

Nonsmooth Optimization Techniques on Riemannian Manifolds

S. Hosseini · M.R. Pouryayevali

Received: 31 March 2012 / Accepted: 12 December 2012
© Springer Science+Business Media New York 2012

Abstract We present the notion of weakly metrically regular functions on manifolds. Then, a sufficient condition for a real valued function defined on a complete Riemannian manifold to be weakly metrically regular is obtained, and two optimization problems on Riemannian manifolds are considered. Moreover, we present a generalization of the Palais–Smale condition for lower semicontinuous functions defined on manifolds. Then, we use this notion to obtain necessary conditions of optimality for a general minimization problem on complete Riemannian manifolds.

Keywords Ekeland variational principle · Contingent cone · Metric regularity · Generalized gradient · Riemannian manifolds

1 Introduction

The notion of metric regularity plays an important role in optimization in the case of linear spaces and, as is well known, strict differentiability is a very useful tool in the analysis of metrically regular functions. Being well-established and recognized, this concept still continues its expansion into new areas of mathematical analysis; see monographs [1, 2] and survey paper [3]. In [4], the notion of metric regularity on linear spaces was compared with a regularity condition which was established

Communicated by René Henrion.

S. Hosseini · M.R. Pouryayevali (✉)
School of Mathematics, Institute for Research in Fundamental Sciences (IPM),
P.O. Box 19395-5746, Tehran, Iran
e-mail: pourya@math.ui.ac.ir

S. Hosseini
e-mail: somayeh-hosseini@hotmail.com

by means of the image space analysis. Another useful notion in optimization is the Palais–Smale condition. For the existence of a minimum, one usually imposes certain generalized convexity conditions while for the existence of other critical points, one needs the Palais–Smale condition. The rapid development of nonsmooth optimization in the recent decades has led to numerous attempts to extend the main facts of classical analysis to nondifferentiable and nonconvex functions. In [5], utilizing Clarke generalized gradient, a generalization of this condition for locally Lipschitz functions on Hilbert spaces, was presented. In [6], the authors applied the concept of a Clarke generalized gradient on Riemannian manifolds to introduce a generalization of Palais–Smale condition for locally Lipschitz functions on Riemannian manifolds.

A manifold, in general, does not have a linear structure, hence the usual techniques, which are often used to study optimization problems on Banach spaces, cannot be applied. Therefore new techniques are needed for dealing with optimization problems posed on manifolds. In the past few years a number of results have been obtained on numerous aspects of optimization on Riemannian manifolds; see, for instance, [7–14]. Moreover, nonsmooth functions on Riemannian manifolds were studied in [15–18].

In spite of the significant role and various applications of metric regularity, it is merely considered on linear spaces. To obtain similar results on manifolds, the first step is to consider the notion of metric regularity on Riemannian manifolds. In this paper, we introduce the concept of strict differentiability in the setting of Riemannian manifolds. Then, we provide a sufficient condition for real valued functions defined on complete Riemannian manifolds to be weakly metrically regular in terms of strict differentiability. In addition, we obtain necessary optimality conditions for optimization problems in this setting. Our key tool is Ekeland’s variational principle; hence we shall work only with complete Riemannian manifolds.

Using the notion of contingent cone to closed subsets of Riemannian manifolds, a version of Palais–Smale condition for lower semicontinuous functions on Riemannian manifolds is introduced; see [19]. This condition reduces to the usual one in the case of locally Lipschitz functions. Then, by using the concept of generalized gradient, we present necessary conditions for an optimization problem on manifolds.

The paper is planned as follows. In Sect. 2, we introduce some basic definitions and notations, widely used in the sequel. Then, the notion of metric regularity of functions on the manifold setting is generalized. Moreover, we characterize the contingent cone to the feasible region of a problem under a constraint qualification. At the end of this section, two optimization problems on Riemannian manifolds are considered. Section 3 is concerned with a general minimization problem on Riemannian manifolds.

2 Metric Regularity

In this paper, we use the standard notations and known results of Riemannian manifolds; see, for instance, [20]. In what follows M is a C^∞ smooth manifold modeled on a Hilbert space H , endowed with a Riemannian metric $\langle \cdot, \cdot \rangle_x$ on the tangent space $T_x M \cong H$. In the case when γ is a minimizing geodesic and $\gamma(t_0) = x$, $\gamma(t_1) = y$,

the parallel translation from $T_x M$ to $T_y M$ along the curve γ is denoted by L_{xy} . The next remark contains a property of parallel translation which will be used in the rest of the paper.

Remark 2.1 Let M be a Riemannian manifold. An easy consequence of the definition of the parallel translation along a curve as a solution to an ordinary linear differential equation, implies that the mapping

$$(x, \xi) \in T_x M \mapsto L_{xx_0}(\xi),$$

where x is in a neighborhood of x_0 , is well defined and continuous at (x_0, ξ_0) ; that is, if $(x_n, \xi_n) \rightarrow (x_0, \xi_0)$ in TM , then $L_{x_n x_0}(\xi_n) \rightarrow L_{x_0 x_0}(\xi_0) = \xi_0$, for every $(x_0, \xi_0) \in TM$; see [15, Remark 6.11].

If S is a nonempty and closed subset of M , we define $d_S : M \rightarrow \mathbb{R}$ by

$$d_S(x) := \inf\{d(x, s) : s \in S\},$$

where d is the Riemannian distance on M .

Let us present some definitions and notions related to nonsmooth analysis on Riemannian manifolds. We refer to [21] for nonsmooth analysis on Banach spaces.

Let $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function on a Riemannian manifold M and $x \in \text{dom}(f) := \{x \in M \mid f(x) < +\infty\}$. Then, the subderivative of f at x in the direction $v \in T_x M$, denoted by $f'(x, v)$, is defined by

$$f'(x, v) := \liminf_{w \rightarrow 0_x, t \downarrow 0} \frac{f(\exp_x(t(v+w))) - f(x)}{t}.$$

Let $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function on a Riemannian manifold M . The contingent derivative of f at $x \in \text{dom}(f)$ in the direction $v \in T_x M$, denoted by $Df(x, v)$, is defined by

$$Df(x, v) := \limsup_{w \rightarrow 0_x, t \downarrow 0} \frac{f(\exp_x(t(v+w))) - f(x)}{t}.$$

Another useful notion in this paper is the generalized directional derivative; see [15].

