# 有界变量约束优化的非单调最优路径内点算法

郭佩华，朱德通

(上海师范大学 数理信息学院，上海 200234)

**摘　要**：采用最优路径结合非单调内点回代算法解有界变量约束的非线性优化问题. 从构建的最优路径解二次模型获得迭代方向，通过线搜索获得步长因子以保证迭代点既落在严格可行域内，又能使目标函数产生足够下降. 基于导出的最优路径的良好性质，在合理的假设下，证明了此算法不仅具有整体收敛性，而且保持局部超线性收敛速率. 引入非单调技术将克服病态问题，从而加速收敛性进程. 数值计算表明了算法的可行性和有效性.

**关键词**：有界变量约束；最优路；内点法；非单调技术

**中图分类号**：O221.2　　**文献标识码**：A　　**文章编号**：1000-5137(2004)03-0023-07

## 1 介　绍

本文研究有界变量约束的非线性优化问题

$$\min_{x\in\Omega} f(x). \tag{1.1}$$

其中

$$\Omega \overset{\text{def}}{=} \{\, x \in \mathscr{R}^n \mid l^i \le x^i \le u^i,\ i = 1,\cdots, n \} \tag{1.2}$$

有界变量约束集 $\Omega$ 是非空集，$x^i$，$l^i$ 和 $u^i$ 分别为 $n$ 维向量 $x$，$l$ 和 $u$ 的第 $i$ 个分量；对所有的 $i$，$l^i \in \mathscr{R} \cup \{-\infty\}$ 和 $u^i \in \mathscr{R} \cup \{+\infty\}$ 满足 $l^i < u^i$；$f: \mathscr{R}^n \to \mathscr{R}$ 是光滑的非线性函数. $\Omega$ 称为可行集，令 $\mathrm{int}(\Omega) \overset{\text{def}}{=} \{\, x \in \mathscr{R}^n \mid l^i < x^i < u^i,\ i = 1,\cdots,n\}$ 为严格内点可行集.

记 $g(x) = \nabla f(x)$ 为目标函数 $f(x)$ 的梯度，$g^i(x)$ 为 $g(x)$ 的第 $i$ 个分量，问题(1.1)的一阶必要性条件为：若 $x_* \in \Omega$ 满足

$$\begin{cases} g^i(x_*) \ge 0 & \text{当 } x_*^i = l^i \text{ 时} \\ g^i(x_*) = 0 & \text{当 } l^i < x_*^i < u^i \text{ 时,} \\ g^i(x_*) \le 0 & \text{当 } x_*^i = u^i \text{ 时} \end{cases} \tag{1.3}$$

称 $x_*$ 为问题(1.1)的稳定点.

通过引入仿射尺度矩阵 $D_k$，Coleman 和 Li 在[3]中提出一种双信赖域内点算法和可行信赖域内点算法来解有界变量约束的非线性优化问题，算法中定义向量 $v(x) \in \mathscr{R}^n$ 的每一个分量为：

$$v_i(x) \overset{\text{def}}{=} \begin{cases} x^i - u^i & \text{若 } g^i < 0,\ \text{且 } u^i < +\infty \\ x^i - l^i & \text{若 } g^i \ge 0,\ \text{且 } l^i > -\infty \\ -1 & \text{若 } g^i < 0,\ \text{且 } u^i = +\infty \\ 1 & \text{若 } g^i \ge 0,\ \text{且 } l^i = -\infty \end{cases}, \tag{1.4}$$

---

**收稿日期**：2004-04-10

**作者简介**：郭佩华(1980-)，女，上海师范大学数理信息学院研究生. 朱德通(1954-)，男，博士，上海师范大学数理信息学院教授.

记对角阵 $D(x) \overset{\text{def}}{=} \mathrm{diag}\{|v_1(x)|^{-\frac{1}{2}},\cdots,|v_n(x)|^{-\frac{1}{2}}\}$，即 $D^{-2}(x) = \mathrm{diag}\{|v(x)|\}$.

