一　维　最　优　化

1. 一维最优化问题

2. 黄金分割法

3. 进退法

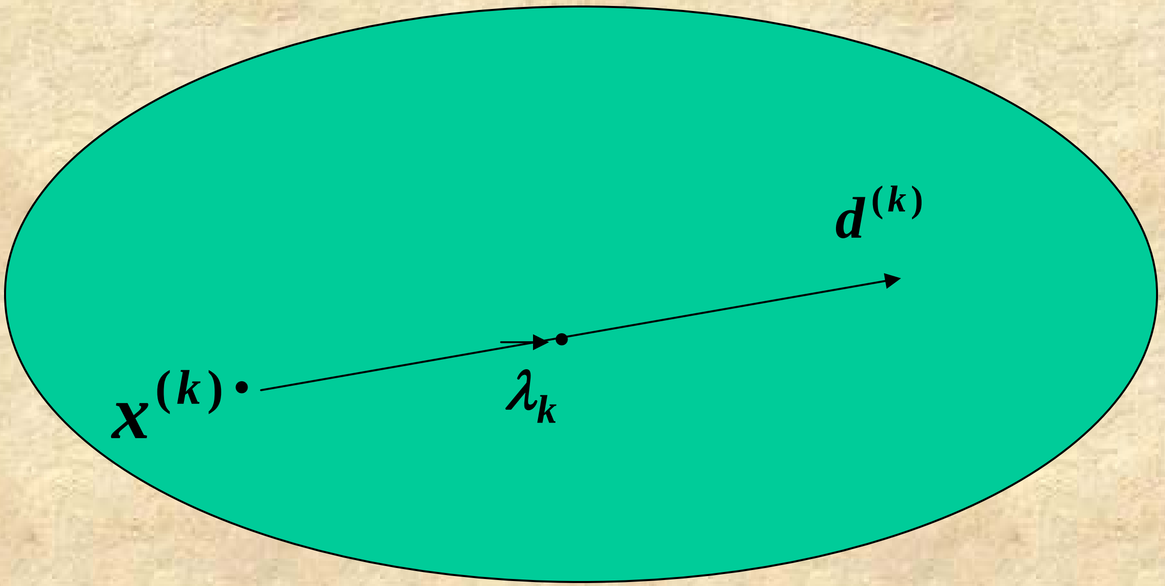
4. 抛物线插值法

5. 三次插值法

一. 一维最优化问题

下降迭代算法中最优步长的确定

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



一维最优化问题: $\min \quad f(x)$

$s.t. \quad x\in R$

极值点的必要条件: $f'(x)=0$

二. 黄金分割法（0.618法）

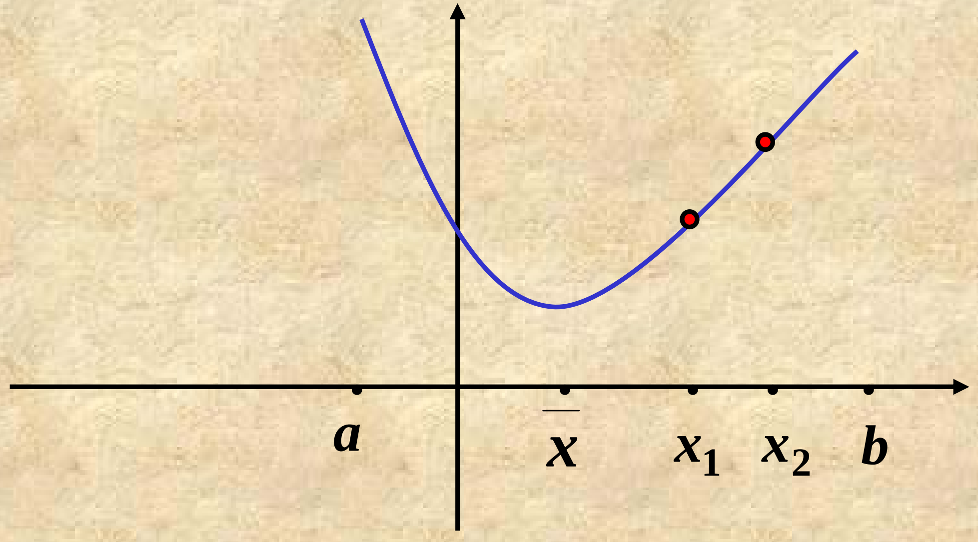
1. 单峰函数

定义：设 $f(x)$ 是区间 $[a,b]$ 上的一元函数，$\bar{x}$ 是 $f(x)$ 在 $[a,b]$ 上的极小点，且对任意的 $x_1,x_2\in[a,b], x_1<x_2$，有