Suppose that $f : M \rightarrow \mathbb{R}$ is a locally Lipschitz function on a Riemannian manifold M and (φ, U) is a chart at $x \in M$. Then, the generalized directional derivative of f at x in the direction $v \in T_x M$, denoted by $f^\circ(x; v)$, is defined by

$$f^\circ(x; v) := \limsup_{y \rightarrow x, t \downarrow 0} \frac{f \circ \varphi^{-1}(\varphi(y) + t d\varphi(x)(v)) - f \circ \varphi^{-1}(\varphi(y))}{t}. \tag{1}$$

Indeed, $f^\circ(x; v) = (f \circ \varphi^{-1})^\circ(\varphi(x); d\varphi(x)(v))$. Considering $0_x \in T_x M$, we have

$$f^\circ(x; v) = (f \circ \exp_x)^\circ(0_x, v). \tag{2}$$

It is worthwhile to mention that, if $f : M \rightarrow \mathbb{R}$ is Lipschitz around x , then

$$f'(x, v) \leq Df(x, v) = f^\circ(x, v), \quad \forall v \in T_x M.$$

The generalized gradient or the Clarke subdifferential of a locally Lipschitz function f at $x \in M$, denoted by $\partial f(x)$, is the subset of $T_x M^*$, whose support function is $f^\circ(x; \cdot)$. Thus $\xi \in \partial f(x)$ if and only if $f^\circ(x; v) \geq \langle \xi, v \rangle$ for all v in $T_x M$. Note that the Clarke subdifferential does not shrink to the classic one under differentiability; it happens, if the function is continuously differentiable.

We define the contingent cone and the tangent cone to a closed set $S \subseteq M$ at the point $x \in S$ denoted by $K_S(x)$ and $T_S(x)$, respectively, as

$$K_S(x) := \{h \in T_x M : \exists t_r \downarrow 0, \exists h_r \in T_x M, h_r \rightarrow h, \exp_x(t_r h_r) \in S, \forall r\},$$

$$T_S(x) := \{h \in T_x M : \forall (z_r) \subset \exp_x^{-1}(S \cap B_\varepsilon(x)); z_r \rightarrow 0_x, \forall t_r \downarrow 0, \exists h_r \in T_x M, h_r \rightarrow h \text{ and } \exp_x(z_r + t_r h_r) \in S, \forall r\},$$

where $B_\varepsilon(x)$ is a geodesic ball around x ; for more details see [18]. We call the set S tangentially regular at the point $x \in S$ iff the contingent and tangent cones coincide. In the case of submanifolds of \mathbb{R}^n , the tangent space and the normal space are orthogonal to one another. In an analogous manner, for a closed subset S of a Riemannian manifold M , the normal cone to S at x , denoted by $N_S(x)$, is defined as the (negative) polar of the tangent cone $T_S(x)$, i.e.

$$N_S(x) := T_S(x)^\circ := \{\xi \in T_x M^* : \langle \xi, z \rangle \leq 0, \forall z \in T_S(x)\}.$$

Note that it is easy to verify that $\partial d_S(x) \subseteq N_S(x)$; see [18].

Let us introduce the notion of weakly metrically regular for functions defined on Riemannian manifolds. Let U be an open subset of a Riemannian manifold M ; a continuous function $h : U \rightarrow \mathbb{R}$ is said to be weakly metrically regular on closed subset $S \subset U$ at the point $x \in S$ iff there is a real constant k such that, for all z in S close to x ,

$$d_{S \cap h^{-1}(h(x))}(z) \leq k|h(z) - h(x)|.$$

Example 2.1 In any sphere the distance function from the north pole is weakly metrically regular at the north pole.

Lemma 2.1 *If h is Fréchet differentiable at the point x in an open subset U of a Riemannian manifold M , then $K_{h^{-1}(h(x))}(x) \subset N(dh(x))$, where $N(dh(x))$ denotes the null space of $dh(x)$.*

Proof Let $p \in K_{h^{-1}(h(x))}(x)$, so there exist a sequence of positive real numbers $\{t_r\}$ and a sequence $\{p_r\} \subset T_x M$ such that $p_r \rightarrow p$ and $t_r \rightarrow 0$ with $h(\exp_x(t_r p_r)) = h(x)$ for all r . By properties of the exponential function, there exists $\varepsilon_1 > 0$ such that $\exp_x : B_{\varepsilon_1}(0_x) \rightarrow B_{\varepsilon_1}(x)$ is diffeomorphism. Let $\varepsilon_2 > 0$ such that $B_{\varepsilon_2}(x) \subseteq U$. Set $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$, and let $h \circ \exp_x : B_\varepsilon(0_x) \rightarrow \mathbb{R}$. Then

$$0 = \lim_{r \rightarrow \infty} \frac{h(\exp_x(t_r p_r)) - h(\exp_x(0_x))}{t_r} = dh(x)(p),$$

as required. □

Let us recall Ekeland’s variational principle on complete Riemannian manifolds. In the rest of this section, M will be a complete finite dimensional Riemannian manifold.

Theorem 2.1 *Suppose that M is a complete Riemannian manifold, and that $f : M \rightarrow \mathbb{R} \cup \{-\infty\}$ is a proper upper semicontinuous function which is bounded above. Let $\epsilon > 0$ be given, and a point $x_0 \in M$ such that*

$$f(x_0) > \sup\{f(x) : x \in M\} - \epsilon.$$

Then, for every $\lambda > 0$ there exists a point $z \in \text{dom}(f)$ such that

- (i) $\frac{\epsilon}{\lambda}d(z, x_0) \leq f(z) - f(x_0)$.
- (ii) $d(z, x_0) \leq \lambda$.
- (iii) $\frac{\epsilon}{\lambda}d(x, z) + f(z) > f(x)$ whenever $x \neq z$.

