第33卷第4期 上海师范大学学报(自然科学版) Vol. 33, No. 4
2004年12月 Journal of Shanghai Normal University (Natural Sciences) 2004, Dec.

仿射内点最优路径法解线性不等式约束的优化问题

王云娟，朱德通

(上海师范大学 数理信息学院，上海 200234)

摘 要：提供了仿射内点回代技术的最优路径法解线性不等式约束的非线性优化问题. 通过构造的最优路径得到搜索迭代方向，结合非单调内点回代线搜索技术获得可接受的步长因子，从而产生保证目标函数值非单调下降的严格内点可行迭代序列. 基于最优路径的良好性质，证明了在合理的假设条件下，算法不仅具有整体收敛性而且保持超线性收敛速率. 引入非单调技术能克服高度非线性的病态问题，加速收敛性进程. 数值计算结果表明了算法的有效性.

关键词：最优路径；不等式约束；内点法；仿射变换；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)04-0017-07

1 介 绍

本文考虑解带有线性不等式约束的非线性最优化问题

$$\begin{aligned} \min_{x\in \mathbf{R}^n} \quad & f(x) \\ \text{s.t.} \quad & Ax \geqslant b \end{aligned} \tag{1.1}$$

其中矩阵 $A^T=[a_1,\cdots,a_m]\in\mathbf{R}^{n\times m}$，向量 $a_i\in\mathbf{R}^n$，$i=1,\cdots,m$，$b=(b_1,\cdots,b_m)^T\in\mathbf{R}^m$. 可行解集记为 $\Omega\overset{\text{def}}{=}\{x\in\mathbf{R}^n\mid Ax\geqslant b\}$，并假设严格内点可行集 $\text{int}(\Omega)\overset{\text{def}}{=}\{x\in\mathbf{R}^n\mid Ax>b\}$ 非空. 这里 $f:\mathbf{R}^n\to\mathbf{R}$ 是在 Ω 上连续可微的非线性函数.

对于向量 $x_*\in\Omega$，问题(1.1)的一阶必要性条件为存在非负的乘子向量 $\lambda_*\in\mathbf{R}^m$ 使得

$$\text{diag}\{Ax_*-b\}\lambda_*=0,\ 和\ \nabla f(x_*)-A^T\lambda_*=0 \tag{1.2}$$

成立，这里 $\text{diag}\{Ax_*-b\}\overset{\text{def}}{=}\text{diag}\{a_1^Tx_*-b_1,a_2^Tx_*-b_2,\cdots,a_m^Tx_*-b_m\}$，称 x_* 为(1.1)的一个稳定点. 若对每个 $i=1,2,\cdots,m$，不等式 $a_i^Tx_*-b_i>0$ 与 $\lambda_*^i>0$ 中都至少有一个成立，即 $|a_i^Tx_*-b_i|+|\lambda_*^i|>0$ 成立，则称在 x_* 处严格互补性条件成立，其中 λ_*^i 表示 λ_* 的第 i 个分量.

许多文章都考虑用具有信赖域思想的逐次二次规划方法来求解问题(1.1)，现存的大部分的方法都是在严格内点可行集中进行迭代[1,5-7]. 信赖域方法的思想非常的直观与简单，但是，不等式约束造成信赖域子问题保持内点的计算困难性. 由于搜索方向必须是严格可行的，这就给计算带来了困难；同时，

收稿日期：2004-04-14

基金项目：上海高校科技发展基金资助项目(2000D12).

作者简介：王云娟(1980-)，女，上海师范大学数理信息学院硕士研究生；朱德通(1954-)，男，上海师范大学数理信息学院教授.

计算每一步新的迭代点工作量也很大. 在[3]中, Coleman & Li 通过引进仿射变换构造了一个近似二次函数与信赖域,从而提供双信赖域方法解(1.1).

为了避免频繁求解信赖域子问题,本文采用最优路径搜索与线搜索相结合的技术,通过最优路径搜索求得模型的迭代方向,然后沿此方向通过非单调线搜索获得步长因子,既能保证迭代点严格可行,又能保证迭代点使得目标函数值是非单调下降的. 基于最优路径的良好性质,在合理的假设条件下,可以证明该算法不仅具有整体收敛性,而且保持局部超线性收敛速率. 采用非单调技术可以更优的解决目标函数图像有多个峡谷形态的一类问题.

