Optimal Theory and Method

最优化理论与方法

程春杰

杭州电子科技大学 自动化学院

科技馆512

Email: cjzhai@hdu.edu.cn

目录

目录

■ 一维搜索的概念

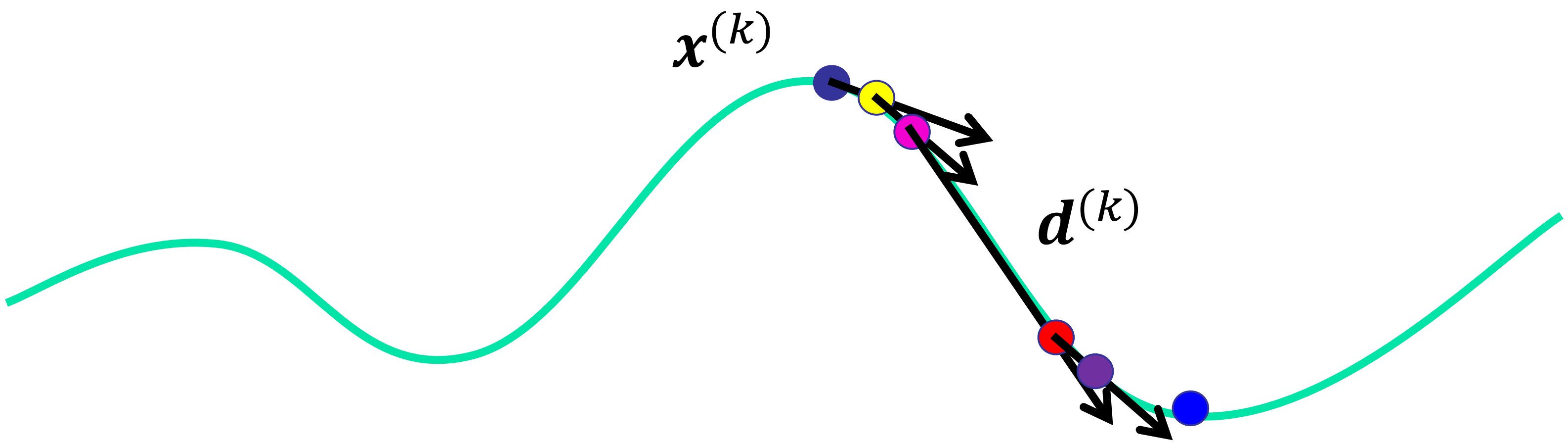
- ➢什么是一维搜索
- ➢一维搜索的三要素
- ➢一维搜索的分类

■ 一维搜索的概念

➢什么是一维搜索

大部分迭代下降算法具有一个共同点：

得到点$x^{(k)}$后，需要按某种规则确定一个方向 $d^{(k)}$，再从$x^{(k)}$出发，沿方向$d^{(k)}$在直线（或射线）上求目标函数的极小点，从而得到$x^{(k)}$的后继点$x^{(k+1)}$

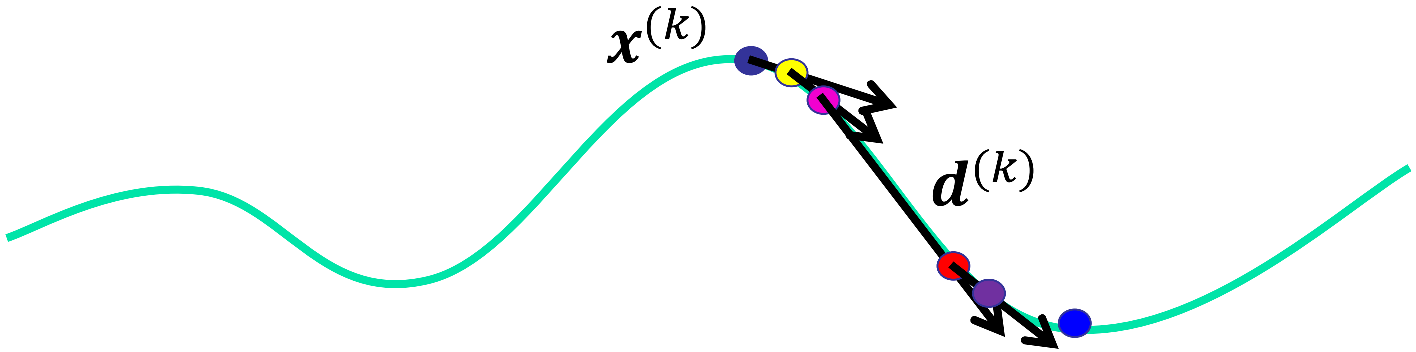


重复以上做法，直至求得问题的解．这里所谓求目标函数在直线上的极小点，称为一维搜索，或称为线搜索

■ 一维搜索的概念

➢一维搜索的三要素

一维搜索可归结为单变量函数的极小化问题



设目标函数为 $f(\boldsymbol{x})$,过点 $\boldsymbol{x}^{(k)}$ 沿方向 $\boldsymbol{d}^{(k)}$ 的直线可用点集来表示,记作

$$L = \{\boldsymbol{x} \mid \boldsymbol{x} = \boldsymbol{x}^{(k)} + \lambda \boldsymbol{d}^{(k)}, -\infty < \lambda < \infty\}$$

