Analysis of an upwind finite difference scheme

Jordan Hoffart

1 Introduction

We consider a certain finite difference approximation of a particular ODE and we analyze it. We consider whether the scheme is consistent, stable, and convergent.

2 Problem

Let $b > 0$ and $f : [0, 1] \to \mathbb{R}$. Consider the following BVP: find $u : [0, 1] \to \mathbb{R}$ such that

$$\begin{aligned} -u''(x) + bu'(x) &= f(x), \quad x \in (0, 1), \\ u(0) &= 0, \\ u(1) &= 0. \end{aligned}$$

3 Discretization

Let $N > 0$, $h = 1/(N + 1)$, and $x_i = ih$ for $i = 0, \dots, N + 1$. Let $f_i = f(x_i)$. We discretize the ODE above using the following finite difference scheme: find a vector $\vec{u} = (u_i)_{i=0}^{N+1}$ such that

$$\begin{aligned} -\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} + b\frac{v_{i+1} - v_i}{h} &= f_i, \quad i \in \{1, \dots, N\} \qquad (1) \\ v_0 &= 0, \\ v_{N+1} &= 0. \end{aligned}$$

4 Existence and uniqueness

We now show that the discretized system has a unique solution. For this, we need a few lemmas.

Lemma 1 (Maximum principle)**.** *Suppose N is large enough so that $h \leq 1/b$. Also suppose that $f_i \leq 0$ for $i \in \{1, \dots, N\}$. Then $v_i \leq \max\{v_0, v_{N+1}\} = 0$ for all i.*

Proof. We rewrite (1) as

$$\underbrace{(1-bh)}_{\geq 0}(v_i - v_{i+1}) + (v_i - v_{i-1}) = h^2 f_i \leq 0.$$

Now if there exists $i \in \{1, \ldots, N\}$ such that

$$v_j \leq v_i \text{ for all } j, \tag{2}$$

then we have that

$$0 \leq \underbrace{(1-bh)}_{\geq 0}\underbrace{(v_i - v_{i+1})}_{\geq 0} + \underbrace{(v_i - v_{i-1})}_{\geq 0} = h^2 f_i \leq 0.$$

Therefore, we have that

$$\underbrace{(1-bh)}_{\geq 0}\underbrace{(v_i - v_{i+1})}_{\geq 0} + \underbrace{(v_i - v_{i-1})}_{\geq 0} = 0.$$

This implies that $v_{i-1} = v_i = v_{i+1}$. Applying this to the smallest such i with property (2) gives us a contradiction, so no such i can exist. In other words, for all j,

$$v_j \leq \max_j v_j = \max\{v_0, v_{N+1}\} = 0.$$

□

Corollary 2 (Minimum principle)**.** *Suppose N is large enough so that $h \leq 1/b$. Also suppose that $f_i \geq 0$ for $i \in \{1, \ldots, N\}$. Then $v_i \geq \min\{v_0, v_{N+1}\} = 0$ for all i.*

Proof. Let $v_i^- = -v_i$ and $f_i^- = -f_i$. Then v_i^- satisfies

$$\begin{aligned} -\frac{v_{i+1}^- - 2v_i^- + v_{i-1}^-}{h^2} + b\frac{v_{i+1}^- - v_i^-}{h} &= f_i^-, \quad i \in \{1, \ldots, N\} \\ v_0^- &= 0, \\ v_{N+1}^- &= 0. \end{aligned}$$

We also have that $f_i^- \leq 0$ for all i. Therefore, by the maximum principle,

$$-v_i = v_i^- \leq \max\{v_0^-, v_{N+1}^-\} = \max\{-v_0, -v_{N+1}\} = -\min\{u_0, u_{N+1}\} = 0$$

for all i. Thus,

$$v_i \geq \min\{v_0, v_{N+1}\} = 0$$

for all i. □

Corollary 3 (Existence and uniqueness)**.** *Let N be large enough so that $h \leq 1/b$. For any vector $\vec{f} = (f_i)_{i=1}^N$, there exists a unique vector $\vec{v} = (v_i)_{i=0}^{N+1}$ such that*

$$-\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} + b\frac{v_{i+1} - v_i}{h} = f_i, \quad i \in \{1, \dots, N\}$$
$$v_0 = 0,$$
$$v_{N+1} = 0.$$

