Optimal Control

Bangyan Liao
liaobangyan@westlake.edu.cn

Oct 2024

WESTLAKE UNIVERSITY

Outline

Outline

Definition (Calculus of Variations)

Let $\mathcal{X}$ denotes some infinite dimensional space, a calculus of variations problem can be defined as:

$$\inf_{\mathrm{x}\in\mathcal{X}} \mathcal{J}[\mathrm{x}] = \int_a^b L(u, x(u), x'(u))du$$
$$\mathrm{x} = \{x(u) : u \in [a, b]\} \tag{1}$$

where $\mathrm{J}[\mathrm{x}] : \mathcal{X} \longrightarrow \mathbb{R}$ is the functional integrating from time $u = a$ to time $u = b$, $L(u, x(u), x'(u))$ defines the Lagrangian cost (e.g. $L = \|x'(u)\|_2^2$) and x defines the general curve indexed by time u.

Theorem (Euler-Lagrange Equations)

Let x *be an extremum of Eq. 1. Then,* x *satisfies the Euler-Lagrange Equations:*

$$\partial_x L(u, x(u), x'(u)) = \frac{\mathrm{d}}{\mathrm{d}u}\partial_{x'} L(u, x(u), x'(u)), u \in [a, b]. \tag{2}$$

Proof.

Let us firstly Taylor expands the functional $\mathcal{J}[x]$ as

$$\delta \mathcal{J} = \int_a^b \left(\frac{\partial L}{\partial u} \delta u + \frac{\partial L}{\partial \dot{u}} \delta \dot{u} \right) dt$$

The term involving $\delta \dot{u}$ can be integrated by parts. Recall that $\delta \dot{u} = \frac{d}{dt}(\delta u)$, so:

$$\int_a^b \frac{\partial L}{\partial \dot{u}} \delta \dot{u} \, dt = \left[\frac{\partial L}{\partial \dot{u}} \delta u \right]_a^b - \int_a^b \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{u}} \right) \delta u \, dt$$

Assume that the variations $\delta u(t)$ vanish at the endpoints, i.e., $\delta u(a) = \delta u(b) = 0$.

$$\delta \mathcal{J} = \int_a^b \left(\frac{\partial L}{\partial u} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{u}} \right) \right) \delta u \, dt$$

Proof.

If we want the variation $\delta\mathcal{J}$ reduces to 0, we have to let the integration part to be 0 as:

$$\frac{\partial L}{\partial u} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{u}}\right) = 0$$

This is the Euler-Lagrange equation:

$$\partial_x L(u, x(u), x'(u)) = \frac{\mathrm{d}}{\mathrm{d}u}\partial_{x'} L(u, x(u), x'(u)), u \in [a, b]. \tag{3}$$

□

Definition (Differential Dynamics defined by ODE)

Let t denotes the system time, $\mathbf{x}(t) \in \mathbb{R}^d$ denotes the state, $\boldsymbol{\theta}(t) \in \Theta \subset \mathbb{R}^m$ denotes the control signal, we can define a trajectory defined by the following ODE:

$$\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), \boldsymbol{\theta}(t)), t \in [t_0, t_1], \mathbf{x}(t_0) = \mathbf{x}_0, \tag{4}$$

where $\mathbf{x}_0$ denotes the given starting state.

Definition (The Bolza Problem of Optimal Control)

$$
\begin{aligned}
&\inf_{\boldsymbol{\theta}} \mathcal{J}[\boldsymbol{\theta}] = \int_{t_0}^{t_1} L(t, \mathbf{x}(t), \boldsymbol{\theta}(t))dt + \Phi(t_1, \mathbf{x}(t_1)) \\
&\text{s.t. } \dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), \boldsymbol{\theta}(t)), t \in [t_0, t_1], \mathbf{x}(t_0) = \mathbf{x}_0,
\end{aligned}
\tag{5}
$$

where $L : \mathbb{R} \times \mathbb{R}^d \times \Theta \longrightarrow \mathbb{R}$ and $\Phi : \mathbb{R} \times \mathbb{R}^d \longrightarrow \mathbb{R}$ are called the running cost and the terminal cost, respectively.

