基本劣微分を用いたKKT条件に対する必要十分制約想定について

島根大学大学院自然科学研究科　大谷浩之 (Hiroyuki Ohtani)
Graduate School of Natural Science and Technology, Shimane University
島根大学総合理工学部　黒岩大史 (Daishi Kuroiwa)
Major in Interdisciplinary Science and Engineering, Shimane University

1　はじめに

本講究録では次のような拡張実数値最適化問題について考える：

$$
\begin{array}{lll}
\text{(P)} & \text{minimize} & f(x) \\
 & \text{subject to} & g_i(x) \leqq 0, \forall i \in I
\end{array}
$$

ただし, I は任意の空でない添字集合, $f, g_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ とする．この問題の KKT 条件に対する必要十分制約想定について観察する．過去の KKT 条件に対する必要十分制約想定の結果として, まず 1971 年に Gould, Tolle が [1] で f, g_i が微分可能関数のときに KKT 条件といわれる

$$\nabla f(\overline{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\overline{x}) = 0, \lambda_i g_i(\overline{x}) = 0 (i \in I)$$

という等式を用いて KKT 条件に対する必要十分制約想定を与えた．次に 2008 年に Li, Ng, Pong が [6] で f, g_i が凸関数のとき

$$\partial f(\overline{x}) + \sum_{i \in I} \lambda_i \partial g_i(\overline{x}) = 0, \lambda_i g_i(\overline{x}) = 0 (i \in I)$$

という等式を用いて KKT 条件に対する必要十分制約想定を与えた．次に 2016 年に Yamamoto, Kuroiwa が [7] で f が凸関数, g_i が locally Lipschitz 関数のとき

$$\partial f(\overline{x}) + \sum_{i \in I} \lambda_i \partial^c g_i(\overline{x}) = 0, \lambda_i g_i(\overline{x}) = 0 (i \in I)$$

という等式を用いて KKT 条件に対する必要十分制約想定を与えた．そこで本講究録では f が locally Lipschitz 関数のときに基本劣微分を用いた KKT 条件に対する必要十分制約想定について考察していく．

2　KKT 条件に対する必要十分制約想定の先行研究と準備

まず f, g_i が微分可能の場合の KKT 条件に対する制約想定と既存の結果について述べる．$I = \{1, \cdots, m\}$ とし, $g_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ は $\overline{x} \in S = \{x \in \mathbb{R}^n \mid g_i(x) \leqq 0, \forall i \in I\}$ で微分可能とする．このとき g_i に Q という条件を仮定したとき, $\overline{x}$ で微分可能な任意の関数 $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ に対して, 以下の「(1) ならば (2)」が成立する：

(1) $\exists r > 0$ s.t. $\forall x \in S \cap B(\overline{x}, r), f(\overline{x}) \leq f(x)$

(2) $\exists \lambda_i \geq 0 (i = 1, \cdots, m)$ s.t. $\begin{cases} \nabla f(\overline{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\overline{x}) = 0 \\ \lambda_i g_i(\overline{x}) = 0 (i = 1, \cdots, m) \end{cases}$

この条件 Q を KKT 条件に対する制約想定という．これまでに KKT 条件に対する制約想定として

- 一次独立制約想定

$$\nabla g_i(\overline{x}) \, (i \in I(\overline{x})) \text{ が一次独立}$$

- Cottle 制約想定

$$\exists y \in \mathbb{R}^n \text{s.t.} \forall i \in I(\overline{x}), \langle \nabla g_i(\overline{x}), y \rangle \leq 0$$

- Abadie 制約想定

$$T(\overline{x}; S) = \{x \in \mathbb{R}^n \mid \langle \nabla g_i(\overline{x}), x \rangle \leq 0, \forall i \in I(\overline{x})\}$$

- Guignard 制約想定

$$N(\overline{x}; S) \subseteq \operatorname{cone} \operatorname{co} \bigcup_{i \in I(\overline{x})} \nabla g_i(\overline{x})$$

が示されてきた．これらの制約想定は

$$\text{一次独立} \Rightarrow \text{Cottle} \Rightarrow \text{Abadie} \Rightarrow \text{Guignard}$$

