Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Linear Algebra and its Applications 394 (2005) 109–118

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Generalized Shannon inequalities based on Tsallis relative operator entropy

Kenjiro Yanagi [a], Ken Kuriyama [a], Shigeru Furuichi [b,*]

[a] *Department of Applied Science, Faculty of Engineering, Yamaguchi University, Tokiwadai 2-16-1, Ube 755-0811, Japan*

[b] *Department of Electronics and Computer Science, Faculty of Science and Engineering, Tokyo University of Science, Yamaguchi, Daigaku-Dori 1-1-1, Onoda 756-0884, Japan*

Received 2 June 2004; accepted 30 June 2004

Submitted by R.A. Brualdi

Abstract

Tsallis relative operator entropy is defined and then its properties are given. Shannon inequality and its reverse one in Hilbert space operators derived by Furuta [Linear Algebra Appl. 381 (2004) 219] are extended in terms of the parameter of the Tsallis relative operator entropy. Moreover the generalized Tsallis relative operator entropy is introduced and then several operator inequalities are derived.

AMS classification: 47A63; 60E15; 26D15

Keywords: Operator inequality; Tsallis relative operator entropy; Shannon inequality

* Corresponding author.

E-mail addresses: yanagi@yamaguchi-u.ac.jp (K. Yanagi), kuriyama@yamaguchi-u.ac.jp (K. Kuriyama), furuichi@ed.yama.sut.ac.jp (S. Furuichi).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.06.025

1. Introduction

Tsallis entropy

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

was defined in [6] for the probability distribution $p(x)$, where q-logarithm function is defined by $\ln_q(x) \equiv \frac{x^{1-q}-1}{1-q}$ for any nonnegative real numbers x and $q \neq 1$. It is easily seen that Tsallis entropy is one parameter extension of Shannon entropy $S_1(X) \equiv -\sum_x p(x) \log p(x)$ and converges to it as $q \to 1$. The study based on Tsallis type entropies has been developed in mainly statistical physics [7]. In the recent work [1], Tsallis type relative entropy in quantum system, defined by

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

for two density operators ρ and σ (i.e., positive operators with unit trace) and $0 \leqslant q < 1$, was investigated.

On the other hand, the relative operator entropy was defined by Fujii and Kamei [3]. Many important results in operator theory and information theory have been published in the relation to Golden–Thompson inequality [2,5]. We are interested in not only the properties of the Tsallis type relative entropy but also the properties before taking a trace, namely, Tsallis type relative operator entropy which is a parametric extension of the relative operator entropy. In this paper, we define the Tsallis relative operator entropy and then show some properties of Tsallis relative operator entropy. To this end, we slightly change the parameter q in Eq. (1) to λ in our definition which will be appeared in the following section. Moreover, in order to make our definition correspond to the definition of the relative operator entropy defined in [3], we change the sign of the original Tsallis relative entropy.

2. Tsallis relative entropy

As mentioned above, we adopt the slightly modified definition of the Tsallis relative entropy in the following.

Definition 1. Let $a = \{a_1, a_2, \dots, a_n\}$ and $b = \{b_1, b_2, \dots, b_n\}$ be two probability vectors satisfying $a_j, b_j > 0$. Then for $0 < \lambda \leqslant 1$

$$S_\lambda(a|b) = \frac{\sum_{j=1}^n a_j^{1-\lambda} b_j^\lambda - 1}{\lambda} \tag{2}$$

is called Tsallis relative entropy between a and b.

We should note that the Tsallis relative entropy is usually defined by

$$D_q(a|b) = \frac{1 - \sum_{j=1}^{n} a_j^q b_j^{1-q}}{1-q}, \tag{3}$$

with a parameter $q \geqslant 0$ in the field of statistical physics [7]. There is the relation between them such that $S_\lambda(a|b) = -D_{1-q}(a|b)$. However, in this paper, we adopt the definition of Eq. (2) in stead of Eq. (3), to study of the properties of the parametrically extended relative operator entropy as a series of the study of the relative operator entropy from the operator theoretical point of view. The opposite sign between the relative entropy defined by Umegaki [8] and the relative operator entropy led us to define the Tsallis relative operator entropy in the above.

