1 NECESSARY OPTIMALITY CONDITIONS FOR CONTROL OF STRONGLY MONOTONE VARIATIONAL INEQUALITIES

Jane J. Ye

Department of Mathematics and Statistics
University of Victoria
Victoria, British Clumbia
Canada V8W 3P4*

janeye@uvic.ca

Abstract: In this paper we derive necessary optimality conditions involving Mordukhovich coderivatives for optimal control problems of strongly monotone variational inequalities.

1.1 INTRODUCTION

The optimal control problem for a system governed by an elliptic variational inequality, first proposed by J.L. Lions (1969,1972) and studied in Barbu (1984) is as follows:

Let V and H be two Hilbert spaces (*state spaces*) such that

$$V \subseteq H = H^* \subseteq V^*$$

where V^* is the dual of V, H^* is the dual of H which is identified with H, and injections are dense and continuous. Let U be another Hilbert space (*control space*). Suppose that $A \in L(V, V^*)$ is *coercive*, i.e., there exists a constant $\mu > 0$ such that

$$\langle Av, v \rangle \geq \mu \|v\|_V^2 \quad \forall v \in V$$

*The research of this paper is partially supported by NSERC.

where $\langle\cdot,\cdot\rangle$ is the duality pairing on $V^* \times V$ and identified with the inner product on H. The optimal control problem for an elliptic variational inequality is the following minimization problem:

$$(P) \qquad \begin{array}{ll} \min & g(y)+h(u) \\ \text{s.t.} & y \in K, u \in U_{ad} \\ & \langle Ay, y'-y\rangle \geq \langle Bu+f, y'-y\rangle \quad \forall y' \in K, \end{array}$$

where $B \in L(U, V^*)$ is compact, $K \subset V$ and $U_{ad} \subseteq U$ are two closed convex subsets in V and U respectively, $f \in V^*$, $g : K \to R_+$ and $h : U_{ad} \to R_+$ are two given functions.

In this paper we study the following optimal control of strongly monotone variational inequality which is more general than the one proposed by Lions:

$$(OCVI) \qquad \begin{array}{ll} \min & J(y,u) \\ \text{s.t.} & y \in K, u \in U_{ad} \\ & \langle F(y,u), y'-y\rangle \geq 0 \quad \forall y' \in K, \end{array}$$

where the following assumptions are satisfied

(A1) K and U_{ad} are closed convex subsets of Asplund spaces (which include all reflexive Banach spaces) V and U respectively. There is a finite codimensional closed subspace M such that $U_{ad} \subseteq M$ and the relative interior of U_{ad} with respect to the subspace M is nonempty.

(A2) $J : V \times U_{ad} \to R$ is Lipschitz near $(\bar{y}, \bar{u})$.

(A3) $F : V \times U_{ad} \to V^*$ is strictly differentiable at $(\bar{y}, \bar{u})$ (see definition given in Remark 2) and locally strongly monotone in y uniformly in u, i.e., there exist $\mu > 0$ and $U(\bar{y}, \bar{u})$, a neighborhood of $(\bar{y}, \bar{u})$ such that

$$\langle F(y',u)-F(y,u), y'-y\rangle \geq \mu\|y'-y\|^2 \;\; \forall (y,u),(y',u) \in U(\bar{y},\bar{u}) \cap (K \times U_{ad}).$$

Our main result is the following theorem:

Theorem 1 *Let $(\bar{y}, \bar{u})$ be a local solution of problem* (OCVI). *Then there exists $\eta \in V$ such that*

$$0 \in \partial J(\bar{y},\bar{u}) + F'(\bar{y},\bar{u})^*\eta + D^*N_K(\bar{y}, -F(\bar{y},\bar{u}))(\eta) \times \{0\} + \{0\} \times N(\bar{u}, U_{ad}) \tag{1.1}$$

where ∂ denotes the limiting subgradient (see Definition 2), F' denotes the strict derivative (see Remark 2), $N(\bar{u}, U_{ad})$ denotes the normal cone of the convex set U_{ad} at $\bar{u}$ and N_K denotes the normal cone operator defined by

$$N_K(y) := \begin{cases} \text{the normal cone of } K \text{ at } y & \text{if } y \in K \\ \emptyset & \text{if } y \notin K \end{cases}$$

and D^* denotes the coderivative of a set-valued map (see Definition 5).

