双対性のための数学の準備

京都大学大学院情報学研究科
山下信雄

参考文献

R.T. Rockerfeller, Convex Analysis, Princeton University Press, 1997.
(初版は1970年)

福島雅夫，非線形最適化の基礎，朝倉書店，2001.
(初版は1980年の「非線形最適化の理論(産業図書)」)

記号と定義

V を**ベクトル空間**, $f: V \to R \cup \{+\infty\}$ とする.

ベクトル空間Vに**内積**が定義されているとする.

$$\langle x, y \rangle \quad (x, y \in V)$$

実行可能定義域(effective domain)

$$\mathrm{dom}\, f = \{x \in V \mid f(x) < +\infty\}$$

真凸関数(proper convex function)：

実行可能定義域が空でない凸関数

記号と定義

閉包(closure)：
集合Sを含む閉集合の中で,他の閉集合を真に含まない集合を集合Sの閉包といい， $\mathrm{cl}\, S$ と表す.
　S が閉集合のときは $S = \mathrm{cl}\, S$

ノルム(norm)：特に断らない限り $\|x\| = \sqrt{\langle x, x \rangle}$

射影(projection)：
集合 Sと x に対して， $\|y - x\|$ を最小とする
$y \in S$ を x の集合 S の射影といい，$P_S(x)$ で表す.

凸集合と凸関数

任意の $x, y \in S$ と $\alpha \in [0,1]$ に対して

$$\alpha x + (1-\alpha) y \in S$$

が成り立つとき，集合 S を**凸集合(convex set)**という．

任意の $x, y \in S$ と $\alpha \in [0,1]$ に対して

$$f(\alpha x + (1-\alpha) y) \le \alpha f(x) + (1-\alpha) f(y)$$

が成り立つとき，f を S 上で**凸関数(convex function)**という．

凸関数の性質

- １回微分可能なとき

$$f\text{ が凸関数} \Leftrightarrow f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle$$

- ２回微分可能なとき

$$f\text{ が凸関数} \Leftrightarrow \nabla^2 f(x)\text{が半正定値}$$

凸関数の劣勾配

凸関数 f が微分可能なとき，すべての $y \in V$ に対して

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$$

f が微分不可能なとき，すべての $y \in V$ に対して

$$f(y) \geq f(x) + \langle \eta, y - x \rangle$$

を満たす $\eta \in V$ を f の x における**劣勾配(subgradient)**という．

劣勾配の集合を**劣微分(subderivative)**といい,以下のようにあらわす

$$\partial f(y) = \left\{ \eta \in V \middle| f(x) \geq f(y) + \langle \eta, x - y \rangle \ \forall x \in V \right\}$$

凸錐

錐(cone)： $\alpha x \in C \quad \text{for all} \quad x \in C, \alpha \in [0,\infty)$

注意：必ず原点を含む.

凸錐(convex cone)：凸集合である錐

錐*C*の極錐(polar cone)，双対錐(dual cone)：

極錐： $C^* = \left\{ y \in V \,\middle|\, \langle y, x \rangle \le 0 \;\forall x \in C \right\}$

双対錐： $C^d = \left\{ y \in V \,\middle|\, \langle y, x \rangle \ge 0 \;\forall x \in C \right\} = -C^*$

錐の例：２次錐，半正定値錐,

共役関数(Legendre変換)

f を凸関数とする．

以下の関数 f^* を f の**共役関数(conjugate function)**という．

$$f^*(v) = \sup\left\{ \langle x, v \rangle - f(x) \,\middle|\, x \in V \right\}$$

$$\Rightarrow \quad x(v) \in \arg\sup\left\{ \langle x, v \rangle - f(x) \,\middle|\, x \in V \right\}$$

f のx(v)における接線の x = 0の値

$$v \in \partial f(x(v)),\ \ f^*(v) = \langle v, x(v) \rangle - f(x(v)) = -\Big(f(x(v)) + \langle v, 0 - x(v) \rangle \Big)$$

凸関数の表現① グラフ $\left\{ (x, y) \,\middle|\, y = f(x) \right\}$

凸関数の表現② 傾きと切片 $\left\{ (v, \alpha) \,\middle|\, \alpha = -f^*(v) \right\}$

共役関数の性質

- (f が凸関数でなくても)共役関数は凸関数
- f が閉凸関数であるとき $(f^*)^* = f$
- (f が凸関数でなくても) $f(x) + f^*(v) \geq \langle x, v \rangle$
- $v \in \partial f(x) \Leftrightarrow x \in \partial f^*(v)$
- f が強凸関数 $\Leftrightarrow f^*$ が微分可能で，∇f^*がリプシッツ連続

