第二章：方程求根

第 (1) 部分：迭代法

目录

问题来源

大量的计算问题可以被理解为求解方程 $f(x)=0$
其中 $x\in\mathbb{R}^d$, $f:\mathbb{R}^d\to\mathbb{R}^m$;
我们一般而言无法找到解析解，因而我们需要数值近似。

例子:

(1) 假设某个势能函数 $V(x)$ 不具有简单的表达式，比如 $V(x)=x^6-x^3+x^2$，而我们需要找到能量最小的位置，即可以通过求解 $V'(x)=0$。对于复杂的势能函数 V，我们无法找到解析解。

(2) 对于常微分方程 $y'(x)=g(x)$ 的隐式格式，如 $y_{k+1}-y_k=(\Delta x)g(y_{k+1})$，其中 y_k 已知，求解 y_{k+1} 的值等于解上述方程 $f(x)=x-y_k-(\Delta x)g(x)$。

(3) 流体力学、量子物理等学科中大量模型涉及偏微分方程，很多时候离散后得到的系统是线性的，即需要求解 $f(x)=Ax+b$ 的形式。

该章节概览

我们将共介绍四种方法:

- 二分法
- 迭代法
- 牛顿法
- 弦截法

目录

核心思想

回顾

对于一个连续函数 f，假设区间 $[a, b]$ 上，$f(a) < 0$, $f(b) > 0$。那么在区间中是否能找到某个位置 $\xi \in (a, b)$，使得 $f(\xi) = 0$?

(A) True

(B) False

核心思想

回顾

对于一个连续函数 f，假设区间 $[a, b]$ 上，$f(a) < 0$, $f(b) > 0$。那么在区间中是否能找到某个位置 $\xi \in (a, b)$，使得 $f(\xi) = 0$?

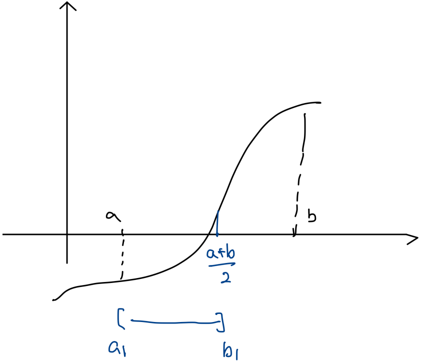
(A) True

(B) False

二分法的核心就是基于该结论!

二分法的算法

在区间 [a.b] 上，假设 $f(a) < 0$，$f(b) > 0$



我们取中点 $x_0 = \frac{a+b}{2}$，

- 若 $f(x_0) = 0$，则找到了！
- 若 $f(x_0) > 0$，则考虑 $[a_1, b_1] = [a, x_0]$；
- 若 $f(x_0) < 0$，则考虑 $[a_1, b_1] = [x_0, b]$。

将 $[a_1, b_1]$ 作为新区间，继续通过不断二分。

二分法的停止条件

将 $[a_1, b_1]$ 作为新区间，继续通过不断二分。我们知道

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \cdots$$

且

$$b_k - a_k = \frac{b-a}{2^k}$$

当二分法的 k 足够大，我们可以看到

$$\left| x^* - \frac{a_k + b_k}{2} \right| \leq \frac{b_k - a_k}{2} = \frac{b-a}{2^{k+1}}$$

例子 1：假设区间为 $a = 0$，$b = 1$，我们希望找到一个近似根使得离真实值 x^* 最多差 $2^{-8} \approx 0.00390625$，则我们做多需要计算函数值 $f(x)$ 多少次?

总结

二分法的好处：简单

二分法的局限：基本局限于一维函数

二分法中计算 f 的次数：最多 $\lceil \log_2 \left(\frac{b-a}{\epsilon}\right) - 1 \rceil$，其中 ϵ 是你需要的误差限，$\lceil \cdot \rceil$ 是向上取整函数。

目录

迭代法的设计原理

另一个常见的等价形式就是求解 $x = \varphi(x)$，即不动点。

问题 1： 对于问题 $f(x) = 0$，对应的 φ 是什么？

问题 2： 具体方法是什么？有什么好处呢？

定理 (不动点原理)

