JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 59, No. 2, NOVEMBER 1988

Auxiliary Problem Principle Extended to Variational Inequalities[1]

G. COHEN[2]

Communicated by D. Q. Mayne

Abstract. The auxiliary problem principle has been proposed by the author as a framework to describe and analyze iterative optimization algorithms such as gradient or subgradient as well as decomposition/coordination algorithms (Refs. 1-3). In this paper, we extend this approach to the computation of solutions to variational inequalities. In the case of single-valued operators, this may as well be considered as an extension of ideas already found in the literature (Ref. 4) to the case of nonlinear (but still strongly monotone) operators. The case of multivalued operators is also investigated.

Key Words. Variational inequalities, monotony, decomposition/coordination algorithms.

1. Introduction

For classical optimization problems, Cohen (Refs. 1, 2) and Cohen and Zhu (Ref. 3) introduced the so-called auxiliary problem principle as a general framework to describe and analyze computational algorithms ranging from gradient or subgradient to decomposition/coordination algorithms. In this paper, we present an extension of this approach to the computation of solutions to variational inequalities.

[1] This research has been supported by the Centre National de la Recherche Scientifique (ATP Complex Technological Systems) and by the Centre National d'Études des Télécommunications (Contract No. 83.5B.034.PAA). It has been conducted partly while the author was visiting the Department of Electrical Engineering of the Pontificia Catholic University of Rio de Janeiro in July-August 1984 under the CAPES/COFECUB scientific exchange program.

[2] Associate Director, Centre d'Automatique et Informatique, École Nationale Supérieure des Mines de Paris, Fontainebleau, France. Also, Scientific Advisor, Institut National de Recherche en Informatique et Automatique, Le Chesnay, France.

More specifically, we study iterative algorithms in which an elementary step amounts to solving a so-called auxiliary problem, which consists in a minimization problem built around an auxiliary convex cost function chosen by the user. This is close to the approach found in Glowinski and associates (Ref. 4), except that these authors consider variational inequalities involving linear single-valued operators only. Also, they deal only with auxiliary problems built upon quadratic cost functions. Therefore, one may consider this paper as an extension of their work to deal with nonlinear operators and strongly convex auxiliary functions. The case of multivalued (point-to-set) operators will also be considered in what follows from the algorithmic point of view. We are not aware of any other similar work for multivalued operators. The assumption of strong monotony of the operator involved remains the key assumption in the convergence proofs presented hereafter.

By adequately choosing the auxiliary cost function, one may induce a decomposition of the auxiliary minimization problem into independent subproblems, provided that a decomposition of the decision space into the product of subspaces and of the constraint set into a product of independent subsets be given (this last feature is also encountered in Pang, Ref. 5). Actually, this work has been motivated by the study of decomposition/coordination procedures for computing Nash equilibria which are amenable, in several ways, to the solution of variational inequalities. This application is discussed in another forthcoming paper (Ref. 6).

2. Auxiliary Problem Principle Extended to Variational Inequalities Involving Single-Valued Operations

2.1. Basic Results on Variational Inequalities. We recall the definition of variational inequalities and some results about existence of their solutions drawn from Ekeland and Temam (Ref. 7). In this section, only the case of single-valued operators is considered. The case of multivalued (point-to-set) operators is postponed in the next section.

Let Ψ be a mapping from a reflexive Banach space U into its dual U^*, and let φ be a proper convex l.s.c. function from U into R. Let U^f be a closed convex subset of U. One looks for $u^* \in U^f$ such that

$$\langle \Psi(u^*), u - u^* \rangle + \varphi(u) - \varphi(u^*) \geqslant 0, \qquad \forall u \in U^f. \tag{1}$$

In Ref. 7, one can find the following existence theorem.

Theorem 2.1. Assume the following:

(A1) φ is a proper convex l.s.c. function from a reflexive Banach space U into R;

(A2) Ψ is a mapping from U into its dual space U^*, which is weakly continuous over every finite-dimensional subspace of U;

(A3) Ψ is monotone, that is,

$$\forall u, u' \in U, \qquad \langle \Psi(u) - \Psi(u'), u - u' \rangle \geq 0; \tag{2}$$

(A4) there exists $w \in \operatorname{dom} \varphi$ such that

$$\lim_{\substack{\|u\| \to +\infty \\ u \in U^f}} \frac{\langle \Psi(u), u - w \rangle + \varphi(u) - \varphi(w)}{\|u\|} = +\infty. \tag{3}$$

Then, there exists a solution u^* to (1).