(a) 当 $x_2\le\bar{x}$ 时，$f(x_1)>f(x_2)$;

(b) 当 $x_1\ge\bar{x}$ 时，$f(x_1)<f(x_2)$;

则称 $f(x)$ 是单峰函数。

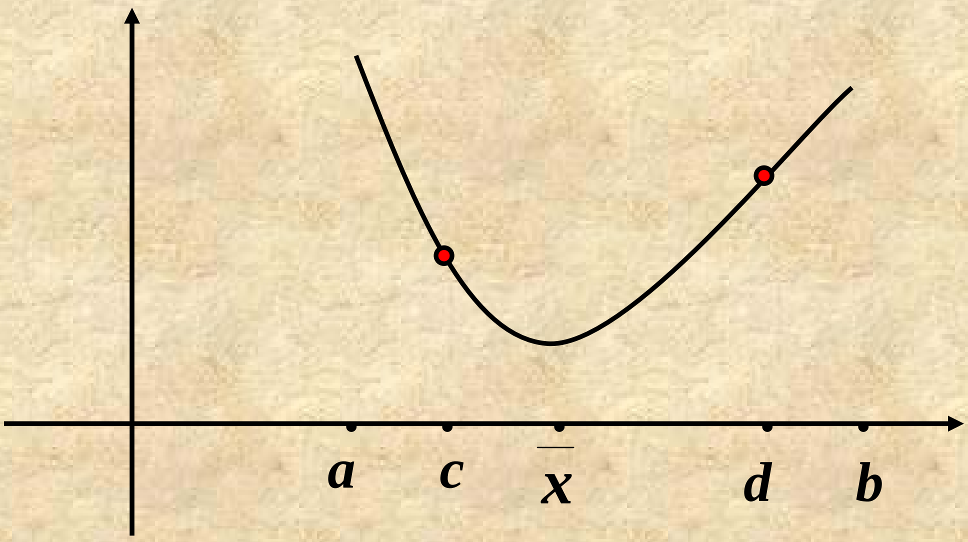


性质：通过计算区间 $[a,b]$内两个不同点的函数值，就可以确定一个包含极小点的子区 间。

定理 设 $f(x)$ 是区间$[a,b]$上的一元函数，$\bar{x}$ 是 $f(x)$ 在$[a,b]$上的极小点。任取点 $c<d\in[a,b]$，则有

(1) 如果 $f(c)>f(d)$，则 $\bar{x}\in[c,b]$；

(2) 如果 $f(c)\le f(d)$，则 $\bar{x}\in[a,d]$；



2. 黄金分割法

思想 通过选取试探点使包含极小点的区间不断缩短，直到区间长度小到一定程度，此时区间上各点的函数值均接近极小值。

下面推导黄金分割法的计算公式。

设 $f(x)$ 在 $[a_1,b_1]$ 上单峰， 极小点 $\bar{x}\in[a_1,b_1]$。假设进行

第 k 次迭代前 $\bar{x}\in[a_k,b_k]$， 取 $\lambda_k,\mu_k\in[a_k,b_k]$，规定 $\lambda_k<\mu_k$。

计算 $f(\lambda_k)$ 和 $f(\mu_k)$，分两种情况：

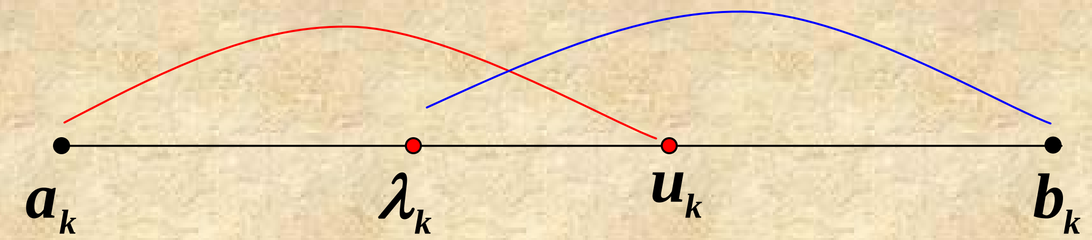
1. 若 $f(\lambda_k)>f(\mu_k)$， 则令 $a_{k+1}=\lambda_k$, $b_{k+1}=b_k$;

2. 若 $f(\lambda_k)\le f(\mu_k)$， 则令 $a_{k+1}=a_k$， $b_{k+1}=\mu_k$.

