Scientific Computing

A practical Companion

4th Notebook

This notebook can be downloaded from the location:
http : //
staff.science.uva.nl / ~walter / SC / Notebooks / SC08 – 4. nb

Author:

Walter Hoffmann (Korteweg-de Vries Institute for Mathematics, UvA)

January - February, 2008

Solving partial differential equations

Discrete form of a PDE

We discuss the problem of solving a Partial Differential Equation (or shortly pde) by numerical means. One way to tackle the problem is to start with a discrete approximate form of the pde and then solve this discrete equation of which we might try to prove that its solution is close to the solution of the original continuous problem.

Solving the discrete equation can be done in various ways; we concentrate on a method which is well suited for parallelization.

Assume that the variables of the pde that we want to solve are time (t) and space (x). In general, x will be a vector containing the spatial coordinates. For simplicity we consider a problem with only one space variable, so for the time being, x will be a scalar variable.

Assume that x has values between bounds a and b: $a \leq x \leq b$ and that we have $t \geq 0$.
Choose an integer n and construct a devision of [a, b] in n subintervals of length $h = \frac{b-a}{n}$.

This defines gridpoints

$$x_0 (= a),\ x_1,\ \ldots,\ x_{n-1},\ x_n\ (= b);$$

$$x_{i+1} - x_i = h;\ i = 0,\ 1,\ \ldots,\ n\ .$$

Define time-steps of length Δt (or k or τ) ;

then we can speak of discrete time points t_j, t_{j+1}, ... with $t_{j+1} = t_j + \Delta t$ for $j = 0, 1, 2, \ldots$,

Having introduced x_i and t_j we can consider a two-dimensional grid of points (x_i, t_j) for all applicable i and j .

Assume that in the original problem a function $f = f(x, t)$ is to be found and that it is defined as the solution of a partial differential equation over the domain $a \leq x \leq b;\ t \geq 0$.

The continuous function f is approximated by a grid-function u which is defined on the points of the grid that we have introduced. The value of u in the point (x_i, t_j) is considered as an unknown in a set of equations to be constructed. All instances of partial derivatives with respect to t or x are replaced by discrete approximations in the gridpoints. Then the original pde can be stated in its discrete form. For that purpose we must be able to approximate partial derivatives by (discrete) divided differences; this is explained in the next section.

Consider the interval [a, b] in the space-domain with subdivision $x_0 (= a),\ x_1,\ \ldots,\ x_{n-1},\ x_n\ (= b)$.

For the simplicity we do not consider variations in time, so the gridfunction u is considered as a function of space only. The value of u in the point x_i is denoted by u_i .

From Taylor's expansion in neighbouring points we find:

$$u_{i+1} = u_i + h\,u'_i + \frac{1}{2}h^2\,u''_i + O\left(h^3\right)$$
$$u_{i-1} = u_i - h\,u'_i + \frac{1}{2}h^2\,u''_i + O\left(h^3\right)$$

Adding these two equations yields:

$$u_{i+1} + u_{i-1} = 2\,u_i + h^2\,u''_i + O\left(h^3\right)$$

from which we find

$$u''_i = \frac{u_{i+1} - 2\,u_i + u_{i-1}}{h^2} + O(h).$$

In fact the approximation of u''_i is more accurate then $O(h)$ which follows from the symmetry in Taylor's expansion for u_{i+1} and u_{i-1}. It is revealed when we consider Taylor's expansion for some more terms:

$$u_{i+1} = u_i + h\,u'_i + \frac{1}{2}h^2\,u''_i + \frac{1}{6}h^3\,u'''_i + \frac{1}{24}h^4\,u''''_i + O\left(h^5\right)$$
$$u_{i-1} = u_i - h\,u'_i + \frac{1}{2}h^2\,u''_i - \frac{1}{6}h^3\,u'''_i + \frac{1}{24}h^4\,u''''_i + O\left(h^5\right)$$

Now addition yields:

$$u_{i+1} + u_{i-1} = 2\,u_i + h^2\,u''_i + \frac{1}{12}h^4\,u''''_i + O\left(h^5\right)$$

from which we find

$$u''_i = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + O(h^2).$$

The above method for solving partial differential equations is applied to the Heat Equation and is explained in the sequel.

Numerical solution of the Heat equation

Consider the time-dependent diffusion equation (which is called the heat equation)

$$\frac{\partial u}{\partial t} = D \nabla^2 u,$$

with appropriate boundary and initial conditions.

We only consider the equation in one space variable:

$$u_t = D u_{xx}$$

Discretization.

Uniform spatial discretization: $\Delta x = h$.

