文章编号:1000-0887(2014)06-0610-10

解信赖域子问题的隐式分段折线算法*

王希云， 李 亮， 于海波

(太原科技大学 应用科学学院，太原 030024)

摘要： 在 Hessian 矩阵正定的前提下,建立了一种最优曲线的微分方程模型.针对此微分方程模型,构造了一条隐式分段折线,从而提出了一种求解信赖域子问题的隐式分段折线算法,并且分析和证明了隐式分段折线路径的合理性.数值结果表明新算法是有效且可行的.

关 键 词： 隐式分段折线算法； 微分方程模型； 信赖域子问题

中图分类号： O221 **文献标志码：** A

doi：10.3879/j.issn.1000-0887.2014.06.003

引 言

在非线性优化中,信赖域方法是一种被广泛应用的方法.它不仅具有可靠性和强适性的优点,而且还具有较强的收敛性.因而,自从其出现之日起就受到非线性优化研究界的广泛重视,成为一个研究热点.目前,信赖域方法已经和线搜索方法并列为求解无约束最优化问题(1)的两大类主要的数值方法.

无约束最优化问题如下：

$$\min f(\boldsymbol{x}), \qquad \boldsymbol{x} \in R^n, \tag{1}$$

其中 $f(\boldsymbol{x}):R^n \to \mathbf{R}$ 是目标函数,$\boldsymbol{x} \in R^n$ 是待求变量.

当用信赖域方法求解无约束优化问题(1)时,关键在于每步迭代时都需求解下面形式的信赖域子问题：

$$\min q(\boldsymbol{\delta}) = \boldsymbol{g}^{\mathrm{T}}\boldsymbol{\delta} + \frac{1}{2}\boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\delta}, \qquad \text{s.t. } \|\boldsymbol{\delta}\|_2 \leqslant \Delta, \tag{2}$$

其中 $\boldsymbol{g} \in R^n$ 为目标函数在当前迭代点的梯度,$\boldsymbol{B} \in R^{n\times n}$ 为目标函数在当前迭代点的 Hessian 矩阵或其近似,$\Delta \in \mathbf{R}$ 为信赖域半径,$\boldsymbol{\delta} \in R^n$ 为待求变量.当 Δ 变化时,上述信赖域子问题(2)的解 $\boldsymbol{\delta}^*$ 形成一条空间曲线,称为最优曲线[1].

对于如何求解信赖域子问题(2),目前已提出很多方法.比如在文献[2]中,后六生、孙文瑜提出了三项预处理共轭梯度法来求解信赖域子问题.赵英良、徐成贤在文献[3]中提出了使用重新开始策略的共轭梯度法求解信赖域子问题.在文献[4]中,陈争、马昌凤提出了一种求解

* **收稿日期：** 2013-11-21；**修订日期：** 2014-05-07
基金项目： 山西省自然科学基金(2008011013);山西省“131”领军人才工程项目
作者简介： 王希云(1964—),女,山西临汾人,教授(E-mail：tykdwxy@126.com)；
李亮(1988—),男,河南驻马店人,硕士生(通讯作者. E-mail：tyustliliang@126.com)；
于海波(1990—),男,山西晋城人,硕士生(E-mail：tyustyuhaibo@126.com).

信赖域子问题光滑 Newton(牛顿)法.关于求解信赖域子问题的折线法的文章可参看文献[5-10].而在 Hessian 矩阵正定的情况下,目前常用的折线法主要有单折线法、双折线法和切线单折线法[8-10].文献[10]的数值实验表明,切线单折线比单折线、双折线更近似于最优曲线,并且比单折线法、双折线法求解信赖域子问题(2)更有效.

受文献[8-10]的启发,本文根据最优曲线的参数方程首先构造出了一个微分方程模型,利用文献[11]中的方法构造一条折线,称为隐式分段折线.进而提出了一种解信赖域子问题(2)的隐式分段折线算法.并且分析和证明了隐式分段折线路径的合理性.通过测试函数进行验证表明,新算法是有效可行的.

1　构造微分方程模型

定理 1.1[12]　$\boldsymbol{\delta}^*$ 是信赖域子问题(2)的解,当且仅当存在$\mu^* \geqslant 0$时,使得式(3)成立,且$(\boldsymbol{B}+\mu^*\boldsymbol{I})$是半正定矩阵.

$$\begin{cases}(\boldsymbol{B}+\mu^*\boldsymbol{I})\boldsymbol{\delta}^* = -\boldsymbol{g},\\ \mu^*(\Delta - \|\boldsymbol{\delta}^*\|_2) = 0,\\ \|\boldsymbol{\delta}^*\|_2 \leqslant \Delta.\end{cases} \tag{3}$$

由定理 1.1 可知,信赖域子问题(2)的精确求解方法的思想即求解下面的方程组:

$$\begin{cases}(\boldsymbol{B}+\mu\boldsymbol{I})\boldsymbol{\delta} = -\boldsymbol{g},\\ \mu(\Delta - \|\boldsymbol{\delta}\|_2) = 0,\\ \|\boldsymbol{\delta}\|_2 \leqslant \Delta.\end{cases} \tag{4}$$

