholes
- Convection Dominated PDE
- Upwinding
- Options on assets with dividends;
- Early exercise and American options;

The $s = 0$ boundary condition for BS PDE

We have seen before that BS PDE degenerates at $s = 0$, i.e. the highest order derivative term coefficient $\frac{\sigma^2 s^2}{2}$ becomes zero on the boundary $s = 0$. Thus, use equation itself and plug $s = 0$ into equation and obtain

$$V_t - rV = 0.$$

In order to discretize it, we apply Crank-Nicolson discretization at $s = s_0$. Luckily, there are no space derivatives and therefore we don't need to think about space discretization, only time derivative:

$$\frac{V_0^{i+1} - V_0^i}{\delta t} - r\left(\frac{1}{2}V_0^i + \frac{1}{2}V_0^{i+1}\right) = 0$$

that after rearrangement becomes

$$\left(1 + \frac{r\delta t}{2}\right) V_0^i = \left(1 - \frac{r\delta t}{2}\right) V_0^{i+1}$$

which we use for the $j = 0$ grid point at every new time level t^i.

Different far boundary conditions comparison

Can we quantify the difference between alternative far boundary conditions? Consider pricing a European call option in Black-Scholes framework with parameters $T = 1, \sigma = .25, r = .05$ and strike $K = 100$. Thus, we are solving Black-Scholes PDE on a bounded domain:

$$V_t + rsV_s + \frac{\sigma^2 s^2}{2} V_{ss} - rV = 0, \quad s \in [0, \hat{s}],\ t \in [0, T)$$

and payoff $f(s) = (s - K)^+$. Let's compare different choices for the far boundary condition:

- $V(t, \hat{s}) = f(\hat{s})$
- $V(t, \hat{s}) = e^{-r(T-t)} f(\hat{s})$
- $V_{ss}(t, \hat{s}) = 0$

Note, the first two carry a benefit of no discontinuity at $t = T$ between the terminal payoff and BC. We compare them quantitatively next.

Comparison results in the explicit FDM

Errors at K when $\hat{s} = 3K$				
m	n	e	e_x	e_t
60	226	0.010520	0.000725	0.017502
120	901	0.002600	0.000173	0.004247
240	3601	0.000653	0.000037	0.000999
480	14401	0.000169	0.000003	0.000191

$V(t, \hat{s}) = f(\hat{s})$

Errors at K when $\hat{s} = 3K$				
m	n	e	e_x	e_t
60	226	0.010533	0.000728	0.017352
120	901	0.002612	0.000176	0.004100
240	3601	0.000665	0.000039	0.000852
480	14401	0.000181	0.000005	0.000044

$V(t, \hat{s}) = e^{-r(T-t)} f(\hat{s})$

Errors at K when $\hat{s} = 3K$				
m	n	e	e_x	e_t
60	226	0.010518	0.000725	0.017526
120	901	0.002597	0.000173	0.004277
240	3601	0.000650	0.000036	0.001031
480	14401	0.000166	0.000003	0.000225

$V_{ss}(t, \hat{s}) = 0$

Convection Dominated PDEs

Type of PDE with large drift μ and small diffusion $\sigma^2/2$ in

$$U_t + \mu(t,x)U_x + \frac{1}{2}\sigma^2(t,x)U_{xx} - rU = 0$$

are called *convection-dominated*. They arise frequently! Some examples are:

- In Black Scholes, where σ is relatively small compare to r;
- Asian, Parisian or many other path dependent PDEs;
- Any other model where stochastic state process is missing $dW(t)$ term;

Example of FDM for hyperbolic equation

Consider first-order linear equation

$$u_t + u_y = 0, \quad y \in \mathbb{R}$$

with boundary condition $f(y)$. This equation has a closed form solution

$$u(t, y) = f\left(y + T - t\right).$$

The solution is simply a translation of $f(y)$. First we answer the following:

- What is the underlying SDE for this equation?
- By Feynman-Kac, this function equal to the expectation, what is it?

Now, consider an explicit finite difference scheme with central difference in space and see there is no way to make this to be a stable scheme with grid sizes!

