CPSC 406 : Computational Optimization

Convex Functions

Convex Functions

- A function $f: C \to \mathbb{R}$ is <u>convex</u> if $C \subseteq \mathbb{R}^n$ is convex and

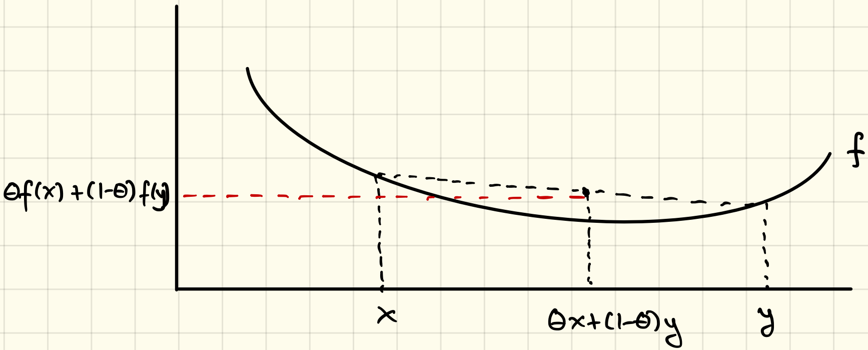
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

for all $x, y \in C$ and $\theta \in [0,1]$.

- f is <u>strictly</u> <u>convex</u> if for all $x, y \in C$ and $\theta \in (0,1)$,

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$$

- f is <u>concave</u> if $-f$ is convex



EXAMPLES

CONVEX FUNCTIONS

- affine : $a^T x + \beta$ for any $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$
- exponential : $e^{\alpha x}$ for any $\alpha \in \mathbb{R}$
- powers : x^α on $\mathbb{R}_{++}$ for all $\alpha \geq 1$ or $\alpha \leq 0$
- abs val : $|x|^p$ for all $p \geq 1$
- neg entropy : $x \log x$ on $\mathbb{R}_{++}$
- norms (use triangle ineq. to show)

CONCAVE FUNCTIONS

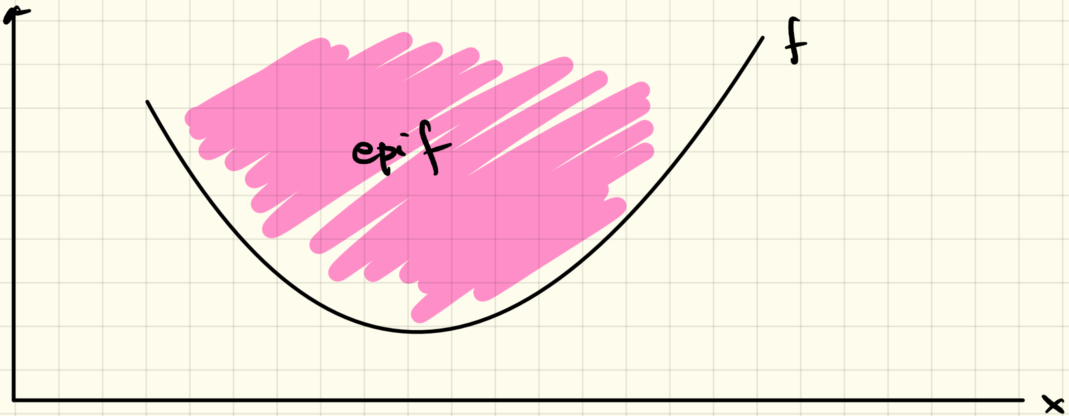
- affine : (see above)
- powers : x^α on $\mathbb{R}_{++}$ for all $0 \leq \alpha \leq 1$
- logarithm : $\log x$ on $\mathbb{R}_{++}$

Equivalent Def'n for convexity of functions:

Def'n (II) A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex if the epigraph,

$$\text{epi} f = \{ (x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \alpha \} \subseteq \mathbb{R}^{n+1}$$

is a convex set.

