October 31, 2024

Announcements

- HW 7
- 4CH 1 } out later today

Goals

- wild things for 1D root solving
- nD
  - ↳ Newton
  - ↳ FPI

Review

- FPI

  $x_{n+1} = g(x_n)$

  linear $\Leftarrow |g'(x^*)| < 1$

  quadratic $\Leftarrow g'(x^*) = 0$

- Newton's method
- Rates of convergence
  - ↳ reliability
  - ↳ quadratic / linear

Newton's Method

Derive Newton's method.

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

$$= g(x_n)$$

m?

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

Demo: Newton's method [cleared]

Convergence and Properties of Newton

What's the rate of convergence of Newton's method?

$$\left(\frac{f}{g}\right)' = \frac{f'g - g'f}{g^2}$$

$$g'(x) = 1 - \underbrace{\frac{f'f'}{(f')^2}}_{1} + \frac{ff''}{(f')^2} = \frac{ff''}{(f')^2}$$

At a root: $f(x^*) = 0$

At a single root: $f(x^*) = 0$, $f'(x^*) \neq 0$,

Newton converges quadratically

Drawbacks of Newton?

- not super robust
- need f'

Demo: Convergence of Newton's Method [cleared]

Convergence and Properties of Newton

What's the rate of convergence of Newton's method?

Drawbacks of Newton?

- Convergence (argument) only *locally*
- Have to know f'!

Demo: Convergence of Newton's Method [cleared]

Secant Method

What would Newton without the use of the derivative look like?

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

$x_0 = (s.g.)$

$x_1 = (s.g.)$

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

Convergence of Properties of Secant

Rate of convergence is $\left(1+\sqrt{5}\right)/2 \approx 1.618$. (proof)

Drawbacks of Secant?

- non-robust
- [illegible]

Demo: Secant Method [cleared]
Demo: Convergence of the Secant Method [cleared]

Secant (and similar methods) are called Quasi-Newton Methods.

Convergence of Properties of Secant

Rate of convergence is $\left(1+\sqrt{5}\right)/2 \approx 1.618$. (proof)

Drawbacks of Secant?

- Convergence argument only good *locally*
  Will see: convergence only local (near root)
- Slower convergence
- Need two starting guesses

Demo: Secant Method [cleared]
Demo: Convergence of the Secant Method [cleared]

Secant (and similar methods) are called Quasi-Newton Methods.

Improving on Newton?

How would we do “Newton + 1” (i.e. even faster, even better)?

Improving on Newton?

How would we do "Newton + 1" (i.e. even faster, even better)?

Easy:

- Use second order Taylor approximation, solve resulting quadratic
- Get cubic convergence!
- Get a method that's *extremely* fast and *extremely* brittle
- Need second derivatives
- What if the quadratic has no solution?

Root Finding with Interpolants

Secant method uses a linear approximation to f based on points $f(x_k)$, $f(x_{k-1})$, could use more points and higher-order approximation:

Put a parabola through $(x_k, f(x_k))$, $(x_{k-1}, f(x_{k-1}))$, $(x_{k-2}, f(x_{k-2}))$

Solve for parabola $= 0$

Muller's method, $r \approx 1.84$

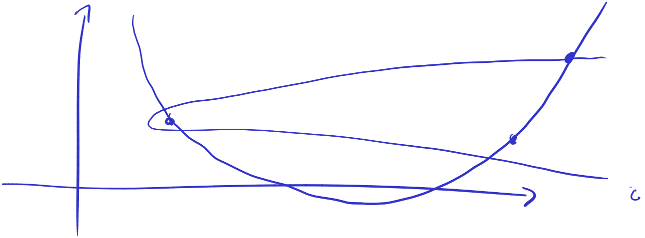
What about existence of roots in that case?

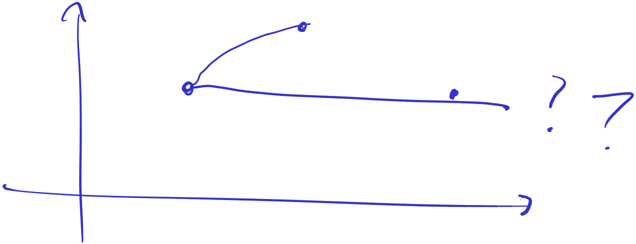
(added after class ⚠)

$(y_k = f(x_k), \underbrace{f^{-1}(y_k)}_{= x_k})$, $(y_{k-1} = f(x_{k-1}), \underbrace{f^{-1}(y_{k-1})}_{= x_{k-1}})$, $(y_{k-2}, \underbrace{f^{-1}(y_{k-2})}_{= x_{k-2}})$

Use a parabola $z(y) \approx f^{-1}(y)$. Evaluate $z(0)$.

