系 统 科 学 与 数 学
J. Sys. Sci. & Math. Scis.
27(6) (2007, 12), 820–836

线性等式约束优化的既约预条件共轭梯度路径法*

林 涛
(上海应用技术学院, 上海 200235)
朱 德 通
(上海师范大学商学院, 上海 200234)

摘要 采用既约预条件共轭梯度路径结合非单调技术解线性等式约束的非线性优化问题. 基于广义消去法将原问题转化为等式约束矩阵的零空间中的一个无约束优化问题, 通过一个增广系统获得既约预条件方程, 并构造共轭梯度路径解二次模型, 从而获得搜索方向和迭代步长. 基于共轭梯度路径的良好性质, 在合理的假设条件下, 证明了算法不仅具有整体收敛性, 而且保持快速的超线性收敛速率. 进一步, 数值计算表明了算法的可行性和有效性.

关键词 共轭梯度路径, 既约预条件, 非单调技术.

MR(2000) 主题分类号 90C30, 65K05

1 引 言

本文研究带有线性等式约束的非线性优化问题

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & A\,x=b, \end{aligned} \tag{1.1}$$

其中 $f:\Re^n\longrightarrow\Re$ 是光滑的非线性函数, 矩阵 $A\in\Re^{m\times n}$, 向量 $b\in\Re^m$, 变量 $x\in\Re^n$, 定义 $\Omega\stackrel{\rm def}{=}\{x|Ax=b\}$ 表示问题 (1.1) 的可行集. Bulteau 和 Vial 在文 [1] 中构造共轭梯度路径解无约束优化问题的思想启迪我们用其来解等式约束优化问题. 利用广义消去法将问题 (1.1) 转化为等式约束矩阵的零空间中的一个无约束优化问题. 由于既约矩阵不保持原矩阵的充分稀疏, 本文通过一个增广系统解既约预条件方程. 由构造既约预条件共轭梯度路径, 结合非单调技术的算法不仅具有整体收敛性, 而且保持快速的超线性收敛速率.

本文概要如下: 第 2 节构造了既约预条件共轭梯度路径并给出其性质; 第 3 节结合既约预条件、共轭梯度路径搜索策略与非单调技术等技巧提供了算法; 第 4 节证明了算法的整体

* 国家自然科学基金项目 (10471094) 和上海市重点学科 (T0401) 项目.
收稿日期: 2005-11-22.

弱收敛性和强收敛性；第 5 节在合理的假设条件下，证明了局部超线性收敛速率；第 6 节，给出了算法的数值实验结果，表明其算法的可行性和有效性.

2 既约预条件共轭梯度路径的构造及其性质

本节将给出既约预条件共轭梯度路径的构造，以及其具有的一些良好性态.

2.1 既约预条件共轭梯度路径的构造

设问题 (1.1) 的拉格朗日函数为

$$L(x,\lambda)=f(x)+\mu^{\mathrm{T}}(Ax-b),$$

其中 $\mu\in\Re^m$ 是拉格朗日乘子. 对于局部最优解 x_*, 问题 (1.1) 的一阶最优性必要条件为存在向量 $\mu_*\in\Re^m$, 使得下列方程成立

$$\begin{aligned} g_*+A^{\mathrm{T}}\mu_*&=0,\\ Ax_*&=b, \end{aligned} \tag{2.1}$$

这里 $g_*\stackrel{\mathrm{def}}{=}\nabla f(x_*)$. 设 (x_k,μ_k) 是牛顿迭代的第 k 次迭代，则方程 (2.1) 在第 k 次迭代中的牛顿步为

$$\begin{bmatrix}\nabla^2 f_k & A^{\mathrm{T}}\\ A & 0\end{bmatrix}\begin{bmatrix}p_k\\ \Delta\mu_k\end{bmatrix}=-\begin{bmatrix}\nabla f_k+A^{\mathrm{T}}\mu_k\\ 0\end{bmatrix}. \tag{2.2}$$

第 $k+1$ 次迭代点定义为

$$\begin{bmatrix}x_{k+1}\\ \mu_{k+1}\end{bmatrix}=\begin{bmatrix}x_k\\ \mu_k\end{bmatrix}+\begin{bmatrix}p_k\\ \Delta\mu_k\end{bmatrix}.$$

对于拉格朗日乘子 μ_k, 可采用最小二乘法求解，即

$$A^{\mathrm{T}}\mu_{k+1}\stackrel{\mathrm{L.S.}}{=}-\nabla f_k.$$

基于牛顿步方程 (2.2), 得到问题 (1.1) 的二次规划子问题 (S_k) 为

$$\begin{cases}\min & \varphi_k(p)\stackrel{\mathrm{def}}{=}f_k+g_k^{\mathrm{T}}p+\dfrac{1}{2}p^{\mathrm{T}}B_kp\\ \mathrm{s.t.} & Ap=0,\end{cases} \tag{2.3}$$

其中 $g_k=\nabla f(x_k)$, $p=x-x_k$, B_k 是 $\nabla^2 f_k$ 或是其近似矩阵， $\varphi_k(p)$ 是 f 的局部二次近似.

假设 A 是行满秩的, 利用广义消去法可得 $p=Z^{\mathrm{T}}p^z$, 其中 Z 的行向量形成零空间 $\mathcal{N}(A)$ 的一组正交基. 由此子问题 (2.3) 等价于如下无约束优化问题

$$\min\quad \varphi_k(p^z)\stackrel{\mathrm{def}}{=}f_k+(Zg_k)^{\mathrm{T}}p^z+\frac{1}{2}(p^z)^{\mathrm{T}}ZB_kZ^{\mathrm{T}}p^z. \tag{2.4}$$

令 $\overline{g}_k\stackrel{\mathrm{def}}{=}Zg_k$, $\overline{H}_k\stackrel{\mathrm{def}}{=}ZB_kZ^{\mathrm{T}}$, (2.4) 式可以重新改写为

$$\min\quad \varphi_k(p^z)\stackrel{\mathrm{def}}{=}f_k+\overline{g}_k^{\mathrm{T}}p^z+\frac{1}{2}(p^z)^{\mathrm{T}}\overline{H}_kp^z. \tag{2.5}$$

将既约共轭方向法应用于 (2.5), 从 $v_1=0$ 开始，得到点列 $v_1,v_2,\cdots,v_{q+1}$ 和共轭方向序列 $d_1,\ d_2,\cdots,\ d_{q+1}$, 使其满足

$$\overline{H}_k\overline{s}_{i+1}=\overline{r}_{i+1},\quad i=1,2,\cdots,q, \tag{2.6}$$

$$d_{i+1}=-\overline{s}_{i+1}+\beta_i d_i,\quad i=1,2,\cdots,q, \tag{2.7}$$

$$v_{i+1}=v_i+\lambda_i d_i,\quad i=1,2,\cdots,q, \tag{2.8}$$

$$d_i^{\mathrm{T}}\overline{H}_k d_i>0,\quad i=1,2,\cdots,q. \tag{2.9}$$

这里

$$\overline{r}_{i+1}=\nabla\varphi_k(v_{i+1})=\overline{r}_i+\lambda_i\overline{H}_k d_i,\quad \beta_i=\frac{\overline{s}_{i+1}^{\mathrm{T}}\overline{H}_k d_i}{d_i^{\mathrm{T}}\overline{H}_k d_i}>0,\quad \lambda_i=\frac{\overline{r}_i^{\mathrm{T}}\overline{s}_i}{d_i^{\mathrm{T}}\overline{H}_k d_i}>0. \tag{2.10}$$

其中 q 表示 $\overline{H}_k$ 相异特征值的个数．当 $\overline{r}_{i+1}=0$ 或者 $\overline{r}_{i+1}\neq 0$ $(d_{i+1}\neq 0)$, 但 $d_{i+1}^{\mathrm{T}}\overline{H}_k d_{i+1}\leq 0$ 时，上述过程终止．

共轭梯度路径 $\mathit{\Gamma}_k^z(\tau)$ 构造如下

$$\mathit{\Gamma}_k^z(\tau)=\sum_{i=1}^{q}t_i(\tau)d_i-t_{q+1}(\tau)d_{q+1}, \tag{2.11}$$

和

$$t_i(\tau)=\min\Big\{\lambda_i,\max\Big\{0,\tau-\sum_{j=1}^{i-1}\lambda_j\Big\}\Big\}. \tag{2.12}$$

其中 d_i 和 λ_i 由 (2.7) 和 (2.10) 给出．当 $i=1$ 时，令 $\sum\limits_{j=1}^{i-1}\lambda_j=0$.

