

Newton-Raphson Method

Appendix to *A Radical Approach to Real Analysis* 2nd edition
©2006 David M. Bressoud

June 20, 2006

A method for finding the roots of an “arbitrary” function that uses the derivative was first circulated by Isaac Newton in 1669. John Wallis published Newton’s method in 1685, and in 1690 Joseph Raphson (1648–1715) published an improved version, essentially the form in which we use it today. The idea is as follows. We are given a function:

$$f(x) = x^3 - 4x^2 - x + 2.$$

Even a rough graph shows that this function has three roots: near -1 , 1 , and 4 (see figure 1). Those are not the exact values of the roots. The value of $f(-1)$ is -2 . We shall have to move a little to the right of -1 . To see how far we should move, we note that the slope of our function at $x = -1$ is $f'(-1) = 10$. Our function does not have a constant slope between $(-1, -2)$ and the root we are seeking, but it looks as if the slope does not change by too much. If we increase x by $1/5$, the y value should increase by about 2 . Our next guess for the location of the root is at $x = -1 + 1/5 = -0.8$ (see figure 2).

Again, we are not quite there: $f(-0.8) = -0.272$. Our slope has changed slightly. At $x = -0.8$ it is only $f'(-0.8) = 7.32$. To increase the y value by 0.272 ($\Delta y = 0.272$), we need to increase x by approximately $0.272/7.32 = 0.037158$ [$\Delta x \approx \Delta y/f'(x)$]. The next guess is pretty good: $f(-0.8 + 0.037158) = -0.0088$, and so the root is close to $x = -0.762842$.

What we have found is an iterative procedure that will get us progressively closer to our root. If x_k was our last guess and neither $f(x_k)$ nor $f'(x_k)$ is zero, then the next approximation will be

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}. \quad (1)$$

This kind of iteration is easily programmed. Starting with $x_1 = -1$, the successive iterations (with ten-digit accuracy) are

$$\begin{aligned} x_2 &= -0.8, \\ x_3 &= -0.7628415301, \\ x_4 &= -0.7615586962, \\ x_5 &= -0.7615571818, \\ x_6 &= -0.7615571818. \end{aligned}$$

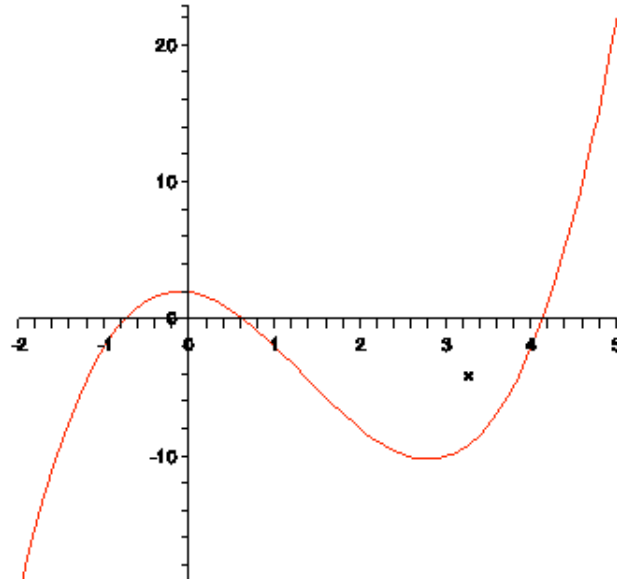


Figure 1: $f(x) = x^3 - 4x^2 - x + 2$.

