Numerical Analysis I
Solutions of Equations in One Variable

Instructor: Wei-Cheng Wang [1]

Department of Mathematics
National TsingHua University

Fall 2010

---

[1] These slides are based on Prof. Tsung-Ming Huang(NTNU)'s original slides

Outline

Bisection

Idea

If $f(x) \in C$

Bisection method algorithm

Given $f(x)$ defined on (a, b), the maximal number of iterations M, and stop criteria δ and ε, this algorithm tries to locate one root of $f(x)$.

```
Compute $fa = f(a)$, $fb = f(b)$, and $e = b - a$
If  $sign(fa) = sign(fb)$, then  stop  End If
For $k = 1, 2, \ldots, M$
    $e = e/2$, $c = (a + b)/2$, $fc = f(c)$
    If $(|e| < \delta$ or $|fc| < \varepsilon)$, then  stop  End If
    If  $sign(fc) \neq sign(fa)$
        $b = c$, $fb = fc$
    Else
        $a = c$, $fa = fc$
    End If
End For
```

Let $\{c_n\}$ be the sequence of numbers produced. The algorithm should stop if one of the following conditions is satisfied.

1. the iteration number $k > M$,
2. $|c_k - c_{k-1}| < \delta$, or
3. $|f(c_k)| < \varepsilon$.

Let $[a_0, b_0], [a_1, b_1], \ldots$ denote the successive intervals produced by the bisection algorithm. Then

$$\begin{aligned} & a = a_0 \leq a_1 \leq a_2 \leq \cdots \leq b_0 = b \\ \Rightarrow \quad & \{a_n\} \text{ and } \{b_n\} \text{ are bounded} \\ \Rightarrow \quad & \lim_{n\to\infty} a_n \text{ and } \lim_{n\to\infty} b_n \text{ exist} \end{aligned}$$

Let $\{c_n\}$ be the sequence of numbers produced. The algorithm should stop if one of the following conditions is satisfied.

1. the iteration number $k > M$,
2. $|c_k - c_{k-1}| < \delta$, or
3. $|f(c_k)| < \varepsilon$.

Let $[a_0, b_0], [a_1, b_1], \ldots$ denote the successive intervals produced by the bisection algorithm. Then

$$\begin{aligned} & a = a_0 \le a_1 \le a_2 \le \cdots \le b_0 = b \\ \Rightarrow \quad & \{a_n\} \text{ and } \{b_n\} \text{ are bounded} \\ \Rightarrow \quad & \lim_{n\to\infty} a_n \text{ and } \lim_{n\to\infty} b_n \text{ exist} \end{aligned}$$

Let $\{c_n\}$ be the sequence of numbers produced. The algorithm should stop if one of the following conditions is satisfied.

1. the iteration number $k > M$,
2. $|c_k - c_{k-1}| < \delta$, or
3. $|f(c_k)| < \varepsilon$.

Let $[a_0, b_0], [a_1, b_1], \ldots$ denote the successive intervals produced by the bisection algorithm. Then

$$\begin{aligned} & a = a_0 \le a_1 \le a_2 \le \cdots \le b_0 = b \\ \Rightarrow \quad & \{a_n\} \text{ and } \{b_n\} \text{ are bounded} \\ \Rightarrow \quad & \lim_{n\to\infty} a_n \text{ and } \lim_{n\to\infty} b_n \text{ exist} \end{aligned}$$

Since

$$\begin{aligned} b_1 - a_1 &= \frac{1}{2}(b_0 - a_0) \\ b_2 - a_2 &= \frac{1}{2}(b_1 - a_1) = \frac{1}{4}(b_0 - a_0) \\ &\vdots \\ b_n - a_n &= \frac{1}{2^n}(b_0 - a_0) \end{aligned}$$

hence

$$\lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty} \frac{1}{2^n}(b_0 - a_0) = 0.$$

Therefore

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \equiv z.$$

Since f is a continuous function, we have that

$$\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n) = f(z) \quad \text{and} \quad \lim_{n\to\infty} f(b_n) = f(\lim_{n\to\infty} b_n) = f(z)$$

Since

$$\begin{aligned}
b_1 - a_1 &= \frac{1}{2}(b_0 - a_0) \\
b_2 - a_2 &= \frac{1}{2}(b_1 - a_1) = \frac{1}{4}(b_0 - a_0) \\
&\vdots \\
b_n - a_n &= \frac{1}{2^n}(b_0 - a_0)
\end{aligned}$$

hence

$$\lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty}\frac{1}{2^n}(b_0 - a_0) = 0.$$

Therefore

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \equiv z.$$

Since f is a continuous function, we have that

$$\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n) = f(z) \quad \text{and} \quad \lim_{n\to\infty} f(b_n) = f(\lim_{n\to\infty} b_n) = f(z).$$

Since

$$\begin{aligned} b_1 - a_1 &= \frac{1}{2}(b_0 - a_0) \\ b_2 - a_2 &= \frac{1}{2}(b_1 - a_1) = \frac{1}{4}(b_0 - a_0) \\ &\vdots \\ b_n - a_n &= \frac{1}{2^n}(b_0 - a_0) \end{aligned}$$

hence

$$\lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty}\frac{1}{2^n}(b_0 - a_0) = 0.$$

Therefore

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \equiv z.$$

Since f is a continuous function, we have that

$$\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n) = f(z) \quad \text{and} \quad \lim_{n\to\infty} f(b_n) = f(\lim_{n\to\infty} b_n) = f(z).$$

Since

$$\begin{aligned} b_1 - a_1 &= \frac{1}{2}(b_0 - a_0) \\ b_2 - a_2 &= \frac{1}{2}(b_1 - a_1) = \frac{1}{4}(b_0 - a_0) \\ &\vdots \\ b_n - a_n &= \frac{1}{2^n}(b_0 - a_0) \end{aligned}$$

hence

$$\lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty}\frac{1}{2^n}(b_0 - a_0) = 0.$$

Therefore

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \equiv z.$$

Since f is a continuous function, we have that

$$\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n) = f(z) \quad \text{and} \quad \lim_{n\to\infty} f(b_n) = f(\lim_{n\to\infty} b_n) = f(z).$$

On the other hand,

$$\begin{aligned} & f(a_n)f(b_n) < 0 \\ \Rightarrow \quad & \lim_{n\to\infty} f(a_n)f(b_n) = f^2(z) \leq 0 \\ \Rightarrow \quad & f(z) = 0 \end{aligned}$$

Therefore, the limit of the sequences $\{a_n\}$ and $\{b_n\}$ is a zero of f in $[a, b]$. Let $c_n = \frac{1}{2}(a_n + b_n)$. Then

$$\begin{aligned} |z - c_n| &= \left| \lim_{n\to\infty} a_n - \frac{1}{2}(a_n + b_n) \right| \\ &= \left| \frac{1}{2}\left[\lim_{n\to\infty} a_n - b_n \right] + \frac{1}{2}\left[\lim_{n\to\infty} a_n - a_n \right] \right| \\ &\leq \max\left\{ \left| \lim_{n\to\infty} a_n - b_n \right|, \left| \lim_{n\to\infty} a_n - a_n \right| \right\} \\ &\leq |b_n - a_n| = \frac{1}{2^n}|b_0 - a_0|. \end{aligned}$$

This proves the following theorem.

On the other hand,

$$\begin{aligned} & f(a_n)f(b_n) < 0 \\ \Rightarrow\quad & \lim_{n\to\infty} f(a_n)f(b_n) = f^2(z) \leq 0 \\ \Rightarrow\quad & f(z) = 0 \end{aligned}$$

Therefore, the limit of the sequences $\{a_n\}$ and $\{b_n\}$ is a zero of f in $[a, b]$.

