Romberg Integration

Let \( f(x) \) be a function with many continuous derivatives on the interval \([a, b]\).

If we break \([a, b]\) into \( n \) subintervals each of length \( h \) (\( h = (b-a)/n \)) we can do a composite trapezoidal approximation to

\[

t_{a}^{b} f(x) \, dx .\quad \text{We'll denote } \int_{a}^{b} f(x) \, dx \text{ by } I(f)
\]

and the trapezoidal approximation to \( I(f) \) using \( n \) subintervals by \( T_{n}(f) \) or just \( T_{n} \).

Thus,

\[
T_{n}(f) = h \left\{ \frac{1}{2} f_{0} + f_{1} + f_{2} + \ldots + f_{n-1} + \frac{1}{2} f_{n} \right\}
\]

using the abbreviation \( f_{j} = f(x_{j}) = f(a + jh) \).

The following fact is true although we will not derive it:

\[
I(f) = T_{n}(f) + C_{2} h^{2} + C_{4} h^{4} + C_{6} h^{6} + C_{8} h^{8} + \ldots
\]

where \( C_{2}, C_{4}, C_{6}, \ldots \) are constants which depend on the function \( f(x) \) and on \( a \) and \( b \) but not on \( h \).
Example

The interpretation of equation (1) is that if

\( h \) is "sufficiently small", then the major part of
the error \( (I(f) - T_n(f)) \) is the term \( c_2 h^3 \).

We seek, of course, to reduce the error. Suppose
we divide \([a, b]\) into twice as many
subintervals \((2n, of \ them)\) and use the
trapezoidal rule. Then

\[
2) \quad I(f) = T_{2n}(f) + c_2 \left( \frac{h}{2} \right)^2 + c_4 \left( \frac{h}{2} \right)^4 + c_6 \left( \frac{h}{2} \right)^6 + \ldots
\]

since the length of each subinterval is now
half of what it was originally, or \( \frac{h}{2} \).
Equation (2) can be written

\[
3) \quad I(f) = T_{2n}(f) + \frac{1}{4} c_2 h^2 + \frac{1}{16} c_4 h^4 + c_6 \frac{1}{64} c_6 h^6 + \ldots
\]

We've reduced the first term by a factor of 4,
but we can do much better with the
following "trick":
Rewrite equations (1) and (3):

1) \( I(f) = T_n(f) + C_2 h^2 + C_4 h^4 + C_6 h^6 + \ldots \)

2) \( I(f) = T_n(f) + \frac{1}{4} C_2 h^2 + \frac{1}{16} C_4 h^4 + \frac{1}{64} C_6 h^6 + \ldots \)

Multiply equation (3) by 4 and subtract equation (3) from equation (1) to get:

4) \[ (1-4T_n) I(f) = T_n(f) - 4T_n(f) + (C_2 - 4 \cdot C_2) h^2 + (C_4 - 4 \cdot \frac{1}{16} C_4) h^4 \]

or

5) \( I(f) = \frac{4T_n(f) - T_n(f)}{3} - \frac{1}{4} C_4 h^4 - \frac{5}{16} C_6 h^6 + \ldots \)

Equation (5) says that the expression

\[ T_n \frac{1}{3} \approx T_n + \frac{T_n - T_n}{3} \]

approximates \( I(f) \)
with an error whose main part is proportional to $h^4$.
If $h$ is small, this is going to be significantly smaller
than the error we had in either of equation (1) or
equation (3). Notice that to carry out this "trick",
we did not need to know what the constant $C_2$ was.

Example: \( f(x) = x^4 \)
\[
\int_{10}^{12} x^4 \, dx = 2,766.4
\]
\( f(10) = 10,000 \); \( f(11) = 14,641 \); \( f(12) = 20,736 \)

\[
T_1(f) = 2 \left\{ \frac{f(10)}{2} + \frac{f(12)}{2} \right\} = 30,736
\]

\[
T_2(f) = 4 \left\{ \frac{f(10)}{2} + f(11) + \frac{f(12)}{2} \right\} = 50,009
\]

\[
\frac{4T_2(f) - T_1(f)}{3} = 29,766.67 \text{ which is much closer}
\]
\[
to \int_{10}^{12} x^4 \, dx = 29,766.4 \text{ than}
\]

\[
\therefore T_1 \text{ or } T_2
\]

We can do this same process repeatedly to reduce the
error; this is called \underline{Romberg Integration}
and yields very accurate approximations to \( \int_a^b f(x) \, dx \)
(for functions that have lots of continuous derivatives)

\[
\text{Let's see what repeating this a couple of times will give us:}
\]
Suppose we compute \( T_{4n} \), that is, the composite trapezoidal rule approximation to \( I(f) \) using \( 4n \) subintervals of length \( h/4 \) each.

