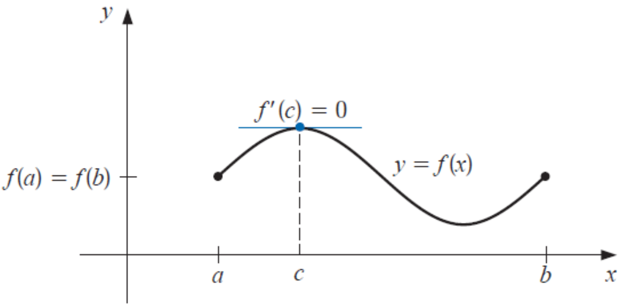
Numerical Analysis

Lecture2: Solutions of Equations in One Variable

Instructor: Prof. Qiwei Zhan
Zhejiang University

Rolle's Theorem

Suppose $f \in C[a, b]$ and f is differentiable on (a, b). If $f(a) = f(b)$, then a number c in (a, b) exists with $f'(c) = 0$.

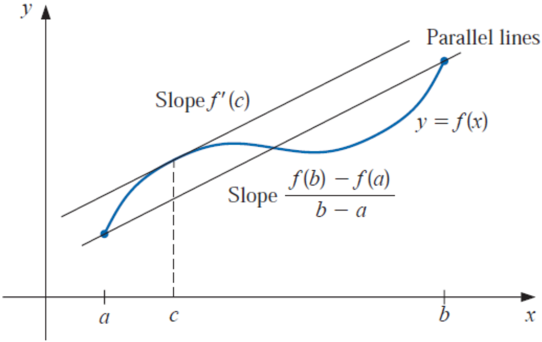


Outline

Mean Value Theorem

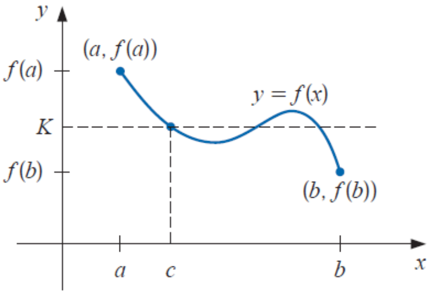
If $f \in C[a, b]$ and f is differentiable on (a, b), then a number c in (a, b) exists with

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$



Intermediate Value Theorem

If $f \in C[a,b]$ and K is any number between $f(a)$ and $f(b)$, then there exists a number c in (a,b) for which $f(c) = K$.



Outline

The Root-Finding Problem

The Root-Finding problem

- This process involves finding a root, or solution, of an equation of the form
$$f(x) = 0$$
for a given function f.
- A root of this equation is also called a zero of the function f.

Outline

The Bisection Method

Assumptions

- Suppose $f \in C[a,b]$ and $f(a) \cdot f(b) < 0$.
- By the IVT, there exists an x in (a,b) with $f(x) = 0$.
- We assume for simplicity that the root in this interval is unique.

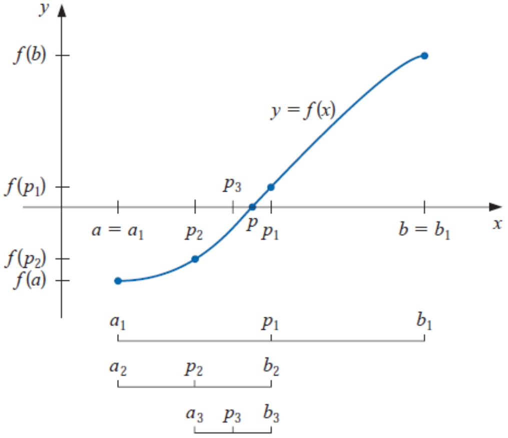
Solution - The Bisection Method

- Divide the interval $[a,b]$ by computing the midpoint

$$p = (a+b)/2$$

- If $f(p)$ has same sign as $f(a)$, consider new interval $[p,b]$.
- If $f(p)$ has same sign as $f(b)$, consider new interval $[a,p]$.
- Repeat until interval small enough to approximate x well.

The Bisection Method



Algorithm

INPUT endpoints a, b; tolerance TOL; maximum number of iterations N_0.

OUTPUT approximate solution p or message of failure.

Step 1 Set $i = 1$;
$FA = f(a)$.

Step 2 While $i \leq N_0$ do Steps 3–6.

Step 3 Set $p = a + (b - a)/2$; *(Compute p_i.)*
$FP = f(p)$.

Step 4 If $FP = 0$ or $(b - a)/2 < TOL$ then
OUTPUT (p); *(Procedure completed successfully.)*
STOP.

Step 5 Set $i = i + 1$.

Step 6 If $FA \cdot FP > 0$ then set $a = p$; *(Compute a_i, b_i.)*
$FA = FP$
else set $b = p$. *(FA is unchanged.)*

Step 7 OUTPUT ('Method failed after N_0 iterations, N_0 =', N_0);
(The procedure was unsuccessful.)
STOP.

Convergence

Theorem

Suppose that $f \in C[a,b]$ *and* $f(a) \cdot f(b) < 0$*. The Bisection method generates a sequence* $\{p_n\}_{n=1}^{\infty}$ *approximating a zero* p *of* f *with*

$$|p_n - p| \le \frac{b-a}{2^n}, \quad \text{when } n \ge 1.$$

Proof

Convergence

Theorem

Suppose that $f \in C[a, b]$ and $f(a) \cdot f(b) < 0$. The Bisection method generates a sequence $\{p_n\}_{n=1}^{\infty}$ approximating a zero p of f with

$$|p_n - p| \leq \frac{b-a}{2^n}, \quad \text{when } n \geq 1.$$

Proof

- For each $n \geq 1$, we have

$$b_n - a_n = \frac{1}{2^{n-1}}(b-a) \quad \text{and} \quad p \in (a_n, b_n).$$

Convergence

Theorem

Suppose that $f \in C[a,b]$ and $f(a) \cdot f(b) < 0$. The Bisection method generates a sequence $\{p_n\}_{n=1}^{\infty}$ approximating a zero p of f with

$$|p_n - p| \leq \frac{b-a}{2^n}, \quad \text{when } n \geq 1.$$

Proof

- For each $n \geq 1$, we have
$$b_n - a_n = \frac{1}{2^{n-1}}(b-a) \quad \text{and} \quad p \in (a_n, b_n).$$
- Since $p_n = \frac{1}{2}(a_n + b_n)$ for all $n \geq 1$, it follows that
$$|p_n - p| \leq \frac{1}{2}(b_n - a_n) = \frac{b-a}{2^n}.$$

Rate of Convergence

Theorem

Suppose that $f \in C[a, b]$ and $f(a) \cdot f(b) < 0$. The Bisection method generates a sequence $\{p_n\}_{n=1}^{\infty}$ approximating a zero p of f with

$$|p_n - p| \leq \frac{b-a}{2^n}, \quad \text{when } n \geq 1.$$

Theorem

The sequence $\{p_n\}_{n=1}^{\infty}$ converges to p with rate of convergence $O(1/2^n)$:

$$p_n = p + O\left(\frac{1}{2^n}\right).$$

Remarks

Remarks

- The Bisection Method has a number of significant drawbacks.
  1. It is very slow to converge in that N may be quite large before $|p - p_N|$ becomes sufficiently small.
  2. It is possible that a good intermediate approximation may be inadvertently discarded.
- It will always converge to a solution however and, for this reason, is often used to provide a good initial approximation for a more efficient procedure.

