广西科学 Guangxi Sciences 2006，13(3)：190～193

非线性互补问题的 Derivative-Free 下降方法

Derivative-free Descent Method For Nonlinear Complementarity Problems

蒋利华[1,2]，马昌凤[1]，徐安农[1]

JIANG Li-hua[1,2]，MA Chang-feng[1]，XU An-nong[1]

（1.桂林电子工业学院计算科学与数学系，广西桂林 541004；2.安徽理工大学数理系，安徽淮南 232001）

（1. Department of Computational Science Mathematics，Guilin University of Electron Technology，Guilin，Guangxi，541004，China；2. Department of Mathematics Physics，Anhui University of Science Technology，Huainan，Anhui，232001，China）

摘要：基于非线性互补问题（$NCP(F)$）的约束极小化变形，构造一种新的 merit 函数，将原始的 $NCP(F)$ 问题转化为约束极小化问题，构造相应的 derivative-free 下降算法. 在 merit 函数严格单调的条件下证明 derivative-free 下降算法的合理性以及整体收敛性.

关键词：非线性互补问题　merit 函数　derivative-free 下降算法　整体收敛性

中图法分类号：O224.2　**文献标识码**：A　**文章编号**：1005-9164(2006)03-0190-04

Abstract：A globally convergent derivative-free descent method for solving nonlinear complementarity problem （$NCP(F)$） is proposed basing on its equivalent formulation of minimization. A new merit function is introduced here. The reasonable and global convergence of the method is verified under the merit function which is monotone function.

Key words：nonlinear complementarity problem，merit function，derivative-free descent algorithm，global convergence

非线性互补问题（$NCP(F)$），就是求 $x \in R^n$ 使得

$$x \geqslant 0, y = F(x) \geqslant 0, x^T y = 0, \tag{1}$$

其中 $F: R^n \to R^n$ 连续可微[1,2].

众所周知，非线性互补问题在数学，经济，物理及工程科学等多个领域中有着广泛的应用. 解决非线性互补问题(1)有很多数值方法，包括投影法，不动点迭代法，B-可微牛顿法，光滑牛顿法等[3~10]. 本文利用互补函数将互补问题转化为约束极小化问题，构造相应的 derivative-free 下降算法，并证明该算法的合理性及整体收敛性.

1　一种新的 merit 函数

为了解决 $NCP(F)$，我们给出如下的 merit 函数，$\Psi_\alpha: R^n \to R$，

收稿日期：2005-11-07
修回日期：2006-03-13
作者简介：蒋利华(1979-)，女，安徽怀远人，硕士，主要从事互补问题的数值方法的研究。

$$\Psi_\alpha(x) = \sum_{i=1}^{n} \phi_\alpha(x_i, F_i(x)), \tag{2}$$

其中 $\phi_\alpha: R^2 \to R$ 是增广 Fischer-Burmeister NCP 函数：

$$\phi_\alpha(a,b) = \frac{1}{2}(ab)_+^2 + \frac{\alpha}{2}(\sqrt{a^2+b^2} - a - b)^2 \tag{3}$$

且 $\alpha \geqslant 0$ 为实的参数.

类似文献[11]，此函数具有如下的一些性质.

命题 1　由(3)定义的 NCP 函数 $\phi_\alpha(a,b)$ 满足

(i) $\phi_\alpha(a,b) = 0 \Leftrightarrow a \geqslant 0, b \geqslant 0, ab = 0$； (4)

(ii) $\phi_\alpha(a,b) \geqslant 0 \; \forall \; (a,b)^T \in R^2$； (5)

(iii) 对 $\forall \; (a,b)^T \in R^2, \phi_\alpha(a,b)$ 连续可导，且

$$\frac{\partial \phi_\alpha}{\partial a}(0,0) = \frac{\partial \phi_\alpha}{\partial b}(0,0) = 0. \tag{6}$$

$$\frac{\partial \phi_\alpha}{\partial a}(a,b) = b(ab)_+ + \alpha\left(\frac{a}{\sqrt{a^2+b^2}} - 1\right) \cdot (\sqrt{a^2+b^2} - a - b), \tag{7}$$

$$\frac{\partial \phi_\alpha}{\partial b}(a,b) = a(ab)_+ + \alpha\left(\frac{b}{\sqrt{a^2+b^2}} - 1\right)(\sqrt{a^2+b^2} - a - b). \tag{8}$$

