Uppsala University Department of Information Technology Division of Scientific Computing

Solutions to the Examination in Analysis of Numerical Methods

2003-10-08

1. a) Taylor expansion of the exact solution u(x, t) around (x_j, t^{n+1}) is easier than around (x_j, t^n) , which is also possible, of course.

$$u(x_j, t_n) = u - ku_t + k^2/2 u_{tt} + \mathcal{O}(k^3)$$

$$u(x_{j\pm 1}, t_{n+1}) = u \pm hu_x + h^2/2 u_{xx} \pm h^3/6 u_{xxx} + \mathcal{O}(k^4)$$

Inserting $u = u(x_j, t_{n+1})$ for v_j^{n+1} into the left hand side of the FDM

$$P_{k,h}v_j^{n+1} = \frac{v_j^{n+1} - v_j^n}{k} + a\frac{v_{j+1}^{n+1} - v_{j-1}^{n+1}}{2h} - bv_j^{n+1} = 0$$
(1)

and using the PDE $u_t + au_x - bu = 0$ yields the truncation error

$$P_{k,h}u = u_t - k/2 u_{tt} + \mathcal{O}(k^2) + a(u_x + h^2/6 u_{xxx} + \mathcal{O}(h^4)) - bu$$

= $-k/2 u_{tt} + \mathcal{O}(k^2) + a(h^2/6 u_{xxx} + \mathcal{O}(h^4))$
= $\mathcal{O}(k) + \mathcal{O}(h^2)$

Thus, the FDM is first order accurate in time and second order accurate in space, i.e. the FDM is consistent.

We perform the von Neumann stability analysis by inserting the ansatz

 $v_j^n = g^n e^{i\omega x_j}$

into the FDM (1). We get

$$\left[\left(1+\frac{\lambda a}{2}(e^{i\omega h}-e^{-i\omega h})-kb\right)g-1\right]\frac{g^n}{k}e^{i\omega x_j}=0.$$
(2)

After dividing and using $e^{i\omega h} - e^{-i\omega h} = 2isin(\omega h)$, we obtain the amplification factor

$$g = \frac{1}{1 - kb + i\lambda asin(\omega h)} \; .$$

We note that

$$|g|^{2} = \frac{1}{(1-kb)^{2} + (\lambda asin(\omega h))^{2}} \le 1$$

because the denominator is larger equal 1 as $b \leq 0$ for all k > 0. Thus, the FDM is unconditionally stable.

The Lax-Richtmyer equivalence theorem states that for a consistent FDM stability and convergence are equivalent. We have shown that the FDM (1) is consistent and unconditionally stable. Thus, it is convergent.

b) In the limit $k \to \infty$, we get

$$|g| = \frac{1}{\sqrt{(1-kb)^2 + (\lambda asin(\omega h))^2}} \longrightarrow \begin{cases} 1 & \text{if } b = 0 \text{ and } \omega h = 0 \text{ or } \pi \\ 0 & \text{otherwise,} \end{cases}$$

Thus, all wave numbers ω , except for the lowest and largest for b = 0, are optimally damped for $k \to \infty$. For $k = \infty$ the time difference in FDM (1) disappears and we get the direct method for the steady state problem, i.e. the steady state solution is obtained in one step.

2. We perform the GKS analysis. The stability of FDM (1) for periodic boundary conditions was shown in task 1.

1. Resolvent equation

Inserting the ansatz $v_j^n = z^n \tilde{v}_j$ into the FDM (1) for $b \equiv 0$ yields

$$\left[(z-1)\tilde{v}_j + z\frac{\lambda a}{2}(\tilde{v}_{j+1} - \tilde{v}_{j-1}) \right] \frac{z^n}{k} = 0$$

For $z \neq 0$, we get the resolvent equation

$$(z-1)\tilde{v}_j + z\frac{\lambda a}{2}(\tilde{v}_{j+1} - \tilde{v}_{j-1}) = 0.$$
(3)