Tsallis relative entropy defined in Eq. (2) has the following properties.

Proposition 1. *We have the following* (1) *and* (2).

(1) $S_\lambda(a|b) \geqslant \sum_{j=1}^{n} a_j \log \frac{b_j}{a_j}$ *for* $0 < \lambda \leqslant 1$.

(2) $\lim_{\lambda \to 0} S_\lambda(a|b) = \sum_{j=1}^{n} a_j \log \frac{b_j}{a_j}$.

Proof. (1) Since $t^\lambda - 1 \geqslant \log t^\lambda$, we have

$$\frac{\sum_{j=1}^{n} a_j^{1-\lambda} b_j^\lambda - 1}{\lambda} = \sum_{j=1}^{n} a_j \frac{\left(\frac{b_j}{a_j}\right)^\lambda - 1}{\lambda} \geqslant \sum_{j=1}^{n} a_j \log \frac{b_j}{a_j}.$$

(2)

$$\begin{aligned}\lim_{\lambda \to 0} \frac{\sum_{j=1}^{n} a_j^{1-\lambda} b_j^\lambda - 1}{\lambda} &= \lim_{\lambda \to 0} \frac{\sum_{j=1}^{n} a_j \left(\frac{b_j}{a_j}\right)^\lambda - 1}{\lambda} \\ &= \sum_{j=1}^{n} a_j \left(\frac{b_j}{a_j}\right)^\lambda \log \frac{b_j}{a_j}\Big|_{\lambda=0} \\ &= \sum_{j=1}^{n} a_j \log \frac{b_j}{a_j}. \qquad \square\end{aligned}$$

Proposition 2

$$0 \geqslant S_\lambda(a|b) \geqslant \frac{\left(\sum_{j=1}^{n} \frac{a_j^2}{b_j}\right)^{-\lambda} - 1}{\lambda}.$$

Proof. Since

$$\sum_{j=1}^{n} a_j^{1-\lambda} b_j^{\lambda} \leqslant \sum_{j=1}^{n} \{(1-\lambda)a_j + \lambda b_j\} = (1-\lambda)\sum_{j=1}^{n} a_j + \lambda \sum_{j=1}^{n} b_j = 1,$$

we have

$$S_\lambda(a|b) \leqslant 0.$$

We also give another inequality. Since

$$\begin{aligned}\sum_{j=1}^{n} a_j^{1-\lambda} b_j^{\lambda} &= \sum_{j=1}^{n} \left(\frac{a_j}{b_j}\right)^{-\lambda} a_j \geqslant \prod_{j=1}^{n} \left(\frac{a_j}{b_j}\right)^{-\lambda a_j} \\ &= \left(\prod_{j=1}^{n} \left(\frac{a_j}{b_j}\right)^{a_j}\right)^{-\lambda} \geqslant \left(\sum_{j=1}^{n} \frac{a_j^2}{b_j}\right)^{-\lambda},\end{aligned}$$

we have

$$S_\lambda(a|b) \geqslant \frac{\left(\sum_{j=1}^{n} \frac{a_j^2}{b_j}\right)^{-\lambda} - 1}{\lambda}. \qquad \square$$

3. Tsallis relative operator entropy

A bounded linear operator T on a Hilbert space H is said to be positive (denoted by $T \geqslant 0$) if $(Tx, x) \geqslant 0$ for all $x \in H$ and also an operator T is said to be strictly positive (denoted by $T > 0$) if T is invertible and positive. We define Tsallis relative operator entropy in the following.

Definition 2. For $A > 0$, $B > 0$ and $0 < \lambda \leqslant 1$,

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

is called Tsallis relative operator entropy between A and B.

In this section we give the Shannon type operator inequality and its reverse one satisfied by Tsallis relative operator entropy.

Theorem 1. *Let $\{A_1, A_2, \ldots, A_n\}$ and $\{B_1, B_2, \ldots, B_n\}$ be two sequences of strictly positive operators on a Hilbert space H. If $\sum_{j=1}^{n} A_j = \sum_{j=1}^{n} B_j = I$, then*

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

We need a lemma before we prove the main theorem.

