CS257 Linear and Convex Optimization

Lecture 5

Bo Jiang

John Hopcroft Center for Computer Science
Shanghai Jiao Tong University

October 10, 2020

Recap: Convex Functions

A function $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is convex if

1. its domain $\mathrm{dom} f = S$ is a convex set
2. for any $\boldsymbol{x}, \boldsymbol{y} \in S$ and $\theta \in [0, 1]$,

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

It is strictly convex if $<$ holds for $\boldsymbol{x} \neq \boldsymbol{y} \in S$ and $\theta \in (0, 1)$.

Properties

- f is convex iff its 1D restrictions $g(t) = f(\boldsymbol{x} + t\boldsymbol{d})$ are convex.
- f is convex iff its epigraph

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

  is a convex set in $\mathbb{R}^{n+1}$.
- Jensen's inequality

$$f\left(\sum_{i=1}^{m} \theta_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^{m} \theta_i f(\boldsymbol{x}_i), \quad \boldsymbol{\theta} \in \Delta_{m-1}$$

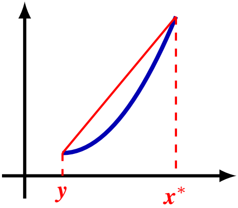
Contents

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 $\boldsymbol{x}_\theta = \theta \boldsymbol{x}^* + \bar{\theta}\boldsymbol{y}$. Then

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



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

$$\|\boldsymbol{x}_\theta - \boldsymbol{x}^*\| = \bar{\theta}\|\boldsymbol{x}^* - \boldsymbol{y}\| \to 0 \quad \text{as} \quad \theta \to 1$$

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.

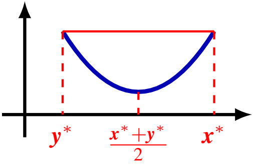
Example. $f(x) = e^x$ has no global or local minimum over $\mathbb{R}$.

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!

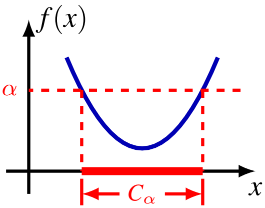
Example. $f(x) = |x|$ has a unique global minimum $x^* = 0$, but it is not strictly convex.

Note. Similar results hold for maxima of concave functions.

Sublevel Sets

The α-sublevel set of a function f is

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



Theorem. Sublevel sets of a convex function are convex.

Proof. Let $\boldsymbol{x}, \boldsymbol{y} \in C_\alpha$, $\theta \in [0, 1]$.

$$f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) \le \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y}) \le \theta\alpha + \bar{\theta}\alpha = \alpha \implies \theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} \in C_\alpha$$

Examples.

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

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

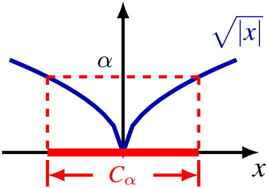
We will see shortly $f(\boldsymbol{x})$ is convex.

Sublevel Sets (cont'd)

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

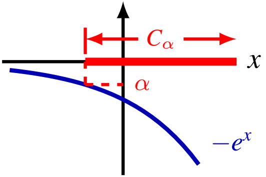
Example. $f(x) = \sqrt{|x|}$ is not convex, but its sublevel sets are all convex,

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



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?

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

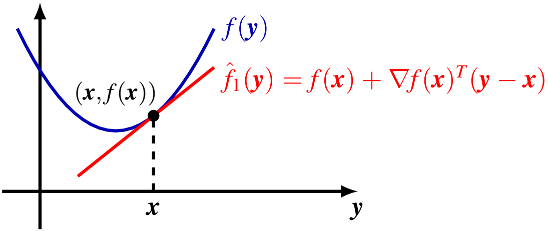
Contents

First-order Condition for Convexity

Theorem. A differentiable f with an open convex domain $\mathrm{dom} f$ is convex iff

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

Note. First-order Taylor approximation underestimates a convex function. Geometrically, all tangent "planes" lie below the graph.



Example. $e^x \geq e^0 + e^0(x-0) = 1 + x$.