We'd get

\[
I(f) = T_{4n} + C_2 (h/4)^2 + C_4 (h/4)^4 + C_6 (h/4)^6 + \cdots
\]

\[
= T_{4n} + \frac{1}{16} C_2 h^2 + \frac{1}{256} C_4 h^4 + \cdots
\]

Recall that we also have Equation (13):

\[
I(f) = T_{2n} + C_2 (\frac{h}{2})^2 + C_4 (\frac{h}{2})^4 + C_6 (\frac{h}{2})^6 + \cdots
\]

\[
= T_{2n} + \frac{1}{4} C_2 h^2 + \frac{1}{16} C_4 h^4 + \cdots
\]

Taking 4 times Equation (7) and subtracting it from Equation (13) gives:

\[
I(f) = \frac{4 T_{4n} - T_{2n}}{4 - 1} + \frac{1}{16} - \frac{4}{256} C_4 h^4 + \cdots
\]

\[
= T_{4n} + \frac{(T_{4n} - T_{2n})}{3} - \frac{1}{64} C_4 h^4 + \cdots
\]
So

\[ (8) \quad I(f) = T_{4n}^1 - \frac{1}{64} C_4 h^4 + O(h^6) \]

We also had (equation 5)

\[ (5) \quad I(f) = T_{2n}^1 - \frac{1}{4} C_4 h^4 + O(h^6) \]

Taking 16 times equation (8) and subtracting from equation (5) gives

\[ (9) \quad -15 \ I(f) = \frac{T_{2n}^1 - 16 T_{4n}^1}{16} + O(h^6) \]

\[ (10) \quad I(f) = \frac{16 T_{4n}^1 - T_{2n}^1}{15} + O(h^6) \]

\[ = \frac{T_{4n}^1 + \frac{T_{4n}^1 - T_{2n}^1}{15}}{15} + O(h^6) \]

\[ = \frac{T_{4n}^2 + O(h^6)}{15} \]
With Romberg Integration we do this repeatedly to form a table.

The table looks like the following:

<table>
<thead>
<tr>
<th>$\frac{\Delta}{3}$</th>
<th>$\frac{\Delta}{15}$</th>
<th>$\frac{\Delta}{63}$</th>
<th>$\frac{\Delta}{127}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_n$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_{2n}$</td>
<td>$T_{2n}^1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_{4n}$</td>
<td>$T_{4n}^1$</td>
<td>$T_{4n}^2$</td>
<td></td>
</tr>
<tr>
<td>$T_{8n}$</td>
<td>$T_{8n}^1$</td>
<td>$T_{8n}^2$</td>
<td>$T_{8n}^3$</td>
</tr>
<tr>
<td>$T_{16n}$</td>
<td>$T_{16n}^1$</td>
<td>$T_{16n}^2$</td>
<td>$T_{16n}^3$</td>
</tr>
</tbody>
</table>

\[ (h^2) \quad (h^4) \quad (h^6) \quad (h^8) \]

\[ \text{\( \text{Error in each column is proportional to} \)} \]

All we've done so far is to calculate the first 3 rows.

We calculated $T_n$, then $T_{2n}$, then $T_{4n}$, and $T_{16n}$.

Then we calculated $T_{8n}$ and $T_{4n}^1$ and $T_{4n}^2$.  

\[ \text{does not mean squared; it's just an identifying tag} \]
Each number in the table is an approximation to the integral of $f(x)$. Smaller subintervals

As you go down a column the approximations get better and better, i.e. the numbers in each column converge to $\int f(x) \, dx$ as you go down the column.

If you look at a column further to the right, the numbers in that column converge to $\int f(x) \, dx$ faster.

If you look at the numbers on the diagonal, i.e. $T_0, T_{0n}^1, T_{0n}^2, T_{0n}^3, \ldots$, they converge to $\int f(x) \, dx$ faster than the numbers in any single column.

We'll use these diagonal numbers to approximate $\int f(x) \, dx$.

This will be our Romberg Integration Rule.
How do we construct the table?

The numbers in the leftmost column, $T_n, T_{2n}, T_{4n}, \ldots$ are just the composite Trapezoidal rule approximations to $I(f)$ using the number of subintervals indicated by the subscript, so $T_n$ uses $2n$ subintervals, $T_{2n}$ uses $4n$ subintervals, etc.

A number, say $T_{4n}^2$, is obtained from two of the numbers in the column just to its left. The two numbers are the one in the same row (here $T_{4n}^1$) and the one a row higher up ($T_{2n}^1$). These two numbers are combined to give $T_{4n}^2$ using the instructions in the table heading $\Delta \over 15$. $\Delta$ stands for difference; what we do to get $T_{4n}^2$ is take $T_{4n}^1$ and add to it the difference $(T_{4n}^1 - T_{2n}^1)$ divided by 15.

So $T_{4n}^2 = T_{4n}^1 + \frac{(T_{4n}^1 - T_{2n}^1)}{15}$.

as we saw before.
Similarly, \( T_{8n}^2 = T_{8n} + \frac{(T_{8n} - T_{4n})}{15} \)

and

\( T_{8n}^3 = T_{8n}^2 + \frac{(T_{8n}^2 - T_{4n}^2)}{63} \) etc.

