MA579H Scientific Computing

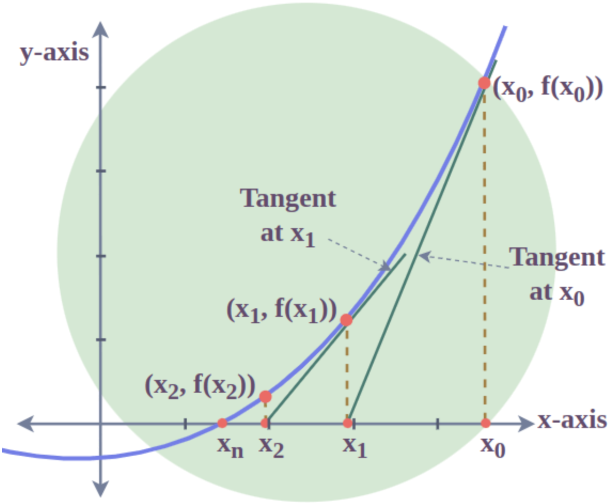
Solutions of Nonlinear equations II

Rafikul Alam
Department of Mathematics
IIT Guwahati

Outline

- Newton method
- Fixed point iteration

Newton's method



Newton's method

Let

$f \longrightarrow$ continuously differentiable;
$\alpha \in \mathbb{R} \longrightarrow f(\alpha) = 0$;
$x_0 \in \mathbb{R} \longrightarrow$ initial guess.

Newton's method

Let

$f \longrightarrow$ continuously differentiable;
$\alpha \in \mathbb{R} \longrightarrow f(\alpha) = 0$;
$x_0 \in \mathbb{R} \longrightarrow$ initial guess.

The tangent to $y = f(x)$ at $(x_0, f(x_0))$ is

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Suppose the tangent is not parallel to the x-axis so that $f'(x_0) \neq 0$. Then it cuts the x-axis at

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

Newton's method

Let

$f \longrightarrow$ continuously differentiable;
$\alpha \in \mathbb{R} \longrightarrow f(\alpha) = 0$;
$x_0 \in \mathbb{R} \longrightarrow$ initial guess.

The tangent to $y = f(x)$ at $(x_0, f(x_0))$ is

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Suppose the tangent is not parallel to the x-axis so that $f'(x_0) \neq 0$. Then it cuts the x-axis at

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

Repeating the process results in Newton's method

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)},\ n = 0, 1, \ldots.$$

Newton's method

Algorithm:

Input: A differentiable function f

Initial guess $c \in \mathbb{R}$ and a tolerance `tol`

A limit `N` for maximum number of iteration.

Output Approximate solution c of $f(x) = 0$ satisfying $|f(c)| \leq$ `tol`.

Newton's method

Algorithm:

Input: A differentiable function f

Initial guess $c \in \mathbb{R}$ and a tolerance $\texttt{tol}$

A limit $\texttt{N}$ for maximum number of iteration.

Output Approximate solution c of $f(x) = 0$ satisfying $|f(c)| \leq \texttt{tol}$.

- IT $= 0$
- **while** $(|f(c)| >$ **tol**) **and** $(\texttt{IT} \leq \texttt{N})$ **do**
- $c = c - f(c)/f'(c)$
- $\texttt{IT} = \texttt{IT} + 1$

Newton's method

Algorithm:

Input: A differentiable function f

Initial guess $c \in \mathbb{R}$ and a tolerance `tol`

A limit `N` for maximum number of iteration.

Output Approximate solution c of $f(x) = 0$ satisfying $|f(c)| \leq$ `tol`.

- $\text{IT} = 0$
- **while** $(|f(c)| >$ **tol**$)$ **and** $(\texttt{IT} \leq \texttt{N})$ **do**
- $c = c - f(c)/f'(c)$
- $\texttt{IT} = \texttt{IT} + 1$

Example: Let $a > 0$. Then the square root $\alpha = \sqrt{a}$ is the zero of $f(x) = x^2 - a$ which gives

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), \; n = 0, 1, 2, \ldots.$$

This scheme converges globally, that is, $x_n \longrightarrow \sqrt{a}$ for any $x_0 \neq 0$. Why?

Newton's method from Secant method

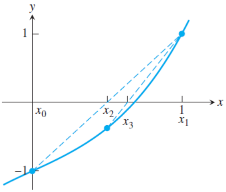


Figure: Secant method for finding the zero of $x^3 + x - 1$.

Given two initial guesses $x_0 \neq x_1$, the secant method generates the iterates

$$x_{n+1} = x_n - \frac{f(x_n)}{f[x_{n-1}, x_n]}, \; n = 1, 2, 3, \ldots.$$

Newton's method from Secant method

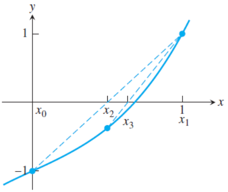


Figure: Secant method for finding the zero of $x^3 + x - 1$.

