Journal of Nonlinear and Convex Analysis
Volume 6, Number 3, 2005, 505–523

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

PROXIMAL POINT ALGORITHMS WITH BREGMAN FUNCTIONS IN BANACH SPACES

FUMIAKI KOHSAKA AND WATARU TAKAHASHI

Abstract. The proximal point algorithm is a well-known method for approximating a zero point of a given maximal monotone operator in Hilbert spaces. In this paper, we propose two proximal-type algorithms with Bregman functions in Banach spaces and, then, discuss the weak and strong convergence of the proposed methods. Our results generalize the recently obtained weak and strong convergence theorems due to Kohsaka–Takahashi and Kamimura–Kohsaka–Takahashi. Applications to a convex minimization problem and a variational inequality problem are also included.

1. Introduction

A well-known method for approximating a zero point of a maximal monotone operator defined in a Hilbert space is the *proximal point algorithm* first introduced by Martinet [25] and generally studied by Rockafellar [37]. This is an iterative procedure, which generates $\{x_n\}$ by $x_1 = x \in E$ and

$$x_{n+1} = J_{r_n} x_n \quad (n = 1, 2, \dots), \tag{1}$$

where $\{r_n\} \subset (0, \infty)$, $T \subset E \times E$ is a maximal monotone operator in a real Hilbert space E and J_r is the resolvent of T defined by $J_r = (I + rT)^{-1}$ for all $r > 0$. In 1976, Rockafellar [37] proved that if $T^{-1}0 \neq \emptyset$ and $\liminf_n r_n > 0$, then the sequence generated by this method converges *weakly* to an element of $T^{-1}0$. In particular, if T is the subdifferential ∂f of a proper lower semicontinuous convex function $f : E \to (-\infty, \infty]$, then (1) is reduced to

$$x_{n+1} = \arg\min_{y \in E} \left\{ f(y) + \frac{1}{2r_n} \|y - x_n\|^2 \right\} \quad (n = 1, 2, \dots). \tag{2}$$

In this case, $\{x_n\}$ converges weakly to a minimizer of f.

Later, many researchers have studied the convergence of the proximal point algorithms in Hilbert spaces; see, for instance, [6, 7, 15, 18, 23, 26, 27, 30, 39] and the references mentioned there. In particular, Motivated by the iterative methods for nonexpansive mappings of Halpern's type [16, 38, 44] and Mann's type [24, 32], Kamimura and Takahashi [18] introduced the following two proximal-type algorithms in Hilbert spaces:

$$x_{n+1} = \alpha_n x_1 + (1 - \alpha_n) J_{r_n} x_n \quad (n = 1, 2, \dots) \tag{3}$$

and

$$x_{n+1} = \alpha_n x_n + (1 - \alpha_n) J_{r_n} x_n \quad (n = 1, 2, \dots), \tag{4}$$

2000 *Mathematics Subject Classification.* Primary 47H05, 47J25.

Key words and phrases. Bregman function, convex minimization problem, maximal monotone operator, proximal point algorithm, uniformly convex function, uniformly smooth function.

where $\{\alpha_n\} \subset [0,1]$ and $\{r_n\} \subset (0,\infty)$. They showed that the sequence generated by (3) is *strongly* convergent to Px and the sequence generated by (4) is *weakly* convergent to the strong $\lim_n Px_n$, where P denotes the metric projection (the nearest point projection) from E onto $T^{-1}0$ defined by

$$P(x) = \arg \min_{y \in T^{-1}0} \|y - x\| \tag{5}$$

for all $x \in E$. Subsequently, these results were generalized to accretive operators in Banach spaces [19, 20]; see also Takahashi [40, 41].

Recently, Kamimura and Takahashi [21] generalized Solodov and Svaiter's strong convergence theorem [39] in a Hilbert space to that in a more general Banach space including L^p $(1 < p < \infty)$; see also Ohsawa and Takahashi [28] for another generalization of Solodov and Svaiter's theorem. More recently, motivated by Kamimura and Takahashi [21] and the convex combination based on Bregman distances due to Censor and Reich [12], Kohsaka and Takahashi [22] and Kamimura, Kohsaka and Takahashi [17] introduced the following proximal-type algorithms in a smooth and uniformly convex Banach space:

$$x_{n+1} = J^{-1}(\alpha_n J(x_1) + (1-\alpha_n)J(J_{r_n}x_n)) \quad (n = 1, 2, \dots) \tag{6}$$

and

$$x_{n+1} = J^{-1}(\alpha_n J(x_n) + (1-\alpha_n)J(J_{r_n}x_n)) \quad (n = 1, 2, \dots), \tag{7}$$

where J is the normalized duality mapping from E into E^* and $J_r = (J+rT)^{-1}J$ for all $r > 0$. If E is a Hilbert space, then $J = I$, where I is the identity operator on E, and hence (6) and (7) are reduced to (3) and (4), respectively. Then they generalized the previous results due to Rockafellar [37] and Kamimura and Takahashi [18] to maximal monotone operators in Banach spaces. They made use of the generalized projection P from E onto $T^{-1}0$ defined by

$$P(x) = \arg \min_{y \in T^{-1}0} \{\|y\|^2 - 2\langle y, Jx\rangle + \|x\|^2\} \tag{8}$$

for all $x \in E$. If E is a Hilbert space, then (8) is reduced to (5).

On the other hand, many researchers have studied the proximal point algorithm with *Bregman functions* [5]; see, for instance, [3, 4, 7, 8, 9, 10, 13, 29, 33] and the references mentioned there. In particular, Otero and Svaiter [29] dealt with a hybrid proximal-type method with Bregman functions in reflexive Banach spaces. Their result extends the previous results according to Solodov and Svaiter [39] and Kamimura and Takahashi [21]. They made use of the Bregman projection P from E onto $T^{-1}0$ is defined by

$$P(x) = \arg \min_{y \in T^{-1}0} \{g(y) - g(x) - \langle y - x, \nabla g(x)\rangle\} \tag{9}$$

for all $x \in E$, where g is a Bregman function. In the case where $g = \|\cdot\|^2$ in a smooth and uniformly convex Banach space, (9) is reduced to (8).

The purpose of the present paper is to generalize the weak and strong convergence theorems in [17, 22] with Bregman functions in reflexive Banach spaces. We propose the following proximal-type algorithms:

$$x_{n+1} = \nabla g^*(\alpha_n \nabla g(x_1) + (1-\alpha_n)\nabla g(J_{r_n}x_n)) \quad (n = 1, 2, \dots) \tag{10}$$

and

$$x_{n+1} = \nabla g^*(\alpha_n \nabla g(x_n) + (1-\alpha_n)\nabla g(J_{r_n} x_n)) \;\; (n = 1, 2, \dots), \tag{11}$$

where $\{\alpha_n\} \subset [0,1]$, $\{r_n\} \subset (0,\infty)$, $g : E \to \mathbb{R}$ is a Bregman function defined in a reflexive Banach space E and $J_r = (\nabla g + rT)^{-1}\nabla g$ for all $r > 0$. In the particular case where $g = \|\cdot\|^2/2$ in a smooth and uniformly convex Banach space, (10) and (11) are reduced to the (6) and (7), respectively.

Our paper is organized as follows. In Section 2, we state several definitions and known results. In Section 4, using some lemmas proved in Section 3, we prove that the sequence generated by (10) converges strongly to $P(x)$, where P is the Bregman projection from E onto $T^{-1}0$ (Theorem 4.1). In Section 5, we also prove that the sequence generated by (11) converges weakly to the strong $\lim_n P(x_n)$ (Theorem 5.2). The obtained results generalize the corresponding results due to Kohsaka and Takahashi [22] and Kamimura, Kohsaka and Takahashi [17]. In Section 6, we apply our results to a convex minimization problem and a variational inequality problem. In Section 7, we give another proof of a lemma used in the definition of the resolvent of a maximal monotone operator.

