Necessary optimality conditions for discrete inclusions

Aurelian Cernea
Faculty of Mathematics and Informatics,
University of Bucharest,
Academiei 14, 010014 Bucharest, Romania
e-mail: acernea@math.math.unibuc.ro

Abstract

The aim of this paper is to present a short survey of several new results concerning optimization of discrete inclusions. We study an optimization problem given by a discrete inclusion with end point constraints and we present several approaches concerning first and second-order necessary optimality conditions for this problem.

Key words: derived cone, discrete inclusion, local controllability, maximum principle.

Consider the problem

$$\text{minimize} \quad g(x_N) \tag{1}$$

over the solutions of the discrete inclusion

$$x_i \in F_i(x_{i-1}), \quad i = 1, 2, ..., N, \quad x_0 \in X_0, \tag{2}$$

with end point constraints of the form

$$x_N \in X_N, \tag{3}$$

where $F_i : \mathbf{R}^n \to \mathcal{P}(\mathbf{R}^n)$, $i = 1, 2, ..., N$, $X_0, X_N \subset \mathbf{R}^n$ and $g : \mathbf{R}^n \to \mathbf{R}$ are given.

The aim of this paper is to announce several new results concerning first and second-order necessary optimality conditions for problem (1)-(3).

At the beginning we obtain necessary optimality conditions for a solution $\overline{x} = (\overline{x}_0, \overline{x}_1, ..., \overline{x}_N)$ to the problem (1)-(3) in terms of a variational inclusion associated to the problem (2) and in terms of the cone of interior directions (Dubovitskij-Miljutin tangent cone) to the set X_N at x_N. Afterwards this result is improved by replacing the cone of interior directions with the concept of derived cone introduced by Hestenes ([5]) and using a remarkable "intersection property" of derived cones obtained by Mirica ([7]). Finally, we present an approach concerning second-order necessary optimality conditions for the problem (1)-(3).

Optimal control problems for systems described by discrete inclusions have been studied by many authors ([1], [6], [8], [9], [10], [12] etc.). In the framework of multivalued problems, necessary optimality conditions for problem (1)-(2) (i.e. without end point constraints) are obtained in [11]. The idea in [11] is to use a special (Warga's) open mapping theorem to obtain a sufficient condition for (2) to be locally controllable around a given trajectory and as a consequence, via a separation result, to obtain the maximum principle.

In contrast with the approach in [11], even if the problem studied in the present paper is more difficult, due to end point constraints, the method in our approach seems to be conceptually very simple, relying only 2-3 clear-cut steps and using a minimum of auxiliary results.

Denote by S_F the solution set of inclusion (2), i.e.

$$S_F := \{x = (x_0, x_1, ..., x_N); \quad x \text{ is a solution of (2)}\}.$$

and by $R_F^N := \{x_N; \quad x \in S_F\}$ the reachable set of inclusion (2).

We consider $\overline{x} = (\overline{x}_0, \overline{x}_1, ..., \overline{x}_N) \in S_F$ a solution of (2).

Since the reachable set R_F^N is, generally, neither a differentiable manifold, nor a convex set, its infinitesimal properties may be characterized only by tangent cones in a generalized sense, extending the classical concepts of tangent cones in Differential Geometry and Convex Analysis, respectively.

From the multitude of the intrinsic tangent cones in the literature, *the contingent, the quasitangent and Clarke's tangent cones*, defined, respectively, by

$$K_x X = \{v \in \mathbf{R}^n; \quad \exists\, s_m \to 0+,\ x_m \in X : \tfrac{x_m - x}{s_m} \to v\}$$
$$Q_x X = \{v \in \mathbf{R}^n; \quad \exists\, c(.) : [0, s_0) \to X,\ c(0) = x,\ c'(0) = v\}$$
$$C_x X = \{v \in \mathbf{R}^n; \quad \forall\, (x_m, s_m) \to (x, 0+),\ x_m \in X,\ \exists\, y_m \in X : \tfrac{y_m - x_m}{s_m} \to v\}$$

seem to be among the most oftenly used in the study of different problems involving nonsmooth sets and mappings.

