

Subgradient Projection Algorithms for Convex Feasibility on Riemannian Manifolds with Lower Bounded Curvatures

X. M. Wang · C. Li · J. C. Yao

Received: 8 October 2013 / Accepted: 3 April 2014 / Published online: 29 April 2014
© Springer Science+Business Media New York 2014

Abstract Under the assumption that the sectional curvature of the manifold is bounded from below, we establish convergence result about the cyclic subgradient projection algorithm for convex feasibility problem presented in a paper by Bento and Melo on Riemannian manifolds (J Optim Theory Appl 152, 773–785, 2012). If, additionally, we assume that a Slater type condition is satisfied, then we further show that, without changing the step size, this algorithm terminates in a finite number of iterations. Clearly, our results extend the corresponding ones due to Bento and Melo and, in particular, we solve partially the open problem proposed in the paper by Bento and Melo.

Keywords Convex feasibility problem · Cyclic subgradient projection algorithm · Riemannian manifold · Sectional curvature

Communicated by Sándor Zoltán Németh.

X. M. Wang · C. Li
Department of Mathematics, Zhejiang University, Hangzhou 310027, People's Republic of China
e-mail: cli@zju.edu.cn

X. M. Wang
e-mail: sci.xmwang@gzu.edu.cn

X. M. Wang
College of Science, Guizhou University, Guiyang 550025, People's Republic of China

J. C. Yao (✉)
Center for General Education, Kaohsiung Medical University, Kaohsiung 80702, Taiwan
e-mail: yaojc@kmu.edu.tw

J. C. Yao
Department of Mathematics, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia

Mathematics Subject Classification 90C25 · 68U05 · 65K05

1 Introduction

The convex feasibility problem (CFP, for short), which consists in finding a point in the intersection of finitely many closed and convex sets in Hilbert spaces, is a very broad and adaptive framework, and its study has a long and rich history in applied mathematics. The main approach to solve this problem is offered by the projection-like algorithms, which probably goes back to the work by von Neumann in the early 1930s for the special case, when there are only two sets. For details, we refer the reader to [1–10] and the references therein. For the special situation, where each convex set is given as a sublevel set of a convex function, Censor and Lent proposed in [11] a cyclic subgradient projection algorithm by using, instead of the orthogonal projections to the convex sets, the projections of the corresponding subgradients. As pointed out in [12], this special approach does not lose the generality because one can certainly choose the functions as the squared Euclidean distance functions to the involved convex sets to cover the general case. Since then, some authors have incorporated acceleration procedures in order to improve the speed of convergence; see, e.g., [2, 8, 12–15].

Recent interests are focused on extensions of some important notions and techniques in Hilbert spaces to Riemannian manifolds. The reason is that, as explained in [16], some optimization problems, arising in various applications, cannot be posed in linear spaces and require Riemannian manifold structures for their formalization and study; while some nonconvex and/or nonsmooth problems of constrained optimization in \mathbb{R}^n can be reduced to convex and smooth unconstrained optimization problems on appropriate Riemannian manifolds; see, for examples, [17–25]. For developments of optimization techniques on Riemannian manifolds, we refer the reader to [16, 26–37] and the bibliographies therein for various results, examples, discussions, and applications. Here, let us particularly mention the CFP on Riemannian manifolds with each convex set being given as a sublevel set of a convex function and the cyclic subgradient projection algorithm (i.e., Algorithm 3.1) for solving it; this algorithm was first proposed in [38], but the idea comes from the book of Udriste, that is, the general descent and/or gradient algorithms studied in [25, Chap. 7, §3 and §4] for finding a minimizer of a C^2 -function f on Riemannian manifolds. Convergence properties of the gradient algorithm have also been studied in [24, 30, 34, 37, 39]. The CFP can be reformulated as a convex minimization problem on Riemannian manifolds, but the induced convex function f on Riemannian manifolds is nonsmooth in general; hence, the known results in [25] and the literature mentioned above cannot be applied to solve the CFP on Riemannian manifolds considered here, nor establish the convergence properties of the cyclic subgradient projection algorithm.

In the special case, when the underlying Riemannian manifold is of non-negative sectional curvature, Bento and Melo proved in [38] that the cyclic subgradient projection algorithm converges to a solution of the CFP, and if, additionally, a Slater type condition is satisfied, then this algorithm has the finite termination property. Moreover, the following open problem was proposed there: How is the CFP solved on Riemannian manifolds with negative sectional curvature?

The main purpose of the present paper is to study the convergence problem of the cyclic subgradient projection algorithm on Riemannian manifolds with possible non-positive sectional curvature. More precisely, under the assumption that the underlying Riemannian manifold is of sectional curvature bounded from below, we show that, employing the same step sizes as in [38], any sequence generated by the cyclic subgradient projection algorithm converges to a solution of the CFP, and if a Slater type condition is additionally satisfied, then this algorithm terminates in a finite number of iterations. Our results extend the corresponding ones in [38], and, in particular, the open problem mentioned above is solved partially.

The paper is organized as follows. We recall some basic notions, notations, and properties of Riemannian geometry, together with some definitions and preliminary results on convex analysis on Riemannian manifolds in Sect. 2. In Sect. 3, the cyclic subgradient projection algorithm considered in [38] is introduced and some useful lemmas are shown. The main results of the present paper, including the convergence theorems and the finite termination theorems of this algorithm, are presented in Sects. 4 and 5, respectively.

2 Notations and Preliminaries

The notations used in the present paper are standard; and the readers are referred to some textbooks for more details, for example, [25,40,41].

