A Bregman projection method for approximating fixed points of quasi-Bregman nonexpansive mappings

Heinz H. Bauschke*, Jiawei Chen†, and Xianfu Wang‡

accepted November 18, 2013 — to appear in *Applicable Analysis*

Dedicated to Boris Mordukhovich on the occasion of his 65th Birthday

Abstract

We introduce an abstract algorithm that aims to find the Bregman projection onto a closed convex set. As an application, the asymptotic behaviour of an iterative method for finding a fixed point of a quasi Bregman nonexpansive mapping with the fixed-point closedness property is analyzed. We also show that our result is applicable to Bregman subgradient projectors.

Key words: Bregman projection, Bregman subgradient projector, fixed point, Legendre function, Moreau envelope, quasi Bregman nonexpansive.
2010 Mathematics Subject Classification: Primary 47H09; Secondary 52B55, 65K10, 90C25.

1 Introduction

Bregman distances provide a general and flexible framework for studying optimization problems both theoretically and algorithmically [1]–[18]. The objective of this paper is to present an iterative method for finding the Bregman projection onto a closed convex set. Our results extends those of [5] from the Euclidean distance to the Bregman distance setting.

The paper is organized as follows. Section 2 contains useful auxiliary results on Bregman distances. In Section 3, we introduce and analyze the iteration scheme for finding the Bregman projection onto a closed convex set. In Section 4, we apply our iteration scheme to quasi Bregman nonexpansive mappings that are fixed-point closed. The iterates are shown to converge to the fixed point which is the Bregman nearest point to the starting point; moreover, the total length of the trajectory in terms of the Bregman distance is finite. We conclude by pointing out that our theory applies to Bregman subgradient projectors.

*Mathematics, Irving K. Barber School, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. Email: heinz.bauschke@ubc.ca.

†School of Mathematics and Statistics, Southwest University, Chongqing 400715, China. Email: J.W.Chen713@163.com.

‡Mathematics, Irving K. Barber School, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. Email: shawn.wang@ubc.ca.

2 Assumptions, notions and facts

2.1 Standing assumptions

We assume throughout this paper that

$$C \text{ is a closed convex subset of a finite dimensional Euclidean space } X$$

with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, and that

$$f\colon X \to \mathbb{R} \text{ is strictly convex and differentiable, with } \operatorname{dom} f^* \text{ open.}$$

We shall assume that the reader is familiar with basic convex and variational analysis and its notation; see, e.g., [16], [19], [22], or [8]. Our assumptions imply that f is a Legendre function as hence is it Fenchel conjugate f^*. When there is no chance of confusion, we will write $x_n \to 0$ instead of $\lim_{n\to\infty} x_n = 0$ or $\lim_{n\to\infty} \|x_n\| = 0$ and analogously for other similar sequential expressions.

2.2 Bregman distance and projection

Definition 2.1 (Bregman distance) (See [10].) The function

$$D\colon X \times X \to \mathbb{R}_+\colon (x,y) \mapsto f(x) - f(y) - \langle \nabla f(y), x - y \rangle$$

is called the *Bregman distance* with respect to f.

It is well known (see, e.g., [2] or [14]) that the Bregman distance allows nonorthogonal projections in our setting:

Definition 2.2 ((left) Bregman projection) For every $y \in X$, there exists a unique point $\overleftarrow{P}_C(y)$ in C, called the *(left) Bregman projection* of y onto C, such that $D(\overleftarrow{P}_C(y), y) = \min_{c\in C} D(c,y)$.

Note that when $f = (1/2)\|\cdot\|^2$, then $\overleftarrow{P}_C$ is the classical orthogonal projector.

2.3 Useful facts

The following results, which are mostly well known and which will be useful later, are recalled here for the reader's convenience.

