CS450: Numerical Analysis
Nonlinear Equation Systems

Howard Hao Yang

Assistant Professor, ZJU-UIUC Institute

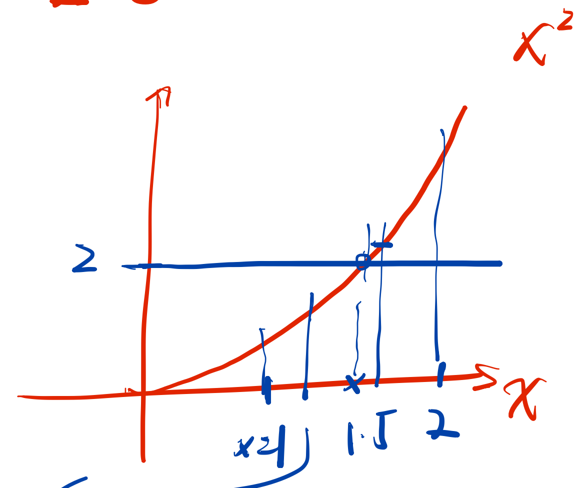
24/11/2023

ZJUI Interview Question

- How to quickly evaluate $\sqrt{2}$?

① let $x = \sqrt{2}$

② solve $x^2 - 2 = 0$



1.25

$x^2|_{x=1} = 1 < 2 = x^2$

$x^2|_{x=2} = 4 > 2 = x^2$

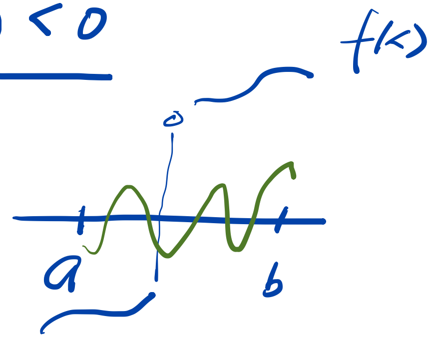
$[1, 2] \to (1.5)^2 > 2$

$[1, 1.5] \to (1.25)^2 < 2$

Solving Nonlinear Equation (1)

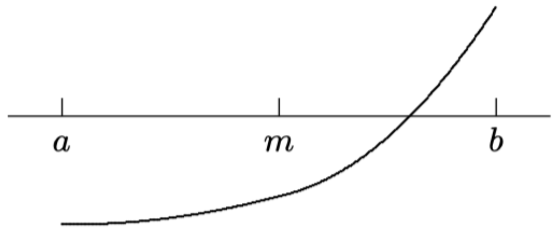
- Given a continuous function $f: R \to R$, how to find x such that $f(x) = 0$?
- Interval Bisection:
  1) start with an interval $[a, b]$ in which f changes sign
  2) cut off half of the interval, till reaches a certain level

$f(a) \cdot f(b) < 0$



Algorithm 5.1 Interval Bisection

while $((b - a) > tol)$ **do**
    $m = a + (b - a)/2$
    if $\text{sign}(f(a)) = \text{sign}(f(m))$ **then**
        $a = m$
    else
        $b = m$
    end
end



