Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 6, No. 1, pp. 187–199, January 2010

ON AN ERGODIC METHOD FOR A CONVEX OPTIMIZATION PROBLEM OVER THE FIXED POINT SET*

Hideaki Iiduka

Abstract: In this paper, we consider a convex optimization problem over the fixed point set of a nonexpansive mapping, and present an ergodic iteration method for this problem together with its convergence analysis. The proposed algorithm has two features: one is that it can be applied to more general case, where the objective function is convex and Fréchet differentiable and has the hemicontinuous gradient; and the other is that as compared with the existing methods for convex optimization problems with Fréchet differentiable objective functions, the proposed algorithm does not require to solve any auxiliary optimization problems. To demonstrate convergence of the proposed method, we present numerical examples for some quadratic optimization problems over the fixed point set.

Key words: *convex optimization problem, convex function, nonexpansive mapping, fixed point, ergodic algorithm, weak convergence*

Mathematics Subject Classification: *47H06, 47J20, 47J25*

1 Introduction

In this paper, we present a new algorithm to the following *convex optimization problem over the fixed point set* [34]: let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. Given a convex, Fréchet differentiable function, $f\colon H \to \mathbb{R}$, and a nonexpansive mapping, $T\colon H \to H$, with $\mathrm{Fix}(T) := \{x \in H\colon T(x) = x\} \neq \emptyset$,

$$\text{find a point } z \in \underset{x \in \mathrm{Fix}(T)}{\operatorname{argmin}} f(x) := \left\{ z \in \mathrm{Fix}(T)\colon f(z) = \min_{x \in \mathrm{Fix}(T)} f(x) \right\}. \tag{1.1}$$

Thanks to introduction of Problem (1.1), we can discuss constrained optimization problems for the cases where the closed form expression of the *metric projection* (see Sec.2) onto the constrained set is not known, for example, important optimization problems for signal processing and inverse problems [6, 27, 28, 35]. Some iterative procedures [6, 15–18, 24, 25, 34] for (1.1) have been presented. In the case where the gradient ∇f of f is strongly monotone and Lipschitz continuous, the following method has been proposed [34]: $x_1 \in H$ and $x_{n+1} = T(x_n) - \mu \alpha_n \nabla f(T(x_n))$ $(n \in \mathbb{N})$, where $\mu > 0$ and $(\alpha_n)_{n \in \mathbb{N}}$ is a slowly diminishing constant sequence. The convergence of $(x_n)_{n \in \mathbb{N}}$ to the uniquely existing solution of (1.1) is also guaranteed [34]. Recently, in order to accelerate the method in [34], iterative algorithms [16, 18] using conjugate gradient directions have been proposed. Other algorithms for

*This work was supported by the Japan Society for the Promotion of Science (JSPS) through a Grant-in-Aid (19001979).

solving Problem (1.1) when ∇f is strongly monotone and Lipschitz continuous have been proposed in [6, 15]. In [6], an effective scheme for solving the signal recovery problem has been proposed and this method converges strongly to the solution without using a diminishing constant sequence. In [15], the *variational inequality problem* (see Sec.2) which contains (1.1) and an iterative algorithm for this problem have been presented. In the case where ∇f is inverse-strongly monotone (see Sec.2), iterative algorithms for (1.1) and its convergence analysis have been proposed in [24, 25]. In the case where C is a nonempty, closed convex subset of $\mathbb{R}^N$ which is simple enough to have a closed form expression of the metric projection P_C and ∇f is a monotone, continuous operator, a projection method for the variational inequality problem has been presented in [29]. This method requires us to solve some auxiliary problem and converges to some solution of Problem (1.1) when $T = P_C$. In the case where f is convex (and is not necessarily differentiable), a subgradient-type method for (1.1) and its convergence analysis have been presented in [17]. As this method requires us to solve an auxiliary maximization problem over the closed ball at every iteration, applications of this method are limited unfortunately.

On the other hand, an *ergodic iterative method* [5] for the variational inequality problem is summarized as: let $f\colon H \to \mathbb{R}$ be lower semicontinuous and convex. Given $x_n \in H$ and $\lambda_n > 0$, choose $\xi_n \in \partial f(x_n)$ (see Sec.2) arbitrarily and compute $x_{n+1} = P_C(x_n - \lambda_n \xi_n)$ and $z_n = \sum_{i=1}^n \lambda_i x_i / \sum_{i=1}^n \lambda_i$. Obviously this method assumes that C is a closed convex set which is simple enough to have a closed form expression of P_C. If $\sum_{n=1}^\infty \lambda_n = \infty$ and if $\sum_{n=1}^\infty \lambda_n^2 \|\xi_n\|^2 < \infty$, then the sequence, $(z_n)_{n\in\mathbb{N}}$, converges weakly to a point in the solution set $\{x \in C\colon \langle v - x, v^* \rangle \geq 0$ for all $v \in C$ and for all $v^* \in \partial f(v)\} \supset \{x \in C\colon \langle v - x, x^* \rangle \geq 0$ for all $v \in C$ and for all $x^* \in \partial f(x)\}$ (see Sec.2). Moreover, if f is Fréchet differentiable and if ∇f is hemicontinuous (see Sec.2), then $(z_n)_{n\in\mathbb{N}}$ converges weakly to a solution of Problem (1.1) when $T = P_C$.

The goal of this paper is to propose a new iteration method to Problem (1.1) which does not require to solve any auxiliary optimization problems. To this goal, we present an ergodic algorithm for (1.1) by combining the ideas of a scheme [34] for a convex optimization problem over the fixed point set and an ergodic iteration [5] for the variational inequality problem. The proposed algorithm can use any nonexpansive mapping T such that $\mathrm{Fix}(T)$ is equal to the constrained set. Hence, our algorithm can be applied to many practical situations where no closed form expression of the constrained set is known. In addition, the conditions on the objective function are weaker than the ones of [6, 15, 16, 24, 25, 34]. Thus, it is anticipated that the proposed algorithm will be used to important problems to which the existing methods [6, 15, 16, 34] are not applied (On an ergodic method for power control for the uplink of code-division multiple-access system, see [19]). In this paper, it is shown that the sequence generated by the proposed algorithm converges weakly to a solution of (1.1) under some assumptions.