2. Characteristic equation

Inserting the ansatz $\tilde{v}_i = \kappa^j$ into the resolvent equation (3), we obtain

$$(z-1)\kappa^{j} + z\frac{\lambda a}{2}(\kappa^{j+1} - \kappa^{j-1}) = 0.$$

Dividing by $z \frac{\lambda a}{2} \kappa^{j-1}$ and ordering, we get the characteristic equation

$$\kappa^2 + \frac{z-1}{z} \frac{2}{\lambda a} \kappa - 1 = 0 \tag{4}$$

3. Determinant condition

The general solution of the resolvent equation (3) is

$$\tilde{v}_j = \begin{cases} \sigma_1 \kappa_1^j + \sigma_2 \kappa_2^j & \text{if } \kappa_1 \neq \kappa_2 \\ \sigma_1 \kappa^j + \sigma_2 j \kappa^{j-1} & \text{if } \kappa_1 = \kappa_2 = \kappa \end{cases}$$

where κ_1 and κ_2 are the two roots of the characteristic equation (4). The coefficients σ_1 and σ_2 are determined such that the boundary condition $v_0^{n+1} = 2v_1^{n+1} - v_2^{n+1}$ and $v^n \in l_2(0, \infty)$ are satisfied for the right half plane problem. (Similar reasoning is used for the left half plane problem with the boundary condition $v_N^{n+1} = 0$, cf. below.) Because of the form of the characteristic equation (4), the product of its roots satisfies the relation

$$\kappa_1\kappa_2=-1.$$

Thus, either $|\kappa_1| < 1$ and $|\kappa_2| > 1$ or $|\kappa_1| = |\kappa_2| = 1$. In either case, we have to set $\sigma_2 = 0$ to secure $v^n \in l_2(0, \infty)$. Now, we check $|\kappa_1| = 1$. Inserting $\kappa = e^{i\xi}$ into the characteristic equation (4), we get for z the same result as for the amplification factor g in the von

Neumann stability analysis (cf. task 1), i.e. $|z| \leq 1$. Thus, there can only be solutions with |z| > 1 for $|\kappa_1| < 1$ and $|\kappa_2| > 1$, and the solution must be of the form

$$v_i^n = z^n \sigma_1 \kappa_1^j$$
.

(Inserting the solution into the boundary condition $v_N^{n+1} = 0$ yields $z^{n+1}\sigma_1\kappa_1^N = 0$. Thus, $\kappa_1 = 0$ for $\sigma_1 \neq 0$, and consequently $v_j^n = 0$. Therefore, the left half plane problem is stable.)

Inserting the solution into the boundary condition $v_0^{n+1} = 2v_1^{n+1} - v_2^{n+1}$, we get

$$z^{n+1}\sigma_1 = z^{n+1}\sigma_1(2\kappa_1 - \kappa_1^2)$$

For $\sigma_1 \neq 0$, we obtain

$$\kappa_1^2 - 2\kappa_1 + 1 = 0 \; .$$

Therefore, the determinant condition is

$$\kappa_1 = 1 . (5)$$

4. Solve equations

We solve the characteristic equation (4) and the determinant condition (5) by inserting $\kappa = 1$ into (4). We get $\frac{z-1}{z} \frac{2}{\lambda_a} = 0$, i.e. z = 1. Thus, the solution is

$$\kappa_1 = 1 , z = 1.$$
(6)

5. Check solutions

If $\kappa = 1$ and z = 1 were a solution in the limit $\kappa \longrightarrow 1_{-}$ and $z \longrightarrow 1_{+}$, the scheme would be unstable. To check that case, we assume that $z = 1 + \delta$ with $\delta > 0$ and $\kappa = 1 + \epsilon$. We insert z and κ into the characteristic equation (4) and check the sign of ϵ as $\delta \longrightarrow 0$. If $\epsilon < 0$, the scheme is unstable.

Inserting z and κ into (4), yields

$$(1+\epsilon)^2 + \frac{1+\delta-1}{1+\delta} \frac{2}{\lambda a} (1+\epsilon) - 1 = 0$$

Using $\frac{1}{1+\delta} = 1 - \delta + \mathcal{O}(\delta^2)$ and neglecting the quadratic and higher order terms, we get

$$\epsilon = -\frac{1}{\lambda a}\delta \; .$$

Thus, $\epsilon > 0$, because $\delta > 0$ and a < 0. As $\epsilon > 0$, the scheme is stable, because then $\kappa = \kappa_2$. But since $\sigma_2 = 0$, there is no critical solution with $z \longrightarrow 1_+$.

