

連続最適化および関連分野に関する夏季学校 2024
数理最適化への数式処理の一応用 — 制約想定が不要な最適性条件の導出 —
解答

2024 年 8 月 24 日–26 日 庵 智幸

■解答 1. 制約付き最適化問題

$$\min_{x \in \mathcal{C}} f(x) \quad \text{subject to} \quad g(x) = 0 \quad (1)$$

の大域的最適解を x^* とする. x_k^* は $P(x; r_k)$ を最小化するので,

$$P(x_k^*, r_k) = f(x_k^*) + r_k g^\top(x_k^*) g(x_k^*) \leq f(x^*) \quad (2)$$

が成り立つ. 仮定より $r_k \rightarrow \infty$ ($k \rightarrow \infty$) であり, 一方で $f(x^*) < \infty$ であるから

$$\lim_{k \rightarrow \infty} g(x_k^*) = 0$$

が成り立つ. いま $\mathcal{C}$ はコンパクトであるから, Bolzano-Weierstrass の定理より $\{x_k^*\}_{k=1}^\infty$ には収束部分列 $\{x_{k_l}^*\}_{l=1}^\infty$ が存在する. $\bar{x} := \lim_{l \rightarrow \infty} x_{k_l}^*$ とすると, $g(\bar{x}) = 0$ であるから, (2) 式より

$$\lim_{l \rightarrow \infty} P(x_{k_l}^*, r_{k_l}) = f(\bar{x}) \leq f(x^*)$$

が成り立つ. 一方で, x^* は問題 (1) の大域最適解であるから $f(\bar{x}) = f(x^*)$ となり, したがって $\bar{x}$ も大域最適解である.

■解答 2. 閉球 $\mathcal{B}_\varepsilon(\hat{x})$ 上の最適化問題

$$\min_{x \in \mathcal{B}_\varepsilon(\hat{x})} f(x) + \|x - \hat{x}\|^2 \quad \text{subject to} \quad g(x) = 0 \quad (3)$$

について, $f(x)$ のヘシアン $H(x)$ を用いてコスト関数 $H(x) + I$ と書けることから, $\hat{x}$ は孤立局所最適解である (もともと $\hat{x}$ は (1) 式の局所最適解なので, $H(\hat{x}) \succeq 0$ であることを用いた). よって, 十分小さい $\varepsilon > 0$ を取れば $\hat{x}$ は問題 (3) の唯一の大域最適解となる.

さて, $\mathcal{C} = \mathcal{B}_\varepsilon(\hat{x})$ として問題 (3) に前問の結果を適用すると, ある単調増加列 $r_k \uparrow \infty$ に対して, $P^{\text{loc}}(x; r_k, \hat{x})$ の大域最適解 $x_k^* \in \mathcal{B}_\varepsilon(\hat{x})$ が存在し, $\lim_{k \rightarrow \infty} x_k^* = \hat{x}$ を満たす. 特に, $\hat{x}$ は $\mathcal{B}_\varepsilon(\hat{x})$ の内点であるから, ある k^* が存在し, 任意の $k \geq k^*$ に対して x_k^* は $\mathcal{B}_\varepsilon(\hat{x})$ の内点となるため, 停留条件 $P_x^{\text{loc}}(x; r_k, \hat{x}) = 0$ を満たす. あとは点列 $\{x_k^*\}_{k=k^*}^\infty$ の添字を適当に付け替えればよい.

■解答 3. ϕ の逆写像を具体的に構成することで全単射性を示す. 写像 $\chi: U_0 \rightarrow \mathbb{R}^n$ を, 任意の $p = (X_0 : \cdots : X_n) \in U_0$ に対して $\chi(p) := \left(\frac{X_1}{X_0}, \dots, \frac{X_n}{X_0} \right)$ で定義する. まず χ が well-defined であることを示す. p の異なる斉次座標 $(X'_0 : \cdots : X'_n)$ に対して, $X_i = \lambda X'_i$ ($i = 0, \dots, n$) なる

$\lambda \neq 0$ が存在する. このとき, $X_i/X_0 = (\lambda X'_i)/(\lambda X'_0) = X'_i/X'_0$ であるため, χ は well-defined である.

任意の $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ に対して,

$$(\chi \circ \phi)(x) = \chi(\phi(x)) = \chi((1 : x_1 : \dots : x_n)) = (x_1, \dots, x_n) = x$$

であり, また任意の $p = (X_0 : \dots : X_n) \in U_0$ に対して,

$$\begin{aligned} (\phi \circ \chi)(p) &= \phi(\chi(p)) = \phi\left(\frac{X_1}{X_0}, \dots, \frac{X_n}{X_0}\right) \\ &= \left(1 : \frac{X_1}{X_0} : \dots : \frac{X_n}{X_0}\right) = (X_0 : X_1 : \dots : X_n) = p \end{aligned}$$

であるから, χ は ϕ の逆写像である. よって, ϕ は全単射である.