如何确定 λ_k 与 μ_k?

要求其满足以下两个条　件：

1. $b_k - \lambda_k = \mu_k - a_k$ (1)



2. 每次迭代区间长度缩短　比率相同，即

$$b_{k+1} - a_{k+1} = \alpha(b_k - a_k) \qquad (0 < \alpha < 1) \tag{2}$$

由式（1）与（2）可得：

$$\begin{cases} \lambda_k = a_k + (1-\alpha)(b_k - a_k) & (3) \\ \mu_k = a_k + \alpha(b_k - a_k) & (4) \end{cases}$$

α 取值的确定？

通过确定 α 的取值，使上一次迭代剩余的迭代点恰与下一次迭代的一个迭代点重合，从而减少算法的计算量。

(1) 设在第 k 次迭代时有 $f(\lambda_k) \le f(u_k)$， 则有

$[a_{k+1}, b_{k+1}] = [a_k, u_k]$。

在第 $k+1$ 次迭代时选取 λ_{k+1}, u_{k+1}， 则由（4）有

$$u_{k+1} = a_{k+1} + \alpha(b_{k+1} - a_{k+1})$$

$$= a_k + \alpha^2 (b_k - a_k)$$

如果令 $\alpha^2 = 1 - \alpha$，则 $u_{k+1} = \lambda_k$，因此 u_{k+1} 不必重新计算 。

$$\alpha^2 = 1 - \alpha \Rightarrow \alpha = \frac{\sqrt{5} - 1}{2} \approx 0.618$$

(2) 若在第 k 次迭代时有 $f(\lambda_k) > f(u_k)$。 同理可得。

算法步骤：

1. 给定初始区间 $[a_1, b_1]$, 精度要求 $\varepsilon > 0$.

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

并计算 $f(\lambda_1)$ 与 $f(\mu_1)$. 令 $k := 1$.

2. 若 $b_k - a_k < \varepsilon$, 停止， 且 $\bar{x} = \dfrac{b_k + a_k}{2}$. 否则，

当 $f(\lambda_k) > f(\mu_k)$时，转 3; 当 $f(\lambda_k) \le f(\mu_k)$时，转 4.

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

$\mu_{k+1} = a_{k+1} + 0.618(b_{k+1} - a_{k+1})$, 计算 $f(\mu_{k+1})$,令 $k := k+1$, 转 2。

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

$\lambda_{k+1} = a_{k+1} + 0.382(b_{k+1} - a_{k+1})$, 计算 $f(\lambda_{k+1})$, 令 $k := k+1$, 转 2。

黄金分割法的迭代效果：第k次后迭代后所得区间长度为

初始区间长度的 $(\frac{\sqrt{5}-1}{2})^k$ 倍。

其它的试探点算法：**Fibonacci**算法。

例 1. 给定函数 $f(x)=x^2-2x+3$ ，区间 $[a_1,b_1]=[-2,2]$ ，问：

（1）求第一次迭代后的区间；

（2）至少几次迭代后，所得区间长度会小于0.5？

解:（1）

$$\because \lambda_1=a_1+0.382(b_1-a_1)=-2+0.382\times 4=-0.472$$

$$\mu_1=a_1+0.618(b_1-a_1)=-2+0.618\times 4=0.472$$

$$\therefore f(\lambda_1)\approx 4.17,\quad f(\mu_1)\approx 2.28$$

$$\because f(\lambda_1)>f(\mu_1)$$

$$\therefore [a_2,b_2]=[-0.472,2]$$

(2) 迭代 k 次后区间长度为

$$l_k = (b_1 - a_1) \cdot 0.618^k$$

要使 l_k 小于 0.5，则

$$(b_1 - a_1) \cdot 0.618^k < 0.5$$

解得：

$$k > 4.32$$

所以至少要5次迭代。

三. 进退法

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。

进退法的计算步骤

1. 给定初始点 $x^{(0)} \in R$, 初始步长 $h_0 > 0$, 令 $h = h_0$, $x^{(1)} = x^{(0)}$, 计算 $f(x^{(1)})$, 并令 $k = 0$.

