Optimization methods.
Seminar 2. Convex functions and matrix calculus

Alexandr Katrutsa

Moscow Institute of Physics and Technology

November 20, 2021

Definitions

Convex function
A function $f : X \subset \mathbb{R}^n \to \mathbb{R}$ is called convex (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
$$\text{epi} f = \{(\mathbf{x}, y) : \mathbf{x} \in \mathbb{R}^n, \ y \in \mathbb{R}, \ y \geq f(\mathbf{x})\} \subset \mathbb{R}^{n+1}$$

Convex function criteria

First order criterion
A function f is convex $\Leftrightarrow$ the function is defined on the convex set X and
$\forall \mathbf{x}, \mathbf{y} \in X \subset \mathbb{R}^n$: $\qquad f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle$

Second order criterion
A continuous and twice differentiable function f is convex $\Leftrightarrow$ the function is defined on the convex set X:

$$f''(\mathbf{x}) \succeq 0.$$

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 $\forall \mathbf{x}, \mathbf{y} \in X \subset \mathbb{R}^n$:
$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|^2$$

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:
$$f''(\mathbf{x}) \succeq m\mathbf{I}.$$

Examples

1. Quadratic function: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$

Examples

1. Quadratic function: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Norms in $\mathbb{R}^n$

Examples

1. Quadratic function: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$

2. 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

Examples

1. Quadratic function: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$

2. 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. Maximum eigenvalue: $f(\mathbf{X}) = \lambda_{\max}(\mathbf{X})$

Matrix calculus: reminder

Consider function $f : D \to E$
- Gradient

Matrix calculus: reminder

Consider function $f : D \to E$

- Gradient
- Hessian in the case of $D \equiv \mathbb{R}^n$ and $E \equiv \mathbb{R}$

Matrix calculus: reminder

Consider function $f : D \rightarrow E$

- Gradient
- Hessian in the case of $D \equiv \mathbb{R}^n$ and $E \equiv \mathbb{R}$
- Jacobi matrix

Examples

1. Linear function: $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$

Examples

1. Linear function: $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$
2. Quadratic form: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}$

Examples

1. Linear function: $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$
2. Quadratic form: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}$
3. Squared ℓ_2 norm of residual vector: $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$

Examples

1. Linear function: $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$
2. Quadratic form: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}$
3. Squared ℓ_2 norm of residual vector: $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$
4. Trace: $f(\mathbf{X}) = \mathrm{trace}(\mathbf{A}\mathbf{X}\mathbf{B})$

Examples

1. Linear function: $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$
2. Quadratic form: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}$
3. Squared ℓ_2 norm of residual vector: $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$
4. Trace: $f(\mathbf{X}) = \mathrm{trace}(\mathbf{A}\mathbf{X}\mathbf{B})$
5. $f(\mathbf{x}) = (\mathbf{x} - \mathbf{A}\mathbf{s})^\top \mathbf{W}(\mathbf{x} - \mathbf{A}\mathbf{s})$