Given two initial guesses $x_0 \neq x_1$, the secant method generates the iterates

$$x_{n+1} = x_n - \frac{f(x_n)}{f[x_{n-1}, x_n]}, \; n = 1, 2, 3, \ldots.$$

Letting $x_{n-1} \longrightarrow x_n$, the secant becomes the tangent to $y = f(x)$ at x_n and yields Newton's method

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \; n = 0, 1, 2, \ldots.$$

Convergence of Newton's method

Suppose that f is twice continuously differentiable near α and that α is a simple zero, that is, $f'(\alpha) \neq 0$.

Convergence of Newton's method

Suppose that f is twice continuously differentiable near α and that α is a simple zero, that is, $f'(\alpha) \neq 0$.

Then we have

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

Convergence of Newton's method

Suppose that f is twice continuously differentiable near α and that α is a simple zero, that is, $f'(\alpha) \neq 0$.

Then we have

$$\begin{aligned} x_{n+1} - \alpha &= x_n - \alpha - \frac{f(x_n)}{f'(x_n)} = (x_n - \alpha)\left(1 - \frac{f(x_n) - f(\alpha)}{(x_n - \alpha) f'(x_n)}\right) \\ &= (x_n - \alpha)\left(1 - \frac{f[x_n, \alpha]}{f[x_n, x_n]}\right) \end{aligned}$$

Convergence of Newton's method

Suppose that f is twice continuously differentiable near α and that α is a simple zero, that is, $f'(\alpha) \neq 0$.

Then we have

$$\begin{aligned} x_{n+1} - \alpha &= x_n - \alpha - \frac{f(x_n)}{f'(x_n)} = (x_n - \alpha)\left(1 - \frac{f(x_n) - f(\alpha)}{(x_n - \alpha)f'(x_n)}\right) \\ &= (x_n - \alpha)\left(1 - \frac{f[x_n, \alpha]}{f[x_n, x_n]}\right) = (x_n - \alpha)^2 \frac{f[x_n, x_n, \alpha]}{f[x_n, x_n]}. \end{aligned}$$

Convergence of Newton's method

Suppose that f is twice continuously differentiable near α and that α is a simple zero, that is, $f'(\alpha) \neq 0$.

Then we have

$$\begin{aligned} x_{n+1} - \alpha &= x_n - \alpha - \frac{f(x_n)}{f'(x_n)} = (x_n - \alpha)\left(1 - \frac{f(x_n) - f(\alpha)}{(x_n - \alpha) f'(x_n)}\right) \\ &= (x_n - \alpha)\left(1 - \frac{f[x_n, \alpha]}{f[x_n, x_n]}\right) = (x_n - \alpha)^2 \frac{f[x_n, x_n, \alpha]}{f[x_n, x_n]}. \end{aligned}$$

Therefore, if $x_n \longrightarrow \alpha$ then

$$\lim_{n \to \infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{f''(\alpha)}{2 f'(\alpha)}.$$

Convergence of Newton's method

Suppose that f is twice continuously differentiable near α and that α is a simple zero, that is, $f'(\alpha) \neq 0$.

Then we have

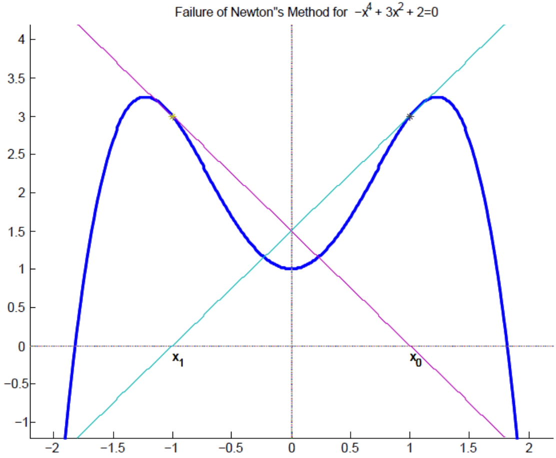
$$\begin{aligned} x_{n+1} - \alpha &= x_n - \alpha - \frac{f(x_n)}{f'(x_n)} = (x_n - \alpha)\left(1 - \frac{f(x_n) - f(\alpha)}{(x_n - \alpha) f'(x_n)}\right) \\ &= (x_n - \alpha)\left(1 - \frac{f[x_n, \alpha]}{f[x_n, x_n]}\right) = (x_n - \alpha)^2 \frac{f[x_n, x_n, \alpha]}{f[x_n, x_n]}. \end{aligned}$$

Therefore, if $x_n \longrightarrow \alpha$ then

$$\lim_{n \to \infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{f''(\alpha)}{2f'(\alpha)}.$$

This shows that the order of convergence of Newton's method is quadratic if $f''(\alpha) \neq 0$.

