Newton’s method

Martin Licht

UC San Diego

Winter Quarter 2021

We want to study the solution of general nonlinear equations

$$G(x) = y.$$

This is typically formulated as the problem of finding the ***root*** of a function, namely, we define $F(x) = G(x) - y$, so we can equivalently ask the solution of

$$F(x) = 0.$$

Because of that argument, solving nonlinear equations is generally (re)formulated as solving $F(x) = 0$ for some x.

Recurring traits in the algorithmic solution of general nonlinear equations:

- ***Iterative algorithms:*** starting with an initial guess x_0, the algorithms construct successive approximations $x_1, x_2, x_3, \ldots$ of the true solution x_*.
- Ideally, $x_1, x_2, x_3, \ldots$ converges to x_*.
- Computing x_* directly generally infeasible.

Newton’s Method in One Variable

Newton’s Method in Several Variables

Convergence Theory for the Newton Method

Newton's Method in One Variable

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function with some unknown root $x_* \in \mathbb{R}$.

So $f(x_*) = 0$.

We will describe a method that starts with an initial approximate root $x_{(0)}$ and constructs a sequence $x_{(1)}, x_{(2)}, x_{(3)}, \ldots$ of approximate roots.

We hope that these approximate roots will converge to the actual root x_*.

Newton's Method in One Variable

Suppose that we already have an approximate root $x_{(0)} \in \mathbb{R}$ that is reasonably close to x_*.

We use the first-order Taylor polynomial around $x_{(0)}$ to approximate f:

$$f(x) \approx f(x_{(0)}) + f'(x_{(0)})(x - x_{(0)}).$$

If we plug in x_*, then $f(x_*) = 0$ gives the approximate relation

$$0 \approx f(x_{(0)}) + f'(x_{(0)})(x_* - x_{(0)}).$$

Assuming that $f'(x_{(0)}) \neq 0$, we rearrange this "equation" and thus find

$$x_* \approx x_{(0)} - f'(x_{(0)})^{-1} f(x_{(0)}).$$

Consequently, we use

$$x_{(1)} := x_{(0)} - f'(x_{(0)})^{-1} f(x_{(0)}).$$

as the next approximate root. Repeating this procedure leads to a sequence

$$\begin{aligned} x_{(1)} &:= x_{(0)} - f'(x_{(0)})^{-1} f(x_{(0)}), \\ x_{(2)} &:= x_{(1)} - f'(x_{(1)})^{-1} f(x_{(1)}), \\ x_{(3)} &:= x_{(2)} - f'(x_{(2)})^{-1} f(x_{(2)}), \\ &\vdots \end{aligned}$$

of approximate roots $x_{(0)}, x_{(1)}, x_{(2)}, \ldots$ that obey the recursive formula

$$x_{(k+1)} := x_{(k)} - f'(x_{(k)})^{-1} f(x_{(k)}).$$

Newton's Method in One Variable

The Newton method is defined by iteration procedure.

$$x_{(k+1)} := x_{(k)} - f'(x_{(k)})^{-1} f(x_{(k)}).$$

The Newton method is defined by iteration procedure.

$$x_{(k+1)} := x_{(k)} - f'(x_{(k)})^{-1} f(x_{(k)}).$$

Remark (Alternative justification)

Instead of solving the problem $f(x) = 0$, we try to solve the "approximate" problem

$$f(x_{(k)}) - f'(x_{(k)})(x - x_{(k)}) = 0.$$

at each step. We hope that the solution to the approximate problem is close to the original problem's solution.

Newton’s Method in One Variable

Note that this requires the function f to be differentiable.

Furthermore, we need that $f'(x_{(k)}) \neq 0$ because otherwise we cannot continue the construction.

And last but not least, even if the sequence is always defined, it is not yet guaranteed that the method converges to the root at all.

Newton's Method in One Variable

Example

Let us have a look at the function $f(x) = x^2 - 2$. Then $f(x) = 0$ if and only if x is a square root of 2.

We have $f'(x) = 2x$. Consequently, the update formula is

$$\begin{aligned} x_{(k+1)} &= x_{(k)} - \frac{(x_{(k)})^2 - 2}{2x_{(k)}} \\ &= \frac{1}{2}\left(x_{(k)} + \frac{2}{x_{(k)}}\right). \end{aligned}$$

This is known as ***Heron's method*** for computing the square root of 2. ***Heron of Alexandria*** was a Greek mathematician who lived around 60 AD and left an explicit account of this method. It is known as ***Babylonian method*** because allegedly the Babylonian's utilized this method.