Proof. The above system is a square linear system of equations in the v_i, so it suffices to show that if $\vec{z}$ solves

$$-\frac{z_{i+1} - 2z_i + z_{i-1}}{h^2} + b\frac{z_{i+1} - z_i}{h} = 0, \quad i \in \{1, \dots, N\}$$
$$z_0 = 0,$$
$$z_{N+1} = 0,$$

then $\vec{z} = 0$. In this case, we can apply both the maximum and the minimum principle to $\vec{z}$ to conclude that

$$0 = \min\{z, z_{N+1}\} \leq z_i \leq \max\{z_0, z_{N+1}\} = 0$$

for all i. Thus $\vec{z} = 0$. □

Therefore, we know that the discrete system always produces a solution, provided that h is sufficiently small. What we will now show is that, whenever we have a smooth solution to the ODE, the corresponding discrete solution approximates it in a certain way. We now make this more precise by discussing consistency, stability, and convergence.

5 Consistency

Now we examine the consistency error of the scheme and its dependence on b and h.

Lemma 4 (Consistency error)**.** *Let $u : [0, 1] \to \mathbb{R}$ be smooth (more precisely, 3 times differentiable with 2 continuous derivatives) and satisfy*

$$-u''(x) + bu'(x) = f(x), \quad x \in (0, 1),$$
$$u(0) = 0,$$
$$u(1) = 0.$$

Let $u_i = u(x_i)$ for all i.

$$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + b\frac{u_{i+1} - u_i}{h} = f_i + hR_i(b, h), \quad i \in \{1, \dots, N\}$$
$$u_0 = 0,$$
$$u_{N+1} = 0,$$

where $f_i = f(x_i)$ *for all* i *and*

$$R_i(b,h) = \frac{b}{2}u''(x_i) + \frac{1}{6}(u'''(\xi_{i-1}) - u'''(\xi_{i+1})) + \frac{bh}{6}u'''(\xi_{i+1}). \tag{3}$$

Proof. From Taylor's theorem with remainder in Lagrange form, for $i \in \{1,\dots,N\}$,

$$\begin{aligned} u_{i+1} &= u_i + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{6}u'''(\xi_{i+1}), \\ u_{i-1} &= u_i - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{6}u'''(\xi_{i-1}), \end{aligned}$$

for some points $\xi_{i\pm1}$ between x_i and $x_{i\pm1}$. Therefore,

$$\begin{aligned} b\frac{u_{i+1}-u_i}{h} &= bu'(x_i) + \frac{bh}{2}u''(x_i) + b\frac{h^2}{6}u'''(\xi_{i+1}), \\ -\frac{u_{i+1}-2u_i+u_{i-1}}{h^2} &= -u''(x_i) + \frac{h}{6}(u'''(\xi_{i-1}) - u'''(\xi_{i+1})). \end{aligned}$$

Thus,

$$-\frac{u_{i+1}-2u_i+u_{i-1}}{h^2} + b\frac{u_{i+1}-u_i}{h} = \underbrace{-u''(x_i) + bu'(x_i)}_{=f_i} + hR_i(b,h). \tag{4}$$

where

$$R_i(b,h) = \frac{b}{2}u''(x_i) + \frac{1}{6}(u'''(\xi_{i-1}) - u'''(\xi_{i+1})) + \frac{bh}{6}u'''(\xi_{i+1}).$$

□

Lemma 5. *If* u *is as above and also* u''' *is bounded, then there is a constant* δ_b *such that for all* $h \le 1/b$,

$$|R_i(b,h)| \le \delta_b$$

for all i.

Proof. Let $M_2 = \max_{x\in[0,1]} |u''(x)|$, $M_3 = \max_{x\in[0,1]} |u'''(x)|$. Then by applying the triangle inequality,

$$|R_i(b,h)| \le \delta_b = (bM_2 + M_3)/2.$$

□

Remark 6 (Consistency). If $\vec{v}$ solves the finite difference approximation with $f_i = f(x_i)$ and u solves the ODE with f, then the error $e_i = v_i - u_i = v_i - u(x_i)$ satisfies

$$\begin{aligned} -\frac{e_{i+1}-2e_i+e_{i-1}}{h^2} + b\frac{e_{i+1}-e_i}{h} &= -hR_i(b,h), \quad i \in \{1,\dots,N\} \\ e_0 &= 0, \\ e_{N+1} &= 0, \end{aligned}$$

The previous lemma implies that $hR_i(b,h) = O(h)$. Therefore, combined with above, this implies that the method is consistent of order h.

