MULTISTAGE HIERARCHICAL OPTIMIZATION PROBLEMS WITH MULTI-CRITERION OBJECTIVES

ROXIN ZHANG, BAO TRUONG AND QINGHONG ZHANG

Department of Mathematics and Computer Science
Northern Michigan University, Marquette, MI 49855, USA

(Communicated by Song Wang)

Abstract. A hierarchical optimization (or bilevel programming) problem consists of a decision maker called the leader who is interested in optimizing an objective function that involves with the decisions from another decision maker called the follower whose decisions are based in part on the policies made by the leader. However, if the planning horizon expands into an extended period of time, it may be unrealistic for either players to commit to the original decisions so there is a desire to break the problem into stages and the leader may wish to reevaluate the follower’s response at each stage. In this article, we propose a multistage hierarchical optimization problem with the leader’s objective consisting of multiple criteria and study the optimality conditions of such problems using an extremal principle of Mordukhovich.

1. Introduction. In a hierarchical optimization problem, two levels of decision-making are involved in such a way that the upper-level decision maker, called the leader, makes the decision to optimize its objectives. However the objectives and the constraints of the leader are subject to change depending on what the other decision maker, the follower, does. The follower’s actions in making decisions are passive and responsive in order to optimize its own objectives based on whatever the leader’s choices of decisions. If the planning horizon includes an extended period of time, it is desirable to break the time into multiple stages. There are also applications in that the leader’s decisions are involved in a design process that evolves over multiple stages and the leader wishes to reevaluate the follower’s reactions to improve its own decisions over a number of time periods. At any stage, the leader is assumed to know the behavior of the follower once a decision is announced. On the other hand, the follower only has the knowledge of the past decisions of the both levels as well as the current leader’s policy, and thus makes decisions in response to the leader’s decisions one stage after another.

The type of hierarchical optimization problems that we are interested in consists of $n$ stages with the upper-level objectives of multicriteria. The decision variables of both decision makers, the leader and the follower, are defined over $n$ stages. At the beginning of stage $i$ for $i \in \{1, 2, \ldots, n\}$, the leader, at the upper-level, announces its decision $x_i$ and the follower, at the lower-level, based on this decision from the
leader and the previous decisions \(x_{i-1}\) and \(y_{i-1}\) from both levels, takes a responsive action with a decision \(y_i\) by optimizing its own objective function of that stage. The leader’s decisions over the entire period of \(n\) stages are made as such to optimize its overall objective with multiple criteria. At the same time, the leader is assumed to be able to establish an agreement with the follower in that if two decisions are indifferent to the follower’s own objective (in a multi-valued solution set scenario) at any stage, a follower’s decision should be chosen in favor of the leader’s benefit. In fact such an agreement is actually mutually beneficial to both of the decision makers because the leader has the advantage of foreseeing the entire time horizon and thus is able to advise the follower to avoid certain decisions at some of the stages that may not lead to any optimal solutions at a later stage.

The hierarchical multistage optimization models described above can often arise from industrial applications. In references \([5]\) and \([4]\), for instance, an energy absorbing process through delamination was considered and a hierarchical optimization model with the evolutionary equilibrium constraints was studied. In their model, the amount of energy absorption is maximized at the upper level and at the lower level, as the constraints, there is a sequence of parametric sub-optimization problems resulting from the time-discretization of the problem of minimizing the elastic stored energy and the dissipation potential. In their treatment of the time-discretization, each sub-problem depends on the solutions to the previous sub-problem as well as the upper level design variable, exhibiting an \(n\)-stage parametric optimization problem.

Specifically, let \(X_i\) and \(Y_i\) be finite dimensional spaces for \(i = 1,\ldots, n\), \(X := X_1 \times \cdots \times X_n\) and \(Y := Y_1 \times \cdots \times Y_n\). The upper- and lower-level decision variables at stage \(i\), for \(i = 1,\ldots, n\), are denoted by \(x_i \in X_i\) and \(y_i \in Y_i\) respectively. We are given a set of real-valued lower-level objective functions \(g_i\) defined on the space \(X_{i-1} \times Y_{i-1} \times X_i\), and a set of lower-level constraint set-valued mappings \(\Omega_i\) defined on \(X_{i-1} \times Y_{i-1} \times X_i\) to \(Y_i\) for \(i = 1,\ldots, n\). We denote by \(\mathbb{R}\) and \(\mathbb{R}_+\) the sets of all real numbers and all nonnegative real numbers respectively, by \(\mathbb{N}\) the set of all natural numbers.

Starting at the first stage the follower minimizes its first objective function \(g_1(x_1, y_1)\) subject to \(y_1 \in \Omega_1(x_1)\) for a given upper-level decision \(x_1 \in X_1\). At stage \(i\), for \(i = 2,3,\ldots, n\), the follower takes on a given set of decision variables \(x_{i-1} \in X_{i-1}, y_{i-1} \in Y_{i-1}\) from the previous stage and the current leader’s decision \(x_i \in X_i\), and wishes to solve the problem

\[
\begin{align*}
(Q_i) & \quad \text{minimize} \quad g_i(x_{i-1}, y_{i-1}, x_i, y_i) \\
\text{subject to} \quad & y_i \in \Omega_i(x_{i-1}, y_{i-1}, x_i).
\end{align*}
\]

The reactions of the follower in optimizing its own objective at each stage is known to the leader and affect the leader’s objectives. To reduce the redundancies in notations, throughout this paper, we will regard \(x_0\) and \(y_0\) as dummy variables and treat \(g_1(x_1, y_1)\) as a function of \((x_0, y_0, x_1, y_1)\), namely \(g_1(x_0, y_0, x_1, y_1) := g_1(x_1, y_1)\).

Suppose we denote the optimal solution set of the follower from the first stage by \(S_1(x_0, y_0, x_1)\), and the solution set from the second stage by \(S_2(x_1, y_1, x_2)\), and so on with the solution set from the last stage by \(S_n(x_{n-1}, y_{n-1}, x_n)\). Then for each given upper-level decision vector \(x := (x_1, x_2, \ldots, x_n)\) the set of lower-level decisions for all stages \(R(x_1, x_2, \ldots, x_n) \in Y\), called the rational response mapping, is defined to be the set of vectors \((y_1, y_2, \ldots, y_n)\) such that

\[
\bar{y}_1 \in S_1(x_0, y_0, x_1), \quad \bar{y}_2 \in S_2(x_1, \bar{y}_1, x_2), \ldots, \quad \bar{y}_n \in S_n(x_{n-1}, \bar{y}_{n-1}, x_n).
\]
The leader, at the upper-level, anticipates the rational responses from the lower-
level to its decisions at each stage, and tries to solve a multiobjective optimization
problem. To clearly state the upper-level problem, we will need the following defi-
nition [6, 7].

