Math. Program., Ser. A 88: 211–216 (2000)

Digital Object Identifier (DOI) 10.1007/s101070000159

Short Communication

Bintong Chen · Xiaojun Chen · Christian Kanzow

A penalized Fischer-Burmeister NCP-function

Received: March 10, 1997 / Accepted: February 15, 2000
Published online May 12, 2000 –

Abstract. We introduce a new NCP-function in order to reformulate the nonlinear complementarity problem as a nonsmooth system of equations. This new NCP-function turns out to have stronger theoretical properties than the widely used Fischer-Burmeister function and other NCP-functions suggested previously. Moreover, numerical experience indicates that a semismooth Newton method based on this new NCP-function performs considerably better than the corresponding method based on the Fischer-Burmeister function.

Key words. nonlinear complementarity problem – Newton's method – semismoothness

1. Introduction

Let F be a continuously differentiable function from $\mathbb{R}^n$ into itself. The *nonlinear complementarity problem* NCP(F) is to find a vector $x \in \mathbb{R}^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0.$$

Many algorithms developed for NCP(F) or related problems are based on reformulating them as a system of equations using so-called NCP-functions. Here, a function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called an *NCP-function* if

$$\phi(a, b) = 0 \Longleftrightarrow ab = 0, a \geq 0, b \geq 0.$$

Given an NCP-function ϕ, let us define

$$\Phi(x) = \text{vec}\{\phi(x_i, F_i(x))\},$$

where vec$\{z_i\}$ denotes a vector whose ith element is given by z_i. By definition, $x^* \in \mathbb{R}^n$ is a solution of NCP(F) if and only if it solves the system of equations $\Phi(x) = 0$.

One popular choice of an NCP-function is the Fischer-Burmeister function [5]:

$$\phi_{FB}(a, b) := a + b - \sqrt{a^2 + b^2}.$$

The Fischer-Burmeister function has many interesting properties. However, it has limitations in dealing with monotone complementarity problems since it is too flat in the

B. Chen: Department of Management and Decision Sciences, Washington State University, Pullman, WA 99164-4736, USA, e-mail: chenbi@wsu.edu

X. Chen: Department of Mathematics and Computer Science, Shimane University, Matsue 690-8504, Japan, e-mail: chen@math.shimane-u.ac.jp

C. Kanzow: University of Hamburg, Institute of Applied Mathematics, Bundesstrasse 55, D-20146 Hamburg, Germany, e-mail: kanzow@math.uni-hamburg.de

Mathematics Subject Classification (1991): 90C33

positive orthant, the region of main interest for a complementarity problem (see the comments before Proposition 4 below).

In view of the above shortcomings of the Fischer-Burmeister function, we introduce the following new NCP-function

$$\phi_\lambda(a, b) := \lambda\phi_{FB}(a, b) + (1 - \lambda)a_+b_+,$$

where $\lambda \in (0, 1)$ is an arbitrary but fixed parameter. Clearly, the new NCP-function is a convex combination of the Fischer-Burmeister function ϕ_{FB} and the term a_+b_+; the latter term penalizes violations of the complementarity condition and will play a significant role from both a theoretical and a practical point of view.

Throughout this note, the equation operator Φ based on this new NCP-function will be denoted by Φ_λ, i.e.,

$$\Phi_\lambda(x) := \mathrm{vec}\{\phi_\lambda(x_i, F_i(x))\}.$$

Moreover, let us write

$$\psi_\lambda(a, b) := \frac{1}{2}\phi_\lambda(a, b)^2.$$

Then a natural merit function of Φ_λ is given by

$$\Psi_\lambda(x) := \frac{1}{2}\Phi_\lambda(x)^T\Phi_\lambda(x) = \sum_{i=1}^{n}\psi_\lambda(x_i, F_i(x)).$$

We denote the Euclidean norm by $\|\cdot\|$. We say that a mapping $G : \mathbb{R}^n \to \mathbb{R}^m$ is an LC^1 function if G' is locally Lipschitzian.

To present our results by this short note, we state all the results without proof. The detailed proofs can be found in the report version [1].

2. Properties of new NCP-function

First we present some properties of ϕ_λ and Φ_λ.

