Numerical Analysis

Matthew Berry

August 26, 2019

For various differential equations, solutions cannot be found analytically. As a result numerical schemes are required to approximate the solution. It is useful to obtain an estimate of how accurate the solution is.To make the analyisis easier it is useful to introduce some notation.
Let P be a differential operator such that,

$$Pu = 0, \tag{1}$$

where $u : (0, \infty) \times \Omega \subset \mathbb{R}^n \to \mathbb{R}$.

1 The Finite Difference Operator

To introduce the finite difference operator we need to consider the space of functions that it operates on. To think about this we first consider the domain $\Omega = \mathbb{R}$. The standard descretisation of this domain is into a grid of equally spaced nodes with spacing h. This can be denoted by,

$$h\mathbb{Z} = \{hz : z \in \mathbb{Z}\}.$$

A similar discretisation is required for the time component as well. We define,

$$k\mathbb{N}_0 = \{kn : n \in \mathbb{N} \cup \{0\}\},$$

where k is the spacing of each nodes. For practicality we consider $k, h > 0$. What we seek is a difference operator, $P_{h,k}$ such that solutions to the equation,

$$P_{k,h} v_m^n = 0, \tag{2}$$

can be used to approximate the solutions to equation 1. A common example of a finite differnce operator is the forward difference operator, δ_+, given by,

$$\delta_+ v_m = \frac{v_{m+1} - v_m}{h}.$$

This is often denoted as δ_{x+} as it is with respect to the x variable. Similarly we can define $\delta_t +$ as,

$$\delta_{t+} v^n = \frac{v^{n+1} - v^n}{k}.$$

We also have a backward and central difference operators δ_- and δ_0, defined by,

$$\delta_- v_m = \frac{v_m - v_{m-1}}{h},$$

$$\delta_0 v_m = \frac{v_{m+1} - v_{m-1}}{2h}.$$

In each of these cases h, k represents the grid spacing of the function v^n_m. The central difference can also ve generated by the average of the forward and backward differences.

These differnces are used to approximate first derivatives. We can find a similar finite difference for higher order derivatives. The central difference operator for the second derivative is $delta_+ \delta_-$ which is denoted by δ^2 .

2 Convergence

We have mentioned that it would be good to use this difference operator to produce approxomate solutions, though we have not set a definition for convergence.

Definition 1. We say that a finite difference scheme approximating a differential operator is convergent, if for a solution $u(t, x)$ of equation 1 and the solution, v^n_m, to equation 2 such that v^0_m converges to $u(0, x)$ as mh approaches x, then v^n_m converges to $u(t, x)$ as (nk, mh) converges to (t, x) as $h, k \to 0$.

This definition does not provide a norm for the convergence and the functions, $u(t_n, x)$ and v^n_m are in different spaces so it is not clear as to how to define convergence.

For finite difference schemes it is useful to to consider consistency. Consistency is the idea that the differential operator is well approximated by the difference operator.

Definition 2. A finite difference operator $P_{k,h}$ is consistent with a differential operator P if for any smooth function ϕ, $|P_{k,h}\phi - P\phi| \to 0$ as $h, k \to 0$, where the convergence is pointwise at each gridpoint.

Example 2.1. We look at the example where $P = \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2} - F$, and F is a function of the argument. The finite difference operator $P_{k,h} = \delta_{t+} - \delta_x^2 - F$. We wish to show consistency of these operators.
To determine consistency we need only to show that the finite difference approximations are consistent with the respective differntial forms since,

$$|P\phi - P_{k,h}\phi| \leq |\frac{\partial \phi}{\partial t} - \delta_{t+}\phi| + |\frac{\partial^2 \phi}{\partial x^2} - \delta_x^2 \phi|. \tag{3}$$

The functional terms $F(\phi)$ drop out of the equation since the difference operator just evaluates the funstion at that point.
For the forward differnce, since ϕ is smooth,

$$\delta_{t+}\phi(t_n, x_m) = \frac{\phi(t_n + k, x_m) - \phi(t_n, x_m)}{k}$$

For simplicity we will drop the x_m variable as this will just be an equation for each x_m. We expand $\phi(t_n + k)$ by it's Taylor series about t_n,

$$\phi(t_n + k) = \phi(t_n) + k \frac{d\phi}{dt}\bigg|_{t=t_n} + \mathcal{O}(k^2)$$

Rearranging this equation, we get,

$$\frac{\phi(t_n + k) - \phi(t_n)}{k} = \frac{d\phi}{dt}\bigg|_{t=t_n} + \mathcal{O}(k). \tag{4}$$

Similarly we can show for the central differnce that,

$$\delta_x^2 \phi - \frac{\partial^2 \phi}{\partial x^2} = \mathcal{O}(h^2). \tag{5}$$

Combining these two components, then,

$$|P\phi - P_{k,h}\phi| \leq \mathcal{O}(k) + \mathcal{O}(h^2). \tag{6}$$

This approaches 0 as h, k approach 0.

