RESEARCH PAPER

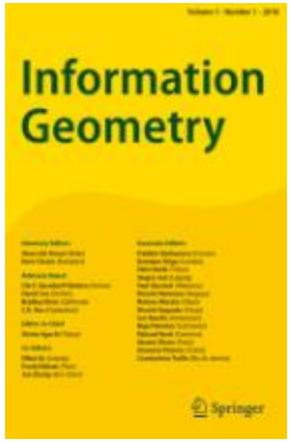
Natural alpha embeddings

Riccardo Volpi[1,2] · Luigi Malagò[1,2]

Received: 30 December 2019 / Revised: 23 November 2020 / Accepted: 2 February 2021 /
Published online: 18 April 2021

Abstract

Learning an embedding for a large collection of items is a popular approach to overcome the computational limitations associated to one-hot encodings. The aim of item embeddings is to learn a low dimensional space for the representations, able to capture with its geometry relevant features or relationships for the data at hand. This can be achieved for example by exploiting adjacencies among items in large sets of unlabelled data. In this paper we interpret in an Information Geometric framework the item embeddings obtained from conditional models. By exploiting the α-geometry of the exponential family, first introduced by Amari, we introduce a family of natural α-embeddings represented by vectors in the tangent space of the probability simplex, which includes as a special case standard approaches available in the literature. A typical example is given by word embeddings, commonly used in natural language processing, such as Word2Vec and GloVe. In our analysis, we show how the α-deformation parameter can impact on standard evaluation tasks.



Volume 4
Issue 1
Article 26

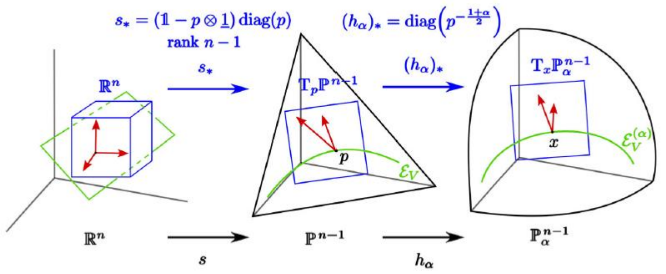


Fig. 1 Illustration of the mapping from the over-parametrization of the natural parameters in input to the softmax s, to the α-representation of the full simplex and of the exponential family $\mathcal{E}_V \subset \mathbb{P}^{n-1}$. Vectors in the tangent space are transported with the pushforward of the composite mapping

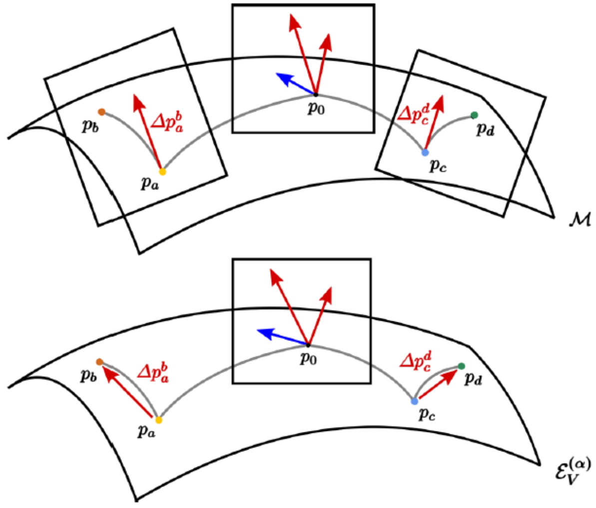


Fig. 2 (top) Geometric resolution of an analogy on a Riemannian manifold in a reference point p_0, by using a metric connection. (bottom) Resolution of an analogy in the α-embeddings framework. The points p should be intended as belonging to $\mathcal{E}_V^{(\alpha)}$, through the h_α-representation. In the bottom figure the vector $\Delta p_a^b = \mathrm{Log}_{p_a}^{(\alpha)} p_b$ does not belong to $\mathrm{T}_{p_a}\mathcal{E}_V^{(\alpha)}$, analogous is true for Δp_c^d

