Lecture 10

Thursday, April 30, 2026 8:55 AM

Review:

Thm: (second order sufficient cond.)

An extremal $y(\cdot)$ is a strict weak minimum of the Basic VP

if $L_{zz}(x,y,y') > 0$, $\forall x \in [a,b]$ and the interval $[a,b]$

contains no points conjugate to a.

Proof: y is extremal $\Rightarrow$ $\delta J|_y(\eta) = 0$

$\Rightarrow J(y+\alpha\eta) = J(y) + \delta^2 J|_y(\eta)\alpha^2 + o(\alpha^2)$

$\delta^2 J|_y(\eta) > 0$ for $\eta \not\equiv 0$.

and we already argued that $\delta^2 J|_y(\eta)$ dominates $o(\alpha^2)$.

So, we just need to argue that

$\rightarrow$ Recall that no conjugate point to a over $[a,b]$ implies existence of

a $W(x): [a,b] \rightarrow \mathbb{R}$ satisfying Riccati equation $P(Q+W') = W^2$ on $[a,b]$,

(I) s.t. $$\delta^2 J|_y[\eta] = \int_a^b \underbrace{P(x)}_{>0} \underbrace{\left(\eta'(x) + \frac{W(x)}{P(x)}\eta(x)\right)^2}_{>0 \text{ if } \eta \not\equiv 0} dx.$$

Now, because $P(x) = L_{zz}(x) > 0$

$P(x) > 0$ on $[a,b]$ $\Rightarrow$ $\exists\, \varepsilon > 0$ s.t. $P(x) > \varepsilon > 0$, $\forall x \in [a,b]$.

$$\text{(I)} \Rightarrow \int_a^b \left[(P(x)-\varepsilon)(\eta'(x))^2 + Q(x)\cdot\eta(x)^2\right] dx > 0,$$

for all $\eta \not\equiv 0$ admissible

$$\Rightarrow \int_a^b \left[P(x)(\eta'(x))^2 + Q(x)\cdot(\eta(x))^2\right] dx > \varepsilon \int_a^b (\eta'(x))^2\, dx$$

because (show this!) $\int_a^b \eta^2(x)\,dx \le \frac{(b-a)^2}{2} \int_a^b (\eta'(x))^2\,dx$

$$\Rightarrow \delta^2 J|_y[\eta] = \int_a^b \left[P(x)\cdot(\eta'(x))^2 + Q(x)\cdot(\eta(x))^2\right] > c\cdot\varepsilon\left[\int_a^b ((\eta')^2+\eta^2)\,dx\right]$$

for some positive constant $c = \frac{1}{1+(b-a)^2/2}$.

Recall from (**) for $\frac{c}{2}\cdot\varepsilon > 0$, $\exists\,\delta>0$ s.t.

$|\bar{P}(x,y,y';\alpha)| < \frac{c}{2}\varepsilon$, $|\bar{Q}(x,y,y';\alpha)| < \frac{c}{2}\varepsilon$ $\forall x$, $\forall \eta$, $\forall |\alpha|<\delta$.

$$\Rightarrow \left|\frac{O(\alpha^2)}{\alpha^2}\right| = \left|\int_a^b \left[\bar{P}(\eta')^2 + \bar{Q}\,\eta^2\right]dx\right| \le \frac{c}{2}\cdot\varepsilon\left[\int_a^b ((\eta')^2+\eta^2)\,dx\right]$$

$$\Rightarrow \quad \delta^2 J|_y[\eta]\cdot\alpha^2 + O(\alpha^2) \ge \frac{c}{2}\varepsilon\cdot\left[\int_a^b ((\eta')^2+\eta^2)\,dx\right] \quad \blacksquare$$

Note: • this argument will fail for 0-norm $\|\eta\|_0$, so our sufficient condition is only for weak minimum.

• For the vector case, everything is similar except that

$P(x) = \frac{1}{2} L_{zz} > 0$ P.D. on $[a,b]$

positive definite matrix

and the Ricatti-type equation reads

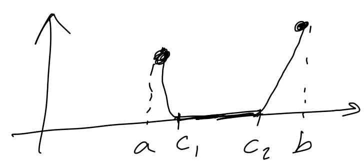
$$\boxed{Q + W' = W P^{-1} W}$$

Chap. 3: From cal. of variations to optimal control!

→ So far, we obtained conditions for weak extermums y to different variational problem: $y \in C^1([a,b]) \Rightarrow \eta \in C^1([a,b])$.

→ In optimal control (and other setting), we are more interested to strong solutions. Here, we focus on the piecewise C^1 curves: $y \in C^0([a,b])$ and y is C^1 except finitely many points

examples: 1) The catenary (chain) problem when the ground is too close to the hanging points compared to the length of the chain:



@ c_1 and c_2 the curve is not C^1!

2) Imagine in controlling a car, sudden break to avoid colissions; a jump in acceleration implies a non-smoothness in the velocity curve!