Let M be a connected n -dimensional Riemannian manifold. We use ∇ to denote the Levi-Civita connection on M . Let $x \in M$, let $T_x M$ denote the tangent space at x to M . We denote by $\langle \cdot, \cdot \rangle_x$ the scalar product on $T_x M$ with the associated norm $\|\cdot\|_x$, where the subscript x is sometimes omitted. For $y \in M$, let $\gamma : [0, 1] \rightarrow M$ be a piecewise smooth curve joining x to y . Then, the arc-length of γ is defined by $l(\gamma) := \int_0^1 \|\gamma'(t)\| dt$, while the Riemannian distance from x to y is defined by $d(x, y) := \inf_{\gamma} l(\gamma)$, where the infimum is taken over all piecewise smooth curves $\gamma : [0, 1] \rightarrow M$ joining x to y . A vector field V is said to be parallel along γ if $\nabla_{\gamma'} V = 0$. In particular, for a smooth curve γ , if γ' is parallel along itself, then γ is called a geodesic, that is, a smooth curve γ is a geodesic if and only if $\nabla_{\gamma'} \gamma' = 0$. A geodesic $\gamma : [0, 1] \rightarrow M$ joining x to y is minimal if its arc-length equals its Riemannian distance between x and y . By the Hopf–Rinow theorem [41], (M, d) is a complete metric space, and there is at least one minimal geodesic joining x to y . The closed metric ball in M centered at the point $x \in M$ with radius $r > 0$ is denoted by $\mathbb{B}(x, r)$, i.e.,

$$\mathbb{B}(x, r) := \{y \in M : d(x, y) \leq r\}.$$

A complete simply connected Riemannian manifold of non-positive sectional curvature is called a *Hadamard manifold*. One important example of Hadamard manifold is the m -dimensional hyperbolic space \mathbb{H}^m , which is defined as follows:

$$\mathbb{H}^m := \{x = (x_1, \dots, x_{m+1}) \in \mathbb{R}^{m+1} : \langle x, x \rangle_{\mathbb{H}} = -1, x_{m+1} > 0\},$$

where $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ denotes the symmetric bilinear form (which is called the Lorentz metric on \mathbb{R}^{m+1}), defined by

$$(x, y)_{\mathbb{H}} := \sum_{i=1}^n x_i y_i - x_{n+1} y_{n+1} \quad \text{for any } x = (x_i), y = (y_i) \in \mathbb{R}^{m+1}.$$

Then, \mathbb{H}^m is a Hadamard manifold with sectional curvature -1 [20,27]. Let $\kappa < 0$ and let M_{κ}^m be the Hadamard manifold obtained from the m -dimensional hyperbolic space \mathbb{H}^m by multiplying the distance function by the constant $\frac{1}{\sqrt{|\kappa|}}$.

Consider two geodesic segments $\gamma_1, \gamma_2 : [0, +\infty[\rightarrow M$ emanating from p (i.e., $\gamma_1(0) = \gamma_2(0) = p$). Let $\angle_p(\gamma_1, \gamma_2)$ denote the angle between γ_1 and γ_2 at p , which is defined to be the angle between the tangent vectors $\gamma_1'(0)$ and $\gamma_2'(0)$. By [27, p. 173, Corollary 1A7.], $\angle_p(\gamma_1, \gamma_2)$ coincides with the Alexandrov angle. Recall that a geodesic triangle $\Delta(p_1 p_2 p_3)$ in M is a figure consisting of three points p_1, p_2, p_3 (the vertices of $\Delta(p_1 p_2 p_3)$) and three geodesic segments γ_i (the edges of $\Delta(p_1 p_2 p_3)$) that join p_{i-1} to p_{i+1} with $i = 1, 2, 3 \pmod{3}$. For each $i = 1, 2, 3 \pmod{3}$, the inner angle of $\Delta(p_1 p_2 p_3)$ at p_i is denoted by $\angle(p_{i-1} p_i p_{i+1})$, which equals $\angle_{p_i}(-\gamma_{i-1}, \gamma_{i+1})$. The following proposition is known in [40, p. 138].

Proposition 2.1 *Let $\kappa < 0$ and let $\Delta(p_1 p_2 p_3)$ be a geodesic triangle in M_{κ}^2 . Then, the following “law of cosines” holds*

$$\begin{aligned} \cosh(\sqrt{|\kappa|}l_2) &= \cosh(\sqrt{|\kappa|}l_1) \cosh(\sqrt{|\kappa|}l_3) \\ &\quad - \sinh(\sqrt{|\kappa|}l_1) \sinh(\sqrt{|\kappa|}l_3) \cos \angle(p_1 p_2 p_3), \end{aligned}$$

where $l_i = d(p_{i-1}, p_{i+1})$ for each $i = 1, 2, 3 \pmod{3}$.

Another important tool that will be used is the comparison theorem described in Proposition 2.2 below, which is known in [40, p. 161, Theorem 4.2].