对于既约预条件方程 (2.6), 线性系统传统的预条件策略依赖于矩阵 $\overline{H}_k=ZB_kZ^{\mathrm{T}}$. 然而尽管矩阵 B_k 很稀疏，在实际应用中既约矩阵 $\overline{H}_k$ 也不一定能保持其充分稀疏．类似于文 [2] 中方法，本文采用增广系统求解 (2.6).

因为 A 是行满秩的， Z 的行向量形成零空间 $\mathcal{N}(A)$ 的一组正交基，取

$$Z=\begin{bmatrix}-A_1^{-1}A_2\\ I_{n-m}\end{bmatrix}^{\mathrm{T}}P,$$

其中 P 是置换矩阵， $AP=(A_1,A_2)$, A_1 是非奇异的．为了简单起见，本文中假设 $P=I_n$. 为了求解 (2.6) 式中的 $\overline{s}_{i+1}$, 定义如下增广系统

$$\begin{bmatrix}\widetilde{H}_k & \widetilde{A}^{\mathrm{T}}\\ \widetilde{A} & 0\end{bmatrix}\begin{bmatrix}s_{i+1}\\ u_{i+1}\end{bmatrix}=\begin{bmatrix}\widetilde{r}_{i+1}\\ 0\end{bmatrix}, \tag{2.13}$$

其中 $\widetilde{H}_k$ 是 B_k 的稀疏对称正定近似， $\widetilde{A}=(\widetilde{A}_1,\widetilde{A}_2)$, 其中 $\widetilde{A}_1$ 是 A_1 的非奇异近似， $\widetilde{A}_2$ 是 A_2 的近似，以及

$$\widetilde{r}_{i+1}=\begin{bmatrix}0^m\\ \overline{r}_{i+1}^{\,n-m}\end{bmatrix},$$

同时在 (2.13) 中 $s_{i+1}=\widetilde{Z}^{\mathrm{T}}\overline{s}_{i+1}$, 其中 $\widetilde{Z}$ 的行向量形成零空间 $\mathcal{N}(\widetilde{A})$ 的一组基，即

$$\widetilde{Z}=\begin{bmatrix}-\widetilde{A}_1^{-1}\widetilde{A}_2\\ I_{n-m}\end{bmatrix}^{\mathrm{T}}.$$

易见通过方程 (2.13) 求解 s_{i+1} 等价于求解如下方程

$$(\widetilde{Z}\widetilde{H}_k\widetilde{Z}^{\mathrm{T}})\overline{s}_{i+1}=\widetilde{Z}\widetilde{r}_{i+1}=\begin{bmatrix}-\widetilde{A}_1^{-1}\widetilde{A}_2\\ I_{n-m}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}0^m\\ \overline{r}_{i+1}^{n-m}\end{bmatrix}=\overline{r}_{i+1}. \tag{2.14}$$

显然逆矩阵 $(\widetilde{Z}\widetilde{H}_k\widetilde{Z}^{\mathrm{T}})^{-1}$ 可看成是 $\overline{H}_k$ 的近似逆.

因为 $\widetilde{Z}$ 是矩阵 $\widetilde{A}$ 的零空间的一组基所构成的矩阵，所以 $\overline{s}_{i+1}=(0,I_{n-m})s_{i+1}$, 因此可以不必求解 $\widetilde{Z}$. 对于方程 (2.13) 中增广矩阵

$$\widetilde{M}_k=\begin{bmatrix}\widetilde{H}_k & \widetilde{A}^{\mathrm{T}}\\ \widetilde{A} & 0\end{bmatrix}$$

利用 Matlab 来求解 (见文 [2]).

2.2 既约预条件共轭梯度路径的性质

下面给出既约预条件共轭梯度路径的一些性质，并加以证明.

引理 2.1 设 d_i 是由 (2.7)–(2.10) 得到的方向序列，则对 $1\le i,j\le l\le q+1$, 下列结论成立

$$\overline{r}_i^{\mathrm{T}}d_j=0,\quad 1\le j<i\le l\le q+1, \tag{2.15}$$

$$d_i^{\mathrm{T}}\overline{H}_kd_j=0,\quad i\ne j, \tag{2.16}$$

$$\overline{r}_i^{\mathrm{T}}M_k^{-1}\overline{r}_j=0,\quad i\ne j, \tag{2.17}$$

$$d_i^{\mathrm{T}}\overline{r}_i=-\overline{r}_i^{\mathrm{T}}M_k^{-1}\overline{r}_i,\quad i=1,2,\cdots,q+1, \tag{2.18}$$

$$d_i^{\mathrm{T}}M_kd_j>0,\quad i\ne j. \tag{2.19}$$

其中 $M_k=\widetilde{Z}\widetilde{H}_k\widetilde{Z}^{\mathrm{T}}$.

证 使用数学归纳法证明 (2.15)–(2.19).

当 $l=1$ 时， (2.18) 成立；当 $l=2$ 时， (2.15), (2.16), (2.17) 和 (2.19) 成立. 假设以上关系式对 l 成立，下面证明对于 $l+1$, 这些关系式也成立.

因为

$$\overline{r}_{l+1}=\overline{H}_kv_{l+1}+\overline{g}_k=\overline{H}_k(v_l+\lambda_ld_l)+\overline{g}_k=\overline{r}_l+\lambda_l\overline{H}_kd_l. \tag{2.20}$$

利用 (2.20) 式，有

$$\overline{r}_{l+1}^{\mathrm{T}}d_j=(\overline{r}_l+\lambda_l\overline{H}_kd_l)^{\mathrm{T}}d_j=\overline{r}_l^{\mathrm{T}}d_j+\lambda_ld_l^{\mathrm{T}}\overline{H}_kd_j. \tag{2.21}$$

当 $1\le j\le l-1$ 时，由归纳法假设可知 (2.21) 右端为零，又

$$\overline{r}_{l+1}^{\mathrm{T}}d_l=\overline{r}_l^{\mathrm{T}}d_l+\frac{\overline{r}_l\overline{s}_l}{d_l^{\mathrm{T}}\overline{H}_kd_l}d_l^{\mathrm{T}}\overline{H}_kd_l=\overline{r}_l^{\mathrm{T}}d_l+\overline{r}_l^{\mathrm{T}}M_k^{-1}\overline{r}_l=0, \tag{2.22}$$

从而 (2.15) 得证.

由 (2.20), (2.14) 和 (2.7) 有

$$\begin{aligned}\overline{r}_{l+1}^{\mathrm{T}}M_k^{-1}\overline{r}_i &= (\overline{r}_l+\lambda_l\overline{H}_k d_l)^{\mathrm{T}}M_k^{-1}\overline{r}_i\\ &= \overline{r}_l^{\mathrm{T}}M_k^{-1}\overline{r}_i+\lambda_l d_l^{\mathrm{T}}\overline{H}_k(-d_i+\beta_{i-1}d_{i-1})\\ &= \overline{r}_l^{\mathrm{T}}M_k^{-1}\overline{r}_i-\lambda_l d_l^{\mathrm{T}}\overline{H}_k d_i+\lambda_l\beta_{i-1}d_l^{\mathrm{T}}\overline{H}_k d_{i-1}.\end{aligned}\tag{2.23}$$

当 $i=1,2,\cdots,l-1$ 时，由归纳法假设可知 (2.23) 右端为零，又

$$\begin{aligned}\overline{r}_{l+1}^{\mathrm{T}}M_k^{-1}\overline{r}_l &= \overline{r}_l^{\mathrm{T}}M_k^{-1}\overline{r}_l-\lambda_l d_l^{\mathrm{T}}\overline{H}_k d_l+\lambda_l\beta_{l-1}d_l^{\mathrm{T}}\overline{H}_k d_{l-1}\\ &= \overline{r}_l^{\mathrm{T}}M_k^{-1}\overline{r}_l-\frac{\overline{r}_l^{\mathrm{T}}\overline{s}_l}{d_l^{\mathrm{T}}\overline{H}_k d_l}d_l^{\mathrm{T}}\overline{H}_k d_l\\ &= \overline{r}_l^{\mathrm{T}}M_k^{-1}\overline{r}_l-\overline{r}_l^{\mathrm{T}}M_k^{-1}\overline{r}_l\\ &= 0,\end{aligned}$$

因此 (2.17) 得证.

