On the convergence rate of an inexact proximal point algorithm for quasiconvex minimization on Hadamard manifolds

N. Baygorrea, E.A. Papa Quiroz and N. Maculan

Federal University of Rio de Janeiro
COPPE-PESC-UFRJ
PO Box 68511, Rio de Janeiro, CEP 21941-972, Brazil.
{nbaygorrea,erik,maculan}@cos.ufrj.br

Abstract

In this paper we present a rate of convergence analysis of an inexact proximal point algorithm to solve minimization problems for quasiconvex objective functions on Hadamard manifolds. We prove that under natural assumptions the sequence generated by the algorithm converges linearly or superlinearly to a critical point of the problem.

Keywords: Proximal point method, quasiconvex function, Hadamard manifolds, nonsmooth optimization, abstract subdifferential, convergence rate.

1 Introduction

The initial work on the proximal point minimization algorithm is due to Martinet [11]. Then that method was extended for finding a zero of an arbitrary maximal monotone operator, in Hilbert spaces, by Rockafellar [16]. In that paper, the convergence of the method was establised under several criteria with conditions amenable to implementation. Moreover, convergence rate is shown to be linear ou superlinear (depending of the positive proximal parameters) whenever the inverse of the operator is Lipschitz continuous at 0. It turns out to be very natural in applications to convex programming. After that, for minimization problems it was shown in Güler[9], that the sequence of the objective function values converges to the optimal value of the minimization problem with a complexity of $O(1/k)$ under minimal assumptions on the regularized parameter.

Since the introduction of the proximal point algorithm by Martinet[11], there have been a geat interest of the optimization community to study the proximal point algorithm in different spaces because it can be viewed as a powerful tool for the regularization of ill-posed convex optimization problems, as a standard tool for solving nonsmooth problems of large-scale, separable optimization problems and its role in multiplier methods based on duality.

Moreover, it must be admitted the importance of the generalization of this algorithm from linear spaces to differentiable manifolds. In particular, on Riemannian manifolds. This importance is based on the fact that, for instance, some nonconvex constrained optimization problems can be solved after being written as convex optimization problems on Hadamard manifolds, see for example da Cruz Neto et al.[5], Ferreira et al.[8], Rapcsák[17] and Udriste[21]. Indeed, the proximal point algorithm, in the setting of Riemannian manifold, has been introduced by Ferreira and Oliveira[7] for solving convex minimization problems on Hadamard manifolds. Recently, some authors focus on studying proximal methods on Riemannian/Hadamard manifolds, see [1, 4, 7, 14, 19] an the references therein.

In the riemannian context there are few results on the rate of convegence. Recently, Tang and Huang[20] estimated the convergence rate of the proximal point algorithm, under a growth condition which is an extension of ones given by Luque[10], for the singularity of maximal monotone vector fields on Hadamard manifolds.

On the other hand, in Baygorrea et al.[3] has been shown two inexact proximal point algorithms for solving quasiconvex minimization problems on Hadamard manifolds. Observe that quasiconvex minimization problems is a larger class than convex minimization problems and has been studied recently motivated by some applications in location theory, economic theory and fractional optimization, see [12, 13, 14, 15].

In this paper, we analyze the convergence rate of an inexact proximal point algorithm on Hadamard manifolds introduced by Baygorrea et al.[3]. The main contribution of this paper is the extension of the linear/superlinear rate of convergence of the proximal point algorithm on Hadamard manifolds from the convex case to the quasiconvex ones. This result is new even in the Euclidean space.

The remainder of the paper is organized as follows: Section 2, we recall some definitions and results about Riemannian geometry, quasiconvex analysis and abstract subdifferential. In Section 3, we present the optimization problem and an inexact algorithm for solving quasiconvex minimization problems. In Section 4, we explore the convergence rate of the proposed algorithm for solving in quasiconvex minimization problems.

