Lecture 3

Monday, April 6, 2026 6:28 PM

Review:

Constrained optimization: $D = \{x \in \mathbb{R}^n : h_1(x) = \cdots = h_m(x) = 0\}$, $h_i \in C^1$.

$$\min_{x \in D} f(x) \quad , \quad f \in C^1.$$

Supp. x^* is a local min of f and a regular point of D.

$\hookrightarrow \{\nabla h_i(x^*), i=1,\dots,m\}$ linearly indep.

• 1st-order nec. cond:

$$\nabla f(x^*) \cdot d = 0 \quad \forall d \in T_{x^*}D = \{d : \nabla h_i(x^*) \cdot d = 0, i=1,\dots,m\}.$$

or, equivalently,

$$\nabla f(x^*) \in \operatorname{span}\{\nabla h_i(x^*), i=1,\dots,m\}$$

or, equivalently

$\rightarrow$ $\exists \lambda^* \in \mathbb{R}^m$ s.t. (x^*, λ^*) is a critical point of Lagrangian

$$L(x,\lambda) = f(x) + \langle \lambda, h(x) \rangle \quad ; \quad \text{i.e., } \nabla L(x^*, \lambda^*) = 0.$$

• 2nd-order nec. cond.:

Beware $\{d^T \nabla^2 f(x^*) d \geq 0, \forall d \in T_{x^*}D\}$ is NOT a nec. cond.. why?

$\rightarrow$ (correct cond.) $\exists \lambda^* \in \mathbb{R}^m$ s.t. $\left\{ \begin{array}{l} \nabla_{x,\lambda} L(x^*, \lambda^*) = 0 \\ \text{and} \\ d^T L_{xx}(x^*, \lambda^*) d \geq 0, \forall d \in T_{x^*}D \end{array} \right.$

• 2nd-order suff. cond:

If $x^* \in D$, $\lambda^* \in \mathbb{R}^m$ s.t. $\left\{ \begin{array}{l} \nabla_{x,\lambda} L(x^*, \lambda^*) = 0 \\ \text{and} \\ d^T L_{xx}(x^*, \lambda^*) d > 0, \forall d \in T_{x^*}D \end{array} \right.$

Then, x^* is a strict local min of f.

<u>Infinite dim. optimization via Calculus of variations:</u>

<u>Domain:</u> • $C^0([a,b], \mathbb{R}^n)$ $\quad y : \underbrace{[a,b]}_{\text{"time"}} \longrightarrow \mathbb{R}^n$ continuous

with $\|y\|_0 = \max_{t \in [a,b]} |y(t)|$.

• $C^1([a,b], \mathbb{R}^n)$ $\quad y : [a,b] \to \mathbb{R}^n$ countinously differentiable

with $\|y\|_1 = \max_{t \in [a,b]} |y(t)| + \max_{t \in [a,b]} |y'(t)|$

<u>Functional:</u> $J : V \to \mathbb{R}$ for $V = C^0$ or C^1.

<u>problem:</u> Given $A \subset V$ feasible set, $\min_{y \in A} J(y)$.

<u>local min:</u> Choose $\|\cdot\|_1$ or $\|\cdot\|_2$.

y^* is a local min of J if

$\exists \varepsilon > 0 : \quad J(y^*) \leq J(y), \quad \forall y : \underbrace{\|y - y^*\|_0}_{} < \varepsilon$ or $\|y - y^*\|_1 < \varepsilon$.

<u>First variation of J (Gateaux derivative):</u>

<u>Def:</u> η is an <u>admissible perturbation/direction</u> if for small enough α

$y + \alpha \eta \in A \rightsquigarrow$ feasible set. $(A \subset V)$

we define $\quad \delta J|_y(\eta) \triangleq \lim_{\alpha \to 0} \dfrac{J(y + \alpha\eta) - J(y)}{\alpha}$.

If $\delta J|_y(\eta)$ is a linear functional on η, then we call it "first variation of J at y in direction η".

show: $J(y+\alpha\eta) = J(y) + \delta J|_y(\eta)\cdot\alpha + o(\alpha)$, $\lim_{\alpha\to 0} \frac{o(\alpha)}{\alpha} = 0$.

