Special issue of the International Conference on Computational and Experimental Science and Engineering (ICCESEN 2014)

Solving Generalized Semi-Infinite Programming Problems with a Trust Region Method

A. Tezel Özturan*
Hacettepe University, Mathematics, Ankara, Turkey

In this paper, a trust region method for generalized semi-infinite programming problems is presented. The method is based on [O. Yi-gui, "A filter trust region method for solving semi-infinite programming problems", J. Appl. Math. Comput. **29**, 311 (2009)]. We transformed the method from standard to generalized semi-infinite programming problems. The semismooth reformulation of the Karush–Kuhn–Tucker conditions using nonlinear complementarity functions is used. Under some standard regularity condition from semi-infinite programming, the method is convergent globally and superlinearly. Numerical examples from generalized semi-infinite programming illustrate the performance of the proposed method.

DOI: 10.12693/APhysPolA.128.B-93
PACS: 02.60.Pn

1. Introduction

The article studies a numerical solution method for generalized semi-infinite optimization problems GSIP. A problem GSIP is of the type

$$\text{GSIP}: \text{minimize } f(x) \text{ subject to } x \in M \tag{1}$$

with $M = \{x \in R^n | g(x,y) \leq 0 \text{ for all } y \in Y(x)\}$ and $Y(x) = \{y \in R^m | v_j(x,y) \leq 0, j \in Q\}$.

All defining functions $fgv_j, j \in Q = \{1,\ldots,q\}$, are assumed to be real-valued and at least twice continuously differentiable on their respective domains. In a GSIP problem, the possibly infinite index set $Y(x)$ is allowed to vary with x, but in a standard semi-infinite optimization problem (SIP) the infinite index set is fixed, that is, we have $Y(x) \equiv Y$.

For surveys and recent results about theory and methods for standard semi-infinite programming let us refer to [1–6]. For introduction and results in generalized semi-infinite programming the reader is referred to [6–10].

Many solution methods in finite optimization is based on the Karush–Kuhn–Tucker (KKT) system, which is, a necessary first order optimality condition. In this paper, we use a semismooth reformulation of KKT system, for GSIP. For SIP, semismooth reformulation of KKT system is obtained in [6, 11].

1.1. Preliminaries

For a locally Lipschitzian function $F: R^n \to R^m$ let $\partial F(x)$ denote Clarke's generalized Jacobian at x [12]. F is called semismooth at $x \in R^n$ if F is directionally differentiable at x and if for all $V \in \partial F(x)$ and $d \to 0$ we have

$$F'(x;d) = Vd + O(\|d\|). \tag{2}$$

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

$$Vd - F'(x;d) = O(\|d\|). \tag{3}$$

A function $\psi: R^2 \to R$ is called a nonlinear complementarity problem (NCP)-function if

$$\psi(a,b) = 0 \text{ if and only if } a \geq 0,\ b \geq 0 \text{ and } ab = 0. \tag{4}$$

An important example of NCP function is the Fischer–Burmeister function

$$\psi_{\mathrm{FB}}(a,b) = \sqrt{a^2+b^2} - a - b. \tag{5}$$

We use NCP function for complementarity conditions in KKT system.

1.2. Semismooth reformulation of KKT conditions for GSIP

The lower level problem of GSIP is defined by:

$$Q(x): \max_{y \in R^m} g(x,y) \text{ subject to } v_j(x,y) \leq 0,\ j \in Q. \tag{6}$$

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.

For $\bar{x} \in M$ let $Y_0(\bar{x}) = \{y \in Y(\bar{x}) | g(\bar{x},y) = 0\}$ denote the set of active indices of $\bar{x}$. If $\bar{x} \in M$ is a local minimizer of GSIP at which the reduction Ansatz without strict complementarity (see [13]) and the extended Mangasarian–Fromovitz constraint qualification hold, then there exist a $p \in \{0,\ldots,n\}$ and multipliers $\bar{\mu}_i \geq 0$, $i \in P = \{1,\ldots,p\}$, such that

$$\nabla f(\bar{x}) + \sum_{i=1}^{p} \bar{\mu}_i \nabla_x \mathcal{L}\left(\bar{x},\bar{y}^i,\bar{\gamma}^i\right) = 0, \tag{7}$$

$$\bar{\mu}_i \geq 0, \quad g\left(\bar{x},\bar{y}^i\right) = 0, i \in P, \tag{8}$$

where $\mathcal{L}\left(\bar{x},\bar{y}^i,\bar{\gamma}^i\right) = g(\bar{x}\bar{y}^i) - \sum_{j=1}^{q} \bar{\gamma}_j^I v_j(\bar{x}\bar{y}^i)$.