**Definition 1.1.** Given a mapping \( \varphi : Z \to H \) between two finite dimensional
vector spaces, an ordering set \( \Theta \subset H \) with \( 0 \in \Theta \), and a constraint \( \Lambda \subset Z \). We say a point \( \bar{z} \in \Lambda \) is locally \((\varphi, \Theta)\)-optimal subject to the constraint \( z \in \Lambda \) if there is a neighborhood \( U \) of \( \bar{z} \) and a sequence of \( a_k \in H \) with \( \|a_k\| \to 0 \) as \( k \to \infty \) such that

\[
\varphi(z) - \varphi(\bar{z}) \notin \Theta - a_k \quad \forall z \in \Lambda \cap U, \forall k,
\]

in this case, we also say that \( \bar{z} \) is a \( \Theta \)-minimizer of \( \varphi \) subject to \( z \in \Lambda \).

In the sequel we will use the term “\( \Theta \)-minimize \( \varphi \)” to represent the multiobjective
problem of finding a \( \Theta \)-minimizer of the mapping \( \varphi \).

The notion of the \( \Theta \)-optimality presented in the above definition is a unified gen-
eralization to a variety of vector optimality concepts in the literature. In particular,
the scalar optimality of minimizing a real-valued function becomes a special case
with \( \Theta := \{ \tau \in \mathbb{R} | \tau \leq 0 \} \). In general, by choosing different types of the ordering
set \( \Theta \), the \((\varphi, \Theta)\)-optimality reverts to various specific notions of vector optimality
or efficiency.

For example, for a given convex cone \( K \subset H \), and the problem of minimizing
\( \varphi(z) \) over \( z \in \Lambda \) with \( \bar{z} \) and \( U \) as defined in Definition 1.1. We have

a.) If \( \Theta = K \), the point \( \bar{z} \) is \((\varphi, \Theta)\)-optimal if and only if \( \bar{z} \) is a strong Pareto
optimal solution in the sense that there is no \( z \in \Lambda \cap U \) and \( z \neq \bar{z} \) such that
\( \varphi(z) - \varphi(\bar{z}) \in K \);

b.) If \( \text{int } K \neq \emptyset \) and \( \Theta = \text{int } K \cup \{ 0 \} \), the point \( \bar{z} \) is \((\varphi, \Theta)\)-optimal if and only
if \( \bar{z} \) is a weak Pareto optimal solution in the sense that there is no \( z \in \Lambda \cap U \) such that
\( \varphi(z) - \varphi(\bar{z}) \in \text{int } K \);

c.) If \( \text{ri } K \neq \emptyset \) and \( \Theta = \text{ri } K \cup \{ 0 \} \), the point \( \bar{z} \) is \((\varphi, \Theta)\)-optimal if and only if
\( \bar{z} \) is an \( \Theta \)-type efficient solution in the sense that there is no \( z \in \Lambda \cap U \) such that
\( \varphi(z) - \varphi(\bar{z}) \in \text{ri } K \).

Now, for \( X, Y \) and \( H \) as defined before, consider the leader’s vector-valued
objective function \( f : X \times Y \to H \) and a constraint set \( D \subset X \times Y \). The leader
wishes to find a \( \Theta \)-minimizer

\[
(\bar{x}, \bar{y}) := (\bar{x}_1, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_n) \in X \times Y
\]
of the function \( f \) subject to the constraints that

\[
(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_n) \in D \quad \text{and} \quad (y_1, \ldots, y_n) \in R(x)
\]
where \( x_i \in X_i, y_i \in Y_i, x \in X, y \in Y \), and \( R(x) \) is the follower’s rational response
mapping.

To proceed with our optimality conditions, we assume in the sequel that for
each given \((x_{i-1}, y_{i-1}, x_i)\) the function \( g_i(x_{i-1}, y_{i-1}, x_i) \) is convex and the set
\( \Omega_i(x_{i-1}, y_{i-1}, x_i) \) is also convex for \( i = 1, 2, \ldots, n \).

Under these assumptions, it is clear that the rational response mapping \( R(x) \) can
be described as such that \((y_1, y_2, \ldots, y_n) \in R(x_1, x_2, \ldots, x_n) \) if and only if

\[
0 \in \partial_{y_i} g_i(x_{i-1}, y_{i-1}, x_i, y_i) + N(y_i | \Omega_i(x_{i-1}, y_{i-1}, x_i)) \quad \forall i = 1, \ldots, n,
\]
where \( \partial_{y_i} g_i \) is the subdifferential of \( g_i \) with respect to \( y_i \) and \( N(y_i | \Omega_i) \) is the normal
cone to the set \( \Omega_i \) at the point \( y_i \) in the sense of convex analysis.
By using the notation that $G_i(x_{i-1}, y_{i-1}, x_i, y_i) := \partial_{y_i} g_i(x_{i-1}, y_{i-1}, x_i, y_i)$ and $Q_i(x_{i-1}, y_{i-1}, x_i, y_i) := N(y_i | \Omega_i(x_{i-1}, y_{i-1}, x_i))$, we can restate the leader’s multiobjective optimization problem (P) as

$$\Theta\text{-minimize } f(x_1, \ldots, x_n, y_1, \ldots, y_n)$$

subject to $0 \in G_i(x_{i-1}, y_{i-1}, x_i, y_i) + Q_i(x_{i-1}, y_{i-1}, x_i, y_i)$,

for $i = 1, \ldots, n$

$(x_1, \ldots, x_n, y_1, \ldots, y_n) \in D$.

Note that if we have only one stage (namely $n = 1$), then the above problem reduces to a bilevel programming problem studied in [1]. The basic structure of the problem (4) is essentially an extension of the proposed problem in [11] to the multiobjective optimization at the upper-level, and here we are able to obtain much stronger results.

Note that in the problem (P), if we remove the multiobjective component of the upper level problem and replace the constraints by the system

$$0 \in \tilde{G}_i(x_i, y_{i-1}, y_i) + \tilde{Q}_i(x_i, y_{i-1}, y_i)$$

where $x := (x_1, \ldots, x_n)$ and $\tilde{G}_i$ and $\tilde{Q}_i$ are defined similarly to $G_i$ and $Q_i$, we arrive at the application model of the delamination process treated in [5] and [4], and the system (5) is referred to as an evolutionary equilibrium constraints.

2. Preliminaries. In this section, we present some of the necessary tools in modern variational analysis that are needed to develop our optimality conditions.

Let $Z$ and $H$ be two finite dimensional vector spaces and consider a set valued mapping $F : Z \Rightarrow H$ and $\bar{z} \in Z$. We denote by $gphF := \{(z, v) | v \in F(z)\}$ the graph of the mapping $F$. The Painlevé-Kuratowski outer limit of $F$ as $z \to \bar{z}$ is defined by

$$\limsup_{z \to \bar{z}} F(z) := \{v \in H | \exists z_k \to \bar{z} \text{ and } v_k \to v \text{ with } v_k \in F(z_k)\}.$$ 

For $C \subset Z$ and $\bar{z} \in C$, the tangent (or contingent) cone $T(\bar{z} | C)$ is defined by

$$T(\bar{z} | C) := \limsup_{\tau \to 0} C - \bar{z} \over \tau.$$ 

The normal (or limiting normal) cone $N(\bar{z} | C)$ is defined by

$$N(\bar{z} | C) := \limsup_{z \to \bar{z}} T^0(z | C)$$

where $T^0(z | C) := \{v | \langle v, w \rangle \leq 0, \forall w \in T(z | C)\}$. Note that in finite dimensional spaces the definition of the normal (limiting) cone defined above is equivalent to the Mordukhovich normal cone defined in [8].