Remark.

For historical reasons, the case where $\Phi = 0$ (no terminal cost) is called a Lagrange problem, where as the case with $L = 0$ (no running cost) is called a Mayer problem.

Outline

Definition (Hamiltonian)

Let us define the Hamiltonian functional $H : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \times \Theta \longrightarrow \mathbb{R}$ as:

$$H(t, \mathbf{x}, \mathbf{p}, \boldsymbol{\theta}) = \mathbf{p}^\top f(t, \mathbf{x}, \boldsymbol{\theta}) - L(t, \mathbf{x}, \boldsymbol{\theta}) \tag{6}$$

Theorem (Pontryagin's Maximum Principle)

Let $\boldsymbol{\theta}^*$ be a bounded, measurable and admissible control, and $\mathbf{x}^*$ be its corresponding state. Then, there exists an a.c. process $\mathbf{p}^* = \{\mathbf{p}^*(t) : t \in [t_0, t_1]\}$ such that

$$\begin{aligned} \dot{\mathbf{x}}^*(t) &= \nabla_{\mathbf{p}} H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}^*(t)), \quad \mathbf{x}^*(t_0) = \mathbf{x}_0 \\ \dot{\mathbf{p}}^*(t) &= -\nabla_{\mathbf{x}} H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}^*(t)), \quad \mathbf{p}^*(t_1) = -\nabla_{\mathbf{x}} \Phi(\mathbf{x}^*(t_1)) \\ H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}^*(t)) &\geq H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}(t)), \quad \forall \boldsymbol{\theta} \in \Theta \text{ and } t \in [t_0, t_1] \end{aligned} \tag{7}$$

Remark.

Pontryagin's Maximum Principle(PMP) can be treated as the necessary condition for optimality. The co-state $\mathbf{p}$ is to propagate back the optimality condition and is the adjoint of the variational equation. In fact, one can also connect the co-state formally to a Lagrange multiplier enforcing the constraint of the ODE. One can regard the PMP as a nontrivial generalization of the Euler-Lagrange equations to handle strong extrema, as well as a generalization of the KKT conditions to non-smooth settings.

Lemma (Dependence on Initial Condition)

Given the time-inhomogeneous ODE as

$$\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{8}$$

we can define the permutation $\mathbf{v}$ *as the solution to the initial permutation* $\mathbf{v}_0$*:*

$$\dot{\mathbf{v}}(s) = \nabla_{\mathbf{x}} f(s, \mathbf{x}(s))\mathbf{v}(s), \quad \mathbf{v}(0) = \mathbf{v}_0. \tag{9}$$

Proof.

Step1: Convert to Mayer Problem.

By going one dimension higher we can rewrite Eq. 5 as the Mayer problem

$$\begin{aligned} &\inf_{\boldsymbol{\theta}} \mathcal{J}[\boldsymbol{\theta}] = \Phi(t_1, \mathbf{x}(t_1)) + y(t_1), \quad t \in [t_0, t_1], \\ &\text{s.t. } \dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), \boldsymbol{\theta}(t)), \mathbf{x}(t_0) = \mathbf{x}_0, \\ &\qquad \dot{y}(t) = L(t, \mathbf{x}(t), \boldsymbol{\theta}(t)), y(t_0) = 0. \end{aligned} \tag{10}$$

For the simplicity, we will only consider this general Mayer problem.

$$\begin{aligned} &\inf_{\boldsymbol{\theta}} \mathcal{J}[\boldsymbol{\theta}] = \bar{\Phi}(t_1, \bar{\mathbf{x}}(t_1)), \quad t \in [t_0, t_1], \\ &\text{s.t. } \dot{\bar{\mathbf{x}}}(t) = \bar{f}(t, \bar{\mathbf{x}}(t), \boldsymbol{\theta}(t)), \bar{\mathbf{x}}(t_0) = \bar{\mathbf{x}}_0. \end{aligned} \tag{11}$$

Proof.