所以,当$\mu = 0$且$\|\boldsymbol{B}^{-1}\boldsymbol{g}\|_2 \leqslant \Delta$时,信赖域子问题(2)的解$\boldsymbol{\delta}^* = -\boldsymbol{B}^{-1}\boldsymbol{g}$;当$\mu > 0$时,则是通过求解一元非线性方程$\|(\boldsymbol{B}+\mu\boldsymbol{I})^{-1}\boldsymbol{g}\|_2 - \Delta = 0$得到解$\mu^*$,然后把$\mu^*$代入方程组(4)的第1个方程,则可以求出信赖域子问题(2)的解$\boldsymbol{\delta}^* = -(\boldsymbol{B}+\mu^*\boldsymbol{I})^{-1}\boldsymbol{g}$.

由信赖域子问题的精确求解方法的思想,得到最优曲线的参数方程如下:

$$\boldsymbol{\delta} = -(\boldsymbol{B}+\mu\boldsymbol{I})^{-1}\boldsymbol{g} \qquad (\mu \geqslant 0). \tag{5}$$

对于参数方程(5)求导,可得

$$\boldsymbol{B}\boldsymbol{\delta}' + \boldsymbol{\delta} + \mu\boldsymbol{\delta}' = \boldsymbol{0},$$

即

$$(\boldsymbol{B}+\mu\boldsymbol{I})\boldsymbol{\delta}' = -\boldsymbol{\delta}. \tag{6}$$

由式(6)可得最优曲线上任意一点的切线斜率如下:

$$\boldsymbol{\delta}' = -(\boldsymbol{B}+\mu\boldsymbol{I})^{-1}\boldsymbol{\delta} \qquad (\mu \geqslant 0).$$

从而构造的微分方程模型如下:

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{\delta}}{\mathrm{d}\mu} = -(\boldsymbol{B}+\mu\boldsymbol{I})^{-1}\boldsymbol{\delta}\\ \boldsymbol{\delta}(0) = -\boldsymbol{B}^{-1}\boldsymbol{g}\end{cases} \qquad (\mu \geqslant 0). \tag{7}$$

2　构造隐式分段折线

由文献[11]可得求解微分方程(7)的公式如下:

$$\boldsymbol{\delta}_{n+1} = \boldsymbol{\delta}_n - h_n(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1} \qquad (n = 0,1,2,\cdots), \tag{8}$$

其中$\tilde{\boldsymbol{\delta}}_{n+1}$由式(9)确定.

$$\tilde{\boldsymbol{\delta}}_{n+1}=\boldsymbol{\delta}_n-h'_n(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n. \tag{9}$$

从初始点 $P_0(\gamma_0,\boldsymbol{\delta}_0)$ 开始,其中 $\gamma_0=\mu_0=0,\boldsymbol{\delta}_0=-\boldsymbol{B}^{-1}\boldsymbol{g}$.利用式(8)和(9)分别求出点 $P_1(\gamma_1,\boldsymbol{\delta}_1),P_2(\gamma_2,\boldsymbol{\delta}_2),\cdots,P_{k-1}(\gamma_{k-1},\boldsymbol{\delta}_{k-1}),\cdots,P_N(\gamma_N,\boldsymbol{\delta}_N)$,其中 N 是满足 $\|\boldsymbol{\delta}_N\|_2\leqslant\Delta$ 的最小整数,$\gamma_i=\gamma_{i-1}+h_{i-1},\mu_i=\mu_{i-1}+h'_{i-1},h_{i-1}\leqslant h'_{i-1},i=1,2,\cdots,N$.

然后分别连接点 $P_0,P_1,\cdots,P_k,\cdots,P_N$,则可以得到隐式分段折线 $P_0P_1\cdots P_k\cdots P_N$,记 $P_0P_1\cdots P_k\cdots P_N$ 为 $\Psi=[P_0,P_1,\cdots,P_N]$.

3 步长 h'_n 和 h_n

为了使构造的隐式分段折线满足文献[13]中的引理6.4.1,则步长 h'_n 和 h_n 需满足如下公式(10)~(13):

$$h'_0=\min\left\{\frac{\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\varepsilon\boldsymbol{I})^{-1}\boldsymbol{\delta}_0}{\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{B}^{-2}\boldsymbol{\delta}_0},\varepsilon\right\}, \tag{10}$$

$$h'_n=\begin{cases}\min\left\{\dfrac{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+(n+1)\varepsilon\boldsymbol{I})^{-1}\boldsymbol{\delta}_n}{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-2}\boldsymbol{\delta}_n},\varepsilon\right\}, & \boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n\leqslant 0,\ n=1,2,3,\cdots,\\ \min\left\{\dfrac{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+(n+1)\varepsilon\boldsymbol{I})^{-1}\boldsymbol{\delta}_n}{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-2}\boldsymbol{\delta}_n},\varepsilon,\dfrac{\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_n-\boldsymbol{\delta}_n^{\mathrm{T}}\boldsymbol{\delta}_n}{\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n}\right\}, & \boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n>0,\ n=1,2,3,\cdots,\end{cases} \tag{11}$$

