technische universität dortmund

EPFL

Optimal Control (EE-736)

Part II.1: Optimal Control Theory

Timm Faulwasser & Yuning Jiang

ie3, TU Dortmund

timm.faulwasser@ieee.org
yuning.jiang@epfl.ch

Block course @ EPFL

Version EE736.2024.I

Overview

Overview

What is an Optimal Control Problem?

Adhoc definition of optimal control:

Determine the control signals that will cause a system to satisfy physical constraints
and—at the same time—minimize some performance criterion.

What is an Optimal Control Problem?

Adhoc definition of optimal control:

Determine the control signals that will cause a system to satisfy physical constraints
and—at the same time—minimize some performance criterion.

Generic Optimal Control Problem (OCP):

$$
\begin{aligned}
&\min_{u(\cdot)} J(u(\cdot)) \\
\text{subject to}& \\
&\dot{x}(t) = f(t, x(t), u(t)), \quad x(t_0) = x_0 \\
&\forall t \in [t_0, t_1] : u(t) \in \mathbb{U} \subseteq \mathbb{R}^{n_u} \\
&\forall t \in [t_0, t_1] : x(t) \in \mathbb{X} \subseteq \mathbb{R}^{n_x}
\end{aligned} \tag{OCP}
$$

What is an Optimal Control Problem?

Generic Optimal Control Problem (OCP):

$$
\begin{aligned}
&\min_{u(\cdot)} \; J(u(\cdot)) \\
&\text{subject to} \\
&\dot{x}(t) = f(t, x(t), u(t)), \quad x(t_0) = x_0 \\
&\forall t \in [t_0, t_1] : u(t) \in \mathbb{U} \subseteq \mathbb{R}^{n_u} \\
&\forall t \in [t_0, t_1] : x(t) \in \mathbb{X} \subseteq \mathbb{R}^{n_x}
\end{aligned}
\qquad \text{(OCP)}
$$

What is an Optimal Control Problem?

Generic Optimal Control Problem (OCP):

$$
\begin{aligned}
&\min_{u(\cdot)} J(u(\cdot)) \\
&\text{subject to} \\
&\dot{x}(t) = f(t, x(t), u(t)), \quad x(t_0) = x_0 \\
&\forall t \in [t_0, t_1] : u(t) \in \mathbb{U} \subseteq \mathbb{R}^{n_u} \\
&\forall t \in [t_0, t_1] : x(t) \in \mathbb{X} \subseteq \mathbb{R}^{n_x}
\end{aligned} \tag{OCP}
$$

Ingredients

- Dynamics?
- Class of input signals? Definition of state and input constraints?
- Performance criterion?

Useful Notions – Admissible Controls

Considered classes of input functions:

- Finitely many discontinuities on compact time intervals: $u : [t_0, t_1] \to \mathbb{R}^{n_u}$

$$u(\cdot) \in \hat{\mathcal{C}}\left([t_0, t_1], \mathbb{R}^{n_u}\right) = \hat{\mathcal{C}}[t_0, t_1]^{n_u}$$

Useful Notions – Admissible Controls

Considered classes of input functions:

- Finitely many discontinuities on compact time intervals: $u : [t_0, t_1] \to \mathbb{R}^{n_u}$

$$u(\cdot) \in \hat{\mathcal{C}}\left([t_0, t_1], \mathbb{R}^{n_u}\right) = \hat{\mathcal{C}}[t_0, t_1]^{n_u}$$

- Finitely many discontinuities on compact time intervals and bounded magnitude $\forall t \in [t_0, t_1] : \quad u(t) \in \mathbb{U}$, $\mathbb{U}$ is a compact subset of $\mathbb{R}^{n_u}$. Brief notation:

$$u(\cdot) \in \hat{\mathbb{U}}[t_0, t_1]^{n_u} \doteq \hat{\mathcal{C}}\left([t_0, t_1], \mathbb{U}\right)$$

Useful Notions – Admissible Controls

Considered classes of input functions:

- Finitely many discontinuities on compact time intervals: $u : [t_0, t_1] \to \mathbb{R}^{n_u}$

$$u(\cdot) \in \hat{\mathcal{C}}\left([t_0, t_1], \mathbb{R}^{n_u}\right) = \hat{\mathcal{C}}[t_0, t_1]^{n_u}$$

- Finitely many discontinuities on compact time intervals and bounded magnitude $\forall t \in [t_0, t_1] : \quad u(t) \in \mathbb{U}$, $\mathbb{U}$ is a compact subset of $\mathbb{R}^{n_u}$. Brief notation:

$$u(\cdot) \in \hat{\mathbb{U}}[t_0, t_1]^{n_u} \doteq \hat{\mathcal{C}}\left([t_0, t_1], \mathbb{U}\right)$$

Remark

Here, we do not consider discrete decision variables, e.g. $u \in \{-1, 0, 1\}$, as they lead to numerical challenges (mixed-integer optimization).

Dynamical System

Time-varying Ordinary Differential Equation (ODE):

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t, x(t), u(t)), \quad x(t_0) = x_0 \tag{Σ}$$

- Brief notation: $\dot{x} = f(t, x, u)$
- $t \in [t_0, t_1] \subset \mathbb{R}$ time
- $x \in \mathbb{X} \subseteq \mathbb{R}^{n_x}$ state
- $u(\cdot) \in \hat{\mathbb{U}}[t_0, t_1]^{n_u} \subseteq \mathbb{R}^{n_u}$ input

Dynamical System

Time-varying Ordinary Differential Equation (ODE):

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t, x(t), u(t)), \quad x(t_0) = x_0 \tag{Σ}$$

- Brief notation: $\dot{x} = f(t, x, u)$
- $t \in [t_0, t_1] \subset \mathbb{R}$ time
- $x \in \mathbb{X} \subseteq \mathbb{R}^{n_x}$ state
- $u(\cdot) \in \hat{\mathbb{U}}[t_0, t_1]^{n_u} \subseteq \mathbb{R}^{n_u}$ input

- Solution to (Σ):
  - $x(\cdot; t_0, x_0, u(\cdot))$ for time-varying ODE
  - $x(\cdot; x_0, u(\cdot))$ for time-invariant ODE

Existence of Solutions

Definition (Local Lipschitz continuity)

A function $f : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ *is said to be locally Lipschitz continuous (locally Lipschitz) at* $x_0 \in \mathbb{R}^{n_x}$ *if there exist constants* $L \geq 0, \eta \geq 0$ *such that*

$$\forall x, y \in \mathbb{B}_\eta(x_0) : \quad \|f(x) - f(y)\| \leq L\|x - y\|. \tag{LC}$$

A function $f : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$, $(t, x) \mapsto f(t, x)$ *is said to be locally Lipschitz at* x_0 *if* (LC) *holds uniformly for all* $t \in [t_0, t_1]$.

Existence of Solutions

Theorem (Local existence and uniqueness of solutions)

Let $f : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ *be piecewise continuous in* t *and Lipschitz at* x_0 *for all* $t \in [t_0, t_1]$. *Then there exists* $\delta > 0$ *such that*

$$\dot{x} = f(t, x), \quad x(t_0) = x_0$$

has a unique solution $x(\cdot; t_0, x_0)$ *over* $[t_0, t_0 + \delta]$.

Example

- $\dot{x} =$

Existence of Solutions

Theorem (Global existence and uniqueness of solutions)

Let $f : \mathbb{R} \times \mathcal{D} \to \mathbb{R}^{n_x}$ be piecewise continuous in t and Lipschitz at x_0 for all $t \geq t_0$ and all $x \in \mathcal{D} \subset \mathbb{R}^{n_x}$. Let $\mathbb{X}$ be a compact subset of $\mathcal{D}$ and $x_0 \in \mathbb{X}$. Suppose that every solution of

$$\dot{x} = f(t, x), \quad x(t_0) = x_0$$

satisfies for all $t \geq t_0 : x(t, t_0, x_0) \in \mathbb{X}$. Then there exists a unique solution that is defined for all $t \geq t_0$.

Performance Criteria

Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Lagrange form

$$J(u(\cdot)) = \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

- $\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}, \quad \ell \in \mathcal{C}^0, \quad \frac{\partial \ell}{\partial x} \in \mathcal{C}^0$

Performance Criteria

Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Lagrange form

$$J(u(\cdot)) = \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

- $\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}, \quad \ell \in \mathcal{C}^0, \quad \frac{\partial \ell}{\partial x} \in \mathcal{C}^0$
- t_0, t_1 can be fixed for free:
  - free end time: $J : \mathbb{R} \times \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}, \quad J(t_1, u(\cdot))$
  - free initial time: $J : \mathbb{R} \times \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}, \quad J(t_0, u(\cdot))$

Performance Criteria

- Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Lagrange form

$$J(u(\cdot)) = \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

$\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}, \quad \ell \in \mathcal{C}^0, \quad \frac{\partial \ell}{\partial x} \in \mathcal{C}^0$

Performance Criteria

- Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Lagrange form

$$J(u(\cdot)) = \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

$\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}, \quad \ell \in \mathcal{C}^0, \quad \frac{\partial \ell}{\partial x} \in \mathcal{C}^0$

- Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Mayer form

$$J(u(\cdot)) = \phi(t_0, x(t_0), t_1, x(t_1))$$

$\phi : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}, \quad \frac{\partial \phi}{\partial x} \in \mathcal{C}^0.$

Performance Criteria

- Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Lagrange form

$$J(u(\cdot)) = \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

$\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}, \quad \ell \in \mathcal{C}^0, \quad \frac{\partial \ell}{\partial x} \in \mathcal{C}^0$

- Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Mayer form

$$J(u(\cdot)) = \phi(t_0, x(t_0), t_1, x(t_1))$$

$\phi : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}, \quad \frac{\partial \phi}{\partial x} \in \mathcal{C}^0.$

- Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Bolza form

$$J(u(\cdot)) = \phi(t_0, x(t_0), t_1, x(t_1)) + \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

Equivalence of Performance Criteria

Physical Constraints

- Point constraints:

Physical Constraints

- Isoperimetric constraints:

Physical Constraints

- Path constraints:

Admissible and Feasible Controls

Definition (Feasible controls and feasible pair)

A control $u(\cdot) \in \hat{\mathbb{U}}[t_0, t_1]^{n_u}$ *is said to be feasible provided that:*

- *the response* $x(\cdot; t_0, x_0, \bar{u}(\cdot))$ *is defined on* $[t_0, t_1]$*;*
- $u(\cdot)$ *and* $x(\cdot; t_0, x_0, u(\cdot))$ *satisfy all constraints (point, path, ...) on* $[t_0, t_1]$*.*

Furthermore, $(u(\cdot), x(\cdot))$ *is called feasible pair.*

Admissible and Feasible Controls

Definition (Feasible controls and feasible pair)

A control $u(\cdot) \in \hat{\mathbb{U}}[t_0, t_1]^{n_u}$ *is said to be feasible provided that:*

- *the response* $x(\cdot; t_0, x_0, \bar{u}(\cdot))$ *is defined on* $[t_0, t_1]$*;*
- $u(\cdot)$ *and* $x(\cdot; t_0, x_0, u(\cdot))$ *satisfy all constraints (point, path, ...) on* $[t_0, t_1]$.

Furthermore, $(u(\cdot), x(\cdot))$ *is called feasible pair.*

Definition (Set of feasible controls)

The set

$$\Omega[t_0, t_1] \doteq \left\{ u(\cdot) \in \hat{\mathbb{U}}[t_0, t_1]^{n_u} \mid u(\cdot) \textit{ is feasible} \right\}$$

is called set of feasible controls.

Optimality Criteria

Definition (Global minimizer)

If for all $u \in \Omega[t_0, t_1]$ *and* $u^\star \in \Omega[t_0, t_1]$: $\quad J(u^\star) \leq J(u)$, *then* $u^\star$ *is called (global) minimizer.*

Optimality Criteria

Definition (Global minimizer)

If for all $u \in \Omega[t_0, t_1]$ *and* $u^\star \in \Omega[t_0, t_1]$: $\quad J(u^\star) \leq J(u)$, *then* $u^\star$ *is called (global) minimizer.*

Definition (Local minimizer)

If there exists $\delta > 0$ *such that for all* $u \in \mathcal{B}_\delta(u^\star) \cap \Omega[t_0, t_1] : J(u^\star) \leq J(u)$, *then* $u^\star$ *is called local minimizer.*

Open-loop Optimal Control

Given an OCP and let $u^\star(t)$ be the optimal control at time t. If $u^\star(t)$ is determined as a function of time for a specified initial state $x(t_0) = x_0$, i.e.

$$u^\star(t) = w_o(t, x(t_0)),$$

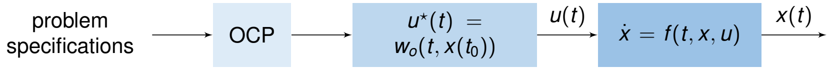
then $w_o : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$ is called *open-loop optimal control*.

