A Decomposition Method Using Bregman Distances to Convex Separable minimization Problems*

E. A. Papa Quiroz †and Philippe. Mahey

Universidad Nacional Mayor de San Marcos.
erikpapa@gmail.com

Laboratoire d'Informatique de Modélisation et d'Optimisation des Systèms
LIMOS- UMRL 6158-Université Blaise Pascal/CNRS
BP 10125 Aubière 63173, France
philippe.mahey@isima.fr

Abstract

In this paper we propose an extension of the proximal decomposition algorithm using Bregman distances to solve convex separable minimization problems. Under some standard assumptions it is proved that the iterations generated by the algorithm are well defined and some convergence results are obtained.

Keywords: Proximal point methods, Bregman distances, monotone vector field, Decomposition methods.

*This paper was developed during a Pos Doctoral visit of the first author to ISIMA-LIMOS Blaise Pascal University of Clermont Ferrand during the months of February to March 2012

†Corresponding author: E.A. Papa Quiroz, telephone: 51(1)5685537//Mobil Phone: 51(1)992332681

1 Introduction

2 Legendre Type Function in $I\!R^n$

Definition 2.1 *A poper lower semicontinuous convex function $h : I\!R^n \to (-\infty, +\infty]$ is called essentially smooth if h satisfies:*

1. *$int(\texttt{dom}\, h) \neq \emptyset$*
2. *h is differentiable on $int(\texttt{dom}\, h)$*
3. *$||\nabla h(x_k)|| \to +\infty$, for all $\{x_k\} \subset int(\texttt{dom}$ such that $x_k \to \bar{x}$, for some $\bar{x} \in \texttt{Bound}(\texttt{dom}\, h)$.*

Remark 2.1 (Rock, theorem 26.1) *h is essentially smooth if and only if*

$$\partial h(x) = \begin{cases} \{\nabla h(x)\}, & x \in int(\texttt{dom}\, h) \\ \emptyset, & otherwise \end{cases}$$

In particular $\texttt{dom}\, \partial h = int(\texttt{dom}\, h)$.

Definition 2.2 *A poper lower semicontinuous convex function $h : I\!R^n \to (-\infty, +\infty]$ is called strictly essentially smooth if:*

1. *h is essentially smooth*
2. *h is striclty convex on every convex subset of $\texttt{dom}\, \partial h$.*

Remark 2.2 (Rock, Section 26) *h is essentially smooth if and only if h^* is essentially strictly convex*

Definition 2.3 *A poper lower semicontinuous convex function $h : I\!R^n \to (-\infty, +\infty]$ is called of Legendre type, denoted by $h \in \mathcal{L}$, if it is essentially smooth and essentially strictly convex.*

Remark 2.3 *If h is of Legendre type, then*

$$(\nabla h)^{-1} = \nabla h^*.$$

Let $h \in \mathcal{L}$. Define the funcion $D_h(.,.) : \bar{S} \times S \to I\!R$ so that

$$D_h(x, y) := h(x) - h(y) - \langle \nabla h(y), \exp_y^{-1} x \rangle_y. \tag{2.1}$$

Definition 2.4 *Let $\Omega \subset I\!R^n$, and let $y \in I\!R^n$. A point $Py \in \Omega$ for which*

$$D_h(Py, y) = \min_{x \in \Omega} D_h(x, y) \tag{2.2}$$

is called a D_h-projection of the point y on the set Ω.
This point also is denoted by

$$Py = \arg\min\{D_h(u, y) : u \in \Omega\}$$

The next Lemma furnishes the existence and uniqueness of D_h-projection for a type Legendre function under an appropriate assumption on Ω.

Lemma 2.1 *Let $\Omega \subset H$ a nonempty closed convex set such that $int(\mathbf{dom}\, h) \cap \mathbf{ir}(\Omega) \neq \emptyset$ and $h \in \mathcal{L}$. Then, for any $y \in H$, there exists a unique D_h-projection Py of the point y on Ω satisfying*

i. $Py \in int(\mathbf{dom}\, h)$.

ii. $\langle \nabla h(y) - \nabla h(Py), c - Py \rangle \leq D_h(c, Py), \forall c \in C$.

