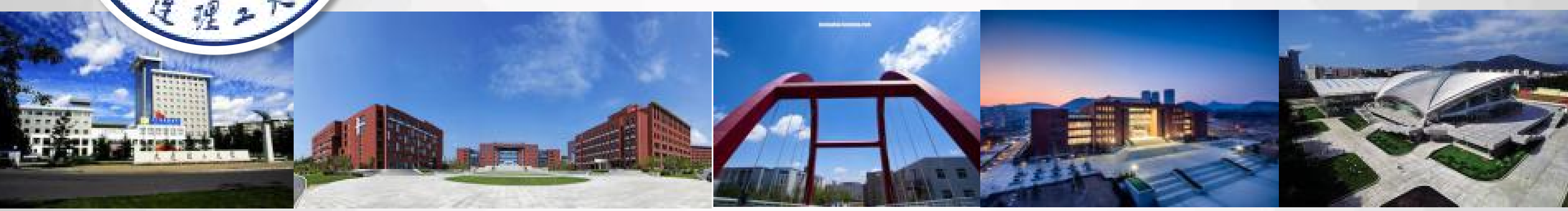
最优化理论
Optimality Theory



目 录 (CONTENT)

09

PART NINE

一维搜索

One-Dimensional Search

■ 主要内容

- 一维搜索的基本概念
- 试探法
- 函数逼近法

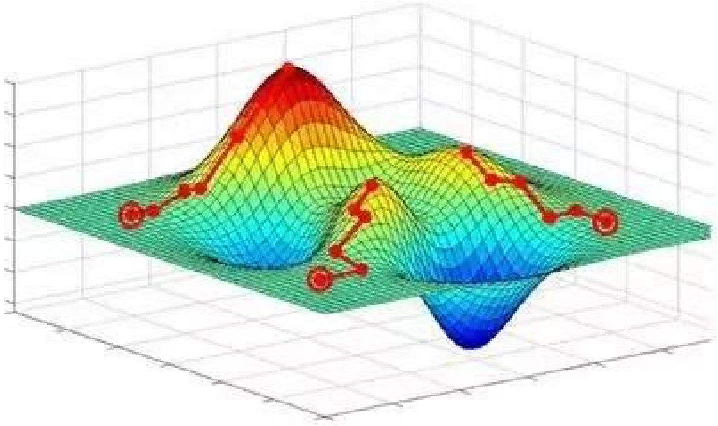
■ 基本概念

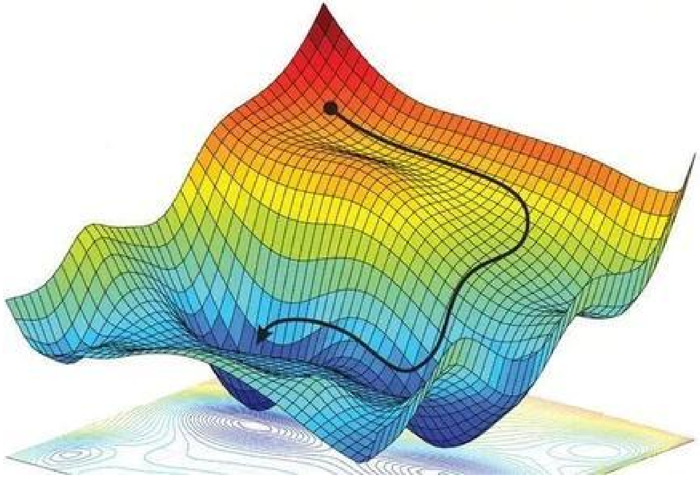
- 最优化问题可归结成如下数学形式:

