Mathematical & Computational Finance II
Lecture Notes

PDE Methods

November 10 2015
Last update: December 4, 2017

1 PDE Type Methods

PDE type methods are particularly useful for pricing American & barrier type options.

1.1 Black-Scholes PDE & Finite Difference Method

The finite difference method is one option to find numerical solutions to option pricing problems (as opposed to other numerical methods to solve PDEs).

Recall that if we let $V(S,t)$ be the value at time t of a European derivative security on an underlying S risky asset with payoff $\Lambda(S)$ at expiry T. We showed that the function $V(S,t)$ satisfies

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV \quad \text{over the interval } (0,\infty) \times [0,T)$$

with terminal condition $V(S,T) = \Lambda(S)$. We need additional boundary conditions for a complete specification, and these conditions depend on the particular form $\Lambda(S)$ takes (i.e. put/call options). For a call option we require, for $K > 0$,

$$\lim_{S\to 0} V(S,t) = 0$$

$$\lim_{S\to\infty} V(S,t) \approx S - Ke^{-r(T-t)}$$

and for a put option we require

$$\lim_{S\to 0} V(S,t) = Ke^{-r(T-t)}$$

$$\lim_{S\to\infty} V(S,t) = 0$$

1.1.1 Derivation of the Black-Scholes PDE

We have derived the Black-Scholes PDE using martingale-type methods, applying Itô's formula, and manipulating the result into known martingales. Let's derive it again another

way.

The no-arbitrage argument implies that any risk free asset must have the same return as the "bank account" S_0. We create a portfolio composed of 1 unit of V and some amount $-\Delta$ of the risky asset S (i.e. short Δ units of S). Then, the value of the portfolio is

$$\Pi(t) = V(S,t) - \Delta S(t)$$

Applying Itô's rule to $V(S,t)$ we get, for some volatility function $\sigma^2(S,t)$,

$$V(S,t) = V(0,0) + \int_0^t V_t(S,u)\,du + \int_0^t V_S(S,u)\,dS_u + \frac{1}{2}\int_0^t V_{SS}(S,u)\,d\langle S\rangle_u$$

$$= \int_0^t V_t(S,u)\,du + \int_0^t V_S(S,u)\,dS_u + \frac{1}{2}\int_0^t V_{SS}(S,u)\sigma^2(S,u)\,du$$

$$= \int_0^t \left(V_t(S,u) + \frac{1}{2}\sigma^2(S,u)V_{SS}(S,t)\right)du + \int_0^t V_S(S,u)\,dS_u$$

$$\implies dV(S,t) = \left(V_t(S,t) + \frac{1}{2}\sigma^2(S,u)V_{SS}(S,t)\right)dt + V_S(S,t)\,dS_t$$

$$\equiv \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S,t)\frac{\partial^2 V}{\partial S^2}\right)dt + \frac{\partial V}{\partial S}\,dS_t$$

so

$$d\Pi(t) = dV(S,t) - \Delta\,dS(t)$$

$$= \left(\frac{\partial V}{\partial S} - \Delta\right)dS(t) + \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S,t)\frac{\partial^2 V}{\partial S^2}\right)dt$$

If $\Delta = \frac{\partial V}{\partial S}$ then

$$\left(\frac{\partial V}{\partial S} - \Delta\right)dS(t) + \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S,t)\frac{\partial^2 V}{\partial S^2}\right)dt = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S,t)\frac{\partial^2 V}{\partial S^2}\right)dt$$

That is, only risk-free terms remain (i.e. only deterministic bounded variation dt terms remain). No arbitrage implies that[1]

$$d\Pi(t) = r\Pi(t)\,dt$$

$$\implies \frac{d\Pi(t)}{dt} = r\Pi(t)$$

and if $\sigma(S,t) = \sigma S$ then we have

$$d\Pi(t) = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right)dt$$

$$\implies \frac{d\Pi(t)}{dt} = \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}$$

$$\implies r\Pi = \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}$$

---

[1]I think this has to do with the fact that the change in portfolio value $d\Pi$ must be only determined by the risk free rate r

2

But $\Pi = V - \Delta S$ so $r\Pi = r(V - \Delta S) = rV - rS\frac{\partial V}{\partial S}$, hence

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

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

as desired. We complete the specification of the Black-Scholes PDE with boundary information on the domain. We have $S > 0$ and $0 \le t \le T$ and we classify the Black-Scholes PDE as a parabolic partial differential equation.

