A convergence theorem by the proximal method with a general perturbation on geodesic spaces with curvature bounded above
測地距離空間上での一般摂動を用いた近接点法による収束定理

東邦大学・理学部 梶村拓豊
Takuto Kajimura
Department of Information Science
Toho University
東邦大学・理学部 木村泰紀
Yasunori Kimura
Department of Information Science
Toho University

Abstract

In this paper, we show some important properties of a resolvent with a general perturbation for convex functions on an admissible complete CAT(1) space. We further investigate an approximation theorem to a minimizer of a convex function by using the proximal point algorithm with a general perturbation in an admissible complete CAT(1) space.

1 Introduction

The proximal point algorithm was started to study by Martinet, Rockafellar, and Brézis&Lions [3, 11, 13]. In 1976, Rockafellar [13] proposed this method in a Hilbert space. The sequence $\{x_n\}$ of a Hilbert space H which is generated by

$$x_{n+1} = \operatorname*{argmin}_{y \in H} \left\{ f(y) + \frac{1}{2\lambda_n} \|y - x_n\|^2 \right\}$$

is weakly convergent to a minimizer of a proper lower semicontinuous convex function.

In a complete CAT(0) space X, Bačák [1] showed that the sequence defined by

$$x_{n+1} = \operatorname*{argmin}_{y \in X} \left\{ f(y) + \frac{1}{2\lambda_n} d(y, x_n)^2 \right\}$$

satisfies the following properties:

- There exists a positive real number K such that

$$f(x_{n+1}) - \inf f(X) \leqq \frac{K}{\sum_{k=1}^{n} \lambda_k};$$

- $\{x_n\}$ is Δ-convergent to an element of $\operatorname{argmin}_X f$;

for more details, see also [12].

In 2017, Kimura and Kohsaka investigated the proximal point algorithm on an admissible complete CAT(1) space. Let X be an admissible complete CAT(1) space. They showed that the proximal mapping

$$Q_\lambda x = \operatorname*{argmin}_{y \in X} \left\{ f(y) + \frac{1}{\lambda} \tan d(y,x) \sin d(y,x) \right\}$$

of a proper lower semicontinuous convex function f is well defined as a single-valued mapping, see [8]. They also showed that the sequence generated by $x_{n+1} = Q_{\lambda_n} x_n$ satisfies

$$f(x_{n+1}) - \inf f(X) \leqq \frac{C}{\sum_{k=1}^{n} \lambda_k} (1 - \cos d(x_1, p))$$

for all $p \in X$, where C is a positive real number. Moreover, $\{x_n\}$ is Δ-convergent to an element of $\operatorname{argmin}_X f$, see also [9].

In this paper, we show some important properties of a resolvent for a convex function with a general perturbation on an admissible complete CAT(1) space. We further investigate the convergence theorem to a minimizer of a convex function by using a proximal point algorithm with a general perturbation.

2 Preliminaries

A mapping T on a metric space X is said to be quasi-nonexpansive if $d(Tx, u) \leqq d(x, u)$ for all $x \in X$ and $u \in \mathcal{F}(T)$, where $\mathcal{F}(T)$ is the set of all fixed points of T. For $x, y \in X$, a mapping γ_{xy} from $[0, \ell]$ into a metric space X is called a geodesic joining x and y if the following conditions hold:

- $\gamma_{xy}(0) = x$;
- $\gamma_{xy}(\ell) = y$;
- $d(\gamma_{xy}(s), \gamma_{xy}(t)) = |s - t|$ for all $s, t \in [0, \ell]$.

Let $D \in \,]0, \infty]$. A metric space X is called a D-geodesic space if there exists a geodesic γ_{xy} for each $x, y \in X$ with $d(x, y) < D$. In this paper, we always assume the uniqueness of γ_{xy}. In this case, we can define a convex combination between x and y by $tx \oplus (1-t)y = \gamma_{xy}((1-t)d(x, y))$ for each $x, y \in X$ and $t \in [0, 1]$.

