Geometric aspects of the non-extensive statistical theory

C. Vignat* and J.-F. Bercher†

*Université Paris-Est, LIGM, Université de Marne la Vallée
†Université Paris-Est, LIGM, ESIEE, 5 bd Descartes, 77454 Marne-la-Vallée Cedex 2, France

Abstract. The family of Tsallis entropies was introduced by Tsallis in 1988. The Shannon entropy belongs to this family as the limit case $q \to 1$. The canonical distributions in $\mathbb{R}^n$ that maximize this entropy under a covariance constraint are easily derived as Student-t ($q < 1$) and Student-r ($q > 1$) multivariate distributions. A nice geometrical result about these Student-r distributions is that they are marginal of uniform distributions on a sphere of larger dimension d with the relationship $p = n + 2 + \frac{2}{q-1}$. As $q \to 1$, we recover the famous Poincaré's observation according to which a Gaussian vector can be viewed as the projection of a vector uniformly distributed on the infinite dimensional sphere. A related property in the case $q < 1$ is also available. Often associated to Rényi-Tsallis entropies is the notion of escort distributions. We provide here a geometric interpretation of these distributions. Another result concerns a universal system in physics, the harmonic oscillator: in the usual quantum context, the waveform of the $n-th$ state of the harmonic oscillator is a Gaussian waveform multiplied by the degree n Hermite polynomial. We show, starting from recent results by Carinena et al., that the quantum harmonic oscillator on spaces with constant curvature is described by maximal Tsallis entropy waveforms multiplied by the extended Hermite polynomials derived from this measure. This gives a neat interpretation of the non-extensive parameter q in terms of the curvature of the space the oscillator evolves on; as $q \to 1$, the curvature of the space goes to 0 and we recover the classical harmonic oscillator in $\mathbb{R}^3$.

Keywords: nonextensive theory, Tsallis entropy, Rényi entropy
PACS: 02.50.Cw

THE FAMILY OF EXTENDED ENTROPIES

The family of Tsallis entropies [1] of a density probability f_X is defined by

$$S_q(X) = \frac{1}{1-q}\int_{\mathbb{R}^n} f_X - f_X^q,\ q \neq 1$$

where the real positive number q is called the nonextensivity parameter. By l'Hospital's rule we recover the Shannon entropy

$$\lim_{q\to 1} S_q(X) = -\int_{\mathbb{R}^n} f_X \log f_X$$

which is thus a limit case. The family of Rényi entropies [2]

$$H_q(X) = \frac{1}{1-q}\log\int_{\mathbb{R}^n} f_X^q$$

is another set of generalized entropies which contains again the Shannon entropy as the limit case $q = 1$. The Rényi and Tsallis have the same extremal distributions: since we will focus on these maximal entropy distributions, we will consider indifferently any of these two families.

In a classical scheme, we study a complex system whose probability density may evolve according to a complicated law - a differential equation for example. We wish to obtain the equilibrium density, that describes the system after relaxation. A general principle states that for many such systems, the Shannon entropy of the system is increasing

$$\frac{dH_1(t)}{dt} \geq 0.$$

Since moreover the system is often subject to a constraint - for example a fixed value of its energy - its equilibrium probability density can be obtained as the one that maximizes the Shannon entropy under this energy constraint.

For some systems, this H theorem does not hold, but it may be shown that a H_q theorem holds [5]: under the same energy constraint, there exists a value of the parameter q such that

$$\frac{dH_q(t)}{dt} \geq 0,$$

so that the equilibrium probability density is now the maximum Rényi - or equivalently Tsallis - entropy under the same constraint, hence our interest in the characterization of these distributions.

Let us begin with the explicit expression of the maximum Rényi entropy distributions under a covariance constraint. We remark first that we may assume all variables centered since, with an infinite support, the entropy in invariant by any shift of the variable.

