Lecture 1 - 2

Monday, March 30, 2026 8:11 PM

Dynamical system: $\dot{x} = f(t, x, u)$, $x(t_0) = x_0$ initial state

time

control input

state

★ How to measure the performance of an "admissible" input $u(t)$ on an interval of time $t \in [t_0, t_f]$, from an initial cond. (t_0, x_0) to a final condition (t_f, x_{t_f}) ?!

Cost functional:

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

t_f: terminal/final time $\rightsquigarrow$ fixed or free

x_{t_f}: " " state $\rightsquigarrow$ fixed or free

Notice that u is a function of time. So $J(\cdot)$ is a real-valued function on space of (admissible input) functions, thus called a "functional".

The optimal Control problem:

$$\min_{u \in \mathcal{U}} J(u)$$

$u \in \mathcal{U} \rightarrow$ admissible control set

subject to $\begin{cases} \dot{x} = f(t, x(t), u(t)) \\ \text{and possibly conditions on } t_f, x_{t_f} \end{cases}$

$\rightsquigarrow$ "An Infinite-dimensional Constrained optimization"

→ Next, we will review some basic necessary and sufficient conditions of optimality for finite and infinite opt. problems

---

Finite-Dimensional optimization:

let $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$, $|\cdot|$ be 2-norm on $\mathbb{R}^n$

→ A point $x^* \in D$ is a local minimum of f over D if

$\exists \varepsilon > 0 \ni \forall x \in D$ with $|x - x^*| < \varepsilon$: $f(x^*) \leq f(x)$.

→ such x^* is a strict local min if $f(x^*) < f(x)$.

→ " " " global minimum if $f(x^*) \leq f(x)$, $\forall x \in D$.

→ A minimum is not unique unless it is both strict and global.

Unconstrained optimization:

$x^* \in \mathcal{N} \subset D$ where $\mathcal{N}$ is neighborhood of x^*.
"open subset containing x^*"

→ simplest case: $D = \mathbb{R}^n$.

→ D is open subset of $\mathbb{R}^n$ containing the local min x^*.

First-order necessary condition for optimality:

- suppose f is C^1 (continuously differentiable) with $x^* \in D$ (open) being a local min.
- pick any admissible direction "d":

$x^* + \alpha d \in D$ for small enough $\alpha \in \mathbb{R}$.

- Define $g(\alpha) := f(x + \alpha d) : [-\varepsilon, \varepsilon] \subset \mathbb{R} \to \mathbb{R}$.
- x^* is a local min for $f \Rightarrow \alpha = 0$ is a minimum of g.
- First order Taylor expansion of g:

$$g(\alpha) = g(0) + g'(0) \cdot \alpha + \underbrace{o(|\alpha|)}_{\text{higher order terms}}$$

$$\rightarrow \lim_{\alpha \to 0} \frac{o(|\alpha|)}{|\alpha|} = 0$$

- Claim: $g'(0) = 0$! why?

proof: we can argue this by contradiction.

w.l.o.g suppose $g'(0) > 0$ and choose $\alpha^* \in [-\varepsilon, 0)$

such that $g'(0).\alpha^* + o(|\alpha^*|) < 0$.

This is possible because $\frac{o(|\alpha|)}{|\alpha|} \to 0$ as $\alpha \to 0$.

But, then

$$g(\alpha^*) = g(0) + g'(0)\alpha^* + o(|\alpha^*|) < g(0)$$

This contradicts the fact that 0 is a min of g. ∎

- Let's translat $g'(0) = 0$ to f:

$$g'(\alpha) = \nabla f(x^* + \alpha d) \cdot d \quad \text{where } \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}.$$

So $\quad 0 = g'(0) = \nabla f(x^*) \cdot d$ for any admissible $\underline{d}$.

$$\Rightarrow \quad \boxed{\nabla f(x^*) = 0} \rightsquigarrow$$ First-order Nec. Cond. for optimality if f is C^1 and x^* is an interior point D.

Second-order necessary cond.:

- supp. f is C^2 (twice countinuously differentiable) and x^* is a local minimum (interior point).

- similarly, by Taylor expansion

$$g(\alpha) = g(0) + g'(0)\alpha + \frac{1}{2} g''(0)\alpha^2 + o(\alpha^2)$$

where $\lim_{|\alpha| \to 0} \frac{o(\alpha^2)}{\alpha^2} = 0$.

- Claim: $g''(0) \geq 0$! why?

proof: we know that $g'(0) = 0$. For contradiction, supp. $g''(0) < 0$. Choose $\alpha^* \in [-\varepsilon, \varepsilon]$ s.t.

$$\frac{1}{2} g''(0) \alpha^{*2} + o(\alpha^{*2}) < 0.$$

This is possible by def. $o(\alpha^2)$. but then

$g(\alpha^*) < g(0) \rightsquigarrow$ a contradiction ∎

- Translate $g''(0) \geq 0$ back to f:

$$g'(\alpha) = \nabla f(x^* + \alpha d) \cdot d = \sum_i d_i \frac{\partial f}{\partial x_i}\Big|_{(x^* + \alpha d)}$$

$$\Rightarrow g''(\alpha) = \sum_{i,j} d_i d_j \frac{\partial^2 f}{\partial x_j \partial x_i}\Big|_{(x^* + \alpha d)}$$

$$\Rightarrow g''(0) = \sum_{i,j} d_i d_j \frac{\partial^2 f}{\partial x_j \partial x_i}\Big|_{x^*} = d^T \nabla^2 f(x^*) d$$

where $\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$ is Hessian of f.

