A new quantum version of f-divergence

Keiji Matsumoto

NII,

ERATO-SORST JST

Classical f-divergence

$$D_f(p||q) := \sum_x q_x f(\frac{p_x}{q_x})$$

$f: [0, \infty) \to R$, convex

Ex

$f(t)$	$D_f(p\|\|q)$	
$t \log t$	$\sum_x p_x \log p_x / q_x$	K-L
$1 - t^{1/2}$	$1 - \sum_x p_x^{1/2} q_x^{1/2}$	Affinity
$1 - t^\alpha \ (\alpha \leq 1)$	$1 - \sum_x p_x^\alpha q_x^{1-\alpha}$	Renyi
$t^\alpha \ (\alpha \geq 1)$	$\sum_x p_x^\alpha q_x^{1-\alpha}$	

Convertibility and f-divergence

Th $\exists P \; p' = Pp, q' = Pq$ P: Stochasitic matrix

$\Leftrightarrow \forall f$: proper, closed, convex

$$D_f(p||q) \geq D_f(p'||q')$$

A Quantum f-Divergence

by Petz and Belavkin independently decades ago

$$\Delta_{\rho,\sigma}(X) := \rho X \sigma^{-1} \quad (\leftrightarrow p/q)$$

$$D_f(\rho||\sigma) := \mathrm{tr}\sqrt{\sigma} f\left(\Delta_{\rho,\sigma}\right)\sqrt{\sigma}$$

Ex

$f(t)$	$D_f(p\|\|q)$	
$t \log t$	$\mathrm{tr}\rho(\log \rho - \log \sigma)$	K-L
$1 - t^{1/2}$	$1 - \mathrm{tr}\sqrt{\rho}\sqrt{\sigma}$	(≠ fidelity)
$1 - t^{\alpha} \quad (\alpha \leq 1)$	$1 - \mathrm{tr}\,\rho^{\alpha}\sigma^{1-\alpha}$	Renyi
$t^{\alpha} \quad (\alpha \geq 1)$	$\mathrm{tr}\,\rho^{\alpha}\sigma^{1-\alpha}$	

Properties of PB Divergence

Condition (F)

$$f(\lambda\rho + (1-\lambda)\rho') \le \lambda f(\rho) + (1-\lambda) f(\rho')$$

f(0)=0, finite on [0,∞)

Th If f satisfies (F), D_f satisfies

- **Normalized:** Coincide with its classical version if commutative
- **Monotonicity**

  $D_f(\rho||\sigma) \ge D_f(\Lambda(\rho)||\Lambda(\sigma))$ Λ:CPTP
- **Jointly convex**

Questions

Other analogues satisfying the same conditions

Operational meanings of them

Characterization of all such quantities

Another possible analogue

$$d(\rho, \sigma) := \sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}} \quad (\leftrightarrow p/q)$$

$$D_f^R(\rho || \sigma) := \mathrm{tr}\, \sigma f(d(\rho, \sigma))$$

When f(t)=t log t, defined by Hiai and Petz, Belavkin decades ago

Th When f satisfies (F), D_f^R satisfies normalization, monotonicity, and joint convexity

Normalized: Coincide with its classical version if commutative

Monotonicity

$$D_f(\rho||\sigma) \geq D_f(\Lambda(\rho)||\Lambda(\sigma)) \quad \Lambda\text{:CPTP}$$

Proof sketch of monotinicity

$$d\big(\Lambda(\rho),\Lambda(\sigma)\big) = \Lambda_\sigma(d(\rho,\sigma)),$$

where

$$\Lambda_\sigma(Z) := \{\Lambda(\sigma)\}^{-1/2}\Lambda\left(\sigma^{1/2}Z\sigma^{1/2}\right)\{\Lambda(\sigma)\}^{-1/2}$$