Theorem 1 *The maximum Rényi entropy distributions under the covariance constraint*

$$E[XX^t] = K$$

where K is an $n \times n$ positive definite matrix are:

- *if $q > 1$, the $n-$variate Student r-distributions*

$$f_{X;q}(X) = \begin{cases} A_q\left(1 - X^t\Sigma^{-1}X\right)^{\frac{1}{q-1}} & X \in \left\{Z \in \mathbb{R}^n | X\Sigma^{-1}X \leq 1\right\} \\ 0 & else \end{cases} \tag{1}$$

  with $A_q = \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{p-n}{2}\right)|\pi\Sigma|^{\frac{1}{2}}}$, $p = \frac{2q}{q-1} + n$ and $\Sigma = pK$.

- *if $\frac{n}{n+2} < q < 1$, the $n-$variate Student t-distributions*

$$f_{X;q}(X) = A_q\left(1 + X^t\Lambda^{-1}X\right)^{\frac{1}{q-1}}; x \in \mathbb{R}^n \tag{2}$$

  with $A_q = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)|\pi\Lambda|^{\frac{1}{2}}}$; $m = \frac{2}{1-q} - n$ and $\Lambda = (m-2)K$.

For obvious reasons, these distributions will be called $q-$Gaussian distributions, remarking that as $q \to 1$, they converge to the classical Gaussian distribution.

A GEOMETRIC CHARACTERIZATION OF THE MAXIMUM ENTROPY DENSITIES

the case $q > 1$

A noticeable property of the above distributions is that they can be simply obtained from a universal one, namely the uniform distribution on the sphere, after some simple geometric transformation. We first recall that a random vector U_p is uniformly distributed on the sphere $\mathscr{S}_{p-1}$ in $\mathbb{R}^p$ if it writes

$$U_p = \frac{N_p}{\|N_p\|}$$

where N_p is any orthogonally invariant random vector. For example, N_p can be chosen Gaussian with identity covariance matrix. We have the following

Theorem 2 *If U_p is uniformly distributed on the sphere $\mathscr{S}_{p-1}$ in $\mathbb{R}^p$ then any marginal vector $X \in \mathbb{R}^n$ of U_p with $n < p$ has a $q-$Gaussian distribution (1) with $K = I$ and $q > 1$ provided that*

$$p - n = \frac{2q}{q-1}.$$

the case $q < 1$

An extension of the preceding result can be given in the case $q < 1$ as follows.

Theorem 3 *If the vector $X = OM$ belongs to the unit ball B_k, then the point M can be considered as the orthogonal projection on B_k of the point $P \in \mathscr{S}_{k+1}$ (see Figure 1). The intersection of the line OM with the hyperplane $H_{k+1} = \left\{Z \in \mathbb{R}^{k+1} | Z_{k+1} = 1\right\}$ defines a unique point N such that the vector $Y = [ON_1, \ldots, ON_k]^t$ follows a Tsallis distribution with nonextensivity parameter $q = \frac{n-2}{n}$.*

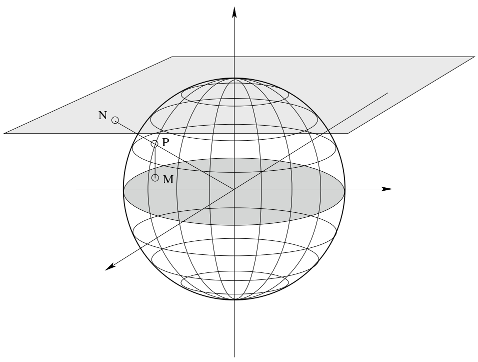

FIGURE 1. the gnomonic projection

We note that the geometric operation that associates the point M to the point N is called a gnomonic projection.

the limit $q \to 1$ case

It is a well-known result that the maximum Shannon entropy $n-$variate distribution under covariance constraint is the Gaussian distribution

$$f_{X;1}(X) = \frac{1}{|2\pi K|^{\frac{1}{2}}} \exp\left(-\frac{X^T K^{-1} X}{2}\right).$$

This limit is obviously recovered in the expressions (1) and (2). However, we may wonder how the geometric approach explicited above extends to this limit case. The answer is given by Poincaré's observation.