Proposition 1. *The function $\phi_\lambda : \mathbb{R}^2 \to \mathbb{R}$ satisfies the following properties:*

1. *ϕ_λ is an NCP-function.*
2. *ϕ_λ is continuously differentiable on $\mathbb{R}^2 \setminus \{(a, b) \mid a \geq 0, b \geq 0, ab = 0\}$.*
3. *ϕ_λ is strongly semismooth on $\mathbb{R}^2$.*
4. *The generalized gradient $\partial\phi_\lambda(a, b)$ is equal to the set of all (v_a, v_b) such that*

$$(v_a, v_b) = \begin{cases} \lambda\left(1 - \frac{a}{\|(a,b)\|}, 1 - \frac{b}{\|(a,b)\|}\right) + (1 - \lambda)(b_+\partial a_+, a_+\partial b_+) \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } (a, b) \neq (0, 0), \\ \lambda(1 - \xi, 1 - \zeta) \qquad\qquad \text{if } (a, b) = (0, 0), \end{cases} \tag{1}$$

   where (ξ, ζ) is any vector satisfying $\|(\xi, \zeta)\| \leq 1$ and

$$\partial z_+ = \begin{cases} 1 & \text{if } z > 0, \\ [0, 1] & \text{if } z = 0, \\ 0 & \text{if } z < 0. \end{cases}$$

5. *Let $\{a^k\}, \{b^k\} \subset \mathbb{R}$ be any two sequences such that either $a_+^k b_+^k \to \infty$, or $a^k \to -\infty$, or $b^k \to -\infty$. Then $|\phi_\lambda(a^k, b^k)| \to \infty$.*

Result 2 implies that the Fischer-Burmeister function is smoother than ϕ_λ. ϕ_{FB} is nondifferentiable only at $(0, 0)$, while ϕ_λ is nondifferentiable on $\{(a, b) \mid a \geq 0, b \geq 0, ab = 0\}$. However, this additional nonsmoothness does not affect the convergence analysis.

The next theorem is a consequence of Result 3 in Proposition 1.

Theorem 1. *Φ_λ is semismooth. Moreover, Φ_λ is strongly semismooth if F is LC^1.*

Based on Result 4 of Proposition 1, we obtain the following overestimation of the C-subdifferential $\partial_C \Phi_\lambda(x)^T := \partial\Phi_{\lambda,1}(x) \times \ldots \times \partial\Phi_{\lambda,n}(x)$.

Proposition 2. *For any $x \in \mathbb{R}^n$, we have $\partial_C \Phi_\lambda(x) \subseteq D_a(x) + D_b(x)F'(x)$, where $D_a(x) = \operatorname{diag}\{a_i(x)\}$ and $D_b(x) = \operatorname{diag}\{b_i(x)\}$ are diagonal matrices with entries $(a_i(x), b_i(x)) \in \partial\phi_\lambda(x_i, F_i(x))$, where $\partial\phi_\lambda(x_i, F_i(x))$ denotes the set from Proposition 1 with (a, b) being replaced by $(x_i, F_i(x))$.*

We now provide a procedure to calculate an element of the C-subdifferential $\partial_C \Phi_\lambda(x)$ at an arbitrary point $x \in \mathbb{R}^n$.

Algorithm 1. *(Procedure to evaluate an element $V \in \partial_C \Phi_\lambda(x)$)*

(S.0) Let $x \in \mathbb{R}^n$ be given, and let V_i denote the ith row of a matrix $V \in \mathbb{R}^{n\times n}$.
(S.1) Set $S_1 = \{i \mid x_i = F_i(x) = 0\}$ and $S_2 = \{i \mid x_i > 0, F_i(x) > 0\}$.
(S.2) Set $z \in \mathbb{R}^n$ such that $z_i = 0$ for $i \notin S_1$ and $z_i = 1$ for $i \in S_1$.
(S.3) For $i \in S_1$, set

$$V_i = \lambda\left(1 - \frac{z_i}{\|(z_i, \nabla F_i(x)^T z)\|}\right) e_i^T + \lambda\left(1 - \frac{\nabla F_i(x)^T z}{\|(z_i, \nabla F_i(x)^T z)\|}\right) \nabla F_i(x)^T.$$

(S.4) For $i \in S_2$, set

$$V_i = \left[\lambda\left(1 - \frac{x_i}{\|(x_i, F_i(x))\|}\right) + (1-\lambda)F_i(x)\right] e_i^T + \left[\lambda\left(1 - \frac{F_i(x)}{\|(x_i, F_i(x))\|}\right) + (1-\lambda)x_i\right] \nabla F_i(x)^T.$$

(S.5) For $i \notin S_1 \cup S_2$, set

$$V_i = \lambda\left(1 - \frac{x_i}{\|(x_i, F_i(x))\|}\right) e_i^T + \lambda\left(1 - \frac{F_i(x)}{\|(x_i, F_i(x))\|}\right) \nabla F_i(x)^T.$$

The following proposition is a consequence of Results 2 and 4 of Proposition 1.