Lemma 2.2 *Let S be a closed subset of a Riemannian manifold M and let h be a real valued continuous function defined on an open subset U of M containing S . Suppose that $x \in S$ and $h(x) = 0$. If h is not weakly metrically regular on S at x , then there is a sequence (v_r) in S such that $v_r \rightarrow x$ and $h(v_r) \neq 0$ for all r . Moreover, there exists a strictly positive sequence $\delta_r \downarrow 0$ such that, if $g_r : S \rightarrow \mathbb{R}$ is defined as $g_r(s) := |h(s)| + \delta_r d(s, v_r)$, then g_r attains a minimum at v_r on S .*

Proof By the definition of weakly metrically regular functions, there exists $\{x_r\} \subset S$ which converges to x such that

$$d_{S \cap h^{-1}(0)}(x_r) > r|h(x_r)| \quad \text{and} \quad d(x_r, x) \leq \frac{1}{r}, \quad \text{for all } r. \tag{3}$$

Now, let

$$\delta_S(x) = \begin{cases} 0, & x \in S, \\ \infty, & \text{otherwise,} \end{cases}$$

and define $f : U \rightarrow \mathbb{R} \cup \{+\infty\}$ by $f(z) := |h(z)| + \delta_S(z)$. Clearly, f is a lower semicontinuous function. Set $x_0 := x_r$, $\epsilon := |h(x_r)|$ and $\lambda := \min\{r\epsilon, \sqrt{\epsilon}\}$ in Ekeland’s variational principle. Therefore, there exists $v_r \in S$ such that $d(v_r, x_r) \leq \lambda$ and the function

$$g_r(s) = f(s) + \frac{\epsilon}{\lambda}d(s, v_r) = |h(s)| + \max\left\{\frac{1}{r}, \sqrt{|h(x_r)|}\right\}d(s, v_r)$$

attains a minimum at v_r on S . It remains to show that $h(v_r) \neq 0$ and $v_r \rightarrow x$. On the contrary, assume that $h(v_r) = 0$. Then

$$r\epsilon = r|h(x_r)| < d_{S \cap h^{-1}(0)}(x_r) < d(v_r, x_r) \leq \lambda \leq r\epsilon,$$

which is a contradiction. Moreover, by (3) and Ekeland’s variational principle, it follows that

$$d(v_r, x) < d(v_r, x_r) + d(x_r, x) \leq \lambda + \frac{1}{r} \leq r\epsilon + \frac{1}{r} \leq d_{S \cap h^{-1}(0)}(x_r) + \frac{1}{r} \rightarrow 0,$$

as required. □

We are going to obtain a sufficient condition for strictly differentiable functions defined on complete Riemannian manifolds to be weakly metrically regular. This is a generalization of a known result on Euclidean spaces; see [22, Theorem 7.1.5]. For this purpose, we need to define strictly differentiable functions on complete Riemannian manifolds.

Definition 2.1 Let U be an open subset of a Riemannian manifold M . A function $h : U \rightarrow \mathbb{R}$ is called strictly differentiable at $x \in U$ iff there exists $\phi \in T_x^*M$ such that

$$\lim_{\substack{y, z \rightarrow x \\ y \neq z}} \frac{h(y) - h(z) - \phi(\exp_x^{-1}(y) - \exp_x^{-1}(z))}{\|\exp_x^{-1}(y) - \exp_x^{-1}(z)\|} = 0.$$

Lemma 2.3 Let U be an open subset of a Riemannian manifold M , and let $h : U \rightarrow \mathbb{R}$ be strictly differentiable at x . Then h is Lipschitz around x .

Proof Let $\varepsilon = 1$, there exists $\delta_1 > 0$ such that, for all $y, z \in B_{\delta_1}(x)$,

$$|h(y) - h(z) - \phi(\exp_x^{-1}(y) - \exp_x^{-1}(z))| \leq \|\exp_x^{-1}(y) - \exp_x^{-1}(z)\|.$$

Thus;

$$\begin{aligned} |h(y) - h(z)| &\leq \|\exp_x^{-1}(y) - \exp_x^{-1}(z)\| + \|\phi\| \|\exp_x^{-1}(y) - \exp_x^{-1}(z)\| \\ &= (1 + \|\phi\|) \|\exp_x^{-1}(y) - \exp_x^{-1}(z)\|. \end{aligned}$$

By [15, Theorem 2.3], there exists $B_{\delta_2}(x)$ such that $\exp_x^{-1} : B_{\delta_2}(x) \rightarrow B_{\delta_2}(0_x)$ is $(1 + \|\phi\|)$ -Lipschitz. Set $\delta = \min\{\delta_1, \delta_2\}$, so h is $(1 + \|\phi\|)^2$ -Lipschitz on $B_\delta(x)$. \square

To prove our next result, we need the following lemma, which can be proved easily.

Lemma 2.4 Let S be a closed subset of a Riemannian manifold M and $x \in S$, suppose that the function $f : M \rightarrow \mathbb{R}$ is Lipschitz around x . If x is a local minimizer of f on S , then, for real number L sufficiently large, x is a local minimizer of $f + Ld_S$.

Theorem 2.2 Let S be a closed subset of a Riemannian manifold M and U be an open subset of M containing S . If $h : U \rightarrow \mathbb{R}$ is strictly differentiable at $x \in S$ and $dh(x)(T_S(x)) = \mathbb{R}$, then h is weakly metrically regular on S at x .

Proof Without any loss of generality, we can suppose $h(x) = 0$. We proceed by contradiction. If h is not weakly metrically regular on S at x , then from Lemma 2.2, there exists a sequence $\{v_r\} \subset S$ converging to x such that $h(v_r) \neq 0$ for all $r \in \mathbb{R}$, and there exists a sequence of positive real numbers $\{\delta_r\}$ converging to zero such that $g_r(s) = |h(s)| + \delta_r d(s, v_r)$ is minimized at v_r on S . By Lemma 2.3, g_r is Lipschitz around x , hence Lemma 2.4 and [18, Proposition 3.1] imply

$$0 \in \partial(g_r + Ld_S)(v_r) \subseteq \partial(|h|)(v_r) + \delta_r B + L\partial d_S(v_r),$$

where B is the unit ball of $T_{v_r} M^*$. Thus, there exist $w_r \in \partial(|h|)(v_r)$ and $u_r \in \partial d_S(v_r)$ such that $L_{v_r, x}(w_r + u_r) \rightarrow 0$. Note that $\{h(v_r)/|h(v_r)|\} \subset \mathbb{R}$ is bounded, hence it has a convergent subsequence to some nonzero element y . Employing the Chain Rule Theorem, we deduce

$$\partial(|h|)(v_r) = \frac{h(v_r)}{|h(v_r)|} dh(v_r).$$

Now, by assumption h is strictly differentiable at x and $v_r \rightarrow x$, therefore Remark 2.1 implies that

$$L_{v_r, x}(w_r) = L_{v_r, x}\left(\frac{h(v_r)}{|h(v_r)|} dh(v_r)\right) \rightarrow dh(x)y.$$