2 Preliminaries

In this section we recall some fundamental properties and notation on Riemannian manifolds. Those basic facts can be seen, for example, in do Carmo [6], Sakai [18], Udriste [21] and Rapcsák [17].

Let M be a n-differential manifold with finite dimension n. We denote by T_xM the tangent space of M at x and $TM = \bigcup_{x\in M} T_xM$. T_xM is a linear space and has the same dimension of M. Because we restrict ourselves to real manifolds, T_xM is isomorphic to $I\!R^n$. If M is endowed with a Riemannian metric g, then M is a Riemannian manifold and we denote it by (M, G) or only by M when no confusion can arise, where G denotes the matrix representation of the metric g. The inner product of two vectors $u, v \in T_xM$ is written as $\langle u, v\rangle_x := g_x(u, v)$, where g_x is the metrics at point x. The norm of a vector $v \in T_xM$ is set by $\|v\|_x := \langle v, v\rangle_x^{1/2}$. If there is no confusion we denote $\langle , \rangle = \langle , \rangle_x$ and $||.|| = ||.||_x$. The metrics can be used to define the length of a piecewise smooth curve $\psi : [t_0, t_1] \to M$ joining $\psi(t_0) = p'$ to $\psi(t_1) = p$ through $L(\psi) = \int_{t_0}^{t_1} \|\psi'(t)\|_{\psi(t)}dt$. Minimizing this length functional over the set of all curves we obtain a Riemannian distance $d(p', p)$ which induces the original topology on M.

Given two vector fields V and W in M, the covariant derivative of W in the direction V is denoted by $\nabla_V W$. In this paper ∇ is the Levi-Civita connection associated to (M, G). This connection defines an unique covariant derivative D/dt, where, for each vector field V, along a smooth curve $\psi : [t_0, t_1] \to M$, another vector field is obtained, denoted by DV/dt. The parallel transport along ψ from $\psi(t_0)$ to $\psi(t_1)$, denoted by P_{ψ,t_0,t_1}, is an application $P_{\psi,t_0,t_1} : T_{\psi(t_0)}M \to T_{\psi(t_1)}M$ defined by $P_{\psi,t_0,t_1}(v) = V(t_1)$ where V is the unique vector field along ψ so that $DV/dt = 0$ and $V(t_0) = v$. Since ∇ is a Riemannian connection, P_{ψ,t_0,t_1} is a linear isometry, furthermore $P^{-1}_{\psi,t_0,t_1} = P_{\psi,t_1,t_0}$ and $P_{\psi,t_0,t_1} = P_{\psi,t,t_1}P_{\psi,t_0,t}$, for all $t \in [t_0, t_1]$. A curve $\psi : I \to M$ is called a geodesic if $D\psi'/dt = 0$.

A Riemannian manifold is complete if its geodesics are defined for any value of $t \in I\!R$. Let $x \in M$, the exponential map $\exp_x : T_xM \to M$ is defined $\exp_x(v) = \gamma(1, x, v)$, for each $x \in M$. If M is complete, then $\exp_x$ is defined for all $v \in T_xM$. Besides, there is a minimal geodesic (its length is equal to the distance between the extremes).

Complete simply-connected Riemannian manifolds with nonpositive curvature are called *Hadamard manifolds*. Throughout the remainder of this paper, we always assume that M is an $n-$ dimensional Hadamard manifold. Some examples of Hadamard manifolds may be found in Section 4 of Papa Quiroz and Oliveira [14].

Given an extended real valued function $f : M \to I\!R \cup \{+\infty\}$ we denote its domain by $\texttt{domf} := \{x \in M : f(x) < +\infty\}$. f is said to be proper if $\texttt{domf} \neq \emptyset$ and $\forall x \in \texttt{domf}$ we have $f(x) > -\infty$ and its epigraph by $\texttt{epi}(f) := \{(x, \beta) \in M \times I\!R \mid f(x) \leq \beta\}$. Moreover f is a lower semicontinuous function if $\texttt{epi}(f)$ is a closed subset of $M \times I\!R$.

