第三章 一维搜索方法

- 3-1 概述
- 3-2 搜索区间的确定与区间消去法原理
- 3-3 黄金分割法
- 3-4 二次插值法
- 3-5 分数法

3-1 概述

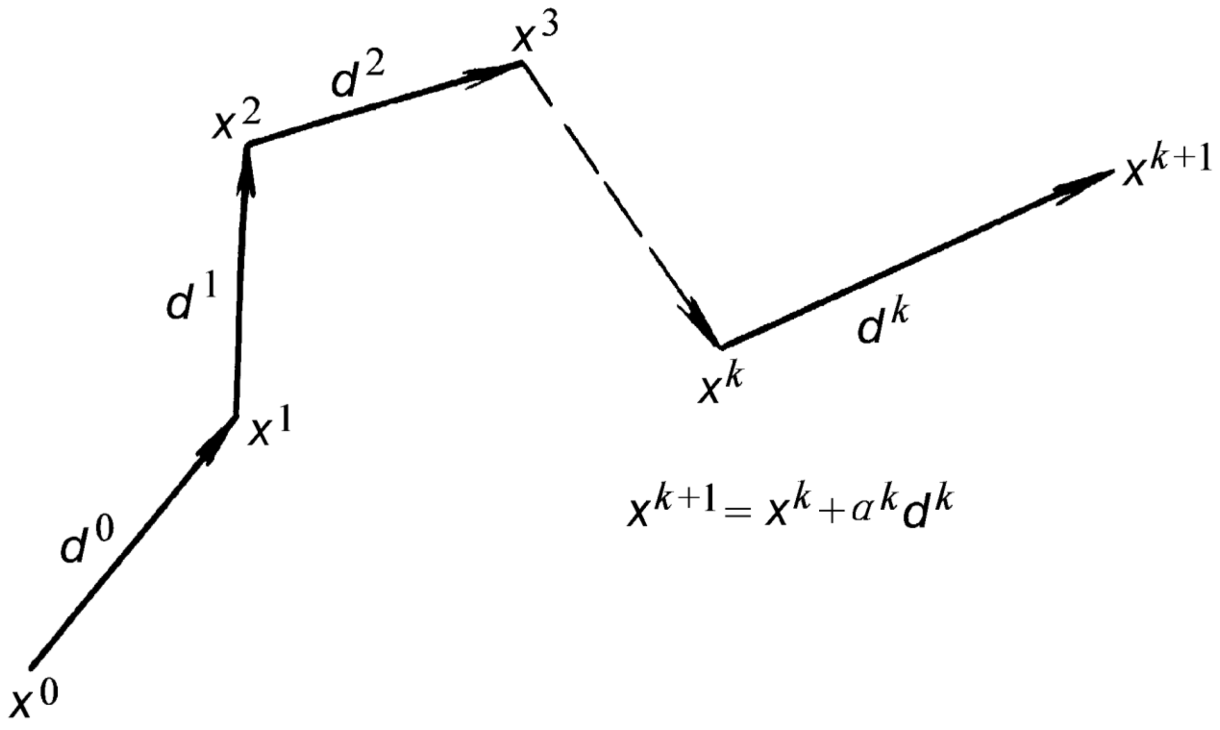
1 一维搜索方法

定义：求解一维目标函数f（α）的极小点和极小值的方法称为一维搜索方法

一维搜索方法是优化搜索方法的基础

当采用数学规划法寻求多元函数的极值点时，一般要进行一系列如下格式的迭代计算：

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^{k} + \alpha_k \boldsymbol{d}^{k} \quad (k = 0,1,2,\cdots)$$

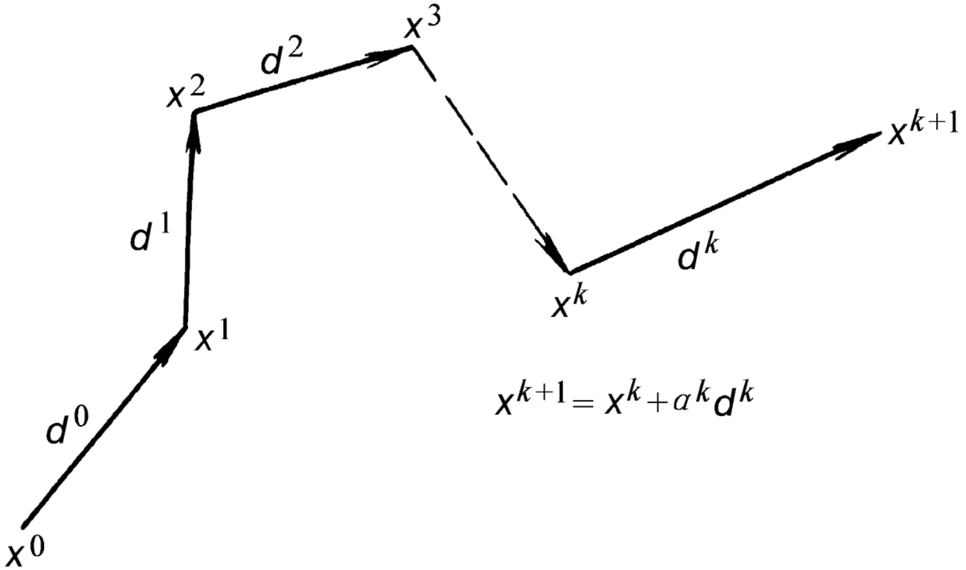


当方向d^k给定，求最佳步长 α 就是求一元函数 ：

$$f(\boldsymbol{x}^{k+1})=f(\boldsymbol{x}^k+\alpha_k\boldsymbol{d}^k)=\varphi(\alpha_k)$$

的极值问题，这一过程被称为一维搜索.

在求多元函数的极值点时，需要进行一系列的一维搜索。可见一维搜索时优化搜索方法的基础



3　求解一元函数 $\varphi(\alpha)$ 的极小点的 α^* 方法

（a）解析解法：需要计算在迭代点的梯度和海赛矩阵

$$\alpha^* = -\frac{d^T \nabla f(x)}{d^T G d}$$

其中$\nabla f(x)$为$x = x^k$点梯度

G为函数在该点的海赛矩阵

解析法的缺点是需要进行求导计算，对于函数关系复杂，求导困难或无法求导的情况，使用解析法非常不便

（b）数值解法

数值法的基本思路是：先确定 α^* 所在的搜索区间，然后根据区间消去法原理不断缩小此区间，从而获得 α^* 的数值近似解

在优化设计中，求解最佳步长因子 α^* 主要使采用数值解法

3-2搜索区间的确定与区间消去法原理

1. 一维搜索的基本思想

欲求一元函数 f(α)的极小点 α^*，必须先确定 α^* 所在的区间。找初始单谷区间是一维搜索的第一步. 第二步使区间缩小。

如何确定包含极小点的一个区间?

