第九章　一维搜索(line search)

$$(LS)\quad \min_{\lambda} f(x^{(k)}+\lambda d^{(k)})$$

如果求得的λ_k，使得

$$f\left(x^{(k)}+\lambda_k d^{(k)}\right)=\min_{\lambda} f\left(x^{(k)}+\lambda d^{(k)}\right)$$

则称该一维搜索为精确一维搜索(Exact Line Search).

称λ_k为最优步长.

如果存在λ_k，使得

$$f\left(x^{(k)}+\lambda_k d^{(k)}\right)<f\left(x^{(k)}\right)$$

则称该一维搜索为非精确一维搜索(Inexact Line Search).

- 一元函数的极值问题, 全局最优步长难求. 通常只能求得局部(近似)最优步长。

- 精确一维搜索通常有两种实现方式:
- （1）试探法：按某种方式找试探点，通过一系列试探点来确定极小点。
- （2）函数逼近法（插值法）：用某种较简单的曲线逼近原来的函数曲线，通过求逼近函数的极小点来估计目标函数的极小点。

基本性质

$$f(x_k+\alpha_k d_k) \leqslant f(x_k+\alpha d_k), \quad \forall\ \alpha > 0$$

$$f'_\alpha(\boldsymbol{x}_k+\alpha_k \boldsymbol{d}_k) = \boldsymbol{d}_k^{\mathrm{T}} \nabla f(\boldsymbol{x}_k+\alpha_k \boldsymbol{d}_k) = 0.$$

一维搜索的闭性

定义：算法映射$M: E^n \times E^n \to 2^{E^n}$定义为：

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

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

定理：设$f(x)$是定义在E^n上的连续函数，$d\neq 0$,
则算法映射M在(x,d)处是闭的。

证明：设$\{x^{(k)}\}$和$\{d^{(k)}\}$是两个序列，且$\left(x^{(k)},d^{(k)}\right)\to(x,d)$,

又设$y^{(k)}\in M\left(x^{(k)},d^{(k)}\right)$,且$y^{(k)}\to y$。

以下证明$y\in M(x,d)$。

对每个k，存在$\lambda_k \geq 0$，使得$y^{(k)} = x^{(k)} + \lambda_k d^{(k)}$ (1)

$\because d \neq 0, \therefore$ 当k充分大时，必有$d^{(k)} \neq 0$，因此

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

$$\Rightarrow \lim_{k \to \infty} \lambda_k = \bar{\lambda} = \frac{\|y - x\|}{\|d\|}$$

（1）中令$k \to \infty$,得$y = x + \bar{\lambda} d$

根据M的定义，对每个k及$\lambda \geq 0$，有$f\left(y^{(k)}\right) \leq f\left(x^{(k)} + \lambda d^{(k)}\right)$

由于f连续，令$k \to \infty$，得$f(y) \leq f(x + \lambda d)$

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

$$\Rightarrow y \in M(x, d).$$

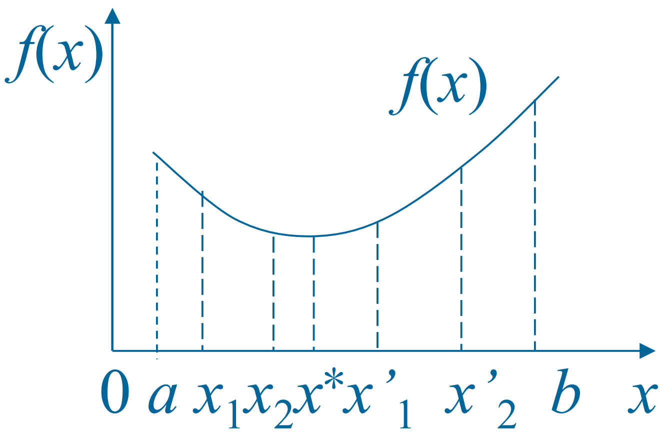
试探法

0.618法(黄金分割法)(Golden Section Method)

