OPTIMAL CONDITIONS FOR ENDPOINT CONSTRAINED OPTIMAL CONTROL

KYUNG-EUNG KIM

ABSTRACT. We deduce the necessary conditions for the optimality of end-point constrained optimal control problem. These conditions comprise the adjoint equation, the maximum principle and the transversality condition. We assume that the cost function is merely differentiable. Therefore the technique under Lipschitz continuity hypothesis is not directly applicable. We introduce Fermat's rule and value function technique to obtain the results.

1. Introduction

We consider the following Mayer type optimal control problem with endpoint constraint

$$\begin{array}{ll} \text{minimize} & \psi(x(S),x(T)) \\ \text{subject to} & \dot{x}(t) = f(t,x(t),u(t)) \text{ a.e. } t \in [S,T] \\ & u(t) \in U(t) \text{ a.e. } t \in [S,T] \\ & (x(S),x(T)) \in C, \end{array}$$

where [S, T] is a time interval,

$$\psi: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$
 and $f: [S, T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$

are two functions.

$$U:[S,T] \leadsto \mathbb{R}^m$$

is a set-valued map (which is also called multifunction) and

$$C \subset \mathbb{R}^n \times \mathbb{R}^n$$

is a set.

If the cost function ψ and the velocity function f are Lipschitz continuous, then the Theorem 6.2.1 of [6] is directly applicable to our problem. According to this theorem, in the case when $\lambda \neq 0$, the following conditions are necessary

Received December 4, 2007.

²⁰⁰⁰ Mathematics Subject Classification. 49K15, 93C10.

Key words and phrases. necessary conditions for optimal control, value function, end-points constraint, Fermat's rule, limiting normal, limiting subdifferential.

for optimality of the arc \bar{x} and the control \bar{u} corresponding to \bar{x} : there exists an absolutely continuous function $p:[S,T]\to\mathbb{R}^n$ such that

$$\begin{split} -\dot{p}(t) &\in co\partial_x H(t,\bar{x}(t),p(t),\bar{u}(t)) \ \text{ a.e. } \ t \in [S,T] \\ H(t,\bar{x}(t),p(t),\bar{u}(t)) &= \max_{u(t) \in U(t)} H(t,\bar{x}(t),p(t),u) \ \text{ a.e. } \ t \in [S,T] \\ (p(S),-p(T)) &\in \partial \psi(\bar{x}(S),\bar{x}(T)) + N_C(\bar{x}(S),\bar{x}(T)), \end{split}$$

where

$$H(t, x, p, u) = p \cdot f(t, x, u)$$

and ∂ and $N_C(\cdot)$ denote the limiting subdifferential and the limiting normal respectively.

But in our control problem the cost function ψ is not Lipscitz continuous (it is merely differentiable). Therefore the theorem of [6] is not directly applicable to our problem, i.e., the limiting subdifferential

$$\partial \psi(\bar{x}(S), \bar{x}(T))$$

is not reduced to the classical derivative

$$\nabla \psi(\bar{x}(S), \bar{x}(T)).$$

This reduction is possible in case that ψ is continuously differentiable. See [5, p.304].

2. Preliminaries

We call an absolutely continuous function an arc and say that an arc satisfying the following control system

(1)
$$\dot{x}(t) = f(t, x(t), u(t)) \text{ a.e. } t \in [S, T]$$

$$u(t) \in U(t) \text{ a.e. } t \in [S, T]$$

$$(x(S), x(T)) \in C$$

is feasible. A feasible arc is called a trajectory. A measurable function $u: [S,T] \to \mathbb{R}^m$ satisfying $u(t) \in U(t)$ a.e. is called a control. We introduce the reachable set in the above system without endpoint constraint

$$R(S,T) = \{(x(S), x(T)) \mid x \text{ is a solution to (1) where } C = \mathbb{R}^n \times \mathbb{R}^n \}.$$

Throughout the whole paper we suppose that

- I) C is a closed set
- II) ψ is differentiable
- III) $f(t,\cdot,u)$ is differentiable
- IV) $f(\cdot, x, u)$ is measurable
- V) $f(t, x, \cdot)$ is continuous
- VI) there exists k > 0 such that for all $t \in [S, T]$, for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$||f(t,x,u) - f(t,y,u)|| \le k||x-y||$$

VII) there exists m > 0 such that for all $t \in [S, T]$,

$$\sup_{v \in f(t, x, U(t))} \|v\| \le m$$

VIII) $U(\cdot)$ is measurable

IX) for all $t \in [S, T]$, U(t) is nonempty and compact

X) for all (t, x), f(t, x, U(t)) is convex.

