

WWW.MATHEMATICSWEB.ORG

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

J. Math. Anal. Appl. 280 (2003) 313-320

www.elsevier.com/locate/jmaa

A note on the existence of zeroes of convexly regularized sums of maximal monotone operators

Regina S. Burachik,^{a,1,3} Susana Scheimberg,^{a,2} and Paulo J.S. Silva^{b,*,3}

^a Engenharia de Sistemas e Computação, COPPE-UFRJ, CP 68511, CEP 21945-970, Rio de Janeiro, RJ, Brazil

^b Computer Science Department, IME-USP, Rua do Matão 1010, CEP 05508-090, São Paulo, SP, Brazil

Received 14 May 2002

Submitted by J. McCarthy

Abstract

Many algorithms for solving the problem of finding zeroes of a sum of two maximal monotone operators T_1 and T_2 , have regularized subproblems of the kind $0 \in T_1(x) + T_2(x) + \partial D(x)$, where D is a convex function. We develop an unified analysis for existence of solutions of these subproblems, through the introduction of the concept of convex regularization, which includes several well-known cases in the literature. Finally, we establish conditions, either on D or on the operators, which assure solvability of the subproblems.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Zeros of sums of maximal monotone operators; Legendre functions; Generalized proximal-like algorithms; Regularization

* Corresponding author.

E-mail addresses: regi@cos.ufrj.br (R.S. Burachik), susana@cos.ufrj.br (S. Scheimberg), rsilva@ime.usp.br (P.J.S. Silva).

¹ Research supported by CNPq Grant 301280/94-0(RE).

² Research supported by CNPq Grant 302393/85-4(RN).

³ Research supported by PRONEX-Optimization.

⁰⁰²²⁻²⁴⁷X/03/\$ – see front matter @ 2003 Elsevier Science (USA). All rights reserved. doi:10.1016/S0022-247X(03)00043-X

1. Introduction

Let *X* be a reflexive real Banach space, with $\langle u, x \rangle$ written in place of u(x), for $x \in X$ and $u \in X^*$. Moreover, let $J: X \rightrightarrows X^*$ be the normalized duality mapping, defined by the property $v \in Jx$ if and only if $||v||_{X^*}^2 = ||x||_X^2 = \langle v, x \rangle$ for any $x \in X$, where $|| \cdot ||_X$ denotes the norm in the space *X*.

A multi-valued mapping $T: X \rightrightarrows X^*$ is said to be monotone if

 $\langle u - v, x - y \rangle \ge 0$ whenever $u \in T(x), v \in T(y)$.

Moreover, it is said to be maximal if its graph is not properly contained in the graph of any other monotone mapping.

Given two maximal monotone operators $T_1, T_2: X \rightrightarrows X^*$, a fundamental problem is the one of finding a zero of $T_1 + T_2$, i.e., the problem of finding an $x \in X$ such that

$$0 \in T_1(x) + T_2(x).$$
(1)

A wide variety of problems can be regarded as special instances of (1). To name a few, linear and convex programming, solving systems of linear equations or inequalities, systems of partial differential equations, finding a point in the intersection of two convex sets, monotone complementarity, variational inequalities, and constrained minimax. When (1) is a model for a variational inequality problem, then T_2 is the normal cone N_C of a closed convex set *C*. Several algorithms for solving this problem have subproblems of the form

$$0 \in T_1(x) + N_C(x) + \partial D(x), \tag{2}$$

where *D* is a regularization associated to *C* [2,3,7,8,10,11,17,19]. Hence general conditions under which subproblems (2) have solutions are important from the algorithmic point of view.

More generally, our aim is to study the existence of solutions to the following problem:

$$0 \in T_1(x) + T_2(x) + \partial D(x), \tag{3}$$

where D is a regularization. For an example in which T_2 is not the normal cone N_C , see, e.g., [1, Section 3] and [5, Section 3].

2. Convex regularizations

We define below the regularizations that will be considered in problem (3). In order to do so, we recall a recent generalization of the concept of a Legendre function [15, Section 26] well suited for reflexive Banach spaces.