Example

Example 1

Show that $f(x) = x^3 + 4x^2 - 10 = 0$ has a root in [1, 2], and use the Bisection method to determine an approximation to the root that is accurate to at least within 10^{-4}.

n	a_n	b_n	p_n	$f(p_n)$
1	1.0	2.0	1.5	2.375
2	1.0	1.5	1.25	−1.79687
3	1.25	1.5	1.375	0.16211
4	1.25	1.375	1.3125	−0.84839
5	1.3125	1.375	1.34375	−0.35098
6	1.34375	1.375	1.359375	−0.09641
7	1.359375	1.375	1.3671875	0.03236
8	1.359375	1.3671875	1.36328125	−0.03215
9	1.36328125	1.3671875	1.365234375	0.000072
10	1.36328125	1.365234375	1.364257813	−0.01605
11	1.364257813	1.365234375	1.364746094	−0.00799
12	1.364746094	1.365234375	1.364990235	−0.00396
13	1.364990235	1.365234375	1.365112305	−0.00194

Outline

The Fixed-Point Problem

The Root-Finding Problem

Given a function $f(x)$ where $a \leq x \leq b$, find values p such that

$$f(p) = 0$$

The Fixed-Point Problem

Given such a function, $f(x)$, we now construct an auxiliary function $g(x)$ such that

$$p = g(p)$$

whenever $f(p) = 0$.

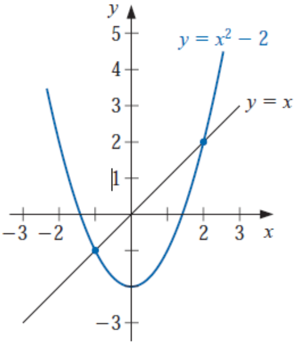
- This construction is not unique.
- The problem of finding p such that $p = g(p)$ is known as the fixed point problem.

The Fixed-Point Problem

A Fixed Point

If g is defined on $[a, b]$ and $g(p) = p$ for some $p \in [a, b]$, then the function g is said to have the fixed point p in $[a, b]$.

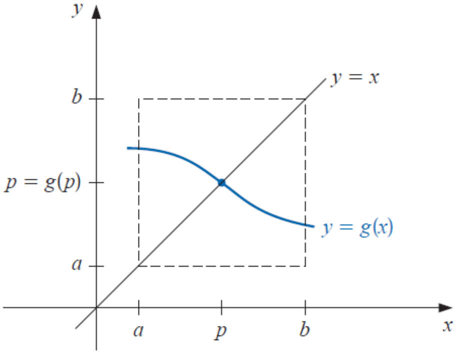
Ex: Determine any fixed points of the function $g(x) = x^2 - 2$.



Existence of Fixed Points

Theorem (Existence of Fixed Points)

If $g \in C[a, b]$ and $g(x) \in [a, b]$ for all $x \in [a, b]$ then the function g has a fixed point in $[a, b]$.



Existence of Fixed Points

Theorem (Existence of Fixed Points)

If $g \in C[a, b]$ and $g(x) \in [a, b]$ for all $x \in [a, b]$ then the function g has a fixed point in $[a, b]$.

Proof

Existence of Fixed Points

Theorem (Existence of Fixed Points)

If $g \in C[a, b]$ and $g(x) \in [a, b]$ for all $x \in [a, b]$ then the function g has a fixed point in $[a, b]$.

Proof

- If $g(a) = a$ or $g(b) = b$, the existence of a fixed point is obvious.

Existence of Fixed Points

Theorem (Existence of Fixed Points)

If $g \in C[a, b]$ and $g(x) \in [a, b]$ for all $x \in [a, b]$ then the function g has a fixed point in $[a, b]$.

Proof

- If $g(a) = a$ or $g(b) = b$, the existence of a fixed point is obvious.
- If not, then it must be true that $g(a) > a$ and $g(b) < b$.

Existence of Fixed Points

Theorem (Existence of Fixed Points)

If $g \in C[a, b]$ and $g(x) \in [a, b]$ for all $x \in [a, b]$ then the function g has a fixed point in $[a, b]$.

Proof

- If $g(a) = a$ or $g(b) = b$, the existence of a fixed point is obvious.
- If not, then it must be true that $g(a) > a$ and $g(b) < b$.
- Define $h(x) = g(x) - x$; h is continuous on $[a, b]$ and, moreover,

$$h(a) = g(a) - a > 0, \qquad h(b) = g(b) - b < 0.$$

Existence of Fixed Points

Theorem (Existence of Fixed Points)

If $g \in C[a,b]$ and $g(x) \in [a,b]$ for all $x \in [a,b]$ then the function g has a fixed point in $[a,b]$.

Proof

- If $g(a) = a$ or $g(b) = b$, the existence of a fixed point is obvious.
- If not, then it must be true that $g(a) > a$ and $g(b) < b$.
- Define $h(x) = g(x) - x$; h is continuous on $[a,b]$ and, moreover,

$$h(a) = g(a) - a > 0, \qquad h(b) = g(b) - b < 0.$$

- The IVT implies that there exists $p \in (a,b)$ for which $h(p) = 0$.

Uniqueness of Fixed Points

Theorem (Uniqueness of Fixed Points)

Let $g \in C[a,b]$ *and* $g(x) \in [a,b]$ *for all* $x \in [a,b]$. *Further if* $g'(x)$ *exists on* $[a,b]$ *and*

$$|g'(x)| \leq k < 1, \forall x \in [a,b],$$

then the function g *has a unique fixed point* p *in* $[a,b]$.

Proof

Uniqueness of Fixed Points

Theorem (Uniqueness of Fixed Points)

Let $g \in C[a,b]$ and $g(x) \in [a,b]$ for all $x \in [a,b]$. Further if $g'(x)$ exists on $[a,b]$ and

$$|g'(x)| \leq k < 1, \forall x \in [a,b],$$

then the function g has a unique fixed point p in $[a,b]$.

Proof

- Suppose that p and q are both fixed point in $[a,b]$ with $p \neq q$.

Uniqueness of Fixed Points

Theorem (Uniqueness of Fixed Points)

Let $g \in C[a,b]$ and $g(x) \in [a,b]$ for all $x \in [a,b]$. Further if $g'(x)$ exists on $[a,b]$ and

$$|g'(x)| \leq k < 1, \forall x \in [a,b],$$

then the function g has a unique fixed point p in $[a,b]$.

Proof

- Suppose that p and q are both fixed point in $[a,b]$ with $p \neq q$.
- By the MVT, a number ξ exists between p and q in $[a,b]$ with

$$\begin{aligned} |p-q| = |g(p)-g(q)| &= |g'(\xi)||p-q| \\ &\leq k|p-q| < |p-q| \end{aligned}$$

  which is a contradiction

Uniqueness of Fixed Points

Theorem (Uniqueness of Fixed Points)

Let $g \in C[a,b]$ and $g(x) \in [a,b]$ for all $x \in [a,b]$. Further if $g'(x)$ exists on $[a,b]$ and

$$|g'(x)| \leq k < 1, \forall x \in [a,b],$$

then the function g has a unique fixed point p in $[a,b]$.