A π-geodesic space X is called a CAT(1) space if the inequality

$$\cos d(tx \oplus (1-t)y, z) \sin d(x, y)$$

$$\geqq \cos d(x,z)\sin(td(x,y)) + \cos d(y,z)\sin((1-t)d(x,y))$$

holds for all $x, y, z \in X$ with $d(x,y) + d(y,z) + d(z,x) < 2\pi$ and $t \in [0,1]$. A CAT(1) space X is said to be admissible if $d(x,y) < \pi/2$ for all $x, y \in X$; for more details, see [2, 4, 14].

Let X be an admissible complete CAT(1) space. An asymptotic center $\mathcal{A}(\{x_n\})$ of a sequence $\{x_n\}$ of X is defined by the set

$$\mathcal{A}(\{x_n\}) = \left\{ u \in X \;\middle|\; \limsup_{n\to\infty} d(u, x_n) = \inf_{y \in X} \limsup_{n\to\infty} d(y, x_n) \right\}.$$

A sequence $\{x_n\}$ is Δ-convergent to an element x in X if $\mathcal{A}(\{x_{n_i}\}) = \{x\}$ for all subsequences $\{x_{n_i}\}$ of $\{x_n\}$. In this case, we call x a Δ-limit of $\{x_n\}$. The set of all Δ-limits of Δ-convergent subsequences of $\{x_n\}$ is denoted by $\omega_\Delta(\{x_n\})$. A sequence $\{x_n\}$ of X is said to be spherically bounded if

$$\inf_{y \in X} \limsup_{n\to\infty} d(y, x_n) < \pi/2.$$

The following crucial properties are well known.

Lemma 2.1 ([5, 10])**.** *Let X be an admissible complete* CAT(1) *space and $\{x_n\}$ a spherically bounded sequence of X. Then $\mathcal{A}(\{x_n\})$ consists of one point and $\{x_n\}$ has a Δ-convergent subsequence.*

Lemma 2.2 ([6, 7])**.** *Let X be an admissible complete* CAT(1) *space and $\{x_n\}$ a spherically bounded sequence of X. If $\{d(z, x_n)\}$ is strongly convergent for each $z \in \omega_\Delta(\{x_n\})$, then $\{x_n\}$ is Δ-convergent to an element of X.*

A function f from X into $]-\infty, \infty]$ is said to be

- proper if $f \not\equiv \infty$;
- lower semicontinuous if
$$f(x) \leqq \liminf_{n\to\infty} f(x_n)$$
whenever a sequence $\{x_n\}$ of X is strongly convergent to x;
- convex if
$$f(tx \oplus (1-t)y) \leqq tf(x) + (1-t)f(y)$$
for all $x, y \in X$ and $t \in]0,1[$.

It is known that if f is a proper lower semicontinuous convex function, then

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

whenever a sequence $\{x_n\}$ of X is Δ-convergent to x, see [8].

3 Properties of a resolvent of a convex function

In this section, we discuss important properties of a resolvent with a general perturbation of convex functions. Through this paper, we always assume that X is an admissible complete CAT(1) space, f is a proper lower semicontinuous convex function and the resolvent

$$R_\lambda x = \operatorname*{argmin}_{y\in X} \left\{ f(y) + \frac{1}{\lambda}\varphi(\cos d(y,x)) \right\}$$

of f is a single-valued mapping, where $\lambda > 0$ and φ a function from $]0,1]$ into $[0,\infty[$ satisfying the following conditions:

- φ is strictly decreasing;
- $\varphi(1) = 0$;
- φ' is differentiable and φ' is continuous.

It is well known that $\mathcal{F}(R_\lambda) = \operatorname{argmin}_X f$. For the sake of completeness, we give the proof. Let $u \in \operatorname{argmin}_X f$ and $y \in X$. Then we have

$$f(u) + \frac{1}{\lambda}\varphi(\cos d(u,u)) = f(u) \leqq f(y) \leqq f(y) + \frac{1}{\lambda}\varphi(\cos d(y,u))$$

and hence $R_\lambda u = u$. Thus $\operatorname{argmin}_X f \subset \mathcal{F}(R_\lambda)$. Conversely, let $u \in \mathcal{F}(R_\lambda)$, $y \in X\backslash\{u\}$ and $t \in]0,1[$. Then we get

