

# Semivectorial Bilevel Optimization on Riemannian Manifolds

Henri Bonnel<sup>1</sup> · Léonard Todjihoundé<sup>2</sup> ·  
Constantin Udriște<sup>3</sup>

Received: 26 November 2014 / Accepted: 24 July 2015 / Published online: 4 August 2015  
© Springer Science+Business Media New York 2015

**Abstract** In this paper, we deal with the semivectorial bilevel problem in the Riemannian setting. The upper level is a scalar optimization problem to be solved by the leader, and the lower level is a multiobjective optimization problem to be solved by several followers acting in a cooperative way inside the greatest coalition and choosing among Pareto solutions with respect to a given ordering cone. For the so-called optimistic problem, when the followers choice among their best responses is the most favorable for the leader, we give optimality conditions. Also for the so-called pessimistic problem, when there is no cooperation between the leader and the followers, and the followers choice may be the worst for the leader, we present an existence result.

**Keywords** Multiobjective optimization on Riemannian manifolds · Semivectorial bilevel optimization problem · Bilevel optimization

---

Communicated by Johannes Jahn.

---

✉ Henri Bonnel  
henri.bonnel@univ-nc.nc  
Léonard Todjihoundé  
leonardt67@gmail.com  
Constantin Udriște  
udriste@mathem.pub.ro

<sup>1</sup> ERIM, Université de la Nouvelle-Calédonie, BP. R4, 98851 Nouméa Cédex, New Caledonia, France

<sup>2</sup> Institut de Mathématiques et de Sciences Physiques, Porto-Novo, Benin

<sup>3</sup> Department of Mathematics-Informatics, Faculty of Applied Sciences, University Politehnica of Bucharest, Splaiul Independentei 313, 060042 Bucharest, Romania

**Mathematics Subject Classification** 49K30 · 49K35 · 49K99 · 58E17 · 91A12 · 91A65 · 90C29

## 1 Introduction

The aim of this paper was to study a semivectorial bilevel optimization problem on Riemannian manifolds.

A semivectorial bilevel optimization problem is a bilevel problem, where the upper level is a scalar optimization problem, and the lower level is a multiobjective optimization problem.

Multiobjective optimization finds its origin in the nineteenth century in the economic works of Edgeworth [1] and Pareto [2], but, in mathematics, this branch commenced in 1951 with the famous paper of Kuhn and Tucker [3]. Due to its various applications in real-life problems, intensively researches have been made in vector optimization in the last 50 years. Dealing with several conflicting objectives, a Pareto solution (called also efficient) is such that none of the objective values can be improved further without deteriorating another.

The lower level of the semivectorial bilevel optimization problem is a parametric multiobjective (vector) optimization problem and may be considered as a single follower having to optimize several objectives, or as several followers each of them having to optimize one (scalar) objective. The last situation corresponds to the so-called greatest coalition multiplayer game. The parameter in the lower level problem is the (vector) variable chosen by the leader, and, for each choice of the leader's variable, the followers choose a Pareto solution.

In a semivectorial bilevel optimization problem, the upper level is a scalar optimization problem to be solved by the leader. The leader objective depends on two (vector) variables: one chosen by the leader, and the second one represents the response of the followers.

If for each choice of the leader the followers choose among their best responses (Pareto solutions) one which is the best for the leader, so, when the followers cooperate with the leader, we deal with the so-called optimistic problem.

In the case when there is no cooperation between the leader and the followers, the leader may consider the worst scenario, i.e., the situation when, for each choice of the leader, the followers choose, among their best responses, one which is the most unfavorable for the leader, leading to the so-called pessimistic problem.

The study of semivectorial optimization problems in Euclidean or Hilbert spaces has begun in 2006; see papers [4, 5], and it was continued by several authors; see papers [6–10]. The case of semivectorial bilevel optimal control problems was considered in [11, 12].

The semivectorial bilevel optimization problem includes as particular cases the following problems which have been intensively studied in the last decades so we will give essentially a few earlier references,

- Optimizing a scalar function over the Pareto set (introduced in [13] and investigated in [14–29] and [30] for a survey);
- Bilevel optimization problems where the upper level and lower level are scalar optimization problems (e.g., [31–36] for an extensive bibliography).

Extension of optimization problem from Euclidean spaces to Riemannian manifolds is natural and nontrivial (see, e.g., [37–41] and the references therein). Such extensions have different advantages. For example, some constrained optimization problems can be seen as unconstrained ones from the Riemannian geometry viewpoint. Moreover, some nonconvex optimization problems in the Euclidean setting may become convex introducing an appropriate Riemannian metric (see, e.g., [37–41]).

In the last years, researchers began the study of the multiobjective optimization problems on Riemannian manifolds [42–44].

It seems that our paper is the first approach in the literature for a semivectorial bilevel problem on Riemannian manifolds. Moreover, with difference of the multiobjective problems on Riemannian manifolds presented in several papers (see, e.g., [42,43]), we consider the multiobjective lower level problem in the more general setting when the partial order is given by an arbitrary closed, convex and pointed cone, and not only by the positive orthant (Pareto cone). This generalization is motivated by the fact that many application problems need the use of arbitrary closed, convex and pointed cones (see, for instance, [45–50] and the references therein).

It is worth to note that some of our results are new also for the Euclidean case. Thus, in our approach in finite dimensions, inspired by the technics used in [11, 12] for semivectorial bilevel convex optimal control problems, where the optimality conditions are related only with the linear-quadratic case for the lower level, we deal now on one hand with the general nonlinear case for both upper and lower level obtaining more explicit optimality conditions, and on the other hand with an arbitrary convex, closed and pointed ordering cone.

Our paper is organized as follows. In Sect. 2, we present first some basic facts in Riemannian geometry and then some aspects of vector optimization on Riemannian manifolds. We deal with weakly as well as properly Pareto solutions with respect to an arbitrary partial order and introduce different kinds of convexity on Riemannian manifolds with respect to the ordering cone. The classical convex scalarization theorem is now proved for the Riemannian setting.

Section 3 contains the main results. It is divided into six subsections. First, we state the optimistic and pessimistic semivectorial bilevel problems. In the second subsection, based on the scalarization theorem for convex multiobjective problems presented in the previous section, we rewrite our problems as bilevel optimization problems with a scalar lower-level problem which has a unique solution. Moreover, we show that under our hypotheses, the weakly Pareto set coincides with the Pareto set. In the third subsection, we give optimality conditions for this scalar lower-level problem. In the fourth subsection, we present optimality conditions for the optimistic semivectorial problem for the case of (weakly) Pareto case as well as for the properly Pareto one. In the fifth subsection, we give an existence result for the more difficult case of the semivectorial bilevel pessimistic problem. In the last subsection, we give an illustrative example, where the lower level is given by a nonconvex multiobjective problem on the initial Euclidean space, but becomes convex for a suitable Riemannian metric. Thus, taking advantage of our results presented in Sects. 2, 3 and 4, we are able to find the optimal solution.

Some concluding remarks are presented the last section.

## 2 Preliminaries

### 2.1 Basics in Riemannian Geometry

An  $n$ -dimensional Riemannian manifold is a pair  $(M, g)$ , where  $M$  stands for an  $n$ -dimensional smooth manifold and  $g$  stands for a smooth, symmetric positive definite  $(0, 2)$ -tensor field on  $M$ , called a Riemannian metric on  $M$ . If  $(M, g)$  is a Riemannian manifold; then, for any point  $p \in M$ , the restriction  $g_p : T_pM \times T_pM \rightarrow \mathbb{R}$  is an inner product on the tangent space  $T_pM$ .

The tangent bundle  $TM$  over  $M$  is  $TM := \bigcup_{x \in M} T_xM$ , and a vector field on  $M$  is a section of the tangent bundle, that is a map  $X : M \rightarrow TM$  such that for any  $p \in M, X(p) \equiv X_p \in T_pM$ . The set of all  $C^\infty$  vector fields on  $M$  is denoted  $\Gamma(TM)$  or  $\chi(M)$ .

By identifying a vector field with its image set, the sets  $\Gamma(TM)$  and  $TM$  can be identified. In all the paper, we use the Einstein summation convention: that is, the summation symbol is omitted when the same index occurs twice in a product, once as an upper index and once as a lower index. Moreover, we will consider only connected Riemannian manifolds.

In local coordinates  $(x^1, \dots, x^n)$  around the point  $x$  on  $M$ , the Riemannian metric  $g$  is written  $g = g_{ij}dx^i dx^j$  with  $g_{ij} = g_x(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$ , where  $(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n})$  is the natural basis of  $T_xM$ , and  $(dx^1, \dots, dx^n)$  is the corresponding dual basis on the cotangent vector space  $T_x^*M$ .

The length of a tangent vector  $v \in T_pM$  is defined by  $\|v\|_p := g_p(v, v)^{\frac{1}{2}}$ .

If  $\gamma : [a, b] \subset \mathbb{R} \rightarrow M$  is a piecewise smooth curve in  $M$ , then its length is defined by  $L(\gamma) := \int_a^b \|\dot{\gamma}(t)\|_{\gamma(t)} dt$ , where  $\dot{\gamma}$  means the first derivative of  $\gamma$  w.r.t to  $t$ .

Let  $p$  and  $q$  be two points in  $(M, g)$  and  $\Gamma_{pq}$  the set of all piecewise smooth curves joining  $p$  and  $q$ . The function

$$d : M \times M \rightarrow \mathbb{R}, \quad d(p, q) := \inf\{L(\gamma) : \gamma \in \Gamma_{pq}\}$$

is a distance on  $M$ , and the induced metric topology on  $M$  coincides with the topology of  $M$  as manifold.

Let  $M_1$  and  $M_2$  be two manifolds. The tangent bundle of the product manifold  $M_1 \times M_2$  can be written as  $T(M_1 \times M_2) = TM_1 \times TM_2$ , and if  $g_1$  and  $g_2$  are two Riemannian metrics on  $M_1$  and  $M_2$ , respectively, then the corresponding Riemannian metric product  $g$  on  $M_1 \times M_2$  is defined by

$$g((X_1, X_2), (Y_1, Y_2)) = g_1(X_1, Y_1) + g_2(X_2, Y_2)$$

for all  $X_1, Y_1 \in TM_1$  and  $X_2, Y_2 \in TM_2$ .

