MATH 335: Numerical Analysis

Problem Set 16, Solutions

10.1.6 Let F be an antiderivative of f in $[a, a+h]$. Then

$$\begin{aligned}
\int_a^{a+h} f(x)\,dx &= F(a+h) - F(a) \\
&= \left[F(a) + hF'(a) + \frac{1}{2}h^2F''(a) + \frac{1}{6}h^3F'''(\xi)\right] - F(a) \\
&= hf(a) + \frac{1}{2}h^2f'(a) + \frac{1}{6}h^3f''(\xi) \\
&= h[f(a) - f(a+h)] + h^2f'(a) + \left[hf(a+h) - \frac{1}{2}h^2f'(a)\right] + \frac{1}{6}h^3f''(\xi)\,,
\end{aligned}$$

where ξ is a number between a and $a+h$. But using Taylor's theorem again, there is a number ζ between a and $a+h$ with

$$f(a) - f(a+h) = f(a) - [f(a) + hf'(a) + \frac{1}{2}h^2f''(\zeta)] = -hf'(a) - \frac{1}{2}h^2f''(\zeta)\,.$$

So

$$\begin{aligned}
\int_a^{a+h} f(x)\,dx &= -h^2f'(a) - \frac{1}{2}h^3f''(\zeta) + h^2f'(a) + \left[hf(a+h) - \frac{1}{2}h^2f'(a)\right] + \frac{1}{6}h^3f''(\xi) \\
&= \left[hf(a+h) - \frac{1}{2}h^2f'(a)\right] + \frac{1}{6}h^3\left[f''(\xi) - 3f''(\zeta)\right] \\
&= hf(a+h) - \frac{1}{2}h^2f'(a) + O(h^3)\,.
\end{aligned}$$

That is, an expression for the truncation error is

$$\frac{1}{6}h^3\left[f''(\xi) - 3f''(\zeta)\right]\,.$$

10.1.20 Since $f(x) = (1+x^3)^{-1/2}$, we have $f''(x) = (15x^4 - 12x)/[4(1+x^3)^{5/2}]$. The next derivative $f'''(x)$ has just one *real* root (source *Mathematica*) inside $[1, 3]$, and the maximum of $|f''(x)|$, which is approximately 0.377587, occurs there. Thus, the error

$$|E| \le \frac{M(b-a)}{12}h^2 = \frac{M(b-a)^3}{12n^2} = \frac{(0.378)(2)^3}{12(100)^2} \doteq 2.52 \times 10^{-5}\,.$$

★33 (a) The actual value of the integral is

$$\int_0^1 x^{1/2}\,dx = \frac{2}{3}x^{3/2}\Big|_0^1 = \frac{2}{3}\,.$$

Since we showed the composite trapezoid rule to be an $O(h^2)$ method, we would expect that each time we "halve" the stepsize h, our error would be cut by a fourth (at least up until the point at which roundoff error begins to be sizeable in comparison to truncation error). The Octave commands

```
octave:102> h = 1/2;
octave:103> x = 0:h:1;
octave:104> y = sqrt(x);
octave:105> format long
octave:106> trapz(x, y)
ans =  0.603553390593274
```

produce a value for the composite trapezoid rule using stepsize $h = 1/2$. Here is a table of relevant values. The column marked n indicates how many subintervals are used ($n = 1/h$); T_n is the composite trapezoid approximation to the integra; E_n is the error in that approximation; $|E_n/E_{n-1}|$ is the ratio of errors between two consecutive step sizes.

h_n	n	T_n	$E_n = 2/3 - T_n$	$\lvert E_n/E_{n-1}\rvert$
2^{-1}	2	0.60355	0.0631133	
2^{-2}	4	0.64328	0.0233836	0.37050
2^{-3}	8	0.65813	0.0085364	0.36506
2^{-4}	16	0.66358	0.0030855	0.36145
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2^{-12}	4096	0.66667	2.233×10^{-6}	0.35401

Given that $h_n/h_{n-1} = 1/2$, we would have expected, if we had *quadratic* convergence, the column $|E_n/E_{n-1}|$ to contain values close to $1/4$; if we had *linear* convergence, these values would be close to $1/2$. Instead, they hover at a value in between, suggesting *superlinear* convergence. The reason our result (theorem?) on the composite trapezoid rule being quadratically convergent does not hold here is that there is no upper bound M on the values of the 2^{nd} derivative of our integrand inside $[0, 1]$ (in fact, the integrand is not even continuously differentiable throughout that *closed* interval).