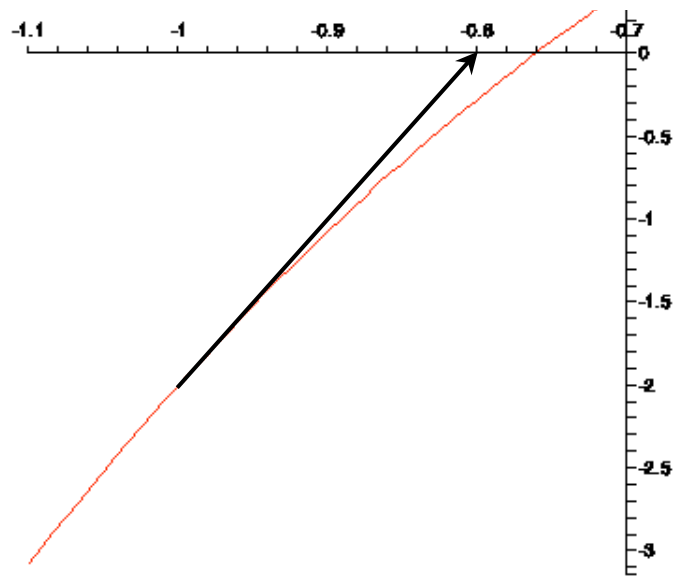
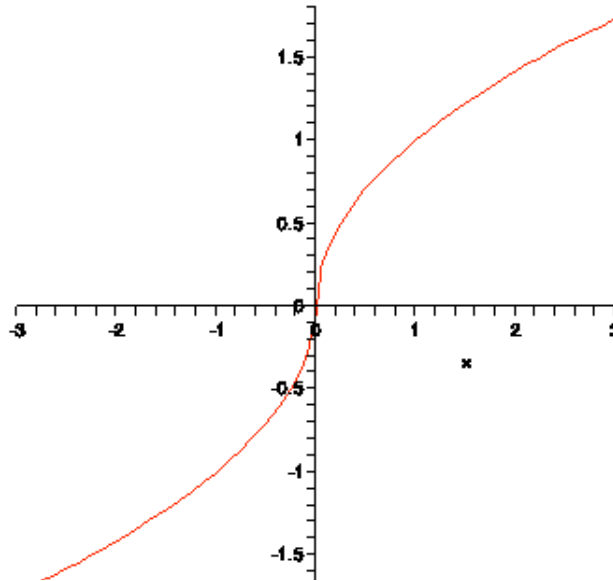


Figure 2: A close-up of $f(x) = x^3 - 4x^2 - x + 2$.

Figure 3: $f(x) = x/\sqrt{|x|}$.

This is as close as we are going to get to the root using a ten-digit decimal approximation. This is the **Newton–Raphson method**.

What is wrong with Newton–Raphson

Most of the time, Newton–Raphson converges very quickly to the root. To find the other two roots, we just start closer to them. But it does not always work. Consider the function defined by

$$f(x) = \frac{x}{\sqrt{|x|}}, \quad x \neq 0,$$

and define $f(0) = 0$ so that it is continuous (see figure 3). The derivative of this function is

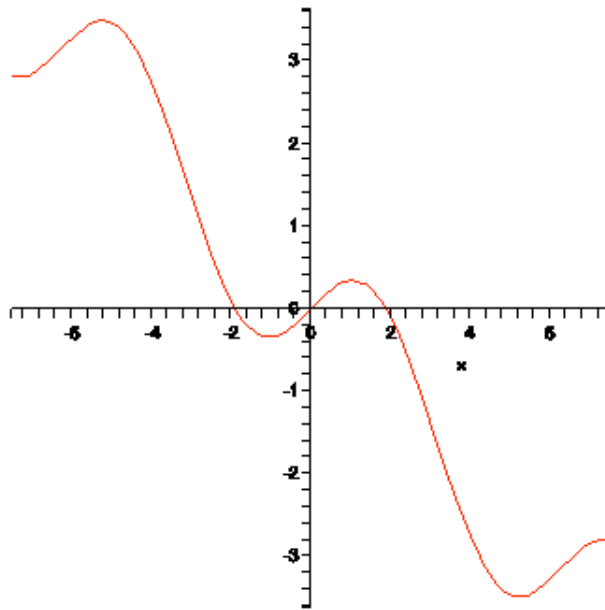
$$f'(x) = \frac{1}{2\sqrt{|x|}}, \quad x \neq 0.$$

If we choose *any* starting point off the actual root, $x_1 = a \neq 0$, then

$$x_2 = a - \frac{a/\sqrt{|a|}}{1/2\sqrt{|a|}} = a - 2a = -a.$$

It follows that $x_3 = -x_2 = a$, $x_4 = -a$, $x_5 = a$, and we keep bouncing back and forth forever.