Let K be a subset of a Banach space X. The positive polar cone of K is defined by

$$K^+ = \{ p \in X' \mid \forall u \in K, \ \langle p, u \rangle \ge 0 \},$$

where X' is the dual space of X. The negative polar cone of K is defined by

$$K^{-} = \{ p \in X' \mid \forall u \in K, \ \langle p, u \rangle \le 0 \}.$$

The contingent cone (Bouligand tangent cone) $T_K(x)$ to K at x is defined by

$$T_K(x) = \{ v \in X \mid \liminf_{h \to 0^+} \frac{\operatorname{dist}(x + hv, K)}{h} = 0 \}.$$

In other words, $T_K(x)$ comprises vectors ξ corresponding to which there exist some sequence v_i in K and some sequence $t_i \setminus 0$ such that $t_i^{-1}(v_i - x) \to \xi$.

In the following, we fix a trajectory-control pair (\bar{x}, \bar{u}) . Let us introduce the following sets.

$$A = \left\{ w \in W^{1,2}([S,T];\mathbb{R}^n) \; \middle| \; \begin{array}{l} \dot{w}(t) = \frac{\partial f}{\partial x}(t,\bar{x}(t),\bar{u}(t))w(t) + v(t), \\ v(t) \in T_{f(t,\bar{x}(t),U(t))}(\dot{\bar{x}}(t)) \; \text{ a.e.} \end{array} \right\}$$

$$\tilde{A} = \left\{ w \in A \; \middle| \; w \in W^{1,\infty}([S,T];\mathbb{R}^n) \right\},$$

where $W^{1,2}([S,T];\mathbb{R}^n)$ (respectively, $W^{1,\infty}([S,T];\mathbb{R}^n)$) is the space of functions $w \in L^2([S,T];\mathbb{R}^n)$ (respectively, $L^{\infty}([S,T];\mathbb{R}^n)$) such that $w' \in L^2([S,T];\mathbb{R}^n)$ (respectively, $L^{\infty}([S,T];\mathbb{R}^n)$).

The next propositions will be used to prove the main theorem.

Proposition 2.1. \tilde{A} is dense in A for the topology of uniform convergence.

Proof. See
$$[3]$$
.

Proposition 2.2. Let γ is a linear continuous function defined by

$$\begin{array}{cccc} \gamma: C([S,T];\mathbb{R}^n) & \to & \mathbb{R}^n \times \mathbb{R}^n \\ w & \mapsto & (w(S),w(T)). \end{array}$$

Then we have

$$\gamma(\tilde{A}) \subset T_{R(S,T)}(\bar{x}(S),\bar{x}(T)).$$

Now consider the following linear continuous operator:

$$\begin{array}{cccc} (1 \times \mathcal{D}): & W^{1,2}([S,T];\mathbb{R}^n) & \to & L^2([S,T];\mathbb{R}^n) \times L^2([S,T];\mathbb{R}^n) \\ & w & \to & (w,\dot{w}) \end{array}$$

where \mathcal{D} denotes the differential operator.

Proposition 2.3. Set

$$\begin{split} L &= \{(x,y) &\in L^2([S,T];\mathbb{R}^n) \times L^2([S,T];\mathbb{R}^n) \mid \\ & y(s) \in \frac{\partial f}{\partial x}(s,\bar{x}(s),\bar{u}(s)) \cdot x(s) + T_{f(s,\bar{x}(s),U(s))}(\dot{\bar{x}}(s)) \\ & a.e. \ in \ [S,T] \}. \end{split}$$

Then

(2)
$$W^{1,2}([S,T];\mathbb{R}^n)' \supset A^+ = (1 \times \mathcal{D})^*(L^+),$$

where A is defined above and $(1 \times D)^*$ denotes the adjoint of $1 \times D$.