RESEARCH PAPER

The dually flat structure for singular models

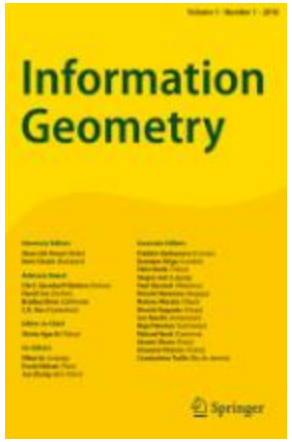
Naomichi Nakajima[1] · Toru Ohmoto[2]

Received: 13 November 2020 / Revised: 2 February 2021 / Accepted: 8 March 2021 /
Published online: 23 March 2021

Abstract

The *dually flat structure* introduced by Amari–Nagaoka is highlighted in information geometry and related fields. In practical applications, however, the underlying pseudo-Riemannian metric may often be degenerate, and such an excellent geometric structure is rarely defined on the entire space. To fix this trouble, in the present paper, we propose a novel generalization of the dually flat structure for a certain class of singular models from the viewpoint of *Lagrange and Legendre singularity theory*—we introduce a *quasi-Hessian manifold* endowed with a possibly degenerate metric and a particular symmetric cubic tensor, which exceeds the concept of statistical manifolds and is adapted to the theory of (weak) contrast functions. In particular, we establish Amari–Nagaoka's extended Pythagorean theorem and projection theorem in this general setup, and consequently, most of applications of these theorems are suitably justified even for such singular cases. This work is motivated by various interests with different backgrounds from Frobenius structure in mathematical physics to Deep Learning in data science.

Keywords Dually flat structure · Canonical divergence · Hessian geometry · Legendre duality · Wavefronts · Caustics · Singularity Theory



Volume 4
Issue 1
Article 27

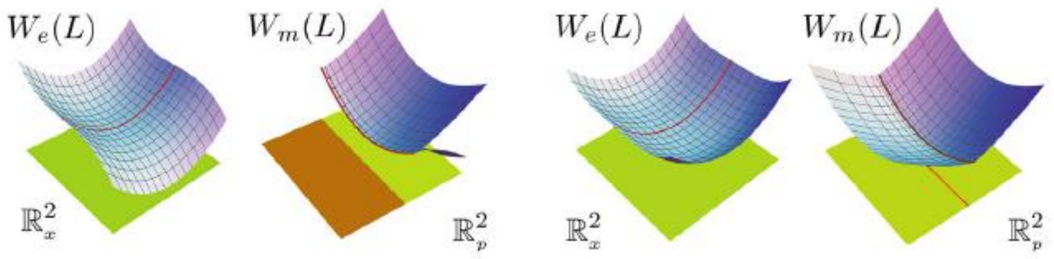
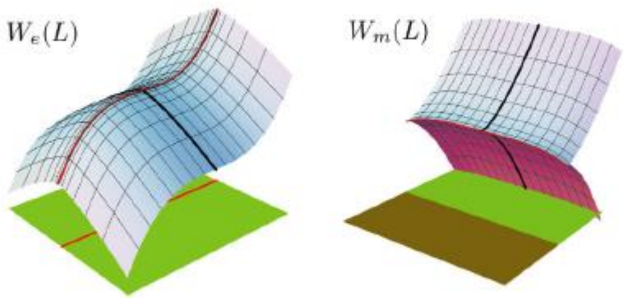


Fig. 1 The e/m-wavefronts and the e/m-caustics (Examples 3.4 and 3.5)

Fig. 2 Both e/m-wavefronts are singular (Example 3.6)



RESEARCH PAPER

Towards a canonical divergence within information geometry

Domenico Felice[1,2] · Nihat Ay[2,3,4,5]

Received: 27 June 2018 / Revised: 22 May 2021 / Accepted: 27 May 2021 /
Published online: 22 June 2021

Abstract
In Riemannian geometry geodesics are integral curves of the Riemannian distance gradient. We extend this classical result to the framework of Information Geometry. In particular, we prove that the rays of level-sets defined by a pseudo-distance are generated by the sum of two tangent vectors. By relying on these vectors, we propose a novel definition of a canonical divergence and its dual function. We prove that the new divergence allows to recover a given dual structure (g, ∇, ∇^*) of a dually convex set on a smooth manifold M. Additionally, we show that this divergence coincides with the canonical divergence proposed by Ay and Amari in the case of: (a) self-duality, (b) dual flatness, (c) statistical geometric analogue of the concept of symmetric spaces in Riemannian geometry. For a dually convex set, the case (c) leads to a further comparison of the new divergence with the one introduced by Henmi and Kobayashi.

