数值分析 Numerical Analysis

Lecture 2：二分法 (Bisection Method) & 迭代法（Fixed-Point Iteration）

骞微著

苏州大学，计算机科学与技术学院

邮箱：wzqian@suda.edu.cn

办公室：理工楼543

Recap

1. 使用数值方法的动机
2. 误差的种类
3. 数值方法的相容性
4. 数值方法的稳定性

Insight

（一元）非线性方程

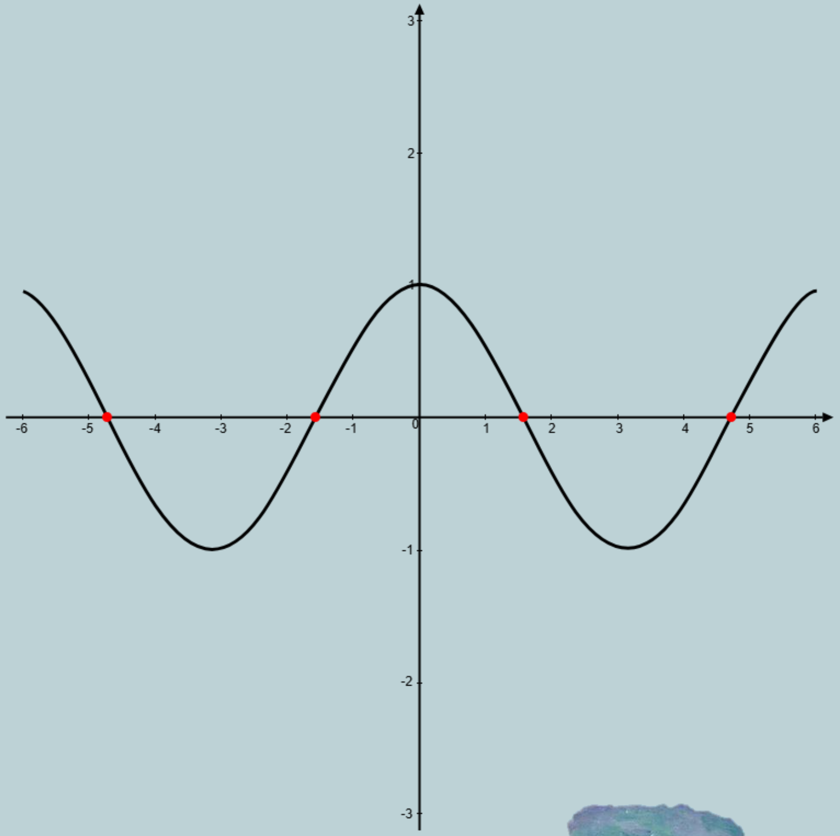
1）函数$f(x)$ 的零点/方程的解或根：Any value x^* s.t. (such that使得)

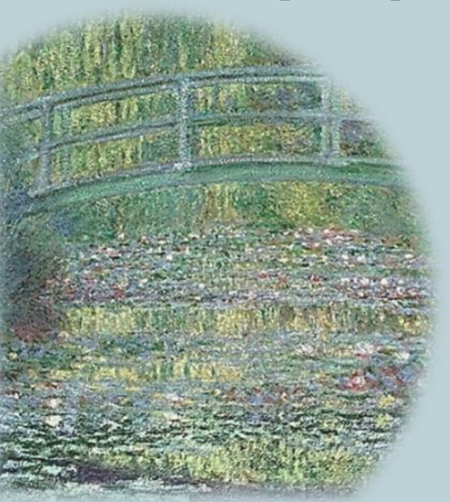
$$f(x^*) = 0$$

2）解的重数: 若 $f(x^*) = (x - x^*)^m g(x)$且 $g(x^*) \neq 0$,

则 x^*为 $f(x) = 0$的 m 重解 (根)。

3）有解区间: 若 [a, b] 内至少存在 $f(x) = 0$的一个**实数解**,
则称 [a, b] 为有解区间.