Open-loop Optimal Control

Given an OCP and let $u^\star(t)$ be the optimal control at time t. If $u^\star(t)$ is determined as a function of time for a specified initial state $x(t_0) = x_0$, i.e.

$$u^\star(t) = w_o(t, x(t_0)),$$

then $w_o : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$ is called *open-loop optimal control*.



Closed-loop Optimal Control

Given an OCP and let $u^\star(t)$ be the optimal control at time t. If a functional relation $w_c : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$

$$u^\star(t) = w_c(t, x(t))$$

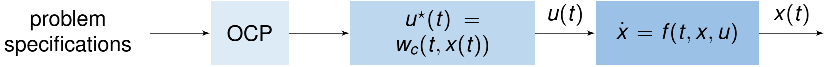
can be found, then w_c is called *closed-loop optimal control*.

Closed-loop Optimal Control

Given an OCP and let $u^\star(t)$ be the optimal control at time t. If a functional relation $w_c : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$

$$u^\star(t) = w_c(t, x(t))$$

can be found, then w_c is called *closed-loop optimal control*.

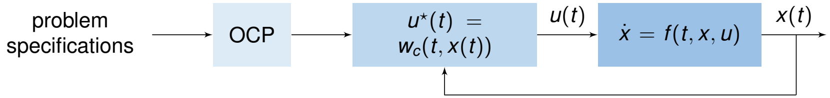


Closed-loop Optimal Control

Given an OCP and let $u^\star(t)$ be the optimal control at time t. If a functional relation $w_c : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$

$$u^\star(t) = w_c(t, x(t))$$

can be found, then w_c is called *closed-loop optimal control*.



Example – Car Control (Part I)

Consider the task to drive optimally from Lausanne to Zürich. The dynamics of the car are modelled as:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \qquad x = \left[p(t), \dot{p}(t)\right]^\top$$

Example – Car Control (Part I)

- Performance criteria (in words)?

- Constraints (in words)?

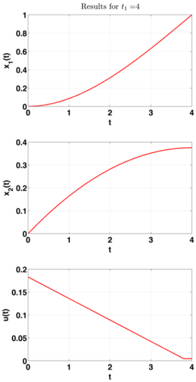
Example – Car Control (Part II)

One possible (simplified) OCP formulation reads:

$$
\begin{aligned}
&\min_{u(\cdot)} \quad \int_{t_0}^{t_1} u^2(t)\mathrm{d}t \\
&\text{subject to} \\
&\qquad \dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad x(t_0) = [p_0, 0]^\top \\
&\qquad x_1(t_1) - p_1 = 0 \\
&\qquad u(\cdot) \in \hat{\mathcal{C}}([t_0, t_1], [0, 1])
\end{aligned}
\tag{1}
$$

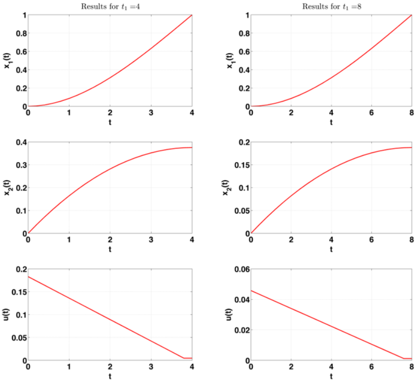
Example – Car Control (Part II)

We solve OCP (1) for $t_0 = 0, t_1 \in \{4, 8, 16, 32\}$, $p_0 = 0, p_1 = 1$



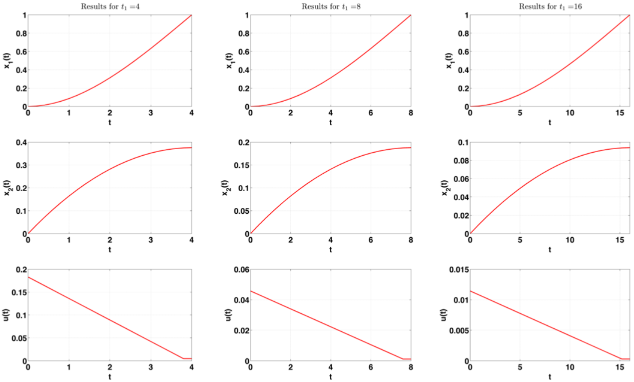
Example – Car Control (Part II)

We solve OCP (1) for $t_0 = 0, t_1 \in \{4, 8, 16, 32\}$, $p_0 = 0, p_1 = 1$



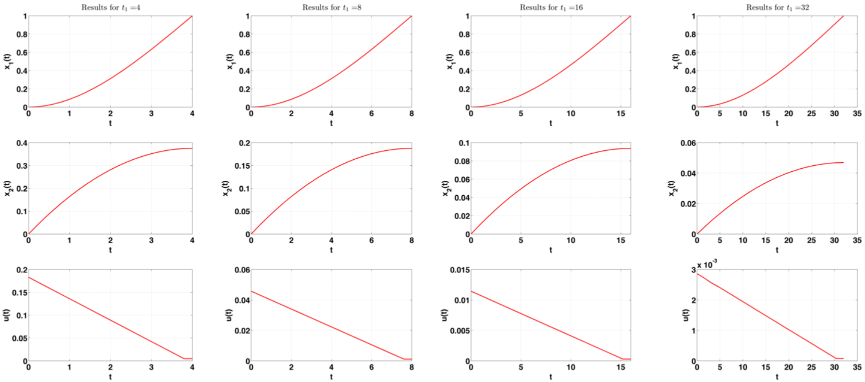
Example – Car Control (Part II)

We solve OCP (1) for $t_0 = 0, t_1 \in \{4, 8, 16, 32\}, p_0 = 0, p_1 = 1$



Example – Car Control (Part II)

We solve OCP (1) for $t_0 = 0, t_1 \in \{4, 8, 16, 32\}, p_0 = 0, p_1 = 1$



t_1	2	4	8	16	32
$J^\star$	0.38	0.05	0.006	0.0007	$9.2 \cdot 10^{-5}$

Example – Car Control (Part III)

Can we do even better solving the following OCP?

$$
\begin{aligned}
&\min_{u(\cdot),t_1} \quad \int_{t_0}^{t_1} u^2(t)\mathrm{d}t \\
&\text{subject to} \\
&\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad x(t_0) = [p_0, 0]^\top \\
&x_1(t_1) - p_1 = 0 \\
&u(\cdot) \in \hat{\mathcal{C}}([t_0, t_1], [0, 1]), \quad t_1 \in [t_0, \infty)
\end{aligned}
\tag{2}
$$

$\rightarrow$ This question will be answered later! $\leftarrow$

Overview

Performance Criteria in OCPs?

Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Lagrange form

$$J(u(\cdot)) = \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

- $\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}, \quad \ell \in \mathcal{C}^0, \quad \frac{\partial \ell}{\partial x} \in \mathcal{C}^0;$

$\to$ How to define a derivative-like operation for $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$? $\leftarrow$

Calculus of Variations – Basic Notions

Problems

$$\begin{aligned} &\min_{x(\cdot)\in\mathbb{X}} \int_{t_0}^{t_1} \ell(t, x(t), \dot{x}(t))\mathrm{d}t \\ &\text{subject to} \\ &x(\cdot) \in \mathcal{D} \end{aligned} \tag{P}$$

of minimizing (P) are called Lagrange problems.

- Set of admissible solutions:
  $$\mathcal{D} \doteq \{x \in \mathbb{X} \mid x \quad \text{is admissible}\}$$
- Boundary constraints:
  $$\mathcal{D} = \{x \in \mathbb{X} \mid x(t_0) = x_0, x(t_1) = x_1\}$$
- Joining a curve:
  $$\mathcal{D} = \{[x, t_1] \in \mathbb{X} \times [t_0, T] \mid x(t_0) = x_0, x(t_1) = g(t_1)\}$$
- ...

Calculus of Variations – Class of Considered Functions

$$\mathbb{X} = \mathcal{C}^1\left([t_0, t_1], \mathbb{R}^{n_x}\right) = \mathcal{C}^1[t_0, t_1]^{n_x},$$

i.e., continuously differentiable function on $x : [t_0, t_1] \to \mathbb{R}^{n_x}$.

$$\mathbb{X} = \hat{\mathcal{C}}^1\left([t_0, t_1], \mathbb{R}^{n_x}\right) = \hat{\mathcal{C}}^1[t_0, t_1]^{n_x},$$

i.e., piecewise continuously differentiable function on $x : [t_0, t_1] \to \mathbb{R}^{n_x}$.

Calculus of Variations – Local and Global Optima

Definition (Global minimizer)

The function $x^\star \in \mathcal{D}$ is said to be a (global) minimizer of $J(x(\cdot))$ on $\mathcal{D}$ if

$$\forall x \in \mathcal{D}: \quad J(x(\cdot)) \geq J(x^\star(\cdot)).$$

Definition (Local minimizer)

The function $x^\star \in \mathcal{D}$ is said to be a local minimizer of $J(x)$ on $\mathcal{D}$ if

$$\exists \epsilon > 0 \text{ such that } \forall x \in \mathbb{B}_\epsilon(x^\star) \cap \mathcal{D}: \quad J(x) \geq J(x^\star).$$

$$\mathbb{B}_\epsilon(\bar{x}) = \{x \in \mathbb{X} \mid \|x - \bar{x}\| < \epsilon\}$$

Calculus of Variations – Norms

- $\mathcal{L}^p$-norm
$$\|x\|_p = \left(\int_{t_0}^{t_1} |x(t)|^p \mathrm{d}t \right)^{\frac{1}{p}};$$
- Strong norm ($\mathcal{L}^\infty$-norm)
$$\|x\|_\infty = \max_{t \in [t_0, t_1]} |x(t)|;$$
- Weak norm
$$\|x\|_{\infty,1} = \max_{t \in [t_0, t_1]} |x(t)| + \max_{t \in [t_0, t_1]} |\dot{x}(t)|.$$

Weak and Strong Minima

Definition (Strong local minimizer)

The function $x^\star \in \mathcal{D}$ is said to be a strong local minimizer of $J(x)$ on $\mathcal{D}$ if

$$\exists \epsilon > 0 \text{ such that } \forall x \in \mathcal{B}_\epsilon^\infty(x^\star) \cap \mathcal{D} : \quad J(x) \geq J(x^\star).$$

Definition (Weak local minimizer)

The function $x^\star \in \mathcal{D}$ is said to be a weak local minimizer of $J(x)$ on $\mathcal{D}$ if

$$\exists \epsilon > 0 \text{ such that } \forall x \in \mathcal{B}_\epsilon^{\infty,1}(x^\star) \cap \mathcal{D} : \quad J(x) \geq J(x^\star).$$

Weak and Strong Minima

Example (Weak minimum $\not\Rightarrow$ strong minimum):

Calculus of Variations – Gateaux Derivative

Definition (First variation (Gâteaux derivative))

Let $J : \mathbb{X} \to \mathbb{R}$ be a functional defined on a linear space $\mathbb{X}$. Then the first variation of J at $x \in \mathbb{X}$ in the direction $\xi \in \mathbb{X}$, also called Gâteaux derivative with respect to ξ at x, is defined as

$$\delta J(x, \xi) \doteq \lim_{\eta \to 0} \frac{J(x + \eta\xi) - J(x)}{\eta} = \frac{\partial}{\partial \eta} J(x + \eta\xi)\Big|_{\eta=0},$$

provided it exists. If the limit exists for all $\xi \in \mathbb{X}$, then J is said to be Gâteaux differentiable.

Calculus of Variations – Gateaux Derivative

Remark

- $\delta\left(J_1(x,\xi) + J_2(x,\xi)\right) = \delta J_1(x,\xi) + \delta J_2(x,\xi)$;
- *Homogenity:* $\quad \delta J(x, \alpha\xi) = \alpha \delta J(x,\xi)$;
- *In general not additive:* $\quad \delta J(x, \xi_1 + \xi_2) \neq \delta J(x,\xi_1) + \delta J(x,\xi_2)$.

Examples – Gateaux Derivative

$$J(x) = \int_{t_0}^{t_1} x^2(t)\mathrm{d}t, \quad x \in \mathcal{C}^1[t_0, t_1]$$

and arbitrary $\xi \in \mathcal{C}^1[t_0, t_1]$:

$$\delta J(x, \xi) = \int_{t_0}^{t_1} 2x(t)\xi(t)\mathrm{d}t$$

J is Gâteaux differentiable at all $x \in \mathcal{C}^1[t_0, t_1]$.

Examples – Gateaux Derivative

$$J(x) = \int_0^1 |x(t)| \mathrm{d}t, \quad x \in \mathcal{C}^1[0,1]$$

and $x_0(t) = 0, \xi_0 = t$:

$$\lim_{\eta \to 0} \frac{J(x + \eta\xi) - J(x)}{\eta} = \frac{1}{2} \frac{|\eta|}{\eta} = \pm \frac{1}{2}$$

J is not Gâteaux differentiable at $x = 0$.

