Classical Numerical Analysis, Chapter 15

Abner J. Salgado and Steven M. Wise

asalgad1@utk.edu swise1@.utk.edu
University of Tennessee

Chapter 15
Solution of Nonlinear Equations

What are we doing here?

In this chapter, we depart from *linear algebra* problems, and concentrate in the study of *nonlinear* problems. We will focus on methods for solving a nonlinear system of equations. In other words, given $m, n \in \mathbb{N}$, and $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ we wish to find a a point $\boldsymbol{\xi} \in \mathbb{R}^n$ such that

$$\mathbf{f}(\boldsymbol{\xi}) = \mathbf{0}. \tag{1}$$

Such a point $\boldsymbol{\xi}$, if it exists, is called a *root of* $\mathbf{f}$. Of course, if $m = n$, and $\mathbf{f}$ so happens to be affine then this problem reduces to a linear one, and can be treated either by the direct methods or, iterative ones. If $m \neq n$, but the function $\mathbf{f}$ is still affine, then the least squares methods apply.

The function $\mathbf{f}$ in this chapter, however, is not assumed to be affine. The importance of (1) cannot be overstated; many problems can be reduced to this.

We must immediately remark that any method that attempts to find a solution to (1) must be iterative, unless a very special structure is assumed on the function $\mathbf{f}$.

Strategy

The general strategy that we will follow can be simply stated as follows:

1. Show that the problem has at least one solution.
2. Isolate a root, that is, find an open region $D \subset \mathbb{R}^n$ for which there is $\boldsymbol{\xi} \in D$ that solves (1) and $\mathbf{f}(\mathbf{x}) \neq \mathbf{0}$ for all $\mathbf{x} \in D \setminus \{\boldsymbol{\xi}\}$.
3. Iterate.

Unfortunately, there is no general strategy to treat the first two points, as these usually require analytical methods or additional **knowledge** about the problem at hand.

Starting from some $\mathbf{x}_0 \in \mathbb{R}^n$, we will construct sequences $\{\mathbf{x}_k\}_{k=1}^{\infty} \subset \mathbb{R}^n$ which, hopefully, converge $\mathbf{x}_k \to \boldsymbol{\xi}$ as fast as possible. We will, as usual, stop the iteration when a prescribed tolerance $\varepsilon > 0$ is reached, i.e.,

$$\|\mathbf{x}_k - \boldsymbol{\xi}\| < \varepsilon.$$

This is not practical. So, how do we **really** know when to stop the iterations? One might be tempted to say that, since $\mathbf{f}(\boldsymbol{\xi}) = \mathbf{0}$ we can stop them whenever

$$\|\mathbf{f}(\mathbf{x}_k)\| < C\varepsilon,$$

for some suitable constant C.

Sensitivity

Assume that $m = n = 1$, and the function f is $k+1$ times continuously differentiable in a neighborhood of its unique root $\xi \in \mathbb{R}$. Moreover, we will assume that $f^{(p)}(\xi) = 0$ for $p < k$ but $f^{(k)}(\xi) \neq 0$. Then, given a perturbation δx, we let η be such that

$$f(\xi + \delta x) = \eta.$$

Taylor's Theorem then shows that

$$\begin{aligned}
\eta &= f(\xi + \delta x) \\
&= f(\xi) + f'(\xi)\delta x + \frac{1}{2} f''(\xi)\delta x^2 + \cdots + \frac{1}{k!} f^{(k)}(\xi)\delta x^k + \mathcal{O}(|\delta x|^{k+1}) \\
&= \frac{1}{k!} f^{(k)}(\xi)\delta x^k + \mathcal{O}(|\delta x|^{k+1}).
\end{aligned}$$

In other words, at least intuitively, the allowed relative size of the perturbation δx, to obtain an output of size η is

$$\left| \frac{\delta x}{\eta} \right| \approx \left| \eta^{1-k} \frac{k!}{f^{(k)}(\xi)} \right|^{1/k}.$$

Sensitivity and Non-Simple Roots

The allowed relative size of the perturbation δx, to obtain an output of size η is

$$\left|\frac{\delta x}{\eta}\right| \approx \left|\eta^{1-k}\,\frac{k!}{f^{(k)}(\xi)}\right|^{1/k}.$$

From this we learn two things: the smaller the value of the first nonzero derivative, the larger δx can be; and, the higher the order of the first nonzero derivative, the larger δx can be. For this reason, of importance to us will be so–called *simple roots*.

Definition (simple root)

Let $f : \mathbb{R} \to \mathbb{R}$ have a root $\xi \in \mathbb{R}$, and assume that f is differentiable at ξ. We say that ξ is a **simple root** if $f'(\xi) \neq 0$. If this is not the case, we say that the root is **non–simple**.

Fixed Points and Contraction Mappings

Fixed Point Iterations

Definition (fixed point iteration)

Suppose $-\infty < a < b < \infty$, $g \in C([a,b])$, *and* $g(x) \in [a,b]$, *for all* $x \in [a,b]$. *In other words,* $g([a,b]) \subseteq [a,b]$. *Given* $x_0 \in [a,b]$, *the algorithm for constructing the recursive sequence* $\{x_k\}_{k=0}^{\infty}$ *via*

$$x_{k+1} = g(x_k), \quad k \geq 0, \tag{2}$$

is called a **simple iteration scheme** *or, sometimes, a* **fixed point iteration scheme**.

Convergence to a Fixed Point

The following fact follows easily from continuity.

Proposition (fixed point)

Suppose $-\infty < a < b < \infty$, $g \in C([a,b])$, and $g([a,b]) \subseteq [a,b]$. Assume that the sequence $\{x_k\}_{k=0}^{\infty}$ obtained by a simple iteration scheme converges to a limit $\xi \in [a,b]$. Then ξ is a fixed point of g, that is, $g(\xi) = \xi$.

Proof.

Indeed, by continuity

$$g(\xi) = g\left(\lim_{k\to\infty} x_k\right) = \lim_{k\to\infty} g(x_k) = \lim_{k\to\infty} x_{k+1} = \xi. \qquad \square$$

Existence of a Fixed Point

Theorem (existence)

Suppose $-\infty < a < b < \infty$, $g \in C([a,b])$, *and* $g([a,b]) \subseteq [a,b]$. *Then, there exists at least one fixed point* $\xi \in [a,b]$ *of* g.

Proof.

Define

$$f(x) = x - g(x), \quad \forall\, x \in [a,b].$$

Then, since $g([a,b]) \subseteq [a,b]$,

$$f(b) = b - g(b) \geq 0 \quad \text{and} \quad f(a) = a - g(a) \leq 0.$$

By the Intermediate Value Theorem, since $f(a) \leq 0 \leq f(b)$, there is a point $\xi \in [a,b]$, such that $f(\xi) = 0$. For this point $\xi = g(\xi)$. □

Contractions

Definition (contraction)

Suppose $-\infty < a < b < \infty$ *and* $g \in C([a,b])$. *We say that* g *is* **Lipschitz continuous** *on* $[a,b]$ *iff there exists a constant* $L > 0$ *such that*

$$|g(x) - g(y)| \leq L|x - y|, \quad \forall\, x, y \in [a,b],$$

and the associated L *is called the* **Lipschitz constant**. *The function* g *is called a* **contraction** *on* $[a,b]$ *iff it is Lipschitz on* $[a,b]$ *and its associated Lipschitz constant,* L, *satisfies* $L \in (0,1)$.

Proposition (translation)

Let $-\infty < a < b < \infty$ and $g \in C([a,b])$ be a contraction on $[a,b]$. There is a constant $m \in \mathbb{R}$ such that the function $\tilde{g} : [a,b] \to \mathbb{R}$, where

$$\tilde{g}(x) = g(x) + m, \quad \forall\, x \in [a,b],$$

is a contraction on $[a,b]$ and, moreover, $\tilde{g}([a,b]) \subseteq [a,b]$.

Proof.

