Proceedings of the International Conference on
Nonlinear Analysis and Convex Analysis &
International Conference on Optimization:
Techniques and Applications -I-
(Hakodate, Japan, 2019), 119–125

HALPERN-TYPE SUBGRADIENT METHODS FOR CONVEX OPTIMIZATION OVER FIXED POINT SETS OF NONEXPANSIVE MAPPINGS

HIDEAKI IIDUKA

ABSTRACT. Convex optimization over fixed point sets has applications such as network resource allocation and machine learning. In this paper, we present methods combining the Halpern fixed point approximation method with subgradient methods for solving the problems and their convergence analyses.

1. INTRODUCTION

In this paper, we consider a convex optimization problem with fixed point constraints of nonexpansive mappings. Solutions to this problem have practical applications such as network resource allocation [7, 8, 9, 13, 15] and machine learning [6, 14].

Iterative methods have been presented for solving the problem. Reference [12] presented incremental proximal methods based on the Krasnosel'skiĭ-Mann fixed point algorithm [16, 17] and the Halpern fixed point algorithm [5, 22]. Reference [20] presented parallel proximal methods based on the Krasnosel'skiĭ-Mann fixed point algorithm and the Halpern fixed point algorithm. Meanwhile, incremental and parallel subgradient methods based on the Krasnosel'skiĭ-Mann fixed point algorithm were presented in [10, 11].

In this paper, we present incremental and parallel optimization methods combining the Halpern fixed point algorithm with subgradient methods for solving the problem and their convergence analyses.

This paper is organized as follows. Section 2 gives the mathematical preliminaries. Section 3 considers the problem of minimizing the sum of convex functions over the intersection of nonexpansive mappings and presents incremental and parallel subgradient methods for solving the problem together with

2010 *Mathematics Subject Classification.* 65K05, 90C25, 90C90.
Key words and phrases. Convex optimization, fixed point, Halpern method, nonexpansive mapping, subgradient method.

their convergence analyses. Section 4 concludes the paper with a brief summary and mention of future work.

2. Mathematical preliminaries

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$ and Id denote the identity mapping on H. Let $\mathbb{N}$ denote the set of all positive integers including zero and $\mathbb{R}$ denote the set of all real numbers.

The *subdifferential* [1, Definition 16.1], [18, Section 23] of a convex function $f \colon H \to \mathbb{R}$ is defined for all $x \in H$ by $\partial f(x) := \{u \in H \colon f(y) \geq f(x) + \langle y - x, u \rangle \ (y \in H)\}$. A point $u \in \partial f(x)$ is called the *subgradient* of f at $x \in H$. $A \colon H \rightrightarrows H$ is said to be *inverse-strongly monotone* (α-inverse-strongly monotone) [3, Definition, p.200] (see [1, Definition 4.4], [4, Definition 2.3.9(e)] for the definition of this operator, which is called a cocoercive operator) if there exists $\alpha > 0$ such that, for all $x, y \in H$, for all $u \in A(x)$, and for all $v \in A(y)$, $\langle x - y, u - v \rangle \geq \alpha \|u - v\|^2$.

$T \colon H \to H$ is said to be Lipschitz continuous (L-Lipschitz continuous) if there exists $L > 0$ such that $\|T(x) - T(y)\| \leq L\|x - y\|$ for all $x, y \in H$. T is said to be *nonexpansive* [1, Definition 4.1(ii)] if T is 1-Lipschitz continuous, i.e., $\|T(x) - T(y)\| \leq \|x - y\|$ for all $x, y \in H$. The *metric projection* [1, Subchapter 4.2, Chapter 28] onto a nonempty, closed convex set C $(\subset H)$, denoted by P_C, is defined for all $x \in H$ by $P_C(x) \in C$ and $\|x - P_C(x)\| = \mathrm{d}(x, C) := \inf_{y \in C} \|x - y\|$. P_C is *firmly nonexpansive*, i.e., $\|P_C(x) - P_C(y)\|^2 + \|(\mathrm{Id} - P_C)(x) - (\mathrm{Id} - P_C)(y)\|^2 \leq \|x - y\|^2$ for all $x, y \in H$, with $\mathrm{Fix}(P_C) = C$ [1, Proposition 4.8, (4.8)], where $\mathrm{Fix}(T)$ is the *fixed point set* of a mapping T defined by $\mathrm{Fix}(T) := \{x \in H \colon x = T(x)\}$.