A linear connection or equivalently a covariant derivative on a  $n$ -dimensional smooth manifold  $M$  is a map  $D : TM \times TM \rightarrow TM$  such that, by putting  $D(X, Y) = D_X Y$ , we have

$D_{X+fY}Z = D_XZ + fD_YZ, D_X(Y+Z) = D_XY + D_XZ$  and  $D_X fY = df(X).Y + f.D_XY$ , for any  $X, Y, Z \in TM$  and any  $f \in C^\infty(M)$ , the set of smooth real functions on  $M$ .

Let  $x \in M, (x^1, \dots, x^n)$  be a local coordinates system on  $M$  around the point  $x$  and let  $(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n})$  be the induced basis of  $T_xM$ . Then for  $i, j = 1, \dots, n, D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$  is a tangent vector to  $M$  at the point  $x$ . Thus, there exist  $n^3$  smooth functions  $\Gamma_{ij}^k, i, j, k = 1, \dots, n$  on  $M$ , such that:  $D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma_{ij}^k \frac{\partial}{\partial x^k}$ .

The functions  $\Gamma_{ij}^k$  are called the Christoffel symbols of the connexion  $D$ .

If  $(M, g)$  is a Riemannian manifold, there exists only one connexion  $\nabla$  on  $M$  and such that  $\nabla_X Y = \nabla_Y X$  and  $X.g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \forall X, Y, Z \in TM$ . This privileged connexion is called Levi-Civita connexion and its Christoffel symbols contain all the geometric information on the Riemannian manifold  $(M, g)$ . In a local coordinates system  $(x^1, \dots, x^n)$ , the Christoffel symbols of the Levi-Civita connexion on  $(M, g)$  are given by:

$\Gamma_{ij}^k = \frac{1}{2}g^{kl} \left( \frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{il}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l} \right)$ , where the matrix  $(g^{kl})$  is the inverse of the matrix  $(g_{ij})$ .

A piecewise smooth curve  $\gamma : [a, b] \rightarrow M$  is said to be parametrized by arclength if  $\|\dot{\gamma}(t)\|_{\gamma(t)}$  is constant on  $[a, b]$ , and  $\gamma$  is called a geodesic joining the points  $\gamma(a)$  and  $\gamma(b)$  if for any  $t \in [a, b], \nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = 0$ . If moreover the length  $L(\gamma)$  is equal to the distance between the points  $\gamma(a)$  and  $\gamma(b)$ , then  $\gamma$  is said to be a minimizing geodesic. A geodesic curve is always parametrized by arclength. In a local coordinates system  $(x^1, \dots, x^n)$ , the curve  $t \mapsto \gamma(t) = (\gamma^1(t), \dots, \gamma^n(t))$  is a geodesic if and only if

$$\ddot{\gamma}^k(t) + \Gamma_{ij}^k(\gamma(t))\dot{\gamma}^i(t)\dot{\gamma}^j(t) = 0, \quad \forall t \in [a, b], \quad \forall k = 1, \dots, n,$$

where  $\dot{\gamma}^k$  means the first derivative of  $\gamma^k$  w.r.t to  $t$ .

Given a point  $p \in M$  and a tangent vector  $v \in T_pM$ , there exists  $\varepsilon > 0$  and precisely one geodesic  $\gamma_v : [0, \varepsilon] \rightarrow M$ , depending smoothly on  $p$  and  $v$ , such that  $\gamma_v(0) = p$  and  $\dot{\gamma}_v(0) = v$ .

Let  $p \in M$  and  $V_p := \{v \in T_pM : \gamma_v \text{ is defined on } [0, 1]\}$ . Then, the map  $\exp_p : V_p \rightarrow M$  defined by  $\exp_p(v) = \gamma_v(1), \forall v \in V_p$ , is called the exponential map of  $M$  at the point  $p$ . The exponential map  $\exp_p$  maps a neighborhood of the null tangent vector  $0 \in T_pM$  diffeomorphically onto a neighborhood of  $p$  in  $M$ .

The Riemannian manifold  $(M, g)$  is said to be geodesically complete if for any point  $p \in M$ , the exponential map  $\exp_p$  is defined on all of  $T_pM$ , that is if any geodesic  $\gamma$  containing the point  $p$  is defined on all  $\mathbb{R}$ . We have the following result known as the theorem of Hopf–Rinow.

**Theorem 2.1** (Hopf-Rinow) *Let  $M$  be a connected Riemannian manifold. The following statements are equivalent:*

- (i)  $M$  is complete as metric space.
- (ii)  $M$  is geodesically complete.

(iii) *The closed and bounded subsets of  $M$  are compact. Furthermore, each of the statements (i) – (iii) implies that any two points of  $M$  can be joined by a minimizing geodesic.*

Let  $(M, g)$  and  $(N, h)$  be two Riemannian manifolds and  $F : M \rightarrow N$  be a differentiable map between  $M$  and  $N$ . The differential  $dF(p)$  of  $F$  at a point  $p \in M$  is a linear map from  $T_pM$  into  $T_{F(p)}N$ . The map  $F$  is said to be conformal if there exists a positive function  $\alpha : M \rightarrow \mathbb{R}$ , called conformal factor, such that for any  $p \in M$  and any  $u, v \in T_pM$ ,  $h_{F(p)}(dF(p)(u), dF(p)(v)) = \alpha(p)g_p(u, v)$ .

A conformal map preserves angles between tangent vectors.

If the conformal factor is a constant function equal to one, then  $F$  is called an isometry.

The gradient of a differentiable function  $f : M \rightarrow \mathbb{R}$  w.r.t. the Riemannian metric  $g$  is the vector field  $\text{grad} f$  defined by  $g(\text{grad} f, X) = df(X)$ ,  $\forall X \in TM$ , where  $df$  denotes the differential of the function  $f$ .

The Hessian of a smooth function  $f : M \rightarrow \mathbb{R}$  is the 2-form  $\text{Hess}(f)$  defined by  $\text{Hess}(f)(X, Y) = \nabla df(X, Y) = g(\nabla_X \text{grad} f, Y)$ ,  $\forall X, Y \in TM$ .

In local coordinates  $(x^1, \dots, x^n)$  around  $p \in M$ , we can write:

$$\begin{aligned} \text{grad} f(p) &= g^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j}, \\ \text{Hess}(f) &= \nabla df = \left( \frac{\partial^2 f}{\partial x^i \partial x^j} - \Gamma_{ij}^k \frac{\partial f}{\partial x^k} \right) dx^i dx^j. \end{aligned}$$

Let  $p \in M$  be fixed. The inverse of the exponential map  $\exp_p^{-1}$  maps diffeomorphically a neighborhood of  $p$  onto a neighborhood of the origin of  $T_pM$ . Considering an orthonormal basis  $(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n})$  in  $T_pM$  with respect to the scalar product  $g_p(\cdot, \cdot)$ , this diffeomorphism establishes a local coordinate system  $(x^1, \dots, x^n)$  around the point  $p$  called *normal coordinate system*. In this normal coordinates system, the geodesics through  $p$  are represented by lines passing through origin. Moreover, the matrix  $(g_{ij})$  associated with the bilinear form  $g$  at the point  $p$  in this orthonormal basis reduces to the identity matrix and the Christoffel symbols vanish. Thus, for any smooth function  $f : M \rightarrow \mathbb{R}$ , in normal coordinates around  $p$ , we obtain

$$\text{grad} f(p) = \sum_i \frac{\partial f}{\partial x^i}(p) \frac{\partial}{\partial x^i},$$

and, for all  $u = u^i \frac{\partial}{\partial x^i}$ ,  $v = v^j \frac{\partial}{\partial x^j} \in T_pM$ ,

$$\text{Hess}_p(f)(u, v) = \frac{\partial^2 f}{\partial x^i \partial x^j}(p) u^i v^j.$$

Now consider a smooth function  $f : M \rightarrow \mathbb{R}$  and the real-valued function  $T_pM \ni v \mapsto f_p(v) := f(\exp_p v)$  defined around 0 in  $T_pM$ .

It is easy to see that

$$\frac{\partial f_p}{\partial x^i}(0) = \frac{\partial f}{\partial x^i}(p), \quad \frac{\partial^2 f_p}{\partial x^i \partial x^j}(0) = \frac{\partial^2 f}{\partial x^i \partial x^j}(p).$$

The Taylor–Young formula (for Euclidean spaces) applied to  $f_p$  around the origin can be written using matrices as

$$f_p(v) = f_p(0) + J_{f_p}(0)v + \frac{1}{2}v^T H_{f_p}(0)v + o(\|v\|^2), \tag{1}$$

where

$$v = [v^1 \ \dots \ v^n]^T, \quad J_{f_p}(0) = \left[ \frac{\partial f}{\partial x^1}(p) \ \dots \ \frac{\partial f}{\partial x^n}(p) \right],$$

$$H_{f_p}(0) = \left( \frac{\partial^2 f}{\partial x^i \partial x^j}(p) \right).$$

In other words, we have the following Taylor–Young expansion for  $f$  around  $p$

$$f(\exp_p v) = f(p) + g_p(\text{grad } f, v) + \frac{1}{2}\text{Hess}_p f(v, v) + o(\|v\|_p^2) \tag{2}$$

which holds in any coordinate system.

A function  $f : M \rightarrow \mathbb{R}$  is said to be (strictly or strongly) convex if its restriction to any geodesic curve  $\gamma : [a, b] \rightarrow M$  is convex in the classical sense; that the one real variable function  $f \circ \gamma : [a, b] \rightarrow \mathbb{R}$  is (strictly) convex.

A function  $f$  on  $M$  is (strictly) convex if and only if its Hessian is (positive definite) positive.

For more detailed and complete information on the fundamentals in Riemannian geometry, we refer to [51, 52].