Proof

- Suppose that p and q are both fixed point in $[a,b]$ with $p \neq q$.
- By the MVT, a number ξ exists between p and q in $[a,b]$ with

$$\begin{aligned} |p-q| = |g(p) - g(q)| &= |g'(\xi)||p-q| \\ &\leq k|p-q| < |p-q| \end{aligned}$$

  which is a contradiction
- Hence, $p = q$ and the fixed point in $[a,b]$ is unique.

Fixed-Point Iteration

A Method to Solve the Fixed-Point Problem

- Choose an initial approximation p_0 and generate the sequence $\{p_n\}_{n=0}^{\infty}$ by letting $p_n = g(p_{n-1})$, for each $n \geq 1$.

Fixed-Point Iteration

A Method to Solve the Fixed-Point Problem

- Choose an initial approximation p_0 and generate the sequence $\{p_n\}_{n=0}^{\infty}$ by letting $p_n = g(p_{n-1})$, for each $n \geq 1$.
- If the sequence converges to p and g is continuous, then

$$p = \lim_{n\to\infty} p_n = \lim_{n\to\infty} g(p_{n-1}) = g\left(\lim_{n\to\infty} p_{n-1}\right) = g(p),$$

and a solution to $x = g(x)$ is obtained.

Fixed-Point Iteration

A Method to Solve the Fixed-Point Problem

- Choose an initial approximation p_0 and generate the sequence $\{p_n\}_{n=0}^{\infty}$ by letting $p_n = g(p_{n-1})$, for each $n \geq 1$.
- If the sequence converges to p and g is continuous, then

$$p = \lim_{n\to\infty} p_n = \lim_{n\to\infty} g(p_{n-1}) = g\left(\lim_{n\to\infty} p_{n-1}\right) = g(p),$$

  and a solution to $x = g(x)$ is obtained.
- This technique is called fixed-point iteration.

Fixed-Point Iteration

Fixed-Point Algorithm

INPUT initial approximation p_0; tolerance TOL; maximum number of iterations N_0.

OUTPUT approximate solution p or message of failure.

Step 1 Set $i = 1$.

Step 2 While $i \leq N_0$ do Steps 3–6.

Step 3 Set $p = g(p_0)$. (*Compute* p_i.)

Step 4 If $|p - p_0| < TOL$ then
OUTPUT (p); (*The procedure was successful.*)
STOP.

Step 5 Set $i = i + 1$.

Step 6 Set $p_0 = p$. (*Update* p_0.)

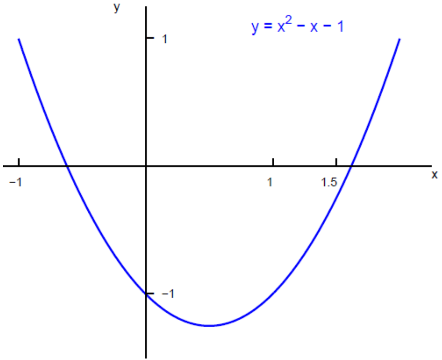
Step 7 OUTPUT ('The method failed after N_0 iterations, $N_0 =$', N_0);
(*The procedure was unsuccessful.*)
STOP.

Examples

Example 1

Find the positive root for the quadratic equation:

$$x^2 - x - 1 = 0$$



Example 1

Solution 1

Convert the quadratic equation $f(x) = x^2 - x - 1 = 0$ to a fixed-point problem.

- Transpose the equation $f(x) = 0$ for variable x:

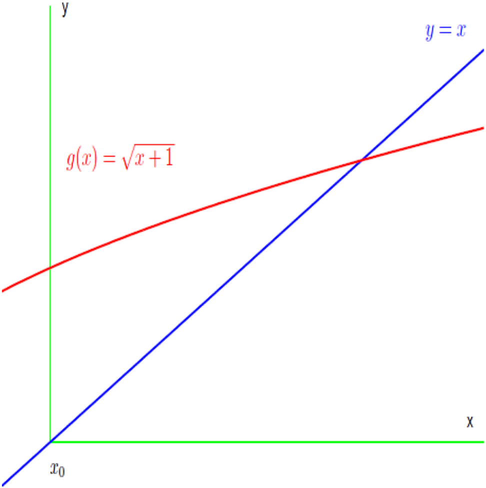
$$\begin{aligned} x^2 - x - 1 &= 0 \\ \Rightarrow \qquad x^2 &= x + 1 \\ \Rightarrow \qquad x &= \pm\sqrt{x+1} \end{aligned}$$

- The constructed fixed-point problem:

$$g(x) = \sqrt{x+1}$$

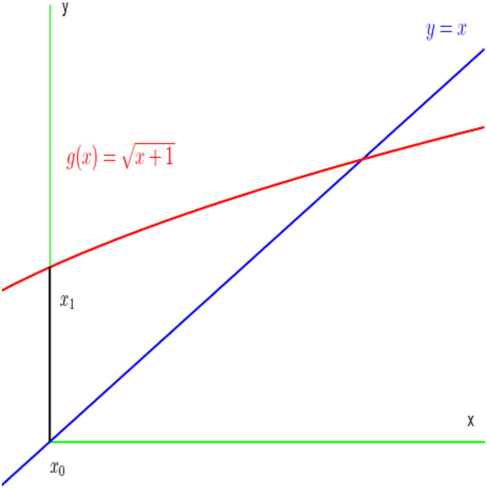
Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



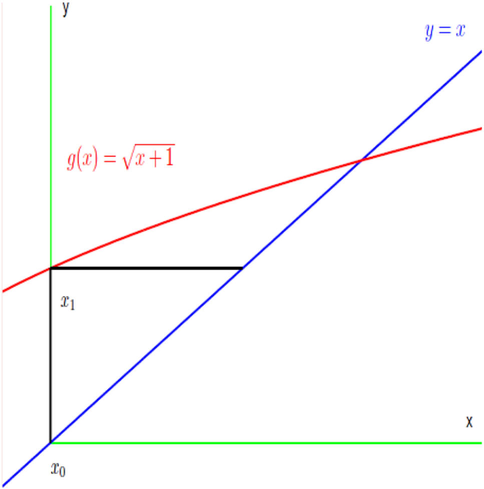
Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



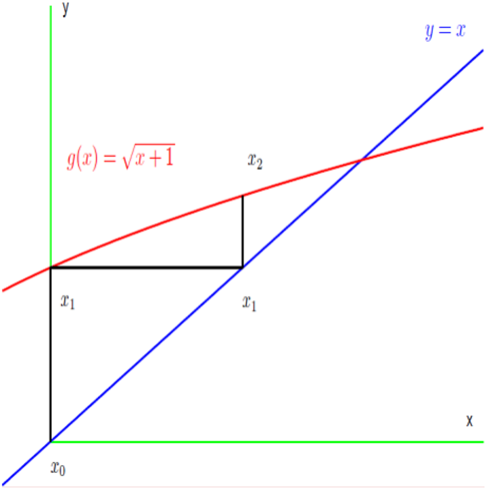
Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