Step 2: Needle Perturbation.

Fix $\tau > 0$ and an admissible control $\mathbf{s} \in \Theta$. Define the needle perturbation to the optimal control

$$\boldsymbol{\theta}_\epsilon(t) = \begin{cases} \mathbf{s}, & \text{if } t \in [\tau - \epsilon, \tau], \\ \boldsymbol{\theta}^*(t), & \text{otherwise} \end{cases} \tag{10}$$

Let $\mathbf{x}_\epsilon(t)$ be the corresponding controlled trajectory, i.e., the solution of

$$\dot{\mathbf{x}}_\epsilon(t) = f(t, \mathbf{x}_\epsilon(t), \boldsymbol{\theta}_\epsilon(t)), \quad \mathbf{x}_\epsilon(t_0) = \mathbf{x}_0. \tag{11}$$

Our goal is to derive necessary conditions for which any such needle perturbation will be sub-optimal, thus resulting in a necessary condition for a strong minimum in the cost functional.

Proof.

Step 3: Variational Equation.

It is clear that $\mathbf{x}_\epsilon(t) = \mathbf{x}^*(t)$ for $t \leq \tau - \epsilon$. Let us define, for $t \geq \tau$

$$\mathbf{v}(t) := \lim_{\epsilon \to 0^+} \frac{\mathbf{x}_\epsilon(t) - \mathbf{x}^*(t)}{\epsilon}. \tag{10}$$

This measures the propagation of the effect of the needle perturbation as time increases. In particular, at $t = \tau$, $\mathbf{v}(\tau)$ is the tangent vector of the curve $\epsilon \mapsto \mathbf{x}_\epsilon(\tau)$, given by

$$\begin{aligned}\mathbf{v}(\tau) &= \lim_{\epsilon \to 0^+} \left(\frac{1}{\epsilon} \int_{\tau-\epsilon}^{\tau} f(t, \mathbf{x}_\epsilon(t), s) dt - \frac{1}{\epsilon} \int_{\tau-\epsilon}^{\tau} f(t, \mathbf{x}^*(t), \boldsymbol{\theta}^*(t)) dt \right) \\ &= f(\tau, \mathbf{x}^*(\tau), s) - f(\tau, \mathbf{x}^*(\tau), \boldsymbol{\theta}^*(\tau)).\end{aligned} \tag{11}$$

Proof.

For the remaining time $t \in [\tau, t_1]$, $\mathbf{x}_\epsilon$ follows the same ODE in Eq. 9.

$$\dot{\mathbf{v}}(t) = \nabla_{\mathbf{x}} f(t, \mathbf{x}^*(t), \boldsymbol{\theta}^*(t))\mathbf{v}(t), \quad t \in [\tau, t_1], \tag{10}$$

with initial condition given by $\mathbf{v}(\tau)$. In particular, the vector $\mathbf{v}(t_1)$ describes the variation in the end point $\mathbf{x}_\epsilon(t_1)$ due to the needle perturbation $\mathbf{v}(\tau)$.

Proof.

Step 4: Optimality Condition at End Point.
By our assumption, the control θ^* is optimal, hence we must have

$$\Phi(\mathbf{x}^*(t_1)) \leq \Phi(\mathbf{x}_\epsilon(t_1)). \tag{10}$$

Thus, we have

$$0 \leq \lim_{\epsilon \to 0^+} \frac{\Phi(\mathbf{x}_\epsilon(t_1)) - \Phi(\mathbf{x}^*(t_1))}{\epsilon} = \frac{d}{d\epsilon}\Phi(\mathbf{x}_\epsilon(t_1))\Big|_{\epsilon=0^+} = \nabla\Phi(\mathbf{x}^*(t_1)) \cdot \mathbf{v}(t_1). \tag{11}$$

In fact, the inequality (2.28) holds for any τ and s that characterizes the needle perturbation.

