NONDEGENERATE NECESSARY CONDITIONS OF OPTIMALITY FOR IMPULSIVE CONTROL PROBLEMS

Fernando Lobo Pereira
Electrical and Computer Engineering Department
Faculdade de Engenharia, Universidade do Porto, Portugal, **flp@fe.up.pt**

Joint research with **A. Arutyunov**[1], **V. Dykhta**[2]
[1]Peoples Friendship University of Russia, Moscow, Russia, **arutun@orc.ru**
[2]Baikal State University of Economy and Law, Irkutsk, Russia, **dykhta@isea.ru**

XXVI CNMAC
26° Congresso de Matemática Aplicada e Computacional
September 8-11, 2003
S. José do Rio Preto, Brazil

Organization of the Presentation

- Introduction
- Statement of the Problem
- Definitions
- Necessary Conditions of Optimality
- Examples
- Outline of the Proof
- Extremal Principle
- Conclusions

Introduction

First and second order necessary conditions are provided.

Main feature: informative for abnormal control processes without a priori normality assumptions.

The proof is based on an extremal principle derived for an abstract minimization problem with equality and inequality type constraints and constraints given by convex cone.

Degeneracy

Let $x^* \in \Re^n$ be a solution to

$$(P_0) \qquad \text{Minimize } f(x) \text{ subject to } W(x) = 0.$$

If $W_x(x^*)$ is not **surjective**, then $\exists \bar{\lambda} \neq 0$ such that:
$\lambda = \text{col}(\lambda_0, \bar{\lambda})$ with $\lambda_0 = 0$ satisfies

$$L_x(x^*, \lambda) = \lambda_0 f_x(x^*) + \langle \bar{\lambda}, W_x(x^*) \rangle = 0.$$

Additional information is sought in second order conditions.
By restricting multipliers so that

$$\text{Index } \Omega^{\lambda}_{x^*}(\delta x) \text{ on Ker } W_x(x^*) \leq \text{ codim Im } W_x(x^*)$$

it is guaranteed that

$$\Omega^{\lambda}_{x^*}(\delta x) \geq 0 \text{ on Ker } W_x(x^*).$$

Applications

Well known applications arise in:

- Aerospace Navigation
- Resources Management
- Friction and Vibro-Impact Mechanics

- D. Lawden, *Optimal Trajectories for Space Navigation*, Butterworths, 1963.
- C. Clark, F. Clarke, G. Munro, "The Optimal Exploitation of Renewable Stocks", Econometrica, 47, 1979, pp. 25-47.
- B. Brogliato, *Nonsmooth Impact Mechanics: Models, Dynamics and Control*, Lect. Notes in Control and Inform. Sci., 220, Springer-Verlag, 1996

The Optimal Control Problem

(**P**) Minimize $J(x_0, u, \mu)$

subject to $dx(t) = f(t, x(t), u(t))dt + \mathbf{G(t, x(t))d\mu(t)},\ t \in [t_0, t_1]$,

$W_1(a) \leq 0,\ \ W_2(a) = 0$,

$\mathbf{d}\mu \in \mathcal{K}$,

where

$a = (x_0, x_1)$, with $x(t_0) = x_0$, and $x(t_1) = x_1$, for some $t_0 < t_1$.

$J(x_0, u, w) := W_0(a)$.

$W_i : R^n \times R^n \to R^{d(W_i)}$, $i = 0, 1, 2$, ($d(W_i)$ is the dimension of W_i, $d(W_0) = 1$).

$f : [t_0, t_1] \times R^n \times R^m \to R^n$, $G : [t_0, t_1] \times R^n \to R^{n \times k}$.

$\mathcal{K} = \{\mu \in C^*([t_0, t_1]; R^k) : \forall$ continuous ϕ s. t. $\phi(t) \in K^{\oplus}\ \forall t$

$\int_B \phi(t) d\mu \geq 0\ \forall$ Borel $B \subset [t_0, t_1]\}$,

K is a given convex, closed, pointed cone from R^k and $K^{\oplus}$ its dual.

Assumptions

(H1) W_0, W_1, and W_2 are twice continuously differentiable.

(H2) f is twice differentiable with respect to x and u for almost all $t \in [t_0, t_1]$.

(H3) f and its first and second order derivatives are measurable with respect to t and bounded on any bounded subset.

(H4) G is continuous in time and twice differentiable in x.

(H5) K is a closed convex pointed cone in $\Re^q$.