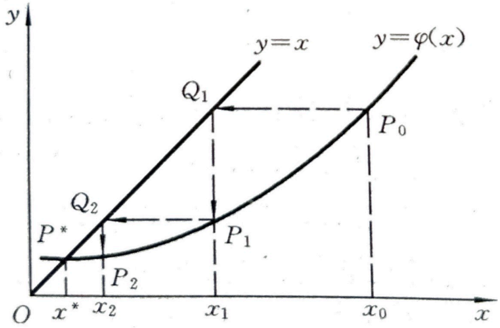
我们考虑区间 $[a, b]$，并且假设 φ 的函数值范围也是 $[a, b]$。考虑迭代过程 $x_{n+1} = \varphi\left(x_n\right), n = 0, 1, 2 \ldots$ 如果

- $|\varphi(x) - \varphi(y)| \leqslant c|x - y| \ \forall x, y \in [a, b]$
- 且 $c < 1$

则 $x^* = \lim_{n\to\infty} x_n$ 存在且 $x^* = \varphi\left(x^*\right)$.

方法一：看 $x_{n+1} - x_n$ 的距离

方法二：几何解释（即画图证明法）



我们具体怎么用?

Algorithm 1: 最简单的算法形式

Input: 取 $\varphi(x) = f(x) + x$

Result: x_N

猜一个值 x_0

$n = 1$

while $n < N$ **do**

  $x_n = \varphi(x_{n-1})$

  $n = n + 1$

end

相比于二分法，该方法也很简单；该方法可以适用于任何维数，例如 $f : \mathbb{R}^{100} \to \mathbb{R}^{100}$。

问题 3： 但是有完美的方法吗？代价是什么？

例子

假设 $\varphi(x) = x^2$，我们想要解 $\varphi(x) = x$ (显然答案是 $x = 0$ 或 $x = 1$)。请测试使用迭代法且初始值为

(1) $x_0 = -\frac{1}{2}$

(2) $x_0 = \frac{1}{2}$

(3) $x_0 = 2$

(4) $x_0 = 1.001$

你能观察到什么现象？目前需要解决的问题是什么？

例子

假设 $\varphi(x)=x^2$，我们想要解 $\varphi(x)=x$ (显然答案是 $x=0$ 或 $x=1$)。请测试使用迭代法且初始值为

(1) $x_0=-\frac{1}{2}$
(2) $x_0=\frac{1}{2}$
(3) $x_0=2$
(4) $x_0=1.001$

你能观察到什么现象? 目前需要解决的问题是什么?

- 猜不同的初值会得到不同的解，因而为了得到所有解，我们或许得多猜几个不同的初值。
- 问题 4：对这个例子，不同初值有时会收敛，有时不会，为什么呢?
- 问题 5：若收敛，何时结束（即终止条件是什么?）

对于使用范围的理解（即解决问题 4）

迭代法无法帮我们找到所有零点，因而我们需要考虑适用范围是什么?

我们来研究一下如下情景:【假设某人】告诉了你 x^* 的近似值，$\varphi'(x^*) > 0$，且 $x_0 \approx x^*$，下一步 $x_1 = \varphi(x_0)$，因而，对于

$$x_1 - x^* = \varphi(x_0) - \varphi(x^*) \;\approx\; \underline{\qquad\qquad} \cdot (x_0 - x^*).$$

如果

- $\underline{\qquad\qquad}$，x_1 会离真实值更远 $\Rightarrow$ 迭代会发散
- $\underline{\qquad\qquad}$，x_1 会离真实值更近 $\Rightarrow$ 迭代会收敛

对于使用范围的理解（即解决问题 4）

迭代法无法帮我们找到所有零点，因而我们需要考虑适用范围是什么?

