Error analysis

Definition (Order of convergence)

Suppose $p_n \to p$. If $\exists\, \lambda, \alpha > 0$ s.t.

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

then $\{p_n\}$ is said to converge to p of **order** α, with asymptotic error constant λ.

Error analysis

Definition (Convergence order of numerical methods)

An iterative method $p_n = g(p_{n-1})$ is of **order** α if the generated $\{p_n\}$ converges to the solution p of $p = g(p)$ at order α.

In particular:

- $\alpha = 1$: **linearly convergent**
- $\alpha = 2$: **quadratically convergent**

Example

Example (Speed comparison: linear vs quadratic)

Suppose p_n (and q_n respectively) converges to 0 linearly (quadratically) with constant 0.5, enumerate the upper bound of $|p_n|$ and $|q_n|$.

Solution. By definition of convergence order, we know

$$\lim_{n\to\infty} \frac{|p_{n+1}|}{|p_n|} = 0.5 \qquad \text{and} \qquad \lim_{n\to\infty} \frac{|q_{n+1}|}{|q_n|^2} = 0.5$$

Suppose that p_0 and q_0 are close enough to 0 s.t. $|p_{n+1}|/|p_n| \approx 0.5$ and $|q_{n+1}|/|q_n| \approx 0.5$ for all n, then

$$|p_n| \approx 0.5|p_{n-1}| \approx 0.5^2|p_{n-2}| \approx \cdots \approx 0.5^n|p_0|$$

$$|q_n| \approx 0.5|q_{n-1}|^2 \approx 0.5 \cdot 0.5^2|q_{n-2}|^4 \approx \cdots \approx 0.5^{2^n-1}|q_0|^{2^n}$$

Example

Example (Speed comparison: linear vs quadratic)

Suppose $p_0, q_0 \approx 0.5$. Then

n	Linear 0.5^n	Quadratic 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}

Convergence rate of fixed point iteration algorithm

Theorem (FPI alg has linear convergence rate)

Suppose $g \in C[a, b]$ *s.t.* $g(x) \in [a, b]$, $\forall x \in [a, b]$. *If* $\exists\, k \in (0, 1)$ *s.t.* $|g'(x)| \le k$, $\forall x \in (a, b)$, *then* $\{p_n\}$ *generated by FPI algorithm converges to the unique FP of* $g(x)$ *on* $[a, b]$ ***linearly***.

Proof.

We already know $p_n \to p$ where p is the unique fixed point of g by FPI theorem. Also $p_{n+1} - p = g(p_n) - g(p) = g'(\xi(p_n))(p_n - p)$ where $\xi(p_n)$ is between p_n and p. So

$$\lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|} = \lim_{n\to\infty} |g'(\xi(p_n))| = |g'(\lim_{n\to\infty} \xi(p_n))| = |g'(p)| \le k < 1$$

So $p_n \to p$ linearly with constant k. □

Improve convergence order of FPI to quadratic

Theorem (Additional condition for quadratic rate)

If $g \in C^2[a, b]$ *and* $g'(p) = 0$ *for a FP* $p \in (a, b)$*, then* $\exists M > 0$ *s.t.* $|g''(x)| \leq M, \forall x \in [a, b]$ *and* $\exists \delta > 0$ *s.t. sequence* $\{p_n\}$ *by FPI stared in* $[p - \delta, p + \delta]$ *satisfies*

$$|p_{n+1} - p| \leq \frac{M}{2}|p_n - p|^2, \qquad \forall n$$

Improve convergence order of FPI

Proof.

$g \in C^2, g(p) = p, g'(p) = 0$ together imply $\exists \delta > 0$ and $k \in (0, 1)$ s.t. $|g'(x)| \leq k < 1$ for all $x \in [p - \delta, p + \delta]$ and $g : [p - \delta, p + \delta] \to [p - \delta, p + \delta]$. Also

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

where $\xi(p_n)$ is between p_n and p.

Since $p_{n+1} = g(p_n)$, $g(p) = p$, and $g'(p) = 0$, we have $p_{n+1} = p + \frac{1}{2}g''(\xi(p_n))(p_n - p)^2$. So

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

□

Improve convergence order of FPI

Suppose we have a fixed point method with $g(x) = x - \phi(x)f(x)$. How to choose ϕ such that FPI converges quadratically?

We need g s.t. $g'(p) = 0$ at a FP p (root of f):

$$g'(p) = 1 - \phi'(p)f(p) - \phi(p)f'(p) = 0$$

Since $f(p) = 0$ we have $\phi(p) = \frac{1}{f'(p)}$. Choose $\phi(x) = \frac{1}{f'(x)}$ s.t.

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

This is exactly Newton's method!