(H6) The matrix G satisfies the so called Frobenius condition, i.e.,

$$\mathbf{G^i_x(t,x)G^j(t,x) - G^j_x(t,x)G^i(t,x) \equiv 0.}$$

Dynamics Interpretation

The function $x \in BV(t_0, t_1)$ is a solution if

$$x(t) = x_0 + \int_{t_0}^{t} f(\theta, x(\theta), u(\theta))d\theta + \int_{t_0}^{t} G(\theta, x(\theta))d\mu_c(\theta) + \sum_{s_i \le t}(z(1; s_i, c^i) - x(s_i^-)), \qquad t > t_0$$

where

- $d\mu(t) = d\mu_c(t) + \sum c^i \delta_{s_i}(t)$, and
- functions $z^i(\tau) = z(\tau; s_i, c_i)$ are solutions to the limiting system

$$\mathbf{\frac{dz^i}{d\tau} = G(s_i, z^i)c_i, \quad z^i(0; s_i, c_i) = x(s_i^-).}$$

Observation: Due to $(H6)$, the solution is unique for any given control (u, μ).

Definitions

Admissible control - It is a pair (u, μ), where $u \in L_\infty^m[t_0, t_1]$ and $\mu \in \mathcal{K}$.

Admissible control process - It is a triple (x_0, u, μ) such that the corresponding trajectory x, defined by the integral equation, satisfies the endpoint constraints.

Local minimizer to (P) - It is an admissible process (x_0^*, u^*, μ^*) satisfying:

$\exists \varepsilon > 0$, and, $\forall$ finite-dimensional subspace $R \subset L_\infty^m[t_0, t_1]$, $\exists \varepsilon_R > 0$, s. t. (x_0^*, u^*, μ^*) is a solution to (P) with additional constraints:

$$\|a - a^*\| < \varepsilon, \quad \|\mu - \mu^*\|_{C^*([t_0,t_1];R^k)} < \varepsilon,$$

$$\|u - u^*\|_{L_\infty^m[t_0,t_1]} < \varepsilon_R, \quad u(\cdot) \in R.$$

Notation

Let $\lambda = (\lambda^0, \lambda^1, \lambda^2)$ be such that

- $\lambda^0 \in R^1$
- $\lambda^1 = (\lambda_1^1, \ldots, \lambda_{d(W_1)}^1) \in R^{d(W_1)}$
- $\lambda^2 \in R^{d(W_2)}$

and ψ be a n-dimensional vector.

- $H = H_0 + H_1$ is the **Pontryagin function** or **Pseudo-Hamiltonian** defined by

$$H_0(t, x, \psi, u) = \langle \psi, f(t, x, u) \rangle$$

$$H_1(t, x, \psi, v) = \langle \psi, G(t, x)v \rangle$$

- l is the **Endpoint Lagrangian** defined by

$$l(\lambda, a) = \lambda_0 W_0(a) + \langle \lambda_1, W_1(a) \rangle + \langle \lambda_2, W_2(a) \rangle$$

Local Maximum Principle

(x_0^*, u^*, μ^*) satisfies the **Euler-Lagrange conditions**

or the **local Maximum Principle** if $\exists \lambda \neq 0$, s.t.

$$\lambda_0 \geq 0,\ \lambda_1 \geq 0,\ \langle \lambda_1, W_1(a^*) \rangle = 0,$$

and a vector function ψ, solution to the adjoint system:

$$\begin{aligned} -d\psi(t) &= H_{0x}(t, x^*(t), \psi(t), u^*(t))dt + (H_{1x})_v(t, x^*(t), \psi(t), \omega^*(t))d\mu^*(t) \\ (-\psi(t_1), \psi(t_0)) &= l_a(a^*, \lambda) \end{aligned}$$

such that

$$\begin{aligned} H_u(t, x^*(t), \psi(t), u^*(t), \omega^*(t)) &= 0 && dt\text{-a.e.} \\ \langle H_v(t, x^*(t), \psi(t), u^*(t), \omega^*(t)), v \rangle &\leq 0 && \forall (t, v) \in [t_0, t_1] \times K \\ \langle H_v(t, x^*(t), \psi(t), u^*(t), \omega^*(t)), \omega^*(t) \rangle &= 0 && d\mu_c^* - \text{a.e.} \\ \langle H_v(t, z_t(s), q_t(s), u^*(t), \omega^*(t)), \omega^*(t) \rangle &= 0 && [0,1] - \text{a.e. } \forall t \in S_d. \end{aligned}$$

Local Maximum Principle (cont.)

