第 21 卷 第 1 期　　应 用 数 学 学 报　　Vol. 21 No. 1
1998 年 1 月　　ACTA MATHEMATICA APPLICATAE SINICA　　Jan., 1998

线性互补问题的阻尼牛顿法的有限终止性*

孙德锋　　韩继业　　赵云彬

(中国科学院应用数学研究所，北京 100080)

摘　要　在 [2] 中， Harker 和 Pang 提出了如下一个公开问题：对于线性互补问题的阻尼牛顿算法，当它收敛时，算法是否能在有限步内终止？本文对此问题给出一个肯定回答，而且进一步给出一个新的求解一般线性互补问题的有限终止算法．这个算法避免了阻尼牛顿算法可能不收敛的情形．

关键词　线性互补问题，阻尼牛顿法，有限终止性， B- 可微

1　引言

最近， Pang[6] 研究了一种求解 B- 可微方程组的全局牛顿法并且讨论了它在数学规划领域中的应用．在 [2] 中， Harker 和 Pang 分析了上述算法应用于线性互补问题的情形．在比较了阻尼牛顿法和分块转轴方法 [4,5] 以后，他们提出了如下的未决问题：如果迭代点列收敛，阻尼牛顿法是否在有限步终止 [2,275 页]？本文中我们将给出确定的答案．我们将首先回顾求 B- 可微方程组的阻尼牛顿法及其应用于线性互补问题的情形，其次我们讨论了解线性互补问题的阻尼牛顿法的有限终止性．最后我们提出了一种新的解线性互补问题的有限终止算法，它避免了阻尼牛顿法的可能不收敛的情形．

2　阻尼牛顿法及其有限终止性

考虑如下线性互补问题 (记作 LCP(q,M)):

$$w=q+Mz\geq 0,\qquad z\geq 0,\quad w^{\mathrm{T}}z=0, \tag{2.1}$$

其中 $q\in R^n$ 是一给定向量， M 是一给定的 $n\times n$ 矩阵．矩阵 M 被称为 P- 矩阵，如果它的每一个主子式为正．如果 M 是 P- 矩阵，则 LCP(q,M) 有唯一解 [5,7]．易知 [6]LCP(q,M) 等价于求解如下的分片线性方程组

$$H(z)=\min\{z,q+Mz\}=0, \tag{2.2}$$

其中 “min” 算子是取各分量最小．虽然函数 $H(z)$ 不是可微的，但它满足所谓的 B- 可微性质．

定义 2.1　函数 $H:R^n\to R^n$ 称为在点 z 处 B- 可微，如果 (i) H 在 z 的一邻域内 Lipschitz 连续， (ii) 存在一正齐次函数 $BH(z,\cdot):R^n\to R^n$(称为 H 在 z 点的 B- 导

本文1995 年 3 月 29 日收到．1997年 5 月 15 日收到修改稿．
*　国家自然科学基金和中国科学院管理、决策和信息系统实验室资助项目．

数), 使得

$$\lim_{v\to 0}\frac{H(z+v)-H(z)-BH(z)v}{\|v\|}=0.$$

函数 H 被称为在集合 S 上 B- 可微，如果 H 在 S 中每一点都 B- 可微.

对于任意 $z\in R^n$, 定义

$$\begin{cases}\alpha(z)=\{i:(q+Mz)_i<z_i\},\\ \beta(z)=\{i:(q+Mz)_i=z_i\},\\ \gamma(z)=\{i:(q+Mz)_i>z_i\}.\end{cases}\tag{2.3}$$

根据 [6] 可知，由 (2.2) 给出的函数 H 在每一点都是 B- 可微，且其 B- 导数的第 i 个分量为

$$(BH(z)v)_i=\begin{cases}M_i^{\mathrm T}v, & \text{如果 } i\in\alpha(z),\\ \min\{M_i^{\mathrm T}v,v_i\}, & \text{如果 } i\in\beta(z),\\ v_i, & \text{如果 } i\in\gamma(z),\end{cases}\tag{2.4}$$

其中 $M_i^{\mathrm T}$ 表示 M 的第 i 行向量.