Second variation of J at y in direction η:

let $$\delta^2 J|_y(\eta) \triangleq \lim_{\alpha\to 0} \frac{J(y+\alpha\eta) - J(y) - \delta J|_y(\eta)\alpha}{\frac{1}{2}\alpha^2}$$

If $\delta^2 J|_y(\eta)$ is a quadratic form in η, then we call it "2nd variation of J at y in direction η".

Recall → "quadratic form": let $B(y,y') : V\times V \to \mathbb{R}$ be a bilinear functional then $Q(y) = B(y,y)$ is quadratic form.

show: $J(y+\alpha\eta) = J(y) + \delta J|_y(\eta)\cdot\alpha + \frac{1}{2}\delta^2 J|_y(\eta)\cdot\alpha^2 + o(\alpha^2)$, $\lim_{\alpha\to 0}\frac{o(\alpha^2)}{\alpha^2} = 0$.

First and second order Nec. Cond. of optimality: $y^* \in A \subset V$

let $g(\alpha) = J(y^* + \alpha\eta)$, the same argument shows

- 1-st order Nec. Cond: $g'(0) = \delta J|_{y^*}(\eta) = 0$, $\forall \eta$ admissible
- 2-nd order Nec. Cond: $g''(0) = \delta^2 J|_{y^*}(\eta) \geq 0$, $\forall \eta$ admissible.

---

what about <u>sufficient</u> Condition ?

Supp. $\delta J|_{y^*}(\eta)=0$, $\delta^2 J|_{y^*}(\eta) > 0$ $\forall \eta$ admissible.

Is it true that y^* is nec. a local min of J? No! because the unit sphere in C^0 or C^1 is not compact, so our previous argument from finite-dimensional opt. fails!

→ we can fix this by using "Fréchet derivative" instead.

- let
$$\underline{DJ|_y}(\eta) \triangleq \lim_{\|\eta\|\to 0} \frac{J(y+\eta)-J(y)}{\|\eta\|}$$

if $DJ|_y(\cdot)$ <u>exists and is linear</u> in η, we call it <u>Fréchet derivative</u>.

→ Every Fréchet derivative is a Gateaux derivative, but not vice versa!

{ Recall the analogy with $\underbrace{\text{directional derivative}}_{\text{Gateaux deriv.}}$ vs $\underbrace{\text{total derivative}}_{\text{Fréchet deriv.}}$ }

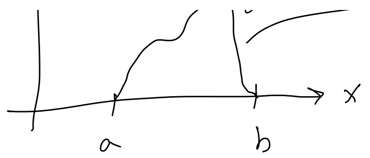
<u>Show:</u> $J(y+\eta) = J(y) + DJ|_y(\eta) + o(\|\eta\|)$, $\lim_{\|\eta\|\to 0} \frac{o(\|\eta\|)}{\|\eta\|} = 0$.

→ For now, we hold our search for suff. Cond., and go back to $\delta J|_y(\cdot)$.

<u>Examples of Variational problems:</u>

1- Dido's Isoperimetric problem:

→ Given a rope (from hide of an ox) with fixed length maximize the enclosed



a fixed length, ...

area on land ($x \geq 0$).

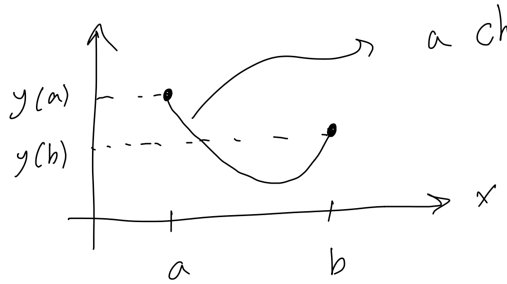
→ Given fixed $[a,b]$, and a length l_0:

$$\max_{\substack{y \in C^1([a,b]) \\ y(a)=y(b)=0}} \underbrace{\int_a^b y(x)\,dx}_{\text{area under the curve}} \quad \text{s.t.} \quad \underbrace{\int_a^b \sqrt{1+(y'(x))^2}\,dx}_{\text{length of curve}} = l_0$$

→ what if <u>b is free</u>?!

$$(ds)^2 = (dx)^2 + (y'(x))^2 (dx)^2$$

2- Catenary (chain) problem:



a chain with uniform mass density.

