計画数理学特論
～第2回：凸解析の基礎～

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

凸関数

(定義)

- 関数$f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$に対して，実効定義域$dom(f)$を以下で定義する.

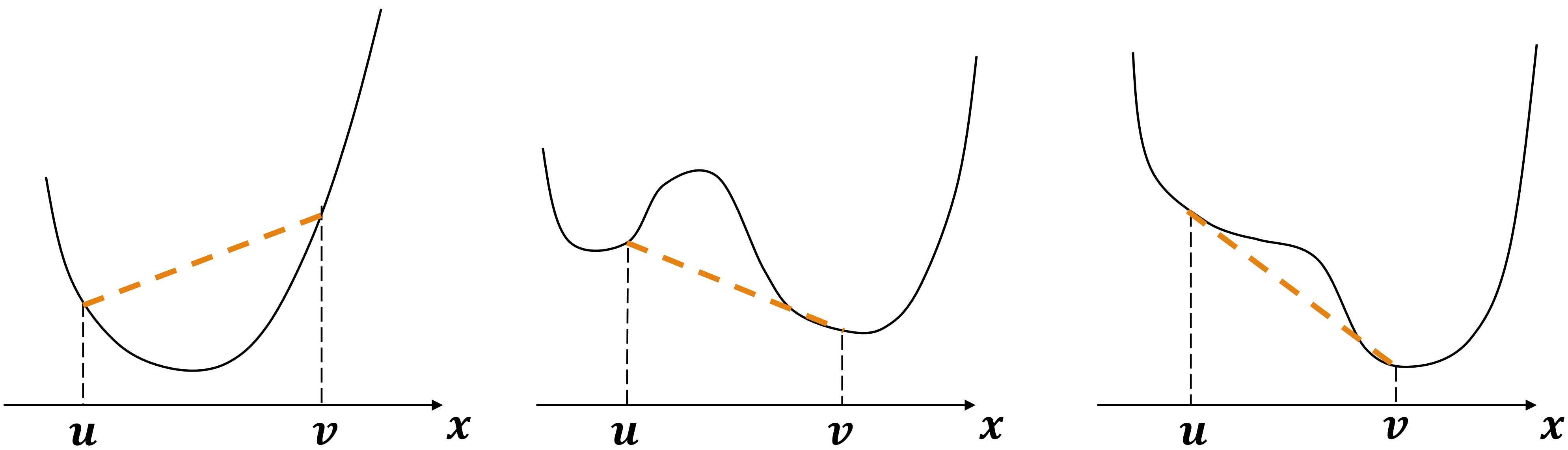
$$dom(f) = \{\boldsymbol{x} \in \mathbb{R}^n | f(\boldsymbol{x}) < \infty\}$$

- 任意の$\boldsymbol{u}, \boldsymbol{v} \in dom(f)$と$\lambda \in [0, 1]$に対して

  が成り立つとき，関数fを**凸関数**(convex function)と呼ぶ.

- 特に，$\boldsymbol{u} \neq \boldsymbol{v}$と$0 < \lambda < 1$で $f\big((1-\lambda)\boldsymbol{u} + \lambda \boldsymbol{v}\big) < (1-\lambda)f(\boldsymbol{u}) + \lambda f(\boldsymbol{v})$が成り立つとき，関数$f$を**狭義凸関数**(strictly convex function)と呼ぶ.
- $-f$が凸関数のとき，関数fを**凹関数**(concave function)と呼ぶ.

凸関数のイメージ



凸関数の例題

(例題)

・$f(x) = |x|$は凸関数であるか？

(解答)

・任意の実数a, bに対して|a + b|≦|a|+|b|が成り立つ.

・よって，任意の実数u, vおよび$\lambda \in [0,1]$に対して，
$f\big((1-\lambda)u + \lambda v\big) = |(1-\lambda)u + \lambda v| \leq |(1-\lambda)u| + |\lambda v| = (1-\lambda)f(u) + \lambda f(v)$
が成り立つので，関数fは凸関数である.