记 $g:R^n\to R$ 为 H 的范数函数，即

$$g(z)=\frac{1}{2}H(z)^{\mathrm T}H(z).\tag{2.5}$$

对于一般的 B- 可微方程组

$$H(z)=0,\tag{2.6}$$

Pang[6] 给出了如下求解 (2.6) 的阻尼牛顿法.

阻尼牛顿法　设 $z^0\in R^n$ 是任一初始向量. 设 s,μ 和 σ 为给定的常数，且 $s\in(0,+\infty)$, $\mu\in(0,1)$ 及 $\sigma\in(0,1/2)$. 若点 z^k 满足 $H(z^k)\neq 0$, 则通过如下两步来得到 z^{k+1}:

第 1 步　解牛顿方程组

$$H(z^k)+BH(z^k)d=0,\tag{2.7}$$

得到方向 d^k.

第 2 步　令 $\lambda_k=\mu^{m_k}s$, 其中 m_k 是使得下面不等式成立的最小非负整数 m,

$$g(z^k)-g(z^k+\mu^m sd^k)\geq 2\sigma\mu^m sg(z^k).\tag{2.8}$$

令 $z^{k+1}=z^k+\lambda_k d^k$.

下面我们假设 H 具有形式 (2.2).

定理 2.1 [2]　设 M 是一 $n\times n$ 的 P- 矩阵，且 q,z^0 是任意向量，则由求解 (2.2) 的阻尼牛顿法所产生的序列 $\{z^k\}$ 被唯一确定、有界且满足

$$\|\min\{z^{k+1},q+Mz^{k+1}\}\|<\|\min\{z^k,q+Mz^k\}\|.\tag{2.9}$$

而且， $\{z^k\}$ 的任一聚点 $\bar z$ 是 LCP (q,M) 的解，当且仅当

$$\bar z_i=(q+M\bar z)_i=0,\qquad \forall\, i\in\beta(\bar z).$$

对 $z\in R^n$, 如果 $\beta(z)$ 为空集，则 z 被称为一非退化向量. 设 z^k 给定且 d^k 是 (2.7) 的解，对每一 k, 定义

$$w^k=q+Mz^k.$$

容易验证序列 $\{w^k\}$ 满足下面的迭代公式:

$$w^{k+1}=(1-\lambda_k)w^k+\lambda_k u^k,$$

其中 $u^k=q+Mv^k$, $v^k=z^k+d^k$, 且 d^k 满足 (2.7). 由定义, 序列 $\{z^k\}$ 满足类似的迭代公式:

$$z^{k+1}=(1-\lambda_k)z^k+\lambda_k v^k. \tag{2.10}$$

定义

$$\rho=\min\{\rho_1,\rho_2\}, \tag{2.11}$$

其中

$$\rho_1=\min\left\{\frac{z_i^k-w_i^k}{z_i^k-w_i^k-v_i^k}:\ i\in\alpha(z^k),\ v_i^k<0\right\}, \tag{2.12}$$

$$\rho_2=\min\left\{\frac{w_i^k-z_i^k}{w_i^k-z_i^k-u_i^k}:\ i\in\gamma(z^k),\ u_i^k<0\right\}. \tag{2.13}$$

引理 2.1 [2] 设 z^k 是一非退化向量, 且不是 LCP(q,M) 的解. 设 $\sigma\in(0,1/2)$, 方向向量 d^k 满足 (2.7), 则存在 $\delta>0$, 使得对每一 $t\in(\rho,\rho+\delta]$, 向量 $z_t=z^k+td^k$ 非退化, 且满足

$$g(z^k)-g(z_t)\geq 2\sigma t g(z^k).$$

现在我们可以叙述 Harker 和 Pang[2] 的求解 (2.2) 的修正阻尼牛顿法.

