Jeffrey Hellrung
Numerical Analysis Qualifying Exam, Winter 2005

1. (5 Pts.) Let $f(x) = \cos(x) - x$.

   (a) Prove that $f(x)$ has exactly one root in the interval $\left[0, \frac{\pi}{2}\right]$.

   (b) Give a good estimate of the minimum number of bisection iterations required to obtain an approximation that is within $10^{-6}\left(\frac{\pi}{2}\right)$ of this root when the initial interval used is $\left[0, \frac{\pi}{2}\right]$.

   Solution

   (a) We note that $f(0) = 1$ while $f\left(\frac{\pi}{2}\right) = -\frac{\pi}{2}$, so, since f is continuous, by the Intermediate Value Theorem, there must be some $x^* \in \left(0, \frac{\pi}{2}\right)$ such that $f(x^*) = 0$. To show uniqueness, suppose some other $x' \in \left(0, \frac{\pi}{2}\right) \setminus \{x^*\}$ with $f(x') = 0$. Then the Mean Value Theorem (applicable since f is also continuously differentiable) provides an x between x' and x^* such that $f'(x) = 0$, and indeed $x \in \left(0, \frac{\pi}{2}\right)$. This is impossible, however, since we must have $f'(x) = -1 - \sin(x) < 0$. It follows that x^* is unique.

   (b) One would need $-\log_2 10^{-6} \approx -\log_2 2^{-20} = 20$ bisection iterations to achieve a precision of $10^{-6}\frac{\pi}{2}$.

2. (5 Pts.) Let I_h be the composite trapezoidal rule approximation to the integral $\int_0^1 f(s)ds$ using N panels of size h (i.e., $h = \frac{1}{N}$).

   (a) Give a derivation of the formula that combines I_h and $I_{h/2}$ to obtain an approximation to the integral that is fourth-order accurate.

   (b) When the trapezoidal method is applied to the function $f(x) = x^{3/2}$, the rate of convergence is approximately 1.7. What is the expected rate of convergence when the formula you derived in (a) is applied to $f(x) = x^{3/2}$?

   Solution

   (a) For sufficiently smooth f,
   $$I_h = I + ah^3 + O(h^4),$$
   for some a independent of h, so
   $$I_{h/2} = I + \frac{1}{8}ah^3 + O(h^4),$$
   hence
   $$\frac{8}{7}I_{h/2} - \frac{1}{7}I_h = I + O(h^4).$$

   (b) The supposition is that
   $$I_h \approx I + ah^{1.7},$$
   so the expression in (a) will not cancel this 1.7-order term; i.e., the rate of convergence will remain approximately 1.7.

3. (5 Pts.) Let A be an $n \times n$ non-singular matrix, and consider iterative methods of the form
   $$Mx^{n+1} = b + Nx^n$$
   where $A = M - N$.

(a) Assuming M is non-singular, state a sufficient condition that ensures convergence of the iterates to the solution of $Ax = b$ for any starting vector x^0.

(b) Describe the matrices M and N for (i) Jacobi iteration and (ii) Gauss-Seidel iteration.

(c) If A is strictly diagonally dominant, prove that Jacobi's method converges.

Solution

(a) Since $Ax = b$,

$$M(x^{n+1} - x) = Mx^{n+1} - Mx = b + Nx^n - (A + N)x = N(x^n - x),$$

so that

$$x^{n+1} - x = M^{-1}N(x^n - x).$$

A sufficient condition for convergence is that all eigenvalues of $M^{-1}N$ be strictly less than 1 in magnitude.

(b) For Jacobi iteration, $M = D$ and $N = L + U$, where D, $-L$, and $-U$ are the diagonal, lower triangular, and upper triangular parts of A, respectively. For Gauss-Seidel, $M = D - L$ and $N = U$.

(c) $A = (a_{ij})$ strictly diagonally dominant implies that

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

It follows that

$$\sum_{j \neq i} \left| \frac{a_{ij}}{a_{ii}} \right| < 1,$$

and $-a_{ij}/a_{ii}$ corresponds precisely with the off-diagonal elements of $M^{-1}N = D^{-1}(D - A)$. Now suppose λ is an eigenvalue of $M^{-1}N$ and x a corresponding eigenvector, normalized such that $\max_i |x_i| = 1$. Let x_i be a component of x equal to ± 1. Then from $M^{-1}Nx = \lambda x$, we have

$$\sum_{j \neq i} -\frac{a_{ij}}{a_{ii}} x_j = \lambda x_i = \pm \lambda.$$

But since each $|x_j| \leq 1$, we obtain

$$\begin{aligned} |\lambda| &= \left| \sum_{j \neq i} -\frac{a_{ij}}{a_{ii}} x_j \right| \\ &\leq \sum_{j \neq i} \left| \frac{a_{ij}}{a_{ii}} \right| |x_j| \\ &\leq \sum_{j \neq i} \left| \frac{a_{ij}}{a_{ii}} \right| \\ &< 1, \end{aligned}$$

hence Jacobi's method converges by the observation in (a).