Crank-Nicolson for Black-Sholes PDE

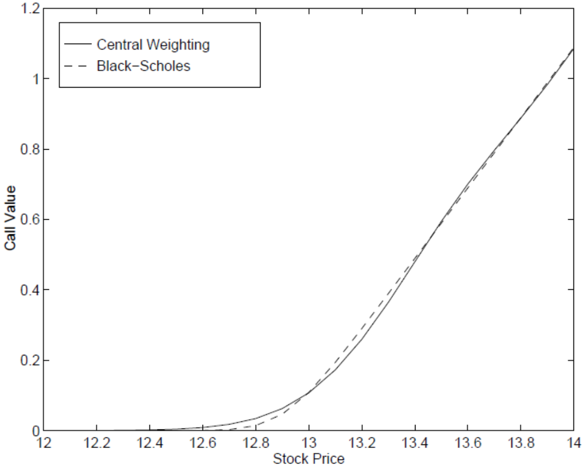
Consider pricing a European option in the Black-Scholes framework:

$$V_t + rsV_s + \frac{1}{2}\sigma^2 s^2 V_{ss} - rV = 0,$$

and terminal condition $V(T,s) = (s-K)^+$. Note that if rs is large compared to $\frac{\sigma^2 s^2}{2}$, the equation is said to be *convection dominated*. When discretized using central difference in space for the first derivative in s, regardless of how it is discretized in time, instability might arise if the payoff is not smooth.

Example in Black Scholes PDE

Crank-Nicolson and central differences discretization for a call with $T = 1, r = 0.15, \sigma = 0.01$ and $K = 13$.



Greeks

However, the greeks have more pronounced issues!

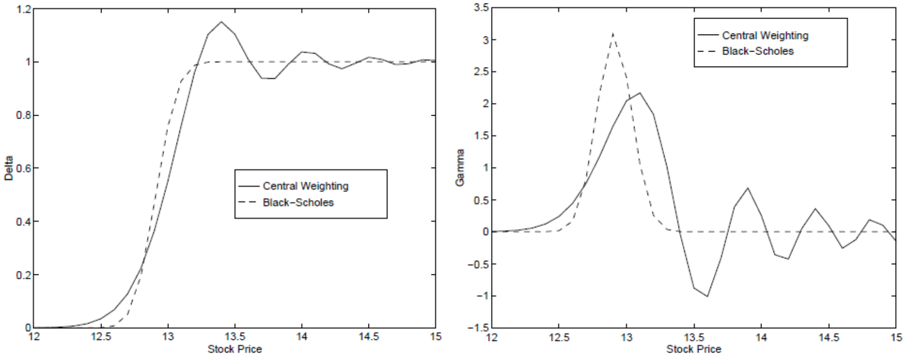


Figure: K=15, r=0.15, σ=0.01 and T=1. $\delta S = 0.1$, $\delta t = 0.01$

Barrier option revisited

Consider a double barrier option: knock-out option, i.e. a call option that expires worthless if one of the two barriers has been hit at a monitoring date. Define a set of monitoring dates $T_0, T_1, \dots, T_n$, lower barrier S_d and upper barrier S_u that are monitored only at times $T_i, i = 0, 1, \dots, n$. At every observation time the function value has a discontinuity at the barriers. The price $V(t, S)$ satisfies the Black-Scholes PDE with initial and boundary conditions:

$$V_t + rSV_S + \frac{1}{2}\sigma^2 S^2 V_{SS} - rV = 0,$$
$$V(T, S) = (S - K)^+ \mathbb{1}_{[S_d, S_u]}(S),$$
$$V(0, T) = 0, V(S_{max}) = 0.$$

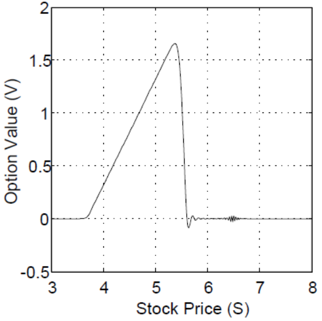
and the discrete monitoring introduces an updating of the solution at the monitoring times:

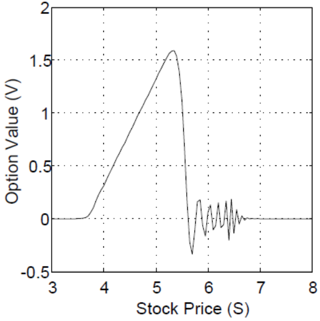
$$V(T_i, S) = V(T_i, S)\mathbb{1}_{[S_d, S_u]}(S)$$

where $\mathbb{1}_{[S_d, S_u]}(S) = 1$ if $S \in [S_d, S_u]$ and 0 otherwise.

