Int J Software Informatics, Volume 8, Issue 3-4 (2014), pp. 265–276
International Journal of Software and Informatics, ISSN 1673-7288

E-mail: ijsi@iscas.ac.cn
http://www.ijsi.org
Tel: +86-10-62661040

Fisher Information and Means: Some Questions in the Classical and Quantum Settings

Paolo Gibilisco

(Department of Economics and Finance, University of Rome Tor Vergata Via Columbia 2, Rome, Italy)

Abstract The algebraic and geometric properties of Fisher information, its relations with the theory of operator means have been very active fields in the last decades. In this paper I try to shed some light on recent results and to address some open questions which could in principle give some new directions to the field.
Key words: Fisher information; operator monotone functions; operator means; uncertainty principle; α-connections; Stam inequality

Gibilisco P. Fisher information and means: some questions in the classical and quantum settings. *Int J Software Informatics*, **Vol.8, No.3-4 (2014): 265–276. http://www.ijsi.org/1673-7288/8/i195.htm**

1 Introduction

The concept of mean is one of the oldest in mathematics and we find in the ancient greeks mathematics several examples of means. Nevertheless the field, in its very different ramifications, is a very active one from the point of view of mathematical research. Fisher information (FI), introduced by the founder of modern statistics in 1925, is certainly a very new concept on the appropriate scale of time.

It is now clear that FI is going to play a role in the scientific and mathematical research comparable to the pervasive role of the much older concept of entropy (to which FI is deeply linked). This is testified, for example, by the fact that FI plays an important, albeit different, role in the works of the two recent Fields Medallists, Perelman and Villani. The fact that entropy and FI share unexpected geometric features was the basis for the birth of information geometry.

Actually there are many 'proofs' that the geometric features of probabilistic-statistical objects are sometimes central to their understanding. One of the most striking one is certainly the passage from classical FI to quantum FI: in this case only by looking at FI as a Riemannian metric can one develop a satisfying non-commutative theory. A crucial instruments for this passage is the theory of operator means by Kubo-Ando.

On the other hand Fisher information appears also rooted in algebra because a quite important result like the Stam inequality is group theoretical in character.

In this paper I will address some of this algebraic, geometric and means-related features of Fisher information discussing some new results and some open problems in the field.

Corresponding author: Paolo Gibilisco, Email: gibilisco@volterra.uniroma2.it
Received 2014-12-12; Accepted 2014-12-17.

2 Preliminary Notion: Means and Fisher information

2.1 Classical Fisher information

If X is a real random variable with a strictly positive differentiable density ρ then the *score* and the *Fisher information* are defined respectively as

$$J_X := J_\rho := \frac{\rho'}{\rho} \quad I_X := I_\rho := \mathrm{Var}_\rho(J_\rho) = \int_{\mathbb{R}} \frac{(\rho')^2}{\rho} \tag{1}$$

2.2 Means for positive numbers

Let $\mathbb{R}^+ = (0, +\infty)$. A *mean* for pair of positive numbers is a function $m(\cdot,\cdot) : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ such that

i) $m(x,x) = x$;

ii) $m(x,y) = m(y,x)$;

iii) $x < y \implies x < m(x,y) < y$;

iv) $x < x' \;\; y < y' \implies m(x,y) < m(x',y')$;

v) $m(\cdot,\cdot)$ is continuous;

vi) for $t > 0$ one has $m(tx, ty) = t \cdot m(x,y)$.

Set

$$\mathcal{M}_{nu} := \{m(\cdot,\cdot) : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+ | m \text{ is a mean } \}$$

$\mathcal{F}_{nu}$ is the class of functions $f(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ such that

i) $f(1) = 1$;

ii) $tf(t^{-1}) = f(t)$;

iii) $t \in (0,1) \implies f(t) \in (0,1)$;

iv) $t \in (0,\infty) \implies f(t) \in (0,\infty)$;

v) f is continuous;

vi) f is monotone increasing.

