Quantum exponential Orlicz space in information geometry

Anna Jenčová
Mathematical Institute, Slovak Academy of Sciences

a talk based on arxiv:math-ph/0511065 and AJ, On quantum information manifolds, in: Algebraic and Geometric Methods in Statistics, 2010

September 28, 2021

Outline

- Introduction: Information geometry
- Quantum relative entropy and perturbation of states
- Young functions and associated norms
- A quantum exponential Orlicz space
- The quantum information manifold

Introduction: Information geometry

The geometry of statistical models

Parametric statistical model: for a measure space (Ω, Σ, μ),

$$M = \{p_\theta,\ \theta \in \Theta \subseteq \mathbb{R}^n\} \text{ - a family of probability measures}$$

Under some regularity of $\theta \mapsto p_\theta$: a differentiable manifold (Rao, 1945; Jeffreys, 1946)

Additional structures:

- Riemannian structure: Fisher information
- affine connections: respecting exponential families

$$\{p_\theta = \exp(\log p + \sum_i \theta_i u_i - c(\theta))\}$$

or mixture families

$$\{p_\theta = \sum_i \theta_i p_i\}$$

The geometry of statistical models

(Amari & Nagaoka)

- duality of the two connections with respect to the metric
- a family of α-connections, $\alpha \in \mathbb{R}$
- geometric structures closely related to divergence measures (relative entropy, α-divergence)
- uniqueness of the structures respecting statistical maps (Cencov, 1982)

Basic references: Amari & Nagaoka, 2000; Ay et al., 2017

Exponential statistical manifold

(Pistone & Sempi, 1995)

- an infinite dimensional manifold M_p of probability measures equivalent to a given p
- exponential arc: $I \ni t \mapsto p_t := \exp(\log p + tu - c_p(t))$
  u a regular random variable, centered at p:

$$u \in L_\rho := \{u, E_p[\exp(tu)] < \infty,\ t \in I, \quad E_p[u] = 0\},$$

- parametrization: $u \mapsto [p^u] := \exp(\log p + u - c_p(u))$, for some $u \in L_\rho$ such that

$$c_p(u) = \log E_p[\exp(u)] < \infty \text{ - cumulant generating functional}$$

- L_ρ is the subspace of centered elements in the exponential Orlicz space $L_{\exp}(\Omega, \Sigma, p)$, with respect to the Young function

$$\Phi(x) = \cosh(x) - 1$$

Exponential statistical manifold

A Banach manifold structure on M_p modelled on $L_{\exp}(\Omega, \Sigma, p)$:

- connected components are the maximal exponential families: all elements p_1, p_2 can be connected by an open exponential arc

$$p_t = \exp(\log p + tu - c_p(t)),\ t \in (a, b),\ p_i = p_{t_i},\ t_1, t_2 \in (a, b)$$

- parametric models are included as submanifolds
- the geometric structures of the parametric models are induced from the exponential manifold

Quantum information geometry

Finite dimensional quantum extensions (matrix algebras)

- quantum Fisher information: a family of monotone metrics (Petz, 1996)
- exponential and mixture connections, α-connections and duality with respect to monotone metrics, divergences,... (Nagaoka, Hasegawa, Gibilisco & Isola, Grasselli & Streater, AJ,...)
- not so clear interpretation in statistics or information theory

Quantum exponential manifold

Relation to statistical physics motivated the construction of an infinite dimensional quantum exponential manifold:
(Labuschagne & Majewski; Streater)

- geometry of the state space induced from L_1 - not suitable: every neighborhood of a state ρ contains some ρ' with $S(\rho'\|\rho) = \infty$
- $(L_{\exp})_* = L\log(L+1)$: generated by states with finite (relative) entropy
- $L_{\exp}$ (extended) space of observables, $L\log(L+1)$ (restricted) space of states
- geometry closely related to state perturbation and relative entropy

I will follow an approach similar to (Streater, 2004), based on a quantum Young function obtained by state perturbation.