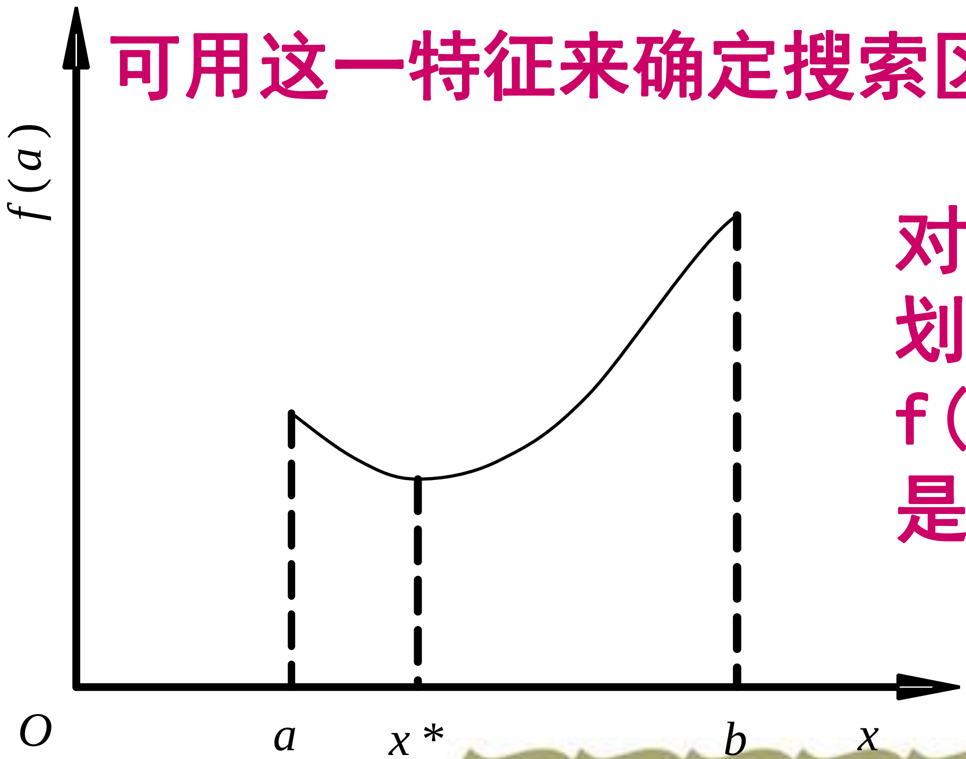
思想:从一点出发,按一定的步长, 试图确定出函数值呈现"高- 低-高"的三点。一个方向不成功，就退回来，再沿相反方向寻找。

(1). 单谷(峰）区间

在给定区间内仅有一个谷(峰)值的函数称为单谷（峰）函数，其区间称为单谷（峰）区间。

单谷函数的性质：在极小点X *左边函数值严格下降，在极小点X *右边函数严格上升

单谷函数具有函数值：“大—小—大”，图形具有高—低—高的特征，可用这一特征来确定搜索区间



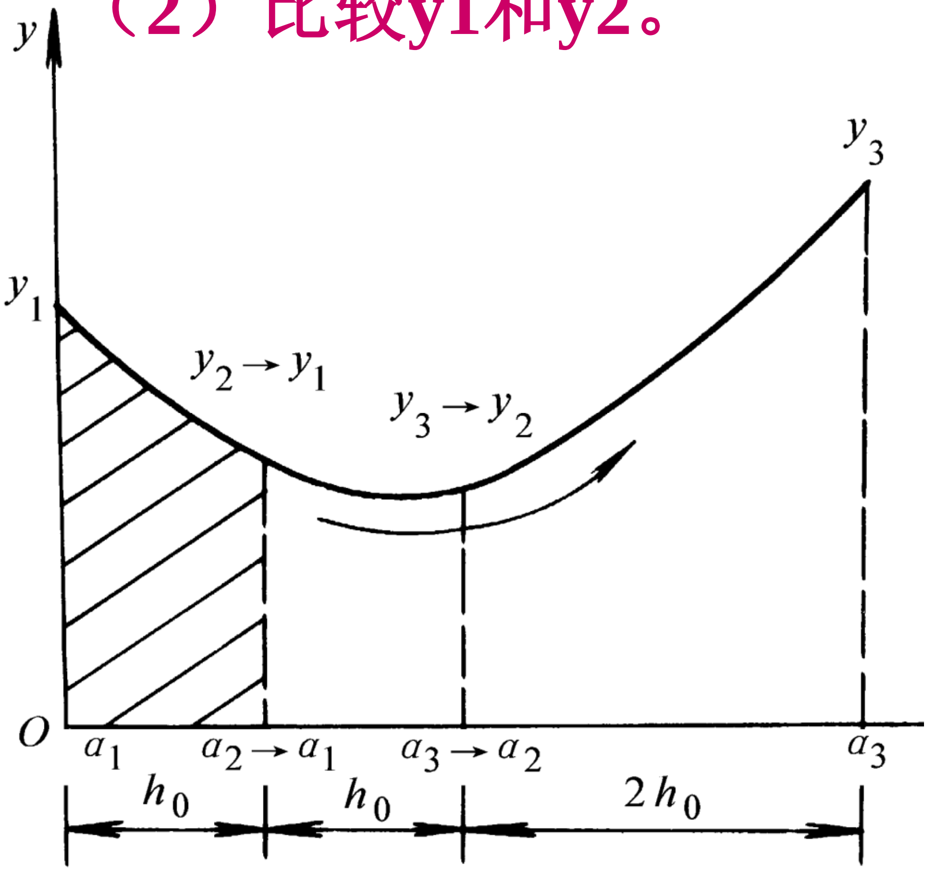
对于多谷函数，只要适当划分[a,b]，就可以使函数f（α）在每一个区间上都是单谷的

2. 确定初始单谷区间的进退法

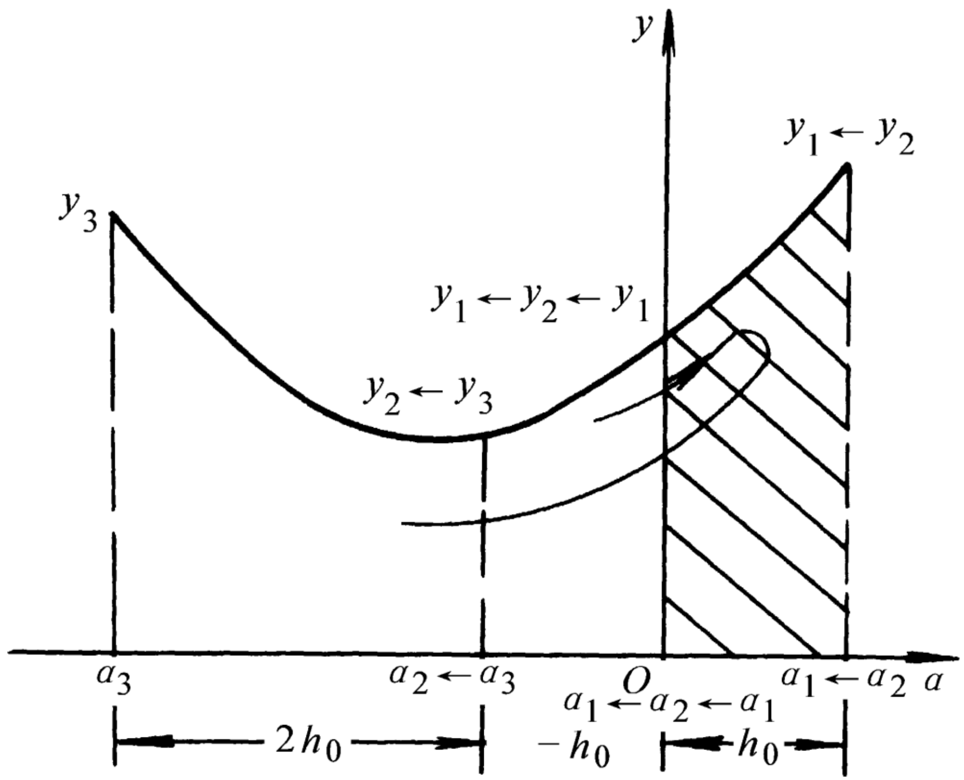
（1）给定初始点$\alpha 0$和初始步长h0

令$\alpha 1=\alpha 0$，h= h0，$\alpha 2=\alpha 1+h$，得$y1=f(\alpha 1)$, $y2=f(\alpha 2)$

（2）比较y1和y2。



书中图3－2



书中图3－3

2. 确定初始单谷区间的进退法

(1)选定初始点$a1$, 初始步长 $h = h_0$,

计算 $y1 = f(a1)$， $y2 = f(a1 + h)$。

(2)比较y1和y2。

(a)如y1>y2, 向右前进；转(3)