Let $c_n = \frac{1}{2}(a_n + b_n)$. Then

$$\begin{aligned} |z - c_n| &= \Big|\lim_{n\to\infty} a_n - \frac{1}{2}(a_n + b_n)\Big| \\ &= \Big|\frac{1}{2}\left[\lim_{n\to\infty} a_n - b_n\right] + \frac{1}{2}\left[\lim_{n\to\infty} a_n - a_n\right]\Big| \\ &\leq \max\left\{\Big|\lim_{n\to\infty} a_n - b_n\Big|, \Big|\lim_{n\to\infty} a_n - a_n\Big|\right\} \\ &\leq |b_n - a_n| = \frac{1}{2^n}|b_0 - a_0|. \end{aligned}$$

This proves the following theorem.

On the other hand,

$$\begin{aligned} & f(a_n)f(b_n) < 0 \\ \Rightarrow \quad & \lim_{n\to\infty} f(a_n)f(b_n) = f^2(z) \le 0 \\ \Rightarrow \quad & f(z) = 0 \end{aligned}$$

Therefore, the limit of the sequences $\{a_n\}$ and $\{b_n\}$ is a zero of f in $[a, b]$. Let $c_n = \frac{1}{2}(a_n + b_n)$. Then

$$\begin{aligned} |z - c_n| &= \Big|\lim_{n\to\infty} a_n - \frac{1}{2}(a_n + b_n)\Big| \\ &= \Big|\frac{1}{2}\Big[\lim_{n\to\infty} a_n - b_n\Big] + \frac{1}{2}\Big[\lim_{n\to\infty} a_n - a_n\Big]\Big| \\ &\le \max\Big\{\Big|\lim_{n\to\infty} a_n - b_n\Big|, \Big|\lim_{n\to\infty} a_n - a_n\Big|\Big\} \\ &\le |b_n - a_n| = \frac{1}{2^n}|b_0 - a_0|. \end{aligned}$$

This proves the following theorem.

On the other hand,

$$\begin{aligned} & f(a_n)f(b_n) < 0 \\ \Rightarrow \quad & \lim_{n\to\infty} f(a_n)f(b_n) = f^2(z) \le 0 \\ \Rightarrow \quad & f(z) = 0 \end{aligned}$$

Therefore, the limit of the sequences $\{a_n\}$ and $\{b_n\}$ is a zero of f in $[a, b]$. Let $c_n = \frac{1}{2}(a_n + b_n)$. Then

$$\begin{aligned} |z - c_n| &= \Big| \lim_{n\to\infty} a_n - \frac{1}{2}(a_n + b_n) \Big| \\ &= \Big| \frac{1}{2}\left[\lim_{n\to\infty} a_n - b_n\right] + \frac{1}{2}\left[\lim_{n\to\infty} a_n - a_n\right] \Big| \\ &\le \max\left\{ \Big|\lim_{n\to\infty} a_n - b_n\Big|, \Big|\lim_{n\to\infty} a_n - a_n\Big| \right\} \\ &\le |b_n - a_n| = \frac{1}{2^n}|b_0 - a_0|. \end{aligned}$$

This proves the following theorem.

Theorem

Let $\{[a_n, b_n]\}$ *denote the intervals produced by the bisection algorithm. Then* $\lim_{n\to\infty} a_n$ *and* $\lim_{n\to\infty} b_n$ *exist, are equal, and represent a zero of* $f(x)$. *If*

$$z = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \quad \text{and} \quad c_n = \frac{1}{2}(a_n + b_n),$$

then

$$|z - c_n| \le \frac{1}{2^n}\,(b_0 - a_0)\,.$$

Remark

$\{c_n\}$ converges to z with the rate of $O(2^{-n})$.

Theorem

Let $\{[a_n, b_n]\}$ *denote the intervals produced by the bisection algorithm. Then* $\lim_{n\to\infty} a_n$ *and* $\lim_{n\to\infty} b_n$ *exist, are equal, and represent a zero of* $f(x)$. *If*

$$z = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \quad \text{and} \quad c_n = \frac{1}{2}(a_n + b_n),$$

then

$$|z - c_n| \leq \frac{1}{2^n}\left(b_0 - a_0\right).$$

Remark

$\{c_n\}$ converges to z with the rate of $O(2^{-n})$.

Example

How many steps should be taken to compute a root of $f(x) = x^3 + 4x^2 - 10 = 0$ on $[1, 2]$ with relative error 10^{-3}?

solution: Seek an n such that

$$\frac{|z - c_n|}{|z|} \le 10^{-3} \;\Rightarrow\; |z - c_n| \le |z| \times 10^{-3}.$$

Since $z \in [1, 2]$, it is sufficient to show

$$|z - c_n| \le 10^{-3}.$$

That is, we solve

$$2^{-n}(2 - 1) \le 10^{-3} \;\Rightarrow\; -n \log_{10} 2 \le -3$$

which gives $n \ge 10$.

Example

How many steps should be taken to compute a root of $f(x) = x^3 + 4x^2 - 10 = 0$ on $[1, 2]$ with relative error 10^{-3}?

solution: Seek an n such that

$$\frac{|z - c_n|}{|z|} \leq 10^{-3} \Rightarrow |z - c_n| \leq |z| \times 10^{-3}.$$

Since $z \in [1, 2]$, it is sufficient to show

$$|z - c_n| \leq 10^{-3}.$$

That is, we solve

$$2^{-n}(2 - 1) \leq 10^{-3} \Rightarrow -n \log_{10} 2 \leq -3$$

which gives $n \geq 10$.

Fixed-Point Iteration

Definition

x is called a fixed point of a given function f if $f(x) = x$.

Root-finding problems and fixed-point problems

- Find x^* such that $f(x^*) = 0$.
  Let $g(x) = x - f(x)$. Then $g(x^*) = x^* - f(x^*) = x^*$.
  $\Rightarrow x^*$ is a fixed point for $g(x)$.
- Find x^* such that $g(x^*) = x^*$.
  Define $f(x) = x - g(x)$ so that $f(x^*) = x^* - g(x^*) = x^* - x^* = 0$
  $\Rightarrow x^*$ is a zero of $f(x)$.

Fixed-Point Iteration

Definition

x is called a fixed point of a given function f if $f(x) = x$.

Root-finding problems and fixed-point problems

- Find x^* such that $f(x^*) = 0$.
  Let $g(x) = x - f(x)$. Then $g(x^*) = x^* - f(x^*) = x^*$.
  $\Rightarrow x^*$ is a fixed point for $g(x)$.
- Find x^* such that $g(x^*) = x^*$.
  Define $f(x) = x - g(x)$ so that $f(x^*) = x^* - g(x^*) = x^* - x^* = 0$
  $\Rightarrow x^*$ is a zero of $f(x)$.

Fixed-Point Iteration

Definition

x is called a fixed point of a given function f if $f(x) = x$.

Root-finding problems and fixed-point problems

- Find x^* such that $f(x^*) = 0$.
  Let $g(x) = x - f(x)$. Then $g(x^*) = x^* - f(x^*) = x^*$.
  $\Rightarrow x^*$ is a fixed point for $g(x)$.
- Find x^* such that $g(x^*) = x^*$.
  Define $f(x) = x - g(x)$ so that $f(x^*) = x^* - g(x^*) = x^* - x^* = 0$
  $\Rightarrow x^*$ is a zero of $f(x)$.