利用 (2.7), (2.20) 和 (2.14) 有

$$\begin{aligned}d_{l+1}^{\mathrm{T}}\overline{H}_k d_i &= (-\overline{s}_{l+1}+\beta_l d_l)^{\mathrm{T}}\overline{H}_k d_i\\ &= -\overline{s}_{l+1}^{\mathrm{T}}\frac{\overline{r}_{i+1}-\overline{r}_i}{\lambda_i}+\beta_l d_l^{\mathrm{T}}\overline{H}_k d_i\\ &= -\frac{\overline{r}_{l+1}^{\mathrm{T}}M_k^{-1}\overline{r}_{i+1}-\overline{r}_{l+1}^{\mathrm{T}}M_k^{-1}\overline{r}_i}{\lambda_i}+\beta_l d_l^{\mathrm{T}}\overline{H}_k d_i,\end{aligned}\tag{2.24}$$

对于 $i=1,2,\cdots,l-1$ 由归纳法假设可知 (2.24) 右端为零，又

$$d_{l+1}^{\mathrm{T}}\overline{H}_k d_l=(-\overline{s}_{l+1}+\beta_l d_l)^{\mathrm{T}}\overline{H}_k d_l=-\overline{s}_{l+1}^{\mathrm{T}}\overline{H}_k d_l+\tfrac{\overline{s}_{l+1}^{\mathrm{T}}\overline{H}_k d_l}{d_l^{\mathrm{T}}\overline{H}_k d_l}d_l^{\mathrm{T}}\overline{H}_k d_l=0.\tag{2.25}$$

由此 (2.16) 得证.

再利用 (2.7) 和 (2.14), 有

$$d_{l+1}^{\mathrm{T}}\overline{r}_{l+1}=(-\overline{s}_{l+1}+\beta_l d_l)^{\mathrm{T}}\overline{r}_{l+1}=-\overline{s}_{l+1}^{\mathrm{T}}\overline{r}_{l+1}=-\overline{r}_{l+1}^{\mathrm{T}}M_k^{-1}\overline{r}_{l+1}.$$

因此 (2.18) 成立.

由 (2.7) 和 (2.14) 得

$$d_{l+1}^{\mathrm{T}}M_k d_i=(-\overline{s}_{l+1}+\beta_l d_l)^{\mathrm{T}}M_k d_i=-\overline{r}_{l+1}^{\mathrm{T}}M_k^{-1}M_k d_i+\beta_l d_l^{\mathrm{T}}M_k d_i=\beta_l d_l^{\mathrm{T}}M_k d_i>0.$$

对于 $i=1,2,\cdots,l$ 成立，于是 (2.19) 得证.

引理 2.2 设迭代步 $p_k^z(\tau)=\mathit{\Gamma}_k^z(\tau)$ 是由既约预条件共轭梯度路径得到的，则满足以下性质

i) 对于 $\tau\in(0,+\infty)$, 路径的范数函数 $\|\mathit{\Gamma}_k^z(\tau)\|_{M_k}$ 是单调递增的，这里 $\|\cdot\|_{M_k}$ 为椭球范数， $\|x\|_{M_k}=(x^{\mathrm{T}}M_k x)^{\frac{1}{2}}, \forall x\in\Re^{n-m}$, 其中 $M_k=\widetilde{Z}\widetilde{H}_k\widetilde{Z}^{\mathrm{T}}$.

ii) 当 $\tau \in (0, +\infty)$ 时，二次近似函数 $\varphi_k(\varGamma_k^z(\tau))$ 沿着既约预条件共轭梯度路径是单调递减的.

iii) 如果 $\overline{H}_k$ 是正定的，则

$$\lim_{\tau \to \infty} \varGamma_k^z(\tau) = -\overline{H}_k^{-1} \overline{g}_k. \tag{2.26}$$

证 为了方便起见，将第 k 次迭代省略.

1) 由共轭梯度路径定义，当 $\sum\limits_{j=1}^{i-1} \lambda_j < \tau \le \sum\limits_{j=1}^{i} \lambda_j (i \le q)$ 时，即 $0 < \tau - \sum\limits_{j=1}^{i-1} \lambda_j \le \lambda_i$ 时，

$$t_i(\tau) = \min\Big\{\lambda_i, \max\Big\{0, \tau - \sum_{j=1}^{i-1} \lambda_j\Big\}\Big\} = \min\Big\{\lambda_i, \tau - \sum_{j=1}^{i-1} \lambda_j\Big\} = \tau - \sum_{j=1}^{i-1} \lambda_j.$$

对于 $l < i$,

$$t_l(\tau) = \min\Big\{\lambda_l, \max\Big\{0, \tau - \sum_{j=1}^{l-1} \lambda_j\Big\}\Big\} = \min\Big\{\lambda_l, \tau - \sum_{j=1}^{l-1} \lambda_j\Big\} = \lambda_l.$$

对于 $i < l \le q$,

$$t_l(\tau) = \min\{\lambda_l, 0\} = 0.$$

所以当 $\sum\limits_{j=1}^{i-1} \lambda_j < \tau \le \sum\limits_{j=1}^{i} \lambda_j (i \le q)$ 时，有

$$\varGamma^z(\tau) = \sum_{j=1}^{i-1} \lambda_j d_j + \Big(\tau - \sum_{j=1}^{i-1} \lambda_j\Big) d_i. \tag{2.27}$$

当 $\tau > \sum\limits_{j=1}^{q} \lambda_j$ 时，可得

$$\varGamma^z(\tau) = \sum_{j=1}^{q} \lambda_j d_j. \tag{2.28}$$

定义 $\varPsi(\tau) \overset{\text{def}}{=} \|\varGamma^z(\tau)\|_{M_k}^2$, 只要证明 $\frac{\mathrm{d}\varPsi(\tau)}{\mathrm{d}\tau} \ge 0$ 即可证得 i) 成立. 由 $\varGamma^z(\tau)$ 的定义有

$$\begin{aligned}
\varPsi(\tau) &= \Big\| \sum_{j=1}^{i-1} \lambda_j d_j + \Big(\tau - \sum_{j=1}^{i-1} \lambda_j\Big) d_i \Big\|_{M_k}^2 \\
&= \Big[\sum_{j=1}^{i-1} \lambda_j d_j + \Big(\tau - \sum_{j=1}^{i-1} \lambda_j\Big) d_i\Big]^{\mathrm{T}} M_k \Big[\sum_{j=1}^{i-1} \lambda_j d_j + \Big(\tau - \sum_{j=1}^{i-1} \lambda_j\Big) d_i\Big] \\
&= \sum_{j=1}^{i-1} \lambda_j d_j^{\mathrm{T}} M_k \Big(\sum_{j=1}^{i-1} \lambda_j d_j\Big) + 2\Big(\tau - \sum_{j=1}^{i-1} \lambda_j\Big) \sum_{j=1}^{i-1} \lambda_j d_i^{\mathrm{T}} M_k d_j + \Big(\tau - \sum_{j=1}^{i-1} \lambda_j\Big)^2 d_i^{\mathrm{T}} M_k d_i,
\end{aligned}$$

对 $\Psi(\tau)$ 关于 τ 求导,

$$\frac{\mathrm{d}\,\Psi(\tau)}{\mathrm{d}\tau}=2\sum_{j=1}^{i-1}\lambda_j d_i^{\mathrm{T}}M_k d_j+2\left(\tau-\sum_{j=1}^{i-1}\lambda_j\right)d_i^{\mathrm{T}}M_k d_i,$$

由 (2.19) 式, 以及 $\tau-\sum\limits_{j=1}^{i-1}\lambda_j>0$ 和 M_k 正定, 可知 $\frac{\mathrm{d}\,\Psi(\tau)}{\mathrm{d}\tau}>0$. 因为 $\|\Gamma^z(\tau)\|_{M_k}$ 在 $\Re^+$ 上连续, 所以 $\|\Gamma^z(\tau)\|_{M_k}$ 在 $\tau\in(0,+\infty)$ 上单调递增, 从而证得 i) 成立.