在不涉及可行性与对偶可行性时,问题(1.1)的一阶必要性条件(1.3)可以等价于下面的非线性方程组:

$$D(x)^{-2}g(x) = 0, \tag{1.5}$$

若第 $k$ 次迭代点 $x_k \in \mathrm{int}(\Omega)$，则用(1.5)解方向 $\delta$ 的牛顿步为:

$$(D_k^{-2}\nabla^2 f(x_k) + \mathrm{diag}(g_k)J_k^v)\delta_k = -D_k^{-2}g_k(x), \tag{1.6}$$

其中 $J^v \in \mathscr{R}^{n\times n}$ 为 $|v(x)|$ 的 Jacobian 矩阵.方程(1.5)是连续的,但不是处处可微的.只当 $v_i = 0$ 时,方程(1.5)是不可微的.为了避免这种情况,可以限制 $x_k \in \mathrm{int}(\Omega)$，从而 $|v_i| > 0$.当 $g^i = 0$ 时,$v_i$ 是不连续的,因此规定当 $g_i = 0$ 时,取定 $J_{ii}^v = 0$.如 Coleman 和 Li 在[3]指出的,虽然 $v_i$ 在这一点是不连续的,但是函数 $v_i \cdot g^i$ 却是连续的,因此 $v_i$ 的值并不重要,故可以不考虑这种类型的不可微性.显然,$J^v$ 是一个对角矩阵,其对角线元素为0或1.当 $l$ 和 $u$ 均不为 $\infty$ 时,$J^v = \mathrm{diag}\{\mathrm{sgn}(g^1),\cdots,\mathrm{sgn}(g^n)\}$;当存在有无限下界或无限上界,或者 $g_i = 0$ 时,$J_{ii}^v = 0$.

基于一阶必要性条件(1.5)的牛顿步(1.6),两边乘上尺度矩阵 $D_k$,经过变换后,Coleman 和 Li 提出了问题(1.1)的信赖域子问题 $(S_k)$:

$$(S_k)\begin{cases}\min & q_k(\delta) \overset{\text{def}}{=} f(x_k) + (g_k)^T\delta + \dfrac{1}{2}\delta^T(B_k + C_k)\delta \\ \text{s.t.} & \|D_k\delta\| \leqslant \Delta_k,\end{cases} \tag{1.7}$$

其中 $g_k = \nabla f(x_k)$,$\delta = x - x_k$,$B_k$ 是 $\nabla^2 f(x_k)$ 或其对称近似矩阵,矩阵 $C_k \overset{\text{def}}{=} D_k\,\mathrm{diag}\{g_k^1,\cdots,g_k^n\}J_k^vD_k$,其中 $g_k^i$ 是 $g_k$ 的第 $i$ 个分量,显然 $C_k$ 是半正定对角矩阵.$q_k(\delta)$ 是 $f$ 的局部二次近似,$\Delta_k$ 是信赖域半径.令 $\hat{g}_k \overset{\text{def}}{=} D_k^{-1}g_k = \mathrm{diag}(|v_k|^{\frac{1}{2}})g_k$,$\hat{M}_k \overset{\text{def}}{=} D_k^{-1}(B_k + C_k)D_k^{-1}$,置 $\hat{\delta} = D_k\delta$,信赖域子问题($S_k$)可以重新改写为($S_k'$):

$$(S_k')\begin{cases}\min & \hat{q}_k(\hat{\delta}) \overset{\text{def}}{=} f(x_k) + \hat{g}_k^T\hat{\delta} + \dfrac{1}{2}\hat{\delta}^T\hat{M}_k\hat{\delta} \\ \text{s.t.} & \|\hat{\delta}\| \leqslant \Delta_k,\end{cases} \tag{1.8}$$

在[3]中,双信赖域方法具有整体收敛性,并且保持局部超线性收敛速率.信赖域方法的优点之一就是它并不要求目标函数是凸的,但为了保持可行性和信赖域接受准则,需要增加计算子问题的难度和计算工作量.本文采用了最优路径算法来代替信赖域算法,最优路径可以用二次模型函数中 Hessian 阵或其近似阵的特征值和特征向量表示,它同样也不要求目标函数是凸的.此外,Grippo 等[4]提出了非单调技术,非单调技术并不要求目标函数在每一步都下降,从而可以避免在具有多个峡谷的非线性函数中,算法被困在谷底的情况.