$$\min_{x\in R^n} f(x) \qquad \text{---目标函数}$$

$$s.t. \qquad g_i(x)\ge 0, i\in I$$

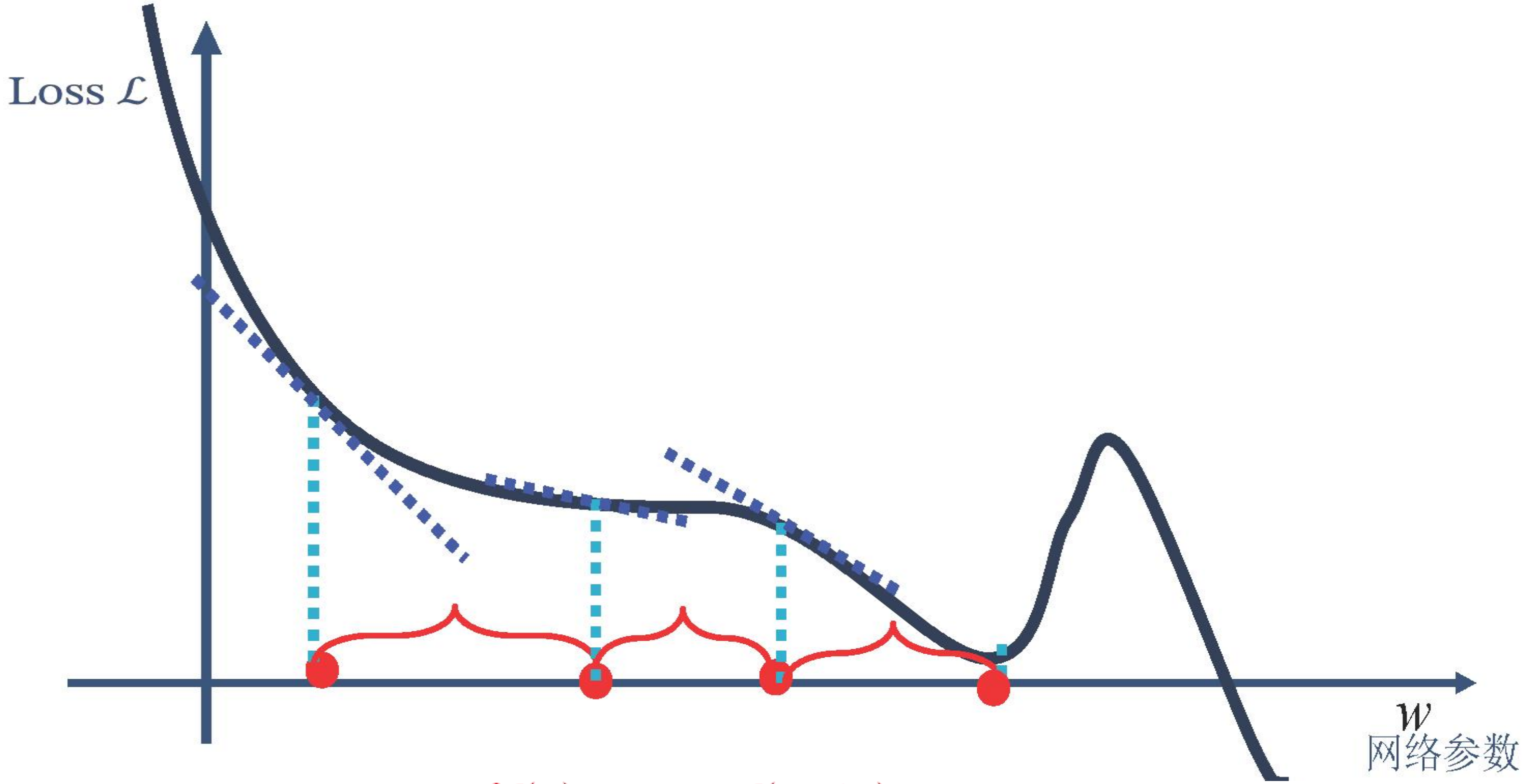
$$h_j(x)=0, j\in E$$





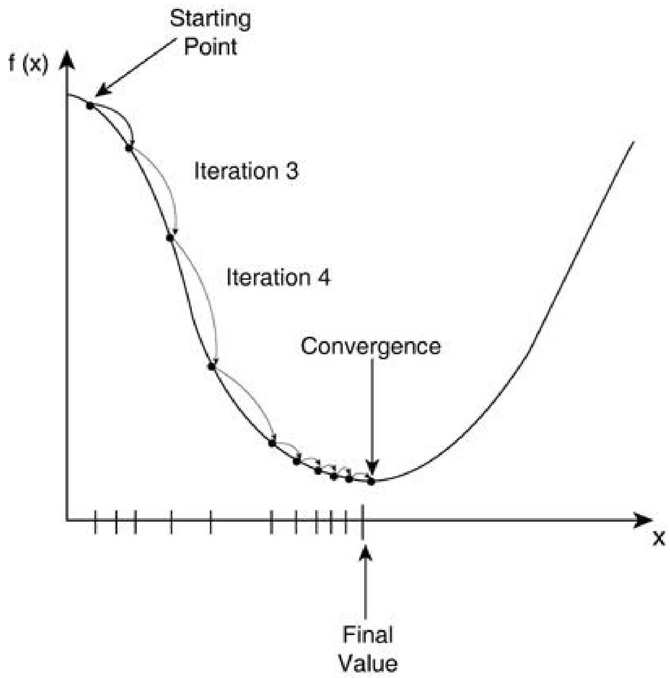
■ 梯度下降 (Gradient Descent)

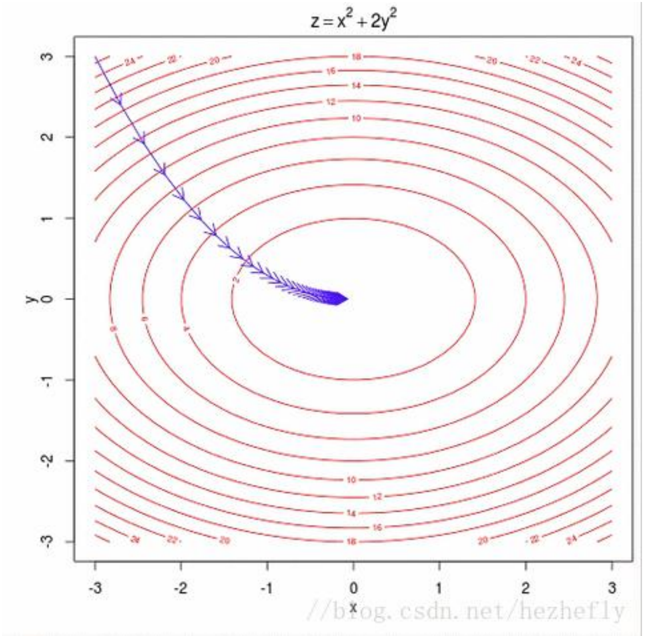
➢ 梯度下降类似盲人下山



梯度: $\frac{\partial f(w)}{\partial w} = \lim_{\Delta w \to 0} \frac{f(w+\Delta w)}{\Delta w}$

■ 梯度下降法(Gradient Descent)





$$\theta_{t+1} = \theta_t - \alpha \frac{\partial \mathcal{R}(\theta)}{\partial \theta_t}$$
$$= \theta_t - \alpha \frac{1}{N} \sum_{i=1}^{N} \frac{\partial \mathcal{L}(\theta_t; x^{(i)}, y^{(i)})}{\partial \theta}.$$

搜索步长α中也叫作学习率(Learning Rate)

■ 概念

迭代优化方法的基本结构：

给定初始点$\mathbf{x}_0$

(a) 确定搜索方向$\mathbf{d}_k$，即按照一定规则，构造f在$\mathbf{x}_k$点处的下降方向作为搜索方向;

(b) 确定步长因子α_k，使目标函数值有某种意义下的下降;

(c) 令 $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$

若$\mathbf{x}_{k+1}$满足某种终止条件则停止迭代，得到近似最优解$\mathbf{x}_{k+1}$，否则，重复上述步骤。

注意到上述迭代算法中,当方向确定后,涉及到求一个步长α_k,使得目标函数值减小(极小化问题),这就是在一直线上求目标函数的极小点,即极小化$f(x_k+\alpha_k d_k)$.这称为α对变量的一维搜索问题,或称为线搜索.

■ 概念

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

$$L=\{x\,|\,x=x^{(k)}+\lambda d^{(k)},-\infty<\lambda<+\infty\} \qquad (9.1.1)$$

求$f(x)$在直线L上的极小点就转化为求

$$\varphi(\lambda)=f(x^{(k)}+\lambda d^{(k)}) \qquad (9.1.2)$$

的极小点.

设$\varphi(\lambda)$的极小点为λ_k,称λ_k为沿方向$d^{(k)}$的步长因子

于是$f(x)$在直线L上的极小点为

$$x^{(k+1)}=x^{(k)}+\lambda_k d^{(k)} \qquad (9.1.3)$$

大连理工大学 Dalian University Of Technology

■ 代表性方法

一维搜索 $\left\{ \begin{array}{l} \text{试探法} \\ \text{函数逼近法/插值法} \end{array} \right.$

- **一维搜索算法的闭性**

假设一维搜索是以x为起点,沿方向为d的进行的,并定义为**算法映射M**

$Df\,9.1.1$ 算法映射$M: \mathrm{R}^n \times \mathrm{R}^n \to \mathrm{R}^n$定义为

$$M(x,d) = \{y \mid y = x + \bar{\lambda} d, \bar{\lambda}\text{满足}$$

$$f(x + \bar{\lambda} d) = \min_{\lambda \ge 0} f(x + \lambda d)\} \qquad (9.1.4)$$

■ 一维搜索算法的闭性

Th9.1.1 设f是定义在R^n的连续函数,$d\neq 0$,则(9.1.4)定义的算法映射M在(x,d)处是闭的

证：设序列$\{x^{(k)}\}$和$\{d^{(k)}\}$满足

$$(x^{(k)},d^{(k)})\to(x,d);\quad y^{(k)}\to y, y^{(k)}\in M(x^{(k)},d^{(k)})$$

下证 $y\in M(x,d)$，注意到,对每个k, $\lambda_k\geq 0$，使

$$y^{(k)}=x^{(k)}+\lambda_k d^{(k)} \qquad (9.1.5)$$

由$d\neq 0$,当k充分大时,必有$d^{(k)}\neq 0$, 于是由(9.1.5)

$$\lambda_k=\frac{\left\|y^{(k)}-x^{(k)}\right\|}{\left\|d^{(k)}\right\|} \qquad (9.1.6)$$

■ 一维搜索算法的闭性

令$k \to \infty$,则$\lambda_k \to \bar{\lambda} = \dfrac{\|y-x\|}{\|d\|}$ (9.1.7)