$$h_0=\min\left\{h'_0,\frac{\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1}{2\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_1}\right\}, \tag{12}$$

$$h_n=\begin{cases}\min\left\{h'_n,\dfrac{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}}{\tilde{\boldsymbol{\delta}}_{n+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{n+1}}\right\}, & \boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}\leqslant 0,\ n=1,2,3,\cdots,\\ \min\left\{h'_n,\dfrac{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}}{\tilde{\boldsymbol{\delta}}_{n+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{n+1}},\dfrac{\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_n-\boldsymbol{\delta}_n^{\mathrm{T}}\boldsymbol{\delta}_n}{\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}}\right\}, & \boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}>0,\ n=1,2,3,\cdots,\end{cases} \tag{13}$$

其中 ε 称为限制步长,即限制每次迭代的步长最大值只能达到 ε.

4 隐式分段折线路径的性质

引理 4.1 设 $\boldsymbol{B}$ 对称正定,记隐式分段折线 $\Psi=[P_0,P_1,\cdots,P_N]$ 为 $\boldsymbol{\delta}(\tau)$,如下所示:

$$\boldsymbol{\delta}(\tau)=\begin{cases}\boldsymbol{\delta}_0-\tau(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1, & \tau\in[\alpha_0,\alpha_1],\\ \boldsymbol{\delta}_1-(\tau-\alpha_1)(\boldsymbol{B}+\mu_2\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_2, & \tau\in(\alpha_1,\alpha_2],\\ \boldsymbol{\delta}_2-(\tau-\alpha_2)(\boldsymbol{B}+\mu_3\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_3, & \tau\in(\alpha_2,\alpha_3],\\ \vdots & \\ \boldsymbol{\delta}_{N-1}-(\tau-\alpha_{N-1})(\boldsymbol{B}+\mu_N\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_N, & \tau\in(\alpha_{N-1},\alpha_N],\end{cases}$$

其中 $\tilde{\boldsymbol{\delta}}_{n+1}=\boldsymbol{\delta}_n-h'_n(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n,(n=0,1,\cdots,N-1),\alpha_0=0,\ \alpha_i=\alpha_{i-1}+h_{i-1},\ i=1,2,\cdots,N$.则当 $n=0,1,\cdots,N-1$ 时,有

(i) $\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}\geqslant 0$;

(ii) $\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_n-\boldsymbol{\delta}_n^{\mathrm{T}}\boldsymbol{\delta}_n\geqslant 0$.

证明 (i) 当 $n=0,1,\cdots,N-1$ 时,因为

$$\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}=\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}[\boldsymbol{\delta}_n-h'_n(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n]=$$
$$\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\boldsymbol{\delta}_n-h'_n\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n,$$

又由式(10)和(11)可得

$$h'_n\leqslant\frac{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+(n+1)\varepsilon\boldsymbol{I})^{-1}\boldsymbol{\delta}_n}{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-2}\boldsymbol{\delta}_n}\leqslant\frac{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\boldsymbol{\delta}_n}{\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n},$$

所以

$$\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}\geqslant 0.$$

(ii) 当 $n=0$ 时结论显然成立.

当 $n=1$ 时,因为

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_1-\boldsymbol{\delta}_1^{\mathrm{T}}\boldsymbol{\delta}_1=\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{\delta}_0-\tau(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1)-$$
$$(\boldsymbol{\delta}_0-\tau(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1)^{\mathrm{T}}(\boldsymbol{\delta}_0-\tau(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1)=$$
$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_0-\tau\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1-\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_0+2\tau\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1-\tau^2\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_1=$$
$$-\tau^2\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_1+\tau\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1,$$

又由式(12)可得

$$h_0\leqslant\frac{\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1}{2\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_1},$$

所以

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_1-\boldsymbol{\delta}_1^{\mathrm{T}}\boldsymbol{\delta}_1\geqslant 0.$$

即当 $n=1$ 时结论成立.

假设当 $1<n\leqslant k$ 时,不等式组 $\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_n-\boldsymbol{\delta}_n^{\mathrm{T}}\boldsymbol{\delta}_n\geqslant 0$ 都成立.则当 $n=k+1$ 时

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_{k+1}-\boldsymbol{\delta}_{k+1}^{\mathrm{T}}\boldsymbol{\delta}_{k+1}=\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{\delta}_k-(\tau-\alpha_k)(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1})-$$
$$(\boldsymbol{\delta}_k-(\tau-\alpha_k)(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1})^{\mathrm{T}}(\boldsymbol{\delta}_k-(\tau-\alpha_k)(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1})=$$
$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_k-\boldsymbol{\delta}_k^{\mathrm{T}}\boldsymbol{\delta}_k-(\tau-\alpha_k)\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}+$$
$$2(\tau-\alpha_k)\boldsymbol{\delta}_k^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}-(\tau-\alpha_k)^2\tilde{\boldsymbol{\delta}}_{k+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{k+1}.$$

下面分 $\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}\leqslant 0$ 和 $\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}>0$ 两种情况,来分别证明 $\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_{k+1}-\boldsymbol{\delta}_{k+1}^{\mathrm{T}}\boldsymbol{\delta}_{k+1}\geqslant 0$.