Find the stationary hanging profile.

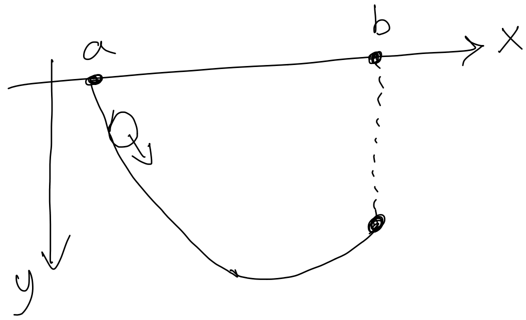
- Given Fixed $y(a)=y_0$, $y(b)=y_1$, length l_0.

$$\min_{\substack{y \in C^1([a,b]) \\ y(a)=y_0,\ y(b)=y_1}} \int_a^b y(x)\cdot\sqrt{1+(y'(x))^2}\,dx \quad \text{s.t.} \quad \int_a^b \sqrt{1+(y'(x))^2}\,dx = l_0$$

→ sol: $y(x) = c\cdot\cosh(x/c)$, for some $c>0$.

3- Brachistochrone (shortest time) problem:

3 - Brachist



- Given fixed a, b, $y(a)=0$, $y(b)=y_b$.
- Consider sliding a ball with no friction, no initial velocity down a curve slide.
- Find the profile of slide that minimizes the time of reach to (b, y_b).

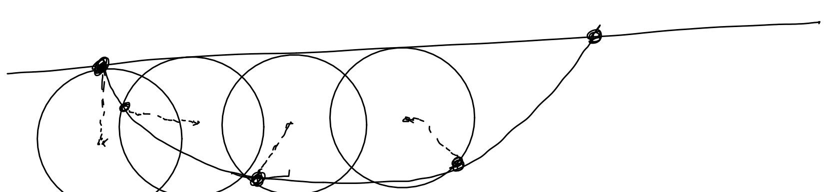
→ Kinetic + potential energy $\equiv \underline{0}$. $\Rightarrow$ $\dfrac{m v^2}{2} - m g y = 0$

⇝ up to constant : $v = \sqrt{y}$.

⇝ time to travel ds : $\dfrac{ds}{v} = \dfrac{\sqrt{1+(y'(x))^2}\,.dx}{\sqrt{y(x)}}$.

$$\min_{\substack{y \in C^1([a,b]) \\ y(a)=0,\ y(b)=y_b}} \int_a^b \frac{\sqrt{1+(y'(x))^2}}{\sqrt{y(x)}}\,.dx$$

→ $\underline{\text{sol}}$: $\underline{\text{cycloids}}$: $\begin{cases} x(\theta) = a + c(\theta - \sin\theta) \\ y(\theta) = c(1-\cos\theta). \end{cases}$ $\theta \in [0, 2\pi]$.



roll a circle without slipping (radius c).

Exercise: Find a variational problem in your field!

Exercise: Find a

Weak vs strong extrema:

Recall $(C^0([a,b],\mathbb{R}^n), \|\cdot\|_0)$ and $(C^1([a,b],\mathbb{R}^n), \|\cdot\|_1)$, and

y^* is an extrema if $\exists \varepsilon > 0 : J(y^*) \underset{(\text{or } \geq)}{\leq} J(y), \forall y : \|y^* - y\|_i < \varepsilon$.

→ If $i = 0$, we call y^* a "strong" local min

→ If $i = 1$, we call y^* a "weak" local min.

Intuition: if y^* is a strong local min then it will be a weak local min; but, not vice versa. why?

Note: weak extrema are often too weak to be useful for optimal control.

Also, requiring $y \in C^1$ is often restrictive.

We will eventually work with piece-wise C^1 functions, meaning

Continuously diff'ble except at finitely many points.