(9.1.5)中令$k \to \infty$并注意到(9.1.7),有

$$y = x + \bar{\lambda} d \qquad (9.1.8)$$

根据M的定义,对每个k及$\lambda_k \geq 0$,有

$$f(y^{(k)}) \leq f(x^{(k)} + \lambda_k d^{(k)}) \qquad (9.1.9)$$

由于f连续,令$k \to \infty$,则由(9.1.9)得

$$f(y) \leq f(x + \lambda d)$$

故 $$f(x + \bar{\lambda} d) = \min_{\lambda \geq 0} f(x + \lambda d)$$

即知 $$y \in M(x, d)$$

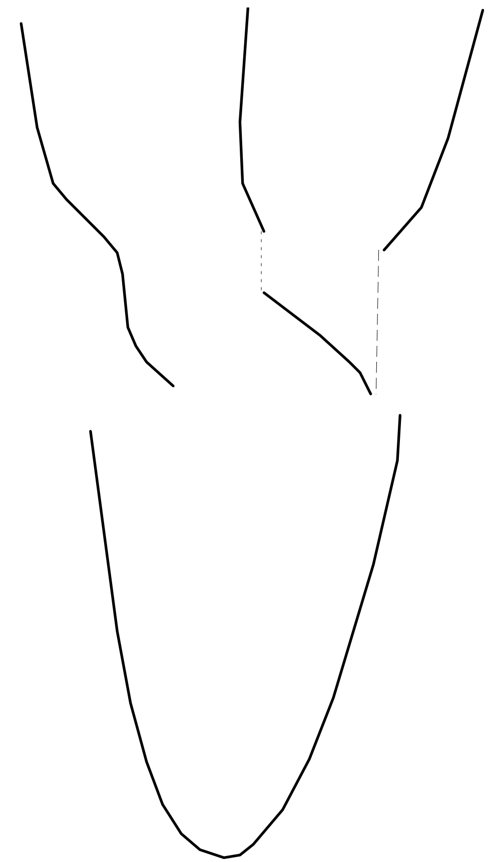
■ 0.618法

Df 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,且对$\forall x^{(1)}, x^{(2)} \in [a,b]$,$x^{(1)} < x^{(2)}$,有

$$\text{当} x^{(2)} \leq \bar{x} \text{时} f(x^{(1)}) > f(x^{(2)})$$

$$\text{当} \bar{x} \leq x^{(1)} \text{时} f(x^{(1)}) < f(x^{(2)})$$

则称f是在闭区间$[a, b]$上的单峰函数.



■ 0.618法

单峰函数的一个等价定义：

设$\boldsymbol{f}:\boldsymbol{R}\to\boldsymbol{R},[\boldsymbol{a},\boldsymbol{b}]\subset\boldsymbol{R}$,若$\exists\alpha^*\in[\boldsymbol{a},\boldsymbol{b}]$, 使得$\boldsymbol{f}(\boldsymbol{x})$在$[\boldsymbol{a},\alpha^*]$上严格递减,在$[\alpha^*,\boldsymbol{b}]$上严格递增,则称$[\boldsymbol{a},\boldsymbol{b}]$是函数$\boldsymbol{f}(\boldsymbol{x})$的单峰区间, $\boldsymbol{f}(\boldsymbol{x})$是$[\boldsymbol{a},\boldsymbol{b}]$上的单峰函数.

单峰函数具有一些很有用的性质：

如果f是$[a,b]$上单峰函数, 则可通过计算此区间内两不同点的函数值, 就能确定一个包含极小点的子区间, 从而缩小了搜索区间.

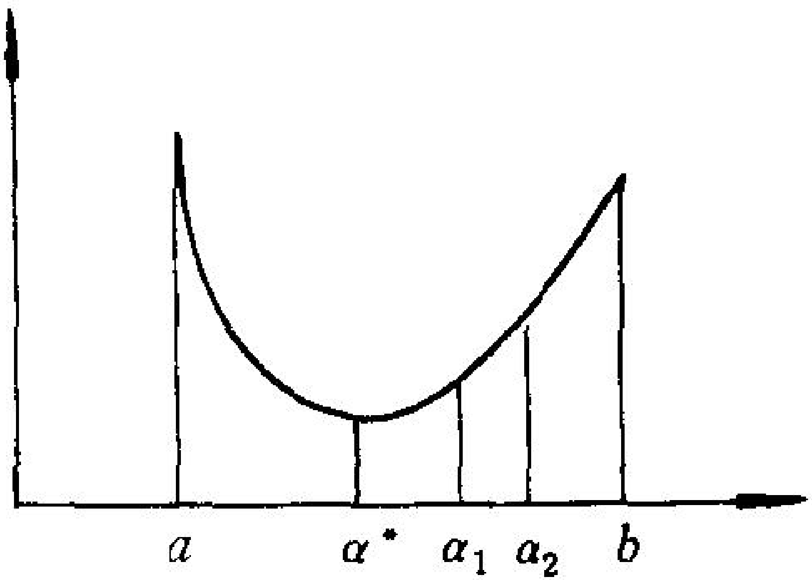
■ 0.618法

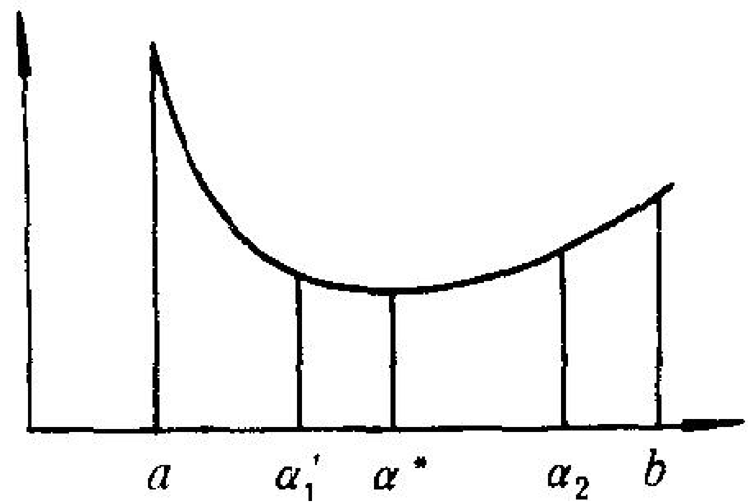
*Th*9.2.1 设f是区间$[a,b]$上的单峰函数,$x^{(1)},x^{(2)}\in[a,b]$.

且$x^{(1)}<x^{(2)}$,则

(1)若$f(x^{(1)})>f(x^{(2)})$,则$\forall x\in[a,x^{(1)}],f(x)>f(x^{(2)})$

(2)若$f(x^{(1)})\le f(x^{(2)})$,则$\forall x\in[x^{(2)},b],f(x)\ge f(x^{(1)})$,





■ 0.618法

证明:仅证(1),反证,如若不然,存在点$x^*\in[a, x^{(1)}]$,使

$$f(x^*)\le f(x^{(2)})$$

显然$x^{(1)}$不是极小点.此时要么极小点$\bar{x}\in[a,x^{(1)}]$

要么$\bar{x}\in[x^{(1)},b]$.

若$\bar{x}\in[a,x^{(1)}]$,则$f(x^{(1)})<f(x^{(2)})$,矛盾.

若$\bar{x}\in[x^{(1)},\mathrm{b}]$,则$f(x^{*})>f(x^{(1)})>f(x^{(2)})$,矛盾.

根据以上定理,只需选择两个点就可缩短包含极小点的

区间: (1)若$f(x^{(1)})>f(x^{(2)})$,则极小点$\bar{x}\in[x^{(1)},\mathrm{b}]$;

(2)若$f(x^{(1)})\le f(x^{(2)})$,则极小点$\bar{x}\in[a,x^{(2)}]$.

■ 0.618法

0.618法的基本思想: 通过取试探点使包含极小点的区间(不确定区间)不断缩小,当区间长度小到一定程度时,区间上各点的函数值均接近极小值,此时该区间内任一点都可以作为极小点的近似值.

