

## 第五讲 非线性方程



### 目录

5.1 对分法

5.2 不动点迭代法

5.3 Steffensen 迭代法

5.4 Newton 迭代法

5.5 割线法与抛物线法

<https://math.ecnu.edu.cn/~jypan/Teaching/NA>

$$f(x) = 0$$

➤ 线性 or 非线性方程

➤ 非线性方程简单应用一: 液体在管道中的摩擦系数

$$\frac{1}{\sqrt{x}} = \frac{1}{k} \ln(\text{Re} \cdot \sqrt{x}) + \left(14 - \frac{5.6}{k}\right)$$

➤ 非线性方程简单应用二: 代数方程求解

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0.$$

## 非线性方程的解法

非线性方程 (组) 一般不存在直接解法, 通常需要使用 **迭代法** 来求数值解.

# 一些基本概念

## ► 解, 根, 零点:

所有满足  $f(x_*) = 0$  的数  $x_*$  都称为方程的解或根, 也称为  $f(x)$  的零点.

## ► 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).

## ► 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.

我们研究内容就是在 **有解** 的前提下求出方程  $f(x) = 0$  的 **近似解**.

 若无特别说明, 本讲我们只考虑实数解, 并假定  $f(x)$  的表达式中也只包含实数

 非线性方程可能存在多个或无穷多个解, 因此要强调 **求解区域**

# 5-1 | 对分法

## 5.1 对分法

### 5.1.1 对分法基本思想

### 5.1.2 对分法的收敛性

# 5-1-1 | 对分法基本思想

## 对分法 (Bisection)

设  $f(x) = 0$  在区间  $[a, b]$  内至少有一个实数解.

对分法的基本思想就是将这个有解区间进行对分, 并找出解所在的小区间, 记为新的有解区间, 然后再对这个小区间进行对分.

依次类推, 直到有解区间的长度足够小为止, 此时有解区间内的任意一点都可以作为  $f(x) = 0$  的近似解 (实际计算中通常取中点).

## 对分法的数学原理

**定理 (介值定理)** 设  $f(x)$  在  $[a, b]$  内连续, 且  $f(a)f(b) < 0$ , 则在  $(a, b)$  内至少存在一点  $\xi$ , 使得  $f(\xi) = 0$ .

# 算法描述

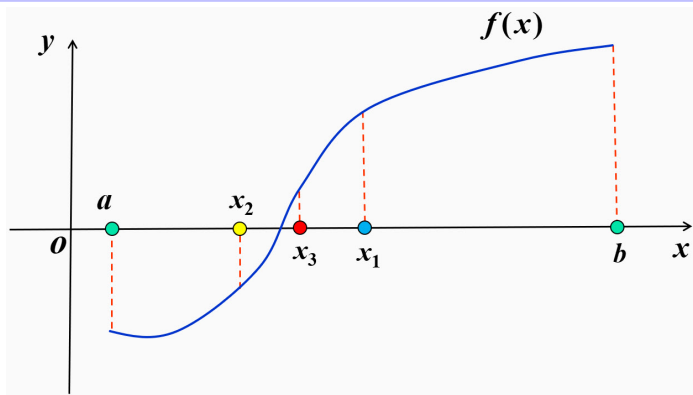
---

## 算法 对分法

---

- 1: 给定函数  $f(x)$  和求解区间  $[a, b]$ , 以及精度要求  $\varepsilon > 0$
  - 2: 令  $a_1 = a, b_1 = b$
  - 3: 计算  $f(a_1)$  和  $f(b_1)$ ,  
    如果  $|f(a_1)| < \varepsilon$ , 则返回数值解  $x = a_1$  并停止计算;  
    如果  $|f(b_1)| < \varepsilon$ , 则返回数值解  $x = b_1$  并停止计算;  
    如果  $f(a_1)f(b_1) > 0$ , 则输出算法失败信息并停止计算
  - 4: **for**  $k = 1, 2, \dots$ , **do**
  - 5:     计算  $x_k = \frac{a_k + b_k}{2}$  和  $f(x_k)$
  - 6:     如果  $|f(x_k)| < \varepsilon$  或者  $|b_k - a_k| < \varepsilon$ , 则返回数值解  $x_k$  并停止计算;
  - 7:     如果  $f(a_k)f(x_k) < 0$ , 则令  $a_{k+1} = a_k, b_k = x_k$ ; 否则令  $a_{k+1} = x_k, b_{k+1} = b_k$ ;
  - 8: **end for**
-

