Lecture 6. Convex functions.

more about separating and supporting hyperplanes.

① why did we emphasis "finite" halfspaces when define polytope.

any closed convex set can be expressed as intersection of a family of halfspaces : all supporting hyperplanes.

proof: $C \subseteq$ intersection trival.

$C \supseteq$ intersection $\forall v \notin C$. suppose $u \in C$. s.t. $\|u-v\| = dist(C, v)$. separating C and $\{v\}$ at u.

② converse theorem of separating hyperplanes.

convex, disjoint $\}$ $\Rightarrow$ separating.

Thm. any convex sets C. and D. at least one of which is open. C. D disjoint iff there exists a separating hyperplane.

Proof: $\Leftarrow$ w.l.o.g. assume C is open and $f(x) = w^T x + b$.

s.t. $f(x) \le 0$ for all $x \in C$. $f(x) \geq 0$ for all $x \in D$.

if $C \cap D \neq \phi$. $\forall x \in C \cap D$. $f(x) = 0$.

C is open $\Rightarrow \exists \varepsilon > 0$. $B(x, \varepsilon) \subseteq C$. $\Rightarrow \exists y \in C$. $f(y) > 0$. contradiction.

③ converse theorem of supporting hyperplane.

Thm. $\forall$ closed set $C \subseteq \mathbb{R}^n$ has nonempty interior.

C has a supporting hyperplane at any $x \in \partial C \Rightarrow$ convex.

New topic: convex functions.

Def. (Jensen's inequality). a function $f: D \xrightarrow{\subseteq \mathbb{R}^n} \mathbb{R}$ is convex.

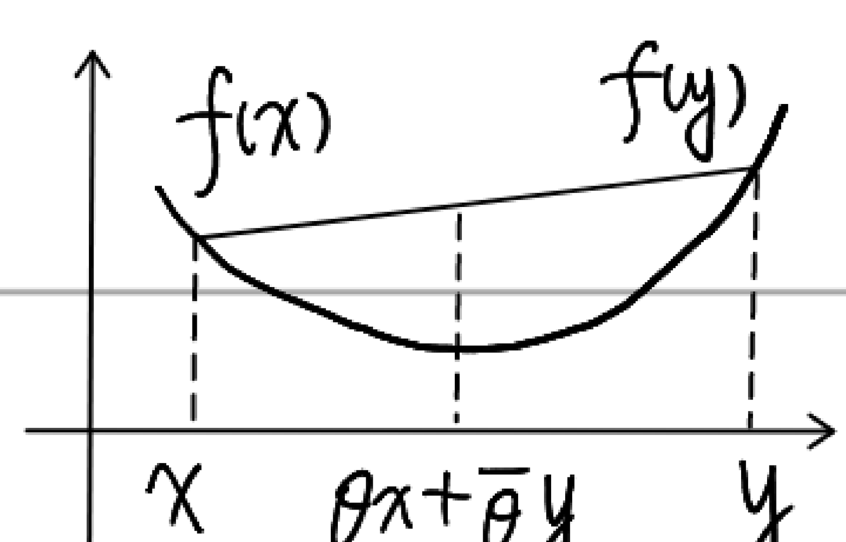
if $\forall x, y \in D \triangleq \mathrm{dom}\, f$, and $\forall \theta \in (0,1)$.

$$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y).$$

strictly convex: if $f(\theta x + (1-\theta) y) < \theta f(x) + (1-\theta) f(y)$

We say f is concave if $-f$ is convex.

strictly concave if $-f$ is strictly convex.



a technical point: domain of f

D must be a convex set.

Extended-value extension: for convienience. extend $\mathbb{R} \to \mathbb{R} \cup \{\infty\}$.