Given an extended real-valued function $\psi : Z \to R \cup \{\infty\}$ and a point $\bar{z} \in Z$ with $|\psi(\bar{z})| < \infty$, the subdifferential of $\psi$ at the point $\bar{z}$ is defined by

$$\partial \psi(\bar{z}) := \{z^* | (z^*, -1) \in N((\bar{z}, \psi(\bar{z})) | epi \psi)\},$$

where the set $epi \psi := \{(z, \tau) | \psi(z) \geq \tau\}$ is the epigraph of $\psi$.

The coderivative of the set-valued mapping $F$ at $(\bar{z}, \bar{v}) \in gphF$ is defined by

$$D^*F(\bar{z}, \bar{v})(v^*) := \{z^* \in Z | (z^*, -v^*) \in N((\bar{z}, \bar{v}) | gphF)\}.$$
For a single-valued mapping \( \varphi : Z \to H \), we shall write \( D^*\varphi(\bar{z})(v^*) \) in place of \( D^*\varphi(\bar{z}, \varphi(\bar{z}))(v^*) \). If, in addition, \( \varphi \) is continuously differentiable at \( \bar{z} \), we will have
\[
D^*\varphi(\bar{z})(v^*) = \nabla \varphi(\bar{z})^\top v^* \quad \forall v^* \in H
\]
where \( \nabla \varphi(\bar{z}) \) and \( \nabla \varphi(\bar{z})^\top \) are the Jacobian matrix of \( \varphi \) at \( \bar{z} \) and its transpose respectively. When \( \varphi \) is locally Lipschitz continuous around the point \( \bar{z} \), then
\[
D^*\varphi(\bar{z})(v^*) = \partial(v^*, \varphi)(\bar{z}), \quad \langle v^*, \varphi(\bar{z}) \rangle = \langle v^*, \varphi(\bar{z}) \rangle.
\]
We refer to [6, 7] and [9] for complete discussions on these concepts.

**Definition 2.1** (extremal system of two sets). Let \( \Omega_1 \) and \( \Omega_2 \) be nonempty subsets of the space \( Z \), and let \( \bar{z} \in \Omega_1 \cap \Omega_2 \). We say \( \bar{z} \) is a locally extremal point of the set system \( \{\Omega_1, \Omega_2\} \) if there is a sequence \( \{a_k\} \subset Z \) and a neighborhood \( U \) of \( \bar{z} \) such that \( \|a_k\| \to 0 \) as \( k \to \infty \) and
\[
\Omega_1 \cap (\Omega_2 + a_k) \cap U = \emptyset \quad \text{for all large } k.
\]
In this case, we call \( \{\Omega_1, \Omega_2, \bar{z}\} \) an extremal system in \( Z \).

The following version of the Mordukhovich extremal principle will be the fundamental vehicle in deriving our optimality condition results.

**Lemma 2.2** (exact extremal principle of two sets). Let \( \{\Omega_1, \Omega_2, \bar{z}\} \) be an extremal system in \( Z \) as defined in Definition 2.1. Then there exists \( z^* \neq 0 \) such that
\[
z^* \in N(\bar{z} | \Omega_1) \text{ and } -z^* \in N(\bar{z} | \Omega_2).
\]

**Proof.** [6, Theorem 2.8].

Given an ordering set \( \Theta \) and a set \( \Xi \subset Z \), we will present a lemma that provides the necessary conditions for multiobjective optimization problems with only the geometric constraint:

\[
\Theta-\text{minimize } \varphi(z) \text{ subject to } z \in \Xi \subset Z,
\]
where \( \varphi : Z \to H \) is a vector-valued function. The result below is actually a specification of Theorem 5.59 of [6, 7] for problems in infinite-dimensional Asplund spaces. Nevertheless, for completeness and the reader’s convenience we present a simplified proof of this lemma, which follows the approach in [1] based on the extremal principle for sets instead of that for multifunctions.

**Lemma 2.3** (necessary conditions in vector-valued optimization with geometric constraints). Let \( \bar{z} \) be a local minimizer to the constrained vector-valued optimization problem (8), where the ordering cone \( \Theta \) is locally closed at the origin. Assume that \( \varphi \) is locally Lipschitz continuous at \( \bar{z} \) and \( \Xi \) is locally closed around \( \bar{z} \). Then there is \( \alpha^* \in N(0 | \Theta) \) with \( \alpha^* \neq 0 \) such that
\[
0 \in D^*\varphi(\bar{z})(\alpha^*) + N(\bar{z} | \Xi) = \partial(\alpha^*, \varphi)(\bar{z}) + N(\bar{z} | \Xi). \tag{9}
\]

**Proof.** We proceed with creating the extremal system of sets generated by the local \((\varphi, \Theta)\)-minimizer \( \bar{z} \) to the problem (8) and then apply the extremal principle, Lemma 2.2, to the system. We first define the sets
\[
\Lambda_1 := \text{gph} \varphi, \quad \Lambda_2 := \Xi \times (\bar{\alpha} + \Theta) \tag{10}
\]
in the product space \( Z \times H \), and then show that \((\bar{z}, \bar{\alpha})\) with \( \bar{\alpha} = \varphi(\bar{z}) \) is a local extremal point of the system \( \{\Lambda_1, \Lambda_2\} \). In fact, it is obvious that \((\bar{z}, \bar{\alpha}) \in \Lambda_1 \cap \Lambda_2 \) and the sets \( \Lambda_1 \) and \( \Lambda_2 \) are locally closed around this point. To justify the condition
(6) for the set system (10), we find a neighborhood $U$ of $\bar{z}$ and a sequence $\{\alpha_k\} \subset H$ with $\|\alpha_k\| \to 0$ by the local $(\varphi, \Theta)$-minimality of $(\bar{z}, \bar{\alpha})$ to (8) such that
\[ \varphi(z) - \varphi(\bar{z}) \notin \Theta - \alpha_k \quad \forall z \in \Xi \cap U \quad \forall k \in \mathbb{N}. \]
This condition can be equivalently written as
\[ \text{gph}\varphi \cap (\Xi \times (\bar{\alpha} + \Theta) + (0, -\alpha_k)) \cap (U \times H) = \emptyset \quad \forall k \in \mathbb{N}. \]
Therefore the local extremality of $\{\varLambda_1, \varLambda_2\}$ at $(\bar{z}, \bar{\alpha})$ is justified with the neighborhood $U \times H$ and the sequence $\{(0, -\alpha_k)\}_{k=0}^\infty$.

