应用数学和力学, 第 25 卷 第 11 期(2004 年 11 月)
Applied Mathematics and Mechanics
应用数学和力学编委会编
重 庆 出 版 社 出 版

文章编号: 1000_0887(2004) 11_1182_07

关于变分不等式的 Kantorovich 定理*

王征宇[1,2], 沈祖和[2]

(1. 南京大学 计算机软件新技术国家重点实验室, 南京 210093;
2. 南京大学 数学系, 南京 210093)

(我刊原编委吴启光推荐)

摘要: 将 Kantorovich 定理推广到变分不等式, 从而使得 Newton 迭代的收敛性、问题解的存在唯一性均可通过初始点处的可计算的条件来判断•

关 键 词: 变分不等式; Newton 迭代; 半局部收敛性; Kantorovich 定理
中图分类号: O224 文献标识码: A

1 引言及问题提出

设 $f: \Omega \subseteq R^n \to R^n$, 变分不等式即是寻求向量 $\boldsymbol{x}^* \in \Omega$ 满足

$$(\boldsymbol{y} - \boldsymbol{x}^*)^{\mathrm{T}} f(\boldsymbol{x}^*) \geqslant 0, \qquad \forall \boldsymbol{y} \in \Omega \cdot \tag{1}$$

我们记条件(1)为 $\mathrm{VI}(\Omega, f)$• 对于变分不等式 Newton 迭代法产生向量序列 $\{\boldsymbol{x}^k\}$ 使得 $\boldsymbol{x}^{k+1}$ 是第 k 个线性化问题 $\mathrm{VI}(\Omega, f^k)$ 的解, 其中

$$f^k(\boldsymbol{x}) = f(\boldsymbol{x}^k) + f'(\boldsymbol{x}^k)(\boldsymbol{x} - \boldsymbol{x}^k),$$
$$(f'(\boldsymbol{x}))_{ij} = (\partial f_i(\boldsymbol{x}) / \partial x_j) \cdot$$

正如我们熟知, 如果初始点 $\boldsymbol{x}^0$ 与(1) 的解 $\boldsymbol{x}^*$ 充分接近, 则 Newton 序列收敛到 $\boldsymbol{x}^*$; 进一步, 如果 f' 在 $\boldsymbol{x}^*$ 附近 Lipschitz 连续, 则 Newton 序列平方收敛• 我们可以在文献中看到许多关于求解变分不等式的 Newton 迭代的局部或者全局的收敛性结果[1~5], 这些结果本质上差异不大, 其条件很难验证, 所以在实践中无法应用• 本文将 Kantorovich 定理[6] 推广到变分不等式上, 结论的条件均可通过计算得到验证•

2 理 论 准 备

引理 1 令 $\boldsymbol{B} \in R^{n \times n}$ 对称, 则 $\boldsymbol{B}$ 正定当且仅当存在对称正定矩阵 $\boldsymbol{A}$ 满足

$$\| \boldsymbol{A} - \boldsymbol{B} \| < \frac{1}{\| \boldsymbol{A}^{-1} \|} \cdot \tag{2}$$

如果 $\boldsymbol{B}^{-1}$ 存在, 则

$$\| \boldsymbol{B}^{-1} \| \leqslant \frac{\| \boldsymbol{A}^{-1} \|}{1 - \| \boldsymbol{I} - \boldsymbol{A}^{-1} \boldsymbol{B} \|} \leqslant \frac{\| \boldsymbol{A}^{-1} \|}{1 - \| \boldsymbol{A}^{-1} \| \, \| \boldsymbol{A} - \boldsymbol{B} \|} \cdot \tag{3}$$

* 收稿日期: 2002_10_01; 修订日期: 2004_05_06
作者简介: 王征宇(1971—), 男, 南京人, 博士(联系人. Tel: + 86_25_83676437; E_mail: wzhengyu@ hotmail.com)•

注 1.1　此处的矩阵范数从属于某个向量范数•

证明　必要性是明显的, 我们证明充分性• 由于 $\boldsymbol{A}$ 对称正定, 设 $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ 为其特征值, 显然 $\boldsymbol{A}^{-1}$ 的特征值为 $\lambda_1^{-1} \geqslant \lambda_2^{-1} \geqslant \cdots \geqslant \lambda_n^{-1}$• 由条件(2)可得

$$\rho(\boldsymbol{A}-\boldsymbol{B}) \leqslant \|\boldsymbol{A}-\boldsymbol{B}\| < \frac{1}{\|\boldsymbol{A}^{-1}\|} \leqslant \frac{1}{\rho(\boldsymbol{A}^{-1})} = \lambda_1,$$