2. Preliminaries

Throughout this paper, we denote the set of real numbers and the set of positive integers by $\mathbb{R}$ and $\mathbb{N}$, respectively. Let E be a (real) Banach space with the dual space E^*. We denote the value of $x^* \in E^*$ at $x \in E$ by $\langle x, x^* \rangle$. We denote the strong convergence and the weak convergence of a sequence $\{x_n\}$ of E to $x \in E$ by $x_n \to x$ and $x_n \rightharpoonup x$, respectively. We also denote the weak* convergence of a sequence $\{x_n^*\}$ of E^* to $x^* \in E^*$ by $x_n^* \overset{*}{\rightharpoonup} x^*$. For $p \in (1,\infty)$, the *duality mapping* J_p from E into 2^{E^*} corresponding to the weight function $\omega(t) = t^{p-1}$ is defined by

$$J_p(x) = \{x^* \in E^* : \langle x, x^* \rangle = \|x\|\|x^*\|, \ \|x^*\| = \|x\|^{p-1}\} \tag{12}$$

for all $x \in E$. The mapping J_2 is called the *normalized duality mapping* from E into 2^{E^*} and it is denoted by J. If E is a Hilbert space, then J is coincident with the identity operator I on E. A Banach space E is said to be *strictly convex* if $\|(x+y)/2\| < 1$ whenever $x, y \in S$ and $x \neq y$, where $S = \{z \in E : \|z\| = 1\}$. The space E is also said to be *uniformly convex* if for all $\varepsilon \in (0, 2]$, there exists $\delta > 0$ such that $x, y \in S$ and $\|x - y\| \geq \varepsilon$ imply $\|(x+y)/2\| \leq 1 - \delta$. It is well-known that every uniformly convex Banach space is strictly convex and reflexive. A Banach space E is said to be *smooth* if

$$\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} \tag{13}$$

exists for all $x, y \in S$. The space E is also said to be *uniformly smooth* if the limit (13) is attained uniformly in $x, y \in S$. It is known that E is uniformly convex if and only if E^* is uniformly smooth. It is also known that if E is reflexive, then E is strictly convex if and only if E^* is smooth; see, for instance, Cioranescu [14] or Takahashi [43] for more details.

For a set-valued mapping $T : E \to 2^{E^*}$ with *domain* $D(T) = \{x \in E : Tx \neq \emptyset\}$ and *range* $R(T) = \bigcup_{x \in E} Tx$, we identify T with its graph $G(T)$ defined by $G(T) = \{(x, x^*) \in E \times E^* : x^* \in Tx\}$. The mapping $T \subset E \times E^*$ is said to be *monotone* if $\langle x-y, x^*-y^*\rangle \geq 0$ whenever $(x,x^*),(y,y^*)\in T$. It is also said to be *maximal monotone* if its graph is not contained in the graph of any other monotone operator on E. If $T\subset E\times E^*$ is maximal monotone, then we can show that the set $T^{-1}0=\{z\in E: 0\in Tz\}$ is closed and convex. A function $f: E\to(-\infty,\infty]$ is said to be *proper* if the domain $D(f)=\{x\in E: f(x)<\infty\}$ is nonempty. It is also said to be *lower semicontinuous* if $\{x\in E: f(x)\leq r\}$ is closed for all $r\in\mathbb{R}$. The function f is also said to be *convex* if

$$f(\alpha x+(1-\alpha)y)\leq \alpha f(x)+(1-\alpha)f(y) \tag{14}$$

for all $x,y\in E$ and $\alpha\in(0,1)$. It is also said to be *strictly convex* if the strict inequality holds in (14) for all $x,y\in D(f)$ with $x\neq y$ and $\alpha\in(0,1)$. For a proper lower semicontinuous convex function $f: E\to(-\infty,\infty]$, the *subdifferential* ∂f of f is defined by

$$\partial f(x)=\{x^*\in E^*: f(x)+\langle y-x, x^*\rangle\leq f(y)\ \ (\forall y\in E)\}$$

for all $x\in E$. Rockafellar's theorem [34, 35] ensures that $\partial f\subset E\times E^*$ is maximal monotone. If $p\in(1,\infty)$ and g is defined by $g(x)=\|x\|^p/p$ for all $x\in E$, then $\partial g=J_p$. The following theorem is also well-known: If $f: E\to(-\infty,\infty]$ is a proper lower semicontinuous convex function and $g: E\to\mathbb{R}$ is a continuous convex function, then $\partial(f+g)=\partial f+\partial g$. For a proper lower semicontinuous convex function $f: E\to(-\infty,\infty]$, the *conjugate function* f^* of f is defined by

$$f^*(x^*)=\sup_{x\in E}\{\langle x,x^*\rangle - f(x)\}$$

for all $x^*\in E^*$. It is known that $f(x)+f^*(x^*)\geq\langle x,x^*\rangle$ for all $(x,x^*)\in E\times E^*$. It is also known that $(x,x^*)\in\partial f$ is equivalent to

$$f(x)+f^*(x^*)=\langle x,x^*\rangle. \tag{15}$$

We also know that if $f: E\to(-\infty,\infty]$ is a proper lower seimicontinuous function, then $f^*: E^*\to(-\infty,\infty]$ is a proper weak* lower semicontinuous convex function; see Phelps [31] or Takahashi [42] for more details on convex analysis.

Let $g: E\to\mathbb{R}$ be a convex function. Then the *directional derivative* $d^+g(x)(y)$ of g at $x\in E$ with the direction $y\in E$ is defined by

$$d^+g(x)(y)=\lim_{t\downarrow 0}\frac{g(x+ty)-g(x)}{t}.$$

The function g is said to be *Gâteaux differentiable* at x if $d^+g(x)\in E^*$ for all $x\in E$. In this case, we denote $d^+g(x)$ by $\nabla g(x)$. The function g is also said to be *Fréchet differentiable* at $x\in E$ if for all $\varepsilon>0$, there exists $\delta>0$ such that $\|y-x\|\leq\delta$ implies

$$|g(y)-g(x)-\langle y-x,\nabla g(x)\rangle|\leq\varepsilon\|y-x\|.$$

A convex function $g: E\to\mathbb{R}$ is said to be *Gâteaux differentiable* (*Fréchet differentiable*, respectively) if it is Gâteaux differentiable everywhere (Fréchet differentiable at everywhere, respectively). We know that if a continuous convex function $g: E\to\mathbb{R}$ is Gâteaux differentiable, then ∇g is norm-to-weak* continuous. We also know that if g is Fréchet differentiable, then ∇g is norm-to-norm continuous; see Butnariu and Iusem [10]. The mapping ∇g is said to be *weakly sequentially continuous* if $z_n\rightharpoonup z$ implies $\nabla g(z_n)\overset{*}{\rightharpoonup}\nabla g(z)$.

Let E be a Banach space and let $g : E \to \mathbb{R}$ be a convex and Gâteaux differentiable function. Then the *Bregman distance* [5, 11] corresponding to g is defined by

$$D(x, y) = g(x) - g(y) - \langle x - y, \nabla g(y) \rangle$$

for all $x, y \in E$. It is obvious that $D(x, y) \geq 0$ for all $x, y \in E$. The function g is also said to be *strongly coercive* if

$$\|z_n\| \to \infty \Longrightarrow \frac{g(z_n)}{\|z_n\|} \to \infty.$$

It is also said to be *bounded on bounded sets* if $g(U)$ is bounded for each bounded subset U of E. The following definition is slightly different from that in Butnariu and Iusem [10]:

Definition 2.1. Let E be a Banach space. Then a function $g : E \to \mathbb{R}$ is said to be a *Bregman function* if the following conditions are satisfied:

(1) g is continuous, strictly convex and Gâteaux differentiable;
(2) the set $\{y \in E : D(x, y) \leq r\}$ is bounded for all $x \in E$ and $r > 0$.

The following lemma follows from Butnariu and Iusem [10] and Zălinescu [46]:

Lemma 2.2. *Let E be a reflexive Banach space and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function. Then*

(1) *$\nabla g : E \to E^*$ is one-to-one, onto and norm-to-weak* continuous;*
(2) *$\langle x - y, \nabla g(x) - \nabla g(y) \rangle = 0$ if and only if $x = y$;*
(3) *$\{x \in E : D(x, y) \leq r\}$ is bounded for all $y \in E$ and $r > 0$;*
(4) *$D(g^*) = E^*$, g^* is Gâteaux differentiable and $\nabla g^* = (\nabla g)^{-1}$;*

If C is a nonempty closed convex subset of a reflexive Banach space E and $g : E \to \mathbb{R}$ is a strongly coercive Bregman function, then for each $x \in E$, there exists a unique $x_0 \in C$ such that

$$D(x_0, x) = \min_{y \in C} D(y, x).$$

The *Bregman projection* P_C from E onto C is defined by $P_C(x) = x_0$ for all $x \in E$. It is well-known that $x_0 = P_C(x)$ if and only if

$$\langle y - x_0, \nabla g(x) - \nabla g(x_0) \rangle \leq 0 \tag{16}$$

for all $y \in C$. It is also known that P_C has the following property:

$$D(u, P_C x) + D(P_C x, x) \leq D(u, x) \tag{17}$$

for all $u \in C$ and $x \in E$; see, for instance, Butnariu and Iusem [10] for more details.

