Distributed Optimization for Machine Learning

School of Electrical and Computer Engineering
University of Tehran
Erfan Darzi

Lecture 2 – Convex Optimization

erfandarzi@ut.ac.ir

Recap

- Global and local optimality for unconstrained optimization
- Necessary and sufficient optimality conditions
- Difference between convex and non-convex optimization
- Convexity + optimality condition ➔ Global optimality

Convex Optimization Problem?

$$\min_{\mathbf{x}} \quad f(\mathbf{x})$$
$$\text{s.t.} \quad \mathbf{x} \in \mathcal{X}$$

Convex

function

Convex Set

- Is a given function convex?

- Is a given set convex?

How to identify a convex function?

- Definition: $f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}), \quad \forall\, \alpha \in [0,1], \ \ x, y$

(Very useful definition: Jensen's inequality)

- Hessian for twice differentiable functions:

Examples:

$$f \ \text{ is convex} \Leftrightarrow \nabla^2 f(\mathbf{x}) \succeq \mathbf{0}, \quad \forall \mathbf{x}$$

$f(x) = x^2$

$f(x) = -\log(x)$, for $x > 0$

$f(x) = 1/x, \ \text{ for } x \geq 1$

- Affine functions:

$$f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$$

- Quadratic functions:

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{x}^T\mathbf{b} + c \ \text{ convex} \Leftrightarrow \mathbf{Q} \succeq \mathbf{0}$$

Convex Functions

- Norms: **Examples:**

  $f:\mathbb{R}^n \mapsto \mathbb{R}$ is a norm if

$$f(\mathbf{x}) \geq 0; f(\mathbf{x}) = 0 \text{ iff } \mathbf{x} = 0$$
$$f(t\mathbf{x}) = |t| f(\mathbf{x})$$
$$f(\mathbf{x}+\mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$$

$\|\mathbf{x}\|_q,\ q \geq 1$

Special cases $1, 2, \infty$

$\|\cdot\|_0$ is not a norm

why ℓ_1 instead of ℓ_0

- A function is convex iff it is convex on every line:

$$f \text{ convex} \Leftrightarrow f(\mathbf{x}_0 + t\mathbf{h}) \text{ convex in } t \text{ for all } \mathbf{x}_0, \mathbf{h}$$

- Positive multiple of convex function f convex, $\alpha > 0 \Rightarrow \alpha f$ convex

- Sum of convex functions: f_1, f_2 convex $\Rightarrow f_1 + f_2$ convex

Extension to integral

Convex Functions

- Pointwise supremum of convex functions

$$f_\alpha \text{ convex}, \ \forall \alpha \in \mathcal{A} \Rightarrow \sup_{\alpha \in \mathcal{A}} f_\alpha \text{ convex}$$

Example

No assumption on $\mathcal{A}$

- Affine transformation of domain

$$f \text{ convex} \Rightarrow f(\mathbf{Ax} + \mathbf{b}) \text{ convex}$$

- Composition: $f(\mathbf{x}) = h(g(\mathbf{x})), \ \text{with } g : \mathbb{R}^n \mapsto \mathbb{R}, \ h : \mathbb{R} \mapsto \mathbb{R}; f \text{ is convex if}$

$$g \text{ convex}; h \text{ convex and nondecreasing}$$

Necessary?

$$g \text{ concave}; h \text{ convex and nonincreasing}$$

•

Examples of Convex Functions on Matrices

- Affine function: $f(\mathbf{X}) = \mathrm{Tr}(\mathbf{AX}+\mathbf{B})$

- Largest eigenvalue on symmetric matrices: $f(\mathbf{X}) = \lambda_{\max}(\mathbf{X}),\ \forall\, \mathbf{X} \succeq \mathbf{0}$

- Largest singular value: $f(\mathbf{X}) = \sigma_1(\mathbf{X}) = \|\mathbf{X}\|_2$

  $f(\mathbf{X}) = \sigma_1(\mathbf{X}) + \sigma_2(\mathbf{X})$

- Log-det-inv on PSD matrices $f(\mathbf{X}) = \log(\det(\mathbf{X}^{-1})),\ \forall\, \mathbf{X} \succeq \mathbf{0}$

$$\begin{aligned} g(t) &\triangleq \log\det(\mathbf{X}_0 + t\mathbf{H})^{-1} \\ &= \log\det \mathbf{X}_0^{-1} + \log\det(\mathbf{I} + t\mathbf{X}_0^{-1/2}\mathbf{H}\mathbf{X}_0^{-1/2})^{-1} \\ &= \log\det \mathbf{X}_0^{-1} - \sum_i \log(1 + t\lambda_i) \end{aligned}$$

$\lambda_i' s$ eigenvalues of $\mathbf{X}_0^{-1/2}\mathbf{H}\mathbf{X}_0^{-1/2}$

From Convex Functions to Convex Sets

- A sublevel set of a convex function is convex set

$$f \text{ convex function} \Rightarrow \mathcal{C}(\alpha) = \{\mathbf{x} \in \text{ dom } f | f(\mathbf{x}) \leq \alpha\} \text{ convex set}$$

Is the converse true?

Examples of Convex Sets

- Affine sets: $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}\}$
- Halfspace:
- Intersection of convex sets

True for union?

$$S_\alpha \text{ is convex for every } \alpha \Rightarrow \bigcap_{\alpha \in \mathcal{A}} S_\alpha \text{ is convex}$$

Example:
Polyhedron: intersection of finite number of halfspaces

Change it to infinite?

- Set of PSD matrices $\{\mathbf{X} \in \mathbb{R}^{n \times n} | \mathbf{X} \succeq \mathbf{0}\}$
- Norm balls

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T\mathbf{x} \leq b\}\{\mathbf{x} \mid \|\mathbf{x} - \mathbf{c}\| \leq \alpha\}$$

Convex Sets

- Affine transformation of convex sets:

$$\mathcal{C} \text{ convex} \Rightarrow \{\mathbf{x} \mid \mathbf{A}\mathbf{x} + \mathbf{b} \in \mathcal{C}\} \text{ convex}$$

$$\mathcal{C} \text{ convex} \Rightarrow \{\mathbf{A}\mathbf{x} + \mathbf{b} \mid \mathbf{x} \in \mathcal{C}\} \text{ convex}$$

- Linear matrix inequalities

$$\mathbf{A}_0, \mathbf{A}_1, \ldots, \mathbf{A}_n \text{ symmetric} \Rightarrow \{\mathbf{x} \mid \mathbf{A}_0 + x_1\mathbf{A}_1 + \ldots + x_n\mathbf{A}_n \succeq \mathbf{0}\} \text{ convex}$$

- A popular way of representing some convex optimization problems:

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_1(\mathbf{x}) \leq 0 \\ & g_2(\mathbf{x}) \leq 0 \\ & \vdots \\ & g_m(\mathbf{x}) \leq 0 \end{aligned}$$