CS 2601 Linear and Convex Optimization

4. Convex functions (part 2)

Bo Jiang

John Hopcroft Center for Computer Science
Shanghai Jiao Tong University

Fall 2022

Outline

- Properties of convex functions

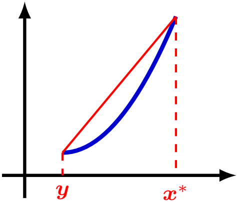
- Convexity-preserving operations

Global minima of convex functions

Theorem. Let f be a convex function defined over a convex set S. If $\boldsymbol{x}^* \in S$ is a local minimum of f, then it is also a global minimum of f over S.

Proof. Suppose there exists $\boldsymbol{y} \in S$ and $\boldsymbol{y} \neq \boldsymbol{x}^*$ s.t. $f(\boldsymbol{y}) < f(\boldsymbol{x}^*)$. For $\theta \in (0,1)$, let $\boldsymbol{x}_\theta = \theta \boldsymbol{y} + \bar{\theta}\boldsymbol{x}^*$. Then

$$f(\boldsymbol{x}_\theta) \leq \theta f(\boldsymbol{y}) + \bar{\theta} f(\boldsymbol{x}^*) < \theta f(\boldsymbol{x}^*) + \bar{\theta} f(\boldsymbol{x}^*) = f(\boldsymbol{x}^*)$$



But $\boldsymbol{x}_\theta \in S$ by convexity of S, and

$$\|\boldsymbol{x}_\theta - \boldsymbol{x}^*\| = \theta \|\boldsymbol{x}^* - \boldsymbol{y}\| \to 0 \quad \text{as} \quad \theta \to 0$$

contradicting the assumption that $\boldsymbol{x}^*$ is a local minimum.

Note. This theorem does not assert the existence of a global minimum in general! It assumes the existence of a local minimum to start with.

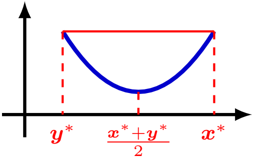
Example. $f(x) = e^x$ has no global or local minimum over $\mathbb{R}$.

Global minima of convex functions (cont'd)

Theorem. Let f be a strictly convex function defined over a convex set S. If $\boldsymbol{x}^* \in S$ is a global minimum of f, then it is unique.

Proof. Suppose there exists $\boldsymbol{y}^* \in S$ and $\boldsymbol{y}^* \neq \boldsymbol{x}^*$ s.t. $f(\boldsymbol{y}^*) = f(\boldsymbol{x}^*)$. By strict convexity,

$$f\left(\frac{\boldsymbol{x}^* + \boldsymbol{y}^*}{2}\right) < \frac{1}{2} f(\boldsymbol{x}^*) + \frac{1}{2} f(\boldsymbol{y}^*) = f(\boldsymbol{x}^*)$$



contradicting the global optimality of $\boldsymbol{x}^*$.

Note. Strict convexity is a sufficient condition for unique global minimum, but it is not necessary!

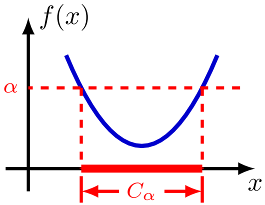
Example. $f(x) = |x|$ has a unique global minimum $x^* = 0$, but it is not strictly convex.

Note. Similar results hold for maxima of concave functions.

Sublevel sets

The α-sublevel set of a function f is

$$C_\alpha = \{\boldsymbol{x} \in \mathrm{dom} f : f(\boldsymbol{x}) \le \alpha\}$$



Theorem. Sublevel sets of a convex function are convex.

Note. For concave f, superlevel set $\{\boldsymbol{x} \in \mathrm{dom} f : f(\boldsymbol{x}) \ge \alpha\}$ is convex.

是对x而言的

Examples.

- Halfspace $H = \{\boldsymbol{x} : \boldsymbol{w}^T\boldsymbol{x} \le b\}$
- Norm ball $\bar{B}(\boldsymbol{x}_0, r) = \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x}_0\| \le r\}$
- Ellipsoid $\mathcal{E} = \{\boldsymbol{x}_0 + \boldsymbol{A}\boldsymbol{u} : \|\boldsymbol{u}\|_2 \le 1\}$, $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{A} \succ \boldsymbol{O}$.