By employing now the extremal principle, Lemma 2.2, to the set system (10), we can find $(z^*, \alpha^*) \neq 0$ satisfying
\[ (z^*, \alpha^*) \in N((\bar{z}, \bar{\alpha}) | \Xi \times (\bar{\alpha} + \Theta)) \cap (-N((\bar{z}, \bar{\alpha}) | \text{gph}\varphi)). \]
Taking into account the definition of coderivative and the normals to Cartesian products, we have
\[ z^* \in N(\bar{z} | \Xi), \quad -z^* \in D^*\varphi(\bar{z})(\alpha^*) \quad \text{and} \quad \alpha^* \in N(\bar{\alpha} | (\bar{\alpha} + \Theta)) = N(0 | \Theta) \quad (11) \]
that clearly proves the necessary condition (9) provided that $\alpha^* \neq 0$. This last nontriviality is true for otherwise if $\alpha^* = 0$, it would imply that $z^* = 0$ because $D^*\varphi(\bar{z})(0) = \{0\}$ due to the locally Lipschitz continuity of $\varphi$ around $\bar{z}$. We arrive at the contradiction that $(z^*, \alpha^*) = 0$, and therefore complete the proof of the lemma. 

3. Optimality conditions. In this section, we first develop necessary conditions for our general multistage hierarchical optimization problem $(P)$, and then apply the results to several problems with certain specific properties.

At the upper-level, the objective function $f : X \times Y \to H$ is a vector-valued mapping involving the decision variables $x = (x_1, \ldots, x_n) \in X$ and $y = (y_1, \ldots, y_n) \in Y$ of both levels, and we wish to find an $(f, \Theta)$-optimal solution for a given ordering set $\Theta$.

On the other hand, at stage $i$ for $i = 1, \ldots, n$, for a given set of vectors $x_{i-1}, y_{i-1}$ and $x_i$, the lower-level objective function and constraint set are $g_i(x_{i-1}, y_{i-1}, x_i, y_i)$ and $\varOmega_i(x_{i-1}, y_{i-1}, x_i)$ respectively with the decision variable $y_i \in Y_i$. We consider two set-valued mappings $G_i$ and $Q_i : X_{i-1} \times Y_{i-1} \times X_i \times Y_i \Rightarrow Y_i$ defined by
\begin{align*}
    G_i(x_{i-1}, y_{i-1}, x_i, y_i) &= \partial_{y_i} g_i(x_{i-1}, y_{i-1}, x_i, y_i), \\
    Q_i(x_{i-1}, y_{i-1}, x_i, y_i) &= N(y_i | \varOmega_i(x_{i-1}, y_{i-1}, x_i)),
\end{align*}
(12)
and the subset $D \subset X \times Y$.

Recall that our underline multiobjective hierarchical optimization problem takes the form
\begin{align*}
    &\Theta\text{-minimize} \quad f(x, y) \\
    \text{subject to} \quad &0 \in G_i(x_{i-1}, y_{i-1}, x_i, y_i) + Q_i(x_{i-1}, y_{i-1}, x_i, y_i), \\
    &\text{for } i = 1, \ldots, n \\
    &\langle x, y \rangle \in D.
\end{align*}
(13)
We wish to establish an optimality condition for the above problem through certain reformulations.

Note that all variables with 0 indices are dummy variables and are for symmetry purposes only; we may simply regard all of them, such as $\bar{x}_0, \bar{y}_0, x_{0}^*, y_0^*$ and later $x_{0,j}^*, y_{0,j}^*$, for whatever $j$, null vectors.
It is obvious that \((x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_n)\) is a feasible solution of (13) if and only if there exists \(v := (v_1, \ldots, v_n) \in Y\) such that
\[
(x_{i-1}, y_{i-1}, x_i, y_i, v_i) \in (gph \, G_i \cap gph(-Q_i)) \quad \text{and} \quad (x, y) \in D
\]
for all \(i = 1, \ldots, n\). Now define \((2n + 1)\) sets \(\{\Lambda_i\}\) in the product space \(W := X \times Y \times Y\) by
\[
\Lambda_i := \{(x, y, v) \in W \mid (x_{i-1}, y_{i-1}, x_i, y_i, v_i) \in gph G_i\},
\Lambda_{n+i} := \{(x, y, v) \in W \mid (x_{i-1}, y_{i-1}, x_i, y_i, v_i) \in gph(-Q_i)\},
\Lambda_{2n+1} := \{(x, y, v) \in W \mid (x_1, \ldots, x_n, y_1, \ldots, y_n) \in D\}
\]
for \(i = 1, \ldots, n\). Then the problem (13) is equivalent to a constrained multiobjective optimization problem with a finite set of geometric constraints
\[
\Theta\text{-minimize} \quad \bar{f}(w)
\]
subject to \(w \in \Xi := \bigcap_{i=1}^{2n+1} \Lambda_i, \quad (15)\)
where \(\bar{f} : W \to H\) is defined by \(\bar{f}(x, y, v) := f(x, y)\).

We can then apply Lemma 2.3 to the above problem (15), and subsequently obtain a set of necessary conditions to the multistage and multiobjective hierarchical optimization problem (13).

In the following theorem, to simplify the notation, we write
\[
N^\Omega_i(x_{i-1}, y_{i-1}, x_i, y_i) := N(y_i \mid \Omega_i(x_{i-1}, y_{i-1}, x_i)), \quad i = 1, \ldots, n.
\]