(b)如y1<y2, 向左后退；h＝－ h0，a1和a2互换，

y1和y2互换，并转(3)向后探测

(c)如y1＝y2，极小点在$a1$和$a1+h$之间。

(3)产生新的探测点$a3 = a2 + h$，$y3 = f(a3)$;

(4) 比较函数值 y2与y3:

(a)如y2<y3, 则初始区间得到;

h>0时，$[a,b]=[a1,a3]$;　　h<0时，$[a,b]=[a3,a1]$;

(b)如y2>y3, 加大步长 h＝2 h ，a1=a2, a2=a3, 转(3)继续探测。

3 算法框图

框图

3-3 一维搜索的试探方法

最常用的一维试探方法是黄金分割方法，又称作0.618法

黄金分割法适用于[*a*, *b*]区间上的任何单谷函数求极小值问题。

黄金分割法也是建立在区间消去法原理基础上的试探方法。它一种压缩区间的方法

1. 黄金分割法的基本原理

在搜索区间内[*a*, *b*]适当插入两点 α_1, α_2 ，将区间分成三段；应用函数的单谷性质，通过函数值大小的比较，删去其中的一段，使搜索区间得以缩短。然后再在保留下来的区间上作同样的处置，如此迭代下去，使搜索区间无限缩小，从而得到极小点的数值近似解。

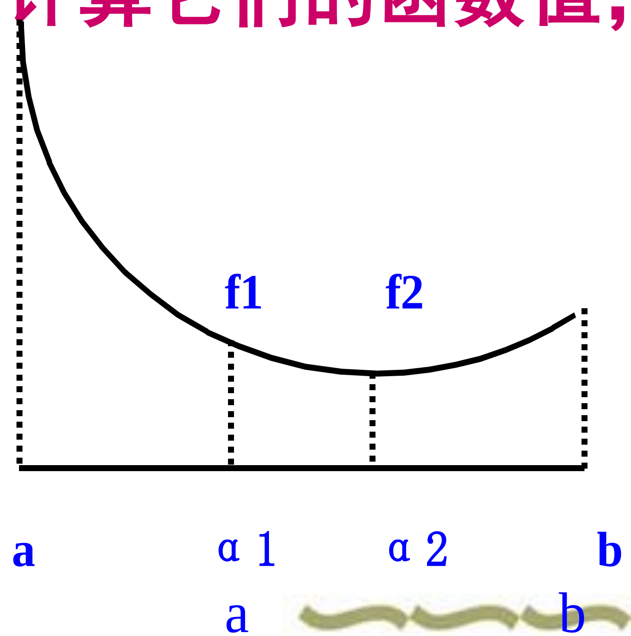
在单谷区间只有一个极小点，黄金分割法它是通过函数值的大小来确定取舍区间

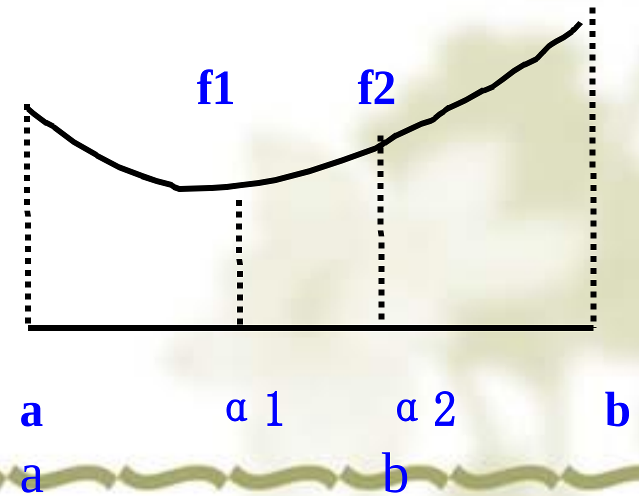
搜索过程如下：

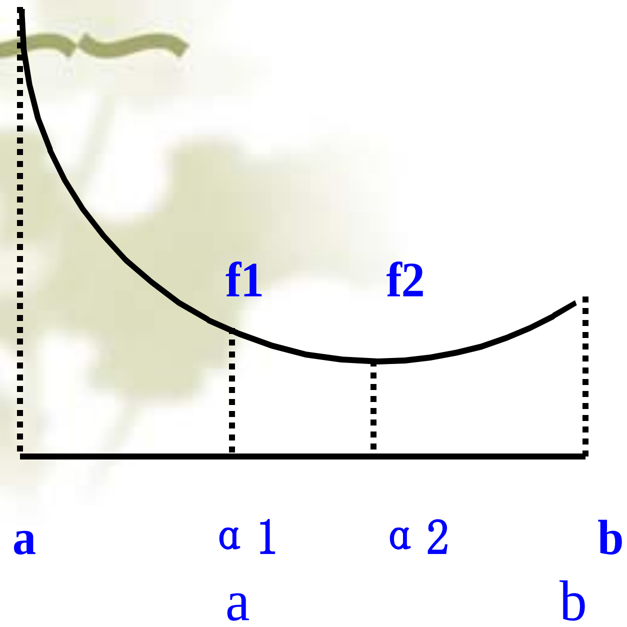
在搜索区间［ a b ］对称地取两点 α_1 α_2：

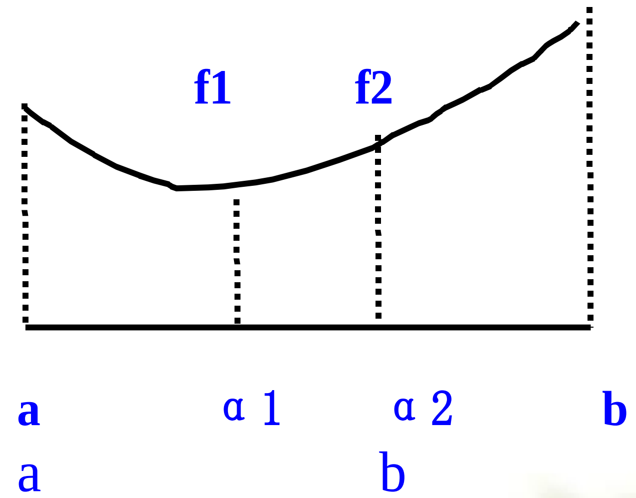
$\alpha_1=a+0.382(b-a)$ $\alpha_2=a+0.618(b-a)$