2) 当 $\sum\limits_{j=1}^{i-1}\lambda_j<\tau\le\sum\limits_{j=1}^{i}\lambda_j(i\le q)$ 时,

$$\begin{aligned}\frac{\mathrm{d}\varphi(\Gamma^z(\tau))}{\mathrm{d}\tau}&=(\overline{g}_k+\overline{H}_k\Gamma^z(\tau))^{\mathrm{T}}\frac{\mathrm{d}\Gamma^z(\tau)}{\mathrm{d}\tau}\\&=\left[\overline{g}_k+\overline{H}_k\left(\sum_{j=1}^{i-1}\lambda_j d_j+\left(\tau-\sum_{j=1}^{i-1}\lambda_j\right)d_i\right)\right]^{\mathrm{T}}d_i\\&=\overline{g}_k^{\mathrm{T}}d_i+\left(\tau-\sum_{j=1}^{i-1}\lambda_j\right)d_i^{\mathrm{T}}\overline{H}_k d_i,\end{aligned}\tag{2.29}$$

由 $\overline{r}_1=\overline{g}_k$, $\tau-\sum\limits_{j=1}^{i-1}\lambda_j\le\lambda_i$, λ_i 的定义以及 (2.14), (2.18) 和 (2.20), 有

$$\begin{aligned}\frac{\mathrm{d}\varphi(\Gamma^z(\tau))}{\mathrm{d}\tau}&=d_i^{\mathrm{T}}\overline{r}_1+\left(\tau-\sum_{j=1}^{i-1}\lambda_j\right)d_i^{\mathrm{T}}\overline{H}_k d_i\\&\le d_i^{\mathrm{T}}\overline{r}_1+\lambda_i d_i^{\mathrm{T}}\overline{H}_k d_i\\&=d_i^{\mathrm{T}}\overline{r}_1+\overline{r}_i^{\mathrm{T}}\overline{s}_i\\&=d_i^{\mathrm{T}}\overline{r}_1+\overline{r}_i^{\mathrm{T}}M_k^{-1}\overline{r}_i\\&=d_i^{\mathrm{T}}(\overline{r}_1-\overline{r}_i)\\&=d_i^{\mathrm{T}}\left(\overline{r}_1-\overline{r}_1-\sum_{j=1}^{i-1}\lambda_j\overline{H}_k d_j\right)\\&=0.\end{aligned}$$

所以 $\varphi(\Gamma^z(\tau))$ 沿着既约预条件共轭梯度路径是单调递减的.

3) 如果 $\overline{H}_k$ 正定, 则上述采用精确搜索的共轭梯度法在 q 步后终止, $\overline{r}_{q+1}=0$, 因为 $\overline{H}_k v_{q+1}+\overline{g}_k=\overline{r}_{q+1}=0$, 所以 $v_{q+1}=-\overline{H}_k^{-1}\overline{g}_k$, 又 $\lambda_{q+1}=\frac{\overline{r}_{q+1}^{\mathrm{T}}\overline{s}_{q+1}}{d_{q+1}^{\mathrm{T}}\overline{H}_k d_{q+1}}$, 所以 $\lambda_{q+1}=0$, 从而有 $\lim\limits_{\tau\to\infty}t_{q+1}(\tau)=0$, 利用 (2.8) 和 (2.28) 可得

$$\lim_{\tau\to\infty}\Gamma_k^z(\tau)=\sum_{j=1}^{q}\lambda_j d_j=v_{q+1}=-\overline{H}_k^{-1}\overline{g}_k.$$

即 (2.26) 成立.

引理 2.3 设迭代步 $p_k^z(\tau)=\Gamma_k^z(\tau)$ 是由既约预条件共轭梯度路径得到的，则

1) 目标函数的既约梯度 $\overline{g}_k$ 和共轭梯度路径产生的迭代步 $p_k^z(\tau)$ 的关系函数 $\Phi_k(\tau)\stackrel{\text{def}}{=}\overline{g}_k^{\mathrm T}p_k^z(\tau)$ 在 $\tau\in(0,+\infty)$ 上是单调递减的；

2) 对于 $0<\tau<+\infty$, $\Phi_k(\tau)\stackrel{\text{def}}{=}\overline{g}_k^{\mathrm T}p_k^z(\tau)$ 满足充分下降方向，即存在 $\overline{\chi}>0,\ \widehat{\chi}>0$ 使得 $\|M_k\|\le\overline{\chi},\ \|M_k^{-1}\|\le\widehat{\chi}$, 有

$$\overline{g}_k^{\mathrm T}p_k^z(\tau)\le-\min\left\{\frac{\tau}{\overline{\chi}}\ ,\ \frac{1}{\overline{\chi}\widehat{\chi}^2\|\overline{H}_k\|}\right\}\|\overline{g}_k\|^2. \tag{2.30}$$

证 1) 由引理 2.2 中的 (2.27) 有

$$\Phi_k(\tau)=(\overline{g}_k)^{\mathrm T}p_k^z(\tau)=\sum_{j=1}^{i-1}\lambda_j\overline{g}_k^{\mathrm T}d_j+\left(\tau-\sum_{j=1}^{i-1}\lambda_j\right)\overline{g}_k^{\mathrm T}d_i, \tag{2.31}$$

上式关于 τ 求导，并利用 $d_1=-M_k^{-1}\overline{g}_k$ 和 (2.19) 可得

$$\frac{\mathrm d\Phi_k(\tau)}{\mathrm d\tau}=\overline{g}_k^{\mathrm T}d_i=-d_1^{\mathrm T}M_kd_i<0, \tag{2.32}$$

所以 $\Phi_k(\tau)$ 在 $\tau\in(0,+\infty)$ 上是单调递减的.

2) 当 $0<\tau\le\lambda_1$ 时，因为 $M_k=\widetilde{Z}\widetilde{H}_k\widetilde{Z}^{\mathrm T}$ 是正定矩阵，所以存在 $\overline{\chi}>0,\ \widehat{\chi}>0$ 使得 $\|M_k\|\le\overline{\chi},\ \|M_k^{-1}\|\le\widehat{\chi}$, 从而

$$\Phi_k(\tau)=\overline{g}_k^{\mathrm T}p_k^z(\tau)=\overline{g}_k^{\mathrm T}(\tau d_1)=-\tau\overline{g}_k^{\mathrm T}M_k^{-1}\overline{g}_k\le-\frac{\tau}{\overline{\chi}}\|\overline{g}_k\|^2. \tag{2.33}$$

当 $\tau>\lambda_1$ 时，由于 $\Phi_k(\tau)$ 在 $\tau\in(\lambda_1,+\infty)$ 上是单调递减的，所以

$$\Phi_k(\tau)<\Phi_k(\lambda_1)=-\lambda_1\|\overline{g}_k\|^2\le-\frac{\overline{g}_k^{\mathrm T}M_k^{-1}\overline{g}_k\|\overline{g}_k\|^2}{\overline{g}_k^{\mathrm T}M_k^{-1}\overline{H}_kM_k^{-1}\overline{g}_k}\le-\frac{1}{\overline{\chi}\widehat{\chi}^2\|\overline{H}_k\|}\|\overline{g}_k\|^2. \tag{2.34}$$

由 (2.33) 和 (2.34) 可得，当 $\tau\in(0,+\infty)$ 时， (2.30) 成立.

(2.30) 表明目标函数的既约梯度 $\overline{g}_k$ 和迭代步长 $p_k^z(\tau)$ 之间的关系具有“足够”下降量.