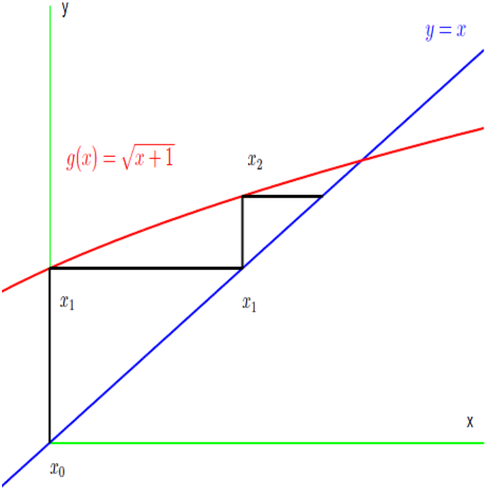
Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



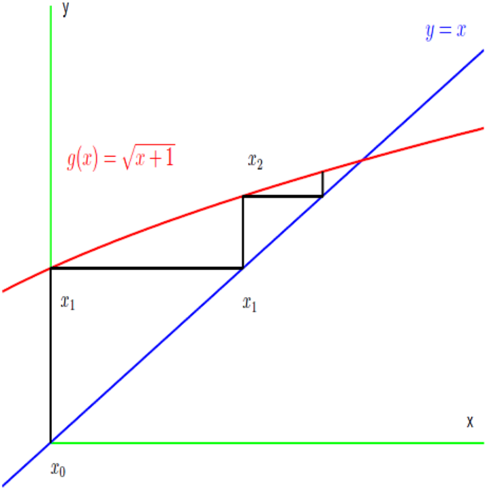
Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



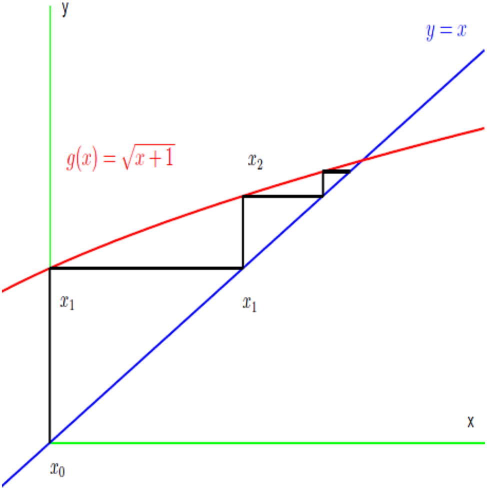
Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



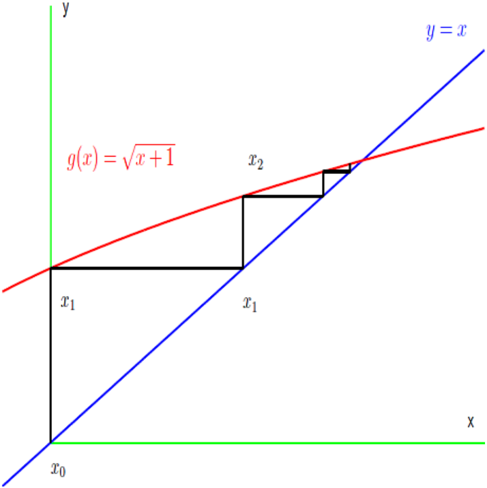
Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



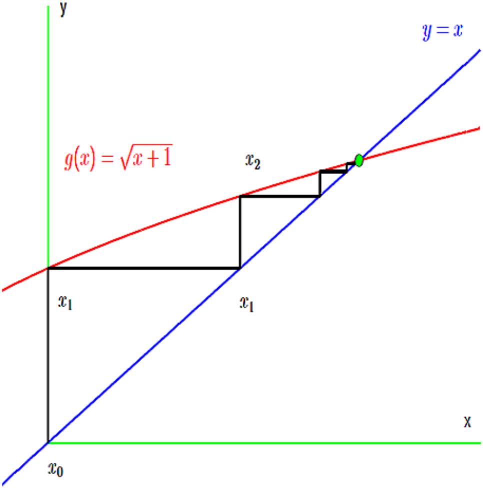
Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



Example 1

Solution 1: $x_{n+1} = g(x_n) = \sqrt{x_n + 1}$ with $x_0 = 0$



Example 1

Solution 2

Convert the quadratic equation $f(x) = x^2 - x - 1 = 0$ to a fixed-point problem.

- Transpose the equation $f(x) = 0$ for variable x:

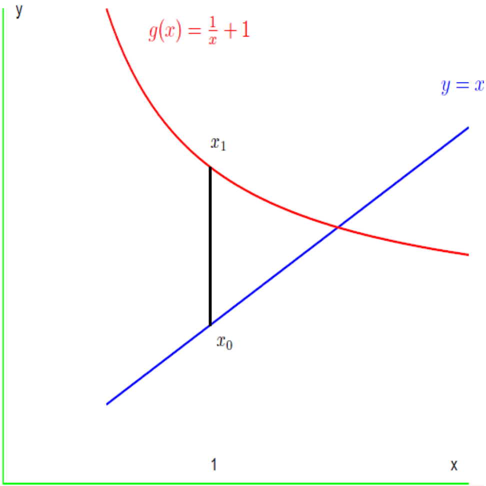
$$\begin{aligned} x^2 - x - 1 &= 0 \\ \Rightarrow \qquad x^2 &= x + 1 \\ \Rightarrow \qquad x &= 1 + \frac{1}{x} \end{aligned}$$

- The constructed fixed-point problem:

$$g(x) = 1 + \frac{1}{x}$$

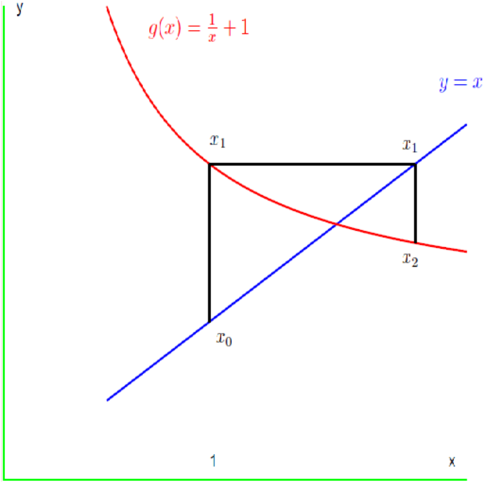
Example 1

Solution 2: $x_{n+1} = g(x_n) = 1 + \frac{1}{x_n}$ with $x_0 = 1$



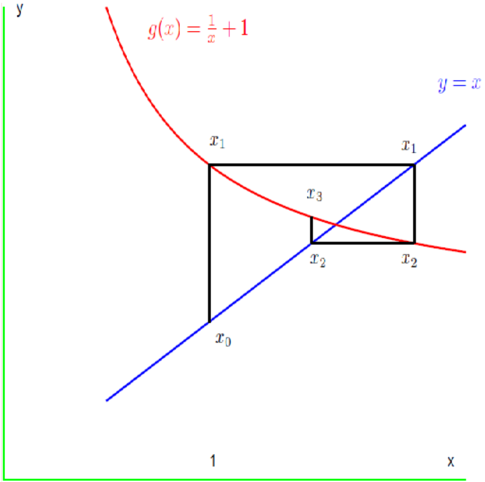
Example 1

Solution 2: $x_{n+1} = g(x_n) = 1 + \frac{1}{x_n}$ with $x_0 = 1$



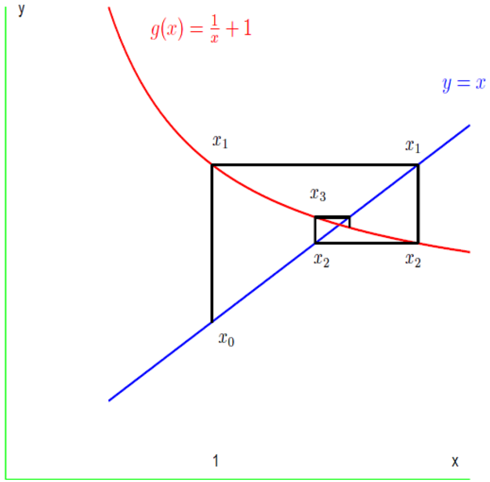
Example 1

Solution 2: $x_{n+1} = g(x_n) = 1 + \frac{1}{x_n}$ with $x_0 = 1$



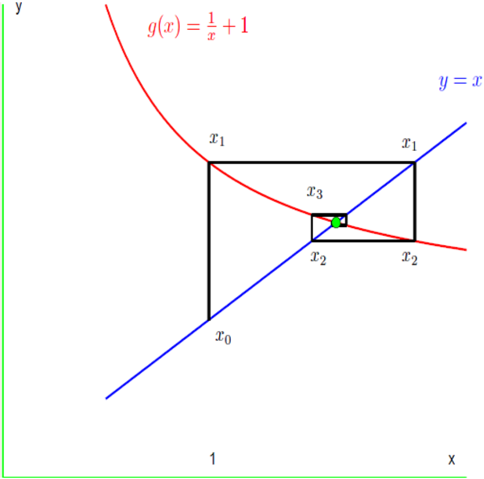
Example 1

Solution 2: $x_{n+1} = g(x_n) = 1 + \frac{1}{x_n}$ with $x_0 = 1$



Example 1

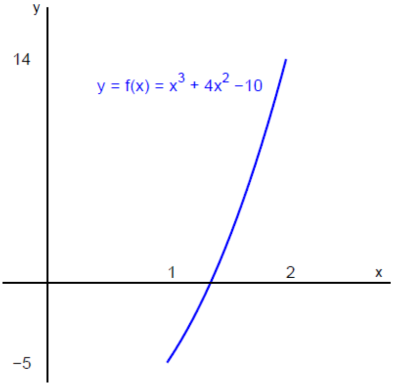
Solution 2: $x_{n+1} = g(x_n) = 1 + \frac{1}{x_n}$ with $x_0 = 1$



Example 2

Find the root for the equation:

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



Example 2

Solutions: $x = g(x)$ with $x_0 = 1.5$

$x = g_1(x) = x - x^3 - 4x^2 + 10$	Does not Converge
$x = g_2(x) = \sqrt{\frac{10}{x} - 4x}$	Does not Converge
$x = g_3(x) = \frac{1}{2}\sqrt{10 - x^3}$	Converges after 31 Iterations
$x = g_4(x) = \sqrt{\frac{10}{4 + x}}$	Converges after 12 Iterations
$x = g_5(x) = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}$	Converges after 5 Iterations

Fixed-Point Theorem

Theorem (Fixed-Point Theorem)

Let $g \in C[a,b]$ *with* $g(x) \in [a,b]$ *for all* $x \in [a,b]$. *Let* $g'(x)$ *exist on* (a,b) *with*

$$|g'(x)| \leq k < 1, \forall x \in [a,b].$$

Then for any point p_0 *in* $[a,b]$, *the sequence defined by*

$$p_n = g(p_{n-1}), n \geq 1$$

will converge to the unique fixed point p *in* $[a,b]$.

Proof

- By the Uniqueness Theorem, a unique fixed point exists in $[a,b]$.

Fixed-Point Theorem

Theorem (Fixed-Point Theorem)

Let $g \in C[a,b]$ *with* $g(x) \in [a,b]$ *for all* $x \in [a,b]$. *Let* $g'(x)$ *exist on* (a,b) *with*

$$|g'(x)| \leq k < 1, \forall x \in [a,b].$$

Then for any point p_0 *in* $[a,b]$, *the sequence defined by*

$$p_n = g(p_{n-1}), n \geq 1$$

will converge to the unique fixed point p *in* $[a,b]$.

Proof

- By the Uniqueness Theorem, a unique fixed point exists in $[a,b]$.
- Since g maps $[a,b]$ into itself, the sequence $\{p_n\}_{n=0}^{\infty}$ is defined for all $n \geq 0$ and $p_n \in [a,b]$ for all n.

Fixed-Point Theorem

Proof

- Using the MVT and the assumption that $|g'(x)| \leq k < 1, \forall x \in [a, b]$, we have

$$
\begin{aligned}
|p_n - p| = |g(p_{n-1}) - g(p)| &= |g'(\xi)||p_{n-1} - p| \\
&\leq k|p_{n-1} - p| \\
&\leq k^2|p_{n-2} - p| \\
&\leq k^n|p_0 - p|
\end{aligned}
$$

where $\xi \in (a, b)$.

Fixed-Point Theorem

Proof

- Using the MVT and the assumption that $|g'(x)| \leq k < 1, \forall x \in [a, b]$, we have

$$\begin{aligned} |p_n - p| = |g(p_{n-1}) - g(p)| &= |g'(\xi)||p_{n-1} - p| \\ &\leq k|p_{n-1} - p| \\ &\leq k^2|p_{n-2} - p| \\ &\leq k^n|p_0 - p| \end{aligned}$$

  where $\xi \in (a, b)$.
- Since $k < 1$,

$$\lim_{n\to\infty} |p_n - p| \leq \lim_{n\to\infty} k^n|p_0 - p| = 0,$$

  and $\{p_n\}_{n=0}^{\infty}$ converges to p.