6. Conclusions

The FDM (1) with the boundary conditions stated above is unconditionally stable.

3. a)

$$\mathbf{A} = \left(\begin{array}{cc} u & c \\ c & u \end{array}\right)$$

Determine eigenvalues λ of **A**:

$$det(\mathbf{A} - \lambda \mathbf{I}) = (u - \lambda)^2 - c^2 = 0 \iff \lambda = u \pm c$$

Thus, the eigenvalues $\lambda_1 = u - c$ and $\lambda_2 = u + c$ are real. Determine corresponding eigenvectors \mathbf{r}_1 and \mathbf{r}_2 , i.e. $(\mathbf{A} - \lambda \mathbf{I})\mathbf{r} = 0$. We obtain after normalizing:

$$\mathbf{r}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
, $\mathbf{r}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Define the transformation matrix as the right eigenvector matrix

$$\mathbf{T} = [\mathbf{r}_1, \mathbf{r}_2] = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1\\ 1 & 1 \end{pmatrix}$$

T is symmetric and orthogonal, i.e. $\mathbf{T}^{-1} = \mathbf{T}$. Therefore,

$$\mathbf{T}^{-1}\mathbf{A}\mathbf{T}=\mathbf{\Lambda}\;,$$

where

$$\mathbf{\Lambda} = \left(\begin{array}{cc} \lambda_1 & 0\\ 0 & \lambda_2 \end{array}\right)$$

i.e. A is diagonalizable.

Summarizing, the first order system

$$\mathbf{V}_t + \mathbf{A} \mathbf{V}_x = 0 , \qquad (7)$$

(8)

where

$$\mathbf{V} = \left(\begin{array}{c} v\\p \end{array}\right)$$

is a hyperbolic system.

b) Multiplying the hyperbolic system (7) by \mathbf{T}^{-1} from the left, using the diagonalization of \mathbf{A} and the definition of the characteristic variables

$$\mathbf{W} = \mathbf{T}^{-1}\mathbf{V} = \frac{1}{\sqrt{2}} \begin{pmatrix} p - v \\ p + v \end{pmatrix} = \begin{pmatrix} w^1 \\ w^2 \end{pmatrix} ,$$
$$\mathbf{W}_t + \mathbf{\Lambda} \mathbf{W}_x = 0 ,$$

we get

i.e. two scalar equations

which correspond to

$$\frac{dw^1}{dt} = 0 \qquad \text{on} \qquad \frac{dx}{dt} = \lambda_1$$
$$\frac{dw^2}{dt} = 0 \qquad \text{on} \qquad \frac{dx}{dt} = \lambda_2 \;.$$

Thus, w^1 and consequently p - v is constant on characteristics $\frac{dx}{dt} = u - c$, and w^2 and consequently p + v is constant on characteristics $\frac{dx}{dt} = u + c$.

c) The leapfrog scheme for solving the hyperbolic system reads

$$\mathbf{V}_{j}^{n+1} = \mathbf{V}_{j}^{n-1} - \lambda \mathbf{A} (\mathbf{V}_{j+1}^{n} - \mathbf{V}_{j-1}^{n}) , \qquad (9)$$

where $\mathbf{V}_{j}^{n} = \begin{pmatrix} v_{j}^{n} \\ p_{j}^{n} \end{pmatrix}$ approximates $\begin{pmatrix} v(x_{j}, t_{n}) \\ p(x_{j}, t_{n}) \end{pmatrix}$ and $\lambda = \frac{k}{h}$ again. The stability analysis is facilitated by performing it for the character

The stability analysis is facilitated by performing it for the characteristic equations (8), i.e. for a scalar equation of the form $w_t + aw_x = 0$, for which the leapfrog scheme reads:

$$w_j^{n+1} = w_j^{n-1} - \lambda a (w_{j+1}^n - w_{j-1}^n) .$$
(10)

 $w = w^l$, $a = \lambda_l$, l = 1, 2, denotes the first or second component of **W** and the first or second eigenvalue of **A**, respectively.

