数理解析研究所講究録 1452 巻 2005 年 206-214

Shorの素因数分解アルゴリズムにおける計算量の精密な評価

栗山憲
(Ken Kuriyama)
山口大・工
(Department of Applied Science, Yamaguchi University)

佐野慎太郎
(Shintaro Sano)
山口大・理工
(Graduate School of Science and Engineering, Yamaguchi University)

古市茂
(Shigeru Furuichi)
山口東京理科大・基礎工
(Department of Electronics and Computer Science, Tokyo University of Science in Yamaguchi)

1 はじめに

1994 年に Shor は量子コンピュータを用いた効率的な素因数分解アルゴリズムを発表した [1, 2]. 現在のコンピュータでの大きな整数の素因数分解には膨大な計算量を要することはよく知られており，インターネット等で広く使われている暗号の安全性はこの計算量の大きさに依存している．そこで，量子コンピュータが実現されるとこの種の暗号の安全性が崩壊するとして Shor の研究は注目を浴びた．

Shor のアルゴリズムの計算量は，素因数分解したい数 n を 2 進数表示したときの桁数 $\log_2 n$ についての多項式時間となる．本研究の目的は，この計算量をより精密に評価することである．従来の評価では，素因数分解したい数 $n = {p_1}^{e_1}{p_2}^{e_2}\ldots{p_k}^{e_k}$ に対して，十分大きい確率で正しく素因数分解するために必要なアルゴリズムの実行回数 N は

$$\forall \varepsilon > 0, \quad N \geq \frac{\log(1/\varepsilon)}{\alpha\beta(1-1/2^{k-1})}(\log_2 n)^2$$

であった．これに対して，本研究の精密な評価では $p_i - 1 = 2^{\tau_i}\sigma_i$ $(i = 1, 2, \ldots, k,\ \tau_i \geq 1,\ \sigma_i$ は奇数$)$, $\tau' = \min(\tau_1, \ldots, \tau_k)$, $\tilde{\tau} = \sum_{i=1}^{k} \tau_i$ とすると

$$\forall \varepsilon > 0, \quad N \geq \frac{\log(1/\varepsilon)}{\alpha\beta\left(1 - \frac{1}{2^k-1}\frac{2^k-2+2^{k\tau'}}{2^{\tilde{\tau}}}\right)}(\log_2 n)^2$$

と表すことができる．これにより，従来の評価が最良の結果であることが明らかになるとともに，素因数分解したい数 n の構成と計算量の関係を考察することができる．

第 2 章では Shor の素因数分解アルゴリズムとその計算量の評価の方法を説明する．第 3 章では従来の計算量の評価を紹介する．第 4 章で本研究での精密な評価に必要な整数論の結果を説明し，第 5 章で実際に精密な評価を行う．最後に第 6 章で従来の評価と本研究の結果の比較をする．

2 Shor のアルゴリズム

Shor の素因数分解のアルゴリズムとその計算量の評価の方法を紹介する．素因数分解したい数を n とする．ここでは簡単のために n は二つの素数の積で $n = pq$ と表されている場合を考える．そうすると素因数分解のアルゴリズムは以下のようにまとめることができる．

1° : $\{1, 2, \ldots, n\}$ からランダムに一つ選び，その数を a とする．
2° : $\gcd(a, n) = 1$ ならば 3° へ行く．$\gcd(a, n) \neq 1$ ならば 1° へ戻る．
3° : a の $\bmod n$ に関する位数 r を求める．(量子コンピュータによる)
4° : 得られた位数 r が偶数ならば 5° へ行く．奇数ならば 1° へ戻る．
5° : $p' = \gcd(a^{r/2} + 1, n)$ と $q' = \gcd(a^{r/2} - 1, n)$ を求める．
6° : p', q' のいずれかが n ならば 1° へ戻る．そうでなければそれらが求める因数 p, q である．

次に，アルゴリズムの計算量の評価の方法を説明する [3]．アルゴリズムを 1 回実行して素因数分解に成功する確率を P_S とすると，十分大きい確率で正しく素因数分解するために必要なアルゴリズムの実行回数 N は

$$\forall \varepsilon > 0, \quad N \geq \log(1/\varepsilon)/P_S \tag{2.1}$$

を満たせばよい．ここで，確率 P_S を評価するために次のような事象を考える．

A_a : $\gcd(a, n) = 1$ となるような n 未満の数 a が得られる事象
A_r : 量子コンピュータによって正しい位数 r が得られる事象
A_e : 量子コンピュータによって得られた位数 r が偶数である事象
A_f : 得られた位数から正しい因数 p, q が得られる事象

これらの事象を用いると，確率 P_S は

$$\begin{aligned} P_S &= P(A_a \cap A_r) P(A_e \cap A_f \mid A_a \cap A_r) + P(A_a \cap A_r) P(A_e \cap A_f \mid \overline{A_a \cap A_r}) \\ &\geq P(A_a \cap A_r) P(A_e \cap A_f \mid A_a \cap A_r) \\ &= P(A_a) P(A_r) P(A_e \cap A_f \mid A_a \cap A_r) \end{aligned} \tag{2.2}$$

となる．ここで，確率 $P(A_a), P(A_r), P(A_e \cap A_f \mid A_a \cap A_r)$ を求めることで，式 (2.1) 及び式 (2.2) から必要なアルゴリズムの実行回数が得られる．

3 従来の計算量の評価

この章では，確率 $P(A_a), P(A_r), P(A_e \cap A_f \mid A_a \cap A_r)$ の従来の評価の方法を紹介する [3]．まず，確率 $P(A_a)$ は，$\gcd(a, n) = 1$ となるような n 未満の数 a が得られる確率であるから，Euler の関数を用いて

$$P(A_a) = \frac{\varphi(n)}{n-1}$$

と表すことができる．Euler の関数については

$$\liminf_{n \to \infty} \frac{\varphi(n) \log\log n}{n} = e^{-\gamma}$$

が成り立つことが知られている．ここで γ は Euler の定数である．したがって，十分大きい n に対して

$$P(A_a) = \frac{\varphi(n)}{n-1} \geq \frac{e^{-\gamma}}{\log\log n} \geq \frac{e^{-\gamma}}{\log n} = \frac{e^{-\gamma}\log_2 e}{\log_2 n} = \frac{\alpha}{\log_2 n} \tag{3.1}$$

が成り立つ．ここで α は n に依存しない定数である．

次に確率 $P(A_r)$ を求める．これは量子コンピュータによって正しい位数 r が得られる確率である．この確率は，r 未満で r と互いに素な数が得られる確率を用いて表すことができる．すなわち，確率 $P(A_a)$ の場合と同様にして

$$P(A_r) \geq \frac{\beta}{\log_2 n} \tag{3.2}$$

と表すことができる．ここで β は n に依存しない定数である．

最後に確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ について説明する．一般に，k 種類の素数の積で $n = p_1{}^{e_1}p_2{}^{e_2}\ldots p_k{}^{e_k}$ と表される n に対して，量子コンピュータによって得られた位数 r が偶数であり，かつ得られた位数から正しい素因数が得られる確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ は

$$P(A_e \cap A_f \mid A_a \cap A_r) \geq 1 - \frac{1}{2^{k-1}} \tag{3.3}$$

であることが知られている [4]．この確率を第 5 章でより精密に記述する．

以上のことから，アルゴリズムの計算量を評価することができる．式 (3.1), (3.2), (3.3) を式 (2.2) に代入すると，アルゴリズムを 1 回実行して素因数分解に成功する確率 P_S は

$$P_S \geq \left(1 - \frac{1}{2^{k-1}}\right)\frac{\alpha\beta}{(\log_2 n)^2} \tag{3.4}$$

となり，式 (2.1) より，十分大きい確率で正しく素因数分解するために必要なアルゴリズムの実行回数 N は

$$\forall \varepsilon > 0, \quad N \geq \frac{\log(1/\varepsilon)}{\alpha\beta(1-1/2^{k-1})}(\log_2 n)^2 \tag{3.5}$$

となる．すなわち，アルゴリズムの実行回数は，素因数分解したい数 n を 2 進数表示したときの桁数 $\log_2 n$ のオーダーになることがわかる．また，位数 r を求める量子コンピュータを構成するのに必要なゲートの数も $O(\log_2 n)$ であることが知られており，総合して Shor のアルゴリズムの計算量は $O(\log_2 n)$ である．