此处 $\rho(\cdot)$ 表示矩阵的谱半径, 由 $\boldsymbol{A}-\boldsymbol{B}$ 对称可知

$$|\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{x}| \leqslant \rho(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x},$$

故对于任意向量 $\boldsymbol{x} \neq \boldsymbol{0}$ 有

$$|\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{x}| < \lambda_1 \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x},$$

所以 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x} > \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} - \lambda_1\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \geqslant 0$, 充分性成立• 不等式(3)可由 Banach 引理直接得到[7]•

引理 2[8]　令 $\Omega \subseteq R^n$ 非空闭凸, 若 f 在 Ω 中严格单调, 则变分不等式 VI(Ω, f) 最多有一解, 若 f 在 Ω 中连续且弹单调, 则 VI(Ω, f) 有唯一解•

3　半局部收敛性分析

对于任意的实方阵 $\boldsymbol{A}$, 记 $\boldsymbol{A} = (\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})/2$, 将 $\|\cdot\|_2$ 简记为 $\|\cdot\|$, 分别记 $S(\boldsymbol{x}, t)$ 与 $S(\boldsymbol{x}, t)$ 为在 $\|\cdot\|$ 下, 球心为 $\boldsymbol{x} \in R^n$, 半径为 $t \geqslant 0$ 的开球与闭球• 以下是本文的主要结论•

定理 1　令 $f: \Omega \subseteq R^n \to R^n$, Ω 非空闭凸, $\boldsymbol{x}^0 \in \Omega$, $f'(\boldsymbol{x}^0)$ 正定, $\boldsymbol{x}^1$ 为子问题 VI(Ω, f^0) 的唯一解, 满足

$$\|f'(\boldsymbol{x}^0)^{-1}\| \leqslant \beta_0, \quad \|\boldsymbol{x}^1 - \boldsymbol{x}^0\| \leqslant \eta_0•$$

如果

$$\|f''(\boldsymbol{x})\| \leqslant \gamma, \quad \forall \boldsymbol{x} \in S(\boldsymbol{x}^0, t), \tag{4}$$

此处

$$t \geqslant t_0 = \frac{1-\sqrt{1-h_0/\alpha}}{h_0}\eta_0,$$

且

$$h_0 = \beta_0\gamma\eta_0 \leqslant \alpha = \frac{7-\sqrt{33}}{4} \approx 0.313\,86, \tag{5}$$

那么以 $\boldsymbol{x}^0$ 为起点, Newton 序列 $\{\boldsymbol{x}^m\} \subset S(\boldsymbol{x}^0, t_0)$ 收敛到 VI(Ω, f) 在 $S(\boldsymbol{x}^0, t_0)$ 中的一个解 $\boldsymbol{x}^*$•

注 3.1　由于 $f'(\boldsymbol{x}^0)$ 正定, 故 $f^0(\boldsymbol{x})$ 强单调[7], 由引理 2 可知 VI(Ω, f^0) 有唯一解, 即 $\boldsymbol{x}^1$ 有定义•

证明　如果 $\gamma = 0$, 那么对于任意的 $\boldsymbol{x} \in S(\boldsymbol{x}^0, t)$ 有 $f'(\boldsymbol{x}) = f'(\boldsymbol{x}^0)$, 故由中值定理[7]可知 $f(\boldsymbol{x}) = f^0(\boldsymbol{x})$, 此时 $\boldsymbol{x}^1$ 就是 VI(Ω, f) 的解, Newton 迭代终止• 如果 $\boldsymbol{x}^1 = \boldsymbol{x}^0$, 由于 $\boldsymbol{x}^1$ 是 VI(Ω, f^0) 的解, 故

$$(\boldsymbol{y}-\boldsymbol{x}^0)^{\mathrm{T}} f(\boldsymbol{x}^0) = (\boldsymbol{y}-\boldsymbol{x}^1)^{\mathrm{T}}[f(\boldsymbol{x}^0) + f'(\boldsymbol{x}^0)(\boldsymbol{x}^1-\boldsymbol{x}^0)] \geqslant 0, \qquad \forall \boldsymbol{y} \in \Omega,$$

