THE COMPLEX NEWTON METHOD—3 DIFFERENT WAYS

MARTIN PERGLER

A well-known heuristic for approximating roots of a polynomial or other "nice function" over **R** or **C** is Newton's method:

Newton's method—Version 1 (formula)**.** *Given a differentiable* $f : \mathbf{R} \to \mathbf{R}$*, or an analytic* $f : \mathbf{C} \to \mathbf{C}$*, the Newton iteration function is*

$$N(x_0) = x_0 - f(x_0)/f'(x_0) \tag{1}$$

or given implicitly by

$$-f(x_0) = f'(x_0)(N(x_0) - x_0)$$

Given any initial guess x_0, we iterate the process to obtain a sequence

$$x_n = N_n(x_0) = (\underbrace{N \circ N \circ \cdots \circ N}_{n \text{ times}})(x)\,(= N(x_{n-1}))$$

If we are lucky, then $N_\infty(x_0) = \lim_{n\to\infty} N_n(x_0)$ exists. In this case, its value is a root of f.

There are various questions that come to mind:

1. Where does formula (1) come from?
2. What does the graph of the limit function N_∞ look like?
3. At what points (initial guesses) and why does the limit function N_∞ fail to exist?

The third question is certainly interesting: clearly there is a problem whenever $f'(x_0) = 0$, but there are many other things that can go wrong, including cycles in the sequence x_n and issues of convergence. This falls within the scope of the study of dynamical systems, and there is actually a fairly extensive qualitative understanding of fixed points, attractors, and cycles of iterated maps over **R** (and to a lesser extent over **C**) in general. However, for Newton's method applied to "reasonable" functions f, the points at which such pathologies occur tend to be isolated and so if one is merely interested in finding any root of f, these considerations are not that relevant.

The second question gives rise to fascinating colourings of the complex plane, which arise as "basin of attraction" graphs for analytic f as simple as $f(z) = z^3 - 1$. These are among the second rank of "most famous" fractal images, after the Mandelbrot set. The situation turns out the be pretty interesting over **R** as well.

The first question is the focus of this note.

1. What does it mean?

Over **R**, the derivation of Newton's method can be found in almost any elementary calculus textbook. The formula (1) is what follows from the following geometric principle:

Date: December 18, 1999.

Newton's method—Version 2 (geometric description)**.** *To obtain $N(x_0)$, replace the graph of $f(x)$ by its tangent line T_{x_0} at $x = x_0$. Let $N(x_0)$ be the point of intersection of T_{x_0} with the x-axis.*

For x close to x_0, T_{x_0} should be a "good" approximation to $f(x)$, so if $N(x_0)$ is close to x_0, $N(x_0)$ should be close to a root of f. The heuristic is that even if $N(x_0)$ is not terribly close to x_0, perhaps $N(x_0)$ is a better approximation to a root of f than x_0 is, and in particular hopefully $N_n(x_0)$ converges to a root $N_\infty(x_0)$.

The slope of T_{x_0} is of course $f'(x_0)$, the limit of the slopes of secant lines through x_0. Thus T_{x_0} consists of points (x,y) which satisfy the equation

$$y - f(x_0) = f'(x_0)(x - x_0).$$

In particular, $(N(x_0), 0)$ is a solution of this equation.

What about the case over **C**? It is hard to make Version 2 make sense with x_0 and $f(x_0)$ complex numbers. Since $\mathbf{C} = \mathbf{R} \times \mathbf{R}$, we now have a picture in 4 dimensions, and the meaning of a tangent "line" T_{x_0} is unclear. If f is analytic, we can calculate $f'(x_0)$ and hence Version 1 makes formal sense. But does it mean anything other than formal manipulation? We show the answer is yes. Namely, we generalize the geometric construction of Version 2 over $\mathbf{C} = \mathbf{R} \times \mathbf{R}$ in two different, natural ways. It turns out that precisely when $f'(x_0)$ exists, these constructions are equivalent, and give the same "answer" (=definition of $N(x_0)$) as Version 1. To do this we present generalizations of Newton's method to vector functions (over **R**).

2. A SINGLE FUNCTIONS OF TWO REAL VARIABLES

Consider a single function $g : \mathbf{R}^2 \to \mathbf{R}$. We apply the reasoning of Version 2 of Newton's method and see what happens.