6 Stability

Our goal in this section is to prove the following stability result.

Theorem 7 (Stability)**.** *There is a constant $C_b > 0$ only dependent on b and a tolerance $0 < h_b \leq 1/b$ only dependent on b such that when N is large enough so that $h = 1/(N+1)$ satisfies $h \leq h_b$ we have that the exact solution $\vec{v}$ to*

$$\begin{aligned} -\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} + b\frac{v_{i+1} - v_i}{h} &= f_i, \quad i \in \{1, \dots, N\} \\ v_0 &= 0, \\ v_{N+1} &= 0, \end{aligned}$$

satisfies the stability estimate

$$\max_{i \in \{1,\dots,N\}} |v_i| \leq C_b \max_{i \in \{1,\dots,N\}} |f_i|.$$

6.1 The standard stability argument fails

For instructional purposes, we show that the standard stability argument fails. We proceed as far as possible, then we mention where exactly the argument breaks down. Since we are in such a simple setting, we can directly solve the ODE system for certain source functions f. We consider the case when $f = 1$.

Lemma 8. *Let*

$$w(x) = \frac{x}{b} - \frac{e^{bx} - 1}{b(e^b - 1)}. \tag{5}$$

Then w satisfies

$$\begin{aligned} -w''(x) + bw'(x) &= 1, \quad x \in (0,1), \\ w(0) &= 0, \\ w(1) &= 0. \end{aligned}$$

Proof. Follows from direct computation. □

Now we find a formula for the consistency error for this particular solution.

Lemma 9. *Let w be as above and $w_i = w(x_i)$. Then*

$$\begin{aligned} -\frac{w_{i+1} - 2w_i + w_{i-1}}{h^2} + b\frac{w_{i+1} - w_i}{h} &= 1 - R_i(h), \quad i \in \{1, \dots, N\} \\ w_0 &= 0, \\ w_{N+1} &= 0, \end{aligned}$$

where

$$R_i(h) = \frac{e^{bh} - 1}{bh} \frac{be^{x_{i-1}}}{e^b - 1} \frac{1 - (1 - bh)e^{bh}}{bh} \geq 0$$

Proof. We have that

$$\begin{aligned} w_{i+1} - w_i &= \frac{h}{b} - \frac{1}{b(e^b - 1)} \left(e^{bx_{i+1}} - e^{bx_i}\right) \\ &= \frac{h}{b} - \frac{e^{bx_i}}{b(e^b - 1)} (e^{bh} - 1). \end{aligned}$$

Thus

$$\begin{aligned} w_{i+1} - w_i - (w_i - w_{i-1}) &= -\frac{e^{bh} - 1}{b(e^b - 1)} (e^{bx_i} - e^{bx_{i-1}}) \\ &= -\frac{(e^{bh} - 1)^2}{b(e^b - 1)} e^{bx_{i-1}}. \end{aligned}$$

Hence,

$$\begin{aligned} b\frac{w_{i+1} - w_i}{h} &= 1 - \frac{be^{bx_i}}{e^b - 1} \frac{e^{bh} - 1}{bh}, \\ -\frac{w_{i+1} - 2w_i + w_{i-1}}{h^2} &= \frac{be^{bx_{i-1}}}{e^b - 1} \left(\frac{e^{bh} - 1}{bh}\right)^2. \end{aligned}$$

Therefore,

$$-\frac{w_{i+1} - 2w_i + w_{i-1}}{h^2} + b\frac{w_{i+1} - w_i}{h} = 1 - R_i(h)$$

where

$$\begin{aligned} R_i(h) &= -\frac{e^{bh} - 1}{bh} \frac{b}{e^b - 1} \left(e^{bx_{i-1}} \frac{e^{bh} - 1}{bh} - e^{bx_i}\right) \\ &= -\frac{e^{bh} - 1}{bh} \frac{be^{x_{i-1}}}{e^b - 1} \frac{(1 - bh)e^{bh} - 1}{bh} \\ &= \frac{e^{bh} - 1}{bh} \frac{be^{x_{i-1}}}{e^b - 1} \frac{1 - (1 - bh)e^{bh}}{bh} \geq 0. \end{aligned}$$