本文基于构建最优路径解二次模型获得迭代方向,通过线搜索获得步长因子以保证迭代点既落在严格可行域内,又能使目标函数产生非单调下降.在合理的假设下,可以证明此算法不仅满足整体收敛性而且具备局部超线性收敛速率.内容概述如下:第二节给出了最优路径的构造及其性质,第三节描述了结合最优路径和线搜索非单调技术的内点算法,第四节证明了整体收敛性,第五节讨论了局部收敛速率,最后,给出了采用最优路径算法的数值实验的结果.

## 2 最优路径及其性质

在信赖域算法中,求解变量有界约束最小化问题基于求解子问题($S_k'$).在信赖域子问题的二次模型中,忽略信赖域半径的约束,采用最优路径法获得迭代步 $\delta_k$,则下一个迭代步是 $x_{k+1} = x_k + \delta_k$.

为使有效地解信赖域子问题($S_k'$),采纳最优路径去近似求解,从而对点 $x_k$ 能产生一系列有效的迭代步 $\delta_k$.为了给出曲线路的具体形式,需要对 $D_k^{-1}(B_k + C_k)D_k^{-1}$ 进行特征值分解.因为 $D_k^{-1}(B_k + C_k)D_k^{-1}$

是对称矩阵，它的特征值是实数，不失一般性，令 $\phi_1 \leqslant \phi_2 \leqslant \cdots \leqslant \phi_n$ 是 $D_k^{-1}(B_k + C_k)D_k^{-1}$ 的特征值，$u^1$，$u^2, \cdots, u^n$ 为相应的特征向量. 根据 $\phi_i > 0$，$\phi_i < 0$ 和 $\phi_i = 0$（$i \in \{1, \cdots, n\}$），将把指标集 $\{1, \cdots, n\}$ 划分为 $\mathscr{I}^+$，$\mathscr{I}^-$ 和 $\mathscr{N}$. 现在来构造最优路解信赖域子问题（$S_k'$）.（参见 Bulteau 和 Vial[2]）

### 2.1 最优路径的构成

令 $\hat{g} \overset{\text{def}}{=} D_k^{-1} g$ . 最优路径 $\Gamma(\tau)$ 可以表示为 $\Gamma(\tau) = \Gamma_1(t_1(\tau)) + \Gamma_2(t_2(\tau))$ . 其中

$$\Gamma_1(t_1(\tau)) = -\left[\sum_{i \in \mathscr{I}} \frac{t_1(\tau)}{\phi_i t_1(\tau) + 1} \hat{g}_k^i u^i + t_1(\tau) \sum_{i \in \mathscr{N}} \hat{g}_k^i u^i\right], \tag{2.1}$$

$$\Gamma_2(t_2(\tau)) = t_2(\tau) u^1, \tag{2.2}$$

这里

$$t_1(\tau) = \begin{cases} \tau, & \text{如果 } \tau < \dfrac{1}{T}, \\ \dfrac{1}{T}, & \text{如果 } \tau \geqslant \dfrac{1}{T}, \end{cases} \tag{2.3}$$

$$t_2(\tau) = \begin{cases} 0, & \text{如果 } \tau < \dfrac{1}{T}, \\ \tau - \dfrac{1}{T}, & \text{如果 } \tau \geqslant \dfrac{1}{T}, \end{cases} \tag{2.4}$$

$\mathscr{I} = \{ i \mid \phi_i \neq 0, i = 1, \cdots, n\}$，$\mathscr{N} = \{ i \mid \phi_i = 0\ i = 1, \cdots, n\}$，$\hat{g}_k^i = (\hat{g}_k)^T u^i$，$i = 1, \cdots, n$，$\hat{g}_k = \sum_{i=1}^{n} \hat{g}_k^i u^i$，$T = \max\{0, -\phi_1\}$ 且当 $T = 0$ 时，$1/T$ 定义为 $+\infty$ . 注意，只有当 $D_k^{-1}(B_k + C_k)D_k^{-1}$ 是不定的，且对所有满足 $\phi_i = \phi_1 < 0$ 的 $i \in \{1, \cdots, n\}$，有 $\hat{g}_k^i = 0$，这时 $\Gamma_2(t_2(\tau))$ 才存在. 这种情况被 Moré 和 D. C. Sorensen 称为最优化问题的困难情况（**hard case**）（见[6]）. 对其他情况 $\Gamma(\tau)$ 只在区间 $0 \leqslant \tau < \dfrac{1}{T}$ 内有定义，也就是 $\Gamma(\tau) = \Gamma_1(t_1(\tau))$ .

