Finite Difference Methods for Financial Problems

Part 2: Accuracy and Stability

Kwant Skool
Copenhagen
March 2011

Jesper Andreasen
Danske Markets
kwant.daddy@danskebank.com

Outline

- Spatial accuracy.
- Accuracy in time.
- Stability.
- Upwinding.
- Sign flip flopping in Crank-Nicolson.
- Stability summary.
- Stability extensions.
- Transition.
- Conclusion.

Spatial Accuracy

- Taylor expansion yields

$$\begin{aligned} f(x+\Delta x)-f(x) &= f_x(x)\Delta x+\frac{1}{2}f_{xx}(x)\Delta x^2+\frac{1}{6}f_{xxx}(x)\Delta x^3+O(\Delta x^4) \\ f(x-\Delta x)-f(x) &= -f_x(x)\Delta x+\frac{1}{2}f_{xx}(x)\Delta x^2-\frac{1}{6}f_{xxx}(x)\Delta x^3+O(\Delta x^4) \end{aligned} \tag{1}$$

- Subtracting and adding these two equations yield respectively

$$\begin{aligned} f_x(x) &= \frac{1}{2\Delta x}[f(x+\Delta x)-f(x-\Delta x)]+O(\Delta x^2) \\ f_{xx}(x) &= \frac{1}{\Delta x^2}[f(x+\Delta x)-2f(x)+f(x-\Delta x)]+O(\Delta x^2) \end{aligned} \tag{2}$$

- From (2) we conclude that if

$$0=\frac{\partial V}{\partial t}+DV \tag{3}$$

- Then

$$0 = \frac{\partial V}{\partial t} + \bar{D}V + O(\Delta x^2) \tag{4}$$

- Conversely if V solves (3) (exactly) then

$$V = v + O(\Delta x^2) \tag{5}$$

- Conclusion: *all our schemes have* $O(\Delta x^2)$ *accuracy*.

Accuracy in Time

- The exact solution to

$$0 = \frac{\partial v}{\partial t} + \bar{D}v \tag{6}$$

can be written as

$$v(t) = e^{\Delta t \bar{D}} v(t + \Delta t) \tag{7}$$

- where we have introduced the matrix exponential

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} \in \mathbb{R}^{n \times n} \quad \mathit{for}\ A \in \mathbb{R}^{n \times n} \tag{8}$$

- The properties of the matrix exponential are

$$
\begin{aligned}
&(e^{A})^{k} = e^{kA} \quad ,k \in \mathbb{Z} \\
&\frac{\partial}{\partial t}[e^{tA}] = Ae^{tA} \quad ,t \in \mathbb{R} \\
&AB = BA \Rightarrow e^{A+B} = e^{A}e^{B}
\end{aligned}
\tag{9}
$$

- Exercise: Prove the properties in (9).

- For the explicit scheme we have

$$
v(t) = e^{\Delta t\bar{D}} v(t+\Delta t) = [I + \Delta t\bar{D}]v(t+\Delta t) + O(\Delta t^{2}) \tag{10}
$$

- For the implicit scheme we have

$$
v(t+\Delta t) = e^{-\Delta t\bar{D}} v(t) = [I - \Delta t\bar{D}]v(t) + O(\Delta t^{2}) \tag{11}
$$

- For the Crank-Nicolson scheme we note that

$$
\begin{aligned}
&e^{-\frac{1}{2}\Delta t\bar{D}}v(t)=v(t+\Delta t/2)=e^{\frac{1}{2}\Delta t\bar{D}}v(t+\Delta t)\\
&\Rightarrow[I-\frac{1}{2}\Delta t\bar{D}+\frac{1}{4}\Delta t^2\bar{D}^2+O(\Delta t^3)]v(t)=[I+\frac{1}{2}\Delta t\bar{D}+\frac{1}{4}\Delta t^2\bar{D}^2+O(\Delta t^3)]v(t+\Delta t)\\
&\Rightarrow[I-\frac{1}{2}\Delta t\bar{D}]v(t)=[I+\frac{1}{2}\Delta t\bar{D}]v(t+\Delta t)+\frac{1}{4}\Delta t^2\bar{D}^2(v(t+\Delta t)-v(t))+O(\Delta t^3)\\
&\Rightarrow[I-\frac{1}{2}\Delta t\bar{D}]v(t)=[I+\frac{1}{2}\Delta t\bar{D}]v(t+\Delta t)+O(\Delta t^3)
\end{aligned}
\tag{12}
$$

