Applied Physics 应用物理, 2019, 9(9), 390-396
Published Online September 2019 in Hans. http://www.hanspub.org/journal/app
https://doi.org/10.12677/app.2019.99047

Hans汉斯

Study on the Stability of (*r*, *s*) Entropy

Tiantian Wang, Jiamei Wang, Wanqing Li

School of Mathematics & Physics, Anhui University of Technology, Ma'anshan Anhui
Email: 1633512926@qq.com

Received: Sep. 6th, 2019; accepted: Sep. 20th, 2019; published: Sep. 27th, 2019

Abstract

In this paper, the stability of Tsallis entropy in reference [1] is generalized, and the compact upper bound of (*r*, *s*) entropy in classical cases is established by using the probabilistic coupling method, and the upper bound of the quantum condition is given in the same way. These results give some quantitative characterizations of uniform continuity and stability of entropy.

Keywords

(*r*, *s*) Entropy, Trace Distance, Probability Distribution, Density Operate

(*r*, *s*)熵的稳定性的研究

王甜甜，汪加梅，李婉晴

安徽工业大学，数理科学与工程学院，安徽 马鞍山
Email: 1633512926@qq.com

收稿日期：2019年9月6日；录用日期：2019年9月20日；发布日期：2019年9月27日

摘 要

本文推广了文献[1]中关于Tsallis熵的稳定性的研究，利用概率耦合方法建立了经典情形下(*r*, *s*)熵的紧上界，并以同样的方法给出了量子情形的紧上界。这些结果给出了(*r*, *s*)熵的一致连续性和稳定性的一些定量刻画。

关键词

(*r*, *s*)熵，迹距离，概率分布，密度算子

文章引用: 王甜甜, 汪加梅, 李婉晴. (*r*, *s*)熵的稳定性的研究[J]. 应用物理, 2019, 9(9): 390-396.
DOI: 10.12677/app.2019.99047

1. 引言

在经典力学和量子论中，用状态泛函来量化状态的信息内容是非常重要的。普遍的经典概率分布的 Shannon 熵和密度算子的 Von Neumann 熵就是两个范例。此外，还从更一般的物理和数学考虑出发，引入了在描述反常现象中起重要作用的各种广义熵量，如 Renyi 熵和 Tsallis 熵等，而这些熵都可由(r,s)熵根据参数的不同取值得到。关于所以这些状态泛函的一个基本问题是它们的连续性或不连续性质：当基本状态发生变化时，它们是如何变化的？

1982 年，Lesche 对 Renyi 熵研究了这个问题，他认为 Renyi *r*-熵对于$r \neq 1$是不稳定的(或者说性质是不强的)，用数学语言来说，它们对于迹距离不是一致连续的。2002 年，安倍再次讨论了 Lesche 型的稳定性问题[2]，他证明了 Tsallis 熵虽然从数学角度来说与 Renyi 熵无关，但对于有限系统来说是稳定的。此外，Curado and Nobre [3]，and Abe *et al.* [4]证明了某些广义熵的稳定性。

具体来说，让我们考虑由所有 m 维向量 $p=\{p_1,p_2,\cdots,p_m\}$ 组成的纯态下 Δ_m 的状态泛函 Φ。假设 p 在 Δ_m 中变化时，$\Phi(p)$ 的最大值为 Φ_m，则稳定性性质数学描述为对任何 $\varepsilon>0$，存在 $\delta>0$，使得对任何 $\sum_{i=1}^{m}|p_i-q_i|<\delta$，都有[2] [3]：

$$\frac{\Phi(p)-\Phi(q)}{\Phi_m}\le\varepsilon$$

在上述稳定性考虑的启发下，本文将通过一些尖锐的不等式，利用迹距离来控制(r,s)熵的变化，从而更精确地量化(r,s)熵的稳定性性质。

2. 经典情形

首先回顾一些概念。设 p_X 为有限维随机变量 X 的概率分布

$$p_X=\{p_X(i):i=1,2,\cdots,m\}$$

则经典(r,s)熵[5]为

$$H_r^s(p_X)=\frac{1}{(1-r)s}\left(\left(\sum_{i=1}^{m}p_X^r(i)\right)^s-1\right) \tag{1}$$