### 2.2 Basics in Vector Optimization on Riemannian Manifolds

Let  $(M, g)$  be a complete Riemannian manifold of dimension  $n$ ; hence, from Hopf–Rinow theorem, for any two points  $p$  and  $q$  in  $M$ , there exists a length minimizing geodesic connecting these two points.

Throughout the paper,  $C \subset \mathbb{R}^r$  will denote a convex and pointed cone (i.e.,  $\mathbb{R}_+ C \subset C$ ,  $C + C \subset C$ ,  $C \cap (-C) = \{0\}$ ). We suppose that  $C$  is closed and  $\text{int}(C) \neq \emptyset$ , where  $\text{int}(A)$  stands for the the topological interior of any subset  $A \subset \mathbb{R}^r$ .

For any  $y, y' \in \mathbb{R}^r$ , we denote

$$y \preceq y' \iff y' - y \in C;$$

$$y \prec y' \iff y' - y \in \text{int}(C); \quad y \not\preceq y' \iff y' - y \in C \setminus \{0\}.$$

Obviously,

$$y \prec y' \implies y \not\preceq y' \implies y \preceq y'.$$

Note that  $\preceq$  is a partial order relation on  $\mathbb{R}^r$ , i.e., a reflexive, antisymmetric and transitive binary relation. Also  $<$  and  $\succcurlyeq$  are transitive relations.

Consider a vector function  $G = (G_1, \dots, G_r) : M \rightarrow \mathbb{R}^r$ , and the multiobjective optimization problem<sup>1</sup>

$$(MOP) \quad \text{MIN}_C G(x) \quad \text{s.t. } x \in M.$$

For (MOP), the point  $a \in M$  is called:

- *Pareto solution* iff there is no  $x \in M$  such that  $G(x) \preceq G(a)$ ;
- *weakly Pareto solution* iff there is no  $x \in M$  such that  $G(x) < G(a)$ ;
- *properly Pareto solution* iff  $a$  is a Pareto solution, and there exists a pointed convex cone  $K$  such that  $C \setminus \{0\} \subset \text{int}(K)$  and  $a$  is a Pareto solution for the problem  $\text{MIN}_K G(x) \quad \text{s.t. } x \in M$ , i.e.,  $G(M) \cap (G(a) - K) = \{G(a)\}$ .

When  $C = \mathbb{R}_+^r := \{\lambda = (\lambda_1, \dots, \lambda_r) \in \mathbb{R}^r : \lambda_i \geq 0, i = 1, \dots, r\}$  (the Pareto cone), the previous definitions can be stated as follows.

- *Pareto solution* iff there is no  $x \in M$  such that, for all  $i \in \{1, \dots, r\}$ ,  $G_i(x) \leq G_i(a)$ , and  $G(x) \neq G(a)$ ;
- *Weakly Pareto solution* iff there is no  $x \in M$  such that, for all  $i \in \{1, \dots, r\}$ ,  $G_i(x) < G_i(a)$ ;
- *Properly Pareto solution* (provided that  $G(M) + \mathbb{R}_+^r$  is convex) iff  $a$  is a Pareto solution, and there exists a real number  $\mu > 0$  such that for each  $i \in \{1, \dots, r\}$  and every  $x \in M$  with  $G_i(x) < G_i(a)$  at least one  $j \in \{1, \dots, r\}$  exists with  $G_j(x) > G_j(a)$  and

$$\frac{G_i(a) - G_i(x)}{G_j(x) - G_j(a)} \leq \mu.$$

We denote the set of all Pareto (resp. weakly Pareto and properly Pareto) solutions by  $\text{ARGMIN}_{x \in M} G(x)$  (resp.  $w\text{-ARGMIN}_{x \in M} G(x)$  and  $p\text{-ARGMIN}_{x \in M} G(x)$ ). Obviously, we have

$$p\text{-ARGMIN}_{x \in M} G(x) \subset \text{ARGMIN}_{x \in M} G(x) \subset w\text{-ARGMIN}_{x \in M} G(x).$$

In order to simplify the notations, we will use the symbol  $\sigma \in \{w, p\}$  for *weak* (if  $\sigma = w$ ) or *proper* (if  $\sigma = p$ ), i.e., we write  $\sigma\text{-ARGMIN}_{x \in M} G(x)$  for the weakly or properly Pareto solutions set.

We recall that a real function  $h : M \rightarrow \mathbb{R}$  is called *convex* if for any two distinct points  $a$  and  $b$  in  $M$ , and for any geodesic segment  $\gamma : [0, 1] \rightarrow M$  with  $\gamma(0) = a$ ,  $\gamma(1) = b$ , the function  $h \circ \gamma$  is convex in the usual way, i.e., for all  $t \in ]0, 1[$

$$h(\gamma(t)) \leq (1 - t)h(a) + th(b).$$

<sup>1</sup> We use the notations ‘‘MIN’’ and ‘‘ARGMIN’’ for vector-valued functions to distinguish from the notations ‘‘min’’ and ‘‘argmin’’ used for scalar-valued functions.



If the last inequality is strict, we say that  $h$  is *strictly convex*.

Now, the *vector-valued function*  $G = (G_1, \dots, G_r) : M \rightarrow \mathbb{R}^r$  is called *C-convex* (resp. *w*-strictly *C*-convex or *p*-strictly *C*-convex) iff for any two distinct points  $a$  and  $b$  in  $M$ , and for any geodesic segment  $\gamma : [0, 1] \rightarrow M$  with  $\gamma(0) = a, \gamma(1) = b$ , we have, respectively,

$$\forall t \in ]0, 1[ \quad G(\gamma(t)) \preceq (1 - t)G(a) + tG(b), \tag{3}$$

$$\forall t \in ]0, 1[ \quad G(\gamma(t)) < (1 - t)G(a) + tG(b), \tag{4}$$

$$\forall t \in ]0, 1[ \quad G(\gamma(t)) \not\preceq (1 - t)G(a) + tG(b). \tag{5}$$

In the case  $C = \mathbb{R}_+^r$ , it is easy to see that

- $G$  is  $\mathbb{R}_+^r$ -convex iff  $G_i$  is convex for all  $i = 1, \dots, r$ ;
- $G$  is *w*-strictly  $\mathbb{R}_+^r$ -convex iff  $G_i$  is strictly convex for all  $i = 1, \dots, r$ ;
- $G$  is *p*-strictly  $\mathbb{R}_+^r$ -convex if  $G$  is  $\mathbb{R}_+^r$ -convex, and there exists  $i \in \{1, \dots, r\}$  such that  $G_i$  is strictly convex.

Throughout the paper,  $\mathbb{R}^r$  is considered with its usual Euclidean structure and identified to its dual space, and we denote by  $\langle \cdot, \cdot \rangle$  its usual inner product (which coincides with the duality product with our identification) and by  $\| \cdot \|$  the induced norm.

The *dual cone of C* (or positive polar cone) is the set

$$C^* := \{ \lambda \in \mathbb{R}^r : \langle \lambda, y \rangle \geq 0 \quad \forall y \in C \},$$

and its *quasi-interior* is given by

$$C_{\sharp}^* := \{ \lambda \in \mathbb{R}^r : \langle \lambda, y \rangle > 0 \quad \forall y \in C \setminus \{0\} \}.$$

Note that

$$(\mathbb{R}_+^r)^* = \mathbb{R}_+^r, \quad \text{and } (\mathbb{R}_+^r)_{\sharp}^* = \text{int } (\mathbb{R}_+^r) = \{ \lambda \in \mathbb{R}^r \mid \lambda_i > 0 \quad i = 1, \dots, r \}.$$

Let us denote

$$\Lambda_{\sigma} = \begin{cases} \{ \lambda \in C^* : \|\lambda\|_1 = 1 \}, & \text{if } \sigma = w, \\ C_{\sharp}^*, & \text{if } \sigma = p. \end{cases} \tag{6}$$

**Proposition 2.1** *The following properties hold.*

- A. *The dual cone  $C^*$  is a closed set in  $\mathbb{R}^r$ .*
- B. *The set  $C_{\sharp}^*$  (the quasi-interior of  $C^*$ ) is a nonempty open set,<sup>2</sup> and it is in fact the topological interior of  $C^*$ .*
- C. *The set  $\Lambda_w$  is compact.*

---

<sup>2</sup> This fact is not true in general, i.e., when  $C$  is a cone in a topological vector space, but in our setting, we take advantage of the finite dimension of  $\mathbb{R}^r$ .

*Proof* A. We can write  $C^*$  as an intersection of a family of closed sets:

$$C^* = \bigcap_{y \in C} \{\lambda \in \mathbb{R}^r : \langle \lambda, y \rangle \geq 0\}.$$

B. Let  $\lambda_0 \in C_\#^*$ . Let  $K = C \cap \{y \in \mathbb{R}^r : \|y\| = 1\}$ . Obviously,  $K$  is a compact set; hence, by Weierstrass’ theorem, there is some real  $\alpha > 0$  such that  $\langle \lambda_0, y \rangle \geq \alpha$  for all  $y \in K$ . Let  $R = \alpha/2$ . For any  $\lambda \in B(\lambda_0; R)$  (the open ball centered in  $\lambda_0$  and of radius  $r$ ), and for any  $y \in C \setminus \{0\}$ , by Cauchy–Schwarz inequality, we have

$$\left\langle \lambda, \frac{y}{\|y\|} \right\rangle = \left\langle \lambda_0, \frac{y}{\|y\|} \right\rangle + \left\langle \lambda - \lambda_0, \frac{y}{\|y\|} \right\rangle \geq \alpha - \|\lambda - \lambda_0\| \geq \alpha - R > 0,$$

hence  $\langle \lambda, y \rangle > 0$ ; therefore  $B(\lambda_0; R) \subset C_\#^*$ . Thus,  $C_\#^*$  is an open set included in  $C^*$ . Let now  $\lambda_0 \in \text{int}(C^*)$ . Arguing by contradiction, if  $\lambda_0 \notin C_\#^*$ , then  $\langle \lambda_0, y \rangle = 0$  for some  $y \in C \setminus \{0\}$ . There exists some real  $R > 0$  such that  $\lambda_0 + h \in C^*$  for all  $h \in \mathbb{R}^r$  with  $\|h\| < R$ . Therefore,

$$0 \leq \langle \lambda_0 + h, y \rangle = \langle h, y \rangle \quad \forall h \in \mathbb{R}^r, \|h\| < R,$$

which implies the contradiction  $y = 0$ .