(iv) $\frac{\partial\phi_\alpha}{\partial a}(a,b)\frac{\partial\phi_\alpha}{\partial b}(a,b)\geqslant 0$，且$\frac{\partial\phi_\alpha}{\partial a}(a,b)\frac{\partial\phi_\alpha}{\partial b}(a,b)=0\Leftrightarrow a\geqslant 0,b\geqslant 0,ab=0$.

$\forall\ (a,b)^T\in R^2$.

(v) $\frac{\partial\phi_\alpha}{\partial a}(a,b)=0\Leftrightarrow\frac{\partial\phi_\alpha}{\partial b}(a,b)=0\Leftrightarrow\phi_\alpha(a,b)=0$.

证明 (i)～(iii) 显然成立，我们只要证明(iv)成立，(v) 也显然成立，因而我们只需考虑(iv). 当$(a,b)=(0,0)$时，(iv) 显然成立. 当$(a,b)\neq(0,0)$时，

$$\frac{\partial\phi_\alpha}{\partial a}(a,b)\frac{\partial\phi_\alpha}{\partial b}(a,b)=ab(ab)_+^2+\alpha(ab)_+(\sqrt{a^2+b^2}-a-b)^2+\alpha^2(\frac{a}{\sqrt{a^2+b^2}}-1)(\frac{b}{\sqrt{a^2+b^2}}-1)(\sqrt{a^2+b^2}-a-b)^2. \quad (9)$$

因为$ab(ab)_+^2\geqslant 0,\alpha(ab)_+(\sqrt{a^2+b^2}-a-b)^2\geqslant 0$，$\frac{a}{\sqrt{a^2+b^2}}-1\leqslant 0,\frac{b}{\sqrt{a^2+b^2}}-1\leqslant 0$，显然$\frac{\partial\phi_\alpha}{\partial a}(a,b)\frac{\partial\phi_\alpha}{\partial b}(a,b)\geqslant 0$成立. 由$\frac{\partial\phi_\alpha}{\partial a}(a,b)\frac{\partial\phi_\alpha}{\partial b}(a,b)=0$可知$\frac{\partial\phi_\alpha}{\partial a}(a,b)=0$或$\frac{\partial\phi_\alpha}{\partial b}(a,b)=0$，下面考虑由$\frac{\partial\phi_\alpha}{\partial a}(a,b)=0$可得$a\geqslant 0,b\geqslant 0,ab=0$. $(a,b)=(0,0)$显然成立，当$(a,b)\neq(0,0)$时，由$\frac{\partial\phi_\alpha}{\partial a}(a,b)=b(ab)_++\alpha(\frac{a}{\sqrt{a^2+b^2}}-1)(\sqrt{a^2+b^2}-a-b)=0$可知

$$b(ab)_+=-\alpha(\frac{a}{\sqrt{a^2+b^2}}-1)(\sqrt{a^2+b^2}-a-b). \quad (10)$$

假设(10)式两边都不等于零，那么我们有$b\neq 0$. 如果$b<0$，(10)式左边非正而右边为正，矛盾；如果$b>0$，为使$b(ab)_+\neq 0$，a必须为正，因而在这种情况下，(10)式左边为正而右边为负，矛盾. 故(10)式两边只能都为零，易见$\phi_\alpha(a,b)=0$成立. 类似可证由$\frac{\partial\phi_\alpha}{\partial b}(a,b)=0$，可得$a\geqslant 0,b\geqslant 0,ab=0$. 而由$a\geqslant 0,b\geqslant 0,ab=0$，易得$\frac{\partial\phi_\alpha}{\partial a}(a,b)\frac{\partial\phi_\alpha}{\partial b}(a,b)=0$. 证明完毕.