当 $\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}\leqslant 0$ 时,有

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_k-\boldsymbol{\delta}_k^{\mathrm{T}}\boldsymbol{\delta}_k-(\tau-\alpha_k)\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}\geqslant 0,$$

又因为

$$h_k\leqslant\frac{\boldsymbol{\delta}_k^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}}{\tilde{\boldsymbol{\delta}}_{k+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{k+1}},$$

所以

$$2(\tau-\alpha_k)\boldsymbol{\delta}_k^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}-(\tau-\alpha_k)^2\tilde{\boldsymbol{\delta}}_{k+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{k+1}\geqslant 0.$$

故

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_{k+1}-\boldsymbol{\delta}_{k+1}^{\mathrm{T}}\boldsymbol{\delta}_{k+1}\geqslant 0.$$

当 $\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}>0$ 时,因为

$$h_k\leqslant\min\left\{\frac{\boldsymbol{\delta}_k^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}}{\tilde{\boldsymbol{\delta}}_{k+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{k+1}},\frac{\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_k-\boldsymbol{\delta}_k^{\mathrm{T}}\boldsymbol{\delta}_k}{\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}}\right\},$$

则可得

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_k-\boldsymbol{\delta}_k^{\mathrm{T}}\boldsymbol{\delta}_k-(\tau-\alpha_k)\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}\geqslant 0,$$

且

$$2(\tau-\alpha_k)\boldsymbol{\delta}_k^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{k+1}-(\tau-\alpha_k)^2\tilde{\boldsymbol{\delta}}_{k+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{k+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{k+1}\geqslant 0,$$

所以

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_{k+1}-\boldsymbol{\delta}_{k+1}^{\mathrm{T}}\boldsymbol{\delta}_{k+1}\geqslant 0.$$

即当 $n=k+1$ 时,$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_{k+1}-\boldsymbol{\delta}_{k+1}^{\mathrm{T}}\boldsymbol{\delta}_{k+1}\geqslant 0$ 成立.

故由第二数学归纳法知原命题成立.证毕.

定理 4.1 设 $\boldsymbol{B}$ 对称正定,且当 $n=2,3,\cdots,N$ 时,式(14)成立:

$$-\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_n-\boldsymbol{\delta}_{n-1}^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_n\geqslant-\boldsymbol{g}^{\mathrm{T}}\boldsymbol{B}^{-1}\tilde{\boldsymbol{\delta}}_n-\boldsymbol{\delta}_{n-1}^{\mathrm{T}}\tilde{\boldsymbol{\delta}}_n. \tag{14}$$

则 $\boldsymbol{\delta}(\tau)$ 满足如下要求:

(i) $\|\boldsymbol{\delta}(\tau)\|_2$ 关于 τ 为单调减函数;

(ii) $q[\boldsymbol{\delta}(\tau)]$ 关于 τ 为单调增函数.

证明 (i) 当 $\tau\in[\alpha_0,\alpha_1]$,即 $\tau\in[0,h_0]$ 时,

$$\|\boldsymbol{\delta}(\tau)\|_2^2=(\boldsymbol{\delta}_0-\tau(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1)^{\mathrm{T}}(\boldsymbol{\delta}_0-\tau(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1)=$$
$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_0-2\tau\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1+\tau^2\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_1,$$

则

$$(\|\boldsymbol{\delta}(\tau)\|_2^2)'=-2\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1+2\tau\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_1.$$

由式(12)可知

$$h_0\leqslant\frac{\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1}{\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_1},$$

则

$$(\|\boldsymbol{\delta}(\tau)\|_2^2)'\leqslant 0,\qquad \tau\in[\alpha_0,\alpha_1],$$

所以 $\|\boldsymbol{\delta}(\tau)\|_2$ 在区间$[\alpha_0,\alpha_1]$上为单调减函数.