4 Shor のアルゴリズムに関連する整数論の結果

この章では，式 (3.3) をより精密に評価するために必要な整数論の結果を用意する．素数 p に対して，体 Z/pZ の invertible element 全体を $(Z/pZ)^\times$ とすると，$(Z/pZ)^\times = \{1, 2, \ldots, p-1\}$ は $p-1$ 個の元をもつ．このとき，よく知られた結果として

$$\left|\{a \in (Z/pZ)^\times ; r_p = d\}\right| = \varphi(d)$$

が成り立つ [5]．但し，$d \mid p-1$，r_p は a の mod$p$ に関する位数，$\varphi(\cdot)$ は Euler の関数である．

補題 4.1　素数 p に対して，$p-1 = 2^\tau\sigma$ $(\tau \geq 1, \sigma : odd)$ と書くことができ，

$$\left|\{a \in (Z/pZ)^\times ; r_p : odd\}\right| = \sigma \tag{4.1}$$

$$\left|\{a \in (Z/pZ)^\times ; r_p = 2^t s\,(s : odd)\}\right| = 2^{t-1}\sigma \tag{4.2}$$

が成り立つ．ここで，t は $1 \leq t \leq \tau$ の固定された数である．

(証明)　まず，$r_p = 2^t s\,(t \geq 0, s : odd)$ と書くと

$$r_p : odd,\ r_p \mid p-1 \Longleftrightarrow r_p \mid \sigma$$

が成り立つ．$r_p \mid p-1 = 2^t s \mid 2^\tau \sigma$ より $t \leq \tau,\ s \mid \sigma$ であり，$r_p = 2^t s$ が奇数であることから $t = 0$．したがって，$r_p = s$ となり，$s \mid \sigma$ から $r_p \mid \sigma$ が得られる．逆に，$r_p \mid \sigma$ ならば，r_p は σ の約数であるから r_p は奇数である．また，$p-1 = 2^\tau \sigma$ より $r_p \mid \sigma \Longrightarrow r_p \mid p-1$ が成り立つ。このことに注意すると

$$\begin{aligned}\left|\{a \in (\boldsymbol{Z}/p\boldsymbol{Z})^\times ; r_p : odd\}\right| &= \sum_{r_p \mid p-1,\, r_p : odd} \varphi(r_p) \\ &= \sum_{r_p \mid \sigma} \varphi(r_p) \\ &= \sigma\end{aligned}$$

となり，式 (4.1) が得られる．次に，$r_p = 2^t s$ のとき

$$r_p \mid p-1 \Longleftrightarrow s \mid \sigma$$

が成り立つ．仮定 $1 \leq t \leq \tau$ より $2^t s \mid 2^\tau \sigma \Longrightarrow s \mid \sigma$ であり，逆に $s \mid \sigma \Longrightarrow 2^t s \mid 2^\tau \sigma$ が成り立つからである．したがって，固定された $t\,(1 \leq t \leq \tau)$ に対して

$$\begin{aligned}\left|\{a \in (\boldsymbol{Z}/p\boldsymbol{Z})^\times ; r_p = 2^t s\,(s : odd)\}\right| &= \sum_{r_p \mid p-1,\, r_p = 2^t s} \varphi(r_p) \\ &= \sum_{s \mid \sigma} \varphi(2^t s) \\ &= \sum_{s \mid \sigma} \varphi(2^t)\varphi(s) \\ &= \varphi(2^t) \sum_{s \mid \sigma} \varphi(s) \\ &= 2^t \left(1 - \frac{1}{2}\right) \sigma \\ &= 2^{t-1} \sigma\end{aligned}$$

となり，式 (4.2) が得られる．

補題 4.2　$n = {p_1}^{e_1} \ldots {p_k}^{e_k}$ (p_i は素数，$i = 1, 2, \ldots, k$) に対して，$p_i - 1 = 2^{\tau_i} \sigma_i$ $(\tau_i \geq 1, \sigma_i : odd)$ と書くことができ，a の $\mathrm{mod} n$ に関する位数を r，$\mathrm{mod} p_i$ に関する位数を $r_{p_i} = 2^{t_{p_i}} s_{p_i}$ とすると

$$\left|\{a \in (\boldsymbol{Z}/n\boldsymbol{Z})^\times ; r : odd\}\right| = \prod_{i=1}^{k} \sigma_{p_i} \tag{4.3}$$

$$\left|\{a \in (\boldsymbol{Z}/n\boldsymbol{Z})^\times ; t_{p_1} = \cdots = t_{p_k} = l\}\right| = 2^{k(l-1)} \prod_{i=1}^{k} \sigma_{p_i} \tag{4.4}$$

が成り立つ．但し，$1 \leq l \leq \min(\tau_{p_1}, \ldots, \tau_{p_k})$ である．

(証明)　Chinese Remainder Theorem より，$(Z/nZ)^\times \cong (Z/p_1Z)^\times \oplus \cdots \oplus (Z/p_kZ)^\times$ が成り立つので，$a \in (Z/nZ)^\times$ に対して

$$r = \mathrm{lcm}\{r_{p_1}, \ldots, r_{p_k}\}$$
$$r : odd \Longleftrightarrow r_{p_1}, \ldots, r_{p_k} : odd$$
$$\left|(Z/nZ)^\times\right| = \left|(Z/p_1Z)^\times\right| \cdots \left|(Z/p_kZ)^\times\right|$$

が成り立つことに注意すると，式 (4.1) を用いて

$$\begin{aligned}
&\left|\{a \in (Z/nZ)^\times ; r : odd\}\right| \\
&= \left|\{a \in (Z/nZ)^\times ; r_{p_1}, \ldots, r_{p_k} : odd\}\right| \\
&= \left|\{a \in (Z/p_1Z)^\times ; r_{p_1} : odd\}\right| \cdots \left|\{a \in (Z/p_kZ)^\times ; r_{p_k} : odd\}\right| \\
&= \prod_{i=1}^{k} \sigma_{p_i}
\end{aligned}$$

となり，式 (4.3) が得られる．また，式 (4.2) を用いて

$$\begin{aligned}
&\left|\{a \in (Z/nZ)^\times ; t_{p_1} = \cdots = t_{p_k} = l\}\right| \\
&= \left|\{a \in (Z/p_1Z)^\times ; t_{p_k} = l\}\right| \cdots \left|\{a \in (Z/p_kZ)^\times ; t_{p_k} = l\}\right| \\
&= 2^{k(l-1)} \prod_{i=1}^{k} \sigma_{p_i}
\end{aligned}$$

となり，式 (4.4) が得られる．

定理 4.3　$\tau' = \min(\tau_{p_1}, \ldots, \tau_{p_k})$, $\tilde{\tau} = \sum_{i=1}^{k} \tau_{p_i}$ とおくと

$$\left|\{a \in (Z/nZ)^\times ; t_{p_1} = \cdots = t_{p_k}\}\right| = \frac{2^k - 2 + 2^{k\tau'}}{2^k - 1} \prod_{i=1}^{k} \sigma_{p_i} \tag{4.5}$$

であり

$$\frac{\left|\{a \in (Z/nZ)^\times ; t_{p_1} = \cdots = t_{p_k}\}\right|}{\left|(Z/nZ)^\times\right|} = \frac{1}{2^k - 1} \frac{2^k - 2 + 2^{k\tau'}}{2^{\tilde{\tau}}} \tag{4.6}$$

が成り立つ．

(証明)　式 (4.5) の左辺を変形し，式 (4.3) および (4.4) を用いると

$$\begin{aligned}
&\left|\{a \in (Z/nZ)^\times ; t_{p_1} = \cdots = t_{p_k}\}\right| \\
&= \left|\bigcup_{l=0}^{\tau'} \{a \in (Z/nZ)^\times ; t_{p_1} = \cdots = t_{p_k} = l\}\right| \\
&= \sum_{l=0}^{\tau'} \left|\{a \in (Z/nZ)^\times ; t_{p_1} = \cdots = t_{p_k} = l\}\right| \\
&= \left|\{a \in (Z/nZ)^\times ; t_{p_1} = \cdots = t_{p_k} = 0\}\right| + \sum_{l=1}^{\tau'} \left|\{a \in (Z/nZ)^\times ; t_{p_1} = \cdots = t_{p_k} = l\}\right| \\
&= \prod_{i=1}^{k} \sigma_{p_i} + \sum_{l=1}^{\tau'} \left(2^{k(l-1)} \prod_{i=1}^{k} \sigma_{p_i}\right)
\end{aligned}$$

$$= \frac{2^k - 2 + 2^{k\tau'}}{2^k - 1} \prod_{i=1}^{k} \sigma_{p_i}$$

となり，右辺が得られる．また，

$$\left|(Z/nZ)^\times\right| = \left|(Z/p_1Z)^\times\right| \cdots \left|(Z/p_kZ)^\times\right| = 2^{\tilde{\tau}} \prod_{i=1}^{k} \sigma_{p_i}$$

