2020年8月12日

量子情報数理特論

(第14回) 量子 f -ダイバージェンス, 量子 Stein の補題

電気通信大学 大学院情報理工学研究科

小川朋宏

13 量子 f -ダイバージェンス, 量子 Stein の補題

(最終レポート, その2) 今週中に Slack に出題します. 締切は 2020 年 9 月 7 日の予定です.

13.1 前回までの復習

Lemma 1 (第12回, Lemma 2). 任意のエルミート作用素 A について,

$$A_+ = A\{A > 0\} \quad (1)$$

$$\text{Tr } A_+ = \text{Tr } A\{A > 0\} = \max_{0 \leq T \leq I} \text{Tr } AT \quad (2)$$

Lemma 2 (第12回, Lemma 6). 任意のエルミート作用素 A, B について,

$$\min_{0 \leq T \leq I} \{\text{Tr } A(I - T) + \text{Tr } BT\} = \text{Tr } A - \text{Tr}(A - B)_+ \quad (3)$$

ここで, 最小化は $S := \{A - B > 0\}$ で達成される.

Lemma 3. 上記を $\rho_n, \sigma_n \in \mathcal{S}(\mathcal{H}_n)$ の仮説検定問題に適用

$$\min_{0 \leq T_n \leq I_n} \{\alpha_n(T_n) + e^{na} \beta_n(T_n)\} = \min_{0 \leq T_n \leq I_n} \{\text{Tr } \rho_n (I_n - T_n) + e^{na} \text{Tr } \sigma_n T_n\} \quad (4)$$

$$= 1 - \text{Tr}(\rho_n - e^{na} \sigma_n)_+ \quad (5)$$

最小化は $S_n(a) := \{\rho_n - e^{na} \sigma_n > 0\}$ (量子 Neyman-Pearson 検定) で達成される.

Definition 1.

$$\psi(s) := \log \operatorname{Tr} \rho^s \sigma^{1-s} \quad (6)$$

$$\varphi(a) := \max_{0 \leq s \leq 1} \{as - \psi(s)\} \quad (7)$$

Lemma 4 (前回, Lemma 7). 任意の $a \in \mathbb{R}$ と任意の $n \in \mathbb{N}$ について,

$$\frac{1}{n} \log \alpha_n(S_n(a)) \leq -\{\varphi(a) - a\} \quad (8)$$

$$\frac{1}{n} \log \beta_n(S_n(a)) \leq -\varphi(a) \quad (9)$$

13.2 量子 Stein の補題 : direct part の証明

Lemma 5. 任意の $a < D(\rho||\sigma)$ について,

$$\alpha_n(S_n(a)) \xrightarrow{n \rightarrow \infty} 0 \quad (\text{Tr } \rho_n S_n(a) \xrightarrow{n \rightarrow \infty} 1) \quad (10)$$

(証明) $\psi(s), \varphi(a)$ のグラフより, $a < D(\rho||\sigma)$ ならば $\varphi(a) - a > 0$ である.
よって, (8) より題意が示される. □

Lemma 6. 任意の $a \in \mathbb{R}$ と任意の $n \in \mathbb{N}$ について,

$$\beta_n(S_n(a)) = \text{Tr } \sigma_n S_n(a) \leq e^{-na} \text{Tr } \rho_n S_n(a) \leq e^{-na} \quad (11)$$

(証明) 定義より,

$$\text{Tr}(\rho_n - e^{na} \sigma_n) S_n(a) = \text{Tr}(\rho_n - e^{na} \sigma_n) \{\rho_n - e^{na} \sigma_n > 0\} \geq 0 \quad (12)$$

であるから, 移項して,

$$\text{Tr } \rho_n S_n(a) \geq e^{na} \text{Tr } \sigma_n S_n(a) \quad (13)$$

両辺を e^{na} で割ればよい. □

Definition 2 (復習).

$$\beta_n^*(\varepsilon) = \min \{ \beta_n(T_n) \mid T_n : \text{test}, \alpha_n(T_n) \leq \varepsilon \} \quad (14)$$

Theorem 1 (量子Steinの補題). 任意の $0 < \varepsilon < 1$ について,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \beta_n^*(\varepsilon) = -D(\rho \parallel \sigma) \quad (15)$$

Theorem 2 (direct part). 任意の $0 < \varepsilon \leq 1$ について,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \leq -D(\rho \parallel \sigma) \quad (16)$$

(証明) 任意の $a < D(\rho \parallel \sigma)$ について, Lemma 5より, 十分大きなすべての n について, $\alpha_n(S_n(a)) \leq \varepsilon$ よって, $\beta_n^*(\varepsilon)$ の定義より,

$$\beta_n^*(\varepsilon) \leq \alpha_n(S_n(a)) \leq e^{-na} \quad (17)$$

ただし, 最後の不等式はLemma 6を用いた. これより,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \leq -a \quad (18)$$

$a < D(\rho \parallel \sigma)$ は任意であったから, $a \nearrow D(\rho \parallel \sigma)$ の極限をとると, 題意が示される. $\square$

13.3 作用素単調関数

Definition 3. 区間 $J \subseteq \mathbb{R}$ における実関数 $f : J \subseteq \mathbb{R} \rightarrow \mathbb{R}$ を考える.