Keywords Classical differential geometry (02.40.Hw) · Riemannian geometries (02.40.Ky) · Inverse problems (02.30.Zz) · Information geometry · Divergence functions

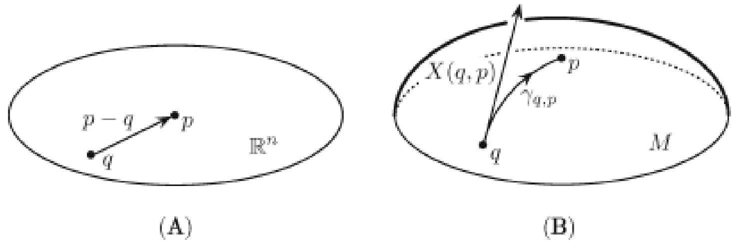
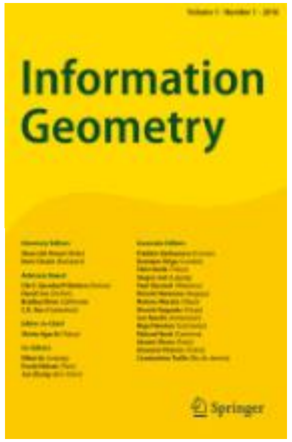


Fig. 1 On the left, (**A**) illustrates the difference vector $p-q$ in the linear vector space $\mathbb{R}^n$; whereas, in (**B**) we can see the difference vector $\mathrm{X}(q,p) = \dot{\gamma}_{q,p}(0)$ in M as the inverse of the exponential map at q (this figure comes from [4])



Volume 4
Issue 1
Article 28

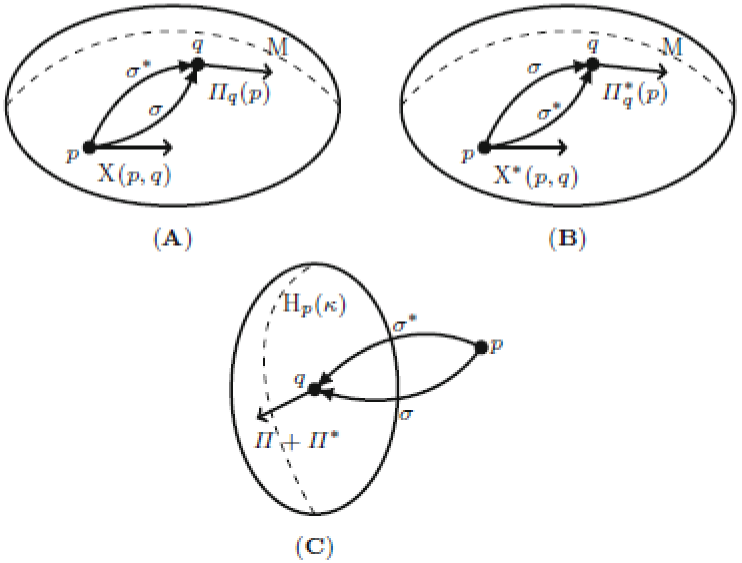


Fig. 2 From the top to the bottom, (**A**) isllustrates the vector Π that is the ∇-parallel transport of $\mathrm{X}(p,q) = \dot{\sigma}(0)$ along the ∇^*-geodesic σ^* from p to q; while (**B**) illustrates the vector Π^* that is the ∇^*-parallel transport of $\mathrm{X}^*(p,q) = \dot{\sigma}^*(0)$ along the ∇-geodesic σ from p to q. Finally, (**C**) shows that the sum $\Pi + \Pi^*$ is orthogonal to the level-hypersurface $\mathrm{H}_p(\kappa)$ of constant pseudo-squared-distance $r_p(q)$