1.1.2 Transformation of the Black-Scholes PDE to the Heat Equation

The canonical example of a parabolic PDE is the diffusion/heat equation, which is typically defined for $t > 0, x \in \mathbb{R}$. That is, we have the heat equation

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2}, \quad k > 0$$

with some initial condition $u(x, 0) = u_0(x)$. To transform the Black-Scholes PDE to the heat equation we first perform the substitution

$$x = \log S \iff S = e^x$$

$$\tau = \frac{\sigma^2}{2}(T - t) \iff t = T - \frac{2\tau}{\sigma^2}$$

and let $V(S, t) = W(x, \tau) = W\left(\log S, \frac{\sigma^2}{2}(T - t)\right)$, then, heavily relying on the chain rule,

$$\frac{\partial V}{\partial t} = \frac{\partial W}{\partial \tau}\frac{\partial \tau}{\partial t} = -\frac{\sigma^2}{2}\frac{\partial W}{\partial \tau}$$

$$\frac{\partial V}{\partial S} = \frac{\partial W}{\partial x}\frac{\partial x}{\partial S} = \frac{1}{S}\frac{\partial W}{\partial x}$$

$$\frac{\partial^2 V}{\partial S^2} = \frac{\partial}{\partial S}\left(\frac{1}{S}\frac{\partial W}{\partial x}\right) = \frac{1}{S^2}\frac{\partial^2 W}{\partial x^2} - \frac{1}{S^2}\frac{\partial W}{\partial x}$$

Substituting these partial derivatives into our original PDE we get

$$\left[-\frac{\sigma^2}{2}\frac{\partial W}{\partial \tau}\right] + \frac{\sigma^2}{2}S^2\left[\frac{1}{S^2}\frac{\partial^2 W}{\partial x^2} - \frac{1}{S^2}\frac{\partial W}{\partial x}\right] + rS\left[\frac{1}{S}\frac{\partial W}{\partial x}\right] - r[W] = 0$$

$$\iff -\frac{\sigma^2}{2}\frac{\partial W}{\partial \tau} + \frac{\sigma^2}{2}\frac{\partial^2 W}{\partial x^2} - \frac{\sigma^2}{2}\frac{\partial W}{\partial x} + \frac{\partial W}{\partial x} - rW = 0$$

$$\iff -\frac{\partial W}{\partial \tau} + \frac{\partial^2 W}{\partial x^2} - \frac{\partial W}{\partial x} + \frac{2r}{\sigma^2}\frac{\partial W}{\partial x} - \frac{2r}{\sigma^2}W = 0$$

$$\iff -\frac{\partial W}{\partial \tau} + \frac{\partial^2 W}{\partial x^2} + \left(\frac{2r}{\sigma^2} - 1\right)\frac{\partial W}{\partial x} - \frac{2r}{\sigma^2}W = 0$$

For sanity let $\kappa := \frac{2r}{\sigma^2}$ so

$$\Longleftrightarrow \quad -\frac{\partial W}{\partial \tau} + \frac{\partial^2 W}{\partial x^2} + (\kappa - 1)\frac{\partial W}{\partial x} - \kappa W = 0$$

We see that in this form we are very close to the classical heat equation. Unfortunately for us $\kappa = \frac{2r}{\sigma^2}$ and so we cannot set this term to 0. To deal with these terms we set $W(x, \tau) = e^{\rho x + \xi \tau} u(x, \tau)$, for some to-be-determined constants ρ and ξ. Then, letting $\psi := e^{\rho x + \xi \tau}$ for brevity, we have

$$\frac{\partial W}{\partial \tau} = \xi \psi u + \psi \frac{\partial u}{\partial \tau}$$

$$\frac{\partial W}{\partial x} = \rho \psi u + \psi \frac{\partial u}{\partial x}$$

$$\frac{\partial^2 W}{\partial x^2} = \rho^2 \psi u + 2\rho \psi \frac{\partial u}{\partial x} + \psi \frac{\partial^2 u}{\partial x^2}$$