A homework exercise. □

Uniqueness of and Convergence to the Fixed Point

Theorem (uniqueness)

Suppose $-\infty < a < b < \infty$, $g \in C([a,b])$, *and* $g([a,b]) \subseteq [a,b]$. *If* g *is a contraction on* $[a,b]$, *then* g *has a unique fixed point* $\xi \in [a,b]$. *Furthermore, the sequence* $\{x_k\}_{k=0}^{\infty}$ *generated by* (2) *converges to* ξ *for any starting value* $x_0 \in [a,b]$.

Proof.

The previous theorem guarantees the existence of at least one fixed point $\xi \in [a,b]$. Suppose $\eta \in [a,b]$ is another fixed point. Since g is a contraction

$$|\xi - \eta| = |g(\xi) - g(\eta)| \leq L|\xi - \eta|.$$

Therefore

$$0 \leq (1-L)|\xi - \eta| \leq 0,$$

which proves that $\xi = \eta$. Hence the fixed point $\xi \in [a,b]$ is unique.

Proof Cont.

Now, suppose that $\{x_k\}_{k=0}^{\infty}$ is generated by (2) for $x_0 \in [a, b]$. Then

$$|\xi - x_k| = |g(\xi) - g(x_{k-1})| \leq L|\xi - x_{k-1}|.$$

By induction, for any $k \in \mathbb{N}$,

$$|\xi - x_k| \leq L^k|\xi - x_0|.$$

By the Squeeze Theorem, since $L \in (0, 1)$, we must have $x_k \to \xi$. □

Local At Least Linear Convergence

Theorem (local (at least linear) convergence)

Suppose $-\infty < a < b < \infty$, $g \in C([a,b])$, *and* $g([a,b]) \subseteq [a,b]$. *Let* $\xi \in [a,b]$ *be a fixed point of* g, *that is,* $\xi = g(\xi)$. *Suppose that there is a constant* $\delta > 0$, *such that* $I_\delta = (\xi - \delta, \xi + \delta) \subset [a,b]$, $g \in C^1(I_\delta)$, *and* $|g'(\xi)| < 1$. *Then the sequence* $\{x_k\}_{k=0}^{\infty}$ *generated by* (2) *converges at least linearly to* ξ, *provided* x_0 *is sufficiently close to* ξ.

Proof.

Suppose that $\xi \in (a,b)$, that is, ξ is in the interior of the set $[a,b]$. The reader can examine the other cases. Since $|g'(\xi)| < 1$, there is an $h \in (0,\delta)$ and an $L \in (0,1)$ such that, for all $x \in I_h = (\xi - h, \xi + h)$,

$$|g'(x)| \leq L < 1.$$

The proof of this last fact is left to the reader as an exercise.

Proof Cont.

Suppose that $x_k \in I_h$. Then, using the Mean Value Theorem,

$$|\xi - x_{k+1}| = |g(\xi) - g(x_k)| = |g'(\eta_k)| \cdot |\xi - x_k| \leq L|\xi - x_k|,$$

for some $\eta_k \in I_h$ between ξ and x_k. This proves that, if $x_k \in I_h$, then $x_{k+1} \in I_h$. Using induction, if $x_0 \in I_h$,

$$|\xi - x_k| \leq L^k |\xi - x_0|.$$

By the squeeze theorem, since $L^k \to 0$, as $k \to \infty$, we have $x_k \to \xi$, as $k \to \infty$. Furthermore, the convergence is at least linear. □

Linear and At Least Linear Convergence

What do we mean by a rate of convergence?

Definition

Let $d \in \mathbb{N}$, and suppose that the sequence $\{\mathbf{x}_k\}_{k=1}^{\infty} \subset \mathbb{C}^d$ converges to the point $\boldsymbol{\xi} \in \mathbb{C}^d$, i.e., $\boldsymbol{\xi} = \lim_{k\to\infty} \mathbf{x}_k$. We say that $\mathbf{x}_k$ **converges to $\boldsymbol{\xi}$ at least linearly** *iff there exists a sequence of positive real numbers $\{\varepsilon_k\}_{k=1}^{\infty}$ that converges to 0 and a real number $\mu \in (0,1)$ such that*

$$\|\mathbf{x}_k - \boldsymbol{\xi}\|_2 \leq \varepsilon_k, \ \forall k \in \mathbb{N}, \qquad \lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \mu. \tag{3}$$

If (3) holds with $\|\mathbf{x}_k - \boldsymbol{\xi}\|_2 = \varepsilon_k$, for $k = 1, 2, \ldots$, we say that $\mathbf{x}_k$ **converges to $\boldsymbol{\xi}$ linearly**, *or* **exactly linearly**.

Order $q > 1$ Convergence, Quadratic and Cubic Convergence

Definition

Let $d \in \mathbb{N}$, and suppose that the sequence $\{\mathbf{x}_k\}_{k=1}^{\infty} \subset \mathbb{C}^d$ converges to the point $\boldsymbol{\xi} \in \mathbb{C}^d$, i.e., $\boldsymbol{\xi} = \lim_{k\to\infty} \mathbf{x}_k$. We say that $\mathbf{x}_k$ **converges to $\boldsymbol{\xi}$ with at least order** *q, $q > 1$, iff there exists a sequence of positive real numbers $\{\varepsilon_k\}_{k=1}^{\infty}$ that converges to 0 and a real number $\mu > 0$ such that*

$$\|\mathbf{x}_k - \boldsymbol{\xi}\|_2 \leq \varepsilon_k, \ \forall k \in \mathbb{N}, \qquad \lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k^q} = \mu. \tag{4}$$

If (4) holds with $\|\mathbf{x}_k - \boldsymbol{\xi}\|_2 = \varepsilon_k$, for $k = 1, 2, \ldots$, we say that $\mathbf{x}_k$ **converges to $\boldsymbol{\xi}$ with order** *q, or with exactly order q. In particular, if $q = 2$, we say that $\mathbf{x}_k$ converges to $\boldsymbol{\xi}$* **quadratically**. *If $q = 3$, we say that $\mathbf{x}_k$ converges to $\boldsymbol{\xi}$* **cubically**.

Relaxation

Definition (relaxation)

Let $I \subseteq \mathbb{R}$ be an interval, $f \in C(I)$, and $x_0 \in I$ be given. The **relaxation method** *is an algorithm for computing the terms of the sequence $\{x_k\}_{k=0}^{\infty}$ via the recursive formula*

$$x_{k+1} = x_k - \lambda f(x_k), \tag{5}$$

where $\lambda \neq 0$. The method is **well–defined** *iff $x_k \in I$ for all $k = 1, 2, 3, \ldots$, and the relaxation method* **converges** *iff there is a $\xi \in I$, with $f(\xi) = 0$, such that $x_k \to \xi$, as $k \to \infty$.*

Convergence of the Relaxation Method

Theorem (convergence)

Let $I \subset \mathbb{R}$ be an interval. Suppose that $f : I \to \mathbb{R}$ and, for some $\xi \in I$, $f(\xi) = 0$, but $f'(\xi) \neq 0$. Assume that, for some $\delta > 0$, $f \in C^1(I_\delta)$, where $I_\delta = [\xi - \delta, \xi + \delta] \subseteq I$. Then, there exists positive real numbers λ and $h \in (0, \delta)$ such that the sequence $\{x_k\}_{k=0}^{\infty}$ defined by the relaxation scheme (5) converges to ξ for any $x_0 \in I_h = [\xi - h, \xi + h]$.

Proof.

Suppose that $f'(\xi) = \alpha > 0$. The case $\alpha < 0$ is analogous, and left to the reader. By continuity, we may assume that

$$0 < \frac{1}{2}\alpha \leq f'(x) \leq \frac{3}{2}\alpha, \quad \forall\, x \in I_\delta.$$

If this is not the case, we can just choose a smaller δ and redefine I_δ. Set

$$M = \max_{x \in I_\delta} f'(x).$$

Proof Cont.