Lemma 1. *For fixed* $t > 0$, *an inequality of* λ $(0 < \lambda \leqslant 1)$ *holds.*

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

Proof. If $t = 1$, then it is clear. If $t \neq 1$, then we put $F(\lambda) = \lambda(t-1) - t^\lambda + 1$. Then we have $F'(\lambda) = t - 1 - t^\lambda \log t$ and $F''(\lambda) = -t^\lambda (\log t)^2 < 0$. Hence $F(\lambda)$ is concave. Since $F(0) = F(1) = 0$, we have the result. □

Proof of Theorem 1. It follows from Lemma 1 that

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

where $A > 0$, $B > 0$ and $0 < \lambda \leqslant 1$. Then we have

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

We also prove another inequality. We apply Proposition 3.1 of Furuta [4] by putting $f(x) = -x^{-\lambda}$, $C_j = A_j^{1/2}$ and $X_j = A_j^{1/2}B_j^{-1}A_j^{1/2}$. Then

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

Hence

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

Then we complete the proof. □

We also obtain the operator version of the Shannon inequality and reverse one given by Furuta [4] as a corollary of Theorem 1 in the following.

Corollary 1 (Furuta [4]). *Let* $\{A_1, A_2, \ldots, A_n\}$ *and* $\{B_1, B_2, \ldots, B_n\}$ *be two sequences of strictly positive operators on a Hilbert space* H. *If* $\sum_{j=1}^n A_j = \sum_{j=1}^n B_j = I$, *then*

$$0 \geqslant \sum_{j=1}^{n} A_j^{1/2}\big(\log A_j^{-1/2} B_j A_j^{-1/2}\big) A_j^{1/2} \geqslant -\log\left[\sum_{j=1}^{n} A_j B_j^{-1} A_j\right].$$

We need the following lemma to prove it.

Lemma 2. *For* $0 < \lambda < 1$, $0 < \alpha < \beta$, *we have the following* (1) *and* (2).

(1) $\lim_{\lambda \to +0} \frac{t^{\lambda} - 1}{\lambda} = \log t$ *uniformly on* $[\alpha, \beta]$.

(2) $\lim_{\lambda \to +0} \frac{t^{-\lambda} - 1}{\lambda} = -\log t$ *uniformly on* $[\alpha, \beta]$.

Proof. We prove it by using Dini's theorem. □

Proof of Corollary 1. By (1) of Lemma 2,

$$\lim_{\lambda \to +0} \frac{(A^{-1/2} B A^{-1/2})^{\lambda} - I}{\lambda} = \log A^{-1/2} B A^{-1/2},$$

where the limit is taken in operator norm. Then

$$\begin{aligned}\lim_{\lambda \to +0} \sum_{j=1}^{n} T_{\lambda}(A_j | B_j) &= \lim_{\lambda \to +0} \sum_{j=1}^{n} \frac{A_j^{1/2}\big(A_j^{-1/2} B_j A_j^{-1/2}\big)^{\lambda} A_j^{1/2} - A_j}{\lambda} \\ &= \sum_{j=1}^{n} A_j^{1/2}\big(\log A_j^{-1/2} B_j A_j^{-1/2}\big) A_j^{1/2}.\end{aligned}$$

On the other hand, by (2) of Lemma 2, we have

$$\lim_{\lambda \to +0} \frac{\big(\sum_{j=1}^{n} A_j B_j^{-1} A_j\big)^{-\lambda} - I}{\lambda} = -\log\left[\sum_{j=1}^{n} A_j B_j^{-1} A_j\right]$$

Therefore Theorem 1 ensures

$$0 \geqslant \sum_{j=1}^{n} A_j^{1/2}\big(\log A_j^{-1/2} B_j A_j^{-1/2}\big) A_j^{1/2} \geqslant -\log\left[\sum_{j=1}^{n} A_j B_j^{-1} A_j\right]. \qquad \square$$

Actually the above Corollary 1 is a part of the Corollary 2.4 in [4]. We will generalize our Tsallis relative operator entropy and derive some generalized operator inequalities by the different way from [4] in the following section.

4. Generalized Tsallis relative operator entropy

We remind of the relative operator entropy and its related operator entropy.