Lemma

Let J be a functional on a normed linear space $(\mathbb{X}, \|\cdot\|)$. Suppose that J has a strictly negative variation $\delta J(\bar{x}, \xi) < 0$ at $\bar{x} \in \mathbb{X}$ in some direction $\xi \in \mathbb{X}$. Then, $\bar{x}$ cannot be a local minimum point for J (in the sense of the norm $\|\cdot\|$).

Proof sketch:

Admissible Directions

Definition

Let J be a functional defined on a subset $\mathcal{D}$ of a linear space $\mathbb{X}$, and let $\bar{x} \in \mathcal{D}$. Then, a direction $\xi \in \mathbb{X}, \xi \neq 0$ is said to be $\mathcal{D}$-admissible at $\bar{x}$ for J, if

(i) $\delta J(\bar{x}, \xi)$ *exists; and*

(ii) $\bar{x} + \eta\xi \in \mathcal{D}$ *for all sufficiently small η, i.e.,*

$$\exists \delta > 0 \text{ such that } \forall \eta\xi \in \mathcal{B}_\delta(\bar{x}) : \quad \bar{x} + \eta\xi \in \mathcal{D}.$$

Geometric Conditions of Optimality

Theorem

Let J be a functional defined on a subset $\mathcal{D}$ of a normed linear space $(\mathbb{X}, \|\cdot\|)$. Suppose that $x^\star \in \mathcal{D}$ is a local minimum point for J on $\mathcal{D}$. Then

$$\delta J(x^\star, \xi) = 0$$

for each $\mathcal{D}$-admissible direction ξ at $x^\star$.

Proof sketch:

First-order Necessary Conditions of Optimality

Theorem (Euler's necessary conditions)

Consider the problem to minimize the functional

$$J(x) = \int_{t_0}^{t_1} \ell(t, x(t), \dot{x}(t))\mathrm{d}t$$

on $\mathcal{D} = \{x \in \mathcal{C}^1[t_0, t_1]^{n_x} \mid x(t_0) = x_0, x(t_1) = x_1\}$ *with* $\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}$ *a continuously differentiable function. Suppose that* $x^\star$ *is a (local) minimum for* J *on* $\mathcal{D}$*. Then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\ell_{\dot{x}_i}(t, x^\star(t), \dot{x}^\star(t)) = \ell_{x_i}(t, x^\star(t), \dot{x}^\star(t)) \tag{3}$$

for each $t \in [t_0, t_1]$ *and* $i = 1, \ldots, n_x$*.*

The proof follows as a special case of the NCOs for OCPs (next section).

Main messages

- Geometric optimality consideration carry over to function spaces naturally!
- Mathematical details can be subtle.
- Further reading $\rightarrow$B. Chachuat. *Nonlinear and Dynamic Optimization - From Theory to Practice*. EPFL, 2009. URL: `https://infoscience.epfl.ch/record/111939/files/Chachuat_07(IC32).pdf`, Chapter 3.

Overview

Example with no Optimal Solution (OCP (2))

$$\min_{u(\cdot),t_1} \quad \int_{t_0}^{t_1} u^2(t)\mathrm{d}t$$

subject to

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad x(t_0) = [p_0, 0]^\top$$

$$x_1(t_1) - p_1 = 0$$

$$u(\cdot) \in \hat{\mathcal{C}}([t_0, t_1], [0, 1]), \quad t_1 \in [t_0, \infty)$$

Example with no Optimal Solution

Example with no Optimal Solution

Example with no Optimal Solution

- Extend class of controls (piecewise continuous & Lipschitz bounds, piecewise constant).
- Focus on necessary conditions of optimality instead of sufficient conditions.

Existence of Optimal Controls?

$$\min_{u(\cdot)} \quad \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(x(t_1))$$

subject to:

$$\dot{x} = f(t, x, u), \quad x(t_0) = x_0$$
$$g(t, x(t), u(t)) \leq 0, \quad \forall t \in [t_0, t_1]$$
$$u(\cdot) \in \hat{\mathcal{C}}([t_0, t_1], \mathbb{U})$$

Reasons for non-existence of optimal controls:

Existence of Optimal Controls?

$$\min_{u(\cdot)} \quad \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(x(t_1))$$

subject to:

$$\dot{x} = f(t, x, u), \quad x(t_0) = x_0$$
$$g(t, x(t), u(t)) \leq 0, \quad \forall t \in [t_0, t_1]$$
$$u(\cdot) \in \hat{\mathcal{C}}([t_0, t_1], \mathbb{U})$$

Reasons for non-existence of optimal controls:

- No feasible solutions exist: $\Omega[t_0, t_1] = \varnothing$
- Finite escape times $\rightarrow$ infinite cost
- Set of feasible controls $\Omega[t_0, t_1]$ is not compact
  - e.g., free end time not bounded $t_1 \in [t_0, \infty)$

Existence of Optimal Controls?

$$\min_{u(\cdot)} \quad \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(x(t_1))$$

subject to:

$$\dot{x} = f(t, x, u), \quad x(t_0) = x_0$$
$$g(t, x(t), u(t)) \leq 0, \quad \forall t \in [t_0, t_1]$$
$$u(\cdot) \in \hat{\mathcal{C}}([t_0, t_1], \mathbb{U})$$

Reasons for non-existence of optimal controls:

- No feasible solutions exist: $\Omega[t_0, t_1] = \varnothing$
- Finite escape times $\rightarrow$ infinite cost
- Set of feasible controls $\Omega[t_0, t_1]$ is not compact
  - e.g., free end time not bounded $t_1 \in [t_0, \infty)$

$\rightarrow$ It is in general difficult to verify the existence of optimal controls! $\leftarrow$

Performance Criteria in OCPs?

Cost functional $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$ in Lagrange form

$$J(u(\cdot)) = \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

$\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}, \quad \ell \in \mathcal{C}^0, \quad \frac{\partial \ell}{\partial x} \in \mathcal{C}^0$

$\to$ How to define a derivative-like operation for $J : \hat{\mathcal{C}}[t_0, t_1]^{n_u} \to \mathbb{R}$? $\leftarrow$

First-order Necessary Conditions of Optimality

Problem Setup

$$\begin{aligned} &\min_{u(\cdot)} \quad \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t \\ &\text{subject to:} \\ &\qquad \dot{x} = f(t, x, u), \quad x(t_0) = x_0 \\ &\qquad u(\cdot) \in \mathcal{C}[t_0, t_1]^{n_u} \end{aligned} \tag{P}$$

$$\begin{aligned} f &: \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, \quad f \in \mathcal{C}^0 \text{ w.r.t. } (t, x, u), \quad f \in \mathcal{C}^1 \text{ w.r.t. } (x, u) \\ \ell &: \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R} \quad , \quad \ell \in \mathcal{C}^0 \text{ w.r.t. } (t, x, u), \quad \ell \in \mathcal{C}^1 \text{ w.r.t. } (x, u) \end{aligned}$$

Short hand notation:

$$\nabla_w z = \left(\frac{\partial z}{\partial w}\right)^\top = z_w, \quad z \in \{f,\, \ell,\, \phi, \dots\} \text{ and } w \in \{x,\, u,\, t\}$$

First-order Necessary Conditions of Optimality

Theorem (First-order necessary conditions)

Suppose that $u^\star(\cdot) \in \mathcal{C}[t_0, t_1]^{n_u}$ *is a local minimizer of Problem* (P) *and* $x^\star(\cdot) \in \mathcal{C}^1[t_0, t_1]^{n_x}$, $x^\star(t) = x(t; t_0, x_0, u^\star(\cdot))$ *is the corresponding solution. Then there exists a function* $\lambda^\star(\cdot) \in \mathcal{C}^1[t_0, t_1]^{n_x}$ *such that, for all* $t \in [t_0, t_1]$*, the triple* $(u^\star(\cdot), x^\star(\cdot), \lambda^\star(\cdot))$ *satisfies:*

$$\begin{aligned}
\dot{x}^\star &= \quad f(t, x^\star, u^\star), & x^\star(t_0) &= x_0 \\
\dot{\lambda}^\star &= -\ell_x(t, x^\star, u^\star) - f_x^\top(t, x^\star, u^\star)\lambda^\star, & \lambda^\star(t_1) &= 0 \\
0 &= \quad \ell_u(t, x^\star, u^\star) + f_u^\top(t, x^\star, u^\star)\lambda^\star. & &
\end{aligned} \tag{E-L}$$

Proof Sketch

Proof Sketch

Proof Sketch

Proof Sketch

Proof Sketch

Proof Sketch

Proof Sketch

First-order Necessary Conditions – Remarks

- (E-L) are also known as Euler-Lagrange equations
- Unknowns: $(u^\star(\cdot), x^\star(\cdot), \lambda^\star(\cdot)) \in \mathcal{C}[t_0, t_1]^{n_u} \times \mathcal{C}^1[t_0, t_1]^{n_x} \times \mathcal{C}^1[t_0, t_1]^{n_x}$
  - 2 n_x ODEs, n_u algebraic equations $\Rightarrow$ complete set of equations
  - Split boundary conditions $(x^\star(t_0), \lambda^\star(t_1)) \Rightarrow$ two-point boundary value problem
- Note (E-L) are **first-order necessary conditions** of (P). Hence any triple $(u(\cdot), x(\cdot), \lambda(\cdot))$ solving (E-L) is also referred to as an *extremal*.

First-order Necessary Conditions – Remarks

- (E-L) are also known as Euler-Lagrange equations
- Unknowns: $(u^\star(\cdot), x^\star(\cdot), \lambda^\star(\cdot)) \in \mathcal{C}[t_0, t_1]^{n_u} \times \mathcal{C}^1[t_0, t_1]^{n_x} \times \mathcal{C}^1[t_0, t_1]^{n_x}$
  - 2 n_x ODEs, n_u algebraic equations $\Rightarrow$ complete set of equations
  - Split boundary conditions $(x^\star(t_0), \lambda^\star(t_1)) \Rightarrow$ two-point boundary value problem
- Note (E-L) are **first-order necessary conditions** of (P). Hence any triple $(u(\cdot), x(\cdot), \lambda(\cdot))$ solving (E-L) is also referred to as an *extremal*.
- If terminal cost (i.e. a Mayer term) ϕ in (P), then

$$\lambda(t_1) = \phi_x(t_1, x(t_1))$$

.

First-order Necessary Conditions – Remarks

- Hamlitonian function: $H : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} \to \mathbb{R}$

  $$H(t, x, u, \lambda) = \ell(t, x, u) + \lambda^\top f(t, x, u) = \ell(t, x, u) + \langle \lambda,\ f(t, x, u) \rangle$$

  Notation for scalar product of $w, z, \in \mathbb{R}^n$: $\langle w,\ z \rangle = w^\top z$
- Alternative formulation of the Euler-Lagrange equations:

$$\begin{aligned}
\dot{x}^\star &= H_\lambda(t, x^\star, u^\star, \lambda^\star), \quad x^\star(t_0) = x_0 \\
\dot{\lambda}^\star &= -H_x(t, x^\star, u^\star, \lambda^\star), \quad \lambda^\star(t_1) = 0 \\
0 &= H_u(t, x^\star, u^\star, \lambda^\star)
\end{aligned}$$

First-order Necessary Conditions – Remarks

- Variation of Hamiltonian along optimal trajectory

$$\frac{\mathrm{d}}{\mathrm{d}t}H(t,x,u,\lambda) = H_t + \langle H_x, f(t,x,u)\rangle + \langle H_u, \dot{u}\rangle + \langle f(t,x,u), \dot{\lambda}\rangle$$

- If (P) is time invariant ($\ell_t = 0, f_t = 0$), then the Hamiltonian is stationary along optimal solutions: $\frac{\mathrm{d}}{\mathrm{d}t}H(t,x,u,\lambda) = 0$.
- Euler-Lagrange equations hold for minimization and maximization.

First-order Necessary Conditions – Remarks

- Variation of Hamiltonian along optimal trajectory

$$\frac{\mathrm{d}}{\mathrm{d}t} H(t,x,u,\lambda) = H_t + \langle H_x, f(t,x,u)\rangle + \langle H_u, \dot{u}\rangle + \langle f(t,x,u), \dot{\lambda}\rangle$$

- If (P) is time invariant ($\ell_t = 0, f_t = 0$), then the Hamiltonian is stationary along optimal solutions: $\frac{\mathrm{d}}{\mathrm{d}t} H(t,x,u,\lambda) = 0$.
- Euler-Lagrange equations hold for minimization and maximization.

The NCOs can be strengthened:

- Legendre-Clebsch condition: for minimization $u^\star(\cdot)$ should minimize Hamlitonian ($H_{uu} \geq 0$)
- B. Chachuat. *Nonlinear and Dynamic Optimization - From Theory to Practice*. EPFL, 2009. URL: `https://infoscience.epfl.ch/record/111939/files/Chachuat_07(IC32).pdf`.