（一元）非线性方程

1）代数方程 algebraic /polynomial equation：

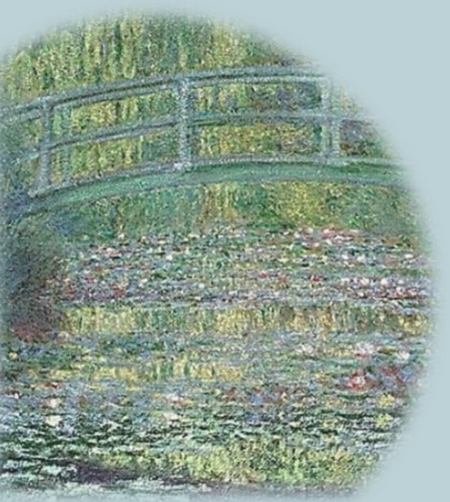
例子1：$Ax^2 + Bx + C = 0$ (looks familiar ☺?)

例子2：$Ax^{200} + Bx^{20} + C = 0$ (how about this?)

2）超越方程 transcendental equation(包含指数函数、对数函数、三角函数等):

例子：$\cos(x) - \ln(x+2) = 0$ (and this?)

一般来说，难以求出精确的解析表达式。



非线性方程的数值解法

数值解法：

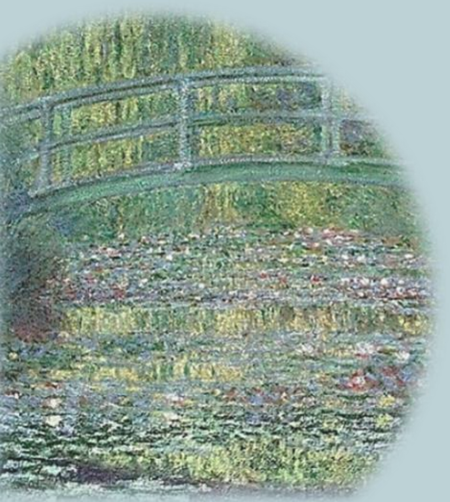
1. 二分法Bisection/Dichotomy method
2. 一般迭代法Iterative method
3. 牛顿法Newton’s method
4. 弦截法Secant method(a quasi Newton’s method)

非线性方程的数值解法

数值计算方法主要分为两大类。

1. 第一类是区间收缩法：
   （1）确定初始含根区间
   （2）收缩含根区间

2. 第二类是迭代法：
   （1）选定根的初始近似值
   （2）按某种原则生成收敛于根的近似点列



非线性方程的数值解法：二分法Bisection Method

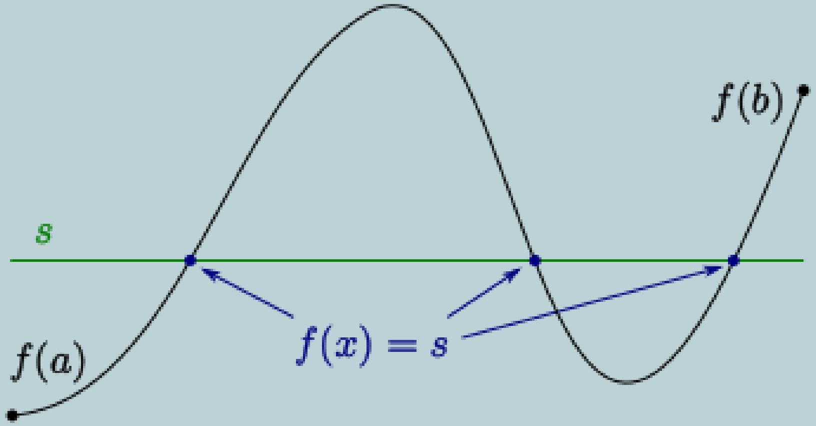
二分法的数学原理

介值定理intermediate value theorem：

在闭区间[a,b]上，$f(x)$连续且$f(\mathrm{a})f(b)<0$，则在 (a, b) 内至少存在一点 x^*, 使得 $f(x^*)=0$ 。

数学表达式：

$$f \in \mathrm{C}([a,b]), f(a)f(b)<0 \Rightarrow \exists x^* \in \mathrm{C}(a,b): f(x^*)=0$$



