Operations Research Letters 42 (2014) 383–387

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Rate of convergence for proximal point algorithms on Hadamard manifolds

Guo-ji Tang [a], Nan-jing Huang [b,*]

[a] School of Science, Guangxi University for Nationalities, Nanning, Guangxi 530006, PR China

[b] Department of Mathematics, Sichuan University, Chengdu, Sichuan 610064, PR China

ARTICLE INFO

Article history:
Received 15 January 2014
Received in revised form
20 June 2014
Accepted 21 June 2014
Available online 26 June 2014

Keywords:
Inexact proximal point algorithm
Hadamard manifold
Convergence rate
Maximal monotone vector field

ABSTRACT

In this paper, an estimate of convergence rate concerned with an inexact proximal point algorithm for the singularity of maximal monotone vector fields on Hadamard manifolds is discussed. We introduce a weaker growth condition, which is an extension of that of Luque from Euclidean spaces to Hadamard manifolds. Under the growth condition, we prove that the inexact proximal point algorithm has linear/superlinear convergence rate. The main results presented in this paper generalize and improve some corresponding known results.

1. Introduction

The proximal point algorithm, first introduced by Martinet [18] for convex minimization and further generalized by Rockafellar [26] to get todays version, is one of basic methods for maximal monotone operators, variational inequalities and minimization problems and so on (see, for example, [26,5,4,33,14,11,13,28,30,31]). The proximal point algorithm has many interesting applications such as signal recovery and others [7,6].

Let us emphasize that the estimate of convergence rate is of fundamental importance in the study of proximal point algorithms. In the case when the space is linear, there are some results about the estimate of convergence rate concerned with the proximal point methods. In a seminal work [26], under the assumption that the inverse operator A^{-1} is Lipschitz continuous at 0, Rockafellar proved that the proximal point algorithm is linearly convergent. In 1999, Solodov and Svaiter [30,31] obtained the same results for a hybrid extragradient-proximal point algorithm and a hybrid projection-proximal algorithm under similar assumptions. We would like to mention that the Lipschitz continuity of the inverse operator A^{-1} around 0 implies that the solution of the variational inclusion problem is unique. Hence, the requirement is rather strong. In order to overcome this drawback, Luque [17] introduced a weaker growth condition to ensure the superlinear convergence. The significance of Luque's work lies in that it does not require the Lipschitz continuity of the inverse operator A^{-1} around 0. For example, when this mapping is polyhedral, the growth condition holds naturally whereas the Lipschitz continuity condition fails. For more information, one also can refer to the work of Dong [11].

Since a seminal work of Rockafellar [26], the proximal point algorithm has been extensively studied in linear spaces ($\mathbb{R}^n$/Hilbert spaces/Banach spaces) by many researchers (see, for example, [15,7,6,26,30,31]). Recently, some authors focus on extending proximal methods from Euclidean spaces to Riemannian/Hadamard manifolds (see, for example, [1–3,10,12,15,16,25,24,23,34,35] and the references therein). Very recently, Tang and Huang [32] extended an inexact proximal point algorithm with the relative error tolerance proposed by Han and He [13] from the Euclidean spaces to the Hadamard manifolds. Under some suitable assumptions, they proved that the sequence generated by the proposed method converges to the singularity of maximal monotone vector fields on Hadamard manifolds. However, to the best of our knowledge, we cannot find any work about the estimate of convergence rate of proximal point algorithms on Riemannian manifolds other than linear spaces. From the point of view of theory and applications, we look forward to obtain some related results concerned with the estimate of convergence rate on Riemannian manifolds or Hadamard manifolds. On the other hand, unlike in the case of Euclidean spaces (or Hilbert spaces), the loss of linear structure in the case of Riemannian manifolds or Hadamard manifolds makes the task more complicated. Therefore, it is interesting and important to explore the convergence rate results for the inexact proximal point algorithms on Riemannian/Hadamard manifolds.

* Corresponding author.
E-mail addresses: nanjinghuang@hotmail.com, njhuang@scu.edu.cn (N.-j. Huang).

http://dx.doi.org/10.1016/j.orl.2014.06.009
0167-6377/

In this paper, we will explore the convergence rate of the proximal point algorithm on Hadamard manifolds. Under a weaker growth condition, we prove that the inexact proximal point algorithm proposed in Tang and Huang [32] admits the superlinear convergence on Hadamard manifolds. The main results of the present paper also provide the estimate of convergence rate for some existing proximal point algorithms on Hadamard manifolds (see, for example, [15,12]) besides the inexact proximal point algorithm of [32].