In general, we get \( T_{2j}^k \) from

\[
T_{2j}^k = T_{2j}^{k-1} + \frac{(T_{2j}^{k-1} - T_j^{k-1})}{4^{k-1} - 1}
\]

\[\Delta \frac{3}{15} \quad \Delta \frac{63}{63}\]

---

Note: We define \( T_{2j}^0 = T_{2j} \)
To use Romberg Integration we proceed as follows:

Choose a number $n$

Calculate and store $T_n$.

Calculate and store $T_{2n}$ and use $T_n$ and $T_{2n}$ to calculate $T_{2n}^1$ (store $T_{2n}^1$).

(You can discard $T_{2n}$ now if you want.)

Calculate and store $T_{4n}$ and use $T_{2n}$, $T_{2n}^1$, and $T_{4n}$ to calculate $T_{4n}^1$ and $T_{4n}^2$.

Store $T_{4n}^1$ and $T_{4n}^2$.

If we've just calculated and stored the numbers $T_{d}$, $T_{d}^1$, $T_{d}^2$, ..., then to get the next row of the table, first calculate the trapezoidal approximation $T_{2d}$ and then use

$$T_{2d}^k = T_{2d}^{k-1} + \frac{(T_{2d}^{k-1} - T_{2d}^{k-1})}{4^{k-1}}$$

for $k = 1, 2, ...$
to get the rest of this row.

Remark 1: Consider \( h = \frac{b-a}{n} \) and calculate trapezoidal approx.

\[
T_n(f) = h \left\{ \frac{f(a)}{2} + f(a+h) + f(a+2h) + \ldots + f(a+(n-1)h) + \frac{f(b)}{2} \right\}
\]

Also,

\[
T_{2n}(f) = \frac{h}{2} \left\{ \frac{f(a)}{2} + f(a+\frac{1}{2}h) + f(a+h) + f(a+\frac{3}{2}h) + f(a+2h) + \ldots + f(a+(n-1)h) + f(a+\frac{3n-1}{2}h) + \frac{f(b)}{2} \right\}
\]

\[
= \frac{1}{2} T_n(f) + \frac{h}{2} \left\{ f(a+\frac{1}{2}) + f(a+\frac{3}{2}h) + \ldots + f(a+\frac{2n-1}{2}h) \right\}
\]

Therefore, if we have \( T_n(f) \) we can get \( T_{2n}(f) \) without evaluating \( f \) at \( a, a+h, a+2h, \ldots, a+(n-1)h \) again.
Using this idea, all the trapezoidal approximations \( T_n(f), T_{2n}(f), T_{4n}(f), \ldots \) for the Romberg Integration table can be computed without ever having to evaluate \( f(x) \) at the same point twice.

Remark 2: When we're using Romberg Integration, how do we know when to stop?

We have to set a tolerance — say we want an approximation to \( I(f) \) with error less than \( 10^{-6} \). We need a means for estimating the error, i.e., the difference between \( I(f) \) and a given item in our table.

It turns out that, in general, if

\[
| T_{2j} - T_{j}^{k} | < 10^{-6},
\]

then

\[
T_{2j}^{k+1} \text{ will be within } 10^{-6} \text{ of } I(f)
\]

and we should therefore accept \( T_{2j}^{k+1} \) as our approximation.
That is, if $|T_{2j}^k - T_{2j}^k| < 10^{-6}$, accept the approximation

$$T_{2j}^{k+1} = T_{2j}^k + \left(\frac{T_{2j}^k - T_{2j}^k}{4^{k+1}}\right).$$

This is not foolproof.

To motivate this test, let's go back to the top of the table. Recall

$$I(f) = \tan + \frac{1}{4} C_2 h^2 + \frac{1}{16} C_4 h^4 + \cdots.$$  

If $h$ is "sufficiently small," the dominant part of the error is the $h^2$ term. The $h^4$ term is much smaller.

We also have

$$I(f) = \tan + \frac{T_{2n} - T_{1n}}{4} - \frac{1}{4} C_4 h^4 + \cdots.$$  

Comparing the last two equations, we see that
if the $h^4$ term is negligible, then

$$\frac{T_{n+1} - T_n}{4-1} \approx \frac{1}{4} C_2 h^4 \approx \text{the error},$$

So, if $|T_{n+1} - T_n|$ were below our error tolerance, we'd have good reason to believe that the error was also below the tolerance.

The reason that this test isn't foolproof is that it's possible for $h$ to have $|T_{n+1} - T_n|$ be very small (by chance) even though $h$ isn't small enough to make the $h^4$ term negligible. In this case, the $h^4$ term still contains a significant portion of the error.

The same kind of reasoning applies to $|T_{j+k} - T_j^k|$. 