Fact 2.3 (See, e.g., [7, Fact 2.3].) For every $x \in X$, the projection $\overleftarrow{P}_C(x)$ is characterized by

$$\overleftarrow{P}_C(x) \in C \quad \text{and} \quad (\forall c \in C) \quad \langle \nabla f(x) - \nabla f(\overleftarrow{P}_C(x)), c - \overleftarrow{P}_C(x) \rangle \leq 0; \tag{1}$$

equivalently, by

$$\overleftarrow{P}_C(x) \in C \quad \text{and} \quad (\forall c \in C) \quad D(c,x) \geq D(c, \overleftarrow{P}_C(x)) + D(\overleftarrow{P}_C(x), x). \tag{2}$$

Moreover, $\overleftarrow{P}_C\colon X \to C$ is continuous.

Lemma 2.4 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X, let $(y_n)_{n\in\mathbb{N}}$ be a bounded sequence in X, and suppose that $D(x_n, y_n) \to 0$. Then $x_n - y_n \to 0$.*

Proof. Combine [12, Remark 2.14] with [12, Theorem 2.10]. ■

Lemma 2.5 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. Then the following are equivalent:*

(i) $(x_n)_{n\in\mathbb{N}}$ *is bounded.*

(ii) $(D(x_n, y))_{n\in\mathbb{N}}$ *is bounded for every* $y \in X$.

(iii) *There exists* $y \in X$ *such that* $(D(x_n, y))_{n\in\mathbb{N}}$ *is bounded.*

Proof. "(i)⇒(ii)": Let $y \in X$ and suppose that $(D(x_n, y))_{n\in\mathbb{N}}$ is not bounded. After passing to a subsequence if necessary, we assume that $x_n \to x \in X$ and that $D(x_n, y) \to +\infty$. On the other hand, $D(x_n, y) \to D(x, y) \in \mathbb{R}_+$. Altogether, we have reached a contradiction.

"(ii)⇒(iii)": This is clear.

"(iii)⇒(i)": Suppose that $(x_n)_{n\in\mathbb{N}}$ is not bounded. After passing to a subsequence if necessary, we assume that $\|x_n\| \to +\infty$. By [2, Theorem 3.7.(iii)], $D(\cdot, y)$ is coercive and thus $D(x_n, y) \to +\infty$, which is absurd. ■

Lemma 2.6 *Let* $x \in X$ *and let* $(y_n)_{n\in\mathbb{N}}$ *be a sequence in* X *such that* $(D(x, y_n))_{n\in\mathbb{N}}$ *is bounded. Then* $(y_n)_{n\in\mathbb{N}}$ *is bounded.*

Proof. This follows readily from [2, Corollary 3.11]. ■

3 Finding the Bregman projection by iteration

In this section, we present an iteration scheme to find the projection $\overleftarrow{P}_C(x_0)$. It will be convenient to set, for every $(x, y) \in X \times X$,

$$\begin{aligned} H(x, y) &:= \big\{z \in X \;\big|\; D(z, y) \leq D(z, x)\big\} \\ &= \big\{z \in X \;\big|\; \langle \nabla f(x) - \nabla f(y), z\rangle \leq f(y) - f(x) + \langle \nabla f(x), x\rangle - \langle \nabla f(y), y\rangle\big\}, \end{aligned}$$

which is either equal to X (if $x = y$) or to a closed halfspace (if $x \neq y$).

Algorithm 3.1 Given $x_0 \in X$ and a nonempty closed convex subset C_0 of X, set $n := 0$.

Step 1. Take $y_n \in X$ and set $C_{n+1} := C_n \cap H(x_n, y_n)$.

Step 2. Compute

$$x_{n+1} := \overleftarrow{P}_{C_{n+1}}(x_0) \tag{3}$$

and stop provided a stopping criterion is satisfied.

Step 3. Set $n := n + 1$ and go to Step 1.

