Optimal Control 2018

Kaoru Yamamoto

Optimal Control 2018

Exercise sessions (20%):
Solve 50% of problems in advance. Or make hand-in later.
Mini-project (20%):
Study and present your own optimal control problem.
Written take-home exam (60%).

Optimal control problem

Find a control $u \in U \subset \mathbb{R}^m$ that minimizes the cost

$$J(u) := \int_{t_0}^{t_f} \underbrace{L(t, x(t), u(t))}_{\text{running cost}}\, dt + \underbrace{K(t_f, x_f)}_{\text{terminal cost}}$$

subject to

$$\dot{x} = f(t, x, u),\ x(t_0) = x_0,\ x \in \mathbb{R}^n.$$

- $J(u)$: a function of a function $u(t)$. $\Rightarrow$ cost **functional**
- How do we find such u?
- Revisit finite-dimensional function minimization.

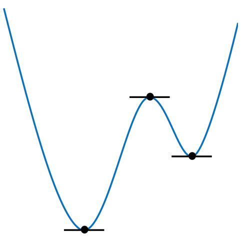
Finite-dimensional function minimization

Consider $f : \mathbb{R}^n \to \mathbb{R}$. Let $D \subset \mathbb{R}^n$. A point x^* is a *local minimum* if $\exists \epsilon > 0$ s.t. $\forall x \in D$ satisfying $|x - x^*| < \epsilon$,

$$f(x^*) \leq f(x).$$

Unconstrained optimization - first-order condition

Consider the case where D is an open subset of $\mathbb{R}^n$.
$\implies$ x^* is an interior point of D.



- Extrema (minima or maxima) occur at *stationary points.*
- **First-order necessary condition** for optimality

$$\boxed{\nabla f(x^*) = 0}$$

where $\nabla f := (f_{x_1}, \dots, f_{x_n})^T$.
- Note: we need $f \in \mathcal{C}^1$.
- A local minimum? $\Rightarrow$ second-order conditions.

Unconstrained optimization - second-order conditions

- Pick an arbitrary vector $d \in \mathbb{R}^n$.
- Since D an open set, $x^* + \alpha d \in D$ for a real parameter $\alpha \approx 0$.
- The Taylor expansion for f around $\alpha = 0$ gives

$$f(x^* + \alpha d) = f(x^*) + \underbrace{\nabla f(x^*)}_{=0} \cdot d\alpha + \frac{1}{2} d^T \nabla^2 f(x^*) d\alpha^2 + o(\alpha^2)$$

where

$$\nabla^2 f := \begin{pmatrix} f_{x_1 x_1} & \dots & f_{x_1 x_n} \\ \vdots & \ddots & \vdots \\ f_{x_n x_1} & \dots & f_{x_n x_n} \end{pmatrix}.$$

- **Second-order necessary condition** for optimality

$$\boxed{\nabla^2 f(x^*) \geq 0}$$

- **Second-order sufficient condition** for optimality

$$\boxed{\nabla f(x^*) = 0 \quad \text{and} \quad \nabla^2 f(x^*) > 0}$$

- Note: we need $f \in \mathcal{C}^2$.

Constrained optimization - Lagrange mulipliers

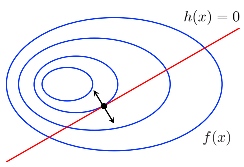
$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && h_1(x) = h_2(x) = \dots = h_m(x) = 0 \end{aligned}$$

where f and $h_i, i = 1, \dots, m$ are $\mathcal{C}^1$ functions from $\mathbb{R}^n$ to $\mathbb{R}$.

First-order necessary condition for constrained optimality:

$$\boxed{\nabla f(x^*) + \lambda_1^* \nabla h_1(x^*) + \dots + \lambda_m^* \nabla h_m(x^*) = 0.}$$

- λ_i^*: Lagrange multipliers.
- Intuition: x^* occurs at a point of tangency.
- Proof: [Liberzon 1.2.2] & E2.



Functional minimization

Minimize a **functional** $J : V \to \mathbb{R}$.