# 对分法的几何表示



## 几点说明

- 这里采用的停机准则是函数值充分小或者求解区间充分小.
- 用对分法求解时, 可以先画出  $f(x)$  草图, 以确定一个大致的有解区间.

# 5-1-2 | 对分法的收敛性

记第  $k$  步得到的有解区间为  $[a_k, b_k]$ , 中点为  $x_k$ . 则  $a_1 = a, b_1 = b$ , 且

$$|x_k - x_*| = \left| \frac{1}{2}(a_k + b_k) - x_* \right| \leq \frac{1}{2}(b_k - a_k).$$

由于每次都是对分有解区间, 因此有

$$b_k - a_k = \frac{1}{2}(b_{k-1} - a_{k-1}) = \frac{1}{2^2}(b_{k-2} - a_{k-2}) = \cdots = \frac{1}{2^{k-1}}(b_1 - a_1).$$

因此

$$|x_k - x_*| \leq \frac{1}{2}(b_k - a_k) \leq \frac{1}{2^k}(b - a) \rightarrow 0 \quad (k \rightarrow \infty)$$

**定理** 设  $f(x) \in C[a, b]$ , 且  $f(a)f(b) < 0$ , 则对分法收敛到  $f(x) = 0$  的一个解.

## 关于对分法的几点注记

- 适用范围: 只适合求连续函数的 **单重实根** 或 **奇数重实根**;
- 优点: 简单易用, 只要满足介值定理的条件, 算法总是收敛的;
- 缺点: (1) 收敛速度较慢; (2) 不能求复根和偶数重根; (3) 只能求一个根;

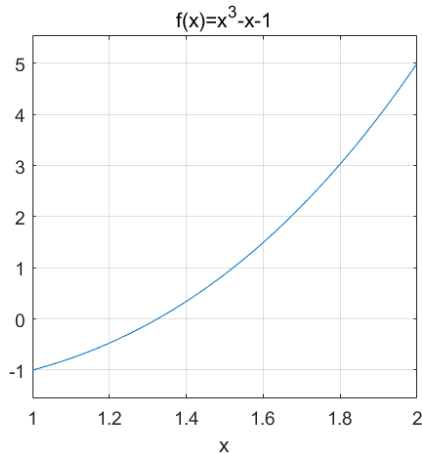
一般可先用对分法来计算解的一个粗糙估计, 然后再用其他方法进行加速, 如 Newton 法.

**例** 用对分法求  $f(x) = x^3 - x - 1 = 0$  在  $[1, 2]$  内的根.

(NLS\_bisection.m)

易知  $f(1) = -1 < 0$ ,  $f(2) = 5 > 0$ , 所以  $f(x)$  在  $[1, 2]$  内存在零点 (见下图).

| $k$ | $a/f(a)$ | $b/f(b)$ | $x$    | $f(x)$  |
|-----|----------|----------|--------|---------|
| 1   | 1.0000/- | 2.0000/+ | 1.5000 | 0.8750  |
| 2   | 1.0000/- | 1.5000/+ | 1.2500 | -0.2969 |
| 3   | 1.2500/- | 1.5000/+ | 1.3750 | 0.2246  |
| 4   | 1.2500/- | 1.3750/+ | 1.3125 | -0.0515 |
| 5   | 1.3125/- | 1.3750/+ | 1.3438 | 0.0826  |
| 6   | 1.3125/- | 1.3438/+ | 1.3281 | 0.0146  |
| 7   | 1.3125/- | 1.3281/+ | 1.3203 | -0.0187 |
| 8   | 1.3203/- | 1.3281/+ | 1.3242 | -0.0021 |
| 9   | 1.3242/- | 1.3281/+ | 1.3262 | 0.0062  |
| 10  | 1.3242/- | 1.3262/+ | 1.3252 | 0.0020  |