求 LCP (q,M) 的修正阻尼牛顿法: 设 z^0 是任一非退化向量. 设 μ 和 σ 同前. 给定非退化向量 z^k 满足 $H(z^k)\neq 0$, 则通过执行如下三步来得到 z^{k+1}:

第 1 步 求解 (2.7) 得 d^k.

第 2 步 检验 $v^k:=z^k+d^k$ 是否为 LCP(q,M) 的解; 如果是, 终止, 否则到下一步.

第 3 步 计算系数 ρ 并令 $\lambda_k=\rho+\varepsilon$, 其中 ε 是一正数使得 $z^{k+1}=z^k+\lambda_k d^k$ 非退化且

$$g(z^k)-g(z^{k+1})\geq 2\sigma\lambda_k g(z^k).$$

在 [2] 中, Harker 和 Pang 讨论了经典的转轴方法 [4,5] 与阻尼牛顿法的关系. 实际上, 虽然线搜索初看起来不像一个转轴法则, 但可以理解为不同于经典方法的一种转轴法则. 然而定理 2.1 的收敛性却是以极限形式给出的. 因而求解 (2.2) 的阻尼牛顿法在收敛情况下是否在有限步终止就成了一个未决问题 [2]. 这里我们将对这个问题给出一确定的答案.

对任意 $z\in\mathbb{R}^n$, 定义 $\partial_b H(z)$ 为所有具有如下形式的矩阵 V 组成的集合:

$$V_i=\begin{cases} M_i, & \text{如果 } i\in\alpha(z),\\ M_i,\ \text{或 } e_i & \text{如果 } i\in\beta(z),\\ e_i, & \text{如果 } i\in\gamma(z),\end{cases} \tag{2.14}$$

其中 e_i 是 $\mathbb{R}^n$ 中的第 i 个单位向量.

定理 2.2 假设 z^* 是 LCP(q,M) 的一个解并且所有 $V^*\in\partial_b H(z^*)$ 非奇异, 则存在 z^* 的一个邻域 N 使得当 $z\in N$ 时, 所有 $V\in\partial_b H(z)$ 非奇异且

$$z^*=z+d, \tag{2.15}$$

其中 $d=-V^{-1}H(z)$, $V\in\partial_b H(z)$.

证　选取 z^* 的一个邻域 N 使得当 $z \in N$ 时，有

$$\alpha(z^*) \subseteq \alpha(z), \qquad \gamma(z^*) \subseteq \gamma(z) \qquad 及 \qquad \beta(z) \subseteq \beta(z^*), \tag{2.16}$$

则对所有 $z \in N$,

$$\partial_b H(z) \subseteq \partial_b H(z^*). \tag{2.17}$$

这意味着所有 $V \in \partial_b H(z)$ 非奇异.

因为 z^* 是 LCP(q,M) 的一个解，所以对每一个 $i \in \{1,\cdots,n\}$,

$$0 = H_i(z^*) = \begin{cases} M_i^{\mathrm{T}} z^* + q_i, & 如果\ i \in \alpha(z^*) \cup \beta(z^*), \\ z_i^*, & 如果\ i \in \gamma(z^*) \cup \beta(z^*). \end{cases} \tag{2.18}$$

对任意 $z \in N$,

$$H_i(z) = \begin{cases} M_i^{\mathrm{T}} z + q_i, & 如果\ i \in \alpha(z) \cup \beta(z), \\ z_i, & 如果\ i \leq \gamma(z) \cup \beta(z), \end{cases} \qquad i = 1,\cdots,n. \tag{2.19}$$

从 (2.16) 及 $\alpha(z) \cup \beta(z) \cup \gamma(z) = \alpha(z^*) \cup \beta(z^*) \cup \gamma(z^*)$ 知，对任意 $z \in N$,

$$\alpha(z) \cup \beta(z) \subseteq \alpha(z^*) \cup \beta(z^*), \qquad \gamma(z^*) \cup \beta(z^*) \subseteq \gamma(z^*) \cup \beta(z^*). \tag{2.20}$$