共役関数の例

線形関数

$$f(x) = c^\top x \quad \Rightarrow \quad f^*(y) = \begin{cases} 0 & \text{if } y = c \\ +\infty & \text{otherwise} \end{cases}$$

凸2次関数

$$f(x) = \frac{1}{2} x^\top A x \quad \Rightarrow \quad f^*(y) = \frac{1}{2} y^\top A^{-1} y$$

平行移動

$$f(x) = g(x - b) \quad \Rightarrow \quad f^*(y) = g^*(y) + b^\top y$$

分離可能な和

$$f(x^1, x^2) = g_1(x^1) + g_2(x^2) \quad \Rightarrow \quad f^*(y^1, y^2) = g_1^*(y^1) + g_2^*(y^2)$$

標示関数と支持関数

$S \subseteq V$ を凸集合とする.

$$\delta_S(x) = \begin{cases} 0 & \text{if} \;\; x \in S \\ +\infty & \text{if} \;\; x \notin S \end{cases}$$

を S の**標示関数(indicator function)**という.
δ_S は凸関数である.

δ_S の共役関数 δ_S^* をS の**支持関数(support function)**という.

$$\delta_S^*(u) = \sup_{x \in V} \left\{ \langle x, u \rangle - \delta_S(x) \right\} = \sup_{x \in S} \left\{ \langle x, u \rangle \right\}$$

共役関数の計算例

L1ノルム： $f(x)=\sum_{i=1}^{n}|x_i|$

$$f^*(v)=\sup_{x\in R^n}\left\{\langle x,v\rangle-f(x)\right\}=\sum_{i=1}^{n}\sup_{x_i\in R}\{x_iv_i-|x_i|\}$$

$$\sup\{x_iv_i-|x_i|\}=\sup\{|x_i||v_i|-|x_i|\}=\begin{cases}0 & \text{if } |v_i|\le 1\\ +\infty & \text{otherwise}\end{cases}$$

凸２次関数： $f(x)=\frac{1}{2}x^\top Qx$ （Qは正定値対称行列）

$$f^*(v)=\sup_{x\in R^n}\left\{\langle x,v\rangle-\frac{1}{2}x^\top Qx\right\}$$

$$=\langle v,Q^{-1}v\rangle-\frac{1}{2}(Q^{-1}v)^\top QQ^{-1}v$$

$$=\frac{1}{2}v^\top Q^{-1}v$$

$$\min_{x\in R^n}-\langle x,v\rangle+\frac{1}{2}x^\top Qx$$

最適性の条件

$$0=-v+Qx$$

$$\Rightarrow x=Q^{-1}v$$

共役関数の計算例

線形関数： $f(x) = \langle a, x \rangle + b$

$$
\begin{aligned}
f^*(v) &= \sup_{x \in V} \{ \langle x, v \rangle - \langle a, x \rangle - b \} \\
&= -b + \sup_{x \in V} \{ \langle v - a, x \rangle \} \\
&= \begin{cases} -b & \text{if } v = a \\ +\infty & \text{if } v \neq a \end{cases}
\end{aligned}
$$

共役関数の計算例

線形関数の共役関数の共役関数：

$$g(x) = \begin{cases} -b & \text{if } x = a \\ +\infty & \text{if } x \neq a \end{cases}$$

$$\begin{aligned} g^*(v) &= \sup_{x \in V} \left\{ \langle x, v \rangle - g(x) \right\} \\ &= \sup_{x = a} \left\{ \langle x, v \rangle - g(x) \right\} \\ &= \langle a, v \rangle + b \end{aligned}$$

適当な仮定のもとで，共役関数の共役関数は元の関数になる

共役関数の計算例

錐Cの支持関数： 錐Cの標示関数 δ_C

$$\delta_C^*(v) = \sup_{x \in C} \langle v, x \rangle = \begin{cases} 0 & \text{if } v \in C^* \\ +\infty & \text{if } v \notin C^* \end{cases} \quad \Rightarrow \quad \delta_C^*(v) = \delta_{C^*}(v)$$

平行移動： $p(x) = f(x-a)$

$$\begin{aligned} p^*(v) &= \sup_{x \in V} \{\langle v, x \rangle - f(x-a)\} = \sup_{y+a \in V} \{\langle v, y+a \rangle - f(y)\} \\ &= \langle v, a \rangle + \sup_{y \in V} \{\langle v, y \rangle - f(y)\} \\ &= \langle v, a \rangle + f^*(v) \end{aligned}$$

ゲージ関数

定義

以下の性質を満たす関数　$\rho : V \to R \cup \{\infty\}$

を**ゲージ関数(gauge function)**という

(i)　非負の関数

(ii)　凸関数

(iii)　正斉次(Positively homogeneous)

$$\rho(tx) = t\rho(x) \quad \forall t \geq 0$$

ノルムとゲージ関数

norm: 標準, 基準…..

gauge: 基準, 規格, 標準寸法….

