A Probabilistic Interpretation of the θ-Method

Per Hörfelt*
Dept. of Mathematics
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
e-mail:perh@math.chalmers.se

Abstract

This paper gives a probabilistic interpretation of a class of finite difference schemes often referred to as the θ-method. In particular, the present paper shows that for some parameter values the θ-method can been seen as a binomial tree with a random time.

Key words: Finite difference method, θ-method, binomial tree.

1 Introduction

It is well-known that the explicit finite difference method for the heat equation is equivalent to a trinomial tree, see e.g. Heston and Zhou [2]. This paper studies the θ-method, which is a class of finite difference methods including, for instance, the Crank-Nicolson method, and shows that for some parameter values the θ-method also has a probabilistic interpretation. To be more precise, the θ-method can for certain parameter values be seen as a binomial tree with a random time.

2 The θ-Method and Its Probabilistic Counterpart

To begin with, consider the initial value problem for the heat equation,

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} & \text{in } (0,T] \times \mathbb{R}, \\ u|_{t=0} = f & \text{on } \mathbb{R}, \end{cases} \tag{1}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a given continuous and bounded function (abbr. $f \in C_b$) and T is a positive constant. There are many various approaches to solve

*The author would like to thank Prof. Christer Borell, Chalmers University of Technology, and Prof. Chris Rogers, University of Bath, for comments and suggestions.

the initial value problem for the heat equation and this section will discuss a finite difference method known as the θ-method.

Let h and λ be positive numbers and set $k = \lambda h^2$. Think of h and k as small increments of the variables x and t. Moreover, suppose $v_n(x) = u(nk, x)$, $x \in \mathbb{R}$, $n \in \mathbb{N}$. The idea behind the θ-method is to approximate the function v_n by $\tilde{v}_n$, where $\tilde{v}_n$ is the solution of the difference equation

$$\begin{cases} \frac{\tilde{v}_{n+1}(x)-\tilde{v}_n(x)}{k} = \frac{1}{2}\Big(\theta \frac{\Delta_h^2 \tilde{v}_{n+1}(x)}{h^2} + (1-\theta)\,\frac{\Delta_h^2 \tilde{v}_n(x)}{h^2}\Big) & x \in \mathbb{R},\ n \in \mathbb{N}, \\ v_0 = f & x \in \mathbb{R}. \end{cases} \tag{2}$$

Here $0 \le \theta \le 1$ and

$$\Delta_h^2 g(x) = g(x+h) - 2g(x) + g(x-h),$$

for any $g \in C_b$.

Now, introduce for each $n \in \mathbb{N}_+$ a finite difference operator $V_n^{(\lambda,h)} : C_b \to C_b$, defined for $n = 1$ by the operator equation

$$V_1^{(\lambda,h)} - 1 = \frac{1}{2}\lambda\big(\theta V_1^{(\lambda,h)} + 1 - \theta\big)\Delta_h^2 \tag{3}$$

and for $n \ge 2$ by

$$V_n^{(\lambda,h)} = \big(V_1^{(\lambda,h)}\big)^n.$$

From equation (2) it is readily seen that the solution of the θ-method may be written as

$$\tilde{v}_n(x) = \big(V_n^{(\lambda,h)} f\big)(x).$$

When $\theta = 0$ the operator $V^{(\lambda,h)}$ is called explicit. Each equation of an explicit operator gives the unknown $\tilde{v}_{n+1}(x)$ directly in terms (finitely many) of the known quantities $\tilde{v}_n(x + jh)$. If θ is not 0, one must solve a set of linear equations to obtain $\tilde{v}_{n+1}(x)$, and the operator $V^{(\lambda,h)}$ is called implicit. The important special case $\theta = 1/2$ is often referred to as the Crank-Nicolson method.