It is straightforward to prove the following

Proposition 2.1. There is bijection between $\mathcal{M}_{nu}$ and $\mathcal{F}_{nu}$ given by the formulas

$$m_f(x,y) := yf(xy^{-1})$$
$$f_m(t) := m(1,t)$$

Here we have some examples of means

Table 1 Means and representing functions

Name	f	m_f
arithmetic	$\frac{1+x}{2}$	$\frac{x+y}{2}$
	$\frac{1}{2}(x^\beta + x^{1-\beta}) \quad \beta \in (0,1)$	$\frac{1}{2}(x^\beta y^{1-\beta} + x^{1-\beta} y^\beta)$
geometric	$\sqrt{x}$	$\sqrt{xy}$
logarithmic	$\frac{x-1}{\log x}$	$\frac{x-y}{\log x - \log y}$
harmonic	$\frac{2x}{x+1}$	$\frac{2}{\frac{1}{x}+\frac{1}{y}}$

2.3 Operator monotone functions and operator means

Let us denote by M_n the complex matrices $n \times n$. A function $f : (0, +\infty) \to R$ is said *operator monotone* iff $\forall A, B \in M_n$ and $\forall n = 1, 2, \ldots$

$$0 \leqslant A \leqslant B \quad \Longrightarrow \quad 0 \leqslant f(A) \leqslant f(B).$$

Often it is useful to restrict to o. m. functions which are: i) normalized (namely $f(1) = 1$); ii) symmetric (namely $tf(t^{-1}) = f(t)$). Let us denote by $\mathcal{F}_{op}$ such family of o.m. functions. The functions in the previous list all belong to $\mathcal{F}_{op}$.

Now set $\mathcal{D}_n := \{A \in M_n | A > 0\}$. An *operator mean* (according Kubo-Ando) is a function $m : \mathcal{D}_n \times \mathcal{D}_n \to \mathcal{D}_n$ such that:

(i) $m(A, A) = A$;

(ii) $m(A, B) = m(B, A)$;

(iii) $A < B\ l; \Longrightarrow \quad A < m(A, B) < B$;

(vi) $A < A', B < B' \quad \Longrightarrow ;\ m(A, B) < m(A', B')$;

(v) m is continuous;

(vi) $Cm(A, B)C^* \leqslant m(CAC^*, CBC^*)$, for all $C \in M_n$.

Let $\mathcal{M}_{op}$ be this family of operator means. Kubo-Ando proved in 1980 the following: there exists a bijection from $\mathcal{M}_{op}$ to $\mathcal{F}_{op}$ given by the formula

$$m_f(A, B) := A^{\frac{1}{2}} f(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}}.$$

Now define

$$\mathcal{F}^r_{op} := \{f \in \mathcal{F}_{op} | f(0) > 0\} \quad \mathcal{F}^n_{op} := \{f \in \mathcal{F}_{op} | f(0) = 0\} \tag{2}$$

Obviously

$$\mathcal{F}_{op} = \mathcal{F}^r_{op} \dot{\cup} \mathcal{F}^n_{op}.$$

Set

$$\tilde{f}(x) := \frac{1}{2}\left[(x+1) - (x-1)^2 \frac{f(0)}{f(x)}\right] \quad x > 0. \tag{3}$$

Theorem 2.1. The correspondence $f \to \tilde{f}$ is a bijection between $\mathcal{F}^r_{op}$ and $\mathcal{F}^n_{op}$.

2.4 Fisher information as Riemannian metric

Rao was the first to realize in 1945 that a statistical model M (a family of densities), under certain conditions, can be considered aa a Riemannian manifold where Fisher information plays the role of the metric. In such approach the ρ' are the tangent vectors and the metric $g_{\rho,F}(\cdot,\cdot)$ is related to Fisher information I_ρ according to the relation

$$g_{\rho,F}(\rho',\rho') := \int_{\mathbb{R}} \frac{(\rho')^2}{\rho} = I_\rho.$$