$$\begin{aligned}\lambda f(u) &= \lambda f(u) + \varphi(\cos d(u,u))\\ &\leqq \lambda f(R_\lambda u) + \varphi(\cos d(R_\lambda u, u))\\ &\leqq \lambda f(ty \oplus (1-t)u) + \varphi(\cos d(ty \oplus (1-t)u, u)\\ &\leqq \lambda t f(y) + \lambda(1-t)f(u) + \varphi(\cos(td(y,u)))\end{aligned}$$

and hence

$$f(u) \leqq f(y) + \frac{\varphi(\cos(td(y,u)))}{\lambda t}.$$

Letting $t \downarrow 0$ and using l'Hospital's rule, we get $f(u) \leqq f(y)$ and hence $u \in \operatorname{argmin}_X f$. Therefore we get $\operatorname{argmin}_X f \supset \mathcal{F}(R_\lambda)$.

We next show important inequalities to consider convergence theorems with the proximal point algorithm.

Theorem 3.1. *Let λ and μ be positive real numbers. Put $C_{\lambda,z} = \cos d(R_\lambda z, z)$ for each $z \in X$. Then*

$$\lambda(f(R_\lambda x) - f(y)) \leqq \varphi'(C_{\lambda,x})\frac{d(R_\lambda x, y)}{\sin d(R_\lambda x, y))}(\cos d(x,y) - C_{\lambda,x}\cos d(R_\lambda x, y))$$

for all $x, y \in X$ *with* $R_\lambda x \neq y$ *and*

$$(\lambda\varphi'(C_{\mu,u})C_{\mu,y} + \mu\varphi'(C_{\lambda,x})C_{\lambda,x})\cos D$$
$$\leqq \lambda\varphi'(C_{\mu,y})\cos d(R_\lambda x, y) + \mu\varphi'(C_{\lambda,x})\cos d(x, R_\mu y)$$

for all $x, y \in X$*, where* $D = d(R_\lambda x, R_\mu y)$*.*

Proof. Let $x, y \in X$ with $R_\lambda x \neq y$ and $t \in]0,1[$, and put $\ell = d(R_\lambda x, y)$. Then, by the convexity of f and the definition of R_λ, we get

$$\begin{aligned}
&\lambda f(R_\lambda x) + \varphi(C_{\lambda,x}) \\
&\leqq \lambda f(ty \oplus (1-t)R_\lambda x) + \varphi(c_\kappa(d(ty \oplus (1-t)R_\lambda x, x))) \\
&\leqq t\lambda f(y) + (1-t)\lambda f(R_\lambda x) + \varphi\left(\frac{\cos d(x,y)\sin(t\ell) + C_{\lambda,x}\sin((1-t)\ell)}{\sin \ell}\right)
\end{aligned}$$

and hence

$$\lambda(f(R_\lambda x) - f(y)) \leqq \left(\varphi\left(\Delta(t)/\sin\ell\right) - \varphi(C_{\lambda,x})\right)/t,$$

where $\Delta(t) = \cos d(x,y)\sin(t\ell) + C_{\lambda,x}\sin((1-t)\ell)$. Then we notice that

$$\Delta'(t) = \ell\cos d(x,y)\cos(t\ell) - \ell C_{\lambda,x}\cos((1-t)\ell).$$

Taking the limit as t tends to 0 and using l'Hospital's rule, we have

$$\lambda(f(R_\lambda x) - f(y)) \leqq \varphi'(C_{\lambda,x})\frac{\ell}{\sin\ell}(\cos d(x,y) - C_{\lambda,x}\cos\ell).$$

This is the first inequality.

We next show the second inequality. Let $x, y \in X$ with $R_\lambda x \neq R_\mu y$. From the first inequality, we obtain

$$\mu\lambda(f(R_\lambda x) - f(R_\mu y)) \leqq \mu\varphi'(C_{\lambda,x})\frac{D}{\sin D}(\cos d(x, R_\mu y) - C_{\lambda,x}\cos D)$$

and

$$\mu\lambda(f(R_\mu y) - f(R_\lambda x)) \leqq \lambda\varphi'(C_{\mu,y})\frac{D}{\sin D}(\cos d(R_\lambda x, y) - C_{\mu,y}\cos D)$$

Adding both sides of these inequalities, we get the desired result. If $R_\lambda x = R_\mu y$ it is clearly satisfied. □

