A smoothing multidimensional filter method for nonlinear complementarity problems

Wu Congwei
Xi'an Research Institute of High-tech
Tongxin Road No.2, 2102 Team,
Baqiao District,
Xi'an City, Shaanxi Province,
P.R.China
Email:wucongweihome@sina.com

Cao Jiping
Xi'an Research Institute of High-tech
Tongxin Road No.2, Baqiao District,
Xi'an City, Shaanxi Province,
P.R.China
Email: xacjp123@163.com

Zhu Yahong
Xi'an Research Institute of High-tech
Tongxin Road No.2, 2102 Team,
Baqiao District,
Xi'an City, Shaanxi Province,
P.R.China
Email: zyhbingchen@163.com

***Abstract*—A smoothing multidimensional filter method for solving NCP is proposed, and the solution of the complementarity problems on the framework of the filter-trust-region method is obtained. The new algorithm does not depend on any extra restoration procedure and the results of numerical experiments show its efficiency.**

***Keywords*—Multidimensional filter techniques; Nonlinear complementarity problems; filter-trust-region method**

I. INTRODUCTION

Let $F(x):\Re^n \to \Re^n$ be a continuous differentiable function. The nonlinear complementarity problem (NCP) is to find a vector $x \in \Re^n$ such that

$$x \ge 0, \quad F(x) \ge 0, \quad x^T F(x) = 0, \tag{1}$$

For convenience, denote $I = \{1, 2, ..., n\}$. Throughout this paper, $\|\cdot\|$ denotes the Euclidean norm.

The traditional approach for NCP involves reformulating the problem as a constrained or unconstrained optimization problem. We discuss for solving this optimization based on the class of trust-region methods and also on that of multidimensional filter methods introduced by Gould and Sainvitu [1].

Definition 1 *A function* $\phi:\Re^2 \to \Re$ *is said to be an NCP function if it possesses the following characterization* $\phi(a,b) = 0 \Leftrightarrow a \ge 0,\ b \ge 0$ *and* $ab = 0$.

In this paper, we will use Fischer-Burmeister functions $\phi:\Re^2 \to \Re$ defined by $\phi(a,b) := \sqrt{a^2 + b^2} - (a+b)$

It is obviously that ϕ is not differentiable at zero point, which is difficult to get the global convergence results, so we use the following smooth approximation for the Fischer-Burmeister function,

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

where $\mu \ge 0$ is a smooth parameter. Then NCP can be approximated by the following nonlinear least square problems,

$$\min_{x \ge 0} f_\mu(x) := \frac{1}{2}\Phi_\mu(x)^T \Phi_\mu(x) \tag{2}$$

where

$$\Phi_\mu(x) = \begin{pmatrix} \phi_\mu(x_1, F_1(x)) \\ \vdots \\ \phi_\mu(x_n, F_n(x)) \end{pmatrix} \tag{3}$$

Obviously, if the complementarity problem (1) is solvable, then the minimization problem (2) and (1) are equivalent when the parameter μ tends to zero. Throughout this paper, to simplify notation we will use the abbreviations $g_\mu(x) = \nabla_x f_\mu(x)$. Beside, $g_{\mu,i}(x)$ denote the i-th component of $g_\mu(x)$.

II. THE MULTIDIMENSIONAL FILTER METHOD

The projected gradient of the objective function f_(x) into the feasible set of the problem (2) is defined componentwise by

$$\bar{g}_i(x) = \begin{cases} g_i(x), & x_i \ge g_i(x), \\ x_i, & x_i < g_i(x), \end{cases} \tag{4}$$

where $i \in I$.

Note that the problem (2) is a nonlinear optimization problem with a positive parameter μ, so we can use x_μ^* denote the KKT point of the problem (2), and then we have the following Conclusion.

Lemma 1 *For all* $\mu > 0$, *a point* x_μ^* *is a KKT point for the problem (2) if and only if* $\bar{g}_\mu(x_\mu^*) = 0$.