Non-convergence of Newton's method



Non-convergence Newton's method

Example: Now consider $f(x) := \sin(x)$ for $|x| < \pi/2$. Then $\alpha = 0$ is the only zero. Newton's method becomes $x_{n+1} = x_n - \tan(x_n)$, $n = 0, 1, \ldots$

If $\hat{x}$ is such that $2\hat{x} = \tan(\hat{x})$ then with $x_0 := \hat{x}$, we have $x_1 = -\hat{x}$ and $x_2 = \hat{x}$. Hence after two iterations Newton's method ends up with where we started. This is called cycle.

Non-convergence Newton's method

Example: Now consider $f(x) := \sin(x)$ for $|x| < \pi/2$. Then $\alpha = 0$ is the only zero. Newton's method becomes $x_{n+1} = x_n - \tan(x_n)$, $n = 0, 1, \ldots$

If $\hat{x}$ is such that $2\hat{x} = \tan(\hat{x})$ then with $x_0 := \hat{x}$, we have $x_1 = -\hat{x}$ and $x_2 = \hat{x}$. Hence after two iterations Newton's method ends up with where we started. This is called cycle.

This shows that Newton's method does not converge for $x_0 = \hat{x}$. However, Newton's method converges for any x_0 such that $|x_0| < |\hat{x}|$. ■

Non-convergence Newton's method

Example: Now consider $f(x) := \sin(x)$ for $|x| < \pi/2$. Then $\alpha = 0$ is the only zero. Newton's method becomes $x_{n+1} = x_n - \tan(x_n)$, $n = 0, 1, \ldots$

If $\hat{x}$ is such that $2\hat{x} = \tan(\hat{x})$ then with $x_0 := \hat{x}$, we have $x_1 = -\hat{x}$ and $x_2 = \hat{x}$. Hence after two iterations Newton's method ends up with where we started. This is called cycle.

This shows that Newton's method does not converge for $x_0 = \hat{x}$. However, Newton's method converges for any x_0 such that $|x_0| < |\hat{x}|$. ■

Observation: If $f \in C^2[a, b]$ is convex (or concave) and $f(a)f(b) < 0$ then the tangents at the endpoints of $[a, b]$ intersects x-axis within $[a, b]$. Hence on geometric ground Newton's method converges globally.

Slow convergence of Newton's method

If α is a zero of f with multiplicity more than 1 then $f'(\alpha) = 0$ and the previous convergence analysis of Newton's Method does not apply.

Slow convergence of Newton's method

If α is a zero of f with multiplicity more than 1 then $f'(\alpha) = 0$ and the previous convergence analysis of Newton's Method does not apply.

Example: For $f(x) = x^m$, the Newton iterations are

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

Since $\alpha := 0$ is the zero of f, we have $e_n := |x_n - \alpha| = |x_n|$ and $\frac{e_{n+1}}{e_n} = \frac{m-1}{m}$. Hence the convergence is linear at the rate $\frac{m-1}{m}$.

Slow convergence of Newton's method

If α is a zero of f with multiplicity more than 1 then $f'(\alpha) = 0$ and the previous convergence analysis of Newton's Method does not apply.

Example: For $f(x) = x^m$, the Newton iterations are

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

Since $\alpha := 0$ is the zero of f, we have $e_n := |x_n - \alpha| = |x_n|$ and $\frac{e_{n+1}}{e_n} = \frac{m-1}{m}$. Hence the convergence is linear at the rate $\frac{m-1}{m}$.

Theorem Let f be an $(m+1)$-times continuously differentiable function on $[a, b]$. Let $\alpha \in (a, b)$ be a zero of f of multiplicity m. Then Newton's Method is locally convergent to α and the convergence is linear at the rate $\frac{m-1}{m}$.

Modified Newton's Method for multiple roots

Example: The function $f(x) = \sin x + x^2 \cos x - x^2 - x$ has a root of multiplicity 3 at 0. With starting guess $x_0 = 1$, $e_0 = 1$ and $e_{n+1} \leq (2/3)^n$ and the number of steps say n, required to achieve accuracy up to 6 decimal digits

$$(2/3)^n < 0.5 * 10^{-6} \Rightarrow n > \frac{\log_{10}(0.5) - 6}{\log_{10}(2/3)} \approx 35.78.$$

Modified Newton's Method for multiple roots

Example: The function $f(x) = \sin x + x^2 \cos x - x^2 - x$ has a root of multiplicity 3 at 0. With starting guess $x_0 = 1$, $e_0 = 1$ and $e_{n+1} \leq (2/3)^n$ and the number of steps say n, required to achieve accuracy up to 6 decimal digits

$$(2/3)^n < 0.5 * 10^{-6} \Rightarrow n > \frac{\log_{10}(0.5) - 6}{\log_{10}(2/3)} \approx 35.78.$$

Theorem Let f be an $(m+1)$-times continuously differentiable function on $[a, b]$. Let $\alpha \in (a, b)$ be a zero of f of multiplicity m. Then the modified Newton's Method

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

converges locally to α and the rate of convergence is quadratic.