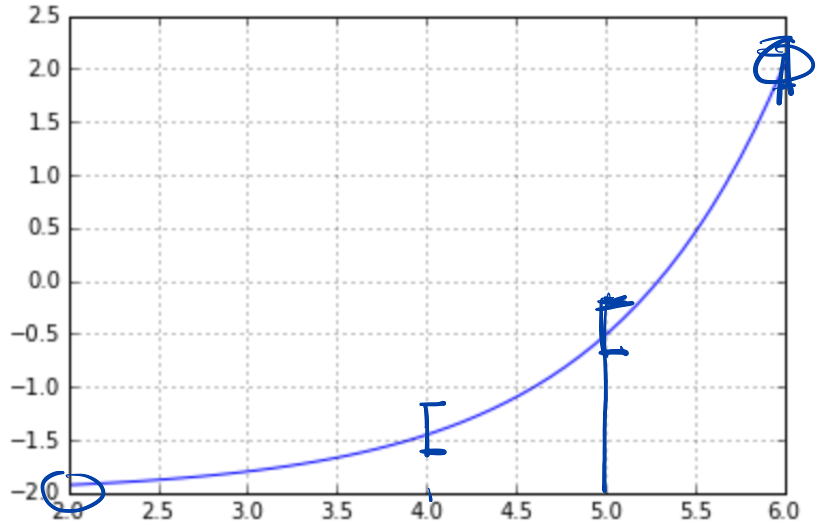
Illustrative Example

$$\frac{e^x}{100} - 2 = 0$$

```
In [1]:  import numpy as np
         import matplotlib.pyplot as pt
```

```
In [2]:  a = 2
         b = 6

         x = np.linspace(a, b)

         def f(x):
             return 1e-2 * np.exp(x) - 2

         pt.grid()
         pt.plot(x, f(x))
```

```
In [25]: #clear
         m = (a+b)/2

         if np.sign(f(a)) == np.sign(f(m)):
             a = m
         else:
             b = m

         print(a, b)

5.298316955566406 5.2983174324035645
```



Interval Bisection: When to Stop?

- Given a continuous function $f: R \to R$, how to find x such that $f(x) = 0$?
- Interval Bisection:
  1) start with an interval $[a, b]$ in which f changes sign
  2) cut off half of the interval, till reaches a certain level
- Different criteria for stopping, define a small ϵ (e.g. $\epsilon = 10^{-3}$)
  1) stop if $|x_n - x_{n-1}| < \epsilon$
  2) stop if $\frac{|x_n - x_{n-1}|}{|x_n|} < \epsilon$
  3) stop if $f(x_n) < \epsilon$

Exercise (1)

- What is the potential issue of stopping at $\frac{|x_n - x_{n-1}|}{\cancel{|x_n|}} < \epsilon$

- Let $x_n = \sum_{k=1}^{n} \frac{1}{k}$, show that $\{x_n\}$ diverges even though $\lim_{n\to\infty} |x_n - x_{n-1}| = 0$

$|x_n - x_{n-1}| = \frac{1}{n} \to 0. \quad n \to \infty$

$x_n \to \infty$

Exercise (2)

- What is the potential issue of stopping at $f(x_n) < \epsilon = 10^{-3}$
- Let $f(x) = (x-1)^{10}$, $x = 1$, and $x_n = 1 + \frac{1}{n}$, show that $|f(x_n)| < 10^{-3}$ whenever $n > 1$ but $|x_n - x| < 10^{-3}$ requires $n > 1000$

$x_2 = 1 + \frac{1}{2}$, $\quad |f(x_2)| = (1 + \frac{1}{2} - 1)^{10} = \frac{1}{2^{10}} = \frac{1}{1024} < 10^{-3}$

$|x_n - x| < 10^{-3} \rightarrow 1 + \frac{1}{n} - 1 < 10^{-3}$

$\Rightarrow n > 1000$

Stopping Criteria

- Possible candidates for stopping the iteration
  - Run until $|f(x)| < \epsilon$, i.e., the residual is small
  - Run until $\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\| < \epsilon$
  - Run until $\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\| / \|\boldsymbol{x}_k\| < \epsilon$
- None of them is bulletproof, depends on the application

$\underline{k > K}$

Interval Bisection

$\frac{b-a}{2^n}$

- Cons: it converges slowly
- Pros: it must converge!
- **Theorem**: Suppose that $f(x)$ is continuous in the interval $[a, b]$ and $f(a) \cdot f(b) < 0$. The Bisection method generates a sequence $\{x_n\}$ approximating a root x of $f(x) = 0$ with

$$|x_n - x| \leq \frac{b-a}{2^n}$$

Solving Nonlinear Equation (2)

x_0
$x_1 = f(x_0)$
$x_2 = f(x_1)$
$x_3 = f(x_2)$
$\vdots$

$x^2 - 2 = 0$

$f(x) = 0 \rightarrow f(x) = x + 2f(x) = x$

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

- Given a continuous function $f: R \rightarrow R$, how to find x such that $f(x) = 0$?

$x = f(x)$

- Fixed Point Iteration:

find a f. such that ?

$x + f(x) = x$

$f(x)$

$x - 10f(x) = x$

$x + x^3 f(x) = x$

1) transform the equation into $g(x) = x$
2) start with an initial guess x_0
3) keep iterating $x_{k+1} = g(x_k)$

$x_1 = f(x_0)$

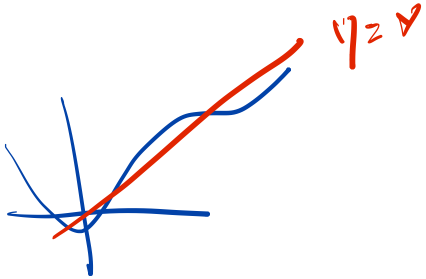
$x_2 = f(x_1)$ ...