### 2.2 最优路径的性质

下面先证明最优路径的一些性质，以使所提供的算法具有整体收敛性，并保持局部收敛速率. 由于篇幅限制，只给出结论，以下的引理和定理都省略了具体的证明过程.

**引理 2.1**　由最优路径得到的迭代步 $\hat{\delta}_k(\tau) = \Gamma_k(\tau)$ 满足以下性质：

(1) 对于 $\tau \in (0, +\infty)$，路径的范数函数 $\|\hat{\delta}_k(\tau)\|$ 是单调递增的. 进一步，有 $\lim\limits_{\tau \to 0} \dfrac{\|\hat{\delta}_k(\tau)\|^2}{\tau^2} = \|\hat{g}_k\|^2$.

(2) 二次近似函数 $q_k(\delta_k(\tau))$ 沿着最优路径是单调递减的.

(3) 如果 $(B_k + C_k)$ 是正定的，则 $\lim\limits_{\tau \to \infty} \hat{\delta}_k(\tau) = -[D_k(B_k + C_k)^{-1} D_k] \hat{g}_k$.

**引理 2.2**　记迭代步 $\hat{\delta}_k(\tau) = \Gamma_k(\tau)$ 是由最优路径得到的，则

(1) 对于 $0 < \tau < \dfrac{1}{T}$，目标函数的梯度 $\hat{g}_k$ 和最优路径产生的迭代步 $\hat{\delta}_k(\tau)$ 的关系函数 $\hat{\psi}_k(\tau) \overset{\text{def}}{=} \hat{g}_k^T \hat{\delta}_k(\tau)$ 是单调递减的. 进一步，当 $\tau \to 0$ 时，$\hat{g}_k^T \dfrac{d\hat{\delta}_k(\tau)}{d\tau} = \dfrac{d\hat{\psi}_k(\tau)}{d\tau} \to -\|\hat{g}_k\|^2$ .

(2) 对于 $0 < \tau < +\infty$，$\hat{\psi}_k(\tau) \overset{\text{def}}{=} \hat{g}_k{}^T \hat{\delta}_k(\tau)$ 满足充分下降方向，即

$$\hat{\psi}_k(\tau) \leqslant -\frac{1}{2} \min\left\{\tau, \frac{1}{\|\hat{M}_k\|}, \frac{1}{T}\right\} \|\hat{g}_k\|^2.$$

**引理 2.3**　令迭代步 $\hat{\delta}_k(\tau_k)$ 是由最优路径得到的，则一定存在 $\mu_k \in (0, +\infty)$，使得下式成立：

$$(\hat{M}_k + \mu_k I)\hat{\delta}_k(\tau_k) = -\hat{g}_k. \tag{2.5}$$

其中

$$\mu_k = \begin{cases} \frac{1}{\tau_k} & 当 0 < \tau < \frac{1}{T} \\ T & 当 \tau \geq \frac{1}{T} \end{cases} \tag{2.6}$$

且 $(\hat{M}_k + \mu_k I)$ 是半正定矩阵.

## 3 算法

在这节中,我们给出求解问题(1.1)的非单调最优路径内点算法.

**初始步**

选定参数 $\beta \in (0, \frac{1}{2})$, $\omega$, $\tilde{w} \in (0,1)$, $\gamma \in (0,1)$, $\epsilon > 0$,正整数 $M$. 令 $m(0) = 0$. 选取对称矩阵 $B_0$,选取初始点 $x_0 \in \mathrm{int}(\Omega)$,再令 $k = 0$ 转主步.

**主步**

(1) 计算 $f_k = f(x_k)$, $g_k = \nabla f(x_k)$, $D_k$ 和 $C_k$.

(2) 如果 $\| D_k^{-1} g_k \| \leq \epsilon$,则停止计算,$x_k$ 作为最优解;否则转下一步.

(3) 计算 $D_k^{-1}(B_k + C_k)D_k^{-1}$ 的特征值和正交特征向量. 构造最优路径 $\Gamma_k(\tau)$,获得路 $\hat{\delta}_k(\tau)$.