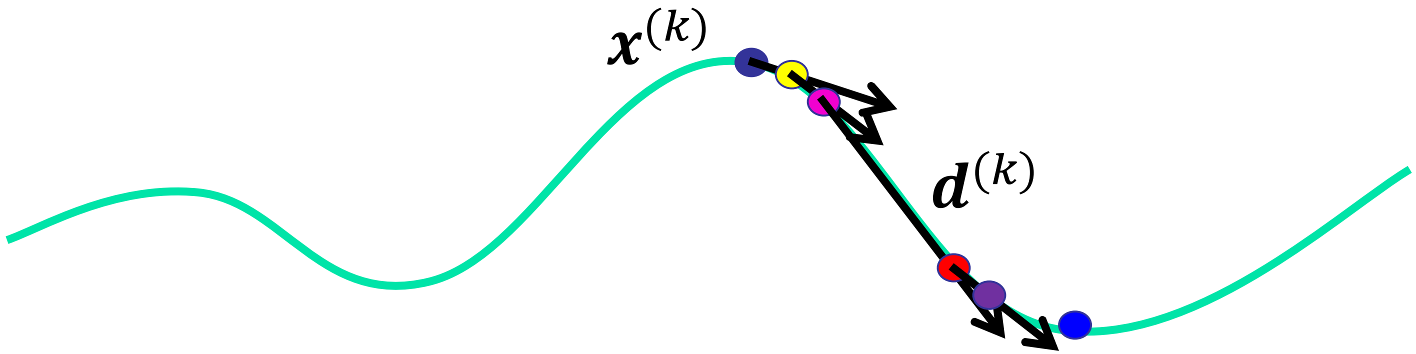
求 $f(\boldsymbol{x})$在直线 L 上的极小点转化为求一元函数

$$\varphi(\lambda) = f(\boldsymbol{x}^{(k)} + \lambda \boldsymbol{d}^{(k)})$$

的极小点.

■ 一维搜索的概念

➢一维搜索的三要素



如果 $\varphi(\lambda)$的极小点为 λ_k ,通常称 λ_k 为沿方向 $\boldsymbol{d}^{(k)}$ 的**步长因子**，或简称为**步长**，$f(\boldsymbol{x})$在直线 L 上的极小点就是

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \lambda_k \boldsymbol{d}^{(k)} .$$

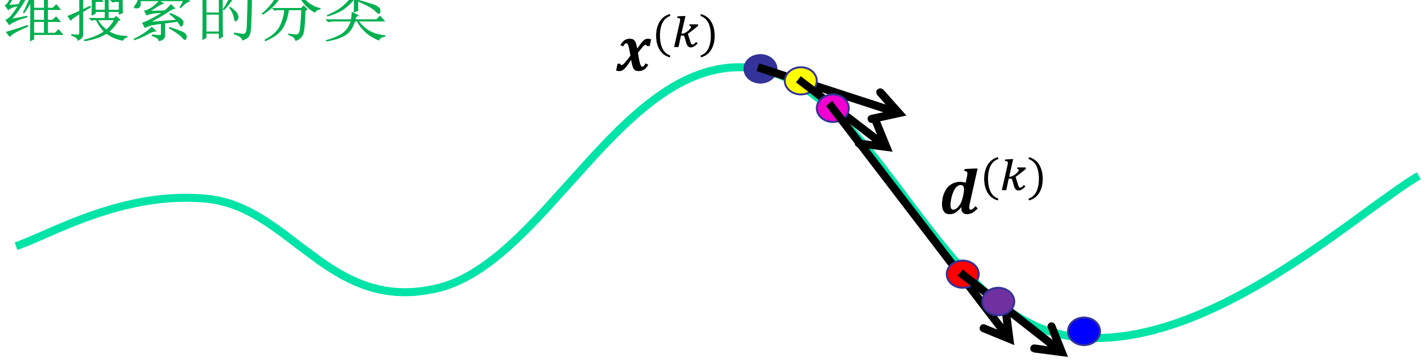
一维搜索的三要素

初始点 **搜索方向** **搜索步长**

不同的搜索方向确定方法衍生了不同的一维搜索算法

■ 一维搜索的概念

➢一维搜索的分类



一维搜索的方法很多，归纳起来，大体可分成两类：

■ 试探法

需要按某种方式找试探点，通过一系列试探点来确定极小点

■ 函数逼近法（插值法）

用某种较简单的曲线逼近本来的函数曲线，通过求逼近函数的极小点来估计目标函数的极小点

目录

- 一维搜索的概念（重）
- 试探法（考）
- 函数逼近法（考）

■ 试探法

- **0.618法（黄金分割法）**
- **Fibonacci 法**
- 二分法

■ 试探法

➢0.618法（黄金分割法）

0.618 法（黄金分割法）适用于单峰函数

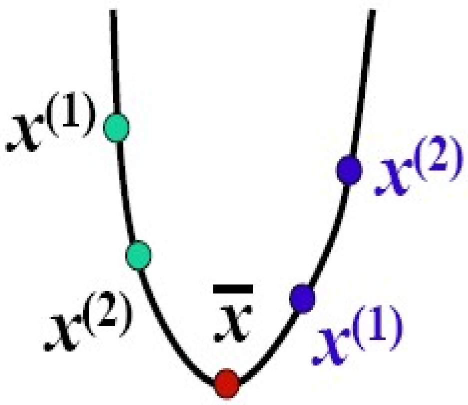
单峰函数

◆定义9．2．1 设f是定义在闭区间[a，b]上的一元实函数，$\bar{x}$是f在[a，b]上的极小点，并且对任意的$x^{(1)}$，$x^{(2)}\in$[a，b]，$x^{(1)}<x^{(2)}$,有