□

Here is where the standard stability argument breaks down. Since $R_i(h) \geq 0$, we have that for w as in Lemma 9,

$$\begin{aligned} -\frac{w_{i+1} - 2w_i + w_{i-1}}{h^2} + b\frac{w_{i+1} - w_i}{h} &= 1 - R_i(h) \leq 1, \quad i \in \{1, \ldots, N\} \\ w_0 &= 0, \\ w_{N+1} &= 0, \end{aligned}$$

At the moment, we cannot apply the maximum principle or the minimum principle to w_i, which hinders the usual stability argument. Indeed, if we let $\vec{v}$ be the unique solution to

$$-\frac{v_{i+1}-2v_i+v_{i-1}}{h^2}+b\frac{v_{i+1}-v_i}{h}=f_i,\quad i\in\{1,\dots,N\}$$
$$v_0=0,$$
$$v_{N+1}=0,$$

then the standard argument is to set $v_i^{\pm}=v_i\pm Mw_i$, where

$$M=\max_{i\in\{1,\dots,N\}}|f_i|.$$

Then we have that $\vec{v}^+$ satisfies

$$-\frac{v_{i+1}^+-2v_i^++v_{i-1}^+}{h^2}+b\frac{v_{i+1}^+-v_i^+}{h}=f_i+(1-R_i(h))M\le f_i+M\ge 0,\quad i\in\{1,\dots,N\}$$
$$v_0=0,$$
$$v_{N+1}=0,$$

and we see that the the inequality needed to apply the minimum principle is facing the wrong way. A similar difficulty also applies to $\vec{v_i}^-$ and the maximum principle, so we cannot proceed this way. We circuvent this difficulty by using the general consistency result from earlier.

6.2 The correct stability argument

Now we are ready to prove Theorem 7, which we restate below.

Theorem 10 (Stability)**.** *There is a constant $C_b>0$ only dependent on b and a tolerance $0<h_b\le 1/b$ only dependent on b such that when N is large enough so that $h=1/(N+1)$ satisfies $h\le h_b$ we have that the exact solution $\vec{v}$ to*

$$-\frac{v_{i+1}-2v_i+v_{i-1}}{h^2}+b\frac{v_{i+1}-v_i}{h}=f_i,\quad i\in\{1,\dots,N\}$$
$$v_0=0,$$
$$v_{N+1}=0,$$

satisfies the stability estimate

$$\max_{i\in\{1,\dots,N\}}|v_i|\le C_b\max_{i\in\{1,\dots,N\}}|f_i|.$$

Proof. We let w be as in Lemma 9. Then since w is analytic, we can apply Lemma 4 to get that

$$-\frac{w_{i+1}-2w_i+w_{i-1}}{h^2}+b\frac{w_{i+1}-w_i}{h}=1+hR_i(b,h),\quad i\in\{1,\dots,N\}$$
$$w_0=0,$$
$$w_{N+1}=0.$$

However, from Lemma 9, we also have that

$$-\frac{w_{i+1}-2w_i+w_{i-1}}{h^2}+b\frac{w_{i+1}-w_i}{h} = 1-R_i(h), \quad i \in \{1,\dots,N\}$$
$$w_0 = 0,$$
$$w_{N+1} = 0.$$

Comparing these and using Lemma 5 tells us that

$$R_i(h) = -hR_i(b,h) \le \delta_b h$$

for all $h \le 1/b$. Thus, if we set

$$h \le h_b := \min\{1/b, 1/(2\delta_b)\},$$

then we have that

$$R_i(h) \le 1/2$$

for all i and all h.

Now we can proceed with our usual stability argument. We let $\vec{v}$ be the unique solution to

$$-\frac{v_{i+1}-2v_i+v_{i-1}}{h^2}+b\frac{v_{i+1}-v_i}{h} = f_i, \quad i \in \{1,\dots,N\}$$
$$v_0 = 0,$$
$$v_{N+1} = 0,$$

then we set

$$v_i^{\pm} = \frac{1}{2}v_i \pm Mw_i,$$

where

$$M = \max_{x\in[0,1]} |f(x)|.$$

Then we have that $\vec{v}^+$ satisfies

$$-\frac{v_{i+1}^+-2v_i^++v_{i-1}^+}{h^2}+b\frac{v_{i+1}^+-v_i^+}{h} = \frac{1}{2}f_i+(1-R_i(h))M \ge \frac{1}{2}(f_i+M) \ge 0, \quad i \in \{1,\dots,N\}$$
$$v_0 = 0,$$
$$v_{N+1} = 0.$$