对 $\forall\tau\in(\alpha_i,\alpha_{i+1}]$,即$(\tau-\alpha_i)\in(0,h_i]$,$i=1,2,\cdots,N-1$ 时,

$$\|\boldsymbol{\delta}(\tau)\|_2^2=(\boldsymbol{\delta}_i-(\tau-\alpha_i)(\boldsymbol{B}+\mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1})^{\mathrm{T}}(\boldsymbol{\delta}_i-(\tau-\alpha_i)(\boldsymbol{B}+\mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1})=$$
$$\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_i-2(\tau-\alpha_i)\boldsymbol{\delta}_i^{\mathrm{T}}(\boldsymbol{B}+\mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1}+(\tau-\alpha_i)^2\tilde{\boldsymbol{\delta}}_{i+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{i+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{i+1},$$

则

$$(\|\boldsymbol{\delta}(\tau)\|_2^2)'=-2\boldsymbol{\delta}_i^{\mathrm{T}}(\boldsymbol{B}+\mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1}+2(\tau-\alpha_i)\tilde{\boldsymbol{\delta}}_{i+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{i+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{i+1}.$$

由式(13)可知

$$h_i\leqslant\frac{\boldsymbol{\delta}_i^{\mathrm{T}}(\boldsymbol{B}+\mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1}}{\tilde{\boldsymbol{\delta}}_{i+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{i+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{i+1}},$$

则

$$(\|\boldsymbol{\delta}(\tau)\|_2^2)'\leqslant 0,\qquad \tau\in(\alpha_i,\alpha_{i+1}],$$

所以 $\|\boldsymbol{\delta}(\tau)\|_2$ 在区间$(\alpha_i,\alpha_{i+1}]$,$i=1,2,\cdots,N-1$ 上为单调减函数.

(ii) 当 $\tau\in[\alpha_0,\alpha_1]$,即 $\tau\in[0,h_0]$ 时,

$$
\begin{aligned}
q[\boldsymbol{\delta}(\tau)] = & \boldsymbol{g}^{\mathrm{T}}(\boldsymbol{\delta}_0 - \tau(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1) + \\
& \frac{1}{2}(\boldsymbol{\delta}_0 - \tau(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1)^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{\delta}_0 - \tau(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1) = \\
& \boldsymbol{g}^{\mathrm{T}}\boldsymbol{\delta}_0 - \tau\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1 + \frac{1}{2}\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\delta}_0 - \tau\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1 + \\
& \frac{1}{2}\tau^2\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\boldsymbol{B}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1,
\end{aligned}
$$

则

$$
\begin{aligned}
q[\boldsymbol{\delta}(\tau)]' = & -\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1 + \\
& \tau\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\boldsymbol{B}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1 \geqslant \\
& -\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{B} + \mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1 \geqslant \\
& -\boldsymbol{g}^{\mathrm{T}}\boldsymbol{B}^{-1}\tilde{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_0^{\mathrm{T}}\tilde{\boldsymbol{\delta}}_1 = \\
& \boldsymbol{\delta}_0^{\mathrm{T}}\tilde{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_0^{\mathrm{T}}\tilde{\boldsymbol{\delta}}_1 = 0, \qquad \tau \in [\alpha_0, \alpha_1].
\end{aligned}
$$

所以 $q[\boldsymbol{\delta}(\tau)]$ 在区间 $[\alpha_0,\alpha_1]$ 上关于 τ 为单调增函数.

对 $\forall\tau \in (\alpha_i,\alpha_{i+1}]$,即 $(\tau - \alpha_i) \in (0,h_i]$,$i = 1,2,\cdots,N-1$ 时,

$$
\begin{aligned}
q[\boldsymbol{\delta}(\tau)] = & \boldsymbol{g}^{\mathrm{T}}(\boldsymbol{\delta}_i - (\tau - \alpha_i)(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1}) + \\
& \frac{1}{2}(\boldsymbol{\delta}_i - (\tau - \alpha_i)(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1})^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{\delta}_i - (\tau - \alpha_i)(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1}) = \\
& \boldsymbol{g}^{\mathrm{T}}\boldsymbol{\delta}_i - (\tau - \alpha_i)\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1} + \frac{1}{2}\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\delta}_i - \\
& (\tau - \alpha_i)\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1} + \\
& \frac{1}{2}(\tau - \alpha_i)^2\tilde{\boldsymbol{\delta}}_{i+1}^{\mathrm{T}}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\boldsymbol{B}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1},
\end{aligned}
$$

则

$$
\begin{aligned}
q[\boldsymbol{\delta}(\tau)]' = & -\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1} - \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1} + \\
& (\tau - \alpha_i)\tilde{\boldsymbol{\delta}}_{i+1}^{\mathrm{T}}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\boldsymbol{B}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1} \geqslant \\
& -\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1} - \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{B} + \mu_{i+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{i+1} \geqslant \\
& -\boldsymbol{g}^{\mathrm{T}}\boldsymbol{B}^{-1}\tilde{\boldsymbol{\delta}}_{i+1} - \boldsymbol{\delta}_i^{\mathrm{T}}\tilde{\boldsymbol{\delta}}_{i+1} = \boldsymbol{\delta}_0^{\mathrm{T}}\tilde{\boldsymbol{\delta}}_{i+1} - \boldsymbol{\delta}_i^{\mathrm{T}}\tilde{\boldsymbol{\delta}}_{i+1} = \\
& \boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_i - \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_i - h_i'\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B} + \mu_i\boldsymbol{I})^{-1}\boldsymbol{\delta}_i + h_i'\boldsymbol{\delta}_i^{\mathrm{T}}(\boldsymbol{B} + \mu_i\boldsymbol{I})^{-1}\boldsymbol{\delta}_i \geqslant \\
& \boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_i - \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_i - h_i'\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B} + \mu_i\boldsymbol{I})^{-1}\boldsymbol{\delta}_i.
\end{aligned}
$$

