Mathematica Aeterna, Vol. 7, 2017, no. 2, 147 - 154

A filter trust region method for solving generalized semi-infinite programming problems

Ruyue Hou
College of Mathematics and Computer Science
Hebei University
Baoding,071000
China

Chun Xu
College of Mathematics and Computer Science
Hebei University
Baoding,071000
China

Lele Ren
College of Mathematics and Computer Science
Hebei University
Baoding,071000
China

Ke Su
College of Mathematics and Computer Science
Hebei University
Baoding,071000
China

Abstract

In this paper, a filter trust region method is proposed for solving generalized semi-infinite programming problem (GSIP). By reformulating the Karush-Kuhn-Tucker conditions, we obtain a system of semismooth equations that is equivalent to the GSIP problem. Also, the NCP function is used to construct the semismooth equations. For solving this equivalent problem, a promising method, called filter method, is introduced. Compared with the existed methods for GSIP, the presented method is more flexible. there is only one system of linear equations needed to be solved at per iteration. And the scale of calculation is reduced to a certain degree. Under some reasonable conditions, the global convergent properties of the presented method are proven.

Mathematics Subject Classification: 90c30

Keywords: GSIP. Filter technique.Trust region method.

1 Introduction

We are concerned with generalized semi-infinite programming problems (GSIP) of the following form:

$$GSIP : \min\ f(x)$$

$$s.t. \begin{cases} x \in X = \{x \in R^n \mid g(x, t(x)) \leq 0, \quad t \in T(x)\} \\ T(x) = \{t \in R^m \mid h_i(x, t) \leq 0, \quad i \in I\} \end{cases} \tag{1}$$

and finite index set I={1,L,q}.In this paper, we assume that the functions $f : R^n \to R$, $g : R^n \times R^m \to R$ and $h_i : R^n \times R^m \to R(i = 1, 2, ...q)$ are real-valued and at least twice continuously differentiable.

For the special case that the index set T=T(x) does not depend on the variable x,i.e. the problem GSIP is a common semi-infinite problem (SIP) and will be abbreviated by SIP. If moreover T is a finite set, then GSIP reduces to a finite programming problem.

The theory and practice of SIP is well-develop since the studies started from the 60th century [1].GSIP is only studied since about 1985 [2-4]. It is meanwhile known that the structure of GSIP is more complicated than SIP. For an introduction to theory, applications, numerical methods for GSIP, we refer to the article [5, 6]. There are many real-life applications of GSIP including robust optimization, design centering, minimax problems, Chebyshev approximation, and disjunctive programming [6, 7].

In all approaches to solve GSIP numerically, one tries to relax GSIP locally or globally to a simpler problem, e.g., to a SIP or a finite program problem. Several different methods are described to transform GSIP to simpler problems in the article[8], including the dual method, the Karush-Kuhn-Kucker (KKT) method, the reduction method and the penalty method.

In this paper, we use a semismooth reformulation of KKT system for GSIP. We firstly reformulate the KKT system into a system of semismooth equations by using NCP functions. Then we solve it by using an ODE method based trust region method, and introduce the nonmonotone technique to modify the criterion for deciding whether accepting a trial point or not. Under certain conditions, we proved that the proposed method is globally convergent. In the classical trust region methods, a quadratic subproblem with a trust region bound has to be solved at per iteration. But in this article, it only solves a system of linear equations at each iteration. Moreover, the value of the filter in the nonmonotone technique is permitted to increase among finite iterations.

This paper is organized as follows. In section 2, we introduce some basic definitions and results which will be frequently used in this paper. In section 3, the GSIP problem will be converted into a system of semismooth equations. This reformulation is equivalent to the given GSIP problem to a certain degree. Then we propose a trust region algorithm based on the ODE-type method.

2 Preliminary Notes

In this section, we summarize some definitions and results, which are useful subsequently.

A locally Lipschitz function $F : R^n \to R^m$ is called semismooth at $x \in R^n$ if F is directionally differentiable at x and for all $V \in \partial F(x + d)$ and $d \to 0$,

$$F'(x; d) - Vd = o(\|d\|^2)$$

where ∂F is the generalized Jacobian of F in the sense of Clarke [9],i.e.

$$\partial F(x) = conv\{\lim_{x_k \to x} \bigtriangledown F(x_k^T) \mid F \ is \ differentiable \ at \ x_k\}$$