令 $\delta_k(\tau) = D_k^{-1}\hat{\delta}_k(\tau)$,

(a) 若 $\hat{M}_k$ 正定或半正定,选取 $\tau_k = \infty$, $\omega^{-n}$, $\omega^{-(n-1)}$, $\cdots$,

(b) 若 $\hat{M}_k$ 不定,且对 $\phi_i = \phi_1$,有 $\hat{g}_k^i \neq 0$,则选取 $\tau_k = \frac{1}{T}\omega$, $\frac{1}{T}\omega^2$, $\frac{1}{T}\omega^2$, $\cdots$,

(c) 若 $\hat{M}_k$ 不定,且对 $\phi_i = \phi_1$,有 $\hat{g}_k^i = 0$,则选取 $\tau_k = \infty$, $\omega^{-n}$, $\omega^{-(n-1)}$, $\cdots$,

直到下列不等式都满足

$$f(x_k) - f(x_k + \delta_k(\tau_k)) \geq \gamma [f(x_k) - q_k(\delta_k(\tau_k))] \tag{3.1}$$

(4) 选取 $\alpha_k = 1, \tilde{w}, \tilde{w}^2, \cdots$,直到下列不等式满足:

$$f(x_k + \alpha_k \delta_k(\tau_k)) \leq f(x_{(l(k))}) + \beta \alpha_k g_k^T \delta_k(\tau_k) \tag{3.2}$$

$$x_k + \alpha_k \delta_k(\tau_k) \in \Omega \tag{3.3}$$

其中 $f(x_{l(k)}) = \max_{0 \leq j \leq m(k)} \{ f(x_{k-j}) \}$.

(5) 令

$$h_k = \begin{cases} \alpha_k \delta_k(\tau_k), & 若 x_k + \alpha_k \delta_k(\tau_k) \in \mathrm{int}(\Omega), \\ \theta_k \alpha_k \delta_k(\tau_k), & 否则. \end{cases} \tag{3.4}$$

其中 $\theta_k \in (\theta_l, 1]$, $0 < \theta_l < 1$ 并且 $\theta_k - 1 = O(\| \hat{\delta}_k \|)$. 从而令 $x_{k+1} = x_k + h_k$.

(6) 计算 $f(x_{k+1})$ 和 $g_{k+1}$. 取 $m(k+1) = \min\{m(k) + 1, M\}$,校正 $B_k$ 得到 $B_{k+1}$,置 $k \leftarrow k + 1$,再转步2.

**注释1** 在子问题 $(S_k)$ 中, $q_k(\delta_k(\tau))$ 是在 $x_k$ 附近函数 $f$ 的局部二次模型,迭代步是沿着最优路径得到的.

**注释2** 若 $\alpha_k$ 表示算法第4步中沿 $\delta_k$ 方向受有界变量约束 $l \leq x_k + \alpha_k \delta_k \leq u$ 的步长,即:

$$\alpha_k^* \stackrel{\mathrm{def}}{=} \min\{ \max\{ \frac{l_i - x_k^i}{\delta_k^i}, \frac{u_i - x_k^i}{\delta_k^i} \}, i = 1, 2, \cdots, n \},$$

其中 $\delta_k^i$ 是向量 $\delta_k$ 中的第 $i$ 个分量,若 $\delta_k^i = 0$,则令 $\frac{l_i - x_k^i}{\delta_k^i} = \frac{u_i - x_k^i}{\delta_k^i} = +\infty$, $\alpha_k$ 的关键特性是取步长 $\alpha_k \delta_k$ 后得到的点都不会超过有界约束的边界. 如果步长因子正好使得 $x_k + \alpha_k \delta_k(\tau_k)$ 在有界变量约束边界

上，则乘上因子 $\theta_k$ 来保证每次迭代 $x_k + h_k$ 均为严格可行内点，即 $l < x_k + h_k < u$.

**注释 3** 单调算法可以看成是非单调算法在 $M = 0$ 时的特殊情形.

**注释 4** 在算法的第5步中，保证了序列 $\{x_k\}$ 为严格可行内点，在以下的收敛性分析中，假设 $\{x_k\}$ 不会充分接近于边界，即存在 $\chi_D > 0$，使得 $\|D_k\| \leq \chi_D$.

## 4 整体收敛性

在这一节中，$f:\mathscr{R}^n \to \mathscr{R}^1$ 是二次连续可微函数，且在 $\Omega$ 上是有下界的. 在整体收敛性分析中，我们通常作如下假设：

**假设(A1)** 给定 $x_0 \in \Omega$，水平集 $\mathscr{L}(x_0) \overset{\mathrm{def}}{=} \{x \in \Omega \mid f(x) \leq f(x_0)\}$ 是紧的，算法产生的序列 $x_k \in \mathscr{L}(x_0)$.