Uniform temporal discretization: $\Delta t = k$.

$$u(x,t) \Longrightarrow u(x_i, T_j) = u(\text{ih}, \text{jk}) := u_i^j$$

$$\frac{\partial}{\partial t} u_i^j = ?$$

Various choices for approximation.

The use of forward differences in time (forward Euler) gives

$$\frac{\partial}{\partial t} u_i^j = \frac{1}{k}\left(u_i^{j+1} - u_i^j\right) + O(k).$$

Using backward differences in time (backward Euler) gives

$$\frac{\partial}{\partial t} u_i^j = \frac{1}{k}\left(u_i^j - u_i^{j-1}\right) + O(k).$$

Other choices are possible;

for instance:

$$\frac{\partial}{\partial t} u_i^j = \frac{1}{2k}\left(u_i^{j+1} - u_i^{j-1}\right) + O(k^2).$$

For the spatial discretization we chose approximation on basis of central differences for the second derivative which yields

$$\frac{\partial^2}{\partial x^2} u_i{}^j = \frac{1}{h^2}\left(u_{i-1}{}^j - 2\,u_i{}^j + u_{i+1}{}^j\right) + O\left(h^2\right).$$

Explicit method

Derivation of the computational scheme:

Discretized equation; forward in time, central in space:

$$\frac{1}{k}\left(u_i{}^{j+1} - u_i{}^j\right) = \frac{D}{h^2}\left(u_{i-1}{}^j - 2\,u_i{}^j + u_{i+1}{}^j\right).$$

Define:

$$\mu := D\,\frac{k}{h^2}\ \text{(Courant}^{*}\text{ number or CFL – number (Courant, Friedrich, Levy))}.$$

Computational scheme:

$$\begin{aligned} u_i{}^{j+1} &= u_i{}^j + \mu\left(u_{i-1}{}^j - 2\,u_i{}^j + u_{i+1}{}^j\right) \\ &= \mu\,u_{i-1}{}^j + (1 - 2\,\mu)\,u_i{}^j + \mu\,u_{i+1}{}^j\,. \end{aligned}$$

* Richard Courant 1888 - 1972

Optimizing the error:

For the temporal discretization we have:

$$\begin{aligned} u_t\,(x_i,\,T_j) &= \frac{u\,(x_i,\,T_{j+1}) - u\,(x_i,\,T_j)}{k} + O\,(k) \\ &= \frac{1}{k}\left(u_i{}^{j+1} - u_i{}^j\right) + O\,(k)\,. \end{aligned}$$

And for the spatial discretization we have:

$$u_{xx}(x_i, T_j) = \frac{1}{h^2}\left(u_{i-1}{}^j - 2\,u_i{}^j + u_{i+1}{}^j\right) + O\left(h^2\right).$$

Combining these we find for the discretized equation:

$$\frac{1}{k}\left(u_i{}^{j+1} - u_i{}^j\right) + O(k) = \frac{D}{h^2}\left(u_{i-1}{}^j - 2\,u_i{}^j + u_{i+1}{}^j\right) + O\left(h^2\right).$$

A closer look at Taylor's expansion, with respect to variation in time and also with respect to variation in space, gives better insight in the approximation error.

For variation in time we have:

$$u^{j+1} = u^j + k\,u_t{}^j + \frac{1}{2}k^2\,u_{tt}{}^j + O\left(k^3\right).$$

And from this we find

$$u_t{}^j = \frac{1}{k}\left(u_i{}^{j+1} - u_i{}^j\right) - \frac{k}{2}\,u_{tt}{}^j + O\left(k^2\right).$$

For variation in space we found already in the introductory part of this notebook:

$$u_{i+1} + u_{i-1} = 2\,u_i + h^2\,u''_i + \frac{1}{12}h^4\,u''''_i + O\left(h^5\right)$$

which results in:

$$u''_i = \frac{1}{h^2}\left(u_{i+1} - 2\,u_i + u_{i-1}\right) - \frac{1}{12}h^2\,u''''_i + O\left(h^3\right)$$

and the discretized equation with a more accurate approximation error reads:

$$\frac{1}{k}\left(u_i{}^{j+1} - u_i{}^j\right) = \frac{D}{h^2}\left(u_{i-1}{}^j - 2\,u_i{}^j + u_{i+1}{}^j\right) - \frac{D}{12}h^2\,u_{xxxx} + \frac{k}{2}\,u_{tt} + \text{'higher order terms'}.$$

We are solving a discrete version of the original equation:

$$u_t = D\,u_{xx}\,.$$

For a solution of the continuous problem it can analytically be proven that:

$$u_{tt} = D^2\,u_{xxxx}\,.$$

Proof:

$$u_{tt} = D\,u_{xxt}\,.$$
$$u_{txx} = D\,u_{xxxx}\,.$$
$$u_{xxt} = u_{txx}\,.$$
$$\Longrightarrow u_{tt} = D\,u_{xxt} = D\,u_{txx} = D^2\,u_{xxxx}$$