Fixed-Point Theorem

Corollary (Corrollary to the Fixed-Point Theorem)

If g satisfies the hypothesis of the Theorem, then

$$|p_n - p| \le \frac{k^n}{1-k}|p_1 - p_0|.$$

Proof

- $|p_{n+1} - p_n| = |g(p_n) - g(p_{n-1})| \le k^n |p_1 - p_0|$

Fixed-Point Theorem

Corollary (Corrollary to the Fixed-Point Theorem)

If g satisfies the hypothesis of the Theorem, then

$$|p_n - p| \leq \frac{k^n}{1-k}|p_1 - p_0|.$$

Proof

- $|p_{n+1} - p_n| = |g(p_n) - g(p_{n-1})| \leq k^n |p_1 - p_0|$
- Thus, for $m > n \geq 1$

$$\begin{aligned}
|p_m - p_n| &= |p_m - p_{m-1} + p_{m-1} - p_{m-2} + \cdots + p_{n+1} - p_n| \\
&\leq |p_m - p_{m-1}| + |p_{m-1} - p_{m-2}| + \cdots + |p_{n+1} - p_n| \\
&\leq k^{m-1}|p_1 - p_0| + k^{m-2}|p_1 - p_0| + \cdots + k^n |p_1 - p_0| \\
&\leq k^n(1 + k + k^2 + \cdots + k^{m-n-1})|p_1 - p_0|.
\end{aligned}$$

Fixed-Point Theorem

Corollary (Corrollary to the Fixed-Point Theorem)

If g satisfies the hypothesis of the Theorem, then

$$|p_n - p| \leq \frac{k^n}{1-k}|p_1 - p_0|.$$

Proof

- $|p_{n+1} - p_n| = |g(p_n) - g(p_{n-1})| \leq k^n |p_1 - p_0|$
- Thus, for $m > n \geq 1$

$$\begin{aligned}
|p_m - p_n| &= |p_m - p_{m-1} + p_{m-1} - p_{m-2} + \cdots + p_{n+1} - p_n| \\
&\leq |p_m - p_{m-1}| + |p_{m-1} - p_{m-2}| + \cdots + |p_{n+1} - p_n| \\
&\leq k^{m-1}|p_1 - p_0| + k^{m-2}|p_1 - p_0| + \cdots + k^n|p_1 - p_0| \\
&\leq k^n(1 + k + k^2 + \cdots + k^{m-n-1})|p_1 - p_0|.
\end{aligned}$$

- Since $\lim_{m\to\infty} p_m = p$, we obtain

$$|p - p_n| = \lim_{m\to\infty} |p_m - p_n| \leq k^n(1+k+k^2+\cdots+k^{m-n-1})|p_1 - p_0| = \frac{k^n}{1-k}|p_1 - p_0|.$$

Example 2

Solutions: $x = g(x)$ with $x_0 = 1.5$

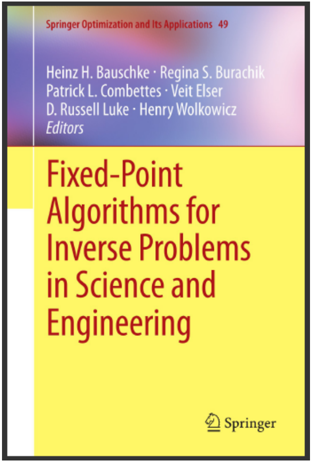
$x = g_1(x) = x - x^3 - 4x^2 + 10$	Does not Converge
$x = g_2(x) = \sqrt{\frac{10}{x} - 4x}$	Does not Converge
$x = g_3(x) = \frac{1}{2}\sqrt{10 - x^3}$	Converges after 31 Iterations
$x = g_4(x) = \sqrt{\frac{10}{4 + x}}$	Converges after 12 Iterations
$x = g_5(x) = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}$	Converges after 5 Iterations

Question

How to construct a fixed-point problem?

Question

How to construct a fixed-point problem?

- 

Outline

Newton's Method

Newton's Method

Newton's (or the Newton-Raphson) method is one of the most powerful and well-known numerical methods for solving a root-finding problem.

Remarks

- Newton's method obtains faster convergence than offered by other types of functional iteration.
- Using Taylor polynomials. We will see there that this particular derivation produces not only the method, but also a bound for the error of the approximation.

Derivation

- Suppose that $f \in C^2[a, b]$. Let $p_0 \in [a, b]$ be an approximation to p such that $f'(p_0) \neq 0$ and $|p - p_0|$ is small.
- Consider the first Taylor polynomial for $f(x)$ expanded about p_0 and evaluated at $x = p$.

$$f(p) = f(p_0) + (p - p_0)f'(p_0) + \frac{(p - p_0)^2}{2} f''(\xi(p)),$$

  where $\xi(p)$ lies between p and p_0.
- Since $f(p) = 0$, this equation gives

$$0 = f(p_0) + (p - p_0)f'(p_0) + \frac{(p - p_0)^2}{2} f''(\xi(p)),$$

Newton's Method

$$0 = f(p_0) + (p - p_0)f'(p_0) + \frac{(p - p_0)^2}{2} f''(\xi(p)),$$

Derivation (cont'd)

- Newton's method is derived by assuming that since $|p - p_0|$ is small, the term involving $(p - p_0)^2$ is much smaller, so

$$0 \approx f(p_0) + (p - p_0)f'(p_0).$$

- Solving for p gives

$$p \approx p_0 - \frac{f(p_0)}{f'(p_0)} \equiv p_1.$$

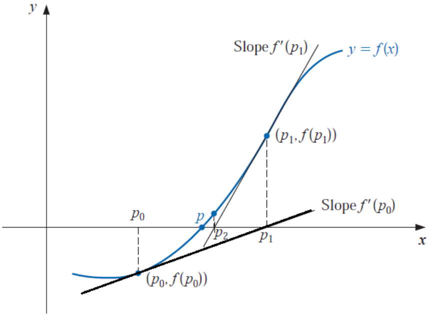
Newton's Method

Newton's Method

Start with an initial approximation p_0 and generate the sequence $\{p_n\}_{n=0}^{\infty}$, by

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \qquad \text{for } \; n \geq 1$$

that is, $0 = (p_n - p_{n-1})f'(p_{n-1}) + f(p_{n-1}) \qquad \text{for } \; n \geq 1$

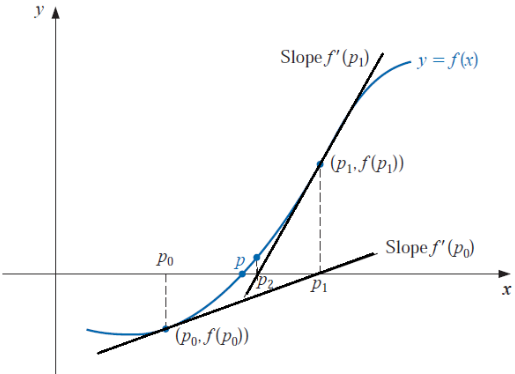


Newton's Method

Newton's Method

Starts with an initial approximation p_0 and generates the sequence $\{p_n\}_{n=0}^{\infty}$, by

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \qquad \text{for } \; n \geq 1$$



Newton's Algorithm

INPUT initial approximation p_0; tolerance TOL; maximum number of iterations N_0.

OUTPUT approximate solution p or message of failure.

Step 1 Set $i = 1$.

Step 2 While $i \leq N_0$ do Steps 3–6.

Step 3 Set $p = p_0 - f(p_0)/f'(p_0)$. (*Compute p_i.*)

Step 4 If $|p - p_0| < TOL$ then
OUTPUT (p); (*The procedure was successful.*)
STOP.

Step 5 Set $i = i + 1$.

Step 6 Set $p_0 = p$. (*Update p_0.*)

Step 7 OUTPUT ('The method failed after N_0 iterations, $N_0 =$', N_0);
(*The procedure was unsuccessful.*)
STOP.