当 $x^{(2)}\leqslant\bar{x}$ 时，$f(x^{(1)})>f(x^{(2)})$

当 $\bar{x}\leqslant x^{(1)}$ 时，$f(x^{(2)})>f(x^{(1)})$

则称f是在闭区间[a，b]上的单峰函数



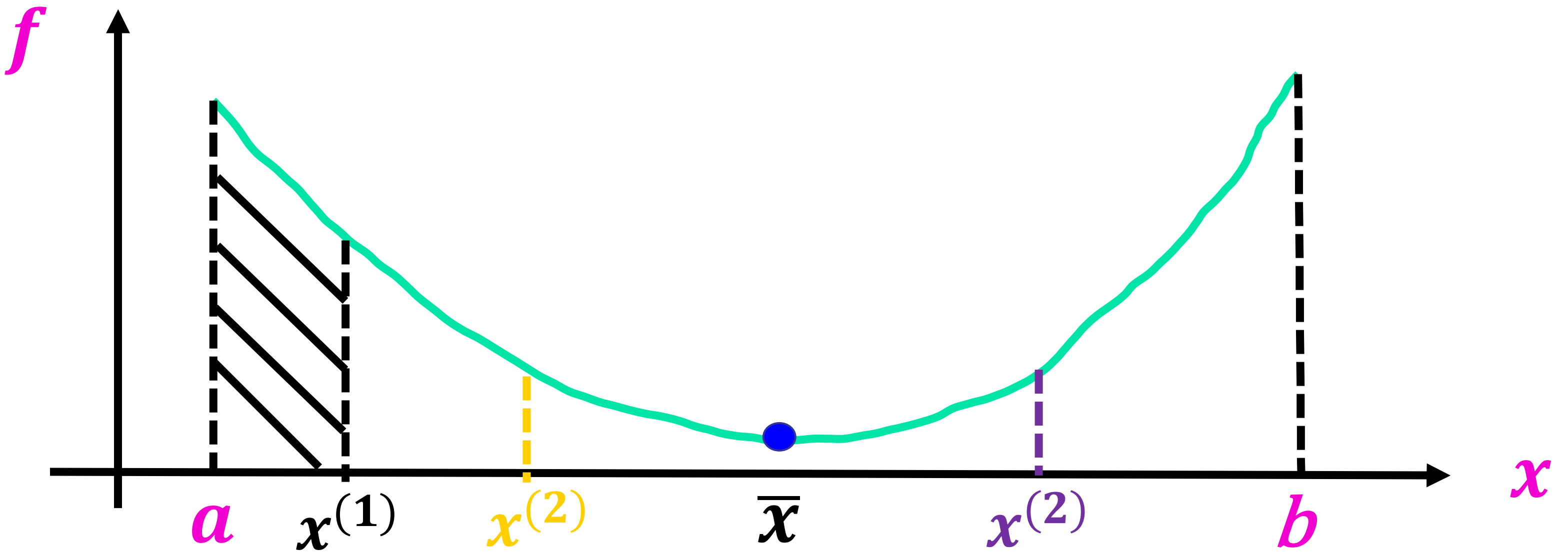
通过计算区间$[a,b]$内两个不同点处的函数值，就能确定一个包含极小点的子区间.

■ 试探法

➢0.618法（黄金分割法）

定理 9.2.1

设 f 是区间$[a, b]$上的单峰函数，$x^{(1)}$，$x^{(2)}\in[a, b]$，且 $x^{(1)}<x^{(2)}$.如果 $f(x^{(1)})>f(x^{(2)})$，则对每一个 $x\in[a, x^{(1)}]$，有 $f(x)>f(x^{(2)})$；如果 $f(x^{(1)})\leqslant f(x^{(2)})$，则对每一个 $x\in[x^{(2)}, b]$，有 $f(x)\geqslant f(x^{(1)})$.

