Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 131938, 8 pages
http://dx.doi.org/10.1155/2013/131938

Research Article

Optimality Conditions for Nonsmooth Generalized Semi-Infinite Programs

Zhangyou Chen[1] and Zhe Chen[2]

[1] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*
[2] *Business School, Sichuan University, Chengdu 610064, China*

Correspondence should be addressed to Zhe Chen; chenzhe@scu.edu.cn

Received 30 July 2013; Accepted 21 August 2013

Academic Editor: Jen-Chih Yao

We consider a class of nonsmooth generalized semi-infinite programming problems. We apply results from parametric optimization to the lower level problems of generalized semi-infinite programming problems to get estimates for the value functions of the lower level problems and thus derive necessary optimality conditions for generalized semi-infinite programming problems. We also derive some new estimates for the value functions of the lower level problems in terms of generalized differentiation and further obtain the necessary optimality conditions.

1. Introduction

Generalized semi-infinite programming problem (GSIP) is of the form

$$\begin{aligned} &\min \quad f(x) \\ &\text{s.t.} \quad x \in \mathbb{R}^n,\ g(x,y) \le 0,\ y \in Y(x), \end{aligned} \tag{1}$$

where $Y(x) := \{y \in \mathbb{R}^m \mid v(x,y) \le 0\}$. GSIP is different from the standard semi-infinite programming in that its index set Y is dependent on x.

This first systematic study of GSIP was Hettich and Still [1] where the reduction method was used to reduce GSIP into standard nonlinear programming problems and second-order optimality conditions were derived. Necessary optimality conditions at an optimal solution $\overline{x}$ for (1) with differentiable data are as follows there exist nonnegative numbers $\lambda_0, \ldots, \lambda_p$, not all zero, such that

$$\lambda_0 \nabla f(\overline{x}) + \sum_{j=1}^{p} \lambda_j \nabla_x L\left(\overline{x}, \overline{y}^j, \alpha^j, \beta^j\right) = 0, \tag{2}$$

where $L(x,y,\alpha,\beta) = \alpha g(x,y) - \langle \beta, v(x,y)\rangle$ and each (α^j, β^j) is the usual FJ-multiplier of the lower level problem at its optimal solution $\overline{y}^j$

$$\begin{aligned} Q(\overline{x}) \quad &\max \quad g(\overline{x}, y) \\ &\text{s.t.} \quad v(\overline{x}, y) \le 0. \end{aligned} \tag{3}$$

This condition was first derived by Jongen et al. [2] in an elementary way without any constraint qualifications or any kind of reduction approaches. They also proposed a constraint qualification under which it follows that $\lambda_0 > 0$ and discussed some geometrical properties of the feasible set which do not appear in standard semi-infinite case. The optimality conditions are further explored by Rückmann and Shapiro [3] and Stein [4].

GSIP in itself is of complex and exclusive structures such as the nonclosedness of the feasible set, nonconvexity, nonsmoothness, and bilevel structure and thus a difficult problem to solve see, for example, [2, 5, 6]. We also refer to [7–10] for some recent study on the structure of GSIP.

It is obvious that GSIP can be rewritten equivalently as the nonlinear programming problem

$$\begin{aligned} &\min \quad f(x) \\ &\text{s.t.} \quad \phi(x) \le 0, \end{aligned} \tag{4}$$

where $\phi(x)$ is the optimal value of $Q(x)$. Then we can relate GSIP to the min-max problem

$$\min_x \max\{f(x) - f(\overline{x}), \phi(x)\}; \tag{5}$$

see [3] for more details. On the other hand, GSIP can be related to the following bilevel problem:

$$\begin{aligned} &\min_{(x,y)} \quad f(x) \\ &\text{s.t.} \quad g(x,y) \le 0, \\ &\qquad y \in \arg\max\{g(x,y) : y \in Y(x)\}. \end{aligned} \tag{6}$$

The problem (6) is a special bilevel optimization problem in that its upper level constraint is the same as the objective function of its lower level problem. However, there is a *slight difference* between GSIP problem (1) and problem (6). The feasible set of (6) is a subset of (1) in that the feasible set of (1) is the combination of the feasible set of (6) and the complement of dom Y. For more comparisons between GSIP and bilevel problems, see Stein and Still [11].

The bilevel problem (6) is equivalent to the following problem:

$$\begin{aligned} &\min_{(x,y)} \quad f(x) \\ &\text{s.t.} \quad (x,y) \in \Omega,\ G(x,y) \le 0, \end{aligned} \tag{7}$$

where $\Omega = \{(x,y) \mid g(x,y) \le 0, v(x,y) \le 0\}$ and $G(x,y) = \phi(x) - g(x,y)$. This approach was used by Dempe and Zemkoho [12] to study bilevel optimization problems. For general references of bilevel optimization, see [13].

In this paper we concentrate on the optimality conditions of nonsmooth GSIP whose defining functions are Lipschitz continuous. Similar works are [14, 15]. We achieve this via the lower level optimal value function reformulation and then derive its necessary optimality conditions via the generalized differentiation. One of the key steps is to estimate the generalized gradients of the lower level optimal value function which involves parametric optimization. We will consider two cases with different approaches related to the two reformulations of GSIP as previously mentioned. Firstly, we develop optimality conditions via the min-max formulation with Lipschitz lower level optimal value function. Secondly, we develop optimality conditions via bilevel formulation under the assumption of partial calmness.