- Many different choices for V.
  – e.g., $\mathcal{C}^k([a, b], \mathbb{R}^n)$ (k-times continuously differentiable)
- "Local minima" of J?
  – **notion of closeness (i.e., norm $\|\cdot\|$) for functions**

Let $A \subset V$ (V equipped with $\|\cdot\|$.) A function $y^* \in A$ is a *local minimum* if $\exists \epsilon > 0$ s.t. $\forall y \in A$ satisfying $\|y - y^*\| < \epsilon$,

$$J(y^*) \leq J(y).$$

- Common norms in function spaces:
  0-norm: $\|y\|_0 := \max_{a \leq x \leq b} |y(x)|$.
  1-norm: $\|y\|_1 := \max_{a \leq x \leq b} |y(x)| + \max_{a \leq x \leq b} |y'(x)|$.

Basic calculus of variations problem

Consider a function $L : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.

Basic Calculus of Variations Problem:
Among all $y \in \mathcal{C}^1([a,b] \to \mathbb{R})$ satisfying given boundary conditions

$$y(a) = y_0,\ y(b) = y_1$$

find (local) minima of the cost functional

$$J(y) := \int_a^b L(x, y(x), y'(x))dx.$$

L: *Lagrangian* or *running cost*

– Note: $\mathcal{C}^1$ assumption not needed if y' does not appear in L.

– multiple-degree-of-freedom (MDOF) case:

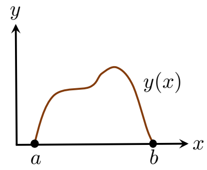
$$y : [a,b] \to \mathbb{R}^n,\ L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}.$$

Dido's isopermetric problem

A legend about the foundation of Carthage around 850 B.C.

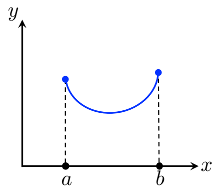
Dido was allowed to have the land along the North Africa coastline that could be enclosed by an oxhide. She sliced the hide into very thin strips so that she was able to enclose a large area.

- Assume a straight coast line.
- Maximize the area given by
  $J(y) = \int_a^b y(x)dx,\ y : [a,b] \to \mathbb{R}.$
- Constraint:
  - $y(a) = y(b) = 0,$
  - $\int_a^b \sqrt{1 + (y'(x))^2}\, dx = C_0.$



Catenary (Galileo, 1630s)

- A fixed length chain with uniform mass density suspended between two fixed points.
- What will be the shape of this chain?
- Galileo claimed parabola - wrong!
- The chain will take the shape of minimal potential energy.



Minimize

$$J(y) = \int_a^b y(x)\sqrt{1 + (y'(x))^2}\, dx, \quad y : [a,b] \to [0, \infty)$$

subject to

$$y(a) = y_0,\ y(b) = y_1 \text{ and } \int_a^b \sqrt{1 + (y'(x))^2}\, dx = C_0.$$

Weak and strong extrema

- Extrema of $J(y)$, $y \in \mathcal{C}^1([a,b] \to \mathbb{R})$ w.r.t
  0-norm $\|y\|_0 := \max_{a \le x \le b} |y(x)|$ – *strong*
  1-norm $\|y\|_1 := \max_{a \le x \le b} |y(x)| + \max_{a \le x \le b} |y'(x)|$ – *weak*
- y^* a strong extremum $\Rightarrow$ y^* a weak extremum (easier)
- Weak minima – not suitable in optimal control.
  - no compelling reason for y'.
  - $\mathcal{C}^1$ requirement too strict.
  - **piecewise** $\mathcal{C}^1$ instead.

First variation and first-order necessary condition

- Recall $f(x^* + \alpha d) = f(x^*) + \nabla f(x^*) \cdot d\alpha + o(\alpha),\ \alpha \approx 0.$
- Consider $y + \alpha\eta$ where $y, \eta \in V,\ \alpha \in \mathbb{R}$. ($\eta$: perturbation.)
- $\eta \in V$ is *admissible* if $y^* + \alpha\eta \in A \subset V,\ \forall \alpha \approx 0.$

A linear functional $\delta J : V \to \mathbb{R}$ is called the *first variation* of J at y if $\forall \eta, \alpha$ we have

$$J(y + \alpha\eta) = J(y) + \delta J(y, \eta)\alpha + o(\alpha).$$

$$\delta J(y, \eta) = \lim_{\alpha \to 0} \frac{J(y + \alpha\eta) - J(y)}{\alpha} \quad \text{(Gateaux derivative of } J\text{)}.$$