Modified Newton's Method for multiple roots

Example: The function $f(x) = \sin x + x^2 \cos x - x^2 - x$ has a root of multiplicity 3 at 0. With starting guess $x_0 = 1$, $e_0 = 1$ and $e_{n+1} \leq (2/3)^n$ and the number of steps say n, required to achieve accuracy up to 6 decimal digits

$$(2/3)^n < 0.5 * 10^{-6} \Rightarrow n > \frac{\log_{10}(0.5) - 6}{\log_{10}(2/3)} \approx 35.78.$$

Theorem Let f be an $(m+1)$-times continuously differentiable function on $[a, b]$. Let $\alpha \in (a, b)$ be a zero of f of multiplicity m. Then the modified Newton's Method

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

converges locally to α and the rate of convergence is quadratic.

Modified Newton's Method for multiple roots

Example: The function $f(x) = \sin x + x^2 \cos x - x^2 - x$ has a root of multiplicity 3 at 0. With starting guess $x_0 = 1$, $e_0 = 1$ and $e_{n+1} \leq (2/3)^n$ and the number of steps say n, required to achieve accuracy up to 6 decimal digits

$$(2/3)^n < 0.5 * 10^{-6} \Rightarrow n > \frac{\log_{10}(0.5) - 6}{\log_{10}(2/3)} \approx 35.78.$$

Theorem Let f be an $(m+1)$-times continuously differentiable function on $[a, b]$. Let $\alpha \in (a, b)$ be a zero of f of multiplicity m. Then the modified Newton's Method

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

converges locally to α and the rate of convergence is quadratic.

The Modified Newton's Method applied to $f(x) = \sin x + x^2 \cos x - x^2 - x$ produces a solution correct up to 6 decimal digits in 5 iterations.

Fixed Point

Consider a function $\phi : [a, b] \to \mathbb{R}$. A point $c \in [a, b]$ is said to be a fixed point of ϕ if $\phi(c) = c$.

Fixed Point

Consider a function $\phi : [a, b] \to \mathbb{R}$. A point $c \in [a, b]$ is said to be a fixed point of ϕ if $\phi(c) = c$.

Examples:

(a) $f(x) = x^2 + 3x + 1$ has a fixed point at -1.

(b) $f(x) = \sin x$ has a fixed point at 0.

(c) If α is a zero of any function $f : [a, b] \to \mathbb{R}$, then α is a fixed point of $g(x) := x + f(x)$.

Fixed Point

Consider a function $\phi : [a, b] \to \mathbb{R}$. A point $c \in [a, b]$ is said to be a fixed point of ϕ if $\phi(c) = c$.

Examples:

(a) $f(x) = x^2 + 3x + 1$ has a fixed point at -1.

(b) $f(x) = \sin x$ has a fixed point at 0.

(c) If α is a zero of any function $f : [a, b] \to \mathbb{R}$, then α is a fixed point of $g(x) := x + f(x)$.

When does a function f have a unique fixed point?

Fixed Point

Consider a function $\phi : [a, b] \to \mathbb{R}$. A point $c \in [a, b]$ is said to be a fixed point of ϕ if $\phi(c) = c$.

Examples:

(a) $f(x) = x^2 + 3x + 1$ has a fixed point at -1.

(b) $f(x) = \sin x$ has a fixed point at 0.

(c) If α is a zero of any function $f : [a, b] \to \mathbb{R}$, then α is a fixed point of $g(x) := x + f(x)$.

When does a function f have a unique fixed point? Ans: Contractive function.

Fixed Point

Consider a function $\phi : [a, b] \to \mathbb{R}$. A point $c \in [a, b]$ is said to be a fixed point of ϕ if $\phi(c) = c$.

Examples:

(a) $f(x) = x^2 + 3x + 1$ has a fixed point at -1.

(b) $f(x) = \sin x$ has a fixed point at 0.

(c) If α is a zero of any function $f : [a, b] \to \mathbb{R}$, then α is a fixed point of $g(x) := x + f(x)$.

When does a function f have a unique fixed point? Ans: Contractive function.

A function $f : [a, b] \to \mathbb{R}$ is said to be a *contractive function or contraction on* $[a, b]$ if there exists $\lambda \in \mathbb{R}$ with $0 < \lambda < 1$ such that $|f(x) - f(y)| \leq \lambda |x - y|$ for all $x, y \in [a, b]$.

Fixed Point

Consider a function $\phi : [a, b] \to \mathbb{R}$. A point $c \in [a, b]$ is said to be a fixed point of ϕ if $\phi(c) = c$.

Examples:

(a) $f(x) = x^2 + 3x + 1$ has a fixed point at -1.

