

計画数理学特論

～第3回：凸解析と最適化・劣勾配～

担当：蓮池隆（経営システム工学科）
e-mail: thasuike@waseda.jp

本日の講義について

- 非線形最適化の中でも重要な凸性について解説
(前回) : 凸性の基礎(凸集合, 凸包, 凸関数)

Q : なぜ凸性, 凸解析が重要なのか?

A : **凸性があれば『局所最適解 = 大域最適解』となる**

→ 効率的な最適化アルゴリズムが設計可能

(状況によっては, 微分=0だけでも大域最適解が見つかる)

→ **今回はこの最適化部分にフォーカスをして解説**

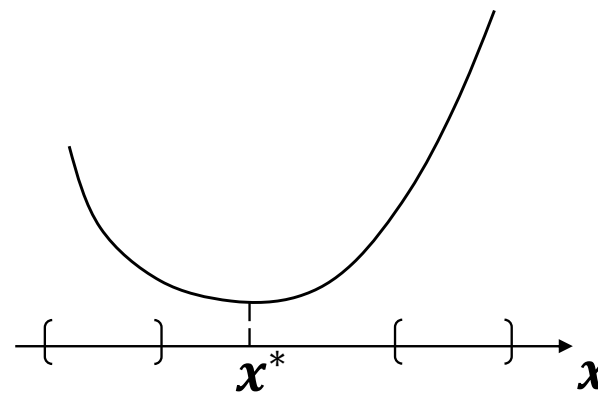
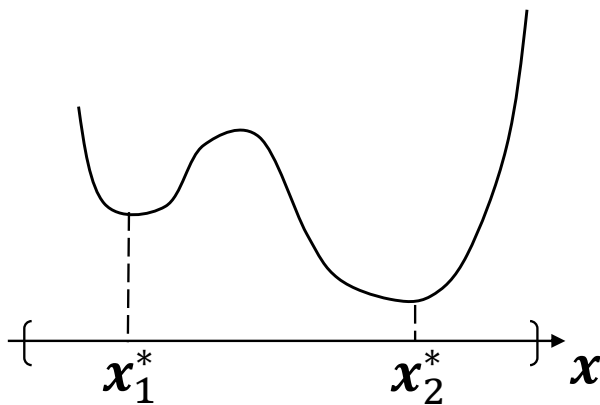
凸関数の最小化

- $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ を凸関数, $S \subset \mathbb{R}^n$ を凸集合とする.
- 最適化問題

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to } x \in S$$

の局所最適解を $x^* \in \text{dom}(f) \cap S$ とする.

このとき, x^* は大域最適解となる.



凸関数の最小化(証明)

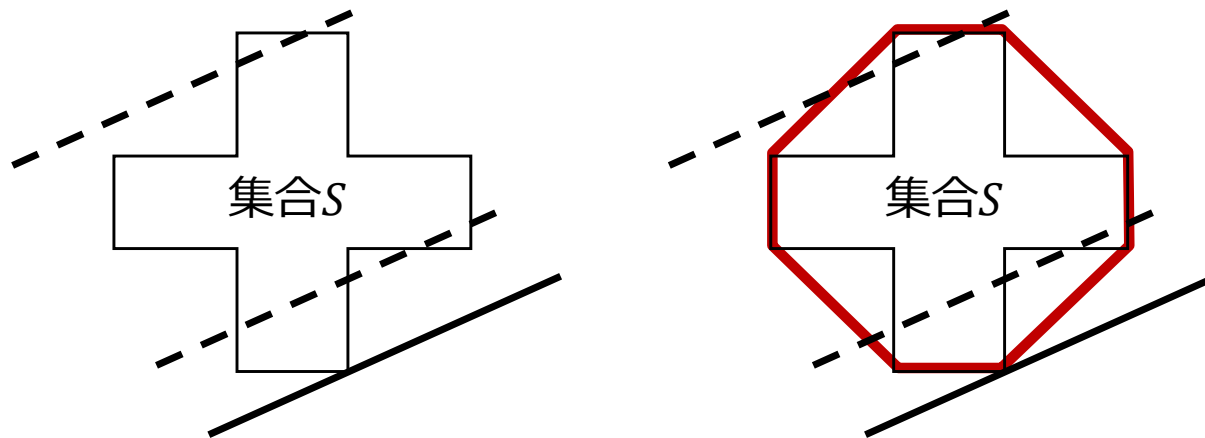
- $T = \text{dom}(f) \cap S$ は, $\text{dom}(f), S$ ともに凸集合であるため, T も凸集合
- $x \in T$ と $\lambda \in (0, 1)$ に対して, $\lambda x^* + (1 - \lambda)x$ も実行可能解
- $f(x) < f(x^*)$ が成り立つなら, f が凸関数であることから,
$$f(\lambda x^* + (1 - \lambda)x) \leq \lambda f(x^*) + (1 - \lambda)f(x) < f(x^*)$$
- ここで, λ を1に近い値にとると, x^* の近傍における $\lambda x^* + (1 - \lambda)x$ での関数の値が, x^* の時の関数の値よりも小さい
→ x^* が局所最適解にならないので, 矛盾
∴ 任意の $x \in T$ に対して, $f(x) \geq f(x^*)$ が成り立つ.