根据$\phi_\alpha(x)$的性质，可知解$NCP(F)$等价于解约束极小化问题.

$$\min_{x\in R^n}\Psi_\alpha(x). \quad (11)$$

2 Derivative-free 下降算法的构造及整体收敛性

为了证明方便，首先给出如下的引理.

引理1 如果$F(x)$在集合$X\subset R^n$连续可微，则

$F(x)$在X上单调$\Leftrightarrow(y-x)^T\nabla F(x)(y-x)\geqslant 0,\forall\ x,y\in X$ (12)

$F(x)$在X上强单调$\Leftrightarrow(y-x)^T\nabla F(x)(y-x)\geqslant\mu\|y-x\|^2,\forall\ x,y\in X,\mu>0$ (13)

假设非线性函数$F(x)$连续可微，引入如下记号：

$$\frac{\partial\phi_\alpha}{\partial a}(x,F(x))=(\frac{\partial\phi_\alpha}{\partial a}(x_1,F_1(x)),\cdots,\frac{\partial\phi_\alpha}{\partial a}(x_n,F_n(x)))^T,$$

$$\frac{\partial\phi_\alpha}{\partial b}(x,F(x))=(\frac{\partial\phi_\alpha}{\partial b}(x_1,F_1(x)),\cdots,\frac{\partial\phi_\alpha}{\partial b}(x_n,F_n(x)))^T.$$

这样，merit 函数$\Psi_\alpha(x)$的梯度可以表示为

$$\nabla\Psi_\alpha(x)=\frac{\partial\phi_\alpha}{\partial a}(x,F(x))+\nabla F(x)^T\frac{\partial\phi_\alpha}{\partial b}(x,F(x)). \quad (14)$$

考虑搜索方向：

$$d^k=-\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k)), \quad (15)$$

根据命题1的(v)，有

$\|d^k\|=0\Leftrightarrow x^k$是$NCP(F)$的解. (16)

下面证明(13)式所定义的搜索方向是 merit 函数$\Psi_\alpha(x)$在$x^k\in R^n$处的下降方向.

引理2 设非线性函数$F(x)$为连续可微的单调函数，且$\alpha>0$. 如果x^k不是问题$NCP(F)$的解，则(13)式所定义的搜索方向为merit 函数$\Psi_\alpha(x)$在点$x^k\in R^n$处的下降方向. 即满足：

$$\nabla\Psi_\alpha(x^k)^Td^k<0. \quad (17)$$

此外，若$F(x)$在R^n上强单调，则存在$\mu>0$，使

$$\nabla\Psi_\alpha(x^k)^Td^k<\mu\|d^k\|^2. \quad (18)$$

证明 如果x^k不是(1)式的解，据命题1的(iv)知，

$$\frac{\partial\phi_\alpha}{\partial a}(x^k,F(x^k))^T\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k))>0.$$

于是，据(12)式，(14)式以及d^k的定义，有

$$\nabla\Psi_\alpha(x^k)^Td^k=-\frac{\partial\phi_\alpha}{\partial a}(x^k,F(x^k))^T\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k))-\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k))^T\nabla F(x^k)\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k))<-\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k))^T\nabla F(x^k)\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k))\leqslant 0. \quad (19)$$

如果$F(x)$强单调，根据(13)式有

$$\nabla\Psi_\alpha(x^k)^Td^k<-\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k))^T\nabla F(x^k)\cdot\frac{\partial\phi_\alpha}{\partial b}(x^k,F(x^k))\leqslant-\mu\|d^k\|^2.$$

2.1 算法

步骤1：给定初始点$x^0\in R^n,\varepsilon>0,\gamma\in(0,1)$，置$k:=0$；

步骤2：如果$\Psi_\alpha(x^k)<\varepsilon$，则$x^k$就是$NCP(F)$的近似解；

步骤3：计算$x^{k+1}=x^k+\gamma^{l_k}d^k$，其中$l_k$为满足下

式的最小非负整数 l,

$$\Psi_\alpha(x^k+\gamma^l d^k)-\Psi_\alpha(x^k)\leqslant-\gamma^{2l}\|d^k\|^2; \quad (20)$$

步骤 4:令 $k:=k+1$ 转到步骤 1.