二分法的计算步骤

问题：求解$f(x)=0$。

基本假设：在闭区间[a,b]上，$f(x)$连续且$f(a)f(b)<0$。

二分法的计算步骤：

二分法的计算步骤是首先选取初始含根区$[a_0,b_0]=[a,b]$，计算区间中点$x_1=\frac{1}{2}(a_0+b_0)$，然后判断x^*是落在$[a_0,x_1]$还是$[x_1,b]$内来确定更小的含根区间$[a_1,b_1]$，其具体做法是：

$$
\text{若 } f(x_1)f(a_0)\begin{cases} >0 & \text{取}[a_1,b_1]=[x_1,b_0] \\ <0 & \text{取}[a_1,b_1]=[a_0,x_1] \\ =0 & \text{则 } x_1=x^*\text{，终止计算} \end{cases}
$$

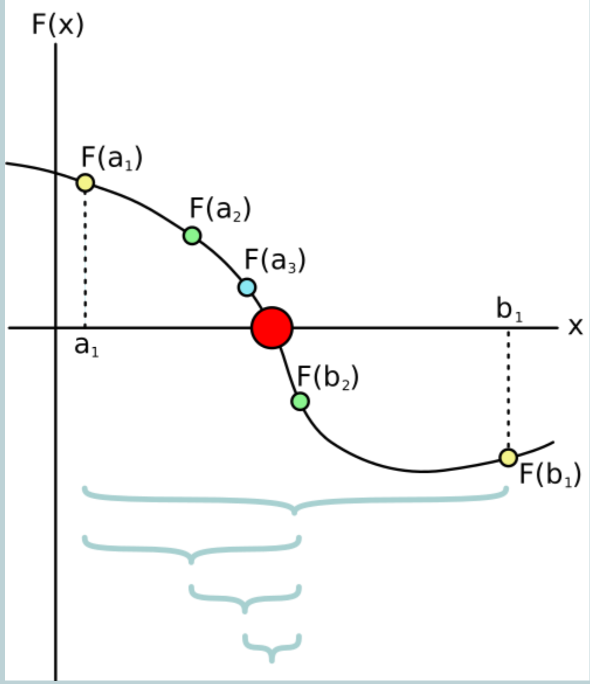
二分法的计算步骤

一般地，如果已计算得到含根区间$[a_k,b_k]$($k=1,2,\cdots$)，则令

$$x_{k+1}=\frac{1}{2}(a_k+b_k),$$

$$\text{若 } f(a_k)f(x_{k+1})\begin{cases}>0 & \text{取}[a_{k+1},b_{k+1}]=[x_{k+1},b_k]\\ <0 & \text{取}[a_{k+1},b_{k+1}]=[a_k,x_{k+1}]\\ =0 & \text{则 } x_{k+1}=x^*\text{，终止计算}\end{cases}$$

这样，除碰巧出现上面第三种情况而得到根x^*外，便构造出长度逐渐减半的含根区间序列$\{[a_k,b_k];\ k=1,2,\cdots\}$。



几何表示

当计算到 $b_k-a_k<2\varepsilon$ 时终止计算，取 $\tilde{x}=\frac{1}{2}(a_k+b_k)$。

二分法的收敛性和事前误差估计

因为 $b_k - a_k = \frac{1}{2}(b_{k-1} - a_{k-1}) = \cdots = \frac{1}{2^k}(b-a) \to 0(k \to \infty\text{时})$

进一步，由 $x_k = \frac{1}{2}(a_k + b_k)$，我们得

$|x_k - x^*| \le \frac{1}{2}(b_k - a_k) = \frac{1}{2^{k+1}}(b-a) \to 0(k \to \infty\text{时})$