Finally,  $C_\#^*$  is nonempty because  $C \setminus \{0\}$  is contained in an open half-space (using a separation theorem since  $C$  is closed convex and pointed).

C. It is obvious. □

Now we give a scalarization theorem on manifolds, which allows us to replace the vector optimization problem with a family of scalar optimization problems.

**Theorem 2.2** *Let  $(M, g)$  be a complete and connected Riemannian manifold in  $\mathbb{R}^n$ , and consider a function  $G = (G_1, \dots, G_r) : M \rightarrow \mathbb{R}^r$ . Then, for each  $\sigma \in \{w, p\}$ , we have*

$$\bigcup_{\lambda \in \Lambda_\sigma} \arg \min_{x \in M} \langle \lambda, G(x) \rangle \subset \sigma\text{-ARGMIN}_C G(x) \tag{7}$$

Moreover, if  $G$  is  $C$ -convex on  $M$ , then the previous inclusion becomes an equality, i.e.,

$$\sigma\text{-ARGMIN}_C G(x) = \bigcup_{\lambda \in \Lambda_\sigma} \arg \min_{x \in M} \langle \lambda, G(x) \rangle. \tag{8}$$

*Proof* Inclusion (7) is a standard result (see, e.g., [46,49]). However, for reader’s convenience, we give a short proof here.

Let  $\sigma = p$ ,  $\lambda \in \Lambda_p$ , and  $a \in \arg \min_{x \in M} \langle \lambda, G(x) \rangle$ . Let us consider the set  $K = \{y \in \mathbb{R}^r \mid \langle \lambda, y \rangle > 0\} \cup \{0\}$ . Obviously,  $K$  is pointed and convex cone, and  $C \setminus \{0\} \subset \text{int}(K)$ . We claim that  $a \in \arg \min_{x \in M} \langle \lambda, G(x) \rangle$ . Indeed, arguing by contradiction, if  $a \notin \arg \min_{x \in M} \langle \lambda, G(x) \rangle$ , then there is some  $x \in M$  such that  $G(a) - G(x) \in K \setminus \{0\}$ ; hence,  $\langle \lambda, G(a) - G(x) \rangle > 0$  which contradicts the choice of  $a$ . Therefore,  $a \in p\text{-ARGMIN}_C G(x)$ .

Let now  $\sigma = w, \lambda \in \Lambda_w$ , and  $a \in \arg \min_{x \in M} \langle \lambda, G(x) \rangle$ . Since  $\lambda \in C^* \setminus \{0\}$ , it is easy to see that  $\lambda(\text{int}(C)) \subset ]0, +\infty[$ . If, by contradiction  $a \notin w\text{-ARGMIN}_C G(x)$ , then we can find  $x \in M$  such that  $G(x) \prec G(a)$ , i.e.,  $G(a) - G(x) \in \text{int}(C)$ ; therefore,  $\langle \lambda, G(a) - G(x) \rangle > 0$  which contradicts the choice of  $a$ . Thus, inclusion (7) holds.

Suppose now that  $G$  is  $C$ -convex. Let us prove that

$$G(M) + C \text{ is convex in } \mathbb{R}^r. \tag{9}$$

Let  $y, y' \in G(M) + C$ , thus  $y = G(x) + c, y' = G(x') + c'$  for some  $x, x' \in M$  and  $c, c' \in C$ . There is a geodesic segment  $\gamma : [0, 1] \rightarrow M$  with  $\gamma(0) = x, \gamma(1) = x'$ . Then, for each  $t \in ]0, 1[$ , we have  $G(\gamma(t)) \preceq (1 - t)G(x) + tG(x')$ , thus

$$c'' := (1 - t)(y - c) + t(y' - c') - G(\gamma(t)) \in C.$$

Therefore,

$$(1 - t)y + ty' = G(\gamma(t)) + (1 - t)c + tc' + c'' \in G(M) + C,$$

hence (9) holds.

Let  $\sigma = w, a \in w\text{-ARGMIN}_C G(x)$ . Thus,  $(G(a) - \text{int}(C)) \cap G(M) = \emptyset$ . Since  $\text{int}(C) + C \subset \text{int}(C)$ , it follows that the convex sets  $(G(a) - \text{int}(C))$  and  $G(M) + C$  are disjoint. Using the separation theorem, we can find  $\lambda \in \mathbb{R}^r \setminus \{0\}$  such that

$$\forall x \in M, c \in \text{int}(C), c' \in C \quad \langle \lambda, G(a) - c \rangle \leq \langle \lambda, G(x) + c' \rangle. \tag{10}$$

Taking above  $x = a, c' = 0$ , we obtain that  $\langle \lambda, c \rangle \geq 0$  for all  $c \in \text{int}(C)$ , hence for all  $c \in C$ , because  $\text{cl}(\text{int}(C)) = \text{cl}(C)$ , where  $\text{cl}(A)$  stands for the topological closure of a set  $A$ . Thus,  $\lambda \in C^* \setminus \{0\}$ . Finally,  $\langle \lambda, G(a) \rangle \leq \langle \lambda, G(x) \rangle$  for all  $x \in M$ , and replacing without loss of generality  $\lambda$  by  $\frac{\lambda}{\|\lambda\|_1}$  we find that  $a \in \arg \min_{x \in M} \langle \lambda, G(x) \rangle$  with  $\lambda \in \Lambda_w$ .

Take now  $\sigma = p$ . Let  $a \in p\text{-ARGMIN}_C G(x)$ . There exists a pointed convex cone  $K$  such that  $C \setminus \{0\} \subset \text{int}(C)$ , and  $a \in \text{ARGMIN}_K G(x)$ . Obviously,  $a \in w\text{-ARGMIN}_K G(x)$ . Using the case  $\sigma = w$  with  $K$  instead of  $C$ , we can find  $\lambda \in K^* \setminus \{0\}$  such that  $a \in \arg \min_{x \in M} \langle \lambda, G(x) \rangle$ . Since  $C \setminus \{0\} \subset \text{int}(K)$ , and  $\lambda \neq 0$ , we have  $\langle \lambda, c \rangle > 0$  for all  $c \in C \setminus \{0\}$ , hence  $\lambda \in C_{\sharp}^*$ .  $\square$

For more details about vector optimization in Euclidean spaces see, e.g., the monographs [45–47,49].

### 3 The Semivectorial Bilevel Problem

#### 3.1 Statement of the Problem

Let  $(M_1, g_1)$  (the leader decision variables set) and  $(M_2, g_2)$  (the follower decision variables set) be two connected Riemannian manifolds of dimension  $m$  and  $n$ , respec-

tively. Moreover,  $(M_2, g_2)$  is supposed to be complete. The corresponding Riemannian metrics will be denoted by  $g_1(\cdot, \cdot)$  and  $g_2(\cdot, \cdot)$ , respectively. Let also  $f : M_1 \times M_2 \rightarrow \mathbb{R}$  be the leader objective function, and let  $F = (F_1, \dots, F_r) : M_1 \times M_2 \rightarrow \mathbb{R}^r$  be the follower multiobjective function.

For each  $x \in M_1$ , we denote by  $\psi(x)$  the *weakly or properly Pareto solution set of the follower multiobjective optimization problem*, i.e.,

$$\psi(x) := \sigma\text{-ARGMIN}_{y \in M_2} C F(x, y)$$

Thus,  $\psi : M_1 \rightrightarrows M_2$  is a set-valued function.

We deal with two semivectorial bilevel problems.

- The “*optimistic semivectorial bilevel problem*”

$$(OSB) \quad \min_{x \in M_1} \min_{y \in \psi(x)} f(x, y).$$

In this case, the follower cooperates with the leader; i.e., for each  $x \in M_1$ , the follower chooses among all its  $\sigma$ -Pareto solutions (his best responses) one which is the best for the leader (assuming that such a solution exists).

- The “*pessimistic semivectorial bilevel problem*”

$$(PSB) \quad \min_{x \in M_1} \sup_{y \in \psi(x)} f(x, y).$$

In this case, there is no cooperation between the leader and the follower, and the leader expects the worst scenario; i.e., for each  $x \in M_1$ , the follower may choose among all its  $\sigma$ -Pareto solutions (his best responses) one which is unfavorable for the leader (in this case we prefer to use “sup” instead of “max”).

Consider the following hypotheses:

- (HC) $_{\sigma}$  For each  $x \in M_1$ , the function  $F(x, \cdot)$  is  $\sigma$ -strictly  $C$ -convex on  $M_2$ ,  $\sigma \in \{w, p\}$ .
- (HCC) $_{\sigma}$  For all  $x \in M_1$  and  $\lambda \in \Lambda_{\sigma}$ , the function  $y \mapsto \langle \lambda, F(x, y) \rangle$  has bounded sublevel sets, i.e, for all reals  $\alpha$ , the set

$$\{y \in M_2 : \langle \lambda, F(x, y) \rangle \leq \alpha\}$$

is bounded.

### 3.2 A Useful Equivalent Form

**Proposition 3.1** *If  $(HC)_{\sigma}$  holds, then, for any  $x \in M_1$  and any  $\lambda \in \Lambda_{\sigma}$ , where  $\Lambda_{\sigma}$  is given in (6), the real-valued function  $M_2 \ni y \mapsto \langle \lambda, F(x, y) \rangle$  is strictly convex.*

*Proof* Let  $a, b \in M_2$   $a \neq b$ , and  $\gamma : [0, 1] \rightarrow M_2$  be the geodesic segment with  $\gamma(0) = a, \gamma(1) = b$ . Let  $t \in ]0, 1[$ .