3 算　法

本节中给出了求解问题 (1.1) 的既约预条件共轭梯度路径的算法.

初始步 选定参数 $\xi\in(0,1),\ \omega\in(0,1),\ \varepsilon>0$. 令 $m(0)=0$, 选取一个对称矩阵 B_0, 选取初始可行点 $x_0\in\Omega\subseteq\Re^n$, 再令 $k=0$, 转主步.

主步

1 计算 $f_k=f(x_k),\ g_k=\nabla f(x_k),\ B_k,\ \widetilde{M}_k,\ \widetilde{A},\ Z$.

2 如果 $\|Zg_k\|\le\varepsilon$, 则停止计算， x_k 作为最优解；否则转下一步.

3 构造共轭梯度路径 $\Gamma_k^z(\tau)$, 获得路 $p_k^z(\tau)$, 则迭代步 $p_k(\tau)=Z^{\mathrm T}p_k^z(\tau)$. 取 $\tau=\infty,\omega^{-n}$, $\omega^{-(n-1)}$, 直到下面的不等式成立

$$f(x_{l(k)})-f(x_k+p_k(\tau))\ge\xi\Big[f(x_k)-\varphi_k(p_k^z(\tau))\Big],\quad \xi\in(0,1), \tag{3.1}$$

其中 $f(x_{l(k)})=\max\limits_{0\le j\le m(k)}\{f(x_{k-j})\}$, 记满足 (3.1) 式的第一个 τ 的值为 τ_k.

4 令

$$x_{k+1}=x_k+p_k(\tau_k). \tag{3.2}$$

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

注 1 把非单调技术和常用的单调技术相比较，当 $Q\ge 1$ 时，迭代步 $p_k(\tau_k)$ 只能保证 $f(x_k+p_k(\tau_k))$ 小于 $f(x_{l(k)})$. 因此，一般情况下， f_k 不是单调递减的. 当 $Q=0$ 时，易见该算法转化为一般意义下的单调算法. 所以单调算法可以看成非单调算法在 $Q=0$ 时的特殊情形.

4 整体收敛性分析

在本节中， $f:\Re^n\longrightarrow\Re$ 是二次连续可微函数，且在可行域 Ω 上是有下界的. 为了证明算法的整体收敛性，需作如下假设

假设 A1 给定 $x_0\in\Omega$, 水平集 $\mathcal{L}(x_0)\stackrel{\rm def}{=}\{x\in\Omega|f(x)\le f(x_0)\}$ 是紧的，算法产生的迭代序列 $x_k\in\mathcal{L}(x_0)$.

假设 A2 $\|B_k\|$ 上有界，即对任意的 k, 存在 $\chi_b>0$, 使得 $\|B_k\|\le\chi_b$.

假设 A3 $\|\nabla^2 f(x)\|$ 上有界，即对任意的 $x\in\mathcal{L}(x_0)$, 存在 $\chi_f>0$, 使得 $\|\nabla^2 f(x)\|\le\chi_f$.

由假设 A2, A3, 以及 Z 是正交矩阵知 $\|\overline{H}_k\|\le\chi_b$, $\|Z\nabla^2 f(x_k)Z^{\rm T}\|\le\chi_f$.

引理 4.1 在第 k 步迭代中，令迭代步 $p_k^z(\tau)=\varGamma_k^z(\tau)$ 是由既约预条件共轭梯度路径得到的，则预计下降量 $f(x_k)-\varphi_k(p_k^z(\tau))$ 满足充分下降条件，即

$$f(x_k)-\varphi_k(p_k^z(\tau))\ge\frac{1}{2}\min\left\{\frac{\tau}{\overline{\chi}},\frac{1}{\chi_b\overline{\chi}^2\widehat{\chi}^2}\right\}\|\overline{g}_k\|^2. \tag{4.1}$$

证 当 $0<\tau\le\lambda_1$ 时,

$$\begin{aligned}f(x_k)-\varphi_k(p_k^z(\tau))&=-\overline{g}_k^{\rm T}(\tau d_1)-\frac{1}{2}(\tau d_1)^{\rm T}\overline{H}_k(\tau d_1)\\&\ge\tau\overline{g}_k^{\rm T}M_k^{-1}\overline{g}_k-\frac{\tau}{2}\lambda_1\overline{g}_k^{\rm T}\overline{H}_k\overline{g}_k\\&=\tau\overline{g}_k^{\rm T}M_k^{-1}\overline{g}_k-\frac{\tau}{2}\frac{\overline{g}_k^{\rm T}M_k^{-1}\overline{g}_k}{d_1^{\rm T}\overline{H}_kd_1}d_1^{\rm T}\overline{H}_kd_1\\&=\frac{\tau}{2}\overline{g}_k^{\rm T}M_k^{-1}\overline{g}_k\\&\ge\frac{1}{2}\frac{\tau}{\overline{\chi}}\|\overline{g}_k\|^2.\end{aligned} \tag{4.2}$$

当 $\tau > \lambda_1$ 时，由于 $\varphi_k(p_k^z(\tau))$ 在 $\tau \in (\lambda_1, +\infty)$ 时是单调递减的，从而

$$\begin{aligned} f(x_k) - \varphi_k(p_k^z(\tau)) &> f(x_k) - \varphi_k(p_k^z(\lambda_1)) \\ &= -\frac{\lambda_1^2}{2}\overline{g}_k^{\mathrm{T}}\overline{H}_k\overline{g}_k + \lambda_1\overline{g}_k^{\mathrm{T}}M_k^{-1}\overline{g}_k \\ &= \frac{(\overline{g}_k^{\mathrm{T}}M_k^{-1}\overline{g}_k)^2}{d_1^{\mathrm{T}}\overline{H}_k d_1} - \frac{(\overline{g}_k^{\mathrm{T}}M_k^{-1}\overline{g}_k)^2}{2(d_1^{\mathrm{T}}\overline{H}_k d_1)^2}d_1^{\mathrm{T}}\overline{H}_k d_1 \\ &= \frac{(\overline{g}_k^{\mathrm{T}}M_k^{-1}\overline{g}_k)^2}{2\overline{g}_k^{\mathrm{T}}M_k^{-1}\overline{H}_k M_k^{-1}\overline{g}_k} \\ &\geq \frac{1}{2\chi_b\overline{\chi}^2\widehat{\chi}^2}\|\overline{g}_k\|^2. \end{aligned} \tag{4.3}$$

由 (4.2) 和 (4.3) 可得，(4.1) 成立.

下面的引理证明了算法第 3 步在有限步内终止，即经过有限步的缩小 τ，使 (3.1) 成立.

引理 4.2　在算法第 3 步中，若 $\|\overline{g}_k\| \neq 0$，则存在 $\widehat{\tau} > 0$，使得

$$\tau_k \geq \widehat{\tau} \overset{\mathrm{def}}{=} \min\left\{\omega\lambda_1, \frac{\omega}{\chi_b\overline{\chi}\widehat{\chi}^2}, \frac{(1-\xi)\omega}{(\chi_f+\chi_b)\overline{\chi}\widehat{\chi}^2}\right\}. \tag{4.4}$$

证　假设引理结论不成立，则

$$\tau_k < \min\left\{\omega\lambda_1, \frac{\omega}{\chi_b\overline{\chi}\widehat{\chi}^2}, \frac{(1-\xi)\omega}{(\chi_f+\chi_b)\overline{\chi}\widehat{\chi}^2}\right\}. \tag{4.5}$$

由算法可知 $\frac{\tau_k}{\omega}$ 不满足 (3.1)，即

$$f(x_{l(k)}) - f\Big(x_k + p_k\Big(\frac{\tau_k}{\omega}\Big)\Big) < \xi\Big[f(x_k) - \varphi_k\Big(p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big)\Big]. \tag{4.6}$$

因为 $f(x_{l(k)}) = \max\limits_{0\leq j\leq m(k)}\{f(x_{k-j})\}$，因而有下式成立

$$f(x_k) - f\Big(x_k + p_k\Big(\frac{\tau_k}{\omega}\Big)\Big) < \xi\Big[f(x_k) - \varphi_k\Big(p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big)\Big], \tag{4.7}$$

