A Characterization of the Multivariate Distributions Maximizing Renyi Entropy

José A. Costa[1], Alfred O. Hero III
E.E.C.S., University of Michigan
Ann Arbor, MI 48109 USA

E-mail: hero@eecs.umich.edu

Christophe Vignat
Université de Marne-la-Vallée, ENST URA 820
77 454 Marne-la-Vallée cedex, France
and E.E.C.S., University of Michigan
Email: vignat@univ-mlv.fr.

Abstract — **We characterize the multivariate probability distributions that maximize the Renyi entropy under covariance constraint. Then, we show that these distributions are stable under a particular type of convolution.**

I. Introduction

Renyi entropy is now a widely used tool in information theory, with applications such as statistical processing or database indexing [1]. In this paper, we characterize the multivariate probability measures that maximize Renyi entropy under a covariance constraint. Then we extend a particular convolution whose stable distributions are these Renyi maximizing laws.

II. Multivariate Renyi maximizing distributions

We first recall that the Renyi entropy H_α of order α of a random variable $\mathbf{X}$ distributed according to $f(x)$ is $H_\alpha = \frac{1}{1-\alpha}\log\int f^\alpha(x)dx$ and that, as $\alpha \to 1$, this entropy converges to the classical Shannon entropy [2]. Let us denote by f_α the n-variate probability densities defined as follows:

$$f_\alpha(\mathbf{x}) = A_\alpha \frac{1}{|2\pi c_\alpha^2 \mathbf{K}|^{\frac{1}{2}}}\left(1 - \mathrm{sign}(\alpha-1)\frac{\mathbf{x}^T\mathbf{K}^{-1}\mathbf{x}}{2c_\alpha^2}\right)^{\frac{1}{\alpha-1}}$$

with $c_\alpha = \left(\mathrm{sign}(\alpha-1)\left(\frac{n}{2}+\frac{1}{\alpha-1}\right)\right)^{1/2}$, $A_\alpha = \frac{\Gamma\left(\frac{1}{\alpha-1}\right)}{\Gamma\left(\frac{1}{\alpha-1}-\frac{n}{2}\right)}$ if $\alpha < 1$, $\frac{\Gamma\left(\frac{\alpha}{\alpha-1}+\frac{n}{2}\right)}{\Gamma\left(\frac{\alpha}{\alpha-1}\right)}$ if $\alpha > 1$ and $\mathbf{x} \in \{\mathbf{x}\,|\,\mathbf{x}^T\mathbf{K}^{-1}\mathbf{x} \le 2c_\alpha^2\}$ for $\alpha > 1$ and $\mathbf{x} \in R^n$ for $\alpha \in [0, 1[$. Our main result expresses as follows:

Theorem 1 *the probability distribution that maximizes H_α under the constraint $E\left[\mathbf{X}\mathbf{X}^T\right] = \mathbf{K}$ is f_α $\forall\alpha > \frac{n}{n+2}$*

This theorem is easily proved using the convexity of a directed version of the Renyi divergence, namely $D_\alpha(f||g) = sign(\alpha-1)\int \frac{f^\alpha}{\alpha} + \frac{\alpha-1}{\alpha}g^\alpha - fg^{\alpha-1}$. In statistics, the distributions f_α are known as the Pearson type II or Student -t [3] laws ($\alpha > 1$) and type VII or Student -s laws ($0 < \alpha < 1$). Moreover, a stochastic representation of these random variables may be deduced from their elliptical symmetry property [4]:

Theorem 2 *if $\frac{n}{n+2} < \alpha < 1$, then $\mathbf{X} \sim f_\alpha$ writes as $\mathbf{X} = \left(2c_\alpha^2\mathbf{K}\right)^{\frac{1}{2}}\mathbf{N}_0/\sqrt{A}$ where $\mathbf{N}_0$ is an n-variate normal random variable, A is a scalar χ^2 random variable with $2c_\alpha^2 = \frac{2}{1-\alpha} - n$ degrees of freedom and independent of $\mathbf{N}_0$. Moreover, if $2c_\alpha^2 \in N$, then denoting $2c_\alpha^2 = m$, $\mathbf{X}$ writes as $\mathbf{X} = \mathbf{K}^{\frac{1}{2}}\frac{\mathbf{N}_0}{\sqrt{\frac{1}{2m}\left(N_1^2+\cdots+N_m^2\right)}}$ where the $N_{0\le i\le m}$ are Gaussian $\mathcal{N}(0,1)$ and mutually independent*

Now if $\alpha > 1$ and if $2c_\alpha^2 \left(= \frac{2}{a-1}+n\right) \in N - \{0\}$ then (denoting $m = 2c_\alpha^2$) $\mathbf{X}$ writes

$$\mathbf{X} = \mathbf{K}^{\frac{1}{2}}\frac{\mathbf{N}_0}{\sqrt{\frac{1}{m}\left(\mathbf{N}_0^T\mathbf{N}_0 + N_1^2 + \cdots + N_m^2\right)}}$$

III. A natural convolution associated with the Renyi entropy

Based on Urbanik's results [6], we provide now an extension of Kingman's convolution [5] that allows to identify the maximum Renyi entropy distributions as stable laws associated with a specific type of convolution.

Theorem 3 *Let X and Y both distributed according to f_α ($\alpha < 1$) and assume that $Z = X * Y \triangleq \left(|X|^{-2} + |Y|^{-2} + 2\lambda|X|^{-1}|Y|^{-1}\right)^{-\frac{1}{2}}$ where λ is a random variable independent of X and Y, and distributed according to f_{α_λ} ($\alpha_\lambda > 1$) with $\alpha = \frac{2\alpha_\lambda}{3\alpha_\lambda - 1}$. Then Z is also distributed according to f_α.*

As remarked by Kingman, among all possible random variables λ in (3), the only choice that makes the operation $Z = X * Y$ associative is exactly $\lambda \sim f_{\alpha_\lambda}$.

IV. Conclusions and perspectives

The preceding results represent a first step toward the characterization of operations having maximizing Renyi entropy distributions as stable laws. An extension to multivariate random variables, and to Student -r laws remains to be exhibited.

References

[1] A. O. Hero, *Divergence matching criteria for registration and indexing*, http://www.eecs.umich.edu/~hero/presentations.html

[2] A. Rényi, *Probability theory*, North-Holland Series in Applied Mathematics and Mechanics, Vol. 10., 1970

[3] A.M.S. de Souza, C. Tsallis, *Student's -t and -r distributions: unified distributions from an entropic variational principle*, Physica A, 236, 52-57, 1997

[4] K.C. Chu, *Estimation and Detection for linear systems with elliptical random variables*, IEEE Tr. on Automatic Control, 18, 500-505, 1973

[5] J. F. C. Kingman, *Random walks with spherical symmetry*, Acta Math 109, 11-53, 1953

[6] K. Urbanik, *Generalized Convolutions I to V*, Studia Mathematica, 23, 45, 80 -1, 83 -2 and 91-2 resp., 1963, 73, 84, 86 and 88 resp., pp. 217-45, 57-70, 167-189, 57-95, 153-178 resp.

[1] J. Costa was supported by Fundacao para a Ciencia e Tecnologia under the project SFRH/BD/2778/2000.

0-7803-7501-7/02/$17.00©2002 IEEE.