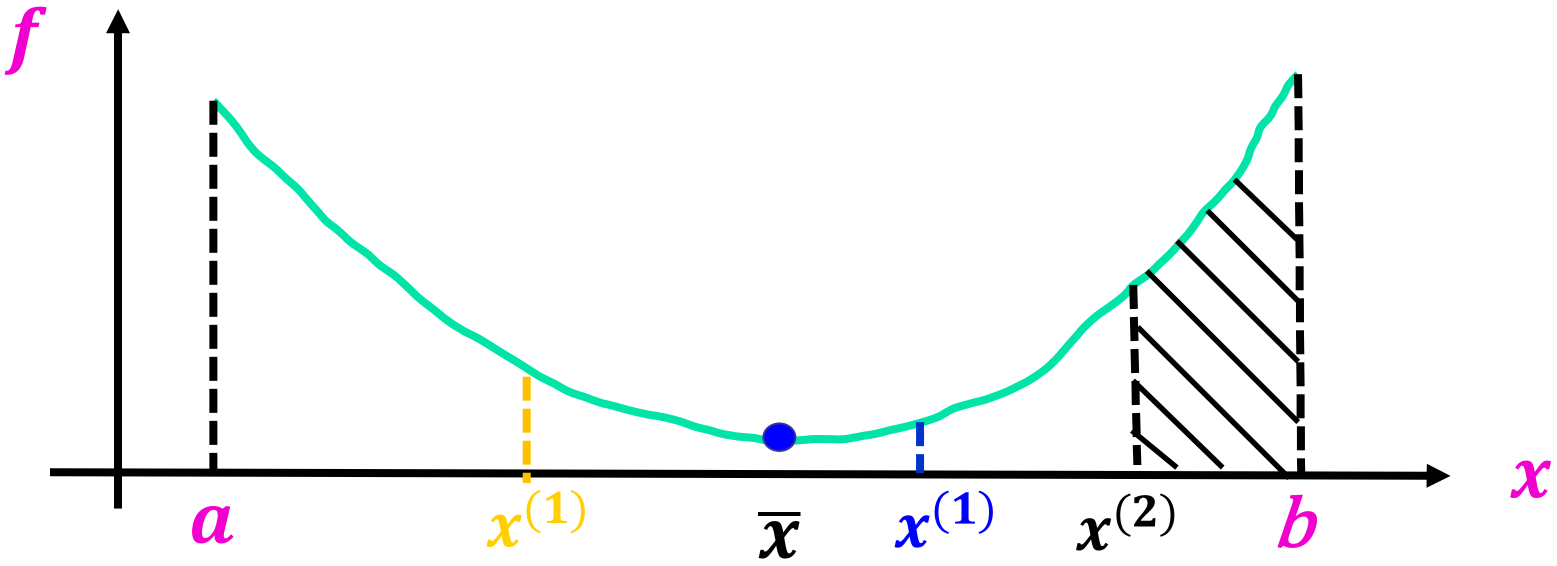


■ 试探法

➢0.618法（黄金分割法）

定理 9.2.1

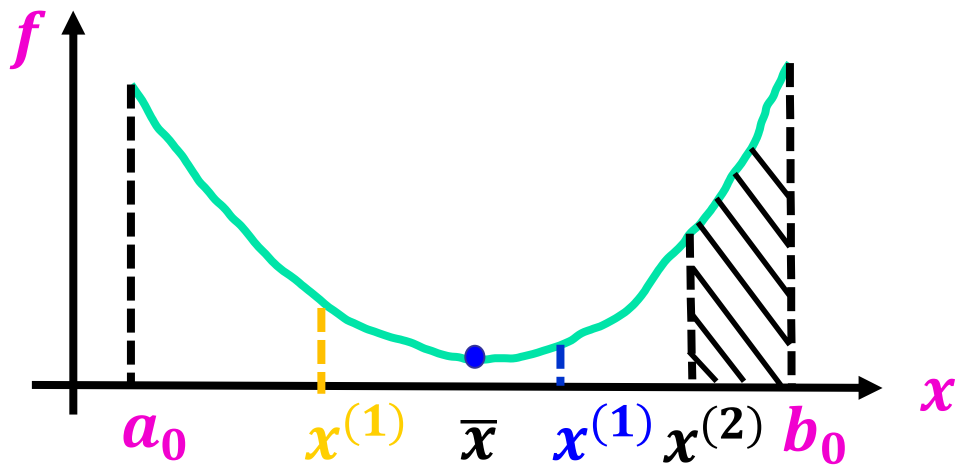
设 f 是区间 $[a, b]$ 上的单峰函数，$x^{(1)}$，$x^{(2)} \in [a, b]$，且 $x^{(1)} < x^{(2)}$。如果 $f(x^{(1)}) > f(x^{(2)})$，则对每一个 $x \in [a, x^{(1)}]$，有 $f(x) > f(x^{(2)})$；如果 $f(x^{(1)}) \leqslant f(x^{(2)})$，则对每一个 $x \in [x^{(2)}, b]$，有 $f(x) \geqslant f(x^{(1)})$.



■ 试探法

➢0.618法（黄金分割法）

0.618法的基本思想:



思路：挑选区间$[a_0, b_0]$中的点，计算对应的目标函数值，通过比较不断缩小极小点所在的区域。

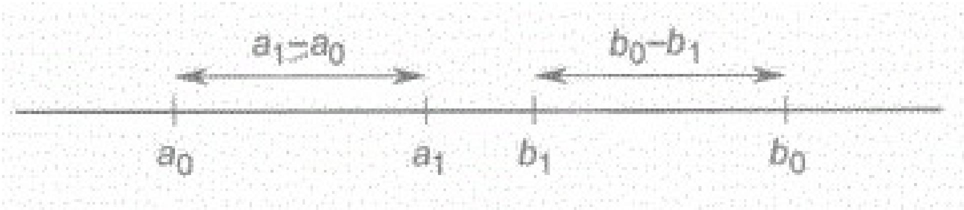
■**关键：如何选择合适的点？**

■ 试探法

➢0.618法（黄金分割法）

◆ **每次选择两个点，采用对称压缩方式缩小极小点所在区域。**

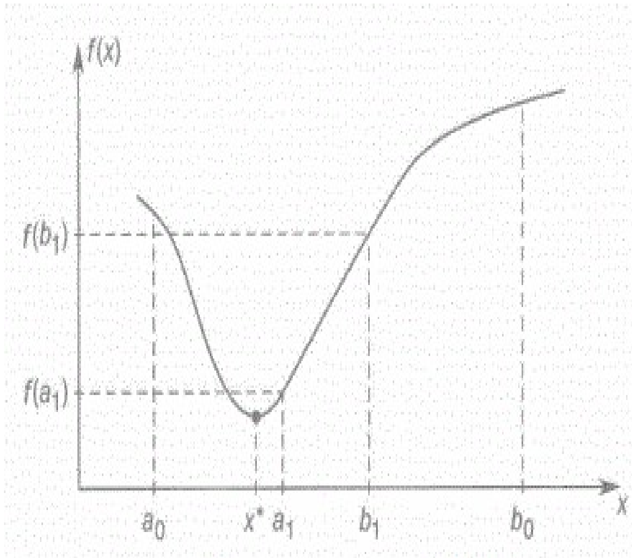
$$a_1 - a_0 = b_0 - b_1 = \rho(b_0 - a_0) \quad \rho < 1/2$$



