Suppose we are given a complicated function $f(x)$ to integrate numerically on the interval $a \leq x \leq b$. It is important to have some idea how accurate our estimate is. The text provides formulas for guaranteed bounds; that is, numbers "Error" such that

$$\left| \int_a^b f(x)\,dx \;-\; \text{Estimate} \right| \;\leq\; \text{Error}.$$

Sometimes this is useless or nearly so. Why is that? Let's take the Trapezoidal rule as an example.

- **Difficulty 1**. The formula for the Trapezoidal rule error bound is $K(b-a)^3/12n^2$ where K is any number such that $|f''(x)| \leq K$ for $a \leq x \leq b$. We may not be able to obtain information about $f''(x)$. This could happen if the values of $f(x)$ are obtained by some measurement rather than by computing values of a mathematically defined function. See Problems 31–38 in Section 7.7.
- **Difficulty 2**. The guaranteed bound could be *much larger* than the actual error. As a result, we would end up taking n much larger than necessary to achieve some desired accuracy and thus we would do a lot of unneeded calculation. Why may the bound be too large? The error formula is based on estimating the maximum possible error in any interval Δx and multiplying that by the number of intervals. Some intervals may have much smaller errors and errors may be of opposite sign and so cancel out. (Recall that the Trapezoidal rule overestimates when $f'' > 0$ and underestimates when $f'' < 0$.) For example the Trapezoidal rule gives the exact result when used to estimate $\int_{-\pi}^{\pi} \sin x\,dx = 0$ (Can you see why?).

The second difficulty simply means more work since we will use a larger value of n than necessary, but the first difficulty means we can't use the error formula at all!

What can we do? The trick is to *estimate* the error directly from the numerical estimates for the integral. This method produces *only estimates, not guaranteed bounds.* (The estimates can be way off if the function $f(x)$ behaves in strange ways, but this is unlikely.) We'll study this approach for the Trapezoidal rule. The Midpoint rule is done exactly the same way. Simpson's rule is done similarly — replace squares of the number of intervals with fourth powers.

The Trapezoidal Rule

We use the notation "$a \approx b$" to mean "a is approximately equal to b."

Recall again that the book gives the error bound $K(b-a)^3/12n^2$ where $K \geq |f''(x)|$ for $a \leq x \leq b$. *If* this were the *actual error* rather than just a bound, the error would be proportional to $1/n^2$, that is, the error in estimating $\int_a^b f(x)\,dx$ with n intervals would be C_T/n^2, where C_T depends on a, b, and $f(x)$. It turns out that this is very nearly true.

This paragraph explains why the error is nearly C_T/n^2. You may skip the paragraph and still be able to read the rest of this note. Although the error bound may be too big for a large interval, it is usually quite accurate for a small interval of width, say Δx. Furthermore, the sign of the error is the opposite of the sign of $f''(x)$.* Thus we have an estimate for the error integrating from x_i to x_{i+1} and it is $-f''(x_i)(\Delta x)^3/12$. (I've omitted the details.) Adding these up over all the n divisions of $[a, b]$ and using $\Delta x = (b - a)/n$ gives

$$-\left(\sum f''(x_i)\Delta x\right)(b-a)^2/12n^2.$$

Since $\sum f''(x_i)\Delta x$ is an estimate for the integral of f'', we see that the error is approximately C_T/n^2 where $C_T \approx -(f'(b) - f'(a))(b - a)^2/12$.

Example A: Finding C_T When the Error is Known. Suppose we have estimated $\int_a^b f(x)\,dx$ using the Trapezoidal rule with $n = 10$ and somehow know that the error is E. How can we estimate the error when $n = 50$? We know the error is about C_T/n^2, but we don't know C_T. Since we know E and $E \approx C_T/10^2$, we have $C_T \approx 100E$. When $n = 50$, the error is about $C_T/50^2 \approx E/25$. Note that the number of steps went up by a factor of 5 and the error went down by a factor of about 5^2. This is true in general and is due the the fact that n is squared in the errror estimate C_T/n^2. This behavior is commented on in the discussion found on the page preceding the error bounds in Stewart.

The problem with the preceding example is that we are not likely to know the value of E. In fact, if someone gave us the error, we could simply add it to our estimate and obtain the exact value! How can we estimate E? Let I be the value of the integral and T_n the Trapezoidal rule estimate with n divisions. If we make two estimates, one with n divisions and one with k, we have the equations

$$I - T_n \approx C_T/n^2 \qquad \text{and} \qquad I - T_k \approx C_T/k^2. \tag{1}$$

The only unknowns are C_T and I. We solve these equations for C_T and I the same way we would if $\approx$ were replaced by $=$. Let's do it for the special case $k = 2n$. We have

$$I - (C_T/n^2) \approx T_n \qquad \text{and} \qquad I - (C_T/n^2)/4 \approx T_{2n}. \tag{2}$$

Solving for the two unknowns C_T and I using standard algebra:

$$C_T \approx \frac{4n^2(T_{2n} - T_n)}{3} \qquad \text{and} \qquad I \approx \frac{4T_{2n} - T_n}{3}. \tag{3}$$

It may appear that the estimate for I is more important than the estimate for C_T. The reverse is usually true: The estimate for C_T allows us to estimate the error for other values of n. The next two examples illustrate these ideas.