の順に弱くなっている．重要な点は, Guignard 制約想定より弱い制約想定は存在しないことで, このことは次の定理で保証されている．

定理 2.1 (Gould, F. J, Tolle, J. W. [1]) I は空でない有限添字集合, $g_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} (i \in I)$ は $\overline{x} \in S$ で微分可能な関数とする．このとき次は同値 :

- $N(\overline{x}; S) \subseteq \operatorname{cone} \operatorname{co} \bigcup_{i \in I(\overline{x})} \nabla g_i(\overline{x})$
- $\overline{x}$ で微分可能な任意の関数 $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ に対して, (1) $\Rightarrow$ (2) が成立 :

  (1) $\exists r > 0$ s.t. $\forall x \in S \cap B(\overline{x}, r), f(\overline{x}) \leq f(x)$

  (2) $\exists \lambda_i \geq 0 \, (i = 1, \cdots, m)$ s.t. $\begin{cases} \nabla f(\overline{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\overline{x}) = 0 \\ \lambda_i g_i(\overline{x}) = 0 (i = 1, \cdots, m) \end{cases}$

ただし,

$$I(\overline{x}) := \{i \in I \mid g_i(\overline{x}) = 0\}, N(\overline{x}; S) := \{z \in \mathbb{R}^n \mid \langle z, y \rangle \leq 0, \forall y \in T(\overline{x}; S)\}$$

$$T(\overline{x}; S) := \left\{ y \in \mathbb{R}^n \,\middle|\, \begin{array}{c} \exists \{x_k\} \subseteq S : x_k \to \overline{x}, \exists \alpha_k \geq 0 (k \in \mathbb{N}) \\ \text{s.t.} \alpha_k (x_k - \overline{x}) \to y \end{array} \right\}$$

すなわち, Guignard の制約想定は $\overline{x}$ で微分可能な任意の関数 $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ に対して, (1) $\Rightarrow$ (2) が成り立つことと同値な条件である. このような制約想定を KKT 条件に対する必要十分制約想定とよぶ. 次に f, g_i が凸関数である場合の KKT 条件に対する必要十分制約想定を述べる. この場合の KKT 条件は凸関数の劣微分を用いて表されたものである.

定理 2.2 (Li, C. Ng, K. F. Pong, T. K. [6]) I は空でない添字集合, $g_i:\mathbb{R}^n\to\mathbb{R}(i\in I)$ は凸関数, $\overline{x}\in S$ とする. このとき次は同値 :

- $\overline{x}$ で BCQ が成立, すなわち,

$$N(\overline{x};S)\subseteq \operatorname{cone}\operatorname{co}\bigcup_{i\in I(\overline{x})}\partial g_i(\overline{x})$$

- 任意の凸関数 $f:\mathbb{R}^n\to\mathbb{R}$ に対して, (1) $\Rightarrow$ (2) が成立 :

(1) $\forall x\in S, f(\overline{x})\leq f(x)$

(2) $\exists J\subseteq I(\overline{x})$: 有限, $\exists\lambda_i\geq 0(i\in J)$ s.t. $\begin{cases}0\in\partial f(\overline{x})+\sum_{i\in J}\lambda_i\partial g_i(\overline{x})\\ \lambda_i g_i(\overline{x})=0,\ i\in J\end{cases}$

ただし, f の $\overline{x}$ における劣微分は次のように与えられている :

$$\partial f(\overline{x}):=\{z\in\mathbb{R}^n\mid f(y)\geq f(\overline{x})+\langle z, y-\overline{x}\rangle, \forall y\in\mathbb{R}^n\}$$

実際に定理を用いて次の問題を解く.