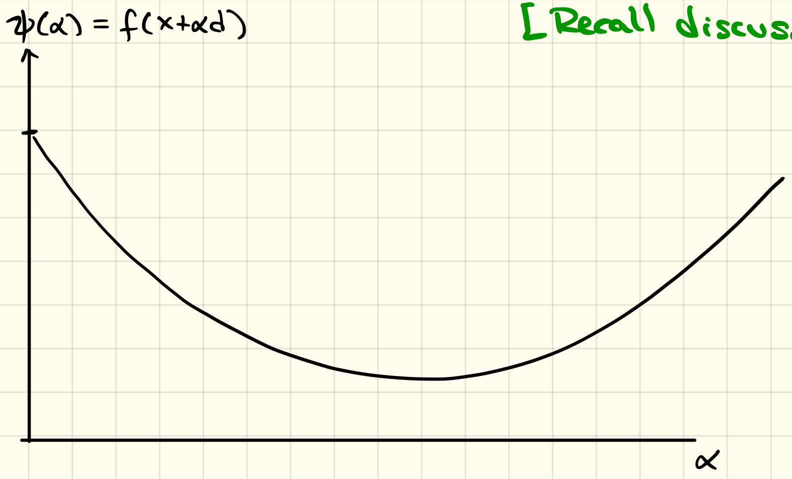


RESTRICTION TO LINES

The function $f : C \to \mathbb{R}$ ($C \subseteq \mathbb{R}^n$ cvx) is convex iff

$$\psi(\alpha) = f(x + \alpha d)$$

is convex for all $x \in C$ and $d \in \mathbb{R}^n$.



[Recall discussion on line search]

Operations that Preserve Convexity

- non-negative multiple

$$\alpha f \text{ is convex if } f \text{ is convex and } \alpha \geq 0.$$

- sum (including infinite sums)

$$f_1 + f_2 \text{ is convex if } f_1, f_2 \text{ are convex.}$$

- composition with affine function

$$f(Ax+b) \text{ is convex if } f \text{ is convex}$$

Examples

- log barrier: $f(x) = -\sum_{i=1}^{m} \log(a_i^T x - b_i)$ convex over $\{x \mid a_i^T x > b_i\} \forall i$

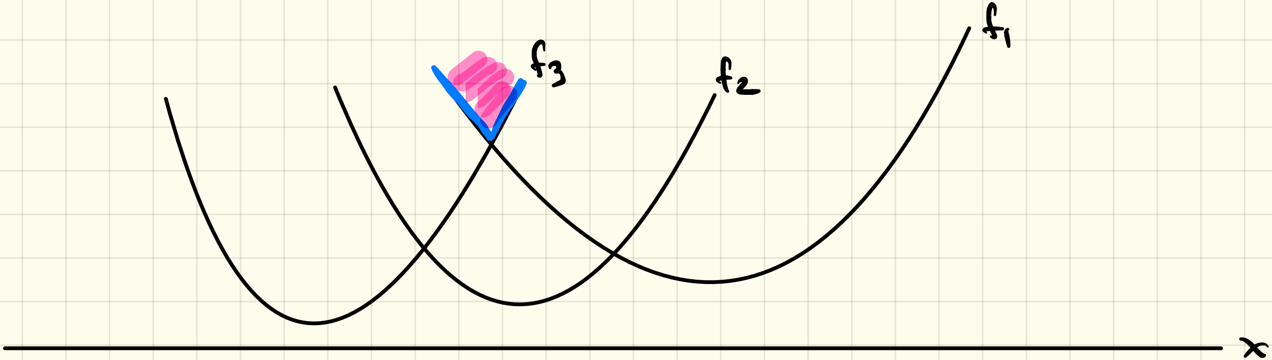
- norm of affine function: $f(x) = \| Ax - b \|$ [composition w/ affine + Δ-inequality]

- $f(x_1, x_2, x_3) = e^{x_1 - x_2 + x_3} + e^{2x_2} + x_1$

Suppose f_i, $i = 1 \cdots m$, are convex functions.

Then $f(x) = \max\{f_1(x), f_2(x), \ldots, f_m(x)\}$

is convex.



Global Optimality

$$(P) \quad \underset{x}{\text{minimize}} \quad f(x) \quad \text{subj to} \quad x \in C$$

where f is a convex function and C is a convex set.