($u = 1, v = 2$のとき，$|(1-\lambda)u + \lambda v| = |(1-\lambda) + 2\lambda| = |1+\lambda| = 1+\lambda$,
$|(1-\lambda)u| + |\lambda v| = (1-\lambda) + 2\lambda = 1+\lambda$,より等号が成立するので，
狭義凸関数ではない)

凸関数の性質

(前提) 関数f, f_1, f_2：凸関数

・$dom(f)$：凸集合

・非負の定数c_1, c_2に対して$c_1 f_1(\boldsymbol{x}) + c_2 f_2(\boldsymbol{x})$も凸関数

・$max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}$も凸関数

・$g: \mathbb{R} \to \mathbb{R}$を非減少な凸関数とするとき，$g\big(f(\boldsymbol{x})\big)$は凸関数

凸関数の性質の一部証明

・非負の定数c_1, c_2に対して$c_1 f_1(\boldsymbol{x}) + c_2 f_2(\boldsymbol{x})$も凸関数

(略証)

・任意の$\boldsymbol{u}, \boldsymbol{v} \in dom\left(c_1 f_1(\boldsymbol{x}) + c_2 f_2(\boldsymbol{x})\right)$と$\lambda \in [0, 1]$に対して，

$$
\begin{aligned}
& c_1 f_1\left((1-\lambda)\boldsymbol{u} + \lambda\boldsymbol{v}\right) + c_2 f_2\left((1-\lambda)\boldsymbol{u} + \lambda\boldsymbol{v}\right) \\
& \leq c_1\left((1-\lambda) f_1(\boldsymbol{u}) + \lambda f_1(\boldsymbol{v})\right) + c_2\left((1-\lambda) f_2(\boldsymbol{u}) + \lambda f_2(\boldsymbol{v})\right) \\
& \quad = (1-\lambda)\left(c_1 f_1(\boldsymbol{x}) + c_2 f_2(\boldsymbol{x})\right) + \lambda\left(c_1 f_1(\boldsymbol{x}) + c_2 f_2(\boldsymbol{x})\right)
\end{aligned}
$$

より，凸関数であることが示せた.

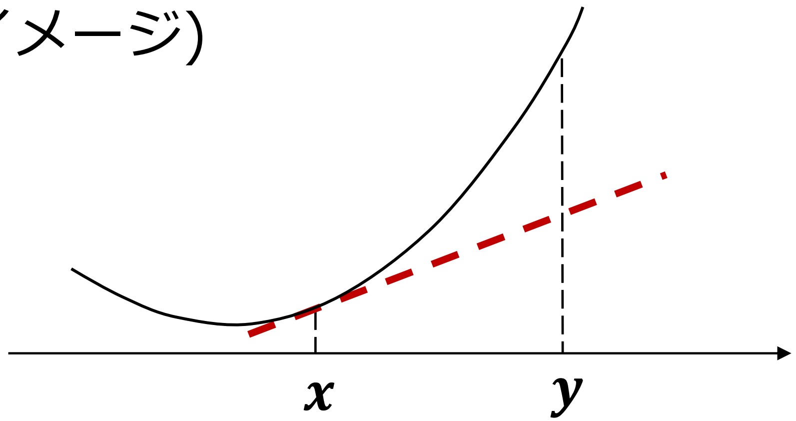
凸関数と勾配との関係性

・集合S：凸集合，関数f：$S \to \mathbb{R}$は微分可能

・このとき，fが凸関数であることと，$\boldsymbol{x} \neq \boldsymbol{y}$である任意の$\boldsymbol{x}, \boldsymbol{y} \in S$に対して

が成り立つことは同値

(イメージ)



つまり，任意のある点における接線と関数fを比較したとき，関数の方が必ず上にある(値が大きい)

凸関数と勾配との関係性

・fが凸関数である

　$\Rightarrow$ $\boldsymbol{x} \neq \boldsymbol{y}$である任意の$\boldsymbol{x}, \boldsymbol{y} \in \mathrm{S}$に対して，＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿が成り立つ

(略証)

凸関数とヘッセ行列との関係性

・集合S：凸集合，関数f：$S \to \mathbb{R}$は連続微分可能

・このとき，fが凸関数であることと，S上でヘッセ行列$\nabla^2 f$が半正定値(これを$\nabla^2 f \succeq \mathbf{0}$と表記する)となることは同値である.