由 (2.18)-(2.20),

$$\begin{aligned} & H_i(z) = H_i(z) - H_i(z^*) \\ =& \begin{cases} M_i^{\mathrm{T}}(z - z^*), & 如果\ i \in \alpha(z) \cup \beta(z), \\ e_i^{\mathrm{T}}(z - z^*), & 如果\ i \in \gamma(z) \cup \beta(z), \end{cases} \qquad i = 1,\cdots,n. \end{aligned} \tag{2.21}$$

结合 (2.14)、(2.16) 及 (2.21), 对任意 $V \in \partial_b H(z)$ 有

$$H_i(z) - V_i^T(z - z^*) = 0, \qquad i \in \alpha(z) \cup \beta(z) \cup \gamma(z),$$

这意味着

$$H(z) = V(z - z^*). \tag{2.22}$$

故

$$z^* = z - V^{-1}H(z) = z + d,$$

这就证明了 (2.15) 式.

推论 2.1　设 M 为一 $n \times n$ 的 P- 矩阵， $q, z \in \mathbb{R}^n$ 为任意的向量. 假设 z^* 是 LCP(q,M) 的解且在阻尼牛顿算法中取 $s = 1$. 则阻尼牛顿序列 $\{z^k\}$ 唯一确定、有界且满足 (2.9), 并且存在 z^* 的一邻域 N 使得当 $z^* \in N$ 且 $H(z^k) \neq 0$, 算法将终止于 $z^{k+1} = z^*$.

证　首先易验证当 M 是一 $n \times n$ 的 P- 矩阵时，所有 $V \in \partial_b H(z^*)$ 非奇异. 对任意 $d \in \mathbb{R}^n$, 易验证存在 $V \in \partial_b H(z)$, 使得

$$BH(z)d = Vd.$$

结合定理 2.1 及 2.2 即得结论.

对修正的阻尼牛顿法，类似于推论 2.1 可给出类似的结果．这里我们将不再写出．注意定理 2.1 和推论 2.1 都没有断言阻尼牛顿法必收敛到 LCP(q,M) 的唯一解 z^*. 推论 2.1 只是说明算法将在有限步终止，如果 $\{z^k\}$ 在极限意义下收敛到 z^*, 这就解决了了 [2] 中的一个未决问题．在定理 2.1 的条件下是否存在一个聚点不是 LCP(q,M) 的解，这是 Harker 和 Pang[2] 提出的另一个未决问题．由于 (2.6) 定义的 $H(z)$ 是非光滑的，全局收敛性一般很难保证，但目前我们还不能给出一个例子来说明这一点．作为一种解决后一未决问题的办法，在下面一节中我们将给出求解一类线性投影方程组 (包括 LCP(q,M)) 的新的处理途径，以避免阻尼牛顿算法可能不收敛的情形．

3 自校正的阻尼牛顿方法

考虑方程

$$H(z)=z-\Pi_X[z-(Mz+q)]=0, \tag{3.1}$$

其中 Π_X 是 X 上的正交投影算子，且 X 是 $\mathbb{R}^n$ 中的非空闭凸集．定义

$$d(z)=(I+M^{\mathrm{T}})H(z). \tag{3.2}$$

何炳生 [3] 在 M 半正定时给出了几类求解 (3.1) 的全局收敛的投影收缩 (PC) 算法．基于 [3] 中的定理 1, 我们容易给出如下的算法．

3.1 求 (3.1) 的 PC 算法:

设 $\gamma\in(0,2)$ 是一常数且 G 为一对称正定矩阵．任给 x^0, 对 $k=0,1,\cdots$, 若 $H(x^k)\neq 0$, 则计算

$$x^{k+1}=x^k-\gamma\rho(x^k)G^{-1}d(x^k), \tag{3.3}$$

其中

$$\rho(x^k)=\frac{\|(H(x^k)\|^2}{\|G^{-1}d(x^k)\|_G^2}.$$

这里对 $y\in\mathbb{R}^n$, 令 $\|y\|_G^2=y^{\mathrm{T}}Gy$.