Cont

We can apply Implicit scheme with central differences and solve numerically for $\delta S = 0.2$ and $\delta S = 0.5$ if $r > \sigma^2$





Example of FDM for hyperbolic equation

Consider the same hyperbolic

$$u_t + u_y = 0, \quad y \in \mathbb{R}$$

with boundary condition $f(y)$.

- Now we discretize using one-sided difference for y variable and get a better result, however, with a space step constraint!
- However, it is also important to consider forward or backward approximation for space derivative!

Stability condition

To avoid oscillations in

$$U_t + \mu(t,x)U_x + \frac{1}{2}\sigma^2(t,x)U_{xx} - rU = 0$$

one can show the following should hold(we omit the derivation of it)

$$\sigma_{k,j}^2 \geq |\mu_{k,j}|\delta x.$$

That condition comes from requiring all u_j and $l_j \geq 0$, where u_j, l_j were defined as entries in space discretization matrix $\mathbf{A}$ defined in Lecture 3.

Ways to handle convection terms

An alternative is to modify the first-order discrete operator D_x s.t. it points in the direction of the large absolute drift. Use the suitably oriented one-sided difference, rather then a central difference whenever the stability condition is violated. This is called *upstream differencing* or *upwinding*:

$$D_x^* V(t,x_j) = \begin{cases} \frac{V(t,x_{j+1})-V(t,x_{j-1})}{2\delta x}, \ |\mu(t,x_j)\delta x \leq \sigma(t,x_j)^2|, \\ \frac{V(t,x_j)-V(t,x_{j-1})}{\delta x}, \ \mu(t,x_j)\delta x < -\sigma(t,x_j)^2, \\ \frac{V(t,x_{j+1})-V(t,x_j)}{\delta x}, \ \mu(t,x_j)\delta x > \sigma(t,x_j)^2. \end{cases}$$

Artificial diffusion

How upwinding actually works? Basically, it introduces enough artificial diffusion into the PDE to satisfy

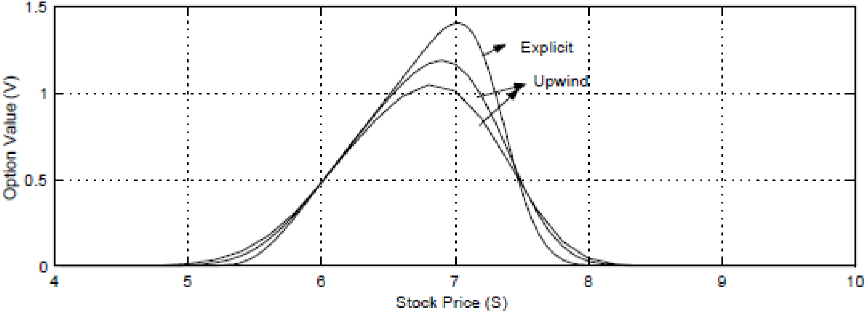
$$\sigma(t, x_j)^2 \geq |\mu(t, x_j)|\delta x.$$

In another way, it can be see from

$$\begin{aligned} \frac{U(x_{j+1}) - U(x_j)}{\delta x} &= \frac{U(x_{j+1}) - U(x_{j-1})}{2\delta x} + \frac{U(x_{j+1}) - 2U(x_j) + U(x_{j+1})}{2\delta x} \\ &\approx \frac{U(x_{j+1}) - U(x_{j-1})}{2\delta x} + \frac{1}{2}\delta x U_{xx} \end{aligned}$$

Upwind results

Below is the application of the upwind method with $\delta S = 0.2, 0.1$ to the discrete barrier call option:



Path dependent options

- FDM are normally limited to Markovian problems and payoffs are simple deterministic functions of the underlying state variables.
- A number of options have payoffs that depend on the entire path $\{S(t), t \in [0, T]\}$

Consider pricing an Asian option, where the payoff depends on the running sum of $S(t)$. Clearly the value function can't be represented as a function of t and $S(t)$, thus an additional state variable is defined

$$Y(t) = \int_0^t S(u)du.$$

Now the function depends on $t, S(t), Y(t)$ and the PDE becomes

$$V_t + rSV_S + \frac{1}{2}\sigma^2 S^2 V_{SS} + SV_Y - rV = 0,$$

However, there is no diffusion in Y dimension!

