Lecture on the Pontryagin Maximum Principle

Soledad Aronna

Università degli Studi di Padova

Young Researchers Workshop
on System Dynamics and Optimal Control
-
Madeira, Portugal. January 2013

ITN SADCO
Initial Training Network
Sensitivity Analysis for Deterministic Controller Design

Università degli Studi di Padova

The problem with free terminal point

Let $[0, T] \subset \mathbb{R}$.

Consider the optimal control problem (P) in the *Mayer form*

$$\max_{u \in \mathcal{U}} \psi(x(T)),$$
$$\dot{x}(t) = f(t, x(t), u(t)),$$
$$x(0) = x_0,$$

where

$$\mathcal{U} := \{u : [0, T] \to U \text{ measurable}\},$$

with $U \subseteq \mathbb{R}^m$ compact set given,

$$x : [0, T] \to \mathbb{R}^n, \quad \psi : \mathbb{R}^n \to \mathbb{R}, \quad f : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}^n.$$

$u^*(\cdot) \in \mathcal{U}$ is an *optimal control* of (P) if

$$\psi(x^*(T)) \geq \psi(x(T)).$$

What is the Pontryagin Maximum Principle?

Necessary condition for optimality: it gives conditions that have to be satisfied by $(x^*(\cdot), u^*(\cdot))$.

Existence of a multiplier: an absolutely continuous function

$$p : [0, T] \to \mathbb{R}^{n,*},$$

that

1. is solution of the *adjoint equation*,
2. verifies a *maximum condition*.

$u^*(\cdot) \in \mathcal{U}$ is an *optimal control* of (P) if

$$\psi(x^*(T)) \geq \psi(x(T)).$$

What is the Pontryagin Maximum Principle?

Necessary condition for optimality: it gives conditions that have to be satisfied by $(x^*(\cdot), u^*(\cdot))$.

Existence of a multiplier: an absolutely continuous function

$$p : [0, T] \to \mathbb{R}^{n,*},$$

that

1. is solution of the *adjoint equation*,
2. verifies a *maximum condition*.

$u^*(\cdot) \in \mathcal{U}$ is an *optimal control* of (P) if

$$\psi(x^*(T)) \geq \psi(x(T)).$$

What is the Pontryagin Maximum Principle?

Necessary condition for optimality: it gives conditions that have to be satisfied by $(x^*(\cdot), u^*(\cdot))$.

Existence of a multiplier: an absolutely continuous function

$$p : [0, T] \to \mathbb{R}^{n,*},$$

that

1. is solution of the *adjoint equation*,
2. verifies a *maximum condition*.

$u^*(\cdot) \in \mathcal{U}$ is an *optimal control* of (P) if

$$\psi(x^*(T)) \geq \psi(x(T)).$$

What is the Pontryagin Maximum Principle?

Necessary condition for optimality: it gives conditions that have to be satisfied by $(x^*(\cdot), u^*(\cdot))$.

Existence of a multiplier: an absolutely continuous function

$$p : [0, T] \to \mathbb{R}^{n,*},$$

that

1. is solution of the *adjoint equation*,
2. verifies a *maximum condition*.

Statement of the PMP

(H1) $f : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}^n$ is continuous in (t, x, u), continuously differentiable with respect to x and

$$|f(t,x,u)| \le C, \quad \|D_x f(t,x,u)\| \le L, \quad \text{for all } (t,x,u) \in \mathbb{R}\times\mathbb{R}^n\times U. \tag{1}$$

$\psi : \mathbb{R}^n \to \mathbb{R}$ is differentiable.

Theorem (Pontryagin Maximum Principle)

Assume (H1). Let $u^*(\cdot)$ be an optimal control for (P), and $x^*(\cdot)$ be its associated trajectory. Denote $p : [0, T] \to \mathbb{R}^{n,*}$ the solution of the *adjoint equation*

$$\dot{p}(t) = -p(t) D_x f(t, x^*(t), u^*(t)), \quad p(T) = \nabla\psi(x^*(T)).$$

Then the *maximum condition*

$$p(t) f(t, x^*(t), u^*(t)) = \max_{w \in U} p(t) f(t, x^*(t), w),$$

holds for almost every $t \in [0, T]$.

Sketch of the proof

Hypothesis: Optimality of $u^*(\cdot)$.

Thesis: The solution $p(\cdot)$ of the adjoint equation satisfies the maximum condition.

Elements of the proof:

1. Variations of the control ('needle' or Weierstrass variations) and Lebesgue Differentiation Theorem.
2. The variational equation.
3. The adjoint equation.

Sketch of the proof

Hypothesis: Optimality of $u^*(\cdot)$.