并计算它们的函数值，$f_1=f(\alpha_1)$ $f_2=f(\alpha_2)$









（1）若f_1> f_2, 极小点必须在区间[α_1 b]内，令a <= α_1, 此时区间收缩一次

（2）若f_2> f_1, 极小点必须在区间[a α_2]内，令b <= α_2, 此时区间收缩一次

将区间依次收缩，直至当b—a≤ ε（设定的某一精度），取

$\alpha^* = \dfrac{\alpha 1+\alpha 2}{2}$ 为极小点

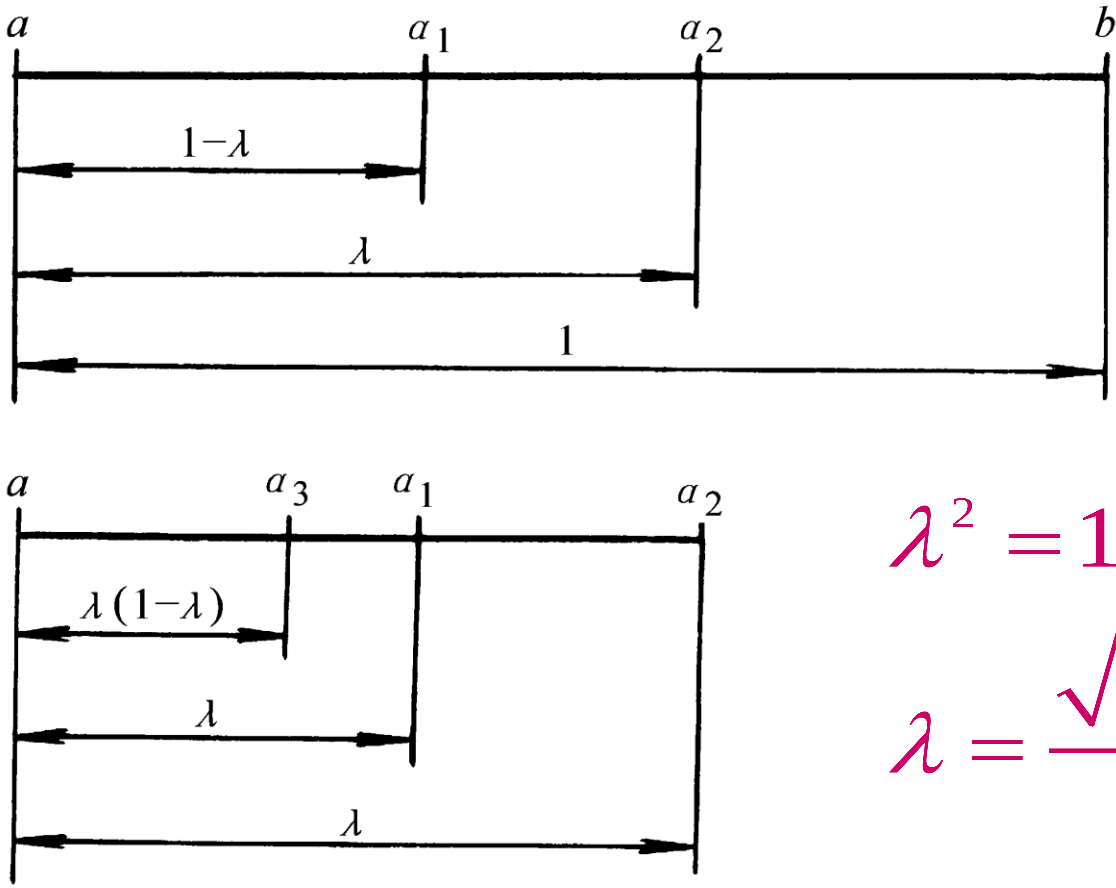
2. 0.618的由来

通过不断收缩搜索区间的长度来寻求一维函数的极小点，它是将搜索区间按比例λ缩小，因λ=0.618，故又称为0.618法

黄金分割法要求插入点的位置相对于区间[a,b]两端点具有对称性

黄金分割法还要下求在保留来的区间内再插入一点所形成的区间新三段，与原来区间的三段具有相同的比例分布

α_1, α_2 将区间分成三段



书中图3—6

$$\lambda^2 = 1 - \lambda$$

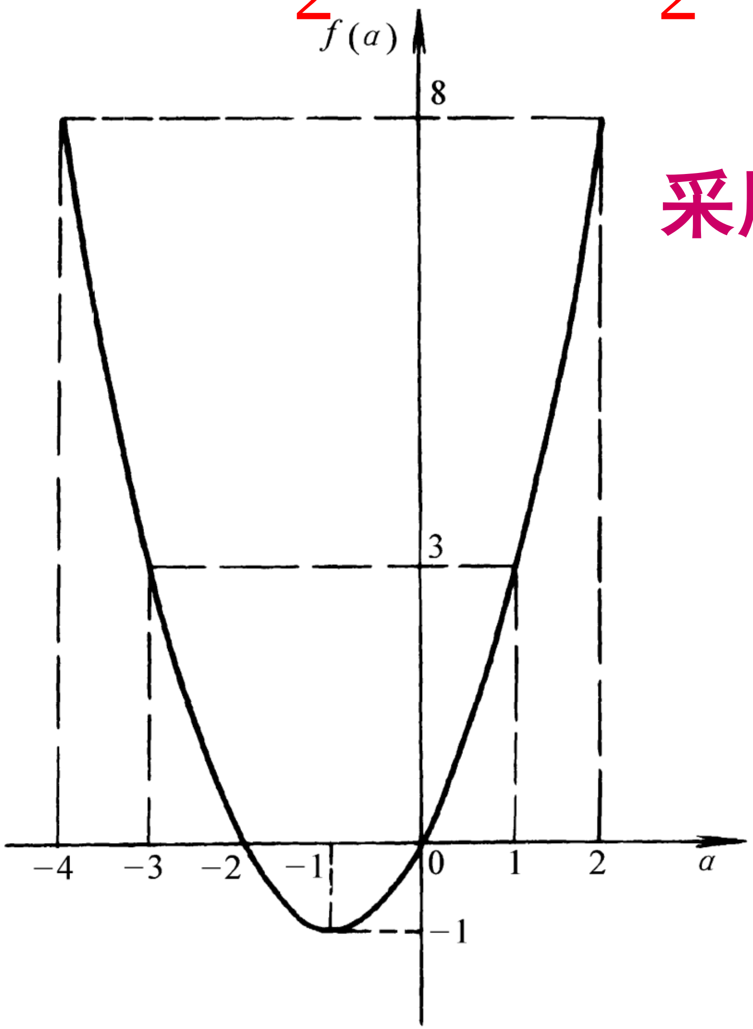
$$\lambda = \frac{\sqrt{5} - 1}{2} \approx 0.618$$

例3.1　对函数 $f(\alpha)=\alpha^2+2\alpha$，当给定搜索区间 $-3\le\alpha\le 5$ 时，试用黄金分割法求极小点。

迭代序号	a	α_1	α_2	b	f_1	比较	f_2
0	−3	0.056	1.944	5	0.115	<	7.667
1	−3	−1.111	0.056	1.944	−0.987	<	0.115
2	−3	−1.832	−1.111	0.056	−0.306	>	−0.987
3	−1.832	−1.111	−0.665	0.056	−0.987	<	−0.888
4	−1.832	−1.386	−1.111	−0.665	−0.851	>	−0.987
5	−1.386	−1.111	−0.940	−0.665			

