Information geometry induced from sandwiched Rényi α-divergence

Akio Fujiwara & Kaito Takahashi
Department of Mathematics, Osaka University, JAPAN
fujiwara@math.sci.osaka-u.ac.jp

Abstract

Information geometrical dualistic structure $(g^{(D_\alpha)}, \nabla^{(D_\alpha)}, \nabla^{(D_\alpha)*})$ induced from the sandwiched Rényi α-divergence $D_\alpha(\rho\|\sigma) := \frac{1}{\alpha(\alpha-1)} \log \mathrm{Tr} \left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho\, \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha$ on the quantum state space $\mathcal{S}(\mathcal{H})$ is studied. It is shown that the Riemannian metric $g^{(D_\alpha)}$ is monotone if and only if $\alpha \in (-\infty - 1] \cup [\frac{1}{2}, \infty)$, and that the quantum statistical manifold $(\mathcal{S}(\mathcal{H}), g^{(D_\alpha)}, \nabla^{(D_\alpha)}, \nabla^{(D_\alpha)*})$ is dually flat if and only if $\alpha = 1$.

Introduction

Let us consider a finite quantum state space

$$\mathcal{S}(\mathcal{H}) := \{\rho \in \mathcal{L}_{++}(\mathcal{H}) \mid \mathrm{Tr}\,\rho = 1\},$$

where $\mathcal{H}$ is a finite dimensional Hilbert space. For $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and $\alpha \in (0,1) \cup (1, \infty)$, let

$$\tilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \log \mathrm{Tr} \left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho\, \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha \tag{1}$$

with the convention that $\tilde{D}_\alpha(\rho\|\sigma) = \infty$ if $\alpha > 1$ and $\ker\sigma \not\subset \ker\rho$. The quantity (1) is called the *quantum Rényi divergence* [8] or the *sandwiched Rényi relative entropy* [11], and is extended to $\alpha = 1$ by continuity, to obtain the *Umegaki-von Neumann relative entropy*:

$$\tilde{D}_1(\rho\|\sigma) = \lim_{\alpha \to 1} \tilde{D}_\alpha(\rho\|\sigma) = \mathrm{Tr}\{\rho(\log\rho - \log\sigma)\}.$$

The quantity (1) has several desirable properties: amongst others, if $\alpha \geq \frac{1}{2}$, it is monotone under completely positive trace preserving maps [8, 11, 2, 4]. This property was successfully used in studying the strong converse properties of the channel capacity [11, 7] and the quantum hypothesis testing problem [6].

For $\alpha < 0$, however, the quantity (1) does not seem to be a reasonable measure of information [10], because it takes negative values. We therefore introduce the following "rescaled" version

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha(\alpha - 1)} \log \mathrm{Tr} \left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho\, \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha, \tag{2}$$

for $\alpha \in \mathbb{R}\backslash\{0,1\}$, which shall be referred to as the *sandwiched Rényi α-divergence*. As a matter of fact, the factor $\frac{1}{\alpha}$ is introduced not only to make $D_\alpha(\rho\|\sigma)$ positive for all α, but also to establish a correspondence to the classical information geometry [1]. Note that (2) is continuously extended to $\alpha = 1$, but cannot be extended to $\alpha = 0$ because $\lim_{\alpha \to 0} D_\alpha(\rho\|\sigma)$ does not always exist.

The objective of the present study is to investigate the information geometrical structure induced from (2) on the quantum state space $\mathcal{S}(\mathcal{H})$.