设 $\varphi(\alpha)$是搜索区间$[a_1,b_1]$上的单峰函数. 设在第k次迭代时搜索区间为$[a_k,b_k]$.取两个试探点$\lambda_k,\mu_k\in[a_k,b_k]$,$\lambda_k<\mu_k$.

计算$\varphi(\lambda_k)$和$\varphi(\mu_k)$,根据$Th\,9.2.1$:

(1),若$\varphi(\lambda_k)\le\varphi(\mu_k)$,令$a_{k+1}=a_k, b_{k+1}=\mu_k$,   (2.1)

(2),若$\varphi(\lambda_k)>\varphi(\mu_k)$,令$a_{k+1}=\lambda_k, b_{k+1}=b_k$,   (2.2)

大连理工大学 Dalian University Of Technology

■ 0.618法

我们要求两个试探点λ_k和μ_k满足下列条件：

(1),λ_k和μ_k到搜索区间$[a_k, b_k]$的端点等距,即

$$b_k - \lambda_k = \mu_k - a_k \ , \qquad (2.3)$$

(2),每次迭代,搜索区间长度的缩短率相同,即

$$b_{k+1} - a_{k+1} = r(b_k - a_k), \qquad (2.4)$$

由(2.3)和(2.4)得到

$$\lambda_k = a_k + (1 - \mathrm{r})(b_k - a_k) \ , \qquad (2.5)$$

$$\mu_k = a_k + r(b_k - a_k), \qquad (2.6)$$

大连理工大学 Dalian University Of Technology

■ 0.618法

考虑(2.1)的情形,此时新的搜索区间为

$$[a_{k+1},b_{k+1}]=[a_k,\mu_k] \tag{2.7}$$

为进一步缩短区间. 需取试探点λ_{k+1}和μ_{k+1}. 由(2.6)

$$\begin{aligned}\mu_{k+1} &= a_{k+1}+r(b_{k+1}-a_{k+1})\\ &=a_k+\mathrm{r}(\mu_k-a_k)\\ &=a_k+\mathrm{r}(a_k+\mathrm{r}(b_k-a_k)-a_k), \qquad (2.8)\\ &=a_k+\mathrm{r}^2(b_k-a_k)\end{aligned}$$

大连理工大学 Dalian University Of Technology

■ 0.618法

若令 $\tau^2=1-\tau$ (2.9)

则 $\mu_{k+1}=a_k+(1-\tau)(b_k-a_k)=\lambda_k$ (2.10)

这样新的试探点μ_{k+1}就不用重新计算只要取λ_k，于是每次迭代中(除第一次)只需取一个试探点.

类似的, 如考虑(2.2)的情形, 新的试探点$\lambda_{k+1}=\mu_k$, 它也不需重新计算.

解方程(2.9)立得区间长度缩短率$\tau=\dfrac{-1\pm\sqrt{5}}{2}$

由于$\tau>0$, 故取$\tau=\dfrac{-1+\sqrt{5}}{2}\approx 0.618$ (2.11)

■ 0.618法

这样,计算公式(2.5)(2.6)可写为

$$\lambda_k = a_k + 0.382(b_k - a_k)\ , \qquad (2.12)$$

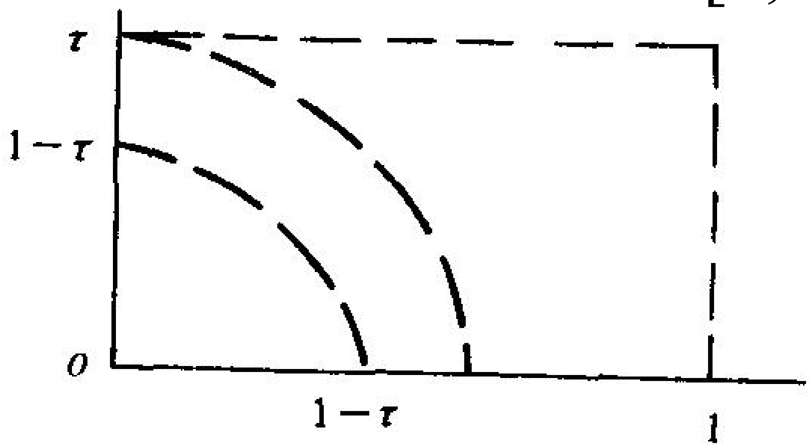
$$\mu_k = a_k + 0.618(b_k - a_k), \qquad (2.13)$$

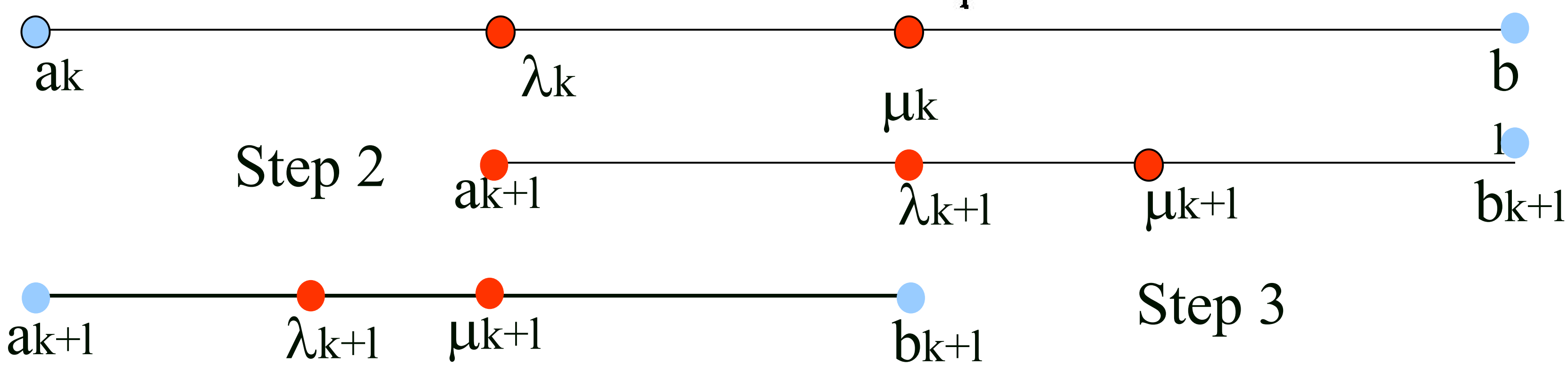
由于每次函数计算后极小区间的缩短率为r,故若初始区间为$[a_1, b_1]$,则最终区间长度为$r^{n-1}(b_1 - a_1)$,因此可知0.618法是线性收敛的。

0.618法也叫黄金分割法,因为缩短率r叫黄金分割数,它满足比率$\frac{r}{1} = \frac{1-r}{r}$,即$\tau^2 + \tau - 1 = 0$。

■ 0.618法

几何意义:黄金分割率τ对应的点在单位长区间[0,1]中的位置相当于其对称点1-τ在区间[0,τ]中的位置





■ 0.618法

算法(0.618法)

步1:选取初始数据,确定初始搜索区间$[a_1,b_1]$和精度要求$\delta>0$.计算最初两试探点λ_1,μ_1:

$$\lambda_1 = \boldsymbol{a}_1 + 0.382(\boldsymbol{b}_1 - \boldsymbol{a}_1),$$

$$\mu_1 = \boldsymbol{a}_1 + 0.618(\boldsymbol{b}_1 - \boldsymbol{a}_1),$$

计算$\varphi(\lambda_1)$和$\varphi(\mu_1)$.