Let $f : M \to I\!R \cup \{+\infty\}$ be a proper function f, it is called quasiconvex if for all $x, y \in M$, $t \in [0, 1]$, it holds that

$$f(\gamma(t)) \leq \max\{f(x), f(y)\},$$

for the geodesic $\gamma : [0, 1] \to I\!R$ such that $\gamma(0) = x$ and $\gamma(1) = y$.

Definition 2.1 *Let $\{z^k\} \subset M$ such that $\{z^k\}$ converges to a point $\bar{z} \in M$. Then, the conver-*

gence is said to be:

i. *linear iff there exist a constant $\theta < 1$ and a positive $N \in I\!N$ such that*

$$d(z^k, \bar{z}) \leq \theta d(z^{k-1}, \bar{z}); \qquad \forall\, k > N.$$

ii. *superlinear iff there exist a sequence $\{\alpha_k\}$ converging to zero and a positive $\bar{N} \in I\!N$ such that*

$$d(z^k, \bar{z}) \leq \alpha_k d(z^{k-1}, \bar{z}). \qquad \forall\, k > \bar{N}.$$

Definition 2.2 *We call abstract subdifferential, denoted by ∂, any operator which associates a subset $\partial f(x)$ of T_xM to any lower semicontinuous function $f : M \to I\!R \cup \{+\infty\}$ and any $x \in M$, satisfying the following properties:*

a. *If f is convex, then $\partial f(x) = \{g \in T_xM \mid \langle g, \exp_x^{-1} z\rangle + f(x) \leq f(z), \quad \forall\, z \in M\}$;*

b. *$0 \in \partial f(x)$, if $x \in M$ is a local minimum of f;*

c. *$\partial(f + g)(x) \subset \partial f(x) + \partial g(x)$, whenever $g : M \to I\!R \cup \{+\infty\}$ is a convex continuous function which is ∂-differentiable at $x \in M$.*

Here, g is ∂-differentiable at x means that both $\partial g(x)$ and $\partial(-g)(x)$ are nonempty. We say that a function f is ∂-subdifferentiable at x when $\partial f(x)$ is nonempty.

As studied in previous works, see for example Aussel[2] and Baygorrea et al.[3], this abstract subdifferential recover a broad range of classical subdifferential. Among them, particulary, we have the Clarke subdifferential defined at the point $a \in M$ as the set

$$\partial^\circ f(x) = \{s \in T_xM \mid \langle s, v\rangle \leq f^\circ(x, v),\ \forall\, v \in T_xM\},$$

where

$$f^\circ(x, v) = \limsup_{\substack{u \to x \\ t \searrow 0}} \frac{f(\exp_u t(D\exp_x)_{\exp_x^{-1} u} v) - f(u)}{t}.$$

In this paper we also use the following (limiting) subdifferential concept of f at $x \in M$, which is defined by

$$\partial^{Lim} f(x) := \{s \in T_xM \mid \exists\, x^k \to x,\ f(x^k) \to f(x), \exists\, s^k \in \partial f(x^k) : P_{\gamma_k,0,1} s^k \to s\},$$

Remark 2.1 *Note that it is an immediate consequence that*

$$\partial f(x) \subseteq \partial^{Lim} f(x). \qquad \forall\, x \in M \tag{2.1}$$

Let $g \in \partial f(x)$. By taking $\{x^k\} = \{x\}$ and $\{g^k\} = \{g\}$ with $g^k \in \partial f(x^k)$, it follows that g^k converges to g. Thus, $g \in \partial^{Lim} f(x)$.

3 Definition of the problem and the Algorithm.

Let M be a Hadamard manifold. We are interested in solving the problem:

$$\min_{x \in M} f(x) \tag{3.2}$$

where $f : M \to I\!R \cup \{+\infty\}$ is an extended real-valued function which satisfies the following assumption:

(H1) f is a proper, bounded from below and lower semicontinuous quasiconvex function.