Let E be a Banach space. The closed unit ball and the unit sphere of E are denoted by B and S, respectively. We also denote rB the set $\{z \in E : \|z\| \leq r\}$ for all $r > 0$. Then a function $g : E \to \mathbb{R}$ is said to be *uniformly convex on bounded sets* [46, pp.203, 221] if $\rho_r(t) > 0$ for all $r, t > 0$, where $\rho_r : [0, \infty) \to [0, \infty]$ is defined by

$$\rho_r(t) = \inf_{x, y \in rB,\, \|x - y\| = t,\, \alpha \in (0,1)} \frac{\alpha g(x) + (1 - \alpha) g(y) - g(\alpha x + (1 - \alpha) y)}{\alpha (1 - \alpha)} \tag{18}$$

for all $t \geq 0$. It is known that ρ_r is a nondecreasing function. The function g is also said to be *uniformly smooth on bounded sets* [46, pp.207, 221] if $\lim_{t \downarrow 0} \sigma_r(t)/t = 0$ for all $r > 0$, where $\sigma_r : [0, \infty) \to [0, \infty]$ is defined by

$$\sigma_r(t) = \sup_{x \in rB,\, y \in S,\, \alpha \in (0,1)} \frac{\alpha g(x + (1-\alpha)ty) + (1-\alpha)g(x - \alpha ty) - g(x)}{\alpha(1-\alpha)} \tag{19}$$

for all $t \geq 0$. We know the following theorems; see Zălinescu [46, Propositions 3.6.3, 3.6.4]:

Theorem 2.3. *Let E be a reflexive Banach space and let $g : E \to \mathbb{R}$ be a continuous convex function which is strongly coercive. Then the following are equivalent:*

1. *g is bounded on bounded sets and uniformly smooth on bounded sets;*
2. *g is Fréchet differentiable and ∇g is uniformly norm-to-norm continuous on bounded sets;*
3. *$D(g^*) = E^*$, g^* is strongly coercive and uniformly convex on bounded sets.*

Theorem 2.4. *Let E be a reflexive Banach space and let $g : E \to \mathbb{R}$ be a continuous convex function which is bounded on bounded sets. Then the following are equivalent:*

1. *g is strongly coercive and uniformly convex on bounded sets;*
2. *$D(g^*) = E^*$, g^* is bounded on bounded sets and uniformly smooth on bounded sets;*
3. *$D(g^*) = E^*$, g^* is Fréchet differentiable and ∇g^* is uniformly norm-to-norm continuous on bounded sets.*

We also know the following theorem; see Zălinescu [46, Theorems 3.7.7, 3.7.8]:

Theorem 2.5. *Let E be a Banach space, let $p \in (1, \infty)$ and let $g = \|\cdot\|^p/p$. Then*

1. *E is uniformly convex if and only if g is uniformly convex on bounded sets;*
2. *E is uniformly smooth if and only if g is uniformly smooth on bounded sets.*

Let E be a reflexive Banach space, let $T \subset E \times E^*$ be a maximal monotone operator and let $g : E \to \mathbb{R}$ be a convex, continuous, strongly coercive and Gâteaux differentiable function which is bounded on bounded sets. Applying Lemma 7.1 in Section 7, for all $x \in E$ and $r > 0$, there exists $z \in E$ such that

$$\nabla g(x) \in \nabla g(z) + rTz; \tag{20}$$

see also Otero and Svaiter [29]. If g is additionally assumed to be strictly convex, then ∇g is strictly monotone, that is $\langle x - y, \nabla g(x) - \nabla g(y) \rangle = 0$ implies $x = y$; see Butnariu and Iusem [10, p.13]. Hence such a z is uniquely determined and hence we can define the *resolvent* $J_r : E \to D(T)$ of T corresponding to g by $J_r x = z$ for all $x \in E$. In other words, $J_r = (\nabla g + rT)^{-1} \nabla g$. The *Yosida approximation* $A_r : E \to E$ is also defined by

$$A_r(x) = (\nabla g(x) - \nabla g(J_r x))/r$$

for all $x \in E$. It is obvious that $(J_r x, A_r x) \in T$ for all $x \in E$ and $r > 0$. It is also known that if $T^{-1}0$ is nonempty, then each J_r has the following property:

$$D(u, J_r x) + D(J_r x, x) \leq D(u, x) \tag{21}$$

for all $u \in T^{-1}0$ and $x \in E$.

3. Lemmas

Using an idea by Kamimura and Takahashi [21], we first prove the following lemma:

Lemma 3.1. *Let E be a Banach space and let $g : E \to \mathbb{R}$ be a Gâteaux differentiable function which is uniformly convex on bounded sets. If $\{x_n\}$ and $\{y_n\}$ are bounded sequences in E and $\lim_n D(x_n, y_n) = 0$, then $\lim_n \|x_n - y_n\| = 0$.*

Proof. Let r be a positive real number such that $\|x_n\| \le r$ and $\|y_n\| \le r$ for all $n \in \mathbb{N}$. Let ρ_r be the function defined as in (18). Then we have

$$g(\alpha x + (1-\alpha)y) + \alpha(1-\alpha)\rho_r(\|x - y\|) \le \alpha g(x) + (1-\alpha)g(y)$$

for all $x, y \in rB$ and $\alpha \in (0,1)$. If $x, y \in rB$ and $\alpha \in (0,1)$, then we have

$$\frac{g(y + \alpha(x-y)) - g(y)}{\alpha} \le g(x) - g(y) - (1-\alpha)\rho_r(\|x-y\|).$$

Tending $\alpha \to 0$, we have

$$\langle x - y, \nabla g(y) \rangle \le g(x) - g(y) - \rho_r(\|x - y\|).$$

This implies that $\rho_r(\|x_n - y_n\|) \le D(x_n, y_n)$ for all $n \in \mathbb{N}$. Thus

$$\lim_{n\to\infty} \rho_r(\|x_n - y_n\|) = 0.$$

Then we can show that $\lim_n \|x_n - y_n\| = 0$. In fact, if not, we have $\varepsilon_0 > 0$ and a subsequence $\{n_i\}$ of $\{n\}$ such that $\|x_{n_i} - y_{n_i}\| \ge \varepsilon_0$ for all $i \in \mathbb{N}$. Since ρ_r is nondecreasing, we have $\rho_r(\|x_{n_i} - y_{n_i}\|) \ge \rho_r(\varepsilon_0)$ for all $i \in \mathbb{N}$. Tending $i \to \infty$, we have $0 \ge \rho_{g,r}(\varepsilon_0)$. This contradicts to the uniform convexity of g on bounded sets. □

We can also prove the following lemmas:

Lemma 3.2. *Let E be a reflexive Banach space, let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function and let V be the function defined by*

$$V(x, x^*) = g(x) - \langle x, x^* \rangle + g^*(x^*) \tag{22}$$

for all $x \in E$ and $x^* \in E^*$. Then

$$D(x, \nabla g^*(x^*)) = V(x, x^*)$$

for all $x \in E$ and $x^* \in E^*$.

Proof. Let $x \in E$ and $x^* \in E^*$. Since $x^* = \nabla g(\nabla g^*(x^*)) = \partial g(\nabla g^*(x^*))$, it follows from from (15) that

$$g(\nabla g^*(x^*)) + g^*(x^*) = \langle \nabla g^*(x^*), x^* \rangle.$$

Thus we have

$$\begin{aligned} D(x, \nabla g^*(x^*)) &= g(x) - g(\nabla g^*(x^*)) - \langle x - \nabla g^*(x^*), x^* \rangle \\ &= g(x) + g^*(x^*) - \langle \nabla g^*(x^*), x^* \rangle - \langle x - \nabla g^*(x^*), x^* \rangle \\ &= g(x) - \langle x, x^* \rangle + g^*(x^*) = V(x, x^*). \end{aligned}$$

This completes the proof. □

Lemma 3.3. *Let E be a reflexive Banach space, let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function and let V be the function defined as in* (22). *Then*

$$V(x, x^*) + \langle \nabla g^*(x^*) - x, y^* \rangle \leq V(x, x^* + y^*)$$

for all $x \in E$ and $x^*, y^* \in E^*$.