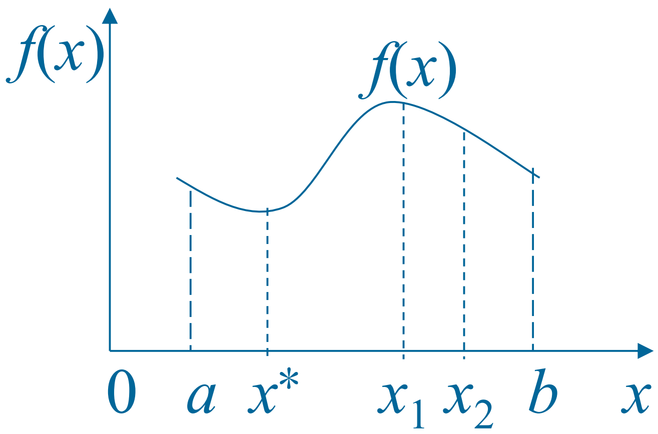
定义: 称$f(x)$在闭区间$[a,b]$上是单峰(unimodal)的，是指$f(x)$在$[a,b]$内有唯一的一点x^*，使得对任意的$x_1,x_2\in[a,b]$, $x_1<x_2$，有

当$x_2\le x^*$时，$f(x_1)>f(x_2)$,

当$x^*\le x_1$时，$f(x_2)>f(x_1)$。



(a)



(b)

性质：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。

定理：设$f(x)$是$[a,b]$上的单峰函数，$x_1, x_2 \in [a,b]$且$x_1 < x_2$，

若$f(x_1) > f(x_2)$，则对任意$x \in [a, x_1]$，有$f(x) > f(x_2)$，

若$f(x_1) \le f(x_2)$，则对任意$x \in [x_2, b]$，有$f(x) \ge f(x_1)$。

0.618法的基本思想：

通过取试探点使包含极小点的区间不断缩短，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，因此任意一点都可以作为极小点的近似。

0.618法的计算公式：

设f在$[a,b]$上单峰，极小点$\bar{x}\in[a,b]$，设进行第k次迭代时，有$\bar{x}\in[a_k,b_k]$，取试探点$\lambda_k,\mu_k\in[a_k,b_k]$，规定$\lambda_k<\mu_k$，计算$f(\lambda_k),f(\mu_k)$。

1．若$f(\lambda_k)>f(\mu_k)$，则有$\bar{x}\in[\lambda_k,b_k]$，令$a_{k+1}=\lambda_k,b_{k+1}=b_k$.

2．若$f(\lambda_k)\le f(\mu_k)$，则有$\bar{x}\in[a_k,\mu_k]$，令$a_{k+1}=a_k,b_{k+1}=\mu_k$.

确定λ_k,μ_k,使它们满足：

（1）λ_k,μ_k在$[a_k,b_k]$中的位置是对称的，即

$$b_k-\mu_k=\lambda_k-a_k.$$

（2）每次迭代区间长度缩短比例相同，即

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

$$\begin{cases} b_k - \mu_k = \lambda_k - a_k & (1) \\ b_{k+1} - a_{k+1} = \alpha(b_k - a_k) & (2) \end{cases}$$

当$f(\lambda_k) > f(\mu_k)$时，$a_{k+1} = \lambda_k$，$b_{k+1} = b_k$，

代入(2)得 $b_k - \lambda_k = \alpha(b_k - a_k)$

当$f(\lambda_k) \le f(\mu_k)$时，$a_{k+1} = a_k$，$b_{k+1} = \mu_k$，

代入(2)得 $\mu_k - a_k = \alpha(b_k - a_k)$

$$\Rightarrow \quad \begin{aligned} \lambda_k &= a_k + (1-\alpha)(b_k - a_k) \\ \mu_k &= a_k + \alpha(b_k - a_k) \end{aligned}$$