This is in fact in the form of the optimality condition given by Shi (1988, 1990) with the paratingent coderivative of the the set-valued map N_K replaced by the Mordukhovich coderivative.

In the case where $J(y,u) = g(y) + h(u)$ and $F(y,u) = Ay - Bu - f$ as in problem (P), Inclusion (1.1) becomes

$$0 \in \partial g(\bar{y}) + A^*\eta + D^*N_K(\bar{y}, -F(\bar{y},\bar{u}))(\eta) \tag{1.2}$$

$$0 \in \partial h(\bar{u}) - B^*\eta + N(\bar{u}, U_{ad}). \tag{1.3}$$

Notice that $N_K(y) = \partial\psi_K(y)$, the coderivative of the set-valued map N_K can be considered as a second order generalized derivative of ψ_K. Hence inclusions (1.2) and (1.3) are in the form of the necessary optimality condition given in Theorem 3.1 of Barbu (1984) with the Clarke subgradient replaced by the limiting subgradient which is in general a smaller set than the Clarke subgradient and with the *notational* second order generalized derivative replaced by the *true* second order generalized derivative D^*N_K.

We organize the paper as follows. §1.2 contains background material on nonsmooth analysis and preliminary results. In §1.3 we derive necessary optimality conditions for (OCVI).

1.2 PRELIMINARIES

This section contains some background material on nonsmooth analysis which will be used in the next section. We only give concise definitions that will be needed in the paper. For more detail information on the subject, our references are Clarke (1983), Mordukhovich and Shao (1996a,b).

First we give some concepts for various normal cones.

Definition 1 *Let Ω be a nonempty subset of a Banach space X and let $\epsilon \geq 0$.*

(i) *Given $\bar{x} \in cl\Omega$, the closure of set Ω, the set*

$$\hat{N}_\epsilon(\bar{x}, \Omega) := \{x^* \in X^* : \limsup_{x \to \bar{x}, x \in \Omega} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq \epsilon\} \tag{1.4}$$

is called the set of Fréchel $\epsilon-$normal to set Ω at point $\bar{x}$. When $\epsilon = 0$, the set (1.4) is a cone which is called the Fréchel normal cone to Ω at point $\bar{x}$ and is denoted by $\hat{N}(\bar{x}, \Omega)$.

(ii) *The following nonempty cone*

$$N(\bar{x}, \Omega) := \{x^* \in X^* | \exists x_k \to \bar{x}, \epsilon_k \downarrow 0, x_k^* \xrightarrow{w^*} x^*, x_k^* \in \hat{N}_{\epsilon_k}(x_k, \Omega) \text{ as } k \to \infty\} \tag{1.5}$$

is called the limiting normal cone to Ω at point $\bar{x}$, As proved in Mordukhovich and Shao (1996a), in Asplund spaces X the normal cone (1.5) admits the simplified representation

$$N(\bar{x},\Omega)=\{x^*\in X^*|\exists x_k\to\bar{x}, x_k^*\xrightarrow{w^*}x^*, x_k^*\in\hat{N}(x_k,\Omega)\text{ as }k\to\infty\}$$

Using the definitions for normal cones, we now give definitions for subgradients of a single-valued map.

Definition 2 *Let X be a Banach space and $f:X\to R\cup\{+\infty\}$ be lower semicontinuous and finite at $\bar{x}\in X$. The limiting subgradient of f at $\bar{x}$ is defined by*

$$\partial f(\bar{x}):=\{x^*\in X^*:(x^*,-1)\in N((\bar{x},f(\bar{x})),epi(f))\}$$

and the singular subdifferential of f at $\bar{x}$ is defined by

$$\partial^\infty f(\bar{x}):=\{x^*\in X^*:(x^*,0)\in N((\bar{x},f(\bar{x})),epi(f))\},$$

where $epi(f):=\{(x,r)\in X\times R:f(x)\le r\}$ is the epigraph of f.

