2021年7月14日

量子情報数理特論
（第13回）量子Steinの補題：順定理

電気通信大学 大学院情報理工学研究科

小川朋宏

12　量子Steinの補題：順定理

(課題：中間レポート) 量子Neyman-Pearson検定$S_n(a)$ $(a \in \mathbb{R})$の誤り確率は漸近的に指数減衰し，
それらの挙動は次式で与えられる．

$$\alpha_n(S_n(a)) \simeq e^{-n\{\varphi(a)-a\}}, \quad \beta_n(S_n(a)) \simeq e^{-n\varphi(a)} \tag{1}$$

ただし，$\varphi(a)$は$\psi(s)$のLegendre変換である．

$$\psi(s) := \log \mathrm{Tr}\, \rho^s \sigma^{1-s}, \quad \varphi(a) := \max_{0\le s\le 1} \{as - \psi(s)\} \tag{2}$$

1. 横軸をs軸として，関数$\psi(s)$のグラフをを描け．その際，以下の特徴を明示せよ．
   (1) 厳密に凸関数であること
   (2) $\psi(0)$, $\psi(1)$の値
   (3) $s=0$, $s=1$における接線の傾き
   (4) $\varphi(a)$および$\varphi(a)-a$の値
2. 横軸をa軸として適切な範囲を設定し，
   $\varphi(a)$および$\varphi(a)-a$のグラフを（ひとつのグラフ中に）を描け．

- 提出方法：手書きでグラフを描いて，写真(jpeg, pdf)による提出を推奨します（iPad等でもOK）
- 締切：2021年7月30日（本日から1週間以内を推奨）

12.1　量子仮説検定の漸近理論（復習）

漸近論の設定

- 量子状態 $\rho \in \mathcal{S}(\mathcal{H})$ または $\sigma \in \mathcal{S}(\mathcal{H})$ のどちらかに準備された系を，（同一の準備状況で）n 回利用できるとする．
- この状況で得られる量子状態は，Hilbert 空間 $\mathcal{H}^{\otimes n} = \underbrace{\mathcal{H} \otimes \mathcal{H} \otimes \cdots \otimes \mathcal{H}}_{n}$ において，

$$\rho^{\otimes n} = \underbrace{\rho \otimes \rho \otimes \cdots \otimes \rho}_{n}, \quad \sigma^{\otimes n} = \underbrace{\sigma \otimes \sigma \otimes \cdots \otimes \sigma}_{n}. \tag{3}$$

  のどちらかである．
- 検定 (test) は $0 \le T_n \le I_n$ を満たす $\mathcal{H}^{\otimes n}$ 上の作用素 T_n で与えられ，POVM $\{T_n, I_n - T_n\}$ と同一視される．

以降，$\mathcal{H}_n = \mathcal{H}^{\otimes n}$, $\rho_n := \rho^{\otimes n}$, $\sigma_n := \sigma^{\otimes n}$ などと書く．

- 真の状態が ρ であるときに，誤って σ であると判定する確率（第一種誤り確率）は

$$\alpha_n(T_n) := \mathrm{Tr}\, \rho_n (I_n - T_n)$$

- 真の状態が σ であるときに，誤って ρ であると判定する確率（第二種誤り確率）は

$$\beta_n(T_n) := \mathrm{Tr}\, \sigma_n T_n$$

12.2　量子Steinの補題：定数制約のトレードオフ（復習）

第一種誤り確率に$\alpha_n(T_n) \le \varepsilon$ $(0 < \varepsilon < 1)$の定数制約を課した上で，第二種誤り確率を可能な限り小さくする問題を考える．

$$\beta_n^*(\varepsilon) = \min \{ \beta_n(T_n) \mid T_n : \text{test}, \alpha_n(T_n) \le \varepsilon \} \tag{4}$$

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

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) = -D(\rho||\sigma) \tag{5}$$

ただし，$D(\rho||\sigma) = \mathrm{Tr}\,\rho(\log\rho - \log\sigma)$ は量子相対エントロピーである．

（意味）nが十分大きいとき，第二種誤りの最適な減衰レートが量子相対エントロピーで与えられる．

$$\beta_n^*(\varepsilon) \simeq e^{-nD(\rho||\sigma)} \tag{6}$$

（講義の目標）この定理を証明するため，順次，次の不等式の証明を与える．

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \le -D(\rho||\sigma) \quad \text{（順定理，direct part：今日はこちらの証明）} \tag{7}$$

$$\liminf_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \ge -D(\rho||\sigma) \quad \text{（逆定理，converse part）} \tag{8}$$

12.3 量子Neyman-Pearson検定（復習）

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

$$\mathrm{Tr}\, A_{+} = \mathrm{Tr}\, A\{A > 0\} = \max_{0 \le T \le I} \mathrm{Tr}\, AT \tag{9}$$