本文内容概述如下:在第 2 节,给出最优路径的构成和性质;在第 3 节,结合回代步、内点仿射变换、既约投影、最优路径搜索策略与非单调线搜索技术等技巧提供了算法;在第 4 节,证明算法具有弱整体收敛性;第 5 节中进一步证明在合理的假设下,算法还具有超线性收敛速率;最后,在第 6 节,通过给出一些数值实验结果来说明算法的有效性.

2 最优路径的构成及其性质

本节中,我们给出最优路径的构成并说明其具有一些良好的性态.

Coleman & Li 在[4]中给出了问题(1.1)的近似二次模型如下

$$\min_{(d;\hat{d})\in \mathbf{R}^{n+m}} \psi_k(d;\hat{d}) \overset{\text{def}}{=} \nabla f_k^T d + \frac{1}{2} d^T B_k d + \frac{1}{2} \hat{d}^T C_k \hat{d} \tag{2.1}$$

$$(S_k) \qquad \text{s.t.} \quad D_k^{\frac{1}{2}} \hat{d} = Ad \tag{2.2}$$

$$\|(d;\hat{d})\|_2 \leqslant \Delta_k. \tag{2.3}$$

其中 $D_k \overset{\text{def}}{=} \operatorname{diag}\{Ax_k - b\}$ 是一个仿射校正矩阵, $C_k \overset{\text{def}}{=} \operatorname{diag}\{|\lambda_k|\}$, $\nabla f_k = \nabla f(x_k)$, $d = x - x_k$, B_k 是 $\nabla^2 f_k$ 或其近似阵, $\nabla f_k^T d + \frac{1}{2} d^T B_k d$ 是 f 在 x_k 的一个局部二次近似, Δ_k 是信赖域半径.

为了避免频繁求解信赖域子问题,考虑将子问题 (S_k) 中的信赖域约束去掉,沿着某条曲线路径去搜索 $(d;\hat{d})$. 受到 Bulteau & Vial 在 [2] 中一般信赖域子问题最优路径的启迪,现在我们构造一条近似于仿射信赖子问题 (S_k) 的最优曲线路径来求解该子问题.

设 $(d_k;\hat{d}_k)$ 为 (S_k) 的一个解. 易见,去掉信赖域约束 (2.3) , (S_k) 的一阶必要性条件为存在乘子 λ_{k+1} 使得:

$$\begin{bmatrix} B_k & 0 \\ 0 & C_k \end{bmatrix} \begin{bmatrix} d_k \\ \hat{d}_k \end{bmatrix} = -\left\{ \begin{bmatrix} \nabla f_k \\ 0 \end{bmatrix} - \begin{bmatrix} A^T \\ -D_k^{1/2} \end{bmatrix} \lambda_{k+1} \right\}. \tag{2.4}$$

对于 Lagrangian 乘子 λ_k ,我们采用最小二乘法来求解,即

$$\begin{bmatrix} A^T \\ -D_k^{1/2} \end{bmatrix} \lambda_k \overset{\text{L.S.}}{=} \begin{bmatrix} \nabla f_k \\ 0 \end{bmatrix}. \tag{2.5}$$

为方便起见,记 $M_k \overset{\text{def}}{=} \begin{bmatrix} B_k & 0 \\ 0 & C_k \end{bmatrix}$. 同时,类似于 [4] ,定义投影梯度为 $g_k \overset{\text{def}}{=} g(x_k) \overset{\text{def}}{=} -(\nabla f_k - A^T \lambda_{k+1})$,并将 (S_k) 中目标函数的增广梯度定义为 $\tilde{g}_k \overset{\text{def}}{=} \begin{bmatrix} \nabla f_k \\ 0 \end{bmatrix} - \begin{bmatrix} A^T \\ -D_k^{1/2} \end{bmatrix} \lambda_{k+1}$. 令 P_k 为到 $[A, -D_k]$ 的零空间的正交投影,则

$$\nabla f_k^T g_k = -\left\| P_k \begin{bmatrix} \nabla f(x_k) \\ 0 \end{bmatrix} \right\|_2^2 = -\left(\|\nabla f_k - A^T \lambda_{k+1}\|_2^2 + \|D_k^{1/2} \lambda_{k+1}\|_2^2 \right) = -\|\tilde{g}_k\|_2^2. \tag{2.6}$$