Theorem 4 *(Poincaré's observation) If U_p is a random vector uniformly distributed on the sphere $\mathscr{S}_{p-1}$ and if X_n with $n < p$ is any vector of U_p then, as $p \to +\infty$, the distribution of X_n (with fixed n) converges to the Gaussian $n-$variate distribution with identity covariance matrix.*

Our geometric characterization can thus be rephrased as follows: under unit covariance constraint, the maximum Shannon entropy distributions are the marginals of the uniform distribution on an infinite dimensional sphere, while the maximum Rényi entropy distributions are the marginals of the uniform distribution of a finite dimensional sphere.

THE ESCORT DISTRIBUTION

In non-extensive statistics, the escort distribution

$$f_q = \frac{f^q}{\int f^q}$$

plays an important role. But its probabilistic meaning is often overlooked; we give here an interpretation of this distribution following a result by Kullback [4, p.40]. Suppose that we have two models for some observed data normalized to a range, say$[-1,+1]$; the first model is that they follow a distribution f, and the second is that they follow the most natural one, i.e. the uniform distribution U on the interval $[-1,+1]$. In the absence of other informations, let us decide that the best choice f_* for the data is a trade-off between these two distributions. To measure this trade-off, we need a divergence, like the Kullback-Leibler divergence, so that a reasonable way to state our problem is to look for the distribution f_* such that

$$f_* = \arg\min_g D(g\|f) \text{ such that } D(g\|f) = \alpha D(g\|U) \tag{3}$$

which expresses the fact that f_* should be as close as possible to f while remaining close to U in a positive ratio α as well. The solution of the problem (3) is given by the escort distribution

$$f_* = f_q$$

where the nonextensivity parameter q is related to the ratio α and to the Lagrange parameter λ of the problem by $q = \frac{1+\lambda}{1+\lambda-\alpha\lambda}$. Two special cases are (i) $\alpha = 0$ which obviously implies $q = 1$ and $f_* = f$, and (ii) $\alpha = +\infty$ for which $q = 0$ and $f_* = U$. This result shows that the best trade-off (in the sense defined above) between a density f with support $[-1, +1]$ and its uniform counterpart is exactly the escort distribution.

GENERALIZED ENTROPIES AND THE QUANTUM HARMONIC OSCILLATOR

The quantum harmonic oscillator in the plane is the solution of the Schrödinger equation associated with a quadratic potential

$$H\psi_{m,n} = E_{m,n}\psi_{m,n}$$

and with Hamiltonian

$$H = -\frac{\hbar^2}{2m_0}\Delta + \frac{1}{2}m_0\omega^2\left(x^2+y^2\right).$$

With energy unit $\hbar\omega$ and length unit $\sqrt{\frac{\hbar}{m_0\omega}}$, the solutions are

$$\psi_{m,n}(x,y) = H_m(x)\,H_n(y)\,e^{-\frac{x^2+y^2}{2}};\; E_{m,n} = m+n+1$$

and the associated probability density is

$$f_{m,n}(x,y) = H_m^2(x)\,e^{-x^2}H_n^2(y)\,e^{-y^2}$$

We remark that the ground state follows the bivariate Gaussian distribution and that the other states are described by the Hermite polynomials, which are orthogonal with respect to the Gaussian measure. These polynomials are described by their Rodrigues formula

$$H_n(X) = (-1)^n e^{X^2}\frac{d^n}{dX^n}e^{-X^2}.$$

In a recent paper [3], Cariñena et al. extend these results to the case where the oscillator evolves in a 3-dimensional space of constant curvature κ: the plane when $\kappa = 0$, the sphere $\mathscr{S}_2$ when $\kappa > 0$ and the hyperbolic plane when $\kappa < 0$. Their results are as follows:

Theorem 5 *The harmonic oscillator on a space of constant negative curvature $-N\,(N>0)$ is*

$$\psi_{m,n}^N(z,y) = \mathfrak{h}_n^{N-m-\frac{1}{2}}(y)\,\mathfrak{h}_m^N(z);\; z = \frac{x}{\sqrt{1+\frac{y^2}{N}}}$$

with

$$\mathfrak{h}_n^N(x) = \left(1+\frac{x^2}{N}\right)^{-\frac{N}{2}-\frac{1}{4}}\mathscr{H}_n^N(x)$$

and $\mathscr{H}_n^N(x)$ *are orthogonal polynomials given by the Rodrigues formula*

$$\mathscr{H}_n^N(x) = (-1)^n \left(1+\frac{x^2}{N}\right)^{N+\frac{1}{2}} \frac{d^n}{dx^n}\left(1+\frac{x^2}{N}\right)^{n-N-\frac{1}{2}}.$$

The harmonic oscillator on a space of constant positive curvature $\nu > 0$ *is*

$$\psi_{m,n}^{\nu}(Z,Y) = \mathfrak{c}_m^{\nu+n+\frac{1}{2}}(Y)\,\mathfrak{c}_n^{\nu}(Z);Z = \frac{X}{\sqrt{1-\frac{Y^2}{\nu}}}$$

with

$$\mathfrak{c}_n^{\nu}(X) = \left(1-X^2\right)^{\frac{\nu}{2}-\frac{1}{4}}\mathscr{C}_n^{\nu}(X)$$

and $\mathscr{C}_n^{\nu}(X)$ *are the Gegenbauer orthogonal polynomials given by the Rodrigues formula*

$$\mathscr{C}_n^{\nu}(X) = \frac{(-2)^n}{n!}\frac{\Gamma(n+\nu)\Gamma(n+2\nu)}{\Gamma(\nu)\Gamma(2n+2\nu)}\left(1-X^2\right)^{\frac{1}{2}-\nu}\frac{d^n}{dX^n}\left(1-X^2\right)^{n+\nu-\frac{1}{2}}.$$

From this result, we observe that the ground states in both - negative and positive curvature - cases are exactly $q-$Gaussian distributions as defined by (1) and (2) respectively. As a conclusion, the Tsallis entropy extends naturally the role of the Shannon entropy for the harmonic oscillator to the case where the underlying space has constant curvature; moreover, the nonextensivity parameter q can be related explicitely to the curvature in such a way that as $q \to 1$, the curvature equals 0; for more details, see [6].

CONCLUSION

We have shown that a geometric approach can be adopted for some of the aspects of the nonextensive statistical theory: (i) the maximum entropy distributions can be obtained in certain cases as geometric projections of the uniform distribution on the sphere (ii) the escort distribution can be considered as a geometric trade-off in the Kullback Leibler divergence sense and (iii) the extension of the quantum harmonic oscillator to constant curvature spaces involves naturally the nonextensive entropy. More results of this kind are under investigation by the authors in order to enhance the understanding of the nonextensive theory.

REFERENCES

1. C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, Journal of Statistical Physics 52, 479-487 (1988).
2. A. Rényi, On measures of entropy and information, in Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I (Univ. California Press, Berkeley, Calif., 1961), pp. 547-561.
3. J. F. Cariñena, M. F. Rañada, M. Santander, The quantum harmonic oscillator on the sphere and the hyperbolic plane, http://arxiv.org/abs/0709.0399v1
4. S. Kullback, Information Theory and Statistics,New York: Wiley, 1959.
5. J. A. S. Lima, R. Silva and A. R. Plastino, Nonextensive Thermostatistics and the H Theorem, Phys. Rev. Lett. 86, 2938 - 2941 (2001)
6. C. Vignat, P.W. Lamberti, Cariñena polynomials are Jacobi polynomials, arXiv:0902.0451v1 [math.CA]