- Definition: The number x is a fixed-point for a given function g if $g(x) = x$

- Exercise: Determine any fixed-point of the function $g(x) = x^2 - 2$

$f(x) = x \Rightarrow x^2 - 2 = x$

$\Rightarrow x^2 - x - 2 = 0 \rightarrow x = 2$ or -1



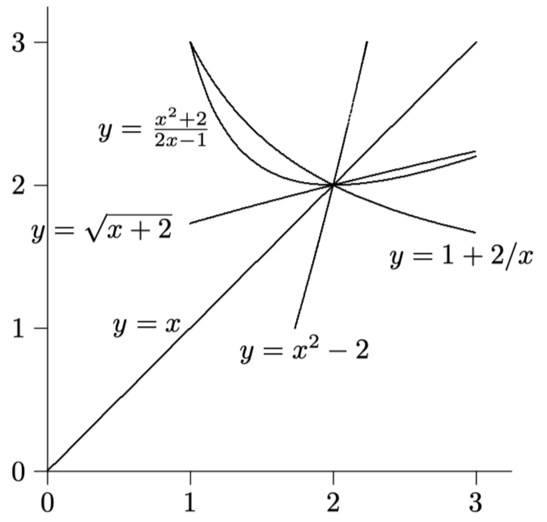
Illustrative Example

- Consider the following nonlinear equation

$$f(x) = x^2 - x - 2 = 0$$

- We want to transform this equation into the form $g(x) = x$
- Equivalent fixed-point problems include

1) $x^2 - 2 = x \quad \Rightarrow \quad g(x) = x^2 - 2$
2) $x^2 = x + 2 \quad \Rightarrow \quad x = \sqrt{x+2} \quad \Rightarrow \quad g(x) = \sqrt{x+2}$
3) $x^2 = x + 2 \quad \Rightarrow \quad x = 1 + \frac{2}{x} \quad \Rightarrow \quad g(x) = 1 + \frac{2}{x}$
4) $x^2 + 2 = 2x^2 - x \quad \Rightarrow \quad x = \frac{x^2+2}{2x-1} \quad \Rightarrow \quad g(x) = \frac{x^2+2}{2x-1}$



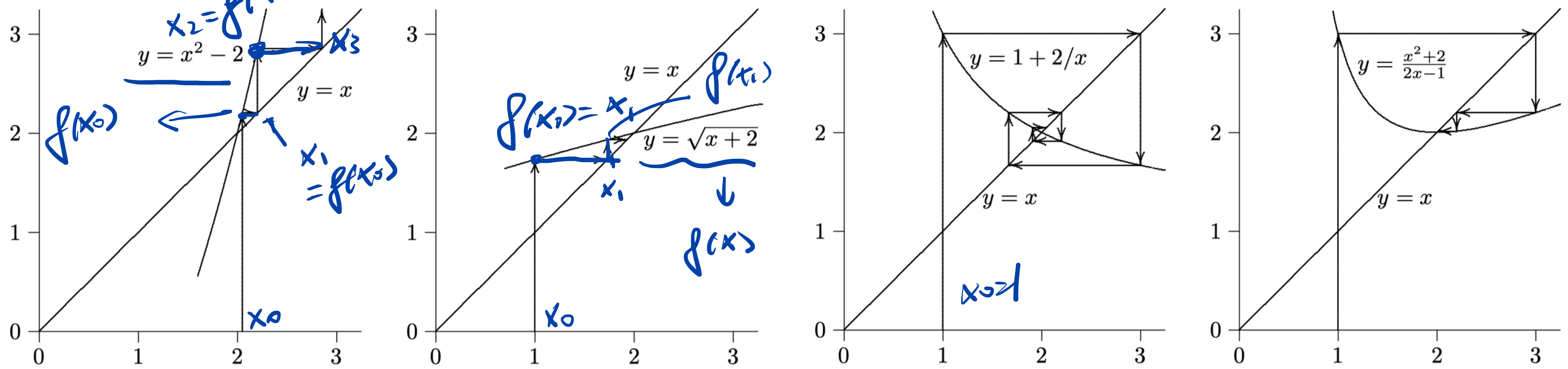
Illustrative Example (Cont'd)

- Consider the following nonlinear equation

$$f(x) = x^2 - x - 2 = 0$$

- Transform this equation into the form $g(x) = x$, and solve by fixed-point iteration—different forms converges (or diverges) differently

$x = x^2 - 2 = f(x_1)$ $x_{n+1} = f(x_n)$, $f(x) = x^2 - 2$



$f(x) = \sqrt{x+2}$, $x = f(x)$

Fixed-Point Iterations: Convergence?