- As the number of time steps in the full solution is $O(1/\Delta t)$ we have now shown that

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

- Exercise: For $r\in[-0.2,0.2], t=1$, and the function $f(t)=e^{-rt}$ plot the errors of the approximations

$$
f_{\theta=0}(t)=1-rt,\ f_{\theta=1}(t)=(1+rt)^{-1},\ f_{\theta=1/2}(t)=\frac{1-rt/2}{1+rt/2}
\tag{13}
$$

Stability

- The accuracy considerations of the previous section are only valid locally.
- If the scheme is not *stable* then the global solution will not converge.
- The main criterion used in FD stability analysis is von Neumann analysis.
- Our theta FD equation is

$$[I-\theta\Delta t\bar{D}]v(t_h)=[I+(1-\theta)\Delta t\bar{D}]v(t_{h+1}) \tag{14}$$

- We consider solutions to (14) of the form

$$e(t_h,x)=g^{-h}e^{ikx} \quad ,i=\sqrt{-1} \quad ,k\in\mathbb{R} \quad ,g\in\mathbb{C} \tag{15}$$

- If $|g|\leq 1$ for all k then we say that the scheme is von Neumann stable.

- For the constant parameter case von Neumann stability is sufficient to conclude that the scheme converges to the right result.

- The reason for this is that for the constant parameter case all eigen solutions to (14) are of the form (15).

- For the general case with level dependent parameters, it is not mathematically proven to be sufficient. There are, however, to my knowledge, no numerical counter examples.

- So general consensus is that von Neumann stability is a sufficient condition for convergence.

- We have

$$
\begin{aligned}
&\bar{D}(e^{ikx}) \\
&= [-r + \mu\delta_x + \frac{1}{2}\sigma^2\delta_{xx}]e^{ikx} \\
&= e^{ikx}[-r + \frac{\mu}{2\Delta x}(e^{ik\Delta x} - e^{-ik\Delta x}) + \frac{1}{2}\frac{\sigma^2}{\Delta x^2}(e^{ik\Delta x} - 2 + e^{-ik\Delta x})] \\
&= -e^{ikx}[r - i\frac{\mu}{\Delta x}\sin(k\Delta x) + \frac{\sigma^2}{\Delta x^2}(1 - \cos(k\Delta x))] \\
&= -e^{ikx}[r - i\frac{\mu}{\Delta x}\sin(k\Delta x) + 2\frac{\sigma^2}{\Delta x^2}\sin(k\Delta x/2)^2]
\end{aligned}
\tag{16}
$$

- So inserting (15) in (14) yields

$$
g = \frac{[I + (1-\theta)\Delta t\bar{D}](e^{ikx})}{[I - \theta\Delta t\bar{D}](e^{ikx})} = \frac{1-(1-\theta)\Delta t[r - i(\mu/\Delta x)\sin(k\Delta x) + 2(\sigma/\Delta x)^2\sin(k\Delta x/2)^2]}{1+\theta\Delta t[r - i(\mu/\Delta x)\sin(k\Delta x) + 2(\sigma/\Delta x)^2\sin(k\Delta x/2)^2]} \tag{17}
$$

- From (17) we can deduce that for $r \geq 0$ and independent of $\Delta t, \Delta x$:

$$
\frac{1}{2} \leq \theta \leq 1 \;\; \Rightarrow \;\; |g| \leq 1 \;\; \forall k \tag{18}
$$

- So the Crank-Nicolson and the fully implicit methods are von-Neumann stable.

- For $r > -1/\theta\Delta t$ and independent of $\Delta t, \Delta x$ we have

$$\frac{1}{2} \le \theta \le 1 \;\Rightarrow\; |g| \le 1 + O(\Delta t) \quad \forall k \tag{19}$$

- …which is also sufficient for stability.