First-order Condition for Convexity (cont'd)

Proof. "$\Rightarrow$". Assume f is convex. Let $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{x}$. By Jensen's inequality,

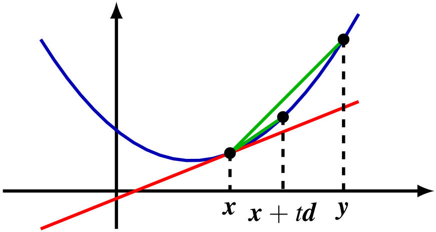
$$f(\boldsymbol{x} + t\boldsymbol{d}) = f(t\boldsymbol{y} + \bar{t}\boldsymbol{x}) \le tf(\boldsymbol{y}) + \bar{t}f(\boldsymbol{x}), \quad t \in (0, 1)$$

Rearranging,

$$\frac{f(\boldsymbol{x} + t\boldsymbol{d}) - f(\boldsymbol{x})}{t} \le f(\boldsymbol{y}) - f(\boldsymbol{x})$$

Letting $t \to 0$,

$$\nabla f(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) = \nabla f(\boldsymbol{x})^T\boldsymbol{d} \le f(\boldsymbol{y}) - f(\boldsymbol{x})$$



Note. $\frac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t\|\boldsymbol{d}\|}$ is the slope of the secant line through $\boldsymbol{x}$ and $\boldsymbol{x} + t\boldsymbol{d}$.

First-order Condition for Convexity (cont'd)

Proof (cont'd). "$\Leftarrow$". Assume the first-order condition holds.

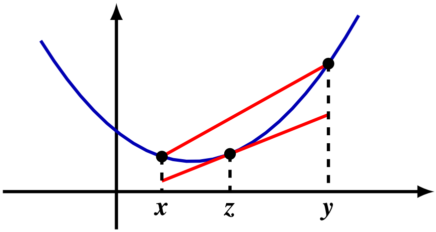
Let $z = \theta x + \bar{\theta} y$. The first-order condition implies

$$f(x) \geq f(z) + \nabla f(z)^T (x - z) \tag{1}$$

$$f(y) \geq f(z) + \nabla f(z)^T (y - z) \tag{2}$$

$\theta \times (1) + \bar{\theta} \times (2)$ yields

$$\theta f(x) + \bar{\theta} f(y) \geq f(z) = f(\theta x + \bar{\theta} y)$$



First-order Condition for Strict Convexity

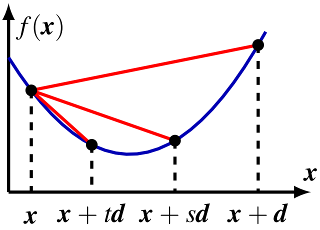
Theorem. A differentiable f with an open convex domain $\operatorname{dom} f$ is strictly convex iff

$$f(\boldsymbol{y}) > f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}), \quad \forall \boldsymbol{x} \neq \boldsymbol{y} \in \operatorname{dom} f$$

Proof. Essentially the same proof with inequalities being strict. The proof of "$\Rightarrow$" requires a further modification. Fix $\boldsymbol{x}$ and $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{x}$.

Add an intermediate point $\boldsymbol{x} + s\boldsymbol{d}$ between $\boldsymbol{x} + t\boldsymbol{d}$ and $\boldsymbol{x} + \boldsymbol{d}$. For $0 < t < s < 1$,

$$\begin{aligned}\frac{f(\boldsymbol{x} + t\boldsymbol{d}) - f(\boldsymbol{x})}{t} &< \frac{f(\boldsymbol{x} + s\boldsymbol{d}) - f(\boldsymbol{x})}{s} \\ &< f(\boldsymbol{x} + \boldsymbol{d}) - f(\boldsymbol{x}),\end{aligned}$$



Now letting $t \to 0$ yield

$$\nabla f(\boldsymbol{x})^T \boldsymbol{d} \leq \frac{f(\boldsymbol{x} + s\boldsymbol{d}) - f(\boldsymbol{x})}{s} < f(\boldsymbol{x} + \boldsymbol{d}) - f(\boldsymbol{x}) = f(\boldsymbol{y}) - f(\boldsymbol{x})$$

Optimality of Stationary Points

Corollary. If $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ for a convex function f, then $\boldsymbol{x}^*$ is a global minimum. If f is strictly convex, then $\boldsymbol{x}^*$ is the unique global minimum.

