A splitting minimization method on geodesic spaces

J.X. Cruz Neto

DM, Universidade Federal do Piauí, Teresina, PI 64049-500, BR

B.P. Lima

DM, Universidade Federal do Piauí, Teresina, PI 64049-500, BR

P.A. Soares Júnior

CCN, Universidade Estadual do Piauí, Piauí, PI 64002-224, BR

Abstract

We present in this paper the alternating minimization method on CAT(0) spaces for solving unconstraints convex optimization problems. Under the assumption that the function being minimize is convex, we prove that the sequence generated by our algorithm converges to a minimize point. The results presented in this paper are the first ones of this type that have been given for theses spaces and alternating type methods. Moreover, the method proposed in this paper is attractive for solving certain range of problems.

Keywords: Proximal point method, geodesic metric spaces, nonpositive sectional curvature, convex functions
2000 MSC: 65K05, 54E50, 32Q05, 46N10

---

1. Introduction

Several convex optimization problems that arise in practice are modeled as a sum of convex functions, see [13, 22, 9, 12]. Minimizing the sum of simple functions or finding a common point to a collection of closed sets is a very active field of research, with applications in approximation theory von

Email addresses: jxavier@ufpi.edu.br (J.X. Cruz Neto), barnabe@ufpi.edu.br (B.P. Lima), pedrosoares@uespi.br or pedroasoaresjr@gmail.com (P.A. Soares Júnior)

Neumann [23], image reconstruction Combettes and Wajs [10]. In the last three decades several authors have proposed the generalized proximal point algorithm that is a method for finding a minimizer of a convex lower semicontinuous function defined on a Hilbert space. Its origin goes back to Martinet, Rockafellar, and Brezis-Lions [19, 21, 7]. In the context of the Hadamard manifold, Ferreira and Oliveira [14] proved convergence of the sequence of proximal points to a minimizer point. Li et al. [18] considered the problem of finding a singularity of a multivalued vector field in a Hadamard manifold and presented a general proximal point method to solve that problem. Papa Quiroz and Oliveira [20] considered this method for quasi-convex functions and proved full convergence of the sequence to a minimizer point in the setting of the Hadamard manifold. Bento et al. [6] considered this method for C^1-lower type functions and obtained local convergence of the generated sequence to a minimizer, in the case of the Hadamard manifold. Also, we should like to mention Zaslavski's very recent paper [25] with a different approach to the proximal point method in metric spaces as well as Băcák [4] that introduced this method in the context of CAT[0] spaces.

To date, in convergence analysis of the exact proximal point algorithm for solving convex or quasiconvex minimization problems, it has been necessary to consider Hadamard type manifolds. This is because the convergence analysis is based on Fejér convergence to the minimizer set of the objective function and because these manifolds, apart from having the same topology and differentiable structures as Euclidean spaces, also have geometric properties that satisfy the characterization of Fejér convergence of the generated sequence by algorithm.

The alternating algorithms has been studied in several setting. The starting fundamental result is due to von Neumann [24] with the purpose of solving convex feasibility problems. More recently Attouch et al. [1] presented the alternating algorithms in the setting nonconvex with applications in decision sciences. Another important approach to algorithm alternating with applications in game theory in the context of Hadamard manifolds, was presented by Cruz Neto et al. [11]. One advantage of the alternating algorithm is that it enables us to monitor what happens in each space of action after a given iteration. Another advantage is that the accounts are quite simplified compared to those in product space. Due to the wide potential range of applications of these algorithms (from engineering to decision sciences) we adopt a quite general terminology. In this paper, we propose and analysis the alternating proximal algorithm in the setting of CAT(0) spaces. Let us describe: consider the following minimization problem

$$\begin{array}{l} \min H(x,y) \\ s.t.\ \ (x,y) \in M \times N, \end{array} \tag{1}$$

where M and N are CAT(0) spaces and $H : M \times N \to \mathbb{R} \cup \{+\infty\}$ is a convex proper lower semicontinuous function bounded from below that has the following structure:

(i) $H(x,y) = f(x) + g(y)$;

(ii) $f : M \to \mathbb{R} \cup \{+\infty\}$, $g : N \to \mathbb{R} \cup \{+\infty\}$ are convex proper lower semicontinuous (lsc).