Thesis: The solution $p(\cdot)$ of the adjoint equation satisfies the maximum condition.

Elements of the proof:

1. Variations of the control ('needle' or Weierstrass variations) and Lebesgue Differentiation Theorem.
2. The variational equation.
3. The adjoint equation.

Lebesgue Differentiation Theorem

Theorem

Let $h : [a, b] \to \mathbb{R}^N$ be integrable. Then, almost every $\tau \in [a, b]$ is a Lebesgue point, i.e.

$$\lim_{\varepsilon \to 0^+} \frac{1}{2\varepsilon} \int_{\tau-\varepsilon}^{\tau+\varepsilon} |h(\tau) - h(s)| \mathrm{d}s = 0.$$

The dynamics

Consider the ordinary differential equation

$$\dot{x}(t) = g(t, x(t)), \quad x(t_0) = x_0. \tag{ODE}$$

(H2) $g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ measurable in t, continuous in x and such that there $\exists$ C, L satisfying

$$|g(t,x)| \leq C, \quad |g(t,x) - g(t,y)| \leq L|x-y| \quad \text{for all } (t,x,y).$$

Theorem (Existence and uniqueness)

Assume (H2). Then, for all $T > t_0$, *the equation* (ODE) *has a global unique solution* $x(\cdot)$ *defined on* $[t_0, T]$.

The variational equation

Theorem (Differentiability respect to initial conditions)

Assume (H2) + 'g continuously differentiable w.r.t. x.'
Let $\hat{x}(t) := x(t; t_0, x_0)$ *be the solution of* (ODE). *For a vector* $v_0 \in \mathbb{R}^n$, *call* $v(\cdot)$ *the solution of the linear Cauchy problem*

$$\dot{v}(t) = D_x g(t, \hat{x}(t))v(t), \quad v(t_0) = v_0. \tag{VE}$$

Then,

$$v(t) = \frac{\partial}{\partial \varepsilon}\Big|_{\varepsilon=0} x(t; t_0, x_0 + \varepsilon v_0) \ \left(= \lim_{\varepsilon \to 0^+} \frac{x(t; t_0, x_0 + \varepsilon v_0) - \hat{x}(t)}{\varepsilon} \right).$$

(VE) is usually called the variational equation associated to (ODE)

Linear adjoint equations

Theorem

$v(\cdot)$ and $p(\cdot)$ solutions of

$$\begin{aligned}\dot{v}(t) &= A(t)v(t),\\ \dot{p}(t) &= -p(t)A(t),\end{aligned}$$

Then $p(t)\cdot v(t)$ is constant in time.

The proof - Part I: needle variations

For $\tau \in (0, T]$, $0 < \varepsilon < \tau$, and $w \in U$, define the needle variations

$$u_\varepsilon(t) := \begin{cases} w & \text{if } t \in [\tau - \varepsilon, \tau], \\ u^*(t) & \text{otherwise.} \end{cases}$$

$x_\varepsilon(\cdot)$ associated with $u_\varepsilon(\cdot)$.

We want to prove the maximum condition at $t = \tau$ (for a.a. τ) and for this arbitrary $w \in U$.

We aim to calculate $\frac{\partial}{\partial \varepsilon}\Big|_{\varepsilon = 0^+} \psi(x_\varepsilon(T))$, which we know is nonpositive.

$$\frac{x_\varepsilon(\tau) - x^*(\tau)}{\varepsilon} = \frac{1}{\varepsilon}\Big\{ \int_{\tau-\varepsilon}^{\tau} f(t, x_\varepsilon(t), w)\mathrm{d}t - \int_{\tau-\varepsilon}^{\tau} f(t, x^*(t), u^*(t))\mathrm{d}t \Big\}.$$

For the first integral:

$$\frac{1}{\varepsilon}\int_{\tau-\varepsilon}^{\tau} \big(f(t, x_\varepsilon(t), w) - f(\tau, x_\varepsilon(\tau), w)\big)\mathrm{d}t$$
$$\leq \frac{1}{\varepsilon}\int_{\tau-\varepsilon}^{\tau} \big(L|x_\varepsilon(t) - x_\varepsilon(\tau)| + f(t, x_\varepsilon(\tau), w) - f(\tau, x_\varepsilon(\tau), w)\big)\mathrm{d}t \underset{\varepsilon\to 0^+}{\longrightarrow} 0.$$

For the second integral: if τ is a Lebesgue point for $u^*(\cdot)$ then

$$\lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\int_{\tau-\varepsilon}^{\tau} f(t, x^*(t), u^*(t))\mathrm{d}t = f(\tau, x^+(\tau), u^*(\tau)).$$