Assumption (A4) is of course useless if U^f is bounded. Moreover, it is also met if we strengthen the monotony assumption (A3) by requiring strong monotony of Ψ over U^f (with modulus a), which means that the following assumption is satisfied:

$$\text{(A5)} \quad \exists a > 0: \forall u, u' \in U^f, \langle \Psi(u) - \Psi(u'), u - u' \rangle \geq a \|u - u'\|^2. \tag{4}$$

Under (A5), u^* is unique. When Ψ is the derivative of a function J [which is convex from (A3)], then u^* minimizes $(J + \varphi)(u)$ over U^f. When J is not differentiable, Ψ must be identified with the subdifferential ∂J, which is now a point-to-set mapping (see next section). When Ψ is not a derivative or a subdifferential, problem (1) cannot generally receive an interpretation in terms of a minimization problem. Nevertheless, such problems are encountered frequently in several fields of applied mathematics [numerical analysis, mechanics, game theory, other equilibrium problems in operations research, etc. (see Pang, Ref. 5)].

2.2. General Algorithm. Let us consider an auxiliary functional $K: U \to R$, that we choose convex and differentiable, and a positive number ϵ. For some $v \in U$, we introduce the following auxiliary problem:

$$\min_{u \in U^f} K(u) + \langle \epsilon \Psi(v) - K'(v), \qquad u \rangle + \epsilon \varphi(u). \tag{5}$$

Let $Q(v)$ denote the solution of this problem (we worry about conditions insuring existence and uniqueness later on). This solution is also characterized by the variational inequality (see Ref. 7)

$$\langle K'[\hat{u}(v)] + \epsilon \Psi(v) - K'(v), u - v \rangle$$
$$+ \epsilon[\varphi(u) - \varphi(\hat{u}(v)] \geq 0, \qquad \forall u \in U^f. \tag{6}$$

Lemma 2.1. If $\hat{u}(v) = v$, then $\hat{u}(v)$ is a solution, denoted u^*, of (1).

Proof. The proof is straightforward using (6). □

This lemma suggests the following fixed-point algorithm.

(i) At $k = 0$, start with some initial u^o.

(ii) At step k, solve the auxiliary problem (5) with $v = u^k$. Let u^{k+1} denote the solution of this problem.

(iii) Stop if $\|u^{k+1} - u^k\|$ is below some threshold. Otherwise, go back to (ii) with $k \leftarrow k+1$.

Theorem 2.2. *Convergence Theorem.* (i) Under (A1), (A2), (A5), there exists a unique solution u^* to (1).

(ii) Moreover, if $K : U \to R$ is a proper convex and differentiable functional and if its derivative K' is strongly monotone with modulus b over U^f, then there exists a unique solution u^{k+1} to (5) or (6), with u^k substituted for v.

(iii) Finally, if in addition Ψ is Lipschitz with modulus L over U^f, that is,

$$\exists L > 0 \colon \forall u, v \in U^f, \ \|\Psi(u) - \Psi(v)\| \leqslant L\|u - v\|, \tag{7}$$

and if we take

$$0 < \epsilon < 2ab/L^2, \tag{8}$$

then the sequence $\{u^k\}$ strongly converges toward u^*.

Proof. (i) φ is being a proper convex l.s.c. functional, for every $w \in \operatorname{int}(\operatorname{dom} \varphi)$, $\partial\varphi(w) \neq \varnothing$; see Ekeland and Temam, Ref. 7. Hence,

$$\varphi(u) \geqslant \varphi(w) + \langle r, u - w\rangle, \qquad \forall r \in \partial\varphi(w), \quad \forall u;$$

therefore, with the help of (4),

$$\begin{aligned}
&\langle \Psi(u), u - w\rangle + \varphi(u)\\
&\geqslant \langle \Psi(w), u - w\rangle + a\|u - w\|^2 + \varphi(w) + \langle r, u - w\rangle\\
&\geqslant (a\|u - w\| - \|r\| - \|\Psi(w)\|)\|u - w\| + \varphi(w), \qquad \forall u \in U^f,
\end{aligned}$$

which implies (A4). From theorem 2.1, u^* exists and is unique from (A5).

(ii) Following the same line, it can be proved that there exists a unique solution u^{k+1} to (6) where u^k has been substituted for v.