Inverse quadratic interpolation


c


? ?

Achieving Global Convergence

The linear approximations in Newton and Secant are only good locally. How could we use that?

"trust region"

Achieving Global Convergence

The linear approximations in Newton and Secant are only good locally. How could we use that?

- Hybrid methods: bisection + Newton
  - Stop if Newton leaves bracket
- Fix a region where they're 'trustworthy' (trust region methods)
- Limit step size
- Sufficient conditions for convergence of Newton (under *strong* assumptions) are available.

Fixed Point Iteration

$$\begin{aligned} \boldsymbol{x}_0 &= \langle\text{starting guess}\rangle \\ \boldsymbol{x}_{k+1} &= \boldsymbol{g}(\boldsymbol{x}_k) \end{aligned}$$

$\vec{x}^*$ = fixed point

When does this converge?

$$\vec{e}_{k+1} = \vec{x}_{k+1} - \vec{x}^* = \vec{g}(\vec{x}_k) - g(\vec{x}^*)$$

$$= J_{\vec{g}}(\vec{x}^*) \cdot \vec{e}_k + O(\|e_k\|^2)$$

$\|e_k\| \to 0$

$$\vec{f}(\underbrace{\vec{x}}_{\vec{x}^*} + \underbrace{\vec{h}}_{\vec{e}_k}) = \vec{f}(\vec{x}) + J_{\vec{f}}(\vec{x}) \cdot \vec{h} + O(\|\vec{h}\|^2)$$

$$J_f = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_n & \cdots & \partial_n f_n \end{pmatrix}$$

$$g(\vec{x}^* + e_n) = g(\vec{x}^*) + \partial g(\vec{x}^*) \cdot \vec{e}_n + O(\|e_n\|^2)$$

$$g(\vec{x}_n) - g(x^*) = \partial g(\vec{x}^*)\, \vec{e}_n + O(\|e_n\|^2)$$

$$\|e_{n+1}\| \text{ "="} \|\partial g(\vec{x}^*)\, \vec{e}_n\|$$

$$\text{"}\le\text{"} \|\partial g(\vec{x}^*)\|\, \|e_n\|$$

$$\|\partial g(\vec{x}^*)\| < 1 \Rightarrow \text{implies linear conv.}$$

For all matrices $A \in \mathbb{C}^{n \times n}$

For all $\varepsilon > 0$

there exists a matrix norm $\|\cdot\|_A$

so that

$$\rho(A) \leq \|A\|_A \leq \rho(A) + \varepsilon$$

So, a sharper criterion for FPI

conv. is $\rho(J_g(x^*)) < 1$.

Newton's Method

What does Newton's method look like in n dimensions?

$$\vec{f}(\vec{x}_k + \vec{h}) \approx \vec{f}(\vec{x}_k) + J_{\vec{f}}(\vec{x}_k) \cdot \vec{h} = \vec{0}$$

$$\boxed{\vec{x}_{k+1} = \vec{x}_k - J_f(\vec{x}_k)^{-1} f(\vec{x}_k)} \qquad J_{\vec{f}}(\vec{x}_k) \cdot \vec{h} = -\vec{f}(\vec{x}_k) \quad \Big| \, J_f(\vec{x}_k)^{-1}$$

$$\vec{h} = -J_{\vec{f}}^{-1}(x_k)\, \vec{f}(x_k)$$

Downsides of n-dim. Newton?

- need J_f
- locally conv.

Demo: Newton's method in n dimensions [cleared]

Secant in n dimensions?

What would the secant method look like in n dimensions?

$$\otimes\ \hat{J}\,\overbrace{(\vec{x}_{k+1}-\vec{x}_k)}^{\vec{h}} = \vec{f}(\vec{x}_{k+1}) - \vec{f}(\vec{x}_k)$$

n^2 unknowns

n equations

start with approx. Jacobian J_0

"Broyden's method"

- update J_n, $\otimes$ minimize $\|J_n - J_{n-1}\|_F$
- update J_n^{-1} (use Sherman-Morrison)

 minimize $\|J_n^{-1} - J_{n-1}^{-1}\|_F$ $\Big\{$ good Broyden / bad Broyden

Numerically Testing Derivatives

Getting derivatives right is important. How can I test/debug them?