2. 令 $x^{(4)} = x^{(1)} + h$, 计算 $f(x^{(4)})$, 令 $k := k + 1$.

3. 若 $f(x^{(4)}) < f(x^{(1)})$, 则转 4, 否则，转 5.

4. 令 $x^{(2)} = x^{(1)}$, $x^{(1)} = x^{(4)}$, $f(x^{(2)}) = f(x^{(1)})$, $f(x^{(1)}) = f(x^{(4)})$, 令 $h := 2h$, 转 2.

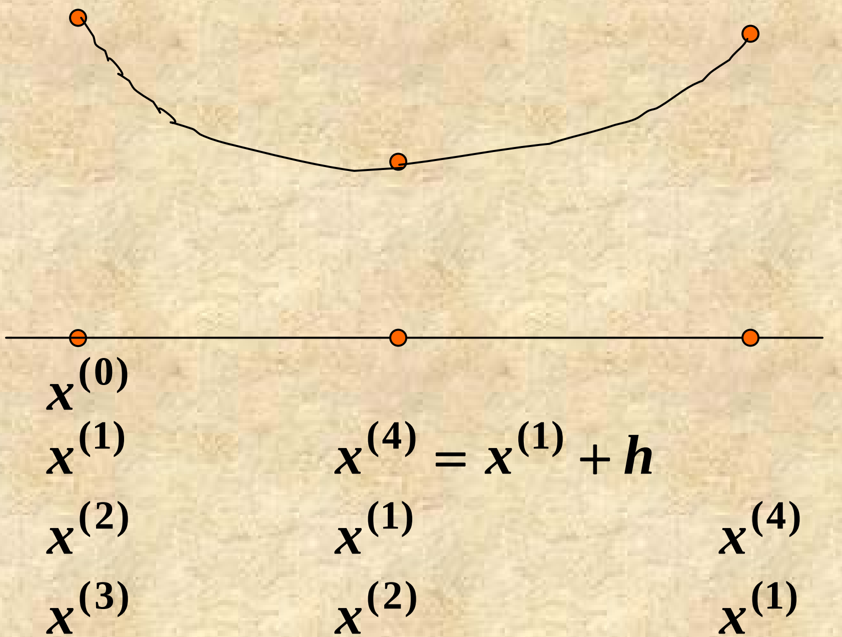
5. 若 $k = 1$, 否则，转 7.

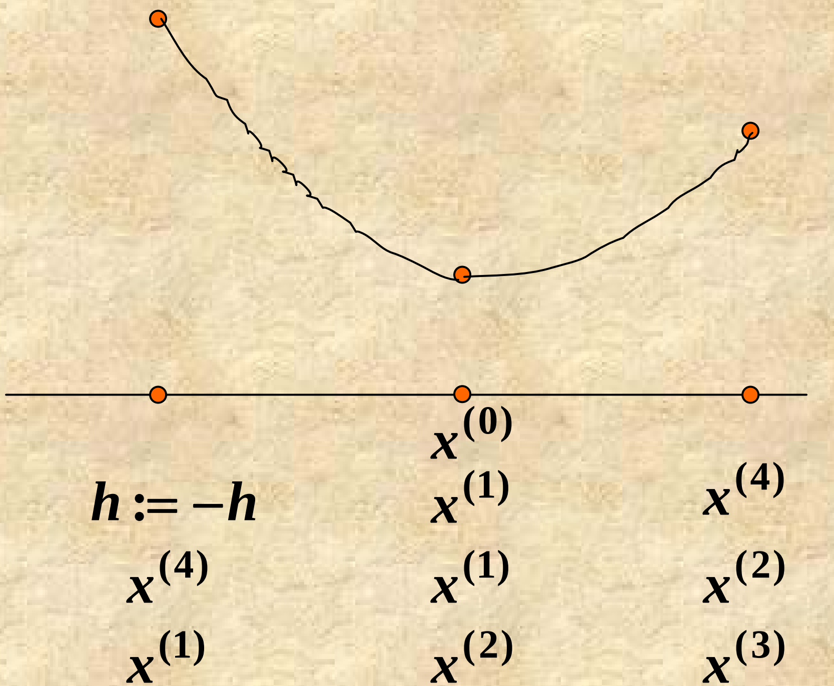
6. 令 $h := -h$, $x^{(2)} = x^{(4)}$, $f(x^{(2)}) = f(x^{(4)})$, 转 2.