定理 3.1 假设 $X^*=\{z^*\in X\mid z^*是(3.1)的解\}\neq\emptyset$ 并且 M 为半正定．则由 PC 方法产生的点列 $\{x_i\}$ 将收敛到 X^* 中的一点 x^*.

证 类似于 [3].

注意上述算法具有全局收敛性，但一般不在有限步终止．下面我们假设

$$X=\{z\in\mathbb{R}^n\mid l\leq z\leq u\}, \tag{3.4}$$

其中 $l,u\in\{\mathbb{R}\cup\{\infty\}\}^n$ 且 $l\leq u$. 当 $X=\mathbb{R}^n_+$ 时，

$$H(z)=z-\Pi_X\left[z-(Mz+q)\right]=\min(z,Mz+q).$$

对任意 $z\in\mathbb{R}^n$, 定义 $\partial_b H(z)$ 为由具有如下形式的矩阵 V 组成:

$$V_i=\begin{cases} M_i, & \text{如果 } l_i<z_i-(Mz+q)_i<u_i,\\ M_i, \text{ 或 } e_i, & \text{如果 } z_i-(Mz+q)_i=l_i\ (\text{或 } u_i),\\ e_i, & \text{如果 } z_i-(Mz+q)_i<l_i\ (\text{或 } u_i).\end{cases}$$

3.2 自校正的阻尼牛顿方法：

第 0 步　设 $z^0 \in \mathbb{R}^n$ 为一任意初始向量．设 ε, μ, σ 及 γ 为给定的常数且满足 $\varepsilon \in (0,1)$，$\mu \in (0,1), \sigma \in (0,1/2)$ 及 $\gamma \in (0,2)$．$k := 0$．

第 1 步　如果 $H(z^k) \neq 0$，选取 $V^k \in \partial_b H(z^k)$．如果 V^k 奇异；转第 5 步；否则解

$$H(z^k) + V^k s^k = 0, \tag{3.5}$$

得 s^k．如果

$$g(z^k + s^k) \leq (1 - 2\sigma) g(z^k). \tag{3.6}$$

令 $z^{k+1} = z^k + s^k$，转第 6 步；否则转第 2 步．

第 2 步　如果 $H(z^k)^{\mathrm{T}} B H(z^k) s^k \leq -2\sigma g(z^k)$，令 $d^k = s^k$ 并转第 4 步；否则转第 3 步．

第 3 步　如果 $H(z^k)^{\mathrm{T}} B H(z^k)(-s^k) \leq -2\sigma g(z^k)$，令 $d^k = -s^k$ 并转第 4 步；否则转第 5 步．

第 4 步 (自校正步)　设 $\lambda_k = \mu^{m_k}$，其中 m_k 是使得

$$g(z^k) - g(z^k + \mu^k d^k) \geq -2\mu^m H(z^k)^{\mathrm{T}} B H(z^k) d^k$$

或

$$\mu^m < \varepsilon$$

成立的最小非负整数 m．如果 $\lambda_k \geq \varepsilon$，令 $z^{k+1} = z^k + \lambda_k d^k$ 且转第 6 步；否则转第 5 步．

第 5 步　令 $x^0 := z^k$ 及 $i := 0$．把 x^0 作为初始向量，利用 PC 方法直到得到序列 $\{x^0, x^1, \cdots, x^{i(k)}\}$ 使得 $i(k)$ 是第 1 个使得下式成立的最小正整数 i，

$$g(x^i) \leq (1 - 2\sigma) g(x^0).$$

令 $z^{k+1} = x^{i(k)}$ 且转第 6 步．

第 6 步　$k := k + 1$．转第 1 步．

注 3.1　如果 $X^* \neq \emptyset$ 且 M 半正定，由定理 3.1 知 $i(k)$ 在有限步内可得．

定理 3.2　假设 X^* 非空且 M 半正定．如果 $X^0 = \{z \in \mathbb{R}^n \mid g(z) \leq g(z^0)\}$ 有界，则由自校正的阻尼牛顿法产生的序列 $\{z^k\}$ 有定义且每一聚点都是 (3.1) 的解．进一步，如果所有 $V^* \in \partial_b H(z^*)$，$z^* \in X^*$ 非奇异，则在有限步内算法将终止于 $z^{k+1} \in X^*$．