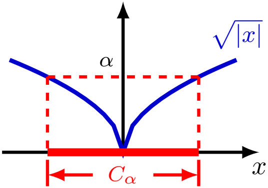
$$\mathcal{E} = \{\boldsymbol{x} : f(\boldsymbol{x}) \le 1\}, \quad f(\boldsymbol{x}) = \|\boldsymbol{A}^{-1}(\boldsymbol{x} - \boldsymbol{x}_0)\|_2^2 = (\boldsymbol{x} - \boldsymbol{x}_0)^T\boldsymbol{A}^{-2}(\boldsymbol{x} - \boldsymbol{x}_0)$$

Sublevel sets (cont'd)

The converse is not true. Nonconvex functions can have convex sublevel sets.

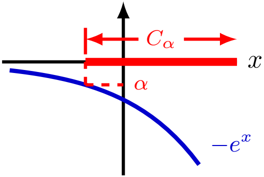
Example. $f(x) = \sqrt{|x|}$ is not convex, but its sublevel sets are all convex,

$$C_\alpha = \begin{cases} \emptyset, & \text{if } \alpha < 0 \\ [-\alpha^2, \alpha^2], & \text{if } \alpha \geq 0 \end{cases}$$



Example. $f(x) = -e^x$ is strictly concave. Its sublevel sets are all convex,

$$C_\alpha = \begin{cases} \emptyset, & \text{if } \alpha \geq 0 \\ [\log(-\alpha), \infty), & \text{if } \alpha < 0 \end{cases}$$



Question. Is the level set $\{\boldsymbol{x} \in \operatorname{dom} f : f(\boldsymbol{x}) = \alpha\}$ convex?

Epigraph 上境图

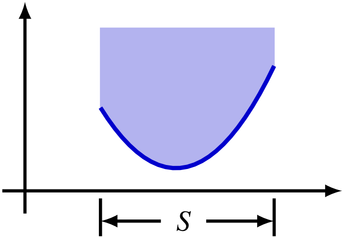
Recall the graph of $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is the set

$$\{(\boldsymbol{x}, f(\boldsymbol{x})) \in \mathbb{R}^{n+1} : \boldsymbol{x} \in S\}$$

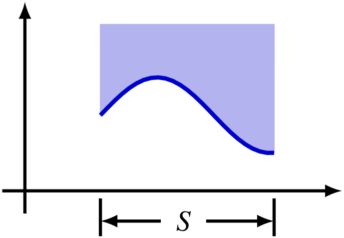
The epigraph[1] of f is

$$\operatorname{epi} f = \{(\boldsymbol{x}, y) \in \mathbb{R}^{n+1} : \boldsymbol{x} \in S, y \geq f(\boldsymbol{x})\}$$

Note. f and its extended-value extension $\tilde{f}$ have the same epigraph.



convex



nonconvex

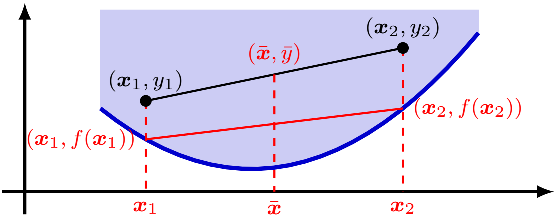
[1] The prefix epi- means “above”, “over”.

Epigraph (cont'd)

充要条件

Theorem. $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function iff $\mathrm{epi} f$ is a convex set.

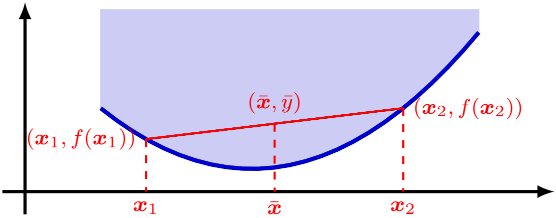
Proof. "$\Rightarrow$". Assume f is convex. Let $(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2) \in \mathrm{epi} f$, $\theta \in [0, 1]$.
Need to show $(\bar{\boldsymbol{x}}, \bar{y}) \triangleq (\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2, \theta y_1 + \bar{\theta} y_2) \in \mathrm{epi} f$.



