Section 5.4 (continued) Repeated Richardson extrapolation.

We use the previous example in numerical derivative to explain the repeated Richardson extrapolation.

Instead of using Taylor polynomial in approximating $f'(x)$ (like one-sided and two-sided approximations done in the previous class notes), we can consider Taylor series:

$$f(x+h) = f(x) + \frac{f(x)}{1!} h + \frac{f'(x)}{2!} h^2 + \frac{f''(x)}{3!} h^3 + \cdots$$

Can view this as:

$$f'(x) = \frac{f(x+h) - f(x)}{h} + \epsilon_h h + \epsilon_h^2 h^2 + \epsilon_h^3 h^3 + \cdots$$

$$u\quad \text{true value}$$
$$u^*\quad \text{approximation}$$
$$\epsilon(u^*) = u - u^*$$
$$\text{error (approximation)} = C_1 h + C_2 h^2 + C_3 h^3 + \cdots$$

In fact, we can view (1) in an inductive manner later.
Recall, bottom of page 10, previous class notes:

\[
\text{true value} = \underbrace{\text{approximation}}_{\text{like } f'(x)} + \underbrace{\text{error}}_{\text{like } f(x+h)-f(x)} \sim K \cdot h_k
\]

depending on \( h \)

Then, by iterating the sequence

\[\begin{align*}
& h_0, \quad h_1 = \frac{1}{2} h_0, \quad h_2 = \frac{1}{2} h_1, \quad h_3 = \frac{1}{2} h_2, \quad \ldots \to 0 \\
& x + h_0, \quad x + h_1, \quad x + h_2, \quad x + h_3, \quad \ldots \to x
\end{align*}\]

Do \( D_0, D_1, D_2, \ldots \to D \)

(2)

\[D = D_0 + \frac{D_n - D_{n-1}}{2^{n-1}}\]

for example, \( D_n = \frac{f(x + h_n) - f(x)}{h_n} \) for \( n = 0, 1, 2, \ldots \)

Now, we combine (1) and (2) in an inductive manner (or \( h_k \))
Assume (original) approximations

\[ D_0, D_1, D_2, \ldots \]

\[ D_n = \frac{f(x + nh) - f(x)}{h} \]

Viewing (1) as true value = approximation + \( c_1 h \), \( (c_2 h^2 + \ldots) \)

So error = \( c_1 h \)

Then, viewing (1) as true value = approximation + \( c_2 h^2 \), \( (c_3 h^3 + \ldots) \)

So error = \( c_2 h^2 \)

Then, viewing (1) as true value = approximation + \( c_3 h^3 \), \( (c_4 h^4 + \ldots) \)

So error = \( c_3 h^3 \)
<table>
<thead>
<tr>
<th>$n$</th>
<th>$D_n$</th>
<th>$D_n^*$</th>
<th>$D_n^{**}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$D_0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$D_1$</td>
<td>$D_1^*$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$D_2$</td>
<td>$D_2^*$</td>
<td>$D_2^{**}$</td>
</tr>
<tr>
<td>3</td>
<td>$D_3$</td>
<td>$D_3^*$</td>
<td>$D_3^{**}$</td>
</tr>
<tr>
<td>4</td>
<td>$D_4$</td>
<td>$D_4^*$</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

$D_n = \frac{f(x + h_n) - f(x)}{h_n}$

$n = 0, 1, 2, \ldots$

$D_n^* = D_n + \frac{D_n - D_{n-1}}{2^1 - 1}$

$D_n^{**} = D_n + \frac{D_n^* - D_{n-1}^*}{2^2 - 1}$

$n = 1, 2, \ldots$

$D_n^{***} = D_n + \frac{D_n^{**} - D_{n-1}^{**}}{2^3 - 1}$

$n = 2, 3, \ldots$
Example 5.4-2:

Compute the triangular Richardson table for the above example. Use $h_0 = 0.1$, so that $h_i = 0.1, 0.05, 0.025, \ldots$. As explained above, the values of $k_j$ for this example are $k_j = 1, 2, 3, \ldots$. The triangular Richardson table is then

\[
\begin{array}{cccc}
0.497363752535389 & 0.519044815722409 & 0.529728186647875 & 0.535029119429193 \\
0.540725878909429 & 0.540411557573841 & 0.540306783794645 & 0.540330034210510 \\
0.540302859736233 & 0.54032299179317 & \\
\end{array}
\]

The errors in these values are

\[
\begin{array}{cccc}
0.429E-0 & -0.424E-3 & -0.109E-3 & -0.277E-4 \\
0.213E-1 & & -0.448E-5 & \\
0.106E-1 & & & -0.554E-6 \\
0.527E-2 & & & 0.669E-8 \\
\end{array}
\]

Note that the errors decrease by a factor of about 2 going down the first column, by a factor of about 4 going down the second column, and by a factor of about 8 going down the third column. Note also the greatly decreasing errors going across each row. Richardson extrapolation has improved the accuracy spectacularly.

\[D \approx 0.590302\ldots\]

\[
0.318(E^{-8})
\]

- Stabilized due to floating-point restriction on the underlying FPN5

- Stabilized when $h = \frac{1}{2^3} h_0 = 0.1 \times 2^{-2} = 2^{-3}$
Conceptually: a problem $P$ is ill-conditioned if, in general, a small change to a typical problem instance $P$, say, small $\Delta P$, can produce a large relative change in the exact mathematical solution $S$,

$$\frac{\Delta S}{S}$$

"vague" — concrete examples in mathematics later.
ill-conditioning of a problem $P$ is an (intrinsic) property of $P$, not due to any algorithm solving $P$.

An algorithm solving $P$ may be limited by restrictions of the underlying FNS, algorithms that cause subtractive cancellation, etc.

The problem is "sensitive" to changes to its problem instances.

Condition number $(P)$ is defined as:

$$ \text{maximum} \left\{ \left| \frac{\text{relative change to } S}{\text{relative change to } P} \right| : P \in P \right\} $$

"large" $\rightarrow \infty$

Condition number $(P)$

not ill-conditioned well
"empirically", for problem $P$ with condition number $K$ (\[ K \approx \max \frac{\text{relative change in solution}}{\text{relative change to problem instance with small } \epsilon} \mid |P(\epsilon)| \])

FPNS

with the "best" possible algorithm solving $P$.

It is expected to lose up to $\log_b(K)$ digits of precision due to "unavoidable subtractive cancellation".

Example: $K = 1000$,

$b = 10, \quad \log_b K = 3$

Expect to lose up to 3 decimal digits of precision/accuracy in the best possible computation due to unavoidable subtractive cancellation.

Or, if we have a small change in a problem instance like changing the "last digit" of a problem instance by 1,

then, approximately, the last $\log_b K$ digits of the corresponding solution will change (in some component(s) of any best possible algorithm/computation).
Problem P

1. Unconditioned
   - K is large

2. Conditioned
   - Small

   "Stable Algorithm"

   "Unstable Algorithm"

   Loss of accuracy in performance

   - Small training data

   - Large training data

   - More information
Section 6.2  Condition Number of Function Computation

Real-valued function \( f \)
real argument \( x \in \text{domain}(f) \)

Problem: compute \( f(x) \)
(problem instance: the input/argument \( x \))
(not the function \( f \))

Condition number (compute \( f \) at argument \( x \))
\( \text{cond}(f, x) \)

By the definition in Section 6.1 earlier:
\[
\text{cond}(f, x) = \left| \frac{\text{relative change in } f(x)}{\text{relative change in } x} \right| \quad \text{with small } \Delta x
\]

\[
= \left| \frac{f(x+\Delta x) - f(x)}{f(x)} \right| \quad \text{when } \Delta x \to 0
\]

\[
= \left| \frac{\frac{f(x+\Delta x) - f(x)}{\Delta x}}{\frac{x}{f(x)}} \right| \quad \text{when } \Delta x \to 0
\]

\[
= \left| f'(x) \cdot \frac{x}{f(x)} \right|
\]
Remarks:

1. \( \text{Cond} (f, x) \) large \( \sim \) combination:
   - large \( f'(x) \)
   - large \( x \)
   - small \( \text{fix} \)

2. In a FPNS radix-\( b \):
   a change to the last digit of \( x \) \( \Rightarrow \) last \( \log_b (\text{cond} (f, x)) \) bits in \( x \) will change.

