AMO - Advanced Modeling and Optimization, Volume 13, Number 2, 2011

A Duality Algorithm for Solving General Variational Inclusions [1]

Abdellatif Moudafi[2]

Abstract. To solve general variational inequalities considered in Robinson [9], we propose a generalized version of an algorithm introduced by Bermudez and Moreno [3]. Our results extend, improve and develop some known results in this field.

AMS subject classification: Primary, 49J53, 65K10; Secondary, 49M37, 90C25.

Keywords. Variational inequalities, proximal algorithm, splitting methods.

1 Introduction and preliminaries

Recently in [2], J.-F. Aujol considered an algorithm introduced by Bermudez and Moreno based on the celebrated work of Rockafellar [10] on the Proximal Point Algorithm and apply it to many image processing problems. The general minimization problem he considered is the following:

$$\inf_{z\in V}\{\frac{1}{2}\langle Az,z\rangle-\langle g,z\rangle+\phi\circ B^*(z)\} \tag{1.1}$$

with V and E two Hilbert spaces, $\phi : E \to I\!R$ a proper convex lower semi continuous function, $B : E \to V$ a bounded linear operator, $B^* : V \to E$ the adjoint of B and $A : V \to V$ a linear symmetric coercive operator.
Problem (1.1) is related, see [5], to the subdifferential inclusion

$$g \in A(u) + B\partial(\phi(B^*u). \tag{1.2}$$

Based on the work by Bermudez and Moreno [3] and motivated by its potential applications in constrained boundary value problems [1], [3] and [7], transportation network [6] and total variation based image restoration [2], we will consider

[1]The work is supported by Centre d'étude et de Recherche en Economie, Gestion, Modélisation et Informatique Appliquée.

[2]Université des Antilles et de Guyane, Ceregmia-DSI 97200 Schoelcher, Martinique, France. E-mail: abdellatif.moudafi@martinique.univ-ag.fr.

AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

the general case where A is no longer a linear operator and we replace the subdifferential by a general maximal monotone operator. More precisely, given two maximal monotone operators A, T and a bounded linear operator $B : E \to V$, our interest is in finding a point that solves the general variational inclusion

$$g \in A(u) + BT(B^*u). \tag{1.3}$$

This problem subsumes a wide spectrum of problems in applied nonlinear analysis. Some important special cases are:
i) By taking $T = \partial\phi$, we recover the mixed variational inequalities considered by Alduncin in [1]. If in addition A is a linear operator, we recover the general minimization problem (1.1) via a qualification condition, see [5].

ii) By setting $A = N_C$ the normal cone to a closed convex set C, $g = 0$ and $T = I - P_Q$, where P_Q stands for the metric projection onto Q and using the fact that $(I + N_C)^{-1} = P_C$ we obtain $u = P_C(u - \gamma B(I - P_Q)B^*u)$ and we recover the split feasibility problem (see [4]):

$$\text{find } u \in C \text{ such } B^*u \in Q. \tag{1.4}$$

It is worth mentioning that the convex feasibility formalism is at the core of the modeling of many inverse problems and has been used to model significant real-world problems, for instance, in sensor networks, in radiation therapy treatment planning, in computerized tomography and data compression.

iii) By taking A a single-valued operator and $T := T_F$, with F a monotone equilibrium function and T_F the associated maximal monotone operator, namely $v \in T_F(x) \Leftrightarrow F(x, y) + \langle v, x - y\rangle \geq 0,\ \forall v \in H$ (see [12]), we immediately get that (1.3) is related to the following mixed equilibrium problem

$$F(B^*u, B^*z) + \langle g - Au, u - z\rangle \geq 0 \quad \forall z.$$

iv) In the particular case in which $E = V$, $B = I$, $T = N_C$ the normal cone to some closed convex subset C and A single-valued, (1.3) is the classical variational inequality of finding $u \in C$ such that $\langle A(u), x - u\rangle \geq 0$ for each $x \in C$.