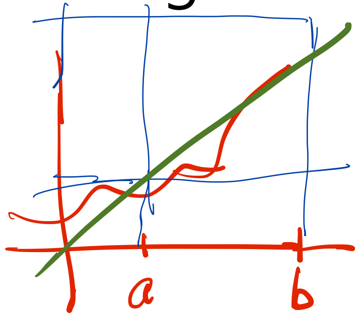
there must be a fixed point of f

$f'(x)$ exists

continuous in $[a,b)$

- **Fixed-Point Theorem**: Suppose that $g(x) \in C[a,b]$ and $a \le g(x) \le b$ for all $x \in [a,b]$. Suppose in addition that $g'(x)$ exists on (a,b) and a constant $0 < c < 1$ exists with $|g'(x)| \le c$ for all $x \in [a,b]$. Then, for any starting point $x_0 \in [a,b]$, the sequence defined by $x_{n+1} = g(x_n)$ converges to the unique fixed point of g.

$$|x_{n+1} - x_n| = |f(x_n) - f(x_{n-1})|$$
$$= |f'(\xi)| \cdot |x_n - x_{n-1}| \le C \cdot |x_n - x_{n-1}|$$



- **Corollary**: The error bounds for the error involved above is

$$|x_n - x| \le c^n \max\{x_0 - a, b - x_0\}$$

Recap: Sensitivity and Conditioning

- What does sensitivity/conditioning capture?
  - Amplification of input error after operation of the system
- Definition: for an input x and its perturbation $\hat{x}$, if the goal is to evaluate $\mathrm{y} = f(x)$
  - The smallest number κ such that

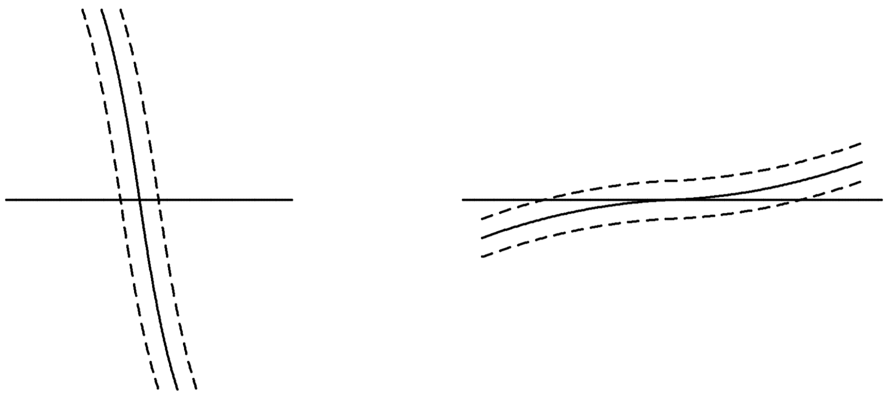
$$\text{(rel. perturbation in output)} \leq \kappa \cdot \text{(rel. perturbation in output)}$$

  - Mathematically, given an input x and perturbation $\hat{x}$

$$\frac{|f(x) - f(\hat{x})|}{|f(x)|} \leq \kappa \cdot \frac{|x - \hat{x}|}{|x|}$$

Sensitivity of Nonlinear Systems

- Comparison of sensitivity conditions between function evaluating (i.e., given an input x, what is the output $y = f(x)$) and root finding (i.e., given an output y, what is the input x such that $y = f(x)$)
  - If the function is insensitive to the value of the argument, then the root will be sensitive
  - If the function is sensitive to the argument, then the root will be insensitive