显然，从(2.6)式可知，由简约值 $\nabla f_k^T g_k$ 的下降量度量的 $\psi_k(d;\hat{d})$ 的一个充分下降将导致满足互补性条件：

$$\lim_{k\to\infty}\|\nabla f_k - A^T\lambda_{k+1}\|_2 = 0 \text{ 和 } \lim_{k\to\infty}\|D_k^{1/2}\lambda_{k+1}\|_2 = 0. \tag{2.7}$$

因为 M_k 是一个对称矩阵，考虑 M_k 的特征分解，则它的 $n+m$ 个特征值全为实数．由于 B_k 是对称矩阵，所以存在单位正交矩阵 U_k 和对角矩阵 Λ_k，使得 $B_k = U_k\Lambda_k U_k^T$．由 M_k 的定义可知可以扩充单位正交矩阵 U_k 和对角阵 Λ_k，使得 $M_k = \begin{bmatrix} U_k & \\ & I\end{bmatrix}\times\begin{bmatrix}\Lambda_k & \\ & C_k\end{bmatrix}\times\begin{bmatrix}U_k^T & \\ & I\end{bmatrix}$，所以 M_k 的 $n+m$ 个特征值 $\phi_1,\phi_2,\cdots,\phi_{n+m}$ 中有 n 个为 Λ_k 的对角元素，有 m 个为 C_k 的对角元素．记 $u^1,u^2,\cdots,u^{n+m}$ 为与 $\phi_1,\phi_2,\cdots,\phi_{n+m}$ 相对应的标准正交特征向量．由 $\tilde{g}_k$ 的定义和(2.5)式可知 $\tilde{g}_k$ 属于 $[A,\ -D_k^{1/2}]$ 的零空间．假设 $[A,\ -D_k^{1/2}]$ 是行满秩的，则 $[A,\ -D_k^{1/2}]$ 的零空间的基是 n 维的．因此全空间的一组标准正交基 $u^1,u^2,\cdots,u^{n+m}$ 中有且仅有 n 个向量属于 $N[A,\ -D_k^{1/2}]$，不妨记为 $\tilde{u}^1,\tilde{u}^2,\cdots,\tilde{u}^n$．相应地，假设 $\tilde{\phi}_1,\tilde{\phi}_2,\cdots,\tilde{\phi}_n$ 为与之对应的特征值，不失一般性，假设 $\tilde{\phi}_1\le\tilde{\phi}_2\le\cdots\le\tilde{\phi}_n$．根据 $\tilde{\phi}_i>0$、$\tilde{\phi}_i<0$ 和 $\tilde{\phi}_i=0$，将集 $\{1,2,\cdots,n\}$ 分成 $\mathscr{T}^+$，$\mathscr{T}^-$ 和 $\mathscr{N}$.

2.1 最优路径

基于 M_k 的特征分解，最优路径 $\Gamma(\tau)$ 可以表示为

$$\Gamma(\tau) = \Gamma_1(t_1(\tau)) + \Gamma_2(t_2(\tau)), \tag{2.8}$$

其中

$$\Gamma_1(t_1(\tau)) = -\Big[\sum_{i\in\mathscr{T}}\frac{t_1(\tau)}{\tilde{\phi}_i t_1(\tau)+1}\tilde{g}_k^i\tilde{u}^i + t_1(\tau)\sum_{i\in\mathscr{N}}\tilde{g}_k^i\tilde{u}^i\Big],$$

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

和

$$t_1(\tau) = \begin{cases}\tau & \text{如果 } \tau<\dfrac{1}{T} \\[2mm] \dfrac{1}{T} & \text{如果 } \tau\ge\dfrac{1}{T}\end{cases},$$

$$t_2(\tau) = \begin{cases}0 & \text{如果 } \tau<\dfrac{1}{T} \\[2mm] \tau-\dfrac{1}{T} & \text{如果 } \tau\ge\dfrac{1}{T}\end{cases},$$