If  $\sigma = w$ , then  $\lambda \in C^* \setminus \{0\}$ , and  $F(x, \gamma(t)) < (1 - t)F(x, a) + tF(x, b)$ , i.e.,  $(1 - t)F(x, a) + tF(x, b) - F(x, \gamma(t)) \in \text{int}(C)$ , hence

$$\langle \lambda, (1 - t)F(x, a) + tF(x, b) - F(x, \gamma(t)) \rangle > 0.$$

If  $\sigma = p$ , then  $\lambda \in C_\#^*$ , and  $F(x, \gamma(t)) \not\leq (1 - t)F(x, a) + tF(x, b)$ , i.e.,  $(1 - t)F(x, a) + tF(x, b) - F(x, \gamma(t)) \in C \setminus \{0\}$ , hence

$$\langle \lambda, (1 - t)F(x, a) + tF(x, b) - F(x, \gamma(t)) \rangle > 0.$$

□

**Proposition 3.2** *Assume that  $(HC)_w$  holds. Then the set of weakly Pareto solutions coincides with the set of Pareto solutions, i.e., for each  $x \in M_1$ ,*

$$w\text{-ARGMIN}_C F(x, y) = \text{ARGMIN}_C F(x, y)$$

*Proof* Let  $x \in M_1$  be fixed, and let  $\hat{y} \in w\text{-ARGMIN}_C F(x, y)$ . Theorem 2.2 and Proposition 3.1 imply that there exists  $\lambda \in \Lambda_w$  such that  $\hat{y}$  is a unique global minimizer of the function  $M_2 \ni y \mapsto \langle \lambda, F(x, y) \rangle$ . If  $\hat{y} \notin \text{ARGMIN}_C F(x, y)$ , then we can find  $\bar{y} \in M_2$  such that  $F(x, \bar{y}) \not\leq F(x, \hat{y})$ . It follows that  $\bar{y} \neq \hat{y}$  and  $\langle \lambda, F(x, \bar{y}) \rangle \leq \langle \lambda, F(x, \hat{y}) \rangle$  which contradicts the fact that  $\hat{y}$  is the unique minimizer of  $\langle \lambda, F(x, \cdot) \rangle$ . □

From now on, we assume for the rest of the paper that  $(HC)_\sigma$  and  $(HCC)_\sigma$  are satisfied.

**Proposition 3.3** *For each  $x \in M_1$  and  $\lambda \in \Lambda_\sigma$ , the minimization problem*

$$\min_{y \in M_2} \langle \lambda, F(x, y) \rangle$$

*admits a unique solution, which will be denoted hereafter  $y(x, \lambda)$ .*

*Proof* Uniqueness is provided by Proposition 3.1. Since any convex function on a Riemannian manifold is continuous, and closed bounded sets on a complete and connected Riemannian manifold are compact, by the hypothesis  $(HCC)_\sigma$  and Weierstrass’ theorem, we can find a minimizer over a nonempty sublevel set of the function  $y \mapsto \langle \lambda, F(x, y) \rangle$  which in fact is a minimizer over  $M_2$ . □

Thus, we obviously have the following.

**Corollary 3.1** *For each fixed  $x \in M_1$ , the map  $\lambda \mapsto y(x, \lambda)$  is a surjection from  $\Lambda_\sigma$  to  $\psi(x)$ , hence*

$$\psi(x) = \bigcup_{\lambda \in \Lambda_\sigma} \{y(x, \lambda)\}. \tag{11}$$

Finally, we obtain the following equivalent form with scalar bilevel problems where the lower level has a unique response.

**Theorem 3.1** *Problem (OSB) is equivalent to the following problem*

$$\begin{aligned} & \min_{x \in M_1} \min_{\lambda \in \Lambda_\sigma} f(x, y(x, \lambda)), \\ & \text{where } y(x, \lambda) \text{ is the unique solution to the problem} \\ & \min_{y \in M_2} \langle \lambda, F(x, y) \rangle. \end{aligned} \tag{12}$$

*Problem (PSB) is equivalent to the following problem*

$$\begin{aligned} & \min_{x \in M_1} \sup_{\lambda \in \Lambda_\sigma} f(x, y(x, \lambda)), \\ & \text{where } y(x, \lambda) \text{ is the unique solution to the problem} \\ & \min_{y \in M_2} \langle \lambda, F(x, y) \rangle. \end{aligned} \tag{13}$$

### 3.3 Optimality Conditions for the Problem $\min_{y \in M_2} \langle \lambda, F(x, y) \rangle$

From now on, we suppose that  $F(\cdot, \cdot)$  is a smooth function.

Let us denote by  $\text{grad}_i$  the gradient operator on  $(M_i, g_i)$ ,  $i = 1, 2$ , and by  $F^a$  (resp.  $\lambda_a$ ),  $a = 1, \dots, r$ , the components functions of the map  $F : M_1 \times M_2 \rightarrow \mathbb{R}^r$  (resp. the canonical coordinates of the vector  $\lambda \in \mathbb{R}^r$ ).

The proof of the following result is well known.

**Proposition 3.4** (Necessary and sufficient conditions for  $y(x, \lambda)$ ) *Let  $\lambda \in \Lambda_\sigma$  and  $x \in M_1$  be given. Then  $y = y(x, \lambda)$  is the unique solution of the problem  $\min_{y \in M_2} \langle \lambda, F(x, y) \rangle$  iff (if and only if)*

$$\lambda_a \text{ grad}_2 F^a(x, y) = 0. \tag{14}$$

□

Now let us consider the map

$$G : \mathbb{R}^r \times M_1 \times M_2 \rightarrow TM_2, \quad G(\lambda, x, y) = \lambda_a \text{ grad}_2 F^a(x, y).$$

Note that, for each  $(\lambda, x) \in \Lambda_\sigma \times M_1$ , the solution  $y = y(x, \lambda)$  to the problem  $\min_{y \in M_2} \langle \lambda, F(x, y) \rangle$  satisfies the equation  $G(\lambda, x, y) = 0$ .

Denote by  $\delta_2 G(\lambda, x, y) : T_y M_2 \rightarrow T_y M_2$  the partial differential of  $G$  w.r.t. to  $y$  at the point  $(\lambda, x, y)$ .

The following is a direct consequence of the implicit functions theorem.

**Proposition 3.5** *Let  $(\lambda_0, x_0) \in \Lambda_\sigma \times M_1$ , and let  $y_0 = y(x_0, \lambda_0)$  be the unique solution of the problem  $\min_{y \in M_2} \langle \lambda_0, F(x_0, y) \rangle$ .*

*Suppose that  $\delta_2 G(\lambda_0, x_0, y_0)$  is an isomorphism.<sup>3</sup> Then, in a neighborhood of  $(\lambda_0, x_0)$ , the function  $y(\cdot, \cdot)$  is smooth and*

$$\frac{\partial}{\partial \lambda} y(\lambda, x) = -(\delta_2 G(\lambda, x, y))^{-1} \circ \frac{\partial G}{\partial \lambda}(\lambda, x, y)$$

and

$$\delta_1 y(\lambda, x) = -(\delta_2 G(\lambda, x, y))^{-1} \circ \delta_1 G(\lambda, x, y),$$

where  $\delta_1$  denotes the partial differential operator w.r.t.  $x \in M_1$ .

*Remark 3.1* Note that for a tangent vector  $u \in T_x M_1$ , we have

$$\delta_1 G(\lambda, x, y)(u) = (\delta_1 G^1(\lambda, x, y)(u), \dots, \delta_1 G^n(\lambda, x, y)(u)),$$

with

$$\delta_1 G^\alpha(\lambda, x, y)(u) = g_1(\text{grad}_1 G^\alpha(\lambda, x, y), u),$$

and for a tangent vector  $v \in T_y M_2$ , we have

$$\delta_2 G(\lambda, x, y)(v) = (d_2 G^1(\lambda, x, y)(v), \dots, d_2 G^n(\lambda, x, y)(v)),$$

with

$$d_2 G^\alpha(\lambda, x, y)(v) = g_2(\text{grad}_2 G^\alpha(\lambda, x, y), v),$$

where  $G^\alpha, \alpha = 1, \dots, n$ , are the component functions of  $G$ .

In local coordinates  $(x^i)_{i=1}^m$  on  $M_1$  and  $(y^\alpha)_{\alpha=1}^n$  on  $M_2$ , we can write

$$\text{grad}_1 G^\alpha(\lambda, x, y) = g_1^{ij} \frac{\partial G^\alpha}{\partial x^i}(\lambda, x, y) \frac{\partial}{\partial x^j},$$

and

$$\text{grad}_2 G^\alpha(\lambda, x, y) = g_2^{\beta\mu} \frac{\partial G^\alpha}{\partial y^\beta}(\lambda, x, y) \frac{\partial}{\partial y^\mu}.$$

### 3.4 Optimality Conditions for the Optimistic Problem

In the sequel, we denote for all  $x \in M_1$  and  $\lambda \in \Lambda_\sigma$

$$\varphi(x, \lambda) = f(x, y(x, \lambda)).$$

In this subsection, we suppose that  $f(\cdot, \cdot)$  is a smooth function.

<sup>3</sup> This hypothesis holds, for example, if we assume that  $G$  is a conformal map, or, for example, if there exists a real number  $c > 0$  such that  $g_2(\delta_2 G(\lambda_0, x_0, y_0)(v), v) \geq c g_2(v, v), \forall v \in T_y M_2$ .

**Theorem 3.2** (Necessary optimality conditions for the optimistic problem when we deal with properly Pareto solutions) *Let  $(x^*, \lambda^*) \in M_1 \times \Lambda_p$  be a (local<sup>4</sup>) solution of the problem*

$$\min_{x \in M_1} \min_{\lambda \in \Lambda_p} f(x, y(x, \lambda)).$$

Let  $y^* = y(x^*, \lambda^*)$ . Then

$$\text{grad}_1 f(x^*, y^*) + \text{grad}_2 f(x^*, y^*) \circ \delta_1 y(x^*, \lambda^*) = 0$$

$$\text{grad}_2 f(x^*, y^*) \circ \frac{\partial y}{\partial \lambda}(x^*, \lambda^*) = 0,$$

where  $\delta_1 y(x^*, \lambda^*)$  and  $\frac{\partial y}{\partial \lambda}(x^*, \lambda^*)$  are given in Proposition 3.5.