To solve (1.3), we usually use a splitting method i.e., an algorithm that uses only the resolvent mappings of A and BTB^*, rather than the resolvent mapping of their sum. In order to apply such method, we have to evaluate the resolvent of BTB^*. To this end D. Gabay [7] and more recently M. Fukushima [6] proved that if $B^* \circ B$ is an isomorphism, the operator BTB^* is maximal monotone. However the formula they proposed, for the associated resolvent, is difficult to evaluate in the practice. Furthermore the assumption is quite restrictive and excludes many cases in which the duality setup can be useful. To overcome this difficulty, we propose a generalization of the Uzawa type algorithm introduced in [3]. It is worth mentioning that in [2], Aujol proved that such algorithm is

very efficient for image restoration. It is also the case for initial and boundary-value constrained problems, see [3] and [7].
Throughout, E, V are real Hilbert spaces, $\langle\cdot,\cdot\rangle$ denotes the associated scalar products and $\|\cdot\|$ stands for the corresponding norms. To begin with, let us recall that an operator with domain $D(A)$ and range $R(A)$ is said to be monotone if $\langle u-v, x-y\rangle \geq 0 \quad \text{whenever} \quad u \in A(x), v \in A(y)$. It is said to be maximal monotone if, in addition, its graph, $gphA := \{(x,y) \in E \times E : y \in A(x)\}$, is not properly contained in the graph of any other monotone operator. It is well-known that for each $x \in E$ and $\lambda > 0$ there is a unique $z \in E$ such that $x \in (I+\lambda A)z$. The single-valued operator $J_\lambda^A := (I+\lambda A)^{-1}$ is called the resolvent of A of parameter λ. It is a nonexpansive mapping which is everywhere defined and is related to its Yosida approximate, namely $A_\lambda(x) := \frac{x - J_\lambda^A(x)}{\lambda}$, by the relation $A_\lambda(x) \in A(J_\lambda^A(x))$. Recall also that the inverse A^{-1} of A is the operator defined by $x \in A^{-1}(y) \Leftrightarrow y \in A(x)$, that the graph of a maximal monotone operator is is weakly-strongly closed and finally that an operator A is said to be α-strongly monotone, if there exists constants $\alpha > 0$ such that

$$\langle A(x_1) - A(x_2), x_1 - x_2\rangle \geq \alpha\|x_1 - x_2\|^2 \ \forall x_1, x_2 \in E.$$

2 The main result

To begin with let us notice that the dual problem associated to (1.3) is given by

$$0 \in (-B^*)A^{-1}(g - B(\cdot))(v) + T^{-1}(v), \tag{2.5}$$

see [9]. To solve problem (1.3), we propose to apply the following algorithm: y_0 being arbitrary, we consider the iterative scheme

$$\begin{cases} u_n \in A^{-1}(g - By_n), \\ y_{n+1} = T_\lambda(B^* u_n + \lambda y_n). \end{cases} \tag{2.6}$$

Algorithm (2.6) corresponds to the Uzawa algorithm associated to (2.5). In the special case with $B = I$, $T = N_C$ and $g = 0, \lambda = 1$, (2.6) reduces to

$$\begin{cases} u_n + A(y_n) = 0, \\ y_{n+1} = (I - P_C)(u_n + y_n), \end{cases}$$

which is the dual projection algorithm considered in [7]. In the setting of split feasibility problem (1.4) with $\lambda = 1$ our algorithm takes the following from

$$\begin{cases} u_n \in \partial\sigma_C(-By_n) = 0, \\ y_{n+1} = \frac{1}{2}(I - P_Q)(B^* u_n + y_n), \end{cases}$$

$\partial\sigma_C$ being the subdifferential of the support function of the convex C.
The key of the convergence proof is the following lemma (see, for instance [10]).

Lemma 2.1 *Let* T *a maximal monotone operator and* λ *a positif real parameter, then one has*

$$\|J_\lambda^T(x_1) - J_\lambda^T(x_1)\|^2 + \lambda^2\|T_\lambda(x_1) - T_\lambda(x_1)\|^2 \leq \|x_1 - x_2\|^2.$$

This follows immediately from definitions of the resolvent and Yosida operator.

Now we are in a position to state our main result.