经过5次迭代后

$$\alpha^* = \frac{1}{2}(a+b) = \frac{1}{2}\times(-1.386-0.665) = -1.0255$$

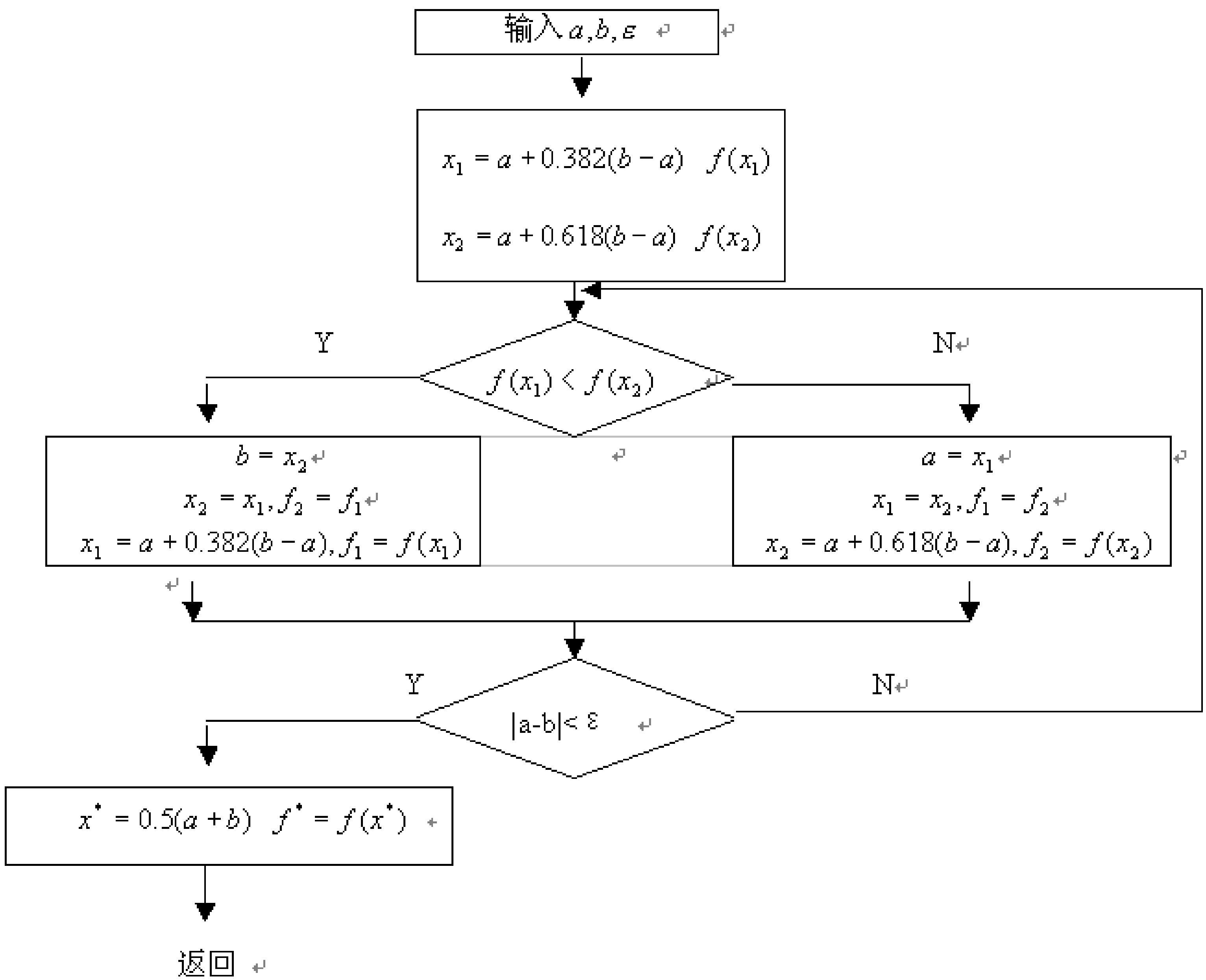


采用解析法可求得其精确解：

$$\alpha^* = -1, f(\alpha^*) = -1$$

可见通过5次迭代已比较接近精确解了

3. 算法框图



输入 a, b, ε
$x_1 = a + 0.382(b-a)$ $f(x_1)$
$x_2 = a + 0.618(b-a)$ $f(x_2)$
Y
N
$f(x_1) < f(x_2)$
$b = x_2$
$x_2 = x_1, f_2 = f_1$
$x_1 = a + 0.382(b-a), f_1 = f(x_1)$
$a = x_1$
$x_1 = x_2, f_1 = f_2$
$x_2 = a + 0.618(b-a), f_2 = f(x_2)$
Y
N
$|a-b| < \varepsilon$
$x^* = 0.5(a+b)$ $f^* = f(x^*)$
返回

3-4 一维搜索的插值类方法

在试探法中实验点位置是由某种给定的规律确定的，不考虑函数值的分布。而在插值法中，实验点是按函数值近似分布的极小点确定的

一、牛顿法

设f(x)为一个连续可微的函数，则在x_0附近，该函数应该与一个二次函数接近，即可在点x_0附近用一个二次函数φ(x)来逼近函数f(x)，即：

$$f(x) \approx \phi(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

用二次函数的φ(x)极小点x_1作为f(x)极小点的一个近似点。根据极值必要条件:

$$\phi'(x_1) = 0$$

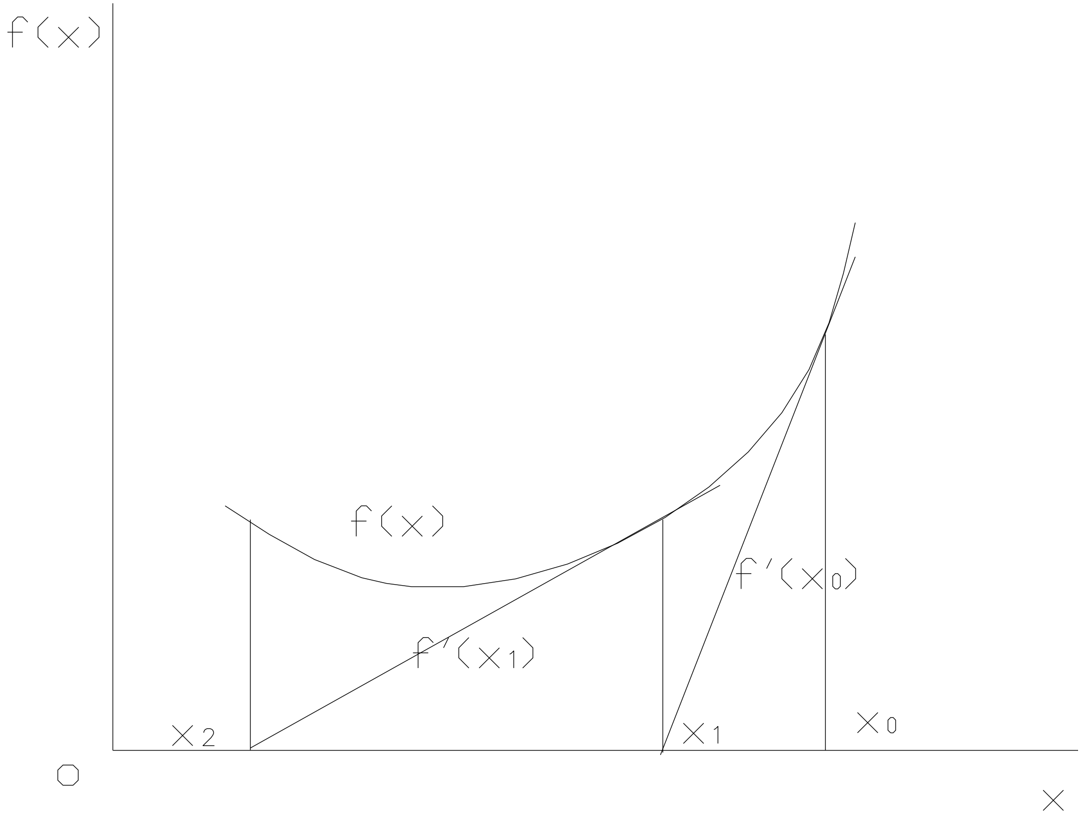
即 $$f'(x_0)+f''(x_0)(x_1-x_0)=0$$

$$x_1=x_0-\frac{f'(x_0)}{f''(x_0)}$$

依次继续下去，可得牛顿迭代公式：

$$x_{k+1}=x_k-\frac{f'(x_k)}{f''(x_k)}$$

在图中，在x_0处用一抛物线$\varphi(x)$代替曲线$f(x)$，相当于用一斜直线$\varphi'(x)$代替曲线$f'(x)$。这样各个近似点是通过对作$f'(x)$切线求得与轴的交点找到的，所以，有时，牛顿法又称作切线法。