这表明 $\boldsymbol{x}^0$ 就是 VI(Ω, f) 的解, 此时 Newton 迭代终止• 由于 $\beta_0 \neq 0$, 故不失一般性假定 $h_0 \neq 0$• 我们分四步证明该定理•

(a) 首先我们证明 $\boldsymbol{x}^2$ 有定义• 由等价性

$$\eta_0 \leqslant t_0 \Leftrightarrow \sqrt{1-\frac{h_0}{\alpha}} \leqslant 1-h_0 \Leftrightarrow h_0 \geqslant 2-\frac{1}{\alpha},$$

以及 $2-1/\alpha=(1-\sqrt{33})/4<0$ 可知 $\eta_0 \leqslant t_0$ 成立, 故 $\boldsymbol{x}^1 \in S(\boldsymbol{x}^0, t_0)$. 因为

$$\begin{aligned}&\|f'(\boldsymbol{x}^1)-f'(\boldsymbol{x}^0)\| = \|(f'(\boldsymbol{x}^1)+f'(\boldsymbol{x}^1)^{\mathrm{T}})/2-(f'(\boldsymbol{x}^0)+f'(\boldsymbol{x}^0)^{\mathrm{T}})/2\| \leqslant \\ &\quad \|(f'(\boldsymbol{x}^1)-f'(\boldsymbol{x}^0)\| \leqslant \gamma\|\boldsymbol{x}^1-\boldsymbol{x}^0\| \leqslant \\ &\quad \gamma\eta_0 = h_0/\beta_0 \leqslant \alpha/\beta_0 < 1/\beta_0,\end{aligned}$$

我们有

$$\|f'(\boldsymbol{x}^1)-f'(\boldsymbol{x}^0)\| < 1/\|f'(\boldsymbol{x}^0)^{-1}\|,$$

故由引理 1 可知 $f'(\boldsymbol{x}^1)$ 正定, 所以 $f'(\boldsymbol{x}^1)$ 也正定. 由引理 2 可知 $\mathrm{VI}(\Omega, f^1)$ 有唯一解, 因此 $\boldsymbol{x}^2$ 有定义.

(b) 我们估计 $\|\boldsymbol{x}^2-\boldsymbol{x}^1\|$ 并证明当 $\boldsymbol{x}^0$ 被 $\boldsymbol{x}^1$ 替换后, 条件(5) 仍旧成立. 由不等式(3)我们得到

$$\|f'(\boldsymbol{x}^1)^{-1}\| \leqslant \frac{\beta_0}{1-\beta_0\eta_0\gamma} = \frac{\beta_0}{1-h_0} = \beta_1.$$

由于 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \Omega$, $\boldsymbol{x}^1$ 是 $\mathrm{VI}(\Omega, f^0)$ 的解, $\boldsymbol{x}^2$ 是 $\mathrm{VI}(\Omega, f^1)$ 的解, 所以

$$\begin{aligned}&(\boldsymbol{x}^2-\boldsymbol{x}^1)^{\mathrm{T}} f^0(\boldsymbol{x}^1) = (\boldsymbol{x}^2-\boldsymbol{x}^1)^{\mathrm{T}}[f(\boldsymbol{x}^0)+f'(\boldsymbol{x}^0)(\boldsymbol{x}^1-\boldsymbol{x}^0)] \geqslant 0, \\ &(\boldsymbol{x}^1-\boldsymbol{x}^2)^{\mathrm{T}} f^1(\boldsymbol{x}^2) = (\boldsymbol{x}^1-\boldsymbol{x}^2)^{\mathrm{T}}[f(\boldsymbol{x}^1)+f'(\boldsymbol{x}^1)(\boldsymbol{x}^2-\boldsymbol{x}^2)] \geqslant 0.\end{aligned}$$