2. Preliminaries

In this section, we present some basic definitions and results from variational analysis [16, 17]. Given a set A in $\mathbb{R}^n$, the regular normal cone $\widehat{N}_A$ of A at $\overline{x} \in A$ is defined by

$$\widehat{N}_A(\overline{x}) = \left\{ v \in \mathbb{R}^n \mid \limsup_{x \xrightarrow{A} \overline{x}} \frac{\langle v, x - \overline{x} \rangle}{\|x - \overline{x}\|} \le 0 \right\}. \tag{8}$$

The (general) normal cone N_A of A at $\overline{x}$ is defined by

$$\begin{aligned} N_A(\overline{x}) &= \limsup_{x \to \overline{x}} \widehat{N}_A(x) \\ &= \left\{ v \mid \exists x^k \longrightarrow \overline{x},\ \exists v^k \longrightarrow v \text{ with } v^k \in \widehat{N}_A\left(x^k\right) \right\}. \end{aligned} \tag{9}$$

Given a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and a point $\overline{x}$ with $f(\overline{x})$ finite, denote by epi f the epigraph of f. The regular subdifferential of f at $\overline{x}$ is defined by

$$\widehat{\partial} f(\overline{x}) := \left\{ u \in \mathbb{R}^n \mid \liminf_{x \to \overline{x}, x \ne \overline{x}} \frac{\left(f(x) - f(\overline{x}) - \langle u, x - \overline{x} \rangle\right)}{\|x - \overline{x}\|} \ge 0 \right\}. \tag{10}$$

The general (basic, limiting) and singular subdifferential of f at $\overline{x}$ are defined, respectively, by

$$\begin{aligned} \partial f(\overline{x}) &:= \limsup_{x \xrightarrow{f} \overline{x}} \widehat{\partial} f(x), \\ \partial^\infty f(\overline{x}) &:= \limsup_{x \xrightarrow{f} \overline{x}, \lambda \downarrow 0} \lambda \widehat{\partial} f(x). \end{aligned} \tag{11}$$

The upper regular subdifferential of f at $\overline{x}$ is defined by $\widehat{\partial}^+ f(\overline{x}) := -\widehat{\partial}(-f)(\overline{x})$, and the upper subdifferential of f is defined by $\overline{x}\partial^+ f(\overline{x}) := \limsup_{x \xrightarrow{f} \overline{x}} \widehat{\partial}^+ f(x)$.

The Clarke (convexified) normal cone can be defined by two different approaches. On the one hand, it can be defined by the polar cone of the Clarke's tangent cone

$$\overline{N}_A(x) := \widehat{T}_A(x)^\circ, \tag{12}$$

where $\widehat{T}_A(x) = \liminf_{y \xrightarrow{A} x, t \searrow 0} (A - y)/t$ or defined via the generalized directional derivative of the (Lipschitzian) distant function dist$(\cdot, A)$; see Clarke [18]. On the other hand, it can be defined by the closed convex hull of the (general) normal cone

$$\overline{N}_A(x) := \operatorname{cl} \operatorname{co} N_A(x). \tag{13}$$

For this definition and also the equivalence of the two definitions, see, for example, Rockafellar and Wets [16].

The Clarke subgradients and Clarke horizon subgradients of f at x are defined by

$$\begin{aligned} \overline{\partial} f(x) &:= \left\{ v \mid (v, -1) \in \overline{N}_{\operatorname{epi} f}(x, f(x)) \right\}, \\ \overline{\partial}^\infty f(x) &:= \left\{ v \mid (v, 0) \in \overline{N}_{\operatorname{epi} f}(x, f(x)) \right\}. \end{aligned} \tag{14}$$

The relationship between the Clarke sub-subdifferentials and basic sub-differentials is also referred to by Mordukhovich [17, Theorem 3.57].

Proposition 1 (see [16]). *Let f be proper and* lsc *around $\overline{x} \in$* dom f. *Then,*

$$\overline{\partial} f(\overline{x}) = \operatorname{cl}\operatorname{co}\left[\partial f(\overline{x}) + \partial^{\infty} f(\overline{x})\right]. \tag{15}$$

If, in particular, f is Lipschitz continuous at $\overline{x}$, then

$$\overline{\partial} f(\overline{x}) = \operatorname{cl}\operatorname{co}\ \partial f(\overline{x}). \tag{16}$$

The normal cone N_A enjoys the robustness property $N_A(\overline{x}) = \limsup_{x\to\overline{x}} N_A(x)$ provided that the setting is finite dimension [17, page 11]. However, this is not true for the convexified cone $\overline{N}_A$; see, for example, Rockafellar [19]. Consider

$$A := \left\{x \in \mathbb{R}^3 \mid x_3 = x_1x_2 \text{ or } x_3 = -x_1x_2\right\}, \tag{17}$$

$$\overline{x} = (0,0,0). \tag{18}$$

The normal cone $\overline{N}_A(\overline{x})$ is just the x_3-axis, but $\overline{N}_A(x)$ is the x_2x_3-plane for all $x = (x_1, 0, 0)$. The following proposition is from Rockafellar [19].

Proposition 2 (see [19]). *If A is convex, or if $\overline{N}_A(\overline{x})$ is pointed, then the multifunction $\overline{N}_A$ is closed at $\overline{x}$; that is, for all $x^k \to \overline{x}$, $y^k \xrightarrow{\overline{N}_A(x^k)} \overline{y}$, one has $\overline{y} \in \overline{N}_A(\overline{x})$.*

Proposition 3. *The Clarke normal cone has the robustness property*

$$\overline{N}_A(x) = \limsup_{y\to x} \overline{N}_A(y) \tag{19}$$

provided that $N_A(x)$ is pointed.

Proof. It suffices to prove that $\limsup_{y\to x}\operatorname{co} N_A(y) \subset \operatorname{cl}\operatorname{co}\limsup_{y\to x} N_A(y) = \overline{N}_A(x)$. Let $v \in \limsup_{y\to x}\operatorname{co} N_A(y)$. Then there are $y_k \in A$, $v_{ik} \in N_A(y_k)$, $i = 1,\ldots,n+1$ such that

$$\sum_{i=1}^{n+1} v_{ik} \longrightarrow v \quad \text{as } k \longrightarrow \infty, \tag{20}$$

since the sets $N_A(y_k)$ are cones. Let $\lambda_k = \sum_{i=1}^{n+1}\|v_{ik}\|$. Then $\{\lambda_k\}$ is bounded; that is also to say, $\{v_{ik}\}$ are bounded for all i. Otherwise, $\sum_{i=1}^{n+1} v_{ik}/\lambda_k \to 0$. That is, $v_1+\cdots+v_{n+1} = 0$, where v_i is the limit of $\{v_{ik}/\lambda_k\}_k$ for each i. Note that $v_i \in N_A(x)$ since $N_A(x) = \limsup_{y\to x} N_A(y)$. Thus $v_1 = \cdots = v_{n+1} = 0$ by the pointedness of $N_A(x)$. On the other hand, $\sum_i \|v_i\| = 1$. This is a contradiction. Thus the sequence $\{v_{ik}\}$ is bounded. By taking subsequences, we may assume that $v_{ik} \to v_i$. Then $v_i \in N_A(x)$ and $v = v_1 + \cdots + v_{n+1}$. This completes the proof. □

The following definitions are required for further development.