Theorem 2.1 *Let* A, T *be two maximal monotone operators,* B *a bounded linear operator. Assume that* A *is* α*-strongly monotone and that* $0 < \frac{1}{\lambda} < \frac{2\alpha}{\|B^*\|^2}$. *Then, the sequence* (u_n) *strongly converges to a solution* u *of problem (1.3). Furthermore,* (y_n) *converges weakly to* y *satisfying* $y \in T(B^*u)$.

Proof. Let us first observe that u solves (1.3) if and only if there exists $y \in T(B^*u)$ with $g - By \in Ay$. In view of

$$y \in T(B^*u) \Leftrightarrow \lambda y + B^*u \in (\lambda I + T^{-1})y \Leftrightarrow y = T_\lambda(B^*u + \lambda y),$$

we obtain that u solves (1.3) if and only if (u, y) is a solution of:

$$\begin{cases} g - By \in A(u), \\ y = T_\lambda(B^*u + \lambda y). \end{cases} \tag{2.7}$$

On the other hand, we successively have

$$\begin{aligned} \|J_\lambda^T(B^*u + \lambda y) - J_\lambda^T(B^*u_n + \lambda y_n)\|^2 &+ \lambda^2\|y - y_{n+1}\|^2 \\ &\leq \|B^*(u - u_n) + \lambda(y - y_m)\|^2 \\ &= \lambda^2\|y - y_n\|^2 + \|B^*(u - u_m)\|^2 \\ &+ 2\lambda\langle B^*(u - u_n), y - y_m\rangle. \end{aligned}$$

The first line of (2.6), the first line of (2.7) and strong monotonicity of A imply

$$\begin{aligned} \langle y - y_n, B^*(u - u_n)\rangle &\leq -\alpha\|u - u_n\|^2 \\ &= -\frac{\alpha}{\|B^*\|^2}\|B^*(u - u_n)\|^2, \end{aligned} \tag{$\star$}$$

which combined with the inequality before yields

$$\begin{aligned} \|J_\lambda^T(B^*u + \lambda y) - J_\lambda^T(B^*u_n + \lambda y_n)\|^2 &+ \lambda^2\|y - y_{n+1}\|^2 \\ &\leq (1 - \frac{2\alpha\lambda}{\|B^*\|^2})\|B^*(u - u_n)\|^2 \qquad (\star\star) \\ &+ \lambda^2\|y - y_n\|^2. \end{aligned}$$

From which we deduce that $(\|y - y_n\|^2)$ is nodecreasing and thus is a convergent sequence in $I\!R$ to some positive real $l(y)$. Passing to the limit in $(\star\star)$, we obtain $\lim_{n\to+\infty}\|B^*(u - u^*)\| = 0$. In the light of $(\star)$, we infer that (u_n) strongly

converges to u.
Furthermore, from $(\star\star)$ we also have

$$\lim_{n\to+\infty} J_\lambda^T(B^*u_n+\lambda y_n) = J_\lambda^T(B^*u+\lambda y).$$

By virtue of the fact that $J_\lambda^T = I - \lambda T_\lambda$, we obtain from the second line of (2.7) that

$$J_\lambda^T(B^*u+\lambda y) = B^*u.$$

Using the second line of (2.6), we infer

$$y_{n+1} = T_\lambda(B^*u_n+\lambda y_n) = y_n + \frac{1}{\lambda}(B^*u_n - J_\lambda^T(B^*u_n+\lambda y_n)).$$

Passing to the limit in the last equality, we obtain that the sequence (y_n) is asymptotically regular, namely $\lim_n \|y_{n+1}-y_n\| = 0$.
Let us rewrite (2.6) as

$$B^*(u_n) + \lambda(y_n - y_{n+1}) \in T^{-1}(y_n) \Leftrightarrow y_n \in T(B^*(u_n) + \lambda(y_n - y_{n+1})),$$