This function is rather unusual. A more common occurrence is that Newton–Raphson works for some choices of starting point but not for others, and when it does work it does not necessarily take you to the closest root. We consider the function defined by (see figure 4)

Figure 4: $f(x) = \sin x - x/2$.

$$f(x) = \sin x - x/2, \quad f'(x) = \cos x - 1/2.$$

This has three roots. One of them is at $x = 0$. The others are near $x = \pm 2$.

For this function, the Newton–Raphson method uses the iteration

$$x_{k+1} = x_k - \frac{\sin x_k - x_k/2}{\cos x_k - 0.5}.$$

If we start with $x_1 = 2$, we quickly approach the rightmost root:

$$\begin{aligned} x_1 &= 2, \\ x_2 &= 1.900995594, \\ x_3 &= 1.898679953, \\ x_4 &= 1.895500113, \\ x_5 &= 1.895494267, \\ x_6 &= 1.895494267. \end{aligned}$$

To within ten digits of accuracy, we have found the positive root of $\sin x - x/2$.

What happens if our initial guess is not so accurate? What happens if we start with $x_1 = 1$? The reader is encouraged to perform these calculations. *They are very sensitive to internal roundoff*

and may vary from one machine to the next.

$$\begin{aligned}
 x_1 &= 1, \\
 x_2 &= -7.47274064, \\
 x_3 &= 14.47852099, \\
 x_4 &= 6.935115429, \\
 x_5 &= 16.63568455, \\
 x_6 &= 8.34393498, \\
 x_7 &= 4.9546244809, \\
 x_8 &= -8.300926412, \\
 x_9 &= -4.816000398, \\
 x_{10} &= 3.764084222, \\
 x_{11} &= 1.88580759, \\
 x_{12} &= 1.895549291, \\
 x_{13} &= 1.895494269.
 \end{aligned}$$

We have again found the rightmost root, but it took considerably longer and we did a lot of bouncing around en route. In fact, starting just a little closer, at $x = 1.1$, my calculator takes me as low as $x_7 = -205.8890057$ and as high as $x_{16} = 323.2663795$ before settling down to the root at $x = 0$.

What if we had started between these two points? If $x_1 = 1.01$, my calculator takes me to the negative root, -1.895494267 . If $x_1 = 1.02$, I move very far away from the origin, eventually reaching a negative number that exceeds the internal storage capacity. Let us call that destination $-\infty$. The table given below lists the terminal values of the Newton–Raphson method with $x_1 = 1.00, 1.01, 1.02, \dots, 1.10$ as found on my calculator. Try this on your own machine. *Do not expect your answers to be the same.*

x_1	terminal value
1.00	1.895494267
1.01	-1.895494267
1.02	$-\infty$
1.03	-1.895494267
1.04	-1.895494267
1.05	$+\infty$
1.06	$-\infty$
1.07	1.895494267
1.08	0
1.09	-1.895494267
1.10	0

This is a good example of **chaos**: extreme sensitivity to initial conditions and machine round-off error.

This does not mean that the Newton–Raphson method is no good. Even today it is one of the most useful and powerful tools available for finding roots. But as we have seen, it can have problems. We need further analysis of how and why it works if we want to determine when we can use it safely and when we must proceed with caution.

What is Happening

Let r be the actual though unknown value of a root that we are trying to approach, and let x_k be our latest guess. We assume that it is “pretty close.” We can calculate $f(x_k)$ and $f'(x_k)$. While we do not know the value of r , the fact that it is a root implies that $f(r) = 0$. Since x_k is “pretty close” to r , $f'(x_k)$ will be “pretty close” to the slope of the line from $(x_k, f(x_k))$ to $(r, 0)$:

$$f'(x_k) \approx \frac{f(r) - f(x_k)}{r - x_k} = \frac{-f(x_k)}{r - x_k}. \quad (2)$$

We can solve this to get something that is “pretty close” to r :

$$r \approx x_k - \frac{f(x_k)}{f'(x_k)}. \quad (3)$$

As our last example showed, “pretty close” can sometimes be not close at all. What do we mean by “pretty close”? Equation (2) shows that we are using an approximation to the derivative. Equation (3) uses this approximation in the *denominator*, and there lies the crux of our problem. While 0.01 and 0.0001 are close to each other, $1/0.01 = 100$ and $1/0.0001 = 10000$ are *not* close. It is not enough to say “pretty close.” We need to know the size of the error in equation (2). Lagrange’s remainder for the Taylor series can help us.