(イメージ：fが$f(x), x \in \mathbb{R}$の場合，fが凸関数であることと$f''(x)$>0であること(つまり，接線の傾きが単調増加)は同値である．)

・また，S上でヘッセ行列$\nabla^2 f$が正定値($\nabla^2 f \succ \mathbf{0}$)ならば，$f$は狭義凸関数

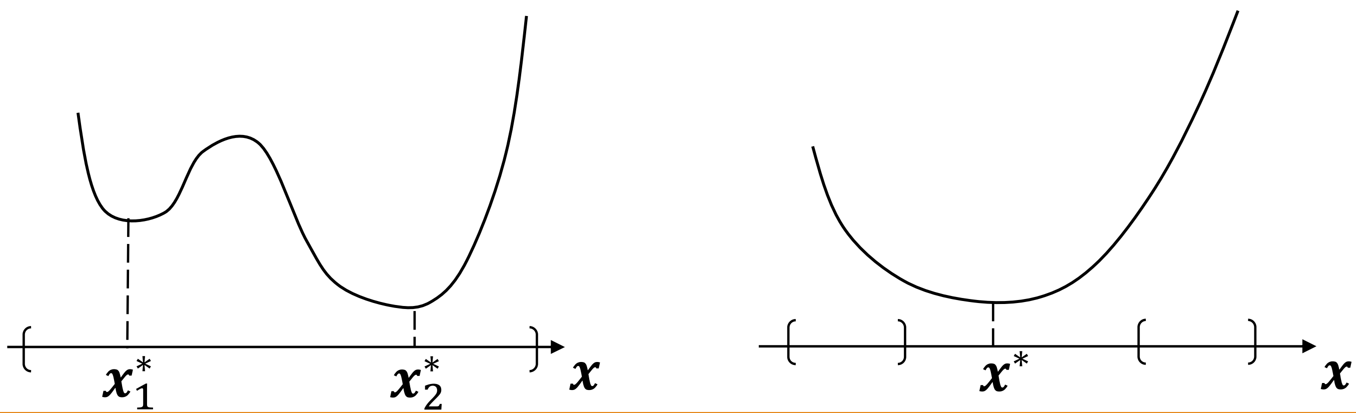
凸関数の最小化

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

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

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

このとき，$\boldsymbol{x}^*$は大域最適解となる.



凸関数の最小化(証明)

- $T = dom(f) \cap S$は，$dom(f), S$ともに凸集合であるため，Tも凸集合
- $\boldsymbol{x} \in T$ と $\lambda \in (0,1)$に対して，$\lambda \boldsymbol{x}^* + (1-\lambda)\boldsymbol{x}$も実行可能解
- $\underline{f(\boldsymbol{x}) < f(\boldsymbol{x}^*)}$<u>が成り立つなら</u>，$f$が凸関数であることから，
$$f(\lambda \boldsymbol{x}^* + (1-\lambda)\boldsymbol{x}) \leq \lambda f(\boldsymbol{x}^*) + (1-\lambda) f(\boldsymbol{x}) < f(\boldsymbol{x}^*)$$
- ここで，λを1に近い値にとると，$\boldsymbol{x}^*$の近傍における$\lambda \boldsymbol{x}^* + (1-\lambda)\boldsymbol{x}$での関数の値が，$\boldsymbol{x}^*$の時の関数の値よりも小さい

  $\rightarrow$ $\boldsymbol{x}^*$が局所最適解にならないので，矛盾

  $\therefore$ 任意の$\boldsymbol{x} \in T$に対して，$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$が成り立つ.

目的関数が1次式の場合

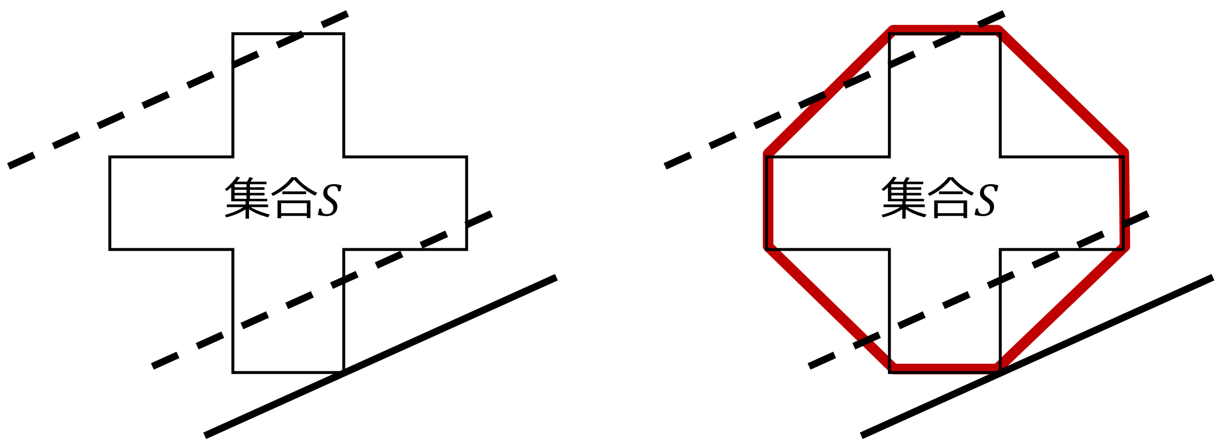
・$S \subset \mathbb{R}^n$を任意の集合(凸集合とは限らない)とする.