We remark that R_λ is quasi-nonexpansive. In fact, we have

$$(\varphi'(C_{\lambda,p}) + \varphi'(C_{\lambda,x}))\cos d(R_\lambda x, p) \leqq \varphi'(C_{\lambda,p})\cos d(R_\lambda x, p) + \varphi'(C_{\lambda,x})\cos(d(x,p))$$

for all $x \in X$ and $p \in \mathcal{F}(R_\lambda)$. Therefore, we get $\cos d(R_\lambda x, p) \geqq \cos d(x,p)$ and thus $d(R_\lambda x, p) \leqq d(x,p)$. Consequently, R_λ is quasi-nonexpansive.

4 The proximal point algorithm

In this section, we show a convergence theorem to a minimizer of a convex function by using the proximal point algorithm.

Theorem 4.1. *Let* $\{\lambda_n\}$ *be a sequence of* $]0,\infty[$ *with* $\sum_{n=1}^{\infty}\lambda_n = \infty$, $x_1 \in X$ *and* $x_{n+1} = R_{\lambda_n} x_n$ *for all* $n \in \mathbb{N}$. *Suppose that* φ' *satisfies either strict increasingness or nonincreasingness. If* $\operatorname{argmin}_X f$ *is nonempty, then there exists a negative real number* L *such that*

$$f(x_{n+1}) - f(p) \leqq \frac{\pi L}{2\sum_{k=1}^{n}\lambda_k}(\cos d(x_1,p) - 1)$$

for all $p \in \operatorname{argmin}_X f$. *Moreover,* $\{x_n\}$ *is* Δ*-convergent to a minimizer of* f.

Proof. If φ' is nonincreasing, then

$$\varphi'(C_{\lambda_k,x_k}) \geqq \inf_{m\in\mathbb{N}} \varphi'(C_{\lambda_m,x_m}) = \varphi'(1) \in \left]-\infty, 0\right[.$$

Suppose that φ' is strictly increasing. Let $p \in \operatorname{argmin}_X f$. By Theorem 3.1, we have

$$\begin{aligned} 0 &\leqq \lambda_m(f(x_{m+1}) - f(p)) \\ &\leqq \varphi'(C_{\lambda_m,x_m})\frac{d(x_{m+1},p)}{\sin d(x_{m+1},p)}(\cos d(x_m,p) - C_{\lambda_m,x_m}\cos d(x_{m+1},p)) \end{aligned}$$

for all $m \in \mathbb{N}$. Since φ is strictly decreasing, we obtain

$$\cos d(x_m,p) - \cos d(x_{m+1},x_m)\cos d(x_{m+1},p) \leqq 0$$

and hence

$$\cos d(x_m,p) \leqq \cos d(x_{m+1},x_m)\cos d(x_{m+1},p) \leqq \cos d(x_{m+1},x_m).$$

Therefore, we get $d(x_{m+1},x_m) \leqq d(x_m,p) \leqq \cdots \leqq d(x_1,p)$ since R_λ is quasi-nonexpansive. It implies that

$$\varphi'(C_{\lambda_k,x_k}) \geqq \inf_{m\in\mathbb{N}} \varphi'(C_{\lambda_m,x_m}) = \varphi'\left(\cos\left(\sup_{m\in\mathbb{N}} d(x_{m+1},x_m)\right)\right) \geqq \varphi'(\cos d(x_1,p)).$$

Thus, there exists a negative real number L such that $\varphi'(C_{\lambda_k,x_k}) \geqq L$. Further, we notice that $t/\sin t \leqq \pi/2$.

On the other hand, by the definition of x_n and R_{λ_n}, we have

$$\begin{aligned} f(p) &\leqq f(x_{n+1}) \\ &\leqq f(x_{n+1}) + \frac{1}{\lambda_n}\varphi(\cos d(x_{n+1},x_n)) \end{aligned}$$

$$\leqq f(x_n) + \frac{1}{\lambda_n}\varphi(\cos d(x_n, x_n)) = f(x_n)$$

for all $n \in \mathbb{N}$. Moreover, by Theorem 3.1, we have

$$\begin{aligned}\lambda_k(f(x_{k+1}) - f(p)) &\leqq \varphi'(C_{\lambda_k, x_k})\frac{d(x_{k+1}, p)}{\sin d(x_{k+1}, p)}(\cos d(x_k, p) - \cos d(x_{k+1}, p)) \\ &\leqq \frac{\pi L}{2}(\cos d(x_k, p) - \cos d(x_{k+1}, p)),\end{aligned}$$

and therefore

$$\sum_{k=1}^{n}\lambda_k(f(x_{k+1}) - f(p)) \leqq \frac{\pi L}{2}(\cos d(x_1, p) - \cos d(x_{n+1}, p)) \leqq \frac{\pi L}{2}(\cos d(x_1, p) - 1).$$