Example

The function $g($
and $x = 2$ since

$$g(-1)$$

Theorem (Existence and uniqueness)

1. *If* $g \in C[a,b]$ *such that* $a \le g(x) \le b$ *for all* $x \in [a,b]$*, then* g *has a fixed point in* $[a,b]$*.*
2. *If, in addition,* $g'(x)$ *exists in* (a,b) *and there exists a positive constant* $M < 1$ *such that* $|g'(x)| \le M < 1$ *for all* $x \in (a,b)$*. Then the fixed point is unique.*

Theorem (Existence and uniqueness)

1. *If* $g \in C[a, b]$ *such that* $a \leq g(x) \leq b$ *for all* $x \in [a, b]$, *then* g *has a fixed point in* $[a, b]$.
2. *If, in addition,* $g'(x)$ *exists in* (a, b) *and there exists a positive constant* $M < 1$ *such that* $|g'(x)| \leq M < 1$ *for all* $x \in (a, b)$. *Then the fixed point is unique.*

Proof

Existence:

- If $g(a) = a$ or $g(b) = b$, then a or b is a fixed point of g and we are done.
- Otherwise, it must be $g(a) > a$ and $g(b) < b$. The function $h(x) = g(x) - x$ is continuous on $[a, b]$, with

  $$h(a) = g(a) - a > 0 \text{ and } h(b) = g(b) - b < 0.$$

  By the Intermediate Value Theorem, $\exists\ x^* \in [a, b]$ such that $h(x^*) = 0$. That is

  $$g(x^*) - x^* = 0 \Rightarrow g(x^*) = x^*.$$

  Hence g has a fixed point x^* in $[a, b]$.

Proof

Existence:

- If $g(a) = a$ or $g(b) = b$, then a or b is a fixed point of g and we are done.
- Otherwise, it must be $g(a) > a$ and $g(b) < b$. The function $h(x) = g(x) - x$ is continuous on $[a, b]$, with

$$h(a) = g(a) - a > 0 \quad \text{and} \quad h(b) = g(b) - b < 0.$$

By the Intermediate Value Theorem, $\exists\, x^* \in [a, b]$ such that $h(x^*) = 0$. That is

$$g(x^*) - x^* = 0 \;\Rightarrow\; g(x^*) = x^*.$$

Hence g has a fixed point x^* in $[a, b]$.

Proof

Existence:

- If $g(a) = a$ or $g(b) = b$, then a or b is a fixed point of g and we are done.
- Otherwise, it must be $g(a) > a$ and $g(b) < b$. The function $h(x) = g(x) - x$ is continuous on $[a, b]$, with

$$h(a) = g(a) - a > 0 \ \text{ and } \ h(b) = g(b) - b < 0.$$

By the Intermediate Value Theorem, $\exists\ x^* \in [a, b]$ such that $h(x^*) = 0$. That is

$$g(x^*) - x^* = 0 \ \Rightarrow \ g(x^*) = x^*.$$

Hence g has a fixed point x^* in $[a, b]$.

Proof

Existence:

- If $g(a) = a$ or $g(b) = b$, then a or b is a fixed point of g and we are done.
- Otherwise, it must be $g(a) > a$ and $g(b) < b$. The function $h(x) = g(x) - x$ is continuous on $[a, b]$, with

$$h(a) = g(a) - a > 0 \text{ and } h(b) = g(b) - b < 0.$$

By the Intermediate Value Theorem, $\exists\ x^* \in [a, b]$ such that $h(x^*) = 0$. That is

$$g(x^*) - x^* = 0 \Rightarrow g(x^*) = x^*.$$

Hence g has a fixed point x^* in $[a, b]$.

Proof

Existence:

- If $g(a) = a$ or $g(b) = b$, then a or b is a fixed point of g and we are done.
- Otherwise, it must be $g(a) > a$ and $g(b) < b$. The function $h(x) = g(x) - x$ is continuous on $[a, b]$, with

  $$h(a) = g(a) - a > 0 \quad \text{and} \quad h(b) = g(b) - b < 0.$$

  By the Intermediate Value Theorem, $\exists\, x^* \in [a, b]$ such that $h(x^*) = 0$. That is

  $$g(x^*) - x^* = 0 \;\Rightarrow\; g(x^*) = x^*.$$

  Hence g has a fixed point x^* in $[a, b]$.

Proof

Uniqueness:
Suppose that $p \neq q$ are both fixed points of g in $[a, b]$. By the Mean-Value theorem, there exists ξ between p and q such that

$$g'(\xi) = \frac{g(p) - g(q)}{p - q} = \frac{p - q}{p - q} = 1.$$

However, this contradicts to the assumption that $|g'(x)| \leq M < 1$ for all x in $[a, b]$. Therefore the fixed point of g is unique. □

Proof

Uniqueness:

Suppose that $p \neq q$ are both fixed points of g in $[a, b]$. By the Mean-Value theorem, there exists ξ between p and q such that

$$g'(\xi) = \frac{g(p) - g(q)}{p - q} = \frac{p - q}{p - q} = 1.$$

However, this contradicts to the assumption that $|g'(x)| \leq M < 1$ for all x in $[a, b]$. Therefore the fixed point of g is unique. $\square$

Proof

Uniqueness:
Suppose that $p \neq q$ are both fixed points of g in $[a, b]$. By the Mean-Value theorem, there exists ξ between p and q such that

$$g'(\xi) = \frac{g(p) - g(q)}{p - q} = \frac{p - q}{p - q} = 1.$$

However, this contradicts to the assumption that $|g'(x)| \leq M < 1$ for all x in $[a, b]$. Therefore the fixed point of g is unique. □

Example

Show that the following function has a unique fixed point.

$$g(x) = (x^2 - 1)/3, \quad x \in [-1, 1].$$

Solution: The Extreme Value Theorem implies that

$$\min_{x\in[-1,1]} g(x) = g(0) = -\frac{1}{3},$$
$$\max_{x\in[-1,1]} g(x) = g(\pm 1) = 0.$$

That is $g(x) \in [-1, 1], \; \forall\, x \in [-1, 1]$.
Moreover, g is continuous and

$$|g'(x)| = \left|\frac{2x}{3}\right| \le \frac{2}{3}, \; \forall\, x \in (-1, 1).$$

By above theorem, g has a unique fixed point in $[-1, 1]$.

Let p be such unique fixed point of g. Then

$$\begin{aligned} p = g(p) = \frac{p^2 - 1}{3} \quad &\Rightarrow \quad p^2 - 3p - 1 = 0 \\ &\Rightarrow \quad p = \frac{1}{2}(3 - \sqrt{13}). \end{aligned}$$

□

Fixed-point iteration or functional iteration

Given a continuous function g, choose an initial point x_0 and generate $\{x_k\}_{k=0}^{\infty}$ by

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

$\{x_k\}$ may not converge, e.g., $g(x) = 3x$. However, when the sequence converges, say,

$$\lim_{k\to\infty} x_k = x^*,$$

then, since g is continuous,

$$g(x^*) = g(\lim_{k\to\infty} x_k) = \lim_{k\to\infty} g(x_k) = \lim_{k\to\infty} x_{k+1} = x^*.$$

That is, x^* is a fixed point of g.

Fixed-point iteration

```
Given $x_0$, tolerance $TOL$, maximum number of iteration $M$.
Set $i = 1$ and $x = g(x_0)$.
While $i \leq M$ and $|x - x_0| \geq TOL$
    Set $i = i + 1$, $x_0 = x$ and $x = g(x_0)$.
End While
```

Example

The equation

$$x^3 + 4x^2 - 10 = 0$$

has a unique root in $[1, 2]$. Change the equation to the fixed-point form $x = g(x)$.