Newton's Algorithm

Stopping Criteria for the Algorithm

- Various stopping criteria can be applied.
- We can select a tolerance $\epsilon > 0$ and generate $p_1, \cdots, p_N$ until one of the following conditions is met:

$$|p_N - p_{N-1}| < \epsilon \tag{1}$$

$$\frac{|p_N - p_{N-1}|}{p_N} < \epsilon, \quad p_N \neq 0, \quad \text{or} \tag{2}$$

$$|f(p_N)| < \epsilon \tag{3}$$

- Note that none of these inequalities give precise information about the actual error $|p_N - p|$.

Newton's Method vs. Fixed-point Iteration

Fixed-Point Iteration (a.k.a Functional Iteration)

- Newton's Method

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \qquad \text{for } \; n \geq 1$$

Newton's Method vs. Fixed-point Iteration

Fixed-Point Iteration (a.k.a Functional Iteration)

- Newton's Method

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \quad \text{for } n \geq 1$$

- It can be written in the form

$$p_n = g(p_{n-1})$$

with

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

Newton’s Method vs. Fixed-point Iteration

Example: Newton’s Method vs. Fixed-point Iteration

Consider the function

$$f(x) = cos(x) - x = 0$$

Approximate a root of f using

1. a fixed-point method
2. Newton’s method

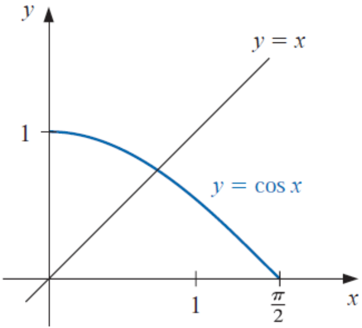
Newton's Method vs. Fixed-point Iteration

(1) Fixed-point Iteration for $f(x) = cos(x) - x$

- A solution to this root-finding problem is also a solution to the fixed-point problem

$$x = cos(x)$$

  and the graph implies that a single fixed-point p lies in $[0, \pi/2]$.



Newton's Method vs. Fixed-point Iteration

(1) Fixed-point Iteration: $x = cos(x)$, $x_0 = \pi/4$

- The following table shows the results of fixed-point iteration with $p_0 = \pi/4$.

n	p_{n-1}	p_n	$\lvert p_n - p_{n-1}\rvert$	e_n/e_{n-1}
1	0.7853982	0.7071068	0.0782914	—
2	0.707107	0.760245	0.053138	0.678719
3	0.760245	0.724667	0.035577	0.669525
4	0.724667	0.748720	0.024052	0.676064
5	0.748720	0.732561	0.016159	0.671826
6	0.732561	0.743464	0.010903	0.674753
7	0.743464	0.736128	0.007336	0.672816

- The best conclusion from these results is that $p \approx 0.74$.

Newton's Method vs. Fixed-point Iteration

(2) Newton's Method for $f(x) = cos(x) - x$

- To apply Newton's method to this problem we need

$$f'(x) = -sin(x) - 1$$

- Starting with $p_0 = \pi/4$, we generate the sequence defined for $n \geq 1$, by

$$\begin{aligned} p_n &= p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \\ &= p_{n-1} - \frac{cos(p_{n-1}) - p_{n-1}}{-sin(p_{n-1}) - 1}. \end{aligned}$$

Newton's Method vs. Fixed-point Iteration

(2) Newton's Method for $f(x) = cos(x) - x$, $x_0 = \pi/4$

- The following table shows the results of Newton's method with $p_0 = \pi/4$.

n	p_{n-1}	$f(p_{n-1})$	$f'(p_{n-1})$	p_n	$\lvert p_n - p_{n-1}\rvert$
1	0.78539816	-0.078291	-1.707107	0.73953613	0.04586203
2	0.73953613	-0.000755	-1.673945	0.73908518	0.00045096
3	0.73908518	-0.000000	-1.673612	0.73908513	0.00000004
4	0.73908513	-0.000000	-1.673612	0.73908513	0.00000000

- An excellent approximation is obtained with $n = 3$!
- Because of the agreement of p_3 and p_4 we could reasonably expect this result to be accurate to the places listed.

Convergence using Newton's Method

Convergence Theorem for Newton's Method

Let $f \in C^2[a,b]$. If $p \in (a,b)$ is such that $f(p) = 0$ and $f'(p) \neq 0$. Then there exists a $\delta > 0$ such that Newton's method generates a sequence $\{p_n\}_{n=1}^{\infty}$, defined by

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}$$

converging to p for any initial approximation

$$p_0 \in [p - \delta, p + \delta].$$

Convergence using Newton's Method

Proof: (1/4)

- The proof is based on analyzing Newton's method as the functional iteration scheme $p_n = g(p_{n-1})$, for $n \geq 1$ with

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

- Let $k \in (0, 1)$. We find an interval $[p - \delta, p + \delta]$ that g maps into itself and for which $g'(x) \leq k$, for all $x \in (p - \delta, p + \delta)$.
- Since f' is continuous and $f'(p) \neq 0$, there exists a $\delta_1 > 0$, such that $f'(x) \neq 0$ for $x \in [p - \delta_1, p + \delta_1] \subseteq [a, b]$.

Convergence using Newton's Method

Proof: (2/4)

- Thus g is defined and continuous on $[p-\delta_1, p+\delta_1]$. 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 $x \in [p-\delta_1, p+\delta_1]$, and since $f \in C^2[a,b]$, we have $g \in C^1[p-\delta_1, p+\delta_1]$.

- By assumption, $f(p) = 0$, so

$$g'(p) = \frac{f(p)f''(p)}{[f'(p)]^2} = 0.$$

Convergence using Newton's Method

$$g'(p) = \frac{f(p)f''(p)}{[f'(p)]^2} = 0.$$

Proof: (3/4)

- Since g' is continuous and $0 < k < 1$, there exists a δ, with $0 < \delta < \delta_1$, and

$$|g'(x)| \leq k, \quad \text{for all} \quad x \in [p-\delta, p+\delta].$$

- It remains to show that g maps $[p-\delta, p+\delta]$ into $[p-\delta, p+\delta]$.
- If $x \in [p-\delta, p+\delta]$, the MVT implies that for some number ξ between x and p, $|g(x) - g(p)| = |g'(\xi)||x-p|$. So

$$|g(x) - p| = |g(x) - g(p)| = |g'(\xi)||x-p| \leq k|x-p| < |x-p|.$$

Convergence using Newton's Method

Proof: (4/4)

- Since $x \in [p - \delta, p + \delta]$, it follows that $|x - p| < \delta$ and that $|g(x) - p| < \delta$. Hence, g maps $[p - \delta, p + \delta]$ into $[p - \delta, p + \delta]$.
- All the hypotheses of the Fixed-Point Theorem are now satisified, so the sequence $\{p_n\}_{n=1}^{\infty}$, defined by

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}$$

converges to p for any $p_0 \in [p - \delta, p + \delta]$.

Remarks of Newton's Method

Choice of Initial Approximation

- The convergence theorem states that, under reasonable assumptions, Newton's method converges if a sufficiently accurate initial approximation is chosen.
- It also implies that the constant k that bounds the derivative of g, and consequently, indicates the speed of convergence of the method, decreases to 0 as the procedure continues.
- This result is important for the theory of Newton's method, but it is seldom applied in practice because it does not tell us how to determine δ.