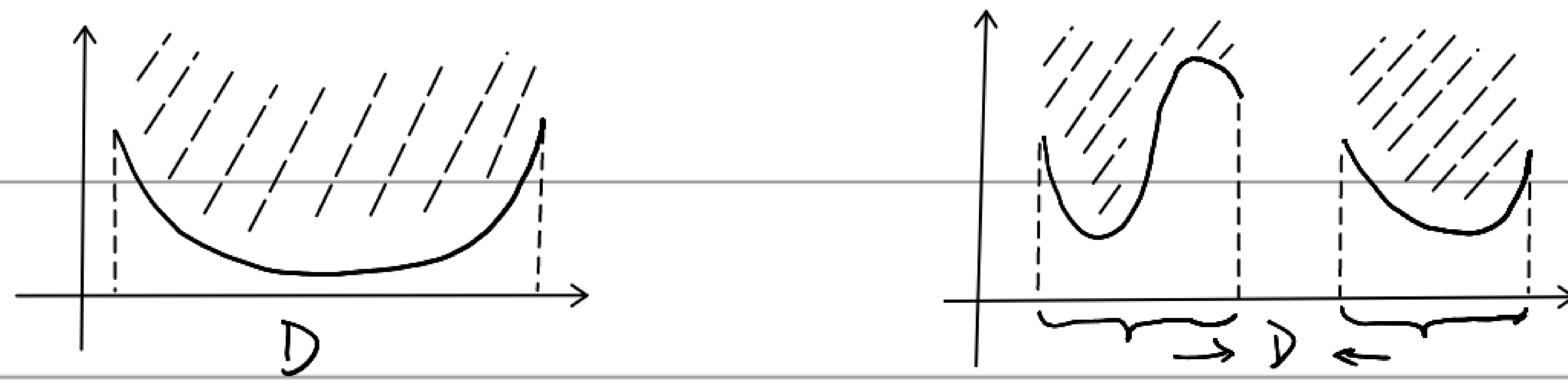
$$\tilde{f}: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} = \begin{cases} f(x) & x \in D \\ \infty & x \notin D. \end{cases} \quad D = \mathrm{dom}\, f.$$

where by convention. $x + \infty = \infty$. $x \cdot \infty = \infty$. $0 \cdot \infty = 0$.

Example: $f(x) = \begin{cases} 1/x & x > 0 \\ \infty & x \leq 0 \end{cases}$ is a convex function.

Epigraph: $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$. graph $\{(\vec{x}, f(\vec{x})) \in \mathbb{R}^{n+1} : \vec{x} \in D\}$.

epigraph of f: $\text{epi} f = \{(\vec{x}, y) \in \mathbb{R}^{n+1} : \vec{x} \in D, y \geq f(\vec{x})\}$



Thm: $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ is a convex function iff epi f convex.

Proof: $\Rightarrow$: f is convex. goal: epi f is convex.

Given $(x_1, y_1), (x_2, y_2) \in \text{epi} f$. $\theta \in [0, 1]$.

let $(x, y) = (\theta x_1 + \bar{\theta} x_2, \theta y_1 + \bar{\theta} y_2)$.

f convex $\Rightarrow x \in D$. $f(x) \leq \theta f(x_1) + \bar{\theta} f(x_2)$.

$(x_i, y_i) \in \text{epi} f \Rightarrow f(x_i) \leq y_i \Rightarrow f(x) \leq \theta y_1 + \bar{\theta} y_2 = y$.

$\Leftarrow$: epi f is convex. goal: f is convex. (first. D is convex).

Given. $x_1, x_2 \in D$. $\theta \in [0, 1]$. let $x = \theta x_1 + \bar{\theta} x_2$

$(x_1, f(x_1)), (x_2, f(x_2)) \in \text{epi} f \Rightarrow (x, \theta f(x_1) + \bar{\theta} f(x_2)) \in \text{epi} f$

$\Rightarrow x \in D$, and $\theta f(x_1) + \bar{\theta} f(x_2) \geq f(x)$. $\square$.

Examples of convex functions: e^{ax}. $-\log x$. x^a if $a \geq 1$ or $a \leq 0$

strictly if $a \neq 0, 1$.

affine functions are convex and concave. not strictly.

Example: any norm $\|\cdot\|: \mathbb{R}^n \to \mathbb{R}$ is convex.

$$\|\theta x + \bar{\theta} y\| \le \|\theta x\| + \|\bar{\theta} y\| = \theta \|x\| + \bar{\theta} \|y\|.$$

but not strictly. why? ($r\|x\| = \|rx\|$).

verify $-\log x$ is convex: $\theta \log x + \bar{\theta} \log y \le \log(\theta x + \bar{\theta} y)$.

$\iff x^{\theta} \cdot y^{\bar{\theta}} \le \theta x + \bar{\theta} y$. if $\theta = 1/2$. easy. general θ ?

more generally. $f(X) = -\log\det(X)$ where $X \in S_{++}^n$

Midpoint convex: $f(\frac{x+y}{2}) \le \frac{f(x)+f(y)}{2}$ $\forall x. y \subseteq \mathrm{dom} f$.

Thm (Jensen, 1905). midpoint convex continuous functions are convex.

remark: if not continuous. conterexample cannot be measurable.

Proof: By contradiction. $\exists x. y. \theta$. $f(\theta x + \bar{\theta} y) > \theta f(x) + \bar{\theta} f(y)$.

let $g(\alpha) = f(\alpha x + \bar{\alpha} y) - \alpha f(x) - \bar{\alpha} f(y)$. $\alpha \in [0, 1]$.

$\exists \alpha \in (0, 1)$. $g(\alpha) > 0$. $M \triangleq \max_{\alpha \in [0,1]} g(\alpha) > 0$. (compact. continuous)

let $\alpha_0 = \arg\max g(\alpha)$. (smallest value of possible α). $g(\alpha_0) = M$.