Furthermore, for solving the problem (3.2), we consider a follows assumptions:

(H2) ($\partial \subset \partial^{D^+}$) or ($\partial \subset \partial^{CR}$ and f is continuous in M). See Section 3 of Baygorrea et al.[3] for a definition of ∂^{D^+} and ∂^{CR} respectively.

HMIP2 Algorithm.

Initialization: Take $x^0 \in M$. Set $k = 0$.

Iterative step: Given $x^{k-1} \in M$, find $x^k \in M$ and $\epsilon^k \in T_{x^k}M$ such that

$$\epsilon^k \in \lambda_k \partial f(x^k) - \exp^{-1}_{x^k} x^{k-1}, \tag{3.3}$$

where

$$d(\exp_{x^k} \epsilon^k, x^{k-1}) \leq \max\left\{||\epsilon^k||, d(x^k, x^{k-1})\right\}, \tag{3.4}$$

$$\|\epsilon^k\| \leq \eta_k d(x^k, x^{k-1}), \tag{3.5}$$

$$\sum_{k=1}^{+\infty} \eta_k^2 < +\infty. \tag{3.6}$$

Stopping rule: If $x^{k-1} = x^k$ or $0 \in \partial f(x^k)$. Otherwise, $k - 1 \leftarrow k$ and go to *Iterative step*.

Remark 3.1 *Throughout this paper, we analyse the assymptotic case of the algorithm, that is, we consider $x^{k-1} \neq x^k$ for all $k \in I\!N$ and $0 \notin \partial f(x^k)$, for all $k \in I\!N$.*

Remark 3.2 *From (3.3), there exists $g^k \in \partial f(x^k)$ such that*

$$\lambda_k g^k = \exp^{-1}_{x^k} x^{k-1} + \epsilon^k. \tag{3.7}$$

4 Convergence rate of the HMIP2 algorithm

We denote the set

$$U := \{x \in M \mid f(x) \leq \inf_k f(x^k)\},$$

which contains the optimal solutions set, whenever it exists.

Remark 4.1 *Through this paper, we will consider U to be a nonempty set. If $U = \emptyset$, as was seen in Baygorrea et al.[3], the sequence $\{x^k\}$ generated by the algorithm will be an unbounded sequence and the sequence of the objective function values $\{f(x^k)\}$ converges to the infimum of f on M.*

The following lemmas and theorems are taken from Baygorrea et al[3]. They will be used often later to discuss the estimation of convergence rate concerned with the proposed algorithm.

Lemma 4.1 *(Baygorrea et al.[3], Lemma 5.2) Let $\{x^k\}$ and $\{\epsilon^k\}$ be two sequences generated by the* HMIP2 *algorithm. If all the assumptions of the problem:* (H1) *and* (H2) *are satisfied, then there exists an integer $k_0 \in I\!N$ such that for all $k \geq k_0$ we have*

$$d^2(x^k, x) \leq \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right) d^2(x^{k-1}, x) - \frac{1}{2} d^2(x^k, x^{k-1}), \quad \forall x \in U. \tag{4.8}$$

Furthermore, $\{x^k\}$ is a bounded sequence and $\lim_{k\to+\infty} d(x^k, x^{k-1}) = 0$.