(固有値が関数 f の定義域 J に含まれている) 任意のエルミート作用素 A, B に対して,

$$A \leq B \implies f(A) \leq f(B) \quad (19)$$

が成り立つとき, f は作用素単調関数 (operator monotone function) であるという.

Example 1.

- $f(t) = \log t$ は作用素単調関数
- $f(t) = t^s$ ($s \in [0, 1]$) は作用素単調関数
- $f(t) = t^s$ ($s \notin [0, 1]$) は作用素単調関数ではない
- $f(t) = -1/t$ は作用素単調関数 (インバースは作用素逆単調)

13.4 作用素凸関数

Definition 4. 区間 $J \subseteq \mathbb{R}$ における実関数 $f : J \subseteq \mathbb{R} \rightarrow \mathbb{R}$ を考える.

(固有値が関数 f の定義域 J に含まれている) 任意のエルミート作用素 A, B に対して,

$$tf(A) + (1 - t)f(B) \geq f(tA + (1 - t)B) \quad (0 \leq \forall t \leq 1) \quad (20)$$

が成り立つとき, f は作用素凸関数 (operator convex function) であるという.

Example 2.

- $f(t) = -\log t$ は作用素凸関数
- $f(t) = t \log t$ は作用素凸関数
- $f(t) = -t^s$ ($s \in [0, 1]$) は作用素凸関数
- $f(t) = t^s$ ($s \in [-1, 0] \cup [1, 2]$) は作用素凸関数
- $f(t) = t^s$ ($s \notin [-1, 0] \cup [1, 2]$) は作用素凸関数ではない
- $f(t) = 1/t$ は作用素凸関数 (インバースは作用素凸関数)

13.5 相対モジュラーオペレーター

Definition 5 (左作用, 右作用).

$$L_A : X \in \mathcal{L}(\mathcal{H}) \mapsto AX \in \mathcal{L}(\mathcal{H}) \quad (\text{左掛け算作用}) \quad (21)$$

$$R_B : A \in \mathcal{L}(\mathcal{H}) \mapsto AB \in \mathcal{L}(\mathcal{H}) \quad (\text{右掛け算作用}) \quad (22)$$

任意の $X \in \mathcal{L}(\mathcal{H})$ に対して, $L_A R_B X = AXB = R_B L_A X$ であるから, L_A と R_B は可換である.

Definition 6. 相対モジュラーオペレーター (relative modular operator) を

$\Delta_{A,B} := L_A R_{B^{-1}}$ で定義する. すなわち,

$$\Delta_{A,B} : X \in \mathcal{L}(\mathcal{H}) \mapsto \Delta_{A,B}(X) = AXB^{-1} \in \mathcal{L}(\mathcal{H}) \quad (23)$$

簡単のため, $\Delta = \Delta_{A,B}$ と書くこともある.

Lemma 7.

- (i) $A \geq 0$ のとき, $L_A \geq 0$
- (ii) $B \geq 0$ のとき, $R_B \geq 0$
- (iii) $A, B \geq 0$ のとき, $\Delta_{A,B} \geq 0$ であり, 特に $\Delta_{A,B}$ はエルミート

(証明) 任意の $X \in \mathcal{L}(\mathcal{H})$ について次式が成り立つことから, (i) (ii) が成立.

$$\langle\langle X, L_A X \rangle\rangle = \text{Tr}[X^* A X] = \text{Tr}[(\sqrt{A} X)^* (\sqrt{A} X)] \geq 0$$

$$\langle\langle X, R_A X \rangle\rangle = \text{Tr}[X^* X A] = \text{Tr}[(X \sqrt{A})^* (X \sqrt{A})] \geq 0$$

(iii) L_A は $R_{B^{-1}}$ は半正定値で可換であるから, $\Delta_{A,B} = L_A R_{B^{-1}} = (L_A^{1/2} R_{B^{-1}}^{1/2})^2 \geq 0$ □

