Nonlinear Systems


Doug James (and Justin Solomon)
Part III: Nonlinear Problems

Not all numerical problems can be solved with \( \backslash \) in Matlab.
Question

Have we already seen a nonlinear problem?
Question

Have we already seen a nonlinear problem?

minimize $\|A\vec{x}\|_2$

such that $\|\vec{x}\|_2 = 1 \leftarrow$ nonlinear!
Root-Finding Problem

Given: $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$

Find: $\vec{x}^* \text{ with } f(\vec{x}^*) = \vec{0}$
Issue: Regularizing Assumptions

\[ f(x) = \begin{cases} 
-1 & \text{when } x \leq 0 \\
1 & \text{when } x > 0 
\end{cases} \]
Issue: Regularizing Assumptions

\[ f(x) = \begin{cases} 
-1 & \text{when } x \leq 0 \\
1 & \text{when } x > 0 
\end{cases} \]

\[ g(x) = \begin{cases} 
-1 & \text{when } x \in \mathbb{Q} \\
1 & \text{when } x \notin \mathbb{Q} 
\end{cases} \]
Typical Regularizing Assumptions

**Continuous**

\[ f(\vec{x}) \rightarrow f(\vec{y}) \text{ as } \vec{x} \rightarrow \vec{y} \]
Typical Regularizing Assumptions

**Continuous**

\[ f(\vec{x}) \rightarrow f(\vec{y}) \text{ as } \vec{x} \rightarrow \vec{y} \]

**Lipschitz**

\[ \| f(\vec{x}) - f(\vec{y}) \|_2 \leq c \| \vec{x} - \vec{y} \|_2 \text{ for all } \vec{x}, \vec{y} \text{ (same } c) \]
Typical Regularizing Assumptions

### Continuous

$$f(\vec{x}) \rightarrow f(\vec{y}) \text{ as } \vec{x} \rightarrow \vec{y}$$

### Lipschitz

$$\|f(\vec{x}) - f(\vec{y})\|_2 \leq c\|\vec{x} - \vec{y}\|_2 \text{ for all } \vec{x}, \vec{y} \text{ (same } c)$$

### Differentiable

$$Df(\vec{x}) \text{ exists for all } \vec{x}$$
Typical Regularizing Assumptions

Continuous

\[ f(\vec{x}) \rightarrow f(\vec{y}) \text{ as } \vec{x} \rightarrow \vec{y} \]

Lipschitz

\[ \| f(\vec{x}) - f(\vec{y}) \|_2 \leq c \| \vec{x} - \vec{y} \|_2 \text{ for all } \vec{x}, \vec{y} \text{ (same } c) \]

Differentiable

\[ Df(\vec{x}) \text{ exists for all } \vec{x} \]

\[ C^k \]

\[ k \text{ derivatives exist and are continuous} \]
Today

\[ f : \mathbb{R} \rightarrow \mathbb{R} \]
Property of Continuous Functions

Intermediate Value Theorem

Suppose that \( f : [a, b] \rightarrow \mathbb{R} \) is continuous and that \( f(a) < u < f(b) \) or \( f(b) < u < f(a) \). Then, there exists \( z \in (a, b) \) such that \( f(z) = u \).
Reasonable Input

- Continuous function $f(x)$

- $\ell, r \in \mathbb{R}$ with $f(\ell) \cdot f(r) < 0$ (why?)
Bisection Algorithm

1. Compute $c = \ell + r/2$.
2. If $f(c) = 0$, return $x^* = c$.
3. If $f(\ell) \cdot f(c) < 0$, take $r \leftarrow c$. Otherwise take $\ell \leftarrow c$.
4. Return to step 1 until $|r - \ell| < \varepsilon$; then return $c$. 
Nonlinearity
Root-finding
Bisection
Fixed Point Iteration
Newton's Method
Secant Method
Conclusion

Bisection: Illustration

\[ f(x) > 0 \]

\[ f(x) < 0 \]
Two Important Questions

1. Does it converge?
Two Important Questions

1. Does it converge?
   Yes! Unconditionally.
Two Important Questions

1. Does it converge?
   Yes! Unconditionally.

2. How quickly?
Examine $E_k$ with

$$|x_k - x^*| < E_k.$$
Bisection: Linear Convergence

\[ E_{k+1} \leq \frac{1}{2} E_k \]

for \( E_k \equiv |r_k - \ell_k| \)
Fixed Points

\[ g(x^*) = x^* \]
Fixed Points

$$g(x^*) = x^*$$

Question:
Same as root-finding?
Simple Strategy

\[ x_{k+1} = g(x_k) \]
Convergence Criterion

\[ E_k \equiv \left| x_k - x^* \right| \]

\[ = \left| g(x_{k-1}) - g(x^*) \right| \]
**Convergence Criterion**

\[ E_k \equiv |x_k - x^*| \]
\[ = |g(x_{k-1}) - g(x^*)| \]
\[ \leq c|x_{k-1} - x^*| \]
\[ \text{if } g \text{ is Lipschitz} \]
\[ = cE_{k-1} \]
Convergence Criterion

\[ E_k \equiv |x_k - x^*| \]
\[ = |g(x_{k-1}) - g(x^*)| \]
\[ \leq c|x_{k-1} - x^*| \]

if \( g \) is Lipschitz

\[ = cE_{k-1} \]

\[ \implies E_k \leq c^k E_0 \]

\[ \to 0 \text{ as } k \to \infty \quad (c < 1) \]
Alternative Criterion

Lipschitz near $x^*$ with good starting point.
Alternative Criterion

Lipschitz near $x^*$ with good starting point.

e.g. $C^1$ with $|g'(x^*)| < 1$
Convergence Rate of Fixed Point

When it converges…
Always linear (why?)
Convergence Rate of Fixed Point

When it converges…
Always linear (why?)

Often quadratic! (→ board)
Approach for Differentiable \( f(x) \)
Newton’s Method

\[ x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \]
Newton’s Method

\[ x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \]

Fixed point iteration on

\[ g(x) \equiv x - \frac{f(x)}{f'(x)} \]
Convergence of Newton

Simple Root

A root $x^*$ with $f'(x^*) \neq 0$. 

Quadratic convergence in this case! (→ board)
Convergence of Newton

**Simple Root**

A root $x^*$ with $f'(x^*) \neq 0$.

Quadratic convergence in this case! ($\rightarrow$ board)
Issue

Differentiation is hard!
Secant Method

\[ x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})} \]
Secant Method

\[ x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})} \]

Trivia:
Converges at rate \( \frac{1+\sqrt{5}}{2} \approx 1.6180339887 \ldots \)
Secant Method

\[ x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})} \]

Trivia:
Converges at rate \( \frac{1 + \sqrt{5}}{2} \approx 1.6180339887 \ldots \)
(“Golden Ratio”)
Hybrid Methods

**Want:** Convergence rate of secant/Newton with convergence guarantees of bisection
Hybrid Methods

**Want:** Convergence rate of secant/Newton with convergence guarantees of bisection

e.g. **Dekker’s Method:** Take secant step if it is in the bracket, bisection step otherwise
Single-Variable Conclusion

- Unlikely to solve exactly, so we settle for iterative methods
- Must check that method converges at all
- Convergence rates:
  - Linear: $E_{k+1} \leq C E_k$ for some $0 \leq C < 1$
  - Superlinear: $E_{k+1} \leq C E_k^r$ for some $r > 1$
  - Quadratic: $r = 2$
  - Cubic: $r = 3$
- Time per iteration also important