Moreover, the point  $(x^*, \lambda^*)$  has the property that the quadratic form associated with the bilinear form  $\text{Hess}(\varphi)(x^*, \lambda^*)$  is positive semidefinite.

*Proof* Note that our problem written in short  $\min_{x \in M_1} \min_{\lambda \in \Lambda_p} \varphi(x, \lambda)$  is (locally) equivalent to the problem

$$\min_{(x, \lambda) \in M_1 \times \Lambda_p} \varphi(x, \lambda),$$

and the set  $M_1 \times \Lambda_p$  is open in  $M_1 \times \mathbb{R}^r$  according to Proposition 2.1. Then, using Fermat’s rule, chain rule and the necessary second-order optimality conditions (which can be justified as in the Euclidean case using Taylor–Young formula (2) around  $(x^*, \lambda^*)$ ), we obtain easily the result. □

**Theorem 3.3** (Necessary optimality conditions for the optimistic problem when we deal with (weakly) Pareto solutions) *Let  $(x^*, \lambda^*) \in M_1 \times \Lambda_w$  be a (local) solution of the problem*

$$\min_{x \in M_1} \min_{\lambda \in \Lambda_w} f(x, y(x, \lambda)).$$

Let  $y^* = y(x^*, \lambda^*)$ . Then

$$\text{grad}_1 f(x^*, y^*) + \text{grad}_2 f(x^*, y^*) \circ \delta_1 y(x^*, \lambda^*) = 0 \tag{15}$$

$$\text{grad}_2 f(x^*, y^*) \circ \frac{\partial y}{\partial \lambda}(x^*, \lambda^*) + N_{\Lambda_w}^C(\lambda^*) \ni 0, \tag{16}$$

where  $N_{\Lambda_w}^C(\lambda^*)$  is the Clarke normal cone to the set  $\Lambda_w$  at the point  $\lambda^*$  (see [53, page 212]), and  $\delta_1 y(x^*, \lambda^*)$ ,  $\frac{\partial y}{\partial \lambda}(x^*, \lambda^*)$  are given in Proposition 3.5.

---

<sup>4</sup> i.e., there exists an open neighborhood  $\mathcal{N} = \mathcal{N}_1 \times \mathcal{N}' \subset M_1 \times \Lambda_p$  of  $(x^*, \lambda^*)$  such that, for each  $x \in \mathcal{N}_1$ , the function  $\mathcal{N}' \ni \lambda \mapsto f(x, y(x, \lambda))$  admits local minimizers, and for all  $x \in \mathcal{N}_1$ ,  $f(x^*, \lambda^*) \leq \min_{\lambda \in \mathcal{N}'} f(x, y(x, \lambda))$ .



*Proof* Note that  $\Lambda_w$  is compact according to Proposition 2.1; hence for each fixed  $x \in M_1$ , the minimum  $\min_{\lambda \in \Lambda_w} \varphi(x, \lambda)$  is attained. Thus, our problem is equivalent to

$$\min_{(x, \lambda) \in M_1 \times \Lambda_w} \varphi(x, \lambda).$$

The function  $\varphi$  is smooth; hence, we can find an open and bounded neighborhood  $O_1 \times O_2$  of  $(x^*, \lambda^*)$  with  $\Lambda_w \subset O_2$  (because  $\Lambda_w$  is compact) and a positive real  $k$  such that  $\varphi$  is Lipschitz of rank  $k$  on  $O_1 \times O_2$ .

Thus,  $(x^*, \lambda^*)$  solves the unconstrained problem

$$\min_{(x, \lambda) \in O_1 \times O_2} (\varphi(x, \lambda) + kd_{\Lambda_w}(\lambda)),$$

where

$$d_{\Lambda_w}(\lambda) = \inf_{\mu \in \Lambda_w} \|\lambda - \mu\|$$

is the distance function associated with the set  $\Lambda_w$ . Then using Clarke’s subgradient calculus (see [53, Proposition 10.36]), we obtain the result.  $\square$

*Remark 3.2* In the particular and important case when  $C = \mathbb{R}_+^r$  (the Pareto cone), the set  $\Lambda_w$  is convex and closed, so the Clarke normal cone is given by (see [53, Theorem 10.39])

$$N_{\Lambda_w}^C(\lambda^*) = \{\alpha \in \mathbb{R}^r \mid \langle \alpha, \lambda - \lambda^* \rangle \leq 0 \quad \forall \lambda \in \Lambda_w\}.$$

Then it is easy to see that,

$$N_{\Lambda_w}^C(\lambda^*) = \{\alpha \in \mathbb{R}^r \mid \alpha_i = \alpha_j \quad \forall i, j \in I_+(\lambda^*); \alpha_k \leq \alpha_i \quad \forall (k, i) \in I_0(\lambda^*) \times I_+(\lambda^*)\},$$

where  $I_+(\lambda^*) = \{i \mid \lambda_i^* > 0\}$  and  $I_0(\lambda^*) = \{k \mid \lambda_k^* = 0\}$ . Therefore, (16) becomes:  
*there exists a real  $v$  such that*

$$\text{grad}_2 f(x^*, y^*) \circ \frac{\partial y}{\partial \lambda_i}(x^*, \lambda^*) = v \quad \forall i \in I_+(\lambda^*) \tag{17}$$

$$\text{grad}_2 f(x^*, y^*) \circ \frac{\partial y}{\partial \lambda_k}(x^*, \lambda^*) \geq v \quad \forall k \in I_0(\lambda^*). \tag{18}$$

These relations can also be obtained by using Karush–Kuhn–Tucker Theorem.

### 3.5 An Existence Result for the Pessimistic Problem

For the more difficult case of the pessimistic problem, we will deal only with (weakly) Pareto solutions.

**Theorem 3.4** *Suppose that the function  $(x, y) \mapsto f(x, y)$  is continuous on  $M_1 \times M_2$  and sublevel bounded in  $x$  uniformly in  $y$ , i.e., there exists a bounded set  $B \subset M_1$  such that for each  $\alpha \in \mathbb{R}$ ,  $\{x \in M_1 : f(x, y) \leq \alpha\} \subset B$  for all  $y \in M_2$  (see [54, Definition 1.16]). Then the pessimistic problem*

$$\min_{x \in M_1} \sup_{\lambda \in \Lambda_w} f(x, y(x, \lambda))$$

has at least one global solution.

*Proof* According to Proposition 2.1, the set  $\Lambda_w$  is compact, and, by Proposition 3.5, the function  $(x, \lambda) \mapsto f(x, y(x, \lambda))$  is continuous; hence for each fixed  $x$ , the supremum  $\sup_{\lambda \in \Lambda_w} f(x, y(x, \lambda))$  is finite and attained by Weierstrass’ theorem.

On the other hand, the function  $x \mapsto \sup_{\lambda \in \Lambda_w} f(x, y(x, \lambda))$  is lower semicontinuous as a supremum of a family of (lower semi)continuous functions. Moreover, the function  $x \mapsto \sup_{\lambda \in \Lambda_w} f(x, y(x, \lambda))$  has bounded sublevel sets by the hypothesis about  $f$ ; hence, by Rockafellar [54, Theorem 1.9], we obtain the conclusion.  $\square$

### 3.6 An Illustrative Example

With the notations used in the Sect. 3.1, we consider in this subsection the particular case  $C = \mathbb{R}_+^2$ ,  $M_1 = \mathbb{R}_{++} := ]0, +\infty[$  with the Euclidean metric, and  $M_2 = \mathbb{R}_{++}^2 := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 > 0, y_2 > 0\}$  with the metric  $g_2$  given in Cartesian coordinates  $(y_1, y_2)$  around the point  $y \in M_2$  by the matrix

$$M_2 \ni y \mapsto (g_{ij})_y = \left( g_2 \left( \frac{\partial}{\partial y_i}, \frac{\partial}{\partial y_j} \right) \right) := \text{diag} \left( y_1^{-2}, y_2^{-2} \right).$$

In other words, for any vectors  $u = (u_1, u_2)$  and  $v = (v_1, v_2)$  in the tangent plane at  $y \in M_2$ , denoted  $T_y M_2$ , which coincides with  $\mathbb{R}^2$ , we have

$$g_2(u, v) = \frac{u_1 v_1}{y_1^2} + \frac{u_2 v_2}{y_2^2}.$$

Let  $a = (a_1, a_2) \in M_2$  and  $v = (v_1, v_2) \in T_a M_2$ . It is easy to see that the (minimizing) geodesic curve  $t \mapsto \gamma(t)$  verifying  $\gamma(0) = a$ ,  $\dot{\gamma}(0) = v$  is given by

$$\mathbb{R} \ni t \mapsto \left( a_1 e^{\frac{v_1}{a_1} t}, a_2 e^{\frac{v_2}{a_2} t} \right).$$

Hence,  $M_2$  is a complete Riemannian manifold. Also, the (minimizing) geodesic segment  $\gamma : [0, 1] \rightarrow M_2$  joining the points  $a = (a_1, a_2)$  and  $b = (b_1, b_2)$ , i.e.,  $\gamma(0) = a$ ,  $\gamma(1) = b$  is given by  $\gamma_i(t) = a_i^{1-t} b_i^t, i = 1, 2$ . Thus, the distance  $d$  on the metric space  $(M_2, g_2)$

$$\begin{aligned}
 d(a, b) &= \int_0^1 \|\dot{\gamma}(t)\|_{\gamma(t)} dt = \int_0^1 \sqrt{\left(\frac{\dot{\gamma}_1(t)}{\gamma_1(t)}\right)^2 + \left(\frac{\dot{\gamma}_2(t)}{\gamma_2(t)}\right)^2} dt \\
 &= \sqrt{\left(\ln \frac{a_1}{b_1}\right)^2 + \left(\ln \frac{a_2}{b_2}\right)^2}.
 \end{aligned}$$

For more details about this particular Riemannian manifold, see for instance [38, 39, 42].