目的関数が1次式の場合

- $S \subset \mathbb{R}^n$ を任意の集合(凸集合とは限らない)とする.
- $\mathbf{c} \in \mathbb{R}^n, a \in \mathbb{R}$ とすると,
$$\inf\{\mathbf{c}^t \mathbf{x} + a \mid \mathbf{x} \in S\} = \inf\{\mathbf{c}^t \mathbf{x} + a \mid \mathbf{x} \in \text{conv}(S)\}$$

が成り立つ.

(イメージ)



関数値の上界・下界

- 関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ が γ -平滑な関数

(参考)

- γ -平滑: 勾配 ∇f がリプシッツ連続性 $\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\| \leq \gamma \|\mathbf{x} - \mathbf{y}\|$ ($\gamma > 0$)
(イメージ: 関数 f の2回微分が有限)

- 関数 f に対して, 次の不等式が成立する

$$f(\mathbf{y}) \leq f(\mathbf{x}) + \nabla f(\mathbf{x})^t (\mathbf{y} - \mathbf{x}) + \frac{\gamma}{2} \|\mathbf{y} - \mathbf{x}\|^2$$

- さらに関数 f が凸関数なら,

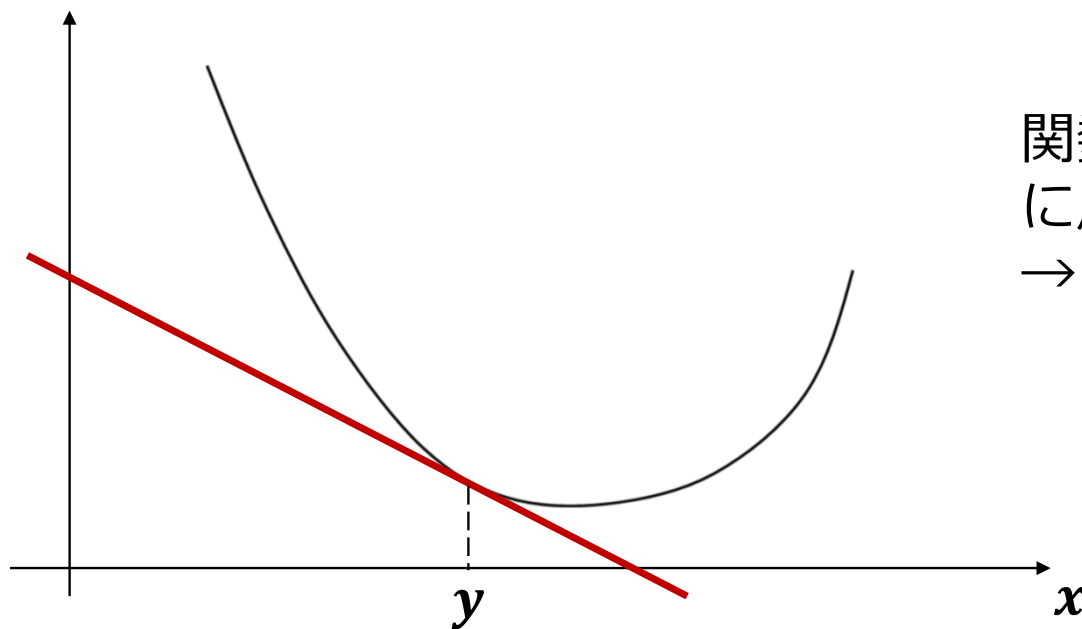
$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^t (\mathbf{y} - \mathbf{x}) + \frac{1}{2\gamma} \|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\|^2$$

共役関数

- 関数 $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ で微分可能
- エピグラフ $\text{epi}(f) \subset \mathbb{R}^{n+1}$
$$\text{epi}(f) = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} \mid f(\mathbf{x}) \leq t\}$$
- 関数 f が凸関数であることと, $\text{epi}(f)$ が凸集合であることは同値
 - 特に $\text{epi}(f)$ が空集合でなければ, 関数 f は真凸関数
 - $\text{epi}(f)$ が閉集合であれば, 関数 f は閉凸関数
- 真凸関数 $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ に対して, 共役関数 $f^*: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ を定義

共役関数のイメージ・関係性

- 真凸関数 $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ に対して, 共役関数 $f^*: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ を定義