Theorem 4.1 *(Baygorrea et al.[3], Theorem 5.4) Let $\{x^k\}$ and $\{\epsilon^k\}$ be sequences generated by* HMIP2 *algorithm. If all the assumptions of the problem:* (H1) *and* (H2) *are satisfied with $\tilde{\lambda} > 0$ such that $\tilde{\lambda} < \lambda_k$ and f be a continuous function, then $\{x^k\}$ converges to some $\bar{x} \in U$ with $0 \in \partial^{Lim} f(\bar{x})$.*

Theorem 4.2 *(Baygorrea et al.[3], Theorem 5.5) Let $\{x^k\}$ and $\{\epsilon^k\}$ be sequences generated by the* HMIP2 *algorithm. If the assumptions* (H1) *and* (H2) *are satisfied with $\tilde{\lambda} > 0$ such that $\tilde{\lambda} < \lambda_k$ and f is a locally Lipschitz function, then $\{x^k\}$ converges to some $\bar{x} \in U$ with $0 \in \partial^o f(\bar{x})$.*

Now we denote the following set

$$Z := U \cap \{x \in M \mid 0 \in \partial^{Lim} f(x)\}.$$

To study the convergence rate of the HMIP2 algorithm, we consider the following assumption:

(H3) For $\bar{x} \in Z$ such that $\lim x^k = \bar{x}$, there exist $\delta := \delta(\bar{x}) > 0$ and $\tau := \tau(\bar{x}) > 0$ such that for all $w \in B(0, \delta) \subset T_{\bar{x}}M$ and all x such that $P_{\psi,1,0}(w) \in \partial^{Lim} f(x)$ with geodesics ψ_k joinning $\psi(0) = x$ and $\psi(1) = \bar{x}$, there holds

$$d(x, \bar{x}) \leq \tau \|w\|_{T_{\bar{x}}M}.$$

Remark 4.2 *The assumptions* (H3) *may be called a growth condition at the point of convergence $\bar{x}$ on Hadamard manifolds. Note that this one is a different condition than one given by Tang and Huang[20] which was given for solving problems of singularity of maximal monotone vector fields on Hadamard manifold (particularly for convex minimization problems on Hadamard manifolds).*

Lemma 4.2 *Let $\{x^k\}$ and $\{\epsilon^k\}$ be the sequences generated by the* **HMIP2** *algorithm. Suppose that assumptions* (H1), (H2) *and* (H3) *hold. Then,*

(i) there exists $\bar{k} \in I\!N$ such that

$$\|g^k\|_{T_{x^k}M} < \delta, \tag{4.9}$$

for all $k \geq \bar{k}$, where g^k is given by (3.7);

(ii) for all $k \geq \bar{k}$, it holds that

$$d(x^k, \bar{x}) \leq \tau \frac{(\eta_k + 1)}{\lambda_k} d(x^k, x^{k-1}). \tag{4.10}$$

Proof. (i) Let $\bar{x} = \lim_{k\to\infty} x^k$ and $g^k \in \partial f(x^k)$ given by (3.7). It follows from (3.5) and since $\tilde{\lambda} < \lambda_k$ that

$$\begin{aligned}
\|g^k\|_{T_{x^k}M} &= \frac{1}{\lambda_k}\|\exp^{-1}_{x^k} x^{k-1} + \epsilon^k\|_{T_{x^k}M} \\
&\leq \frac{1}{\lambda_k}\left(\|\exp^{-1}_{x^k} x^{k-1}\|_{T_{x^k}M} + \|\epsilon^k\|_{T_{x^k}M}\right) \\
&\leq \left(\frac{\eta_k + 1}{\lambda_k}\right) d(x^k, x^{k-1}), && (4.11) \\
&\leq \left(\frac{\eta_k + 1}{\tilde{\lambda}}\right) d(x^k, x^{k-1}), \qquad k \geq 1. && (4.12)
\end{aligned}$$

Since $\eta_k \to 0$, $d(x^k, x^{k-1}) \to 0$ (see Lemma 4.1) and taking $\varepsilon = \delta$ then, there exists $\bar{k} \in I\!N$ such that $\|g^k\|_{T_{x^k}M} < \delta$ for all $k \geq \bar{k}$.