Quantum relative entropy and perturbation of states

Basic setting and notation

- $\mathcal{M}$ a von Neumann algebra (σ-finite)
- $\mathcal{M}_*$ the predual
- $\mathcal{M}^s = \{a = a^* \in \mathcal{M}\}$ self-adjoint part
- $\mathcal{M}_*^s = \{\psi(a^*) = \overline{\psi(a)}\}$ hermitian normal functionals
- $\mathcal{M}_*^+$ the positive cone in $\mathcal{M}_*$
- $\mathfrak{S}_*(\mathcal{M})$ the set of normal states

We fix a faithful normal state $\rho \in \mathfrak{S}_*(\mathcal{M})$.

Haagerup L_p-spaces and a standard form

Haagerup L_p-spaces: $L_p(\mathcal{M})$, $1 \le p \le \infty$, the norm: $\|\cdot\|_p$

- If $\mathcal{M} = B(\mathcal{H})$, can be replaced by Schatten classes:

$$S_p = \{a \in B(\mathcal{H}),\ \mathrm{Tr}\,|a|^p < \infty\}, \quad \|a\|_p = (\mathrm{Tr}\,|a|^p)^{1/p}$$

- $\mathcal{M} \cong L_\infty(\mathcal{M})$, $\mathcal{M}_* \cong L_1(\mathcal{M})$: $\psi \mapsto h_\psi$ ("density operators")
- $L_p(\mathcal{M})^* \cong L_q(\mathcal{M})$, $1/p + 1/q = 1$
- $L_2(\mathcal{M})$ a Hilbert space

Standard form: $(\lambda, L_2(\mathcal{M}), J, L_2(\mathcal{M})^+)$, $\lambda : \mathcal{M} \to B(L_2(\mathcal{M}))$

$$\lambda(a)\xi = a\xi, \quad J\xi = \xi^*, \qquad a \in \mathcal{M},\ \xi \in L_2(\mathcal{M}).$$

$h_\omega^{1/2}$ - (unique) vector representative of $\omega \in \mathcal{M}_*^+$ in $L_2(\mathcal{M})^+$.

Kosaki L_p spaces with respect to ρ

Let $\eta \in [0,1]$. The Kosaki L_p-space with respect to ρ:

$$L_p^\eta(\mathcal{M},\rho) = \{h_\rho^{\eta/q} k h_\rho^{(1-\eta)/q},\ k \in L_p(\mathcal{M})\} \subseteq \mathcal{M}_*$$

$$\text{the norm: } \|h_\rho^{\eta/q} k h_\rho^{(1-\eta)/q}\|_{p,\sigma}^{(\eta)} = \|k\|_p$$

- defined by complex interpolation
- for $1 < p < p' < \infty$: continuous embeddings

$$\mathcal{M} \sqsubseteq L_{p'}^\eta(\mathcal{M},\rho) \sqsubseteq L_p^\eta(\mathcal{M},\rho) \sqsubseteq \mathcal{M}_*$$

- $L_p^\eta(\mathcal{M},\rho)^* \cong L_q^\eta(\mathcal{M},\rho)$, $1/p + 1/q = 1$.

Symmetric L_p-spaces: $\eta = 1/2$

$$L_p(\mathcal{M},\rho) := L_p^{1/2}(\mathcal{M},\rho), \quad \|\cdot\|_{p,\rho} := \|\cdot\|_{p,\rho}^{(1/2)}$$

Relative entropy

For $\omega \in \mathcal{M}_*^+$, the relative modular operator: $\Delta_{\rho, h_\omega^{1/2}}$ (unbounded operator acting on $L_2(\mathcal{M})$)

- Araki relative entropy:

$$S(\omega \| \rho) = -\langle \log(\Delta_{\rho, h_\omega^{1/2}}) h_\omega^{1/2}, h_\omega^{1/2} \rangle$$

- in finite dimensions, this is the same as the Umegaki relative entropy:

$$S(\omega \| \rho) = \mathrm{Tr}\, h_\omega (\log h_\omega - \log h_\rho)$$

Properties of the relative entropy

$\omega \mapsto S(\omega\|\rho)$ is strictly convex, lower semicontinuous

Lower bound:
$S(\omega\|\rho) \geq \omega(1)\log\omega(1)$, with equality iff $\omega = \lambda\rho$, $\lambda \geq 0$