证　由自校正的阻尼牛顿方法知，

$$g(z^{k+1}) \leq (1 - 4\sigma^2 \varepsilon) g(z^k) \leq (1 - 4\sigma^2 \varepsilon)^{k+1} g(z^0).$$

故

$$\lim_{k \to \infty} g(z^k) = 0. \tag{3.7}$$

由 (3.7) 及 X^0 的有界性知 $\{z^k\}$ 有界且 $\{z^k\}$ 的每一聚点都是 (3.1) 的解．

如果所有 $V^* \in \partial_b H(z^*)$，$z^* \in X^*$ 为非奇异，则类似于定理 2.2 的证明我们可知对每一 $z^* \in X^*$ 存在一个邻域 $N(z^*)$ 使得当 $^k \in N(z^*)$ 时，算法将终止于 $z^{k+1} = z^*$．故本定理结论成立．

推论 3.1　如果 M 是正定阵，则自校正的阻尼牛顿方法将在有限步内终止．

注意到在自校正的阻尼牛顿方法中每步仅需解一线性方程组，并且初始点任意点．但修正的阻尼牛顿法需要一非退化的初始点．在 [1] 中，Coleman 和 Hulbert 在 $X = [l, u]$ ($l = (-1, \cdots, -1)^{\mathrm{T}}$，$u = (1, \cdots, -1)^{\mathrm{T}}$) 时提出一种求解 (3.1) 的全局收敛的算法，其

中 M 为对称正定阵. 在严格互补性条件下，他们证明了其算法的超线性收敛性. 这里如果 M 正定 (不一定对称) 我们的算法就将在有限步内终止 (不需要严格互补性).

参 考 文 献

[1] T.F. Coleman and L.A. Hulbert. A Globally and Superlinearly Convergent Algorithm for Convex Quadratic Programming with Simple Bounds. *SIAM J. on Optimization*, 1993, 3: 298–321.

[2] P.T. Harker and J.-S. Pang. A Damped-Newton Method for the Linear Complementarity Problem. Lectures in Applied Mathematics, 1990, 26: 265–284.

[3] B.S. He. Solving a Class of Linear Projection Equations. *Numerische Mathematik*, 1994, 68: 71–80.

[4] M.M. Kostreva. Block Pivot Methods for Solving the Complementarity Problem. *Linear Algebra and Application*, 1978, 21: 207–215.

[5] K.G. Murth. Linear Complementarity, Linear and Nonlinear Programming. Berlin: Helderman-Verlag, 1988.

[6] J.-S. Pang. Newton's Method for B-differentiable Equations. *Mathematics of Operations Research*, 1990, 15: 311–341.

[7] R.W. Cottle, J.S. Pang and R.E. Stone. The Linear Complementarity Problem. New York: Academic, 1992.

ON THE FINITE TERMINATION OF THE DAMPED-NEWTON ALGORITHM FOR LINEAR COMPLEMENTARITY PROBLEM

SUN DEFENG HAN JIYE ZHAO YUNBIN

(*Institute of Applied Mathematics, the Chinese Academy of Sciences, Beijing* 100080)

Abstract In [2], Harker and Pang proposed the following open question: whether or not the damped-Newton algorithm for solving the linear complementarity problems is finite if it converges. This paper gives an affirmative answer to this question. Moreover, a new finite termination algorithm for solving general linear complementarity problems is developed to avoid the possibility of the non-convergence of the previous damped-Newton method.

Key words Linear complementarity problem, damped-Newton method, finite termination, B-differential