5 Nonlinear equations

Develop machinery to solve $f(\vec{x}) = \vec{0}$ ← find $\vec{x}$.

$$f: \mathbb{R}^n \to \mathbb{R}^m$$

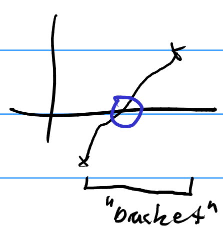
What if we're really looking for $\tilde{f}(x) = y$?

Intuition: Each of the m equations describes a surface.

Demo

Existence? Not easy — many different ways.

- Intermediate value theorem (1D-only, unhelpful in nD)

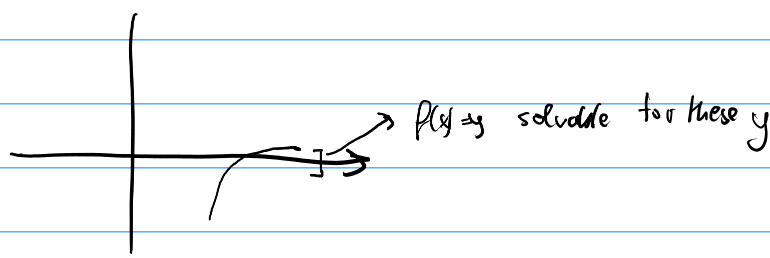


f continuous

- Inverse function theorem

LEFT OUT

$f \in C^1$, have solution $f(x^*) = 0$, $J_f(x)$ invertible ($\Rightarrow$ no extremum)
$\Rightarrow$ solutions also exist for $f(x) = y$ with $\|y - 0\| < \varepsilon$.



- Contraction mapping thm.

$g: \mathbb{R}^n \to \mathbb{R}^n$ contractive : there exists a $0<\gamma<1$ s.t.

$$\| g(x)-g(y) \| \leq \gamma \| x-y \|$$

Fixed point of g: $g(x)=x$

Real-world example: map

Closed set $S \subseteq \mathbb{R}^n$, $g(S) \subseteq S$
$\Rightarrow$ there exists a unique fixed point.

Uniqueness?

o [Hard to say anything general

No (usually)

Sensitivity

cond (root finding) = cond (evaluation of inverse)

Multiple roots (1D)

$$\left.\begin{aligned} &f(x)=0 \\ &f'(x)=0 \\ &\vdots \\ &f^{(m-1)}(x)=0 \\ &f^{(m)}(x)\neq 0 \end{aligned}\right\} \text{ "multiplicity } m\text{"}$$

o [Why are these a headache?

$\rightarrow$ conditioning of root finding

WS16 pb1

5.2 Convergence rates of iterative procedures

Let $e_k = \hat{u}_k - u$ be the error in the kth iterate $\hat{u}_k$.

Seen convergence behavior: $\|e_{k+1}\| \leq \underset{<1}{C} \cdot \|e_k\|$

"Gain constant # of digits each step"

Faster convergence $\|e_{k+1}\| \leq C \cdot \|e_k\|^2$

"Double # of digits each step"

An iterative method converges with rate r if

$$\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|^r} = C \begin{smallmatrix} <\infty \\ >0 \end{smallmatrix}$$

$r=1$ → "linear"

$r=2$ → "quadratic"

$r=3$ → 'cubic'

$\left.\begin{matrix} r=2 \\ r=3 \end{matrix}\right\} >1$: superlinear

WS16p2a

<Lec 18>

Recap: conv. rates

$$\|e_{k+1}\| \leq C \cdot \|e_k\|^r \longrightarrow$$ rate matters when $\|e_k\|$ is small → 0

① ⇒ $\dfrac{\|e_{k+1}\|}{\|e_k\|^r} \leq C$ ② > 0 (but that doesn't mean anything!)

③ $0 < C_{low} \leq \dfrac{\|e_{k+1}\|}{\|e_k\|^r} \leq C_{high}$ > 0

④ $\lim_{k\to\infty} \dfrac{\|e_{k+1}\|}{\|e_k\|^r} = C > 0$

5.3 Stopping criteria

- $\|f(x)\| < tol$ "residual"
- $\|x_{k+1} - x_k\| < tol$
- $\|x_{k+1} - x_k\| / \|x_k\| < tol$.