Proof. *C'est facil.*

Now, we define the proximal normal cone $N_\Omega^P(x)$ of Ω at $x \in \Omega$.

$$N_\Omega^P(x) = \{t(y - x) : t \geq 0, x \in P_\Omega(y), y \in I\!R^L\}$$

where P_Ω denotes the usual projection on Ω and each vector is called a proximal normal to Ω at x

Proposition 2.1 *Suppose that $h \in \mathcal{L}$ is twice continuously differentiable on $S = \mathbf{intdom}\, h$, let $y \in S$ and suppose that $x = P_{D_h}(y)$ on Ω, then*

$$\nabla h(y) - \nabla h(x) \in N_\Omega^P(x).$$

3 A Proximal Decomposition Method with Bregman Distances

Consider the problem:

$$\text{find } (x, y) \in A \times B \text{ such that } y \in T(x) \tag{3.3}$$

Along this section we use the notation $S = int(\mathtt{dom}\ h)$

Definition 3.1 *Let T be a monotone multivalued operator on $I\!R^n$ and $(x, y) \in S \times I\!R^n$. The proximal decomposition with the factor $\lambda > 0$ of (x, y) on the graph of T and the function $h \in \mathcal{L}$, is the unique pair $(u, v) \in S \times I\!R^n$ such that*

$$\nabla h(x) + \lambda y = \nabla h(u) + \lambda v; v \in T(u)$$

Remark 3.1 *The above definition is well given. In fact, suppose that there exists another $(u', v') \in Gr(T)$ satisfying the condition of the definition, from the convexity of h, the above definition and the monotonicity of T we have that*

$$0 \leq \langle \nabla h(u) - \nabla h(u'), u - u' \rangle = \langle v' - v, u - u' \rangle \leq 0$$

Now, from the strictly convexity of h we obtain that $u' = u$ and using this fact in the definition we obtain that $v' = v$.

Remark 3.2 *A simple manipulation of the definition provides the following expression:*

$$u = (\nabla h + \lambda T)^{-1} (\nabla h(x) + \lambda y)$$

$$v = \left(I + \frac{1}{\lambda}\nabla h \circ T^{-1}\right)^{-1} \left(\frac{1}{\lambda}\nabla h(x) + \lambda y\right)$$

PDB Algorithm

Let $h \in \mathcal{L}$ and let D_h be the function associate to h and defined by (2.1).

Initialization:
Let $\lambda > 0$ and

$$(x^0, y^0) \in (S \cap A) \times B \tag{3.4}$$

Main Steps:
For $k = 1, 2, 3, ...,$

$$u^k = (\nabla h + \lambda T)^{-1} \left(\nabla h(x^k) + \lambda y^k\right) \tag{3.5}$$

$$v^k = -\frac{1}{\lambda}\nabla_1 D_h(u^k, x^k) + y^k \tag{3.6}$$

If $(u^k, v^k) \in A \times B$ then stop.
Otherwise,

$$x^{k+1} = P_{D_h}(u^k) \text{ and } y^{k+1} = v^k_B \tag{3.7}$$

Take $k \leftarrow k + 1$.

Remark 3.3 *$P_{D_h}(u^k)$ is the projection of u^k on A with respect to the Bregman distance D_h, that is,*

$$x^{k+1} = \arg\min\{D_h(x, u^k) : x \in A\}$$

Remark 3.4 *If $h(x) = \frac{1}{2}||x||^2$ then*

$$u^k = (I - \lambda T)^{-1}(x^k + \lambda y^k)$$

$$v^k = \frac{1}{\lambda}(x^k + \lambda y^k - u^k)$$

$$x^{k+1} = u^k_A, y^{k+1} = v^k_B$$

which is the proximal decomposition method.

Remark 3.5 *If $T = \partial f$ where f is a proper, lsc and convex function then the problem (3.3) becomes in*

$$\min\{f(x) : x \in A\}$$

and the algorithm is:

Algorithm for Minimization

Let $h \in \mathcal{L}$ with $S = \texttt{intdom}\, h$, as defined in Section 2, and let D_h be the function associate to h and defined by (2.1).

Initialization:
Let $\lambda > 0$ *and*

$$(x^0, y^0) \in (S \cap A) \times B$$

Main Steps:
For $k = 1, 2, 3, ...,$

$$u^k = \arg\min\{f(u) - \langle u, y^k\rangle + \frac{1}{\lambda} D_h(u, x^k)\} \tag{3.8}$$

$$v^k = -\frac{1}{\lambda}\nabla_1 D_h(u^k, x^k) + y^k$$

If $(u^k, v^k) \in A \times B$ *then stop.*
Otherwise,

$$x^{k+1} = P_{D_h}(u^k) \text{ and } y^{k+1} = v_B{}^k$$

Take $k \leftarrow k + 1$.