# 5-2 | 不动点迭代法

## 5.2 不动点迭代法

5.2.1 基本思想

5.2.2 收敛性分析

5.2.3 收敛性阶

# 5-2-1 | 基本思想

## 不动点迭代法

$$f(x) = 0 \iff \varphi(x) - x = 0 \quad \text{或} \quad x = \varphi(x)$$

基于该等价方程, 构造出不动点迭代的 **一般迭代格式**:

$$x_{k+1} = \varphi(x_k), \quad k = 0, 1, 2, \dots,$$

其中  $x_0$  为迭代初始值, 可以任意选取.

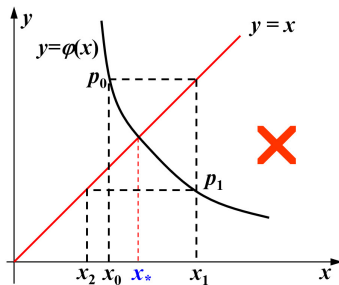
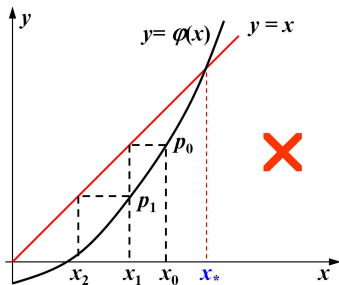
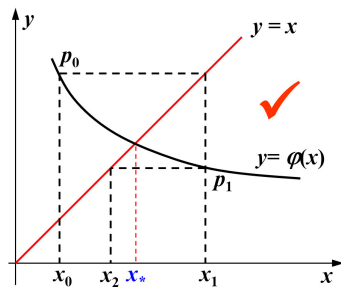
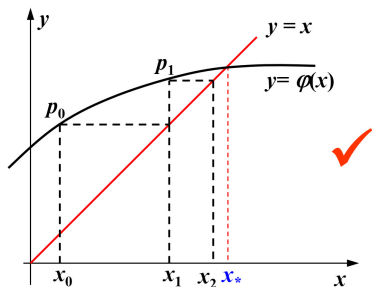
这就是 **不动点迭代法** (Fixed-Point iteration), 简称 **迭代法**,  $\varphi(x)$  称为 **迭代函数**.

# 不动点迭代法

## 注记

- $x_*$  是  $f(x) = 0$  的解当且仅当  $x_* = \varphi(x_*)$ , 即  $\varphi(x)$  的不动点.
- 不动点迭代将 **方程求解** 转化为 **函数求值**, 后者显然要容易很多.

# 几何含义: 曲线 $y = \varphi(x)$ 与直线 $y = x$ 的交点



# 5-2-2 | 收敛性分析

## 收敛 or 发散

设  $\varphi(x)$  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

$$\lim_{k \rightarrow \infty} x_k = x_*,$$

则由连续函数的性质可知

$$x_* = \lim_{k \rightarrow \infty} x_{k+1} = \lim_{k \rightarrow \infty} \varphi(x_k) = \varphi\left(\lim_{k \rightarrow \infty} x_k\right) = \varphi(x_*).$$

因此  $x_*$  是  $\varphi(x)$  的一个不动点, 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.

# 存在唯一性

**定理 (不动点的存在唯一性)** 设  $\varphi(x) \in C[a, b]$  且满足

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

$$|\varphi(x) - \varphi(y)| \leq L|x - y|,$$

则  $\varphi(x)$  在  $[a, b]$  上存在唯一的不动点.

(板书)

 条件  $|\varphi(x) - \varphi(y)| \leq L|x - y|$  称为 **Lipschitz 条件**.

当  $L < 1$  时, 称  $\varphi(x)$  为 **压缩映射**.

# 全局收敛性

**定理 (不动点迭代的全局收敛性)** 设  $\varphi(x) \in C[a, b]$  且满足

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

$$|\varphi(x) - \varphi(y)| \leq L|x - y|.$$

则对任意初始值  $x_0 \in [a, b]$ , 不动点迭代收敛, 且

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

其中  $x_*$  是  $\varphi(x)$  在  $[a, b]$  内的唯一不动点.