(a) $x = g_1(x) \equiv x - f(x) = x - x^3 - 4x^2 + 10$

(b) $x = g_2(x) = \left(\frac{10}{x} - 4x\right)^{1/2}$

$$x^3 = 10 - 4x^2 \ \Rightarrow\ x^2 = \frac{10}{x} - 4x \ \Rightarrow\ x = \pm\left(\frac{10}{x} - 4x\right)^{1/2}$$

(c) $x = g_3(x) = \frac{1}{2}\left(10 - x^3\right)^{1/2}$

$$4x^2 = 10 - x^3 \quad \Rightarrow \quad x = \pm\frac{1}{2}\left(10 - x^3\right)^{1/2}$$

(d) $x = g_4(x) = \left(\frac{10}{4+x}\right)^{1/2}$

$$x^2(x+4) = 10 \quad \Rightarrow \quad x = \pm\left(\frac{10}{4+x}\right)^{1/2}$$

(e) $x = g_5(x) = x - \frac{x^3+4x^2-10}{3x^2+8x}$

$$x = g_5(x) \equiv x - \frac{f(x)}{f'(x)}$$

Results of the fixed-point iteration with initial point $x_0 = 1.5$

Theorem (Fixed-point Theorem)

Let $g \in [a,b]$ *be such that* $g(x) \in [a,b]$ *for all* $x \in [a,b]$. *Suppose that* g' *exists on* (a,b) *and that* $\exists\ k$ *with* $0 < k < 1$ *such that*

$$|g'(x)| \leq k, \ \forall\ x \in (a,b).$$

Then, for any number x_0 *in* $[a,b]$,

$$x_n = g(x_{n-1}), \ n \geq 1,$$

converges to the unique fixed point x *in* $[a,b]$.

Proof:

By the assumptions, a unique fixed point exists in $[a, b]$. Since $g([a, b]) \subseteq [a, b]$, $\{x_n\}_{n=0}^{\infty}$ is defined and $x_n \in [a, b]$ for all $n \geq 0$. Using the Mean Values Theorem and the fact that $|g'(x)| \leq k$, we have

$$|x - x_n| = |g(x_{n-1}) - g(x)| = |g'(\xi_n)||x - x_{n-1}| \leq k|x - x_{n-1}|,$$

where $\xi_n \in (a, b)$. It follows that

$$|x_n - x| \leq k|x_{n-1} - x| \leq k^2|x_{n-2} - x| \leq \cdots \leq k^n|x_0 - x|. \tag{1}$$

Since $0 < k < 1$, we have

$$\lim_{n\to\infty} k^n = 0$$

and

$$\lim_{n\to\infty} |x_n - x| \leq \lim_{n\to\infty} k^n|x_0 - x| = 0.$$

Hence, $\{x_n\}_{n=0}^{\infty}$ converges to x.

Corollary

If g satisfies the hypotheses of above theorem, then

$$|x - x_n| \leq k^n \max\{x_0 - a, b - x_0\}$$

and

$$|x_n - x| \leq \frac{k^n}{1-k}|x_1 - x_0|, \ \forall \ n \geq 1.$$

Proof: From (**??**),

$$|x_n - x| \leq k^n |x_0 - x| \leq k^n \max\{x_0 - a, b - x_0\}.$$

For $n \geq 1$, using the Mean Values Theorem,

$$|x_{n+1} - x_n| = |g(x_n) - g(x_{n-1})| \leq k|x_n - x_{n-1}| \leq \cdots \leq k^n|x_1 - x_0|.$$

Thus, for $m > n \geq 1$,

$$\begin{aligned}
|x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} - \cdots + x_{n+1} - x_n| \\
&\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \cdots + |x_{n+1} - x_n| \\
&\leq k^{m-1}|x_1 - x_0| + k^{m-2}|x_1 - x_0| + \cdots + k^n|x_1 - x_0| \\
&= k^n|x_1 - x_0| \left(1 + k + k^2 + \cdots + k^{m-n-1}\right).
\end{aligned}$$

It implies that

$$\begin{aligned}
|x - x_n| &= \lim_{m\to\infty} |x_m - x_n| \leq \lim_{m\to\infty} k^n|x_1 - x_0| \sum_{j=0}^{m-n-1} k^j \\
&\leq k^n|x_1 - x_0| \sum_{j=0}^{\infty} k^j = \frac{k^n}{1-k}|x_1 - x_0|.
\end{aligned}$$

Example

For previous example,

$$f(x) = x^3 + 4x^2 - 10 = 0.$$

For $g_1(x) = x - x^3 - 4x^2 + 10$, we have

$$g_1(1) = 6 \quad \text{and} \quad g_1(2) = -12,$$

so $g_1([1,2]) \not\subseteq [1,2]$. Moreover,

$$g_1'(x) = 1 - 3x^2 - 8x \quad \Rightarrow \quad |g_1'(x)| \geq 1 \ \forall \ x \in [1,2]$$

- DOES NOT guarantee to converge or not

For $g_3(x) = \frac{1}{2}(10 - x^3)^{1/2}, \ \forall\ x \in [1, 1.5]$,

$$g_3'(x) = -\frac{3}{4}x^2(10 - x^3)^{-1/2} < 0, \ \forall\ x \in [1, 1.5],$$

so g_3 is strictly decreasing on $[1, 1.5]$ and

$$1 < 1.28 \approx g_3(1.5) \leq g_3(x) \leq g_3(1) = 1.5, \ \forall\ x \in [1, 1.5].$$

On the other hand,

$$|g_3'(x)| \leq |g_3'(1.5)| \approx 0.66, \ \forall\ x \in [1, 1.5]$$

Hence, the sequence is convergent to the fixed point.

For $g_4(x) = \sqrt{10/(4+x)}$, we have

$$\sqrt{\frac{10}{6}} \le g_4(x) \le \sqrt{\frac{10}{5}},\ \forall\ x \in [1,2] \quad \Rightarrow \quad g_4([1,2]) \subseteq [1,2]$$

Moreover,

$$|g_4'(x)| = \left|\frac{-5}{\sqrt{10}(4+x)^{3/2}}\right| \le \frac{5}{\sqrt{10}(5)^{3/2}} < 0.15,\ \forall\ x \in [1,2].$$

The bound of $|g_4'(x)|$ is much smaller than the bound of $|g_3'(x)|$, which explains the more rapid convergence using g_4.

Newton's method

Suppose that $f : \mathbb{R} \to \mathbb{R}$ and $f \in C^2[a, b]$, i.e., f'' exists and is continuous. If $f(x^*) = 0$ and $x^* = x + h$ where h is small, then by Taylor's theorem

$$\begin{aligned} 0 = f(x^*) &= f(x+h) \\ &= f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + \frac{1}{3!}f'''(x)h^3 + \cdots \\ &= f(x) + f'(x)h + O(h^2). \end{aligned}$$

Since h is small, $O(h^2)$ is negligible. It is reasonable to drop $O(h^2)$ terms. This implies

$$f(x) + f'(x)h \approx 0 \quad \text{and} \quad h \approx -\frac{f(x)}{f'(x)}, \quad \text{if} \quad f'(x) \neq 0.$$

Hence

$$x + h = x - \frac{f(x)}{f'(x)}$$

is a better approximation to x^*.

This sets the stage for the Newton-Rapbson's method, which starts with an initial approximation x_0 and generates the sequence $\{x_n\}_{n=0}^{\infty}$ defined by

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

Since the Taylor's expansion of $f(x)$ at x_k is given by

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(x_k)(x - x_k)^2 + \cdots .$$

At x_k, one uses the tangent line

$$y = \ell(x) = f(x_k) + f'(x_k)(x - x_k)$$

to approximate the curve of $f(x)$ and uses the zero of the tangent line to approximate the zero of $f(x)$.