(iii) Let us now study the functional

$$\Lambda(u) \triangleq K(u^*) - K(u) - \langle K'(u), u^* - u\rangle \geqslant (b/2)\|u - u^*\|^2, \tag{9}$$

where the last inequality classically derives from the strong monotony of K'. We have

$$\begin{aligned}\Lambda(u^k)-\Lambda(u^{k+1}) &= K(u^{k+1})-K(u^k)-\langle K'(u^k), u^{k+1}-u^k\rangle \\ &\quad +\langle K'(u^{k+1})-K'(u^k), u^*-u^{k+1}\rangle \\ &\geq (b/2)\|u^k-u^{k+1}\|^2+\epsilon\langle\Psi(u^k), u^{k+1}-u^*\rangle \\ &\quad +\epsilon[\varphi(u^{k+1})-\varphi(u^k)]\end{aligned}$$

the inequality above resulting from (6) with $v=u^k$, $\hat{u}(v)=u^{k+1}$, and $u=u^*$. On the other hand, if we place u^{k+1} in (1) and combine it with the above, we get

$$\begin{aligned}\Lambda(u^k)-\Lambda(u^{k+1}) &\geq (b/2)\|u^k-u^{k+1}\|^2 \\ &\quad +\epsilon\langle\Psi(u^k)-\Psi(u^*), u^{k+1}-u^*\rangle \\ &= (b/2)\|u^k-u^{k+1}\|^2 \\ &\quad +\epsilon\langle\Psi(u^k)-\Psi(u^*), u^{k+1}-u^k\rangle \\ &\quad +\epsilon\langle\Psi(u^k)-\Psi(u^*), u^k-u^*\rangle \\ &\geq (b/2)\|u^k-u^{k+1}\|^2+\epsilon a\|u^k-u^*\|^2 \\ &\quad -\epsilon\|\Psi(u^k)-\Psi(u^*)\|\,\|u^{k+1}-u^k\| \\ &\geq (b/2)\|u^k-u^{k+1}\|^2+\epsilon a\|u^k-u^*\|^2 \\ &\quad -(\epsilon^2/2b)\|\Psi(u^k)-\Psi(u^*)\|^2 \\ &\quad -(b/2)\|u^{k+1}-u^k\|^2,\end{aligned}$$

that is, finally

$$\Lambda(u^k)-\Lambda(u^{k+1})\geq\epsilon(a-\epsilon L^2/2b)\|u^k-u^*\|^2. \tag{10}$$

Thanks to (8), (10) shows that the sequence $\{\Lambda(u^k)\}$ is strictly decreasing (unless $u^k=u^*$), it is nonnegative from (9); hence, it converges to some number. Therefore, the difference of two successive terms of this sequence goes to zero, and we conclude that u^k strongly converges toward u^* as k goes to infinity by looking again at (8)-(10). This completes the proof.

□

Remark 2.1. It is interesting to compare condition (8) with the corresponding one obtained by Cohen (Ref. 2) for minimization problems (that is, when Ψ is the derivative of some convex functional). This condition is, with the present notation,

$$0<\epsilon<2b/L. \tag{11}$$

We see that both conditions coincide when $a=L$. In general, a is of course smaller than L and condition (8) is more severe than (11).

Remark 2.2. When $a = 0$ (that is, when Ψ is monotone but not strongly monotone), (8) can no longer stand. On the other hand, for minimization problems, even when $a = 0$, still under (11), we proved convergence of u^k toward some solution u^* (which may now be nonunique), but the convergence took place in the weak topology (actually, the precise result in Ref. 2 is that the sequence $\{u^k\}$ is bounded and every cluster point in the weak topology is a solution; this result has been further refined in Cohen (Ref. 8) under mild additional assumptions to prove convergence of the whole sequence toward some solution. It does not seem possible to obtain such a result when Ψ is not the derivative of some convex functional (i.e., when Ψ is not symmetric) and when it is monotone but not strongly monotone, at least without any additional (questionable) assumption, as shown by the following example, due to Chaplais. Let $U = R^2$ and Ψ be the linear operator defined by

$$\Psi_1(u_1, u_2) = -u_2 \qquad \text{and} \qquad \Psi_2(u_1, u_2) = u_1.$$

It is checked that Ψ meets (2) but not (4). Assuming that $U^f = U$, $\varphi = 0$, and $K(\cdot) = \|\cdot\|^2/2$, algorithm (5) yields

$$u_1^{k+1} = u_1^k + \epsilon u_2^k,$$

$$u_2^{k+1} = u_2^k - \epsilon u_1^k,$$

from which it is seen that the norm of (u_1, u_2) increases with k for every positive value of ϵ, and thus the algorithm does not converge to the solution $(0, 0)$.