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

$$\min_{0 \le T \le I} \{\mathrm{Tr}\, A(I - T) + \mathrm{Tr}\, BT\} = \mathrm{Tr}\, A - \mathrm{Tr}(A - B)_{+} \tag{10}$$

ここで，最小化は$S := \{A - B > 0\}$で達成される．

Lemma 3 (前回Lemma 9). $A = \rho$, $B = e^a \sigma$ $(a \in \mathbb{R})$とおくことで，

$$\min_{0 \le T \le I} \{\alpha(T) + e^a \beta(T)\} = \min_{0 \le T \le I} \{\mathrm{Tr}\, \rho\, (I - T) + e^a\, \mathrm{Tr}\, \sigma T\} = 1 - \mathrm{Tr}(\rho - e^a \sigma)_{+} \tag{11}$$

最小化は$S(a) := \{\rho - e^a \sigma > 0\}$（量子Neyman-Pearson検定）で達成される．

Lemma 4 (前回 Theorem 2). 任意の非負定値作用素 A, B と，$0 \le \forall s \le 1$ について，

$$\min_{0 \le T \le I} \{\mathrm{Tr}\, A(I-T) + \mathrm{Tr}\, BT\} = \mathrm{Tr}\, A - \mathrm{Tr}(A-B)_+ \le \mathrm{Tr}\, A^s B^{1-s}. \tag{12}$$

（前回の式から，s, $1-s$ の入れ替えを行っている）

$\rho, \sigma \in \mathcal{S}(\mathcal{H})$ の仮説検定問題に適用：$A = \rho$, $B = e^a \sigma$ $(a \in \mathbb{R})$ とおく．Lemma 3 と合わせて，

$$\begin{aligned}\alpha(S(a)) + e^a \beta(S(a)) &= \min_{0 \le T \le I} \{\alpha(T) + e^a \beta(T)\} = \min_{0 \le T \le I} \{\mathrm{Tr}\, \rho\, (I-T) + e^a\, \mathrm{Tr}\, \sigma T\} \\ &= 1 - \mathrm{Tr}(\rho - e^a \sigma)_+ \le e^{a(1-s)}\, \mathrm{Tr}\, \rho^s \sigma^{1-s}\end{aligned}$$

よって，

$$\alpha(S(a)) \le \alpha(S(a)) + e^a \beta(S(a)) \le e^{a(1-s)}\, \mathrm{Tr}\, \rho^s \sigma^{1-s} \tag{13}$$

$$e^a \beta(S(a)) \le \alpha(S(a)) + e^a \beta(S(a)) \le e^{a(1-s)}\, \mathrm{Tr}\, \rho^s \sigma^{1-s} \tag{14}$$

Lemma 5. 任意の $a \in \mathbb{R}$ と $0 \le s \le 1$ について，

$$\alpha(S(a)) \le e^{a(1-s)}\, \mathrm{Tr}\, \rho^s \sigma^{1-s}, \quad \beta(S(a)) \le e^{-as}\, \mathrm{Tr}\, \rho^s \sigma^{1-s} \tag{15}$$

12.5　大偏差理論 (Large Deviation Theory)

$\rho \leftarrow \rho_n = \rho^{\otimes n}$, $\sigma \leftarrow \sigma_n = \sigma^{\otimes n}$, $a \leftarrow na$ の置き換えをする．

$S_n(a) := \{\rho_n - e^{na}\sigma_n > 0\}$ を量子Neyman-Pearson検定として，(15)より，

$$\alpha_n(S_n(a)) \le e^{na(1-s)} \operatorname{Tr} \rho_n^s \sigma_n^{1-s}, \quad \beta_n(S_n(a)) \le e^{-nas} \operatorname{Tr} \rho_n^s \sigma_n^{1-s} \tag{16}$$

$\operatorname{Tr} \rho_n^s \sigma_n^{1-s} = \operatorname{Tr} \left(\rho^{\otimes n}\right)^s \left(\sigma^{\otimes n}\right)^{1-s} = \operatorname{Tr} \left(\rho^s\right)^{\otimes n} \left(\sigma^{1-s}\right)^{\otimes n} = \operatorname{Tr} \left(\rho^s \sigma^{1-s}\right)^{\otimes n} = \left(\operatorname{Tr} \rho^s \sigma^{1-s}\right)^n$ に注意．

Definition 1. $\psi(s) := \log \operatorname{Tr} \rho^s \sigma^{1-s}, \quad \varphi(a) := \max_{0 \le s \le 1} \{as - \psi(s)\}$　（Legendre変換）

$$\frac{1}{n} \log \alpha_n(S_n(a)) \le -a(s-1) + \frac{1}{n} \log \operatorname{Tr} \rho_n^s \sigma_n^{1-s} = -\{as - \psi(s)\} + a \tag{17}$$

$$\frac{1}{n} \log \beta_n(S_n(a)) \le -as + \frac{1}{n} \log \operatorname{Tr} \rho_n^s \sigma_n^{1-s} = -\{as - \psi(s)\} \tag{18}$$