2.2 算法的合理性和整体收敛性

定理 1 设 $F(x)$ 是 R^n 上的连续可微函数,且 $\alpha\geqslant 0$,则算法是定义合理的.

证明 我们只需证明步骤 3 是定义合理的便可,即只要证明存在有限的非负整数 l 使得(20)式成立.若不然,对任意的非负整数都有

$$\Psi_\alpha(x^k+\gamma^l d^k)-\Psi_\alpha(x^k)>-\gamma^{2l}\|d^k\|^2.$$

两边同除以 γ^l 并令 $l\to\infty$ 可得

$$\lim_{l\to\infty}\frac{\Psi_\alpha(x^k+\gamma^l d^k)-\Psi_\alpha(x^k)}{\gamma^l}>-\lim_{l\to\infty}\gamma^l\|d^k\|^2=0,$$

即

$$\lim_{l\to\infty}\frac{\Psi_\alpha(x^k+\gamma^l d^k)-\Psi_\alpha(x^k)}{\gamma^l d^k}d^k=\nabla\Psi_\alpha(x^k)d^k>0.$$

这与引理 2 矛盾,因此存在一个最小的非负整数 l_k 使得(20)式成立.算法是合理的.

下面说明算法所生成的序列 x^k 的每一个聚点都是 $NCP(F)$ 的解.

定理 2 设 $F(x)$ 连续单调,且 $\alpha\geqslant 0$,则算法所产生的序列 x^k 的每一个聚点都是 $NCP(F)$ 的解.

证明 不失一般性地设算法所生成的序列 x^k 收敛于 x^*.由于每次迭代 $\Psi_\alpha(x^k)$ 都下降,则 k 趋于无穷大时,(20)式的左边趋于零.我们考虑两种情况:

情况 1,$\{l_k\}$ 为有界的非负整数列.此时,$k\to+\infty$ 时,γ^{l_k} 不趋于零,故式(20)两边同时令 $k\to+\infty$,有 $\|d^k(x^*)\|=0$,从而 x^* 是 $NCP(F)$ 的解.

情况 2,$\{l_k\}$ 为无界的非负整数列.此时,根据定理 1 的证明,我们可以假设非负整数列 l_k 单调上升趋于正无穷,且

$$\Psi_\alpha(x^k+\gamma^{l_k-1}d^k)-\Psi_\alpha(x^k)>-\gamma^{2l_k-1}\|d^k\|^2.$$

由于 $\Psi_\alpha(x)$ 连续可微,上式两边同除以 γ^{l_k} 并令 $k\to+\infty$,可得

$$\nabla\Psi_\alpha(x^*)^T\left(-\frac{\partial\phi_\alpha}{\partial b}(x^*,F(x^*))\right)\geqslant 0. \quad (21)$$

结合引理 2 的结论,我们有

$$0=\nabla\Psi_\alpha(x^*)^T\frac{\partial\phi_\alpha}{\partial b}(x^*,F(x^*))=\frac{\partial\phi_\alpha}{\partial a}(x^*,F(x^*))^T\frac{\partial\phi_\alpha}{\partial b}(x^*,F(x^*))+\frac{\partial\phi_\alpha}{\partial b}(x^*,F(x^*))^T\nabla F(x^*)\frac{\partial\phi_\alpha}{\partial b}(x^*,F(x^*)). \quad (22)$$

这样必然可得 $\frac{\partial\phi_\alpha}{\partial a}(x^*,F(x^*))^T\frac{\partial\phi_\alpha}{\partial b}(x^*,F(x^*))=0$,由命题 1 的(iv),可知 x^* 是 $NCP(F)$ 的解.

为此可以得到算法的整体收敛性定理.

定理 3 如果 $F(x)$ 在 R^n 上强单调,或者 $F(x)$ 在 R^n 上单调且 $NCP(F)$ 严格可行,则算法产生的序列 $\{x^k\}$ 至少有一个聚点,且每个聚点都是 $NCP(F)$ 的解.