・$\boldsymbol{c} \in \mathbb{R}^n, a \in \mathbb{R}$とすると,

$$inf\{\boldsymbol{c}^t\boldsymbol{x} + a | \boldsymbol{x} \in S\} = inf\{\boldsymbol{c}^t\boldsymbol{x} + a | \boldsymbol{x} \in conv(S)\}$$

が成り立つ.

(イメージ)



関数値の上界・下界

- 関数$f: \mathbb{R}^n \to \mathbb{R}$が$\gamma$-平滑な関数
(参考)
  - γ-平滑：勾配∇fがリプシッツ連続性　$\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\| \leq \gamma \|\boldsymbol{x} - \boldsymbol{y}\|$ $(\gamma > 0)$
  (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
$$f(\boldsymbol{y}) \leq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^t(\boldsymbol{y} - \boldsymbol{x}) + \frac{\gamma}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2$$
- さらに関数fが凸関数なら，
$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^t(\boldsymbol{y} - \boldsymbol{x}) + \frac{1}{2\gamma}\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|^2$$

共役関数

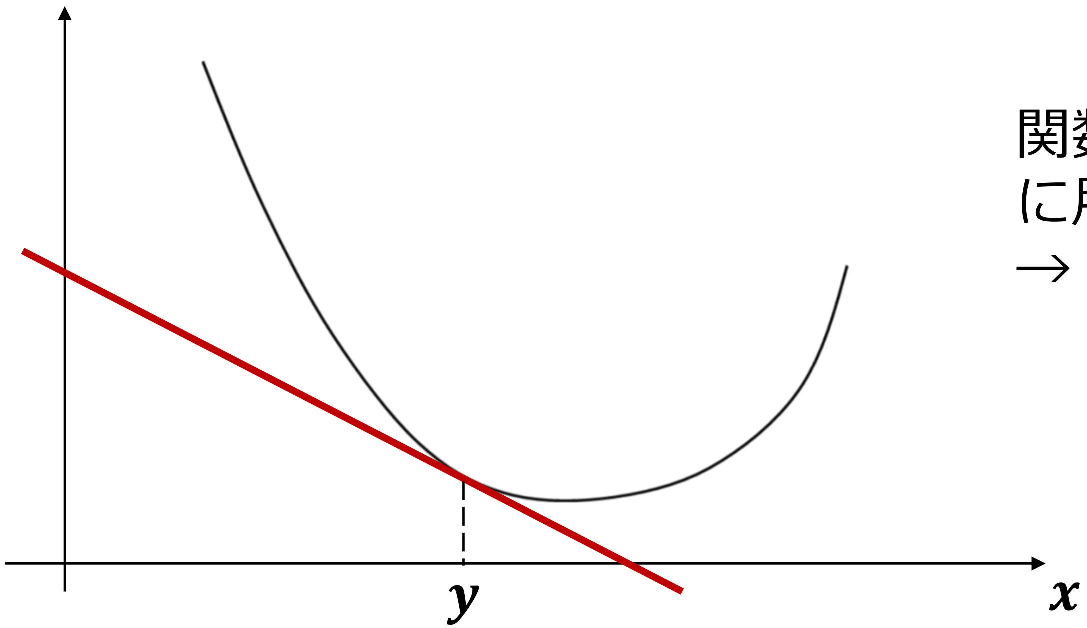
- 関数$f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$で微分可能
- エピグラフ$epi(f) \subset \mathbb{R}^{n+1}$

$$epi(f) = \{(\boldsymbol{x}, t) \in \mathbb{R}^n \times \mathbb{R} | f(\boldsymbol{x}) \leq t\}$$