1. f convex $\implies \bar{\boldsymbol{x}} \in S$ and $f(\bar{\boldsymbol{x}}) \le \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2)$
2. $(\boldsymbol{x}_i, y_i) \in \mathrm{epi} f \implies f(\boldsymbol{x}_i) \le y_i \implies \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2) \le \theta y_1 + \bar{\theta} y_2 = \bar{y}$
3. By 1 and 2, $\bar{\boldsymbol{x}} \in S$ and $f(\bar{\boldsymbol{x}}) \le \bar{y} \implies (\bar{\boldsymbol{x}}, \bar{y}) \in \mathrm{epi} f$

Epigraph (cont'd)

Proof (cont'd). "$\Leftarrow$". Assume $\mathrm{epi} f$ is convex. Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, $\theta \in [0,1]$. Need to show $\bar{\boldsymbol{x}} \triangleq \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2 \in S$ and $f(\bar{\boldsymbol{x}}) \leq \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2) \triangleq \bar{y}$.



1. $f(\boldsymbol{x}_i) \leq f(\boldsymbol{x}_i) \implies (\boldsymbol{x}_i, f(\boldsymbol{x}_i)) \in \mathrm{epi} f$ by definition
2. $\mathrm{epi} f$ convex $\implies (\bar{\boldsymbol{x}}, \bar{y}) = \theta(\boldsymbol{x}_1, f(\boldsymbol{x}_1)) + \bar{\theta}(\boldsymbol{x}_2, f(\boldsymbol{x}_2)) \in \mathrm{epi} f$
3. $\bar{\boldsymbol{x}} \in S$, $f(\bar{\boldsymbol{x}}) \leq \bar{y} = \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2)$ by definition of $\mathrm{epi} f$

Note. The same proof shows the following result: the projection $\{\boldsymbol{x} : (\boldsymbol{x}, \boldsymbol{y}) \in C \text{ for some } \boldsymbol{y}\}$ of a convex set C is convex.

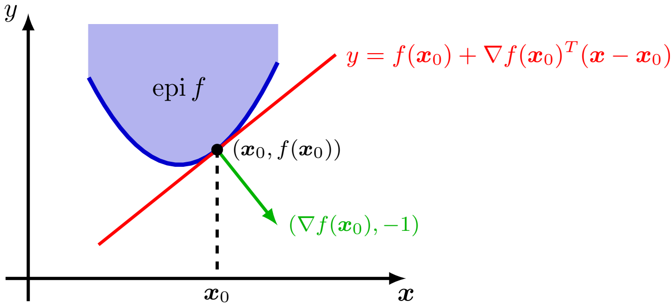
Epigraph (cont'd)

The first-order condition

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^T(\boldsymbol{x} - \boldsymbol{x}_0)$$

implies $y = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^T(\boldsymbol{x} - \boldsymbol{x}_0)$ is a supporting hyperplane of $\mathrm{epi} f$ at $(\boldsymbol{x}_0, f(\boldsymbol{x}_0))$, i.e.

$$\begin{bmatrix} \nabla f(\boldsymbol{x}_0) \\ -1 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{x} \\ y \end{bmatrix} \leq \begin{bmatrix} \nabla f(\boldsymbol{x}_0) \\ -1 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{x}_0 \\ f(\boldsymbol{x}_0) \end{bmatrix}, \quad \forall (\boldsymbol{x}, y) \in \mathrm{epi} f$$



Jensen's inequality

For convex function f, $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$, $\theta \in [0, 1]$

$$f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$$

More generally, for $\boldsymbol{x}_i \in \operatorname{dom} f$, $\theta_i \geq 0$, and $\sum_{i=1}^m \theta_i = 1$,

$$f\left(\sum_{i=1}^m \theta_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^m \theta_i f(\boldsymbol{x}_i)$$

Example. $f(x) = x^2$ is convex over $\mathbb{R}$.

$$\left(\sum_{i=1}^n \frac{1}{n} x_i\right)^2 \leq \sum_{i=1}^n \frac{1}{n} x_i^2 \implies \frac{1}{n} \sum_{i=1}^n x_i \leq \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$$

Example. $f(x) = \log x$ is concave over $(0, \infty)$. For $x_i > 0$,

$$\log\left(\sum_{i=1}^n \frac{1}{n} x_i\right) \geq \sum_{i=1}^n \frac{1}{n} \log x_i \implies \frac{1}{n} \sum_{i=1}^n x_i \geq \sqrt[n]{\prod_{i=1}^n x_i}$$