And therefore the approximate solution can be expected to be more accurate if the latter equation is satisfied. If we substitute $u_{tt} = D^2\,u_{xxxx}$ in the 'leading error term' of the discrete equation we find:

$$-\frac{D}{12} h^2\,u_{xxxx} + \frac{k}{2}\,u_{tt} =$$
$$-\frac{D}{12} h^2\,u_{xxxx} + \frac{k}{2}\,D^2\,u_{xxxx} =$$
$$-\frac{D}{2} h^2\,u_{xxxx}\left(\frac{1}{6} - \frac{D\,k}{h^2}\right).$$

So the optimal choice will be:

$$\mu = D\,\frac{k}{h^2} = \frac{1}{6}.$$

Accuracy and amount of work:

Assume that for a more accurate approximation the spatial discretization is divided in 2; 2 times as many points are lying on a grid line. If the approximation is of order 2, this will have the effect that the error is divided by 4 .

Assume that the Courant number

$$\mu\left(= D\,\frac{k}{h^2}\right)$$

will be kept fixed. To realize this, the temporal discretization should be divided by 4. This means 4 times as many grid lines in the temporal domain. Each having twice as many grid points, giving 8 times as much work (just for the 1-D problem)!

Conclusion: 4 times as accurate for 8 times the amount of work.

Stability of the explicit method:

Consider vector notation:

$$u^j = \left(u_0{}^j, u_1{}^j, \ldots, u_{n-1}{}^j, u_n{}^j\right)^T$$

then:

$$u^{j+1} = A\, u^j$$

where A is a tridiagonal matrix of which the first four rows are given by:

$$\begin{bmatrix} 1-2\mu & \mu & & & \\ \mu & 1-2\mu & \mu & & \\ & \mu & 1-2\mu & \mu & \\ & & \mu & 1-2\mu & \mu \end{bmatrix}$$

Assume:

$$\mathrm{fl}\left(u^k\right) = u^k + e$$

then

$$\begin{aligned} A\left(u^k + e\right) &= A\, u^k + A\, e \\ &= u^{k+1} + A\, e \end{aligned}$$

and after m steps we arrive at

$$u^{k+m} + A^m\, e\,.$$

The method is stable if the error does diminish and in the limit approaches 0:

$$\left\|A^m\right\| \to 0 \quad .$$

This only holds if $\rho(A)$ (spectral radius of A) is not greater than one.

For a symmetrical matrix, the spectral radius is bounded by the largest eigenvalue. According to Gerschgorin's first theorem we have:

$$(1-2\mu) - 2\mu \le \lambda \le (1-2\mu) + 2\mu\,.$$

▿ Theorem (Gerschgorin's first theorem):

Let A be an n×n matrix with eigenvalues λ_i . The eigenvalues of A

are contained in the union of the discs D_i defined by

$$D_i = \left\{ z \in \mathbb{C} \;\middle|\; |z - a_{ii}| \leqslant \sum_{j=1, j\neq i}^{n} |a_{ij}| \right\}.$$

Proof:

Let λ be any eigenvalue of A and $A\,x = \lambda\,x$.

Assume that x has been scaled that $\|x\|_\infty = 1$.

Let i be an index such that $|x_i| = 1$.

Then $\lambda\,x_i = \sum_{j=1}^{n} a_{ij}\,x_j$

$(\lambda - a_{ii})\,x_i = \sum_{j\neq i} a_{ij}\,x_j$

$|\lambda - a_{ii}| \leqslant |\sum_{j\neq i} a_{ij}\,x_j| \leqslant \sum_{j\neq i} |a_{ij}||x_j| \leqslant \sum_{j\neq i} |a_{ij}|$.

Therefore λ is contained in D_i

So, for all eigenvalues to lie between -1 and +1 we should have

$$1 - 4\mu \geqslant -1$$

leading to the well known stability condition for the explicit scheme:

$$\mu \leq \frac{1}{2}$$

NOTA BENE :

The bound $\mu \leq \frac{1}{2}$ (giving $\Delta t \leq \frac{h^2}{2D}$) is derived for a single space dimension.