The rest of this paper is divided into three sections. In Section 2, we state preliminaries on fixed points, nonexpansive mappings, metric projections, monotone operators, and variational inequality problems. In Section 3, we present an ergodic iteration method (Algorithm 3.1) for a convex optimization problem over the fixed point set of a nonexpansive mapping together with its convergence analysis (Theorem 3.2) for the problem. Numerical examples on the proposed algorithm for a quadratic optimization problem over the fixed point set are presented in Section 4.

2 Preliminaries

2.1 Convexity, Continuity, and Monotonicity

A function $f\colon H \to \mathbb{R}$ is said to be *convex* if, for any $x, y \in H$ and for any $\lambda \in [0,1]$, $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$. It is well known that any convex, Fréchet differentiable function $f\colon H \to \mathbb{R}$ is continuous and that, for all $x \in H$,

$$\partial f(x) := \{z \in H\colon f(y) \geq f(x) + \langle y - x, z\rangle \text{ for all } y \in H\} \neq \emptyset.$$

The set-valued operator $\partial f\colon H \to 2^H$ is called a *subdifferential* of f. A subdifferential of a convex function f satisfies the *monotonicity*, that is, $\langle x - y, z - w\rangle \geq 0$ for all $x, y \in H$ and for all $z \in \partial f(x)$ and $w \in \partial f(y)$. A convex, Fréchet differentiable function f implies that, for all $x \in H$, $\partial f(x) = \{\nabla f(x)\}$ holds. An operator $A\colon H \to H$ is said to be *hemicontinuous* (for example, see [32]) if, for any $x, y, z \in H$, a mapping $g\colon [0,1] \to \mathbb{R}$ defined by $g(t) := \langle z, A(tx + (1-t)y)\rangle$ $(t \in [0,1])$ is continuous. Define a quadratic function $f\colon \mathbb{R}^N \to \mathbb{R}$ by $f(x) := (1/2)\langle x, Q(x)\rangle + \langle b, x\rangle$ for all $x \in \mathbb{R}^N$, where $Q \in \mathbb{R}^{N\times N}$ is positive semidefinite and $b \in \mathbb{R}^N$. Then, $\nabla f(\cdot) = Q(\cdot) + b$ is monotone and hemicontinuous. $A\colon H \to H$ is said to be *Lipschitz continuous* if there exists $L > 0$ such that $\|A(x) - A(y)\| \leq L\|x - y\|$ for all $x, y \in H$. In such a case, A is called L-Lipschitz continuous. An operator $A\colon H \to H$ is said to be *inverse-strongly-monotone* [1,4,8,22,37,38] if there exists $\alpha > 0$ such that $\langle x - y, A(x) - A(y)\rangle \geq \alpha\|A(x) - A(y)\|^2$ for all $x, y \in H$. Suppose that $f\colon H \to \mathbb{R}$ is convex and continuously Fréchet differentiable and that $\nabla f\colon H \to H$ is L-Lipschitz continuous. Then, ∇f is $1/L$-inverse-strongly-monotone [1]. Suppose that $Q \in \mathbb{R}^{N\times N}$ is a positive semidefinite, that the maximum eigenvalue $\lambda_{\max}$ of Q is positive, and that $b \in \mathbb{R}^N$. Define $f(x) := (1/2)\langle x, Q(x)\rangle + \langle b, x\rangle$ $(x \in \mathbb{R}^N)$. Then, $\nabla f(\cdot) := Q(\cdot) + b$ is $\lambda_{\max}$-Lipschitz continuous and $1/\lambda_{\max}$-inverse-strongly monotone [1,37,38].

2.2 Fixed Point and Nonexpansivity

A *fixed point* of a mapping $T\colon H \to H$ is a point $x \in H$ satisfying $T(x) = x$. The set $\mathrm{Fix}(T) := \{x \in H\colon T(x) = x\}$ is called the *fixed point set* of T. A mapping $T\colon H \to H$ is said to be *nonexpansive* [2,3,12,13,26,31,32] if, for all $x, y \in H$, $\|T(x)-T(y)\| \leq \|x-y\|$. It is well known that the fixed point set of a nonexpansive mapping is closed and convex [2,13,32]. Given a nonempty, closed convex subset C of H, the mapping that assigns every point in H to its unique nearest point in C is called the *metric projection* onto C; and denoted by P_C, that is, $P_C(x) \in C$ and $\|x - P_C(x)\| = \inf_{y\in C}\|x - y\|$. The metric projection P_C is a typical nonexpansive mapping satisfying $\mathrm{Fix}(P_C) = C$. Some closed convex set C is simple in the sense that the closed form expression of P_C is known, which implies that P_C can be computed within a finite number of arithmetic operations. This will be the case, for example, when C is a linear variety, a closed ball, a closed cone, or a closed polytope [2,7,33]. Let $Q \in \mathbb{R}^{N\times N}$ be a positive semidefinite matrix with $\lambda_{\max} > 0$, $\lambda \in (0, 2/\lambda_{\max}]$, and $b \in \mathbb{R}^N$. We define a function $f\colon \mathbb{R}^N \to \mathbb{R}$ and a mapping $T\colon \mathbb{R}^N \to \mathbb{R}^N$ by $f(x) := (1/2)\langle x, Q(x)\rangle + \langle b, x\rangle$ for all $x \in \mathbb{R}^N$ and $T(x) := P_C(x - \lambda\nabla f(x))$ for all $x \in \mathbb{R}^N$, respectively. By the inverse-strong monotonicity of ∇f, we can prove that T is nonexpansive and $\mathrm{Fix}(T) = \mathrm{argmin}_{x\in C} f(x)$ [20].

2.3 Variational Inequality Problem

Problem (1.1) can be formulated equivalently as the *variational inequality problem* [9–11, 21,23,30,32,34,36] over $\mathrm{Fix}(T)$: find a point $z \in \mathrm{Fix}(T)$ such that $\langle v - z, \nabla f(z)\rangle \geq 0$ for all

$v \in \mathrm{Fix}(T)$. Suppose that $\nabla f \colon H \to H$ is monotone and hemicontinuous. Then, the problem is equivalent to the following problem: find a point $z \in \mathrm{Fix}(T)$ such that $\langle v - z, \nabla f(v) \rangle \geq 0$ for all $v \in \mathrm{Fix}(T)$. This implies that the set of solutions of the variational inequality problem is closed and convex.