例 2.1 次の問題を考える. この問題は解をもつとする.

$$\begin{array}{ll}\text{minimize} & 2|x|\\ \text{subject to} & x-1\leq 0\\ & -x-2\leq 0\end{array}$$

与えられている関数が全て凸関数なので定理 2.2 を用いてこの問題の解を考えていく.

$$f(x)=2|x|, g_1(x)=x-1, g_2(x)=-x-2$$

とすると,

$$\partial f(\overline{x})=\begin{cases}\{2\} & (\overline{x}>0)\\ [-2,2] & (\overline{x}=0)\\ \{-2\} & (\overline{x}<0)\end{cases}, \partial g_1(\overline{x})=\{1\}, \partial g_2(\overline{x})=\{-1\}$$

BCQ が成立していることは簡単にわかるので, $\overline{x}$ が問題の最適解であるとして定理を用いると以下のことが成り立つ :

$$\exists\lambda_1,\lambda_2\geq 0 \text{ s.t.}\begin{cases}0\in\partial f(\overline{x})+\lambda_1\partial g_1(\overline{x})+\lambda_2\partial g_2(\overline{x})\\ \lambda_1 g_1(\overline{x})=0, \lambda_2 g_2(\overline{x})=0\end{cases}$$

つまり

$$\exists\lambda_1,\lambda_2\geq 0 \text{ s.t.}\begin{cases}0\in\partial f(\overline{x})+\lambda_1\{1\}+\lambda_2\{-1\}\\ \lambda_1(x-1)=0, \lambda_2(-x-2)=0\end{cases}$$

(1) $\overline{x}=1$ のとき

$$\lambda_1 \geq 0, \lambda_2 = 0, 0 \in \{2\} + \lambda_1\{1\}$$

これは矛盾.

(2) $\overline{x}=-2$ のとき

$$\lambda_1 = 0, \lambda_2 \geq 0, 0 \in \{-2\} + \lambda_2\{-1\}$$

これは矛盾.

(3) $\overline{x}\neq 1, -2$ のとき

$$\lambda_1 = 0, \lambda_2 = 0, 0 \in \partial f(\overline{x})$$

$0 \in \partial f(\overline{x})$ をみたす $\overline{x}$ は $\overline{x}=0$ のみ.

よって 0 がこの問題の解となる.

次に f が凸関数, g_i が locally Lipschitz 関数である場合の KKT 条件に対する必要十分制約想定を述べる. この場合の KKT 条件は凸関数の劣微分および locally Lipschitz 関数の Clarke の劣微分を用いて表されたものである.

定理 2.3 (Yamamoto, S. Kuroiwa, D. [7]) I は空でない添字集合, $g_i : \mathbb{R}^n \to \mathbb{R}(i \in I)$ は locally Lipschitz 関数, $\overline{x} \in S$, S は凸集合とする. このとき次は同値 :

- $N(\overline{x}; S) \subseteq \operatorname{cone}\operatorname{co} \bigcup_{i\in I(\overline{x})} \partial^c g_i(\overline{x})$
- 任意の凸関数 $f : \mathbb{R}^n \to \mathbb{R}$ に対して, (1) $\Rightarrow$ (2) が成立 :
  - (1) $\forall x \in S, f(\overline{x}) \leq f(x)$
  - (2) $\exists \lambda_i \geq 0 (i \in I)$ s.t. $\begin{cases} 0 \in \partial f(\overline{x}) + \sum_{i\in I} \lambda_i \partial^c g_i(\overline{x}) \\ \lambda_i g_i(\overline{x}) = 0,\ i \in I \end{cases}$

ただし, f の $\overline{x}$ における Clarke の劣微分は次のように与えられている :

$$\partial^c f(\overline{x}) := \left\{ z \in \mathbb{R}^n \,\middle|\, \langle z, d\rangle \leq \limsup_{\substack{y\to\overline{x}, y\in\mathbb{R}^n\setminus\{\overline{x}\} \\ t\searrow 0, t\in(0,+\infty)}} \frac{f(y+td)-f(y)}{t}, \forall d \in \mathbb{R}^n \right\}$$

以上のことから定理 2.1, 2.2, 2.3 では法線錐 $N(\overline{x}; S)$ と微分 $\nabla f(\overline{x})$, 劣微分 $\partial f(\overline{x})$, Clarke の劣微分 $\partial^c f(\overline{x})$ といったように用いられている微分が異なる. 本講究録では基本法線錐と基本劣微分を用いるので, それらの概念を定義する.