Now we prove item (ii). Taking $w = P_{\psi_k,0,1} g^k$ in assumption (H3) and due to the isometry of parallel transport $P_{\psi_k,0,1}$, for all $k \geq \bar{k}$, we have

$$\begin{aligned}
d(x^k, \bar{x}) &\leq \tau\|w\|_{T_{\bar{x}}M} \\
&= \tau\|P_{\psi_k,0,1} g^k\|_{T_{\bar{x}}M} \\
&= \tau\|g^k\|_{T_{x^k}M}
\end{aligned}$$

Therefore, the relation (4.10) follows from the last inequality combined with (4.11). ■

We now give a rate of convergence theorem, for the **HMIP2** algorithm, which completes the convergence result given by Theorem 4.1.

Theorem 4.3 *Let* $\{x^k\}$ *and* $\{e^k\}$ *be the sequences generated by the* **HMIP2** *algorithm. Suposse that assumptions* (H1), (H2) *and* (H3) *such that* $\lambda_k \in [\tilde{\lambda}, +\infty)$ *with* $\tilde{\lambda} > 0$ *are satisfied and assume* f *be a continuous function. Then* $\{x^k\}$ *converges linearly to* $\bar{x} \in Z$*. Moreover, if* $\lambda_k \nearrow +\infty$*, then this convergence is superlinear.*

Proof. Let $\bar{x} \in Z$ be the limit point of the sequence $\{x^k\}$ and $g^k \in \partial f(x^k)$ given by (3.7). Define

$$w^k := P_{\psi_k,0,1} g^k,$$

where ψ_k is the geodesic joining x^k to $\bar{x}$. Due to the isometric property of the parallel transport $P_{\psi_k,0,1}$ and the relation (4.9), we have that

$$\|w^k\|_{T_{\bar{x}}M} = \|P_{\psi_k,0,1} g^k\|_{T_{\bar{x}}M} = \|g^k\|_{T_{x^k}M} < \delta,$$

for $k \geq \bar{k}$. That is, $w^k \in B(0,\delta) \subset T_{\bar{x}}M$, for $k \geq \bar{k}$.

Furthermore,

$$P_{\psi_k,1,0} w^k = P_{\psi_k,1,0}(P_{\psi_k,0,1} g^k) = g^k \in \partial f(x^k).$$

Applying this relation to (2.1), we have $g^k \in \partial^{Lim} f(x^k)$. Thus, $P_{\psi_k,1,0} w^k \in \partial^{Lim} f(x^k)$, $\forall\, k \geq \bar{k}$.

Moreover, applying (4.10) to relation (4.8), for all $k \geq \max\{k_0, \bar{k}\}$, it follows that

$$d^2(x^k,\bar{x}) \leq \left(1 + \frac{2\eta_k^2}{1-2\eta_k^2}\right) d^2(x^{k-1},\bar{x}) - \frac{1}{2}\left(\frac{\lambda_k}{\tau(\eta_k+1)}\right)^2 d^2(x^k,\bar{x}).$$

and so

$$\left(1 + \frac{\lambda_k^2}{2\tau^2(\eta_k+1)^2}\right) d^2(x^k,\bar{x}) \leq \frac{1}{1-2\eta_k^2} d^2(x^{k-1},\bar{x}).$$

This implies that

$$d^2(x^k,\bar{x}) \leq \alpha_k^2 d^2(x^{k-1},\bar{x}), \qquad k \geq \max\{k_0,\bar{k}\} \tag{4.13}$$

where

$$\alpha_k^2 = \frac{1}{1-2\eta_k^2}\left(\frac{2\tau^2(\eta_k+1)^2}{2\tau^2(\eta_k+1)^2 + \lambda_k^2}\right). \tag{4.14}$$

Since $0 < \tilde{\lambda} < \lambda_k$, for all $k \in I\!N$, we show that

$$\alpha_k \leq r_k, \tag{4.15}$$

where

$$r_k = \frac{1}{1-2\eta_k^2}\left(\frac{2\tau^2(\eta_k+1)^2}{2\tau^2(\eta_k+1)^2 + \tilde{\lambda}^2}\right).$$

