一般化モーメント問題と切除平面法

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

一般化モーメント問題

コンパクトなHausdorff空間 X に対して，X 上の有限な符号つきBorel測度からなるBanach空間を $M(X)$，また X 上の有限なBorel測度からなる凸錐を $M(X)_+$ と書くことにし，以下のような $M(X)$ 上の最適化問題（一般化モーメント問題）を考える。

$$\begin{cases} \displaystyle\min_{\mu\in\mathcal{M}} \int_X f(x)\,d\mu \\ \text{subject to} \quad \displaystyle\int_X \psi(x,z)\,d\mu = h(z) \quad \forall z\in Z \end{cases}$$

ただし，$\mathcal{M}$ を $M(X)$ の凸集合，$h: Z\to\mathbb{R}$ とし，関数 $f: X\to\mathbb{R}$ および，各 $z\in Z$ に対して，関数 $\psi(\cdot,z): X\to\mathbb{R}$ は任意の $\mu\in\mathcal{M}$ に関して可積分であるとする。Z は必ずしも有限集合でなくてよい。

応用上重要なのは $\mathcal{M}$ が $\mathcal{M}\subset M(X)_+$ かつ線形不等式制約条件で規定されるとき，すなわち

$$\mathcal{M} = \left\{ \mu\in M(X)_+ \;\middle|\; \int_X \varphi(x,y)\,d\mu \ge g(y) \quad \forall y\in Y \right\}$$

のように与えられるときである。ここで，$g: Y\to\mathbb{R}$ とし，各 $y\in Y$ に対して，関数 $\varphi(\cdot,y): X\to\mathbb{R}$ は任意の $\mu\in M(X)_+$ に関して可積分であるとする。Y も有限集合であるとは限らない。このとき，一般化モーメント問題は $M(X)$ 上の等式および不等式制約条件つき線形計画問題として以下のように書ける。

$$\begin{cases} \displaystyle\min_{\mu\in M(X)} \int_X f(x)\,d\mu \\ \text{subject to} \quad \displaystyle\int_X \varphi(x,y)\,d\mu \ge g(y) \quad \forall y\in Y \\ \qquad\qquad\quad \displaystyle\int_X \psi(x,z)\,d\mu = h(z) \quad \forall z\in Z \\ \qquad\qquad\quad \mu\ge 0 \end{cases} \tag{P}$$

特に，X, Y, Z がコンパクトなHausdorff空間で，関数 f, g, h, φ, ψ がすべて連続であるとき，以下のような双対問題を考えることができる。

$$\begin{cases} \displaystyle\max_{\substack{\nu\in M(Y)\\ \xi\in M(Z)}} \int_Y g(y)\,d\nu + \int_Z h(z)\,d\xi \\ \text{subject to} \quad \displaystyle\int_Y \varphi(x,y)\,d\nu + \int_Z \psi(x,z)\,d\xi \le f(x) \quad \forall x\in X \\ \qquad\qquad\quad \nu\ge 0 \end{cases} \tag{D}$$

$M(X)$ などのBanach空間は回帰的ではないが，関数 f, g 等の連続性を仮定することにより回帰的なBanach空間に準じた双対理論を展開できる（双対問題 (D) の双対が主問題 (P) となることに注意されたい）。また，Y, Z が有限集合ならば，双対問題 (D) は半無限線形計画問題となる。

双方向切除平面法

実際にはほとんどの場合，Z は少数の点からなる有限集合

$$Z = \{z_1, z_2, \dots, z_l\}$$

であろう。問題 (P) および (D) において，X, Y をそれぞれ有限部分集合

$$X_k := \{x_1^k, x_2^k, \dots, x_{n_k}^k\}\subset X, \quad Y_k := \{y_1^k, y_2^k, \dots, y_{m_k}^k\}\subset Y$$