◆ **缩小极小点区域：计算目标函数在这些中间点处的值，**

若 $f(a_1) < f(b_1)$，极小点应位于 $[a_0, b_1]$

若 $f(a_1) > f(b_1)$，极小点应位于 $[a_1, b_0]$。



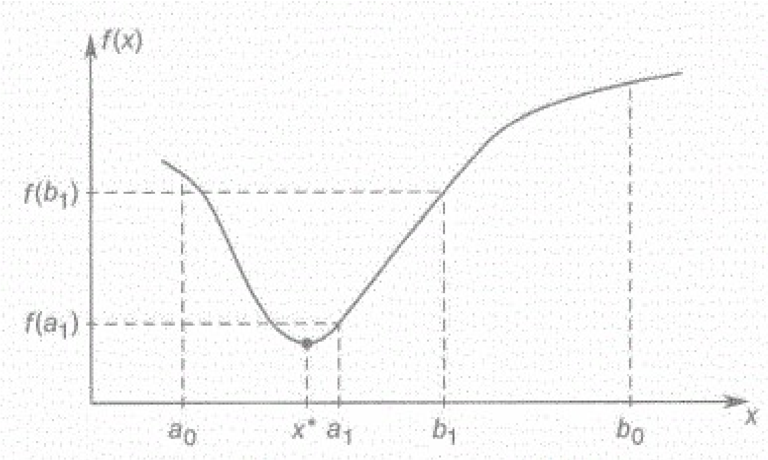
◆ **重复上述过程，确定两个新的中间点 a_2, b_2，重复计算。**

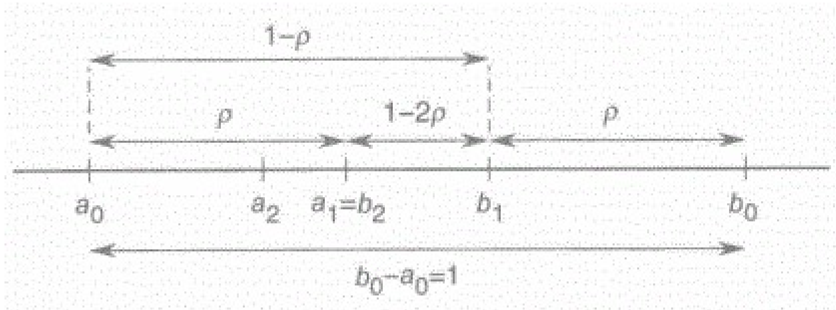
确定合适的参数 ρ

■ 试探法

➢0.618法（黄金分割法）

■**确定合适的参数** ρ **每次迭代只需计算一次目标函数 的值**





■**极小值从区域** $[a_0, b_0]$ **压缩至** $[a_0, b_1]$

■**由于** $a_1 \in [a_0, b_0]$**，且** $f(a_1)$ **已知,令** $b_2 = a_1$**，则有**

$$b_1 - b_2 = \rho(b_1 - a_0)$$
$$\Rightarrow 1 - 2\rho = \rho(1 - \rho)$$
$$\Rightarrow 1 - \rho = 0.618$$

■ 试探法

小端点不变，大端点更新，小探点计算，大探点继承

➢0.618法（黄金分割法）

◆ **黄金分割法在迭代中计算试探点**

$$\lambda_{k+1} = a_k + 0.382(b_k - a_k) \quad k = 0,1,\cdots,N$$

$$\mu_{k+1} = a_k + 0.618(b_k - a_k) \quad k = 0,1,\cdots,N$$

◆ **计算试探点的目标函数值并比较**

若 $f(\lambda_{k+1}) \le f(\mu_{k+1})$，**则** $a_{k+1} = a_k, b_{k+1} = \mu_{k+1}, \mu_{k+2} = \lambda_{k+1}$ ，

计算 $\lambda_{k+2} = a_{k+1} + 0.382(b_{k+1} - a_{k+1})$

若 $f(\lambda_{k+1}) > f(\mu_{k+1})$，**则** $a_{k+1} = \lambda_{k+1}, b_{k+1} = b_k, \lambda_{k+2} = \mu_{k+1}$ ，

计算 $\mu_{k+2} = a_{k+1} + 0.618(b_{k+1} - a_{k+1})$

区间按 $1-\rho=0.618$ **比例逐步压缩，每步只需计算一个新点。经过N步压缩后，总压缩比** $(0.618)^N$

■ 试探法

小端点不变，大端点更新，小探点计算，大探点继承

➢Fibonacci 法

■问题：搜索**单峰函数** f 在区间$[a_0, b_0]$的局部极小点。

黄金分割法中 ρ 始终保持不变 ⟶ ■参数 ρ 可调?