Next, the upper level first order condition is complemented by a lower level first order condition. In fact,

*e-mail: aozturan@hacettepe.edu.tr

since the active indices $\bar{y}^i \in Y_0(\bar{x}), i \in P$ are global solutions of $Q(\bar{x})$, under some constraint qualification like the Slater condition in the lower level problem, there exist vectors of the Lagrange multipliers $\bar{\gamma}^i \in R^q$ such that

$$\nabla_y g\left(\bar{x}, \bar{y}^i\right) - \sum_{j=1}^{q} \bar{\gamma}_j^i \nabla_y v_j\left(\bar{x}, \bar{y}^i\right) = 0,\ i \in P, \tag{9}$$

$$\bar{\gamma}_j^i \geq 0,\ v_j\left(\bar{x}, \bar{y}^i\right) \leq 0,\ \bar{\gamma}_j^i v_j\left(\bar{x}, \bar{y}^i\right) = 0,$$

$$i \in P,\ j \in Q. \tag{10}$$

Now, with any NCP function ψ, the solution of upper and lower level first order condition is seen to be equivalent to finding a zero of the following function:

$$T(z) = \begin{pmatrix} \nabla f(x) + \sum\limits_{i=1}^{p} \mu_i \nabla_x \mathcal{L}\left(x, y^i, \gamma^i\right) \\ \psi(\mu_1, -g(x, y^1)) \\ \vdots \\ \psi(\mu_p, -g(x, y^p)) \\ \nabla_y g\left(x, y^1\right) - \sum\limits_{j=1}^{q} \gamma_j^1 \nabla_y v_j(x, y^1) \\ \psi(\gamma_1^1, -v_1(x, y^1)) \\ \vdots \\ \psi(\gamma_q^1, -v_q(x, y^1)) \\ \vdots \\ \nabla_y g\left(x, y^p\right) - \sum\limits_{j=1}^{q} \gamma_j^p \nabla_y v_j(x, y^p) \\ \psi(\gamma_1^p, -v_1(x, y^p)) \\ \vdots \\ \psi(\gamma_q^p, -v_q(x, y^p)) \end{pmatrix} \tag{11}$$

T is strongly semismooth under our assumptions. Define some functions as follows:

$$\mathcal{L}(z) = \nabla f(x) + \sum_{i=1}^{p} \mu_i \nabla_x \mathcal{L}\left(x, y^i, \gamma^i\right),$$

$$\Psi(z) = \left(\psi_1(z), \ldots, \psi_p(z)\right)^{\mathrm{T}},$$

$$\psi_i(z) = \psi\left(\mu_i, -g\left(x, y^i\right)\right),\ i \in P,$$

$$l(z) = \left(l_1(z)^{\mathrm{T}}, \ldots, l_p(z)^{\mathrm{T}}\right)^{\mathrm{T}},$$

$$l_i(z) = \nabla_y g\left(x, y^i\right) - \sum_{j=1}^{q} \gamma_j^i \nabla_y v_j\left(x, y^i\right),\ i \in P,$$

$$\Phi(z) = \left(\Phi_1(z)^{\mathrm{T}}, \ldots, \Phi_p(z)^{\mathrm{T}}\right)^{\mathrm{T}},$$

$$\Phi_i(z) = \left(\psi_{i1}(z), \ldots, \psi_{iq}(z)\right)^{\mathrm{T}},\ i \in P,$$

$$\psi_{ij}(z) = \psi\left(\gamma_j^i, -v_j\left(x, y^i\right)\right),\ i \in P,\ j \in Q. \tag{12}$$

1.3. FTR algorithm

In this section, a filter trust-region (FTR) algorithm is presented to solve $T(z) = 0$. Let $F(z) = \frac{1}{2} T(z)^{\mathrm{T}} T(z)$ be the merit function. In order to use a filter technique, we define the function θ as follows: $\theta(z) = (\theta_1(z)\, \theta_2(z)\, \theta_3(z)\, \theta_4(z))$ with $\theta_1(z) = \|L(z)\|, \theta_2(z) = \|\Psi(z)\|, \theta_3(z) = \|l(z)\|, \theta_4(z) = \|\Phi(z)\|$. A filter is a list F of 4-tuples of the form $(\theta_{1,k}, \ldots, \theta_{4,k})$ such that $\theta_{j,k} < \theta_{j,l}$ for at least one $j \in \{1,2,3,4\}$. A new trial iterate z_k^+ is acceptable for the filter F if and only if for all $\theta_l = \theta(z_l) \in F$, there exists $j \in \{1,2,3,4\}$ such that

