6. Assignment 6

Due Wednesday, March 14

Gregory Beylkin, ECOT 323

(1) (Note: you will need the code of this assignment for another HW).
Implement trapezoidal rule to solve the initial value problem

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$$

where $\mathbf{y} = (y_1, y_2)^t$, $\mathbf{f}(t, \mathbf{y}) = (f_1(t, \mathbf{y}), f_2(t, \mathbf{y}))^t$ and $\mathbf{y}(0) = \mathbf{y}_0$. Use repeated Richardson extrapolation to improve the results.
Using your code to solve

$$t^2 y'' + t y' + \left(t^2 - 1\right) y = 0,$$

with the initial conditions $y(0) = 0$ and $y'(0) = 1/2$ on the interval $[0, 3\pi]$. Use repeated Richardson extrapolation to compute $y(3\pi)$ with 10 accurate digits. *Hint: For constructing the first order system, determine the first few terms of the Taylor expansion of the solution $y(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \mathscr{O}\left(t^4\right)$, and then substitute $y(t) = t u(t)$ to obtain the first order system for u.* The exact solution is $J_1(t)$, the Bessel function of the first kind of order 1.

(2) Show that the two step method

$$\mathbf{y}_{n+1} = \frac{1}{2}(\mathbf{y}_n + \mathbf{y}_{n-1}) + \frac{h}{4}[4\mathbf{f}(x_{n+1}, \mathbf{y}_{n+1}) - \mathbf{f}(x_n, \mathbf{y}_n) + 3\mathbf{f}(x_{n-1}, \mathbf{y}_{n-1})].$$

is second order.

(3) Determine order of the multistep method

$$\mathbf{y}_{n+1} = 4\mathbf{y}_n - 3\mathbf{y}_{n-1} - 2h\mathbf{f}(x_{n-1}, \mathbf{y}_{n-1}),$$

and illustrate with an example that the method is unstable.

(4) Show that the multistep method

$$\begin{gathered}\mathbf{y}_{n+3} + a_2 \mathbf{y}_{n+2} + a_1 \mathbf{y}_{n+1} + a_0 \mathbf{y}_n = \\ h[b_2 \mathbf{f}(t_{n+2}, \mathbf{y}_{n+2}) + b_1 \mathbf{f}(t_{n+1}, \mathbf{y}_{n+1}) + b_0 \mathbf{f}(t_n, \mathbf{y}_n)]\end{gathered}$$

is fourth order only if $a_0 + a_2 = 8$ and $a_1 = -9$. Deduce that this method cannot be both fourth order and convergent. (This is problem 2.6 in Iserles).

mine the three coefficients a, b, c. One finds, however, that the formula holds also for $f(x) = x^3$; hence one must go to $f(x) = x^4$ to get the error estimate. Generally speaking, one can expect similar experiences in problems which have a certain symmetry.

Example 7.2.3

We know the values of a function at the equidistant points $x_n = x_0 + nh$. Suppose that for a certain n we have that $f_n \leq f_{n-1}$ and $f_n \leq f_{n+1}$. Determine an approximate value for both the minimum point $\hat{x}$ and the minimum value of the function f in $[x_{n-1}, x_{n+1}]$.

Approximate f using a second-degree polynomial $\bar{f}$,

$$\bar{f}(x) = a + b(x - x_n) + \tfrac{1}{2}c(x - x_n)^2,$$

where we have (from Taylor's formula and Example 7.1.7),

$$a = \bar{f}_n,$$

$$b = \bar{f}'_n = \frac{\bar{f}_{n+1} - \bar{f}_{n-1}}{2h},$$

$$c = \bar{f}''_n = \frac{\bar{f}_{n+1} - 2\bar{f}_n + \bar{f}_{n-1}}{h^2}.$$

$$\bar{f}'(x) = b + c(x - x_n) = 0 \quad \text{for} \quad x = \hat{x} = x_n - \frac{b}{c},$$

$$\bar{f}(\hat{x}) = a - \frac{b^2}{c} + \frac{1}{2}c\left(\frac{b}{c}\right)^2 = a - \frac{1}{2}\frac{b^2}{c}.$$

Thus, having chosen $\bar{f}$ to be the interpolating polynomial which agrees with f at $x = x_{n-1}, x_n, x_{n+1}$, we get the *result*

$$\min f(x) \approx a - \frac{1}{2}\frac{b^2}{c} \quad \text{for} \quad x \approx \hat{x} = x_n - \frac{b}{c}, \tag{7.2.7}$$

where

$$a = f_n, \qquad b = \frac{f_{n+1} - f_{n-1}}{2h}, \qquad c = \frac{f_{n+1} - 2f_n + f_{n-1}}{h^2}.$$