and let y be a weak cluster point of (y_n). By passing to the limit (on a subsequence) in the last inclusion and by using the fact that the graph of the maximal monotone operator T is weakly-strongly closed, we obtain $y \in T(B^*u)$. It remains to prove that there is no more than one cluster point for (y_n), our argument follows that given in Rockafellar [10] and is presented here for completeness.
Let $\bar{y}$ be another cluster of $\{y_n\}$, we will show that $\bar{y} = y$. This is a consequence of $(\star\star)$. Indeed,

$$l(y) = \lim_{n\to+\infty} \|y_n - y\|^2 \text{ and } l(\bar{y}) = \lim_{n\to+\infty} \|y_n - \bar{y}\|^2,$$

from

$$\|y_n - \bar{y}\|^2 = \|y_n - y\|^2 + \|y - \bar{y}\|^2 + 2\langle y_n - y, y - \bar{y}\rangle,$$

we see that the limit of $\langle y_n - y, y - \bar{y}\rangle$ as $n \to +\infty$ must exists. This limit has to be zero, because $\bar{y}$ is a cluster point of $\{y_n\}$. Hence at the limit, we obtain $l(\bar{y}) = l(y) + \|y - \bar{y}\|^2$.
Reversing the role of $\bar{y}$ and y, we also have $l(y) = l(\bar{y}) + \|y - \bar{y}\|^2$. That is $\bar{y} = y$, which completes the proof.

3 Application to discrete image denoising

Having in mind that the resolvent of the subdifferential of a proper convex lower semicontinuous function ϕ is nothing else than the so-called proximal mapping given by $prox_{\lambda\phi}(g) = argmin_u\{\phi(u) + \frac{1}{2\lambda}\|u-g\|^2\}$, let us consider digital images defined on $\{1, \cdots, n\} \times \{1, \cdots, n\}$, reshape them columnwise into vectors $g \in I\!R^N$ with $N = n^2$ and apply our algorithm to the following problem

$$\min_{u \in I\!R^N} \{\phi_i(Bu) + \frac{1}{2}\|u - g\|_2^2\}, \ B \in I\!R^{M,N} \text{ with } M \geq N, \tag{3.8}$$

where for $i = 1$, $\phi_1(u) = \|\Lambda u\|_1$ with $\Lambda = diag(\beta_j)_{j=1}^M, \beta_j \geq 0$ and for $i = 2$, $\phi_2(u) = \|\tilde{\Lambda}|u|\|_1$ with

$\tilde{\Lambda} = diag(\tilde{\beta}_j)_{j=1}^N, \tilde{\beta}_j \geq 0, \ |u| = (\|\mathbf{u}_j\|_2)_{j=1}^N$ for $\mathbf{u}_j = (u_{j+kN})_{k=0}^{p-1}$ and $M = pN$.

The corresponding proximal mappings with $\lambda = 1$, see for instance [11], are given by

$$prox_{\phi_1}(g) = S_\Lambda(g) \text{ and } prox_{\phi_2}(g) = \tilde{S}_{\tilde{\Lambda}}(g),$$

where S_Λ stands for the soft shrinkage function given componentwise by

$$S_{\beta_j}(g_j) = \begin{cases} 0 & \text{if } |g_j| \leq \beta_j, \\ g_j - \beta_j sgn(g_j) & \text{if } |g_j| > \beta_j, \end{cases}$$

and $\tilde{S}_{\tilde{\Lambda}}$ denotes the coupled shrinkage function given by

$$\tilde{S}_{\tilde{\beta}_j}(\mathbf{g}_j) = \begin{cases} 0 & \text{if } \|\mathbf{g_j}\|_2 \leq \tilde{\beta}_j, \\ \mathbf{g}_j - \tilde{\beta}_j \frac{\mathbf{g_j}}{\|\mathbf{g_j}\|_2} & \text{if } \|\mathbf{g_j}\|_2 > \tilde{\beta}_j. \end{cases}$$

Now, the optimality condition for (3.8) is given by

$$g \in u + \partial(\phi_i \circ B)u,$$

thus our algorithm for soft shrinkage takes the following form

$$\begin{cases} u_n = g - B^T y_n, \\ y_{n+1} = \frac{1}{\lambda}(I - S_{\lambda\Lambda})(B(g - B^T y_n) + \lambda y_n). \end{cases} \tag{3.9}$$