由引理 4.1 可知

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_i - \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_i \geqslant 0,$$

又由式(11)可知

$$\boldsymbol{\delta}_0^{\mathrm{T}}\boldsymbol{\delta}_i - \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_i - h_i'\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B} + \mu_i\boldsymbol{I})^{-1}\boldsymbol{\delta}_i \geqslant 0,$$

则

$$q[\boldsymbol{\delta}(\tau)]' \geqslant 0, \qquad \tau \in (\alpha_i,\alpha_{i+1}].$$

所以 $q[\boldsymbol{\delta}(\tau)]$ 在区间 $(\alpha_i,\alpha_{i+1}]$,$i = 1,2,\cdots,N-1$ 上为单调增函数.证毕.

由定理 4.1 可得定理 4.2.

定理 4.2　设 $\boldsymbol{B}$ 对称正定,则对任意给定的信赖域半径 Δ,存在 N 使得 $\|\boldsymbol{\delta}_N\|_2 \leqslant \Delta$.

定理 4.1 和定理 4.2 说明对任意给定的信赖域半径 Δ,隐式分段折线 $\boldsymbol{\delta}(\tau)$ 上的近似解存

在且唯一.并且当沿隐式分段折线$\boldsymbol{\delta}(\tau)$求信赖域子问题(2)的最优解时,最优解$\boldsymbol{\delta}^*$在信赖域边界上达到.

5 隐式分段折线算法

通过上面的讨论,下面给出隐式分段折线算法的具体步骤:

步 0 给定梯度$\boldsymbol{g}$,正定矩阵$\boldsymbol{B}$,信赖域半径Δ.

步 1 令$\boldsymbol{\delta}_0=-\boldsymbol{B}^{-1}\boldsymbol{g}$,如果$\Delta\geqslant\|\boldsymbol{\delta}_0\|_2$,则取$\boldsymbol{\delta}^*=\boldsymbol{\delta}_0$.停止计算,否则转步 2.

步 2 令$\tilde{\boldsymbol{\delta}}_1=\boldsymbol{\delta}_0-h_0'(\boldsymbol{B}+\mu_0\boldsymbol{I})^{-1}\boldsymbol{\delta}_0$,其中$\mu_0=0$,$h_0'$为式(10).转步 3.

步 3 令$\boldsymbol{\delta}_1=\boldsymbol{\delta}_0-h_0(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1$,其中$\mu_1=h_0'$,$h_0$为式(12).如果$\Delta\geqslant\|\boldsymbol{\delta}_1\|_2$,则取

$$\boldsymbol{\delta}^*=\boldsymbol{\delta}_0-\eta(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1,$$

其中

$$\eta=\frac{b-\sqrt{b^2-ac}}{a},\ a=\tilde{\boldsymbol{\delta}}_1^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_1,\ b=\boldsymbol{\delta}_0^{\mathrm{T}}(\boldsymbol{B}+\mu_1\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_1,\ c=\|\boldsymbol{\delta}_0\|_2^2-\Delta^2.$$

停止计算,否则令$n:=1$,转步 4.

步 4 令$\tilde{\boldsymbol{\delta}}_{n+1}=\boldsymbol{\delta}_n-h_n'(\boldsymbol{B}+\mu_n\boldsymbol{I})^{-1}\boldsymbol{\delta}_n$,其中$\mu_n=\mu_{n-1}+h_{n-1}'$,$h_n'$为式(11).转步 5.

步 5 令$\boldsymbol{\delta}_{n+1}=\boldsymbol{\delta}_n-h_n(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1}$,其中$h_n$为式(13).如果$\Delta\geqslant\|\boldsymbol{\delta}_{n+1}\|_2$,则取

$$\boldsymbol{\delta}^*=\boldsymbol{\delta}_n-\eta(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1},$$

其中

$$\eta=\frac{b-\sqrt{b^2-ac}}{a},\ a=\tilde{\boldsymbol{\delta}}_{n+1}^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-2}\tilde{\boldsymbol{\delta}}_{n+1},$$

$$b=\boldsymbol{\delta}_n^{\mathrm{T}}(\boldsymbol{B}+\mu_{n+1}\boldsymbol{I})^{-1}\tilde{\boldsymbol{\delta}}_{n+1},\ c=\|\boldsymbol{\delta}_n\|_2^2-\Delta^2.$$

停止计算,否则令$n:=n+1$,转步 4.