Definition 4 (see [17, Definition 1.63]). Let $S : X \rightrightarrows Y$ be a set-valued mapping with $\overline{x} \in \operatorname{dom} S$, the domain of S.

(i) Given $\overline{y} \in S(\overline{x})$, we say that the mapping S is inner semicontinuous at $(\overline{x}, \overline{y})$ if for every sequence $x_k \to \overline{x}$ there is a sequence $y_k \in S(x_k)$ converging to $\overline{y}$ as $k \to \infty$.

(ii) S is inner semicompact at $\overline{x}$ if for every sequence $x_k \to \overline{x}$ there is a sequence $y_k \in S(x_k)$ that contains a convergent subsequence as $k \to \infty$.

(iii) S is μ-inner semicontinuous at $(\overline{x}, \overline{y})$ (μ-inner semicompact at $\overline{x}$) if in above two cases, $x_k \to \overline{x}$ is replaced by $x_k \to \overline{x}$ with $\mu(x_k) \to \mu(\overline{x})$.

Here the concept of μ-inner semicontinuity/semicompactness is important for our considerations. It is typical that the value function ϕ of the lower level problem $Q(x)$ of GSIP is not continuous, even taking value $-\infty$.

Theorem 5 (subdifferentiation of maximum functions [17, Theorem 3.46]). *Consider the maximum function of the form*

$$(\max \phi_i)(x) = \max\{\phi_i(x) \mid i = 1,\ldots,l\}. \tag{21}$$

Let ϕ_i be lower semicontinuous around $\overline{x}$ for $i \in I(\overline{x}) := \{i \mid \max(\phi_i)(\overline{x}) = \phi_i(\overline{x})\}$ and be upper semicontinuous at $\overline{x}$ for $i \notin I(\overline{x})$. Assume that the qualification holds:

$$\left[\sum_{i\in I(\overline{x})} x_i^* = 0,\ x_i^* \in \partial^{\infty}\phi_i(\overline{x})\right] \Longrightarrow x_i^* = 0, \quad i \in I(\overline{x}). \tag{22}$$

Then

$$\partial(\max\phi_i)(\overline{x}) \subset \bigcup\left\{\sum_{i\in I(\overline{x})} \lambda_i \partial\phi_i(\overline{x}) \mid (\lambda_1,\ldots,\lambda_l) \in \Lambda(\overline{x})\right\}, \tag{23}$$

where $\Lambda(\overline{x}) = \{(\lambda_1,\ldots,\lambda_l) \mid \lambda_i \geq 0, \sum_{i=1}^{l}\lambda_i = 1, \lambda_i(\phi_i(\overline{x}) - (\max\phi_i)(\overline{x})) = 0\}$.

Note that qualification (22) always holds if all related functions are locally Lipschitz.

The following two results are about continuity properties and estimates of subdifferentials of marginal functions which are crucial to our analysis for GSIP problems.

Proposition 6 (limiting subgradients of marginal functions [20]). *Consider the parametric optimization problem*

$$\mu(x) := \inf\{\varphi(x,y) \mid y \in G(x)\}, \tag{24}$$

and let $M(x) := \{y \in G(x) \mid \mu(x) = \varphi(x,y)\}$, $G(x) := \{y \in \mathbb{R}^m \mid \varphi_i(x,y) \leq 0,\ i = 1,\ldots,l\}$. For simplicity, one does not consider the case with equality constraints involved.

(i) *Assume that M is μ-inner semicontinuous at $(\overline{x},\overline{y}) \in \operatorname{gph} M$ (the graph of M), that φ and all φ_i are Lipschitz continuous around $(\overline{x},\overline{y})$, and that the following qualification condition is satisfied:*

$$\begin{aligned} &\text{only the vector } (\lambda_1,\ldots,\lambda_l) = 0 \in \mathbb{R}^l \\ &\text{satisfies the relation } 0 \in \sum_{i=1}^{l} \lambda_i \partial\varphi_i(\overline{x},\overline{y}) \\ &\text{for some } (\lambda_1,\ldots,\lambda_l) \in \mathbb{R}^l_+ \\ &\text{with } \lambda_i\varphi_i(\overline{x},\overline{y}) = 0, \quad i = 1,\ldots,l. \end{aligned} \tag{25}$$

One has the inclusions

$$\begin{aligned} \partial\mu(\overline{x}) \subset \Bigg\{ & u^* \in X^* \mid (u^*,0) \in \partial\varphi(\overline{x},\overline{y}) \\ & + \sum_{i=1}^{l} \lambda_i \partial\varphi_i(\overline{x},\overline{y}) \ \text{ for some } (\lambda_1,\ldots,\lambda_l) \in \mathbb{R}^l_+ \\ & \text{with } \lambda_i\varphi_i(\overline{x},\overline{y}) = 0,\ i = 1,\ldots,l \Bigg\}. \end{aligned} \tag{26}$$

(ii) *Assume that M is μ-inner semicompact at $\overline{x}$, and all φ and φ_i are Lipschitz continuous around $(\overline{x},\overline{y})$ for all $\overline{y} \in M(\overline{x})$, and qualification (25) holds for all $(\overline{x},\overline{y})$, $\overline{y} \in M(\overline{x})$. Then*

$$\begin{aligned} \partial\mu(\overline{x}) \subset \bigcup_{\overline{y}\in M(\overline{x})} \Bigg\{ & u^* \in X^* \mid (u^*,0) \in \partial\varphi(\overline{x},\overline{y}) \\ & + \sum_{i=1}^{l} \lambda_i \partial\varphi_i(\overline{x},\overline{y}), \text{ for some } (\lambda_1,\ldots,\lambda_l) \in \mathbb{R}^m_+ \\ & \text{with } \lambda_i\varphi_i(\overline{x},\overline{y}) = 0,\ i = 1,\ldots,l \Bigg\}. \end{aligned} \tag{27}$$

Proposition 7 (Lipschitz continuity of marginal functions [21, Theorem 5.2]). *Continue to consider the parametric problem (24) in Proposition 6. Then the following assertions hold.*