**假设(A2)** $\|B_k\|$ 是上有界. 即对任意的 $k$，存在 $\chi_B > 0$，使得 $\|B_k\| \leq \chi_B$.

**假设(A3)** $\|\nabla^2 f(x)\|$ 是上有界. 即对任意的 $x \in \mathscr{L}(x_0)$，存在 $\chi_G > 0$，使得 $\|\nabla^2 f(x)\| \leq \chi_G$.

**假设(A4)** $\|g(x)\|_\infty$ 是上有界. 即对任意的 $x \in \mathscr{L}(x_0)$，存在 $\chi_g > 0$，使得 $\|g(x)\|_\infty \leq \chi_g$.

**假设(A5)** $\|\hat{\delta}_k(\tau_k)\|$ 是上有界. 即对任意的 $k$，存在 $\hat{\chi}_\delta > 0$，使得 $\|\hat{\delta}_k(\tau_k)\| \leq \hat{\chi}_\delta$.

由假设(A1)可知，尺度矩阵 $\|D_k^{-1}\|$ 是有界的，即 $\|D_k^{-1}\| \leq \chi_{D^{-1}}$；由假设(A1)～(A3)可知，存在 $\hat{\chi}_M, \chi, \hat{\chi}_g$，满足 $\|\hat{M}_k\| \leq \hat{\chi}_M$，$\|D_k^{-1}\nabla^2 f(x) D_k^{-1}\| \leq \chi$，$\|\hat{g}_k\| \leq \hat{\chi}_g$.

又因为在算法中假设 $\|D_k\| \leq \chi_D$，结合假设(A4)可得：$\|C_k\| \leq \chi_C$，从而 $\|M_k\|$ 也是上有界的，即存在 $\chi_M > 0$，使得 $\|M_k\| \leq \chi_M$.

**引理 4.1** 在第 $k$ 步迭代中，令 $\delta_k(\tau)$ 是由最优路径得到的，则预计下降量 $f(x_k) - q_k(\delta_k(\tau))$ 满足充分下降条件，即：$f(x_k) - q_k(\delta_k(\tau)) \geq \frac{1}{4}\min\{\tau, \frac{1}{\|\hat{M}_k\|}, \frac{1}{T}\}\|\hat{g}_k\|^2$.

在非单调最优路径算法中，下面的引理保证了算法第三步是有限步可行的.

**引理 4.2** 假设 $f:\mathscr{R}^n \to \mathscr{R}^1$ 是二次连续可微函数，给定起始点 $x_0$，利用第三节中给出的算法，则在第 $k$ 迭代步中，或者 $\hat{g}_k = 0$ 或者在算法的第三步，经过有限次迭代后，产生一个 $\tau_k$，满足(3.1)式.

在保证预计下降量满足充分下降条件以及算法有限步终止的前提下，定理4.3进一步说明了此算法具有弱收敛性质.

**定理 4.3** 假设(A1)～(A5)成立，令 $\{x_k\} \subset \mathscr{R}^n$ 是由算法产生的序列. 则 $\liminf_{k\to\infty}\|D_k^{-1}g_k\| = 0$.

## 5 局部收敛性

上述定理只说明算法具有弱收敛性质，为使获得强收敛和超线性的收敛速率，需要先给出如下假设：

**假设(A6)** 假设 $\lim_{k\to\infty}\frac{\|[B_k - \nabla^2 f(x_k)]\delta_k(\tau_k)\|}{\|\delta_k(\tau_k)\|} = 0$.

**假设(A7)** $M_k = B_k + C_k$ 是一致正定的，即存在 $\rho > 0$，使得：$p^T M_k p \geq \rho\|p\|^2$.

**定理 5.1** 若假设(A1)－(A7)都满足，设 $\{x_k\}$ 为算法产生的迭代序列，则 $\lim_{k\to\infty}\|\hat{g}_k\| = 0$.

**定理 5.2** 若假设(A1)～(A7)都满足，设 $\{x_k\}$ 为算法产生的迭代序列，则 $\{x_k\}$ 超线性收敛于 $x_*$，即：

$$\lim_{k\to\infty}\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0. \tag{5.1}$$

定理5.1说明了算法具有强整体收敛性，定理5.2阐述了算法的局部收敛速率依赖于目标函数Hessian矩阵的性质，当迭代步 $h_k$ 为牛顿步时，由算法产生的序列 $\{x_k\}$ 超线性收敛于 $x_*$.