Proof. See [2].
$$\Box$$

3. Necessary conditions for optimality

In this section, we first assume that the process (\bar{x}, \bar{u}) is optimal for the problem without endpoint constraint. Note that $\tilde{A} \neq \emptyset$ because $0 \in \tilde{A}$. The next lemma which is called Fermat's rule is the main idea to obtain the necessary conditions for optimality.

Lemma 3.1. If \bar{x} is optimal for the problem without endpoint constraint, then

$$\nabla \psi(\bar{x}(S), \bar{x}(T)) \in \left(T_{R(S,T)}(\bar{x}(S), \bar{x}(T))\right)^{+}.$$

Proof. Let $(u, v) \in T_{R(S,T)}(\bar{x}(S), \bar{x}(T))$. Then there exist sequences $h_i \longrightarrow 0^+$ and $(u_i, v_i) \longrightarrow (u, v)$ such that

$$(\bar{x}(S), \bar{x}(T)) + h_i(u_i, v_i) \in R(S, T) \quad \forall i.$$

Since \bar{x} is optimal, we have

$$\psi(\bar{x}(S) + h_i u_i, \bar{x}(T) + h_i v_i) \ge \psi(\bar{x}(S), \bar{x}(T))$$

and thereby

$$\langle \nabla \psi(\bar{x}(S), \bar{x}(T)), (u, v) \rangle \ge 0.$$

Since (u, v) is arbitrary, we have

$$\nabla \psi(\bar{x}(S), \bar{x}(T)) \in \left(T_{R(S,T)}(\bar{x}(S), \bar{x}(T))\right)^{+}.$$

Now we can deduce the necessary conditions for optimality in the problem without endpoint constraint.

Proposition 3.2. Suppose that (\bar{x}, \bar{u}) is optimal. Then there exists an absolutely continuous function p such that

i) (Adjoint equation)

$$-p'(t) = \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))^* p(t) \quad a.e. \quad in \quad [S, T]$$

ii) (Maximum principle)

$$\max_{u \in U(t)} \left\langle p(t), f(t, \bar{x}(t), u) \right\rangle = \left\langle p(t), f(t, \bar{x}(t), \bar{u}(t)) \right\rangle \quad a.e. \quad in \quad [S, T]$$

iii) (Transversality conditions)

$$(p(S), -p(T)) \in \nabla \psi(\bar{x}(S), \bar{x}(T)).$$

Proof. By Lemma 3.1 and Proposition 2.2,

$$\nabla \psi(\bar{x}(S), \bar{x}(T)) \in \left(T_{R(S,T)}(\bar{x}(S), \bar{x}(T))\right)^{+} \subset \left(\gamma\left(\tilde{A}\right)\right)^{+}$$

i.e., for all $w \in \tilde{A} \subset C([S,T];\mathbb{R}^n)$,

$$\langle \nabla \psi(\bar{x}(S), \bar{x}(T)), \gamma(w) \rangle = \langle \gamma^* \nabla \psi(\bar{x}(S), \bar{x}(T)), w \rangle \ge 0.$$

This implies that

(3)
$$\gamma^* \nabla \psi(\bar{x}(S), \bar{x}(T)) \in \tilde{A}^+.$$

Using the fact that $\tilde{A}^+ = A^+$ (Proposition 2.1), let $\xi \in A^+ \subset C([S,T];\mathbb{R}^n)'$ be such that

$$\gamma^* \nabla \psi(\bar{x}(S), \bar{x}(T)) = \xi.$$

Note that

$$C([S,T];\mathbb{R}^n)' \subset W^{1,2}([S,T];\mathbb{R}^n)'.$$

Therefore we have

(4)
$$\langle \gamma^* \nabla \psi(\bar{x}(S), \bar{x}(T)), w \rangle = \langle \xi, w \rangle \ge 0 \quad \forall w \in W^{1,2}([S, T]; \mathbb{R}^n).$$