(i) *Assume that M is μ-inner semicontinuous at $(\overline{x},\overline{y}) \in \operatorname{gph} M$ and φ is locally Lipschitz around this point. Then μ is Lipschitz around $\overline{x}$ provided that it is* lsc *around $\overline{x}$ and G is Lipschitz-like around $(\overline{x},\overline{y})$.*

(ii) *Assume that M is μ-inner compact at $\overline{x}$ and φ is locally Lipschitz around $(\overline{x},\overline{y})$ for all $\overline{y} \in M(\overline{x})$. Then μ is Lipschitz around $\overline{x}$ provided that it is* lsc *around $\overline{x}$ and G is Lipschitz-like around $(\overline{x},\overline{y})$ for all $\overline{y} \in M(\overline{x})$.*

3. Main Results

Now we are prepared to develop the optimality conditions for GSIP problem (1). Given a local solution $\overline{x}$ of problem (1), associate it with the following min-max problem

$$\min_x \max\{f(x) - f(\overline{x}), \phi(x)\}, \tag{28}$$

where $\phi(x) := \sup_{y\in Y(x)} g(x,y)$. Let $Y_0(x) := \{y \in Y(x) : \phi(x) = g(x,y)\}$. Denote by

$$F(x) = \max\{f(x) - f(\overline{x}), \phi(x)\}. \tag{29}$$

If $\overline{x}$ solves GSIP (1), then $\overline{x}$ also solves problem

$$\min\{F(x) : x \in \mathbb{R}^n\}, \tag{30}$$

and thus by generalized Fermat's rule (cf. [16, Theorem 10.1]), we have

$$0 \in \partial F(\overline{x}). \tag{31}$$

So, calculus for the maximum function and the estimate of subdifferentials are essential to proceed. From (31) and Theorem 5, there exists $\mu \in [0,1]$ such that (if ϕ is Lipschitz)

$$0 \in \mu\partial f(\overline{x}) + (1-\mu)\partial\phi(\overline{x}) \subset \mu\partial f(\overline{x}) + (1-\mu)\operatorname{co}\partial\phi(\overline{x}). \tag{32}$$

Note that for a Lipschitz function ϕ,

$$\operatorname{co}\partial\phi(\overline{x}) = -\operatorname{co}\partial(-\phi)(\overline{x}). \tag{33}$$

Theorem 8 (optimality for GSIP with Lipschitz lower level optimal value function). *Consider the GSIP problem (1), and let $\overline{x} \in M$ be its locally optimal solution. Assume that all functions f, g, and v are Lipschitz continuous, Y_0 is ϕ-inner semicompact at $\overline{x}$, and Y is Lipschitz-like at $(\overline{x},\overline{y})$ for all $\overline{y} \in Y_0(\overline{x})$. Then there is $\overline{y}^j \in Y_0(\overline{x})$ and $\overline{\lambda}_0 \geq 0$, $\overline{\lambda}_j \geq 0$, $\overline{\alpha}_i^j \geq 0$, $i = 1,\ldots,l$, $j = 1,\ldots,k$ such that $\sum_{j=1}^{k} \overline{\lambda}_j = 1$ and*

$$0 \in \left(\overline{\lambda}_0\partial f(\overline{x}), 0\right) + \sum_{j=1}^{k} \overline{\lambda}_j \left[-\partial(-g)(\overline{x},\overline{y}^j) - \sum_{i=1}^{l} \overline{\alpha}_i^j \partial v_i(\overline{x},\overline{y}^j) \right]. \tag{34}$$

If in addition g and all components of v are regular at all $(\overline{x},\overline{y}^j)$, then the optimality is of the form

$$\begin{aligned} 0 &\in \overline{\lambda}_0\partial f(\overline{x}) + \sum_{j=1}^{k} \overline{\lambda}_j \left[-\partial_x(-g)(\overline{x},\overline{y}^j) - \sum_{i=1}^{l} \overline{\alpha}_i^j \partial_x v_i(\overline{x},\overline{y}^j) \right], \\ 0 &\in \sum_{j=1}^{k} \overline{\lambda}_j \left[-\partial_y(-g)(\overline{x},\overline{y}^j) - \sum_{i=1}^{l} \overline{\alpha}_i^j \partial_y v_i(\overline{x},\overline{y}^j) \right]. \end{aligned} \tag{35}$$

Note that $-\partial(-g) = \partial^+ g$.

Proof. Under regularity and Lipschitz continuity, since the following calculus rule for basic subgradients holds (see [16, Corollary 10.11]):

$$\partial_x f(x,y) = \{u \mid \exists v \text{ s.t. } (u,v) \in \partial f(x,y)\}, \quad (36)$$

the last two equations follow directly from the first one. Note that the function $-\phi$ is in position of μ in Proposition 6. Under our assumptions, by Proposition 7, $-\phi$ is Lipschitz continuous and the estimate of $\partial(-\phi)(\overline{x})$ is

$$\begin{aligned}\partial(-\phi)(\overline{x}) \subset \bigcup_{\overline{y}\in Y_0(\overline{x})} \Big\{ & u^* \in X^* \mid (u^*,0) \in \partial(-g)(\overline{x},\overline{y}) \\ & + \sum_{i=1}^{l} \alpha_i \partial v_i(\overline{x},\overline{y}) \\ & \text{for some } (\alpha_1,\dots,\alpha_l) \in \mathbb{R}^l_+ \\ & \text{with } \alpha_i v_i(\overline{x},\overline{y}) = 0,\ i=1,\dots,l \Big\}.\end{aligned} \quad (37)$$

If $\overline{x} \in M$ solves GSIP, then it also solves $\min_x F(x)$. By (31), (32), and (33), there exists $\mu \in [0,1]$ such that

$$0 \in \mu\partial f(\overline{x}) - (1-\mu)\operatorname{co}\partial(-\phi)(\overline{x}). \quad (38)$$

Combining (37) and (38), there is $\overline{y}^j \in Y_0(\overline{x})$ and $\lambda_j \geq 0$, $\alpha_i^j \geq 0$, $i=1,\dots,l$, $j=1,\dots,k$ such that $\sum_{j=1}^k \lambda_j = 1$ and

$$\begin{aligned}0 \in & (\mu\partial f(\overline{x}),0) \\ & + (1-\mu)\sum_{j=1}^{k}\lambda_j\left[-\partial(-g)(\overline{x},\overline{y}^j) - \sum_{i=1}^{l}\alpha_i^j\partial v_i(\overline{x},\overline{y}^j)\right].\end{aligned} \quad (39)$$