Similarity reduction

In some cases it is possible to make a transformation(in variables or probability measure) which reduces it to a one-dimensional PDE. This method called *similarity reduction*. One can see that the equation

$$U_t + U_{xx} = 0$$

can be reduced to

$$V_{zz} + \frac{1}{2} z V_z = 0$$

where $U(t, x) = V(z)$ and z is defined by

$$z = \frac{x}{\sqrt{T - t}}$$

Asian option example

Define a new variable

$$Z = \frac{1}{S} e^{-r(T-t)} \left(\frac{Y}{T} - K \right) + \frac{1}{rT}(1 - e^{-r(T-t)}).$$

Then we can seek a solution to the PDE

$$V_t + rSV_S + \frac{1}{2}\sigma^2 S^2 V_{SS} + SV_Y - rV = 0,$$

of the form

$$V(t, S, Y) = Sg(t, Z)$$

and g the satisfies a simpler PDE

$$g_t - \frac{1}{2}\sigma^2 \left(Z - \frac{1}{rT}(1 - e^{-r(T-t)}) \right)^2 g_{ZZ} = 0.$$

However, this is not very general and one has to use numerical methods to solve the PDE for most of the path dependent options.

Path-dependent pricing framework

Consider a general payoff

$$V(T) = f(S(T), Y(T)),$$

where Y is a path integral of the type

$$Y(t) = \int_0^t h(S(u))du,$$

for some deterministic function h. In particular, for Asian option

$$Y(t) = \int_0^t S(u)du$$

and $Y(0) = 0$ and the payoff depends on $Y(T)/T$.
In the classical BS framework,

$$dS(t) = rS(t)dt + \sigma S(t)d\widetilde{W}(t),$$

and

$$dY(t) = h(S(t))dt.$$

PDE in a path dependent case

the pricing function $V(t, s, y)$ that satisfies

$$V_t + rsV_s + \frac{1}{2}\sigma^2 V_{ss} + h(s)V_y - rV = 0,$$

subject to the terminal condition $V(T, s, y) = f(s, y)$. That PDE gives rise to a few complications:

- It involves *two* spacial variables and therefore use of a two-dimensional PDE solver;
- Contains no second derivative in the variable I, i.e. it is convection-dominated in y-direction;
- the term $h(s)$ multiplying V_y may be of different order of magnitude that the other coefficients;

Discrete Sampling

One normally samples the function h at a discrete set of dates. Thus we replace $Y(T)$ with

$$Y(T) = \sum_{i=1}^{k} h(S(T_i))(T_i - T_{i-1}),$$

where $T_0 < T_1 < \ldots < T_k$ are the discrete dates schedule, or monitoring dates. In a PDE setting, we incorporate a process through appropriate jump condition:

$$V(T_i-, s, y) = V(T_i+, s, y + h(s)(T_i - T_{i-1})).$$

and between those dates we solve

$$V_t + rsV_s + \frac{1}{2}\sigma^2 V_{ss} - rV = 0,$$

which has no y term.

Coupon-paying securities and Dividends

Many fixed income securities are coupon-bearing and involve transfer of a cash amount between the buyer and the seller. This can be incorporated into a finite difference grid through a jump condition.
Consider a security that pays its owner a single cash amount $p(T^*, x)$ at time $T^* < T$, then we apply the following condition

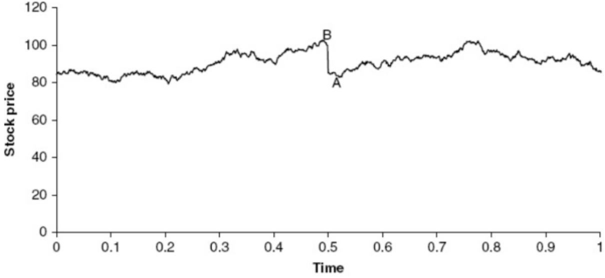
$$V(T^*-, x) = V(T^*+, x) + p(T^*, x).$$

which is performed at each time step where the cash is paid.