であることに注意すれば，式 (4.5) より式 (4.6) は明らか．

5 精密な計算量の評価

第 3 章の式 (3.3) で表される確率は，n の素因数の個数のみで表現されている．この章では，第 4 章での整数論の結果を用いて，素因数の個数だけでなく，素因数から定まる数を使って確率を精密に評価する．

補題 5.1 $n = {p_1}^{e_1} \ldots {p_k}^{e_k}$ (p_i は素数，$i = 1, 2, \ldots, k$) に対して，$p_i - 1 = 2^{\tau_i}\sigma_i$ ($\tau_i \geq 1$, σ_i は奇数)，$\tau' = \min(\tau_1, \ldots, \tau_k)$，$\tilde{\tau} = \sum_{i=1}^{k} \tau_i$ とすると

$$P(A_e \cap A_f \mid A_a \cap A_r) = 1 - \frac{1}{2^k - 1} \frac{2^k - 2 + 2^{k\tau'}}{2^{\tilde{\tau}}} \tag{5.1}$$

が成り立つ．

(証明)　アルゴリズムのステップ 5° と 6° と位数の性質に注意すると

$$\begin{aligned} P(A_e \cap A_f \mid A_a \cap A_r) &= P\left(\{r : even\} \cap \{a^{r/2} \neq \pm 1 \pmod n\} \middle| A_a \cap A_r\right) \\ &= P\left(\{r : even\} \cap \{a^{r/2} \neq -1 \pmod n\} \middle| A_a \cap A_r\right) \\ &= 1 - P\left(\{r : odd\} \cup \{a^{r/2} = -1 \pmod n\} \middle| A_a \cap A_r\right) \end{aligned}$$

となる．ここで a の mod$n$ に関する位数を $r = 2^t s$，modp_i に関する位数を $r_i = 2^{t_i} s_i$ とする．ただし，$i = 1, 2, \ldots, k$，$t, t_i \geq 1$，s, s_i は奇数とする．すると確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ はさらに変形でき

$$\begin{aligned} &P(A_e \cap A_f \mid A_a \cap A_r) \\ &= 1 - P\left(\{r : odd\} \cup \left(\bigcap_{i=1}^{k} \{a^{r/2} = -1 \pmod{p_i}\}\right) \middle| A_a \cap A_r\right) \\ &= 1 - P\left(\{t_1 = \cdots = t_k = 0\} \cup \left(\bigcap_{i=1}^{k} \{t_i = t\}\right) \middle| A_a \cap A_r\right) \\ &= 1 - P\left(t_1 = \cdots = t_k \mid A_a \cap A_r\right) \end{aligned}$$

となる．ここで式 (4.6) を用いると式 (5.1) が得られる．

定理 5.2 $n = {p_1}^{e_1} \ldots {p_k}^{e_k}$ (p_i は素数，$i = 1, 2, \ldots, k$) に対して，$p_i - 1 = 2^{\tau_i}\sigma_i$ ($\tau_i \geq 1$, σ_i は奇数)，$\tau' = \min(\tau_1, \ldots, \tau_k)$，$\tilde{\tau} = \sum_{i=1}^{k} \tau_i$ とすると，アルゴリズムを 1 回実行して素因数分解に成功する確率 P_S は

$$P_S \geq \left(1 - \frac{1}{2^k - 1} \frac{2^k - 2 + 2^{k\tau'}}{2^{\tilde{\tau}}}\right) \frac{\alpha\beta}{(\log_2 n)^2} \tag{5.2}$$

であり，十分大きい確率で正しく素因数分解するために必要なアルゴリズムの実行回数 N は

$$\forall \varepsilon > 0, \quad N \geq \frac{\log(1/\varepsilon)}{\alpha\beta\left(1-\frac{1}{2^{k}-1}\frac{2^{k}-2+2^{k\tau'}}{2^{\tilde{\tau}}}\right)}(\log_2 n)^2 \tag{5.3}$$

である．ここで α, β は n に依存しない定数である．

(証明)　式 (5.1) を式 (2.1) および (2.2) に用いる．

6　結果の比較

本論文で精密に評価した確率は，量子コンピュータによって得られた位数 r が偶数であり，かつ得られた位数から正しい素因数が得られる確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ である．従来の評価では，$n = {p_1}^{e_1}{p_2}^{e_2}\ldots{p_k}^{e_k}$ とすると

$$P(A_e \cap A_f \mid A_a \cap A_r) \geq 1 - \frac{1}{2^{k-1}}$$

とされていたものに対して，$p_i - 1 = 2^{\tau_i}\sigma_i$ $(i = 1,2,\ldots,k,\ \tau_i \geq 1,\ \sigma_i$ は奇数$)$，$\tau' = \min(\tau_1,\ldots,\tau_k)$，$\tilde{\tau} = \sum_{i=1}^{k}\tau_i$ とすることで

$$P(A_e \cap A_f \mid A_a \cap A_r) = 1 - \frac{1}{2^{k}-1}\frac{2^{k}-2+2^{k\tau'}}{2^{\tilde{\tau}}}$$

と精密に表すことができた．これらの式の間には次のような関係がある．

定理 6.1　$n = {p_1}^{e_1}\ldots{p_k}^{e_k}$ (p_i は素数，$i = 1,2,\ldots,k$) に対して，$p_i - 1 = 2^{\tau_i}\sigma_i$ $(\tau_i \geq 1,\ \sigma_i$ は奇数$)$，$\tau' = \min(\tau_1,\ldots,\tau_k)$，$\tilde{\tau} = \sum_{i=1}^{k}\tau_i$ とすると

$$P(A_e \cap A_f \mid A_a \cap A_r) = 1 - \frac{1}{2^{k}-1}\frac{2^{k}-2+2^{k\tau'}}{2^{\tilde{\tau}}} \geq 1 - \frac{1}{2^{k-1}} \tag{6.1}$$

が成り立つ．等号成立は $\tau_1 = \cdots = \tau_k = 1$ のとき．

(証明)　式 (6.1) の不等式について

$$\begin{aligned}
&\frac{1}{2^{k-1}} - \frac{1}{2^{k}-1}\frac{2^{k}-2-2^{k\tau'}}{2^{\tilde{\tau}}} \\
&\geq \frac{1}{2^{k-1}} - \frac{1}{2^{k}-1}\frac{2^{k}-2-2^{k\tau'}}{2^{k\tau'}} \\
&= \left(\frac{1}{2^{k}} - \frac{1}{2^{k\tau'}}\right)\left(1 - \frac{1}{2^{k}-1}\right) \\
&\geq 0
\end{aligned}$$

が成り立つ．

この定理により，従来の評価が最良の結果であり，確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ の下限を $1 - \frac{1}{2^{k-1}}$ より大きくはできないことがわかる．また，$\tau_1 = \cdots = \tau_k = 1$ のときに確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ が最小になることから，n を構成する素数が全て $(2 \cdot$ 奇数 $+ 1)$ の形をしているときに最も計算量を要することになる．

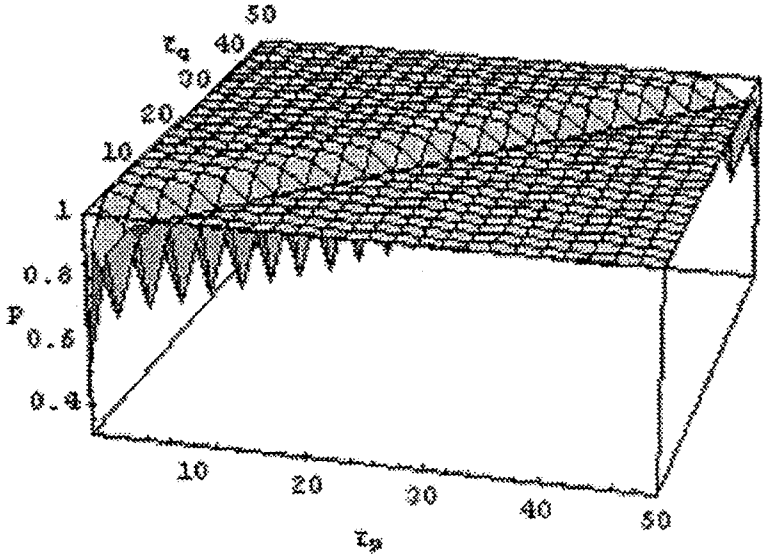

図1　τ_p および τ_q と確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ の関係

n を構成する素数と確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ の関係をグラフ（図 1）に示す．ここでは，n は 2 つの素数の積で $n = pq$ と表されている場合を考え，$p-1 = 2^{\tau_p}\sigma_p$, $q-1 = 2^{\tau_q}\sigma_q$ （$\tau_p, \tau_q \geq 1$，σ_p, σ_q は奇数）とする．このときの確率は，従来の評価では式 (3.3) より

$$P(A_e \cap A_f \mid A_a \cap A_r) \geq \frac{1}{2}$$

である．一方，本研究の精密な評価では，式 (5.1) より

$$P(A_e \cap A_f \mid A_a \cap A_r) = 1 - \frac{1}{3}\frac{2 + 2^{2\min(\tau_p,\tau_q)}}{2^{\tau_p+\tau_q}}$$