In the discrete case on the simplex

$$\mathcal{P}_n^1 := \left\{ \rho \in R^n \;\middle|\; \sum_i \rho_i = 1, \quad \rho_i > 0 \right\},$$

we have as tangent space

$$T\mathcal{P}_n^1 = \left\{ u \in R^n \;\middle|\; \sum_i u_i = 0 \right\}.$$

The so called *Fisher-Rao metric* is the scalar product

$$g_{\rho,F}(u,v) := \sum_i \frac{u_i v_i}{\rho_i}.$$

2.5 The quantum Fisher information

On the simplex $\mathcal{P}_n^1$, Chentsov theorem proves that Fisher metric is the unique Riemannian metric contracting under any "coarse graining" T (trace-preserving, positive, linear map).

This means that for any tangent vector X in the point ρ we have

$$g^m_{T(\rho)}(TX,TX) \leqslant g^n_\rho(X,X)$$

In the non-commutative case we have a similar result. Set

$$D_n^1 := \{\rho \in M_n | \mathrm{Tr}(\rho) = 1, \;\; \rho > 0\} = \text{Space of faithful states}$$

then we call *quantum Fisher information* (QFI) a Riemannian metric on D_n^1 contracting under any quantum "coarse graining" T (a trace-preserving, completely positive, linear map), namely we ask that

$$g^m_{T(\rho)}(TA,TA) \leqslant g^n_\rho(A,A).$$

Introducing left and right multiplication operators

$$L_\rho(A) := \rho A, \quad R_\rho(A) := A\rho.$$

we can state Petz theorem (1996) saying that the formula

$$\langle A,B\rangle_{\rho,f} := \mathrm{Tr}(A \cdot m_f(L_\rho,R_\rho)^{-1}(B)).$$

establishes a bijection between QFI and operator monotone functions.

3 Stam Inequality

3.1 Fisher information and Stam inequality on $\mathbb{R}$

Let us suppose that $I_X,\ I_Y < \infty$. The Stam inequality says that if $X, Y : (\Omega, \mathcal{F}, p) \to \mathbb{R}$ are independent random variables then

$$\frac{1}{I_{X+Y}} \geq \frac{1}{I_X} + \frac{1}{I_Y}, \tag{4}$$

with equality if and only if X, Y are Gaussian.

Stam inequality has been generalized few years ago by Madiman and Barron. Moreover Voiculescu gave a free version of the result: in the free case one have equality for the Wigner semicircular distribution. This is not surprising because the Wigner distribution maximizes free entropy the same way the Gaussian maximize ordinary entropy. Despite it was clear the group theoretical character of the inequality only in 1993 it appears a result pointing in this direction: Papathanasiou prove a version of Stam inequality where the additive group of the real numbers $\mathbb{R}$ is substituted by the group of the integers $\mathbb{Z}$. In such case the equality case correspond to the Poisson distribution.

3.2 Stam inequality on Lie groups

The first result for a Lie group different from $\mathbb{R}$ appear in Ref. [7].

If $\rho : S^1 \to \mathbb{R}$ is a density on the circle the tangential derivative is defined as

$$D_T\rho(z) := \lim_{h\to 0} \frac{1}{h}\left[\rho(ze^{ih}) - \rho(z)\right].$$

It is straightforward to define score and Fisher information as

$$J_\rho := \frac{D_T\rho}{\rho}, \quad I_\rho := I_X := \mathrm{Var}_\rho(J_\rho) = \mathbb{E}_\rho[J_\rho^2]. \tag{5}$$

It is possible to prove that if $X, Y : (\Omega, \mathcal{F}, p) \to \mathbb{S}^1$ are independent random variables with densities ρ, σ then

$$\frac{1}{I_{\rho*\sigma}} \geqslant \frac{1}{I_\rho} + \frac{1}{I_\sigma}$$

where one has equality for the uniform distribution case.

How far can we go in this direction?

In Ref. [2] Chirikjian has proved a Stam-like inequality for commuting densities on a unimodular Lie group. But Jupp has sketched in Ref. [9] a proof that would hold in *any* Lie group!