Remark 1 *Let Ω be a closed set of a Banach space and ψ_Ω denote the indicator function of Ω. Then it follows easily from the definition that*

$$\partial\psi_\Omega(\bar{x})=\partial^\infty\psi_\Omega(\bar{x})=N_\Omega(\bar{x}).$$

The following fact is also well-known and follows easily from the definition:

Proposition 1 *Let X be a Banach space and $f:X\to R\cup\{+\infty\}$ be lower semicontinuous. If f has a local minimum at $\bar{x}\in X$, then*

$$0\in\partial f(\bar{x}).$$

To ensure that the sum rule holds in an infinite dimensional Asplund space, we need the following definitions.

Definition 3 *Let X be a Banach space and Ω a closed subset of X. Ω is said to be sequentially normally compact at $\bar{x}\in\Omega$ if any sequence (x_k,x_k^*) satisfying*

$$x_k^*\in\hat{N}(x_k,\Omega), x_k\to\bar{x}, x_k^*\xrightarrow{w^*}0 \text{ as } k\to\infty$$

contains a subsequence with $\|x_{k_\nu}^*\|\to 0$ as $\nu\to 0$.

Definition 4 *Let Let X be a Banach space and $f:X\to R\cup\{+\infty\}$ be lower semicontinuous and finite at $\bar{x}\in X$. f is said to be sequentially normally epi-compact around $\bar{x}$ if its epigraph is sequentially normally compact at $\bar{x}$.*

Proposition 2 *Let Let X be a Banach space and $f:X\to R\cup\{+\infty\}$ be directionally Lipschitz in the sense of Clarke (1983) at $\bar{x}\in X$. Then f is sequentially normally epi-compact around $\bar{x}$.*

Proof. By Proposition 3.1 of Borwein (1987), if f is directionally Lipschitz at $\bar{x}$, then it is compactly Lipschitz at $\bar{x}$, i.e., its epigragh is compactly epi-Lipschitz at $(\bar{x}, f(\bar{x}))$ in the sense of Borwein and Strojwas (1985). By Proposition 3.7 of Loewen (1992), a compactly epi-Lipschitz set is sequentially normally compact. Hence the proof of the proposition is complete.

■

The following is the sum rule for limiting subgradients.

Proposition 3 *[Corollary 3.4 of Mordukhovich and Shao (1996b)] Let X be an Asplund space, functions $f_i : X \to R \cup \{\infty\}$ be lower semicontinuous and finite at $\bar{x}$, $i = 1, 2$ and one of them be sequentially normally epi-compact around $\bar{x}$. Then one has the inclusion*

$$\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x})$$

provided that

$$\partial^\infty f_1(\bar{x}) \cap (-\partial^\infty f_2(\bar{x})) = \{0\}.$$

For set-valued maps, the definition for limiting normal cone leads to the definition of coderivative of a set-valued map introduced by Mordukhovich.

Definition 5 *Let $\Phi : X \rightrightarrows Y$ be an arbitrary set-valued map (assigning to each $x \in X$ a set $\Phi(x) \subset Y$ which may be empty) and $(\bar{x}, \bar{y}) \in$ cl gphΦ where gphΦ is the graph of the set-valued Φ defined by*

$$gph\Phi := \{(x, y) \in X \times Y : y \in \Phi(x)\}$$

and clΩ denotes the closure of the set Ω. The set-valued map $D^*\Phi(\bar{x}, \bar{y})$ from Y^* into X^* defined by

$$D^*\Phi(\bar{x}, \bar{y})(y^*) = \{x^* \in X^* : (x^*, -y^*) \in N((\bar{x}, \bar{y}), gph\Phi)\},$$

is called the coderivative of Φ at point $(\bar{x}, \bar{y})$. By convention for $(\bar{x}, \bar{y}) \notin$ clgphΦ we define $D^*\Phi(\bar{x}, \bar{y})(y^*) = \emptyset$. The symbol $D^*\Phi(\bar{x})$ is used when Φ is single-valued at $\bar{x}$ and $\bar{y} = \Phi(\bar{x})$.