As an error estimate for the minimum value of f we get:

$$|\min f(x) - \min \bar{f}(x)| \leq \max |f(x) - \bar{f}(x)|, \quad x \in [x_{n-1}, x_{n+1}].$$

Using the remainder term for interpolation (Theorem 4.3.3), we have, if $|f'''(x)| < M$ (see Problem 3 at the end of this section),

$$\max |f(x) - \bar{f}(x)| \leq \frac{1}{6} M \max |(x - x_n - h)(x - x_n)(x - x_n + h)|$$

$$= \frac{Mh^3}{(243)^{1/2}}.$$

M can be estimated by using third differences; see Theorem 7.1.4.

Example 7.2.4

Estimate f'_0, when $f_{-2}, f_{-1}, f_0, f_1, f_2$ are known $(x_n = x_0 + nh)$. The function f is judged to be so regular that a local fit with a second-degree polynomial (in $[x_{-2}, x_2]$) is reasonable, but the values of the function are contaminated by random errors; hence the method of least squares is motivated. Approximate $f(x)$ by $\bar{f}(x) = \sum_{j=0}^{2} c_j \varphi_j(x)$, where $\varphi_j(x) = (x - x_0)^j$. Then $f'_0 \approx \bar{f}'_0 = c_1$. Use the scalar product

$$(f, g) = \sum_{i=-2}^{2} f_i g_i.$$

From the normal equations of Eq. (4.2.6), we get:

$$(\varphi_0, \varphi_1)c_0 + (\varphi_1, \varphi_1)c_1 + (\varphi_2, \varphi_1)c_2 = (f, \varphi_1).$$

Since $(\varphi_0, \varphi_1) = (\varphi_2, \varphi_1) = 0$, we have

$$c_1 = \frac{(f, \varphi_1)}{(\varphi_1, \varphi_1)} = \frac{\sum_{i=-2}^{2} hif_i}{\sum_{i=-2}^{2} h^2 i^2}.$$

Hence

$$f'_0 \approx c_1 = \frac{2(f_2 - f_{-2}) + (f_1 - f_{-1})}{10h}. \tag{7.2.8}$$

Later in this chapter we shall see other approximate formulas for numerical differentiation.

7.2.2. Repeated Richardson Extrapolation

In many calculations what one would really like to know is the limiting value of a certain quantity as the step length approaches zero. Let $F(h)$ denote the value of the quantity obtained with step length h. The work to compute $F(h)$ often increases sharply as $h \longrightarrow 0$. In addition, the effects of round-off errors often set a practical bound for how small h can be chosen.

Often, one has some knowledge of how the truncation error $F(h) - F(0)$ behaves when $h \longrightarrow 0$. If

$$F(h) = a_0 + a_1 h^p + O(h^r), \quad (h \longrightarrow 0, r > p),$$

where $a_0 = F(0)$ is the quantity we are trying to compute and a_1 is unknown, then a_0 and a_1 can be estimated if we compute F for two step lengths, h and qh, $(q > 1)$:

$$F(h) = a_0 + a_1 h^p + O(h^r),$$

$$F(qh) = a_0 + a_1 (qh)^p + O(h^r),$$

from which we get

$$F(0) = a_0 = F(h) + \frac{F(h) - F(qh)}{q^p - 1} + O(h^r). \tag{7.2.9}$$

This process is called **Richardson extrapolation**, or *the deferred approach to the limit*. An example of this was mentioned in Chap. 1—the application of the above process to the trapezoidal rule for numerical integration (where $p = 2$, $q = 2$).

Suppose that, as in Theorem 7.1.5, one knows the form of a more complete expansion of $F(h)$ in powers of h. Then one can, even if the values of the coefficients in the expansion are unknown, **repeat the use of Richardson extrapolation** in a way to be described below. *This process is, in many numerical problems—especially in the numerical treatment of integrals and differential equations—the simplest way to get results which have negligible truncation error.* The application of this process becomes especially simple when the step lengths form a geometric series,

$$h_0,\ q^{-1}h_0,\ q^{-2}h_0,\ \dots .$$

THEOREM 7.2.1

Suppose that

$$F(h) = a_0 + a_1h^{p_1} + a_2h^{p_2} + a_3h^{p_3} + \dots, \tag{7.2.10}$$

where $p_1 < p_2 < p_3 < \dots$, *and set*

$$F_1(h) = F(h), \qquad F_{k+1}(h) = F_k(h) + \frac{F_k(h) - F_k(qh)}{q^{p_k} - 1}. \tag{7.2.11}$$