3 Ergodic Iteration Method for a Convex Optimization Problem over the Fixed Point Set of a Nonexpansive Mapping

In this section, we assume that

(A1) $f \colon H \to \mathbb{R}$ is a convex, Fréchet differentiable function;

(A2) $\nabla f \colon H \to H$ is hemicontinuous;

(A3) $T \colon H \to H$ is a nonexpansive mapping with $\mathrm{Fix}(T) \neq \emptyset$;

(A4) $\mathrm{argmin}_{x \in \mathrm{Fix}(T)} f(x) \neq \emptyset$.

On some examples of f and T satisfying Conditions (A1)–(A4), see Section 4.

We present the following algorithm for Problem (1.1):

Algorithm 3.1 (Ergodic algorithm for convex optimization problem).

Step 0. Choose $x_1 \in H$ and $\lambda_1 \in (0, \infty)$ arbitrarily, and let $n := 1$.

Step 1. Given $x_n \in H$, choose $\lambda_n \in (0, \infty)$ (see Theorem 3.2) and compute $x_{n+1} \in H$ and $z_n^{(k)} \in H$ as

$$x_{n+1} := T(x_n - \lambda_n \nabla f(x_n))$$

and

$$z_n^{(k)} := \frac{\sum_{i=k}^{n} \lambda_i x_i}{\sum_{i=k}^{n} \lambda_i} \quad (k = 1, 2, \ldots, n). \tag{3.1}$$

Update $n := n + 1$ and go to Step 1.

For Algorithm 3.1, we present the following convergence analysis:

Theorem 3.2. *Assume that the sequence,* $(\nabla f(x_n))_{n \in \mathbb{N}}$, *in* Algorithm 3.1 *is bounded, and that there exists* $n_0 \in \mathbb{N}$ *such that* $\arg\min_{x \in \mathrm{Fix}(T)} f(x) \subset \Omega := \bigcap_{n=n_0}^{\infty} \{x \in \mathrm{Fix}(T) \colon f(x) \leq f(x_n)\}$ (For details, see Remark 3.3 and Section 4). *If we use* $(\lambda_n)_{n \in \mathbb{N}} \subset (0, \infty)$ *with* (i) $\lambda_{n+1} \leq \lambda_n$ $(n \in \mathbb{N})$, (ii) $\sum_{n=1}^{\infty} \lambda_n = \infty$, *and* (iii) $\sum_{n=1}^{\infty} \lambda_n^2 < \infty$ (An example of $(\lambda_n)_{n \in \mathbb{N}}$ is $\lambda_n := 1/n$), *then* $(x_n)_{n \in \mathbb{N}}$ *and* $(z_n^{(n_0)})_{n \geq n_0}$ *generated by* Algorithm 3.1 *satisfy the following:*

(a) [Boundedness] $(x_n)_{n \in \mathbb{N}}$ *and* $(z_n^{(n_0)})_{n \geq n_0}$ *are bounded.*

(b) [Relation between (x_{n+1}) and $(T(x_n))$] $\lim_{n \to \infty} \|x_{n+1} - T(x_n)\| = 0$.

(c) [Convergence of (z_n)] *The sequence* $(z_n^{(n_0)})_{n \geq n_0}$ *converges weakly to a solution of* (1.1).

Proof. (a) Fix $u \in \Omega$ arbitrarily. By the definition of ∂f, we have $\langle x_n - u, \nabla f(x_n)\rangle \geq f(x_n) - f(u) \geq 0$ for all $n \geq n_0$. From Condition (A3), we also have that, for all $n \geq n_0$,

$$\begin{aligned} \|x_{n+1} - u\|^2 &= \|T(x_n - \lambda_n \nabla f(x_n)) - T(u)\|^2 \\ &\leq \|(x_n - u) - \lambda_n \nabla f(x_n)\|^2 \\ &= \|x_n - u\|^2 - 2\lambda_n \langle x_n - u, \nabla f(x_n)\rangle + \lambda_n^2 \|\nabla f(x_n)\|^2 \\ &\leq \|x_n - u\|^2 + M^2 \lambda_n^2, \end{aligned} \tag{3.2}$$

where $M := \sup\{\|\nabla f(x_n)\| : n \in \mathbb{N}\} < \infty$. By $\sum_{n=1}^{\infty} \lambda_n^2 < \infty$ and (3.2), there exists $\lim_{n\to\infty} \|x_n - u\|$, and hence, $(x_n)_{n\in\mathbb{N}}$ is bounded. By (3.1), the sequence $(z_n^{(n_0)})_{n\geq n_0}$ is also bounded.

(b) By Condition (A3) and the definition of x_n, we get that, for every $n \in \mathbb{N}$,

$$\begin{aligned} \|x_{n+1} - T(x_n)\| &= \|T(x_n - \lambda_n \nabla f(x_n)) - T(x_n)\| \\ &\leq \lambda_n \|\nabla f(x_n)\| \leq M\lambda_n. \end{aligned}$$

Since Condition (iii) implies $\lim_{n\to\infty} \lambda_n = 0$, we deduce

$$\lim_{n\to\infty} \|x_{n+1} - T(x_n)\| = 0. \tag{3.3}$$

(c) Put $z_n := z_n^{(n_0)}$ for every $n \geq n_0$. By the boundedness of $(z_n)_{n\geq n_0}$, there exist a subsequence $(z_{n_i})_{i\in\mathbb{N}}$ of $(z_n)_{n\geq n_0}$ and $z \in H$ such that, for all $w \in H$, $\lim_{i\to\infty}\langle z_{n_i} - z, w\rangle = 0$. The proof is divided into the following four steps:

(I) Proof of $z \in \mathrm{Fix}(T)$.