Def: we call the points where y is not C^1, corner points.

Q: Suppose y is a piecewise-C^1 curve and is a strong exteremal for the Basic VP. what do we know about y?

1) it is a weak extremal between each two corner points.

$\Rightarrow$ must satisfy E-L eq.

more formally, it should be a weak* solution w.r.t. the generalized 1-norm :

$$\|y\|_{1,*} \triangleq \max_{x\in[a,b]} |y(x)| + \max_{x\in[a,b]} \max\{ |y'(x^-)|, |y'(x^+)| \}.$$

2) we will obtain two new conditions for strong extremums:

I) Weierstrass - Erdmann corner condition for strong extremum y :

$p = L_z$ and $H = p \cdot y' - L$ (momentum and Hamiltonian)

are continuous at corner points.

II) we can also generalize the Legendre's condition to corner points for a strong minimum :

$$L_{zz}(x, y(x), y^{\pm}(x)) \geq 0 .$$

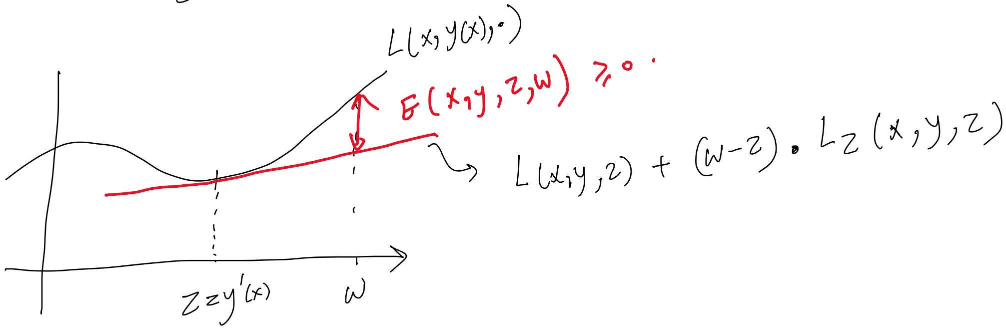
III) Weierstrass - excess condition for strong minimum y :

define $E(x, y, z, w) \triangleq L(x, y, w) - L(x, y, z) - (w - z) \cdot L_z(x, y, z)$

$\forall x \in [a,b]$ non - corner points .

then $E(x, y(x), y'(x), w) \geq 0$, $\forall x \in [a,b]$
$\forall w \in \mathbb{R}^n$

Geometrically, this means $L(x, y(x), \cdot)$ is locally convex at $y'(x)$.



Exercise: show that in Hamiltonian terminology

$$E(x,y,y',w) = H(x,y,y',p) - H(x,y,w,p)$$

where $p = L_z(x,y,y')$ is the momentum. And, by treating p and y' as independent variables,

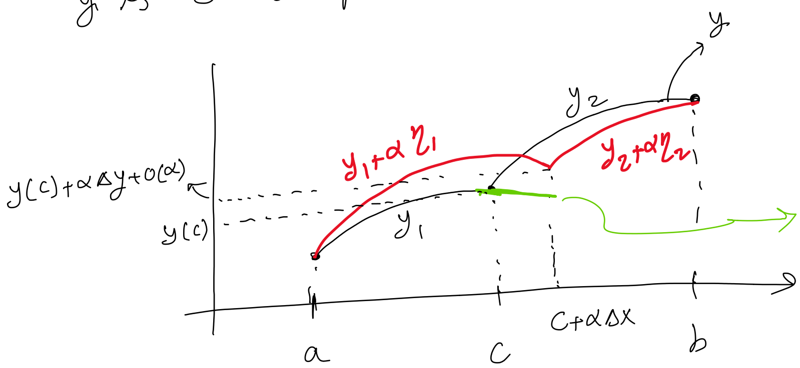
$E \geq 0$ reads that $H^*(z) = H^*(x,y,z,p)$ is maximized at $z = y'(x)$!

Note: The Weierstrass condition for strong minimum is so powerful that implies the Weierstrass-Erdmann corner conditions and the Legendre's condition as collaries.

Exercise: show this! see Exercise 3.3.

→ Let's try to argue the Weierstrass-Erdmann corner conditions:

y is C^1 except at one point $c \in [a,b]$.



$y_1 : [a,c] \to \mathbb{R}$

$y_2 : [c,b] \to \mathbb{R}$

linear extension of y_1.

$\Rightarrow y_1'(c) = y_1'(c^-)$

→ $y_1 + \alpha\eta_1$ is the perturbed curved on $[a, c+\alpha\Delta x]$ where y_1 is extended linearly beyond c (green line) and $\eta_1(a) = 0$

→ $y_2 + \alpha\eta_2$ is defined similarly and $\eta_2(b) = 0$

→ To get a piecewise C^1 curve:

I $\quad y_1(c+\alpha\Delta x) + \alpha\,\eta_1(c+\alpha\Delta x) = y_2(c+\alpha\Delta x) + \alpha\,\eta_2(c+\alpha\Delta x)$.

$$\triangleq y(c) + \alpha\Delta y + o(\alpha).$$

where Δy is defined, and is independent of Δx.