The alternating proximal algorithm to solve optimization problems of the form (1) generates, for a starting point $z_0 = (x_0, y_0) \in M \times N$, a sequence (z_k), with $z_k = (x_k, y_k) \in M \times N$, as it follows:

$$(x_k, y_k) \to (x_{k+1}, y_k) \to (x_{k+1}, y_{k+1})$$

$$\begin{cases} x_{k+1} = \arg\min\{H(x, y_k) + \frac{1}{2\lambda_k} d_M^2(x_k, x);\ x \in M\} \\ \\ y_{k+1} = \arg\min\{H(x_{k+1}, y) + \frac{1}{2\mu_k} d_N^2(y_k, y),\ y \in N\} \end{cases} \tag{2}$$

where d_M, d_N are distances associated with the spaces M and N respectively, (λ_k) and (μ_k) are sequences of positive numbers and $H(x,y) = f(x) + g(y)$ is a separable function. Previous related works can be found in Attouch et al. [2, 1, 3], but there, the setting is $M = \mathbb{R}^m$, $N = \mathbb{R}^n$. Lewis and Malick [17] studied the method of alternating projections in the context in which $M \subset \mathbb{R}^n$ and $N \subset \mathbb{R}^n$ are two smooth manifolds intersect transversally.

In each iteration we must solve the following subproblems:

$$\begin{array}{c} \min H(x, y_k) + \frac{1}{2\lambda_k} d_M^2(x_k, x), \\ s.t.\ \ x \in M, \end{array} \tag{3}$$

and

$$\begin{array}{c} \min H(x_{k+1}, y) + \frac{1}{2\mu_k} d_N^2(y_k, y), \\ s.t.\ \ y \in N. \end{array} \tag{4}$$

To solve the subproblem of the form (3) or (4), we use the exact proximal point algorithm that generates, for a starting point $x_0 \in M$, a sequence (x_k), with $x_k \in M$ as it follows:

$$x_{k+1} = \arg\min_{x \in M} \left\{ f(x) + \frac{1}{2\lambda_k} d_M^2(x_k, x) \right\},$$

where $f(x) = H(x, y_k)$ e d_M is the distance in M.

The organization of our paper is as follows. In Section 2, present some elements and terminology concerning CAT(0) spaces, projection, weak convergence and convexity. In Section 3, present the main result of the paper. Section 4 contains a conclusion.

2. Elements of CAT(0) spaces

In this section we introduce some fundamental properties and notations concerning CAT(0) spaces, which may be found in [8], projection, weak convergence that will be used later.

Let (M, d) be a metric space, where M is a set and d a metric in M. A geodesic path joining $x \in M$ to $y \in M$ (or, more briefly, a geodesic from x to y) is a map γ from a closed interval $[0, \ell] \subset \mathbb{R}$ such that $\gamma(0) = x$, $\gamma(\ell) = y$ and $d(\gamma(t), \gamma(t')) = |t - t'|$ for all $t, t' \in [0, \ell]$. Let M be a geodesic space, i.e., a metric space for which every two points $x, y \in M$ can be joined by a geodesic segment, and $\Delta(x, y, z)$ a geodesic triangle in M, which is a union of three geodesics. Let $[x, y]$ denote the geodesic side between x, y. A comparison triangle for Δ is a triangle $\bar{\Delta}(\bar{x}, \bar{y}, \bar{z})$ in $\mathbb{R}^2$ with the same side lengths as Δ. The interior angle of $\bar{\Delta}$ at $\bar{x}$ is called the comparison angle between y and z at x, and is denoted $\alpha'(y, x, z)$. Let p be a point on a side of Δ, say, $[x, y]$. A comparison point in $\bar{\Delta}$ is a point $\bar{p} \in [\bar{x}, \bar{y}]$ with $d(x, p) = d_{\mathbb{R}^2}(\bar{x}, \bar{p})$. $\bar{\Delta}$ satisfies the $CAT(0)$ inequality if for any $p, q \in \Delta$ and their comparison points $\bar{p}, \bar{q} \in \bar{\Delta}$,

$$d(p, q) \leq d_{\mathbb{R}^2}(\bar{p}, \bar{q}).$$

M is a CAT(0) space iff all geodesic triangles in M satisfy the CAT(0) inequality. (M, d) is said to be a geodesic metric space (or, more briefly, a geodesic space) if every two points in M are joined by a geodesic.

Let (M, d) be a $CAT(0)$ space. Having two points $x, y \in M$, we denote the geodesic segment from x to y by $[x, y]$. We usually do not distinguish between a geodesic and its geodesic segment, as no confusion can arise. A subset A of a metric space (M, d) is said to be convex iff every pair of points $x, y \in A$ can be joined by a geodesic in M and the image of every such geodesic is contained in A.