Here,

- $a^* = (x^*(t_0), x^*(t_1))$
- $\omega^*(t) = \frac{d\mu^*(t)}{d|\mu^*(t)|}$ in the sense of Radon-Nicodym.
- S_d is the support of the atomic component of the optimal control measure.
- $\begin{cases} \dot{z}^t(s) = G(t, z^t(s))\omega^*(t) & [0,1] - \text{a.e.}, \ \forall t \in S_d \\ z^t(0) = x(t^-) \end{cases}$
- $\begin{cases} -\dot{q}^t(s) = (H_{1x})_v(t, z^t(s), q^t(s), \omega^*(t))\omega^*(t) & [0,1] - \text{a.e.}, \ \forall t \in S_d \\ q^t(1) = \psi(t) \end{cases}$

Local Maximum Principle (cont.)

Denote the set of all normalized (say, $\|\lambda\| = 1$) Lagrange multipliers λ satisfying the Local Maximum Principle by

$$\Lambda(x_0^*, u^*, w^*).$$

First order necessary condition for a weak local minimum for (P):

$$\Lambda(x_0^*, u^*, w^*) \neq \emptyset.$$

For short, we denote $\Lambda(x_0^*, u^*, w^*)$ by Λ.

The Critical Cone

$\mathcal{K}_{cr}$, the **Cone of Critical Variations**, is the set of all variations

$$(\delta x_0, \delta u, \delta\mu) \in R^n \times L_\infty^m \times BV^k$$

with state trajectories $\delta x \in BV^n(S_d)$, satisfying:

$$\langle W_{ia}(a^*), \delta a\rangle \begin{cases} \leq 0, & i = 0, 1, \\ = 0, & i = 2 \end{cases}$$

$$\delta a = (\delta x(t_0), \delta x(t_1)),$$

$$d(\delta x) = [f_x(t)\delta x + f_u(t)\delta u]dt + \left(\sum_{i=1}^{k} g_x^i(t) d\mu_i^*(t)\right) \delta x + G(t)d(\delta\mu)$$

$$d(\delta\mu) \in T_{\mathcal{K}}(d\mu^*) = \mathcal{K} + \mathrm{Lin}\{d\mu^*\}, \qquad \delta\mu(t_0) = 0$$

$$\delta x(t) = \delta q(1; t, \mu^*(\{t\})), \qquad \forall t \in S_d.$$

Notation

Here,

- Given some function $Q(t, y, z)$, $Q(t, y)$ denotes $Q(t, y, z^*(t))$.
- $\delta q(\tau; t, \mu^*(\{t\})) := \delta q^t(\tau)$ is the solution to

$$\begin{cases} \frac{d(\delta q^t)}{d\tau} = H_{1\psi x}(t, z^t(\tau), \mu^*(\{t\}))\delta q^t \\ \delta q^t(0) = \delta x(t^-) & t > t_0. \\ \delta q^{t_0}(0) = \delta x_0 \end{cases}$$

- $z^t(\tau)$ is the solution to

$$\begin{cases} \frac{dz^t}{d\tau} = G(t, z^t(\tau))\mu^*(\{t\}) \\ z^t(0) = x(t^-). \end{cases}$$

IMPULSIVE CONTROL

The Quadratic Form

For any $\lambda \in \Lambda$ define the quadratic form

$$\Omega^\lambda(\delta x_0, \delta u, \delta\mu) = \delta a^T l_{aa}(a^*, \lambda)\delta a + Q_1^\lambda(\delta a, \delta\mu_1) - \int_{t_0}^{t_1} Q^\lambda(\delta x, \delta u, \delta\mu)(t)dt$$

where Q^λ and Q_1^λ are the following quadratic forms:

$$Q^\lambda(\delta x, \delta u, \delta\mu) = \delta u^T H_{uu}^\lambda \delta u + 2\delta x^T H_{xu}^\lambda \delta u - 2\delta\mu^T(\dot{H}_v^\lambda)_u \delta u - \delta\mu^T(\ddot{H}_v^\lambda)_v \delta\mu - 2\delta\mu^T(\dot{H}_v^\lambda)_x \delta x + \delta x^T H_{xx}^\lambda \delta x$$

$$Q_1^\lambda(\delta x(\cdot), \delta\mu_1) = 2\delta x(t_0)^T l_{x_0 x_1}^\lambda(a^*)G(t_1)\delta\mu_1 - 2\delta x(t_1)^T H_{xv}^\lambda(t_1)\delta\mu_1 + \delta\mu_1^T G^T(t_1)[L_{x_1x_1}^\lambda(a^*)G(t_1) - H_{xv}^\lambda(t_1)]\delta\mu_1 - \sum_{s \in S_d}[\delta x^T(s)\Psi^\lambda(s)\delta x(s) - \delta x^T(s^-)\Psi^\lambda(s^-)\delta x(s^-)].$$

The Quadratic Form (cont.)