Remark 2 *Recalled that a single-valued mapping $\Phi : X \to Y$ is called strictly differentiable at $\bar{x}$ with the derivative $\Phi'(\bar{x})$ if*

$$\lim_{x, x' \to \bar{x}} \frac{\Phi(x) - \Phi(x') - \Phi'(\bar{x})(x - x')}{\|x - x'\|} = 0$$

In the special case when a set-valued map is single-valued and $\Phi : X \to Y$ is strictly differentiable at $\bar{x}$, the coderivative coincides with the adjoint linear operator to the classical strict derivative, i.e.,

$$D^*\Phi(\bar{x})(y^*) = \Phi'(\bar{x})^* y^* \quad \forall y^* \in Y^*,$$

where $\Phi'(\bar{x})^*$ denotes the adjoint of $\Phi'(\bar{x})$.

The following proposition is a sum rule for coderivatives when one mapping is single-valued and strictly differentiable.

Proposition 4 [Theorem 3.5 of Mordukhovich and Shao (1996b)] *Let X, Y be Banach spaces, $f : X \to Y$ be strictly differentiable at $\bar{x}$ and $\Phi : X \rightrightarrows Y$ be an arbitrary closed set-valued map. Then for any $\bar{y} \in f(\bar{x}) + \Phi(\bar{x})$ and $y^* \in Y^*$ one has*

$$D^*(f+\Phi)(\bar{x},\bar{y})(y^*) = f'(\bar{x})^* y^* + D^*\Phi(\bar{x}, \bar{y} - f(\bar{x}))(y^*).$$

1.3 PROOF OF THE NECESSARY OPTIMALITY CONDITION

The purpose of this section is to derive the necessary optimality conditions involving coderivatives for (OCVI) as stated in Theorem 1.

Proof of Theoerm 1. Since K is a convex set, by the definition of a normal cone in the sense of convex analysis, it is easy to see that problem (OCVI) can be rewritten as the optimization problem with generalized equation constraints (GP):

$$\begin{array}{lll} \text{(GP)} & \min & J(y,u) \\ & \text{s.t.} & (y,u) \in V \times U_{ad}. \\ & & 0 \in F(y,u) + N_K(y), \end{array}$$

where

$$N_K(y) := \begin{cases} \text{the normal cone of } K \text{ at } y & \text{if } y \in K \\ \emptyset & \text{if } y \notin K \end{cases}$$

is the normal cone operator.

Let $\Phi(y,u) : V \times U \rightrightarrows V^*$ be the set-valued map defined by

$$\Phi(y,u) := F(y,u) + N_K(y).$$

By local optimality of the pair $(\bar{y},\bar{u})$ we can find $U(\bar{y},\bar{u})$, a neighborhood of $(\bar{y},\bar{u})$, such that

$$\begin{array}{rcl} J(\bar{y},\bar{u}) & \leq & J(y^*,u) \quad \forall (y^*,u) \in U(\bar{y},\bar{u}) \cap (V \times U_{ad}) \text{ s.t. } 0 \in \Phi(y^*,u), \\ & = & J(y,u) + J(y^*,u) - J(y,u) \\ & & \qquad \forall (y^*,u) \in U(\bar{y},\bar{u}) \cap (V \times U_{ad}) \text{ s.t. } 0 \in \Phi(y^*,u), \\ & \leq & J(y,u) + L_J \|y^* - y\| \\ & & \qquad \forall (y^*,u),(y,u) \in U(\bar{y},\bar{u}) \cap (V \times U_{ad}) \text{ s.t. } 0 \in \Phi(y^*,u) \\ & \leq & J(y,u) + \dfrac{L_J}{\mu} \dfrac{\langle F(y^*,u) - F(y,u), y^* - y \rangle}{\|y^* - y\|} \\ & & \qquad \forall (y^*,u),(y,u) \in U(\bar{y},\bar{u}) \cap (K \times U_{ad}) \text{ s.t. } 0 \in \Phi(y^*,u). \end{array}$$

Let $y, y^* \in K, u \in U_{ad}$ be such that $0 \in \Phi(y^*,u)$ and $v \in \Phi(y,u)$. Then by definition of the normal cone, we have

$$\begin{array}{rcl} \langle v - F(y,u), y' - y \rangle & \leq & 0 \quad \forall y' \in K \\ \langle -F(y^*,u), y' - y^* \rangle & \leq & 0 \quad \forall y' \in K. \end{array}$$