In two dimensions the matrix is different which then leads to $\mu \leq \frac{1}{4}$ (giving $\Delta t \leq \frac{h^2}{4D}$)

(In our analysis we used k for Δ t)

Implicit method

Derivation of the computational scheme:

Discretized equation: backward in time, central in space:

$$\frac{1}{k}\left(u_i^{\,j} - u_i^{\,j-1}\right) = \frac{D}{h^2}\left(u_{i-1}^{\,j} - 2\,u_i^{\,j} + u_{i+1}^{\,j}\right)$$

With use of the Courant number:

$$\mu = D \frac{k}{h^2}$$

this yields:

$$-\mu\, u_{i-1}{}^{j} + (1 + 2\mu)\, u_i{}^{j} - \mu\, u_{i+1}{}^{j} = u_i{}^{j-1} \quad .$$

The values at the new level j are implicitly given. To get their numerical values, a linear system has to be solved. For the 1-D case that we are treating here, this is a system with a tridiagonal coefficient matrix. In matrix notation the system is given by

$$T\, u^j = u^{j-1}$$

where the four first rows of matrix T are given by

$$\begin{bmatrix} 1+2\mu & -\mu & & & & \\ -\mu & 1+2\mu & -\mu & & & \\ & -\mu & 1+2\mu & -\mu & & \\ & & -\mu & 1+2\mu & -\mu & \end{bmatrix}$$

Stability of the implicit method:

As in the stability analysis of the explicit scheme assume:

$$\mathrm{fl}\left(u^k\right) = u^k + e$$

Although u^{k+1} is calculated by solving a linear system, we may write

$$\begin{aligned} u^{k+1} &= T^{-1}\left(u^k + e\right) \\ &= u^{k+1} + T^{-1}\, e \end{aligned}$$

and after m steps we arrive at

$$u^{k+m} + T^{-m}\, e \;.$$

The method is stable if the error does diminish and in the limit approaches 0.

$$\left\| T^{-m} \right\| \to 0 \quad .$$

This only holds if $\rho\,(\mathbf{T}^{-1})$ (spectral radius of $\mathbf{T}^{-1}$) is not greater than one, or alternatively, if for the eigenvalues of T we do have $|\lambda| \geq 1$.

According to Gerschgorin's first theorem we have for the eigenvalues of T:

$$(1 + 2\mu) - 2\mu \leq \lambda \leq (1 + 2\mu) + 2\mu$$

Which says $|\lambda| \geq 1$, independent of μ .

From this we conclude that the implicit method is unconditionally stable. The size of μ only determines the accuracy of the approximate solution!

Optimizing the error:

As a final remark we mention that the analysis which told us that the relation

$$k = \frac{1}{6} h^2$$

gives an approximation with a higher order of accuracy for the explicit scheme, is also valid for the implicit scheme that we are treating here.

Crank-Nicolson method

Derivation of the computational scheme:

To realize a higher order of approximation for the time derivative, we make the following observation.

Consider the forward difference formula for the time derivative.

$$\frac{\partial}{\partial t} u_i{}^j = \frac{1}{k} \left(u_i{}^{j+1} - u_i{}^j\right) + O(k).$$

As we know, this formula has first order of accuracy when it is used as an approximation of the first derivative in point $u_i{}^j$

The same difference quotient acts as approximation for the first derivative in time in point $u_i{}^{j+\frac{1}{2}}$ with accuracy $O(k^2)$.

For this same point $u_i{}^{j+\frac{1}{2}}$ we then also consider the second derivative in space.

We take that approximation as the average of the central difference approximations for the points $u_i{}^{j+1}$ and $u_i{}^j$ yielding:

$$\frac{1}{2} \frac{D}{h^2} \left(\left(u_{i-1}{}^{j+1} - 2 u_i{}^{j+1} + u_{i+1}{}^{j+1}\right) + \left(u_{i-1}{}^j - 2 u_i{}^j + u_{i+1}{}^j\right) \right)$$

Using again $\mu = D \frac{k}{h^2}$, the discrete analogon of the heat equation (stated for the fictitious point $u_i^{j+\frac{1}{2}}$) becomes

$$u_i^{j+1} - u_i^j = \frac{1}{2} \mu \left(\left(u_{i-1}^{j+1} - 2 u_i^{j+1} + u_{i+1}^{j+1} \right) + \left(u_{i-1}^j - 2 u_i^j + u_{i+1}^j \right) \right)$$

After multiplying left and right hand side by 2 this yields:

$$-\mu u_{i-1}^{j+1} + (2 + 2\mu) u_i^{j+1} - \mu u_{i+1}^{j+1} = \mu u_{i-1}^j + (2 - 2\mu) u_i^j + \mu u_{i+1}^j$$

The values at the new level j+1 are implicitly given. They follow from the solution of a linear system with a tridiagonal coefficient matrix. The right hand side has to be calculated as the matrix product of a tridiagonal matrix and the known vector u^j . This scheme is known as the Crank-Nicolson* scheme.

* John Crank 1916

Phyllis Nicolson 1917 - 1968

Stability and accuracy of the Crank-Nicolson method:

This method has better accuracy than both the explicit and the implicit method and like the implicit method, is unconditional stable .

Close all sections