- 関数fが凸関数であることと，$epi(f)$が凸集合であることは同値
  - 特に$epi(f)$が空集合でなければ，関数fは真凸関数
  - $epi(f)$が閉集合であれば，関数fは閉凸関数
- 真凸関数$f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$に対して，共役関数$f^*: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$を定義

共役関数のイメージ・関係性

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



関数の変数を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$f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$に対して，共役関数$f^*: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$を定義

- 閉凸かつ真凸な関数(＝閉真凸関数)$f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$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$\boldsymbol{x} \in dom(f)$とし，ベクトル$\boldsymbol{g} \in \mathbb{R}^n$が任意の点$\boldsymbol{y} \in \mathbb{R}^n$に対して，

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

劣微分の性質

- $\mathbf{0} \in \partial f(\boldsymbol{x})$のとき，任意の$\boldsymbol{y} \in \mathbb{R}^n$に対して，$f(\boldsymbol{y}) \geq f(\boldsymbol{x})$
- 実数$\lambda > 0$に対して，$\partial(\lambda \cdot f)(\boldsymbol{x}) = \lambda \partial f(\boldsymbol{x})$

- 関数$f: \mathbb{R}^n \to \mathbb{R}$は微分可能な凸関数ならば，$\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}$
- 関数$f: \mathbb{R}^n \to \mathbb{R}$が微分可能な凸関数$g_1, g_2, \ldots, g_p$を用いて，

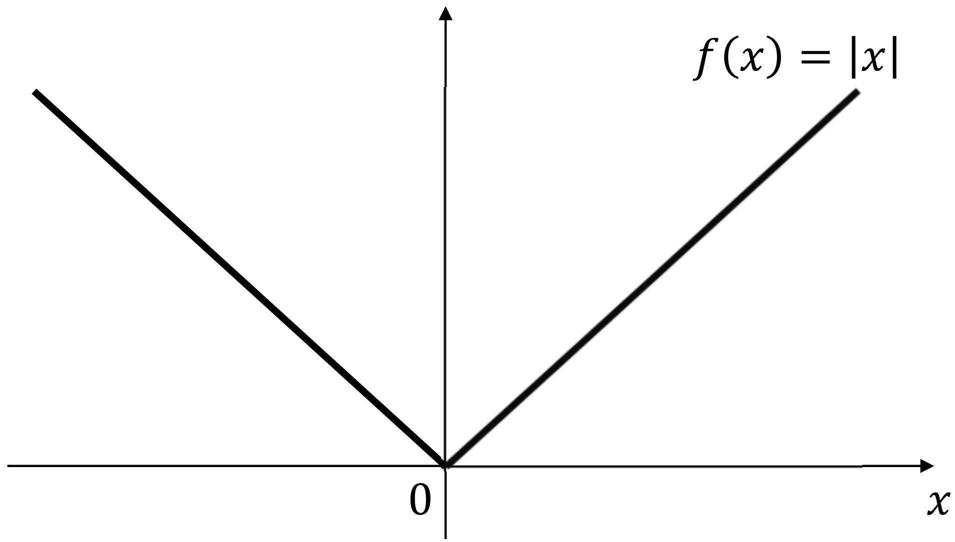
$$f(\boldsymbol{x}) = max_{i=1,\ldots,p} g_i(\boldsymbol{x})$$

  と表せるとする場合，添字の集合$J(\boldsymbol{x}) = \{i | f(\boldsymbol{x}) = g_i(\boldsymbol{x})\}$とすると，

$$\partial f(\boldsymbol{x}) = conv(\{\nabla g_i(\boldsymbol{x}) | i \in J(\boldsymbol{x})\})$$

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$f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$に対して，以下の3つは同値関係である
  (a) 劣勾配$\boldsymbol{g} \in \partial f(\boldsymbol{x})$
  (b) $\boldsymbol{x} \in \partial f^*(\boldsymbol{g})$
  (c) $f(\boldsymbol{x}) + f^*(\boldsymbol{g}) = \boldsymbol{g}^t \boldsymbol{x}$

- (a)と(b)を見ると，$\partial f(\boldsymbol{x})$と$\partial f^*(\boldsymbol{g})$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