Consequently, $L_{v_r, x}(u_r) \rightarrow -dh(x)y$. Thus from [18, Theorem 2.9], we have $-dh(x)y \in N_S(x)$. Let $0 \neq p \in T_S(x)$ such that $dh(x)(p) = -y$, hence

$$0 \geq -dh(x)y(p) = -dh(x)(p)y = y^2 > 0,$$

which is a contradiction. □

The following examples illustrate some weakly metrically regular functions which arise on Riemannian manifolds.

Example 2.2 A closed subset S of a finite dimensional Hadamard manifold M is called φ -convex iff there exists a continuous function $\varphi : S \rightarrow (0, +\infty)$ such that

$$\langle \zeta, \exp_x^{-1}(y) \rangle_x \leq \varphi(x) \|\zeta\|_x d(x, y)^2, \quad \text{for every } x, y \in S \text{ and } \zeta \in N_S^F(x),$$

where $N_S^F(x)$ denotes the Fréchet normal cone to S at x , see [23]. Let S be a nonempty and φ -convex subset of M . Then, by [23, Theorem 3.6] there exists an open neighborhood U of S such that $d_S^2 : M \rightarrow \mathbb{R}$ is C^1 on $U \setminus S$ and $\nabla d_S^2|_{U \setminus S} \neq 0$, where ∇d_S^2 denotes the gradient vector field of d_S^2 . Therefore, Theorem 2.2 implies d_S^2 is weakly metrically regular on M at every point $x \in U \setminus S$.

Example 2.3 Let $\text{Sym}_n(\mathbb{R})$ be the set of all $n \times n$ symmetric matrices and let $\text{Pos}_n(\mathbb{R})$ be the set of all $n \times n$ positive definite matrices endowed with Riemannian metric $g_X(U, V) = \text{tr}(VX^{-1}UX^{-1})$, where $X \in \text{Pos}_n(\mathbb{R})$, U and V are in $T_X(\text{Pos}_n(\mathbb{R})) = \text{Sym}_n(\mathbb{R})$. It is easy to see that $(\text{Pos}_n(\mathbb{R}), g)$ is a complete finite dimensional Riemannian manifold. Now we define $h : (\text{Pos}_n(\mathbb{R}), g) \rightarrow \mathbb{R}$ by $h(X) := \ln \det X$. Note that h is C^1 and the gradient vector field of h on $\text{Pos}_n(\mathbb{R})$ is nonzero. So by Theorem 2.2 h is weakly metrically regular on $\text{Pos}_n(\mathbb{R})$.

Example 2.4 Let $\mathbb{H} := \{z = x + iy \in \mathbb{C} : \text{Im}g(z) = y > 0\}$ with the Riemannian metric $g_z(a, b) = (1/(\text{Im}g(z)^2))\langle a, b \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the scalar product on \mathbb{R}^2 . It is easy to see (\mathbb{H}, g) is a complete finite dimensional Riemannian manifold. We define $h : (\mathbb{H}, g) \rightarrow \mathbb{R}$ by $h(z) = \ln \text{Im}g(z)$, which is C^1 and the gradient vector field of h on \mathbb{H} is nonzero. Hence Theorem 2.2 implies h is weakly metrically regular on \mathbb{H} .

Now we arrive at one of the main results of this paper. We extend Liusternik’s theorem to the functions defined on complete Riemannian manifolds; see [22, Theorem 7.1.6].

Theorem 2.3 *Let U be an open subset of a Riemannian manifold M , and let $h : U \rightarrow \mathbb{R}$ be strictly differentiable at the point x and weakly metrically regular at x on $h^{-1}(h(x))$. Then the set $h^{-1}(h(x))$ is tangentially regular at x and $K_{h^{-1}(h(x))}(x) = N(dh(x))$, where $N(dh(x))$ denotes the null space of $dh(x)$.*

Proof It is sufficient to prove $N(dh(x)) \subseteq T_{h^{-1}(h(x))}(x)$. Indeed, from the definitions of the contingent cone and the tangent cone and Lemma 2.1, we obtain $T_{h^{-1}(h(x))}(x) \subseteq K_{h^{-1}(h(x))}(x) \subseteq N(dh(x))$. Let $p \in N(dh(x))$ and suppose that the arbitrary sequences $\{z_r\} \subset T_x M$ and $\{t_r\} \subset \mathbb{R}$ satisfying $z_r \rightarrow 0_x$, $t_r \downarrow 0$ are given. Since h is weakly metrically regular at x , hence there is $k \in \mathbb{R}$ such that

$$d_{h^{-1}(h(x))}(\exp_x(z_r + t_r p)) \leq k|h(\exp_x(z_r + t_r p)) - h(x)|, \quad \text{for all large } r.$$

Moreover, there exists $\varepsilon > 0$ such that $\exp_x^{-1} : B_\varepsilon(x) \rightarrow B_\varepsilon(0_x)$ is Lipschitz of rank $k + 1$; see [15, Theorem 2.3]. Note that there is $w_r \in h^{-1}(h(x))$ such that

$$d(w_r, \exp_x(z_r + t_r p)) \leq (k^2 + k)|h(\exp_x(z_r + t_r p)) - h(x)|.$$

Since $w_r \rightarrow 0_x$, it follows that, for all r large enough, $w_r \in B_\varepsilon(0_x)$; hence

$$\begin{aligned} \|z_r + t_r p - \exp_x^{-1}(w_r)\| &\leq (k + 1)d(w_r, \exp_x(z_r + t_r p)) \\ &\leq k(k + 1)^2|h(\exp_x(z_r + t_r p)) - h(x)|. \end{aligned}$$