设在第k次迭代得出$f(\lambda_k)\le f(\mu_k)$，则

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

在第$k+1$次迭代中，需要取试探点λ_{k+1}，μ_{k+1}

$$\begin{aligned}\mu_{k+1}&=a_{k+1}+\alpha(b_{k+1}-a_{k+1})=a_k+\alpha(\mu_k-a_k)\\&=a_k+\alpha(a_k+\alpha(b_k-a_k)-a_k)\\&=a_k+\alpha^2(b_k-a_k).\end{aligned}$$

$$\boxed{\begin{aligned}\lambda_k&=a_k+(1-\alpha)(b_k-a_k)\\\mu_k&=a_k+\alpha(b_k-a_k)\end{aligned}}$$

若令$\alpha^2=1-\alpha$，则$\mu_{k+1}=\lambda_k$，

$\Rightarrow \mu_{k+1}$不必重新计算，减少计算量。

设在第k次迭代得出$f(\lambda_k) > f(\mu_k)$，则

$$[a_{k+1}, b_{k+1}] = [\lambda_k, b_k]$$

在第$k+1$次迭代中，需要取试探点λ_{k+1}，μ_{k+1}

$$\begin{aligned}
\lambda_{k+1} &= a_{k+1} + (1-\alpha)(b_{k+1} - a_{k+1}) = \lambda_k + (1-\alpha)(b_k - \lambda_k) \\
&= a_k + (1-\alpha)(b_k - a_k) + (1-\alpha)(b_k - a_k - (1-\alpha)(b_k - a_k)) \\
&= a_k + (1-\alpha)(b_k - a_k) + (1-\alpha)\alpha(b_k - a_k) \\
&= a_k + (1-\alpha^2)(b_k - a_k).
\end{aligned}$$

$$\boxed{\begin{aligned} \lambda_k &= a_k + (1-\alpha)(b_k - a_k) \\ \mu_k &= a_k + \alpha(b_k - a_k) \end{aligned}}$$

若令$\alpha = 1-\alpha^2$，即$\alpha^2 = 1-\alpha$，则$\lambda_{k+1} = \mu_k$，

$\Rightarrow \lambda_{k+1}$不必重新计算，减少计算量。

解方程 $\alpha^2 = 1-\alpha$ 得

$$\alpha = \frac{-1 \pm \sqrt{5}}{2}$$

$$\because \alpha > 0, \quad \therefore \alpha = \frac{-1+\sqrt{5}}{2} \approx 0.618$$

$$\lambda_k = a_k + 0.382(b_k - a_k)$$

$$\mu_k = a_k + 0.618(b_k - a_k)$$

步骤：

1.置初始区间$[a_1, b_1]$及精度要求$L>0$，计算

$$\lambda_1 = a_1 + 0.382(b_1 - a_1), \mu_1 = a_1 + 0.618(b_1 - a_1),$$

计算$f(\lambda_1)$和$f(\mu_1)$。令$k=1$。

2.若$b_k - a_k < L$，则停止计算；否则当$f(\lambda_k) > f(\mu_k)$时，转3；当$f(\lambda_k) \leq f(\mu_k)$时，转4。