Proof. By the first-order condition and the assumption $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$,

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*) + \nabla f(\boldsymbol{x}^*)^T (\boldsymbol{x} - \boldsymbol{x}^*) = f(\boldsymbol{x}^*), \quad \forall \boldsymbol{x} \in \operatorname{dom} f$$

so $\boldsymbol{x}^*$ is a global minimum.

Similarly, if f is strictly convex,

$$f(\boldsymbol{x}) > f(\boldsymbol{x}^*) + \nabla f(\boldsymbol{x}^*)^T (\boldsymbol{x} - \boldsymbol{x}^*) = f(\boldsymbol{x}^*), \quad \forall \boldsymbol{x}^* \neq \boldsymbol{x} \in \operatorname{dom} f$$

so $\boldsymbol{x}^*$ is the unique global minimum.

Note. For concave functions, similar results hold with all inequalities reversed, and min replaced by max.

Contents

Second-order Condition for Convexity

Theorem. A twice continuously differentiable f with an open convex domain $\mathrm{dom} f$ is convex iff $\nabla^2 f(\boldsymbol{x}) \succeq \boldsymbol{O}$ is positive semidefinite at every $\boldsymbol{x} \in \mathrm{dom} f$.

Proof. "$\Rightarrow$". Assume f is convex. Fix $\boldsymbol{x} \in \mathrm{dom} f$ and $\boldsymbol{d} \in \mathbb{R}^n$.

- $\mathrm{dom} f$ is open $\implies \boldsymbol{x} + t\boldsymbol{d} \in \mathrm{dom} f$ for small t
- By the second-order Taylor expansion with Peano remainder,

$$f(\boldsymbol{x} + t\boldsymbol{d}) = f(\boldsymbol{x}) + t\nabla f(\boldsymbol{x})^T\boldsymbol{d} + \frac{1}{2}t^2\boldsymbol{d}^T\nabla^2 f(\boldsymbol{x})\boldsymbol{d} + o(t^2).$$

- By the first-order condition for convexity,

$$f(\boldsymbol{x} + t\boldsymbol{d}) \geq f(\boldsymbol{x}) + t\nabla f(\boldsymbol{x})^T\boldsymbol{d} \implies \frac{1}{2}\boldsymbol{d}^T\nabla^2 f(\boldsymbol{x})\boldsymbol{d} + o(1) \geq 0$$

- Setting $t \to 0 \implies \boldsymbol{d}^T\nabla^2 f(\boldsymbol{x})\boldsymbol{d} \geq 0 \implies \nabla^2 f(\boldsymbol{x}) \succeq \boldsymbol{O}$

Second-order Condition for Convexity (cont'd)

Proof (cont'd). "$\Leftarrow$". Assume the second-order condition holds. Fix $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$ and let $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{x}$.

- By the second-order Taylor expansion with Lagrange remainder,

$$f(\boldsymbol{y}) = f(\boldsymbol{x} + \boldsymbol{d}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \frac{1}{2} \boldsymbol{d}^T \nabla^2 f(\boldsymbol{x} + s\boldsymbol{d}) \boldsymbol{d}, \quad s \in (0,1)$$

- $\operatorname{dom} f$ is convex $\implies \boldsymbol{x} + s\boldsymbol{d} = s\boldsymbol{y} + \bar{s}\boldsymbol{x} \in \operatorname{dom} f$
- By the second-order condition,

$$\nabla^2 f(\boldsymbol{x} + s\boldsymbol{d}) \succeq \boldsymbol{O} \implies f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d}$$

which is the first-order condition for convexity, so f is convex.

Second-order Condition for Convexity (cont'd)

Theorem. A twice continuously differentiable f with an open convex domain $\mathrm{dom}\, f$ is strictly convex if $\nabla^2 f(\boldsymbol{x})$ is positive definite at every $\boldsymbol{x} \in \mathrm{dom}\, f$.

Proof. Replace $\succeq$ and $\geq$ by $\succ$ and $>$ respectively in "$\Leftarrow$" part.

Note. Positive definiteness is sufficient but not necessary.

Example. $f(x) = x^4$ is strictly convex, but $f''(x) = 0$ at $x = 0$

Example. $f(\boldsymbol{x}) = f(x_1, x_2) = x_1^2 + x_2^4$ is strictly convex, but $\nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{pmatrix}$ is not positive definite for $x_2 = 0$.