Proposition 3. *The matrix V calculated by Algorithm 1 is an element of $\partial_C \Phi_\lambda(x)$.*

To ensure fast local convergence for a semismooth algorithm, we require all elements in the C-subdifferential $\partial_C \Phi_\lambda(x^*)$ to be nonsingular at a solution point x^* of NCP(F). The next result shows that the R-regularity condition [8] is sufficient for this purpose.

Theorem 2. *If x^* is an R-regular solution of NCP(F), all $V \in \partial_C \Phi_\lambda(x^*)$ are nonsingular.*

Next, we present some properties of ψ_λ and Ψ_λ.

Theorem 3. *The merit function Ψ_λ is continuously differentiable with $\nabla\Psi_\lambda(x) = V^T\Phi_\lambda(x)$ for any $V \in \partial_C\Phi_\lambda(x)$. In addition, if x^* is a stationary point of Ψ_λ such that the Jacobian $F'(x^*)$ is a P_0-matrix, then x^* is a solution of NCP(F).*

We next consider the level sets

$$\mathcal{L}(c) := \{x \in \mathbb{R}^n \mid \Psi_\lambda(x) \leq c\}.$$

It is known [3] that the level sets of the Fischer-Burmeister function are compact if F is a uniform P-function. The following result shows that this is also true for the new merit function Ψ_λ. More importantly, this result states that the level sets of Ψ_λ are also bounded for many monotone complementarity problems, which is not true for the merit function Ψ_{FB}, as pointed out in [6].

Proposition 4. *The level sets $\mathcal{L}(c)$ of the merit function Ψ_λ are compact for any $c \geq 0$ if one of the following two conditions are satisfied:*

1. *F is a uniform P-function.*
2. *F is monotone and NCP(F) has a strictly feasible vector.*

The next result shows that $\sqrt{\Psi_\lambda}$ provides a global error bound for a complementarity problem with a uniform P-function.

Theorem 4. *If F is a uniform P-function, then there exists a constant $\tau > 0$ such that*

$$\|x - x^*\|^2 \leq \tau\Psi_\lambda(x)$$

for all $x \in \mathbb{R}^n$, where x^* is the unique solution of NCP(F) .

Notice that the above result does not require F to be Lipschitz continuous, a condition often needed for similar results based on merit functions derived from the Fischer-Burmeister function or related NCP-functions. See [6] for another NCP-function with this property.

Now we elaborate on Result 2 of Proposition 4. More precisely, we show that the assumption on the existence of a strictly feasible point is also necessary for $\mathcal{L}(c)$ to be bounded for monotone complementarity problems.

Proposition 5. *If F is a P_0-function and NCP(F) has a nonempty and bounded solution set, then there is a strictly feasible point for NCP(F).*

Note that Proposition 5 for a continuous P_0-function was independently derived by Ravindran and Gowda in their very recent paper [7]. As a consequence of previous discussions, we have the following interesting result about monotone NCPs.

Corollary 1. *If F is a monotone function, the following three statements are equivalent:*

1. *NCP(F) has a nonempty and bounded solution set.*
2. *NCP(F) has a strictly feasible point.*
3. *$\mathcal{L}(c)$ is bounded for all $c \geq 0$.*

3. Numerical results

Based on the discussion of the previous section, the NCP-function ϕ_λ as well as the merit function Ψ_λ possess all the nice features of the Fischer-Burmeister function and the corresponding merit function Ψ_{FB}. Moreover, Ψ_λ has some stronger properties than Ψ_{FB}. Therefore, by replacing ϕ_{FB} and Ψ_{FB} by ϕ_λ and Ψ_λ, respectively, in any semismooth based algorithm designed for the former functions, we can preserve and in some cases improve the convergence properties of the algorithm. Here, we choose the modified damped Newton method proposed by De Luca, Facchinei and Kanzow [2] to test the new NCP-function. This algorithm, based on the new function, is guaranteed to solve not only complementarity problems with uniform P-functions, but also monotone problems with a strictly feasible point.