となる．図からわかるように，$\tau_p = \tau_q = 1$ のとき確率 1/2 で最小となっている．また，一般に $\tau_p = \tau_q$ のときに確率 $P(A_e \cap A_f \mid A_a \cap A_r)$ は比較的小さくなり，$\tau_p \neq \tau_q$ で急激に確率 1 に近づくことがわかる．

また，n が 2 つの素数の積で表される場合に限り，式 (3.1) の $\gcd(a, n) = 1$ となるような n 未満の数 a が得られる確率は簡単にすることができる．

補題 6.2　$n = pq$ （p, q は素数）とする．十分大きい n に対して，$\gcd(a, n) = 1$ となるような n 未満の数 a が得られる確率 $P(A_a)$ は

$$\forall \varepsilon > 0, \quad P(A_a) = \frac{\varphi(n)}{n} \geq \frac{1}{2} - \varepsilon$$

で与えられる．

（証明）　$n = pq$ （p, q は素数）とすると，Euler の関数は $\varphi(n) = n\left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right)$ であるから

$$\frac{\varphi(n)}{n} = \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right) \geq \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{q}\right) = \frac{1}{2}\left(1 - \frac{1}{q}\right)$$

となる．q についても同様にすると

$$\frac{\varphi(n)}{n} \geq \frac{1}{2}\left(1-\frac{1}{p}\right),\ \frac{1}{2}\left(1-\frac{1}{q}\right)$$

が得られる．ここで，$n \to \infty \Longleftrightarrow (p \to \infty$ または $q \to \infty)$ に注意すると，任意の ε に対して，n を十分大きくすると

$$\frac{\varphi(n)}{n} > \frac{1}{2} - \varepsilon$$

となる．

系 6.3　$n = pq$ (p, q は素数) とする．$p-1 = 2^{\tau_p}\sigma_p$, $q-1 = 2^{\tau_q}\sigma_q$, $\tau' = \min(\tau_p, \tau_q)$ (但し，$\tau_p, \tau_q \geq 1$, σ_p, σ_q は奇数) とすると，アルゴリズムを 1 回実行して素因数分解に成功する確率 P_S は

$$P_S \geq \frac{\alpha}{2\log_2 n}\left(1 - \frac{1}{3}\frac{2+2^{2\tau'}}{2^{\tau_p+\tau_q}}\right)$$

であり，十分大きい確率で確率で正しく素因数分解するために必要なアルゴリズムの実行回数 N は

$$\forall \varepsilon > 0, \quad N \geq \frac{2\log(1/\varepsilon)}{\alpha\left(1 - \frac{1}{3}\frac{2+2^{2\tau'}}{2^{\tau_p+\tau_q}}\right)}\log_2 n$$

である．

参考文献

[1] P.W.Shor, Algorithms for quantum computation: Discrete log and factoring, Proc. of the 35th Annual IEEE Symp. on Foundations of Computer Science, pp.124-134, 1994.

[2] P.W.Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM Journal on Computing, vol.26, no.5, pp.1484-1509, 1997.

[3] 上坂吉則, 量子コンピュータの基礎数理, コロナ社, 2000.

[4] A.Ekert and R.Jozsa, Quantum computation and Shor's factoring algorithm, Rev.Mod.Phys., 68, 3, pp.733-753, 1996.

[5] G.H.Hardy and E.M.Wright, An introduction to the Theory of Numbers, Fifth Edition, Oxford Science Publications, 1979.

数理解析研究所講究録 1452 巻 2005 年 215-226

Tsallis の relative entropy と relative operator entropy

柳　研二郎 (Kenjiro Yanagi)
山口大・工
(Department of Applied Science, Yamaguchi University)
古市　茂 (Shigeru Furuichi)
山口東京理科大
(Tokyo University of Science in Yamaguchi)
栗山　憲 (Ken Kuriyama)
山口大・工
(Department of Applied Science, Yamaguchi University)

1 Classical Tsallis relative entropy

Definition 1 次で定義される $S_q(X)$ を *Tsallis entropy* という.

$$S_q(X) = -\sum_x p(x)^q \ln_q p(x),$$

ただし $p(x) = P(X = x)$ は *random variable* X の *probability distribution* でありまた

$$\ln_q(x) = \frac{x^{1-q} - 1}{1-q}, \quad x \geq 0,\ q \geq 0$$

とする.

このとき

$$\lim_{q\to 1} S_q(X) = S(X) = -\sum_x p(x) \log p(x)$$

である. すなわち Shannon entropy に収束する.

Definition 2 $A = \{a_1, a_2, \ldots, a_n\}$, $B = \{b_1, b_2, \ldots, b_n\}$ を 2 つの *probability distribution* とする. ただし $a_j > 0, b_j > 0$ を仮定する. このとき次で定義される $D_q(A|B)$ を *Tsallis relative entropy* という.

$$D_q(A|B) = -\sum_{j=1}^n a_j \ln_q \frac{b_j}{a_j} = \frac{1 - \sum_{j=1}^n a_j^q b_j^{1-q}}{1-q},$$

ただし $0 \ln_q \infty = 0$ と定義する.

このとき

$$\lim_{q\to 1} D_q(A|B) = D_1(A|B) = \sum_{j=1}^n a_j \log \frac{a_j}{b_j}$$

である. すなわち Kullback-Leibler information に収束する.

Proposition 1 *Tsallis relative entropy* の性質は次の通りである.

(1)(Nonnegativity): $D_q(A|B) \geq 0$.

(2)(Symmetry):

$$D_q(a_{\pi(1)},\ldots,a_{\pi(n)}|b_{\pi(1)},\ldots,b_{\pi(n)}) = D_q(a_1,\ldots,a_n|b_1,\ldots,b_n).$$

(3)(Possibility of extension):

$$D_q(a_1,\ldots,a_n,0|b_1,\ldots,b_n,0) = D_q(a_1,\ldots,a_n|b_1,\ldots,b_n).$$

(4)(Pseudoadditivity):

$$\begin{aligned} & D_q(A^{(1)}\times A^{(2)}|B^{(1)}\times B^{(2)}) \\ = \; & D_q(A^{(1)}|B^{(1)}) + D_q(A^{(2)}|B^{(2)}) + (q-1)D_q(A^{(1)}|B^{(1)})D_q(A^{(2)}|B^{(2)}), \end{aligned}$$

ただし

$$A^{(1)}\times A^{(2)} = \{a_j^{(1)}a_j^{(2)}|a_j^{(1)}\in A^{(1)}, a_j^{(2)}\in A^{(2)}\},$$
$$B^{(1)}\times B^{(2)} = \{b_j^{(1)}b_j^{(2)}|b_j^{(1)}\in B^{(1)}, b_j^{(2)}\in B^{(2)}\}.$$

(5)(Joint convexity): $0\leq\lambda\leq 1$, $q\geq 0$ とする. $A^{(i)} = \{a_j^{(i)}\}$, $B^{(i)} = \{b_j^{(i)}\}$, $(i=1,2)$ に対して次が成り立つ.

$$\begin{aligned} & D_q(\lambda A^{(1)} + (1-\lambda)A^{(2)}|\lambda B^{(1)} + (1-\lambda)B^{(2)}) \\ \leq \; & \lambda D_q(A^{(1)}|B^{(1)}) + (1-\lambda)D_q(A^{(2)}|B^{(2)}). \end{aligned}$$

(6)(Strong additivity):

$$\begin{aligned} & D_q(a_1,\ldots,a_{i-1},a_{i_1},a_{i_2},a_{i+1},\ldots,a_n|b_1,\ldots,b_{i-1},b_{i_1},b_{i_2},b_{i+1},\ldots,b_n) \\ = \; & D_q(a_1,\ldots,a_n|b_1,\ldots,b_n) + b_i^{1-q}a_i^q D_q(\frac{a_{i_1}}{a_i},\frac{a_{i_2}}{a_i}|\frac{b_{i_1}}{b_i},\frac{b_{i_2}}{b_i}), \end{aligned}$$

ただし $a_i = a_{i_1} + a_{i_2}, b_i = b_{i_1} + b_{i_2}$.