Example: Euler-Lagrange Equations

$$\min_{u(\cdot)} \quad \int_0^1 \frac{1}{2}u^2(t) - x(t)\mathrm{d}t$$

$$\dot{x}(t) = 2(1 - u(t)), \quad x(0) = 1$$

$$u(\cdot) \in \mathcal{C}[0, 1]$$

Piecewise Continuous Extremals in OCP (P)

- Consider piecewise continuous controls $u(\cdot) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u}$ in (P)
- Allow for finitely many discontinuities at corner points $c_i \in [t_0, t_1],\ i \in \{1, \ldots, c_n\}$

Piecewise Continuous Extremals in OCP (P)

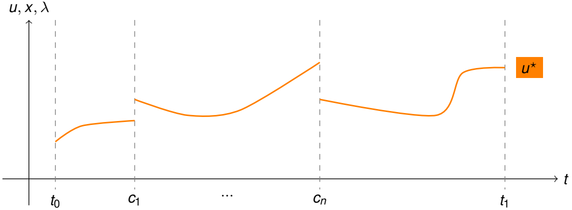
- Consider piecewise continuous controls $u(\cdot) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u}$ in (P)
- Allow for finitely many discontinuities at corner points $c_i \in [t_0, t_1],\ i \in \{1, \dots, c_n\}$

$\rightarrow$ State and adjoint trajectories remain continuous

$$\lim_{t \uparrow c_i} x^\star(t) = \lim_{t \downarrow c_i} x^\star(t) \qquad \lim_{t \uparrow c_i} \lambda^\star(t) = \lim_{t \downarrow c_i} \lambda^\star(t)$$

$\rightarrow$ Hamiltonian is also continuous

$$\lim_{t \uparrow c_i} H(t, x^\star(t), \lambda^\star(t), u^\star(t)) = \lim_{t \downarrow c_i} H(t, x^\star(t), \lambda^\star(t), u^\star(t))$$



Piecewise Continuous Extremals in OCP (P)

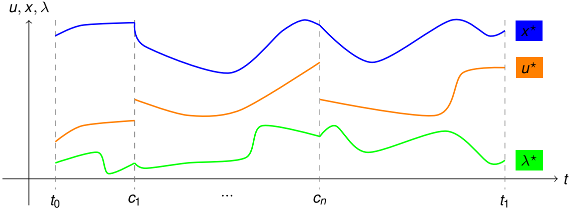
- Consider piecewise continuous controls $u(\cdot) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u}$ in (P)
- Allow for finitely many discontinuities at corner points $c_i \in [t_0, t_1],\ i \in \{1, \dots, c_n\}$

$\rightarrow$ State and adjoint trajectories remain continuous

$$\lim_{t \uparrow c_i} x^\star(t) = \lim_{t \downarrow c_i} x^\star(t) \qquad \lim_{t \uparrow c_i} \lambda^\star(t) = \lim_{t \downarrow c_i} \lambda^\star(t)$$

$\rightarrow$ Hamiltonian is also continuous

$$\lim_{t \uparrow c_i} H(t, x^\star(t), \lambda^\star(t), u^\star(t)) = \lim_{t \downarrow c_i} H(t, x^\star(t), \lambda^\star(t), u^\star(t))$$



- These additional conditions are also known as *Erdman's Corner Conditions*.

Example – Piecewise Continuous Extremals

$$\min_{u(\cdot)} \quad \int_0^1 (u^2(t) - u^4(t) - x(t))\mathrm{d}t$$

$$\dot{x} = -\,u(t), \quad x(0) = 1$$

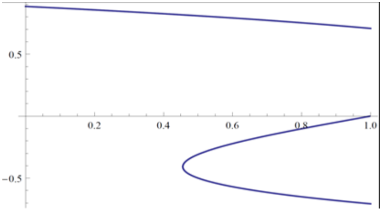
$$u(\cdot) \in \hat{\mathcal{C}}[0,1]$$

Example – Piecewise Continuous Extremals

$$\min_{u(\cdot)} \quad \int_0^1 (u^2(t) - u^4(t) - x(t))\mathrm{d}t$$

$$\dot{x} = -\,u(t), \quad x(0) = 1$$

$$u(\cdot) \in \hat{\mathcal{C}}[0,1]$$



Overview

Optimal Value Function of an OCP

Consider Problem (P):

$$\begin{aligned}\min_{u(\cdot)} J(u(\cdot)) &= \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t \\ \text{subject to:}& \\ \dot{x} = f(t, x, u), &\quad x(t_0) = x_0 \\ u(\cdot) \in \mathcal{C}[t_0, t_1]^{n_u}.\end{aligned} \tag{P}$$

Definition (Optimal value function)

Let $u^\star(\cdot)$ be the minimizer of Problem (P). *The function $V : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}$, $(x_0, t_0) \mapsto J(u^\star(\cdot))$*

$$V(x_0, t_0) \doteq J^\star(u^\star(\cdot))$$

*is called **optimal value function** of Problem* (P).[1]

[1] Note: Often the optimal value function is written as $V_{t_1 - t_0}(x_0, t_0)$ whereby the subscript $t_1 - t_0$ refers to the considered horizon in the OCP.

Interpretation of Adjoint Variables – Setting

- Consider as before Problem (P).
- Assume $u^\star(\cdot) \in \mathcal{C}[t_0, t_1]^{n_u}$ is the unique optimal control and $x^\star(\cdot) \in \mathcal{C}^1[t_0, t_1]^{n_x} \quad \lambda^\star(\cdot) \in \mathcal{C}^1[t_0, t_1]^{n_x}$ are the corresponding trajectories.

- NLPs:

  Lagrange multipliers $\rightarrow$ sensitivity of objective w.r.t. to constraints
- OCPs:

  Interpretation of adjoints $\lambda^\star(\cdot)$?

Interpretation of Adjoint Variables – Part I

Interpretation of Adjoint Variables – Part I

Interpretation of Adjoint Variables – Part I

$$\left(\frac{\partial V(y(t_0), t_0)}{\partial \xi}\right)^\top\Bigg|_{\xi=0} = V_x(x_0, t_0) = \lambda^\star(t_0)$$

$\rightarrow$ Adjoint variable $\lambda^\star$ at time t_0 is the gradient of $V(x_0)$ w.r.t. the initial condition x_0! $\leftarrow$

Principle of Optimality

Theorem (Principle of optimality [2])

Let $u^\star(\cdot) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u}$ be an optimal control for the problem

$$\begin{aligned} \min_{u(\cdot)} \quad & \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(x(t_1)) \\ \textit{subject to} \quad & \\ & \dot{x} = f(t, x, u), \quad x(t_0) = x_0 \\ & u(\cdot) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u} \end{aligned} \tag{P_{t_1}}$$

and $x^\star(\cdot; t_0, x_0, u^\star(\cdot))$ is the corresponding trajectory.

Principle of Optimality

Theorem (Principle of optimality [2])

Let $u^\star(\cdot) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u}$ *be an optimal control for the problem*

$$\begin{aligned} \min_{u(\cdot)} \quad & \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(x(t_1)) \\ \textit{subject to} \quad & \\ & \dot{x} = f(t, x, u), \quad x(t_0) = x_0 \\ & u(\cdot) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u} \end{aligned} \tag{P_{t_1}}$$

and $x^\star(\cdot; t_0, x_0, u^\star(\cdot))$ *is the corresponding trajectory.*
Then, for any $\tau \in [t_0, t_1]$, *the restriction of* $u^\star(\cdot)$ *to* $[\tau, t_1]$ *is optimal for the problem*

$$\begin{aligned} \min_{u(\cdot)} \quad & \int_{\tau}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(x(t_1)) \\ \textit{subject to} \quad & \\ & \dot{x} = f(t, x, u), \quad x(\tau) = x^\star(\tau; t_0, x_0, u^\star(\cdot)) \\ & u(\cdot) \in \hat{\mathcal{C}}[\tau, t_1]^{n_u} \end{aligned} \tag{P_{τ}}$$

Principle of Optimality

"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

R. Bellman. Dynamic Programming. Princeton University Press, 1957

Proof sketch:

Proof Sketch

Proof Sketch

Interpretation of Adjoint Variables – Part II

Consider Problem (P_τ) and use same arguments as before:

$$\begin{aligned}
\forall \tau \in [t_0, t_1]: \quad & \lambda^\star(\tau) && = \nabla_x V_{t_1 - t_0 - \tau}(x^\star(\tau), \tau) \\
\forall t \in [t_0, t_1]: \quad & \lambda^\star(t) && = \nabla_x V_{t_1 - t_0 - t}(x^\star(t), t)
\end{aligned}$$

$\rightarrow$ The adjoint variable $\lambda^\star(t)$ can be interpreted as the change of the objective function for small state pertubations and re-optimization.

$\rightarrow$ Reoptimization means that indeed for $t \in (t_0, t_1)$ the optimal value function refers to the truncated horizon, i.e.

$$\nabla_x V_{t_1 - t_0 - t}(x^\star(t), t) = \lambda^\star(t)$$

$\rightarrow$ Sensitivity/gradient of the optimal value function V w.r.t. to x.

Interpretation of Adjoint Variables – Part II

Consider Problem (P_τ) and use same arguments as before:

$$\begin{aligned} \forall \tau \in [t_0, t_1]: \quad & \lambda^\star(\tau) && = \nabla_x V_{t_1 - t_0 - \tau}(x^\star(\tau), \tau) \\ \forall t \in [t_0, t_1]: \quad & \lambda^\star(t) && = \nabla_x V_{t_1 - t_0 - t}(x^\star(t), t) \end{aligned}$$

$\rightarrow$ The adjoint variable $\lambda^\star(t)$ can be interpreted as the change of the objective function for small state pertubations and re-optimization.

$\rightarrow$ Reoptimization means that indeed for $t \in (t_0, t_1)$ the optimal value function refers to the truncated horizon, i.e.

$$\nabla_x V_{t_1 - t_0 - t}(x^\star(t), t) = \lambda^\star(t)$$

$\rightarrow$ Sensitivity/gradient of the optimal value function V w.r.t. to x.

$\rightarrow$ Observe that in problems without Mayer term ($\phi(x(t_1)) = 0$), there is no influence of disturbance at final time:

$$\lambda^\star(t_1) = \phi_x(x^\star(t_1)) = 0 \qquad \Rightarrow \qquad V_x(x^\star(t_1), t_1) = 0$$

Overview

OCPs with Terminal Equality Constraints

Problem Setup

$$\begin{aligned} \min_{u(\cdot)} \quad & \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(x(t_1)) \\ \text{subject to} \quad & \\ \dot{x} =& f(t, x, u), \quad x(t_0) = x_0 \\ u(\cdot) \in& \hat{\mathcal{C}}[t_0, t_1]^{n_u} \\ \psi_k&(t_1, x(t_1)) = 0, \quad k = 1, \dots, n_\psi \end{aligned} \tag{P_{eq}}$$

OCPs with Terminal Equality Constraints

Problem Setup

$$\min_{u(\cdot)} \quad \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(x(t_1))$$

subject to (P$_{eq}$)

$$\begin{aligned} \dot{x} &= f(t, x, u), \quad x(t_0) = x_0 \\ u(\cdot) &\in \hat{\mathcal{C}}[t_0, t_1]^{n_u} \\ \psi_k&(t_1, x(t_1)) = 0, \quad k = 1, \dots, n_\psi \end{aligned}$$

$$\Rightarrow \qquad P_k(u(\cdot)) \doteq \psi_k(t_1, x(t_1)) = 0, \quad k = 1, \dots, n_\psi$$

OCPs with Terminal Equality Constraints

Theorem

Suppose that $u^\star(\cdot) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u}$ *is the local minimizer of Problem* (P_{eq}) *and* $x^\star(\cdot) \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x}$ *is the corresponding solution* $x^\star(\cdot) = x^\star(\cdot; t_0, x_0, u^\star(\cdot))$. *Furthermore, assume that the terminal constraint* $\psi_k(t_1, x(t_1)) = 0$, $k = 1, \dots, n_\psi$ *is regular at* $u^\star(\cdot)$.
Then there exists a function $\lambda^\star(\cdot) \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x}$ *and a vector* $\nu^\star \in \mathbb{R}^{n_\psi}$ *such that* $(u^\star(\cdot),\, x^\star(\cdot),\, \lambda^\star(\cdot),\, \nu^\star)$ *satisfy for all* $t \in [t_0, t_1]$

$$
\begin{aligned}
\dot{x}(t) &= \quad H_\lambda(t, x, u, \lambda), \qquad x(t_0) = x_0 \\
\dot{\lambda}(t) &= -\, H_x(t, x, u, \lambda), \quad \lambda(t_1) = \Phi_x(t_1, x(t_1)) \\
0 &= \quad H_u(t, x, u, \lambda) \\
0 &= \quad \psi_k(t_1, x(t_1)), \quad k = 1, \dots, n_\psi \\
\Phi(t_1, x(t_1)) &= \quad \phi + \sum_{k=1}^{n_\psi} \nu_k \psi_k(t_1, x(t_1)).
\end{aligned}
$$