関数の変数とその微分に変えるために用いられる変換(**Legendre変換**)
→ 元の関数を接線の傾きと切片で指定される接線の集合で表現

共役関数のイメージ・関係性

- 真凸関数 $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ に対して, 共役関数 $f^*: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ を定義
- 閉凸かつ真凸な関数 (= 閉真凸関数) $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ の共役関数 f^* も閉真凸関数
- 元々の関数と共役関数の間には, 最適化問題の観点で, 主問題と双対問題の関係性がある

劣勾配

- 関数 f : 微分不可能な点を持つ凸関数
- 点 $x \in \text{dom}(f)$ とし, ベクトル $g \in \mathbb{R}^n$ が任意の点 $y \in \mathbb{R}^n$ に対して,

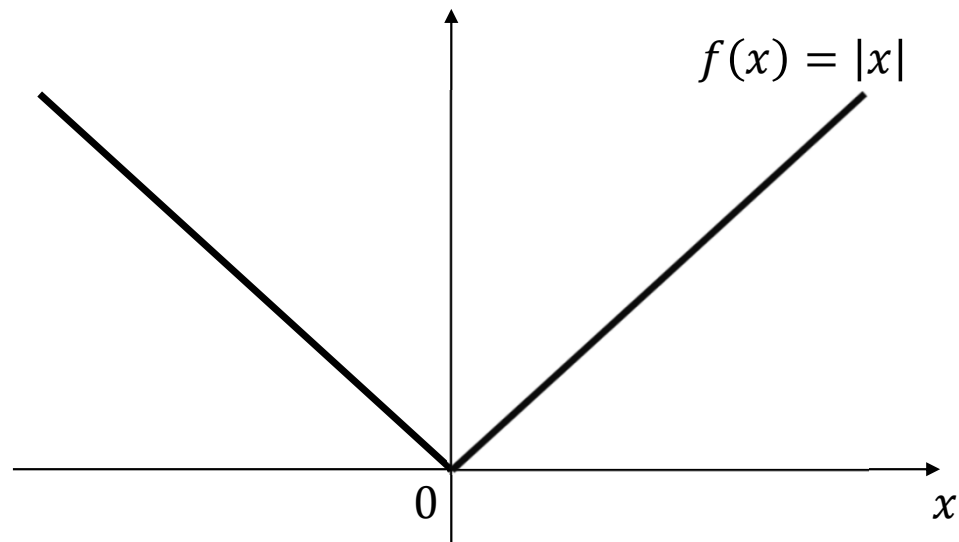
を満たすとき, g を f の x における**劣勾配**(subgradient)と呼ぶ.
- 点 x におけるすべての劣勾配を集めた集合を $\partial f(x)$ と表し, **劣微分**(subdifferential)と呼ぶ.

劣微分の性質

- $\mathbf{0} \in \partial f(\mathbf{x})$ のとき, 任意の $\mathbf{y} \in \mathbb{R}^n$ に対して, $f(\mathbf{y}) \geq f(\mathbf{x})$
- 実数 $\lambda > 0$ に対して, $\partial(\lambda \cdot f)(\mathbf{x}) = \lambda \partial f(\mathbf{x})$
- 関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ は微分可能な凸関数ならば, $\partial f(\mathbf{x}) = \{\nabla f(\mathbf{x})\}$
- 関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ が微分可能な凸関数 g_1, g_2, \dots, g_p を用いて,
$$f(\mathbf{x}) = \max_{i=1, \dots, p} g_i(\mathbf{x})$$
と表せるとする場合, 添字の集合 $J(\mathbf{x}) = \{i \mid f(\mathbf{x}) = g_i(\mathbf{x})\}$ とすると,
$$\partial f(\mathbf{x}) = \text{conv}(\{\nabla g_i(\mathbf{x}) \mid i \in J(\mathbf{x})\})$$

劣微分を例題から学ぶ

(例) 関数 $f(x) = |x|$ の $x = 0$ における劣微分



劣勾配と共役関数の関係

- 閉真凸関数 $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ に対して, 以下の3つは同値関係である
 - (a) 劣勾配 $\mathbf{g} \in \partial f(\mathbf{x})$
 - (b) $\mathbf{x} \in \partial f^*(\mathbf{g})$
 - (c) $f(\mathbf{x}) + f^*(\mathbf{g}) = \mathbf{g}^t \mathbf{x}$
- (a)と(b)を見ると, $\partial f(\mathbf{x})$ と $\partial f^*(\mathbf{g})$ がいわゆる「逆関数」の関係にある
→ 元の関数 f と共役関数 f^* の間の共役関係の本質

今日の講義のまとめ

- 凸性を有する最適化, およびそれぞれ劣勾配・劣微分について学習
- 最適化問題を凸関数・凸集合で定式化しようとすることは非常に重要
 - 全体が無理であったとしても, 局所的に凸関数・凸集合で表現できるだけでも意味があることが多い