Newton's Method

```
Given $x_0$, tolerance $TOL$, maximum number of iteration $M$.
Set $i = 1$ and $x = x_0 - f(x_0)/f'(x_0)$.
While $i \le M$ and $|x - x_0| \ge TOL$
    Set $i = i + 1$, $x_0 = x$ and $x = x_0 - f(x_0)/f'(x_0)$.
End While
```

Three stopping-technique inequalities

$$\begin{aligned}
&(a). \quad |x_n - x_{n-1}| < \varepsilon, \\
&(b). \quad \frac{|x_n - x_{n-1}|}{|x_n|} < \varepsilon, \quad x_n \neq 0, \\
&(c). \quad |f(x_n)| < \varepsilon.
\end{aligned}$$

Note that Newton's method for solving $f(x) = 0$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad \text{for} \quad n \geq 1$$

is just a special case of functional iteration in which

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

Example

The following table shows the convergence behavior of Newton's method applied to solving $f(x) = x^2 - 1 = 0$. Observe the quadratic convergence rate.

n	x_n	$\|e_n\| \equiv \|1 - x_n\|$
0	2.0	1
1	1.25	0.25
2	1.025	2.5e-2
3	1.0003048780488	3.048780488e-4
4	1.0000000464611	4.64611e-8
5	1.0	0

Theorem

Assume $f(x^*) = 0,\ f'(x^*) \neq 0$ *and* $f(x)$, $f'(x)$ *and* $f''(x)$ *are continuous on* $N_\varepsilon(x^*)$. *Then if* x_0 *is chosen sufficiently close to* x^*, *then*

$$\left\{ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \right\} \to x^*.$$

Proof: Define

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

Find an interval $[x^* - \delta, x^* + \delta]$ such that

$$g([x^* - \delta, x^* + \delta]) \subseteq [x^* - \delta, x^* + \delta]$$

and

$$|g'(x)| \leq k < 1,\ \forall\, x \in (x^* - \delta, x^* + \delta).$$

Since f' is continuous and $f'(x^*) \neq 0$, it implies that $\exists\ \delta_1 > 0$ such that $f'(x) \neq 0\ \forall\ x \in [x^* - \delta_1, x^* + \delta_1] \subseteq [a, b]$. Thus, g is defined and continuous on $[x^* - \delta_1, x^* + \delta_1]$. Also

$$g'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2},$$

for $x \in [x^* - \delta_1, x^* + \delta_1]$. Since f'' is continuous on $[a, b]$, we have g' is continuous on $[x^* - \delta_1, x^* + \delta_1]$.

By assumption $f(x^*) = 0$, so

$$g'(x^*) = \frac{f(x^*)f''(x^*)}{|f'(x^*)|^2} = 0.$$

Since g' is continuous on $[x^* - \delta_1, x^* + \delta_1]$ and $g'(x^*) = 0$, $\exists\ \delta$ with $0 < \delta < \delta_1$ and $k \in (0, 1)$ such that

$$|g'(x)| \leq k,\ \forall\ x \in [x^* - \delta, x^* + \delta].$$

Claim: $g([x^* - \delta, x^* + \delta]) \subseteq [x^* - \delta, x^* + \delta]$.
If $x \in [x^* - \delta, x^* + \delta]$, then, by the Mean Value Theorem, $\exists\ \xi$ between x and x^* such that

$$|g(x) - g(x^*)| = |g'(\xi)||x - x^*|.$$

It implies that

$$\begin{aligned} |g(x) - x^*| &= |g(x) - g(x^*)| = |g'(\xi)||x - x^*| \\ &\leq k|x - x^*| < |x - x^*| < \delta. \end{aligned}$$

Hence, $g([x^* - \delta, x^* + \delta]) \subseteq [x^* - \delta, x^* + \delta]$.
By the Fixed-Point Theorem, the sequence $\{x_n\}_{n=0}^{\infty}$ defined by

$$x_n = g(x_{n-1}) = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, \quad \text{for} \quad n \geq 1,$$

converges to x^* for any $x_0 \in [x^* - \delta, x^* + \delta]$.

Example

When Newton's method applied to $f(x) = \cos x$ with starting point $x_0 = 3$, which is close to the root $\frac{\pi}{2}$ of f, it produces $x_1 = -4.01525, x_2 = -4.8526, \cdots$, which converges to another root $-\frac{3\pi}{2}$.

Secant method

Disadvantage of Newton's method

In many applications, the derivative $f'(x)$ is very expensive to compute, or the function $f(x)$ is not given in an algebraic formula so that $f'(x)$ is not available.

By definition,

$$f'(x_{n-1}) = \lim_{x \to x_{n-1}} \frac{f(x) - f(x_{n-1})}{x - x_{n-1}}.$$

Letting $x = x_{n-2}$, we have

$$f'(x_{n-1}) \approx \frac{f(x_{n-2}) - f(x_{n-1})}{x_{n-2} - x_{n-1}} = \frac{f(x_{n-1}) - f(x_{n-2})}{x_{n-1} - x_{n-2}}.$$

Using this approximation for $f'(x_{n-1})$ in Newton's formula gives

$$x_n = x_{n-1} - \frac{f(x_{n-1})(x_{n-1} - x_{n-2})}{f(x_{n-1}) - f(x_{n-2})},$$

which is called the Secant method.

From geometric point of view, we use a secant line through x_{n-1} and x_{n-2} instead of the tangent line to approximate the function at the point x_{n-1}. The slope of the secant line is

$$s_{n-1} = \frac{f(x_{n-1}) - f(x_{n-2})}{x_{n-1} - x_{n-2}}$$

and the equation is

$$M(x) = f(x_{n-1}) + s_{n-1}(x - x_{n-1}).$$

The zero of the secant line

$$x = x_{n-1} - \frac{f(x_{n-1})}{s_{n-1}} = x_{n-1} - f(x_{n-1}) \frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})}$$

is then used as a new approximate x_n.

Secant Method

```
Given $x_0, x_1$, tolerance $TOL$, maximum number of iteration $M$.
Set $i = 2$; $y_0 = f(x_0); y_1 = f(x_1)$;
    $x = x_1 - y_1(x_1 - x_0)/(y_1 - y_0)$.
While $i \leq M$ and $|x - x_1| \geq TOL$
    Set $i = i + 1$; $x_0 = x_1; y_0 = y_1; x_1 = x; y_1 = f(x)$;
        $x = x_1 - y_1(x_1 - x_0)/(y_1 - y_0)$.
End While
```

Method of False Position

1. Choose initial approximations x_0 and x_1 with $f(x_0)f(x_1) < 0$.
2. $x_2 = x_1 - f(x_1)(x_1 - x_0)/(f(x_1) - f(x_0))$
3. Decide which secant line to use to compute x_3:
   If $f(x_2)f(x_1) < 0$, then x_1 and x_2 bracket a root, i.e.,

   $$x_3 = x_2 - f(x_2)(x_2 - x_1)/(f(x_2) - f(x_1))$$

   Else, x_0 and x_2 bracket a root, i.e.,

   $$x_3 = x_2 - f(x_2)(x_2 - x_0)/(f(x_2) - f(x_0))$$

   End if

Method of False Position

```
Given $x_0, x_1$, tolerance $TOL$, maximum number of iteration $M$.
Set $i = 2$; $y_0 = f(x_0); y_1 = f(x_1)$; $x = x_1 - y_1(x_1 - x_0)/(y_1 - y_0)$.
While $i \le M$ and $|x - x_1| \ge TOL$
    Set $i = i + 1$; $y = f(x)$.
    If $y \cdot y_1 < 0$, then set $x_0 = x_1; y_0 = y_1$.
    Set $x_1 = x; y_1 = y$; $x = x_1 - y_1(x_1 - x_0)/(y_1 - y_0)$.
End While
```

Error analysis for iterative methods

Definition

Let $\{x_n\} \to x^*$. If there are positive constants c and α such that

$$\lim_{n\to\infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^\alpha} = c,$$

then we say the rate of convergence is of order α.