(板书)

 **全局收敛:** 收敛性与迭代初值的选取无关.

# 一阶连续可导情形

设  $\varphi(x)$  在  $[a, b]$  上连续, 在  $(a, b)$  内可导, 则由 Lagrange 中值定理可得

$$\varphi(y) - \varphi(x) = \varphi'(\xi)(y - x), \quad \xi \in (a, b)$$

因此

$$|\varphi(y) - \varphi(x)| \leq \max_{\xi \in [a, b]} |\varphi'(\xi)| \cdot |y - x|.$$

**推论** 设  $\varphi(x) \in C^1[a, b]$  且对  $\forall x \in [a, b]$ , 都有  $|\varphi'(x)| \leq L < 1$ . 若存在常数  $L$ , 使得

$$|\varphi'(x)| \leq L < 1, \quad \forall x \in [a, b],$$

则对任意初始值  $x_0 \in [a, b]$ , 不动点迭代收敛, 且

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

# 不动点迭代法举例

**例** 试构造不动点迭代格式, 计算  $f(x) = x^3 - x - 1$  在  $[1, 2]$  中的零点.

(NLS\_fixpoint\_01.m)

➤  $\varphi(x) = x^3 - 1$       无法判断

(1)  $\varphi(x) \in [1, 2]$ ?      → **NO**

(2)  $|\varphi'(x)| \leq L < 1$ ?      → **NO**

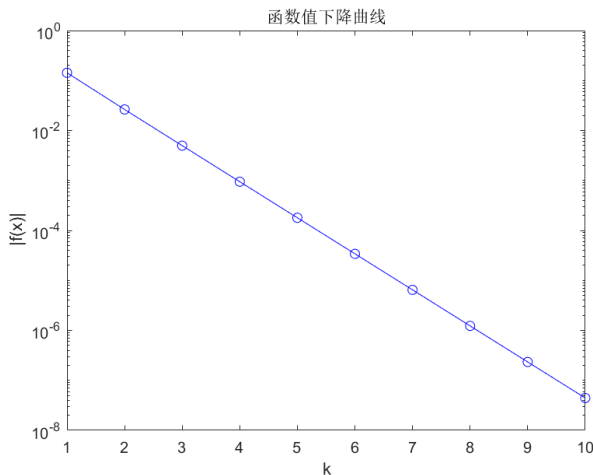
➤  $\varphi(x) = \sqrt[3]{x+1}$       全局收敛

(1)  $1 \leq \varphi(x) \leq 2, \quad \forall x \in [1, 2]$

(2)  $|\varphi'(x)| = \left| \frac{1}{3}(x+1)^{-2/3} \right| \leq \frac{1}{3} \sqrt[3]{0.25} < 1$

取中点为初始值, 不动点迭代  $x_{k+1} = \varphi(x_k) = \sqrt[3]{x_k + 1}$  的计算结果:

| $k$ | $x$    | $ f(x) $ |
|-----|--------|----------|
| 1   | 1.3572 | 1.43e-01 |
| 2   | 1.3309 | 2.63e-02 |
| 3   | 1.3259 | 4.98e-03 |
| 4   | 1.3249 | 9.44e-04 |
| 5   | 1.3248 | 1.79e-04 |
| 6   | 1.3247 | 3.41e-05 |
| 7   | 1.3247 | 6.47e-06 |
| 8   | 1.3247 | 1.23e-06 |
| 9   | 1.3247 | 2.33e-07 |
| 10  | 1.3247 | 4.43e-08 |



# 局部收敛性

**定义** 设  $x_*$  是  $\varphi(x)$  的不动点, 若存在  $x_*$  的某个  $\delta$ -邻域

$$U_\delta(x_*) \triangleq \{x \in \mathbb{R} : |x - x_*| < \delta\},$$

使得对任意  $x_0 \in U_\delta(x_*)$ , 不动点迭代均收敛, 则称该迭代是 **局部收敛** 的.