Remark 2.3. The general algorithm of Section 2.2 is an interesting way of computing a solution to (1) as long as (5) is easier to solve than (1). This depends crucially on the choice of the auxiliary cost function K. We first note that (5) is a minimization problem, which make several methods available to solve it. Moreover, K may be chosen, for example, quadratic [Pang and Chan (Ref. 9) call the algorithm a projection algorithm in this instance]. Finally, in terms of decomposition, if we assume that U is a product of subspaces U_i, that U^f is itself a product of closed convex subsets U_i^f of U_i (that is constraints are decoupled), and if φ is additive with respect to this decomposition ($\varphi(u)$ can be written as a sum of $\varphi_i(u_i)$), then, by choosing K additive too, problem (5) splits up into independent subproblems in each u_i. That is, coupling through Ψ can be managed by the algorithm. Coupling through the constraints would require handling the constraints through some dual tools (as done in Refs. 1–3). This is a direction for further investigation.

Remark 2.4. Although the fact that the auxiliary problem (5) is a minimization problem may be considered as an advantage in general, one may be interested in formulating the auxiliary problem directly as a variational inequality (6) with some auxiliary operator Γ standing for K'. This may open the possibility of making Γ be a closer approximation of Ψ with the hope of a faster convergence. Extending the above convergence proof to this situation is an open problem [see, however, Pang (Ref. 5), who provides conditions of the fixed-point type for less general situations].

3. Case of Multivalued Operators

In this section, we investigate briefly the case when Ψ is a point-to-set operator. From the point of view of existence results, this situation is somewhat more involved, and we refer the reader to Aubin and Ekeland (Ref. 10) for corresponding statements. We simply recall that problem (1) must now be stated as follows: Ψ has values which are subsets of U^*, and one looks for $u^* \in U^f$ such that

$$\exists r^* \in \Psi(u^*): \langle r^*, u-u^*\rangle + \varphi(u) - \varphi(u^*) \geq 0, \ \forall u \in U^f. \tag{12}$$

Correspondingly, (A5) must be changed into the following assumption:

(A5′) $\exists a > 0$: $\forall u, u' \in U^f, \forall r \in \Psi(u), r' \in \Psi(u')$,

$$\langle r - r', u - u'\rangle \geq a\|u - u'\|^2. \tag{13}$$

An analogous observation holds for (A3).

We come back to the general algorithm of Section 2.2, but we make the following modifications. First, since $\Psi(u^k)$ is now a set, we pick any r^k in this set to play the role of $\Psi(u^k)$ in Section 2.2. Moreover, as for nonsmooth minimization problems (see Ref. 3 or Ref. 11, for example), we replace the large step ϵ (which may depend on k but which remains away from zero) by small steps ϵ^k with the following conditions:

$$\epsilon^k > 0, \qquad \sum_{k=0}^{+\infty} \epsilon^k = +\infty, \qquad \sum_{k=0}^{+\infty} (\epsilon^k)^2 < +\infty. \tag{14}$$

To summarize, the auxiliary problem at step k is now

$$\min_{u \in U^f} K(u) + \langle \epsilon^k r^k - K'(u^k), \qquad u\rangle + \epsilon^k \varphi(u), \tag{15}$$

with $r^k \in \Psi(u^k)$. The solution is denoted u^{k+1}. Before stating our convergence theorem, we also have to introduce a new assumption instead of the

Lipschitz condition (7). This assumption is

$$\exists \alpha > 0, \exists \beta > 0: \forall u \in U^f, \forall r \in \Psi(u), \|r\| \leq \alpha \|u\| + \beta. \tag{16}$$

This essentially means that the norm of Ψ does not increase faster than linearly with the norm of u.

Theorem 3.1. (i) We assume that problem (12) does have a solution; that (13) holds (hence u^* is unique); that φ is a proper convex l.s.c. functional;

(ii) if, moreover, K is a proper convex and differentiable functional and if its derivative K' is strongly monotone with modulus b over U^f, then there exists a unique solution u^{k+1} to (15);

(iii) finally, if, in addition, Ψ meets condition (16), and if the sequence $\{\epsilon^k\}$ verifies (14), then the sequence $\{u^k\}$ strongly converges toward u^*.