Proof. Let $x \in E$ be given and let p be the function defined $p(x^*) = V(x, x^*)$ for all $x^* \in E^*$. Then we have $\partial p(x^*) = -x + \nabla g^*(x^*)$ for all $x^* \in E^*$. Hence we have the desired inequality. □

4. Strongly Convergent Proximal-Type Algorithm

Now, we can prove the following strong convergence theorem:

Theorem 4.1. *Let E be a reflexive Banach space, let $T \subset E \times E^*$ be a maximal monotone operator and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function which is uniformly convex on bounded sets and bounded on bounded sets. Let J_r be the resolvent of T for all $r > 0$ and let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$x_{n+1} = \nabla g^*(\alpha_n \nabla g(x) + (1 - \alpha_n)\nabla g(J_{r_n} x_n)) \ \ (n = 1, 2, \dots),$$

where $\{\alpha_n\} \subset [0, 1]$ and $\{r_n\} \subset (0, \infty)$ satisfy $\lim_n \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = \infty$ and $\lim_n r_n = \infty$. If $T^{-1}0 \neq \emptyset$, then the sequence $\{x_n\}$ converges strongly to $P(x)$, where P is the Bregman projection from E onto $T^{-1}0$.

Proof. Put $y_n = J_{r_n} x_n$ for all $n \in \mathbb{N}$. By Lemma 3.2, we have

$$D(Px, x_{n+1}) = V(Px, \alpha_n \nabla g(x) + (1 - \alpha_n)\nabla g(y_n)) \tag{23}$$

for all $n \in \mathbb{N}$. By Lemma 3.2 and the convexity of V in the second variable, we have

$$\begin{aligned} V(Px, \alpha_n \nabla g(x) + (1 - \alpha_n)\nabla g(y_n)) &\leq \alpha_n V(Px, \nabla g(x)) + (1 - \alpha_n) V(Px, \nabla g(y_n)) \\ &= \alpha_n D(Px, x) + (1 - \alpha_n) D(Px, y_n) \end{aligned} \tag{24}$$

Using (21), we have

$$D(Px, y_n) \leq D(Px, x_n). \tag{25}$$

It follows from (23), (24) and (25) that

$$D(Px, x_{n+1}) \leq \alpha_n D(Px, x) + (1 - \alpha_n) D(Px, x_n)$$

for all $n \in \mathbb{N}$. Thus we have

$$D(Px, x_n) \leq D(Px, x)$$

for all $n \in \mathbb{N}$. By assumption, the set $\{y \in E : D(Px, y) \leq D(Px, x)\}$ is bounded and hence $\{x_n\}$ is bounded. It also follows from (25) that $\{y_n\}$ is bounded.

Put $z_n = x_{n+1}$ for all $n \in \mathbb{N}$. We next show that

$$\limsup_{n \to \infty} \langle x_n - Px, \nabla g(x) - \nabla g(Px) \rangle \leq 0. \tag{26}$$

In fact, we have a subsequence $\{z_{n_i}\}$ of $\{z_n\}$ and $z \in E$ such that $z_{n_i} \rightharpoonup z$ and

$$\limsup_{n \to \infty} \langle x_n - Px, \nabla g(x) - \nabla g(Px) \rangle = \langle z - Px, \nabla g(x) - \nabla g(Px) \rangle.$$

In view of (16), to show (26), it is sufficient to show that z is an element of $T^{-1}0$. By the definition of $\{x_n\}$, we have

$$\nabla g(z_n) - \nabla g(y_n) = \alpha_n(\nabla g(x) - \nabla g(y_n))$$

for all $n \in \mathbb{N}$. Since g is bounded on bounded sets, ∇g is also bounded on bounded sets; see Butnariu and Iusem [9, p.16]. This implies that $\{\nabla g(y_n)\}$ is bounded. Hence

$$\lim_{n\to\infty} \|\nabla g(z_n) - \nabla g(y_n)\| = 0.$$

Since g is strongly coercive and uniformly convex on bounded sets, it follows from Theorem 2.4 that ∇g^* is uniformly norm-to-norm continuous on every bounded subset of E^*. Thus we have

$$\lim_{n\to\infty} \|z_n - y_n\| = \lim_{n\to\infty} \|\nabla g^*(\nabla g(z_n)) - \nabla g^*(\nabla g(y_n))\| = 0.$$

This implies that $\{y_{n_i}\}$ also converges weakly to z. On the other hand, we have from $\lim_n r_n = \infty$ that

$$\lim_{n\to\infty} \|A_{r_n} x_n\| = \lim_{n\to\infty} \frac{1}{r_n} \|\nabla g(x_n) - \nabla g(y_n)\| = 0.$$

If $(w, w^*) \in T$, then it follows from $(y_n, A_{r_n} x_n) \in T$ and the monotonicity of T that

$$\langle w - y_n, w^* - A_{r_n} x_n \rangle \geq 0$$

for all $n \in \mathbb{N}$. Tending $n_i \to \infty$, we have

$$\langle w - z, w^* \rangle \geq 0.$$

By the maximality of T, we have $z \in T^{-1}0$.

Fix $\varepsilon > 0$. Then, using (26), we have $m \in \mathbb{N}$ such that

$$\langle x_n - Px, \nabla g(x) - \nabla g(Px) \rangle < \varepsilon \tag{27}$$

for all $n \geq m$. By Lemma 3.3, (25) and (27) we have

$$\begin{aligned}
&D(Px, x_{n+1}) \\
&= V(Px, \alpha_n \nabla g(x) + (1-\alpha_n)\nabla g(y_n)) \\
&\leq V(Px, \alpha_n \nabla g(x) + (1-\alpha_n)\nabla g(y_n) - \alpha_n(\nabla g(x) - \nabla g(Px))) \\
&\quad - \langle x_{n+1} - Px, -\alpha_n(\nabla g(x) - \nabla g(Px)) \rangle \\
&\leq V(Px, \alpha_n \nabla g(Px) + (1-\alpha_n)\nabla g(y_n)) + \alpha_n \varepsilon \\
&\leq \alpha_n V(Px, \nabla g(Px)) + (1-\alpha_n) V(Px, \nabla g(y_n)) + \alpha_n \varepsilon \\
&= (1-\alpha_n) D(Px, y_n) + \alpha_n \varepsilon \\
&\leq \varepsilon\{1 - (1-\alpha_n)\} + (1-\alpha_n) D(Px, x_n)
\end{aligned}$$

for all $n \geq m$. Using this, we can show that

$$\begin{aligned}
&D(Px, x_{n+1}) \\
&\leq \varepsilon\{1 - (1-\alpha_n)\} + (1-\alpha_n)\Big[\varepsilon\{1 - (1-\alpha_{n-1})\} + (1-\alpha_{n-1}) D(Px, x_{n-1})\Big] \\
&= \varepsilon\{1 - (1-\alpha_n)(1-\alpha_{n-1})\} + (1-\alpha_n)(1-\alpha_{n-1}) D(Px, x_{n-1})
\end{aligned}$$

$$\leq \cdots \leq \varepsilon\Big\{1 - \prod_{i=m}^{n}(1-\alpha_i)\Big\} + \prod_{i=m}^{n}(1-\alpha_i)D(Px, x_m)$$

for all $n \geq m$. On the other hand, $\sum_{i=1}^{\infty} \alpha_i = \infty$ implies $\prod_{i=1}^{\infty}(1-\alpha_i) = 0$. Hence we obtain

$$\limsup_{n\to\infty} D(Px, x_n) \leq \varepsilon.$$

and hence $\limsup_n D(Px, x_n) \leq 0$. Thus $\lim_n D(Px, x_n) = 0$. Using Lemma 3.1, we have $\lim_n \|Px - x_n\| = 0$. This completes the proof. □

Using Theorem 4.1, we can prove the following strong convergence theorem in a smooth and uniformly convex Banach space. In the case of $p = 2$, this corollary is reduced to the strong convergence theorem due to Kohsaka and Takahashi [22].