Taking into account that $\{\eta_k\}$ converges to zero, we have

$$r_k \to \frac{2\tau^2}{2\tau^2 + \tilde{\lambda}^2}.$$

Thus, there exists a positive number $k_1 \in I\!N$ with $k \geq k_1$ such that

$$r_k < \frac{1}{2}\left(1 + \frac{2\tau^2}{2\tau^2 + \tilde{\lambda}^2}\right), \qquad \forall\ k \geq k_1 \tag{4.16}$$

Combining (4.15) and (4.16), we get

$$\alpha_k^2 < \frac{1}{2}\left(1 + \frac{2\tau^2}{2\tau^2 + \tilde{\lambda}^2}\right) := \theta < 1, \qquad \forall\ k \geq k_1 \tag{4.17}$$

It follows from (4.13) and (4.17) that

$$d(x^k, \bar{x}) \leq \theta^{1/2} d(x^{k-1}, \bar{x}),$$

for all $k \geq \max\{\bar{k}, k_0, k_1\}$. Thus, the sequence $\{x^k\}$ converges linearly to $\bar{x}$.
To obtain the superlinear convergence of the sequence generated by the **HMIP2** algorithm, consider $\lambda_k \nearrow +\infty$ and since $\eta_k \to 0$, it follows from relation (4.14) that $\alpha_k \to 0$. This completes the proof. ■

Now, the following theorem shows the convergence rate of the convergence results of Theorem 4.2

Theorem 4.4 *Let $\{x^k\}$ and $\{e^k\}$ be the sequences generated by the* **HMIP2** *algorithm. Suposse that assumptions* (H1), (H2) *and* (H3) *such that $\lambda_k \in [\tilde{\lambda}, +\infty)$ with $\tilde{\lambda} > 0$ are satisfied and assume f be a locally Lipschitz function. Then $\{x^k\}$ converges linearly to $\bar{x} \in Z$. Moreover, if $\lambda_k \nearrow +\infty$, then this convergence is superlinear.*

Proof. Note that locally Lipschitz condition implies continuity of f. Taking $\partial = \partial^\circ$ and from relation (2.1), it follows that $\partial^\circ f(\bar{x}) \subseteq \partial^{Lim} f(\bar{x})$. Therefore, rate of convergence results of Theorem 4.2 is a particular case of Theorem 4.3 . ■

5 Conclusions.

- Motivated by the work of Tang and Huang[20], we extend the linear/superlinear rate of convergence of the proximal point method for quasiconvex functions on Hadamard manifolds introducing the condition (H3). This condition is different to the weak growth condition worked by Tang and Huang[20]. It allows us to obtain an authentic linear and superlinear rate of convergence of the algorithm to the point of convergence. Note that this result is new even for the case of Euclidean spaces and it improves the convergence studed by Tang and Huang[20] for convex minimization.

- In the rate of convergence analysis of the proposed algorithm, we define the set $Z = U \cap \{x \in M \mid 0 \in \partial^{Lim} f(x)\}$, which is a nonempty set whenever U is nonempty. If Z is a convex set, then assuming the weaker growth condition (H_2) given by Tang and Huang[20] and following the same ideas of that paper, it is possible to obtain the same rate of convergence of the proposed algorithm. In this sense, we would generalize the convergence of Tang and Huang[20] for quasiconvex minimization problems on Hadamard manifolds.

- Finally, it is need to make a comparison between the rate of convegence results obtained by Tang and Huang([20]) and our paper. Indeed, they obtained the linear/superlinear convergence of the sequence generate by the proximal point algorithm with respect to the solution set of the poblem. In this paper, we obtain the linear/superlinear convergence of the sequence generated by the proposed algorithm to a critical point of the problem (optimal solution in the convex case) under the assumption (H3).