Then $F_n(h)$ *has an expansion of the form*

$$F_n(h) = a_0 + a_n^{(n)}h^{p_n} + a_{n+1}^{(n)}h^{p_{n+1}} + \dots .$$

Proof. (By induction.) From Eq. (7.2.10), the theorem holds for $n = 1$. Suppose the above expansion holds for $n = k$. Then, using Eq. (7.2.11), we see that $F_{k+1}(h)$ has an expansion containing the same powers of h as the expansion for $F_k(h)$. In the expansion for $F_{k+1}(h)$, the coefficient of h^{p_k} is

$$a_k^{(k)} + \frac{a_k^{(k)} - a_k^{(k)}q^{p_k}}{q^{p_k} - 1} = a_k^{(k)} - a_k^{(k)} = 0.$$

Thus the theorem holds for $n = k + 1$, and the proof is complete.

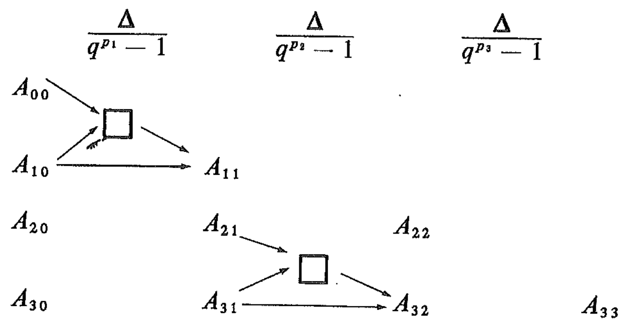
Hence if an expansion of the form of Eq. (7.2.10) is known, the above theorem gives a way to compute increasingly better estimates of a_0. The first term in the expression for the truncation error in $F_n(h)$ is $a_n^{(n)}h^{p_n}$, so we get truncation errors which begin with increasingly higher powers of h, as n increases (recall $p_1 < p_2 < p_3 < \dots$).

A moment's reflection on Eq. (7.2.11) will convince the reader that (using the notation of the theorem) $F_{k+1}(h)$ is determined by the $k + 1$ values $F_1(h)$, $F_1(qh), \dots, F_1(q^kh)$. One gets (with some slight changes in notation) the following algorithm:

Algorithm. For $m = 0, 1, 2, \dots$, set $A_{m,0} = F(q^{-m}h_0)$, and compute, for $k = 1, 2, \dots, m$,

$$A_{m,k} = A_{m,k-1} + \frac{A_{m,k-1} - A_{m-1,k-1}}{q^{p_k} - 1}. \tag{7.2.12}$$

The value $A_{m,k+1}$ is accepted as an estimate of a_0 when $|A_{m,k} - A_{m-1,k}|$ is less than the permissible error. The computations can be conveniently set up in the scheme:



(7.2.13)

Thus one extrapolates until two values *in the same column* agree to the desired accuracy. In most situations, the magnitude of the difference between two values in the same column gives (if h is sufficiently small), with a large margin, a bound for the truncation error in the lower of the two values. One cannot, however, get a guaranteed error bound in all situations.

The most common special case is to take $q = 2$ when one has an expansion of the form

$$F(h) = a_0 + a_1h^2 + a_2h^4 + a_3h^6 + \dots, \tag{7.2.14}$$

where clearly $p_k = 2k$. Then in (7.2.13), the headings of the columns become

$$\frac{\Delta}{3}, \frac{\Delta}{15}, \frac{\Delta}{63}, \dots, \quad \text{when} \quad p_k = 2k, \quad q = 2. \tag{7.2.15}$$

We now illustrate the application of the above in *numerical differentiation.* Using Theorem 7.1.5, one has an expansion of the form

$$\frac{f(a+h) - f(a-h)}{2h} = f'(a) + a_1h^2 + a_2h^4 + \dots .$$

Example 7.2.5

Compute $f'(3)$ for $f(x) = \ln(x)$ using values for $\ln(x)$ taken from a six-place table. Choose $h_0 = 0.8$. Then,

$$A_{m0} = \frac{\ln(3+h) - \ln(3-h)}{2h} \quad \text{with} \quad h = 2^{-m}h_0.$$

h		$\frac{\Delta}{3}$		$\frac{\Delta}{15}$		$\frac{\Delta}{63}$
0.8	$A_{00} = 0.341590$					
		$-2{,}087$				
0.4	$A_{10} = 0.335330$		$A_{11} = 0.333243$			
		-500		$+6$		
0.2	$A_{20} = 0.333830$		$A_{21} = 0.333330$		$A_{22} = 0.333336$	
		-125		0		0
0.1	$A_{30} = 0.333455$		$A_{31} = 0.333330$		$A_{32} = 0.333330$	

Using the stopping criterion of the algorithm, one accepts $A_{32} = 0.333330$, where $|R_T| \lesssim \frac{1}{2}\cdot 10^{-6}$. Since $f'(x) = 1/x$, the correct answer is $f'(3) = 0.333333$. The actual error is thus $-3\cdot 10^{-6}$. Here, round-off error is the dominant source of error, something more typical of numerical differentiation than of Richardson extrapolation.