Remark 3.2 Since each $H(x_n, y_n)$ is equal to either X or some closed halfspace, we note that the each set C_{n+1} is closed and convex. To compute (3) may in general be a challenging problem; however, our assumptions are similar to those found in the literature (see, e.g., [21, Algorithm (3.6)]). Furthermore, if C_0 is a polyhedron (e.g., $C_0 = X$), then so is C_{n+1} which makes the computations more tractable. It would be interesting to obtain variants of this algorithm that would allow for the removal of some of the earlier halfspaces. Finally, the sequence $(y_n)_{n\in\mathbb{N}}$ is later (see Section 4) chosen to be equal to $(Tx_n)_{n\in\mathbb{N}}$, where T is some well behaved operator.

Let us collect some basic properties of Algorithm 3.1.

Proposition 3.3 *Suppose that* $(x_n)_{n\in\mathbb{N}}$ *is a sequence generated by Algorithm 3.1. Then the following hold:*

(i) **(decreasing sets)** $(\forall n \in \mathbb{N})\ C_n \supseteq C_{n+1}$.

(ii) **(increasing distances)** $(\forall n \in \mathbb{N})\ D(x_n, x_0) \leq D(x_{n+1}, x_0)$.

(iii) $(\forall k \in \mathbb{N}) \sum_{n=0}^{k} D(x_{n+1}, x_n) \leq D(\overleftarrow{P}_{C_{k+1}} x_0, x_0)$.

(iv) *The constant*

$$\beta := \lim_{n \in \mathbb{N}} D(x_n, x_0) = \sup_{n \in \mathbb{N}} D(x_n, x_0) \tag{4}$$

is well defined.

(v) *For all nonnegative integers m and n such that $m < n$, we have*

$$\langle \nabla f(x_0) - \nabla f(x_m), x_n - x_m \rangle \leq 0 \tag{5}$$

and

$$D(x_n, y_m) \leq D(x_n, x_m). \tag{6}$$

Proof. We only show (iii) and (v) because the other properties are clear.

(iii): In (2), put $x = x_0$, $C = C_n$. For $x_{n+1} \in C_{n+1} \subseteq C_n$, we have $D(x_{n+1}, x_0) \geq D(x_{n+1}, x_n) + D(x_n, x_0)$, i.e., $D(x_{n+1}, x_0) - D(x_n, x_0) \geq D(x_{n+1}, x_n)$. Now sum the last inequality over $n \in \{0, 1, \dots, k\}$.

(v): In (1), put $x = x_0$, $c = x_n$, $C = C_m$, noting that $x_m = \overleftarrow{P}_{C_m} x_0$ and $x_n \in C_n \subseteq C_m$ when $n > m$. This gives (5). Finally (6) follows because $x_n \in C_n \subseteq C_{m+1} = C_m \cap H(x_m, y_m)$. ■

We now begin the convergence analysis of Algorithm 3.1.

Lemma 3.4 *Suppose that $(x_n)_{n \in \mathbb{N}}$ is a sequence generated by Algorithm 3.1 and that $(x_n)_{n \in \mathbb{N}}$ is bounded. Then the following hold:*

(i) $\sum_{n \in \mathbb{N}} D(x_{n+1}, x_n) < +\infty$.

(ii) $x_{n+1} - x_n \to 0$.

Proof. (i): By Lemma 2.5, $(D(\overleftarrow{P}_{C_n} x_0, x_0))_{n \in \mathbb{N}} = (D(x_n, x_0))_{n \in \mathbb{N}}$ is bounded. Now apply Proposition 3.3(iii).

(ii): Since $D(x_{n+1}, x_n) \to 0$ by (i), we deduce from Lemma 2.4 that $x_{n+1} - x_n \to 0$. ■

Lemma 3.5 *Suppose that $(x_n)_{n \in \mathbb{N}}$ is a sequence generated by Algorithm 3.1 such that for every subsequence $(x_{k_n})_{n \in \mathbb{N}}$ of $(x_n)_{n \in \mathbb{N}}$, we have*

$$\left. \begin{array}{r} x_{k_n} \to \bar{x} \\ x_{k_n} - y_{k_n} \to 0 \end{array} \right\} \Rightarrow \bar{x} \in C. \tag{7}$$

Then every bounded subsequence of $(x_n)_{n \in \mathbb{N}}$ must converge to a point in C.