3.置$a_{k+1} = \lambda_k, b_{k+1} = b_k, \lambda_{k+1} = \mu_k,$

$\mu_{k+1} = a_{k+1} + 0.618(b_{k+1} - a_{k+1})$,计算$f(\mu_{k+1})$，转5。

4.置$a_{k+1} = a_k, b_{k+1} = \mu_k, \mu_{k+1} = \lambda_k,$

$\lambda_{k+1} = a_{k+1} + 0.382(b_{k+1} - a_{k+1})$,计算$f(\lambda_{k+1})$，转5。

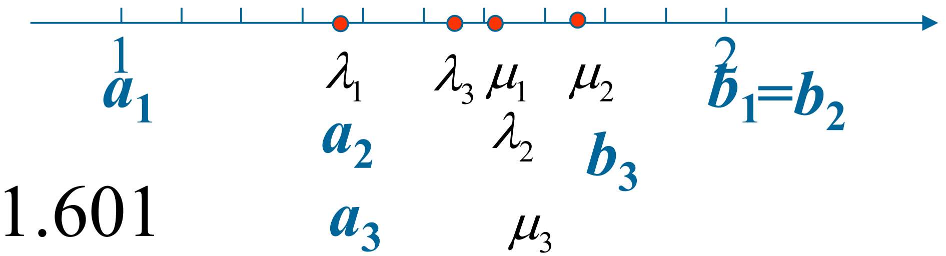
5.置$k = k+1$，返回2。

优点:不要求函数可微,甚至当函数不连续时，
0．618法仍可应用。

缺点：收敛比较慢，**0．618**法只适用于单峰函数，所以需要先确定单峰区间，再使用**0．618**法的计算公式。

例： $\min e^x - 5x \quad (1 \le x \le 2), L \le 0.04.$

k	a_k	b_k	λ_k	μ_k	$f(\lambda_k)$	$f(\mu_k)$	$b_k - a_k$
1	1	2	1.382	1.618	−2.928	−3.048	1
2	1.382	2	1.618	1.764	−3.048	−2.985	0.618
3	1.382	1.764	1.528	1.618	−3.032	−3.048	0.382
4	1.528	1.764	1.618	1.674	−3.048	−3.037	0.236
5	1.528	1.674	1.584	1.618	−3.046	−3.048	0.146
6	1.584	1.674	1.618	1.640	−3.048	−3.046	0.09
7	1.584	1.640	1.605	1.618	−3.048	−3.048	0.056
8	1.584	1.618					0.036



$$x^* = \frac{1.584 + 1.618}{2} = 1.601$$

函数逼近法

一．牛顿法

基本思想：在极小点附近用二阶**Taylor**多项式近似。

$$\min \quad f(x)$$

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

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

得$\varphi(x)$的驻点，记作$x^{(k+1)}$，则

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

算法收敛定理

定理： 设A为X上的一个算法，Ω为解集合，给定初点$x^{(1)} \in X$，进行如下迭代：

若$x^{(k)} \in \Omega$，则停止迭代；否则，令$x^{(k+1)} \in A(x^{(k)})$，用$k+1$代替k，重复该过程。

如果下面的条件成立：

1．序列$\{x^{(k)}\}$含于X的某个紧子集中；

2．存在一个连续函数α，它是关于Ω和A的下降函数；

3．映射A在Ω的补集上是闭的。

则序列$\{x^{(k)}\}$的任一收敛子序列的极限属于Ω。

定理：设$f(x)$存在连续三阶导数，$\bar{x}$满足$f'(\bar{x})=0$，$f''(\bar{x})\neq 0$，初点$x^{(1)}$充分接近$\bar{x}$，则牛顿法产生的序列$\{x^{(k)}\}$至少以二级收敛速度收敛于$\bar{x}$。

证明：牛顿法可定义为算法映射

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

设解集合为$\Omega=\{\bar{x}\}$

定义函数 $\alpha(x)=|x-\bar{x}|$

设 $x^{(k)}\neq\bar{x}, x^{(k+1)}\in A(x^{(k)})$

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

$$\alpha(x^{(k+1)}) = | x^{(k+1)} - \bar{x} |$$

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

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

$$= \frac{1}{| f''(x^{(k)}) |} \times \frac{1}{2} (\bar{x} - x^{(k)})^2 \, | f'''(\xi) |$$ 其中ξ在x和$\bar{x}$之间。

$\because f''(x)$和$f'''(x)$连续，$f''(\bar{x}) \neq 0, \therefore$ 当$x^{(k)}$接近$\bar{x}$时，存在 $k_1, k_2 > 0$使得在包含$\bar{x}$和$x^{(k)}$的闭区间上的每一点x处，有