其中 r 是可以取任意值的，取$r=1$时可取$r\to1$，但是为了避免其他难题，我们只考虑具有特殊物理意义的$r>1,s\ge0$的情况。当$s\to0$，$r\to1$时$H_r^s(p_X)$是由常数因子$\log_e 2$决定的 Shannon 熵；

$$H(p_X)=-\sum_i p_X(i)\log p_X(i)$$

对数为以 2 为底。

很容易看出，m 固定时，当且仅当对任意的 i，$p_X(i)=\frac{1}{m}$ 时，$H_r^s(p_X)$取最大值，最大值为

$$H_{r,m}^s=\frac{m^{(1-r)s}-1}{(1-r)s} \tag{2}$$

对于一对可能相关的随机变量(X,Y)，联合概率分布为p_{XY}，联合(r,s)熵为：

$$H_r^s(p_{XY})=\frac{1}{(1-r)s}\left(\left(\sum_{i,j=1}^{m}p_{XY}{}^r(i,j)\right)^s-1\right)$$

因为$r>1$，则有：

$$p_X^r(i)=\left(\sum_j p_{XY}(i,j)\right)^r\geq\sum_j p_{XY}^r(i,j)$$

则

$$\left(\sum_{i=1}^{m}p_X^r(i)\right)^s\geq\left(\sum_{i,j}^{m}p_{XY}^r(i,j)\right)^s \tag{3}$$

且$1-r<0$，则$H_r^s(p_X)\leq H_r^s(p_{XY})$。

对于两个随机变量X和Y在相同范围内概率分布分别为p_X和p_Y，它们的迹距离定义为：

$$\|p_X-p_Y\|_K=\frac{1}{2}\sum_{i=1}^{m}|p_X(i)-p_Y(i)| \tag{4}$$

因子$\frac{1}{2}$可保证迹距离的范围在 0 到 1 之间且将简化以后的公式。这个距离在描述概率分布的可分辨性方面起着重要的作用[6]，也叫作 Kolmogorov 距离，因为

$$\|p_X-p_Y\|_K=\sup_A\left|\sum_{i\in A}p_X(i)-\sum_{i\in A}p_Y(i)\right|$$

是被 Kolmogorov 提出的[7]。

有了以上的准备，我们可以给出以下第一个主要的结论：

定理 1. 设$\varepsilon=\|p_X-p_Y\|_K$是由等式(4)定义的迹距离，且$H_{r,m-1}^s=\frac{(m-1)^{(1-r)s}-1}{(1-r)s}$，则对于$r>1$，有

$$|H_r^s(p_X)-H_r^s(p_Y)|\leq H_r^s(\varepsilon)+\varepsilon^{rs}H_{r,m-1}^s \tag{5}$$

成立。其中$H_r^s(\varepsilon)=\frac{1}{(1-r)s}\left(\varepsilon^r+(1-\varepsilon)^r-1\right)$为二元$(r,s)$熵。

注 1　不等式(5)可以取等号，例如当

$$p_X(1)=1-\varepsilon, p_X(2)=p_X(3)=\cdots=p_X(m)=\frac{\varepsilon}{m-1}$$

和

$$p_Y(1)=1, p_Y(2)=p_Y(3)=\cdots=p_Y(m)=0$$

时取等号。

注 2　注意不等式(5)的左边与p_X或p_Y中的元素的顺序无关，但是迹距离$\varepsilon=\|p_X-p_Y\|_k$与元素的顺序有关。因此不等式(5)可以进一步加强：

$$|H_r^s(p_X)-H_r^s(p_Y)|\leq H_r^s(\varepsilon_\tau)+\varepsilon_\tau^{rs}H_{r,m-1}^s \tag{6}$$

其中$\varepsilon_\tau=\frac{1}{2}\sum_{i=1}^{m}|p_X(\tau(i))-p_Y(i)|$，$\tau$是集合$\{1,2,\cdots,m\}$的任意排列。

定理 1 的证明：我们分两步进行。若 $p_X = p_Y$，结果显然成立，因此证明中我们都假设 $p_X \neq p_Y$。$I=\{i: p_X(i)\le p_Y(i)\}$ 和 $I^c=\{i: p_X(i)> p_Y(i)\}$ 是非空的。

1) 首先，对于任意两个概率分布 p_X 和 p_Y，存在一个二元概率分布 p_{XY} 使得

$$p_{XY}(i,i)=\min\{p_x(i),p_Y(i)\}, i=1,2,\cdots,m$$

令 $\varepsilon = p(X\neq Y)$，则

$$\begin{aligned}P\{X\neq Y\} &= 1-P(X=Y)\\ &=1-\sum_{i=1}^{m}\min\{p_x(i),p_Y(i)\}\\ &=\frac{1}{2}\sum_{i=1}^{m}\left(p_x(i)+p_Y(i)-2\min\{p_x(i),p_Y(i)\}\right)\\ &=\frac{1}{2}\sum_{i=1}^{m}\left|p_x(i)-p_Y(i)\right|\\ &=\varepsilon\end{aligned}$$