For coupled shrinkage we have to replace $S_{\lambda\Lambda}$ by $\tilde{S}_{\lambda\tilde{\Lambda}}$.
According to theorem 2.1, if $0 < \frac{1}{\lambda} < \frac{2}{\|B\|^2}$, then the sequence (u_n) converges to a solution of (3.8) and (y_n) converges in turn to y satisfying $y \in \partial\phi(Bu)$ or equivalently $Bu \in (\partial\phi_i)^{-1}(y) = \partial\phi_i^*(y)$, where ϕ_i^* stands for the conjugate function of ϕ_i. Since by (2.7) we have that $u = g - B^T y$, we obtain

$$0 \in -B(g - B^T y) + \partial\phi_i^*(y) \Leftrightarrow 0 \in \partial(\frac{1}{2}\|g - B^T(\cdot)\|_2^2 + \phi_i^*(\cdot))(y).$$

In others words y solves the dual problem associated to (3.8), namely

$$\min_{y \in I\!R^M} \{\frac{1}{2}\|g - B^T(y)\|_2^2 + \phi_i^*(y)\}, \tag{3.10}$$

which is in fact a constrained least-squares problem, since ϕ_1^* and ϕ_2^* are respectively the indicator functions of $C = \{v \in I\!R^M; \ |v_i| \leq \beta_j, j = 1, \cdots, M\}$ and

$\tilde{C} = \{v \in I\!R^M;\ \|\mathbf{v}_j\|_2 \leq \tilde{\beta}_j, j = 1, \cdots, N\}$, see for example [11].

Our work is related to significant real-world applications, see for instance [1], [2] and [7], where such methods were applied to some relevant mechanical models, many image processing problems and several problems of continuum mechanics. Based on the work by Bermudez-Moreno [3], we give an extension of their unified framework and obtain a convergence result of our algorithm in the context of general variational inequalities and state an application to discrete image denoising problems.

Remark 3.1 *For a study of algorithms for more general variational inclusions including comments about their applications see, for example, M. A. Noor [8].*

Acknowledgment: The author would like to thank the anonymous referees and Professor Mohamed Tawhid for their careful reading of the paper and for their comments and suggestions which permitted to improve its presentation. ■

References

[1] G. Alduncin, *On Gabay's algorithms for mixed variational inequalities*, Applied Mathematics Optimization 35 (1997) p. 21-44.

[2] J. F. Aujol, *Some first-order algorithms for total variation based image restoration*, J. Math. Imaging and vision Vol. 34 no 3 (2009) p. 307-327.

[3] A. Bermudez and C. Moreno, *Duality methods for solving variational inequalities*, Comp. & Math. with applications 7 (1981) p. 43-58.

[4] Ch. Byrne, *Iterative oblique projection onto convex sets and the split feasibility problem*, Inverse Problems (2002) 18 441.

[5] I. Ekland and R. Temam, *Analyse convexe et inéquations variationnelles.* Dunod, Paris (1974).

[6] M. Fukushima, *The primal Douglas-Rachford splitting algorithm for a class of monotone mappings with application to the traffic equilibrium problem*, Math. Programming Vol. 72, no1 (1996), p. 1-15

[7] D. Gabay, and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*, Computers & Mathematics with applications Vol. 2, Issue 1, (1976) p. 17-40.

[8] M. A. Noor, *Generalized set-valued variational inclusions and resolvent equations*, J. Math. Anal. Appl. 228 (1998) p. 206-220.

[9] S. Robinson, *Composition duality and maximal monotonocity*, Math. Programming 85 (1999) p. 1-13.

[10] R. T. Rockafellar, *Monotone operator and the proximal point algorithm*, SIAM J. Control Optim 14 (1976), p. 877-898.

[11] S. Setzer, *Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage*, Lecture Notes in Computer Science, Vol. 5567 Springer (2009), p. 464-476.

[12] W. Takahashi, K. Aoyama and Y. Kimura, *Maximal monotone operators and Maximal monotone functions for equilibrium problems*, J. Convex Anal. 15-2 (2008), p. 395-409.