由上式可得

$$(1-\xi)\Big[\varphi_k\Big(p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big) - f(x_k)\Big] + f\Big(x_k + p_k\Big(\frac{\tau_k}{\omega}\Big)\Big) - \varphi_k\Big(p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big) > 0. \tag{4.8}$$

由 (4.1) 和 (4.5) 可得

$$\begin{aligned} (1-\xi)\Big[\varphi_k\Big(p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big) - f(x_k)\Big] &\leq -\frac{1}{2}(1-\xi)\min\left\{\frac{\tau_k}{\omega\overline{\chi}}, \frac{1}{\chi_b\overline{\chi}^2\widehat{\chi}^2}\right\}\|\overline{g}_k\|^2 \\ &= -\frac{1}{2}(1-\xi)\frac{\tau_k}{\omega\overline{\chi}}\|\overline{g}_k\|^2. \end{aligned} \tag{4.9}$$

又因为 $\|p_k^z(\frac{\tau_k}{\omega})\|^2=\|p_k(\frac{\tau_k}{\omega})\|^2$, 从而

$$\begin{aligned}&f\Big(x_k+p_k\Big(\frac{\tau_k}{\omega}\Big)\Big)-\varphi_k\Big(p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big)\\&=\frac{1}{2}\Big(p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big)^{\mathrm{T}}Z\nabla^2 f\Big(x_k+\theta p_k\Big(\frac{\tau_k}{\omega}\Big)\Big)Z^{\mathrm{T}}p_k^z\Big(\frac{\tau_k}{\omega}\Big)-\frac{1}{2}\Big(p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big)^{\mathrm{T}}\overline{H}_k p_k^z\Big(\frac{\tau_k}{\omega}\Big)\\&\le\frac{1}{2}(\chi_f+\chi_b)\Big\|p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big\|^2,\end{aligned}\tag{4.10}$$

其中 $0<\theta<1$. 由 (4.5) 知， $\tau_k<\omega\lambda_1$, 即 $0<\frac{\tau_k}{\omega}<\lambda_1$, 因而有 $p_k^z(\frac{\tau_k}{\omega})=\frac{\tau_k}{\omega}d_1=-\frac{\tau_k}{\omega}M_k^{-1}\overline{g}_k$, 从而

$$\Big\|p_k^z\Big(\frac{\tau_k}{\omega}\Big)\Big\|^2=\frac{\tau_k^2}{\omega^2}\|M_k^{-1}\overline{g}_k\|^2\le\frac{\tau_k^2\widehat{\chi}^2}{\omega^2}\|\overline{g}_k\|^2,\tag{4.11}$$

由 (4.8)–(4.11) 得

$$-\frac{1}{2}(1-\xi)\frac{\tau_k}{\omega\overline{\chi}}\|\overline{g}_k\|^2+\frac{1}{2}(\chi_f+\chi_b)\frac{\tau_k^2\widehat{\chi}^2}{\omega^2}\|\overline{g}_k\|^2>0,\tag{4.12}$$

因为 $\|\overline{g}_k\|\neq 0$, 从而有

$$\tau_k>\frac{(1-\xi)\omega}{(\chi_f+\chi_b)\overline{\chi}\widehat{\chi}^2}.\tag{4.13}$$

这与 (4.5) 相矛盾. 从而假设不成立，所以 (4.4) 成立.

在引理 4.1, 引理 4.2 的前提下，定理 4.3 进一步说明了此算法具有整体弱收敛性.

定理 4.3 假设 A1–A3 成立. 令 $\{x_k\}\subseteq\Re^n$ 是由算法产生的序列，则

$$\liminf_{k\to\infty}\|Zg_k\|=0.\tag{4.14}$$

证 假设定理的结论不成立，则存在 $\varepsilon>0$, 使得对于所有的 k,

$$\|\overline{g}_k\|\stackrel{\text{def}}{=}\|Zg_k\|\ge\varepsilon.\tag{4.15}$$

由接受准则 (3.1) 有

$$f(x_{l(k)})-f(x_k+p_k(\tau))\ge\xi\Big[f(x_k)-\varphi_k(p_k^z(\tau))\Big],$$

所以 $f(x_{k+1})\le f(x_{l(k)})$. 注意到 $m(k+1)=\min\{m(k)+1,Q\}\le m(k)+1$, 则有

$$\begin{aligned}f(x_{l(k+1)})&\le\max_{0\le j\le m(k)+1}\{f(x_{k+1-j})\}\\&=\max\{f(x_{l(k)}),f(x_{k+1})\}\\&=f(x_{l(k)}),\end{aligned}$$

这表明序列 $\{f(x_{l(k)})\}$ 是单调非增的. 又因为 $f(x_k)$ 是下有界的，从而序列 $\{f(x_{l(k)})\}$ 是收敛的. 从而由 (4.1) 可得

$$\begin{aligned}f(x_{l(k)}) &= f(x_{l(k)-1}+p_{l(k)-1}(\tau_{l(k)-1}))\\ &\le \max_{0\le j\le m(k)}\{f(x_{l(k)-1-j})\}-\xi\Big[f(x_{l(k)-1})-\varphi_{l(k)-1}(p^z_{l(k)-1}(\tau_{l(k)-1}))\Big]\\ &\le \max_{0\le j\le m(k)}\{f(x_{l(k)-1-j})\}-\frac{\xi}{2}\min\Big\{\frac{\tau_{l(k)-1}}{\overline{\chi}},\frac{1}{\chi_b\overline{\chi}^2\widehat{\chi}^2}\Big\}\varepsilon^2,\end{aligned}\tag{4.16}$$

由于 $\{f(x_{l(k)})\}$ 是收敛的，由上式可得

$$\lim_{k\to\infty}\tau_{l(k)-1}=0.\tag{4.17}$$

类似文 [4] 可证得

$$\lim_{k\to\infty}f(x_{l(k)})=\lim_{k\to\infty}f(x_k).\tag{4.18}$$

由接受准则 (3.1) 得

$$f(x_{k+1})-f(x_{l(k)})\le-\xi\Big[f(x_k)-\varphi_k(p^z_k(\tau))\Big]\le-\frac{\xi}{2}\min\Big\{\frac{\tau_k}{\overline{\chi}},\frac{1}{\chi_b\overline{\chi}^2\widehat{\chi}^2}\Big\}\varepsilon^2.$$

这意味着

$$\lim_{k\to\infty}\tau_k=0.\tag{4.19}$$

这与引理 4.2 相矛盾，从而假设 (4.15) 不成立，所以定理结论成立.

5 局部收敛速率

定理 4.3 证明了算法的整体弱收敛性，本节给出此算法不仅具有整体强收敛性，而且保持快速的超线性收敛速率. 首先需如下假设.

假设 A4 问题 (1.1) 的解 x_* 满足强二阶充分条件，即存在 $\zeta>0$, 使得

$$(p^z)^{\mathrm T}Z\nabla^2f(x_*)Z^{\mathrm T}p^z\ge\zeta\|p^z\|^2,\quad\forall p^z\in\Re^{n-m}.\tag{5.1}$$

假设 A5

$$\lim_{k\to\infty}\frac{\|(\overline{H}_k-Z\nabla^2f(x_*)Z^{\mathrm T})Zp_k(\tau_k)\|}{\|p_k(\tau_k)\|}=0.\tag{5.2}$$

假设 A5 意味着

$$p^z_k(\tau_k)^{\mathrm T}(Z\nabla^2f(x_*)Z^{\mathrm T}-\overline{H}_k)p^z_k(\tau_k)=o(\|p_k(\tau_k)\|\|p^z_k(\tau_k)\|).\tag{5.3}$$