Now, we define $p_r := (\exp_x^{-1}(w_r) - z_r)/(t_r) \in T_x M$, thus $w_r = \exp_x(z_r + t_r p_r)$ is in $h^{-1}(h(x))$. Consequently

$$\|p_r - p\| = \frac{\|z_r + t_r p - \exp_x^{-1}(w_r)\|}{t_r} \leq k(k + 1)^2 \frac{|h(\exp_x(z_r + t_r p)) - h(x)|}{t_r}.$$

By letting r go to $+\infty$, we see that the right hand of the inequality converges to $k(k + 1)^2|dh(x)(p)| = 0$, so p_r converges to p and the proof is complete. \square

Corollary 2.1 *Let U be an open subset of a Riemannian manifold M . If $h : U \rightarrow \mathbb{R}$ is strictly differentiable at the point x and $dh(x)$ is surjective, then the set $h^{-1}(h(x))$ is tangentially regular at x and $K_{h^{-1}(h(x))}(x) = N(dh(x))$, where $N(dh(x))$ denotes the null space of $dh(x)$.*

Theorem 2.4 *Let S be a closed subset of a Riemannian manifold M and U be an open subset of M containing S . If the point \bar{x} is a local minimizer of $f : U \rightarrow \mathbb{R}$ on S and the function f is Fréchet differentiable at \bar{x} , then $-df(\bar{x}) \in K_S(\bar{x})^\circ$, where \circ denotes the negative polar.*

Proof On the contrary, assume that $-df(\bar{x}) \notin K_S(\bar{x})^\circ$. Therefore, there exists $p \in K_S(\bar{x})$ such that $df(\bar{x})(p) < 0$. Since $p \in K_S(\bar{x})$, there are $p_r \rightarrow p$ and $t_r \downarrow 0$ such that $\exp_{\bar{x}}(t_r p_r) \in S$ for all r . On the other hand, f is Fréchet differentiable at \bar{x} , so

$$\lim_{r \rightarrow \infty} \frac{f(\exp_{\bar{x}}(t_r p_r)) - f(\bar{x}) - df(\bar{x})(t_r p_r)}{t_r \|p_r\|} = 0.$$

We obtain $f(\exp_{\bar{x}}(t_r p_r)) < f(\bar{x})$ for all r large enough. Since \bar{x} is a local minimizer of f , we get a contradiction. □

Now, we consider the following problem with constraints on a complete finite dimensional Riemannian manifold M :

$$\inf\{f(x) \mid x \in U, h(x) = 0\}, \tag{4}$$

where U is an open subset of M and $f : U \rightarrow \mathbb{R}$ is objective function and $h : U \rightarrow \mathbb{R}$ is continuous. We indicate a method to prove a version of Lagrange multipliers rule on complete Riemannian manifolds by means of metric regularity.

Remark 2.2 Any linear maps $A, G : X \rightarrow \mathbb{R}$, where X is a finite dimensional linear space, satisfy $\{x \in N(G) \mid Ax \leq 0\}^\circ = \mathbb{R}_+ A + \mathbb{R} G$, where \circ denotes the negative polar, $N(G)$ denotes null space of G and \mathbb{R}_+ is the set of nonnegative real numbers.

Theorem 2.5 *Suppose that the point \bar{x} is a local minimizer for the problem (4) and the objective function f is Fréchet differentiable at \bar{x} . If h is weakly metrically regular and strictly differentiable at the point $\bar{x} \in h^{-1}(0)$, then there exists a multiplier $\mu \in \mathbb{R}$ satisfying $df(\bar{x}) + \mu dh(\bar{x}) = 0$.*

Proof From Theorem 2.4 and Theorem 2.3,

$$-df(\bar{x}) \in K_S(\bar{x})^\circ = N(dh(\bar{x}))^\circ = \mathbb{R} dh(\bar{x}). \tag{5}$$

Now, we consider the following problem with constraints on a complete finite dimensional Riemannian manifold M ,

$$\inf\{f(x) \mid x \in U, h(x) = 0, g_i(x) \leq 0, \text{ for } i = 1, \dots, m\}, \tag{5}$$

where U is an open subset of M and $f : U \rightarrow \mathbb{R}$ is objective function and $h : U \rightarrow \mathbb{R}$ and $g_i : U \rightarrow \mathbb{R}$ (for $i = 1, \dots, m$) are all continuous.

Let us introduce a regularity condition named Mangasarian–Fromovitz constraint qualification: The functions g_i (for i in $I(\bar{x})$) are Fréchet differentiable at the point \bar{x} and h is strictly differentiable at \bar{x} , $dh(\bar{x}) : T_{\bar{x}}M \rightarrow \mathbb{R}$ is surjective, and there exists $p \in T_{\bar{x}}M$ such that

$$dh(\bar{x})(p) = 0, \quad dg_i(\bar{x})(p) < 0, \quad \text{for all } i \text{ in } I(\bar{x}),$$

where $I(\bar{x}) = \{i \mid g_i(\bar{x}) = 0, 1 \leq i \leq m\}$.

Now, we suppose that the Mangasarian–Fromovitz constraint qualification holds and we find a characterization of the contingent cone to the feasible region.

Theorem 2.6 *Let the Mangasarian–Fromovitz constraint qualification hold and S be the feasible region of problem (5). Then*

$$K_S(x) = \{p \in N(dh(x)) \mid dg_i(x)(p) \leq 0, \forall i \in I(x)\}.$$

Proof Consider the sets $K := \{p \in N(dh(x)) \mid dg_i(x)(p) \leq 0, \forall i \in I(x)\}$ and

$$\tilde{K} := \{p \in N(dh(x)) \mid dg_i(x)(p) < 0, \forall i \in I(x)\}.$$

First, we claim that

- (a) $K_S(x) \subseteq K$.
- (b) $\overline{\tilde{K}} = K$.
- (c) $\tilde{K} \subseteq K_S(x)$.

Then, since $K_S(x)$ is a closed set, (b) and (c) imply that $K \subseteq K_S(x)$ and the proof is complete.