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

$\therefore \left| x^{(k+1)} - \bar{x} \right| \leq \dfrac{k_2}{2k_1} (\bar{x} - x^{(k)})^2 \Rightarrow \left\{ x^{(k)} \right\}$至少是二级收敛的。

$$\therefore \left| x^{(k+1)} - \bar{x} \right| \le \frac{k_2}{2k_1} (\bar{x} - x^{(k)})^2$$

取初点$x^{(1)}$充分接近$\bar{x}$，使得

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

$$\therefore \{x^{(k)}\} \subset X = \{x \mid |x - \bar{x}| \le |x^{(1)} - \bar{x}|\}$$

且有$|x^{(k+1)} - \bar{x}| < |x^{(k)} - \bar{x}|$

$\therefore \alpha$是下降函数，且X为紧集，$A(x)$在X上连续

$\therefore \{x^{(k)}\} \to \bar{x}$(由算法收敛定理)。

步骤：

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

2.若$|f'(x^{(k)})| < \varepsilon$，则停止计算，得点$x^{(k)}$；否则转3.

3.计算点$x^{(k+1)}$

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

置$k = k+1$，返回2。

缺点：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。

例： $\min e^{x}-5x \quad (1\le x\le 2), \varepsilon < 0.01.$

解： 取初始点$x^{(0)}=2$

k	$x^{(k)}$	$f'\left(x^{(k)}\right)$	$f''\left(x^{(k)}\right)$
0	2	2.388	7.388
1	1.677	0.349	5.349
2	1.612	0.012	5.012
3	1.6096	-0.00002	

$\therefore \quad x=1.6096$为近似解。

例： $f(x)=\min\int_0^x \operatorname{arctg} t\, dt$　　最优解 $x^*=0.$

用**Newton**法求解： $f'(x)=arctg\ x,\quad f''(x)=\dfrac{1}{1+x^2}$

k	x_k	$f'(x_k)$	$1/f''(x_k)$
1	1	0.7854	2
2	-0.5708	-0.5178	1.3258
3	0.1169	0.1163	1.0137
4	-0.001061		

k	x_k	$f'(x_k)$	$1/f''(x_k)$
1	2	1.1071	5
2	-3.5357	-1.2952	13.50
3	13.95		

二．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，令$\varphi(x)$与$f(x)$在三点$x^{(1)} < x^{(0)} < x^{(2)}$处有相同的函数值，并假设

$$f(x^{(1)}) > f(x^{(0)}),\quad f(x^{(0)}) < f(x^{(2)}).$$

令 $$\varphi(x) = a + bx + cx^2$$

令
$$\varphi(x^{(0)}) = a + bx^{(0)} + cx^{(0)2} = f(x^{(0)})$$
$$\varphi(x^{(1)}) = a + bx^{(1)} + cx^{(1)2} = f(x^{(1)})$$
$$\varphi(x^{(2)}) = a + bx^{(2)} + cx^{(2)2} = f(x^{(2)})$$

求$\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$。

记 $B_1=({x^{(0)}}^2-{x^{(2)}}^2)f(x^{(1)})$ $\quad B_2=({x^{(2)}}^2-{x^{(1)}}^2)f(x^{(0)})$

$B_3=({x^{(1)}}^2-{x^{(0)}}^2)f(x^{(2)})$ $\quad C_1=(x^{(0)}-x^{(2)})f(x^{(1)})$

$C_2=(x^{(2)}-x^{(1)})f(x^{(0)})$ $\quad C_3=(x^{(1)}-x^{(0)})f(x^{(2)})$

$D=(x^{(1)}-x^{(0)})(x^{(0)}-x^{(2)})(x^{(2)}-x^{(1)})$

则 $b=\dfrac{B_1+B_2+B_3}{D}$, $\quad c=-\dfrac{C_1+C_2+C_3}{D}$.

令 $\varphi'(x)=b+2cx=0, \quad \Rightarrow \quad x=-\dfrac{b}{2c}$.

