Nonlinear Analysis 75 (2012) 1655–1678

Contents lists available at SciVerse ScienceDirect

Nonlinear Analysis

journal homepage: www.elsevier.com/locate/na

Necessary optimality conditions for bilevel minimization problems

Alexander J. Zaslavski

Department of Mathematics, The Technion-Israel Institute of Technology, 32000 Haifa, Israel

ARTICLE INFO

Communicated by Ravi Agarwal

MSC:
90C29
49J52
49J53

Keywords:
Bilevel minimization problem
Generalized differentiation
Implicit function theorem
Necessary optimality conditions

ABSTRACT

In this paper we study bilevel minimization problems. Using the implicit function theorem, variational analysis and exact penalty results we establish necessary optimality conditions for these problems.

1. Introduction

The study of bilevel minimization problems has recently been a rapidly growing area of research. See, for example, [1–6] and the references mentioned therein. The goal of this paper is to obtain necessary optimality conditions for these problems. The main difficulty is that a bilevel minimization problem is, in general, nonconvex and nonsmooth even if all functions which appear in its description are convex and smooth. Necessary optimality conditions for bilevel minimization problems were obtained in [2]. They are based on a partial calmness property. Note that there are simple examples of bilevel minimization problems for which the partial calmness property of [2] does not hold. In the present paper we obtain necessary optimality conditions of a different kind which are based on the implicit function theorem, variational analysis and exact penalty results. We study bilevel minimization problems with unconstrained lower level problems. Note that in the literature [1–6], bilevel minimization problems are also studied such that their lower level problems possess constraints. Here for simplicity we consider only the case with unconstrained lower level problems. It should be mentioned that in our results we use the assumption that a solution of a bilevel problem is not a critical point of a constraint function. This allow us to apply exact penalty results of [7,8].

The paper has the following structure. Section 2 contains preliminaries while Section 3 contains auxiliary results. In Section 4 we obtain necessary optimality conditions for the problem

$$g(x) + \inf\{F(x, y) : y \in R^m \text{ and } f(x, y) = \phi(x)\} \to \min, \quad x \in R^n$$

where $g : R^n \to R^1 \cup \{\infty\}$, $F : R^n \times R^m \to R^1$, $f : R^n \times R^m \to R^1$ and

$$\phi(x) = \inf\{f(x, y) : y \in R^m\}, \quad x \in R^n.$$

We obtain these necessary optimality conditions using the notion of Mordukhovich basic subdifferential introduced in [9] (see also [10, page 82]).

E-mail address: ajzasl@tx.technion.ac.il.

0362-546X/$ – see front matter
doi:10.1016/j.na.2011.03.049

For $x \in R^n$ set

$$\mathrm{Argmin}(f, x) = \{y \in R^m : f(x, y) = \phi(x)\}.$$

In Section 6 using the notion of Mordukhovich basic subdifferential [9,10] and exact penalty results of [7] which are discussed in Section 5 we obtain necessary optimality conditions for the problems

$$\text{Minimize } \inf\{F(x, y) : y \in \mathrm{Argmin}(f, x)\} \quad \text{subject to } g(x) \le 0$$

and

$$\text{Minimize } \inf\{F(x, y) : y \in \mathrm{Argmin}(f, x)\} \quad \text{subject to } g(x) = 0.$$

In Section 8 we continue to study these two bilevel minimization problems and obtain necessary optimality conditions using the notion of Clarke's generalized gradient [11] and exact penalty results of [8] which are discussed in Section 7.

In Sections 9 and 10 we consider classes of bilevel minimization problems which are identified with corresponding complete metric spaces of functions. We show that for most problems the necessary optimality conditions established in Sections 4, 6 and 8 hold. In Section 11 we discuss the partial calmness property of [2]. We present two examples of problems for which necessary optimality conditions obtained in the present paper hold and for which the partial calmness property does not hold. We also prove Theorem 11.1 which shows that many problems do not possess the partial calmness property.

2. Preliminaries

Let $k \ge 1$ be an integer and let us consider the space R^k equipped with the inner product $\langle \cdot, \cdot \rangle$ which induces the Euclidean norm $\|\cdot\|$. For each $z \in R^k$ and each $r > 0$ set

$$B(z, r) = \{v \in R^k : \|v - z\| \le r\},$$
$$B^0(z, r) = \{v \in R^k : \|v - z\| < r\}.$$

For each $A \subset R^k$ and each $x \in R^k$ set

$$d(x, A) = \inf\{\|x - y\| : y \in A\}.$$

For each pair of nonempty sets $A, B \subset R^k$ put

$$H(A, B) = \max\{\sup_{x \in A} d(x, B), \sup_{z \in B} d(z, A)\}.$$

For each function $h : R^k \to R^1 \cup \{\infty\}$ and each nonempty set $A \subset R^k$ set

$$\inf(h) = \inf\{h(z) : z \in R^k\}, \qquad \inf(h; A) = \inf\{h(z) : z \in A\}.$$

Assume that $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$ and $(Z, \|\cdot\|)$ are Banach spaces, W is an open set in $X \times Y$ and that a mapping $\psi : W \to Z$ belongs to $C^1(W, Z)$. We denote by ψ_x and ψ_y the partial derivatives of ψ with respect to $x \in X$ and $y \in Y$ if they exist. We denote by ψ_{xx}, ψ_{xy} and ψ_{yy} the partial derivatives of ψ of the second order if they exist. For each mapping $\phi \in C^1(V, Z)$ where V is an open subset of X, denote by $\phi'(x)$ the Frechet derivative of ϕ at a point $x \in V$.

For each $x \in X$, each $y \in Y$ and each $r > 0$ set

$$B_X(x, r) = \{u \in X : \|u - x\| \le r\},$$
$$B_X^0(x, r) = \{u \in X : \|u - x\| < r\},$$
$$B_Y(y, r) = \{u \in Y : \|u - y\| \le r\},$$
$$B^0(y, r) = \{u \in Y : \|u - y\| < r\}.$$

In this paper we use the following implicit function theorem [12].

Theorem 2.1. *Let $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$, $(Z, \|\cdot\|)$ be Banach spaces, $(x_0, y_0) \in X \times Y$, W be an open neighborhood of (x_0, y_0) in $X \times Y$ and let a mapping $\psi : W \to Z$ belongs to $C^1(W, Z)$. Assume that $\psi(x_0, y_0) = 0$ and that there exists an inverse operator $[\psi_y(x_0, y_0)]^{-1} : Z \to Y$ which is linear and continuous. Then there exist $\epsilon > 0$, $\delta > 0$ and a mapping $\phi : B_X^0(x_0, \delta) \to Y$ such that:*

$$\phi \in C^1(B_X^0(x_0, \delta), Y);$$
$$\phi(x_0) = y_0;$$
$$\psi(x, \phi(x)) = 0 \quad \text{for all } x \in B_X^0(x_0, \delta);$$
$$\|\phi(x) - y_0\| < \epsilon \quad \text{for each } x \in B_X^0(x_0, \delta);$$

for each $x \in B_X^0(x_0, \delta)$ and each $y \in B_Y^0(y_0, \epsilon)$ the equality $\psi(x, y) = 0$ holds if and only if $y = \phi(x)$;

$$\phi'(x) = -[\psi_y(x, \phi(x))]^{-1} \psi_x(x, \phi(x)) \quad \text{for all } x \in B_X^0(x_0, \delta).$$