3.3 Stam inequality on finite groups

What happens in the discrete group case? For the cyclic group $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ score and Fisher information are defined as

$$J_\rho(k) := \frac{\rho(k) - \rho(k-1)}{\rho(k)}, \quad I_\rho := I_X := \mathrm{Var}_f(J_\rho) = \mathbb{E}_\rho[J_\rho^2]. \tag{6}$$

Also in this case if $X, Y : (\Omega, \mathcal{F}, p) \to \mathbb{Z}_n$ are independent random variables then

$$\frac{1}{I_{X+Y}} \geq \frac{1}{I_X} + \frac{1}{I_Y}, \tag{7}$$

with equality for the uniform ditribution (maximum entropy distribution as usual). The proof given is based on a standard property of the score: if $Z := X + Y$ then

$$J_Z(Z) = \mathbb{E}_p[J_X(X)|Z] = \mathbb{E}_p[J_Y(Y)|Z]. \tag{8}$$

A straightforward extension can be give for finite abelian group but it is an open problem how to produce a similar result for arbitrary finite groups. Possibly the ideas from the Jupp work on general Lie groups can help in this direction.

3.4 Quantum Stam inequality

Recently also a quantum version has been given in Ref. [10]. The result appears promising but it use only (so to say) a particular version of the quantum Fisher information: the authors consider the Hessian of the Umegaki relative entropy. This implies they are talking about the Bogoliubov-Kubo-Mori metric.

It is natural to ask: can a general Stam inequality holds for an arbitrary QFI?

3.5 Stam inequality for the Gamma distribution: a mean inequality?

Let $\mathbb{R}^+ := (0 + \infty)$ and $f : \mathbb{R}^+ \to \mathbb{R}$ be a differentiable, strictly positive density. Define

$$I_X := \int_{\mathbb{R}^+} (f'(x)/f(x))^2 f(x) dx. \tag{9}$$

Let us suppose that $I_X,\ I_Y < \infty$. We may ask: if $X, Y : (\Omega, \mathcal{F}, p) \to \mathbb{R}^+$ are independent random variables is the inequality

$$\frac{1}{I_{X+Y}} \geq \frac{1}{I_X} + \frac{1}{I_Y}, \tag{10}$$

true?

Let's see what happens for a particular case. Let $\alpha > 0$, the *Gamma function* is defined as

$$\Gamma(\alpha) := \int_0^{+\infty} x^{\alpha-1} e^{-x} \mathrm{d}x.$$

We have the property

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha).$$

Let $\lambda > 0$. The *Gamma density* is defined for $x > 0$ as

$$f(x) := \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}.$$

and 0 otherwise. If $X \sim \Gamma(\alpha, \lambda)$ and $Y \sim \Gamma(\beta, \lambda)$ are independent then $X + Y \sim \Gamma(\alpha + \beta, \lambda)$. The case $\alpha = 1$ is known as the *exponential distribution*.

Proposition 3.1. If $\alpha = 1$ or $\alpha > 2$ then

$$X \sim \Gamma(\alpha, \lambda) \implies I_X = \frac{\lambda^2}{|\alpha - 2|}. \tag{11}$$

If $\alpha \neq 1$ and $\alpha < 2$ then $I_X = +\infty$.

Proof i) Case $\alpha = 1$ that is $|\alpha - 2| = 1$. Suppose X has as density $f(x) = \lambda e^{-\lambda x}$. We have $f'(x) = -\lambda^2 e^{-\lambda x}$ and therefore

$$I_X = \int_0^{+\infty} \frac{f'(x)^2}{f(x)} \mathrm{d}x = \int_0^{+\infty} \lambda^2 f(x) \mathrm{d}x = \lambda^2,$$

that was to be proved.