Following [40, p. 161], a generalized geodesic hinge $\Lambda(p; \gamma, \tau)$ in M is a figure consisting of a point $p \in M$ (the vertex of the hinge) and two geodesic segments γ, τ (the edges of the hinge) emanating from p with one being minimal. Moreover, a hinge $\Lambda(\bar{p}; \bar{\gamma}, \bar{\tau})$ in M_{κ}^2 is called a comparison hinge of $\Lambda(p; \gamma, \tau)$ if it satisfies

$$l(\bar{\gamma}) = l(\gamma), l(\bar{\tau}) = l(\tau), \quad \text{and} \quad \angle_{\bar{p}}(\bar{\gamma}, \bar{\tau}) = \angle_p(\gamma, \tau).$$

Proposition 2.2 *Let $\kappa < 0$ and suppose that M is a complete Riemannian manifold of sectional curvature bounded from below by κ . Let $\Lambda(p; \gamma, \tau)$ be a generalized geodesic hinge in M . Then, there exists a comparison hinge $\Lambda(\bar{p}; \bar{\gamma}, \bar{\tau})$ in M_{κ}^2 such that $d(q_{\gamma}, q_{\tau}) \leq d(\bar{q}_{\bar{\gamma}}, \bar{q}_{\bar{\tau}})$, where $q_{\gamma}, q_{\tau}, \bar{q}_{\bar{\gamma}},$ and $\bar{q}_{\bar{\tau}}$ denote the end points of $\gamma, \tau, \bar{\gamma},$ and $\bar{\tau}$, respectively.*

Let Γ_{xy} denote the set of all geodesics $\gamma : [0, 1] \rightarrow M$ with $\gamma(0) = x$ and $\gamma(1) = y$. Then, Γ_{xy} is nonempty for all $x, y \in M$. Consider next a real-valued function $f : M \rightarrow \mathbb{R}$. We say that f is convex on M if, for any $x, y \in M$ and any $\gamma \in \Gamma_{xy}$, the composition $(f \circ \gamma) : [0, 1] \rightarrow \mathbb{R}$ is a convex function on $[0, 1]$, i.e.,

$$f(\gamma(t)) \leq (1 - t)f(x) + tf(y) \quad \text{for all } t \in [0, 1].$$

For the remainder, we always assume that $f : M \rightarrow \mathbb{R}$ is convex on M . Then, as pointed out in [21, p. 642] (see also [25, p. 70]), f is Lipschitz continuous on any compact subset of M . This, in particular, implies that f is continuous on M . Let $x \in M$. We define the subdifferential $\partial f(x)$ of f at x by

$$\partial f(x) := \{g \in T_x M : f(y) \geq f(x) + \langle g, \gamma'(0) \rangle, \quad \forall y \in M \text{ and } \gamma \in \Gamma_{xy}\}.$$

Then, $\partial f(x)$ is a nonempty, closed convex set in the space $T_x M$. Furthermore, by the Lipschitz continuity, one sees that

$$\partial f(\cdot) \text{ is uniformly bounded on any bounded subset of } M \tag{1}$$

(see also [38]).

We end this section with the notions of the distance function and the projection associated to a subset Q of M . The distance function $d_Q(\cdot)$ of Q is defined by

$$d_Q(x) := \inf_{y \in Q} d(x, y) \quad \text{for each } x \in M;$$

while the projection $P(\cdot|Q)$ on Q is defined by

$$P(x|Q) := \{y \in Q : d(x, y) = d_Q(x)\} \quad \text{for each } x \in M.$$

Note that, in general, $P(\cdot|Q)$ is a set-valued map even if Q is convex in the sense that, for any $y, z \in Q$, any minimal geodesic joining y to z is in Q . For example, consider $M := \mathbb{S}^2$, the two dimensional sphere, and

$$Q := \{(t_1, t_2, t_3) \in M : \text{each } t_i \geq 0\}.$$

Let $x_0 := (0, 0, -1)$. Then, $d_Q(x_0) = \frac{\pi}{2}$ and $P(x_0|Q) = \{(t_1, t_2, 0) \in Q\}$.

3 Algorithm and Auxiliary Results

Let $I := \{1, 2, \dots, m\}$, where $m \in \mathbb{N}$, and let $f_i : M \rightarrow \mathbb{R}$ be a convex function for each $i \in I$. Consider a family of closed and convex subsets $\{Q_i : i \in I\}$ in M with each Q_i given by

$$Q_i := \{x \in M : f_i(x) \leq 0\}.$$

The CFP on Riemannian manifold M considered here consists in finding a point x^* satisfying

$$x^* \in Q := \bigcap_{i=1}^m Q_i.$$

Let $f : M \rightarrow \mathbb{R}$ be the sup-function of $\{f_i : i \in I\}$ defined by

$$f(x) := \max_{i \in I} f_i(x) \quad \text{for each } x \in M.$$

Then, f is convex on M , and the solution set Q is re-expressed as

$$Q = \{x \in M : f(x) \leq 0\}.$$

We assume throughout the whole paper that $Q \neq \emptyset$.

The following algorithm was proposed in [38] by Bento et al. for solving the CFP on Riemannian manifolds, which is actually suggested by the book of Udriste (see [25, Chap. 7, §3 and §4]); in particular, this algorithm coincides with “the gradient method” described in [25, p. 263] in the case when f is smooth.

Algorithm 3.1 Step 1. Select $x^0 \in M$ and put $k := 0$.

Step 2. If $f(x^k) \leq 0$, then stop; otherwise, go to Step 3.

Step 3. Put $i_k := (k \bmod m) + 1$.

Step 4. If $f_{i_k}(x^k) \leq 0$, then set $x^{k+1} := x^k$; otherwise, go to Step 5.

Step 5. Select $g^k \in \partial f_{i_k}(x^k)$, $t_k \in (0, +\infty)$ and construct the geodesic γ_k such that

$$\gamma_k(0) = x^k \quad \text{and} \quad \gamma'_k(0) = -\frac{g^k}{\|g^k\|}. \tag{2}$$

Step 6. Set $x^{k+1} := \gamma_k(t_k)$.

Step 7. Replace k by $k + 1$ and go to Step 2.

Remark 3.1 (i) Note that the selected $g^k \in \partial f_{i_k}(x^k)$ in Step 5 is nonzero (otherwise, x^k is a minimum of f_{i_k} and then, $f_{i_k}(x^k) \leq 0$). This means that Algorithm 3.1 is well defined.