这里 $\mathscr{T}=\{\tilde{\phi}_i\ne 0, i=1,2,\cdots,n\}$，$\mathscr{N}=\{\tilde{\phi}_i=0, i=1,2,\cdots,n\}$，$\tilde{g}_k^i=(\tilde{g}_k)^T\tilde{u}^i$，$i=1,2,\cdots,n$，$\tilde{g}_k=\sum_{i=1}^n\tilde{g}_k^i\tilde{u}^i$，$T=\max\{0,-\tilde{\phi}_1\}$，若 $T=0$，则定义 $\frac{1}{T}=+\infty$．注意到，只有当 M_k 不定且对所有的满足 $\tilde{\phi}_i=\tilde{\phi}_1<0$ 的 $i\in\{1,2,\cdots,n\}$ 均有 $\tilde{g}_k^i=0$ 时，$\Gamma_2(t_2(\tau))$ 才有定义，而上面这种情况称为困难情况；在其他情况下，$\Gamma(\tau)$ 仅定义在 $0\le\tau<\frac{1}{T}$ 上，即 $\Gamma(\tau)=\Gamma_1(t_1(\tau))$.

2.2 最优路径的性质

基于[8]中引理2.3和引理2.5等的证明，可以得到如下的最优路径的性质（由于篇幅限制，文中以下的证明过程都省略）.

引理2.1 设步 $\begin{bmatrix}\delta_k(\tau)\\ \hat{\delta}_k(\tau)\end{bmatrix}=\Gamma_k(\tau)$ 是由最优路得到的．则

(1) $\begin{bmatrix}\delta_k(\tau)\\ \hat{\delta}_k(\tau)\end{bmatrix}$ 属于 $[A,\ -D_k^{1/2}]$ 的零空间，即有 $[A,\ -D_k^{1/2}]\begin{bmatrix}\delta_k(\tau)\\ \hat{\delta}_k(\tau)\end{bmatrix}=0$.

(2) 路的范数函数 $\|\Gamma_k(\tau)\|$ 在 $\tau \in (0,+\infty)$ 是单调递增的.

(3) 目标函数的梯度 ∇f_k 与由最优路径得到的步 $\delta_k(\tau)$ 的关系函数 $\hat{\psi}_k(\tau) \stackrel{\text{def}}{=} \nabla f_k^T\delta_k(\tau)$ 在 $0<\tau<\frac{1}{T}$ 上是单调递减的且满足一个充分下降方向,即

$$\hat{\psi}_k(\tau) \leqslant -\frac{1}{2}\min\{\frac{1}{\|M_k\|},\tau,\frac{1}{T}\}|-\nabla f_k^T g_k|. \tag{2.9}$$

(4) 进一步,当 $\tau \to 0$ 时有

$$\nabla f_k^T \frac{\mathrm{d}\delta_k(\tau)}{\mathrm{d}\tau} = \frac{d\hat{\psi}_k(\tau)}{d\tau} \to -|-\nabla f_k^T g_k|. \tag{2.10}$$

(5) 如果 M_k 在 $[A,-D_k^{1/2}]$ 的零空间上正定,则

$$\lim_{\tau\to\infty} M_k\Gamma_k(\tau) = -\tilde{g}_k. \tag{2.11}$$

(6) 函数

$$\psi_k(\delta_k(\tau);\hat{\delta}_k(\tau)) \stackrel{\text{def}}{=} \nabla f_k^T\delta_k(\tau) + \frac{1}{2}\delta_k(\tau)^T B_k\delta_k(\tau) + \frac{1}{2}\hat{\delta}_k(\tau)^T C_k\hat{\delta}_k(\tau) \tag{2.12}$$

在 $\tau \in (0,+\infty)$ 上单调递减,并且对任意的 $\tau \in (0,+\infty)$,有

$$-\psi_k(\delta_k(\tau);\hat{\delta}_k(\tau)) \geqslant \frac{1}{4}\min\{\frac{1}{\|M_k\|},\tau,\frac{1}{T}\}|-\nabla f_k^T g_k|. \tag{2.13}$$

(7) 对于任意的 $\tau \in (0,+\infty)$,有

$$(M_k+\nu_k I)\begin{bmatrix}\delta_k(\tau)\\ \hat{\delta}_k(\tau)\end{bmatrix} = -\tilde{g}_k, \tag{2.14}$$

其中 $\nu_k \geqslant 0$ 定义如下:

$$\nu_k = \frac{1}{\tau},\ 如果\ \tau < \frac{1}{T}; \tag{2.15}$$

$$\nu_k = T,\ t_2(\tau) = \tau - \frac{1}{T},\ 如果\ \tau \geqslant \frac{1}{T}. \tag{2.16}$$

这里 $T=\max\{0,-\tilde{\phi}_1\}$,而且矩阵 $M_k+\nu_k I$ 在零空间 $N[A,-D_k^{1/2}]$ 上至少是半正定的.