$$\theta_j\left(z_k^+\right) \leq \theta_{j,l} - \gamma_\theta \delta(\|\theta_l\|, \|\theta_k^+\|), \tag{13}$$

where $\gamma_\theta \in \left(0, \frac{1}{2}\right)$ 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 we simply perform the operation $F := F \cup \theta_k^+$. From a current iterate z_k one computes a trial step d_k by solving a system of linear equation

$$\left(W_k^{\mathrm{T}} W_k + \frac{1}{h_k}\right) d = -W_k^{\mathrm{T}} T(z_k), \tag{14}$$

thus avoiding solving a quadratic programming subproblem with a trust region bound.

Algorithm [14]

Step 1. Let $z_1 \in R^N, h_1 > 0,\ \varepsilon \geq 0,\ 0 < \rho < 1,\ 0 < \gamma_\theta < 1$ *and* $F = \emptyset,\ k = 1$.

Step 2. Choose $W_k \in \partial T(z_k)$. *If* $\|W_k^{\mathrm{T}} T(z_k)\| \leq \varepsilon$, *stop.*

Step 3. Solve $\left(W_k^{\mathrm{T}} W_k + \frac{1}{h_k}\right) d = -W_k^{\mathrm{T}} T(z_k)$ *to obtain* d_k. *Set* $z_k^+ = z_k + d_k$.

Step 4. Calculate $\rho_k = \frac{Ared_k}{Pred_k} = \frac{F(z_k) - F(z_k^+)}{F(z_k) - q_k(d_k)}$ *where* $q_k(d) = \frac{1}{2}\|T(z_k) + W_k d\|^2$.

Step 5. If θ_k^+ *is not acceptable for the current filter, go to Step 6. Otherwise, set* $z_{k+1} = z_k^+$ *(called a succesful iteration) and go to Step 7.*

Step 6. If $\rho_k \geq \rho_0$, *then set* $z_{k+1} = z_k^+$, $h_{k+1} = 2h_k$, *and go to Step 8. Otherwise, set* $h_k := \frac{1}{2} h_k$ *and go to Step 3 (called inner cycle).*

Step 7. If $\rho_k \geq \rho_0$, *then set* $h_{k+1} = 2h_k$, *and go to Step 8. Otherwise, add* θ_k^+ *to the filter F$ and set* $h_{k+1} := \frac{1}{2} h_k$, *and go to Step 8.*

Step 8. Set $k := k + 1$ *and go to Step 2 (called outer cycle).*

Under convexity of lower level and conditions for the Clarke subdifferential regularity of generalized Jacobian at solution point (see [15]) it can be shown that every accumulation point of the sequence z_k is a solution of $T(z) = 0$ and thus a stationary point of GSIP. Moreover, it is also possible to show that the algorithm is superlinearly convergent under a bound condition on $\|d\|$ (see [14]).

1.4. Numerical results

The Algorithm is implemented in Matlab 7.8. Throughout the computational experiments, the parameters used in the algorithm are $h_1 = 10$, $\varepsilon = 10^{-6}$, $\rho_0 = 0.1$, $\gamma_\theta = 0.001$. The algorithm is terminated when $\|W_k^{\mathrm{T}} T(z_k)\| < 10^{-6}$.

Example 1. In a general design centering problem, the aim is to maximize some measure (e.g., the volume $Vol(B(x))$) of a body $B(x)$ under the constraint that

$B(x)$ is contained in a given fixed body G. Let fixed body is given by $G = \{y \in R^2 | g(y) \leq 0\}$ with

$$g(y) = \left(-y_1 - y_2^2, \frac{y_1}{4} + y_2 - \frac{3}{4}, -y_2 - 1\right)^{\mathrm{T}}. \quad (15)$$

The GSIP formulation of the general design centering problem is as follows:

$$\max_{x \in R^n} Vol(B(x)) \text{ s.t. } g(y) \leq 0 \ \forall y \in B(x). \quad (16)$$

Problem 1. Consider the problem of finding the largest disc with free center and radius inscribed in G. We then have $n = 3$ and $B(x) = \{y \in R^2 | (y_1 - x_1)^2 + (y_2 - x_2)^2 - x_3^2 \leq 0\}$, $Vol(B(x)) = \pi x_3^2$. The FTR method obtains the optimal value 1.8606 after 5 iterations within 0.44 s of CPU time with $\|T(\bar{z})\| = 5.1023 \times 10^{-9}$.