Letting $\overline{\lambda}_0 = \mu$, $\overline{\lambda}_j = (1-\mu)\lambda_j$, $\overline{\alpha}_i^j = \alpha_i^j$, $i=1,\dots,l$, $j=1,\dots,k$, we obtain the desired result. □

Corollary 9. *In addition to the assumptions in Theorem 8, assume that f, g, and v are continuously differentiable. Then the optimality condition at the optimal point $\overline{x}$ is that there exist $\overline{y}^j \in Y_0(\overline{x})$ and $\overline{\lambda}_0 \geq 0$, $\overline{\lambda}_j \geq 0$, $\overline{\alpha}_i^j \geq 0$, $i=1,\dots,l$, $j=1,\dots,k$ such that $\sum_{j=1}^k \overline{\lambda}_j = 1$ and*

$$\begin{aligned}0 &= \overline{\lambda}_0\nabla_x f(\overline{x}) + \sum_{j=1}^{k}\overline{\lambda}_j\left[\nabla_x g(\overline{x},\overline{y}^j) - \sum_{i=1}^{l}\overline{\alpha}_i^j\nabla_x v_i(\overline{x},\overline{y}^j)\right], \\ 0 &= \sum_{j=1}^{k}\overline{\lambda}_j\left[\nabla_y g(\overline{x},\overline{y}^j) - \sum_{i=1}^{l}\overline{\alpha}_i^j\nabla_y v_i(\overline{x},\overline{y}^j)\right].\end{aligned} \quad (40)$$

Next we consider the case where the lower level value function ϕ may fail to be Lipschitz and give estimates for the subdifferentials of ϕ and thus further derive the optimality conditions for GSIP. However, it requires to use the Clarke subdifferentials.

Proposition 10. *Consider the parametric optimization problem same as* (24)*:*

$$\mu(x) := \inf\{\varphi(x,y) \mid y \in G(x)\}, \quad (41)$$

with corresponding solution mapping $M : X \rightrightarrows Y$. Let $\overline{x} \in \operatorname{dom} M$. Assume that the following conditions hold:

(i) *φ is lower semicontinuous at $\overline{x}$;*

(ii) *M is μ-inner semicompact at $\overline{x}$ and $M(\overline{x})$ is nonempty and compact;*

(iii) *if $(u_i,w_i) \in \overline{\partial}^{\infty}\varphi(\overline{x},\overline{y}_i)$, $(v_i,-w_i) \in \overline{N}_{\operatorname{gph} G}(\overline{x},\overline{y}_i)$, $\overline{y}_i \in M(\overline{x})$, $i \leq n+1$, and $\sum_i u_i + v_i = 0$, then $u_i = v_i = 0$, $w_i = 0$;*

(iv) *the cones $N_{\operatorname{gph} G}(\overline{x},\overline{y})$ and $N_{\operatorname{epi}\varphi}(\overline{x},\overline{y},\varphi(\overline{x},\overline{y}))$ for all $\overline{y} \in M(\overline{x})$ are pointed.*

Then one has the inclusion

$$\overline{\partial}\mu(\overline{x}) \subset \operatorname{co}\left(\bigcup_{\overline{y}\in M(\overline{x})}\left\{u \mid (u,0) \in \overline{\partial}\varphi(\overline{x},\overline{y}) + \overline{N}_{\operatorname{gph} G}(\overline{x},\overline{y})\right\}\right). \quad (42)$$

Proof (sketch: the definition of $\overline{\partial} = \operatorname{cl}\operatorname{co}\{\partial + \partial^{\infty}\}$). The proof is divided into two parts. First the set on the right hand of the required inclusion, denoted by Λ, is closed. The second step is to justify that $\partial + \partial^{\infty} \subset \Lambda$.

Let $\sum_{i=1}^{n+1}\lambda_i^k u_i^k \to u$, $\lambda_i^k \geq 0$, $\sum_{i=1}^{n+1}\lambda_i^k = 1$, and

$$(u_i^k,0) = (x_{1i}^k,y_{1i}^k) + (x_{2i}^k,y_{2i}^k), \quad i=1,2,\dots,n+1, \quad (43)$$

where $(x_{1i}^k,y_{1i}^k) \in \overline{\partial}\varphi(\overline{x},\overline{y}_i^k)$, $(x_{2i}^k,y_{2i}^k) \in \overline{N}_{\operatorname{gph} G}(\overline{x},\overline{y}_i^k)$, $\overline{y}_i^k \in M(\overline{x})$, $i=1,\dots,n+1$. We have to show that $u \in \Lambda$. We may assume that $\overline{y}_i^k \to \overline{y}_i$, $i=1,\dots,n+1$, as $M(\overline{x})$ is compact. We show first that the sequence $\{z^k := (z_i^k)_i = (x_{1i}^k,y_{1i}^k,x_{2i}^k,y_{2i}^k)\}_k$ is bounded. Suppose on the contrary that $\|z^k\| \to \infty$. Then for each i,

$$\begin{aligned}\frac{1}{\|z^k\|}(u_i^k,0) &= \frac{1}{\|z^k\|}(x_{1i}^k,y_{1i}^k) + (u_{2i}^k,v_{2i}^k), \text{ with } (u_{2i}^k,v_{2i}^k) \\ &= \frac{1}{\|z^k\|}(x_{2i}^k,y_{2i}^k).\end{aligned} \quad (44)$$

Multiplying (44) by λ_i^k and taking summation over i, we have

$$\frac{1}{\|z^k\|}\left(\sum_{i=1}^{n+1}\lambda_i^k x_{1i}^k, \sum_{i=1}^{n+1}\lambda_i^k y_{1i}^k\right) + \left(\sum_{i=1}^{n+1}\lambda_i^k u_{2i}^k, \sum_{i=1}^{n+1}\lambda_i^k v_{2i}^k\right) \longrightarrow (0,0). \quad (45)$$

Note that for each i, by definition of $\overline{\partial}$,

$$\left(\frac{1}{\|z^k\|}(x_{1i}^k,y_{1i}^k), \frac{-1}{\|z^k\|}\right) \in \overline{N}_{\operatorname{epi}\varphi}(\overline{x},\overline{y}_i^k,\varphi(\overline{x},\overline{y}_i^k)). \quad (46)$$