3 算 法

本节中将给出求解子问题 (S_k) 的内点仿射变换非单调最优路径算法.

初始步: 选取参数 $\beta \in (0,\frac{1}{2})$, $w,\tilde{w},\gamma \in (0,1)$ 和一个正整数 Q. 令 $m(0)=0$. 选取一个对称矩阵 B_0,给定一个初始可行点 $x_0 \in \mathbf{R}^n$. 令 $k=0$,进入主步.

主步: (1) 计算 $f_k=f(x_k)$, $\nabla f(x_k)$, $D_k=\mathrm{diag}\{Ax_k-b\}$ 和 $C_k=\mathrm{diag}\{|\lambda_k|\}$. 选取一个对称矩阵 $B_k \approx \nabla^2 f(x_k)$. 由(2.5)式计算最小二乘 Lagrange 乘子近似 λ_k.

(2) 如果 $\|g_k\|+\|D_k^{1/2}\lambda_k\| \leqslant \varepsilon$,则停止迭代并取 x_k 为近似解.

(3) 计算 B_k 的特征值与标准正交特征向量. 构造一条最优路径 $\Gamma_k(\tau)$,得到步 $\delta_k(\tau)$,其中

$$\begin{bmatrix}\delta_k(\tau)\\ \hat{\delta}_k(\tau)\end{bmatrix} = \Gamma_k(\tau).$$

若 M_k 正定,选取 $\tau=\infty, w^{-n}, w^{-(n-1)},\cdots$;否则,取一个充分大的 N_0,选取 $\tau = w^{-N_0}, w^{-(N_0-1)},\cdots$, 直到下面的不等式成立

$$f(x_k)-f(x_k+\delta_k(\tau)) \geqslant -\gamma\psi_k(\delta_k(\tau);\hat{\delta}_k(\tau)). \tag{3.1}$$

这里 $\psi_k(\delta_k(\tau);\hat{\delta}_k(\tau))$ 由(2.12)式定义. 记满足上式的第一个 τ 的值为 τ_k,即获得 $\delta_k(\tau_k)$.

(4) 选取 $\alpha_k = 1, \bar{w}, \bar{w}^2, \cdots$，直到下面的不等式成立

$$f(x_k + \alpha_k\delta_k(\tau_k)) \leqslant f(x_{l(k)}) + \alpha_k\beta\nabla f_k^T\delta_k(\tau_k), \tag{3.2}$$

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

其中 $f(x_{l(k)}) = \max\limits_{0\leqslant j\leqslant m(k)}\{f(x_{k-j})\}$.

(5) 置

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

其中 $\theta_k \in (\theta_0, 1]$，$0 < \theta_0 < 1$ 且 $\theta_k - 1 = O(\|\delta_k(\tau_k)\|^2)$.

令

$$x_{k+1} = x_k + h_k \tag{3.4}$$

(6) 取 $m(k+1) = \min\{m(k)+1, Q\}$，校正 B_k 至 B_{k+1}. 令 $k \leftarrow k+1$，转步 2.

4 整体收敛性

为了证明算法具有整体收敛性，需要一些常见的假设.

假设 A0　给定一个初始点 $x_0 \in \mathrm{int}(\Omega)$；假设水平集 $L(x_0)$ 是紧集，由算法产生的序列 $\{x_k\}$ 仍在紧集 $L(x_0)$ 中，其中 $L(x_0) \overset{\text{def}}{=} \{x \in \mathbf{R}^n \mid f(x) \leqslant f(x_0), Ax - b \geqslant 0\}$；矩阵 $[A, -D_k^{1/2}]$ 对任意 $x \in L(x_0)$ 都是行满秩的.

假设 A1　存在正数 κ_f 和 κ_g，使得对于任意的 $x \in L(x_0)$，均有 $\|\nabla f(x)\| \leqslant \kappa_f$ 和 $\|g(x)\| \leqslant \kappa_g$ 成立.