Proof.

(1): $-\ln_q(x)$ は convex function であるので次を得る.

$$D_q(A|B) \equiv -\sum_{j=1}^n a_j \ln_q \frac{b_j}{a_j} \geq -\ln_q(\sum_{j=1}^n a_j\frac{b_j}{a_j}) = 0.$$

(2), (3) 及び (4): 明らか.
(5): [3] の generalized log-sum inequality より任意の $\alpha_i,\beta_i\geq 0 (i=1,2,\ldots,n), q\geq 0$ に対して次の不等式が成り立つ.

$$\sum_{i=1}^n \alpha_i \ln_q(\frac{\beta_i}{\alpha_i}) \leq (\sum_{i=1}^n \alpha_i)\ln_q(\frac{\sum_{i=1}^n\beta_i}{\sum_{i=1}^n\alpha_i}). \tag{1}$$

これを用いればよい.
(6): $q\geq 0$ に対して function L_q を次のように定義する.

$$L_q(x,y) \equiv -x\ln_q\frac{y}{x}.$$

また次の記号を導入する.

$$a_{i_1} = a_i(1-s),\ a_{i_2} = a_i s,\ b_{i_1} = b_i(1-t),\ b_{i_2} = b_i t.$$

このとき

$$L_q(x_1x_2, y_1y_2) = x_2 L_q(x_1, y_1) + x_1 L_q(x_2, y_2) + (q-1)L_q(x_1, y_1)L_q(x_2, y_2).$$

これを用いればよい. q.e.d.

Remark 1 次が成り立つ.

(1) *Proposition 1 の (1)* より $S_q(A) \leq \ln_q n$.

(2) *Proposition 1 の (4)* より

$$S_q(A^{(1)} \times A^{(2)}) = S_q(A^{(1)}) + S_q(A^{(2)}) + (1-q)S_q(A^{(1)})S_q(A^{(2)}).$$

(3) *Proposition 1 の (5)* より

$$S_q(\lambda A^{(1)} + (1-\lambda)A^{(2)}) \geq \lambda S_q(A^{(1)}) + (1-\lambda)S_q(A^{(2)}).$$

(4) *Proposition 1 の (6)* より

$$\begin{aligned} & S_q(a_1, \ldots, a_{i-1}, a_{i_1}, a_{i_2}, a_{i+1}, \ldots, a_n) \\ = \ & S_q(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n) + a_i^q S_q(\frac{a_{i_1}}{a_i}, \frac{a_{i_2}}{a_i}). \end{aligned}$$

$\mathcal{A}, \mathcal{B}$ を 2 つの finite alphabet sets とする. $W = \{W_{ji}\}, (i = 1, \ldots, n, j = 1, \ldots, m)$ を $\mathcal{A}$ から $\mathcal{B}$ への transition probability matrix とする. すなわち $\sum_{j=1}^m W_{ji} = 1, i = 1, 2, \ldots, n$ である. また $A = \{a_i^{(in)}\}, B = \{b_i^{(in)}\}$ を $\mathcal{A}$ における 2 つの異なる probability distribution とする. このとき $\mathcal{B}$ における probability distribution $WA = \{a_j^{(out)}\}, WB = \{b_j^{(out)}\}$ は次のように定義される.

$$a_j^{(out)} = \sum_{i=1}^n a_i^{(in)} W_{ji}, \quad b_j^{(out)} = \sum_{i=1}^n b_i^{(in)} W_{ji}.$$

Proposition 2 任意の $q \geq 0$ に対して次が成り立つ.

$$D_q(WA|WB) \leq D_q(A|B).$$

Proof. generalized log-sum inequality (1) を適用して次を得る.

$$\begin{aligned} D_q(WA|WB) &= -\sum_{j=1}^m a_j^{(out)} \ln_q \frac{b_j^{(out)}}{a_j^{(out)}} \\ &= -\sum_{j=1}^m \sum_{i=1}^n a_j^{(in)} W_{ji} \ln_q \frac{\sum_{i=1}^n b_i^{(in)} W_{ji}}{\sum_{i=1}^n a_i^{(in)} W_{ji}} \\ &\leq -\sum_{j=1}^m \sum_{i=1}^n a_i^{(in)} W_{ji} \ln_q \frac{b_i^{(in)} W_{ji}}{a_i^{(in)} W_{ji}} \\ &= -\sum_{i=1}^n a_i^{(in)} \ln_q \frac{b_i^{(in)}}{a_i^{(in)}} \\ &= D_q(A|B). \end{aligned}$$

q.e.d.

2 Quantum Tsallis relative entropy

Definition 3 ρ,σ を2つの *density operators* とする. $0 \le q < 1$ に対して 次で定義される $D_q(\rho|\sigma)$ を *quantum Tsallis relative entropy* という.

$$D_q(\rho|\sigma) = \frac{1 - Tr[\rho^q \sigma^{1-q}]}{1-q}.$$

また Umegaki [14] によって quantum relative entropy が次のように定義されていることに注意する.

$$U(\rho|\sigma) = Tr[\rho(\log\rho - \log\sigma)].$$

$D_q(\rho|\sigma)$ を $0 \le q \le 2$ に拡張して定義すると都合がよい. 即ち $0 \le q < 1$ に対しては $D_{2-q}(\rho|\sigma)$, $1 < q \le 2$ に対しては $D_q(\rho|\sigma)$ が定義されるとする.

Proposition 3 次の (1), (2) が成り立つ.

(1) $D_q(\rho|\sigma) \le U(\rho|\sigma) \le D_{2-q}(\rho|\sigma)$ for $0 \le q < 1$.

(2) $D_{2-q}(\rho|\sigma) \le U(\rho|\sigma) \le D_q(\rho|\sigma)$ for $1 < q \le 2$.

Proof. 任意の $x > 0, t > 0$ に対して

$$\frac{1 - x^{-t}}{t} \le \log x \le \frac{x^t - 1}{t}$$

が成り立つので任意の $a, b, t > 0$ に対して次の不等式を得る.

$$a(\frac{1 - a^{-t}b^t}{t}) \le a\log\frac{a}{b} \le a(\frac{a^t b^{-t} - 1}{t}). \tag{2}$$

$\rho = \sum_i \lambda_i P_i$, $\sigma = \sum_i \mu_j Q_j$ をスペクトル分解とすると $\sum_i P_i = \sum_j Q_j = I$ だから次の不等式を得る.

$$\begin{aligned} & Tr[\frac{\rho^{1+t}\sigma^{-t} - \rho}{t} - \rho(\log\rho - \log\sigma)] \\ = & \sum_{i,j} Tr[P_i\{\frac{\rho^{1+t}\sigma^{-t} - \rho}{t} - \rho(\log\rho - \log\sigma)\}Q_j] \\ = & \sum_{i,j} (\frac{1}{t}\lambda_i^{1+t}\mu_j^{-t} - \frac{1}{t}\lambda_i - \lambda_i\log\lambda_i + \lambda_i\log\mu_j) Tr[P_i Q_j] \ge 0. \end{aligned}$$

最後の不等式は (2) の右側の不等式から得られる. したがって

$$Tr[\rho(\log\rho - \log\sigma)] \le \frac{1}{t} Tr[\rho^{1+t}\sigma^{-t} - \rho].$$

左側の不等式も同様にして得られる. したがって $1 - q = t(> 0)$ とおけば Proposition 5 の (1) を得る. また $q - 1 = t(> 0)$ とおけば Proposition 5 の (2) を得る. q.e.d.