Acknowledgments.

This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) of the Federal University of Rio de Janeiro(UFRJ). Brazil.

References

[1] Ahmadi P. and Khatibsadeh H., On the convergence of inexact proximal point algorithm on Hadamard manifolds, Taiwanese Journal of Mathematics, vol 18(2), 419-433 (2014)

[2] Aussel, D., Corvellec, J.N. and Lassonde, M., Mean-value property and subdifferential criteria for lower semicontinuous functions, Trans. of the American Math. Society, 347, 4147-4161 (1995)

[3] Baygorrea N., Papa Quiroz E.A. and Maculan N., Inexact proximal point methods for quasiconvex minimization on Hadamard manifolds. Submitted to Journal of Optimization, Theory and Applications. JOTA-D-15-00197 (2015)

[4] Bento G.C., Ferreira O.P. and Oliveira P.R., Proximal point method for a special class of nonconvex functions on Hadamard manifolds, Optimization, 64(2), 289-319 (2015)

[5] Da Cruz Neto, J.X., Ferreira, O.P. and Lucâmbio Pérez L.R., Contribution to the study of monotone vector fields, Acta Math. Hungarica, 94(4), 307-320 (2002)

[6] Do Carmo, M. P., Riemannian Geometry, Bikhausen, Boston (1992)

[7] Ferreira, O. P., and Oliveira, P. R., Proximal Point Algorithm on Riemannian Manifolds, Optimization, 51(2) 257-270 (2002)

[8] Ferreira, O.P., Lucambio Prez, L.R., Nmeth, S.Z., Singularities of monotone vector fields and an extragradient-type algorithm. J. Global Optim. 31, 133151 (2005)

[9] Güler, O. On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim., 29(2):403419 (1991)

[10] Luque F.J., Asymptotic convegence analysis of the proximal point algorithm, SIAM J. Control Optim. 22, 277-293 (1984)

[11] Martinet, B., Regularisation d'inéquations variationelles par approximations successives, Revue Française Autom. Informatique et Recherche Opérationnelle, 4, 154-159 (1970)

[12] Papa Quiroz, E.A. and Oliveira, P.R., Proximal point methods for minimizing quasiconvex locally lipschitz functions on Hadamard manifolds, Nonlinnear Analysis. 75, 5924-5932, (2012)

[13] Papa Quiroz, E. A., O. P. and Oliveira, P. R., Proximal point methods for quasiconvex and convex functions with Bregman distances on Hadamard manifolds. J. Convex Anal. 16(1) 49-69 (2009)

[14] Papa Quiroz, E.A. and Oliveira, P.R., Full convergence of the proximal point method for quasiconvex function on Hadamard manifolds, ESAIM: Control, Optimisation and Calculus of Variations, 18, 483-500 (2012)

[15] Papa Quiroz, E.A., Mallma Ramirez L. and Oliveira, P.R., An Inexact Proximal Method for Quasiconvex Minimizations, http://www.optimization-online.org/DB_FILE/2013/08/3982.pdf, acepted for publication in EJOR (2015)

[16] Rockafellar, R.T., Monotone operators and the proximal point algorithm, SIAM Journal of control and Opt., 877-898 (1976)

[17] Rapcsák, T., Smooth nonlinear optimization, Kluwer Academic Publishers (1997)

[18] Sakai, T. Riemannian Geometry, American Mathematical Society, Providence, RI. (1996)

[19] Tang G.J., Zhou L.W. and Huang N.J., The proximal point algorithm for pseudomonotone variational inequalities on Hadamard manifolds. Operations Research Letters 7(4),779-790 (2012)

[20] Tang G.J. and Huang N.J., Rate of convergence for proximal point algorithms on Hadamard manifolds. Operations Research Letters, (42), 383-387 (2014)

[21] Udriste, C., Convex function and optimization methods on Riemannian manifolds, Kluwer Academic Publishers (1994)