By Condition (A3), we have that, for every $y \in H$ and for every $n \in \mathbb{N}$, $\|x_{n+1} - T(y)\| \leq \|x_{n+1} - T(x_n)\| + \|x_n - y\|$. So, it holds from the boundedness of $(x_n)_{n\in\mathbb{N}}$ and Condition (i) that, for all $n \geq n_0$,

$$\begin{aligned} \lambda_{n+1}\|x_{n+1} - T(y)\|^2 &\leq \lambda_{n+1}(\|x_n - y\| + \|x_{n+1} - T(x_n)\|)^2 \\ &\leq \lambda_n \|x_n - y\|^2 + (2\|x_n - y\| + \|x_{n+1} - T(x_n)\|)\lambda_n \|x_{n+1} - T(x_n)\| \\ &\leq \lambda_n \|x_n - y\|^2 + K\lambda_n \|x_{n+1} - T(x_n)\|, \end{aligned}$$

where $K := \sup\{2\|x_n - y\| + \|x_{n+1} - T(x_n)\| \colon n \in \mathbb{N}\} < \infty$. Hence, we obtain that, for all $n \geq n_0$,

$$\begin{aligned} 0 &\leq \lambda_n \|x_n - y\|^2 - \lambda_{n+1}\|x_{n+1} - T(y)\|^2 + K\lambda_n \|x_{n+1} - T(x_n)\| \\ &= \lambda_n \|x_n - T(y)\|^2 + 2\lambda_n \langle x_n - T(y), T(y) - y\rangle + \lambda_n \|T(y) - y\|^2 \\ &\quad - \lambda_{n+1}\|x_{n+1} - T(y)\|^2 + K\lambda_n \|x_{n+1} - T(x_n)\|, \end{aligned}$$

which implies

$$\begin{aligned} 0 \leq \lambda_{n_0}\|x_{n_0} - T(y)\|^2 - \lambda_{m+1}\|x_{m+1} - T(y)\|^2 + \|T(y) - y\|^2 \sum_{k=n_0}^{m} \lambda_k \\ + K \sum_{k=n_0}^{m} \lambda_k \|x_{k+1} - T(x_k)\| + 2\left\langle \sum_{k=n_0}^{m} \lambda_k x_k - \sum_{k=n_0}^{m} \lambda_k T(y), T(y) - y \right\rangle \end{aligned}$$

for all $m \geq n_0$. Therefore, we get

$$0 \leq \frac{\lambda_{n_0}\|x_{n_0} - T(y)\|^2}{\sum_{k=n_0}^{m} \lambda_k} + 2\langle z_m - T(y), T(y) - y\rangle + \|T(y) - y\|^2 + K\frac{\sum_{k=n_0}^{m} \lambda_k \|x_{k+1} - T(x_k)\|}{\sum_{k=n_0}^{m} \lambda_k}.$$

Taking $m := n_i$, from Condition (ii) and (3.3), we have $0 \leq 2\langle z - T(y), T(y) - y\rangle + \|T(y) - y\|^2$ as $i \to \infty$. Putting $y := z$, we get $0 \leq 2\langle z - T(z), T(z) - z\rangle + \|T(z) - z\|^2 = -\|T(z) - z\|^2$. Thus, we obtain $z \in \mathrm{Fix}(T)$.

(II) Proof of $z \in \mathrm{argmin}_{x \in \mathrm{Fix}(T)} f(x)$.

Let $v \in \mathrm{Fix}(T)$. By the monotonicity of ∇f, we get that, for all $n \geq n_0$,

$$\begin{aligned} \|x_{n+1} - v\|^2 &\leq \|(x_n - v) - \lambda_n \nabla f(x_n)\|^2 \\ &= \|x_n - v\|^2 - 2\lambda_n\langle x_n - v, \nabla f(x_n)\rangle + M^2\lambda_n^2 \\ &\leq \|x_n - v\|^2 - 2\lambda_n\langle x_n - v, \nabla f(v)\rangle + M^2\lambda_n^2. \end{aligned}$$

So, we deduce that, for all $m \geq n_0$,

$$\|x_{m+1} - v\|^2 - \|x_{n_0} - v\|^2 \leq -2\left\langle \sum_{k=n_0}^{m} \lambda_k x_k - \sum_{k=n_0}^{m} \lambda_k v, \nabla f(v) \right\rangle + M^2 \sum_{k=n_0}^{m} \lambda_k^2,$$

which implies that

$$-\frac{\|x_{n_0} - v\|^2}{\sum_{k=n_0}^{m} \lambda_k} \leq -2\langle z_m - v, \nabla f(v)\rangle + M^2 \frac{\sum_{k=n_0}^{m} \lambda_k^2}{\sum_{k=n_0}^{m} \lambda_k}.$$

Taking $m := n_i$, from Conditions (ii) and (iii), we obtain $0 \leq -2\langle z - v, \nabla f(v)\rangle$ as $i \to \infty$, and hence, $\langle v - z, \nabla f(v)\rangle \geq 0$ for all $v \in \mathrm{Fix}(T)$. Condition (A2) ensures that $\langle v - z, \nabla f(z)\rangle \geq 0$ for all $v \in \mathrm{Fix}(T)$, that is, $z \in \mathrm{argmin}_{x \in \mathrm{Fix}(T)} f(x)$.

By Condition (A1) and the closedness and convexity of $\mathrm{Fix}(T) \subset H$, it is shown that $\Omega \subset H$ is closed and convex. So, we can define $u_n := P_\Omega(x_n)$ for every $n \in \mathbb{N}$.

(III) Proof of convergence of $(u_n)_{n \in \mathbb{N}}$

Fix $w \in \Omega$ arbitrarily. It holds from (3.2) that, for every $n, m \geq n_0$,

$$\begin{aligned} \|x_{n+m} - w\|^2 &\leq \|x_{n+m-1} - w\|^2 + M^2\lambda_{n+m-1}^2 \\ &\leq \|x_{n+m-2} - w\|^2 + M^2(\lambda_{n+m-1}^2 + \lambda_{n+m-2}^2) \\ &\leq \|x_n - w\|^2 + M^2 \sum_{i=n}^{n+m-1} \lambda_i^2 \leq \|x_n - w\|^2 + M^2 \sum_{i=n}^{\infty} \lambda_i^2. \end{aligned}$$

So, we have

$$\|x_{n+m} - u_n\|^2 \leq \|x_n - u_n\|^2 + M^2 \sum_{i=n}^{\infty} \lambda_i^2 \tag{3.4}$$

for every $n, m \geq n_0$. By $u_{n+m} = P_\Omega(x_{n+m})$ and the convexity of Ω, we also have

$$\left\| x_{n+m} - \frac{u_n + u_{n+m}}{2} \right\| \geq \|x_{n+m} - u_{n+m}\|. \tag{3.5}$$