Problem 2. The aim is to find the largest ellipse with free center and axis lengths inscribed in G. We have $n = 4$ and $B(x) = \{y \in R^2 | \frac{(y_1 - x_1)^2}{x_3^2} + \frac{(y_2 - x_2)^2}{x_4^2} - 1 \leq 0\}$, $Vol(B(x)) = \pi x_3 x_4$. The FTR method obtains the optimal value 3.484 after 7 iterations within 0.53 s of CPU time with $\|T(\bar{z})\| = 1.3814 \times 10^{-9}$.

Problem 3. Consider the problem of finding largest ellipsoid in a given simple polyhedron G. The ellipsoid is defined by $B(x) = \{y \in R^2 | \frac{(y_1 - x_1)^2}{x_4^2} + \frac{(y_2 - x_2)^2}{x_5^2} + \frac{(y_3 - x_3)^2}{x_6^2} - 1 \leq 0\}$, $Vol(B(x)) = \frac{4}{3}\pi x_4 x_5 x_6$. The FTR method obtains the optimal value 11890 after 8 iterations within 0.61 s of CPU time with $\|T(\bar{z})\| = 3.1045 \times 10^{-7}$.

Problem 4. Let us find the largest simple diamond inscribed in a given simple polyhedron. The diamond is described by $B(x) = \{y \in R^3 | v(x, y) \leq 0\}$ where $v(x, y) = (v_1(x, y), \ldots, v_{16}(x, y) v_{17}(xy))$ with $v_1, \ldots, v_8$ defines upper planes, $v_9, \ldots, v_{16}$ defines lower planes and v_{17} is the capping plane for diamond. The volume is given by $Vol(B(x)) = \frac{8}{3} \tan\left(\frac{\pi}{16}\right) \left[x_3^3 \left(\frac{1}{x_1} - \frac{1}{x_2}\right) - (x_3 - x_4 x_1)^2 \left(\frac{x_3}{x_1} - x_4\right)\right]$. The FTR method obtains the optimal value 1.398 after 5 iterations within 0.25 s of CPU time with $\|T(\bar{z})\| = 9.8142 \times 10^{-8}$.

Example 2. In robust optimization problems the data are uncertain and only known to belong to some uncertainty set which may be taken as infinite index set in semi-infinite programming. The following robust portfolio optimization problem is originally taken from [16].

Problem 5. Let 1€ be invested in a portfolio comprised of K shares. At the end of a given period the return of share i is $y_i > 0$. The goal is to determine the amount x_i to be invested in share i, $i = 1, \ldots, K$, so as to maximize the end-of-period portfolio value $y^{\mathrm{T}} x$. GSIP formulation is

$$\max_{x,y} z \text{ s.t. } z - y^{\mathrm{T}} x \leq 0 \ \forall y \in Y(x), \ \sum_{i=1}^{K} x_i = 1,$$
$$x \geq 0, \quad (17)$$

where $Y(x) = \left\{y \in R^K | \sum_{i=1}^{K} \frac{(y_i - \bar{y}_i)^2}{\sigma_i^2} \leq \Theta(x)\right\}$, $\Theta(x) = \theta\left(1 + \sum_{i=1}^{K} (x_i - \frac{1}{N})^2\right)$.

Here, the uncertainty set $Y(x)$ depends on x in which the risk aversion of the decision maker depends on the point x. The columns of Table is labelled as follows: K is the number of shares, ov is the optimal value, $\|T(\bar{z})\|$ is the Euclidean norm of $T(z)$ at the last iteration point, CPU is the CPU time for iterations in seconds, and iter is the number of iterations.

TABLE

Results of FTR method for optimal portfolio problem.

K	ov	$\|T(\bar{z})\|$	CPU	iter
10	0.7033	3.1265×10^{-9}	0.39	7
50	0.9638	1.2811×10^{-8}	0.54	9
100	1.0259	3.1504×10^{-7}	1.67	13
150	1.0535	7.6355×10^{-9}	5.12	19

It can be checked that in the problems in Example 1 and Example 2 strict complementarity holds in the upper and lower level problems, so that we actually have a smooth system. Now, for an illustration of the case that strict complementarity is violated in the upper level or lower level, we give the following examples from GSIP.

Problem 6. Let us consider the following GSIP from [7]:

$$\min x_1^2 + x_2^2 \text{ s.t. } x_2 - y \leq 0 \ \forall y \in Y(x)$$
$$\text{where } Y(x) = \{y \in R | x_1 \leq 0, -x_1 - y \leq 0\}.$$

Note that the strict complementarity is violated in the upper level. $T(z)$ is nonsmooth at the solution. The FTR method obtains optimal value 0 after 5 iterations within 0.16 s of CPU time with $\|T(\bar{z})\| = 1.7023 \times 10^{-8}$.