7. 令 $x^{(3)} = x^{(2)}$, $x^{(2)} = x^{(1)}$, $x^{(1)} = x^{(4)}$, 停止计算。

区间 $[x^{(1)}, x^{(3)}]$ 或 $[x^{(3)}, x^{(1)}]$ 即为包含极小点的区间 。

例:





极小点区间: $[x^{(3)}, x^{(1)}]$

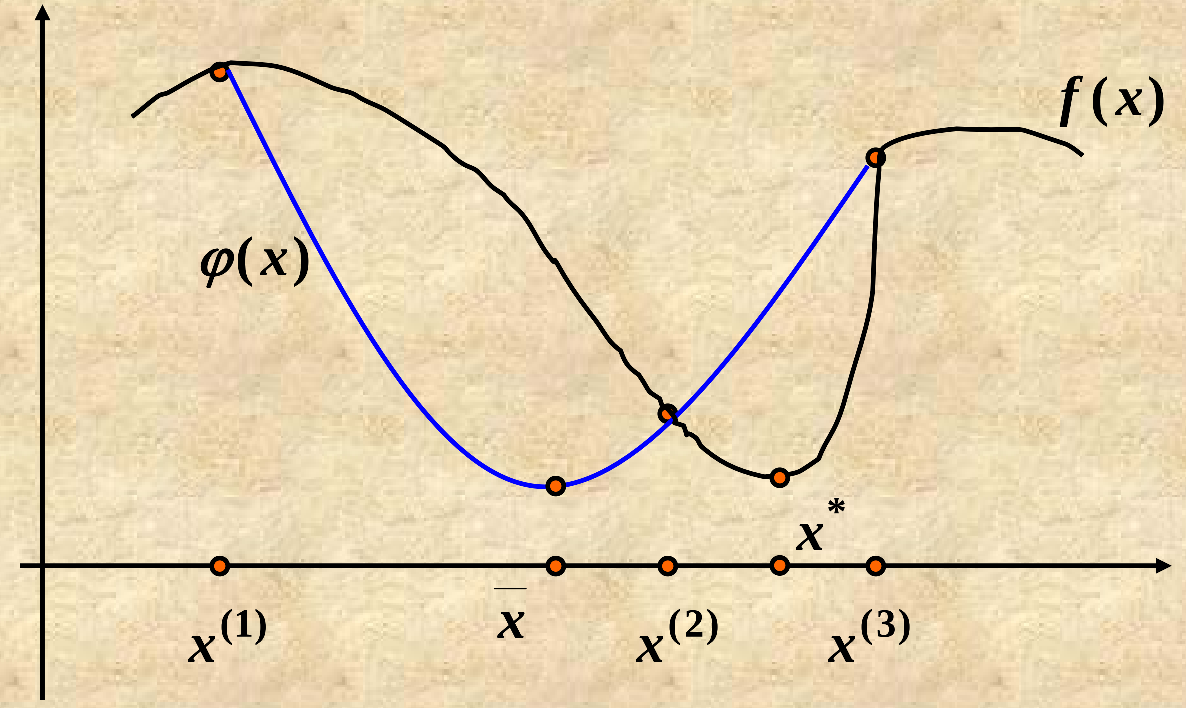
极小点区间: $[x^{(1)}, x^{(3)}]$

四. 抛物线插值法

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

令

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

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

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

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 $\bar{x}=-\dfrac{b}{2c}$

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

以 $\bar{x}$ 作为 $f(x)$ 的极小点的估计值 。

抛物线插值算法步骤：

(1) 给定初始区间 $[x^{(1)}, x^{(3)}]$，设 $f(x^{(1)}) > f(x^{(2)})$， $f(x^{(2)}) < f(x^{(3)})$.