Lemma 2 *Assume that* $F(x)$ *is a twice-continuously differentiable function on an closed and bounded set, then the KKT point of (2) converges to the KKT point of the following problem*

$$\min_{x \ge 0} f_0(x) := \frac{1}{2}\Phi_0(x)^T \Phi_0(x) \tag{5}$$

when the parameter μ *tends to zero.*

To solve the optimization problem (2), we compute a trial step d_k at a given iterate x_k by the following trust-region subproblem

$$\begin{aligned} \min \quad & Q_k(d) = f_{\mu_k}(x_k) + \nabla_x f_{\mu_k}(x_k)^T d + \frac{1}{2} d^T B_k d, \\ s.t. \quad & x_k + d \ge 0, \\ & \| d \|_\infty \le \Delta_k \end{aligned} \tag{6}$$

Where Δ_k is the trust region radius, $B_k = \nabla_x \Phi_{\mu_k}(x_k) \nabla_x \Phi_{\mu_k}(x_k)^T$, and

$$\nabla_x \Phi_\mu(x) = \begin{pmatrix} \frac{\partial \Phi_{\mu,1}(x)}{\partial x_1} & \cdots & \frac{\partial \Phi_{\mu,n}(x)}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Phi_{\mu,1}(x)}{\partial x_n} & \cdots & \frac{\partial \Phi_{\mu,n}(x)}{\partial x_n} \end{pmatrix} \tag{7}$$

The positive parameters μ_k tends to zero during the iterate of algorithm. A trial point x_k^+ is then computed by the trial step d_k, denote $x_k^+ = x_k + d_k$.

In our context, Our aim is to encourage convergence to first-order critical points by driving every component of the projected gradient $\overline{g}_\mu(x) = (\overline{g}_{\mu,1}(x), \overline{g}_{\mu,2}(x), \ldots \overline{g}_{\mu,n}(x))^T$ to zero.

Definition 2 *A iterate point* x_k *is said to dominate another point* x_l *if and only if* $|\overline{g}_{\mu_{k-1},i}(x_k)| \le |\overline{g}_{\mu_{l-1},i}(x_l)|, \ \forall i \in I$.

Definition 3 *A filter set* $\mathcal{F}$ *is a set of points such that no pair dominates any other.*

We say that a new trial point x_k^+ is acceptable for the filter $\mathcal{F}_k$ if and only if

$$\forall x_l \in \mathcal{F}_k, \exists j \in I, |\overline{g}_{\mu_k,j}(x_k^+)| \le |\overline{g}_{\mu_{l-1},j}(x_l)| - \gamma_g \|\overline{g}_{\mu_{l-1}}(x_l)\|, \tag{8}$$

Where $\gamma_g = (0, 1/\sqrt{n})$.

If an iterate x_k is acceptable for the filter $\mathcal{F}_k$, we add it to the filter and remove from it every $x_l \in \mathcal{F}_k$ such that $|\overline{g}_{\mu_k,i}(x_k^+)| \le |\overline{g}_{\mu_{l-1},i}(x_l)|$ for all $i \in I$, i.e.

$$\mathcal{F}_{k+1} = \mathcal{F}_k \bigcup \{x_k^+ \setminus \mathcal{D}_k\} \tag{9}$$

where,

$$\mathcal{D}_k = \{x_l \in \mathcal{F}_k \mid |\overline{g}_{\mu_k,i}(x_k^+)| \le |\overline{g}_{\mu_{l-1},i}(x_l)|, \forall i \in I\} \tag{10}$$

Note that the element of the set $\mathcal{F}_k$ is not x_k^+ but $(|\overline{g}_{\mu_k,1}(x_k^+)|, \ldots, |\overline{g}_{\mu_k,n}(x_k^+)|)^T$. when we say a new trial point x_k^+ is acceptable for the filter $\mathcal{F}_k$. Nevertheless, it is convenient to say that we add point x_k^+ into $\mathcal{F}_k$.

III. The Multidimensional Filter Algorithm

At each step, after a subproblem is solved, the filter and traditional trust region criterion are all employed to determine whether to accept the trial point or not. In trust region criterion, if there is a good agreement between the model and the objective function value at the current trail point $x_k^+ = x_k + d_k$, it is said to be successful iteration; otherwise, it is said to be unsuccessful iteration.

In following algorithm, the multidimensional filter criterion is a relaxation for the trust region criterion to a certain extent because we even accept the set of iterations which is not accepted by trust-region criterion but filter criterion.