4 Some Preliminar Results

Alog this section we have the following assumptions:

Assumption 1: T is a multivalued monotone operator satisfying $\texttt{dom}T \cap A \cap S \neq \emptyset$

Lemma 4.1 *Suppose that* $\lambda > 0$ *and* $h \in \mathcal{L}$,. *Then under assumption* 1, *the mapping* $(\nabla h + \lambda T)^{-1}$ *is single valued.*

Proof. [Eckstein, 1993] Since that h **is strictly convex on** S we have for $x \neq y$

$$\langle \nabla h(x) - \nabla h(y), x - y\rangle > 0$$

As T is monotone then $\nabla h + \lambda T$ is strictly monotone and therefore $(\nabla h + \lambda T)(x)$ and $(\nabla h + \lambda T)(y)$ do not intersect to $x \neq y$. Therefore $(\nabla h + \lambda T)^{-1}$ is single valued. ∎

Theorem 4.1 *Under assumption 1, and* $h \in \mathcal{L}$, *the sequence* (x^k, y^k) *generated by the (PBD) algorithm is well definided for each* k *and* $(x^k, y^k) \in (A \cap S) \times B$.

Proof. We proced by induction. It holds for $l = 0$, due to (3.4). Assume that $(x^l, y^l) \in (A \cap S) \times B$. As $h \in \mathcal{L}$ from Lemma 4.1 there exist (u^l, v^l) satisfying (3.5) and (3.8) respectively. As $h \in \mathcal{L}$ the set

$$\arg\min\{D_h(x, u^k) : x \in A\}$$

is nonempty and has an unique element in $A \cap S$, thus there exists $x^{l+1} \in A \cap S$ such that

$$x^{l+1} = arg\min\{D_h(x, u^k) : x \in A\},$$

On the other hand, always there exists y^{l+1} :

$$y^{l+1} = v_B^l.$$

is well defined.
On the other hand, Therefore we obtain the aimed resultd. ∎

Remark 4.1 *Consider $\lambda = 1$ and define*

$$L : H \times H \to H \text{ such that } L(x, y) = \nabla h(x) + y,$$

$$F : H \to H \times H : F(z) = F(\nabla h(x) + y) = (u, v)$$

where $\nabla h(x) + y = \nabla h(u) + v, v \in T(u)$,

$$P : H \times H \to A \times B : P(x, y) = (P_{D_h}(x), y_B).$$

Then, one iteration of the algoritm may be view as the application of the operator

$$Z = P \circ F \circ L$$

on $A \times B$.

Lemma 4.2 *All fixed point of the operator $Z = P \circ F \circ L$ is a solution of the problem (3.3).*

Proof. Let $(x, y) \in \texttt{dom Z}$ such that:

$$(x, y) = Z(x, y).$$

This implies that

$$(x, y) = P(u, v) = (P_{D_h}(u), v_B) \tag{4.9}$$

where

$$\nabla h(x) + y = \nabla h(u) + v, v \in T(u) \tag{4.10}$$

Thus, from (4.10)

$$(\nabla h(u) - \nabla h(x), v - y) \in \{(a, b) : a + b = 0\} \tag{4.11}$$

On the other hand, from (4.9) we obtain that

$$v = v_A + v_B = v_A + y$$

which implies that

$$v - y \in A \tag{4.12}$$

Also, from (4.9) we will prove that

$$\nabla h(u) - \nabla h(x) \in B \tag{4.13}$$

Of fact, as $x = P_{D_h} u$, on A then from Proposition 2.1 then

$$\nabla h(u) - \nabla h(x) \in N_A^P(x).$$

So, there exists $y \in I\!R^n$ and $t \geq 0$ such that $x = y_A$ and

$$\nabla h(u) - \nabla h(x) = t(y - x) = t(y_A + y_B - x) = t y_B \in B. \tag{4.14}$$

From (4.13) and (4.12) we obtain that

$$(\nabla h(u) - \nabla h(x), v - y) \in B \times A \tag{4.15}$$

which jointly with (4.11) gives

$$\nabla h(x) = \nabla h(u) \text{ and } u = v$$

as $h \in \mathcal{L}$ this implies that

$$x = u \text{ and } u = v$$

References

[1] Auslender A., Teboulle M., Entropic Proximal Decomposition Methods for Convex Programs and Variational Inequalities, Math. Program., Ser A, 91, 33-47, 2001.

[2] Auslender A., Teboulle M., Interior gradient and proximal methods for convex and conic optimization. SIAM J. Optim., 16, 3: 697-725, 2006.

[3] Chen, G., and Teboulle, M., A proximal-Based Decomposition Method for Convex Minimization Problems, Mathematical Programming, Serie A, 1, Vol. 64, 81-101, 1994.

[4] Kyono M., and Fukushima M., Nonlinear Proximal Decomposition Method for Convex Programming, JOTA, Vol 106, N2, 357-372, 2000.

[5] Mahey P., Proximal Decomposition of the Graph of a Maximal Monotone Operator, SIAM J. Optim. 5, 2, 454-466, 1995

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