で緩和すると，有限次元の線形計画問題

$$\begin{cases} \displaystyle\min_{\mu\in\mathbb{R}^{n_k}} \sum_{i=1}^{n_k} f(x_i^k)\,\mu_i \\ \text{subject to} \quad \displaystyle\sum_{i=1}^{n_k} \varphi(x_i^k, y_j^k)\,\mu_i \ge g(y_j^k), \quad j=1,2,\dots,m_k \\ \qquad\qquad\quad \displaystyle\sum_{i=1}^{n_k} \psi(x_i^k, z_j)\,\mu_i = h(z_j), \quad j=1,2,\dots,l \\ \qquad\qquad\quad \mu_i\ge 0, \quad i=1,2,\dots,n_k \end{cases} \tag{P_k}$$

$$\begin{cases} \displaystyle\max_{\substack{\nu\in\mathbb{R}^{m_k}\\ \xi\in\mathbb{R}^l}} \sum_{j=1}^{m_k} g(y_j^k)\,\nu_j + \sum_{j=1}^{l} h(z_j)\,\xi_j \\ \text{subject to} \quad \displaystyle\sum_{j=1}^{m_k} \varphi(x_i^k, y_j^k)\,\nu_j + \sum_{j=1}^{l} \psi(x_i^k, z_j)\,\xi_j \le f(x_i^k), \\ \qquad\qquad\qquad\qquad\qquad\qquad i=1,2,\dots,n_k \\ \qquad\qquad\quad \nu_j\ge 0, \quad j=1,2,\dots,m_k \end{cases} \tag{D_k}$$

が得られ，このとき (P_k), (D_k) の最適解 $(\mu^k, \nu^k, \xi^k)\in\mathbb{R}^{n_k}\times\mathbb{R}^{m_k}\times\mathbb{R}^l$ に対して

$$\delta(\mu^k) := \min_{y\in Y}\left\{ \sum_{i=1}^{n_k} \varphi(x_i^k, y)\,\mu_i^k - g(y) \right\}$$

$$\gamma(\nu^k, \xi^k) := \max_{x\in X}\left\{ \sum_{j=1}^{m_k} \varphi(x, y_j^k)\,\nu_j^k + \sum_{j=1}^{l} \psi(x, z_j)\,\xi_j^k - f(x) \right\}$$

により問題 (P), (D) に対する最適性を判定，また右辺の最適解 $\bar{y}^k$, $\bar{x}^k$ を用いて X_{k+1}, Y_{k+1} を構成することにより，双方向切除平面法のアルゴリズムを考えることができる。

初期近似 X_1, Y_1 が粗すぎないこと，すなわち

$$\begin{cases} \displaystyle\sum_{i=1}^{n_1} \varphi(x_i^1, y)\,\mu_i > g(y) \quad \forall y\in Y \\ \displaystyle\sum_{i=1}^{n_1} \psi(x_i^1, z_j)\,\mu_i = h(z_j), \quad j=1,2,\dots,l \\ \displaystyle\sum_{j=1}^{m_1} \varphi(x, y_j^1)\,\nu_j + \sum_{j=1}^{l} \psi(x, z_j)\,\xi_j < f(x) \quad \forall x\in X \end{cases}$$

を満たす $(\mu,\nu,\xi)\in\mathbb{R}_+^{n_1}\times\mathbb{R}_+^{m_1}\times\mathbb{R}^l$ が存在することを仮定すると，緩和問題 (P_k), (D_k) の可解性（最適解の存在）と強双対性（最適目的関数値の一致）および問題 (P), (D) の可解性と強双対性が保証され，さらには上記の切除平面法の大域的収束性が保証される（μ^k, ν^k, ξ^k に対応する離散測度の列を考えたとき，$M(X)$, $M(Y)\times M(Z)$ において問題 (P), (D) の最適解に∗弱収束する部分列が存在する）。

同様のアルゴリズムは，目的汎関数が凸2次

$$\frac{1}{2}\int_X\int_X q(x_1,x_2)\,d\mu(x_1)\,d\mu(x_2) + \int_X f(x)\,d\mu$$

（ただし，$q: X\times X\to\mathbb{R}$ は半正定値性

$$\int_X\int_X q(x_1,x_2)\,d\mu(x_1)\,d\mu(x_2)\ge 0 \quad \forall\mu\in M(X)$$

を持つものとする）など非線形の場合に拡張することができる。