(b) $f(x) = \sin x$ has a fixed point at 0.

(c) If α is a zero of any function $f : [a, b] \to \mathbb{R}$, then α is a fixed point of $g(x) := x + f(x)$.

When does a function f have a unique fixed point? Ans: Contractive function.

A function $f : [a, b] \to \mathbb{R}$ is said to be a *contractive function or contraction on* $[a, b]$ if there exists $\lambda \in \mathbb{R}$ with $0 < \lambda < 1$ such that $|f(x) - f(y)| \leq \lambda |x - y|$ for all $x, y \in [a, b]$.

Examples:

(a) $f(x) = \frac{x^2+2x+4}{4}$ is a contraction on $(-2, 0.75)$.

(b) $f(x) = \sin x$ is a contraction on $[\frac{\pi}{6}, \frac{5\pi}{6}]$.

(c) $f(x) = \cos x$ is a contraction on $[0, \frac{\pi}{3}]$.

Fixed-point iteration

Let $\phi : [a, b] \to \mathbb{R}$. Suppose that ϕ has a fixed-point in $[a, b]$. For $x_0 \in [a, b]$, consider the fixed-point iteration (FPI)

$$x_n = \phi(x_{n-1}), \; n = 1, 2, \ldots.$$

If (x_n) converges to α then $\phi(\alpha) = \alpha$. Thus α is a fixed-point.

Fixed-point iteration

Let $\phi : [a, b] \to \mathbb{R}$. Suppose that ϕ has a fixed-point in $[a, b]$. For $x_0 \in [a, b]$, consider the fixed-point iteration (FPI)

$$x_n = \phi(x_{n-1}), \; n = 1, 2, \ldots.$$

If (x_n) converges to α then $\phi(\alpha) = \alpha$. Thus α is a fixed-point.

Remark: The function ϕ is called the iteration function for the fixed-point iteration.

Fixed-point iteration

Let $\phi : [a, b] \to \mathbb{R}$. Suppose that ϕ has a fixed-point in $[a, b]$. For $x_0 \in [a, b]$, consider the fixed-point iteration (FPI)

$$x_n = \phi(x_{n-1}),\ n = 1, 2, \ldots.$$

If (x_n) converges to α then $\phi(\alpha) = \alpha$. Thus α is a fixed-point.

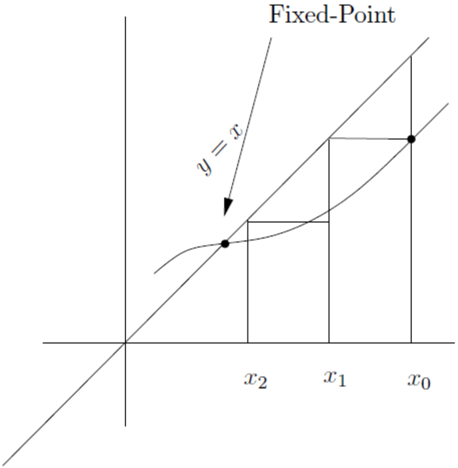
Remark: The function ϕ is called the iteration function for the fixed-point iteration.

Example: Suppose we wish to solve $\cos(e^x) = 0$. Note that

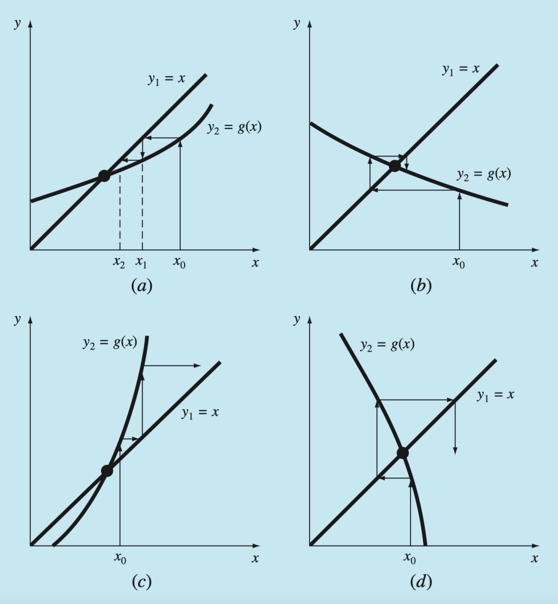
$$\cos(e^c) = 0 \iff \cos(e^c) + c = c.$$

Setting $\phi(x) := \cos(e^x) + x$, we have $\phi(c) = c$. Hence the fixed-point iteration $x_n = \phi(x_{n-1})$ can be used to solve $\cos(e^x) = 0$.

Fixed-point iteration



Fixed-point iteration



Convergence of fixed-point iteration

Question: When does a function ϕ have a fixed point?

Convergence of fixed-point iteration

Question: When does a function ϕ have a fixed point? When does a fixed-point iteration converge?