Remark 1 (別の考え方). 密度作用素 $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ に対して,

$$\langle X, Y \rangle_\rho^- := \text{Tr} \rho X^* Y, \quad \langle X, Y \rangle_\sigma^+ := \text{Tr} \sigma Y X^* \quad (24)$$

とおくと, これらは $\mathcal{L}(\mathcal{H})$ 上の半双線型形式 (sesquilinear form) で, $\rho > 0, \sigma > 0$ のとき内積になる. 相対モジュラーオペレーター $\Delta_{\sigma,\rho}$ は次式を満たす作用素として定義してもよい.

$$\langle X, \Delta_{\sigma,\rho} Y \rangle_\rho^- = \text{Tr}[\rho \cdot X^* \cdot \sigma Y \rho^{-1}] = \text{Tr} \sigma Y X^* = \langle X, Y \rangle_\sigma^+ \quad (25)$$

上式で $X = Y$ とすれば,

$$\langle X, \Delta_{\sigma,\rho} X \rangle_\rho^- = \langle X, X \rangle_\sigma^+ \geq 0 \quad (26)$$

であるから, $\Delta_{\sigma,\rho} \geq 0$ が分かる. 特に, $\Delta_{\sigma,\rho}$ はエルミート作用素である.

13.6 量子 f -ダイバージェンス

Definition 7.

$$D_f^X(\rho||\sigma) := \langle X, f(\Delta)(X) \rangle_{\rho}^{-} \quad (X \in \mathcal{L}(\mathcal{H})) \quad (27)$$

$$D_f(\rho||\sigma) := D_f^I(\rho||\sigma) = \langle I, f(\Delta)(I) \rangle_{\rho}^{-} = \text{Tr}[\rho \cdot f(\Delta_{\sigma, \rho})(I)] = \left\langle\left\langle \rho^{1/2}, f(\Delta_{\sigma, \rho}) \rho^{1/2} \right\rangle\right\rangle \quad (28)$$

密度作用素 $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ に対して、スペクトル分解

$$\rho = \sum_i a_i E_i, \quad \sigma = \sum_j b_j F_j \quad (29)$$

を考える。(簡単のため $\rho, \sigma > 0$ を仮定)

$$\Delta(F_j E_i) = \sigma F_j E_i \rho^{-1} = \frac{b_j}{a_i} F_j E_i \quad (30)$$

であるから、 $\Delta_{\sigma, \rho}$ の固有値は $\frac{b_j}{a_i}$ で、対応する固有ベクトルは $F_j E_i$ である。

$$\Delta^n(F_j E_i) = \left(\frac{b_j}{a_i}\right)^n F_j E_i \quad (n = 0, 1, 2, \dots) \quad (31)$$

であるから、関数 $f(t)$ のテーラー展開を考えれば、

$$f(\Delta)(F_j E_i) = f\left(\frac{b_j}{a_i}\right) F_j E_i \quad (32)$$

が成り立つ。これより,

$$f(\Delta)(I) = f(\Delta) \left(\sum_i \sum_j F_j E_i \right) = \sum_i \sum_j f(\Delta)(F_j E_i) = \sum_i \sum_j f\left(\frac{b_j}{a_i}\right) (F_j E_i) \quad (33)$$

$$D_f(\rho||\sigma) = \text{Tr}[\rho \cdot f(\Delta)(I)] = \sum_i \sum_j a_i f\left(\frac{b_j}{a_i}\right) \text{Tr} F_j E_i = \sum_i \sum_j (a_i \text{Tr} E_i F_j) \cdot f\left(\frac{b_j \text{Tr} E_i F_j}{a_i \text{Tr} E_i F_j}\right) \quad (34)$$

ここで,

$$p(i, j) = a_i \text{Tr} E_i F_j, \quad q(i, j) = b_i \text{Tr} E_i F_j \quad (35)$$

とおけば, これらは確率分布である。量子 f -ダイバージェンスは, これらの古典 f -ダイバージェンスでもある。

$$D_f(\rho||\sigma) = D_f(p||q) \quad (36)$$

Theorem 3 (単調性). $f(t)$ が作用素凸関数であるとき, 任意の CPTP 写像 $\mathcal{E}$ について,

$$D_f(\rho||\sigma) \geq D_f(\mathcal{E}(\rho)||\mathcal{E}(\sigma)) \quad (37)$$

(証明は省略)