步2:比较函数值.若$\varphi(\lambda_k) > \varphi(\mu_k)$,则转步3,否则,$\varphi(\lambda_k) \le \varphi(\mu_k)$转步4

■ 0.618法

步3: 若$\boldsymbol{b}_k-\lambda_k\le\delta$,停止计算,输出$\mu_k$;否则,令

$$\boldsymbol{a}_{\mathrm{k+1}}:=\lambda_k,\boldsymbol{b}_{\mathrm{k+1}}:=\boldsymbol{b}_k,\lambda_{k+1}:=\mu_k,$$

$$\varphi(\lambda_{\mathrm{k+1}})=\varphi(\mu_k),\mu_{\mathrm{k+1}}=\boldsymbol{a}_{\mathrm{k+1}}+0.618(\boldsymbol{b}_{\mathrm{k+1}}-\boldsymbol{a}_{\mathrm{k+1}}),$$

计算$\varphi(\mu_{k+1})$转步5

步4:若$\mu_k-\boldsymbol{a}_k\le\delta$,停止计算,输出$\lambda_k$;否则,令

$$\boldsymbol{a}_{\mathrm{k+1}}:=\boldsymbol{a}_k,\boldsymbol{b}_{\mathrm{k+1}}:=\mu_k,\mu_{k+1}:=\lambda_k,$$

$$\varphi(\mu_{\mathrm{k+1}}):=\varphi(\lambda_k),\lambda_{\mathrm{k+1}}=\boldsymbol{a}_{\mathrm{k+1}}+0.382(\boldsymbol{b}_{\mathrm{k+1}}-\boldsymbol{a}_{\mathrm{k+1}}),$$

计算$\varphi(\lambda_{k+1})$,转步5

步5:$\mathrm{k}:=\mathrm{k}+1$,转步2。

大连理工大学 Dalian University Of Technology

■ Fibonacci法

Fibonacci法是与0.618法类似的一种方法。

它与0.618法的主要区别之一在于：搜索区间长度的缩短率不是采用黄金分割数，而是采用Fibonacci数：

$$F_{n+1}=F_n+F_{n-1}\ ,\quad F_0=F_1=1,\quad n=1,2,\ldots$$

Fibonacci法中计算公式为：

$$\lambda_k=a_k+\left(1-\frac{F_{n-k}}{F_{n-k+1}}\right)(b_k-a_k)$$

$$=a_k+\frac{F_{n-k-1}}{F_{n-k+1}}(b_k-a_k),k=1,2,..,n-1$$

$$\mu_k=a_k+\frac{F_{n-k}}{F_{n-k+1}}(b_k-a_k),\ k=1,2,...,n-1$$

■ Fibonacci法

显然, 这里F_{n-k}/F_{n-k+1}相当于0.618法(1.5)-(1.6)中的τ,
每次的缩短率满足

$$b_{k+1}-a_{k+1}=\frac{F_{n-k}}{F_{n-k+1}}(b_k-a_k)$$

这里n时计算函数值的次数,即要求经过n次计算函数值后,最后区间的长度不超过δ,即$b_n-a_n\le\delta.$

由于

$$b_n-a_n=\tfrac{F_1}{F_2}(\,b_{n-1}-a_{n-1})$$

$$=\frac{F_1}{F_2}\cdot\frac{F_2}{F_3}\cdot\ldots\cdot\frac{F_{n-1}}{F_n}(b_1-a_1)=\tfrac{1}{F_n}(b_1-a_1)$$

大连理工大学 Dalian University Of Technology

■ Fibonacci法

故 $$\frac{1}{F_n}(b_1-a_1)\le\delta\Rightarrow F_n\ge\frac{b_1-a_1}{\delta}\qquad(1.20)$$

给出最终区间长度得的上界δ,由(1.20)求出$Fibonacci$数F_n,再跟据F_n确定出n, 从而搜索一直进行到第n 个搜索点为止．注意到

$$F_n=\frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^{k+1}-\left(\frac{1-\sqrt{5}}{2}\right)^{k+1}\right\}$$

从而 $$\lim_{k\to\infty}\frac{F_{k-1}}{F_k}=\frac{\sqrt{5}-1}{2}=\tau.$$

■ Fibonacci法

上式表明当n 趋于无穷时, Fibonacci法与0.618法的区间缩短率相同, 因而也是以收敛比r线性收敛。

可以证明Fibonacci法是分割方法求一维极小化问题的最优策略, 而0.618法是近似最优的。

■ Fibonacci法

$Step1$,给定初始区间$[a_1,b_1]$和最终区间长度L.求计算函数值的次数n,使得$F_n \geq \dfrac{b_1-a_1}{L}$

置辨别常数$\delta>0$.计算试探点λ_1和μ_1:

$$\lambda_1 = a_1 + \frac{F_{n-2}}{F_n}(b_1-a_1),\mu_1 = a_1 + \frac{F_{n-1}}{F_n}(b_1-a_1)$$

■ Fibonacci法

计算函数值$\varphi(\lambda_1)$,$\varphi(\mu_1)$.置k=1

$Step2$,若$\varphi(\lambda_k)>\varphi(\mu_k)$,转3； 若$\varphi(\lambda_k)\le\varphi(\mu_k)$,转4。

$Step3$,令$a_{k+1}=\lambda_k, b_{k+1}=b_k, \lambda_{k+1}=\mu_k$.计算试探点$\mu_{k+1}$

$$\mu_{k+1}=a_{k+1}+\frac{F_{n-k-1}}{F_{n-k}}(b_{k+1}-a_{k+1}).$$

若$k=n-2$,转6,否则， 计算$\varphi(\mu_{k+1})$,转5。

$Step4$,令$a_{k+1}=\lambda_k, b_{k+1}=\mu_k, \mu_{k+1}=\lambda_k$.计算试探点$\lambda_{k+1}$

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

若$k=n-2$,转6,否则,计算$\varphi(\lambda_{k+1})$,转5.

■ Fibonacci法

Step 5, 置k:=k+1,转2.

*Step*6, 令$\lambda_n=\lambda_{\text{n-1}},\mu_{\text{n}}=\lambda_{n-1}+\delta$.计算$\varphi(\lambda_n),\varphi(\mu_n)$.

若$\varphi(\lambda_n)>\varphi(\mu_n)$,则令$a_n=\lambda_{\text{n}},b_{\text{n}}=b_{n-1}$;

若$\varphi(\lambda_n)\le\varphi(\mu_n)$,则令$a_n=a_{\text{n-1}},b_{\text{n}}=\lambda_n$.

停止计算,极小点含于$[a_n,b_n]$.

■ 进退法

确定搜索区间的一种简单方法叫进退法，其基本思想是从一点出发，按一定步长，试图确定出函数值呈现“高 - 低 - 高”的三点．一个方向不成功，就退回来，再沿相反方向寻找．具体地说，就是给出初始点 α_0, 初始步长 $h_0 > 0$, 若

$$\varphi(\alpha_0 + h_0) < \varphi(\alpha_0),$$

则下一步从新点 $\alpha_0 + h_0$ 出发，加大步长，再向前搜索．若

$$\varphi(\alpha_0 + h_0) > \varphi(\alpha_0),$$

则下一步仍以 α_0 为出发点，沿反方向同样搜索，直到目标函数上升就停止．这样便得到一个搜索区间，这种方法叫进退法．

■ 进退法

步 1 选取初始数据. $\alpha_0 \in [0, \infty)$, $h_0 > 0$, 加倍系数 $t > 1$ (一般取 $t = 2$), 计算 $\varphi(\alpha_0)$, $k := 0$.