The *second-order quasitangent set* to X at x relative to $v \in Q_xX$ is defined by

$$Q^2_{(x,v)}X = \{w \in \mathbf{R}^n; \quad \forall s_m \to 0+,\ \exists w_m \to w:\ x + s_m v + s_m^2 w_m \in X\}.$$

We recall that, in contrast with K_xX, Q_xX, the cone C_xX is convex and one has $C_xX \subset Q_xX \subset K_xX$.

Another important tangent cone is the *cone of interior directions* (Dubovitskij-Miljutin tangent cone) defined by

$$I_xX := \{v \in \mathbf{R}^n; \quad \exists\, s_0,\, r > 0:\ x + sB(v,r) \subset X\ \forall\, s \in [0,s_0)\},$$

$$B(v,r) := \{w \in \mathbf{R}^n; \quad ||w - v|| < r\}, \quad \overline{B}(v,r) := \mathrm{cl}B(v,r).$$

From the properties of the cone of interior directions we recall only the following:

$$Q_xX_1 \cap I_xX_2 \subset Q_x(X_1 \cap X_2). \tag{4}$$

Definition 1. ([5]) A subset $M \subset \mathbf{R}^n$ is said to be a *derived set to* $X \subset \mathbf{R}^n$ at $x \in X$ if for any finite subset $\{v_1, ..., v_k\} \subset M$, there exist $s_0 > 0$ and a continuous mapping $a(.) : [0, s_0]^k \to X$ such that $a(0) = x$ and $a(.)$ is (conically) differentiable at $s = 0$ with the derivative $\mathrm{col}[v_1, ..., v_k]$ in the sense that

$$\lim_{\mathbf{R}^k_+ \ni \theta \to 0} \frac{||a(\theta) - a(0) - \sum_{i=1}^k \theta_i v_i||}{||\theta||} = 0.$$

We shall write in this case that the derivative of $a(.)$ at $s = 0$ is given by

$$Da(0)\theta = \sum_{i=1}^k \theta_j v_j, \quad \forall \theta = (\theta_1, \ldots, \theta_k) \in \mathbf{R}^k_+ := [0, \infty)^k.$$

A subset $C \subset \mathbf{R}^n$ is said to be a *derived cone* of X at x if it is a derived set and also a convex cone.

For the basic properties of derived sets and cones we refer to Hestenes ([5]); we recall that if M is a derived set then $M \cup \{0\}$ as well as the convex cone generated by M, defined by

$$cco(M) = \{\sum_{i=1}^k \lambda_j v_j; \quad \lambda_j \geq 0,\ k \in \mathbf{N},\ v_j \in M,\ j = 1, ..., k\}$$

is also a derived set, hence a derived cone.

The fact that the derived cone is a proper generalization of the classical concepts in Differential Geometry and Convex Analysis is illustrated by the following results ([5]): if $X \subset \mathbf{R}^n$ is a differentiable manifold and $T_x X$ is the tangent space in the sense of Differential Geometry to X at x

$$T_x X = \{v \in \mathbf{R}^n; \quad \exists\, c : (-s, s) \to X, \text{ of class } C^1, c(0) = x, c'(0) = v\}$$

then $T_x X$ is a derived cone; also, if $X \subset \mathbf{R}^n$ is a convex subset then the tangent cone in the sense of Convex Analysis defined by

$$TC_x X = cl\{t(y - x); \quad t \geq 0,\, y \in X\}$$

is also a derived cone. By $cl\ A$ we denote the closure of the set $A \subset \mathbf{R}^n$.

Since any convex subcone of a derived cone is also a derived cone, such an object may not be uniquely associated to a point $x \in X$; moreover, simple examples show that even a maximal with respect to set-inclusion derived cone may not be uniquely defined: if the set $X \subset \mathbf{R}^2$ is defined by

$$X = C_1 \bigcup C_2, \quad C_1 = \{(x, x); x \geq 0\}, \quad C_2 = \{(x, -x), x \leq 0\}$$

then C_1 and C_2 are both maximal derived cones of X at the point $(0, 0) \in X$.