■解答 4. ある点 $p \in \mathbb{P}^n$ の斉次座標 $(X_0 : \dots : X_n)$ について $f(X_0, \dots, X_n) = 0$ が成り立つとする. p の任意の斉次座標 $(X'_0 : \dots : X'_n)$ に対し, 定義より $X'_i = \lambda X_i$ なる $\lambda \neq 0$ が存在し,

$$f(X'_0, \dots, X'_n) = f(\lambda X_0, \dots, \lambda X_n) = \lambda^d f(X_0, \dots, X_n) = 0$$

となる. ここで d は f の全次数であり, 2 つ目の等号では f が斉次多項式であることを用いた.

■解答 5. 直線の収束の定義より, パラメータ表示 $L(v)$ および $L_i(v_i)$ であって, $\lim_{i \rightarrow \infty} v_i = v$ を満たすものが存在する. また, $q_i \in L_i$ ($i = 1, 2, \dots$) であるから, ある非零実数 t_i が存在し, $q_i = t_i v_i$ と書ける. 特に $\lim_{i \rightarrow \infty} q_i = 0$ と $v_i \neq 0$ ($i = 1, 2, \dots$) より, $t_i \neq 0$ かつ $\lim_{i \rightarrow \infty} t_i = 0$ が言える.

いま, $f(x)$ を V 上で消える多項式, すなわち $f(x) \in \mathbf{I}(V)$ とする. f の次数を d とし, 斉次成分の和として表すと,

$$f(x) = \sum_{l=d_{\min}}^d f_l(x)$$

と書ける. ここで, $d_{\min}$ は f の最小次数であり, $f_{d_{\min}} = f^{\min}$ である. $q_i = t_i v_i \in V$ より,

$$0 = \sum_{l=d_{\min}}^d f_l(t_i v_i) = \sum_{l=d_{\min}}^d t_i^l f_l(v_i)$$

が成り立つ. $t_i \neq 0$ より両辺を $t_i^{d_{\min}}$ で割ることができ,

$$0 = \sum_{l=d_{\min}}^d t_i^{l-d_{\min}} f_l(v_i) = f^{\min}(v_i) + \sum_{l=1}^{d-d_{\min}} t_i^l f_{d_{\min}+l}(v_i) \rightarrow f^{\min}(v) \quad (t_i \rightarrow 0)$$

を得る. $f^{\min}$ は斉次多項式であるから, 任意の $t \in \mathbb{R}$ について $f^{\min}(tv) = t^{d_{\min}} f^{\min}(v) = 0$ が成り立ち, したがって $L \subset C_0(V)$ が示された.

■解答 6.

$$f^{\min} \notin \langle g_1^{\min}, \dots, g_t^{\min} \rangle \quad (4)$$

となるような $f \in \mathbf{I}(V)$ であって, $\text{LT}(f^{\min})$ の順序が最も小さいものを選ぶ (項順序には最小元があるので, このような f は必ず存在する). $G_1, \dots, G_t$ が Gröbner 基底であることから, $\text{LT}(f^{\text{hom}}) \in \langle \text{LT}(G_1), \dots, \text{LT}(G_t) \rangle$ が成り立ち, 特に $\text{LT}(G_i)$ は $\text{LT}(f^{\text{hom}})$ を割り切る. したがって, ある実数 $c \in \mathbb{R}$ と単項式 $X^{\tilde{\alpha}}$ ($\tilde{\alpha} \in \mathbb{N}^{n+1}$) が存在し, $\text{LT}(f^{\min}) = cX^{\tilde{\alpha}}\text{LT}(G_i)$ と書ける. この関係式に $X_0 = 1$ および $X_i = x_i$ ($i = 1, \dots, n$) を代入すると, $\text{LT}(f^{\text{hom}}) = \text{LT}(f^{\min})X_0^{\alpha}$ および $\text{LT}(G_i) = \text{LT}(g_i^{\min})X_0^b$ と書けることから, 関係式

$$\text{LT}(f^{\min}) = cx^{\alpha}\text{LT}(g_i^{\min}) \quad (5)$$

が従う. ここで, $\alpha \in \mathbb{N}^n$ は $\tilde{\alpha}$ の第 1 成分を取り除いた指数ベクトルである.