and so our PDE in W becomes the PDE in u

$$-\left[\xi \psi u + \psi \frac{\partial u}{\partial \tau}\right] + \left[\rho^2 \psi u + 2\rho \psi \frac{\partial u}{\partial x} + \psi \frac{\partial^2 u}{\partial x^2}\right] + (\kappa - 1)\left[\rho \psi u + \psi \frac{\partial u}{\partial x}\right] - \kappa \psi u = 0$$

$$\Longleftrightarrow \quad -\xi u - \frac{\partial u}{\partial \tau} + \rho^2 u + 2\rho \frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} + (\kappa - 1)\rho u + (\kappa - 1)\frac{\partial u}{\partial x} - \kappa u = 0$$

$$\Longleftrightarrow \quad -\frac{\partial u}{\partial \tau} + \frac{\partial^2 u}{\partial x^2} + (\kappa - 1 + 2\rho)\frac{\partial u}{\partial x} + (-\xi + \rho^2 + (\kappa - 1)\rho - \kappa)u = 0$$

We see that we are still very close to the canonical heat equation from the two leftmost terms. In order to rid ourselves of the pesky convection term u_x and heat source term u we must solve the following system for ρ and ξ,

$$0 = \kappa - 1 + 2\rho$$

$$0 = -\xi + \rho^2 + (\kappa - 1)\rho - \kappa$$

which has the solution

$$\rho = -\frac{1}{2}(\kappa - 1) = -\frac{1}{2}\left(\frac{2r}{\sigma^2} - 1\right)$$

$$\xi = -\frac{1}{4}(\kappa + 1)^2 = -\frac{1}{4}\left(\frac{2r}{\sigma^2} + 1\right)^2$$

and so, for these values of ρ and ξ, we finally have the PDE

$$-\frac{\partial u}{\partial \tau} + \frac{\partial^2 u}{\partial x^2} = 0 \quad \Longleftrightarrow \quad \frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}$$

which is precisely the heat equation in x and τ. We now take a moment to notice that our substitution in $\tau := T - \frac{2t}{\sigma^2}$ has led to a time reversal making our terminal condition into an initial condition

$$V(S, T) = \Lambda(S) \implies u(x, 0) = e^{\rho x}\Lambda(S) =: \lambda(x)$$

For a call option we have the boundary conditions

$$\lim_{S \to 0} V(S,t) = 0 \implies \lim_{x \to -\infty} u(x,\tau) = 0$$

$$\lim_{S \to \infty} V(S,t) \approx S - Ke^{-r(T-t)} \implies \lim_{x \to \infty} u(x,\tau) \approx e^{\rho x + \xi t}\left(e^x - Ke^{-\frac{2rt}{\sigma^2}}\right)$$

and for a put option

$$\lim_{S \to 0} V(S,t) = Ke^{-r(T-t)} \implies \lim_{x \to -\infty} u(x,\tau) = e^{\rho x + \xi t}Ke^{-r\frac{2\tau}{\sigma^2}}$$

$$\lim_{S \to \infty} V(S,t) = 0 \implies \lim_{x \to \infty} u(x,\tau) = 0$$

Then, for a general option we want to solve

$$u_t = u_{xx}, \quad x \in \mathbb{R}, t > 0$$
$$u(x,0) = \lambda(x)$$

with boundary conditions

$$u(x,t) \approx \alpha(x,t) \quad x \to -\infty, \ \forall_{t>0}$$
$$u(x,t) \approx \beta(x,t) \quad x \to +\infty, \ \forall_{t>0}$$

for some λ, α, β corresponding to the specification of a particular option. We approximate this numerically and implement the algorithm on a computer by discretizing in x and t. To approximate this we must truncate the x domain, leading to some error directly caused by the truncation. That is, we approximate

$$u_t = u_{xx}, \quad x \in (x_L, x_R), \ 0 < t \leq t_{max}$$
$$u(x,0) = \lambda(x), \quad x \in [x_L, x_R)$$

and

$$u(x_L, t) = \alpha(x_L, t) \quad 0 < t \leq t_{max}$$
$$u(x_R, t) = \beta(x_R, t) \quad 0 < t \leq t_{max}$$