First-order necessary condition for optimality: $\forall \eta$ admissible,

$$\boxed{\delta J(y^*, \eta) = 0}$$

Euler-Lagrange equation

- For basic calculus of variations, we have
  $V = \mathcal{C}^1([a,b] \to \mathbb{R}),\ A = \{y \in V : y(a) = y_0,\ y(b) = y_1\},$
  $J(y) = \int_a^b L(x, y(x), y'(x))dx.$
- $J(y + \alpha\eta) = J(y) + \delta J(y, \eta)\alpha + o(\alpha).$
- For η to be admissible, we need $\eta(a) = \eta(b) = 0.$

$$\begin{aligned} J(y + \alpha\eta) &= \int_a^b L(x, y(x) + \alpha\eta(x), y'(x) + \alpha\eta'(x))dx \\ &= \int_a^b (L(x, y(x), y'(x)) + L_y(x, y(x), y'(x))\alpha\eta(x) \\ &\quad + L_{y'}(x, y(x), y'(x))\alpha\eta'(x) + o(\alpha))dx. \end{aligned}$$

Euler-Lagrange equation (cont.)

$$\begin{aligned} \underbrace{\delta J(y, \eta)}_{\text{must be 0}} &= \int_a^b (L_y(x, y(x), y'(x))\eta(x) + L_{y'}(x, y(x), y'(x))\eta'(x))\, dx \\ &= \int_a^b \underbrace{\left(L_y(x, y(x), y'(x)) - \frac{d}{dx} L_{y'}(x, y(x), y'(x))\right)}_{=:\xi(x)} \eta(x)\, dx \\ &\quad + \underbrace{L_{y'}(x, y(x), y'(x))\eta(x)|_a^b}_{=0}. \end{aligned}$$

$$\int_a^b \xi(x)\eta(x)dx = 0, \xi \text{ continuous } \forall \eta \text{ admissible} \implies \xi(x) \equiv 0.$$

Euler-Lagrange equation:

$$\boxed{L_y = \frac{d}{dx} L_{y'}}$$

Shortest path between two points

Find the shortest path between two points a and b in the plane.

$$J(y) = \int_a^b \sqrt{1 + (y'(x))^2}\, dx.$$

Since $L_y = 0,\ \frac{d}{dx} L_{y'} = 0$ (Euler-Lagrange equation).

$$0 = \frac{d}{dx} L_{y'}(x, y(x), y'(x)) = \frac{d}{dx} \frac{y'(x)}{\sqrt{1 + (y'(x))^2}}.$$

$\implies$ y' constant.

$\implies$ y a straight line.

Euler-Lagrange equation – MDOF case

$$y = (y_1, \dots y_n)^T \in \mathbb{R}^n.$$

Euler-Lagrange equation:

$$\boxed{L_{y_i} = \frac{d}{dx} L_{y_i'}, \quad i = 1, \dots n.}$$

Detailed look at Euler-Lagrange equation

$$\underbrace{\delta J(y, \eta)}_{\text{must be } 0} = \int_a^b \underbrace{\Big(L_y(x, y(x), y'(x)) - \frac{d}{dx} L_{y'}(x, y(x), y'(x)) \Big)}_{=:\xi(x)} \eta(x)\, dx.$$

$$\frac{d}{dx} L_{y'}(x, y(x), y'(x)) = L_{y'x}(x, y(x), y'(x)) + L_{y'y}(x, y(x), y'(x)) y'(x) + L_{y'y'}(x, y(x), y'(x)) y''(x).$$

For ξ to be continuous, $L \in \mathcal{C}^2$ & $y \in \mathcal{C}^2$ (!)

Alternative formulation

$$\underbrace{\delta J(y, \eta)}_{\text{must be } 0} = \int_a^b (L_y(x, y(x), y'(x))\eta(x) + L_{y'}(x, y(x), y'(x))\eta'(x))\, dx$$

$$= \int_a^b \underbrace{\Big(L_{y'}(x, y(x), y'(x)) - \int_a^x L_y(w, y(w), y'(w)) dw \Big)}_{=:\xi(x)} \eta'(x)\, dx + \underbrace{\eta(x) \int_a^x L_y(w, y(w), y'(w)) \Big|_a^b}_{=0}.$$

$\int_a^b \xi(x)\eta'(x) dx = 0$, ξ continuous, $\forall \eta$ admissible $\Rightarrow \xi(x)$ constant.