The θ-method is not stable for all values of $\lambda > 0$ and $0 \le \theta \le 1$. Stability means that the collection of operators $\{V_n^{(\lambda,h)},\ 0 < h \le h_0,\ 0 < n\lambda h^2 \le T\}$, where $h_0 > 0$ and λ are fixed, is uniformly bounded with respect to the (operator) norm in C_b, i.e. there is a constant K such that

$$\|V_n^{(\lambda,h)}\|_{C_b} \le K$$

uniformly for all $0 < h \le h_0$ and all n such that $0 < n\lambda h^2 \le T$. Here $\|f\|_{C_b} = \sup_x |f(x)|$. Stability is a necessary and sufficient condition for uniform convergence in connection with the θ-method. However, it can be shown that the θ-method is stable if, and only if,

$$\begin{aligned} &\lambda \leq \frac{1}{1-2\theta}, && \text{if } 0 \leq \theta < 1/2, \\ &\text{no restriction,} && \text{if } 1/2 \leq \theta \leq 1. \end{aligned}$$

For a further discussion about the θ-method and other finite difference methods, see Atkinsson et al. [1] or Rictmeyer et al. [3].

In certain cases it is possible to give a probabilistic interpretation of the θ-method. To see this, assume that ζ is a lattice random variable with span 1, i.e. $P(\zeta \in \mathbb{Z}) = 1$, expectation zero, and bounded second moment. Furthermore, suppose $\zeta_1, \ldots, \zeta_n$ are independent stochastic copies of ζ and set

$$(U_n^{(\zeta,h)} f)(x) = E\big[f(x + h\sum_{j=1}^{n} \zeta_j)\big], \quad n \in \mathbb{N}_+.$$

The operator $U_n^{(\zeta,h)}$ will henceforth be referred to as a *lattice method* or a *lattice tree*. In particular, if ε is a Rademacher random variable, i.e. $P(\varepsilon = -1) = P(\varepsilon = 1) = 1/2$, then $U_n^{(\varepsilon,h)}$ will be called a *binomial tree*. The aim in this section is to prove that for some values of θ and λ there is a lattice random variable ζ with $\mathrm{Var}(\zeta) = \lambda$ such that

$$U_n^{(\zeta,h)} f = V_n^{(\lambda,h)} f \tag{4}$$

for all $f \in C_b$, $h > 0$ and each $n \in \mathbb{N}_+$ (abbr. $U^{(\zeta,\cdot)} = V^{(\lambda,\cdot)}$).

To prove the statement in equation (4), note that it is sufficient to assume that $n = 1$ since $U_n^{(\zeta,h)} = (U_1^{(\zeta,h)})^n$. Moreover, if $\theta = 0$ and $\lambda \leq 1$ then it is obvious that equation (4) follows by simply letting ζ be defined by

$$P(\zeta = 0) = 1 - \lambda \quad \text{and} \quad P(\zeta = -1) = P(\zeta = 1) = \frac{\lambda}{2}.$$

Next we will consider the case $0 < \theta \leq 1$. An approximation argument gives that it is only necessary to show that equation (4) holds for all $f \in C_b$ such that f is integrable with respect to the Lesbegue measure. Now, it follows as in Richtmeyer et al. [3] p. 17, that the Fourier transform of the function $x \mapsto (V_1^{(\lambda,h)} f)(x)$ is given by

$$\int_{-\infty}^{\infty} (V_1^{(\lambda,h)} f)(x) e^{-i\xi x} dx = \phi(h\xi)\hat{f}(\xi), \quad \xi \in \mathbb{R},$$

where $\hat{f}$ is the Fourier transform of f and

$$\phi(\xi) = \frac{1 - \lambda(1-\theta)(1-\cos\xi)}{1 + \theta\lambda(1-\cos\xi)}.$$

Indeed, this follows at once from equation (3) and the fact that $f(x+h)$ possesses the Fourier transform $\hat{f}(\xi)\exp(ih\xi)$. The function ϕ above is often referred to as the *symbol* or the *characteristic polynomial* of the θ-method.