所以，二分法**总是**收敛的

故对给定的精度要求 $|x_k - x^*| < \varepsilon$，

可要求 $\frac{1}{2^{k+1}}(b-a) < \varepsilon$

预先估计出所需迭代步数为 $K = \left[\frac{\lg\frac{b-a}{\varepsilon}}{\lg 2}\right]$。

推导

二分法Bisection Method

特别当 $\varepsilon = 10^{-m}$ 时，$K = \left[\dfrac{\lg(b-a)+m}{\lg 2}\right]$

例 2.1 试用二分法求$f(x) = x^3 - 2x - 5 = 0$的一个正根，使误差小于10^{-3}：

解 因为 $f(0) = -5, f(1) = -6, f(2) = -1, f(3) = 16,$

故可取初始区间$[a_0, b_0] = [2,3]$

这时 $K = \left[\dfrac{\lg(3-2)+3}{\lg 2}\right] = [9.97] = 9$

$\tilde{x} = \frac{1}{2}(b_9 + a_9)$ 即为所求近似值。

二分法的计算结果

表 2-1　二分法的计算结果

k	a_k	b_k	x_k	$f(x_k)$的符号
0	2	3	2.5	+
1	2	2.5	2.25	+
2	2	2.25	2.125	+
3	2	2.125	2.062 5	−
4	2.062 5	2.125	2.093 8	−
5	2.093 8	2.125	2.109 4	+
6	2.093 8	2.109 4	2.101 6	+
7	2.093 8	2.101 6	2.097 7	+
8	2.093 8	2.097 7	2.095 8	+
9	2.093 8	2.095 8	2.094 9	

所以 $\tilde{x}=2.095$。

二分法评价

优点：简单可靠，易于编程实现，它对函数要求低，适用于的奇数重根情形。

缺点：不能直接用于求偶重根，不能用于求复根，也难以向方程组推广使用，收敛速度慢。

非线性方程的数值解法: 迭代法Fixed-point Iteration

(不动点)迭代法的算法思想： 对 $f(x) = 0 \qquad (2-2)$

例：

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

$$\downarrow$$

$$x = g(x) = x - \frac{x^2 - 2}{2x} = \frac{1}{2}(x + \frac{2}{x})$$

Step 1) 把（2-2）等价变换为如下形式:

$$x = g(x) \qquad (2-2)'$$

从而 $x^* = g(x^*)$，x^* 称为**迭代函数**$g(x)$ 的**不动点Fixed Point**。

Step 2) 建立迭代格式**（单步迭代法）**：

$$x_{k+1} = g(x_k) \qquad (2-3)$$

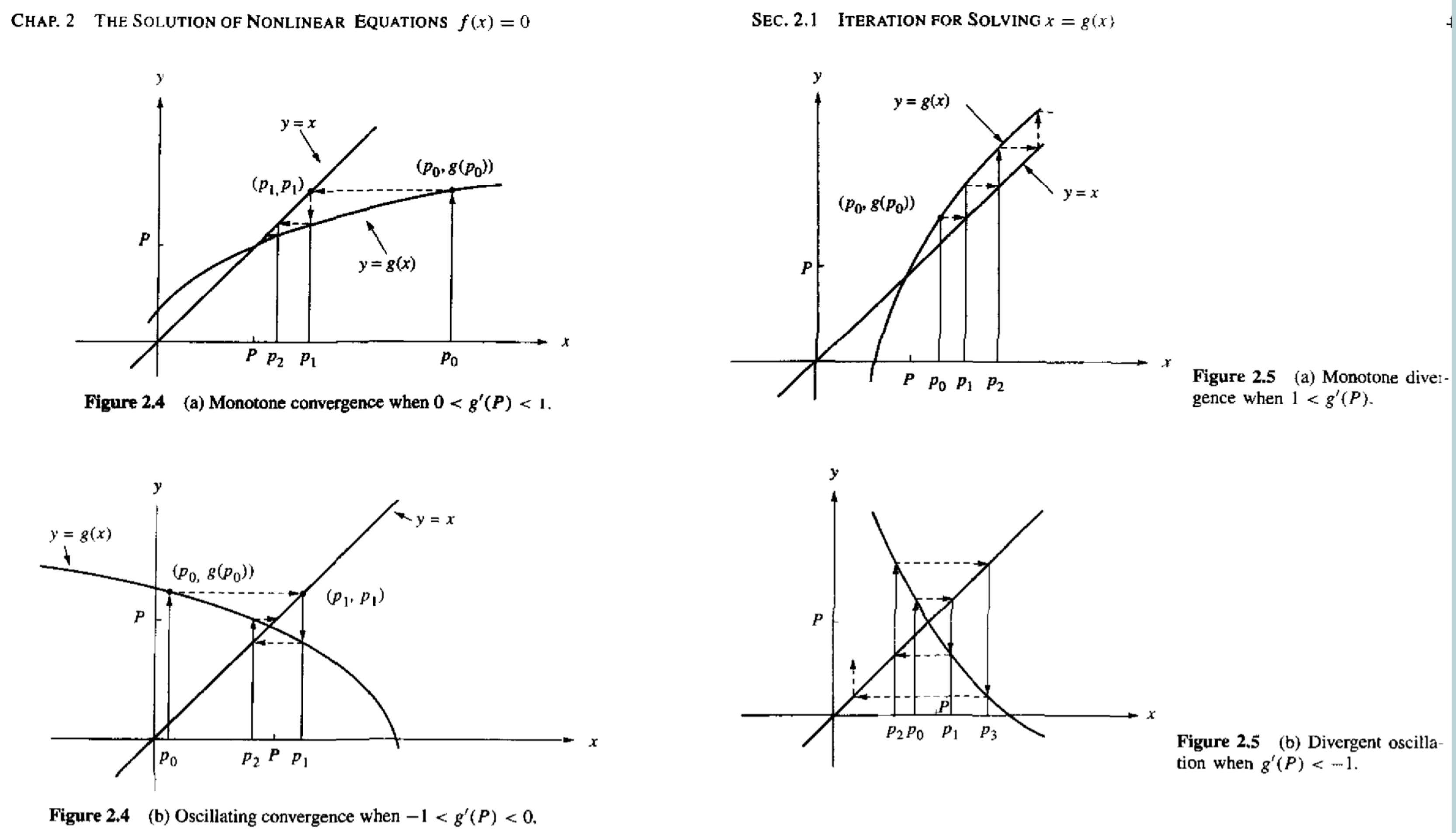
或更一般地建立迭代格式**（多步迭代法）**：

$$x_{k+1} = g(x_k, x_{k-1}, \cdots x_{k-m}) \ (m \geq 1) \qquad (2-3)'$$

Step 3) 适当选取初始值，递推计算出所需的解。

迭代法的几何含义

曲线$g(x)$ 与直线$y = x$ 的交点。



迭代法的收敛性

假设$g(x)$ 连续，不动点迭代生成的点列为$\{x_0, x_1, x_2, \dots, x_k, \dots\}$，如果存在$x^*$ 使得

$$\lim_{k\to\infty} x_k = x^*$$

则由连续函数的性质可知

$$x^* = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} g(x_k) = g(\lim_{k\to\infty} x_k) = g(x^*)$$

因此 x^* 是$g(x)$ 的一个不动点即方程的一个根，此时我们称迭代法**收敛**。
如果点列$\{x_k\}_{k=0}^{\infty}$不收敛, 则称不动点迭代是**发散**的。