([Liberzon, Lemma 2.2., proof: E1])

Alternative formulation (cont.)

$$L_{y'}(x, y(x), y'(x)) = \int_a^x L_y(w, y(w), y'(w)) dw + C$$

$$\implies \frac{d}{dx} L_{y'}(x, y(x), y'(x)) = L_y(w, y(w), y'(w))$$

Euler-Lagrange equation recovered without extra assumptions.

Variable-endpoint problems

$$\underbrace{\delta J(y, \eta)}_{\text{must be } 0} = \int_a^b \underbrace{\Big(L_y(x, y(x), y'(x)) - \frac{d}{dx} L_{y'}(x, y(x), y'(x)) \Big)}_{=:\xi(x)} \eta(x)\, dx + \underbrace{L_{y'}(x, y(x), y'(x))\eta(x)\Big|_a^b}_{=L_{y'}(b, y(b), y'(b))\eta(b)}$$

- RHS must be 0 $\forall \eta$ admissible, including η s.t. $\eta(b) = 0$.
- $\Rightarrow$ The Euler-Lagrange equation must hold.
- $\Rightarrow$ The entire integral is 0 $\forall \eta$ admissible. (*not only for the ones with* $\eta(b) = 0$)
- The final term must vanish!

An additional necessary condition for optimality:

$$\boxed{L_{y'}(b, y(b), y'(b)) = 0}$$

Shortest path from a point to a vertical line

Find the shortest path from a point a to a vertical line.

$$J(y) = \int_a^b \sqrt{1 + (y'(x))^2}\, dx.$$

$$L_{y'}(b, y(b), y'(b)) = \frac{y'(x)}{\sqrt{1 + (y'(x))^2}} \Big|_{x=b} = 0 \implies y'(b) = 0.$$

- a horizontal tangent at the final point.
- the optimal path between two points - a straight line
- y a horizontal line

Hamilton's canonical equations

The Euler-Lagrange equation $L_y - \frac{d}{dx} L_{y'} = 0$ is equivalent to

$$\frac{d}{dx} (L_{y'} \cdot y' - L) - L_x = 0$$

proof: homework. (hint: $\frac{dL}{dx} = L_x + L_y \cdot y' + L_{y'} \cdot y''$)

momentum $p := L_{y'}(x, y, y')$
Hamiltonian $H(x, y, y', p) := p \cdot y' + L(x, y, y')$

Hamiltonian's canonical equations:

$$\boxed{y' = H_p, \quad p' = -H_y}$$

Maximizing the Hamiltonian

momentum $p := L_{y'}(x, y, y')$
Hamiltonian $H(x, y, y', p) := p \cdot y' - L(x, y, y')$

Observe that

$$H_{y'}(x, y, y', p) = p - L_{y'}(x, y, y') = 0.$$

- $\Rightarrow H$ has a stationary point as a function of y' along an optimal curve $(x, y(x), p$ fixed).
- $H^*(z) := p \cdot z - L(x, y, z)$ then $\frac{dH^*}{dz}(y'(x)) = 0$.
- This stationary point is actually a maximum ($\Rightarrow$ the maximum principle, L3 – L4)

Principle of least action

Hamilton's principle of least action

Trajectories of mechanical systems are extremals of the functional

$$\int_{t_0}^{t_1} (T - U)dt$$

which is called *action integral*.

$$L := \underbrace{\frac{1}{2}m(\dot{q}\cdot\dot{q})}_{T} - U(q), \quad q = (x, y, z)^T$$

$$\frac{d}{dt}L_{\dot{q}} = L_q \iff \frac{d}{dt}\underbrace{(m\dot{q})}_{\text{momentum}} = \underbrace{-U_q}_{\text{force}}$$

$$H = L_{\dot{q}}\cdot\dot{q} - L = T + U \quad \text{(total energy)}$$