A Pitfall in Using Newton's Method

David E. Dobbs

Introduction

Newton's Method (NM) is the most familiar numerical method for root approximation in a first course on calculus (Stewart 2001, 324–29). Not only is NM rather straightforward to implement, it is also often effective, yielding quadratic convergence to a root of a given function f under reasonably mild assumptions (Wade 1995, 206). In addition to providing practice with NM, calculus texts often mention pitfalls that may arise in using NM. For instance, Stewart (2001, 325) indicates that a sequence generated via NM may not converge (let alone, to a root). Additional pitfalls associated with NM appear in *Calculus-Concepts and Contexts*, Exercises 21–23 (Stewart 2001, 328), such as the "sequence" $\{x_n\}$ generated via NM stopping with its m-th term, provided that $f'(x_m) = 0$. Our main purpose here is to make explicit another pitfall that affects NM due to the fact that any computing device has an upper bound for the decimal place accuracy that it can calculate or report.

The pitfall in question concerns the following advice offered by Stewart (2001, 326) if one wishes to use NM to approximate a root of f to k decimal place accuracy: "The rule of thumb that is generally used is (to use x_n or x_{n+1} to approximate a root of f) if x_n and x_{n+1} agree to (at least) k decimal places." In this regard, we show in Examples 2.1 and 2.2 that the user of any computing device, working in conjunction with NM, can be utterly misled and defeated when following the above "rule of thumb," in the following sense. Suppose that we are given a differentiable function f, positive integers m and N, and a real number x_1 such that $f(x_1)$ is "small" in the sense that $|f(x_1)| \leq 10^{-m}$; and suppose that NM responds to input x_1 with output $x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$ satisfying $x_2 \approx x_1$ in the sense that $|x_2 - x_1| \leq 10^{-N}$. (To see how these assumptions relate to the above rule of thumb, consider $N > k$; consider also N such that the computing device at hand can report at most the first $N - 1$ decimal places, or nonzero significant digits, of calculated numbers.) Under these assumptions, is it certain that x_1 (or x_2) is "near" a root of f? Absolutely not! Indeed, given m and N as above, Examples 2.1 and 2.2 each produce a differentiable (in fact, polynomial) function f and a real number x_1 as above such that x_1 is arbitrarily far (say, at least 1 unit away) from any real root of f.

Rather than casting doubt on the effectiveness of NM (or any other numerical method for root approximation), results of the kind in Examples 2.1 and 2.2 are intended to make for better-informed users of technology, as one seeks to understand both the strengths and the inherent limitations of a particular numerical method. The material in this note could be used to enrich courses on calculus, real analysis, advanced calculus or numerical analysis.

Results

We begin by describing a simple construction that has all the desired properties.

Example 2.1. As in the introduction, let m and N be positive integers. Observe that if n and r are positive integers and c is a nonzero real number, then the n-th degree polynomial function $f(x) = c(x - r)^n$ has r as its only root (and r has multiplicity n as a root of f). Begin an application of NM by choosing $x_1 = 0$; this ensures (as desired) that $|x_1 - r| = r \geq 1$. Consider

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = -\frac{f(0)}{f'(0)} = -\frac{c(-r)^n}{cn(-r)^{n-1}} = \frac{r}{n}.$$

To arrange that $x_2 \approx x_1$ in the sense that $|x_2 - x_1| \leq 10^{-N}$, we need n such that $\frac{r}{n} \leq 10^{-N}$; that is, such that $n \geq 10^N r$. Does some such choice of n lead to $f(x_1)$ being "small" in the sense that $|f(x_1)| \leq 10^{-m}$, that is, such that $r^n c \leq 10^{-m}$? Certainly: given $(m, N,)$ r and $n \geq 10^N r$ as above, it suffices to take $c = r^{-n} 10^{-m}$.

It could be argued that the n-th degree polynomial in Example 2.1, having just one root (with multiplicity n), is somewhat pathological. Thus, one could ask for a polynomial f, which not only has the above properties but also only simple (that is, multiplicity 1) roots. Such an example is constructed in Example 2.2. As background, we recall two facts. The first of these is the Linear Factor Theorem (Dobbs and Hanks 1992, 39–40): if $r_1,\ldots,r_n$ are all the roots (counted with multiplicity) of an n-th degree polynomial and g and c is the leading coefficient of g, then $g(x) = c\prod_{i=1}^{n}(x - r_i)$. The second fact needed for Example 2.2 is the divergence of the harmonic series (see, for instance, [Stewart 2001, 577, Example 7] for a particularly accessible proof of this fact).

Example 2.2. Once again, let m and N be positive integers. Fix any positive number E (for "error"), with the requirement that our example must satisfy $|x_1 - s| \geq E$ for each root s of f. There is no loss of generality in also supposing that E is an integer. For convenience of notation, we once again choose $x_1 = 0$. Next come the three key steps. Since the harmonic series diverges, we can pick a positive integer n such that $\sum_{j=E}^{E+n-1} \frac{1}{j} \geq 10^N$.

Further, define

$$c = \frac{10^{-m}}{E(E+1)(E+2)\cdots(E+n-1)} = \frac{(E-1)!}{(E+n-1)!\,10^m}$$

and

$$f(x) = c(x-E)(x-E-1)\cdots(x-E-n+1) = c\prod_{j=E}^{E+n-1}(x-j).$$

We shall prove that the above construction has produced data x_1 and f with the desired properties.