Hence, Lebesgue Differentiation Theorem implies

$$\lim_{\varepsilon\to 0} \frac{x_\varepsilon(\tau) - x^*(\tau)}{\varepsilon} = f(\tau, x^*(\tau), w) - f(\tau, x^*(\tau), u^*(\tau)), \quad \text{for a.a. } \tau \in [0, T].$$

Part II: the variational equation

Let $v : [\tau, T] \to \mathbb{R}^m$ the solution of

$$\dot{v}(t) = D_x f(t, x^*(t), u^*(t)) v(t),$$

with initial condition

$$v(\tau) = \lim_{\varepsilon \to 0^+} \frac{x_\varepsilon(\tau) - x^*(\tau)}{\varepsilon} = f(\tau, x^*(\tau), w) - f(\tau, x^*(\tau), u^*(\tau)).$$

We know that

$$0 \geq \lim_{\varepsilon \to 0^+} \frac{\psi(x_\varepsilon(T)) - \psi(x^*(T))}{\varepsilon} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon = 0^+} \psi(x_\varepsilon(T)) = D\psi(x^*(T)) v(T).$$

Part III: the adjoint equation

From the last equation we have,

$$p(T)v(T) \leq 0.$$

Hence,

$$p(\tau)v(\tau) \leq 0,$$

which yields

$$p(\tau)[f(\tau, x^*(\tau), w) - f(\tau, x^*(\tau), u^*(\tau))] \leq 0,$$

and thus the maximum condition follows for a.a. $\tau \in (0, T]$ and for all $winU$.

□

The problem with terminal constraints

Consider now the problem (CP),

$$\max_{u \in \mathcal{U}} \; \psi(x(T)),$$
$$\dot{x}(t) = f(t, x(t), u(t)),$$
$$x(0) = x_0,$$
$$x(T) \in \mathcal{S},$$

with $\mathcal{S} \subset \mathbb{R}^n$.

Approximating cones

Definition

Let $\mathcal{S} \subset \mathbb{R}^n$, $s \in \mathcal{S}$ and K a convex cone in $\mathbb{R}^n$. We say that K is an approximating cone to $\mathcal{S}$ at s if there exists a neighbourhood $W \subset \mathbb{R}^n$ of 0, and a continuous map $G : W \cap K \to \mathcal{S}$, s.t.

$$G(v) = s + v + o(|v|)$$

as $K \ni v \to 0$.

Remark (Tangent cones of nonsmooth analysis)

Any convex cone of a Bouligand tangent cone, or of a Clarke tangent cone, or of a tangent cone of Convex Analysis, is an approximating cone.

Definition

Polar cone If $K \subseteq X$, with X a real vector space, then the polar cone is

$$K^{\perp} := \{p \in X^* : \left\langle p, w \right\rangle \geq 0, \quad \text{for all } \ w \in K\}.$$

The PMP with endpoint constraints

Theorem

Assume the hypotheses (H1). Let $u^*(\cdot)$ be an optimal control for problem (CP), and $x^*(\cdot)$ be its associated trajectory. Let C be an approximating cone to $\mathcal{S}$ at point $x^*(T)$. Then, there exists an absolutely continuous function $p : [0, T] \to \mathbb{R}^{n,*}$ and $\alpha \geq 0$ s.t.

$$\dot{p}(t) = -p(t) D_x f(t, x^*(t), u^*(t)),$$

with terminal transversality condition

$$p(T) - \alpha \nabla \psi(x^*(T)) \in C^{\perp},$$

the maximum condition holds

$$p(t) f(t, x^*(t), u^*(t)) = \max_{w \in U} p(t) f(t, x^*(t), w),$$

holds for almost all $t \in [0, T]$.

Sketch of the proof

1. Prove the maximum condition for a finite number of times and values in U :
   - 1.a Composition of finitely many needle variations.
   - 1.b Defining an approximating cone of the reachable set.
   - 1.c Using a set separation result that gives the transversality condition.
   - 1.d Transporting the maximum condition back in time.

2. By some topological arguments, extend the maximum condition to all U and a.a. $t \in [0, T]$.

References

- *Introduction to the mathematical theory of control*, by A. Bressan and B. Piccoli.
- *Lecture notes in optimal control*, by F. Rampazzo.
- *Geometry and Optimal Control*, by H. Sussmann.
- *The Pontryagin Maximum Principle*, lecture notes by F. Bonnans.
- *An introduction to mathematical control theory*, lecture notes by L. Evans.
- *The Mathematical Theory of Optimal Processes*, by L. Pontryagin and V. Boltyanski and R. Gamkrelidze and E. Michtchenko, 1962.

THANK YOU FOR YOUR ATTENTION.