We say that the rate of convergence is

1. linear if $\alpha = 1$ and $0 < c < 1$.
2. superlinear if
$$\lim_{n\to\infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|} = 0;$$
3. quadratic if $\alpha = 2$.

Suppose that $\{x_n\}_{n=0}^{\infty}$ and $\{\tilde{x}_n\}_{n=0}^{\infty}$ are linearly and quadratically convergent to x^*, respectively, with the same constant $c = 0.5$. For simplicity, suppose that

$$\frac{|x_{n+1} - x^*|}{|x_n - x^*|} \approx c \quad \text{and} \quad \frac{|\tilde{x}_{n+1} - x^*|}{|\tilde{x}_n - x^*|^2} \approx c.$$

These imply that

$$|x_n - x^*| \approx c|x_{n-1} - x^*| \approx c^2|x_{n-2} - x^*| \approx \cdots \approx c^n|x_0 - x^*|,$$

and

$$\begin{aligned} |\tilde{x}_n - x^*| &\approx c|\tilde{x}_{n-1} - x^*|^2 \approx c\left[c|\tilde{x}_{n-2} - x^*|^2\right]^2 = c^3|\tilde{x}_{n-2} - x^*|^4 \\ &\approx c^3\left[c|\tilde{x}_{n-3} - x^*|^2\right]^4 = c^7|\tilde{x}_{n-3} - x^*|^8 \\ &\approx \cdots \approx c^{2^n-1}|\tilde{x}_0 - x^*|^{2^n}. \end{aligned}$$

Remark

Quadratically convergent sequences generally converge much more quickly thank those that converge only linearly.

Theorem

Let $g \in C[a,b]$ *with* $g([a,b]) \subseteq [a,b]$*. Suppose that* g' *is continuous on* (a,b) *and* $\exists\ k \in (0,1)$ *such that*

$$|g'(x)| \leq k,\ \forall\ x \in (a,b).$$

If $g'(x^*) \neq 0$*, then for any* $x_0 \in [a,b]$*, the sequence*

$$x_n = g(x_{n-1}), \quad \text{for} \quad n \geq 1$$

converges only linearly to the unique fixed point x^* *in* $[a,b]$*.*

Proof:

- By the Fixed-Point Theorem, the sequence $\{x_n\}_{n=0}^{\infty}$ converges to x^*.
- Since g' exists on (a, b), by the Mean Value Theorem, $\exists\ \xi_n$ between x_n and x^* such that

$$x_{n+1} - x^* = g(x_n) - g(x^*) = g'(\xi_n)(x_n - x^*).$$

- $\because \{x_n\}_{n=0}^{\infty} \to x^* \quad \Rightarrow \quad \{\xi_n\}_{n=0}^{\infty} \to x^*$
- Since g' is continuous on (a, b), we have

$$\lim_{n\to\infty} g'(\xi_n) = g'(x^*).$$

- Thus,

$$\lim_{n\to\infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|} = \lim_{n\to\infty} |g'(\xi_n)| = |g'(x^*)|.$$

Hence, if $g'(x^*) \neq 0$, fixed-point iteration exhibits linear convergence.

Theorem

Let x^* be a fixed point of g and I be an open interval with $x^* \in I$. Suppose that $g'(x^*) = 0$ and g'' is continuous with

$$|g''(x)| < M, \ \forall \ x \in I.$$

Then $\exists \ \delta > 0$ such that

$$\{x_n = g(x_{n-1})\}_{n=1}^{\infty} \ \rightarrow \ x^* \quad \textit{for} \quad x_0 \in [x^* - \delta, x^* + \delta]$$

at least quadratically. Moreover,

$$|x_{n+1} - x^*| < \frac{M}{2}|x_n - x^*|^2, \quad \textit{for sufficiently large} \quad n.$$

Proof:

- Since $g'(x^*) = 0$ and g' is continuous on I, $\exists\ \delta$ such that $[x^* - \delta, x^* + \delta] \subset I$ and

$$|g'(x)| \leq k < 1,\ \forall\ x \in [x^* - \delta, x^* + \delta].$$

- In the proof of the convergence for Newton's method, we have

$$\{x_n\}_{n=0}^{\infty} \subset [x^* - \delta, x^* + \delta].$$

- Consider the Taylor expansion of $g(x_n)$ at x^*

$$\begin{aligned} x_{n+1} = g(x_n) &= g(x^*) + g'(x^*)(x_n - x^*) + \frac{g''(\xi)}{2}(x_n - x^*)^2 \\ &= x^* + \frac{g''(\xi)}{2}(x_n - x^*)^2, \end{aligned}$$

where ξ lies between x_n and x^*.

- Since

$$|g'(x)| \leq k < 1, \ \forall \ x \in [x^* - \delta, x^* + \delta]$$

and

$$g([x^* - \delta, x^* + \delta]) \subseteq [x^* - \delta, x^* + \delta],$$

it follows that $\{x_n\}_{n=0}^{\infty}$ converges to x^*.

- But ξ_n is between x_n and x^* for each n, so $\{\xi_n\}_{n=0}^{\infty}$ also converges to x^* and

$$\lim_{n\to\infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^2} = \frac{|g''(x^*)|}{2} < \frac{M}{2}.$$

- It implies that $\{x_n\}_{n=0}^{\infty}$ is quadratically convergent to x^* if $g''(x^*) \neq 0$ and

$$|x_{n+1} - x^*| < \frac{M}{2}|x_n - x^*|^2, \quad \text{for sufficiently large} \quad n. \qquad \square$$

For Newton's method,

$$g(x) = x - \frac{f(x)}{f'(x)} \;\Rightarrow\; g'(x) = 1 - \frac{f'(x)}{f'(x)} + \frac{f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}$$

It follows that $g'(x^*) = 0$. Hence Newton's method is locally quadratically convergent.

Error Analysis of Secant Method

Reference: D. Kincaid and W. Cheney, "Numerical analysis"
Let x^* denote the exact solution of $f(x) = 0$, $e_k = x_k - x^*$ be the errors at the k-th step. Then

$$\begin{aligned}
e_{k+1} &= x_{k+1} - x^* \\
&= x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} - x^* \\
&= \frac{1}{f(x_k) - f(x_{k-1})}\left[(x_{k-1} - x^*)f(x_k) - (x_k - x^*)f(x_{k-1})\right] \\
&= \frac{1}{f(x_k) - f(x_{k-1})}\left(e_{k-1}f(x_k) - e_k f(x_{k-1})\right) \\
&= e_k e_{k-1}\left(\frac{\frac{1}{e_k}f(x_k) - \frac{1}{e_{k-1}}f(x_{k-1})}{x_k - x_{k-1}} \cdot \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}\right)
\end{aligned}$$

To estimate the numerator $\frac{\frac{1}{e_k}f(x_k)-\frac{1}{e_{k-1}}f(x_k)}{x_k-x_{k-1}}$, we apply the Taylor's theorem

$$f(x_k)=f(x^*+e_k)=f(x^*)+f'(x^*)e_k+\frac{1}{2}f''(x^*)e_k^2+O(e_k^3),$$

to get

$$\frac{1}{e_k}f(x_k)=f'(x^*)+\frac{1}{2}f''(x^*)e_k+O(e_k^2).$$

Similarly,

$$\frac{1}{e_{k-1}}f(x_{k-1})=f'(x^*)+\frac{1}{2}f''(x^*)e_{k-1}+O(e_{k-1}^2).$$

Hence

$$\frac{1}{e_k}f(x_k)-\frac{1}{e_{k-1}}f(x_{k-1})\approx\frac{1}{2}(e_k-e_{k-1})f''(x^*).$$