ノルムはゲージ関数

非負性： $\|x\| \geq 0$

正斉次性： $\|tx\| = t\|x\| \quad (t \geq 0)$

凸性： $\alpha \in [0,1]$ に対して

$$\|\alpha x + (1-\alpha)y\| \leq \|\alpha x\| + \|(1-\alpha)y\| = \alpha\|x\| + (1-\alpha)\|y\|$$

三角不等式

ノルムの例（有名なもの）

ベクトルのノルム

- pノルム　$\|x\|_p = \sqrt[p]{\sum |x_i|^p}$
- 無限大ノルム　$\|x\|_\infty = \max_i |x_i|$
- 1ノルム　$\|x\|_1 = \sum |x_i|$

⇒　正則化やスパース化によく使われる

行列のノルム　$\sigma(A)$を行列Aの特異値を並べたベクトルとする

- Frobenius ノルム　$\|A\|_F = \|\sigma(A)\|_2$
- Nuclear ノルム　$\|A\|_* = \|\sigma(A)\|_1$

⇒　Nuclearノルムは rank(A) の 凸緩和として使われる

ノルムの例(有名でないもの?)

$x_{(i)}$ ： $x \in R^n$ の成分を，絶対値の大きい順に並べたときの i 番目

$$|x_{(1)}| \geq |x_{(2)}| \geq \cdots \geq |x_{(n)}|$$

利用例 $\|x\|_{\infty} = |x_{(1)}| = \sum_{i=1}^{1} |x_{(i)}|, \quad \|x\|_1 = \sum_{i=1}^{n} |x_{(i)}|$

CVaRノルム (largest-k ノルム)

$$x \quad _{\alpha} = \sum_{i=1}^{n-n\alpha} |x_{(i)}|$$

ある線形計画問題の最適値として表せるため，最適化モデルで使いやすい

応用： 金融のCVaR，ν -SVM

ノルムでないゲージ関数

- 0とのMax関数

$$\sigma(x) = \sum_{i=1}^{n} \max\{0, x_i\}$$

＊損失関数やペナルティ関数に使われる

- 凸錐 C の標示関数 $\delta_C(x)$

凸関数の標示関数 ⇒ 非負の凸関数

錐の標示関数 ⇒ 正斉次関数

ゲージ関数の極関数

ゲージ関数　f の**極関数(polar function)**

$$f^{\circ}(y) = \sup\left\{ \langle x, y \rangle \mid f(x) \le 1 \right\}$$

f がノルムのとき，f° は双対ノルム

参考：共役関数

$$f^{*}(y) = \sup_x \left\{ \langle y, x \rangle - f(x) \right\}$$

双対ノルムの例

- L1ノルム $\|x\|_1$ $\Longleftrightarrow$ 無限大ノルム $\|x\|_\infty$
- pノルム $\|x\|_p = \sqrt[p]{\sum |x_i|^p}$ $\Longleftrightarrow$ $\|x\|_q$ ただし$\dfrac{1}{p}+\dfrac{1}{q}=1$
- CVaRノルム x_{α} $\Longleftrightarrow$ $x^{\circ}_{\alpha} = \max\left\{ \dfrac{\|x\|_1}{n-n\alpha}, \|x\|_\infty \right\}$

極関数の例

- 0とのMax関数

$$g(x) = \sum_{i=1}^{m} \max\{0, x_i\} \quad \Longleftrightarrow \quad g^{\circ}(y) = \begin{cases} \max_i y_i & \text{if} \quad y \geq 0 \\ +\infty & \text{otherwise} \end{cases}$$

- 錐 C の標示関数 $\Longleftrightarrow$ 極錐 C° の標示関数 $\delta_{C^{\circ}}(x)$

$$\delta_C(x)$$

ただし

$$C^{\circ} = \{y \mid \langle x, y \rangle \leq 0 \ \forall x \in C\}$$

極関数の諸性質

共役関数の諸性質(再掲)

- $f^{**} = (f^*)^* = f$
- $f(x) + f^*(y) \geq \langle x, y \rangle$
- f^* は凸関数

ゲージ関数の極関数の諸性質

- $f^{\circ\circ} = (f^\circ)^\circ = f$
- $f(x) \times f^\circ(y) \geq \langle x, y \rangle \quad \forall x \in \mathrm{dom}\, f, \forall y \in \mathrm{dom}\, f^\circ$
- f° はゲージ関数

Gauge双対のGauge双対
を
考える上で重要

Gauge双対の
双対性で重要

Gauge双対の
凸性で重要