**Theorem 3.1.** Suppose \((\bar{x}, \bar{y}, \bar{v}) := (\bar{x}_1, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_n, \bar{v}_1, \ldots, \bar{v}_n)\) solves the problem (15), in which case \((\bar{x}, \bar{y})\) is automatically a solution to the problem (13), \(\Theta\) is a locally closed ordering set containing the origin and \(f\) is locally Lipschitz around \((\bar{x}, \bar{y})\). Assume that the following qualification condition is fulfilled: the system of extended generalized equations
\[
\begin{align*}
x^*_{i,i} + x^*_{i,i+1} + x^*_{i,n+i} + x^*_{i,n+i+1} + x^*_{i,2n+1} = 0, & \quad i = 1, \ldots, n - 1 \\
y^*_{i,i} + y^*_{i,i+1} + y^*_{i,n+i} + y^*_{i,n+i+1} + y^*_{i,2n+1} = 0, & \quad i = 1, \ldots, n - 1 \\
x^*_{n,n} + x^*_{n,2n} + x^*_{n,2n+1} = 0, & \\
y^*_{n,n} + y^*_{n,2n} + y^*_{n,2n+1} = 0, & \\
\end{align*}
\]
with
\[
\begin{align*}
(ax^*_{i-1,1}, y^*_{i-1,1}, x^*_{i,i}, y^*_{i,i}) & \in D^* \partial y_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, \bar{v}_i)(v^*_i), \\
(ax^*_{i-1,n+i}, y^*_{i-1,n+i}, x^*_{i,n+i}, y^*_{i,n+i}) & \in D^* N^\Omega_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, \bar{v}_i)(v^*_i), \\
(ax^*_{1,2n+1}, \ldots, x^*_{n,2n+1}, y^*_{1,2n+1}, \ldots, y^*_{n,2n+1}) & \in N((\bar{x}, \bar{y}) \mid D),
\end{align*}
\]
for \(i = 1, \ldots, n\)
has only the trivial solutions:
\[
\begin{align*}
x^*_{i,i} = x^*_{i,i+1} = x^*_{i,n+i} = x^*_{i,n+i+1} = x^*_{i,2n+1} = 0, & \quad i = 1, \ldots, n - 1 \\
y^*_{i,i} = y^*_{i,i+1} = y^*_{i,n+i} = y^*_{i,n+i+1} = y^*_{i,2n+1} = 0, & \\
x^*_{n,n} = x^*_{n,2n} = x^*_{n,2n+1} = 0, & \\
y^*_{n,n} = y^*_{n,2n} = y^*_{n,2n+1} = 0, & \\
v^* := (v^*_1, \ldots, v^*_n) = 0, & \quad for \ i = 1, \ldots, n - 1.
\end{align*}
\]
Then there exist $\alpha^* \in N(0 \mid \Theta) \setminus \{0\}$, $v_i^* \in Y_i$, and
\[
(x^*_{i-1,1}, y^*_{i-1,1}, x^*_{i,1}, y^*_{i,1}) \in D^e \partial_y g_i(x_{i-1,1}, y_{i-1,1}, x_i, y_i) (v_i^*),
\]
\[
(x^*_{i-1,n+1}, y^*_{i-1,n+1}, x^*_{i,n+1}, y^*_{i,n+1}) \in D^e N^\Theta_i(\bar{x}_{i-1,1}, \bar{y}_{i-1,1}, \bar{x}_i, \bar{y}_i, -\bar{v}_i) (v^*_i),
\]
\[
(x^*_{1,2n+1}, \ldots, x^*_{n,2n+1}, y^*_{1,2n+1}, \ldots, y^*_{n,2n+1}) \in N((\bar{x}, \bar{y}) \mid D),
\]
\[
(x^*_{1,2n+2}, \ldots, x^*_{n,2n+2}, y^*_{1,2n+2}, \ldots, y^*_{n,2n+2}) \in D^e f(\bar{x}, \bar{y}) (\alpha^*),
\]
for $i = 1, \ldots, n$ satisfying the following system of equations
\[
\begin{cases}
  x^*_{i,i} + x^*_{i,i+1} + x^*_{i,n+i} + x^*_{i,n+i+1} + x^*_{i,2n+1} + x^*_{i,2n+2} = 0,
  \quad y^*_{i,i} + y^*_{i,i+1} + y^*_{i,n+i} + y^*_{i,n+i+1} + y^*_{i,2n+2} + y^*_{i,2n+2} = 0,
  \quad x^*_{i,n+n} + x^*_{i,2n+1} + x^*_{i,2n+2} = 0,
  \quad y^*_{i,n+n} + y^*_{i,2n+1} + y^*_{i,2n+2} = 0,
  \quad i = 1, \ldots, n-1.
\end{cases}
\]

**Proof.** First observe that a decision pair $(\bar{x}, \bar{y})$ solves the problem (13) if and only if there exists $\bar{v} \in Y$ such that $(\bar{x}, \bar{y}, \bar{v})$ is a solution to the problem (15). Now suppose $\tilde{w} := (\bar{x}, \bar{y}, \bar{v}) = (x_1, \ldots, x_n, \bar{y}_1, \ldots, \bar{y}_n, v_1, \ldots, v_n)$ is a $\Theta$-minimizer of $\tilde{f}$ to the problem (15). Then by employing Lemma 2.3 to the later problem at $\tilde{w}$, we know that there exists $\alpha^* \in N(0 \mid \Theta) \setminus \{0\}$ satisfying
\[
0 \in D^e \tilde{f}(\tilde{w}) (\alpha^*) + N(\tilde{w} \mid \Xi) = D^e f(\bar{x}, \bar{y}) (\alpha^*) \times \{0\} + N(\tilde{w} \mid \bigcap_{i=1}^{2n+1} \Lambda_i)
\]
where the equality holds true due to the definition of $\tilde{f}$.

Next, we further elaborate on the normals to the set $\Xi$ that is generated by the intersection of $(2n+1)$ sets $\{\Lambda_i \mid i = 1, \ldots, 2n+1\}$ by using the intersection rule [6, Theorem 3.4]. To do this, we need to check the fulfillment of the qualification condition that for $w_i^* \in N(\tilde{w} \mid \Lambda_i)$ with $i = 1, \ldots, 2n+1$ one has
\[
\sum_{i=1}^{2n+1} w_i^* = 0 \implies w_1^* = \ldots = w_{2n+1}^* = 0.
\]
Taking into account the constructions of the sets $\{\Lambda_i\}$ in (14), and the assumptions that $x^*_{0,1} = 0$ and $y^*_{0,1} = 0$, we have for $i = 1, \ldots, n$ that
\[
w_i^* = (0, \ldots, 0, x^*_{i-1,1}, x^*_{i,i}, 0, \ldots, 0, y^*_{i-1,i}, y^*_{i,i}, 0, \ldots, 0, -v^*_{i}, 0, \ldots, 0),
w_{n+i}^* = (0, \ldots, 0, x^*_{i-1,n+i}, x^*_{i,n+i}, 0, \ldots, 0, y^*_{i-1,n+i}, y^*_{i,n+i}, 0, \ldots, 0, v^*_{i}, 0, \ldots, 0),
w_{2n+1}^* = (x^*_{1,2n+1}, \ldots, x^*_{n,2n+1}, y^*_{1,2n+1}, \ldots, y^*_{n,2n+1}, 0, \ldots, 0)
\]
with
\[
(x^*_{i-1,1}, y^*_{i-1,1}, x^*_{i,i}, y^*_{i,i}) \in D^e G_i(x_{i-1,1}, y_{i-1,1}, x_i, y_i) (v_i^*),
\]
\[
(x^*_{i-1,n+i}, y^*_{i-1,n+i}, x^*_{i,n+i}, y^*_{i,n+i}) \in D^e Q_i(x_{i-1,1}, y_{i-1,1}, x_i, y_i, -\bar{v}_i) (v_i^*),
\]
\[
(x^*_{1,2n+1}, \ldots, x^*_{n,2n+1}, y^*_{1,2n+1}, \ldots, y^*_{n,2n+1}) \in N((\bar{x}, \bar{y}) \mid D).
\]
By direct substitutions we obtain the following form of implication that is equivalent to (20):

\[
\begin{bmatrix}
  x_{i,i} + x_{i,i+1}^* + x_{i,n+i}^* + x_{i,n+i+1}^* + x_{i,2n+1}^* = 0,
  y_{i,i} + y_{i,i+1}^* + y_{i,n+i}^* + y_{i,n+i+1}^* + y_{i,2n+1}^* = 0,
  (x_{n,n}^*, y_{n,n}^*) + (x_{n,2n}^*, y_{n,2n}^*) + (x_{n,2n+1}^*, y_{n,2n+1}^*) = (0, 0),
\end{bmatrix}
\]
for \( i = 1, \ldots, n \)

\[
\implies \begin{bmatrix}
  x_{i,i} = x_{i,i+1}^* = x_{i,n+i}^* = x_{i,n+i+1}^* = x_{i,2n+1}^* = 0,
  y_{i,i} = y_{i,i+1}^* = y_{i,n+i}^* = y_{i,n+i+1}^* = y_{i,2n+1}^* = 0,
  x_{n,n}^* = x_{n,2n}^* = x_{n,2n+1}^* = 0,
  y_{n,n}^* = y_{n,2n}^* = y_{n,2n+1}^* = 0,
\end{bmatrix}
\]

