Math 373: 01 — Fall 2000
MW8 SC–205
Prof. Bumby

Steffensen Acceleration

Given a function g, for which we seek a *fixed point*, we define a new function

$$S(x) = x - \frac{\big(g(x) - x\big)^2}{g\big(g(x)\big) - 2g(x) + x}.$$

This fails to be defined if $g(g(x)) - 2g(x) + x = 0$, but otherwise, this definition gives $S(x) = x$ if $g(x) = x$. Our goal is to show that the process of iterating S is usually quadratically convergent, and to estimate the rate of convergence in terms of properties of the function g. We will also consider how the accuracy with which $g(x)$ can be computed affects the accuracy of the rootfinding by this method.

The main tool will be Taylor series about a fixed point of g which will be denoted x_∞. However, we don't want to specify in advance how many terms of the Taylor series are to be calculated, so we need some way to find the effect of computing more terms of a series in a computation without repeating the computation. This can be done by adding a subscript to the ξ in the Taylor series error term that will identify the series that was used in its computation. When the series is refined, most of the terms in the new series will be the same as the terms of the old series, and the remaining terms contain the known part of the error term as a factor. This should be clear enough when it is done even if it is awkward to describe in the abstract.

The first form of Taylor's theorem is the *Mean Value Theorem*. For the function g it takes the form

$$\begin{aligned} g(x) &= g(x_\infty) + g'(\xi_1)\cdot(x - x_\infty) \\ &= x_\infty + g'(\xi_1)\cdot(x - x_\infty) \\ g(x) - x_\infty &= g'(\xi_1)\cdot(x - x_\infty) \\ g(x) - x &= g'(\xi_1)\cdot(x - x_\infty) - (x - x_\infty) \\ &= \big(g'(\xi_1) - 1\big)\cdot(x - x_\infty) \end{aligned}$$

where the second form of the equation uses the assumption that x_∞ is a fixed point of g, and the other expressions use simple algebraic manipulation. Note that this expression shows that $g(x) - x$ and $x - x_\infty$ are of comparable size if g' is bounded away from 1 near x_∞. Since the iteration must stop if $g(x)$ and x cannot be distinguished, this shows that a convergent iteration will usually approximate the fixed point about as well as possible. Although we often pretend that the expressions that we write can be evaluated exactly, it is more realistic to say that one always has a fixed computational accuracy ϵ (typically about 10^{-10} on a calculator or 10^{-16} using the floating point registers of a computer, although better (i.e., smaller) values can be simulated in software) such that numbers x_0 and x_1 cannot be distinguished if $|x_0 - x_1| < \epsilon x_0$. When a difference of very close quantities appears in a computation, the difference is known to reduced relative accuracy because the accuracy of a computation is relative to the largest number appearing in the calculation, not to the answer.

Applying the same result with $g(x)$ in place of x gives

$$\begin{aligned} g\big(g(x)\big) - x_\infty &= g'(\xi_2)\cdot\big(g(x) - x_\infty\big) \\ &= g'(\xi_2)\cdot g'(\xi_1)\cdot(x - x_\infty) \end{aligned}$$

Thus

$$\begin{aligned}
g\big(g(x)\big) - 2g(x) + x &= \Big(g\big(g(x)\big) - x_\infty\Big) - 2\Big(g(x) - x_\infty\Big) + \Big(x - x_\infty\Big) \\
&= \big(g'(\xi_2)\cdot g'(\xi_1) - 2\cdot g'(\xi_1) + 1\big)\cdot(x - x_\infty) \\
&= \Big(\big(g'(\xi_2) - 1\big)\cdot g'(\xi_1) - \big(g'(\xi_1) - 1\big)\Big)\cdot(x - x_\infty)
\end{aligned}$$

From this one sees that, if $g'(\xi_1)$ and $g'(\xi_2)$ are both close to some number other than 1, the denominator in the calculation of $S(x)$ will be a *reasonable* multiple of $(x - x_\infty)$, so the denominator will only appear to be zero if the root has been found to the full accuracy allowed in our computation. In computing the numerator, one first finds $g(x) - x$, which also retains accuracy until the relative accuracy of the root is close to that allowed in the computation. Although this quantity becomes much smaller when squared, there will only be a small loss in *relative* accuracy in that computation. This suggests that this computation can be performed as long as one can detect any difference between x and $g(x)$.