Main Results

For each $\alpha \in \mathbb{R}\backslash\{0\}$, the quantity (2) enjoys the property:

$$D_\alpha(\rho\|\sigma) \geq 0 \quad (\forall \rho, \sigma \in \mathcal{S}), \quad \text{and} \quad D_\alpha(\rho\|\sigma) = 0 \quad \text{if} \quad \rho = \sigma.$$

This fact allows us to introduce, using Eguchi's method [3], a Riemannian metric:

$$g_\rho^{(D_\alpha)}(X, Y) := D_\alpha((XY)_\rho\|\sigma)\Big|_{\sigma=\rho},$$

and a pair of affine connections:

$$g_\rho^{(D_\alpha)}(\nabla_X^{(D_\alpha)} Y, Z) := -D_\alpha((XY)_\rho\|(Z)_\sigma)\Big|_{\sigma=\rho}, \quad g_\rho^{(D_\alpha)}(\nabla_X^{(D_\alpha)*} Y, Z) := -D_\alpha((Z)_\rho\|(XY)_\sigma)\Big|_{\sigma=\rho}.$$

A Riemannian metric g on a quantum state space $\mathcal{S}(\mathcal{H})$ is called a *monotone metric* [9] if

$$g_\rho(X, Y) \geq g_{\gamma(\rho)}(\gamma_* X, \gamma_* Y) \tag{3}$$

holds for all completely positive trace preserving maps $\gamma : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ and all vector fields $X, Y \in T\mathcal{S}(\mathcal{H})$. It is commonly believed that any physical process is represented by a trace preserving completely positive map. Therefore the monotonicity (3), which implies that the infinitesimal distance between two nearby states always shrinks by a physical process γ, is a natural requirement for a physical information processing. In this sense, characterizing the monotone metric is of fundamental importance in quantum information theory.

The main result of the present study is the following.

Theorem 1. *The induced Riemannian metric $g^{(D_\alpha)}$ is monotone under completely positive trace preserving maps if and only if $\alpha \in (-\infty, -1] \cup [\frac{1}{2}, \infty)$.*

As a by-product, we arrive at the following corollary, the latter part of which was first observed by numerical evaluation [8].

Corollary 2. *The sandwiched Rényi α-divergence $D_\alpha(\rho\|\sigma)$ is not monotone under completely positive trace preserving maps if $\alpha \in (-1, 0) \cup (0, \frac{1}{2})$. Consequently, the original sandwiched Rényi relative entropy $\tilde{D}_\alpha(\rho\|\sigma)$ is not monotone if $\alpha \in (0, \frac{1}{2})$.*

We also studied the dualistic structure $(g^{(D_\alpha)}, \nabla^{(D_\alpha)}, \nabla^{(D_\alpha)*})$ on the quantum state space $\mathcal{S}(\mathcal{H})$, and obtained the following.

Theorem 3. *The quantum statistical manifold $(\mathcal{S}(\mathcal{H}), g^{(D_\alpha)}, \nabla^{(D_\alpha)}, \nabla^{(D_\alpha)*})$ is dually flat if and only if $\alpha = 1$.*

Sketch of Proof of Theorem 1

Let $X^{(m)}$ be the m-representation of a tangent vector $X \in T_\rho\mathcal{S}$ defined by $X^{(m)} := X\rho$, and let $X_f^{(e)}$ be the e-representation of X defined by

$$X_f^{(e)} := f(\Delta_\rho)^{-1}\left\{(X\rho)\rho^{-1}\right\},$$

where $f : \mathbb{R}_{++} \to \mathbb{R}_{++}$ is a symmetric monotone function satisfying $f(1) = 1$, and Δ_ρ is the *modular operator* associated with $\rho \in \mathcal{S}(\mathcal{H})$ defined by

$$\Delta_\rho : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}) : A \mapsto \rho A \rho^{-1}.$$

Lemma 4. *For each $\alpha \in \mathbb{R}\backslash\{0,1\}$, the metric $g^{(D_\alpha)}$ is represented in the form*

$$g_\rho^{(D_\alpha)}(X, Y) = \mathrm{Tr}\left\{X^{(m)} Y_{f^{(D_\alpha)}}^{(e)}\right\}$$

where

$$f^{(D_\alpha)}(t) := (\alpha - 1)\frac{t^{\frac{1}{\alpha}} - 1}{1 - t^{\frac{1-\alpha}{\alpha}}} \tag{4}$$

with the convention that $f^{(D_\alpha)}(1) := \lim_{t \to 1} f^{(D_\alpha)}(t) = 1$.