Monotonicity
For a positive unital normal map $T : \mathcal{N} \to \mathcal{M}$, with preadjoint $T_* : \mathcal{M}_* \to \mathcal{N}_*$,

$$S(T_*(\omega)\|T_*(\rho)) \leq S(\omega\|\rho), \qquad \omega \in \mathcal{M}_*^+$$

Relation to Kosaki (symmetric) L_p-spaces

The sandwiched Rényi relative entropy: for $\alpha > 1$

$$\tilde{D}_\alpha(\omega\|\rho) = \frac{1}{\alpha - 1} \log \frac{\|h_\omega\|_{\alpha,\rho}}{\omega(1)}$$

Let $h_\omega \in L_p(\mathcal{M},\rho)^+ = L_p(\mathcal{M},\rho) \cap \mathcal{M}_*^+$, $p > 1$.

- $\tilde{D}_\alpha(\omega\|\rho) < \infty$ for $\alpha \in (1, p]$
- $\alpha \mapsto \tilde{D}_\alpha(\omega\|\rho)$ is nondecreasing on $(1, p]$
- $\lim_{\alpha\downarrow 1} \tilde{D}_\alpha(\omega\|\rho) = \frac{S(\omega\|\rho)}{\omega(1)}$

In particular, $S(\omega\|\rho) < \infty$ for $h_\omega \in L_p(\mathcal{M},\rho)^+$, $p > 1$.

Sets with finite relative entropy

We define

$$\mathcal{P}_\rho := \{\omega \in \mathcal{M}_*^+,\ S(\omega\|\rho) < \infty\}$$
$$\mathcal{S}_\rho := \{\omega \in \mathfrak{S}_*(\mathcal{M}),\ S(\omega\|\rho) < \infty\}$$

Donald's identity: for $\omega_i \in \mathcal{M}_*^+$, $\omega = \sum_i \omega_i$

$$S(\omega\|\rho) + \sum_i S(\omega_i\|\omega) = \sum_i S(\omega_i\|\rho)$$

- $\mathcal{P}_\rho$ is a convex cone, face of $\mathcal{M}_*^+$
- $\mathcal{S}_\rho$ is a base of $\mathcal{P}_\rho$, face of $\mathfrak{S}_*(\mathcal{M})$
- $L_p(\mathcal{M},\rho)^+ \subseteq \mathcal{P}_\rho$ for $p > 1$

Perturbation of states and relative entropy

Let $a \in \mathcal{M}^s$ and

$$c_\rho(a) = \sup_{\omega \in \mathfrak{S}_*(\mathcal{M})} \omega(a) - S(\omega \| \rho)$$

- $c_\rho(a) < \infty$ for all $a \in \mathcal{M}^s$,
- c_ρ is convex and continuous.

The perturbed state

the unique state $[\rho^a] \in \mathcal{S}_\rho$ such that the supremum is attained:

$$c_\rho(a) = [\rho^a](a) - S([\rho^a] \| \rho).$$

Perturbation of states and relative entropy

For all $\omega \in \mathfrak{S}_*(\mathcal{M})$, we have

$$\omega(a) - S(\omega\|\rho) = c_\rho(a) - S(\omega\|[\rho^a])$$

- $\mathcal{S}_\rho = \mathcal{S}_{[\rho^a]}$, $\mathcal{P}_\rho = \mathcal{P}_{[\rho^a]}$
- $[\rho^a]$ is faithful
- In finite dimensions, we have

$$[\rho^a] = \exp(\log\rho + a - c_\rho(a)), \quad c_\rho(a) = \log \mathrm{Tr}\,[\exp(\log\rho + a)].$$

Quantum exponential arc: $\{[\rho^{ta}], t \in I \subseteq \mathbb{R}\}$

Young functions and associated norms

Young functions

Let X be a real vector space.

A Young function: $\Phi : X \to [0, \infty]$ such that

1. Φ is convex,
2. $\Phi(x) = \Phi(-x)$, $\Phi(0) = 0$,
3. if $x \neq 0$, $\lim_{t\to\infty} \Phi(tx) = \infty$.

Young functions

For a Young function Φ, we put

- $C_\Phi := \{x \in X,\ \Phi(x) \leq 1\}$,
- $L_\Phi := \{x \in X,\ \exists s > 0,\ \Phi(sx) < \infty\} = \mathrm{span}(\mathrm{dom}(\Phi))$.