We recall that two cones $C_1, C_2 \subset \mathbf{R}^n$ are said to be *separable* if there exists $q \in \mathbf{R}^n \backslash \{0\}$ such that:

$$< q, v > \leq 0 \leq < q, w > \quad \forall v \in C_1, w \in C_2.$$

We denote by C^+ the positive dual cone of $C \subset \mathbf{R}^n$

$$C^+ = \{q \in \mathbf{R}^n; \quad < q, v > \geq 0, \quad \forall\, v \in C\}$$

The negative dual cone of $C \subset \mathbf{R}^n$ is $C^- = -C^+$.

The following "intersection property" of derived cones, obtained by Mirică ([7]), is a key tool in the proof of necessary optimality conditions.

Lemma 2. ([7]) *Let $X_1, X_2 \subset \mathbf{R}^n$ be given sets, $x \in X_1 \cap X_2$, and let C_1, C_2 be derived cones to X_1, resp. to X_2 at x. If C_1 and C_2 are not separable, then:*

$$cl(C_1 \cap C_2) = (cl(C_1)) \cap (cl(C_2)) \subset Q_x(X_1 \cap X_2).$$

For a mapping $g(.) : X \subset \mathbf{R}^n \to \mathbf{R}$ which is not differentiable, the classical (Fréchet) derivative is replaced by some generalized directional derivatives. We recall only the upper right-contingent derivatives, defined by

$$\overline{D}_K g(x;v) = \limsup_{(\theta,w)\to(0+,v)} \frac{g(x+\theta w)-g(x)}{\theta}, \quad v \in K_x X$$

and in the case when $g(.)$ is locally-Lipschitz at $x \in int(X)$ by Clarke's generalized directional derivative, defined by:

$$D_C g(x;v) = \limsup_{(y,\theta)\to(x,0+)} \frac{g(y+\theta v)-g(y)}{\theta}, \quad v \in \mathbf{R}^n.$$

The first and second order uniform lower Dini derivative are defined as follows

$$\underline{D}_K g(x;v) = \liminf_{(v',\theta)\to(v,0+)} \frac{g(x+\theta v')-g(x)}{\theta},$$

$$D^2_K g(x,v;w) = \liminf_{(w',\theta)\to(w,0+)} \frac{g(x+\theta v+\theta^2 w')-g(x)-\theta \underline{D}_K g(x;v)}{\theta^2}.$$

When $g(.)$ is of class C^2 one has

$$\underline{D}_K g(x,v) = g'(x)^T v, \quad D^2_K g(x,v;w) = g'(x)^T z + \frac{1}{2} v^T g''(x) v.$$

The results in the next section will be expresed, in the case where $g(.)$ is locally-Lipschitz at x, in terms of the Clarke generalized gradient, defined by:

$$\partial_C g(x) = \{q \in \mathbf{R}^n; \quad < q, v > \leq D_C g(x;v) \quad \forall\, v \in \mathbf{R}^n\}.$$

By $\mathcal{P}(\mathbf{R}^n)$ we denote the family of all subsets of $\mathbf{R}^n$.

Corresponding to each type of tangent cone, say $\tau_x X$ one may introduce a *set-valued directional derivative* of a multifunction $G(.) : X \subset \mathbf{R}^n \to \mathcal{P}(\mathbf{R}^n)$ (in particular of a single-valued mapping) at a point $(x,y) \in Graph(G)$ as follows

$$\tau_y G(x;v) = \{w \in \mathbf{R}^n;\ (v,w) \in \tau_{(x,y)} \mathrm{Graph}(G)\}, \quad v \in \tau_x X.$$

Similarly one may define second-order directional deivatives of the set-valued map $G(.)$. For example the second-order quasitangent derivative of G at (x,u) relative to $(y,v) \in Q_{(x,u)}(graph(G(.))$ is the set-valued map $Q^2_{(u,v)} G(x,y,.)$ defined by

$$graph Q^2_{(u,v)} G(x,y;.) = Q^2_{((x,u),(y,v))}(graph G(.)).$$

We recall that a set-valued map, $A(.) : \mathbf{R}^n \to \mathcal{P}(\mathbf{R}^n)$ is said to be a *convex* (respectively, closed convex) *process* if Graph$(A(.)) \subset \mathbf{R}^n \times \mathbf{R}^n$ is a convex (respectively, closed convex) cone.