步 2 比较目标函数值. 令 $\alpha_{k+1} = \alpha_k + h_k$, 计算 $\varphi_{k+1} = \varphi(\alpha_{k+1})$, 若 $\varphi_{k+1} < \varphi_k$, 转步 3, 否则转步 4.

步 3 加大探索步长. 令 $h_{k+1} := th_k$, $\alpha := \alpha_k$, $\alpha_k := \alpha_{k+1}$, $\varphi_k := \varphi_{k+1}$, $k := k + 1$, 转步 2.

步 4 反向探索. 若 $k = 0$, 转换探索方向, 令 $h_k := -h_k$, $\alpha_k := \alpha_{k+1}$, 转步 2; 否则, 停止迭代, 令

$$a = \min\{\alpha, \alpha_{k+1}\}, \quad b = \max\{\alpha, \alpha_{k+1}\},$$

输出[a,b]

■ 牛顿法

考虑问题 $\min f(x), x \in R$ (3.1)

令

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

又令

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

得到$\varphi(x)$的驻点，记做$x^{(k+1)}$，则

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})} \qquad (3.2)$$

■ 牛顿法

在点$x^{(k)}$附近，$f(x)\approx\varphi(x)$，因此可用$\varphi(x)$的极小点作为目标函数$f(x)$的极小点的估计。

如果$x^{(k)}$是$f(x)$的极小点,则利用(3.2)可以得到极小点的一个进一步的估计.于是得到一个序列$\{x^{(k)}\}$.

*Th*3.1设$f(x)$存在连续三阶导数,$\bar{x}$满足

$$f'(\bar{x})=0, f''(\bar{x})\neq 0$$

初始点$x^{(1)}$充分接近$\bar{x}$,则牛顿法产生的序列$\{x^{(k)}\}$至少以2阶收敛速率收敛于$\bar{x}$.

■ 牛顿法

证明:牛顿法可定义为算法映射

$$A(x) = x - \frac{f'(x)}{f''(x)} \qquad (3.3)$$

设解集合$\Omega=\{\bar{x}\}$,定义函数$\alpha(x) = |x - \bar{x}|$

下证α是关于解集合Ω和算法A的下降函数

■ 牛顿法

设$x^{(k)} \neq \bar{x}$，$x^{(k)} \in A(x^{(k)})$.由$f'(\bar{x})=0$,有

$$\alpha(x^{(k+1)}) = \left| x^{(k)} - \bar{x} \right| = \left| x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})} - \bar{x} \right|$$

$$= \frac{1}{\left|f''(x^{(k)})\right|} \left| x^{(k)} f''(x^{(k)}) - f'(x^{(k)}) - \bar{x} f''(x^{(k)}) \right|$$

$$= \frac{1}{\left|f''(x^{(k)})\right|} \left| f'(\bar{x}) - [f'(x^{(k)}) + (\bar{x} - x^{(k)}) f''(x^{(k)})] \right|$$

$$= \frac{1}{\left|f''(x^{(k)})\right|} \frac{1}{2} (\bar{x} - x^{(k)})^2 \left| f'''(\xi) \right| \qquad (3.4)$$

其中ξ在$\bar{x}$与$x^{(k)}$之间。

■ 牛顿法

由于$f'(x)$和$f''(x)$ 连续，$f''(\bar{x})\neq 0$， 故当$x^{(k)}$接近$\bar{x}$时，必存在$k_1,k_2>0$使得在包含$x^{(k)}$和$\bar{x}$的闭区间上的每一点x处有

$$|f''(x)|\geq k_1,|f'''(x)|\leq k_2 \qquad (3.5)$$

代入(3.4)，则 $\left|x^{(k+1)}-\bar{x}\right|\leq \dfrac{k_2}{2k_1}(\bar{x}-x^{(k)})^2.$ $\qquad$ (3.6)

取初始点$x^{(1)}$充分接近$\bar{x}$，使得

$$\frac{k_2}{2k_1}\left|\bar{x}-x^{(1)}\right|<1$$

由此推得 $\{x^{(k)}\}\subset X=\{x\,\|x-\bar{x}|\leq|x^{(1)}-\bar{x}|\}$

且 $\left|x^{(k+1)}-\bar{x}\right|<\left|x^{(k)}-\bar{x}\right| \qquad (3.7)$

■ 牛顿法

由此知,α是关于解集合Ω和算法A的下降函数, 且X为紧集,

$A(x)$在X上连续. 根据$Th8.2.1$,$\{x^{(k)}\}$收敛于$\bar{x}$. 由(3.6)知收敛阶为2.

算法(牛顿法)

Step1, 给定初始点$x^{(0)}$, 允许误差$\varepsilon>0$, 置$k=0$;

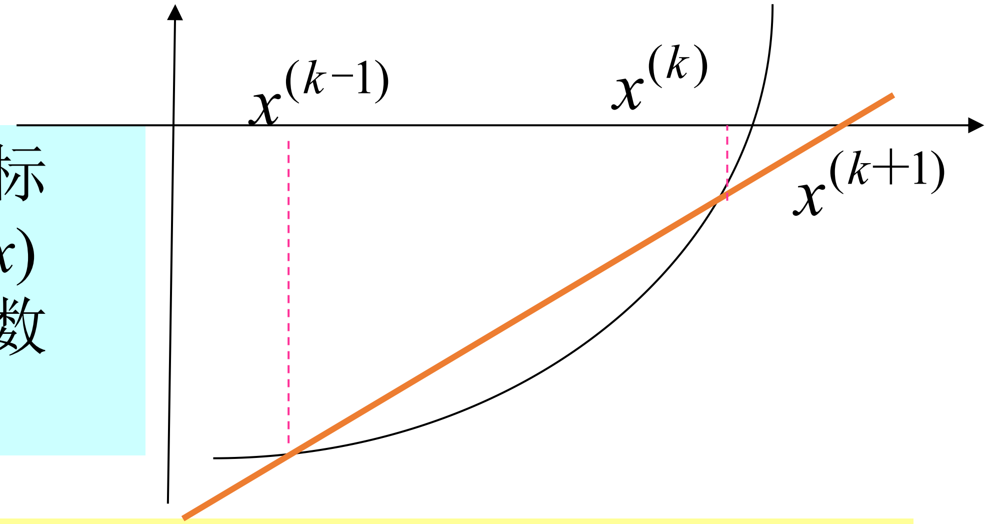
Step2, 若$\left|f'(x^{(k)})\right|<\varepsilon$, 停止，得$x^{(k)}$。

Step3, 计算点$x^{(k+1)}$:

$$x^{(k+1)}=x^{(k)}-\frac{f'(x^{(k)})}{f''(x^{(k)})}\quad 置k=k+1.转2$$

■ 割线法

基本思想：用割线逼近目标函数的导函数的曲线$y=f'(x)$把割线的零点作为目标函数的驻点的估计。



设在点$x^{(k)}$和$x^{(k-1)}$处的导数分别为$f'(x^{(k)})$和$f'(x^{(k-1)})$。令

$$\varphi(x)=f'(x^{(k)})+\frac{f'(x^{(k)})-f'(x^{(k-1)})}{x^{(k)}-x^{(k-1)}}(x-x^{(k)})=0$$

$$\Rightarrow x^{(k+1)}=x^{(k)}-\frac{x^{(k)}-x^{(k-1)}}{f'(x^{(k)})-f'(x^{(k-1)})}f'(x^{(k)}) \qquad (3.8)$$

用公式(3.8)进行迭代，得到序列$\{x^{(k)}\}$.