Proof Idea

- Consider first variation of
$$L(u(\cdot)) = J(u(\cdot)) + \sum_{k=1}^{n_\psi} \nu_k \underbrace{P_k(u(\cdot))}_{\psi_k(t_1, x(t_1))}$$

OCPs with Terminal Equality Constraints

$$\min_{u(\cdot)} \int_0^1 \frac{1}{2} u^2(t) \mathrm{d}t$$

subject to

$$\dot{x}(t) = u(t) - x(t), \quad x(0) = 1,\ x(1) = 0$$

$$u(\cdot) \in \mathcal{C}([0,1], \mathbb{R})$$

Example – OCPs with Terminal Equality Constraints

OCPs with Terminal Inequality Constraints

Problem setup (free end time and terminal constraints)

$$\min_{u(\cdot),t_1} \quad \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t + \phi(t_1, x(t_1))$$

subject to

$$\begin{aligned}
&\dot{x} = f(t, x, u), \quad x(t_0) = x_0 \\
&u(\cdot) \in \mathcal{C}[t_0, t_1]^{n_u} \\
&t_1 \in [t_0, T] \\
&\psi_k(t_1, x(t_1)) \leq 0, \quad k = 1, \ldots, n_\psi
\end{aligned}$$

$+$ differentiability assumptions:

$$\begin{aligned}
f &: \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, \quad f \in \mathcal{C}^0 \text{ w.r.t. } (t, x, u), \quad f \in \mathcal{C}^1 \text{ w.r.t. } (x, u) \\
\ell &: \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R} \quad , \quad \ell \in \mathcal{C}^0 \text{ w.r.t. } (t, x, u), \quad \ell \in \mathcal{C}^1 \text{ w.r.t. } (x, u) \\
\psi_k &: \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}, \quad \psi_k \in \mathcal{C}^1 \text{ w.r.t. } (t, x) \\
\phi &: \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}, \quad \phi \in \mathcal{C}^1 \text{ w.r.t. } (t, x).
\end{aligned}$$

Summary – Necessary Conditions of Optimality

Necessary conditions of optimality

$$
\begin{aligned}
\dot{x}(t) &= \quad H_\lambda(t, x, u, \lambda), \qquad x(t_0) = x_0 \\
\dot{\lambda}(t) &= -\, H_x(t, x, u, \lambda), \quad \lambda(t_1) = \mathbf{\Phi}_x(t_1, x(t_1)) \\
0 &= \quad H_u(t, x, u, \lambda) \\
0 &\leq \quad H_{uu}(t, x, u, \lambda) \\
0 &= \quad H(t, x, u, \lambda)|_{t_1} + \mathbf{\Phi}_t|_{t_1} \\
0 &\geq \quad \psi_k(t_1, x(t_1)), \quad k = 1, \ldots, n_\psi \\
0 &\leq \quad \nu \\
0 &= \quad \left\langle (\nu_1, \ldots, \nu_{n_\psi})^T,\ (\psi_1(t_1, x(t_1)), \ldots, \psi_{n_\psi}(t_1, x(t_1)))^T \right\rangle \\
\mathbf{\Phi}(t_1, x(t_1)) &= \phi + \left\langle (\nu_1, \ldots, \nu_{n_\psi})^T,\ (\psi_1(t_1, x(t_1)), \ldots, \psi_{n_\psi}(t_1, x(t_1)))^T \right\rangle
\end{aligned}
$$

- Extension to input constraints?$\rightarrow$Next lecture!
- Extension to state constraints?$\rightarrow$In general difficult, not covered here.

Overview

Lev Pontryagin

- Russian/Soviet mathematican
- 1908-1988
- Blind since the age of 14
- Seminal contributions to optimal control:
  - *Boltyanskii, V. G.; Gamkrelidze, R. V.; Pontryagin, L. S. (1956): Towards a Theory of Optimal Processes. Dokl. Akad. Nauk SSSR (in Russian)*
  - *Pontryagin, L. S.; Boltyanskii, V. G.; Gamkrelidze, R. V.; Mishchenko, E. F. (1962). The Mathematical Theory of Optimal Processes. English translation.*

Source: A. Terentiev. CC BY-SA 3.0

Pontryagin's Maximum Principle (Preliminaries)

Problem setup (free end time & terminal constraint)

$$\min_{u(\cdot),t_1} \quad \int_{t_0}^{t_1} \ell(x(t),u(t))\mathrm{d}t$$

subject to

$$\begin{aligned} &\dot{x} = f(x,u), \quad x(t_0) = x_0 \\ &u(\cdot) \in \hat{\mathcal{C}}\left([t_0,t_1],\mathbb{U}\right), \quad \mathbb{U} \subseteq \mathbb{R}^{n_u} \\ &t_1 \in [t_0,T], \quad T < \infty \\ &x(t_1) = x_1 \end{aligned} \tag{P_{PMP}}$$

$$f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, \quad f \in \mathcal{C}^0 \text{ w.r.t. } (x,u), \quad f \in \mathcal{C}^1 \text{ w.r.t. } (x)$$

$$\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R} \quad , \quad \ell \in \mathcal{C}^0 \text{ w.r.t. } (x,u), \quad \ell \in \mathcal{C}^1 \text{ w.r.t. } (x)$$

Pontryagin's Maximum Principle (Preliminaries)

Reformulation

$$
\begin{aligned}
c(t) &= \int_{t_0}^{t_1} \ell(x(t), u(t))\mathrm{d}t \\
\tilde{x}(t) &= \begin{bmatrix} c(t) \\ x(t) \end{bmatrix} \\
\dot{\tilde{x}}(t) &= \tilde{f}(\tilde{x}(t), u(t)) = \begin{bmatrix} \ell(x(t), u(t)) \\ f(x(t), u(t)) \end{bmatrix}, \quad \tilde{x}(t_0) = [0, x_0]^\top
\end{aligned}
$$

Pontryagin's Maximum Principle

Theorem

Suppose that $(u^\star(\cdot), t_1^\star) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u} \times [t_0, T]$ *is a local minimizer of Problem* (P_{PMP}) *and let* $\tilde{x}^\star(\cdot) \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$ *be the corresponding extended solution* $\tilde{x}^\star(\cdot) = \tilde{x}^\star(\cdot; t_0, \tilde{x}_0, u^\star(\cdot))$*.*

Pontryagin's Maximum Principle

Theorem

Suppose that $(u^\star(\cdot), t_1^\star) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u} \times [t_0, T]$ *is a local minimizer of Problem* (P_{PMP}) *and let* $\tilde{x}^\star(\cdot) \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$ *be the corresponding extended solution* $\tilde{x}^\star(\cdot) = \tilde{x}^\star(\cdot; t_0, \tilde{x}_0, u^\star(\cdot))$.

Then there exists a function $\tilde{\lambda}^\star(\cdot) = (\lambda_0^\star(\cdot), \lambda^\star(\cdot))^\top \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$, $\tilde{\lambda}^\star(t) \neq (0, \ldots, 0)^\top$ *for all* $t \in [t_0, t_1]$, *such that* $(u^\star(\cdot),\ \tilde{x}^\star(\cdot),\ \tilde{\lambda}^\star(\cdot))$ *satisfy for all* $t \in [t_0, t_1]$

$$
\begin{aligned}
\dot{\tilde{x}}^\star(t) &= H_{\tilde{\lambda}}\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right), \qquad \tilde{x}(t_0) = (0, x_0)^\top \\
\dot{\tilde{\lambda}}^\star(t) &= -H_{\tilde{x}}\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right)
\end{aligned}
$$

with $H(x, u, \tilde{\lambda}) = \left\langle \tilde{\lambda},\ \tilde{f}(x, u)\right\rangle$.

Pontryagin's Maximum Principle

Theorem

Suppose that $(u^\star(\cdot), t_1^\star) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u} \times [t_0, T]$ *is a local minimizer of Problem* (P_{PMP}) *and let* $\tilde{x}^\star(\cdot) \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$ *be the corresponding extended solution* $\tilde{x}^\star(\cdot) = \tilde{x}^\star(\cdot; t_0, \tilde{x}_0, u^\star(\cdot))$.

Then there exists a function $\tilde{\lambda}^\star(\cdot) = (\lambda_0^\star(\cdot), \lambda^\star(\cdot))^\top \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$, $\tilde{\lambda}^\star(t) \neq (0, \ldots, 0)^\top$ *for all* $t \in [t_0, t_1]$, *such that* $(u^\star(\cdot), \tilde{x}^\star(\cdot), \tilde{\lambda}^\star(\cdot))$ *satisfy for all* $t \in [t_0, t_1]$

$$\begin{aligned} \dot{\tilde{x}}^\star(t) &= \quad H_{\tilde{\lambda}}\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right), \qquad \tilde{x}(t_0) = (0, x_0)^\top \\ \dot{\tilde{\lambda}}^\star(t) &= -\, H_{\tilde{x}}\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) \end{aligned}$$

with $H(x, u, \tilde{\lambda}) = \left\langle \tilde{\lambda}, \tilde{f}(x, u)\right\rangle$.
Moreover:

i) *The function* $H\left(x^\star(t), v, \tilde{\lambda}^\star(t)\right)$ *attains its minimum on* $\mathbb{U}$ *at* $v = u^\star(t)$:

$$\forall t \in [t_0, t_1^\star]: \quad H\left(x^\star(t), v, \tilde{\lambda}^\star(t)\right) \geq H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right).$$

Pontryagin's Maximum Principle

Theorem

Suppose that $(u^\star(\cdot), t_1^\star) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u} \times [t_0, T]$ *is a local minimizer of Problem* (P_{PMP}) *and let* $\tilde{x}^\star(\cdot) \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$ *be the corresponding extended solution* $\tilde{x}^\star(\cdot) = \tilde{x}^\star(\cdot; t_0, \tilde{x}_0, u^\star(\cdot))$.

Then there exists a function $\tilde{\lambda}^\star(\cdot) = (\lambda_0^\star(\cdot), \lambda^\star(\cdot))^\top \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$, $\tilde{\lambda}^\star(t) \neq (0, \ldots, 0)^\top$ *for all* $t \in [t_0, t_1]$, *such that* $(u^\star(\cdot),\ \tilde{x}^\star(\cdot),\ \tilde{\lambda}^\star(\cdot))$ *satisfy for all* $t \in [t_0, t_1]$

$$
\begin{aligned}
\dot{\tilde{x}}^\star(t) &= \quad H_{\tilde{\lambda}}\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right), \qquad \tilde{x}(t_0) = (0, x_0)^\top \\
\dot{\tilde{\lambda}}^\star(t) &= -\, H_{\tilde{x}}\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right)
\end{aligned}
$$

with $H(x, u, \tilde{\lambda}) = \left\langle \tilde{\lambda}, \tilde{f}(x, u) \right\rangle$.
Moreover:

i) *The function* $H\left(x^\star(t), v, \tilde{\lambda}^\star(t)\right)$ *attains its minimum on* $\mathbb{U}$ *at* $v = u^\star(t)$*:*

$$\forall t \in [t_0, t_1^\star]: \quad H\left(x^\star(t), v, \tilde{\lambda}^\star(t)\right) \geq H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right).$$

ii) *It holds for all* $t \in [t_0, t_1^\star]$ *that* $\lambda_0^\star(t) = const. \geq 0$ *and* $H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) = const.$

Pontryagin's Maximum Principle

Theorem

Suppose that $(u^\star(\cdot), t_1^\star) \in \hat{\mathcal{C}}[t_0, t_1]^{n_u} \times [t_0, T]$ *is a local minimizer of Problem* (P_{PMP}) *and let* $\tilde{x}^\star(\cdot) \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$ *be the corresponding extended solution* $\tilde{x}^\star(\cdot) = \tilde{x}^\star(\cdot; t_0, \tilde{x}_0, u^\star(\cdot))$*.*

Then there exists a function $\tilde{\lambda}^\star(\cdot) = (\lambda_0^\star(\cdot), \lambda^\star(\cdot))^\top \in \hat{\mathcal{C}}^1[t_0, t_1]^{n_x+1}$*,* $\tilde{\lambda}^\star(t) \neq (0, \ldots, 0)^\top$ *for all* $t \in [t_0, t_1]$*, such that* $(u^\star(\cdot),\ \tilde{x}^\star(\cdot),\ \tilde{\lambda}^\star(\cdot))$ *satisfy for all* $t \in [t_0, t_1]$

$$\begin{aligned}
\dot{\tilde{x}}^\star(t) &= \quad H_{\tilde{\lambda}}\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right), \qquad \tilde{x}(t_0) = (0, x_0)^\top \\
\dot{\tilde{\lambda}}^\star(t) &= -\,H_{\tilde{x}}\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right)
\end{aligned}$$

with $H(x, u, \tilde{\lambda}) = \left\langle \tilde{\lambda}, \tilde{f}(x, u)\right\rangle$*.*