(ii) The sequence $\{t_k\}$ is called the step size sequence. By Algorithm 3.1, we have that

$$d(x^k, x^{k+1}) \leq t_k \quad \text{for each } k \in \mathbb{N}. \tag{3}$$

To analyze the convergence of Algorithm 3.1, we first establish the basic inequality on Riemannian manifolds. For simplicity, we define the function $\tilde{h} :]0, \infty[\rightarrow \mathbb{R}$ by

$$\tilde{h}(t) := \frac{\tanh t}{t} \quad \text{for any } t \in]0, \infty[.$$

It is easy to verify that \tilde{h} is strictly decreasing on $]0, \infty[$.

For the remainder of the paper, we assume that the Riemannian manifold M is of sectional curvature bounded from below by some negative number κ .

Lemma 3.1 *Let $\{x^k\}$ be a sequence generated by Algorithm 3.1. Let $z \in M$ and let $k \geq 0$ be such that $f_{i_k}(x^k) > 0$. Then, the following inequality holds:*

$$\begin{aligned} \cosh\left(\sqrt{|\kappa|}d(x^{k+1}, z)\right) &\leq \cosh\left(\sqrt{|\kappa|}d(x^k, z)\right) \\ &\quad + \sqrt{|\kappa|} \cosh\left(\sqrt{|\kappa|}d(x^k, z)\right) \sinh(\sqrt{|\kappa|}t_k) \\ &\quad \cdot \left(\frac{t_k}{2} - \hbar\left(\sqrt{|\kappa|}d(x^k, z)\right) \frac{f_{i_k}(x^k) - f_{i_k}(z)}{\|g^k\|}\right). \end{aligned} \tag{4}$$

Proof Without loss of generality, we assume that $\kappa = -1$. Take $z \in M$ and let $\tau_k : [0, 1] \rightarrow M$ be a minimal geodesic segment joining x^k and z with $\tau_k(0) = x^k$ and $\tau_k(1) = z$. Let $\gamma_k : [0, t_k] \rightarrow M$ be the geodesic segment defined by Algorithm 3.1. Then,

$$l(\gamma_k) = t_k, \quad l(\tau_k) = d(x^k, z). \tag{5}$$

Consider the generalized geodesic hinge $\Lambda(x^k; \gamma_k, \tau_k)$ in M , and let $\alpha := \angle_{x^k}(\gamma_k, \tau_k)$. Then α coincides with the angle between the tangent vectors $\gamma'_k(0)$ and $\tau'_k(0)$ in the tangent space $T_{x^k}M$. This, together with (2), implies that

$$\langle g^k, \tau'_k(0) \rangle = -\|g^k\| \cdot d(x^k, z) \cos \alpha. \tag{6}$$

By Proposition 2.2, there exists a comparison geodesic hinge $\Lambda(\bar{x}^k; \bar{\gamma}_k, \bar{\tau}_k)$ of the hinge $\Lambda(x^k; \gamma_k, \tau_k)$ in M_{-1}^2 such that

$$d(x^{k+1}, z) \leq d(\bar{x}^{k+1}, \bar{z}), \tag{7}$$

where \bar{x}^{k+1} and \bar{z} are the end points of $\bar{\gamma}_k$ and $\bar{\tau}_k$, respectively. Consider the geodesic triangle $\Delta(\bar{x}^k \bar{x}^{k+1} \bar{z})$ in M_{-1}^2 . By the definition of the comparison geodesic hinge and using (5), we have that

$$d(\bar{x}^k, \bar{x}^{k+1}) = t_k, \quad d(\bar{x}^k, \bar{z}) = d(x^k, z), \quad \text{and} \quad \angle_{\bar{x}^k}(\bar{\gamma}_k, \bar{\tau}_k) = \alpha.$$

Then, applying the “law of cosines” to the geodesic triangle $\Delta(\bar{x}^k \bar{x}^{k+1} \bar{z})$ in M_{-1}^2 , we get

$$\cosh d(\bar{x}^{k+1}, \bar{z}) = \cosh d(x^k, z) \cosh t_k - \sinh d(x^k, z) \sinh t_k \cos \alpha.$$

Combing this and (7) yields that

$$\cosh d(x^{k+1}, z) \leq \cosh d(x^k, z) \cosh t_k - \sinh d(x^k, z) \sinh t_k \cos \alpha. \tag{8}$$

By elementary calculus, one can easily verify that

$$\cosh t \leq 1 + \frac{t}{2} \sinh t \quad \text{for any } t \geq 0.$$

It follows from (8) that

$$\begin{aligned} \cosh d(x^{k+1}, z) &\leq \cosh d(x^k, z) \left(1 + \frac{t_k \sinh t_k}{2} \right) - \sinh d(x^k, z) \sinh t_k \cos \alpha \\ &= \cosh d(x^k, z) + \cosh d(x^k, z) \sinh t_k \left[\frac{t_k}{2} - \tanh d(x^k, z) \cos \alpha \right]. \end{aligned} \tag{9}$$

Recall that $g^k \in \partial f_{i_k}(x^k)$. By equality (6), one has that

$$f_{i_k}(z) \geq f_{i_k}(x^k) + \langle g^k, \tau'_k(0) \rangle = f_{i_k}(x^k) - \|g^k\| \cdot d(x^k, z) \cos \alpha;$$

hence,

$$-\cos \alpha \leq -\frac{f_{i_k}(x^k) - f_{i_k}(z)}{d(x^k, z) \|g^k\|}.$$

Substituting this into (9) and recalling the definition of the function \tilde{h} , we see that (4) holds, and the proof is complete. □