若令 $q_j=\sum_{i:i\neq j} p_{XY(i,j)}$，则

$$\sum_j q_j=\sum_{i\neq j} p_{XY}(i,j)=p(X\neq Y)=\varepsilon \tag{7}$$

2) 不失一般性，我们假设 $H_r^s(p_X)\ge H_r^s(p_Y)$，则根据不等式(3)有

$$\left|H_r^s(p_X)-H_r^s(p_Y)\right|=H_r^s(p_X)-H_r^s(p_Y)\le H_r^s(p_{XY})-H_r^s(p_Y) \tag{8}$$

令 $\Theta=\begin{cases}0, X=Y\\ 1, X\neq Y\end{cases}$，则明显有 $P(\Theta)=P(X\neq Y)=\varepsilon$，而

$$H_r^s(p_{XY})-H_r^s(p_Y)=\left(H_r^s(p_{XY})-H_r^s(p_{\Theta Y})\right)+\left(H_r^s(p_{\Theta Y})-H_r^s(p_Y)\right)$$

因为

$$H_r^s(p_{\Theta Y})=\frac{1}{(1-r)s}\left[\left(\sum_{\theta,j}p_{\Theta Y}^r(\theta,j)\right)^s-1\right]$$

$$p_{\Theta Y}^r(0,j)=P(\Theta=0,Y=j)=P(X=Y,Y=j)=P_{XY}(j,j)$$

$$p_{\Theta Y}^r(1,j)=P(\Theta=1,Y=j)=P(X\neq Y,Y=j)=\sum_{i:i\neq j}p_{XY}(i,j)$$

则

$$\begin{aligned}&H_r^s(p_{XY})-H_r^s(p_{\Theta Y})\\ &=\frac{1}{(1-r)s}\left[\left(\sum_{i,j}p_{XY}^r(i,j)\right)^s-1\right]-\frac{1}{(1-r)s}\left[\left(\sum_j p_{XY}^r(j,j)\right)^s+\sum_j\left(\sum_{i:i\neq j}p_{XY}^r(i,j)\right)^s-1\right]\\ &=\frac{1}{(1-r)s}\left[\left(\sum_{i,j}p_{XY}^r(i,j)\right)^s-\sum_j\left(\sum_{i:i\neq j}p_{XY}^r(i,j)\right)^s\right]\\ &=\frac{1}{(1-r)s}\left(\sum_j q_j^r\right)^s\left[\left(\sum_{i:i\neq j}\frac{p_{XY}^r(i,j)}{q_j^r}\right)^s-1\right]\end{aligned}$$

这里 $q_j=\sum_{i:i\neq j}p_{XY}(i,j)$，因为 $\frac{1}{(1-r)s}\left[\left(\sum_{i:i\neq j}\frac{p_{XY}^r(i,j)}{q_j^r}\right)^s-1\right]$ 是概率分布

$$\left\{\frac{p_{XY}(i,j)}{q_j},i-1,2,\cdots,m,i\neq j\right\}$$

的 (r,s) 熵，对于 $m-1$ 维成立，以 $H_{r,m-1}^s$ 为上界。因此，结合等式(9)，对于 $r>1,s\geq 0$ 有

$$H_r^s(p_{XY})-H_r^s(p_{\Theta Y})\leq\left(\sum_j q_j^r\right)^s H_{r,m-1}^s\leq\left(\sum_j q_j\right)^{rs} H_{r,m-1}^s=\varepsilon^r H_{r,m-1}^s \tag{9}$$

根据 $r>1$，有：

$$\sum_j\left(\frac{q_j}{\sum_k q_k}\right)^r\leq\sum_j\frac{q_j}{\sum_k q_k}=1$$

根据指数函数 $y=a^x$ 的单调性可知，当 $0<a<1,x\geq 0$ 时，$y\leq 1$，则

$$\frac{\left(\sum_j q_j^r\right)^s}{\left(\sum_j q_j\right)^{rs}}\leq 1$$

则第二个不等式成立。

注意到 $H_r^s(p_\Theta)=H_r^s(\varepsilon,1-\varepsilon)$，则我们只需证

$$H_r^s(p_{\Theta Y})-H_r^s(p_Y)\leq H_r^s(p_\Theta) \tag{10}$$

而 $p_{\Theta Y}(\theta,j)=P(\Theta=\theta,Y=j)$，$p_{\Theta|Y}(\theta|j)=P(\Theta=\theta|Y=j)=\frac{p_{\Theta Y}(\theta,j)}{p_Y(j)}$，则