Applying the von Neumann stability analysis, we insert the ansatz $w_j^n = g^n e^{i\omega x_j}$ into (10) to get

$$g^{n+1}e^{i\omega x_j} = [g^{n-1} - \lambda a(e^{i\omega h} - e^{-i\omega h})g^n]e^{i\omega x_j}$$

With the same arguments as in task 1, we get the quadratic equation for the amplification factor g

$$g^2 + 2i\lambda asin(\omega h)g - 1 = 0$$
.

Its roots are

$$g_{1,2} = -i\lambda asin(\omega h) \pm \sqrt{1 - (\lambda asin(\omega h))^2}$$

We have three cases:

Case 1: $|\lambda a| < 1$

Then, $(\lambda a sin(\omega h))^2 < 1$ and the square root is real and positive. Consequently, $g_1 \neq g_2$. As $|g_{1,2}| = 1$ and the roots are simple, the scheme is then stable.

Case 2: $|\lambda a| = 1$

For $\omega h = \pi/2$, $g_1 = g_2 = -i\lambda a$ and $|g_1| = |g_2| = 1$. As g_1 is a double root, but not $|g_1| < 1$, the scheme is then unstable.

Case 3: $|\lambda a| > 1$

For $\omega h = \pi/2$, either $|g_1| > 1$ or $|g_2| > 1$, since $|g_{1,2}| = |-\lambda a \pm \sqrt{(\lambda a)^2 - 1}|$. Therefore, the scheme is unstable then.

Thus, the leapfrog scheme is stable for $|\lambda a| < 1$. As here $a = u \pm c$ and $max|u \pm c| = |u| + c$, the stability condition becomes

$$\lambda(|u|+c) < 1. \tag{11}$$

4. a) Applying the von Neumann stability analysis, the ansatz $v_j^n = g^n e^{i\omega x_j}$ is inserted into the upwind method. We get

$$g = 1 - a\lambda(1 - e^{-i\omega h}) = 1 - a\lambda(1 - \cos(\omega h) - ia\lambda\sin(\omega h))$$
(12)

Using trigonometric identities, we obtain

$$|g|^2 = 1 - 4a\lambda(1 - a\lambda)\sin^2(\frac{\omega h}{2}) .$$

We see that the necessary and sufficient stability condition

$$|g(\omega h)| \le 1$$

is equivalent to the CFL condition

$$|a\lambda| \le 1$$

A scheme is total variation diminishing (TVD), if

$$\sum_{j=1}^{N} |v_j^{n+1} - v_{j-1}^{n+1}| \le \sum_{j=1}^{N} |v_j^n - v_{j-1}^n| .$$

Inserting the upwind method, we get

$$\begin{split} \sum_{j=1}^{N} |v_{j}^{n+1} - v_{j-1}^{n+1}| &= \sum_{j=1}^{N} |v_{j}^{n} - a\lambda(v_{j}^{n} - v_{j-1}^{n}) - [v_{j-1}^{n} - a\lambda(v_{j-1}^{n} - v_{j-2}^{n})| \\ &= \sum_{j=1}^{N} |(1 - a\lambda)(v_{j}^{n} - v_{j-1}^{n}) + a\lambda(v_{j-1}^{n} - v_{j-2}^{n})| \\ &\leq (1 - a\lambda) \sum_{j=1}^{N} |v_{j}^{n} - v_{j-1}^{n}| + a\lambda \sum_{j=1}^{N} |v_{j-1}^{n} - v_{j-2}^{n}| \\ &= \sum_{j=1}^{N} |v_{j}^{n} - v_{j-1}^{n}| \end{split}$$

where the triangle inequality and $0 \le a\lambda \le 1$ were used for the inequality and the periodic boundary conditions for the last equality.

b) The conservative upwind method for the inviscid Burgers' equation becomes

$$v_j^{n+1} = v_j^n - a\lambda (h_{j+1/2}^n - h_{j-1/2}^n), \qquad (13)$$

where

$$h_{j+1/2}^n = h(v_j^n, v_{j+1}^n) = \begin{cases} (v_j^n)^2/2 & \text{if } a_{j+1/2}^n \ge 0\\ (v_{j+1}^n)^2/2 & \text{if } a_{j+1/2}^n < 0 \end{cases}$$

with $a_{j+1/2}^n = \frac{1}{2}(v_j^n + v_{j+1}^n)$. By definition, the upwind scheme (13) is conservative. It is also consistent, because $h(u, u) = u^2/2$.