Since we now have confidence that we can compute the expression for $S(x)$ as it is written, we can apply algebraic simplification to this expression to see what it is that we have computed.

$$\begin{aligned}
S(x) - x_\infty &= x - x_\infty - \frac{\big(g'(\xi_1)^2 - 2g'(\xi_1) + 1\big)\cdot\big(x - x_\infty\big)^2}{\big(g'(\xi_2)\cdot g'(\xi_1) - 2\cdot g'(\xi_1) + 1\big)\cdot\big(x - x_\infty\big)} \\
&= \frac{g'(\xi_2)\cdot g'(\xi_1) - g'(\xi_1)^2}{g'(\xi_2)\cdot g'(\xi_1) - 2\cdot g'(\xi_1) + 1}\cdot\big(x - x_\infty\big) \\
&= \frac{g'(\xi_1)\cdot\big(g'(\xi_2) - g'(\xi_1)\big)}{g'(\xi_2)\cdot g'(\xi_1) - 2\cdot g'(\xi_1) + 1}\cdot\big(x - x_\infty\big)
\end{aligned}$$

Although this could be used to identify functions g for which the corresponding S is contracting, it does not appear to lead to any useful criteria. However, the right side contains factors of both $x - x_\infty$ and $g'(\xi_2) - g'(\xi_1)$ indicating that quadratic convergence should be expected. To obtain such a bound, additional terms of the Taylor series of g are needed. From

$$\begin{aligned}
g(x) &= g(x_\infty) + g'(\xi_1)\cdot(x - x_\infty) \\
g(x) &= g(x_\infty) + g'(x_\infty)\cdot(x - x_\infty) + \frac{g''(\xi_3)}{2}(x - x_\infty)^2,
\end{aligned}$$

it follows that

$$g'(\xi_1) = g'(x_\infty) + \frac{g''(\xi_3)}{2}(x - x_\infty).$$

Similarly,

$$\begin{aligned}
g'(\xi_2) &= g'(x_\infty) + \frac{g''(\xi_4)}{2}(g(x) - x_\infty) \\
&= g'(x_\infty) + \frac{g''(\xi_4)}{2}g'(\xi_1)(x - x_\infty).
\end{aligned}$$

Thus,

$$g'(\xi_2) - g'(\xi_1) = \frac{1}{2}\big(g''(\xi_4)g'(\xi_1) - g''(\xi_3)\big)\cdot\big(x - x_\infty\big)$$

and this shows that $S(x) - x$ is the product of a bounded quantity and $(x - x_\infty)^2$ if g' is bounded away from 1 for the values needed in the calculation.

Recall that quadratic convergence gives a doubling of accuracy at each step when you get close enough to the fixed point. That is, a statement of the form

$$|S(x) - x_\infty| \leq M\,|x - x_\infty|$$

can be rewritten as

$$M\,|S(x) - x_\infty| \leq \big(M\,|x - x_\infty|\big)^2,$$

so that each iteration of S doubles the number of correct digits of x_∞ known *when measured in units of* $1/M$. For $S(x)$, the value of M becomes $\frac{g'(\xi)g''(\xi)}{2(g'(\xi)-1)}$ if we are confined to an interval where $\left|g''(\xi)\right|$ is small, which forces $g'(\xi)$ to change very slowly.

If the iteration of g is itself quadratically convergent, this analysis shows that the accelerated iteration S will have cubic convergence. This is misleading since it depends on two iterations of something that is quadratically convergent, and the second iterate has an error that is a *fourth power* of the original error. In this case the Steffensen extrapolation has slowed the rate of convergence.