Problem 7. Let us consider the following GSIP:

$$\min x_1^2 + x_2^2 \text{ s.t. } (y_1 - x_1)^2 - (y_2 - x_2)^2 \leq 0$$
$$\forall y \in Y(x) \text{ where } Y(x) = \{y \in R | y_1 - x_1 \leq 0,$$
$$y_2 - x_2 \leq 0\}.$$

Note that the strict complementarity is violated in the lower level. $T(z)$ is nonsmooth at the solution. The FTR method obtains optimal value 0 after 6 iterations within 0.17 s of CPU time with $\|T(\bar{z})\| = 3.5123 \times 10^{-7}$.

Problem 8. Let us consider the following GSIP:

$$\min x_1^2 + x_2^2 + x_3^2 \text{ s.t. } -\frac{1}{2}(y_1 - x_1)^2 - (y_2 - x_2)^2$$
$$-x_3 \leq 0 \ \forall y \in Y(x) \text{ where } Y(x) = \{y \in R^2 | -y_1$$
$$-y_2 - x_1 \leq 0, -y_2 - x_2 \leq 0, y_1^2 + y_2^2 - 1 \leq 0\}.$$

The strict complementarity for both levels is violated at the solution. The FTR method obtains optimal value 0 after 5 iterations within 0.18 s of CPU time with $\|T(\bar{z})\| = 1.5223 \times 10^{-8}$.

In the numerical examples, $\psi_{\min}(a, b) = -\min\{a, b\}$ NCP function is also tested and it is observed that the performance of the method does not change significantly if ψ_{FB} is replaced by $\psi_{\min}$. Moreover, it is also observed that computational results are not very sensitive to the choice of the parameter h_1.

2. Conclusion

In this paper a filter trust region method is applied for solving the generalized semi-infinite programming problems. An advantage of the method is that at each iteration, only a system of linear equations is solved to get search direction. We point out that to get feasible points lower level problem must be convex. We transformed the method from standard to generalized semi-infinite programming problems. For convergence Clarke subdifferential regularity of generalized Jacobian at the solution point (for conditions see [15]) is needed and it holds under standard conditions from semi-infinite programming. In the approach in [14], strict complementarity is a part of their assumptions. We also point out that nonsmoothness is caused by a possible violation of strict complementarity slackness in the lower level or/and in the upper level. Comparing the results of FTR method for common problems in [13] shows that the proposed method is comparable to given paper in computational effort. For future work more sophisticated numerical examples should be presented. However, in the case of more variables (e.g. design centering), the system of semismooth nonlinear equations is large-scale and ill-conditioned. As a future work, preconditioning techniques can be considered to solve mentioned systems.

References

[1] M.A. Goberna, M. Lopez, *Linear Semi-Infinite Optimization*, Wiley, Chichester 1998.

[2] R. Hettich, K. Kortanek, *SIAM Rev.* **35**, 380 (1993).

[3] M. Lopez, G. Still, *Europ. J. Oper. Res.* **180**, 491 (2007).

[4] R.M. Reemtsen, S. Görner, in: *Semi-Infinite Programming*, Eds. R.M. Reemtsen, J. Rückmann, in: *Kluwer Nonconvex Optim. Appl.*, Vol. 25, Springer, p. 195.

[5] *Semi-Infinite Programming*, Eds. R. Reemtsen, J.-J. Rückmann, Kluwer, Boston 1998.

[6] O. Stein, *Bilevel Strategies in Semi-Infinite Programming*, Kluwer, Boston 2003.

[7] G. Still, *Europ. J. Operat. Res.* **119**, 301 (1999).

[8] G. Still, *Optimization* **49**, 223 (2001).

[9] F.G. Vazquez, J.-J. Rückmann, O. Stein, G. Still, *J. Comput. Appl. Math.* **217**, 394 (2008).

[10] G.-W. Weber, *Generalized Semi-Infinite Optimization and Related Topics*, Heldermann Verlag, Lemgo (Germany) 2003.

[11] L. Qi, S.-Y. Wu, G. Zhou, *J. Global Optim.* **27**, 215 (2003).

[12] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York 1983.

[13] O. Stein, A. Tezel, *J. Global Optimiz.* **41**, 245 (2008).

[14] O. Yi-gui, *J. Appl. Math. Comput.* **29**, 311 (2009).

[15] O. Stein, A. Tezel, *SIAM J. Optimiz.* **20**, 1052 (2009).

[16] A. Ben-Tal, A. Nemirovski, *Operat. Res. Lett.* **25**, 1 (1999).