◆ 用于单峰函数

◆ 第一次迭代需要计算两个试探点，以后每次迭代只需新计算一点，另一点取自上次迭代。

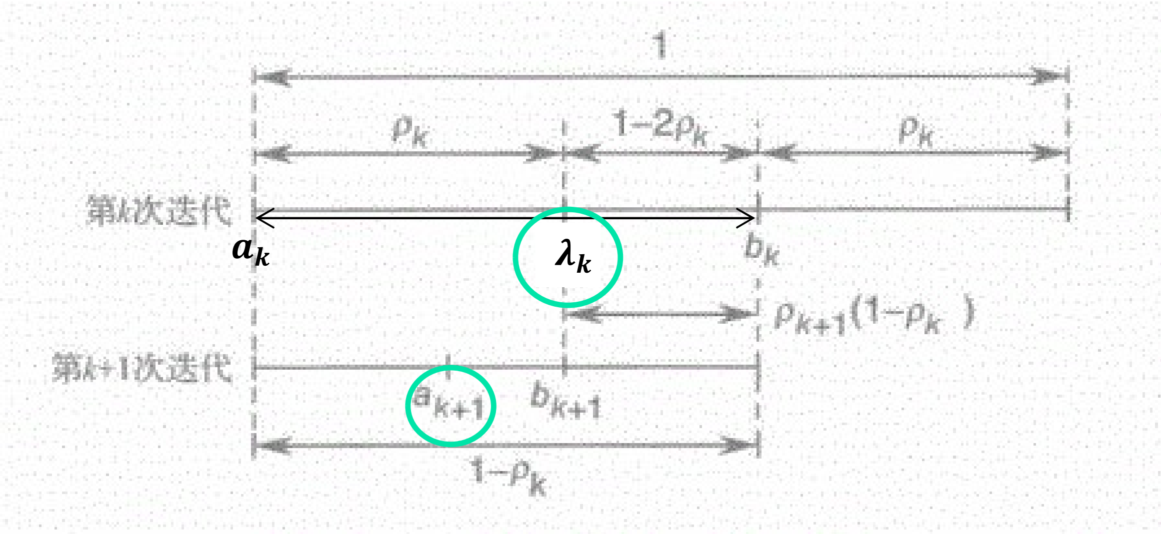
◆ 区间长度缩短比率 ρ 不是常数，而是由所谓的**Fibonacci数确定**

■ 试探法

小端点不变，大端点更新，小探点计算，大探点继承

➢Fibonacci 法

■区间长度压缩比可调 ➡ ■确定一个参数序列 ρ_k



$$a_{k+1}-a_k=b_k-b_{k+1}=\rho_{k+1}(b_k-a_k)$$

$$b_k-a_k=1-\rho_k \qquad b_k-b_{k+1}=1-2\rho_k$$

$$\rho_{k+1}(1-\rho_k)=1-2\rho_k \quad \Rightarrow \quad \rho_{k+1}=1-\frac{\rho_k}{1-\rho_k}$$

采用Fibonacci数列

■ 试探法

➢Fibonacci 法

$$\rho_{k+1}=1-\frac{\rho_k}{1-\rho_k}$$

Fibonacci数列

◆定义9.2.2 设有数列$\{F_k\}$，满足条件：

1. $F_0=F_1=1$
2. $F_{k+1}=F_k+F_{k-1}$ k=1，2，…

则称$\{F_k\}$为Fibonacci数列

k	0	1	2	3	4	5	6	7	8	9	10	…
F_k	1	1	2	3	5	8	13	21	34	55	89	…

■采用Fibonacci数列

$$\rho_1=1-\frac{F_N}{F_{N+1}}=\frac{F_{N-1}}{F_{N+1}}$$

$$\rho_2=1-\frac{F_{N-1}}{F_N}=\frac{F_{N-2}}{F_N}$$

$$\vdots$$

$$\rho_k=1-\frac{F_{N-k+1}}{F_{N-k+2}}=\frac{F_{N-k}}{F_{N-k+2}}$$

$$\vdots$$

$$\rho_N=1-\frac{F_1}{F_2}=\frac{F_0}{F_2}$$

◆ **区间长度缩短比率不是常数，而是由所谓的Fibonacci数确定**

■ 试探法

小端点不变，大端点更新，小探点计算，大探点继承

➢Fibonacci 法

◆ **Fibonacci法在迭代中计算试探点**

$$\lambda_{k+1} = a_k + \frac{F_{N-k-1}}{F_{N-k+1}}(b_k - a_k) \quad k = 0,1,\cdots,N$$

$$\mu_{k+1} = a_k + \frac{F_{N-k}}{F_{N-k+1}}(b_k - a_k) \quad k = 0,1,\cdots,N$$

◆ **计算试探点的目标函数值并比较**

若 $f(\lambda_{k+1}) \le f(\mu_{k+1})$，**则** $a_{k+1} = a_k, b_{k+1} = \mu_{k+1}, \mu_{k+2} = \lambda_{k+1}$，

计算 $\lambda_{k+2} = a_{k+1} + \frac{F_{N-k-2}}{F_{N-k}}(b_{k+1} - a_{k+1})$

若 $f(\lambda_{k+1}) > f(\mu_{k+1})$，**则** $a_{k+1} = \lambda_{k+1}, b_{k+1} = b_k, \lambda_{k+2} = \mu_{k+1}$，

计算 $\mu_{k+2} = a_{k+1} + \frac{F_{N-k-1}}{F_{N-k}}(b_{k+1} - a_{k+1})$

总压缩比 $(1-\rho_1)(1-\rho_1)\cdots(1-\rho_N) = \frac{F_N}{F_{N+1}}\frac{F_{N-1}}{F_N}\cdots\frac{F_1}{F_2} = \frac{1}{F_{N+1}}$