Definition 1 [4, Definition 5.2]. We say that a proper convex lower semicontinuous function $f: X \to (-\infty, +\infty]$ is

(a) essentially smooth, if ∂f is both locally bounded and single valued on its domain;

- (b) essentially strictly convex, if ∂f⁻¹ is locally bounded on its domain and f is strictly convex on every convex sub-set of dom ∂f. Note that the local boundedness of ∂f⁻¹ is equivalent to rge ∂f (= dom ∂f⁻¹) be open [4, Corollary 2.19];
- (c) Legendre, if it is both essentially smooth and essentially strictly convex.

We can now introduce the concept of a convex regularization.

Definition 2. Let *C* be subset of *X* with nonempty and convex interior. Let $D: X \rightarrow (-\infty, +\infty]$ be a proper, convex and lower semi-continuous function. We will say that *D* is a *convex regularization associated to C* when it satisfies the following conditions:

- (a) int dom $D = \operatorname{int} C = \operatorname{dom} \partial D$,
- (b) D attains its minimum on int C,
- (c) *D* is a Legendre function.

Some examples are:

• $X = \mathbb{R}^n$ and $C = \mathbb{R}^n_+$. Any φ -divergence [18] with center in $y \in \mathbb{R}^n_{++}$,

$$D(x) = \sum_{j=1}^{n} y_i \varphi\left(\frac{x_i}{y_i}\right),$$

is a convex regularization associated to \mathbb{R}^n_+ . In particular, the Kullback–Leibler relative entropy $D(x) = \sum_{j=1}^n (x_j \log(x_j/y_i) + y_j - x_j)$. Another related example is the recently introduced family of distances based in an second-order homogeneous kernels. They are defined as $D_{\theta}(x) = d_{\theta}(x, y) = \sum_{j=1}^n y_j^2 \theta(x_j/y_j)$, where $\theta(\cdot)$ is a function with logarithmic-quadratic behavior [2, Section 2]. The same holds for their generalizations presented in [17].

• *X* is a reflexive smooth and rotund space, so that $(1/2)||x||^2$ is a Legendre function. Consider *C* a subset of *X* such that int $C = \{x \in X : ||x|| < 1\}$. The functions $D_1(x) = -\sqrt{1 - ||x||^2}$ if $||x|| \le 1$ and $+\infty$ otherwise, and $D_2(x) = (1 - ||x||^2)^{-1}$ if ||x|| < 1 and $+\infty$ otherwise, are convex regularizations associated to *C*. Actually, in [4] it is proved that these are Legendre functions. They achieve their minimum at x = 0. We point out that in [12], the function $1 + D_1(x)$ is considered as a penalty function in the closed unit ball of center zero *B*. This last work deals with the variational inequality problem *VIP*(*T*, *B*) on an uniformly convex and uniformly smooth Banach space.

Another nontrivial and important example is given in Proposition 6.

In view of (3) and condition (a) in Definition 2, it will be natural to consider $C = \text{dom } T_2$. In this case, it should be clear that any solution to (3) must lie in dom $T_1 \cap \text{int dom } T_2$. Hence, as our objective is to study the existence of such solutions, we will assume throughout this paper that

$$\operatorname{dom} T_1 \cap \operatorname{int} \operatorname{dom}(T_2) \neq \emptyset. \tag{4}$$

This assumption also ensures that $T_1 + T_2$ is maximal monotone [16, Theorem 1]. Note that, by strict convexity of D, if (3) has a zero, it must be unique.

3. Existence of solutions for special convex regularizations

In this section, we will show some reasonable extra assumptions on D, the convex regularization associated to int dom T_2 that can ensure the existence of solutions to (3).

We start by recalling two auxiliary results:

Lemma 3 [14, Lemma 2.1]. Let T be a maximal monotone operator and $F \subset X^*$ such that

$$\forall u \in F, \ \exists y \in X, \quad \sup_{(z,v) \in \text{gph}T} \langle v - u, y - z \rangle < \infty, \tag{5}$$

then conv $F \subset \overline{\operatorname{rge} T}$ *and* $\operatorname{int}(\operatorname{conv} F) \subset \operatorname{rge} T$.