With that example it is useful to introduce the notion of order of accuracy.

Definition 3. A finite differnce scheme $P_{k,h}v = 0$ that is consistent with the differnetial equation $Pu = 0$ is accurate of order p in time and q in space, if for any smooth function ϕ,

$$|P\phi - P_{k,h}\phi| = \mathcal{O}(k^p) + \mathcal{O}(h^q).$$

The next concept that needs to be defined is stability.

Definition 4. A linear finite difference scheme $P_{k,h}v_m^n = 0$ is stable in a stability region Λ if there exists am integer J such that for any positive time T, there exists a constant C_T such that,

$$\sum_{m=-\infty}^{\infty} |v_m^n|^2 \leq C_T \sum_{j=0}^{J} \sum_{m=-\infty}^{\infty} |v_m^j|^2, \tag{7}$$

for all $nk \leq T$ for $(h,k) \in \Lambda$.

In general for single-step time schemes, the constant J is taken as 0.

Theorem 1 (Lax-Richtmyer Equivilanve Theorem)**.** *A finite difference scheme consistent with the differential equation, for which the initial value problem is well-posed. Is convergent if and only if it is stable.*

The proof of this theorem requiress the use of fourier transforms and also a more appropriate sense of convergence.

To discuss the convergence of the finite differnce schemes, we need to consider functions in the same space. To do this we introduce two operators.

Definition 5. The Truncation T maps funtions in $L^2(\mathbb{R})$ to function in $L^2(h\mathbb{Z})$. given $u \in L^2(\mathbb{R})$, such that,

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ix\xi} \hat{u}(\xi) d\xi. \tag{8}$$

Tu is defined as,

$$Tu_m = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{imh\xi} \hat{u}(\xi) d\xi, \tag{9}$$

for each gridpoint $mh \in h\mathbb{Z}$.

When we consider the discrete fourier transform of Tu,

$$\widehat{Tu}(\xi) = \hat{u}(\xi) \qquad\qquad |\xi| \leq \frac{\pi}{h}$$

Definition 6. The interpolation operator S maps functions from $L^2(h\mathbb{Z})$ to functions in $L^2(\mathbb{R})$. Given $v \in L^2(h\mathbb{Z})$, with

$$v_m = \frac{1}{\sqrt{2\pi}} \int_{\frac{-\pi}{h}}^{\frac{\pi}{h}} e^{imh\xi} \hat{v}(x)$$

then Sv is defined as,

$$Sv(x) = \frac{1}{\sqrt{2\pi}} \int_{\frac{-\pi}{h}}^{\frac{\pi}{h}} e^{ix\xi} \hat{v}(x)$$

These operators are also defined in terms of the grid spacing, h. though this is left out of the notation. Using these two operators we can then define convergence in terms of the interpolation operator.

Definition 7. A finite difference scheme approximating the homogeneous initial value problem for a partial differential equation is a convergent scheme if Sv^n converges to $u(t_n, \cdot)$ in $L^2(\mathbb{R})$, where $t_n = nk$, for every solution u to the differential equation and every set of solutions to the difference scheme v. depending on h, and k, for which Sv^0 converges to $u(0, \cdot)$ in $L^2(\mathbb{R})$ as h and k tend to 0

3 Nonlinear Convergence

The convergence theory in the previous section related to linear problems rather than nonlinear. For the equation of interest we have a nonlinear problem which requires more specific techniques to deal with. A simple theorem cannot provide convergence for all types of nonlinear problems. In this section we look at a convergence theorem for a general parabolic function on a 1-dimensional domain with direchlet conditions. The convergence results presented are from Reynolds.

Consider the Partial differential equation given by,

$$u_t = f(t, x, u, u_x, u_x x) \quad \text{in} \quad [0, T] \times (a, b), \tag{10}$$

$$u(0, x) = \psi(x), \quad u(t, a) = \psi_0(t), \quad \text{and} \quad u(t, b) = \psi_1(t), \tag{11}$$

where $\psi(a) = \psi_0(0)$ and $\psi(b) = \psi_1(0)$.