Proof. Direct computation using methods of the Gâteaux differentiation. □

Example 5. *When $\alpha = \frac{1}{2}$, the function $f^{(D_{1/2})}(t) = \frac{1+t}{2}$ corresponds to the SLD metric, and when $\alpha = -1$, the function $f^{(D_{-1})}(t) = \frac{2t}{1+t}$ corresponds to the real RLD metric. Further, the limiting function $f^{(D_1)}(t) := \lim_{\alpha \to 1} f^{(D_\alpha)}(t) = \frac{t-1}{\log t}$ gives the Bogoliubov metric: this is consistent to the fact that $D_1(t) := \lim_{\alpha \to 1} D_\alpha(t)$ is the von Neumann relative entropy. It is well known that these three functions are operator monotone. Note that the limiting function $f^{(D_{\pm\infty})}(t) := \lim_{\alpha \to \pm\infty} f^{(D_\alpha)}(t) = \frac{t \log t}{t-1} = t/f^{(D_1)}(t)$ is also operator monotone.*

To prove Theorem 1, we must specify all the values of α that make the function (4) operator monotone [9]. In what follows, we change the parameter α into $\beta := \frac{1}{\alpha}$, and denote the corresponding function $f^{(D_\alpha)}(t)$ by $f_\beta(t)$, i.e.,

$$f_\beta(t) := \frac{\beta - 1}{\beta}\,\frac{t^\beta - 1}{t^{\beta-1} - 1}$$

where $\beta \notin \{0, 1\}$. We extend this function to $\beta = 0$ and 1 by continuity, to obtain

$$f_0(t) := \lim_{\beta \to 0} f_\beta(t) = \frac{t \log t}{t - 1}, \qquad f_1(t) := \lim_{\beta \to 1} f_\beta(t) = \frac{t - 1}{\log t}.$$

Lemma 6. *The function $f_\beta(t)$ is operator monotone if and only if $-1 \leq \beta \leq 2$.*

Proof. The proof of 'if' part is divided into two steps: we first observe the identity $f_{\frac{1}{2}-\delta}(t) = \frac{t}{f_{\frac{1}{2}+\delta}(t)}$, and then represent $f_\beta(t)$ for $\frac{1}{2} \leq \beta \leq 2$ as a composition of some known operator monotone functions (See also [5]). The 'only if' part is proved by showing that $\frac{2t}{1+t} \leq f_\beta(t) \leq \frac{1+t}{2}$ holds for $t > 0$ only when $-1 \leq \beta \leq 2$. □

References

[1] S.-I. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of Mathematical Monographs 191 (AMS and Oxford, RI, 2000).

[2] S. Beigi, J. Math. Phys., **54**, 122202 (2013); arXiv:1306.5920.

[3] S. Eguchi, Hiroshima Math. J., **22**, 631-647 (1992).

[4] R. L. Frank and E. H. Lieb, "Monotonicity of a relative Rényi entropy," arXiv:1306.5358.

[5] T. Furuta, Lin. Alg. Appl., **429**, 972-980 (2008).

[6] M. Mosonyi and T. Ogawa, Commun. Math. Phys., **334**, 1617-1648 (2015); arXiv:1309.3228.

[7] M. Mosonyi and T. Ogawa, arXiv:1409.3562.

[8] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr and M. Tomamichel, J. Math. Phys., **54**, 122203 (2013): arXiv:1306.3142.

[9] D. Petz, Lin. Alg. Appl., **224**, 81-96 (1996).

[10] A. Rényi, in *Proc. Symp. on Math. Stat. and Probability*, 547-561 (University of California Press, Berkeley, 1961).

[11] M. M. Wilde, A. Winter, and D. Yang, Commun. Math. Phys., **331**, 593-622 (2014): arXiv:1306.1586.

Acknowledgement

The authors are grateful to Prof. Tomohiro Ogawa for helpful discussions.