13.7 量子 f -ダイバージェンスの例

Lemma 8. エルミート作用素 $A, B \in \mathcal{L}(\mathcal{H})$ に対して,

$$f(L_A) = L_{f(A)}, \quad f(R_B) = R_{f(B)} \quad (38)$$

(証明) 左作用 L_A について (38) を示す (右作用 R_A の場合も同様). 任意の $X \in \mathcal{L}(\mathcal{H})$ について,

$$(L_A)^n X = A^n X = L_{A^n} X \quad (n = 0, 1, 2, \dots) \quad (39)$$

であるから, $f(t) = t^n$ ($n = 0, 1, 2, \dots$) のとき (38) が成り立つ. 線形性

$$L_{aA+bB} = aL_A + bL_B \quad (a, b \in \mathbb{C}, A, B \in \mathcal{L}(\mathcal{H})) \quad (40)$$

が成り立つので, $f(t)$ が多項式の場合も (38) が成立. 一般の場合, $f(t)$ のテーラー展開を考えればよい.

Example 3. $f(t) = -\log t$ のときは, 量子相対エントロピー $D(\rho||\sigma)$ である.

左作用と右作用の可換性から,

$$-\log \Delta = -\log L_\sigma R_{\rho^{-1}} = -(\log L_\sigma + \log R_{\rho^{-1}}) \stackrel{(38)}{=} -(L_{\log \sigma} + R_{\log \rho^{-1}})$$

よって,

$$D_f(\rho||\sigma) = \text{Tr } \rho (-\log \Delta)(I) = -\text{Tr } \rho (\log \sigma + \log \rho^{-1}) = \text{Tr } \rho (\log \rho - \log \sigma) = D(\rho||\sigma)$$

Example 4. $f(t) = t \log t$ のときは, 量子相対エントロピー $D(\sigma||\rho)$ である.

$$\Delta \log \Delta = (L_\sigma R_{\rho^{-1}}) \log(L_\sigma R_{\rho^{-1}}) = (L_\sigma R_{\rho^{-1}})(\log L_\sigma + \log R_{\rho^{-1}}) \stackrel{(38)}{=} (L_\sigma R_{\rho^{-1}})(L_{\log \sigma} + R_{\log \rho^{-1}})$$

であるから,

$$(\Delta \log \Delta)(I) = (L_\sigma R_{\rho^{-1}})(\log \sigma + \log \rho^{-1}) = \sigma(\log \sigma - \log \rho)\rho^{-1}$$

よって,

$$D_f(\rho||\sigma) = \text{Tr} \rho (\Delta \log \Delta)(I) = \text{Tr} [\rho \cdot \sigma(\log \sigma - \log \rho)\rho^{-1}] = \text{Tr} \sigma(\log \sigma - \log \rho) = D(\sigma||\rho)$$

Example 5. $f(t) = t^{1-s}$ のとき,

$$D_f(\rho||\sigma) = \text{Tr} \rho (\Delta^{1-s})(I) = \text{Tr} \rho \sigma^{1-s} \rho^{-(1-s)} = \text{Tr} \rho^s \sigma^{1-s} \quad (41)$$

Lemma 9 (単調性の系). 任意の CPTP 写像について,

$$\text{Tr} \rho^s \sigma^{1-s} \leq \text{Tr} \mathcal{E}(\rho)^s \mathcal{E}(\sigma)^{1-s} \quad (s \in [0, 1]) \quad (42)$$

$$\text{Tr} \rho^s \sigma^{1-s} \geq \text{Tr} \mathcal{E}(\rho)^s \mathcal{E}(\sigma)^{1-s} \quad (s \in [-1, 0] \cup [1, 2]) \quad (43)$$

13.8 量子Steinの補題 : converse part

Theorem 4 (量子Steinの補題). 任意の $0 < \varepsilon < 1$ について,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \beta_n^*(\varepsilon) = -D(\rho \parallel \sigma) \quad (44)$$

Theorem 5 (converse part). 任意の $0 \leq \varepsilon < 1$ について,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \geq -D(\rho \parallel \sigma) \quad (45)$$

を証明する.

$\beta_n^*(\varepsilon)$ を達成する検定 $0 \leq T_n \leq I_n$ について, $\alpha_n(T_n) \leq \varepsilon$ であるから, 以下を示せばよい.