4. Consider the following finite difference scheme for solving $y' = f(y)$:

$$y_{n+1} = y_n + hf\left((1 - \theta)y_n + \theta y_{n+1}\right),$$

where $\theta \in [0, 1]$ is a parameter.

(a) Find the order of the scheme for $\theta \in [0,1]$.

(b) Determine the region of linear stability.

(c) Determine all the values of $\theta \in [0,1]$ for which the method is A-stable.

Solution

(a) We begin by supposing that $y_n = y(t_n)$. Then, applying Taylor's Thereom several times,

$$\begin{aligned}
y_{n+1} &= y_n + hf\left((1-\theta)y_n + \theta y_{n+1}\right) \\
&= y_n + h\left(f\left((1-\theta)y_n + \theta y(t_{n+1})\right) + \theta f'(\beta)(y_{n+1} - y(t_{n+1}))\right) \\
&= y_n + hf\left(y_n + \theta(y(t_{n+1}) - y_n)\right) + \theta f'(\beta)h(y_{n+1} - y(t_{n+1})) \\
&= y_n + h\left(f(y_n) + \theta f'(y_n)f(y_n)h + O(h^2)\right) + \theta f'(\beta)h(y_{n+1} - y(t_{n+1})) \\
&= y_n + f(y_n)h + \theta f'(y_n)f(y_n)h^2 + O(h^3) + \theta f'(\beta)h(y_{n+1} - y(t_{n+1})) \\
&= y(t_{n+1}) + \left(\theta - \frac{1}{2}\right) f'(y_n)f(y_n)h^2 + O(h^3) + \theta f'(\beta)h(y_{n+1} - y(t_{n+1})),
\end{aligned}$$

so that the truncation error is

$$y(t_{n+1}) - y_{n+1} = \frac{\left(\frac{1}{2} - \theta\right) f'(y_n)f(y_n)}{1 - \theta f'(\beta)h}h^2 + O(h^3).$$

Thus, the scheme is second-order for $\theta = \frac{1}{2}$ and first-order otherwise.

(b) To analyze stability, we apply the scheme to the model problem $y'(t) = f(y(t)) = \lambda y(t)$:

$$\begin{aligned}
&y_{n+1} = y_n + h\lambda\left((1-\theta)y_n + \theta y_{n+1}\right) \\
&\quad\Rightarrow\quad (1 - \theta\lambda h)y_{n+1} - (1 + (1-\theta)\lambda h)y_n = 0.
\end{aligned}$$

The characteristic polynomial is thus

$$\rho(\phi) = (1 - \theta\lambda h)\phi - (1 + (1-\theta)\lambda h)$$

which has the single root

$$\zeta = \frac{1 + (1-\theta)\lambda h}{1 - \theta\lambda h}.$$

The region of absolute stability corresponds the set of complex λh such that

$$\left|\frac{1 + (1-\theta)\lambda h}{1 - \theta\lambda h}\right| < 1.$$

(c) The method is A-stable if it is stable whenever $\Re(\lambda h) < 0$. Thus we must determine when

$$|1 + (1-\theta)z| < |1 - \theta z|$$

for all $\Re(z) < 0$. Clearly, if $\theta \geq \frac{1}{2}$, $|\Re(1+(1-\theta)z)| < |\Re(1-\theta z)|$ and $|\Im(1+(1-\theta)z)| < |\Im(1-\theta z)|$, so $|1+(1-\theta)z| < |1-\theta z|$. On the other hand, suppose $\theta < \frac{1}{2}$, and consider z on the imaginary axis. Then $|1+(1-\theta)z| > |1-\theta z|$ since $1-\theta > \theta$, so by continuity, some z with $\Re(z) < 0$ is also such that $|1+(1-\theta)z| > |1-\theta z|$. Therefore, the method is A-stable if and only if $\frac{1}{2} \leq \theta \leq 1$.

5. (10 Pts.) Consider the equation

$$u_{tt} = u_{xx} + u_x,$$

to be solved for $t > 0$, $0 \leq x \leq 1$.

(a) Give initial data and boundary data that make this a well-posed problem. Do not assume periodicity in x.

(b) Give a stable and convergent finite difference approximation to this initial-boundary value problem. Justify your answers.