ii) Case $\alpha > 2$, let $X \sim \Gamma(\alpha, \lambda)$. We have

$$f'(x) = f(x) \left[\frac{\alpha - 1}{x} - \lambda \right],$$

so that

$$\frac{f'(x)^2}{f(x)} = f(x) \left(\frac{\alpha - 1}{x} - \lambda \right)^2.$$

For $\alpha - 1 > 0$, using the functional equation of Γ-function, we have

$$\begin{aligned}\int_0^{+\infty} f(x) \left(\frac{\alpha - 1}{x} \right) \mathrm{d}x &= (\alpha - 1) \int_0^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{(\alpha-1)-1} e^{-\lambda x} \\ &= (\alpha - 1) \frac{\lambda}{\alpha - 1} \int_0^{+\infty} \frac{\lambda^{\alpha-1}}{\Gamma(\alpha - 1)} x^{(\alpha-1)-1} e^{-\lambda x} = \lambda.\end{aligned}$$

Instead for $\alpha - 2 > 0$ we have

$$\begin{aligned}\int_0^{+\infty} f(x) \left(\frac{\alpha - 1}{x} \right)^2 \mathrm{d}x &= (\alpha - 1)^2 \int_0^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{(\alpha-2)-1} e^{-\lambda x} \\ &= \frac{(\alpha - 1)^2 \lambda^2}{(\alpha - 1)(\alpha - 2)} \int_0^{+\infty} \frac{\lambda^{\alpha-2}}{\Gamma(\alpha - 2)} x^{(\alpha-2)-1} e^{-\lambda x} = \lambda^2 \frac{\alpha - 1}{\alpha - 2}.\end{aligned}$$

Therefore if $\alpha > 2$ one gets

$$\begin{aligned}I_X &= \int_0^{+\infty} \frac{f'(x)^2}{f(x)} \mathrm{d}x = \int_0^{+\infty} f(x) \left(\frac{\alpha - 1}{x} - \lambda \right)^2 \mathrm{d}x \\ &= \int_0^{+\infty} f(x) \left(\frac{\alpha - 1}{x} \right)^2 \mathrm{d}x + \lambda^2 \int_0^{+\infty} f(x) \mathrm{d}x - 2\lambda \int_0^{+\infty} f(x) \left(\frac{\alpha - 1}{x} \right) \mathrm{d}x \\ &= \lambda^2 \left(\frac{\alpha - 1}{\alpha - 2} \right) + \lambda^2 - 2\lambda^2 = \frac{\lambda^2}{\alpha - 2}.\end{aligned}$$

iii) Case $\alpha < 2, \alpha \neq 1$. We have

$$\int_0^{+\infty} f(x) \left(\frac{\alpha - 1}{x} \right)^2 \mathrm{d}x = (\alpha - 1)^2 \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} x^{(\alpha-2)-1} e^{-\lambda x} = +\infty$$

and this ends the proof. $\square$

Proposition 3.2 If $X \sim \Gamma(\alpha, \lambda)$, $Y \sim \Gamma(\beta, \lambda)$, X, Y are independent and $I_X, I_Y, I_{X+Y} < \infty$ then

$$\frac{1}{I_{X+Y}} \geq \frac{1}{I_X} + \frac{1}{I_Y}. \tag{12}$$

We have the equality iff only one between X and Y has exponential distribution.

Proof First case: $\alpha, \beta > 2$. We have

$$\frac{1}{I_{X+Y}} = \frac{\alpha+\beta-2}{\lambda^2} > \frac{\alpha-2}{\lambda^2} + \frac{\beta-2}{\lambda^2} = \frac{1}{I_X} + \frac{1}{I_Y} \tag{13}$$

Second case: $\alpha = 1, \beta > 2$. We have

$$\frac{1}{I_{X+Y}} = \frac{(\beta+1)-2}{\lambda^2} = \frac{1}{\lambda^2} + \frac{\beta-2}{\lambda^2} = \frac{1}{I_X} + \frac{1}{I_Y} \tag{14}$$

□

Note that exponential can be characterized by the maximum entropy property as it should be in this case.