定理 5.1 假设 A1–A5 成立，且假设存在 $\eta>0$, 使得 $\|p^z_k(\tau_k)\|\le\eta\|\overline{g}_k\|$, 其中 $\|\overline{g}_k\|\overset{\text{def}}{=}\|Zg_k\|$. 令 $\{x_k\}\subseteq\Re^n$ 是由算法产生的序列，则

$$\lim_{k\to\infty}\|\overline{g}_k\|=0.\tag{5.4}$$

证 假设存在 $\varepsilon_1 \in (0,1)$ 和 $\{\overline{g}_k\}$ 中的一个子列 $\{\overline{g}_{m_i}\}$, 使得对所有的 $m_i(i=1,2,\cdots)$ 都有

$$\|\overline{g}_{m_i}\| \geq \varepsilon_1. \tag{5.5}$$

定理 4.3 保证了对任意的 $\varepsilon_2 \in (0,\varepsilon_1)$, 存在另一个子序列 $\{\overline{g}_{n_i}\}$, 当 i 充分大时，有下列两式成立

$$\|\overline{g}_k\| \geq \varepsilon_2,\ m_i \leq k < n_i \text{ 和 } \|\overline{g}_{n_i}\| < \varepsilon_2.$$

由 (4.1) 和 (4.4), 以及定理假设条件得

$$\begin{aligned} f(x_k) - \varphi_k(\Gamma_k^z(\tau)) &\geq \frac{1}{2\eta^2}\min\left\{\frac{\widehat{\tau}}{\overline{\chi}}, \frac{1}{\chi_b\overline{\chi}^2\widehat{\chi}^2}\right\}\|p_k^z(\tau_k)\|^2 \\ &= \widehat{\alpha}\|p_k^z(\tau_k)\|^2 \\ &= \widehat{\alpha}\|p_k(\tau_k)\|^2, \end{aligned} \tag{5.6}$$

其中 $\widehat{\alpha} = \frac{1}{2\eta^2}\min\{\frac{\widehat{\tau}}{\overline{\chi}}, \frac{1}{\chi_b\overline{\chi}^2\widehat{\chi}^2}\}$. 由 (4.16) 和 (5.6) 可得

$$\begin{aligned} f(x_{l(k)}) &\leq f(x_{l(l(k)-1)}) - \xi\Big[f(x_{l(k)-1}) - \varphi_{l(k)-1}(p^z_{l(k)-1}(\tau_{l(k)-1}))\Big] \\ &\leq f(x_{l(l(k)-1)}) - \xi\widehat{\alpha}\|p_{l(k)-1}(\tau_{l(k)-1})\|^2. \end{aligned} \tag{5.7}$$

类似于文 [4] 中的证明有

$$\lim_{k\to\infty} f(x_{l(k)}) = \lim_{k\to\infty} f(x_k). \tag{5.8}$$

从而有

$$f(x_{k+1}) - f(x_{l(k)}) \leq -\xi\widehat{\alpha}\|p_k(\tau_k)\|^2, \tag{5.9}$$

(5.8) 和 (5.9) 意味着

$$\lim_{k\to\infty} \|p_k(\tau_k)\| = 0. \tag{5.10}$$

由假设 A5 可得

$$\begin{aligned} &\Big|\varphi_k(p_k^z(\tau_k)) - f(x_k + p_k(\tau_k))\Big| \\ =&\Big|\overline{g}_k^{\mathrm{T}}p^z + \frac{1}{2}(p^z)^{\mathrm{T}}\overline{H}_kp^z - g_k^{\mathrm{T}}p_k + \frac{1}{2}(p_k)^{\mathrm{T}}\overline{H}_kp_k + o(\|p_k\|^2)\Big| \\ =&\Big|\frac{1}{2}(p_k^z)^{\mathrm{T}}(\overline{H}_k - Z\nabla^2 f(x_*)Z^{\mathrm{T}})p_k^z + o(\|p_k\|^2)\Big| \\ =&\, o(\|p_k\|^2). \end{aligned} \tag{5.11}$$

从而对于充分大的 i, $m_i < k < n_i$ 可得

$$\begin{aligned}\frac{f(x_k)-f(x_k+p_k)}{f(x_k)-\varphi_k(p_k^z)} &= 1+\frac{\varphi_k(p_k^z)-f(x_k)+f(x_k)-f(x_k+p_k)}{f(x_k)-\varphi_k(p_k^z)}\\ &= 1+\frac{\varphi_k(p_k^z)-f(x_k+p_k)}{f(x_k)-\varphi_k(p_k^z)}\\ &\geq 1-\frac{o(\|p_k\|^2)}{\widehat{\alpha}\|p_k\|^2},\end{aligned}\tag{5.12}$$

因此存在 $\rho\in(0,1)$ 使得

$$\frac{f(x_k)-f(x_k+p_k)}{f(x_k)-\varphi_k(p_k^z)}\geq\rho,\tag{5.13}$$

这意味着对充分大的 i, $m_i < k < n_i$ 使得

$$\begin{aligned}f(x_k)-f(x_k+p_k) &\geq \rho(f(x_k)-\varphi_k(p_k^z))\\ &\geq \frac{\rho}{2}\min\Big\{\frac{\widehat{\tau}}{\overline{\chi}},\frac{1}{\chi_b\overline{\chi}^2\widehat{\chi}^2}\Big\}\varepsilon_2^2\\ &= \overline{\alpha}\varepsilon_2^2,\end{aligned}\tag{5.14}$$

其中 $\overline{\alpha}=\frac{\rho}{2}\min\{\frac{\widehat{\tau}}{\overline{\chi}},\frac{1}{\chi_b\overline{\chi}^2\widehat{\chi}^2}\}$. 从而由 $n_i-m_i\geq 1$ 得

$$\sum_{k=m_i}^{n_i-1}[f(x_k)-f(x_k+p_k)]\geq(n_i-m_i)\overline{\alpha}\varepsilon_2^2\geq\overline{\alpha}\varepsilon_2^2,\tag{5.15}$$

这意味着对于充分大的 i, 有

$$\overline{\alpha}\varepsilon_2^2\leq f(x_{m_i})-f(x_{n_i})\to 0.\tag{5.16}$$

显然 (5.16) 产生矛盾，因此假设 (5.5) 不成立，从而定理结论成立.

引理 5.2 假设 $f(x_k)-\varphi_k(p_k^z)\geq-\lambda_{\min}(\overline{H}_k)\|p_k^z\|^2$, 其中 $\lambda_{\min}(\overline{H}_k)$ 表示 $\overline{H}_k$ 的最小特征值，设 x_* 是序列 $\{x_k\}$ 的一个极限点，则 $\overline{H}_*$ 是半正定的.

证　假设 $\overline{H}_*$ 不是半正定的，则存在 $\varepsilon>0$, 使得

$$\lambda_{\min}(\overline{H}_*)=-\varepsilon<0,\tag{5.17}$$

则当 x_k 充分接近 x_* 时，有

$$-\lambda_{\min}(\overline{H}_k)\geq\frac{\varepsilon}{2}.\tag{5.18}$$

因而由定理假设条件得

$$f(x_k)-\varphi_k(p_k^z)\geq\frac{\varepsilon}{2}\|p_k^z\|^2,\tag{5.19}$$

由接受准则 (3.1) 有

$$\begin{aligned} f(x_{l(k)}) &\leq f(x_{l(l(k)-1)}) - \xi\Big[f(x_{l(k)-1}) - \varphi_{l(k)-1}(p^z_{l(k)-1}(\tau_{l(k)-1}))\Big] \\ &\leq f(x_{l(l(k)-1)}) - \frac{\xi\varepsilon}{2}\|p^z_{l(k)-1}\|^2, \end{aligned} \tag{5.20}$$

因为 $f(x_{l(k)})$ 是收敛的，从而有

$$\lim_{k\to\infty}\|p^z_{l(k)-1}\| = 0. \tag{5.21}$$

类似于文 [4] 中的证明有

$$\lim_{k\to\infty} f(x_{l(k)}) = \lim_{k\to\infty} f(x_k). \tag{5.22}$$

从而由 (3.1) 和 (5.19) 得

$$f(x_{k+1}) - f(x_{l(k)}) \leq -\xi\Big[f(x_k) - \varphi_k(p^z_k)\Big] \leq -\frac{\xi\varepsilon}{2}\|p^z_k\|^2, \tag{5.23}$$

(5.22) 和 (5.23) 意味着

$$\lim_{k\to\infty}\|p^z_k\| = 0. \tag{5.24}$$

因为 M_k 是正定矩阵，易得 $\|p^z_k\|$ 和 $\|p^z_k\|_{M_k}$ 是等价的. 又 $\|p^z_k\|_{M_k}$ 在 $(0,+\infty)$ 上是严格单调递增的，由 p^z_k 的定义可知

$$\lim_{k\to\infty}\tau_k = 0. \tag{5.25}$$

类似引理 4.2 可证得存在 $\hat{\tau} > 0$, 使得 $\tau_k \geq \hat{\tau}$. 这意味着 (5.25) 不成立，从而 $\lambda_{\min}(\overline{H}_k) \geq 0$, 因此 $\overline{H}_*$ 是半正定的.