Therefore, we can apply the minimum principle and conclude that

$$\frac{1}{2}v_i + Mw_i \ge 0$$

for all i, i.e.

$$-v_i \le 2w_iM \le 2\|w\|_\infty M \le \frac{2}{b}M,$$

where the last equality comes from explicitly bounding $w(x)$ above on $[0,1]$ from (5). Similarly, we have that

$$-\frac{v_{i+1}^- - 2v_i^- + v_{i-1}^-}{h^2} + b\frac{v_{i+1}^- - v_i^-}{h} = \frac{1}{2}f_i - (1-R_i(h))M \leq \frac{1}{2}(f_i - M) \leq 0, \quad i \in \{1,\dots,N\}$$
$$v_0 = 0,$$
$$v_{N+1} = 0.$$

Therefore, we can apply the maximum principle and conclude that

$$\frac{1}{2}v_i - Mw_i \leq 0$$

for all i, i.e.

$$v_i \leq \frac{2}{b}M.$$

In summary, when N is large enough so that $h = 1/(N+1)$ satisfies

$$h \leq h_b = \min\left\{1/b, 1/(2\delta_b)\right\},$$

we have that the exact solution $\vec{v}$ to

$$-\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} + b\frac{v_{i+1} - v_i}{h} = f(x_i), \quad i \in \{1,\dots,N\}$$
$$v_0 = 0,$$
$$v_{N+1} = 0,$$

satisfies the stability estimate

$$\max_{i\in\{1,\dots,N\}} |v_i| \leq \frac{2}{b}\max_{i\in\{1,\dots,N\}} |f_i|.$$

□

Remark 11. The particular tolerance h_b found above is far from optimal. Numerical experiments show that the solution remains stable simply for $h \leq 1/b$.

7 Convergence

Now we combine our consistency and stability results in order to prove convergence.

Theorem 12 (Convergence)**.** *Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Suppose that we have a solution $u : [0,1] \to \mathbb{R}$ to*

$$-u'' + bu' = f \text{ on } (0,1)$$
$$u(0) = 0$$
$$u(1) = 0$$

Let N be large enough so that $h = 1/(N+1) \leq 1/b$. Let $x_i = ih$ for $i \in \{0, \ldots, N+1\}$ and let $f_i = f(x_i)$ for all i. Let $\vec{v}$ be the unique solution to

$$-\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} + b\frac{v_{i+1} - v_i}{h} = f_i, \quad i \in \{1, \ldots, N\}$$
$$v_0 = 0$$
$$v_{N+1} = 0$$

that is smooth (of class C^2 and with a bounded third derivative). Let $u_i = u(x_i)$ for all i, and let $e_i = v_i - u_i$ for all i. Then there is a constant C_b depending only on b and a tolerance $0 < h_b \leq 1/b$ only depending on b such that for all $h \leq h_b$,

$$\max_{i \in \{1,\ldots,N\}} |e_i| \leq C_b h.$$

Proof. We use the consistency error in Lemma 4 and conclude that the u_i satisfy

$$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + b\frac{u_{i+1} - u_i}{h} = f_i + hR_i(b,h), \quad i \in \{1, \ldots, N\}$$
$$u_0 = 0,$$
$$u_{N+1} = 0,$$

Therefore, the errors e_i satisfy

$$-\frac{e_{i+1} - 2e_i + e_{i-1}}{h^2} + b\frac{e_{i+1} - e_i}{h} = -hR_i(b,h), \quad i \in \{1, \ldots, N\}$$
$$e_0 = 0,$$
$$e_{N+1} = 0,$$

Let

$$g_i = -hR_i(b,h)$$

for each i. Then the general stability result from above as well as the general consistency error lets us conclude that

$$\max_{i \in \{1,\ldots,N\}} |e_i| \leq C_b \max_{i \in \{1,\ldots,N\}} h|R_i(b,h)| \leq C_b \delta_b h$$

for all $h \leq h_b$. ☐

8 Conclusion

We analyzed an upwind difference approximation of a particular ODE and showed that, provided that h is small enough to guarantee stability, the method converges pointwise uniformly with order 1 rate of convergence.