Set

$$f^* := \inf_{x \in M} f(x).$$

Then, $f^* \leq 0$ by the nonemptiness assumption of Q . Throughout the whole paper, let $\{f_i^* : i \in I\} \subseteq \mathbb{R}$ satisfy

$$f^* \leq f_i^* \leq 0 \text{ for each } i \in I$$

and choose the step size sequence $\{t_k\}$ as follows:

$$t_k := \begin{cases} \alpha_k \frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|}, & \text{if } f_{i_k}(x^k) > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

where $i_k = (k \bmod m) + 1$. Set

$$C := \{x \in M : f_i(x) \leq f_i^*, \forall i \in I\}.$$

Note that $C \subseteq Q$ and the equality holds if each $f_i^* = 0$. Furthermore, let $x^0 \in M$ and write $r_0 := d_C(x^0)$.

Lemma 3.2 *Let $x^0 \in M$ and suppose that*

$$0 < \inf_k \alpha_k \leq \sup_k \alpha_k < 2\tilde{h}(2\sqrt{|\kappa|r_0}). \tag{11}$$

Let $\{x^k\}$ be a sequence generated by Algorithm 3.1 with initial point x^0 . Set

$$\tilde{C} := C \cap \left(\bigcup_{\omega \in P(x^0|C)} \mathbb{B}(\omega, r_0) \right)$$

and let $z \in \tilde{C}$. Then, the following estimate holds for each $k \in \mathbb{N}$:

$$d^2(x^{k+1}, z) \leq d^2(x^k, z) - \frac{\sigma t_k^2}{4\beta}, \tag{12}$$

where

$$\beta := \frac{\sinh(2\sqrt{|\kappa|r_0})}{4\sqrt{|\kappa|r_0}} \quad \text{and} \quad \sigma := \min \left\{ \inf_k \alpha_k, 2 - \frac{\sup_k \alpha_k}{\tilde{h}(2\sqrt{|\kappa|r_0})} \right\}.$$

Proof As in the proof of Lemma 3.1, we also assume that $\kappa = -1$. By the definition of σ , assumption (11) implies that $\sigma \in]0, 1[$ and

$$\sigma \leq \alpha_k \leq (2 - \sigma)\tilde{h}(2r_0) \quad \text{for each } k \in \mathbb{N}. \tag{13}$$

Without loss of generality, we assume that $f_{i_k}(x^k) > 0$ (otherwise, $x^{k+1} = x^k$ and (12) holds trivially by (10)). Then,

$$f_{i_k}(x^k) - f_{i_k}(z) \geq f_{i_k}(x^k) - f_{i_k}^* > 0 \tag{14}$$

(noting that $z \in \tilde{C}$). By the choice of t_k and noting (13), one sees that

$$\frac{t_k}{2} = \frac{\alpha_k}{2} \frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|} \leq \left[\tilde{h}(2r_0) - \frac{\sigma \tilde{h}(2r_0)}{2} \right] \frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|}.$$

This, together with (14), implies that

$$\begin{aligned} \frac{t_k}{2} - \tilde{h} \left(d(x^k, z) \right) \frac{f_{i_k}(x^k) - f_{i_k}(z)}{\|g^k\|} \\ \leq \left[\tilde{h}(2r_0) - \tilde{h} \left(d(x^k, z) \right) - \frac{\sigma \tilde{h}(2r_0)}{2} \right] \frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|}. \end{aligned}$$

Applying (4), we get that

$$\begin{aligned} \cosh d(x^{k+1}, z) &\leq \cosh d(x^k, z) \\ &+ \cosh d(x^k, z) \sinh t_k \left[\tilde{h}(2r_0) - \tilde{h} \left(d(x^k, z) \right) - \frac{\sigma \tilde{h}(2r_0)}{2} \right] \frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|}. \end{aligned} \tag{15}$$

Since $\sigma > 0$, we have that

$$\begin{aligned} \cosh d(x^{k+1}, z) &\leq \cosh d(x^k, z) \\ &+ \cosh d(x^k, z) \sinh t_k \left[\hbar(2r_0) - \hbar(d(x^k, z)) \right] \frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|}. \end{aligned} \tag{16}$$

Below, we use mathematical induction to show that

$$d(x^k, z) \leq 2r_0 \tag{17}$$

holds for each $k \in \mathbb{N}$. By the choice of z and the definition of \tilde{C} , we can choose $\omega \in P(x^0|C)$ such that $d(z, \omega) \leq r_0$. Thus, (17) is clear for $k = 0$, because

$$d(x^0, z) \leq d(x^0, \omega) + d(\omega, z) \leq 2r_0.$$

Now, suppose that (17) is valid for $k = n$, that is,

$$d(x^n, z) \leq 2r_0. \tag{18}$$

Then,

$$\hbar(2r_0) \leq \hbar(d(x^n, z)) \tag{19}$$

(as \hbar is decreasing on $]0, +\infty[$). It follows from (16) that

$$\cosh d(x^{n+1}, z) \leq \cosh d(x^n, z);$$

hence, $d(x^{n+1}, z) \leq d(x^n, z)$. This, together with (18), yields that (17) holds for $k = n + 1$. Therefore, (17) is checked for all $k \in \mathbb{N}$, and so (19) holds for all $n \in \mathbb{N}$. Thus, combining (15) and (19) (with k in place of n), one concludes that

$$\cosh d(x^{k+1}, z) \leq \cosh d(x^k, z) - \cosh d(x^k, z) \sinh t_k \frac{\sigma \hbar(2r_0) (f_{i_k}(x^k) - f_{i_k}^*)}{2\|g^k\|}. \tag{20}$$