In particular one has

$$\begin{aligned} \langle v - F(y,u), y^* - y \rangle &\leq 0 \\ \langle -F(y^*,u), y - y^* \rangle &\leq 0 \end{aligned}$$

which implies that

$$\langle v + F(y^*,u) - F(y,u), y^* - y \rangle \leq 0.$$

Hence we have

$$J(\bar{y},\bar{u}) \leq J(y,u) + \frac{L_J}{\mu}\|v\| \quad \forall (y,u,v) \in Gr\Phi, (y,u) \in U(\bar{y},\bar{u}) \cap (V \times U_{ad}).$$

That is, $(\bar{y},\bar{u},0)$ is a local solution to the penalized problem of (GP):

$$\begin{aligned} \min \quad & J(y,u) + \frac{L_J}{\mu}\|v\| \\ \text{s.t.} \quad & (y,u) \in V \times U_{ad}. \\ & (y,u,v) \in Gr\Phi. \end{aligned}$$

Let $\psi_\Omega(x)$ denote the indicate function of Ω. Then it is easy to see that $(\bar{y},\bar{u},0)$ is a local minimizer of the lower semicontinuous function

$$f(y,u,v) := J(y,u) + \frac{L_J}{\mu}\|v\| + \psi_{Gr\Phi}(y,u,v) + \psi_{U_{ad}}(u).$$

It follows from Propositions 1 that

$$0 \in \partial f(\bar{y},\bar{u},0). \tag{1.6}$$

Since J is Lipschitz near $(\bar{y},\bar{u})$, $g(y,u,v) := J(y,u) + \frac{L_J}{\mu}\|v\|$ is Lipschitz at $(\bar{y},\bar{u},0)$. Hence it is directionally Lipschitz by Theorem 2.9.4 of Clarke (1983) and $\partial^\infty g(\bar{y},\bar{u},0) = \{0\}$ by Proposition 2.5 of Mordukhovich and Shao (1996a). Consequently by Proposition 3 we have

$$\begin{aligned} \partial f(\bar{y},\bar{u},0) &\subseteq \partial g(\bar{y},\bar{u},0) + \partial(\psi_{Gr\Phi} + \psi_{U_{ad}})(\bar{y},\bar{u},0) \\ &\subseteq \partial J(\bar{y},\bar{u}) \times \frac{L_J}{\mu}B + \partial(\psi_{Gr\Phi} + \psi_{U_{ad}})(\bar{y},\bar{u},0), \end{aligned} \tag{1.7}$$

where B is the closed unit ball of V. Next we shall prove that

$$\partial(\psi_{Gr\Phi} + \psi_{U_{ad}})(\bar{y},\bar{u},0) \subseteq \partial\psi_{Gr\Phi}(\bar{y},\bar{u},0) + \{0\} \times \partial\psi_{U_{ad}}(\bar{u}) \times \{0\}$$

by using the sum rule Propostion 3. By (vii) of Theorem 1 and Remark 3 of Borwein, Lucet and Mordukhovich (1998), the assumption (A1) implies that U_{ad} is compactly Epi-Lipschitz. Hence the epigragh of the function $\psi_{U_{ad}}$ is also compactly Epi-Lipschitz. By Proposition 3.7 of Loewen (1992), a compactly epi-Lipschitz set is sequentially normally compact. Therefore the function $\psi_{U_{ad}}$

is sequentially normlly epi-compact around every point in U_{ad}. Now we check the condition

$$\partial^{\infty}\psi_{Gr\Phi}(\bar{y},\bar{u},0)\cap(-\{0\}\times\partial^{\infty}\psi_{U_{ad}}(\bar{u})\times\{0\})=\{0\}.$$

Let $(0,\xi_2,0)\in\partial^{\infty}\psi_{Gr\Phi}(\bar{y},\bar{u},0)\cap(-\{0\}\times\partial^{\infty}\psi_{U_{ad}}(\bar{u})\times\{0\})$. Then

$$(0,\xi_2,0)\in\partial^{\infty}\psi_{Gr\Phi}(\bar{y},\bar{u},0)=N_{Gr\Phi}(\bar{y},\bar{u},0)$$