Example 2.1. *Standard examples of CAT(0) spaces, see [8] .*

- *Euclidean space,* $\mathbb{R}^n$.
- *Hyperbolic spaces,* H^n.
- *Symmetric spaces of non-compact type. For example,* $SL(n, \mathbb{R})/SO(n)$.
- *Hadamard manifolds, i.e., complete, simply connected Riemannian manifolds of non-positive sectional curvature.*
- *Products of CAT(0) spaces.*
- *When endowed with the induced metric, a convex subset of Euclidean space* $\mathbb{R}^n$ *is CAT(O).*

Remark 2.1. *A characterization for a geodesic metric space* (M, d) *to be CAT(0) space is that for any* $x \in M$ *and any geodesic* $\gamma : [0, 1] \to M$ *and any* $t \in [0, 1]$*:*

$$d(x, \gamma(t))^2 \leq (1 - t)d(x, \gamma(0))^2 + td(x, \gamma(1))^2 - t(1 - t)d(\gamma(0), \gamma(1))^2. \quad (5)$$

See [4].

For any metric space (M, d) and $A \subset M$, define the distance function by

$$d_A(x) = \inf_{a \in A} d(x, a), \ x \in M$$

Let us note that the function d_A is convex and continuous provided M is $CAT(0)$ and A is convex and complete [8, Cor. 2.5, p.178].

Lemma 2.1. *Let* (M, d) *be a* $CAT(0)$ *space and* $A \subset M$ *be complete and convex. Then:*

(i) For every $x \in M$*, there exists an unique point* $P_A(x) \in A$ *such that* $d(x, P_A(x)) = d_A(x)$.

(ii) If $y \in [x, P_A(x)]$ *, then* $P_A(x) = P_A(y)$.

(iii) If $x \in M \backslash A$ and $y \in A$ such that $P_A(x) \neq y$, then $\alpha(x, P_A(x), y) \geq \pi/2$
(iv) The mapping P_A is a non-expansive retraction from M onto A.

Proof. See [8, Proposition 2.4, p.176]. □

The mapping P_A is called the metric projection onto A.

Let M and N be complete $CAT(0)$ spaces. We define the distance in $M \times N$ as follows:

Definition 2.1. *Let*

$$d : (M \times N) \times (M \times N) \to \mathbb{R}_+$$

be given by

$$d(z_1, z_2) = [d_M^2(x_1, x_2) + d_N^2(y_1, y_2)]^{1/2}$$

for all $z_1 = (x_1, y_1),\ z_2 = (x_2, y_2)$ in $M \times N$, where d_M, d_N are distances in M, N respectively. It is easy to see that d is a distance in $M \times N$.

A point $x_0 \in M$ is called the weak limit of a sequence $(x_n)_{n \in \mathbb{N}} \subset M$ iff for every geodesic arc γ starting at x_0, $P_\gamma(x_n)$ converges to x_0. In this case, we say that (x_n) converges to x_0 weakly. We use the notation $x_n \xrightarrow{w} x$. Note that if $x_n \to x$, then $x_n \xrightarrow{w} x$. Moreover, if there is a subsequence (x_{n_k}) of (x_n) such that $x_{n_k} \xrightarrow{w} \bar{x}$ for some $\bar{x} \in M$, we say that $\bar{x}$ is a weak cluster point of the sequence (x_n). Every bounded sequence has a weak cluster point, see [15, Theorem 2.1], or [16, p. 3690] .

A function $f : M \to (-\infty, \infty]$ is called weakly lower semicontinuous (lsc) at a given point $x \in M$ iff

$$\liminf_{n \to \infty} f(x_n) \geq f(x)$$

for each sequence $x_n \xrightarrow{w} x$. A sequence $(x_n) \subset M$ is Fejér monotone with respect to A if, for any $a \in A$,

$$d(x_{n+1}, a) \leq d(x_n, a), \quad n \in \mathbb{N}.$$

3. Principal results

We begin with a useful proposition, whose proof follows easily from the fact that a closed convex subset of a complete $CAT(0)$ space is (sequentially) weakly closed [5, Lemma 3.1].