Moreover:

i) *The function* $H\left(x^\star(t), v, \tilde{\lambda}^\star(t)\right)$ *attains its minimum on* $\mathbb{U}$ *at* $v = u^\star(t)$*:*

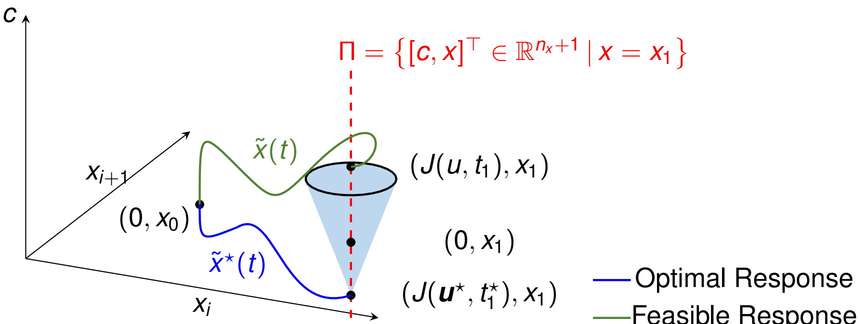
$$\forall t \in [t_0, t_1^\star]: \quad H\left(x^\star(t), v, \tilde{\lambda}^\star(t)\right) \geq H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right).$$

ii) *It holds for all* $t \in [t_0, t_1^\star]$ *that* $\lambda_0^\star(t) = const. \geq 0$ *and* $H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) = const.$

iii) *If the final time* t_1 *is unspecified, the following transversality condition holds*

$$H\left(x^\star(t_1^\star), u^\star(t_1^\star), \tilde{\lambda}^\star(t_1^\star)\right) = 0.$$

Rudimentary Proof Sketch



Detailed proofs of the PMP:

- D. Liberzon. *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, 2012. URL: `http://liberzon.csl.illinois.edu/teaching/cvoc.pdf`
- E.R. Pinch. *Optimal Control and the Calculus of Variations*. Oxford University Press, 1995

Example – Pontryagin's Maximum Principle

$$\min_{u(\cdot)} \quad \int_0^1 \frac{1}{2} u^2(t) \mathrm{d}t$$

subject to

$$\dot{x}(t) = u(t) - x(t), \quad x(0) = 1,\ x(1) = 0$$

$$u(\cdot) \in \hat{\mathcal{C}}([0,1], [-0.6,\ 0])$$

Example – Pontryagin's Maximum Principle

Arcs of Optimal Solutions

Due to the minimization of the Hamiltonian (PMP, part i), distinctive parts (arcs) of the optimal input trajectory $u^\star(\cdot)$ can be identified.

Arcs of Optimal Solutions

Due to the minimization of the Hamiltonian (PMP, part i), distinctive parts (arcs) of the optimal input trajectory $u^\star(\cdot)$ can be identified.

Let the time horizon of an OCP be $[t_0, t_1]$.

- If on $[\tau_1, \tau_2] \subseteq [t_0, t_1]$ one specific $\{$input, path$\}$ constraint is active, $u^\star(\cdot)$ on $[\tau_1, \tau_2]$ is said to be a *constrained arc*.
- Likewise, if at $t_1 - \tau_1$ the optimal input $u^\star(\cdot)$ changes its behavior in order to meet a state terminal constraint at t_1, $u^\star(\cdot)$ on $[t_1 - \tau_1, t_1]$ is said to be a *constrained arc*.

Arcs of Optimal Solutions

Due to the minimization of the Hamiltonian (PMP, part i), distinctive parts (arcs) of the optimal input trajectory $u^\star(\cdot)$ can be identified.

Let the time horizon of an OCP be $[t_0, t_1]$.

- If on $[\tau_1, \tau_2] \subseteq [t_0, t_1]$ one specific $\{$input, path$\}$ constraint is active, $u^\star(\cdot)$ on $[\tau_1, \tau_2]$ is said to be a *constrained arc*.
- Likewise, if at $t_1 - \tau_1$ the optimal input $u^\star(\cdot)$ changes its behavior in order to meet a state terminal constraint at t_1, $u^\star(\cdot)$ on $[t_1 - \tau_1, t_1]$ is said to be a *constrained arc*.
- If on $[\tau_1, \tau_2] \subseteq [t_0, t_1]$ no constraints are active and thus $u^\star(\cdot)$ on $[\tau_1, \tau_2]$ is determined by the unconstrained minimum of the Hamiltonian H, then $u^\star(\cdot)$ on $[\tau_1, \tau_2]$ is said to be a *sensitivity-seeking arc*.

Arcs of Optimal Solutions

Due to the minimization of the Hamiltonian (PMP, part i), distinctive parts (arcs) of the optimal input trajectory $u^\star(\cdot)$ can be identified.

Let the time horizon of an OCP be $[t_0, t_1]$.

- If on $[\tau_1, \tau_2] \subseteq [t_0, t_1]$ one specific $\{$input, path$\}$ constraint is active, $u^\star(\cdot)$ on $[\tau_1, \tau_2]$ is said to be a *constrained arc*.
- Likewise, if at $t_1 - \tau_1$ the optimal input $u^\star(\cdot)$ changes its behavior in order to meet a state terminal constraint at t_1, $u^\star(\cdot)$ on $[t_1 - \tau_1, t_1]$ is said to be a *constrained arc*.
- If on $[\tau_1, \tau_2] \subseteq [t_0, t_1]$ no constraints are active and thus $u^\star(\cdot)$ on $[\tau_1, \tau_2]$ is determined by the unconstrained minimum of the Hamiltonian H, then $u^\star(\cdot)$ on $[\tau_1, \tau_2]$ is said to be a *sensitivity-seeking arc*.
- *Singular arcs* will be introduced later.

Pontryagin's Maximum Principle – Remarks

- If $\lambda_0 > 0$, then $\lambda_0, \lambda_1(t), \dots, \lambda_{n_x}(t)$ are defined up to a common multiple ($\rightarrow$ **normal case**).
- Often one normalizes λ_0, such that $\forall t: \quad \lambda_0 = 1$.
- If $\lambda_0 = 0$, then $\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ is not present in the NCOs ($\rightarrow$ **abnormal case**).

Pontryagin's Maximum Principle – Remarks

- If $\lambda_0 > 0$, then $\lambda_0, \lambda_1(t), \ldots, \lambda_{n_x}(t)$ are defined up to a common multiple ($\rightarrow$ **normal case**).
- Often one normalizes λ_0, such that $\forall t: \quad \lambda_0 = 1$.
- If $\lambda_0 = 0$, then $\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ is not present in the NCOs ($\rightarrow$ **abnormal case**).
- Usually, OCPs without terminal constraints are normal, i.e., $\lambda_0 \neq 0$.
  Observation:

Pontryagin's Maximum Principle – Remarks

- If $\lambda_0 > 0$, then $\lambda_0, \lambda_1(t), \ldots, \lambda_{n_x}(t)$ are defined up to a common multiple ($\rightarrow$ **normal case**).
- Often one normalizes λ_0, such that $\forall t : \quad \lambda_0 = 1$.
- If $\lambda_0 = 0$, then $\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ is not present in the NCOs ($\rightarrow$ **abnormal case**).
- Usually, OCPs without terminal constraints are normal, i.e., $\lambda_0 \neq 0$. Observation:
- **Note**: minimization of the Hamiltonian is global w.r.t. u.
- Non-trivial solution , i.e., $\forall t \in [t_0, t_1]$ $(\lambda_0, \lambda_1(t), \ldots, \lambda_{n_x}(t)) \neq (0, \ldots, 0)$ required.

Pontryagin's Maximum Principle – Remarks

- If constraints are not active, one obtains the Euler-Lagrange equations (E-L).
- In case of free end-time $H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) = 0$
- In case of fixed end-time $H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) = const.$

Pontryagin's Maximum Principle – Remarks

- If constraints are not active, one obtains the Euler-Lagrange equations (E-L).
- In case of free end-time $H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) = 0$
- In case of fixed end-time $H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) = const.$
- The PMP allows consideration of minimum-time problems.

Pontryagin's Maximum Principle – Remarks

- If constraints are not active, one obtains the Euler-Lagrange equations (E-L).
- In case of free end-time $H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) = 0$
- In case of fixed end-time $H\left(x^\star(t), u^\star(t), \tilde{\lambda}^\star(t)\right) = const.$
- The PMP allows consideration of minimum-time problems.
- State constraints can be considered $\rightarrow$ see lecture notes B. Chachuat.
- PMP for time-varying systems $\rightarrow$ see lecture notes B. Chachuat.

Extensions of the PMP

- OCPs with fixed end time and without terminal constraints

$$J(u(\cdot)) = \int_{t_0}^{t_1} \ell(x(t), u(t))\mathrm{d}t + \phi(x(t_1))$$

lead to the usual terminal condition for the adjoint

$$\lambda^\star(t_1) = \phi_x(x^\star(t_1)).$$

Extensions of the PMP

- OCPs with generalized terminal constraint $x(t_1) \in \mathbb{X}_1$ lead to a transversality condition

$$\forall d \in \mathcal{T}(x^\star(t_1)) : \quad \langle \lambda^\star(t_1),\, d \rangle = 0$$

where $\mathcal{T}(x^\star(t_1))$ is the tangent plane to $\mathbb{X}_1$ at $x^\star(t_1)$.

- OCPs with generalized terminal constraint $x(t_1) \in \mathbb{X}_1$ lead to a transversality condition

$$\forall d \in \mathcal{T}(x^\star(t_1)) : \quad \langle \lambda^\star(t_1),\, d \rangle = 0$$

where $\mathcal{T}(x^\star(t_1))$ is the tangent plane to $\mathbb{X}_1$ at $x^\star(t_1)$.

- For terminal conditions $\Psi(x) = \begin{bmatrix} \psi_1(x) & \dots & \psi_{n_\psi}(x) \end{bmatrix} = 0$ with $\operatorname{rank}(\Psi_x(x^\star(t_1))) = n_\psi$ we have

$$\mathcal{T}(x^\star(t_1)) = \left\{ d \in \mathbb{R}^{n_x} : \quad d^\top \Psi_x(x^\star(t_1)) = 0 \right\}.$$

The terminal condition for the adjoint becomes

$$\lambda^\star(t_1) = \phi_x(x^\star(t_1)) + \nu^\top \Psi_x(x^\star(t_1)).$$

Time Optimal Control of Linear Systems

$$\min_{u(\cdot),t_1} \quad \int_0^{t_1} 1\mathrm{d}t$$

subject to

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \quad x(t_1) = 0 \qquad (\mathrm{P}_{minT})$$

$$u(\cdot) \in \hat{\mathcal{C}}\left([0, t_1], [\underline{u}, \overline{u}]\right), \quad \underline{u} < 0 < \overline{u}$$

Time Optimal Control of Linear Systems

Time Optimal Control of Linear Systems

Definition (Normal system)

Consider $\dot{x} = Ax + Bu, \quad x(0) = x_0$*. The pair* (A, B) *with* $B = [b_1, b_2, \ldots, b_m]$ *is called* ***normal*** *if for all* $i \in \{1, \ldots, m\}$ *the pairs* (A, b_i) *are controllable.*

Time Optimal Control of Linear Systems

Bang-Bang Solutions

Theorem (Bang-bang solutions for linear systems)

Consider Problem (P_{minT})*. Suppose that the input constraint is given by*

$$\mathbb{U} = [-1,\, 1] \times \cdots \times [-1,\, 1] \subset \mathbb{R}^{n_u}$$

and that (A, B) *is normal.*
Then

- *the optimal control* $u^\star(\cdot)$ *has finitely many switching points, i.e.* $u^\star(\cdot) \in \hat{\mathcal{C}}([0, t_1^\star], \mathbb{U})$*;*
- $u^\star(\cdot) \in \hat{\mathcal{C}}([0, t_1^\star], \mathbb{U})$ *is unique everywhere except the switching points;*
- *and* $u^\star(\cdot) \in \hat{\mathcal{C}}([0, t_1^\star], \mathbb{U})$ *takes only values on the vertices of* $\mathbb{U} \subset \mathbb{R}^{n_u}$*.*

Example – Time Optimal Control

$$\min_{u(\cdot),t_1} \quad t_1$$

subject to

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = u, \quad x(0) = x_0, \quad x(t_1) = 0$$

$$u(\cdot) \in \hat{\mathcal{C}}\left([0, t_1], [-1, 1]\right)$$

Example – Time Optimal Control

Overview

Singular Optimal Control Problems

$$\min_{u(\cdot)} \int_{t_0}^{t_1} \ell(t, x(t), u(t))\mathrm{d}t$$

subject to

$$\dot{x} = f(t, x, u), \quad x(t_0) = x_0 \qquad (\mathrm{P}_s)$$

$$u(\cdot) \in \hat{\mathcal{C}}\left([t_0, t_1], \mathbb{U}\right), \quad \mathbb{U} \subseteq \mathbb{R}^{n_u}$$