■ 试探法

➢二分法

探点导数大于0，大端点更新，探点导数小于0，小端点更新

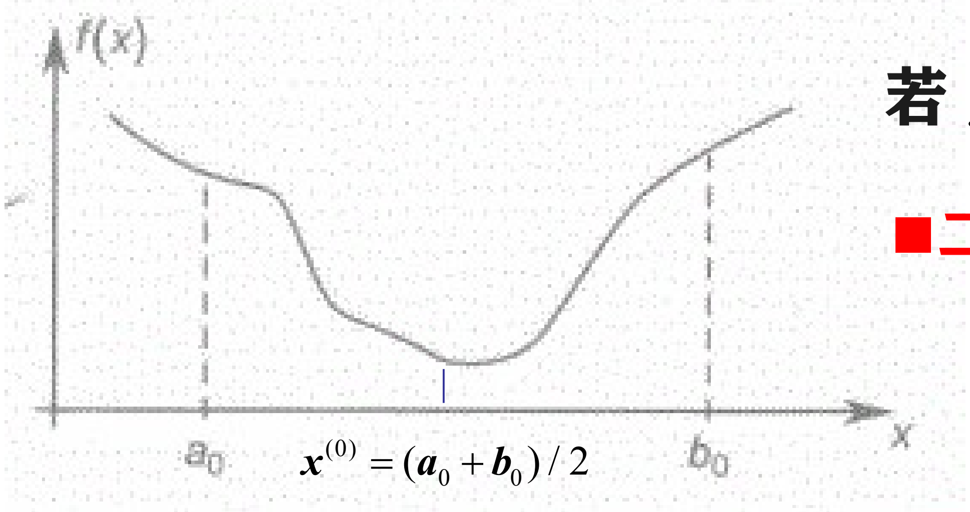
用途：搜索单峰函数 f 在区间 $[a_0, b_0]$ 的局部极小点。

要求：函数 f 连续可微

- **确定初始区间的中点** $x^{(0)} = (a_0 + b_0)/2$
- **计算函数 f 在 $x^{(0)}$ 处的一阶导数** $f'(x^{(0)})$

若 $f'(x^{(0)}) > 0$，极小点位于中点左侧，最优区间为 $\left[a_0, x^{(0)}\right]$

若 $f'(x^{(0)}) < 0$，极小点位于中点右侧，最优区间为 $\left[x^{(0)}, b_0\right]$



若 $f'(x^{(0)}) = 0$，极小点为 $x^{(0)}$

■二分法的总压缩比

$$(1/2)^N$$

■ 试探法

➢总结

- 黄金分割法（0.618法）和Fibonacci数列法使用**目标函数值**，二分法使用**目标函数的导数**
- 0.618可作为Fibonacci法的极限形式

  当 $N \geq 7$ 时，$F_{N-1}/F_N \approx 0.618$
- 理论上**Fibonacci法精度高于黄金分割法**
- Fibonacci法**需要事先知道计算函数的次数N**，黄金分割法（0.618法）更简单

目录

■ 函数逼近法

- ➢牛顿法
- ➢割线法
- ➢抛物线法
- ➢三次插值法

■ 函数逼近法

➢牛顿法

基本思想：在极小点附近用二阶Taylor多项式近似目标函数 $f(x)$ ，进而求出极小点的估计值。

前提：函数 $f(x)$ 连续二阶可微

函数 $f(x)$ 的近似

$$\varphi(x) = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2} f''(x^{(k)})(x - x^{(k)})^2$$

近似函数的导函数

$$\varphi'(x) = f'(x^{(k)}) + f''(x^{(k)})(x - x^{(k)})$$

求近似函数的极小点?

令 $\varphi'(x) = 0$ ，**求得驻点** $x^{(k+1)} = x^{(k)} - \dfrac{f'(x^{(k)})}{f''(x^{(k)})}$

当 $\left| x^{(k+1)} - x^{(k)} \right| \le \varepsilon$ **时，算法迭代停止。**

■ 函数逼近法

➢牛顿法

牛顿法的计算步骤如下：

(1) 给定初点 $x^{(0)}$，允许误差 $\varepsilon>0$，置 $k=0$.

(2) 若 $|f'(x^{(k)})|<\varepsilon$，则停止迭代，得到点 $x^{(k)}$.

(3) 计算点 $x^{(k+1)}$，

$$x^{(k+1)}=x^{(k)}-\frac{f'(x^{(k)})}{f''(x^{(k)})}$$

置 $k:=k+1$，转步骤(2).

运用牛顿法时，初点选择十分重要. 如果初始点靠近极小点，则可能很快收敛；如果初始点远离极小点，迭代产生的点列可能不收敛于极小点.

■ 函数逼近法

➢牛顿法

利用牛顿法求解如下函数的极小点

$$f(x)=\frac{1}{2}x^2-\sin x$$

初始值为 $x^{(0)}=0.5$，精度为 $\varepsilon=10^{-5}$ 。

解：函数的一阶和二阶导数

$$f'(x)=x-\cos x \qquad f''(x)=1+\sin x$$

由此 $$x^{(1)}=0.5-\frac{0.5-\cos 0.5}{1+\sin 0.5}=0.7552$$

$$x^{(2)}=x^{(1)}-\frac{f'(x^{(1)})}{f''(x^{(1)})}=0.7552-\frac{0.0271}{1.685}=0.7391$$

$$x^{(3)}=x^{(2)}-\frac{f'(x^{(2)})}{f''(x^{(2)})}=0.7391-\frac{9.461\times10^{-5}}{1.673}=0.7390$$

$$x^{(4)}=x^{(3)}-\frac{f'(x^{(3)})}{f''(x^{(3)})}=0.7390-\frac{1.17\times10^{-9}}{1.673}=0.7390$$

■ 函数逼近法

➢割线法

如果函数的二阶导不存在，采用不同点处的一阶导对其近似

$$f''(x^{(k)}) = \frac{f'(x^{(k)}) - f'(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}$$

代入牛顿法迭代公式

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})}$$

得到新的迭代公式

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{f'(x^{(k)}) - f'(x^{(k-1)})} f'(x^{(k)})$$

或

$$x^{(k+1)} = \frac{f'(x^{(k)})x^{(k-1)} - f'(x^{(k-1)})x^{(k)}}{f'(x^{(k)}) - f'(x^{(k-1)})}$$

■ 函数逼近法

➢割线法

割线法的计算步骤如下：

(1) 给定初点 $x^{(0)}$，允许误差 $\varepsilon>0$，置 $k=0$.

(2) 若 $|f'(x^{(k)})|<\varepsilon$，则停止迭代，得到点 $x^{(k)}$.