- Complex components in the growth factor relate to the drift.

- If there is no drift then the growth factor is real: $\mu = 0 \Rightarrow g \in \mathbb{R}$

- Further, for zero drift we have

$$\begin{aligned} &\theta = 1, \mu = 0 \Rightarrow 0 \le g \le 1 \\ &\theta = 1/2, \mu = 0 \Rightarrow -1 \le g \le 1 \end{aligned} \tag{20}$$

- For zero volatility we have

$$g = \frac{1 + i(1-\theta)\Delta t(\mu/\Delta x)\sin(k\Delta x)}{1 - i\theta\Delta t(\mu/\Delta x)\sin(k\Delta x)}$$

- For this case the growth factor has no (real) damping and oscillations will tend to persist through the solution.

- This indicates that an FD problem solution potentially has two sources of oscillations:

  - CN diffusion induced sign flipping of the growth factor.

  - General drift induced complex ringing in the growth factor.

- The latter is more problematic, the higher the drift is relative to the diffusion.

- The CN sign flipping can be cured by taking a sufficient number of time steps (we will compute how many).

- The explicit method, $\theta = 0$, is von Neumann stable if

$$(1 - r\Delta t - \frac{\sigma^2 \Delta t}{\Delta x^2})^2 + (\frac{\mu \Delta t}{\Delta x})^2 \leq 1 \qquad (21)$$

- This creates very strict restrictions on the combination of time and spatial discretisation and effectively means that the number of time steps has to be an order higher than the number of spatial steps. If for example $r = \mu = 0$, then the following condition needs to be satisfied

$$\Delta t \leq \frac{2}{\sigma^2} \Delta x^2 \qquad (22)$$

- The conditions on the time and spatial steps in the explicit case ensure that the entries in the transition matrix $I - \Delta t \bar{D}$ are positive

$$I-\Delta t\bar{D}=I-\Delta t\cdot\begin{bmatrix} -r_0-\frac{\mu_0}{\Delta x} & \frac{\mu_0}{\Delta x} & & & & & \\ -\frac{\mu_1}{2\Delta x}+\frac{\sigma_1^2}{2\Delta x^2} & -r_1-\frac{\sigma_1^2}{\Delta x^2} & \frac{\mu_1}{2\Delta x}+\frac{\sigma_1^2}{2\Delta x^2} & & & & \\ & -\frac{\mu_2}{2\Delta x}+\frac{\sigma_2^2}{2\Delta x^2} & -r_2-\frac{\sigma_2^2}{\Delta x^2} & \frac{\mu_2}{2\Delta x}+\frac{\sigma_2^2}{2\Delta x^2} & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -\frac{\mu_{n-2}}{2\Delta x}+\frac{\sigma_{n-2}^2}{2\Delta x^2} & -r_{n-2}-\frac{\sigma_{n-2}^2}{\Delta x^2} & \frac{\mu_{n-2}}{2\Delta x}+\frac{\sigma_{n-2}^2}{2\Delta x^2} & \\ & & & & -\frac{\mu_{n-1}}{2\Delta x}+\frac{\sigma_{n-1}^2}{2\Delta x^2} & -r_{n-1}-\frac{\sigma_{n-1}^2}{\Delta x^2} & \frac{\mu_{n-1}}{2\Delta x}+\frac{\sigma_{n-1}^2}{2\Delta x^2} \\ & & & & & -\frac{\mu_n}{\Delta x} & -r_n+\frac{\mu_n}{\Delta x} \end{bmatrix}$$

(23)

Upwinding

- One remedy for high drift terms is upwinding.