Thus $\frac{1}{2}\alpha \leq M \leq \frac{3}{2}\alpha$. For any $\lambda > 0$, it follows that

$$1 - \lambda M \leq 1 - \lambda f'(x) \leq 1 - \frac{1}{2}\lambda\alpha, \quad \forall\, x \in I_\delta.$$

We now choose, if possible, $\lambda > 0$ such that

$$1 - \lambda M = -\theta \quad \text{and} \quad 1 - \frac{1}{2}\lambda\alpha = \theta.$$

These equations are satisfied iff

$$\lambda M - 1 = 1 - \frac{1}{2}\lambda\alpha, \quad \theta = 1 - \frac{1}{2}\lambda\alpha$$

iff

$$\lambda = \frac{4}{2M + \alpha}, \quad \theta = \frac{2M - \alpha}{2M + \alpha}.$$

Now, define the iteration function g via

$$g(x) = x - \lambda f(x) = x - \frac{4f(x)}{2M + \alpha}.$$

Proof Cont.

With g defined via

$$g(x) = x - \lambda f(x) = x - \frac{4f(x)}{2M + \alpha},$$

the next step is to show the following: there is a constant $\delta > 0$, such that $I_\delta = (\xi - \delta, \xi + \delta) \subset [a, b]$, $g \in C^1(I_\delta)$, and $|g'(\xi)| < 1$. If this holds, then a previous theorem guarantees that $x_k \to \xi$, provided x_0 is sufficiently close to ξ. The convergence is at least linear. □

The Chord Method

Definition (chord method)

Let $I = [a, b]$ be an interval and $x_0 \in I$. The sequence $\{x_k\}_{k=0}^{\infty}$ is obtained by the **chord method** *iff is constructed via the relaxation method* (5) *with*

$$\lambda = \frac{1}{\frac{f(b)-f(a)}{b-a}}.$$

Theorem (convergence)

Let $I = [a, b] \subset \mathbb{R}$ be an interval, and $f \in C^1([a, b])$ is such that it has a unique simple root $\xi \in [a, b]$. If $b - a$ is sufficiently small, then the sequence $\{x_k\}_{k=1}^{\infty}$ obtained by the chord method converges linearly to ξ as $k \to \infty$.

Proof.

A homework exercise. □

Simplified Newton's Method

Definition (simplified Newton)

Let I be an interval and $x_0 \in I$ be such that $f'(x_0) \neq 0$. The sequence $\{x_k\}_{k=0}^{\infty}$ is obtained by **simplified Newton's method** *if is constructed via the relaxation method* (5) *with*

$$\lambda = \frac{1}{f'(x_0)}.$$

Proposition (convergence)

Let $I \subset \mathbb{R}$ be an interval, and let $f \in C(I)$ be such that there is $\xi \in I$ for which $f(\xi) = 0$. Define, for $\delta > 0$, $I_\delta = [\xi - \delta, \xi + \delta] \subseteq I$ and assume that there is $\delta > 0$ such that $f \in C^1(I_\delta)$. Finally, assume that $f'(\xi) = \alpha > 0$. Then, simplified Newton' method converges linearly to ξ.

Proof.

A homework exercise. We'll come back to this method later. □

Newton's Method in One Space Dimension

Newton's Method (The Big Gun)

Definition (Newton's method)

Let $I \subseteq \mathbb{R}$ be an interval, $f \in C^1(I)$, and $x_0 \in I$, with $f'(x_0) \neq 0$, be given. **Newton's method** *is an algorithm for computing the terms of the sequence $\{x_k\}_{k=0}^{\infty}$ via the recursive formula*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}. \tag{6}$$

We say that the method is **well–defined** *iff $x_k \in I$ and $f'(x_k) \neq 0$, for all $k = 1, 2, 3, \ldots$. We say that Newton's method* **converges** *iff there is a $\xi \in I$, with $f(\xi) = 0$, such that $x_k \to \xi$, as $k \to \infty$.*

At Least Linear Convergence

Proposition (linear convergence)

Let $I \subseteq \mathbb{R}$ be an interval and $f \in C^2(I)$ with $|f'(x)| \geq \alpha > 0$, for all $x \in I$. Assume that there is $\xi \in I$ for which $f(\xi) = 0$. There is a constant $h > 0$, such that, if $x_0 \in I$ and $|x_0 - \xi| < h$, then Newton's method converges at least linearly to ξ.

Proof.

For this proof, let us apply the theory of fixed iterations to the function

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

Notice that, since $f \in C^2(I)$, then

$$g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2} \quad \Longrightarrow \quad g'(\xi) = 0.$$

Proof Cont.

Thus, by continuity, there is a constant $\delta > 0$, such that, if $x \in I_\delta = [\xi - \delta, \xi + \delta]$, then

$$|g'(x)| < 1.$$

In addition notice that, by Taylor's Theorem we have, for any $x \in I$,

$$0 = f(\xi) = f(x) + f'(x)(\xi - x) + \frac{1}{2} f''(\eta)(\xi - x)^2,$$

for some η between x and ξ. Let us define

$$A = \frac{\max_{x \in I} |f''(x)|}{\alpha}, \quad h = \min\left\{\delta\,,\ \frac{1}{A}\right\}.$$

Then, for any $x \in I_h = [\xi - h, \xi + h]$, we have $|g'(x)| < 1$ and

$$|g(x) - \xi| = \left| x - \xi - \frac{f(x)}{f'(x)} \right| = \frac{1}{2} \left| \frac{f''(\eta)}{f'(x)} (x - \xi)^2 \right| \leq \frac{1}{2} A h^2 \leq \frac{1}{2} h,$$

so that $g(x) \in I_h$. We conclude the (at least linear) convergence of the fixed point iteration. □

Quadratic Convergence

Theorem (quadratic convergence)

Let $I \subset \mathbb{R}$ be an interval. Suppose that $f : I \to \mathbb{R}$ and, for some $\xi \in I$, $f(\xi) = 0$, but $f'(\xi) \neq 0$ and $f''(\xi) \neq 0$. Assume that, for some $\delta > 0$, $f \in C^2(I_\delta)$, where $I_\delta = [\xi - \delta, \xi + \delta] \subseteq I$, and $0 < \alpha \leq |f'(x)|$, for all $x \in I_\delta$. Set

$$A = \frac{\max_{x \in I_\delta} |f''(x)|}{\alpha} \ , \quad h = \min\left\{\delta \ , \ \frac{1}{A}\right\} . \tag{7}$$

If $|\xi - x_0| \leq h$, then the sequence $\{x_k\}_{k=0}^{\infty}$ defined by Newton's method (6) converges quadratically, as $k \to \infty$, to the root ξ.

Proof.

(Well-possedness): Suppose that $x_k \in I_\delta$. Then, by Taylor's Theorem,

$$0 = f(\xi) = f(x_k) + (\xi - x_k) f'(x_k) + \frac{(\xi - x_k)^2}{2} f''(\eta_k) \ ,$$

for some η_k between x_k and ξ.

Proof Cont.

Note that $f'(x_k) \neq 0$, and we have, using Equation (6) (the definition of Newton's method),

$$x_{k+1} - \xi = \frac{(\xi - x_k)^2 f''(\eta_k)}{2f'(x_k)}. \tag{8}$$

This is the fundamental error equation for Newton's method, and is the key to our analysis. Now, if $x_k \in I_h = [\xi - h, \xi + h]$,

$$|\xi - x_{k+1}| = \frac{1}{2}\frac{|f''(\eta_k)|}{|f'(x_k)|} \cdot |\xi - x_k| \cdot |\xi - x_k| \leq \frac{A}{2} \cdot h \cdot |\xi - x_k| \leq \frac{1}{2}|\xi - x_k|,$$

and $x_{k+1} \in I_h$ as well. The algorithm is, therefore, well-defined, since $x_{k+1} \in I_h$ and $f'(x_{k+1}) \neq 0$.

(Linear convergence): By induction, it is clear from the contraction estimate that

$$|\xi - x_k| \leq \frac{h}{2^k},$$

which proves that the sequence converges to the root ξ as $k \to \infty$, at least linearly.

Proof Cont.