## 6 数值结果

对上述最优路算法,我们利用数学软件 Matlab 进行编程,在型号为奔腾 IV - 1.70GHz 的电脑上实现了具体的数值结果.选取参数如下:$\epsilon = 10^{-6}$ , $\gamma = 0.02$ , $\gamma_3 = 2$ , $\beta = 0.4$ , $w1 = 0.5$ , $w2 = 0.5$ .

应用了8个标准数值测试题,引自于文献[15]和[19].每题既有文献[15]和[19]给出的有界约束(或无界约束),也有另外给出的.数值计算结果由表给出,其中NF表示目标函数的计算次数,NG表示其梯度函数的计算次数,$M$表示非单调控制参数.数值结果表示所提供的算法的有效性和可靠性,通过对数值结果的分析,可以看到非单调技术在 Rosenbrock 问题的求解时,提高了计算速度,优于单调搜索.但对其他问题并没有太大的改善.此外,当最优解落在边界上时,内点算法不是很理想,因而需进一步改善.

表1 数值结果

| 编号 | 有界约束 | M = 0 | | M = 5 | | M = 10 | |
|---|---|---|---|---|---|---|---|
| | | NF | NG | NF | NG | NF | NG |
| 1 | $l, u$ | 59 | 24 | 56 | 14 | 56 | 14 |
| | $\hat{l}, \hat{u}$ | 57 | 23 | 115 | 12 | 115 | 12 |
| 2 | $l, u$ | 163 | 35 | 263 | 25 | 263 | 25 |
| | $\hat{l}, \hat{u}$ | 140 | 33 | 234 | 20 | 234 | 20 |
| 3 | $l, u$ | 2 | 2 | 2 | 2 | 2 | 2 |
| | $\hat{l}, \hat{u}$ | 2 | 2 | 2 | 2 | 2 | 2 |
| 4 | $l, u$ | 52 | 27 | 52 | 27 | 52 | 27 |
| | $\hat{l}, \hat{u}$ | 52 | 27 | 52 | 27 | 52 | 27 |
| 5 | $l, u$ | 121 | 24 | 236 | 22 | 236 | 22 |
| | $\hat{l}, \hat{u}$ | 363 | 28 | 1312 | 47 | 1312 | 47 |
| 6 | $l, u$ | 28 | 15 | 28 | 15 | 28 | 15 |
| | $\hat{l}, \hat{u}$ | 28 | 15 | 28 | 15 | 28 | 15 |
| 7 | $l, u$ | 307 | 41 | 404 | 39 | 404 | 39 |
| | $\hat{l}, \hat{u}$ | 6 | 4 | 6 | 4 | 6 | 4 |
| 8 | $l, u$ | 7 | 5 | 7 | 5 | 7 | 5 |
| | $\hat{l}, \hat{u}$ | 7 | 5 | 7 | 5 | 7 | 5 |

(1) (Rosenbrock C=10): $f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$;

$l = (-2, -2)^T$, $u = (2,2)^T$; $\hat{l} = (-\infty, -\infty)^T$, $\hat{u} = (+\infty, +\infty)^T$.

(2) (Rosenbrock C=100): $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$;

$l = (-\infty, -\infty)^T$, $u = (+\infty, +\infty)^T$; $\hat{l} = (-2, -2)^T$, $\hat{u} = (2,2)^T$.

(3) (s201): $f(x) = 4(x_1 - 5)^2 + (x_2 - 6)^2$;

$l = (-\infty, -\infty)^T$, $u = (+\infty, +\infty)^T$; $\hat{l} = (-1, -1)^T$, $\hat{u} = (10,10)^T$.