$$\begin{aligned}
H_r^s(p_{\Theta Y})-H_r^s(p_Y)&\leq\frac{1}{(1-r)s}\left[\left(\sum_{\theta,j}p_{\Theta Y}^r(\theta,j)\right)^s-\left(\sum_j p_Y^r\right)^s\right]\\
&=\frac{1}{(1-r)s}\left(\sum_j p_Y^r\right)^s\left[\frac{\left(\sum_{\theta,j}p_{\Theta Y}^r(\theta,j)\right)^s}{\left(\sum_j p_Y^r\right)^s}-1\right]\\
&=\frac{1}{(1-r)s}\left(\sum_j p_Y^r\right)^s\left[\left(\sum_\theta\frac{p_{\Theta Y}^r(\theta,j)}{p_Y^r}\right)^s-1\right]\\
&=\frac{1}{(1-r)s}\left(\sum_j p_Y^r\right)^s\left[\left(\sum_\theta p_{\Theta|Y}^r(\theta|j)\right)^s-1\right]
\end{aligned}$$

根据函数 $f(t)=t^r$　$(r>1)$的凸性和单调性知

$$\left(\sum_j p_Y(j)p_{\Theta|Y}(\theta|j)\right)^r\leq\sum_j p_Y(j)p_{\Theta|Y}^r(\theta|j)$$

$$\left(\sum_j p_Y(j)\,p_{\Theta|Y}(\theta|j)\right)^{rs} \le \left(\sum_j p_Y(j)\,p_{\Theta|Y}^r(\theta|j)\right)^s$$

因此，根据$1-r<0$，有

$$\begin{aligned}
H_r^s(p_\Theta) &= \frac{1}{(1-r)s}\left[\left(\sum_\theta p_\Theta^r(\theta)\right)^s - 1\right] \\
&= \frac{1}{(1-r)s}\left\{\left[\sum_\theta\left(\sum_j p_Y(j)\,p_{\Theta|Y}(\theta|j)\right)^r\right]^s - 1\right\} \\
&\ge \frac{1}{(1-r)s}\left[\left(\sum_\theta\sum_j p_Y(j)\,p_{\Theta|Y}^r(\theta|j)\right)^s - 1\right] \\
&= \frac{1}{(1-r)s}\left(\sum_j p_Y(j)\right)^s\left[\left(\sum_\theta p_{\Theta|Y}^r(\theta|j)\right)^s - 1\right] \\
&\ge \frac{1}{(1-r)s}\left(\sum_j p_Y^r(j)\right)^s\left[\left(\sum_\theta p_{\Theta|Y}^r(\theta|j)\right)^s - 1\right] \\
&= H_r^s(p_{\Theta Y}) - H_r^s(p_Y)
\end{aligned}$$

由式可得到定理 1 成立。

3. 量子延伸

文献[1]中作者根据经典情形给出了量子情形 Tsallis 熵在量子情形下的稳定性;对于两个密度算子ρ和σ，定义它们的迹距离为

$$\|\rho-\sigma\|_K = \frac{1}{2}tr|\rho-\sigma| \tag{11}$$

令$\gamma = \|\rho-\sigma\|_K$，且$S_{r,m-1}^s = \dfrac{(m-1)^{(1-r)s}-1}{(1-r)s}$，则对于$\alpha>1$，有

$$|S_\alpha(\rho)-S_\alpha(\sigma)| \le H_\alpha(\gamma)+\gamma^\alpha S_{\alpha,m-1}$$

在这一部分中，我们会也将不等式(5)和(6)延伸到量子情形中，用密度算子代替概率密度，求和改为取迹。更具体地说，令ρ为密度算子，类比于等式(1)，它的(r,s)熵为

$$S_r^s(\rho) = \frac{1}{(1-r)s}\left[\left(tr\rho^r\right)^s - 1\right]$$

当$s=1, r\to 1$时，$S(\rho) = -tr\rho\log\rho$为 Von Neumann 熵。

当密度算子被限制在m维复希尔伯特空间上，当$p=\dfrac{1}{m}$，即最大混合态时$S_r^s(\rho)$有最大值，且最大值与等式(2)所定义的一样

$$S_{r,m}^s = H_{r,m}^s = \frac{m^{(1-r)s}-1}{(1-r)s}$$

定理 2. 令$\gamma = \|\rho-\sigma\|_K$为(11)式所定义的迹距离，$S_{r,m-1}^s = \dfrac{(m-1)^{(1-r)s}-1}{(1-r)s}$，则对于$r>1, s\ge 0$，有：

$$\left|S_r^s(\rho)-S_r^s(\sigma)\right|\le H_r^s(\gamma)+\gamma^{rs}S_{r,m-1}^s \tag{12}$$