Lemma 4 [8, Lemma 2.7]. Let *T*, *S* be monotone operators. Suppose that they satisfy the following conditions:

(1) *S* is regular, i.e., $\forall u \in \text{rge } S$, $y \in \text{dom } S$,

$$\sup_{(z,v)\in \mathrm{gph}\,S}\langle v-u,\,y-z\rangle<\infty;$$

(2) dom $T \cap$ dom $S \neq \emptyset$ and rge $S = X^*$;

(3) T + S is maximal monotone.

Then, $rge(T + S) = X^*$.

Now, it is easy to show a generalization of [7, Theorem 1] and part (2) of [2, Proposition 2]. This is basically [8, Corollary 3.1] presented in a general setting, not limited to Bregman distances [8]:

Theorem 5. If rge $\partial D = X^*$, then (3) admits a unique solution.

Proof. Since we assumed that $\operatorname{rge} \partial D = X^*$, we simply use Lemma 4 with $S := \partial D$ and $T := T_1 + T_2$, remembering that ∂D is regular [5, Example 1], $(T_1 + T_2) + \partial D$ is maximal monotone [16, Theorem 1] and that assumption (4) holds. \Box

Before we state our next result, we show that an important and well-known kind of regularization is a convex regularization.

Proposition 6. Let $f: X \to (-\infty, +\infty]$ be a proper closed strictly convex function which has \overline{C} as effective domain. Assume that f is differentiable on int C. For $y \in \text{int } C$, define

$$\mathcal{D}_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle \quad \text{for all } x \in C,$$

a Bregman-like distance. Then $\mathcal{D}_f(\cdot, y)$ is a convex regularization associated to C if and only if f is Legendre.

Proof. Assume that $\mathcal{D}_f(\cdot, y)$ is a convex regularization associated to *C*, then by using Definition 1, we conclude that *f* must be Legendre, since $\mathcal{D}_f(\cdot, y)$ and *f* differ only

by an affine term. Conversely, assume that f is Legendre. Since part (c) of Definition 1 trivially holds, we only have to check conditions (a) and (b). Note that, int $C = \operatorname{int} \operatorname{dom} f = \operatorname{dom} \partial \mathcal{D}_f(\cdot, y)$, where the second equality holds because f is Legendre. Then condition (a) holds. Condition (b) is verified because the minimum of $\mathcal{D}_f(\cdot, y)$ is attained at y. \Box

Corollary 7. Assume that the hypotheses of Proposition 6 hold. The Bregman-like distance $\mathcal{D}_f(\cdot, y)$ is a convex regularization associated to *C* if

(a) rge ∂f is open, and

(b) \mathcal{D}_f is boundary coercive [9, Assumption B6].

Condition (a) above holds if X is finite dimensional and both conditions are valid whenever f is zone coercive [8].

Proof. By Proposition 6, it is enough to prove that f is Legendre.

Assume that (a) and (b) hold. Clearly, (a) and the strict convexity of f states that it is essentially strictly convex.

Moreover, let $\{y^k\}$ be a sequence in int *C* converging to some point in the boundary of \overline{C} . As \mathcal{D}_f is boundary coercive we know that for every $x \in \text{int } C$

 $\langle \nabla f(y^k), x - y^k \rangle \to -\infty.$

It follows that $\|\nabla f(y^k)\| \to +\infty$. Using [4, Theorem 5.6] we conclude that f is essentially smooth.

Let us prove the last assertion. Assume that X is finite dimensional and f is strictly convex. By [15, Theorem 26.3], f^* is essentially smooth. Using [15, Theorem 26.1] we conclude that rge $\partial f = \operatorname{dom} \partial f^*$ is open.