(3) 计算点 $x^{(k+1)}$，

$$x^{(k+1)}=x^{(k)}-\frac{x^{(k)}-x^{(k-1)}}{f'(x^{(k)})-f'(x^{(k-1)})}f'(x^{(k)})$$

置 $k:=k+1$，转步骤(2).

割线法与牛顿法相比，收敛速率较慢，但不需要计算二阶导数. 它的缺点与牛顿法有类似之处，都不具有全局收敛性，如果初点选择得不好，可能不收敛

■ 函数逼近法

➢抛物线法

算法思路：在极小点附近，用二次三项式 $\varphi(x)$ 逼近目标函数 $f(x)$

$$\varphi(x)=a+bx+cx^2$$

令 $\varphi(x)$ 与函数 $f(x)$ 在三点处 $x^{(1)}<x^{(2)}<x^{(3)}$ 有相同的函数值

$$\begin{cases} a+bx^{(1)}+c(x^{(1)})^2=f(x^{(1)}) \\ a+bx^{(2)}+c(x^{(2)})^2=f(x^{(2)}) \\ a+bx^{(3)}+c(x^{(3)})^2=f(x^{(3)}) \end{cases}$$

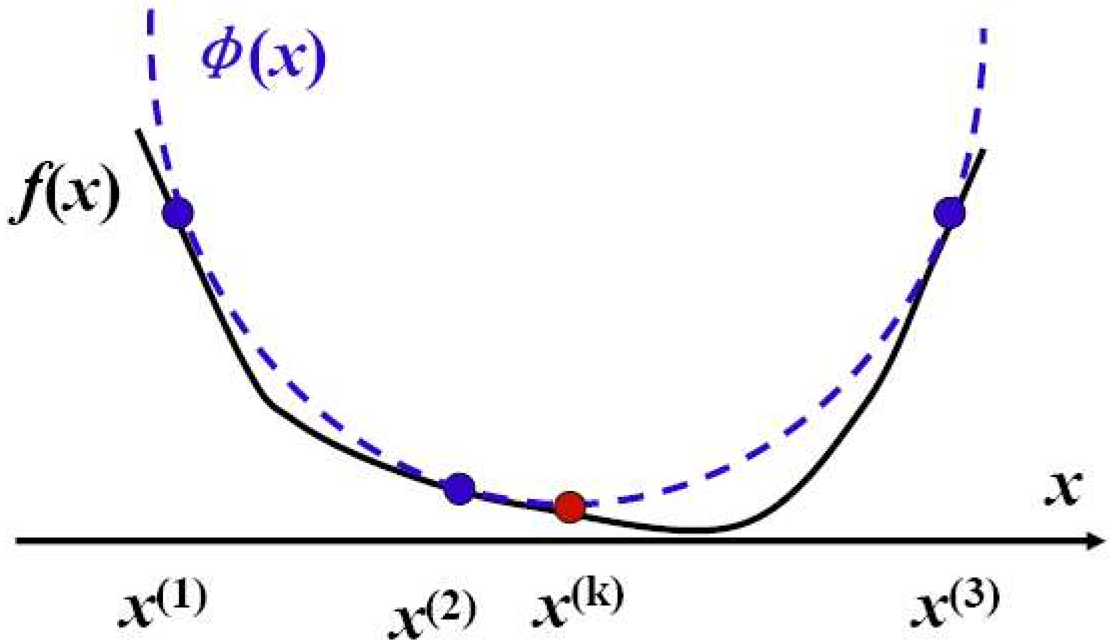
联立求解系数 b 和 c，得到近似函数 $\varphi(x)$ 。

$$b=\frac{B_1+B_2+B_3}{D}, c=-\frac{C_1+C_2+C_3}{D}$$

由最优性条件 $\varphi'(x)=0 \Rightarrow b+2cx=0 \Rightarrow \boxed{x=-b/2c}$

■ 函数逼近法

➢抛物线法



令 $x^k = -b/2c$ ，从 $x^{(1)}, x^{(2)}, x^{(3)}, x^k$ 中选出目标函数值最小的点及其左右两点，构造新的方程组求解系数，计算新的估计值 x^{k+1} ，上述过程迭代，直到收敛 $\left|x^{k+1} - x^k\right| \le \varepsilon$

■ 函数逼近法

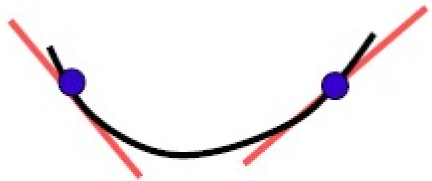
➢三次插值法

算法思路：在极小点附近，用三次多项式 $\varphi(x)$ **逼近目标函数 f (x)** $\varphi(x)=a(x-x^{(1)})^3+b(x-x^{(1)})^2+c(x-x^{(1)})+d$

令 $\varphi(x)$ **与函数** $f(x)$ **在两点处 有相同的函数值及相同的导数**

选择初始点 $x^{(1)}<x^{(2)}$ **，使得** $f'(x^{(1)})<0,\ f'(x^{(2)})>0$

$$\begin{cases}\varphi(x^{(1)})=f(x^{(1)})\\ \varphi'(x^{(1)})=f'(x^{(1)})\\ \varphi(x^{(2)})=f(x^{(2)})\\ \varphi'(x^{(2)})=f'(x^{(2)})\end{cases}$$



联立求解系数 a, b, c, d**，得到多项式** $\varphi(x)$ **。**

用多项式 $\varphi(x)$ **的极小点估计** $f(x)$ **的极小点。**

■ 作业

P280：

2