Definition (Singular optimal control problem)

Consider Problem (P_s) *from above and its corresponding Hamiltonian*

$$H(t, x, u, \lambda_0, \lambda) = \lambda_0 \ell(t, x, u) + \langle \lambda, f(t, x, u) \rangle .$$

If there exist $(t, x, u, \lambda_0, \lambda) \neq 0$ *for which* $\det H_{uu} = 0$*, then* (P_s) *is said to be a singular optimal control problem.*

Singular Optimal Control Problems

OCP with dynamics and objective affine in the scalar input $u \in \mathbb{R}$:

$$\min_{u(\cdot)} \quad \int_{t_0}^{t_1} \ell_0(t,x) + \ell_1(t,x)u(t)\mathrm{d}t$$

subject to

$$\dot{x} = f_0(t,x) + f_1(t,x)u, \quad x(t_0) = x_0$$

$$u(\cdot) \in \hat{\mathcal{C}}\left([t_0, t_1], [\underline{u}, \overline{u}]\right), \quad \underline{u} < 0 < \overline{u}$$

Singular Optimal Control Problems

OCP with dynamics and objective affine in the scalar input $u \in \mathbb{R}$:

$$\min_{u(\cdot)} \quad \int_{t_0}^{t_1} \ell_0(t,x) + \ell_1(t,x)u(t)\mathrm{d}t$$

subject to

$$\dot{x} = f_0(t,x) + f_1(t,x)u, \quad x(t_0) = x_0$$

$$u(\cdot) \in \hat{\mathcal{C}}\left([t_0,t_1],[\underline{u},\overline{u}]\right), \quad \underline{u} < 0 < \overline{u}$$

Hamiltonian is also affine in $u \in \mathbb{R}$

$$\begin{aligned} H(t,x,u,\lambda_0,\lambda) &= \left\langle \begin{pmatrix}\lambda_0\\ \lambda\end{pmatrix}, \begin{pmatrix}\ell_0 + \ell_1 u\\ f_0 + f_1 u\end{pmatrix}\right\rangle \\ &= \underbrace{\left\langle \begin{pmatrix}\lambda_0\\ \lambda\end{pmatrix}, \begin{pmatrix}\ell_0\\ f_0\end{pmatrix}\right\rangle}_{H_0(t,x,\lambda_0,\lambda)} + \underbrace{\left\langle \begin{pmatrix}\lambda_0\\ \lambda\end{pmatrix}, \begin{pmatrix}\ell_1\\ f_1\end{pmatrix}\right\rangle}_{H_1(t,x,\lambda_0,\lambda)} u \end{aligned}$$

$\rightarrow$ How to obtain the optimal input $u^\star(\cdot)$? $\leftarrow$

Singular Optimal Control Problems

Hamiltonian is affine in $u \in \mathbb{R}$

$$H(t, x, u, \lambda_0, \lambda) = H_0(t, x, \lambda_0, \lambda) + H_1(t, x, \lambda_0, \lambda)u.$$

Singular Optimal Control Problems

Hamiltonian is affine in $u \in \mathbb{R}$

$$H(t, x, u, \lambda_0, \lambda) = H_0(t, x, \lambda_0, \lambda) + H_1(t, x, \lambda_0, \lambda)u.$$

- **Case I**: $H_1(t, x, \lambda_0, \lambda)$ vanishes only at isolated points in time.
  $\Rightarrow$ Bang-bang control, e.g., time-optimal control of linear systems.

Singular Optimal Control Problems

Hamiltonian is affine in $u \in \mathbb{R}$

$$H(t, x, u, \lambda_0, \lambda) = H_0(t, x, \lambda_0, \lambda) + H_1(t, x, \lambda_0, \lambda)u.$$

- **Case I**: $H_1(t, x, \lambda_0, \lambda)$ vanishes only at isolated points in time.
  $\Rightarrow$ Bang-bang control, e.g., time-optimal control of linear systems.
- **Case II**: $H_1(t, x, \lambda_0, \lambda)$ vanishes on some interval $[\tau_0, \tau_1] \subseteq [t_0, t_1]$.
  $\Rightarrow$ Any value of u minimizes the Hamiltonian on $[\tau_0, \tau_1]$.
  $\Rightarrow$ **Singular arc** on $[\tau_0, \tau_1]$.

Resulting problem on $[\tau_0, \tau_1]$:

$$\dot{x} = H_\lambda(t, x, u, \lambda_0, \lambda) \tag{4a}$$

$$\dot{\lambda} = -H_x(t, x, u, \lambda_0, \lambda) \tag{4b}$$

$$0 = H_1(t, x, \lambda_0, \lambda) \tag{4c}$$

$\rightarrow$ Note: (4) is a system of **Differential Algebraic Equations (DAE)**. $\leftarrow$

Singular Optimal Control Problems

General OCP with Hamiltonian affine in $u \in \mathbb{R}$

$$H(t, x, u, \lambda_0, \lambda) = H_0(t, x, \lambda_0, \lambda) + H_1(t, x, \lambda_0, \lambda)u.$$

and along singular arc it holds that:

$$\begin{aligned} \dot{x} &= \quad H_\lambda(t, x, u, \lambda_0, \lambda) \\ \dot{\lambda} &= -H_x(t, x, u, \lambda_0, \lambda) \\ 0 &= \quad H_u(t, x, u, \lambda_0, \lambda) = H_1(t, x, \lambda_0, \lambda) \end{aligned}$$

$\rightarrow$ Treat $0 = H_u(t, x, u, \lambda_0, \lambda)$ as algebraic constraint.

$\rightarrow$ Obtain $u^\star(\cdot)$ by computation of time derivatives of $0 = H_u(t, x, u, \lambda_0, \lambda)$. $\leftarrow$

Singular Optimal Control Problems

Compute $\frac{\mathrm{d}^k}{\mathrm{d}t^k} H_u(t, x, u, \lambda_0, \lambda)$ until u appears:

$$\begin{aligned} H_u(t, x, u, \lambda_0, \lambda) \quad &= 0 \\ \tfrac{\mathrm{d}}{\mathrm{d}t} H_u(t, x, u, \lambda_0, \lambda) \quad &= 0 \\ \vdots \quad & \\ \tfrac{\partial}{\partial u} \left(\tfrac{\mathrm{d}^\sigma}{\mathrm{d}t^\sigma} H_u(t, x, u, \lambda_0, \lambda) \right) \quad &\neq 0 \end{aligned} \tag{5}$$

Solve (5) for the optimal control $u^\star(\cdot)$ on the singular arc.

$\rightarrow$ **Hint:** If the objective is separable in x and u and strictly quadratic in u, and moreover no state constraints are present (active), singular arcs cannot occur. $\leftarrow$

Example – Singular OCP

Example – Singular OCP

Overview

How should the PMP for OCPs with $T = \infty$ read?

$$
\begin{aligned}
V_\infty(x_0, t_0) = \min_{u(\cdot)} \quad & \int_{t_0}^{\infty} \ell(x(t), u(t)) \mathrm{d}t \\
\text{subject to:} \quad & \dot{x} = f(x, u), \quad x(t_0) = x_0 \\
& u(\cdot) \in \hat{\mathcal{C}}[t_0, t_0 + \infty]^{n_u}
\end{aligned}
$$

Solution?

Example – Halkin's Problem

$$
\begin{aligned}
&\min_{u(\cdot)} \quad \int_{t_0}^{\infty} -(1-x)u \,\mathrm{d}t \\
&\text{subject to:} \\
&\qquad \dot{x} = (1-x)u, \quad x(0) = x_0 \in [-\hat{x}, 1), \hat{x} < \infty \\
&\qquad u(\cdot) \in \hat{\mathcal{C}}([0,\infty],[0,1])
\end{aligned}
\tag{6}
$$

Solution?

Example – Halkin's Problem (cont'd)

- For any $u \in [0, 1]$ and $x_0 < 1$, $\dot{x} > 0$

[2] We see later in Part III that dissipativity notions for OCP allow to overcome this issue.

Example – Halkin's Problem (cont'd)

- For any $u \in [0, 1]$ and $x_0 < 1$, $\dot{x} > 0$
- To minimize the objective in (6) means to maximize $\lim_{t\to\infty} x^\star(t)$
- Hence we have $u^\star(t) \equiv 1$

[2] We see later in Part III that dissipativity notions for OCP allow to overcome this issue.

Example – Halkin's Problem (cont'd)

- For any $u \in [0, 1]$ and $x_0 < 1$, $\dot{x} > 0$
- To minimize the objective in (6) means to maximize $\lim_{t\to\infty} x^\star(t)$
- Hence we have $u^\star(t) \equiv 1$
- Solving the adjoint ODE gives

$$\lambda^\star(t) = (\lambda^\star(0) - \lambda_0^\star)e^t - \lambda_0^\star$$

- Normalizing $-\lambda_0^\star = \lambda^\star(0)$ implies

$$\lambda^\star(t) \equiv -\lambda_0^\star \neq 0$$

[2] We see later in Part III that dissipativity notions for OCP allow to overcome this issue.

Example – Halkin's Problem (cont'd)

- For any $u \in [0, 1]$ and $x_0 < 1$, $\dot{x} > 0$
- To minimize the objective in (6) means to maximize $\lim_{t\to\infty} x^\star(t)$
- Hence we have $u^\star(t) \equiv 1$
- Solving the adjoint ODE gives

$$\lambda^\star(t) = (\lambda^\star(0) - \lambda_0^\star)e^t - \lambda_0^\star$$

- Normalizing $-\lambda_0^\star = \lambda^\star(0)$ implies

$$\lambda^\star(t) \equiv -\lambda_0^\star \neq 0$$

Halkin's observation [7]

The usual transversality condition for the adjoint variable does not hold for infinite-horizon problems!
The PMP on infinite horizons remains partially open.[2]

[2] We see later in Part III that dissipativity notions for OCP allow to overcome this issue.

Overview

Towards closed-loop optimal control?

The PMP

- Constitutes necessary conditions of optimality
- In most cases provides *open-loop* optimal control
- Implies to solve a two-point boundary value problems for a differential algebraic equation

How to switch to closed-loop optimal control?

Problem Setting

$$V_T(x_0, t_0) = \min_{u(\cdot)} \quad \int_{t_0}^{t_0+T} \ell(t, x(t), u(t))\mathrm{d}t + \phi(t_0 + T, x(t_0 + T))$$

subject to: (P)

$$\dot{x} = f(t, x, u), \quad x(t_0) = x_0$$

$$u(\cdot) \in \hat{\mathcal{C}}[t_0, t_0 + T]^{n_u}$$

$$f : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, \quad f \in \mathcal{C}^0 \text{ w.r.t. } (t, x, u), \quad f \in \mathcal{C}^1 \text{ w.r.t. } (x, u)$$

$$\ell : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R} \quad , \quad \ell \in \mathcal{C}^0 \text{ w.r.t. } (t, x, u), \quad \ell \in \mathcal{C}^1 \text{ w.r.t. } (x, u)$$

Deriving the Hamilton-Jacobi-Bellman-Equation

Consider

$$V_T(x_0, t_0) = \int_{t_0}^{t_0+\delta t} \ell(t, x^\star(t), u^\star(t))\mathrm{d}t + \underbrace{\int_{t_0+\delta t}^{t_0+T} \ell(t, x^\star(t), u^\star(t))\mathrm{d}t}_{V_{T-\delta t}(x^\star(t_0+\delta t),\, t_0+\delta t)}$$

Deriving the Hamilton-Jacobi-Bellman-Equation

Consider

$$V_T(x_0, t_0) = \int_{t_0}^{t_0+\delta t} \ell(t, x^\star(t), u^\star(t))\mathrm{d}t + \underbrace{\int_{t_0+\delta t}^{t_0+T} \ell(t, x^\star(t), u^\star(t))\mathrm{d}t}_{V_{T-\delta t}(x^\star(t_0+\delta t),\, t_0+\delta t)}$$

The principle of optimality can be written as follows

$$V_T(x_0, t_0) = \min_{u(\cdot)} \left\{ \int_{t_0}^{t_0+\delta t} \ell(t, x(t), u(t))\mathrm{d}t + V_{T-\delta t}(x^\star(t_0+\delta t), t_0+\delta t) \right\}$$

Deriving the Hamilton-Jacobi-Bellman-Equation

Suppose that V_T is $\mathcal{C}^1$ in (t, x), then

$$V_{T-\delta t}(x^\star(t_0 + \delta t), t_0 + \delta t) = V_T(x_0, t_0) + \frac{\partial V_T}{\partial t}\delta t + \frac{\partial V_T}{\partial x}\dot{x}\delta t + h.o.t.$$

Deriving the Hamilton-Jacobi-Bellman-Equation

Suppose that V_T is $\mathcal{C}^1$ in (t, x), then

$$V_{T-\delta t}(x^\star(t_0 + \delta t), t_0 + \delta t) = V_T(x_0, t_0) + \frac{\partial V_T}{\partial t}\delta t + \frac{\partial V_T}{\partial x}\dot{x}\delta t + h.o.t.$$