Convergence of fixed-point iteration

Question: When does a function ϕ have a fixed point? When does a fixed-point iteration converge?

A function $\phi : [a, b] \to \mathbb{R}$ is said to be a contraction (or contractive) if there exists a constant $0 < \lambda < 1$ such that

$$|\phi(x) - \phi(y)| \leq \lambda |x - y| \text{ for all } x, y \in [a, b].$$

Convergence of fixed-point iteration

Question: When does a function ϕ have a fixed point? When does a fixed-point iteration converge?

A function $\phi : [a, b] \to \mathbb{R}$ is said to be a contraction (or contractive) if there exists a constant $0 < \lambda < 1$ such that

$$|\phi(x) - \phi(y)| \leq \lambda |x - y| \text{ for all } x, y \in [a, b].$$

Theorem: If $\phi : [a, b] \to \mathbb{R}$ is a contraction then it has a unique fixed point. Further, for any $x_0 \in [a, b]$, the iteration $x_n = \phi(x_{n-1})$ converges to the unique fixed point.

Proof: Consider $x_n = \phi(x_{n-1}), n = 1, 2, \ldots$. Then
$|x_{n+1} - x_n| \leq \lambda |x_n - x_{n-1}| \Rightarrow (x_n)$ is a Cauchy sequence (Check).

Convergence of fixed-point iteration

Question: When does a function ϕ have a fixed point? When does a fixed-point iteration converge?

A function $\phi : [a, b] \to \mathbb{R}$ is said to be a contraction (or contractive) if there exists a constant $0 < \lambda < 1$ such that

$$|\phi(x) - \phi(y)| \leq \lambda |x - y| \text{ for all } x, y \in [a, b].$$

Theorem: If $\phi : [a, b] \to \mathbb{R}$ is a contraction then it has a unique fixed point. Further, for any $x_0 \in [a, b]$, the iteration $x_n = \phi(x_{n-1})$ converges to the unique fixed point.

Proof: Consider $x_n = \phi(x_{n-1}), n = 1, 2, \ldots$. Then $|x_{n+1} - x_n| \leq \lambda |x_n - x_{n-1}| \Rightarrow (x_n)$ is a Cauchy sequence (Check).

Hence $x_n \to c$ for some $c \in [a, b]$. Now $x_n = \phi(x_{n-1}) \Rightarrow c = \phi(c)$. The uniqueness of c is immediate. ■

Observation: The fixed point iteration converges linearly at the rate λ.

Sufficient condition for contraction

Fact: Let ϕ be differentiable on $[a, b]$ such that $\max_{x \in [a,b]} |\phi'(x)| < 1$. Then ϕ is a contraction on $[a, b]$.

Sufficient condition for contraction

Fact: Let ϕ be differentiable on $[a, b]$ such that $\max_{x\in[a,b]} |\phi'(x)| < 1$. Then ϕ is a contraction on $[a, b]$. Indeed, by mean value theorem $|\phi(x) - \phi(y)| = |\phi'(c)|\,|x - y|$ shows that $\phi(x)$ is a contraction on $[a, b]$.

Sufficient condition for contraction

Fact: Let ϕ be differentiable on $[a, b]$ such that $\max_{x\in[a,b]} |\phi'(x)| < 1$. Then ϕ is a contraction on $[a, b]$. Indeed, by mean value theorem $|\phi(x) - \phi(y)| = |\phi'(c)|\,|x - y|$ shows that $\phi(x)$ is a contraction on $[a, b]$.

Example: Consider $\phi(x) = \cos(e^x) + x$. Then

$$\phi'(x) = 1 - e^x \sin(e^x) \Rightarrow \lambda := \max_{x\in[0,1]} |\phi'(x)| < 1.$$

Hence $\phi(x)$ is a contraction on $[0, 1]$.

Sufficient condition for contraction

Fact: Let ϕ be differentiable on $[a, b]$ such that $\max_{x\in[a,b]} |\phi'(x)| < 1$. Then ϕ is a contraction on $[a, b]$. Indeed, by mean value theorem $|\phi(x) - \phi(y)| = |\phi'(c)|\,|x - y|$ shows that $\phi(x)$ is a contraction on $[a, b]$.

Example: Consider $\phi(x) = \cos(e^x) + x$. Then

$$\phi'(x) = 1 - e^x \sin(e^x) \Rightarrow \lambda := \max_{x\in[0,1]} |\phi'(x)| < 1.$$

Hence $\phi(x)$ is a contraction on $[0, 1]$.

Newton method as a fixed-point iteration: Consider $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. Then setting $\phi(x) := x - \frac{f(x)}{f'(x)}$, we have $x_{n+1} = \phi(x_n), n = 0, 1, \ldots$.