Proposition 2.3 implies that there exists $(r,q) \in L^+$ such that

(5)
$$\xi = (1 \times \mathcal{D})^*(r, q).$$

The equations (4) and (5) imply that for all $w \in W^{1,2}([S,T];\mathbb{R}^n)$,

$$\langle \gamma^* \nabla \psi(\bar{x}(S), \bar{x}(T)), w \rangle = \langle (1 \times \mathcal{D})^*(r, q), w \rangle.$$

Thus for all $w \in W^{1,2}([S,T];\mathbb{R}^n)$,

(6)
$$\int_{S}^{T} r(t)w(t)dt + \int_{S}^{T} q(t)\dot{w}(t)dt = \langle \nabla \psi(\bar{x}(S), \bar{x}(T)), (w(S), w(T)) \rangle.$$

On the other hand, by integrating by parts, we have

$$\int_{S}^{T} r(t)w(t)dt = -\int_{S}^{T} \dot{w}(t) \int_{S}^{t} r(s)dsdt + \left\langle w(t), \int_{S}^{t} r(s)ds \right\rangle \bigg|_{S}^{T}.$$

Now we set

$$W_0^{1,2}([S,T];\mathbb{R}^n) = \{ w \in W^{1,2}([S,T];\mathbb{R}^n) \mid w(S) = w(T) = 0 \}.$$

Then (6) becomes, for all $w \in W_0^{1,2}([S,T];\mathbb{R}^n)$,

$$\int_{S}^{T} \dot{w}(t) \Big(q(t) - \int_{S}^{t} r(s) ds \Big) dt = 0.$$

By Dubois-Reymond Lemma ([1, p.42]), there exists a constant $c_0 \in \mathbb{R}^n$ such that

$$q(t) = c_0 + \int_S^t r(s)ds.$$

We define

(7)
$$p(t) := -c_0 - \int_S^t r(s) ds.$$

Then

$$\dot{p}(t) = -r(t)$$

and

$$q(t) = -p(t).$$

On the other hand, for all $w \in W^{1,2}([S,T];\mathbb{R}^n)$ such that w(S) = 0, we have by (6),

$$\left\langle \int_{S}^{T} r(t)dt, w(T) \right\rangle + \int_{S}^{T} \dot{w}(t) \left(q(t) - \int_{S}^{t} r(s)ds \right) dt$$
$$- \left\langle \nabla_{2} \psi(\bar{x}(S), \bar{x}(T)), w(T) \right\rangle$$
$$= \left\langle \int_{S}^{T} r(t)dt + c_{0} - \nabla_{2} \psi(\bar{x}(S), \bar{x}(T)), w(T) \right\rangle$$
$$= 0$$

Therefore we have

(8)
$$\int_{S}^{T} r(t)dt = \nabla_2 \psi(\bar{x}(S), \bar{x}(T)) - c_0.$$

This implies that

$$\nabla_2 \psi(\bar{x}(S), \bar{x}(T)) = -p(T),$$

where ∇_2 denote the derivative with respect to second variable. Similarly, for all $w \in W^{1,2}([S,T];\mathbb{R}^n)$ such that w(T) = 0, we have by (6)

$$\nabla_1 \psi(\bar{x}(S), \bar{x}(T)) = p(S),$$

where ∇_1 denote the derivative with respect to first variable.

Now, recall that $(r,q) \in L^+$. For all $v \in L^2([S,T];\mathbb{R}^n)$ which verifies

$$v(t) \in T_{f(t,\bar{x}(t),U(t))}(\dot{\bar{x}}(t))$$

we have

$$(0,v)\in L$$
.

Therefore

(9)
$$\langle (r,q), (0,v) \rangle = \int_{S}^{T} \langle q(t), v(t) \rangle dt \ge 0.$$

From the measurable selection theorem and the fact that $f(s, \bar{x}(s), U(s)) - \dot{\bar{x}}(s) \subset T_{f(s,\bar{x}(s),U(s))}(\dot{\bar{x}}(s))$, we deduce that

$$\sup\{\langle -q(s), \xi \rangle | \xi \in f(s, \bar{x}(s), U(s)) - f(s, \bar{x}(s), \bar{u}(s))\} \le 0$$
 a.e.