Cont

Another example is a derivative that doesn't pay coupons, but is written on a security(e.g. stock) that does. Thus, a constant dividend yield payment at time T^*, then it would be associated with a discontinuity in the state variable

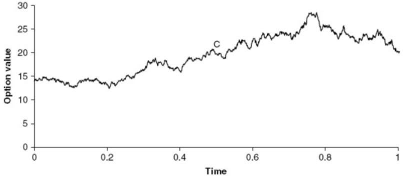
$$S(T^*+) = S(T^*-) - d * S(T^*-) = S(T^*-)(1-d).$$

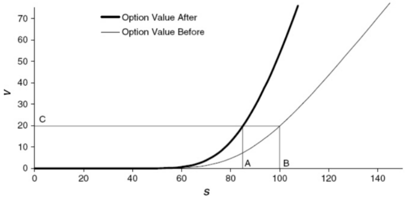


Continuity condition

However, by the arbitrage-free arguments, the price of a derivative must be continuous, i.e.

$$V(T^*-, S) = V(T^*+, S(1-d)).$$





Review of Linear Interpolation

If the function f is linear in an interval $[a, b]$ its value at $x \in [a, b]$ can be expressed in terms of its values at the endpoints by

$$f(x) = \lambda_1 f(a) + \lambda_2 f(b)$$

where

$$\lambda_1 = \frac{b-x}{b-a}, \quad \lambda_2 = \frac{x-a}{b-a}$$

Note, $\lambda_1 + \lambda_2 = 1$ and that both are non-negative. If f is not linear we can use the same formula to get an estimate of the value $f(x)$. The error of this approximation is bounded by $(b-a)^2 M/8$, where M is a bound for the magnitude of the second derivative of f in the interval (a, b). Note, if f is linear, there is no error(as second derivative is then zero).

Implementing the jump condition

Assume there is k s.t. $t^k = T^*$ and we incorporate the jump condition

$$V(T^*-, S) = V(T^*+, S(1-d)).$$

Suppose we calculated the numerical solution $\hat{V}$ by using a FDM down to t^k. Then, at the same grid level, these values to be adjusted according to the jump condition. At each grid point (t^k, S_j), the new value is given by the value of $\hat{V}$ at $S_j(1-d)$.
Assume that value falls between S_I and S_{I+1} for some value of I. Then define

$$\lambda_1 = \frac{S_{I+1} - S_j(1-d)}{S_{I+1} - S_I}, \quad \lambda_2 = 1 - \lambda_1$$

Then, by linear interpolation

$$V_j^k = \lambda_1 \hat{V}_I^k + \lambda_2 \hat{V}_{I+1}^k$$

which applied for all $j = 1, \ldots, m$. The resulting V_j^k is treated as a final condition to continue the solution to $k-1$ and so on down to $t = t^0$.

Securities with Early Exercise

Option holders make optimal decisions in the sense of maximizing expectations. This means

$$V(t,s) = \max_{\tau} \widetilde{\mathbb{E}}[e^{-r(T-t)} f(S(\tau))]$$

Functions τ are called *stopping times*. But this is a harder task compare to European option where τ is only T, the optimal strategy must be found as part of the option price computation. For example, for an American put, optimal exercise strategy is given as a curve $s^*(t)$ such that at time t the holder should exercise if

$$V(t,s) \le f(s^*(t))$$

This curve $s^*(t)$ is called an *exercise boundary*.

Implementation of Early Exercise

American and Bermudan securities allow for early exercise feature and finite difference methods can handle this feature very well. Assume exercise values are determined by a deterministic function $h(t, x)$.

- **Bermudan option:** exercise opportunities are restricted to the finite set $\{T_k\}_{k=1}^{K}$. Thus, it suffices to implement a simple just condition for each $k = 1, \ldots, K$:

$$V(T_k-, x) = \max(V(T_k+, x), h(T_k, x))$$

- **American option:** exercise opportunities are present at each time step. Thus, apply condition at each step. But me mindful that it is likely to decrease the order of convergence!

Recommended reading

- [TR] Relevant parts of Chapter 4.