Theorem 6. 任意の検定列 $\{T_n\}_{n=1}^{\infty}$ ($0 \leq T_n \leq I_n$) と, 任意の $0 \leq \varepsilon < 1$ について,

$$\alpha_n(T_n) \leq \varepsilon \implies \liminf_{n \rightarrow \infty} \frac{1}{n} \log \beta_n(T_n) \geq -D(\rho \parallel \sigma) \quad (46)$$

(証明) 任意の $1 \leq s \leq 2$ について, 単調性 Lemma 9 より,

$$\mathrm{Tr} \rho_n^s \sigma_n^{1-s} \geq \{\mathrm{Tr} \rho_n T_n\}^s \{\mathrm{Tr} \sigma_n T_n\}^{1-s} + \{\mathrm{Tr} \rho_n (I_n - T_n)\}^s \{\mathrm{Tr} \sigma_n (I_n - T_n)\}^{1-s} \quad (47)$$

$$\geq \{\mathrm{Tr} \rho_n T_n\}^s \{\mathrm{Tr} \sigma_n T_n\}^{1-s} \quad (48)$$

$$\geq (1 - \varepsilon)^s \{\mathrm{Tr} \sigma_n T_n\}^{1-s} \quad (49)$$

よって,

$$s \log(1 - \varepsilon) + (1 - s) \log \mathrm{Tr} \sigma_n T_n \leq \log \mathrm{Tr} \rho_n^s \sigma_n^{1-s} = \log \mathrm{Tr} (\rho^s \sigma^{1-s})^n = n \cdot \psi(s) \quad (50)$$

$1 < s \leq 2$ について, $1 - s < 0$ と $\psi(1) = 0$ より,

$$\frac{1}{n} \cdot \frac{s}{1-s} \log(1 - \varepsilon) + \frac{1}{n} \log \mathrm{Tr} \sigma_n T_n \geq -\frac{\psi(s) - \psi(1)}{s-1} \quad (51)$$

両辺の $\liminf_{n \rightarrow \infty}$ をとることで,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathrm{Tr} \sigma_n T_n \geq -\frac{\psi(s) - \psi(1)}{s-1} \quad (52)$$

$s \searrow 1$ の極限 (右極限) をとることで,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathrm{Tr} \sigma_n T_n \geq -\lim_{s \searrow 1} \frac{\psi(s) - \psi(1)}{s-1} = -D(\rho \parallel \sigma) \quad (53)$$

□

13.9 情報スペクトルの極限定理（その1）

Theorem 7.

$$\mathrm{Tr} \rho_n S_n(a) = \mathrm{Tr} \rho^{\otimes n} \{ \rho^{\otimes n} - e^{na} \sigma^{\otimes n} > 0 \} \xrightarrow{n \rightarrow \infty} \begin{cases} 1 & a < D(\rho || \sigma) \\ 0 & a > D(\rho || \sigma) \end{cases} \quad (54)$$

$a < D(\rho || \sigma)$ の場合は Lemma 5 で示した.

Lemma 10. 任意の $a > D(\rho || \sigma)$ について,

$$\alpha_n(S_n(a)) \xrightarrow{n \rightarrow \infty} 1 \quad (\mathrm{Tr} \rho_n S_n(a) \xrightarrow{n \rightarrow \infty} 0) \quad (55)$$

(証明) 任意の $1 \leq s \leq 2$ について,

$$\begin{aligned} \mathrm{Tr} \rho_n S_n(a) &= \{\mathrm{Tr} \rho_n S_n(a)\}^s \{\mathrm{Tr} \rho_n S_n(a)\}^{1-s} \\ &\leq e^{na(1-s)} \{\mathrm{Tr} \rho_n S_n(a)\}^s \{\mathrm{Tr} \sigma_n S_n(a)\}^{1-s} \quad (\because 1-s \leq 0) \\ &\leq e^{na(1-s)} [\{\mathrm{Tr} \rho_n S_n(a)\}^s \{\mathrm{Tr} \sigma_n S_n(a)\}^{1-s} + \{\mathrm{Tr} \rho_n (I_n - S_n(a))\}^s \{\mathrm{Tr} \sigma_n (I_n - S_n(a))\}^{1-s}] \\ &\leq e^{na(1-s)} \mathrm{Tr} \rho_n^s \sigma_n^{1-s} \end{aligned} \quad (56)$$

$$\leq e^{na(1-s)} (\mathrm{Tr} \rho^s \sigma^{1-s})^n \quad (57)$$

ただし, (56)の不等式では単調性Lemma 9を用いた. よって, $\psi(s) = \log \rho^s \sigma^{1-s}$ とおけば,

$$\mathrm{Tr} \rho_n S_n(a) \leq \exp[-n \cdot \max_{1 \leq s \leq 2} \{a(s-1) - \psi(s)\}] \quad (58)$$

$\psi'(1) = D(\rho||\sigma)$ であるから,

$$a > D(\rho||\sigma) \implies \max_{1 \leq s \leq 2} \{a(s-1) - \psi(s)\} > 0 \quad (59)$$

となって, $\mathrm{Tr} \rho_n S_n(a) = 1 - \alpha_n(S_n(a)) \xrightarrow{n \rightarrow \infty} 0$ が示される. □