定理 5.3 如果 $\overline{H}_k$ 正定，假设 A1–A5 都满足，设 $\{x_k\}$ 为算法产生的迭代序列， x_* 为 $\{x_k\}$ 的极限点，则序列 $\{x_k\}$ 超线性收敛于 x_*, 即

$$\lim_{k\to\infty}\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0. \tag{5.26}$$

证 当 $\overline{H}_k$ 正定时，由引理 2.2 中的 iii) 可知 $p^z_k(\infty) = -\overline{H}_k^{-1}\overline{g}_k$, 下面验证当 k 充分大时，取 $p^z_k(\infty) = -\overline{H}_k^{-1}\overline{g}_k$ 满足算法中的 (3.1).

因为 $\overline{g}_k^{\mathrm{T}}p^z_k(\infty) = -p^z_k(\infty)^{\mathrm{T}}\overline{H}_kp^z_k(\infty)$, 由 (5.1) 和 (5.3) 得

$$\begin{aligned} &\varphi_k(p^z_k(\infty)) - f(x_k) \\ =& \overline{g}_k^{\mathrm{T}}p^z_k(\infty) + \frac{1}{2}p^z_k(\infty)^{\mathrm{T}}\overline{H}_kp^z_k(\infty) \\ =& -p^z_k(\infty)^{\mathrm{T}}\overline{H}_kp^z_k(\infty) + \frac{1}{2}p^z_k(\infty)^{\mathrm{T}}\overline{H}_kp^z_k(\infty) \\ =& \frac{1}{2}p^z_k(\infty)^{\mathrm{T}}(Z\nabla^2 f(x_k)Z^{\mathrm{T}} - \overline{H}_k)p^z_k(\infty) - \frac{1}{2}p^z_k(\infty)^{\mathrm{T}}Z\nabla^2 f(x_k)Z^{\mathrm{T}}p^z_k(\infty) \\ \leq& -\zeta\|p^z_k(\infty)\|^2 + o(\|p_k(\infty)\|\|p^z_k(\infty)\|). \end{aligned} \tag{5.27}$$

因为 $f(x)$ 二次连续可微， $\xi\in(0,1)$, 由 (5.3) 和 (5.27) 可得

$$
\begin{aligned}
f(x_k+p_k(\infty)) &= f(x_k)+g_k^{\mathrm T}p_k(\infty)+\frac{1}{2}p_k(\infty)^{\mathrm T}\nabla^2 f(x_k)p_k(\infty)+o(\|p_k(\infty)\|^2)\\
&= f(x_k)+\overline{g}_k^{\mathrm T}p_k^z(\infty)+\frac{1}{2}p_k^z(\infty)^{\mathrm T}Z\nabla^2 f(x_k)Z^{\mathrm T}p_k^z(\infty)+o(\|p_k^z(\infty)\|^2)\\
&= f(x_k)+\overline{g}_k^{\mathrm T}p_k^z(\infty)+\frac{1}{2}p_k^z(\infty)^{\mathrm T}\overline{H}_k p_k^z(\infty)+\frac{1}{2}p_k^z(\infty)^{\mathrm T}Z\nabla^2 f(x_k)Z^{\mathrm T}p_k^z(\infty)\\
&\quad -\frac{1}{2}p_k^z(\infty)^{\mathrm T}\overline{H}_k p_k^z(\infty)+o(\|p_k^z(\infty)\|^2)\\
&= f(x_k)+\xi\Big[\varphi_k(p_k^z(\infty))-f(x_k)\Big]+(1-\xi)\Big[\varphi_k(p_k^z(\infty))-f(x_k)\Big]\\
&\quad +\frac{1}{2}p_k^z(\infty)^{\mathrm T}(Z\nabla^2 f(x_k)Z^{\mathrm T}-\overline{H}_k)p_k^z(\infty)+o(\|p_k^z(\infty)\|^2)\\
&\le f(x_{l(k)})+\xi\Big[\varphi_k(p_k^z(\infty))-f(x_k)\Big]-(1-\xi)\zeta\|p_k^z(\infty)\|^2+o(\|p_k^z(\infty)\|^2)\\
&\le f(x_{l(k)})+\xi\Big[\varphi_k(p_k^z(\infty))-f(x_k)\Big]. \qquad (5.28)
\end{aligned}
$$

所以当 k 充分大时， $p_k^z(\infty)=-\overline{H}_k^{-1}\overline{g}_k$ 满足 (3.1).

这说明当 $\overline{H}_k$ 正定时，算法在零空间中变成牛顿步或拟牛顿步．从而 (5.26) 成立．

6 数值结果

对上述既约预条件共轭梯度路径算法，在型号为奔腾 IV-1.6GHz 的电脑上，利用数学软件 Matlab 进行编程，实现了具体的数值结果．为了检验算法的有效性，选取参数 $\varepsilon=10^{-6}$, $\xi=0.02$, $w=0.5$, 非单调参数 $Q=0$.

选择了 5 个标准数值测试题．数值计算结果由表 1 给出． HS 表示测试题引自于文 [5]. 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，同时迭代次数和 NG 相同，在此省略列出．非单调技术对求解以下测试题并没有太大的改善，因此本文只给出当 $Q=0$ 的情况．

表 1　测试题数值结果

测试题	Q=0	
	NF	NG
HS028	2	2
HS048	2	2
HS049	14	14
HS050	10	10
HS051	2	2

参 考 文 献

[1] Bulteau J P and Vial J Ph. Curvilinear path and trust region in unconstrained optimization, a convergence analysis. *Mathematical Programming Study*, 1987, **30**: 82–101.

[2] Coleman T F and Verma Arun. A preconditioned conjugate gradient approach to linear equality constrained minimization. *Computational Optimization and Applications*, 2001, **20**: 61–72.

[3] Coleman T F and Li Y. An interior trust region approach for minimization subject to bounds. *SIAM J. Optimization*, 1996, **6**(2): 418–445.

[4] Grippo L, Lampariello F and Lucidi S. A nonmonotonic line search technique for Newton's methods. *SIAM Journal on Numerical Analysis*, 1986, **23**: 707–716.

[5] Hock W and Schittkowski K. Test Examples for Nonlinear Programming Codes. Lecture Notes in Economics and Mathematical Systems 187, Springer-Verlag, Berlin 1981.

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

A REDUCED PRECONDITIONAL CONJUGATE GRADIENT PATH METHOD FOR LINEAR EQUALITY CONSTRAINED OPTIMIZATION

Lin Tao

(*Shanghai Institute of Technology, Shanghai* 200235)

Zhu Detong

(*Business College, Shanghai Normal University, Shanghai* 200234)

Abstract A reduced preconditional conjugate gradient path method with nonmonotonic technique for linear equality constrained optimization problem is proposed. By using the generalized elimination method, the subproblem is equivalent to an unconstrained optimization problem in the null space of constrained matrix. We develop preconditioners based on an extended system. By employing the reduced preconditional conjugate gradient path search strategy, we obtain an iterative direction by solving the quadratic model as well as the iterative step. Based on the good properties of the conjugate gradient path, the global convergence results of the proposed algorithm are proved while fast local superlinear convergence rate is established under some reasonable conditions. Furthermore, numerical results indicate that the algorithm is feasible and effective.

Key words Conjugate gradient path, reduced preconditional, nonmonotonic technique.