日期	科目	班级	姓名	学号
2022 年 4 月 16 日	最优化方法	强基数学 002	吴天阳	2204210460

第三章

4. 设 $\varphi(t)=e^{-t}+e^{t}$，区间为 $[-1,1]$。

(1) 用 0.618 法极小化 $\varphi(t)$。（计算两步）

(2) 用 Fibonacci 法极小化 $\varphi(t)$。（计算两步）

解答. (1) 初始化：

$$a_1=-1,\ b_1=1,\ \lambda_1=a_1+\frac{3-\sqrt{5}}{2}(b_1-a_1)=2-\sqrt{5},\ \mu_1=a_1+\frac{\sqrt{5}-1}{2}(b_1-a_1)=\sqrt{5}-2$$

第一步：由于 $\varphi(t)$ 为偶函数，所以 $\varphi(\lambda_1)=\varphi(\mu_1)$，于是

$$a_2=a_1=-1,\ b_2=\mu_1=\sqrt{5}-2,\ \mu_2=\lambda_1=2-\sqrt{5},\lambda_2=a_2+\frac{3-\sqrt{5}}{2}(b_2-a_2)=2\sqrt{5}-5$$

第二步：由于 $2.285\approx\varphi(\lambda_2)>\varphi(\mu_2)\approx 2.056$，所以极小化值为 $\varphi(\mu_2)\approx 2.056$。

(2) 初始化：取 $n=4$，则 $a_1=-1,\ b_1=1,\ \lambda_1=a_1+\frac{1}{3}(b_1-a_1)=-\frac{1}{3},\ \mu_1=a_1+\frac{2}{3}(b_1-a_1)=\frac{1}{3}$

第一步：由于 $\varphi(\lambda_1)=\varphi(\mu_1)$，于是

$$a_2=a_1=-1,\ b_2=\mu_1=\frac{1}{3},\ \mu_2=\lambda_1=-\frac{1}{3},\ \lambda_2=a_2+\frac{1}{2}(b_2-a_2)=-\frac{1}{3}$$

第二步：由于 $\mu_2=\lambda_2$，于是极小化值为 $\varphi(\mu_2)\approx 2.112$。

6. 设 $\varphi(t)=1-te^{-t^2}$，区间为 $[0,1]$。试用三点二次插值法极小化 $\varphi(t)$。（计算一步）

解答. 设 $x_1=0,\ x_2=0.5,\ x_3=1$，下面列出三阶差商表

x_i	$\varphi[x_i]$	$\varphi[x_i,x_{i+1}]$	$\varphi[x_i,x_{i+1},x_{i+2}]$
x_1	1		
x_2	$1-\frac{1}{2}e^{-\frac{1}{4}}$	$-e^{-\frac{1}{4}}$	
x_3	$1-e^{-1}$	$e^{-\frac{1}{4}}-2e^{-1}$	$2e^{-\frac{1}{4}}-2e^{-1}$

由 Newton 插值多项式可知

$$N_2(x)=1-e^{-\frac{1}{4}}x+(2e^{-\frac{1}{4}}-2e^{-1})x(x-\frac{1}{2})=(2e^{-\frac{1}{4}}-2e^{-1})x^2+(-2e^{-\frac{1}{4}}+e^{-1})x+1$$

则 $\overline{x}=\dfrac{2e^{-\frac{1}{4}}-e^{-1}}{4(e^{-\frac{1}{4}}-e^{-1})}\approx 0.7238>x_2$，又由于 $0.61\approx\varphi(x_2)>\overline{\varphi}=\varphi(\overline{x})\approx 0.57$，所以令 $x_1\Leftarrow\overline{x}$，$\varphi(x_1)\Leftarrow\overline{\varphi}$，极小化值为 $\overline{\varphi}\approx 0.57$。

第四章

1. 设 $f(x)=\dfrac{3}{2}x_1^2+\dfrac{1}{2}x_2^2-x_1x_2-2x_1$，设初始点 $x^{(0)}=(-2,4)^T$。试用最速下降法和牛顿法极小化 $f(x)$。（计算两步）

解答. 最速下降法：已知 $\nabla f=[3x_1-x_2-2,x_2-x_1]^T$，$x^{(0)}=(-2,4)^T$。

第一步：$d_0=-g_0=-\nabla f(x^{(0)})=[12,-6]^T$，求线性步长因子

$$\nabla f(x^{(0)}+\alpha_0 d_0)^T d_0=\nabla f(-2+12\alpha_0,4-6\alpha_0)^T\begin{bmatrix}12\\-6\end{bmatrix}=-5+16\alpha_0=0$$

则 $\alpha_0=\dfrac{5}{17}$，于是 $x^{(1)}=x^{(0)}+\alpha_0 d_0=(\dfrac{26}{17},\dfrac{38}{17})^T$。

第二步：$d_1=-g_1=-\nabla f(x^{(1)})=-\left[\frac{6}{17},\frac{12}{17}\right]^T$，求线性步长因子

$$\nabla f(x^{(1)}+\alpha_1 d_1)^T d_1=\nabla f\left(\frac{26}{17}-\frac{6}{17}\alpha_1,\frac{38}{17}-\frac{12}{17}\alpha_1\right)^T\begin{bmatrix}-\frac{6}{17}\\-\frac{12}{17}\end{bmatrix}=\left(\frac{6}{17}\right)^2(3\alpha_1-5)=0$$