Hölder's inequality

Dual 对偶

Let $p, q \in (1, \infty)$ be conjugate exponents, i.e. $p^{-1} + q^{-1} = 1$. For $\boldsymbol{x} = (x_1, \ldots, x_n)^T$, $\boldsymbol{y} = (y_1, \ldots, y_n)^T$, Hölder's inequality holds,

$$\sum_{i=1}^{n} |x_i y_i| \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$$

Proof. Assume $\boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{y} \neq \boldsymbol{0}$; otherwise trivial. Let $\tilde{\boldsymbol{x}} = \boldsymbol{x}/\|\boldsymbol{x}\|_p$ and $\tilde{\boldsymbol{y}} = \boldsymbol{y}/\|\boldsymbol{y}\|_q$. The above inequality is equivalent to $\sum_{i=1}^{n} |\tilde{x}_i \tilde{y}_i| \leq 1$.

1. Show $x^{\frac{1}{p}} y^{\frac{1}{q}} \leq \frac{1}{p}x + \frac{1}{q}y$ for $x, y \geq 0$.
   - trivial if $xy = 0$
   - if $xy > 0$, $\log x$ is concave $\implies \log\left(\frac{1}{p}x + \frac{1}{q}y\right) \geq \frac{1}{p}\log x + \frac{1}{q}\log y$
2. Let $x = |\tilde{x}_i|^p$ and $y = |\tilde{y}_i|^q$ in the inequality in 1,
$$|\tilde{x}_i| \cdot |\tilde{y}_i| \leq p^{-1}|\tilde{x}_i|^p + q^{-1}|\tilde{y}_i|^q$$
3. Sum over i and note $\|\tilde{\boldsymbol{x}}\|_p = \|\tilde{\boldsymbol{y}}\|_q = 1$,
$$\sum_{i=1}^{n} |\tilde{x}_i \tilde{y}_i| \leq \frac{1}{p}\|\tilde{\boldsymbol{x}}\|_p^p + \frac{1}{q}\|\tilde{\boldsymbol{y}}\|_q^q = \frac{1}{p} + \frac{1}{q} = 1$$

Minkowski's inequality

For $1 < p < \infty$,

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p \leq \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$$

Proof. Only need to consider case $\|\boldsymbol{x}+\boldsymbol{y}\|_p > 0$.

- $\|\boldsymbol{x}+\boldsymbol{y}\|_p^p = \sum_i |x_i+y_i|^p \leq \sum_i |x_i| \cdot |x_i+y_i|^{p-1} + \sum_i |y_i| \cdot |x_i+y_i|^{p-1}$
- Let $\frac{1}{p}+\frac{1}{q}=1$. Applying Hölder's inequality and $(p-1)q=p$,

$$\sum_i |x_i| \cdot |x_i+y_i|^{p-1} \leq \|\boldsymbol{x}\|_p \left(\sum_i |x_i+y_i|^{(p-1)q}\right)^{1/q} = \|\boldsymbol{x}\|_p \|\boldsymbol{x}+\boldsymbol{y}\|_p^{p/q}$$

- Interchanging x and y, $\sum\limits_i |y_i| \cdot |x_i+y_i|^{p-1} \leq \|\boldsymbol{y}\|_p \|\boldsymbol{x}+\boldsymbol{y}\|_p^{p/q}$
- Combining above inequalities,

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p^p \leq (\|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p) \|\boldsymbol{x}+\boldsymbol{y}\|_p^{p/q}$$

- Cancel $\|\boldsymbol{x}+\boldsymbol{y}\|_p^{p/q}$ and note $p - p/q = 1$.

Outline

- Properties of convex functions
- Convexity-preserving operations

Convexity-preserving operations

- nonnegative combinations

$$f(\boldsymbol{x}) = \sum_{i=1}^{m} c_i f_i(\boldsymbol{x})$$

- composition with affine functions

$$f(\boldsymbol{x}) = g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b})$$

- composition of convex/concave functions

$$f(\boldsymbol{x}) = h(g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))$$

- pointwise maximum/supremum

$$f(\boldsymbol{x}) = \sup_{i \in I} f_i(\boldsymbol{x})$$

- partial minimization

$$f(\boldsymbol{x}) = \inf_{\boldsymbol{y} \in C} g(\boldsymbol{x}, \boldsymbol{y})$$

Nonnegative combinations

Proposition. Let $f_i : \mathbb{R}^n \to (-\infty, \infty]$, $i = 1, \ldots, m$, be convex functions. Then for any $c_1, \ldots, c_m \geq 0$, the function $f : \mathbb{R}^n \to (-\infty, \infty]$ defined by

$$f(\boldsymbol{x}) = \sum_{i=1}^{m} c_i f_i(\boldsymbol{x})$$

is convex, with $\operatorname{dom} f = \bigcap_{i=1}^{m} \operatorname{dom} f_i$ [2].