Proof.

Step 5: The Adjoint Equation and the Maximum Principle.
To this end, we define $p^*(t)$ as the solution of the backward Cauchy problem

$$\dot{\mathbf{p}}^*(t) = -\nabla_{\mathbf{x}} f(t, \mathbf{x}^*(t), \boldsymbol{\theta}^*(t))^\top \mathbf{p}^*(t), \quad \mathbf{p}^*(t_1) = -\nabla \Phi(\mathbf{x}^*(t_1)). \tag{10}$$

Then, observe that we indeed have

$$\frac{d}{dt}[\mathbf{p}^*(t)^\top \mathbf{v}(t)] = 0 \quad \forall t \in [\tau, t_1] \implies \mathbf{p}^*(\tau)^\top \mathbf{v}(\tau) = \mathbf{p}^*(t_1)^\top \mathbf{v}(t_1) \leq 0, \tag{11}$$

which implies that for any $\tau \in (t_0, t_1]$ we have

$$[\mathbf{p}^*(\tau)]^\top f(\tau, \mathbf{x}^*(\tau), \mathbf{s}) \geq [\mathbf{p}^*(\tau)]^\top f(\tau, \mathbf{x}^*(\tau), \boldsymbol{\theta}^*(\tau)) \quad \forall \mathbf{s} \in \Theta. \tag{12}$$

By continuity this also holds for $t = t_0$.

Proof.

By undoing the conversion in Step 1, we can go back to a general Bolza problem by sending $\bar{\mathbf{p}}^* \to (\mathbf{p}^*, p_y^*)$. In particular, observe that $p_y^*(t_1) = -1$ and $\dot{p}_y^*(t) = -\nabla_y L(t, \mathbf{x}(t), \boldsymbol{\theta}(t))^\top p_y^*(t) = 0$. Hence, $p_y^*(t) \equiv -1$. Hence, we get from the optimality condition that

$$\mathbf{p}^*(\tau)^\top f(\tau, \mathbf{x}^*(\tau), \boldsymbol{\theta}^*(\tau)) - L(\tau, \mathbf{x}^*(\tau), \boldsymbol{\theta}^*(\tau)) \geq \mathbf{p}^*(\tau)^\top f(\tau, \mathbf{x}^*(\tau), \mathbf{s}) - L(\tau, \mathbf{x}^*(\tau), \mathbf{s}), \tag{10}$$

where $\mathbf{p}^*$ satisfies the adjoint equation

$$\dot{\mathbf{p}}^*(t) = -\nabla_{\mathbf{x}} H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}^*(t)), \quad \mathbf{p}^*(t_1) = -\nabla \Phi(\mathbf{x}^*(t_1)). \tag{11}$$

□

Outline

Definition (Value Function)

The value function $V : [t_0, t_1] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ is the minimum cost attainable starting from the initial state $\mathbf{z}$ at time s.

$$\begin{aligned} V(s,\mathbf{z}) = \inf_{\boldsymbol{\theta}} \int_s^{t_1} L(t,\mathbf{x}(t),\boldsymbol{\theta}(t))dt + \Phi(t_1,\mathbf{x}(t_1)) \\ \text{s.t. } \dot{\mathbf{x}}(t) = f(t,\mathbf{x}(t),\boldsymbol{\theta}(t)), t \in [s,t_1], \mathbf{x}(s) = \mathbf{z}, \end{aligned} \tag{12}$$

Theorem (Dynamic Programming Principle)

For every $\tau, s \in [t_0, t_1], s \leq \tau$, *and* $\mathbf{z} \in \mathbb{R}^d$, *we have*

$$\begin{aligned} V(s,\mathbf{z}) = \inf_{\boldsymbol{\theta}} \left\{ \int_s^{\tau} L(t,\mathbf{x}(t),\boldsymbol{\theta}(t))dt + V(\tau,\mathbf{x}(\tau)) \right\} \\ s.t. \ \dot{\mathbf{x}}(t) = f(t,\mathbf{x}(t),\boldsymbol{\theta}(t)), t \in [s,\tau], \mathbf{x}(s) = \mathbf{z}, \end{aligned} \tag{13}$$

Some Remarks about the DPP

Dynamic Programming Principle (DPP)

Remark.