Sufficient condition for contraction

Fact: Let ϕ be differentiable on $[a, b]$ such that $\max_{x\in[a,b]} |\phi'(x)| < 1$. Then ϕ is a contraction on $[a, b]$. Indeed, by mean value theorem $|\phi(x) - \phi(y)| = |\phi'(c)|\,|x - y|$ shows that $\phi(x)$ is a contraction on $[a, b]$.

Example: Consider $\phi(x) = \cos(e^x) + x$. Then

$$\phi'(x) = 1 - e^x \sin(e^x) \Rightarrow \lambda := \max_{x\in[0,1]} |\phi'(x)| < 1.$$

Hence $\phi(x)$ is a contraction on $[0, 1]$.

Newton method as a fixed-point iteration: Consider $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. Then setting $\phi(x) := x - \frac{f(x)}{f'(x)}$, we have $x_{n+1} = \phi(x_n), n = 0, 1, \ldots$.

This shows that if ϕ is a contraction on $[a, b]$ then for any $x_0 \in [a, b]$, the Newton method will converge to a solution of $f(x) = 0$.

Sufficient condition for contraction

Fact: Let ϕ be differentiable on $[a, b]$ such that $\max_{x\in[a,b]} |\phi'(x)| < 1$. Then ϕ is a contraction on $[a, b]$. Indeed, by mean value theorem $|\phi(x) - \phi(y)| = |\phi'(c)|\,|x - y|$ shows that $\phi(x)$ is a contraction on $[a, b]$.

Example: Consider $\phi(x) = \cos(e^x) + x$. Then

$$\phi'(x) = 1 - e^x \sin(e^x) \Rightarrow \lambda := \max_{x\in[0,1]} |\phi'(x)| < 1.$$

Hence $\phi(x)$ is a contraction on $[0, 1]$.

Newton method as a fixed-point iteration: Consider $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. Then setting $\phi(x) := x - \frac{f(x)}{f'(x)}$, we have $x_{n+1} = \phi(x_n), n = 0, 1, \ldots$.

This shows that if ϕ is a contraction on $[a, b]$ then for any $x_0 \in [a, b]$, the Newton method will converge to a solution of $f(x) = 0$. Now

$$\phi'(x) = 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}$$

yields a sufficient condition for convergence of Newton method.

Choice of iteration function

Consider the equation $f(x) := xe^x - 1 = 0$. Then fixed-point iterations are obtained as follows:

- $xe^x - 1 = 0 \iff x = e^{-x} \Rightarrow x_n = e^{-x_{n-1}} = \phi(x_{n-1}), n = 1, 2, \ldots.$ With the initial guess $x_0 = 0.5$, x_n converges to $c = 0.5671432\ldots$

Choice of iteration function

Consider the equation $f(x) := xe^x - 1 = 0$. Then fixed-point iterations are obtained as follows:

- $xe^x - 1 = 0 \iff x = e^{-x} \Rightarrow x_n = e^{-x_{n-1}} = \phi(x_{n-1})$, $n = 1, 2, \ldots$. With the initial guess $x_0 = 0.5$, x_n converges to $c = 0.5671432\ldots$.
- $xe^x - 1 = 0 \iff xe^x + x = 1 + x \iff x = \frac{1+x}{1+e^x}$ which gives

$$x_n = \frac{1 + x_{n-1}}{1 + e^{x_{n-1}}} = \phi(x_{n-1}), n = 1, 2, \ldots.$$

This time the convergence is much faster - we need only three iterations to obtain a 10-digit approximation of c. Indeed, $x_1 = 0.5663110032, x_2 = 0.5671431650$ and $x_3 = 0.5671432904$.

Choice of iteration function

Consider the equation $f(x) := xe^x - 1 = 0$. Then fixed-point iterations are obtained as follows:

- $xe^x - 1 = 0 \iff x = e^{-x} \Rightarrow x_n = e^{-x_{n-1}} = \phi(x_{n-1}), n = 1, 2, \ldots.$ With the initial guess $x_0 = 0.5$, x_n converges to $c = 0.5671432\ldots.$
- $xe^x - 1 = 0 \iff xe^x + x = 1 + x \iff x = \frac{1+x}{1+e^x}$ which gives

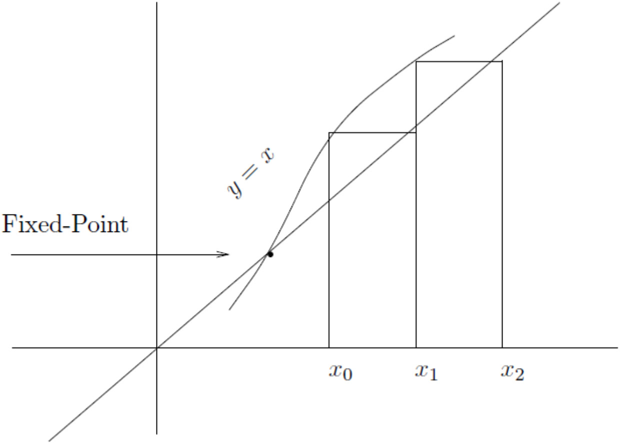
$$x_n = \frac{1 + x_{n-1}}{1 + e^{x_{n-1}}} = \phi(x_{n-1}), n = 1, 2, \ldots.$$