将这两个不等式相加并整理, 我们得到

$$\begin{aligned}&(\boldsymbol{x}^1-\boldsymbol{x}^2)^{\mathrm{T}} f'(\boldsymbol{x}^0)(\boldsymbol{x}^1-\boldsymbol{x}^2) \leqslant \\ &\quad (\boldsymbol{x}^1-\boldsymbol{x}^2)^{\mathrm{T}}[f(\boldsymbol{x}^1)-f(\boldsymbol{x}^0)+f'(\boldsymbol{x}^1)(\boldsymbol{x}^2-\boldsymbol{x}^1)-f'(\boldsymbol{x}^0)(\boldsymbol{x}^2-\boldsymbol{x}^0)] = \\ &\quad (\boldsymbol{x}^1-\boldsymbol{x}^2)^{\mathrm{T}}[(f'(\boldsymbol{x}^1)-f'(\boldsymbol{x}^0))(\boldsymbol{x}^2-\boldsymbol{x}^0)- \\ &\quad (f(\boldsymbol{x}^0)-f(\boldsymbol{x}^1)-f'(\boldsymbol{x}^1)(\boldsymbol{x}^0-\boldsymbol{x}^1))] \leqslant \\ &\quad \|\boldsymbol{x}^1-\boldsymbol{x}^2\|[\|f'(\boldsymbol{x}^1)-f'(\boldsymbol{x}^0)\|\|\boldsymbol{x}^2-\boldsymbol{x}^0\|+ \\ &\quad \|f(\boldsymbol{x}^0)-f(\boldsymbol{x}^1)-f'(\boldsymbol{x}^1)(\boldsymbol{x}^0-\boldsymbol{x}^1)\|].\end{aligned}$$

因为 $\boldsymbol{x}^1 \in S(\boldsymbol{x}^0, t_0)$, 故由 Taylor 定理[7]我们知道

$$\|f(\boldsymbol{x}^0)-f(\boldsymbol{x}^1)-f'(\boldsymbol{x}^1)(\boldsymbol{x}^0-\boldsymbol{x}^1)\| \leqslant \gamma\|\boldsymbol{x}^1-\boldsymbol{x}^0\|^2/2.$$

由于

$$\begin{aligned}&(\boldsymbol{x}^1-\boldsymbol{x}^2)^{\mathrm{T}} f'(\boldsymbol{x}^0)(\boldsymbol{x}^1-\boldsymbol{x}^2) = (\boldsymbol{x}^1-\boldsymbol{x}^2)^{\mathrm{T}} f'(\boldsymbol{x}^0)(\boldsymbol{x}^1-\boldsymbol{x}^0) \geqslant \\ &\quad \|\boldsymbol{x}^1-\boldsymbol{x}^2\|^2/\|f'(\boldsymbol{x}^0)^{-1}\| \geqslant \|\boldsymbol{x}^1-\boldsymbol{x}^2\|^2/\beta_0,\end{aligned}$$

我们有

$$\begin{aligned}&\|\boldsymbol{x}^1-\boldsymbol{x}^2\|^2/\beta_0 \leqslant \\ &\quad \|\boldsymbol{x}^1-\boldsymbol{x}^2\|\left(\frac{\gamma}{2}\|\boldsymbol{x}^1-\boldsymbol{x}^0\|^2+\gamma\|\boldsymbol{x}^1-\boldsymbol{x}^0\|+\|\boldsymbol{x}^2-\boldsymbol{x}^0\|\right) = \\ &\quad \gamma\|\boldsymbol{x}^1-\boldsymbol{x}^2\|\|\boldsymbol{x}^1-\boldsymbol{x}^0\|\left(\frac{1}{2}\|\boldsymbol{x}^1-\boldsymbol{x}^0\|+\|\boldsymbol{x}^2-\boldsymbol{x}^0\|\right) \leqslant \\ &\quad \gamma\|\boldsymbol{x}^1-\boldsymbol{x}^2\|\|\boldsymbol{x}^1-\boldsymbol{x}^0\|\left(\frac{1}{2}\|\boldsymbol{x}^1-\boldsymbol{x}^0\|+\|\boldsymbol{x}^2-\boldsymbol{x}^1\|+\|\boldsymbol{x}^1-\boldsymbol{x}^0\|\right) \leqslant \\ &\quad \gamma\|\boldsymbol{x}^1-\boldsymbol{x}^2\|\eta_0\left(\frac{3}{2}\eta_0+\|\boldsymbol{x}^2-\boldsymbol{x}^1\|\right),\end{aligned}$$

整理得

$$\|x^2-x^1\| \leqslant \frac{3h_0\eta_0}{2(1-h_0)}=\eta_1.$$

注意到

$$\rho=\frac{3}{2(1-\alpha)}=\frac{1-\alpha}{\alpha},\quad \sigma=\frac{3}{2(1-\alpha)^2}=\frac{1}{\alpha},$$

显然有 $\eta_1\leqslant\rho h_0\eta_0$. 因此由假设 $h_0\leqslant\alpha$ 我们可以得到

$$h_1=\beta_1\eta_1\gamma=\frac{\beta_0}{1-h_0}\,\frac{3}{2}\,\frac{h_0\eta_0}{1-h_0}\gamma=\frac{3}{2}\,\frac{h_0^2}{(1-h_0)^2}\leqslant\frac{3\alpha^2}{2(1-\alpha)^2}=\alpha,$$

即当 x^0 被替换成 x^1 后, 条件(5)仍旧成立.