牛顿法所作的几何解释

牛顿法的计算步骤是：

给定初始点 x_0，控制误差 ε，并令 $k=0$。

1）计算 $f'(x_k)$，$f''(x_k)$。

2）求 $x_{k+1}=x_k-\dfrac{f'(x_k)}{f''(x_k)}$。

3）若 $|x_{k+1}-x_k|\le\varepsilon$ 则求得近似解 $x^*=x_{k+1}$，停止计算，否则作 4。

4）令 $k\Leftarrow k+1$ 转 1。

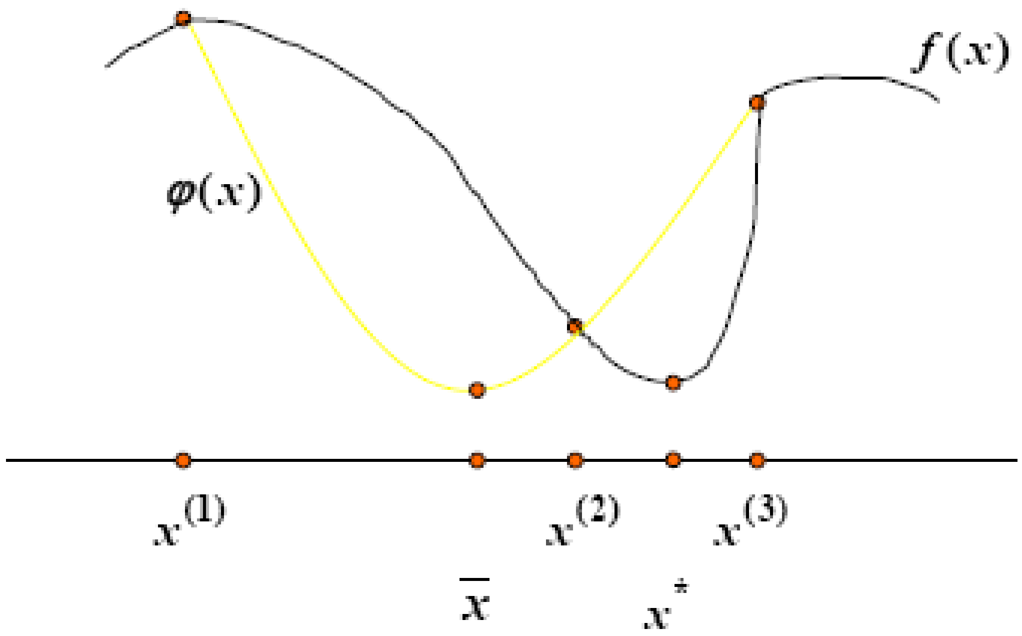
牛顿法的优点是收敛速度快。缺点是要计算函数的一阶和二阶导数，因而增加了每次迭代的工作量。如果用数值微分计算函数的二阶导数，其舍入误差将严重影响牛顿法的收敛速度， **f’(x)**的值越小，这个问题就越严重。另外，牛顿法要求初始点选的比较好，也就是说应离极小点不太远，否则有可能使极小化序列发散或收敛到非极小点。

二、抛物线法（二次插值法）

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



如何计算函数 $\varphi(x)$?

设 $$\varphi(x)=a+bx+cx^2,$$

$$\varphi(x^{(1)})=a+bx^{(1)}+cx^{(1)^2}=f(x^{(1)})$$

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

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

解上述方程组，可得逼 近函数 $\varphi(x)$ 的系数 b 和 c.

再求函数 $\varphi(x)$ 的极小点，　令

$$\varphi'(x)=b+2cx=0,$$

解得 $\bar{x}=-\dfrac{b}{2c}.$

以 $\bar{x}$ 作为 $f(x)$ 的极小点的估计值 。

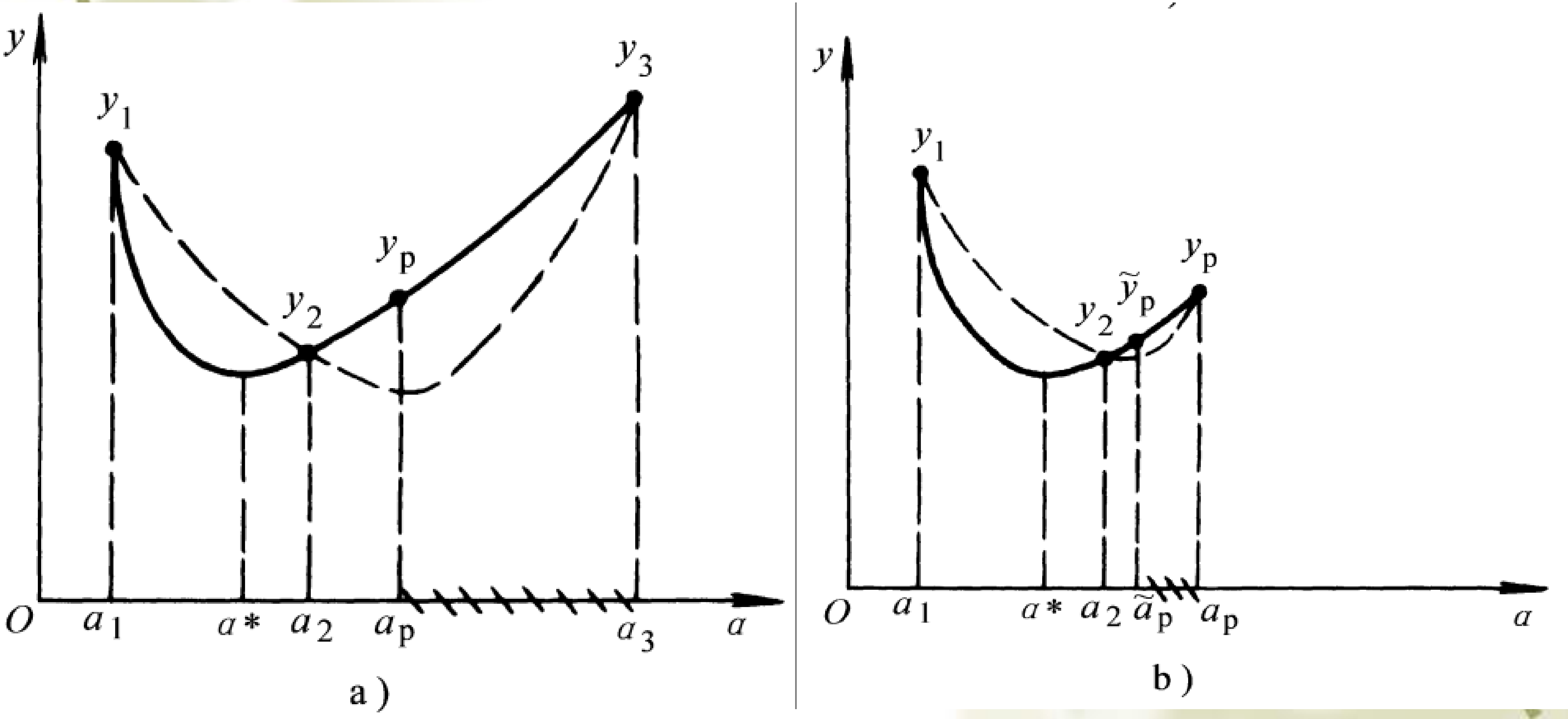
在求解一元函数f(α)的极小点时，常常利用一个低次插值多项式P(α)来逼近原目标函数，然后求出该多项式的极小点，并以此作为目标函数f(α)的近似极小点