令 $k := 1$。$\overline{x}^{(0)} = x^{(1)}$。给定精度 ε。

(2) 设 $\varphi(x) = a + bx + cx^2$。令 $\varphi(x^{(i)}) = f(x^{(i)}), i = 1, 2, 3$。解出

系数 a, b, c。解出 $\varphi(x)$ 的极小点 $\overline{x}^{(k)} = -\dfrac{b}{2c}$。

若 $|f(\overline{x}^{(k)}) - f(\overline{x}^{(k-1)})| < \varepsilon$，则算法停止，否则转（3）。

(3) 从 $x^{(1)}, x^{(2)}, x^{(3)}$ 和 $\overline{x}^{(k)}$ 中选择 $f(x)$ 函数值最小的一个点 及其

左、右两点，重新标记为 $x^{(1)}, x^{(2)}, x^{(3)}$。令 $k := k + 1$。转（2）。

例 2. 给定如下一维搜索问题

$$\min_{x\in[0,2\pi]} \quad f(x)=\sin x$$

设初始区间为 $[x^{(1)},x^{(3)}]=[\frac{\pi}{2},2\pi]$，而 $x^{(2)}=\frac{7\pi}{6}$.

试用抛物线插值算法求解次问题。计算一次迭代后的近似极小点 $\bar{x}^{(1)}$。

解: 设 $\varphi(x)=a+bx+cx^2$.

令
$$\begin{cases}\varphi(\frac{\pi}{2})=a+\frac{\pi}{2}b+\frac{\pi^2}{4}c=\sin\frac{\pi}{2}\\ \varphi(\frac{7\pi}{6})=a+\frac{7\pi}{6}b+\frac{49\pi^2}{36}c=\sin\frac{7\pi}{6}\\ \varphi(2\pi)=a+2\pi b+4\pi^2c=\sin 2\pi\end{cases}$$

解方程组可得

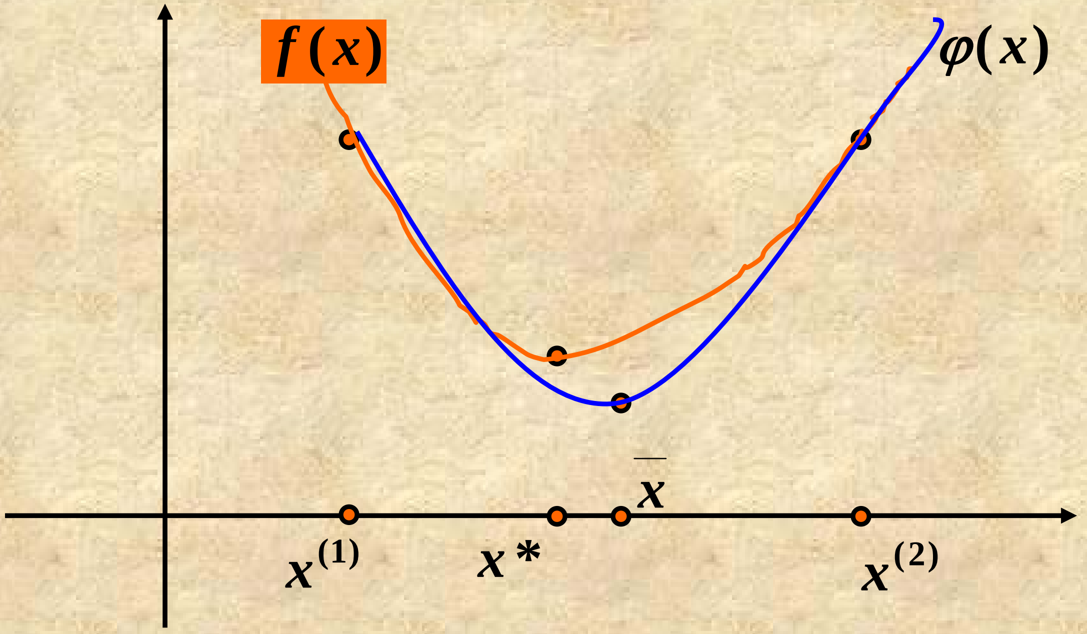
$$b=-\frac{65}{12\pi},c=\frac{19}{10\pi^2}.$$