Corollary 4.2. *Let E be a smooth and uniformly convex Banach space and let $T \subset E \times E^*$ be a maximal monotone operator. Let $p \in (1, \infty)$, let J_p be the duality mapping from E into E^* corresponding to the weight function $\omega(t) = t^{p-1}$ and let $Q_r = (J_p + rT)^{-1}J_p$ for all $r > 0$. Let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$x_{n+1} = J_p^{-1}(\alpha_n J_p(x) + (1-\alpha_n)J_p(Q_{r_n}x_n)) \ \ (n = 1, 2, \dots),$$

where $\{\alpha_n\} \subset [0, 1]$ and $\{r_n\} \subset (0, \infty)$ satisfy $\lim_n \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = \infty$ and $\lim_n r_n = \infty$. If $T^{-1}0 \neq \emptyset$, then the sequence $\{x_n\}$ converges strongly to $P(x)$, where P is the Bregman projection from E onto $T^{-1}0$ corresponding to the Bregman function $g = \|\cdot\|^p/p$.

Proof. Since E is strictly convex and smooth, g is also strictly convex and Gâteaux differentiable. It is obvious that g is bounded on bounded sets and strongly coercive. Let $x \in E$ and $r > 0$. Then we next show that $A = \{y \in E : D(x, y) \leq r\}$ is bounded. In fact, if not, then A contains a sequence $\{y_n\}$ such that $\lim_n \|y_n\| = \infty$. On the other hand, we have

$$\begin{aligned} D(x, y_n) &= \|x\|^p/p - \|y_n\|^p/p - \langle x - y_n, J_p(y_n)\rangle \\ &= \|x\|^p/p - \|y_n\|^p/p - \langle x, J_p(y_n)\rangle + \|y_n\|\|J_p y_n\| \\ &\geq \|x\|^p/p + (1 - 1/p)\|y_n\|^p - \|x\|\|y_n\|^{p-1} \\ &= \|x\|^p/p + \|y_n\|^{p-1}\{(1 - 1/p)\|y_n\| - \|x\|\} \end{aligned}$$

for all $n \in \mathbb{N}$. This implies that $\lim_n D(x, y_n) = \infty$. This is a contradiction. Hence A is bounded and hence g is a Bregman function in the sense of Definition 2.1. By Theorem 2.5, g is uniformly convex on bounded sets. We also know that $\nabla g^* = (\nabla g)^{-1} = J_p^{-1}$. Therefore, we have the desired result from Theorem 4.1. □

5. Weakly Convergent Proximal-Type Algorithm

In this section, we first prove the following lemma:

Lemma 5.1. *Let E be a reflexive Banach space, let $T \subset E \times E^*$ be a maximal monotone operator such that $T^{-1}0 \neq \emptyset$ and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function which is uniformly convex on bounded sets and bounded on bounded sets.*

Let P be the Bregman projection from E onto $T^{-1}0$, let J_r be the resolvent of T for all $r > 0$ and let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and

$$x_{n+1} = \nabla g^*(\alpha_n \nabla g(x_n) + (1-\alpha_n)\nabla g(J_{r_n} x_n)) \ \ (n = 1, 2, \dots),$$

where $\{\alpha_n\} \subset [0,1]$ and $\{r_n\} \subset (0,\infty)$. Then the sequence $\{P(x_n)\}$ converges strongly to an element of $T^{-1}0$, which is the unique element z satisfying

$$\lim_{n\to\infty} D(z, x_n) = \min_{y\in T^{-1}0} \lim_{n\to\infty} D(y, x_n).$$

Proof. Let $u \in T^{-1}0$ be given. As in the proof of Theorem 4.1, we can show that

$$\begin{aligned} D(u, x_{n+1}) &\le \alpha_n D(u, x_n) + (1-\alpha_n) D(u, y_n) \\ &\le D(u, x_n) \end{aligned} \tag{28}$$

for all $n \in \mathbb{N}$. Hence the limit $\lim_n D(u, x_n)$ exists and hence $\{x_n\}$ and $\{y_n\}$ are bounded. By the definition of P and (28), we can show that

$$D(Px_{n+1}, x_{n+1}) \le D(Px_n, x_n)$$

for all $n \in \mathbb{N}$ and hence the limit $\lim_n D(Px_n, x_n)$ exits. If $m > n$, then it follows from (17) and (28) that

$$\begin{aligned} D(Px_n, Px_m) &\le D(Px_n, x_m) - D(Px_m, x_m) \\ &\le D(Px_n, x_n) - D(Px_m, x_m). \end{aligned} \tag{29}$$

Then we show that $\{Px_n\}$ is a Cauchy sequence. In fact, if not, we have $\varepsilon_0 > 0$ and subsequences $\{n_i\}$ and $\{m_i\}$ of $\{n\}$ such that

$$\|Px_{n_i} - Px_{m_i}\| \ge \varepsilon_0 \tag{30}$$

for all $i \in \mathbb{N}$. It follows from (17) that $\{Px_n\}$ is bounded. Let $r = \sup_n \|Px_n\|$. Then, as in the proof of Lemma 3.1, we have

$$\rho_r(\|Px_n - Px_m\|) \le D(Px_n, Px_m) \tag{31}$$

for all $m, n \in \mathbb{N}$. On the other hand, by the existence of $\lim_n D(Px_n, x_n)$ and (29), we have $n_0 \in \mathbb{N}$ such that

$$D(Px_n, Px_m) < \rho_r(\varepsilon_0) \tag{32}$$

for all $m, n \in \mathbb{N}$ with $m > n \ge n_0$. Thus it follows from (31) and (32) that

$$\rho_r(\|Px_n - Px_m\|) < \rho_r(\varepsilon_0)$$

for all $m, n \in \mathbb{N}$ with $m > n \ge n_0$. This contradicts to (30) and hence $\{Px_n\}$ is a Cauchy sequence. Since $T^{-1}0$ is closed, $\{Px_n\}$ converges strongly to an element z of $T^{-1}0$.

We next show that z is a minimizer of the continuous convex function $h : T^{-1}0 \to \mathbb{R}$ defined by

$$h(y) = \lim_{n\to\infty} D(y, x_n)$$

for all $y \in T^{-1}0$. If $n \in \mathbb{N}$ and $y \in T^{-1}0$, then we have

$$h(Px_n) = \lim_{m\to\infty} D(Px_n, x_m) \le D(Px_n, x_n) \le D(y, x_n).$$

Tending $n \to \infty$, we have

$$h(z) \le \lim_{n\to\infty} D(y, x_n) = h(y).$$

Thus z is a minimizer of h. We finally show that z is the unique minimizer of g. If not, there exists $y_1, y_2 \in T^{-1}0$ such that $y_1 \neq y_2$ and

$$h(y_1) = h(y_2) = \min_{y \in T^{-1}0} h(y) = \ell.$$

Then we have

$$g\Big(\frac{y_1 + y_2}{2}\Big) \le \frac{1}{2}(g(y_1) + g(y_2)) - \frac{1}{4}\rho_s(\|y_1 - y_2\|),$$

where $s = \max\{\|y_1\|, \|y_2\|\}$. This implies that

$$\begin{aligned}
h\Big(\frac{y_1 + y_2}{2}\Big) &= \lim_{n\to\infty} D\Big(\frac{y_1 + y_2}{2}, x_n\Big) \\
&= \lim_{n\to\infty} \Big\{ g\Big(\frac{y_1 + y_2}{2}\Big) - g(x_n) - \Big\langle \frac{y_1 + y_2}{2} - x_n, \nabla g(x_n) \Big\rangle \Big\} \\
&\le \lim_{n\to\infty} \Big\{ \frac{1}{2}(D(y_1, x_n) + D(y_2, x_n)) - \frac{1}{4}\rho_s(\|y_1 - y_2\|) \Big\} \\
&= \ell - \frac{1}{4}\rho_s(\|y_1 - y_2\|).
\end{aligned}$$

Therefore we have $h((y_1 + y_2)/2) < \ell \le h((y_1 + y_2)/2)$. This is a contradiction. □

Using Lemma 5.1, we next show the following weak convergence theorem:

Theorem 5.2. *Let E be a reflexive Banach space, let $T \subset E \times E^*$ be a maximal monotone operator and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function which is uniformly smooth on bounded sets, uniformly convex on bounded sets and bounded on bounded sets. Let J_r be the resolvent of T for all $r > 0$ and let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$x_{n+1} = \nabla g^*(\alpha_n \nabla g(x_n) + (1 - \alpha_n)\nabla g(J_{r_n} x_n)) \ \ (n = 1, 2, \dots),$$

where $\{\alpha_n\} \subset [0, 1]$ and $\{r_n\} \subset (0, \infty)$ satisfy $\limsup_n \alpha_n < 1$ and $\liminf_n r_n > 0$. If $T^{-1}0 \neq \emptyset$, then the sequence $\{x_n\}$ is bounded and each weak subsequential limit of $\{x_n\}$ belongs to $T^{-1}0$. Further, if ∇g is weakly sequentially continuous, then $\{x_n\}$ converges weakly to the strong $\lim_n P(x_n)$, where P is the Bregman projection from E onto $T^{-1}0$.

