Newton's Method for Nonlinear Equations

CS/SE 4X03

Ned Nedialkov

McMaster University

November 11, 2021

Outline

Scalar case

- Given a scalar function f find a zero/root of f, i.e. an r such that $f(r) = 0$
- f may have no zeros, one, or many
- Let r be a root of f and let $x_n \approx r$
  From

$$0 = f(r) = f(x_n) + f'(x_n)(r - x_n) + O(|r - x_n|^2)$$
$$0 = f(r) \approx f(x_n) + f'(x_n)(r - x_n)$$

we find x_{n+1} by solving

$$f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0 \quad (1)$$

Scalar case cont.

- That is

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{2}$$

- We start with an initial guess x_0 and compute $x_1, x_2, \ldots$
- How to choose x_0, does it converge to a root, when to stop iterating...?

Interpretation

Given x_0, we compute

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

The tangent line at $(x, f(x_0))$ is

$$l(x) = f(x_0) + f'(x_0)(x - x_0)$$

We find x_1 such that $l(x)$ crosses the x axis, $l(x_1) = 0$:

$$0 = l(x_1) = f(x_0) + f'(x_0)(x_1 - x_0)$$

Similarly for x_2, x_3, ...

Examples

Square root

- Given $a > 0$, compute $\sqrt{a}$
- Write $x = \sqrt{a}$, $f(x) = x^2 - a$
- Apply (2):

$$
\begin{aligned}
x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - a}{2x_n} \\
&= x_n - \frac{x_n}{2} + \frac{a}{2x_n} \\
&= 0.5\left(x_n + \frac{a}{x_n}\right)
\end{aligned}
$$

- Let $a = 2$ and $x_0 = 3$
- We compute

i	x_i	$\lvert x_i - \sqrt{2}\rvert$
1	1.8333333333333333	4.19e-01
2	1.4621212121212122	4.79e-02
3	1.4149984298948031	7.85e-04
4	1.4142137800471977	2.18e-07
5	1.4142135623731118	1.67e-14
6	1.4142135623730949	2.22e-16

Examples cont.

Dividing without division operation

- How to obtain a/b without division?
- $a/b = a * (1/b)$
- Find $1/b$. Write $f(x) = 1/x - b$ and apply (2)

$$\begin{aligned} x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{1/x_n - b}{-1/x_n^2} \\ &= x_n + x_n - bx_n^2 \\ &= x_n(2 - bx_n) \end{aligned}$$

Examples cont.

- With $b = 3$ and $x_0 = 0.3$, we compute

i	x_i	$\lvert x_i - 1/3 \rvert$
1	0.330000000000000	3.33e-03
2	0.333300000000000	3.33e-05
3	0.333333330000000	3.33e-09
4	0.333333333333333	5.55e-17

Convergence

Theorem 1. If f, f', and f'' are continuous in a neighbourhood of a root r of f and $f'(r) \neq 0$, then $\exists \delta > 0$ such that if $|r - x_0| \leq \delta$, then all x_n satisfy

$$|r - x_n| \leq \delta, \tag{3}$$

$$|r - x_{n+1}| \leq c(\delta)|r - x_n|^2, \tag{4}$$

where $c(\delta)$ is defined in (6), and x_n converges to r

Let $e_n = r - x_n$. (4) is

$$|e_{n+1}| \leq c(\delta)|e_n|^2 \tag{5}$$

If e.g. $|e_n| \approx 10^{-4}$, $|e_{n+1}| \lessapprox c(\delta)10^{-8}$

If sufficiently close to r, each iteration $\approx$ doubles the number of accurate digits

Quadratic convergence $|e_{n+1}| \leq \text{constant} \cdot |e_n|^2$

Order of convergence is 2

Convergence cont.

Proof. From Taylor series,

$$
\begin{aligned}
0 = f(r) &= f(x_n) + f'(x_n)(r - x_n) + \frac{f''(\xi)}{2}(r - x_n)^2 \\
&= f(x_n) + f'(x_n)e_n + \frac{f''(\xi)}{2}e_n^2
\end{aligned}
$$

$$
f(x_n) + f'(x_n)e_n = -\frac{f''(\xi)}{2}e_n^2, \quad \xi \text{ is between } r \text{ and } x_n
$$

The error in x_{n+1} is

$$
\begin{aligned}
e_{n+1} = r - x_{n+1} &= r - \left(x_n - \frac{f(x_n)}{f'(x_n)} \right) = r - x_n + \frac{f(x_n)}{f'(x_n)} \\
&= e_n + \frac{f(x_n)}{f'(x_n)} = \frac{f(x_n) + e_n f'(x_n)}{f'(x_n)} \\
&= -\frac{1}{2}\frac{f''(\xi)}{f'(x_n)}e_n^2
\end{aligned}
$$

Convergence cont.