From (3.4) and (3.5), we get

$$\begin{aligned}\|u_{n+m}-u_n\|^2 &= \|(u_{n+m}-x_{n+m})+(x_{n+m}-u_n)\|^2 \\ &= 2\|u_{n+m}-x_{n+m}\|^2+2\|x_{n+m}-u_n\|^2-4\Big\|x_{n+m}-\frac{u_n+u_{n+m}}{2}\Big\|^2 \\ &\leq 2\|x_{n+m}-u_n\|^2-2\|u_{n+m}-x_{n+m}\|^2 \\ &\leq 2\|x_n-u_n\|^2-2\|u_{n+m}-x_{n+m}\|^2+2M^2\sum_{i=n}^{\infty}\lambda_i^2,\end{aligned} \tag{3.6}$$

and hence,

$$\limsup_{m\to\infty}\|x_m-u_m\|^2 \leq \|x_n-u_n\|^2+M^2\sum_{i=n}^{\infty}\lambda_i^2.$$

So, by Condition (iii), there exists $\lim_{n\to\infty}\|x_n-u_n\|$. Thus, by (3.6) and Condition (iii), $(u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since $\Omega\subset H$ is closed, $(u_n)_{n\in\mathbb{N}}$ converges strongly to a point $\hat{z}\in\Omega$.

(IV) Proof of $z=\hat{z}$

By $u_n=P_\Omega(x_n)$ and $z\in\operatorname{argmin}_{x\in\mathrm{Fix}(T)} f(x)\subset\Omega$, we have $\langle z-u_n, u_n-x_n\rangle\geq 0$ for all $n\geq n_0$. Then, we get

$$\begin{aligned}\langle z-\hat{z}, x_n-u_n\rangle &= \langle z-u_n, x_n-u_n\rangle+\langle u_n-\hat{z}, x_n-u_n\rangle \\ &\leq \|u_n-\hat{z}\|\|x_n-u_n\|\leq L\|u_n-\hat{z}\|,\end{aligned}$$

where $L:=\sup\{\|x_n-u_n\|: n\in\mathbb{N}\}<\infty$. Hence, we deduce that, for every $m\geq n_0$,

$$\left\langle z-\hat{z}, z_m-\frac{\sum_{k=n_0}^m\lambda_k u_k}{\sum_{k=n_0}^m\lambda_k}\right\rangle \leq L\frac{\sum_{k=n_0}^m\lambda_k\|u_k-\hat{z}\|}{\sum_{k=n_0}^m\lambda_k}.$$

Taking $m:=n_i$, by Condition (ii) and $\lim_{n\to\infty}\|u_n-\hat{z}\|=0$, we obtain $\langle z-\hat{z}, z-\hat{z}\rangle\leq 0$ as $i\to\infty$, that is, $z=\hat{z}$. This implies that the sequence (z_n) converges weakly to the point $\hat{z}=z\in\operatorname{argmin}_{x\in\mathrm{Fix}(T)} f(x)$. □

Remark 3.3. Assume that there exists $n_0\in\mathbb{N}$ such that $x_n\in\mathrm{Fix}(T)$ for all $n\geq n_0$. Then, the convergence condition, $\operatorname{argmin}_{x\in\mathrm{Fix}(T)} f(x)\subset\Omega$, holds. On numerical examples for the relation $x_n\in\mathrm{Fix}(T)$, see Section 4.

4 Numerical Examples

In this section, we present numerical examples for an algorithm presented in the previous section. We consider the following quadratic optimization problem:

Problem 4.1 (Quadratic optimization problem over complex constrained set).

$$\text{minimize } f(x):=\frac{1}{2}\langle x, Q(x)\rangle \text{ subject to } x\in C,$$

where $Q\in\mathbb{R}^{64\times 64}$ is a positive semidefinite matrix generated in [14] and $C\subset\mathbb{R}^{64}$ is a nonempty, closed convex set such that no closed form expression of P_C is known, or its computation is not easy.

By the definitions of f and Q, a function f and its gradient $\nabla f := Q$ satisfy Conditions (A1) and (A2).

We first discuss the case where C is the intersection of two closed balls. Let $C_1 := \{x \in \mathbb{R}^{64} \colon \|x\|^2 \leq 1\}$ and $C_2 := \{x \in \mathbb{R}^{64} \colon \|x - (1,1,0,\ldots,0)^T\|^2 \leq 1\}$, and define

$$C := C_1 \cap C_2 \neq \emptyset.$$

In such a problem, the exact solutions of Problem 4.1 cannot be described. We note that the computations of P_{C_1} and P_{C_2} are easy, but the computation of P_C is not easy. To relax this complexity, we use a computable mapping, $T \colon \mathbb{R}^{64} \to \mathbb{R}^{64}$, defined by

$$T(x) := \frac{1}{2} P_{C_1}(x) + \frac{1}{2} P_{C_2}(x) \text{ for all } x \in \mathbb{R}^{64}.$$

Such a mapping T satisfies the nonexpansivity and $\mathrm{Fix}(T) = C \neq \emptyset$, that is, Condition (A3) is satisfied. Moreover, by the continuity of f and the compactness of C, Condition (A4) holds. In this case, we used $x_1 := (-0.5, -0.5, \ldots, -0.5)^T \in \mathbb{R}^{64}$ and $\lambda_n := 1/(10^4(n+1))$. FIGURE 1 shows the behaviors for $(x_n)_{n=1}^{100}$ and $(z_n^{(100)})_{n \geq 100}$. It is seen from FIGURE 1 that each point x_n ($n \geq 100$) is in $\mathrm{Fix}(T) = C$. So, the convergence condition, $\mathrm{argmin}_{x \in C} f(x) \subset \Omega$, is satisfied. At the same time, it can be observed from FIGURE 1 that the behavior of $f(z_n^{(100)})$ ($n \geq 100$) is stable.