Substitute the first-order Taylor series

$$V_T(x_0, t_0) = \min_{u(\cdot)} \left\{ V_T(x_0, t_0) + \frac{\partial V_T}{\partial t}\delta t + \frac{\partial V_T}{\partial x}\dot{x}\delta t + \int_{t_0}^{t_0+\delta t} \ell(t, x(t), u(t))\mathrm{d}t \right\}$$

Deriving the Hamilton-Jacobi-Bellman-Equation

Suppose that V_T is $\mathcal{C}^1$ in (t, x), then

$$V_{T-\delta t}(x^\star(t_0+\delta t), t_0+\delta t) = V_T(x_0, t_0) + \frac{\partial V_T}{\partial t}\delta t + \frac{\partial V_T}{\partial x}\dot{x}\delta t + h.o.t.$$

Substitute the first-order Taylor series

$$V_T(x_0, t_0) = \min_{u(\cdot)}\left\{ V_T(x_0, t_0) + \frac{\partial V_T}{\partial t}\delta t + \frac{\partial V_T}{\partial x}\dot{x}\delta t + \int_{t_0}^{t_0+\delta t} \ell(t, x(t), u(t))\mathrm{d}t \right\}$$

Substract $V_T(x_0, t_0)$, divide by δt, and let $\delta t \to 0$ to obtain

$$0 = \min_{u(\cdot)}\left\{ \frac{\partial V_T}{\partial t} + \frac{\partial V_T}{\partial x}\dot{x} + \ell(t, x(t), u(t)) \right\}$$

The Hamilton-Jacobi-Bellman-Equation for finite T

Consider problem (P) and let V_T be $\mathcal{C}^1$ in (t, x), we have

$$\begin{aligned} -\frac{\partial V_T}{\partial t} &= \min_u \{\langle \nabla_x V_T, f(t,x,u)\rangle + \ell(t,x,u)\} \\ V_T(x(t_0+T)) &= \phi(t_0+T, x(t_0+T)) \end{aligned} \tag{HJBE}$$

- The HJBE is a nonlinear partial differential equation

The Hamilton-Jacobi-Bellman-Equation for finite T

Consider problem (P) and let V_T be $\mathcal{C}^1$ in (t, x), we have

$$
\begin{aligned}
-\frac{\partial V_T}{\partial t} &= \min_u \{\langle \nabla_x V_T, f(t,x,u)\rangle + \ell(t,x,u)\} \\
V_T(x(t_0+T)) &= \phi(t_0+T, x(t_0+T))
\end{aligned}
\tag{HJBE}
$$

- The HJBE is a nonlinear partial differential equation
- The assumption that $V_T \in \mathcal{C}^1$ is quite restrictive
- Hence the analysis and numerical solution has evolved to so-called viscosity solutions [5]

The HJBE for $T = \infty$

$$V_\infty(x_0, t_0) = \min_{u(\cdot)} \quad \int_{t_0}^{\infty} \ell(x(t), u(t))\mathrm{d}t$$

$$\text{subject to:} \qquad (\mathrm{P}_\infty)$$

$$\dot{x} = f(x, u), \quad x(t_0) = x_0$$

$$u(\cdot) \in \hat{\mathcal{C}}[t_0, t_0 + \infty]^{n_u}$$

$$f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, \quad f \in \mathcal{C}^1 \text{ w.r.t. } (x, u)$$

$$\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R} \quad , \quad \ell \in \mathcal{C}^1 \text{ w.r.t. } (x, u)$$

The HJBE for $T = \infty$

Consider problem (P_∞) and let V_∞ be $\mathcal{C}^1$ in (t, x).

The HJBE for $T = \infty$

Consider problem (P_∞) and let V_∞ be $\mathcal{C}^1$ in (t, x).

Principle of optimality: end pieces of optimal trajectories are optimal.

The HJBE for $T = \infty$

Consider problem (P_∞) and let V_∞ be $\mathcal{C}^1$ in (t, x).

Principle of optimality: end pieces of optimal trajectories are optimal.

Hence end pieces of infinite-horizon optimal solutions are optimal on the infinite horizon, i.e. on $[t_0 + \delta t, \infty)$.

Therefore, for time-invariant infinite-horizon problems $\dfrac{\partial V_\infty}{\partial t} = 0$

The HJBE for $T = \infty$

Consider problem (P_∞) and let V_∞ be $\mathcal{C}^1$ in (t, x).

Principle of optimality: end pieces of optimal trajectories are optimal.

Hence end pieces of infinite-horizon optimal solutions are optimal on the infinite horizon, i.e. on $[t_0 + \delta t, \infty)$.

Therefore, for time-invariant infinite-horizon problems $\dfrac{\partial V_\infty}{\partial t} = 0$ and thus

$$0 = \min_u \left\{ \langle \nabla_x V_\infty, f(x,u) \rangle + \ell(x,u) \right\} \tag{7}$$

The HJBE for $T = \infty$

Consider problem (P_∞) and let V_∞ be $\mathcal{C}^1$ in (t, x).

Principle of optimality: end pieces of optimal trajectories are optimal.

Hence end pieces of infinite-horizon optimal solutions are optimal on the infinite horizon, i.e. on $[t_0 + \delta t, \infty)$.

Therefore, for time-invariant infinite-horizon problems $\dfrac{\partial V_\infty}{\partial t} = 0$ and thus

$$0 = \min_u \left\{ \langle \nabla_x V_\infty, f(x, u) \rangle + \ell(x, u) \right\} \tag{7}$$

Observation

In contrast to the PMP, which becomes more tricky for infinite horizons, the HJBE simplifies for $T = \infty$.

How to get the optimal control?

How to get the optimal control?

Suppose that V_T is known, then

$$u^\star(t,x) \in \arg\min_u \langle \nabla_x V_T, f(t,x,u)\rangle + \ell(t,x,u)$$

The Link between the HJBE and the PMP

Recall that

$$H(\lambda_0, \lambda, x, u) = \lambda_0 \ell + \langle \lambda, f \rangle$$

suppose that $\lambda_0 = 1$ (normal problem). Moreover, on Slide II.72 we have shown that

$$\nabla_x V_{T-t} = \lambda(t_0 + t)$$

The Link between the HJBE and the PMP

Recall that

$$H(\lambda_0, \lambda, x, u) = \lambda_0 \ell + \langle \lambda, f \rangle$$

suppose that $\lambda_0 = 1$ (normal problem). Moreover, on Slide II.72 we have shown that

$$\nabla_x V_{T-t} = \lambda(t_0 + t)$$

Hence

$$H(1, \nabla V_T, x, u) = \ell + \langle \nabla_x V_T, f \rangle$$

and thus (HJBE) can be written as

$$-\frac{\partial V_T}{\partial t} = \min_u H(1, \nabla_x V_T, x, u)$$
$$V_T(x(t_0 + T) = \phi(t_0 + T, x(t_0 + T))$$

The Link between the HJBE and the PMP

- The PMP and the HJBE are two side of the same medal
- Formal statements obtained via the HJBE are stronger (necessary and sufficient)
- HJBE provides optimal feedbacks
- Assumptions needed to work with the HJBE are stronger (differentiability, can be relaxed but quite technical)
- The HJBE for time-invariant problems becomes easier for infinite-horizons, cf. (HJBE) vs. (7)
- The PMP as such (cf. Slide 85) becomes tricky for $T = \infty$, i.e., the adjoint transversality for $\lambda^\star(t_1)$ creates problems [7]
- There exist interesting links between the HJBE and dissipativity properties of dynamical systems [6]
- Strict dissipativity also allows to characterize $\lim_{t\to\infty} \lambda^\star(t)$, see [6]

Overview

Summary Part II.1 – Optimal Control Theory

- Optimal control $\rightarrow$ optimization in function spaces

Summary Part II.1 – Optimal Control Theory

- Optimal control $\to$ optimization in function spaces
- Euler-Lagrange equations (E-L) $\leftrightarrow$ first-order NCOs for OCPs without input constraints

Summary Part II.1 – Optimal Control Theory

- Optimal control $\rightarrow$ optimization in function spaces
- Euler-Lagrange equations (E-L) $\leftrightarrow$ first-order NCOs for OCPs without input constraints
- Pontryagin's Maximum Principle (Slide 85) $\leftrightarrow$ first-order NCOs for OCPs with input constraints (extension to state constraints is possible)

Summary Part II.1 – Optimal Control Theory

- Optimal control $\rightarrow$ optimization in function spaces
- Euler-Lagrange equations (E-L) $\leftrightarrow$ first-order NCOs for OCPs without input constraints
- Pontryagin's Maximum Principle (Slide 85) $\leftrightarrow$ first-order NCOs for OCPs with input constraints (extension to state constraints is possible)
- NCOs for OCPs $\rightarrow$ boundary value problems (difficult to solve)
- Direct application of NCOs often difficult (small-scale examples)

Summary Part II.1 – Optimal Control Theory

- Optimal control $\rightarrow$ optimization in function spaces
- Euler-Lagrange equations (E-L) $\leftrightarrow$ first-order NCOs for OCPs without input constraints
- Pontryagin's Maximum Principle (Slide 85) $\leftrightarrow$ first-order NCOs for OCPs with input constraints (extension to state constraints is possible)
- NCOs for OCPs $\rightarrow$ boundary value problems (difficult to solve)
- Direct application of NCOs often difficult (small-scale examples)
- **Understanding of NCOs $\rightarrow$ insight into the structure (arcs) of optimal solutions**

Summary Part II.1 – Optimal Control Theory

- Optimal control $\rightarrow$ optimization in function spaces
- Euler-Lagrange equations (E-L) $\leftrightarrow$ first-order NCOs for OCPs without input constraints
- Pontryagin's Maximum Principle (Slide 85) $\leftrightarrow$ first-order NCOs for OCPs with input constraints (extension to state constraints is possible)
- NCOs for OCPs $\rightarrow$ boundary value problems (difficult to solve)
- Direct application of NCOs often difficult (small-scale examples)
- **Understanding of NCOs $\rightarrow$ insight into the structure (arcs) of optimal solutions**
- Bellman's optimality principle exploits the coupling through time evolution

Literature and References for Part II.1

- Chapter 4 of:
  B. Chachuat. *Nonlinear and Dynamic Optimization - From Theory to Practice*. EPFL, 2009. URL:
  `https://infoscience.epfl.ch/record/111939/files/Chachuat_07(IC32).pdf`
- D. Liberzon. *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, 2012. URL:
  `http://liberzon.csl.illinois.edu/teaching/cvoc.pdf`
- E.R. Pinch. *Optimal Control and the Calculus of Variations*. Oxford University Press, 1995
- A.E. Bryson **and** Y.-C. Ho. *Applied Optimal Control*. Ginn **and** Company, Waltham, Massachusetts, 1969
- M. Athans **and** P.L. Falb. *Optimal Control - An Introduction to Theory and its Applications*. McGraw-Hill Book Company, 1966
- H.J. Sussmann **and** J.C. Willems. “300 years of optimal control: from the brachystochrone to the maximum principle”. **in** *IEEE Control Systems*: 17.3 (1997), **pages** 32–44
- ...

$\rightarrow$ How to solve optimal control problems numerically? Next lectures. $\leftarrow$

References I

[1] M. Athans **and** P.L. Falb. *Optimal Control - An Introduction to Theory and its Applications*. McGraw-Hill Book Company, 1966.

[2] R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.

[3] A.E. Bryson **and** Y.-C. Ho. *Applied Optimal Control*. Ginn **and** Company, Waltham, Massachusetts, 1969.

[4] B. Chachuat. *Nonlinear and Dynamic Optimization - From Theory to Practice*. EPFL, 2009. URL: `https://infoscience.epfl.ch/record/111939/files/Chachuat_07(IC32).pdf`.

[5] M.G. Crandall **and** P.-L. Lions. “Two approximations of solutions of Hamilton-Jacobi equations”. **in***Mathematics of Computation*: 43.167 (1984), **pages** 1–19.

[6] T. Faulwasser **and** C.M. Kellett. “On Continuous-Time Infinite Horizon Optimal Control – Dissipativity, Stability and Transversality”. **in***Automatica*: 134 (2021), **page** 109907. DOI: `10.1016/j.automatica.2021.109907`.

References II

[7] H. Halkin. "Necessary conditions for optimal control problems with infinite horizons". **in** *Econometrica: Journal of the Econometric Society*: 42.2 (1974), **pages** 267–272.

[8] D. Liberzon. *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, 2012. URL: `http://liberzon.csl.illinois.edu/teaching/cvoc.pdf`.

[9] E.R. Pinch. *Optimal Control and the Calculus of Variations*. Oxford University Press, 1995.

[10] H.J. Sussmann **and** J.C. Willems. "300 years of optimal control: from the brachystochrone to the maximum principle". **in** *IEEE Control Systems*: 17.3 (1997), **pages** 32–44.