The graph of g is now a surface in $\mathbf{R}^3$. We let x_0 (now a 2-vector) be a guess for a root of g. We replace the graph of $g(x)$ by its tangent plane T_{x_0} at x_0. This tangent plane consists of points (x,y) (where x is a 2-vector and y a scalar) such that

$$y - g(x_0) = \nabla g(x_0) \cdot (x - x_0).$$

Here ∇g is the gradient of g and on the right we are taking the dot product of vectors. We seek solutions of the form $(x, 0)$, or x such that

$$-g(x_0) = \nabla g(x_0) \cdot (x - x_0). \tag{2}$$

However, now we see that there is no unique solution which we could call $N(x_0)$. In fact, unless $\nabla g = 0$, in which case we have the same problem as before, the solution set consists exactly of the (1-dimensional) straight line L_{x_0} in which T_{x_0} intersects the x-plane. Let x^* and x^{**} be two points on L_{x_0}. By plugging into equation (2) and subtracting, we see that $\nabla g(x_0) \cdot (x^* - x^{**}) = 0$. In other words, L_{x_0} is perpendicular to $\nabla g(x_0)$.

Which x on L_{x_0} do we choose for $N(x_0)$? Without further information, any one should work equally well. It seems reasonable to choose the closest one. In this case, $N(x_0) - x_0 = \alpha \nabla g(x_0)$ for some α. Substitution in equation (2) yields

$$\alpha = -g(x_0)/||\nabla g(x_0)||^2 \quad \text{or} \quad N(x_0) - x_0 = -g(x_0)\nabla g(x_0)/||\nabla g(x_0)||^2.$$

We remark that $\nabla g(x_0)$ points in the direction of steepest ascent, $N(x_0) - x_0$ points opposite to it if $g(x_0)$ is positive, and $||N(x_0) - x_0|| = g(x_0)/||\nabla g(x_0)||$ is minimal. We remark that this process works just as well in higher dimension, and thus we have

Newton's method—Version 3 (steepest descent)**.** *Suppose $g : \mathbf{R}^k \to \mathbf{R}$. The Newton iteration function is*

$$N(x_0) = x_0 - \frac{g(x_0)}{||\nabla g(x_0)||}\,\mathrm{direction}(\nabla g(x_0)). \tag{3}$$

3. Two functions of two variables

In our case, $f : \mathbf{C} \to \mathbf{C}$ is actually a function $f : \mathbf{R}^2 \to \mathbf{R}^2$. Convention dictates that we label the variable z and real and complex parts of f by $u = u(z)$ and $v = v(z)$. We seek a common root of u and v. Thus $N(z_0)$ should satisfy equation (2) with both $g = u$ and $g = v$. We can encode this using the Jacobian matrix

$$Df(z_0) = \begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix}$$

using which the two instances of equation (2) can be written together as

$$-f(z_0) = Df(z_0)(N(z) - z_0).$$

Here z_0, $N(z_0)$ and $f(z_0)$ are taken as real column 2-vectors, and the right hand side is a matrix product. If $Df(z_0)$ is nonsingular, there is now a unique solution. We state it in greater generality, in higher dimension and forgetting the complex structure:

Newton's method—Version 4 (simultaneous equations)**.** *Suppose $f : \mathbf{R}^k \to \mathbf{R}^k$ and let $Df = \left(\partial f_i/\partial x_j\right)_{i,j}$ be its Jacobian. The Newton iteration function is*

$$N(x_0) = x_0 - (Df(x_0))^{-1} f(x_0) \tag{4}$$

(the Jacobian must be nonsingular).

The above Version of Newton's method is the standard higher-dimensional generalization of Version 1, found in some advanced calculus textbooks, but not as well known as it perhaps should be. Further generalizations in this vein are of use in dynamical systems for more general operators.

We remark that the Jacobian Dg of a scalar function $g : \mathbf{R}^k \to \mathbf{R}$ is a row matrix, the transpose of the column vector ∇g.

An alternative approach to Version 4 is to apply Version 3 to $||f||$. We continue considering $f : \mathbf{R}^k \to \mathbf{R}^k$ and easily verify using the chain rule that

$$||f||\nabla||f|| = (f^T Df)^T = (Df)^T f, \tag{5}$$

and hence we get

Newton's method—Version 5 (norm steepest descent)**.** *Suppose $f : \mathbf{R}^k \to \mathbf{R}^k$ and let $Df = \left(\partial f_i/\partial x_j\right)_{i,j}$ be its Jacobian. The Newton iteration function is*

$$N(x_0) = x_0 - \frac{||f||^2}{||(Df)^T f||^2}(Df)^T f \text{ (evaluated at } x_0\text{)}. \tag{6}$$

4. Application to complex functions

We now show that if f is a complex analytic function, then Versions 1, 4, and 5 give the same formula for $N(z_0)$. We blithely and without warning switch back an forth from considering f as a scalar function $f : \mathbf{C} \to \mathbf{C}$ and as a vector function $f : \mathbf{R}^2 \to \mathbf{R}^2$, and cheerfully exchange the notations $|\cdot|$ and $||\cdot||$ for modulus and norm.