Since

$$I_{\alpha Z} = \frac{1}{\alpha^2} I_Z$$

The Stam inequality for the Gamma family takes the form

$$m_h(I_X, I_Y) = \frac{2}{\frac{1}{I_X} + \frac{1}{I_Y}} \geqslant 2I_{X+Y} = 2 \cdot \frac{1}{4} \cdot I_{\frac{X+Y}{2}} = \frac{1}{2} I_{m_a(X,Y)}$$

Now let $f(x) := \frac{x+1}{2}$ so that $f(0) = \frac{1}{2}$. The above inequality takes the form

$$m_{\tilde{f}}(I_X, I_Y) \geqslant f(0) \cdot I_{m_f(X,Y)}$$

Is this only a coincidence? Or is this just an example of a more general theorem?

4 Rao Inequality for the Harmonic Mean

Let x, y be positive real numbers. The arithmetic, geometric and harmonic means are defined as

$$m_a(x,y) = \frac{x+y}{2}, \quad m_g(x,y) = \sqrt{xy}, \quad m_h(x,y) = \frac{2}{\frac{1}{x} + \frac{1}{y}}.$$

Suppose $X, Y \colon \Omega \to (0, +\infty)$ are positive random variables. Then trivially we have that

$$\mathbb{E}(m_a(X,Y)) = m_a(\mathbb{E}(X), \mathbb{E}(Y)).$$

On the other hand the Cauchy-Schwartz inequality implies

$$\mathbb{E}(m_g(X,Y)) \leqslant m_g(\mathbb{E}(X), \mathbb{E}(Y)).$$

Working on a result by Fisher on ancillary statistics Rao proved the following result using Hölder inequality (and the harmonic-geometric inequality).

Proposition 4.1.[16,17]

$$\mathbb{E}(m_h(X,Y)) \leqslant m_h(\mathbb{E}(X), \mathbb{E}(Y)). \tag{15}$$

It is natural to ask how general is this result. In the paper[8] it has been recently proved that the result is true (in a sense to be specified) for *any* means in the commutative as well non-commutative setting (the matrix case for the harmonic mean was anticipated by Prakasa Rao and Rao in Refs. [15,18]).

5 Quantum Uncertainty and Quantum Fisher Information

In Ref. [13] S. Luo started from the Wigner-Yanase information and from an orthonormal basis $\{H_j\}$ to introduce the quantity

$$Q^{WY}(\rho) := \sum_j I_\rho^{WY}(H_j)$$

as a measure to describe the quantum uncertainty of the state ρ. Luo proved that $Q^{WY}(\rho)$ is basis independent and moreover that

$$Q^{WY}(\rho) = \sum_{j<k} \left(\sqrt{\lambda_j} - \sqrt{\lambda_k}\right)^2$$

where $\{\lambda_j\}$ is the spectrum of ρ. Applications of the function $Q^{WY}(\rho)$ appear also in Ref. [14].

Remembering that the WY information is the QFI associated to the functions

$$f_{WY}(x) := \left(\frac{1+\sqrt{x}}{2}\right)^2 \qquad \tilde{f}_{WY} = \sqrt{x}$$

One has

$$Q^{WY}(\rho) = 2\sum_{j<k}\left[\frac{\lambda_j+\lambda_k}{2} - \sqrt{\lambda_j\lambda_k}\right] = 2\sum_{j<k}\left[m_a(\lambda_j,\lambda_k) - m_{\tilde{f}_{WY}}(\lambda_j,\lambda_k)\right]$$

From the above remarks one is leaded to the following questions.

i) For a regular $f \in \mathcal{F}_{op}^r$ does the definition

$$Q^f(\rho) := \sum_j I_\rho^f(H_j)$$

give a basis independent function of the state ρ?

ii) If i) has a positive answer one may ask if

$$Q^f(\rho) = 2\sum_{j<k}\left[m_a(\lambda_j,\lambda_k) - m_{\tilde{f}}(\lambda_j,\lambda_k)\right].$$

iii) Finally one should study if the properties of the function Q^{WY} are specific or general along the lines of (for example) Ref. [12] where it is proved that to detect entaglement Wigner-Yanase information and SLD-information have very different behaviour.