RESEARCH PAPER

Information geometry and asymptotic geodesics on the space of normal distributions

Wolfgang Globke[1] · Raul Quiroga-Barranco[2]

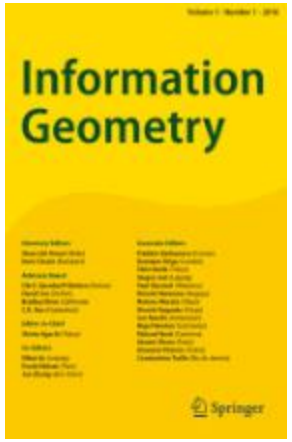
Received: 6 May 2019 / Revised: 10 June 2021 / Accepted: 17 June 2021 / Published online: 25 June 2021

Abstract

The family $\mathcal{N}$ of n-variate normal distributions is parameterized by the cone of positive definite symmetric $n \times n$-matrices and the n-dimensional real vector space. Equipped with the Fisher information metric, $\mathcal{N}$ becomes a Riemannian manifold. As such, it is diffeomorphic, but not isometric, to the Riemannian symmetric space $\mathrm{Pos}_1(n+1, \mathbb{R})$ of unimodular positive definite symmetric $(n+1) \times (n+1)$-matrices. As the computation of distances in the Fisher metric for $n > 1$ presents some difficulties, Lovrič et al. (J Multivar Anal 74:36–48, 2000) proposed to use the Killing metric on $\mathrm{Pos}_1(n+1, \mathbb{R})$ as an alternative metric in which distances are easier to compute. In this work, we survey the geometric properties of the space $\mathcal{N}$ and provide a quantitative analysis of the defect of certain geodesics for the Killing metric to be geodesics for the Fisher metric. We find that for these geodesics the use of the Killing metric as an approximation for the Fisher metric is indeed justified for long distances.

Keywords Gaussian distributions · Fisher metric · Cone of positive definite matrices · Symmetric spaces · Geodesics.

Mathematics Subject Classification Primary 53C35; Secondary 53C30 · 62H05 · 62B10



Volume 4
Issue 1
Article 29

RESEARCH PAPER

1-Conformal geometry of quasi statistical manifolds

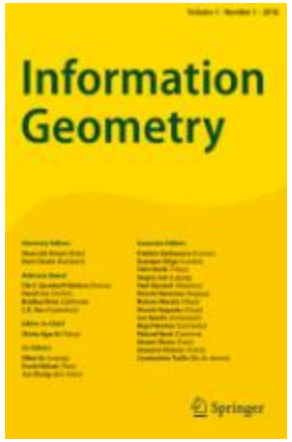
Keisuke Haba[1]

Received: 13 July 2019 / Revised: 20 September 2020 / Accepted: 25 September 2020 /
Published online: 7 October 2020

Abstract

A quasi statistical manifold is a generalization of a statistical manifold. The notion of quasi statistical manifolds was introduced to formulate the geometry of non-conservative estimating functions in statistics. Later, it was showed that quasi statistical manifolds are induced from affine distributions in the same way as statistical manifolds are induced from affine immersions. Here, an affine distribution is a non-integrable version of an affine immersion, and it is useful in quantum information geometry. On the other hand, it is known that generalized conformal geometry is useful for the study of statistical manifolds from the viewpoint of affine differential geometry. In particular, 1-conformal geometry of statistical manifolds gives a relation with the notion of affine immersions. Although generalized conformal geometry of quasi statistical manifolds is also expected to be useful, the geometry has not been cleared yet. The aim of this paper is to formulate 1-conformal geometry of quasi statistical manifolds. We research a relation between 1-conformal geometry of quasi statistical manifolds and the notion of affine distributions. As the main result, we show the fundamental theorems for affine distributions. We also formulate a hypersurface theory of quasi statistical manifolds.