(Quadratic convergence): From (8) we obtain

$$\frac{|\xi - x_{k+1}|}{|\xi - x_k|^2} = \frac{1}{2}\frac{|f''(\eta_k)|}{|f'(x_k)|}.$$

Since $x_k \to \xi$, by the Squeeze Theorem, $\eta_k \to \xi$ as $k \to \infty$ as well. Passing to limits, we have

$$\lim_{k\to\infty} \frac{|\xi - x_{k+1}|}{|\xi - x_k|^2} = \frac{1}{2}\frac{|f''(\xi)|}{|f'(\xi)|} = \sigma \in (0, \infty),$$

which establishes quadratic convergence. □

Global Convergence

Theorem (global convergence)

Let $[a, b] \subset \mathbb{R}$ be an interval and $f \in C^2([a, b])$ be such that, for some $\xi \in [a, b]$, $f(\xi) = 0$. Assume further that f' and f'' are strictly positive on the interval $[a, b]$. For any starting value $x_0 \in (\xi, b]$, the sequence $\{x_k\}_{k=0}^{\infty}$ defined by Newton's method (6) converges quadratically to the root ξ as $k \to \infty$. Moreover, $x_k > \xi$, for all $k \in \mathbb{N}$.

Proof.

Since f is monotonically increasing on $[a, b]$, ξ is the only root in $[a, b]$. Otherwise, by Rolle's Theorem, one could find a point where f' is zero, contradicting the assumptions. Also note that $f(x) > 0$ for all $x \in (\xi, b]$, and, likewise, $f(x) < 0$ for all $x \in [a, \xi)$.

Assume that $x_k > \xi$. Employing Newton's method (6), and using the positivity of $f(x_k)$ and $f'(x_k)$ we immediately obtain that

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

Proof Cont.

Furthermore, from the error equation,

$$x_{k+1} - \xi = \frac{(\xi - x_k)^2 f''(\eta_k)}{2f'(x_k)}, \tag{9}$$

using the positivity of $f''(\eta_k)$ and $f'(x_k)$, we find out that $\xi - x_{k+1} < 0$. In other words, $\xi < x_{k+1} < x_k$. Thus, $\{x_k\}_{k=0}^{\infty}$ is a bounded, monotonically decreasing sequence in $[a, b]$. By the Monotone Convergence Theorem, it must therefore have a limit point in $[a, b]$, call it η, as $k \to \infty$. But, this limit must be a fixed point of the function

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

Therefore, $f(\eta) = 0$. But since ξ is the unique root of f in $[a, b]$, it must be that $\xi = \eta$. This shows that $x_k \to \xi$ as $k \to \infty$. Quadratic convergence follows as in the proof of the last theorem. □

Example

Let us use Newton's method to compute the square root of a positive real number. Suppose that we want to compute $\sqrt{5}$. Define $f(x) = x^2 - 5$. There are two solutions to $f(x) = 0$, namely $\xi_{\pm} = \pm\sqrt{5}$. Let us pick $x_0 = 5$. The last theorem guarantees that this is a suitable choice if we want to compute the zero $\xi_+ = \sqrt{5}$. The sequence of approximations for Newton's method is defined by

$$x_{k+1} = x_k - \frac{x_k^2 - 5}{2x_k}, \quad k = 0, 1, \ldots. \tag{10}$$

Below, the correct digits are indicated using boldface.

k	x_k
0	5.000000000000000
1	3.000000000000000
2	**2**.333333333333333
3	**2.23**8095238095238
4	**2.23606**8895643363
5	**2.23606797749**978
6	**2.236067977499790**

Example (Cont.)

This example illustrates an empirical fact that is a consequence of quadratic convergence: Newton's method doubles the number of correct digits with each iteration. A partial explanation for this fact is as follows: From the proof of the convergence theorem we see that

$$|\xi - x_{k+1}| \leq C|\xi - x_k|^2,$$

so that, by taking base–10 logarithms (which essentially counts the number of correct digits) we have

$$\log_{10}|\xi - x_{k+1}| \leq 2\log_{10}|\xi - x_k| + \log_{10} C.$$

Example

Define $f : [1, 5] \to \mathbb{R}$ via $f(x) = (x - 3)^3$. Observe that, for this simple example, $\xi = 3$ is a non–simple root, i.e., $f(3) = f'(3) = f''(3) = 0$, which is something that the theory (up to this point) cannot handle. Nevertheless, Newton's method will still work. In particular, if Newton's method is employed with the starting point $x_0 = 4$ to approximate the root $\xi = 3$, then one can show directly that the convergence is exactly linear.

Reduced Convergence Rate for Non-Simple Roots

Theorem (non–simple roots)

Let $m \geq 2$ be a positive integer and I be a closed and bounded interval. Suppose that $f \in C^m(I)$, is such that there is $\xi \in I$, for which

$$f(\xi) = f'(\xi) = \cdots = f^{(m-1)}(\xi) = 0,$$

but $f^{(m)}(\xi) \neq 0$. If $|\xi - x_0|$ is sufficiently small, the sequence $\{x_k\}_{k=0}^{\infty}$ defined by Newton's method (6) is well–defined and converges to ξ exactly linearly with

$$\lim_{k\to\infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|} = \frac{m-1}{m} = \sigma \in (0, 1).$$

Proof.

We give a sketch of the proof. The details of well–definedness and convergence, in particular, are left for an exercise.

Proof Cont.

By Taylor's Theorem

$$f(x_k) = f(\xi) + f'(\xi)(x_k - \xi) + \cdots + f^{(m-1)}(\xi)\frac{(x_k-\xi)^{m-1}}{(m-1)!} + f^{(m)}(\eta_k)\frac{(x_k-\xi)^m}{m!}$$
$$= f^{(m)}(\eta_k)\frac{(x_k-\xi)^m}{m!},$$

for some η_k between x_k and ξ. Another application of Taylor's Theorem gives

$$f'(x_k) = f'(\xi) + f''(\xi)(x_k - \xi) + \cdots + f^{(m-1)}(\xi)\frac{(x_k-\xi)^{m-2}}{(m-2)!}$$
$$+ f^{(m)}(\zeta_k)\frac{(x_k-\xi)^{m-1}}{(m-1)!}$$
$$= f^{(m)}(\zeta_k)\frac{(x_k-\xi)^{m-1}}{(m-1)!},$$

for some ζ_k between x_k and ξ.

Proof Cont.

Thus, assuming $x_k \neq \xi$,

$$x_{k+1} - \xi = x_k - \xi - \frac{f(x_k)}{f'(x_k)} = x_k - \xi - \frac{f^{(m)}(\eta_k)\frac{(x_k-\xi)^m}{m!}}{f^{(m)}(\zeta_k)\frac{(x_k-\xi)^{m-1}}{(m-1)!}},$$

or

$$\frac{x_{k+1} - \xi}{x_k - \xi} = 1 - \frac{f^{(m)}(\eta_k)}{m \cdot f^{(m)}(\zeta_k)}.$$

Since $\eta_k, \zeta_k \to \xi$ as $k \to \infty$,

$$\lim_{k\to\infty} \frac{x_{k+1} - \xi}{x_k - \xi} = 1 - \frac{1}{m} = \frac{m-1}{m}.$$

□

Quasi-Newton Methods

Simplified Newton's Method

Recall, this method is defined by

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}.$$

Theorem (convergence)

Let $I \subseteq \mathbb{R}$ be an interval. Assume that $f \in C(I)$ is such that there is an $\xi \in I$ for which $f(\xi) = 0$, but $f'(\xi) \neq 0$ and $f''(\xi) \neq 0$. Set, for $\delta > 0$, $I_\delta = [\xi - \delta, \xi + \delta] \subseteq I$ and assume that for some $\delta > 0$, $f \in C^2(I_\delta)$, and $0 < \alpha \leq |f'(x)|$, for all $x \in I_\delta$. Set

$$A = \frac{\max_{x \in I_\delta} |f''(x)|}{\alpha}, \quad h = \min\left\{\delta, \frac{1}{3A}\right\}. \tag{11}$$

If $|\xi - x_0| \leq h$, then the sequence $\{x_k\}_{k=0}^{\infty}$ defined by the simplified Newton's method converges linearly to the zero ξ of f as $k \to \infty$.