我们来研究一下如下情景:【假设某人】告诉了你 x^* 的近似值，$\varphi'(x^*) > 0$，且 $x_0 \approx x^*$，下一步 $x_1 = \varphi(x_0)$，因而，对于

$$x_1 - x^* = \varphi(x_0) - \varphi(x^*) \approx \varphi'(x^*) \cdot (x_0 - x^*).$$

如果

- $|\varphi'(x^*)| > 1$，x_1 会离真实值更远 $\Rightarrow$ 迭代会发散
- $|\varphi'(x^*)| < 1$，x_1 会离真实值更近 $\Rightarrow$ 迭代会收敛

这个解释了问题 4。

定理

假定函数 φ 满足

1° 对任意 $x \in [a, b]$ ，有 $a \leq \varphi(x) \leq b$;

2^0 存在正数 $L < 1$ 使得对任意 $x \in [a, b]$ 有

$$|\varphi'(x)| \leq L < 1$$

则迭代过程 $x_{k+1} = \varphi(x_k)$ 对任意初值 $x_0 \in [a, b]$ 均收敛到 $x = \varphi(x)$ 的根 x^* 且

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

注明:（1）其证明的核心基本就是不动点原理的证明;（2）由 $|\varphi'(x)| \leq L$, 通过微分中值定理，我们可以证明 $|\varphi(x) - \varphi(y)| \leq L|x - y| \ \forall x, y \in [a, b]$。

练习

例子 2：假设我们考虑例子 $[a,b]=[-\frac{\pi}{2},\frac{\pi}{2}]$，$\varphi(x)=\frac{\pi}{2}\sin(x)$，一个解为 $x^*=0$。假设 $x_0=10^{-2}$，迭代法能否收敛到 $x^*=0$?

例子 3：假设我们考虑例子 $[a,b]=[-\frac{\pi}{2},\frac{\pi}{2}]$，$\varphi(x)=\frac{1}{2}\sin(x)$，假设 $x_0=10^{-2}$，

(1) 该方法能收敛到 $x^*=0$ 吗?

(2) 通过上述估计，来确认 L 的范围，并且若我们需要误差限 10^{-4}，则我们最多需要迭代几次?

【可同步参考代码部分】

停止迭代的判断方法（部分解答问题 5）

停止迭代的判断方法

$$\begin{aligned}\left|x_{k+p}-x_k\right| &\leqslant \left(L^{p-1}+L^{p-2}+\cdots+1\right)\left|x_{k+1}-x_k\right| \\ &\leq \frac{1}{1-L}\left|x_{k+1}-x_k\right|\end{aligned}$$

令 $p \to \infty$，

$$\left|x^*-x_k\right| \leqslant \frac{1}{1-L}\left|x_{k+1}-x_k\right|$$

因此只要相邻结果 $\left|x_{k+1}-x_k\right|$ 是够小，且 L 不是非常靠近 1，即可保证近似精确。

定理

若存在 x^* 的某于邻域 $R: |x - x^*| \le \delta$ 使迭代过程 $x_{k+1} = \varphi(x_k)$ 对于任意初值 $x_0 \in R$ 均收敛，则称迭代过程在根 x^* 邻近具有局部收敛性。

定理

设 x^* 为方程 $x = \varphi(x)$ 的根，$\varphi'(x)$ 在 x^* 附近连续且 $|\varphi'(x^*)| < 1$ ，则迭代过程 $x_{k+1} = \varphi(x_k)$ 在 x^* 邻近具有局部收敛性。

迭代法的收敛速度

对于简单的迭代法（假设收敛），

$$x_{k+1} - x^* = \varphi(x_k) - \varphi(x^*) \approx \varphi'(x^*)(x_k - x^*)$$

假设 x_k 已经非常接近 x^*，因而误差 $e_{k+1} \approx |\varphi'(x^*)|e_k$。由于 $|\varphi'(x^*)| < 1$，$e_{k+1} < e_k$（即误差不断减少）

$$e_{k+1} \approx |\varphi'(x^*)|e_k \qquad e_{k+2} \approx |\varphi'(x^*)|e_{k+1} \qquad \cdots$$