2. Preliminaries

In this section, we recall some fundamental definitions, properties and notations concerned with the Riemannian manifolds, which can be found in any textbook on Riemannian geometry, for example, [27].

Let $\mathbb{M}$ be a connected $m-$dimensional manifold and let $x \in \mathbb{M}$. We always assume that $\mathbb{M}$ can be endowed with a Riemannian metric to become a Riemannian manifold. The tangent space of $\mathbb{M}$ at x is denoted by $T_x\mathbb{M}$. We denote by $\langle \cdot, \cdot \rangle_x$ the scalar product on $T_x\mathbb{M}$ with the associated norm $\| \cdot \|_x$, where the subscript x is sometimes omitted. The tangent bundle of $\mathbb{M}$ is denoted by $T\mathbb{M} = \cup_{x\in\mathbb{M}} T_x\mathbb{M}$, which is naturally a manifold. Given a piecewise smooth curve $\gamma : [a, b] \to \mathbb{M}$ joining x to y (i.e. $\gamma(a) = x$ and $\gamma(b) = y$), we can define the length of γ by $l(\gamma) = \int_a^b \|\gamma'(t)\|dt$. Then the Riemannian distance $d(x, y)$, which induces the original topology on $\mathbb{M}$, is defined by minimizing this length over the set of all such curves joining x to y.

Let ∇ be a Levi-Civita connection associated with the Riemannian metric. Let γ be a smooth curve in $\mathbb{M}$. A vector field X is said to be parallel along γ iff $\nabla_{\gamma'}X = 0$. If γ' itself is parallel along γ, we say that γ is a geodesic (this notion is different from the corresponding one in the calculus of variations), and in this case $\|\gamma'\|$ is a constant. When $\|\gamma'\| = 1$, γ is said to be normalized. A geodesic joining x to y in $\mathbb{M}$ is said to be minimal if its length equals $d(x, y)$.

A Riemannian manifold is complete if, for any $x \in \mathbb{M}$, all geodesics emanating from x are defined for all $-\infty < t < +\infty$. By the Hopf–Rinow Theorem, we know that, if $\mathbb{M}$ is complete, then any pair of points in $\mathbb{M}$ can be joined by a minimal geodesic. Moreover, $(\mathbb{M}, d)$ is a complete metric space and any bounded closed subsets are compact.

Assuming that $\mathbb{M}$ is complete, the exponential map $\exp_x : T_x\mathbb{M} \to \mathbb{M}$ at x is defined by $\exp_x v = \gamma_v(1, x)$ for each $v \in T_x\mathbb{M}$, where $\gamma(\cdot) = \gamma_v(\cdot, x)$ is the geodesic starting at x with velocity v. Then $\exp_x tv = \gamma_v(t, x)$ for each real number t. Note that, the mapping $\exp_x$ is differentiable on $T_x\mathbb{M}$ for any $x \in \mathbb{M}$.

A complete, simply connected Riemannian manifold of nonpositive sectional curvature is called an Hadamard manifold. Throughout the remainder of this paper, we will always assume that $\mathbb{M}$ is an m-dimensional Hadamard manifold.

Let $\mathcal{X}(\mathbb{M})$ denote the set of all multivalued vector fields $A : \mathbb{M} \to 2^{T\mathbb{M}}$ such that $A(x) \subseteq T_x\mathbb{M}$ for each $x \in \mathbb{M}$ and the domain $\mathcal{D}(A)$ of A is closed and convex, where the domain $\mathcal{D}(A)$ of A is defined by

$$\mathcal{D}(A) = \{x \in \mathbb{M} : A(x) \neq \emptyset\}.$$

The notions of monotonicity in the following definition were given and studied extensively in [8,9,15,19–21], which are extensions of the corresponding ones in the Euclidean spaces. In particular, concepts (i) and (ii) were introduced and studied in [21] for the single-valued case and in [8] for the multivalued case.