Solution

(a)

$$\begin{aligned} u(0,x) &= u_0(x),\ x \in [0,1]; \\ u_t(0,x) &= u_1(x),\ x \in [0,1]; \\ a_i(t)u(t,i) + b_i(t)u_x(t,i) &= c_i(t),\ t > 0,\ i = 0,1. \end{aligned}$$

(b) (F04.5(b))

6. (10 Pts.) Consider the equation

$$u_t = u_{xx} + u_{yy} + 2au_{xy},$$

where a is a real number, to be solved for $0 \le x \le 1$, $0 \le y \le 1$, $t \ge 0$, with initial data $u(x,y,0) = u_0(x,y)$ and periodicity in x and y: $u(x+1,y,t) \equiv u(x,y,t)$, $u(x,y+1,t) \equiv u(x,y,t)$.

(a) For which values of a would you expect good behavior of the solution?

(b) Write a stable and convergent finite difference approximation to this problem. Justify your answers.

Solution

(a) We first compute the symbol $p(s,\xi,\eta)$ of the differential operator $P = \partial_t - \partial_x^2 - \partial_y^2 - 2a\partial_{xy}$:

$$\begin{aligned} p(s,\xi,\eta) &= P\left(e^{st}e^{i(\xi x+\eta y)}\right) \Big/ e^{st}e^{i(\xi x+\eta y)} \\ &= s + \xi^2 + \eta^2 + 2a\xi\eta. \end{aligned}$$

The root of the symbol (as a function of s) is then

$$q(\xi,\eta) = -\left(\xi^2 + \eta^2 + 2a\xi\eta\right).$$

Well-posedness requires that $\Re(q_\pm)$ be bounded above for all ξ,η, i.e., in this case, due to homogeneity of q in ξ,η,

$$\xi^2 + \eta^2 + 2a\xi\eta \ge 0.$$

Considering $\xi = \eta$ gives the requirement $a \ge -1$; considering $\xi = -\eta$ gives the requirement $a \le 1$. We now show that these bounds are not only necessary, but also sufficient. So suppose $-1 \le a \le 1$; then

$$\begin{aligned} \xi^2 + \eta^2 + 2a\xi\eta &\ge \xi^2 + \eta^2 - 2|\xi\eta| \\ &= (|\xi| - |\eta|)^2 \\ &\ge 0, \end{aligned}$$

as desired. Therefore, the problem is well-posed for $-1 \le a \le 1$.

(b) We consider using forward differencing for u_t and centered differences for the spatial derivatives:

$$\begin{aligned}
P_{k,h_x,h_y}u^n_{\ell,m} &= D_{t+}u^n_{\ell,m} - D_x^2 u^n_{\ell,m} - D_y^2 u^n_{\ell,m} - 2aD_{x0}D_{y0}u^n_{\ell,m} \\
&= \frac{u^{n+1}_{\ell,m} - u^n_{\ell,m}}{k} - \frac{u^n_{\ell+1,m} - 2u^n_{\ell,m} + u^n_{\ell-1,m}}{h_x^2} - \frac{u^n_{\ell,m+1} - 2u^n_{\ell,m} + u^n_{\ell,m-1}}{h_y^2} \\
&- 2a\frac{u^n_{\ell+1,m+1} - u^n_{\ell+1,m-1} - u^n_{\ell-1,m+1} + u^n_{\ell-1,m-1}}{4h_xh_y}; \\
R_{k,h_x,h_y}f^n_{\ell,m} &= f^n_{\ell,m}.
\end{aligned}$$

The symbols $p_{k,h_x,h_y}(s,\xi,\eta)$ and $r_{k,h_x,h_y}(s,\xi,\eta)$ for these difference operators are

$$\begin{aligned}
p_{k,h_x,h_y}(s,\xi,\eta) &= P\left(e^{skn}e^{i(\xi h_x\ell+\eta h_y m)}\right)\Big/ e^{skn}e^{i(\xi h_x\ell+\eta h_y m)} \\
&= \frac{1}{k}\left(e^{sk}-1\right) + \frac{2}{h_x^2}(1-\cos\xi h_x) + \frac{2}{h_y^2}(1-\cos\eta h_y) + \frac{2a}{h_xh_y}\sin\xi h_x \sin\eta h_y; \\
r_{k,h_x,h_y}(s,\xi,\eta) &= R\left(e^{skn}e^{i(\xi h_x\ell+\eta h_y m)}\right)\Big/ e^{skn}e^{i(\xi h_x\ell+\eta h_y m)} \\
&= 1.
\end{aligned}$$