■ 割线法

在一定的条件下，这个序列收敛于解：

$Th3.2$ 设$f(x)$存在连续三阶导数,$\bar{x}$满足

$$f'(\bar{x})=0,\ f''(\bar{x})\neq 0$$

若$x^{(1)}$和$x^{(2)}$充分接近$\bar{x}$,则割线法产生的序列$\{x^{(k)}\}$收敛于$\bar{x}$.

收敛阶为1.618.

证明：设$\Delta=\{x\,\big|\,|x-\bar{x}|\leq\delta\}$是包含$\bar{x}$的某个充分小的闭区间，

使得对每一个$x\in\Delta$,有$f''(x)\neq 0$,取$x^{(1)},x^{(2)}\in\Delta$.

以$x^{(k)},x^{(k-1)}\in\Delta$为节点构造插值多项式

$$\varphi(x)=f'(x^{(k)})+\frac{f'(x^{(k)})-f'(x^{(k-1)})}{x^{(k)}-x^{(k-1)}}(x-x^{(k)}) \qquad (3.9)$$

■ 割线法

$\Rightarrow$ 插值余项

$$f'(x)-\varphi(x)=\frac{f'''(\xi_1)}{2}(x-x^{(k)})(x-x^{(k-1)}) \qquad (3.10)$$

其中$\xi_1 \in \Delta$。

由于$f'(\bar{x})=0$,因此由(3.10)得到

$$\varphi(\bar{x})=-\frac{f'''(\xi_1)}{2}e_k e_{k-1} \qquad (3.11)$$

$$e_k=x^{(k)}-\bar{x},\ e_{k-1}=x^{(k-1)}-\bar{x}$$

另一方面,由(3.8)知$\varphi(x^{(k)})=0$,由(3.9)知

大连理工大学 Dalian University Of Technology

■ 割线法

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

$$\Rightarrow \varphi(\bar{x})=\varphi(\bar{x})-\varphi(x^{(k+1)})=\varphi'(\xi_2)(\bar{x}-x^{(k+1)})$$

$$=\frac{f'(x^{(k)})-f'(x^{(k-1)})}{x^{(k)}-x^{(k-1)}}(\bar{x}-x^{(k+1)})=f''(\xi_3)(\bar{x}-x^{(k+1)})$$

$$=-f''(\xi_3)e_{k+1}, \qquad (3.12)$$

其中$e_{k+1}=x^{(k+1)}-\bar{x}$，$\xi_3$在$x^{(k)}$和$x^{(k-1)}$之间，$f''(\xi_3)\neq 0.$

由(3.11)和(3.12)得到

$$e_{k+1}=\frac{f'''(\xi_1)}{2f''(\xi_3)}e_k e_{k-1} \qquad (3.13)$$

■ 割线法

上式两端取绝对值，则

$$\left|e_{k+1}\right| = \frac{\left|f'''(\xi_1)\right|}{2\left|f''(\xi_3)\right|}\left|e_k\right|\left|e_{k-1}\right| \qquad (3.14)$$

令 $$\mathrm{M} = \frac{\max\limits_{x\in\Delta}\left|f'''(x)\right|}{2\min\limits_{x\in\Delta}\left|f''(x)\right|}$$

则 $$\left|e_{k+1}\right| \le \mathrm{M}\left|e_k\right|\left|e_{k-1}\right| \qquad (3.15)$$

取充分小的Δ.使得 $$M\delta < 1 \qquad (3.16)$$

则 $$\left|e_{k+1}\right| \le \mathrm{M}\delta\delta < \delta$$

于是, $x^{(k)}, x^{(k-1)} \in \Delta \Rightarrow x^{(k+1)} \in \Delta$, 进而由$x^{(1)}, x^{(2)} \in \Delta$

$\Rightarrow x^{(k)} \in \Delta, \forall \mathrm{k}$, 由(3.15) $\Rightarrow \{x^{(k)}\}$收敛于$\bar{x}$.

■ 割线法

下面考虑收敛速率，考虑k取充分大的情形. 根据(3.14)

$$|e_{k+1}| \approx \bar{M}\,|e_k|\,|e_{k-1}| \qquad (3.17)$$

其中 $\bar{\mathrm{M}}=\dfrac{|f'''(\bar{x})|}{2|f''(\bar{x})|}$

令 $|e_k| = a^{y_k}/\bar{\mathrm{M}} \qquad (3.18)$

代入 (3.17).于是可考虑差分方程

$$y_{k+1} = y_k + y_{k-1}$$

它的特征方程是$\tau^2 - \tau - 1 = 0$

其两个根是$\tau_1 = \dfrac{1+\sqrt{5}}{2} \approx 1.618, \tau_2 = \dfrac{1-\sqrt{5}}{2} \approx -0.618$

■ 割线法

于是 $$|e_k| = \frac{1}{\bar{M}} a^{c_1\tau_1^{\,k} + c_2\tau_2^{\,k}} \approx \frac{1}{\bar{M}} a^{c_1\tau_1^{\,k}} \ (k\text{远大于}1) \qquad (3.19)$$

$$|e_{k+1}| \approx \frac{1}{\bar{M}} a^{c_1\tau_1^{\,k+1}} (k\text{远大于}1) \qquad (3.20)$$

$$\Rightarrow \frac{|e_{k+1}|}{|e_k|^{\tau_1}} \approx \bar{M}^{\tau_1 - 1} (k\text{远大于}1)$$

注意： 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，如果初始点选择得不好，可能不收敛。

■ 抛物线法

- 基本思想：在极小点附近用二次三项式$\varphi(x)$逼近目标函数$f(x)$, 令$\varphi(x)$与$f(x)$在三点$x^{(1)}<x^{(2)}<x^{(3)}$处有相同的函数值，并假设

$$f(x^{(1)})>f(x^{(2)}),\ f(x^{(2)})<f(x^{(3)})$$

- 令 $\varphi(x)=a+bx+cx^2$ (9.3.21)
- 又令 $\varphi(x^{(1)})=a+bx^{(1)}+c(x^{(1)})^2=f(x^{(1)})$ (9.3.22)

$\varphi(x^{(2)})=a+bx^{(2)}+c(x^{(2)})^2=f(x^{(2)})$ (9.3.23)

$\varphi(x^{(3)})=a+bx^{(3)}+c(x^{(3)})^2=f(\ ^{(3)})$ (9.3.24)

解方程组(9.3.22-24),求二次逼近函数$\varphi(x)$的系数a，b，c

为书写方便，记

抛物线法

$$B_1=((x^{(2)})^2-(x^{(3)})^2)f(x^{(1)}),B_2=((x^{(3)})^2-(x^{(1)})^2)f(x^{(2)})$$

$$B_3=((x^{(1)})^2-(x^{(2)})^2)f(x^{(3)}),C_1=(x^{(2)}-x^{(3)})f(x^{(1)}),$$

$$C_2=(x^{(3)}-x^{(2)})f(x^{(2)}),C_3=(x^{(1)}-x^{(2)})f(x^{(3)});$$

$$D=(x^{(1)}-x^{(2)})(x^{(2)}-x^{(3)})(x^{(3)}-x^{(1)})$$

则 $\Rightarrow b=\dfrac{B_1+B_2+B_3}{D},$ (3.25)

$$c=\frac{C_1+C_2+C_3}{D} \qquad (3.26)$$

为求φ(x)的极小点，令 $\varphi'(\mathrm{x})=b+2cx=0$

解得 $x=-\dfrac{b}{2c}$ (3.27)

大连理工大学 Dalian University Of Technology

■ 抛物线法

把φ(x)的驻点记做$\bar{x}^{(k)}$.则

$$\bar{x}^{(k)}=\frac{B_1+B_2+B_3}{2(C_1+C_2+C_3)} \qquad (3.28)$$

这样把$\bar{x}^{(k)}$作为$f(x)$的极小点的一个估计.