★ Note: we have perturbations $\eta_1, \eta_2, \Delta x, \Delta y$ and they are all conditioned on these equalities to be … we will simplify these later!

admissible!

$\star$ Note: the entire curve is parameterized by the <u>same α</u>,

and we denote it by $y(\cdot,\alpha): [a,b] \to \mathbb{R}$

which is C^1 except at $c+\alpha\Delta x$.

$\to$ Lets consider the first varations of J with respect to $y(\cdot,\alpha)$:

$$J(y(\cdot,\alpha)) = \underbrace{\int_a^{c+\alpha\Delta x} L(x, y_1, y_1')\, dx}_{\triangleq J_1(y_1+\alpha\eta_1)} + \underbrace{\int_{c+\alpha\Delta x}^{b} L(x, y_2, y_2')\, dx}_{\triangleq J_2(y_2+\alpha\eta_2)}$$

Use <u>Leibniz rule</u> to get:

$$\delta J_1\big|_{y_1}(\eta_1) = \frac{d}{d\alpha}\Big|_{\alpha=0} J_1(y_1+\alpha\eta_1)$$

$$= \int_a^c \left(L_y\big|_{y_1} \eta_1 + L_z\big|_{y_1} \eta_1' \right) dx + L(c, y(c), y'(c)) \cdot \Delta x$$

Integration by parts and $\begin{cases} \eta_1(c) \neq 0 \\ y_1'(c) = y_1'(c^-) \end{cases}$

$$= \int_a^b \left(L_y(x, y_1, y_1') - \frac{d}{dx} L_z(x, y_1, y_1') \right) \cdot \eta_1 \, dx$$

$$+ L_z(c, y_1(c), y_1'(c^-))\, \eta_1(c) + L(c, y_1(c), y_1'(c^-)) \cdot \Delta x$$

Similarly,

$$\delta J_2\big|_{y_2}(\eta_2) = \int_c^b \left(L_y(x, y_2, y_2') - \frac{d}{dx} L_z(x, y_2, y_2') \right) \cdot \eta_2 \, dx$$

$$- L_2(c, y(c), y'(c^+))\, \eta_2(c) - L(c, y(c), y'(c^+))\, \Delta x .$$

So :

$$0 = \frac{d}{d\alpha}\Big|_{\alpha=0} J(y(\cdot,\alpha)) = \frac{d}{d\alpha}\Big|_{\alpha=0} \left(J_1(y_1+\alpha\eta_1) + J_2(y_2+\alpha\eta_2) \right)$$

$$= \delta J_1\big|_{y_1}(\eta_1) + \delta J_2\big|_{y_2}(\eta_2) ,$$

(I)

$\forall$ admissible perturbations $(\eta_1, \eta_2, \Delta x, \Delta y)$.

→ Choose the subset of perturbations with $\Delta x = \Delta y = \eta_1(c) = \eta_2(c) = 0$,

then we must have E-L eq for y_1 on $[a, c]$ and y_2 on $[c, b]$.

→ but, now that y_1 and y_2 should satisfy E-L. eq.

for all perturbations, Condition (II) reads :

(III) $$L_2(c, y(c), y'(c^-))\, \eta_1(c) - L_2(c, y(c), y'(c^+))\, \eta_2(c)$$

$$+ L(c, y(c), y'(c^-))\, \Delta x - L(c, y(c), y'(c^+))\, \Delta x = 0$$

$\forall$ admissible $(\eta_1, \eta_2, \Delta x, \Delta y)$.

→ Now, we need to characterize the admissible set better from (I):

look at first variations of (I) w.r.t. α at 0 :

$y_1'(c) = y_1'(c^-)$

$$\boxed{y'(c^-)\, \Delta x + \eta_1(c) = y'(c^+)\, \Delta x + \eta_2(c) = \Delta y}$$

this is necessary for perturbations to be admissible,
but is enough to eliminate η_1, η_2 from (III') to get:

$$- L_z(x, y(x), y'(x)) \Big]_{c^-}^{c^+} \cdot \Delta y$$

$$+ \Big(L_z(x, y(x), y'(x)) \cdot y'(x) - L(x, y(x), y'(x)) \Big) \Big]_{c^-}^{c^+} \cdot \Delta x = 0$$

$\forall$ admissible $\Delta x, \Delta y$ which are independent and totally arbitrary!

$$\Rightarrow L_z \Big]_{c^-}^{c^+} = 0 \quad \text{and} \quad (L_z \cdot y' - L) \Big]_{c^-}^{c^+} = 0 \quad \blacksquare$$