Note. For concave functions, replace "positive (semi)definite" by "negative (semi)definite" in previous theorems.

Examples

Example. Exponential $f(x) = e^{ax}$ is convex for $a \in \mathbb{R}$

Proof. $f''(x) = a^2 e^{ax} \geq 0$

Example. Logarithm $f(x) = \log x$ is strictly concave over $(0, \infty)$

Proof. $f''(x) = -x^{-2} < 0$

Example. Power $f(x) = x^a$ is convex over $(0, \infty)$ for $a \geq 1$ or $a \leq 0$, and concave over $(0, \infty)$ for $0 \leq a \leq 1$

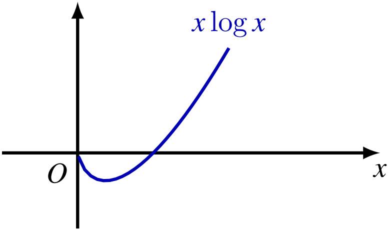
Proof. $f''(x) = a(a-1)x^{a-2} \gtreqless 0$ depending on a

Note. Domain is important. $f(x) = x^{-1}$ is concave over $(-\infty, 0)$, but neither convex nor concave over $(-1, 0) \cup (0, 1)$.

Example: Negative Entropy

The negative entropy $f(x) = x \log x$ is strictly convex over $(0, \infty)$.

Proof. $f'(x) = \log x + 1, f''(x) = x^{-1} > 0$



Note. We typically extend the definition of f to $x = 0$ by continuity, i.e.

$$f(0) \triangleq \lim_{x \to 0_+} f(x) = 0$$

f is still strictly convex with this extension.

Example: Quadratic Functions

A quadratic function

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} + \boldsymbol{c}$$

with symmetric $\boldsymbol{Q}$ is convex iff $\boldsymbol{Q} \succeq \boldsymbol{O}$, and strictly convex iff $\boldsymbol{Q} \succ \boldsymbol{O}$.

Proof. For convexity, $\nabla^2 f(\boldsymbol{x}) = \boldsymbol{Q}$ and use second-order condition.

For strict convexity, note $\nabla f(\boldsymbol{x}) = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}$ and

$$f(\boldsymbol{x}+\boldsymbol{d}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T\boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T\boldsymbol{Q}\boldsymbol{d}$$

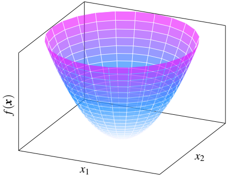
By first-order condition, f is strictly convex iff for $\boldsymbol{d} \neq \boldsymbol{0}$

$$f(\boldsymbol{x}+\boldsymbol{d}) > f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T\boldsymbol{d} \iff \boldsymbol{d}^T\boldsymbol{Q}\boldsymbol{d} \iff \boldsymbol{Q} \succ \boldsymbol{O}$$

Note. Recall in general $\nabla^2 f(\boldsymbol{x}) \succ \boldsymbol{O}$ is not a necessary condition for strict convexity, but it is necessary when f is quadratic.

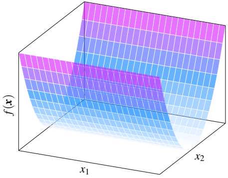
Example: Quadratic Functions (cont’d)

Quadratic function $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$ in $\mathbb{R}^2$



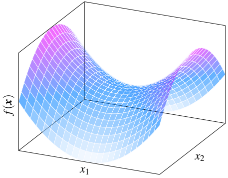
$\boldsymbol{Q} = \text{diag}\{1, 1\}$

strictly convex



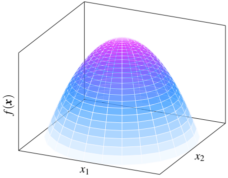
$\boldsymbol{Q} = \text{diag}\{0, 1\}$

convex



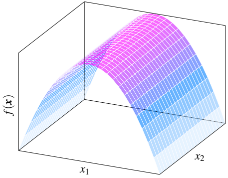
$\boldsymbol{Q} = \text{diag}\{1, -1\}$

neither convex nor concave



$\boldsymbol{Q} = \text{diag}\{-1, -1\}$

strictly concave



$\boldsymbol{Q} = \text{diag}\{-1, 0\}$

concave

Example: Least Squares Loss

The least squares loss

$$f(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2$$

is always convex.