Performing one time step with the given initial condition, the solution does not change. Thus, instead of approximating the correct entropy satisfying exact solution, i.e. a rarefaction wave, the upwind scheme (13) yields an entropy violating steady expansion shock. The numerical flux function of the upwind scheme (13) can also be written as

$$h_{j+1/2}^{n} = h(v_{j}^{n}, v_{j+1}^{n}) = \frac{1}{2} \left((v_{j}^{n})^{2}/2 + (v_{j+1}^{n})^{2}/2 - |a_{j+1/2}^{n}|(v_{j+1}^{n} - v_{j}^{n}) \right)$$
(14)

At x = 0, we get $a_{0-1/2} = \frac{1}{2}(v_{-1} + v_0) = \frac{1}{2}(-1 + 1) = 0$. Thus, at the discontinuity, where we need numerical viscosity, we do not get any with the upwind scheme (13). Therefore, Harten (1983) suggested as entropy fix to replace $|a_{j+1/2}^n|$ in (14) by $Q(a_{j+1/2}^n)$, where

$$Q(a) = \begin{cases} |a| & \text{if } |a| \ge \delta\\ (a^2 + \delta^2)/(2\delta) & \text{if } |a| < \delta \end{cases}$$

with e.g. $\delta = max(0, v_{j+1}^n - v_j^n)$ for $Q(a_{j+1/2}^n)$.

5. a) The multigrid method is based on the following two ideas:

(a) Iterative improvement

Suppose $\tilde{\mathbf{u}}$ is an approximation to the exact solution \mathbf{u} to the linear system $\mathbf{A}\mathbf{u} = \mathbf{f}$. Compute the residual $\mathbf{r} = \mathbf{A}\tilde{\mathbf{u}} - \mathbf{f}$ and solve the linear system $\mathbf{A}\mathbf{v} = \mathbf{r}$. Then, correct the approximate solution to get the exact solution

$$\mathbf{u}=\tilde{\mathbf{u}}-\mathbf{v}\;,$$

because $\mathbf{A}\mathbf{u} = \mathbf{A}(\tilde{\mathbf{u}} - \mathbf{v}) = \mathbf{A}\tilde{\mathbf{u}} - \mathbf{A}\mathbf{v} = \mathbf{A}\tilde{\mathbf{u}} - \mathbf{r} = \mathbf{A}\tilde{\mathbf{u}} - (\mathbf{A}\tilde{\mathbf{u}} - \mathbf{f}) = \mathbf{f}$.

Even if $\mathbf{A}\mathbf{v} = \mathbf{r}$ is only solved approximately by $\tilde{\mathbf{v}}$, $\tilde{\mathbf{u}} - \tilde{\mathbf{v}}$ will yield an improved solution. (b) Residual smoothing

The eigenvalues of the the iteration matrix $\mathbf{G}_{J,\omega} = \mathbf{I} - \omega \frac{\hbar^2}{2} \mathbf{A}$ of the damped Jacobi method $\mathbf{u}^{k+1} = \mathbf{S}_1 \mathbf{u}^k = \mathbf{G}_{J,\omega} \mathbf{u}^k + \omega \frac{\hbar^2}{2} \mathbf{f}$ for $\omega = \frac{1}{2}$ are

$$\lambda_{\mu} = 1 - \sin^2(\frac{\mu \pi h}{2})$$
 , $\mu = 1, ..., n$

i.e. the high wave numbers $\mu = \frac{n+1}{2}, ..., n$ are quickly damped by a few damped Jacobi iterations, since $|\lambda_{\mu}| \leq \frac{1}{2}$ for those μ . The low wave numbers $\mu = \frac{n+1}{4}, ..., \frac{n+1}{2} - 1$ can be damped on a coarser grid with $h_0 = 2h$, where they become high wave numbers.