If x^* is a local minimizer of (P) $[f(x^*) \le f(x) \ \forall\ x \in C \cap \mathbb{B}_\varepsilon(x^*)]$
it is also a <u>global minimizer</u> of (P) $[f(x^*) \le f(x) \ \forall\ x \in C \quad]$

<u>Proof</u> Suppose $\bar{x}$ is a <u>local</u>, but not global, minimizer. Then,

- there exists $y \in C$ such that $f(y) < f(\bar{x})$.

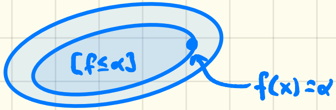
- $$\begin{aligned} f(\theta\bar{x} + (1-\theta)y) &\le \theta f(\bar{x}) + (1-\theta) f(y) \\ &< \theta f(\bar{x}) + (1-\theta) f(\bar{x}) \\ &= f(\bar{x}) \end{aligned}$$

which contradicts hypothesis.

Level Sets

The level set of a function $f : C \to \mathbb{R}$ is the set

$$[f \leq \alpha] := \{ x \in C \mid f(x) \leq \alpha \}$$



If f is convex $\Longrightarrow$ all of its level sets are convex.

(↑ note that not $\Longleftarrow$)

Proof

Take $x, y \in [f \leq \alpha]$. Then $f(x) \leq \alpha$ and $f(y) \leq \alpha$.

Because f is convex, for $\theta \in [0,1]$

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

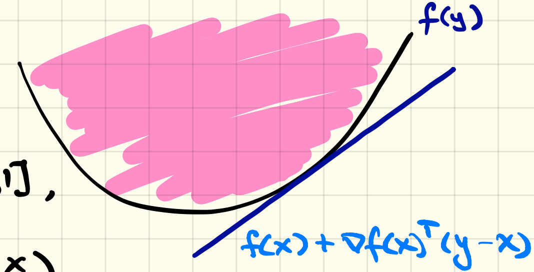
Thus, $\theta x + (1-\theta) y \in [f \leq \alpha]$.

Corollary The set of minimizers of (P) is convex.

First-Order Characterization

Let $f: C \to \mathbb{R}$ be cont. diff over $C \subseteq \mathbb{R}^n$ (cvx). Then f is convex iff

$$f(x) + \nabla f(x)^T (y-x) \leq f(y) \quad \text{for all } x, y \in C$$



Proof sketch:

By convexity of f, $\forall\, x, y \in C$, $\theta \in [0,1]$,

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

$$\Rightarrow \quad \frac{f(\theta y + (1-\theta)x) - f(x)}{\theta} \leq f(y) - f(x)$$

Take lim $\theta \searrow 0$, then

$$\left.\begin{array}{l} \underbrace{f'(x; y-x)}_{= \nabla f(x)^T (y-x)} \leq f(y) - f(x) \end{array}\right\} \Rightarrow f(y) \geq f(x) + \nabla f(x)^T (y-x)$$

FIRST-ORDER SUFFICIENCY

By convexity

$$f(z) \geq f(x) + \nabla f(x)^T (z-x) \qquad \forall\, x, z.$$

Suppose $x = x^*$ $\Rightarrow$ $\nabla f(x^*) = 0$. Then

$$f(z) \geq f(x^*) \qquad \forall\, z.$$

Second-Order Characterization

For $f : \mathbb{R}^n \to \mathbb{R}$ twice continuously differentiable on an open convex set $C \subseteq \mathbb{R}^n$, then f is convex over C <u>iff</u>

$$\nabla^2 f(x) \succeq 0 \qquad \forall x \in C$$

Proof sketch (only $\Leftarrow$)

- By the second-order Taylor expansion of f:

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(z)(y-x)$$

for $x, y \in C$ and some $z \in [x, y]$.

- Because $\nabla^2 f(x) \succeq 0$, $(y-x)^T \nabla^2 f(z)(y-x) \geq 0$
- Hence,

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \qquad \forall x, y \in C$$

and f is convex over C.

<u>Example</u> $f(x) = x^\alpha$ for $x \geq 0$ $\alpha \geq 0$

- Differentiating twice over $x > 0$:

$$f''(x) = \alpha(\alpha-1)\, x^{\alpha-2}$$

- $f''(x) \geq 0$ if $\alpha \geq 1$
- $f''(x) \leq 0$ if $0 \leq \alpha \leq 1$