Next we consider the case where C is the solution set of a convex optimization problem over a simple constrained set. Let $D := \{x \in \mathbb{R}^{64} \colon \|x\|^2 \leq 1\}$, $R \in \mathbb{R}^{64 \times 64}$ the diagonal matrix that has eigenvalues $0, 1, \ldots, 63$, $b \in \mathbb{R}^{64}$, and

$$C := \left\{ \hat{x} \in D \colon \frac{1}{2}\langle \hat{x}, R(\hat{x}) \rangle + \langle b, \hat{x} \rangle = \min_{x \in D} \left[\frac{1}{2} \langle x, R(x) \rangle + \langle b, x \rangle \right] \right\} \neq \emptyset.$$

Problem 4.1 with this constrained set C is a *three-stage convex optimization problem* and the exact solutions of this problem are more than one, and these solutions cannot be described. As the closed form expression of P_C is not known, we cannot compute P_C. So, we define a mapping, $T \colon \mathbb{R}^{64} \to \mathbb{R}^{64}$, by

$$T(x) := P_D(x - \alpha(R(x) + b)) \text{ for all } x \in \mathbb{R}^{64},$$

where $\alpha \in (0, 2/63]$. Then, T is a nonexpansive mapping with $\mathrm{Fix}(T) = C \neq \emptyset$ (see Sec.2), and hence, Condition (A3) holds. By $C \subset D$ and the compactness of $D \subset \mathbb{R}^{64}$, Condition (A4) is also satisfied. In this problem, we used $x_1 := (0, 0, \ldots, 0)^T \in \mathbb{R}^{64}$, $\alpha := 2/63$, and $\lambda_n := 1/(10^5(n+1))$. The behaviors for $(x_n)_{n=1}^{2000}$ and $(z_n^{(2000)})_{n \geq 2000}$ when $b := (-0.1, -0.1, \ldots, -0.1)^T \in \mathbb{R}^{64}$ are presented in FIGURE 2. From FIGURE 2, we note that x_n ($n \geq 2000$) is in $\mathrm{Fix}(T) = C$, and hence, the convergence condition, $\mathrm{argmin}_{x \in C} f(x) \subset \Omega$, is satisfied. FIGURE 2 shows that the behaviors of $f(x_n)_{n=500}^{2000}$ and $f(z_n^{(2000)})$ ($n \geq 2000$) are stable, and that its values are the same. Therefore, it is considered that the proposed algorithm converges to some solution of Problem 4.1. The behaviors for $(x_n)_{n=1}^{100}$ and $(z_n^{(100)})_{n \geq 100}$ when $b := (1, 1, \ldots, 1)^T \in \mathbb{R}^{64}$ are presented in FIGURE 3. FIGURE 3 shows that the behavior of $f(x_n)$ ($n = 1, 2, \ldots, 100$) is unstable, but, by computing the mean of x_n, the behavior of $f(z_n^{(100)})$ ($n \geq 100$) is stable. Hence, it is considered that the proposed algorithm converges to some solution of Problem 4.1.

Let $C_1 := \{x \in \mathbb{R}^{64} \colon \|x\|^2 \leq 1\}$, $D_1 := \{x := (x_1, x_2, \ldots, x_{64})^T \in \mathbb{R}^{64} \colon x_1 \geq a\}$, and $D_2 := \{x := (x_1, x_2, \ldots, x_{64})^T \in \mathbb{R}^{64} \colon x_2 \geq a\}$, where $a \geq 0$. Finally we consider the case where $C := C_1 \cap (D_1 \cap D_2)$. When $a \geq 1$, the condition $C = \emptyset$ holds. To resolve this situation, we define the following *generalized convex feasible set* [34]:

$$C_\Phi := \left\{ \hat{x} \in C_1 \colon \Phi(\hat{x}) = \min_{x \in C_1} \Phi(x) \right\} \neq \emptyset,$$

where $d(x, D_i) := \inf\{\|x - y\| \colon y \in D_i\}$ $(i = 1, 2, x \in \mathbb{R}^{64})$ and $\Phi(x) := (1/2)[(1/2)d(x, D_1)^2 + (1/2)d(x, D_2)^2]$ $(x \in \mathbb{R}^{64})$. If $C \neq \emptyset$, that is, $a < 1$, then $C_\Phi = C$ holds. Even if $C = \emptyset$, the set C_Φ is well defined as the set of all minimizers of Φ over C_1. It follows from the compactness of C_1 that $C_\Phi \neq \emptyset$ is satisfied; for more details, see [34]. Problem 4.1 with C_Φ is also a three-stage convex optimization problem and no closed form expression of P_Φ is known. So, we define a mapping, $T \colon \mathbb{R}^{64} \to \mathbb{R}^{64}$, as follows:

$$T(x) := P_{C_1}\left[\frac{1}{2} P_{D_1}(x) + \frac{1}{2} P_{D_2}(x)\right] \text{ for all } x \in \mathbb{R}^{64}.$$

Then, T is nonexpansive and $\mathrm{Fix}(T) = C_\Phi \neq \emptyset$ [34]. We used $x_1 := (0.5, 0.5, \ldots, 0.5)^T \in \mathbb{R}^{64}$ and $\lambda_n := 1/(10^4(n+1))$. Figure 4 shows the behaviors for $(x_n)_{n=1}^{100}$ and $(z_n^{(100)})_{n \geq 100}$ when $a := 1/2$. From Figure 4, we note that $x_n \in \mathrm{Fix}(T) = C_1 \cap (D_1 \cap D_2)$ $(n \geq 100)$ and that the behavior of $f(z_n^{(100)})$ $(n \geq 600)$ is stable. The behaviors for $(x_n)_{n=1}^{100}$ and $(z_n^{(100)})_{n \geq 100}$ when $a := 1$ are presented in Figure 5. It can be observed from Figure 5 that $x_n \in \mathrm{Fix}(T)$ $(n \geq 100)$. Moreover, it is seen from Figure 5 that the behavior of $f(z_n^{(100)})$ $(n \geq 100)$ is stable by using the ergodic method. Therefore, it is considered that the proposed method converges to some solution of Problem 4.1.

Remark 4.2. By Theorem 3.2, the proposed algorithm requires us to choose $n_0 \in \mathbb{N}$ such that $\arg\min_{x \in \mathrm{Fix}(T)} f(x) \subset \Omega := \bigcap_{n=n_0}^{\infty} \{x \in \mathrm{Fix}(T) \colon f(x) \leq f(x_n)\}$. A choice of the number $n_0 \in \mathbb{N}$ depends on a nonexpansive mapping T, the objective function f, and an initial point x_1.

Acknowledgments

The author thanks Professor I. Yamada for his discussions on the algorithms.