Proof. f is a quadratic function,

$$f(\boldsymbol{x}) = (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y})^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x} + \boldsymbol{y}^T\boldsymbol{y}.$$

$\boldsymbol{A}^T\boldsymbol{A} \succeq \boldsymbol{O}$ since

$$\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{A}\boldsymbol{x})^T(\boldsymbol{A}\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x}\|_2^2 \geq 0$$

Question. When is it strictly convex?

Answer. When $\boldsymbol{A}^T\boldsymbol{A} \succ \boldsymbol{O}$, which is true iff $\boldsymbol{A}$ has full column rank.

$$\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = 0 \iff \|\boldsymbol{A}\boldsymbol{x}\|_2 = 0 \iff \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$$

Example: Log-sum-exp Function

Log-sum-exp function defined below is convex

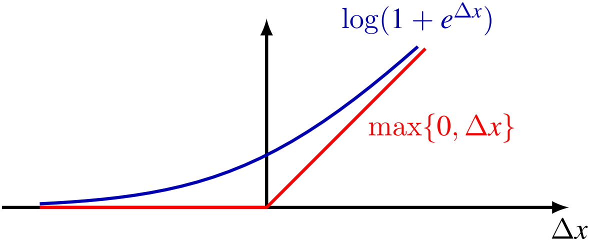
$$f(\boldsymbol{x}) = \log\left(\sum_{i=1}^{n} e^{x_i}\right)$$

Note. Also called "soft max", as it smoothly approximates $\max\limits_{1\le i\le n} x_i$.

For $n = 2$, $\Delta x = x_2 - x_1$,

$$f(x_1, x_2) = \log(e^{x_1} + e^{x_2}) = \log[e^{x_1}(1 + e^{\Delta x})] = x_1 + \log(1 + e^{\Delta})$$

$$\approx x_1 + \max\{0, \Delta x\} = \max\{x_1, x_2\}$$



Example: Log-sum-exp Function (cont'd)

Proof. Let $s(\boldsymbol{x}) = \sum_{k=1}^{n} e^{x_k}$, so $f(\boldsymbol{x}) = \log s(\boldsymbol{x})$.

$$g_i \triangleq \frac{\partial f(\boldsymbol{x})}{\partial x_i} = \frac{1}{s}\frac{\partial s}{\partial x_i} = \frac{e^{x_i}}{s}$$

$$H_{ij} \triangleq \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} = \frac{\partial g_i}{\partial x_j} = \frac{e^{x_i}}{s}\delta_{ij} - \frac{e^{x_i}e^{x_j}}{s^2} = g_i\delta_{ij} - g_i g_j, \quad \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

For any $\boldsymbol{d} \in \mathbb{R}^n$,

$$\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}) \boldsymbol{d} = \sum_{i=1}^{n} \sum_{j=1}^{n} d_i d_j H_{ij} = \sum_{i=1}^{n} g_i d_i^2 - \left(\sum_{i=1}^{n} g_i d_i \right)^2 \geq 0$$

by the convexity of $x \mapsto x^2$ (why?), so $\nabla^2 f(\boldsymbol{x}) \succeq \boldsymbol{O}$.

Contents

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

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

Extended-value Extension

Given convex $f : S \subset \mathbb{R}^n \to \mathbb{R}$, define its extended-value extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ by

$$\tilde{f}(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}), & \boldsymbol{x} \in S \\ \infty, & \boldsymbol{x} \notin S \end{cases}$$

with extended arithmetic and ordering

$$a + \infty = \infty \text{ for } a > -\infty; \quad a \cdot \infty = \infty \text{ for } a > 0; \quad 0 \cdot \infty = 0$$

- The (effective) domain of $\tilde{f}$, also the domain of f, is

$$\operatorname{dom} \tilde{f} = \operatorname{dom} f = S = \{\boldsymbol{x} : \tilde{f}(\boldsymbol{x}) < \infty\}$$

- f is convex iff $\tilde{f}$ is convex, i.e. for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\theta \in [0, 1]$,

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

- f and $\tilde{f}$ have the same epigraph; we will identify f with $\tilde{f}$

Note. We can similarly extend a concave function by $f(\boldsymbol{x}) = -\infty$ for $\boldsymbol{x} \notin \operatorname{dom} f$.