6 α-Connections, L^p Spheres and Complete Integrability for the Generalized Proudman-Johnson Equation

Recently Fisher information and α–geometries has been introduced into the realm of diffeomorphism group, see Ref. [11] and reference therein. In particular, in this context has been proved that the geodesic equation of the α-connections are given by a generalized α-Proudman-Johnson equation

$$u_{txx} + (2-\alpha)u_x u_{xx} + u u_{xxx} = 0$$

and the complete integrability of the case $\alpha = 0, \pm 1$ has been established.

One should integrate this results with approach in Ref. [3] where α–geometries are shown to correspond to the geometry of the L^p sphere. In this way one should be able to explicitly describe geodesics and to discuss complete integrability for any α.

7 α-Geometries and the Diffeomorphism Group

A version of Chentsov theorem on uniqueness of Fisher information ha been proved in Ref. [1] by Bauer, Bruveris and Michor. The authors establish (under specific conditions) that invariance under the action of the diffeomorphism group characterizes the Fisher metric on the densities over a manifold. Since also α-connections appear in this context (see Ref. [11]) possibly one can imagine an α-version of the result.

8 The Dynamical Uncertainty Principle in a Classical Setting

Let $A_1, \ldots, A_n$ be observables namely s.a. $n \times n$ matrices and let ρ be a state. The expectation is defined as $\mathbb{E}_\rho(A) := \mathrm{Tr}(\rho A)$. We may define covariance an variance as

$$\mathrm{Cov}_\rho(A,B) := \mathbb{E}_\rho\left(\frac{AB+BA}{2}\right) - \mathbb{E}_\rho(A)\mathbb{E}_\rho(B), \qquad \mathrm{Var}_\rho(A) := \mathrm{Cov}_\rho(A,A).$$

The more general form of the uncertainty principle has been given by Robertson in 1934 as

$$\det\{\mathrm{Cov}_\rho(A_h, A_j)\} \geqslant \det\left\{-\frac{i}{2}\mathrm{Tr}(\rho[A_h, A_j])\right\}, \tag{16}$$

where $h, j = 1, \ldots, N$. For $N = 2$ one gets the Schrödinger uncertainty principle

$$\mathrm{Var}_\rho(A) \cdot \mathrm{Var}_\rho(B) - \mathrm{Cov}_\rho(A,B)^2 \geqslant \frac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2, \tag{17}$$

which implies the Heisenberg uncertainty principle

$$\mathrm{Var}_\rho(A) \cdot \mathrm{Var}_\rho(B) \geqslant \frac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2. \tag{18}$$

Since the matrix $\{-\frac{i}{2}\mathrm{Tr}(\rho[A_h, A_j])\}$ é is antisymmetric the uncertainty principle, should be formulated as

$$\det\{\mathrm{Cov}_\rho(A_h, A_j)\} \geqslant 0 \quad N = 2n+1 \tag{19}$$

$$\det\{\mathrm{Cov}_\rho(A_h, A_j)\} \geqslant \det\{-\frac{i}{2}\mathrm{Tr}(\rho[A_h, A_j])\} \quad N = 2n \tag{20}$$

The left hand side is known as *generalized variance* and for N odd the uncertainty principle say the trivial fact that it cannot be negative, a classical fact.

One can prove an uncertainty principle which does not have such inconvenient:

$$\det\{\mathrm{Cov}_\rho(A_j, A_k)\} \geqslant \det\left\{\frac{f(0)}{2}\langle i[\rho, A_j], i[\rho, A_k]\rangle_{\rho,f}\right\}. \tag{21}$$

where $\langle\cdot,\cdot\rangle_{\rho,f}$ is the QFI associated to the function f according to Petz theorem.

The inequality (21) is known as *dynamical uncertainty principle* since the r.h.s. is a measure of the dissimilarity of the quantum trajectories generated by the observables $A_1, \ldots, A_N$. Please note that it is the first inequality of this kind giving a non-trivial bound for an odd number of observables (see Ref. [6]).