Now, we prove the claim. If $p \in K_S(x)$, then there exist $t_r \downarrow 0$ and $p_r \rightarrow p$ with $\exp_x(t_r p_r) \in S$. Thus, for $i = 1, \dots, m$, $g_i(\exp_x(t_r p_r)) \leq 0$ and $h(\exp_x(t_r p_r)) = 0$. Since $S \subseteq h^{-1}(0)$, it follows that $K_S(x) \subseteq K_{h^{-1}(0)}(x)$ and from Lemma 2.1, $K_{h^{-1}(0)}(x) \subseteq N(dh(x))$. Moreover, for $i \in I(x)$ we have $g_i(\exp_x(t_r p_r)) \leq 0$ and $g_i(x) = 0$, so

$$0 \geq \lim_{t_r \rightarrow 0} \frac{g_i(\exp_x(t_r p_r)) - g_i(x)}{t_r \|p_r\|} = dg_i(x) \left(\frac{p}{\|p\|} \right),$$

hence $K_S(x) \subseteq K$.

Now, we prove part (b). Since $\tilde{K} \subseteq K$ and K is closed, then $\overline{\tilde{K}} \subseteq K$. If $d \in K$ and $dg_i(x)d < 0$ for all $i \in I(x)$, then $d \in \tilde{K}$. Otherwise, $d \in K$ and $dg_i(x)d = 0$, for some $i \in I(x)$, then by the Mangasarian–Fromovitz constraint qualification $\tilde{K} \neq \emptyset$, so there exists $\hat{d} \in \tilde{K}$ such that $dg_i(x)\hat{d} < 0$ for all $i \in I(x)$. Clearly, for $k \in \mathbb{N}$ we get $d_k = d + (1/k)\hat{d} \in \tilde{K}$ such that d_k converges to d , hence $d \in \overline{\tilde{K}}$; that is, $K \subseteq \overline{\tilde{K}}$.

For the last part, suppose that $p \in \tilde{K}$, Corollary 2.1 implies $p \in K_{h^{-1}(0)}(x)$, so there are sequences $t_r \downarrow 0$ and $p_r \rightarrow p$ with $h(\exp_x(t_r p_r)) = 0$ for all r . We claim that there exists $N \in \mathbb{N}$ such that $g_i(\exp_x(t_r p_r)) < 0$ for all $i = 1, \dots, m$ and $\exp_x(t_r p_r) \in U$, provided that $r \geq N$. Then the proof is complete. Since U is an open set and $x \in U$, clearly $\exp_x(t_r p_r) \in U$ for all large r . By continuity the claim follows for indices $i \notin I(x)$. It remains to prove the claim for $i \in I(x)$. On the contrary, suppose that for a subsequence $g_i(\exp_x(t_r p_r)) \geq 0, i \in I(x)$; then

$$0 = \lim_{r \rightarrow \infty} \frac{g_i(\exp_x(t_r p_r)) - g_i(x) - dg_i(x)(t_r p_r)}{\|t_r p_r\|} \geq -dg_i(x) \left(\frac{p}{\|p\|} \right) > 0,$$

and it is a contradiction. □

Finally, the Karush–Kuhn–Tucker condition is a direct consequence of the previous theorems.

Theorem 2.7 (Karush–Kuhn–Tucker Condition) *Suppose that the point \bar{x} is a local minimizer for the problem (5) and the objective function f is Fréchet differentiable at \bar{x} . If the Mangasarian–Fromovitz constraint qualification holds, then there exist multipliers λ_i in \mathbb{R}_+ (for $i \in I(\bar{x})$) and $\mu \in \mathbb{R}$ satisfying*

$$df(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i dg_i(\bar{x}) + \mu dh(\bar{x}) = 0.$$

Proof From Theorems 2.4 and 2.6,

$$\begin{aligned} -df(\bar{x}) \in K_S(\bar{x})^\circ &= \{p \in N(dh(\bar{x})) : dg_i(\bar{x})(p) \leq 0, \forall i \in I(\bar{x})\}^\circ \\ &= \sum_{i \in I(\bar{x})} \mathbb{R}_+ dg_i(\bar{x}) + \mathbb{R} dh(\bar{x}). \end{aligned} \quad \square$$

Example 2.5 Let us consider the set $\text{Sym}_2(\mathbb{R})$ of symmetric 2×2 matrices endowed with the Frobenius metric $g_X(U, V) = \text{tr}(UV)$ where $X \in (\text{Pos}_2 \mathbb{R})$ and $U, V \in T_X(\text{Pos}_2 \mathbb{R}) = \text{Sym}_2(\mathbb{R})$. The set $\text{Pos}_2(\mathbb{R})$ is a Hadamard manifold; see[20]. Consider the following problem on $\text{Pos}_2(\mathbb{R})$:

$$\begin{aligned} (\mathbf{P}_1) \quad & \min f(X) = x_1 \\ & \text{s.t. } g(X) = x_2 + x_3 - 7 \leq 0, \\ & h(X) = -x_1 + 1 \leq 0, \\ & X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \in \text{Pos}_2(\mathbb{R}). \end{aligned}$$

It is easy to check that

$$\bar{X} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$$

is a global optimal solution for (\mathbf{P}_1) . Using the Riemannian metric g and its inverse we have

$$\text{grad } f(\bar{X}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence, $df(\bar{X})(B) = b_1$ for all $B = \begin{bmatrix} b_1 & b_2 \\ b_2 & b_3 \end{bmatrix} \in T_{\bar{X}}(\text{Pos}_2 \mathbb{R}) = \text{Sym}_2(\mathbb{R})$.

Note that

$$\text{grad } h(\bar{X}) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix},$$

and $dh(\bar{X}) : \text{Sym}_2(\mathbb{R}) \rightarrow \mathbb{R}$ is onto. At the point $P = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$ we have $dg(\bar{X})(P) = -1$ and $dh(\bar{X})(P) = 0$. Therefore, the Mangasarian–Fromovitz constraint qualification holds. It is trivial that for $\lambda = 0$ and $\mu = 1$ we have $df(\bar{X}) + \lambda dg(\bar{X}) + \mu dh(\bar{X}) = 0$.

3 A General Minimization Problem

In what follows, M is a C^∞ smooth manifold modeled on a separable Hilbert space H , either finite dimensional or infinite dimensional. We first generalize the Palais–Smale condition for lower semicontinuous functions on Riemannian manifolds, which has been introduced on Hilbert spaces in [19]. Then, by means of the notion of generalized gradient, we present the necessary conditions of optimality for the problem

$$\min_{x \in S} f(x), \tag{6}$$

where $f : M \rightarrow \mathbb{R}$ is a locally Lipschitz function and S is an arbitrary nonempty subset of a Riemannian manifold M .