所以有

$$\overline{x}^{(1)}=-\frac{b}{2c}=\frac{325\pi}{228}$$

五. 三次插值法

思想： 已知 $f(x)$ 在点 $x^{(1)}$、$x^{(2)}$（$x^{(1)} < x^{(2)}$）的函数值，且 $f'(x^{(1)}) < 0$，$f'(x^{(2)}) > 0$。则区间（$x^{(1)}, x^{(2)}$）内存在 $f(x)$ 的极小点。利用这两点的函数值和导数构造一个 三次多项式 $\varphi(x)$，使它与 $f(x)$ 在这两点有相同的函数值和导数，用 $\varphi(x)$ 逼近 $f(x)$，并用 $\varphi(x)$ 的极小点近似 $f(x)$ 的极小点。



确定函数 $\varphi(x)$的方法:

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

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

则有

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

求解方程组得出系数 a,b,c,d的值,从而确定函数 $\varphi(x)$。

求解满足 $\begin{cases}\varphi'(x)=0\\ \varphi''(x)>0\end{cases}$ 的极小点 $\bar{x}$ 。

令 $\varphi'(x)=3a(x-x^{(1)})^2+2b(x-x^{(1)})+c=0 \qquad (3)$

而 $\varphi''(x)=6a(x-x^{(1)})+2b$ 。

解方程（3），有两种情况:

1. 当$a=0$时,解得

$$\bar{x}-x^{(1)}=-\frac{c}{2b} \qquad (4)$$

由（2）可知

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

所以 $b>0$ 。

因此 $\varphi''(\bar{x})=2b>0$,即 $\bar{x}$是$\varphi(x)$的极小点。

2. 当$a \neq 0$时，解得

$$\bar{x} - x^{(1)} = \frac{-b \pm \sqrt{b^2 - 3ac}}{3a} \qquad (5)$$

此时有

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

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

为使 $\varphi''(\bar{x}) > 0$，则在(5)式中应取

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

$$= \frac{-c}{b + \sqrt{b^2 - 3ac}} \qquad (6)$$

而当$a = 0$时，由(6)式可知 $\bar{x} - x^{(1)} = -\dfrac{c}{2b}$，所以 (6)式

是1、2两种情况下极小点的统一表达式。

极小点的计算公式:

令 $$s=\frac{3[f(x^{(2)})-f(x^{(1)})]}{x^{(2)}-x^{(1)}} \qquad (7)$$

$$t=s-f'(x^{(1)})-f'(x^{(2)}) \qquad (8)$$

$$w^2=t^2-f'(x^{(1)})f'(x^{(2)}) \qquad (9)$$

则有 $$\bar{x}=x^{(1)}+(x^{(2)}-x^{(1)})\left(1-\frac{f'(x^{(2)})+t+w}{f'(x^{(2)})-f'(x^{(1)})+2w}\right) \qquad (10)$$

算法步骤:

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$\delta > 0, \varepsilon > 0$。

2.按照公式(7)、(8)、(9)、(10)计算s、t、w和$\bar{x}$。

3.若$|x^{(2)} - x^{(1)}| < \delta$,则停止计算，得到点$\bar{x}$;否则，转4。

4.计算$f(\bar{x}), f'(\bar{x})$。若$|f'(\bar{x})| < \varepsilon$,停止计算，得到点$\bar{x}$。

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

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

例　用三次插值算法求解下 述问题：

$$\min \quad f(x)=x^3-3x+1$$
$$s.t. \quad 0\le x\le 2$$

解：取初始点 $x^{(1)}=0, x^{(2)}=2$。则有

$$x^{(2)}-x^{(1)}=2,\quad f(x^{(1)})=1,\quad f(x^{(2)})=3,$$

由 $f'(x)=3x^2-3$,得 $f'(x^{(1)})=-3,\quad f'(x^{(2)})=9$。

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

$$t=s-f'(x^{(1)})-f'(x^{(2)})=-3,$$

$$w^2=t^2-f'(x^{(1)})f'(x^{(2)})=6 \quad \Rightarrow w=\sqrt{6}。$$

则 $\bar{x}=0+2\times(1-\dfrac{9-3+\sqrt{6}}{9+3+2\sqrt{6}})=1$。

而 $f'(\bar{x})=0$，所以 $\bar{x}=1$是极小点。

收敛速度：三次插值算法的收敛速度为 2 阶。