

「作用素スケーリング入門」演習問題

2024 年 8 月 6 日

問題 1. 非負行列 $P, A \in \mathbb{R}_+^{m \times n}$ に対する KL ダイバージェンスを以下のように定める.

$$D_{KL}(P : A) = \sum_{i,j} \left(p_{ij} \log \frac{p_{ij}}{a_{ij}} - p_{ij} + a_{ij} \right).$$

ただし, $0 \log \frac{0}{0} = 0$ とし, $a_{ij} = 0$ かつ $p_{ij} \neq 0$ を満たす (i, j) が存在するときは $D_{KL}(p : a) = +\infty$ と定める. 非負ベクトル $r \in \mathbb{R}_+^m, c \in \mathbb{R}_+^n$ に対し, KL ダイバージェンス最小化問題

$$\inf_P D_{KL}(P : A) \quad \text{subject to} \quad P \geq 0, \quad P\mathbf{1}_n = r, \quad P^\top \mathbf{1}_m = c$$

の Lagrange 双対問題が

$$\sup_{x \in \mathbb{R}^m, y \in \mathbb{R}^n} - \sum_{i,j} a_{ij} e^{x_i + y_j} + \sum_i r_i x_i + \sum_j c_j y_j$$

であることを示せ.

問題 2 (Csiszár, 1975). 非負行列 $A \in \mathbb{R}^{m \times n}$ と非負ベクトル $r \in \mathbb{R}_+^m, c \in \mathbb{R}_+^n$ ($\sum_i r_i = \sum_j c_j$) に対する Sinkhorn 反復 $A^{(t)}$ が以下を満たすことを示せ.

$$\begin{aligned} A^{(2t+1)} &= \operatorname{argmin}_{P \geq 0} \{ D_{KL}(P : A^{(2t)}) : P\mathbf{1}_n = r \}, \\ A^{(2t+2)} &= \operatorname{argmin}_{P \geq 0} \{ D_{KL}(P : A^{(2t+1)}) : P^\top \mathbf{1}_m = c \}. \end{aligned}$$

問題 3 (Sinkhorn, 1964). $n \times n$ 正行列 A に対して, 正対角行列 L, R で $(LAR)\mathbf{1} = \mathbf{1}$, $(LAR)^\top \mathbf{1} = \mathbf{1}$ となるものが存在することを示せ. (ヒント: 正行列と $r = c = \mathbf{1}$ に対して, KL ダイバージェンス最小化問題は Slater 条件を満たす)

問題 4. $n \times n$ 非負行列 A に対して, CP 写像 $\Phi : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$ で, $\Phi(I) = \operatorname{Diag}(A\mathbf{1})$, $\Phi^*(I) = \operatorname{Diag}(A^\top \mathbf{1})$ となるものを構成せよ.

以下, $n \times n$ 正定値エルミート行列の全体を $\operatorname{PD}(n)$ と書く.

問題 5. $X \in \operatorname{PD}(n)$ に対して定義される関数 $f(X) = \log \det X$ について, $\nabla f(X) = X^{-1}$ であることを示せ.

問題 6 (フレームスケーリング (Barthe, 1998)). ベクトル $u_1, \dots, u_m \in \mathbb{C}^n$ と $\alpha_1, \dots, \alpha_m \in \mathbb{R}_+$ に対して, 次の問題を考える.

$$\text{正則行列 } P \in \mathbb{C}^{n \times n} \text{ で } \sum_{\ell=1}^m \alpha_\ell \frac{(Pu_\ell)(Pu_\ell)^\dagger}{\|Pu_\ell\|_2^2} = I_n \text{ となるものは存在するか?}$$

この問題を周辺分布つき作用素スケーリングとして定式化せよ.

問題 7. $\Phi : \mathbb{C}^{n_2 \times n_2} \rightarrow \mathbb{C}^{n_1 \times n_1}$ を CP 写像とし, $p \in \mathbb{R}_+^{n_1}, q \in \mathbb{R}_+^{n_2}$ ($p_1 \geq \dots \geq p_{n_1}, q_1 \geq \dots \geq q_{n_2}$) とする. 以下の関数 $f : \text{PD}(n_1) \times \text{PD}(n_2) \rightarrow \mathbb{R}$ を考える.

$$f(X, Y) = X \bullet \Phi(Y) - \sum_{i=1}^{n_1} (p_i - p_{i+1}) \log \Delta_i(X) - \sum_{j=1}^{n_2} (q_j - q_{j+1}) \log \Delta_j(Y)$$

ここで, $\Delta_i(X)$ は X の i 番目の首座小行列式. また $p_{n_1+1} := 0, q_{n_2+1} := 0$.

(1)

$$\nabla_X f(X, Y) = \Phi(Y) - \sum_{i=1}^{n_1} (p_i - p_{i+1}) E_i^\top X_i^{-1} E_i$$

を示せ. ここで, X_i は X の i 番目の首座小行列, E_i は $i \times n_1$ 行列で,

$$E_i = \begin{bmatrix} \overset{i}{1} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \begin{bmatrix} \overset{n-i}{O} \\ \\ \\ \end{bmatrix}$$

である.

(2) $\nabla_X f(X, Y) = O$ とする. $X = L^\dagger L, Y = R^\dagger R$ とコレスキー分解したとき, $\Phi_{L,R}(I) = \text{Diag}(p)$ であることを示せ.

問題 8. Brascamp–Lieb 不等式の $(\mathbf{B}, p)$, f_ℓ を適当に設定することで, 次の不等式を示せ. ただし, これらの不等式における BL 定数が 1 以下であることは認めて良い. また, $\mathbf{B}$ に対応する BL 多面体を求めよ.