$\Lambda_\sigma(1) = 1$, and completely positive

fact : $f(\Lambda_\sigma(d)) \leq \Lambda_\sigma(f(d))$

Th If f satisfies (F), and D_f^Q is satisfies normalization and monotonicity,

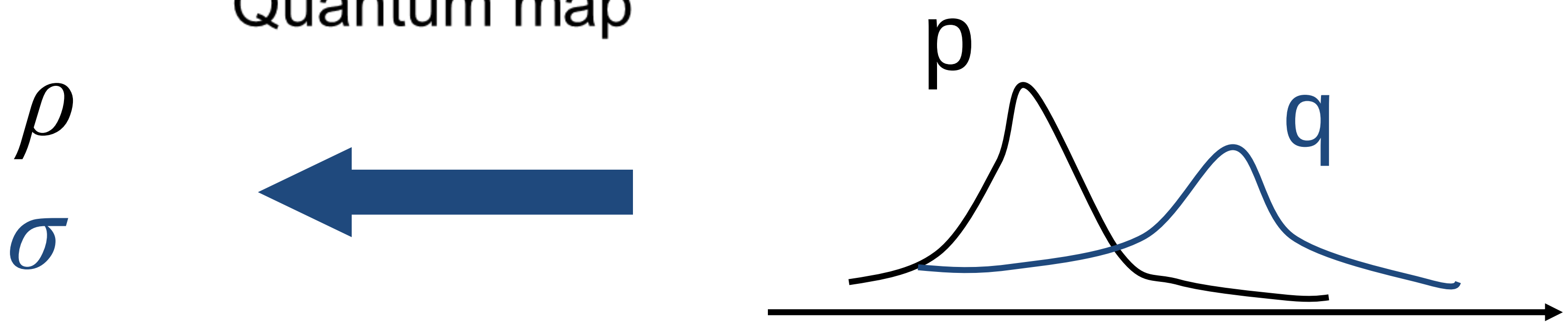
$$D_f^Q \leq D_f^R$$

In other words, D_f^R is the maximal monotone analogue of classical D_f

In the following slides, we give the proof sketch

Key tool: Classical-Quantum map

Γ:Classical
Quantum map



$(\Gamma, \{p, q\})$ is optimized
To minimize $D_f(p||q)$

Th.

$$\min D_f(p||q) = D_f^R(\rho||\sigma)$$

Composition of C-Q map

$d(\rho,\sigma) = \sum_x d_x P_x$ P_x :projector, d_x: eigenvalue

$$q(x) := \mathrm{tr}\,\sigma P_x, \qquad p(x) := d_x q(x)$$

$$\Gamma(\delta_a) := \frac{1}{q_a}\sqrt{\sigma} P_a \sqrt{\sigma}$$

δ_a : delta distribution consentrated at a

Note 1: if commutative, dx=p(x)/q(x)

Note 2: the same (Γ,{p,q}) is the optimal for all f

Proof of $D_f^Q \leq D_f^R$

$$D_f^Q(\rho||\sigma) = D_f^Q(\Gamma(p)\big|\big|\Gamma(q)) \quad \text{(def of Γ,p,q)}$$

$$\leq D_f^Q\ (p||q) \quad \text{(monotonicity)}$$

$$= D_f\ (p||q) \quad \text{(normalization)}$$

$$= D_f^R(\rho||\sigma) \quad \text{(def of Γ,p,q)}$$

Classical-Quantum convertibility

Th There is a CPTP map Γ with

$$\Gamma(p) = \rho, \Gamma(q) = \sigma$$

if and only if

$$D_f^R(\rho||\sigma) \geq D_f(p||q)$$

holds for all f satisfying (F)

Key facts for the proof:

1. For any f, the optimal C-Q map is the same
2. C-C convertibility is characterized by f-divergence

When $\mathbf{Supp}\,\boldsymbol{\rho} \not\leq \mathbf{Supp}\,\boldsymbol{\sigma}$

$$d(\rho, \sigma) := \sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}}$$

- $\rho = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$
- $\rho^* = \rho_{11} - \rho_{12} \rho_{22}^{-1} \rho_{21}$

$$D_f^R(\rho || \sigma) = \operatorname{tr} \sigma f\big(d(\rho^*, \sigma)\big) + \lim_{y \to \infty} f(y)/y$$

How about lower bound?