For a $\delta > 0$, let

$$c(\delta) = \frac{1}{2} \frac{\max_{|r-x|\leq\delta} |f''(x)|}{\min_{|r-x|\leq\delta} |f'(x)|} \tag{6}$$

Then (4) follows from

$$\begin{aligned}
|e_{n+1}| &= \frac{1}{2} \frac{|f''(\xi)|}{|f'(x_n)|} e_n^2 \leq \frac{1}{2} \frac{\max_{|r-x|\leq\delta} |f''(x)|}{\min_{|r-x|\leq\delta} |f'(x)|} e_n^2 \\
&\leq c(\delta) e_n^2
\end{aligned}$$

Choose δ such that $c(\delta)\delta < 1$. This is possible since

$$c(\delta) \to \frac{1}{2} \left| \frac{f''(r)}{f'(r)} \right| \quad \text{as } \delta \to 0$$

and $f'(r) \neq 0$ by assumption

Convergence cont.

If $|e_n| = |r - x_n| \le \delta$, then

$$\begin{aligned} |e_{n+1}| &\le c(\delta)e_n^2 = c(\delta)\cdot e_n \cdot e_n \le c(\delta)\delta \cdot e_n \\ &< \rho e_n, \quad \text{where } \rho = \delta c(\delta) < 1 \end{aligned}$$

and (3) follows

Hence

$$|e_n| \le \rho|e_{n-1}| \le \rho^2|e_{n-2}| \le \cdots \le \rho^n|e_0|$$

Since $\rho < 1$, $|e_n| \to r$ as $n \to \infty$

Subtleties

We require $f'(r) \neq 0$

If $f'(r) = 0$ and $f''(r) \neq 0$, r is a double root, e.g. $f(x) = (x-1)^2$

A root r is of multiplicity m if $f^{(k)}(r) = 0$ for all $k = 1, 2, \dots m-1$ and $f^{(m)}(r) \neq 0$. In this case

$$x_{n+1} = x_n - m\frac{f(x_n)}{f'(x_n)}$$

is quadratically convergent

If $f'(x_n)$ is not available, we can approximate $f'(x_n) \approx \frac{f(x_n)-f(x_{n-1})}{x_n - x_{n-1}}$
Then

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}$$

This is the secant method. Order of convergence is $(1+\sqrt{5})/2 \approx 1.618$ (golden ratio)

Newton for systems of equations

- Consider a system of n equations in n variables

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0 \\ f_2(x_1, x_2, \ldots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, x_2, \ldots, x_n) &= 0 \end{aligned}$$

- Denote $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and $F = (f_1, f_2, \ldots, f_n)$
- Find $\mathbf{x}^*$ (if it exists) such that $F(\mathbf{x}^*) = 0$

Newton for systems of equations cont.

- Assume $\mathbf{x}^*$ is such that $F(\mathbf{x}^*) = 0$ and $\mathbf{x}^{(k)} \approx \mathbf{x}^*$
- From

$$0 = F(\mathbf{x}^*) \approx F(\mathbf{x}^{(k)}) + F'(\mathbf{x}^{(k)})(\mathbf{x}^* - \mathbf{x}^{(k)})$$

find $\mathbf{x}^{(k+1)}$ by solving (cf. (1))

$$F(\mathbf{x}^{(k)}) + F'(\mathbf{x}^{(k)})(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = 0 \tag{7}$$

- $F'(\mathbf{x}^{(k)})$ is the Jacobian of F at $\mathbf{x}^{(k)}$, an $n \times n$ matrix

Newton for systems of equations cont.

- Let $s = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$
- Solve (assuming $F'(\mathbf{x}^{(k)})$ nonsingular) linear system

$$F'(\mathbf{x}^{(k)})s = -F(\mathbf{x}^{(k)}) \tag{8}$$

and set

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + s \tag{9}$$

- (8,9) is basic Newton for systems of equations

Example

- Consider

$$0 = F(\mathbf{x}) = \begin{cases} x_1^2 + x_2^2 - 25 \\ x_1^2 - x_2 - 1 \end{cases}$$

- Jacobian is

$$F'(\mathbf{x}) = \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{pmatrix}$$

- Let $x_0 = (5, 1)^T$

- Then

$$F(\mathbf{x}^{(0)}) = (1, 23)^T$$
$$J(\mathbf{x}^{(0)}) = \begin{pmatrix} 10 & 2 \\ 10 & -1 \end{pmatrix}$$

- Solve $J(\mathbf{x}^{(0)})s = -F(\mathbf{x}^{(0)})$ **for s**
- $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + s$ and so on
- We compute

i	x_1	x_2	$\|F(\mathbf{x})\|$
1	3.433333333333334	8.333333333333332	5.63e+01
2	2.632585333089088	5.289308176100628	9.93e+00
3	2.358810087435537	4.489032143454986	7.19e-01
4	2.329316858408983	4.424847176309882	5.06e-03
5	2.329040359270796	4.424428918660463	2.63e-07
6	2.329040339044829	4.424428900898053	7.11e-15