Since $x_k-x_{k-1}=e_k-e_{k-1}$ and

$$\frac{x_k-x_{k-1}}{f(x_k)-f(x_{k-1})}\to\frac{1}{f'(x^*)},$$

we have

$$\begin{aligned} e_{k+1} &\approx e_k e_{k-1}\left(\frac{\frac{1}{2}(e_k - e_{k-1})f''(x^*)}{e_k - e_{k-1}} \cdot \frac{1}{f'(x^*)}\right) = \frac{1}{2}\frac{f''(x^*)}{f'(x^*)} e_k e_{k-1} \\ &\equiv C e_k e_{k-1} \end{aligned} \tag{2}$$

To estimate the convergence rate, we assume

$$|e_{k+1}| \approx \eta |e_k|^{\alpha},$$

where $\eta > 0$ and $\alpha > 0$ are constants, i.e.,

$$\frac{|e_{k+1}|}{\eta |e_k|^{\alpha}} \to 1 \quad \text{as} \quad k \to \infty.$$

Then $|e_k| \approx \eta |e_{k-1}|^{\alpha}$ which implies $|e_{k-1}| \approx \eta^{-1/\alpha} |e_k|^{1/\alpha}$. Hence (**??**) gives

$$\eta |e_k|^{\alpha} \approx C |e_k| \eta^{-1/\alpha} |e_k|^{1/\alpha} \quad \Longrightarrow \quad C^{-1}\eta^{1+\frac{1}{\alpha}} \approx |e_k|^{1-\alpha+\frac{1}{\alpha}}.$$

Since $|e_k| \to 0$ as $k \to \infty$, and $C^{-1}\eta^{1+\frac{1}{\alpha}}$ is a nonzero constant,

$$1 - \alpha + \frac{1}{\alpha} = 0 \quad \Longrightarrow \quad \alpha = \frac{1+\sqrt{5}}{2} \approx 1.62.$$

This result implies that $C^{-1}\eta^{1+\frac{1}{\alpha}} \to 1$ and

$$\eta \to C^{\frac{\alpha}{1+\alpha}} = \left(\frac{f''(x^*)}{2f'(x^*)}\right)^{0.62}.$$

In summary, we have shown that

$$|e_{k+1}| = \eta |e_k|^{\alpha}, \quad \alpha \approx 1.62,$$

that is, the rate of convergence is superlinear.
Rate of convergence:

- secant method: superlinear
- Newton's method: quadratic
- bisection method: linear

Each iteration of method requires

- secant method: one function evaluation
- Newton's method: two function evaluation, namely, $f(x_k)$ and $f'(x_k)$.
  $\Rightarrow$ two steps of secant method are comparable to one step of Newton's method. Thus

$$|e_{k+2}| \approx \eta |e_{k+1}|^{\alpha} \approx \eta^{1+\alpha} |e_k|^{\frac{3+\sqrt{5}}{2}} \approx \eta^{1+\alpha} |e_k|^{2.62}.$$

  $\Rightarrow$ secant method is more efficient than Newton's method.

Remark

Two steps of secant method would require a little more work than one step of Newton's method.

Accelerating convergence

Aitken's Δ^2 method

- Accelerate the convergence of a sequence that is linearly convergent.
- Suppose $\{y_n\}_{n=0}^{\infty}$ is a linearly convergent sequence with limit y. Construct a sequence $\{\hat{y}_n\}_{n=0}^{\infty}$ that converges more rapidly to y than $\{y_n\}_{n=0}^{\infty}$.

For n sufficiently large,

$$\frac{y_{n+1} - y}{y_n - y} \approx \frac{y_{n+2} - y}{y_{n+1} - y}.$$

Then

$$(y_{n+1} - y)^2 \approx (y_{n+2} - y)(y_n - y),$$

so

$$y_{n+1}^2 - 2y_{n+1}y + y^2 \approx y_{n+2}y_n - (y_{n+2} + y_n)y + y^2$$

and

$$(y_{n+2} + y_n - 2y_{n+1})y \approx y_{n+2}y_n - y_{n+1}^2.$$

Solving for y gives

$$\begin{aligned} y &\approx \frac{y_{n+2}y_n - y_{n+1}^2}{y_{n+2} - 2y_{n+1} + y_n} \\ &= \frac{y_n y_{n+2} - 2y_n y_{n+1} + y_n^2 - y_n^2 + 2y_n y_{n+1} - y_{n+1}^2}{y_{n+2} - 2y_{n+1} + y_n} \\ &= \frac{y_n(y_{n+2} - 2y_{n+1} + y_n) - (y_{n+1} - y_n)^2}{(y_{n+2} - y_{n+1}) - (y_{n+1} - y_n)} \\ &= y_n - \frac{(y_{n+1} - y_n)^2}{(y_{n+2} - y_{n+1}) - (y_{n+1} - y_n)}. \end{aligned}$$

Aitken's Δ^2 method

$$\hat{y}_n = y_n - \frac{(y_{n+1} - y_n)^2}{(y_{n+2} - y_{n+1}) - (y_{n+1} - y_n)}. \tag{3}$$

Example

The sequence $\{y_n = \cos(1/n)\}_{n=1}^{\infty}$ converges linearly to $y = 1$.

n	y_n	$\hat{y}_n$
1	0.54030	0.96178
2	0.87758	0.98213
3	0.94496	0.98979
4	0.96891	0.99342
5	0.98007	0.99541
6	0.98614	
7	0.98981	

- $\{\hat{y}_n\}_{n=1}^{\infty}$ converges more rapidly to $y = 1$ than $\{y_n\}_{n=1}^{\infty}$.

Definition

For a given sequence $\{y_n\}_{n=0}^{\infty}$, the forward difference Δy_n is defined by

$$\Delta y_n = y_{n+1} - y_n, \quad \text{for} \quad n \geq 0.$$

Higher powers of Δ are defined recursively by

$$\Delta^k y_n = \Delta(\Delta^{k-1} y_n), \quad \text{for} \quad k \geq 2.$$

The definition implies that

$$\Delta^2 y_n = \Delta(y_{n+1} - y_n) = \Delta y_{n+1} - \Delta y_n = (y_{n+2} - y_{n+1}) - (y_{n+1} - y_n).$$

So the formula for $\hat{y}_n$ in (**??**) can be written as

$$\hat{y}_n = y_n - \frac{(\Delta y_n)^2}{\Delta^2 y_n}, \quad \text{for} \quad n \geq 0.$$

Theorem

Suppose $\{y_n\}_{n=0}^{\infty} \to y$ *linearly and*

$$\lim_{n\to\infty} \frac{y_{n+1} - y}{y_n - y} < 1.$$

Then $\{\hat{y}_n\}_{n=0}^{\infty} \to y$ *faster than* $\{y_n\}_{n=0}^{\infty}$ *in the sense that*

$$\lim_{n\to\infty} \frac{\hat{y}_n - y}{y_n - y} = 0.$$

- Aitken's Δ^2 method constructs the terms in order:

$$y_0, \quad y_1 = g(y_0), \quad y_2 = g(y_1), \quad \hat{y}_0 = \{\Delta^2\}(y_0), \quad y_3 = g(y_2),$$
$$\hat{y}_1 = \{\Delta^2\}(y_1), \quad \ldots.$$

$\Rightarrow$ Assume $|\hat{y}_0 - y| < |y_2 - y|$

- Steffensen's method constructs the terms in order:

$$\begin{aligned} &y_0^{(0)} \equiv y_0, && y_1^{(0)} = g(y_0^{(0)}), && y_2^{(0)} = g(y_1^{(0)}), \\ &y_0^{(1)} = \{\Delta^2\}(y_0^{(0)}), && y_1^{(1)} = g(y_0^{(1)}), && y_2^{(1)} = g(y_1^{(1)}), && \ldots. \end{aligned}$$

Steffensen's method (To find a solution of $y = g(y)$)

```
Given $y_0$, tolerance $TOL$, maximum number of iteration $M$.
Set $i = 1$.
While $i \le M$
    Set $y_1 = g(y_0)$; $y_2 = g(y_1)$; $y = y_0 - (y_1 - y_0)^2/(y_2 - 2y_1 + y_0)$.
    If $|y - y_0| < Tol$, then STOP.
    Set $i = i + 1$; $y_0 = y$.
End While
```

Theorem

Suppose that $x = g(x)$ has the solution x^* with $g'(x^*) \neq 1$. If $\exists\ \delta > 0$ such that $g \in C^3[x^* - \delta, x^* + \delta]$, then Steffensen's method gives quadratic convergence for any $x_0 \in [x^* - \delta, x^* + \delta]$.