Definition 3.1 The lower semicontinuous function $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ is said to satisfy the Palais–Smale condition, abbreviated as (P.S.), at the level $c \in \mathbb{R}$ on the set $S \subset M$ iff every sequence $\{x_n\} \subset S$ along which $f(x_n) \rightarrow c$, and for all $v \in K_S(x_n)$, $Df(x_n, v) \geq -\epsilon_n \|v\|$, for a sequence $\epsilon_n \rightarrow 0^+$, possesses a convergent subsequence.

It is easy to verify that a locally Lipschitz function $f : M \rightarrow \mathbb{R}$ satisfies the (P.S.) condition on M if any sequence $\{x_n\}$, along which $f(x_n)$ is bounded and $\lambda(x_n) = \min_{w \in \partial f(x_n)} \|w\|_{T_{x_n} M^*} \rightarrow 0$, possesses a convergent subsequence.

The following theorem proves the existence result for the problem (6).

Theorem 3.1 *Let S be a closed subset of a complete Riemannian manifold M and $f : M \rightarrow \mathbb{R} \cup +\infty$ be a bounded below lower semicontinuous function on S with $\text{dom}(f) \cap S \neq \emptyset$. Assume that f satisfies the (P.S.) condition on S at level $c = \inf_S f$. Then the problem (6) has at least a solution $s \in S$ and*

$$Df(s; v) \geq 0, \quad \text{for all } v \in K_S(s). \tag{7}$$

Proof By Theorem 2.1, there exists a sequence $\{s_n\}$ in S such that

$$f(s) \geq f(s_n) - \frac{1}{n} d(s, s_n), \quad \forall s \in S. \tag{8}$$

Considering an arbitrary $v \in K_S(s_n)$, one can find sequences $\{t_k\} \subset \mathbb{R}^+$ and $\{w_k\} \subset T_{s_n} M$ such that $t_k \downarrow 0$ and $w_k \rightarrow 0_{s_n}$, and $\exp_{s_n}(t_k(v + w_k)) \in S$ for all k . From (8) it follows that

$$f(\exp_{s_n}(t_k(v + w_k))) - f(s_n) \geq \frac{-1}{n} d(\exp_{s_n}(t_k(v + w_k)), s_n) = \frac{-1}{n} \|t_k(v + w_k)\|.$$

By letting k go to $+\infty$, we get

$$Df(s_n, v) \geq \liminf_{k \rightarrow +\infty} \frac{f(\exp_{s_n}(t_k(v + w_k))) - f(s_n)}{t_k} \geq \frac{-1}{n} \|v\|, \quad \forall v \in K_S(s_n).$$

Moreover, $c = \inf_S f = \lim f(s_n)$. So that the (P.S.) condition provides a relabeled subsequence $\{s_n\}$ which converges to some point s in the closed set S . By lower semicontinuity of f , $f(s) = \inf_S f$.

It remains to show that the inequality (7) is true. For every $v \in K_S(s)$, there exist sequences $\{t_k\} \subset \mathbb{R}$ and $\{w_k\} \subset T_s M$ such that $t_k \downarrow 0$ and $w_k \rightarrow 0_s$, and $\exp_s(t_k(v + w_k)) \in S$ for all k . Hence $f(s) \leq f(\exp_s(t_k(v + w_k)))$ for all k , and one can deduce $Df(s; v) \geq 0$ for all $v \in K_S(s)$. □

The following result is a generalization of [5, Theorem 3.5] which can be proved by the previous theorem.

Corollary 3.1 *Let M be a complete Riemannian manifold and $f : M \rightarrow \mathbb{R}$ be a bounded below locally Lipschitz function. Assume that f satisfies the (P.S.) condition. Then problem $\min_{x \in M} f(x)$ has at least a solution $x \in M$ and $0 \in \partial f(x)$.*

Now, we get necessary conditions of optimality for the problem (6) in the case where f is locally Lipschitz.

Theorem 3.2 *Let S be a closed subset of a complete Riemannian manifold M and $f : M \rightarrow \mathbb{R}$ be a locally Lipschitz function. Let $s \in S$ be a solution of the problem (6). Assume that, for every $s' \in S$, there exists a linear operator*

$$A_{s'} : \text{dom}(A_{s'}) \subset T_{s'} M \rightarrow Y_{s'},$$

such that the domain of $A_{s'}$ is dense in $T_{s'} M$, $A_{s'}$ is a closed operator and the range of $A_{s'}$ is closed in $Y_{s'}$, where $Y_{s'}$ is a Banach space. Moreover, suppose that the null space of $A_{s'}$ is a subset of $K_S(s')$. Then, there is p in the domain of A_s^ such that $A_s^*(p) \in \partial f(s)$, where $A_s^* : \text{dom}(A_s^*) \rightarrow T_s M^*$ is the adjoint operator of A_s .*

Proof Let v be an arbitrary point of the null space of A_s , then there exist sequences $\{t_k\} \subset \mathbb{R}^+$ and $\{w_k\} \subset T_s M$ such that $t_k \downarrow 0$ and $w_k \rightarrow 0_s$, and $\exp_s(t_k(v + w_k)) \in S$ for all k . Since $s \in S$ is a solution of the problem (6) it follows that

$$f(s) \leq f(\exp_s(t_k(v + w_k))),$$

and one can deduce $f^\circ(s, v) \geq 0$ for all v in the null space of A_s . Hahn–Banach theorem provides some point $\xi \in T_s M^*$ such that $\langle \xi, v \rangle = 0$ for all v in the null space A_s , and $\langle \xi, y \rangle \leq f^\circ(s, y)$ for all $y \in T_s M$, which means $\xi \in \partial f(s)$. On the other hand, since range of A_s is closed, it follows that range of A_s^* is closed too. Hence $\xi \in [N(A_s)]^\perp = \overline{R(A_s^*)} = R(A_s^*)$, where N and R denote null and range spaces, respectively, which means there is p in domain A_s^* such that $\xi = A_s^*(p) \in \partial f(s)$, as required. □

We conclude this section with an application of the necessary conditions of optimality that demonstrates the generality of our results.