Let $(u_{1i}^k, v_{1i}^k) := 1/\|z^k\|(x_{1i}^k, y_{1i}^k) \to (u_{1i}, v_{1i})$. Then by assumption (iv) and Proposition 3 one has $(u_{1i}, v_{1i}, 0) \in \overline{N}_{\operatorname{epi}\varphi}(\overline{x}, \overline{y}_i, \varphi(\overline{x}, \overline{y}_i))$, and thus $(u_{1i}, v_{1i}) \in \overline{\partial}^{\infty}\varphi(\overline{x}, \overline{y}_i)$. Then

$$\frac{1}{\|z^k\|}\left(\sum_{i=1}^{n+1}\lambda_i^k x_{1i}^k, \sum_{i=1}^{n+1}\lambda_i^k y_{1i}^k\right) \longrightarrow \sum_{i=1}^{n+1}\lambda_i\left(u_{1i}, v_{1i}\right), \tag{47}$$

where $\lambda_i := \lim_{k\to\infty}\lambda_i^k$, $\sum_{i=1}^{n+1}\lambda_i = 1$. Let $(u_{2i}^k, v_{2i}^k) \to (u_{2i}, v_{2i}) \in \overline{N}_{\operatorname{gph}G}(\overline{x}, \overline{y}_i)$. Then $v_{1i} + v_{2i} = 0$ from (43). Combining (45) and (47), we have

$$\sum_{i=1}^{n+1}\lambda_i\left(\left(u_{1i}, v_{1i}\right) + \left(u_{2i}, v_{2i}\right)\right) = 0. \tag{48}$$

Based on the assumption (iii), we have $(u_{1i}, v_{1i}) = (0, 0)$, $(u_{2i}, v_{2i}) = (0, 0)$. This contradicts the fact that $(u_{1i}, v_{1i}, u_{2i}, v_{2i})$ is of norm 1, and thus $\{z^k\}$ is bounded. Then $(x_{1i}^k, y_{1i}^k), (x_{2i}^k, y_{2i}^k)$ have convergent subsequences, say $(x_{1i}^k, y_{1i}^k) \to (x_{1i}, y_{1i}) \in \overline{\partial}\varphi(\overline{x}, \overline{y}_i), (x_{2i}^k, y_{2i}^k) \to (x_{2i}, y_{2i}) \in \overline{N}_{\operatorname{gph}G}(\overline{x}, \overline{y}_i)$. Thus $u = \sum\lambda_i u_i$ with $u_i = x_{1i} + x_{2i}$. That is to say, Λ is closed.

Next, we justify that $\partial + \partial^{\infty} \subset \Lambda$. Assume that $u_1 \in \partial\mu(\overline{x})$, $u_2 \in \partial^{\infty}\mu(\overline{x})$. Under the semicompactness assumption, invoking [17, Theorem 1.108], one gets that

$$\begin{aligned}
\partial\mu\left(\overline{x}\right) &\subset \left\{u \mid (u, 0) \in \cup_{\overline{y}\in M(\overline{x})}\partial\left(\varphi\left(\overline{x}, \overline{y}\right) + \delta\left(\left(\overline{x}, \overline{y}\right), \operatorname{gph}G\right)\right)\right\},\\
\partial^{\infty}\mu\left(\overline{x}\right) &\subset \left\{u \mid (u, 0) \in \cup_{\overline{y}\in M(\overline{x})}\partial^{\infty}\left(\varphi\left(\overline{x}, \overline{y}\right) + \delta\left(\left(\overline{x}, \overline{y}\right), \operatorname{gph}G\right)\right)\right\}.
\end{aligned} \tag{49}$$

Employing the sum rule from [17, Theorem 3.36] to the two above leads to

$$\begin{aligned}
\left(u_1 + u_2, 0\right) &\in \left[\partial\varphi\left(\overline{x}, \overline{y}\right) + N_{\operatorname{gph}G}\left(\overline{x}, \overline{y}\right)\right]\\
&\quad + \left[\partial^{\infty}\varphi\left(\overline{x}, \overline{y}\right) + N_{\operatorname{gph}G}\left(\overline{x}, \overline{y}\right)\right]\\
&\subset \operatorname{cl}\operatorname{co}\left[\partial\varphi\left(\overline{x}, \overline{y}\right) + \partial^{\infty}\varphi\left(\overline{x}, \overline{y}\right)\right] + \overline{N}_{\operatorname{gph}G}\left(\overline{x}, \overline{y}\right)\\
&= \overline{\partial}\varphi\left(\overline{x}, \overline{y}\right) + \overline{N}_{\operatorname{gph}G}\left(\overline{x}, \overline{y}\right),
\end{aligned} \tag{50}$$

which completes the proof. □

Theorem 11 (convexified normal cone to inequality system [22]). *Consider G defined by inequality system $G(x) = \{y \mid \phi(x, y) \le 0\}$. Let ϕ be Lipschitz, and the qualification (nonsmooth MFCQ) at $(\overline{x}, \overline{y})$ holds:*

$$\begin{aligned}
&\left[\sum_{i\in I(\overline{x},\overline{y})}\lambda_i w_i = 0 \text{ with } w_i \in \partial\phi_i\left(\overline{x}, \overline{y}\right),\ \lambda_i \ge 0\right] \Longrightarrow \lambda_i\\
&\quad = 0 \quad \text{for } i \in I\left(\overline{x}, \overline{y}\right),
\end{aligned} \tag{51}$$

where $I(\overline{x}, \overline{y}) := \{i \mid \phi_i(\overline{x}, \overline{y}) = 0\}$. Then

$$\overline{N}_{\operatorname{gph}G}\left(\overline{x}, \overline{y}\right) \subset \left\{\sum_{i\in I(\overline{x},\overline{y})}\lambda_i\overline{\partial}\phi_i\left(\overline{x}, \overline{y}\right) \mid \lambda_i \ge 0\right\}. \tag{52}$$

As mentioned, GSIP can be relaxed into the following bilevel programming problem:

$$\begin{aligned}
&\min_{(x,y)} \quad f\left(x\right)\\
&\text{s.t.} \quad g\left(x, y\right) \le 0,\ v\left(x, y\right) \le 0,\ \phi\left(x\right) - g\left(x, y\right) \le 0.
\end{aligned} \tag{53}$$