如果p(α)为二次插值多项式，则称为二次插值法，若如果p(α)为三次插值多项式，则称为三次插值法

最常用一维搜索的插值方法是二次插值法，又称抛物线法

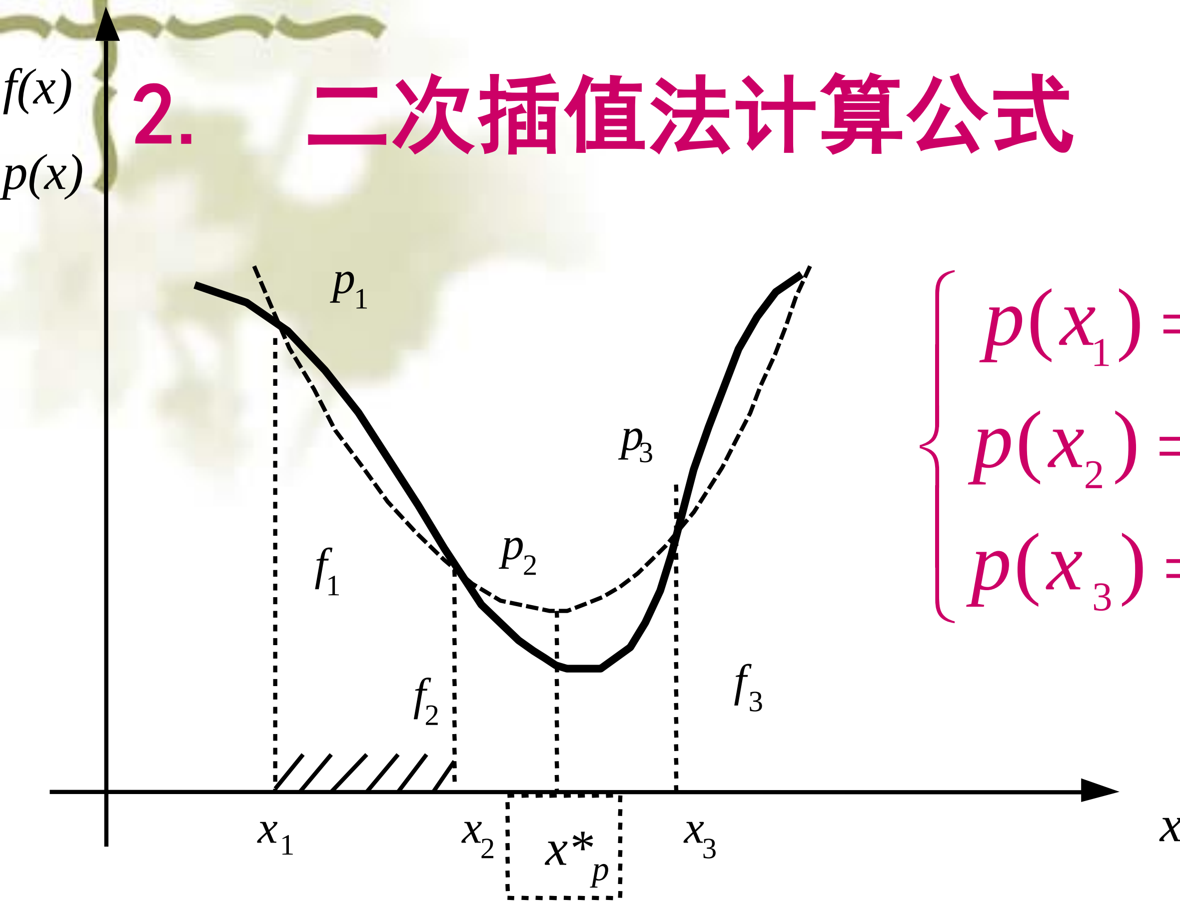
1. 二次插值法的基本原理

二次插值法的基本思想是利用目标函数在不同3点的函数值构成一个与原函数 $f(x)$ 相近似的二次多项式 $p(x)$，以函数 $p(x)$ 的极值点 α_p（即 $p(x)=0$ 的根）作为目标函数 $f(x)$ 的近似极值点。若不满足要求，缩短区间，以新的三点重新构造二次插值多项式，直至满足精度要求为止


y
y_1
y_3
y_p
y_2
O
a_1
a^*
a_2
a_p
a_3
a
a)
y
y_1
y_p
$\tilde{y}_p$
y_2
O
a_1
a^*
a_2
$\tilde{a}_p$
a_p
a
b)

书中图3—10

2. 二次插值法计算公式



$$\begin{cases} p(x_1) = a_0 + a_1 x_1 + a_2 x_1^2 = f_1 \\ p(x_2) = a_0 + a_1 x_2 + a_2 x_2^2 = f_2 \\ p(x_3) = a_0 + a_1 x_3 + a_2 x_3^2 = f_3 \end{cases}$$

求待定系数a_0, a_1和a_2, 并代入上式，得：

$$x_p^* = \frac{1}{2}\frac{(x_2^2 - x_3^2) f_1 + (x_3^2 - x_1^2) f_2 + (x_1^2 - x_2^2) f_3}{(x_2 - x_3) f_1 + (x_3 - x_1) f_2 + (x_1 - x_2) f_3}$$

3. 二次插值法程序框图

见文档

例3.2 用二次插值法求 $f(x)=x^4-4x^3-6x^2-16x+4$ 的极值点。初始搜索区间$[x_1\ x_3]=[-1\ \ 6]$，$\varepsilon=0.05$

解：　取x2点为区间[x1, x3]的中点

$$x_2=0.5\times(x_1+x_3)=2.5$$

计算x1, x2, x3 3点处的函数值

f1=19，f2=-96.9375，f3=124。

可见函数值满足“高—低—高”形态。

以x1, x2, x3为插值点构造二次曲线，

求第一次近似的二次曲线p(x)的极小值点，由公式

$$x_p^* = \frac{1}{2}\frac{(x_2^2 - x_3^2) f_1 + (x_3^2 - x_1^2) f_2 + (x_1^2 - x_2^2) f_3}{(x_2 - x_3) f_1 + (x_3 - x_1) f_2 + (x_1 - x_2) f_3}$$

得 $x_p^* = 1.9545$

$$f(x_p^*) = -65.4648 > f(x_2) = -96.9375$$

比较数据, 应消除左边区段 $[x_1, x_p^*]$。然后用 x_p^*, x_2, x_3 作为*x*1, *x*2, *x*3新3点, 重新构造二次曲线*p*(*x*)，如此反复计算，直到 $|x_2 - x_p^*| < \varepsilon$ 为止。

从表3-1中可见，经7次迭代后, $|x_2 - x_p^*| = 0.0378 < \varepsilon = 0.05$ 终止迭代。故最优点 $x^* = x_p^* = 3.9501$

表3－1二次插值法的计算过程示例

	x_1	x_2	x_3	x_p^*	$f(x_p^*)$
1	-1.0	2.5	6	1.9545	-65.4648
2	1.9545	2.5	6	3.1932	-134.5394
3	2.5	3.1932	6	3.4952	-146.7761
4	3.1932	3.4952	6	3.7268	-153.1043
5	3.4952	3.7268	6	3.8403	-154.9771
6	3.7268	3.8402	6	3.9123	-155.6850
7	3.8403	3.9123	6	3.9501	-155.8969

例3. 3 用二次插值法求 $f(\alpha)=\sin\alpha$ 在 $4\le\alpha\le5$ 上的极小点

解：初始的搜索区间为[4, 5], 取 $\alpha1=4, \alpha2=4.5, \alpha3=5$

经过两次的计算过程及结果如表3－2所示

表3－2二次插值法的计算过程示例

	1	2
$\alpha1$	4	4.5
$\alpha2$	4.5	4.705120
$\alpha3$	5	5
y1	-0.756802	-0.977590
y2	-0.977590	-0.999974
y3	-0.958924	-0.958924
αp	4.705120	4.710594
yp	-0.999974	-0.999998

经过两次计算解得最优解为

$$\alpha^*=4.710594$$

$$f(\alpha^*)=-0.999998$$

这和精确值－1已经非常接近

例 3.4 用二次插值法求函数$f(x)=3x^3-4x+2$的极小点，给定 $x_0=0$, $\varepsilon=0.2$。

- 解

1）确定初始区间

初始区间$[a,b]=[0,2]$, 中间点$x_2=1$。

2）用二次插值法逼近极小点

相邻三点的函数值: $x_1=0$, $x_2=1$, $x_3=2$; $f_1=2$, $f_2=1$, $f_3=18$. 代入公式:

$$x_p^* = \frac{1}{2}\frac{(x_2^2-x_3^2)f_1+(x_3^2-x_1^2)f_2+(x_1^2-x_2^2)f_3}{(x_2-x_3)f_1+(x_3-x_1)f_2+(x_1-x_2)f_3}$$

$x_p^*=0.555$, $f_p=0.292$

由于$f_p<f_2$, $x_p*<x_2$, 新区间$[a,b]=[a,x_2]=[0,1]$
$|x_2-x_p*|=1-0.555=0.445>0.2$, 应继续迭代。

- 在新区间，相邻三点的函数值: $x_1=0$, $x_2=0.555$, $x_3=1$; $f_1=2$, $f_2=0.292$, $f_3=1$.
- $x_p*=0.607$, $f_p=0.243$