By the Linear Factor Theorem, we can label the roots of f as $r_1 = E, r_2 = E + 1,\ldots, r_j = E + j - 1,\ldots, r_n = E + n - 1$. Notice that each r_j is a simple root of f and $\min\{|x_1 - r_j|: 1 \leq j \leq n\} = \min|-r_j| = E$. Moreover, $f(x_1)$ is approximately "small," since $|f(x_1)| = |f(0)| = |c(-E)(-E-1)\cdots(-E-n+1)| = |(-1)^n 10^{-m}| = 10^{-m}$.

It remains to explain why $x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} \approx x_1$ in the sense that $|x_2 - x_1| \leq 10^{-N}$. As in Example 2.1, our task is to verify that $\left|\frac{f(0)}{f'(0)}\right| \leq 10^{-N}$. It will be notationally easier to work with the polynomial

$$g(x) = c^{-1}f(x) = (x-E)(x-E-1)\cdots(x-E-n+1) = \prod_{j=E}^{E+n-1}(x-j).$$

As $g'(x) = c^{-1}f'(x)$, it follows that $\frac{g(x)}{g'(x)} = \frac{f(x)}{f'(x)}$, and so it will suffice to show that $\left|\frac{g(0)}{g'(0)}\right| \leq 10^{-N}$.

Since $g(x) = \prod_{j=1}^{n}(x - r_j)$, we can find $g'(x)$ by the Product Rule:

$$g'(x) = \sum_{i=1}^{n}\prod_{j\neq i}(x - r_j)$$

and so

$$\frac{g'(x)}{g(x)} = \sum_{i=1}^{n}\frac{1}{x-r_i} \text{ for } x \neq r_1,\ldots,r_n.$$

More to the point, we have that

$$\frac{g(x)}{g'(x)} = \frac{1}{\sum_{i=1}^{n}\frac{1}{x-r_i}} \text{ for } x \neq r_1,\ldots,r_n \text{ such that } g'(x) \neq 0.$$

In particular,

$$\left|\frac{g(0)}{g'(0)}\right| = \left|\frac{1}{\sum_{i=1}^{n}\frac{1}{-r_i}}\right| = \frac{1}{\sum_{j=E}^{E+n-1}\frac{1}{j}} \leq \frac{1}{10^N} = 10^{-N}$$

to complete the proof.

Remark 2.3. (a) The simplest example of a polynomial resulting from the construction in Example 2.1 is $f(x) = 10^{-m}(x-1)^{10N} = \frac{(x-1)^{10^N}}{10^m}$. Is it realistic to expect to be able to compute $f(x)$? Certainly, $f(0)$ can be computed, for we have supposed that it is known that $|f(0)| \leq 10^{-m}$. For any fixed x such that $0 < x < 2$, does $f(x)$ "look like" 0 to your computing device? The answer is yes if N is much larger than m, since $\lim_{N\to\infty}(x-1)^{10^N} = 0$. However, the answer is no if m is much larger than N (provided that your computing device can calculate the ratio of the two typically "small" numbers $(x-1)^{10^N}$ and 10^m). Thus, it is important to understand the relative sizes of N and m (as well as the relevant domain of x values) if one is to make practical use of Example 2.1.

(b) We next address similar practicality issues in regard to Example 2.2. The simplest example of a polynomial resulting from the construction in Example 2.2 is

$$f(x) = \frac{(x-1)(x-2)\cdots(x-n)}{n!\,10^m}.$$

The practicality of this construction is somewhat compromised by the fact that the harmonic series diverges notoriously slowly. For instance, most, if not all, of today's graphing calculators would fail to determine a suitable value of n in case $N = 2$.

What about the coefficients of $f(x)$? Is it realistic to expect your computing device to calculate these coefficients? Certainly, the constant coefficient, say k, of $f(x)$ presents no problem, since $k = (-1)^n 10^{-m}$ and we have supposed that 10^{-m} is known (as an upper bound for $|f(0)|$). Similarly, it is not difficult to show that d, the coefficient of x in $f(x)$, is given by

$$d = \frac{(-1)^{n-1}(1 + \frac{1}{2} + \ldots + \frac{1}{n})}{10^m}$$

and so $|d| \geq \frac{10^N}{10^m}$. Bounding the general coefficient of $f(x)$ would take us too far afield, but it is already clear from the above bound on $|d|$ that the practicality of Example 2.2, just like that of

Example 2.1, is impacted by the relative sizes of N and m.

In closing, we raise two questions. Is it possible to construct more effectively computable functions having the properties of the constructions in Examples 2.1 and 2.2? Are all numerical root approximation methods subject to the sort of pitfall that we have identified for NM?

References

Dobbs, D. E., and R. Hanks. *A Modern Course on the Theory of Equations*. 2d ed. Washington, N.J.: Polygonal Publishing House, 1992.

Stewart, J. *Calculus-Concepts and Contexts*. 2d ed. Pacific Grove, Calif.: Brooks/Cole, 2001.

Wade, W. R. *An Introduction to Analysis*. Englewood Cliffs, N.J.: Prentice Hall, 1995.

Two Problems from Isaac Newton (1642–1727)

A geometric sequence has three terms. The sum of these terms is 19, and the sum of their squares is 133. What are the terms of the sequence?

A geometric sequence has four terms. The sum of the outer terms (that is, the first and fourth) is 13 and the sum of the two middle terms (that is, the second and third) is 4. What are the terms of the sequence?