迭代法的收敛性:李普希兹连续Lipschitz Continuity

定义2.1　设在某个区间Δ内，函数$g(x)$满足下述**李普希兹条件**：

$$|g(x)-g(y)| \leq L|x-y| \quad (\forall x, y \in \Delta, L \geq 0\text{为常数})$$

则称在Δ内$g(x)$**李普希兹连续**。

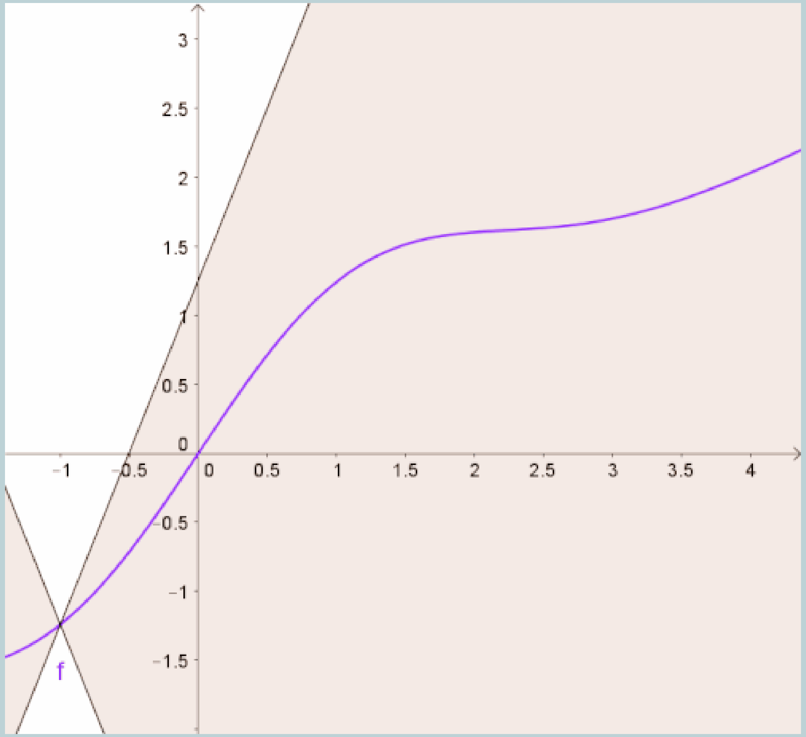
注（教材上定义不严谨）： Lipschitz continuity 只要求Lipschitz Constant $L < \infty$。$0 < L < 1$时，称$g(x)$ 为压缩映射Contraction Mapping，此时Δ内存在唯一不动点，迭代法收敛。

李普希兹连续Lipschitz Continuity

Intuitively, L is a measure of how fast the function can change.

Questions:

1) Examples of Lipschitz continuous functions?
2) Can I assume that neural networks are Lipschitz continuous functions?
3) 有没有其他度量方法?它们之间的区别?



For a Lipschitz continuous function, there exists a double cone (white) whose origin can be moved along the graph so that the whole graph always stays outside the double cone.

迭代法的全局收敛性

假设$g(x) \in C[a,b]$ 且满足

1）对于任意$x \in [a,b]$，都有$g(x) \in [a,b]$，

2）存在常数 L，满足$0 < L < 1$，使得对于任意$x, y \in [a,b]$都有

$$|g(x) - g(y)| \le L|x - y|$$

Metric space
Banach Fixed Point Theorem
Cauchy sequence

则对任意初始值 $x_0 \in [a,b]$，不动点迭代收敛, 且

$$|x^* - x_k| \le \frac{L^k}{1-L}|x_k - x_{k-1}| \le \frac{L^k}{1-L}|x_1 - x_0|$$

板书

其中， x^*是$g(x)$在$[a,b]$内的唯一不动点。

全局收敛：收敛性与初始迭代值的选取无关。

一阶连续可导情况

命题2.1　若$g'(x) \in \mathrm{C}^1[a,b]$（在闭区间内一阶连续可导）且

$$|g'(x_0)| \le L < 1\ (\forall x \in [a,b])$$

则$g(x)$在$[a,b]$内李普希兹连续。

证明 拉格朗日中值定理：$g(x)-g(y)=g'(\xi)(x-y)$

$$|g(x)-g(y)| = |g'(\xi)||x-y| \le L|x-y|$$

推论　设 $x^* = g(x^*)$，若$g(\mathrm{x})$在 x^*附近连续可微且 $|g'(x^*)| < 1$，则迭代格式 $x_{k+1} = g(x_k)$ 在 x^* 附近局部收敛。

注　由于 x^* 事先未知，故实际应用时，代之以近似判则 $|g'(x_0)| < 1$。 但需注意，这实际上是假设了x_0充分接近 x^*，若 x_0 离 x^*较远，迭代格式可能不收敛。