上記の $c \in \mathbb{R}$ および x^{α} を用いて定まる多項式 $\tilde{f} = f - cx^{\alpha}g_i$ を考える. $f, g_i \in \mathbf{I}(V)$ なので, $\tilde{f} \in \mathbf{I}(V)$ である. また (4) 式より, $\tilde{f}^{\min} = f^{\min} - cx^{\alpha}g_i^{\min} \neq 0$ かつ $\tilde{f}^{\min} \notin \langle g_1^{\min}, \dots, g_t^{\min} \rangle$ が成り立つ. (もし $f^{\min} - cx^{\alpha}g_i^{\min} = 0$ ならば, $f^{\min} = cx^{\alpha}g_i^{\min}$ と書けることになり, (4) 式に反する). 一方, (5) 式より $f^{\min}$ と $cx^{\alpha}g_i^{\min}$ の先頭項同士は互いに打ち消し合うため, $\text{LT}(\tilde{f}) < \text{LT}(f)$ が成り立たねばならない. しかしながら, これは $\text{LT}(f)$ が最小であるという仮定に矛盾する. よって, 題意が示された.

■解答 7. 多項式最適化問題

$$\min_{x \in \mathbb{R}^3} \frac{1}{2} \|x\|^2 \quad \text{subject to} \quad (x_2 - 2x_1^2 + 2)^2 - (x_3 - 1)^3 = 0$$

に対応する局所ペナルティ関数は以下で書ける.

$$P^{\text{loc}}(x; r, y) = \frac{1}{2} \|x\|^2 + \|x - y\|^2 + r \left((x_2 - 2x_1^2 + 2)^2 - (x_3 - 1)^3 \right)^2$$

ここで, y は特定の局所最適解に対応する補助変数である. 停留条件は

$$\begin{aligned} F(x, y, r) &:= P_x^{\text{loc}}(x; r, y) \\ &= x + 2(x - y) + 2r \left((x_2 - 2x_1^2 + 2)^2 - (x_3 - 1)^3 \right) \begin{pmatrix} -4x_1(x_2 - 2x_1^2 + 2) \\ 2(x_2 - 2x_1^2 + 2) \\ 3(x_3 - 1)^2 \end{pmatrix} = 0 \end{aligned}$$

となる. 例えば第 1 成分 $F_1(x, y, r)$ の全次数は 8 であり, これに対して斉次化

$$F_1^{\text{hom}}(X_0, \dots, X_7) = X_0^8 F_1 \left(\frac{(X_1, X_2, X_3)^{\top}}{X_0}, \frac{(X_4, X_5, X_6)^{\top}}{X_0}, \frac{X_7}{X_0} \right)$$

を行うと、 $\mathbb{P}^7$ 上の方程式

$$F_1^{\text{hom}} = X_0^7 X_1 + 2X_0^7 (X_1 - X_4) + 2X_7 ((X_0 X_2 - 2X_1^2 + 2X_0^2)^2 - X_0 (X_3 - X_0)^3) \\ \times (-4X_1 (X_0 X_2 - 2X_1^2 + 2X_0^2)) = 0$$

が得られる。この左辺は 8 次の斉次多項式となっている。続いて非斉次化

$$\mathcal{F}_1(\rho, \xi, \eta) = F_1^{\text{hom}}(\rho, \xi_1, \dots, \xi_3, \eta_1, \dots, \eta_3, 1)$$

によって ρ - ξ - η 空間上の方程式

$$\mathcal{F}_1(\rho, \xi, \eta) = \rho^7 \xi_1 + 2\rho^7 (\xi_1 - \eta_1) \\ + 2 ((\rho \xi_2 - 2\xi_1^2 + 2\rho^2)^2 - \rho (\xi_3 - \rho)^3) (-4\xi_1 (\rho \xi_2 - 2\xi_1^2 + 2\rho^2)) = 0$$

を得る。 $F_2 = 0$ および $F_3 = 0$ にも同様の操作を行うことで、それぞれに対応する ρ - ξ - η 空間上の方程式 $\mathcal{F}_2 = 0$ および $\mathcal{F}_3 = 0$ が得られる。

得られた方程式 $\mathcal{F}(\rho, \xi, \eta)$ に対して Tangent Cone アルゴリズムを適用することで、原点における接錐 $C_0(\mathbf{V}(\mathcal{F}))$ の定義方程式 $\mathcal{G}(\rho, \xi, \eta) = 0$ を求めることができる。今回の例では、 $\mathcal{G}$ は 14 個の多項式からなっており、そのうち次数が最大のは 16 次、最小のものは 8 次の多項式である。接線の傾きを求めるため $\rho = 1$ を代入し、また収束先が一致するという条件から $\xi = \eta = x$ を代入すると、最適性必要条件 $G(x) = \mathcal{G}(1, x, x) = 0$ が得られる。今回の問題においてこの最適性条件は 5 つの実数解 (図 1: “+” の点) を持ち、そのうち 2 つが大域最適解 (図 1: “○” の点) である。