Here,

$$\Psi^\lambda(t) = -Z^T(1;t)\left(\int_0^1 Z^{-1T}(\tau,t)H_{1xx}(\tau;t)Z^{-1}(\tau;t)d\tau\right)Z(1;t)$$

where the $n \times n$ matrix $Z(\tau;t)$ satisfies

$$-\frac{dZ}{d\tau} = ZH_{1\psi x}(t, z^*(\tau;t), q^*(\tau;t), \mu^*(t)), \qquad Z(0;t) = I.$$

Observation. No limits extraction to compute, for example, $\Psi^\lambda(t^-)$.
We just need to solve first

$$\begin{cases} \frac{dz^*}{d\tau} = G(t,z^*)(\mu^*(t) - \mu^*(\{t\})), & z^*(1;t) = x^*(t) \\ -\frac{dq^*}{d\tau} = H_{1x}(t, z^*, q^*, \mu^*(t) - \mu^*(\{t\})), & q^*(1;t) = \Psi^\lambda(t) \end{cases}$$

and then $Z(\tau, t^-)$ is obtained by solving

$$-\frac{dZ}{d\tau} = ZH_{1\Psi x}(t, z^*(\tau,t), \mu^*(t) - \mu^*(\{t\})), \quad Z(0;t) = I.$$

Λ_d

Let $d := \text{codimIm}(\mathcal{A})$ where

- $\mathcal{A} : R^n \times L_\infty^m \times L_\infty^k \times R^k \to \Re^{d(W)}$ defined by

$$\mathcal{A}(\delta x(0), \delta u, \delta\mu, h) := W_{x_0}(a^*)\delta x_0 + W_{x_1}(a^*)[\delta x_1 + G(t_1)\pi h]$$

- $\pi : \Re^k \to N := K \bigcap (-K)$ $\quad$ ($C^*([t_0, t_1], N)$ is the maximal subspace of $\mathcal{K}$)
- $\mathcal{K}_\pi := \{(\delta x(0), \delta u, \delta\mu, h) \in Ker(\mathcal{A})$: solution to

$$\dot{\delta x} = F_x(t)\delta x + F_u(t)\delta u - (\dot{H}_v)_\psi(t)\pi\delta\mu, \quad t \notin S_d\}.$$

Then $\qquad \Lambda_d := \{\lambda \in \Lambda : \text{ Index}(\Omega^\lambda) \text{ on } \mathcal{K}_\pi \le d\}.$

Observation: d is the dimension of the kernel of $\left[A^T|B^T|G(t_1)\pi^T\right]^T$, where

$$A = W_{x_0}(a^*) + \Phi(t_1)W_{x_1}(a^*)$$

$$B = W_{x_1}(a^*)^T\Phi(t_1)\int_{t_0}^{t_1}\Phi^{-1}(t)\Gamma(t) \times \Gamma(t)^T\Phi^{-1}(t)^T dt\Phi(t_1)^T W_{x_1}(a^*).$$

Here, $\Gamma(t) = [F_u(t)| - (\dot{H}_v)_\psi(t)\pi]$ and Φ is the solution to $\dot{\Phi} = F_x(t)\Phi$, $\Phi(0) = I$.

The Main Result

Main Theorem (Necessary conditions of optimality). Let the control process (x^*, u^*, μ^*) be a local optimal to the problem (MP). Then, $\Lambda_d \neq \emptyset$ and, for any $(\delta x_0, \delta u, \delta\mu) \in \mathcal{K}_{cr}$, we have

$$\max_{\lambda\in\Lambda_d} \Omega^\lambda(\delta x_0, \delta u, \delta\mu) \geq 0.$$

Here,

- $\Omega^\lambda(\delta x_0, \delta u, \delta\mu)$ is a quadratic form
- Λ_d is an appropriate set of multipliers
- $\mathcal{K}_{cr}$ is the cone of critical variations

as defined in previous slides.