ρ,σ: strictly positive operators に対して relative operator entropy $S(\rho|\sigma)$ が Fujii-Kamei [4] によって次のように定義された.

$$S(\rho|\sigma) = \rho^{1/2}\log(\rho^{-1/2}\sigma\rho^{-1/2})\rho^{1/2}.$$

ここで ρ,σ が commutative のとき

$$U(\rho|\sigma) = -Tr[S(\rho|\sigma)]$$

であることは明らかである. 一方 Hiai-Petz [8] によって次の関係が成り立つことが示された.

$$U(\rho|\sigma) \leq -Tr[S(\rho|\sigma)].$$

また Yanagi-Furuichi-Kuriyama [15] によって $0 \leq q < 1$ に対して Tsallis relative operator entropy $T_q(\rho|\sigma)$ が次のように定義された.

$$T_q(\rho|\sigma) = \frac{\rho^{1/2}(\rho^{-1/2}\sigma\rho^{-1/2})^{1-q}\rho^{1/2} - \rho}{1-q}.$$

このとき次が成り立つ.

$$\lim_{q\to 1} T_q(\rho|\sigma) = S(\rho|\sigma).$$

また ρ,σ が commutative のとき次が成り立つことは明らかである.

$$D_q(\rho|\sigma) = -Tr[T_q(\rho|\sigma)].$$

さらに次が成り立つ.

$$\lim_{q\to 1} D_q(\rho|\sigma) = U(\rho|\sigma).$$

Theorem 1 ρ,σ を *strictly positive density operators* とする. このとき $0 \leq q < 1$ に対して次が成り立つ.

$$D_q(\rho|\sigma) \leq -Tr[T_q(\rho|\sigma)].$$

Proof. α-power mean $\sharp_\alpha$ は次のように定義される.

$$A\sharp_\alpha B \equiv A^{1/2}(A^{-1/2}BA^{-1/2})^\alpha A^{1/2}.$$

[8] の Theorem 3.4 より任意の $\alpha \in [0,1]$ に対して

$$Tr[e^A \sharp_\alpha e^B] \leq Tr[e^{(1-\alpha)A+\alpha B}].$$

$A = \log\rho$, $B = \log\sigma$ とおくと次を得る.

$$Tr[\rho\sharp_\alpha\sigma] \leq Tr[e^{\log\rho^{1-\alpha}+\log\sigma^\alpha}].$$

ここで Golden-Thompson inequality より任意の Hermitian operators A,B に対して $Tr[e^{A+B}] \leq Tr[e^A e^B]$ が成り立つので次が得られる.

$$Tr[e^{\log\rho^{1-\alpha}+\log\sigma^\alpha}] \leq Tr[e^{\log\rho^{1-\alpha}}e^{\log\sigma^\alpha}] = Tr[\rho^{1-\alpha}\sigma^\alpha].$$

したがって

$$tr[\rho^{1/2}(\rho^{-1/2}\sigma\rho^{-1/2})^\alpha\rho^{1/2}] \leq Tr[\rho^{1-\alpha}\sigma^\alpha].$$

ここで $\alpha = 1-q$ とおけばよい. q.e.d.

Corollary 1 (Hiai-Petz [8]) ρ,σ を *strictly positive density operators* とするとき次が成り立つ.

$$Tr[\rho(\log\rho - \log\sigma)] \leq Tr[\rho\log(\rho^{1/2}\sigma^{-1}\rho^{1/2})].$$

Proposition 4 ρ,σ を *density operators* とする. $0 \leq q < 1$ に対して次の (1) ~ (4) が成り立つ.

(1)(Nonnegativity): $D_q(\rho|\sigma) \geq 0$.

(2)(Pseudoadditivity):

$$D_q(\rho_1 \otimes \rho_2|\sigma_1 \otimes \sigma_2) = D_q(\rho_1|\sigma_1) + D_q(\rho_2|\sigma_2) + (q-1)D_q(\rho_1|\sigma_1)D_q(\rho_2|\sigma_2).$$

(3)(Joint convexity):

$$D_q(\sum_j \lambda_j\rho_j|\sum_j \lambda_j\sigma_j) \leq \sum_j \lambda_j D_q(\rho_j|\sigma_j).$$

(4)(Invariance): unitary transformation U に対して

$$D_q(U\rho U^*|U\sigma U^*) = D_q(\rho|\sigma).$$

Proof. 任意の $x \geq 0, y \geq 0, 0 \leq q < 1$ に対して

$$f(q;x,y) \equiv \frac{x - x^q y^{1-q}}{1-q} - (x-y) \geq 0$$

が成り立つので次を得る.

$$D_q(\rho|\sigma) \geq Tr[\rho - \sigma].$$

ρ,σ が density operators であるので (1) が得られる.
(2): 直接の計算で得られる.
(3): Lieb's theorem より任意の operator Z と任意の $0 \leq t \leq 1$ に対して functional $f(A,B) \equiv Tr[Z^*A^tZB^{1-t}]$ は positive operators A, B について jointly concave である.
(4): Stone-Weierstrass approximation theorem を用いると明らかである. q.e.d.

Theorem 2 任意の *trace-presearving completely positive linear map* Φ と任意の *density operators* ρ,σ と $0 \leq q < 1$ に対して次が成り立つ.

$$D_q(\Phi(\rho)|\Phi(\sigma)) \leq D_q(\rho|\sigma).$$

Proof. [9] と同様にすればよい.

まず composite system AB において partial trace Tr_B に対して $D_q(\rho|\sigma)$ の monotonicity を証明する. ρ^{AB},σ^{AB} を composite system AB における density operators とする. [10] より次のような unitary operators U_j と probability p_j が存在する.

$$\rho^A \otimes \frac{1}{n} = \sum_j p_j (I \otimes U_j)\rho^{AB}(I \otimes U_j)^*,$$

ただし n は system B の次元, I は system B の identity operator, $\rho^A = Tr_B[\rho^{AB}], \sigma^A = Tr_B[\sigma^{AB}]$ である. Tsallis relative entropy の joint convexity と unitary invarinace より次の関係式を得る.

$$\begin{aligned} & D_q(\rho^A \otimes \frac{1}{n}) \\ \leq & \sum_j p_j D_q((I \otimes U_j)\rho^{AB}(I \otimes U_j)^*|(I \otimes U_j)\sigma^{AB}(I \otimes U_j)^*) \\ = & \sum_j p_j D_q(\rho^{AB}|\sigma^{AB}) \\ = & D_q(\rho^{AB}|\sigma^{AB}). \end{aligned}$$

ここで

$$D_q(\rho^A \otimes \frac{1}{n}|\sigma^A \otimes \frac{1}{n}) = D_q(\rho^A|\sigma^A),$$

より次を得る.

$$D_q(Tr_B[\rho^{AB}]|Tr_B[\sigma^{AB}]) \leq D_q(\rho^{AB}|\sigma^{AB}). \tag{3}$$

[11] より任意の trace presearving completely positive linear map Φ は total system AB 上の unitary operator U^{AB} と subsystem B 上の projection(pure state) P^B でつぎのように表現される.

$$\Phi(\rho^A) = Tr_B[U^{AB}(\rho^A \otimes P^B)(U^{AB})^*].$$

したがって (3) と $D_q(\rho|\sigma)$ の unitary invariance を再度用いると次を得る.

$$\begin{aligned} & D_q(\Phi(\rho^A)|\Phi(\sigma^A)) \\ \leq\ & D_q(U^{AB}(\rho^A \otimes P^B)(U^{AB})^*|U^{AB}(\sigma^A \otimes P^B)(U^{AB})^*) \\ =\ & D_q(\rho^A \otimes P^B|\sigma^A \otimes P^B) \\ =\ & D_q(\rho^A|\sigma^A). \end{aligned}$$

q.e.d.

Corollary 2 任意の *trace-presearving completely positive linear map* Φ と 任意の *density operator* ρ と $0 \leq q < 1$ に対して次が成り立つ.

$$H_q(\Phi(\rho)) \geq H_q(\rho),$$

ただし

$$H_q(X) = \frac{Tr[X^q] - 1}{1 - q}$$

は *quantum Tsallis entropy* である.

3 Generalized Tsallis relative entropy

Definition 4 任意の *positive operators* A, B と任意の実数 $q \in [0, 1)$ に対して $D_q(A||B)$ を次のように定義する.

$$D_q(A||B) = \frac{Tr[A] - Tr[A^q B^{1-q}]}{1 - q}.$$

Lieb's concavity theorem より次が成り立つ.

$$D_q(\sum_j \lambda_j A_j || \sum_j \lambda_j B_j) \leq \sum_j \lambda_j D_q(A_j||B_j), \tag{4}$$

ただし $\lambda_j > 0 (\sum_j \lambda_j = 1)$ である.

Theorem 3 任意の *positive operators* A_1, A_2, B_1, B_2 と任意の $0 \leq q < 1$ に対して次の *subadditivity* が成り立つ.

$$D_q(A_1 + A_2||B_1 + B_2) \leq D_q(A_1||B_1) + D_q(A_2||B_2). \tag{5}$$

Proof. 任意の実数 α, β, 任意の positive operators A, B に対して次が成り立つことに注意する.

$$D_q(\alpha A||\beta B) = \alpha D_q(A||B) - \alpha \ln_q \frac{\beta}{\alpha} Tr[A^q B^{1-q}].$$

(4) より任意の positive operators X_1, X_2, Y_1, Y_2 と任意の $\lambda_1, \lambda_2 (\lambda_1 + \lambda_2 = 1)$ に対して次を得る.

$$D_q(\lambda_1 X_1 + \lambda_2 X_2||\lambda_1 Y_1 + \lambda_2 Y_2) \leq \lambda_1 D_q(X_1||Y_1) + \lambda_2 D_q(X_2||Y_2).$$

ここで $A_i = \lambda_i X_i, B_i = \lambda_i Y_i$ $(i = 1, 2)$ とおくと

$$D_q(A_1 + A_2||B_1 + B_2) \leq \lambda_1 D_q(\frac{A_1}{\lambda_1}||\frac{B_1}{\lambda_1}) + \lambda_2 D_q(\frac{A_2}{\lambda_2}||\frac{A_2}{\lambda_2}).$$

したがって (5) より結論が得られる. q.e.d.