- The idea is to replace the central difference operator by the signed approximations

$$
\begin{aligned}
&\frac{\partial}{\partial x} f(x) = 1_{\mu>0}\,\delta_x^{+} f(x) + 1_{\mu<0}\,\delta_x^{-} f(x) + O(\Delta x) \\
&\delta_x^{+} f(x) = \frac{1}{\Delta x}\big(f(x+\Delta x) - f(x)\big) \\
&\delta_x^{-} f(x) = \frac{1}{\Delta x}\big(f(x) - f(x-\Delta x)\big)
\end{aligned}
\tag{24}
$$

- Performing the von Neumann analysis for the $r=\sigma=0$ case, yields

$$
g = \frac{1-(1-\theta)(\mu\Delta t/\Delta x)[i\sin(k\Delta x)+(1-\cos(k\Delta x))]}{1+\theta(\mu\Delta t/\Delta x)[i\sin(k\Delta x)+(1-\cos(k\Delta x))]} \tag{25}
$$

- Now the drift term becomes self-dampening which will dampen out oscillations (high order modes).

- In the growth factor sense it adds diffusion to the problem.

- However, this comes at a cost: loss of accuracy to $O(\Delta x)$ rather than $O(\Delta x^2)$.

- This cost may be too high a price to pay. If the payoff is smooth we might instead use a higher order scheme (five points instead of three) in combination with very few points.

- For five point schemes a general band diagonal solver is needed. Here NRC::bandec() and NRC::banbks() can be used.

- This is the direction we chose when implementing the Cheyette model.

- A better alternative is probably to use winding in combination with a five point scheme, so

$$
\begin{aligned}
&\frac{\partial}{\partial x} f(x) = 1_{\mu>0}\delta_x^+ f(x) + 1_{\mu<0}\delta_x^- f(x) + O(\Delta x^2) \\
&\delta_x^+ f(x) = \frac{1}{2\Delta x}(-f(x+2\Delta x) + 4f(x+\Delta x) - 3f(x)) \\
&\delta_x^- f(x) = \frac{1}{2\Delta x}(3f(x) - 4f(x-\Delta x) + f(x-2\Delta x))
\end{aligned}
\tag{26}
$$

- …but I have actually never tried it.

- There are reports (Ratcliffe) that high order winding schemes work well for the Cheyette model.

Crank-Nicolson Sign Flip Flopping

- Consider again the growth factor for CN in the zero drift and zero interest rate case

$$g = \frac{1 - \Delta t[(\sigma / \Delta x)^2 \sin(k\Delta x / 2)^2]}{1 + \Delta t[(\sigma / \Delta x)^2 \sin(k\Delta x / 2)^2]} \tag{27}$$

- Assuming $\sin(k\Delta x / 2) \approx 1/2$, we have for large Δt:

$$g \approx -1 + \frac{\Delta x^2 / 4\sigma^2}{\Delta t} \tag{28}$$

- After m time steps we have

$$g^m \approx (-1 + \frac{\Delta x^2 / 4\sigma^2}{\Delta t})^m = (-1)^m (1 - \frac{\Delta x^2 / 4\sigma^2}{\Delta t})^m \approx (-1)^m e^{-m(\Delta x^2 / 4\sigma^2)/\Delta t} \tag{29}$$

- So the sign flipping dies out at an exponential rate in the number of time steps.

- In fact, the higher frequency it is, the quicker it will die out.

- Example: So if the $\Delta x \approx 10\sigma\sqrt{T}/n$ and $\Delta t \approx T/m$, then we get

$$\begin{aligned} g^m &\approx (-1)^m e^{-m(\Delta x^2/4\sigma^2)/\Delta t} \\ &= (-1)^m e^{-m((10\sigma\sqrt{T}/n)^2/4\sigma^2)/(T/m)} \\ &= (-1)^m e^{-25\cdot(m/n)^2} \end{aligned} \tag{30}$$

- This suggest that with $m/n \approx 1/2$ we should be fine and with $m/n \approx 1$ there are virtually no oscillations left.

Stability Generalisations – NEW!

- Assume that $\bar{D}$ has *distinct* complex eigenvalues $\lambda_1,\ldots,\lambda_n \in \mathbb{C}$.

- If so, the eigenvectors of $\bar{D}$ have full rank and span $\mathbb{C}^n$. Hence, any vector in $\mathbb{C}^n$ can be written as a linear combination of eigenvectors for $\bar{D}$.