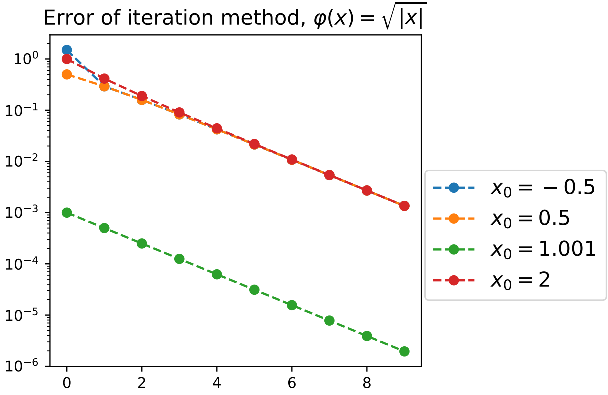
找规律可知 $e_N \approx |\varphi'(x^*)|^{N-k} e_k$ 即

$$\log(e_N) \approx \log(e_k) + (N-k)\underbrace{\log|\varphi'(x^*)|}_{<0}$$

当我们画 $\log(e_N)$ 关于迭代次数 N 的曲线时，我们预计得到线性关系。

数值实验

例子 4：考虑 $\varphi(x) = \sqrt{|x|}$ 以及初始值 x_0



且经验的斜率在 -0.6907 至 -0.6956 范围之间，而理论值是 $\log(1/2) \approx -0.6931$。

总结

- 迭代法通过将 $f(x)=0$ 的形式变成 $\varphi(x)=x$ 的形式求解
- 核心的原理就是 不动点原理
- 如果要求得到多个解，则需要尝试不同的初值
- 当 $|\varphi'(x^*)|<1$，迭代法可以在初值 x_0 取得很 " 聪明 " 的情况下有机会得到该 x^* 的近似解；反之，迭代法无法帮助我们得到 x^* 的近似。
- 迭代法的终止条件可以通过 $|x_{k+1}-x_k|$ 的大小来判断。
- 误差在取 $\log$ 后（即 $\log(e_N)$）和迭代次数 N 线性负相关。

目录

Aitken 加速法

猜测值 x_0，我们通过迭代得到 $x_1 = \varphi(x_0)$，$x_2 = \varphi(x_1)$。由于

$$\frac{\varphi(x_0) - x^*}{x_0 - x^*} \approx \frac{\varphi(x_1) - x^*}{x_1 - x^*}$$

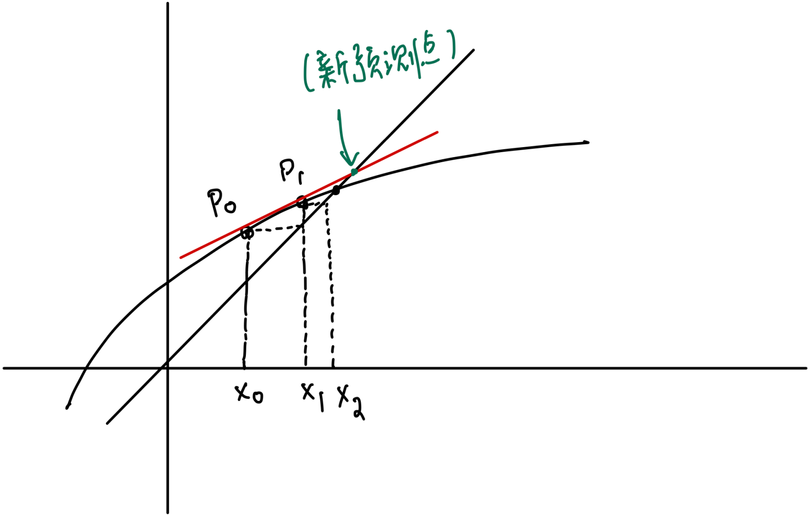
我们对于 x^* 的值通过解该方程得到，即找到 x^* 使得 $(x_0, \varphi(x_0))$，$(x_1, \varphi(x_1))$，(x^*, x^*) 在同一直线上。解方程可知

$$x^* \approx x_2 - \frac{(x_2 - x_1)^2}{x_0 + x_2 - 2x_1}$$

因而该算法可写为：在第 k 步

$$\begin{aligned}
\tilde{x}_{k+1} &= \varphi(x_k) \\
\bar{x}_{k+1} &= \varphi(\tilde{x}_{k+1}) \\
x_{k+1} &= \bar{x}_{k+1} - \frac{(\bar{x}_{k+1} - \tilde{x}_{k+1})^2}{\bar{x}_{k+1} - 2\tilde{x}_{k+1} + x_k}
\end{aligned}$$

Aitken 加速法的几何解释



总结

该部分探究了二分法和迭代法。

迭代法中最重要的是了解**迭代法的收敛条件与收敛速度**。