The following is the Halpern fixed point approximation method [5, 22] for finding a fixed point of a nonexpansive mapping $T \colon H \to H$: for all $n \in \mathbb{N}$,

$$x_{n+1} := \alpha_n x_0 + (1 - \alpha_n) T(x_n), \tag{2.1}$$

where $x_0 \in H$ and $(\alpha_n)_{n \in \mathbb{N}} \subset (0, 1]$. The sequence $(x_n)_{n \in \mathbb{N}}$ generated by (2.1) with $(\alpha_n)_{n \in \mathbb{N}}$ satisfying $\lim_{n \to +\infty} \alpha_n = 0$ and $\sum_{n=0}^{+\infty} \alpha_n = +\infty$ converges strongly to the minimizer of $\| \cdot - x_0\|^2$ over $\mathrm{Fix}(T)$ [2, Theorem 6.19].

Thanks to [19, Proposition 12.60] and [1, Theorem 18.15], we have the following proposition.

Proposition 2.1. *Let $f \colon H \to \mathbb{R}$ be convex and continuous. Then, the following properties are equivalent:*

(i) *∂f is $(1/L)$-inverse-strongly monotone;*
(ii) *f is Fréchet differentiable and ∇f is L-Lipschitz continuous.*

3. Subgradient methods

In this paper, we consider the following problem (see also [11, Problem 2.1] and [12, Problem 2.1]):

Problem 3.1.

$$\text{Minimize } f(x) := \sum_{i \in \mathcal{I}} f_i(x) \text{ subject to } x \in \bigcap_{i \in \mathcal{I}} \mathrm{Fix}(T_i),$$

where we assume that

(A1) $T_i \colon H \to H$ $(i \in \mathcal{I} := \{1, 2, \ldots, I\})$ is firmly nonexpansive with $\bigcap_{i \in \mathcal{I}} \mathrm{Fix}(T_i) \neq \emptyset$;

(A2) $f_i \colon H \to \mathbb{R}$ is convex and continuous with $\mathrm{dom}(f_i) := \{x \in H \colon f_i(x) < +\infty\} = H$, $\partial f_i \colon H \rightrightarrows H$ $(i \in \mathcal{I})$ is $(1/L)$-inverse-strongly monotone, and the subgradient of f_i at any $x \in H$ can be efficiently computed.

We present the following subgradient method based on the Halpern fixed point approximation method (2.1) (step 6 in Algorithm 1) for solving Problem 3.1.

Algorithm 1 Incremental subgradient method for solving Problem 3.1

Require: $(\alpha_n)_{n \in \mathbb{N}} \subset (0, 1]$, $(\lambda_n)_{n \in \mathbb{N}} \subset (0, +\infty)$

1: $n \leftarrow 0$, $x_0 = x_{0,0} \in H$, $\bar{x}_i \in H$ $(i \in \mathcal{I})$