上界を $0 \le s \le 1$ について最適化すると，

Lemma 6. 任意の $a \in \mathbb{R}$ について，

$$\frac{1}{n} \log \alpha_n(S_n(a)) \le -\{\varphi(a) - a\} \tag{19}$$

$$\frac{1}{n} \log \beta_n(S_n(a)) \le -\varphi(a) \tag{20}$$

12.6　ポテンシャル関数 $\psi(s)$ と f-ダイバージェンス

$$\psi(s) = \log \operatorname{Tr} \rho^s \sigma^{1-s} \tag{21}$$

- スペクトル分解を考える．

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

- $P(i,j) := a_i \operatorname{Tr} E_i F_j, \quad Q(i,j) := b_j \operatorname{Tr} E_i F_j$ とおくと，$P(i,j)$, $Q(i,j)$ は確率分布である．

$$\sum_i \sum_j P(i,j) = \operatorname{Tr}\left(\sum_i a_i E_i\right)\left(\sum_j F_j\right) = \operatorname{Tr} \rho = 1 \quad \text{（}Q(i,j)\text{についても同様）}$$

- $P(i,j)$, $Q(i,j)$ で表すと，

$$\operatorname{Tr} \rho^s \sigma^{1-s} = \sum_i \sum_j a_i^s b_j^{1-s} \operatorname{Tr} E_i F_j = \sum_i \sum_j (a_i \operatorname{Tr} E_i F_j)^s (b_j \operatorname{Tr} E_i F_j)^{1-s} = \sum_i \sum_j P(i,j)^s Q(i,j)^{1-s}$$

- $x = (i,j)$ と変数の書き換えを行えば，

$$\psi(s) = \log \sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s}$$

（log の中身）$\sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s}$ は古典 f-ダイバージェンスとみなせる．

12.7 f-ダイバージェンス

Definition 2. 確率分布 $P(x)$, $Q(x)$ について，

$$D_f(P||Q) := \sum_{x \in \mathcal{X}} P(x) f\left(\frac{Q(x)}{P(x)}\right) \tag{22}$$

Example 1. $f(t) = -\log t$ のときは，相対エントロピー $D(P||Q)$ である．

$$D_f(P||Q) = -\sum_{x \in \mathcal{X}} P(x) \log\left(\frac{Q(x)}{P(x)}\right) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)} \tag{23}$$

Example 2. $f(t) = t \log t$ のときは，相対エントロピー $D(Q||P)$ である．

$$D_f(P||Q) = \sum_{x \in \mathcal{X}} P(x) \left(\frac{Q(x)}{P(x)}\right) \log\left(\frac{Q(x)}{P(x)}\right) = \sum_{x \in \mathcal{X}} Q(x) \log \frac{Q(x)}{P(x)} \tag{24}$$

Example 3. $f(t) = -t^{1-s}\ (0 \leq s \leq 1)$ のとき，

$$D_f(P||Q) = -\sum_{x \in \mathcal{X}} P(x) \left(\frac{Q(x)}{P(x)}\right)^{1-s} = -\sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s} \tag{25}$$

12.8　f-ダイバージェンスの単調性

条件付き確率 $W(y|x)$ $(x \in \mathcal{X}, y \in \mathcal{Y})$ から定まる古典通信路を考える：

$$P \in \mathcal{P}(\mathcal{X}) \longmapsto PW \in \mathcal{P}(\mathcal{Y}) \quad \left(PW(y) := \sum_{x \in \mathcal{X}} P(x) W(y|x) \right)$$

Theorem 2 (単調性). $f(t)$ が凸関数であるとき，

$$D_f(P||Q) \geq D_f(PW||QW) \tag{26}$$

$f(t)$ が厳密に凸関数であるとき，等号成立の必要十分条件は

$$\exists V(x|y)\,(\text{条件付き確率}), \quad PWV = P, \quad QWV = Q \tag{27}$$

（証明）　入力のシンボル x が，$P(x)$ または $Q(x)$ で発生する場合，入出力 (x, y) の同時分布は，

$$P_{XY}(x,y) := P(x)W(y|x), \quad Q_{XY}(x,y) := Q(x)W(y|x) \tag{28}$$

で与えられる．y の分布は周辺化により，

$$P_Y(y) = \sum_{x \in \mathcal{X}} P_{XY}(x,y) = \sum_{x \in \mathcal{X}} P(x)W(y|x) = PW(y) \tag{29}$$

$$Q_Y(y) = \sum_{x \in \mathcal{X}} Q_{XY}(x,y) = \sum_{x \in \mathcal{X}} Q(x)W(y|x) = QW(y) \tag{30}$$