Definition 3. For $A > 0$, $B > 0$

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

is called relative operator entropy between A and B. It was defined by Fujii and Kamei [3] originally. For $A > 0$, $B > 0$ and $\lambda \in \mathbb{R}$, the generalized relative operator entropy was defined by Furuta [4]

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

and

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

In particular we remark that $S_0(A|B) = S(A|B)$, $A\natural_0 B = A$ and $A\natural_1 B = B$.

We generalize the definition of the Tsallis relative operator entropy.

Definition 4. For $A > 0$, $B > 0$, $\lambda, \mu \in \mathbb{R}$, $\lambda \neq 0$ and $k \in \mathbb{Z}$,

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

is called generalized Tsallis relative operator entropy. In particular we remark that for $\lambda \neq 0$

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

We state the relationship among $S_{\mu\pm k\lambda}(A|B)$, $S_{\mu\pm(k+1)\lambda}(A|B)$ and $\tilde{T}_{\mu,k+1,\pm\lambda}(A|B)$.

Proposition 3. *If* $\lambda > 0$, $\mu \in \mathbb{R}$ *and* $k = 0, 1, 2, \ldots,$ *then*

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

Proof. If $\lambda > 0$, $\mu \in \mathbb{R}$ and $k = 0, 1, 2, \ldots$, then it is easy to give the following inequalities for any $t > 0$:

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

Then replace t by $A^{-1/2}BA^{-1/2}$ and multiply $A^{1/2}$ on both sides so we get the desired results. □

By putting $k = 0$ or 1, we get the following.

Corollary 2. *For* $A > 0,\ B > 0,\ \mu \in \mathbb{R}$ *and* $\lambda > 0$,

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

In particular by putting $\mu = 0$, $\lambda = 1$, we get the following.

Corollary 3. *For* $A > 0,\ B > 0$,

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

We rewrite the following:

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

Similarly we state the relationship among $\sum_{j=1}^{n} S_{\mu\pm k\lambda}(A_j|B_j)$, $\sum_{j=1}^{n} S_{\mu\pm(k+1)\lambda}(A_j|B_j)$ and $\sum_{j=1}^{n} \tilde{T}_{\mu,k+1,\pm\lambda}(A_j|B_j)$, where $A_j > 0$, $B_j > 0$ satisfying $\sum_{j=1}^{n} A_j = \sum_{j=1}^{n} B_j = I$. If $\lambda > 0$, $\mu \in \mathbb{R}$ and $k = 0, 1, 2, \ldots$, then we have the following:

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

By putting $k = 0$ or 1, we get the following.

Corollary 4. *For* $A > 0,\ B > 0,\ \mu \in \mathbb{R}$ *and* $\lambda > 0$,

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

In particular by putting $\mu = 0, \lambda = 1$, we get the following result which is somewhat different type from Corollary in [4].

Corollary 5. *For* $A_j > 0,\ B_j > 0$ *satisfying* $\sum_{j=1}^{n} A_j = \sum_{j=1}^{n} B_j = I$,

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

Acknowledgments

The authors thank Professor T. Furuta for reading our manuscript and for valuable comments.

References

[1] S. Abe, Monotonic decrease of the quantum nonadditive divergence by projective measurements, Phys. Lett. A 312 (2003) 336–338.

[2] N. Bebiano, J. da Providencia Jr., R. Lemos, Matrix inequalities in statistical mechanics, Linear Algebra Appl. 376 (2004) 265–273.

[3] J.I. Fujii, E. Kamei, Relative operator entropy in noncommutative information theory, Math. Japon. 34 (1989) 341–348.

[4] T. Furuta, Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators, Linear Algebra Appl. 381 (2004) 219–235.

[5] F. Hiai, D. Petz, The Golden–Thompson trace inequality is complemented, Linear Algebra Appl. 181 (1993) 153–185.

[6] C. Tsallis, Possible generalization of Boltzman–Gibbs statistics, J. Statist. Phys. 52 (1988) 479–487.

[7] C. Tsallis et al., in: S. Abe, Y. Okamoto (Eds.), Nonextensive Statistical Mechanics and Its Applications, Springer-Verlag, Heidelberg, 2001, Comprehensive list of references available from: <http://tsallis.cat.cbpf.br/biblio.htm>.

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