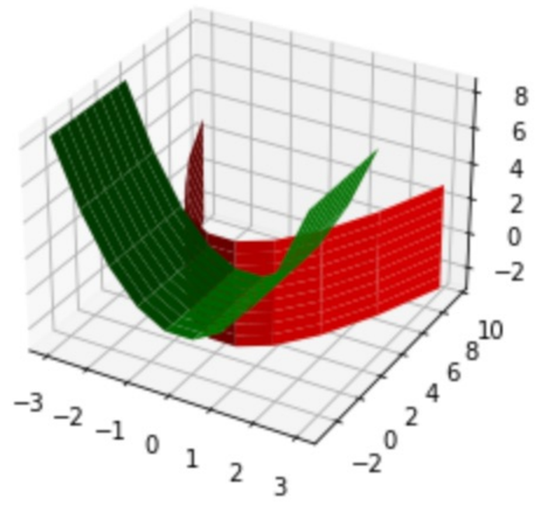
well-conditioned ill-conditioned

Recap: Linear Equation Systems

- Given an $m \times n$ matrix $\boldsymbol{A}$ and an m-dimensional vector $\boldsymbol{b}$, we aim to find n-dimensional vector $\boldsymbol{x}$ such that $\boldsymbol{Ax} = \boldsymbol{b}$
  - Linear combination of columns of $\boldsymbol{A}$ to yield $\boldsymbol{b}$
- Geometric interpretation: Line/plane intersection
- How many solutions might it have?
  - Depends on $\boldsymbol{A}$
  - But always three possible cases: zero, one, or infinity

Nonlinear Equation Systems

- Given a mapping $\boldsymbol{f}: R^n \rightarrow R^n$, we aim to find n-dimensional vector $\boldsymbol{x}$ such that $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$
  - If we look for solution to $\tilde{\boldsymbol{f}}(\boldsymbol{x}) = \boldsymbol{y}$, simply consider $\boldsymbol{f}(\boldsymbol{x}) = \tilde{\boldsymbol{f}}(\boldsymbol{x}) - \boldsymbol{y}$
- Geometric interpretation: Curve intersection
- How many solutions might it have?
  - Depends on the equations
  - Can be any possible integers

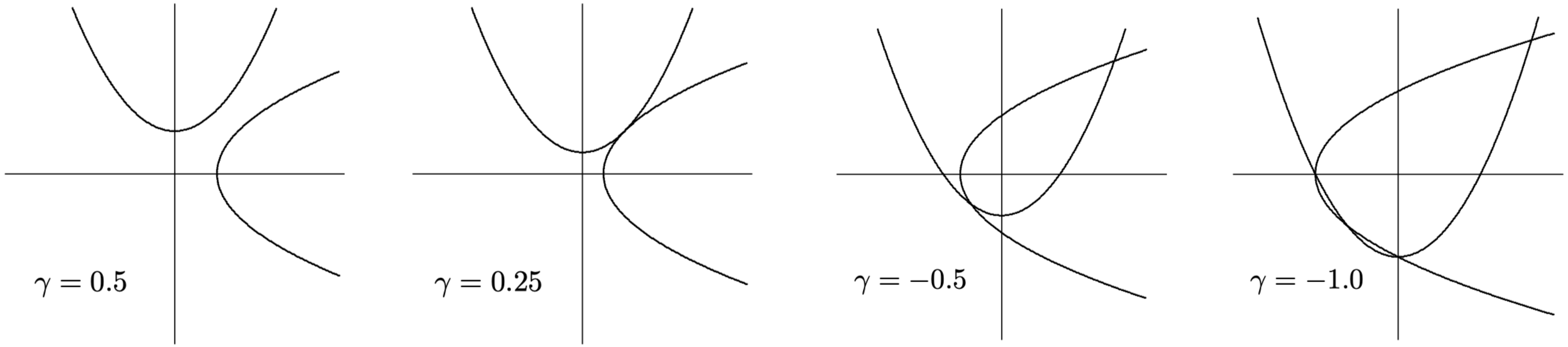


Illustrative Example

- Consider the system of equations in two dimensions

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} x_1^2 - x_2 + \gamma \\ -x_1 + x_2^2 + \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- The number of solutions depends on the particular value of γ



Existence (1)

- Given $\boldsymbol{f}$, when would there exist a variable $\boldsymbol{x}$ that solves $\boldsymbol{f(x) = 0}$?
- 1-Dimension case
  - (**Intermediate Value Theorem**) if $f: R \to R$ is continuous on $[a, b]$, and c lies between $f(a)$ and $f(b)$, then there is a value $x^* \in [a, b]$, such that $f(x^*) = c$
  - Q: Is continuity necessary? Is closed interval necessary?
  - How can we use this to determine the existence of a root?
  - Can we determine the number of roots?
  - In general, how to find such an interval?