Proof. Let $u \in T^{-1}0$ be given and put $y_n = J_{r_n} x_n$ for all $n \in \mathbb{N}$. Then, as in the proof of Lemma 5.1, we can show that $\lim_n D(u, x_n)$ exists and $\{x_n\}$ is bounded. Using (21) and Lemmas 3.2 and 3.3, we can also show that

$$\begin{aligned}
D(u, x_{n+1}) &\le \alpha_n D(u, x_n) + (1 - \alpha_n) D(u, y_n) \\
&\le \alpha_n D(u, x_n) + (1 - \alpha_n)(D(u, x_n) - D(y_n, x_n)) \\
&= D(u, x_n) - (1 - \alpha_n) D(y_n, x_n)
\end{aligned}$$

for all $n \in \mathbb{N}$. Hence we have

$$(1 - \alpha_n) D(y_n, x_n) \le D(u, x_n) - D(u, x_{n+1})$$

for all $n \in \mathbb{N}$. By the existence of $\lim_n D(u, x_n)$ and $\limsup_n \alpha_n < 1$, we have $\lim_n D(y_n, x_n) = 0$. Using Lemma 3.1, we obtain

$$\lim_{n\to\infty} \|y_n - x_n\| = 0. \tag{33}$$

Since g is bounded on bounded sets and uniformly smooth on bounded sets, by Theorem 2.3, ∇g is uniformly norm-to-norm continuous on every bounded subsets of E and hence

$$\lim_{n\to\infty} \|\nabla g(y_n) - \nabla g(x_n)\| = 0.$$

By $\liminf_n r_n > 0$, we have

$$\lim_n \|A_{r_n} x_n\| = \lim_{n\to\infty} \frac{1}{r_n} \|\nabla g(x_n) - \nabla g(y_n)\| = 0.$$

Let $\{x_{n_i}\}$ be a subsequence of $\{x_n\}$ converging weakly to an element w of E. Then it follows from (33) that $\{y_{n_i}\}$ is also weakly convergent to w. If $(v, v^*) \in T$, then we have from $(y_n, A_{r_n} x_n) \in T$ and the monotonicity of T that

$$\langle v - y_n, v^* - A_{r_n} x_n \rangle \geq 0$$

for all $n \in \mathbb{N}$. Tending $n_i \to \infty$, we have $\langle v - w, v^* \rangle \geq 0$. Since T is maximal, we obtain $w \in T^{-1}0$.

We finally show that if ∇g is weakly sequentially continuous, then $\{x_n\}$ converges weakly to $w = \lim_n Px_n$. By (16) and $w \in T^{-1}0$, we have

$$\langle w - Px_n, \nabla g(x_n) - \nabla g(Px_n) \rangle \leq 0$$

for all $n \in \mathbb{N}$. Lemma 5.1 implies that $\{Px_n\}$ is strongly convergent to an element z of $T^{-1}0$. Since $\{x_{n_i}\}$ is weakly convergent to w and ∇g is weakly sequentially continuous, $\{\nabla g(x_{n_i})\}$ is also weakly convergent to $\nabla g(w)$. Thus we have

$$\langle w - z, \nabla g(w) - \nabla g(z) \rangle \leq 0.$$

Using Lemma 2.2, we have $w = z$. Therefore $\{x_n\}$ converges weakly to $z = \lim_n Px_n$. □

If $\alpha_n = 0$ for all $n \in \mathbb{N}$ in Theorem 5.2, we have the following corollary for the proximal point algorithm of Rockafellar's type [37]:

Corollary 5.3. *Let E be a reflexive Banach space, let $T \subset E \times E^*$ be a maximal monotone operator and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function which is uniformly smooth on bounded sets, uniformly convex on bounded sets and bounded on bounded sets. Let J_r be the resolvent of T for all $r > 0$ and let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$x_{n+1} = J_{r_n} x_n \quad (n = 1, 2, \dots),$$

where $\{\alpha_n\} \subset [0, 1]$ and $\{r_n\} \subset (0, \infty)$ satisfy $\limsup_n \alpha_n < 1$ and $\liminf_n r_n > 0$. If $T^{-1}0 \neq \emptyset$, then the sequence $\{x_n\}$ is bounded and each weak subsequential limit of $\{x_n\}$ belongs to $T^{-1}0$. Further, if ∇g is weakly sequentially continuous, then $\{x_n\}$ converges weakly to the strong $\lim_n P(x_n)$, where P is the Bregman projection from E onto $T^{-1}0$.

As in the proof of Corollary 4.2, we can deduce the following corollary from Theorem 5.2. In the case of $p = 2$, this corollary is reduced to the weak convergence theorem due to Kamimura, Kohsaka and Takahashi [17].

Corollary 5.4. *Let E be a uniformly smooth and uniformly convex Banach space and let $T \subset E \times E^*$ be a maximal monotone operator. Let $p \in (1, \infty)$, let J_p be the duality mapping from E into E^* corresponding to the weight function $\omega(t) = t^{p-1}$ and let $Q_r = (J_p + rT)^{-1} J_p$ for all $r > 0$. Let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$x_{n+1} = J_p^{-1}(\alpha_n J_p(x_n) + (1 - \alpha_n) J_p(Q_{r_n} x_n)) \quad (n = 1, 2, \dots),$$

where $\{\alpha_n\} \subset [0, 1]$ and $\{r_n\} \subset (0, \infty)$ satisfy $\limsup_n \alpha_n < 1$ and $\liminf_n r_n > 0$. If $T^{-1}0 \neq \emptyset$, then the sequence $\{x_n\}$ is bounded and each weak subsequential limit of $\{x_n\}$ belongs to $T^{-1}0$. Further, if J_p is weakly sequentially continuous, then $\{x_n\}$ converges weakly to the strong $\lim_n P(x_n)$, where P is the Bregman projection from E onto $T^{-1}0$ corresponding to the Bregman function $g = \| \cdot \|^p / p$.

6. Applications

Using Theorem 4.1, we first study a convex minimization problem in a Banach space.

Corollary 6.1. *Let E be a reflexive Banach space, let $f : E \to (-\infty, \infty]$ be a proper lower semicontinuous convex function and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function which is uniformly convex on bounded sets and bounded on bounded sets. Let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$\begin{cases} y_n = \arg\min_{y \in E} \Big\{ f(y) + \dfrac{1}{r_n} D(y, x_n) \Big\}; \\ x_{n+1} = \nabla g^*(\alpha_n \nabla g(x) + (1 - \alpha_n) \nabla g(y_n)) \quad (n = 1, 2, \dots), \end{cases}$$

where $\{\alpha_n\} \subset [0, 1]$ and $\{r_n\} \subset (0, \infty)$ satisfy $\lim_n \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = \infty$ and $\lim_n r_n = \infty$. If $(\partial f)^{-1}(0) \neq \emptyset$, then the sequence $\{x_n\}$ converges strongly to $P_{(\partial f)^{-1}(0)}(x)$.