则 $\alpha_1=\dfrac{5}{3}$，于是 $x^{(2)}=x^{(1)}+\alpha_1 d_1=\left(\dfrac{16}{17},\dfrac{18}{17}\right)^T$，所以通过最速下降法两步得到的极小化值为

$$f(x^{(2)})=\frac{3}{2}\cdot\left(\frac{16}{17}\right)^2+\frac{1}{2}\left(\frac{18}{17}\right)^2-\frac{16}{17}\cdot\frac{18}{17}-2\cdot\frac{16}{17}=-\frac{286}{289}\approx-0.9896$$

牛顿法：已知 $\nabla^2 f(x)=\begin{bmatrix}3&-1\\-1&1\end{bmatrix}$，$x^{(0)}=(-2,4)^T$，则

$$G_0d_0=-g_0\Rightarrow\begin{bmatrix}-3&-1\\-1&1\end{bmatrix}d_0=\begin{bmatrix}12\\-6\end{bmatrix}\Rightarrow d_0=\begin{bmatrix}3\\-3\end{bmatrix}$$

求线性步长因子

$$\nabla f(x^{(0)}+\alpha_0 d_0)^T d_0=\nabla f(-2+3\alpha_0,4-3\alpha_0)^T\begin{bmatrix}3\\-3\end{bmatrix}=54(\alpha_0-1)=0$$

则 $\alpha_0=1$，于是 $x^{(1)}=x^{(0)}+\alpha_0 d_0=(1,1)^T$，由于 $g_1=\nabla f(x^{(1)})=\mathbf{0}$，所以 $f(x)$ 的极小化值为 $f(x^{(1)})=-1$。

定理 4.3.3. 设 $x_0\in\mathbb{R}^n$ 是任意初始点。对于极小化二次函数 $\dfrac{1}{2}x^TGx-b^Tx$，共轭方向至多经 n 步精确线性搜索终止；且每一 x_{i+1} 都是 $f(x)$ 在 x_0 和方向 $d_0,\cdots,d_i$ 所张成的线性流形 $\{x:x=x_0+\sum\limits_{j=0}^{i}\alpha_jd_j,\ \forall\alpha_j\}$ 中的极小点。

证明. 设 $U=\{x: x=x_0+\sum_{j=0}^{i}\alpha_j d_j,\ \forall\alpha_j\}$，则 $x_{i+1}=\min_{x\in U} f(x)$ 等价于

$$\nabla f(x_{i+1})=0 \iff [g_{i+1}^T d_1, g_{i+1}^T d_2, \cdots, g_{i+1}^T d_i]=0$$

所以，只需证明 $\forall 1\leqslant i\leqslant n$，有 $g_i^T d_j=0\ (\forall 0\leqslant j<i)$ 成立，由于

$$g_{k+1}-g_k=G(x_{k+1}-x_k)=\alpha G d_k \text{ 且 } g_{k+1}^T d_k=0$$

则，任意的 $j<i$ 有

$$\begin{aligned} g_i^T d_j &= (g_{j+1}+(g_i-g_{i-1}+g_{i-1}-g_{i-2}+\cdots+g_{j+2}-g_{j+1}))^T d_j \\ &= \left(g_{j+1}+\sum_{k=j+1}^{i-1}\alpha_k d_j^T G d_k\right)^T d_j = g_{j+1}d_j+\sum_{k=j+1}^{i-1}\alpha_k d_j^T G d_k \\ &= 0 \quad (\text{第一式由于精确线性搜索为 } 0\text{，第二式由于共轭性为 } 0) \end{aligned}$$

由于 $\mathbb{R}^n$ 空间可以视为由 n 个共轭向量所张成的, 当 $i=n$ 时, 有 $g_n^T d_j=0\ (j=0,1,\cdots,n-1)$, 所以 $g_n\equiv 0$，所以共轭方向法至多经 n 步精确线性搜索终止。 □

4. 证明: 当极小化正定二次函数时, 共轭梯度法 FR 公式, PRP 公式和 Dixon 公式是等价的。

证明. 设正定二次函数为 $f(x)=\frac{1}{2}x^T G x-b^T x$，第 k 步的搜索方向为 $d_k=-g_k+\beta_{k-1}d_{k-1}$，由精确线性搜索可知 $g_{k+1}^T d_k=0$，则 $g_{k+1}^T(-g_k+\beta_{k-1}d_{k-1})=0$，通过**定理 4.3.3** 可知 $g_{k+1}^T g_k=\beta_{k-1}g_{k+1}^T d_{k-1}=0$，于是

$$\begin{aligned} (\text{Dixon 公式}) \quad \beta_{k-1} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}} &= -\frac{g_k^T g_k}{(-g_{k-1}+\beta_{k-2}d_{k-2})^T g_{k-1}} \xlongequal{\text{精确线性搜索条件}} \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \quad (\text{FR 公式}) \\ &= \frac{g_k^T g_k+g_k^T g_{k-1}}{g_{k-1}^T g_{k-1}} = \frac{g_k^T(g_k-g_{k-1})}{g_{k-1}^T g_{k-1}} \quad (\text{PRP 公式}) \end{aligned}$$

□