Proof of Convergence of Simplified Newton's Method.

Suppose $x_0, x_k \in [\xi - h, \xi + h]$. Then we have

$$\begin{aligned} x_{k+1} - \xi &= \frac{1}{f'(x_0)}\Big[f'(x_0)\,(x_k - \xi) - f(x_k)\Big] \\ &= \frac{1}{f'(x_0)}\Big[\big(f'(x_0) - f'(\xi)\big)\,(x_k - \xi) - f(x_k) + f'(\xi)\,(x_k - \xi)\Big]\,. \end{aligned}$$

By the Mean Value Theorem, there is $\beta \in [\xi - h, \xi + h]$ between x_0 and ξ for which

$$f''(\beta)(x_0 - \xi) = f'(x_0) - f'(\xi).$$

Furthermore, for some $\eta_k \in [\xi - h, \xi + h]$ between ξ and x_k, we have

$$f(x_k) = f(\xi) + f'(\xi)(x_k - \xi) + \frac{f''(\eta_k)}{2}(x_k - \xi)^2$$

from Taylor's Theorem. Rearranging terms and using $f(\xi) = 0$ yields

$$-f(x_k) + f'(\xi)(x_k - \xi) = -\frac{f''(\eta_k)}{2}(x_k - \xi)^2.$$

Proof Cont.

Putting things together, we find

$$\begin{aligned} x_{k+1} - \xi &= \frac{1}{f'(x_0)} \left[f''(\beta)(x_0 - \xi)(x_k - \xi) - \frac{f''(\eta_k)}{2}(x_k - \xi)^2 \right] \\ &= \frac{1}{f'(x_0)} \left[f''(\beta)(x_0 - \xi) - \frac{f''(\eta_k)}{2}(x_k - \xi) \right] (x_k - \xi). \end{aligned}$$

Taking absolute values,

$$\begin{aligned} |x_{k+1} - \xi| &= \frac{1}{|f'(x_0)|} \left| f''(\beta)(x_0 - \xi) - \frac{f''(\eta_k)}{2}(x_k - \xi) \right| \cdot |x_k - \xi| \\ &\leq \frac{1}{|f'(x_0)|} \left(|f''(\beta)| \cdot |x_0 - \xi| + \frac{1}{2}|f''(\eta_k)| \cdot |x_k - \xi| \right) |x_k - \xi| \\ &\leq \left(A|x_0 - \xi| + \frac{1}{2}|x_k - \xi| A \right) |x_k - \xi| \\ &\leq \left(A \cdot \frac{1}{3A} + \frac{1}{2}\frac{1}{3A} A \right) |x_k - \xi| \\ &= \frac{1}{2}|x_k - \xi|. \end{aligned}$$

Proof Cont.

Hence $x_{k+1} \in [\xi - h, \xi + h]$, for any $k \in \mathbb{N}$, as long as $x_0, x_k \in [\xi - h, \xi + h]$. The simplified Newton algorithm is well–defined. Furthermore, it is clear that

$$|x_k - \xi| \leq \frac{h}{2^k},$$

which proves that $x_k \to \xi$ as $k \to \infty$.
The convergence is exactly linear, as can be seen from the error equation:

$$\lim_{k\to\infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|} = \frac{|f''(\beta)| \cdot |x_0 - \xi|}{|f'(x_0)|} = \mu.$$

By our assumptions $0 < \mu \leq 1/3 < 1$. □

Steffensen's Method

Definition (Steffensen's method)

Let $I \subseteq \mathbb{R}$ be an interval, $f \in C(I)$, and $x_0 \in I$. **Steffensen's method** is an algorithm for computing the terms of the sequence $\{x_k\}_{k=0}^{\infty}$ via

$$x_{k+1} = x_k - \frac{f(x_k)}{s_k}, \qquad s_k = \frac{f(x_k + f(x_k)) - f(x_k)}{f(x_k)}.$$

We say that this method is **well–defined** iff $x_0 \in I$ implies $x_k \in I$ for all $k = 1, 2, \ldots$. We say that this method **converges** iff there is $\xi \in I$, with $f(\xi) = 0$, such that $x_k \to \xi$ as $k \to \infty$.

Theorem (Convergence of Steffensen's Method)

Let $I \subseteq \mathbb{R}$ be an interval, and $f \in C(I)$ be such that, for some $\xi \in I$, $f(\xi) = 0$. Define, for $\delta > 0$, $I_\delta = [\xi - \delta, \xi + \delta] \subseteq I$. Assume that there is $\delta > 0$ for which $f \in C^2(I_\delta)$, $f'(\xi) \neq 0$, and $f''(\xi) \neq 0$. If $|\xi - x_0|$ is sufficiently small, then the sequence $\{x_k\}_{k=0}^{\infty}$ defined by Steffensen's method is well–defined and converges quadratically to the zero ξ as $k \to \infty$.

Proof.

See the textbook. □

Two-Step Newton's Method

Definition (two–step Newton)

Let $I \subseteq \mathbb{R}$ be an interval, and $f \in C^1(I)$. For $x_0 \in I$, with $f'(x_0) \neq 0$, the sequence $\{x_k\}_{k=0}^{\infty}$ defined by

$$y_k = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad x_{k+1} = y_k - \frac{f(y_k)}{f'(x_k)}, \tag{12}$$

is called the **two–step Newton's** method. We say that the method is **well–defined** if $x_k \in I$, and $f'(x_k) \neq 0$ for all $k \geq 0$. We say that the method **converges** if there is $\xi \in I$ such that $f(\xi) = 0$ and $x_k \to \xi$ as $k \to \infty$.

Theorem (Convergence of Two-Step Newton's Method)

Let $I \subseteq \mathbb{R}$ be an interval, $f \in C(I)$ is such that there is $\xi \in I$ for which $f(\xi) = 0$, but $f'(\xi) \neq 0$, and $f''(\xi) \neq 0$. Set, for $\delta > 0$, $I_\delta = [\xi - \delta, \xi + \delta] \subseteq I$ and assume there is $\delta > 0$ for which $f \in C^2(I_\delta)$, and $0 < \alpha \leq |f'(x)|$ for all $x \in I_\delta$. Set

$$A = \frac{\max_{x \in I_\delta} |f''(x)|}{\alpha}, \quad h = \min\left\{\delta, \frac{1}{A}\right\}.$$

If $|\xi - x_0| \leq h$, then the sequence $\{x_k\}_{k=0}^{\infty}$ defined by the two–step Newton's method, converges exactly cubically to the zero ξ as $k \to \infty$.

Proof.

Suppose $x_k \in [\xi - h, \xi + h] \subseteq I_\delta$. Then by Taylor's Theorem

$$0 = f(\xi) = f(x_k) + f'(x_k)(\xi - x_k) + \frac{f''(\eta_k)}{2}(\xi - x_k)^2,$$

for some η_k between x_k and ξ.

Proof Cont.

Note that $f'(x_k) \neq 0$ and, using the first equation in (12), we have that

$$\xi - y_k = -\frac{(\xi - x_k)^2}{2}\frac{f''(\eta_k)}{f'(x_k)} \ . \tag{13}$$

Hence

$$|\xi - y_k| = \frac{1}{2}\frac{|f''(\eta_k)|}{|f'(x_k)|} \cdot |\xi - x_k| \cdot |\xi - x_k| \leq \frac{A}{2} \cdot h \cdot |\xi - x_k| \leq \frac{1}{2}|\xi - x_k| \ \leq \frac{h}{2}.$$

We can conclude that if $x_k \in [\xi - h, \xi + h]$ then $y_k \in [\xi - h, \xi + h]$ as well. Now, using the second equation in (12), we have

$$\begin{aligned} x_{k+1} - \xi &= \frac{1}{f'(x_k)}\Big[f'(x_k)\,(y_k - \xi) - f(y_k)\Big] \\ &= \frac{1}{f'(x_k)}\Big[\big(f'(x_k) - f'(\xi)\big)\,(y_k - \xi) - f(y_k) + f'(\xi)\,(y_k - \xi)\Big]. \end{aligned}$$

Proof Cont.