Proposition 3.1. *Let M be a complete $CAT(0)$ space. If $f: M \to (-\infty, \infty]$ is a lsc convex function, then it is weakly lsc.*

Proof. See [4, Lemma 3.1]. □

From now on fix our areas of activity. We consider M and N complete $CAT(0)$ spaces and $H: M \times N \to \mathbb{R} \cup \{+\infty\}$ a function , where H has the following structure:

(i) $H(x, y) = f(x) + g(y)$;

(ii) $f: M \to \mathbb{R} \cup \{+\infty\}$, $g: N \to \mathbb{R} \cup \{+\infty\}$ are proper and lsc convex functions.

Now, let A be the set of minimizers of H,

$$A := \left\{ a \in M \times N;\ H(a) = \inf_{a \in M \times N} H(z) \right\}.$$

We can assume that $H(a) = 0$, $\forall a \in A$.

Lemma 3.1. *Let (z_k) be the sequence generated by* (2) *and $a \in A$. Then,*

(i) $d(z_{k+1}, a) \leq d(z_k, a),\ \forall k \in \mathbb{N}$;

(ii) $2\lambda_k H(z_{k+1}) \leq d^2(z_k, a) - d^2(z_{k+1}, a)$.

Proof. Let $\gamma: [0, 1] \to M \times N$ be a geodesic, with $\gamma(0) = a$ and $\gamma(1) = z_{k+1}$. From of (2), we obtain

$$f(x_{k+1}) + g(y_k) + \frac{1}{2\lambda_k} d_M^2(x_k, x_{k+1}) \leq f(x) + g(y_k) + \frac{1}{2\lambda_k} d_M^2(x_k, x),$$

$$f(x_{k+1}) + g(y_{k+1}) + \frac{1}{2\lambda_k} d_N^2(y_k, y_{k+1}) \leq f(x_{k+1}) + g(y) + \frac{1}{2\lambda_k} d_N^2(y_k, y).$$

Adding the last two inequalities (member to member), we obtain

$$H(z_{k+1}) + \frac{1}{2\lambda_k} d^2(z_k, z_{k+1}) \leq H(z) + \frac{1}{2\lambda_k} d^2(z_k, z). \tag{6}$$

Thus,

$$H(z_{k+1}) + \frac{1}{2\lambda_k} d^2(z_k, z_{k+1}) \leq H(\gamma(t)) + \frac{1}{2\lambda_k} d^2(z_k, \gamma(t)).$$

Now, $H(\gamma(t)) \le (1-t)H(\gamma(0)) + tH(\gamma(1))$, it follows that,

$$\frac{1}{2}d^2(z_k, z_{k+1}) - \frac{1}{2}d^2(z_k, \gamma(t)) \le \lambda_k(H(\gamma(t)) - H(z_{k+1})) \le \lambda_k(t-1)H(z_{k+1}).$$

Using the last inequality and (5), with $x = z_k$, we have

$$\frac{(1-t)}{2}d^2(z_k, z_{k+1}) - \frac{(1-t)}{2}d^2(z_k, a) + \frac{t(1-t)}{2}d^2(a, z_{k+1}) \le \lambda_k(t-1)H(z_{k+1}).$$

Thus,

$$((1-t)/2)[d^2(z_k, z_{k+1}) - d^2(z_k, a) + td^2(a, z_{k+1})] \le \lambda_k(t-1)H(z_{k+1}),$$

therefore,

$$d(z_k, z_{k+1})^2 - d(z_k, a)^2 + td(a, z_{k+1})^2 \le 0,$$

making $t = 1$ this proves the item (i). Now, splitting up the penultimate inequality by $(1-t)$, making $t = 1$ and using that $H(a) = 0$, we obtain (ii). □

Theorem 3.1. *Let $(M \times N, d)$ be a complete $CAT(0)$ space and H be a lsc convex function. We assume that H has a minimizer. Then, for an starting point $z_0 \in M \times N$, and a sequence of positive numbers (λ_k) such that $\sum_1^\infty \lambda_k = \infty$, the sequence $(z_k) \subset M \times N$ defined by (2) weakly converges to a minimizer of H.*

Proof. From Lemma 3.1 inequalities (i) and (ii) , we have

$$H(z_n)\sum_{k=1}^{n-1}\lambda_k \le \sum_{k=1}^{n-1}\lambda_k H(z_{k+1}) \le \frac{1}{2}d^2(z_1, a) - \frac{1}{2}d^2(z_n, a)$$

and

$$H(z_n) \le \frac{d^2(z_1, a)}{2\sum_{k=1}^{n-1}\lambda_k}$$

Since the right side of the last inequality tends to zero as $n \to \infty$, it follows that (z_k) is a minimizing sequence. Therefore, $H(z_k)$ tends to zero as $k \to \infty$. To finish the proof, let z^* be a weak cluster point of (z_k). From Proposition 3.1 we have that H is weakly lsc. Thus, $H(z^*) = 0$, and therefore $z^* \in A$. Using [5, Proposition 3.3] it follows $z_k \xrightarrow{w} z^*$.