(c) 在此我们证明 $S(x^1,t_1)\subseteq S(x^0,t_0)$, 这一事实将保证当 x^0 被替换成 x^1 后, 条件(4)仍旧成立, 此处

$$t_1=\frac{1-\sqrt{1-h_1/\alpha}}{h_1}\eta_1=\frac{\sqrt{2\alpha}-\sqrt{(2\alpha-3)h_0^2-4\alpha h_0+2\alpha}}{\sqrt{2\alpha}h_0}\eta_0-\eta_0.$$

由于

$$t_1\leqslant t_0-\eta_0\Leftrightarrow\frac{\sqrt{2\alpha}-\sqrt{(2\alpha-3)h_0^2-4\alpha h_0+2\alpha}}{\sqrt{2\alpha}h_0}\eta_0\leqslant\frac{1-\sqrt{1-h_0/\alpha}}{h_0}\eta_0\Leftrightarrow$$

$$\sqrt{(2\alpha-3)h_0^2-4\alpha h_0+2\alpha}\geqslant\sqrt{2\alpha-2h_0}\Leftrightarrow$$

$$(3-2\alpha)h_0^2+(4\alpha-2)h_0\leqslant 0\Leftrightarrow$$

$$0\leqslant h_0\leqslant\frac{2-4\alpha}{3-2\alpha}=\alpha, \tag{6}$$

故 $t_1\leqslant t_0-\eta_0$. 所以对于任意 $x\in S(x^1,t_1)$ 有

$$\|x-x^0\|\leqslant\|x-x^1\|+\|x^1-x^0\|\leqslant t_1+\eta_0\leqslant t_0$$

即 $S(x^1,t_1)\subseteq S(x^0,t_0)$.

(d) 我们证明 Newton 序列收敛到 VI(Ω,f) 的一个解, 此解存在于 $S(x^0,t_0)$. 由数学归纳法我们知道以 x^0 为初始点, Newton 序列 $\{x^m\}\subset S(x^0,t_0)$ 有定义, 并满足

$$\beta_m=\frac{\beta_{m-1}}{1-h_{m-1}},\quad \eta_m=\frac{3}{2}\,\frac{h_{m-1}\eta_{m-1}}{1-h_{m-1}},\quad h_m=\frac{3}{2}\,\frac{h_{m-1}^2}{(1-h_{m-1})^2},$$

此处 $m=1,2,\cdots$. 我们有 $h_m\leqslant\sigma h_{m-1}^2$, 即 $\sigma h_{m-1}\leqslant(\sigma h_{m-1})^2$, 这表明 $\sigma h_m\leqslant(\sigma h_0)^{2^m}$, 即

$$h_m\leqslant\frac{1}{\sigma}(\sigma h_0)^{2^m}. \tag{7}$$

重复使用不等式 $\eta_m\leqslant\rho h_{m-1}\eta_{m-1}$, 我们有

$$\eta_m\leqslant\rho h_{m-1}\eta_{m-1}\leqslant\rho^2h_{m-1}h_{m-2}\eta_{m-2}\leqslant\cdots\leqslant\rho^m h_{m-1}h_{m-2}\cdots h_0\eta_0.$$

由(7)可得

$$\eta_m\leqslant\rho^m h_{m-1}h_{m-2}\cdots h_0\eta_0\leqslant\frac{\rho^m}{\sigma^m}(\sigma h_{m-1})(\sigma h_{m-2})\cdots(\sigma h_0)\eta_0\leqslant$$

$$\left(\frac{\rho}{\sigma}\right)^m(\sigma h_0)^{2^{m-1}}(\sigma h_0)^{2^{m-2}}\cdots(\sigma h_0)\eta_0=\left(\frac{\rho}{\sigma}\right)^m(\sigma h_0)^{2^m-1}\eta_0. \tag{8}$$

由归纳法可知 $S(x^m,t_m)\subseteq S(x^{m-1},t_{m-1})$, 此处 $m=1,2,\cdots$,

$$t_m=\frac{1-\sqrt{1-h_m/\alpha}}{h_m}\eta_m.$$

很明显 $\forall p \in N$, $\boldsymbol{x}^{m+p} \in S(\boldsymbol{x}^m, t_m)$, 所以

$$\|\boldsymbol{x}^{m+p} - \boldsymbol{x}^m\| \leqslant t_m = \frac{1-\sqrt{1-h_m/\alpha}}{h_m}\eta_m \leqslant \frac{\eta_m}{\alpha}.$$