Recalling (17), one can use Taylor expansion for the function $\cosh(\cdot)$ to verify the following estimate:

$$\cosh d(x^{k+1}, z) - \cosh d(x^k, z) \geq \beta \left(d^2(x^{k+1}, z) - d^2(x^k, z) \right),$$

where $\beta := \frac{\sinh(2r_0)}{4r_0}$. This, together with (20), implies that

$$d^2(x^{k+1}, z) \leq d^2(x^k, z) - \frac{\sigma}{2\beta} \hbar(2r_0) \cosh d(x^k, z) \sinh t_k \frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|}.$$

Since $\sinh t_k > t_k$ and $\cosh d(x^k, z) \geq 1$, it follows that

$$d^2(x^{k+1}, z) \leq d^2(x^k, z) - \frac{\sigma \hbar(2r_0)t_k (f_{i_k}(x^k) - f_{i_k}^*)}{2\beta \|g^k\|}. \tag{21}$$

On the other hand, using (10) and (13), we see that

$$\frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|} = \frac{t_k}{\alpha_k} \geq \frac{t_k}{(2 - \sigma)\hbar(2r_0)} \geq \frac{t_k}{2\hbar(2r_0)}.$$

This, together with (21), yields (12). The proof is complete. □

Remark 3.2 Suppose that assumption (11) holds. Then, the following assertions hold by Lemma 3.2:

- (i) The sequence $\{x^k\} \subseteq \mathbb{B}(\omega, r_0)$ for any point $\omega \in P(x^0|C)$, and is Fejér convergent to \tilde{C} in the sense that

$$d(x^{k+1}, z) \leq d(x^k, z) \quad \text{for each } k \in \mathbb{N} \quad \text{and } z \in \tilde{C}. \tag{22}$$

- (ii) By (1), we have that

$$L := \sup\{\|g\| : g \in \partial f_i(x^k), (i, k) \in I \times \mathbb{N}\} < +\infty. \tag{23}$$

- (iii) By (3) and (12), we have that

$$\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) = \lim_{k \rightarrow \infty} t_k = 0. \tag{24}$$

4 Convergence Results

We assume in this section that $f_i^* = 0$ for each $i \in I$ in the step size sequence $\{t_k\}$, that is Algorithm 3.1 employs the following step sizes

$$t_k := \begin{cases} \alpha_k \frac{f_{i_k}(x^k)}{\|g^k\|}, & \text{if } f_{i_k}(x^k) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then, one has that $C = Q$. The main result of this section is as follows.

Theorem 4.1 *Let $x^0 \in M$ and suppose that assumption (11) holds. Then, any sequence $\{x^k\}$ generated by Algorithm 3.1 with initial point x^0 converges to a point $x^\infty \in Q$.*

Proof By assumption (11), Remark 3.2 is applicable. Then, $\{x^k\}$ satisfies (22), (24), and that

$$d(x^k, \omega) \leq r_0 \quad \text{for any } \omega \in P(x^0|C). \tag{25}$$

Thus, there exists a subsequence $\{x^{k_j}\}$ of $\{x^k\}$ such that $x^{k_j} \rightarrow x^\infty \in M$. Below, we verify that $x^\infty \in Q$, that is, $f_l(x^\infty) \leq 0$ for each $l \in I$. To do this, recall that $i_{k_j} = (k \bmod m) + 1$ and so $i_k \in I$. Thus, without loss of generality, we assume that $i_{k_j} = m$ for each j . Let $l \in I$ and consider the subsequence $\{x^{k_j+l}\}_{j=0}^\infty$. Then,

$$x^{k_j+l} \rightarrow x^\infty \quad \text{and} \quad i_{k_j+l} = l, \tag{26}$$

where the first assertion is true by (24). Recalling the choice of t_k , we have that, if $f_l(x^{k_j+l}) > 0$, then

$$t_{k_j+l} = \alpha_{k_j+l} \frac{f_l(x^{k_j+l})}{\|g^{k_j+l}\|} \geq \frac{\sigma}{L} f_l(x^{k_j+l}),$$

where $L > 0$ is given by (23); hence,

$$f_l(x^{k_j+l}) \leq \frac{L}{\sigma} t_{k_j+l}. \tag{27}$$

This inequality holds trivially if $f_l(x^{k_j+l}) \leq 0$. Thus, (27) holds for each $j \in \mathbb{N}$. Since f_l is continuous on M , it follows from (26) that

$$f_l(x^\infty) = \lim_{j \rightarrow \infty} f_l(x^{k_j+l}) \leq 0.$$

Therefore, $x^\infty \in Q$ as $l \in I$ is arbitrary. Moreover, by (25), we get that $d(x^\infty, \omega) \leq r_0$, and so $x^\infty \in \tilde{C}$ (noting that $C = Q$). Applying (22), one concludes that $\{x^k\}$ converges to x^∞ and the proof is complete. □

As a consequence of Theorem 4.1, we get the following corollary, which was proved in [38].

Corollary 4.1 *Suppose that the sectional curvature of M is non-negative and that*

$$0 < \inf_k \alpha_k \leq \sup_k \alpha_k < 2. \tag{28}$$

Then, any sequence $\{x^k\}$ generated by Algorithm 3.1 with initial point $x^0 \in M$ converges to a point $x^\infty \in Q$.