By the Mean Value Theorem, there is $\beta_k \in [\xi - h, \xi + h]$ between x_k and ξ for which

$$f''(\beta_k)(x_k - \xi) = f'(x_k) - f'(\xi).$$

Furthermore, for some $\gamma_k \in [\xi - h, \xi + h]$ between ξ and y_k, we have

$$f(y_k) = f(\xi) + f'(\xi)(y_k - \xi) + \frac{f''(\gamma_k)}{2}(y_k - \xi)^2$$

from Taylor's Theorem. Rearranging terms and using $f(\xi) = 0$ yields

$$-f(y_k) + f'(\xi)(y_k - \xi) = -\frac{f''(\gamma_k)}{2}(y_k - \xi)^2.$$

Putting things together, we find

$$x_{k+1} - \xi = \frac{1}{f'(x_k)}\left[f''(\beta_k)(x_k - \xi)\,(y_k - \xi) - \frac{f''(\gamma_k)}{2}(y_k - \xi)^2\right]. \qquad (14)$$

Proof Cont.

Taking absolute values

$$
\begin{aligned}
|x_{k+1} - \xi| &= \frac{1}{|f'(x_k)|}\left|f''(\beta_k)\,(x_k - \xi)\,(y_k - \xi) - \frac{f''(\gamma_k)}{2}(y_k - \xi)^2\right| \\
&\leq A|x_k - \xi| \cdot |y_k - \xi| + \frac{1}{2}|y_k - \xi| \cdot |y_k - \xi| A \\
&\leq A|x_k - \xi|\frac{h}{2} + \frac{1}{2} \cdot \frac{h}{2} \cdot \frac{1}{2}|x_k - \xi| A \\
&\leq \frac{1}{2}|x_k - \xi| + \frac{1}{8}|x_k - \xi| \\
&= \frac{5}{8}|x_k - \xi|.
\end{aligned}
$$

We can conclude that if $x_k \in [\xi - h, \xi + h]$ then $x_{k+1} \in [\xi - h, \xi + h]$ as well. More importantly, we see by induction that

$$
|\xi - x_k| \leq \left(\frac{5}{8}\right)^k h,
$$

which proves that $x_k \to \xi$ as $k \to \infty$. Using this fact, it is easy to see that y_k, β_k, γ_k, $\eta_k \to \xi$ as $k \to \infty$ as well.

Proof Cont.

Now, using (14) we see that

$$\frac{x_{k+1}-\xi}{(x_k-\xi)(y_k-\xi)} = \frac{f''(\beta_k)}{f'(x_k)} - \frac{(y_k-\xi)}{2(x_k-\xi)}\frac{f''(\gamma_k)}{f'(x_k)}.$$

Making use of (13)

$$\frac{(x_{k+1}-\xi)}{(x_k-\xi)^3} = \frac{f''(\beta_k)f''(\eta_k)}{2(f'(x_k))^2} - \frac{1}{8}(x_k-\xi)\frac{(f''(\eta_k))^2}{(f'(x_k))^2}\frac{f''(\gamma_k)}{f'(x_k)}.$$

Taking limits, we have

$$\lim_{k\to\infty}\frac{|x_{k+1}-\xi|}{|x_k-\xi|^3} = \frac{1}{2}\left|\frac{f''(\xi)}{f'(\xi)}\right|^2 = \sigma \in (0,\infty).$$

This shows that the convergence is exactly cubic. □

Secant Method

Definition (secant method)

Let $I \subseteq \mathbb{R}$ be an interval, $f \in C(I)$, and $x_0 \in I$. **secant method** is an algorithm for computing the terms of the sequence $\{x_k\}_{k=0}^{\infty}$ via

$$x_{k+1} = x_k - \frac{f(x_k)}{s_k}, \quad s_k = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}, \quad k \geq 1, \tag{15}$$

We say that this method is **well–defined** iff $x_0, x_1 \in I$ implies $x_k \in I$ for all $k = 2, 3, \ldots$. We say that this method **converges** iff there is $\xi \in I$, with $f(\xi) = 0$, such that $x_k \to \xi$ as $k \to \infty$.

Theorem (Convergence of the Secant Method)

Let $I \subseteq \mathbb{R}$ be an interval, and $f \in C(I)$ be such that there is $\xi \in I$ for which $f(\xi) = 0$. Set, for $\delta > 0$, $I_\delta = [\xi - \delta, \xi + \delta] \subseteq I$ and assume there is $\delta > 0$ for which $f \in C^1(I_\delta)$ and, for simplicity, $f'(\xi) > 0$. The sequence $\{x_k\}_{k=0}^{\infty}$ defined by the secant method, converges (at least) linearly to the root ξ as $k \to \infty$, provided x_0 and x_1 are sufficiently close to ξ.

Proof.

Set $f'(\xi) = \alpha > 0$. By continuity there is no loss in generality in assuming that, for all $x \in I_\delta$,

$$0 < \frac{3\alpha}{4} \leq f'(x) \leq \frac{5\alpha}{4}.$$

Suppose now that $x_k, x_{k-1} \in I_\delta$. By the Mean Value Theorem, there is η_k, between x_k and x_{k-1}, such that $s_k = f'(\eta_k)$. Then,

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

Proof Cont.

By Taylor's Theorem, there is a γ_k between x_k and ξ, such that

$$f(x_k) = f(\xi) + f'(\gamma_k)(x_k - \xi) = f'(\gamma_k)(x_k - \xi).$$

Thus,

$$x_{k+1} - \xi = x_k - \xi - \frac{f'(\gamma_k)(x_k - \xi)}{f'(\eta_k)} = (x_k - \xi)\left[1 - \frac{f'(\gamma_k)}{f'(\eta_k)}\right].$$

Since

$$-\frac{2}{3} \leq 1 - \frac{f'(\gamma_k)}{f'(\eta_k)} \leq \frac{2}{5},$$

it follows that

$$|x_{k+1} - \xi| \leq \frac{2}{3}|x_k - \xi|.$$

If $|x_0 - \xi| \leq \delta$ and $|x_1 - \xi| \leq \delta$ then, by induction, we see that for $k \geq 2$,

$$|x_k - \xi| \leq \left(\frac{2}{3}\right)^{k-1} \delta.$$

This proves that $x_k \to \xi$ at least linearly. □

Super-Linear Convergence of the Secant Method

Theorem (super–linear convergence)

Let $I \subseteq \mathbb{R}$ be an interval and $f \in C(I)$. In the setting of the previous theorem, assume, in addition, that $f \in C^2(I_\delta)$ and $f''(\xi) > 0$. Then, the sequence $\{x_k\}_{k=0}^{\infty}$ generated by the secant method converges to ξ at the rate $q = \frac{1+\sqrt{5}}{2}$.

Proof.

A homework exercise. □

Newton's Method in Several Dimensions

Newton's Method in Several Dimensions

Definition (Newton's method)

Suppose $d \in \mathbb{N}$, $\Omega \subseteq \mathbb{R}^d$ *is an open, convex set,* $\mathbf{x}_0 \in \Omega$ *is given, and* $\mathbf{f} \in C^1(\Omega; \mathbb{R}^d)$. **Newton's method in** d**–dimensions** *is an algorithm for computing the terms of the sequence* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *via the recursive iteration*

$$\mathrm{J}_{\mathbf{f}}(\mathbf{x}_k)\,(\mathbf{x}_{k+1} - \mathbf{x}_k) = -\mathbf{f}(\mathbf{x}_k), \tag{16}$$

where $\mathrm{J}_{\mathbf{f}}$ *is the Jacobian matrix of* $\mathbf{f}$. *We say that the method is* **well–defined** *iff* $\mathbf{x}_k \in \Omega$, *and* $\mathrm{J}_{\mathbf{f}}(\mathbf{x}_k)$ *is non–singular, for all* $k \in \mathbb{N}$. *We say that Newton's method* **converges** *iff there is a* $\boldsymbol{\xi} \in \Omega$, *with* $\mathbf{f}(\boldsymbol{\xi}) = \mathbf{0}$, *such that* $\mathbf{x}_k \to \boldsymbol{\xi}$, *as* $k \to \infty$.