(4) (s213): $f(x) = (10(x_1 - x_2)^2 + (x_1 - 1)^2)^4$;

$l = (-\infty, -\infty)^T$, $u = (+\infty, +\infty)^T$; $\hat{l} = (0,0)^T$, $\hat{u} = (3.5,1.5)^T$.

(5) (s312): $f(x) = (x_1^2 + 12x_2 - 1)^2 + (49x_1^2 + 49x_2^2 + 84x_1 + 232x_2 - 681)^2$;

$l = (-\infty, -\infty)^T$, $u = (+\infty, +\infty)^T$; $\hat{l} = (-1, -1)^T$, $\hat{u} = (4,4)^T$.

(6) (Powell): $f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$;

$l = (-2, -2, -2, -2)^T$, $u = (2,2,2,2)^T$; $\hat{l} = (-\infty, -\infty, -\infty, -\infty)^T$, $\hat{u} = (+\infty, +\infty, +\infty, +\infty)^T$.

(7) (Wood): $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 + 90(x_4 - x_3^2)^2 + (1 - x_3)^2$

$+ 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1)$;

$l = (-\infty, -\infty, -\infty, -\infty)^T$, $u = (+\infty, +\infty, +\infty, +\infty)^T$; $\hat{l} = (-3, -3, -3, -3)^T$, $\hat{u} = (1,1,1,1)^T$.

(8) (Hs05): $f(x) = \sin(x_1 + x_2) + (x_1 - x_2)^2 - 1.5x_1 + 2.5x_2 + 1$;

$l = (-1.5, -3)^T$, $u = (4,3)^T$; $\hat{l} = (-2, -2)^T$, $\hat{u} = (1,1)^T$.

## 参考文献：


[1] BRANCH M A, COLEMAN T F, LI Y. A subspace, interior and conjugate gradient method for large-scale bound-constrained minimization problems[J]. SIAM J Sci Comput, 1999, 21(1):1-23.

[2] BULTEAU J P, VIAL J PH. Curvilinear path and trust region in unconstrained optimization, a convergence analysis[J]. Mathematical Programming Study, 1987, 30:82-101.

[3] COLEMAN T F, LI Y. An interior trust region approach for nonlinear minimization subject to bounds[J]. SIAM J Optimization, 1996, 6(2):418-445.

[4] GRIPPO F LAMPARIELLO, LUCIDI S. A nonmonotonic line search technique for Newton's methods[J]. SIAM Journal on Numerical Analysis, 1986, 23:707-716.

[5] SCHITTKOWSKI K. More Test Examples for Nonlinear Programming Codes[J]. Berlin: Lecture Notes in Economics and Mathematical Systems, 282, Springer-Verlag, 1987.

[6] MORE J J, SORENSEN D C. Computing a trust region step[J]. SIAM Journal on Science and Statistical Computing, 1983, 4:553-572.

[7] HOCK W, SCHITTKOWSKI K. Test examples fro nonlinear programming code[J]. Berlin: Lecture Notes in Economics and Mathematical Systems 187, Springer-Verlag, 1981.

[8] ZHU D T. Curvilinear paths and trust region methods with nonmonotonic back tracking technique for unconstrained optimization[J]. J of Computational Mathematics, 2001, 19:241-258.

[9] ZHU D T. Nonmonotone preconditional curvilinear path algorithms for unconstrained optimization[J]. Numerical Mathematics, 2003, 12:99-110.


# A nonmonotonic interior point algorithm via optimal path for nonlinear optimization with bounds to variables


GUO Pei-hua, ZHU De-tong

(Mathematics and Sciences College, Shanghai Normal University, Shanghai 200234, China)



**Abstract**: This paper proposes a nonmonotonic interior point algorithm via optimal path for nonlinear optimization subject to bounds to variables. Based on the properties of the optimal path, the iterative direction is obtained by solving the quadratic model via the path. Using the nonmonotonic line search technique, we find an acceptable trial step length along this direction which is strictly feasible and makes the objective function monotonically decreasing. Theoretical analyses are given which prove that the proposed algorithm is globally convergent and has a local superlinear convergence rate under some reasonable conditions. The nonmonotonic criterion is used to speed up the convergence progress in the contours of objective function with large curvature. Numerical results indicate that the algorithm is effective in practice.



**Key words**: bounded variable; optimal path; interior point; nonmonotonic technique