令 $\overline{x}^{(k)}=x$, 则

$$x^{(3)}=\overline{x}^{(k)}=\frac{B_1+B_2+B_3}{2(C_1+C_2+C_3)}.$$

收敛准则： $\quad |x^{(3)}-x^{(0)}|<\varepsilon$

给定$x_0, x_1, x_2, x_1 < x_0 < x_2, \varepsilon > 0$

$$F_0 = f(x_0), F_1 = f(x_1), F_2 = f(x_2)$$

$$F_0 \le F_1, \quad F_0 \le F_2$$

$$A = 2[(x_0 - x_2)F_1 + (x_2 - x_1)F_0 + (x_1 - x_0)F_2]$$

$$A = 0$$

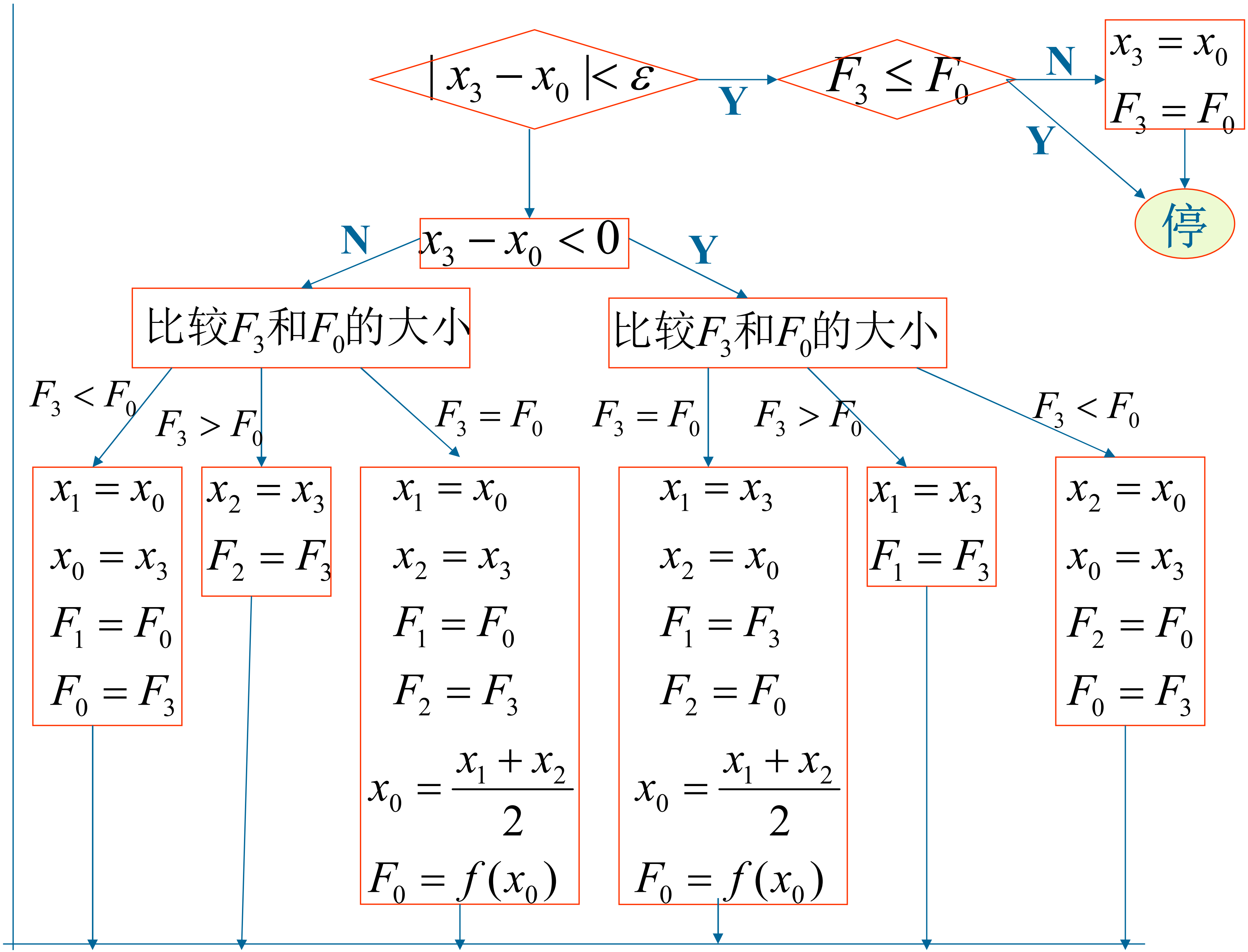
Y

停

N

$$x_3 = \frac{1}{A}[(x_0^2 - x_2^2)F_1 + (x_2^2 - x_1^2)F_0 + (x_1^2 - x_0^2)F_2]$$

$$F_3 = f(x_3)$$



$|x_3 - x_0| < \varepsilon$
Y
$F_3 \le F_0$
N
$x_3 = x_0$
$F_3 = F_0$
Y
停
N
$x_3 - x_0 < 0$
Y
比较F_3和F_0的大小
比较F_3和F_0的大小
$F_3 < F_0$
$F_3 > F_0$
$F_3 = F_0$
$F_3 = F_0$
$F_3 > F_0$
$F_3 < F_0$
$x_1 = x_0$
$x_0 = x_3$
$F_1 = F_0$
$F_0 = F_3$
$x_2 = x_3$
$F_2 = F_3$
$x_1 = x_0$
$x_2 = x_3$
$F_1 = F_0$
$F_2 = F_3$
$x_0 = \frac{x_1 + x_2}{2}$
$F_0 = f(x_0)$
$x_1 = x_3$
$x_2 = x_0$
$F_1 = F_3$
$F_2 = F_0$
$x_0 = \frac{x_1 + x_2}{2}$
$F_0 = f(x_0)$
$x_1 = x_3$
$F_1 = F_3$
$x_2 = x_0$
$x_0 = x_3$
$F_2 = F_0$
$F_0 = F_3$

例：$\min e^x - 5x \quad (1 \le x \le 2), \varepsilon = 0.04.$

解： 取初始点$x_1 = 1, x_0 = 1.5, x_2 = 2$