Theorem (Convergence in Several Dimensions)

Let $\boldsymbol{\xi} \in \mathbb{R}^d$ and $r > 0$ be given. Suppose that

$$\mathbf{f} \in C^2(\overline{B}(\boldsymbol{\xi}, r); \mathbb{R}^d),$$

$\mathbf{f}(\boldsymbol{\xi}) = \mathbf{0}$, *and for every* $\mathbf{x} \in \overline{B}(\boldsymbol{\xi}, r)$ *the Jacobian matrix* $\mathrm{J_f}(\mathbf{x})$ *is invertible, with the estimate*

$$\left\| [\mathrm{J}(\mathbf{x})]^{-1} \right\|_2 \leq \beta.$$

Then, the sequence $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *defined by Newton's method* (16) *converges (at least) quadratically to the root* $\boldsymbol{\xi}$ *as* $k \to \infty$, *provided* $\mathbf{x}_0$ *is sufficiently close to* $\boldsymbol{\xi}$.

Proof.

By Taylor's Theorem, for each $i = 1, \ldots, d$, there is a point $\boldsymbol{\eta}_{k,i} \in B(\boldsymbol{\xi}, r)$, such that,

$$0 = f_i(\boldsymbol{\xi}) = f_i(\mathbf{x}_k) + \nabla f_i(\mathbf{x}_k)^{\mathsf{T}} (\boldsymbol{\xi} - \mathbf{x}_k) + c_{k,i},$$

where $c_{k,i} = \frac{1}{2} (\boldsymbol{\xi} - \mathbf{x}_k)^{\mathsf{T}} \mathrm{H}_i(\boldsymbol{\eta}_{k,i}) (\boldsymbol{\xi} - \mathbf{x}_k)$, and H_i is the Hessian matrix of f_i.

Proof Cont.

To simplify notation in the proof, let us set $\mathbf{c}_k = [c_{k,1}, \ldots, c_{k,d}]^\mathsf{T}$, $\mathsf{J}_k = \mathsf{J}_\mathsf{f}(\mathbf{x}_k)$, and $\mathsf{H}^{(k,i)} = \mathsf{H}_i(\boldsymbol{\eta}_{k,i})$.

Using the definition of Newton's method together with our Taylor expansion, we get

$$\nabla f_i(\mathbf{x}_k)^\mathsf{T}(\mathbf{x}_{k+1} - \mathbf{x}_k) = -f_i(\mathbf{x}_k) = \nabla f_i(\mathbf{x}_k)^\mathsf{T}(\boldsymbol{\xi} - \mathbf{x}_k) + c_{k,i},$$

which simplifies to

$$\boldsymbol{\xi} - \mathbf{x}_{k+1} = -\mathsf{J}_k^{-1}\mathbf{c}_k.$$

Proof Cont.

Using the Cauchy–Schwarz and other basic inequalities,

$$\begin{aligned}
\left\| \mathrm{J}_k^{-1} \mathbf{c}_k \right\|_2 &\leq \left\| \mathrm{J}_k^{-1} \right\|_2 \left\| \mathbf{c}_k \right\|_2 \\
&\leq \beta \sqrt{\sum_{i=1}^{d} c_{k,i}^2} \\
&= \frac{\beta}{2} \sqrt{\sum_{i=1}^{d} \left| (\boldsymbol{\xi} - \mathbf{x}_k)^{\mathsf{T}} \, \mathrm{H}^{(k,i)} \, (\boldsymbol{\xi} - \mathbf{x}_k) \right|^2} \\
&\leq \frac{\beta}{2} \sqrt{\sum_{i=1}^{d} \left\| \boldsymbol{\xi} - \mathbf{x}_k \right\|_2^2 \left\| \mathrm{H}^{(k,i)} \, (\boldsymbol{\xi} - \mathbf{x}_k) \right\|_2^2} \\
&= \frac{\beta}{2} \left\| \boldsymbol{\xi} - \mathbf{x}_k \right\|_2 \sqrt{\sum_{i=1}^{d} \left\| \mathrm{H}^{(k,i)} \, (\boldsymbol{\xi} - \mathbf{x}_k) \right\|_2^2}.
\end{aligned}$$

Proof Cont.

Another application of Cauchy–Schwarz gives

$$\begin{aligned}
\left\| \mathrm{H}^{(k,i)} \left(\boldsymbol{\xi} - \mathbf{x}_k \right) \right\|_2^2 &= \sum_{j=1}^{d} \left| \sum_{m=1}^{d} \frac{\partial^2 f_i}{\partial x_j \partial x_m} (\boldsymbol{\eta}_{k,i}) (\xi_m - x_{k,m}) \right|^2 \\
&\leq \sum_{j=1}^{d} \left[\sum_{m=1}^{d} \left| \frac{\partial^2 f_i}{\partial x_j \partial x_m} (\boldsymbol{\eta}_{k,i}) \right|^2 \left\| \boldsymbol{\xi} - \mathbf{x}_k \right\|_2^2 \right] \\
&\leq \left\| \boldsymbol{\xi} - \mathbf{x}_k \right\|_2^2 \sum_{j=1}^{d} \left[\sum_{m=1}^{d} A^2 \right] \\
&= \left\| \boldsymbol{\xi} - \mathbf{x}_k \right\|_2^2 A^2 d^2,
\end{aligned}$$

where A is a upper bound on the absolute values of the second derivatives of $\mathbf{f}$, which is available because $\mathbf{f} \in C^2 \left(\overline{B}(\boldsymbol{\xi}, r); \mathbb{R}^d \right)$. We finally get the fundamental error estimate:

$$\left\| \boldsymbol{\xi} - \mathbf{x}_{k+1} \right\|_2 \leq \frac{\beta A d^{3/2}}{2} \left\| \boldsymbol{\xi} - \mathbf{x}_k \right\|_2^2 .$$

Proof Cont.

Starting with the fundamental error estimate,

$$\|\boldsymbol{\xi} - \mathbf{x}_{k+1}\|_2 \leq \frac{\beta A d^{3/2}}{2} \|\boldsymbol{\xi} - \mathbf{x}_k\|_2^2 ,$$

if $\|\boldsymbol{\xi} - \mathbf{x}_0\|_2 \leq \frac{1}{\beta A d^{3/2}} = h$, then $\|\boldsymbol{\xi} - \mathbf{x}_1\|_2 \leq \frac{1}{2} \|\boldsymbol{\xi} - \mathbf{x}_0\|_2$. By induction, it follows that

$$\|\boldsymbol{\xi} - \mathbf{x}_k\|_2 \leq h \left(\frac{1}{2}\right)^{2^k - 1} = \varepsilon_k .$$

Thus $\{\mathbf{x}_k\}_{k=0}^{\infty}$ is well–defined, $\mathbf{x}_k \to \boldsymbol{\xi}$, and the order is at least quadratic. □

Theorem (Kantorovich Theorem)

Let $d \in \mathbb{N}$. Suppose that $\Omega \subset \mathbb{R}^d$ is an open, bounded, convex set, $\mathbf{x}_0 \in \Omega$, and $\mathbf{f} \in C^1\left(\overline{\Omega}; \mathbb{R}^d\right)$. Assume, additionally, with $\mathrm{J_f}$ denoting the Jacobian matrix of $\mathbf{f}$, that there is $\gamma > 0$ such that

$$\left\|\mathrm{J_f}(\mathbf{x}) - \mathrm{J_f}(\mathbf{y})\right\|_2 \leq \gamma \left\|\mathbf{x} - \mathbf{y}\right\|_2, \quad \forall \mathbf{x}, \mathbf{y} \in \Omega.$$

Theorem (Kantorovich Theorem (Cont.))