**Condition:** 
\[
\text{Cond} (f, x) = \left| \frac{\Delta f(x)}{f(x)} \right| = \left| \frac{\Delta x}{x} \right|
\]

Small \( f(x) \): \( \frac{\Delta f(x)}{f(x)} \) (large)

Large \( f(x) \): \( \frac{\Delta x}{x} \) (large)

large \( x \): \( \frac{\Delta x}{x} \) small
Example 1: \( f = \sin (\text{ sine function}) \)

\( x = \frac{\pi}{2} \)

\[
\text{Cond} (f, x) = \left| \frac{f'(x)}{f(x)} \right| = \frac{\sin x}{\cos x}
\]

\( f'(x) = \cos x \)

\[
= \left| \cos x \cdot \frac{x}{\sin x} \right|_{x = \frac{\pi}{2}} = 0
\]

perfectly well-conditioned!

\[
\text{Cond} (f, x) = \left| \frac{\Delta f(x)}{\Delta x} \right|_{\Delta x \to 0} = 0
\]

so, regardless of \( \frac{\Delta x}{x} \), \( \frac{\Delta f(x)}{\Delta x} = 0 \)

(i.e., \( \Delta f(x) = 0 \) "flat"

\( f'(\frac{\pi}{2}) = 0 \)

\( f(\frac{\pi}{2}) = \cos x \)

\( x = \frac{\pi}{2} \)
Example 2: \( f = \sin x \), \( x = 0 \)

\[
\lim_{x \to 0} \frac{f'(x) - \frac{x}{f(x)}}{x} = \lim_{x \to 0} \frac{\frac{df}{dx}}{f(x)} = \frac{f'(0) - \frac{0}{f(0)}}{0} = \frac{\cos 0 - \frac{0}{\sin 0}}{0} = \frac{1 - 0}{0} = \text{undefined}
\]

So, consider \( \lim_{x \to 0} \frac{\cos x \cdot \frac{x}{\sin x}}{x} \)

\[
= \lim_{x \to 0} \left( \cos x \cdot \frac{x}{\sin x} \right) = \cos 0 \cdot \frac{x}{\sin x} = 1 \cdot \frac{x}{\sin x} = \frac{1}{\cos 0} = 1
\]

\[
\lim_{x \to 0} \frac{\sin x}{x} = 1
\]

well-conditioned

\[
\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x) = \cos x
\]

\[
f'(0) = 1
\]

\[
f(0) = \sin x
\]
Example 3: \( f = \sin \quad x = \pi \)

\[
\text{cond}(f, x) = \left| \frac{f'(x)}{f(x)} \right| = \frac{\cos x}{\sin x}
\]

\[
= \left| \cos \pi \cdot \frac{\pi}{\sin \pi} \right| = \left| -1 \cdot \frac{\pi}{0} \right| \to \infty
\]

\( f = \sin \) ill-conditioned

\( x = 0 \)

\[
\text{cond}(f, 0) = \left| \frac{f'(0)}{f(0)} \right| = \left| \frac{0}{\sin 0} \right| = +1
\]

indeterminate form

"saved" by \( x = 0 \)

\[
\text{cond}(f, \pi) = \left| \frac{f'(\pi)}{f(\pi)} \right| = \left| \frac{\pi}{\sin \pi} \right| = \left| \frac{\pi}{0} \right| = +\infty
\]

not indeterminate form

why the difference?

\[\text{cond}(f, 0) = +1 \quad \text{well-conditioned} \]
\[\text{cond}(f, \pi) = +\infty \quad \text{ill-conditioned} \]