One can show (for $p_k = 2k, q = 2$) that *if the values in the first column—i.e., $A_{00}, A_{10}, A_{20}, \ldots$—are afflicted with errors whose magnitudes are less than ϵ, then the errors caused later in the extrapolation scheme have magnitudes which nowhere exceed 2ϵ*. The reader is recommended to verify this, at least for $k = 1, 2, 3$. In the example above, the error in A_{m0} is at most $10^{-6}/2h \leq 5\cdot 10^{-6}$, which gives $|R_A| \leq 10^{-5}$. When *choosing the precision to be retained* in the values in the first column, one should consider *what precision one hopes to attain in the final result of the extrapolations*. In the example, the truncation error in A_{10} is—if the value is understood as an approximation to the derivative—about 0.002, but it would be wasteful to round A_{10} to three decimals. The extrapolation process also uses the information contained in the digits which are afflicted with truncation error.

The *idea* of a deferred approach to the limit (Richardson extrapolation) is much more general than the theorem and the algorithm given above. It is, for instance, not necessary that the step sizes form a geometric progression. Note that if $p_j = j\cdot p$ in Eq. (7.2.10), then the partial sums of the expansion are **polynomial** functions of h^p. If $k + 1$ values $F(q_0h), F(q_1h), \ldots, F(q_kh)$ are known, then by Theorem 4.3.2 a kth-degree interpolation polynomial is determined uniquely by the conditions:

$$Q((q_ih)^p) = F(q_ih), \quad i = 0, 1, 2, \ldots, k.$$

One can prove that $Q(0) - F(0) = O(h^{(k+1)p}), h \longrightarrow 0$.

There are many other variations. It is essentially a problem of estimating the coefficients in some theoretically motivated expression for $F(h)$, when some values of F are numerically known. The program for the extrapolation in these generalizations is more complicated, though quite practicable. One generalization that has proved advantageous in the numerical solution of differential equations (see the Bulirsch-Stoer method, Sec. 8.3.1) is to fit a **rational** function with almost the same degree in the numerator and the denominator to the given values of F.

Moreover, it is not at all necessary that the parameter be a step size. For instance, the same idea can be used when one has some theoretical knowledge of how the remainder in an infinite series depends asymptotically on the number of terms. Aitken extrapolation (see Sec. 3.2.3) is in this sense an application of the same basic idea.

Finally, the idea of a deferred approach to the limit is sometimes used in the experimental sciences—for example, when some quantity is to be measured in complete vacuum (difficult or expensive to produce). It can then be more practical to measure the quantity for several different values of the pressure. Expansions analogous to Eq. (7.2.10) can sometimes be motivated by the kinetic theory of gases, and the deferred approach to the limit can be used.

REVIEW QUESTION

Give the theory behind repeated Richardson extrapolation and explain its use in numerical differentiation.

PROBLEMS

1. Simpson's rule is occasionally written in the form:

$$\int_a^b f(x)\,dx \approx \frac{h}{3}(f_0 + 4U + 2E + f_n),$$

where $U = f_1 + f_3 + \ldots + f_{n-1}$, $E = f_2 + f_4 + \ldots + f_{n-2}$, for n even. Show that this agrees with the formula given in Example 7.2.2.

2. (a) Derive a formula

$$(2h)^{-1/2}\int_0^{2h} x^{-1/2}f(x)\,dx \approx (A_0f(0) + A_1f(h) + A_2f(2h))$$

which is exact when $f(x)$ is any second-degree polynomial.

(b) Give an asymptotically correct error term.

3. In Example 7.2.3 it is asserted that

$$\max \frac{1}{6}|(x - x_n - h)(x - x_n)(x - x_n + h)| = \frac{h^3}{243^{1/2}}, \quad |x - x_n| \leq h.$$

Prove this.

4. $\{f_n\}$ is a sequence of function values at equidistant points. Set

$$g_n = af_{n+1} + bf_n + cf_{n-1}.$$