□

4. Conclusion

We present and analyze the alternating proximal algorithm on CAT[0] spaces for minimizing sums of separable convex functions. We use intrinsic structure this space, convexity of the objective function and Férje monotonicity of the sequence generated by the algorithm alternating to derive important theoretical result of convergence (Theorem 3.1).

References

[1] H. Attouch, J. Bolte, P. Redont A. Soubeyran, Proximal Alternating Minimization and Projection Methods for Nonconvex Problems: An Approach Based on the Kurdyka-Lojasiewicz Inequality, Math. Oper. Res. 35(2) 2010 438–457.

[2] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's. J. Convex Anal. 15(3) (2008) 485–506.

[3] H. Attouch, P. Redont, A. Soubeyran, A new class of alternating proximal minimization algorithms with costs to move, SIAM J. Optim. 18(3) (2007) 1061–1081.

[4] M. Băcák, The proximal point algorithm in metric spaces, preprint (2012), to appear in Israel J. Math., avaliable at http://arxiv.org/pdf/1206.7074.pdf.

[5] M. Băcák, I. Searston, B. Sims, Alternating projections in CAT(0) spaces, J. Math. Anal. Appl., 385 (2012), pp. 599–607.

[6] G.C. Bento, O.P. Ferreira, P.R. Oliveira, Local convergence of the proximal point method for a special class of nonconvex functions on Hadamard manifolds, Nonlinear Anal. 73 (2010) 564–572.

[7] H. Brézis, P. Lions, Produits infinis de résolvantes, Israel J. Math. 29 (1978) 329–345.

[8] M. Bridson, A. Haefliger, Metric spaces of non-positive curvature, Springer-Verlag, Berlin, 1999.

[9] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inf. Theory, 52 (2006) 489–509.

[10] P.L. Combettes, V. Wajs, Signal recovery by proximal forward-backward splitting, Multiscale Model. Simulation 4(4) (2005) 1168–1200.

[11] J.X. Cruz Neto, P.R. Oliveira, P.A. Soares Júnior, A. Soubeyran, Learning how to play Nash, potential games and alternating minimization method for structured nonconvex problems on Riemannian manifolds, preprint (2012), available in J. Convex Anal., at http://www.heldermann.de/JCA/JCA20/JCA202/jca20025.htm

[12] D. Donoho, Compressed sensing, IEEE Trans. Inf. Theory, 52 (2006) 1289–1306.

[13] D. Goldfarb, S. Ma, Fast multiple splitting algorithms for convex optimization, preprint (2011), available online at http://arxiv.org/abs/0912.4570v2

[14] O.P. Ferreira, P.R. Oliveira, Proximal point algorithm on Riemannian manifolds, Optimization 51 (2) (2002) 257–270.

[15] J. Jost, Equilibrium maps between metric spaces, Calc. Var. Partial Differential Equations, 2 (1994) 173–204.

[16] W. Kirk, B. Panyanak, A concept of convergence in geodesic spaces, Nonlinear Anal. 68 (2008) 3689–3696.

[17] A.S. Lewis, J. Malick, Alternating projection on manifolds, Math. Oper. Res. 33(1) (2008) 216–234.

[18] C. Li, G. Lopez, V. Martin-Marquez, Vector fields and the proximal point algorithm on Hadamard manifolds, J. London Math. Soc. 79 (2) (2009) 663–683.

[19] B. Martinet, Régularisation d inéquations variationnelles par approximations successives, Rev. Française Informat. Recherche Opérationnelle, 4 (1970) 154–158.

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

[21] T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976), pp. 877–898.

[22] A.N. Tikhonov, V.Y. Arsenin, Solutions of ill-posed problems, Winston & Sons, Washington, D.C., 1977.

[23] J. von Neumann, Functional operators, Annals of Mathematics Studies, vol. 22, Princeton University Press, Princeton, NJ, 1950.

[24] J. von Neumann, Functional Operators, II. The Geometry of Orthogonal Spaces, Annals of Mathematics Studies vol 22, Princeton University Press, Princeton, NJ, 1950.

[25] A. Zaslavski, Inexact proximal point methods in metric spaces, Set-Valued Anal. 19 (2011) 589–608.