When $f(t) = 1 - \sqrt{t}$, any D_f^Q satisfying normalization and monotonicity satisfies

$$D_f^Q(\rho||\sigma) \geq 1 - F(\rho, \sigma)$$

F is fidelity $F(\rho, \sigma) := \mathrm{tr}\left(\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}\right)^{\frac{1}{2}}$

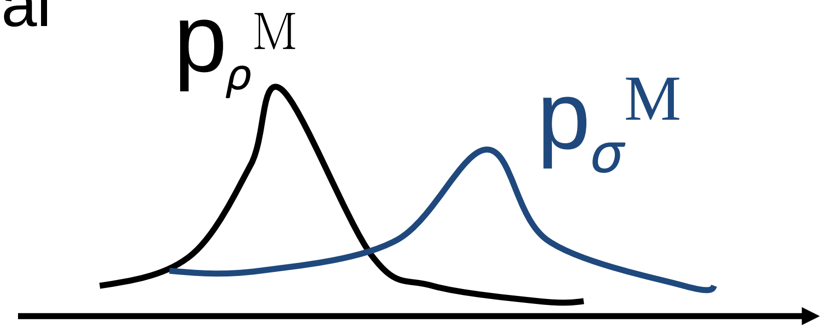
In general, lower bound is given by considering Qtuantum-to-classical CPTP map, or measrurement. (Dual of C-Q map)

Test :Distillation of relative entropy

ρ

σ

M: Measurement
=Quantum Classical map



M is optimized

To maximize $D(p_\rho^M || p_\sigma^M)$

Th. (Hiai-Petz)

$$\max_M \frac{1}{n} D(p^M_{\rho^{\otimes n}} || p^M_{\sigma^{\otimes n}}) = D(\rho||\sigma) + o(1)$$

Relation to RLD Fisher informaton

$$\begin{aligned}
\left.\frac{\mathrm{d}^2}{\mathrm{d}t\,\mathrm{d}s}\mathrm{D}'_f\left(\rho+sX||\rho-tY\right)\right|_{t=0,s=0} &= \left.\frac{\mathrm{d}^2}{\mathrm{d}t\,\mathrm{d}s}\mathrm{D}'_f\left(\rho||\rho+sX+tY\right)\right|_{t=0,s=0} \\
&= \left.\frac{\mathrm{d}^2}{\mathrm{d}t\,\mathrm{d}s}\mathrm{D}'_f\left(\rho+sX+tY||\rho\right)\right|_{t=0,s=0} \\
&= f''(1)\,\Re\,J^R_\rho(X,Y),
\end{aligned}$$

where J^R_ρ is the RLD Fisher metric,

$$J^R_\rho(X,Y) := \mathrm{tr}\,X\rho^{-1}Y.$$

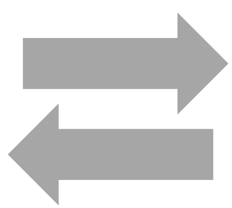
Summary

1. A new analogue of f-divergence is proposed.
2. This gives upper bound to any reasonable quantum analogue of f-divergence
3. In proving the lower bund, classical-to-quantum CPTP map played the central role
4. In composition of C-Q map and the def of D_f^R, $d(\rho,\sigma) := \sigma^{-1/2}\rho\sigma^{-1/2}$, q-analogue of p/q, plays an important role

Outline

1. Asymptotic characterization of quantum relative entropy
2. Non-asymptotic characterization of quantum relative entropy
3. A new version of fidelity
4. On generalized fidelity

Throughout the talk, conversion between quantum state family and probability distribution family plays the key role.

ρ σ  p q