Proof. Suppose that $(x_{k_n})_{n \in \mathbb{N}}$ is a bounded subsequence of $(x_n)_{n \in \mathbb{N}}$. By Lemma 2.5, $(D(x_{k_n}, x_0))_{n \in \mathbb{N}}$ is bounded. Hence Proposition 3.3(iv) implies that the constant β defined in (4) belongs to $\mathbb{R}_+$. Let m and n be in $\mathbb{N}$ such that $m < n$. Then $x_{k_n} \in C_{k_n} \subseteq C_{k_m}$. Using Fact 2.3 (applied to $x = x_0, C = C_{k_m}$) and (4), we have

$$D(x_{k_n}, x_{k_m}) \leq D(x_{k_n}, x_0) - D(x_{k_m}, x_0) \to \beta - \beta = 0 \quad \text{as} \quad n > m \to +\infty.$$

It now follows from Lemma 2.4 that $x_{k_n} - x_{k_m} \to 0$ as $n > m \to +\infty$, i.e., $(x_{k_n})_{n \in \mathbb{N}}$ is a Cauchy sequence. Therefore,

$$x_{k_{n+1}} - x_{k_n} \to 0 \tag{8}$$

and there exists $\bar{x} \in X$ such that

$$x_{k_n} \to \bar{x}.$$

It follows from Remark 3.2 and (3) that $\bar{x} \in C_{k_n}$ and that $\bar{x} \in H(x_{k_n}, y_{k_n})$ for every $n \in \mathbb{N}$. By the definition of H, one has

$$D(\bar{x}, y_{k_n}) \leq D(\bar{x}, x_{k_n}) = f(\bar{x}) - f(x_{k_n}) - \langle \nabla f(x_{k_n}), \bar{x} - x_{k_n} \rangle \to 0.$$

Hence $(D(\bar{x}, y_{k_n}))_{n\in\mathbb{N}}$ is bounded. By Lemma 2.6, $(y_{k_n})_{n\in\mathbb{N}}$ is bounded too. Now, from $x_{k_{n+1}} \in C_{k_{n+1}} \subseteq H(x_{k_n}, y_{k_n})$, we obtain

$$D(x_{k_{n+1}}, y_{k_n}) \leq D(x_{k_{n+1}}, x_{k_n}) \to 0.$$

Again from Lemma 2.4, one has

$$x_{k_{n+1}} - y_{k_n} \to 0. \tag{9}$$

Combining (8) with (9), we deduce that

$$\|x_{k_n} - y_{k_n}\| \leq \|x_{k_n} - x_{k_{n+1}}\| + \|x_{k_{n+1}} - y_{k_n}\| \to 0.$$

This and (7) yield the result. ■

Lemma 3.4 and Lemma 3.5 allow us to derive the following dichotomy result.

Theorem 3.6 (dichotomy) Suppose that $(x_n)_{n\in\mathbb{N}}$ is generated by Algorithm 3.1, that $(\forall n \in \mathbb{N})\ C \subseteq C_n$, and that for every subsequence $(x_{k_n})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$, we have

$$\left.\begin{array}{r} x_{k_n} \to \bar{x} \\ x_{k_n} - y_{k_n} \to 0 \end{array}\right\} \Rightarrow \bar{x} \in C. \tag{10}$$

Then exactly one of the following holds:

(i) $C \neq \varnothing$, $x_n \to \overleftarrow{P}_C x_0$, and $\sum_{n\in\mathbb{N}} D(x_{n+1}, x_n) < +\infty$.

(ii) $C = \varnothing$ and $\|x_n\| \to +\infty$.