(b) A similar table (to the one above) for Simpson's 1/3 rule is given below. Given our conclusions in part (a) about bounding derivatives of the integrand, it is not surprising that the column marked $|E_n/E_{n-1}|$ does not have values close to $1/16 = 0.0625$ like we would expect if truncation error $= \mathcal{O}(h^4)$ were realized. Nevertheless, it is perhaps surprising that there is virtually no speed-up in convergence rate for the composite Simpson 1/3 rule over the composite trapezoid rule for *this particular integral*.

h_n	n	S_n	$E_n = 2/3 - S_n$	$\lvert E_n/E_{n-1}\rvert$
2^{-1}	2	0.63807	0.0285955	
2^{-2}	4	0.65653	0.0101404	0.35462
2^{-3}	8	0.66308	0.0035874	0.35377
2^{-4}	16	0.66540	0.0012685	0.35359
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2^{-11}	2048	0.66667	8.759×10^{-7}	0.35355

10.2.20 (a) The derivatives of $f(x) = e^{2x}\sin(3x)$ up to 5$^{\text{th}}$ order are

$$\begin{aligned} f'(x) &= e^{2x}[3\cos(3x) + 2\sin(3x)]\,, & f^{(4)}(x) &= -e^{2x}[120\cos(3x) + 119\sin(3x)]\,,\\ f''(x) &= e^{2x}[12\cos(3x) - 5\sin(3x)]\,, & f^{(5)}(x) &= e^{2x}[-597\cos(3x) + 122\sin(3x)]\,,\\ f'''(x) &= e^{2x}[9\cos(3x) - 46\sin(3x)]\,. \end{aligned}$$

One could immediately say, from the expression for $f''(x)$, that $|f''(x)| \le 17e^4 \doteq 928.17$ for all $x \in [0,2]$. Being a little more careful (i.e., finding extrema of $|f''(x)|$ through solving for the zeros of $f'''(x)$), we get that the maximum of $|f''(x)|$ occurs at the right endpoint $x = 2$, where $M = |f''(2)| = e^4(12\cos 6 - 5\sin 6) \doteq 705.4$. Since $|E| \le M(b-a)h^2/12 = M(b-a)^2/(12n^2)$, it will suffice to solve the inequality

$$\frac{(705.4)(2)^3}{12n^2} \le 10^{-4} \qquad \Rightarrow \qquad n \ge \sqrt{\frac{(705.4)(2)^3(10)^4}{12}} \doteq 2168.6\,.$$

That is, we must use a minimum of 2169 subintervals.

(b) From the graph of $|f^{(4)}(x)|$, it appears it attains its maximum value at $x = 2$ as well, where $M = |f^{(4)}(2)| = e^4(120\cos 6 + 119\sin 6) \doteq 4475.42$. We have $|E| \le M(b-a)h^4/180 = M(b-a)^5/(180n^4)$, so as above it suffices to solve the inequality

$$\frac{(4475.42)(2)^5}{180n^4} \le 10^{-4} \qquad \Rightarrow \qquad n \ge \sqrt[4]{\frac{(4475.42)(2)^5(10)^4}{180}} \doteq 53.1\,,$$

or, at least 54 subintervals.

10.3.8 If T_{2^k} represents the composite trapezoid approximation with 2^k subintervals, then

k	$R_{k,1} = T_{2^k}$	$R_{k,2}$	$R_{k,3}$	$R_{k,4}$
0	25950			
1	21691	20271.33		
2	20625.5	20270.33	20270.27	
3	20359.25	20270.5	20270.51	20270.51

I don't quite understand the answer in the back of the text. Certainly the approximations keep changing from column to column. If we require an answer accurate in the ones digit, it seems we need to use all provided data values.

★34 (a) Here is my routine:

```
function intVal = glQuad (f, a, b, n)
% function intVal = glQuad (f, a, b, n)
%
%  This routine computes a numeric approximation of the definite integral
%  of f from a to b using Gauss-Legendre quadrature with n nodes.
%
%  INPUTS:
%     f     function handle for the integrand
%     a, b  limits of integration
%     n     number of nodes to use
%
%  OUTPUT:  approximate value of the integral

  [xis, wts] = GLNodeWt(n);        % wts is a column vector
  uis = ((b - a)*xis + (a + b))/2;
  fuis = [ ];
  for jj = 1:length(uis)
    fuis = [fuis f(uis(jj))];      % construct a row vector of f-values
  end
  intVal = fuis * wts * (b - a)/2;
end
```

(b) It would seem the answer in the back of the text is incorrect. The command

```
octave:222> quad(@(x) tan(pi*x/2), 0, 0.999)
ans = 4.11012737602947
```

gives us a better indication of the true value of the integral. Using 4 nodes, my routine yields the approximation 2.3399397. The lack of accuracy is due to the fact that $\int_0^1 \tan(\pi x/2)\,dx$ is *divergent* (the integrand has a vertical asymptote at $x = 1$), and the area it represents is infinite. Stopping at upper limit $x = 0.999$ makes the area finite, but the integrand is growing very fast towards the right end of the interval of integration.