假设 A2　$\|B_k\|$ 上有界，即存在某个正数 κ_B，使得对所有的 k，都有 $\|B_k\| \leqslant \kappa_B$.

假设 A3　$\|\nabla^2 f(x)\|$ 上有界，即对任意的 $x \in L(x_0)$，存在某个正数 κ_G，使得 $\|\nabla^2 f(x)\| \leqslant \kappa_G$.

由假设 A0 可知，$\lambda(x)$ 在 $x \in L(x_0)$ 上连续且有界；再由假设 A2，立即可知 $\|M_k\|$ 也是上有界的，设 $\kappa_M > 0$ 为它的一个上界.

为了使算法合理的适定和收敛，我们将证明下述引理与定理.

引理 4.1　假设 $f: \mathbf{R}^n \to \mathbf{R}$ 是二次连续可微的函数，对于给定的初始可行点 $x_0 \in \mathbf{R}^n$，序列 $\{x_k\}$ 由算法产生. 则在第 k 步，或者 $|-\nabla f_k^T g_k| = 0$，或者经过有限次的缩小 τ 之后，存在 τ_k 满足下面的式子

$$f(x_k) - f(x_k + \delta_k(\tau)) \geqslant -\gamma\psi_k(\delta_k(\tau); \hat{\delta}_k(\tau)) \quad \gamma \in (0,1). \tag{4.1}$$

定理 4.2　设序列 $\{x_k\}$ 由算法产生. 假设 A0 – A3 都成立且在 $\{x_k\}$ 的任一极限点 x_* 处，问题 (1.1) 的严格互补性条件成立. 则

$$\liminf_{k\to\infty} |\nabla f_k^T g_k| = 0. \tag{4.2}$$

5 局部收敛速率

定理 4.2 表明 $\{x_k\}$ 至少有一个极限点是稳定点. 在这一节，我们推广这一结论到一个更强的结果并说明它具有局部超线性收敛速率，但这需要进一步给出下列的一些假设条件.

假设 A4　问题 (1.1) 的解 x_* 满足强二阶充分性条件，即存在 $\alpha > 0$ 以使

$$p^T(Z_*^T H_* Z_*)p \geqslant \alpha\|p\|^2, \ \forall\, p \tag{5.1}$$

其中 Z_* 是 $[A, -D_*^{\frac{1}{2}}]$ 的零空间的标准正交基张成的矩阵，而 $H_* \overset{\text{def}}{=} \begin{bmatrix}\nabla^2 f(x_*) & 0 \\ 0 & C_*\end{bmatrix}$.

假设 A5
$$\lim_{k\to\infty}\frac{\|(B_k-\nabla^2 f_k)\delta_k(\tau_k)\|}{\|\delta_k(\tau_k)\|}=0. \tag{5.2}$$

引理 5.1 假设 A0 ~ A5 都成立. 设 $\{x_k\}$ 为由算法产生的序列. 若在 $\{x_k\}$ 的任意极限点 x_* 处问题(1.1)的严格互补性条件总成立,则 $\delta_k(\tau_k)\to 0$.

定理 5.2 假设 A0 ~ A5 都成立. 设 $\{x_k\}$ 为由算法产生的序列. 若在 $\{x_k\}$ 的任意极限点 x_* 处问题(1.1)的严格互补性条件总成立,则
$$\lim_{k\to\infty}|\nabla f_k^T g_k|=0. \tag{5.3}$$

现在讨论算法的局部收敛速率. 为此,要证明对充分大的 k ,步长 $\alpha_k\equiv 1$,且 $\delta_k(\infty)=B_k^{-1}g_k$.

定理 5.3 假设 A0 ~ A5 都成立, $\{x_k\}$ 为由算法产生的序列, x_* 为 $\{x_k\}$ 的一个局部极小值点,则 $\{x_k\}$ 超线性收敛到 x_* ,即
$$\lim_{k\to\infty}\frac{\|x_{k+1}-x_*\|}{\|x_k-x_*\|}=0. \tag{5.4}$$
定理 5.3 表明算法的局部收敛速率取决于目标函数在 x_* 处的 Hessian 阵以及步 $\delta_k(\infty)$ 的局部收敛速率. 如果 $\delta_k(\infty)$ 成为零子空间 $N[A, -D_k^{1/2}]$ 的拟牛顿步,则算法产生的序列 $\{x_k\}$ 超线性收敛到 x_* . 进一步,在相同的条件下,可以证明[4]中的局部收敛速率的结论.