Furthermore, let us assume the following:

1. *For all* $\mathbf{x} \in \Omega$, *the Jacobian matrix* $J_f(\mathbf{x})$ *is invertible, and there is* $\beta > 0$ *such that*
$$\left\| [J_f(\mathbf{x})]^{-1} \right\|_2 \leq \beta, \quad \forall \mathbf{x} \in \Omega.$$
2. *The initial iterate,* $\mathbf{x}_0 \in \Omega$, *satisfies*
$$\left\| [J_f(\mathbf{x}_0)]^{-1} \mathbf{f}(\mathbf{x}_0) \right\|_2 \leq \alpha.$$
3. *The parameters satisfy*
$$h = \frac{\alpha\beta\gamma}{2} < 1.$$
4. *The initial iterate is well inside* Ω, *in the sense that*
$$\overline{B}(\mathbf{x}_0, r) \subseteq \Omega,$$
where $r = \frac{\alpha}{1-h}$.

Theorem (Kantorovich Theorem (Cont.))

In this setting, the sequence $\mathbf{x}_k$ *defined by Newton's method* (16) *is well–defined, and, in particular,* $\mathbf{x}_k \in B(\mathbf{x}_0, r)$, *for each* $k \in \mathbb{N}$. *Moreover, there exists a point* $\boldsymbol{\xi} \in \overline{B}(\mathbf{x}_0, r)$, *such that* $\lim_{k\to\infty} \mathbf{x}_k = \boldsymbol{\xi}$, *with the convergence estimate*

$$\|\mathbf{x}_k - \boldsymbol{\xi}\|_2 \leq \alpha \frac{h^{2^k-1}}{1-h^{2^k}}, \qquad \forall\, k \in \mathbb{N}.$$

Since $0 < h < 1$, *convergence is at least quadratic. Finally, the point* $\boldsymbol{\xi}$ *is a zero of the function* $\mathbf{f}$, *that is* $\mathbf{f}(\boldsymbol{\xi}) = \mathbf{0}$.

Proof.

We split the proof into several steps.
(1): Since $[\mathrm{J_f}(\mathbf{x})]^{-1}$ exists for all $\mathbf{x} \in \Omega$, we will have that $\mathbf{x}_{k+1}$ is defined if $\mathbf{x}_k \in B(\mathbf{x}_0, r)$.

Proof Cont.

Suppose that, for all $j = 0, 1, \ldots, k$, $\mathbf{x}_j \in B(\mathbf{x}_0, r)$. Then,

$$\begin{aligned}
\|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2 &= \left\| [\mathrm{J_f}(\mathbf{x}_k)]^{-1} \mathbf{f}(\mathbf{x}_k) \right\|_2 \\
&\leq \left\| [\mathrm{J}(\mathbf{x}_k)]^{-1} \right\|_2 \|\mathbf{f}(\mathbf{x}_k)\|_2 \\
&\leq \beta \|\mathbf{f}(\mathbf{x}_k)\|_2 \\
&= \beta \|\mathbf{f}(\mathbf{x}_k) - \mathbf{f}(\mathbf{x}_{k-1}) - \mathrm{J}(\mathbf{x}_{k-1})(\mathbf{x}_k - \mathbf{x}_{k-1})\|_2 \\
&\leq \frac{\beta\gamma}{2} \|\mathbf{x}_k - \mathbf{x}_{k-1}\|_2^2 ,
\end{aligned} \tag{17}$$

using the result of a theorem in the last step. We claim that (17) implies that, for all $k \geq 0$,

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2 \leq \alpha h^{2^k - 1}. \tag{18}$$

The proof is by induction. The case $k = 0$ holds because of assumption 2:

$$\|\mathbf{x}_1 - \mathbf{x}_0\|_2 = \left\| [\mathrm{J_f}(\mathbf{x}_0)]^{-1} \mathbf{f}(\mathbf{x}_0) \right\|_2 \leq \alpha.$$

Proof Cont.

For the induction step we suppose that (18) is valid for $k = j - 1$:

$$\|\mathbf{x}_j - \mathbf{x}_{j-1}\|_2 \leq \alpha h^{2^{j-1}-1}.$$

Let $k = j$ now. Using (17) and the induction hypothesis,

$$\begin{aligned}\|\mathbf{x}_{j+1} - \mathbf{x}_j\|_2 &\leq \frac{\beta\gamma}{2}\|\mathbf{x}_j - \mathbf{x}_{j-1}\|_2^2 \leq \frac{\beta\gamma}{2}\alpha^2\left(h^{2^{j-1}-1}\right)^2 = \frac{\beta\gamma}{2}\alpha^2 h^{2^j-2} \\ &= \frac{\alpha\beta\gamma}{2}\alpha h^{2^j-2} = \alpha h^{2^j-1}.\end{aligned}$$

Hence, estimate (18) follows by induction. Now, by the triangle inequality,

$$\begin{aligned}\|\mathbf{x}_{k+1} - \mathbf{x}_0\|_2 &\leq \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2 + \cdots + \|\mathbf{x}_1 - \mathbf{x}_0\|_2 \\ &\leq \alpha\left(1 + h + h^3 + h^7 + \cdots + h^{2^k-1}\right) \\ &< \alpha\left(1 + h + h^2 + \cdots\right) \\ &= \frac{\alpha}{1-h} \\ &= r.\end{aligned}$$

Proof Cont.

Thus, $\mathbf{x}_{k+1} \in B(\mathbf{x}_0, r)$. By induction, $\mathbf{x}_k \in B(\mathbf{x}_0, r)$, for all $k \in \mathbb{N}$.
(2): Using (18), we can prove that $\{\mathbf{x}_k\}_{k=0}^{\infty}$ is a Cauchy sequence. Suppose $m > n \geq 0$. Then

$$\begin{aligned} \|\mathbf{x}_m - \mathbf{x}_n\|_2 &\leq \|\mathbf{x}_m - \mathbf{x}_{m-1}\|_2 + \cdots + \|\mathbf{x}_{n+1} - \mathbf{x}_n\|_2 \\ &\leq \alpha h^{2^n-1}\left(1 + h^{2^n} + h^{3\cdot 2^n} + h^{5\cdot 2^n} + \cdots\right) \\ &< \frac{\alpha h^{2^n-1}}{1 - h^{2^n}} \\ &< \varepsilon, \end{aligned} \tag{19}$$

provided n is sufficiently large. Since $\mathbf{x}_k$ is Cauchy, it converges to a unique limit point $\boldsymbol{\xi} \in \overline{B}(\mathbf{x}_0, r)$, appealing to the fact that $\overline{B}(\mathbf{x}_0, r)$ is closed. It follows on taking $m \to \infty$ in (19) that

$$\|\boldsymbol{\xi} - \mathbf{x}_n\|_2 < \frac{\alpha h^{2^n-1}}{1 - h^{2^n}}.$$

From this estimate it follows that convergence is at least quadratic.

Proof Cont.

(3): Finally, we prove that $\mathbf{f}(\boldsymbol{\xi}) = \mathbf{0}$. Since $\mathbf{x}_k \in B(\mathbf{x}_0, r)$,

$$\|\mathrm{J_f}(\mathbf{x}_k) - \mathrm{J_f}(\mathbf{x}_0)\|_2 \leq \gamma \|\mathbf{x}_0 - \mathbf{x}_k\|_2 \leq \gamma r.$$

Thus,

$$\|\mathrm{J_f}(\mathbf{x}_k)\|_2 = \|\mathrm{J_f}(\mathbf{x}_k) - \mathrm{J_f}(\mathbf{x}_0) + \mathrm{J_f}(\mathbf{x}_0)\|_2 \leq \gamma r + \|\mathrm{J_f}(\mathbf{x}_0)\|_2 = R.$$

As a consequence,

$$\|\mathbf{f}(\mathbf{x}_k)\|_2 = \|-\mathrm{J_f}(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k)\|_2 \leq R \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2,$$

which implies that $\lim_{k\to\infty} \|\mathbf{f}(\mathbf{x}_k)\|_2 = 0$. It follows that $\mathbf{f}(\boldsymbol{\xi}) = \mathbf{0}$. The proof is complete. □