The two-grid multigrid algorithm can be summarized as follows:

$$\begin{split} TMG(u, f) \\ \mathbf{u} &:= \mathbf{S}_{1}^{\nu} \mathbf{u} \\ \mathbf{r} &:= \mathbf{R}(\mathbf{A}\mathbf{u} - \mathbf{f}) \\ \mathbf{v} &:= \mathbf{A}_{0}^{-1} \mathbf{r} \\ \mathbf{u} &:= \mathbf{u} - \mathbf{P} \mathbf{v} \end{split} \qquad \begin{array}{l} \nu \text{ iterations with damped Jacobi on fine grid} \\ residual on fine grid is restricted to coarse grid} \\ \mathbf{A}_{0} \mathbf{v} &= \mathbf{r} \text{ is solved exactly on coarse grid} \\ \mathbf{u} &:= \mathbf{u} - \mathbf{P} \mathbf{v} \end{split}$$

Since only the matrices are relevant for the iteration matrix, we can set \mathbf{f} to zero as for investigating the error $\mathbf{e}^{k+1} = \mathbf{G}\mathbf{e}^k$ to get the TMG iteration matrix. Thus, we get from the TMG algorithm above:

$$\mathbf{e}^{k+1} = (\mathbf{I} - \mathbf{P}\mathbf{A}_0^{-1}\mathbf{R}\mathbf{A})\mathbf{G}_{J,\omega}^{\nu}\mathbf{e}^k$$
.

Therefore, the iteration matrix of the TMG algorithm is

$$\mathbf{G}(\nu) = (\mathbf{I} - \mathbf{P}\mathbf{A}_0^{-1}\mathbf{R}\mathbf{A})\mathbf{G}_{J,\omega}^{\nu}$$

b) If the FDM is multiplied by h^2 , it can be written as

$$-(u_{i+1,j}-2u_{i,j}+u_{i-1,j})-(u_{i,j+1}-2u_{i,j}+u_{i,j-1}) = h^2 f_{i,j} .$$

Thus, the FDM for i = 1, ..., n and j = 1, ..., n can be expressed as the linear system

$$Au = h^2 f$$

where $U = [u_{1,1}, u_{2,1}, ..., u_{n,1}, u_{1,2}, u_{2,2}, u_{n,2}, ..., u_{1,n}, u_{2,n}, ..., u_{n,n}]^T$ and A is the $n \times n$ block-tridiagonal matrix

with the $n \times n$ tridiagonal matrix

and the $n \times n$ diagonal matrix $\boldsymbol{C} = -\boldsymbol{I}_n$.

The direct solution of a linear system with a $N \times N$ band matrix **A** of bandwidth d takes $\mathcal{O}(Nd^2)$ flops with LU decomposition and forward and backward substitution. Since here $N = n^2$ and d = n, the direct solution of the above system takes $\mathcal{O}(n^4)$ flops.

An iterative method like Jacobi or Gauss-Seidel needs $\mathcal{O}(n^2)$ iterations to converge, because the number of iterations k to reach a tolerance tol is

$$k \ge \frac{\log(tol/||\mathbf{e}^0||_2)}{\log(||\mathbf{G}||_2)}$$

and $log(||\mathbf{G}||_2) \approx log(1 - \mathcal{O}(h^2)) \approx -\mathcal{O}(h^2) \approx -\frac{1}{n^2}$. One Jacobi or Gauss-Seidel iteration takes $\mathcal{O}(N)$ flops. Thus, to reach a certain tolerance takes $\mathcal{O}(n^2)\mathcal{O}(N) = \mathcal{O}(n^4)$ flops.

The multigrid method takes $\mathcal{O}(1)$ iterations to converge. According to Achi Brandt, it should not take more than 10 multigrid cycles. Since one multigrid cycle takes $\mathcal{O}(N)$ flops, the multigrid method needs $\mathcal{O}(1)\mathcal{O}(N) = \mathcal{O}(n^2)$ flops.