2: **loop**

3:   **for** $i = 1, 2, \ldots, I$ **do**

4:     $g_{n,i} \in \partial f_i(x_{n,i-1})$

5:     $y_{n,i} := T_i(x_{n,i-1} - \lambda_n g_{n,i})$

6:     $x_{n,i} := \alpha_n \bar{x}_i + (1 - \alpha_n) y_{n,i}$

7:   **end for**

8:   $x_{n+1} = x_{n,I} = x_{n+1,0}$

9:   $n \leftarrow n + 1$

10: **end loop**

Consider a network system with I users and suppose that user i has its own private objective function f_i and firmly nonexpansive mapping T_i. Furthermore, assume that user i can communicate with user $(i - 1)$, where user 0 is user I. This implies that user i can use $x_{n,i-1}$, which is computed by user $(i - 1)$. Since user i tries to minimize f_i over $\mathrm{Fix}(T_i)$, user i computes $y_{n,i} = T_i(x_{n,i-1} - \lambda_n g_{n,i})$ (step 5 in Algorithm 1) using $x_{n,i-1}$ and $g_{n,i} \in \partial f_i(x_{n,i-1})$. User i then computes $x_{n,i} = \alpha_n \bar{x}_i + (1 - \alpha_n) y_{n,i}$ (step 6 in Algorithm 1) to find a fixed point of T_i. Accordingly, each user in the network system can implement Algorithm 1. Problem 3.1 in such a network system includes network resource allocation [9, 13] and machine learning [6, 14].

We assume the following:

Assumption 3.2. The decreasing sequences $(\alpha_n)_{n\in\mathbb{N}}$ and $(\lambda_n)_{n\in\mathbb{N}}$ converge to 0 and satisfy the following conditions[1]:

$$\text{(C1)}\ \sum_{n=0}^{+\infty}\alpha_n = +\infty,\ \text{(C2)}\ \lim_{n\to+\infty}\frac{1}{\alpha_{n+1}}\left|\frac{1}{\lambda_{n+1}}-\frac{1}{\lambda_n}\right| = 0,\ \text{(C3)}\ \lim_{n\to+\infty}\frac{\alpha_n}{\lambda_n} = 0,$$

$$\text{(C4)}\ \lim_{n\to+\infty}\frac{1}{\lambda_{n+1}}\left|1-\frac{\alpha_n}{\alpha_{n+1}}\right| = 0,\ \text{(C5)}\ \frac{\lambda_n}{\lambda_{n+1}} \leq \sigma \text{ for some } \sigma \geq 1.$$

Moreover, $(y_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathcal{I})$ is bounded[2].

The following is a convergence analysis of Algorithm 1.

Proposition 3.3. *Consider Problem 3.1 and suppose that Assumption 3.2 holds. Then, any weak sequential cluster point of the sequence $(x_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathcal{I})$ generated by Algorithm 1 belongs to the solution set of Problem 3.1.*

Proof. Assumptions (A1), (A2), and 3.2 and Proposition 2.1 imply that the assumptions in [9, Lemma 3.2] hold. Accordingly, the proof of [9, Lemma 3.2] ensures that

$$\lim_{n\to+\infty}\|x_n - T_i(x_n)\| = 0\ (i \in \mathcal{I}) \text{ and } \limsup_{n\to+\infty} f(x_n) \leq f^\star,$$

where $f^\star$ is the optimal value of Problem 3.1. Let $x^\star$ be any weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$. Then, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ which converges weakly to $x^\star$. From $\lim_{n\to+\infty}\|x_n - T_i(x_n)\| = 0$ $(i \in \mathcal{I})$ (see also the proof of [9, Lemma 3.2]), we have that $x^\star \in \bigcap_{i\in\mathcal{I}}\mathrm{Fix}(T_i)$. Moreover, from $\limsup_{n\to+\infty} f(x_n) \leq f^\star$ and the continuity of f, we have

$$f(x^\star) \leq \liminf_{k\to+\infty} f(x_{n_k}) \leq \limsup_{k\to+\infty} f(x_{n_k}) \leq \limsup_{n\to+\infty} f(x_n) \leq f^\star,$$

which implies that $x^\star$ is a solution of Problem 3.1. Here, let $j \in \{1, 2, \ldots, I-1\}$ be fixed arbitrarily and let x_j^* be any weak sequential cluster point of $(x_{n,j})_{n\in\mathbb{N}}$. Then, there exists a subsequence $(x_{n_l,j})_{l\in\mathbb{N}}$ of $(x_{n,j})_{n\in\mathbb{N}}$ which converges weakly to x_j^*. From $\lim_{n\to+\infty}\|x_n - x_{n,i-1}\| = 0$ $(i \in \mathcal{I})$ (see [9, Lemma 3.2(iii)]), we have that $(x_{n_l})_{l\in\mathbb{N}}$ weakly converges to x_j^*. A discussion similar to the one

[1] Examples of $(\lambda_n)_{n\in\mathbb{N}}$ and $(\alpha_n)_{n\in\mathbb{N}}$ are $\lambda_n = 1/(n+1)^a$ and $\alpha_n = 1/(n+1)^b$, where $a \in (0, 1/2)$ and $b \in (a, 1-a)$.