The meaning of the DPP is that the optimization problem defining $V(s, \mathbf{z})$ can be split into two parts:

1. First, solve the optimization problem on $[\tau, t_1]$ with the usual running cost L and terminal cost Θ, but for all initial state $\mathbf{z}' \in \mathbb{R}^d$. This gives us the value function $V(\tau, \cdot)$.
2. Second, solve the optimization problem on $[s, \tau]$ with running cost L and terminal cost $V(\tau, \cdot)$ given by the step before.

Theorem (Hamilton-Jacobi-Bellman Equations)

The value function V in Eq. 12 is the unique viscosity solution of the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} &\partial_t V(t, \mathbf{x}) + \inf_{\boldsymbol{\theta}} \left\{ L(t, \mathbf{x}, \boldsymbol{\theta}) + [\nabla_{\mathbf{x}} V(t, \mathbf{x})]^\top f(t, \mathbf{x}, \boldsymbol{\theta}) \right\} = 0 \\ &V(t_1, \mathbf{x}) = \Phi(\mathbf{x}), \quad (t, \mathbf{x}) \in [t_0, t_1] \times \mathbb{R}^d \end{aligned} \tag{14}$$

Remark.

HJB equation establishes the necessary and sufficient conditions for optimal control problem. Provided we can solve the HJB, the optimal control solution is of feed-back or closed-loop form, meaning that it tells how to steer the system by just observing the state trajectory. We can contrast with the PMP, where we obtain open-loop controls that are pre-computed and cannot be applied on-the-fly.

Proof.

To begin with, we can derive the infinitesimal version of the dynamic programming principle defined in Eq. 13. Let $\tau = s + \Delta s$, then

$$
\begin{aligned}
V(s, \mathbf{z}) &= \inf_{\boldsymbol{\theta}} \left\{ \int_s^{s+\Delta s} L(t, \mathbf{x}(t), \boldsymbol{\theta}(t)) dt + V(s + \Delta s, \mathbf{x}(s + \Delta s)) \right\} \\
&\approx \inf_{\boldsymbol{\theta}} \left\{ L(s, \mathbf{z}, \boldsymbol{\theta}(s)) \Delta s + V(s + \Delta s, \mathbf{x}(s + \Delta s)) \right\} \\
&\approx \inf_{\boldsymbol{\theta}} \{ L(s, \mathbf{z}, \boldsymbol{\theta}(s)) \Delta s + V(s, \mathbf{x}(s)) \\
&\quad + \partial_s V(s, \mathbf{z}) \Delta s + [\nabla_{\mathbf{z}} V(s, \mathbf{z})]^\top f(s, \mathbf{z}, \boldsymbol{\theta}(s)) \Delta s \} \\
\dot{\mathbf{x}}(t) &= f(t, \mathbf{x}(t), \boldsymbol{\theta}(t)), \quad t \in [s, \tau], \quad \mathbf{x}(s) = \mathbf{z}
\end{aligned}
\tag{15}
$$

Proof.