It follows easily that the closed ball  $B(a; R)$  centered in  $a \in M_2$  of radius  $R \geq 0$  verifies

$$[a_1 e^{-\frac{R}{\sqrt{2}}}, a_1 e^{\frac{R}{\sqrt{2}}}] \times [a_2 e^{-\frac{R}{\sqrt{2}}}, a_2 e^{\frac{R}{\sqrt{2}}}] \subset B(a; R), \tag{19}$$

thus every closed rectangle  $[\rho_1, \eta_1] \times [\rho_2, \eta_2]$  ( $\rho_1 > 0, \rho_2 > 0$ ) is bounded in the metric space  $(M_2, g_2)$  with the distance  $d$ .

Consider now the functions  $F : M_1 \times M_2 \rightarrow \mathbb{R}^2$  and  $f : M_1 \times M_2 \rightarrow \mathbb{R}$  given for any  $(x, y) \in M_1 \times M_2$  by

$$\begin{aligned}
 F(x, y) &= (F_1(x, y), F_2(x, y)) = (\ln^2(y_1) + x\sqrt{y_2}, x \ln(y_1) - \ln(y_2)), \\
 f(x, y) &= \frac{1}{x} \ln(y_1) - x - \frac{1}{\sqrt{y_2}}(1 + y_2) \ln(y_1).
 \end{aligned}$$

Let us solve the optimistic semivectorial bilevel optimization problem in the properly Pareto case

$$(OSB)_p \quad \min_{x \in M_1} \min_{y \in \psi(x)} f(x, y),$$

where, for each  $x \in M_1$ ,

$$\psi(x) := p\text{-ARGMIN}_{y \in M_2} F(x, y)$$

It is obvious that for each fixed  $x \in M_1$ , none of the functions  $F_1(x, \cdot)$  and  $F_2(x, \cdot)$  is convex on  $M_2$  with the Euclidean metric. Therefore, the vector function  $F(x, \cdot)$  is not  $C$ -convex on  $\mathbb{R}_{++}^2$  with the Euclidean metric.

It is easy to see that, for any geodesic segment  $\gamma : [0, 1] \rightarrow M_2$  with  $\gamma(0) = a, \gamma(1) = b$ , the functions  $F_i(x, \cdot) \circ \gamma : [0, 1] \rightarrow \mathbb{R}$  are convex ( $i = 1, 2$ ). Moreover, the function  $F_1(x, \cdot) \circ \gamma$  is strictly convex; hence, the function  $F(x, \cdot)$  is  $p$ -strictly  $C$ -convex on the Riemannian manifold  $(M_2, g_2)$  (see Sect. 2.2). So the hypothesis  $(HC)_p$  holds.

On the other hand,  $\Lambda_p = \text{int}(C^*) = \mathbb{R}_{++}^2$ , so for every  $\lambda = (\lambda_1, \lambda_2) \in \Lambda_p$  and  $(x, y) \in M_1 \times M_2$ , we have

$$\langle \lambda, F(x, y) \rangle = \lambda_1 F_1(x, y) + \lambda_2 F_2(x, y).$$

Let  $\alpha \in \mathbb{R}$ ,  $x \in M_1$  and  $\lambda \in \Lambda_p$  be fixed and consider the sublevel set

$$A := \{y \in M_2 \mid \langle \lambda, F(x, y) \rangle \leq \alpha\}.$$

The inequality  $\langle \lambda, F(x, y) \rangle \leq \alpha$  can be written  $\phi_1(y_1) + \phi_2(y_2) \leq \alpha$ , with

$$\phi_1(y_1) := \lambda_1 \ln^2(y_1) + \lambda_2 x \ln(y_1); \quad \phi_2(y_2) = \lambda_1 x \sqrt{y_2} - \lambda_2 \ln(y_2).$$

A simple computation shows that the variations of  $\phi_1$  and  $\phi_2$  are given by

$y$	0	$\xi_1$	$+\infty$	$y$	0	$\xi_2$	$+\infty$
$\phi'_1(y)$	–	0	+	$\phi'_2(y)$	–	0	+
$\phi_1(y)$	$+\infty \searrow$	$\mu_1$	$\nearrow +\infty$	$\phi_2(y)$	$+\infty \searrow$	$\mu_2$	$\nearrow +\infty$

where  $\xi_1 = e^{-\frac{\lambda_2 x}{2\lambda_1}}$ ,  $\mu_1 = \phi_1(\xi_1)$ ,  $\xi_2 = \frac{4\lambda_2^2}{\lambda_1^2 x^2}$ ,  $\mu_2 = \phi_2(\xi_2)$ . Thus, if  $y \in A$ , then  $\phi_1(y_1) \leq \alpha - \mu_2$  and  $\phi_2(y_2) \leq \alpha - \mu_1$ . Hence, if  $\alpha < \mu_1 + \mu_2$ , the set  $A$  is empty; otherwise,  $A \subset [\eta_1, \nu_1] \times [\eta_2, \nu_2]$  with  $0 < \eta_i \leq \xi_i \leq \nu_i$ ,  $\phi_i(\eta_i) = \phi_i(\nu_i) = \alpha - \mu_j$ ,  $j = i + (-1)^{i-1}$ ,  $i = 1, 2$ . Hence,  $A$  is bounded with respect to the distance  $d$  of the metric space  $(M_2, g_2)$ . Therefore, the hypothesis  $(HCC)_p$  is satisfied. Notice also that the functions  $\phi_1$  and  $\phi_2$  are not convex on  $\mathbb{R}_{++}$  in the usual sense, but the scalar function  $y \mapsto \langle \lambda, F(x, y) \rangle = \phi_1(y_1) + \phi_2(y_2)$  is strictly convex on the Riemannian manifold  $(M_2, g_2)$ . Since  $f$  and  $F$  are smooth, all the hypotheses of the previous sections are fulfilled.

So we can take advantage of the fact that the lower level (followers) objective is convex on this suitable Riemannian manifold, and therefore, we can transform equivalently our semivectorial bilevel problem into a usual bilevel problem (i.e., upper level and lower level are scalar problems) such that, for each choice of the new leader, the best response of the new lower level is unique (Theorem 3.1)!

Let us now find the function  $(x, \lambda) \mapsto y(x, \lambda)$ , which represents the unique solution of the problem  $\min_{y \in M_2} \langle \lambda, F(x, y) \rangle$ . To simplify the computations, let us remark that this minimization problem is equivalent to the problem

$$\min_{y \in M_2} \frac{1}{\lambda_1 + \lambda_2} \langle \lambda, F(x, y) \rangle = \min_{y \in M_2} (\theta F_1(x, y) + (1 - \theta)F_2(x, y)),$$

where  $\theta := \frac{\lambda_1}{\lambda_1 + \lambda_2} \in ]0, 1[$ .

Thus,  $y(x, \lambda) = y(x, \theta)$ , where  $y(x, \theta)$  is the unique solution of the problem  $\min_{y \in M_2} (\theta F_1(x, y) + (1 - \theta)F_2(x, y))$ .

Since  $\{(\frac{\lambda_1}{\lambda_1 + \lambda_2}, \frac{\lambda_2}{\lambda_1 + \lambda_2}) \mid \lambda \in \Lambda_p\} = \{(\theta, 1 - \theta) \mid \theta \in ]0, 1[\}$ , by Theorem 3.1 problem  $(OSB)_p$  is equivalent to the problem

$$\min_{x \in M_1} \min_{\theta \in ]0, 1[} f(x, y(x, \theta)). \tag{20}$$

So  $y(x, \theta)$  solve the equation  $\text{grad}(y \mapsto (\theta F_1(x, y) + (1 - \theta) F_2(x, y))) = 0$ , i.e.,

$$y_1^2 \left( \theta \frac{\partial F_1}{\partial y_1}(y) + (1 - \theta) \frac{\partial F_2}{\partial y_1}(y) \right) \frac{\partial}{\partial y_1} + y_2^2 \left( \theta \frac{\partial F_1}{\partial y_2}(y) + (1 - \theta) \frac{\partial F_2}{\partial y_2}(y) \right) \frac{\partial}{\partial y_2} = 0.$$

Thus,  $y(x, \theta) = (y_1(x, \theta), y_2(x, \theta))$  solves the systems of equations

$$\begin{aligned} \theta \frac{\partial F_1}{\partial y_1}(y) + (1 - \theta) \frac{\partial F_2}{\partial y_1}(y) &= 0 \\ \theta \frac{\partial F_1}{\partial y_2}(y) + (1 - \theta) \frac{\partial F_2}{\partial y_2}(y) &= 0, \end{aligned}$$

and some simple computations give

$$y(x, \theta) = \left( \exp \left( - \frac{1 - \theta}{2\theta} x \right), \frac{4(1 - \theta)^2}{\theta^2 x^2} \right).$$

To simplify notations, put  $\eta = \frac{1 - \theta}{2\theta}$  and notice that the function  $\theta \mapsto \eta$  is a bijection from  $]0, 1[$  to  $]0, +\infty[$ . Thus, if we denote

$$\hat{y}(x, \eta) = y(x, \theta) = \left( e^{-\eta x}, \frac{16\eta^2}{x^2} \right),$$

problem (20) becomes

$$\min_{x \in M_1} \min_{\eta \in ]0, +\infty[} f(x, \hat{y}(x, \eta)) = \min_{(x, \eta) \in M_1 \times ]0, +\infty[} f(x, \hat{y}(x, \eta)).$$

But

$$f(x, \hat{y}(x, \eta)) = \frac{x^2}{4} + 4\eta^2 - x - \eta,$$

which obviously has  $(2, \frac{1}{8})$  as unique global minimizer over  $M_1 \times ]0, +\infty[ = \mathbb{R}_+^2$ .