[Why is none of these foolproof?

5.3 Methods in 1D

Bisection

Demos

O [Rate of convergence? Constant?

Fixed point iteration

$x_0 = $ <some starting value>

$x_{k+1} = g(x_k)$

Demo

When does FPI converge?

Assume:

- g smooth
- $|g'(x^*)| < 1$ at fixed point $x^* = g(x^*)$

Error:

$$e_{k+1} = x_{k+1} - x^* = g(x_k) - g(x^*)$$

Mean Value Theorem: $g(x_k) - g(x^*) = g'(\theta_k)(x_k - x^*)$

$$= \underbrace{g'(\theta_k)}_{1.1 \leq c < 1} e_k$$

in region around x^*

→ assume entire iteration in that region

$\Rightarrow \quad |e_{k+1}| \leq C|e_k| \leq C^2 |e_{k-1}| \leq \cdots \leq C^k |e_0|$

$\Rightarrow \quad |e_k| \to 0$ linearly with rate C.

If $g'(x^*) = 0$, then convergence is faster than linear.

$$g(x_k) - g(x^*) = g''(\xi_k)\,(x_k - x^*)^2/2 \qquad \text{(Taylor)}$$

$\Rightarrow$ quadratic convergence!

WS17 p1

Would like to systematically find FPI w/ $g'(x^*)=0$.

Newton's method

Idea:

- Approximate f at current iterate.

$$f(x_k+h) \approx \underbrace{f(x_k) + f'(x_k)\,h}$$

- Find root of

$$f(x) + f'(x)\,h = 0 \iff h = -\frac{f(x_k)}{f'(x_k)}$$

x_0 = initial guess

$$x_{k+1} = x_k + h = x_k - \frac{f(x_k)}{f'(x_k)}$$

Can view Newton as FPI:

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

$$\underbrace{f(x^*) = 0, \; f'(x^*) \neq 0}_{\text{What does this mean?}} \Rightarrow \underbrace{g'(x^*) = 0}_{\text{What does this imply?}}$$

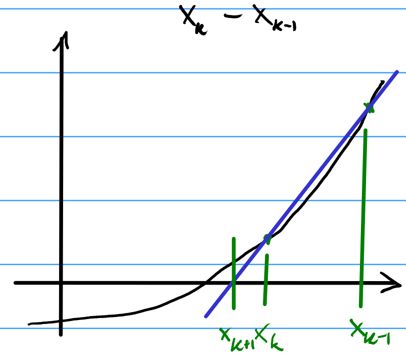
Drawback? → Have to have derivative!
→ locally convergent

Secant method

exam grades
go over exam on Wed
hw3 due next Wed.

<lec20> Use approximate derivative in Newton:

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$



Rate of convergence: $(1+\sqrt{5})/2 \approx 1.618$

Drawbacks? → slower conv.
→ two starting guesses
→ locally convergent

Demo: Newton and secant

Where should you "trust" a Newton/secant approximation?
→ limit steps to that region
→ "trust region" methods

WS18p1

5.4 Methods in nD ("Systems of equations")

Fixed point iteration

$x_0 = \langle \text{starting guess} \rangle$

$x_{k+1} = g(x_k)$

Converges (locally) if $\rho(J_g(x^*)) < 1$.

Jacobian matrix: $J_g(x) = \begin{pmatrix} \end{pmatrix}$ Also: $\nabla g = g' = J_g$

$J_g(x^*) = 0 \Rightarrow$ at least quadratic convergence.

Newton

$$f(x+s) = f(x) + J_f(x)s \overset{!}{=} 0$$

$$\Rightarrow \quad J_f(x)s = -f(x)$$

$$s = -(J_f(x))^{-1} f(x)$$

$$\left[\begin{array}{l} x_0 = \langle \text{starting guess} \rangle \\ x_{k+1} = x_k - (J_f(x_k))^{-1} f(x_k) \end{array}\right.$$

Downsides: → locally convergent

→ computing / inverting J_f expensive

Fix: Use approximate Jacobian matrix satisfying

$$\tilde{J}(x_{k+1} - x_k) = f(x_{k+1}) - f(x_k)$$

How many equations?

How many unknowns?

One choice: Broyden's method (minimizes change to $\tilde{J}$)