Finally, if f is zone coercive, this means that rge $\partial f = X^*$, which is open. For checking (b), we show first that dom $\partial f = \operatorname{int} C$. Otherwise there would be two points, one in the boundary of \overline{C} and the other int C, sharing a sub-gradient, which would contradict the strict convexity of f [4, Lemma 5.1(ii)]. Hence, dom $\partial f = \operatorname{int} C = \operatorname{int} \operatorname{dom} f$ and ∂f is single valued on its domain. Now, take $x \in \operatorname{int} C$ and let $\{y^k\}$ be a sequence in int C converging to \overline{y} in the boundary of \overline{C} . Using [4, Theorem 5.6 (iii) and (v)], we conclude that $\|\nabla f(y^k)\| \to +\infty$. Using the Fenchel equality,

$$\left\langle \nabla f(y^k), x - y^k \right\rangle = \left\langle \nabla f(y^k), x \right\rangle - f^* \left(\nabla f(y^k) \right) - f(y^k).$$

We observe that $f^*(\cdot) - \langle \cdot, x \rangle$ is coercive, as $x \in \text{int dom } f$ [4, Fact 3.1]. Taking limits above it follows that

$$\lim_{k \to +\infty} \left\langle \nabla f(x^k), x - y^k \right\rangle \leqslant -\infty - f(\bar{y}) = -\infty. \quad \Box$$

In the light of the previous lemma and the examples presented in Section 2, the next theorem is specially interesting. Actually, it generalizes several results that proved the existence of solutions to proximal subproblems in the context of variational inequalities. In particular, we drop the standard requirement of zone coerciveness for Bregman distances

and replace it by the weaker conditions (a) and (b) from Corollary 7 above. We highlight that condition (a) is void in finite dimension, and condition (b), boundary coerciveness, is rather natural for interior point methods. The following theorem generalizes [11, Theorem 4(i)], [7, Theorems 1 and 2], and [8, Corollary 3.1]. Another result, not based on Bregman functions, that is extended below is [3, Theorem A.1], whose proof inspired our approach.

Theorem 8. Let
$$0 \in \overline{\operatorname{rge}(T_1 + T_2)}$$
 and, for all $y \in \operatorname{int} \operatorname{dom} T_2$ and all $w \in \operatorname{dom} T_2$,

$$\sup_{z \in \operatorname{int} \operatorname{dom} T_2} \left\{ \langle \nabla D(z) - \nabla D(y), w - z \rangle \right\} < \infty.$$
(6)

Then, (3) admits a unique solution. Moreover, if D is finite on dom T_2 , (6) holds.

Proof. The proof is divided into four steps:

(1) rge ∂D is an open neighborhood of 0.

It holds that $0 \in \text{rge} \partial D$, as we assumed that D attains its minimum. Moreover, since D is essentially strictly convex and X is a reflexive Banach space, D^* is essentially smooth. Hence dom $\partial D^* = \text{rge} \partial D$ is open and nonempty.

(2) $0 \in \operatorname{int}(\operatorname{rge}(T_1 + T_2) + \operatorname{rge} \partial D).$

Since $0 \in \text{int} \operatorname{rge} \partial D$, there is an $\varepsilon > 0$ such that $B(0, \varepsilon) \subset \operatorname{rge} \partial D$. Using also that $0 \in \overline{\operatorname{rge}(T_1 + T_2)}$, we have that $B(0, \varepsilon/2) \subset \operatorname{rge}(T_1 + T_2) + \operatorname{rge} \partial D$.

(3) $0 \in \operatorname{intrge}(T_1 + T_2 + \partial D)$.

Let us apply Lemma 3 with $F \doteq \operatorname{rge}(T_1 + T_2) + \operatorname{rge} \partial D$ and $T \doteq T_1 + T_2 + \partial D$. Since $\operatorname{dom}(T_1 + T_2 + \partial D) = \operatorname{dom} T_1 \cap \operatorname{int} \operatorname{dom} T_2$, we need to show that

$$\forall u \in \operatorname{rge}(T_1 + T_2) + \operatorname{rge} \partial D, \ \exists y \in X$$

 $\sup\{\langle v-u, y-z\rangle \mid z \in \operatorname{dom} T_1 \cap \operatorname{int} \operatorname{dom} T_2, \ v \in (T_1+T_2+\nabla D)(z)\} < \infty.$ (7)

