Physica A 510 (2018) 365–374

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Towards an information geometric characterization/classification of complex systems. II. Critical parameter values from the (c,d)-manifold

Demetris P.K. Ghikas *, Fotios D. Oikonomou

Department of Physics, University of Patras, Patras 26500, Greece

HIGHLIGHTS

- Our starting point is the two-parameter generalized entropy of Hanel and Thurner.
- We construct the normalized associated probability distribution.
- We use the two parameters as coordinates of a manifold.
- We construct the information metric of the above manifold.
- We compute the scalar curvature and present it graphically.
- We conjecture that the maxima and minima of it correspond to systems with special properties.

ARTICLE INFO

Article history:
Received 18 March 2018
Received in revised form 22 June 2018
Available online 30 June 2018

Keywords:
Complex systems
Generalized entropies
Information geometry

ABSTRACT

In our previous paper (I) we derived information geometric objects from the two parameter generalized entropy of Hanel and Thurner (2011), using the c,d parameters as labels of the corresponding manifolds. Here we follow a completely different approach by considering these parameters as coordinates of a single information manifold. This gives a manageable two-dimensional manifold amenable to easy manipulations but most importantly it offers a direct characterization of complex systems in terms of the pair of the c,d values. As a result we obtain certain characteristic values from the scalar curvature which we could conjecture that they represent complex systems with specific behavior. It is further observed that the boundary values of the c,d parameters which characterize the Hanel–Thurner classification are in some sense singular. This asks for a regularization scheme which we try to establish.

1. Introduction

Classification of complex systems and generalized entropies

In our previous paper (I) [1] we have discussed both the problems of definition, characterization and classification of complex systems as well as the use of generalized entropies giving the most relevant references. Here our starting point is again the two-parameter distribution related to the (c,d)-entropy of Hanel and Thurner [2] but we consider these parameters as coordinates of our information manifold. This offers a direct association of complex systems with particular values of these parameters where the scalar curvature is extremized.

* Corresponding author.
E-mail address: ghikas@physics.upatras.gr (D.P.K. Ghikas).

https://doi.org/10.1016/j.physa.2018.06.113
0378-4371/

Information manifolds from probability distributions

Information Geometry emerged as a geometric framework in the theory of parameter estimation in mathematical statistics [3]. For a given statistical model, that is a given class of probability measures, there is associated an information manifold and a Riemannian metric. This metric enters in the estimation procedure through the Cramer–Rao Inequality giving the possible accuracy of an estimator of the parameters of the model. Further on, one may define non-Riemannian connections which offer a deeper analysis of the estimation procedure. Our work is based on the geometric quantities emerging in the Information Geometry which is based on the two-parameter entropy functional of Hanel and Thurner [2]. In this paper the Hanel–Thurner parameters are used as coordinates of the information manifold. Thus we obtain a two-dimensional manifold without any approximation, and we may use both analytical and numerical results to establish the structure of this.

In Section 2 we introduce in a minimal way the necessary definitions and geometric quantities of Information Geometry. In Section 3 the generalized entropy of Hanel and Thurner [2] is introduced with few comments on its derivation and properties. In Section 4 we present our results. First we state some properties which establish the appropriateness of the generalized distribution function. Then we compute the Riemannian metric for the (c,d)-distribution function and the scalar curvature. The first three graphs give aspects of the distribution function. The next three present the metric elements as functions of the coordinates c,d giving the first impression of their non-trivial form. Then the numerical construction of the scalar curvature is given in the seventh graph. From this it is evident that some areas of the manifold are not typical. To reveal this structure we present certain sections of the curvature surface as functions of the c parameter for some values of the d parameter. We also give the contour plots of the curvature and the locus of minima as a function of c and d. We could conjecture that these extreme values correspond to complex systems of particular properties.

2. Basic concepts of information geometry

Geometry from probability distributions and the Cramer–Rao inequality

Here we present only the necessary concepts in order to establish the notation. We refer to the bibliography for the details [3]. Let $\mathcal{X}$ be a measurable space on which probability distributions are defined. Let Ξ be a subset of R^n which contains the parameters that characterize a set of probability distributions. Then

$$S = \{p_\xi = p(x;\xi) | \xi = [\xi^1, \dots, \xi^n] \in \Xi\} \tag{1}$$

is a parametric family of probability distributions on $\mathcal{X}$. This is an n-dimensional parametric statistical model. Given the N observations $x_1, \dots, x_N$ the Classical Estimation Problem concerns the statistical methods that may be used to detect the true distribution, that is to estimate the parameters ξ. To this purpose, an appropriate estimator is used for each parameter. These estimators are functions of data x, precisely, we define an estimator as a mapping $\hat{\xi} = [\hat{\xi}^1, \dots, \hat{\xi}^n] : \chi \to R^n$. Since x is randomly generated, we can consider $\hat{\xi}$ as function of the random variable X . The quality of the estimation is measured by the variance–covariance matrix $V_\xi(\hat{\xi}) = [v_\xi^{ij}]$ where

$$v_\xi^{ij} = E_\xi[(\hat{\xi}^i(X) - \xi^i)(\hat{\xi}^j(X) - \xi^j)] \tag{2}$$

Suppose that the estimators are unbiased, namely

$$E_\xi[\hat{\xi}(X)] = \xi, \quad \forall \xi \in \Xi \tag{3}$$

Then a lower bound for the estimation error is given by the Cramer–Rao inequality

$$V_\xi(\hat{\xi}) \geq G(\xi)^{-1} \tag{4}$$

in the sense that $V_\xi(\hat{\xi}) - G(\xi)^{-1}$ is positive semi-definite, where $G(\xi) = [g_{ij}(\xi)]$

$$g_{ij}(\xi) = E_\xi[\partial_i l(x;\xi)\partial_j l(x;\xi)] \tag{5}$$

the Classical Fisher Matrix with

$$l_\xi = l(x;\xi) = \ln p(x;\xi) \tag{6}$$

the score function. As it has been shown the Fisher Matrix provides a metric on the manifold of classical probability distributions. This metric, according to the theorem of Cencov [4], is the unique metric which is monotone under the transformations of the statistical model. This means that if the map $F : \mathcal{X} \to \mathcal{Y}$ induces a model $S_F = \{q(y;\xi)\}$ on $\mathcal{Y}$ then

$$G_F(\xi) \leq G(\xi) \tag{7}$$

That is, the distance of the transformed distributions is smaller than the original distributions. Thus monotonicity of the metric is intuitively related to the fact that in general we lose distinguishability of the distributions from any transformation of the information.