Convergence Analysis

let the function f have continuous first and second derivatives f' and f'', and r be a simple root of f with $f'(r) \neq 0$. If x_0 is *sufficiently* close to r, then Newton's method converges to r quadratically.

$$|x_{n+1} - r| \leq c\,|x_n - r|^2$$

if x_n differs from r by at most one unit in the kth decimal place, i.e.,

$$|x_n - r| \leq 10^{-k}$$

then, for $c = 1$, we have

$$|x_{n+1} - r| \leq 10^{-2k}$$

the number of correct decimal digits doubled after another iteration

Convergence Proof

let $e_n = r - x_n$, Newton's method gives a sequence $\{x_n\}$ such that

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

using Taylor's expansion, there exists a point ξ_n between x_n and r for which

$$\begin{aligned} 0 &= f(r) = f(x_n + e_n) \\ &= f(x_n) + e_n\, f'(x_n) + \frac{1}{2}\, e_n^2\, f''(\xi_n) \end{aligned}$$

it follows that

$$e_n f'(x_n) + f(x_n) = -\frac{1}{2}\, e_n^2\, f''(\xi_n)$$

Convergence Proof Cont.

we thus have

$$e_{n+1} = -\frac{1}{2}\left(\frac{f''(\xi_n)}{f'(x_n)}\right) e_n^2$$

define a upper bound

$$c(\delta) = \frac{1}{2}\frac{\max_{|x-r|\le\delta}|f''(x)|}{\max_{|x-r|\le\delta}|f'(x)|}, \qquad (\delta > 0)$$

we can choose δ small so that

$$|e_n| = |x_n - r| \le \delta \quad \text{and} \quad |\xi_n - r| \le \delta$$

this is to guarantee that x_n is close to r within a distance of δ

Convergence Proof Cont.

for very small $\delta > 0$, we have

$$|e_{n+1}| = \frac{1}{2}\left|\frac{f''(\xi_n)}{f'(x_n)}\right| e_n^2 \leq c(\delta)\, e_n^2$$

$$\leq \delta\, c(\delta)\, |e_n| = \rho\, |e_n|$$

with $\rho = \delta\, c(\delta) < 1$ if δ is small enough, therefore

$$|x_{n+1} - r| = |e_{n+1}| \leq \rho\, |e_n| \leq |e_n| \leq \delta$$

x_{n+1} is also close to r within a distance of δ. By recursion, if x_0 is close to r, then

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

Since $\rho < 1$, this is to say

$$\lim_{n\to\infty} |e_n| = 0 \quad \text{as} \quad n \to \infty$$

Weakness of Newton's Method

Newton's method converges fast, only when x_0 is chosen close to r. In practice, there might be a number of problems also

1.) needs derivative value and availability

2.) starting point must be close to r

3.) lose quadratic convergence if multiple root

4.) iterates may runaway (not in convergence domain)

5.) flat spot with $f'(x_n) = 0$

6.) cycling iterates around r

Systems of Nonlinear Equations

Newton's method is really useful for finding zero of a system of nonlinear equations

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

written in vector form as

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}$$

where

$$\begin{aligned} \mathbf{f} &= (f_1, f_2, \ldots, f_n)^T \\ \mathbf{x} &= (x_1, x_2, \ldots, x_n)^T \end{aligned}$$

we have

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - [\mathbf{f}'(\mathbf{x}^{(k)})]^{-1}\mathbf{f}(\mathbf{x}^{(k)})$$

$\mathbf{f}'(\mathbf{x}^{(k)}) = \mathbf{J}(\mathbf{x}^{(k)})$ is the Jacobian matrix

A 3 Equation Example

$$
\begin{aligned}
f_1(x_1, x_2, x_3) &= 0 \\
f_2(x_1, x_2, x_3) &= 0 \\
f_3(x_1, x_2, x_3) &= 0
\end{aligned}
$$

using Taylor expansion

$$
\begin{aligned}
& f_i(x_1 + h_1, x_2 + h_2, x_3 + h_3) \\
& = f_i(x_1, x_2, x_3) + \\
& \qquad h_1 \frac{\partial f_i}{\partial x_1} + h_2 \frac{\partial f_i}{\partial x_2} + h_3 \frac{\partial f_i}{\partial x_3} + \cdots
\end{aligned}
$$

let $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)})^T$ be an approximate solution and the computed correction be $\mathbf{h} = (h_1, h_2, h_3)^T$. Hence

$$0 \approx \mathbf{f}(\mathbf{x}^{(0)} + \mathbf{h}) = \mathbf{f}(\mathbf{x}^{(0)}) + \mathbf{f}'(\mathbf{x}^{(0)})\mathbf{h}$$

Example Cont.

the Jacobian matrix is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{bmatrix}$$

it follows that

$$\mathbf{h} \approx -[\mathbf{f}'(\mathbf{x}^{(0)})]^{-1}\mathbf{f}(\mathbf{x}^{(0)})$$

hence, the new iterate is

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - [\mathbf{f}'(\mathbf{x}^{(0)})]^{-1}\mathbf{f}(\mathbf{x}^{(0)})$$

in practice, we solve the Jacobian matrix in

$$[\mathbf{J}(\mathbf{x}^{(k)})]\mathbf{h}^{(k)} = -\mathbf{f}(\mathbf{x}^{(k)})$$

so that

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{h}^{(k)}$$

Secant Method

in Newton's method

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

we need to evaluate $f(x_n)$ and $f'(x_n)$ at each iteration

we can approximate the derivative at $x = x_n$ by

$$f'(x_n) \approx \frac{f(x_{n-1}) - f(x_n)}{x_{n-1} - x_n}$$

thus, the secant method generates iterates

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

only one functional evaluation at each iteration

Comments

secant method needs to iterates to start with, can use bisection method to generate the second iterate

secant method does not need to know the derivative of $f(x)$

if $|f(x_n) - f(x_{n-1})|$ is small, the computation may lose significant digits and becomes unstable

the convergence rate of secant method is superlinear

$$|e_{n+1}| \leq C\, |e_n|^{\alpha}$$

with $\alpha = \frac{1}{2}(1 + \sqrt{5}) \approx 1.62$. Its convergence rate is between that of bisection method and the Newton's method