同時分布について，yを条件とする形で無理やり書き換える．

$$P_{XY}(x,y) = P_Y(y)V_P(x|y), \quad Q_{XY}(x,y) = Q_Y(y)V_Q(x|y) \tag{31}$$

$V_P(x|y)$, $V_Q(x|y)$ は入力分布 P, Q に依存した条件付き分布になることに注意．

$$\begin{aligned} D_f(P_{XY}||Q_{XY}) &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_{XY}(x,y) f\left(\frac{Q_{XY}(x,y)}{P_{XY}(x,y)}\right) = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P(x)W(y|x) f\left(\frac{Q(x)W(y|x)}{P(x)W(y|x)}\right) \\ &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P(x)W(y|x) f\left(\frac{Q(x)}{P(x)}\right) = \sum_{x\in\mathcal{X}} P(x) f\left(\frac{Q(x)}{P(x)}\right) = D_f(P||Q) \end{aligned} \tag{32}$$

一方，Jensenの不等式（凸性の不等式）より，

$$\begin{aligned} D_f(P_{XY}||Q_{XY}) &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_Y(y)V_P(x|y) f\left(\frac{Q_Y(y)V_Q(x|y)}{P_Y(y)V_P(x|y)}\right) \\ &= \sum_{y\in\mathcal{Y}} P_Y(y) \sum_{x\in\mathcal{X}} V_P(x|y) f\left(\frac{Q_Y(y)V_Q(x|y)}{P_Y(y)V_P(x|y)}\right) \\ &\geq \sum_{y\in\mathcal{Y}} P_Y(y) f\left(\sum_{x\in\mathcal{X}} V_P(x|y)\frac{Q_Y(y)V_Q(x|y)}{P_Y(y)V_P(x|y)}\right) \end{aligned} \tag{33}$$

$$= \sum_{y\in\mathcal{Y}} P_Y(y) f\left(\frac{Q_Y(y)}{P_Y(y)}\right) = D_f(PW||QW) \tag{34}$$

- 等号成立条件について，(27) $\Rightarrow$ (26) の等号成立を示す．(27) を仮定すると，既に証明した単調性の不等式 (26) を二回使うことで，

$$D_f(P||Q) \geq D_f(PW||QW) \geq D_f(PWV||QWV) = D_f(P||Q) \tag{35}$$

両辺が一致して，(26) の等号が成立することが示された．
- $f(t)$ が厳密に凸関数であるとき，(26) の等号成立 $\Rightarrow$ (27) を示す．$f(t)$ が厳密に凸関数である場合，(26) の等号成立するのは，各 $y \in \mathcal{Y}$ について (33) における関数の引数がすべて等しい場合に限られる．すなわち，

$$\forall y \in \mathcal{Y}, \exists C_y, \forall x \in \mathcal{X}, \frac{Q_Y(y)V_Q(x|y)}{P_Y(y)V_P(x|y)} = \frac{Q_Y(y)}{P_Y(y)} = C_y \tag{36}$$

これより，

$$\forall y \in \mathcal{Y}, \forall x \in \mathcal{X}, \frac{V_Q(x|y)}{V_P(x|y)} = 1 \tag{37}$$

$V(x|y) := V_Q(x|y) = V_P(x|y)$ とおけば，(31) より，

$$P_{XY}(x,y) = P_Y(y)V(x|y), \quad Q_{XY}(x,y) = Q_Y(y)V(x|y) \tag{38}$$

この周辺分布をとれば，

$$P(x) = \sum_{y\in\mathcal{Y}} P_{XY}(x,y) = \sum_{y\in\mathcal{Y}} P_Y(y)V(x|y), \quad Q(x) = \sum_{y\in\mathcal{Y}} Q_{XY}(x,y) = \sum_{y\in\mathcal{Y}} Q_Y(y)V(x|y) \tag{39}$$

□

Lemma 7 (正値性). $f(t)$ が凸関数であるとき，

$$D_f(P||Q) \geq D_f(P_0||P_0) = f(1) \tag{40}$$

$f(t)$ が厳密に凸な関数であるとき，等号成立の必要十分条件は $P = Q$ である．

- $f(t) = -t^{1-s}$ は $0 \leq s \leq 1$ で凸関数であるから，

$$D_f(P||Q) = -\sum_{x \in \mathcal{X}} P(x) \left(\frac{Q(x)}{P(x)} \right)^{1-s} = -\sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s} \geq f(1) = -1 \tag{41}$$

すなわち，

$$\psi(s) = \log \sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s} \leq 0 \quad (0 \leq s \leq 1) \tag{42}$$