Keywords Information geometry · Statistical manifold admitting torsion · Quasi statistical manifold · Affine distribution · 1-Conformal geometry



Volume 4
Issue 1
Article 30

RESEARCH PAPER

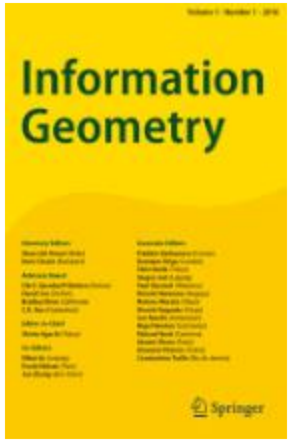
A characterization of the alpha-connections on the statistical manifold of normal distributions

Hitoshi Furuhata[1] · Jun-ichi Inoguchi[2] · Shimpei Kobayashi[1]

Received: 3 September 2019 / Revised: 2 September 2020 / Accepted: 12 October 2020 /
Published online: 20 October 2020

Abstract
We show that the statistical manifold of normal distributions is homogeneous. In particular, it admits a 2-dimensional solvable Lie group structure. In addition, we give a geometric characterization of the Amari–Chentsov α-connections on the Lie group.

Keywords Statistical manifolds · The Amari–Chentsov α-connection · Lie groups



Volume 4
Issue 1
Article 31

RESEARCH PAPER

Check for updates

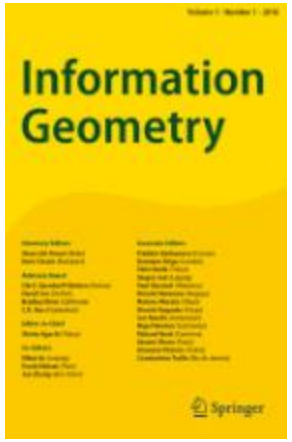
Statistical submanifolds from a viewpoint of the Euler inequality

Naoto Satoh[1] · Hitoshi Furuhata[1] · Izumi Hasegawa[2] · Toshiyuki Nakane[4] · Yukihiko Okuyama[4] · Kimitake Sato[4] · Mohammad Hasan Shahid[3] · Aliya Naaz Siddiqui[3]

Received: 1 October 2019 / Revised: 15 April 2020 / Accepted: 21 August 2020 /
Published online: 4 September 2020

Abstract
We generalize the Euler inequality for statistical submanifolds. Several basic examples of doubly autoparallel statistical submanifolds in warped product spaces are described, for which the equality holds at each point. Besides, doubly totally-umbilical submanifolds are also illustrated.

Keywords Statistical manifolds · Warped product · The Euler inequality · Doubly autoparallel submanifolds · Doubly totally-umbilical submanifolds



Volume 4
Issue 1
Article 32

RESEARCH PAPER

Harmonic exponential families on homogeneous spaces

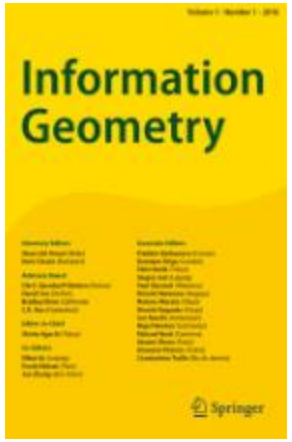
Koichi Tojo[1] · Taro Yoshino[2]

Received: 30 November 2019 / Revised: 27 August 2020 / Accepted: 4 September 2020 /
Published online: 6 October 2020

Abstract

Exponential families play an important role in the field of information geometry. By definition, there are infinitely many exponential families. However, only a small part of them are widely used. We want to give a framework to deal with these "good" families. In the light of the observation that the sample spaces of most of them are homogeneous spaces of certain Lie groups, we propose a method to construct exponential families on homogeneous spaces G/H by taking advantage of representation theory. Families obtained by this method are G-invariant exponential families. Then the following question naturally arises: are any G-invariant exponential families on G/H obtained by this method? We give an affirmative answer to this question. More precisely, any G-invariant exponential family on G/H can be realized as a subfamily of a family obtained by our method.