6 数 值 结 果

选取不同的信赖域半径Δ,取限制步长$\varepsilon=0.3$,采用文献[10]中的测试函数 function 1 和 function 2,然后将隐式分段折线算法利用 MATLAB 进行了实验.并且用该算法求得的测试函数在近似最优解的函数值与切线单折线法求得的测试函数在近似最优解的函数值进行了比较.数值结果列在表 1,2 中,其中Δ表示信赖域半径,q_{TDL}表示使用切线单折线法所求得的测试函数在近似最优解的函数值,q_{ISD}表示使用隐式分段折线算法所求得的测试函数在近似最优解的函数值.$q_{\mathrm{TDL}}-q_{\mathrm{ISD}}$表示使用切线单折线法和隐式分段折线算法所求得的测试函数在近似最优解的函数值之差,如果该值大于 0,则表明隐式分段折线算法比切线单折线法好.

从表 1,2 的数值结果可以看出,总体来说,隐式分段折线算法要优于切线单折线法.对于文献[10]中的测试函数 function 1,当信赖域半径$6.4\leqslant\Delta\leqslant8.5$时,切线单折线法要稍微好于隐式分段折线算法,当信赖域半径$\Delta\geqslant10.20$时,隐式分段折线算法和切线单折线法的效果一样,而对于其他给定的信赖域半径Δ,隐式分段折线算法所求得的数值结果都要好于切线单折线法;对于文献[10]中的测试函数 function 2,当信赖域半径$5.8\leqslant\Delta\leqslant8.3$时,切线单折线法所求得的数值结果稍微好于隐式分段折线算法,当信赖域半径$\Delta\geqslant10.02$时,隐式分段折线算法和切线单折线法所求得的数值结果相同,而对于其他给定的信赖域半径Δ,隐式分段折线算法所求得的数值结果都要优于切线单折线法.因此,本文构造的隐式分段折线比切线单折线更好

地近似了最优曲线,而且隐式分段折线算法是一个比切线单折线法更好的求解信赖域子问题的折线法.

Function 1[11]: $\|\boldsymbol{\delta}_{zp}\|_2 = 2.36$, $\|\boldsymbol{B}^{-1}\boldsymbol{g}\|_2 = 10.20$;

Function 2[12]: $\|\boldsymbol{\delta}_{zp}\|_2 = 0.67$, $\|\boldsymbol{B}^{-1}\boldsymbol{g}\|_2 = 10.02$.

表 1　测试 function 1 的数值结果

Table 1　The numerical results of test function 1

Δ	q_{TDL}	q_{ISD}	$q_{TDL} - q_{ISD}$
1.0	−12.642 135 624	−12.706 491 521	0.064 355 897
1.5	−17.838 203 436	−18.209 825 575	0.371 622 139
1.8	−20.595 844 123	−21.282 503 284	0.686 659 161
2.36	−25.033 982 776	−26.605 402 413	1.571 419 637
3.0	−31.442 439 446	−32.093 838 865	0.651 399 419
4.0	−39.354 809 845	−39.536 873 516	0.182 063 671
6.0	−50.744 075 760	−50.749 216 892	0.005 141 132
6.3	−52.036 515 879	−52.036 951 807	0.000 435 928
6.4	−52.444 829 034	−52.444 305 661	−0.000 523 373
6.5	−52.841 989 407	−52.840 753 509	−0.001 235 898
7.0	−54.662 229 476	−54.658 468 267	−0.003 761 209
8.0	−57.490 211 014	−57.486 745 678	−0.003 465 336
8.5	−58.505 028 511	−58.502 380 447	−0.002 648 064
8.6	−58.676 394 829	−59.939 446 057	1.263 051 228
9.0	−59.257 074 687	−59.964 568 507	0.707 493 820
9.5	−59.748 170 054	−59.987 367 773	0.239 197 719
10.0	−59.979 757 639	−59.998 933 693	0.019 176 054
≥10.20	−60	−60	0

表 2　测试 function 2 的数值结果

Table 2　The numerical results of test function 2

Δ	q_{TDL}	q_{ISD}	$q_{TDL} - q_{ISD}$
0.3	−3.770 140 687	−3.825 648 982	0.055 508 295
0.67	−7.118 505 868	−7.892 320 985	0.773 815 117
1.0	−10.896 622 252	−11.186 904 994	0.290 282 742
2.0	−20.026 890 040	−20.070 641 529	0.043 751 489
3.0	−27.722 865 703	−27.733 242 737	0.010 377 034
3.5	−31.153 497 775	−31.158 703 979	0.005 206 204
5.0	−39.879 056 251	−39.879 490 101	0.000 433 850
5.7	−43.163 177 442	−43.163 188 623	0.000 011 181
5.8	−43.591 780 528	−43.591 749 219	−0.000 031 309
6.5	−46.309 000 235	−46.308 881 419	−0.000 118 816
7.0	−47.947 248 925	−47.947 118 759	−0.000 130 166
8.0	−50.468 975 959	−50.468 893 766	−0.000 082 193
8.3	−51.029 518 712	−51.029 454 033	−0.000 064 679
8.4	−51.196 282 593	−52.440 848 945	1.244 566 352
9.0	−51.986 082 151	−52.475 203 821	0.489 121 670
9.5	−52.368 347 997	−52.493 325 055	0.124 977 058
≥10.02	−52.5	−52.5	0