For any $u \in \operatorname{rge}(T_1 + T_2) + \operatorname{rge} \partial D$, let $u_0 \in X^*$, $y \in \operatorname{dom} T_1 \cap \operatorname{dom} T_2$ and $\tilde{y} \in \operatorname{int} \operatorname{dom} T_2$ be such that

 $u = u_0 + \nabla D(\tilde{y}), \quad u_0 \in (T_1 + T_2)(y).$

For $z \in \text{dom } T_1 \cap \text{int dom } T_2$ and $v \in (T_1 + T_2 + \nabla D)(z)$, let $v_0 \in (T_1 + T_2)(z)$ be such that $v = v_0 + \nabla D(z)$. Then,

$$\langle v - u, y - z \rangle = \langle v_0 + \nabla D(z) - u_0 - \nabla D(\tilde{y}), y - z \rangle$$

= $\langle v_0 - u_0, y - z \rangle + \langle \nabla D(z) - \nabla D(\tilde{y}), y - z \rangle$
 $\leq \langle \nabla D(z) - \nabla D(\tilde{y}), y - z \rangle \quad (T_1 + T_2 \text{ is monotone}).$

Taking the supremum for all $z \in \text{dom } T_1 \cap \text{int dom } T_2$ and $v \in (T_1 + T_2 + \nabla D)(z)$ above, we conclude that (6) implies (7).

Therefore Lemma 3 states that int conv(rge($T_1 + T_2$) + rge ∂D) \subset int rge($T_1 + T_2 + \partial D$). Finally, using Step 2 above, we learn that $0 \in$ int rge($T_1 + T_2 + \partial D$). This establishes the first assertion of the thesis.

(4) Property (6) holds if D is finite on dom T_2 .

Let $w \in \text{dom } T_2$ and $y, z \in \text{int dom } T_2$, then

$$\begin{split} \langle \nabla D(z) - \nabla D(y), w - z \rangle &\leq D(w) - D(z) - \langle \nabla D(y), w - z \rangle \\ &= D(w) - D(z) - \langle \nabla D(y), w \rangle + \langle \nabla D(y), z \rangle \\ &\leq D(w) - D(z) - \langle \nabla D(y), w \rangle + D(z) - D(y) + \langle \nabla D(y), y \rangle \\ &= D(w) - D(y) + \langle \nabla D(y), y - w \rangle, \end{split}$$

where we used the gradient inequality. Since the right-hand side does not depend on z, the supremum in (6) is bounded above. \Box

4. Existence of solutions for special problems

In this section, we change the extra assumptions on D for the following assumption on problem (1):

$$h(x) := \sup\left\{ \langle v, x - y \rangle \mid y \in \operatorname{dom} T_2, \ v \in (T_1 + T_2)(y) \right\} < \infty$$
(8)

for all $x \in \text{dom } T_1 \cap \text{dom } T_2$. This inequality was studied in [6, Chapter 3] and [7, Theorem 2] in the context of variational inequalities in Hilbert spaces. Here we use it in the context of reflexive Banach spaces for problem (1). We should stress that (8) holds whenever $T_1 = \partial f$ for some convex function bounded below and $T_2 = N_C$ for a nonempty, convex, and closed set *C*. Therefore, the next theorem also extends [19, Lemma 3.1], [10, Proposition 4.1], and [17, Lemma 3.2]. For other conditions that ensure (8) see [6, Proposition 3.1].

Theorem 9. Assume that D is a convex regularization associated to dom T_2 . If (8) holds, then (3) admits a unique solution.