Algorithm 1 The Multidimensional Filter Algorithm for NCP

Step 0: *Initialization.* An initial point and an initial trust-region radius Δ_0 are Let an initial point x_0, an initial trust-region radius $\Delta_0 > 0$, and an initial filter set $\mathcal{F}_0 = (10^5, \ldots, 10^5)^T$ be given, as well as constants $\gamma_g = (0, 1/\sqrt{n})$, $0 < \gamma_1 < \gamma_2 < 1 < \gamma_3$, $0 < \Delta_0 < \Delta_{\max}$. Compute $f_{\mu_0}(x_0)$, $g_{\mu_0}(x_0)$, $\overline{g}_{\mu_0}(x_0)$, B_0, set $k_0 := 0$.

Step 1: *Test for optimality.* If $\|\overline{g}_{\mu_k}(x_k)\| + \mu_k < \varepsilon$, stop.

Step 2: *Determine a trial step.* Compute a solution d_k of the subproblem (6).

Step 3: If $d_k = 0$, set $x_{k+1} = x_k$, $\mu_{k+1} = \theta \mu_k$, $B_{k+1} = B_k$, set $k := k+1$, and go to Step1; else set $x_k^+ = x_k + d_k$, and compute $f_{\mu_k}(x_k^+)$, $g_{\mu_k}(x_k^+)$.

Step 4: *Test for optimality.* If $\|\overline{g}_{\mu_k}(x_k^+)\| + \mu_k < \varepsilon$, stop; else, compute

$$\rho_k = \frac{f_{\mu_k}(x_k) - f_{\mu_k}(x_k^+)}{Q_k(0) - Q_k(d_k)} \tag{11}$$

Step 5: *Tests to a accept the trial step.*

- If $\rho_k \ge \eta_1$, set $x_{k+1} = x_k^+$, $\mathcal{F}_{k+1} = \mathcal{F}_k \bigcup \{x_k^+ \setminus \mathcal{D}_k\}$;
- Else if $\rho_k < \eta_1$ and x_k^+ is acceptable for the filter $\mathcal{F}_k$, set $x_{k+1} = x_k^+$, $\mathcal{F}_{k+1} = \mathcal{F}_k \bigcup \{x_k^+ \setminus \mathcal{D}_k\}$,
- Otherwise, set $x_{k+1} = x_k$, $\mathcal{F}_{k+1} = \mathcal{F}_k$.

Step 6: *Update the trust-region radius and the smooth parameter.*

$$\Delta_{k+1} = \begin{cases} [\gamma_1 \Delta_k, \gamma_2 \Delta_k], & \text{if } \rho_k < \eta_1, \\ (\gamma_2 \Delta_k, \Delta_k], & \text{if } \rho_k \in [\eta_1, \eta_2), \\ \min\{\Delta_{\max}, \gamma_3 \Delta_k\}, & \text{if } \rho_k \ge \eta_2. \end{cases} \tag{12}$$

$$\mu_{k+1} = \begin{cases} \theta \mu_k, & \text{if } \mu_k > 0.1 \|\overline{g}_{\mu_k}(x_{k+1})\|, \\ \mu_k, & \text{otherwise.} \end{cases} \tag{13}$$

Step 7: Compute $f_{\mu_{k+1}}(x_{k+1})$, $g_{\mu_{k+1}}(x_{k+1})$, $\overline{g}_{\mu_{k+1}}(x_{k+1})$, B_{k+1}, set $k := k+1$, and go to step1.

It seems that we have to compute the value of the projected gradient and the objective function twice in every iteration. In practice, μ_k seldom update because it is a sufficient small parameter. Thus, it is rarely to compute the value of the projected gradient and the objective function twice in every iteration. There is an advantage to choosing a large $\Delta_{k+1} = \gamma_3 \Delta_k$ when $\rho_k \ge \eta_2$, but it may be unwise to choose it to be to large, so we give a upper bound $\Delta_{\max}$, and set $\Delta_{k+1} = \min\{\Delta_{\max}, \gamma_3 \Delta_k\}$ when $\rho_k \geq \eta_2$.

Observes that the subproblem (6) is compatible during in algorithm, so we do not need any extra feasibility restoration phase[2] in our algorithm, which differentiated our paper from those filter algorithm for nonlinear complementarity problems[2]. By the way, we are surprised to find that $B_{k+1} = \nabla_x \Phi_{\mu_k}(x_k) \nabla_x \Phi_{\mu_k}(x_k)^T$ can be computed easily instead of updating B_{k+1} with higher numerical expenditure, because of (3) and (7).