迭代法的局部收敛性

定理2.1 设 $x^* = g(x^*)$，$g(x_k)$ 在闭区间$\Delta: |x - x^*| < \delta$ 内李普希兹连续，则对任何初值 $x_0 \in \Delta$由迭代格式x_{k+1}=$g(x_k)$计算得到的解序列 $\{x_k\}$ 收敛于x^*(这时我们称迭代格式$x_{k+1} = g(x_k)$在 x^*的邻域Δ上**局部收敛**）。

全局收敛与局部收敛 局部收敛意味着只有当初值离真解足够近时。才能保证收敛。由于真解是不知道的。因此如果只具有局部收敛性。则初值选取要求较高，很有可能无法保证收敛。这也是局部收敛与全局收敛的最大区别。实际计算中可以用具有全局收敛性的方法（比如二分法）获取一个近似解。然后再进行迭代。

迭代法的误差估计

$$|x^* - x_k| \le \frac{L}{1-L}(x_k - x_{k-1}) \qquad |x^* - x_k| \le \frac{L^k}{1-L}|x_1 - x_0|$$

由此，对给定的精度$|\tilde{x} - x^*| < \varepsilon$可进行:

(1) 事前误差估计a priori error estimate:

在数值计算之前，通过分析问题的性质（比如函数的光滑性、方法的阶数）来推导误差的理论上界

$$K = \left[\frac{1}{|\ln L|}\left(\ln\frac{|x_1 - x_0|}{1-L} - \ln\varepsilon\right)\right] + 1$$

(2) 事后误差估计a posteriori error estimate:

在数值计算之后，根据计算结果本身来评估误差

$$\frac{L}{1-L}|x_k - x_{k-1}| < \varepsilon$$

简单地代之以$|x_k - x_{k-1}| < \varepsilon$

迭代法：例子

例 2.2 试建立收敛的迭代格式求解$f(x)=x^3-2x-5=0$
在区间(2,3)内满足精度要求$\varepsilon=10^{-8}$的根。取 x_0 =2 。

解 方法1）首先可简单的把$x^3-2x-5=0$ 等价化为$x=\frac{x^3-5}{2}$

由此建立迭代格式$x_{k+1}=\tilde{g}(x_k)$， $\tilde{g}(x)=\frac{x^3-5}{2}$

$\tilde{g}'(x)=3/2x^2$，可验证$\boldsymbol{\tilde{g}'(x_\theta)>1}$

所以该迭代格式在内不收敛，不可取。

迭代法：例子

方法2）为建立收敛的迭代格式，我们把 $x^3-2x-5=0$ 等价化为：

$$x=\sqrt[3]{2x+5}$$

从而建立迭代格式：

$$x_{k+1}=g(x_k),\quad g(x)=\sqrt[3]{2x+5}$$

可验证 $\tilde{g}'(x_\theta)<1$，故该迭代格式收敛。

取 $x_0=2$ 计算，结果见表 2-2。

表 2-2 采用迭代格式 $x_{k+1}=\sqrt[3]{2x_k+5}$ 计算的结果

k	x_k	k	x_k
1	2.080 083 817	6	2.094 550 371
2	2.092 350 721	7	2.094 551 325
3	2.094 217 062	8	2.094 551 563
4	2.094 500 780	9	2.094 551 563
5	2.094 543 695		

拓展： Lipschitz Continuity在机器学习中的应用: 泛化性、稳定性、健壮性， e. g., adversarial robustness, Wasserstein GAN, avoiding exploding gradients。

迭代法的收敛速度

定义2.2 设迭代格式$x_{k+1}=g(x_k)$的解序列$\{x_k\}$收敛于$x=g(x)$的根x^*，如果迭代误差 $e_k=x_k-x^*$ 当$k\to\infty$时满足渐近关系式

$$\frac{e_{k+1}}{e_k^p}\to C\ (C\neq 0\text{为常数})\qquad\qquad (2-11)$$

则称该迭代格式是**p 阶收敛**的。特别地，p=1时称为**线性收敛**，1<p<2 时称为**超线性收敛**，p=2 时称为**平方收敛**。

例：取$e_{k+1}=C\,e_k^p$；$e_0=1$；　$C=0.1$

分别验证线性收敛和平方收敛。

迭代法的加速收敛: Aitken加速技巧

设序列$\{x_k\}$线性收敛于x^*，即有$\lim\limits_{n\to\infty}\frac{e_{k+1}}{e_k}=\mathrm{C}$

$$\begin{cases} x_1 - x^* \approx \mathrm{C}(x_0 - x^*) \\ x_2 - x^* \approx \mathrm{C}(x_1 - x^*) \end{cases}$$