- $f(t) = t^{1-s}$ は $s \leq 0$ または $s \geq 1$ で凸関数であるから，

$$D_f(P||Q) = \sum_{x \in \mathcal{X}} P(x) \left(\frac{Q(x)}{P(x)} \right)^{1-s} = \sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s} \geq f(1) = 1 \tag{43}$$

すなわち，

$$\psi(s) = \log \sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s} \geq 0 \quad (s \leq 0, s \geq 1) \tag{44}$$

12.10　ポテンシャル関数$\psi(s)$の性質

Lemma 8. $\psi(s) = \log \sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s}$ について，以下の性質が成り立つ．

(i) $\psi(0) = 0$, $\psi(1) = 0$
(ii) $\psi(s) \le 0\ (0 \le s \le 1)$, $\psi(s) \ge 0\ (s \le 0, s \ge 1)$.
(iii) $\psi'(0) = -D(Q||P)$, $\psi'(1) = D(P||Q)$.
(iv) $P \ne Q$ならば，$\psi''(s) > 0$である（$\psi(s)$は厳密に凸な関数）

（証明）(i)は自明，(ii)は既に示した．
定義より，$e^{\psi(s)} = \sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s}$であり，

$$P_s(x) := e^{-\psi(s)} P(x)^s Q(x)^{1-s} = \frac{P(x)^s Q(x)^{1-s}}{\sum_{x \in \mathcal{X}} P(x)^s Q(x)^{1-s}} \tag{45}$$

とおくと，$P_s(x)$は確率分布で

$$P_0(x) = Q(x), \quad P_1(x) = P(x) \tag{46}$$

微分を計算する．

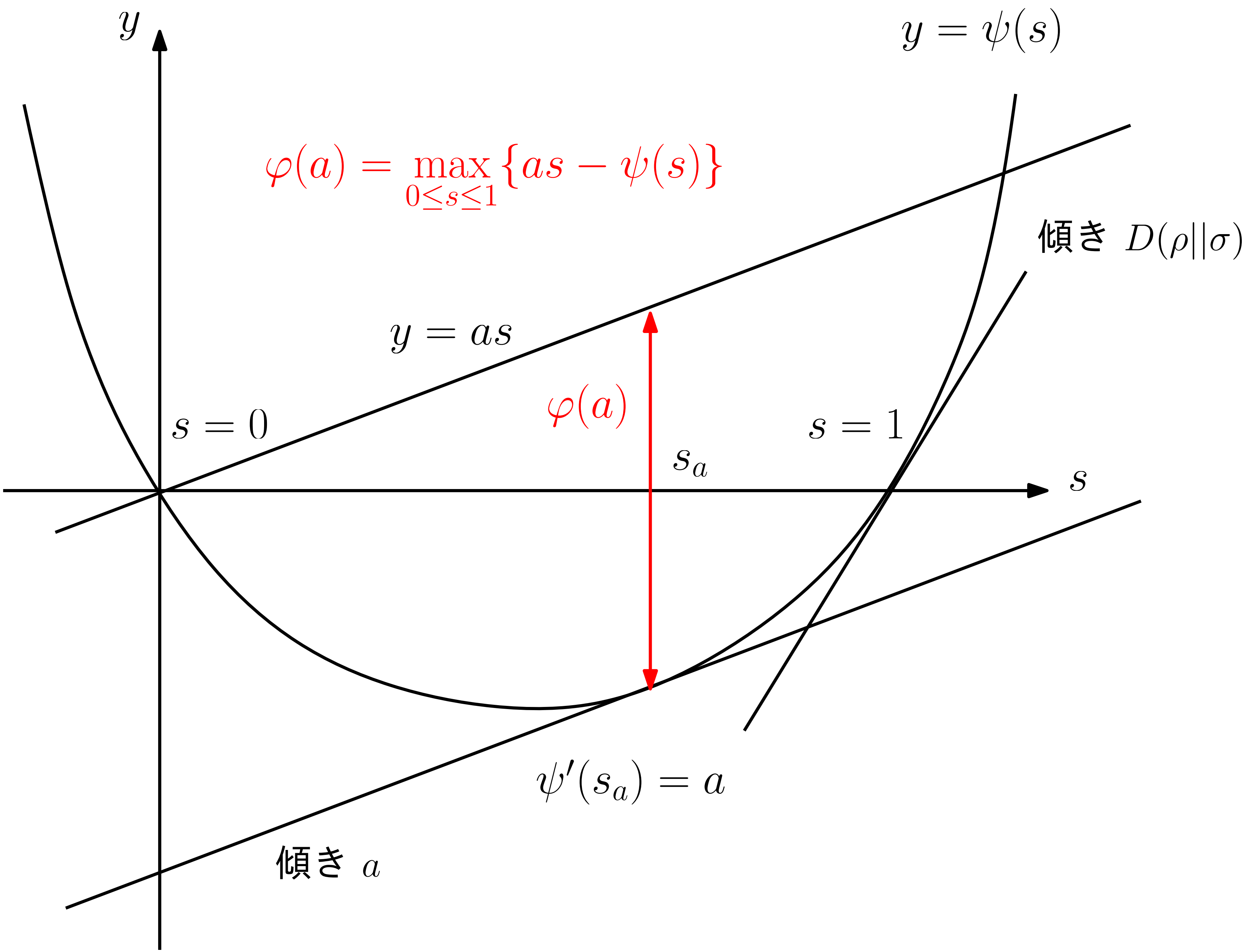
$$
\begin{aligned}
\psi'(s) &= \frac{\sum_{x\in\mathcal{X}} P(x)^s Q(x)^{1-s}\{\log P(x) - \log Q(x)\}}{\sum_{x\in\mathcal{X}} P(x)^s Q(x)^{1-s}} \\
&= e^{-\psi(s)} \sum_{x\in\mathcal{X}} P(x)^s Q(x)^{1-s}\{\log P(x) - \log Q(x)\} \qquad (47) \\
&= \sum_{x\in\mathcal{X}} P_s(x)\{\log P(x) - \log Q(x)\} \\
&= \mathrm{E}\Big[\log P(X) - \log Q(X) \;\Big|\; X \sim P_s\Big], \qquad (48)
\end{aligned}
$$