After cancelling the term $V(s, \mathbf{z})$ on both slides and taking the limit $\Delta s \longrightarrow 0$, the infimum over paths $\boldsymbol{\theta}$ on $t \in [s, s + \Delta s]$ becomes an infimum over a scalar $\theta = \theta(s)$, thus we obtain the Hamilton-Jacobi-Bellman equation for the value function.

$$0 = \partial_s V(s, \mathbf{z}) + \inf_{\boldsymbol{\theta}} \left\{ L(s, \mathbf{z}, \boldsymbol{\theta}(s)) + [\nabla_{\mathbf{z}} V(s, \mathbf{z})]^\top f(s, \mathbf{z}, \boldsymbol{\theta}(s)) \right\} \tag{15}$$

Then, combine with the boundary condition $V(t_1, \mathbf{x}) = \Phi(\mathbf{x})$, we can result the full HJB equations. □

The Necessary Condition

Dynamic Programming Principle (DPP)

Proof.

By the assumption of global optimality, we can perform Taylor expanding and comparing with the usual dynamic programming principle as:

$$
\begin{aligned}
-\partial_t V(t, \mathbf{x}^*) &= \inf_{\boldsymbol{\theta}} \left\{ L(t, \mathbf{x}^*, \boldsymbol{\theta}) + [\nabla_{\mathbf{x}} V(t, \mathbf{x}^*)]^\top f(t, \mathbf{x}^*, \boldsymbol{\theta}) \right\} \\
&= L(t, \mathbf{x}^*, \boldsymbol{\theta}^*) + [\nabla_{\mathbf{x}} V(t, \mathbf{x}^*)]^\top f(t, \mathbf{x}^*, \boldsymbol{\theta}^*)
\end{aligned} \tag{16}
$$

Then, recall the Hamiltonian formulation as

$$
H(t, \mathbf{x}, \mathbf{p}, \boldsymbol{\theta}) = \mathbf{p}^\top f(t, \mathbf{x}, \boldsymbol{\theta}) - L(t, \mathbf{x}, \boldsymbol{\theta}) \tag{17}
$$

Finally, we can rewrite it as a similar statement of th PMP

$$
H(t, \mathbf{x}^*, -\nabla_{\mathbf{x}} V(t, \mathbf{x}^*), \boldsymbol{\theta}^*) = \max_{\boldsymbol{\theta}} H(t, \mathbf{x}^*, -\nabla_{\mathbf{x}} V(t, \mathbf{x}^*), \boldsymbol{\theta}) \tag{18}
$$

Proof.

Let us now assume that a continuously differentiable function V satisfies the HJB equation and moreover that a control $\hat{\boldsymbol{\theta}} : [t_0, t_1] \to \Theta$ satisfies

$$H(t, \hat{\mathbf{x}}(t), -\nabla_{\mathbf{x}} V(t, \hat{\mathbf{x}}(t)), \hat{\boldsymbol{\theta}}(t)) = \max_{\boldsymbol{\theta} \in \Theta} H(t, \hat{\mathbf{x}}(t), -\nabla_{\mathbf{x}} V(t, \hat{\mathbf{x}}(t)), \boldsymbol{\theta}), \tag{19}$$

for all $t \in [t_0, t_1]$, where $\hat{\mathbf{x}}(t)$ is the state process corresponding to the control $\hat{\boldsymbol{\theta}}$, then $\hat{\boldsymbol{\theta}}$ is a globally optimal control that solves the dynamic programming principle with optimal cost $V(t_0, x_0)$.

To show this, observe that if we set $x = \hat{\mathbf{x}}(t)$ in the HJB equation for V, noting the condition, we have

$$\partial_t V(t, \hat{\mathbf{x}}(t)) + [\nabla_{\mathbf{x}} V(t, \hat{\mathbf{x}}(t))]^T f(t, \hat{\mathbf{x}}(t), \hat{\boldsymbol{\theta}}(t)) + L(t, \hat{\mathbf{x}}(t), \hat{\boldsymbol{\theta}}(t)) = 0, \tag{20}$$

Proof.

which means

$$\frac{d}{dt}V(t, \hat{\mathbf{x}}(t)) + L(t, \hat{\mathbf{x}}(t), \hat{\boldsymbol{\theta}}(t)) = 0. \tag{19}$$