证明 根据文献[11]可知,当 $F(x)$ 在 R^n 上强单调,或者 $F(x)$ 在 R^n 上单调且 $NCP(F)$ 严格可行时,水平集 $L(\Psi_\alpha,x^0)=\{x\in R^n\mid\Psi_\alpha(x)\leqslant\Psi_\alpha(x^0)\}$ 有界,又由于 $\Psi_\alpha(x^k)$ 单调下降,即 $\{x^k\}\subset L(\Psi_\alpha,x^0)$,则序列 $\{x^k\}$ 为有界序列,故至少存在一个聚点,而定理 2 则说明 x^k 的每一个聚点都是 $NCP(F)$ 的解.

3 数值试验

下面用 3 个数值例子来验证算法的实效性.

例 1 $NCP(F)$ 的试验函数为:

$$F(x)=\begin{bmatrix}x_1^2-x_2^2+x_3^2\\2x_1^2-x_2-x_3-1\\-x_1^2-x_2^2-x_3^2+9\end{bmatrix}$$

这个问题在迭代区间$[0,+\infty)$上有一个解:$x^*=(1,1,0)^T$.选取初始点 $x_0=(0,0,0)^T$ 和参数 $\alpha=0.25,\gamma=0.1,\varepsilon=10^{-12}$,得到数值结果:$x=(1.0000,1.0000,0.0002)^T,F(x)=(0.0000,0.0001,7.0000)^T$.

例 2 $NCP(F)$ 的试验函数为:

$$F(x)=\begin{bmatrix}3x_1^2+2x_1x_2+2x_2^2+x_3+3x_4-6\\2x_1^2+x_1+x_2^2+3x_3+2x_4-2\\3x_1^2+x_1x_2+2x_2^2+2x_3+3x_4-1\\x_1^2+3x_2^2+2x_3+3x_4-3\end{bmatrix}$$

这个问题在迭代区间$[0,+\infty)$上有唯一解:

$$x^*=\left(\frac{\sqrt{6}}{2},0,0,\frac{1}{2}\right)^T$$

选取初始点 $x_0=(1,1,1,1)^T$ 和参数 $\alpha=0.25,\gamma=0.1,\varepsilon=10^{-12}$,得到数值结果:

$$x=(1.2247,0.0002,0.0000,0.5000)^T$$

$$F(x)=(0.0003,3.2247,4.9999,0.0002)^T.$$

例 3 线性互补函数:

$$F(x)=Mx+q,$$

其中矩阵 M 和向量 q 由下式给出,

$$M=\begin{bmatrix}3&0&-1&0\\-1&3&-1&0\\0&-1&4&-2\\-1&-1&-1&5\end{bmatrix},q=\begin{bmatrix}-2\\-3\\-4\\5\end{bmatrix}$$

这个问题在迭代区间$[0,+\infty)$上有唯一解

$$x^*=(1,0,1,0)^T.$$

选取初始点 $x_0=(10,10,10,10)^T$ 和参数 $\alpha=0.25,\gamma=0.1,\varepsilon=10^{-12}$,得到数值结果:

$x = (1.0000, 0.0000, 1.0000, 0.0000)^T$,

$F(x) = (0.0000, 1.0000, 0.0000, 3.0000)^T$.

参考文献：

[1] HARKER P T,PANG J S. Finite-dimensional variational inequality and nonlinear complementarity problem: asurvey of theory algorithms and applications[J]. Math Programming,1990,48:161-220.

[2] PANG J S. Complementarity problems[M]//HORST R, PARDALOSE P,eds. Handbook of Global Optimization. Boston:Kluwer Academic Publishers,1995:271-338.

[3] NOOR M A,Al-Said E A. An iterative techniquefor generalized strongly nonlinear complementarty problems [J]. Appl Math Lett,1999(12):75-79.

[4] NOOR M A. Fixed point approach for complementarty-problems[J]. J Math Anal Appl,1988,133:437-448.

[5] NOOR M A. Iteraitve methods for a class of complementarty problems[J]. Engineering Analysis,1986,3(4):221-224.

[6] NOOR M A,ZARAE S. An iteraitve scheme for complementarty problems[J]. Math Anal Appl,1988,133:366-382.