We should note that we have reversed the time axis in the transformation of V to u, where $0 =$ "option expiry" and $t_{max} =$ "option initialization". We create a grid of points in (x,t) covering

$$[x_L, x_R] \times [0, t_{max}]$$

where the points on our grid $(x_j, t_n), j = 0, ..., J, n = 0, ..., N$ are

$$x_j = x_L + j\Delta x$$
$$\Delta x = \frac{x_R - x_L}{J}$$
$$t_n = n\Delta t$$
$$\Delta t = \frac{t_{max}}{N}$$

We approximate the differential terms via finite difference. We approximate $\frac{\partial u}{\partial t}$ at point (x_j, t_n) by

$$\partial_t u_j^n \approx \frac{U_j^{n+1} - U_j^n}{\Delta t} =: \delta_j^+ U_j^n$$

where δ_j^+ is our <u>forward finite difference operator</u>. The Taylor approximation gives the error of this as

$$\approx \frac{1}{2} \Delta t \frac{\partial^2 u}{\partial t^2}(x_j, t_n)$$

and we approximate $\frac{\partial^2 u}{\partial x^2}$ at (x_j, t_n) by, applying the finite difference method twice,

$$\partial_x^2 u_j^n \approx \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{(\Delta x)^2} = \delta_x^- \delta_x^+ U_j^n$$

In general, we approximate the second derivative at a point x by

$$\frac{f'\left(x + \frac{1}{2}h\right) - f'\left(x - \frac{1}{2}h\right)}{\Delta h} \approx \frac{1}{\Delta h} \frac{f\left((x + \frac{1}{2}h) + \frac{1}{2}h\right) - f\left((x + \frac{1}{2}h) - \frac{1}{2}h\right)}{\Delta h} -$$
$$\frac{1}{\Delta h} \frac{f\left((x - \frac{1}{2}h) + \frac{1}{2}h\right) - f\left((x - \frac{1}{2}h) - \frac{1}{2}h\right)}{\Delta h}$$
$$\approx \frac{f(x+h) - f(x)}{(\Delta h)^2} - \frac{f(x) - f(x-h)}{(\Delta h)^2}$$
$$\approx \frac{f(x+h) - 2f(x) + f(x-h)}{(\Delta h)^2}$$

Thus, to approximate $\frac{\partial^2 f}{\partial x^2}$ we end up with evaluating the function on both sides of the original point x.

1.2   Forward in Time Central in Space (FTCS) Algorithm

If we replace $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ by the finite difference approximations U_j^n at each node we get the balance equation

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{(\Delta x)^2}$$

Letting $\nu = \frac{\Delta t}{(\Delta x)^2}$ and rearranging a little bit yields

$$U_j^{n+1} = \nu U_{j+1}^n + (1 - 2\nu)U_j^n + \nu U_{j-1}^n$$

Note that we have managed to express the point (x_j, t_{n+1}) in terms of $(x_{j-1}, t_n), (x_j, t_n)$, and (x_{j+1}, t_n). That is, we have managed to express a point <u>forward in time and central in space</u> with respect to the points used to compute it. At $n = 0, j = 0, j = J$ we impose the boundary conditions

$$U_j^0 = \lambda(x_j) \quad 0 \le j < J$$
$$U_0^n = \alpha(x_L, t_n) \quad 0 < n \le N$$
$$U_J^N = \beta(x_R, t_n) \quad 0 < n \le N$$

This is the FTCS finite difference algorithm. We should note that there are other ways to specify the boundary conditions. We have just given some examples of "Dirichlet" boundary conditions for u (i.e. specified the boundary values). We may also provide "von Neumann" boundary conditions which instead specify the value of a derivative, for example

$$u_x(x_R, t_n) = 0 \implies u_J^n = u_{J-1}^n$$

or

$$u_{xx} = 0 \implies \text{the solution is linear}$$

1.2.1 Stability

An interesting question to think about is the stability of the solution. We want to discuss stability in order to say something about error estimates (i.e whether or not a particular solution will eventually explode and become unusable). The goal is to have the solution be bound by some initial data. From the triangle inequality we have

$$|U_j^{n+1}| \leq \nu|U_{j-1}^n| + |1 - 2\nu||U_j^n| + \nu|U_{j+1}^n|$$
$$\leq (2\nu + |1 - 2\nu|) \|u^n\|_\infty$$