由于$f_p<f_2$, $x_p>x_2$, 新区间$[a,b]=[x_2, b]=[0.555,1]$
$|x_2-x_p*|=|0.555-0.607|=0.052<0.2$, 迭代终止。

$$x_p*=0.607, \quad f*=0.243$$

4 二次插值法和黄金分割法的比较

	0.618法	二次插值法
实验点的确定方式	按规定确定	按函数值近似分布的极小点确定
效率	较低（仅用2点信息）	较高
分类	试探法	插值法
信息量	少（两点）	大（三点）

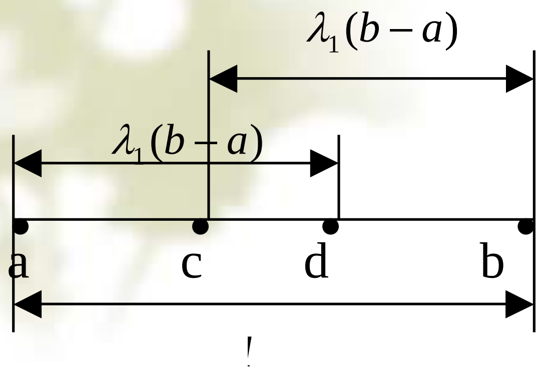
3-5 分数法

1.费波那契数列{Fn}：

费波那契数列满足以下递推关系：

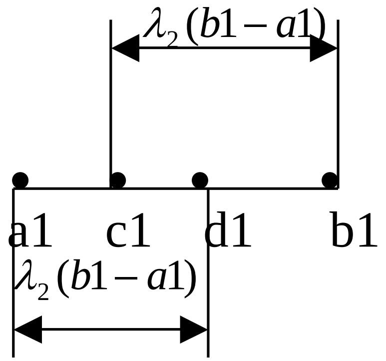
n	0	1	2	3	4	5	6	7	8	9	10	11	12	……
Fn	1	1	2	3	5	8	13	21	34	55	89	144	…	

以试点n-5为例：



$$\lambda 1=\frac{Fn-1}{Fn}=\frac{F4}{F5}=5/8$$

$$\lambda 2=\frac{Fn-2}{Fn-1}=\frac{F3}{F4}=3/5$$



$$\lambda 3=\frac{Fn-3}{Fn-2}=\frac{F2}{F3}=2/3$$

$$\lambda 4=\frac{Fn-4}{Fn-3}=\frac{F1}{F2}=1/2$$

经过五次计算后，总收缩率λ=λ1·λ2·λ3·λ4=1/8

计算n次，函数值所获得的区间总缩短率为

λ=λ1·λ2...·λn−1=F1/Fn

1.迭代步骤：

(1)用进退法确定区间 [a , b]

(2)确定试点的个数n

(b-a)经过n次计算（b_{n-1}－a_{n-1}）,通过b_{n-1}－a_{n-1}≤ ε ， 来确定n

$$b_{n-1}-a_{n-1}=\frac{b-a}{Fn}\leqslant \varepsilon \rightarrow Fn\geqslant \frac{b-a}{\varepsilon}\rightarrow n$$

(3)计算试点：

(a)计算 [a b]内的x1 , x2

$$x1=a+(b-a)\frac{Fn-1}{Fn} \qquad x2=a+b-x1$$

(b)比较x1 x2的函数值f(x1)和f(x2)

若f1>f2 , b←x1   x1←x2

x2=a+b—x1

若f1<f2 ,a←x2   x2←x1

x1=a+b—x2

(c)当b—a≤ε,终止迭代， X*=(b+a)/2

否则转到（3）

思考：

分数法和黄金分割法，二次插值法有何不同？

例3.4　一个圆柱螺旋压缩弹簧，不考虑共振，要求重量W最轻。

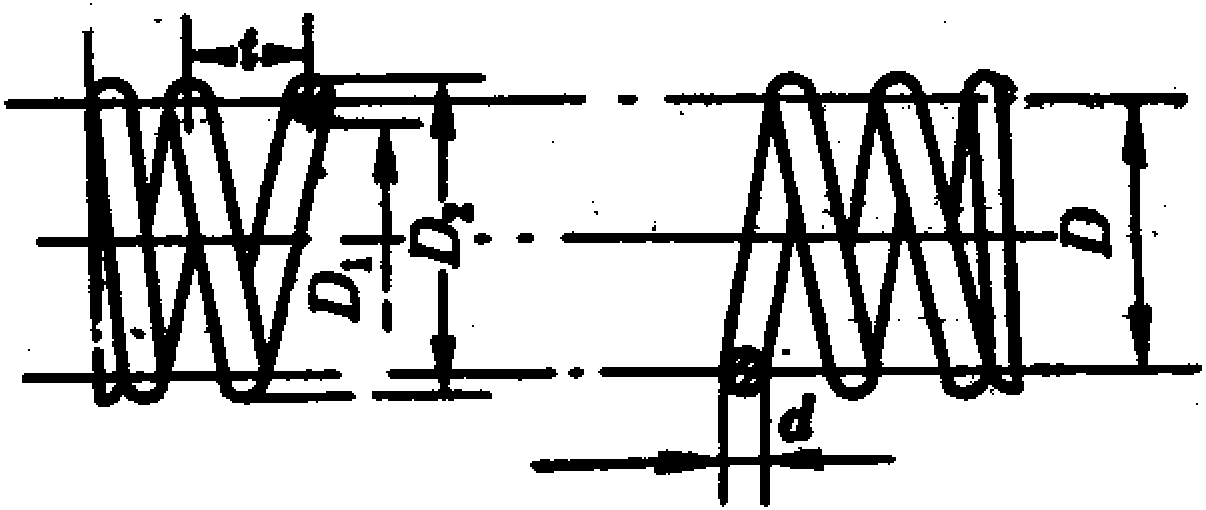
解：先列出弹簧的有关设计计算公式： $\tau = \dfrac{8KFC}{\pi d^2}$　式（3—1）

其中　τ – 最大工作负荷所产生的最大切应力

F – 最大工作负荷

D – 弹簧中径

d – 弹簧丝直径



C – 缠绕比　$C = \dfrac{D}{d}$

K – 曲度系数

$$K = \frac{4C-1}{4C-4} + \frac{0.615}{C}$$

将式（3—1）$\tau = \frac{8KFC}{\pi d^2}$ 变形

$$d = \sqrt{\frac{8KFC}{\pi \tau}}$$

$$d_{\min} = \sqrt{\frac{8KFC}{\pi[\tau]}}$$ 式（3—2）

$[\tau]$为弹簧许用应力

弹簧的重量公式，也即目标函数为

$$W=(n+n2)\pi^2(\frac{Cd^3}{4})\gamma \quad \text{式（3—3）}$$

其中 $n-$工作有效圈数

n_2-不起作用圈数（总圈数与工作有效圈数之差）

γ—材料密度

$C-$缠绕比

$$n=\frac{Gd\lambda}{8FC^3} \quad \text{式（3—4）}$$

$G-$常温下剪切弹性模量

λ—弹簧的总变形量

$F-$最大工作负荷

$d-$弹簧丝直径

将式（3—2），式（3—4），带入式（3—3）得，

$$W=\left(\frac{\lambda G}{8FC^{3}}\sqrt{\frac{8KFC}{\pi[\tau]}}+n2\right)\pi^{2}\frac{C}{4}\left(\frac{8KFC}{\pi[\tau]}\right)^{\frac{3}{2}}\gamma$$

从上式可看出，除变量C以外，其余参数都可以由弹簧设计要求和选定材料来给出。所以问题变为minW=f(C)

可见其为一维搜索问题

第三章作业：

1. 黄金分割法中内部两个点的选择有何特点？并阐述黄金分割法的基本原理。

2用二次插值法求函数f(x)=3x^3−4x+2的极小点，给定 x_1=0，x_2=1，x_2=2 ， ε =0.2。