Since the qualification condition of the intersection rule is fulfilled by the assumed qualification condition (16), this rule yields

\[
N(\bar{w} \mid \bigcap_{i=1}^{2n+1} \Lambda_i) \subset N(\bar{w} \mid \Lambda_1) + \ldots + N(\bar{w} \mid \Lambda_{2n+1}).
\]

Again, by substituting this upper estimate into (19) and combining it with (20), we have for \( i = 1, \ldots, n - 1 \)

\[
\left\{ \begin{array}{l}
  x_{i,i} + x_{i,i+1}^* + x_{i,n+i}^* + x_{i,n+i+1}^* + x_{i,2n+1}^* = 0, \\
  y_{i,i} + y_{i,i+1}^* + y_{i,n+i}^* + y_{i,n+i+1}^* + y_{i,2n+1}^* = 0, \\
  (x_{n,n}^*, y_{n,n}^*) + (x_{n,2n}^*, y_{n,2n}^*) + (x_{n,2n+1}^*, y_{n,2n+1}^*) = (0, 0)
\end{array} \right.
\]

where

\[
(x_{i,i}^*, y_{i,i}^*, x_{i,i+1}^*, y_{i,i+1}^*, x_{i,n+i}^*, y_{i,n+i}^*, x_{i,n+i+1}^*, y_{i,n+i+1}^*, x_{i,2n+1}^*, y_{i,2n+1}^*) \in D^*G_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, \bar{v}_i)(v_i^*),
\]

\[
(x_{i,n+i}^*, y_{i,n+i}^*, x_{i,n+i+1}^*, y_{i,n+i+1}^*) \in D^*Q_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, -\bar{v}_i)(v_i^*),
\]

\[
(x_{2n,2n+1}^*, \ldots, x_{2n,2n+1}^*, y_{2n,2n+1}^*, \ldots, y_{2n,2n+1}^*) \in N((\bar{x}, \bar{y}) \mid D),
\]

\[
(x_{1,2n+2}^*, \ldots, x_{1,2n+2}^*, y_{1,2n+2}^*, \ldots, y_{1,2n+2}^*) \in D^*f(\bar{x}, \bar{y})(\alpha^*)
\]

for \( i = 1, \ldots, n \). We arrived at the conditions (17) and (18), and therefore conclude the proof. \( \square \)

The next result provides a much simplified form of the necessary condition in Theorem 3.1 where we assume our problem (13) possesses the following specific set of properties:

\[
f(x_1, \ldots, x_n, y_1, \ldots, y_n) := f_1(x_1, y_1) + \ldots + f_n(x_n, y_n),
\]

\[
\Omega_i(x_i-1, y_i-1, x_i) \equiv \Omega_i \subset Y_i \quad \forall (x_i-1, y_i-1, x_i),
\]

\[
g_i(x_i-1, y_i-1, x_i, y_i) \equiv g_i(x_i, y_i) \quad \forall i = 1, \ldots, n.
\]

In the following proposition, for a given set \( C \) and \( z \in C \), \( \delta(z \mid C) \) is the indicator function of \( C \) at \( z \), namely \( \delta(z \mid C) = 0 \) if \( z \in C \) and \( +\infty \) otherwise.

**Proposition 1.** Let \((\bar{x}, \bar{y}, \bar{v})\) solve the problem (15), in which case \((\bar{x}, \bar{y})\) is a \((f, \Theta)\)-minimizer of the problem (13), where \( \Theta \) is a locally closed ordering set containing
the origin, and \( f \) is locally Lipschitz continuous at the point \((\bar{x}, \bar{y})\). Assume further that the following simplified qualification condition is fulfilled:

\[
\begin{bmatrix}
(x^*_{i,i}, y^*_{i,i}) + (0, y^*_{i,n+i}) + (x^*_i, \ldots, x^*_{i,2n+1}, y^*_{2n+1}) = (0, 0) \\
\text{with } (x^*_{i,i}, y^*_{i,i}) \in D^*\partial g_i(\bar{x}_i, \bar{y}_i, \bar{v}_i) v^*_i,
\end{bmatrix}
\]

\[
y^*_{i,n+i} \in \partial^2 \delta(\cdot | \Omega_i)\bar{v}_i(v^*_i),
\]

\[
(x^*_i, \ldots, x^*_{i,2n+1}, y^*_i, \ldots, y^*_{2n+1}) \in N((\bar{x}, \bar{y}) | D),
\]

for \( i = 1, \ldots, n \).

Then there are \( \alpha^* \in N(0 | \Theta) \setminus \{0\}, v^*_i \in Y_i \) for \( i = 1, \ldots, n \), and

\[
(x^*_i, y^*_i) \in D^* f_i(\bar{x}_i, \bar{y}_i)(\alpha^*)
\]

\[
(x^*_{i,i}, y^*_{i,i}) \in D^* \partial g_i(\bar{x}_i, \bar{y}_i, \bar{v}_i) v^*_i, \quad i = 1, \ldots, n,
\]

\[
y^*_{i,n+i} \in \partial^2 \delta(\cdot | \Omega_i)(\bar{v}_i, \bar{v}_i)(v^*_i), \quad i = 1, \ldots, n,
\]

\[
(x^*_i, \ldots, x^*_{i,2n+1}, y^*_i, \ldots, y^*_{2n+1}) \in N((\bar{x}, \bar{y}) | D),
\]

such that the system of equations

\[
(x^*_i, y^*_i) + (x^*_{i,i}, y^*_{i,i}) + (0, y^*_{i,n+i}) + (x^*_i, \ldots, x^*_{i,2n+1}, y^*_i, \ldots, y^*_{2n+1}) = (0, 0)
\]

holds true for \( i = 1, \ldots, n \).