Proof. We proceed as in the proof of Theorem 2.2 and we indicate only the major changes. By making use of (12) and the analog of (6), we can get, with similar calculations to those leading to (10)

$$\Lambda(u^k) - \Lambda(u^{k+1}) \geq \epsilon^k a \|u^k - u^*\|^2 - [(\epsilon^k)^2/2b] \|r^k - r^*\|^2. \tag{17}$$

Now, using the inequality

$$\|x + y\| \leq 2(\|x\|^2 + \|y\|^2)$$

repeatedly, also using (16), and summing up (17) from $k=0$ to $N-1$, we get, for all N,

$$\begin{aligned} (b/2)\|u^N - u^*\|^2 &\leq \Lambda(u^N) \\ &\leq \Lambda(u^o) + \sum_{k=0}^{N} [-\epsilon^k a \|u^k - u^*\|^2 + (\epsilon^k)^2(\gamma \|u^k - u^*\|^2 + \delta)], \end{aligned} \tag{18}$$

where γ and δ are some positive constants. Considering the two extreme sides of (18), ignoring the negative term in the right-hand side for the time being and using Lemma 5 in Ref. 3 and (14), we conclude that $\{u^k\}$ is a bounded sequence. Therefore, on a bounded convex hull of this sequence, the function $[u \to \|u - u^*\|^2]$ is Lipschitz. Moreover, from (18), (14), and the boundedness of $\{u^k\}$, we see that

$$\sum_{k=0}^{+\infty} \epsilon^k \|u^k - u^*\|^2 < +\infty.$$

Finally, with (14) and the above considerations, we can apply Lemma 4 of Ref. 3 to conclude that u^k strongly converges toward u^*. □

4. Conclusions

In this paper, we have shown how to extend the so-called auxiliary problem principle, which has proved to be a useful tool in studying iterative computational algorithms (including decomposition/coordination algorithms) in the case of optimization problems, to encompass more general variational problems. This work meets similar ideas and results already found in the literature (Refs. 4, 5, 9, among others) and sometimes generalizes them in some directions (in particular, for multivalued operators). Some open questions and topics of further investigations have been mentioned in Remarks 2.1 to 2.4.

References

1. COHEN, G., *Optimization by Decomposition and Coordination: A Unified Approach*, IEEE Transactions on Automatic Control, Vol. AC-23, No. 2, pp. 222–232, 1978.
2. COHEN, G., *Auxiliary Problem Principle and Decomposition of Optimization Problems*, Journal of Optimization Theory and Applications, Vol. 32, No. 3, pp. 277–305, 1980.
3. COHEN, G., and ZHU, D. L., *Decomposition Coordination Methods in Large-Scale Optimization Problems: The Nondifferentiable Case and the Use of Augmented Lagrangians*, Advances in Large-Scale Systems, Theory and Applications, Edited by J. B. Cruz, JAI Press, Greenwich, Connecticut, Vol. 1, pp. 203–266, 1984.
4. GLOWINSKI, R., LIONS, J. L., and TREMOLIERES, R., *Analyse Numérique des Inéquations Variationnelles, Tome 1: Théorie Générale, Premières Applications*, Dunod, Paris, France, 1976.
5. PANG, J. S., *Asymmetric Variational Inequality Problems over Product Sets: Applications and Iterative Methods*, Mathematical Programming, Vol. 31, No. 2, pp. 206–219, 1985.
6. COHEN, G., *Nash Equilibria: Gradient and Decomposition Algorithms*, Large Scale Systems, Vol. 12, No. 2, pp. 173–184, 1987.
7. EKELAND, I., and TEMAM, R., *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, Holland, 1976.
8. COHEN, G., *Two Lemmas and Their Use in Convergence Analysis of Some Optimization Algorithms*, Internal Report No. E/68, Centre d'Automatique et Informatique, Ecole des Mines de Paris, Fontainebleau, France, 1982.
9. PANG, J. S., and CHAN, D., *Iterative Methods for Variational and Complementarity Problems*, Mathematical Programming, Vol. 24, No. 3, pp. 284–313, 1982.
10. AUBIN, J. P., and EKELAND, I., *Applied Nonlinear Analysis*, Wiley, New York, New York, 1984.
11. AUSLENDER, A., *Optimisation, Méthodes Numériques*, Masson, Paris, France, 1976.