So Newton's method converges quadratically.

Convergence of Newton's method when $f'(p) = 0$

We mentioned condition $f'(p) \neq 0$ at the root p of f in the convergence proof of Newton's method above.

What if $f'(p) = 0$? When will this happen and how to address it?

Multiple roots

$f'(p) = 0$ at root p means p is not a "simple root".

Definition (Root multiplicity)

A solution p of $f(x)$ is a **root (zero) of multiplicity** m if $f(x) = (x-p)^m q(x)$ for some q s.t. $\lim_{x\to p} q(x) \neq 0$.

Definition (Simple root)

p is a **simple root (zero)** of f if its multiplicity $m = 1$.

Multiple roots

Theorem (S.N.C. for simple root)

$f \in C^1[a,b]$ has a simple root $p \in (a,b)$ iff $f(p) = 0$ and $f'(p) \neq 0$.

Proof.

"$\Longrightarrow$": $f(x) = (x-p)q(x)$ where $\lim_{x\to p} q(x) \neq 0$. Then $f'(x) = q(x) + (x-p)q'(x)$. So $f \in C^1$ implies

$$f'(p) = \lim_{x\to p} f'(x) = \lim_{x\to p} (q(x) + (x-p)q'(x)) \neq 0$$

"$\Longleftarrow$": $f(x) = f(p) + f'(\xi(x))(x-p)$ where $\xi(x)$ between x and p. Define $q(x) = f'(\xi(x))$ then

$$\lim_{x\to p} q(x) = \lim_{x\to p} f'(\xi(x)) = f'(\lim_{x\to p} \xi(x)) = f'(p) \neq 0$$

So f has a simple root at p. □

Multiple roots

Theorem (S.N.C. for multiple root)

$f \in C^m[a,b]$ has a zero p of multiplicity m iff

$$f(p) = f'(p) = \cdots = f^{(m-1)}(p) = 0 \quad \textit{and} \quad f^{(m)}(p) \neq 0$$

Proof.

Hint: Follow the proof above and use

$$(uv)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} u^{(k)} v^{(n-k)}$$

□

Example

Example (Multiple root)

Let $f(x) = e^x - x - 1$, show that $f(x)$ has a zero of multiplicity 2 at $x = 0$.

Solution. $f(x) = e^x - x - 1$, $f'(x) = e^x - 1$, and $f''(x) = e^x$. So $f(0) = f'(0) = 0$ and $f''(0) = 1 \neq 0$. By Theorem above f has root (zero) at $x = 0$ of multiplicity 2.

Modified Newton's method

Instead of using $f(x)$ in Newton's method, we can replace f by

$$\mu(x) := \frac{f(x)}{f'(x)}$$

We need to show:

$$\boxed{p \text{ is a root (simple or not) of } f \implies p \text{ is a simple root of } \mu}$$

Modified Newton's method

Recall that f has a root p of multiplicity m if $f(x) = (x-p)^m q(x)$ for some q with $\lim_{x \to p} q(x) \neq 0$.

Now there is

$$\mu(x) = \frac{f(x)}{f'(x)} = \frac{(x-p)^m q(x)}{m(x-p)^{m-1} q(x) + (x-p)^m q'(x)}$$
$$= (x-p) \cdot \frac{q(x)}{mq(x) + (x-p)q'(x)}$$

where $\frac{q(x)}{mq(x)+(x-p)q'(x)} \to \frac{1}{m} \neq 0$ as $x \to p$.

By definition, $\mu(x)$ has simple root at p, i.e., $\mu(p) = 0$ and $\mu'(p) \neq 0$.

Modified Newton's method

Now we use $\mu(x)$ instead of $f(x)$ in Newton's method:

$$g(x) = x - \frac{\mu(x)}{\mu'(x)} = x - \frac{(f(x)/f'(x))}{(f(x)/f'(x))'} = \cdots = x - \frac{f(x)f'(x)}{(f'(x))^2 - f(x)f''(x)}$$

The **modified Newton's method** is

$$p_n = p_{n-1} - \frac{f(p_{n-1})f'(p_{n-1})}{(f'(p_{n-1}))^2 - f(p_{n-1})f''(p_{n-1})}$$

Drawbacks of the modified Newton's method:

- Needs f'' in computation.
- Denominator approximates 0 as $p_n \to p$, so round-off may degrade convergence.

Accelerating convergence

We showed that FPI generally has linear convergence only. How to improve?