We use the equality

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(c)}{2!}(x - a)^2, \quad (4)$$

where c is some unknown constant between a and x . If we solve for $f'(a)$ we see that this equation is equivalent to

$$f'(a) = \frac{f(x) - f(a)}{x - a} - \frac{f''(c)}{2!}(x - a). \quad (5)$$

The error is precisely $-f''(c)(x - a)/2$. Although we do not know the value of c , it may be possible to find bounds on $f''(c)$ when c is between a and x , and thus find bounds on the error.

We replace a with x_k and x with r , and then solve for r , keeping the $r - x_k$ term in the error:

$$\begin{aligned} f'(x_k) &= \frac{-f(x_k)}{r - x_k} - \frac{f''(c)}{2!}(r - x_k), \\ r - x_k &= \frac{-f(x_k)}{f'(x_k)} - \frac{f''(c)}{2f'(x_k)}(r - x_k)^2, \\ r &= x_k - \frac{f(x_k)}{f'(x_k)} - \frac{f''(c)}{2f'(x_k)}(r - x_k)^2 \\ &= x_{k+1} - \frac{f''(c)}{2f'(x_k)}(r - x_k)^2. \end{aligned} \quad (6)$$

Are we getting closer?

We now have to decide what we want to mean when we say we are getting “pretty close” to r . A reasonable criterion that is easy to use is to ask that x_{k+1} be strictly closer to r than x_k was. Equation (6) gives us a relationship between $|r - x_{k+1}|$ and $|r - x_k|$:

$$|r - x_{k+1}| = \left| \frac{f''(c)}{2f'(x_k)} \right| |r - x_k|^2. \quad (7)$$

We shall get closer to r provided

$$\left| \frac{f''(c)}{2f'(x_k)} \right| |r - x_k| < 1,$$

or, equivalently,

$$|f''(c)| < 2 \left| \frac{f'(x_k)}{r - x_k} \right|. \quad (8)$$

For the example we have been using, $f(x) = \sin x - x/2$, we want

$$|\sin c| < 2 \left| \frac{\cos x_k - 0.5}{r - x_k} \right|. \quad (9)$$

Since we do not know how c depends on x_k and r , we do not know the exact values of x_k for which Newton–Raphson gets progressively closer to r , but this inequality is satisfied if the right side is larger than 1. If we graph the right side of inequality (9) near (but not at) the root $r = 1.89\dots$ (figure 5), we see that it is larger than 1 for $1.35 \leq x \leq 4.07$. If we start with any point in this range, we are guaranteed to converge to the positive root.

Exercises

The symbol **M&M** indicates that *Maple* and *Mathematica* codes for this problem are available in the **Web Resources** at www.macalester.edu/aratra, Chapter 3.

1. Lagrange’s form of the remainder tells us that if we use $1 + 2/1 + 2^2/2! + 2^3/3! + \dots + 2^{k-1}/(k-1)!$ to approximate e^2 , then the difference between the true value and this approximation is