这里的 $H_r^s(\bullet)$ 是定理 1 中定义的二元 (r,s) 熵。

定理 2 的证明 令 ρ 和 σ 有谱分解：

$$\rho=\sum_{i=1}^m p_i\left|\phi_i\right\rangle\left\langle\phi_i\right|,\quad \sigma=\sum_{i=1}^m q_i\left|\psi_i\right\rangle\left\langle\psi_i\right|$$

谱集 $p=\{p_i\}$ 和 $q=\{q_i\}$ 分别构成两个经典概率分布，并且有

$$S_r^s(\rho)=H_r^s(p),\quad S_r^s(\sigma)=H_r^s(q)$$

所以根据定理 1 或者更确切地说根据不等式(7)，我们有对于任何排列的 τ

$$\left|S_r^s(\rho)-S_r^s(\sigma)\right|=\left|H_r^s(\rho)-H_r^s(\sigma)\right|\le\left|S(\rho)-S(\sigma)\right|\le H_r^s(\gamma_\tau)+\gamma_\tau^r H_{r,m-1}^s \tag{13}$$

这里 $\gamma_\tau=\frac{1}{2}\sum_{i=1}^m\left(p_\tau(i)-q(i)\right)$。

由[8]的定理III.4.4 知，存在排列 u 和 v，使得 $\gamma_u\le\gamma\le\gamma_v$，这里 γ_u 和 γ_v 与 γ_z 的定义一样。函数 $f(\gamma)=S_r^s(\gamma)+\gamma^\alpha S_{r,m-1}^s$ 是单峰函数(先单调增加，后单调递减)，我们得到 $f(\gamma_v)\le f(\gamma)$ 或 $f(\gamma_u)\le f(\gamma)$，将该不等式代入到不等式(13)中得到不等式(12)。定理 2 得证。

对比定理 1 和定理 2 可以清晰地看出量子情形是在经典情形的基础上将概率密度改为密度算子，将求和改为取迹，证明过程也基本类似，无论是经典情形还是量子情形，(r,s) 熵都有紧上界，也即满足稳定性的结论。

4. 总结

本文利用概率耦合的方法讨论了经典 (r,s) 熵的 Lesche 稳定性，通过一些尖锐的不等式，利用迹距离来控制 (r,s) 熵的变化，从而更精确地量化经典 (r,s) 熵的稳定性性质，同时也推出了相应的量子 (r,s) 熵的稳定性结论，当 r，s 取不同值时可以直接得到 Tsallis 熵等其他的熵的稳定性性质。

参考文献

[1] Zhang, Z.M. (2007) Uniform Estimates on the Tsallis Entropies. *Letters in Mathematical Physics*, **80**, 171-181. https://doi.org/10.1007/s11005-007-0155-1

[2] Abe, S. (2005) Stability of Tsallis Entropy and Instabilities of Rényi and Normalized Tsallis Entropies: A Basis for q-Exponential Distribution. *Physical Review E Statistical Nonlinear & Soft Matter Physics*, **72**, Article ID: 046134.

[3] Abe, S., Kaniadakis, G. and Scarfone, A.M. (2004) Stabilities of Generalized Entropies. *Journal of Physics A*: *Mathematical and General*, **37**, 10513-10519. https://doi.org/10.1088/0305-4470/37/44/004

[4] Curado, E.M.F. and Nobre, F.D. (2004) On the Stability of Analytic Entropic Forms. *Physica A*: *Statistical Mechanics and its Applications*, **335**, 94-106. https://doi.org/10.1016/j.physa.2003.12.026

[5] Hu, X.H. and Ye, Z.X. (2006) Generalized Quantum Entropy. *Journal of Mathematical Physics*, **47**, Article ID: 023502. https://doi.org/10.1063/1.2165794

[6] Fuchs, C.A. and van de Graaf, J . (1999) Cryptographic Distinguishability Measures for Quantum-Mechanical States. *IEEE Transactions on Information Theory*, **45**, 1216-1227. https://doi.org/10.1109/18.761271

[7] Kolmogorov, A.N. (1963) On the Approximation of Distributions of Sums of Independent Summands by Infinitely Divisible Distributions. *Sankhyā*: *The Indian Journal of Statistics*, *Series A* (1961-2002), **25**, 159-174.

[8] Bhatia, R. (2011) Matrix Analysis. World Book Publishing Company.