k	x_1	x_0	x_2	F_1	F_0	F_2	A	x_3	F_3
0	1	1.5	2	-2.282	-3.019	-2.612	0.569	1.573	-3.044
1	1.5	1.573	2	-3.019	-3.044	-2.612	0.084	1.608	-3.048

$\because |x_3 - x_0| = 0.035 < 0.04$

且$F_3 \le F_0$

$\therefore \quad x = 1.608$为近似解。

三次插值法

选取两个初始点x_1和$x_2\,(x_1 < x_2)$,使得$f'(x_1) < 0$及$f'(x_2) > 0$.

令$\varphi(x) = a(x - x_1)^3 + b(x - x_1)^2 + c(x - x_1) + d$使得

$$\varphi(x_1) = f(x_1), \quad \varphi'(x_1) = f'(x_1)$$

$$\varphi(x_2) = f(x_2), \quad \varphi'(x_2) = f'(x_2)$$

得
$$\begin{cases} d = f(x_1) \\ c = f'(x_1) < 0 \\ a(x_2 - x_1)^3 + b(x_2 - x_1)^2 + c(x_2 - x_1) + d = f(x_2) \\ 3a(x_2 - x_1)^2 + 2b(x_2 - x_1) + c = f'(x_2) \end{cases}$$

求$\varphi(x)$的极小点，

$$\varphi'(x)=3a(x-x_1)^2+2b(x-x_1)+c$$

$$\varphi''(x)=6a(x-x_1)+2b.$$

若$a=0$，则有$\bar{x}-x_1=-\dfrac{c}{2b}$，$\varphi''(\bar{x})=2b$，

由于$c=f(x')<0$，$2b(x_2-x_1)+c=f'(x_2)>0$

$$\therefore 2b=\frac{f'(x_2)-c}{x_2-x_1}>0,$$

$\Rightarrow \bar{x}$为极小点。

求$\varphi(x)$的极小点，

$$\varphi'(x)=3a(x-x_1)^2+2b(x-x_1)+c$$

$$\varphi''(x)=6a(x-x_1)+2b.$$

若$a\neq 0$,则有$\bar{x}-x_1=\dfrac{-b\pm\sqrt{b^2-3ac}}{3a}$, $\varphi''(\bar{x})=\pm 2\sqrt{b^2-3ac}$,