Proof. Note first that

$$(\forall n \in \mathbb{N}) \quad D(x_n, x_0) = \inf_{c \in C_n} D(c, x_0) \leq \inf_{c \in C} D(c, x_0). \tag{11}$$

(i): Assume that $C \neq \varnothing$. Then $(x_n)_{n\in\mathbb{N}}$ is bounded by (11) and Lemma 2.5. By Lemma 3.5,

$$\bar{x} := \lim_{n\in\mathbb{N}} x_n \in C.$$

On the other hand, (11) yields

$$D(\bar{x}, x_0) \leq \inf_{c\in C} D(c, x_0).$$

Altogether, $\bar{x} = \overleftarrow{P}_C x_0$. Finally, $\sum_{n\in\mathbb{N}} D(x_{n+1}, x_n) < +\infty$ because of Lemma 3.4.

(ii): Suppose that $\|x_n\| \not\to +\infty$. Then $(x_n)_{n\in\mathbb{N}}$ contains a bounded subsequence which, by Lemma 3.5, must converge to a point in C. Therefore if $C = \varnothing$, then $\|x_n\| \to +\infty$. ■

4 Fixed points of quasi Bregman nonexpansive mappings

In this section, we shall apply the results of Section 3 to find the Bregman nearest fixed point of a quasi Bregman nonexpansive mapping.

4.1 Quasi Bregman nonexpansive (QBNE) mappings

Let E be a nonempty closed convex subset of X. The *fixed point set* of $T : E \to X$ is defined and denoted by $\operatorname{Fix} T := \{x \in E \mid Tx = x\}$.

Definition 4.1 Let E be a nonempty closed convex subset of X, and let $T\colon E \to X$. Then T is said to be:

(i) *fixed-point closed* if, for every sequence $(x_n)_{n\in\mathbb{N}}$ in E, $\left.\begin{array}{c} x_n \to \bar{x} \\ x_n - Tx_n \to 0 \end{array}\right\} \Rightarrow \bar{x} \in \operatorname{Fix} T$.

(ii) *quasi Bregman nonexpansive (QBNE)* if $(\forall x \in \operatorname{Fix} T)(\forall y \in E)\ D(x, Ty) \leq D(x, y)$.

It is easy to see that if $T : E \to X$ is QBNE, then $\operatorname{Fix} T \subseteq \bigcap_{x\in E} H(x, Tx)$.

Fact 4.2 Let E be a nonempty closed convex subset of X, and let $T\colon E \to X$ be QBNE. Then $\operatorname{Fix} T$ is closed and convex.

Proof. Inspect the [18, proof of Lemma 15.5], or combine [4, Proposition 3.3(iv)&(vii)]. ■

4.2 Finding the Bregman nearest fixed point

When applied to a quasi Bregman nonexpansive mapping with the fixed-point closedness property, Algorithm 3.1 and Theorem 3.6 together provide an iterative method for finding the Bregman nearest fixed point.

Theorem 4.3 (trichotomy) Let E be a nonempty closed convex subset of X, let $T\colon E \to X$ be QBNE and fixed-point closed, let $x_0 \in X$, and let C_0 be a closed convex nonempty subset of X containing $\operatorname{Fix} T$. Define sequences $(C_n)_{n\in\mathbb{N}}$ and $(x_n)_{n\in\mathbb{N}}$ by

$$(\forall n \in \mathbb{N}) \quad C_{n+1} := C_n \cap H(x_n, Tx_n) \text{ and } x_{n+1} = \overleftarrow{P}_{C_{n+1}} x_0.$$

Then exactly one of the following holds:

(i) $\operatorname{Fix} T \neq \varnothing$, $x_n \to \overleftarrow{P}_{\operatorname{Fix} T} x_0$ and $\sum_{n\in\mathbb{N}} D(x_{n+1}, x_n) < +\infty$.

(ii) $\operatorname{Fix} T = \varnothing$ and $\|x_n\| \to +\infty$.