[7] CHEN X,QI L,SUN D. Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities[J]. Math of Computation,1998,67:519-540.

[8] KANZOW C,PIEPER H. Jacobian smoothing methods for nonlinear complementarity problems[J]. SIAM J Optim,1999,9:342-373.

[9] PANG J S. Newton's method for B-differentiable equations[J]. Math Oper Res,1990,15:311-341.

[10] HARKER P T,XIAO B. Newton's method for the nonliner complementarity problems: a B-differentiable equation approach[J]. Math Oper Res,1990,48:339-358.

[11] YMADA K,YAMASHITA N,FUKUSHIMA M. A new derivative-free descent method for the nonlinear complementarity problem [M]//DIPILLO G, GINNESSI F,eds. Nonlinear Optimization and Related Topics. Boston: Kluwer Academic Publishers,2000:436-487.

（责任编辑：邓大玉　凌汉恩）

（上接第186页 Continue from page 186）

$$\frac{dM_2(c)}{dc} = 2c(tr\Phi'_1 Cov\beta_1(T)\Phi_1 + \|\Phi'_1\beta_1\|^2) - 2\|\Phi'_1\beta_1\|^2. \tag{2.5}$$

所以当$\frac{\|\Phi'_1\beta_1\|^2}{tr\Phi'_1 Cov(\bar{\beta}_1(T))\Phi_1 + \|\Phi'_1\beta_1\|^2} < c < 1$时，$\frac{dM_2(c)}{dc} > 0$，又$M_2(1)=0$，故在上述区间$M_2(c) < 0$，即$MSE\beta_1^*(c,T) < MSE\beta_1^*(T)$，

$tr\Phi'_1 Cov(\bar{\beta}_1(T))\Phi_1 = tr[\sigma_{11}\Phi'_1(X'_1X_1)^{-1}\Phi_1 - \delta\Phi'_1(X'_1X_1)^{-1}X'_1N_2X_1(X'_1X_1)^{-1}\Phi_1] \geqslant \sigma_{11}tr\Phi'_1(X'_1X_1)^{-1}\Phi_1 - \delta tr\Phi'_1(X'_1X_1)^{-1}X'_1X_1(X'_1X_1)^{-1}\Phi_1 = (\sigma_{11} - \delta)tr\Phi'_1(X'_1X_1)^{-1}\Phi_1 = (\sigma_{11}-\delta)\sum_{i=1}^{r}\lambda_i^{(1)-1}$，

所以，
$$\frac{\|\Phi'_1\beta_1\|^2}{tr\Phi'_1 Cov(\bar{\beta}_1(T))\Phi_1 + \|\Phi'_1\beta_1\|^2} \leqslant \frac{\|\beta_1\|^2}{(\sigma_{11}-\delta)\sum_{i=1}^{r}\lambda_i^{(1)-1} + \|\beta_1\|^2}, \tag{2.6}$$

且由引理2可得$\sigma_{11} - \delta = \sigma_{11}(1-\rho_{12}^2)(1+\frac{1}{n-r-2})$，命题得证.

同样方法可以证明[4]：

定理6　在定理5的条件下，两步Stein型主成分改进估计优于协方差改进估计，即$MSE(\beta_1^*(c,T)) < MSE\beta_1(T)$.

3　结束语

对于线性回归系统(1)的第二个方程的回归系数β_2，当设计阵X_2呈病态时同样可以构造相应的Stein型主成分改进估计.

致谢

作者感谢山东大学博士生导师林路教授的宝贵意见.

参考文献：

[1] 王松桂，线性回归系统回归系数的一种新估计[J].中国科学：A辑，1988(10):1033-1040.

[2] 刘爱义.相依回归方程组主成分估计的改进[J].数理统计与应用概率，1993(2):70-75.

[3] 于义良，宋卫星.回归系数的Stein型主成分估计[J].数学研究，1995(3):85-88.

[4] 林路.相依非线性回归系统中的附加信息Bayes拟似然[J].数学学报，2002(6):1227-1234.

（责任编辑：邓大玉　凌汉恩）