再从$x^{(1)},x^{(2)},x^{(3)},\bar{x}^{(k)}$中选择目标函数值最小的点及其左右两点,

给予相应的上标,代入公式,求出极小点的新的估计值$\bar{x}^{(k+1)}$.

以此类推,产生点列$\{x^{(k)}\}$,

在一定条件下, 这个点列收敛于问题的解, 其收敛级为1.3.

■ 三次插值法

基本思想：首先选取两个点$x^{(1)}, x^{(2)}(x^{(1)} < x^{(2)})$，使得$f'(x^{(1)}) < 0, f'(x^{(2)}) > 0$.于是区间$(x^{(1)}, x^{(2)})$内存在极小点利用在这两点的函数值和导数构造一个三次多项式$\varphi(x)$，使它与$f(x)$在$x^{(1)}$和$x^{(2)}$有相同的函数值和相同的导数，用$\varphi(x)$逼近$f(x)$.进而用$\varphi(x)$的极小点估计$f(x)$的极小点.

■ 三次插值法

令

$$\varphi(x)=a(x-x^{(1)})^3+b(x-x^{(1)})^2+c(x-x^{(1)})+d \qquad (3.29)$$

$$\varphi(x^{(1)})=f(x^{(1)}), \qquad (3.30); \qquad \varphi'(x^{(1)})=f'(x^{(1)}), \qquad (331)$$

$$\varphi(x^{(2)})=f(x^{(2)}), \qquad (3.32); \qquad \varphi(x^{(2)})=f(x^{(2)}), \qquad (3.33)$$

将(3.30)-(33)依次代入(3.29)得

$$\begin{cases} d=f(x^{(1)}) \\ c=f'(x^{(1)}) \\ a(x^{(2)}-x^{(1)})^3+b(x^{(2)}-x^{(1)})^2+c(x^{(2)}-x^{(1)})+d=f(x^{(2)}) \\ 3a(x^{(2)}-x^{(1)})^2+2b(x^{(2)}-x^{(1)})+c=f'(x^{(2)}) \end{cases} \qquad (3.34)$$

■ 三次插值法

我们目的是求$\varphi(x)$的极小点,期望用它来逼近极小点,或者基于此再确定新的迭代.为此,求出满足极值条件的点,即满足$\varphi'(x)=0$, $\varphi''(x)>0$ 的点.

$$\varphi'(x)=3a(x-x^{(1)})^2+2b(x-x^{(1)})+c \qquad (3.35)$$

$$\varphi''(x)=6a(x-x^{(1)})+2b, \qquad (3.36)$$

令 $\varphi'(x)=0\Rightarrow$

$$a(x-x^{(1)})^2+2b(x-x^{(1)})+c=0, \qquad (3.37)$$

解此方程

$$(1)a=0\Rightarrow \bar{x}-x^{(1)})=-\frac{c}{2b}, \qquad (3.38)$$

$$(2)a\neq 0 \ \Rightarrow \ \bar{x}-x^{(1)})=-\frac{-b\pm\sqrt{b^2-3ac}}{3a} \qquad (3.33)$$

■ 三次插值法

第一种情形,有$\varphi''(x)=2b>0$(由假设及(3.24)可得)

故$\bar{x}$是$\varphi(x)$的极小点.

第二种情形,将方程的根代入(3.36)

$$\varphi''(\bar{x})=6a(\bar{x}-x^{(1)})+2b=6a\frac{-b\pm\sqrt{b^2-3ac}}{3a}+2b$$

$$=\pm2\sqrt{b^2-3ac}$$

要使$\varphi''(x)>0$

$$\Rightarrow \text{应取}\ \bar{x}=\frac{-b+\sqrt{b^2-3ac}}{3a}=\frac{-\mathrm{c}}{b+\sqrt{b^2-3ac}} \qquad (3.40)$$

■ 三次插值法

注意到当a=0时,b>0,故当a=0时由(3.40)得

$$\bar{x} - x^{(1)} = -\frac{c}{2b}$$

这个结果恰好是(3.38).这表明(3.40)是在a=0和a≠0两种情形下极小点的统一表达式.

这样可以解方程组来求出系数a,b,c再代入(3.40),从而可得φ(x)的极小点x*.

下面给出用$f(x^{(1)}), f'(x^{(1)}), f(x^{(2)})$和$f'(x^{(2)})$表示的$\varphi(x)$的极小点$\bar{x}$的表达式.

■ 三次插值法

记

$$s = \frac{3[f(x^{(2)}) - f(x^{(1)})]}{x^{(2)} - x^{(1)}} \qquad (3.41)$$

$$z =, s - f'(x^{(1)}) - f'(x^{(2)}) \qquad (3.42)$$

$$w^2 = z^2 - f'(x^{(1)})\, f'(x^{(2)}), \qquad (3.43)$$

$$\Rightarrow \bar{x} = x^{(1)} + (x^{(2)} - x^{(1)})(1 - \frac{f'(x^{(2)}) + w + z}{f'(x^{(2)}) - f'(x^{(1)}) + 2w}) \qquad (3.44$$

■ 三次插值法

式(3.44)中必有$f'(x^{(2)})-f'(x^{(1)})+2w\neq 0$

(事实上, $f'(x^{(1)})<0, f'(x^{(2)})>0, w$由(3.43)确定,取算术根, 故 $w>0$)

于是利用公式 (3.41)-(43)求出w,z,再由 (3.44) 求得极小点$\bar{x}$.

若$|f'(x)|$充分小, $\bar{x}$就可作为$f(x)$的可接受的极小点, 否则,

可从$x^{(1)}, x^{(2)}$和$\bar{x}$中确定两个插值点, 再利用上述公式进行计算.

■ 三次插值法

算法(两点三次插值法)

$Step1:$ 给定初点 $x^{(1)}, x^{(2)}$, 计算 $f(x^{(1)}), f'(x^{(1)}), f(x^{(2)}), f'(x^{(2)})$

要求满足 $x^{(2)} > x^{(1)}, f'(x^{(1)}) < 0, f'(x^{(2)}) > 0$, 给定允许误差 δ.

$Step2:$ 按公式(3.41)-(3.44)式计算 $s, z, w, \bar{x}$

$Step3:$ 若 $\left|x^{(2)} - x^{(1)}\right| \le \delta$ =,停止,得点 $\bar{x}$ 否则转 4

$Step4:$ 计算 $f(\bar{x}), f'(\bar{x})$.若 $f'(\bar{x}) = 0$ 停止,得点 $\bar{x}$.

若 $f'(\bar{x}) < 0$, 令 $x^{(1)} = \bar{x}, f(\bar{x}) = f(x^{(1)}), f'(\bar{x}) = f'(x^{(1)})$ 转 2

若 $f'(\bar{x}) > 0$, 令 $x^{(2)} = \bar{x}, f(\bar{x}) = f(x^{(2)}), f'(\bar{x}) = f'(x^{(2)})$ 转 2.

- 第六次作业

第280页第九章1.2.3.4

大连理工大学 Dalian University Of Technology

小结

- 一维搜索的基本概念
- 试探法
- 函数逼近法