Newton's Method in One Variable

Example

Suppose we start with $x_{(0)} = 2$ as an approximation of $\sqrt{2}$. We get

$$
\begin{aligned}
x_{(0)} &= 2,\\
x_{(1)} &= \frac{3}{2} = 1.5,\\
x_{(2)} &= 17/12 = 1.4166\overline{6},\\
x_{(3)} &= 577/408 = 1.41421\overline{5686274509803921},\\
x_{(4)} &= 665857/470832 \approx 1.4142135623746899106262955788901 3.....,
\end{aligned}
$$

After four calculation steps, we already have a very good approximation of $x_* = 1.4142135623730950488 0....$

Newton's Method in One Variable

However, there are many examples where Newton's method does not converge. The iteration may diverge to infinity or get trapped in a cycle.

Example

Consider the function $f(x) = \arctan(x)$, so

$$f'(x) = (1 + x^2)^{-1}.$$

This function has a root at $x = 0$. There exists a number

$$R \approx 1.3917...$$

such that starting Newton's method with any $x_{(0)} \in \mathbb{R}$ has the following behavior:

- If $|x_{(0)}| < R$, then the iteration converges to zero.
- If $|x_{(0)}| = R$, then the iteration cycles between R and $-R$.
- If $|x_{(0)}| > R$, then the iteration diverges to infinity, switching between positive and negative values.

Newton's Method in One Variable

Example

We might accidentally hit a plateau of the function, so that Newton's method is no longer defined.

Applying Newton's method to find a root of the cubic polynomial

$$f(x) = x^3 - x^2 + 1$$

with $x_{(0)} = 1$, we calculate

$$f'(x) = 3x^2 - 2x$$

and get

$$x_{(1)} = x_{(0)} - f(x_{(0)})/f'(x_{(0)}) = 1 - 1 = 0.$$

But $f'(0) = f'(x_{(1)}) = 0$, so Newton's method is ill-defined from there on.

Newton's Method in One Variable

Example

Consider the function $f(x) = x^3 - 5x$. Clearly, this function has three different roots.

We have $f'(x) = 3x^2 - 5$.

Newton's method with starting value $x_{(0)} = 1$ produces

$$\begin{aligned} x_{(1)} &= x_{(0)} - f(x_{(0)})/f'(x_{(0)}) = 1 - (-4)/(-2) = -1, \\ x_{(2)} &= x_{(1)} - f(x_{(1)})/f'(x_{(1)}) = (-1) - (4)/(-2) = 1. \end{aligned}$$

So we get stuck in a loop $1, -1, 1, -1, \ldots$ while never getting closer to any of the roots 0, $\sqrt{5}$ or $-\sqrt{5}$.

Newton’s Method in One Variable

These examples demonstrate some general properties of the Newton method in one variable.

Typically, there are neighborhoods of the root inside of which the Newton method will converge to the root, and the convergence will be *quadratic*.

Practically, that means that the number of correct digits *doubles* at each iteration step.

On the other hand, outside of these convergence areas, the Newton method might actually diverge. Right at the boundary of the convergence area, the Newton will get stuck in a loop.

Newton's Method in Several Variables

Newton's Method in Several Variables

Suppose that $F : \mathbb{R}^n \to \mathbb{R}^n$ is a differentiable function with some unknown root $x_* \in \mathbb{R}^n$.

So $F(x_*) = 0$.

We will describe a method that starts with an initial approximate root $x_{(0)}$ and constructs a sequence $x_{(1)}, x_{(2)}, x_{(3)}, \ldots$ of approximate roots.

We hope that these approximate roots will converge to the actual root x_*.

We will use the multivariate Newton method, which is a natural extension of the Newton method in one real variable.

Newton’s Method in Several Variables

We need to linearize the function F around some point $a \in \mathbb{R}^n$.

We use the first-order Taylor polynomial:

$$F(x) \approx F(a) + \sum_{i=1}^{n} \partial_i F(a) \cdot (x - a)_i.$$

The partial derivatives $\partial_i F(a)$ are the columns of the total derivative $F'(a)$ at the point a. Hence we can rewrite the first-order Taylor polynomial as

$$F(x) \approx F(a) + F'(a)(x - a).$$

Here, $F'(a) \in \mathbb{R}^{n \times n}$. This notation is more compact and leads immediately to the multivariate Newton method.