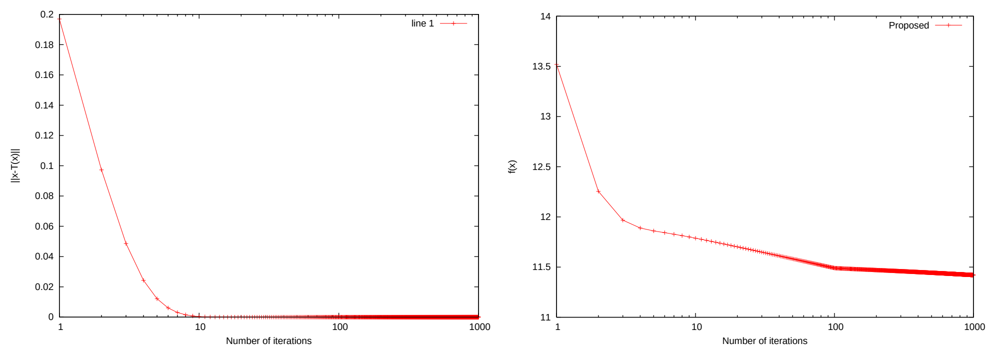

Figure 1: Behaviors of $\|z_n^{(100)} - T(z_n^{(100)})\|$ and $f(z_n^{(100)})$ when a positive semidefinite matrix Q which is given in [14] and the constrained set $C := C_1 \cap C_2$.

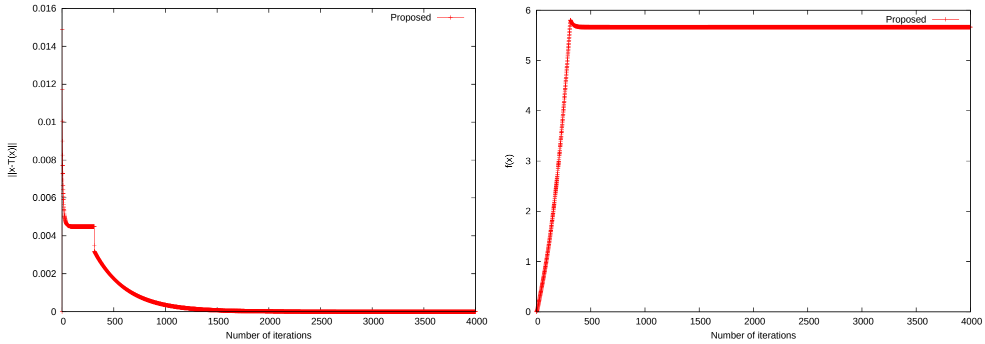

Figure 2: Behaviors of $\|z_n^{(2000)} - T(z_n^{(2000)})\|$ and $f(z_n^{(2000)})$ when a positive semidefinite matrix Q which is given in [14], $b := (-0.1, -0.1, \ldots, -0.1)^T \in \mathbb{R}^{64}$, and the constrained set $C := \mathrm{argmin}_{x \in D}[(1/2)\langle x, R(x)\rangle + \langle b, x\rangle]$.

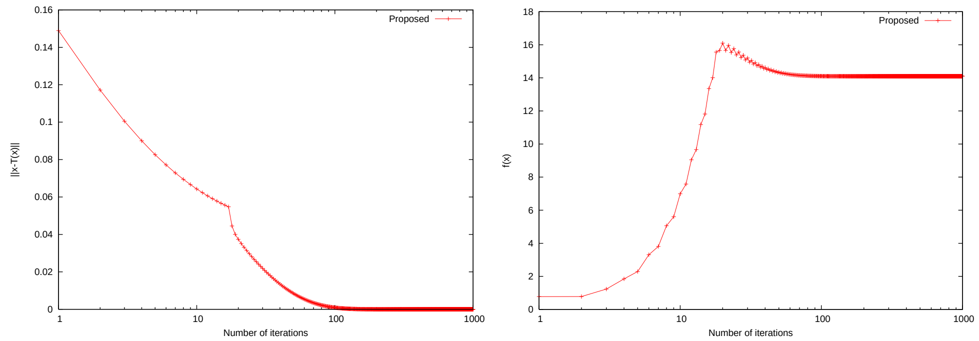

Figure 3: Behaviors of $\|z_n^{(100)} - T(z_n^{(100)})\|$ and $f(z_n^{(100)})$ when a positive semidefinite matrix Q which is given in [14], $b := (1, 1, \ldots, 1)^T \in \mathbb{R}^{64}$, and the constrained set $C := \mathrm{argmin}_{x \in D}[(1/2)\langle x, R(x)\rangle + \langle b, x\rangle]$.

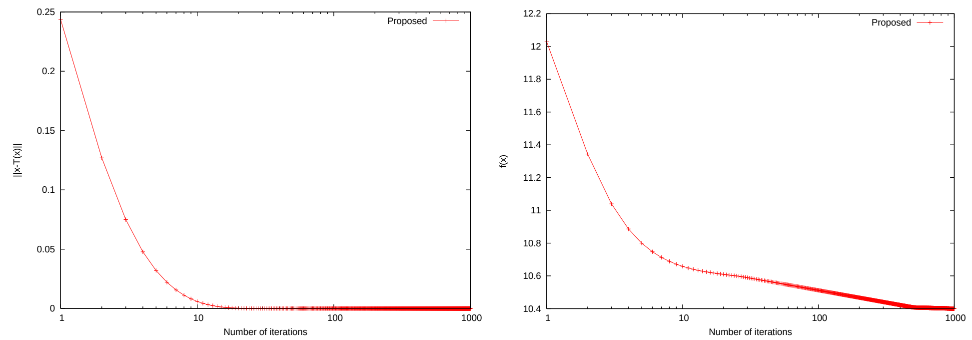

Figure 4: Behaviors of $\|z_n^{(100)} - T(z_n^{(100)})\|$ and $f(z_n^{(100)})$ when a positive semidefinite matrix Q which is given in [14], $a := 1/2$, and the constrained set $C_\Phi = C_1 \cap (D_1 \cap D_2)$.