Theorem 4 任意の *positive operators* A, B と任意の $0 \leq q < 1$ に対して次の不等式が成り立つ.

$$D_q(A||B) \geq \frac{Tr[A] - (Tr[A])^q (Tr[B])^{1-q}}{1-q}.$$

Proof. Holder' inequality より $Tr[|X|^s] < \infty, Tr[|Y|^t] < \infty$ を満たす任意の bounded linear operators X, Y と $1/s + 1/t = 1$ を満たす任意の $1 < s < \infty, 1 < t < \infty$ に対して次が成り立つ.

$$|Tr[XY]| \leq Tr[|X|^s]^{1/s} Tr[|Y|^t]^{1/t}.$$

ここで $X = A^q, Y = B^{1-q}$, $s = 1/q, t = 1/(1-q)$ とおくと

$$Tr[A^q B^{1-q}] \leq (Tr[A])^q (Tr[B])^{1-q}$$

が得られ結論に至る. q.e.d.

4 Tsallis relative operator entropy

Hilbert space H 上の bounded linear operator T は 任意の $x \in H$ に対して $(Tx, x) \geq 0$ を満たすとき positive といい $T \geq 0$ と表わす. また T が invertible かつ positive であるとき strictly positive といい $T > 0$ と表わす. Tsallis relative operator entropy は次のように定義される.

Definition 5 ([5]) $A > 0, B > 0$ と $0 < \lambda \leq 1$ に対して

$$T_\lambda(A|B) = \frac{A^{1/2}(A^{-1/2} B A^{-1/2})^\lambda A^{1/2} - A}{\lambda}$$

を A と B の間の *Tsallis relative operator entropy* と定義する.

$T_\lambda(A|B)$ の基本的性質は [5] で与えられている. ここではTsallis relative operator entropy を用いて Shannon type の operator inequality とその逆の inequality を与える.

Theorem 5 $\{A_1, A_2, \ldots, A_n\}$ と $\{B_1, B_2, \ldots, B_n\}$ を *Hilbert space* H 上の *strictly positive operator* からなる 2 つの列で $\sum_{j=1}^n A_j = \sum_{j=1}^n B_j = I$ を満たすとする. このとき次が成り立つ.

$$0 \geq \sum_{j=1}^n T_\lambda(A_j|B_j) \geq \frac{(\sum_{j=1}^n A_j B_j^{-1} A_j)^{-\lambda} - I}{\lambda}.$$

証明にあたり次の Lemma を必要とする.

Lemma 1 $t>0$ を固定すると次の $\lambda\,(0<\lambda\leq 1)$ に関する *inequality* が成り立つ.

$$\frac{t^{\lambda}-1}{\lambda}\leq t-1.$$

Proof. $t=1$ のときは明らか. $t\neq 1$ とする. $F(\lambda)=\lambda(t-1)-t^{\lambda}+1$ とおく. このとき $F'(\lambda)=t-1-t^{\lambda}\log t$ かつ $F''(\lambda)=-t^{\lambda}(\log t)^2<0$. したがって $F(\lambda)$ は concave function である. $F(0)=F(1)=0$ だから結論を得る. q.e.d.

Proof of Theorem 5. Lemma 1 より

$$\begin{aligned}\frac{A^{1/2}(A^{-1/2}BA^{-1/2})^{\lambda}A^{1/2}-A}{\lambda} &= A^{1/2}\frac{(A^{-1/2}BA^{-1/2})^{\lambda}-I}{\lambda}A^{1/2}\\ &\leq A^{1/2}(A^{-1/2}BA^{-1/2}-I)A^{1/2}\\ &= B-A,\end{aligned}$$

ただし $A>0, B>0$ かつ $0<\lambda\leq 1$. したがって

$$\begin{aligned}\sum_{j=1}^{n}T_{\lambda}(A_j|B_j) &= \sum_{j=1}^{n}\frac{A_j^{1/2}(A_j^{-1/2}B_jA_j^{-1/2})^{\lambda}A_j^{1/2}-A_j}{\lambda}\\ &\leq \sum_{j=1}^{n}(B_j-A_j)=0.\end{aligned}$$

別の inequality も証明する. $f(x)=-x^{-\lambda}$, $C_j=A_j^{1/2}$ かつ $X_j=A_j^{1/2}B_j^{-1}A_j^{1/2}$ とおくことにより Furuta [6] の Proposition 3.1 を適用すると次を得る.

$$-(\sum_{j=1}^{n}A_j^{1/2}(A^{1/2}B_j^{-1}A_j^{1/2})A_j^{1/2})^{-\lambda}\geq-\sum_{j=1}^{n}A_j^{1/2}(A_j^{1/2}B_j^{-1}A_j^{1/2})^{-\lambda}A_j^{1/2}.$$

したがって

$$(\sum_{j=1}^{n}A_jB_j^{-1}A_j)^{-\lambda}\leq\sum_{j=1}^{n}A_j^{1/2}(A_j^{-1/2}B_jA_j^{-1/2})^{\lambda}A_j^{1/2}.$$

ゆえに証明を完了する. q.e.d.

Theorem 5 の corolary として Furuta [6] によって得られた Shannon inequalty とその逆の inequality の operator 版に相当するものが得られる.

Corollary 3 (Furuta [6]) $\{A_1,A_2,\ldots,A_n\}$ と $\{B_1,B_2,\ldots,B_n\}$ を *Hilbert space* H 上の *strictly positive operator* からなる 2 つの列で $\sum_{j=1}^{n}A_j=\sum_{j=1}^{n}B_j=I$ を満たすとする. このとき次が成り立つ.

$$0\geq\sum_{j=1}^{n}A_j^{1/2}(\log A_j^{-1/2}B_jA_j^{-1/2})A_j^{1/2}\geq-\log[\sum_{j=1}^{n}A_jB_j^{-1}A_j].$$

Corollary 3 は [6] の Corollary 2.4 の 1 部分である. 次の section では一般化された Tsallis relative operator entropy を新しく定義し, その性質を調べる.

5 Generalized Tsallis relative operator entropy

relative operator entropy と related operator entropy を思い出そう.

Definition 6 $A>0, B>0$ に対して

$$S(A|B)=A^{1/2}(\log A^{-1/2}BA^{-1/2})A^{1/2}$$

は A と B の間の *relative operator entropy* と定義される. これは *Fujii and Kamei* [4] によって定義された. また $A>0. B>0$ と $\lambda\in\mathbb{R}$ に対して, *generalized relative operator entropy* が *Furuta* [6] によって次のように定義された.

$$S_\lambda(A|B)=A^{1/2}(A^{-1/2}BA^{-1/2})^\lambda(\log A^{-1/2}BA^{-1/2})A^{1/2},$$

$$A\natural_\lambda B=A^{1/2}(A^{-1/2}BA^{-1/2})^\lambda A^{1/2}.$$

特に $S(A|B)=S_0(A|B), A\natural_0 B=A, A\natural_1 B=B$ となることに注意する.

Tsallis relative operator entropy を次のように一般化する.

Definition 7 $A>0, B>0$, $\lambda,\mu\in\mathbb{R}, \lambda\neq 0$, $k\in\mathbb{Z}$ に対して

$$\tilde{T}_{\mu,k,\lambda}(A|B)=\frac{A\natural_{\mu+k\lambda}B-A\natural_{\mu+(k-1)\lambda}B}{\lambda}$$

を *generalized Tsallis relative operator entropy* と定義する. 特に $\lambda\neq 0$ に対して

$$\tilde{T}_{0,1,\lambda}(A|B)=\frac{A\natural_\lambda B-A\natural_0 B}{\lambda}=\frac{A^{1/2}(A^{-1/2}BA^{-1/2})^\lambda A^{1/2}-A}{\lambda}=T_\lambda(A|B).$$

$S_{\mu\pm k\lambda}(A|B), S_{\mu\pm(k+1)\lambda}(A|B)$ と $\tilde{T}_{\mu,k+1,\pm\lambda}(A|B)$ の間の関係について調べる.

Proposition 5 $\lambda>0, \mu\in\mathbb{R}$, $k=0,1,2,\ldots$ のとき次が成り立つ.