Proof. If dom $(T_1 + T_2 + \partial D)$ is bounded, then by [13, Theorem 4.1], $T_1 + T_2 + \partial D$ is onto, and hence (3) has a solution. Otherwise, let \bar{x} be the point in int dom T_2 where D attains its minimum and let $\alpha = |h(\bar{x})|$. As the level sets of D are bounded, there must be $\delta > 0$ such that

 $x\in X, \quad \|x-\bar{x}\|>\delta \quad \Rightarrow \quad D(x)-D(\bar{x})>\alpha.$

Since dom $(T_1 + T_2 + \partial D)$ is unbounded, there exists $x \in \text{dom}(T_1 + T_2 + \partial D) = \text{dom } T_1 \cap \text{int dom } T_2$, such that $||x|| > ||\bar{x}|| + \delta$. For such x, take $v \in (T_1 + T_1 + \partial D)(x)$, which may be written in the form $v_0 + \nabla D(x)$ for some $v_0 \in (T_1 + T_2)(x)$. Note that x verifies $||x - \bar{x}|| > \delta$, hence,

$$\langle v, x - \bar{x} \rangle = \langle v_0, x - \bar{x} \rangle + \left\langle \nabla D(x), x - \bar{x} \right\rangle \ge -\alpha + \left\langle \nabla D(x), x - \bar{x} \right\rangle \ge -\alpha + \alpha = 0$$

It follows from [16, Theorem 5(c)] applied to the zeroes of $(T_1 + T_2 + \partial D)$, that (3) has a solution. \Box

References

- A. Auslender, M. Teboulle, Lagrangian duality related multiplier methods for variational inequality problems, SIAM J. Optim. 10 (2000) 1097–1115.
- [2] A. Auslender, M. Teboulle, S. Ben-Tiba, A logarithmic–quadratic proximal method for variational inequalities, Comput. Optim. Appl. 12 (1999) 31–40.
- [3] A. Auslender, M. Haddou, An interior-proximal method for convexly linearly constrained problems an its extension to variational inequalities, Math. Programming 71 (1995) 77–100.
- [4] H.H. Bauschke, J.M. Borwein, P.L. Combettes, Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, preprint.
- [5] H. Brézis, A. Haraux, Image d'une somme d'operateurs monotones et applications, Israel J. Math. 23 (1976) 165–186.
- [6] R.S. Burachik, Generalized proximal point methods for the variational inequality problem, Ph.D. thesis, IMPA, Rio de Janeiro, Brazil, 1995.
- [7] R.S. Burachik, 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, S. Scheimberg, A proximal point method for the variational inequality problem in Banach spaces, SIAM J. Control Optim. 39 (2000) 1633–1649.
- [9] A.N. Iusem, Some properties of generalized proximal point methods for quadratic and linear programming, J. Optim. Theory Appl. 85 (1995) 593–612.
- [10] A.N. Iusem, B.F. Svaiter, M. Teboulle, Entropy-like proximal methods in convex programming, Math. Oper. Res. 19 (1994) 790–814.
- [11] J. Eckstein, Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, Math. Oper. Res. 18 (1993) 202–226.
- [12] R.G. Otero, Inexact versions of proximal point and cone-constrained augmented Lagrangians in Banach spaces, Ph.D. thesis, IMPA, Rio de Janeiro, Brazil, 2001.
- [13] D. Pascali, S. Sburlan, Nonlinear Mappings of Monotone Type, Editura Academiei, Bucharest, 1978.
- [14] S. Reich, The range of sums of accretive and monotone operators, J. Math. Anal. Appl. 68 (1979) 310–317.[15] R.T. Rockafellar, Convex Analysis, Princeton Univ. Press, 1970.
- [16] R.T. Rockafellar, On the maximality of sums of nonlinear monotone operators, Trans. Amer. Math. Soc. 149 (1970) 75–88.
- [17] P.J.S. Silva, J. Ecktein, C. Humes Jr., Rescaling and stepsize selection in proximal method using generalized distances. SIAM J. Optim. 12 (2001) 238–261.
- [18] M. Teboulle, Entropic proximal mappings with applications to nonlinear programming, Math. Oper. Res. 17 (1992) 670–690.
- [19] M. Teboulle, Convergence of proximal-like algorithms, SIAM J. Optim. 7 (1997) 1069-1083.

Further reading

[1] R.T. Rockafellar, Local boundedness of nonlinear monotone operators, Michigan Math. J. 16 (1969) 397-407.