**Proof.** This is a straightforward special case of Theorem 3.1 by taking into account the special structures of the objective functions \( f \) and \( g_i \), and that of the constrained sets \( \Omega_i \) for all \( i \in \{1, \ldots, n\} \). In particular, we have a reduced form of the mapping

\[
N^\Omega(x_{i-1}, y_{i-1}, z, y_i) = N^\Omega(y_i) := N(\bar{y}_i | \Omega_i) = \partial \delta(\bar{y}_i | \Omega_i)
\]

and consequently

\[
D^* N_i^\Omega(\bar{y}_i, -\bar{v}_i)(v^*_i) = \partial^2 \delta(\cdot | \Omega_i)(\bar{v}_i, \bar{v}_i)(v^*_i)
\]

for all \( i = 1, \ldots, n \). We also have that

\[
D^* f_i(\bar{x}_i, \bar{y}_i, \ldots, \bar{x}_n, \bar{y}_n)(\alpha^*) = D^* f_i(\bar{x}_i, \bar{y}_i)(\alpha^*) + \ldots + D^* f_i(\bar{x}_n, \bar{y}_n)(\alpha^*).
\]

The conclusion follows. \( \square \)

We now turn to the case where the lower-level constraint sets \( \Omega_i \) are assumed to be polyhedral with respect to the lower level decision variable \( y_i \) for all \( i \), and obtain appropriate optimality conditions. We first state the coderivative representation of the normal cone to a polyhedral set obtained in [3].

Given a finite index set \( T \subset \mathbb{N} \), a finite dimensional vector space \( Z \), and a set of vectors \( \{c_j \in Z | j \in T\} \), define

\[
\Delta := \{ z \in Z | \langle c_j, z \rangle \leq 0, j \in T \}.
\]

For any \( \bar{z} \in \Delta \), we may regard the normal cone \( N(\bar{z} | \Delta) \) to \( \Delta \) at \( \bar{z} \) as a set-valued map and write \( F_\Delta(\bar{z}) := N(\bar{z} | \Delta) \). The coderivative of this mapping is denoted by \( D^* F_\Delta \). We use \( I(\bar{z}) \) to represent the set of indices of all active constraints of \( \Delta \) at \( \bar{z} \), namely \( I(\bar{z}) := \{ j \in T | \langle c_j, \bar{z} \rangle = 0 \} \).
Lemma 3.2. Let \( \bar{z}^* \in N(\bar{z} \mid \Delta) \) and assume the generating vectors \( \{ c_j \} \) are linearly independent. Then for each \( u \in \text{Dom} D^* \mathcal{F}_\Delta(\bar{z}, \bar{z}^*) \), one has
\[
D^* \mathcal{F}_\Delta(\bar{z}, \bar{z}^*)(u) = \sum_{j \in I_0(u), \mu_j \in \mathbb{R}} \mu_j c_j + \sum_{j \in I_\sim(u), \eta_j \in \mathbb{R}_+} \eta_j c_j
\]  
(23)
where \( I_0(u) := \{ j \in I(\bar{z}) \mid \langle c_j, u \rangle = 0 \} \), \( I_\sim(u) := \{ j \in I(\bar{z}) \mid \langle c_j, u \rangle > 0 \} \).

Proof. [3, Theorem 4.6]. \qed

In what follows, we consider the special case of the multistage hierarchical problem where the lower-level problem \((Q_i)\) has a smooth objective function and a polyhedral constrained set at each stage. Assume, for each \( i \in \{1, \ldots, n\} \), the objective function \( g_i \) is twice continuously differentiable, and there is an index set \( T_i \) and a set of linearly independent vectors \( \{ a_i^j \mid j \in T_i \} \) such that the lower-level constraint set \( \Omega_i \) takes the form:
\[
\Omega_i(x_{i-1}, y_{i-1}, x_i) = \Omega_i := \{ y_i \in \mathcal{Y}_i \mid \langle y_i, a_i^j \rangle \leq 0 \forall j \in T_i \}.
\]  
(24)
Let \( \mathcal{F}_i(y_i) := N(y_i \mid \Omega_i) \). Then in the optimality conditions (17) of Theorem 3.1, for any \( i \in \{1, \ldots, n\} \) the inclusion
\[
(x_{i-1,n+i}^*, y_{i-1,n+i}^*, x_{i,n+i}^*, y_{i,n+i}^*) \in D^* N_i^\Omega(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, -\bar{v}_i)(v_i^*)
\]  
(25)
holds if and only if \( x_{i-1,n+i}^* = 0, y_{i-1,n+i}^* = 0, x_{i,n+i}^* = 0 \) and
\[
y_{i,n+i}^* \in D^* \mathcal{F}_i(\bar{y}_i, -\bar{v}_i)(v_i^*)
\]  
(26)
due to the fact that \( N_i^\Omega \) is independent of \( x_{i-1}, y_{i-1} \) and \( x_i \), and consequently
\[
N ((\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, -\bar{v}_i) \mid \text{gph} N_i^\Omega) = (0, 0, 0) \times N ((\bar{y}_i, -\bar{v}_i) \mid \text{gph} \mathcal{F}_i).
\]
By applying Lemma 3.2 to \( D^* \mathcal{F}_i(\bar{y}_i, -\bar{v}_i)(v_i^*) \), we have that (25) holds if and only if there exist constants \( \mu_i^j \in \mathbb{R} \) and \( \eta_i^j \in \mathbb{R}_+ \) such that
\[
y_{i,n+i}^* = \sum_{j \in I_0^i(v_i^*)} \mu_i^j a_i^j + \sum_{j \in I_\sim^i(v_i^*)} \eta_i^j a_i^j
\]  
(27)
where \( I_0^i(v_i^*) := \{ j \in I_i^i(\bar{y}_i) \mid \langle a_i^j, v_i^* \rangle = 0 \} \), \( I_\sim^i(v_i^*) := \{ j \in I_i^i(\bar{y}_i) \mid \langle a_i^j, v_i^* \rangle > 0 \} \) and \( I_i^i(\bar{y}_i) := \{ j \in T_i \mid \langle a_i^j, \bar{y}_i \rangle = 0 \} \).

On the other hand, since \( \{ g_i \mid i = 1, \ldots, n \} \) are twice continuously differentiable, we have that the condition
\[
(x_{i-1,i}^*, y_{i-1,i}^*, x_{i,i}^*, y_{i,i}^*) \in D^* \partial_{g_i} g_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, \bar{v}_i)(v_i^*)
\]  
(28)
appearing in the optimality conditions of Theorem 3.1 is equivalent to
\[
(x_{i-1,i}^*, y_{i-1,i}^*, x_{i,i}^*) = [\nabla_{x_{i-1,i}, y_{i-1,i}} g_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, \bar{v}_i)]^\top v_i^*
\]  
(29)
and
\[
y_{i,i}^* = [\nabla_{y_{i,i}}^2 g_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}_i, \bar{y}_i, \bar{v}_i)]^\top v_i^*
\]  
(30)
where \( \nabla_{x_{i-1,i}, y_{i-1,i}} \nabla_{y_i} g_i \) and \( \nabla_{y_{i,i}}^2 g_i \) are the Jacobian matrix of the mapping \( \nabla_{y_i} g_i \) with respect to \( (x_{i-1,i}, y_{i-1,i}) \) and \( y_i \) respectively.