定義 2.1 (基本法線錐 [4]) X をバナッハ空間, $(\emptyset \neq) S \subseteq X$, $\varepsilon \geq 0$ とする. このとき $x \in S$ に対して

$$N^b_\varepsilon(x; S) := \left\{ x^* \in X^* \,\middle|\, \limsup_{\substack{y\in S\setminus\{x\} \\ y\to x}} \frac{\langle x^*, y-x\rangle}{||y-x||} \leq \varepsilon \right\}$$

を S の x における ε-法線錐という．また $\overline{x} \in S$ に対して

$$N^b(\overline{x}; S) := \mathop{\mathrm{Limsup}}_{\substack{\varepsilon \to 0 \\ x \to \overline{x} \\ x \in S \setminus \{\overline{x}\}}} N^b_\varepsilon(x; S)$$

を S の $\overline{x}$ における基本法線錐という．

定義 2.2 (基本劣微分 [4]) X はバナッハ空間, $f : X \to \mathbb{R} \cup \{+\infty\}$, $\overline{x} \in \mathrm{dom}\, f$ とする．

$$\partial^b f(\overline{x}) := \{x^* \in X^* \mid (x^*, -1) \in N^b((\overline{x}, f(\overline{x})); \mathrm{epi}\, f)\}$$

を $\overline{x}$ における基本劣微分という．

3 主結果

f が locally Lipschitz 関数, g_i は任意の関数の場合の, 基本劣微分を用いた KKT 条件に対する必要十分制約想定を述べる．

定理 3.1 (Kuroiwa, D. Ohtani, H. [8]) I は空でない添字集合, $g_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} (i \in I)$, $\overline{x} \in S$, S は凸集合とする．このとき次は同値 :

- $N^b(\overline{x}; S) \subseteq \mathrm{cone}\,\mathrm{co} \bigcup_{i \in I(\overline{x})} \partial^b g_i(\overline{x})$
- $\overline{x}$ で locally Lipschitz な任意の関数 $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ に対して, (1)⇒(2) が成立 :

  (1) $\forall x \in S, f(\overline{x}) \le f(x)$

  (2) $\exists J \subseteq I(\overline{x})$: 有限, $\exists \lambda_i \ge 0 (i \in J)$ s.t. $\begin{cases} 0 \in \partial^b f(\overline{x}) + \sum_{i \in J} \lambda_i \partial^b g_i(\overline{x}) \\ \lambda_i g_i(\overline{x}) = 0,\ i \in J \end{cases}$

この結果の良い点は, 定理 2.1, 2.2, 2.3 と比較して問題に使われている関数の範囲が広がっていることである．実際,

$$f \text{ は凸関数} \Rightarrow \partial f(\overline{x}) = \partial^b f(\overline{x})$$

$$f \text{ は凸関数} \Rightarrow f \text{ は locally Lipschitz 関数}$$

が成り立つので定理 3.1 は定理 2.2 の完全な拡張である．次に

$$\partial^c f(\overline{x}) \ne \partial^b f(\overline{x})$$

という locally Lipschitz 関数 f が存在するので, 定理 3.1 は定理 2.3 の完全な拡張とはいえない．しかし定理 3.1 は g_i の locally Lipschitz 性の仮定が不要かつ

$$f \text{ は凸関数} \Rightarrow f \text{ は locally Lipschitz 関数}$$

が成り立つので, 定理 3.1 が適用できる最適化問題の範囲が定理 2.3 よりも極めて広い．次に

$$\{\nabla f(\overline{x})\} \ne \partial^b f(\overline{x})$$

という微分可能な関数 f が存在し, さらに定理 2.1 には S が凸の仮定がないので, 定理 3.1 は定理 2.1 の完全な拡張とはいえない. しかし

$$f \text{ は微分可能な関数} \Rightarrow f \text{ は locally Lipschitz 関数}$$

が成り立つので, S が凸である条件の下では, 定理 3.1 が適用できる最適化問題の範囲が定理 2.1 よりも極めて広い. 次に実際に主定理を用いた例を述べる.