[2] See the discussion in [9, Assumption 3.2] for examples satisfying the boundedness of $(y_{n,i})_{n\in\mathbb{N}}$ $(i \in \mathcal{I})$.

for showing that $x^\star$ is a solution of Problem 3.1 guarantees that x_j^* is also a solution of Problem 3.1. This completes the proof. □

Next, we present the following algorithm.

Algorithm 2 Parallel subgradient method for solving Problem 3.1

Require: $(\alpha_n)_{n\in\mathbb{N}} \subset (0,1]$, $(\lambda_n)_{n\in\mathbb{N}} \subset (0,+\infty)$
1: $n \leftarrow 0$, $x_0 \in H$, $\bar{x}_i \in H$ $(i \in \mathcal{I})$
2: **loop**
3: **for** $i = 1, 2, \ldots, I$ **do**
4: $g_{n,i} \in \partial f_i(x_n)$
5: $y_{n,i} := T_i(x_n - \lambda_n g_{n,i})$
6: $x_{n,i} := \alpha_n \bar{x}_i + (1-\alpha_n) y_{n,i}$
7: **end for**
8: $x_{n+1} = \dfrac{1}{I} \displaystyle\sum_{i\in\mathcal{I}} x_{n,i}$
9: $n \leftarrow n+1$
10: **end loop**

Consider a network system with I users and suppose that user i has its own private objective function f_i and firmly nonexpansive mapping T_i. We also assume the existence of the operator managing the network system. This implies that the operator can use $x_{n,i}$, which is computed by user i and that each user knows x_n transmitted from the operator. Since user i tries to minimize f_i over $\mathrm{Fix}(T_i)$, user i computes $y_{n,i} = T_i(x_n - \lambda_n g_{n,i})$ (step 5 in Algorithm 2) using x_n and $g_{n,i} \in \partial f_i(x_n)$. User i then computes $x_{n,i} = \alpha_n \bar{x}_i + (1-\alpha_n) y_{n,i}$ (step 6 in Algorithm 2) to find a fixed point of T_i. The operator can compute $x_{n+1} = (1/I)\sum_{i\in\mathcal{I}} x_{n,i}$, since the operator knows all of $x_{n,i}$ (step 8 in Algorithm 2). Accordingly, the operator and each user in the network system can implement Algorithm 2. Problem 3.1 in such a network system with an operator includes storage allocation [15]. Reference [21] researched the actual computation times of parallel and incremental subgradient methods by using parallel computing on multi-core processors for a concrete convex optimization problem.

The following is a convergence analysis of Algorithm 2.

Proposition 3.4. *Consider Problem 3.1 and suppose that Assumption 3.2 holds. Then, any weak sequential cluster point of the sequence $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 2 belongs to the solution set of Problem 3.1.*

Proof. Assumptions (A1), (A2), and 3.2 and Proposition 2.1 imply that the assumptions in [9, Lemma 4.2] hold. Accordingly, [9, Lemma 4.2(iii), (4.8)] ensures that

$$\lim_{n\to+\infty} \|x_n - T_i(x_n)\| = 0 \ (i \in \mathcal{I}) \text{ and } \limsup_{n\to+\infty} f(x_n) \leq f^\star,$$

where $f^\star$ is the optimal value of Problem 3.1. A discussion similar to the one showing that $x^\star$ is a solution of Problem 3.1 (the proof of Proposition 3.3) leads to the assertion in Proposition 3.4. □

4. Conclusion and future work

This paper presented two subgradient methods, based on the Halpern fixed point approximation method, for solving the problem of minimizing the sum of convex functions over the intersection of fixed point sets of nonexpansive mappings in a real Hilbert space. It also presented their convergence analyses under the condition that the subdifferential of each convex function is inverse-strongly monotone. Since the condition implies that each convex function has the Lipschitz gradient, it would be strong. Accordingly, we should develop Halpern-type subgradient methods without assuming this condition.