(iii) $\operatorname{Fix} T = \varnothing$ and the sequence is not well defined (i.e., $C_{n+1} = \varnothing$ for some $n \in \mathbb{N}$).

Proof. Suppose that $C = \operatorname{Fix} T$ and set $(y_n)_{n\in\mathbb{N}} = (Tx_n)_{n\in\mathbb{N}}$ when $(x_n)_{n\in\mathbb{N}}$ is well defined. In this case, it is clear that (10) holds because T is fixed-point closed.

(i): Assume that $C \neq \varnothing$. We show inductively that $(\forall n \in \mathbb{N})\ C \subseteq C_n$. Note that $C \subseteq C_0 \neq \varnothing$. Suppose that $C \subseteq C_n$ for some $n \in \mathbb{N}$. Then x_n is well defined and $C \subseteq H(x_n, Tx_n)$ because T is QBNE. Moreover, $C \subseteq C_n \cap H(x_n, Tx_n) = C_{n+1}$. Therefore $(\forall n \in \mathbb{N})\ C \subseteq C_n$, and C_n is nonempty, closed, and convex by Remark 3.2. Hence, the sequence $(x_n)_{n\in\mathbb{N}}$ is well defined. The conclusion thus follows from Theorem 3.6.

(ii)&(iii): Assume that $C = \varnothing$. If $(x_n)_{n\in\mathbb{N}}$ is not well defined, then (iii) happens. Finally, if $(x_n)_{n\in\mathbb{N}}$ is well defined, then (ii) occurs, again by Theorem 3.6. ■

4.3 Bregman subgradient projectors

Let us now show that every Bregman subgradient projector is QBNE and that it has the fixed point closedness property. We can also arrange that C is its fixed point set. This guarantees that Theorem 4.3 is applicable to Bregman subgradient projectors.

For the remainder of this paper, we assume that

$$g : X \to \mathbb{R} \text{ is a continuous and convex with } \operatorname{lev}_{\leq 0} g := \big\{x \in X \,\big|\, g(x) \leq 0\big\} \neq \varnothing,$$

and that

$$(\forall z \in X)(\forall z^* \in \partial g(z)) \quad H_g(z, z^*) := \big\{x \in X \,\big|\, g(z) + \langle z^*, x - z\rangle \leq 0\big\}.$$

The following result follows directly from the definitions.

Proposition 4.4 *Let $z \in X$ and let $z^* \in \partial g(z)$. Then the following hold:*

(i) $\operatorname{lev}_{\leq 0} g \subseteq H_g(z, z^*)$.

(ii) *$H_g(z, z^*)$ is convex, closed, and nonempty; it is a halfspace when $z^* \neq 0$.*

(iii) $z \in H_g(z, z^*) \Leftrightarrow z \in \operatorname{lev}_{\leq 0} g$.

Definition 4.5 Let $s\colon X \to X$ be a selection of ∂g, i.e., $(\forall z \in X)\ s(z) \in \partial g(z)$. The associated *(left) Bregman subgradient projector* onto $\operatorname{lev}_{\leq 0} g$ is

$$Q_s\colon X \to X\colon z \mapsto \overleftarrow{P}_{H_g(z,s(z))}(z). \tag{12}$$

The following result is known.

Lemma 4.6 (See [4, Propositions 3.3 and 3.38].) *Q_s is QBNE with* $\operatorname{Fix} Q_s = \operatorname{lev}_{\leq 0} g$.

We now show that Q_s is fixed-point closed.

Lemma 4.7 *Q_s is fixed-point closed.*

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X such that $x_n \to \bar{x}$ and

$$x_n - Q_s(x_n) \to 0. \tag{13}$$

We must show that $\bar{x} \in \operatorname{Fix} Q_s$. In view Lemma 4.6, it suffices to show that $g(\bar{x}) \leq 0$.