Questions. Let f_1, f_2 be convex functions.

- Is $f_1 - f_2$ convex?
- Is $f_1 \cdot f_2$ convex?
- Is $\dfrac{f_1}{f_2}$ convex?

---

[2] Often we require $\operatorname{dom} f \neq \emptyset$ to preclude the trivial case $f(\boldsymbol{x}) = \infty, \forall \boldsymbol{x}$

Affine composition

Proposition. Let $g : \mathbb{R}^n \to (-\infty, \infty]$ be convex , $\boldsymbol{A} \in \mathbb{R}^{n\times m}$ and $\boldsymbol{b} \in \mathbb{R}^m$.
Then $f : \mathbb{R}^m \to (-\infty, \infty]$ defined by

$$f(\boldsymbol{x}) = g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b})$$

is convex, with $\operatorname{dom} f = \{\boldsymbol{x} \in \mathbb{R}^m : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b} \in \operatorname{dom} g\}$.

Example. $f(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|$ is convex

Example. The following log-sum-exp function is convex

$$f(\boldsymbol{x}) = \log\left(\sum_{i=1}^{n} e^{\boldsymbol{w}_i^T\boldsymbol{x}+b_i}\right).$$

Take $g(\boldsymbol{y}) = \log(\sum_{i=1}^{n} e^{y_i})$ and $\boldsymbol{y} = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)^T\boldsymbol{x} + \boldsymbol{b}$.

Example. $f(x_1, x_2) = (x_1 - 2x_2)^4 + 2e^{3x_1+2x_2-5}$ is convex.
Take $g(y_1, y_2) = y_1^4 + 2e^{y_2}$ and $(y_1, y_2) = (x_1 - 2x_2, 3x_1 + 2x_2 - 5)$.

Scalar composition

Proposition. Consider $g : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R} \to \mathbb{R}$ and $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$.

- f is convex if h is convex and increasing, g is convex
- f is convex if h is convex and decreasing, g is concave
- f is concave if h is concave and increasing, g is concave
- f is concave if h is concave and decreasing, g is convex

Note. When h is increasing, f, h, g have the same convexity property.
When h is decreasing, f, h, $-g$ have the same convexity property.
For $n = 1$ and differentiable g and h, this can be seen from

$$f''(x) = h''(g(x))[g'(x)]^2 + h'(g(x))g''(x)$$

Proof of the first case. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\theta \in [0, 1]$.

1. g convex $\implies g(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) \leq \theta g(\boldsymbol{x}) + \bar{\theta} g(\boldsymbol{y})$
2. h increasing $\implies h \circ g(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) \leq h(\theta g(\boldsymbol{x}) + \bar{\theta} g(\boldsymbol{y}))$
3. h convex $\implies h(\theta g(\boldsymbol{x}) + \bar{\theta} g(\boldsymbol{y})) \leq \theta h(g(\boldsymbol{x})) + \bar{\theta} h(g(\boldsymbol{y}))$
4. 2 and 3 $\implies f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$

Scalar composition (cont'd)

Example. $f(\boldsymbol{x}) = e^{\|\boldsymbol{x}\|}$ is convex for any norm $\|\cdot\|$. Take $g(\boldsymbol{x}) = \|\boldsymbol{x}\|$, $h(x) = e^x$.

Example. $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$ is convex. Take $g(\boldsymbol{x}) = \|\boldsymbol{x}\|$, $h(x) = x^2$ for $x \geq 0$ and $h(x) = 0$ for $x < 0$. (Can we take $h(x) = x^2$ here?)

Example. $f(\boldsymbol{x}) = e^{\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}}$ is convex if $\boldsymbol{Q} \succeq \boldsymbol{O}$. Take $g(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}$, $h(x) = e^x$.

Note. When the domains of g and h are not the entire $\mathbb{R}^n$ or $\mathbb{R}$, similar results hold for their extended-value extensions.

Note. When the conditions fail, we need case-by-case analysis.

Example. $g(x) = x^2$ is convex, $h(x) = e^{-x}$ is convex and decreasing, $f(x) = e^{-x^2}$ is neither convex nor concave.