## 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 (比如对分法) 获取一个近似解, 然后再进行迭代.

# 局部收敛性

**定理** 设  $x_*$  是  $\varphi(x)$  的不动点, 若  $\varphi'(x)$  在  $x_*$  的某个邻域内连续且

$$|\varphi'(x_*)| < 1,$$

则不动点迭代  $x_{k+1} = \varphi(x_k)$  局部收敛.

(板书)

# 5-2-3 | 收敛阶

收敛阶是衡量迭代法收敛速度快慢的一个重要指标.

**定义** 设迭代  $x_{k+1} = \varphi(x_k)$  收敛到不动点  $x_*$ . 记  $e_k \triangleq x_k - x_*$ , 若

$$\lim_{k \rightarrow \infty} \frac{|e_{k+1}|}{|e_k|^p} = c,$$

其中  $p \geq 1$ , 常数  $c$  与  $k$  无关, 则称该迭代是  $p$  阶收敛的.

- (1) 若  $p = 1$  且  $0 < c < 1$ , 则称 **线性收敛**;
- (2) 若  $p = 2$ , 则称 **二次收敛** 或 **平方收敛**;
- (3) 若  $1 < p < 2$  或  $p = 1$  且  $c = 0$ , 则称 **超线性收敛**.

# 线性收敛

**定理** 设  $x_*$  是  $\varphi(x)$  的不动点, 若  $\varphi'(x)$  在  $x_*$  的某个邻域内连续且

$$|\varphi'(x_*)| < 1,$$

则不动点迭代至少 **局部线性收敛**.

# $p$ 阶收敛

**定理** 设  $x_*$  是  $\varphi(x)$  的不动点且  $p \geq 2$  是正整数. 若  $\varphi^{(p)}(x)$  在  $x_*$  的某个邻域内连续且

$$\varphi'(x_*) = \varphi''(x_*) = \cdots = \varphi^{(p-1)}(x_*) = 0, \quad \varphi^{(p)}(x_*) \neq 0, \quad (5.1)$$

则不动点迭代是  $p$  阶局部收敛 的, 且有

$$\lim_{k \rightarrow \infty} \frac{x_{k+1} - x_*}{(x_k - x_*)^p} = \frac{\varphi^{(p)}(x_*)}{p!}.$$

(板书)

**例** 试构造不同的不动点迭代格式, 计算  $f(x) = x^2 - 3$  的正的零点  $x_* = \sqrt{3}$ .

- (1) 构造  $\varphi(x) = x^2 - 3 + x$ , 则  $\varphi'(x_*) = 2\sqrt{3} + 1 > 1$ , 无法判断其收敛性, 所以该迭代函数不能用.
- (2) 构造  $\varphi(x) = x - \frac{x^2 - 3}{4}$ , 则  $\varphi'(x_*) = 1 - \frac{\sqrt{3}}{2} \approx 0.134 < 1$ , 所以该迭代函数可以采用, 且一阶局部收敛.
- (3) 构造  $\varphi(x) = \frac{1}{2} \left( x + \frac{3}{x} \right)$ , 则  $\varphi'(x_*) = 0$ ,  $\varphi''(x_*) = \frac{2}{\sqrt{3}} \neq 0$ , 所以该迭代函数可以采用, 且二阶局部收敛.

(NLS\_fixpoint\_02.m)

## 几点说明

- 一般来说,  $|\varphi'(x_*)|$  越小, 不动点迭代收敛越快!
- 在前面的介绍的迭代法中, 计算  $x_{k+1}$  时, 只用到点  $x_k$  的值. 有时, 我们为了利用前面多个点的信息, 比如前  $\ell \geq 2$  个点, 可以设计下面的迭代法:

$$x_{k+1} = \varphi(x_k, x_{k-1}, \dots, x_{k-\ell+1}), \quad k = \ell - 1, \ell, \ell + 1, \dots \quad (5.2)$$

我们称之为 **多点迭代法**. 比如后面会提到的割线法和抛物线法, 都是多点迭代. 显然, 多点迭代法一开始需要给定  $\ell$  个初始值, 即  $x_0, x_1, \dots, x_{\ell-1}$ .