例 3.1 次の問題を考える.

$$\begin{array}{ll} \text{minimize} & -|x|+2 \\ \text{subject to} & x-1 \leq 0 \\ & -x-2 \leq 0 \end{array}$$

この問題は目的関数が locally Lipschitz 関数なので定理 3.1 を用いて解く (目的関数が locally Lipschitz 関数なので, 我々の知る限り, 直接的には過去の結果を用いることができない). まず

$$f(x) = -|x|+2, g_1(x) = x-1, g_2(x) = -x-2$$

とおくと基本劣微分は

$$\partial^b f(\overline{x}) = \begin{cases} \{-1\} & (\overline{x} > 0) \\ \{-1, 1\} & (\overline{x} = 0) \\ \{1\} & (\overline{x} < 0) \end{cases}, \partial^b g_1(\overline{x}) = \{1\}, \partial^b g_2(\overline{x}) = \{-1\}$$

このとき

$$N^b(\overline{x}; S) \subseteq \operatorname{cone} \operatorname{co} \bigcup_{i \in I(\overline{x})} \partial^b g_i(\overline{x})$$

が成立することは簡単にわかるので, $\overline{x}$ が問題の最適解であるとして主定理を用いると以下のことが成り立つ :

$$\exists \lambda_1, \lambda_2 \geq 0 \,\text{s.t.} \begin{cases} 0 \in \partial^b f(\overline{x}) + \lambda_1 \partial^b g_1(\overline{x}) + \lambda_2 \partial^b g_2(\overline{x}) \\ \lambda_1 g_1(\overline{x}) = 0, \lambda_2 g_2(\overline{x}) = 0 \end{cases}$$

つまり

$$\exists \lambda_1, \lambda_2 \geq 0 \,\text{s.t.} \begin{cases} 0 \in \partial^b f(\overline{x}) + \lambda_1\{1\} + \lambda_2\{-1\} \\ \lambda_1(\overline{x}-1) = 0, \lambda_2(-\overline{x}-2) = 0 \end{cases}$$

が成り立つ.

(1) $\overline{x} = 1$ のとき

$$\lambda_1 \geq 0, \lambda_2 = 0, 0 \in \{-1\} + \lambda_1\{1\}$$

これをみたすのは $\lambda_1 = 1$.

(2) $\overline{x} = -2$ のとき

$$\lambda_1 = 0, \lambda_2 \geq 0, 0 \in \{1\} + \lambda_2\{-1\}$$

これをみたすのは $\lambda_2 = 1$.

(3) $\overline{x} \neq 1, -2$ のとき

$$\lambda_1 = 0, \lambda_2 = 0, 0 \in \partial^b f(\overline{x})$$

$0 \in \partial^b f(\overline{x})$ をみたす $\overline{x}$ は存在しないので矛盾.

以上より主定理を使うことで, この問題の解は $x = 1$ もしくは $x = -2$ であり, $f(-2) < f(1)$ より $x = -2$ のみが解になる可能性があることがわかる. 実際に $x = -2$ がこの問題の解である.

参考文献

[1] Gould, F. J.; Tolle, J. W. A necessary and sufficient qualification for constrained optimization SIAM J. Appl. Math. 20 (1971), 162-172.

[2] Mordukhovich, B. S. Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions. Trans. Amer. Math. Soc. 340 (1993), no. 1, 1-35.

[3] Mordukhovich, B. S. Generalized differential calculus for nonsmooth and set-valued mappings. J. Math. Anal. Appl. 183 (1994), no. 1, 250-288.

[4] Mordukhovich, B. S. Variational analysis and generalized differentiation. I. Springer-Verlag, Berlin, (2006).

[5] Goberna, M. A.; Jeyakumar, V.; López, M. A. Necessary and sufficient constraint qualifications for solvability of systems of infinite convex inequalities. Nonlinear Anal. 68 (2008), no. 5, 1184-1194.

[6] Li, C.; Ng, K. F.; Pong, T. K. Constraint qualifications for convex inequality systems with applications in constrained optimization. SIAM J. Optim. 19 (2008), no. 1, 163-187.

[7] Yamamoto, S.; Kuroiwa, D. Constraint qualifications for KKT optimality condition in convex optimization with locally Lipschitz inequality constraints. Linear Nonlinear Anal.2 (2016), no. 1, 101-111.

[8] Kuroiwa, D.; Ohtani, H. Constraint qualification for KKT condition by basic subdifferential, preprint.

[9] 今野 浩.; 山下 浩. 非線形最適化. 日科技連 (1978).