Set $(\forall n \in \mathbb{N})\ p_n := Q_s(x_n)$. For every $n \in \mathbb{N}$, by the definition of $Q_s(x_n)$, p_n minimizes the function

$$y \mapsto D(y, x_n) = f(y) - f(x_n) - \langle \nabla f(x_n), y - x_n\rangle$$

over the set $H_g(x_n, s(x_n)) = \big\{x \in X \,\big|\, g(x_n) + \langle s(x_n), x - x_n\rangle \leq 0\big\}$, where $s(x_n) \in \partial g(x_n)$; hence

$$g(x_n) + \langle s(x_n), p_n - x_n\rangle \leq 0. \tag{14}$$

Since $x_n \to \bar{x}$, it follows that $g(x_n) \to g(\bar{x})$ and $(s(x_n))_{n\in\mathbb{N}}$ is bounded. It therefore follows from (13) and (14) that $g(\bar{x}) \leq 0$, as required. ■

Combining Theorem 4.3, Lemma 4.6, and Lemma 4.7, we obtain the following result.

Theorem 4.8 Let $x_0 \in X$, and let C_0 be a closed convex subset of X such that $\operatorname{lev}_{\leq 0} g \subseteq C_0$. Define sequence $(x_n)_{n\in\mathbb{N}}$ and $(C_n)_{n\in\mathbb{N}}$ via

$$(\forall n \in \mathbb{N}) \quad C_{n+1} := C_n \cap H\big(x_n, Q_s(x_n)\big) \text{ and } x_{n+1} := \overleftarrow{P}_{C_{n+1}} x_0.$$

Then $x_n \to \overleftarrow{P}_{\operatorname{lev}_{\leq 0} g} x_0$ and $\sum_{n\in\mathbb{N}} D(x_{n+1}, x_n) < +\infty$.

We conclude with a few examples of Bregman subgradient projectors illustrating that this class is quite large.

Example 4.9 Suppose that $f = \frac{1}{2}\|\cdot\|^2$ and that g is differentiable on $X \smallsetminus \operatorname{lev}_{\leq 0} g$. The Bregman subgradient projector (see (12)) then turns into the classical subgradient projector

$$Q\colon X \to X\colon x \mapsto \begin{cases} x, & \text{if } g(x) \leq 0; \\ x - \frac{g(x)}{\|\nabla g(x)\|^2}\nabla g(x), & \text{otherwise.} \end{cases} \tag{15}$$

By Lemma 4.6 and 4.7, Q is QBNE and fixed-point closed. We single out two special cases:

(i) Suppose $g = \frac{1}{p} d_C^p$, where $1 \leq p < +\infty$ and $(\forall x \in X)\ d_C(x) := \min_{c\in C}\|x - c\|$. Then

$$Q = \big(1 - \tfrac{1}{p}\big)\operatorname{Id} + \tfrac{1}{p} P_C, \tag{16}$$

where $\operatorname{Id} := P_X$. Indeed, if $x \notin C$, then $\nabla g(x) = d_C^{p-2}(x)(x - P_C(x))$ and (16) follows from (15).

(ii) Suppose $g = e_h$, where $h : X \to \mathbb{R}$ is convex, lower semicontinuous, proper, with $\operatorname{lev}_{\leq 0} h \neq \varnothing$ and where e_h is the Moreau envelope of h, i.e.,

$$(\forall x \in X) \quad e_h(x) := \inf_{w\in X}\big(h(w) + \tfrac{1}{2}\|w - x\|^2\big). \tag{17}$$