两式相除得 $\frac{x_1-x^*}{x_2-x^*}\approx\frac{x_0-x^*}{x_1-x^*}$

解得

$$x^* - \left[x_2 - \frac{(x_2 - x_1)^2}{x_2 - 2x_1 + x_0}\right] \approx 0$$

取修正值:

$\tilde{\mathrm{x}}_2 = x_2 - \frac{(x_2-x_1)^2}{x_2-2x_1+x_0}$作为$x^*$新的近似值。

迭代法的加速收敛: Steffensen迭代法

Steffensen 迭代法：将 Aitken 加速技巧与不动点迭代相结合。

算法 2.1 （Steffensen 算法）

(1)输入精度要求 ε 和 $\bar{\varepsilon}$,初值 x_0,最大迭代步数 N;

(2)令 $k=0$;

(3)计算 $\alpha=g(x_k)$,$\beta=g(\alpha)$;

(4)计算 $\gamma=\beta-2\alpha+x_k$,若 $|\gamma|<\bar{\varepsilon}$,则终止计算,输出 β 作为近似值;否则转入(5);

(5)计算 $x_{k+1}=\beta-\dfrac{(\beta-\alpha)^2}{\gamma}$(加速处理);

(6)若 $|x_{k+1}-x_k|<\varepsilon$,则终止计算,输出 x_{k+1}作为近似值;否则转入(7);

(7)令 $k=k+1$,若 $k>N$,则停止计算,输出失败信息(这时须另选初值计算);否则转入(3)。

需要注意的是,埃特肯加速技巧对于超线性收敛的解序列并不能明显改善收敛速度。

计算结果对比

表 2-2 采用迭代格式 $x_{k+1}=\sqrt[3]{2x_k+5}$计算的结果

k	x_k	k	x_k
1	2.080 083 817	6	2.094 550 371
2	2.092 350 721	7	2.094 551 325
3	2.094 217 062	8	2.094 551 563
4	2.094 500 780	9	2.094 551 563
5	2.094 543 695		

表 2-3 采用 Steffensen 算法的计算结果

k	x_k	α	β
0	2	2.080 083 823	2.092 350 678
1	2.094 577 839	2.094 555 487	2.094 552 090
2	2.094 551 481	2.094 551 481	2.094 551 481
3	2.094 551 481		

课后作业Homework 2

1. 对于方程 $f(x)=x^3-7.84x-7.68=0$，取步长 $h=1$ 搜索正根所在区间，并对求出的含根区间估计用二分法求正根时所需的步数（精度要求为 $\varepsilon=10^{-4}$）。

3. 为求方程 $x^3-x^2-1=0$ 在 $x=1.5$ 附近的一个根，将方程改写为下列等价形式，并建立相应的迭代格式。

(1) $x=1+\dfrac{1}{x^2}$，迭代格式为 $x_{k+1}=1+\dfrac{1}{x_k^2}$

(2) $x^3=1+x^2$，迭代格式为 $x_{k+1}=\sqrt[3]{1+x_k^2}$

(3) $x^2=\dfrac{1}{x-1}$，迭代格式为 $x_{k+1}=\dfrac{1}{\sqrt{x_k-1}}$

讨论每种格式的收敛性，并用格式(2)求出精度为 10^{-2} 的根的近似值。

作业以word格式将推导，代码，（可视化）运行结果等（控制文件大小，不要粘贴代码图片，文件名：姓名_学号_作业2）发到助教（小王）邮箱：20245227027@stu.suda.edu.cn。

注意：在作业中自己编写函数，不要直接使用相关库中的现成函数。


Thanks