Example. $g(x) = 1 + e^x$ and $h(x) = -\log x$ are convex, the extension $\tilde{h}$ is decreasing, $f(x) = h(g(x)) = -\log(1 + e^x)$ is concave.

Vector composition

$h : \mathbb{R}^m \to \mathbb{R}$ is increasing (decreasing) if

$$\boldsymbol{x} \geq \boldsymbol{y} \text{ (componentwise)} \implies h(\boldsymbol{x}) \geq (\leq) h(\boldsymbol{y})$$

Proposition. Let $h : \mathbb{R}^m \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, and define $f : \mathbb{R}^n \to \mathbb{R}$ by

$$f(\boldsymbol{x}) = h(g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))$$

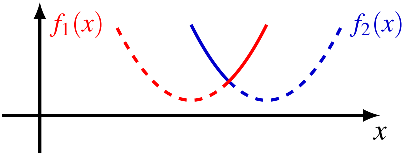
- f is convex if h is convex and increasing, g_i are convex
- f is convex if h is convex and decreasing, g_i are concave
- f is concave if h is concave and increasing, g_i are concave
- f is concave if h is concave and decreasing, g_i are convex

Pointwise maximum

Proposition. If $f_i : \mathbb{R}^n \to (-\infty, \infty]$ be convex functions, $i = 1, 2, \ldots, m$. Then the pointwise maximum $f : \mathbb{R}^n \to (-\infty, \infty]$ defined by

$$f(\boldsymbol{x}) = \max_{1 \le i \le m} f_i(\boldsymbol{x})$$

is convex, with domain $\operatorname{dom} f = \bigcap_{i=1}^m \operatorname{dom} f_i$.



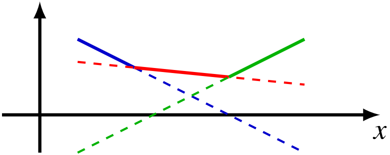
Proof. $\operatorname{dom} f$ is convex. Fix $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$, $\theta \in [0, 1]$. For each i,

$$\begin{aligned} f_i(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) &\le \theta f_i(\boldsymbol{x}) + \bar{\theta} f_i(\boldsymbol{y}) && \text{by convexity of } f_i \\ &\le \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y}) && \text{by definition of } f,\ f_i \le f \end{aligned}$$

Maximizing over i yields $f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) \le \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$.

Pointwise maximum (cont'd)

Example. $f(\boldsymbol{x}) = \max_{1 \le i \le m} (\boldsymbol{w}_i^T \boldsymbol{x} + b_i)$ is convex and piecewise linear.



Special cases.

- Hinge function $(x)^+ = \max\{x, 0\}$
-
$$f(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\} = \max_{1 \le i \le n} \boldsymbol{e}_i^T \boldsymbol{x},$$

  where $\boldsymbol{e}_i$ is i-th standard basis vector.

Question. Is $f(\boldsymbol{x}) = \min_{1 \le i \le m} f_i(\boldsymbol{x})$ convex?

Pointwise supremum

Proposition. Let $f_i : \mathbb{R}^n \to (-\infty, \infty]$, $i \in I$, be convex functions. Then $f : \mathbb{R}^n \to (-\infty, \infty]$ defined by

$$f(\boldsymbol{x}) = \sup_{i \in I} f_i(\boldsymbol{x})$$

is convex.

Proof. Recall the intersection of convex sets is convex.

$$\operatorname{epi} f = \left\{(\boldsymbol{x}, y) : y \geq f(\boldsymbol{x}) = \sup_{i \in I} f_i(\boldsymbol{x})\right\} = \bigcap_{i \in I} \{(\boldsymbol{x}, y) : y \geq f_i(\boldsymbol{x})\} = \bigcap_{i \in I} \operatorname{epi} f_i$$

Example. Given $f : \mathbb{R}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^m$,

$$\phi(\boldsymbol{\lambda}) = \sup_{\boldsymbol{x} \in \mathbb{R}^n} \{f(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x})\}$$

is convex, since $\phi_{\boldsymbol{x}}(\boldsymbol{\lambda}) \triangleq f(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x})$ is affine in $\boldsymbol{\lambda}$ for fixed $\boldsymbol{x} \in \mathbb{R}^n$.

