2015年6月19日　統計数理研究所　オープンハウス

ヒルベルト空間における局所リプシッツ最適化

伊藤　聡　数理・推論研究系，統計的機械学習研究センター，統計思考院　教授

一般方向微係数と一般勾配

X を実ヒルベルト空間，D を X の非空開部分集合，f を D 上の実汎関数とするとき，f の $x \in D$ における方向 $s \in X$ への一般方向微係数 $f^\circ(x;s)$ と $x \in D$ における一般勾配 $\partial^\circ f(x)$ を以下のように定義する (Clarke, 1981)。

$$f^\circ(x;s) := \limsup_{\substack{y \to x \\ \alpha \downarrow 0}} \frac{f(y+\alpha s) - f(y)}{\alpha}$$
$$\partial^\circ f(x) := \{\, z \in X \mid f^\circ(x;s) \geq \langle z, s \rangle \quad \forall s \in X \,\}$$

f が $x \in D$ の近傍でリプシッツ連続ならば，任意の方向 $s \in X$ への一般方向微係数 $f^\circ(x;s)$ が存在し，$f^\circ(x;\cdot)$ は X 上の連続準線形（正斉次劣加法的すなわち凸）汎関数となる。また，このとき一般勾配 $\partial^\circ f(x)$ は X の非空弱コンパクト凸部分集合となり，次の関係が成立する。

$$f^\circ(x;s) = \max_{z \in \partial^\circ f(x)} \langle z, s \rangle \quad \forall s \in X$$

局所リプシッツ汎関数の例として，D が特に凸集合であるとき D 上の連続凸汎関数が挙げられるが，このとき一般方向微係数と一般勾配はそれぞれ方向微係数と劣勾配に一致する。すなわち一般勾配は劣勾配の概念を非凸汎関数に拡張したものである。

一般化されたFarkasの定理

S が X の非空部分集合であるとき，$\bigcup_{\lambda \geq 0} \lambda S$ および $\bigcup_{\lambda \geq 0} \lambda \operatorname{co} S = \operatorname{co} \bigcup_{\lambda \geq 0} \lambda S$ はそれぞれ S の錐包（S を含む最小の錐），S の凸錐包（S を含む最小の凸錐）である。閉凸包の錐包 $\bigcup_{\lambda \geq 0} \lambda \overline{\operatorname{co}}\, S$ は凸錐であるが必ずしも閉凸錐であるとは限らないことに注意されたい。

連続準線形不等式系に対してFarkas型の二者択一定理は以下のように一般化される。

> A, B_i, $i = 1, 2, \ldots, m$ を X の非空弱コンパクト集合とする。このとき
> $$\bigcup_{\lambda \geq 0} \lambda \overline{\operatorname{co}} \bigcup_{i=1}^{m} B_i \text{ が閉} \tag{$*$}$$
> ならば，
> $$\max_{a \in A} \langle a, x \rangle < 0; \quad \max_{b_i \in B_i} \langle b_i, x \rangle \leq 0, \quad i = 1, 2, \ldots, m \tag{I}$$
> を満たす $x \in X$ が存在するか，あるいは
> $$0 \in \overline{\operatorname{co}}\, A + \bigcup_{\lambda \geq 0} \lambda \overline{\operatorname{co}} \bigcup_{i=1}^{m} B_i \tag{II}$$
> が成り立つ（すなわち $0 \in \overline{\operatorname{co}}\, A + \sum_{i=1}^{m} \lambda_i \overline{\operatorname{co}}\, B_i$ を満たす $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$ が存在する）。しかし (I), (II) が同時に成立することはない。