In what follows, we shall assume the following hypothesis.

Hypothesis 1. *i) $X_0, X_N \subset \mathbf{R}^n$ are closed sets.*
ii) There exists $L > 0$ such that $F_i(.)$ is Lipschitz with the Lipschitz constant L, $\forall i \in \{1, ..., N\}$.
iii) There exists $A_i(.) : \mathbf{R}^n \to \mathcal{P}(\mathbf{R}^n)$, $i = 1, 2, ..., N$ a family of closed convex processes such that

$$A_i(v) \subset Q_{\overline{x}_i} F_i(\overline{x}_{i-1}; v) \quad \forall v \in \mathbf{R}^n, \forall i \in \{1, ..., N\}.$$

Let $A_0 \subset Q_{\overline{x}_0} X_0$ be a closed convex cone. To the problem (2) we associate the linearized problem

$$w_i \in A_i(w_{i-1}), \quad i = 1, 2, ..., N, \quad w_0 \in A_0. \tag{5}$$

Denote by S_A the solution set of inclusion (5) and by R_A^N the reachable set of inclusion (5).

We recall that if $A : \mathbf{R}^n \to \mathcal{P}(\mathbf{R}^n)$ is a set-valued map then the adjoint of A is the multifunction $A^* : \mathbf{R}^n \to \mathcal{P}(\mathbf{R}^n)$ defined by

$$A^*(p) = \{q \in \mathbf{R}^n; \ < q, v > \leq < p, v' > \quad \forall (v, v') \in graph A(.)\}.$$

Using the property in (4), the fact that $R_A^N \subset Q_{\overline{x}_N} R_F^N$ and the duality results in [11] we obtain a Maximum Principle for problem (1)-(3).

Theorem 1. *Let $\overline{x} = (\overline{x}_0, \overline{x}_1, ..., \overline{x}_N) \in S_F$ be an optimal solution for problem (1)-(3) such that Hypothesis 1 is satisfied and let $g(.) : \mathbf{R}^n \to \mathbf{R}$ be a locally Lipschitz function.*

Then for any closed convex cone $A_0 \subset Q_{\overline{x}_0} X_0$ and any convex cone $C_1 \subset I_{\overline{x}_N} X_N$ there exist $\lambda \in \{0, 1\}$ and $p = (p_0, p_1, ..., p_N) \in \mathbf{R}^{(N+1)n}$ such that

$$p_0 \in A_1^*(p_1), \ p_1 \in A_2^*(p_2), \ ..., \ p_{N-1} \in A_N^*(p_N), \ p_N = w_N, \tag{6}$$

$$p_N \in \lambda \partial_C g(\overline{x}_N) - C_1^+, \quad p(0) \in A_0^+ \tag{7}$$

$$\begin{aligned} &< -p_0, \overline{x}_0 > = \max\{< -p_0, v >; \quad v \in X_0\}, \\ &< -p_i, \overline{x}_i > = \max\{< -p_i, v >; \quad v \in F_i(\overline{x}_{i-1})\}, \quad i = 1, ..., N, \end{aligned} \tag{8}$$

$$(9) \qquad \lambda + ||p|| > 0.$$

For the details of the proof see [2].

In Theorem 1 an important hypothesis is that the terminal set X_N is assumed to have a nonempty cone of interior directions. Such type of assumptions may be overcome using the concept of derived cone.