一方で、この問題の KKT 条件を計算すると

$$\begin{cases} (x_2 - 2x_1^2 + 2)^2 - (x_3 - 1)^3 = 0 \\ x + (\lambda_1 \quad \lambda_2 \quad \lambda_3) \begin{pmatrix} -4x_1(x_2 - 2x_1^2 + 2) \\ 2(x_2 - 2x_1^2 + 2) \\ 3(x_3 - 1)^2 \end{pmatrix} = 0 \end{cases}$$

となり、その解は 2 つ (図 1: “□”) のみである。つまり、本問題においては KKT 条件を解くだけでは大域最適解は求まらない。

また、FJ 条件を計算すると

$$\begin{cases} (x_2 - 2x_1^2 + 2)^2 - (x_3 - 1)^3 = 0 \\ \lambda_0 x + (\lambda_1 \quad \lambda_2 \quad \lambda_3) \begin{pmatrix} -4x_1(x_2 - 2x_1^2 + 2) \\ 2(x_2 - 2x_1^2 + 2) \\ 3(x_3 - 1)^2 \end{pmatrix} = 0 \end{cases} \quad (6)$$

となる。これは

$$x_2 - 2x_1^2 + 2 = 0 \quad \text{and} \quad x_3 - 1 = 0 \quad (7)$$

を満たす任意の x においては $\lambda_0 x = 0$ と等価である。よって、 $\lambda_0 = 0$ であるような任意の $(\lambda_0, \lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^4$ は (6) 式を満たすこととなり、(7) 式を満たす点全体 (図 1: オレンジの曲線)

が FJ 条件を満たしてしまう。

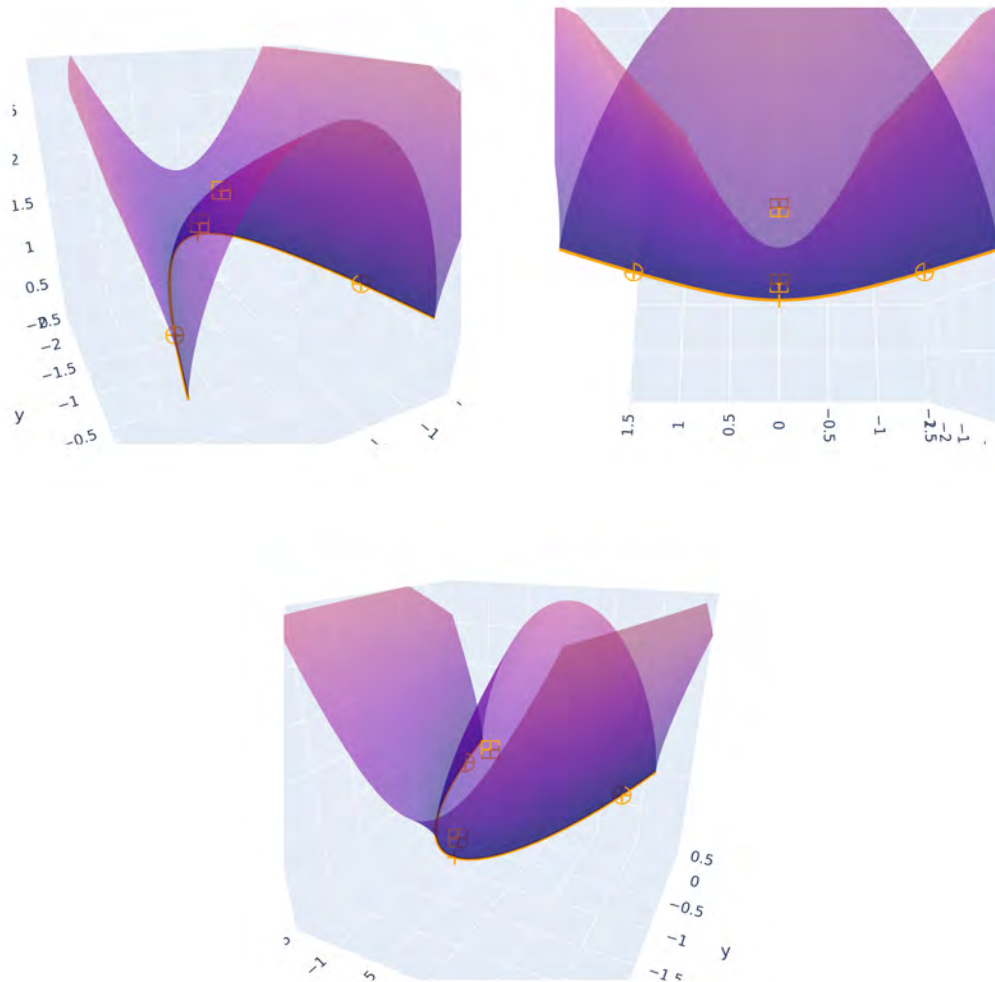


図1 実行可能集合と最適解候補