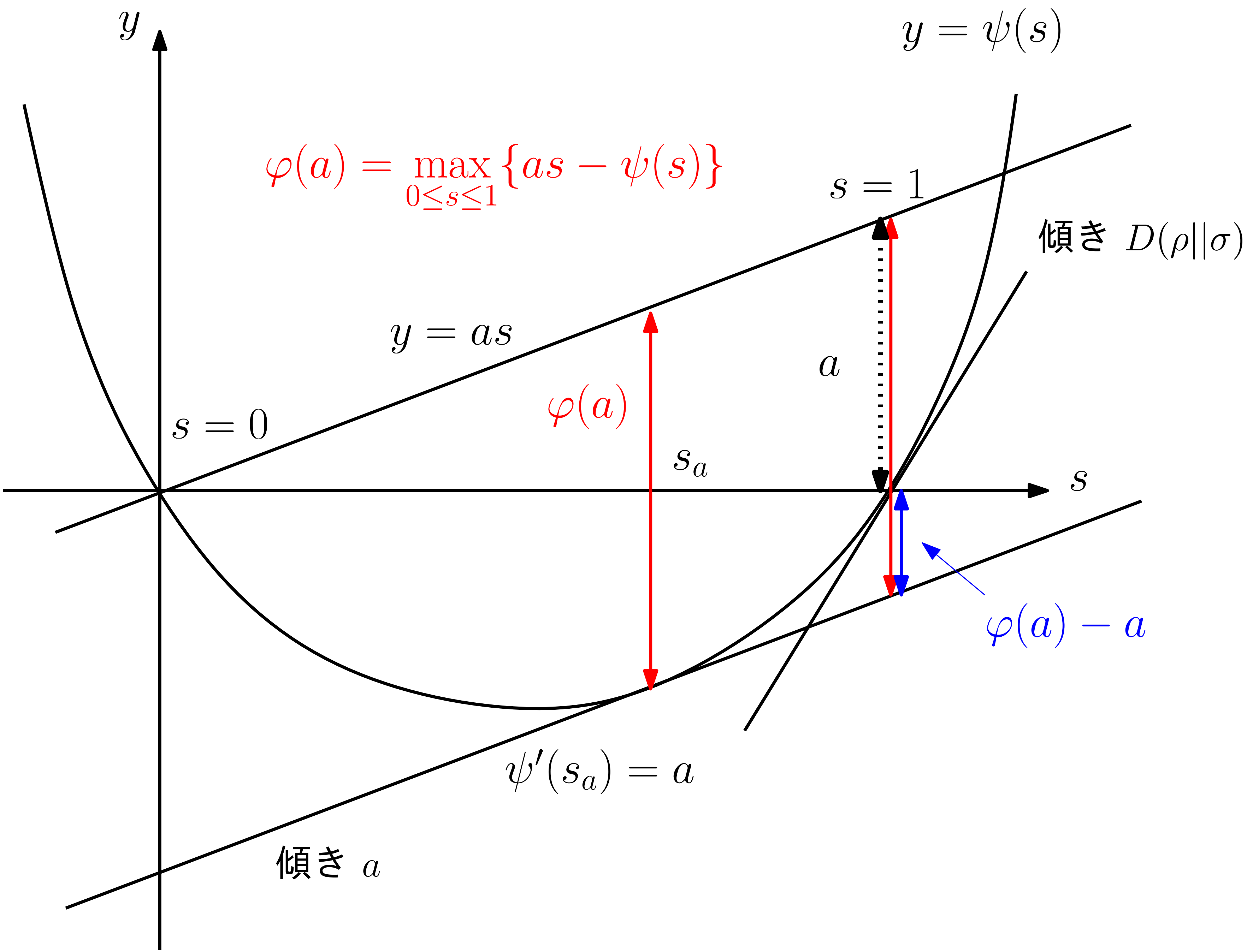
ここで，$s=0$, $s=1$ とすると(iii)が示される．(47)より，

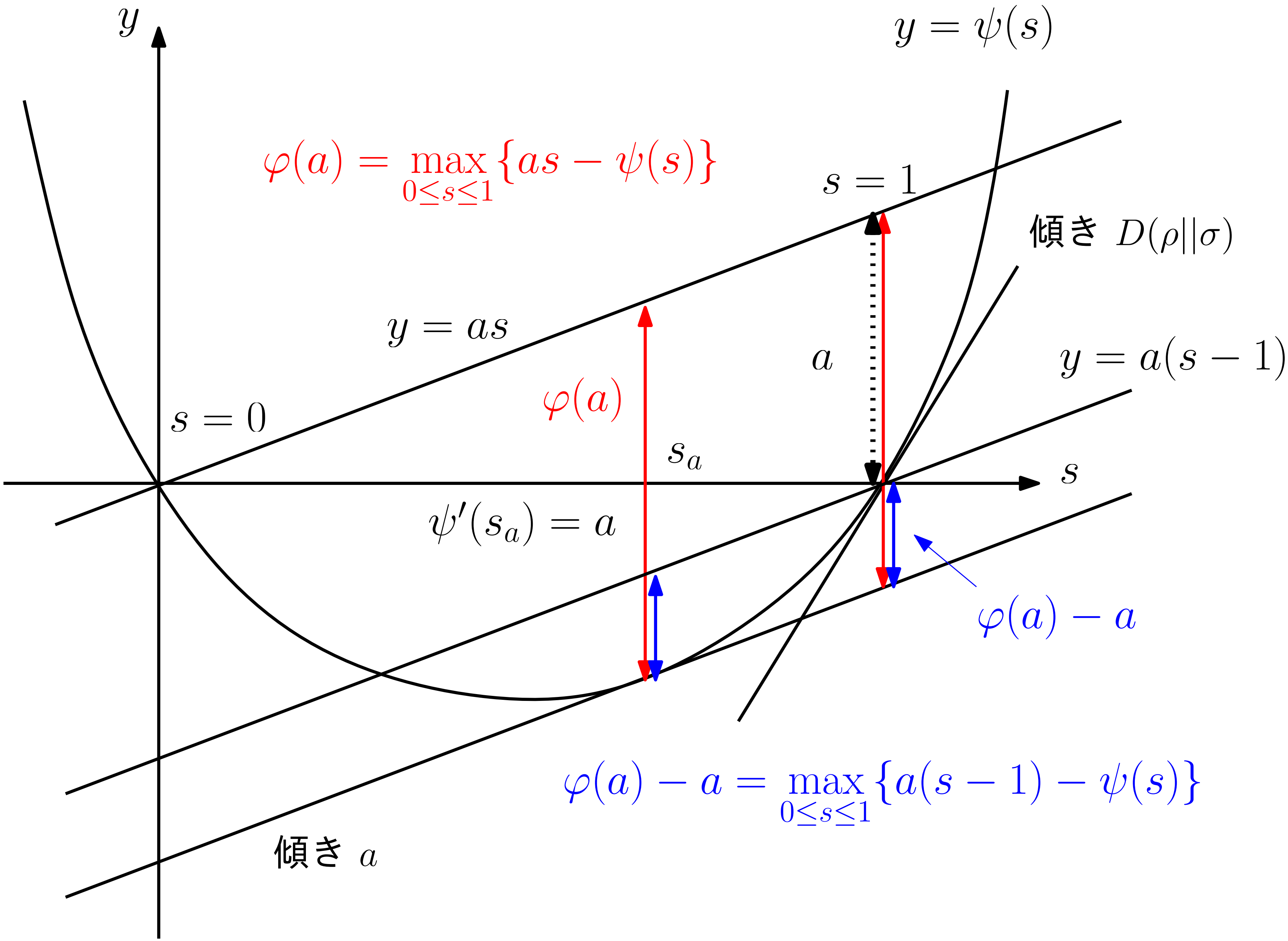
$$
\begin{aligned}
\psi''(s) &= -\psi'(s) e^{-\psi(s)} \sum_{x\in\mathcal{X}} P(x)^s Q(x)^{1-s}\{\log P(x) - \log Q(x)\} \\
&\quad + e^{-\psi(s)} \sum_{x\in\mathcal{X}} P(x)^s Q(x)^{1-s}\{\log P(x) - \log Q(x)\}^2 \\
&= \sum_{x\in\mathcal{X}} P_s(x)\{\log P(x) - \log Q(x)\}^2 - \psi'(s) \sum_{x\in\mathcal{X}} P_s(x)\{\log P(x) - \log Q(x)\} \\
&= \mathrm{E}\Big[\{\log P(X) - \log Q(X)\}^2\Big] - \Big\{\mathrm{E}\left[\log P(X) - \log Q(X)\right]\Big\}^2 \\
&= V\Big[\log P(X) - \log Q(X) \;\Big|\; X \sim P_s\Big] > 0 \qquad (49)
\end{aligned}
$$