Then L_Φ is a subspace, $C_\Phi \subset L_\Phi$ absolutely convex and absorbing.

The Minkowski functional of C_Φ defines a norm in L_Φ:

$$\begin{aligned} \|x\|_\Phi &= \inf\{s > 0,\ x \in sC_\Phi\} \\ &= \inf\{s > 0, \Phi(x/s) \leq 1\} \end{aligned}$$

Let B_Φ denote the completion of $(L_\Phi, \|\cdot\|_\Phi)$.

Examples

Classical Orlicz spaces:

For $(\Omega, \mathcal{B}, \mu)$ a measure space, $\varphi : \mathbb{R} \to [0, \infty]$ a Young function, put

$$X = \{[f]_\mu, f : \Omega \to \mathbb{R} \text{ measurable}\}, \ \Phi([f]_\mu) = \int_\Omega \varphi(|f|) d\mu.$$

Kosaki L_p^η-spaces:

For $\mathcal{M}$ a von Neumann algebra, ρ a state, $1 < p < \infty$, $\eta \in [0, 1]$:

$$X = \mathcal{M}^s, \quad \Phi(a) = \|h_\rho^{\eta/p} a h_\rho^{(1-\eta)/p}\|_p$$

Trunov, 1979; Kosaki, 1984

Examples

Noncommutative Orlicz spaces with respect to trace

For $(\mathcal{M}, \tau)$ semifinite, $\varphi : \mathbb{R} \to [0, \infty]$ a Young function:

$$X = \{\tau\text{-measurable operators}\}, \quad \Phi(a) = \tau(\varphi(|a|)).$$

Muratov, 1978; Dodds, Dodds & de Pagter, 1989; Kunze 1990

Noncommutative Orlicz space with respect to ρ

For $(\mathcal{M}, \tau)$ semifinite, ρ a state, $\varphi :\to [0, \infty]$ a Young function, $\eta \in [0, 1]$:

$$X = \{\tau\text{-measurable operators}\},$$
$$\Phi(a) = \tau(\varphi(\varphi^{-1}(\rho)^{\eta} a \varphi^{-1}(\rho)^{1-\eta}))$$

Al-Rashed & Zegarlinski, 2007, 2011

The conjugate function and dual space

Further assumptions:

- (X, Y) locally convex spaces in separating duality, with corresponding weak topologies
- Φ lower semicontinuous, continuous at 0
- $\mathrm{dom}(\Phi)$ is dense in X.

Conjugate function

$$\Phi^*(y) = \sup_{x \in X} \langle\, y, x \,\rangle - \Phi(x)$$

is a lower semicontinuous Young function on Y.

The dual space

Under the assumptions:

- $L_\Phi = X$, $\|\cdot\|_\Phi$ is continuous (on X)
- $X \sqsubseteq B_\Phi$
- For the Banach space dual, we have

$$B_\Phi^* \simeq B_{\Phi^*} = L_{\Phi^*} \sqsubseteq Y$$

- Hölder inequality:

$$|\langle y, x \rangle| \leq 2\|y\|_{\Phi^*}\|x\|_\Phi, \qquad y \in B_{\Phi^*}, x \in B_\Phi.$$

A quantum exponential Orlicz space

The exponential Young function and its conjugate

The exponential Young function: put $X = \mathcal{M}^s$, $Y = \mathcal{M}^s_*$

$$\Phi_\rho(a) := \frac{\exp(c_\rho(a)) + \exp(c_\rho(-a))}{2} - 1, \qquad a \in \mathcal{M}^s.$$

- Φ_ρ is a Young function $X \to [0, \infty)$,
- (X, Y) and Φ_ρ satisfy the additional assumptions.
- the conjugate function is

$$\Phi^*_\rho(v) = \frac{1}{2} \inf_{\substack{\omega_1,\omega_2 \in \mathcal{M}^+_* \\ 2v = \omega_1 - \omega_2}} [S(\omega_1 \| \rho) - \omega_1(1) + S(\omega_2 \| \rho) - \omega_2(1)] + 1$$

The quantum exponential Orlicz space

If $\mathcal{M}$ is commutative: $\mathcal{M} \cong L_\infty(\Omega, \Sigma, \rho)$ for a probability space,

$$\Phi_\rho(u) = \int_\Omega \big(\cosh(|u|) - 1\big)\, d\rho, \qquad u \in \mathcal{M}^s.$$

Then B_{Φ_ρ} is the closure of $L_\infty(\Omega, \Sigma, \rho)^s$ in $L_{\exp}(\Omega, \Sigma, \rho)$, denoted by $E_{\exp}(\Omega, \Sigma, \rho)$. We have

- $u \in E_{\exp}$ if and only if $\Phi_\rho(tu) < \infty$ for all $t \in \mathbb{R}$
- $E_{\exp}^{**} = L_{\exp}$.