Furthermore, F is called strongly semismooth at x if F is semismooth at x and for all $V \in \partial F(x + d)$ and $d \to 0$,

$$F'(x; d) - Vd = O(\|d\|^2)$$

Assume that $F : R^n \to R^m$ be locally Lipschitzian and semismooth at x , then for any $V \in \partial F(x + h)$ and $h \to 0$, we have

$$F(x + h) = F(x) + Vh + o(\|h\|) \quad (2)$$

Furthermore, if F is strongly semismooth,

$$F(x + h) = F(x) + Vh + O(\|h\|^2) \quad (3)$$

Suppose N_x is a neighborhood of $x \in X$ and $\varphi : R^n \to R^m \in LC^1$, i.e. φ is continuously differentiable and its gradient is locally Lipschitzian. Then there exists a positive constant γ such that

$$|\varphi(x + d) - \varphi(x) - \bigtriangledown\varphi(x)^T d| \leq \frac{\gamma\|d\|^2}{2} \quad (4)$$

for all $x + d \in N_x$.

The Fisher-Burmeister function is defined as follows

$$\phi(a,b) = \sqrt{a^2+b^2} - a - b$$

which satisfies the following conditions.
(1) $\phi(a,b) = 0 \Leftrightarrow a \geq 0, b \geq 0, and\ ab = 0$;
(2) ϕ^2 is continuously differentiable;
(3) ϕ is second-order continuously differentiable at any points excepting for zero, and it is continuous at origin. Moreover ϕ is semismooth function.

We use NCP function for complementarity conditions in KKT system.

Next we briefly describe the reduction approach which can be applied to SIP and GSIP as well. Since this approach has been detailed in [2], here we only give the results.

For $\overline{x} \in X$,we define the set of active points $T_0(\overline{x}) = \{\overline{t} \in T(\overline{x}|g(\overline{x},\overline{t}) = 0)\}$

Obviously, for feasible $\overline{x} \in X$,any point $\overline{t} \in T_0(\overline{x})$ is a (global) maximum of the following parametric optimization problem (the so-called lower level problem):

$$Q(x) : \max_{t \in R^n} \{g(x,t)\}$$

$$s.t.\ \ h_i(x,t) \leq 0,\ i \in I \tag{5}$$

The main computational problem in semi-infinite programming is that lower level problem has to be solved to global optimality, even if only a stationary point of the upper level problem is sought. Since we replace lower level problem by its KKT conditions, lower level problem must be convex.

If $\overline{x} \in X$,is a local minimize of GSIP at which the reduction Ansatz without strict complementarity [7] and the extended Mangasarian-Fromovitz constraint qualification hold, then there exist p positive numbers $\overline{u}_j,\ j \in P = \{1, L, p\}$ such that

$$\nabla f(\overline{x}) + \sum_{\jmath=1}^{p} \overline{\mu}_j \nabla_x L(\overline{x}, \overline{t}^j, \overline{\lambda}^j) = 0 \tag{6}$$

$$\overline{\mu}_j > 0, g(\overline{x}, \overline{t}^j) = 0, j \in P \tag{7}$$

where $L(\overline{x}, \overline{t}^j, \overline{\lambda}^j) = g(\overline{x}, \overline{t}^j) - \sum_{\imath=1}^{q} \overline{\lambda_i}^p h_i(\overline{x}, \overline{t}^j), \overline{\lambda}^i \geq 0, i \in I.$

Then the upper level first order condition is complemented by a lower level first order condition. In fact, the active indices $\overline{t}^j \in T_0(\overline{x}), j \in P$ are global solutions of $Q(\overline{x})$.Under some constraint qualification like the Slater condition in the lower level problem, there exist vectors of the Lagrange multipliers $\overline{\lambda}^j \in R^q$ such that

$$\nabla_t g(\overline{x}, \overline{t}^j) - \sum_{\imath=1}^{q} \overline{\lambda}^j \nabla_t h_i(\overline{x}, \overline{t}^j) = 0, j \in P \tag{8}$$

$$\overline{\lambda}^j \geq 0, h_i(\overline{x}, \overline{t}^j) \leq 0, \overline{\lambda i}^j h_i(\overline{x}, \overline{t}^j) = 0, j \in P, i \in I. \tag{9}$$