Proof Let $x^0 \in M$. Recall that function \tilde{h} is continuous on $]0, \infty[$ and $\tilde{h}(0) = 1$. Then, by assumption (28), one can choose $\kappa < 0$ such that assumption (11) holds. Furthermore, by assumption, it is clear that the sectional curvature of M is bounded from below by κ . Thus, Theorem 4.1 is applicable to complete the proof. □

5 Finite Convergence

Recall that the step size sequence is given as in (10):

$$t_k := \begin{cases} \alpha_k \frac{f_{i_k}(x^k) - f_{i_k}^*}{\|g^k\|}, & \text{if } f_{i_k}(x^k) > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $i_k = (k \bmod m) + 1$ and $f^* \leq f_i^* \leq 0$ for each $i \in I$. We suppose in this section that the CFP on M satisfies the Slater condition:

$$f^* < 0,$$

and that $\{f_i^* : i \in I\}$ satisfies

$$f^* \leq f_i^* < 0 \quad \text{for each } i \in I.$$

The following theorem shows that Algorithm 3.1 terminates in a finite number of iterations.

Theorem 5.1 *Let $x^0 \in M$ and suppose that assumption (11) holds. Then, any sequence $\{x^k\}$ generated by Algorithm 3.1 with initial point x^0 terminates in a finite number of iterations.*

Proof Suppose on the contrary that Algorithm 3.1 generates an infinite sequence $\{x^k\}$, which satisfies that

$$\sup_{i \in I} f_i(x_k) > 0 \quad \text{for each } k \in \mathbb{N}.$$

This means that there exists a subsequence $\{x^{k_j}\}$ such that

$$f_{i_{k_j}}(x^{k_j}) > 0 \quad \text{for each } j \in \mathbb{N}, \tag{29}$$

where $i_{k_j} = (k_j \bmod m) + 1$ as before. Moreover, by assumption (11), Remark 3.2 is applicable. Therefore, $\{x^k\}$ is bounded and (24) holds. Thus, without loss of generality, we may further assume that $\{x^{k_j}\}$ converges to some point $x^\infty \in M$ and $i_{k_j} = 1$ for each j (as each i_{k_j} is in the finite set I). Since f_1 is continuous on M , it follows from (29) that

$$f_1(x^\infty) = \lim_{j \rightarrow +\infty} f_1(x^{k_j}) \geq 0. \tag{30}$$

Furthermore, we have that

$$t_{k_j} = \alpha_{k_j} \frac{f_1(x^{k_j}) - f_1^*}{\|g^{k_j}\|} \geq \frac{\sigma}{L} (f_1(x^{k_j}) - f_1^*) \quad \text{for each } j \in \mathbb{N},$$

where $L > 0$ is given by (23). Passing to the limit and making use of (24), we get that $f_1(x^\infty) - f_1^* \leq 0$; hence, $f_1(x^\infty) \leq f_1^* < 0$, which contradicts (30). Hence, the proof is complete. \square

With a similar argument that we did for Corollary 4.1, but using Theorem 5.1 in place of Theorem 4.1, we can show the following corollary, which was proved in [38, Theorem 5.1].

Corollary 5.1 *Suppose that the sectional curvature of M is non-negative and that*

$$0 < \inf_k \alpha_k \leq \sup_k \alpha_k < 2.$$

Then, any sequence $\{x^k\}$ generated by Algorithm 3.1 with initial point $x^0 \in M$ terminates in a finite number of iterations.

6 Conclusions

By using a different approach, we generalize the results in [38] for the Riemannian manifolds with non-negative sectional curvature to those with sectional curvature bounded from below. In particular, the open problem proposed in [38] has been solved partially. It is still an open problem how to establish similar convergence results on Riemannian manifolds without the assumption that the sectional curvature is bounded from below. Note that, in the Hilbert space setting, if the CFP satisfies a Slater type condition, then the cyclic subgradient projection algorithm is linearly convergent [1]. Our future work is to establish the linear convergence result about the algorithm on Riemannian manifolds under a Slater type condition. Moreover, recalling that every complete, connected Riemannian manifold is a special geodesic metric space (as pointed out in [27]), we expect that our approach in the present paper can be used to study similar problems in the setting of geodesic spaces (with curvature bounded from below).

Acknowledgments Research of the second author is supported in part by the National Natural Science Foundation of China (Grants 11171300, 11371325) and by Zhejiang Provincial Natural Science Foundation of China (Grant LY13A010011). Research of the third author was partially supported by the National Science Council of Taiwan under Grant NSC 99-2115-M-037-002-MY3.

References

1. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. *SIAM Rev.* **38**(3), 367–426 (1996)
2. Bauschke, H.H., Combettes, P.L., Kruk, S.G.: Extrapolation algorithm for affine–convex feasibility problems. *Numer. Algorithms* **41**, 239–274 (2006)
3. Censor, Y.: *Mathematical Optimization for the Inverse Problem of Intensity-Modulated Radiation Therapy*. Medical Physics Publisher, Madison (2003)
4. Censor, Y., Altschuler, M.D., Powlis, W.D.: On the use of Cimmino’s simultaneous projections method for computing a solution of the inverse problem in radiation therapy treatment planning. *Inverse Probl.* **4**, 607–623 (1988)