Since $f(p) \leqq f(x_{k+1}) \leqq f(x_k)$ for all $k \in \{1, 2, \dots, n\}$, we obtain

$$(f(x_{n+1}) - f(p))\sum_{k=1}^{n}\lambda_k \leqq \frac{\pi L}{2}(\cos d(x_1, p) - 1)$$

and thus we get the desired result.

We next show that $\{x_n\}$ is Δ-convergent to $p \in \operatorname{argmin}_X f$. By the inequality above, $\{f(x_n)\}$ is strongly convergent to $f(p)$ as n tends to ∞. Take a Δ-convergent subsequence $\{x_{n_i}\}$ of $\{x_n\}$. Let z be the Δ-limit of $\{x_{n_i}\}$. Then, we know that

$$f(z) \leqq \liminf_{i\to\infty} f(x_{n_i}) = f(p)$$

and thus $z \in \operatorname{argmin}_X f = \mathcal{F}(R_{\lambda_n})$. Since R_{λ_n} is quasi-nonexpansive, we have $0 \leqq d(x_{n+1}, z) \leqq d(x_n, z)$. Therefore, $\{d(x_n, z)\}$ is strongly convergent. Consequently, by Lemma 2.2, $\{x_n\}$ is Δ-convergent to $x_0 \in X$. From the argument above, x_0 is an element of $\operatorname{argmin}_X f$. □

Acknowledgment. This work was partially supported by JSPS KAKENHI Grant Number JP21K03316.

References

[1] M. Bačák, *The proximal point algorithm in metric spaces*, Isreal J. Math. **29** (2013), 689–701.

[2] M. Bačák, *Convex Analysis and Optimization in Hadamard Spaces*, De Gruyter, Würzbrung, 2014.

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

[4] M. R. Bridson and A. Haefliger, *Metric Spaces of Non-Positive Curvature*, Springer-Verlag, Berlin, 1999.

[5] R. Espínola and A. Fernández-León, CAT(κ)-*spaces, weak convergence and fixed points*, J. Math. Anal. Appl. **65** (2006), 762–772.
[6] Y. Kimura, *Convergence of a sequence of sets in a Hadamard space and the shrinking projection method for a real Hilbert ball*, Abstr. Appl. Anal. **2010** (2010), 11pp.
[7] Y. Kimura, S. Saejung, and P. Yptkaew, *The Mann algorithm in a complete geodesic space with curvature bounded above*, Fixed Point Theory Appl. **2013** (2013), doi: 10.1186/1687-1812-2013-336, 13 pages.
[8] Y. Kimura and F. Kohsaka, *Spherical nonspreadingness of resolvents of convex functions in geodesic spaces*, J. Fixed Point Theory Appl. **18** (2016), 93–115.
[9] Y. Kimura and F. Kohsaka, *The proximal point algorithm in geodesic spaces with curvature bounded above*, Linear and Nonlinear Analysis **3**, No. 1 (2017), 73–86.
[10] W. A. Kirk and B. Panyanak, *A concept of convergence in geodesic spaces*, Nonlinear Anal. **68** (2008), 3689–3696.
[11] B. Martinet, *Régularisation d'inéquations variationnelles par approximations successives*, Rev. Francaise Informat. Recherche Opérationnelle, **4** (1970), pp. 154–158.
[12] U. F. Mayer, *Gradient flows on nonpositively curved metric spaces and harmonic maps*, Comm. Anal. Geom. **6** (1998), 199–206.
[13] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. **14** (1976), 877–898.
[14] S. Sudo, *Parallelogram laws on geodesic spaces*, Master theses, Toho University, 2023.