The quantum exponential Orlicz space:

$$E_{\exp}(\mathcal{M}, \rho) := B_{\Phi_\rho}, \qquad L_{\exp}(\mathcal{M}, \rho) := B_{\Phi_\rho}^{**}$$

We will mostly work with $E_{\exp}$.

The dual space

We know that

$$E^*_{\exp}(\mathcal{M},\rho) = B^*_{\Phi_\rho} \simeq B_{\Phi^*_\rho} \sqsubseteq \mathcal{M}^s_*$$

From

$$\Phi^*_\rho(v) = \frac{1}{2} \inf_{\substack{\omega_1,\omega_2\in\mathcal{M}^+_* \\ 2v=\omega_1-\omega_2}} [S(\omega_1\|\rho) - \omega_1(1) + S(\omega_2\|\rho) - \omega_2(1)] + 1$$

we infer

- $B_{\Phi^*_\rho} = \mathcal{P}_\rho - \mathcal{P}_\rho$, $B_{\Phi^*_\rho} \cap \mathcal{M}^+_* = \mathcal{P}_\rho$
- The unit ball in $B_{\Phi^*_\rho}$:

$$U_\rho := \{\frac{1}{2}(\omega_1-\omega_2),\ \omega_1,\omega_2 \in \mathcal{M}^+_*, \\ S(\omega_1\|\rho) + S(\omega_2\|\rho) \leq \omega_1(1) + \omega_2(1)\}$$

Properties of the quantum exponential Orlicz space

Continuous embeddings: for $p > 1$

$$\mathcal{M}^s \sqsubseteq E_{\exp}(\mathcal{M}, \rho) \sqsubseteq L_p(\mathcal{M}, \rho)^s \sqsubseteq E^*_{\exp}(\mathcal{M}, \rho) \sqsubseteq \mathcal{M}^s_*$$

Alternative definition: let $K_\rho = \{\omega \in \mathcal{M}_*^+, S(\omega\|\rho) \leq \omega(1)\}$.

- K_ρ is convex and weakly compact
- $E_{\exp}(\mathcal{M}, \rho) \simeq A(K_\rho)$ continuous affine functions on K_ρ
- $L_{\exp}(\mathcal{M}, \rho) \simeq A_b(K_\rho)$ bounded affine functions on K_ρ

Positive unital normal maps: $T : \mathcal{N} \to \mathcal{M}$

extends to a contraction $E_{\exp}(\mathcal{N}, T_*\rho) \to E_{\exp}(\mathcal{M}, \rho)$.

The quantum information manifold

The quantum information manifold

Let $\mathcal{F}(\mathcal{M}) \subset \mathfrak{S}_*(\mathcal{M})$ be the set of all faithful normal states on $\mathcal{M}$.

A Banach manifold structure on $\mathcal{F}(\mathcal{M})$: a C^∞-atlas

a family of pairs $\{U_i, e_i\}$ such that

- $U_i \subset \mathcal{F}(\mathcal{M})$, $\cup_i U_i = \mathcal{F}(\mathcal{M})$
- $e_i : U_i \to e_i(U_i)$ a bijection onto an open subset of a Banach space B_i
- for all i, j, $e_i(U_i \cap U_j)$ is open in B_i
- for all i, j, $e_j e_i^{-1} : e_i(U_i \cap U_j) \to e_j(U_i \cap U_j)$ is a C^∞-isomorphism

We will construct a C^∞ atlas using the map $a \mapsto [\rho^a]$.