Applying (28), (29) and (30) to Theorem 3.1, we have the following proposition.
Proposition 2. Suppose $(\bar{x}, \bar{y}, \bar{v}) := (\bar{x}_1, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_i, \ldots, \bar{v}_n, \ldots, \bar{v}_n)$ solves the problem (15). Assume, in addition to the assumptions of Theorem 3.1, that at each stage $i \in \{1, \ldots, n\}$ the lower level objective function $g_i$ is twice continuously differentiable in all variables, and the constraint set $\Omega_i$ has the form (24). Then there exist $\alpha^* \in N(0 \mid \Theta) \setminus \{0\}$, $v_i^* \in Y_i$ for $i = 1, \ldots, n$, the constants $\mu_i^j \in \mathbb{R}$ and $\eta_i^j \in \mathbb{R}_+$, and

\[
(x^*_{i-1}, i, y^*_{i-1}, i, x^*_i) = \left[ \nabla_{x_{i-1}, y_{i-1}, x} g_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}, \bar{y}, \bar{v}) \right] v_i^*,
\]

\[
y^*_{i,i} = \left[ \nabla_{y_{i-1}}^2 g_i(\bar{x}_{i-1}, \bar{y}_{i-1}, \bar{x}, \bar{y}, \bar{v}) \right] v_i^*,
\]

\[
y^*_{i,n+i} = \sum_{j \in I^*_i(v_i^*)} \mu^j_i a^j + \sum_{j \in I^*_i(v_i^*)} \eta^j_i a^j,
\]

\[
(x^*_{1,2n+1}, \ldots, x^*_{n,2n+1}, y^*_{1,2n+1}, \ldots, y^*_{n,2n+1}) \in N((\bar{x}, \bar{y}) \mid D),
\]

\[
(x^*_{1,2n+2}, \ldots, x^*_{n,2n+2}, y^*_{1,2n+2}, \ldots, y^*_{n,2n+2}) \in D^* f(\bar{x}, \bar{y})(\alpha^*),
\]

\[
i = 1, \ldots, n,
\]

satisfying the following system of equations

\[
x^*_{i,i} + x^*_{i,i+1} + x^*_{i-1,2n+1} + x^*_{i-1,2n+2} = 0, \quad i = 1, \ldots, n - 1,
\]

\[
y^*_{i,i} + y^*_{i,i+1} + y^*_{i,n+i} + y^*_{i-1,2n+2} + y^*_{i-1,2n+2} = 0, \quad i = 1, \ldots, n - 1,
\]

\[
x^*_{n,n} + x^*_{n,2n+1} + x^*_{n,2n+2} = 0,
\]

\[
y^*_{n,n} + y^*_{n,2n+2} + y^*_{n,2n+2} = 0.
\]

Proof. Substituting (25), (27), (28), (29) and (30) into the condition (17) of Theorem 3.1.

To conclude this section let us compare the approach used in this paper with the value function approach to bilevel programming. In order to study nonsmooth bilevel/multi-level programming we convert the given problem to an optimization problem with an equilibrium constraint. Since $i = 1$ we will drop all indexes in (13) and thus it reduces to

\[
\Theta \text{-minimize } \quad f(x, y)
\]

subject to \quad $0 \in \partial_y g(x, y) + N(y; \Omega(x)), \quad (31)$

\[
(x, y) \in D.
\]

Then the optimal solution (if exists) $(\bar{x}, \bar{y})$ ensures the existence of $\bar{v} \in \partial_y g(\bar{x}, \bar{y}) \cap \left( N(\bar{y} \mid \Omega(\bar{x})) \right)$. It has been shown that under very mild assumptions imposed on initial data at the triple $(\bar{x}, \bar{y}, \bar{v})$ the necessary conditions in this paper and the paper [1] hold. In contrast, the known value function approach requires the so-called calmness condition; see, for example, [2]. Such a condition is quite restrictive but unavoidable since the latter approach concerns the following equivalent problem

\[
\Theta \text{-minimize } \quad f(x, y)
\]

subject to \quad $g(x, y) - \phi(x) \leq 0, \quad (32)$

\[
y \in \Omega(x), \quad (x, y) \in D,
\]

where $\phi : X \to \mathbb{R} \cup \{\infty\}$ is the value function of the lower-level optimization problem, and imposes assumptions at the optimal solution $(\bar{x}, \bar{y})$. In the other words, it allows some kind of freedom in choosing $\bar{v}$ at each $(\bar{x}, \bar{y})$ due to the fact that $\partial_y g(x, y)$ is in general not a singleton, and thus the calmness condition is needed to guarantee
a well-behavior of $\bar{v}$. Note finally that both approaches might yield the classical necessary conditions for bilevel programming with differentiation data.

4. Conclusions. A multistage hierarchical optimization problem with multicriterion upper-level objective function is proposed and its distinctive structure is described in details. The optimality conditions are obtained for the problem where the lower-level solution sets are allowed to be multi-valued. The results are then applied to several special cases where the underline problem possesses some additional properties.

REFERENCES

[1] T. Bao, P. Gupta and B. Mordukhovich, Necessary conditions in multiobjective optimization with equilibrium constraints, Journal of Optimization Theory and Applications, 135 (2007), 179–203.

[2] S. Dempe, J. Dutta and B. Mordukhovich, New necessary optimality conditions in optimistic bilevel programming, Optimization, 56 (2007), 577–604.

[3] R. Henrion, B. Mordukhovich and N. Nam, Second-order analysis of polyhedral systems in finite and infinite dimension with applications to robust stability of variational inequalities, SIAM J. Optim., 20 (2009), 2199–2227.

[4] M. Kočvara, M. Kružík and J. Outrata, On the control of an evolutionary equilibrium in micromagnetics, in “Optimization with Multivalued Mappings” (eds. S. Dempe and V. Kalashnikov), Springer, New York, (2006), 143–168.

[5] M. Kočvara and J. Outrata, On the modeling and control of delamination processes, in “Control and Boundary Analysis” (eds. J. Cagnol and J.-P. Yolesion), Marcel Dekker, (2004), 171–190.

[6] B. Mordukhovich, “Variational Analysis and Generalized Differentiation, Vol. I: Basic Theory,” Springer, Berlin, 2006.

[7] B. Mordukhovich, “Variational Analysis and Generalized Differentiation, Vol. 2: Applications,” Springer, Berlin, 2006.

[8] B. Mordukhovich, Maximum principle in problems of time optimal control with nonsmooth constraints, J. Appl. Math. Mech., 40 (1976), 960–969.

[9] R. Rockafellar and R. Wets, “Variational Analysis,” Springer, Berlin, 1998.

[10] G. Tseng, S. Tsau and J. Chen, Application of hierarchy multistage-multiobjective approach to network design: case of express road, in “Global Engineering Management: Emerging Trends in the Asia Pacific,” Proceedings of 1995 IEEE Annual International Publication, (1995), 370–377.

[11] R. Zhang, Multistage bilevel programming problems, Optimization, 52 (2003), 605–616.

Received October 2009; 1st revision April 2010; final revision October 2010.

E-mail address: rzhang@nmu.edu
E-mail address: btruong@nmu.edu
E-mail address: qzhang@nmu.edu