Now, by using F-B function ϕ, the solution of upper and lower level first order condition is seen to be equivalent to the following function:

$$F(z) = \begin{bmatrix} \nabla f(x) + \sum_{j=1}^{p} u_j \nabla_x L(x, t^j, \lambda^j \\ \phi(\mu_1, -g(x, t^1) \\ M \\ \phi(\mu_p, -g(x, t^p) \\ \nabla_t g(x, t^1) - \sum_{i=1}^{q} \lambda i^1 \nabla_t h_i(x, t^1) \\ \phi(\lambda_1^1, -h_1(x, t^1) \\ M \\ \phi(\lambda_q^1, -h_q(x, t^1) \\ M \\ \nabla_t g(x, t^p) - \sum_{i=1}^{q} \lambda i^p \nabla_t h_i(x, t^p) \\ \phi(\lambda_1^p, -h_1(x, t^p) \\ M \\ \phi(\lambda_q^p, -h_q(x, t^p) \end{bmatrix} = 0 \tag{10}$$

where $z = (x^T, \mu^T, t^T, \lambda^T)^T$,T is strongly semismooth under our assumptions. Define some functions as follows:

$$1(z) = \nabla f(x) + \sum_{j=1}^{p} \mu_j \nabla_x L(x, t^j, \lambda^j, \tag{11}$$

$$\Psi(z) = (\psi_1(z), L, \psi_p(z))^T, \tag{12}$$

$$\psi_j(z) = \phi(\mu_j, -g(x, t^j)), j \in P \tag{13}$$

$$l(z) = (l_1(z)^T, L, l_p(z)^T)^T, \tag{14}$$

$$l_j(z) = \nabla_t g(x, t^j) - \sum_{i=1}^{q} \lambda i^j \nabla_t h_i(x, t^j), i \in P \tag{15}$$

$$\Phi(z) = (\Phi_1(z)^T, L, \Phi_p(z)^T)^T, \tag{16}$$

$$\Phi_j(z) = (\psi_{1j}(z), L, \psi_{qj}(z)), j \in P \tag{17}$$

$$\Psi_{ij}(z) = \phi(\lambda i^j, -h_i(x, t^j)), i \in I, j \in P \tag{18}$$

Then we have

$$F(z) = \begin{bmatrix} 1(z) \\ \Psi(z) \\ l(z) \\ \Phi(z) \end{bmatrix} = 0 \tag{19}$$

It is clear that the above equation (19) can be converted equivalently into the least square problem:

$$\min f(x) = \frac{1}{2}\|F(z)\|^2 \tag{20}$$

So far there have been many methods to deal with problem (20). An efficient approach for solving the nonsmooth equations (19) is to use the ODE-type trust region method[10].At the k-th iteration, this method is to obtain the search direction $d(k)$ by solving a system of linear equation

$$(J_k^T J_k + B_k + \frac{1}{h_k} I)d = -J_k^T F(z_k) \tag{21}$$

where $h_k > 0$ is the stepsize, $B_k \in R^{(n+l)\times(n+l)}$ is a symmetric matrix which carries out the second order information of f , $J_k \in \partial F(z_k)$.

Here is no assumption that $\{B_k\}$ is positive definite. We can use SR1 or BFGS formulas.

In order to use a filter technique, we define the function as follows:

$$\theta(z) = (\theta_1(z), \theta_2(z), theta_3(z), \theta_4(z))$$

where $\theta_1(z) = \|1(z)\|$,$\theta_2(z) = \|B(z)\|$,$\theta_3(z) = \|l(z)\|$,$\theta_4(z) = \|\phi(z)\|$.

A filter is a list $\overline{F}$ of 4-tuples of the form $(\theta_{1,k}, L, \theta_{4,k})$ such that $\theta_{j,k} < \theta_{j,l}$ for any $j \neq k$,where $\theta_{j,k} = \theta_j(z_k)$,$j \in \{1, 2, 3, 4\}$.A new trial iterate z_k^+ is acceptable for the filter $\overline{F}$ if and only if for all $\theta_l = \theta_z(l) \in \overline{F}$ such that

$$\theta_j(z_k^+) \leq \theta_{j,l} - \tau_\theta \delta(\|\theta_l\|, \|\theta_k^+\|) \tag{22}$$

where $\theta_k^+ = \theta(z_k^+)$,$\tau_\theta \in (0, 1)$is a small positive constant, and

$$\delta(\ \|\theta_l\|, \|\theta_k^+\|\ ) = \min\{\ \|\theta_l\|, \|\theta_k^+|\ \}$$

If a new trial point z_k^+ is acceptable in the sense of(22),we wish to add it to the filter, we simply perform the operation: $\overline{F} \leftarrow \overline{F} \cup (\theta_j^+)$, we also remove all from the filter that are dominated by (θ_k^+) .

3 FTR Algorithm

Now we give the formal description of our algorithm, then we discuss the convergence of the algorithm in this section.

Step0. Initialization: Given $z_1, h_1, \epsilon \geq 0, 0 < \rho_0 < 1, 0 < \tau_\theta < 1$. Set $\mathcal{F} = \emptyset, k := 1$

Step1. compute J_k İf $\|J_k^T F(z_k)\| \leq \epsilon$, stop.