Newton's Method in Several Variables

Suppose that we already have an approximate root $x_{(0)} \in \mathbb{R}$ that is reasonably close to x_*.

We use the first-order Taylor polynomial around $x_{(0)}$ to approximate F:

$$F(x) \approx F(x_{(0)}) + F'(x_{(0)})(x - x_{(0)}).$$

If we plug in x_*, then $F(x_*) = 0$ gives the approximate relation

$$0 \approx F(x_{(0)}) + F'(x_{(0)})(x_* - x_{(0)}).$$

Assuming that $F'(x_{(0)})$ is invertible, we rearrange this "equation" and thus find

$$x_* \approx x_{(0)} - F'(x_{(0)})^{-1} F(x_{(0)}).$$

Newton's Method in Several Variables

Consequently, we use

$$x_{(1)} := x_{(0)} - F'(x_{(0)})^{-1}F(x_{(0)}).$$

as the next approximate root. Repeating this procedure leads to a sequence

$$\begin{aligned}
x_{(1)} &:= x_{(0)} - F'(x_{(0)})^{-1}F(x_{(0)}),\\
x_{(2)} &:= x_{(1)} - F'(x_{(1)})^{-1}F(x_{(1)}),\\
x_{(3)} &:= x_{(2)} - F'(x_{(2)})^{-1}F(x_{(2)}),\\
&\vdots
\end{aligned}$$

of approximate roots $x_{(0)}, x_{(1)}, x_{(2)}, \ldots$ that obey the recursive formula

$$x_{(k+1)} := x_{(k)} - F'(x_{(k)})^{-1}F(x_{(k)}).$$

Newton's Method in Several Variables

A sequence $x_{(0)}, x_{(1)}, x_{(2)}, \dots$ is constructed according to the recursive rule

$$x_{(k+1)} := x_{(k)} - F'(x_{(k)})^{-1} F(x_{(k)}).$$

Note that this requires F to be differentiable.

Furthermore, we need that the Jacobian $F'(x_{(k)})$ is an invertible matrix because otherwise we cannot continue the construction.

Even if the sequence is always defined, it is not guaranteed that the method converges to the root at all.

Newton’s Method in One Variable

Newton’s Method in Several Variables

Convergence Theory for the Newton Method

Convergence Theory for the Newton Method

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function and that the approximations $x_{(0)}, x_{(1)}, x_{(2)}, \ldots$ have been defined:

$$x_{(k+1)} = x_{(k)} - f'(x_{(k)})^{-1} f(x_{(k)})$$

and $f'(x_{(k)}) \neq 0$ in particular.

Suppose that x_* solves $f(x) = 0$ and define the error terms

$$e_{(k)} = x_{(k)} - x_*.$$

We want to estimate $|e_{(k)}|$. This can only be done with additional assumptions on f and the starting value $x_{(0)}$. One example theorem is the following.

Convergence Theory for the Newton Method

Theorem

Let $f : \mathbb{R} \to \mathbb{R}$ be a twice differentiable function with a root x^*.

Suppose that f'' is continuous and that $f'(x_*) \neq 0$.

Then there exists $\delta > 0$ such that for all $x_{(0)} \in \mathbb{R}$ with $|x_{(0)} - x_*| < \delta$ we have that Newton's method converges. Moreover, there exists $C > 0$ such that

$$|x_{(k+1)} - x_*| \leq C|x_{(k)} - x_*|^2.$$

Convergence Theory for the Newton Method

With additional conditions on f we can guarantee the convergence of Newton's method even for ***arbitrary*** starting points.

Theorem

Let $f : \mathbb{R} \to \mathbb{R}$ be a twice differentiable function with a root x^*.

Suppose that f is convex and that f'' is continuous.

Then Newton's method will converge from any starting value.

Convergence Theory for the Newton Method

Example

Quadratic convergence is not guaranteed if $f'(x_*) = 0$. For example, if $f(x) = x^2$, so that $f'(x) = 2x$, then Newton's method gives

$$x_{(k+1)} = x_{(k)} - f(x_{(k)})/f'(x_{(k)}) = x_{(k)} - \frac{1}{2}x_{(k)} = \frac{1}{2}x_{(k)}.$$

So for any starting value $x_{(0)}$ we have

$$x_{(k)} = 2^{-k}x_{(0)}.$$

The convergence is not quadratic but merely linear:

$$\frac{x_* - x_{(k+1)}}{x_* - x_{(k)}} = \frac{1}{2}$$