Suppose N is large, and p_n, p_{n+1}, p_{n+2} satisfy

$$\begin{aligned}
& \frac{p_{n+1}-p}{p_n-p} \approx \frac{p_{n+2}-p}{p_{n+1}-p} \\
\Longleftrightarrow \quad & (p_{n+1}-p)^2 \approx (p_n-p)(p_{n+2}-p) = p_n p_{n+2} - p(p_{n+2}+p_n) + p^2 \\
\Longleftrightarrow \quad & p \approx \frac{p_n p_{n+2} - p_{n+1}^2}{p_{n+2} - 2p_{n+2} + p_n} = \cdots = p_n - \frac{(p_{n+1}-p_n)^2}{p_{n+2} - 2p_{n+1} + p_n}
\end{aligned}$$

Aitken's Δ^2 method

Denote $\Delta p_n := p_{n+1} - p_n$, called **forward difference**, and

$$\begin{aligned}\Delta^2 p_n :=& \Delta(\Delta p_n) = \Delta(p_{n+1} - p_n) \\ =& (p_{n+2} - p_{n+1}) - (p_{n+1} - p_n) \\ =& p_{n+2} - 2p_{n+1} + p_n\end{aligned}$$

So the result above can be written as $p \approx p_n - \frac{(\Delta p_n)^2}{\Delta^2 p_n}$.

Aitken's Δ^2 method:
Given $\{p_n\}$ generated by FPI, set $\hat{p}_n = p_n - \frac{(\Delta p_n)^2}{\Delta^2 p_n}$. Then $\hat{p}_n \to p$ faster than p_n.

Aitken's Δ^2 method

What does it mean by "faster"?

Theorem (Faster convergence by Aitken's Δ^2 method)

If $p_n \to p$ linearly with $\lim_{n\to\infty} \frac{p_{n+1}-p}{p_n-p} < 1$, then $\hat{p}_n$ computed by Aitken's Δ^2 method satisfy

$$\lim_{n\to\infty} \frac{\hat{p}_n - p}{p_n - p} = 0$$

Proof.

Hint: Define $e_n := p_n - p$, then $\Delta e_n = \Delta p_n$, $\Delta^2 e_n = \Delta^2 p_n$, and $\frac{e_{n+1}}{e_n} \to \lambda < 1$. Then

$$\frac{\hat{p}_n - p}{p_n - p} = \frac{p_n - \frac{(\Delta p_n)^2}{\Delta^2 p_n} - p}{p_n - p} = \frac{e_n - \frac{(\Delta e_n)^2}{\Delta^2 e_n}}{e_n} = \frac{\frac{e_{n+2}}{e_{n+1}} - \frac{e_{n+1}}{e_n}}{\frac{e_{n+2}}{e_{n+1}} - 2 + \frac{e_n}{e_{n+1}}} \to \frac{\lambda - \lambda}{\lambda - 2 + \frac{1}{\lambda}} = 0$$

$\square$

Steffenson's method

Aitken's method computes $\hat{p}_n$ separately from p_n. Steffenson's method makes use of $\hat{p}_n$ to compute future p_n.

Steffenson's method: given g for FPI, compute

$$
\begin{aligned}
&p_0^{(0)}, && p_0^{(0)} = g(p_0^{(0)}), \quad p_0^{(0)} = g(p_0^{(0)}) \\
&p_0^{(1)} = p_0^{(0)} - \frac{(\Delta p_0^{(0)})^2}{\Delta^2 p_0^{(0)}}, && p_1^{(1)} = g(p_0^{(1)}), \quad p_2^{(1)} = g(p_1^{(1)}) \\
&p_0^{(2)} = p_0^{(1)} - \frac{(\Delta p_0^{(1)})^2}{\Delta^2 p_0^{(1)}}, && p_1^{(2)} = g(p_0^{(2)}), \quad p_2^{(2)} = g(p_1^{(2)}) \\
&\vdots
\end{aligned}
$$

Steffenson's method

Steffenson's method

- **Input.** Initial guess p_0, ϵ_{tol}, $N_{\max}$. Set $N = 1$.
- While $N \le N_{\max}$, do :
  1. Set $p_1 = g(p_0)$, $p_2 = g(p_1)$ and $p = p_0 - \frac{(p_1 - p_0)^2}{p_2 - 2p_1 + p_0}$
  2. If $|p - p_0| < \epsilon_{\text{tol}}$, STOP
  3. $p_0 = p$
  4. Set $N = N + 1$
- **Output.** Return p. If $N \ge N_{\max}$, print("Max iteration reached.").

Steffenson's method

Theorem

Suppose $g(x)$ has a fixed point p and $g'(p) \neq 1$. If $\exists\, \delta > 0$, s.t. $f \in C^3[p-\delta, p+\delta]$, then Steffenson's method generates a sequence $\{p_n\}$ converging to p quadratically for any initial $p_0 \in [p-\delta, p+\delta]$.