(1) $S_{\mu-(k+1)\lambda}(A|B)\leq\tilde{T}_{\mu,k+1,-\lambda}(A|B)\leq S_{\mu-k\lambda}(A|B)$.

(2) $S_{\mu+k\lambda}(A|B)\leq\tilde{T}_{\mu,k+1,\lambda}(A|B)\leq S_{\mu+(k+1)\lambda}(A|B)$.

Proof. $\lambda>0, \mu\in\mathbb{R}$, $k=0,1,2,\ldots$ のとき 任意の $t>0$ に対して次の inequalities を得る.

$$t^{\mu-(k+1)\lambda}\log t\leq\frac{t^{\mu-(k+1)\lambda}-t^{\mu-k\lambda}}{-\lambda}\leq t^{\mu-k\lambda}\log t,$$

$$t^{\mu+k\lambda}\log t\leq\frac{t^{\mu+(k+1)\lambda}-t^{\mu+k\lambda}}{\lambda}\leq t^{\mu+(k+1)\lambda}\log t.$$

したがって t を $A^{-1/2}BA^{-1/2}$ で置き換え両辺に $A^{1/2}$ をかけることにより目標の式を得る. q.e.d.

$k=0$ 又は 1 のとき次のようになる.

Corollary 4 $A>0, B>0, \mu\in\mathbb{R}$, $\lambda>0$ のとき

$$\begin{aligned}&S_{\mu-2\lambda}(A|B)\leq\tilde{T}_{\mu,2,-\lambda}(A|B)\leq S_{\mu-\lambda}(A|B)\\ \leq\ &\tilde{T}_{\mu,1,-\lambda}(A|B)\leq S_\mu(A|B)\leq\tilde{T}_{\mu,1,\lambda}(A|B)\\ \leq\ &S_{\mu+\lambda}(A|B)\leq\tilde{T}_{\mu,2,\lambda}(A|B)\leq S_{\mu+2\lambda}(A|B).\end{aligned}$$

特に $\mu=0,\lambda=1$ とおくと次を得る.

Corollary 5 $A>0,B>0$ に対して

$$\begin{aligned} & S_{-2}(A|B)\le \tilde{T}_{0,2,-1}(A|B)\le S_{-1}(A|B)\\ \le\ & \tilde{T}_{0,1,-1}(A|B)\le S_0(A|B)\le \tilde{T}_{0,1,1}(A|B)\\ \le\ & S_1(A|B)\le \tilde{T}_{0,2,1}(A|B)\le S_2(A|B). \end{aligned}$$

書き直すと

$$\begin{aligned} & S_{-2}(A|B)\le AB^{-1}A-AB^{-1}AB^{-1}A\le S_{-1}(A|B)\\ \le\ & A-AB^{-1}A\le S(A|B)\le B-A\\ \le\ & S_1(A|B)\le BA^{-1}B-B\le S_2(A|B). \end{aligned}$$

同様にして $\sum_{j=1}^n S_{\mu\pm k\lambda}(A_j|B_j), sum_{j=1}^n S_{\mu\pm(k+1)\lambda}(A_j|B_j)$ と $\sum_{j=1}^n \tilde{T}_{\mu,k+1,\pm\lambda}(A_j|B_j)$ の間の関係を求める. ただし $A_j>0,B_j>0$ は $\sum_{j=1}^n A_j=\sum_{j=1}^n B_j=I$ を満たすと仮定する. $\lambda>0,\mu\in\mathbb{R}$, $k=0,1,2,\dots$ のとき次を得る.

$$\sum_{j=1}^n S_{\mu-(k+1)\lambda}(A_j|B_j)\le\sum_{j=1}^n \tilde{T}_{\mu,k+1,-\lambda}(A_j|B_j)\le\sum_{j=1}^n S_{\mu-k\lambda}(A_j|B_j),$$

$$\sum_{j=1}^n S_{\mu+k\lambda}(A_j|B_j)\le\sum_{j=1}^n \tilde{T}_{\mu,k+1,\lambda}(A_j|B_j)\le\sum_{j=1}^n S_{\mu+(k+1)\lambda}(A_j|B_j).$$

$k=0$ 又は 1 とおくと次を得る.

Corollary 6 $A>0,B>0,\mu\in\mathbb{R}$, $\lambda>0$ のとき

$$\begin{aligned} & \sum_{j=1}^n S_{\mu-2\lambda}(A_j|B_j)\le\sum_{j=1}^n \tilde{T}_{\mu,2,-\lambda}(A_j|B_j)\le\sum_{j=1}^n S_{\mu-\lambda}(A_j|B_j)\\ \le\ & \sum_{j=1}^n \tilde{T}_{\mu,1,-\lambda}(A_j|B_j)\le\sum_{j=1}^n S_{\mu}(A_j|B_j)\le\sum_{j=1}^n \tilde{T}_{\mu,1,\lambda}(A_j|B_j)\\ \le\ & \sum_{j=1}^n S_{\mu+\lambda}(A_j|B_j)\le\sum_{j=1}^n \tilde{T}_{\mu,2,\lambda}(A_j|B_j)\le\sum_{j=1}^n S_{\mu+2\lambda}(A_j|B_j). \end{aligned}$$

特に $\mu=0,\lambda=1$ とおくと [6] の Corollary とは違った結果を得る.

Corollary 7 $\sum_{j=1}^n A_j=\sum_{j=1}^n B_j=I$ を満たす $A_j>0,B_j>0$ に対して次を得る.

$$\begin{aligned} & \sum_{j=1}^n S_{-2}(A_j|B_j)\le\sum_{j=1}^n A_jB_j^{-1}A_j-\sum_{j=1}^n A_jB_j^{-1}A_jB_j^{-1}A_j\\ \le\ & \sum_{j=1}^n S_{-1}(A_j|B_j)\le I-\sum_{j=1}^n A_jB_j^{-1}A_j\le\sum_{j=1}^n S(A_j|B_j)\le 0\\ \le\ & \sum_{j=1}^n S_1(A_j|B_j)\le\sum_{j=1}^n B_jA_j^{-1}B_j-I\le\sum_{j=1}^n S_2(A_j|B_j). \end{aligned}$$

References

[1] S Abe,Monotonic decrease of the quantum nonadditive divergence by projective measurements, Physics Letters A, vol.312, pp.336-338, 2003.

[2] N.Bebiano,J.da Providencia Jr. and R.Lemos, Matrix inequalities in statistical mechanics, Linear Algebra and its Applications, vol.376, pp.265-273, 2004.

[3] L.Borland, A.R.Plastino and C.Tsallis, Information gain within nonextensive thermostatistics, J. Math. Phys., vol.39, pp.6490-6501, 1998, and its Erratum, vol.40, pp.2196, 1999.

[4] J.I.Fujii and E.Kamei, Relative operator entropy in noncommutative information theory, Math. Japonica, vol.34, pp.341-348, 1989.

[5] S.Furuichi, K.Yanagi and K.Kuriyama, Fundamental properties of Tsallis relative entropy, J. Math. Phys. vol.45, pp.4868-4877, 2004.

[6] T.Furuta, Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators, Linear Algebra and its Applications, vol.381, pp.219-235, 2004.

[7] F.Hansen and G.K.Pedersen, Jensen's operator inequality, Bull. London Math. Soc., vol.35, pp.553-564, 2004.

[8] F.Hiai and D.Petz, The Golden-Thompson trace inequality is complemented, Linear Algebra and its Applications, Linear Algebra and its Applications, vol.181, pp.153-185, 1993.

[9] G.Lindblad, Completely positive maps and entropy inequalities, Comm. Math. Phys., vol.40, pp.147-151, 1975.

[10] M.A.Nielsen and I.Chuang, Quantum computation and quantum information, Cambridge Press, 2000.

[11] B.Schmacher, Sending entanglement through noisy quantum channel, Phys. Review A, vol.54, pp.2614-2628, 1996.

[12] C.Tsallis, Possible generalization of Bolzmann-Gibbs statistics, J.Stat.Phys., vol.52, pp.479-487, 1988.

[13] C.Tsallis et al., Nonextensive Statistical Mechanics and Its Applications, edited by S. Abe and Y. Okamoto (Springer-Verlag, Heidelberg,2001); see also the comprehensive list of references at http://tsallis.cat.cbpf.br/biblio.htm.

[14] H.Umegaki, Conditional expectation in an operator algebra, IV (entropy and information), Kodai Math. Sem. Rep., vol.14, pp.59-85, 1962.

[15] K.Yanagi, S.Furuichi and K.Kuriyama, Generalized Shannon inequalities based on Tsallis relative operator entropy, Linear Algebra and its Applications, vol.394, pp.109-118, 2005.