Next, observe that the Fourier transform of $x \mapsto (U_1^{(\zeta,h)} f)(x)$ equals

$$\int_{-\infty}^{\infty} (U_1^{(\zeta,h)} f)(x) e^{-i\xi x} dx = E\big[e^{ih\xi\zeta}\big]\hat{f}(\xi), \quad \xi \in \mathbb{R}.$$

Hence, equation (4) follows if $\xi \mapsto \phi(\xi)$ is a characteristic function of a lattice random variable with span 1 and variance λ. To examine under what conditions this is true, note that ϕ may be written as

$$\phi(\xi) = \frac{1 - \lambda(1-\theta)(1-\cos\xi)}{(1+\theta\lambda)(1-\kappa\cos\xi)},$$

where $\kappa = \lambda\theta/(1+\lambda\theta)$. Moreover, if $\theta > 0$ then a Taylor expansion yields

$$\begin{aligned} \phi(\xi) &= \frac{1 - \lambda\big(1-\theta\big)\big(1-\cos\xi\big)}{1+\theta\lambda} \sum_{j=0}^{\infty} \kappa^j \cos^j \xi \\ &= \sum_{j=0}^{\infty} \rho_j \cos^j \xi \end{aligned} \tag{5}$$

where

$$\rho_j = \begin{cases} p, & \text{if } j = 0, \\ q(1-p)(1-q)^{j-1}, & \text{if } j \geq 1, \end{cases}$$

with

$$p = 1 - \frac{\lambda}{1+\lambda\theta} \quad \text{and} \quad q = \frac{1}{1+\lambda\theta}.$$

If λ and θ is chosen such that $p \geq 0$, then $\rho_j \geq 0$ for all j. Moreover, since $\phi(0) = 1$ we have $\sum_0^\infty \rho_j = 1$. Recall that $\xi \mapsto \cos\xi$ is the characteristic function of a Rademacher random variable ε. Consequently, if $p \geq 0$ then[1]

$$\phi(\xi) = E\big[e^{i\xi\sum_{j=1}^{\nu}\varepsilon_j}\big]$$

where $\{\varepsilon_j\}_{j=1}^\infty$ is a series of independent stochastic copies of ε and ν is a random variable that is independent of $\{\varepsilon_j\}_{j=1}^\infty$ with $P(\nu = j) = \rho_j$ for each $j \geq 0$. In particular, ϕ is the characteristic function of a lattice random variable with span 1. In addition, since $\phi'(0) = 0$ and $\phi''(0) = -\lambda$ the random variable $\sum_{j=1}^{\nu} \varepsilon_j$ has expectation zero and variance λ.

[1] We assume $\sum_{j=1}^{0} = 0$

Theorem 1. *Suppose $0 \le \theta \le 1$. There exists a random variable ν, independent of $\{\varepsilon_j\}_{j=1}^{\infty}$, such that the random variable*

$$\zeta = \sum_{j=1}^{\nu} \varepsilon_j \tag{6}$$

satisfies

$$U^{(\zeta,\cdot)} = V^{(\lambda,\cdot)}$$

if, and only if, $\lambda > 0$ is chosen such that

$$\begin{aligned} &\lambda \le \frac{1}{1-\theta}, && \textit{if } \theta < 1, \\ &\textit{no restriction}, && \textit{if } \theta = 1. \end{aligned}$$

Proof. We have already proven that the conditions on λ are sufficient but it remains one extra argument to prove that the conditions also are necessary. To be more precise, we need to show that the only choice of parameters d_j, $j \in \mathbb{N}$, satisfying

$$\phi(\xi) = \sum_{j=0}^{\infty} d_j \cos^j \xi, \quad \text{for all } \xi \in \mathbb{R},$$

are $d_j = \rho_j$, $j \in \mathbb{N}$. By substituting $x = \cos\xi$ this follows directly from the theory of analytic functions. $\square$

Finally in this section necessary conditions will be given such that $U^{(\zeta,\cdot)} = V^{(\lambda,\cdot)}$ for some lattice random variable ζ. Note that Theorem 1 only yields sufficient conditions.