So by definition of coderivatives,

$$(0,\xi_2)\in D^*\Phi(\bar{y},\bar{u},0)(0).$$

By the sum rule for coderivatives Proposition 4, we have

$$D^*\Phi(\bar{y},\bar{u},0)(0) \quad\subset\quad F'(\bar{y},\bar{u})^*0+D^*N_K(\bar{y},-F(\bar{y},\bar{u}))(0)\times\{0\}$$

which implies that $\xi_2=0$. Hence by Proposition 3 we have

$$\begin{aligned}\partial(\psi_{Gr\Phi}+\psi_{U_{ad}})(\bar{y},\bar{u},0) &\subseteq \partial\psi_{Gr\Phi}(\bar{y},\bar{u},0)+\{0\}\times\partial\psi_{U_{ad}}(\bar{u})\times\{0\}\\ &= N_{Gr\Phi}(\bar{y},\bar{u},0)+\{0\}\times N(\bar{u},U_{ad})\times\{0\}.(1.8)\end{aligned}$$

By (1.6), (1.7) and (1.8), we have

$$0\in\partial J(\bar{y},\bar{u})\times\frac{L_J}{\mu}B+N_{Gr\Phi}(\bar{y},\bar{u},0)+\{0\}\times N(\bar{u},U_{ad})\times\{0\}.$$

That is, there exist $\eta\in\frac{L_J}{\mu}B$, $(\xi_1,\xi_2)\in\partial J(\bar{y},\bar{u})$ and $\zeta\in N(\bar{u},U_{ad})$ such that

$$(-\xi_1,-\xi_2-\zeta,-\eta)\in N_{Gr\Phi}(\bar{y},\bar{u},0).$$

Hence by the definition of coderivatives and the sum rule for coderivatives Proposition 4, we have

$$\begin{aligned}(-\xi_1,-\xi_2-\zeta) &\in D^*\Phi(\bar{y},\bar{u},0)(\eta)\\ &\subset F'(\bar{y},\bar{u})^*\eta+D^*N_K(\bar{y},-F(\bar{y},\bar{u}))(\eta)\times\{0\}\end{aligned}$$

The proof of the theorem is complete. ■

References

Barbu, V. (1984). *Optimal Control of Variational Inequalities* Pitman, Boston.
Borwein, J.M. (1987). Epi-Lipschitz-like sets in Banach space:theorems and examples, *Nonlinear Analysis: TMA.* **11** 1207-1217.
Borwein, J.M., Lucet, Y. and Mordukhovich, B. (1998). Compactly Epi-Lipschitz convex sets, Preprint.
Borwein, J.M. and Strojwas (1985), H.M., Tangential approximations, *Nonlinear Analysis TMA* **9** 1347-1366.

Clarke, F.H. (1983). *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York.

Lion, J.L. (1969). *Quelwues méthodes de résolution des probleémes aux limites non-linéaires*, Dunod-Gauthier-Villars, Paris.

Lion, J.L. (1972). Some aspects of the optimal control of distributed parameter systems CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Loewen, P.D. (1992). Limits of Fréchet Normals in nonsmooth analysis, in: A.D. Ioffe *et al.* (eds.), Optimization and Nonlinear Analysis, Pitman Research Notes in Math. Series **244**, Longman, Harlow, Essex, 178-188.

Mordukhovich, B.S. and Shao, Y. (1996a). Nonsmooth sequential analysis in Asplund spaces *Trans. Amer. Math. Soc.* **348** 1235-1280.

Mordukhovich, B.S. and Shao, Y. (1996b). Nonconvex differential calculus for infinite-dimensional multifunctions, *Set-Valued Analysis* **4** 205-236.

Shi, S. (1988). Optimal control of strongly monotone variational inequalities, *SIAM J. on Control Optim.* **26** 274-290.

Shi, S. (1990). Erratum: Optimal control of strongly monotone variational inequalities, *SIAM J. on Control Optim.* **28** 243-249.