Example B: Getting Estimates From Tables. Example 5 in Section 7.7 provides a table of $r(t)$ from 1981 through 1997 and estimates $\int_{1981}^{1997} r(t)\,dt$ using Simpson's rule. Let's

---

* It has same sign as $f''(x)$ for the Midpoint rule.

estimate the integral *and the error* if one uses the Trapezoidal rule instead of Simpson's rule. With $2n = 16$, which uses all the data, we obtain the estimate 61.9 for the integral. Using $n = 8$, we obtain 63 instead. From (3), $I \approx (4 \times 61.9 - 63)/3 \approx 61.5$.

Alternatively, we could use the estimate for C_T in (3) to determine that $C_T \approx -100$. Thus the error for the approximation with $n = 16$ is about $-100/16^2 \approx 0.4$. In other words, 61.9 to be about 0.4 too high and so our better estimate for I is $61.9 - 0.4 = 61.5$. It's no accident that 61.5 is the same estimate that (3) gives for I: Equation (3) gives us our best guess for C_T and I. Using our best guess for C_T gives our best guess for the error in 61.9 and making this correction gives us our best guess for I.

Example C: Estimating $\ln 2$ Accurately. After his Example 1, Stewart gives a table of results for approximating $\ln 2 = \int_1^2 (1/x)\,dx$. With Trapezoidal rule, he obtains 0.695635 with 5 divisions and 0.693771 with ten. We can use (3) with $n = 5$. Our estimate for C_T is

$$C_T \approx \frac{100(0.693771 - 0.695635))}{3} \approx -0.062.$$

Thus, with $n = 20$, we would expect an error of about $-0.062/20^2 = -0.0016$, which agrees with the actual error found by Stewart for $n = 20$.

We can use C_T to estimate how large a value of n is needed to obtain an estimate of some desired accuracy. For example, to obtain an error less than 0.5×10^{-6}, we would choose n so that $0.062/n^2 < 0.5 \times 10^{-6}$. In other words $n > (0.14 \times 10^6)^{1/2} \approx 400$.

Remark There is no reason why we had to take $k = 2n$ in (1). For example, we could have used $k = 3n$. There is even no reason to require that k be an integer multiple of n. In fact, we could solve (1) with k still present. In that case, we obtain

$$C_T \approx \frac{(kn)^2(T_k - T_n)}{k^2 - n^2}. \tag{6}$$

The Midpoint Rule and Simpson's Rule

Example D: Connections with Simpson's Rule and the Midpoint Rule. What about the estimate for I given by the second equation in (3)? This is the same formula as Simpson's rule! (You should verify this.) In other words, by trying to eliminate the error we have obtained Simpson's rule, which generally has a much smaller error than the Trapezoidal rule.* Put another way, we've derived the formula for Simpson's rule from that for T_n.

Simpson's rule can also be written as $(2M_n + T_n)/3$, where M_n is the midpoint rule estimate. (You should verify this.) Since this approximation to I has a much smaller error than T_n and M_n, we can write $I \approx (2M_n + T_n)/3$ and the error will be much smaller than

* From the text, the error in Simpson's rule has the form C_S/n^4, which goes to zero faster that the Trapezoidal rule because there is an n^4 in the denominator instead of n^2.

writing either $I \approx M_n$ or $I \approx T_n$. So what? Multiplying $I \approx (2M_n + T_n)/3$ by 3 and rearranging, we have

$$I - T_n \approx -2(I - M_n).$$

In words, the Trapezoidal rule error is about twice the Midpoint rule error and of opposite sign. Of course, you knew that for the error bounds from the theorem in the text. Now you know it for the actual errors as well as the bounds, and you have a way of seeing why it should be true.

Example E: Redoing Example B with Simpson's Rule. Let's repeat the process using Simpson's rule. We need the formulas that correspond to (2) and (3). As indicated earlier, we replace n^2 with n^4 in (2) to obtain

$$I - (C_S/n^4) \approx S_n \qquad \text{and} \qquad I - C_S/(2n)^4 = I - (C_S/n^4)/16 \approx S_{2n} \tag{4}$$

and so

$$C_S \approx \frac{16n^4(S_{2n} - S_n)}{15} \qquad \text{and} \qquad I \approx \frac{16S_{2n} - S_n}{15}. \tag{5}$$

Stewart obtained 61.533 using Simpson's rule with 16 divisions, so we have $2n = 16$. I obtained 60.4 by using Simpson's rule with $n = 8$. From (5), $C_s \approx 16 \times 8^4(61.533 - 60.4)/15$ and so the error estimate for 61.533 is approximately

$$\frac{16 \times 8^4(61.533 - 60.4)/15}{16^4} = \frac{1.133}{15} \approx 0.075.$$

Thus we estimate that 61.533 is about 0.075 too low and so our improved estimate is 61.608.

Exercises Based on Section 7.7 Exercises

31T. Use the table in Exercise 31 and (3) in this note to estimate the error in the Trapezoidal rule approximation to the integral.

32T. Use the table in Exercise 32 and (3) in this note to estimate the error in the Trapezoidal rule approximation.

31S. Use the table in Exercise 31 and (5) in this note to estimate the error in the Simpson's rule approximation.

32S. Explain why you cannot use the table in Exercise 31 and (5) in this note to estimate the error in the Simpson's rule approximation.

Eq6. Derive (6) in this note and derive a similar result for Simpson's rule.