5. Censor, Y., Zenios, S.A.: *Parallel Optimization: Theory, Algorithms, and Applications*. Oxford University Press, New York (1997)
6. Combettes, P.L.: The foundations of set theoretic estimation. *Proc. IEEE* **81**, 182–208 (1993)
7. Combettes, P.L.: The convex feasibility problem in image recovery. *Adv. Imaging Electron Phys.* **95**, 155–270 (1996)
8. Combettes, P.L.: Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections. *IEEE Trans. Image Process.* **6**(4), 493–506 (1997)
9. Herman, G.: *Image Reconstruction from Projections. The Fundamentals of Computerized Tomography*. Academic Press, New York (1980)
10. Marks, L.D., Sinkler, W., Landree, E.: A feasible set approach to the crystallographic phase problem. *Acta Crystallogr.* **55**(4), 601–612 (1999)
11. Censor, Y., Lent, A.: A Cyclic Subgradient Projections Method for the Convex Feasibility Problems. Technical Report. University of Haifa, Israel (1980)
12. Butnariu, D., Censor, Y., Gurfil, P., Hadar, E.: On the behavior of subgradient projections methods for convex feasibility problems in Euclidean spaces. *SIAM J. Optim.* **19**, 786–807 (2008)
13. Kiwiel, K.C.: The efficiency of subgradient projection. *SIAM J. Control Optim.* **34**, 677–697 (1996)
14. Kiwiel, K.C.: The efficiency of subgradient projection methods for convex optimization I: General level methods. *SIAM J. Control Optim.* **34**, 660–676 (1996)
15. dos Santos, L.T.: A parallel subgradient method for the convex feasibility problem. *J. Comput. Appl. Math.* **18**, 307–320 (1987)
16. Li, C., Mordukhovich, B.S., Wang, J.H., Yao, J.C.: Weak sharp minima on Riemannian manifolds. *SIAM J. Optim.* **21**(4), 1523–1560 (2011)
17. Rapsák, T.: *Smooth Nonlinear Optimization in \mathbb{R}^n . Nonconvex Optimization and its Applications*, vol. 19. Kluwer, Dordrecht (1997)
18. Adler, R., Dedieu, J.P., Margulies, J., Martens, M., Shub, M.: Newton’s method on Riemannian manifolds and a geometric model for the human spine. *IMA J. Numer. Anal.* **22**, 359–390 (2002)
19. Burke, J.V., Lewis, A., Overton, M.: Optimal stability and eigenvalue multiplicity. *Found. Comput. Math.* **1**, 205–225 (2001)
20. Ferreira, O.P., Pérez, L.R.L., Nemeth, S.Z.: Singularities of monotone vector fields and an extragradient-type algorithm. *J. Glob. Optim.* **31**(1), 133–151 (2005)
21. Greene, R., Wu, H.: On the subharmonicity and plurisubharmonicity of geodesically convex functions. *Indiana Univ. Math. J.* **22**, 641–653 (1972)
22. Mahony, R.E.: The constrained Newton method on a Lie group and the symmetric eigenvalue problem. *Linear Algebr. Appl.* **248**, 67–89 (1996)
23. Miller, S.A., Malick, J.: Newton methods for nonsmooth convex minimization: connections among U-Lagrangian, Riemannian Newton and SQP methods. *Math. Program.* **104**, 609–633 (2005)
24. Smith, S.T.: *Geometric Optimization Methods for Adaptive Filtering*. Division of Applied Sciences. PhD Thesis, Harvard University, Cambridge, MA (1993)
25. Udriste, C.: *Convex Functions and Optimization Methods on Riemannian Manifolds. Mathematics and Its Applications*, vol. 297. Kluwer, Dordrecht (1994)
26. Absil, P.A., Baker, C.G., Gallivan, K.A.: Trust-region methods on Riemannian manifolds. *Found. Comput. Math.* **7**(3), 303–330 (2007)
27. Bridson, M., Haefliger, A.: *Metric Spaces of Non-positive Curvature*. Springer, Berlin (1999)
28. Ferreira, O.P., Oliveira, P.R.: Subgradient algorithm on Riemannian manifolds. *J. Optim. Theory Appl.* **97**(1), 93–104 (1998)
29. Ferreira, O.P., Oliveira, P.R.: Proximal point algorithm on Riemannian Manifold. *Optimization* **51**(2), 257–270 (2002)
30. Gabay, D.: Minimizing a differentiable function over a differential manifold. *J. Optim. Theory Appl.* **37**(2), 177–219 (1982)
31. 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)
32. Li, C., Wang, J.H.: Newton’s method for sections on Riemannian manifolds: generalized covariant α -theory. *J. Complex.* **24**, 423–451 (2008)
33. Li, C., Yao, J.C.: Variational inequalities for set-valued vector fields on Riemannian manifolds: convexity of the solution set and the proximal point algorithm. *SIAM J. Control Optim.* **50**(4), 2486–2514 (2012)
34. Luenberger, D.G.: The gradient projection method along geodesics. *Manag. Sci.* **18**, 620–631 (1972)

35. Wang, J.H., Huang, S.C., Li, C.: Extended Newton's algorithm for mappings on Riemannian manifolds with values in a cone. *Taiwan J. Math.* **13**, 633–656 (2009)
36. 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)
37. Yang, Y.: Globally convergent optimization algorithms on Riemannian manifolds: uniform framework for unconstrained and constrained optimization. *J. Optim. Theory Appl.* **132**(2), 245–265 (2007)
38. Bento, G.C., Melo, J.G.: Subgradient algorithm for convex feasibility on Riemannian manifolds. *J. Optim. Theory Appl.* **152**(3), 773–785 (2012)
39. da Cruz Neto, J.X., de Lima, L., Oliverira, P.R.: Geodesic algorithms in Riemannian geometry. *Balkan J. Geom. Appl.* **3**, 89–100 (1998)
40. Sakai, T.: *Riemannian Geometry*. Translations of Mathematical Monographs, vol. 149. American Mathematical Society, Providence (1996)
41. do Carmo, M.P.: *Riemannian Geometry*. Birkhauser, Boston (1992)