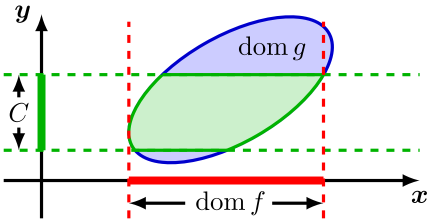
Partial minimization

Proposition. If $g : \mathbb{R}^n \times \mathbb{R}^m \to (-\infty, \infty]$ is convex, and $\emptyset \neq C \subset \mathbb{R}^m$ is convex, then f defined below is convex provided $f(\boldsymbol{x}) > -\infty$ for all $\boldsymbol{x}$,

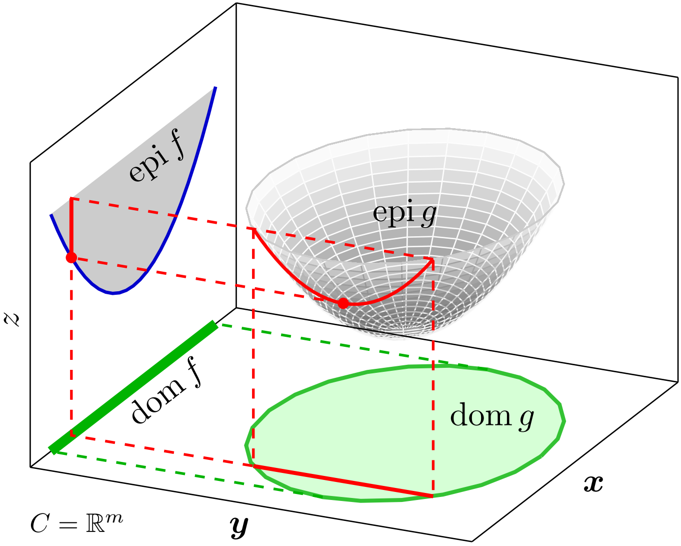
$$f(\boldsymbol{x}) = \inf_{\boldsymbol{y} \in C} g(\boldsymbol{x}, \boldsymbol{y})$$

Note $\operatorname{dom} f$ is the projection of $\operatorname{dom} g \cap (\mathbb{R}^n \times C)$ onto the $\boldsymbol{x}$ coordinates.

$$\begin{aligned}\operatorname{dom} f &= \{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}, \boldsymbol{y}) < \infty \text{ for some } \boldsymbol{y} \in C\} \\ &= \{\boldsymbol{x} \in \mathbb{R}^n : (\boldsymbol{x}, \boldsymbol{y}) \in \operatorname{dom} g \text{ for some } \boldsymbol{y} \in C\}\end{aligned}$$



Partial minimization (cont'd)



Partial Minimization (cont'd)

Proof of proposition. Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \operatorname{dom} f$ and $\theta \in [0, 1]$.

1. By definition of f, for any $\epsilon > 0$, there exists $\boldsymbol{y}_i \in C$ s.t.

$$g(\boldsymbol{x}_i, \boldsymbol{y}_i) < f(\boldsymbol{x}_i) + \epsilon, \quad i = 1, 2$$

2. By convexity of g and C, $\theta \boldsymbol{y}_1 + \bar{\theta} \boldsymbol{y}_2 \in C$, and

$$\begin{aligned}
f(\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2) &= \inf_{\boldsymbol{y} \in C} g(\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2, \boldsymbol{y}) \\
&\leq g(\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2, \theta \boldsymbol{y}_1 + \bar{\theta} \boldsymbol{y}_2) && \theta \boldsymbol{y}_1 + \bar{\theta} \boldsymbol{y}_2 \in C \\
&\leq \theta g(\boldsymbol{x}_1, \boldsymbol{y}_1) + \bar{\theta} g(\boldsymbol{x}_2, \boldsymbol{y}_2) && g \text{ is convex} \\
&< \theta [f(\boldsymbol{x}_1) + \epsilon] + \bar{\theta} [f(\boldsymbol{x}_2) + \epsilon] && \text{by step 1} \\
&= \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2) + \epsilon
\end{aligned}$$

3. Since ϵ is arbitrary,

$$f(\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2) \leq \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2)$$

   so f is convex

Partial minimization (cont'd)

Example. Distance to convex set is convex, $\mathrm{dist}(\boldsymbol{x}, C) = \inf_{\boldsymbol{y} \in C} \|\boldsymbol{x} - \boldsymbol{y}\|$