参考文献(References):

[1] 徐成贤，陈志平，李乃成. 近代优化方法[M]. 北京：科学出版社，2002.(XU Cheng-xian, CHEN Zhi-ping, LI Nai-cheng. *Modern Optimization Method*[M]. Beijing: Science Press, 2002.(in Chinese))

[2] 后六生，孙文瑜. 三项预处理共轭梯度法与信赖域子问题[J]. 南京师大学报(自然科学版)，2001，**24**(3)：1-6.(HOU Liu-sheng, SUN Wen-yu. Three-term preconditioned conjugate gradient method and trust region subproblem[J]. *Journal of Nanjing Normal University* (*Natural Science*), 2001, **24**(3): 1-6.(in Chinese))

[3] 赵英良，徐成贤. 信赖域子问题使用重新开始策略的共轭梯度法[J]. 高校应用数学学报(A辑)，2003，**18**(3)：341-349.(ZHAO Ying-liang, XU Cheng-xian. Conjugate gradient method using restart strategy for solving trust region subproblems[J]. *Applied Mathematics: A Journal of Chinese Universities*(*Ser A*), 2003, **18**(3): 341-349.(in Chinese))

[4] 陈争，马昌风. 求解信赖域子问题的一个光滑牛顿法[J]. 福建师范大学学报(自然科学版)，2011，**27**(4)：31-35.(CHEN Zheng, MA Chang-feng. A smoothing Newton method for trust region subproblem[J]. *Journal of Fujian Normal University*(*Natural Science Edition*), 2011, **27**(4): 31-35.(in Chinese))

[5] 张立，唐志强. 解信赖域子问题的混合折线法[J]. 南京师大学报(自然科学版)，2001，**24**(1)：28-32.(ZHANG Li, TANG Zhi-qiang. The hybrid dogleg method to solve subproblems of trust region[J]. *Journal of Nanjing Normal University* (*Natural Science*), 2001, **24**(1): 28-32.(in Chinese))

[6] PU Ding-guo, HAN Bo-shun, YAO Lin, ZHENG Guang-hua. A class of dogleg trust region methods[J]. *Operations Research Transactions*, 2003, **7**(1): 1-10.

[7] CHEN Jun, SUN Wen-yu. Nonmonotone adaptive trust region algorithms with indefinite dogleg path for unconstrained minimization[J]. *Northeast Math Journal*, 2008, **24**(1): 19-30.

[8] Powell M J D. A hybrid method for nonlinear equations[C]//Rabonowitz P. *Numerical Methods for Nonlinear Algebraic Equations*. London: Gordon and Breach, 1970: 87-114.

[9] Dennis J E, Mei H H W. Two new unconstrained optimization algorithms which use function and gradient values[J]. *Journal of Optimization Theory and Applications*, 1979, **28**(4): 453-482.

[10] 赵英良，徐成贤. 解信赖域子问题的切线单折线法[J]. 数值计算与计算机应用，2000，**21**(1)：77-80.(ZHAO Ying-liang, XU Cheng-xian. Tangent single dogleg method for trust region subproblem[J]. *Journal of Numerical Methods and Computer Applications*, 2000, **21**(1): 77-80.(in Chinese))

[11] 颜庆津. 数值分析[M]. 北京：北京航空航天大学出版社，2006.(YAN Qing-jin. *Numerical Analysis*[M]. Beijing: Beihang University Press, 2006.(in Chinese))

[12] 袁亚湘，孙文瑜. 最优化理论与方法[M]. 北京：科学出版社，2007.(YUAN Ya-xiang, SUN Wen-yu. *Optimization Theory and Methods*[M]. Beijing: Science Press, 2007.(in Chinese))

[13] 李董辉，童小娇，万中. 数值最优化算法与理论[M]. 北京：科学出版社，2010.(LI Dong-hui, TONG Xiao-jiao, WAN Zhong. *Numerical Optimization Algorithms and Theory*[M]. Beijing: Science Press, 2010.(in Chinese))

An Implicit Piecewise Dogleg Algorithm for Solving Trust-Region Subproblems

WANG Xi-yun, LI Liang, YU Hai-bo

(*School of Applied Science, Taiyuan University of Science and Technology, Taiyuan* 030024, *P.R.China*)

Abstract: Based on the premise that Hessian matrix was positive definite, a differential equation model was established for the optimal curve. Then an implicit piecewise dogleg was constructed according to the differential equation. In turn, the implicit piecewise dogleg algorithm for solving trust-region subproblems was presented. And the rationality of the implicit piecewise dogleg path was analyzed and demonstrated. Numerical results indicate that the new algorithm is effective and practicable.

Key words: implicit piecewise dogleg algorithm; differential equation model; trust-region subproblem