Zeros of polynomials and Müller's method

- Horner's method:

Let

$$\begin{aligned} P(x) &= a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1} + a_n x^n \\ &= a_0 + x\left(a_1 + x\left(a_2 + \cdots + x\left(a_{n-1} + a_n x\right)\cdots\right)\right) \end{aligned}$$

If

$$\begin{aligned} b_n &= a_n, \\ b_k &= a_k + b_{k+1} x_0, \quad \text{for} \quad k = n-1, n-2, \ldots, 1, 0, \end{aligned}$$

then

$$b_0 = a_0 + b_1 x_0 = a_0 + \left(a_1 + b_2 x_0\right) x_0 = \cdots = P(x_0).$$

Consider

$$Q(x) = b_1 + b_2 x + \cdots + b_n x^{n-1}.$$

Then

$$\begin{aligned} & b_0 + (x - x_0)Q(x) = b_0 + (x - x_0)\left(b_1 + b_2 x + \cdots + b_n x^{n-1}\right) \\ = \; & (b_0 - b_1 x_0) + (b_1 - b_2 x_0)x + \cdots + (b_{n-1} - b_n x_0)x^{n-1} + b_n x^n \\ = \; & a_0 + a_1 x + \cdots + a_n x^n = P(x). \end{aligned}$$

Differentiating $P(x)$ with respect to x gives

$$P'(x) = Q(x) + (x - x_0)Q'(x) \quad \text{and} \quad P'(x_0) = Q(x_0).$$

Use Newton-Raphson method to find an approximate zero of $P(x)$:

$$x_{k+1} = x_k - \frac{P(x_k)}{Q(x_k)}, \; \forall \; k = 0, 1, 2, \ldots .$$

Similarly, let

$$\begin{aligned} c_n &= b_n = a_n, \\ c_k &= b_k + c_{k+1}x_k, \quad \text{for} \quad k = n-1, n-2, \ldots, 1, \end{aligned}$$

then $c_1 = Q(x_k)$.

Horner's method (Evaluate $y = P(x_0)$ and $z = P'(x_0)$)

```
Set $y = a_n$; $z = a_n$.
For $j = n-1, n-2, \ldots, 1$
    Set $y = a_j + yx_0$; $z = y + zx_0$.
End for
Set $y = a_0 + yx_0$.
```

If x_N is an approximate zero of P, then

$$\begin{aligned} P(x) &= (x - x_N)Q(x) + b_0 = (x - x_N)Q(x) + P(x_N) \\ &\approx (x - x_N)Q(x) \equiv (x - \hat{x}_1)Q_1(x). \end{aligned}$$

So $x - \hat{x}_1$ is an approximate factor of $P(x)$ and we can find a second approximate zero of P by applying Newton's method to $Q_1(x)$. The procedure is called deflation.

- Müller's method for complex root:

Theorem

If $z = a + ib$ is a complex zero of multiplicity m of $P(x)$ with real coefficients, then $\bar{z} = a - bi$ is also a zero of multiplicity m of $P(x)$ and $(x^2 - 2ax + a^2 + b^2)^m$ is a factor of $P(x)$.

Secant method: Given p_0 and p_1, determine p_2 as the intersection of the x-axis with the line through $(p_0, f(p_0))$ and $(p_1, f(p_1))$.

Müller's method: Given p_0, p_1 and p_2, determine p_3 by the intersection of the x-axis with the parabola through $(p_0, f(p_0))$, $(p_1, f(p_1))$ and $(p_2, f(p_2))$.

- Müller's method for complex root:

Theorem

If $z = a + ib$ is a complex zero of multiplicity m of $P(x)$ with real coefficients, then $\bar{z} = a - bi$ is also a zero of multiplicity m of $P(x)$ and $(x^2 - 2ax + a^2 + b^2)^m$ is a factor of $P(x)$.

Secant method: Given p_0 and p_1, determine p_2 as the intersection of the x-axis with the line through $(p_0, f(p_0))$ and $(p_1, f(p_1))$.

Müller's method: Given p_0, p_1 and p_2, determine p_3 by the intersection of the x-axis with the parabola through $(p_0, f(p_0))$, $(p_1, f(p_1))$ and $(p_2, f(p_2))$.

Let

$$P(x) = a(x - p_2)^2 + b(x - p_2) + c$$

that passes through $(p_0, f(p_0))$, $(p_1, f(p_1))$ and $(p_2, f(p_2))$. Then

$$\begin{aligned} f(p_0) &= a(p_0 - p_2)^2 + b(p_0 - p_2) + c, \\ f(p_1) &= a(p_1 - p_2)^2 + b(p_1 - p_2) + c, \\ f(p_2) &= a(p_2 - p_2)^2 + b(p_2 - p_2) + c = c. \end{aligned}$$

It implies that

$$\begin{aligned} c &= f(p_2), \\ b &= \frac{(p_0 - p_2)^2 [f(p_1) - f(p_2)] - (p_1 - p_2)^2 [f(p_0) - f(p_2)]}{(p_0 - p_2)(p_1 - p_2)(p_0 - p_1)}, \\ a &= \frac{(p_1 - p_2) [f(p_0) - f(p_2)] - (p_0 - p_2) [f(p_1) - f(p_2)]}{(p_0 - p_2)(p_1 - p_2)(p_0 - p_1)}. \end{aligned}$$

To determine p_3, a zero of P, we apply the quadratic formula to $P(x) = 0$ and get

$$p_3 - p_2 = \frac{2c}{b \pm \sqrt{b^2 - 4ac}}.$$

Choose

$$p_3 = p_2 + \frac{2c}{b + sgn(b)\sqrt{b^2 - 4ac}}$$

such that the denominator will be largest and result in p_3 selected as the closest zero of P to p_2.

Müller's method (Find a solution of $f(x) = 0$)

```
Given $p_0, p_1, p_2$; tolerance $TOL$; maximum number of iterations $M$
Set $h_1 = p_1 - p_0$; $h_2 = p_2 - p_1$;
    $\delta_1 = (f(p_1) - f(p_0))/h_1$; $\delta_2 = (f(p_2) - f(p_1))/h_2$;
    $d = (\delta_2 - \delta_1)/(h_2 + h_1)$; $i = 3$.
While $i \leq M$
    Set $b = \delta_2 + h_2 d$; $D = \sqrt{b^2 - 4f(p_2)d}$.
    If $|b - D| < |b + D|$, then set $E = b + D$ else set $E = b - D$.
    Set $h = -2f(p_2)/E$; $p = p_2 + h$.
    If $|h| < TOL$, then STOP.
    Set $p_0 = p_1$; $p_1 = p_2$; $p_2 = p$; $h_1 = p_1 - p_0$; $h_2 = p_2 - p_1$;
        $\delta_1 = (f(p_1) - f(p_0))/h_1$; $\delta_2 = (f(p_2) - f(p_1))/h_2$;
        $d = (\delta_2 - \delta_1)/(h_2 + h_1)$; $i = i + 1$.
End while
```