B_i, $i = 1, 2, \ldots, m$ が有限集合であるか，あるいは

$$\max_{b_i \in B_i} \langle b_i, x \rangle < 0, \quad i = 1, 2, \ldots, m$$

を満たす $x \in X$ が存在すれば，条件 $(*)$ は満たされる。

局所リプシッツ最適化問題

f, g を X 全体で定義された実汎関数とする。まず，無制約最適化問題

$$\min_x \; f(x) \tag{UP}$$

に対して，f が $x^* \in X$ の近傍でリプシッツ連続であるとき，x^* が問題(UP)の局所的最適解ならば $0 \in \partial^\circ f(x^*)$ が成り立ち，また f が X 上で連続かつ凸であるとき，これは x^* が問題(UP)の大域的最適解であるための必要十分条件となる。

次に，不等式制約つき最適化問題

$$\min_x \; f(x) \quad \text{subject to} \quad g(x) \leq 0 \tag{CP}$$

に対しては，一般化されたFarkasの定理を用いて，以下のようにKKT型の最適性必要条件が得られる。ここで，$I(x) := \{\, i = 1, 2, \ldots, m \mid g_i(x) = 0 \,\}$ は $x \in X$ において活性な制約式の添字集合である。

> f, g は $x^* \in X$ の近傍でリプシッツ連続であり，
> $$g_i^\circ(x^*; s) \left(= \max_{z \in \partial^\circ g_i(x^*)} \langle z, s \rangle \right) < 0, \quad i \in I(x^*)$$
> を満たす $s \in X$ が存在する（g は x^* でCottle型の制約想定を満たす）とする。このとき x^* が問題(CP)の局所的最適解ならば，次式を満たす $\lambda \in R^m$ が存在する。
> $$0 \in \partial^\circ f(x^*) + \sum_{i=1}^{m} \lambda_i \partial^\circ g_i(x^*), \quad g(x^*) \leq 0, \quad \langle \lambda, g(x^*) \rangle = 0, \quad \lambda \geq 0$$

f, g の連続凸性のもとでは，上式を満たす乗数 λ の存在は x^* が問題(CP)の大域的最適解であるための必要十分条件となる。

最大値汎関数の一般勾配

X を実ヒルベルト空間，D を X の非空開部分集合，Y をハウスドルフ空間，f を $D \times Y$ 上の実汎関数とし，最大値汎関数 $v : D \to R$ を

$$v(x) := \sup_{y \in Y} \; f(x,y)$$

と定義する。いま，Y がコンパクトで，f が $D \times Y$ 上で連続ならば，最大解写像 $\widehat{Y} : D \to 2^Y$ を

$$\widehat{Y}(x) := \arg\max_{y \in Y} f(x,y) = \{\, y \in Y \mid f(x,y) = v(x) \,\}$$

により定義することができるが，最大解集合 $\widehat{Y}(x)$ は任意の $x \in D$ に対して非空コンパクト，$v : D \to R$ は D 上で連続となる。

さらに $f : D \times Y \to R$ が任意の $y \in Y$ に対し D 上でフレッシェ偏微分可能，かつ $\nabla_x f : D \times Y \to X$ が $D \times Y$ 上で連続ならば，最大値汎関数 $v : D \to R$ は D の任意の点の近傍でリプシッツ連続であり，任意の $x \in D$ における一般勾配 $\partial^\circ v(x)$ は

$$\begin{aligned}\partial^\circ v(x) &= \overline{\operatorname{co}} \{\, \nabla_x f(x,y) \mid y \in \widehat{Y}(x) \,\} \\ &= \left\{ \int_Y \nabla_x f(x,y)\, d\mu \;\middle|\; \int_Y d\mu = 1, \;\; \mu \geq 0, \;\; \int_{\widehat{Y}(x)^c} d\mu = 0 \right\}\end{aligned}$$

で与えられる（μ は最大解集合 $\widehat{Y}(x)$ 上でのみ正となる確率測度）。

この確率測度による一般勾配の表現を用いれば，実ヒルベルト空間上で定義されたmin-max問題やロバスト最適化問題など，最大値汎関数を含む最適化問題の最適性条件が得られ，またこれらを数値的に解くためのアルゴリズムを構成することができる。