Proof. By Rockafellar's theorem [34, 35], ∂f is a maximal monotone operator. Let $r > 0$ and let J_r be the resolvent of ∂f. Then $x_r = J_r x$ is equivalent to

$$\begin{aligned} 0 &\in \partial f(x_r) + \frac{1}{r}(\nabla g(x_r) - \nabla g(x)) \\ &= \partial \Big(f + \frac{1}{r}(g - \nabla g(x)) \Big)(x_r), \end{aligned}$$

which is also equivalent to

$$\begin{aligned} x_r &= \arg\min_{y \in E} \Big\{ f(y) + \frac{1}{r}(g(y) - \langle y, \nabla g(x) \rangle) \Big\} \\ &= \arg\min_{y \in E} \Big\{ f(y) + \frac{1}{r} D(y, x) \Big\}. \end{aligned}$$

This implies that $y_n = J_{r_n} x_n$ for all $n \in \mathbb{N}$. Thus, by Theorem 4.1, we have the desired result. ∎

We next study a variational inequality problem in a Banach space. Let C be a nonempty closed convex subset of a Banach space E and let A be a single-valued monotone operator form C into E^*. The operator A is said to be *hemicontinuous* if it is continuous along each line segment contained in C with respect to the weak* topology of E^*. A point u of C is said to be a solution to the *variational inequality* for A if

$$\langle y-u, Au\rangle \geq 0$$

for all $y \in C$. We denote the set of all solutions to the variational inequality for A by $VI(C,A)$. The *normal cone* for C at $x \in C$ is also defined by

$$N_C(x) = \{x^* \in E^* : \langle y-x, x^*\rangle \leq 0 \ \ (\forall y \in C)\}.$$

Corollary 6.2. *Let C be a nonempty closed convex subset of a reflexive Banach space E, let $A : C \to E^*$ be a single-valued, monotone and hemicontinuous operator and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function which is uniformly convex on bounded sets and bounded on bounded sets. Let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$\begin{cases} y_n = VI\Big(C, A + \dfrac{1}{r_n}(\nabla g - \nabla g(x_n))\Big); \\ x_{n+1} = \nabla g^*(\alpha_n \nabla g(x) + (1-\alpha_n)\nabla g(y_n)) \quad (n = 1, 2, \dots), \end{cases}$$

where $\{\alpha_n\} \subset [0,1]$ and $\{r_n\} \subset (0,\infty)$ satisfy $\lim_n \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = \infty$ and $\lim_n r_n = \infty$. If $VI(C,A) \neq \emptyset$, then the sequence $\{x_n\}$ converges strongly to $P_{VI(C,A)}(x)$.

Proof. Let $T \subset E \times E^*$ be a set-valued mapping defined by

$$Tx = \begin{cases} Ax + N_C(x) & (x \in C); \\ \emptyset & (\text{otherwise}). \end{cases} \tag{34}$$

By Rockafellar's theorem [36], T is a maximal monotone operator and $T^{-1}0 = VI(C,A)$. Let $r > 0$ and let J_r be the resolvent of T. Then $x_r = J_r x$ is equivalent to

$$-Ax_r - \frac{1}{r}\nabla g(x_r) + \frac{1}{r}\nabla g(x) \in N_C(x_r),$$

which is also equivalent to

$$\Big\langle y - x_r, Ax_r + \frac{1}{r}\nabla g(x_r) - \frac{1}{r}\nabla g(x)\Big\rangle \geq 0$$

for all $y \in C$. Thus we have

$$x_r = VI\Big(C, A + \frac{1}{r}(\nabla g - \nabla g(x))\Big).$$

Hence $y_n = J_{r_n} x_n$ for all $n \in \mathbb{N}$. Thus, by Theorem 4.1, we have the desired result. □

As in the proofs of Corollaries 6.1 and 6.2, we have the following corollaries from Theorem 5.2:

Corollary 6.3. *Let E be a reflexive Banach space, let $f : E \to (-\infty, \infty]$ be a proper lower semicontinuous convex function and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function which is uniformly smooth on bounded sets, uniformly convex on bounded sets and bounded on bounded sets. Let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$\begin{cases} y_n = \arg\min_{y\in E}\Big\{f(y) + \dfrac{1}{r_n}D(y, x_n)\Big\}; \\ x_{n+1} = \nabla g^*(\alpha_n \nabla g(x_n) + (1-\alpha_n)\nabla g(y_n)) \quad (n = 1, 2, \dots), \end{cases}$$

where $\{\alpha_n\} \subset [0, 1]$ and $\{r_n\} \subset (0, \infty)$ satisfy $\limsup_n \alpha_n < 1$ and $\liminf_n r_n > 0$. If $(\partial f)^{-1}(0) \neq \emptyset$, then the sequence $\{x_n\}$ is bounded and each weak subsequential limit of $\{x_n\}$ belongs to $(\partial f)^{-1}(0)$. Further, if ∇g is weakly sequentially continuous, then $\{x_n\}$ converges weakly to the strong $\lim_n P_{(\partial f)^{-1}(0)}(x_n)$.

Corollary 6.4. *Let C be a nonempty closed convex subset of a reflexive Banach space E, let $A : C \to E^*$ be a single-valued, monotone and hemicontinuous operator and let $g : E \to \mathbb{R}$ be a strongly coercive Bregman function which is uniformly smooth on bounded sets, uniformly convex on bounded sets and bounded on bounded sets. Let $\{x_n\}$ be a sequence generated by $x_1 = x \in E$ and*

$$\begin{cases} y_n = VI\Big(C, A + \dfrac{1}{r_n}(\nabla g - \nabla g(x_n))\Big); \\ x_{n+1} = \nabla g^*(\alpha_n \nabla g(x_n) + (1-\alpha_n)\nabla g(y_n)) \quad (n = 1, 2, \dots), \end{cases}$$

where $\{\alpha_n\} \subset [0, 1]$ and $\{r_n\} \subset (0, \infty)$ satisfy $\limsup_n \alpha_n < 1$ and $\liminf_n r_n > 0$. If $VI(C, A) \neq \emptyset$, then the sequence $\{x_n\}$ is bounded and each weak subsequential limit of $\{x_n\}$ belongs to $VI(C, A)$. Further, if ∇g is weakly sequentially continuous, then $\{x_n\}$ converges weakly to the strong $\lim_n P_{VI(C,A)}(x_n)$.

7. Appendix

In this section, we give another proof of the existence of a solution to the equation (20). Let E be a Banach space and let $B : E \to E^*$ be a single-valued operator. Then B is said to be *coercive* if

$$\|z_n\| \to \infty \Longrightarrow \frac{\langle z_n, Bz_n\rangle}{\|z_n\|} \to \infty.$$

We know the following lemma; see Barbu [2] and Takahashi [42]:

Lemma 7.1. *Let K be a nonempty closed convex subset of a reflexive Banach space E and let $A \subset E \times E^*$ be a monotone operator such that $D(A) \subset K$ and $0 \in D(A)$. Let $B : E \to E^*$ be a single-valued, monotone and hemicontinuous operator such that B is bounded on bounded sets and coercive. Then there exists $x_0 \in K$ such that $\langle z - x_0, z^* + Bx_0\rangle \geq 0$ for all $(z, z^*) \in A$.*

Using Lemma 7.1, we prove the following:

Lemma 7.2. *Let E be a reflexive Banach space and let $T \subset E \times E^*$ be a maximal monotone operator. Let $g : E \to \mathbb{R}$ be a convex, continuous, strongly coercive and Gâteaux differentiable function which is bounded on bounded sets. Then*

$$R(\nabla g + rT) = E^*$$

for all $r > 0$.

Proof. Let $r > 0$ and $x^* \in E^*$ be given. Since T is maximal monotone, the graph of T is nonempty. Fix $(q, q^*) \in T$. Then we can define a maximal monotone operator $A \subset E \times E^*$ by $A(x) = T(x + q)$ for all $x \in E$. We can also define a single-valued monotone hemicontinuous operator $B : E \to E^*$ by

$$B(x) = \frac{1}{r}(\nabla g(x + q) - x^*)$$

for all $x \in E$. By Butnariu and Iusem [9], ∇g is bounded on bounded sets if and only if g is bounded on bounded sets. Hence B is also bounded on bounded sets. We next show that B is coercive. If $\|z_n\| \to \infty$, then we have

$$g(z_n + q) + \langle y - (z_n + q), \nabla g(z_n + q)\rangle \le g(y)$$

for all $y \in E$ and $n \in \mathbb{N}$. Putting $y = q$, we have

$$\langle z_n, \nabla g(z_n + q)\rangle \ge g(z_n + q) - g(q) \tag{35}$$

for all $n \in \mathbb{N}$. Then it follows from (35) and the strong coercivity of g that