From this we quickly see that the scheme is consistent:

$$p_{k,h_x,h_y}(s,\xi,\eta) - r_{k,h_x,h_y}(s,\xi,\eta)p(s,\xi,\eta) = O(k) + O(h_x^2) + O(h_y^2).$$

By the Lax-Richtmyer Equivalence Theorem, stability of the scheme will imply convergence. Thus we replace $g = e^{sk}$ in $p_{k,h_x,h_y}(s,\xi,\eta) = 0$ and solve for g to determine the roots of the amplification polynomial:

$$\begin{aligned}
&\frac{1}{k}(g-1) + \frac{2}{h_x^2}(1-\cos\xi h_x) + \frac{2}{h_y^2}(1-\cos\eta h_y) + \frac{2a}{h_xh_y}\sin\xi h_x\sin\eta h_y = 0 \\
&\Rightarrow \quad g - 1 + 2\mu_x(1-\cos\theta) + 2\mu_y(1-\cos\phi) + 2a\sqrt{\mu_x\mu_y}\sin\theta\sin\phi = 0 \\
&\Rightarrow \quad g = 1 - 2\left(\mu_x(1-\cos\theta) + \mu_y(1-\cos\phi) + a\sqrt{\mu_x\mu_y}\sin\theta\sin\phi\right).
\end{aligned}$$

Let $c = \mu_x(1 - \ldots \sin\phi$ to simplify the notation, so that $g = 1-2c$. Then $|g| \le 1$ if and only if $0 \le c \le 1$. The lower bound on c is established as follows:

$$\begin{aligned}
c &= \mu_x(1-\cos\theta) + \mu_y(1-\cos\phi) + a\sqrt{\mu_x\mu_y}\sin\theta\sin\phi \\
&= \left(\sqrt{\mu_x}\sin\theta\right)^2\frac{1-\cos\theta}{\sin^2\theta} + \left(\sqrt{\mu_y}\sin\phi\right)^2\frac{1-\cos\phi}{\sin^2\phi} + a\left(\sqrt{\mu_x}\sin\theta\right)\left(\sqrt{\mu_y}\sin\phi\right) \\
&= \left(\sqrt{\mu_x}\sin\theta\right)^2\frac{1}{1+\cos\theta} + \left(\sqrt{\mu_y}\sin\phi\right)^2\frac{1}{1+\cos\phi} + a\left(\sqrt{\mu_x}\sin\theta\right)\left(\sqrt{\mu_y}\sin\phi\right) \\
&\ge \frac{1}{2}\left(\left(\sqrt{\mu_x}\sin\theta\right)^2 + \left(\sqrt{\mu_y}\sin\phi\right)^2 + 2a\left(\sqrt{\mu_x}\sin\theta\right)\left(\sqrt{\mu_y}\sin\phi\right)\right) \\
&\ge 0
\end{aligned}$$

if $-1 \le a \le 1$ (same argument as in (a)). We now show that $c \le 1$ if $\sqrt{\mu_x} + \sqrt{\mu_y} \le 1/\sqrt{2}$ (remembering that $|a| \le 1$):

$$\begin{aligned}
c &= \mu_x(1-\cos\theta) + \mu_y(1-\cos\phi) + a\sqrt{\mu_x\mu_y}\sin\theta\sin\phi \\
&\le 2\mu_x + 2\mu_y + \sqrt{\mu_x}\sqrt{\mu_y} \\
&\le 2\left(\mu_x + \mu_y + 2\sqrt{\mu_x}\sqrt{\mu_y}\right) \\
&= 2\left(\sqrt{\mu_x} + \sqrt{\mu_y}\right)^2 \\
&\le 1,
\end{aligned}$$

as desired. It follows that for $-1 \le a \le 1$ and $\sqrt{\mu_x} + \sqrt{\mu_y} \le 1/\sqrt{2}$, the scheme is stable, hence convergent.

7. (10 Pts.) Consider the boundary value problem

$$\begin{aligned} -\Delta u + u &= f(x,y),\ (x,y) \in \Omega = [0,1] \times [0,1], \\ u &= 0 \text{ for } (x,y) \in \partial\Omega,\ x = 0,1, \\ u_y &= 0 \text{ for } (x,y) \in \partial\Omega,\ y = 0,1. \end{aligned}$$

(a) Give a weak variational formulation of the problem.

(b) Analyze the existence and uniqueness of the solution to this problem. Justify your answers. (Assume $f \in L^2(\Omega)$.)

(c) Formulate a finite element approximation of the elliptic problem using piecewise-linear elements. Discuss the form and properties of the stiffness matrix and the existence and uniqueness of the solution of the linear system thus obtained. Justify your answers.

Solution

(W06.7)