(1) (Hölder の不等式) 可測関数 $f, g : \mathbb{R}^n \rightarrow \mathbb{R}, p, q \in \mathbb{R}_+ (1/p + 1/q = 1)$ に対し,

$$\int_{\mathbb{R}^n} |f(x)g(x)| dx \leq \left(\int_{\mathbb{R}^n} |f(x)|^p dx \right)^{1/p} \left(\int_{\mathbb{R}^n} |g(x)|^q dx \right)^{1/q}$$

(2) (Loomis–Whitney の不等式の特殊ケース) 可測集合 $E \subseteq \mathbb{R}^3$ に対し, E_1, E_2, E_3 をそれぞれ E の xy, yz, zx 平面への射影とする. このとき,

$$\text{vol}(E) \leq \sqrt{\text{area}(E_1) \text{area}(E_2) \text{area}(E_3)}.$$

(ここで, vol, area はそれぞれ体積と面積を表す)

(3) (物足りない人向けの問) 上の 2 つの間で用いた $(\mathbf{B}, p)$ に対応する BL 定数が 1 以下であることを示せ.

問題 9. (1) (Barthe (1998)) ベクトル $u_1, \dots, u_m \in \mathbb{R}^n$ が $\langle u_1, \dots, u_m \rangle = \mathbb{R}^n$ を満たすとする. ランク 1 行列 $B_\ell = u_\ell^\top \in \mathbb{R}^{1 \times n} (\ell = 1, \dots, m)$ に関する BL 多面体が, u_ℓ に関する線形マトロイドの基多面体であることを示せ. ただし, 線形マトロイドの基多面体が, 次

の線形不等式系で表されることを用いて良い.

$$\begin{aligned}\sum_{\ell \in S} x_\ell &\leq \dim \langle u_\ell : \ell \in S \rangle \quad (S \subseteq [m]) \\ \sum_{\ell=1}^m x_\ell &= \dim \langle u_1, \dots, u_m \rangle \\ x &\geq \mathbf{0}\end{aligned}$$

(2) (Garg, Gurvits, Oliveira, and Wigderson (2018)) ベクトル $u_1, \dots, u_m, v_1, \dots, v_m \in \mathbb{R}^n$ が $\langle u_1, \dots, u_m \rangle = \langle v_1, \dots, v_m \rangle = \mathbb{R}^n$ を満たすとする. ランク 2 のブロック行列

$$B_\ell = \begin{bmatrix} u_\ell^\top & \\ & v_\ell^\top \end{bmatrix} \in \mathbb{R}^{2 \times 2n} \quad (\ell = 1, \dots, m)$$

に関する BL 多面体が, u_ℓ, v_ℓ に関する 2 つの線形マトロイドの共通基多面体であることを示せ. ここで, 共通基多面体が次の線形不等式系で表されることを用いて良い.

$$\begin{aligned}\sum_{\ell \in S_1} x_\ell &\leq \dim \langle u_\ell : \ell \in S_1 \rangle \quad (S_1 \subseteq [m]) \\ \sum_{\ell \in S_2} x_\ell &\leq \dim \langle v_\ell : \ell \in S_2 \rangle \quad (S_2 \subseteq [m]) \\ \sum_{\ell=1}^m x_\ell &= \dim \langle u_1, \dots, u_m \rangle = \dim \langle v_1, \dots, v_m \rangle \\ x &\geq \mathbf{0}\end{aligned}$$

以上

補足

以下, X は $n \times n$ エルミート行列とする.

- X はユニタリ行列 U を用いて

$$X = U \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} U^\dagger$$

と対角化できる (**固有分解**). ここで, $\lambda_1 \geq \dots \geq \lambda_n$ は X の固有値で, $\lambda_i(X)$ と書く. また, U の列ベクトルを u_i とすると,

$$X = \sum_{i=1}^n \lambda_i u_i u_i^\dagger$$

と書ける.

- X が**正定値** ($X \succ O$ と書く) $\stackrel{\text{def}}{\iff} v^\top X v > 0 \ (v \in \mathbb{C}^n)$
 $\iff X$ の固有値が全て正
 $\iff$ 正則な上三角行列 C を用いて $X = C^\dagger C$ (**コレスキー分解**). この C は, C の対角成分を非負実数に制限すると, 一意である. $\text{Chol}(X)$ と書く.
 $\iff$ ある正定値エルミート行列 M に対して $X = M^2$. この M は一意. $X^{1/2}$ と書く (**行列の平方根**).
- X の**トレース**: $\text{tr } X = \sum_{i=1}^m x_{ii} = \sum_{i=1}^n \lambda_i(X)$
- $\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$
- エルミート行列 A, B に対し, $A \bullet B := \text{tr}(AB)$ はエルミート内積を定める (**トレース内積**).
- $\mathbb{C}^n$ の部分空間 U に対し, $n \times n$ 半正定値エルミート行列 Π_U で, $\text{im } \Pi_U = U, \Pi_U^2 = \Pi_U$ を満たすものが一意に存在する (**直交射影行列**).

References

- F. Barthe (1998). “On a reverse form of the Brascamp-Lieb inequality”. *Inventiones Mathematicae* 134.2, pp. 335–361.
- I. Csiszár (1975). “ I -divergence geometry of probability distributions and minimization problems”. *The Annals of Probability* 3, pp. 146–158.
- A. Garg, L. Gurvits, R. Oliveira, and A. Wigderson (2018). “Algorithmic and optimization aspects of Brascamp-Lieb inequalities, via Operator Scaling”. *Geometric and Functional Analysis* 28.1, pp. 100–145.
- R. Sinkhorn (1964). “A relationship between arbitrary positive matrices and doubly stochastic matrices”. *The Annals of Mathematical Statistics* 35.2, pp. 876–879.