The extended functional and perturbed state

For $a \in E_{\exp}(\mathcal{M}, \rho)$, put

$$c_\rho(a) := \sup_{\omega \in \mathcal{S}_\rho} \omega(a) - S(\omega\|\rho)$$

Then c_ρ is well defined and

- finite valued (over $E_{\exp}$)
- attained at a unique state $[\rho^a] \in \mathcal{S}_\rho$, faithful
- for all $a \in E_{\exp}(\mathcal{M}, \rho)$:

$$\omega(a) - S(\omega\|\rho) = c_\rho(a) - S(\omega\|[\rho^a])$$

- $\mathcal{P}_\rho = \mathcal{P}_{[\rho^a]}$, $E_{\exp}(\mathcal{M}, \rho) \simeq E_{\exp}(\mathcal{M}, [\rho^a])$
- for $\lambda \in \mathbb{R}$: $[\rho^{a+\lambda}] = [\rho^a]$, $c_\rho(a+\lambda) = c_\rho(a) + \lambda$.

Subspace of centered elements

We define the subspace of centered elements:

$$E_{\exp,0}(\mathcal{M},\rho) = \{a \in E_{\exp}(\mathcal{M},\rho), \rho(a) = 0\}$$

The cumulant generating functional: $\bar{c}_\rho := c_\rho|_{E_{\exp,0}}$

- $\bar{c}_\rho(a) = S(\rho\|[\rho^a])$
- $[\rho^a](a) = S([\rho^a]\|\rho) + S(\rho\|[\rho^a])$
- $\bar{c}_\rho$ is strictly convex and Gateaux differentiable
- The chain rule: for $a, b \in E_{\exp,0}(\mathcal{M},\rho)$

$$\bar{c}_\rho(a+b) = \bar{c}_\rho(b) + \bar{c}_{[\rho^a]}(b), \quad [\rho^{a+b}] = [[\rho^a]^b]$$

- $[\rho^a] = [\rho^b]$ if and only if $a = b$.

Subspace of centered elements

An equivalent norm in $E_{\exp,0}(\mathcal{M},\rho)$: put

$$\Psi_\rho(a) = \bar{c}_\rho(a) + \bar{c}_\rho(-a), \quad a \in E_{\exp,0}(\mathcal{M},\rho)$$

Then

- Ψ_ρ is a Young function on $E_{\exp,0}(\mathcal{M},\rho)$
- $B_{\Psi_\rho} \simeq E_{\exp,0}(\mathcal{M},\rho)$.

The dual space: $E^*_{\exp,0} = E^*_{\exp}|_{\{\rho\}}$, with unit ball

$$\{[\omega_1 - \omega_2],\ S(\omega_1\|\rho) + S(\omega_2\|\rho) \le 1\}$$

The map $\mathcal{S}_\rho \to E^*_{\exp,0}$, $\omega \mapsto [\omega - \rho]$ is one-to-one

The C^∞-atlas

For $\rho \in \mathcal{F}(\mathcal{M})$, let s_ρ be the map $a \mapsto [\rho^a]$, $a \in E_{\exp,0}(\mathcal{M}, \rho)$. Put

- $U_\rho = s_\rho(V_\rho)$, V_ρ the open unit ball in $E_{\exp,0}(\mathcal{M}, \rho)$
- $e_\rho = s_\rho^{-1} : U_\rho \to V_\rho$

Using the chain rule, we can show that

- $\{e_\rho, U_\rho\}_{\rho \in \mathcal{F}(\mathcal{M})}$ is a C^∞- atlas on $\mathcal{F}_\rho$
- connected components are of the form

$$\{[\rho^a],\ a \in E_{\exp,0}(\mathcal{M}, \rho)\}$$

for $\rho \in \mathcal{F}(\mathcal{M})$.

Further developments and open questions

- an extension to $L_{\mathrm{exp}}(\mathcal{M}, \rho)$
- a description of the connected components

$$\{[\rho^a], a \in E_{\mathrm{exp},0}(\mathcal{M}, \rho)\}$$

- a mixture manifold structure, obtained from the map

$$\mathcal{S}_\rho \to E^*_{\mathrm{exp},0}(\mathcal{M}, \rho), \qquad \omega \mapsto [\omega - \rho]$$

- compatibility of the two structures
- the induced topologies on $\mathcal{F}(\mathcal{M})$