Numerical Methods 46.950

Lecture 6

October 4, 2016

Overview of the Lecture:

- Non-equidistant grid
- ADI and predictor corrector methods overview
- Operator splitting and it is application to path dependent options
- Review

Mesh Refinements

- There are maybe situations where there are a few exercise dates that are of a particular importance and we would like to have a finer mesh around those dates.
- Before we would choose a confidence interval for $X(T)$ and divide the interval in a number of intervals; Can lead to bad resolution at earlier dates!
- Consider an example of two dates T_1 and T_2, s.t. the first one is close to the valuation date 0 and the second is close to the expiration T. Thus, we might have few points between 0 and T_1. Also, the grid size based on the distribution of $X(T)$ could be very large for the possible confidence interval for $X(T_1)$.

Cont

There are a couple of ways to refine x discretization around points of interest:

- use non-rectangular $(t, x) -$ *domain* and link together different equidistant rectangular meshes; It has requires interpolation in between regions and therefore there are disadvantages:
  - slows down the solver;
  - not clear how it might affect stability and convergence;
  - not very general
- Use *non-equidistant* discretization, i.e. concentrate more points around the initial point $t = 0, x = x(0)$

Non-equidistant discretization

There are a few reasons why we would like to have non-unifrom grid:

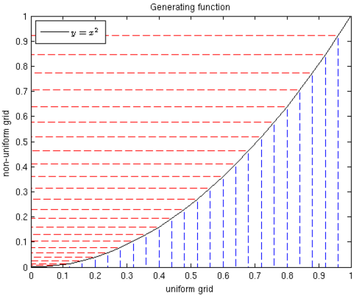
- We like to align the grid to particular dates: coupon, dividend is paid, etc.
- Concentrate computational effort on domains of particular importance to the solution of the PDE;

Cont

The non-uniform grid can be created to be more fine(coarse, dense) in a particular region, i.e. one can choose a grid generating function g, that is monotonically increasing and smooth. Then choose the domain $0 = x_0 < x_1 < \ldots < x_n = 1$ that is uniformly discretized. Then, we can define a non-uniform grid as:

$$y_i = cg(x_i) + d$$

for all $i = 1, 2, \ldots, n$. For example, if $g(x_i) = x_i^2$ we have



Cont

So, how do we approximate partial derivatives on such a grid?
Define $\delta x_j^+ = x_{j+1} - x_j$ and $\delta x_j^- = x_j - x_{j-1}$, and set

$$D_x^+ V(t, x_j) = \frac{V(t, x_{j+1}) - V(t, x_j)}{\delta x_j^+}$$

$$D_x^- V(t, x_j) = \frac{V(t, x_j) - V(t, x_{j-1})}{\delta x_j^-}.$$

Cont

Then by Taylor expansion we get,

$$D_x^+ V(t,x_j) = V_x(t,x_j) + \frac{1}{2} V_{xx}(t,x_j)\delta x_j^+ + \frac{1}{6} V_{xxx}(t,x_j)(\delta x_j^+)^2 + O\left((\delta x_j^+)^3\right),$$

$$D_x^- V(t,x_j) = V_x(t,x_j) - \frac{1}{2} V_{xx}(t,x_j)\delta x_j^- - \frac{1}{6} V_{xxx}(t,x_j)(\delta x_j^-)^2 + O\left((\delta x_j^-)^3\right).$$

Approximation of the first derivative

To achieve the maximum accuracy on the first-order derivative approximation, select a weighted combination of $D_x^+ V(t, x_j)$ and $D_x^+ V(t, x_j)$:

$$D_x V(t, x_j) = \frac{\delta x_j^-}{\delta x_j^- + \delta x_j^+} D_x^+ V(t, x_j) + \frac{\delta x_j^+}{\delta x_j^- + \delta x_j^+} D_x^- V(t, x_j)$$

which has a truncation error

$$O\left(\frac{(\delta x_j^+)^2 \delta x_j^- + (\delta x_j^-)^2 \delta x_j^+}{\delta x_j^+ + \delta x_j^-}\right)$$

which is second-order accurate.