Finally, the global minimizer of our original problem  $(OSB)_p$  is given by

$$(x, y_1, y_2) = \left( 2, e^{-\frac{1}{4}}, \frac{1}{16} \right).$$

### 4 Conclusions

We have considered the generalization in three directions of the technics presented in [11, 12] for the semivectorial bilevel convex optimal control problems. Thus, our problem has nonlinear upper and lower level obtaining more explicit optimality conditions, the order given by a convex pointed closed cone, and, the most important thing, we deal now with the Riemannian manifolds setting. The different results obtained are

closely related to the geometric structure of leader's Riemannian manifold as well as the followers one. Thus, explicit solutions could be obtained on specific Riemannian manifolds. The results presented are also important for the Euclidean case.

Even if some constrained optimization problem on linear spaces may become unconstrained for a suitable Riemannian manifold, we consider that further research avenues should include the case of explicit constraints in the upper and lower level.

**Acknowledgments** The authors are very grateful to the anonymous referees for their useful comments and suggestions which have improved the quality of the paper.

## References

1. Edgeworth, F.Y.: *Mathematical Psychics*. Kegan Paul, London (1881)
2. Pareto, V.: *Cours d'Economie Politique*. F. Rouge, Lausanne (1896)
3. Kuhn, H.W., Tucker, A.W.: Nonlinear programming. In: Neyman, J. (ed.) *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 481–492. University of California Press, Berkeley (1951)
4. Bonnel, H., Morgan, J.: Semivectorial bilevel optimization problem: penalty approach. *J. Optim. Theory Appl.* **131**(3), 365–382 (2006)
5. Bonnel, H.: Optimality conditions for the semivectorial bilevel optimization problem. *Pac. J. Optim.* **2**(3), 447–468 (2006)
6. Ankhili, Z., Mansouri, A.: An exact penalty on bilevel programs with linear vector optimization lower level. *Eur. J. Oper. Res.* **197**, 36–41 (2009)
7. Eichfelder, G.: Multiobjective bilevel optimization. *Math. Program. Ser. A* **123**, 419–449 (2010)
8. Calvete, H., Galé, C.: On linear bilevel problems with multiple objectives at the lower level. *Omega* **39**, 33–40 (2011). Elsevier
9. Zheng, Y., Wan, Z.: A solution method for semivectorial bilevel programming problem via penalty method. *J. Appl. Math. Comput.* **37**, 207–219 (2011)
10. Dempe, S., Gadhi, N., Zemkoho, A.B.: New optimality conditions for the semivectorial bilevel optimization problem. *J. Optim. Theory Appl.* **157**(1), 54–74 (2013)
11. Bonnel, H., Morgan, J.: Semivectorial bilevel convex optimal control problems: an existence result. *SIAM J. Control Optim.* **50**(6), 3224–3241 (2012)
12. Bonnel, H., Morgan, J.: Optimality conditions for semivectorial bilevel convex optimal control problems. In: Bailey, D.H., Bauschke, H.H., Borwein, P., Garvan, F., Théra, M., Vanderwerff, J.D., Wolkowicz, H. (eds.) *Computational and Analytical Mathematics in Honor of Jonathan Borwein's 60th Birthday*, pp. 43–74. Springer, Berlin (2013). ISBN: 978-1-4614-7620-7
13. Philip, J.: Algorithms for the vector maximization problem. *Math. Program.* **2**, 207–229 (1972)
14. Benson, H.P.: Optimization over the efficient set. *J. Math. Anal. Appl.* **98**, 562–580 (1984)
15. Craven, B.D.: Aspects of multicriteria optimization. In: *Recent Developments in Mathematical Programming*, pp. 93–100 (1991)
16. Benson, H.P.: A finite, non-adjacent extreme point search algorithm for optimization over the efficient set. *J. Optim. Theory Appl.* **73**, 47–64 (1992)
17. Bolintineanu, S.: Minimization of a quasi-concave function over an efficient set. *Math. Program.* **61**, 89–110 (1993)
18. Bolintineanu, S.: Optimality conditions for minimization over the (weakly or properly) efficient set. *J. Math. Anal. Appl.* **173**(2), 523–541 (1993)
19. Bolintineanu, S.: Necessary conditions for nonlinear suboptimization over the weakly-efficient set. *J. Optim. Theory Appl.* **78**, 579–598 (1993)
20. Dauer, J.P.: Optimization over the efficient set using an active constraint approach. *J. Oper. Res.* **35**, 185–195 (1991)
21. Fülöp, J.: A cutting plane algorithm for linear optimization over the efficient set. In: Komlósi, S., Rapcsák, T., Schaible, S. (eds.) *Generalized Convexity. Lecture Notes in Economics and Mathematical System*, vol. 405, pp. 374–385. Springer, Berlin (1994)
22. Dauer, J.P., Fosnauth, T.A.: Optimization over the efficient set. *J. Glob. Optim.* **7**, 261–277 (1995)

23. Konno, H., Thach, P.T., Yokota, D.: Dual approach to minimization on the set of pareto-optimal solutions. *J. Optim. Theory Appl.* **88**, 689–707 (1996)
24. Bolintineanu, S., El Maghri, M.: Pénalisation dans l'optimisation sur l'ensemble faiblement efficient. *RAIRO Oper. Res.* **31**(3), 295–310 (1997)
25. Horst, R., Thoai, N.V.: Maximizing a concave function over the efficient or weakly-efficient set. *Eur. J. Oper. Res.* **117**, 239–252 (1999)
26. Horst, R., Thoai, N.V., Yamamoto, Y., Zenke, D.: On optimization over the efficient set in linear multicriteria programming. *J. Optim. Theory Appl.* **134**, 433–443 (2007)
27. Bonnel, H., Kaya, C.Y.: Optimization over the efficient set of multi-objective control problems. *J. Optim. Theory Appl.* **147**(1), 93–112 (2010)
28. Bonnel, H., Collonge, J.: Stochastic optimization over a pareto set associated with a stochastic multi-objective optimization problem. *J. Optim. Theory Appl.* **162**(2), 405–427 (2014)
29. Bonnel, H., Collonge, J.: Optimization over the pareto outcome set associated with a convex bi-objective optimization problem: theoretical results, deterministic algorithm and application to the stochastic case. *J. Glob. Optim.* (in press). doi:[10.1007/s10898-014-0257-0](https://doi.org/10.1007/s10898-014-0257-0)
30. Yamamoto, Y.: Optimization over the efficient set: an overview. *J. Glob. Optim.* **22**, 285–317 (2002)
31. Morgan, J.: Constrained well-posed two-level optimization problems. In: Clarke, F., Dem'yanov, V.F., Giannessi, F. (eds.) *Nonsmooth Optimization and Related Topics*, Ettore Majorana International Sciences Series, pp. 307–326. Plenum Press, New York (1989)
32. Loridan, P., Morgan, J.: New results on approximate solutions in two-level optimization. *Optimization* **20**, 819–836 (1989)
33. Lignola, M.B., Morgan, J.: Stability of regularized bilevel programming problems. *J. Optim. Theory Appl.* **93**, 575–596 (1997)
34. Dempe, S.: *Foundations of Bilevel Programming*. Kluwer, Dordrecht (2002)
35. Colson, B., Marcotte, P., Savard, G.: An overview of bilevel optimization. *Ann. Oper. Res.* **153**, 235–256 (2007)
36. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization* **52**, 333–359 (2003)
37. Udriște, C.: *Convex Functions and Optimization Methods on Riemannian Manifolds*, Mathematics and Its Applications, vol. 297. Kluwer, Dordrecht (1994)
38. Da Cruz Neto, J.X., Ferreira, O.P., Lucámbio Pérez, L.R., Németh, S.Z.: Convex-and monotone-transformable mathematical programming problems and a proximal-like point method. *J. Glob. Optim.* **35**, 53–69 (2006)
39. Bento, G.C., Melo, J.G.: Subgradient method for convex feasibility on riemannian manifolds. *J. Optim. Theory Appl.* **152**, 773–785 (2012)
40. Colao, V., López, G., Marino, G., Martín-Márquez, V.: Equilibrium problems in Hadamard manifolds. *J. Math. Anal. Appl.* **388**, 61–77 (2012)
41. Ferreira, O.P., Iusem, A.N., Németh, S.Z.: Concepts and techniques of optimization on the sphere. *TOP* **22**, 1148–1170 (2014). doi:[10.1007/s11750-014-0322-3](https://doi.org/10.1007/s11750-014-0322-3)
42. Bento, G.C., Ferreira, O.P., Olivera, P.R.: Unconstrained steepest descent method for multicriteria optimization on Riemannian manifolds. *J. Optim. Theory Appl.* **154**, 88–107 (2012)
43. Bento, G.C., Cruz Neto, J.X.: A subgradient method for multiobjective optimization on Riemannian manifolds. *J. Optim. Theory Appl.* **159**, 125–137 (2013)
44. Bento, G.C., Cruz Neto, J.X., Santos, P.S.M.: An inexact steepest descent method for multicriteria optimization on Riemannian manifolds. *J. Optim. Theory Appl.* **159**, 108–124 (2013)
45. Yu, P.L.: *Multiple-Criteria Decision Making: Concepts, Techniques, and Extensions*. Mathematical Concepts and Methods in Science and Engineering, vol. 30. Plenum Press, New York (1985)
46. Luc, D.T.: *Theory of Vector Optimization*. Lecture Notes in Econom. and Math. Systems, vol. 319. Springer, Berlin (1989)
47. Miettinen, K.M.: *Nonlinear Multiobjective Optimization*. Kluwer, Dordrecht (1998)
48. Ehrgott, M.: *Multicriteria Optimization*. Springer, Berlin (2000)
49. Jahn, J.: *Vector Optimization*. Springer, Berlin (2004)
50. Eichfelder, G.: *Adaptive Scalarization Methods in Multiobjective Optimization*. Springer, Berlin (2008)
51. Do Carmo, M.P.: *Riemannian Geometry*. Birkhäuser, Boston (1993)
52. Jost, J.: *Riemannian Geometry and Geometric Analysis*. Springer, Berlin (2005)
53. Clarke, F.: *Functional Analysis, Calculus of Variations and Optimal Control*. Springer, London (2013)
54. Rockafellar, R.T., Wets, R.J.-B.: *Variational Analysis*. Springer, Berlin (1998)