Let M be a complete Riemannian manifold. Assume now that the subset S in the problem (6) is given by

$$S = \{x \in M \mid \exists j \in J \text{ such that } G_j(x) = 0\},$$

where for each $j \in J$, $G_j : M \rightarrow \mathbb{R}$ is a C^1 mapping and $dG_j(x) : T_x M \rightarrow \mathbb{R}$ is surjective and the null space of $dG_j(x)$ has a topological complement whenever $G_j(x) = 0$. Moreover, assume that

$$G_i^{-1}(0) \cap G_j^{-1}(0) = \emptyset, \quad \text{if } i \neq j.$$

For every $s \in S$, there is a unique $j \in J$ such that $s \in G_j^{-1}(0)$. Suppose that $A_s = dG_j(s)$, using Corollary 2.1 one can deduce that the assumption of the previous theorem is satisfied. Consequently, there exists $p \in \mathbb{R}$ such that $dG_j(s)^*(p) \in \partial f(s)$. In the particular case when J is singleton, that is, $S = G^{-1}(0)$ where $G : M \rightarrow \mathbb{R}$ is a C^1 mapping, then for some $\lambda \in \mathbb{R}$, $\lambda dG(s) \in \partial f(s)$, which is the classical Lagrange multiplier rule for locally Lipschitz functions.

4 Conclusions

The concept of metric regularity on manifolds is of fundamental importance for both theoretical developments and numerically oriented studies. In this paper, we establish a sufficient condition for real valued functions defined on complete Riemannian manifolds to be weakly metrically regular in terms of strict differentiability. Some examples which arise naturally on Riemannian manifolds are presented. Using the notion of contingent cone, a version of Palais–Smale condition for lower semicontinuous functions on Riemannian manifolds is introduced. Then, by means of the notion of generalized gradient, we present necessary conditions of optimality for a general minimization problem.

Acknowledgements The second author was partially supported by the Center of Excellence for Mathematics, University of Shahrekord, Iran.

References

1. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, I: Basic Theory. Grundlehren Series (Fundamental Principles of Mathematical Sciences), vol. 331. Springer, Berlin (2006)
2. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, New York (1998)
3. Ioffe, A.D.: Metric regularity and subdifferential calculus. Russ. Math. Surv. **55**, 501–558 (2000)
4. Moldovan, A., Pellegrini, L.: On regularity for constrained extremum problems. Part 1: Sufficient optimality condition. J. Optim. Theory Appl. **142**, 147–163 (2009)
5. Chang, K.C.: Variational methods for non-differentiable functional and their applications to partial differential equations. J. Math. Anal. Appl. **80**, 102–129 (1981)
6. Mirenchi, E., Salvatore, A.: A non-Smooth two points boundary value problem on Riemannian manifolds. Ann. Mat. Pura Appl. **166**, 253–265 (1994)
7. Ferreira, O.P., Oliveira, P.R.: Proximal point algorithm on Riemannian manifolds. Optimization **51**, 257–270 (2002)

8. Bento, G.C., Ferreira, O.P., Oliveira, P.R.: Local convergence of the proximal point method for a special class of nonconvex functions on Hadamard manifolds. *Nonlinear Anal.* **73**, 564–572 (2010)
9. Ferreira, O.P.: Dini derivative and a characterization for Lipschitz and convex functions on Riemannian manifolds. *Nonlinear Anal.* **68**, 1517–1528 (2008)
10. Ferreira, O.P.: Proximal subgradient and a characterization of Lipschitz functions in Riemannian manifolds. *J. Math. Anal. Appl.* **313**, 587–597 (2006)
11. Li, S.L., Li, C., Liou, Y.C., Yao, J.C.: Existence of solutions for variational inequalities on Riemannian manifolds. *Nonlinear Anal.* **71**, 5695–5706 (2009)
12. Li, C., López, G., Martín-Márquez, V.: Monotone vector fields and the proximal point algorithm on Hadamard manifolds. *J. Lond. Math. Soc.* **79**, 663–683 (2009)
13. Li, C., Mordukhovich, B.S., Wang, J.H., Yao, J.C.: Weak sharp minima on Riemannian manifolds. *SIAM J. Optim.* **21**, 1523–1560 (2011)
14. Wang, J.H., López, G., Martín-Márquez, V., Li, C.: Monotone and accretive vector fields on Riemannian manifolds. *J. Optim. Theory Appl.* **146**, 691–708 (2010)
15. Azagra, D., Ferrera, J., López-Mesas, F.: Nonsmooth analysis and Hamilton–Jacobi equations on Riemannian manifolds. *J. Funct. Anal.* **220**, 304–361 (2005)
16. Azagra, D., Ferrera, J.: Proximal calculus on Riemannian manifolds, with applications to fixed point theory. *Mediterr. J. Math.* **2**, 437–450 (2005)
17. Ledyae, Yu.S., Zhu, Q.J.: Nonsmooth analysis on smooth manifolds. *Trans. Am. Math. Soc.* **359**, 3687–3732 (2007)
18. Hosseini, S., Pouryayevali, M.R.: Generalized gradients and characterization of epi-Lipschitz sets in Riemannian manifolds. *Nonlinear Anal.* **74**, 3884–3895 (2011)
19. Le, K., Motreanu, D.: Some properties of general minimization problems with constraints. *Set-Valued Anal.* **14**, 413–424 (2009)
20. Lang, S.: *Fundamentals of Differential Geometry*. Graduate Texts in Mathematics, vol. 191. Springer, New York (1999)
21. Clarke, F.H., Ledayaev, Yu.S., Stern, R.J., Wolenski, P.R.: *Nonsmooth Analysis and Control Theory*. Graduate Texts in Mathematics, vol. 178. Springer, New York (1998)
22. Borwein, J.M., Lewis, A.S.: *Convex Analysis and Nonlinear Optimization*. CMS Books in Mathematics. Springer, New York (2000)
23. Barani, A., Hosseini, S., Pouryayevali, M.R.: On the metric projection onto φ -convex subsets of Riemannian manifolds. *Rev. Mat. Complut.* (to appear). doi:[10.1007/s13163-011-0085-4](https://doi.org/10.1007/s13163-011-0085-4)