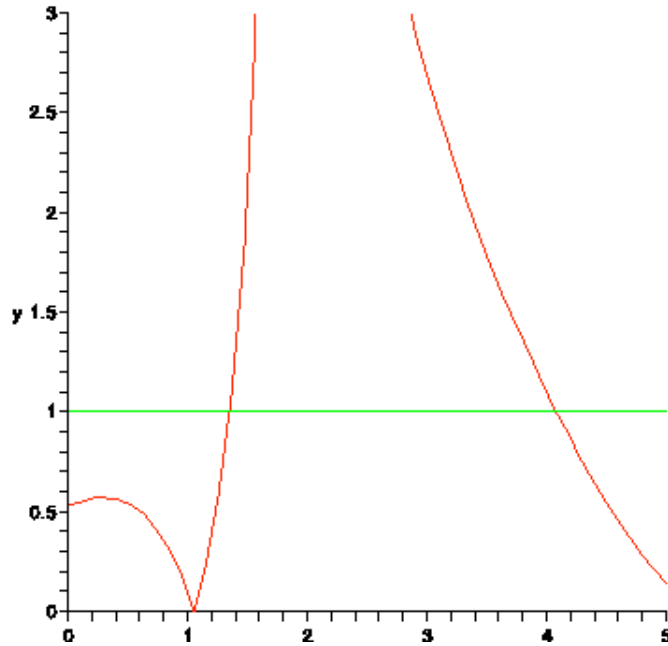
$$E(k) = e^2 - \left(1 + \frac{2}{1} + \frac{4}{2!} + \dots + \frac{2^{k-1}}{(k-1)!} \right) = \frac{2^k e^c}{k!}$$

for some c between 0 and 2. Stirling’s formula tells us that $k! > k^k e^{-k} \sqrt{2\pi k}$, and therefore the error is less than

$$\frac{2^k e^2}{k!} < \left(\frac{2e}{k} \right)^k \frac{e^2}{\sqrt{2\pi k}}.$$

Using this bound, find a small value of k that guarantees an error of less than 0.001, an error of less than 0.000001.

2. **M&M** To ten-digit accuracy, find the other two roots of $x^3 - 4x^2 - x + 2$.

Figure 5: $2|\cos x - 0.5|/|1.89 - x|$.

3. **(M&M)** For the function $f(x) = \sin x - x/2$, run ten iterations of the Newton–Raphson method at every value of x from $x = 0.75$ to $x = 1.25$ in steps of 0.01. The outcome will suggest which initial values between 0.75 and 1.25 wind up at which roots. In regions of transition, you may want to decrease the step size and increase the number of iterations.

4. **(M&M)** The Newton–Raphson method is equally valid for complex-valued functions. Compare the results of the previous exercise with the results obtained if x ranges from $0.75 + 0.01i$ to $1.25 + 0.01i$.

5. Prove that if $f(x) = x/\sqrt{|x|}$, $x \neq 0$, then $f'(x) = 1/2\sqrt{|x|}$, $x \neq 0$.

6. The example $f(x) = x/\sqrt{|x|}$ works the way it does because this function satisfies the differential equation

$$\frac{f(x)}{f'(x)} = 2x.$$

Describe all functions that satisfy this differential equation.

7. Describe all functions that satisfy the differential equation

$$\frac{f(x)}{f'(x)} = 3x.$$

Choose one of them and implement the Newton–Raphson method on it. What happens? Discuss why it happens.

8. Describe all complex-valued functions that satisfy the differential equation

$$\frac{f(x)}{f'(x)} = (1 - i)x.$$

Choose one of them and implement the Newton–Raphson method on it. What happens? Discuss why it happens.

9. (M&M) Use the Newton–Raphson method to find all the roots (to within ten-digit accuracy) of $f(x) = \cos(3x) - x$.

10. There is one positive root for $f(x) = \cos(3x) - x$. Using the methods of this section, find an interval around this root such that if Newton–Raphson is initiated at any point in this interval, each iteration will take you closer to the positive root.

11. For the function $f(x) = \sin x - x/2$, find an interval containing the origin and as large as reasonably possible so that if we start the Newton–Raphson method with any x in this interval, the iterations will always converge to 0.

12. Prove that if we have an interval I around a root r and a positive number $\alpha < 1$ such that

$$\left| \frac{f''(c)}{2f'(x)} \right| |r - x| \leq \alpha$$

for all $x \in I$ and c between x and r , then the Newton–Raphson method *must* converge to r .

13. Is it possible that from some starting point each iteration of Newton–Raphson takes us closer to the root r and yet from this starting point we can never get arbitrarily close to the value of r ? Explain your answer.