On the other hand,

$$f(s, \bar{x}(s), \bar{u}(s)) \in f(s, \bar{x}(s), U(s)),$$

hence

$$\max\{\langle -q(s),\xi\rangle|\xi\in f(s,\bar{x}(s),U(s))-f(s,\bar{x}(s),\bar{u}(s))\}=0\quad\text{a.e.},$$

therefore we obtain the maximum principle:

$$\langle -q(s), f(s, \bar{x}(s), \bar{u}(s)) \rangle = \max_{u \in U(s)} \langle -q(s), f(s, \bar{x}(s), u) \rangle \quad \text{a.e.}$$

i.e.,

$$\left\langle p(s), f(s, \bar{x}(s), \bar{u}(s)) \right\rangle = \max_{u \in U(s)} \left\langle p(s), f(s, \bar{x}(s), u) \right\rangle \quad \text{a.e.}$$

Since

$$0 \in T_{f(s,\bar{x}(s),U(s))}(\dot{\bar{x}}(s)) \quad \forall s \in [S,T],$$

we have

$$\left(w, \frac{\partial f}{\partial x}(\cdot, \bar{x}(\cdot), \bar{u}(\cdot))w\right) \in L \quad \forall w \in L^2([S, T]; \mathbb{R}^n).$$

Since $(r,q) \in L^+$, we have for all $w \in L^2([S,T]; \mathbb{R}^n)$,

$$\left\langle (r,q), (w, \frac{\partial f}{\partial x}(\cdot, \bar{x}(\cdot), \bar{u}(\cdot))w) \right\rangle = \left\langle r + \frac{\partial f}{\partial x}(\cdot, \bar{x}(\cdot), \bar{u}(\cdot))^*q, w \right\rangle \geq 0.$$

Therefore

$$r(t) = -\frac{\partial f}{\partial x}(t,\bar{x}(t),\bar{u}(t))^*q(t) \ \text{ a.e.,}$$

i.e.,

$$-\dot{p}(t) = \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))^* p(t)$$
 a.e.

Next we return to our problem with endpoint constraint.

Theorem 3.3. Suppose that (\bar{x}, \bar{u}) is optimal to the endpoint constrained optimal control. Then there exists an absolutely continuous function p such that

i) (Adjoint equation)

$$-p'(t) = \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))^* p(t) \quad a.e. \quad in \quad [S, T]$$

ii) (Maximum principle)

$$\max_{u \in U(t)} \left\langle p(t), f(t, \bar{x}(t), u) \right\rangle = \left\langle p(t), f(t, \bar{x}(t), \bar{u}(t)) \right\rangle \quad a.e. \quad in \quad [S, T]$$

iii) (Transversality condition)

$$(p(S), -p(T)) \in \nabla \psi(\bar{x}(S), \bar{x}(T)) + N_C(\bar{x}(S), \bar{x}(T)).$$

Proof. Note that there exists $\delta > 0$ such that the process (\bar{x}, \bar{u}) is optimal with respect to all feasible processes (x, u) satisfying $||x - \bar{x}||_{L^{\infty}} < \delta$. We consider the following perturbated problem P(a) for some $a \in \mathbb{R}^n \times \mathbb{R}^n$:

$$\begin{array}{ll} \text{minimize} & \psi(x(S),x(T)) \\ \text{subject to} & \dot{x}(t) = f(t,x(t),u(t)) \text{ a.e. } t \in [S,T] \\ & u(t) \in U(t) \text{ a.e. } t \in [S,T] \\ & (x(S),x(T)) \in C + \{a\} \\ & \|x-\bar{x}\| < \delta. \end{array}$$

Denote the infimal cost of P(a) by V(a) which is called the value function. The process (\bar{x}, \bar{u}) is a minimizer for P(0), in other words,

(10)
$$V(0) = \psi(\bar{x}(S), \bar{x}(T)).$$

By [4, p.121], the value function V is lower semi-continuous at $0 \in \mathbb{R}^n \times \mathbb{R}^n$.