为使$\varphi''(\bar{x})>0$,应取

$$\bar{x}-x_1=\frac{-b+\sqrt{b^2-3ac}}{3a}=\frac{-c}{b+\sqrt{b^2-3ac}}.$$

$\Rightarrow$ $\bar{x}$为极小点。

$$\bar{x}-x_1=\frac{-b+\sqrt{b^2-3ac}}{3a}=\frac{-c}{b+\sqrt{b^2-3ac}}.$$

步骤：

1.给定初始点$x^{(1)}, x^{(2)}$，计算$f(x^{(1)}), f(x^{(2)}), f'(x^{(1)}), f'(x^{(2)})$

要求满足条件：

$$x^{(2)} > x^{(1)}, f'(x^{(1)}) < 0, f'(x^{(2)}) > 0,$$

给定允许误差$\delta > 0.$

2.计算$s = \dfrac{3[f(x^{(2)}) - f(x^{(1)})]}{x^{(2)} - x^{(1)}}$，$\quad z = s - f'(x^{(1)}) - f'(x^{(2)}),$

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

$$\bar{x} = x^{(1)} + (x^{(2)} - x^{(1)})\left(1 - \frac{f'(x^{(2)}) + w + z}{f'(x^{(2)}) - f'(x^{(1)}) + 2w}\right)。$$

3.若$\left|x^{(2)}-x^{(1)}\right|\leq\delta$，则停止计算，得点$\overline{x}$；否则转4。

4.计算$f(\overline{x})$，$f'(\overline{x})$。若$f'(\overline{x})=0$，则停止计算，得点$\overline{x}$。

若$f'(\overline{x})<0$，则令$x^{(1)}=\overline{x}$，$f(x^{(1)})=f(\overline{x})$，$f'(x^{(1)})=f'(\overline{x})$，转2。

若$f'(\overline{x})>0$，则令$x^{(2)}=\overline{x}$，$f(x^{(2)})=f(\overline{x})$，$f'(x^{(2)})=f'(\overline{x})$，转2。

例：$\min e^x - 5x \quad (1 \le x \le 2), \varepsilon = 0.01.$

解： 取初始点$x_1 = 1, x_2 = 2, \delta = 0.01.$

k	x_1	x_2	$f(x_1)$	$f(x_2)$	$f'(x_1)$	$f'(x_2)$	$\bar{x}$	$f(\bar{x})$	$f'(\bar{x})$
0	1	2	-2.282	-2.612	-2.282	2.388	1.606	-3.048	-0.018
1	1.606	2	-3.048	-2.612	-0.018	2.388	1.6096	-3.048	-0.00002

$$\because |f'(\bar{x})| < \delta$$

$\therefore \quad x = 1.6096$为近似解。

最优步长小结

$$\alpha_k = \arg\min_{\alpha \geq 0} f(x_k + \alpha d_k)$$

- ✔ 目标函数沿搜索方向下降量达到最大

  理想化策略，计算量大，很难取到

- ✔ 便于算法理论分析（如二次终止性、收敛速率等）

- ✘ 该步长搜索对全局最优解的帮助不大

非精确线搜索方法

最优步长规则的初衷是在每一迭代步使目标函数的下降量达到最大. 尽管最优步长的计算是一单元函数的极值问题，但计算量太大，而且不好求.

最优化关注目标函数在整个可行域上的最优值点. 集中于某个方向上的线搜索似乎没有必要. 这就引出了非精确线搜索步长规则.

常见的线搜索步长规则

Armijo步长规则、 Goldstein步长规则、 Wolfe步长规则