Remarks of Newton's Method

In a practical application ...

- an initial approximation is selected
- and successive approximations are generated by Newton's method.
- These will generally either converge quickly to the root, or it will be clear that convergence is unlikely.

Remarks of Newton's Method

Weakness of Newton's Method

- It needs to know the value of the derivate of f at each approximation.
- Frequently, $f'(x)$ is far more difficult and needs more arithmetic operations to calculate than $f(x)$.
- Adjoint method is suggested for the gradient...
- Try the software tool Mathematica...

Outline

Error Analysis for Iterative Methods

Order of Convergence

Suppose $\{p_n\}_{n=0}^{\infty}$ is a sequence that converges to p, with $p_n \neq p$ for all n. If positive constants λ and α exist with

$$\lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|^{\alpha}} = \lambda$$

then $\{p_n\}_{n=0}^{\infty}$ converges to p of order α, with asymptotic error constant λ.

- If $\alpha = 1$, the sequence is linearly convergent.
- If $\alpha = 2$, the sequence is quadratically convergent.

Error Analysis for Iterative Methods

linearly convergent vs. quadratically convergent

Suppose that $\{p_n\}_{n=0}^{\infty}$ is linearly convergent to 0 with

$$\lim_{n\to\infty} \frac{|p_{n+1}|}{|p_n|} = 0.5$$

and that $\{\tilde{p}_n\}_{n=0}^{\infty}$ is quadratically convergent to 0 with the same asymptotic error constant,

$$\lim_{n\to\infty} \frac{|\tilde{p}_{n+1}|}{|\tilde{p}_n|^2} = 0.5.$$

For simplicity we assume that for each n we have

$$\frac{|p_{n+1}|}{|p_n|} \approx 0.5 \quad \text{and} \quad \frac{|\tilde{p}_{n+1}|}{|\tilde{p}_n|^2} \approx 0.5.$$

For the linearly convergent scheme, this means that

$$|p_n - 0| = |p_n| \approx 0.5|p_{n-1}| \approx (0.5)^2|p_{n-2}| \approx \cdots \approx (0.5)^n|p_0|,$$

whereas the quadratically convergent procedure has

$$\begin{aligned} |\tilde{p}_n - 0| = |\tilde{p}_n| &\approx 0.5|\tilde{p}_{n-1}|^2 \approx (0.5)[0.5|\tilde{p}_{n-2}|^2]^2 = (0.5)^3|\tilde{p}_{n-2}|^4 \\ &\approx (0.5)^3[(0.5)|\tilde{p}_{n-3}|^2]^4 = (0.5)^7|\tilde{p}_{n-3}|^8 \\ &\approx \cdots \approx (0.5)^{2^n-1}|\tilde{p}_0|^{2^n}. \end{aligned}$$

Error Analysis for Iterative Methods

linearly convergent vs. quadratically convergent

n	Linear Convergence Sequence $\{p_n\}_{n=0}^{\infty}$ $(0.5)^n$	Quadratic Convergence Sequence $\{\tilde{p}_n\}_{n=0}^{\infty}$ $(0.5)^{2^n-1}$
1	5.0000×10^{-1}	5.0000×10^{-1}
2	2.5000×10^{-1}	1.2500×10^{-1}
3	1.2500×10^{-1}	7.8125×10^{-3}
4	6.2500×10^{-2}	3.0518×10^{-5}
5	3.1250×10^{-2}	4.6566×10^{-10}
6	1.5625×10^{-2}	1.0842×10^{-19}
7	7.8125×10^{-3}	5.8775×10^{-39}

Error Analysis for Iterative Methods

Theorem (Fixed Point Method)

Let $g \in C[a,b]$ be such that $g(x) \in [a,b]$ for all $x \in [a,b]$. Suppose, in addition, that $g'(x)$ is continuous on (a,b) and a positive constant $k < 1$ exists with

$$|g'(x)| \le k, \forall x \in [a,b].$$

If $\underline{g'(p) \neq 0}$, then for any number $p_0 \neq p$ in $[a,b]$, the sequence

$$p_n = g(p_{n-1}), n \ge 1$$

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

Error Analysis for Iterative Methods

Proof

- Since g' exists on (a, b), applying the MVT, we have

$$p_{n+1} - p = g(p_n) - g(p) = g'(\xi_n)(p_n - p)$$

- Since $\{p_n\}_{n=0}^{\infty}$ converges to p, $\{\xi_n\}_{n=0}^{\infty}$ also converges to p.
- Thus,

$$\lim_{n\to\infty} \frac{p_{n+1} - p}{p_n - p} = \lim_{n\to\infty} g'(\xi_n) = g'(p)$$

  Hence, if $g'(p) \neq 0$, fixed-point iteration exhibits linear convergence with asymptotic error constant $|g'(p)|$.

Error Analysis for Iterative Methods

Theorem

Let p be a solution of the equation $x = g(x)$. Suppose that $g'(p) = 0$ and g'' is continuous with $g''(x) < M$ on an open interval I containing p. Then there exists a $\delta > 0$ such that, for $p_0 \in [p - \delta, p + \delta]$, the sequence defined by $p_n = g(p_{n-1})$, when $n \geq 1$, converges at least quadratically to p. Moreover, for sufficiently large values of n,

$$|p_{n+1} - p| < \frac{M}{2}|p_n - p|^2.$$

Error Analysis for Iterative Methods

Proof (1/2)

- Choose k in $(0, 1)$ and $\delta > 0$ such that on the interval $[p - \delta, p + \delta]$ contained in I, we have $|g'(x)| \leq k$ andg'' continuous.
- Expanding $g(x)$ in a linear Taylor polynomial for $x \in [p - \delta, p + \delta]$ gives

$$g(x) = g(p) + g'(p)(x - p) + \frac{g''(\xi)}{2}(x - p)^2$$

- The hypotheses $g(p) = p$ and $g'(p) = 0$ imply that

$$g(x) = p + \frac{g''(\xi)}{2}(x - p)^2$$

Error Analysis for Iterative Methods

Proof (2/2)

- When $x = p_n$, we have

$$p_{n+1} = g(p_n) = p + \frac{g''(\xi)}{2}(p_n - p)^2$$

- Thus,

$$\lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|^2} = \frac{|g''(p)|}{2} < \frac{M}{2}$$

Error Analysis for Iterative Methods

Remarks

- For a fixed point method to converge quadratically we need to have both $g(p) = p$, and $g'(p) = 0$.
- If $f(p) = 0$ and $f'(p) \neq 0$, then for starting values sufficiently close to p, Newton's method will converge at least quadratically.

Convergence using Newton's Method

Proof: (2/4)

- Thus g is defined and continuous on $[p-\delta_1, p+\delta_1]$. 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 $x \in [p-\delta_1, p+\delta_1]$, and since $f \in C^2[a,b]$, we have $g \in C^1[p-\delta_1, p+\delta_1]$.

- By assumption, $f(p) = 0$, so

$$g'(p) = \frac{f(p)f''(p)}{[f'(p)]^2} = 0.$$

Assignment

- Numerical Analysis, Chapter 1 & 2;
- Homework 1.