Using the fact that if A_0 is a derived cone to X_0 at $\overline{x}_0$ then the reachable set R_A^N is a derived cone to R_F^N at $\overline{x}_N$, Lemma 1 and the the duality results in [11] we have the next version of the Maximum Principle for problem (1)-(3).

Theorem 2. *Let $X_N \subset \mathbf{R}^n$ be a closed set, let $X_0 \subset \mathbf{R}^n$, $F_i, i = 1, ..., N$ satisfy Hypothesis 1 and are convex valued, let $\overline{x} = (\overline{x}_0, \overline{x}_1, ..., \overline{x}_N) \in S_F$ be an optimal solution for problem (1)-(3) such that Hypothesis 1 is satisfied and let $g(.) : \mathbf{R}^n \to \mathbf{R}$ be a locally Lipschitz function.*

Then for any derived cones A_0 of X_0 at $\overline{x}_0$ and C_1 of X_N at $\overline{x}_N$ there exist $\lambda \in \{0, 1\}$ and $p = (p_0, p_1, ..., p_N) \in \mathbf{R}^{(N+1)n}$ such that (6)-(9) hold true.

The proof can be found in [3].

In particular, when F_i are expressed in the parametrized form

$$F_i(x_{i-1}) = \bigcup_{u_i \in U_i} f_i(x_{i-1}, u_i) \quad \forall x_{i-1} \in \mathbf{R}^n, i = 1, ..., N$$

and $X_0 = \{\overline{x}_0\}$, i.e. inclusion (2) becames the nonlinear discrete system

$$(10) \qquad x_i = f_i(x_{i-1}, u_i), \quad u_i \in U_i, \quad i = 1, ..., N, \quad x_0 = \overline{x}_0$$

taking $A_i(v) = \frac{\partial f_i}{\partial x}(\overline{x}_{i-1}, \overline{u}_i)v$, $i = 1, ..., N$ we obtain the following consequence of Theorem 2.

Corollary 1. *Let $X_N \subset \mathbf{R}^n$, $U_i \subset \mathbf{R}^n$ be compact set, let $f_i(.,.) : \mathbf{R}^n \times U_i \to \mathbf{R}^n$ be such that $f_i(., u_i)$ is differentiable and the multifunction F_i satisfy Hypothesis 1, $i = 1, ..., N$, let $\overline{x} = (\overline{x}_0, \overline{x}_1, ..., \overline{x}_N) \in \mathbf{R}^{(N+1)n}$ be an optimal solution for problem (1),(10),(3) and $\overline{u} = (\overline{u}_0, \overline{u}_1, ..., \overline{u}_N)$ be a control corresponding to solution $\overline{x}$. Consider $g(.) : \mathbf{R}^n \to \mathbf{R}$ a locally Lipschitz function.*

Then for any derived cone C_1 of X_N at $\overline{x}_N$ there exist $\lambda \in \{0,1\}$ and $p = (p_0, p_1, ..., p_N) \in \mathbf{R}^{(N+1)n}$ such that

$$p_0 \in (\frac{\partial f_1}{\partial x}(\overline{x}_0, \overline{u}_1))^*(p_1), \ ..., \ p_{N-1} \in (\frac{\partial f_N}{\partial x}(\overline{x}_{N-1}, \overline{u}_N))^*(p_N),$$

$$p_N \in \lambda \partial_C g(\overline{x}_N) - C_1^+,$$

$$< -p_i, \overline{x}_i >= \max\{< -p_i, f_i(\overline{x}_{i-1}, u_i) >, \quad u_i \in U_i\}, \quad i = 1, ..., N,$$

$$\lambda + ||p|| > 0.$$

Denote by R_Q^N the reachable set of the discrete inclusion.

$$w_i \in Q_{\overline{x}_i} F_i(\overline{x}_{i-1}, w_{i-1}), \quad i = \overline{1, N}, \quad w_0 \in Q_{\overline{x}_0} X_0. \tag{11}$$

Let $\overline{y} = (\overline{y}_0, \overline{y}_1, ..., \overline{y}_N)$ satisfy (11) and let R_Q^2 denote the reachable set of the discrete inclusion

$$v_i \in Q^2_{(\overline{x}_i, \overline{y}_i)} F_i(\overline{x}_{i-1}, \overline{y}_{i-1}; v_{i-1}), \quad i = \overline{1, N}, \quad w_0 \in Q^2_{(\overline{x}_0, \overline{y}_0)} X_0.$$

In the next result we obtain second-order necessary optimality conditions for problem (1)-(3).