Example 1

Let $t_0 = 0$, $t_1 = 1$, $u \in L^1([0,1], R^1)$, $x = \text{col}(x_1, x_2, x_3, x_4) \in R^4$ and $K = R^+ \times R^+$.

$$\begin{aligned} \text{Minimize} \quad & x_4(1) \\ \text{subject to} \quad & dx_1 = x_1 dt + d\mu_1, \quad \dot{x}_2 = x_1 + u, \\ & dx_3 = (x_1 - e)^2 dt - d\mu_2, \quad dx_4 = u^2 dt + d\mu_1 + d\mu_2, \\ & x(0) = 0, \quad x_3(1) = 0, \text{ and } x_1(1) = x_2(1). \end{aligned}$$

The optimal control process is:

$$u^*(t) = \alpha \ \forall t \in [0,1], \quad d\mu^*(t) = (\alpha\delta_0(t), d\mu_2^*(t))$$

$$x_1^*(t) = \alpha e^t, \qquad x_3^*(t) = \frac{1}{2}\alpha^2(e^{2t} - 1) - 2\alpha e(e^t - 1) + e^2 t - \int_{[0,t]} d\mu_2^*(s)$$

$$x_2^*(t) = \alpha((e^t - 1) + t), \qquad x_4^*(t) = \alpha^2 t + \alpha + \int_{[0,t]} d\mu_2^*(s)$$

where $\alpha = \frac{2e^2 - 2e - 1}{e^2 + 1}$, and $\int_{[0,1]} d\mu_2^*(t) = \frac{1}{2}\alpha^2(e^2 - 1) - 2\alpha e(e - 1) + e^2$.

Example 1 (cont.)

- $H = \psi_1 x_1 + \psi_2(x_1 + u) + \psi_3(e - x_1)^2 + \psi_4 u^2$
- $l = \lambda_1(x_1(1) - x_2(1)) + \lambda_2 x_3(1) + \lambda_0 x_4(1)$

Clearly $\psi(1) = \text{col}(-\lambda_1, \lambda_1, -\lambda_2, -\lambda_0)$, $\psi_i(t) = \psi_i(1)$, $i = 2, 3, 4$, and $\psi_1(t) = -\lambda_1 e^{1-t} + \int_t^1 e^{-(t-s)}(\lambda_1 + 2\lambda_2 e - 2\lambda_2 \alpha e^s)ds$.

From $\frac{\partial H}{\partial u} = 0$, we have $u^*(t) \equiv \frac{\lambda_1}{2\lambda_0}$ and, thus, $\alpha = \frac{\lambda_1}{2\lambda_0}$.

From $\psi(t)G(t) = (\psi_1(t) - \lambda_0, \lambda_2 - \lambda_0)$ we have:

- $\lambda_2 = \lambda_0$ as $d\mu_2^*$ has support on any subset of $[0, 1]$.
- $\psi_1(0) = \lambda_0$ as $\dot{\psi}_1(t) < 0$, $\forall t \in [0, 1)$, and the first order conditions are satisfied.

This is achieved for $\alpha = \frac{2e^2 - 2e - 1}{e^2 + 1}$. Conclusion: $\Lambda = \{(1, 2\alpha, 1)\}$.

$\Omega_\lambda(\delta x(0), \delta u) = \int_0^1([\delta x_1]^2 + [\delta u]^2)dt \geq 0$ for $\lambda = (1, 2\alpha, 1)$ and $\Lambda_d = \Lambda$.

Example 2

Take $x, u \in R^n$, $y \in R^k$, $z \in R^1$, $K = R^+$

$$\begin{array}{ll} \text{Minimize} & z(0) \\ \text{subject to} & \dot{x} = zu, \quad dy = z(2Q[x,u] + a)dt - ad\mu, \quad \dot{z} = 0, \quad t \in [0,1], \\ & x(0) = 0, \quad y(0) = -a, \text{ and } y(1) = 0. \end{array}$$

Here, $a \notin Q(R^n)$ ($Q(x) = Q[x, x]$), and $Q : R^n \times R^n \to R^k$ is a bilinear symmetric mapping s.t.:

a) $\exists y \in R^k$ s.t. $Q(x) \neq y \ \forall x \in R^n$.
b) There is no $\lambda \in R^k$, s.t.$\langle \lambda, Q(x) \rangle \geq 0 \ \forall x \in R^n$.

Optimality of $(x^*(t), y^*(t), z^*(t), u^*(t), \mu^*) = (0, ta, 1, 0, 0)$:

$$y(1) = \int_0^1 \frac{d}{dt}Q(x)dt + a(z - 1 - \int_0^1 d\mu) = a(z - 1 - \int_0^1 d\mu) + Q(x(1)) = 0.$$

Since $a \notin Q(R^n)$, we have $z \geq 1$ and, as z is to be minimized, $\mu \equiv 0$ and $z = 1$.