故

$$\|\boldsymbol{x}^{m+p} - \boldsymbol{x}^m\| \leqslant \frac{1}{\alpha}\left(\frac{\rho}{\sigma}\right)^m (\sigma h_0)^{2^m-1}\eta_0. \tag{9}$$

此处 $\sigma h_0 = h_0/\alpha \leqslant 1$, $\rho/\sigma = 1-\alpha < 1$, 所以$\{\boldsymbol{x}^m\}$是 Cauchy 序列, 其极限 $\boldsymbol{x}^*$ 包含于 $S(\boldsymbol{x}^m, t_m)$, 此处 $m = 0, 1, \cdots$. 由于

$$(\boldsymbol{y} - \boldsymbol{x}^{m+1})^{\mathrm{T}}[f(\boldsymbol{x}^m) + f'(\boldsymbol{x}^m)(\boldsymbol{x}^{m+1} - \boldsymbol{x}^m)] \geqslant 0, \ \forall \boldsymbol{y} \in \Omega,$$

且 f 与 f' 是连续的, 故当 $m \to \infty$ 时我们得到

$$(\boldsymbol{y} - \boldsymbol{x}^*)^{\mathrm{T}} f(\boldsymbol{x}^*) \geqslant 0, \qquad \forall \boldsymbol{y} \in \Omega$$

即 $\boldsymbol{x}^*$ 为 $\mathrm{VI}(\Omega, f)$ 的解. □

关于 Newton 迭代我们有以下进一步的结论.

定理 2 如果定理 1 的条件满足, 则以任意的 $\boldsymbol{y}^0 \in S(\boldsymbol{x}^0, \eta_0) \cap \Omega$ 为初始向量, Newton 序列$\{\boldsymbol{y}^m\}$将收敛到 $\mathrm{VI}(\Omega, f)$ 在 $S(\boldsymbol{x}^0, t^*) \cap \Omega$ 中的解 $\boldsymbol{x}^*$.

证明 由定理 3 的证明部分(a) 可知 $f'(\boldsymbol{y}^0)$ 正定, 所以子问题 $\mathrm{VI}(\Omega, f^0)$ 有唯一解, 此处

$$f^0(\boldsymbol{x}) = f(\boldsymbol{y}^0) + f'(\boldsymbol{y}^0)(\boldsymbol{x} - \boldsymbol{y}^0).$$

令 $\boldsymbol{y}^1$ 表示 $\mathrm{VI}(\Omega, f^0)$ 的唯一解, 由于 $\boldsymbol{y}^1 \in \Omega$, 所以

$$(\boldsymbol{y}^1 - \boldsymbol{x}^1)^{\mathrm{T}}[f(\boldsymbol{x}^0) + f'(\boldsymbol{x}^0)(\boldsymbol{x}^1 - \boldsymbol{x}^0)] \geqslant 0,$$
$$(\boldsymbol{x}^1 - \boldsymbol{y}^1)^{\mathrm{T}}[f(\boldsymbol{y}^0) + f'(\boldsymbol{y}^0)(\boldsymbol{y}^1 - \boldsymbol{y}^0)] \geqslant 0.$$

使用与定理 1 的证明部分(b) 类似的方法可以得到

$$\begin{aligned}(\boldsymbol{y}^1 - \boldsymbol{x}^1)^{\mathrm{T}} f'(\boldsymbol{x}^0)(\boldsymbol{y}^1 - \boldsymbol{x}^1) \leqslant \\ &(\boldsymbol{y}^1 - \boldsymbol{x}^1)^{\mathrm{T}}[f(\boldsymbol{x}^0) - f(\boldsymbol{y}^0) - f'(\boldsymbol{y}^0)(\boldsymbol{x}^0 - \boldsymbol{y}^0) + \\ &(f'(\boldsymbol{x}^0) - f'(\boldsymbol{y}^0))(\boldsymbol{y}^1 - \boldsymbol{x}^0)]\end{aligned}$$

以及

$$\|\boldsymbol{y}^1 - \boldsymbol{x}^1\| \leqslant \frac{3}{2}\frac{\beta_0\gamma\eta_0^2}{1-\beta_0\gamma\eta_0} = \eta_1$$

即 $\boldsymbol{y}^1 \in S(\boldsymbol{x}^1, \eta_1)$. 由数学归纳法可知对于任意的 $m = 0, 1 \cdots$ 有 $\boldsymbol{y}^m \in S(\boldsymbol{x}^m, \eta_m)$. 不等式(8) 表明 $\eta_m \to 0$, 此时显然有$\{\boldsymbol{y}^m\} \to \boldsymbol{x}^*$. □

由不等式(9) 我们立即得到以下的误差估计式.