  This time the convergence is much faster - we need only three iterations to obtain a 10-digit approximation of c. Indeed, $x_1 = 0.5663110032, x_2 = 0.5671431650$ and $x_3 = 0.5671432904$.
- $xe^x - 1 = 0 \iff x = x + 1 - xe^x$ which yields the iteration

$$x_n = x_{n-1} + 1 - x_{n-1}e^{x_{n-1}} = \phi(x_{n-1}),\ n = 1, 2, \ldots.$$

  However, this iteration function ϕ does not generate a convergent sequence.

Non-convergence of Fixed-point iteration



Finding zeros via fixed-point iterations

Note that if α is a solution α of $f(x) = 0$ then α is a fixed point of $g(x) := f(x) + x$.

In fact, solutions of nonlinear equations may be viewed as fixed points of infinitely many functions.

Finding zeros via fixed-point iterations

Note that if α is a solution α of $f(x) = 0$ then α is a fixed point of $g(x) := f(x) + x$.

In fact, solutions of nonlinear equations may be viewed as fixed points of infinitely many functions.

Example: A solution of $x^3 + x - 1 = 0$ may be found as a fixed point of some $g(x)$

$$(a)\ \ g(x) = 1 - x^3 \quad (b)\ \ g(x) = (1-x)^{\frac{1}{3}} \quad (c)\ \ g(x) = \frac{1+2x^3}{1+3x^2}.$$

Finding zeros via fixed-point iterations

Note that if α is a solution α of $f(x) = 0$ then α is a fixed point of $g(x) := f(x) + x$.

In fact, solutions of nonlinear equations may be viewed as fixed points of infinitely many functions.

Example: A solution of $x^3 + x - 1 = 0$ may be found as a fixed point of some $g(x)$

$$(a)\ \ g(x) = 1 - x^3 \quad (b)\ \ g(x) = (1-x)^{\frac{1}{3}} \quad (c)\ \ g(x) = \frac{1+2x^3}{1+3x^2}.$$

With $x_0 = 0.5$ the iterations $x_{i+1} = g(x_i)$

- *do not converge* for (a),
- *converge very slowly* for (b).
- **converge in 7 iterations** for (c).

Finding zeros via fixed-point iterations

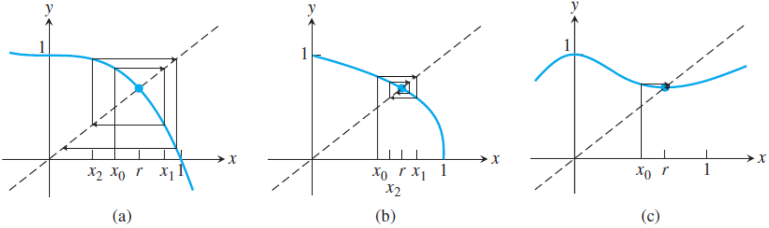


Figure: Diagrams for FPIs to find a zero of $x^3 + x - 1$ showing (a) no convergence with $g(x) = 1 - x^3$ (b) slow convergence for $g(x) = (1 - x)^{1/3}$ and (c) fast convergence for $g(x) = \frac{1+2x^3}{1+3x^2}$.

Finding zeros via fixed-point iterations

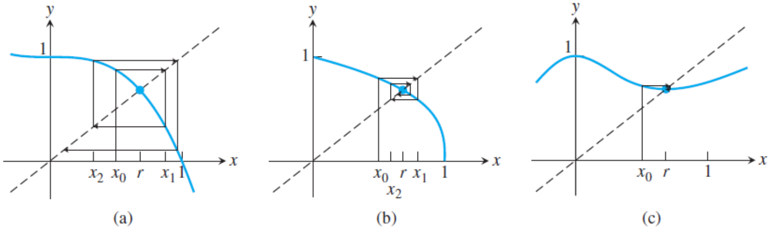


Figure: Diagrams for FPIs to find a zero of $x^3 + x - 1$ showing (a) no convergence with $g(x) = 1 - x^3$ (b) slow convergence for $g(x) = (1-x)^{1/3}$ and (c) fast convergence for $g(x) = \frac{1+2x^3}{1+3x^2}$.

Here $r \approx 0.6823$ is the only real solution of $x^3 + x - 1 = 0$ and

- $|g'(r)| \approx 1.3966 > 1$ for $g(x) = 1 - x^3$;
- $|g'(r)| \approx 0.716 < 1$ for $g(x) = (1-x)^{1/3}$;
- $|g'(r)| = 0$ for $g(x) = \frac{1+2x^3}{1+3x^2}$.