where the max norm, or L infinity norm, of a vector $\vec{v}$ is defined as

$$\|\vec{v}\|_\infty = \max_{0 \leq j \leq J} |v_j|$$

Hence

$$\max_{0 \leq j \leq J} |U_j^{n+1}| \leq (2\nu + |1 - 2\nu|) \max_{0 \leq j \leq J} |U_j^n|$$

and if we have $(2\nu + |1 - 2\nu|) \leq 1)$ then we see that our error estimate is shrinking at each future point in t. Note that

$$2\nu + |1 - 2\nu| \leq 1 \quad \text{if } 0 < \nu \leq \frac{1}{2}$$
$$2\nu + |1 - 2\nu| > 0 \quad \text{else}$$

Hence, if $\nu \leq \frac{1}{2}$ then the solution at time step $n+1$ is bound by the data point before it. This ensures that the output won't explode. The stability condition $\nu \leq \frac{1}{2}$ is equivalent to

$$\frac{\Delta t}{(\Delta x)^2} \leq \frac{1}{2}$$
$$2\Delta t \leq (\Delta x)^2$$
$$\implies \frac{2T}{N} \leq \frac{(x_R - x_L)^2}{J^2}$$
$$\implies \frac{2TJ^2}{(x_R - x_L)^2} \leq N$$

Therefore, if we increase J by a factor of 2 we must increase N by a factor of $2 \cdot 4 = 8$ to satisfy stability, that is, we require an 8-fold increase in work for a doubling of J!.

1.2.2 Error Tracking

We can track the error in our solution by, letting the error at node (x_j, t_n) as E_j^n,

$$E_j^n = U(x_j, t_n) - U_j^n$$

where $u(x_j, t_n)$ is the true solution and U_j^n is our approximation. We can show that (via Taylor's theorem)[2] and balance equations that

$$E_j^{n+1} = \nu E_{j-1}^n + (1 - 2)\nu_j^n + \nu E_{j+1}^n + \Delta t T_j^n$$

where

$$T_j^n = \frac{\Delta t}{2} \frac{\partial^2}{\partial t^2} u(x_j, \tau_n) - \frac{(\Delta x)^2}{12} \frac{\partial^4}{\partial x^4} u(\xi_j, t_n)$$

for some constants $\xi_j \in (x_{j-1}, x_{j+1})$ and $\tau_n \in (t_n, t_{n+1})$. The idea is that if $\left|\frac{\partial^2 u}{\partial t^2}\right|$ and $\left|\frac{\partial^4 u}{\partial x^4}\right|$ are bounded in the region of interest and we satisfy the stability condition $\nu \leq \frac{1}{2}$, then we have

$$|T_j^n| \leq C(\Delta x)^2 \quad \text{for some constant } C$$

Given Dirichlet boundary conditions we know that the error along the boundaries will be zero (assuming we specified it without error), that is,

$$E_0^n = E_J^n = 0$$

Therefore, the error at a particular point in time is

$$\left\|E^{n+1}\right\|_\infty \leq \|E^n\|_\infty + C\Delta t(\Delta x)^2$$

where here $\|\cdot\|$ was defined as above so that

$$\max_{0 \leq n \leq N} \|E^n\|_\infty \leq CT(\Delta x)^2$$

This implies that reducing Δx will reduce the bound on the error. We can show that if $\nu = \frac{1}{2}$ we recover the binomial model. We say that the binomial model is "on the edge of stability". Recall the FTCS algorithm

$$U_j^{n+1} = \nu U_{j+1}^n + (1 - 2\nu)u_j^n + \nu U_{j-1}^n$$

and with $\nu = \frac{1}{2}$ we have

$$U_j^{n+1} = \frac{1}{2} U_{j+1}^n + \frac{1}{2} U_{j-1}^n$$

That is, a future[3] point in time depends on the adjacent two points above & below it from the previous point in time. If we plot the domain in (x, t)-space we get something that looks like a binomial tree with nodes branching in $\pm x$ as t progresses.

---

[2]This was done in Assignment 5: Question 3.

[3]We have a future point with respect to u, so a past point with respect to V