Acknowledgments. The author would like to thank Professors Yasunori Kimura, Masakazu Muramatsu, Wataru Takahashi, and Akiko Yoshise for giving me a chance to submit my paper to this journal. The author is sincerely grateful to the anonymous reviewers for helping me improve the original manuscript. This work was supported by a JSPS KAKENHI Grant, Number JP18K11184.

References

[1] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2nd ed., 2017.

[2] V. Berinde, *Iterative Approximation of Fixed Points*, Springer, Berlin, 2007.

[3] F. E. Browder and W. V. Petryshyn, *Construction of fixed points of nonlinear mappings in Hilbert space*, J. Math. Anal. Appl. **20** (1967), 197–228.

[4] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems I*, Springer, New York, 2003.

[5] B. Halpern, *Fixed points of nonexpanding maps*, Bull. Am. Math. Soc. **73** (1967), 957–961.

[6] Y. Hayashi and H. Iiduka, *Optimality and convergence for convex ensemble learning with sparsity and diversity based on fixed point optimization*, Neurocomputing, **273** (2018), 367–372.

[7] H. Iiduka, *Fixed point optimization algorithm and its application to network bandwidth allocation*, J. Comput. Appl. **236** (2012), 1733–1742.

[8] H. Iiduka, *Fixed point optimization algorithm and its application to power control in CDMA data networks*, Math. Program. **133** (2012), 227–242.

[9] H. Iiduka, *Fixed point optimization algorithms for distributed optimization in networked systems*, SIAM J. Optim. **23** (2013), 1–26.

[10] H. Iiduka, *Parallel computing subgradient method for nonsmooth convex optimization over the intersection of fixed point sets of nonexpansive mappings*, Fixed Point Theory Appl. **2015** (2015), 72.

[11] H. Iiduka, *Incremental subgradient method for nonsmooth convex optimization with fixed point constraints*, Optim. Methods Softw. **31** (2016), 931–951.

[12] H. Iiduka, *Proximal point algorithms for nonsmooth convex optimization with fixed point constraints*, Eur. J. Oper. Res. **253** (2016), 503–513.

[13] H. Iiduka, *Distributed optimization for network resource allocation with nonsmooth utility functions*, IEEE Trans. Control. Netw. Syst. **6** (2019), 1354–1365.

[14] H. Iiduka, *Stochastic fixed point optimization algorithm for classifier ensemble*, IEEE Trans. Cybern. **50** (2020), 4370–4380.

[15] H. Iiduka and K. Hishinuma, *Acceleration method combining broadcast and incremental distributed optimization algorithms*, SIAM J. Optim. **24** (2014), 1840–1863.

[16] M. A. Krasnosel'skiĭ, *Two remarks on the method of successive approximations*, Uspekhi Mat. Nauk **10** (1955), 123–127.

[17] W. R. Mann, *Mean value methods in iteration*, Proc. Am. Math. Soc. **4** (1953), 506–510.

[18] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, New Jersey, 1970.

[19] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 3rd ed., 2010.

[20] K. Sakurai, T. Jimba, and H. Iiduka, *Iterative methods for parallel convex optimization with fixed point constraints*, Journal of Nonlinear and Variational Analysis **3** (2019), 115–126.

[21] K. Shimizu and H. Iiduka, *Computation time of iterative methods for nonsmooth convex optimization with fixed point constraints of quasi-nonexpansive mappings*, Linear and Nonlinear Analysis **6** (2020), 281–286.

[22] R. Wittmann, *Approximation of fixed points of nonexpansive mappings*, Arch. Math. **58** (1992), 486–491.

H. Iiduka
Department of Computer Science, Meiji University, 1-1-1 Higashimita, Tama-ku, Kawasaki-shi, Kanagawa 214-8571, Japan
E-mail address: iiduka@cs.meiji.ac.jp