Keywords Exponential family · Representation theory · Homogeneous space · Harmonic analysis



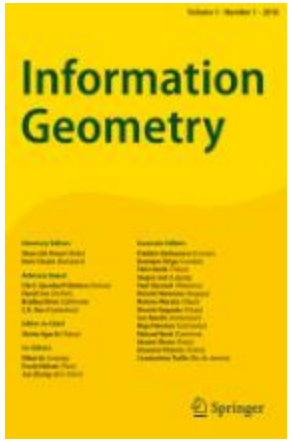
Volume 4
Issue 1
Article 33

SURVEY PAPER

Koszul lecture related to geometric and analytic mechanics, Souriau's Lie group thermodynamics and information geometry

Frédéric Barbaresco[1]

Received: 30 November 2019 / Revised: 6 November 2020 / Accepted: 8 November 2020 /
Published online: 13 January 2021



Volume 4
Issue 1
Article 34

Abstract

This paper deals with Jean-Louis Koszul's works related to Geometric and Analytic Mechanics, and to Souriau's Lie Group Thermodynamics that have appeared over time as elementary structures of Information Geometry. The 2nd Koszul form has been extended by Jean-Marie Souriau in his Symplectic model of Statistical Physics called "Lie Groups Thermodynamics" providing an extension of Fisher metric for homogeneous Symplectic manifolds, associated to KKS (Kirillov-Kostant-Souriau) 2-form in case of non-null cohomology. Jean-Louis Koszul has developed mathematical foundation of Souriau model in the Lecture "Introduction to Symplectic Geometry".

Keywords Koszul forms · Affine representation of lie algebra and lie group · Lie groups thermodynamics

Fig. 1 Jean-Louis Koszul and Hirihiko Shima at GSI'13 "Geometric Science of Information" conference in Ecole des Mines ParisTech in Paris, October 2013

RESEARCH PAPER

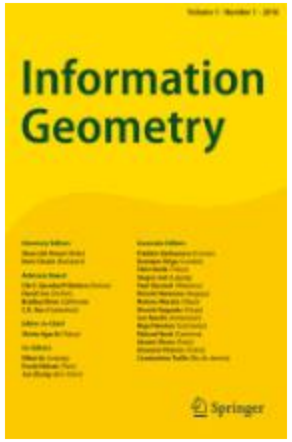
The last formula of Jean-Louis Koszul

Michel Nguiffo Boyom[1]

Received: 6 April 2020 / Revised: 1 September 2020 / Accepted: 11 September 2020 /
Published online: 29 September 2020

Abstract
Since the first international conference in France on the Information Geometry, GSI2013, Jean-Louis Koszul's interest in Information Geometry went increasing. Motivated by the impact of the cohomology of Koszul-Vinberg algebras on the Information Geometry and moved by some issues raised by Albert Nijenhuis, Jean-Louis Koszul undertook another rewriting of the **Brut formula** of the coboundary operator of the KV complex. The source of other motivations of Jean-Louis Koszul was the relationships between the theory of KV cohomology and the theory of deformation of locally flat manifolds.In 2015 Jean-Louis Koszul sent me his *Last formula* of the KV boundary operator. In that *Last formula* Jean-Louis Koszul dealt with the case where spaces of coefficients are trivial modules of KV algebras. A part of the present work is devoted to extending the *Last formula* of Jean-Louis Koszul to KV cochain complexes whose spaces of coefficients are non-trivial two-sided modules of KV algebras. At another side, I also aim to highlight other significant impacts of the theory of KV cohomology of Koszul-Vinberg algebras. In particular I will use the KV cohomology to widely revisit the theory of statistical models of measurable sets. The reader will see why the source of the theory of statistical models is of homological nature.I also intend to highlight several impacts of the KV cohomology on the quantitative differential topology. I am particularly concerned with problems regarding the existence of Riemannian foliations, the existence of symplectic foliations as well as the existence of multi-dimensional webs. The homological theory of statistical models is presented as branches of rooted trees whose roots are *weakly Jensen* random cohomology classes.

Keywords Lie algebroids · KV cohomology · Canonical characteristic class · Koszul geometry · Functor of Amari · Locally flat manifolds · Complex systems



Volume 4
Issue 1
Article 35