Since V is lower semi-continuous and $V(0) < \infty$, V has a proximal subdifferential ζ at 0. This means that there exist $\alpha > 0$ and M > 0 such that for all e satisfying $||e|| < \alpha$,

(11)
$$V(e) - V(0) \ge \zeta \cdot e - M ||e||^2.$$

Define

$$J((x,u),c) = \psi(x(S),x(T)) - \zeta \cdot ((x(S),x(T)) - c) + M \|(x(S),x(T)) - c\|^{2}.$$

Since $(x(S), x(T)) \in C + ((x(S), x(T)) - c \text{ for any } c \in C,$

(12)
$$\psi(x(S), x(T)) \ge V((x(S), x(T)) - c).$$

We have from (11), (12), and (10)

$$J((x,u),c) - J((\bar{x},\bar{u}),(\bar{x}(S),\bar{x}(T)))$$

$$> V((x(S), x(T)) - c) - V(0) - \zeta \cdot ((x(S), x(T)) - c) + M ||(x(S), x(T)) - c||^2$$

for all $c \in C$ and all (x, u) satisfying

$$\begin{split} \dot{x}(t) &= f(t, x(t), u(t)) \quad \text{a.e.} \quad t \in [S, T] \\ u(t) &\in U(t) \quad \text{a.e.} \quad t \in [S, T] \\ \|x - \bar{x}\| &< \min\{\delta, \alpha\}. \end{split}$$

Set
$$(x, u) = (\bar{x}, \bar{u})$$
 in (3). Then for all $c \in \{c \in C | \|c - (\bar{x}(S), \bar{x}(T))\|^2 < \alpha\}$

$$J((\bar{x}, \bar{u}), c) - J((\bar{x}, \bar{u}), (\bar{x}(S), \bar{x}(T))) > 0$$

and so

$$-\zeta \cdot ((\bar{x}(S), \bar{x}(T)) - c) \le M \|(\bar{x}(S), \bar{x}(T)) - c\|^2.$$

This inequality implies that

$$-\zeta \in N_C^P(\bar{x}(S), \bar{x}(T)).$$

Now set $c = (\bar{x}(S), \bar{x}(T))$ in (3), Then by (11)

$$J((x, u), (\bar{x}(S), \bar{x}(T))) - J((\bar{x}, \bar{u}), (\bar{x}(S), \bar{x}(T))) > 0.$$

We see that the process (\bar{x}, \bar{u}) is a minimizer for the following problem:

$$\begin{array}{ll} \text{minimize} & \psi(x(S), x(T)) - \zeta \cdot ((x(S), x(T)) - (\bar{x}(S), \bar{x}(T))) \\ & + M \| (x(S), x(T)) - (\bar{x}(S), \bar{x}(T)) \|^2 \\ \text{subject to} & \dot{x}(t) = f(t, x(t), u(t)) \text{ a.e. } t \in [S, T] \\ & u(t) \in U(t) \text{ a.e. } t \in [S, T] \\ & \|x - \bar{x}\| < \min\{\delta, \alpha\}. \end{array}$$

This is an endpoint constraint-free problem. We can so apply the above Proposition 3.2 to obtain the conclusions.

References

- [1] L. Cesari, Optimization-Theory and Applications, Problems with ordinary differential equations. Applications of Mathematics (New York), 17. Springer-Verlag, New York, 1983.
- [2] H. Frankowska, Contingent cones to reachable sets of control systems, SIAM J. Control Optim. 27 (1989), no. 1, 170-198.
- [3] K. E. Kim, Necessary conditions for optimal control problem under state constraints, J. Korean Math. Soc. 42 (2005), no. 1, 17–35.
- [4] P. D. Loewen, Optimal Control via Nonsmooth Analysis, CRM Proceeding & Lecture Notes, American Mathematical Society, 1993.
- [5] R. T. Rockafellar and R. J.-B. Wets, Variational Analysis, Grundlehren der Mathematischen Wissenschaften, 317. Springer-Verlag, Berlin, 1998.
- [6] R. B. Vinter, Optimal Control Birkhäuser, Boston, 2000.

DEPARTMENT OF MATHEMATICS

SCHOOL OF THE LIBERAL ARTS

SEOUL NATIONAL UNIVERSITY OF TECHNOLOGY

SEOUL 139-743, KOREA

E-mail address: kimke@snut.ac.kr