- In that case it makes sense to consider application of our scheme to functions of the form

$$v(t_h) = g^{-h} e \quad ,g \in \mathbb{C} \tag{31}$$

  where e is an eigenvector for $\bar{D}$:

$$\bar{D}e = \lambda e \tag{32}$$

- Inserting (31) into the theta scheme yields

$$
\begin{aligned}
& g[I-\theta\Delta t\bar{D}]e=[I+(1-\theta)\Delta t\bar{D}]e \\
& \Downarrow \\
& g[1-\theta\Delta t\lambda]e=[1+(1-\theta)\Delta t\lambda]e \\
& \Downarrow \\
& g=\frac{1+(1-\theta)\Delta t\lambda}{1-\theta\Delta t\lambda}
\end{aligned}
\tag{33}
$$

- We conclude that $|g|\leq 1$ for $\theta\in[1/2,1]$, if all eigenvalues λ satisfy

$$
\mathrm{Re}\lambda\leq 0 \tag{34}
$$

- It has been proven [by Huge (2010) and others] that if

$$
|\mu|\Delta x\leq\sigma^2 \quad or \quad upwind \Rightarrow \quad \lambda_i\in]-\infty,0] \quad ,\lambda_i\neq\lambda_j \;\forall i\neq j \quad \Rightarrow|g|\leq 1 \quad ,\theta\in[1/2,1] \tag{35}
$$

- …for general $\mu(x),\sigma(x)$ and non-uniform discretisation Δx.

- We suspect that this line of research may provide further generalisations.

Transition – NEW!

- The theta scheme

$$v(t_h) = [I - \theta\Delta t\bar{D}]^{-1}[I + (1-\theta)\Delta t\bar{D}]v(t_{h+1}) \tag{36}$$

- So the entries of the matrix

$$A = [I - \theta\Delta t\bar{D}]^{-1}[I + (1-\theta)\Delta t\bar{D}] \tag{37}$$

are the transition weights for going from x_i to x_j over one time step.

- That is, we have

$$v(t_h, x_i) = \sum_j A_{ij} v(t_{h+1}, x_j) \tag{38}$$

- …and state prices (or conditional probabilities) are given by

$$A_{ij} = \Pr(x(t_{h+1}) = x_j \mid x(t_h) = x_i) \tag{39}$$

- Here, the state prices include discounting and are not necessarily positive.
- We can compute A to give us more insight into the finite difference scheme.
- In fact, there exist an $O(n)$ algorithm that produces a decomposition of A.
- It can be shown that for the fully implicit scheme, $\theta = 1$, we have the following

$$|\mu| \Delta x \leq \sigma^2 \quad or \quad upwind \Rightarrow \quad A^{-1} \geq 0 \tag{40}$$

- Hence, in this case the rows of the transition matrix are indeed probabilities.
- So the properties that provide stability are linked to the positivity of the transition matrix.
- $A^{-1} \geq 0$ guarantees that the solution method is stable, but further we have that positivity and convexity is preserved over each time step.

- The implicit method can thus be seen as a discrete arbitrage free financial model in itself.

- …and can thus be seen as a discrete model and not only a discretisation method for numerical solution of the PDE.

- Exercise: how can this be used for simulation and other things?

Conclusion

- Crank-Nicolson has $O(\Delta t^2 + \Delta x^2)$ accuracy.

- Explicit and implicit methods $O(\Delta t + \Delta x^2)$.

- Crank-Nicolson and the fully implicit methods are unconditionally stable.

- However, in both cases we need to be aware that there can be ringing in the solution that we need to keep under control by an appropriate number of times steps and/or by upwinding.

- Drift dominated problems are a much bigger practical problem than CN sign-flip-flop.

- Crank-Nicolson should be the weapon of choice for the pro.

- The explicit method is only conditionally stable, generally $\Delta t^{-1} = O(\Delta x^{-2})$ is needed to ensure stability. The explicit method is not for professional use.

- Stability results *can* be generalised to the case state dependent parameters and non-uniform grids.

- Stability results are linked to the positivity of transition matrices for the fully implicit case and the fully implicit scheme can be seen as an arbitrage free financial model in itself not necessarily only as a discretisation scheme for numerical solution of the continuous PDE.

- Questions?