Theorem 3. *Assume that Hypothesis 1 is satisfied, let $g(.) : \mathbf{R}^n \to \mathbf{R}$ be a locally Lipschitz function, let $C_0 \subset Q_{\overline{x}_0} X_0$ be a closed convex cone, let $\overline{x} = (\overline{x}_0, \overline{x}_1, ..., \overline{x}_N) \in S_F$ be an optimal solution for problem (1)-(3) and assume that the following constraint qualification is satisfied*

$$\{-w_N; \exists p = (p_0, p_1, ..., p_N) \in \mathbf{R}^{(N+1)n} \text{ such that } p_0 \in C_0^+, p_0 \in (C_{\overline{x}_1} F_1(\overline{x}_0,$$

$$.))^* p_1, ..., p_{N-1} \in (C_{\overline{x}_N} F_N(\overline{x}_{N-1}, .))^* p_N, p_N = w_N\} \cap (C_{\overline{x}_N} X_N)^+ = \{0\}.$$

Then we have the first-order necessary condition

$$\underline{D}_K g(\overline{x}_N; y_N) \geq 0 \quad \forall y_N \in R_Q^N \cap Q_{\overline{x}_N} X_N.$$

Furthermore, if equality holds for some $\overline{y}_N$, then we have the second-order necessary condition

$$D_K^2 g(\overline{x}_N, \overline{y}_N; w_N) \geq 0 \quad \forall w_N \in R_Q^2 \cap Q^2_{(\overline{x}_N, \overline{y}_N)} X_N.$$

The proof, that can be find in [4], is based on a general (abstract) optimality condition formulated by Zheng ([13]) and use also several first and second-order approximations of the reachable set R_F^N at $\overline{x}_N$ ([4]).

References

[1] V. G. Boltjanskii, *"Optimal control for discrete systems"*, Nauka, Moskow (in Russian), 1973.

[2] A. Cernea, On the maximum principle for discrete inclusions with end point constraints, *Math. Reports*, **7(57)**, 2005.

[3] A. Cernea, Derived cones to reachable sets of discrete inclusions, *Nonlinear Studies*, submitted.

[4] A. Cernea, Second-order necessary conditions for discrete inclusions with end point constraints, *Discuss. Math., Diff. Incl., Control Optimization*, to appear.

[5] M.R. Hestenes, *"Calculus of Variations and Optimal Control Theory"*, Wiley, New York, 1966.

[6] A.G. Kusraev, Discrete maximum principle, *Math. Notes*, **34**(1984), 617-619.

[7] Ş. Mirică, New proof and some generalizations of the Minimum Principle in Optimal Control, *J. Optim. Theory Appl.*, **74**(1992), 487-508.

[8] V.N. Phat, Controllability of nonlinear discrete systems without differentiability assumption, *Optimization*, **19**(1988), 133-142.

[9] A.I. Propoi, The maximum principle for discrete control systems, *Automat. Remote Control*, **26**(1965), 451-461.

[10] R. Pytlak, A variational approach to discrete maximum principle, *IMA J. Math. Control and Information*, **9**(1992), 197-220.

[11] H. D. Tuan and Y. Ishizuka, On controllability and maximum principle for discrete inclusions, *Optimization*, **34**(1995), 293-316.

[12] N. D. Yen and T. C. Dien, On local controllability of nondifferentiable discrete time systems with nonconvex constraints on control, *Optimization*, **20**(1989), 889-899.

[13] H.Zheng, Second-order necessary conditions for differential inclusion problems, *Appl. Math. Opt.*, **30** (1994), 1-14.