Integrating from t_0 to t_1 and using the boundary condition $V(t_1, \mathbf{x}) = \Phi(\mathbf{x})$, we have

$$V(t_0, x_0) = \int_{t_0}^{t_1} L(t, \hat{\mathbf{x}}(t), \hat{\boldsymbol{\theta}}(t))dt + \Phi(\hat{\mathbf{x}}(t_1)) = J[\hat{\boldsymbol{\theta}}]. \tag{20}$$

On the other hand, if θ be any other control whose trajectory is x, we would have

$$\partial_t V(t, \mathbf{x}(t)) + [\nabla_{\mathbf{x}} V(t, \mathbf{x}(t))]^T f(t, \mathbf{x}(t), \boldsymbol{\theta}(t)) + L(t, \mathbf{x}(t), \boldsymbol{\theta}(t)) \geq 0, \tag{21}$$

Proof.

which yields

$$0 \leq \int_{t_0}^{t_1} L(t, \mathbf{x}(t), \boldsymbol{\theta}(t))dt + V(t_1, x(t_1)) - V(t_0, x_0), \tag{19}$$

or

$$J[\hat{\boldsymbol{\theta}}] = V(t_0, x_0) \leq J[\boldsymbol{\theta}]. \tag{20}$$

This shows that $\hat{\boldsymbol{\theta}}$ is globally optimal, with cost $V(t_0, x_0)$.

□

Outline

Outline

Theorem (Euler-Lagrange Equations)

Let x *be an extremum of Eq. 1. Then,* x *satisfies the Euler-Lagrange Equations:*

$$\partial_x L(u, x(u), x'(u)) = \frac{\mathrm{d}}{\mathrm{d}u}\partial_{x'} L(u, x(u), x'(u)), u \in [a, b]. \tag{21}$$

Theorem (Pontryagin's Maximum Principle)

Let $\boldsymbol{\theta}^*$ *be a bounded, measurable and admissible control, and* $\mathbf{x}^*$ *be its corresponding state. Then, there exists an a.c. process* $\mathbf{p}^* = \{\mathbf{p}^*(t) : t \in [t_0, t_1]\}$ *such that*

$$\begin{aligned}
\dot{\mathbf{x}}^*(t) &= \nabla_{\mathbf{p}} H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}^*(t)), \quad \mathbf{x}^*(t_0) = \mathbf{x}_0 \\
\dot{\mathbf{p}}^*(t) &= -\nabla_{\mathbf{x}} H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}^*(t)), \quad \mathbf{p}^*(t_1) = -\nabla_{\mathbf{x}} \Phi(\mathbf{x}^*(t_1)) \\
H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}^*(t)) &\geq H(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\theta}(t)), \quad \forall \boldsymbol{\theta} \in \Theta \text{ and } t \in [t_0, t_1]
\end{aligned} \tag{22}$$

Theorem (Hamilton-Jacobi-Bellman Equations)

The value function V in Eq. 12 is the unique viscosity solution of the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} &\partial_t V(t,\mathbf{x}) + \inf_{\boldsymbol{\theta}} \left\{ L(t,\mathbf{x},\boldsymbol{\theta}) + \left[\nabla_{\mathbf{x}} V(t,\mathbf{x})\right]^\top f(t,\mathbf{x},\boldsymbol{\theta}) \right\} = 0 \\ &V(t_1,\mathbf{x}) = \Phi(\mathbf{x}), \quad (t,\mathbf{x}) \in [t_0,t_1] \times \mathbb{R}^d \end{aligned} \tag{21}$$

Remarks on PMP

PMP establishes the necessary conditions for optimal control problem. PMP obtain open-loop controls that are pre-computed and cannot be applied on-the-fly.

Remarks on HJB

HJB equation establishes the necessary and sufficient conditions for optimal control problem. Provided we can solve the HJB, the optimal control solution is of feed-back or closed-loop form, meaning that it tells how to steer the system by just observing the state trajectory.

Outline

Reference

Dynamical System and Machine Learning