Recall that a $k \times k$ real matrix M is orthogonal if $M^T M = \mathrm{Id}$. We'll call M *quasi-orthogonal* if $M^T M$ is any scalar matrix, i.e. a scalar multiple of Id. It is easy to see that this happens precisely when M is some scalar r times an orthogonal matrix, and then $M^T M = r^2\,\mathrm{Id}$. We now use these two delightful but not-very-well known facts:

Lemma 1. *There is a one-to-one correspondence between complex numbers z and 2×2 quasi-orthogonal matrixes M, given by*

$$x + iy \longleftrightarrow \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$
$$re^{i\theta} \longleftrightarrow \begin{pmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{pmatrix}$$

Furthermore, under this correspondence, if w is any other complex number, then $zw = Mw$. Here we have complex multiplication on the left, and matrix multiplication on the right, with w a column 2-vector.

Lemma 2. *If $f : \mathbf{C} \to \mathbf{C}$ is a function, then $f'(z_0)$ exists precisely when the Jacobian $Df(z_0)$ is a quasi-orthogonal matrix, and then $f'(z_0)$ is the complex number to which it corresponds.*

The first of these Lemmas is easily verfied by computation. In the conventional treatment of complex analysis, the second Lemma is a restatement of the Cauchy-Riemann equations. However, in Tristan Needham's wonderful exposition of complex analysis from a geometric vein [**?**], this Lemma is taken as the defining property of complex differentiability. The property that the Jacobian Df be quasi-orthogonal (which Needham calls *amplitwist*) means precisely that level curves of $u = \Re(f)$ and $v = \Im(f)$ are orthogonal, or that $f : \mathbf{R}^2 \to \mathbf{R}^2$ takes infinitesimal squares to infinitesimal squares. The local rotation and expansion (or "twisting" and "amplification") of f near z_0 are encoded in the polar representation of the complex number $f'(z_0)$.

These two Lemmas immediately show that Versions 1 and 4 of Newton's method are identical. Next we show that Versions 4 and 5 are identical precisely when Df is quasi-orthogonal. Let ρ be the determinant of $Df(z_0)$ and write $Df(z_0) = \rho M$, where M is a matrix of determinant 1. Then equations (4) and (6) become

$$N(z_0) = z_0 - (1/\rho)\, M^{-1} f, \text{ and} \tag{7}$$

$$N(z_0) = z_0 - \frac{||f||^2}{\rho ||M^T f||^2}\, M^T f, \tag{8}$$

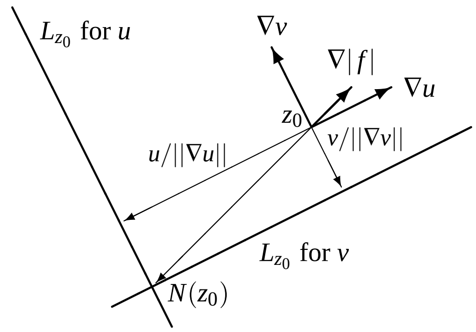
which are identical precisely when $M^{-1} = M^T$, i.e. when M is orthogonal.

Figure 1 summarizes the geometry of computing $N(z_0)$ when f is complex analytic.

5. Remarks

1. It is possible to consider Newton's method for functions $f : \mathbf{C} \to \mathbf{C}$ which are not complex analytic. However, there are different ways of doing it, in particular Versions 4 and 5. In fact, these methods work for arbitrary nicely-differentiable functions $f : \mathbf{R}^k \to \mathbf{R}^k$, and it would be interesting to see what sorts of interesting fractal pictures arise from "simple" functions of this sort somehow analogous to $f(z) = z^m - 1$.

FIGURE 1. Determination of $N(z_0)$. Diagram shows the z_0-plane only, and assumes $u = u(z_0)$ and $v = v(z_0)$ are both > 0.



2. Our discussion shows that for complex analytic f, $f(z)/f'(z)$, considered as a vector, points in the direction of $\nabla|f|$, something which does not seem *a priori* obvious. In fact, if M is the matrix corresponding by Lemma 1 to a complex number z, then the matrix corresponding to $\bar{z}$ is M^T. Thus we can rewrite equation (5) as
$$|f|\nabla|f| = \overline{Df}\, f = \overline{f'}\, f$$
(matrix multiplication in the middle, complex multiplication on the right) and in particular
$$\text{angle}(\nabla|f|) = \text{argument}(f) - \text{argument}(f').$$
3. One could apply Version 3 to $|f|^2$ or $|f|^\alpha$ for any $\alpha > 0$. Calculation then shows that (provided f is complex analytic), we get
$$N(z_0) = z_0 - (1/\alpha)f(z_0)/f'(z_0).$$
For $\alpha > 1$ this has been considered (purely formally) as a so-called "damped Newton method" (see [**?**]). This shows it has geometric foundation.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, CHICAGO IL 60637
E-mail address: `pergler@math.uchicago.edu`