定理 3 如果定理 1 的条件成立, 则

$$\|\boldsymbol{x}^m - \boldsymbol{x}^*\| \leqslant \frac{1}{\alpha}\left(\frac{\rho}{\sigma}\right)^m (\sigma h_0)^{2^{m-1}}\eta_0.$$

证明 在式(9) 中令 $p \to \infty$ 即可得证. □

定理 4 如果定理 1 的条件成立, 且 Newton 序列$\{\boldsymbol{x}^m\}$的极限存在于开球 $S(\boldsymbol{x}^0, t_0)$ 中, 则 $\boldsymbol{x}^*$ 是变分不等式 $\mathrm{VI}(\Omega, f)$ 在 $S(\boldsymbol{x}^0, t_0)$ 中的唯一解.

证明 注意到

$$\|f'(\boldsymbol{x}^0) - f'(\boldsymbol{y})\| \leqslant \gamma\|\boldsymbol{y} - \boldsymbol{x}^0\|,$$

以及

$$\gamma t_0 = \frac{1 - \sqrt{1 - h_0/\alpha}}{h_0}\gamma\eta_0 \leqslant \frac{\eta_0}{\beta_0\alpha} \leqslant \frac{1}{\beta_0},$$

故对于任意向量 $\boldsymbol{y} \in S(\boldsymbol{x}^0, t_0)$ 我们有

$$\| f'(\boldsymbol{x}^0) - f'(\boldsymbol{y}) \| < \gamma t_0 \leqslant \frac{1}{\| f'(\boldsymbol{x}^0)^{-1} \|}.$$

由引理1 可知 $f'(\boldsymbol{y})$ 正定, 所以 f 在 $S(\boldsymbol{x}^0, t_0)$ 中严格单调, 故由引理2 可知不等式在 $S(\boldsymbol{x}^0, t_0)$ 解唯一.

定理 5 若定理 1 的条件成立且 $h_0 < \alpha$, 则 $\boldsymbol{x}^*$ 是 $\mathrm{VI}(\Omega, f)$ 在 $S(\boldsymbol{x}^0, t_0)$ 中的唯一解.

证明 如果 $h_0 < \alpha$, 则由等价关系(6) 可知 $t_1 < t_0 - \eta_0$ 以及 $S(\boldsymbol{x}^1, t_1) \subseteq S(\boldsymbol{x}^0, t_0)$. 由
$$\{\boldsymbol{x}^m, \boldsymbol{x}^{m+1}, \dots\} \subset S(\boldsymbol{x}^m, t_m)$$
可知 $\boldsymbol{x}^* \in S(\boldsymbol{x}^1, t_1)$, 故 $\boldsymbol{x}^* \in S(\boldsymbol{x}^0, t_0)$, 由定理4 我们知道 $\boldsymbol{x}^*$ 是 $\mathrm{VI}(\Omega, f)$ 在 $S(\boldsymbol{x}^0, t_0)$ 中的唯一解.

定理 6 如果定理 1 的条件成立且 $h_0 < \alpha$, 则以 $S(\boldsymbol{x}^0, \eta_0)$ 中任意向量为初始点 Newton 序列 $\{\boldsymbol{y}^m\}$ 平方收敛.

证明 对于任意初始向量 $\boldsymbol{y}^0 \in S(\boldsymbol{x}^0, \eta_0)$, 由定理 2 可知 Newton 序列 $\{\boldsymbol{y}^m\}$ 收敛到 $\mathrm{VI}(\Omega, f)$ 的解 $\boldsymbol{x}^*$. 由于 $\boldsymbol{x}^* \in \Omega$, $\boldsymbol{y}^{m+1}$ 是 $\mathrm{VI}(\Omega, f^m)$ 的解, 此处

$$f^m(\boldsymbol{x}) = f(\boldsymbol{y}^m) + f'(\boldsymbol{y}^m)(\boldsymbol{x} - \boldsymbol{y}^m).$$