Step2. Construct a symmetric matrix B_kİf $h_k(J_k^T J_k + B_k) + I$ is positive definite,go to Step 4;Otherwise,let m_k be the smallest positive integer such that

$$2^{-m_k} h_k(J_k^T J_k + B_k) + I$$

is positive definite. Set $h_k = 2^{-m_k} h_k$.

Step3. Solve equation (10) to obtain d_k .Set $z_k^+ = z_k + d_k$ .Let M be a nonnegative integer. Compute

$$\rho_k = \frac{Are\ d_k}{Pre\ d_k} = \frac{f(z_{k-j}) - f(z_k + d_k)}{f(z_k) - q_k(d_k)}$$

where $q_k(d_k) = \frac{1}{2}\ \|F(z_k) + J_k d\|^2 + \frac{1}{2} d^T B_k d$

Step4. If θ_k^+ is not acceptable for the current filter,go to Step5;Otherwise, set $z_{k+1} = z_k^+$ (called a successful iteration)and go to Step 6.

Step5. If $\rho_k \geq \rho_0$,then set $z_{k+1} = z_k^+, h_{k+1} = 2h_k$,go to Step7;Otherwise, set$h_k := \frac{1}{2} h_k$, and go to Step2(called inner cycle).

Step6. If $\rho_k \geq \rho_0$,then set $z_{k+1} = z_k^+, h_{k+1} = 2h_k$,go to Step7;Otherwise, add(θ_k^+) to the filter $\mathcal{F}$ and set $h_k := \frac{1}{2} h_k$, go to Step7.

Step7. Update the matrix B_k to B_{k+1} and set $k := k + 1$ (called outer cycle), go to Step2.

If $h_k(J_k^T J_k + B_k) + I$ is not positive definite, there must exist m_k which is the smallest positive integer such that $2^{-m_k} h_k(J_k^T J_k + B_k) + I$ is positive definite. Therefore, the system of the equation (10) has a unique solution.

Under convexity of lower level and conditions for the Clarke subdifferential regularity of generalized Jacobian at solution point, it can be shown that every accumulation point of the sequence z_k is a solution of $F(z) = 0$ and thus a stationary point of GSIP[11]. Moreover, it is also possible to show that the algorithm is superlinearly convergent under a bound condition on $\|d\|$[12].

ACKNOWLEDGEMENTS. This work is supported by the National Natural Science Foundation of China (No. 1101115), the Natural Science Foundation of Hebei Province (No. A2014201033), the Key Research Foundation of Education Bureau of Hebei Province (No. ZD2015069) and Post-graduates Innovation Fund Project of Hebei University (No. X2016060).

References

[1] R.Hettich and G.Still, Second order optimality conditions for generalized semi-infinite programming problems, Optimization, 1995, 34,195-211.

[2] R.Hettich and K.Kortanek, Semi-infinite programming: theory, methods and applications, SIAM Review,1993, 35(3)380-429.

[3] H.Th jongen, J.J.Rckmann and O.Stein. Generalized semi-infinite optimization: a first order optimality condition and examples. Mathematical Programming, 1998,83,145-158.

[4] J.J.Rckmann and A Shapiro. Fisrt order optimality conditions in generalized semi-infinite programming, JOTA,1999, 101(2), 677-691.

[5] F.Guerra Vazquez, J.J.Rckmann, O.Stein and G.Still. Generalized semi-infinite programming: a tutorial, J.Comput.Appl.Math,2008,217(2), 394-419.

[6] O.Stein,Bi-level Strategies in semi-infinite programming, Kluwer Academic, Boston ,2003.

[7] G.Still, Generalized semi-infinite programming: theory and methods, European Journal of Operational Research,1999, 119(2), 301-313.

[8] G.still,Solving generalized semi-infinite programs by reduction to simple problems, Optimization,2004,53(1),19-38.

[9] Clarke F.H,Optimization and nonsmooth analysis, 1983.

[10] Ou Y.G,A superlinearly convergent ODE-type trust region algorithm for nonsmooth nonlinear equations, J.Appl.Math.Comput,2006,22(3), 371-380.

[11] O.Stein, A.Tezel,The semismooth approach for semi-infinite programming under the Reduction Ansatz, Journal of Global Optimization,2008,41(2),245-266.

[12] Ou Y.G,A filter trust region method for solving semi-infinite programming problems, J.Appl.Math.Comput,2009,29(1-2),311-324.

Received: April 19, 2017