$$\begin{aligned}\frac{\langle z_n, Bz_n\rangle}{\|z_n\|} &= \frac{1}{r} \cdot \frac{\langle z_n, \nabla g(z_n + q) - x^*\rangle}{\|z_n\|} \\ &= \frac{1}{r}\Big\{\frac{\|z_n + q\|}{\|z_n\|} \cdot \frac{\langle z_n, \nabla g(z_n + q)\rangle}{\|z_n + q\|} - \frac{\langle z_n, x^*\rangle}{\|z_n\|}\Big\} \\ &\ge \frac{1}{r}\Big\{\frac{\|z_n + q\|}{\|z_n\|} \cdot \frac{g(z_n + q) - g(q)}{\|z_n + q\|} - \frac{\langle z_n, x^*\rangle}{\|z_n\|}\Big\} \to \infty\end{aligned}$$

as $n \to \infty$. Thus B is coercive. By Lemma 7.1, we have $x_0 \in E$ such that

$$\langle z - x_0, z^* + Bx_0\rangle \ge 0$$

for all $(z, z^*) \in A$. Since A is maximal monotone, we have $(x_0, -Bx_0) \in A$. Thus

$$x^* \in \nabla g(x_0 + q) + rT(x_0 + q).$$

Therefore we obtain $R(\nabla g + rT) = E^*$. □

References

[1] Y. I. Alber, *Metric and generalized projections in Banach spaces: properties and applications*, in Theory and Applications of Nonlinear Operators of Accretive and Monotone Type (A. G. Kartsatos Ed.), Dekker, New York 1996, pp. 15–20.

[2] V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach spaces*, Editura Academiei R. S. R., Bucuresti (1976).

[3] H. H. Bauschke, J. M. Borwein and P. L. Combettes, *Bregman monotone optimization algorithms*, SIAM J. Control Optim. **42** (2003), 596–636.

[4] H. H. Bauschke and P. L. Combettes, *Construction of best Bregman approximations in reflexive Banach spaces*, Proc. Amer. Math. Soc. **131** (2003), 3757–3766.

[5] L. M. Bregman, *The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Comput. Math. Math. Phys. **7** (1967), 200–217.

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

[7] R. S. Burachik and A. N. Iusem, *A generalized proximal point algorithm for the variational inequality problem in a Hilbert space*, SIAM J. Optim. **8** (1998), 197–216.

[8] R. S. Burachik and S. Scheimberg, *A proximal point method for the variational inequality problem in Banach spaces*, SIAM J. Control. Optim. **39** (2000), 1633–1649.

[9] D. Butnariu and A. N. Iusem, *On a proximal point method for convex optimization in Banach spaces*, Numer. Funct. Anal. Optim. **18** (1997), 723–744.

[10] D. Butnariu and A. N. Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization*, Kluwer Academic Publishers, Dordrecht (2000).

[11] Y. Censor and A. Lent, *An iterative row-action method for interval convex programming*, J. Optim. Theory Appl. **34** (1981), 321–353.

[12] Y. Censor and S. Reich, *Iterations of paracontractions and firmly nonexpansive operators with applications to feasibility and optimization*, Optimization **37** (1996), 323–339.

[13] Y. Censor and S. A. Zenios, *Proximal minimization algorithm with D-functions*, J. Optim. Theory Appl. **73** (1992), 451–464.

[14] I. Cioranescu, *Geometry of Banach Spaces, duality mappings and nonlinear problems*, Kluwer Academic Publishers Group, Dordrecht (1990).

[15] O. Güler, *On the convergence of the proximal point algorithm for convex minimization*, SIAM J. Control Optim. **29** (1991), 403–419.

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

[17] S. Kamimura, F. Kohsaka and W. Takahashi, *Weak and strong convergence theorems for maximal monotone operators in a Banach space*, Set-Valued Anal. **12** (2004), 417–429.

[18] S. Kamimura and W. Takahashi, *Approximating solutions of maximal monotone operators in Hilbert spaces*, J. Approx. Theory **106** (2000), 226–240.

[19] S. Kamimura and W. Takahashi, *Iterative schemes for approximating solutions of accretive operators in Banach spaces*, Sci. Math. **3** (2000), 107–115.

[20] S. Kamimura and W. Takahashi, *Weak and strong convergence of solutions to accretive operator inclusions and applications*, Set-Valued Anal. **8** (2000), 361–374.

[21] S. Kamimura and W. Takahashi, *Strong convergence of a proximal-type algorithm in a Banach space*, SIAM J. Optim. **13** (2002), 938–945.

[22] F. Kohsaka and W. Takahashi, *Strong convergence of an iterative sequence for maximal monotone operators in a Banach space*, Abstr. Appl. Anal. **2004** (2004), 239–249.

[23] P. L. Lions, *Une méthode itérative de résolution d'une inéquation variationnelle*, Israel J. Math. **31** (1978), 204–208.

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

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

[26] K. Nakajo and W. Takahashi, *Approximation of a zero of maximal monotone operators in Hilbert spaces*, Proceedings of the Second International Conference on Nonlinear Analysis and Convex Analysis (W. Takahashi and T. Tanaka Eds., Hirosaki, 2001), 303–314, Yokohama Publishers, Yokohama 2003.

[27] K. Nakajo and W. Takahashi, *Strong convergence theorems for nonexpansive mappings and nonexpansive semigroups*, J. Math. Anal. Appl. **279** (2003), 372–379.

[28] S. Ohsawa and W. Takahashi, *Strong convergence theorems for resolvents of maximal monotone operators in Banach spaces*, Arch. Math. **81** (2003), 439–445.

[29] R. G. Otero and B. F. Svaiter, *A strongly convergent hybrid proximal method in Banach spaces*, J. Math. Anal. Appl. **289** (2004), 700–711.

[30] G. B. Passty, *Ergodic convergence to a zero of the sum of monotone operators in Hilbert space*, J. Math. Anal. Appl. **72** (1979), 383–390.

[31] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Mathematics, 1364, Springer-Verlag (1989).

[32] S. Reich, *Weak convergence theorems for nonexpansive mappings in Banach spaces*, J. Math. Anal. Appl. **67** (1979), 274–276.

[33] S. Reich, *A weak convergence theorem for the alternating method with Bregman distance*, in Theory and Applications of Nonlinear Operators of Accretive and Monotone Type (A. G. Kartsatos Ed.), 313–318, Dekker, New York 1996.

[34] R. T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pacific J. Math. **17** (1966), 497–510.
[35] R. T. Rockafellar, *On the maximal monotonicity of subdifferential mappings*, Pacific J. Math. **33** (1970), 209–216.
[36] R. T. Rockafellar, *On the maximality of sums of nonlinear monotone operators*, Trans. Amer. Math. Soc. **149** (1970), 75–88.
[37] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. **14** (1976), 877–898.
[38] N. Shioji and W. Takahashi, *Strong convergence of approximated sequences for nonexpansive mappings in Banach spaces*, Proc. Amer. Math. Soc. **125** (1997), 3641–3645.
[39] M. V. Solodov and B. F. Svaiter, *Forcing strong convergence of proximal point iterations in a Hilbert space*, Math. Program. **87** (2000), 189–202.
[40] W. Takahashi, *Iterative methods for approximation of fixed points and their applications*, J. Oper. Res. Soc. Japan **43** (2000), 87–108.
[41] W. Takahashi, *Fixed point theorems and proximal point algorithms*, Proceedings of the Second International Conference on Nonlinear Analysis and Convex Analysis (W. Takahashi and T. Tanaka Eds., Hirosaki, 2001), 471–481, Yokohama Publishers, Yokohama, 2003.
[42] W. Takahashi, *Convex Analysis and Approximation of Fixed Points*, Yokohama Publishers, Yokohama (2000) (Japanese).
[43] W. Takahashi, *Nonlinear Functional Analysis -Fixed Point Theory and its Applications*, Yokohama Publishers, Yokohama (2000).
[44] R. Wittmann, *Approximation of fixed points of nonexpansive mappings*, Arch. Math. **58** (1992), 486–491.
[45] C. Zălinescu, *On uniformly convex functions*, J. Math. Anal. Appl. **95** (1983), 344–374.
[46] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing Co. Inc., River Edge NJ, (2002).

Manuscript received May 20, 2005
revised August 12, 2005

Fumiaki Kohsaka
Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-okayama, Meguro-ku, Tokyo, 152-8552, Japan
E-mail address: `kohsaka9@is.titech.ac.jp`

Wataru Takahashi
Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-okayama, Meguro-ku, Tokyo, 152-8552, Japan
E-mail address: `wataru@is.titech.ac.jp`