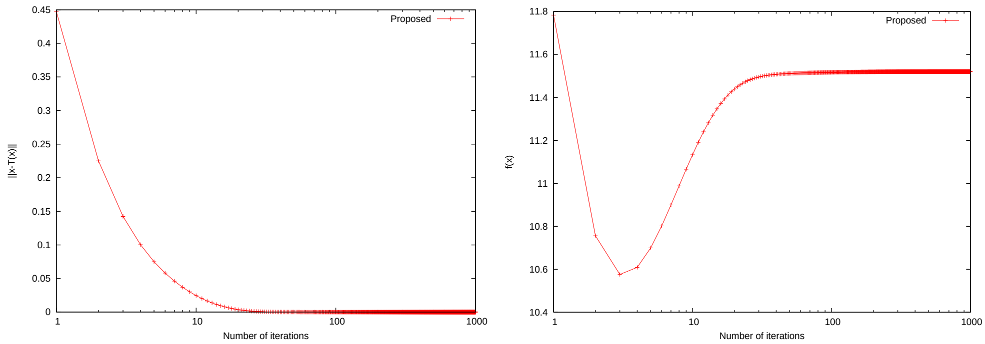

Figure 5: Behaviors of $\|z_n^{(100)} - T(z_n^{(100)})\|$ and $f(z_n^{(100)})$ when a positive semidefinite matrix Q which is given in [14], $a := 1$, and the constrained set C_Φ $(C_1 \cap (D_1 \cap D_2) = \emptyset)$.

References

[1] J.B. Baillon and G. Haddad, Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones, *Israel J. Math.* 26 (1977) 137–150.

[2] H.H. Bauschke and J.M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review* 38 (1996) 367–426.

[3] H.H. Bauschke, J.M. Borwein and A.S. Lewis, The method of cyclic projections for closed convex sets in Hilbert space, *Contemp. Math.* 204 (1997) 1–38.

[4] 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.

[5] R.E. Bruck, Jr., On the weak convergence of an ergodic iteration for the solution of variational inequalities for monotone operators in Hilbert space, *J. Math. Anal. Appl.* 61 (1977) 159–164.

[6] P.L. Combettes, A block-iterative surrogate constraint splitting method for quadratic signal recovery, *IEEE Trans. Signal Process.* 51 (2003) 1771–1782.

[7] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer, New York, NY, 2001.

[8] J.C. Dunn, Convexity, monotonicity, and gradient processes in Hilbert space, *J. Math. Anal. Appl.* 53 (1976) 145–158.

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

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

[11] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer, New York, NY, 1984.

[12] K. Goebel and W.A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge Univ. Press, New York, 1990.

[13] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Dekker, New York and Basel, 1984.

[14] P.C. Hansen, *Regularization Tools Version 4.1 (for Matlab Version 7.3)*, Available: http://www2.imm.dtu.dk/~pch/Regutools/

[15] S.A. Hirstoaga, Iterative selection methods for common fixed point problems, *J. Math. Anal. Appl.* 324 (2006) 1020–1035.

[16] H. Iiduka, Hybrid conjugate gradient method for a convex optimization problem over the fixed-point set of a nonexpansive mapping, *J. Optim. Theory Appl.* 140 (2009) 463–475.

[17] H. Iiduka and I. Yamada, A subgradient type method for the equilibrium problem over the fixed point set and its applications, *Optimization* 58 (2009) 251–261.

[18] H. Iiduka and I. Yamada, A use of conjugate gradient direction for the convex optimization problem over the fixed point set of a nonexpansive mapping, *SIAM J. Optim.* 19 (2009) 1881–1893.

[19] H. Iiduka and I. Yamada, An ergodic algorithm for the power control games for CDMA data networks, *J. Math. Model. Algorithms* 8 (2009) 1–18.

[20] H. Iiduka, W. Takahashi and M. Toyoda, Approximation of solutions of variational inequalities for monotone mappings, *PanAmer. Math. J.* 14 (2004) 49–61.

[21] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and their Applications*, Academic Press, New York, 1980.

[22] F. Liu and M. Z. Nashed, Regularization of nonlinear ill-posed variational inequalities and convergence rates, *Set-Valued Anal.* 6 (1998) 313–344.

[23] Z.Q. Luo, J.S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, New York, 1996.

[24] P.E. Maingé and A. Moudafi, Strong convergence of an iterative method for hierarchical fixed-point problems, *Pacific J. Optim.* 3 (2007) 529–538.

[25] A. Moudafi, Krasnoselski-Mann iteration for hierarchical fixed-point problems, *Inverse Problems* 23 (2007) 1635–1640.

[26] S. Reich, Some problems and results in fixed point theory, *Contemp. Math.* 21 (1983) 179–187.

[27] K. Slavakis, I. Yamada and K. Sakaniwa, Computation of symmetric positive definite Toeplitz matrices by the hybrid steepest descent method, *Signal Processing* 83 (2003) 1135–1140.

[28] K. Slavakis and I. Yamada, *Robust wideband beamforming by the hybrid steepest descent method*, IEEE Trans. Signal Process. 55 (2007) 4511–4522.

[29] M.V. Solodov and B.F. Svaiter, A new projection method for variational inequality problems, *SIAM J. Control Optim.* 37 (1999) 765–776.

[30] G. Stampacchia, Formes bilinéaires coercitives sur les ensembles convexes, *C. R. Acad. Sci. Paris* 258 (1964) 4413–4416.

[31] H. Stark and Y. Yang, *Vector Space Projections: A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics*, John Wiley & Sons Inc, New York, 1998.

[32] W. Takahashi, *Nonlinear Functional Analysis*, Yokohama Publishers, Yokohama, 2000.

[33] P. Wolfe, Finding the nearest point in a polytope, *Math. Program.* 11 (1976) 128–149.

[34] I. Yamada, The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings, in *Inherently Parallel Algorithms for Feasibility and Optimization and Their Applications*, D. Butnariu, Y. Censor and S. Reich (eds.), Elsevier, New York, 2001, pp. 473–504.

[35] I. Yamada, N. Ogura and N. Shirakawa, A numerically robust hybrid steepest descent method for the convexly constrained generalized inverse problems, *Contemp. Math.* 313 (2002) 269–305.

[36] E. Zeidler, *Nonlinear Functional Analysis ans its Applications III Variational Methods and Optimization*, Springer, New York, NY, 1985.

[37] D. Zhu and P. Marcotte, New classes of generalized monotonicity, *J. Optim. Theory Appl.* 87 (1995) 457–41.

[38] D.L. Zhu and P. Marcotte, Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities, *SIAM J. Optim.* 6 (1996) 714–726.

Manuscript received 8 December 2008
accepted for publication 29 May 2009

Hideaki Iiduka
Network Design Research Center, Kyushu Institute of Technology
Hibiya Kokusai Bldg. 1F 107, 2-2-3 Uchisaiwai-cho, Chiyoda-ku
Tokyo 100-0011, Japan
E-mail address: iiduka@ndrc.kyutech.ac.jp