Definition 2.1. Let $A \in \mathcal{X}(\mathbb{M})$ be a vector field. Then A is said to be

(i) monotone iff the following condition holds for any $x, y \in \mathcal{D}(A)$,

$$\langle u, \exp_x^{-1} y\rangle \le \langle v, -\exp_y^{-1} x\rangle, \quad \forall u \in A(x) \text{ and } \forall v \in A(y);$$

(ii) maximal monotone iff it is monotone and the following implication holds for any $x \in \mathcal{D}(A)$ and $u \in T_x\mathbb{M}$:

$$\langle u, \exp_x^{-1} y\rangle \le \langle v, -\exp_y^{-1} x\rangle, \quad \forall y \in \mathcal{D}(A) \text{ and } v \in A(y) \Rightarrow u \in A(x).$$

3. Convergence rate of an inexact proximal point algorithm

We first recall the inexact proximal point algorithm proposed by Tang and Huang [32] for singularities of maximal monotone vector fields on Hadamard manifolds. Let $A : \mathbb{M} \to 2^{T\mathbb{M}}$ be a maximal monotone vector field such that $A(x) \subseteq T_x\mathbb{M}$ for each $x \in \mathbb{M}$. A point $x \in \mathbb{M}$ is said to be a singularity of the maximal monotone vector field A if $0 \in A(x)$. We denote the set of all singularities of the maximal monotone vector field A by Z, that is, $Z = \{x \in \mathbb{M} : 0 \in A(x)\}$.

Algorithm 3.1. Initialization: choose an initial point $x^0 \in \mathcal{D}(A)$. Set $k = 0$ and let $\{c_k\} \subset [c, +\infty)$ be a sequence of scalars with $c > 0$.

Iterative step: at stage k, given x^k, compute x^{k+1} such that

$$e^{k+1} \in c_kA(x^{k+1}) - \exp_{x^{k+1}}^{-1} x^k, \tag{3.1}$$

where e^{k+1} is regarded as an error term and conforms to the following condition:

$$\|e^{k+1}\| \le \eta_k d(x^{k+1}, x^k) \quad \text{with} \sum_{k=0}^{\infty} \eta_k^2 < +\infty. \tag{3.2}$$

The following lemmas and theorem are taken from [32] which are useful for the discussion of the estimation of convergence rate concerned with the proposed algorithm.

Lemma 3.1. *Let $\{x^k\}$ and $\{e^k\}$ be sequences that conform to recursion (3.1). Then for any $x^* \in Z$ and all $k \ge 0$ we have*

$$\langle e^{k+1} + \exp_{x^{k+1}}^{-1} x^k, \exp_{x^{k+1}}^{-1} x^*\rangle \le 0 \tag{3.3}$$

and

$$d^2(x^{k+1}, x^*) \le d^2(x^k, x^*) - d^2(x^{k+1}, x^k) + 2\langle e^{k+1}, -\exp_{x^{k+1}}^{-1} x^*\rangle. \tag{3.4}$$

Lemma 3.2. *Let $\{x^k\}$ and $\{e^k\}$ be sequences generated by Algorithm 3.1. Then there exists an integer $k_0 \ge 0$ such that, for all $k \ge k_0$,*

$$d^2(x^{k+1}, x^*) \le \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right) d^2(x^k, x^*) - \frac{1}{2} d^2(x^{k+1}, x^k), \tag{3.5}$$

where x^* is any singularity of the vector field A. Furthermore, $\{x^k\}$ is a bounded sequence and

$$\lim_{k\to\infty} d(x^{k+1}, x^k) = 0. \tag{3.6}$$

Theorem 3.1. *Let A be a maximal monotone vector field on $\mathbb{M}$, and $\{x^k\}$ and $\{e^k\}$ be sequences generated by Algorithm 3.1. Then $\{x^k\}$ converges to some x^∞ with $0 \in A(x^\infty)$.*

For the convenience of readers, we shall recall some concepts about the convergence rate and the parallel transport on the tangent bundle $T\mathbb{M}$.

Definition 3.1. Let $\{x^k\}$ be a sequence converging to x^*. The convergence is said to be:

(i) at least linear iff there exist a constant $\theta < 1$ and a positive N such that

$$d(x^{k+1}, x^*) \le \theta d(x^k, x^*), \quad \forall k > N;$$