δ sufficiently small. s.t. $(\alpha_0 - \delta, \alpha_0 + \delta) \subseteq (0, 1)$.

midpoint convexity of $f \Rightarrow g(\alpha_0 - \delta) + g(\alpha_0 + \delta) \geq 2 g(\alpha_0)$.

However. $g(\alpha_0 - \delta) < M$. $g(\alpha_0) = M$. $g(\alpha_0 + \delta) \le M$. contradiction. $\square$

proof of convexity of $f(X) = -\log\det(X)$. $X \subseteq S^n_{++}$:

midpoint convexity: $\det((X+Y)/2) \geq \det(X)^{1/2} \det(Y)^{1/2}$.

$\Longleftrightarrow \left|\frac{I + X^{-1}Y}{2}\right| \geq |X^{-1}Y|^{1/2}$ (X positive definite $\Rightarrow \exists X^{-1}$)

$\Longleftrightarrow \prod\left(\frac{1+\lambda_i(X^{-1}Y)}{2}\right) \geq \sqrt{\prod \lambda_i(X^{-1}Y)}$. $\Longleftarrow \lambda_i(X^{-1}Y) \geq 0$.

$\lambda_i(X^{-1}Y) = \lambda_i(X^{1/2}X^{-1}YX^{-1/2}) = \lambda_i(X^{-1/2}YX^{-1/2})$.

$X^{-1/2}$ symmetric. $Y \succ 0 \Rightarrow w^T X^{-1/2} Y X^{-1/2} w = \tilde{w}^T Y \tilde{w} > 0$. $\square$.

Zeroth-order condition (restriction to lines).

Thm: f is convex iff for any $x \in \mathrm{dom}\, f$. and any direction $d \in \mathbb{R}^n$

the function $g(t) = f(x+td)$ is convex. where $x+td \in \mathrm{dom}\, f$.

Proof: "$\Rightarrow$". assume f is convex. given any $x \in \mathrm{dom}\, f$. $d \in \mathbb{R}^n$.

let $t_1, t_2 \in \mathrm{dom}\, g$. $\theta \in [0,1]$. $x_i = x + t_i d$.

$\tilde{x} = \theta x_1 + \bar{\theta} x_2 = x + (\theta t_1 + \bar{\theta} t_2) d$ $\triangleq \tilde{t}$

$t_i \in \mathrm{dom}\, g \Rightarrow x_i \in \mathrm{dom}\, f \Rightarrow \tilde{x} \in \mathrm{dom}\, f. \Rightarrow \tilde{t} \in \mathrm{dom}\, g$.

convexity of $f \Rightarrow f(\tilde{x}) \leq \theta f(x_1) + \bar{\theta} f(x_2)$.

$$\Rightarrow g(\tilde{t}) \leq \theta g(t_1) + \bar{\theta} g(t_2).$$

"$\Leftarrow$". assume g is convex. given any $x, y \in \mathrm{dom}\, f$. let $d = y - x$.

$g(t) = f(x + td)$. $x, y \in \mathrm{dom}\, f \Rightarrow 0, 1 \in \mathrm{dom}\, g$.

$\forall \theta \in [0,1]$. convexity of $g \Rightarrow \theta \in \mathrm{dom}\, g \Rightarrow x + \theta d \in \mathrm{dom}\, f$.

$\Leftrightarrow (1-\theta)x + \theta y \in \mathrm{dom}\, f \Rightarrow \mathrm{dom}\, f$ is convex.

$$g(\theta) \le (1-\theta) g(0) + \theta g(1).$$

$$\Rightarrow f((1-\theta)x + \theta y) = f(x + \theta d) \le (1-\theta) f(x) + \theta f(y). \quad \square$$

Example: $f(x): \mathbb{R}^n \to \mathbb{R} = e^{x_1 + x_2 + \cdots + x_n}$ is convex.

$\forall u, v \in \mathbb{R}^n$. $g(t) = f(u + tv) = e^{u_1 + u_2 + \cdots + u_n} \cdot e^{(v_1 + v_2 + \cdots + v_n)t}$ $\square$.

First-order condition.

Thm. Let f differentiable. Then f is convex iff $\mathrm{dom}\, f$ convex. and

$$f(y) \ge f(x) + \nabla f(x)^T (y - x). \quad \forall x, y \in \mathrm{dom}\, f.$$

Second-order condition.

Thm. Let f twice continuously differentiable on an open domain.

Then f is convex iff $\mathrm{dom}\, f$ is convex. and its Hessian

$$H(x) = \nabla^2 f(x) \succeq 0$$ is positive semidefinite. $\forall x \in \mathrm{dom}\, f$.