6 数值结果

我们已经将所提供的仿射内点最优路径算法,利用数学软件 Matlab 编程进行数值实现. 我们选取参数如下: $\epsilon=10^{-6}$, $\beta=0.4$, $w=0.5$, $\bar{w}=0.5$, $\gamma=0.02$,非单调参数 $Q=0$, 5 和 10 .

应用了 5 个标准数值测试题. 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. 其迭代次数与梯度函数取值次数一样,这里省略. 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(表 1),可以看到非单调技术在求解 Rosenbrock 等问题时,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. 当 $Q=0$ 时,该算法就是通常的算法.

表 1 数值结果

编号	初始点	$Q=0$		$Q=5$		$Q=10$	
		NF	NG	NF	NG	NF	NG
1	x_0	339	33	138	21	138	21
2	x_0	12	7	12	7	12	7
3	x_0	95	96	95	96	95	96
4	x_0	75	22	102	22	102	22
5	x_0	48	10	47	10	47	10

(1) $\min f(x)=100(x_2-x_1^2)^2+(1-x_1)^2$;

s.t. $\frac{1}{3}x_1+x_2\geq 0.1$; $-\frac{1}{3}x_1+x_2\geq -0.1$.

其最优解为 $x_*=(1,1)^T, f(x_*)=0$.

(2) $\min f(x)=0.1(12+x_1^2+(1+x_2^2)/(x_1^2)+(x_1^2x_2^2+100)/(x_1^4x_2^4))$;

s.t. $x_1\geq 1$; $x_2\geq 1$; $-x_1\geq -3$; $-x_2\geq -3$.

其最优解为 $x_*=(1.743, 2.030)^T$, $f(x_*)=1.74415$.

(3) $\min f(x)=-(9-(x_1-3)^2)((x_2^3)/(27t))$; $t=\sqrt{3}$,

s.t. $\frac{1}{t}x_1 - x_2 \geqslant 0;\ x_1 + t \geqslant 0;\ -x_1 - tx_2 \geqslant -6;\ x_1 \geqslant 0;\ x_2 \geqslant 0.$

其最优解为 $x_* = (3,\sqrt{3})^T, f(x_*) = -1.$

(4) $\min f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2;$

s.t. $x_1 \geqslant -2;\ x_2 \geqslant -2;\ -x_1 \geqslant -2;\ -x_2 \geqslant -2.$

其最优解为 $x_* = (1,1)^T, f(x_*) = 0.$

(5) $\min f(x) = (x_1^2 + x_2^2 + x_1x_2)^2 + (\sin(x_1))^2 + (\cos(x_2))^2;$

s.t. $x_1 \geqslant -\infty;\ x_2 \geqslant -\infty.$

其最优解为 $x_* = (-0.1554, 0.6946), f(x_*) = 0.773199.$

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 P. 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. Combining trust region and affine scaling for linearly constrained nonconvex minimization[J]. In: Advances in Nonlinear Programming, Yuan Y,ed,Kluwer Academic Publishers,1998: 219-250.

[4] COLEMAN T F, LI Y. A trust region and affine scaling interior point method for nonconvex minimization with linear inequality constraints[J]. Math Program Ser A, 2000, 88:1-31.

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

[6] DENNIS J E JR, SCHNABLE R B. Numerical methods for unconstrained optimization and non – linear equations[M]. Prentice Hall,New Jersey,1983.

[7] DIKIN I I. Iterative solution of problems of linear and quadratic programming[J]. Soviet Math Dokl, 1967, 8:18-35.

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

Interior affine scaling curvilinear path algorithm for nonlinear optimizations subject to linear inequality constraints

WANG Yun-juan, ZHU De-tong

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

Abstract: We propose an interior affine scaling optimal path algorithm with nonmonotonic interior back – tracking technique for nonlinear optimization subject to linear inequality constraints. Using both optimal path search strategy and line search technique, the quadratic model at each iteration generates a backtracking step to obtain a new accepted step. The global convergence and fast local convergence rate of the proposed algorithm are established 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 useful and effective in practice.

Key words: optimal path; linear inequality constraint; interior point method; affine scaling ; nonmonotonic technique