The feasible set of above problem is a subset of the feasible set M of (1). Thus, if $\overline{x}$ solves GSIP and $\phi(\overline{x}) = 0$ and $Y(\overline{x}) \ne \emptyset$, then $\overline{x}$ also solves problem (6). The perturbed version of the above bilevel problem is

$$\begin{aligned}
&\min_{(x,y)} \quad f\left(x\right)\\
&\text{s.t.} \quad g\left(x, y\right) \le 0, v\left(x, y\right) \le 0, \phi\left(x\right) - g\left(x, y\right) \le u.
\end{aligned} \tag{54}$$

Problem (6) is said to be *partially calm* at $(\overline{x}, \overline{y})$ [23] if

$$\begin{aligned}
&\text{there is } \kappa > 0 \text{ and a neighborhood } V \text{ of } \left(\overline{x}, \overline{y}, 0\right) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}\\
&\text{such that for all } \left(x, y, u\right) \in V \text{ feasible for (54) we have}\\
&f\left(x\right) - f\left(\overline{x}\right) \ge -\kappa\left|u\right|.
\end{aligned} \tag{55}$$

Under the partial calmness condition, problem (6) can be transformed into the problem below, for some constant $\kappa > 0$,

$$\begin{aligned}
&\min_{(x,y)} \quad f\left(x\right) + \kappa\left[\phi\left(x\right) - g\left(x, y\right)\right]\\
&\text{s.t.} \quad g\left(x, y\right) \le 0, \quad v\left(x, y\right) \le 0.
\end{aligned} \tag{56}$$

Theorem 12 (necessary conditions of optimality of GSIP). *Let $\overline{x}$ be an optimal solution of GSIP with $\phi(\overline{x}) = 0$, and $Y(\overline{x}) \ne \emptyset$. Let the data functions f, g, and v be Lipschitz and the partial calmness condition (55) hold at $(\overline{x}, \overline{y})$ for some $\overline{y} \in Y_0(\overline{x})$. Assume that the following conditions hold:*

(i) *Qualification (51) holds for $\Omega := \{(x, y) \mid v(x, y) \le 0, g(x, y) \le 0\}$ at $(\overline{x}, \overline{y})$.*

(ii) *Y_0 is ϕ-inner semicompact at $\overline{x}$ and $Y_0(\overline{x}) \ne \emptyset$.*

(iii) *If $(u_i, w_i) \in \overline{\partial}^{\infty}g(\overline{x}, \overline{y}_i)$, $(v_i, -w_i) \in \overline{N}_{\operatorname{gph}Y}(\overline{x}, \overline{y}_i)$, $\overline{y}_i \in Y_0(\overline{x})$, $i \le n + 1$, and $\sum_i u_i + v_i = 0$, then $u_i = v_i = 0$ and $w_i = 0$.*

(iv) *The cones $N_{\operatorname{gph}Y}(\overline{x}, y)$ and $N_{\operatorname{epi}g}(\overline{x}, y, g(\overline{x}, y))$ for all $y \in Y_0(\overline{x})$ are pointed.*

Then there is $\kappa > 0$, $\lambda_i \ge 0$, $\overline{y}_i \in Y_0(\overline{x})$, $i = 1, \ldots, r$ such that $\sum_{i=1}^r\lambda_i = 1$, and

$$\begin{aligned}
0 &\in \left(\overline{\partial}f\left(\overline{x}\right), 0\right) + \kappa\sum_{i=1}^{r}\lambda_i\left[\overline{\partial}g\left(\overline{x}, \overline{y}_i\right) - \overline{\partial}g\left(\overline{x}, \overline{y}\right)\right]\\
&\quad + \sum_{i=1}^{r}\overline{N}_{\operatorname{gph}Y}\left(\overline{x}, \overline{y}_i\right) + \overline{N}_{\Omega}\left(\overline{x}, \overline{y}\right).
\end{aligned} \tag{57}$$

Proof. Let $\Omega = \{(x, y) \mid v(x, y) \le 0,\ g(x, y) \le 0\}$. Under our assumptions, GSIP can be relaxed into problem

(6) and $(\overline{x}, y)$ also solves (6) for all $y \in Y_0(\overline{x})$. Due to the partial calmness of (6), we also have that $(\overline{x}, \overline{y})$ solves (56). Under the qualification assumption, the necessary optimality condition for problem (56) is (see, e.g., [24, Theorem 5.1] or [22, Theorem 6.2])

$$0 \in (\partial f(\overline{x}), 0) + \kappa(\partial\phi(\overline{x}), 0) + \kappa\partial(-g)(\overline{x}, \overline{y}) + N_\Omega(\overline{x}, \overline{y}). \tag{58}$$

If $v \in \partial^+(-\phi)(x)$, then $-v \in \overline{\partial}\phi(x)$. Indeed, by definition,

$$\begin{aligned}\partial^+(-\phi)(x) &= \{v \mid (-v, 1) \in N_{\mathrm{hypo}(-\phi)}(x, -\phi(x))\} \\ &= \{v \mid (-v, -1) \in N_{\mathrm{epi}\,\phi}(x, \phi(x))\} \\ &\subset \{v \mid (-v, -1) \in \overline{N}_{\mathrm{epi}\,\phi}(x, \phi(x))\}.\end{aligned} \tag{59}$$

Thus, $\partial\phi(x) = -\partial^+(-\phi)(x) \subset \overline{\partial}\phi(x) = -\overline{\partial}(-\phi)(x)$, and from (58),

$$\begin{aligned}0 \in{}& (\overline{\partial} f(\overline{x}), 0) - \kappa(\overline{\partial}(-\phi)(\overline{x}), 0) + \kappa\overline{\partial}(-g)(\overline{x}, \overline{y}) \\ &+ \overline{N}_\Omega(\overline{x}, \overline{y}).\end{aligned} \tag{60}$$

Applying Proposition 10 to $-\phi$, there are $\lambda_i \geq 0$, $\overline{y}_i \in Y_0(\overline{x})$, $i = 1, \ldots, r$ such that $\sum_{i=1}^r \lambda_i = 1$ and

$$(\overline{\partial}(-\phi)(\overline{x}), 0) \subset \sum_{i=1}^r \lambda_i \left[\overline{\partial}(-g)(\overline{x}, \overline{y}_i) + \overline{N}_{\mathrm{gph}\,Y}(\overline{x}, \overline{y}_i)\right]. \tag{61}$$