Approximation of the second derivative

And we can also approximate the second derivative as:

$$D_{xx}V(t,x_j) = \frac{D_x^+ V(t,x_j) - D_x^- V(t,x_j)}{\frac{1}{2}\left(\delta x_j^+ + \delta x_j^-\right)}$$

which has a truncation error

$$O\left(\frac{(\delta x_j^+)^2 - (\delta x_j^-)^2}{\delta x_j^+ + \delta x_j^-} + \frac{(\delta x_j^+)^3 - (\delta x_j^-)^3}{\delta x_j^+ + \delta x_j^-}\right)$$

which is first-order accurate unless $\delta x_j^- = \delta x_j^+$! Despite this, the global error typically remain second order accurate in spatial step(there exists a proof for that result).

Two-dimensional PDE without mixed derivatives

Consider two state processes $X(t)$ and $Y(t)$ that are independent. Then the value function is to satisfy the PDE of the form:

$$V_t + (\mathcal{L}_X + \mathcal{L}_Y)V = 0,$$

where

$$\mathcal{L}_X = \mu_x \frac{\partial}{\partial x} + \frac{1}{2}\gamma_x^2 \frac{\partial^2}{\partial x^2} - \frac{1}{2}r$$

and

$$\mathcal{L}_Y = \mu_y \frac{\partial}{\partial y} + \frac{1}{2}\gamma_y^2 \frac{\partial^2}{\partial y^2} - \frac{1}{2}r$$

ADI

Alternating Direction Implicit(ADI) is an example of *operator splitting* methods, where the application of two operators is split into two *sequential* operator applications.

$$\left(1 - \frac{1}{2}\delta t \left(\widehat{\mathcal{L}}_X + \widehat{\mathcal{L}}_Y \right) \right) \approx \left(1 - \frac{1}{2}\delta t \widehat{\mathcal{L}}_X\right)\left(1 - \frac{1}{2}\delta t \widehat{\mathcal{L}}_Y\right)$$

and

$$\left(1 + \frac{1}{2}\delta t \left(\widehat{\mathcal{L}}_X + \widehat{\mathcal{L}}_Y \right) \right) \approx \left(1 + \frac{1}{2}\delta t \widehat{\mathcal{L}}_X\right)\left(1 + \frac{1}{2}\delta t \widehat{\mathcal{L}}_Y\right)$$

The operator on RHS has the same order truncation error as LHS.

Cont

But the equation

$$\left(1-\frac{1}{2}\delta t\widehat{\mathcal{L}}_X\right)\left(1-\frac{1}{2}\delta t\widehat{\mathcal{L}}_Y\right)\widehat{V}_{j_1,j_2}(t_i)$$
$$=\left(1+\frac{1}{2}\delta t\widehat{\mathcal{L}}_X\right)\left(1+\frac{1}{2}\delta t\widehat{\mathcal{L}}_Y\right)\widehat{V}_{j_1,j_2}(t_{i+1})$$

can be decomposed into the system

$$\left(1-\frac{1}{2}\delta t\widehat{\mathcal{L}}_X\right)U_{j_1,j_2}=\left(1+\frac{1}{2}\delta t\widehat{\mathcal{L}}_Y\right)\widehat{V}_{j_1,j_2}(t_{i+1}),$$
$$\left(1-\frac{1}{2}\delta t\widehat{\mathcal{L}}_Y\right)\widehat{V}_{j_1,j_2}(t_i)=\left(1+\frac{1}{2}\delta t\widehat{\mathcal{L}}_X\right)U_{j_1,j_2},$$

with intermediate value U_{j_1,j_2}. The main advantage there is only one operator on LHS of each of the equations!

Graphical illustration of ADI

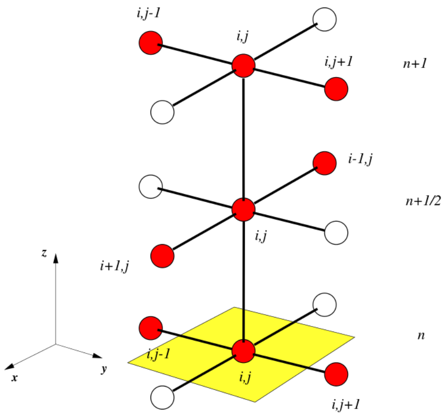


Figure: ADI

PDE with Mixed Derivatives

Consider now the casae where the 2-dimensional PDE has a mixed partial derivative,

$$V_t + (\mathcal{L}_X + \mathcal{L}_Y + \mathcal{L}_{XY})V = 0,$$

with

$$\mathcal{L}_{XY} = s_{xy}\frac{\partial^2}{\partial x \partial y} = \rho\gamma_x\gamma_y\frac{\partial^2}{\partial x \partial y}.$$

where we define ρ to be an instantaneous correlation between $X(t)$ and $Y(t)$. There are two main approaches to add the mixed derivative term:

- Orthogonalize the PDE;
- Predictor-Corrector scheme;

We consider only the first method next.