So, $g''(0) \geq 0$ implies $d^T \nabla^2 f(x^*) d \geq 0$ $\forall$ admissible d

$\Rightarrow \boxed{\nabla^2 f(x^*) \succeq 0 \text{ P.S.D}}$ ⇝ 2nd-order Necessary Cond. of opt. if f is C^2 and x^* is interior point.

P.S.D: positive semi-definite.

Second-order sufficient condition:

claim: If f is C^2, x^* is interior and

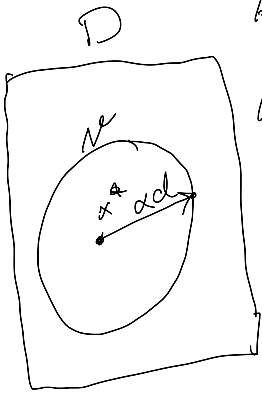
$\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) \succ 0$ P.D. (positive definite)

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

proof: similarly, we have

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

And, for any $d \neq 0$: $d^T \nabla^2 f(x^*) d > 0$.



we wanna show that $f(x^* + \alpha d) > f(x^*)$ for all possible directions d with small enough α, because then x^* will be a strict local minimum.

Fix d. we can pick $\varepsilon > 0$ small enough so that

$$|\alpha| < \varepsilon, \ \alpha \neq 0 \Rightarrow \frac{1}{2} d^T \nabla^2 f(x^*) d \, \alpha^2 + o(\alpha^2) > 0$$

and thus $f(x^* + \alpha d) > f(x^*)$.

→ we need this to be true for all possible d. But the choice of $\varepsilon = \varepsilon(d)$ depends continuously on d.

and there is infinitly many d, and some "d" can be very small.

(small "d" means we need to pick smaller ε to work!)

So it is not clear if such $\varepsilon > 0$ exists uniformly for all d.

→ Luckily, we can update this argument.

Instead, we consider all possible d with $\|d\|_2 = 1$.

Then, $d \in S^1 \rightsquigarrow$ the unit sphere in $\mathbb{R}^n$ is compact.

and the choice of $\varepsilon = \varepsilon(d)$ is a continuous function.

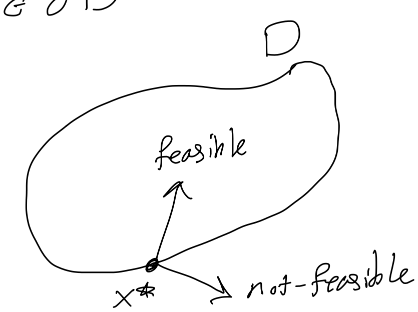
$\Rightarrow$ by Weierstrass thm. $\inf_{d \in S^1} \varepsilon(d) = \varepsilon(d^*) > 0$. ∎

Feasible/Admissible Directions :

suppose D has a boundary and $x^* \in \partial D$

→ d is feasible direction at x^* if

$x^* + \alpha d \in D$ for small enough $\alpha > 0$



Exercise 1: If all directions are not feasible then $\nabla f(x^*) = 0$ is not a necessary condition for optimality of x^*. why?

Instead, we can show that:

- x^* is a local minimum $\Rightarrow$ $\boxed{\nabla f(x^*) . d \geq 0 \quad \text{for all feasible } d}$

"1-st order Nec. Cond."

- x^* is a local minimum $\Rightarrow$

$$\boxed{d^T \nabla^2 f(x^*) d \geq 0 \quad \text{for all feasible direction } d \text{ with } \nabla f(x^*).d = 0}$$

"2-nd order Nec. Cond."

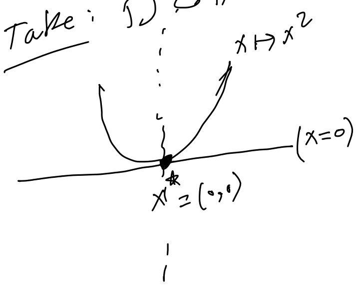
Exercise 2: Supp. f is C^2, $x^* \in D$ and

$\nabla f(x^*) . d \geq 0$ and $d^T \nabla^2 f(x^*) d > 0$ for all admissible d $(d \neq 0)$.

Is x^* necessarily a local min of f?

(prove or give a counter example).

Take: $D \subseteq \mathbb{R}^2$



$D = \{(x,0) : x \in \mathbb{R}\} \cup \{(x, x^2) : x \in \mathbb{R}\}$ and $f(x,y) = x^2 - 2y$

$$\nabla f(x,y) = \begin{bmatrix} 2x \\ -2 \end{bmatrix}$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

- Show feasible direction at x^* is only $d = (a,0)$ for any $a \in \mathbb{R}$.
- show $\nabla f(x^*) \cdot d \geq 0$ for any feasible d.
- show $d^T \nabla^2 f(x^*) d \geq 0$ " " " .
- show that x^* is not a local min. of f.

what was the issue up here? what if D has to be convex?