Optimization methods. Seminar 6. Convex functions

Alexandr Katrutsa

Moscow Institute of Physics and Technology
Department of Control and Applied Mathematics

October 17, 2017

Reminder

- Derivative by scalar
- Derivative by vector
- Derivative by matrix
- Chain rule

Functions definitions

Convex function

A function $f : X \subset \mathbb{R}^n \to \mathbb{R}$ is called conves (strictly convex), if X is a convex set and $\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ and $\alpha \in [0, 1]$ ($\alpha \in (0, 1)$):

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \; (<) \; \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$$

Concave function

A function f is concave (strictly concave), if $-f$ is convex (strictly convex).

Strongly convex function

A function $f : X \subset \mathbb{R}^n \to \mathbb{R}$ is called strongly convex with constant $m > 0$, if X is a convex set and $\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ и $\alpha \in [0, 1]$:

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) - \frac{m}{2}\alpha(1-\alpha)\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

Sets definitions

Epigraph

An epigraph of a function f is called a set
$\mathrm{epi} f = \{(\mathbf{x}, y) : \mathbf{x} \in \mathbb{R}^n,\ y \in \mathbb{R},\ y \geq f(\mathbf{x})\} \subset \mathbb{R}^{n+1}$

Sublevel set

A sublevel set of a function f is called a set

$$C_\gamma = \{\mathbf{x} | f(\mathbf{x}) \leq \gamma\}.$$

Quasi-convex function

A function f is called quasi-convex, if its domain is convex set and sublevel set for any γ is convex.

Convex function criteria

First order criterion

A function f is convex $\Leftrightarrow$the function is defined on the convex set X and $\forall \mathbf{x}, \mathbf{y} \in X \subset \mathbb{R}^n$:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + (\nabla f(\mathbf{x}))^\mathsf{T} (\mathbf{y} - \mathbf{x})$$

Second order criterion

A continuous and twice differentiable function f is convex $\Leftrightarrow$ the function is defined on the convex set X and $\forall \mathbf{x}, \mathbf{y} \in \mathbf{relint}(X) \subset \mathbb{R}^n$:

$$\nabla^2 f(\mathbf{x}) \succeq 0.$$

Relation to the epigraph property

A function is convex $\Leftrightarrow$ its epigraph is convex set.

Restriction to the line

A function $f : X \to \mathbb{R}$ is convex iff X is a convex set and the univariate function $g(t) = f(\mathbf{x} + t\mathbf{v})$ defined on the set $\{t | \mathbf{x} + t\mathbf{v} \in X, \ \forall \mathbf{x}, \mathbf{v}\}$ is convex.

Strongly convexity criteria

First order criterion

A function f is strongly convex with constant $m \Leftrightarrow$ the function is defined on the convex set X and $\forall \mathbf{x}, \mathbf{y} \in X \subset \mathbb{R}^n$:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + (\nabla f(\mathbf{x}))^\mathsf{T} (\mathbf{y} - \mathbf{x}) + \frac{m}{2}\|\mathbf{y} - \mathbf{x}\|^2$$

Second order criterion

A continuous and twice differentiable function f is strongly convex with constant $m \Leftrightarrow$ the function is defined on the convex set X and $\forall \mathbf{x} \in \mathbf{relint}(X) \subset \mathbb{R}^n$:

$$\nabla^2 f(\mathbf{x}) \succeq m\mathbf{I}.$$

Examples

1. Quadratic function: $f(x) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Proper norms in $\mathbb{R}^n$
3. $f(\mathbf{x}) = \log\left(e^{x_1} + \ldots + e^{x_n}\right)$, $\mathbf{x} \in \mathbb{R}^n$ — smooth approximation of maximum
4. Log determinant: $f(\mathbf{X}) = -\log\det \mathbf{X}$, $\mathbf{X} \in \mathbf{S}^n_{++}$
5. A set of the convex functions is a convex cone
6. Element-wise maximum of convex functions: $f(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}$, $\mathrm{dom}\ f = \mathrm{dom}\ f_1 \cap \mathrm{dom}\ f_2$
7. Extension to the infinite set of functions: if for any $\mathbf{y} \in \mathcal{A}$ a function $f(\mathbf{x}, \mathbf{y})$ is convex on $\mathbf{x}$, then $\sup\limits_{\mathbf{y} \in \mathcal{A}} f(\mathbf{x}, \mathbf{y})$ is convex on $\mathbf{x}$
8. Leading eigenvalue: $f(\mathbf{X}) = \lambda_{\max}(\mathbf{X})$

Jensen inequality

Jensen inequality

For any convex function f we have the following inequality:

$$f\left(\sum_{i=1}^n \alpha_i \mathbf{x}_i\right) \leq \sum_{i=1}^n \alpha_i f(\mathbf{x}_i),$$

where $\alpha_i \geq 0$ and $\sum\limits_{i=1}^n \alpha_i = 1$.

Or in the case of infinitely many $\mathbf{x}_i$: $p(x) \geq 0$ и $\int\limits_X p(x) = 1$

$$f\left(\int\limits_X p(x) x dx\right) \leq \int\limits_X f(x)p(x)dx$$

if integrals exist.

Examples

1. Hölder inequality
2. Arithmetic mean vs. geometric mean
3. $f(\mathbf{E}(x)) \leq \mathbf{E}(f(x))$
4. Convexity of the hyperbolic set $\{\mathbf{x} | \prod\limits_{i=1}^{n} x_i \geq 1\}$

Recap

- Convex function
- Epigraph and sublevel set of function
- Convex function criteria
- Jensen inequality