It is evident that if $\theta = 0$, then $\lambda \le 1$ is equivalent to the existence of lattice random variable ζ such that $U^{(\zeta,\cdot)} = V^{(\lambda,\cdot)}$. Now, suppose that $0 < \theta \le 1$. A Fourier expansion of ϕ yields

$$\phi(\xi) = \frac{1}{2\pi} c_0 + \frac{1}{\pi} \sum_{k=1}^{\infty} c_k \cos k\xi$$

where

$$c_k = \int_0^{2\pi} \phi(\xi) \cos k\xi \, d\xi. \tag{7}$$

Recall that $\phi(0) = 1$, and thus, the function ϕ is a characteristic function of a lattice random variable ζ if, and only if, $c_k \ge 0$ for all k and, in that case, ζ is symmetric with $P(\zeta = k) = c_k/2\pi$, $k \ge 0$.

First we prove that $c_k \geq 0$ for all $k \geq 1$ and each $\lambda > 0$, $0 \leq \theta \leq 1$. Namely, equation (5) yields

$$c_k = \sum_{j=0}^{\infty} \rho_j \int_0^{2\pi} \cos^j \xi \cos k\xi \, d\xi.$$

In addition,

$$\begin{aligned} \cos^j \xi &= E\big[e^{i\xi \sum_{k=0}^{j} \varepsilon_k}\big] \\ &= \sum_{l=0}^{j} d_{l,j} \cos l\xi, \end{aligned}$$

where $d_{0,j} = P(\sum_{k=0}^{j} \varepsilon_k = 0)$ and $d_{l,j} = 2P(\sum_{k=0}^{j} \varepsilon_k = l)$ for $l \geq 1$. Thus, orthogonality implies

$$c_k = \pi \sum_{j=k}^{\infty} \rho_j d_{k,j}, \quad k \geq 1.$$

Recall that $\rho_j \geq 0$ for all $j \geq 1$ and hence, it is enough to find conditions such that $c_0 \geq 0$.

By substituting $x = \tan \xi/2$ in equation (7) follows

$$c_0 = 2 \int_0^{\infty} \frac{1 + \big(1 - 2\lambda(1-\theta)\big)\, x^2}{1 + (1 + 2\lambda\theta)\, x^2} \frac{2dx}{1+x^2}$$

which after some elementary calculus yields

$$c_0 = \frac{2\pi}{\theta} \left(\theta - 1 + \sqrt{\frac{1}{1 + 2\theta\lambda}} \right).$$

To sum up, we have shown

Theorem 2. *Suppose* $0 \leq \theta \leq 1$. *There exists a lattice random variable* ζ *such that*

$$U^{(\zeta,\cdot)} = V^{(\lambda,\cdot)}$$

if, and only if, $\lambda > 0$ *is chosen such that*

$$\begin{aligned} &\lambda \leq 1, && \text{if } \theta = 0, \\ &\lambda \leq \frac{1}{1-\theta} \frac{2-\theta}{2(1-\theta)}, && \text{if } 0 < \theta < 1, \\ &\text{no restriction}, && \text{if } \theta = 1. \end{aligned}$$

It is obvious that the collection $\{U_n^{(\zeta,h)},\ 0 < h \leq h_0,\ 0 < n\lambda h^2 \leq T\}$ is stable, to be more precise, for any ζ and any $h_0 > 0$ we have

$$\|U_n^{(\zeta,h)}\|_{C_b} \leq 1$$

uniformly for all $0 < h \leq h_0$ and all n such that $0 < n\lambda h^2 \leq T$. Thus, the θ-method provide us with an example showing that the class of all stable finite difference methods is strictly larger than the class of all lattice methods. In addition, it shows that the class of all lattice methods with symmetrical lattice random variables is strictly larger than the class of all binomial trees with independent random times.

References

[1] Atkinson, K. & Weimin, H. (2001) *Theoretical Numerical Analysis - A Functional Analysis Framework* Springer Verlag

[2] Heston, S. & Zhou, G. (2000) *On the Rate of Convergence of Discrete-Time Contingent Claims* Math. Finance **10**, 53-75

[3] Richtmeyer, R. & Morton, K.W. (1967) *Difference Methods for Initial Value Problems* John Wiley & Sons