The first cases of this inequality are due to S. Luo.

Can the dynamical uncertainty principle have a classical counterpart?

Now, while it is difficult to imagine a classical counterpart for the standard uncertainty principle the same is not true for the dynamical one, namely for the inequality 21. Indeed the right hand side of the inequality is the volume given by the tangent vectors associated to different evolutions of the state ρ with respect to the different "Hamiltonians" A_j. The volume is that associated to the Fisher information seen as Riemannian metric. All these object can, in principle, exist in a classical setting where Poisson brackets are associated to a dynamics of states. What if the termodynamical uncertainty relations (see Ref. [19] for example) could be derived by a similar mechanism?

Acknowledgement

It is a pleasure to thank M. De Gosson, F. Hansen, F. Luef, S. Luo and G. Misiolek for stimulating discussions on the subjects of the present paper. Of course only the author should be blamed for mistakes and very speculative ideas.

References

[1] Bauer M, Bruveris M, Michor PW. Uniqueness of the Fisher-Rao metric on the space of smooth densities. 2014. arXiv:1411.5577.

[2] Chirikjian GS. Information-theoretic inequalities on unimodular Lie groups. J. Geom. Mech., 2010, 2: 119–158.

[3] Gibilisco P, Isola T. Connections on statistical manifolds of density operators by geometry of non-commutative L^p-spaces. Inf. Dim. Anal. Quan. Prob. Rel. Top., 1999, 2(1): 169-178.

[4] Gibilisco P, Isola T. Fisher information and Stam inequality on a finite group. Bull. Lond. Math. Soc., 2008, 40(5): 855–862. See also Erratum, Bull. Lond. Math. Soc., 2010, 42(6): 973.

[5] Gibilisco P, Imparato D, Isola T. Stam inequality on Z_n. Stat. and Prob. Lett., 2008, 78(13): 1851–1856.

[6] Gibilisco P, Imparato D, Isola T. A Robertson-type uncertainty principle and quantum Fisher information. Lin. Alg. Appl., 2008, 428(7): 1706–1724.

[7] Gibilisco P, Imparato D, Isola T. A characterization of the uniform distribution on the circle by Stam inequality. Jour. Ineq. Pure Appl. Math., 2009, 10(2): Article 34, 7.

[8] Gibilisco P, Hansen F. Extensions of Rao's inequality for the harmonic mean of. Preprint. 2014.

[9] Jupp P. General versions of the information inequalities of van Trees and of Stam. Communications at the Conference, Information Geometry and its Applications III. University of Leipzig. 2010.

[10] König R, Smith G. The entropy power inequality for quantum systems. IEEE Trans. Inform. Theory, 2014, 60(3): 1536–1548.

[11] Lenells J. and Misiolek, G.: Amari-Chentsov connections and their geodesics on homogeneous spaces of diffeomorphism groups. J. Math. Sci., 2014, 196(2): 144–151.

[12] Li N, Luo S. Entanglement detection via quantum Fisher information. Phys. Rev. A, 2013, 88: 014301.

[13] Luo S. Quantum uncertainty of mixed states based on skew information. Phys. Rev. A, 2006, 73: 022324.

[14] Luo S, Fu S, Oh CH. Quantifying correlations via the Wigner-Yanase skew information. Phys. Rev. A, 2012, 85: 032117.

[15] Prakasa Rao BLS. An inequality for the expectation of harmonic mean of random matrices [Technical Report]. Indian Statistical Institute, Delhi. 1998.

[16] Rao CR. Fisher: the founder of modern statistics. Stat. Sci., 1992, 7: 34–38.

[17] Rao CR. Seven inequalities in statistical estimation theory. Student, 1996, 1: 149–158.

[18] Rao CR. Statistical proofs of some matrix inequalities. Lin. Alg. Appl, 2000, 321: 307–320.

[19] Uffink J, van Lith J. Thermodynamic uncertainty relations. Found. Phys., 1999, 29(5): 655–692.