ここで分散は（一点分布の場合を除いて）正であることを用いた．よって(iv)が示された．

12.11 $\psi(s)$ と $\varphi(a)$ のグラフ(1)




y
$y = \psi(s)$
$\varphi(a) = \max_{0 \le s \le 1} \{as - \psi(s)\}$
$s = 1$
傾き $D(\rho\|\sigma)$
$y = as$
a
$\varphi(a)$
$s = 0$
s_a
s
$\varphi(a) - a$
$\psi'(s_a) = a$
傾き a


y
$y = \psi(s)$
$\varphi(a) = \max_{0\le s\le 1}\{as - \psi(s)\}$
$s = 1$
傾き $D(\rho||\sigma)$
$y = as$
a
$y = a(s-1)$
$\varphi(a)$
$s = 0$
s_a
s
$\psi'(s_a) = a$
$\varphi(a) - a$
$\varphi(a) - a = \max_{0\le s\le 1}\{a(s-1) - \psi(s)\}$
傾き a

Definition 3 (復習)**.**

$$\psi(s) := \log \operatorname{Tr} \rho^s \sigma^{1-s} \tag{50}$$

$$\varphi(a) := \max_{0 \le s \le 1} \{as - \psi(s)\} \tag{51}$$

Lemma 9 (復習Lemma 6)**.** 任意の$a \in \mathbb{R}$と任意の$n \in \mathbb{N}$について,

$$\frac{1}{n} \log \alpha_n(S_n(a)) \le -\{\varphi(a) - a\} \qquad \left[\alpha_n(S_n(a)) \le e^{-n\{\varphi(a)-a\}}\right] \tag{52}$$

$$\frac{1}{n} \log \beta_n(S_n(a)) \le -\varphi(a) \qquad \left[\beta_n(S_n(a)) \le e^{-n\varphi(a)}\right] \tag{53}$$

Remark 1. 上式の評価は漸近的に最適であることが知られている.

$$\lim_{n\to\infty} \frac{1}{n} \log \alpha_n(S_n(a)) = -\{\varphi(a) - a\} \tag{54}$$

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n(S_n(a)) = -\varphi(a) \tag{55}$$

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

$$\lim_{n\to\infty} \alpha_n(S_n(a)) = 0 \quad \left(\lim_{n\to\infty} \mathrm{Tr}\,\rho_n S_n(a) = 1\right) \tag{56}$$

（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，(52)より題意が示される． □

Lemma 11. 任意の$a \in \mathbb{R}$について，

$$\varphi(a) \geq a \tag{57}$$

よって，(53)より，

$$\beta_n(S_n(a)) \leq e^{-n\varphi(a)} \leq e^{-na} \quad (\forall n \in \mathbb{N}) \tag{58}$$

（証明） $\varphi(a)$の定義より，

$$\varphi(a) = \max_{0\leq s\leq 1} \{as - \psi(s)\} \geq a \cdot 1 - \psi(1) = a \tag{59}$$

□

Definition 4 (復習).

$$\beta_n^*(\varepsilon) = \min\{\, \beta_n(T_n) \mid T_n : \text{test},\ \alpha_n(T_n) \le \varepsilon \,\} \tag{60}$$

Theorem 3 (量子Steinの補題：direct part). 任意の$0 < \varepsilon \le 1$について，

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \le -D(\rho||\sigma) \tag{61}$$

（証明） 任意の$a < D(\rho||\sigma)$について，Lemma 10より，

$$\exists N \in \mathbb{N},\ \forall n \ge N\ (\text{十分大きなすべての}n\text{について}),\quad \alpha_n(S_n(a)) \le \varepsilon \tag{62}$$

よって，$\beta_n^*(\varepsilon)$の定義より，

$$\forall n \ge N, \quad \beta_n^*(\varepsilon) \le \beta_n(S_n(a)) \le e^{-na} \tag{63}$$

ただし，最後の不等式は(58)を用いた．これより，

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \le -a \tag{64}$$

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