所以

$$(\boldsymbol{y}^{m+1} - \boldsymbol{x}^*)^{\mathrm{T}} f(\boldsymbol{x}^*) \geqslant 0,$$
$$(\boldsymbol{x}^* - \boldsymbol{y}^{m+1})^{\mathrm{T}}[f(\boldsymbol{y}^m) + f'(\boldsymbol{y}^m)(\boldsymbol{y}^{m+1} - \boldsymbol{y}^m)] \geqslant 0.$$

将二式相加, 使用与定理 1 证明部分(b)类似的方法可以得到

$$(\boldsymbol{y}^{m+1} - \boldsymbol{x}^*)^{\mathrm{T}} f'(\boldsymbol{y}^m)(\boldsymbol{y}^{m+1} - \boldsymbol{x}^*) \leqslant$$
$$(\boldsymbol{y}^{m+1} - \boldsymbol{x}^*)^{\mathrm{T}}[f(\boldsymbol{x}^*) - f(\boldsymbol{y}^m) - f'(\boldsymbol{y}^m)(\boldsymbol{x}^* - \boldsymbol{y}^m)],$$

所以

$$\| \boldsymbol{y}^{m+1} - \boldsymbol{x}^* \| \leqslant \| f'(\boldsymbol{y}^m)^{-1} \| \frac{\gamma}{2} \| \boldsymbol{y}^m - \boldsymbol{x}^* \|^2.$$

如果 $h_0 < \alpha$, 由定理2的证明过程可以看出 $\boldsymbol{y}^m \in S(\boldsymbol{x}^m, t_m)$, 所以必存在常数 M 满足 $\| f'(\boldsymbol{y}^m)^{-1} \| \leqslant M$. 设 $K = \gamma M/2$, 我们有

$$\| \boldsymbol{y}^{m+1} - \boldsymbol{x}^* \| \leqslant K \| \boldsymbol{y}^m - \boldsymbol{x}^* \|^2,$$

即 Newton 序列是平方收敛的. □

使用自动微分技术[9] 与区间计算技术[10] 本文所给出的收敛性条件可以完全通过计算得到验证, 所以在实践中易于使用. 由于非线性互补问题、广义互补问题、凸规划等问题均可归结为变分不等式, 故文中结论对于这些数学规划问题也是适用的.

[参 考 文 献]

[1] Eaves B C. A locally quadratically convergent algorithm for computing stationary point[R]. Department of Operations Research, Stanford University, 1978.

[2] Josephy N H. Newton's method for generalized equations[R]. Mathematics Research Center, University of Wisconsin, 1979.

[3] Harker P T, Pang J S. Finite_dimensional variational inequalities and non_linear complementarity problems: a survey of theory, algorithms and applications[J]. Mathematical Programming, 1990, **48**(2): 161—220.

[4] Robinson S M. Generalized equations and their solutions[J]. Mathematical Programming Study, 1979, **10**(1): 128—141.

[5] Robinson S M. Strongly regular generalized equations[J]. Mathematics of Operations Research, 1980, **5**(1): 43—62.

[6] Kantorovich L V. Functional analysis and applied mathematics[J]. Uspehi Mat Nauk, 1948, **3**(6): 89—185.

[7] Ortega J M, Rheinboldt W C. Iterative Solutions of Nonlinear Equations in Several Variables[M]. New York: Academic Press, 1970.

[8] Stampacchia G. Variational inequalities[A]. In: Stampacchia G Ed. Theory and Applications of Monotone Operators, Proceedings of the NATO Advanced Study Institute [C]. Venice: Edizioni Oderisi, Gubbio, 1968, 102—192.

[9] Rall L B. Computational Solution of Nonlinear Operator Equations[M]. New York: Wiely, 1969.

[10] Alefeld G E, Herzberger J. Introduction to Interval Computations[M]. New York and London: Academic Press, 1983.

Kantorovich Theorem for Variational Inequalities

WANG Zheng_yu[1,2], SHEN Zu_he[2]

(1. State Key Laboratory for Novel Software Technology, Nanjing University,
2. Department of Mathematics, Nanjing University, Nanjing 210093, P.R. China)

Abstract: Kantorovich theorem was extended to variational inequalities by which the convergence of Newton iteration, the existence and uniqueness of the solution of the problem can be tested via computational conditions at the initial point.

Key words: variational inequality; Newton iteration; semilocal convergence; kantorovich theorem