So, noting that $\overline{\partial}(-g) = -\overline{\partial}g$,

$$\begin{aligned}0 \in{}& (\overline{\partial} f(\overline{x}), 0) + \kappa\sum_{i=1}^r \lambda_i \overline{\partial} g(\overline{x}, \overline{y}_i) - \kappa\overline{\partial} g(\overline{x}, \overline{y}) \\ &+ \kappa\sum_{i=1}^r \lambda_i \overline{N}_{\mathrm{gph}\,Y}(\overline{x}, \overline{y}_i) + \overline{N}_\Omega(\overline{x}, \overline{y}) \\ \subset{}& (\overline{\partial} f(\overline{x}), 0) + \kappa\sum_{i=1}^r \lambda_i \left[\overline{\partial} g(\overline{x}, \overline{y}_i) - \overline{\partial} g(\overline{x}, \overline{y})\right] \\ &+ \sum_{i=1}^r \overline{N}_{\mathrm{gphY}}(\overline{x}, \overline{y}_i) + \overline{N}_\Omega(\overline{x}, \overline{y}).\end{aligned} \tag{62}$$

This completes the proof. □

Conflict of Interests

The authors declare that there is no conflict of interests.

Acknowledgments

This work is supported by the National Science Foundation of China (11001289) and the Key Project of Chinese Ministry of Education (211151).

References

[1] R. Hettich and G. Still, "Second order optimality conditions for generalized semi-infinite programming problems," *Optimization*, vol. 34, no. 3, pp. 195–211, 1995.

[2] H. Th. Jongen, J.-J. Rückmann, and O. Stein, "Generalized semi-infinite optimization: a first order optimality condition and examples," *Mathematical Programming A*, vol. 83, no. 1, pp. 145–158, 1998.

[3] J. J. Rückmann and A. Shapiro, "First-order optimality conditions in generalized semi-infinite programming," *Journal of Optimization Theory and Applications*, vol. 101, no. 3, pp. 677–691, 1999.

[4] O. Stein, "First-order optimality conditions for degenerate index sets in generalized semi-infinite optimization," *Mathematics of Operations Research*, vol. 26, no. 3, pp. 565–582, 2001.

[5] G. Still, "Generalized semi-infinite programming: theory and methods," *European Journal of Operational Research*, vol. 119, no. 2, pp. 301–313, 1999.

[6] O. Stein, *Bi-Level Strategies in Semi-Infinite Programming*, vol. 71 of *Nonconvex Optimization and its Applications*, Kluwer Academic, Boston, Mass, USA, 2003.

[7] H. Günzel, H. Th. Jongen, and O. Stein, "On the closure of the feasible set in generalized semi-infinite programming," *Central European Journal of Operations Research*, vol. 15, no. 3, pp. 271–280, 2007.

[8] H. Günzel, H. Th. Jongen, and O. Stein, "Generalized semi-infinite programming: the symmetric reduction ansatz," *Optimization Letters*, vol. 2, no. 3, pp. 415–424, 2008.

[9] F. Guerra-Vázquez, H. Th. Jongen, and V. Shikhman, "General semi-infinite programming: symmetric Mangasarian-Fromovitz constraint qualification and the closure of the feasible set," *SIAM Journal on Optimization*, vol. 20, no. 5, pp. 2487–2503, 2010.

[10] H. Th. Jongen and V. Shikhman, "Generalized semi-infinite programming: the nonsmooth symmetric reduction ansatz," *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 193–211, 2011.

[11] O. Stein and G. Still, "On generalized semi-infinite optimization and bilevel optimization," *European Journal of Operational Research*, vol. 142, no. 3, pp. 444–462, 2002.

[12] S. Dempe and A. B. Zemkoho, "The generalized Mangasarian-Fromowitz constraint qualification and optimality conditions for bilevel programs," *Journal of Optimization Theory and Applications*, vol. 148, no. 1, pp. 46–68, 2011.

[13] K. Shimizu, Y. Ishizuka, and J. F. Bard, *Nondifferentiable and Two-level Mathematical Programming*, Kluwer Academic, Boston, Mass, USA, 1997.

[14] N. Kanzi and S. Nobakhtian, "Necessary optimality conditions for nonsmooth generalized semi-infinite programming problems," *European Journal of Operational Research*, vol. 205, no. 2, pp. 253–261, 2010.

[15] N. Kanzi, "Lagrange multiplier rules for non-differentiable DC generalized semi-infinite programming problems," *Journal of Global Optimization*, vol. 56, no. 2, pp. 417–430, 2013.

[16] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, vol. 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Springer, Berlin, Germany, 1998.

[17] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation. I: Basic Theory*, vol. 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Springer, Berlin, Germany, 2006.

[18] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons, New York, NY, USA, 1983.

[19] R. T. Rockafellar, *The Theory of Subgradients and Its Applications to Problems of Optimization. Convex and Nonconvex*, vol. 1 of *R & E*, Heldermann, Berlin, Germany, 1981.

[20] B. S. Mordukhovich, N. M. Nam, and N. D. Yen, "Subgradients of marginal functions in parametric mathematical programming," *Mathematical Programming B*, vol. 116, no. 1-2, pp. 369–396, 2009.

[21] B. S. Mordukhovich and N. M. Nam, "Variational stability and marginal functions via generalized differentiation," *Mathematics of Operations Research*, vol. 30, no. 4, pp. 800–816, 2005.

[22] B. S. Mordukhovich, N. M. Nam, and H. M. Phan, "Variational analysis of marginal functions with applications to bilevel programming," *Journal of Optimization Theory and Applications*, vol. 152, no. 3, pp. 557–586, 2012.

[23] J. J. Ye and D. L. Zhu, "Optimality conditions for bilevel programming problems," *Optimization*, vol. 33, no. 1, pp. 9–27, 1995.

[24] S. Dempe, J. Dutta, and B. S. Mordukhovich, "New necessary optimality conditions in optimistic bilevel programming," *Optimization*, vol. 56, no. 5-6, pp. 577–604, 2007.