Existence (2)

- Given $\boldsymbol{f}$, when would there exist a variable $\boldsymbol{x}$ that solves $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$?
- n-Dimension case
  - (**Inverse Function Theorem**) if $\boldsymbol{f}: R^n \rightarrow R^n$ is continuously differentiable, if the Jacobian matrix $\{\boldsymbol{J}_f(\boldsymbol{x})\}_{ij} = \boldsymbol{\partial f}_i(\boldsymbol{x})/\boldsymbol{\partial x}_j$ is nonsingular at a point $\boldsymbol{x}^*$. Then, there is a neighborhood of $\boldsymbol{f}(\boldsymbol{x}^*)$ in which $\boldsymbol{f}^{-1}$ exists. That is, $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}$ has a solution for any $\boldsymbol{y}$ in that neighborhood of $\boldsymbol{f}(\boldsymbol{x}^*)$.
  - Issue: only local, not global. Besides, relies on calculating the Jacobian matrix.

Existence (3)

- Given $\boldsymbol{f}$, when would there exist a variable $\boldsymbol{x}$ that solves $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$?
- n-Dimension case
  - (**Contraction Mapping Theorem**) a function $\boldsymbol{g}: R^n \to R^n$ is called contractive if there exists $0 < \gamma < 1$ such that $||\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})|| \leq \gamma||\boldsymbol{x} - \boldsymbol{y}||$. A fixed point of $\boldsymbol{g}$ is a point where $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x}$
  - On a closed set $S \subset R^n$ with $\boldsymbol{g}(S) \subset S$ there exists a unique fixed point (why?)
  - Example: real-world map

Rate of Convergence

- Consider an iterative algorithm to solve $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$, where $\boldsymbol{e}_k = \hat{\boldsymbol{x}}_k - \boldsymbol{x}^*$ is the error in the k-th iteration. Assume $\boldsymbol{e}_k \to 0$ as $k \to \infty$, an iterative algorithm converges with rate r if

$$\lim_{k\to\infty} \frac{\|\boldsymbol{e}_{k+1}\|}{\|\boldsymbol{e}_k\|^r} = C \begin{cases} > 0 \\ < \infty \end{cases}$$

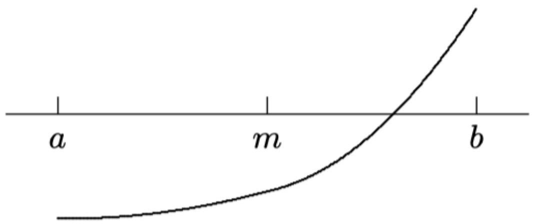
  - If $r = 1$, it is called *linear convergence* (Example: Power Method)
  - If $r > 1$, it is called *superlinear convergence*
  - If $r = 2$, it is called *quadratic convergence* (Example: Rayleigh Quotient Iteration)

Example (1)

- What is the convergence rate of Bisection Interval?

Algorithm 5.1 Interval Bisection

while $((b-a) > tol)$ **do**
  $m = a + (b-a)/2$
  if $\text{sign}(f(a)) = \text{sign}(f(m))$ **then**
    $a = m$
  else
    $b = m$
  end
end



- Linear with constant $\frac{1}{2}$ (can you see this?)

Example (2)

- What is the convergence rate of Fixed Point Iteration?
- Mean value theorem says: There exists a $\boldsymbol{\theta}_k$ between $\boldsymbol{x}_k$ and $\boldsymbol{x}^*$ such that

$$g(\boldsymbol{x}_k) - g(\boldsymbol{x}^*) = g'(\boldsymbol{\theta}_k)(\boldsymbol{x}_k - \boldsymbol{x}^*)$$

- Because under FPI, $g(\boldsymbol{x}) = \boldsymbol{x}$. Therefore, the error decreases by

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

- If $|g'(\boldsymbol{\theta}_k)|$<1 around $\boldsymbol{x}^*$, then FPI converges at a linear rate
- Q: what if $g(\boldsymbol{x}^*) = \boldsymbol{0}$?
- By Taylor expansion, $\boldsymbol{e}_{k+1} = g''(\boldsymbol{\xi}_k)\|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2/2$ → quadratic convergence

Learning Objectives

- Nonlinear equation systems: Finding intersections of curves
- Existence: different criteria to judge
- Sensitivity: duality to the evaluation of a function
- Solving nonlinear equations: Bisection Interval and Fixed-Point Iteration
- Stopping criteria—no general bulletproof
- Convergence rate