# 5-3 | Steffensen 迭代法

## 5.3 Steffensen 迭代法

### 5.3.1 Aitken 加速技巧

### 5.3.2 Steffensen 迭代格式

# 5-3-1 | Aitken 加速技巧

迭代两次:  $x_{k+1} = \varphi(x_k)$ ,  $x_{k+2} = \varphi(x_{k+1})$ , 可得

$$x_{k+1} - x_* = \varphi(x_k) - \varphi(x_*) = \varphi'(\xi_1)(x_k - x_*),$$

$$x_{k+2} - x_* = \varphi(x_{k+1}) - \varphi(x_*) = \varphi'(\xi_2)(x_{k+1} - x_*).$$

 假定迭代收敛, 则当  $k$  较大时,  $x_k$  和  $x_{k+1}$  都接近  $x_*$ , 因此  $\xi_1$  与  $\xi_2$  也相近, 于是  $\varphi'(\xi_1) \approx \varphi'(\xi_2)$ :

$$\frac{x_{k+1} - x_*}{x_{k+2} - x_*} \approx \frac{x_k - x_*}{x_{k+1} - x_*} \implies$$

$$x_* \approx x_k - \frac{(x_{k+1} - x_k)^2}{x_{k+2} - 2x_{k+1} + x_k}$$

因此, 上式右端可能是  $x_*$  的一个更好的近似.

# Aitken 加速方法

$$y_{k+1} = x_k - \frac{(x_{k+1} - x_k)^2}{x_{k+2} - 2x_{k+1} + x_k}$$

这时我们就得到两个迭代序列:  $\{x_k\}_{k=0}^{\infty}$  和  $\{y_k\}_{k=0}^{\infty}$ .

**定理** 假定原不动点迭代收敛, 且  $\varphi'(x_*) \neq 1$ , 则

$$\lim_{k \rightarrow \infty} \frac{y_{k+1} - x_*}{x_k - x_*} = 0.$$

这意味着  $y_k$  比  $x_k$  更快地收敛到  $x_*$ .

## 5-3-2

## Steffensen 迭代格式

将 Aitken 加速技巧与不动点迭代相结合, 就得到 **Steffensen 迭代法**:

$$y_k = \varphi(x_k), \quad z_k = \varphi(y_k), \quad x_{k+1} = x_k - \frac{(y_k - x_k)^2}{z_k - 2y_k + x_k}, \quad k = 0, 1, 2, \dots$$



写成不动点迭代形式可得

$x_{k+1} = \psi(x_k)$ , 其中迭代函数  $\psi(x)$  为

$$\psi(x) = x - \frac{(\varphi(x) - x)^2}{\varphi(\varphi(x)) - 2\varphi(x) + x}$$

# 收敛性

**定理** 设  $x_*$  是  $\psi(x)$  的不动点, 则  $x_*$  是  $\varphi(x)$  的不动点. 反之, 若  $x_*$  是  $\varphi(x)$  的不动点,  $\varphi''(x)$  存在且  $\varphi'(x_*) \neq 1$ , 则  $x_*$  是  $\psi(x)$  的不动点. 另外, 若原不动点迭代是线性收敛的, 则对应的 Steffensen 迭代二阶收敛.

 如果原迭代法是  $p$  阶收敛的, 则 Steffensen 迭代  $p + 1$  阶收敛.

 若原迭代法不收敛, Steffensen 迭代可能收敛.

# Steffensen 迭代举例

**例** 用 Steffensen 迭代法求  $f(x) = x^3 - x - 1$  在区间  $[1, 2]$  内的零点

(取  $\varphi(x) = x^3 - 1$ )

([NLS\\_Steffensen\\_01.m](#))

**例** 用 Steffensen 迭代法求  $f(x) = 3x^2 - e^x$  在区间  $[3, 4]$  内的零点

(取  $\varphi(x) = 2 \ln(x) + \ln 3$ )

([NLS\\_Steffensen\\_02.m](#))

谢谢  
THANK YOU