Then

$$Q\colon X \to X\colon x \mapsto \begin{cases} x, & \text{if } e_h(x) \leq 0; \\ \dfrac{e_h(x) - 2h(P_h(x))}{2(e_h(x) - h(P_h(x)))}x + \dfrac{e_h(x)}{2(e_h(x) - h(P_h(x)))}P_h(x), & \text{otherwise,} \end{cases} \tag{18}$$

where $P_h(x) := \operatorname{argmin}_{w\in X}(h(w) + \frac{1}{2}\|w - x\|^2)$ denotes the proximal mapping of h. To see (18), we start by observing that $\operatorname{lev}_{\leq 0} h \subseteq \operatorname{lev}_{\leq 0} e_h \neq \varnothing$ because $\operatorname{lev}_{\leq 0} h \neq \varnothing$ and $e_h \leq h$. By e.g. [20, Theorem 2.26], P_h is single-valued and continuous, and e_h is convex and continuously differentiable with $\nabla e_h = (\operatorname{Id} - P_h)$. From (17) it follows that

$$e_h(x) = h(P_h(x)) + \frac{1}{2}\|x - P_h(x)\|^2; \tag{19}$$

thus,

$$\|x - P_h(x)\|^2 = 2(e_h(x) - h(P_h(x))). \tag{20}$$

Combining (15), (19), and (20), we obtain (18). Note that $e_h(x) \leq 0 \Rightarrow h(P_h(x)) \leq 0$.

Acknowledgments

We thank the referee for some helpful suggestions and pertinent comments. HHB was partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program. JC was partially supported by the Natural Science Foundation of China, the Doctor Fund of Southwest University and the Fundamental Research Fund for the Central Universities. XW was partially supported by the Natural Sciences and Engineering Research Council of Canada.

References

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

[2] H.H. Bauschke and J.M. Borwein, Legendre functions and the method of random Bregman projections, *J. Convex Anal.* 4 (1997), 27–67.

[3] H.H. Bauschke, J.M. Borwein, and P.L. Combettes, Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, *Commun. Contemp. Math.* 3 (2001), 615–647.

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

[5] H.H. Bauschke, J. Chen, and X. Wang, A projection method for approximating fixed points of quasi nonexpansive mappings without the usual demiclosedness condition, *J. Nonlinear Convex Anal.*, to appear.

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

[7] H.H. Bauschke and P.L. Combettes, Iterating Bregman retraction, *SIAM J. Optim.* 13 (2003), 1159–1173.

[8] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

[9] H.H. Bauschke and D. Noll, The method of forward projection, *J. Nonlinear Convex Anal.* 3 (2002), 191–205.

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

[11] D. Butnariu and A.N. Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization, Applied Optimization*, vol. 40, Kluwer Academic, Dordrecht, 2000.

[12] D. Butnariu and E. Resmerita, Bregman distances, totally convex functions, and a method for solving operator equations in Banach spaces, *Abstr. Appl. Anal.* 84919 (2006), 39 pages.

[13] Y. Censor and S. Reich, The Dykstra algorithm with Bregman projections, *Commun. Appl. Anal.* 2 (1998), 407–419.

[14] Y. Censor and S.A. Zenios, *Parallel Optimization*, Oxford University Press, New York, 1997.

[15] G. Kassay, S. Reich, and S. Sabach, Iterative methods for solving systems of variational inequalities in reflexive Banach spaces, *SIAM J. Optim.* 21 (2011), 1319–1344.

[16] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation I*, Springer, New York, 2006.

[17] S. Reich and S. Sabach, A projection method for solving nonlinear problems in reflexive Banach spaces, *J. Fixed Point Theory Appl.* 9 (2011), 101–116.

[18] S. Reich and S. Sabach, Existence and approximation of fixed points of Bregman firmly nonexpansive operators in reflexive Banach spaces, in: *Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Optimization and its Applications*, Springer, New York, 2011, pp. 301–316.

[19] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[20] R.T. Rockafellar and R.J-B Wets, *Variational Analysis*, corrected 3rd printing, Springer, Berlin, 2009.

[21] S. Sabach, Products of finitely many resolvents of maximal monotone mappings in reflexive Banach spaces, *SIAM J. Optim.* 21 (2011), 1289–1308.

[22] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing, 2002.