We implemented this algorithm in MATLAB using the parameters $\lambda = 0.95$, $\beta = 0.5, \sigma = 10^{-4}, \rho = 10^{-10}, p = 2.1$. The main termination criterion is $\Psi_\lambda(x) \leq 10^{-12}$. We applied Algorithm 1 to compute an element in $\partial_C \Phi_\lambda(x^k)$. In addition, we incorporated some strategies to improve the numerical behaviour of the algorithm to some extent, namely a nonmonotone line search as well as a backtracking strategy in order to avoid possible domain violations. In order to compare the behaviour of our new function with the Fischer-Burmeister function, we also run the algorithm with $\lambda = 1$.

We tested our algorithm on all test problems from the MCPLIB and GAMSLIB libraries (see [4]), using all the different starting points which are available within the MATLAB framework. In Table 1, we present the results only for those test problems where the number of iterations used by the two methods differ by at least 5. In particular, the columns in Table 1 have the following meanings:

Table 1. Results for some MCPLIB and GAMSLIB test problems

			Ψ_λ				Ψ_{FB}			
problem	n	SP	k	F	G	$\Psi_\lambda(x^f)$	k	F	G	$\Psi_{FB}(x^f)$
bertsekas	15	2	19	72	0	1.1e-14	31	231	0	6.7e-16
bertsekas	15	3	27	34	0	2.7e-16	17	28	0	3.4e-16
billups	1	1	441	5260	0	4.2e-13	–	–	–	–
colvdual	20	1	14	16	0	5.0e-15	34	82	0	4.0e-19
colvdual	20	2	37	49	0	2.6e-19	–	–	–	–
hanskoop	14	5	11	15	1	3.1e-14	16	20	1	2.4e-20
josephy	4	1	7	8	0	2.7e-20	100	128	0	6.1e-17
josephy	4	3	15	16	0	1.1e-23	81	98	0	3.2e-14
pgvon106	106	1	500	3864	121	5.2e-11	–	–	–	–
powell	16	3	15	18	1	1.2e-19	27	112	9	7.0e-13
scarfbsum	40	2	26	28	1	2.1e-14	46	427	10	1.3e-16
sppe	27	1	10	32	2	2.1e-18	15	24	2	5.7e-29
tobin	42	2	10	13	0	5.9e-16	15	16	0	5.5e-24
dmcmge	170	1	22	51	0	7.4e-16	–	–	–	–
harkmcp	32	1	9	10	1	3.7e-14	14	16	1	8.5e-15
nsmge	212	1	18	21	0	6.2e-18	12	16	0	9.8e-16
transmcp	11	1	13	38	2	9.0e-17	20	37	1	1.5e-27
vonthmcp	125	1	58	335	10	1.5e-14	–	–	–	–
vonthmge	80	1	37	324	4	9.9e-22	51	243	1	3.2e-23

problem:	name of test example,
n :	dimension of test example,
SP:	number of starting point (see the M-file `cpstart.m`),
Ψ_λ/Ψ_{FB}:	results for the new function/FB-function,
k:	number of iterations,
F:	number of function evaluations,
G:	number of gradient steps,
$\Psi_\lambda(x^f)/\Psi_{FB}(x^f)$:	function value of Ψ_λ/Ψ_{FB} at the final iterate x^f,

Table 1 clearly indicates the superior behaviour of the new function. In particular, we were able to solve all test problems, while the same method based on the Fischer-Burmeister function has a couple of failures. Moreover, the new method usually needs fewer number of iterations even on the other test problems.

References

1. Chen, B., Chen, X., Kanzow, C. (1997): A penalized Fischer-Burmeister NCP-function: Theoretical investigation and numerical results. Preprint 126, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany
2. De Luca, T., Facchinei, F., Kanzow, C. (1996): A semismooth equation approach to the solution of nonlinear complementarity problems. Math. Program. **75**, 407–439
3. Facchinei, F., Soares, J. (1997): A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J. Optim. **7**, 225–247
4. Ferris, M.C., Rutherford, T.F. (1996): Accessing realistic mixed complementarity problems within MATLAB. In: Di Pillo, G., Giannessi, F. (eds.): Nonlinear Optimization and Applications. Plenum Press, New York, 1996, 141–153
5. Fischer, A. (1992): A special Newton-type optimization method. Optimization **24**, 269–284
6. Kanzow, C., Yamashita, N., Fukushima, M. (1997): New NCP-functions and their properties. J. Optim. Theory Appl. **94**, 115–135
7. Ravindran, G., Gowda, M.S. (1997): Regularization of P_0-functions in box variational inequality problems. Technical Report, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD
8. Robinson, S.M. (1980): Strongly regular generalized equations. Math. Oper. Res. **5**, 43–62