Orthogolization of the PDE

Define a new process

$$Z(t) = -\rho \frac{\gamma_y}{\gamma_x} X(t) + Y(t),$$

then $v(t, x, z) = V(t, x, y)$ satisfies

$$\begin{aligned} v_t + \mu_x v_x + \mu_z v_z + \frac{1}{2}\gamma_x^2 v_{xx} \\ \frac{1}{2}\left(1-\rho^2\right)\gamma_y^2 v_{zz} - rv = 0 \end{aligned}$$

where μ_z is the new drift coefficient in the space of processes $X(t)$ and $Z(t)$. Let's derive them next.

Cont

In fact, assume we have two independent Brownian motions $W(t)$ and $B(t)$ and dynamics

$$
\begin{aligned}
dX(t) &= \mu_x dt + \gamma_x dW(t) \\
dY(t) &= \mu_y dt + \gamma_y \left(\rho dW(t) + \sqrt{1-\rho^2} dB(t) \right)
\end{aligned}
$$

and it can be seen $dX(t)dY(t) = \rho\gamma_x\gamma_y dt$. Applying Ito formula to the process $Z(t)$:

$$
\begin{aligned}
dZ(t) =& -\rho\frac{\gamma_y}{\gamma_x}\mu_x dt - \rho\frac{\gamma_y}{\gamma_x}\gamma_x dW(t) \\
& + \mu_y dt + \gamma_y \left(\rho dW(t) + \sqrt{1-\rho^2} dB(t) \right) \\
=& \left(-\rho\frac{\gamma_y}{\gamma_x}\mu_x + \mu_y \right) dt + \gamma_y \sqrt{1-\rho^2} dB(t).
\end{aligned}
$$

Cont

Thus we can write

$$\mu_z = -\rho \frac{\gamma_y}{\gamma_x} \mu_x + \mu_y$$

and

$$\begin{aligned} dX(t) =& \mu_x dt + \gamma_x dW(t), \\ dZ(t) =& \mu_z dt + \gamma_y \sqrt{1-\rho^2} dB(t). \end{aligned}$$

Processes $X(t)$ and $Z(t)$ define Markov SDEs with Brownian motions in those being *independent*, and therefore $v(t, x, z)$ satisfies the PDE without a mixed derivative!

Splitting methods

Operator splitting means that the spacial operator appearing in the PDE is split into a sum of different sub operators that have simpler forms. Then, we can use an appropriate numerical tool to solve sub-problem individually. Consider a Cauchy problem

$$\begin{aligned} u_t + \mathcal{L}_x u + \mathcal{L}_y u &= 0, \\ u(T, x, y) &= f(x, y) \end{aligned}$$

we can apply successively the so-called “Lie splitting” defined by

$$\begin{aligned} u^{[1/2]} &= u^{[n]} + \widehat{\mathcal{L}}_x u^{[n]}, \\ u^{[n-1]} &= u^{[1/2]} + \widehat{\mathcal{L}}_y u^{[1/2]} \end{aligned}$$

where in the first step y is a parameter and in the second x is a parameter.

Application of splitting to path dependent options

PDEs for path dependent options often have a form:

$$u_t + \mathcal{L}_S u + h(s, y) u_y = 0$$

and therefore can be split in the system of one-dimensional problems:

$$u_t + \mathcal{L}_S u = 0$$

and

$$u_t + h(s, y) u_y = 0.$$

Exam Review

- Feynman-Kac application
- Barrier options
- Boundary conditions and degenerate boundaries
- Discretization
- Crank-Nicolson oscillations
- Path dependent options and similarity transformation
- Incorporation of early exercise and decision features