We consider a discrete mesh over the domain $[0, T] \times (a, b)$ defined by the set of points (t_i, x_j) such that,

$$\begin{aligned} h &= \frac{b-a}{m+1} \\ x_j &= a + ih \\ \Delta t &= T/n \\ t_i &= i\Delta t \end{aligned}$$

for $i = 1, 2, ..., n$ and j=$0, 1, ..., m+1$.
We can also define functions on this mesh. Let $v(t, x)$, be a functioned defined on the domain $[0, T] \times (a, b)$. Then the function defined on the mesh is denoted v_j^i, with $v_j^i = v(t_i, x_j)$, for $i = 1, 2, ..., n$ and j=$0, 1, ..., m+1$. The notation used here is to be consistent with the previous sections.
We use the finite differences that were defined in section 1. to replace our derivatives. We shall also simplify our notation by denoting $f(t_i, x_j, v_j^i, \delta_0 v_j^i, \delta^2 v_j^i)$ by $f(v_j^i)$, where the differences are taken in space. For $i = 1, 2, ..., n$ and j=$0, 1, ..., m+1$ let $\alpha_j(t_i)$ and $\beta_j(t_i)$ be defined by $u_x(t_i, x_j) = \delta_0(u_j^i) + \alpha_j(t_i)$ and $u_x x(t_i, x_j) = \delta^2 u_j^i + \beta_j(t_i)$, where we have assumed that u(t,x) is the unique solution to 10-11. assuming that u has sufficient regularity then then α_j and β_j are $\mathcal{O}(h^2)$. By replacing the derivatives in equation 10, we can consider the discrete problem given by,

$$\delta_{t-} v_j^i = \theta f(v_j^i) + (1-\theta) f(v_j^{i-1}) \quad \text{for } 1 \le i \le n \text{ and } 1 \le j \le m, \tag{12}$$

$$v_0^i = \psi_0(t_i), \quad v_{n+1}^i = \psi_1(t_i) \quad \text{and} \quad v_j^0 = \psi(x_j). \tag{13}$$

We now have enough information to state the main convergence result in the paper.

Theorem 2. *let u denote the unique solution to equatios 10 and 11, have bounded derivatives, u_{tt} and u_{xxxx}.*
Suppose there exists such constants $\alpha, A, B \ge 0$ and constants C and C' such that f satisfies,

$$\alpha\,(\bar{r} - r) \le f(t, x, z, p, \bar{r}) - f(t, x, z, p, r) \le A(\bar{r} - r), \quad \bar{r} \ge r, \tag{14}$$

$$|f(t, x, z, \bar{p}, r) - f(t, x, z, p, r)| \le B|\bar{p} - p|, \tag{15}$$

$$-C\,(\bar{z} - z) \le f(t, x, \bar{z}, p, r) - f(t, x, z, p, r) \le C'(\bar{z} - z). \tag{16}$$

Suppose the grid parameters, Δt and h, satisfy,

$$\theta \Delta t C' \leq 1, \tag{17}$$

$$\alpha - \frac{hB}{2} \geq 0, \tag{18}$$

$$(1-\theta)\, 2 \left(A \frac{\Delta t}{h^2} \right) \leq 1 + (1-\theta)\,(-C)\,. \tag{19}$$

Suppose there exists a function ω defined on $[0,T] \times \mathbb{R}^3$ and a function $\rho > 0$, $\rho \in C^1([0,T])$, such that,

$$f(t,x,\bar{z},\bar{p},\bar{r}) - f(t,x,z,p,r) \leq \omega(t, \bar{z}-z, |\bar{p}-p|, |\bar{r}-r|) \quad \textit{for } \bar{z} \geq z, \tag{20}$$

$$\rho'(t') \geq 2\omega(\bar{t}, \rho(\bar{t}), |\alpha_j|, |\beta_j|) \textit{ on } [0,T] \text{ for } i = 1,2,...,n \text{ and } |\bar{t}-t'| \geq \Delta t, \tag{21}$$

$$\rho'(t') > \Delta t \sup_{0<\bar{t}\leq T, 1\leq j\leq m} |u_{tt}(\bar{t}, x_j)| \quad \text{for} |\bar{t}-t| \leq \Delta t. \tag{22}$$

Then, $\sup_{0\leq j\leq m+1} |(u(t_i,x_j) - v_j^i| \leq \rho(t_i)$ for $i = 0,1,...,n$, where v_j^i is the solution to 12

Whilst this theorem doesn't explicitly state that the equations converge, it provides an upper bound on the difference at each node. As a result if the solution oscillates on a much finer mesh,this detail is not necessarilly captured. What is not stated in this theorem is how the function ρ behaves. It is possible to construct a function ρ that is $\mathcal{O}(\Delta t) + \$\mathcal{O}(h^2)$ which provides convergence as $\Delta t, h$ approach 0.
The paper also provides a method of construction for the functions ω and ρ such that conditions 17-22 are sufficient to provide convergence.