IV. CONCLUSIONS AND NUMERICAL EXPERIMENTS

The new algorithm 1 presented in this paper combines with the multidimensional filter technique and the trust region method, which has a good numerical calculation results.

Now, we give some numerical results for the following 14 complementarity test problems in TABLE I. The values for the constants used in our tests are $\mu_0 = 10^{-5}$, $\varepsilon = 10^{-5}$, $\gamma_g = 10^{-3}$, $\gamma_1 = 0.25$, $\gamma_3 = 2$, $\eta_1 = 0.25$, $\eta_2 = 0.95$, $\theta = 0.1$, $\Delta_{\max} = 10^3$, $\mathscr{F}_0 = (10^5, ..., 10^5)^{\mathrm{T}}$. The trust-region radius update is implemented as

$$\Delta_{k+1} = \begin{cases} \gamma_1 \Delta_k, & if\ \rho_k < \eta_1, \\ \Delta_k, & if\ \rho_k \in [\eta_1, \eta_2), \\ \min\{\Delta_{\max}, \gamma_3 \Delta_k\}, & if\ \rho_k \geq \eta_2. \end{cases}$$

Example 1.1 (Kojima-Shindo Nonlinear complementarity test problem)We choose the degenerate example in [2, 3].

Example 1.2 (Kojima-Shindo Nonlinear complementarity test problem)We choose the non-degenerate example in [3].

Example 2 (Kanzow Nonlinear complementarity test problem[2, 4]).

Example 3.1 We choose Example 2 with the constant $b = ((-1)^1, ..., (-1)^i, ..., (-1)^n)^T$ in [5].

Example 3.2 We choose Example 2 with the constant $b = ((-1)^1\sqrt{1}, ..., (-1)^i\sqrt{i}, ..., (-1)^n\sqrt{n})^T$ in [5].

Example 4 We choose Example 2 in [6].

The computational results are listed in Table 1, in which *iter* denotes the number of iterations, and $\mathrm{Re}sf$ stand for the computing accuracy, i.e. $\mathrm{Re}sf = f_{\mu_k}(x_k)$. The numerical results show that Algorithm 1 is robust and efficient. The number of iterations and computing accuracy for most problems are satisfactory.

TABLE I. NUMERICAL RESULT OF TEST PROBLEMS

Example	***Start point***	***iter***	***Resf***
1.1	$(0,0,0,0)^T$	7	$4.43e-12$
1.1	$(2,1,0.5,2)^T$	9	$1.02e-12$
1.2	$(0,0,0,0)^T$	9	$7.98e-15$
1.2	$(2,1,0.5,2)^T$	8	$2.84e-10$
2	$(3,2,1,2,3)^T$	2	$1.43e-13$
2	$(-2,...,-2)^T$	3	$9.17e-13$
3.1	$(2,...,2)^T$	8	$1.66e-11$
3.2	$(2,...,2)^T$	5	$1.05e-14$
4	$(10^{-2},1,0.5,10^{-2},10^{-2})^T$	4	$1.68e-11$
4	$(12,-12,12,-12,12)^T$	9	$5.84e-11$

References

[1] N. I. M. Gould, C. Sainvitu, and P. L. Toint, "A filter-trust-region method for unconstrained optimization," SIAM J. Optim., vol.16, 2005, pp.341-357.

[2] J. Long and S. Y. Zeng, "A new Filter-Levenberg-Marquardt method with disturbance for solving nonlinear complementarity problems," Appl. Math. Comput., vol. 216, 2010, pp. 677-688.

[3] J. D. Ding and H. Y. Yin, "A new homotopy method for nonlinear complementarity problems," Numer. Math. J. Chinese Univ. (English Ser.), vol. 16, 2007, pp. 155-163.

[4] T. Ni and P. Wang, "A smoothing-type algorithm for solving nonlinear complementarity problems with a non-monotone line search, "Appl. Math. Comput., vol.216, 2010, pp. 2207-2214.

[5] C. C. Huang and S. Wang, "A power penalty approach to a nonlinear complementarity problem, " Oper. Res. Letters, vol.38, 2010,pp. 72-76.

[6] Y. Xia, H. Leung, and J. Wang, "A projection neural network and its application to constrained optimization problems," IEEE Transactions on Circuits and Systems-I, vol. 49, 2002,pp. 447-458.